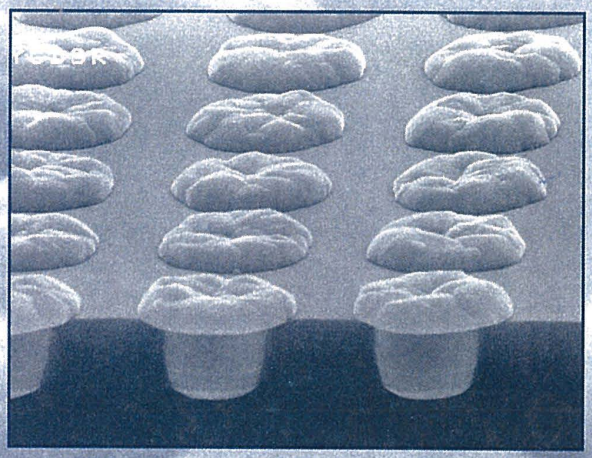
ULSI TECHNOLOGY



C.Y. CHANG and S.M. SZE

Micron Ex. 1023, p. 1
Micron v. YMTC
IPR2024-00909

ULSI Technology

Micron Ex. 1023, p. 2
Micron v. YMTC
IPR2024-00909

McGraw-Hill Series in Electrical and Computer Engineering

SENIOR CONSULTING EDITOR

Stephen W. Director, Carnegie Mellon University

Circuits and Systems
Communications and Signal Processing
Computer Engineering
Control Theory
Electromagnetics
Electronics and VLSI Circuits
Introductory
Power and Energy
Radar and Antennas
ULSI Technology

Previous Consulting Editors

Ronald N. Bracewell, Colin Cherry, James F. Gibbons, Willis W. Harman, Hubert Heffner, Edward W. Herold, John G. Linvill, Simon Ramo, Ronald A. Rohrer, Anthony E. Siegman, Charles Susskind, Frederick E. Terman, John, G. Truxal, Ernst Weber, and John R. Whinnery

Micron Ex. 1023, p. 3
Micron v. YMTC
IPR2024-00909

ULSI Technology

EDITED BY

C. Y. Chang

Chair Professor, College of Electrical Engineering and Computer Science
National Chiao Tung University
Director, National Nano Device Laboratories
Hsinchu, Taiwan, ROC

S. M. Sze

UMC Chair Professor
Department of Electronics Engineering
Director, Microelectronics and Information Systems Research Center
National Chiao Tung University
Hsinchu, Taiwan, ROC

THE McGRAW-HILL COMPANIES, INC.

New York St. Louis San Francisco Auckland Bogotá Caracas Lisbon
London Madrid Mexico City Milan Montreal New Delhi
San Juan Singapore Sydney Tokyo Toronto

Micron Ex. 1023, p. 4
Micron v. YMTC
IPR2024-00909

McGraw-Hill

A Division of The McGraw-Hill Companies



ULSI TECHNOLOGY

Copyright ©1996 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

Acknowledgments begin on page 724 and appear on this page by reference.

This book is printed on acid-free paper.

2 3 4 5 6 7 8 9 0 DOC DOC 9 0 9 8 7 6

ISBN 0-07-063062-3

This book was set in Times Roman by Publication Services, Inc.
The editors were Lynn Cox and John M. Morriss;
the production supervisor was Denise L. Puryear.
The cover was designed by Farenga Design Group.
Project supervision was done by Publication Services, Inc.
R. R. Donnelley & Sons Company was printer and binder.

Cover photo: Electron micrograph of contact holes filled with CVD tungsten plugs (see Chapter 8). The diameter is 0.25 micron. *Courtesy of the National Nano Device Laboratories, National Science Council, R.O.C.*

Library of Congress Catalog Card Number: 95-81366

Micron Ex. 1023, p. 5
Micron v. YMTC
IPR2024-00909

ABOUT THE EDITORS

C. Y. CHANG is Chair professor at the National Chiao Tung University and Director of the National Nano Device Laboratories, Taiwan, R.O.C. He has been the Dean of the College of Engineering and the Dean of the College of Electrical Engineering and Computer Science. He has written a book, *GaAs High Speed Devices,* has published more than 200 technical papers in international journals, and holds eight patents in ULSI, VHIC, and optoelectronics. He has taught many electronics engineers who have contributed significantly to the rapid growth of ULSI industries in Taiwan. Dr. Chang was elected a Fellow of IEEE in 1988 for his "contributions to semiconductor development and to education."

S. M. SZE is UMC Chair Professor of the Electronics Engineering Dept. and Director of Microelectronics and Information Systems Research Center, the National Chiao Tung University, Taiwan, R.O.C. For many years he was a member of the technical staff at AT&T Bell Laboratories, Murray Hill, New Jersey. Dr. Sze has made pioneering contributions to semiconductor devices and processing technologies, including the invention of the nonvolatile memory (1967) and the fabrication of MOSFETs in the 0.1μm regime (1982). Author or coauthor of more than 100 technical papers, Dr. Sze has written three books on semiconductor devices and edited six books on VLSI technology, high-speed devices, semiconductor sensors, and related topics. He has been elected Fellow of IEEE, member of the Academia Sinica, and member of the National Academy of Engineering.

Micron Ex. 1023, p. 6
Micron v. YMTC
IPR2024-00909

To Our Colleagues and Students—
Past, Present, and Future
On the Centennial of Our University—
The National Chiao Tung University

Micron Ex. 1023, p. 7
Micron v. YMTC
IPR2024-00909

CONTENTS

List of Contributors xiii
Preface xv
Introduction xvii

1 Cleanroom Technology 1
H. P. Tseng and R. Jansen
1.1 Introduction 1
1.2 Cleanroom Classification 4
1.3 Cleanroom Design Concept 8
1.4 Cleanroom Installation 24
1.5 Cleanroom Operations 32
1.6 Automation 34
1.7 Related Facility Systems 38
1.8 Summary and Future Trends 54
References 55
Problems 58

2 Wafer-Cleaning Technology 60
C. Y. Chang and T. S. Chao
2.1 Introduction 60
2.2 Basic Concepts of Wafer Cleaning 60
2.3 Wet-Cleaning Technology 92
2.4 Dry-Cleaning Technology 93
2.5 Summary and Future Trends 100
References 101
Problems 104

3 Epitaxy 105
P. Wang
3.1 Introduction 105
3.2 Fundamental Aspects of Epitaxy 107
3.3 Conventional Si Epitaxy 115
3.4 Low-Temperature Epitaxy of Si 125
3.5 Selective Epitaxial Growth of Si 131

Micron Ex. 1023, p. 8
Micron v. YMTC
IPR2024-00909

3.6 Characterization of Epitaxial Films 135
3.7 Summary and Future Trends 138
References 139
Problems 143

4 Conventional and Rapid Thermal Processes 144
R. B. Fair
4.1 Introduction 144
4.2 Requirements for Thermal Processes 151
4.3 Rapid Thermal Processing 159
4.4 Summary and Future Trends 196
References 198
Problems 203

5 Dielectric and Polysilicon Film Deposition 205
H. C. Cheng
5.1 Introduction 205
5.2 Deposition Processes 205
5.3 Atmospheric-Pressure Chemical-Vapor-Deposited (APCVD) and Low-Pressure Chemical-Vapor-Deposited (LPCVD) Silicon Oxides 215
5.4 LPCVD Silicon Nitrides 225
5.5 LPCVD Polysilicon Films 226
5.6 Plasma-Assisted Depositions 240
5.7 Other Deposition Methods 247
5.8 Applications of Deposited Polysilicon, Silicon Oxide, and Silicon Nitride Films 257
5.9 Summary and Future Trends 262
References 264
Problems 268

6 Lithography 270
K. Nakamura
6.1 Introduction 270
6.2 Optical Lithography 272
6.3 Electron Lithography 295
6.4 X-Ray Lithography 312
6.5 Ion Lithography 320

Micron Ex. 1023, p. 9
Micron v. YMTC
IPR2024-00909

6.6 Summary and Future Trends 323
References 324
Problems 327

7 Etching 329
Y. J. T. Lii
7.1 Introduction 329
7.2 Low-Pressure Gas Discharge 330
7.3 Etch Mechanisms, Selectivity, and Profile Control 334
7.4 Reactive Plasma Etching Techniques and Equipment 343
7.5 Plasma Processing Processes 353
7.6 Diagnostics, End Point Control, and Damage 362
7.7 Wet Chemical Etching 364
7.8 Summary and Future Trends 366
References 367
Problems 369

8 Metallization 371
R. Liu
8.1 Introduction 371
8.2 Metal Deposition Techniques 379
8.3 Silicide Process 395
8.4 CVD Tungsten Plug and Other Plug Processes 406
8.5 Multilevel Metallization 412
8.6 Metallization Reliability 448
8.7 Summary and Future Trends 451
References 457
Problems 468

9 Process Integration 472
C. Y. Lu and W. Y. Lee
9.1 Introduction 472
9.2 Basic Process Modules and Device Considerations for ULSI 473
9.3 CMOS Technology 489
9.4 Bipolar Technology 495
9.5 BiCMOS Technology 502
9.6 MOS Memory Technology 510

Micron Ex. 1023, p. 10
Micron v. YMTC
IPR2024-00909

9.7 Process Integration Considerations in ULSI Fabrication Technology 519
9.8 Summary and Future Trends 523
References 525
Problems 529

10 Assembly and Packaging 530
T. Tachikawa
10.1 Introduction 530
10.2 Package Types 531
10.3 ULSI Assembly Technologies 540
10.4 Package Fabrication Technologies 552
10.5 Package Design Considerations 557
10.6 Special Package Considerations 569
10.7 Other ULSI Packages 573
10.8 Summary and Future Trends 580
References 583
Problems 585

11 Wafer Fab Manufacturing Technology 587
T. F. Shao and F. C. Wang
11.1 What Is Manufacturing? 587
11.2 Wafer Fab Manufacturing Considerations 593
11.3 Manufacturing Start-Up Technology 609
11.4 Volume Ramp-Up Considerations 639
11.5 Continuous Improvement 646
11.6 Summary and Future Trends 652
References 653
Problems 655

12 Reliability 656
J. T. Yue
12.1 Introduction 656
12.2 Hot Carrier Injection 657
12.3 Electromigration 663
12.4 Stress Migration 674
12.5 Oxide Breakdown 679
12.6 Effect of Scaling on Device Reliability 684

Micron Ex. 1023, p. 11
Micron v. YMTC
IPR2024-00909

12.7 Relations between DC and AC Lifetimes 686
12.8 Some Recent ULSI Reliability Concerns 690
12.9 Mathematics of Failure Distribution 692
12.10 Summary and Future Trends 696
References 697
Problems 703

Appendixes 705
A. Properties of Si at 300 K 705
B. List of Symbols 707
C. International System of Units 709
D. Physical Constants 710

Index 711

Micron Ex. 1023, p. 12
Micron v. YMTC
IPR2024-00909

LIST OF CONTRIBUTORS

C. Y. Chang
National Chiao Tung University
Hsinchu, Taiwan, ROC

T. S. Chao
National Nano Device Laboratories
Hsinchu, Taiwan, ROC

H. C. Cheng
National Nano Device Laboratories
Hsinchu, Taiwan, ROC

R. B. Fair
Microfabrication Technology
Center for Microelectronic Systems Technologies
Microelectronics Center of North Carolina
Raleigh, North Carolina, USA

R. Jansen
Crystal Consulting
Eindhoven. The Netherlands

W. Y. Lee
Electronics Research and Service Organization
Hsinchu, Taiwan, ROC

Y. J. T. Lii
Motorola
Austin, Texas, USA

R. Liu
AT& T Bell Labs
Murray Hill, New Jersey, USA

C. Y. Lu
Vanguard International Semiconductor Corporation
Hsinchu, Taiwan, ROC

K. Nakamura
National Nano Device Laboratories
Hsinchu, Taiwan, ROC

T. F. Shao
Texas Instruments, Inc.
Dallas, Texas, USA

T. Tachikawa
Mitsubishi Electrical Corp.
Mizuhara Itami, Hyogo, Japan

H. P. Tseng
Vanguard International Semiconductor Corporation
Hsinchu, Taiwan, ROC

F. C. Wang
Texas Instruments, Inc.
Dallas, Texas, USA

P. J. Wang
National Nano Device Laboratories
Hsinchu, Taiwan, ROC

J. T. Yue
Advanced Micro Devices
Sunnyvale, California, USA

Micron Ex. 1023, p. 14
Micron v. YMTC
IPR2024-00909

PREFACE

ULSI Technology describes the theoretical and practical aspects of the most advanced state of electronics technology—ultralarge-scale integration (ULSI), where an integrated circuit (IC) chip contains over 10 million semiconductor devices. With ULSI technology, the cost of electronics products will decrease while the system functionality and performance will increase. The ULSI chips will result in the realization of smart and brilliant electronic systems, and in the improvement of quality of life and global productivity.

To fabricate IC chips with such complexity, we have to employ the most sophisticated process equipment, to follow the most precise process steps, and to adopt the most stringent cleanroom specifications. The basic process steps for ICs were considered in *VLSI Technology, 2nd Edition* (McGraw-Hill, 1988). Because of their importance to ULSI circuits, topics such as cleanroom technology, wafer-cleaning technology, manufacturing technology, and the rapid thermal process, which were essentially not covered in the 1988 VLSI book, are extensively discussed in *ULSI Technology.* In addition, many key processes, such as lithography, etching, metallization, and process integration, have been totally revised and updated. However, because of space limitations, certain classic topics such as crystal growth, conventional thermal processes, analytical technologies, and yield are covered only briefly or not covered at all. We suggest that our readers consult *VLSI Technology, 2nd Edition* for details.

In *ULSI Technology,* each chapter has an introduction that provides a general discussion of a specific aspect of ULSI processing. Subsequent sections present the basic science underlying individual process steps, the necessity for particular steps in achieving required parameters, and the tradeoffs in optimizing device performance and manufacturability. The problems at the end of each chapter form an integral part of the development of the topic.

The book is intended as a textbook for senior undergraduate or first-year graduate students in applied physics, electronics engineering, and materials science; it assumes that the reader has already acquired an introductory understanding of the physics and technology of semiconductor devices. Because it elaborates on IC processing technology in a detailed and comprehensive manner, it can also serve as a reference for those actively involved in integrated circuit fabrication and process development.

In the course of writing this text, many people have assisted us and offered their support. First, we express our appreciation to the management of our industrial and academic institutions, without whose help this book could not have been written. We have benefited from suggestions made by our reviewers: Dr. K. M. Brown of Digital Equipment Corporation, Drs. P. Chang and I. D. Liu of United Microelectronics Corporation, Drs. J. Chen, N. S. Tsai, R. Tsai, and F. C. Tseng of Taiwan Semiconductor Manufacture Company, Dr. L. P. Chen of National Nano Device Laboratories, Dr. T. C. Chen of International Business Machines, Dr. P. Fang of Advanced Micro

Micron Ex. 1023, p. 16
Micron v. YMTC
IPR2024-00909

Devices, Dr. B. J. Lin of Linovation Incorporated, Prof. S. Murarka of Rensselaer Polytechnic Institute, Prof. K. Ohtsuka of Meisei University, Dr. S. Okazaki of Hitachi Ltd, Prof. C. Osbum of North Carolina State University, Dr. J. Sung of AT&T Bell Laboratories, Dr. N. Tsai of MOSEL-Vitalic, Inc., Prof. C. Y. Yang of Santa Clara University, Dr. J. L. Yeh of HP Laboratories, Prof. G. Declerck of IMEC, and Prof. T. Ohmi of Tohoku University.

We are further indebted to Mr. N. Erdos of AT& T Bell Laboratories for technical editing of the manuscript. We also thank Mrs. Shenn-May Lee Chang and Ms. C. C. Chang for handling the correspondence with our contributors and reviewers, Mrs. T. W. Sze for preparing the Appendixes, and Ms. P. L. Huang, Ms. L. J. Chang, Ms. Y. M. Chen, Ms. F. F. Fang, Ms. S. L. Hsiau, Ms. S. Y. Teng, and Mr. S. Y. Wu for preparing the Index.

We wish to thank the Ministry of Education, ROC, the National Science Council, ROC, and the Spring Foundation of the National Chiao Tung University for their financial support. One of the editors (S. M. Sze) would especially thank the United Microelectronics Corporation (UMC), Taiwan, ROC, for the UMC Chair Professorship grant that provided the environment to work on this book.

C. Y. Chang
S. M. Sze

Micron Ex. 1023, p. 17
Micron v. YMTC
IPR2024-00909

INTRODUCTION

GROWTH OF THE INDUSTRY

The United States has the largest electronics industry in the world, with a global market share of over 40%. Since 1958, the beginning of the integrated-circuit (IC) era, the factory sales of electronic products have increased by about thirty times [see Fig. 1, curve (a)[1,2]]. Electronics sales, which were $303 billion in 1993, are projected to increase at an average annual rate of 8.5% and reach a half-trillion-dollar level by the year 2000. In the same period, the IC market itself has increased at an even higher rate [see Fig. 1, curve (b)[1,2]].* IC sales in the United States were $28 billion in 1993 and are expected to grow by 13% annually, reaching $65 billion by the year 2000. The main impetuses for such phenomenal market growth are the intrinsic pervasiveness of electronic products and the continued technological breakthroughs in integrated circuits.

The world markets of electronics and semiconductor industries will grow at comparable rates. Figure 2 shows the 1993 world electronics industry with a global sales volume of $679.7 billion. Also shown are the market shares of the six major electronics applications: computer and peripherals equipment at 32.3%, consumer electronics at 21.2%, telecommunication equipment at 16.5%, industrial electronics at 14.3%, defense and space at 11.5%, and transportation at 4.2%. By the year 2000, the world electronics industry is projected to reach $1200 billion, which will surpass the automobile, chemical, and steel industries in sales volume.

Figure 3 shows the 1993 world semiconductor industry, with total sales of $85.6 billion. Only 14% is related to optoelectronics and discrete semiconductor devices. IC sales constitute 86% of the total volume, with the largest segment being memory ICs, followed by microprocessor and microcontroller units, logic ICs, and analog ICs. In 2000, the semiconductor industry is projected to reach $200 billion, with over $170 billion in integrated circuits.

Figure 4 shows the market shares of the three major IC groups: MOSFET, bipolar transistor, and ICs made from III–V compound semiconductors.[3] At the beginning of the IC era, the IC market was broadly based on bipolar transistors. However, because of the advantages in device miniaturization, low power consumption, and high yield, sales volume of MOS-based ICs has increased steadily and in 1993 amounted to 75% of the total IC market. By the year 2000, MOS ICs will capture the largest market share (88%) of all ICs sold. This book, therefore, emphasizes MOS-related ULSI technology.

*There were only two years in which the growths were negative: in 1974, due to the Middle East oil embargo, and in 1985, due to overproduction of personal computers.

Micron Ex. 1023, p. 18
Micron v. YMTC
IPR2024-00909

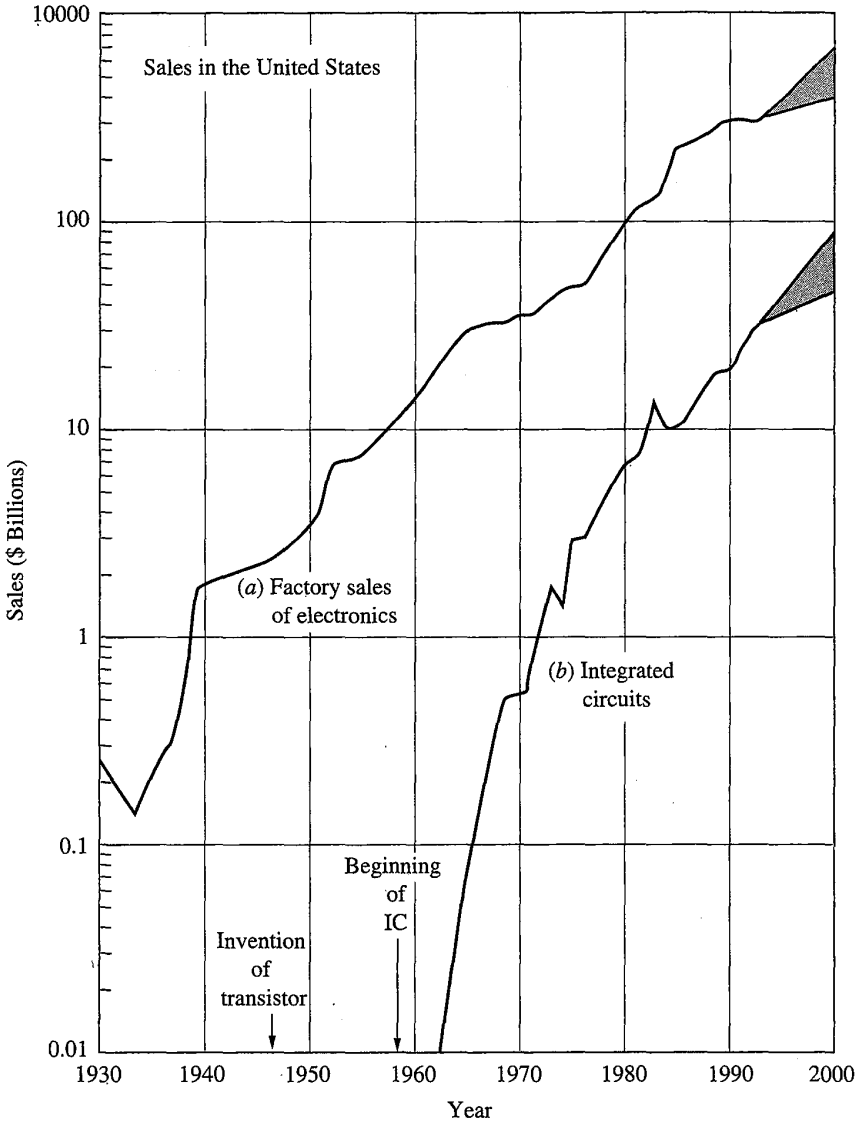


FIGURE 1
(*a*) Factory sales of electronics in the United States for the 64 years between 1930 and 1993 and projected to 2000. (*b*) Integrated circuit market in the United States for 32 years between 1962 and 1993 and projected to 2000. *(After Refs. 1 and 2.)*

DEVICE MINIATURIZATION

Figure 5, curve (*a*), shows the rapid growth in the number of components per MOS memory chip.[4,5] Note that the MOS IC complexity has advanced from small-scale integration (SSI), to medium-scale integration (MSI), to large-scale integration (LSI), to very-large-scale integration (VLSI), and finally to ultralarge-scale integration (ULSI), which has 10^7 or more components per chip. We note that since 1975 the growth has been maintained at a rate of about 40% annually; in other words, the

Micron Ex. 1023, p. 19
Micron v. YMTC
IPR2024-00909

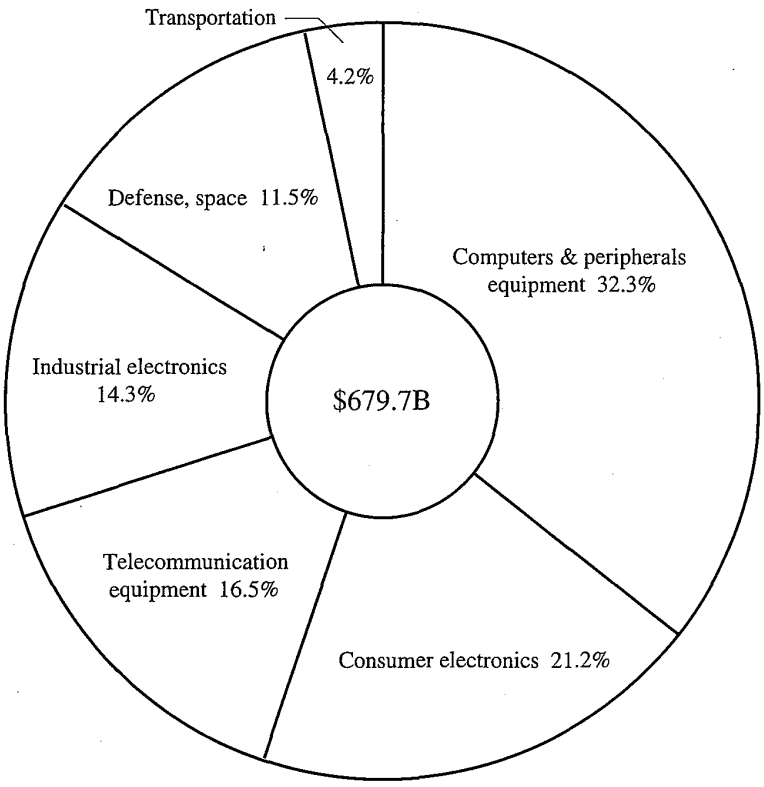


FIGURE 2
1993 world electronics industry. *(After Dataquest, 1994.)*

number of components has doubled every two years. At this rate, over 100 million components per chip will be available before the year 2000; in the early 21st century we will move into the gigabit range, with IC chips having more than one billion components.[6,7] Also shown in Fig. 5 is the growth of the number of components for bipolar, MESFET, and MODFET ICs. They are about two orders of magnitude lower in complexity compared with MOS-based ICs.

The most important factor in achieving the ULSI complexity is the continued reduction of the minimum device-feature length [see Fig. 6, curve (*a*)]. Since 1960, the annual rate of reduction has been 13%, which corresponds to a reduction by a factor of two every six years. At this rate, the minimum feature length will shrink from its present length of 0.5 μm to 0.2 μm in the year 2000. The junction depth of the source and drain junctions, and the gate oxide thickness are also being reduced at a similar rate as shown in curves (*b*) and (*c*) of Fig. 6, respectively.

The reduction of the device feature length and related dimensions has resulted in reduced overall device size and unit price per function. Figure 7 shows the relative price and size reductions.[8] In the past fifty years, prices have gone down by 100 million times, and the size has been reduced by a factor of one billion. By 2000 the price per bit is expected to be less than 0.1 millicent for a 64-megabit memory chip. Similar price reductions are expected for logic ICs. Additional benefits from device

Micron Ex. 1023, p. 20
Micron v. YMTC
IPR2024-00909

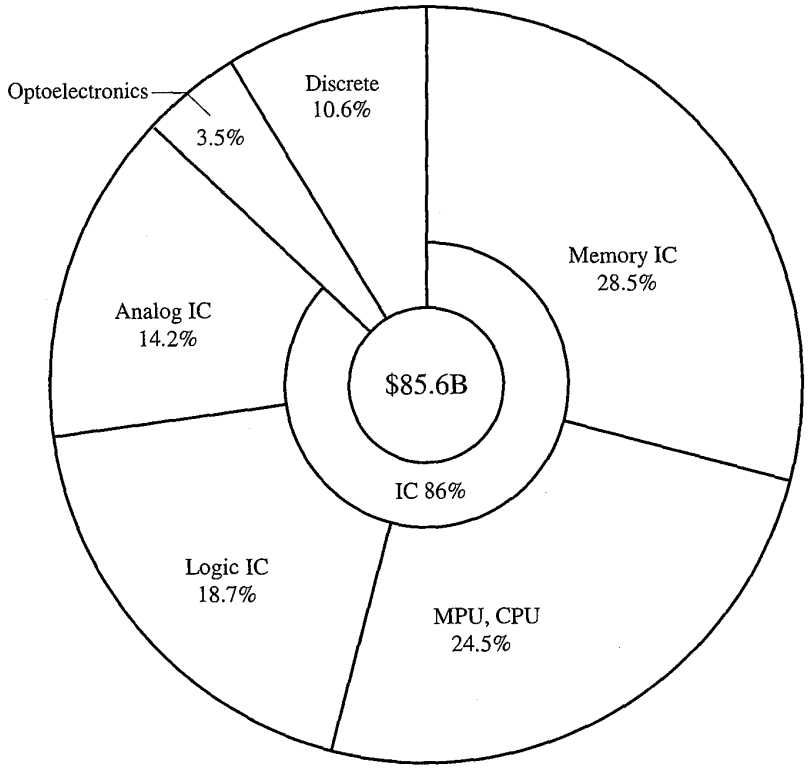


FIGURE 3
1993 world semiconductor industry. *(After Dataquest,1994.)*

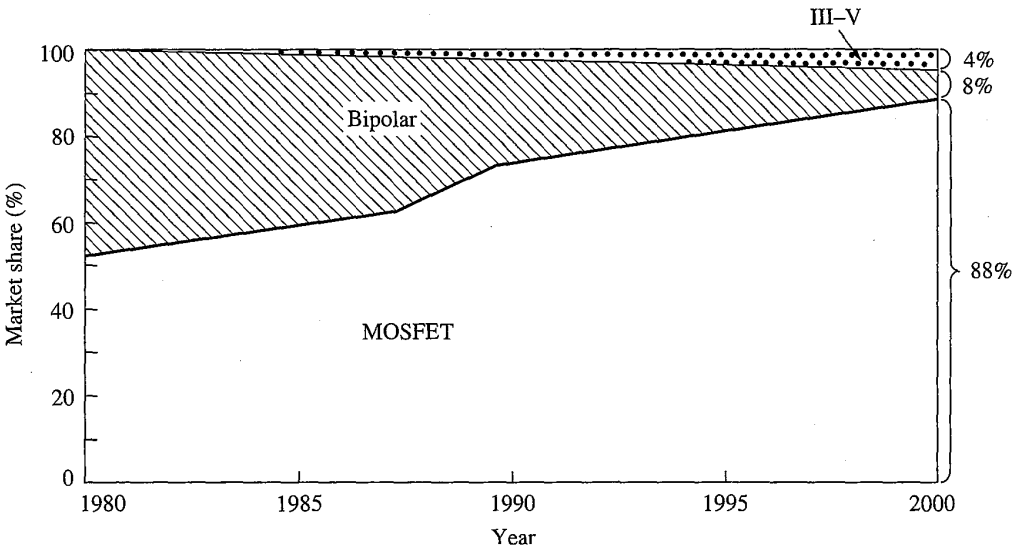


FIGURE 4
World IC market (1980–2000). *(After Zdebel, Ref.3.)*

Micron Ex. 1023, p. 21
Micron v. YMTC
IPR2024-00909

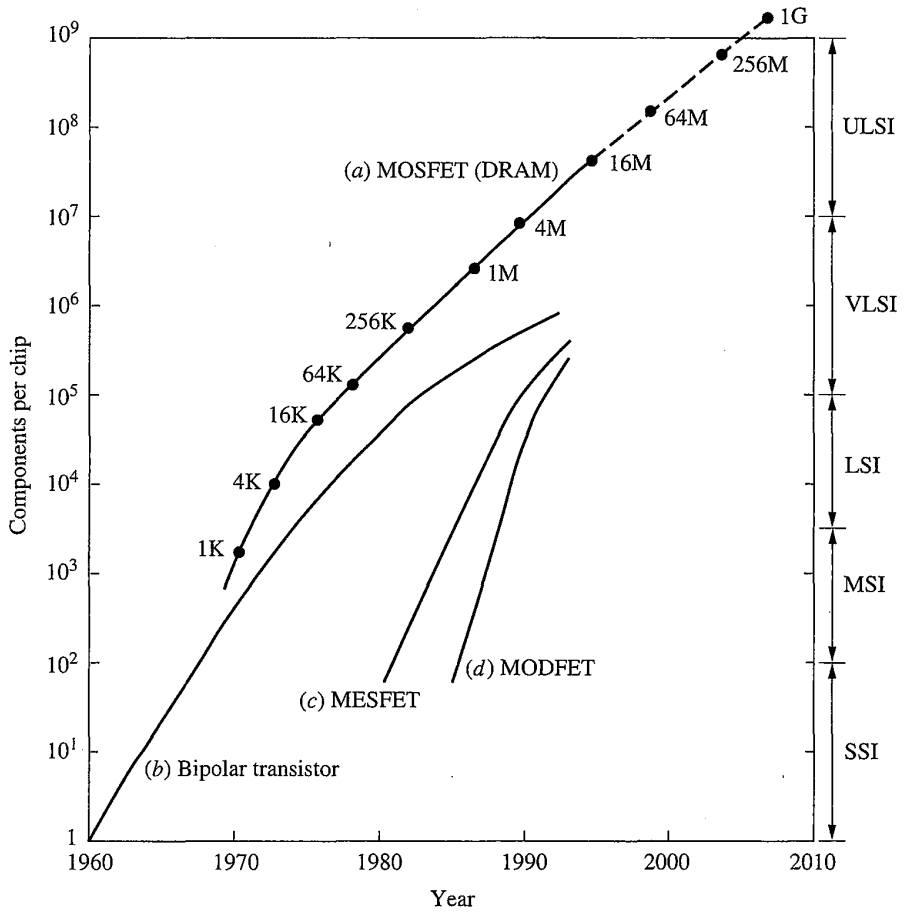


FIGURE 5
(*a*) Exponential growth of the number of components per MOS IC chip. *(After Moore, Ref. 4, and Myers, Ref. 5.)* (*b*), (*c*). and (*d*) Components per chip versus year for bipolar, MESFET, and MODFET ICs, respectively.

miniaturization include improvement of device speed (which varies inversely with the device feature length) and reduction of power consumption (which varies approximately with the square of the feature length). Higher speeds lead to expanded IC functional throughput rates, so that future ICs can perform data processing, numerical computation, and signal conditioning at 100 and higher gigabit-per-second rates.[9] Reduced power consumption results in lowering the energy required for each switching operation. Since 1960 the required energy, called the *power-delay product,* has decreased by six orders of magnitude.[10]

ORGANIZATION OF THE BOOK

Figure 8 shows how the 12 chapters of this book are organized. Chapter 1 considers cleanroom technology. The continued miniaturization in ULSI devices implies more

Micron Ex. 1023, p. 22
Micron v. YMTC
IPR2024-00909

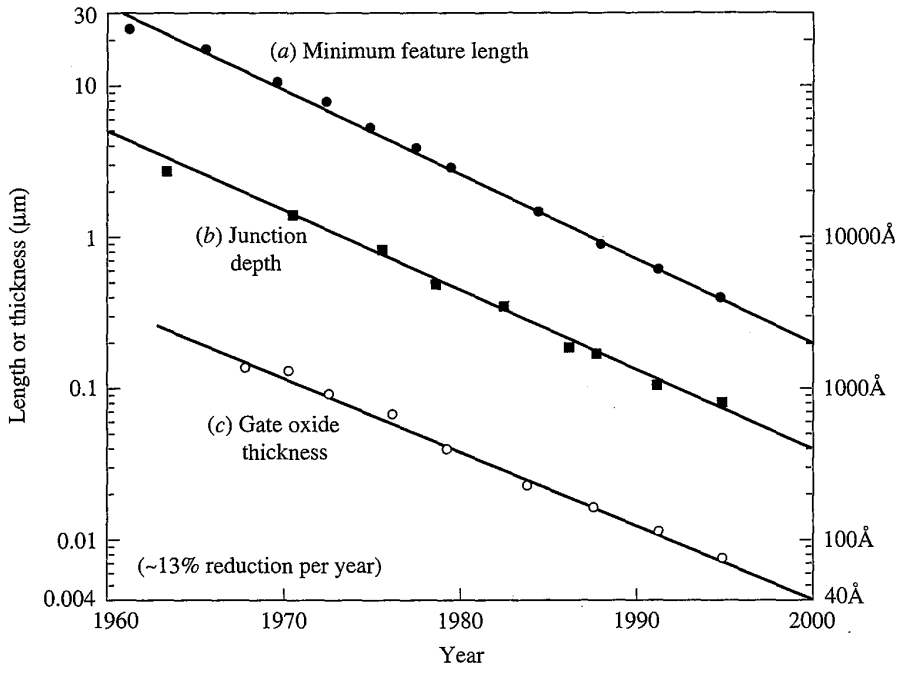


FIGURE 6
Exponential decrease of (*a*) minimum feature length, (*b*) junction depth, and (*c*) gate oxide thickness of MOSFET.

stringent requirements with respect to contamination control. Without an ultraclean processing environment, ULSI circuits simply cannot be realized.[11]

ULSI technology is synonymous with *silicon* ULSI technology. The unique combination of silicon's adequate bandgap, stable oxide, and abundance in nature ensures that in the foreseeable future no other semiconductor will seriously challenge its preeminent position in ULSI applications. Some important properties of silicon are listed in Appendix A.

Once the silicon wafers are in the cleanroom, we enter into the wafer-processing sequence, described in Chapters 2 through 8 and depicted in the wafer-shaped central circle of Fig. 8. Each of these chapters considers a specific process step. Of course, many processing steps are repeated many times in IC fabrication; for example, lithography and etching steps may be repeated 10 to 20 times. In ULSI technology the wafer-cleaning technology is as important as the cleanroom technology. Without a contamination-free wafer surface, the ICs will suffer from low yield and poor reliability. Because of limitations on the total length of the book, many classic topics, such as crystal growth, oxidation, diffusion, and ion implantation, are only briefly mentioned. The reader may consult textbooks on VLSI technology for details.[12]

The individual processing steps described in Chapters 2 through 8 are combined in Chapter 9 to form devices and integrated circuits. Chapter 9 considers the fundamental building process modules and four important IC families: CMOS (complementary MOSFET), bipolar ICs, BiCMOS (a combination of bipolar and CMOS),

Micron Ex. 1023, p. 23
Micron v. YMTC
IPR2024-00909

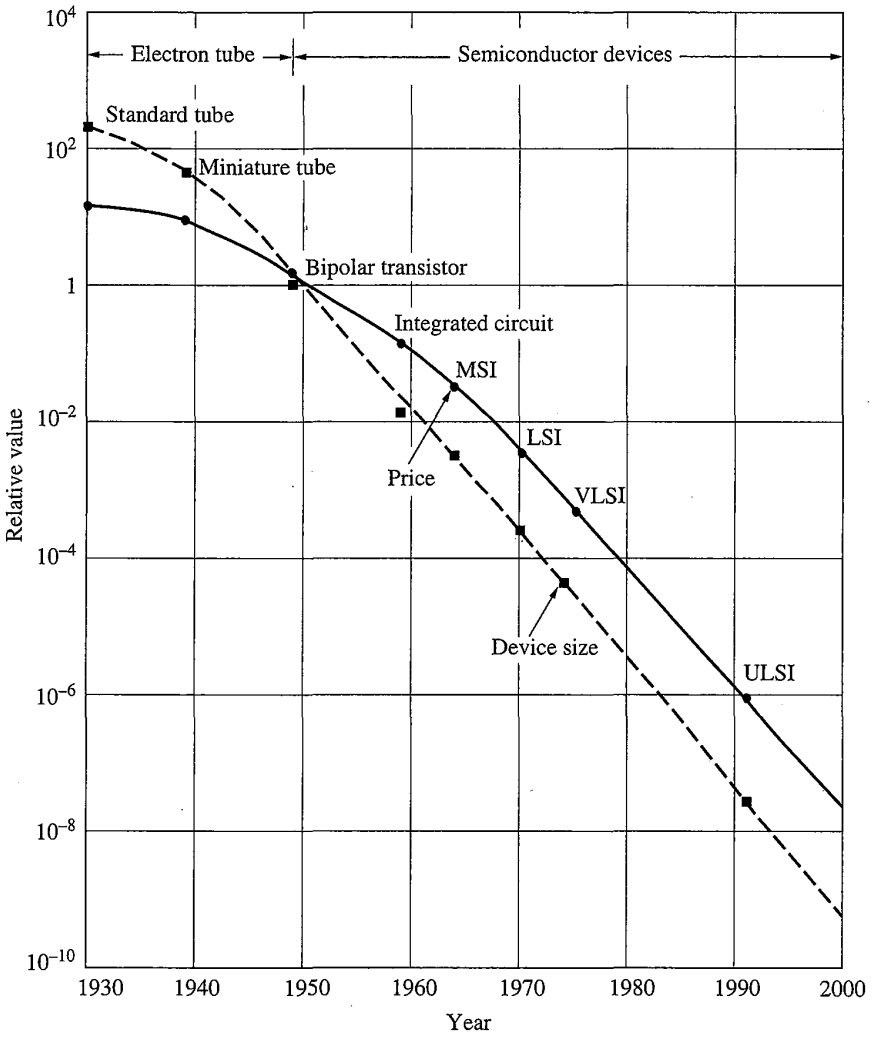


FIGURE 7
Price and size reduction of active electronic compents. *(After Shoda, Ref. 8.)*

and MOS memory ICs. After the completely processed wafers are tested, those chips that pass the tests are ready to be packaged. Chapter 10 describes the assembly and packaging of ULSI chips. Chapter 11 considers the manufacturing technology, that is, the strategy and logistics to implement various technologies to produce ULSI chips that meet customers' specifications in a timely fashion and to generate adequate return on investment for the IC manufacturer. Chapter 12 describes a multitude of reliability issues related to ULSI processes. As device dimensions move to the sub-half-micron and sub-quarter-micron regime, ULSI processing becomes more automated, resulting in tighter control of all processing parameters. At every step of production, from wafer cleaning to device packaging, numerous requirements are being imposed to improve the device performance and reliability.

To keep the notation simple in this book, we sometimes found it necessary to use a symbol more than once, with different meanings. However, within each chapter a

Micron Ex. 1023, p. 24
Micron v. YMTC
IPR2024-00909

FIGURE 8
Organization of this book.

Micron Ex. 1023, p. 25
Micron v. YMTC
IPR2024-00909

symbol has only one meaning and is defined the first time it appears. Many symbols do have the same or similar meanings consistently throughout this book; they are summarized in Appendix B.*

ULSI technology is presently moving at a rapid pace. The number of ULSI publications has doubled every year since 1990, the beginning of the ULSI era. Many topics, such as lithography, rapid thermal processing, and metallization, are still under intensive study. Their ultimate capabilities are not fully understood. The material presented in this book is intended to serve as a foundation. The references listed at the end of each chapter can supply more information.

REFERENCES

1. *1994 Electronic Market Data Book,* Electronic Industries Association, Washington, D.C., 1994.
2. *1994 Annual Report of Semiconductor Industry,* Industrial Technology Research Institute, Hsinchu, Taiwan, ROC, 1994.
3. P. J. Zdebel, "Current Status of High Performance Silicon Bipolar Technology," *14th Annual IEEE GaAs IC Symp. Tech. Digest,* 15 (1992).
4. G. Moore, "VLSI, What Does the Future Hold," *Electron Aust.,* **42,** 14 (1980).
5. W. Myers, "The Drive to the Year 2000," *IEEE Micro,* **11,** 10 (1991).
6. P. K. Chatterjee and G. B. Larrabee, "Gigabit Age Microelectronics and Their Manufacture," *IEEE Trans. VLSI Syst.* **1,** 7 (1993).
7. K. Mori, H. Yamada, and S. Takizawa, "System on Chip Age," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications,* k15 (1993).
8. K. Shoda, "Home Electronics in the 1990s," *Proceedings of the International Symposium on VLSI Technology, Systems, and Applications,* 1(1991).
9. H. Komiya, M. Yoshimoto, and H. Ishikura, "Future Technological and Economic Prospects for VLSI," *IEICE Trans. Electron.* **E76-C,** 1555 (1993).
10. R. W. Keyes, "Limitations of Small Devices and Large Systems," in N. G. Einspruch, Ed., *VLSI Electronics,* Academic, New York, 1981, Vol. 1, p. 186.
11. T. Ohmi, "ULSI Reliability through Ultraclean Processing," *Proc. IEEE,* **81,** 716 (1993).
12. For example, S. M. Sze, Ed., *VLSI Technology, 2nd Ed.,* McGraw-Hill, New York, 1988.

*Also included are the International System of Units (Appendix C) and Physical Constants (Appendix D).

Micron Ex. 1023, p. 26
Micron v. YMTC
IPR2024-00909

CHAPTER 1

Cleanroom Technology

H. P. Tseng and R. Jansen

1.1 INTRODUCTION

Microtechnologies are developing in a way that makes both the production tool and the process environment increasingly critical for the manufacturing process. The issue is, above all, to prevent defects in the product or, in other words, to increase the yield. This is achieved by the control of the wafer environment. The environmental factors include the control of temperature, relative humidity, electrostatic discharges, airborne particles, chemical contamination, electromagnetic fields, oxygen, and vibration.

The continuous miniaturization in device technology implies more and more stringent requirements with respect to contamination control.[1] Table 1 lists the changes in requirements from 1980 projected to 2004. In the 1980s a cleanliness class of 100–1000 was sufficient for the feature size of 2 μm. Today, a cleanliness class of 0.1 is the requirement when addressing a feature size of 0.5 μm down to 0.25 μm. In the future, inert minienvironments may be the only applicable solution.[2,3] To eliminate microcontamination and reduce native oxide growth on silicon wafers, the wafer processing and loading/unloading sections of a process tool are enclosed in an extremely high cleanliness minienvironment flushed with ultrapure nitrogen containing no oxygen and moisture. Similar developments can be observed with respect to the purities of the process utilities: from 1 ppm in the 1980s to 1 ppb in the very near future.

Reducing defect density is the most important aspect of increasing yield. Defect density is defined as the number of defects per cm^2 wafer that may occur during the processing due to all kinds of contamination. Particle contamination is still responsible for over 80% of defects. Table 2 summarizes the defect density requirements with respect to different design rules. Critical particle sizes, above which killing defects may be generated on the wafer surfaces, follow technology developments,

Micron Ex. 1023, p. 28
Micron v. YMTC
IPR2024-00909

TABLE 1
Evolution of IC processing features

Mass production started	1980	1984	1987	1990	1993	1996	1999	2004
Wafer size, mm	75	100	125	150	200	200	200	300
Technology (DRAM)	64K	256K	1M	4M	16M	64M	256M	1G
Chip size (cm^2)	0.3	0.4	0.5	0.9	1.4	2.0	3.0	4.5
Feature size(μm)	2.0	1.5	1.0	0.8	0.5	0.35	0.25	0.2–0.1
Process Steps	100	150	200	300	400	500	600	700–800
Cleanroom class (to be defined in Sec. 1.2)	1,000–100	100	10	1	0.1	0.1	0.1/ mini-environment	0.1/inert mini-environment
Utility impurity (ppb)	1,000	500	100	50	5	1	0.1	0.01

Source: Courtesy of KLA.

decreasing from 0.12 μm at a 0.8-μm-feature size to 0.03 μm at a 0.25-μm-feature size. As the process technology migrates from 0.8 μm 4M DRAM (dynamic random access memory) to 0.25 μm 256M DRAM over a time span of one decade, the maximum tolerable defect densities per critical layer measured at the defect size of 0.12 μm must be reduced from 0.28 D/cm^2 to 0.004 D/cm^2, representing an improvement factor of 65 times.

It is evident that the requirements for the control of the wafer environment are of great importance to the yield. In practice, the control of the wafer environment is a crucial aspect of the wafer manufacturing processes. It constitutes the basis for the design of the cleanroom systems and process utilities systems.

TABLE 2
Minimum defect density requirement with respect to design rule

DRAM process technology	Unit	4M	16M	64M	256M
Design rule	Micron	0.8	0.5	0.35	0.25
Critical layer	Each	9	10	11	13
Critical particle size	Micron	0.12	0.09	0.05	0.03
Killing defect size	Micron	0.27	0.18	0.1	0.06
Defect density measured at the killing defect size	D/cm^2	0.50	0.40	0.32	0.22
Defect density measured at 0.12 μm particle size	D/cm^2	2.53	0.90	0.22	0.055
Defect density/critical level measured at 0.12 μm	D/cm^2	0.28	0.09	0.02	0.004
Defect improvement factor		1	3	14	65

Source: Courtesy of KLA.

Micron Ex. 1023, p. 29
Micron v. YMTC
IPR2024-00909

Besides the extreme high-quality demands, issues such as flexibility, safety, reliability, and an accountable cost level are equally important to the design. Moreover, rapid developments within the market and process technology call for swift actions. At first glance, a number of these actions seem to be contradictory.

In the past, a typical reaction of designers and engineers to this rather complex problem was to rule out, as much as possible, all conceivable risks. This attitude caused unnecessarily high investments. Cost-effectiveness is the only acceptable solution. Cost-effectiveness can be understood as the optimum integration of appropriate quality, quantity, safety, flexibility, reliability, and time to achieve a competitive cost level. To achieve such cost-effectiveness, four conditions are of utmost importance, namely, know-how, approach, organization, and control.

The available know-how determines whether a design is or is not cost-effective. Inadequate knowledge often leads to overkill, resulting in excessive cost.

The project approach, combined with the project organization, constitutes the most important control mechanism. Modifications and surprises introduced at a late stage are caused mostly by an improper project approach.

The project organization is extremely important for efficiency, information, communication, and chain of responsibility. If projects are not properly organized, they tend to drift away from their objectives.

Control focuses on furnishing continuous, up-to-date information about the costs, timing, and quality. Such information is vital as a control tool.

In the past decade, submicrometer technology has enormously enriched the know-how and experience in both technique and approach. The learning curve is vitally important to each individual who is involved with the cleanroom technology in any way.

In an attempt to give some practical insights, a global clarification of the most important technological developments will be provided. The following subjects will be covered:

Section 1.2, on cleanroom classification, defines cleanroom cleanliness classes according to the current standard and explains their application.

Section 1.3, on cleanroom design concept, deals with the considerations needed to define the optimum process layout, and the design of a cleanroom. In addition, the common cleanroom performance criteria are listed, such as cleanliness class, temperature and humidity control, air quality, etc. Principal technical data that are essential for sizing the cleanroom facilities of a 200-mm-wafer fab are also presented.

Section 1.4, on cleanroom installation, provides basic information concerning the cleanroom subsystems such as exhaust, make-up air, recirculation air, chilled water, hot water, steam, filter ceiling, partition walls, and raised floor. Moreover, the recommended project approach, an overall time schedule for the design and realization of a cleanroom installation project, as well as a cost analysis are presented.

Section 1.5, on cleanroom operation, explains common cleanroom gowning procedures and cleanroom maintenance practices.

Section 1.6, on automation, clarifies various levels of process equipment control, from shop-floor-control software to the application of equipment interfaces.

Section 1.7, on related facility systems, provides basic information about the design of the deionized water system, process chemicals and gases, and the importance of space-management design.

Micron Ex. 1023, p. 30
Micron v. YMTC
IPR2024-00909

Finally, in Section 1.8, we present conclusions on all the main issues along with a look of the future trends.[4]

1.2 CLEANROOM CLASSIFICATION

The laminar-flow cleanroom concept was developed at Sandia National Laboratories, Albuquerque, New Mexico, in 1961 to provide a particle-free environment for thin-film deposition and the assembly of delicate mechanisms. This cleanroom concept was swiftly adopted by the then-burgeoning semiconductor industry.

TABLE 3
Evolution of Fed. Std 209 Series for specifying cleanliness of air

Date	Fed. Std.	Highlights of the original and revised contents
Dec. 1963	209	Cleanroom operation principles
Aug. 1966	209 A	Cleanroom design and testing methods Air flow pattern Laminar flow and turbulent flow Air velocity, 90 ± 20 ft/min Pressure, temperature, humidity, vibration Audio frequency noise, air exchange rate Air cleanliness classification specified as the number of particles at sizes larger than 0.5 μm per cubic foot; class 100, 10,000, and 100,000
Apr. 1973	209 B	Changed air velocity from 90 ± 20 ft/min to 90 ± 20% ft/min and changed humidity from 45% to 40 ± 5%
May 1977	209B amendment	Added cleanliness class 1,000
Oct. 1987	209C	Major revision of cleanroom classification and testing method Added classes 1 and 10 Extended the particle measurements from 5 μm and 0.5 μm down to 0.3 μm and 0.2 μm for class 100, and down to 0.3 μm, 0.2 μm, and 0.1 μm for class 10 and class 1 Clearly defined particulate sampling locations and numbers of sampling and measuring time
June 1988	209D	Corrected several typographical errors found in Fed. Std. 209C
Sep. 1992	209E	Adopted metric system Added descriptor to specify the maximum allowable number of ultrafine particles per cubic meter Added sequential airborne particle sampling plan to the single air sampling plan specified in Fed. Std. 209D

Micron Ex. 1023, p. 31
Micron v. YMTC
IPR2024-00909

TABLE 4
Metric definition of airborne particulate cleanliness classes per Fed. Std. 209E

	Particles/m³				
Class	**0.1 μm**	**0.2 μm**	**0.3 μm**	**0.5 μm**	**5 μm**
M1	3.50×10^2	7.57×10^1	3.09×10^1	1.00×10^1	
M1.5	1.24×10^3	2.65×10^2	1.06×10^2	3.53×10^1	
M2	3.50×10^3	7.57×10^2	3.09×10^2	1.00×10^2	
M2.5	1.24×10^4	2.65×10^3	1.06×10^3	3.53×10^2	
M3	3.50×10^4	7.57×10^3	3.09×10^3	1.00×10^3	
M3.5		2.65×10^4	1.06×10^4	3.53×10^3	
M4		7.57×10^4	3.09×10^4	1.00×10^4	
M4.5				3.53×10^4	2.47×10^2
M5				1.00×10^5	6.18×10^2
M5.5				3.53×10^5	2.47×10^3
M6				1.00×10^6	6.18×10^3
M6.5				3.53×10^6	2.47×10^4
M7				1.00×10^7	6.18×10^4

In December 1963, the U.S. government issued Fed. Std. (Federal Standard) 209 to standardize cleanroom design and operation guidelines. Over the past three decades, U.S. Fed. Std. 209 has undergone several revisions as illustrated in Table 3. In the last revision, Fed. Std. 209E, the original title used for nearly 30 years, "Cleanroom and Work Station Requirements, Controlled Environment," had been changed to "Airborne Particulate Cleanliness Classes in Cleanrooms and Clean Zones."

The recent addition of metric system units ensures that the Fed. Std. 209 series will be used by the international technical community as a world standard technical document for the characterization of the cleanliness of air. The definition of the metric airborne particulate cleanliness classes and the relevant air sampling method specified in Fed. Std. 209E are quite different than those of Fed. Std. 209D and, hence, warrant further discussion.

Tables 4 and 5 illustrate air cleanliness classes expressed in SI and English units, respectively. The numerical designation of the class in SI units is taken from the

TABLE 5
English definition of airborne particulate cleanliness classes per Fed. Std. 209E

	Particles/ft³				
Class	**0.1 μm**	**0.2 μm**	**0.3 μm**	**0.5 μm**	**5 μm**
1	3.50×10	7.50	3.00	1.00	
10	3.50×10^2	7.50×10	3.00×10	1.00×10^1	
100		7.50×10^2	3.00×10^2	1.00×10^2	
1000				1.00×10^3	7.00
10,000				1.00×10^4	7.00×10
100,000				1.00×10^5	7.00×10^2

Micron Ex. 1023, p. 32
Micron v. YMTC
IPR2024-00909

logarithm (base 10) of the maximum allowable number of particles, 0.5 μm and larger, per cubic meter. The numerical designation of the class in English units is taken from the maximum allowable number of particles, 0.5 μm and larger, per cubic foot. Note that the concentration of the particle, in general, is roughly inversely proportional to the particle size to a 2.2 power.

In addition to the metric definition of new and existing classes, standard nomenclature for the verification of a class was introduced in Fed. Std. 209E. It is no longer sufficient to state that air meets class 100 or class 1; Fed. Std. 209E requires that classes are expressed by using the format

Class X (at Y μm)

where X represents the numerical designation of the airborne particulate cleanliness class and Y represents the particle size.

In Fed. Std. 209E, a new air cleanliness definition, the U descriptor is added. The U descriptor specifies the maximum allowable numbers of ultrafine particles in a cubic meter of air in the size range from approximately 0.02 μm to the upper detectable limit of a discrete-particle counter. Without referring to a specific particle size, the U descriptor defines the cleanliness as

U(X)

where X is the maximum allowable number of ultrafine particles per cubic meter of air.

The description of air cleanliness is illustrated by the following examples:

"Class M 2.5 (at 0.3 μm)" describes air with no more than 1060 particles/m^3 with a particle size of 0.3 μm and larger.

"Class 100 (at 0.5 μm)" describes air with no more than 100 particles/ft^3 with a particle size of 0.5 μm and larger.

"U(20)" describes air with no more than 20 ultrafine particles/m^3.

Figure 1 shows the maximum allowable particle concentrations of each cleanliness class in U.S. Fed. Std. 209E as a function of particle size. The air cleanliness class in Japanese standard JIS B9920 rev. (revised) is shown in Fig. 2 for comparison.

Similar to Fed. Std. 209E, JIS B9920 rev. also defines cleanliness class by employing the logarithm of the maximum allowable particles per cubic meter. However, instead of specifying the cleanliness class at particle size of 0.5 μm, the JIS B9920 rev. specified the cleanliness class at particle size of 0.1 μm.

Because it uses metric units, the Fed. Std. 209E has made the necessary change in the specification of air sample location for verification of air cleanliness.[5,6] The sample location should be uniformly spaced throughout the clean zone operated in either a unidirectional (laminar) airflow or a nonunidirectional airflow environment. The minimum number of sample locations required for verification in a clean zone is specified below:

Unidirectional airflow (the lesser of (a) or (b)):

SI units: (a) $A/2.32$
(b) $A \times 64/(10^M)^{0.5}$

Micron Ex. 1023, p. 33
Micron v. YMTC
IPR2024-00909

English units: (a) A/25
(b) $A/(Nc)^{0.5}$

Nonunidirectional airflow:

SI units: $A \times 64/(10^M)^{0.5}$

English units: $A/(Nc)^{0.5}$

where A is the floor area of the clean zone (or entrance plane) in ft^2 for English units and m^2 for SI units, and M and Nc are the numerical designations of the class in SI and English units, respectively.

Note that number of locations must always be rounded to the next higher integer.

One of the most useful new features of Fed. Std. 209E is the sequential sampling option.[7] The single sampling plan in Fed. Std. 209E requires, as it did in Fed. Std. 209D, that the air volume to be sampled should anticipate 20 particles estimated from the air cleanliness class to be verified. For example, the sampling air volumes required for the verification of a cleanroom of class 1 should be at least 20 ft^3. Using a standard air sampling rate of 1 ft^3/min, it will take 20 minutes to complete a class

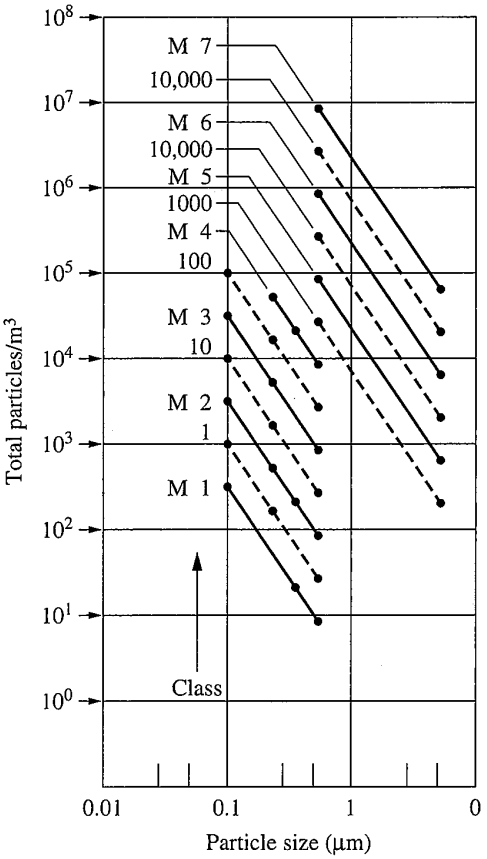


FIGURE 1
Air cleanliness according to U.S. Fed. Std. 209E.

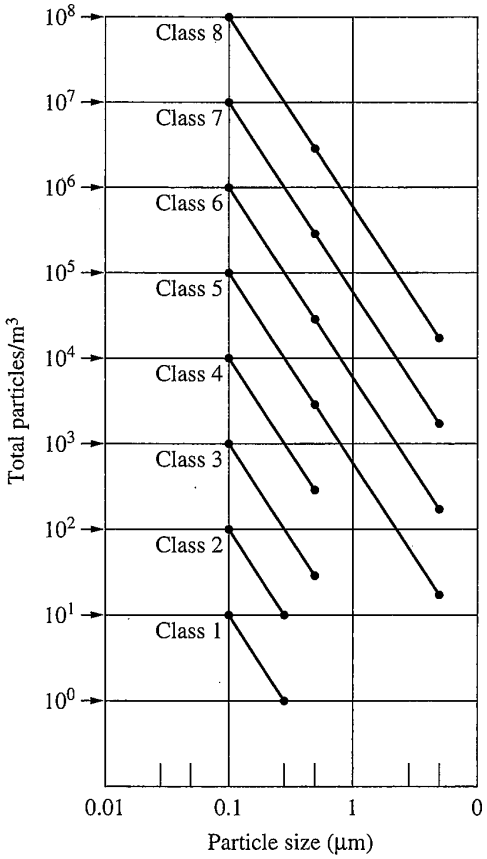


FIGURE 2
Air cleanliness according to Japanese Std. B9920 rev.

Micron Ex. 1023, p. 34
Micron v. YMTC
IPR2024-00909

1 cleanliness verification test. For air that is very clean, this statistical requirement demands that air be sampled for a very long time at each location. Thus, measuring the cleanliness of air that contains very low concentrations of particles can be unacceptably expensive. However, if air at a location is sampled while a running count of particles is recorded as a function of the time, frequently one can obtain a clear-cut indication of whether the air at that location passes or fails in a very short time by applying statistical criteria to the pattern of the count. This procedure allows verification of an airborne-particulate cleanliness class using samples of shorter duration than is required by the original single-sampling plan.

To use the sequential-sampling plan, record the number of particles observed as a function of time. Compare the count, as sampling continues, with the upper and lower limits, which are calculated by the two equations

$$\text{Upper} = 3.96 + 1.03E \tag{1.1}$$

$$\text{Lower} = -3.96 + 1.03E \tag{1.2}$$

where E is the expected count estimated from the air cleanliness class.

If the cumulative observed count for the sample exceeds the upper limit, then sampling is stopped and the air is judged to have failed. If the cumulative observed count drops below the lower limit, then sampling is stopped and the air is judged to have passed. If the cumulative observed counts equal 20 or less at the end of the sample duration, which is calculated from the air cleanliness class to anticipate a maximum of 20 particles, the air is also judged to have met the requirement.

1.3 CLEANROOM DESIGN CONCEPT

The design process starts with the translation of the process technology and the production volume into the specifications for the architectural, electrical, HVAC (heating, ventilation, and air-conditioning), central, and process utilities systems. The process technology is an important standard for quality requirement. For example, the cleanroom relates strongly to the feature size (design rules) of the technology in question. The quantity requirements are derived from the production volume combined with the technology. As a matter of course, attention should be paid to future developments of process tools and their impact on the cleanroom design.

With the aid of a rough computation, a first process tool list is developed. The process utilities requirements are established on the basis of this list. Electrical, mechanical, exhaust, gas, chemical, and other demands are listed in a matrix according to equipment tool. An accurate estimation of the average consumption of each process utility is essential for sizing the utility capacity properly. The facilities are subdivided into the following main systems:

Site infrastructure, site facilities, and buildings
Power supply
Electrical and communication
Cleanrooms and HVAC

Micron Ex. 1023, p. 35
Micron v. YMTC
IPR2024-00909

Central utilities
Process utilities
Environment and safety

The design criteria of each facility system and its subsystems are formulated, parallel to the definition of the specifications. These criteria include the quality and quantity characteristics as well as the schematic set-up. It should be clear that the process utilities requirements and the design criteria are determining factors for cost-effectiveness and that ignorance of these factors will almost always lead to overkill.

The design of a cleanroom normally starts with the optimization of cleanroom layout. Based on the targeted process technologies and the building architecture, a proper cleanroom system is selected and, subsequently, the relevant cleanroom performance criteria are defined.

1.3.1 Cleanroom Layout

Most cleanrooms have a rectangular shape with a central aisle. The central aisle serves as the corridor to each process tunnel extending sideways into the remaining portions of the cleanroom. The ideal width of the central aisle and process tunnel should be 2.4 meters to allow most process equipment to be transported in or out easily. The ideal height of the filter ceiling should be 3.4 to 3.6 meters above the raised floor to facilitate the transportation and installation of vertical furnaces, which are approximately 3 meters high.

Most 200-mm-wafer process equipment has the same throughput as the 150-mm-wafer equipment, except for steppers and high-current implanters. The footprints of the furnace, chemical stations, stepper, and photoresist coater/developer are also larger. Therefore a 200-mm-wafer fabrication plant requires 10% to 20% more cleanroom space than the 150-mm-wafer fab. A 200-mm, 16-Mbit, DRAM wafer fab with a monthly output capacity of 20,000 wafers requires approximately 5000 m^2 cleanroom area to accommodate the necessary process equipment and another 1000 m^2 cleanroom of lower classification to incorporate supporting facilities such as a gowning area, equipment-cleaning areas, and parts-cleaning areas, as well as a tube-cleaning area.

The entire IC wafer manufacturing process consists of 14 to 20 process modules. A typical process module, illustrated in Fig. 3, in general consists of 10 to 20 individual process steps. Within each process module, a layer of thin dielectric or conductive film is applied to the silicon wafer via oxidation, LPCVD (low-pressure chemical vapor deposition), APCVD (atmospheric-pressure chemical vapor deposition), PECVD (plasma-enhanced chemical vapor deposition), or metal-deposition techniques. Circuit patterns are then formed on a layer of photosensitive material in the photolithography area. Subsequently, the wafers are sent to either the etching or the ion implantation area. In the etching area, circuit patterns are etched into the thin dielectric or conductive film whereas in the implantation area ions such as As, P, or B are implanted into wafers to form devices. After stripping the photoresists, the implanted wafers are loaded into a diffusion furnace to anneal out the crystal defects

Micron Ex. 1023, p. 36
Micron v. YMTC
IPR2024-00909

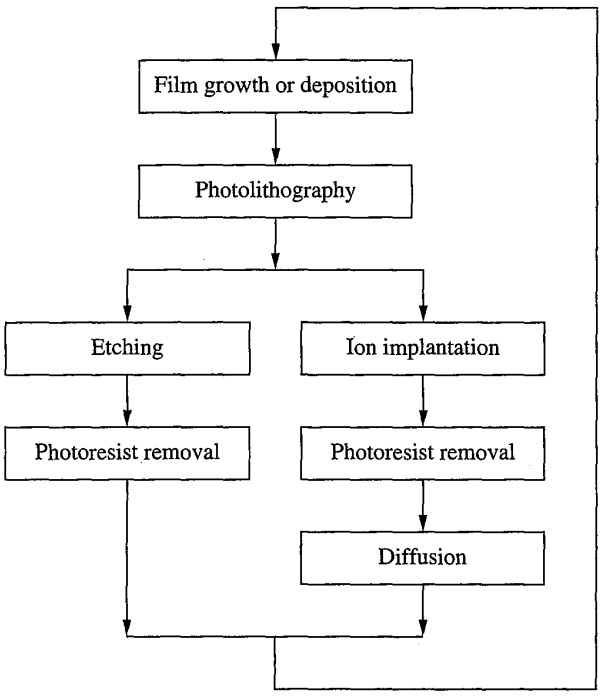


FIGURE 3
IC wafer manufacturing process module.

resulting from the ion implantation. The percentages of cleanroom space taken up by each process are summarized as follows:

Photolithography	25%
Diffusion and LPCVD	20%
Thin film	20%
Dry etching	15%
Implantation	10%
Wet process	10%

The following cleanroom layout principles are often used to optimize the IC production logistics.

Minimization of wafer transfer distance between each process area by putting photo, etching, and photo-resist removal close together.

Reduction of cross contamination by separating the back-end processes containing wafers exposed to metallic films from the front-end furnace and etching areas.

Flexibility for replacement and extension of process equipment.

Provision for interbay and intrabay automation and installation of equipment enclosures to encircle the wafer loading and unloading area in a highly classified environment (minienvironment).

Minimization of wafer transfer frequency between each process area by arranging process equipment in the following minimodule:

Furnace area:	Furnace, chemical station, and film measurement

Micron Ex. 1023, p. 37
Micron v. YMTC
IPR2024-00909

Photo area:	Stepper, resist coater, developer, resist-baking hot plates, optical inspection, and SEM (scanning electron microscope)
Etching area:	Resist-baking hot plates, etcher, and optical inspection
Resist stripping area:	Resist asher, wet chemical station

An example of cleanroom layout is given in Fig. 4. The film growth and deposition areas are located upstream of the photo area, whereas the etching, ion implantation, and photo-resist strip areas are located downstream. All process areas are

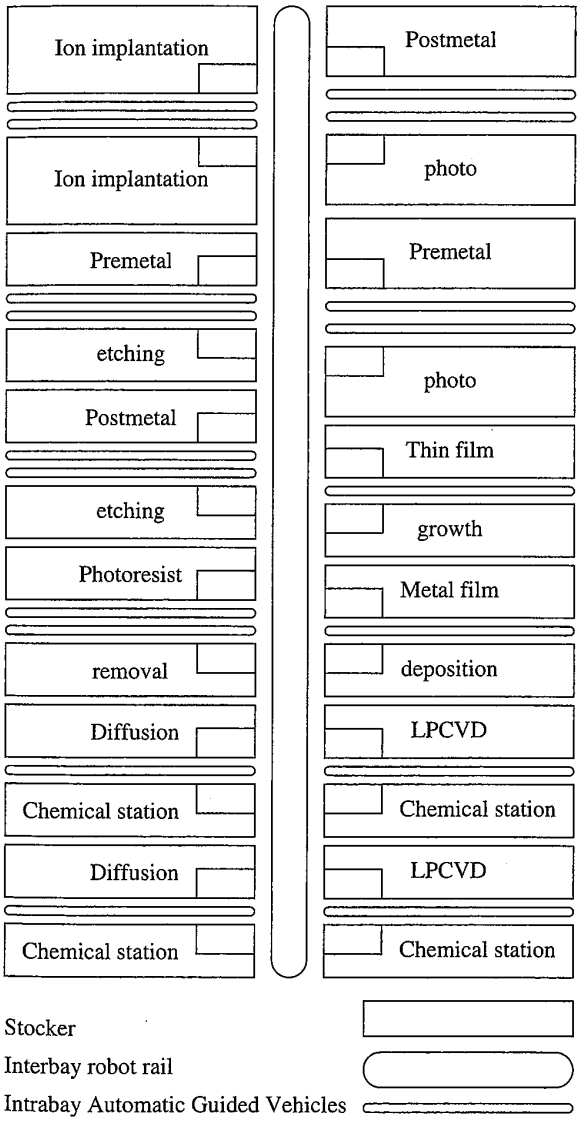


FIGURE 4
Cleanroom layout for a fully-automatic IC fab.

Micron Ex. 1023, p. 38
Micron v. YMTC
IPR2024-00909

connected via an interbay wafer transfer system. Within each process area, wafers are retrieved from a stocker, which can store up to 150 wafer cassettes, and then transported to process equipment using intrabay automatically guided vehicles (AGVs).

1.3.2 Cleanroom Design[8–17]

Design can be developed on the basis of the formulated specifications and the design criteria. The prepared process layout is all-important. In principle, this layout is derived from the required process tools on the one hand and the logistic requirements on the other. In addition, the potential future developments within the process technologies have to be considered. This implies that the following trend must be taken into account.

In the 1970s, the manufacturing process required an environment that provided an overall cleanliness class 10,000 and a local class 100. The "ballroom" concept of one large room was created.

As a consequence of the developments of submicron technologies, a need for a highly classified area originated in the 1980s. This led to the introduction of the "tunnel" concept, in which a corridor separates the process area from the service area. To achieve the required air cleanliness, the majority of the equipment maintenance functions were accommodated in the low-classified service areas. At the same time, the costly high-classified process tunnels where the wafers were handled were reduced.

Today, the process technologies are focused on 16-Mbit and 64-Mbit DRAM products. The process environment conditions required by these technologies are so stringent that the enclosure of the process environment for each process tool is being considered, moving the process tool in the direction of "clean machines."

This requirement resulted in the creation of the "minienvironment" concept as shown in Fig. 5.[18–23] Within the enclosure of the minienvironment, an extremely high cleanliness class (0.1 at 0.1 μm) is realized, whereas the overall production area has a cleanliness class 1000. The wafer processing and loading/unloading sections of process equipment will be affected automatically by input/output devices. The transport of wafers outside the minienvironments is by means of SMIF pods (standard mechanical interfaces), a technique originally developed by Hewlett Packard.

In general, flexibility is a key issue within this framework. Modifications to the process layout or the process environment conditions should be easy to make at any time.

Another consideration in the initial design is the selection of the cleanroom system, since it is normative with respect to the building structure. The cleanroom principle is based on the isolation of the process environment from external influences using a laminar airstream. This airstream is filtered so that it contains only a few particles of a determined size.

In the current manufacturing environment, only 1 particle at 0.1 μm or larger per cubic foot air is regarded as the upper threshold value. In addition, the temperature, relative humidity, and speed of the airstream is controlled within very tight tolerances.

Micron Ex. 1023, p. 39
Micron v. YMTC
IPR2024-00909

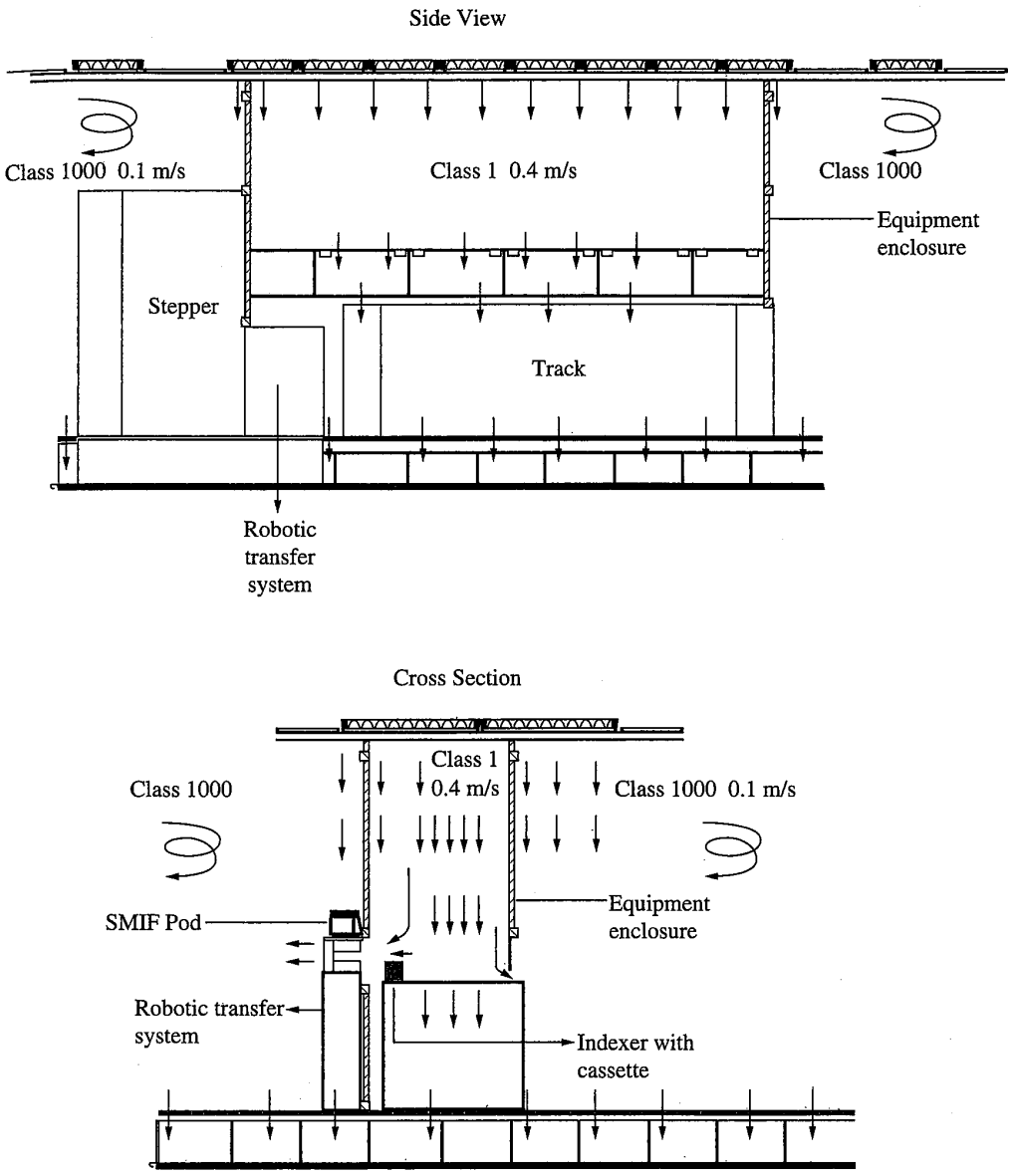


FIGURE 5
Creating an ultraclean minienvironment for a process tool using SMIF.

In the past, the common air speed was set at 0.45 m/sec. Currently, this value has been reduced to 0.35 m/sec based on experience. The reduced air speed offers a considerable cost reduction. In addition to the airflow, issues such as material selection, design details, and pressurization hierarchy are of great importance for the correct operation of the cleanroom.

A laminar airflow allows the creation of a cleanliness class 100, 10, 1, or even better. As a matter of course, the area in which such laminar airflow is established depends on the layout.

Micron Ex. 1023, p. 40
Micron v. YMTC
IPR2024-00909

To guarantee a highly classified process environment, the conditions in the background (overall production area) have to be controlled. In general, a factor of 10 is adopted between the cleanliness class of the critical area (foreground) and the background. The cleanliness classes 100,000, 10,000, and 1000, up to 100, can be reached with a turbulent airflow. This means that air turbulence is acceptable, and, as a consequence, local air-supply diffusers can be applied and low air-speed values are permissible. The dividing line between laminar and turbulent flow, an important transition point with respect to the investment level, lies with cleanliness class 100. This is illustrated with the denomination of a turbulent or laminar class 100. Naturally, the generation of particles inside the cleanroom must be prevented as much as possible. For this reason, gowning procedures should be defined and implemented for persons entering the cleanroom and decontamination procedures must be followed for incoming material.

1.3.3 Cleanroom Concept

There are four basic concepts applicable to the cleanroom system.[24] All current designs can be derived from them. The first and least expensive concept is illustrated in Fig. 6. The supplied air is distributed in the room via a filter ceiling and is returned to the air-handling units via the lower sides of the walls. Both process and service functions are accommodated in one level. This approach has very low flexibility; layout modifications are difficult and, thus, relatively expensive. Consequently, this concept is seldom used. However, the introduction of filter fan units allows for a one-level concept with a certain degree of flexibility. The disadvantages with respect to the location of remote equipment and the connection to the utility distribution networks still remain.

The second, third, and fourth basic concepts depart from separated process and service levels. The process level is completely classified. Costs are reduced by accommodating only the primary process functions on this level. The service level has all piping, ductwork, and cabling systems, and it is also used for the installation of remote equipment. Moreover, the air will be returned via this level to the air-handling units, by which the process level receives a high cleanliness class without extra costs. All potential contamination sources should be enclosed and exhausted to prevent cross-contamination. All three concepts feature a high degree of flexibility. The difference between these concepts lies in the air-handling system, namely,

Option two, shown in Fig. 7, is based on the installation of centrifugal fan units on top of the process level. Air-return shafts are located at both sides of the building. An air-supply plenum is included between the fan units and the filter ceiling.

Option three, illustrated in Fig. 8, is based on the installations of axial fan units at both sides of the building. In this case the space between the roof and the filter ceiling serves as the air-supply plenum.

Option four, shown in Fig. 9, is based on the installation of filter fan units on top of the ceiling grid. Air-return shafts are located at both sides of the building. The space between the roof and the filter fan units serves as an air-supply

Micron Ex. 1023, p. 41
Micron v. YMTC
IPR2024-00909

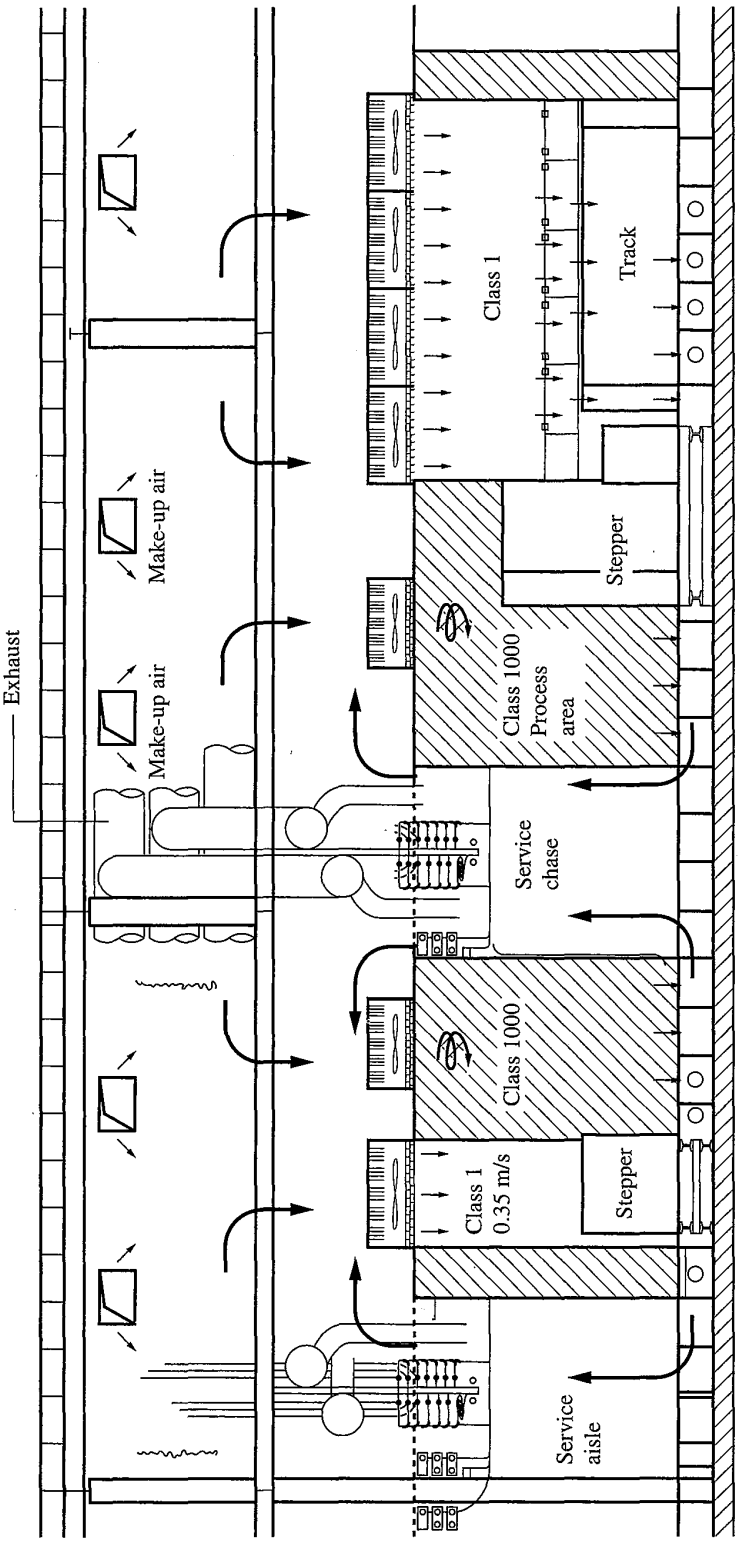


FIGURE 6
Ballroom-type cleanroom with process and service areas located on the same floor.

Micron Ex. 1023, p. 42
Micron v. YMTC
IPR2024-00909

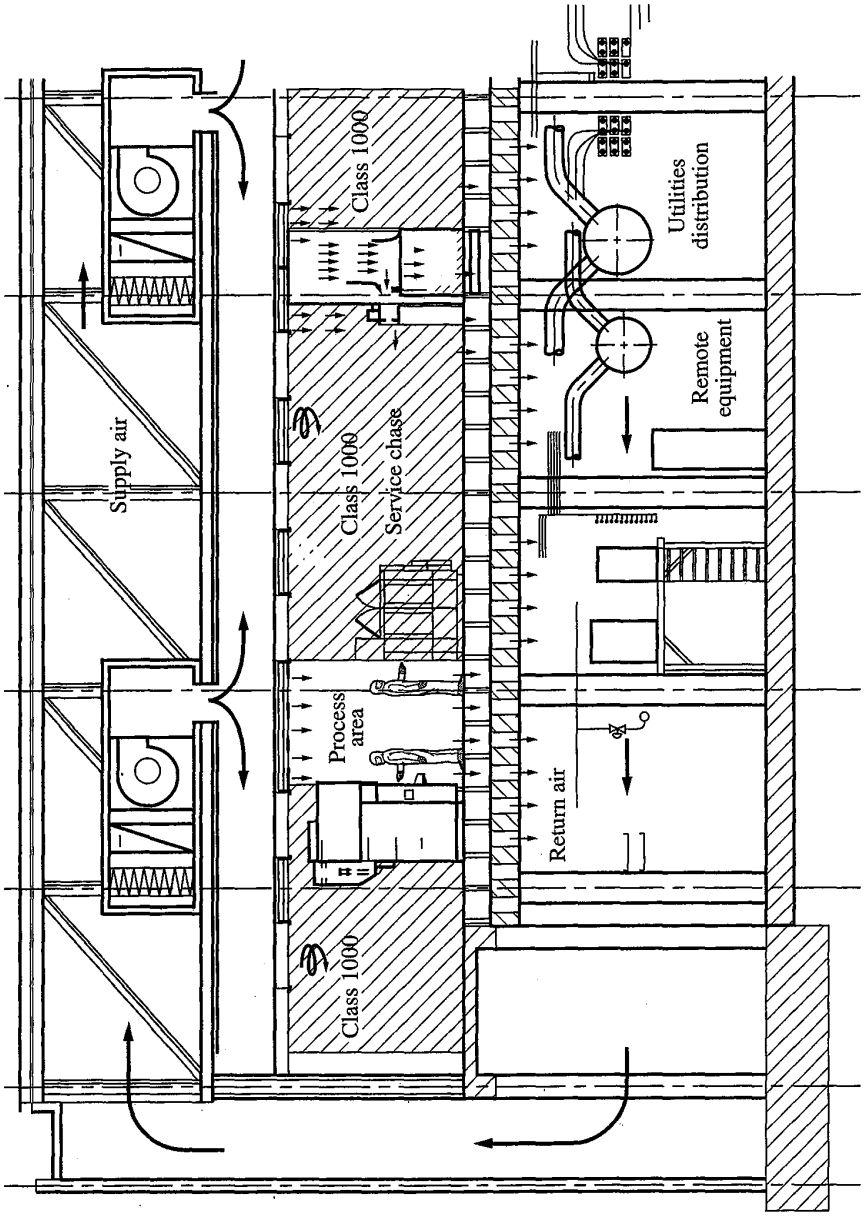


FIGURE 7
Cleanroom with centrifugal fan units installed on top of process level.

Micron Ex. 1023, p. 43
Micron v. YMTC
IPR2024-00909

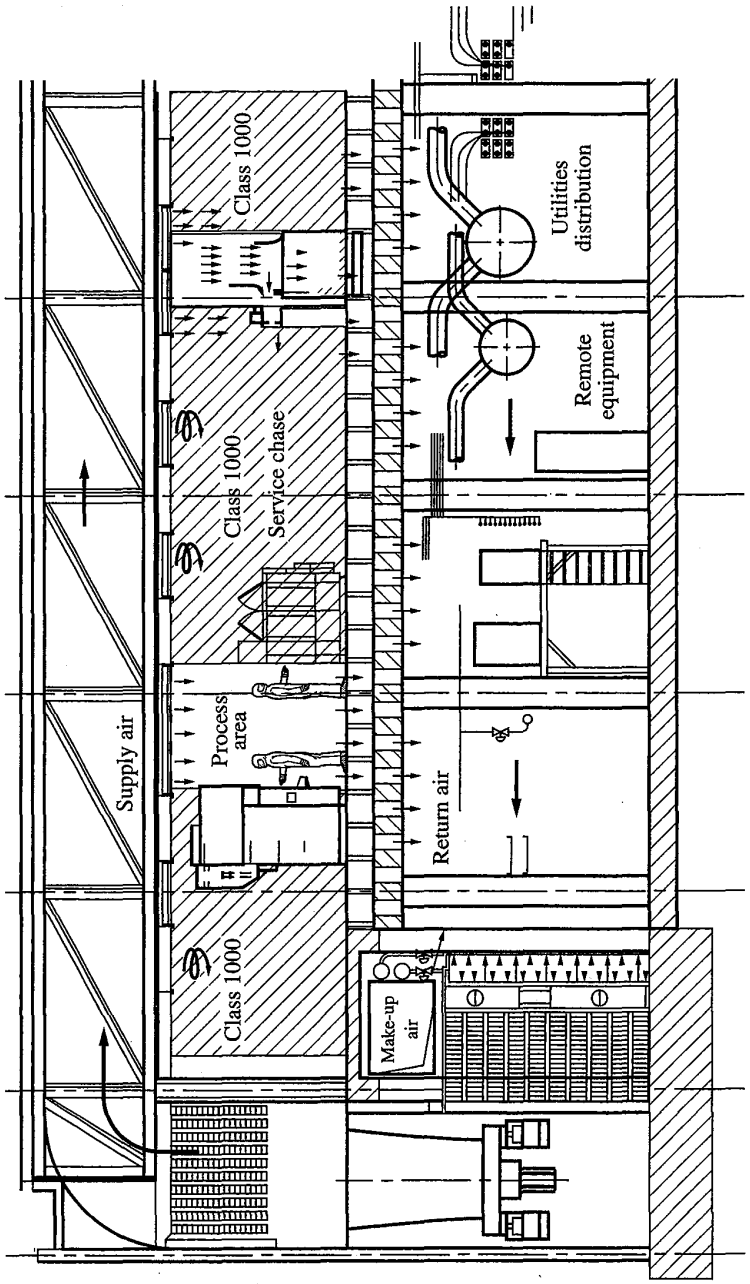


FIGURE 8
Cleanroom with axial fan units installed sideways connecting the air-supply plenum at the top and the air-return plenum at the bottom.

Micron Ex. 1023, p. 44
Micron v. YMTC
IPR2024-00909

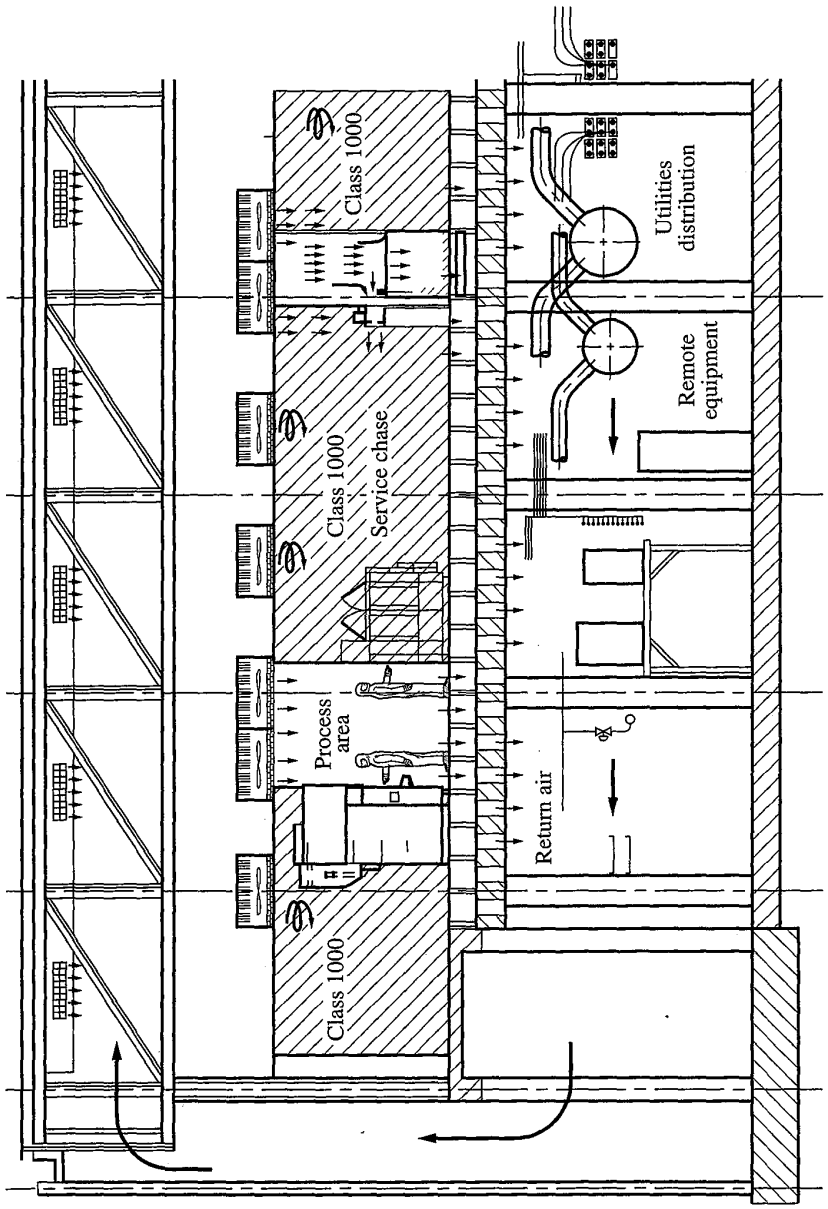


FIGURE 9
Cleanroom with filter fan units installed on the top of process area.

Micron Ex. 1023, p. 45
Micron v. YMTC
IPR2024-00909

plenum. Since the air pressure inside the cleanroom is higher than that above the filter ceiling, airborne particles cannot infiltrate the cleanroom area. As a consequence, less expensive "dry" ceiling systems can be used. (Referred also to the filter ceiling system described in Section 1.4.1.)

1.3.4 Cleanroom Performance Criteria

Some common parameters[25,26] and values for the quality characteristics of a typical cleanroom are summarized in Table 6 and the details are elaborated in the following section.

Cleanliness class

The cleanest class (M-1) defined in the Fed. Std. 209E allows for

10 particles/m^3 or 0.28 particles/ft^3 at 0.5 μm
350 particles/m^3 or 9.91 particles/ft^3 at 0.12 μm

The actual technical goal is far beyond the definition of the Fed. Std. 209E. In practice, cleanliness class is often defined for the MPPS (most penetrating particle

TABLE 6
Principal cleanroom design criteria

Design parameters	Criteria
Cleanliness, pieces/ft^3	
Process area ≥ 0.12 μm	≤ 1
Service area ≥ 0.30 μm	≤ 1,000
Temperature control, °C	
Photolithography area	22 ± 0.1
Other cleanroom area	22 ± 0.5
Humidity control, %	
Photolithography area	43 ± 2
Other cleanroom area	43 ± 5
Air quality, ppb	
Total hydrocarbons (THC)	< 100
NO_x	< 0.5
SO_2	< 0.5
Cleanroom interior surface outgassing rate, torr-L/cm^2/sec	6.3×10^8
Cleanroom pressure, relative to outside ambient, Pa	> 30
Acoustical noise, dB	< 60
Vibration (8 to 100 Hz), μm/sec	< 3
Grounding resistance, Ω	< 1
Magnetic field variation, mG	< ±1
Charging voltage, V	< ± 50

Micron Ex. 1023, p. 46
Micron v. YMTC
IPR2024-00909

size) of 0.10 μm to 0.12 μm. Figure 10[27] illustrates the filtration efficiencies of HEPA (high-efficiency particulate air) filters with respect to different particle sizes. Note that around 0.1 μm the filtration efficiency is the lowest.

The current cleanroom classification is, in general, defined as

Class 1	(at 0.12 μm) inside the process area
Class 1000	(at 0.3 μm) in the surrounding area

During cleanroom operation, in-line monitoring of particle density should be performed routinely. A practical goal is less than one particle/ft^3 with a size greater than 0.12 μm for critical process areas.

Temperature and humidity control

The recommended temperature control requirements are

metrology/lithography area	22 ± 0.1°C
other cleanroom area	22 ± 0.5°C

A relative humidity of 43% at 22°C can be readily achieved by using regular chilled-water systems with the chilled water inlet and outlet temperature controlled

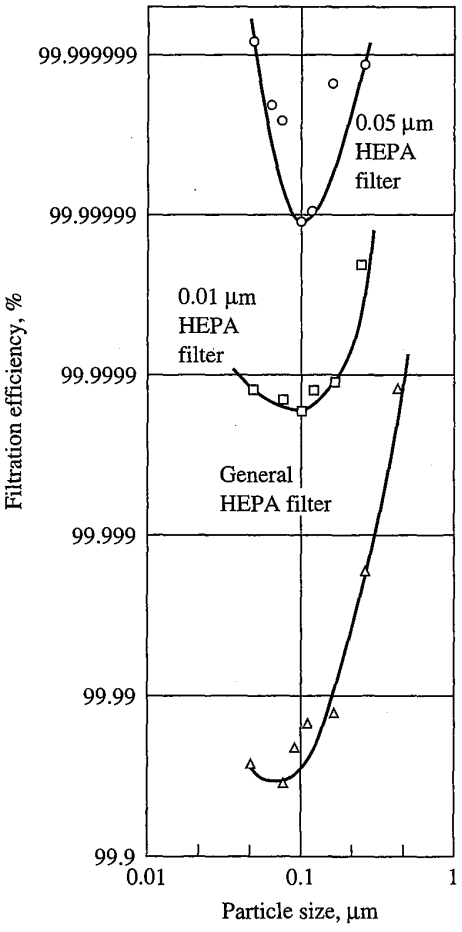


FIGURE 10
Filtration efficiencies of HEPA filters with respect to particle sizes.

Micron Ex. 1023, p. 47
Micron v. YMTC
IPR2024-00909

at 5°C and 12°C, respectively. Tolerances of 2% at 22 ± 0.1°C are achievable via dew point control in the make-up air units. Typical relative humidity requirements are

metrology/lithography area	43% ± 2%
other cleanroom area	43% ± 5%

Air quality

Hydrocarbons comprise ethylene, benzene, methane, xylene, etc. The outside hydrocarbon contamination, typically in a range of 1.5 ppm to 50 ppm, can be reduced approximately 90% by use of active-carbon absorbers in the make-up air units. An approximate 98% reduction can be achieved with new active pleated filters. The inside hydrocarbon contamination can be reduced by the selection of low outgassing material finishes for the walls and floor. Chemically treated carbon composite filters located in the air recirculation can achieve sub-ppm THC (total hydrocarbon) results of 75 to 80 ppb.

Other external molecular contaminants are SO_2, SO_3, NO_2, halogen (CCl_4, BCl_3, Br_2, Cl_2), and sulfur compounds such as H_2S. The typical measured values of NO_x and SO_x in outside air, depending on local conditions and the location, are 5 to 100 and 15 to 50 ppb, respectively. The measured NO_x and SO_x levels inside a cleanroom using ion chromatography, in general, are less than 0.5 ppb.

Outgassing

The maximum outgassing rate of 6.3×10^8 torr-L/cm^2-sec is a criterion for the selection of all materials in contact with the airflow in the make-up air and the recirculation air system. Outgassing is the evolution of a volatile compound from cleanroom materials or process tools, which may condense on the wafer and create problems due to organic and ionic contamination.

Cleanroom pressurization

The cleanroom pressure should be positive (e.g., +30 Pa) relative to the outside pressure. Cleanroom pressure is automatically controlled by adjusting the make-up air volume via variable-frequency drives in the make-up air units.

The dynamic pressure (velocity pressure) of the air should be as low as possible to minimize pressure variation. The pressure variation and low-frequency sound waves of 0.2 to 8 Hz are caused by acceleration (air return) and deceleration (supply plenum) of airflow. Turbulent flow is reduced by eliminating sound attenuators. It is recommended that recirculation air velocity be kept below 2.5 m/sec.

Acoustical noise

The technical goal is to have a cleanroom in an "as-built" condition designed to meet the PNC (Pressure National Code) 60-dB sound levels as defined by the 1987 American Society of Heating Refrigerating and Air-Conditioning Engineers (ASHRAE) recommendation. A cleanroom in an "as-built" situation means no process tools are operating, but all fans and HVAC systems are running. The challenge is to meet PNC 60 sound levels without installing sound attenuators in the recirculation fan units.

Micron Ex. 1023, p. 48
Micron v. YMTC
IPR2024-00909

Vibration isolation

Vibration criteria apply in both the vertical and horizontal directions at the top of the cleanroom waffle slab. Maximum vibration velocity should be no greater than 3 μm/sec, measured in 1/3-octave bands between 8 to 100 Hz. Below 8 Hz the criteria vary by 6 dB per octave. The criteria apply only at the photography and metrology area. Most steppers will perform with floor vibrations (waffle slab level) of 5 μm/sec, assuming that the criterion is met at the low end of the band at 8 and 10 Hz. Sensitive metrology equipment (such as the SEM) and defect measurement will determine vibration design criteria.

The entire clean production area should be provided with a waffle slab subdivided by joints. To meet the vibration limitations in the photolithography and metrology areas, the waffle slab is supported by 500 mm × 500 mm columns every 3.6 m. The column spacing in the other areas can be 7.2 m. The waffle slab typically has 500-mm-wide main ribs and 200-mm subribs with heights of 800 mm and 400 mm, respectively. The waffle grid is 600 mm × 1200 mm to match the raised floor and filter ceiling grids.

Grounding

Grounding is required for safety to protect people against a shock hazard. One overall grounding-grid system, buried approximately 400 mm below grade, should be installed. The grounding grids of the different buildings must be interconnected to form an equal potential grounding network.

The wye point (secondary side) of the transformers should be grounded via copper bars connected to the grounding grid. All major electrical equipment, such as power distribution panels and MCCs (motor control centers), must be grounded to the main grounding grid. Cable trays, pipe racks, metal tanks, handrails, supports, etc. should also be grounded.

A ground-grid bar is required for each power and distribution panel. All receptacles should be provided with a grounding pole. The grounding of the control system must be connected in a way that permits only a single path to the ground. The lightning-protection installation and the low-impedance grounding grid (see EMI) must also be connected to the grounding grid. The grounding-grid system should have a grounding resistance not exceeding 1Ω.

Electromagnetic interference (EMI)

Process tools are sensitive to electromagnetic interference reaching them through cables, pipes, or electrical and magnetic fields. At the same time, the tools generate conducted and radiated interference. Compatibility is achieved when the generated interference is kept below the sensitivity of process tools. The level of interference in the building is caused by external sources as well as internal sources. The effect of external sources can be limited by the building structure; the building should act as a shield against external sources of interference. This can be achieved by making the building a Faraday cage. Usually reinforced steel panels are installed over the roof and along the walls to provide the necessary shielding.

Doors and windows create leakage in the Faraday cage, so extra shielding is necessary. This can be achieved by creating other Faraday cages within our existing

Micron Ex. 1023, p. 49
Micron v. YMTC
IPR2024-00909

cage of the building. In this way, a number of zones can be provided to increase shielding.

The most important measure with respect to the floors, walls, and roofs is the use of planar grounds. A planar ground or low-impedance grounding grid should be installed under the raised floor in the clean production area.

Electrostatic discharge (ESD)

Static charges can develop on surfaces due to touching and rubbing (*triboelectricity*). The discharge of that electricity to machines and to people damages semiconductors, disturbs process tools, and may cause injury. ESD can be controlled by grounding all machines, by controlling the relative humidity, and by constructing walls and floor coverings of slightly conductive materials to route electrical charges to ground. The control of triboelectricity enhances dust and particulate contamination control. All metal racks, pipelines, cabinets, cable trays, and rails are grounded to an equal potential bar or planar ground. The metal pedestals of the raised floor connect to the planar ground under the raised floor. The metal framework of the cleanroom wall systems should be connected to this planar ground. To enhance ESD control, air ionization systems might be installed at certain locations in the process area.[28–33]

1.3.5 Principal Technical Data for Sizing the Cleanroom Facilities

Several key design parameters that are essential for sizing the cleanroom facilities of a 200-mm-wafer fab are summarized below. The dimensions of these parameters are all converted to units per cleanroom area.

Production capacity	85–120	mask × (wafer/month)/m^2
Exhaust	80	(m^3/hr)/m^2
Recirculation air	670	(m^3/hr)/m^2
Power for process equipment	800	W/m^2
Process cooling water	240	W/m^2
Heat load	450	W/m^2

The size of a cleanroom increases with process complexity and the photo mask level. Typically, the monthly production capacity of one square meter of cleanroom is around 85 to 120 mask steps, with DRAMs at the lower end of the range and ASICs (application-specific ICs) at the higher end.

The exhaust air rate is high; about 12% of the recirculation air is replaced by fresh make-up air. Approximately 65% of the exhaust rate is attributed to furnaces, chemical stations, and tube-cleaning equipment.

The average power requirement of process equipment is approximately 800 W/m^2. One-third of the heat generated is carried away by process-cooling water, and another third by the exhaust air. The rest of the heat dissipated by equipment, recirculating fans, and lighting is removed by the cooling coils installed in the recirculation air units.

Micron Ex. 1023, p. 50
Micron v. YMTC
IPR2024-00909

The power consumptions of major facilities are listed below:

Facility	Consumption
Office building	6%
Deionized (DI) water system	3%
Process tool	30%
Testing equipment	10%
Utility equipment	15%
Support cleanroom	3%
Fab recirculation fans	7%
Boilers	8%
Chillers	18%

The power consumption of the office building is relatively small, whereas up to 40% of the power demand is used to run the process tools and testing equipment and the remaining 50% is used to support cleanroom operations, including pumps, recirculation fans, lighting, boilers, and chillers.

Other quantities of facilities required to complete the processing of a 200-mm, 16 Mbit DRAM wafer are as follows:

Chemicals	10 kg
Deionized water	4.5 ton
Compressed dry air	55 m^3
N_2	25 m^3
O_2	0.9 m^3
H_2	0.1 m^3
Power	470 Kwh

1.4 CLEANROOM INSTALLATION

The cleanroom installation section starts with an introduction of the basic information on the cleanroom subsystems followed by the discussion of the cleanroom installation approach. The investment cost required for constructing a semiconductor fab is analyzed last.

1.4.1 Cleanroom Subsystems[25, 34–36]

The air-handling principle of cleanrooms is based on the interrelated design of the process exhaust systems, the make-up air system, and the recirculation air system. Chilled water and hot water systems are required for temperature and relative humidity control. The filter ceiling system, the cleanroom partition system, and the raised floor system constitute the enclosing of a cleanroom area. The features of all these systems are described in the following paragraphs.

Micron Ex. 1023, p. 51
Micron v. YMTC
IPR2024-00909

Process exhaust systems

The following main process exhaust systems can be identified:

General exhaust
Scrubbed exhaust
Solvent exhaust

The general exhaust and the solvent exhaust systems typically comprise ductworks, exhaust fans, bypass, and stacks. In addition, the scrubbed exhaust system includes scrubbers. The bypass allows entry of outside air if the pressure at the suction side of the operating fans exceeds a preset value.

The general exhaust system removes heat dissipated by the process equipment. This exhaust air should not contain acids, caustics, or solvents. The solvent exhaust system removes air containing solvents from the process equipment. The exhaust fans employed should be explosion-proof. The scrubbed exhaust system removes air containing acids and/or caustics from the process equipment. Ammonia, toxic, and silane exhaust systems are, in most cases, also connected to the scrubbed exhaust system upstream of the scrubber.

There are two types of scrubbers, namely, dry and wet scrubbers. A wet scrubber is an exhaust air treatment unit that removes the acids and caustics from the exhaust air by washing them out. The water used is typically city water. Waste water from the scrubber is routed to the neutralization plant of the waste treatment area.

A dry scrubber removes toxicants from the exhaust air by absorbing them into the scrubber material at a desired temperature. If the scrubber material is saturated, it has to be replaced.

Make-up air system

The make-up air system is designed to maintain the cleanroom pressure, to compensate for the air losses from building enclosure leaks, and to compensate for the process exhaust. Another function of the make-up air system is control of humidity in the cleanrooms, i.e., by humidification or dehumidification, depending on the process environment and ambient conditions. For exact adjustment of the make-up air volume with the exhaust rates, frequency convectors are incorporated in both process exhaust systems and the make-up air system.

A make-up air unit, illustrated in Fig. 11, withdraws outside air through an air intake grill and removes the majority of airborne particles via a prefilter. After passing through an air-cooling coil to remove excess moisture, the cold air is heated by an air heater prior to entering an air fan sandwiched between two sound attenuators. Fine particles are removed by a HEPA filter. The air is regulated by an air-volume controlling damper before it is fed into the make-up air ductwork, which distributes the make-up air to the suction side of the recirculation air units. During winter or dry season when the humidity is low, steam generated by a humidifier is added to the air stream to maintain the moisture at a designated level.

Recirculation air system

The recirculation air system serves three purposes:

Micron Ex. 1023, p. 52
Micron v. YMTC
IPR2024-00909

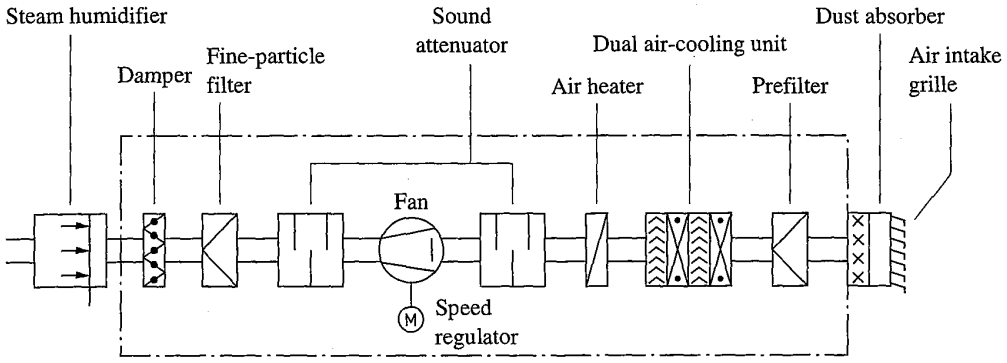


FIGURE 11
Schematic diagram of the make-up air unit.

Temperature control
Particle control
Air flow control

Depending on the cleanroom design principle, three types of recirculation air units can be used:

Filter fan units
Centrifugal fan units
Axial fan units

A filter fan unit consists of an enclosure with a fan and a final filter assembly. If required, the filter fan unit can be provided with a dry-cooling coil to control the temperature in the cleanroom area within very tight margins. Filter fan units can provide a high degree of flexibility; however, the cost for a given air-handling capacity is also the highest among the three types of recirculation units. The filter fan units are best suited for a small cleanroom without enough space to accommodate the big recirculation air units and the necessary ductwork.

A centrifugal fan unit recirculates air by using a centrifugal fan and removes particles and heat by passing air through air filters and a set of cooling coils. The entire air-handling unit is sandwiched between a pair of sound attenuators to cut down the sound level. The conditioned air is routed to the filters in the ceiling grid via either ductwork (more expensive solution) or a plenum.

In the case of an axial fan unit, the recirculating air is withdrawn through a series of components: a prefilter, a cooling coil, a sound attenuator, and an air-volume controlling damper, all located in the basement of a cleanroom for easy maintenance. The conditioned air is then routed through a vertical mounted axial fan and the second sound attenuator before it is distributed directly into the air supply plenum.

Both centrifugal and axial fan systems can be applied successfully within the current cleanroom performance criteria. Naturally, the applications of both systems have pros and cons when they are compared. This means that each application requires its dedicated measures. In general, the ultimate selection of centrifugal or axial fans is not determined by the comparison of advantages and disadvantages but by the chosen cleanroom concept.

Micron Ex. 1023, p. 53
Micron v. YMTC
IPR2024-00909

The application of axial fans is most suitable when a fan bay concept is adopted. The selection of a cleanroom system with recirculation air units on top of the cleanroom leads to centrifugal fan systems and a separate air supply plenum. The key pros and cons of axial fans compared with centrifugal fans can be summarized as follows:

Higher investment level but lower running cost
Higher sound pressure levels at relatively high frequencies
More frequent and more comprehensive maintenance
Easier mounting due to compact sizes
Easier air flow adjustment
Directly driven by a motor instead of by particle-generating belts
Vibration generated at 19 to 30 Hz instead of below 17 Hz

Chilled water system[37]

The total cooling capacity of all make-up air units, recirculation air units, air coolers, ventilation units, and central and process utility systems is provided by a chilled water system. The dehumidification operation in the make-up air unit requires a 5 to 11°C temperature range to remove the excess moisture from the air (wet-cooling principle). On the other hand, the temperature of the cooling coils employed in the recirculation air units needs to be controlled above the dew point of 9°C to prevent any moisture from condensing so that the humidity inside the cleanroom can be maintained at a constant level (dry-cooling principle). By mixing the cooling water supply and return, a temperature of 14 to 18°C can be obtained.

A chiller consists of a condenser and an evaporator. In most cases, the chiller will be a water-cooled type. The required water cooling is provided by a cooling tower. Inside the cooling tower, the cooling water is sprayed downward to meet the uprising outside air drawn in by a fan mounted on top of the cooling tower. A small portion of water evaporates as the water travels through the cooling tower. The evaporation of water demands heat and, thus, heat is removed from the cooling water. The evaporated cooling water must be replaced.

In many cases, the cooling water required for process equipment will also be taken from the chilled water system using a heat exchanger. The primary side of this process cooling water system connects to the chilled water system, the secondary side is designed as an "open" system to keep the water pressure in the cooling water return lines as low as possible. For certain process tools such as sputters, the low return cooling water pressure is essential for preventing water from leaking into the process chamber.

In an open process cooling water system, the recirculating cooling water is returned to a holding tank opened to atmosphere. The level control sensors installed in the water holding tank allow a supply of deionized water to compensate for water losses due to evaporation. To avoid accumulation of particles in the systems, filter units are incorporated in the return lines.

Hot water system

The total heating capacity of all make-up air units, ventilation units, and certain process utility systems is provided by a hot water system. The hot boilers heat water to 90°C. The return temperature is, typically, 70°C.

Micron Ex. 1023, p. 54
Micron v. YMTC
IPR2024-00909

Filter ceiling system

The filter ceiling system of the clean production area comprises ceiling grid framework, filters, blanks, and auxiliary steel structures fastened to the steel trusses or the concrete ceiling. The ceiling grid framework covers the entire clean production area and is suspended from the auxiliary steel structure by height-adjustable hanging rods to permit leveling. The grid is based on a 1200 × 600 mm system, i.e., on the size of the filters and blanks. The filter ceiling framework offers full flexibility with respect to the location of cleanroom partitions, filters, blanks, sprinkler heads, and lighting fixtures.

There are two kinds of filter ceiling systems, namely, the fluid filter ceiling system and the dry filter ceiling system. The designation of fluid or dry ceiling is related to the type of sealing. Figure 12 illustrates that the airtight sealing between filters and a fluid ceiling grid is provided by a self-healing, highly viscous, silicon-based sealing gel. In the case of a dry filter ceiling, the airtight sealing of filter assemblies and blanks onto the ceiling grid is provided by an elastic gasket. The filters and blanks are secured with fasteners.

The fluid ceiling system is typically applied in a cleanroom with an overpressure above the filter ceiling (air supply plenum) which must be kept absolutely airtight to prevent the unfiltered air from entering into the highly classified cleanroom. If the pressure inside the cleanroom area is higher than that above the filter ceiling, there is no risk of particulate infiltration, and a dry filter ceiling, which is not completely leaktight, can be applied. In the dry filter ceiling system, the pressurized air required for each filter is supplied either by a fan located right above the filter fan unit or by a flexible duct connected to each filter housing.

Although various manufacturers use different designations for filters, the names "HEPA" and "ULPA" (Ultra Low Penetration Air) are of the most frequent occurrence. In general, ULPA is used for filters having an efficiency of 99.9995% and more at particulate diameters > 0.12 μm. Filters with lower efficiencies are typically designated as HEPA.[38] For this reason, we recommend specifying a type of filter stating the efficiency and particle size. Details of HEPA filter construction are illustrated in Fig. 13.

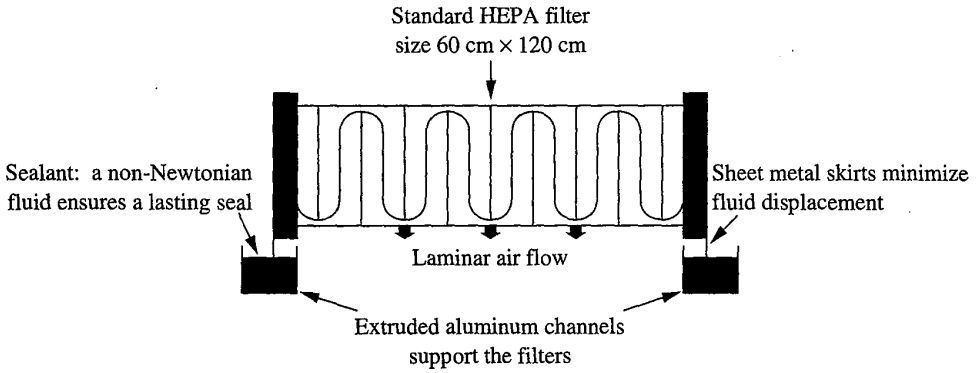


FIGURE 12
Positive fluid-sealing method for mounting HEPA filters.

Micron Ex. 1023, p. 55
Micron v. YMTC
IPR2024-00909

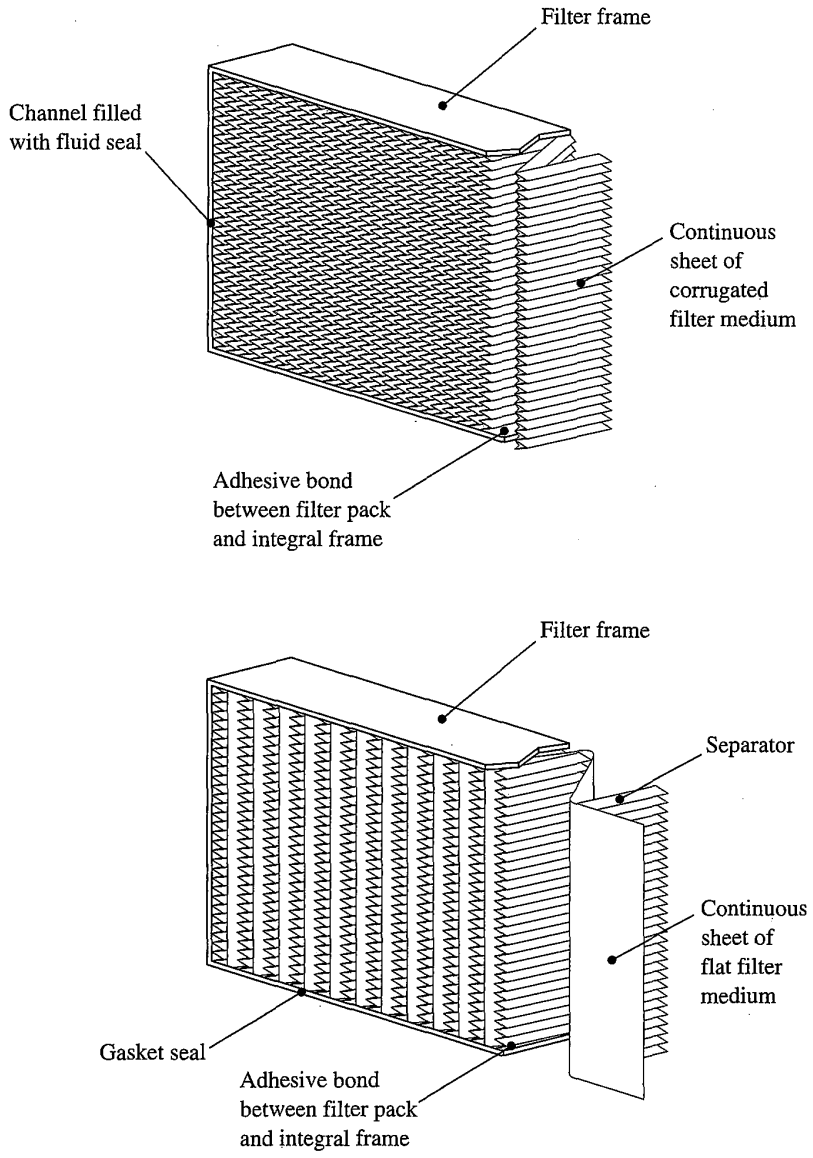


FIGURE 13
Details of HEPA filter construction.

The selection of filters depends on the level of cleanliness required. The cleanliness level in a cleanroom is controlled not only by the grade of air filter used but also by the amount of dust generated and the air flow pattern in the cleanroom. Therefore, filters of the same grade may result in different actual performance.

Cleanroom partition system

Partition walls separate the process area from the service area in the cleanroom. The entire cleanroom partition system should be grounded to meet ESD requirements. A typical cleanroom partition system consists of a framework, single-walled

Micron Ex. 1023, p. 56
Micron v. YMTC
IPR2024-00909

or double-walled panels, glass panels, and door units. The framework is constructed of extruded aluminum. The wall panels are typically metal sheet coated with a conductive paint. The glass panels are made of conductive clear glass or laminated plastic complete with associated sealings. The door units with glass have gaskets around the door posts and at floor side. The surface of the metal panels is smooth, cleanable, impact resistant, and does not generate or shed particles.

The partition wall system is constructed with flexible connections to minimize the transmission of vibrations between the filter ceiling grid and the partitions. All panels are easily dismountable and interchangeable to ensure optimum flexibility with respect to layout changes.

The cleanroom partition system has a total height equal to the distance between the raised floor and the filter ceiling (3.4 to 3.6 meters) and is 50 to 60 mm thick depending on the ceiling grid framework width. The cleanroom partition system, including all its components, must meet the isolation resistance requirement of 10^6 to $10^8\Omega$.

Raised floor system

The raised floor should be installed 450 to 600 mm above the finished concrete waffle slab, and in general cover the entire clean production area. The grid of the raised floor is based on a 600×600 mm system and should be aligned with the centerlines of the filter ceiling grid. Some of the floor tiles should be perforated. The adjustment of the air-pressure drop and the balancing of the air flow in the clean production area can be achieved by selecting the floor tiles with proper perforating ratio.

The floor tiles should be static-dissipative and made of noncombustible vinyl that is chemical- and abrasion-resistant, impact-resistant, and smooth and meets the ESD isolation resistance requirement of approximately $10^6\ \Omega$.

The raised floor should be laterally stable in all directions with or without tiles present. Floor tiles are supported at each corner by height-adjustable pedestals. These pedestals are glued to or bolted on the finished concrete waffle slab. An insulation plate, placed on top of each pedestal, attenuates footstep sound and ensures electrical conductivity. To avoid horizontal vibrations, some of the pedestals are reinforced with bracing to increase the rigidity.

1.4.2 Construction Project Approach and Schedule

Cost effectiveness is the key starting point in the design of a semiconductor plant. This means that the final result in terms of quality, cost, and timing must be ensured during the entire course of the project.

The baseline-management method offers an excellent guideline. According to this method, the entire project is subdivided into clearly identifiable phases, from broad to detailed, and the progress in each phase is recorded. This provides a total picture of the project at any time to help project managers to make critical decisions.

A project can be subdivided into the following main phases: the initial design, the tendering, the elaboration (preparation for the execution, construction, documents, etc.), the realization, and the process-tool hookup.

Micron Ex. 1023, p. 57
Micron v. YMTC
IPR2024-00909

The initial design constitutes the first and most important project step. During this phase, all starting points will be translated into an integrated building and services concept with the associated specifications, cost, and time indicated. The target project result is, thus, defined clearly at the beginning. All subsequent phases are a further elaboration of these initial concepts.

In general, the construction of an IC fab, from the start of the design activities to the first wafer out, extends over approximately two years. Design of facility systems and tendering of the relevant contracts takes 4 to 6 months. The construction of buildings and installation of all the utility systems take another 12 to 14 months. The final 4 to 6 months will be required to certify the facility systems, to install process equipment, and to test or tune processes.

The enclosed time schedule contains the timing of each project phase, subdivided into its essential parts. The development of site and buildings must receive the appropriate priority in the initial phase. When the fab-module shell is closed and the interior thoroughly cleaned, the installation of the cleanroom systems can start. In addition, the site infrastructure (roads) must be partly completed at a very early stage to reduce dust and to provide access to the fab.

The time schedule can be summarized by the following milestones:

Event	Cumulative months
Start initial design	0
Release conceptual design of the site/buildings	2
Release initial design document	3
Release preliminary design of the site/buildings	4
Receive approved building permit	6
Award contract for site/buildings	6
Release preliminary design of all facilities	6
Break ground	6
Start construction of buildings	7
Award contract for cleanroom/utilities	8
Award contract for mechanical/electrical work	8
Start cleanroom/utilities	15
Complete installation of cleanroom/utilities	20
Start equipment move-in/hookup	20
Complete testing/certification	21
Complete process-tool hookup	23
Complete functional testing	24
Start-up	24

1.4.3 Cost Analysis

After the specifications, the design criteria, the layout, and the conceptual designs of all subsystems are decided, the costs and time schedule can be determined accurately. Both are highly dependent on the dimensions and quality level. It is, therefore, incorrect to apply general characteristic values, e.g., the investment sum per square

Micron Ex. 1023, p. 58
Micron v. YMTC
IPR2024-00909

meters of cleanroom. Such standard figures are only applicable within a clearly defined frame of reference, e.g., IC fabs with a clean production area of some 3000 square meters projected for submicrometer technologies.

Approximately 75% of the investment is attributed to process equipment, whereas the remaining 25% is required for all the facilities including process utilities. The following table outlines the cost of each facility system:

Building	25%
Cleanroom systems	25%
Central utilities	10%
Process utilities	35%
Others	5%

A few years ago, the cleanroom systems constituted the most expensive item. To date the cost of the cleanroom systems has decreased mainly because of better knowledge, whereas the cost of process systems in the last years has increased substantially due to more stringent purity requirements and technology developments. Except for the building, the cleanroom, which houses process tools and test equipment, is the most expensive item. The average life span for the process equipment, the facility system, and the building is estimated at 5, 10, and 20 years, respectively. The depreciation costs can be calculated as follows:

Process equipment	88%
Facility system	10%
Building	2%

It is obvious that any significant reduction of facility costs, which may ultimately impair the yield of the wafers, has little effect on the manufacturing costs. It is, however, important to pursue realistic and soundly based expectations for quality and quantity and to exclude any unrealistic goals and overkilling, in order to make the facility systems cost-effective.

The facility systems should be designed to support at least three generations of process technology without major renovation. The following design principles are recommended:

- The quality of cleanroom, process chemicals, DI water, and process gases should be of state-of-the-art.
- The facility system should be highly flexibile with respect to future modification, tool change, and automation.
- The facility system should be capable of continuous operation while the facilities are undergoing expansion or modification.

1.5 CLEANROOM OPERATIONS

Preventive maintenance of all the electrical and mechanical components is absolutely necessary to prevent unscheduled shutdown in the cleanroom. Continuous monitoring of particles, temperature, and humidity inside the cleanroom is required to alert process engineers to changes in the cleanroom environment so that steps

Micron Ex. 1023, p. 59
Micron v. YMTC
IPR2024-00909

may be taken to prevent temperature- and humidity-sensitive processes from drifting out of control. Proper gowning procedure and cleanroom maintenance practices are critical to prevent microcontamination.

1.5.1 Cleanroom Gowning

People are major sources of cleanroom contaminants. Not only do they generate a large number of contaminants but they are also in close proximity to the wafers at many stages of the IC manufacturing process. Therefore, an appropriate gowning procedure is necessary to minimize the exposure of hair, bare skin, and street clothes.

It has become a common practice for many IC manufacturers to require its personnel to change from street clothes and street shoes into company-provided clean undergowning and shoes before entering a cleanroom gowning area. In the gowning area, personnel are obliged to cover hair, nose, mouth, and even the eyes with a hair net, a face mask, and goggles. Gloves, a hood with an integral face mask, a cleanroom bunny suit, and booties are then worn to achieve as complete body coverage as possible.[39] Good cleanroom suit material is normally made of woven fabrics consisting of long synthetic fibers covered with a layer of PTFE (polytetrafluoroethylene) laminate material. This semiporous PTFE material can prevent particles from passing through but at the same time allow a high degree of vapor transmission.[40] The cleanroom suits and shoes should be washed regularly using deionized water and sodium-free detergent. The in-house or externally contracted laundry service should take appropriate measures to limit contamination while washing, packaging, transporting, and storing cleanroom garments.[41] All cleanroom clothing, including gloves and boots, should be electrically conductive to minimize the accumulation of electrical charges. Before entering a process area, a hand wash using DI water to remove particles from gloves may be necessary to prevent contamination.

1.5.2 Cleanroom Maintenance

Proper cleanroom maintenance is critical to minimize microcontamination. Some common industrial cleanroom maintenance practices are listed below.[42–47]

1. All personnel working in the cleanroom should receive special training in gowning, wafer handling, cleaning, and safety procedures.
2. Makeup and body powder is prohibited in the cleanroom.
3. Wool, fur collars, and jewelry should never be worn inside the cleanroom.
4. Pencils, clip ballpoint pens, noncleanroom paper, woods, aerosol spray cans, and boxes made of cardboard are prohibited in cleanroom areas.
5. Documents or photos that cannot be reproduced on cleanroom paper have to be laminated in a conductive PP (polypropylene) film before they are brought into the cleanroom.
6. Parts, boxes containing wafers, chemical bottles, and equipment should be thoroughly cleaned in a classified decontamination room adjacent to the cleanroom before they are brought into the cleanroom.

Micron Ex. 1023, p. 60
Micron v. YMTC
IPR2024-00909

7. Ceiling grids and cleanroom walls should be mopped with DI water regularly. The floor should be washed with DI water several times each day. On average, one person is needed to clean each 1000 square meters of cleanroom.
8. Free-standing shield panels should be set up around process equipment during installation and maintenance to contain as much as possible any contamination that may be generated.
9. After equipment maintenance, dust or debris should be cleaned with a central vacuum cleaning system or a portable vacuum cleaner provided with a HEPA filter.
10. Maintenance hand tools should be cleaned in an ultrasonic bath to remove particles attached to their surfaces.
11. After they are detached from equipment, dirty parts should be covered before removal from the cleanroom.
12. Wafer cassettes or wafer boxes should be hand-carried one at a time. Movement of wafer cassettes by cart causes potential contamination problems due to vibration.
13. Pumps and equipment that have potential to generate contamination should be installed in a cabinet provided with an exhaust pipe connected to the general exhaust system.

1.6 AUTOMATION

With the ever-increasing process complexity, a process technology comprising some 20 mask levels and between 300 to 400 recipes becomes quite common. The number of WIPs (wafers in process) inside an IC fab routinely exceeds 30,000 pieces. It becomes critically important to implement automation to resolve the common IC manufacturing problems listed in Table 7.

TABLE 7
Common IC manufacturing problems and relevant solutions

Problem	Suggested actions
Wrong lot went to the equipment	Online lot validation
Operator cannot find cassette required	Automated stoker control
Operator selects wrong recipe	Real-time recipe download
Errors in manually recorded data	Data upload from equipment
Lot tracking not performed promptly	Real-time lot tracking
Equipment not correctly monitored	Real-time equipment monitoring
Schedule does not reflect fab condition	Real-time scheduling (just in time)
Inefficient recipe/bay setup	Recipe scheduling
Equipment capacity not fully utilized	Automation
Operators fail to report problem	Real-time process data management

Micron Ex. 1023, p. 61
Micron v. YMTC
IPR2024-00909

Three levels of automation can be distinguished:[48,49]

1. *Shop-floor-control software.*[50] A shop-floor-control software resident in a computer is used to assist the operations to manually control process equipment and transfer wafer cassettes between stations.
2. *Computer Integrated Manufacturing (CIM).*[51] Automatic control of process equipment but manual transfer of wafer cassettes by operators.
3. *Mechanization.*[52–55] Automatic control of equipment and transfer of wafer cassettes directly to process tools by robots.

1.6.1 Shop-Floor-Control Software

The shop-floor-control software developed either in-house or purchased from outside vendors has been used extensively by almost every manufacturer to replace process run cards on which process steps and recipes were listed. This software usually has the following features:

1. Recipe management
2. Lot management
   - WIP control
   - Line balancing
   - Exception lot handling (holds/releases, splits/merges, scraps)
3. Scheduling and dispatching
   - Material priority handling and coordination
   - Just-in-time scheduling and dispatching
   - Dynamic rescheduling based on changing conditions
4. Data collection
   - Equipment status
   - Measurement data
5. Cost analysis
6. Analysis and reporting
   - Lot history
   - SPC (statistical process control) chart
   - Production report

A good shop-floor-control program can significantly improve cycle time and reduce WIP by optimizing lot batching, scheduling, and dispatching.

1.6.2 Computer Integrated Manufacturing

To reduce human-induced errors and enhance the equipment productivity, many fabs built in the late 1980s started to link process equipment to a host computer. The following equipment control features have been implemented:

Micron Ex. 1023, p. 62
Micron v. YMTC
IPR2024-00909

Bar code reader
Equipment setup
Recipe download
Process data acquisition and upload
Equipment state and status
Equipment alarm management
Equipment maintenance support

The semiautomatic fab operation mode can improve equipment utilization rate and fab yield considerably. Real-time equipment monitoring can minimize equipment idle time and provide early warning of potential equipment failures and problems. Automatic recipe download can improve wafer yield by eliminating recipe errors. It can also reduce equipment downtime resulting from incorrect operations.

1.6.3 Mechanization

The interbay automation system has been adopted by many IC manufacturers to automatically transfer and track wafer cassettes. Automatic stokers located in each process tunnel are linked by an overhead clean rail system. Wafers are transferred automatically from one process tunnel to the other on a vehicle driven by a linear motor. An interbay automation system can avoid confusion and save considerable time in tracking a wafer within a fab.

Intrabay automation is the last step to achieve a fully automatic operation mode. An AGV (automatic guided vehicle), shown in Fig. 14, is used to transfer wafer cassettes automatically from a stocker directly to process equipment. Intrabay automation can reduce the number of operators required. It can also significantly reduce wafer scratches and breakage by eliminating manual handling of the wafers.

However, the intrabay automation system has not been very popular outside Japan due to relatively high installation and maintenance costs. In the 300-mm wafer fab of the future, the wafer cassette might be so heavy that the intrabay automation system becomes absolutely necessary.

1.6.4 SECS Equipment Interface

The establishment of communication between a host computer and process equipment for downloading the processes recipe and monitoring of equipment status is a necessary step for factory automation. Programming the communication software, in general, follows the SECS (Semiconductor Equipment Communication Standard) protocol published by SEMI (Semiconductor Equipment Manufacturing Institute).

The SECS standard allows manufacturers to produce equipment that can be linked to any host computer without specific knowledge of the computer. The standard also allows host system designers to program a host computer and to plan a computer network without specific knowledge of the equipment functions. In addition, system designers can use the standard to connect different pieces of semiconductor equipment to each other.

Micron Ex. 1023, p. 63
Micron v. YMTC
IPR2024-00909



FIGURE 14
Automatic wafer transfer to process tool using AGV. (*Courtesy of Murata of Japan.*)

There are two different sets of SECS standards: SECS-I and SECS-II. The SECS-I standard describes the physical connections, signal levels, data rate, and logical protocols used to exchange messages between a host computer and process equipment. The SECS-II standard defines the contents of the messages passed between the host computer and the process equipment.

The SECS-I stipulates that equipment should be connected to the host computer by a serial communication cable with an RS-232C connector. Through the RS-232C cable, data signals can be transmitted over a distance of more than 15 meters.

An SECS-II message consists of 13 to 257 bytes of 8-bit data. Every byte of data represents a character or a single-digit number. Each message block begins with a single byte giving the total length of the message block, followed by a 10-byte message header. The message header contains the ID number of a particular piece of equipment with which the host computer intends to establish communication. It also identifies the type of message data, such as an alarm code, equipment status code, etc. Up to a maximum of 244 bytes of message data follows immediately after the message header. Every SECS-II message ends with a 2-byte checksum.

The checksum computed by the host computer adds the binary value of all the preceding data bytes in an SECS-II message. After receiving the incoming SECS-II message, the equipment repeats the computation. If the data transmission is noise-free, the checksum computed by the equipment should match exactly that sent by the host computer.

Micron Ex. 1023, p. 64
Micron v. YMTC
IPR2024-00909

1.6.5 Cost Analysis

For every $100 in equipment investment, an additional amount of approximately $15 is needed to implement a fully automatic system in a 200-mm fab. A cost breakdown of each level of automation is illustrated below:

Shop-floor-control software and hardware	10%
Interbay automation	20%
Semiautomatic process equipment interfacing	30%
Intrabay automation	40%

When an IC fab is operated in a fully automatic mode, the fab utilization rate can be increased up to 10%, the fab yield can be increased about 5% and the number of operators can be reduced by 75%. Therefore, the additional investment in fab automation is justified economically.

1.7 RELATED FACILITY SYSTEMS

Apart from the cleanroom subsystems described in Section 1.4, other process utility systems can significantly affect the wafer yield. The quality characteristics and the evolution of the specifications and the features are discussed in the following subsections.

1.7.1 DI Water[56–63]

Large quantities of DI (deionized) water are required to process IC wafers. The consumption of DI water increases with the wafer size. The DI water and chemical tanks of the earlier-generation 200-mm-wafer wet-benches are twice as large as those of a 150-mm-wafer wet-bench. As a result, the DI water and chemical consumption of a 200-mm-wafer IC fab are also doubled.

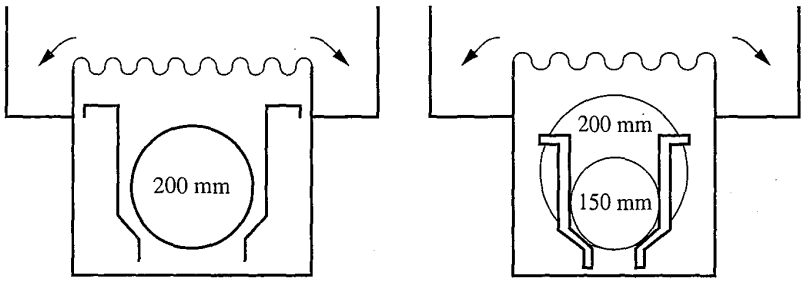


FIGURE 15
Comparison of the DI water tanks of cassette and cassetteless chemical stations.

Micron Ex. 1023, p. 65
Micron v. YMTC
IPR2024-00909

Since the early 1990s, the cassetteless wet-benches have been developed. In these wet-benches, wafers are unloaded from wafer cassettes to a quartz boat, which is subsequently picked up by a robot arm. The robot arm dips the boat into a series of chemical and DI water tanks in a way similar to a conventional wet-bench. Figure 15 illustrates that the sizes of the DI water tanks of a 200-mm-wafer cassetteless wet-bench are about the same as a conventional 150-mm-wafer cassette type wet-bench. Table 8 also confirms that the consumption of DI water and chemicals of a 200-mm-wafer cassetteless wet-bench is close to that of a conventional 150-mm-wafer wet-bench. For each 200-mm, 16-Mbit, DRAM wafer produced, approximately 4.5 tons of DI water are required.

A state-of-the-art DI-water installation consists of two loops:

The make-up loop removes particles and colloids, total organic carbons, microorganisms, ionic impurities, and total dissolved solids from the raw water. In other words, it has a pretreatment section for removal of gross particles of sizes larger than 1 μm and a purifier section for removal of ionic impurities, bacteria, and dissolved gases.

The polishing loop removes the last traces of contaminants that originate in the construction materials or come from secondary microbial contamination.

There are many contaminants in raw water, such as particles, organic material, inorganic material, microorganisms, bacteria, and dissolved gases. As the raw water is fed through the DI water system, all of these contaminants are removed sequentially by a series of different types of filters, a high-vacuum degasifier, and an ion-exchange unit. Figure 16 shows the filtration spectrum of different types of filters employed in a DI water system.

The operating principles of multimedia filters, percoated filters, and microfilters in the make-up loop are similar. The filters let raw water pass through and trap the suspended solids and colloids in the filter media. When the pressure drop across the filter reaches a predefined threshold value, the dirty filtration media is regenerated

TABLE 8
Comparison between a 150-mm-wafer chemical station and a 200-mm-wafer chemical station

	Wafer Size			
	150 mm		**200 mm**	
Wafer transfer	Cassette	Cassette	Cassette	Cassetteless
Number of cassettes	Double	Single	Double	Double
Chemical tank, L	27.6	29.5	57.5	25
DI water tank, L	19	18	45	20
Particle				
DI water batch				
Density, pieces/L	30	30	30	10
Size, μm	0.3	0.3	0.3	0.16
Spin dry				
Density, pieces/L	10	10	10	10
Size, μm	0.3	0.3	0.3	0.16

Micron Ex. 1023, p. 66
Micron v. YMTC
IPR2024-00909

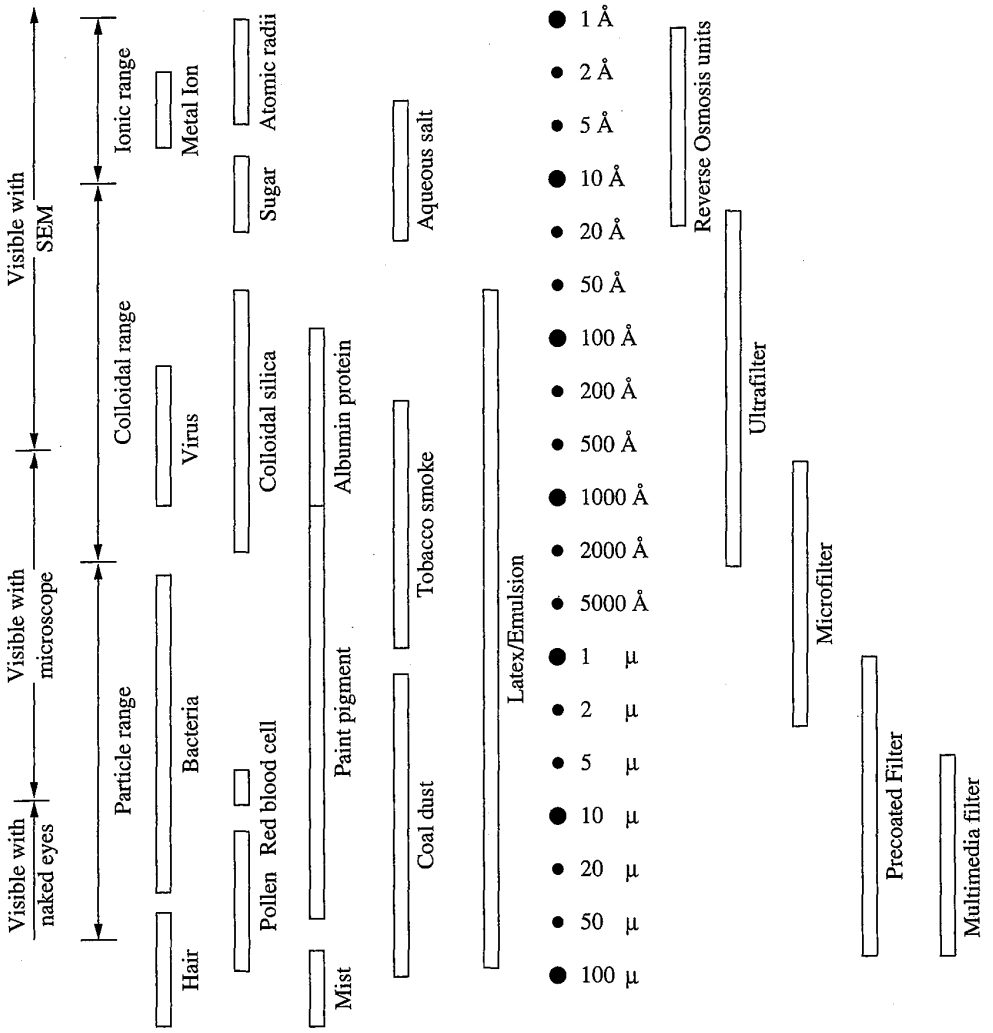


FIGURE 16
Filtration spectrum of different types of filters. (*Courtesy of Christ AG. Switzerland.*)

through a back-washing operation. Chemicals such as $FeCl_3$ are often injected at the filter inlet to enhance the agglomeration of suspended solids and thereby improve the filtration efficiency.

RO (reverse osmosis) units are used to remove smaller particles and metallic ions. The operating principle of RO units is illustrated in Fig. 17. As water enters the spiral-wound modules, water molecules diffuse through the porous membrane but microparticles, microorganisms, and ions are held back by the membrane. The rejected impurities are drained in the concentrated filtration effluent, and very few impurities are held back inside the membranes of the RO units. Conversely, multimedia filters and microfilters keep the captured particles in the filter media. RO units are regenerated by injecting RO permeate water on their feed side to wash away the sediments attached to the porous membrane.

Ultrafilters are used to remove submicrometer-size particles. During normal filtration mode, water is fed to the outside of a band of hollow fibers. As the water

Micron Ex. 1023, p. 67
Micron v. YMTC
IPR2024-00909

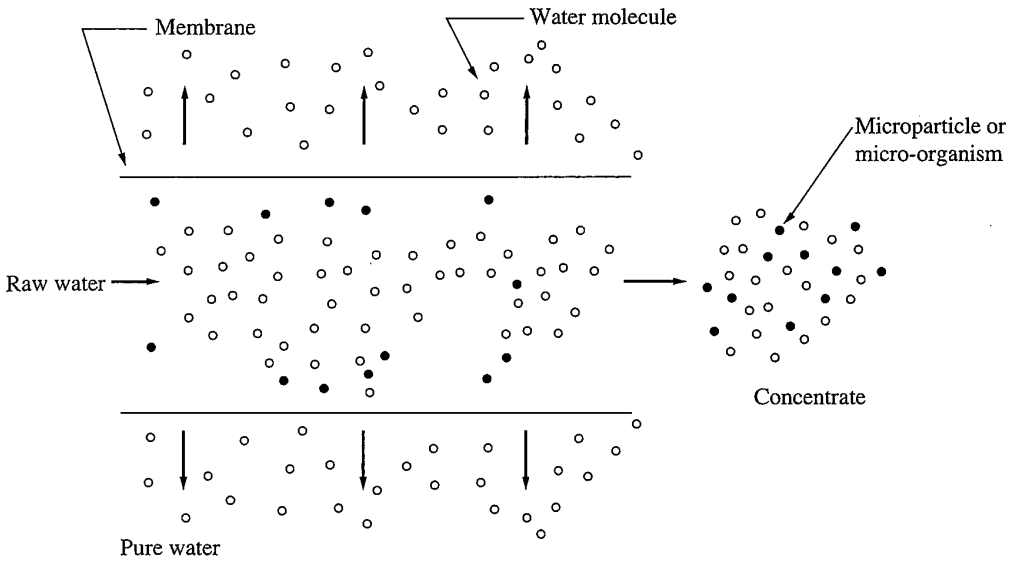


FIGURE 17
Operation principle of reverse osmosis filters.

penetrates through the porous fibers, particles are left behind on the exterior surface. During regeneration, the particles are washed away by the purified water fed in the reverse direction. A cross section of a hollow ultrafiltration fiber and a schematic diagram of an ultrafiltration cartridge are illustrated in Fig. 18.

The majority of TOC (total organic carbons) is removed in the make-up loop via the reverse osmosis units and the remaining portion of TOC is removed in the polishing loop by first ionizing it in the UV lamp modules and later capturing the ionized TOC in the mixed-bed polishing cartridges.

In the polishing loop, the DI water will be continuously circulated at a velocity greater than 1.5 m/sec from the process area back to the DI water installation to maintain low levels of contamination. All piping behind the polishing cartridge is constructed of PVDF (polyvinylidene fluoride) to minimize the release of particles and TOC.

PEEK (polyether ether ketone), a plastic tubing, and GOLDEP, an electropolished stainless steel piping passivated by Cr_2O_3 film, are reported to have much lower TOC- and metallic-ion dissolving characteristics compared to PVDF, and they have also been used in the polishing loop.

The evolution of DI water system design, shown in Fig. 19, can be characterized by the introduction of new techniques and materials.

For 256K DRAM devices, the DI water system includes the following features:

Pretreatment filters for removing particles in the micrometer range
RO units for removing ions, TOC, and particles
A vacuum degasifier for removing dissolving gases
Use of 254 nm UV lamps for bacterial sterilization
Use of mixed-bed ion exchange unit for removing residual ions
Use of membrane filters to capture particles

Micron Ex. 1023, p. 68
Micron v. YMTC
IPR2024-00909

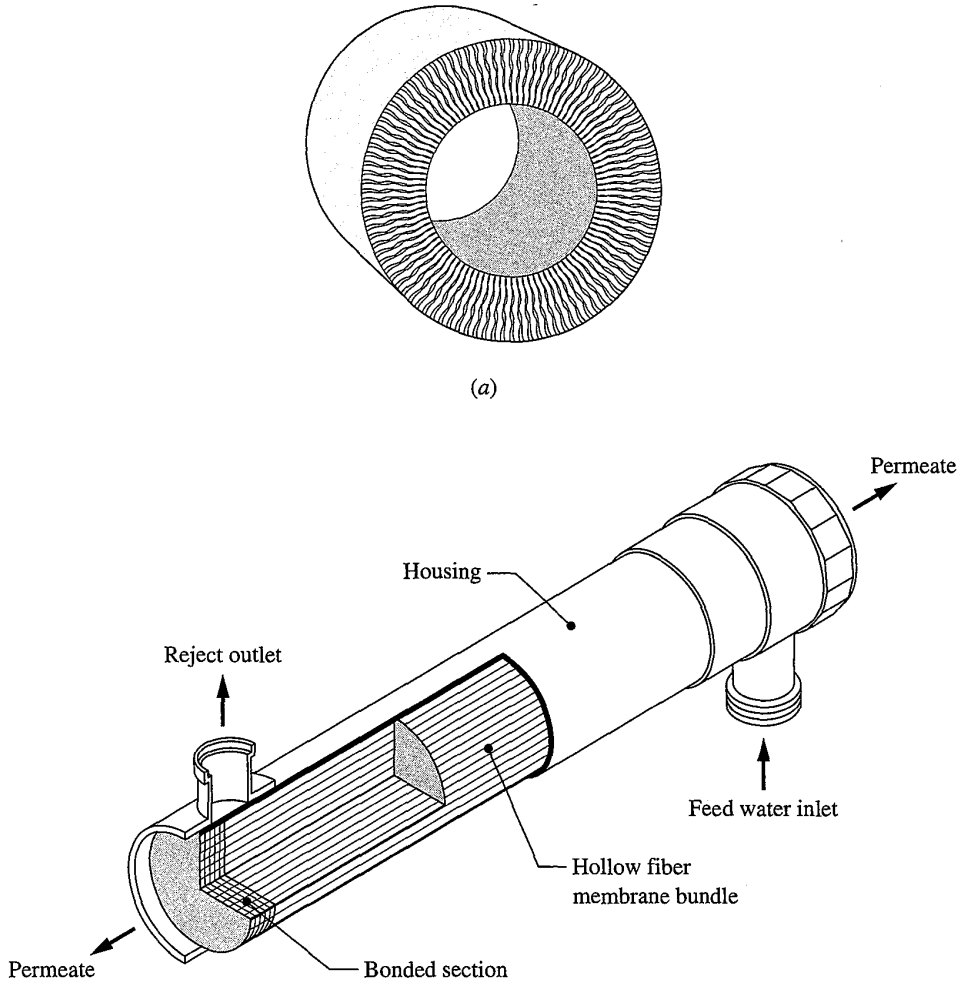


FIGURE 18
(*a*) Schematic diagram of the cross section of an ultrafiltration fiber. (*b*) Schematic diagram of an ultrafiltration cartridge. (*courtesy of Asahi Chemicals of Japan*)

For 1M DRAM devices, the following additional features are included:

Use of polishing cartridges with low TOC leachables
Use of ultrafiltration units as final filter
PVDF supply piping instead of PVC (polyvinyl chloride) or PP
Periodical sterilization of the distribution loop with ozone.

For 4M DRAM applications, the use of 185 nm UV lamp sterilization to reduce for TOC was adopted. For 16M DRAM, the DI water system process flow is similar to that for the 4M DRAM except that the RO and ultrafiltration filters are of better quality. For 64M DRAM applications, the use of secondary RO units in the make-up is forecast.

Micron Ex. 1023, p. 69
Micron v. YMTC
IPR2024-00909

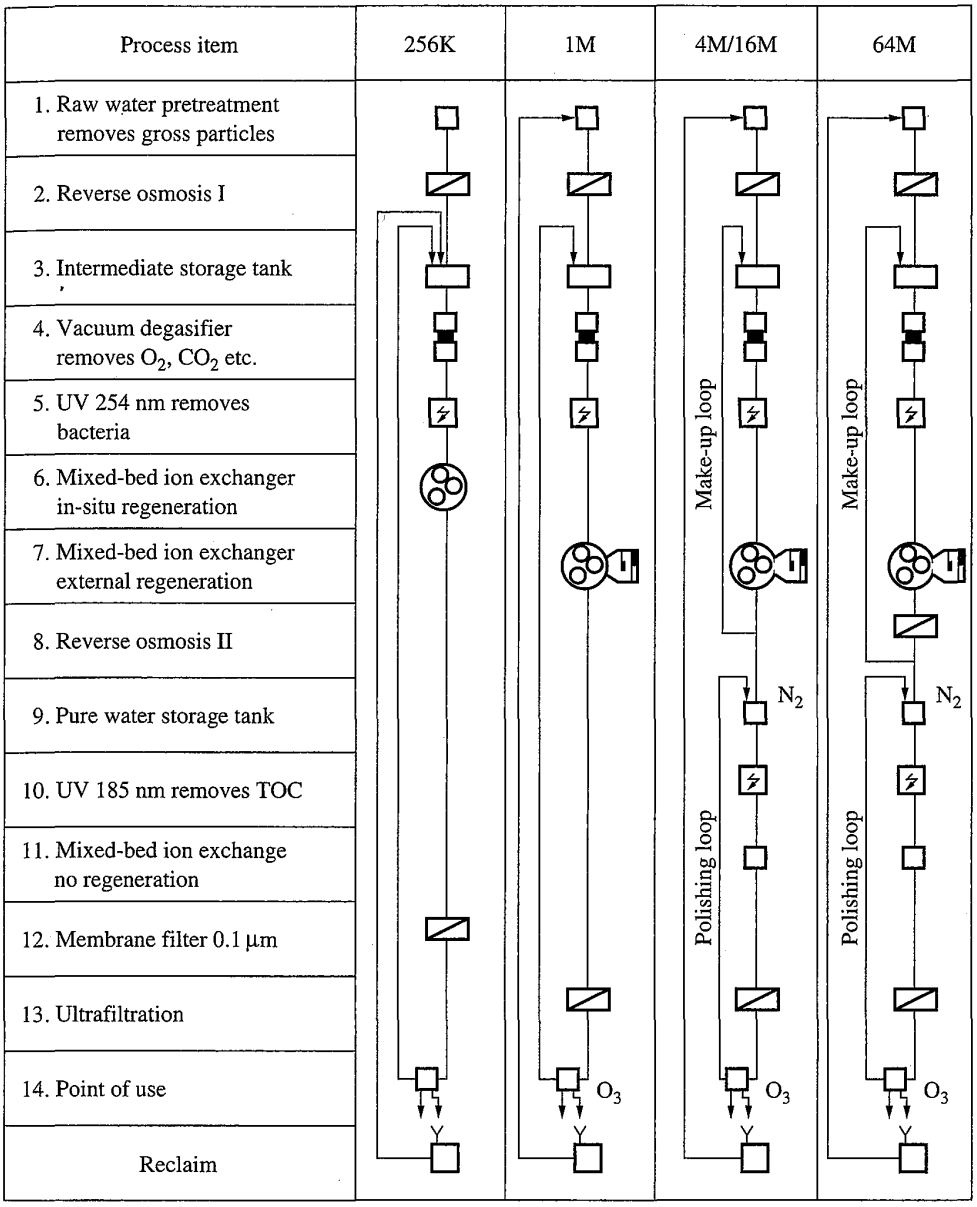


FIGURE 19
Evolution of ultrapure water system design. (*Courtesy of Christ AG. Switzerland.*)

Evolution of the DI water specifications is illustrated in Table 9. The specifications of most parameters, except for resistivity, have been tightened one order of magnitude for the 1-Mbit DRAM to the 16-Mbit DRAM.

One of the most critical impurities in DI water is the dissolved oxygen, which has to be kept as low as possible to prevent native oxide growth on bare silicon wafers. Much work has been done to improve the performance of vacuum degasifiers. There are three popular degasification processes: the hot-water process involves heating DI water to a temperature above 55°C; the nitrogen purging process entails injecting nitrogen counter flow into a DI water vessel; and the catalytic process involves

Micron Ex. 1023, p. 70
Micron v. YMTC
IPR2024-00909

TABLE 9
Evolution of DI water specifications

	Process technology				
	256K DRAM	**1M DRAM**	**4M DRAM**	**16M DRAM**	**64M DRAM**
Design rule, μm	2.0	1.2	0.8	0.5	0.3
Resistivity, MΩ-cm	> 17.0	> 17.5	> 18.0	> 18.1	> 18.2
Particle (pieces/cc)					
> 0.2 μm	< 30	< 10			
> 0.1 μm	< 50	< 20	< 5	< 1	
> 0.085 μm			< 10	< 2	< 0.5
> 0.05 μm				< 5	< 1
Bacteria, unit/L	< 200	< 50	< 10	< 1	< 1
TOC, ppb	< 100	< 50	< 20	< 5	< 1
Oxygen, ppb	< 100	< 100	< 50	< 10	< 5
Silica, ppb	< 10	< 5	< 3	< 1	< 0.2
Na, ppb	1	< 1	< 0.1	< 0.05	< 0.01
Cl, ppb	1	< 1	< 0.1	< 0.05	< 0.01
Metal ion, ppb	1	< 1	< 0.1	< 0.05	< 0.01

Source: Courtesy of Christ AG. Switzerland.

feeding DI water through a vessel packed with palladium catalytic resin. These three degasifiers are expensive. Recently, a new degasifier operated at a room temperature of 23°C and employing neither nitrogen nor a catalyst has been developed. A vacuum level beyond 700 mm Hg is generated externally. The dissolved oxygen is removed by the boiling water vapor bubbling through the DI water. The dissolved oxygen level can be kept around 2 ppb.

Over 70% of the DI water supplied to the points-of-use can be recovered.[60] The reclaimed water can be fed directly to the raw water storage tank or pumped to an RO permeate storage tank after passing through various treatment steps such as activated carbon filters, a TOC UV lamp module, ion-exchange units, and RO units.

1.7.2 Chemical Supply System[64–65]

Close to 20 types of ultrapure chemicals are used in IC manufacturing processes. About two-thirds of them are acid and caustic and the remaining are organic solvents. These chemicals are used for the following processes:

Process	**Chemicals**
Prefurnace cleaning	NH_4OH, H_2SO_4, HCl, H_2O_2 and IPA
Thin film wet etching	HNO_3, H_3PO_4, HF, BHF (buffered HF)
Photoresist coating and developing	Photoresist, developer, acetone
Residual photoresist removal	H_2SO_4, H_2O_2, organic resist stripper
Postmetal process cleaning	Organic solvent

Particles and especially metallic impurities in the chemicals will significantly effect the wafer yield. The metallic impurities will cause degradation of gate oxide integrity and thereby result in early breakdown. The metallic impurities also enhance OSIF (oxidation induced stacking faults), resulting in junction leakage and a decrease of carrier lifetime.

Micron Ex. 1023, p. 71
Micron v. YMTC
IPR2024-00909

TABLE 10
Evolution of chemicals specifications

	Process technology				
	256K DRAM	**1M DRAM**	**4M DRAM**	**16M DRAM**	**64M DRAM**
Design rule, μm	2.0	1.2	0.8	0.5	0.3
Particles (pcs/cc)					
> 0.5 μm	< 100				
> 0.3 μm		< 50	< 10		
> 0.2 μm		< 500	< 50	< 10	< 1
> 0.1 μm				< 100	< 10
Anion, ppb	< 2000	< 1000	< 500	< 100	< 50
Metal, ppb	< 100	< 50	< 10	< 1	< 0.2

Source: Courtesy of Merck-Kanto.

With the ever increasing chip size and process complexity, the continuing reduction in gate oxide thickness, and the ever tightening device leakage current requirement, the demand on the qualities of chemicals becomes very severe. Table 10 shows the evolution of the chemical specifications with respect to process technologies.

In submicron fabs, all bulk chemicals, except the highly viscous photoresist, are centrally delivered to points of use via chemical supply systems. A chemical supply system, as shown in Fig. 20,[65] consists of three parts: the chemical delivery modules normally located in a separate room adjacent to the cleanroom, the chemical delivery piping network, and the central controller.

There are two types of chemical supply modules. One is a big storage tank with a 1000 to 9000 L capacity equipped with external pumps and filters. This type of chemical supply module is used to supply chemicals consumed in large quantities, such as H_2SO_4, NH_4OH, H_2O, IPA, and photoresist developer. Chemicals are pumped from chemical drums supplied by the vendors into the big holding tank and then are delivered to the points of use as needed.

The other type of chemical supply module, used for delivering chemicals used in smaller volume, is a cabinet unit that houses two chemical drums, each with a 200 L capacity. One of the two chemical drums serves as a supply tank and the other as a standby. The chemical supply module automatically switches to a standby tank when the supply tank is empty. HEPA filters are often installed inside the chemical supply cabinet to provide a clean, laminar-flow environment.

It is a common practice to periodically recycle the chemicals inside the chemical drum through an external filter, when the tank is not delivering the chemicals, to reduce the particles. When the chemicals are delivered by a pneumatically driven bellows pump, particles trapped inside the filter may be pushed out due to the pulsating action of the pumps. To overcome this problem, some advanced chemical-supply modules employ a combination of the pumped and pressurized delivery techniques.[64] When the chemical supply module is in the supply mode, the chemical is delivered out of the holding tank by a constant nitrogen source, and when the chemical supply module is in the standby mode, the chemical is recycled through external pumps to filter out particles.

The piping networks delivering acid and caustic chemicals are made of PFA (perfluoroalkoxy) pipe enclosed in a seamless clear PVC tube that is connected to

Micron Ex. 1023, p. 72
Micron v. YMTC
IPR2024-00909

FIGURE 20
Schematic diagram of a chemical sypply system.

Micron Ex. 1023, p. 73
Micron v. YMTC
IPR2024-00909

a valve box. Sensors are installed at the bottom of the valve box to detect chemical leakage. The piping for solvent is constructed of electropolished SS316L tubing.

Each chemical delivery module and all interface controllers, each connected to a point of use, are linked to a master controller that monitors the status of every delivery module and activates the chemicals supply function on receiving the request from any point of use.

1.7.3 Process Gas Piping System[66–73]

Over 20 different ultrapure gases are used in semiconductor manufacturing processes. Among these gases, N_2, O_2, H_2, and Ar are often called bulk gases because large quantities are used, whereas other gases, including He, AsH_3, PH_3, SiH_4, NH_3, and NF_3, are classified as special gases because only small quantities are needed.

In general, N_2 comes from three sources: an N_2 pipeline delivered from a remote air-separation plant, a liquefied-nitrogen storage tank located in the gas yard, and an on-site N_2 generation plant. O_2 and Ar are supplied from liquefied gas storage tanks. H_2 is delivered from either a liquefied-gas storage tank or a bank of high-pressure gas cylinders. These bulk gases pass through purifiers and gas filters to remove impurities and particles before entering a gas-distribution piping system located inside the cleanroom. The special gases are sent directly to process tools from gas cylinders located inside gas cabinets.

The gas cabinet is an exhausted safety enclosure that contains the gas cylinders and the associated gas-handling panels. The basic function of the gas cabinet is to allow purging and safe exchange of the cylinders. The gas panels incorporate all components for control and monitoring of the high-purity gases. The gas cabinet usually contains two process cylinders to allow automatic switch-over when one cylinder is empty. One nitrogen cylinder is also available for purging the piping line.

With a few exceptions, such as Cl_2 and SiH_2Cl_2, the supply pressure for bulk and special gases is normally kept around 5 kg/cm^2. At each point of use, the pressure of each gas has to be independently and locally controlled by a series of valves, pressure regulators, pressure sensors, and particle filters located inside a gas manifold box. Each gas is distributed at a special pressure dictated by the process tool. For each process tool there are one or more gas manifold boxes installed nearby for gas distribution and control.

The evolution of bulk-gas specifications with respect to process technology is shown in Table 11. From the 1M DRAM to the 4M DRAM, particles and impurities levels have to be reduced by one order of magnitude. To achieve the tight process-gas quality specifications required by a 16M DRAM process technology, the selection of gas components, piping-system design, and installation procedures becomes extremely critical.

Ultraclean gas components with the following features are used extensively in the submicrometer level process-gas piping systems:

Minimum dead space
Metal seal to prevent any leak to outside
Interior surface with submicrometer-level roughness

Micron Ex. 1023, p. 74
Micron v. YMTC
IPR2024-00909

TABLE 11
Evolution of bulk gas specifications

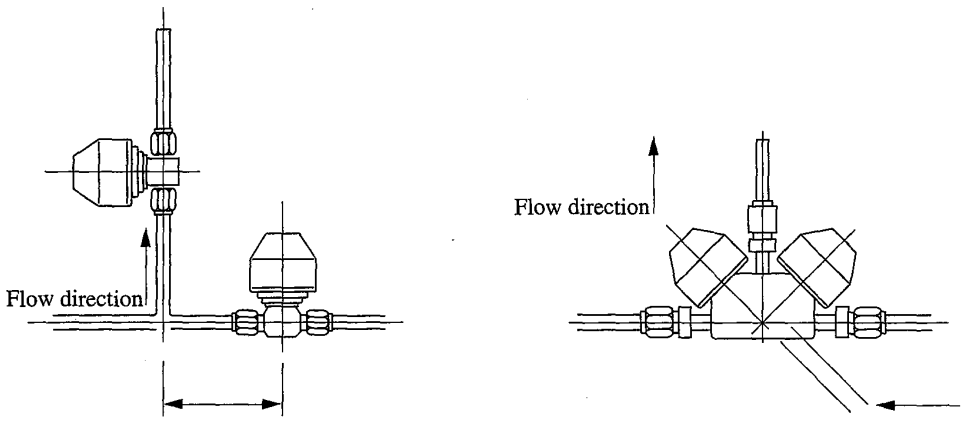
	DRAM technology				
	256K	**1M**	**4M**	**16M**	**64M**
Design rule, μm	2.3	1.3	0.8	0.5	0.3
Particle level					
Size, μm	0.3	0.2	0.1	0.1	0.05
Density, pcs/cc	< 10	< 10	< 5	< 5	< 5
Impurity level, ppb					
O_2	< 100	< 50	< 10	< 5	< 1
CO			< 10	< 5	< 1
CO_2			< 10	< 5	< 1
CH_4			< 10	< 5	< 1
H_2O level					
Dew point, °C	< −76	< −80	< −90	< −100	< −120
Concentration, ppb	1,000	500	100	10	0.13
Metallic level, μg/m³	–	1	1	0.1	0.01

Source: Courtesy of BOC of England.

No thread seal and no moving parts
Fittings with bearing or welding applied to all joints

All of these features are aimed at minimizing particles and adsorbed moisture from the interior wall that makes contact with the ultrapure process gases, reducing the trapping of unwanted gas in the dead space, and preventing ambient moisture and oxygen from leaking into the piping systems. The performance characteristics of the ultraclean gas components can be illustrated by the following examples.

Figure 21 shows the schematic diagrams of a conventional branch piping method and an integrated gas valve.[67] The dead space of the former is 1.3 cc and of the latter



(*a*) Two-way piping valve (*b*) Integrated valve

FIGURE 21
(*a*) Conventional branch piping method and (*b*) newly developed integrated gas valve.

Micron Ex. 1023, p. 75
Micron v. YMTC
IPR2024-00909

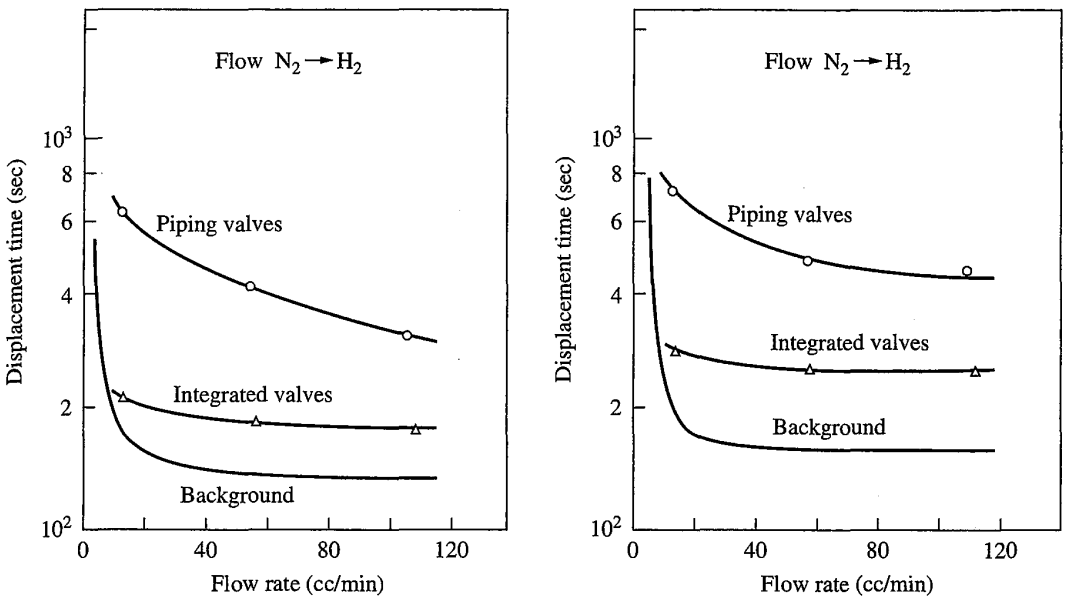


FIGURE 22
Gas displacement time (*a*) for branch valve and (*b*) for integrated valve.

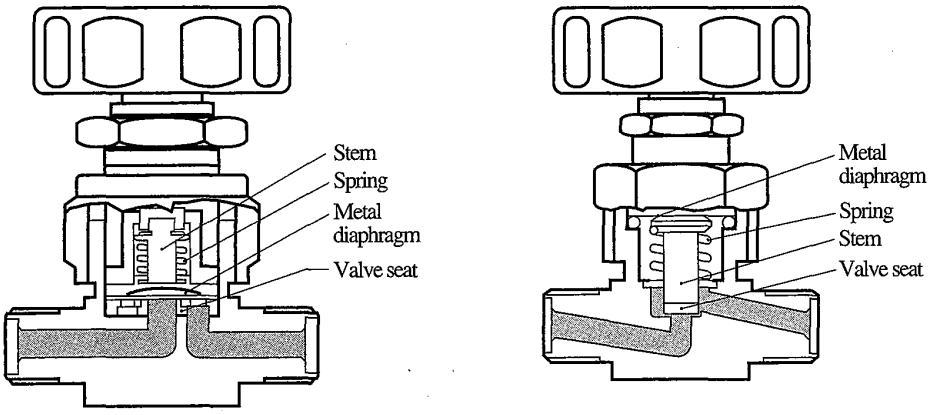
is 0.01 cc. As illustrated in Fig. 22, the integrated valve, which has a dead space two orders of magnitude lower, offers much better gas displacement characteristics than the conventional branch piping method. Recently, block valves that have several integrated single valves in a single metal block have been used by some IC manufacturers in their process-gas distribution systems.

Hundreds of diaphragm valves and bellows valves are used in a process-gas piping system. Schematic diagrams of these two valves are shown in Fig. 23. Compared with the bellows valve, the diaphragm valve has much lower dead space, and therefore exhibits superior gas-displacement characteristics.[67]

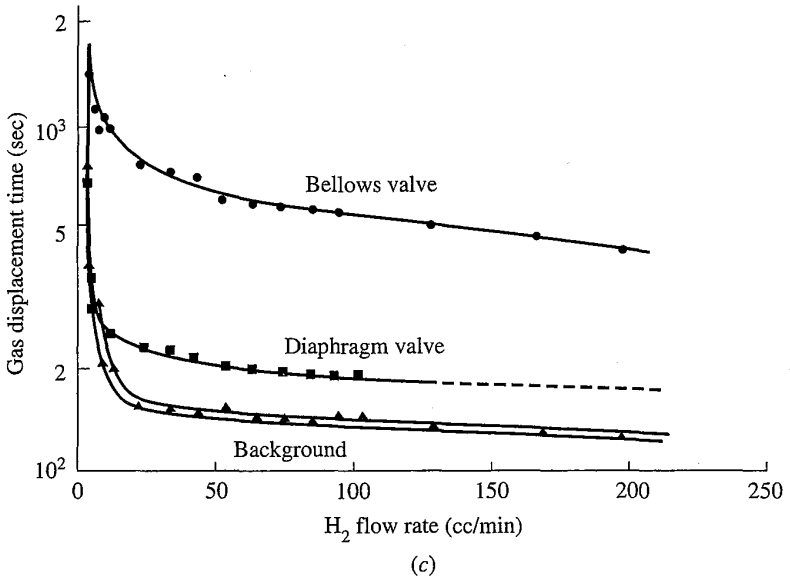
In the past decade, gas components with an interior surface electropolished to a roughness below 1 μm have become a standard. The moisture outgassing characteristics of an EP (electropolished) surface can be improved further by growing a Cr_2O_3 passivation layer in an oxygen ambient. The oxygen-passivated EP piping has a golden color and is also known as GOLDEP piping. Figure 24 shows the moisture outgassing characteristics observed in the start-up phases of three different gas-piping systems.[71] Within a few hours of purging, the moisture level of the gas piping system employing GOLDEP tubing and gas components with all-metal seals (see case C) drops to a few ppb. It will take more than a month of continuous N_2 purging for the moisture level of the gas-piping system using conventional EP tubing and gas components with plastic seals to drop below 5 ppb. Figure 24 also illustrates the importance of purging continuously after the welding work is completed.

Table 12 summarizes the gas components and their relevant characteristics, such as interior surface roughness and external leakage rate, with respect to process technology ranging from 256K DRAM to 64M DRAM. The surface roughness of

Micron Ex. 1023, p. 76
Micron v. YMTC
IPR2024-00909



(*a*) Diaphragm valve (*b*) Bellows valve



(*c*)

FIGURE 23
(*a*) Diaphragm valve, (*b*) bellows valve, (*c*) gas displacement characteristics.

gas piping and the external leakage of gas components has been improved by several orders of magnitude. The bellows valves have given way to diaphragm valves. The only plastic parts remaining in the gas-piping system are the membrane gas filters installed at the beginning of the main gas piping. Compared with the total gas flow rate in the main pipe, the outgassing rate of a membrane filter is negligibly small. The conventional Borden tube pressure gauge contains a spiral hollow coil that moves the pressure indicator needle by expanding and contracting as the pressure fluctuates. This type of spiral coil has a large dead space and also generates particles. The Borden tube pressure gauges are being replaced by flow-through type pressure

Micron Ex. 1023, p. 77
Micron v. YMTC
IPR2024-00909

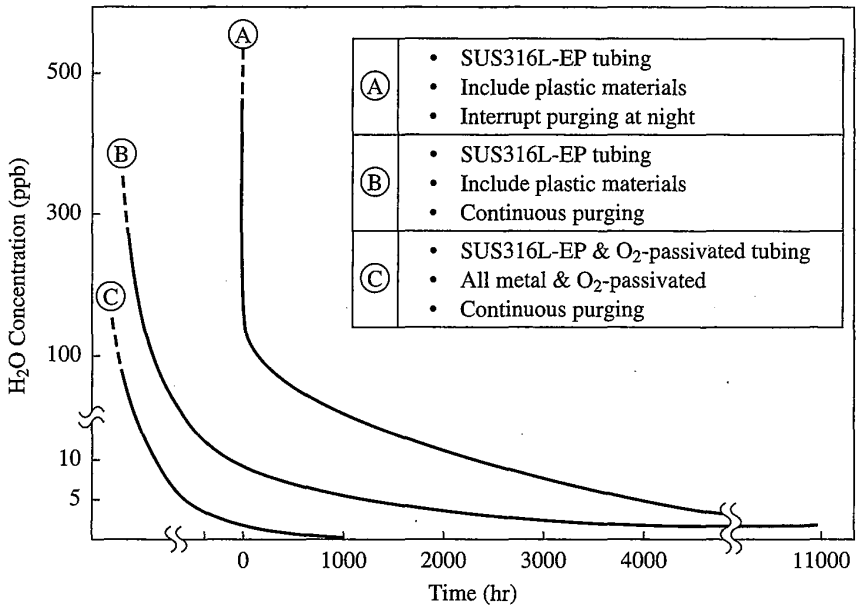


FIGURE 24
Moisture outgassing characteristics of three piping systems during the start-up stages.

TABLE 12
Evolution of process-gas components

DRAM technology	256K	1M	4M	16M	64M
Piping material	SS316LBA	SS316LBA	SS316LEP	SS316LEP	SS316LGEP
Internal surface	SCH-10S	SCH-10S	SCH-5S	SCH-5S	SCH-5S
Roughness, R_{max}, μm	< 3 – 5	< 1.0	< 0.7	< 0.3	< 0.1
Tube fitting	Swagelog	VCR	VCR	VCR	MCG/JSK
Ring			Metal	Metal	Metal
External leak, atm-cc/s	1 E−4	1 E−6	1 E−8	1 E−11	1 E−13
Valve					
Large	Bellows	Bellows	Bellows	Diaphragm	Diaphragm
Small	Bellows	Bellows and diaphragm	Diaphragm	All-metal diaphragm	All-metal diaphragm
Filter					
Cutoff size, μm	0.02	0.02	0.01	0.01	0.005
Large flow rate	Membrane	Membrane	Membrane	Membrane	Membrane
Small flow rate	Membrane	Membrane	Ceramic	Metal	Metal
Pressure regulator	Diaphragm	Diaphragm	Diaphragm	Diaphragm	Diaphragm
Pressure gauge	Borden tube	Borden tube	Borden tube	Flow-through pressure sensor	Flow-through pressure sensor

Source: Courtesy of BOC of England.

Micron Ex. 1023, p. 78
Micron v. YMTC
IPR2024-00909

sensors that monitor gas pressure with a built-in metallic diaphragm attaching to the interior wall of the gas tubing.

Improvement has also been made to gas cylinder valves. The conventional CGA (Compressed Gas Association) cylinder valve has been replaced by a newly developed DISS (Diameter Index Safety System) cylinder valve. The DISS gas cylinder valve employs a metallic gasket to provide a better seal at the contact between the male gas cylinder outlet and the female connector attaching to the gas-distribution system. Special provisions are also made to minimize the torque force, and thereby the abrasion, generated at the surface of the gas cylinder outlet when the male and female DISS connectors are tightened.

The design and construction concepts of an ultraclean process-gas piping system is similar to the ultraclean gas components:

Leak free
Particle free
Dead space free
Outgas free
Corrosion free
Plastic material free

Purge valves are often installed at the ends of the subbranches extended sideways from the bulk-gas main piping. A small volume of gas is continuously purging out of the piping system to ensure that the process gas is not stagnant inside any section of the piping system.

In addition to the proper selection of gas components, the actual field installation procedures plays a decisive role in determining the final gas quality. All the components and tubing should be wrapped in a nitrogen-filled plastic package prior to shipment from the vendors. The on-site subassembly of gas components should be performed in a clean booth with a cleanliness better than class 100. The installation of the subassembled parts should also be in a clean environment, preferably after the installation of HEPA filters. During welding, Ar containing 3 to 5% H_2 is used as an arc gas and back-seal gas to prevent the hot welding spot from oxidizing with ambient oxygen and moisture. After welding, the piping should be purging continuously with N_2 to get rid of the adsorbed moisture. Helium leak tests are also applied to all the welding spots to ensure there will be no leakage due to defective welding.

1.7.4 Space Management

Space management refers to the arrangement of process utilities including make-up air and exhaust ductwork, process pipings, electrical cable tray and bus bars, etc. in both the utility building and the cleanroom area. It comprises the design of the process utilities distribution network and the design of the pipe rack on which the process utilities are installed. Space management, which is often overlooked during cleanroom design and construction, results in an optimum and flexible process utilities distribution.

Each process utility piping network can be subdivided into main pipings and branch pipings. Figure 25 illustrates the arrangement of main piping and branch

Micron Ex. 1023, p. 79
Micron v. YMTC
IPR2024-00909

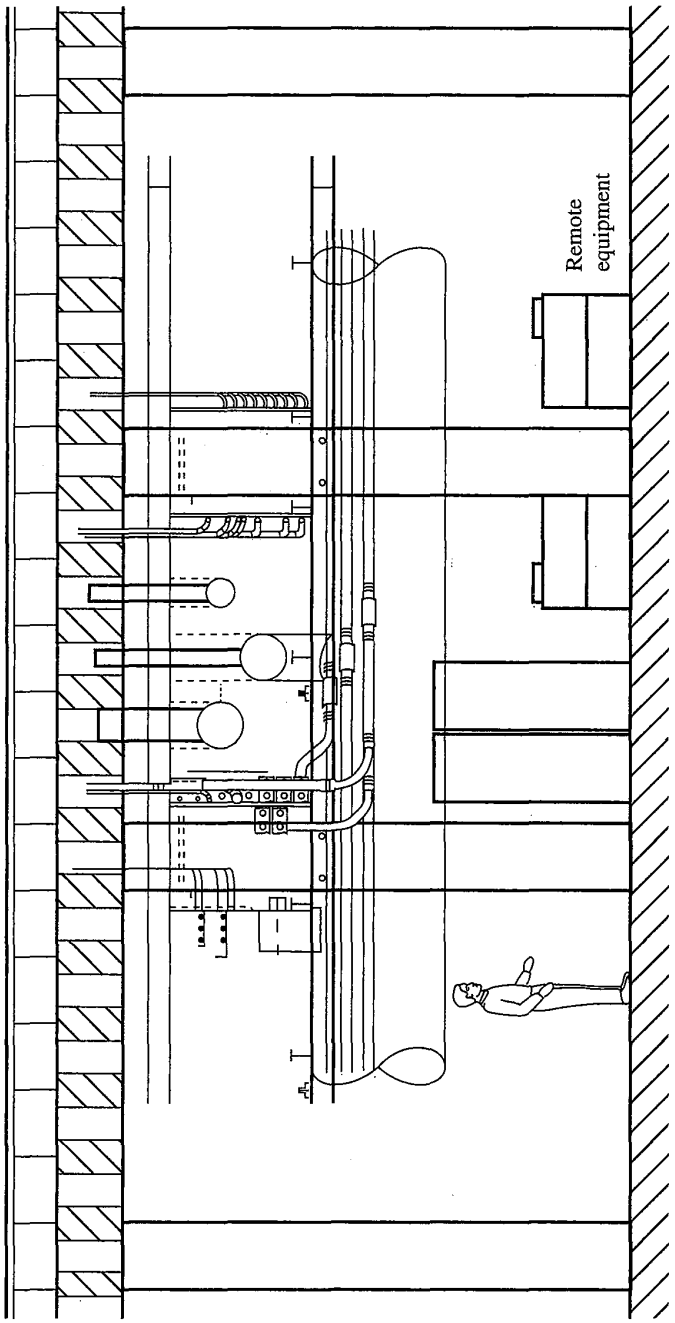


FIGURE 25
Example of the arrangement of the process utilities main and branch headers in the basement of a cleanroom.

Micron Ex. 1023, p. 80
Micron v. YMTC
IPR2024-00909

headers in the basement of a cleanroom. The main headers of the process utilities are based on a horizontal distribution of piping, cabling, and exhaust ductwork. The branch headers of the process utilities, which crisscross over the main header, are based on a vertical distribution of piping, cabling, and exhaust ductwork. From the branch header, process utilities are connected upward to the process equipment located at the second floor or downward to the pumps and power cabinets located on the first floor.

The optimal height of the cleanroom basement, from the floor level to ceiling, should be approximately 6 meters to allow enough clearance for people to walk freely under the main utilities headers.

The pipe rack is required to support the main and branch headers of the process utilities and occasionally the central utilities, such as chilled water for the recirculation air units.

1.8 SUMMARY AND FUTURE TRENDS

Because devices are continuously miniaturized, the design of the cleanroom system and process utilities is increasingly critical for the IC manufacturing process. As feature sizes decrease and wafer sizes increase, purity requirements become more stringent. A low defect density demands a substantial improvement in the control of the wafer environment. The current cleanroom classification is based on 1 particle with a size equal to or larger than 0.1 μm per ft^3 in the wafer environment, obtained in a downward laminar air flow at a maximum speed of 0.35 m/sec.

The design of a cleanroom system must be flexible to allow any desired process layout. The two-level cleanroom concept is recommended, in which the upper level is the clean production area with the process equipment, and the lower level is a utilities distribution area and a return air plenum as well as a site for installing remote equipment. All cleanroom subsystems should be integrated for optimum performance. Material selection and basic setup for each subsystem must satisfy the requirements of the manufacturing process. Vibration, temperature, relative humidity, ESD and EMI, and particle and chemical contamination must all be controlled to ensure high quality cleanroom performance.

The baseline management method, which divides the entire project into clearly identifiable phases, is the recommended approach for a cleanroom construction project. The time span from beginning the design to starting up the process is estimated to be approximately 2 years, divided into a 6-month preparation period, a 14-month realization period, and a 4-month hookup and final-testing period. Of the total investment for a semiconductor fab, 75% purchases the process equipment and 25% pays for the facilities and utilities. Of the latter amount, approximately 35% pays for the process utilities and 25% finances the cleanroom system.

The generation of particles inside the cleanroom must be prevented as much as possible. Gowning procedures are imposed on persons entering the cleanroom and decontamination procedures must be followed for incoming materials. Sources of microcontamination inside the cleanroom are minimized by proper cleanroom maintenance.

Micron Ex. 1023, p. 81
Micron v. YMTC
IPR2024-00909

Other related process utilities require the highest possible quality level. DI water, chemical supplies and process gases require attention and investment.

Total integration of process utilities requires space management—the optimal arrangement of the various utilities distribution networks and the supporting pipe rack. Trends in cleanroom design are toward minienvironments in which the wafer loading/unloading and processing areas are enclosed. Wafers are transferred automatically in and out of the minienvironment. Outside these minienvironments, the wafers are sealed inside a portable pod for transport. To eliminate microcontamination and reduce native oxide growth on silicon wafers, the nitrogen tunnel has been proposed for manufacturing 1G DRAM. In this future cleanroom, the wafers will be transported by robots, one piece at a time, inside a nitrogen tunnel that connects to each piece of process equipment. An alternative solution may be the use of cluster tools with dry HF etch chambers to clear the native oxide before a native-oxide-sensitive process is performed. In designing the cleanroom layout, enough space should be kept between process tools to accommodate the cluster tools, which usually have large footprints.

During the last three decades, the wafer size has increased from 75 mm to 200 mm, and the standard loading capacity of a wafer cassette has remained at 25 pieces. The lot size for the 300-mm wafer will become much smaller. Smaller lot size can significantly reduce the wafer processing cycle time; however, it can also dramatically increase the wafer handling frequency. Automatic transfer of the delicate wafers between process tools is the only applicable solution. Ultimately, the process tool manufacturing community must come to an agreement on standardizing the robotics interface so that total automation can be implemented smoothly.

Total automation is a starting point in the design of the process layout. This means that the required space and logistics must be planned. Three basic levels of automation can be distinguished: tool automation, intrabay automation, and interbay automation. Interbay automation is now a standard feature. Intrabay automation is becoming popular and will become absolutely necessary in the future 300-mm wafer fab.

REFERENCES

1. T. Ohmi, "Future Trends and Applications of Ultra-Clean Technology," *Tech Dig. IEDM,* 49 (1989).
2. T. Ohmi, "Ultraclean Cleanroom Environments: Closed System Essential for High-Quality Processing within Advanced Semiconductor Lines," *Microcontamination,* **8**(6), 27 (1990).
3. T. Ohmi and T. Shibata, "Developing a Fully Automated, Closed Wafer Manufacturing System," *Microcontamination,* **8**(7), 25 (1990).
4. T. Ohmi, "Breakthrough for Scientific Semiconductor Manufacturing in 2001. A Proposal from Tohoku University," Realize, Inc. (Japan) 1992.
5. D. W. Cooper, S. J. Grotzinger, L. R. Ackman, and V. Srinivasan, "Selecting Nearly Optimal Sampling Locations Throughout an Area. Application to Cleanrooms and Federal Standard 209," *Institute of Environmental Sciences, Annual Technical Meeting, New Orleans, LA,* 1990, p. 257.
6. S. J. Grotzinger and D. W. Cooper, "Selecting a Cost-Effective Number of Samples to Use at Preselected Locations," *Journal of the IES,* **35**(1), 41 (1992).

Micron Ex. 1023, p. 82
Micron v. YMTC
IPR2024-00909

7. D. W. Cooper and D. C. Milholland, "Sequential Sampling Plan for Federal Standard 209," *Journal of the IES,* **33**(5), 28 (1990).
8. R. Simon, "Clean Room Technology for Semiconductor Manufacturing," (German), *Stroemungsmechanik und Stroemungsmaschinen,* **39** 139 (Nov., 1988).
9. K. Unno, "Cleanroom," *Denski Zairyo (Electronic Material) 1988 Sup. Vol.,* 134 (1988).
10. T. Kawamata, "Cleanroom," *Denski Zairyo (Electronic Material) 1989 Sup. Vol.,* 125 (1989).
11. N. Namiki, "Cleanroom," *Denski Zairyo (Electronic Material) 1990 Sup. Vol.,* 136 (1990).
12. B. Patel, J. Greiner, and T. R. Huffman, "Constructing a High-Performance, Energy-Efficient Cleanroom," *Microcontamination,* **9** (2), 29 (1991).
13. B. Newboe, "Successful Cleanroom Designs," *Semic. Intl.,* **14** (6), 106 (1991).
14. J. R. Weaver, "Cleanroom Reverse Design Principle," *Semic. Intl.,* **14** (10), 4 (1991).
15. R. W. Frick and M. C. Miller, "Designing NASA's Largest Cleanroom," *Microcontamination,* **10** (1), 37 (1992).
16. A. Saiki, "Cleanroom," *Denski Zairyo (Electronic Material) 1993 Sup. Vol.,* 143 (1993).
17. T. Kawamata, "Cleanroom," *Denski Zairyo (Electronic Material) 1994 Sup. Vol.,* 133 (1994).
18. R. A. Hughes, G. B. Moslehi, D. M. Campbell, K. J. Radigan, W. Lukaszek, and E. D. Castel, "Eliminating the Cleanroom: More Experiences with an Open-Area SMIF Isolation Site (OASIS)," *Microcontamination,* **8** (4), 35 (1990).
19. J. Inbody and V. E. Bradley, "Upgrading an Existing Wafer Fab Facility with SMIF Technology," *Microcontamination,* **8** (9), 25 (1990).
20. C. Y. Hsu and L. C. Tu, "Designing, Operating a Submicron Facility with Isolation Technology," *Microcontamination,* **10** (3), 29 (1992).
21. T. Baechle, G. Marvell, W. Fosnight, and M. Lynch, "Assessing the Capabilities of a Minienvironment to Meet Increasing Environmental Specifications," *Microcontamination,* **10** (5), 25 (1992).
22. T. Baechle, G. Marvell, and M. Lynch, "Evaluating the Capabilities of Mini-Environments Using Polished Silicon Monitor Wafers," *Microcontamination,* 10 (5), 35 (1992).
23. W. C. Grande, "Upgrading a Class 100 Fab through Use of Manual-Access Microenvironments," *Microcontamination,* **11** (1), 25 (1993).
24. A. Morizuki, "Cleanroom for Semiconductor Industries," *Denski Zairyo (Electronic Material),* 35 (Aug., 1990).
25. T. Ohmi, Y. Kasama, K. Sugiyama, Y. Mizuguchi, Y. Yagi, H. Inaba, and M. Kawakami, "Controlling Wafer Surface Contamination in Air Conditioning, Particle Removal Subsystems," *Microcontamination,* **8** (2), 45 (1990).
26. T. Ohmi, H. Inaba, "Cleanroom Special Edition — Environmental Control Technique," (Japanese), *Nikkei Microdevices,* 115, (Sep., 1989).
27. Suzuki, "Supercleanroom," (Japanese), *Hitachi Hyoron (Commentary),* **68** (9), 737 (1986).
28. M. J. Bader, "Ionizers for Clean Rooms, Eliminating Both Electrostatic Charging and Dust," (German), *Plastverarbeiter,* **39** (12), 44 (1988).
29. J. R. Turner, D. K. Liguras, and H. J. Fissan, "Clean Room Applications of Particle Deposition from Stagnation Flow: Electrostatic Effects," *Journal of Aerosol Science,* **20** (4), 403 (1989).
30. P. C. D. Hobbs, V. P. Gross, and K. D. Murray, "Suppression of Particle Generation in a Modified Clean Room Corona Air Ionizer," *J. Aerosol Science,* **21** (3), 463 (1990).
31. A. Steinman, "Evaluating Air Ionization Systems," *Evaluation Engineering,* **29** (4), (1990).

Micron Ex. 1023, p. 83
Micron v. YMTC
IPR2024-00909

32. P. C. D. Hobbs, V. P. Gross, and K. D. Murray, "Reviewing Clean Corona Discharge: Laser-Produced Plasma Ionization Technologies," *Microcontamination,* **9** (6), 19 (1991).
33. T. Sebald, "Continuously Working Monitoring System for Clean Room Ionizers," *Environmental Engineering,* **5** (4), 16 (1992).
34. T. Takenami, T. Ohmi, and S. Fukuda, "Air Conditioning and Particle Filtration Systems for Energy Saving," *Sol. State Technol.,* **32** (4), 161 (1989).
35. P. Naughton, "HVAC Systems for Semiconductor Cleanrooms. Part 1: System Components," *ASHRAE Trans., Part 2,* 620 (1990).
36. P. Naughton, "HVAC Systems for Semiconductor Cleanrooms. Part 2: Total System Dynamics," *ASHRAE Trans., Part 2,* 620 (1990).
37. T. Ohmi, H. Inaba, and T. Takenami, "Using Water-Based Cooling Systems in Cleanroom Environments," *Microcontamination,* **7** (12), 27 (1989).
38. Y. Suzuki, S. Oikawa, and T. Sekiguchi, "Super Cleanroom Technology: A High-Tech Balancing Act," *Microcontamination,* **6** (9), 59 (1988).
39. B. Brandt and L. A. Wright, "Analyzing Particle Release of Cleanroom Headcoverings," *Microcontamination,* **8** (10), 53 (1990).
40. R. Spector, C. Berndt, and C. W. Berndt, "Reviewing Methods for Evaluating Cleanroom Garment Fabrics," *Microcontamination,* **11** (3), 31 (1993).
41. P. Ravis, G. H. Ranta, "How to Choose a Clean Room Laundry," *Institute of Environmental Sciences, Annual Technical Meeting, New Orleans, LA,* 1990, p. 355.
42. K. Skidmore, "Keep Your Clean Room Particle Free," *Semic. Intl.,* **12** (9), 94 (1989).
43. W. K. Kwok and J. T. Summers, "Characterization of Cleanroom Wipers: Particle Generation," *Institute of Environmental Sciences, Annual Technical Meeting, New Orleans, LA,* 365 (1990).
44. R. K. Schneider, "Developing and Implementing a Cleanroom Construction Protocol," *Microcontamination,* **8** (8), 35 (1990).
45. H. H. Schicht, "Contamination Control. An Indispensable Factor in High-Technology Manufacturing Tasks," *J. Aerosol Science,* **21,** Supp 1, 719 (1990).
46. J. Greiner and M. O'Halloran, "The Forgotten Functions: Support Facilities for the Modern Microenvironment," *Microcontamination,* **9** (3), 45 (1991).
47. A. M. Dixon, "Protesting Your High-Tech Investment through Sound Cleanroom Maintenence Practices," *Microcontamination,* **10** (11), 38 (1992).
48. T. Ohmi, "Proposal for Advanced Semiconductor Manufacturing Equipment: An Approach to Automated IC Manufacturing," *Proceedings of the 5th Symposium on Automated Integrated Circuits Manufacturing,* IEEE, 1989, p. 3.
49. T. Ohmi, T. Shibata, "Requirements of CAM in IC Technology," *Microelectron. Eng.,* **10** (3–4), 177 (1991).
50. P. K. John, "Optimal Partitions for Shop Floor Control in Semiconductor Wafer Fabrication," *Europ. J. Oper. Res.,* **59** (2), 294 (1992).
51. Y. Mizokami, "Total CIM System for Semiconductor Plants," IEEE/SEMI International Semiconductor Manufacturing Science Symposium—ISMSS '90, May 21–23, 1990.
52. R. S. Weiss, "Automation and Particulate Control within an IC Manufacturing Facility," *Proceedings of the 5th Symposium on Automated Integrated Circuits Manufacturing,* IEEE, 1989, p. 199.
53. H. Tabata, T. Yamashita, M. Murata, M. Onishi, and T. Tsubaki, "Autonomous Mobile Robot Transport System for Clean Rooms," (Japanese), *R&D, Research and Development, Kobe Steel, Ltd.,* **40** (3), 36 (1990).
54. I. Y. Wang, A. S. Li, and B. V. Ilene, "A Magnetic Levitation Transport Path," *IEEE Trans. Semic. Manuf.,* **4** (2), 145 (1991).

Micron Ex. 1023, p. 84
Micron v. YMTC
IPR2024-00909

55. T. Araki, "Autotransfer System," Japanese, *Denski Zairyo (Electronic Material) 1994 Sup. Vol.,* 141 (1994).
56. H. Sato, M. Hashimoto, T. Shinoda, and Y. Hiratsuka, "Ultrapure Water System for 16M DRAM," Japanese, *Semiconductor World,* **8** (6), 54 (1989).
58. A. Houzuki and K. Ushigoe, "Ultrapure Water Manufacturing System," (Japanese), *Denski Zairyo (Electronic Material) 1989 Sup. Vol.,* 118 (1989).
59. M. Furuichi, "Ultrapure Water Manufacturing System," (Japanese), *Denski Zairyo (Electronic Material),* 69 (Aug., 1990).
60. M. Toto, "A Closed Ultrapure Water Manufacturing System," (Japanese), *Denski Zairyo (Electronic Material),* 59 (Aug., 1991).
61. K. Oda and T. Ozaki, "Ultrapure Water Manufacturing System," (Japanese), *Denski Zairyo (Electronic Material) 1991 Sup. Vol.,* 108 (1991).
62. Y. Yagi, T. Imaoka, Y. Kasama, and T. Ohmi, "Advanced Ultrapure Water Systems with Low Dissolved Oxygen for Native Oxide Free Wafer Processing," *IEEE Trans. Semic. Manuf.,* **5** (2), 7 (1992).
63. T. Doi, "Ultrapure Water Manufacturing System," (Japanese), *Denski Zairyo (Electronic Material) 1993 Sup. Vol.,* 137 (1993).
64. K. T. Pate, "Examining the Design, Capabilities, and Benefits of Bulk Chemical Delivery Systems," *Microcontamination,* **9** (10), 25 (1991).
65. K. Kobayashi, M. Tamura, T. Shimada, and H. Sakai, "Chemical Autodelivery System," *Denski Zairyo (Electronic Material) 1993 Sup. Vol.,* 157 (1993).
66. T. Ohmi, Y. Kasama, K. Sugiyama, Y. Mizuguchi, Y. Yagi, H. Inaba, and M. Kawakami, "Examining Performance of Ultra-High-Purity Gas, Water, and Chemical Delivery Subsystems," *Microcontamination,* **8** (3), 27 (1990).
67. T. Ohmi, K. Sugiyama, F. Nakahara, M. Tsuda, Y. Sugano, and N. Onaga, "The Reduction of Residual Gas After Purging by Eliminating the Dead Space in the Valve," (Japanese), *Nikkei Microdevices,* 126 (Jul., 1988).
68. Y. Mukagawa and K. Tamura, "The New Trend of Semiconductor Process and Supporting Facility Technologies," (Japanese), *Denski Zairyo (Electronic Material),* 22 (Aug., 1990).
69. K. Sugiyama, F. Nakahara, and T. Ohmi, "Designing a Gas Delivery System for Lower Submicron ULSI Processes," *Microcontamination,* **7** (7), 29 (1989).
70. D. H. Hope, R. J. Markle, T. F. Fisher, J. B. Goddard, J. Notaro, and R. D. Woodward, "Installing and Certifying SEMATECH'S Bulk-Gas Delivery Systems," *Microcontamination,* **8** (5), 31 (1990).
71. M. Nakamura, T. Ohmi, and K. Kawada, "All Metal and Oxygen Passivation Tubing Technology for Ultra Clean Gas Delivery System," *Institute of Environmental Science Technical Meeting,* San Diego, 605 (1991).
72. S. Kamoki, K. Sugiyama and M. Nakamura, "Gas Delivery System for Semiconductor Processes," (Japanese), *Electronic Material 1993 Sup. Vol.,* p. 151 (1993).
73. T. Aida, "Gas Delivery System for Semiconductor Processes," (Japanese), *Electronic Material 1994 Sup. Vol.,* 147 (1994).

PROBLEMS

1. Due to contamination that occurred in the cleanroom, the wafer defect density, measured at sizes above 0.3 μm, has increased fivefold from 0.2 D/cm^2 to 1.0 D/cm^2. Use the following equation along with the data provided in Tables 1 and 2 to estimate the yield loss of a 4M DRAM and a 16M DRAM wafer.

$$Y = e^{-DA}$$

Micron Ex. 1023, p. 85
Micron v. YMTC
IPR2024-00909

where D is the defect density and A is the chip area. Note that the density of defects is roughly inversely proportional to the defect size to the second power.

2. The class 1 according to Japanese Std. B9920 rev. describes air cleanliness with no more than 10 particle/m^3 of a size 0.1 μm or larger. Calculate the equivalent air cleanliness class according to U.S. Fed. Std. 209E.

3. Estimate the numbers of sampling locations and the total air sampling time required to certify a submicrometer fab having 2000 m^2 class M1 process area and a 2000 m^2 class M3 service area. Use the single sampling plan and sample the air at 28.3 L/min.

4. An accidental spillage of an IPA bottle raised the cleanliness level of a 10 m^2 cleanroom from class 1000 to class 100,000. Assuming the air exchange rate of the cleanroom is 60 times per hour, estimate the time required for the air cleanliness to recover to class 1000.

5. The cleanliness of outside ambient air is around class M7 according to Fed. Std. 209E. Estimate the cleanliness of the air leaving the make-up air unit having a HEPA filter of 99.97% filtration efficiency. What is the minimum filtration efficiency of the ULPA filters required for the process area to reach a class M1 (at 0.1 μm) cleanliness?

6. The cleanliness of a cleanroom is achieved by recirculating air through HEPA filters. Estimate the annual electricity requirement for maintaining the air cleanliness of a 4000 m^2 submicrometer cleanroom. The HEPA filter coverage of the entire cleanroom is 60%. The average air speed leaving HEPA filters is 0.35 m/sec. The average pressure drop across the HEPA filter is 120 Pa. The efficiency of the axial fan unit is 85%.

7. By adopting the minienvironment concept, the HEPA coverage of the conventional cleanroom described in Problem 6 can be reduced by half. The total investment of implementing SMIF is around $25 million. Estimate the annual saving in electricity and compare that with the SMIF investment cost. The cost of electricity is $0.1 per kWh. The amortization period of SMIF investment is 7 years and the annual interest of the capital is 6%.

8. To maintain a constant air velocity of 0.35m/sec, the pressure drop across the HEPA filter has to be increased over time. Use the data provided in Problem 6 to estimate the annual increment in electricity cost for a HEPA filter having an initial pressure drop of 100 Pa and an annual pressure increment of 10 Pa.

9. To control the relative humidity level of a cleanroom at 43%, the outside air has to be cooled to 9°C to remove the excess moisture before being heated to the operating temperature of 22°C. Since a large amount of heat is released by the process tools located inside the cleanroom, significant energy saving on chillers and boilers can be realized by lowering the outlet temperature of the make-up air to, for example, 14°C instead of 22°C. Use the data provided in Section 1.3.5 to estimate the annual energy saving achieved in a submicrometer fab with a monthly production capacity of 15,000 wafers. Assume the average mask level is 20 and the average heat removal efficiency of a chiller is 5, i.e., 5 kW heat removed per kW of electricity required. The cost of electricity is $0.1 per kWh.

10. Estimate the gas displacement time of a 1/4-in. bellows valve at the flow rate of 50 cc/min. Repeat the same calculations for a diaphragm valve. The dead space of the bellows valve and the diaphragm is 2 cc and 0.2 cc, respectively. The gas displacement time is defined as the time required for the displaced gas to reach 0.01% of its initial concentration.

Micron Ex. 1023, p. 86
Micron v. YMTC
IPR2024-00909

CHAPTER 2

Wafer-Cleaning Technology

C. Y. Chang and T. S. Chao

2.1 INTRODUCTION

ULSI technology requires more stringent and reliable means to control the surface smoothness and to get rid of particles and contamination, such as metallic and organic residues on the silicon wafer surface, than does VLSI technology.

The well-known RCA wet clean[1] processes have been used extensively since the 1970s. The basic mechanisms of cleaning will be discussed in the following section. This discussion is followed by a presentation of recent developments that modify the standard RCA process to meet the more severe requirements and various applications of ULSI processes. These processes include thinner (< 5 nm) gate oxide formation, metal contact, and high-crystal-quality epitaxy. Next, we will address dry cleaning processes that meet the future requirements of integrated cluster-processing systems. Recent developments in surface electronic-state configurations, surface termination (e.g., H-terminated or F-terminated), surface kinetics, and reactions will be discussed in order to understand the surface-cleaning processes.

Supporting technologies, such as O_2-content-free deionized water, pure gas delivery systems, special designs for avoiding particulate generation in the process chamber, cleanroom design, etc., that are crucial to avoid surface contamination of the processing wafer have been discussed in detail in Chapter 1 and will be readdressed where they are needed.

2.2 BASIC CONCEPTS OF WAFER CLEANING

The wafer-cleaning process has been the important and critical step in semiconductor manufacturing for over 30 years. The most common process used today is the RCA wet cleaning process.[1] The RCA cleaning process was first developed by Kern and

Micron Ex. 1023, p. 87
Micron v. YMTC
IPR2024-00909

Puotinen in 1960 at RCA and was published in 1970. Two sequential cleaning solutions, NH_4OH–H_2O_2–H_2O (called *standard cleaning 1* (SC-1), a composition of 1 : 1 : 5 to 1 : 2 : 7 at 70 to 80°C) and HCl–H_2O_2–H_2O (called *standard cleaning 2* (SC-2), a composition of 1 : 1 : 6 to 1 : 2 : 8 at 70 to 80°C), were used. Both of these cleaning processes are based on hydrogen peroxide. At a high pH value SC-1 can effectively remove organic contamination and particles by oxidation. At a low pH value SC-2 can desorb metal contamination by forming a soluble complex. These basic concepts to remove contamination have not changed since they were developed. However, as the dimensions of devices have shrunk to submicrometer scale, additional requirements have been proposed for an ultraclean wafer surface in VLSI and ULSI. For complete fulfillment of an ultraclean wafer surface, it should be free from particles, organic contamination, metal contamination, surface microroughness, and native oxide.

Particle, organic, and metal contamination have been reduced with a controlled, super-cleanroom technology and improved wet cleaning technology. To improve the integrity of ultrathin gate oxide and to grow an epi-layer in a low-temperature process, the surface must be free of native oxide, or perfectly hydrogen-terminated. Another issue is the microroughness. This issue is not so serious when the thickness of the gate oxide is larger than 200 Å using 0.7 μm technology. However, as the thickness is reduced to less than 100 Å for 0.35 μm or smaller technology, surface microroughness at the SiO_2/Si interface becomes a prominent influence on channel electron mobility as well as on ultrathin gate oxide quality. The sources and the related effects of the contaminations are shown in Table 1. Different contaminants cause different effects on device reliability, which will be discussed in the next section. All these problems are summarized in Fig. 1 from the point of view of

TABLE 1
Sources and related effects of various contaminations

Contamination	Possible source	Effects
Particles	Equipment, ambient, gas, deionized (DI) water, chemical	Low oxide breakdown Poly-Si and metal bridging–induced low yield
Metal	Equipment, chemical, reactive ion etching (RIE), implantation ashing	Low breakdown field Junction leakage Reduced minority lifetime V_t shift
Organic	Vapor in room, residue of photoresist, storage containers, chemical	Change in oxidation rate
Microroughness	Initial wafer material, chemical	Low oxide breakdown field Low mobility of carrier
Native oxide	Ambient moisture, DI water rinse	Degraded gate oxide Low quality of epi-layer High contact resistance Poor silicide formation

Micron Ex. 1023, p. 88
Micron v. YMTC
IPR2024-00909

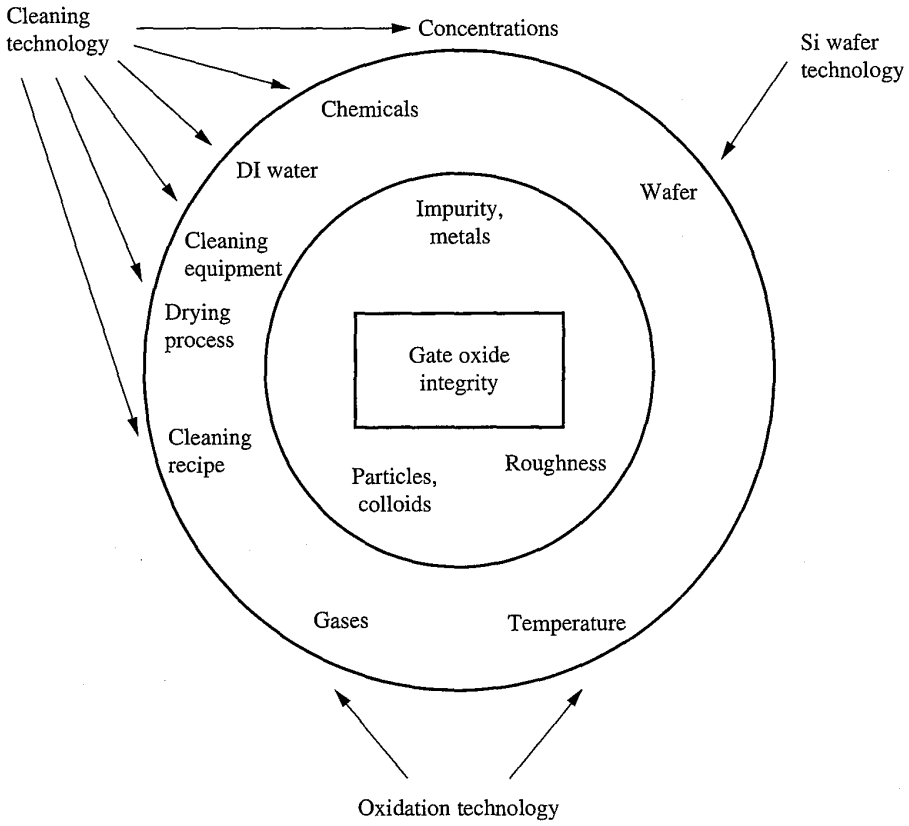


FIGURE 1
Interrelation of different factors on the gate oxide. (*After Verhaverbeke, Ref. 2.*)

gate oxide integrity.[2] These problems are important because gate oxide quality is one of the critical steps that determine the yield, reliability, and performance of a ULSI circuit. Problems due to roughness, impurity, and particles can be the result of cleaning technology (concentrations, chemicals, water, cleaning equipment, dry process, and cleaning recipe), oxidation technology (gases, temperature), and Si wafer technology (initial wafer). Because chemicals, deionized (DI) water, and Si wafers are available in high purity, the sources of most contaminations are tools and processes rather than the environment and operators. Although many cleaning processes have been developed to remove contaminants, the most important thing is to avoid contamination rather than to clean it up during processing. The basic concepts of each issue are discussed in the following section. The reliability problems related to each contaminant will also be discussed.

2.2.1 Mechanisms for Removing Particles and Contaminations

The mechanisms of removing particles and contaminations are addressed in the following paragraphs.

Micron Ex. 1023, p. 89
Micron v. YMTC
IPR2024-00909

TABLE 2
Particle concentration (number/mL) in ULSI-grade semiconductor chemicals[6]

	≥ 0.2 μm	≥ 0.5 μm
NH_4OH	130–240	15–30
H_2O_2	20–100	5–20
HF	0–1	0
HCl	2–7	1–2
H_2SO_4	180–1150	10–80

Particles

Particle adhesion to a silicon surface usually occurs during the process from the equipment, ambient, gas, chemicals, and DI water. The particle contamination in ULSI-grade semiconductor chemicals is shown in Table 2. H_2SO_4 has the highest number of particles and HF the lowest. Figure 2 shows that the yield of gate oxide is reduced as the number of particles is increased.[3] The requirement of the particle size is scaled down with the minimum feature size of the technology.[4] The general guide for particle size is one-tenth of the device feature size. By the year 2000, 0.2-μm device features will require a particle size smaller than 0.02 μm.

To control particles, it is necessary to understand the mechanisms of adhesion and removal of particles. Particle adhesion is dominated by several mechanisms: forces due to static charge on the particle or to van der Waals forces, forces due to the formation of an electrical double layer, forces due to capillary action around the particle, and a chemical bond between the particle and the surface. Particle removal mechanisms can be classified into four types:

1. Dissolution
2. Oxidizing degradation and dissolution
3. Liftoff by slight etching of the wafer surface
4. Electric repulsion between particles and the wafer surface

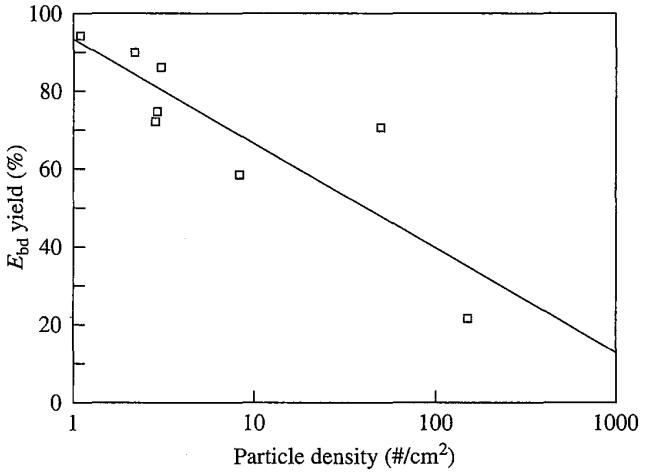


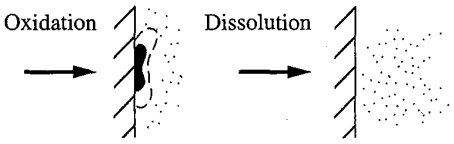
FIGURE 2
Dependence of E_{bd} yield on particle density. (*After Werkhoven et al., Ref. 3.*)

Micron Ex. 1023, p. 90
Micron v. YMTC
IPR2024-00909

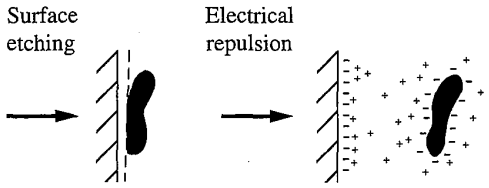
The oxidation and electrical repulsion mechanisms are illustrated in Figs. 3*a* and *b*.[5] Figure 3*a* shows the oxidation of a particle to make it soluble. Figure 3*b* shows the particle being removed from the surface by the electric repulsion. The SC-1 chemistry has both mechanisms. Hydrogen peroxide can oxidize the silicon surface, and OH^- from the NH_4OH can provide the negative charges on the silicon surface and the particle.

The deposition of the particle is a strong function of the pH value of the solution. Figure 4 shows the number of particles deposited as a function of pH.[6] Increasing the pH to 10 results in low particle deposition. Clearly, maximum particle deposition is observed in a highly acidic solution. A comparison of the particle-removal efficiency of various cleaning processes is shown in Table 3; the SC-1 was found to have the highest removal efficiency.

Some methods have been developed to improve the efficiency of removing particles. The most effective commercial method is the *megasonic* cleaning process.[7] SC-1 in conjunction with the megasonic cleaning process removes organic and inorganic particles from the surface at a temperature of less than 40°C. Figure 5*a* shows the detailed schematic of a static megasonic cleaning tank, which is noncontact and brushless, for removing particles (< 1 μm) simultaneously from both sides of a silicon wafer. The model of the megasonic action on a particle on a silicon wafer surface is shown in Fig. 5*b*. Sonic energy from the transducer is directed parallel to the silicon surfaces, which are immersed in the SC-1 solution. The high-power (up to 300 W) and high-frequency (850 to 900 kHz) sonic pressure waves wet the particle first, and the solvent diffuses into the interface. Finally, the particle is totally wetted and is removed from the surface as a floating free particle. This particle-cleaning process is very useful in cleaning after the newly developed chemical-mechanical polishing (CMP) process to remove the slurry on the surface. In addition to particle removal, the megasonic process with SC-1 is reported to remove copper contamination effectively.[7] This is because the metal exchanges ions readily with the ammonia and therefore can be rinsed off.



(*a*) in a powerful oxidizing agent



(*b*) in an alkaline solution

FIGURE 3
Acidic and alkaline wet chemical mechanisms to remove particles, organics, and metallics: (*a*) in a powerful oxidizing agent, (*b*) in an alkaline solution. (*After Singer, Ref. 5.*)

Micron Ex. 1023, p. 91
Micron v. YMTC
IPR2024-00909

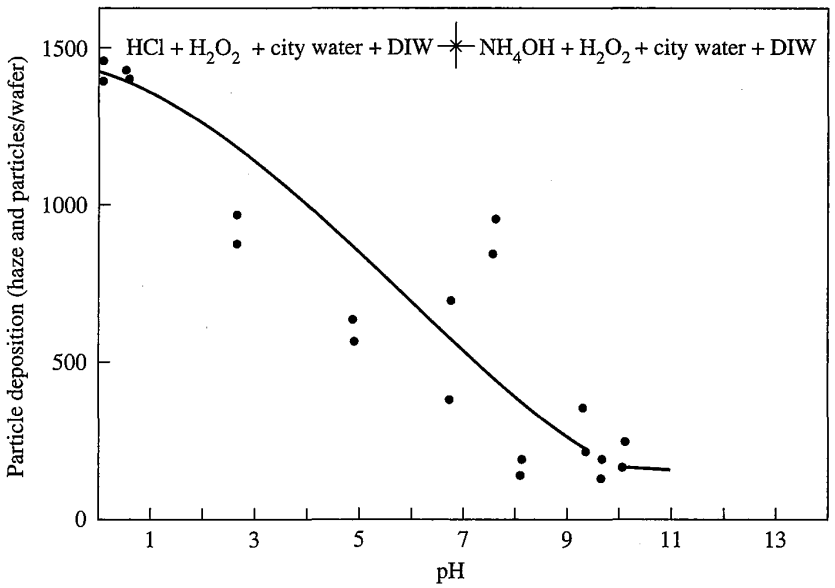


FIGURE 4
Deposition of tap water particles onto wafer surfaces as a function of seed solution pH. 10 percent by volume of tap water was mixed with ultrapure water. Acidity/basicity was controlled by the addition of either HCl with H_2O_2 or NH_4OH with H_2O_2. (*After Itano et al., Ref. 6.*)

Two points must be considered in using this process. One is that enough energy must be imparted to the particle so that it moves far enough from the surface not to readhere; this problem can be reduced by adding a continuous filtration system to trap the free particles from the surface into the solution. The other problem is that the concentration of the H_2O_2 will be reduced after long operation. Consequently, the smoothness of the silicon surface will be degraded by etching by the NH_4OH. This will be discussed in the microroughness section.

TABLE 3
Particle contamination removal efficiencies for various cleaning processes[6]

	Cleaning			
Pollution source	**SPM†**	**SC-1‡**	**SC-2§**	**PM¶**
City water	98.4	98.9	86.0	42.4
SiO_2	83.3	98.4	97.1	94.7
PSL††	91.7	99.2	55.2	7.2
Environmental air	95.8	96.3	86.9	0.

† SPM (Sulfuric–peroxide mixture): H_2SO_4: H_2O_2: (4:1) 5 min.
‡ SC-1: NH_4OH: H_2O_2: H_2O (0.1:1.5) 10 min at 80–90°C.
§ SC-2: HCl: H_2O_2: H_2O (1:1:6) 10 min at 80–90°C.
¶ PM (peroxide mixture): H_2O_2: H_2O (1:5) 10 min at 80–90°C.
†† Polystyrene latex sphere.

Micron Ex. 1023, p. 92
Micron v. YMTC
IPR2024-00909

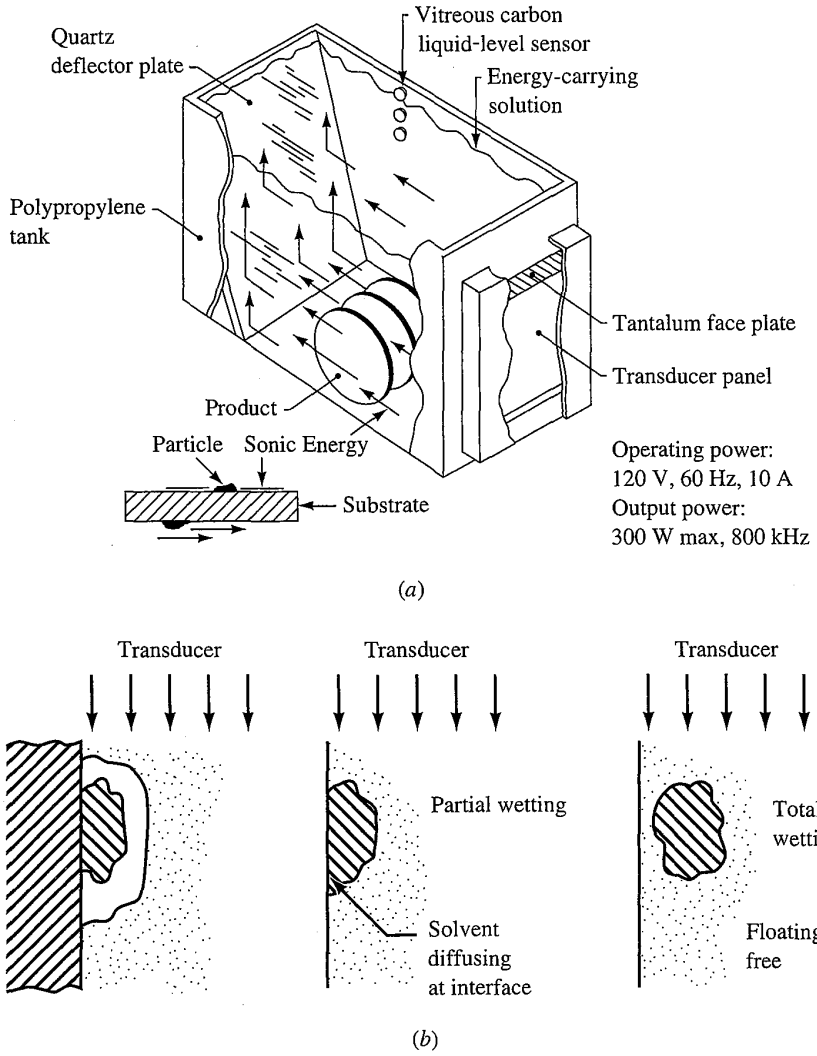


FIGURE 5
(*a*) Detailed schematic of static megasonic cleaning tank. The transducer plate incorporates eight transducers. The water holder is not shown, for clarity. The megasonic action for particle removal is indicated schematically. In an alternative system design the transducers are mounted in the shape of strips, eliminating the need for the reflector. (*b*) Model of megasonic action on a particle held on a silicon wafer surface: frequency 900 kHz; pressure $9.1 \times 10^5 N/m^2$; maximum instantaneous velocity 30 cm/s; wavelength 1.3 mm; acceleration $\sim 10^5$ g; motion of H_2O molecule ~ 0.1 μm. (*After Shwartzman et al., Ref. 7.*)

Metal contamination

The source of metal contamination can be the chemical solution, the ion implantation, or the reactive ion etching (RIE) process. Metal contaminations start at a level of 10^{10} atom/cm^2. Such a low level is below the lowest limit of measurement of the commonly used total-reflection x-ray fluorescence (TRXF).[8] The effects of metal

Micron Ex. 1023, p. 93
Micron v. YMTC
IPR2024-00909

contamination on devices are structural defects at the interface,[9] stacking faults during later oxidation or epitaxial process,[10] increased leakage current of p-n junctions, and reduced minority carrier lifetime.[11] Figure 6 shows the metallic impurities on a silicon wafer surface resulting from different processes.[12,13] Steam oxidation has the lowest level, less than 10^{10} atom/cm^2. Ion implantation exhibits the largest metal contamination, at a level of 10^{12} to 10^{13} atom/cm^2.

There are two mechanisms for precipitation of metal impurities onto the Si surface. The first is the direct binding to the Si surface by the charge exchange between a metallic ion and hydrogen atom that terminated on the Si substrate.[13] This type of impurity is not easy to remove in the wet cleaning process. Metals of this kind are usually noble metal ions, such as Au, whose electronegativity is higher than that of Si and that tend to be neutralized by taking an electron from Si and to precipitate on the Si surface. The electronegativity of various metals is shown in Table 4.

The second mechanism for precipitation of metal impurities occurs when oxide forms on the surface and metallic impurities are included simultaneously. Metals such as Al, Cr, and Fe tend to oxidize when the Si surface is oxidized and are included in the oxide film. These metallic impurities can be removed by etching the oxide with dilute HF. The enthalpy of metallic oxide formation is shown in Table 5.

Currently, the wet cleaning process is the most effective method for removing metallic contaminations. HF–H_2O_2(HF : 0.5%, H_2O_2 : 10%) cleaning is effective in removing metallic impurities that bind directly with Si surface. Some additives, such as hydrocarbon-type surfactants, which are added to BHF (buffered HF) to improve its wettability of the Si surface, are also found to be effective in removing Cu at room temperature.[14]

Both SC-1 and SC-2 have the capability of removing metallic impurities on the Si wafer. This is believed to result from the high oxidizing mechanism of the H_2O_2. Metallic contamination on the surface also induces microroughness,[15] which is shown in Fig. 7. Compared to the RCA-cleaned wafer, surfaces contaminated with Ca show a pronounced increase in haze after annealing in Ar at 950°C for 5 min.

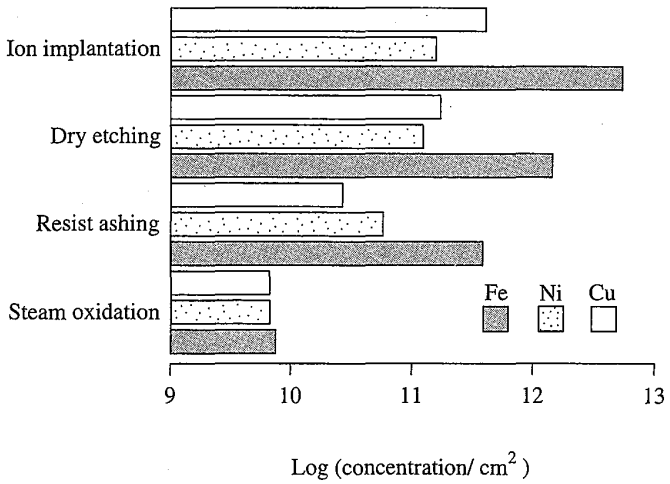


FIGURE 6
Contamination induced by various processes. (*After Anzai et al., Ref. 12, and Ohmi et al., Ref. 13.*)

Micron Ex. 1023, p. 94
Micron v. YMTC
IPR2024-00909

TABLE 4
Electronegativity scale of metals:[13] $M^+ + e^- \rightarrow M$

Element	Electronegativity (Pauling)	Half-cell reduction potential (V)	
Au	2.4	1.68	↑
Pt	2.2	1.19	
Ag	1.9	0.80	
Hg	1.9	0.79	
Cu	1.9	0.34	
Si	1.8	0.10	Tendency to be precipitated on bare Si
Pb	1.8	−0.13	
Sn	1.8	−0.14	
Ni	1.8	−0.23	
Fe	1.8	−0.41	
Zn	1.6	−0.76	
Al	1.5	−1.66	
Mg	1.2	−2.34	
Ca	1.0	−2.87	
Na	0.9	−2.71	
K	0.8	−2.92	

The defect density for these contaminated wafers after oxidation is shown in Fig. 8. Fortunately, Cu contamination has little effect on the oxide quality and the surface roughness. The most detrimental effect is from the Ca contamination, which results in a rougher surface during the annealing process and also increases defect densities in the oxide. Figure 9 shows the yield as the function of Ca contamination.[16] As the Ca contamination increases to 10^{11} atom/cm^2, the yield drops abruptly from 80% to 20%. This suggests that the threshold value for Ca contamination is 10^{11} atom/cm^2 for a reasonable yield. Figure 10 shows a typical *I-V* curve of an oxide grown on a Ca-contaminated wafer.[17] In the low-field region the current is enhanced, because of the rough surface caused by Ca contamination, and results in early breakdown. Although

TABLE 5
Enthalpy of oxide formation:[13] $M + O_2 \rightarrow MO_2, \Delta H < 0$ (heat-releasing process)

Oxide	ΔH_{25}^{293} (kJ/mol)	
Al_2O_3	−1675	
Cr_2O_3	−1130	↑
CrO_2	−583	
CrO_3	−580	Tendency to be included in the oxide film
Fe_3O_4	−1118	
Fe_2O_3	−822	
SiO_2	−909	
NiO	−241	
CuO	−155	

Micron Ex. 1023, p. 95
Micron v. YMTC
IPR2024-00909

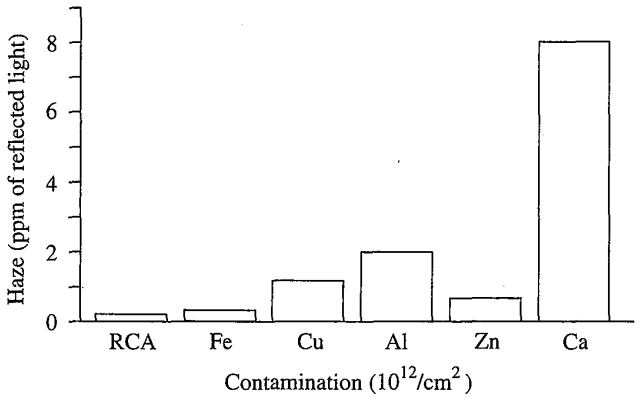


FIGURE 7
Haze number on contaminated Si wafers after annealing in Ar at 950°C for 5 min. (*After Verhaverbeke et al., Ref. 15.*)

Ca contamination will cause serious problems in the gate oxide, most of it can be removed by the HF-last cleaning steps and DI water rinsing. Table 6 shows typical metallic contamination after different last cleaning steps.[18] HF-dipping reduces the metal contamination most. The HF–H_2O_2–H_2O cleaning process can remove more Ca than HF can, but larger amounts of the other contaminants remain.

Another heavy metal prevalent on the Si is Fe, due to the common use of stainless steel in the equipment and facility. Figure 11 shows the histogram of oxide failure for Fe contamination levels of 1×10^{13} to 5×10^{14} atom/cm^3.[9] When the Fe contamination is less than 1×10^{10} atom/cm^3, there is no effect on the failure rate, but when Fe contamination increases to 5×10^{13} atom/cm^3, all oxides exhibit an early breakdown. The contamination threshold decreases as the oxide thickness is decreased[9] as shown in Fig. 12. For an oxide only 100 Å thick, the threshold of Fe concentration is as low as 10^{11} atom/cm^3. Therefore, the Fe contamination level must be controlled after each process to achieve a high gate oxide yield.

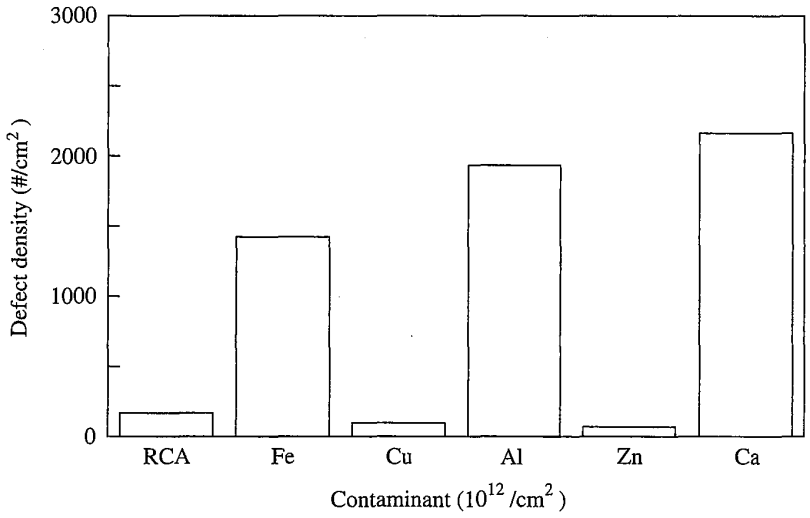


FIGURE 8
Oxide defect density on the RCA-cleaned and contaminated samples as obtained from extreme value statistics. The oxide thickness was 15 nm. (*After Verhaverbeke et al., Ref. 15.*)

Micron Ex. 1023, p. 96
Micron v. YMTC
IPR2024-00909

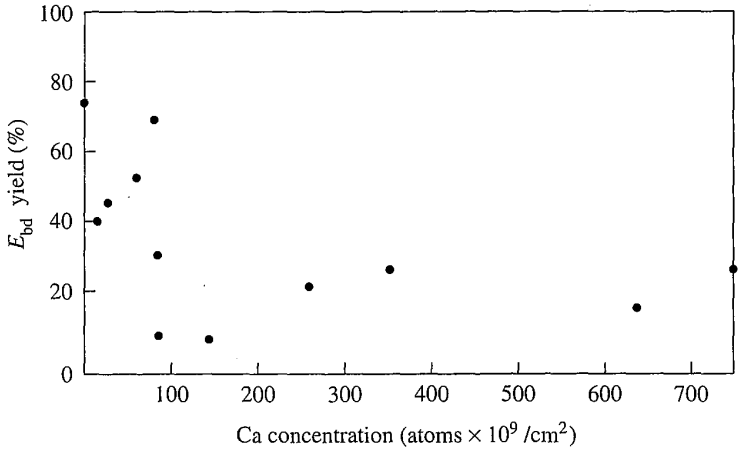


FIGURE 9
Relation between the Ca concentration and the yield. (*After Verhaverbeke et al., Ref. 16.*)

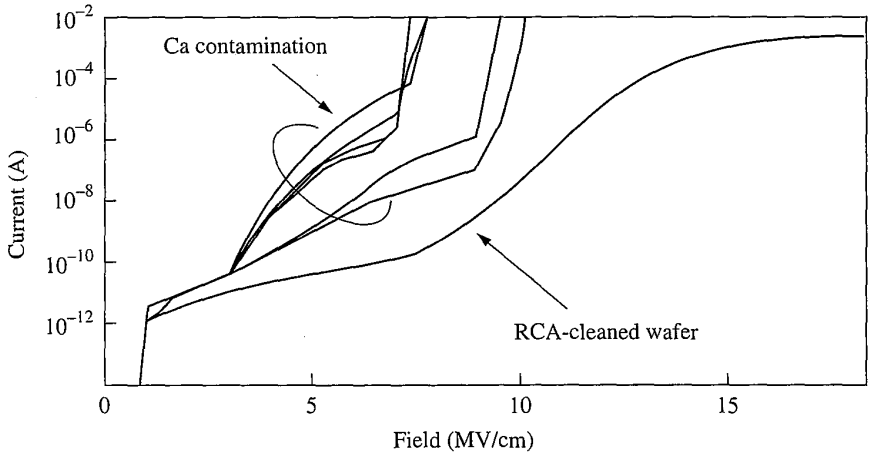


FIGURE 10
A typical *I-V* curve obtained from a Ca-contaminated wafer. (*After Verhaverbeke, Ref. 17.*)

TABLE 6
Typical metallic contamination after the last cleaning step followed by rinsing ($\times 10^{10}$ atom/cm^2), measured by VPD (vapor-phase decomposition)-TRXF[18]

	K	Ca	Cr	Fe	Ni	Cu	Zn
RCA	0.3	8.6	0.2	5.1	3.3	0.3	0.4
HF	0.1	3.8	0.05	0.3	0.1	0.06	0.1
HF–H_2O_2	0.6	1.6	0.3	2.2	0.2	0.09	1.2
BHF	0.2	1.4	0.4	2.6	0.3	3.7	0.7

Micron Ex. 1023, p. 97
Micron v. YMTC
IPR2024-00909

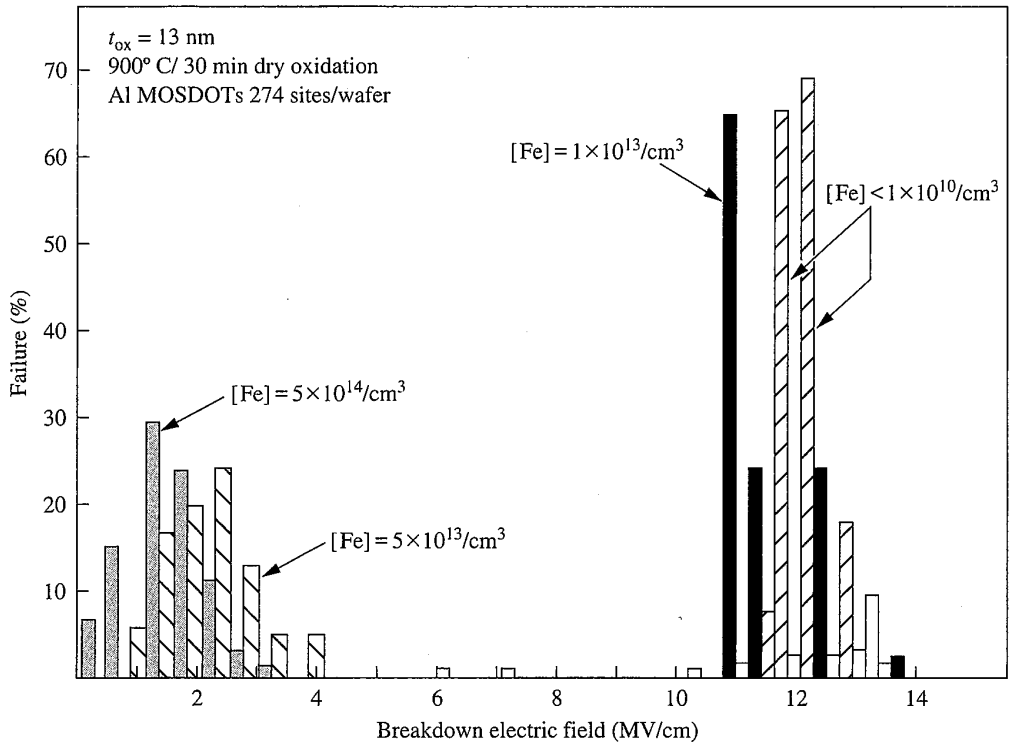


FIGURE 11
Combined summary breakdown field histograms for 13 nm oxides. All oxides were grown in dry O_2 at 900°C. The threshold contamination level for this oxide thickness is between 1×10^{13} and 5×10^{13} cm^{-3}. (*After Henley et al., Ref. 9.*)

Characterization of metal contamination level is important when one wants to control the metallic contamination. Characterization techniques are compiled in Table 7. Note that the sensitivity limits are quoted in cm^{-2} or cm^{-3}, corresponding to surface or bulk concentration. Atomic absorption spectroscopy (AAS), surface photovoltage (SPV), and total reflection x-ray fluorescence (TRXF) have the highest sensitivity. The SPV measurement can identify only Fe contamination. On the other hand, it is suitable for routine monitoring of Fe because it is nondestructive.

Organic contamination

Organic contaminations on the silicon surface are usually from organic vapor in the ambient, storage containers, and the residue of photoresist. The presence of organic impurities on the surface will cause incomplete cleaning of the surface, leaving contaminants such as the native oxide or metal impurities,[19,20] and will also cause a micromasking effect in the later RIE process. Residual photoresist is the main source of organic contamination in the IC processes. In the present cleaning processes, photoresist is usually stripped by a combination of dry ashing by ozone with subsequent wet cleaning in H_2SO_4–H_2O_2 = (3 : 1 to 4 : 1 at 120 to 130°C). Most of the organic residue is removed in the first dry ashing process; the wet process is employed to clean the wafer completely. However, the high cleaning temperature depletes the concentration of H_2O_2 and makes the process difficult to control. Recently, ozone-injected ultrapure water[21] or the sequence of UV and a filter system[22] to

Micron Ex. 1023, p. 98
Micron v. YMTC
IPR2024-00909

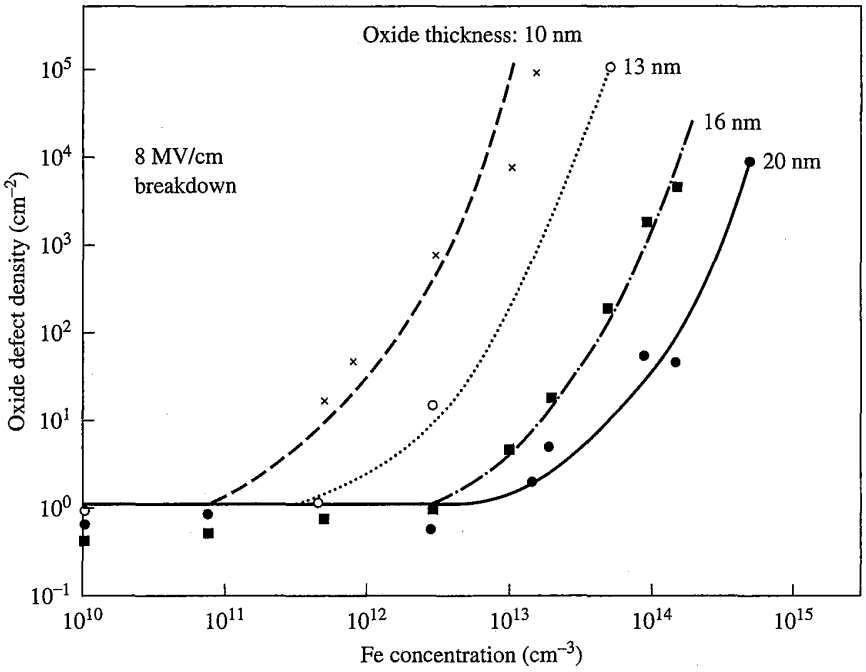


FIGURE 12
MOSDOT capacitor experimental data: oxide defect density versus iron contamination level for 900°C dry oxides. A defect is defined as breakdown occurring at less than 8 MV/cm during a ramp voltage *I-V* test. (*After Henley et al., Ref. 9.*)

remove organic impurities was proposed and designed. The mechanism is as follows: When ozone is dissolved in ultrapure water, it decomposes and becomes a strong oxidizing agent that decomposes organic impurities. This process has the advantages of lower temperature, simplicity, and reduced chemical consumption and chemical waste. However, this method causes native oxide to grow on the Si wafer because of

TABLE 7
Analytical techniques for the detection of metallic impurities in silicon[9]
Sensitivities are typical reported values for routine operation.

Techniques	Sensitivity	Chemical ID?	Destructive?	SPC†† line monitor?
SIMS†	10^{15} cm^{-3}	Yes	Yes	No
TRXF‡	10^{10} cm^{-2}	Yes	No	Maybe
AAS§	10^{9} cm^{-2}	Yes	Yes	No
DLTS¶	10^{11} cm^{-3}	Yes	Yes	No
SPV††	10^{9} cm^{-3}	Iron only	No	Yes

† Secondary-ion mass spectroscopy
‡ Total-reflection x-ray fluorescence
§ Atomic absorption spectroscopy
¶ Deep-level transient spectroscopy
†† Surface photovoltage
†† Standard process control

Micron Ex. 1023, p. 99
Micron v. YMTC
IPR2024-00909

the ozone. The oxide thickness increases as the immersion time increases, and it also increases with the concentration of ozone. Figure 13 shows the native oxide growth caused by ozone-injected ultrapure water.[21] Rinsing in 2 ppm ozoned ultrapure water results in a very rapid growth of native oxide. The surface roughness using the ozone-injected cleaning process is the same as that from the H_2SO_4–H_2O_2 cleaning process.

Figure 14 shows the influence of residual surfactant molecule on the silicon wafer surface.[21] From the measurement of x-ray photoelectron (Si_{2p}) spectra, the wafer with the surfactant still has native oxide on the surface after the dilute HF dip (DHF) process. The other sample, with ozone treatment, exhibits an almost native-oxide-free surface after the DHF process. The organic contaminant–free surface resulting from ozone treatment helps the subsequent cleaning to function properly.

The contact angle of a droplet of DI water on the silicon surface is usually used to indicate the chemical condition of the surface. With poor passivation the surface free energy is high and results in a low contact angle (Fig. 15*a*), whereas a surface with good passivation exhibits a high contact angle (Fig. 15*b*). Figure 15*c* shows the contact angle as the function of the ozone concentration on a Si surface with absorbed surfactant.[21] The contact angle of the absorbed surfactant on Si surface is around 56.6°. After 10 min rinsing in ultrapure water, the contact angle is only 62°, which is lower than the initial contact angle, 78.9°. However, if the ozone concentration increases to as little as 0.28 ppm, the contact angle increases to the initial value at the same time. Increasing the concentration to 1.5 ppm reduces the clean time. Hence, the ozone-injected ultrapure water is an effective cleaning process to remove

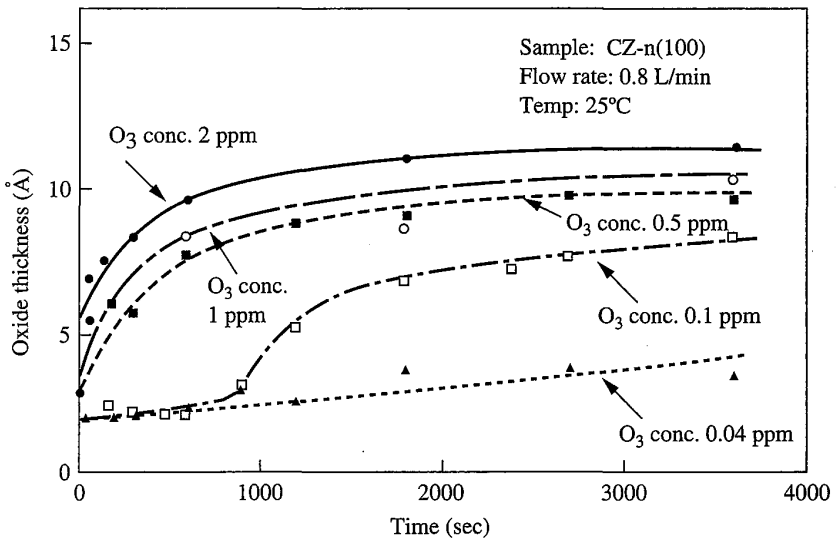


FIGURE 13
Native oxide growth on the wafer surface by ozone-injected ultrapure water as a function of dipping time and ozone concentrations. The horizontal axis shows immersion time, and the vertical axis shows the thickness of the native oxide formed in the ozone-injected ultrapure water. Ozone concentrations are 2, 1, 0.5, 0.1, and 0.04 ppm. Very rapid growth of native oxide is confirmed in the 2 ppm ozonized ultrapure water. (*After Ohmi et al., Ref. 21.*)

Micron Ex. 1023, p. 100
Micron v. YMTC
IPR2024-00909

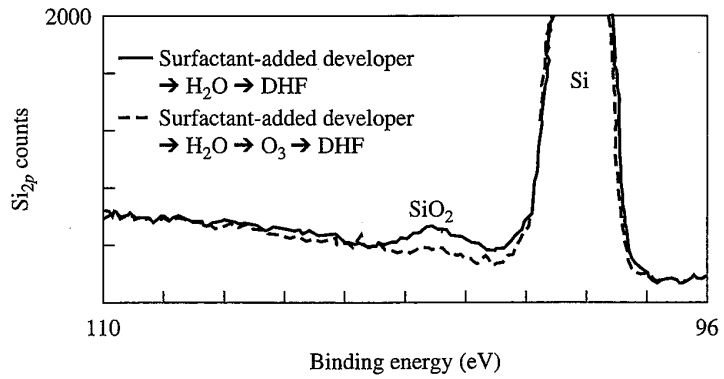


FIGURE 14
Influence of residual surfactant molecules on the silicon wafer surface. The solid line is the sample without the ozone-injected ultrapure water cleaning; the broken line is the cleaned sample. (*After Ohmi et al., Ref. 21.*)

the residual organic impurities on Si surface. It is a room-temperature process and can replace the conventional H_2SO_4–H_2O_2 process.

Surface microroughness

Surface microroughness is an important factor in the manufacture of high-performance and high-reliability submicrometer and deep submicrometer ULSIs.[23–26] This is because the thickness of the thin dielectric, which is grown after the conventional RCA cleaning process, is reduced to less than 100 Å for 0.35 μm technology. For the 0.1 μm technology of the year 2000, the thickness of the gate oxide will be reduced to only 40 Å. Then the surface requirement should be atomic flatness. In the conventional RCA cleaning processes, NH_4OH–H_2O_2–H_2O is used as the first step to remove particle, organic, and some metal impurities. It has been reported that this conventional process with NH_4OH–H_2O_2–H_2O = 1 : 1 : 5 at 70°C for 10 to 15 min results in a rough surface. The mechanism producing this rough surface is that the NH_4OH acts as the etchant of the oxide while H_2O_2 acts as the oxidant. The processes of etching and oxidizing progress simultaneously in the SC-1 solution; consequently, the surface becomes rough. Methods of reducing the microroughness can be summarized as

1. Reduce the proportion of NH_4OH (the etchant).
2. Reduce the temperature of the bath.
3. Reduce the cleaning time.

The scanning tunneling microscope (STM) is the most commonly used instrument to investigate the topography of the silicon wafer. Note that the native oxide on the silicon surface affects the results of an STM study. The existing native oxide reduces the tunneling current and influences the resolution. The atomic force microscope (AFM) is also used to measure the topography of an insulator, such as thin oxide. Figure 16 shows the surface microroughness of an n-type Czochralski (CZ) wafer as a function of various etching solutions (10 min immersion) such as HF (0.5%), conventional BHF (NH_4F: 35 to 38%), and advanced surface-active BHF

Micron Ex. 1023, p. 101
Micron v. YMTC
IPR2024-00909

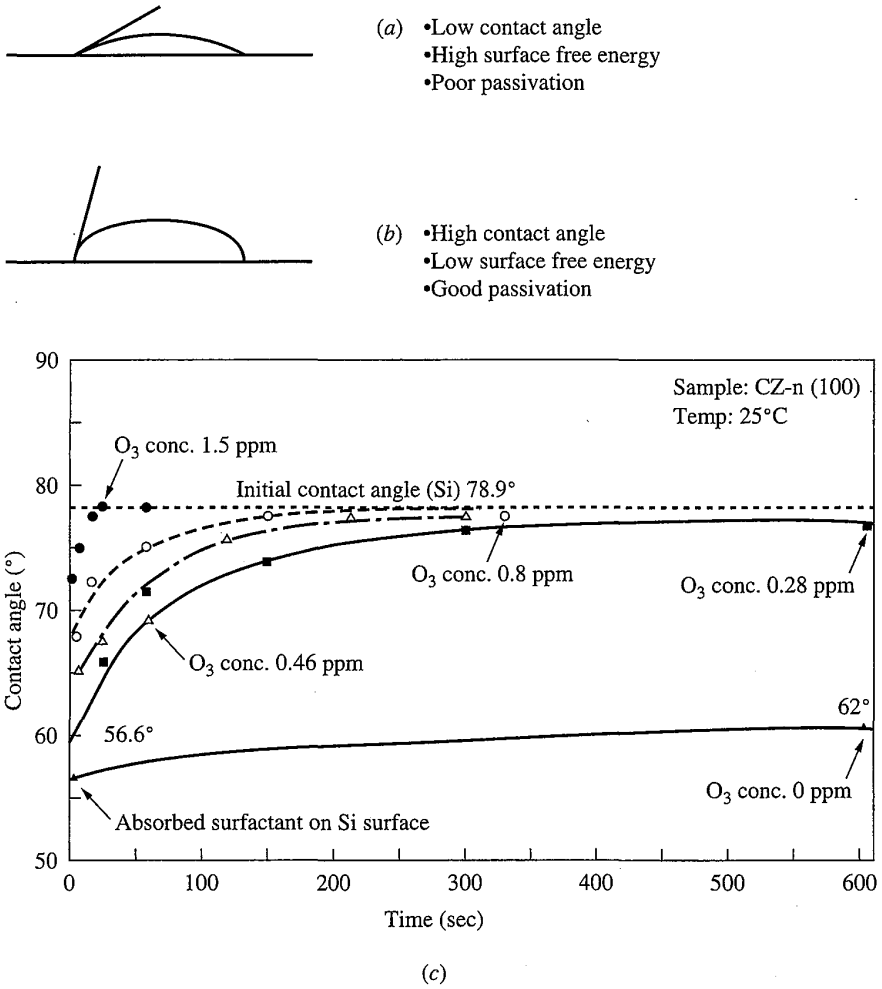


FIGURE 15
Schematic representation of the contact angle measurements for (*a*) poor passivation and (*b*) good passivation. (*c*) Time dependence for eliminating absorbed surfactant molecule from the silicon surface using ozonized ultrapure water, as measured by the contact angle. The horizontal axis shows immersion time; the vertical axis shows contact angle. Ozone concentration in the ozone-injected ultrapure water was varied. Very rapid recovery, up to the initial contact angle, has been observed under 1.5 ppm ozone concentration. (*After Ohmi et al., Ref. 21.*)

(NH_4F : 17%) with surfactant.[25] It is clearly seen from the figure that the average roughness, R_a, is largest with conventional BHF, followed by the HF. The advanced BHF has the same level as the control sample treated in advanced BHF for 1 min.

The R_a values of the other conventional cleaning processes are shown in Fig. 17:

H_2SO_4(98%)–H_2O_2(30%) (4 : 1) (SPM)
NH_4OH(28%)–H_2O_2(30%)–H_2O (1 : 1 : 5) (SC-1)
HCl(36%)–H_2O_2(30%)–H_2O (1 : 1 : 6) (SC-2)

Micron Ex. 1023, p. 102
Micron v. YMTC
IPR2024-00909

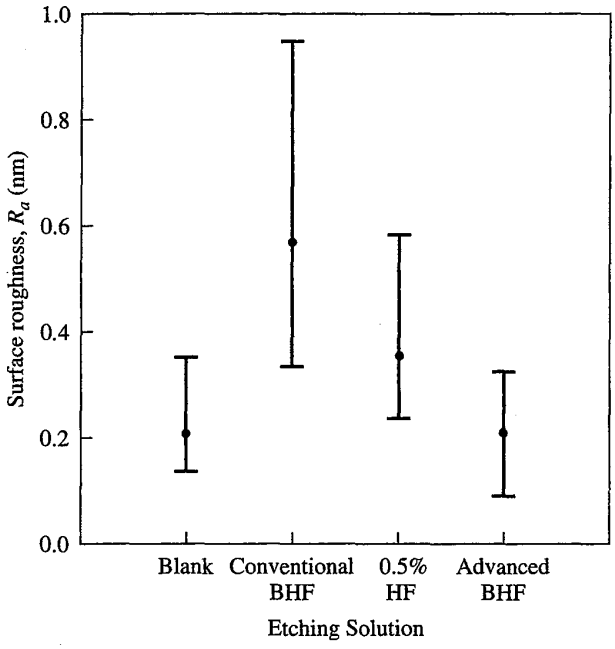


FIGURE 16
Relationship between the average surface microroughness (R_a) of an n-type CZ wafer and various etching solutions. A "blank" wafer is a wafer treated in advanced BHF for 1 min. (*After Miyashita et al., Ref. 25.*)

No significant increase of roughness is found for the SPM and SC-2 cleaning processes. However, the SC-1 cleaning process has the largest R_a and deviation.[24] The proportion of NH_4OH in the cleaning solution has a significant effect on microroughness. Figure 18*a* shows the dependence of the surface microroughness of CZ and floating-zone (FZ) wafers on the SC-1 cleaning process with various NH_4OH ratios.[25] The figure shows that the R_a is reduced to the level of SC-2 cleaning when

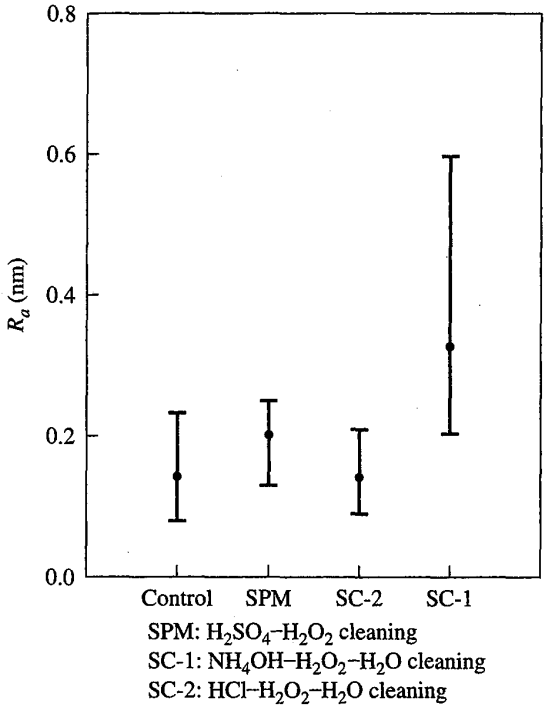
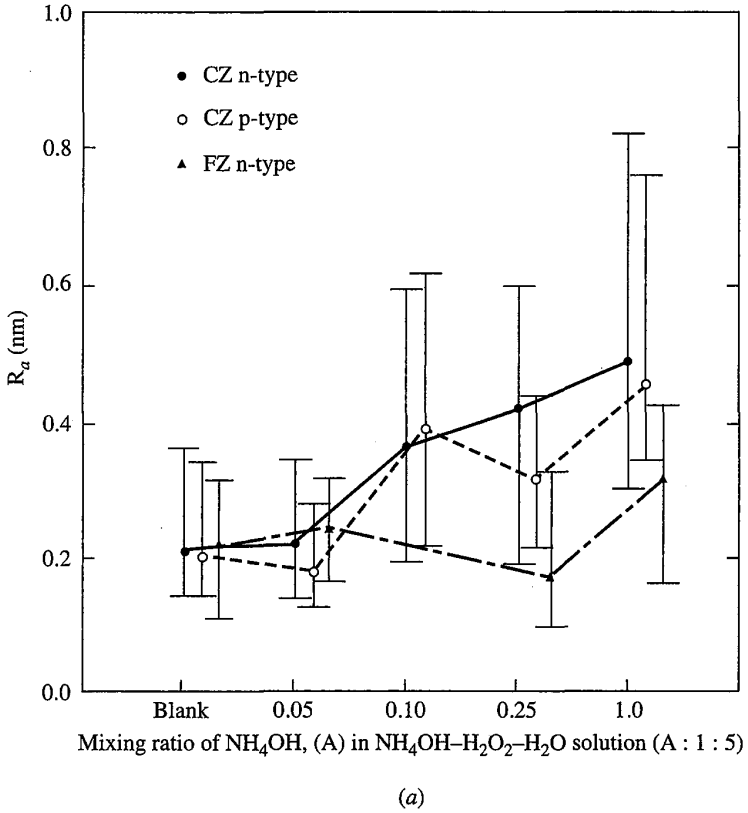


FIGURE 17
Surface microroughness on a silicon wafer treated with SC-2 cleaning, SC-1 cleaning, and fourth-time cyclic SPM cleaning. (*After Ohmi et al., Ref. 24.*)

Micron Ex. 1023, p. 103
Micron v. YMTC
IPR2024-00909



(*a*)

FIGURE 18
Relationship of surface microroughness of (*a*) CZ and FZ wafers having various NH_4OH mixing ratios in SC-1 solution. (*After Miyashita et al., Ref. 25.*)

the ratio of NH_4OH is reduced from 1 to 0.05. The FZ wafer exhibits weak dependence on the NH_4OH ratio, because of the lower Si vacancy concentration in an FZ wafer than in the CZ wafer; a 4-hr wet oxidation at 1000°C can overcome this problem.[25] The surface microroughness of a wafer surface treated by a complete RCA cleaning process with four different mixing ratios in SC-1 is shown in Fig. 18*b*. As the ratio is increased, the surface becomes rougher. The succeeding steps of SC-1 cleaning also influence the surface condition. Rinsing at high temperature results in a larger R_a.[24] Hence, for the advanced SC-1 cleaning process the ratio of NH_4OH must be reduced and the wafers rinsed in DI water at room temperature for an atomically flat surface.

One problem that must be taken into consideration is that when the concentration of NH_4OH is reduced, the efficiency of particle and metal impurity removal is reduced. The particle removal efficiencies as a function of NH_4OH mixing ratio are shown in Fig. 18*c*; in fact, the particle removal efficiency optimizes at a ratio from 0.05 to 0.10. Nor does the metal-impurities removal efficiency change when the ratio of NH_4OH is reduced from 1.0 to 0.05 (Fig. 18*d*).

The etched thermal oxide depth in various SC-1 mixtures as a function of NH_4OH concentration at 70°C for 10 min is shown in Fig. 19. The etching rate

Micron Ex. 1023, p. 104
Micron v. YMTC
IPR2024-00909

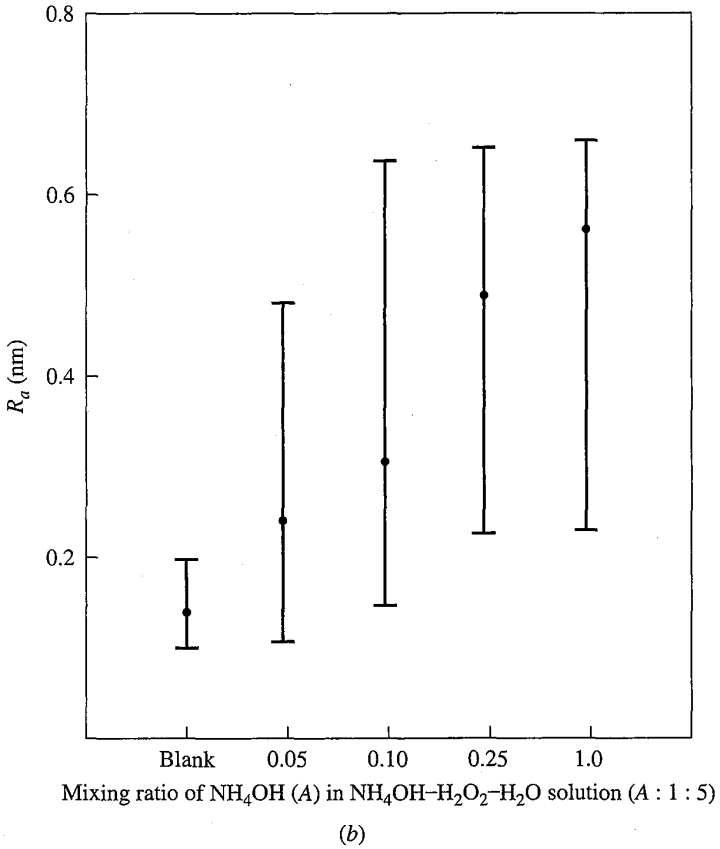


FIGURE 18 (*continued*)
(*b*) silicon wafers treated in an entire RCA cleaning process with four different NH_4OH mixing ratios in SC-1 solution. (*After Miyashita et al., Ref. 25.*)

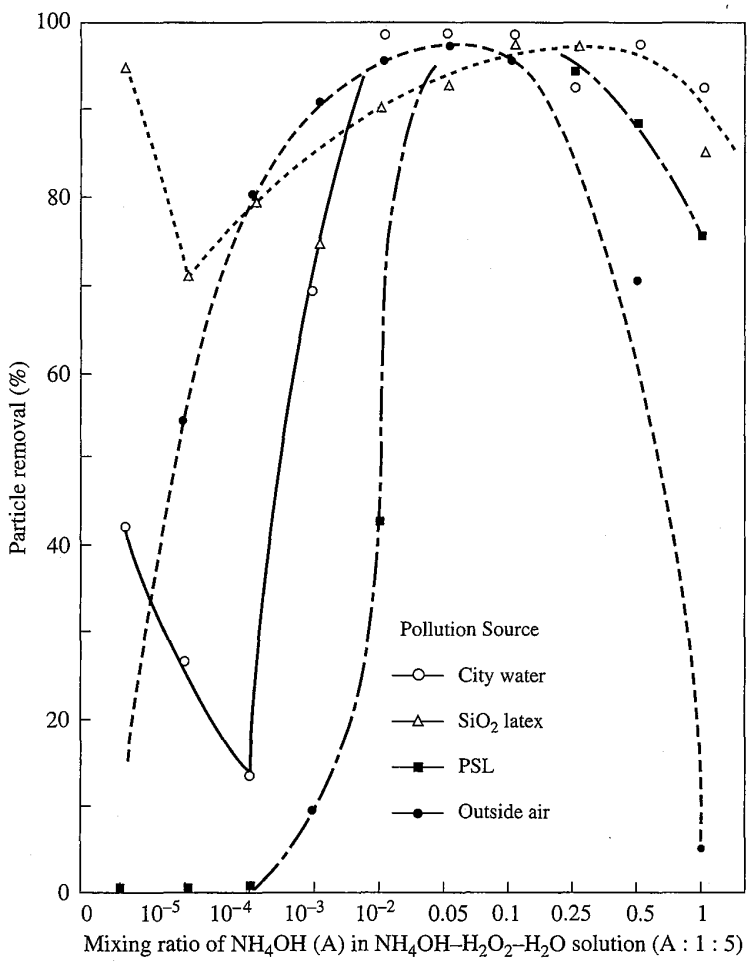
increases as the concentration of NH_4OH increases. For the conventional concentration the etching rate is 3 nm for a 10-min immersion at 70°C. The higher the bath temperature, the higher the etching rate.

Hence, to avoid microroughness the ratio of NH_4OH can be reduced to between 0.05 and 0.25. This ratio can still maintain the efficiency of removing particles and metallic contaminations.

The effects on reliability of the gate oxide and the device performance resulting from the microroughness have been discussed.[26] The dependence of the breakdown field on microroughness R_a is shown in Fig. 20*a*. The larger the R_a, the lower the breakdown field. The trend is the same for oxide thicknesses of 9.8 and 8.2 nm. Figure 20*b* shows the charge to breakdown, Q_{bd}, versus the surface microroughness.[26] As R_a increases to 0.8 nm, the Q_{bd} value decreases almost an order of magnitude.

In addition to the influence on the quality of oxide, microroughness also has a significant influence on the performance of the devices. Typical current–voltage characteristics for two MOS transistors with average surface microroughness of 0.6 and 0.3 nm are shown in Fig. 21. The transistor with an R_a of 0.6 nm has lower current derivability than that with an R_a of 0.3 nm. The channel electron mobility,

Micron Ex. 1023, p. 105
Micron v. YMTC
IPR2024-00909



(*c*)

FIGURE 18 (*continued*)
(*c*) Particle removal efficiencies as a function of NH_4OH mixing ratio in SC-1 cleaning for four different particles: PSL (polystyrene latex spheres), silica latex, particles from city water, and particles from outside air. (*After Ohmi et al., Ref. 24.*)

as a function of SiO_2–Si interface microroughness, is shown in Fig. 22. The mobility of electrons in the sample with $R_a = 1.5$ nm reduces to 35% of the reference value, 0.2 nm. These results clearly indicate the importance of surface microroughness for MOS transistor performance, especially in speed and reliability.

Native oxide

Native oxide that is free on a silicon surface becomes very important in the advanced ULSI process.[27–29] A silicon surface with native oxide results in uncontrollable ultrathin oxide growth, high contact resistance, and the inhibition of selective chemical vapor deposition (CVD) or epitaxy. In addition, some metallic impurities are included in native oxide during the cleaning process. If this native oxide is not completely removed, it serves as the source of metallic impurities, which diffuse

Micron Ex. 1023, p. 106
Micron v. YMTC
IPR2024-00909

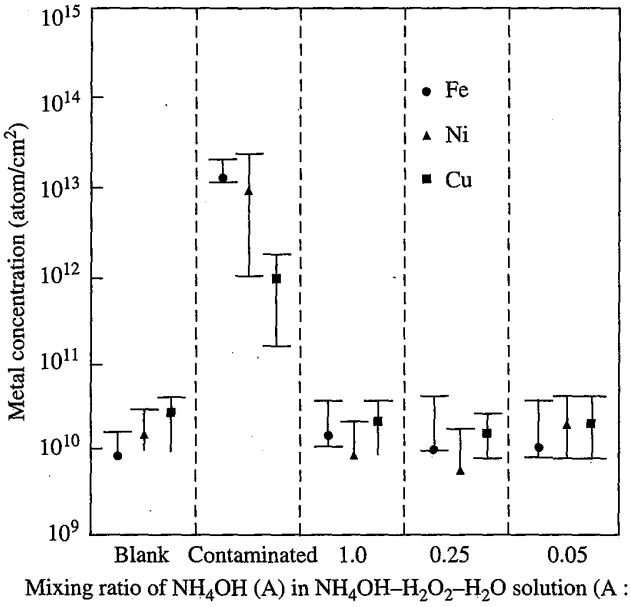


FIGURE 18 (*continued*)
(*d*) Metallic impurities Fe, Ni, and Cu removal efficiency by SC-1 cleaning with three different NH_4OH mixing ratios. (*After Ohmi et al., Ref. 24.*)

into the silicon or precipitate at the interface of the SiO_2–Si, resulting in defects in the succeeding high-temperature process. Therefore, a silicon surface that is native oxide–free is a key factor in obtaining high performance and reliability of semiconductor devices. Native oxide thickness is usually determined by x-ray photoelectron spectroscopy (XPS) and ellipsometry. It is difficult to determine the

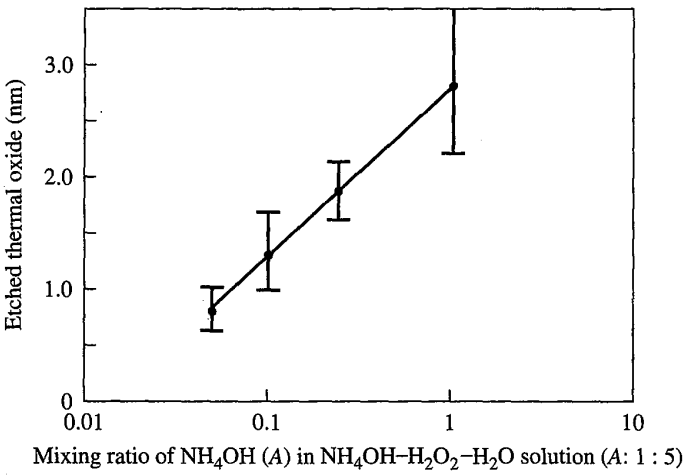


FIGURE 19
Etch depth of thermal oxide after 10 min in various SC-1 mixtures as a function of ammonia concentration of the mixture at 70°C. (*After Verhaverbeke, Ref. 2.*)

Micron Ex. 1023, p. 107
Micron v. YMTC
IPR2024-00909

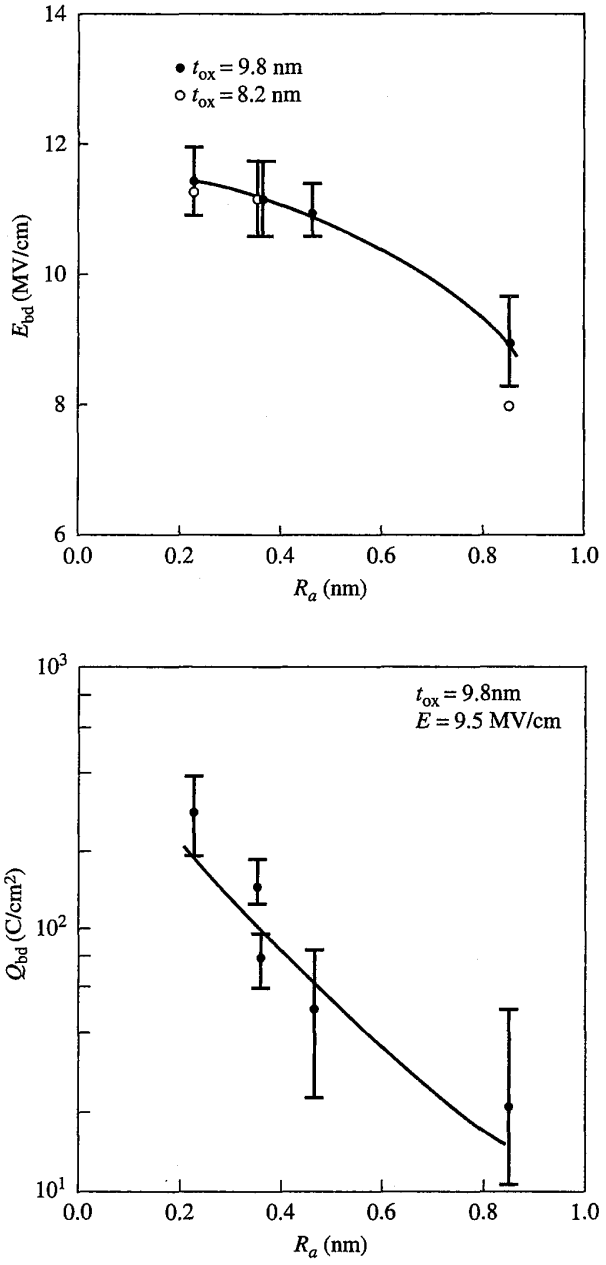


FIGURE 20
(*a*) Dielectric breakdown field intensity (E_{bd}) plotted as a function of average surface microroughness for two oxide films, having thicknesses of 9.8 and 8.2 nm formed on an n-type substrate. (*b*) Surface microroughness dependence of Q_{bd} (C/cm^2) under a constant field of 9.5 MV/cm. (*After Ohmi et al., Ref. 26.*)

thickness and refractive index simultaneously with an ellipsometric measurement when the thickness is less than 200 Å. Hence, the native oxide thickness is determined by using a fixed refractive index of 1.46. The Si_{2p} spectra of XPS is also used to measure the thickness of native oxide. However, this method does not obtain the result directly. The area ratio of the signal from thicker oxide to that of the silicon substrate must be calibrated first by ellipsometry. Then the thickness of the native oxide can be obtained from the area of the XPS.

The growth of native oxide in air at room temperature (23.7°C) with a humidity of 42% is covered in Ref. 30. The oxide thickness increases as the exposure time increases. The thickness of the native oxide shows a steplike increase of 2 to 3 Å. This

Micron Ex. 1023, p. 108
Micron v. YMTC
IPR2024-00909

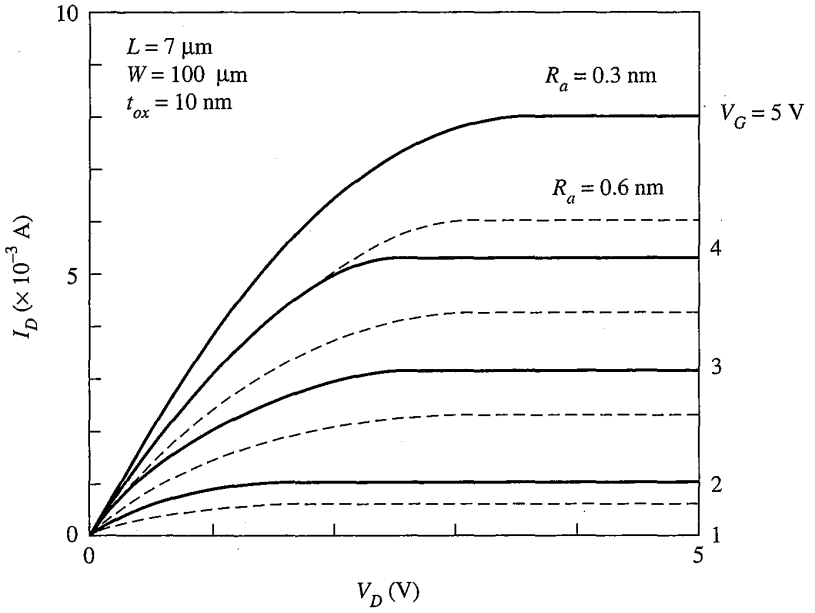


FIGURE 21
Typical current-voltage characteristics for two MOS transistors with average surface microroughness of 0.6 nm (dashed line) and 0.3 nm (solid line). (*After Ohmi et al., Ref. 26.*)

can be considered as a layer-by-layer growth of native oxide on a wafer at room temperature. N-type silicon with a high doping concentration (10^{20} atoms cm^{-3}) grows thicker native oxide than less-doped silicon (10^{15} cm^{-3}). After 7 days of exposure to cleanroom air containing 1.2% H_2O, the thickness of native oxide is 6.7 Å, which is larger than that of samples in an $N_2 + O_2$(1.7Å) or N_2(1.9Å) environment with moisture less than 0.1 ppm.[27]

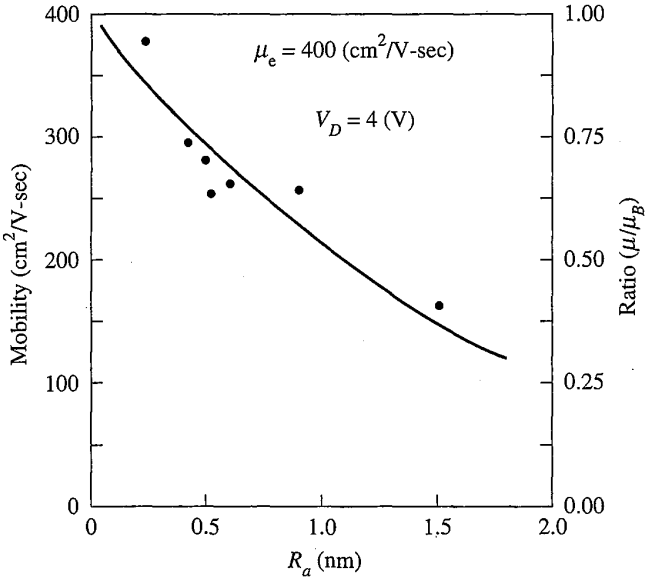


FIGURE 22
Relationship of channel electron mobility (μ_e) to the Si–SiO_2 interface microroughness obtained in the current saturation region. The channel electron mobility normalized by the bulk electron mobility is also plotted. (*After Ohmi et al., Ref. 26.*)

Micron Ex. 1023, p. 109
Micron v. YMTC
IPR2024-00909

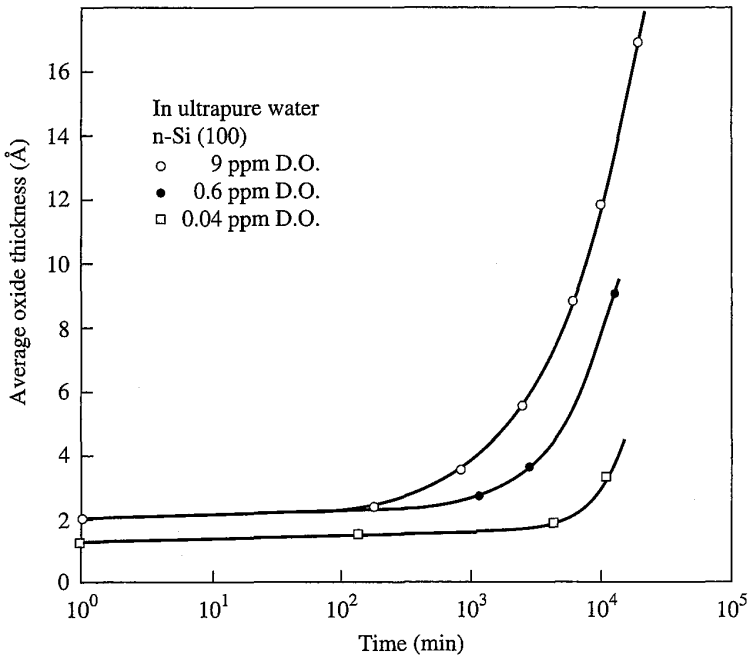


FIGURE 23
Oxide thicknesses of wafers as a function of immersion time in ultrapure water at room temperature for different dissolved oxygen (D.O.) concentrations. (*After Morita et al., Ref. 30.*)

The dissolved oxygen in DI water enhances the growth of native oxide.[30] Figure 23 shows the oxide thickness of wafers as a function of immersion time in ultrapure water for different dissolved oxygen concentrations. The thickness of native oxide increases as the dissolved oxygen concentration increases. The curves increase parabolically, unlike those of layer-by-layer growth in air. Si dissolves into the water simultaneously during native oxide growth. The number of dissolved Si atoms is larger than the number of Si atoms in native oxide by one order of magnitude. This will result in a rougher surface.

A model of how native oxide grows in air and in ultrapure water has been proposed.[30] Figure 24 shows the model of native oxide grown in air. The oxygen

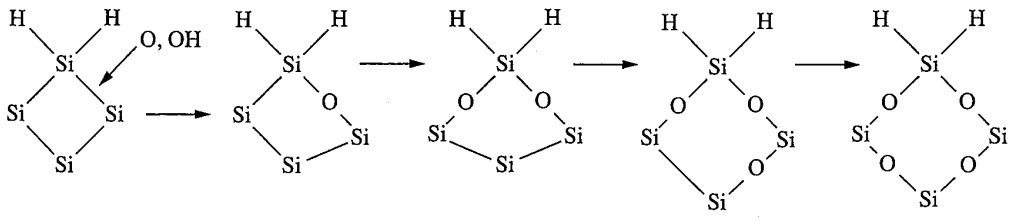


FIGURE 24
A model of native oxide growth in air. (*After Morita et al., Ref. 30.*)

Micron Ex. 1023, p. 110
Micron v. YMTC
IPR2024-00909

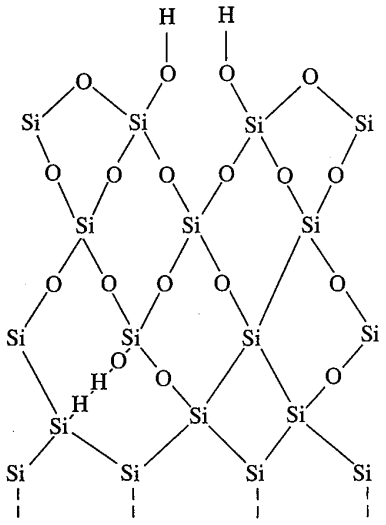


FIGURE 25
A model of native oxide growth in ultrapure water. (*After Morita et al., Ref. 30.*)

species break Si–Si bonds first. The surface still remains hydrogen-terminated and hydrophobic. For the surface grown in ultrapure water shown in Fig. 25, the surface is terminated by O or OH; the Si–H bond exists only at the SiO_2–Si interface.

Some methods have been proposed to reduce the dissolved oxygen.[31] The first stage is a degassing-membrane module. Oxygen in the water rapidly diffuses through the membrane, which is surrounded by a low-pressure vacuum. After this stage, the oxygen concentration is reduced to 300 to 400 ppb; then two methods can be employed. The first is reduction with catalytic resins by N_2H_4 or H_2 (Pd catalyst) according to the following reactions:

$$O_2 + N_2H_4 \rightarrow 2H_2O + N_2$$

or

$$O_2 + 2H_2 \rightarrow 2H_2O$$

The residual concentration of dissolved oxygen can be reduced to 5 ppb or less by using this method. The other method is N_2 gas bubbling, which reduces the partial pressure of oxygen in the water. Theoretically, the oxygen concentration can be reduced to less than 0.1 ppb with this method.

An Al–Si Schottky diode contact made with a native oxide–free process exhibits a great improvement, such as a low contact resistance of 0.4 $\mu\Omega$-cm^2, and an ideality factor, n, of 1.02 without any thermal treatment.[32] The native oxide plays a more important role when the gate oxide thickness is only 50 to 60 Å.[33,34] A *preoxide* process, to prevent the increase of surface microroughness during the high-temperature treatment of Si wafers, has been reported. The silicon surface is intensively oxidized at 300°C in ultraclean oxygen to form one molecular layer of oxide as a passivation layer after the desorption of hydrogen from the surface by the HF-clean. Dielectric breakdown histograms of this process are shown in Fig. 26*a* and, with the control samples, in Fig. 26*b*. Early breakdown of the dielectric is found for controlled

Micron Ex. 1023, p. 111
Micron v. YMTC
IPR2024-00909

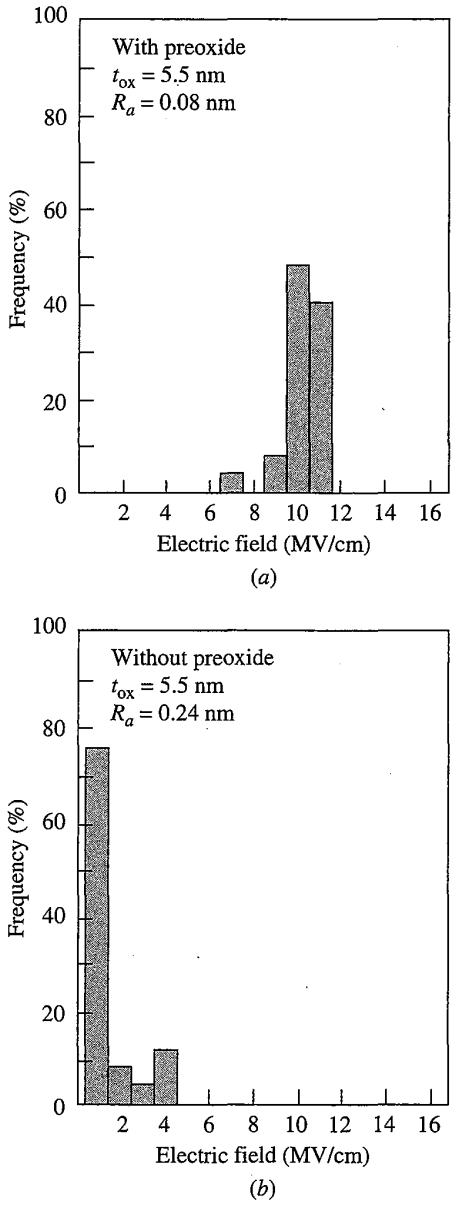


FIGURE 26
Dielectric breakdown histograms for $Al/SiO_2/n\text{-}Si(100)$ diodes (*a*) with and (*b*) without preoxide. (*After Makihara et al., Ref. 34.*)

samples. Figure 27 shows the current density–voltage characteristics of 5.5 and 9.0 nm oxide with and without the preoxide processes.[34] The leakage current in the conventional dry oxide is much higher than that in the ultraclean preoxide. This indicates that the preoxide-controlled, ultraclean oxide has a higher electrical insulating capability, particularly for ultrathin oxide.

Last step of wafer cleaning for gate oxidation/epitaxy

Recently, epitaxy has become more important in high-speed devices, such as high-electron-mobility transistors (HEMTs) or SiGe devices.[35] Gate oxide yield with HF-last cleaning increases to 74%, comparing to the conventional RCA-last cleaning

Micron Ex. 1023, p. 112
Micron v. YMTC
IPR2024-00909

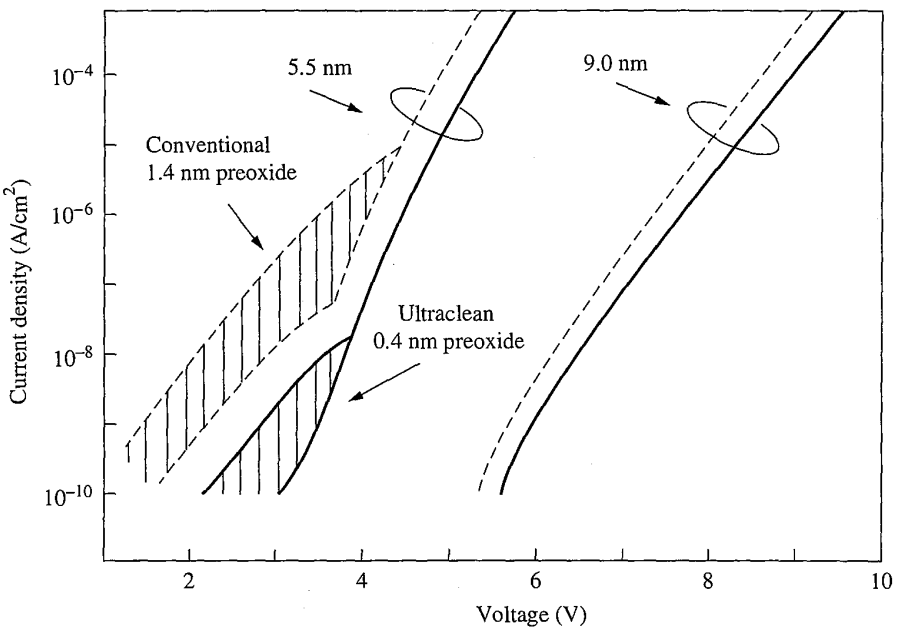


FIGURE 27
Current density-voltage characteristics of n^+-polycrystalline $Si/SiO_2/p$-Si(100) diodes under negative-biased metal electrodes with and without preoxide. (*After Makihara et al., Ref. 34.*)

(47%). This increase is due to the reduced metallic contamination as contaminants included in the oxide can be removed in the HF-last cleaning.[36] Native oxide–free silicon is the key requirement for producing a high-quality epitaxy film.[37,38] Dilute HF solution is usually used to remove the native oxide on the semiconductor surface. After the HF dipping, most of the silicon surface (80%) is covered by H-terminated bonds, and the rest are covered by F, O, and C terminations. However, if the hydrogen termination of the surface is incomplete, subsequent exposure to air causes absorption of carbonic impurities. The F coverage also depends on the orientation, HF concentration, and immersion time.[39] Figure 28 shows the ideal hydrogen termination on a Si(100) and a Si(111) surface. On a Si(100) surface, a surface Si atom has two dangling bonds that can be passivated with hydrogen; on a Si(111) surface, each surface Si atom has only one dangling bond for hydrogen passivation. The real surface is always terminated by both monohydrides and dihydrides, resulting from an imperfect, defective, or rough surface.

High-resolution electron energy loss spectroscopy (HREELS) is commonly used to analyze the chemical composition of a treated surface. Figure 29 shows the vibrational spectra of Si(100) and Si(111) surfaces after a 1 min dip in 40% HF.[39] The Si(100) is covered by dihydride at 2100 cm^{-1} (stretching mode), 900 cm^{-1} (scissor mode), and a wagging or bending mode at 640 cm^{-1}. On the other hand, the Si(111) surface is covered predominantly by monohydride (stretching mode at 2080 cm^{-1} and bending mode at 640 cm^{-1}). CH_x bonds can be found on both surfaces. They have wavenumbers of 2950 and 1350 cm^{-1} for stretching and deformation vibration. The spectra of the surfaces after they are rinsed in DI water are shown in

Micron Ex. 1023, p. 113
Micron v. YMTC
IPR2024-00909

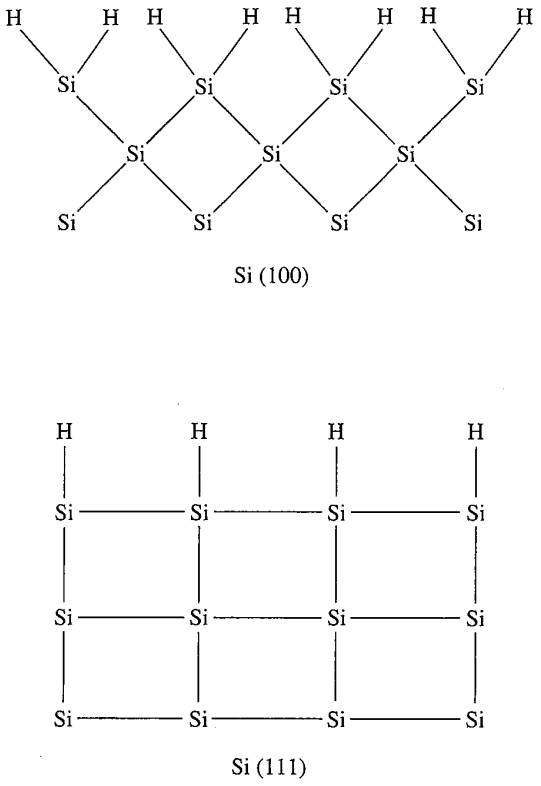


FIGURE 28
Schematic representation of the ideal hydrogen terminations on Si(100) and Si(111) surfaces.

Fig. 30. A Si–OH vibration is found after only a 30-sec water rinse. This presence of Si–OH is explained by the fast exchange of Si–F groups with water. The silicon surface, covered with hydrogen from the HF-last process, is hydrophobic. The HREELS spectrum of a hydrophilic surface after the SC-1 and SC-2 wet cleaning process has a strong asymmetric Si–O–Si vibration. Hence, the OH groups are responsible for the hydrophilic nature of the Si surface.[39]

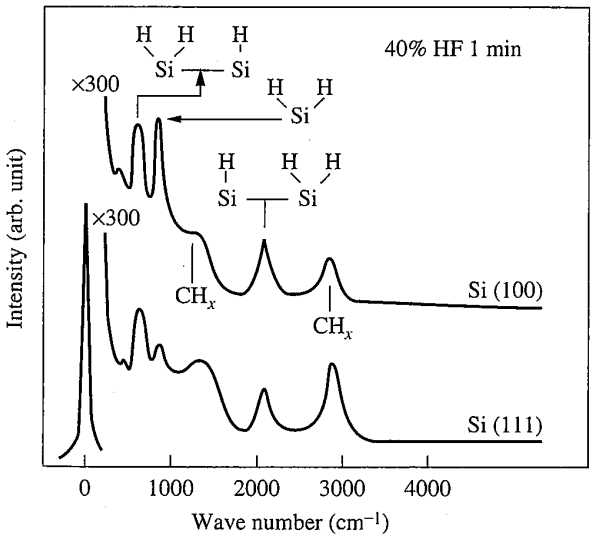


FIGURE 29
Vibrational spectra of Si(100) and (111) surfaces after a 1-min dip in 40% HF. Only Si–H (640 cm^{-1}, 900 cm^{-1}, and 2100 cm^{-1}) and hydrocarbon vibrations (2950 cm^{-1} and 1350 cm^{-1}) contribute to the spectra. (*After Grundner et al., Ref. 39.*)

Micron Ex. 1023, p. 114
Micron v. YMTC
IPR2024-00909

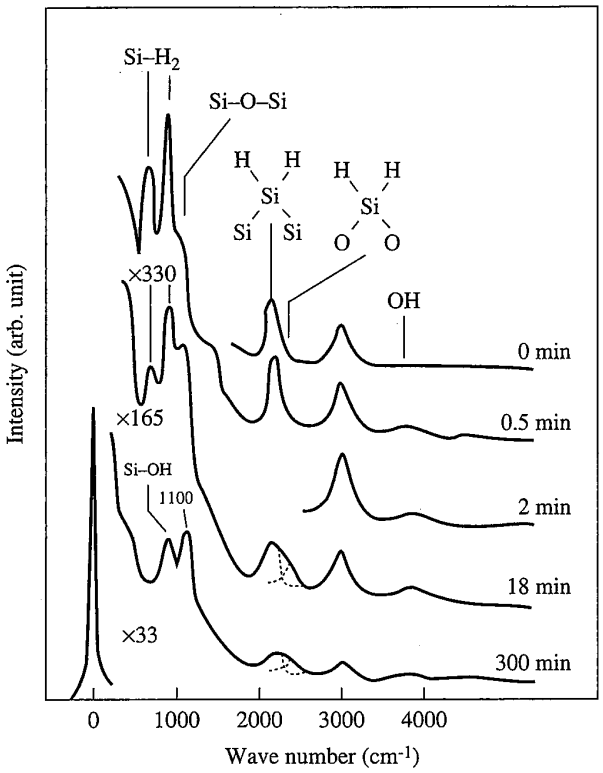


FIGURE 30
Vibrational spectra of a Si(100) sample (treated in 40% HF for 1 min) after different water rinsing times. Changes in the spectra occur around the OH, Si–O–Si, and Si–H stretching frequencies. The latter develops a schematically shifted component, indicated as O_2–Si–H_2. After 300 min there is still an unshifted contribution, suggesting that part of the surface has not yet oxidized. (*After Grundner et al., Ref. 39.*)

Contact angle is the simplest and most effective method to indicate the level of passivation. High passivation results in a high contact angle. Figure 31 shows the contact angle on a silicon surface after HF dipping (0.5%) with and without a DI water rinse.[36] A dip of about 30 seconds increases the contact angle to a maximum of 70.0°. The surface with the DI water rinse has a lower contact angle because of

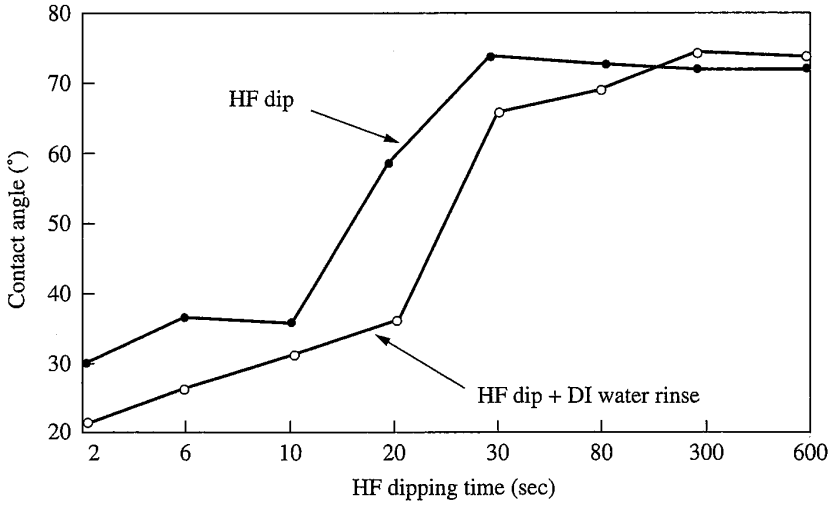


FIGURE 31
Contact angle on Si after HF dipping (0.5%) with and without a DI water rinse. (*After Verhaverbeke et al., Ref. 36.*)

Micron Ex. 1023, p. 115
Micron v. YMTC
IPR2024-00909

partial reoxidation. The contact angle is also a function of the dipping time and the temperature of water. Hot HF (70.0°C) can be used effectively to get the surface completely oxide-free. Some modified HF-last processes have been proposed:

HF/IPA: A small amount of isopropyl alcohol (IPA) (< 1000 ppm) mixed with HF (0.5%) prevents the addition of particles to the Si surface in the traditional HF-last process and preserves the chemical state of the surface. Figure 32 shows the particle density after the HF/IPA cleaning process. When 1000 ppm of IPA is added to the HF mixture, the number of particles decreases to the level of the standard RCA-clean. After a 60-sec dip in HF/IPA, the contact angle of the silicon surface reaches the same maximum value as that achieved in the HF-last process. The correlation between XPS and contact angle measurement for an HF/IPA-dipped Si surface is shown in Fig. 33. The intensity of O_{1s} to SiO_2, as determined by XPS, decreases as the contact angle increases. Therefore, the contact angle is very sensitive to the degree of oxidation of a clean surface.[40]

HF/H_2O_2: H_2O_2 (10%) added to HF (0.5%) can remove metallic impurities to a level of 10^{10}atom/cm^2 in 10 min and can also reduce the number of particles. This has been discussed in connection with metal impurities.[14] This process has been proposed to replace the conventional HF-last process because it improves gate oxide integrity. However, a high H_2O_2 ratio results in a slightly oxidized silicon surface and, consequently, reduces the contact angle.

There are some other additives that can be used in the HF solution, such as HF/ethanol and HF/acetone. However, neither ethanol nor acetone can reduce the number of particles; hence, they are not suitable for the HF-last process.

Conventionally, prior to epitaxy in ultrahigh vacuum (UHV), a high-temperature (1200°C) thermal etching of native oxide is used. Figure 34 shows the in situ Auger

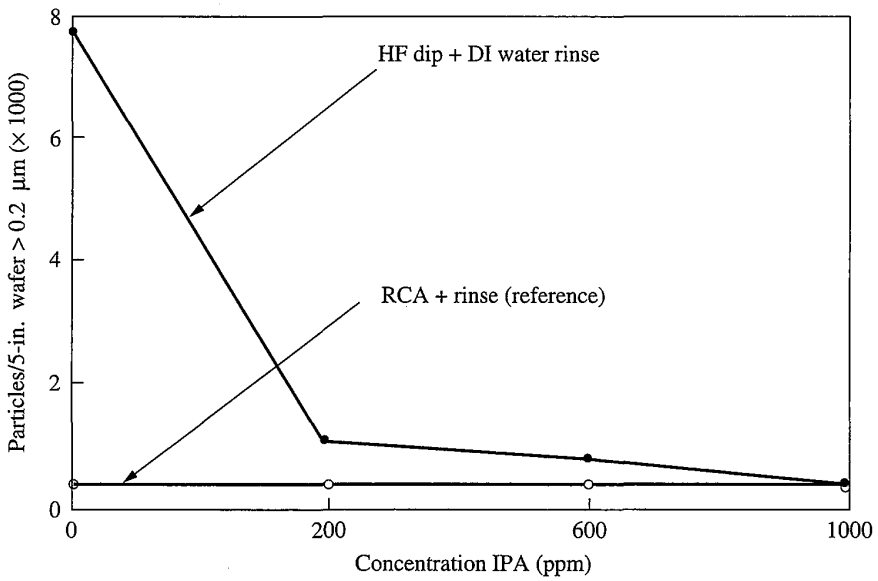


FIGURE 32
Particle density on a 5-inch wafer after dipping in HF with various amounts of IPA added and subsequent DI-water rinsing. (*After Verhaverbeke et al., Ref. 36.*)

Micron Ex. 1023, p. 116
Micron v. YMTC
IPR2024-00909

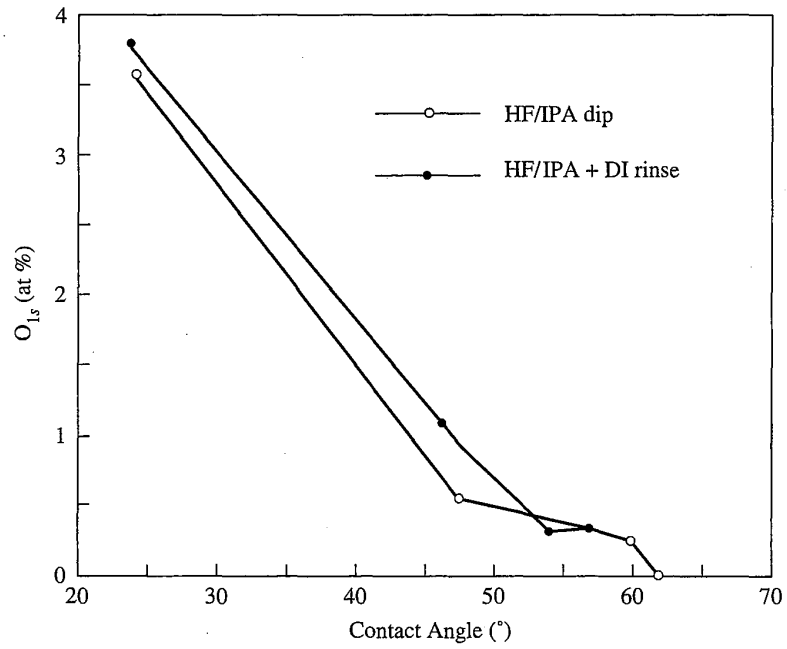


FIGURE 33
Correlation between XPS O_{1s} concentration and contact-angle measurement for HF/IPA-dipped Si surfaces. (*After Alay et al., Ref. 40.*)

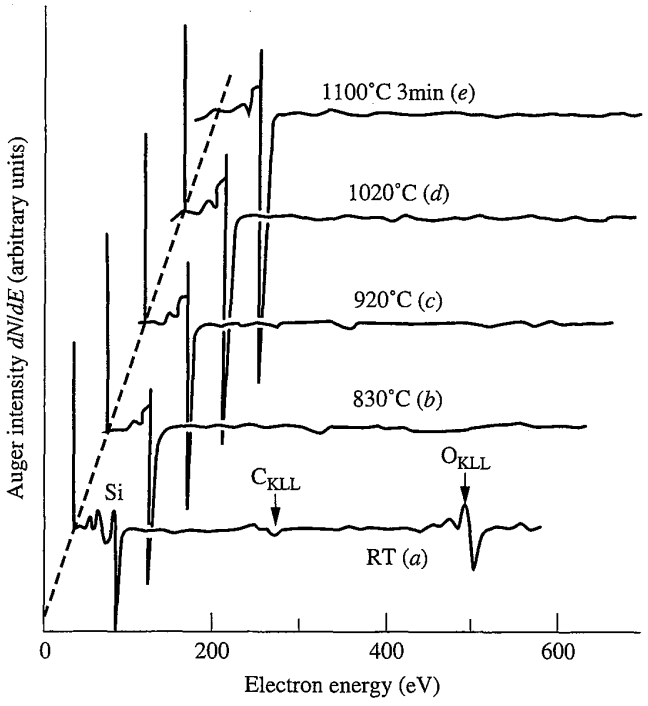


FIGURE 34
In-situ Auger spectra of a Si surface cleaned by the conventional high-temperature thermal etching method: (*a*) freshly oxidized, room temperature (RT), (*b*) after heating at 830°C, (*c*) 920°C, (*d*) 1020°C, and (*e*) 1100°C. (*After Ishizaka et al., Ref. 41.*)

Micron Ex. 1023, p. 117
Micron v. YMTC
IPR2024-00909

spectra of the Si surface cleaned by the conventional high-temperature etching method.[41] Line *a* is the surface after conventional chemical treatment; O_{KLL} (515 eV) and C_{KLL} (272 eV) Auger signals, which originate from the surface contamination, are observed. The O_{KLL} signal disappears as the temperature rises to 850°C, but the C_{KLL} peak still remains. Lander et al.[42] first reported the removal of oxide at temperatures between 800 and 1000°C in a high vacuum by

$$Si + SiO_2 \rightarrow 2SiO$$

The C_{KLL} signal disappears after the temperature rises to 1100°C. That is why the conventional epitaxy process needs such a high-temperature treatment. Hence, if a carbon-free, oxidized Si surface can be prepared, low-temperature thermal etching of an oxide-free surface is possible (< 800°C). The high-temperature treatment causes impurity diffusion and changes the designed impurities concentration profile, crystal defects, such as dislocation and stacking faults, high wearout of the substrate-heating elements, and low throughput. Therefore, a low-temperature surface cleaning of silicon has been proposed (Shiraki clean process).[41] It includes

1. Degreasing
2. HNO_3 boiling + DHF dip (2.5%)
3. NH_4OH–H_2O_2–H_2O(1 : 1 : 3) + DHF dip (2.5%)
4. HCl–H_2O_2–H_2O(3 : 1 : 1)
5. Rinsing + spin dry to let the surface become uniformly hydrophilic

Figure 35 shows the in situ Auger spectra after heating at 785°C for 15 min. The surface becomes oxide- and carbon-free. This method results in an atomically clean surface, and the epitaxy can be achieved at a low temperature of 785°C. There is

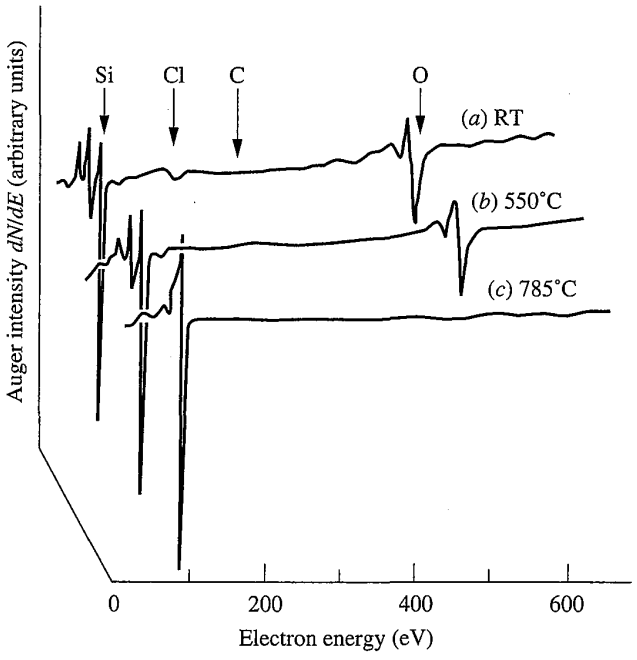


FIGURE 35
In-situ Auger spectra of Si surface cleaned by the low-temperature thermal etching method: (*a*) as freshly oxidized, (*b*) after heating at 550°C for 15 min, (*c*) after heating at 785°C for 15 min. (*After Ishizaka et al., Ref. 41.*)

Micron Ex. 1023, p. 118
Micron v. YMTC
IPR2024-00909

also another method to reduce carbon. A modified Shiraki-clean process has been proposed that omits the last HCl boiling and uses 4% HF as the last step to produce a hydrogen-terminated surface without native oxide. The growth temperature using this process can be reduced to as low as 650°C and maintains the same high quality as that of the Shiraki cleaning process.[43]

2.3 WET CLEANING TECHNOLOGY

Some optional processes steps have been added into the conventional RCA cleaning process.[44] A preliminary cleanup treatment with a hot $H_2SO_4 : H_2O_2$(2 : 1 to 3 : 1, at 120°C) mixture is used to remove the greasy contamination, which may be from the cassette or residues from the photoresist layers. After the SC-1 process, a 15-sec immersion in 1% HF–H_2O solution may be beneficial for removing any trace impurity. Some techniques have been proposed for advanced cleaning processes.[44]

Immersion technique: Vessels of fused silicon have been introduced to prevent leaching of Al, B, and alkalis, which can result if Pyrex glass is used. A DI water overflow-quenching system is also used for the automatic processing.

Megasonic cleaning: This cleaning process can remove particles and contamination with diluted SC-1 solution at a temperature of only 35 to 42°C.

Centrifugal spray cleaning: The chemical solutions, including SC-1, SC-2, and DI water, are pressure-fed directly onto the spinning wafer. This process can reduce the volume of fresh chemical and is faster than immersion. The spin-drying process requires antistatic protection to prevent static-induced particle deposition on the wafers.[45]

Closed-system chemical cleaning: In this process wafers are cleaned in a closed system during the entire cleaning, rinsing, and drying process to avoid contamination from the operator and environments.

Advanced wet cleaning processes. The steps of the advanced wet-chemical cleaning process proposed by Ohmi,[46] and the contaminants each removes, are

$H_2O + O_3$	Organic contamination
$NH_4OH + H_2O_2$	Particle, organic, and metallic impurities
$HF + H_2O_2$	Native oxide, metallic impurities
Ultrapure water	Rinsing

This process omits the conventional RCA SC-2 step when highly purified HF solution is used. In the clean process, most of the organic contamination is removed first in the ozoned, ultrapure water (low oxygen dissolved < 1 ppb) cleaning. Then NH_4OH–H_2O_2–H_2O = 0.05 : 1 : 5 is used to remove the particles, organic, and metallic impurities, maintaining an atomically flat surface. A HF (0.5%)–H_2O_2 (10%) cleaning is proposed as the final step to remove the native oxide and noble metals.

Micron Ex. 1023, p. 119
Micron v. YMTC
IPR2024-00909

TABLE 8
Concept of IMEC cleaning[47]

Step 1† Oxide growth	Step 2‡ Oxide removal	Step 3§ Optional	Drying¶
H_2SO_4/H_2O_2	Wet DHF	Chemical oxide	IPA vapor
H_2SO_4/O_3	HF vapor + rinsing		Hot DI water
H_2O/O_3			Spin dry
UV-O_3			N_2 blow

†Step 1: growth of about 1.5 nm of chemical oxide.
‡Step 2: removal of the oxide: drying procedure.
§Step 3 (optional): growth of chemical oxide.
¶Step 4: easier drying due to the growth of a hydrophilic Si surface.

Another new cleaning concept, IMEC (for Interuniversity Microelectronic Center) cleaning, is illustrated in Table 8. The basic concept of this new cleaning process is to separate the oxidizing and the etching action.[47] In step 1 the organic contamination is removed and a thin chemical oxide is grown; then this oxide is etched in step 2. By using this method, the total etch depth should be sufficient to remove the particles. The advantage of separating the cleaning into two steps is that the formation of the oxide layer is a self-limiting process of the chemical reaction. This layer is easy to control by oxide removal to obtain minimum surface roughness. Step 3, a chemical oxide growth, is optional. A cleaning process based on the IMEC cleaning concept is as follows:

H_2SO_4–H_2O_2 = 4 : 1	at 90°C for 10 min
HF(0.5%)/IPA(0.1%)	at room temperature for 2 min

It has been shown that the particle removal efficiency using the IMEC clean is better than SC-1 on the native oxide, and slightly weaker on the thermal oxide. IMEC cleaning is also quite superior to the other cleaning process in removing metallic contamination.[48] The yields of the gate oxide for four different clean methods are

IMEC + RCA(spray) + HF(bath)	80%
IMEC + RCA(spray)	65%
RCA(bath)	60%
IMEC(bath)	83%

The IMEC (bath) clean has the highest yield. The yield is defined as the percentage of gate oxide with a breakdown field larger than 12 MV/cm.

2.4 DRY CLEANING TECHNOLOGY

Wet cleaning technology still remains the major method for wafer cleaning. However, the ULSI industry has begun to recognize problems associated with wet cleaning.

Micron Ex. 1023, p. 120
Micron v. YMTC
IPR2024-00909

Among them are particle generation, drying difficulties, cost, chemical waste disposal, incompatibility with advanced concepts in integrated processing, and general inflexibility.[49] The dry cleaning process may provide a way to solve these problems, although currently it is not completely successful. The integrated dry cleaning processing system is shown in Fig. 36, in which the dry cleaning chamber is connected to the "cluster" processing system. The dry cleaning process, a kind of gas-phase chemistry, usually requires excitation energy to enhance the chemical reaction at low temperature. This added energy may be plasma, particle beam, short-wavelength radiation, or thermal heating, which enhances the surface cleaning result but must avoid the side effects of damage to the wafer.

Recently, many dry cleaning technologies have been developed, namely HF/H_2O vapor cleaning,[49–52] ultraviolet-ozone cleaning (UVOC),[53–56] H_2/Ar plasma cleaning,[29,57–59] and thermal cleaning.[37] These processes will be discussed in the following paragraphs. Usually, wafers receive a wet cleaning first. Although RCA clean[1,44] has been used for more than 30 years in semiconductor manufacturing because it is effective in removing organic, particulate, and metallic contaminations, the resulting Si wafer surface is covered with 4 to 6 nonuniform monolayers (ML) of silicon oxide resulting from liquid surface tension, bubbles, and generated particles.[49] The diluted HF dip cleaning (1 to 4%HF in H_2O) usually suffers some problems such as contamination by hydrocarbons on the surface.[56] It is desirable to have a hydrogen-passivated surface[3,38] for the subsequent processing steps, such as poly-Si or epitaxial silicon growth. However, in dry cleaning, after the HF/H_2O vapor process, the surface is fluorine (F)-terminated.[49–50]

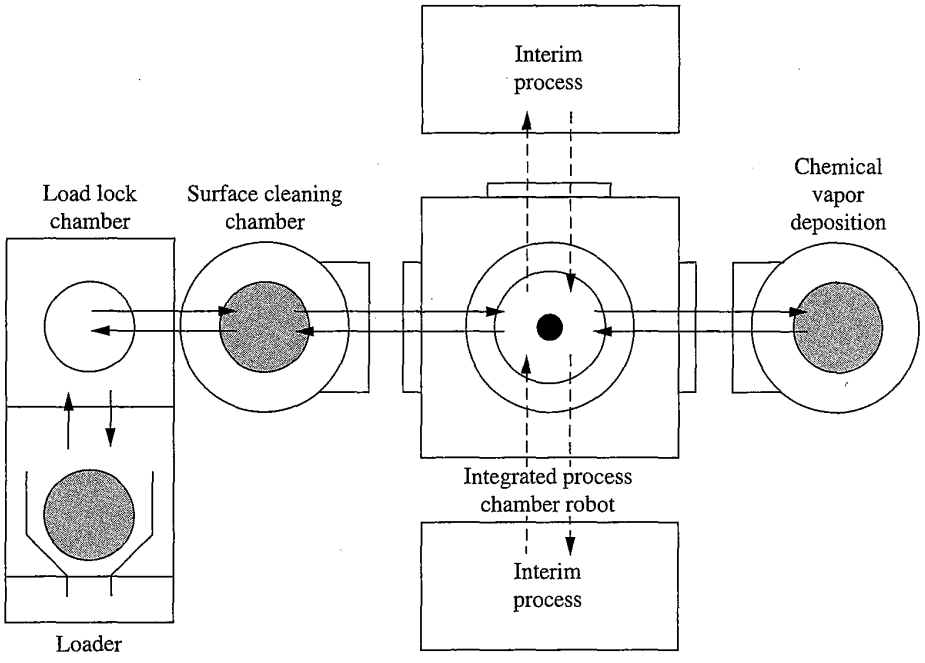


FIGURE 36
An integrated processing system concept incorporating a dry cleaning chamber.

Micron Ex. 1023, p. 121
Micron v. YMTC
IPR2024-00909

Ultraviolet-ozone clean (UVOC)[53–56]

This is a very effective way to remove hydrocarbons, although the surface is oxide-passivated. The mechanism of UVOC can be explained by the following process equations.

$$\text{Absorbed impurity} + h\nu \xrightarrow{2000\text{–}3000\ \text{Å UV}} \text{Excited impurity} \qquad (2.1)$$

In addition, oxygen atoms are excited by UV light to form atomic oxygen. This is a surface excitation process.

$$O_2 + h\nu \xrightarrow{1849\ \text{Å UV}} 2O \qquad (2.2)$$

$$O + O_2 \rightarrow O_3 \qquad (2.3)$$

$$O_3 + h\nu \xrightarrow{2537\ \text{Å UV}} O + O_2 \qquad (2.4)$$

$$\text{Then} \quad \text{Excited impurity} + (O, O_3) \rightarrow \text{volatile compound} \qquad (2.5)$$

The reaction in Eq. (2.1) is a surface-excitation process, whereas the reactions in Eqs. (2.2) through (2.5) are gas-phase excitation processes that convert impurities, such as hydrocarbons, into volatile gases.

Figure 37 shows the XPS spectra of a HF-dipped surface and a UV/O_3 exposed surface, which indicates the formation of surface oxide.[29] The high resolution of surface energy loss spectroscopy (HREELS) is used to examine the decomposition of hydrocarbons. As shown in Fig. 38, after a 300 sec exposure in UV/O_2, peaks of δ(CH) and ν(CH) disappear, while ν(SiO) and δ(SiO) peaks are intensified. The

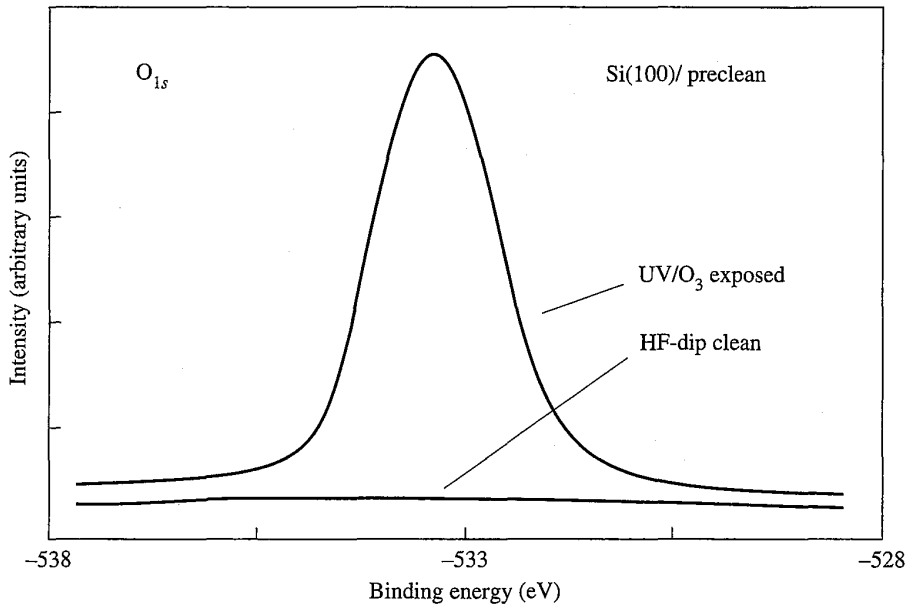


FIGURE 37
X-ray photoemission spectra displaying the O_{1s} regions of Si(100) wafers after 15 min UV-ozone and after 10 sec 10% HF dip clean. (*After Offenberg et al., Ref. 29.*)

Micron Ex. 1023, p. 122
Micron v. YMTC
IPR2024-00909

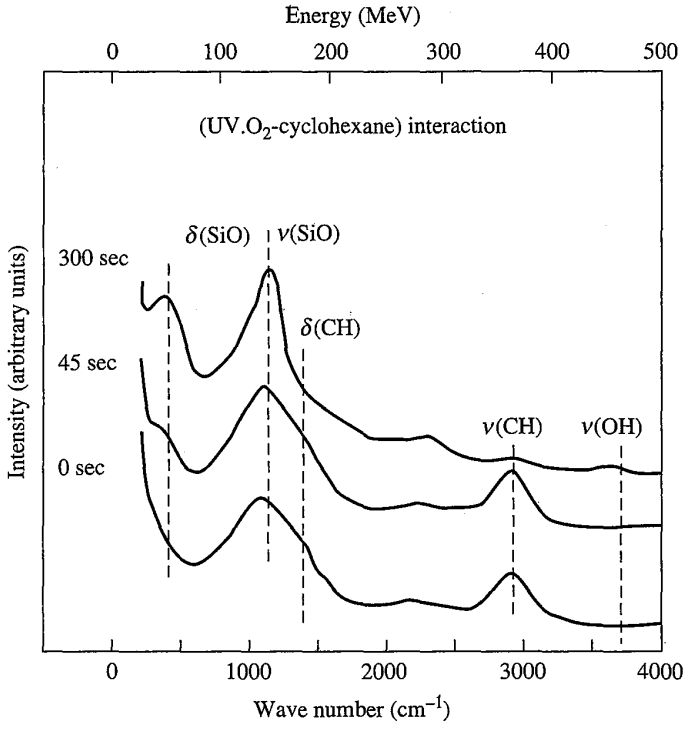


FIGURE 38
HREELS spectra for a Si(100) wafer that was HF-dip cleaned, subsequently contaminated with cyclohexane, then cleaned with UV/O_2 for 45 sec, and for 300 sec. (*After Kasi and Liehr, Ref. 50.*)

subsequent HF/H_2O vapor or Ar/H_2 plasma cleaning processes in vacuum would be helpful in removing the surface oxide.

HF/H_2O vapor clean

The diluted HF dip promotes a hydrogen-passivated surface, but the HF/H_2O vapor clean induces a fluorine-terminated surface.[51,52] Table 9 indicates that "clean 1" ($2H_2SO_4 + 1H_2O_2$) results in a 5.8-monolayer (ML) oxide on the surface, "clean 2" (HF dip, then H_2O rinse) results in an OH termination with 0.9 ML oxide cover on the surface, and "clean 3" (1% HF in H_2O) results in a hydrogen-terminated surface with only 0.3 ML oxide. The contact angle measured is 19° for clean 1, 138° for clean 2, and 155° for clean 3. This indicates a tendency to change from a hydrophilic surface to a hydrophobic surface. When wafers receive a HF/H_2O vapor clean, the content of water vapor varies with the composition ratios. Although the surfaces are F-terminated, the [C], [O], and [F] concentrations are different with different process conditions. Contact angles decrease from 140° for 0.4 ML of [O] to 20° for 5.9 ML of [O], indicating surface-state transformations.[51,52] Figure 39 indicates the concentration of fluorine, C[F], versus the concentration of oxygen C[O], where $F_1 = C$[Si–F bond], $F_2 = C$[O–Si–F bond]. Clean 3, with a 155° contact angle, is slightly more hydrophobic than process 1, with 140° of contact angle.[51]

Micron Ex. 1023, p. 123
Micron v. YMTC
IPR2024-00909

TABLE 9
XPS surface residues and contact angle of Si wafers[51, 52]

	Termination	[C], ML	[O], ML	[F], ML	Contact angle (°)
Wet clean					
$2H_2SO_4 + 1H_2O_2$ (clean 1)	O	1.4 ML	5.8		19
HF, then H_2O (clean 2)	OH	1.6	0.9		138
HF/H_2O last (1:100) (clean 3)	H	0.9	0.3	0.08	155
HF/H_2O vapor clean					
Process 1	F	0.4	0.4	0.2	140
Process 2, water-rich	F	0.37	0.7	0.48	60
Process 3	F	0.5	1.4	0.4	—
Process 4	F	0.7	5.9	0.7 ML	20

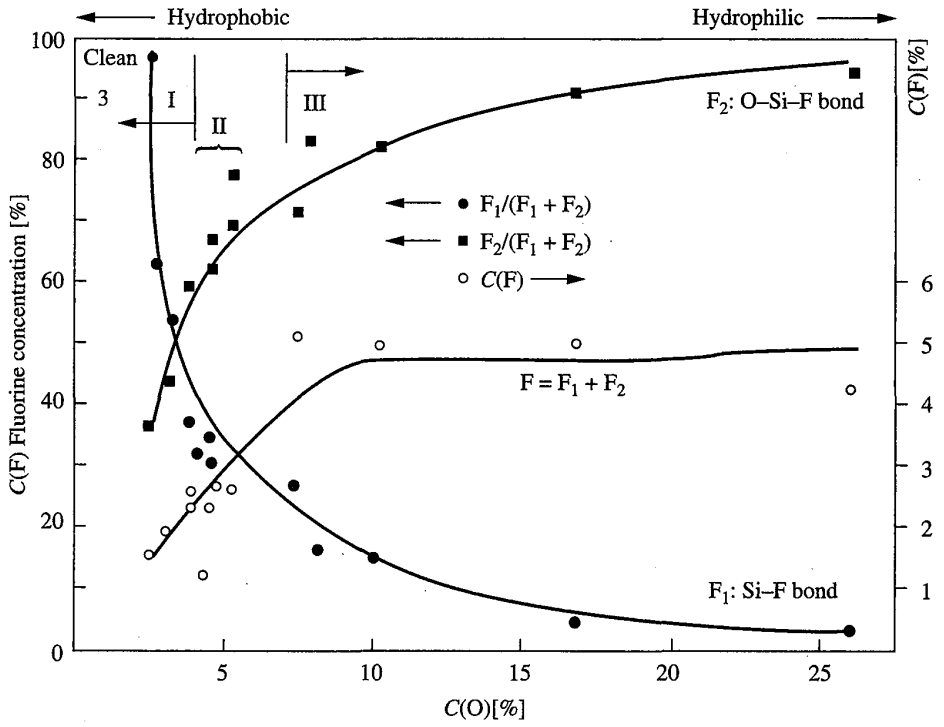


FIGURE 39
Fluorine concentration versus oxygen concentration and respective F–Si and O–Si–F concentrations versus oxygen concentration *C*[O]. (*After Ermolieff et al., Ref. 51.*)

Micron Ex. 1023, p. 124
Micron v. YMTC
IPR2024-00909

AR/H_2 plasma cleaning

Usually a remote plasma is used to reduce the bombardment damage to the wafer. Gases, Ar or H_2 or both, are passed through an rf (13.6 MHz) inductive-coil or capacitive-coupled region where a plasma region is formed. At a pressure of 1 torr, the Ar or H_2 gases molecules are excited or ionized and generate a plasma. These gaseous molecules are carried to another region of the wafer, where a cleaning process begins. Excited Ar ions physically sputter the surface impurities away, while the excited H_2 ions chemically etch the surface. By proper adjustment of the physical/chemical etching (cleaning) ratio, an optimum cleaning condition can be obtained that produces minimum damage of the surface.

Instead of RF plasma excitation, electron cyclotron resonance (ECR) can do the same job at a higher plasma density and at a lower kinetic energy of the excited atom or molecules (~100 eV). ECR occurs at a frequency of $f_c = eB/2\pi m^*$. The f_c is 2.45 GHz, for a B of 875 gauss. The pressure is about 1 mtorr. The detailed construction of ECR will be discussed in Chapter 7. Native oxide and hydrocarbons are removed by H_2-plasma cleaning or by ECR Ar/H_2 plasma[60] at room temperature or at a temperature below 250°C. Figure 40 shows the effective carbon and oxygen removal by a remote hydrogen plasma[58] at several temperatures up to 305°C, monitored by Auger spectroscopy.

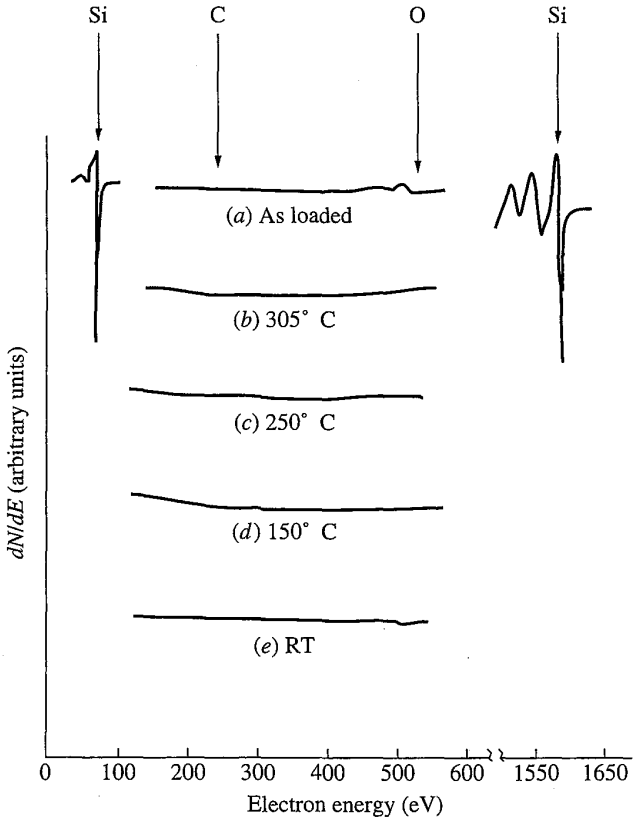


FIGURE 40
AES analysis of a Si(100) surface after (*a*) a 40 : 1 H_2O : HF (49%) dip and remote hydrogen plasma clean for 45 min at (*b*) 305°C, (*c*) 250°C, (*d*) 150°C, and (*e*) room temperature (RT). (*After Hsu et al., Ref. 58.*)

Micron Ex. 1023, p. 125
Micron v. YMTC
IPR2024-00909

Thermal cleaning

The native oxide can be removed by heating the wafer to 800°C or above in ultrahigh vacuum ($< 10^{-10}$ torr) to vaporize the oxide. The native oxide is SiO_x, in which x depends on the previous cleaning process. In the Shiraki cleaning process,[41] an 850°C, 10 min thermal clean is necessary, whereas in the HF dip process (4% HF) a prebake of 200°C, or no bake process, is used for a 650°C epitaxial growth.

In all these processes, high-temperature cleaning should be carefully examined because at high temperature (> 800°C), the following reaction process occurs at low oxygen partial pressure,[52]

$$Si + SiO_2 \rightarrow 2SiO \tag{2.6}$$

The SiO is volatile at temperature above 750°C. When the SiO_2 film is removed, the silicon wafer starts to etch:

$$Si + O_2 \rightarrow 2SiO \tag{2.7}$$

produces aggravated etching, which induces microroughness on the surface. The deposited gate oxide, therefore, has a low breakdown voltage (~1 MV/cm) due to microroughness. The low oxygen partial-pressure oxidation kinetics can be explained by the work of Smith and Ghidini,[61] as shown in Fig. 41. At low temperature and

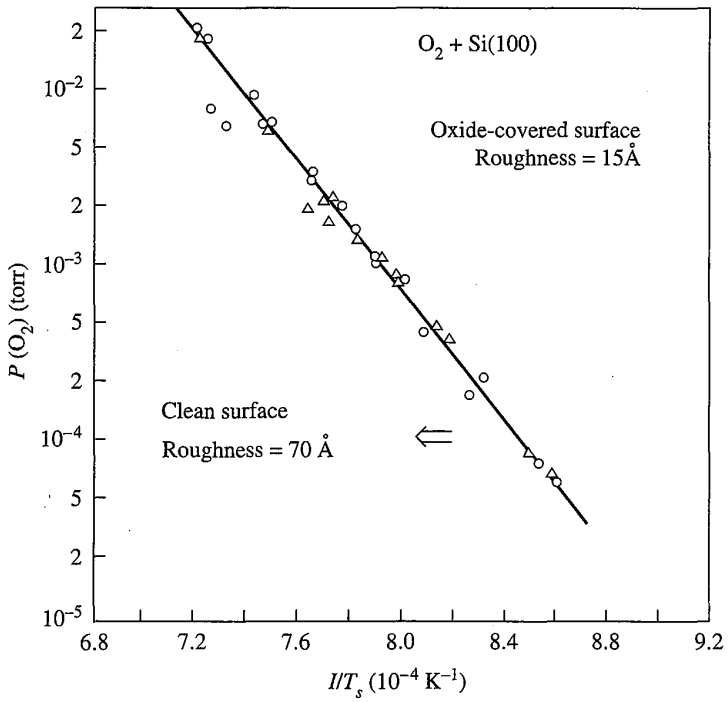


FIGURE 41
Boundary separating oxide-covered surface (upper right) from clean surfaces (lower left) for Si(100) exposed to a partial pressure of O_2 at elevated temperature. (*After Smith and Ghidini, Ref. 61.*)

Micron Ex. 1023, p. 126
Micron v. YMTC
IPR2024-00909

TABLE 10
Comparison of wet and dry cleaning

Method	Surface termination	Main effects	Remarks	References
Wet				
RCA	4~6 ML SiO_2	Removal of organic, metallic, and other particulate residues	Nonuniformity due to bubble, particle, surface tension, etc.	1, 44
HF-last	H^-	Oxide removal		3, 40
Dry				
HF + H_2O (vapor)	F^-	Oxide removal		49–52
UVOC	OH^-, SiO^-	CH removal		53–56
H_2/Ar plasma	H^-	H_2 chemical clean Ar physical clean	1 torr	57–59
Remote plasma	H^-	C/O removal	1 torr	
ECR	H^-	C/O removal	1 mtorr	
Thermal clean		Microroughness	> 800°C, UHV	52

a high O_2 partial pressure a smooth surface with a roughness of 1.5 nm is obtained, whereas at high temperature and a low O_2 partial pressure a roughness of 7 nm indicates that a thermal etching process Eq.(2.6) is taking place.

Table 10 summarizes the effects of various cleaning processes. In surface termination, the RCA clean results in 4 to 6 monolayers of SiO_2 while removing organic and metallic residues on the surface. The HF-last cleaning process forms a hydrogen-passivated surface and removes the oxide on the surface. These processes suffer the result of incomplete cleaning due to forming gas bubbles, surface tension, and remaining particles. The dry cleaning processes summarized in Table 10 have therefore been developed; they produce different surface terminations and impurities removals. The HF + H_2O vapor cleaning forms a F-terminated surface. The UVOC results in an OH^-, SiO^- covered surface; it is an effective way to remove hydrocarbons. The H_2/Ar plasma processes result in a hydrogen-passivated surface; however, the surface residues are removed by H_2 chemical etch and by Ar physical sputtering processes. Good control of H_2/Ar ratio will give a good result of "soft" cleaning with minimum damage. Remote plasma and ECR plasma cleaning processes are effective in carbon and oxide removal. Finally, a thermal clean is presented that removes oxide, but sometimes a microrough surface is obtained if oxygen remains in the vacuum chamber.

2.5 SUMMARY AND FUTURE TREND

Removal of silicon surface contaminations such as heavy metals, organics, and particles can be achieved by wet cleaning with a modified RCA cleaning process. Recent studies of surface-state physics have increased the efficiency of cleaning methods. An oxygen-free DI water rinse promotes hydrogen surface termination by eliminat-

Micron Ex. 1023, p. 127
Micron v. YMTC
IPR2024-00909

ing the Si–Si bond breakage resulting from oxygen atoms, thus making the surface more hydrophobic. IPA added to the HF solution enhances particle removal. An O_3-containing DI water rinse helps to remove the organic and hydrocarbon contaminants. Advanced wet cleaning processes have been proposed for the more sophisticated ULSI processes. Microroughness is crucial to ULSI device quality, and a detailed mechanism and method of obtaining a good microsmooth surface has been proposed.

To meet future requirements for an integrated processing system, dry etching processes have been introduced. A summary of these technologies were given, which should help readers make a judgment about using these processes in their individual process applications.

The present state of these technologies still does not meet all the future requirements for more advanced ULSI devices. More in-depth understanding of the surface-state physics and chemistry is required, not only to eliminate the surface contaminations but also to provide the atomically flat surface and interface required for future semiconductor manufacturing. In addition, the development of more efficient dry etching and wet cleaning processes is essential for future integrated processing systems.

REFERENCES

1. W. Kern and D. A. Puotinen, "Cleaning Solutions Based on Hydrogen Peroxide for Use in Silicon Semiconductor Technology," *RCA Rev.* **31,** 187 (1970).
2. S. Verhaverbeke, *Dielectric Breakdown in Thermally Grown Oxide Layers,* Katholieke Universiteit, Leuven, 1993.
3. C. Werkhoven, E. Granneman, R. De Blank, S. Verhaverbeke, P. Mertens, M. Meuris, W. Vandervorst, and M. Heyns, "Wet and Dry HF-Last Cleaning Process for High-Integrity Gate Oxides," *Tech. Dig., IEDM,* 633 (1992).
4. M. Liehr, "Science Issues Related to Wafer Cleaning in Silicon Technology," *Mat. Res. Soc., Symp. Pro.,* **259,** 3 (1992).
5. P. H. Singer, "Trends in Wafer Cleaning," *Semic. Intl.* 36 (Dec., 1992).
6. M. Itano, F. W. Kern, Jr., R. Reed, W. Rosenberg, M. Miyashita, I. Kawanabe, and T. Ohmi, "Particle Deposition and Removal in Wet Cleaning Processes for ULSI Manufacturing," *IEEE Trans. Semic. Manuf.* **5,** 114 (1992).
7. S. Shwartzman, A. Mayer, and W. Kern, "Megasonic Particle Removal from Solid-State Wafers," *RCA Rev.* **46,** 81 (1985).
8. S. Verhaverbeke, C. Werkhoven, M. Meuris, H. F. Schmidt, K. Dillenbeck, P. Mertens, M. Heyns, and A. Philipossian, "Investigation of Sources of Metallic Contamination by TXRF," *Institute of Environmental Science Annual Technical Meeting, Las Vegas, Nevada,* 1993, p.423.
9. W. B. Henley, L. Jastrzebski, and N. F. Haddad, "Monitoring Iron Contamination in Silicon by Surface Photovoltage and Correlation to Gate Oxide Integrity," *Mat. Res. Soc., Symp. Proc.,* **315,** 299 (1993).
10. H. Park, O. R. Helms, D. Ko, M. Tran, and B. B. Triplett, "Effect of Surface Iron on Gate Oxide Integrity and Its Removal From Silicon Surfaces," *Mat. Res. Soc., Symp. Proc.,* **315,** 353 (1993).

Micron Ex. 1023, p. 128
Micron v. YMTC
IPR2024-00909

11. L. Jastrzebski, O. Milic, M. Dexter, J. Lagowski, D. Debusk, K. Nauka, R. Witowski, M. Gordon, and E. Persson, "Monitoring of Heavy Metal Contamination During Chemical Cleaning with Surface Photovoltage," *J. Electrochem. Soc.,* **140,** 1152 (1993).
12. N. Anzai, Y. Kureishi, S. Shimizu, and T. Nitta, *Proceeding of the 1st Workshop on ULSI Ultra Clean Technology,* 1989, p.75.
13. T. Ohmi, T. Imaoka, I. Sugiyama, and T. Kezaka, "Metallic Impurities Segregation at the Interface between Si Wafer and Liquid during Wet Cleaning," *J. Electrochem. Soc.,* **139,** 3317 (1992).
14. T. Ohmi, T. Imaoka, T. Kezuka, J. Takano, and M. Kogure, "Segregation and Removal of Metallic Impurity at Interface of Silicon and Fluorine Etchant," *J. Electrochem. Soc.,* **140,** 811 (1993).
15. S. Verhaverbeke, P. W. Mertens, M. Meuris, M. M. Heyns, A. Schnegg, and A. Philipossian, "The Influence of Metallic Impurities on the Dielectric Breakdown of Oxide and New Processes to Avoid Them," Technical Conference SEMICON, Europe, 1992.
16. S. Verhaverbeke, M. Meuris, P. W. Mertens, A. Kelleher, M. M. Heyns, R. F. Keersmaecker, M. Murrell, and C. J. Sofield, "The Effect of Metallic Contamination on Void Formation, Dielectric Breakdown and Hole Trapping in Thermal SiO_2 Layers," *Proceeding, in Cleaning Technology in Semiconductor Device Manufacturing,* Electrochemical Society, 1991, p.187.
17. Ref. 2, p. 76.
18. Ref. 2, p. 178.
19. K. Hashimoto, K. Egashira, M. Suzuki, D. Matsunaga, "Gate Oxide Deterioration Caused by Organic Contamination onto the Oxide," *1991 International Conference on Solid-State Devices and Materials, Yokohama,* 1991, p.143.
20. M. Kogure, T. Futatsuki, J. Takano, T. Isagawa, K. Kimura, Y. Ogato, F. Tanaka, and T. Ohmi, "Influence of Organic Impurities in Hydrogen Peroxide for Advanced Wet Chemical Processing," *J. Electrochem. Soc.,* **140,** 3321 (1993).
21. T. Ohmi, T. Isagawa, M. Kogure, and T. Imaoka, "Native Oxide Growth and Organic Impurity Removal on Si Surface with Ozone-Injected Ultrapure Water," *J. Electrochem. Soc.,* **140,** 805 (1993).
22. R. A. Governal, A. Bonner, and F. Shadman, "Effect of Component Interactions on the Removal of Organic Impurities in Ultra Pure Water Systems," *IEEE Trans. Semic. Manuf.,* **4,** 298 (1991).
23. Ref. 2, p. 112.
24. T. Ohmi, T. Tsuga, J. Takano, M. Kogure, K. Makihara, and T. Imaoka, "Influence of Vacancy in Silicon Wafer of Various Types on Surface Microroughness in Wet Chemical Process," *IEICE Trans. Electron.,* **7,** 800 (1992).
25. M. Miyashita, T. Tsuga, K. Makihara, and T. Ohmi, "Dependence of Surface Microroughness of CZ, FZ, and Epi Wafers on Wet Chemical Processing," *J. Electrochem. Soc.,* **139,** 526 (1992).
26. T. Ohmi, M. Miyashita, M. Itano, T. Imaoka, and I. Kawanabe, "Dependence of Thin Oxide Films Quality on Surface Microroughness," *IEEE Trans. Electron. Dev.,* **39,** 537 (1992).
27. M. Hirose, T. Yasaka, M. Takakura, S. Miyazaki, "Initial Oxidation of Chemically Cleaned Silicon Surfaces," *Sol. State Technol.,* 43 (Dec. 1991).
28. J. M. de Larios, D. B. Kao, C. R. Helms, B. E. Deal, "Effect of SiO_2 Surface Chemistry on the Oxidation of Silicon," *Appl. Phys. Lett.,* **54,** 715 (1989).
29. M. Offenberg, M. Liehr, and G. W. Rubloff, "Surface Etching and Roughening in Integrated Processing of Thermal Oxides," *J. Vac. Sci. Technol. A,* **9,** 1058 (1991).

Micron Ex. 1023, p. 129
Micron v. YMTC
IPR2024-00909

30. M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami, and M. Ohwada, "Growth of Native Oxide on a Silicon Surface," *J. Appl. Phys.* **68,** 1272 (1990).
31. Y. Yagi, T. Imaoka, Y. Kasama, and T. Ohmi, "Advanced Ultrapure Water Systems with Low Dissolved Oxygen for Native Oxide Free Wafer Processing," *IEEE Trans. Semic. Manuf.,* **5,** 121 (1992).
32. M. Miyawaki, S. Yoshitaka, and T. Ohmi, "Improvement of Al-Si contact Performance in Native Oxide Free Processing," *IEEE Electron Dev. Lett.,* **11,** 448 (1990).
33. T. Ohmi, M. Morita, A. Teramoto, K. Makihara, and K. S. Tseng, "Very Thin Oxide Film on a Silicon Surface by Ultraclean Oxidation," *Appl. Phys. Lett.,* **60,** 2126 (1992).
34. K. Makihara, A. Teramoto, K. Nakamura, M. Y. Kwon, M. Morita, and T. Ohmi, "Preoxide-Controlled Oxidation for Very Thin Oxide Films," *Jpn. J. Appl. Phys.,* **32,** 294 (1993).
35. S. M. Sze, Ed., *High-Speed Semiconductor Devices,* John Wiley and Sons, Inc., New York, 1990.
36. S. Verhaverbeke, M. Meuris, M. Schaekers, L. Haspeslagh, P. Mertens, M. M. Heyns, R. De Blank, and A. Philipossian, "A New Modified HF-Last Cleaning Process for High-Performance Gate Dielectrics," 1992 Symposium on VLSI Technology, Washington, June 22, 1992.
37. T. Yamazaki, N. Miyata, T. Aoyama, and T. Ito, "Investigation of Thermal Removal of Native Oxide from Si(100) Surface in Hydrogen for Low-Temperature Si CVD Epitaxy," *J. Electrochem. Soc.,* **139,** 1175 (1992).
38. T. C. Chang, C. Y. Chang, T. G. Jung, and W. C. Tsai, "Nanometer Thick Si/SiGe Strained-Layer Superlattice Grown by an UHVCVD Technique," *J. Appl. Phys.,* **75,** 3441 (1994).
39. M. Grundner, D. Graf, P.O. Hahn, and A. Schnegg, "Wet Chemical Treatments of Si Surface: Chemical Composition and Morphology," *Sol. State Technol.,* 69 (Feb. 1991).
40. J. L. Alay, S. Verhaverbeke, W. Vandervorst, and M. Heyns, "Critical Parameters for Obtaining Low Particle Densities in an HF-Last Process," *1992 International Conference on Solid-State Devices and Materials, Yokohama,* 1992, p.123.
41. A. Ishizaka and Y. Shiraki, "Low Temperature Surface Cleaning of Silicon and Its Application to Silicon MBE," *J. Electrochem. Soc.,* **133,** 666 (1986).
42. T. J. Lander and J. Morrison, "Low Voltage Electron Diffraction Study of Oxidation and Reduction of Si," *J. Appl. Phys.,* **33,** 2089 (1962).
43. P. E. Thompson, M. E. Twigg, D. J. Godbey, and K. D. Hobart, "Low Temperature Cleaning Processes for Si Molecular Beam Epitaxy," *J. Vac. Sci. Technol. B,* **11(3),** 1077 (1993).
44. W. Kern, "The Evolution of Silicon Wafer Cleaning Technology," *J. Electrochem. Soc.,* **137,** 1887 (1990).
45. H. Inada, S. Sakata, T. Yoshida, T. Okado, and T. Ohmi, "Antistatic Protection in Wafer Drying Process by Spin-Drying," *IEEE Trans. on Semic. Manuf.,* **5,** 541 (1992).
46. T. Ohmi, "Advanced Wet Chemical Cleaning for Future ULSI Fabrication," *Electrochem. Soc., 184th Meeting, New Orleans,* 1993, p.495.
47. M. M. Heyns, S. Verhaverbeke, M. Meuris, P. W. Mertens, H. Schmidt, M. Kubota, A. Philipossian, K. Dillenbeck, D. Graf, A. Schnegg, and R. De Blank, "New Wet Cleaning Strategies for Obtaining Highly Reliable Thin Oxides," *Mat. Res. Soc., Symp. Proc.,* **315,** 35 (1993).
48. S. Verhaverbeke, M. Meuris, P. Mertens, H. Schmidt, M. M. Heyns, A. Philipossian, D. Graf, and K. Dillenbeck, "Advanced Wet Cleaning Technology for High Reliable Thin Oxides," *Proceedings of the 4th International Symposium on ULSI Science and Technology,* Electrochemical Society, 1993, p.199.

Micron Ex. 1023, p. 130
Micron v. YMTC
IPR2024-00909

49. B. E. Deal, M. A. McNeilly, D. B. Kao, J. M. de Larios, "Vapor Phase Wafer Cleaning Processing for the 1990s," *Sol. State Technol.,* 73 (July 1990).
50. S. R. Kasi and M. Liehr, "Vapor Phase Hydrocarbon Removal for Si Processing," *Appl. Phys. Lett.,* **57,** 2095 (1990).
51. A. Ermolieff, F. Martin, A. Amouroux, S. Marthon, and J. F. M. Westendorp, "Surface Composition Analysis of HF-Vapor-Cleaned Silicon by X-ray Photoelectron Spectroscopy," *Semic. Sci. Technol.,* **6,** 98 (1991).
52. X. Xu, R. T. Kuehn, M. C. Ozturk, J. J. Wortman, R. J. Nemanich, G. S. Harris, and D. M. Maher, "Influence of Dry and Wet Cleaning on the Properties of Rapid Thermal Grown and Deposited Gate Dielectrics," *J. Electron. Matls.,* **22,** 335 (1993).
53. J. R. Vig, "UV/Ozone Cleaning of Surfaces," *J. Vac. Sci. Technol. A,* **3,** 1027 (1985).
54. J. Ruzyllo, G. T. Duranko, and A. M. Hoff, "Preoxidation UV Treatment of Silicon Wafers," *J. Electrochem. Soc.,* **134,** 2052 (1987).
55. M. Suemitsu, T. Kaneko, and N. Miyamoto, "Low Temperature Silicon Surface Cleaning by HF Etching/Ultraviolet Ozone Cleaning (HF/UVOC) Method (I)" *Jpn. J. Appl. Phys.,* **12,** 2421 (1989).
56. T. Kaneko, M. Suemitsu, and N. Miyamoto, "Low Temperature Silicon Surface Cleaning by HF Etching/Ultraviolet Ozone Cleaning Method (II): In Situ UVOC," *Jpn. J. Appl. Phys.,* **12,** 2425 (1989).
57. T. R. Yew and R. Reif, "Low-Temperature In Situ Surface Cleaning of Oxide Patterned Wafers by Ar/H_2 Plasma Sputter," *J. Appl. Phys.,* **68,** 4681 (1990).
58. T. Hsu, B. Anthony, R. Qian, J. Irby, S. Banerjee, A. Tasch, S. Lin, H. Marcus, and C. Magee, "Cleaning and Passivation of the Si (100) Surface by Low Temperature Remote Hydrogen Plasma Treatment for Si Epitaxy," *J. Electron. Matls.,* **20,** 279 (1991).
59. J. Cho, T. P. Schneider, J. Vander Weide, H. Jeon, and R. J. Nemanich, "Surface Electronic States of Low-Temperature H-Plasma Cleaned Si (100)," *Appl. Phys. Lett.,* **59,** 1995 (1991).
60. S. J. Pearton, F. Ren, A. Katz, U. K. Chakrabarti, E. Lane, W. S. Hobson, R. F. Kopf, C. R. Abernathy, C. S. Wu, D. A. Bohling, and J. C. Ivankovits, "Dry Surface Cleaning of Plasma-Etched High Electron Mobility Transistor," *J. Vac. Sci. Technol. B,* **11,** 546 (1993).
61. F. W. Smith and G. Ghidini, "Reaction of Oxygen with Si(111) and (100)-Critical Conditions for the Growth of SiO_2," *J. Electrochem. Soc.,* **129,** 1300 (1982).

PROBLEMS

1. What is needed for the complete achievement of an ultraclean wafer surface?

2. What are the advantages of the HF-last step?

3. The RCA SC-1 process results in a rough silicon surface. What can be done to reduce the roughness of the silicon surface?

4. What are the advantages of the dry cleaning processes?

Micron Ex. 1023, p. 131
Micron v. YMTC
IPR2024-00909

CHAPTER 3

Epitaxy

Pei-Jih Wang

3.1 INTRODUCTION

Epitaxy is one of the most fundamental processes for semiconductor device fabrication. It is the foundation on which most devices are built. The term *epitaxy* means "upon-ordered" in Greek. In reality, epitaxy is simply a process to grow a single-crystal layer on a single-crystal substrate. When the single-crystal layer and the single-crystal substrate are of exactly the same material, it is called *homoepitaxy;* if the grown single-crystal layer and its single-crystal substrate are different in any aspects, the process is called *heteroepitaxy.* Strictly speaking, almost all epitaxial processes are heteroepitaxial, because there are always some differences between the layer and the substrate, even if they are of the same material chemically. Such microscopic differences may result from variations in doping or defect density. However, it is widely accepted that homoepitaxy takes place when the two materials have the same crystalline and chemical structures, whereas heteroepitaxy involves different materials.

In principle, it is much easier to achieve high-quality epitaxial layers with homoepitaxy, because there are no problems associated with lattice mismatch. In contrast, it is much more difficult to grow a high-quality layer with heteroepitaxy, because two dissimilar crystals of different lattice constants or orientations are thermodynamically unfavorable to combine without generating defects, such as dislocations. This is not to say that perfect heteroepitaxial layers are not obtainable. Most of the theories, together with experimental results, agree that the strain due to lattice mismatch can be accommodated elastically if the layer is thinner than a certain critical thickness, h_c.[1] Figure 1 shows schematic drawings of a typical unstrained and strained heteroepitaxy. It has also been reported by many researchers that the value of h_c is inversely proportional to the misfit strain between two

Micron Ex. 1023, p. 132
Micron v. YMTC
IPR2024-00909

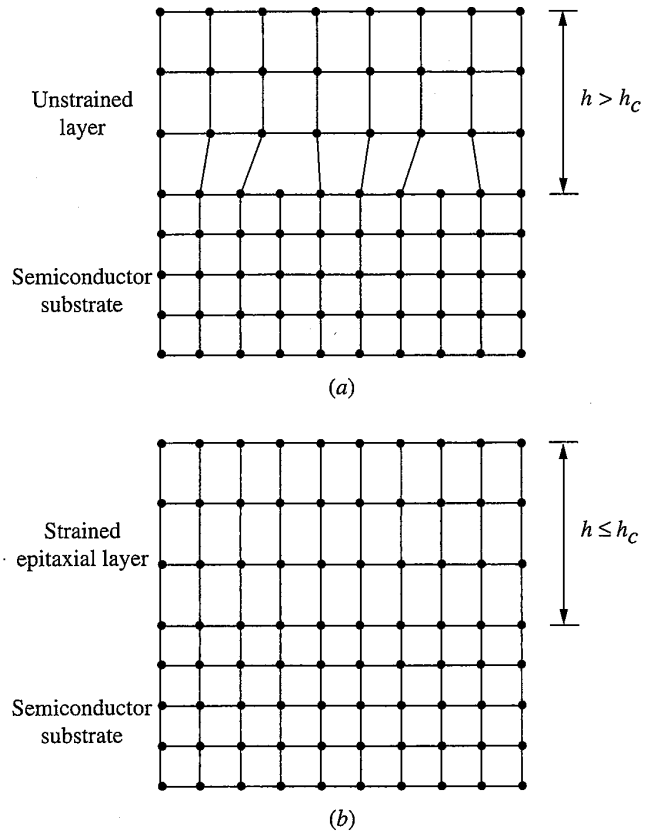


FIGURE 1
Schematic drawings of typical (*a*) unstrained and (*b*) strained heteroepitaxy.

materials and is greatly increased with reduced growth temperatures. At low growth temperatures the epitaxial growth process is far from equilibrium and is dominated by a kinetic reaction, resulting in a larger h_c.[2]

The term *epitaxy* is applied to all kinds of thin-film depositions as long as they are "arranged on in order." This process has been widely employed to prepare III–V and II–VI compound semiconductor materials, such as GaAs, InP, AlGaAs, InGaAsP, CdSe, and HgCdTe, and devices, such as the long-wavelength infrared detector, light-emitting diode (LED), semiconductor laser, heterojunction bipolar transistor (HBT), and modulation-doped field-effect transistor (MODFET). However, in this chapter we will focus only on silicon epitaxy. For readers who are interested in pursuing work in the field of III–V or II–VI compound semiconductors, please refer to other articles.[3]

Silicon epitaxy provides flexibility for a device designer to tailor or optimize the device performance. Unlike the diffusion or ion implantation process, epitaxy is a much easier method for controlling doping layer thickness, doping concentration, and doping profile. In addition, it can produce a buried layer structure or a lightly doped epitaxial layer on top of a heavily doped substrate, a combination that is desirable for tuning device performance. For example, a bipolar device built directly on a heavily doped substrate has a low collector resistance but suffers from a low collector/substrate junction breakdown voltage. On the other hand, if a lightly doped epitaxial layer is deposited on top of a heavily doped substrate, the collector/substrate

Micron Ex. 1023, p. 133
Micron v. YMTC
IPR2024-00909

junction breakdown voltage of a bipolar device can be greatly improved without sacrificing the advantage of low collector resistance, which is essential for high-speed operation. An epitaxial wafer with a lightly doped layer on top of a heavily doped substrate has also been employed for advanced complementary metal-oxide-silicon (CMOS) devices to minimize latch-up effects.[4]

An epitaxial process involves many complex physical and chemical steps. In this chapter, we will first introduce the fundamental aspects of epitaxy, including basic growth kinetics and various growth models, in Section 3.2. Once the background concepts are established, conventional Si epitaxy is introduced in Section 3.3. The term *conventional* is used to differentiate it from other novel approaches such as low-temperature epitaxy, described in Section 3.4, and selective Si epitaxy, described in Section 3.5. These novel growth techniques have received a great deal of attention recently, because of the demands for a low thermal budget and much more miniaturized devices for ULSI technology. Device performance is ultimately influenced by the quality of an epitaxial layer; in Section 3.6 we will introduce various characterization techniques that are commonly used to evaluate the quality of an epitaxial layer. Finally, the trends and future developments in the field of Si epitaxy will be summarized in Section 3.7.

3.2 FUNDAMENTAL ASPECTS OF EPITAXY

Epitaxial growth can be achieved from solid-phase, liquid-phase, vapor-phase, and molecular-beam deposition. For a Si epi-layer, vapor-phase epitaxy (VPE) has received the widest acceptance of all these techniques. Therefore, in this section, we will focus on the basics of the VPE process. However, some of the models and mechanisms, which are also applicable to other epitaxial growth processes, will be discussed as well.

A typical VPE process is illustrated schematically in Fig. 2. Prior to the layer deposition, the growth system is purged by nitrogen or hydrogen for a short period, and followed by a vapor HCl etching. The deposition process is then initiated by

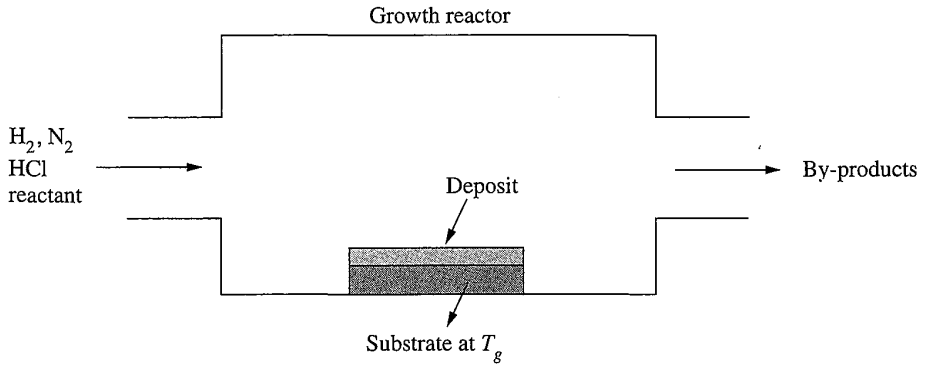


FIGURE 2
Schematic illustration of a typical VPE process.

Micron Ex. 1023, p. 134
Micron v. YMTC
IPR2024-00909

directing the reactant gases into the growth reactor, where the substrate is located and kept at a desired growth temperature, T_g. Once the reactant gases are fed into the growth reactor, these chemical species undergo a series of physical and chemical reactions, which result in the layer deposition. In a steady state, the sequence of the growth process[5] is best illustrated by Fig. 3. According to this figure, a growth process can be broken down into the following steps:

1. Introduction of the reactant species to the substrate region
2. Transfer of the reactant species to the substrate surface
3. Adsorption of the reactant species on the substrate surface
4. Surface diffusion, site accommodation, chemical reaction, and layer deposition
5. Desorption of residual reactants and by-products
6. Transfer of residual reactants and by-products from the substrate surface
7. Removal of residual reactants and by-products from the substrate region

During growth these steps may occur consecutively or in series, depending on the growth conditions such as growth temperature and growth pressure. These sequential steps describe only the fundamental reactions involving the VPE growth process. As with all chemical reactions, the slowest step of these reactions determines the overall rate of the growth process. This slowest step is called the *rate-limiting step.* This concept is very important to describe a growth process. If a growth process is dominated by step 3, 4, or 5 in the list above, one may say it is a *surface-controlled* process. On the other hand, if a growth process is dominated by step 1 or 2, it is called a *mass-transport-controlled* process. In all cases a growth process can be either surface-controlled or mass-transport-controlled at different growth conditions. We will discuss this point in some detail in the next section.

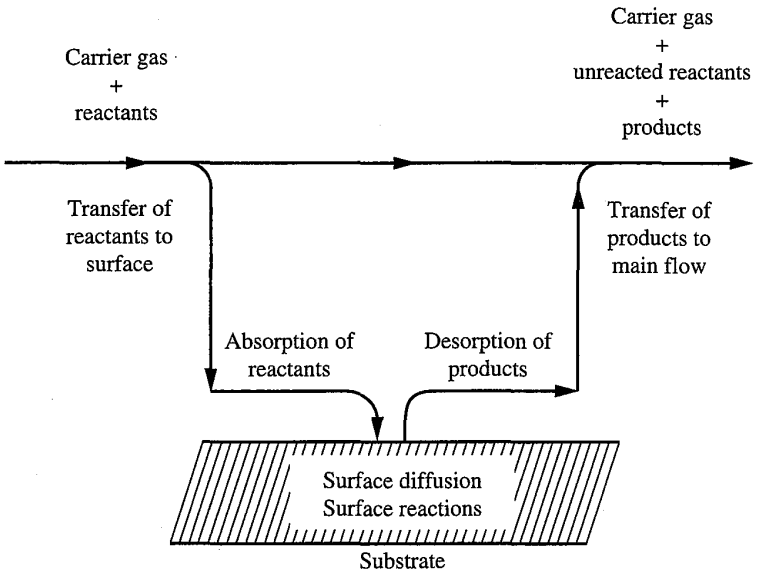


FIGURE 3
Schematic description of the sequential steps of a VPE process. (*After Shaw, Ref. 5.*)

Micron Ex. 1023, p. 135
Micron v. YMTC
IPR2024-00909

This growth sequence is only applicable to cases where the input reactant gases are flowing in one direction and passing over the substrate to the exhaust. This kind of flow pattern is called *laminar flow* and is a characteristic of a low-growth-pressure process. Conventionally, the growth system is operated at atmospheric pressure. Under this condition, turbulent flow inevitably occurs, which complicates the growth process. In addition, desorption of the reactant species from the walls of the tubing and furnace and the back-diffusion of gases from the exhaust make the growth process even more complicated. Even though a VPE growth process is too complicated to understand fully, many models have been proposed to describe the basic phenomena observed in this process. Before introducing these models, we shall briefly review the basic growth kinetics.

3.2.1 Basic Growth Kinetics of Epitaxy

Deposition of a thin epitaxial layer on the substrate can be described by two processes: nucleation and growth. In the nucleation process a small region of a new phase is created on another phase, in this case the surface of the substrate. When the newly formed phase grows to approximately a critical size, it becomes stable and is usually called a *nucleus;* below the critical size it is called an *embryo* and is unstable. A nucleation process proceeds in two ways: homogeneously and heterogeneously. The former is characterized by the formation of an embryo without the aid of other phases, such as foreign materials or surfaces, whereas the heterogeneous nucleation forms in the presence of other objects.[6] For a VPE process the former represents the formation of an embryo via gas-phase reactions; the latter, via the substrate surface.

In homogeneous nucleation, for a VPE process, a number of embryos are formed in the vapor phase. Subsequently, these embryos coalesce and condense into the solid phase. Such a process involves the release of free energy (ΔG_v), resulting from the condensation process, and an increase of free energy (γ_s), due to the formation of surfaces in the coalescence process. The former is a function of reactant gas partial pressure and increases as the reactant gas partial pressure increases. The latter depends on the surface energy of the associated phases in the growth region, such as vapor, substrate surface, and nucleus. The total free-energy change (ΔG_t) for an embryo of a radius r with homogeneous nucleation can be expressed as

$$\Delta G_t = \left(\frac{4\pi r^3}{3}\right)\Delta G_v + (4\pi r^2)\gamma_s \tag{3.1}$$

From the thermodynamic viewpoint, the critical radius r_c of an embryo can be obtained at maximum total free-energy change, where

$$r_c = \frac{-2\gamma_s}{\Delta G_v} \tag{3.2}$$

After substituting r_c in Eq. (3.2) into r in Eq. (3.1), the critical formation free-energy change ΔG_{fc} of a stable embryo is then written as

Micron Ex. 1023, p. 136
Micron v. YMTC
IPR2024-00909

$$\Delta G_{\text{fc}} = \frac{16\pi\gamma_s^3}{3(\Delta G_v)^2} \tag{3.3}$$

An embryo with a radius larger than r_c will tend to lower its free energy by growing. Under a quasi-equilibrium condition, the number of stable embryos, n_c, can be expressed by the following equation:

$$n_c = n_o \exp(-\Delta G_{\text{fc}}/kT) \tag{3.4}$$

where n_o is the total number of embryos, k is Boltzmann's constant, and T is the substrate temperature in kelvins. Finally, the nucleation rate of a homogeneous nucleation process, R_n, is given by

$$R_n = (n_o/t_e) \exp(-\Delta G_{\text{fc}}/kT) \tag{3.5}$$

where t_e is the average lifetime for stable embryos.

As already mentioned, most nucleation processes may occur heterogeneously on a substrate in the VPE process. Figure 4 illustrates the initial wetting characteristics on a substrate surface, where γ_{sv}, γ_{nv}, and γ_{ns} are surface energies associated with surface (s)/vapor (v), nucleus (n)/vapor (v), and nucleus (n)/surface (s), respectively, and θ is the contact angle. A mathematical derivation[6] gives the critical energy for heterogeneous nucleation:

$$\Delta G_{\text{hetero}} = \Delta G_{\text{homo}}[(2 + \cos\theta)(1 - \cos\theta)^2/4] \tag{3.6}$$

According to Eq. (3.6), with the aid of a substrate surface, ΔG_{hetero} is lower than ΔG_{homo}. Therefore, thermodynamically, heterogeneous nucleation is always preferable to homogeneous nucleation.

It can be noted from Eq. (3.6) that surface wetting (i.e., the contact angle θ) is a predominant factor for a heterogeneous nucleation process. Thus, surface preparation prior to the epitaxy is one of the most important steps that determine the quality of an epitaxial layer. In addition, since the value of ΔG_v decreases with decreasing reactant partial pressure and increasing temperature, performing the VPE process at reduced partial pressure and higher temperature is preferable.

An epitaxial process proceeds by impingement, adsorption, diffusion, reaction, and desorption of the reactant species. Many models have been proposed to describe this process. However, none of them can account for all aspects of the various growth

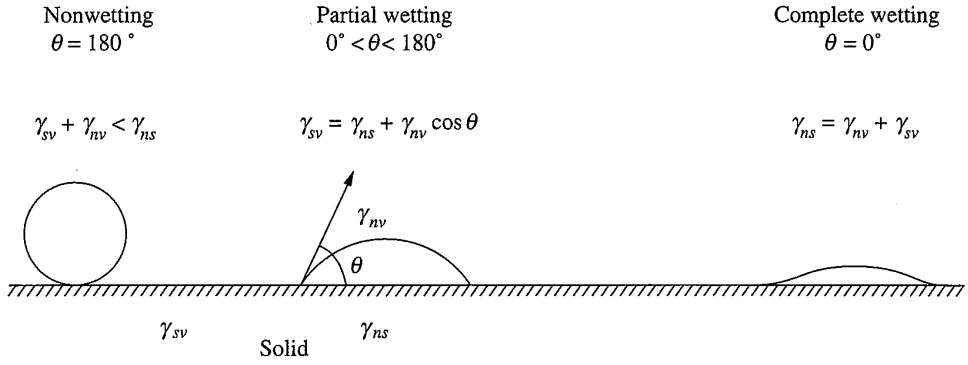


FIGURE 4
Wetting characteristics of a heterogeneous nucleation process. (*After Ghandhi, Ref. 3.*)

Micron Ex. 1023, p. 137
Micron v. YMTC
IPR2024-00909

results obtained by different growth techniques. In the next section we will introduce a few generally accepted models to explain some characteristics of an epitaxial growth process.

3.2.2 Growth Models for Epitaxy

A model[7] based on the assumption that only transport of the reactant species and chemical reactions are operative is shown in Fig. 5. In the schematic drawing of the model, F_i is the flux difference between the reactant in the input gas stream, C_g, and that at the substrate surface, C_s, and F_o is the consumption of the reactant at the substrate surface. Assuming a linear change from C_g to C_s, F_i can be expressed by

$$F_i = h_g(C_g - C_s) \tag{3.7}$$

where h_g is the gas-phase mass transfer coefficient and is relatively independent of temperature. For simplicity, by assuming F_o proportional to C_s, F_o is expressed as

$$F_o = k_s C_s \tag{3.8}$$

where k_s is the chemical surface reaction constant. The value of k_s, on the other hand, is a strong function of temperature and is usually expressed in the form of $k_s = k_o \exp(-E_a/kT)$, where k_o is a temperature-independent constant and E_a is the activation energy of the reaction.

At steady-state conditions the two fluxes must be equal. By equating Eqs. (3.7) and (3.8), one obtains

$$C_s = \frac{C_g}{1 + (k_s/h_g)} \tag{3.9}$$

If $h_g \gg k_s$, then $C_s \approx C_g$. This condition is called a surface-controlled process. On the other hand, if $h_g \ll k_s$, $C_s \approx 0$, and the process is called mass-transport-controlled. Simple mathematical derivations lead to the growth rate G of a VPE process given by

$$G = C_t k_s M/N \qquad \text{(surface-controlled case)} \tag{3.10}$$

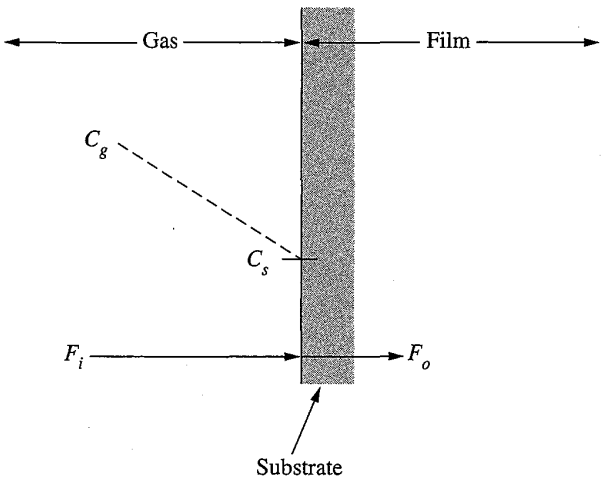


FIGURE 5
Schematic representation of the Grove (Ref. 7) model.

Micron Ex. 1023, p. 138
Micron v. YMTC
IPR2024-00909

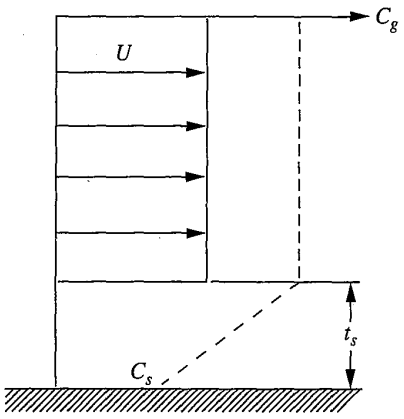
or by

$$G = C_t h_g M/N \qquad \text{(mass-transport-controlled case)} \tag{3.11}$$

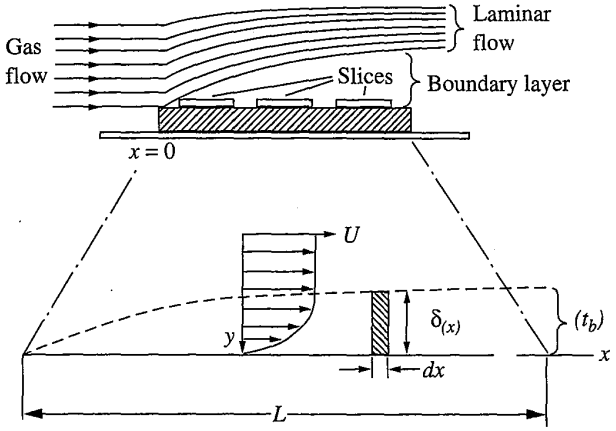
where C_t is the total number of molecules per cm^3 in the gas stream, M is the mole fraction of the reactant species in the gas stream, and N is the number of atoms incorporated into a unit volume of the layer.

Another model assumes that the gas stream above the substrate surface is divided into two regions.[8] As shown in Fig. 6*a*, in the region away from the substrate surface, the gas stream moves with a constant speed V in parallel to the substrate surface, where in the region next to the substrate surface, a stagnant layer exists in which the gas velocity is zero. In this model, mass transfer of the reactant species from the upper region across the stagnant layer to the substrate surface is dominated by a diffusion process. With this assumption, F_i in Eq. (3.7) can be rewritten as

$$F_i = D_g[(C_g - C_s)/t_s] \tag{3.12}$$



(*a*)



(*b*)

FIGURE 6
Schematic representations of (*a*) stagnant-layer model and (*b*) boundary-layer model. (*After Wolf and Tauber, Ref. 20.*)

Micron Ex. 1023, p. 139
Micron v. YMTC
IPR2024-00909

where D_g is the diffusion coefficient of the reactant species and t_s is the thickness of the stagnant layer. From Eqs. (3.7) and (3.12) we find

$$h_g = D_g/t_s \tag{3.13}$$

A model similar to the stagnant-layer model was developed to provide a more realistic estimate of the mass transfer coefficient h_g.[8] As shown in Fig. 6*b*, a boundary layer is present between the substrate, where the gas velocity is zero, and a free gas stream, where the gas velocity is u. In this model, u is not a constant value but changes rapidly across the entire range of substrate or susceptor. The boundary layer thickness t_b is a function of distance along the flat plate, x. Therefore, it can be expressed as

$$t_b(x) = \left(\frac{\mu x}{\rho u}\right)^{1/2} \tag{3.14}$$

where μ is the gas viscosity and ρ is the gas density. Upon integration, the average boundary layer thickness over an entire flat plate with length L is expressed as

$$t_b = \frac{(2/L)}{3(\mathrm{Re})^{1/2}} \tag{3.15}$$

where Re is the Reynolds number for the gas, given by

$$\mathrm{Re} = \frac{\rho u L}{\mu} \tag{3.16}$$

For $\mathrm{Re} < 2000$ the gas flow has a laminar pattern, whereas for larger Re values the gas flow is turbulent.

The average boundary layer thickness, t_b in Eq. (3.15), can be substituted for t_s in Eq. (3.12), resulting in an expression for h_g as

$$h_g = \frac{3D_g\mathrm{Re}^{1/2}}{2L} \tag{3.17}$$

According to this equation, in the mass-transport-controlled regime the growth rate of a layer depends on the square root of the free gas stream velocity u. This predicted dependency is observed in a typical experimental result for epi-Si grown at low flow rates, as shown in Fig. 7. At higher flow rates the growth rate reaches a maximum and becomes independent of the flow rate. Figure 8 shows the formation of boundary layers in a horizontal reactor. The spatial dependencies of reactant concentration, gas velocity, and temperature across the entire susceptor are illustrated. The fluxes of the various reactant species that cross this layer to the substrate surface are therefore dependent on factors such as boundary layer thickness, reactant concentration, growth pressure, and growth temperature.

Once the reactant species are adsorbed on the substrate surface (in which condition they are often called *adatoms*), the growth of the epitaxial layer is believed to occur via the migration of adatoms (position A in Fig. 9) to steps on the substrate surface (positions B, C in Fig. 9). From the thermodynamic viewpoint, the kink position (position C in Fig. 9) along a step is the most favorable site for growth to occur. Such a layer growth is assumed to proceed by lateral movements of the steps. In

Micron Ex. 1023, p. 140
Micron v. YMTC
IPR2024-00909

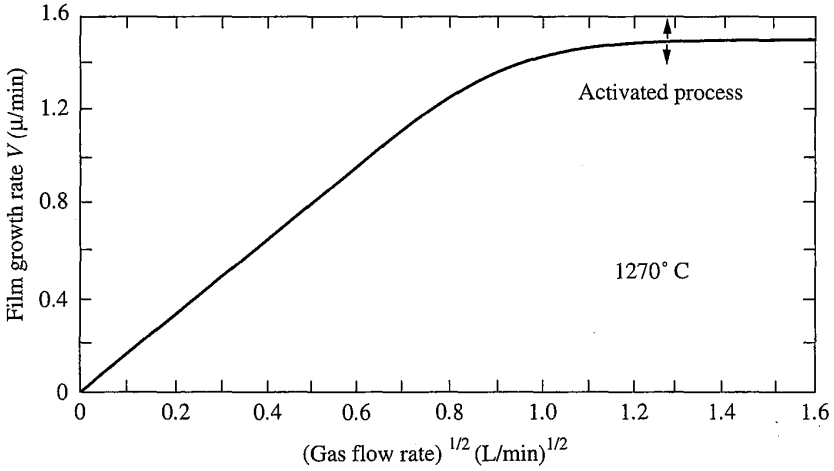


FIGURE 7
The growth rate of an epi-Si versus the free gas flow rate. (*After Wolf and Tauber, Ref. 20.*)

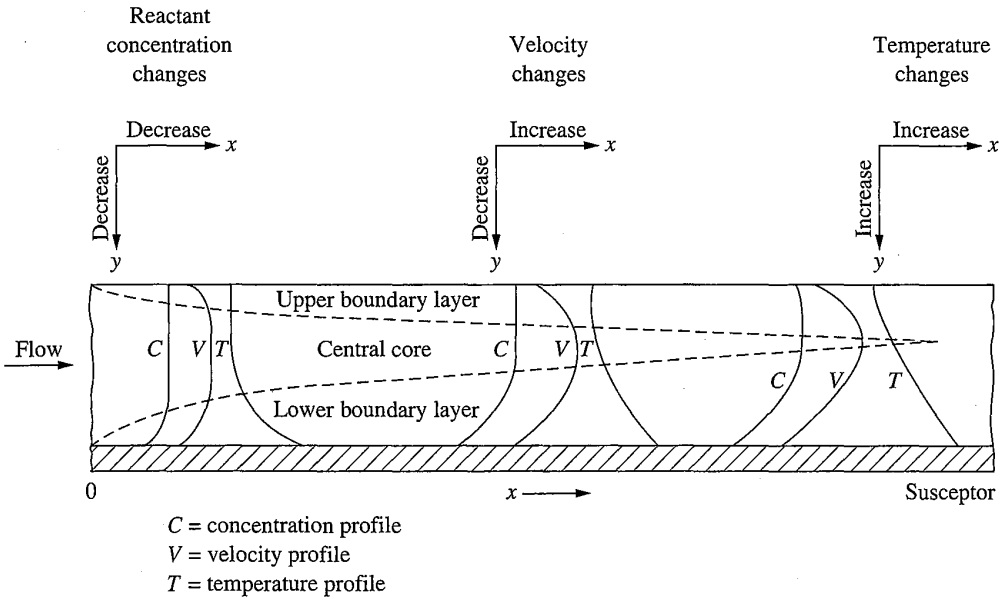


FIGURE 8
Schematic illustration of the formation of boundary layers in a horizontal reactor. (*After Ban, Ref. 8.*)

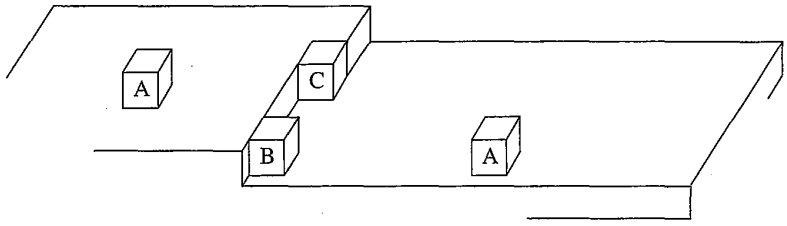


FIGURE 9
Schematic illustration of epitaxial growth with adatoms. (*After Wolf and Tauber, Ref. 20.*)

Micron Ex. 1023, p. 141
Micron v. YMTC
IPR2024-00909

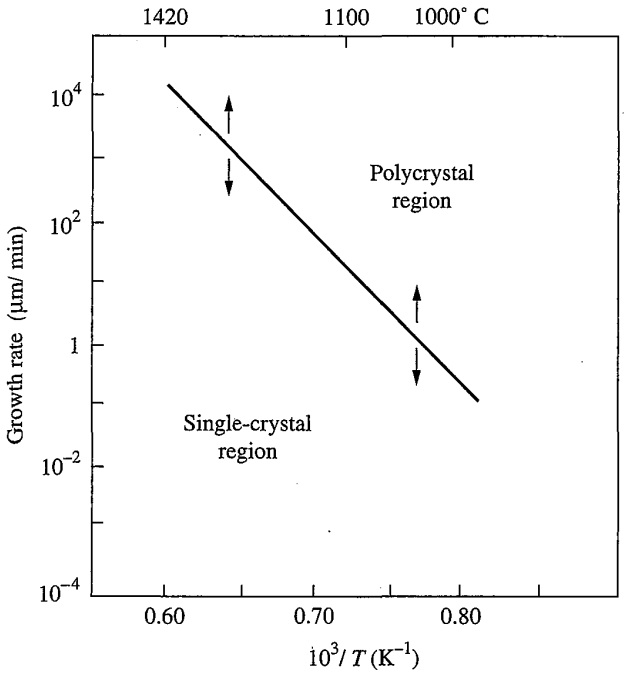


FIGURE 10
Crystallinity with respect to growth rate and growth temperature. (*After Bloem, Ref. 10.*)

this growth model, a "maximum growth rate" is imposed to ensure the crystallinity of the resulting layer. At high growth rates, not enough time is available for the adatoms to migrate, resulting in a polycrystal layer. However, as the growth temperature increases, the surface migration rate increases, facilitating single-crystal layer growth. Figure 10 shows the crystallinity of a layer with respect to temperature and growth rate. For a fixed temperature, there is a maximum growth rate for an epi-layer. Likewise, for a fixed growth rate, a minimum temperature is required to achieve an epi-layer. For Si epitaxy, an activation energy of $\approx$5 eV is observed. This value is comparable to the activation energy of self-diffusion in Si. Therefore, self-diffusion in Si is believed to be the dominant mechanism for the growth of Si epi-layers.[10]

3.3 CONVENTIONAL Si EPITAXY

The term *conventional Si epitaxy* is used to differentiate it from the other approaches, such as low-temperature Si epitaxy and selective Si epitaxy, which will be discussed in the following sections. Conventional Si epitaxy was originally developed to overcome the problems associated with high concentrations of impurities present in the substrate. This technique was then used extensively to optimize the performance of bipolar transistors and bipolar integrated circuits. It enables the growth of a lightly doped Si epi-layer over a heavily doped Si substrate. As a result, a higher collector–substrate breakdown voltage is achieved with minimal increase of collector resistance. Similarly, if CMOS devices or circuits are built in the lightly doped Si epi-layer over the heavily doped Si substrate, the latch-up effect is greatly reduced.[4]

Micron Ex. 1023, p. 142
Micron v. YMTC
IPR2024-00909

The advantages of conventional Si epitaxy are its flexibility and accuracy in controlling the doping concentration and doping profile. However, since this technique is usually carried out at temperatures above 1100°C, it suffers from problems such as autodoping and pattern shift. In addition, because the epitaxy is an "arranged on" process, it is susceptible to defect generation due to strain, impurities, and particles on the substrate prior to or during the growth process. In this section, we present (*a*) growth chemistry; (*b*) doping and autodoping; (*c*) defect generation and minimization; (*d*) pattern shift, distortion, and washout; and (*e*) epitaxial deposition equipment.

3.3.1 Growth Chemistry

The success of an epitaxial growth relies heavily on the chemistry used in the process. The major source gases used to deposit Si epi-layer commercially contain hydrogen and chlorosilanes, such as silicon tetrachloride ($SiCl_4$), trichlorosilane ($SiHCl_3$), and dichlorosilane (SiH_2Cl_2). In addition, silane (SiH_4) has also been used when a lower growth temperature is desired. The choice of each source gas is based on the growth conditions and layer requirements. One of the most important parameters that affect the selection of the Si source gas is the growth rate. Figure 11 shows the growth rate of Si layers as a function of growth temperature for various silicon sources.[11] Two distinct growth regimes are noted in this figure. At low growth temperatures (region A) the growth rate of Si layer is exponentially dependent on temperature, which is evidence of a surface-controlled reaction. In the high-temperature regime (region B), the growth rate is nearly independent of the growth temperature, which is typical of a mass-transport or diffusion-controlled

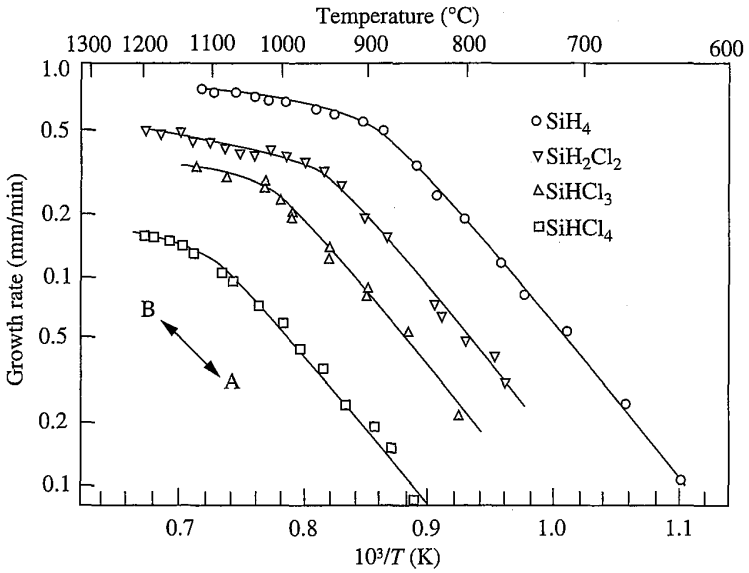


FIGURE 11
Growth rate of epi-Si as a function of growth temperature for a variety of Si sources. (*After Eversteyn, Ref. 11.*)

Micron Ex. 1023, p. 143
Micron v. YMTC
IPR2024-00909

reaction. It is important to note that at the lower temperature, as-grown Si layers are in polycrystal form. Epi-Si layers are formed at temperatures above the transition point of each curve seen in Fig. 11. The position of the transition point varies with the mole fraction of reactant, gas flow rate, and reactor type. However, it can be inferred from this figure that epi-Si can be grown from the SiH_4 gas at temperatures above 900°C, whereas for $SiCl_4$ gas it is above 1100°C.

Reductions of $SiCl_4$ and SiH_4 to solid Si encompass two extremes of the chemical reactions involving conventional Si epitaxy. The overall reaction for the reduction of $SiCl_4$ in H_2 can be written as

$$SiCl_4 + 2H_2 \longrightarrow Si + 4HCl \tag{3.18}$$

In reality, a study of thermodynamics of the H–Cl–Si system showed that a number of chemical reactions were occurring either simultaneously or sequentially in the deposition process. It is postulated[12] that the net reaction of Eq. (3.18), may proceed as follows:

$$SiCl_4 + H_2 \longleftrightarrow SiHCl_3 + HCl \tag{3.19}$$

$$SiHCl_3 + H_2 \longleftrightarrow SiH_2Cl_2 + HCl \tag{3.20}$$

$$SiH_2Cl_2 \longleftrightarrow SiCl_2 + H_2 \tag{3.21}$$

$$SiHCl_3 \longleftrightarrow SiCl_2 + HCl \tag{3.22}$$

$$SiCl_2 + H_2 \longleftrightarrow Si + 2HCl \tag{3.23}$$

Note that all the above reactions, Eqs. (3.19) to (3.23), are reversible. Therefore, both reduction and etching processes compete with each other, depending on the mole fraction of reactants and growth temperature. Figure 12 shows the boundary between etching and deposition in a $SiCl_4$ and H_2 mixture at atmospheric pressure as a function of temperature and partial pressure of $SiCl_4$.[13] A separate study also showed that the growth rate of silicon varied with the growth temperature with $SiCl_4 + H_2$ chemistry. Figure 13 shows that an etching process is favored at both low and elevated

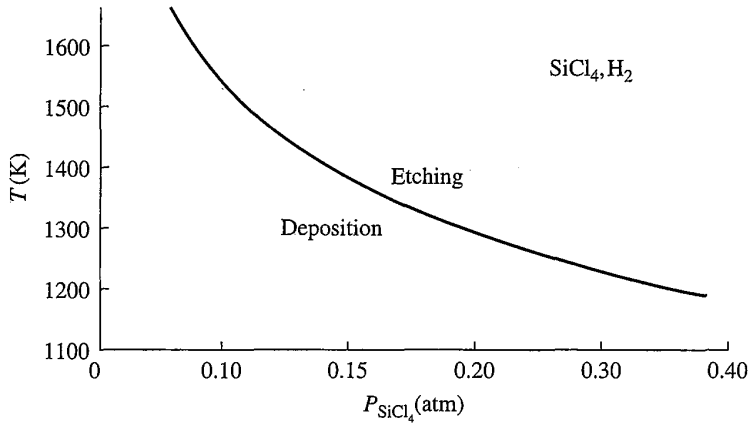


FIGURE 12
Growth temperature as a function of the partial pressure of $SiCl_4$ in a $SiCl_4 + H_2$ mixture at one atmosphere. (*After Ban and Gilbert, Ref. 13.*)

Micron Ex. 1023, p. 144
Micron v. YMTC
IPR2024-00909

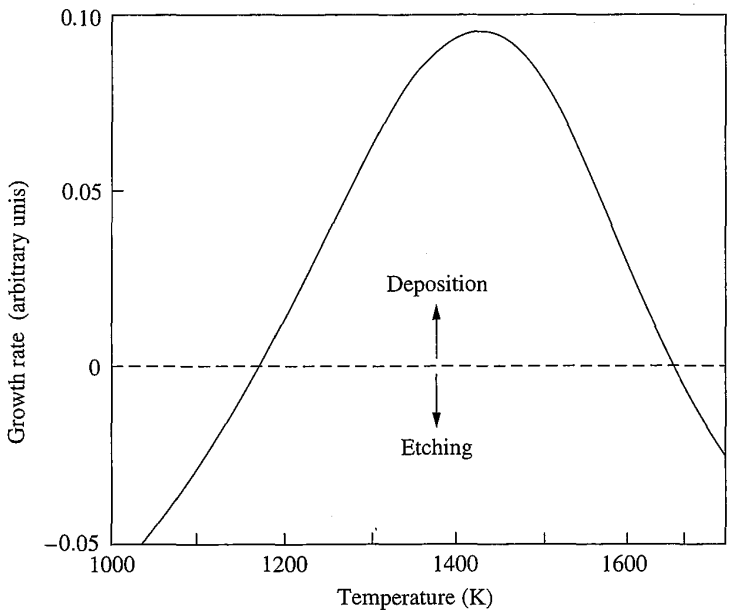


FIGURE 13
Growth rate of Si films as a function of growth temperature for $SiCl_4 + H_2$ chemistry. (*After Sirtl, Hunt, and Sawyer, Ref. 14.*)

temperatures.[14] Therefore, the deposition process is usually performed at temperatures between 1100 and 1300°C in this case. For thicker layer growth $SiHCl_3$ is often used because its growth rate is higher than that of $SiCl_4$.

In contrast to $SiCl_4$ chemistry, the reduction reaction of SiH_4 is not reversible and proceeds by the following overall reaction:

$$SiH_4 \longrightarrow Si + 2H_2 \tag{3.24}$$

As noted in Fig. 11, the major advantage of the SiH_4 is that epi-Si can be obtained at relatively lower temperatures than any of the chlorosilane chemistries. However, it is very difficult to avoid some degree of gas-phase nucleation of Si via the homogeneous reaction of SiH_4. As a result, Si particulate is formed, which results in poor surface morphology or even polycrystal Si growth. This problem can be minimized by raising the gas velocity through the growth region via a low-pressure operation or by controlling the deposition temperature. In constrast to the chlorosilane chemistry, SiH_4 is more susceptible to oxidation and is explosive. Therefore, it is not commonly used in the conventional Si epitaxial process. Because there is no HCl in the SiH_4 growth system, there is no etching process. This may result in higher metallic impurities in the film grown with SiH_4 unless a careful preclean process is used.

3.3.2 Doping and Autodoping

The conductivity type and electrical resistivity of an epitaxial layer are ultimately dependent on the doping characteristics. In an epitaxial process, dopants are incorporated into the grown layer simultaneously or intermittently during the layer

Micron Ex. 1023, p. 145
Micron v. YMTC
IPR2024-00909

deposition reaction. Typically, hydrides of various types are used to introduce these dopants. For Si epitaxy, diborane (B_2H_6) is used to incorporate the p-type dopant B, whereas phosphine (PH_3) and arsine (AsH_3) are used for the n-type dopants P and As, respectively. These gases are extremely toxic and are unstable above room temperature. Therefore, they are normally diluted with a large amount of H_2. Because of this, the hydride doping gases would not follow a simple rule for incorporation into the growing Si film from the gas phase. The dopant/Si ratio in solid phase corresponding to that in gas phase can only be obtained through experiments for each growth condition and each growth reactor. The many factors that affect the dopant incorporation include the growth temperature, growth rate, dopant/Si ratio in gas phase, and reactor geometry.

In addition, the interactions or competing reactions between the dopant gases and Si source gases make the doping process even more complicated. It has been found that B_2H_6 and PH_3 have opposite influences on the silicon deposition rate.[15] The former enhances the silicon deposition rate, whereas the latter suppresses it. It has been found that PH_3 chemisorbs dissociatively with the formation of Si–H bonds at high temperatures. Hydrogen can be desorbed at temperatures above 400°C to form a phosphorus layer, which has a maximum adsorption at 550°C. Once adsorbed, PH_3 is stable under a SiH_4 flux, effectively passivating the silicon surface and suppressing the growth rate in the phosphorus-doped Si epitaxy.

On the other hand, B_2H_6 has a very small sticking coefficient and is easy to decompose directly on the silicon surface to deposit boron. The adsorbed B_2H_6 facilitates the heterogeneous reaction of SiH_4, resulting in enhanced growth rate. The amount of dopant incorporation in epi-Si is influenced by the growth rate. Figure 14

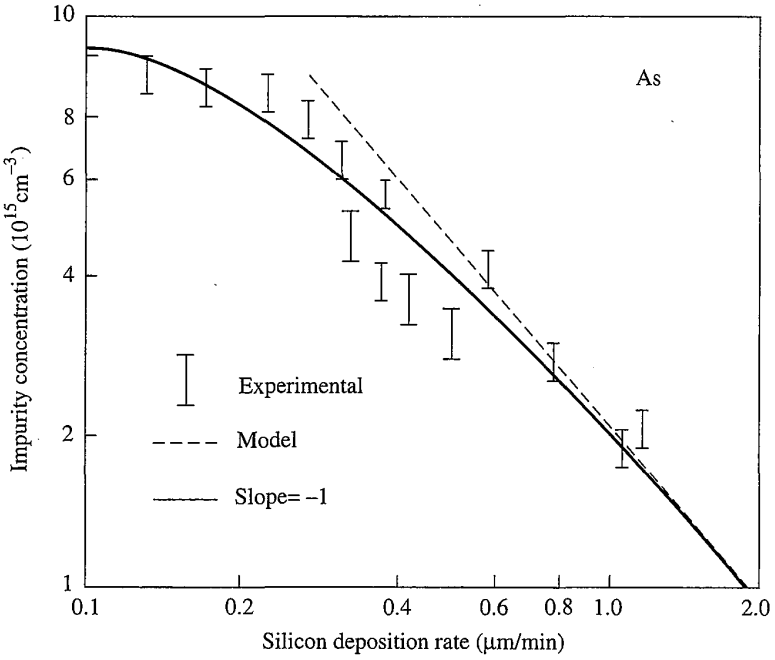


FIGURE 14
A typical plot of arsenic doping concentration as a function of silicon deposition rate. (*After Reif, Kamins, and Saraswat, Ref. 16.*)

Micron Ex. 1023, p. 146
Micron v. YMTC
IPR2024-00909

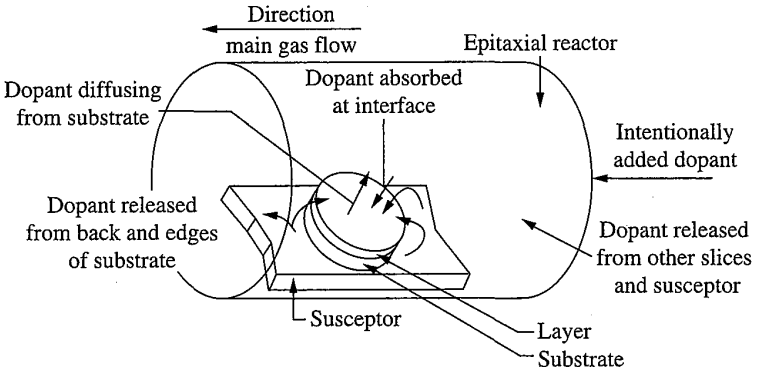


FIGURE 15
Schematic illustration of various sources of dopants for a VPE growth process. (*After Srinivasan, Ref. 17.*)

shows typical results of impurity incorporation as a function of Si deposition rate for arsenic doping.[16] At high growth rates the incorporation of As is dominated by surface kinetic processes such as surface adsorption and surface diffusion, while at low growth rates it is controlled by the mass transport process.

The foregoing discussion applies only to the doping process that is performed intentionally to obtain different conductivity types and various resistivities in the epi-Si layer. In practice, the doping process is complicated by an unintentional process named *autodoping*. Figure 15 shows a schematic drawing of various sources of dopants for an epitaxial growth.[17] Dopants can be released from the substrate or susceptor through outdiffusion or outgassing. The outdiffusion is in the solid state, whereas the outgassing is in gas phase and is often termed the *memory effect*. Figure 16 shows a plot of a typical doping profile of an epitaxial layer grown over a heavily

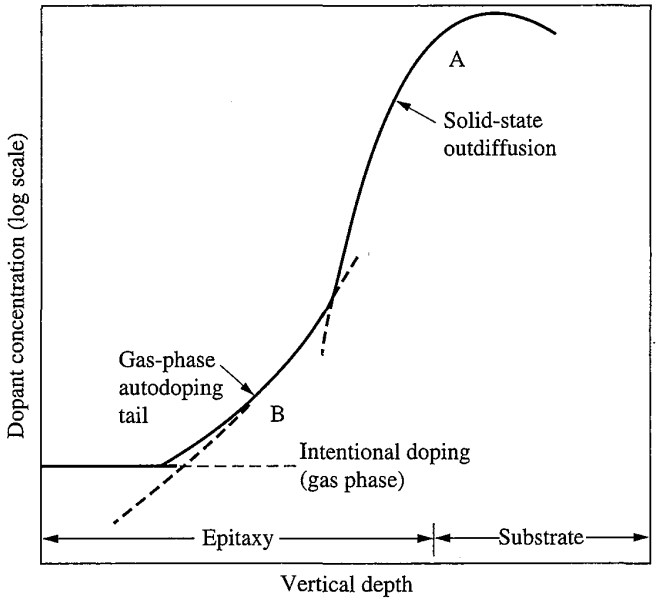


FIGURE 16
A typical doping profile of an epi-Si grown over a heavily doped Si substrate. (*After Srinivasan, Ref. 18.*)

Micron Ex. 1023, p. 147
Micron v. YMTC
IPR2024-00909

doped substrate.[18] Several regions of autodoping are noted. Close to the substrate/epi interface, outdiffusion in solid phase from the substrate is predominant, resulting in a broad transition region. This effect is soon limited, because the growth rate is normally much higher than the speed of solid-state outdiffusion. Thereafter, the dopant incorporation is controlled by the sources in gas phase. If the outgassing dopant exceeds that intentionally introduced, an autodoping tail develops. However, this effect becomes insignificant as the growth process proceeds, because of the coverage of a low-doped layer over substrate and susceptor, which slows down the supply of the dopant to the surface (i.e., the outgassing speed). The extent of the autodoping tail is a function of substrate dopant species and its concentration, reactor geometry, growth temperature, growth rate, and growth pressure. Once the autodoping is minimized, the intentional doping process predominates.

For most ULSI applications lightly doped epi-Si layers over a heavily doped substrate are required. Autodoping imposes a constraint on the minimum layer thickness and minimum doping level for an epi-Si layer with controlled doping. Therefore, it is a subject of great technological importance. Various techniques are proposed to minimize this effect. Low growth temperature seems to reduce boron autodoping, whereas lower growth pressure works well for phosphorus and arsenic. In-situ vapor-phase etching, using HCl, or cap-seal, with a thin lightly doped Si over the substrate and susceptor, are also employed. With the advent of low-temperature epitaxial growth techniques, the autodoping effect has been greatly minimized.

3.3.3 Defect Generation and Minimization

In an epitaxial process various types of defects are often observed. These defects may result from substrate imperfection, contamination or inclusion in the reactor or substrate, or any mismatch between the substrate and epi-layer. In principle, the defect density of an epi-layer cannot be lower than that of its parent substrate. Defects due to the substrate are related to the surface conditions prior to the deposition process. Therefore, care must be taken to remove organic and metallic impurities and particulate on the front and rear substrate surfaces. In addition, a defect-free wafer is preferred for an epitaxial process.

There are four types of crystalline defects in an epitaxial layer: *point defects,* such as impurities and vacancies; *line defects,* such as dislocations; *area defects,* such as stacking faults; and *volume defects,* such as voids and precipitates. Figure 17 illustrates the defects that commonly occur in epitaxial layers.[19] In this figure, 1 represents a dislocation line propagated from the substrate through the growing layer; 2 is a stacking fault nucleated from the impurity or precipitate on the substrate surface; 3 is an impurity precipitate incorporated by contamination in the growth system; 4 is a *growth hillock* caused by the process or surface finish of the substrate; and 5 is a bulk stacking fault or void caused by an existing surface defect in the substrate. In addition to these, defects such as dislocations or slip lines can be generated from the thermal stresses set up by the temperature gradients between the front and rear substrate surfaces during the deposition process, as illustrated in Fig. 18.[20] Furthermore, the existence of lattice mismatch between the substrate and epi-layer leads to the formation of dislocations to accommodate the misfit strain.

Micron Ex. 1023, p. 148
Micron v. YMTC
IPR2024-00909

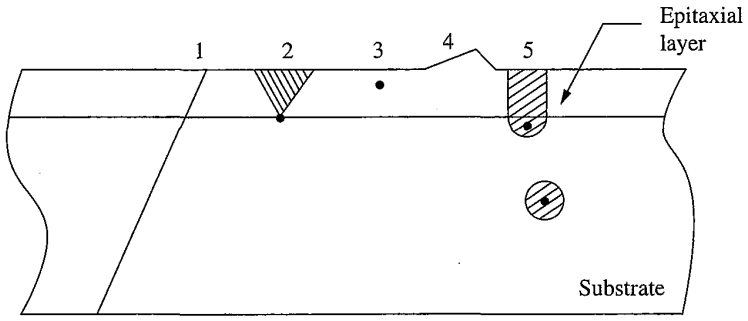


FIGURE 17
Schematic illustration of common defects in epitaxial layers. (*After Ravi, Ref. 19.*)

One of the most important measures for minimizing the defects in epi-Si is the substrate cleaning. The pre-epitaxy substrate cleaning (see Chapter 2) consists of a wet clean followed by a dilute HF dip and an in-situ HCl, HF or SF_6 vapor etch.[21] This preclean process is especially important for low-temperature Si epitaxy and will be discussed separately in Section 3.4.2. In addition to the wafer cleaning, defects in epitaxial layers can be reduced by employing "denuded" substrates.[22] Such a substrate is obtained by using an oxygen intrinsic gettering technique to minimize the defect and metallic impurity levels on the substrate surface and thus reduce the defect density in the growing layer. For heteroepitaxy, misfit dislocations due to lattice mismatch are inevitably formed. It has been found that a strain field near the misfit dislocation serves as a gettering site to trap impurities.[23] A strained-layer structure grown between the substrate and active device layers was also found to be effective in gettering heavy metals without introducing misfit dislocations into the device layer, leading to better device performances.

3.3.4 Pattern Shift, Distortion, and Washout

Pattern shift, distortion, and washout (the extreme case of pattern shift) occurs when an epitaxial layer is deposited over a buried layer in the fabrication of modern bipolar integrated circuits. Schematic representations of these phenomena[24] are shown

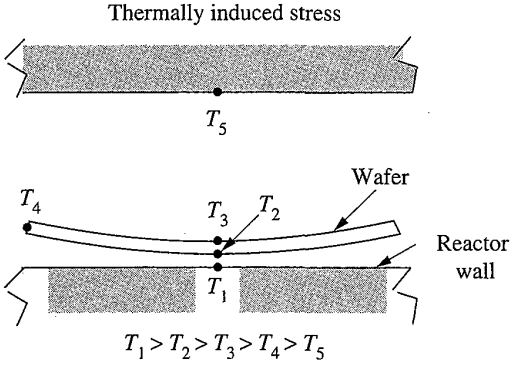


FIGURE 18
Schematic drawing of wafer bow, due to a temperature gradient between front and back of a wafer during a growth process. (*After Wolf and Tauber, Ref. 20.*)

Micron Ex. 1023, p. 149
Micron v. YMTC
IPR2024-00909

in Fig. 19. The main cause of these effects is the different growth rates on various crystal orientations. However, the amount of shift or distortion also depends on deposition rate, growth temperature, growth pressure, and silicon source. Therefore, to compensate these effects the circuit designer must adjust the position of features on the subsequent mask levels.

It has been found that pattern shift is at a minimum when the pattern is oriented on-axis with respect to a (100)-oriented substrate. For the case of (111)-oriented substrates, a misorientation of 2 to 5 degrees from the [111] to the [110] direction is needed to minimize these effects. For both cases the amount of shift seems to decrease with increasing growth temperature and decreasing growth rate.[25] Pattern shift can be completely prevented if SiH_4 is used. In addition, reduced growth pressure seems to minimize the amount of shift. In contrast to pattern shift, the amount of pattern distortion increases with increasing temperature and decreasing growth rate,[26] and the use of SiH_4 results in more severe pattern distortion than use of chlorosilanes. Moreover, pattern distortion occurs in both (100) and (111) substrates. Reduced growth pressure may be effective for controlling such a distortion for the (111)-oriented substrate, but not for the (100)-oriented substrate. However, increased layer thickness seems to reduce the distortion for both cases. Therefore, an optimized epitaxial process is required to minimize both pattern shift and pattern distortion.

3.3.5 Epitaxial Deposition Equipment

A typical epitaxial growth equipment consists of a gas distribution system, a high-purity quartz bell jar or reactor tube, a susceptor, a substrate heating system, an electrical control system, a cooling system, and an exhaust system.

The gas-handling system is used to feed the reactant gases, such as silane or chlorosilanes, H_2, and HCl; doping gases, such as AsH_3, PH_3, and B_2H_6; and purging gas, H_2 or N_2. Note that except for N_2, all other gases are explosive, corrosive, or highly toxic. Therefore, built-in safety interlocks, together with hydrogen and

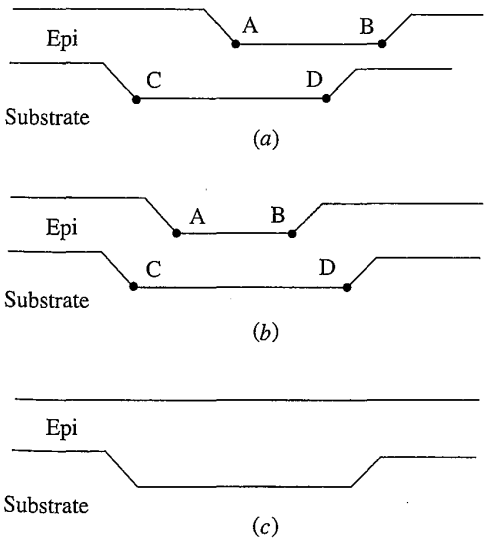


FIGURE 19
Schematic illustrations of (*a*) pattern shift, (*b*) pattern distortion, and (*c*) pattern washout. (*After Wolf and Tauber, Ref. 24.*)

Micron Ex. 1023, p. 150
Micron v. YMTC
IPR2024-00909

toxic-gas monitors, are required for preventing accidents. In addition, the entire growth system must be operated under leaktight conditions to ensure safety. In order to meet the stringent requirements of an epitaxial process for ULSI applications, high-precision mass-flow controllers with a vent-run configuration are used to obtain tight control on gas flows. Electropolished stainless steel tubing is used to transport all gases fed into the growth reactor. All gases must be of high purity or passed through a gas purifier (e.g., Pd-diffused H_2) to ensure minimum impurity incorporation. An exhaust system, consisting of a burn box and a wet or dry scrubber, is necessary to safely remove or treat unreacted by-products from the growth system.

A high-purity quartz bell jar is normally used as the growth reactor. The susceptor in the reactor is used to support substrates and to supply thermal energy for the reaction. Therefore, it should be mechanically strong and not reactive to the reactants and their by-products. In addition, a susceptor must be noncontaminating to the growth system. Graphite is commonly used as the primary material for susceptors. Silicon nitride– or silicon carbide–coated polysilicon or quartz susceptors are often used as alternatives. Coating is also required for the graphite susceptor, because it is impure and soft. In addition to SiC and Si_3N_4, glassy carbon and porolytic graphite are often used to coat the susceptor. Care must be taken to avoid any crack or scratch on the susceptor and to eliminate any contaminating source from the susceptor.[27]

Commercially, Si epitaxy is carried out in reactors of the types shown in Fig. 20. Susceptors of horizontal, pancake, and barrel shapes are widely used.[27] The horizontal reactor is the most commonly used system. It offers high capacity and throughput. However, it suffers from an inability to achieve a uniform deposition over the entire susceptor. Tilting the susceptor a few degrees (1.5 to 3 degrees) seems to reduce this problem greatly. In contrast, the vertical pancake reactor is capable of very uniform growth with minimal autodoping problems. Disadvantages of this system include its mechanical complexity, its low throughput, and its susceptibility to particulate incorporation. The barrel reactor is an expanded version of the horizontal reactor

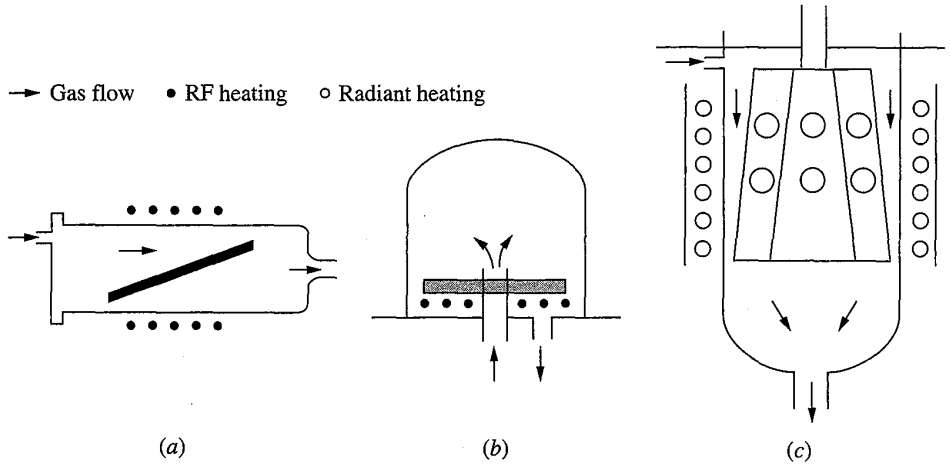


FIGURE 20
Types of conventional epitaxial reactors: (*a*) horizontal reactor; (*b*) vertical reactor; (*c*) barrel reactor. (*After Pearce, Ref. 27.*)

Micron Ex. 1023, p. 151
Micron v. YMTC
IPR2024-00909

in a different configuration. When used with a tilted susceptor, this system allows high-volume production and uniform growth.

Two types of substrate-heating systems are used to provide the thermal energy to the reaction: inductive heating and radiant heating. In the former case the susceptor is heated inductively by the underlying water-cooled RF coils. These coils are usually set outside the reaction chamber but close enough to the susceptor to obtain RF coupling. Once the susceptor is heated, the thermal energy is transported to the substrate by conduction and radiation. In the latter case the energy is supplied by radiant heating using banks of quartz-halogen lamps via IR absorption. This method provides more uniform heating than inductive heating does.[28] Temperature measurement is usually done with an optical or IR pyrometer. To avoid contaminating the growth environment, thermocouples should not be used in the reactor.

3.4 LOW-TEMPERATURE EPITAXY OF Si

Low-temperature epitaxy (LTE)[29] of Si is a process that produces epitaxial growth of the Si layer at temperatures of 550°C or less, much lower than the conventional Si epitaxy process does. Because tight controls over interfacial quality, spatial dopant distribution, and layer thickness are required for many advanced Si devices such as heterojunction bipolar transistors (HBTs), low processing temperature becomes necessary to minimize thermal diffusion and/or mass-transport-controlled processes. This section will start with an introduction of the chemical aspects of the LTE, followed by a discussion of a key step for this technique: surface preparation. There are many growth techniques that can be classified as LTE; this section will introduce the techniques that attract the most research interest. Finally, device applications of the LTE process are reviewed.

3.4.1 Chemical Aspects

The reason a Si layer can be deposited epitaxially onto a substrate at low temperature is primarily the growth conditions, such as substrate surface material, purity of Si source gas, and growth pressure.

To grow an epitaxial layer, the seed substrate surface should be of single-crystal material. It is well known that SiO_2 is readily formed on Si substrate under equilibrium conditions and becomes unstable at a temperature higher than 1000°C at atmospheric growth pressure. It has been found that the stability of the SiO_2 film depends on the partial pressures of oxygen and water vapor in the growth environment. As shown in Fig. 21, the crossover between the oxidation and the oxide-free regions is a function of water-vapor partial pressure and temperature.[29] Similar results have been observed for oxygen. Therefore, one of the most important criteria to realize LTE is to control the oxygen and water-vapor partial pressures in the growth reactor. In order to grow the Si epi-layer at a temperature lower than 700°C, the partial pressure of water vapor should be lower than 10^{-8} torr, maintaining a bare single-crystal

Micron Ex. 1023, p. 152
Micron v. YMTC
IPR2024-00909

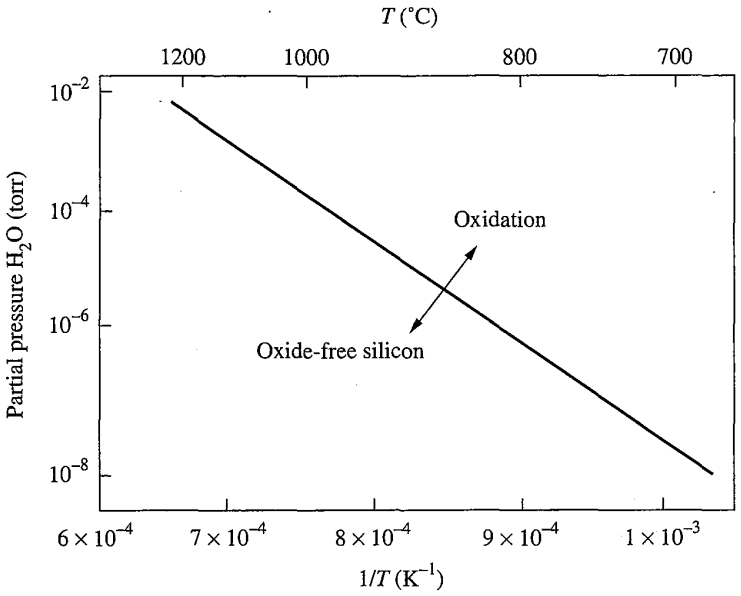


FIGURE 21
Water vapor pressure vs. reciprocal temperature for oxide-free and oxidized Si surfaces. (*After Meyerson, Ref. 29.*)

Si substrate surface. Presumably, in the oxidation region, the stable SiO_2 is formed on the silicon surface by the chemical reaction

$$\text{Si (solid)} + 2H_2O\text{ (gas)} \rightarrow SiO_2\text{ (solid)} + 2H_2\text{ (gas)} \tag{3.25}$$

In the oxide-free region, on the other hand, the water vapor serves to etch the silicon surface by the following chemical reaction:

$$\text{Si(solid)} + H_2O\text{(gas)} \rightarrow \text{SiO(gas)} + H_2\text{(gas)} \tag{3.26}$$

There are many ways to achieve a bare Si single-crystal surface prior to the epitaxial growth process. As mentioned above, the straightforward approach is to keep the growth environment at an ultrahigh vacuum. Growth techniques employing this approach are ultrahigh vacuum/chemical vapor deposition (UHV/CVD) and molecular beam epitaxy (MBE). Another approach is to employ ultraclean gases with ultralow oxygen and water-vapor contents. Growth techniques using this approach are atmospheric-pressure chemical vapor deposition (APCVD) and ultraclean low-pressure CVD. In addition, plasma-enhanced CVD (PECVD) and photochemical CVD (photo-CVD) have been proposed to achieve LTE-grown epi-Si layers.

The surface preparation of the Si substrate prior to the layer deposition is a key step in a successful LTE process. This subject will be discussed in the next section. To date, most LTE studies use SiH_4 as the Si source gas; other gases, such as Si_2H_6 and $SiCl_2H_2$, have also been used, but to a lesser extent. The major advantage of the SiH_4 source is its low decomposition temperature, which facilitates the LTE process. In addition, H_2 plays an important role in the LTE process. It has been reported that the desorption rate of H_2 dominates the growth rate of the Si layer in the UHV/CVD LTE process.[30]

Micron Ex. 1023, p. 153
Micron v. YMTC
IPR2024-00909

3.4.2 Surface Preparation

Surface preparation is a key step in the LTE process. The conditions of the initial growth interface determine the quality and the content of unintended impurities at addition to providing an ultraclean growth environment, all LTE growth techniques rely strongly on the preparation of an ultraclean substrate surface prior to growth. A wide variety of surface cleaning techniques have been proposed and implemented in the LTE process. Those techniques can be classified into two groups, namely ex-situ and in-situ. A combination of both ex-situ and in-situ methods is frequently used. The ex-situ techniques perform the surface cleaning process outside the growth reactor, whereas the in-situ techniques perform it inside the growth reactor.

The most widely used ex-situ cleaning technique involves the removal of organic impurities and particles using H_2SO_4: H_2O_2 and NH_4OH: H_2O_2 and the removal of metallic impurities by growth of a thick protective SiO_2 layer using a Si oxidizer, such as a HCl: H_2O_2 solution. This method is often called the *RCA clean* or *Shiraki clean* in the literature.[31] Following the wet chemical ex-situ cleaning process, two methods have been used to remove the oxide. One is to dip the substrate in the HF: H_2O (e.g., 1 : 10) solution ex-situ, and the other is to desorb the surface oxide in the growth reactor. In the case of the former, it has been observed that after a dip in HF solution, a hydrophobic (i.e., hydrogen-passivated) silicon surface is formed. This hydrogen-passivated silicon surface is air-stable and remains oxide-free for more than 10 minutes.[32] Using this method, device-quality Si and SiGe epi-layers were obtained at 550°C or lower without employing any in-situ surface-cleaning process. In the case of the latter, the thick oxide was removed by thermally desorbing it at a temperature above 950°C in ultrahigh vacuum or by etching it at a temperature above 800°C in H_2 inside the growth reactor.[33] The drawback to this approach is the use of a high-temperature process, which results in rapid impurity diffusion, extensive outgassing, and wafer warpage. In addition, the high-temperature process complicates integrated device fabrication; hence, it is not desirable from a manufacturing viewpoint.

Recently, a number of low-temperature Si-cleaning techniques have been developed. Among them *Si beam cleaning,* which employs a small flux of Si to desorb a thin Si oxide, has brought the temperature down to 700°C. Furthermore, with an in-situ exposure of a clean Si surface to H atoms, a successful Si epitaxy can be performed at 600°C. Etching in gaseous HF has also been reported to result in a H-passivated surface that allows the Si epitaxy at 550°C. However, a 200°C prebake is essential for low-temperature epitaxy on HF-dipped Si.[34] By incorporating this process, Si epitaxy has been performed at a temperature of 370°C.

3.4.3 Low-Temperature Si Epitaxial Growth Techniques

There are many ways that Si epi-layers can be grown at low temperatures. All of these techniques rely on both an ultraclean growth environment and a unique Si surface-cleaning process. CVD[29] and MBE[35] are the most popular methods used by researchers. However, solid-phase epitaxy (SPE) has also been reported.[36]

CVD has received the most research interest because of its simplicity and manufacturability. Many CVD approaches have been reported, such as ultrahigh

Micron Ex. 1023, p. 154
Micron v. YMTC
IPR2024-00909

vacuum/chemical vapor deposition (UHV/CVD),[29] ultraclean low-pressure CVD,[37] rapid-thermal-processing CVD (RTPCVD),[38] limited-reaction-processing CVD (LRP-CVD),[38] atmospheric-pressure CVD (APCVD),[39] plasma-enhanced CVD (PECVD),[40] photo-CVD,[41] and excimer laser–assisted CVD.[42] These techniques can be divided into three groups based on their characteristics, namely ultrahigh vacuum and/or ultraclean environment, rapid thermal processing, and enhanced-mode processing. In this section, we will only consider the two most promising techniques: UHV/CVD and MBE. For details of the other techniques, see the references herein.

A schematic drawing of a UHV/CVD system[29] is shown in Fig. 22. The wafers used in this technique are subjected to an RCA clean followed by a 10 : 1 H_2O/HF dip for 10 seconds. These precleaned wafers are placed coaxially into a quartz boat and loaded into the load-lock chamber. The load-lock chamber is kept at approximately 100°C and is pumped down to a pressure below 10^{-6} torr. Once this pressure is reached, the wafers are introduced into the growth chamber together with flowing hydrogen. The growth chamber is kept at a constant temperature of 550°C and is pumped down to a base pressure of below 10^{-9} torr. The source gases are subsequently fed into the growth chamber prior to the desorption of hydrogen atoms from the HF-passivated surface. Typical source gases used for this system are silane (SiH_4) and germane (GeH_4) for Si and SiGe epi-layer growths. The dopant gases are diborane (B_2H_6) for p-type and phosphine (PH_3) for n-type. Nominally, the operating growth pressure is maintained at 10^{-3} torr as the growth process proceeds. As a result of the low-pressure process, the reactant residence time within the growth zone is greatly reduced. This technique makes an abrupt control of the Ge composition and doping profile possible.

A typical stepwise boron profile in epi-Si achieved by a UHV/CVD technique is shown in Fig. 23. Abrupt transitions between each doping step are clearly evident. In addition, this growth technique yields a growth rate of about 3 Å/min for intrin-

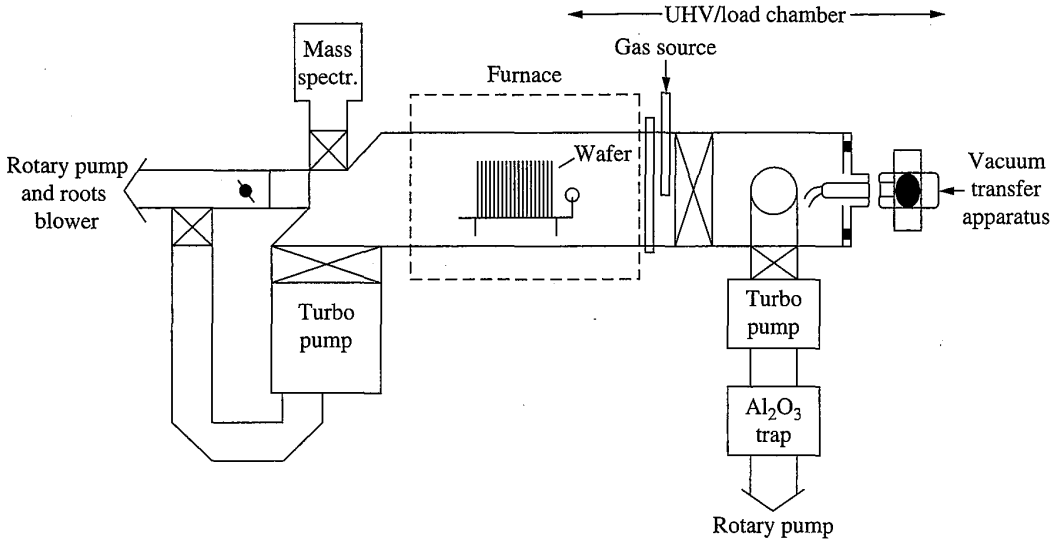


FIGURE 22
Schematic drawing of a UHV/CVD system. (*After Meyerson, Ref. 29.*)

Micron Ex. 1023, p. 155
Micron v. YMTC
IPR2024-00909

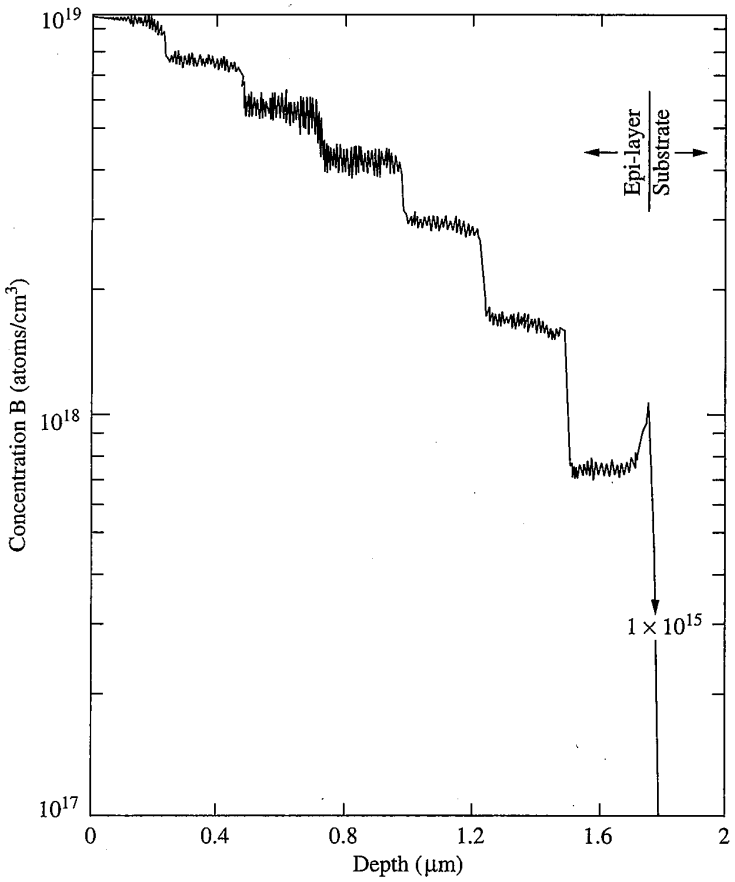


FIGURE 23
A typical stepwise boron profile in epi-Si achieved by a UHV/CVD technique. (*After Meyerson, Ref. 29.*)

sic or boron-doped silicon epi-layers. Therefore, it enables a precise control of layer thickness for advanced device applications. Since the growth of epi-layers of this technique is thermally activated, with an activation energy of about 1.6 eV, it can be increased rapidly at higher growth temperature. By using this growth technique, many novel heterostructures have been realized for Si-based materials.

MBE is a technique that was developed to fabricate devices that demand very tight controls over epi-layer thickness and doping uniformity. A conventional MBE technique, as shown in Fig. 24, is similar to that of an evaporation process.[43] For Si-MBE, the silicon and desired dopants are simply evaporated onto the Si substrate under ultrahigh-vacuum conditions. Typically, the base pressure of the MBE process is kept at 10^{-11} torr or higher. The evaporated atoms or molecules impinge on the heated substrate and then condense to form layers epitaxially. The Si-MBE process is usually performed in a temperature range of 500 to 900°C. Therefore, it has been, and remains, one of the most sophisticated low-temperature growth techniques for epi-Si. Because of its complexity and noncompatibility with device fabrication requirements, MBE has received little attention until very recently, when a low thermal budget has become more and more important for ULSI device fabrication.

Micron Ex. 1023, p. 156
Micron v. YMTC
IPR2024-00909

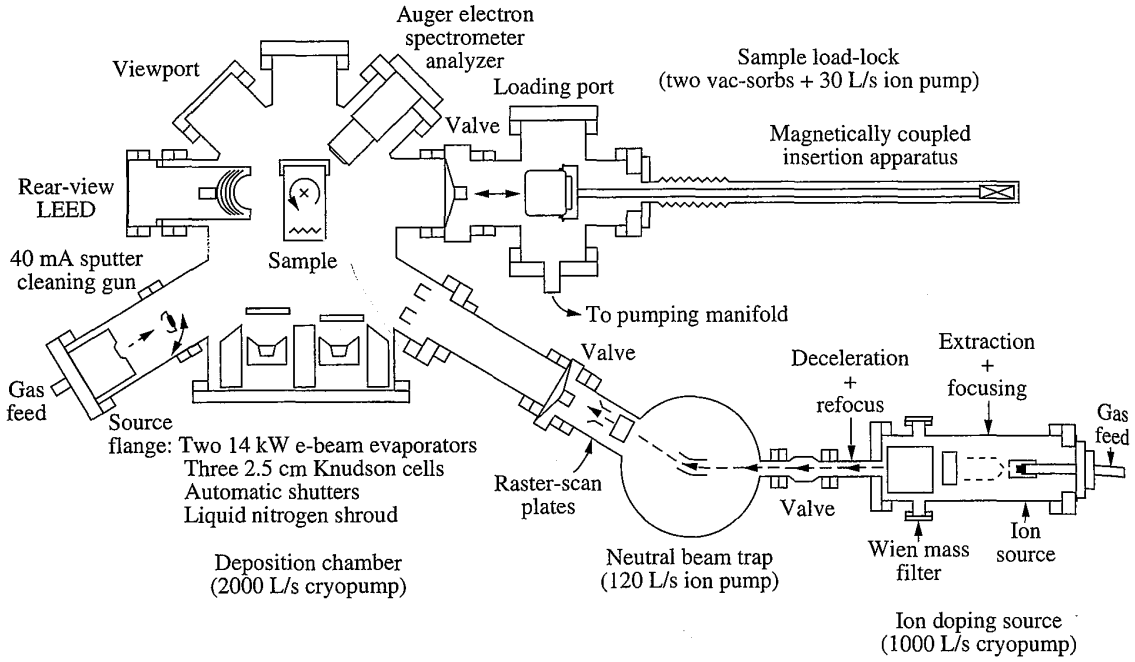


FIGURE 24
Schematic drawing of a typical MBE growth system. (*After Bean, Ref. 43.*)

Over the last two decades the MBE process has been improved significantly. However, its conventional electron-gun evaporator configuration still suffers from problems, such as clustering and spitting defects. To solve these problems, a modified MBE system, called gas-source MBE (GSMBE),[44] has been reported. This growth technique employs gaseous rather than solid sources to deposit epi-Si. As a result, spitting-type defects are eliminated. Conceptually, GSMBE is similar to UHV/CVD, except that the former is a cold-wall system whereas the latter is a hot-wall system. From the manufacturing point of view, MBE still suffers from low throughput. However, GSMBE opens a new approach toward the development of LTE for ULSI applicatio. s.

3.4.4 Device Applications

The LTE process has been employed to fabricate advanced bipolar transistors, MOSFETs, resonant tunneling diodes, infrared waveguide detectors, and other devices. Because the scope of this book is ULSI, only bipolar and MOSFET devices using an LTE process will be discussed. For readers who are interested in the other device applications of the LTE process, please refer to the references herein.

Both UHV/CVD and MBE techniques have been used to fabricate SiGe-based heterojunction bipolar transistors (HBTs). It is well known that the speed of a bipolar transistor can be improved by scaling the vertical profile and by reducing the parasitic resistance and capacitance of the extrinsic portion of the device. Theoretically, with a proper doping profile, a homojunction Si bipolar transistor can reach unity-current-gain cutoff frequencies (f_T) as high as 50 GHz. However, such a device

Micron Ex. 1023, p. 157
Micron v. YMTC
IPR2024-00909

would be practically of no use, because of the limitations imposed by base resistance, punch-through, and breakdown voltage. These limitations are mitigated with a heterojunction device. This concept has been implemented extensively for III–V compound semiconductor devices for more than a decade. However, only very recently, it was applied to the Si-based devices, because it is now possible to make device-quality Si epi-layers at low temperatures. In the case of the Si-based bipolar transistor, the heterojunction is formed by using SiGe as the base material.[45, 46]

Improved transport properties of electrons and holes have been demonstrated by Si/SiGe heterostructures grown with UHV/CVD and MBE techniques.[47, 48] N-channel and p-channel FETs of various types have been fabricated to exploit the advantages of Si/SiGe heterostructures for high-speed devices.[49, 50] Modulation-doped SiGe p-channel MOSFETs, utilizing the enhanced hole mobility characteristics of Si/SiGe heterostructures, have achieved a hole mobility of 220 cm^2/V-s at 300 K and 980 cm^2/V-s at 82 K in the SiGe channel, which is much higher than those of conventional p-channel MOSFETs. A similar study, which simply used high-quality UHV/CVD-grown SiGe layers to form p-channel MOSFETs, achieved a 70% higher transconductance at 300 K than conventional p-channel devices.[51] To avoid degradation of LTE-grown layers during device processing, a PECVD oxide was selected to form the gate at low temperatures. However, it was found that thermal and PECVD gate oxides gave comparable device characteristics. Therefore, the choice of a gate oxide formation process depends on the design of the MOSFET device. Even though enhanced electron mobility has been reported for Si/SiGe heterostructures grown by the UHV/CVD technique,[47] n-channel MOSFETs have been fabricated only by using MBE,[50] presumably because of the difficulties with n-type doping control using the UHV/CVD technique. Enhanced electron mobility has been reported by incorporating an n-type modulation-doped structure.

3.5 SELECTIVE EPITAXIAL GROWTH OF Si

Selective epitaxial growth (SEG) of Si is a process that allows the deposition of a Si epitaxial layer on a bare Si-substrate surface without the simultaneous growth of amorphous Si thin film on the silicon dioxide or silicon nitride surface. Therefore, the SEG process has attracted a great deal of interest for novel device applications.[52]

The SEG process was initially developed to achieve an advanced dielectric isolation structure. This approach is intended to replace the conventional local oxidation of Si (LOCOS) process, because of the scalability limits of LOCOS.[53] $SiCl_4$ has often been used for the SEG of Si. Both silicon dioxide and silicon nitride have been employed as the mask materials. The addition of HCl to the $SiCl_4$ is believed to prevent spurious nucleation and growth of silicon on the silicon dioxide mask surface; hence, it increases the selectivity of the growth process.[54] To a lesser extent, SiH_4 chemistry has also been used for the SEG process.[55]

SEG of Si was performed using a commerical cylindrical radiant-heating epitaxy reactor at a growth temperature below 1000°C using $SiCl_2H_2$ + HCl at a reduced pressure of 80 torr.[56] A lower growth temperature plus an increased HCl additive flow were found to lower the stacking fault density in the Si epi-layer. This

Micron Ex. 1023, p. 158
Micron v. YMTC
IPR2024-00909

work marked the major advance in SEG of Si. Since then, most of the SEG of Si has been performed at reduced temperatures and pressures. In 1985 a successful SEG-Si deposition down to 826°C, using SiH_2Cl_2 at 25 torr, was reported. Proper epitaxial precleaning and deposition techniques would prevent the oxide sidewall lifting (i.e., oxide undercutting) problem.[57] In 1989 the SEG-Si process was carried out with a growth temperature down to 650°C at an operating pressure less than 2 torr using SiH_2Cl_2. Over a temperature range of 650 to 1100°C no loading effect was observed.[58] This process was further performed at a growth temperature down to 600°C using the $SiH_4 + H_2$ chemistry at low deposition pressures. Even at a deposition pressure as low as 1 mtorr an ultraclean growth environment was required to achieve a successful SEG process.[37, 59] In 1991 a UHV/CVD system was employed to use the SEG process in novel device fabrication.[60]

A novel approach, called epitaxial lateral overgrowth (ELO), using the SEG-Si process, was invented to obtain silicon-on-insulator (SOI) substrates.[61] As IC technology advances, the quality of SEG-Si layers becomes crucial for device applications. The use of a sacrificial oxide was found effective for removing the contamination and damage on the Si substrate, greatly improving the SEG-Si surface quality. The defects at the interface between the silicon dioxide and the SEG-Si layer could be minimized when the mask edges were oriented in a [100] direction.[62]

3.5.1 Fundamental Aspects

Strictly speaking, SEG-Si symbolizes a process that deposits Si only on the bare Si substrate, with no Si in any form deposited on the silicon dioxide or silicon nitride mask. However, in practice, the SEG-Si process can be divided into two types. *Type 1* is the standard SEG-Si process, in which Si epitaxy occurs only on the bare Si substrate surface within the openings in the dielectric mask film, with no Si deposition whatsoever on the dielectric films. *Type 2* SEG-Si is a process that allows simultaneous deposition of a Si epitaxial layer on the single-crystal Si substrate within the openings in the dielectric film and deposition of polycrystalline Si (poly-Si) on the dielectric mask surface.[63]

Type 1 SEG-Si is achieved when Si atoms possessing high surface mobility are deposited from the silicon sources. Si atoms are able to migrate to favored sites on the single-crystal substrate, where nucleation occurs. It has been found that the surface mobility of Si atoms is affected by the presence of halides in the silicon sources. The degree of selectivity increases as the number of chlorine atoms in the silicon sources increases. Accordingly, $SiCl_4$ becomes the silicon source of choice for Type 1 selective deposition. However, as shown in Fig. 11, $SiCl_4$ requires the highest reduction temperature among all silicon sources; thus, the use of this source gas has been limited because of the degradation of the silicon dioxide mask at high temperatures. As an alternative, the silicon nitride mask, which can sustain higher temperature than the silicon dioxide mask, was employed for Type 1 SEG-Si process using $SiCl_4$ source gas. However, silicon nitride is a better nucleation source for silicon deposition than silicon dioxide is, and its use results in poor selectivity. As a result, $SiHCl_3$ becomes the optimum silicon source of choice to perform

Micron Ex. 1023, p. 159
Micron v. YMTC
IPR2024-00909

Type 1 SEG-Si, simply because of its high chlorine content and lower reduction temperature compared to $SiCl_4$. SiH_2Cl_2 and SiH_4 have also been used frequently as the source gases for Type 1 SEG-Si process. In these cases, HCl or Cl_2 has been added to the growth system to increase selectivity. The presence of crystallographic defects, such as microtwins and stacking faults, is commonly observed at epi/dielectic mask interface. It has been reported that these defects are greatly reduced, because the wall of the silicon dioxide step is coated with poly-Si prior to the SEG deposition.

In contrast, the Type 2 SEG-Si process can be achieved more easily by using SiH_4, since selectivity is not strictly required and a low deposition temperature is needed. This process results in an interface between the epitaxial layer on the substrate surface and polysilicon on the dielectric mask film. The angle of this interface, relative to the film-growth direction, depends on the crystallographic orientation of the substrate. For example, this angle would be 90 degrees for (110), 72 degrees toward poly-Si for (100), and 70 degrees tapered toward the single-crystal silicon for (111)-oriented substrates.

Many factors affect the selective nature of Si deposition, including growth pressure, growth temperature, HCl concentration, silicon source, Si substrate surface condition, dielectric opening size, and the surface ratio of bare Si to oxide (dielectric film). Applications of the SEG-Si process are discussed in the next section.

3.5.2 Applications

One of the most important applications of the SEG process is for device isolation at submicron levels. The conventional LOCOS isolation process has a problem known as "bird's beak encroachment." Therefore, the scalability of the LOCOS is limited to about the μm range. To increase the device integration level, different isolation techniques are required. A typical SEG process for device isolation[57] is shown in Fig. 25. This approach offers many advantages, such as near-zero lateral encroachment, deep submicron-device separation, excellent planarity, and excellent latch-up immunity. In addition, it allows the formation of independent n and p wells, making retrograde wells possible by buried-layer epitaxy or graded epitaxy techniques.[57] For CMOS devices, the SEG process can be used to tailor the substrate and well resistance, increasing the latch-up immunity. Recently, low-temperature epitaxy was used in the SEG process, resulting in a sharp epi/substrate transition region due to minimal dopant outdiffusion and autodoping.

Another application of the SEG process is epitaxial lateral overgrowth (ELO), which is one way to form silicon-on-insulator (SOI) structures.[64] A schematic illustration of ELO process is shown in Fig. 26. In this process the lateral-to-vertical growth ratio depends on the step height and the substrate orientation.[65] It has been found that for the same substrate orientation, the lateral-to-vertical growth ratio decreases as the step height increases. For the same step height the lateral-to-vertical growth ratio decreases according to orientation, (100) > (110) > (111). To avoid facet formation, both lateral and vertical growth should be in the [100] directions.[52]

Device applications of the SEG process have been reported by many researchers. It has been used to develop an epitaxy-over-trench (EOT) process for dynamic

Micron Ex. 1023, p. 160
Micron v. YMTC
IPR2024-00909

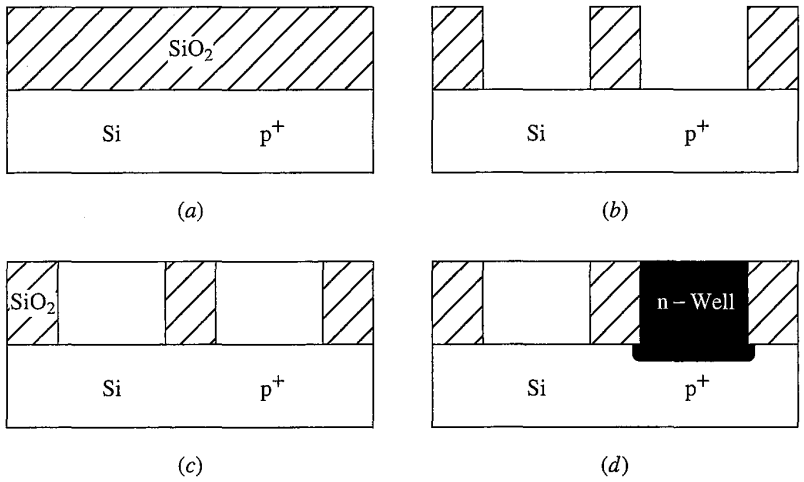


FIGURE 25
Schematic illustration of a typical SEG process for device isolation: (*a*) oxide deposition; (*b*) window formation; (*c*) epi growth; (*d*) n-well drive-in. (*After Borland and Drowley, Ref. 57.*)

random-access memory (DRAM) technology.[66] This approach permits the transfer transistor to be fabricated directly over the storage capacitor, resulting in a high-density DRAM. For bipolar technology a method using the SEG-Si process to form epitaxial bases has been proposed. Such a selective-epitaxy-base transistor (SEBT) allows the scaling of bipolar devices to less than that obtainable with the ion implantation technique to produce a narrow base width in the deep submicron regime.[67] In addition, a selective-epitaxy emitter-window (SEEW) process has been used. In this case, the SEG-Si is deposited with high boron concentration to define the geometry of the emitter window opening and to provide electrical contact to the extrinsic base of the transistor.[60]

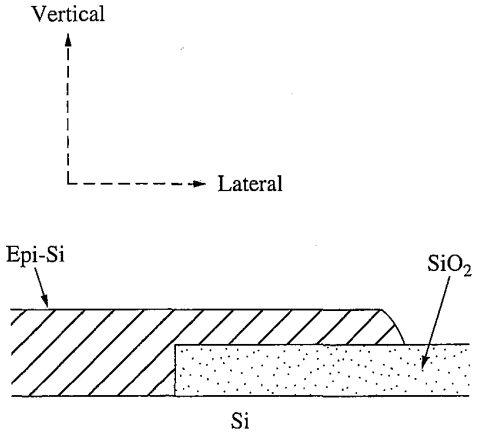


FIGURE 26
Schematic illustration of an ELO process.

Micron Ex. 1023, p. 161
Micron v. YMTC
IPR2024-00909

3.6 CHARACTERIZATION OF EPITAXIAL FILMS

To develop and then implement an epitaxial process would not be successful without employing proper characterization techniques. In this section, we will review some of the measurements that are essential to evaluate epitaxial films. By nature, these techniques can be used to investigate the physical, chemical, and electrical properties of an epi-layer.

3.6.1 Physical Properties

The physical properties of an epitaxial layer include surface morphology, crystallographic defects, stress, and layer thickness.

Surface morphology is normally examined by eyes under high-intensity illumination or microscopically by conventional optical microscopy (in the brightfield or darkfield mode), Nomarski interference-contrast microscopy, and scanning electron microscopy (SEM). The selection of a proper technique depends solely on the need for each application. In an industrial environment, human eyes with high-intensity illumination are frequently used, because the process is simple and has a high screening speed. When a detailed surface morphology is required, microscopic methods of various modes are employed. In the brightfield mode of optical microscopy the image is formed by the reflection of the light received by the sample; in the darkfield mode only light that is reflected by the features on the wafer surface is collected. The former gives the best overall image and information about the surface and is, therefore, commonly used. In contrast, the latter is very powerful in resolving small structural features on the surface and is often used to scan a wide range of defects (including crystallographic defects after etching) on the surface. Using a splitting polarized light source, Nomarski interference-contrast microscopy is capable of resolving surface features of 3 to 5 nm in depth, far beyond what can be achieved by ordinary optical microscopy. SEM operates by scanning a focused electron beam over a surface and sensing the secondary electrons emitted from the surface. This technique has a large depth of field and high resolution and is capable of resolving submicron surface features at a magnification much higher than that obtainable with any optical microscopy.[68]

An epitaxial layer may contain a variety of crystallographic defects. In an as-deposited epi-layer, these defects are not always readily visible. Etchants are used to delineate defects preferentially prior to microscopic examination. For Si epitaxy, Wright etch, a mixture of HF, HNO_3, CrO_3, $Cu(NO_3)_2$, acetic acid, and deionized water, is the most frequently used etchant. Stacking faults in an epi-layer can be revealed with a 5-min Wright etch. Other solutions, such as the Secco or Sirtl etches, are also often used to delineate dislocations and pits.[69]

When a thin epi-layer is deposited on a substrate of another material, it is under either tensile or compressive stress, causing the substrate to deflect downward or upward, respectively. A variety of techniques may be used to measure the wafer

Micron Ex. 1023, p. 162
Micron v. YMTC
IPR2024-00909

center deflection and, hence, the stress in the epi-layer. The simplest method is to place the wafer on the stage of an optical microscope and calibrate the vertical adjustment for the in-focus vertical positions at the wafer edge and the wafer center.[70] A commercial wafer deflection gauge, using the degree of light reflection, is also used to evaluate the amount of wafer deflection.

Depending on the nature of a film, a measurement of the film's thickness can be fairly simple or quite complex. In general, a thickness measurement technique can be either destructive or nondestructive. Although there are many thickness measurement techniques, only a few, especially nondestructive ones, are preferable for Si epitaxial layers. Destructive methods include *angle-lap and stain* and *groove and stain.*[71] These techniques can be used only for an epi-layer that has a different doping type or doping concentration than the substrate. For example, the cross section of the p-n or n-n^+ junction is differentiated by chemical etching (staining) and then examined with optical microscopy or SEM to measure the epi-layer thickness.

The most widely used nondestructive epi-layer thickness measurement techniques are based on optical properties of the Si epi-layer. In the near infrared region (2.5 to 50 μm), lightly doped Si is relatively transparent, whereas heavily doped Si behaves as a reflective surface. Therefore, when an electromagnetic radiation with various wavelengths is applied to an epi-layer on a Si substrate, interference and optical patterns are formed because of the interaction between the reflections from the air/epi-layer surface and epi-layer/substrate interface. Based on these optical interference patterns, the epi-layer thickness can be obtained. The infrared (IR) interference spectrophotometer is, therefore, often used as a quick and nondestructive thickness measurement technique for Si epi-layers.[72] A Fourier-transform infrared (FTIR) spectroscopy procedure using a Michelson interferometer has been developed to measure Si epi-layer thickness. This technique is capable of a rapid ($\approx$30 sec), accurate ($\pm$0.05 μm), and precise ($\pm$0.005 μm) measurement and, therefore, is widely accepted.[73] More recently, a high-resolution double-crystal x-ray diffraction (HRXRD) method[74] has been employed to measure the layer thickness of Si/SiGe multilayer structures with an accuracy within $\pm$2%.

3.6.2 Chemical and Electrical Properties

The chemical and electrical properties of an epi-layer are usually closely related. The resistivity of an epi-layer is proportional to the activated dopant concentration. When a dopant is incorporated into an epi-layer, it can be either electrically active or nonactive, depending on the lattice site. Electrically active dopant concentration can be determined only by electrical measurement techniques, such as the four-point probe, capacitance-voltage (C-V), and spreading resistance. On the other hand, total dopant concentrations, including electrically active and nonactive species, are usually measured by surface spectroscopic techniques, such as secondary-ion mass spectroscopy (SIMS), Auger electron spectroscopy (AES), electron spectroscopy for chemical analysis (ESCA), and Rutherford backscattering spectroscopy (RBS).

The four-point probe is one of the fastest techniques to measure the dopant concentration.[75] However, it is applicable only to an epi-layer that is grown on a

Micron Ex. 1023, p. 163
Micron v. YMTC
IPR2024-00909

lightly doped substrate of an opposite conductivity type to the layer. In this technique, the resistivity ρ of the epi-layer is expressed by

$$\rho = R_s TF \tag{3.27}$$

where R_s is the sheet resistivity, T is the epi-layer thickness, and F is the geometric correction factor. Low current is preferred, to avoid ohmic heating or reaching punch-through voltage. When a thin layer is measured, care must be taken to avoid the tips penetrating the layer. Overall, this technique can achieve an accuracy of only $\pm 10\%$ and is not capable of measuring doping profiles.

The C-V technique is performed by using a mercury contact to form a Schottky barrier diode with the epi-layer.[76] This is a rapid, nondestructive technique to measure the doping profile in the layer. By applying a reverse bias, the capacitance, as a function of voltage, is recorded, and the doping profile is obtained according to the following equation:

$$N(x) = -C^3[q\epsilon A^2(dC/dV)]$$

and

$$x = \epsilon A/C(V) \tag{3.28}$$

where C is the capacitance, q is the electric charge, ϵ is the permittivity of silicon, A is the junction area, x is the distance from the junction, and V is the bias voltage. This technique works well with doping in the range of 10^{14} to $10^{17}/\text{cm}^3$.

For higher doping concentrations the spreading resistance technique is used. This technique employs a cleaved and beveled sample. A two-point probe is used to measure the resistivities sequentially at a number of points, starting from the epi-layer surface through the layer/substrate interface.[77] By correlating the change of resistivity between each point with carrier concentration, the doping concentration and its profile are deduced. This technique is capable of measuring a wide range of doping concentrations (10^{14} to $10^{20}/\text{cm}^3$). However, it suffers from poor accuracy and complex data reduction. Besides, special care must be taken to compensate for the effect of the wear of the probe tip, which significantly affects the accuracy of this measurement technique.

In contrast to these electrical measurements, surface spectroscopic techniques are used when a total doping profile is needed. All surface spectroscopic techniques are carried out under high-vacuum conditions. Therefore, they are not useful for routine process control. However, as Si technology becomes more and more miniaturized, these techniques become more and more important because of their excellent depth-profiling capability.

Secondary-ion mass spectroscopy (SIMS) uses ions, that are accelerated by high voltage (500 to 5000 eV), as the probe beam to sputter ions, called secondary ions, from the surface of interest.[78] These secondary ions are then analyzed with a mass spectrometer to characterize the elemental constitutes of the surface. Both oxygen and cesium ion beams are used for SIMS measurements. The former is more effective for electropositive elements such as B, Al, and Cr, and the latter for electronegative elements such as C, O, and As. Since this technique involves sputtering of the surface, depth profiling can be obtained readily. Depth resolution in the 2 to 5 nm

Micron Ex. 1023, p. 164
Micron v. YMTC
IPR2024-00909

range can be achieved routinely. Layers of ≈1 μm thickness can be depth-profiled with this technique. For Si with doping levels down to 10^{15} cm^{-3}, SIMS is one of the most powerful techniques for measuring the total doping profile.

Auger electron spectroscopy (AES) employs an energetic electron beam (up to 10 keV) to probe the surface.[79] The energetic electron can ionize an atom by dislodging an inner-core electron. When an outer-core electron falls back to the inner core to replace the ejected electron, the atom gives up its excess energy by emitting an x-ray. Alternatively, it can eject a second, "Auger" electron. The energy of this Auger electron is characterized by the atom from which it comes, allowing the elemental composition to be measured. Although the electron beam can penetrate many hundreds of angstroms into the surface, only Auger electrons that escape from the top 1 to 3 nm of the surface can be collected. When combined with a sputtering process, AES is also capable of depth profiling.

Electron spectroscopy for chemical analysis (ESCA) is also known as x-ray photoelectron spectroscopy (XPS). This technique uses a beam of low-energy x-rays to bombard a sample.[80] The x-ray photon energy, when it is absorbed, removes an outer-shell electron from an atom. This electron is then emitted with a kinetic energy of a magnitude equal to the difference between the x-ray and the bonding energy of the electron. The energy of the emitted electron is analyzed to determine the type of atom, and the number of electrons at this energy is used to obtain the density of such atoms in the layer. Like AES, this technique is truly a surface analysis technique (1 to 3 nm from the surface). Since an x-ray beam is used, ESCA allows rapid depth profiling. In addition, it is often used to obtain information about chemical bonding around an atom, which is not achievable with other surface analysis techniques.

Rutherford backscattering spectroscopy (RBS) employs high-energy He^{2+} ions (≥ 1 MeV) to penetrate the sample surface to obtain the depth profile in the layer.[81] This technique is nondestructive and is based on the elastic collision theory to obtain the chemical information in a layer. When He^{2+} ions impinge on a surface, penetrate to the subsurface, and collide with atoms, some ions are backscattered with an energy loss that is characteristic of the atom struck and the distance from the sample surface where the collision occurs. Therefore, both the elemental composition and its depth profile can be obtained simultaneously. RBS is a surface-analysis technique that gives quantitative chemical information without using standards. However, it has relatively poor depth resolution (≈2 nm) compared to other techniques.

3.7 SUMMARY AND FUTURE TRENDS

Epitaxy will remain as a unique process that provides doping capability and flexibility for IC technology. As the industry moves into the ULSI era, LTE will play an important role in preparing the active parts of IC devices, such as the intrinsic base for a bipolar transistor. In addition, low temperature is required to minimize autodoping effects, which become more and more critical for ULSI circuits. In the near future, LTE will undoubtedly receive most attention in the field of epitaxy. Among all LTE techniques, UHV/CVD seems to be most promising for being incorporated into ULSI manufacturing. However, APCVD also shows a high potential for this appli-

Micron Ex. 1023, p. 165
Micron v. YMTC
IPR2024-00909

cation. In any case, an ultraclean environment is inevitably required for any epitaxial processes in the future. Because of the limitation imposed by the Si material and Si homojunction, the development of Si/SiGe heteroepitaxy will receive a great deal of attention in the foreseeable future. Defect minimization remains as one of the major topics in the pursuit of heteroepitaxy. In terms of growth chemistry, SiH_4 will be the preferred chemical in the future, because LTE will be a dominant technology. Finally, more sophisticated instrumentation, such as the FTIR and SIMS apparatus, will be widely used for characterizing physical, chemical, and electrical properties of an epi-layer to meet the stringent requirements of ULSI circuits.

REFERENCES

1. J. W. Matthews, "Defects Associated with the Accommodation of Misfit Between Crystals," *J. Vac. Sci. Technol.,* **12,** 126 (1975).
2. R. People and J. C. Bean, "Calculation of Critical Layer Thickness versus Lattice Mismatch for GeSi/Si Strained-Layer Heterostructures," *Appl. Phys. Lett.,* **47,** 322 (1985).
3. S. K. Ghandhi, *VLSI Fabrication Principles: Silicon and Gallium Arsenide,* Wiley, New York, 1983.
4. S. M. Sze, *Physics of Semiconductor Devices,* 2nd ed., Wiley, New York, 1981; C. J. Hu et al., "A Self-Aligned 1 μm CMOS Technology for VLSI," *Tech. Dig. IEDM,* 731 (1983).
5. D. W. Shaw, "Fundamental Aspects of Epitaxy," in *Crystal Growth, Vol.1,* C. H. L. Goodman, Ed., Plenum, New York (1974).
6. R. E. Reed-Hill, *Physical Metallurgy Principles,* Van Nostrand, New York, 1973.
7. A. S. Grove, "Mass Transfer in Semiconductor Technology," *Ind. Eng. Chem.,* **58,** 48 (1966).
8. V. S. Ban, "Mass Spectrometric Studies of Chemical Reactions and Transport Phenomena in Silicon Epitaxy," *Proceedings of the 6th International Conference on Chemical Vapor Deposition,* The Electrochemical Society 1977, p. 66.
9. J. Bloem and L. J. Gilling, "Epitaxial Growth by Chemical Vapor Deposition," in *VLSI Electronics,* N. G. Einspruch and H. Huff, Eds., *Vol. 12,* 89, Academic Press, Orlando, Florida, 1985.
10. J. Bloem, "Nucleation and Growth of Silicon by CVD," *J. Cryst. Growth,* **50,** 581 (1980).
11. F. C. Eversteyn, "Chemical-Reaction Engineering in the Semiconductor Industry," *Philips Res. Rept.,* **29,** 45 (1974).
12. J. Nishizawa and M. Saito, "Growth Mechanism of Chemical Vapor Deposition of Silicon," *Proceedings of the 8th International Conference on Chemical Vapor Deposition,* The Electrochemical Society 1981, p. 317.
13. V. S. Ban and S. L. Gilbert, "Chemical Processes in Vapor Deposition of Silicon," *J. Electrochem. Soc.,* **122,** 1382 (1975).
14. E. Sirtl, L. P. Hunt, and D. H. Sawyer, "High Temperature Reactions in Silicon-Hydrogen-Chlorine System," *J. Electrochem. Soc.,* **121,** 919 (1974).
15. M. L. Yu, D. J. Vitkavage, and B. S. Meyerson, "Doping Reaction of PH_3 and B_2H_6 with Si(100)," *J. Appl. Phys.,* **59,** 4032 (1986).
16. R. Reif, T. I. Kamins, and K. C. Saraswat, "A Model for Dopant Incorporation into Growing Silicon Epitaxial Films," *J. Electrochem. Soc.,* **126,** 644 (1979).
17. G. R. Srinivasan, "Autodoping Effects in Silicon Epitaxy," J. Electrochem. Soc., **127,** 1334 (1980).

Micron Ex. 1023, p. 166
Micron v. YMTC
IPR2024-00909

18. G. R. Srinivasan, "Kinetics of Lateral Autodoping in Silicon Epitaxy," *J. Electrochem. Soc.,* **125,** 146 (1978).
19. K. V. Ravi, *Imperfection and Impurities in Semiconductor Silicon,* Wiley, New York, 1981.
20. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era, Vol.1,* Lattice Press, California, 1986, p. 140.
21. B. J. Baliga, "Defect Control During Silicon Epitaxial Growth Using Dichlorosilane," *J. Electrochem. Soc.,* **129,** 1078 (1982).
22. W. Dyson, L. G. Hellwig, J. M. Moody, and J. A. Rossi, "n^+ and p^+ Substrate Effects on Epitaxial Silicon Properties," in *VLSI Science and Technology,* K. E. Bean and G. Rozgonyi, Eds., The Electrochemical Society, 1984, p. 107.
23. Nalih, A. S., Rozgonyi, R. F. Davis, and H. J. Kim, "Extrinsic Gettering with Epitaxial Misfit Dislocations," *Semiconductor Processing,* ASTM, STP 850, Gupta (1984).
24. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era, Vol.1,* Lattice Press, California, 1986, p. 144.
25. P. H. Lee, M. T. Wauk, R. S. Rosler, and W. C. Benzing, "Epitaxial Pattern Shift Comparisons in Vertical, Horizontal, and Cylindrical Reactor Geometries," *J. Electrochem. Soc.,* **124,** 1824 (1977).
26. S. P. Weeks, "Pattern Shift and Pattern Distortion During CVD Epitaxy on <111> and <100> Silicon," *Sol. State Technol.,* **24,** 111 (1981).
27. C. W. Pearce, "Epitaxy," *VLSI Technology,* 2nd edn., S. M. Sze, Ed., McGraw-Hill, 1988, p. 51.
28. M. L. Hammond, "Silicon Epitaxy," *Sol. State Technol.* **21,** 68 (1978).
29. B. S. Meyerson, "Low Temperature Si and Si:Ge Epitaxy by Ultrahigh Vacuum/Chemical Vapor Deposition: Process Fundamentals," *IBM J. Res. Devel.,* **34,** 806 (1990); T. G. Jung, C. Y. Chang, T. C. Chang, H. C. Lin, T. Wang, W. C. Tsai, G. W. Huang, and P. J. Wang, "Low-temperature Epitaxial Growth of Silicon and Silicon-Germanium Alloy by Ultrahigh-Vacuum Chemical Vapor Deposition," *Jpn. J. Appl. Phys.,* **33,** 224 (1994).
30. B. S. Meyerson, K. J. Uram, and F. K. LeGoues, "Cooperative Growth Phenomena in Silicon/Germanium Low Temperature Epitaxy," *Appl. Phys. Lett.,* **53,** 2555 (1988).
31. W. Kern and D. A. Puotinen, "Cleaning Solutions Based on Hydrogen Peroxide for Use in Silicon Semiconductor Technology," *RCA Rev.,* **31,** 187 (1970).
32. B. S. Meyerson, F. J. Himpsel, and K. J. Uram, "Bistable Conditions for Low-temperature Silicon Epitaxy," *Appl. Phys. Lett.,* **57,** 1034 (1990).
33. T. O. Sedgwick and P. D. Agnello, "Atmospheric Pressure Chemical Vapor Deposition of Si and SiGe at Low Temperatures," *J. Vac. Sci. Technol.,* **A10,** 1913 (1992).
34. D. J. Eaglesham, G. S. Higashi, and M. Cerullo, "370°C Clean for Si Molecular Beam Epitaxy Using a HF Dip," *Appl. Phys. Lett.,* **59,** 685 (1991).
35. J. C. Bean, "Silicon Molecular Beam Epitaxy: 1984–1986," *J. Cryst. Growth,* **81,** 411 (1987).
36. E. D. Ahlers, R. M. Ostrom, P. Datta, and F. G. Allen, "Crystal Quality of Solid Phase Epitaxially Grown Silicon," *J. Vac. Sci. Technol. B,* **6,** 723 (1988).
37. J. Murota, N. Nakamura, M. Kato, N. Mikoshiba, and T. Ohmi, "Low-Temperature Silicon Selective Deposition and Epitaxy on Silicon Using the Thermal Decomposition of Silane Under Ultraclean Environment," *Appl. Phys. Lett.,* **54,** 1007 (1989).
38. T. Y. Hsieh, K. H. Jung, D. L. Kwong, and S. K. Lee, "Silicon Homoepitaxy by Rapid Thermal Processing Chemical Vapor Deposition (RPTCVD)—A Review," *J. Electrochem. Soc.,* **138,** 1188 (1991).
39. C. M. Gronet, et al., "Thin, Highly Doped Layers of Epitaxial Silicon Deposited by Limited Reaction Processing," *Appl. Phys. Lett.,* **48,** 1012 (1986).

Micron Ex. 1023, p. 167
Micron v. YMTC
IPR2024-00909

40. K. Baert, P. Deschepper, J. Poortmans, J. Nijs, and R. Mertens, "Selective Si Epitaxial Growth by Plasma-Enhanced Chemical Vapor Deposition at Very Low Temperature," *Appl. Phys. Lett.,* **60,** 442 (1992).
41. A. Yamada, Y. Jia, M. Kohagai, and K. Takahashi, "Heavily P-doped Si and SiGe Films Grown by Photo-CVD at 250°C," *J. Electron. Matls.,* **19,** 1083 (1990).
42. A. Yamada, A. Satoh, M. Konagai, and K. Takahashi, "Low Temperature (600–650°C) Silicon Epitaxy by Excimer Laser-Assisted Chemical Vapor Deposition," *J. Appl. Phys.,* **65,** 4268 (1989).
43. J. C. Bean, "Silicon Molecular Beam Epitaxy as a VLSI Processing Technique," *IEEE Proc. IEDM,* 1981, p. 6.
44. H. Hirayama, T. Tatsumi, and N. Aizaki, "Gas Source Molecular Beam Epitaxy Using Disilane," *Appl. Phys. Lett.,* **52,** 1484 (1988).
45. G. L. Patton, J. H. Comfort, B. S. Meyerson, E. F. Crabbé, G. J. Scilla, Z. De Frésart, J. M. C. Stork, J. Y. C. Sun, D. L. Harame, and J. N. Burghartz, "75-GHz f_T SiGe-Base Heterojunction Bipolar Transistors," *IEEE Electron Dev. Lett.,* **11,** 171 (1990); G. L. Patton, S. S. Iyer, S. L. Delage, S. Tiwari, and J. M. C. Stork, "Silicon-Germanium-Base Heterojunction Bipolar Transistors by Molecular Beam Epitaxy," *IEEE Electron Dev. Lett.,* **9,** 165 (1988).
46. E. F. Crabbé, B. S. Meyerson, J. M. C. Stork, and D. L. Harame, "113-GHz f_T Graded Base SiGe HBT's" 1993 Device Research Conference, Paper #IIA-3, June, 1993.
47. K. Ismail, B. S. Meyerson, and P. J. Wang, "High Electron Mobility in Modulation-doped Si/SiGe," *Appl. Phys. Lett.,* **58,** 2117 (1991).
48. P. J. Wang, B. S. Meyerson, F. F. Fang, J. Nocera, and B. Parker, "High Hole Mobility in Si/SiGe/Si P-type Modulation-Doped Double Heterostructures," *Appl. Phys. Lett.,* **55,** 2333 (1989).
49. S. Verdonckt-Vandebroek, E. F. Crabbé, B. S. Meyerson, D. L. Harame, P. J. Restle, J. M. Stork, A. C. Megdanis, C. L. Stanis, A. A. Bright, G. M. Kroesen, and A. C. Warren, "High-Mobility Modulation-Doped Graded SiGe-Channel P-MOSFET's," *IEEE Electron Dev. Lett.,* **12,** 447 (1991).
50. H. Daembkes, H. J. Herzog, H. Jorke, H. Kibbel, and E. Kaspar, "The N-Channel SiGe/Si Modulation-Doped Field-Effect Transistor," *IEEE Trans. Electron. Dev.,* **ED-33,** 633 (1986).
51. S. Subbanna, V. P. Kesan, M. J. Tejwani, P. J. Restle, D. J. Mis, and S. S. Iyer, "Si/SiGe P-channel MOSFET's," *Symposium on VLSI Technology, Japan,* 1991, p. 103.
52. B. J. Ginsberg, J. Burghartz, G. B. Bronner, and S. R. Moder, "Selective Epitaxial Growth of Silicon and Some Potential Applications," *IBM J. Res. Devel.,* **34,** 816 (1990).
53. A. Ishitani, H. Kitajima, K. Tanno, and H. Tsuya, "Selective Silicon Epitaxial Growth for Device-Isolation Technology," *Microelectron. Eng.,* **4,** 3 (1986).
54. D. M. Jackson, "Advanced Epitaxial Process for Monolithic Integrated-Circuit Application," *Trans. Met. Soc. AIME,* **233,** 596 (1965).
55. P. Rai-Choudhury and D. K. Schroder, "Selective Silicon Epitaxy and Orientation Dependence of Growth," *J. Electrochem. Soc.,* **120,** 664 (1973).
56. K. Tanno, N. Endo, H. Kitajima, Y. Kurogi, and H. Tsuya, "Selective Silicon Epitaxy Using Reduced Pressure Technique," *Jpn. J. Appl. Phys.,* **21,** L564 (1982).
57. J. O. Borland and C. I. Drowley, "Advanced Dielectric Isolation through Selective Epitaxial Growth Techniques," *Sol. State Technol.,* **28,** 141 (1985).
58. J. L. Regolini, D. Bensahd, E. Scheid, and J. Mercier, "Selective Epitaxial Silicon Growth in the 650–1100°C Range in a Reduced Pressure Chemical Vapor Deposition Reactor Using Dichlorosilane," *Appl. Phys. Lett.,* **54,** 658 (1989).

Micron Ex. 1023, p. 168
Micron v. YMTC
IPR2024-00909

59. T. R. Yew and R. Reif, "Selective Silicon Epitaxial Growth at 800°C by Ultralow-Pressure Chemical Vapor Deposition Using SiH_4 and SiH_4/H_2," *J. Appl. Phys.*, **15,** 2509 (1989).
60. J. N. Burghartz, S. R. Mader, B. S. Meyerson, B. J. Ginsberg, J. M. Stork, C. Stanis, and Y. C. Sun, "Self-Aligned Bipolar NPN Transistor with 60nm Epitaxial Base," *Technical Digest, IEEE International Electron Device Meeting,* 229 (1989).
61. L. Jastrzebski, J. F. Corboy, and R. Pagliaro, Jr., "Growth of Electronic Quality Silicon Over SiO_2 by Epitaxial Lateral Overgrowth Technique," *J. Electrochem. Soc.,* **129,** 2645 (1982).
62. J. T. McGinn, L. Jastrzebski, and F. Corboy, "Defect Characterization in Monocrystalline Silicon Grown over SiO_2," *J. Electrochem. Soc.,* **131,** 398 (1984).
63. S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era, Vol.1,* 155, Lattice Press, California (1986).
64. J. Borland, M. Gangani, R. Wise, S. Fong, Y. Oka, and Y. Matsumoto, "Silicon Epitaxial Growth for Advanced Device Structures," *Sol. State Technol.,* **31,** 111 (1988).
65. D. D. Rathman, D. J. Silversmith, and J. A. Burns, "Lateral Epitaxial Overgrowth of Silicon on SiO_2," *J. Electrochem. Soc.,* **129,** 2303 (1982).
66. G. B. Bronner, N. C. C. Lu, T. V. Rajeevakumar, B. Ginsberg, and B. Machesney, "Epitaxy over Trench for ULSI DRAMs," *Technical Digest, IEEE Symposium on VLSI Technology,* 1988, p. 21.
67. J. N. Burghartz, B. J. Ginsberg, S. R. Mader, T. C. Chen, and D. L. Harame, "Selective Epitaxy Base Transistor (SEBT)," *IEEE Electron Dev. Lett.,* **9,** 259 (1988).
68. P. R. Thronton, *Scanning Electron Microscopy,* Chapman and Hall, London, 1968.
69. D. G. Schimmel, "A Comparison of Chemical Etches for Revealing ⟨100⟩ Silicon Crystal Defects," *J. Electrochem. Soc.,* **123,** 734 (1976).
70. R. Glang, R. Holmwood, and R. Rosenfeld, "Determination of Stress in Films on Single Crystalline Silicon Substrates," *Rev. Sci. Inst.,* **36,** 7 (1965).
71. "Standard Test Method for Thickness of Epitaxial and Diffused Layers in Silicon by the Angle Lapping and Staining Technique," *1984 Annual Book of ASTM Standards,* F110-84, Vol.10.05, 1984, p. 230.
72. "Thickness of Epitaxial Layers of Silicon on Substrates of Same Type by Infrared Reflectance," *1984 Annual Book of ASTM Standards,* F95, Vol.10.05, 1984, p. 213.
73. A. L. Smith, *Applied Infrared Spectroscopy,* Wiley, New York, 1979.
74. P. J. Wang, M. S. Goorsky, B. S. Meyerson, F. K. Le Goues, and M. J. Tejwani, "Characterization of Si/SiGe Strained-Layer Superlattices Grown by Ultrahigh Vacuum/Chemical Vapor Deposition Technique," *Appl. Phys. Lett.,* **59,** 814 (1991); T. C. Chang, C. Y. Chang, T. G. Jung, W. C. Tsai, P. J. Wang, T. L. Lee, and L. J. Chen, "Nanometer Thick Si/SiGe Strained-Layer Superlattices Grown by an Ultrahigh Vacuum/Chemical Vapor Deposition Technique," *J. Appl. Phys.,* **75,** 1 (1994).
75. Am. Soc. Test. Mater., ASTM Standard, F374, Part 43.
76. D. L. Rehrig and C. W. Pearce, "Production Mercury Probe Capacitance-Voltage Testing," *Semic. Intl.,* **3,** 151 (1972).
77. Y. Isda, H. Abe, and M. Knodo, "Impurity Profile Measurement of Thin Epitaxial Wafers by Multilayer Spreading Resistance Analysis," *J. Electrochem. Soc.,* **124,** 1118 (1977).
78. J. A. Mchugh, "Secondary Ion Mass Spectroscopy," in *Methods of Surface Analysis,* A. W. Czanderna, Ed., Elsevier, New York, 1975, p. 223.
79. A. Joshi, I. E. Davis, and P. W. Palmberg, "Auger Electron Spectroscopy," in *Methods of Surface Analysis,* A. W. Czanderna, Ed., Elsevier, New York, 1975, p. 164.
80. K. F. J. Heinrich, *Electron-Beam X-Ray Microanalysis,* Van Nostrand Reinhold Co., New York, 1981.

Micron Ex. 1023, p. 169
Micron v. YMTC
IPR2024-00909

81. W. K. Chu, J. Mayer, and M. Nicolet, *Backscattering Spectroscopy,* Academic, New York, 1978.

PROBLEMS

1. A heteroepitaxial layer with a lattice constant of a_1 is deposited on a (001) substrate with a lattice constant of a_s. Assuming that the layer is in a perfectly strained condition, what are its lattice constants along the directions normal and parallel to the substrate? Is the layer in tensile or compressive stress along these two directions if $a_1 > a_s$?

2. Following Problem 1, assume that the lattice constants of the layer normal and parallel to the substrate are c and a, respectively. How are a, c, and a_1 correlated?

3. Derive mathematically the relationship between the critical layer thickness h_c and the lattice-misfit strain ϵ under equilibrium conditions. Explain why a larger h_c value is obtained at lower growth temperatures.

4. Derive Eq. (3.6) in the text. Explain mathematically why reduced pressure and high temperature are preferable for a VPE process from a thermodynamic viewpoint.

5. Intrinsic diffusivity of B in Si can be expressed by $D = 0.76e^{-3.46(\text{eV})/kT}$. Calculate the minimum growth rate that is required when a Si-epi is grown on a heavily B-doped Si substrate for 20 min at 1200°C. Discuss why a low growth temperature is essential to achieve a submicron Si-epi layer.

6. The UHV/CVD technique is capable of achieving a boron doping concentration two orders higher than the solubility of boron in Si at a growth temperature of 550°C. Why? Discuss why a conventional MBE technique is not able to achieve such a high boron doping concentration.

7. Use a schematic drawing to illustrate how facets can be avoided when a Si-epi is grown with the SEG process on a (100)-oriented Si substrate.

8. Develop a method to determine a dislocation density of $10^2/\text{cm}^2$ in a Si-epi routinely using an optical microscopy.

Micron Ex. 1023, p. 170
Micron v. YMTC
IPR2024-00909

CHAPTER 4

Conventional and Rapid Thermal Processes

Richard B. Fair

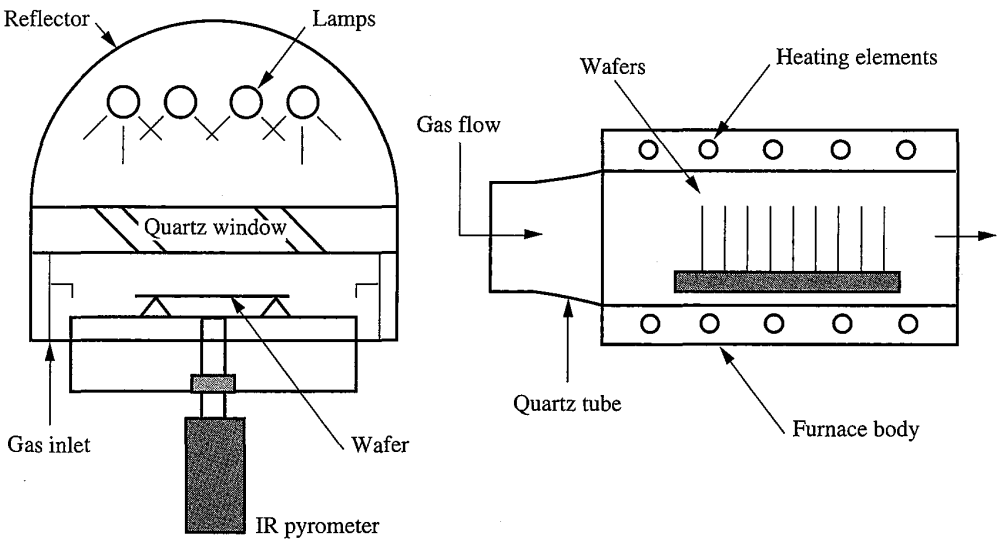
4.1 INTRODUCTION

The chemical and physical processes applied to silicon (Si) wafers are, in general, thermally activated; therefore, the wafers require heating at elevated temperatures so that these processes proceed at reasonable rates. Typical silicon-based semiconductor factories use *batch furnaces* for thermal fabrication steps, where a batch consists of 20 to 100 wafers that are simultaneously processed together in a single system. Batch-furnace thermal processes include thermal oxidation, dopant diffusion, annealing, glass reflow, and various chemical vapor deposition (CVD) processes. Such conventional thermal processing involves heating by convection and conduction where the wafers are in thermal equilibrium with the furnace surroundings. For example, batch wafer processing may typically occur in a clean, high-purity fused-quartz tube that is inserted into the core of a furnace that is heated resistively. Wafers are placed inside the open tube and pushed into the hot zone of the furnace where a uniform temperature is maintained. Thus, the heating elements surrounding the quartz tube heat the tube, the ambient inside the tube, and the wafers with long-wavelength thermal radiation (spectral distribution peak at 2 μm). In this approach, many wafers can be processed simultaneously, and the cost of the operation is divided among the number of wafers in each batch. However, in this "hot wall" process, the wafers are susceptible to contamination from the hot processing chamber on which deposited films may accumulate and flake off or through which impurities may diffuse and contaminate the wafers inside. For additional details on conventional, batch-furnace thermal processes, the reader is referred to the appropriate chapters on oxidation, diffusion, epitaxy, ion implantation, and dielectric and polysilicon film deposition in the previous edition of this book, *VLSI Technology,* (2nd ed.).[1]

Micron Ex. 1023, p. 171
Micron v. YMTC
IPR2024-00909

Processing of ULSI wafers requires tight control of contamination, process parameters, and reduced manufacturing costs. Toward this end, producers are beginning to use microprocessing environments for single wafers rather than batches of wafers. In single-wafer machines, it is necessary to extract a wafer out of a carrier and present it to the process chamber. If multiple process chambers are connected together, wafer transport among the different modules can be done best in a modest vacuum. Thus, a compatible heating technology is required that allows thermal processing to occur in closed microenvironments and that matches the wafer throughput of the conventional open-tube, batch-furnace systems. Rapid thermal processing (RTP) using transient lamp heating can meet this requirement nicely. With RTP, a single wafer is heated quickly under atmospheric conditions or at low pressure under isothermal conditions. The processing chamber is made of either quartz, silicon carbide, stainless steel, or aluminum and has quartz windows through which the optical radiation passes to illuminate the wafer. The wafer holder is often made of quartz and contacts the wafer in a minimum number of places. A measurement system is placed in a control loop to set wafer temperature. The RTP system is interfaced with a gas-handling system and a computer that controls system operation.[2] A schematic diagram of a basic RTP system is shown in Fig. 1*a* and is compared with a batch-furnace system that is resistively heated, shown in Fig. 1*b*. Typically, wafer temperature in an RTP system is measured with a noncontact optical pyrometer that determines temperature from radiated infrared energy.

The small thermal mass inherent in the RTP heating system stems from the fact that the wafer heats up but the surroundings remain cool. This results from



(*a*) RTP system (*b*) Batch-furnace system

FIGURE 1
(*a*) Rapid thermal processing (RTP) system that is optically heated compared with (*b*) batch-furnace system that is resistively heated.

Micron Ex. 1023, p. 172
Micron v. YMTC
IPR2024-00909

the Si wafers selectively absorbing radiation from the lamps, which produce short-wavelength radiation (spectral distribution peaks in the range of 0.5 to 1.0 μm). Thus, RTP is based on the energy transfer between a radiant heat source and a wafer, and the wafer is not, typically, in thermal equilibrium with the reactor wall. The response of a processed Si wafer to near-surface radiant energy deposition is a highly complex process in which the absorption of the energy depends on its wavelength, reflections of energy in the RTP chamber, as well as light diffusion, refraction, and transmission. Current knowledge is limited for the purposes of accurate modeling of wafer heating dynamics, so equipment manufacturers must rely on empirical methods of measurement and design. Some of the basic concepts are presented here.

4.1.1 Thermophysics in Rapid Thermal Processing (RTP)

The basic physical laws that describe radiative heat flow begin with the total radiant exitance, M, expressed in units of watts emitted per square meter of surface by an object with absolute surface temperature T, and given by the Stefan–Boltzmann law[3]

$$M(T) = \varepsilon\sigma T^4 \tag{4.1}$$

where ε is the emissivity and σ is the Stefan–Boltzmann radiation constant with a value of $5.6697 \times 10^{-8}\text{W/m}^2\text{K}^4$. In an RTP system with lamps and reflectors removed some distance from the wafer and with an intervening quartz optical window in place, it is necessary to produce a high level of radiative heat flow. From Eq. (4.1), this requires that the temperature of the source be very high. In addition, each type of radiation source has the wavelength of the maximum intensity related to its temperature according to Wien's displacement law[3]

$$\lambda_{\max} \cdot T = 2.89783 \times 10^{-3} \tag{4.2}$$

where $\lambda_{\max}$ is in units of meters. Typical lamps in practical RTP systems have theoretical blackbody color temperatures from Eq (4.2) between 6000 K (arc lamp) and 2900 K (tungsten filament).[4] However, the color temperature of the heated wafer is much lower, usually 900 to 1400 K. Thus, in the optical-heating reactor, the absorption and emission spectra of the Si wafer never overlap. The result is that temperature nonuniformity is created in the wafers, and this is the greatest challenge to the acceptance of RTP as a manufacturing technology. By contrast, conventional furnaces that are resistively heated produce a blackbody color temperature of 1450 K, which allows for the thermal-uniformity characteristic of wafers that have seen furnace heating.

The total emissivity of Si in an RTP system depends on the optical properties of the starting wafer (*intrinsic emissivity*), the dielectric or conducting layers that may exist on top of the wafer surface or on the wafer bottom, buried layers within the wafer (*extrinsic emissivity*), and the specific optical properties of the reflective chamber and the components contained therein (*effective emissivity*).[4] The intrinsic emissivity of a perfect blackbody is $\varepsilon = 1$, causing it to emit the maximum possible amount of radiation [see Eq. (4.1)]. Also, under thermal equilibrium conditions, a diffuse radiator will have an emissivity that is equal to its absorptivity.

Micron Ex. 1023, p. 173
Micron v. YMTC
IPR2024-00909

The emissivity, ε, of a 1.8-mm-thick, 15-ohm-cm, phosphorus-doped silicon wafer is shown in Fig. 2 as a function of wafer temperature and radiation wavelength.[5] It can be seen that for $\lambda > 1.5$ μm, ε varies from near zero at low temperatures up to 0.7 as the temperature increases. The value of 0.7 is characteristic of an ideal gray body. The deviation from ideal gray body behavior is caused by electronic valence-to-conduction band absorption, which is efficient for wavelengths below the silicon bandgap of 1.2 μm. Above 6 μm, absorption due to lattice vibrations becomes efficient. Between these two wavelengths (1.2 to 6 μm) and below 600°C, the main Si wafer heating process is by intrinsic free-carrier absorption. As a

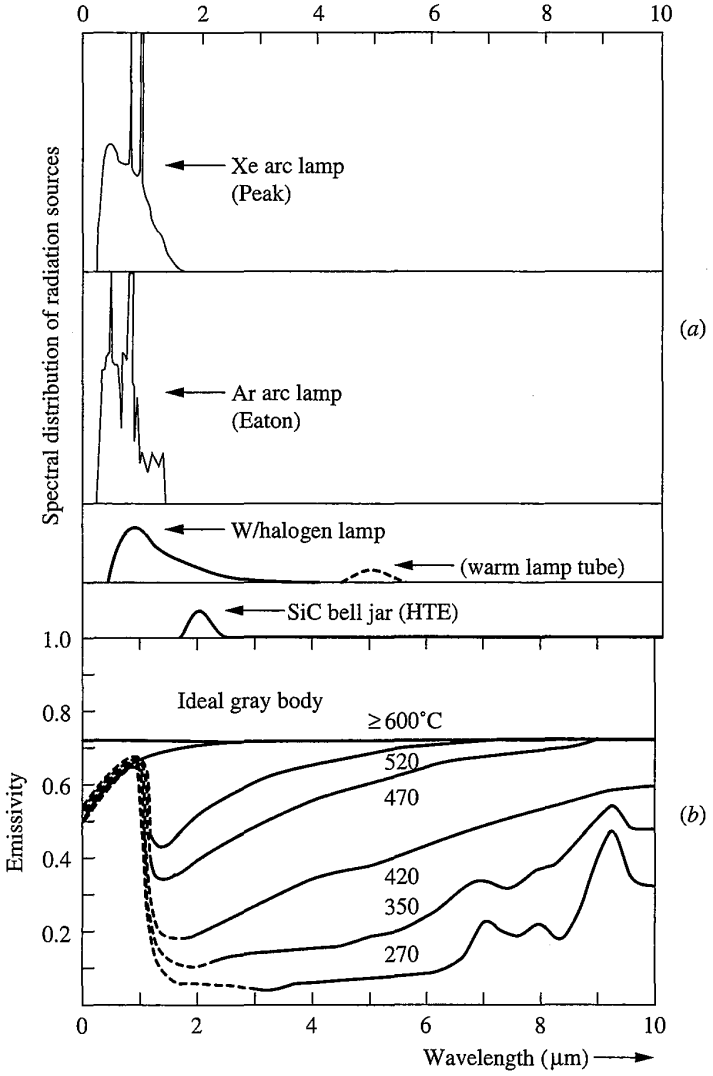


FIGURE 2
Spectral features in RTP systems: (*a*) heat sources (*After Roozeboom, Ref. 6.*); (*b*) temperature dependence of emissivity of a 1.8-mm-thick, 15 ohm-cm, P-doped Si wafer. (*After Sato, Ref. 5.*)

Micron Ex. 1023, p. 174
Micron v. YMTC
IPR2024-00909

result, the emissivity in this range depends strongly on the free-carrier density and, therefore, dopant concentration, wafer thickness, and temperature.

From this discussion it is evident that proper selection of the spectral output of the heating lamps in an RTP system is important, since this determines the response of the processed semiconductor to the rapid near-surface energy deposition. For comparison, the spectral distributions of several popular radiation sources used in commercial RTP systems are shown[6] in Fig. 2*a*. Since a portion of these spectra occur at wavelengths above 1.2 μm, the initial heating rate of Si below 600°C depends on free-carrier absorption and, thus, on doping concentration. And the fraction of radiation that falls in this range is different for the various lamps. For example, the fraction is smallest for arc lamps since they emit their spectra with a peak wavelength of 0.5 μm. Tungsten-halogen lamps emit radiation with a peak at 1.0 μm, producing a larger fraction of energy above the silicon bandgap. Finally, resistively heated silicon carbide bell jars exhibit a peak at 2.0 μm. Thus, the influence of lamp spectral characteristics on wafer heating is only important up to 600°C, at which point the wafer is no longer transparent to radiation but rather becomes an opaque gray body with $\varepsilon = 0.7$.

Device wafers contain layers of deposited films and patterns in those films, which further complicates the concept of emissivity. Parameters that affect the extrinsic emissivity include film thicknesses (front and backside), material optical properties, and etched features. Optical interference will produce changes in emissivity, and thus absorptivity, during the deposition of films. For example, in Fig. 3, emissivity changes are shown for Si with 0.5 μm of SiO_2 and variable thicknesses of deposited polycrystalline Si.[7] This graph reveals that the spectral emissivity oscillates as a function of film thickness. The implications for such variations are serious, since the surface temperature varies in a similar fashion. It has been reported that temperature variations can exceed 100°C during poly-Si deposition in an RTP system, causing alternating layers of amorphous and polycrystalline Si if the temperature swings below the polycrystalline-to-amorphous transition value.[8]

The walls of an RTP chamber are made highly reflective to overcome large variations in extrinsic emissivity as described above. For example, Fig. 3[7] shows two emissivity curves for poly-Si thickness changes; one in a chamber with 0% chamber reflectivity and one with 70% reflectivity. The incidence of direct and reflected radiation on a wafer is shown schematically for a typical RTP reactor in Fig. 4.[9] It can be seen that reflections occur off of all surfaces onto both the front and the backside of the Si. This phenomenon complicates the determination of the emissivity. Thus, the extrinsic wafer parameters together with the RTP chamber and its internal components make up the overall effective emissivity.[4] Effective emissivity is, therefore, dependent on the specific reactor and must be measured at the wafer surface in conjunction with an infrared pyrometer to determine the temperature at any given time during an RTP process.

If a wafer is patterned, an additional complication occurs during lamp heating, since the various surfaces absorb energy differently, creating a lateral temperature distribution across the wafer. This nonisothermal nature is due to larger differences in the time constants of the physical processes involved. For instance, the process of photon absorption has a very short reaction time (10^{-12} to 10^{-14} sec), whereas the

Micron Ex. 1023, p. 175
Micron v. YMTC
IPR2024-00909

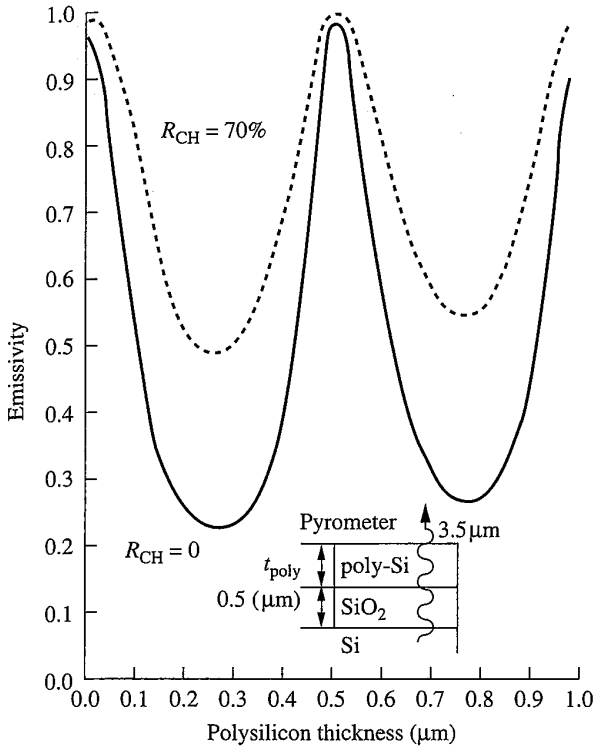


FIGURE 3
Emissivity changes with varying polycrystalline Si thickness deposited on 0.5 μm SiO_2 as measured with a pyrometer at 3.5 μm in a black chamber (reflectivity, R_{CH} = 0) and in a 70% reflecting chamber. (*After Hill, Ref. 7.*)

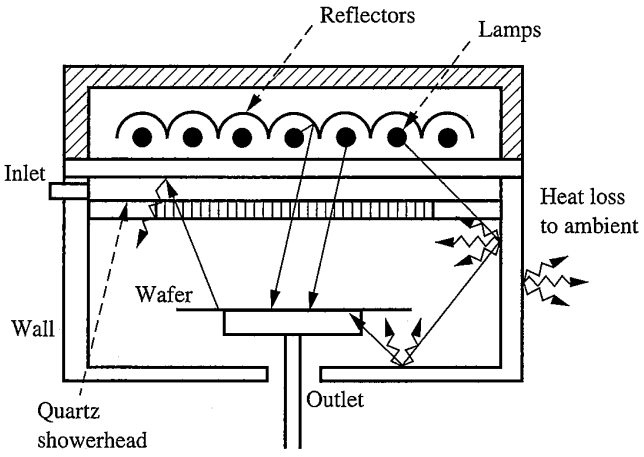


FIGURE 4
Modes of radiation heat transfer in the system and to the ambient in a typical RTP reactor. (*After Merchant et al., Ref. 9.*)

time constant for thermal conductivity is a function of the square of the distance. Consequently, there is a great difference between the vertical response time (temperature changes in 5 to 50 msec) and the lateral thermal response time (temperature changes in 1 to 10 sec) of the wafer among different features on the surface.[10] Thus, RTP is rarely an isothermal process in practice.

Micron Ex. 1023, p. 176
Micron v. YMTC
IPR2024-00909

4.1.2 Other RTP Approaches

While the challenges inherent in optical heating are being tackled by many equipment companies, alternative approaches have been implemented that use wafer susceptors with lamps to improve wafer temperature nonuniformity[11] or that use continuous heat sources. Rapid wafer heating can be accomplished with continuous, resistively heated sources when the wafer is moved quickly in and out of the vicinity of the heat source. The difference between optical heating and this approach is that thermal equilibrium is maintained within the hot wall–ambient system while the wafer is heating or cooling. A diagram of the system is shown in Fig. 5.[12] The wafer is supported on three quartz pins on an elevator that is rapidly inserted into a silicon carbide bell jar. The top of the bell jar is resistively heated and the bottom of the bell jar is water-cooled, establishing a smooth temperature gradient from top to bottom. Annealing temperatures are dependent on the position of the elevator in the jar, and control in the range of 2 to 4°C is possible across a 200 mm wafer. Temperature ramp rates are 10 to 100°C/sec, whereas optical systems can operate at 50 to 300°C/sec.

4.1.3 Comparisons between RTP and Furnaces

A comparison between conventional furnace and RTP technologies is shown in Table 1. In order to achieve short processing times, tradeoffs must be made in temperature

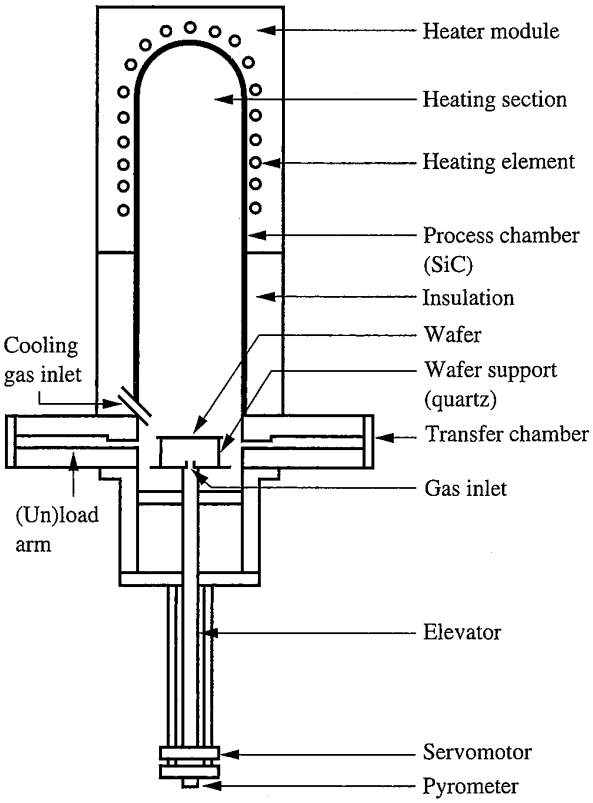


FIGURE 5
A schematic view of a continuous heat source, vertical furnace system for rapid thermal heating. (*After Lee and Chizinsky, Ref. 12.*)

Micron Ex. 1023, p. 177
Micron v. YMTC
IPR2024-00909

TABLE 1
Technology comparisons

Furnace	RTP
Batch	Single-wafer
Hot wall	Cold wall
Long time	Short time
Small dT/dt	Large dT/dt
High cycle time	Low cycle time
Environment temperature measurement	Wafer temperature measurement
Issues	Issues
Large thermal budget	Uniformity
Particles	Repeatability
Atmosphere control	Throughput
	Wafer stress
	Absolute temperature measurement

and process uniformity, temperature measurement and control, and wafer stress and throughput. In addition, there are concerns over the introduction of electrically active wafer defects during the very fast (100 to 300°C/sec) thermal transients.[13] Rapid heating with temperature gradients in the wafers can cause wafer damage in the form of slip dislocations induced by thermal stress. On the other hand, conventional furnace processing brings with it significant problems such as particle generation from the hot walls, limited ambient control in an open system, and a large thermal mass that restricts controlled heating times to tens of minutes. In fact, it is these latter issues with the added concern of manufacturing costs that have driven semiconductor producers to consider RTP over conventional furnace technologies. Requirements on contamination, process control, and cost of manufacturing floor space are driving a paradigm shift to a single-wafer, microprocessing methodology.

To date, RTP has been integrated into some noncritical ULSI manufacturing applications where the advantages of a reduced thermal cycle are obtained. For RTP to replace existing manufacturing technologies it must demonstrate technical and cost advantages such as higher process yield, higher throughput, and no lost processing potential. The major limitation on RTP has been uncertainty in temperature uniformity and repeatability. And high process yield can only be obtained by improving the repeatability and uniformity intrinsic to the equipment.[14] We now investigate some additional manufacturing requirements for thermal processes.

4.2 REQUIREMENTS FOR THERMAL PROCESSES

ULSI processing equipment must produce chips with high reliability, tight tolerances, and high yield. The specific requirements on thermal processes can be stated in terms of a thermal budget, which is the allowed time-at-temperature for a given step, as well as a manufacturing budget that deals with temperature control, reproducibility, process step yield, and uniformity across a wafer. For example, typical requirements on equipment to meet the thermal and manufacturing budget needs of junction formation and oxidation include the following:

Micron Ex. 1023, p. 178
Micron v. YMTC
IPR2024-00909

Mechanical control budget for processing and positioning large-diameter wafers
Temperature uniformity and reproducibility budget within each wafer and from run to run
Atmospheric budget for controlling gases used in the equipment
Time control budget
Absolute, repeatable temperature measurements on wafers

4.2.1 Thermal Budget

Dopants are introduced into Si by processes involving ion implantation, gaseous sources, solid sources, or crystal growth. Each method creates n- or p-type conductivity layers with a desired spatial distribution depending on the application. The desired time-at-temperature (thermal budget) to produce pn junctions at the desired depth from the Si surface for p-channel metal-oxide-silicon (MOS) transistor sources and drains is shown in Fig. 6. Assuming a gaseous dopant source is used that gives a boron (B) surface concentration of 1×10^{20} atoms/cm^3, then the thermal budget as a function of temperature can be determined. In Fig. 6, the requirements for four MOS technologies specified by the minimum patterned feature sizes (2, 1, 0.5, and 0.25 μm) are plotted.[15] Thus, for each generation of technology the time allowed at each temperature to form a junction at the necessary depth can be determined from the calculated lines. For example, for 0.25-μm devices, the time allowed to form the required 700 Å deep source-drain junction at 1000°C is only 24 seconds! This compares to junction-formation times of over 1000 sec at 1000°C for 2-μm technology.

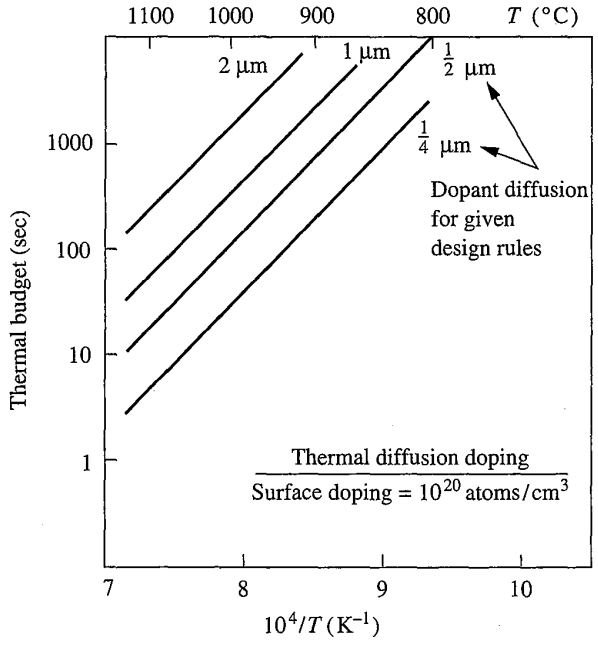


FIGURE 6
Available time-at-temperature (thermal budget) to produce scaled p^+n junction depths for four MOS technologies. Thermal diffusion sources are assumed with a fixed surface concentration of 1×10^{20} atoms/cm^3. (*After Fair, Ref. 15.* © 1990 IEEE)

Micron Ex. 1023, p. 179
Micron v. YMTC
IPR2024-00909

This example shows us that the thermal budget requirements for MOS technology are driving a change in the type of furnaces needed to form pn junctions by diffusion. The natural question is, Why not keep the conventional batch furnaces and reduce the temperature to get more thermal budget time? If this were done for the last example, we could operate at 800°C and have 1000 seconds of processing time available. Although this certainly is feasible, gaseous source diffusion is generally not used, because it is difficult to get rid of the source after the junctions are formed. Traditional diffusion sources form doped glasses in the open windows that are defined in SiO_2. These glasses etch marginally faster than SiO_2, making them difficult to remove without losing the definition of the window itself. In addition, doping control is difficult with gaseous sources. The preferred process for introducing dopants into Si is ion implantation, since it is controllable and there is no residual source material to remove. However, implanting high-energy ions into a Si substrate produces extensive damage to the crystal through collisions between the incoming ions and the lattice atoms.[16] Once the temperature of the implanted wafer is raised sufficiently, the implant damage begins to anneal or change form, producing large quantities of point defects, i.e., Si vacancies and Si self-interstitial atoms. If the implant dose exceeds approximately 2×10^{14} ions/cm^2, more stable, extended defects may form, such as dislocation lines and loops. As we will see, an 800°C anneal is not sufficient to remove these extended forms of damage, which may cause incomplete dopant activation or leaky junctions. Thus, the appropriate annealing temperatures for ion-implanted junctions are above 900°C, which constrains the thermal budgets for shallow junctions for ULSI devices to tens of seconds.

Perhaps the most significant factor associated with the use of ion implantation in junction formation for ULSI is the production of large quantities of point defects during the damage anneal. The basic process of thermal diffusion of dopants requires the presence of such point defects, usually found in the Si crystal at equilibrium concentrations. However, damage annealing can produce point defects in concentrations that may reach many orders of magnitude above equilibrium levels at the annealing temperature. As a result, dopant diffusion can be significantly enhanced during ion-implantation damage annealing when excess point defects are present.[17]

The time required to remove the simplest forms of damage produced by ion implantation (point defect clusters) is measured in seconds at temperatures above 1000°C. This time is considered insignificant for older MOS technologies with deep source/drain junctions that require considerable diffusion times. However, this is not the case for ULSI technologies, where junctions must be established in seconds. As a result, thermal budgets can be dominated by the effects of implantation damage annealing. The thermal budgets for ion-implanted p^+ junctions for the same four generations of MOS technologies previously discussed in Fig. 6 are now shown in Fig. 7. Implants consisted of 6- to 10-keV BF_2 ions to form B-doped layers in Si with surface concentrations of 1×10^{20} atoms/cm^3.[5] The thermal-budget curves were calculated using the PREDICT simulator,[18] which contains models for damage effects on diffusion. For 0.25-μm technology, only 2.5 seconds are available for a diffusion temperature of 1000°C, which is a factor of 10 less than the thermal budget for a comparable gaseous-source process in Fig. 6. And, as we have seen, reducing the diffusion temperature to buy more process time is a limited option. It is usually

Micron Ex. 1023, p. 180
Micron v. YMTC
IPR2024-00909

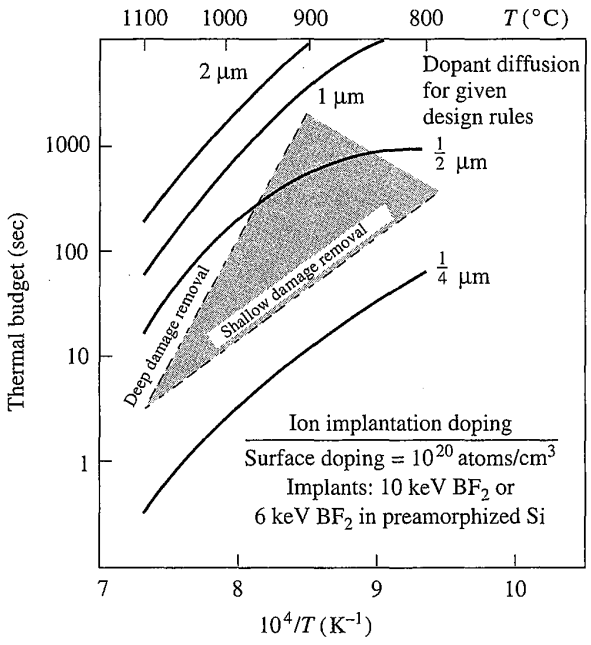


FIGURE 7
Available time-at-temperature (thermal budget) to produce p^+ scaled junction depths for four MOS technologies. Ion-implantation doping using BF_2 ions is assumed with a final surface concentration of 1×10^{20} atoms /cm^3. Shown for reference is the range of times required to remove shallow and deep extended dislocations caused by implantation damage. (*After Fair, Ref. 15.* © 1990 IEEE)

important to remove the implantation damage before the desired junction depth has been achieved. Otherwise, the annealing of the damage with the resulting generation of excess point defects in the Si will dominate the junction formation process rather than normal, thermally assisted diffusion. This results in poor control, because implantation damage is not very reproducible from one ion implanter to the next. The variables in damage production include wafer temperature during implantation, ion beam currents (dose rates), and wafer orientation with respect to the beam. Additional uncertainties follow when the wafer is heated. The production of excess point defects depends on temperature ramp rates as well as the final annealing temperature. It follows, then, that the preferred junction formation process should be controlled by thermally assisted diffusion (no damage) rather than by damage-assisted diffusion.

Times to remove shallow and deep dislocations produced by ion implantation are superimposed on the diffusion-time curves in Fig. 7.[15] Such dislocations result when heavy ion implants are performed at doses sufficient to turn the implanted crystal amorphous. The shallow dislocation boundary is based on data for damage 400 Å deep, and the deep dislocation boundary is for damage 1000 Å deep. The time to remove damage continues to decrease as the proximity to the surface increases. For example, the time required to form 700-Å-deep junctions will be less than the minimum time needed to remove the 400-Å-deep dislocations. Thus, the damage will remain after annealing, unless the dislocations are within 200 Å of the Si surface. To do this requires ultralow-energy ion implants.

Experience has shown that rapid thermal annealing will be required to achieve defect-free junctions for ULSI. The technology allows rapid heating to high temperatures for short times, which is best for removing damage and limiting diffusion of dopants. In addition, the technology road map for shallow junctions shown in Fig. 8 requires the use of either very-low-energy ion implantation ($E < 1$ keV for B—see Section 4.3.4, "RTA of ion-implanted junctions")[19] or solid diffusion sources that

Micron Ex. 1023, p. 181
Micron v. YMTC
IPR2024-00909

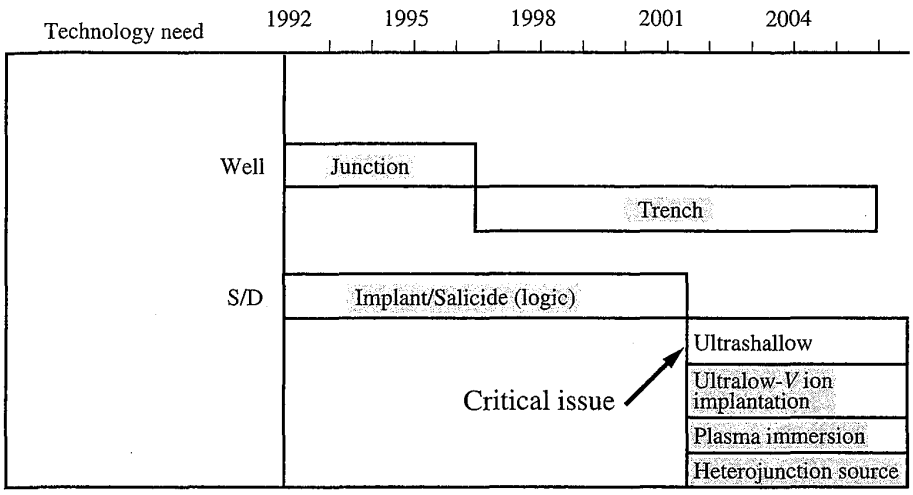


FIGURE 8
Semiconductor Industry Association front-end-of-the-line road map for isolation wells and source /drain junction formation showing current and future technologies.

become part of the device structure. Both approaches will yield diffusion-controlled junction processing, which is necessary for manufacturability.

4.2.2 Ambient Control Budget

The smallest fabricated dimension in an MOS transistor is the gate-oxide thickness, which is produced by thermal oxidation. For 0.25-μm devices, a 70-Å gate oxide controlled to within 7 Å is needed; this meets stringent reliability requirements with defect densities on the order of $0.5/cm^2$. To manufacture devices with ultrathin oxides requires careful process control and oxidation-furnace optimization. For example, the use of 150 to 200-mm diameter Si wafers requires large-diameter furnace tubes in which careful ambient control is difficult. The backstreaming of air from the large open ends of the furnace tubes is a major problem. In addition, the partial pressures of the oxidants, dry O_2 or wet O_2, would have to be maintained within 6% of nominal.[15] For dry O_2 oxidation, there is the additional requirement that trace amounts of water be controlled. A relatively small amount of water (125 ppm) in O_2 can increase the oxidation rate at 980°C, enough to throw the oxidation process out of spec for a sub-0.5 μm technology.[20]

Silicon surface control prior to the growth of thin oxide layers is also important because a native oxide grows on bare Si at room temperature. Thus, 70-Å film growth can be controlled if the Si surface is cleaned with HF, which leaves the Si surface bonds terminated with H.[21] Desorption of the H at 300°C in a highly pure Ar gas ambient followed by the formation of one monolayer of oxide passivates the surface for subsequent gate-oxide growth.

For ambient control, single-wafer RTP chambers offer a microenvironment approach that can satisfy the stringent requirements for ultrathin growth using rapid thermal oxidation (RTO).

Micron Ex. 1023, p. 182
Micron v. YMTC
IPR2024-00909

4.2.3 Contamination Budget

Scaling transistors to dimensions suitable for ULSI has a profound effect on the manufacturing yield and reliability of integrated circuits. Processing complexity naturally increases with each generation of chips, caused in part by the need for additional mask levels to interconnect the increased number of subcircuits on the chip. This trend makes devices more susceptible to contamination introduced by particulate and chemical impurities at each process step. In addition, smaller devices are vulnerable to smaller defects and smaller amounts of chemical impurities that may result in chip loss.[22]

Particles can cause yield loss through the presence of random defects in the patterning of film levels. Chip yield is expressed in terms of the defect density through various statistical models such as the Poisson distribution,[23]

$$\text{Yield} = e^{-A\rho} \tag{4.3}$$

where A is the chip area and ρ is the density of defects per unit area. As devices are scaled down in their dimensions, the density of potentially fatal defects increases rapidly. This result follows because killer defects are at least one-third the size of a lithographic feature or one-half of a film thickness. With a gate oxide 70 Å thick, a 35-Å particle could be fatal.

Measured particle densities are plotted in Fig. 9 versus particle size for bulk gases, semiconductor chemicals, deionized water, and cleanroom air at MCNC (Microelectronics Center of North Carolina).[15, 22] All of these distributions show increasing particle densities with decreasing particle size. And this is the environment in which smaller devices will be made. Device scaling by a factor of two occurs in a processing environment in which the number of potentially fatal particles increases by four to eight times.[22] The impact on yield may be devastating. Under these conditions, a process that yields 25% and is limited by particle contamination would yield nothing after scaling it down by a factor of two. Consequently, improved means must be found for controlling particle size and numbers in the processing environment.

Improved contamination control is possible in single wafer, in-situ processing where process modules are clustered together and interconnected by vacuum transport means. An example of such an arrangement is shown in Fig. 10,[24] where the connected processing modules include wafer cleaning, silicon oxidation, and low-pressure chemical vapor deposition (LPCVD). Rapid thermal processing technology allows for the use of cold-wall systems, which can better meet the stringent requirements for low-level contamination. It should be noted that recent developments of vertical batch furnaces can help reduce the contamination problem inherent in hot-wall, conventional furnaces. However, there is no methodology to integrate such furnaces with an ultralow-contamination environment that protects the wafers during pre- and post-processing.

4.2.4 Device Electrical Budget

The electrical budget refers to the control of electrical device parameters that are determined by device-scaling rules. Included in this budget are specifications

Micron Ex. 1023, p. 183
Micron v. YMTC
IPR2024-00909

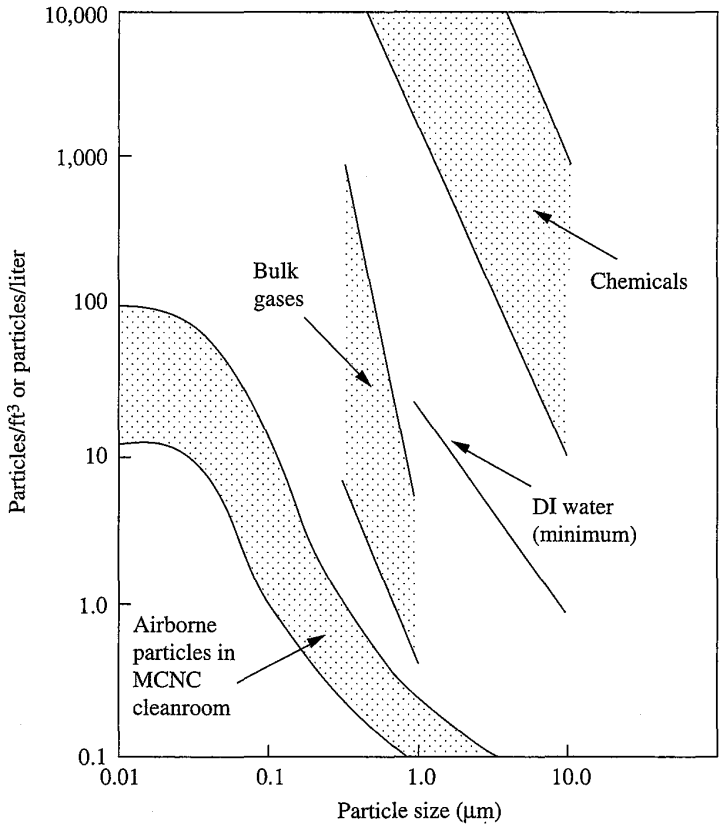


FIGURE 9
Comparison of particle size distributions in bulk gases, chemicals, deionized water, and processed air in the MCNC semiconductor cleanroom. The particle densities increase rapidly with decreasing particle size; this is the environment in which smaller silicon devices are being scaled. (*After Osburn et al., Ref. 22.*)

for contacts, interconnections, electric-field levels, and process-induced electric charge.

For ULSI, it is desirable that contacts with low specific resistance (1×10^{-8} ohm-cm^2) to semiconductor junctions be made with high yield. These contacts must also be reliable, serving as barriers to unwanted metal reactions with Si. Barrier layers of titanium/tungsten and titanium nitride have proven to be good choices.

Low-resistance contacts are imperative. If contact dimensions are halved, contact resistance increases by a factor of four. Thus, specific contact resistances must be decreased by factors of 10 or better for ULSI.

Refractory silicides of transition metals have been used to improve contact resistances. Recent work with Al–TiW–$TiSi_2$ contacts to shallow n^+ junctions has been reported.[25] By performing sputter etching of the $TiSi_2$ surface to remove any oxides prior to TiW deposition in the same vacuum environment, specific contact resistances below 2×10^{-8}ohm-cm^2 can be achieved. Thus, the contact-resistance budget is driving the use of in-situ vacuum processing in an RTP, single-wafer module.

Micron Ex. 1023, p. 184
Micron v. YMTC
IPR2024-00909

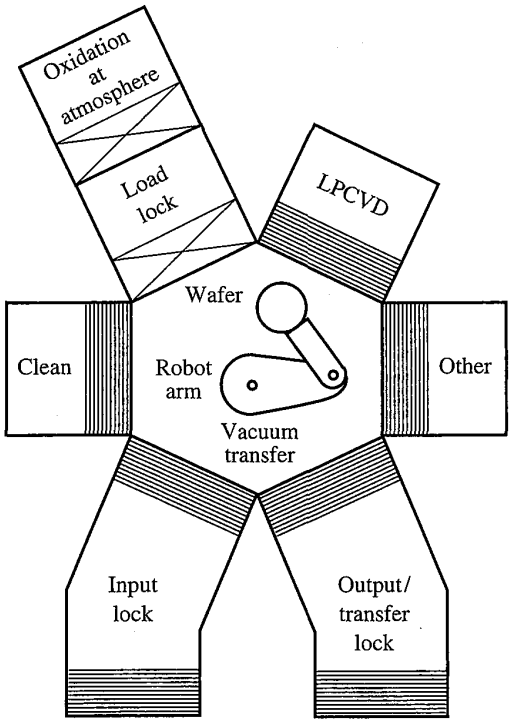


FIGURE 10
Example of rapid thermal processing modules in a radial, multichamber cluster designed for better particle exclusion, interprocess ambient control, and flexible operation. (*After Rosser, Moynagh, and Affolter, Ref. 24.*)

The intrinsic electrical-parameter budget of submicrometer semiconductor devices becomes increasingly bounded as device dimensions decrease. For example, submicrometer MOS transistors must be designed and manufactured without exceeding physical limits imposed by drain-junction avalanche breakdown, bulk punchthrough, short-channel effects, and hot-electron effects.[26–29] In addition, device design for one-transistor DRAM cells is constrained by noise-margin requirements that are dominated by the threshold-voltage mismatch at the input of the on-chip sense amplifier.[27] Although there are several sources of threshold-voltage mismatch, one that will become much more important as device dimensions approach 0.1-μm minimum feature size is the variation resulting from channel doping distribution statistics. Indeed, channel doping by ion implantation causes scattering in the location of the impurity atoms in the Si, which causes a one-sigma variation in V_t of 15 mV for a 0.1-μm channel length.[30] However, this is a control problem that can be solved by rapid thermal epitaxy and atomic-layer doping in a single-wafer processing module (see Section 4.3.1). Such new doping technologies will also affect control of the subthreshold conduction current between source and drain of a MOSFET. This current can vary in a 0.5-μm device by a factor of ten with a 500-Å change in channel profile depth. In addition, such process parameter variations also change the subthreshold slope (the change in gate voltage produces a change in drain current) by 25 mV/decade.[31]

4.2.5 Cost Issues and the Manufacturing Budget Crisis

The cost of manufacturing submicrometer ULSI chips is increasing in inverse proportion to the decrease in device feature sizes. These expenses are driven by a man-

Micron Ex. 1023, p. 185
Micron v. YMTC
IPR2024-00909

ufacturing budget crisis associated with the technology and complexity limits of integrated circuit (IC) design, costs of research and development to address manufacturing issues, and facility capitalization.[6] Thus, examination of the rate of progress in microelectronics over the past 30 years suggests that the primary challenge in reaching the gigachip age in the year 2001 will be the semiconductor industry's ability to change the cost trend lines; i.e., to change the economics of how ICs are developed and manufactured.[2]

Several approaches have been suggested for reducing the costs associated with developing and manufacturing ULSI chips. Requirements on contamination, process control (manufacturing-parameter budgets), and cost of manufacturing floor space are driving a paradigm shift to a microprocessing methodology. Thus, single-wafer processing environments with highly controlled, ultraclean ambients clustered together in specialty process modules are being considered. Today, in-situ vacuum processing equipment accounts for 40% of the total equipment, and it is projected that by the year 2001, 80% of ULSI equipment could be in-situ vacuum-based modules clustered together and controlled by a factory information system.[2]

4.3 RAPID THERMAL PROCESSING

Perhaps the most important feature of a rapid thermal processing system is its generation and delivery of radiant energy to the wafer from tungsten-halogen lamps in a wavelength band of 0.3 to 4.0 μm. Because of the optical character and wavelength of the energy transfer, the quartz walls do not absorb light efficiently, whereas the Si wafer does. Thus, the wafer is not in thermal equilibrium with the cold walls of the system. This allows for short processing times (seconds to minutes) compared to minutes to hours for conventional furnaces.

With the addition of reactive ambients to RTP reactors, other applications have opened up besides simply heating wafers. The merger of single-wafer RTP and single-wafer chemical vapor deposition (CVD) has brought continuous improvements to film deposition technology. Thus, rapid thermal CVD (RTCVD) technology provides a means of controlling chemical vapor deposition processes with temperature ramps rather than by turning gas flows on and off. This methodology is called *limited reaction processing* (LRP).[32] The use of LRP in silicon epitaxy is demonstrated in Fig. 11 where temperature-versus-time profiles are drawn comparing conventional furnace epitaxy to the LRP growth process.[33] The reduction in temperature-time exposure afforded by the RTP-based method is dramatic.

4.3.1 Epitaxy

Processing demands on the growth of high-purity epitaxial Si include ambient purity (oxygen and water levels in the parts per billion range), optimization of gas flow patterns, minimum wall deposition, and vacuum compatibility. Thus, an RTP system suitable for epitaxy will look different from other RTP systems. A schematic

Micron Ex. 1023, p. 186
Micron v. YMTC
IPR2024-00909

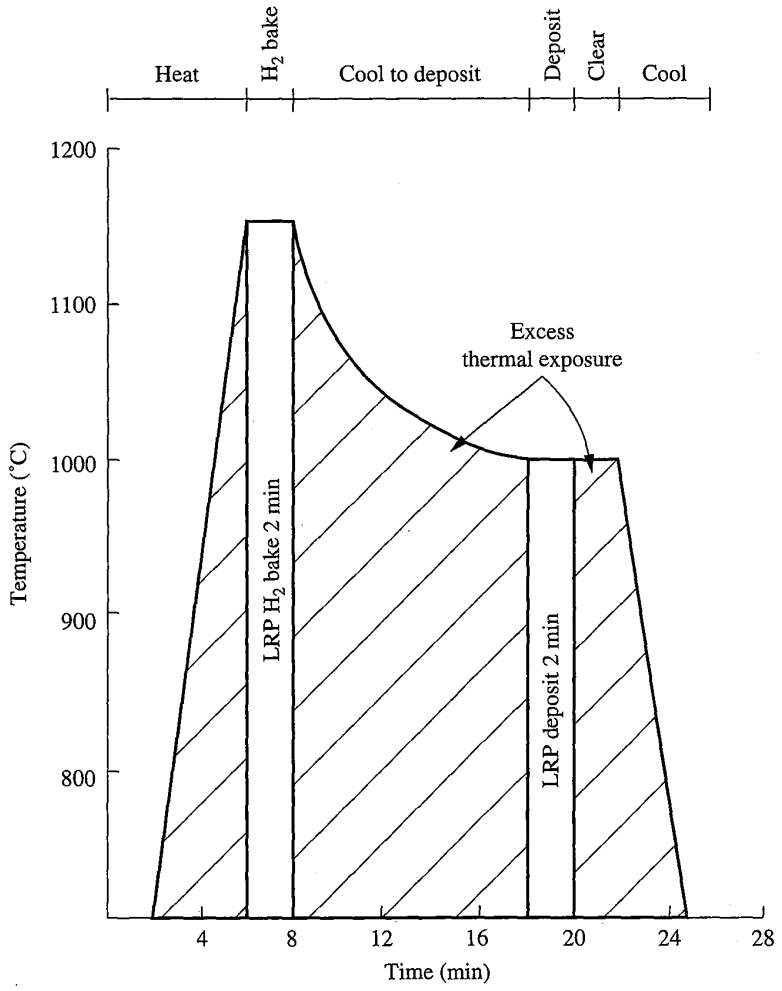


FIGURE 11
Schematic temperature versus time profile comparing conventional furnace, silicon epitaxy (hatched regions) to the LRP growth process. (*After Hoyt, Ref. 33.*)

diagram of an advanced, low-pressure, epitaxial Si reactor based on RTP technology is shown in Fig. 12.[33] This technology is an alternative to other growth methods such as plasma-enhanced CVD, ultrahigh-vacuum (UHV) CVD, and molecular beam epitaxy (MBE).

Conventional epitaxial growth of Si is usually performed in a quartz reaction chamber with a batch of wafers placed flat against a graphite susceptor. Deposition occurs with the wafers held at an elevated temperature while the reactant gas stream flows across the wafers' surfaces. Typical Si source gases include silane and dichlorosilane, and a hydrogen carrier gas provides the required gas velocity. Overall, the deposition process comprises a mass-transport process with a weak temperature dependence and a sequential surface-reaction process that is exponentially dependent on wafer temperature (see Chapter 3). Since the properties of epitaxial

Micron Ex. 1023, p. 187
Micron v. YMTC
IPR2024-00909

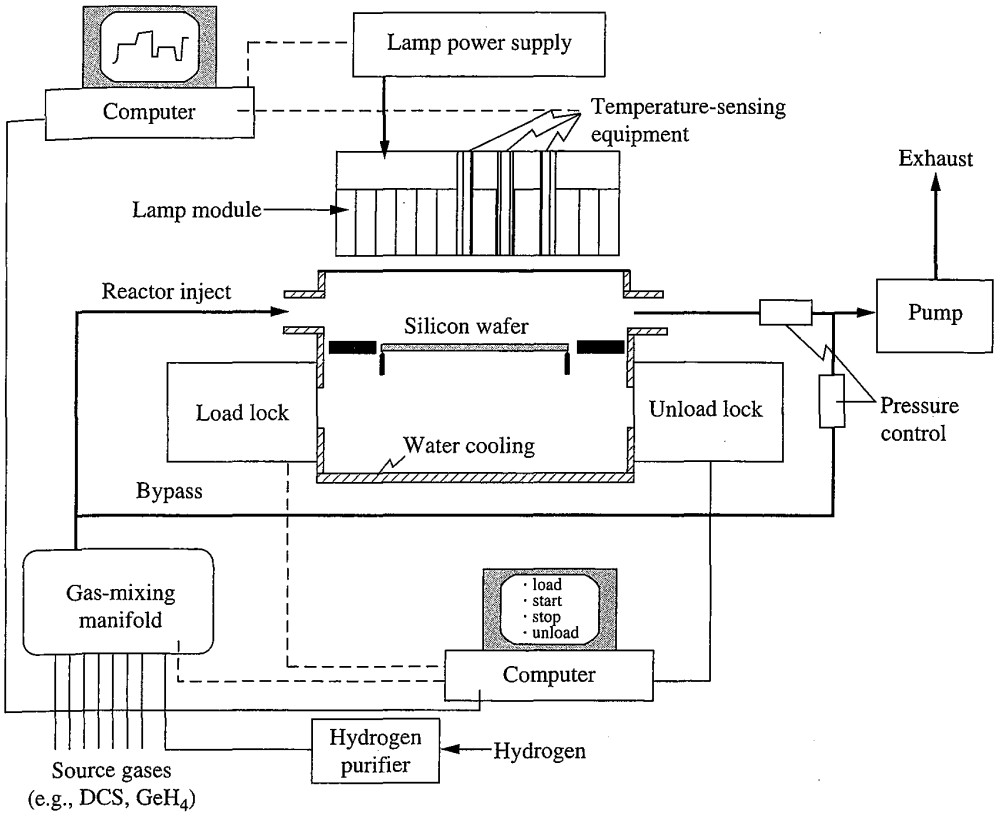


FIGURE 12
Diagram of an advanced, low-pressure, RTP, epitaxial, Si reactor. (*After Hoyt, Ref. 33.*)

layers improve as the growth pressure decreases, most conventional systems operate at either atmosphere (760 torr) or reduced pressures (20 to 100 torr). Typically, the graphite susceptor is heated inductively or by infrared lamps to around 1150°C, and the quartz chamber is water- or air-cooled or both.[33]

The requirements for ULSI devices include thinner epitaxial layers, better dopant profile control, and selective epitaxial growth.[34] In addition, lower cost-per-wafer requirements translate into factors such as equipment costs, energy consumption, and raw materials. If such costs remain constant, then the wafer cost varies inversely with the reactor throughput.[35] But throughput in conventional reactors is achieved primarily by the number of wafers that can be accommodated on the susceptor. As wafer sizes increase, fewer wafers can be processed in a single run. Thus, the next-generation machine must be purchased with a larger susceptor and a higher price. On the other hand, for RTP, throughput depends minimally on wafer size and mostly on processing time, which can be reduced with improved automatic wafer-handling tools and high-rate processes.

Selective growth

To achieve full selective growth of silicon epitaxy from dichlorosilane in a conventional tool requires the addition of HCl. The presence of HCl is important for

Micron Ex. 1023, p. 188
Micron v. YMTC
IPR2024-00909

etching small Si nuclei and clusters that will try to grow over the masking layer where growth is undesirable. The widely accepted vapor phase reaction for dichlorosilane is

$$SiH_2Cl_2 \rightarrow SiCl_2 + H_2 \tag{4.4}$$

Under conditions of limited reaction processing, full, selective, epitaxial growth of Si has been observed over the temperature range from 650 to 1100°C without the addition of HCl.[33] It has been proposed that the reaction kinetics are different under LRP conditions and that HCl is produced as a by-product of the dichlorosilane reaction.[36] Thus, LRP deposition kinetics are controlled by this reaction:

$$SiH_2Cl_2 \rightarrow SiHCl + HCl \tag{4.5}$$

The dissociation energy of Eq. (4.5) is 2.7 eV, whereas Eq. (4.4) yields 1.4 eV. Measurements of growth rate versus temperature for three deposition conditions are shown in Fig. 13.[36] Unlike conventional epitaxy, where surface reaction energies are independent of gas-phase species,[37] LRP conditions produce variable reaction energies that depend on Cl concentration in the gas stream. The activation energies for the three curves are (*a*) 1.7 eV, (*b*) 2.6 eV, and (*c*) 3.2 eV. The temperature dependence of the surface reaction-limited deposition rate is exponential and follows the Arrhenius relation

$$R = Ae^{-E_a/kT} \tag{4.6}$$

where E_a is the activation energy in eV, T is the absolute temperature in K, and A is the frequency factor. One conclusion from this result is that good temperature control is even more important for low-temperature LRP reactions because of the higher activation energies and, consequently, the larger dependence of epitaxial growth on

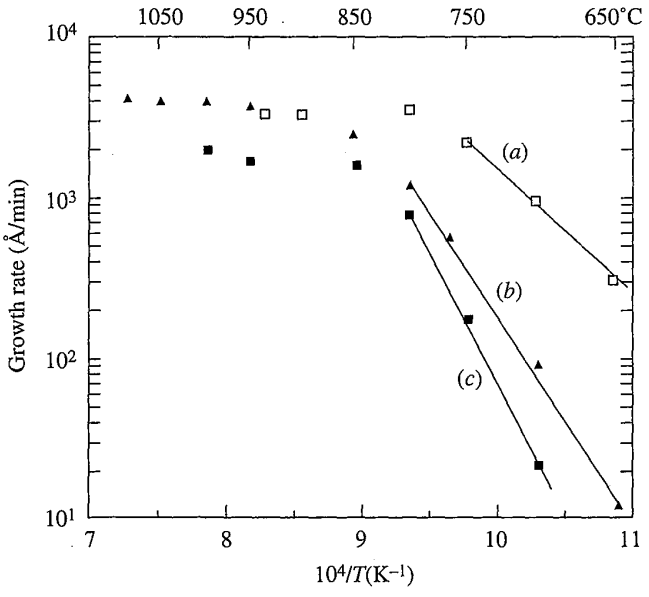


FIGURE 13
Silicon growth rate in an RTP reactor operating at 2 torr with 2 slm H_2 and (*a*) 40 sccm (standard cm^3/min) of SiH_4, (*b*) 80 sccm of SiH_2Cl_2, and (*c*) 20 sccm of SiH_4 and 2 sccm of HCl (*After Regolini et al., Ref. 36.*)

Micron Ex. 1023, p. 189
Micron v. YMTC
IPR2024-00909

temperature. However, temperature control and uniformity of today's RTP systems do not approach the levels of conventional systems, which is a major drawback for production control.

Doping profiles

The growth of thin epitaxial layers on heavily doped substrates requires that autodoping and outdiffusion from the substrate be minimized. Otherwise the required doping in the epi-layer can be lost by diffusion of impurities from the epi/substrate interface or from vapor-phase doping of impurities that have evaporated from the substrate. Diffusion decreases with temperature and time, and autodoping decreases similarly through lower gas-residence time, lower dopant surface segregation, and dopant evaporation rates. RTCVD has been shown to produce intrinsic Si layers grown on n^+ Si with results comparable to molecular beam epitaxy,[32] i.e., abrupt doping-transition regions. An example of boron delta-doping spikes obtained by epitaxial Si layer growth at 900°C in a UHV-RTCVD reactor is shown in Fig. 14.[38] The use of B_2H_6 with Si_2H_6 was preferred over SiH_4 in this application because higher growth rates could be obtained at lower growth temperatures, allowing higher throughput and better control of gas impurities. Precision temperature and time profile control allows the growth of such films with dimensional control within tens of angstroms.

The ability of RTCVD to produce abrupt n^+ doping layers has met with mixed success. Whereas abrupt i-n-i layers have been demonstrated using dichlorosilane

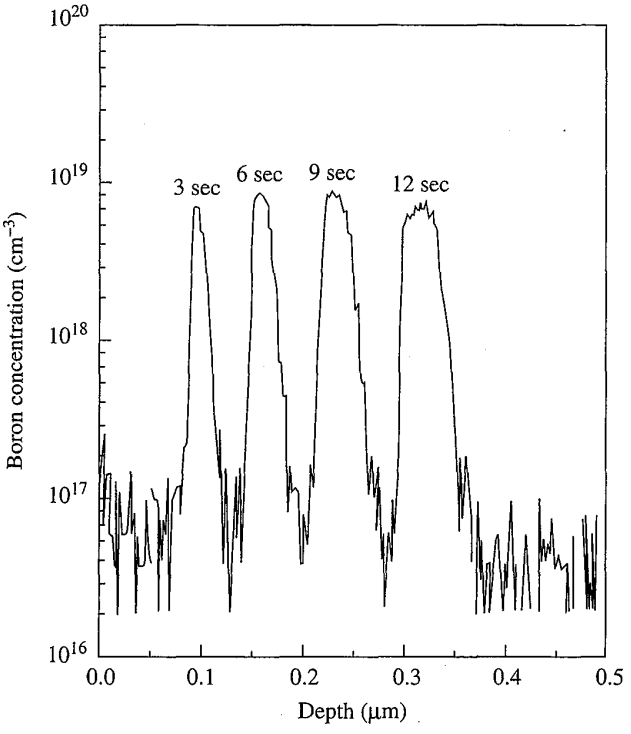


FIGURE 14
Boron delta-doping spikes obtained by epitaxial growth using Si_2H_6, B_2H_6, and H_2 in a UHV-RTCVD reactor. The growth times in seconds for each spike are shown. (*After Ozturk and Wortman, Ref. 38.*)

Micron Ex. 1023, p. 190
Micron v. YMTC
IPR2024-00909

and phosphine in a single-wafer reactor,[39] similar arsenic-doped layers have not been possible, because of the accumulation of arsenic (As) in the growing Si surface.[40] As the layer is grown, the As atoms available for doping the layer are slowly released, causing a gradual buildup of doping to its steady-state level.

4.3.2 Thin-Film Deposition

Chemical vapor deposition can be performed over a wide pressure range from atmospheric to ultrahigh vacuum. Low-pressure CVD (LPCVD), atmospheric CVD (APCVD), and ultrahigh-vacuum CVD (UHVCVD) all rely on the pyrolysis of source gases at elevated temperatures. At present, LPCVD is the preferred conventional deposition technology for many thin films because of high deposition rates and excellent film-thickness uniformity.

The application of RTP to CVD processes is particularly well-suited to single-wafer cluster tool technology. A typical design of an RTCVD system appears in Fig. 15, which shows a cross-sectional view of a RAPRO™ RTCVD chamber.[41] The chamber walls are water-cooled, thus eliminating deposition on the chamber walls. In LPCVD, the chamber walls are hot and in equilibrium with the rest of the system. Another difference between conventional CVD and RTCVD is the sequence of required tasks. The process sequences for LPCVD and RTCVD are shown in Fig. 16.[41] It can be seen that RTCVD requires higher temperatures but shorter times to deposit the same amount of material as LPCVD. In addition, temperature is the switch that turns the RTP deposition process on or off, avoiding the long ramp-up and ramp-down times required in the conventional method.

Oxide deposition processes

To be commercially viable, RTCVD systems must be able to process a single wafer in 1 to 2 min. The corresponding throughput is 30 to 60 wafers per hour. For relatively thick deposited films of 1000 to 2000 Å, this requires deposition rates greater than 1000 Å/min. However, conventional batch LPCVD processes are much slower (100 Å/min). Therefore, if one plans on transferring a recipe for an LPCVD deposition step to an RTCVD system, it will be necessary to raise the temperature in the RTCVD system to increase the system throughput. However, there are limits on the deposition rate for a given gas flow once the temperature is raised sufficiently, so the deposition kinetics become mass-transport-limited. For example, in Fig. 17 the mass transport region occurs when the deposition temperature exceeds 850°C, and further increases in temperature do not yield further increases in the rate.

As an example, consider the deposition of thick SiO_2 layers as either an interlevel dielectric between metallization layers or as a sidewall spacer on a polysilicon gate. These applications require oxides that are several thousands of angstroms in thickness. An important requirement for both applications is good oxide step coverage. The physical parameter indicative of step coverage is the sticking coefficient of the CVD process.[42] A small sticking coefficient of absorbed atoms means that they will tend to diffuse on the surface of the deposited layer or be reemitted. A CVD process with a small sticking coefficient provides better step coverage

Micron Ex. 1023, p. 191
Micron v. YMTC
IPR2024-00909

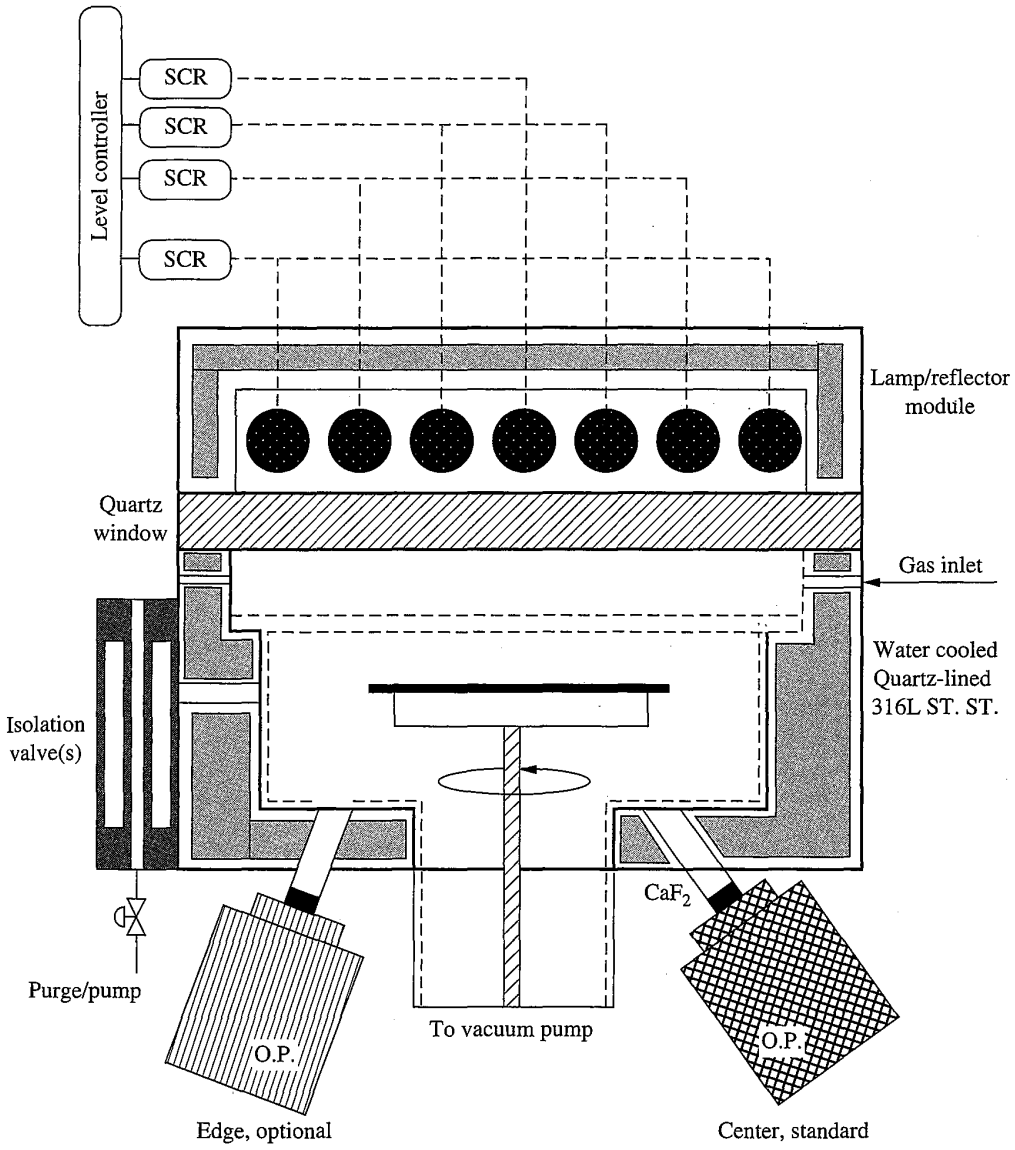


FIGURE 15
Cross-sectional view of the RAPRO™ RTCVD chamber (*After Ozturk, Ref. 41.*)

and conformality.[41] This requirement favors the use of TEOS (tetraethoxysilane or tetraethylorthosilicate) and TMCTS (tetramethylcyclotrasiloxane) in the CVD process. Thus the RTCVD of SiO_2 by pyrolysis of TEOS is believed to occur by the reaction[43]

$$Si(OC_2H_5)_4 \rightarrow SiO_2 + 2H_2O + 4C_2H_4 \tag{4.7}$$

The temperature dependence of the oxide deposition rate obtained in an RTCVD system according to Eq. (4.7) is shown in Fig. 18.[44] Below 800°C, the deposition rate is controlled by surface reaction processes with a 3.3 eV activation energy. This large sensitivity to temperature would require tight temperature control for thin

Micron Ex. 1023, p. 192
Micron v. YMTC
IPR2024-00909

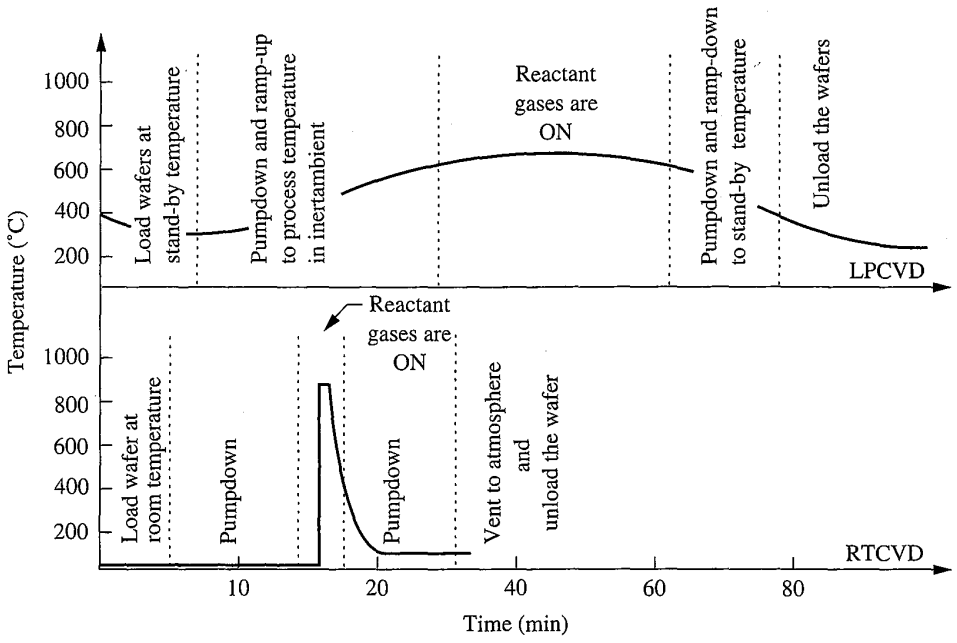


FIGURE 16
Process sequences for conventional LPCVD and RTCVD. The major difference in the two technologies lies in the way the deposition is initiated; in LPCVD, the gas flow is turned on and off for control, whereas in RTCVD, temperature is used for control with the gas flowing. (*After Ozturk, Ref. 41.*)

oxide deposition at low temperatures. Above 800°C, the deposition rate of SiO_2 approaches 1000 Å/min with lower activation energy, which meets the throughput and deposition control requirements of the RTCVD system. Such a high operating temperature, however, restricts the process to oxide deposition steps at the front end of the line. The process is unsuited for depositing interlevel SiO_2 at the back end of the line in multilevel-metallization technologies that use aluminum. For this purpose, a process is needed that can provide acceptable throughput at temperatures below approximately 450°C. Such an RTCVD process has not yet been found.[41]

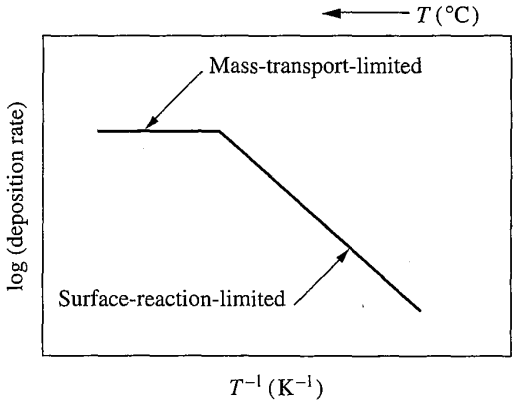


FIGURE 17
Temperature dependence of the deposition rate for a typical CVD process. (*After Ozturk, Ref. 41.*)

Micron Ex. 1023, p. 193
Micron v. YMTC
IPR2024-00909

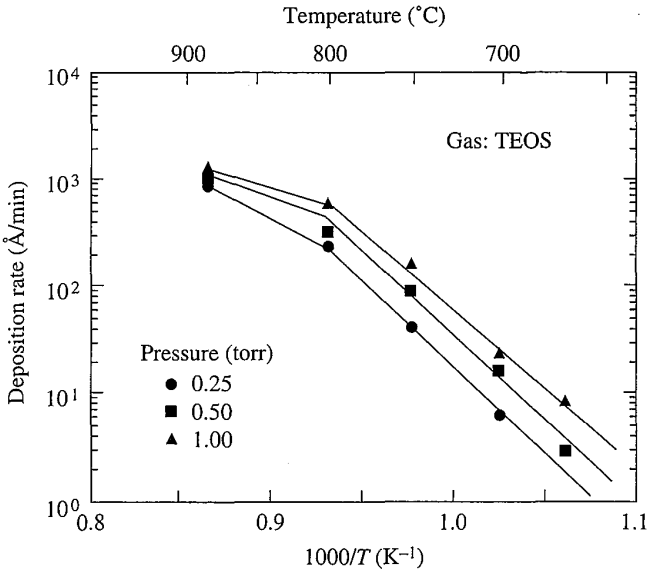


FIGURE 18
Temperature dependence of the SiO_2 deposition rate at different pressures using TEOS in an RTCVD reactor. The activation energy below 800°C is 3.3 eV. (*After Miller et al., Ref. 44.*)

Thin oxide applications include in-situ deposition of MOS gate structures where Si surface cleaning, gate oxide formation, and polysilicon gate electrode deposition would all occur in a low-pressure, multichamber cluster tool. This process would require the low-temperature deposition of SiO_2 as a substitute for the preferred, thermally grown oxide. However, oxidation or rapid thermal oxidation (RTO) is performed at atmospheric pressure, which would not be consistent with the low-pressure cluster tool approach. In addition, RTO is performed at temperatures exceeding 950°C, which requires process times up to 2 min. Therefore, both throughput and thermal-budget issues could become important for RTO. As a result, work is being done to bring the quality of low-temperature CVD oxides up to that attainable with thermal oxides.

RTCVD offers a promising technology for high-quality oxide deposition through the use of ultraclean gases in a chamber with a highly controlled ambient. Most conventional LPCVD furnaces use silane and oxygen in the oxide deposition reaction described in Eq. (4.8).[45]

$$SiH_4 + O_2 \rightarrow SiO_2 + 2H_2 \tag{4.8}$$

Typical deposition temperatures are in the 300 to 450°C range. However, Eq. (4.8) results in excessive particulate generation, which may contaminate the wafers. Thus, alternate gas combinations are being explored for RTCVD. One potential gate-quality RTCVD oxide-deposition process being studied utilizes SiH_4 and N_2O:[41]

$$SiH_4 + 2N_2O \rightarrow SiO_2 + 2N_2 + 2H_2O \tag{4.9}$$

At 800°C and with $SiH_4 : N_2O$ ratios less than 4%, stoichiometric films are obtained[46] with activation energies in the range of 1.5 to 1.75 eV and deposition rates around 50 Å/min, which are nearly 100 times larger than the thermal oxidation rate at atmospheric pressure. The films also have acceptable electrical properties.

Micron Ex. 1023, p. 194
Micron v. YMTC
IPR2024-00909

Silicon nitride deposition

Silicon nitride has found important uses in integrated circuit processing: as a mask against O_2 diffusion during local oxidation, as a passivation layer because it is a barrier to contaminants, as a gate dielectric in MNOS memory transistors, and a new use as an interlevel dielectric in oxide-nitride-oxide stacked-gate structures.[47]

Deposition of Si_3N_4 in conventional systems involves the reaction of ammonia with either silane in APCVD or dichlorosilane in LPCVD. This latter reaction has become the preferred approach and is described as follows:

$$3SiH_2Cl_2 + 10NH_3 \rightarrow Si_3N_4 + 6NH_4Cl + 6H_2 \tag{4.10}$$

This reaction is unsuitable in a cold-wall RTCVD system. The reaction by-product, ammonium chloride (NH_4Cl), deposits in the form of a fine powder on cool surfaces in the chamber. Although this is a minor concern in hot-wall LPCVD systems, it would be a major problem in an RTCVD system, in which deposits would occur essentially everywhere, including the quartz windows and chamber walls.[41]

Alternative reactions have been studied for application to RTCVD systems, including the use of SiH_4 and NH_3.[48] This reaction can be described as

$$3SiH_4 + 4NH_3 \rightarrow Si_3N_4 + 12H_2 \tag{4.11}$$

It is interesting that Eq. (4.11) has been tried in LPCVD systems with little success. Wafers with deposited nitride films showed radial nonuniformities in which less material was deposited in the wafer center than on the perimeter. This "bullseye" effect was believed to be caused by mass-transfer deficiencies.[41] However, the reaction in Eq. (4.11) works well in a single-wafer RTCVD system, where reactant-depletion effects are not observed.[48]

It has been observed[48] that stoichiometric nitride films are obtained for depositions using NH_3 : SiH_4 ratios greater than 120 : 1. The Si_3N_4 deposition rate versus NH_3 : SiH_4 reactant gas ratio at 785°C is shown in Fig. 19. The useful deposition rates at high flow ratios are very small (100 Å/min), and therefore this process is

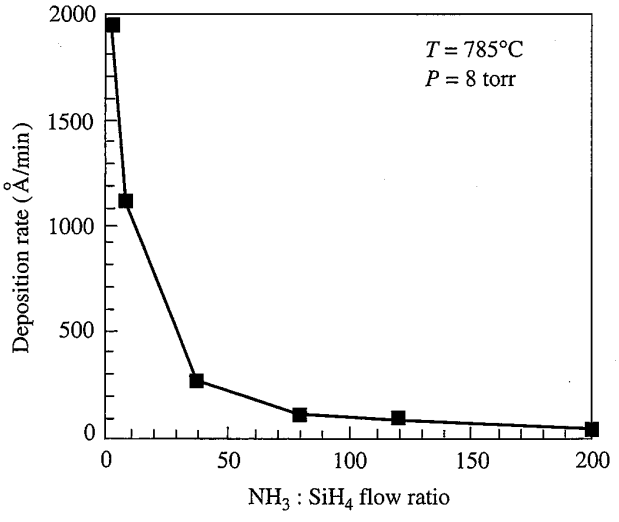


FIGURE 19
Si_3N_4 deposition rate versus reactant gas ratio. Deposition conditions were NH_3 and 10% SiH_4 in Ar at 785°C and a total pressure of 8 torr. (*After Johnson et al., Ref. 48.*)

Micron Ex. 1023, p. 195
Micron v. YMTC
IPR2024-00909

suited only for thin-film applications such as stacked-gate structures. However, no suitable single-wafer RTCVD processes exist yet for typical thick-film uses of silicon nitride.

Polycrystalline silicon deposition

Polysilicon is the preferred gate electrode material in ULSI MOS devices. It is also used as an interconnect layer with a low-resistivity, silicide-forming metal such as titanium or cobalt. In addition, polysilicon resistors are used in analog ICs as well as in static random-access memory (SRAM) chips. In conventional LPCVD systems, polysilicon is deposited by the pyrolysis of silane:[49]

$$SiH_4(gas) \rightarrow SiH_4(absorbed) \rightarrow Si + 2H_2 \qquad (4.12)$$

Deposition usually occurs around 600°C at 0.2 to 1 torr with rates of 100 to 200 Å/min. Since these rates are too low for single-wafer processing, it becomes necessary to raise the temperature of the reaction above 700°C with the goal of attaining rates in the 2000 Å/min range. This approach, however, has been tried in LPCVD systems with the result that higher deposition temperatures resulted in polysilicon films with rough surfaces.[50] Conversely, very smooth polysilicon films have been deposited in RTCVD systems at elevated temperatures and pressures. The main factors that control polysilicon deposition rates are temperature and the partial pressure of the reactants. For example, the temperature dependence of deposition using 10% SiH_4 in a carrier gas of Ar is shown in Fig. 20.[49a] The surface reaction activation energy of 1.7 eV is similar to that achieved in LPCVD systems, and Fig. 20 shows that 1000 Å/min rates are possible at 700°C. The figure also

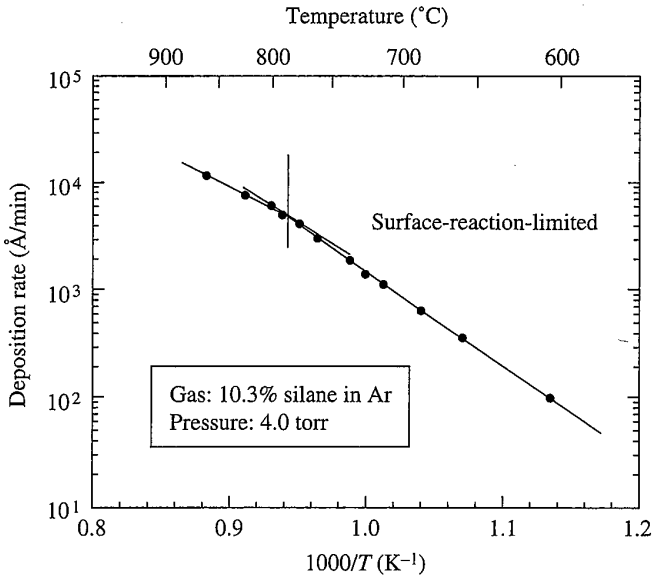


FIGURE 20
The deposition rate of polysilicon from SiH_4 in an RTCVD reactor. Silane was diluted with Ar to 10.3%. (*After Ozturk et al., Ref. 49a.*)

Micron Ex. 1023, p. 196
Micron v. YMTC
IPR2024-00909

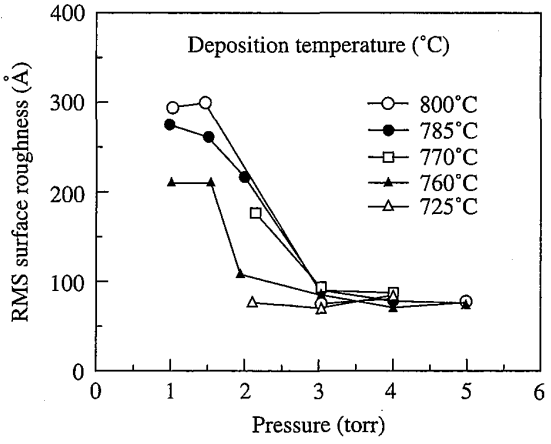


FIGURE 21
Root-mean-square surface roughness of 2000-Å-thick polysilicon deposited on SiO_2 at several temperatures and pressures. Data were obtained using either cross-section TEM or UV surface-reflectance measurements. (*After Ren et al., Ref. 51.*)

shows that there is no true mass-transport-limited deposition region above 780°C. With regard to surface roughness of the polysilicon films, it appears that very smooth films can be achieved in RTCVD systems at deposition temperatures above 725°C and at elevated pressures (above 3 to 4 torr).[51] Surface roughness measurements on 2000-Å-thick films are shown in Fig. 21, where root-mean-square roughness data measured on 2000-Å-thick polysilicon layers are illustrated. These data were obtained from either cross-section TEM or UV surface-reflectance measurements.

Perhaps the biggest challenge for RTCVD systems designed to deposit polysilicon is film-thickness uniformity. Deposition requirements for ULSI manufacturers range from ±1% to ±3% film uniformity across a 200-mm wafer. The influence of the deposition system on uniformity depends on the process chemistry through the activation energy. If a process has zero activation energy, then there is no temperature dependence and perfect uniformity is possible. Assuming a 1.6-eV process for polysilicon deposition, the deposition uniformity as a function of the percentage of temperature uniformity across a wafer is plotted in Fig. 22 for three temperatures.[51a] An assumed ±3% film uniformity extrapolates to ±0.2% temperature uniformity.

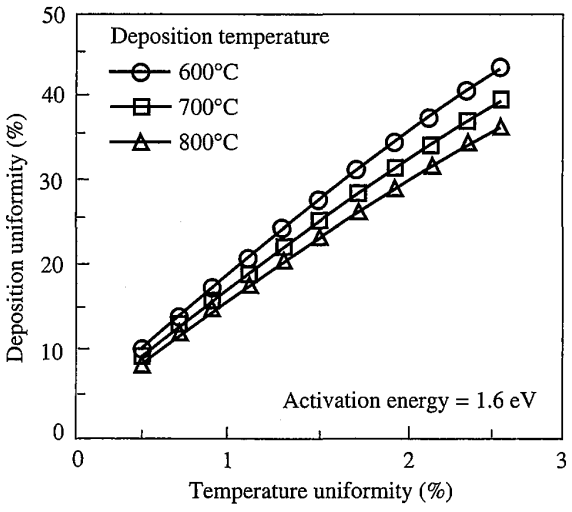


FIGURE 22
Polysilicon deposition-thickness uniformity versus the percentage of temperature uniformity across the wafer for three deposition temperatures and an activation energy of 1.6 eV. (*After Ozturk et al., Ref. 51a.* © 1991 IEEE)

Micron Ex. 1023, p. 197
Micron v. YMTC
IPR2024-00909

This means that at 700°C the temperature would have to be held to within ±1.4°C across the wafer. Only a very few RTP tools have this level of control today, whereas this level of control is routine in conventional furnaces.

4.3.3 Oxidation

Rapid thermal oxidation (RTO) technology is well-suited to the growth of thin, high-quality dielectric layers in ULSI devices. An ideal RTO system is also an ideal RTP system, but with some additional requirements.[52] The system must be designed to handle reactive gases without introducing contamination or particles onto the wafers, and the gas-handling system must be vacuum-compatible and capable of switching quickly from one process ambient to another. In addition, the heating system must have a radiation spectrum that is compatible with the absorption properties of the wafer to allow uniform heating across the wafer.

The primary issues that differentiate RTO from conventional thermal oxidation are the more complex chamber design in RTO compared to the quartz tube in a furnace, the design of the radiation source, and the temperature measurement and calibration. From the point of view of oxide-growth kinetics, it is believed that oxide growth during RTO may be influenced by both thermally activated processes and a nonthermal, photon-induced component observed in the ultraviolet and visible range of the optical spectrum.[53] The basic photochemistry is described by the following reactions:[54]

$$O_2 + h\nu \rightarrow O + O \tag{4.13a}$$

$$O + O_2 + O_2 \rightarrow O_3 + O_2 \tag{4.13b}$$

$$O + O_3 \rightarrow 2O_2 \tag{4.13c}$$

$$O_3 + h\nu \rightarrow O + O_2 \tag{4.13d}$$

Enhanced Si oxidation has been observed at 450 to 550°C resulting from the presence of UV radiation. The cause of this is speculated to be the absorption of UV light by the oxygen gas, leading to photodissociation and the production of the species in Eq. (4.13). The highly reactive, monatomic O atoms produced are believed to create a parallel oxidation reaction that dominates at low temperatures.[54] However, it is not known whether this reaction is important at normal RTO temperatures above 900°C.

Thermally induced stresses in wafers processed in RTO systems have also been shown to affect thin-oxide growth[55] by up to 10%. It has been found that regions of the Si wafer in tensile stress yield thicker oxides than compressive regions. A tensile stress in Si stretches the Si atom-to-atom spacing, making oxidation easier. In conventional oxidation furnaces, thermal stresses are controlled by uniform heating and slow-temperature ramping.

Oxide growth comparisons

An example of a two-step RTO process is shown in Fig. 23.[56] The wafer is ramped up in O_2 to a preset temperature below the oxidation temperature, and then ramped again to the oxide growth temperature. A second, high-temperature, rapid

Micron Ex. 1023, p. 198
Micron v. YMTC
IPR2024-00909

thermal annealing (RTA) step is performed in N_2 to anneal and improve the electrical properties of the Si-SiO_2 interface. Figure 23 compares operation of the system under two temperature control methods: pyrometer temperature control and preset control of the heating source power. In conventional furnaces, the temperature is fixed and the wafers are inserted under N_2 ambient conditions. Switching the O_2 gas flow on and off determines the duration of oxidation.

RTO growth kinetics exhibit activation energies that differ from those measured in conventionally grown oxides. In the initial 20 seconds, the RTO growth is observed to be linear followed by nonlinear growth.[57] The duration of the linear region is found to depend on the type of heating lamp used (tungsten halogen vs. water-cooled arc), and the linear region showed activation energies of 1.44 eV and 1.71 eV, respectively, for the two lamps. This result compares to a linear-rate activation energy of 1.76 eV reproduced in conventional furnaces. Numerous other studies of linear-growth kinetics in RTO systems have been conducted, each producing different activation energies. The origin of these differences has not been modeled, and work in this direction would not be warranted now because of uncertainties about wafer temperatures in different RTO systems. As an example, discrepancies in RTO growth rates on (100) Si in dry O_2 at 950°C are shown in Fig. 24.[52] The data were taken from a number of published reports in which these specific oxidation conditions were used. For comparison, the analytical growth relationship of Massoud and Plummer,[58] using rate constants for conventional oxidation, is also plotted at 950°C and at 1000°C. Figure 24 shows that only one RTO growth experiment was consistent with the calculated 950°C curve. This result implies an uncertainty of up to 50°C in the RTO systems used.

Temperature control in closed-loop RTO systems is particularly difficult. Such systems use pyrometry as a temperature-sensing method, but this method is viable only if the effective emissivity of the wafer is known. Section 4.1.1 stated that many

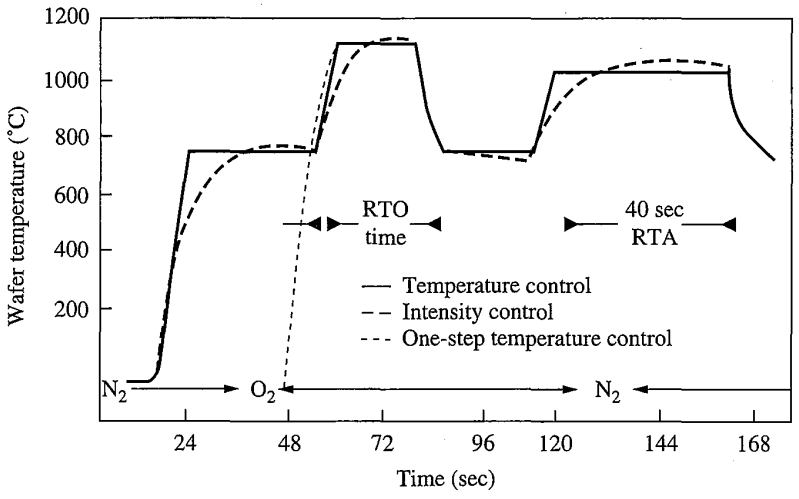


FIGURE 23
Example of a two-step RTO cycle showing wafer temperatures and ambients for RTO followed by RTA. (*After Nulman, Krusius, and Renteln, Ref. 56.*)

Micron Ex. 1023, p. 199
Micron v. YMTC
IPR2024-00909

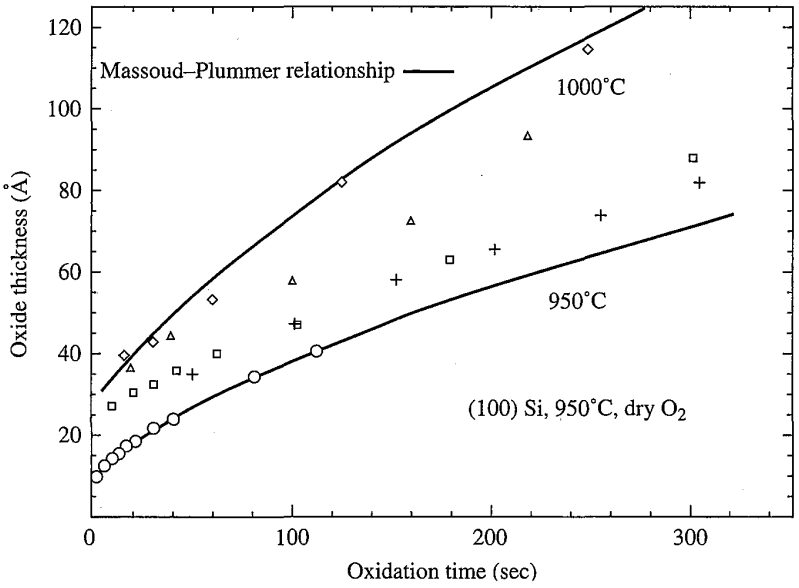


FIGURE 24
Discrepancy in RTO growth kinetics on (100) Si in dry O_2 at 950°C. Data symbols represent oxide growth in a number of different RTO systems. Solid lines are the analytical thermal-oxidation-growth curves calculated from the model of Massoud and Plummer.[58] (*After Massoud, Ref. 52.*)

factors influence emissivity, such as absorption in the ambient gas and surface film thicknesses. Wafer emissivity also depends on the presence of dielectric films which may have been deposited or grown on the reverse of the wafer during processing. The layers act as interference filters, leading to constructive or destructive interference of the emitted radiation in the wavelength range of the pyrometer. This causes a critical emissivity dependence on thickness, stacking sequence, and optical constants of these layers.[14] An example of this effect is shown in Fig. 25. RTO was performed on low-doped Si wafers using closed-loop pyrometric feedback. A six-layer stack of deposited films was put on the reverse side of the wafer in the following sequence:[14]

1. Oxide—4900 Å
2. Nitride—1600 Å
3. Oxide—variable
4. Polysilicon—2100 Å
5. Oxide—4900 Å
6. Nitride—250 Å

RTO was performed on the bare, front side for layer-three oxide thicknesses of 4000 to 8300 Å, and the resulting RTO oxid thicknesses are shown in Fig. 25. The calculated temperatures corresponding to the oxide thickness grown vary by 1081 ± 85°C because of the changes in layer-three oxide thickness. Significant improvements in wafer temperature control are possible when the RTO system is operated in a power-control mode in which lamp power, cooling air flow, and optical characteristics of

Micron Ex. 1023, p. 200
Micron v. YMTC
IPR2024-00909

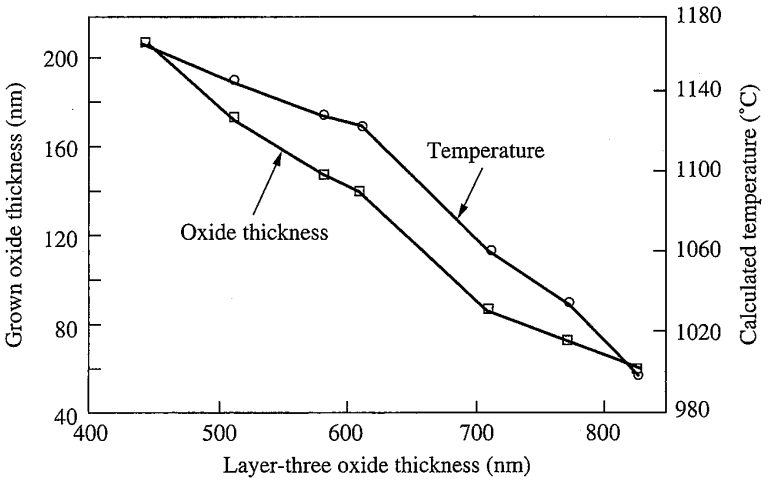


FIGURE 25
Rapid thermal oxidation results obtained with closed-loop pyrometric feedback on the front side of a wafer with a six-layer stack of deposited films on the reverse. The third deposited oxide layer thickness is the variable. Emissivity variations result in significant wafer-temperature changes and RTO film thickness. (*After Nakos, Ref. 14.*)

the system are controlled. Indeed, ±3.5°C temperature control was achieved in the same experiment described above under a power-control mode.

Properties of rapid, thermally processed oxides

Rapid thermal processing is used to grow dielectric films as well as to postprocess the films to impart the desired chemical or electrical properties. For example, in Fig. 23 we saw a two-step RTO process followed by an RTA process to improve the interface electrical properties. Other RTP processes used in thin-gate dielectric formation for MOS devices include rapid thermal nitridation (RTN), rapid thermal oxynitridation, and all the RTCVD processes previously discussed.

Oxides annealed in an RTA system generally possess electrical and dielectric properties that are comparable to or better than those of furnace oxides.[52] Parameters included in this statement are oxide fixed charge, interface traps, and charge-to-breakdown levels, Q_{BD}. Indeed, it has been shown that Q_{BD} increases from 20 C/cm^2 for conventional furnace oxides to 80 C/cm^2 for RTO-RTA oxides of 100 Å thickness.[59] The charge-to-breakdown is directly related to the trapping of electrons that are injected into the oxide. A higher Q_{BD} implies that RTO-RTA oxides have fewer trapping centers, resulting in better reliability. In studies of the annealing of interface traps by RTA, it was shown that the RTA temperature was dominant in reducing traps. Since RTA can be performed in the RTO chamber at higher temperatures than postfurnace annealing, the expected improvement in interface traps is observed.[60]

High-frequency, quasi-static capacitance-voltage measurements on oxides yield interface trap distributions and oxide-fixed charges at the Si-SiO_2 interface. A comparison of interface trap distributions in furnace oxides (FO) and RTO oxides is

Micron Ex. 1023, p. 201
Micron v. YMTC
IPR2024-00909

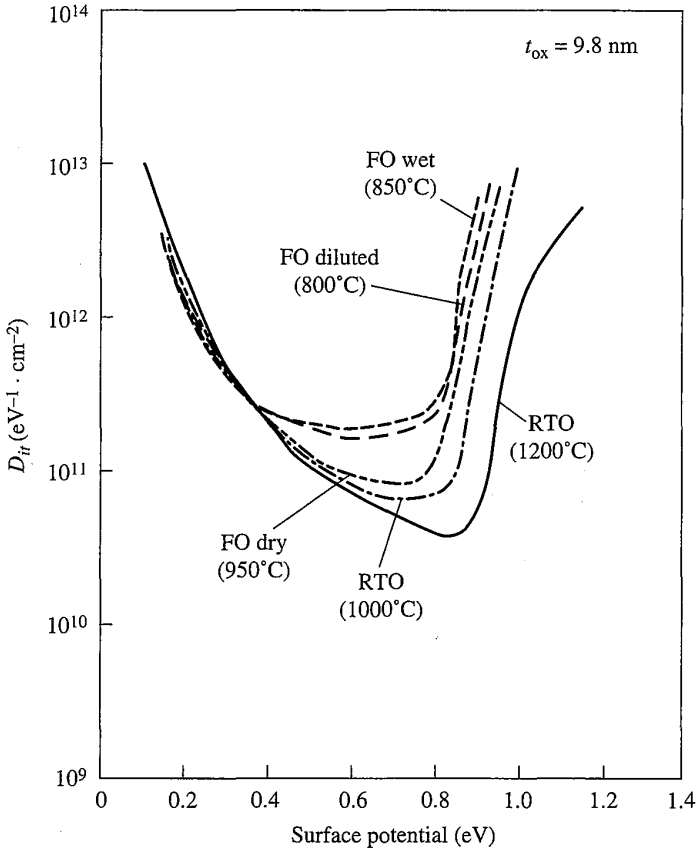


FIGURE 26
Comparison of interface-trap distributions in furnace oxides (FO) and RTO oxides. (*After Fukuda, Ref. 61.*)

shown in Fig. 26.[61] It should be pointed out that the results shown in Fig. 26 are not state-of-the-art, but do illustrate the possible advantages of RTO over furnace annealing. The growth temperatures and ambient conditions for FO growth are noted in the figure. All RTO processes were performed in dry O_2. The higher RTO temperature results in lower mid-gap trap levels, but the trap densities for both FO and RTO are similar at lower temperatures. Thus, the apparent advantage of RTO over FO may occur only because of the higher oxidation temperatures used in RTO processing, which cause a reduction in the amount of unoxidized, excess Si at the Si-SiO_2 interface and, thus, the fixed charge.

RTO oxides seem to exhibit improved dielectric breakdown properties over FO oxides. Time-zero dielectric breakdown (TZDB) testing involves applying a negative gate voltage to an MOS capacitor and recording the breakdown voltage when the measured current is 1 μA. The TZDB distributions of FO and RTO oxides of the same thickness are compared in Fig. 27. The FO samples were grown at 950°C in dry O_2 and the RTO samples were grown at 1000°C in dry O_2.[61] The TZDB data for the RTO samples are better in two respects. First, no low- or medium-field breakdowns (1 to 8 MV/cm) were seen; second, the maximum breakdown frequency occurs at higher fields than in the FO samples.[52, 61] Both of these observations are related to

Micron Ex. 1023, p. 202
Micron v. YMTC
IPR2024-00909

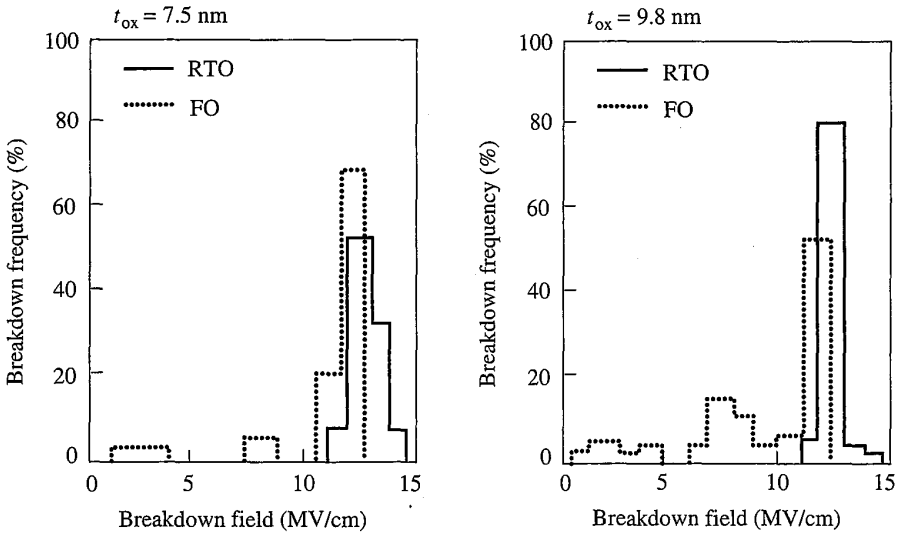


FIGURE 27
Time-zero dielectric breakdown (TZDB) histograms of FO and RTO MOS capacitors with 7.5- and 9.8-nm oxide thickness. (*After Fukuda, Ref. 61.*)

the lower charge-trapping properties of RTO oxides. It is believed that there is a reduction in strained Si–O–Si and Si dangling bonds at the Si-SiO_2 interface when the RTO oxide is grown at a higher temperature.

4.3.4 Rapid Thermal Annealing for Junction Formation and Defect Removal

It has been shown that RTA differs from furnace annealing in the way the wafers are heated (radiative versus convective and conductive) and in the rates at which wafers are heated and cooled (10 to 200°C/sec versus 5 to 50°C/min, respectively). Even though RTA rates are rapid and optical excitation of the semiconductor produces large numbers of electron-hole pairs, the relaxation of the excess free carriers is much shorter than the temperature ramp rates. Indeed, it has been shown theoretically that, under high-intensity light irradiation, nonradiative processes limit electron-hole recombination times to the 10^{-9} to 10^{-12}/sec range. For these values, it has been found that an excess carrier density will be too small to change the position of the Fermi level in Si at 1000°C.[62] Thus, carrier levels are considered to be at equilibrium, and processes such as impurity diffusion and defect annealing in the semiconductor are believed to be in thermal equilibrium during the RTA temperature cycle.

The application of RTA to producing pn junctions by annealing ion-implanted layers has been studied since 1983. However, the semiconductor industry has put RTA into limited production because only a few vendors of equipment have provided systems that meet the production requirements for reproducibility and uniformity, and these systems were not available until 1992. Noncritical processing steps with wide processing latitude, such as silicide formation, have found their place in production. However, there have been more fundamental reasons why RTA has

Micron Ex. 1023, p. 203
Micron v. YMTC
IPR2024-00909

not gained widespread acceptance in processes such as junction formation. Some of these are listed below:

High RTA temperatures produce slip lines in wafers caused by thermal stresses.
RTA-annealed, ion-implanted junctions show excess diffusion for a given thermal budget that is dominated by defect annealing rather than by thermally assisted diffusion.
The excess diffusion of dopants during RTA has not been modeled adequately to allow engineers to design devices and processes with confidence.
Localized heating on patterned wafers causes higher wafer temperature than on bare control wafers, making process control difficult.

Nevertheless, RTA has several important advantages over furnace systems:

Low thermal budget control for shallow junctions
High level of dopant activation and defect annealing

In the following two sections, the application of RTA to the annealing of ion implantation defects and to the diffusion of dopants for junction formation is discussed.

Defect annealing

Nuclear-stopping processes of ions implanted into Si are responsible for producing displacement damage. The types and amount of damage produced depend on implant species, energy, dose, wafer temperature and orientation, dose rate, and materials covering the substrate. The as-implanted defect morphologies then change during wafer annealing, depending on temperature, time, furnace ambient, and ramp-up/ramp-down rates. Many of the point-defects (vacancy and self-interstitial defects) are generated during the changes from as-implanted damage to stable or dissolved-damage structures as a result of annealing. The impact of the damage evolution during annealing increases as the lateral and vertical dimensions of Si devices decrease, because interactions between point defects and extended defects that form during annealing dominate dopant diffusion. Thus, it is important to have doping and annealing technologies that deal with these effects and produce defect-free junctions of the desired depth and controllability.

Depending on the dose and species, ion implantation can produce different types of damage, as shown schematically in Fig. 28. For example, when the implant dose is below a critical level required for the formation of extended defects such as dislocations (at a critical dose level of approximately 2×10^{14}ions/cm^2), the dominant damage species are isolated point defects or point-defect clusters. Figure 28*a* shows an example in which I-type (interstitial) defects are distributed throughout the depth of the implanted region and beyond.

Defects also arise whenever an amorphous layer is generated by a high-dose implant, as shown in Fig. 28*b*. V-type (vacancy) defects are observed forming near the Si surface. I-type defects occur outside the amorphous region and are probably interstitial clusters. Type II dislocation loops are located just beyond the amorphous/crystal interface in the heavily damaged, but still crystalline, substrate. They are also called end-of-range (EOR) damage.[63] The EOR defects may be significant to successful junction formation if they are present in the resulting space-charge

Micron Ex. 1023, p. 204
Micron v. YMTC
IPR2024-00909

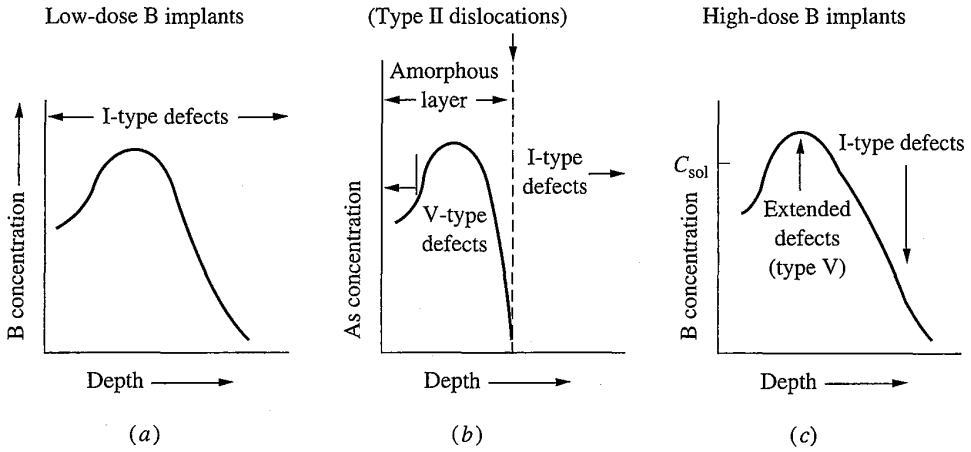


FIGURE 28
Ion-implantation defect production for three classes of implants, showing dopant concentrations versus depth. I-type defects are of the self-interstitial type and V-type defects are of the vacancy type. (*a*) I-type defects are distributed through a lightly implanted B layer. (*b*) An amorphous layer forms after high-dose implantation—V-type defects form in the Si surface, I-type defects form in the crystalline region beyond the amorphous layer, and dislocation loops form at the amorphous-crystalline interface. (*c*) Projected range dislocations are shown at the peak of a high-dose B implant, and I-type defects are shown in the Si bulk when the peak B concentration exceeds the B solubility limit, C_{sol}. (*After Fair, Ref. 66.*)

region, and their presence may affect the diffusion process occurring during RTA. Other types of defects related to the amorphous-layer regrowth may be avoided by choosing implantation and annealing conditions carefully, but the EOR defect is unavoidable once an amorphous layer is formed. Therefore, when the dose exceeds the amorphization threshold (1×10^{14} to 1×10^{15}ions/cm^2), implantation at room temperature of most Group III and V elements would result in extrinsic dislocation loops upon annealing.

The third type of implantation defect, shown in Fig. 28*c*, is the projected-range defect, which is caused by the formation of precipitates of the implanted species distributed near the projected range when solid solubility at the annealing temperature is exceeded. These defects are most noticeable in high-dose B implants where no amorphous layer is formed. The defects are in the form of extrinsic dislocation loops extending to the Si surface. It is difficult to remove these defects by annealing since they are very stable at temperatures up to 1050°C. During annealing by RTA or conventional furnace, self-interstitial defects are generated by the dissolving projected-range dislocations, causing enhanced diffusion for extended periods.

Annealing characteristics of EOR defects may be the most important process to understand for ULSI, since the preamorphization of the Si surface is performed before shallow junctions form, to prevent dopant channeling during low-energy implantation. In addition, high-dose implantations to reduce sheet resistance are likely

Micron Ex. 1023, p. 205
Micron v. YMTC
IPR2024-00909

to form surface amorphous layers when the implantation is done in a crystalline substrate. The location and annealing properties of the EOR defects are important factors in determining the implanted dopant diffusion and the quality of implanted junctions. The annealing rate and the inverse of the time required to completely remove the loops were found to have approximately 5 eV activation energy, indicating that the self-diffusion of Si atoms limits the annealing process.[63–65] This high activation energy is very significant and is greater than the activation energy for dopant diffusion, which is 3.5 to 4.2 eV. Thus, effective removal of the defects with minimum dopant diffusion would occur at high temperatures for short times, favoring RTA[66] over furnace annealing.

RTA of ion-implanted junctions

Any modern junction formation process must be capable of creating "good quality" pn junctions without excessive process complexity. In a "good quality" junction the final junction depth must be controllable and reproducible. These characteristics are important for manufacturing. Additionally, the pn junctions must have low reverse-bias leakage current and good (near ideal) forward characteristics. This requirement translates into a need for well-controlled, junction-region defect densities. Both the number and the location of any process-induced defects (i.e., EOR damage from implantation) are important. Diffused or implanted regions must have low sheet resistance. Thus, high carrier mobility and a high level of dopant activation are required. Finally, there are compatibility issues. An advanced junction process should be as compatible as possible with existing, well-established processing steps. A suitable contact technology must be available, either independent of the junction formation process or as a part of it. Thus, the designer is faced with numerous tradeoffs in what is described in Fig. 29 as the "junction processing circus."[65] Implant-damage annealing may increase dopant diffusivity and decrease dopant electrical activation, whereas increased activation increases diffusivity in additional ways, with solid solubility levels determining the rates of damage dissolution.

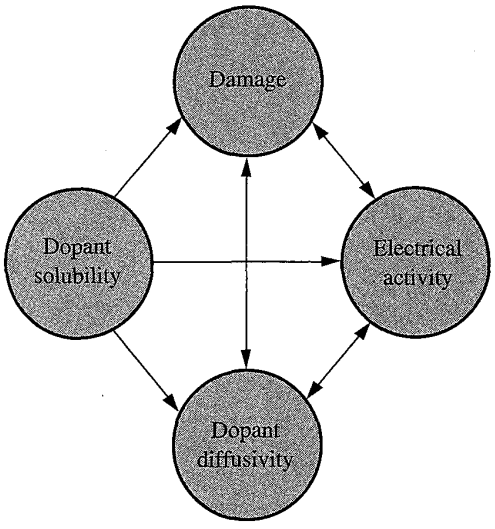


FIGURE 29
The junction processing circus, otherwise known as the implantation-diffusion interaction matrix, showing the interactions between damage, dopant solubility limits, dopant diffusivity, and electrical activity of the implanted dopants. (*After Jones, Ref. 65.*)

Micron Ex. 1023, p. 206
Micron v. YMTC
IPR2024-00909

It has been widely observed that the diffusivity of dopants in Si may be greatly enhanced during implant-damage annealing because of the generation of point defects. The extent of the observed transient diffusion enhancement is believed to be related to which point defects are required by a dopant for diffusion (vacancies or self-interstitials) and which of the two types of point defects is generated. For example, B requires Si self-interstitials for diffusion, whereas As diffusion is predominantly vacancy-based.

The nuclear collisions between incoming, energetic ions and the Si lattice atoms produce atomic displacements and, subsequently, vacant lattice sites (vacancies). The displaced Si atoms become energetic self-interstitials that displace additional lattice atoms and produce cascades of Si self-interstitials. The distributions of vacancies and self-interstitials throughout the implanted region follow the general shape of the ion-implanted impurity distribution. An example of vacancy and interstitial distributions following an ion implantation is shown in Fig. 30. The distributions were determined using Boltzmann transport damage calculations and also show the net defect distribution, which is obtained by subtracting the vacancy concentrations from the self-interstitial concentrations.[67] It can be seen that vacancy-rich and interstitial-rich regions are predicted, and this result has been verified experimentally. However, the self-interstitials can move in Si at low temperatures and tend to clump together to form clusters during the implantation step, causing a distortion in the interstitial profile shown in Fig. 30. As soon as the Si wafer is heated, these interstitial clusters

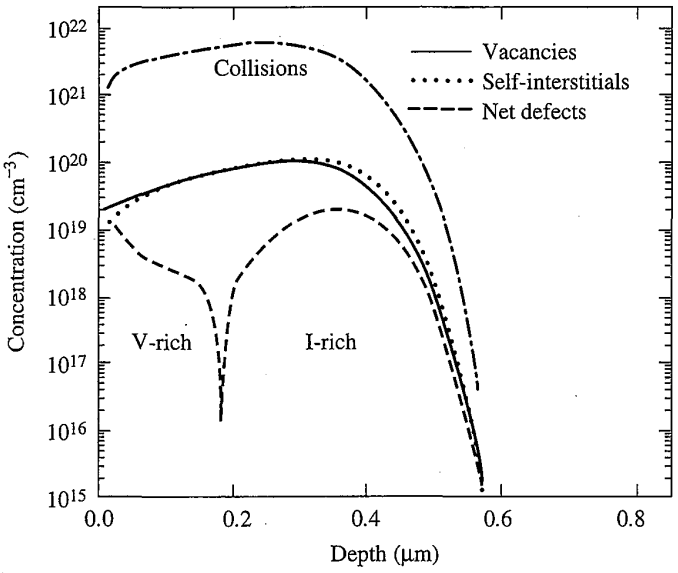


FIGURE 30
Ion implantation damage calculations, calculated with the Boltzmann transport equation, to determine production of vacancy and self-interstitial distributions. The total distribution of collisions, the resulting distributions of vacancies (V) and self-interstitials (I), and the net defects are shown. (*After Giles, Ref. 67.*)

Micron Ex. 1023, p. 207
Micron v. YMTC
IPR2024-00909

begin to dissolve, producing an excess interstitial transient that can enhance the diffusion of dopants such as B that are transported in Si via an interstitialcy mechanism.

From Fig. 28*a,* low-dose implants introduce isolated interstitial-type cluster defects (the origin of which has been described). *Low-dose,* in this case, refers to implants below 2×10^{14}ions/cm^2, which is the threshold dose for creating extended dislocations in Si. Such low-dose implants might be used in setting the base doping in a bipolar transistor process or in doping the channel of an MOS transistor. The number of such isolated defects, $N(E)$, depends on the energy, E, and dose of the implant through the production of displaced atoms for each incident ion:

$$N(E) = \frac{E_n}{2E_{oo}} \tag{4.14}$$

where E_n is the energy that goes into nuclear collisions and E_{oo} is the energy required to displace a Si-lattice atom. Thus, if 25 eV is required to displace a Si-lattice atom, a single 50 keV ion would produce 1000 displaced atoms along its trajectory in the Si. Some of these displaced atoms coalesce in stable clusters of various sizes and remain until the Si is heated, whereupon they dissolve.

The energy effect on the time-averaged pn-junction displacement can be seen for low-dose phosphorus (P) implants in Fig. 31.[66] The Δx_j data represent the difference between the initial, as-implanted junction depth at a level of 1×10^{17}atoms/cm^3, and the final depth of the profile after annealing. Data are shown for both RTA and furnace anneals. The diffusion coefficient of P, D_p, is directly proportional to the number of point defects produced, $N(E)$, and the junction displacement is proportional to $D_p^{1/2}$. This accounts for the $E^{1/2}$ dependence of the curves in Fig. 31. The magnitude of the displacement is defined at a P concentration of 1×10^{17} atoms/cm^3. Similar results are observed for B implants. Both RTA- and furnace-annealed samples show damage-assisted diffusion with the same $E^{1/2}$ dependence. The difference in slopes between the RTA curve and the furnace annealing curve is due to the differences in temperatures of the anneals and associated damage annealing times,

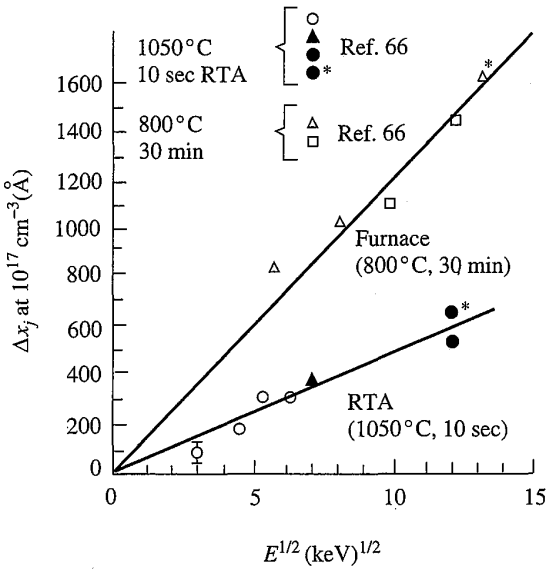


FIGURE 31
Transient junction displacement during RTA and furnace annealing for low-dose, phosphorus implants versus ion-implantation energy. Both annealing methods show an $E^{1/2}$ dependence on junction displacement. (*After Fair, Ref. 66.*)

Micron Ex. 1023, p. 208
Micron v. YMTC
IPR2024-00909

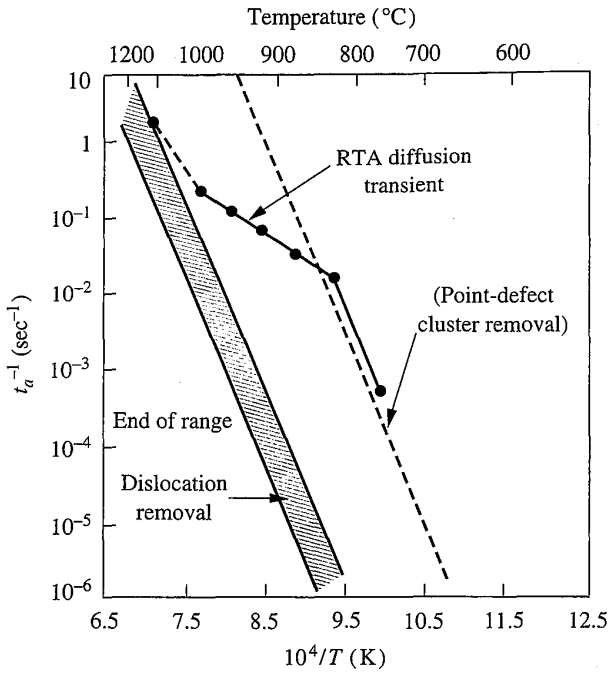


FIGURE 32
Measured boron diffusion transients during RTA and damage annealing times in ion-implanted Si. The annealing of both extended dislocations and clusters shows a 5-eV activation energy, and the diffusion transient data make a transition from domination by point-defect generation from extended dislocations at high temperatures, to domination by point-defect cluster dissolution at low temperatures. (*After Fair, Ref. 71.*)

t_a. This implies that the $D_p \cdot t_a$ product is lower for RTA even though the RTA was performed 250°C higher in temperature.

To understand this anomalous result, note that during normal, thermally assisted diffusion of low-concentration P, D_p is given by the following expression that has been obtained from experimental data:[68]

$$D_p = 3.85 \exp(-3.66 \text{ eV}/kT) \text{ (cm}^2\text{/sec)} \tag{4.15}$$

During damage-assisted P diffusion, the activation energy is reduced from 3.66 eV to approximately 2.2 eV as measured in transient-diffusion experiments:[69]

$$D_p(\text{enhanced}) = 0.007 \exp(-2.2 \text{ eV}/kT) \tag{4.16}$$

It is believed that the 2.2 eV may be the migration energy of a phosphorus self-interstitial pair. In the presence of excess self-interstitials caused by ion-implantation damage annealing, the diffusion-activation energy is reduced by approximately the amount of energy needed for the Si crystal to thermally generate equilibrium levels of self-interstitials. Estimates of this energy vary from 1.5 to 2.5 eV.

The times for damage annealing, t_a, can be measured by studying the dissolution of damage for various annealing time increments with an electron microscope. It is also believed that t_a is related to the time constant for transient dopant diffusion. The time constants for transient diffusion of low-dose B implants (1×10^{14} ions/cm^2, 150 keV) were measured and fit to an exponential function associated with the decay of excess point defects with annealing time:[70]

$$D(t) = D_i + D_o \exp(-t/\tau_o) \tag{4.17}$$

where D_i is the intrinsic B diffusivity, τ_o is the decay time of generated point defects, and D_o is an empirical prefactor for the transient diffusion coefficient.

Micron Ex. 1023, p. 209
Micron v. YMTC
IPR2024-00909

Decay-time data are plotted in Fig. 32 along with the dissolution rate curves for end-of-range dislocations and point-defect clusters.[71] Other diffusion transient data are also shown for low-dose B implants annealed at temperatures as low as 750°C and as high as 1100°C. These results show that short-time anneals below 1050°C start to be dominated by relatively rapid dissolution of point-defect clusters, with good agreement between diffusion-coefficient time constants and damage-annealing times. At higher temperatures, enhanced diffusion is assisted by the annealing of extended dislocations. The t_o data show how the transition occurs between the dominance in dopant diffusion of the two different sources of point defects as a function of temperature.

Other data for the enhanced diffusion time of low-dose B implants versus annealing temperature are shown in Fig. 33.[66] If we arbitrarily take the data from Ref. 72 as the most representative, then t_a is proportional to exp (3.7 eV/kT), and now the $D_p \cdot t_a$ product becomes

$$D_p(\text{enhanced}) \cdot t_a \propto \exp\,(1.5\ \text{eV}/kT) \tag{4.18}$$

The result in Eq. (4.18) can be used to explain the differences in the slopes of the RTA and furnace diffusion junction-displacement curves in Fig. 31. Higher anneal

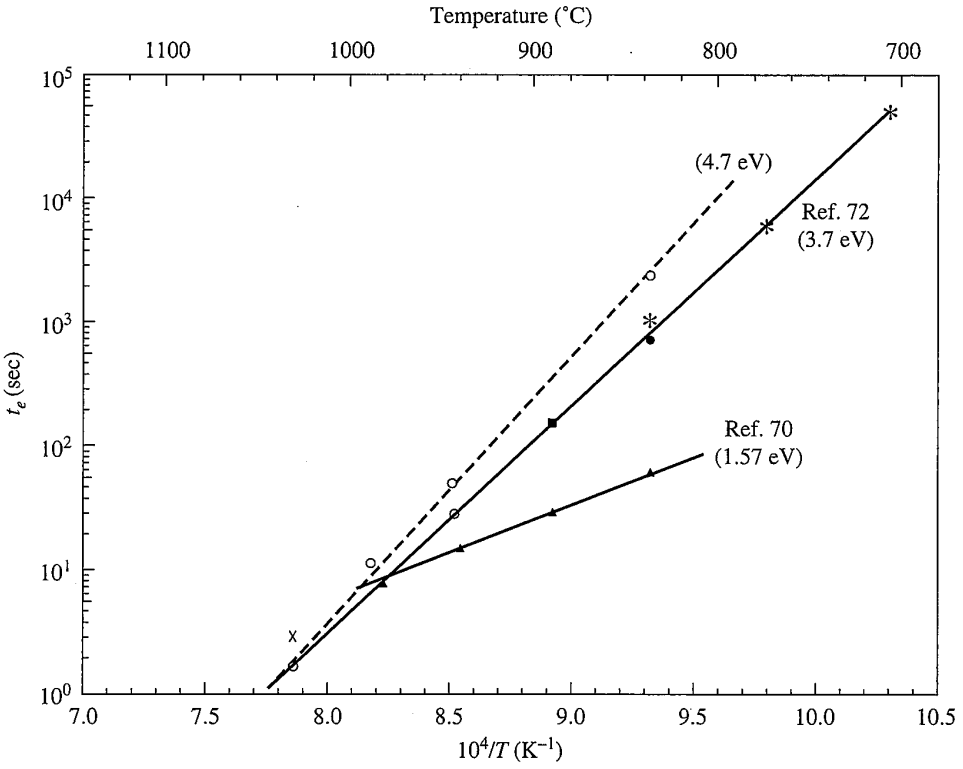


FIGURE 33
Transient-enhanced diffusion time, t_e, for low-dose boron implants versus the inverse annealing temperature showing the measured results of different researchers. (*After Miyake, Aoyama, and Kurchi, Ref. 70, and Solmi, Baruffaldi, and Canteri, Ref. 72.*) The data symbols are identified in Ref. 66; their disparity indicates uncertainties in the annealing temperatures and detection of the transient durations. (*After Fair, Ref. 66.*)

Micron Ex. 1023, p. 210
Micron v. YMTC
IPR2024-00909

temperatures result in higher dopant diffusion (Eq. 4.16), but they also resultin correspondingly smaller transient time constants, with the latter dominating the product of the terms in Eq. (4.18). The result is less diffusion for the damage annealing times associated with high-temperature RTA than for the low-temperature furnace anneals.

The formation of shallow emitters in bipolar transistors or source-drains in MOS transistors requires high-dose implants in the range of 5×10^{14} to 1×10^{16}atoms/cm^2, which are sufficient to create extended defects. Annealing conditions for meeting the requirements of shallow, highly doped junctions must, therefore, include removing extended defects from junction space-charge regions as well as controlling junction depths. The implantation damage produced for amorphizing implants and high-dose B implants is shown in Fig. 28*b* and 28*c*, respectively. Therefore, as the implant dose increases, more defects are created that can contribute to additional damage-assisted diffusion for longer times and, subsequently, diminished control of junction depths. The onset of the longer diffusion transient time with increasing B dose is illustrated in Fig. 34.[73] At 1×10^{14} ions/cm^2, the transient saturates in less than 10 seconds at 1000°C. As the dose reaches 5×10^{14} ions/cm^2, lattice damage increases and the transient becomes 20 seconds long. Finally, above a dose of 2×10^{15} ions/cm^2, the diffusion transient is over 40 seconds. In Fig. 34*d*, the B concentration in the peak of the implant has exceeded the solid solubility level of B in Si at 1000°C, and projected range dislocations (type V) have formed, as illustrated in Fig. 28*c*. High-temperature RTA is better suited to remove these defects than low-temperature furnace anneals, since at 950°C these types of defects can pump out Si self-interstitials for hours, causing continued enhanced B diffusion.[73, 74, 75]

If the implant dose is sufficient to transform the crystalline Si into the amorphous phase, then different point-defect generation processes occur when the wafer is annealed. During heating, the amorphous region will rapidly regrow to single-crystal Si at temperatures above 500°C by epitaxially aligning with the underlying crystalline substrate. The recrystallized Si is shown schematically in Region A of Fig. 35. However, damage produced near the original amorphous/crystalline Si interface evolves into end-of-range (EOR) dislocation loops by absorbing self-interstitials that are produced by the dissolution of interstitial clusters in Region B of Fig. 35. The clusters are formed during ion implantation. The annealing of the clusters produces an excess of self-interstitials in Region B, but not in Region A, since the growing EOR loops absorb the interstitials, thus screening them out of A. Once the clusters have dissolved (times of one sec at 1000°C), then the loops begin to dissolve into the Si bulk, producing a second self-interstitial transient that will remain as long as the loops are present (times of seconds to hours at 1000°C, depending on the distance from the Si surface). The resulting distribution of self-interstitials is shown in Fig. 35.

An example of implant-damage-assisted diffusion resulting from BF_2 implants that have formed amorphous layers in the first 500 Å of the surface is shown in Fig. 36.[76] RTA was performed first to show that diffusion saturated within 15 seconds at 900°C. However, when this same sample was subsequently placed in a furnace and annealed at 850°C for 1 hr, substantial additional diffusion occurred. This result highlights the compounding effect of damage-assisted diffusion from the annealing of point-defect clusters (short transient) and EOR dislocations produced near the amorphous/crystalline interface (long transient). If the EOR dislocations are more than a few hundred angstroms from the Si surface, then it becomes difficult to remove

Micron Ex. 1023, p. 211
Micron v. YMTC
IPR2024-00909

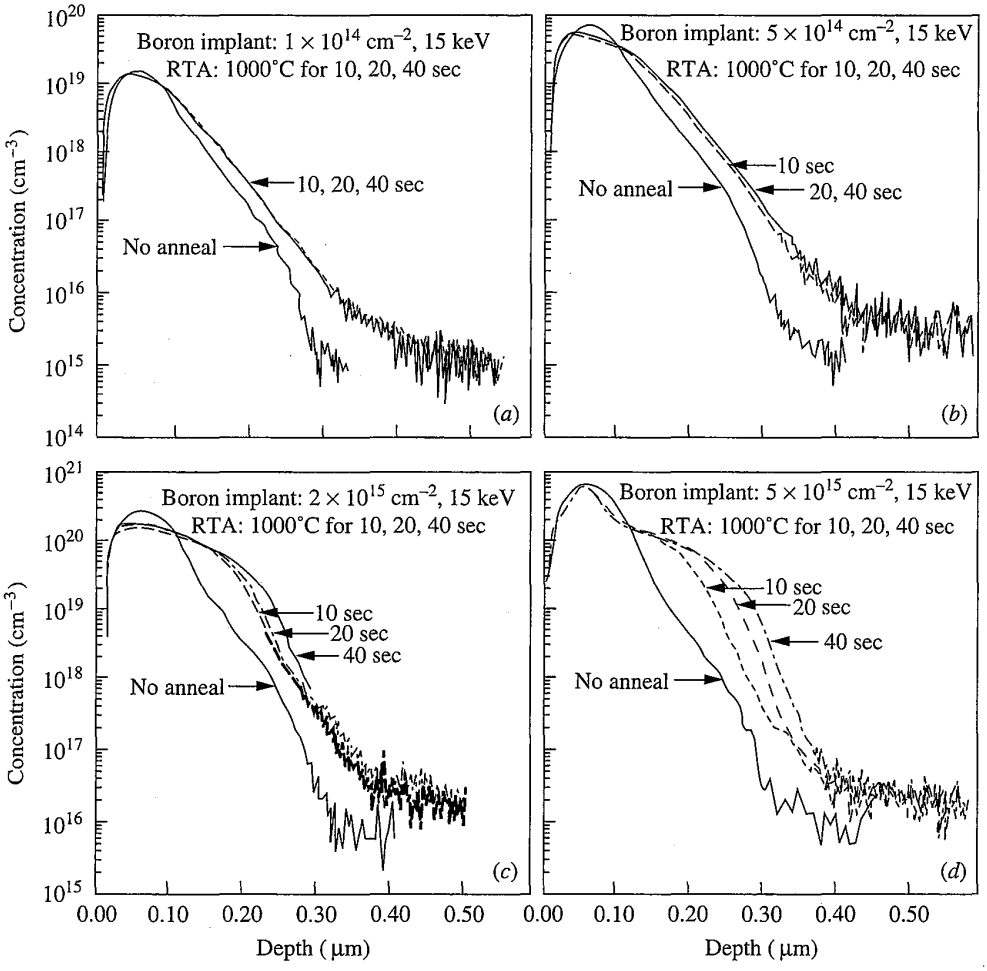


FIGURE 34
Boron secondary-ion mass spectroscopy (SIMS) profiles before and after RTA at 1000°C for 10, 20, and 40 sec at 15 keV. (*a*) 1×10^{14} cm^{-2}; (*b*) 5×10^{14} cm^{-2}; (*c*) 2×10^{15} cm^{-2}; (*d*) 5×10^{15} cm^{-2}. (*After Kim, Konoshita, and Kwong, Ref. 73.*)

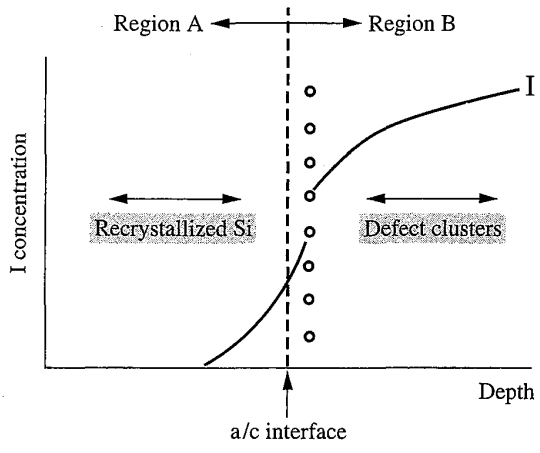


FIGURE 35
Schematic drawing of the annealing of an amorphizing implant with the resulting production of loop defects and self-interstitials. The amorphous layer in Region A rapidly regrows to single-crystal Si upon heating, and interstitial clusters in Region B dissolve, producing excess interstitials. Loops at the amorphous/crystalline interface grow in response, but screen the interstitials out of Region A by absorption. Finally, the loops begin to dissolve and create excess interstitials in Region B.

Micron Ex. 1023, p. 212
Micron v. YMTC
IPR2024-00909

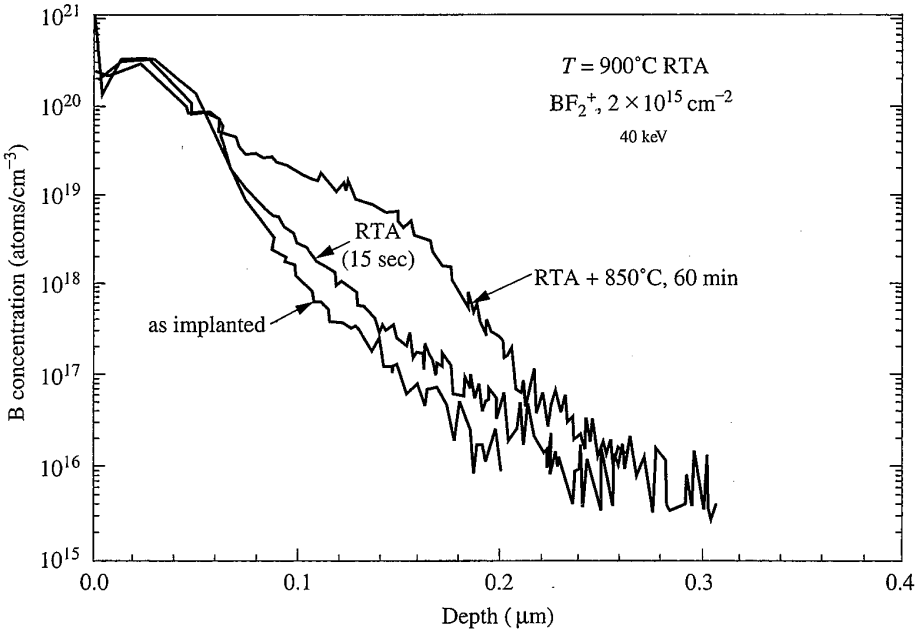


FIGURE 36
SIMS boron profiles for 2×10^{15} cm^{-2} BF_2 implants in (100) Si at 40 keV, annealed by RTA and RTA plus furnace annealing. (*After Kim, Massoud, and Fair, Ref. 76.*)

them within the time period allowed for junction anneal. For illustration, the first and second diffusion transients are identified in Fig. 37[77] for P with deep EOR damage. Annealing was performed at 900°C in a furnace. The first transient-enhanced diffusion (TED) is completed within 30 sec, whereas the second TED continues for hours while the EOR damage anneals. At higher temperatures and with shallower implants, the first and second TEDs can be made very short so that they become a minor part of the total junction diffusion time. This requires RTA and very low-energy implants.

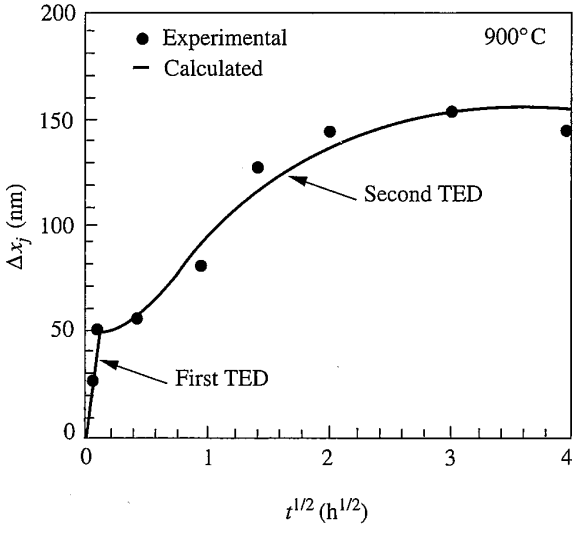


FIGURE 37
Junction depth displacement as a function of the square root of annealing time at 900°C in P-doped specimens, showing the time ranges for the first and second transient-enhanced diffusions (TED). The continuous line is calculated. (*After Solmi, Ref. 77.*)

Micron Ex. 1023, p. 213
Micron v. YMTC
IPR2024-00909

The conditions for creating defect-free, minimum depth p^+ junctions using RTA for 10 seconds are shown in Fig. 38.[19] Simulations were performed to determine the annealing temperatures necessary to remove implant damage at each energy and for each implant species (B or BF_2 in crystalline and Ge-preamorphized Si). For all combinations of implants, energies below 2 keV and RTA are required to form the 500- to 800-Å-deep junctions for ULSI. These results also show that, although preamorphization is important for reducing junction depths of B and BF_2 implants above 5 keV, little advantage exists at lower energies except that higher dopant activation is achieved for preamorphized implants.

The advantages of very shallow implants combined with RTA are

Lowered excess point-defect levels because of less implant damage and the proximity of the Si surface, which *sinks* point defects
Lowered point-defect transient time because of rapid damage annihilation
Junction depth control by thermal diffusion, not by damage annealing

Dopant activation

The electrical activation of implanted impurities by RTA or furnace heating is an important issue, especially for shallow junctions created by low-thermal-budget processing. Activation is related to the fraction of impurity atoms that are on substi-

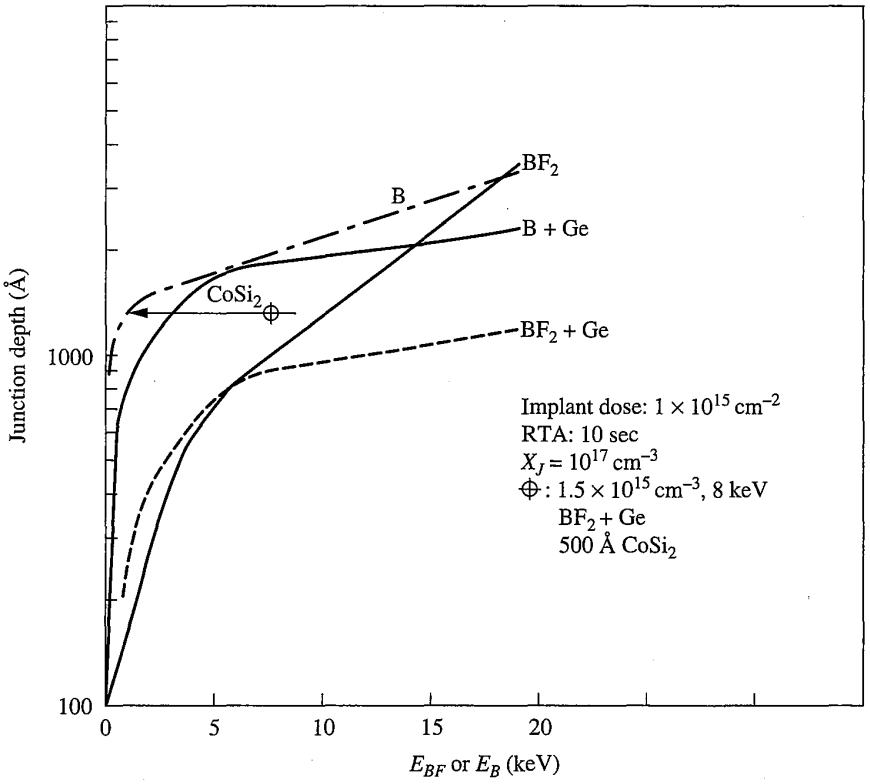


FIGURE 38
Minimum depth, damage-free p^+ junctions versus implant energy. The curves are simulated with the annealing temperatures necessary to remove damage from the junction region in 10 sec. (*After Fair, Ref. 19.* © 1990 IEEE)

Micron Ex. 1023, p. 214
Micron v. YMTC
IPR2024-00909

tutional lattice sites acting as donors or acceptors. Thus, free carrier concentrations as well as carrier mobility are indicators of activation as manifested in the measured sheet resistance of a junction layer.

A key question is, How can a high activation be achieved without loss of junction depth control? It is desirable to achieve 100% electrical activation up to the solubility limit of the dopant. High electrical activity is needed to ensure low contact resistances to junctions, since contact resistance is sensitive to active surface concentrations. In addition, high activation allows for the lowest possible sheet resistances of the implanted and diffused layers. The attainment of high carrier mobility is also important and requires that residual defects be minimized in the doped layer, i.e., the junction processing circus shown in Fig. 29.

The activation of B in high-dose implants into Si proceeds slowly at low annealing temperatures of 800 to 850°C. An example of activation at 800°C is shown in Fig. 39.[77] The total as-implanted B profile and the free-carrier profiles result from annealing in a furnace for 5 min and 2 hr. The carrier profiles have spread into the Si by diffusion. Although the hole concentration increases with time, it largely remains below the solid solubility level of 3.5×10^{19} atoms/cm^3 at 800°C. The inactive B above this level is either in the form of B precipitates or interstitial B.

By contrast, improved activation of B is possible by performing RTA at or above 900°C. An example is shown in Fig. 40.[77] After 5-sec and 15-sec RTA at 1000°C, a maximum electrical activity is achieved for a 2×10^{15} cm^{-2} implant that is 2 to 3 times greater than the equilibrium solubility. Similar results are seen at 900°C.[78] These results demonstrate that RTA, with its high-temperature and short-time capability, is a preferred technology in the trade-off between activation and junction-depth.

The basic mechanism of activation involves the transition of B atoms from interstitial positions in the Si lattice to substitutional positions. Since this process involves point defects, the activation energy is approximately 5 eV (diffusion energy for Si atoms), which makes it highly temperature-dependent. In addition, the damage

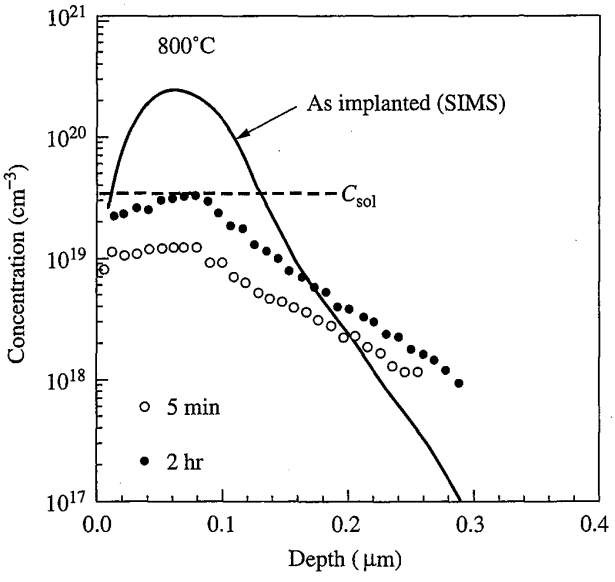


FIGURE 39
Carrier concentration profiles obtained in Si that was B-implanted at 20 keV with a dose of 2×10^{15} cm^{-2} and then annealed at 800°C twice. The as-implanted profile is total B obtained by SIMS. Hole concentration increases with time but remains below the solubility limit. (*After Solmi, Ref. 77.*)

Micron Ex. 1023, p. 215
Micron v. YMTC
IPR2024-00909

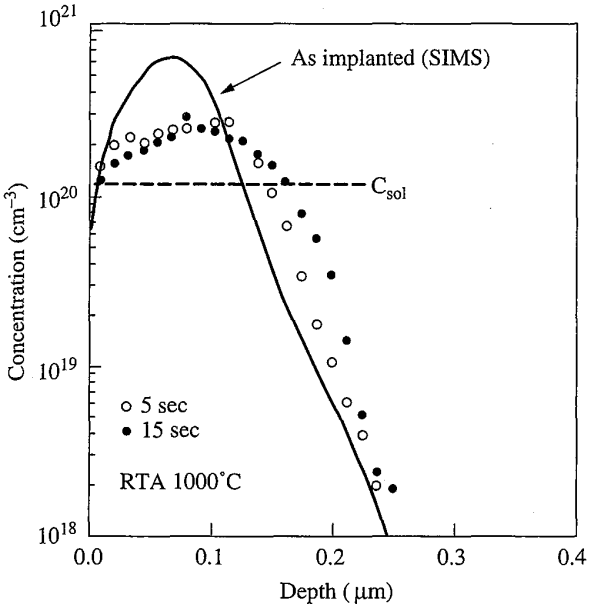


FIGURE 40
Carrier concentration profiles obtained in samples implanted at 30 keV with a dose of 1×10^{16} cm^{-2} after RTA for 5 and 15 sec. The as-implanted B profile was measured by SIMS. Hole concentrations over twice the equilibrium solubility limit are obtained. (*After Solmi, Ref. 77.*)

produced by the implantation is known to retard B activation,[78] and damage removal also proceeds with a 5-eV energy (see Fig. 32). Consequently, high temperature favors damage removal and activation. On the other hand, B precipitation is a 2.1-eV process[77] and will prevail over activation at lower temperatures, whereas the opposite is true at high temperatures.

The competition between activation and precipitation is diagrammed in Fig. 41, where peak free-carrier concentrations are plotted against annealing time.[77] The dashed curve shows the decrease in free carriers resulting from precipitation, and the solid curve shows the increase in carriers resulting from activation. Figure 41*a* shows that, at high temperature, activation can occur quickly up to the limit allowed by precipitation, which exceeds the solubility limit for short times. Figure 41*b* shows that, at low temperature, by the time the activated carrier concentration reaches its highest level, all the B above solid solubility is precipitated.

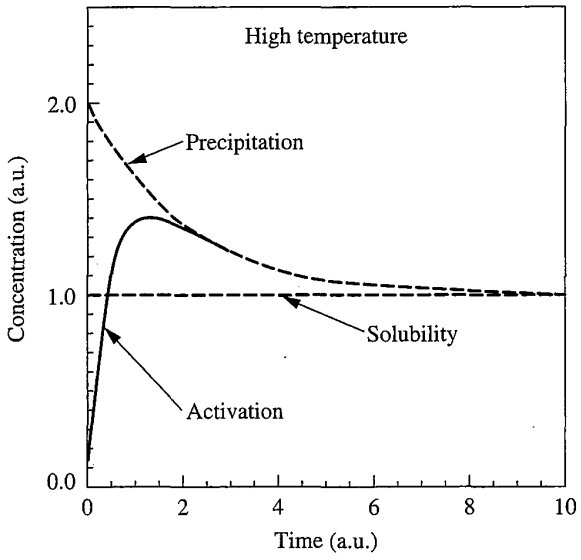
A special case exists for low-energy B or BF_2 implanted into preamorphized Si. The fraction of activated B as a function of annealing temperature depends on the depth of the amorphous layer relative to the implant depth. Two effects have been observed if the amorphous layer is shallower than the B implant:[79]

Reduced B activation occurs if EOR loops exist in the implanted region. A decrease in electrically active B occurs as the dislocation loops coalesce from defect clusters at moderate annealing temperatures (600 to 800°C).

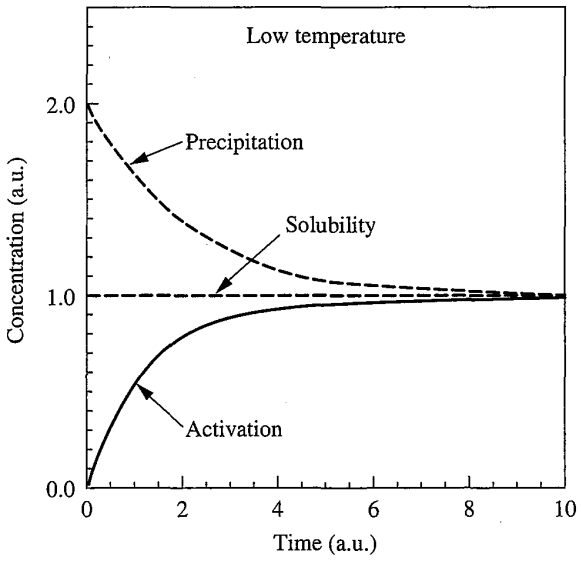
Decreasing the preamorphization depth decreases the fraction of activated dopant, because fewer B atoms reside within the regrown layer.

Over the annealing range of 500 to 900°C, preamorphization produces higher levels of activation for B implants than implants into crystalline Si.[80] This is caused by the incorporation of B atoms onto substitutional sites during the process of solid phase epitaxial regrowth of the amorphous layer, and this process begins at temperatures as low as 550°C.

Micron Ex. 1023, p. 216
Micron v. YMTC
IPR2024-00909



(*a*)



(*b*)

FIGURE 41
Diagrams of the peak free-carrier concentrations allowed by precipitation of B and by activation are plotted versus time in arbitrary units (a.u.) for (*a*) high-temperature annealing and (*b*) low-temperature annealing in samples implanted with concentrations above solid solubility. (*After Solmi, Ref. 77.*)

Activation limits of P-doped Si have been defined in terms of a solubility of electrically active P. Studies have shown that the electrically inactive P is in the form of phosphide precipitates.[81] However, it has also been shown that 100% activation of P implants is possible using laser annealing.[82] Free-carrier concentrations of up to 5×10^{21} cm^{-3} were obtained in laser-annealed layers free of structural imperfections. All of the P was on substitutional lattice sites, since the high wafer temperature produced by the laser caused rapid activation. Subsequent thermal anneals reduced the carrier densities to equilibrium values dependent only on temperature.

Micron Ex. 1023, p. 217
Micron v. YMTC
IPR2024-00909

For the annealing time associated with high-temperature RTA, the equilibrium, active P concentrations are achieved[83] rapidly for diffusions from spin-on phosphosilicate glasses. Activation of P-implanted layers at lower temperatures is complicated by the presence of lattice defects. In general, the following observations have been made:

Implanted P in very-low-damage regions can be activated by annealing at temperatures as low as 400°C.

For damage densities of about 20%, carrier trapping centers may be formed to compensate the activation. Temperatures above 500°C are needed to achieve electrical activity.

For damage densities up to 75%, amorphous clusters are formed by ion-implantation overlap, producing larger damage regions. Decomposition of these regions during annealing may produce vacancies that enhance activation of P atoms onto substitutional sites.

For fully amorphized regions (100% damage density), movement of the amorphous/crystalline (a/c) interface during epitaxial regrowth incorporates P atoms into lattice sites. However, the activation of P atoms deeper than the a/c interface depends on the annealing temperature.

The activation of As is similar to that of P. However, controversy still remains regarding the form of electrically inactive As. It has been proposed that complex point defects exist in thermal equilibrium that result in electrically inactive As species. The effect of low-thermal-budget anneals and furnace ramps on the electrical activation of As was recently studied.[84] The equilibrium between active and inactive As can generally be described by a rate equation such as

$$\frac{\partial C_{\text{active}}}{\partial t} = K_D F(C_{\text{inactive}}, C_{\text{chemical}}) - K_c R(C_{\text{active}}, C_{\text{chemical}}) \tag{4.19}$$

where K_D is a declustering or deprecipitation coefficient, K_c is a clustering or precipitation coefficient, and F and R are functions of the concentrations indicated and depend on the model chosen. Computer simulation of activation or deactivation requires that Eq. (4.19) be solved simultaneously with the diffusion equation for As,[84] given as follows:

$$\frac{\partial C_{\text{chemical}}}{\partial t} = \frac{\partial}{\partial x} D \frac{\partial C_{\text{active}}}{\partial x} + Z\mu C_{\text{active}} \frac{\partial \Psi}{\partial x} \tag{4.20}$$

where D and μ are the mobile As diffusivity and mobility, respectively, Z is the charge state, and Ψ is the crystal potential. The second term on the right-hand side of Eq. (4.20) is the electric field dependence.

Depending on the initial conditions chosen, the solution to Eq. (4.20) will yield different solutions for active As concentration versus anneal time. The activation model of Tsai et al.[85] shows that, for high As doping, less than 5 minutes are required to achieve equilibrium at 1000°C. However, at 900°C, the time approaches 60 minutes. Thus, high-temperature RTA is favored for As as well as for B and P. During most common RTA cycles, the activated As concentration for high-dose

Micron Ex. 1023, p. 218
Micron v. YMTC
IPR2024-00909

As implants is in transition from the initial value after epitaxial regrowth to a final equilibrium value. This also means that the activation depends on the ramp-up and ramp-down rates of the RTA cycle.

4.3.5 Metallization

The application of RTCVD as an alternative to LPCVD or PECVD of metal films has received very limited attention. On the other hand, rapid thermal processing has been widely applied to forming metal silicides for contacts to junctions and gate structures. RTA has been found to be superior to conventional furnace techniques for silicides because of the limited thermal stability of many silicide materials and because of their extreme sensitivity to trace impurities in the annealing ambient.[86] Thus, RTA is used more for control in heating and ambient than for deposition. Current applications include Ti, Co, W, Ta, and Pt silicides and TiN formation and annealing for local interconnections on the chip.

RTA of silicides

The early development of self-aligned silicide technology was performed in furnace systems that required a significant amount of attention to avoiding oxygen contamination. The mechanics of oxygen control basically rendered the approach impractical. Two features of RTP enabled silicide technology to reach production:

Atmosphere control
Short, high-temperature annealing, which allows the nucleation and growth of the desired high-conductivity silicide phase

In addition, the reaction of thin metal films with Si to form silicides proceeds rapidly at moderate temperatures and thus provides a self-limiting, uniform, thin-film process that is relatively independent of temperature variations inherent in RTP technology. Indeed, silicide processes are tolerant of thermal gradients and temperature uncertainties that occur in the wafers. The temperature dependencies of the resulting sheet resistances of silicide films are independent of the annealing system.[86]

One-and two-step processes used to form self-aligned silicides in RTP systems are shown in Fig. 42.[86] The metal reacts only where it comes in contact with Si, so after the first annealing step the unreacted metal is etched away selectively. One early difficulty with using silicides in patterned MOS devices was the lateral migration of Si from the substrate through the formed silicide to the unreacted metal, where it then formed additional silicide outside the patterned window. When the selective metal etch was applied, there would remain a silicide layer that stretched beyond the contact window in the form of an electrical short to the surrounding features of the device, i.e., the gate. This lateral overgrowth is minimized by the anneal-etch-anneal process shown in Fig. 42*b*. RTP enables the use of the first low-temperature, short-time step to form only a modest amount of Si/metal interdiffusion. The etch step then removes the excess metal outside the patterned contact windows, and a subsequent high-temperature step forms the high-conductivity silicide phase. This approach is unique to RTA, since it requires short annealing times.

Micron Ex. 1023, p. 219
Micron v. YMTC
IPR2024-00909

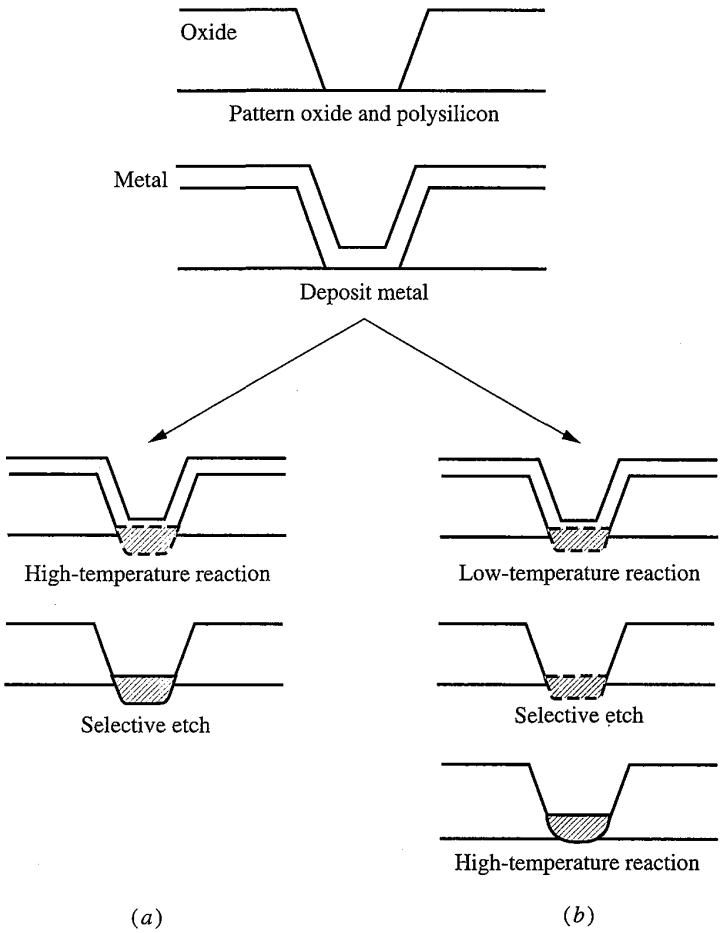


FIGURE 42
Thermal reaction of a metal thin film with Si to form a self-aligned silicide: (*a*) single-step anneal and (*b*) two-step anneal. (*After Osburn, Ref. 86.*)

The process windows giving the acceptable range of initial RTA temperatures and times needed to avoid lateral bridging in MOSFETs are shown in Fig. 43*a* for $TiSi_2$[87] and Fig. 43*b* for $CoSi_2$.[88] Those processing conditions within the acceptable window avoid lateral bridging of the junction and the gate in MOSFETs that have oxide spacers separating the poly-Si gate and the junction. Figure 43*b* shows the data on spacer-bridging yield for $CoSi_2$ formed from 38 nm of cobalt as a function of the initial 30-sec RTA temperature. The graphs demonstrate that improved yield is obtained for lower initial RTA temperatures, with 600°C being an upper limit.

Silicide growth kinetics—RTA versus furnace

The reactions of metals with Si are either diffusion-controlled or limited by nucleation of the phase that is undergoing growth. Diffusion-controlled kinetics are observed in most silicides, and the growth rate increases with the square root of time.[86] For example, the silicide growth-rate parameter of $TiSi_2$ measured in terms of the

Micron Ex. 1023, p. 220
Micron v. YMTC
IPR2024-00909

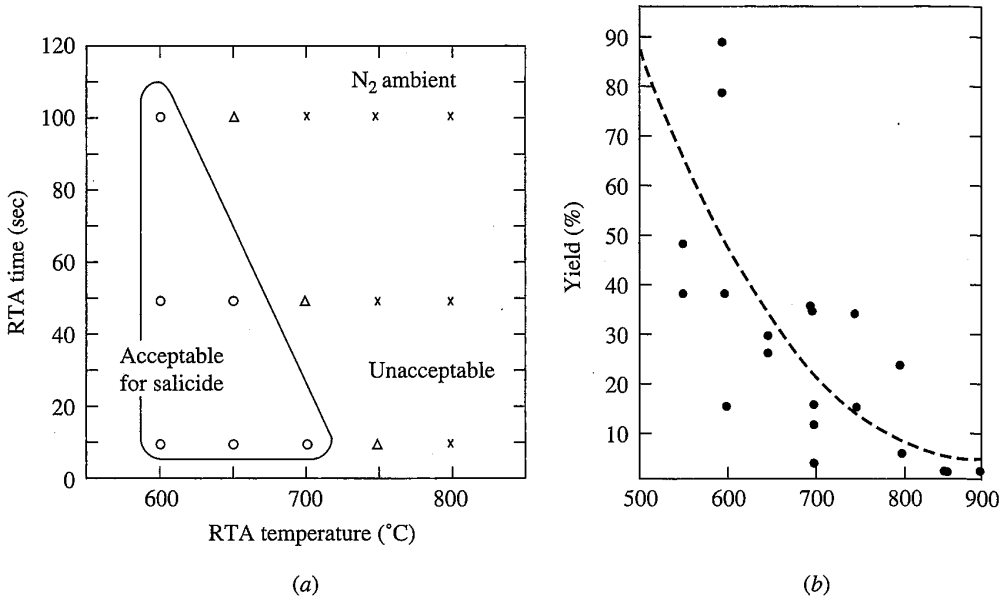


FIGURE 43
Process windows for (*a*) $TiSi_2$[87] and (*b*) $CoSi_2$[88] showing the acceptable range of RTA temperatures and times required to avoid lateral bridging (contact-to-gate shorts). The devices tested were MOSFETs with oxide spacers separating the polysilicon gate from the contact windows.

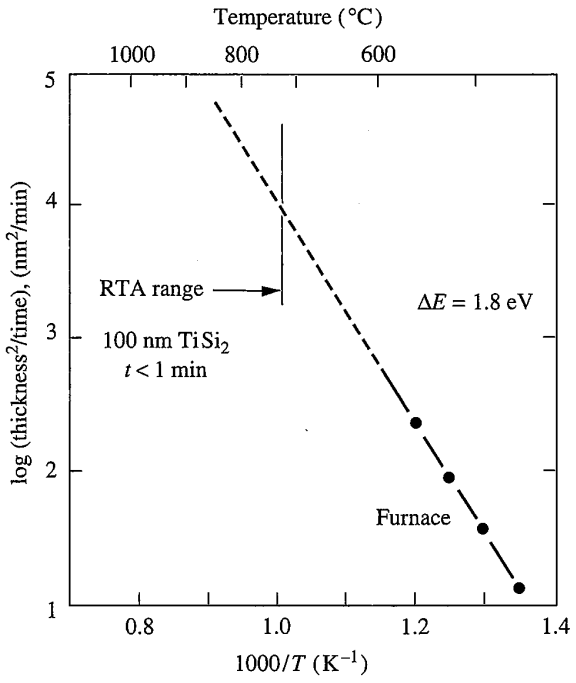


FIGURE 44
Kinetics of silicide formation in the furnace and RTA regimes. (*After d'Heurle, Ref. 89.*)

Micron Ex. 1023, p. 221
Micron v. YMTC
IPR2024-00909

log of thickness^2/time is plotted in Fig. 44 over a wide range of times, from seconds to hours, and a wide range of temperatures, including the low-temperature furnace range and the higher-temperature RTA range.[89] Since a single activation energy of 1.8 eV characterizes the reaction, the use of RTA at moderate temperatures (500 to 800°C) can fully react the metal film in a few seconds. Extended annealing times are undesirable after silicide formation, since agglomeration of the silicide may occur, destroying film uniformity.

Factors that can influence silicide kinetics are the presence of a thin native-oxide film on the Si surface, dopant impurities, and oxidation of the surface of the deposited metal film by oxygen or by moisture that competes with the silicide reaction at the metal/Si interface. The sensitivity of the sheet resistance of titanium silicide to trace amounts of oxygen is shown in Table 2.[90] Keeping oxygen contamination below these levels is difficult in large-diameter furnace tubes with air backstreaming from the outlet end. However, the microenvironment achievable within an RTA system along with the low impurity levels in high-purity ULSI gases, N_2 or Ar, are suitable for silicides.

Finally, the application of RTA to silicide film growth is not without some issues. Among the most serious are the following:[86]

- Nonuniform silicide/Si interface, which limits the depth of the silicide relative to the pn junction depth
- Silicide agglomeration and instability
- Forces at the silicide/Si interface that can lead to dislocation generation at the higher RTA temperatures
- Depletion of dopants on the Si side of the interface via grain-boundary diffusion through the silicide, which leads to irregular junctions

The effects of surface and junction roughening and dopant depletion are illustrated in Fig. 45.[91] During silicide growth, silicon is consumed in the reaction with metal. Dopants in the Si can diffuse out through the silicide by grain-boundary transport toward the surface where they either accumulate or evaporate. Dopants may be introduced by ion implanting through the layer once the silicide is formed. However, the resulting damage produced by the process can increase the sheet resistance of the silicide. Last, the annealing step causes top and bottom surface roughening as a result of silicide agglomeration. The angles that occur between the grain boundaries present in the layer change in response to surface-tension forces, producing thermal grooving. Continued heating can cause increased roughness and the formation of silicide islands, and this result drives annealing requirements in the direction of controlled, low-thermal-budget processing, achievable with RTA.

TABLE 2
Effect of oxygen impurities on $TiSi_2$ sheet resistance[90]

Gas	Impurities (ppm of oxygen)	Silicide sheet resistance (Ω / □)
Nitrogen	0.1	2.6
Nitrogen	1.0	3.4
Nitrogen	10.0	100.0
Argon	<0.1	1.8

Micron Ex. 1023, p. 222
Micron v. YMTC
IPR2024-00909

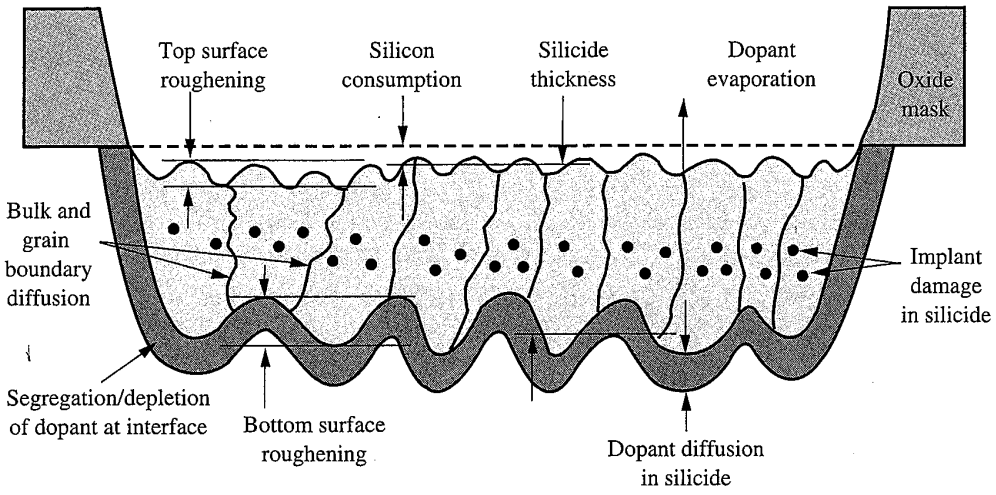


FIGURE 45
Key phenomena in silicide processes. (*After Osburn et al., Ref. 91.*)

4.4 SUMMARY AND FUTURE TRENDS

Rapid thermal processing (RTP) has the advantages over conventional furnace technology of operating in a controlled microenvironment with low thermal mass and rapid heating rates. Because high temperatures are attainable in a short time, single-wafer thermal processing can occur in multiple-use chambers or in clusters of single-use chambers, thermal budgets can be achieved for defect-free junctions with high dopant activation, and sharp interfaces can be created in layered structures.

A road map showing the introduction of RTP into manufacturing DRAMs is depicted in Table 3. Each of the key processing areas that will be affected by RTP is listed along with estimated time for the transition from batch to single-wafer technology. The road map shows that the potential of RTP increases the further the technology moves into the sub-half-micrometer design rule regime. For example, the "processing" row shows that batch processing will give way to single-wafer processing to meet the requirements of 0.5-μm manufacturing technologies. Once this choice is made, there are no alternatives left except to implement RTP for wafer throughput. As more and more processes are converted to single-wafer tools, it eventually becomes economically feasible and necessary to begin integrating the tools into clusters, as is expected by the end of this decade for 0.15-μm complementary MOS.

Within the next few years, all batch processes will have been demonstrated in RTP. However, the introduction of these RTP processes into manufacturing will occur only when the process is absolutely needed because of technical and cost requirements.[4] For example, existing CVD processes in conventional, batch CVD furnaces do not always work well in RTP systems. New chemistries are needed, especially for thick-film deposition, to obtain the deposition rates that are required for wafer throughput. Otherwise, the processes will be too costly. However, there may

Micron Ex. 1023, p. 223
Micron v. YMTC
IPR2024-00909

TABLE 3
RTP technology road map

	1985		1990		1995		2000	
Design rule (μm)	2	1.3	0.8	0.5	0.3	0.2	0.15	
DRAM equivalent	256k	1M	4M	16M	64M	256M	1G	
Wafer size (mm)	125	125	150–200	200	200–250		300	
t_{ox}(nm)	35	25	20	15	12	10	8	
L_{eff}(μm)	2	1.3	0.8	0.5	0.3	0.2	0.15	
x_j(μm)	0.3	0.25	0.2	0.15	0.10	0.08	0.06	
Repeatability (°C)	±15		±5		±2		±1	
Uniformity (°C)	±7–10		±3–5		±2–3		< ±2	
Accuracy (°C)	±20		±5–10		±3–5		< ±3	
Thermal budget			Furnace				RTA	
t_{ox} control			Furnace				RTO	
Thin dielectrics			LPCVD				RTCVD-ONO	
Silicides			Furnace		RTA			
Epitaxy			LPCVD, MBE,† APCVD, etc.		RTCVD			
Polysilicon			LPCVD		RTCVD-SiGe			
Metals			LPCVD sputtered, evaporated		RTCVD			
Cleaning			Wet chemistry		RTC			
Processing	Batch				Single wafer		Clusters	
Atmospheric budget	0.1/cm²	> 0.5 μm	0.01/cm²	> 0.3 μm	0.005/cm²	> 0.2 μm	0.001/cm²	> 0.1 μm

†Molecular beam epitaxy

Micron Ex. 1023, p. 224
Micron v. YMTC
IPR2024-00909

be good justification for introducing thin-film CVD processing in the near future. As single-wafer technology evolves, CVD processes may first become integrated into single-process, multichamber tools with some level of in-situ etch cleaning. For fabricating MOSFETs in a single chamber, it is desirable to have a high-quality CVD process for depositing the gate dielectric rather than using RTO with its attendant problems of temperature control and uniformity. Thus, the use of RTCVD-ONO (oxide-nitride-oxide) deposition precedes the advent of RTO in Table 3.

Much of the time-to-introduction of the various technologies listed in Table 3 depends on the ability of equipment manufacturers to design RTP tools with suitable process chambers, temperature measurement and control systems, sensors, and reproducible processes. Technology advancement will probably focus on RTP, single-wafer chambers with multizone heating reflectors, and on real-time, independent control of the zones in the multilamp heating systems. End point detection techniques for process control should be introduced, rather than relying on temperature measurement. End point detection involves a measurement of some property associated with a process, such as a film thickness or the chemical composition of the chamber ambient. Typical end point detection methods include ellipsometry, reflectometry, interferometry, and possibly emissivity measurements.[4] Although end point detection will not solve the problems of temperature control in RTP, it may improve the repeatability of a process.

Finally, it is evident that RTP technologies are evolving without any major breakthroughs, but rather with constant and diligent attention to engineering details. Indeed, radiant heating had a deceptively simple veneer that covered all the underestimated problems. And no producer of RTP tools has solved the fundamental issues of depositing radiant energy into the materials common to semiconductor devices.[92] As a result, progress is made by fixing and patching. The preferred methodology would be one in which the basic issues and science were well understood so that problems could be avoided.

REFERENCES

1. S. M. Sze, Ed., *VLSI Technology,* 2nd ed., McGraw Hill, New York, 1988.
2. R. B. Fair, "Rapid Thermal Processing—A Justification," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993, p. 1.
3. J. K. Roberts, *Heat and Thermodynamics,* 4th ed., Blackie and Son, Ltd., Glasgow, 1955.
4. F. Roozeboom, "Manufacturing Equipment Issues in Rapid Thermal Processing," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993, p. 349.
5. T. Sato, "Spectral Emissivity of Silicon," *Jpn. J. Appl. Phys.* **6,** 339 (1967).
6. F. Roozeboom, "Temperature Control and System Design Aspects in RTP," in *Mater. Res. Soc. Symp. Proc.,* **224,** 9 (1991).
7. C. Hill, S. Jones, and D. Boys, "Rapid Thermal Annealing—Theory and Practice," in R. A. Levy, Ed., *Reduced Thermal Processing for ULSI,* Plenum Press, New York, 1989, p. 143.

Micron Ex. 1023, p. 225
Micron v. YMTC
IPR2024-00909

8. J. C. Liao and T. I. Kamins, "Power Absorption during Deposition of Polycrystalline Silicon in a Lamp-Heated CVD Reactor," *J. Appl. Phys.*, **67,** 3848 (1990).
9. T. P. Merchant, K.-H. Lee, J. V. Cole, and K. F. Jensen, "Strategies for Modeling of Rapid Thermal Processing Systems," in R. B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ* 1993, p. 376.
10. Z. Nenyei and A. Tillmann, "Reaction Time Analysis in Rapid Thermal Annealing," in R. B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 429.
11. I. Beinglass and S. S. Schwartz, "Single Wafer Integrated Processing System for High Temperature Thin Films Formation," in R.B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 416.
12. C. Lee and G. Chizinsky, "Rapid Thermal Processing Using a Continuous Heat Source," *Solid State Tech.,* **32 (1),** (1989), p. 43.
13. A. Kamgar and S. Hillenius, "Rapid Thermal Anneal Induced Effects in Polycrystalline Silicon Gate Structures," *Appl. Phys. Lett.,* **51,** 1251 (1987).
14. J.S. Nakos, "Application of Rapid Thermal Processing in Manufacturing: The Effect of Emissivity and Coupling," in R. B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 421.
15. R. B. Fair, "Challenges in Manufacturing Submicron, Ultra-Large Scale Integrated Circuits," *Proc. IEEE,* **78,** 1990, p. 1687.
16. R. S. Nelson, "The Physical State of Ion Implanted Solids," in G. Dearnaley, J. H. Freeman, R. S. Nelson, and J. Stephen, Eds., *Ion Implantation,* North Holland Pub. Co., Amsterdam, 1973, p. 154.
17. A. E. Michel, "Diffusion Modeling of the Redistribution of Ion Implanted Impurities," in T. O. Sedgwick, T. E. Seidel, and B. Y. Tsaur, Eds., *Rapid Thermal Processing,* Mat. Res. Soc., Pittsburgh, **52,** 3 (1986).
18. *PREDICT®—PRocess Estimator for the Design of IC Technologies,* MCNC, Research Triangle Park, NC, 1987.
19. R. B. Fair, "Damage Removal/Dopant Diffusion Trade-offs in Ultra-Shallow Implanted p^+n Junctions," *IEEE Trans. Electron Dev.,* **37,** 2237 (1990).
20. E. A. Irene and R. Ghez, "Silicon Oxidation Studies: The Role of H_2O," *J. Electrochem. Soc.,* **124,** 1757 (1977).
21. N. Yabumoto, K. Saito, M. Morita, and T. Ohmi, "Oxidation Process of Hydrogen-Terminated Silicon Substrate Studied by Thermal Desorption Spectroscopy," *Jpn. J. Appl. Phys.,* **30,** 1419 (1991).
22. C. M. Osburn, H. Berger, R. Donovan, and G. Jones, "The Effect of Contamination on Semiconductor Manufacturing Yield," *Proc. of Institute of Environmental Science,* **31,** 45 (1988).
23. B. T. Murphy, "Cost-Size Optima of Monolithic Integrated Circuits," *Proc. IEEE,* **52,** 1537 (1963).
24. P. J. Rosser, P. B. Moynagh, and K. B. Affolter, "Multi-Chamber Rapid Thermal Processing," *Soc. Photo-Opt. Instrum. Eng. Symp. Proc.* **1393,** 49 (1990).
25. K. Shenai, P. A. Piacente, and B. J. Baliga, "Reliable Low Contact Resistance Al-TiW-$TiSi_2$-n^+ Si Contacts for Silicon VLSI and Smart Power Applications," in M. Scott, Y. Arkasaka, and R. Rief, Eds., *Advanced Materials for VLSI,* Electrochem. Soc., Pennington, NJ, 155 (1988).
26. Y. El-Mansey, "MOS Device and Technology Constraints in VLSI," *IEEE Trans. Electron. Dev.,* **ED-29,** 567 (1982).
27. E. Takeda, G. A. C. Jones, and H. Ahmed, "Constraints of the Application of 0.5 μm MOSFETs to VLSI Systems," *IEEE Trans. Electron. Dev.,* **ED-32,** 322 (1985).

Micron Ex. 1023, p. 226
Micron v. YMTC
IPR2024-00909

28. E. Sun, J. Moll, J. Berger and B. Alders, "Breakdown Mechanism in Short-Channel MOS Transistors," in *Tech. Dig. IEDM,* 478 (1983).
29. W. H. Lee, T. Osakama, K. Asada, and T. Sugano, "Design Methodology and Size Limitations of Submicrometer MOSFETs for DRAM Application," *IEEE Trans. Electron. Dev.,* **35,** 1876 (1988).
30. S. Asai, "The Material Matters in Future ULSIs," in *Symposium on Adv. Sci. and Tech. of Si Materials,* Japan Society for Promotion of Science, Kona, 1 (1991), (unpublished).
31. K. M. Cham and S. Y. Chiang, "Device Design for the Sub-Micrometer p-Channel FET with n^+ Polysilicon Gate," *IEEE Trans. Electron. Dev.,* **ED-31,** 964 (1984).
32. J. F. Gibbons, C. M. Gronet, and K. E. Williams, "Limited Reaction Processing: Silicon Epitaxy," *Appl. Phys. Lett.* **47,** 721 (1986).
33. J. L. Hoyt, "Rapid Thermal Processing-Based Epitaxy," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993.
34. J. O. Borland, "Historical Review of SEG and Future Trends in Silicon Epi Technology," in G.W. Cullen, Ed., *Proceedings of the 10th International Conference on Chemical Vapor Deposition, Pennington, NJ,* Electrochemical Society, 87 (1987).
35. M. Ogirima and R. Takahashi, "Some Problems and Future Trends in Silicon Epitaxial Technology," in G.W. Cullen, Ed., *Proceedings of the 10th International Conference on Chemical Vapor Deposition, Pennington, NJ,* Electrochemical Society, 87 (1987).
36. J. L. Regolini, D. Bensahel, J. Mercier, and E. Scheid, "Selective Epitaxial Silicon Growth in the 650–1100°C Range in a Reduced Pressure Chemical Vapor Deposition Reactor Using Dichlorosilane," *Appl. Phys. Lett.* **54,** 658 (1989).
37. F. C. Eversteyn, "Chemical-Reaction Engineering in the Semiconductor Industry," *Philips Res. Rept.* **29,** 45 (1974).
38. M. C. Ozturk and J. J. Wortman, "Rapid Thermal Chemical Vapor Deposition for Silicon Based Microelectronics Manufacturing," in R.B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 297.
39. T. O. Sedgwick, P. D. Agnello, D. Nguyen Ngoc, T. S. Kuan, and G. Scilla, "High Phosphorus Doping of Epitaxial Silicon at Low Temperature and Atmospheric Pressure," *Appl. Phys. Lett.* **58,** 1896 (1991).
40. J.-E. Sundgren, J. Knall, W.-X. Ni, M.-A. Hasan, L. C. Markert, and J. E. Green, "Dopant Incorporation Kinetics and Abrupt Profiles during Silicon Molecular Beam Epitaxy," *Thin Solid Films,* **183,** 281 (1989).
41. M. C. Ozturk, "Thin-Film Deposition," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993, p. 79.
42. L. Y. Cheng, J. C. Rey, J. P. McVitte, and K. Saraswat, *IEEE VMIC Conference, June 12–14, Extended Abstracts,* 1990, p. 404.
43. F. S. Becker, D. Pawlik, H. Anzinger, and A. Spitzer, "Low Pressure Deposition of High Quality SiO_2 films by Pyrolysis of Tetraethylorthosilicate," *J. Vac. Sci. Technol.* **B5,** 1555 (1987).
44. R. Miller, M. C. Ozturk, J. J. Wortman, F. S. Johnson, and D. T. Grider, "LPCVD of Silicon Dioxide by Pyrolysis of TEOS in a Rapid Thermal Processor," *Mat. Lett.* **8,** 353 (1989).
45. A. C. Adams, "Dielectric and Polysilicon Film Deposition," in S. M. Sze, Ed., *VLSI Technology,* 2nd ed., McGraw-Hill, New York, 1988, p. 233.
46. X. Xu, R. Kuehn, J. J. Wortman, and M. C. Ozturk, unpublished (1992).
47. W. Ting, S. N. Lin, and D. L. Kwong, "Thin Stacked Oxide/Nitride Dielectrics Formation by in situ Multiple Reactive Rapid Thermal Processing," *Appl. Phys. Lett.* **55,** 2312 (1989).
48. F. S. Johnson, R. M. Miller, M. C. Ozturk, and J. J. Wortman, "Characterization of LPCVD of Silicon Nitride in a Rapid Thermal Processor," in *Rapid Thermal Annealing/*

Micron Ex. 1023, p. 227
Micron v. YMTC
IPR2024-00909

Chemical Vapor Deposition and Integrated Processing Symposium, MRS Symposia Proceedings, **146** 1989, p. 345.

49. R. F. C. Farrow, "The Kinetics of Silicon Deposition by Pyrolysis of Silane," *J. Electrochem. Soc.,* **121,** 899 (1974).

49a. M. C. Ozturk, J. J. Wortman, Y. Zhong, X. Ren, R. M. Miller, F. S. Johnson, D. T. Grider, and D. A. Abercrombie, in *Rapid Thermal Annealing/Chemical Vapor Deposition and Integrated Processing Symposium,* MRS Symposia Proceedings, **146** 1989, p. 109.

50. K. L. Chiang, C. J. Dell'Oca and F. N. Schwettmann, "Optical Evaluation of Polycrystalline Silicon Surface Roughness," *J. Electrochem. Soc.,* **126,** 2267 (1979).

51. X. Ren, M. C. Ozturk, J. J. Wortman, B. Zhang, and D. Maher, "Deposition and Characterization of Polysilicon Films Deposited by Rapid Thermal Processing," *J. Vac. Sci. Technol. B,* **10,** 1081 (1992).

51a. M. C. Ozturk, F. Y. Sorrell, J. J. Wortman, F. S. Johnson, and D. T. Grider, "Manufacturability Issues in Rapid Thermal Chemical Vapor Deposition," *IEEE Trans. Semicond. Manuf.* **4,** 155 (1991).

52. H. Z. Massoud, "Rapid Thermal Growth and Processing of Dielectrics," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993.

53. A. Kazor and I. W. Boyd, "UV-assisted Growth of 100 Å Thick SiO_2 at 550°C," *Electron. Lett.,* **27,** 909 (1991).

54. V. Craciun and I. W. Boyd, "Low Temperature UV Induced Oxidation of Silicon and Silicon-Germanium Strained Layers," in R. B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 363.

55. R. Deaton and H. Z. Massoud, "Effect of Thermally-Induced Stresses on the Rapid Thermal Oxidation of Silicon," *J. Appl. Phys.* **70,** 3588 (1991).

56. J. Nulman, J. P. Krusius, and P. Renteln, "Material and Electrical Properties of Gate Dielectrics Grown by Rapid Thermal Processing," *Mat. Res. Soc., Symp. Proc.,* **52,** 341 (1985).

57. C. A. Paz de Araujo, J. C. Gelpey, Y. P. Huang, and R. Kwor, "Comparison of the Growth Kinetics of Oxides Grown in Tungsten-Halogen and a Water-cooled Arc Lamp System," *Mat. Res. Soc., Symp. Proc.,* **92,** 133 (1987).

58. H. Z. Massoud and J. D. Plummer, "Analytical Relationship for the Oxidation of Silicon in Dry Oxygen in the Thin-Film Regime," *J. Appl. Phys.,* **62,** 3416 (1987).

59. J. Nulman, J. Scarpulla, T. Mele, and J. P. Krusius, "Electrical Characteristics of Thin Gate Implanted MOS Channels Grown by Rapid Thermal Processing," in *Tech. Dig. IEDM,* Cat. No. 85CH2252-5, 376, IEEE, New York (1985), p. 376.

60. M. L. Reed and J. D. Plummer, "Kinetic Studies of Silicon-Silicon Dioxide Interface Trap Annealing Using Rapid Thermal Annealing," in *Mat. Res. Soc., Symp. Proc.* **52,** 333 (1986).

61. H. Fukuda, "Thin-Gate SiO_2 Films Formed by in situ Multiple Rapid Thermal Processing," T. Arakawa and S. Ohno, Eds., *IEEE Trans. Electron. Dev.* **39,** 127 (1992).

62. B. Lojek, "Point Defect Behavior During Rapid Thermal Annealing," unpublished (1992).

63. K. S. Jones, S. Prussin, and E. R. Weber, "Enhanced Elimination of Implantation Damage Upon Exceeding the Solid Solubility," *J. Appl. Phys.* **62,** 4114 (1987).

64. T. E. Seidel, D. J. Lischner, C. S. Pai, R. V. Knoell, D. M. Maher, and D. C. Jacobson, " A Review of Rapid Thermal Annealing (RTA) of B, BF_2 and As Implanted into Silicon," *Nucl. Inst. Methods Phys. Res. B,* **7/8,** 251 (1985).

65. K. S. Jones, "Extended Defects from Ion Implantation and Annealing," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993.

Micron Ex. 1023, p. 228
Micron v. YMTC
IPR2024-00909

66. R. B. Fair, "Junction Formation in Silicon by Rapid Thermal Annealing," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993.
67. M. D. Giles, unpublished (1990).
68. R. B. Fair, "Concentration Profiles of Diffused Dopants in Silicon," in F. F. Y. Wang, Ed., *Impurity Doping Processes in Silicon,* North Holland Press, Amsterdam, 1981.
69. F. Morehead and R. Hodgson, "A Simple Model for the Transient Enhanced Diffusion of Ion-Implanted Phosphorus in Silicon," in D. K. Biegelsen, G. A. Rozgonyi, and C. V. Shanks, Eds., *Energy Beam-Solid Interactions and Transient Thermal Processing,* **35,** Mat. Res. Soc., Pittsburgh, 1985.
70. M. Miyake, S. Aoyama, and K. Kurchi, "Transient Enhanced Diffusion of Ion Implanted B in Si during Rapid Thermal Annealing," in *Extended Abstracts,* Meeting of the Electrochem. Soc., Abs. #691, Honolulu (1987).
71. R. B. Fair, "The Role of Transient Damage Annealing in Shallow Junction Formation," *Nucl. Instr. Meth. Phys. Res. B,* **37/38,** 371 (1989).
72. S. Solmi, F. Baruffaldi, and R. Canteri, "Diffusion of Boron in Silicon During Post-implantation Annealing," *J. Appl. Phys.,* **69,** 2135 (1991).
73. Y. M. Kim, G. Q. Lo, H. Konoshita, and D. L. Kwong, "Roles of Extended Defect Evolution on the Anomalous Diffusion of Boron in Si during Rapid Thermal Annealing," *J. Electrochem. Soc.* **138,** 1122 (1991).
74. D. K. Sadana, S. C. Shatas and A. Gat, in *Microscopy of Semiconducting Materials,* Institute of Physics Conference Series, Institute of Physics, London (1983).
75. R. B. Fair, "Low-Thermal-Budget Process Modeling with the PREDICT™ Computer Program," *IEEE Trans. Electron. Dev.,* **35,** 285 (1988).
76. Y. Kim, H. Z. Massoud, amd R. B. Fair, "The Effect of Ion-Implantation Damage on Dopant Diffusion in Silicon During Shallow-Junction Formation," *J. Electron. Matls.,* **18,** 143 (1989).
77. S. Solmi, "Dopant Diffusion and Activation in Implanted Silicon," in R. B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 179.
78. S. Solmi, F. Baruffaldi, and R. Canteri, "Diffusion of Boron in Silicon During Post-Implantation Annealing," *J. Appl. Phys.*, **69,** 2135 (1991).
79. S. N. Hong, G. A. Ruggles, J. J. Wortman, and M. C. Ozturk, "Material and Electrical Properties of Ultra-Shallow p^+-n Junctions Formed by Low-Energy Ion Implantation and Rapid Thermal Annealing," *IEEE Trans. Electron. Dev.,* **38,** 476 (1991).
80. J. J. Wortman, unpublished (1991).
81. M. Finetti, P. Negrini, S. Solmi, and D. Nobili, "Electrical Properties and Stability of Supersaturated Phosphorus-Doped Silicon Layers," *J. Electrochem. Soc.,* **128,** 1313 (1981).
82. D. Nobili, A. Armigliato, M. Finetti, and S. Solmi, "Precipitation as the Phenomenon Responsible for the Electrically Inactive Phosphorus in Silicon," *J. Appl. Phys.,* **53,** 1484 (1982).
83. J. Kato and Y. Ono, "Phosphorus Diffusion Using Spin-On Phosphosilicate-Glass Source and Halogen Lamps," *J. Electrochem. Soc.*, **132,** 1730 (1985).
84. M. Orlowski, R. Subrahmanyan, and G. Huffman, "The Effect of Low Thermal Budget Anneals and Furnace Ramps on the Electrical Activation of Arsenic," *J. Appl. Phys.*, **71,** 164 (1992).
85. M. Y. Tsai, F. F. Morehead, J. E. E. Baglin, and A. E. Michel, "Shallow Junctions by High-Dose As Implants in Si: Experiments and Modeling," *J. Appl. Phys.,* **51,** 3230 (1980).

Micron Ex. 1023, p. 229
Micron v. YMTC
IPR2024-00909

86. C. M. Osburn, "Silicides," in R. B. Fair, Ed., *Rapid Thermal Processing: Science and Technology,* Academic Press, Boston, 1993, p. 227.
87. V. Q. Ho and D. Poulin, "Formation of Self-aligned $TiSi_2$ for VLSI Contacts and Interconnects," *J. Vac. Sci. Technol.,* **A5,** 1396 (1987).
88. R. D. J. Verhaar, A. A. Bos, H. Kraaij, R. A. M. Wolters, K. Maex, and L. Van denhove, "Self-aligned $CoSi_2$ in a submicron CMOS Process," in A. Heuberger, H. Ryssel, and P. Lange, Eds., *Proc. ESSDERC '89,* Springer-Verlag, Heidelberg, 229 (1989).
89. F. M. d' Heurle, "Material Properties of Silicides for VLSI Technology," in P. Balk and O. G. Folberth, Eds., *Solid State Devices,* Elsevier, Amsterdam, 1986, p. 213.
90. C. M. Osburn, H. Berger, R. P. Donovan, and G. W. Jones, "The Effects of Contamination on Semiconductor Manufacturing Yield," *J. Environ. Sci.,* **31,** 45 (1988).
91. C. M. Osburn, J. Y. Tsai, Q. F. Wang, J. Rose, and A. Cowen, "PREDICT 1.6: Modeling of Metal Silicide Processes," *Electrochem. Soc. Meeting, Honolulu,* May, 1993, p. 352.
92. B. Lojek, "Rapid Thermal Processing of Semiconductors 1963–1993. Where To from Here?" in R.B. Fair and B. Lojek, Eds., *1st International Rapid Thermal Processing Conference, RTP'93, Scottsdale, AZ,* 1993, p. 2.

PROBLEMS

1. Define emissivity and give examples of factors that influence intrinsic, extrinsic, and effective emissivity in an RTP chamber.

(*a*) Why does the intrinsic emissivity of a Si wafer become relatively independent of the wavelength of incident radiation at temperatures greater than 600°C ?

(*b*) Explain why the temperature of a deposited overlayer on a wafer undergoing RTA might depend on the overlayer's thickness.

(*c*) In (*b*), what would be the difficulty associated with trying to control wafer temperature during film growth by using an optical pyrometer in the temperature-control feedback loop?

(*d*) In (*c*), what factor does the reflectivity of the RTP chamber play in temperature measurement?

2. In a 0.25-μm CMOS process, what is the thermal budget for forming ion-implanted p^+ source/drain layers with surface doping concentrations of 1×10^{20} B/cm^3 if the RTA temperature is 900°C ?

(*a*) What kinds of device performance issues might a process engineer encounter with this thermal treatment?

(*b*) What must be done to create defect-free junctions for the 0.25-μm ion-implantation process without increasing the junction depth? (Assume 700 Å junction depth.)

(*c*) Why is it that doping processes such as epitaxy and CVD of solid dopant sources onto Si have larger thermal budgets than ion-implantation processes?

3. Discuss the pros and cons of batch CVD epitaxy versus LRP, single-wafer epitaxy from the standpoint of throughput, control, Si film uniformity, and cost.

Micron Ex. 1023, p. 230
Micron v. YMTC
IPR2024-00909

4. RTCVD of dielectric films requires high deposition rates for single-wafer processing to achieve high wafer throughput.

(*a*) What are the issues for applying RTCVD to thin-gate dielectric film deposition?

(*b*) What are the issues for applying RTCVD to interlevel dielectric formation at the back-end of the process?

(*c*) If SiH_4 and N_2O react to deposit SiO_2, what temperature uniformity across the wafer is required to deposit films with a 2% uniformity, assuming a 1.8-eV activation energy and a maximum wafer temperature of 800°C ?

5. Conventional batch-furnace oxidation has been used by manufacturers for over thirty years to grow high-quality, MOS-grade gate oxides, which are highly uniform across the wafer. Is there a compelling case to be made for replacing such a successful technology with single-wafer RTO for ULSI? Describe your analysis.

6. Explain why high-temperature RTA is preferable to low-temperature RTA for defect-free shallow junction formation.

(*a*) Why do you obtain shallower junctions with a high-temperature RTA (1000°C, 20 sec) than with a low-temperature RTA (800°C, 20 sec)? Assume a P implant.

(*b*) What is the basis for the activation of high-dose, ion-implanted B or P by high-temperature RTA when the resulting free-carrier concentrations exceed the solid solubility limit at the RTA temperature?

7. Assuming BF_2 ion implantation in Si with a dose of $1 \times 10^{15}\ cm^{-2}$, what is the maximum implant energy allowed to produce a shallow, defect-free p^+ junction less than 1000 Å deep using a 10-sec RTA?

(*a*) Does preamorphization using Ge implantation provide any benefits?

(*b*) What type of ion-implantation defects are produced by the BF_2 implant?

(*c*) Describe the two steps involved in the transient diffusion of the BF_2-implanted Si during RTA.

8. Why has silicide formation using RTA been easily incorporated into the manufacturing of ICs?

(*a*) Why is it necessary to use a two-step RTA process in silicide formation?

(*b*) What are the limitations of silicides for ultrashallow junction contacts?

(*c*) Using the concept of surface tension, explain top silicide surface roughening. *Hint:* The process of roughening is referred to as thermal grooving.

9. What are the primary design considerations in building an RTA system that yields uniform wafer temperature distributions (not necessarily uniform radiation)?

Micron Ex. 1023, p. 231
Micron v. YMTC
IPR2024-00909

CHAPTER 5

Dielectric and Polysilicon Film Deposition

H. C. Cheng

5.1 INTRODUCTION

Dielectric and polysilicon film deposition have been used extensively in integrated circuits (ICs) since the device planarization technique was developed. Dielectric films, including silicon dioxide and silicon nitride, are used as the isolation, mask, and passivation layers. Silicon dioxide deposited by chemical vapor deposition (CVD) has been shown to fit the requirements of gate dielectrics in ultralarge-scale integration (ULSI) circuits. On the other hand, polysilicon film can be used as the conducting layer, semiconductor, or resistor by proper doping with different impurities. Various deposition processes, such as CVD and physical vapor deposition (PVD), are described in Section 5.2. Different deposition schemes and the film qualities resulting from these deposition methods are described in Sections 5.3, 5.4, and 5.5 for silicon dioxide, silicon nitride, and polysilicon films. Other deposition techniques, e.g., plasma-assisted deposition, photo CVD, laser CVD, rapid-thermal-processing CVD (RTPCVD), and electron-cyclotron-resonance (ECR) CVD, are demonstrated. Finally, the applications of these deposited films and their future trends are presented in Sections 5.8 and 5.9, respectively.

5.2 DEPOSITION PROCESSES

5.2.1 Reaction Theories of Chemical Vapor Deposition (CVD) and Physical Vapor Deposition (PVD)

There are many deposition techniques for the material formations in ULSI technology. These methods can be classified into two main reaction mechanisms:

Micron Ex. 1023, p. 232
Micron v. YMTC
IPR2024-00909

chemical vapor deposition and physical vapor deposition. These two technologies are described as follows.

Chemical vapor deposition (CVD)

Chemical vapor deposition is defined as the formation of a nonvolatile solid film on a substrate by the reaction of vapor-phase chemicals (reactants) that contain the required constituents. It is most often used for semiconductor processing. It is a material synthesis process whereby the constituents of the vapor phases react chemically near or on a substrate surface to form a solid product.

Several steps must occur in every CVD reaction:[1,2]

1. Transport of reacting gaseous species to the substrate surface
2. Absorption, or chemisorption, of the species on the substrate surface
3. Heterogeneous surface reaction catalyzed by the substrate surface
4. Desorption of gaseous reaction products
5. Transport of reaction products away from the substrate surface

The sequence of reaction steps in a CVD process is illustrated in Fig. 1. In practice, the chemical reactions of the reactant gases leading to the formation of a solid material may take place not only on (or very close to) the wafer surface (heterogeneous reaction) but also in the gas phase (homogeneous reaction). Heterogeneous reactions are much more desirable, because such reactions occur selectively only on the heated surfaces and produce good-quality films. Homogeneous reactions, on the other hand, are undesirable, because they form gas-phase clusters of the depositing material, resulting in poorly adhering, low-density films with defects. In addition, such reactions also consume the reactants and can cause a decrease in deposition rates. Thus, one important issue of a chemical reaction for CVD application is the degree to which heterogeneous reactions are favored over gas-phase reactions.

The most common deposition methods are atmospheric-pressure CVD (APCVD), low-pressure CVD (LPCVD), and plasma-enhanced CVD (PECVD).[3,4,5] A comparison of APCVD and LPCVD shows that the benefits of the low-pressure deposition processes are uniform step coverage, precise control of composition and structure, low-temperature processing, high enough deposition rates and throughput, and low

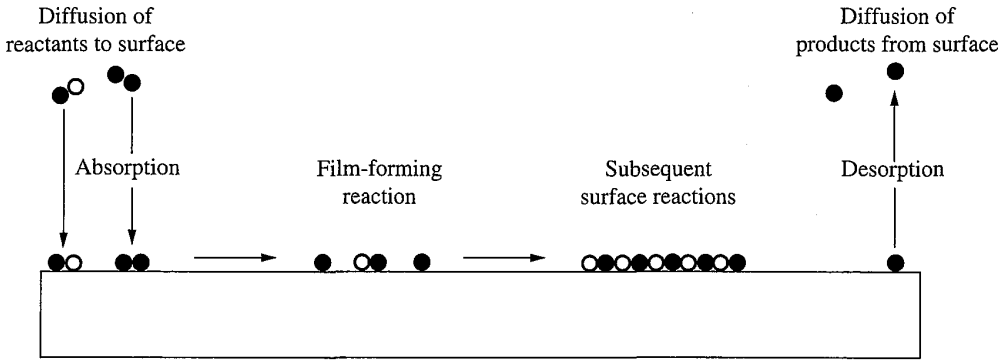


FIGURE 1
The sequence of reaction steps in a CVD reaction.

Micron Ex. 1023, p. 233
Micron v. YMTC
IPR2024-00909

processing costs.[6] Furthermore, no carrier gases are required in LPCVD, which reduces particle contamination. Hence, LPCVD is used widely in the highly cost-competitive semiconductor industry for film depositions. Epitaxial growth of silicon at a reduced pressure minimizes autodoping (contamination of the substrate by the dopant), a major problem in atmospheric-pressure epitaxy. The most serious disadvantage of LPCVD and APCVD is that their operating temperatures are high. PECVD is an appropriate method to solve this problem, as illustrated later.

Because the deposition process includes force convection, boundary-layer diffusion, surface absorption, decomposition, surface diffusion, and incorporation, there are several variables to be controlled. Temperature, pressure, flow rate, position, and reactant ratio all are important factors for high-quality films. The industry has optimized these conditions to improve the film properties.

Since the aforementioned steps for a CVD process are sequential, the one that occurs at the slowest rate will determine the deposition rate. The rate-determining steps can be grouped into gas-phase processes and surface processes. For the gas-phase process, the concern is the rate at which gases impinge on the substrate. This model considers the rate at which gases cross the boundary layer that separates the bulk regions of flowing gas and substrate surface. Such transport processes occur by gas-phase diffusion, which is proportional to the diffusivity of the gas and the concentration gradient across the boundary layer. The rate of mass transport is only relatively weakly influenced by the deposition temperature.

On the other hand, at low temperatures the surface reaction rate is reduced, and eventually the arrival rate of reactants exceeds the rate at which they are consumed by the surface reaction process. Under such conditions, the deposition rate is surface-reaction-rate-limited. Thus, at high temperatures, the deposition is usually *mass-transport-limited,* while at low temperatures it is *surface-reaction-rate-limited,* as shown in Fig. 2. In actual processes the temperature at which the deposition condition moves from one of these growth regimes to the other depends on the activation energy of the reaction and the gas flow conditions in the reactor. In processes that are under surface-reaction-rate-limited conditions, the deposition temperature is an important parameter. That is, uniform deposition rates throughout a reactor require conditions that maintain a constant reaction rate. This, in turn, implies that a constant temperature must also exist everywhere at all wafer surfaces. On the other hand, under such conditions the rate at which reactant species arrive at the surface is not so important, because their concentrations do not limit the growth rate. Thus,

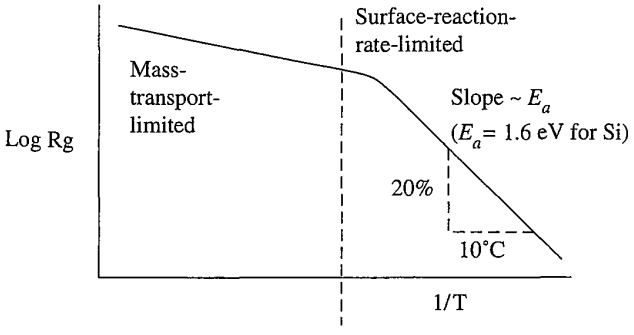


FIGURE 2
The deposition rate R_g is a rapid varying function of temperature T in the surface-reaction-limited regime of operation (low temperatures), whereas it changes only slowly with temperature in the mass-transport-limited regime (higher temperatures).

Micron Ex. 1023, p. 234
Micron v. YMTC
IPR2024-00909

it is not so critical that a reactor be designed to supply an equal flux of reactants to all locations of a wafer surface. It will be seen that in low-pressure CVD (LPCVD) reactors wafers can be stacked vertically and at very close spacing, because such systems operate in a surface-reaction-rate-limited mode. Under the low pressure of an LPCVD reactor (~1 torr), the diffusivity of the gas species is increased by a factor of 1000 over that at atmospheric pressure, and this is only partially offset by the fact that the boundary layer (the distance across which the reactants must diffuse) increases by less than the square root of the pressure. The net effect is that there is more than an order-of-magnitude increase in the transport of reactants to (and by-products away from) the substrate surface, and the rate-limiting step is thus the surface reaction. Since the surface reaction rate also generally increases with increasing surface concentration, nonuniform gas-phase concentrations produced by local depletion of reactants within the reactor can result in depletion nonuniformities. An example of such an effect is the depletion of reactants by their deposition on wafers located at the inlet of an end-feed reactor tube. Wafers near the outlet end are consequently exposed to lower concentrations of reactants than those at the inlet end of the tube. For film uniformity, the deposition temperatures at these two ends will thus have to be properly adjusted.

In deposition processes that are mass-transport-limited, the temperature control is not so critical. The mass transport process, which limits the growth rate, is only weakly dependent on the temperature. On the other hand, it is very important that the same concentration of reactants be present in the bulk gas regions adjacent to all locations of a wafer, because the arrival rates of reactants are directly proportional to the concentration gradient in the bulk gas. Thus, to ensure that films are uniform across a wafer, reactors operated in the mass-transport-limited regime must be designed so that all locations of the wafer surface and all wafers in a run are supplied with an equal flux of reactant species. Atmospheric-pressure reactors that deposit SiO_2 at ~400°C are operated in the mass-transport-limited regime. The most widely used APCVD reactor designs provide a uniform supply of reactants by horizontally positioning the wafers and moving them under a gas stream.

Physical vapor deposition (PVD)

Physical vapor depositions are different from CVD in the deposition mechanism. The deposition rate for CVD will be proportional to the deposition temperature. However, the deposition rate for PVD will generally decrease with increasing deposition temperature. The PVD technologies usually used are described as follows.

Evaporation technology. This is one of the oldest techniques used for depositing thin films. A vapor is first generated, by evaporating a source material in a vacuum chamber, then transported from the source to the substrate and condensed to a solid film on the substrate surface.

Molecular beam epitaxy (MBE). This is a sophisticated, finely controlled method for growing epitaxial films in an ultrahigh vacuum at 10^{-8} to 10^{-10} torr. The most widely studied materials are epitaxial layers of III–V semiconductor compounds and SiGe compounds. Figure 3 shows the schematic diagram of a molecular

Micron Ex. 1023, p. 235
Micron v. YMTC
IPR2024-00909

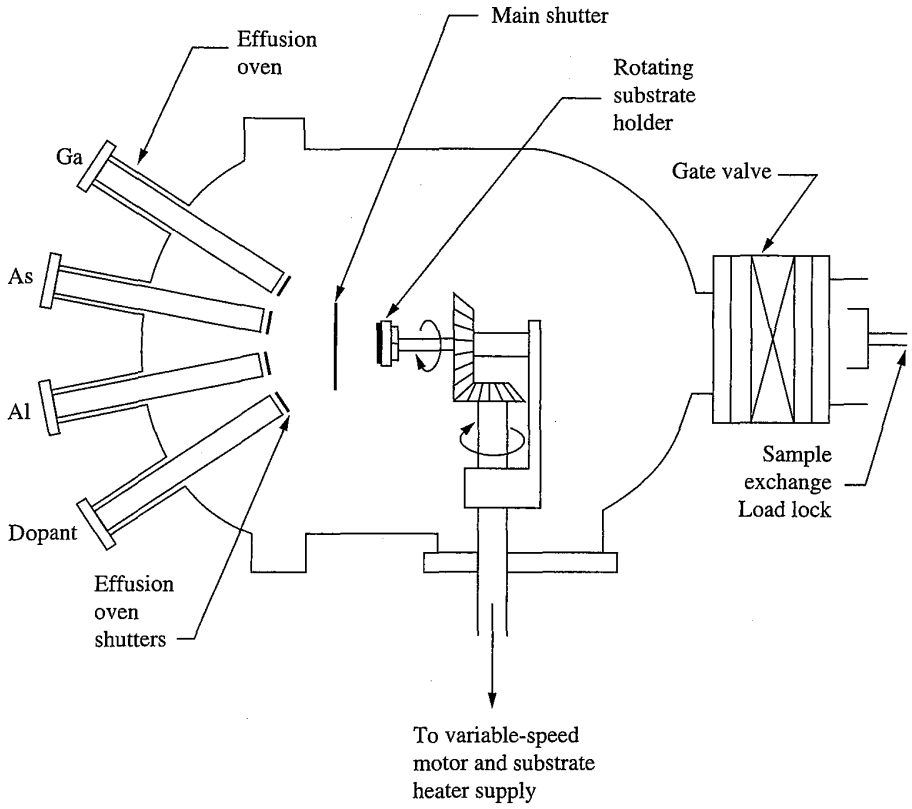


FIGURE 3
Schematic of a molecular beam epitaxy growth chamber.

beam epitaxy growth chamber. The evaporated species are transported at a very high velocity to the substrate. The relatively low vapor pressures of dopants and silicon ensures the condensation of deposited films on a low-temperature substrate.

Sputtering. This process involves the ejection of surface atoms from an electrode surface by momentum transfer from the bombarding ions to the electrode surface atoms. The generated vapor of electrode material is then deposited on the substrate. Sputtering processes, unlike evaporation, are very well controlled and are generally applicable to all materials, such as metals, insulators, semiconductors, and alloys. The schematic diagram of a sputtering system is shown in Fig. 4.

5.2.2 Equipment

The design and operation of CVD reactors depend on a variety of factors; hence, these reactors can be categorized in several ways. One way of grouping CVD reactors depends on the method used to heat the wafers. Another criterion used to distinguish reactor types is the pressure regime of operation (atmospheric-pressure reactors or reduced-pressure reactors). The reduced-pressure group can be further split into

Micron Ex. 1023, p. 236
Micron v. YMTC
IPR2024-00909

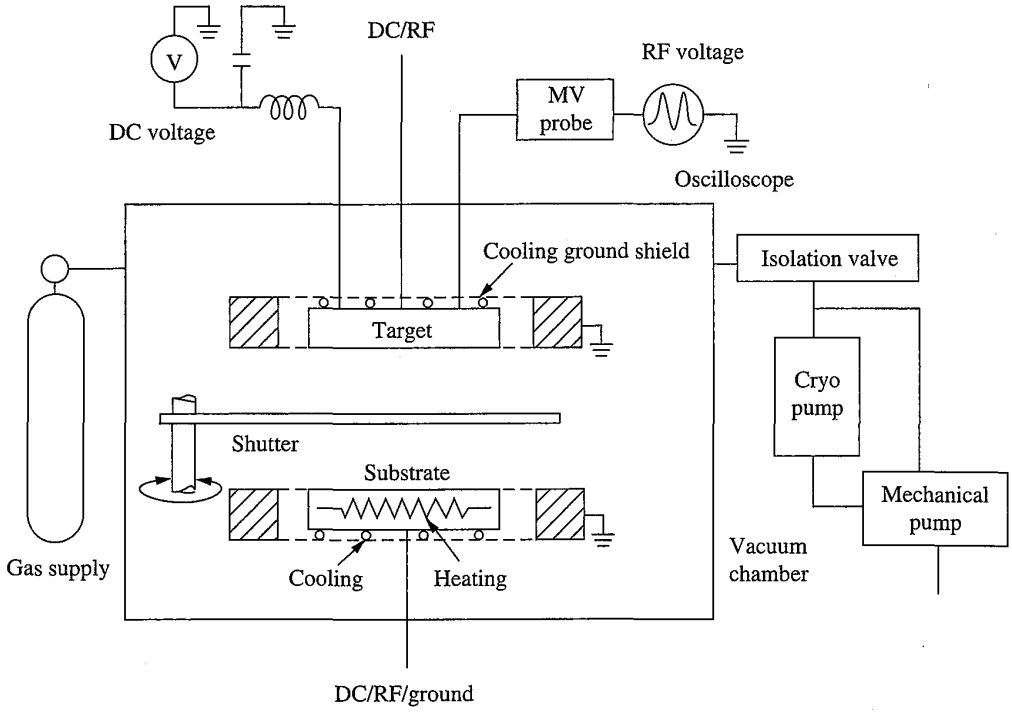


FIGURE 4
Schematic drawing showing some of the components of a sputtering system.

1. Low-pressure reactors, called low pressure-CVD (LPCVD) reactors, in which the energy input is entirely thermal
2. Those in which energy is partially supplied by a plasma as well, known as plasma-enhanced CVD or PECVD reactors

Each of the reactor types in these two pressure regimes is further divided into subgroups defined by reactor configuration and method of heating.

Four methods of heating the wafers have been adopted:

1. Resistance heating
2. RF induction heating
3. Heating by energy from a glow discharge (plasma)
4. Heating by photon energy

Energy may be transferred either to the reactant gases themselves or to the substrate. When radiant heating, from resistance-heated coils surrounding the reaction tube, is utilized, not only the wafer but also the reaction chamber walls become hot, and hence such designs are known as hot-wall reactors. In these systems, film-forming reactions occur on reaction chamber walls as well as on the substrate. This implies that such systems will require frequent cleaning to avoid particle contamination. On the other hand, energy input via RF induction or infrared lamps mounted within the reactor will heat the wafers as well as susceptors and does not cause appreciable heating of the reaction chamber walls. Consequently, systems heated by these methods are called cold-wall reactors. In some cold-wall systems, however, significant

Micron Ex. 1023, p. 237
Micron v. YMTC
IPR2024-00909

chamber wall heating can still take place, so means for cooling the walls must be implemented (e.g., water cooling) to prevent reactions or depositions on the wall.

Reactor geometry is constrained by the pressure regime and energy source and is an important factor in throughput. Since atmospheric-pressure reactors operate in the mass-transport-limited regime, they must be designed so that an equal flux of reactants is delivered to each wafer. As a result, wafers are never stacked vertically at close spacing but rather are laid flat on a horizontal surface. An undesirable consequence of this design is the high susceptibility of the wafers to incorporated falling particles. LPCVD reactors are not constrained by the mass-transfer-rate limitation, allowing designs that accommodate a large number of wafers per run. Wafers can be stacked side by side, only a few mm apart in a quartz reaction tube. Quartz wafer holders (*boats*) can hold up to 200 wafers. However, LPCVD reactors must be capable of precise temperature control, because they operate in the surface-reaction-limited mode. Table 1 summarizes the characteristics and applications of various CVD reactor designs.

For high-volume production of integrated circuits, the number of wafers processed simultaneously should be as large as possible. The requirement for high wafer capacity led to the development of the commercially important hot-wall, low-pressure CVD reactor, shown in Fig. 5. For maximum wafer capacity, the wafers are held vertically and separated from each other by a narrow space. As in a diffusion or oxidation furnace, the wafers are placed perpendicular to the gas flow in a quartz tube with a circular cross section. The gas flows by forced convection through the annular space between the wafers and the walls of the deposition chamber and then must travel along the narrow space between the walls. The pressures in the reaction chamber are typically 0.1 to 5.0 torr; the temperatures are between 300 and 900°C; and the gas flow rates range between 100 and 1000 standard cm^3/min (sccm). Special inserts that alter the gas flow dynamics are sometimes used. Large load space, good uniformity, and the ability to feed large-diameter wafers are the major advantages of this reactor. However, low deposition rates and the frequently used special gases are still problems.

TABLE 1
Characteristics and applications of CVD processes[58]

Process	Advantages	Disadvantages	Applications
APCVD (low-temperature)	Simple reactor, fast deposition, low temperature	Poor step coverage, particle contamination, low throughput	Low-temperature oxides (both doped and undoped)
LPCVD	Excellent purity and uniformity, conformal step coverage, large wafer capacity, high throughput	High temperature, low deposition rate	High-temperature oxides (both doped and undoped), silicon nitride, poly-Si, W, WSi_2
PECVD	Low temperature, fast deposition, good step coverage	Chemical (e.g., H_2) and particle contamination	Low-temperature insulators over metals, passivation (nitride)

Micron Ex. 1023, p. 238
Micron v. YMTC
IPR2024-00909

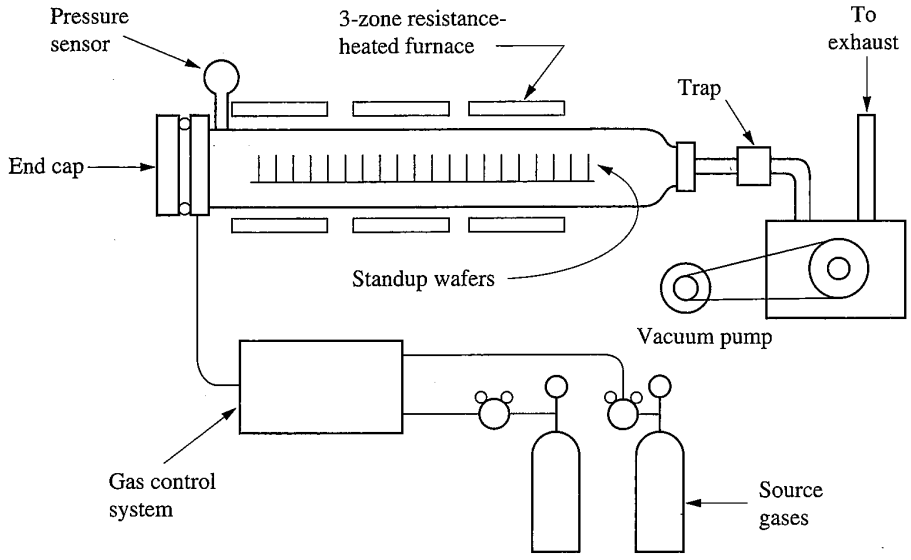


FIGURE 5
The hot-wall, low-pressure reactor is used for routine deposition of polysilicon because of its high wafer capacity and simplicity.

Atmospheric-pressure CVD (APCVD) reactors were the first to be used in the microelectronics industry.[7,8] Operation at atmospheric pressures keeps reactor design simple and allows high film deposition rates. APCVD, however, is susceptible to gas-phase reactions, and the films typically exhibit poor step coverage. Since APCVD is generally conducted in the mass-transport-limited regime, the reactant flux to all parts of every substrate in the reactor must be precisely controlled. Figure 6*a* shows the horizontal tube type. Such systems consist of a horizontal quartz tube, with the wafers lying flat on a fixed horizontal plate, while gas flows parallel to the wafer surface. Figures 6*b* and *c* show continuous-processing APCVD reactors. These configurations are the most widely used designs for depositing low-temperature CVD oxide films in production applications.[8] The samples are carried through the reactor on a conveyor belt. Reactant gases flowing through the center of the reactor are contained by gas curtains formed by a fast flow of nitrogen. These continuous reactors possess the benefits of high throughput, good uniformity, and the capability to process large-diameter wafers. However, they have the problems of high gas consumption and frequent need of reactor cleaning.

Figure 7 shows the diagram of a typical commercial PECVD system. Rather than relying solely on thermal energy to sustain chemical reactions, PECVD systems use an rf-induced glow discharge to transfer energy into the reactant gases, allowing the substrate to remain at a lower temperature than those in APCVD or LPCVD processes.[9,10,11] Low deposition temperature is the major advantage of PECVD, and in fact, PECVD provides a method of depositing films on substrates that do not have the thermal stability to accept coatings by other methods. In addition, PECVD can enhance the deposition rate, as compared with thermal reactions alone, and can produce films of unique compositions and properties. However, the limited capacity

Micron Ex. 1023, p. 239
Micron v. YMTC
IPR2024-00909

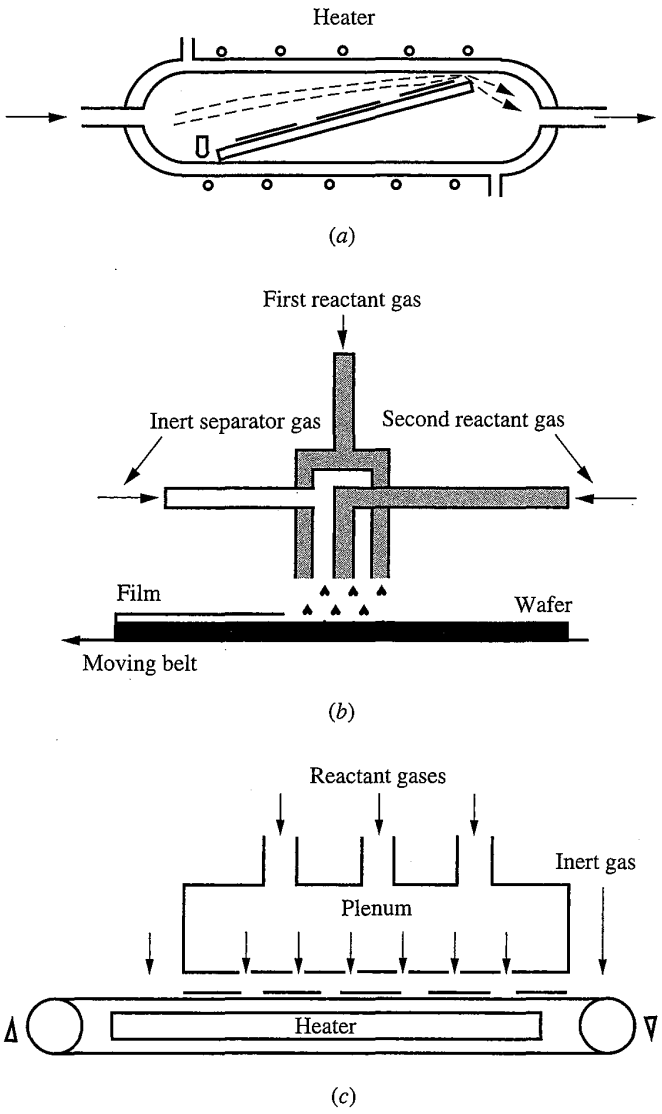


FIGURE 6
(*a*) Horizontal tube APCVD reactor. (*b*) Gas injection–type continuous-processing APCVD reactors. (*c*) Plenum-type continuous-processing APCVD reactor.

(especially for large-diameter wafers) and the possibility of particle contamination by loosely adhering deposits are still major concerns.

5.2.3 Safety Issues

Most of the gases used for film deposition are toxic. Those used for LPCVD are more unsafe than those used for atmospheric-pressure CVD, because pure silane is used

Micron Ex. 1023, p. 240
Micron v. YMTC
IPR2024-00909

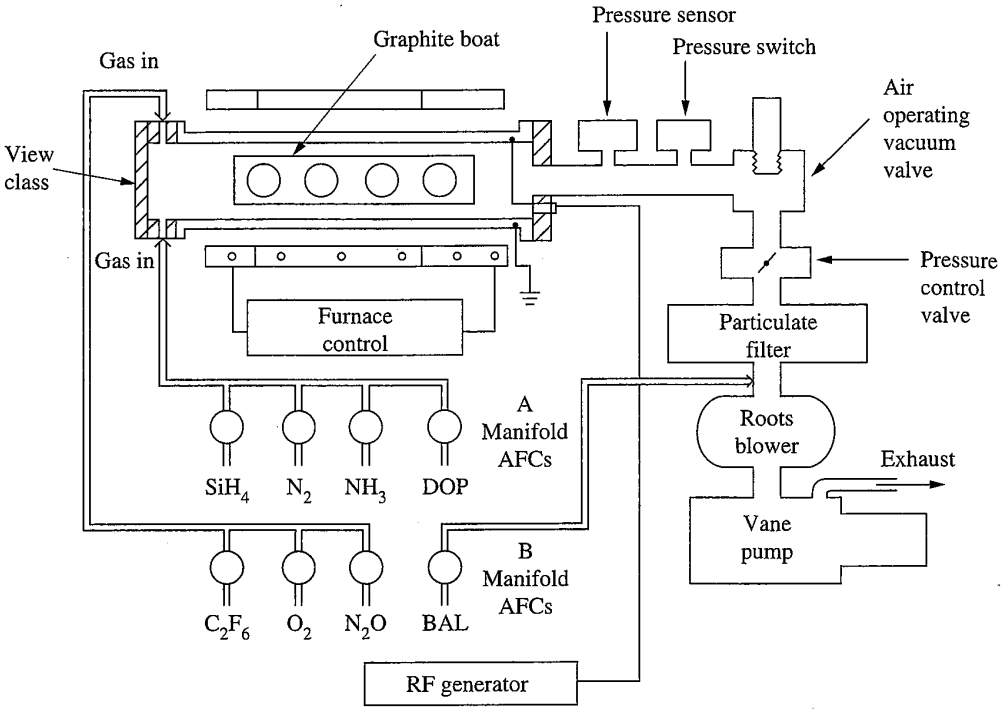


FIGURE 7
Diagram of a typical commercial PECVD system. (*Courtesy of Pacific Western Systems.*)

in LPCVD systems. In addition, hazardous gases can also cause a reaction with the vacuum pump oil.

These hazardous gases can be divided into four general classes: pyrophoric (flammable or explosive); poisonous; corrosive; and dangerous combinations of gases. Gases commonly used in CVD are listed in Table 2. Gas combinations such

TABLE 2
Gases commonly used in CVD

Gas	Formula	Hazard	Flammable limits in air (vol%)	Exposure limit (ppm)
Ammonia	NH_3	Toxic, corrosive	16–25	25
Argon	Ar	Inert	—	—
Arsine	AsH_3	Toxic	—	0.05
Diborane	B_2H_6	Toxic, flammable	1–98	0.1
Dichlorosilane	SiH_2Cl_2	Flammable, toxic	4–99	5
Hydrogen	H_2	Flammable	4–74	—
Hydorgen chloride	HCl	Corrosive, toxic	—	5
Nitrogen	N_2	Inert	—	—
Nitrogen oxide	N_2O	Oxidizer	—	—
Oxygen	O_2	Oxidizer	—	—
Phosphine	PH_3	Toxic, flammable	Pyrophoric	0.3
Silane	SiH_4	Flammable, toxic	Pyrophoric	0.5

Micron Ex. 1023, p. 241
Micron v. YMTC
IPR2024-00909

as silane with halogens, silane with hydrogen, and oxygen with hydrogen will cause safety problems.[12]

In addition, silane will react with air to form solid products, causing particle contaminations within the gas lines. These particles can plug the pipes and perhaps create combustion. Hence, the piping of the gas supply system must be carefully installed and maintained. The gas piping should be monitored and inspected before each operation to prevent the leakage of residual gases. To avoid the plugging of the gas products, burn-box systems for the hydride gases have been widely used in ULSI operations.

5.3 ATMOSPHERIC-PRESSURE CHEMICAL-VAPOR-DEPOSITED (APCVD) AND LOW-PRESSURE CHEMICAL-VAPOR-DEPOSITED (LPCVD) SILICON OXIDES

Dielectric layers have been inevitably used as the gate materials in transistors and capacitors. Dielectric layers can also be used to electrically isolate one level of conductor from another in multilevel-interconnect systems. The requirements of dielectric layers in multilevel interconnects are stringent, including

1. No residual constituents that will outgas during later processing
2. Good step coverage
3. Ease of etching
4. Stability up to temperatures of 500°C
5. Good adhesion
6. Low dielectric constant for frequencies up to ~20 MHz (to keep capacitance low between metal lines)
7. High breakdown field strength (>5 MV/cm)
8. No moisture absorption or permeability to moisture

Furthermore, there is a significant difference between the dielectric films used as poly–metal interlevel dielectrics (e.g., between poly, or other local interconnect–level materials, and metal 1) and the dielectric films employed between the metal layers (intermetal dielectrics). Poly–metal interlevel dielectric (PMD) films can be deposited (and densified if necessary) at a higher temperature than is possible for the intermetal dielectric layers. Furthermore, PMD films can be flowed and reflowed at temperatures in excess of 800°C. On the other hand, the maximum temperature of the intermetal dielectric layers is limited to 450°C or below, because Al is present on the wafer surface.

In general, silicon dioxides can be formed by a thermal annealing in oxygen ambient or by a chemical reaction (SiH_4 with O_2, $SiCl_2H_2$ with N_2O, or decomposition of $Si(OC_2H_5)_4$). The former can attain excellent electrical properties and is commonly used to fabricate the gate dielectrics in ULSI devices. The latter, called chemical vapor deposition (CVD), is often used for PMD, intermetal isolation, dopant masks, capping or passivation layers, diffusion sources, and even gate dielectrics in ULSI processes. This is because CVD can achieve more flexible characteristics, including large thickness and low-temperature processing.[13,14] Doped or undoped

Micron Ex. 1023, p. 242
Micron v. YMTC
IPR2024-00909

silicon dioxide films are widely used in various applications. Undoped silicon dioxide is used as an ion-implantation or diffusion mask, an insulating layer between multilevel metallization, a capping layer over doped regions to prevent outdiffusion, and an oxide layer to increase the thickness of thermally grown oxides. Undoped CVD oxides used as gate dielectrics and insulators between conducting layers are deposited and then densified by annealing. A solution containing fluoride or a CHF_3 plasma can etch the oxides to open the contact windows and via holes. Phosphorus-doped silicon dioxide is used as an insulator between metal layers, a final passivation over devices, and a gettering source.[15] Oxides doped with phosphorus, arsenic, or boron are used occasionally as diffusion sources.

Silicon nitride films have generally not been used as stand-alone PMD layers because they possess a much higher dielectric constant than SiO_2 films and cannot be flowed or reflowed. Hence, doped CVD SiO_2 films have found the widest application as PMD layers in MOS ICs. Although high-temperature CVD oxide films (deposited at 900°C by the reaction of dichlorosilane and nitrous oxide) possess properties almost identical to those of thermally grown SiO_2 films, these films cannot, unfortunately, be doped under the high temperatures at which they are deposited.[16] As a result, they cannot be flowed or reflowed at temperatures lower than 1100°C. Generally, phosphorus-doped CVD oxides are heated to a temperature between 950 and 1100°C so that the oxides soften and flow and form a smooth surface to obtain a good step coverage for the next metallization.[15,17] Boron-doped phosphosilicate oxides will further lower the reflow temperatures.

5.3.1 Deposition Methods

The methods used to deposit the silicon dioxide are classified by heating source, chemical reactants, reactor, and deposition temperature.

We first consider a method in which oxide films are deposited at temperatures below 500°C by reacting silane with oxygen. If phosphorus doping is included, the chemical reactions are expressed as

$$SiH_4(g) + O_2(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{5.1}$$

$$4PH_3(g) + 5O_2(g) \rightarrow 2P_2O_5(s) + 6H_2(g) \tag{5.2}$$

This process can be conducted in an atmospheric-pressure chamber (atmospheric-pressure chemical vapor deposition, APCVD) or in a reduced-pressure chamber (low-pressure chemical vapor deposition, LPCVD). Owing to the advantages of low deposition temperatures this method can be used to deposit silicon dioxides on the metal interconnect. However, the poor step coverage and high particle concentration characteristic of this technique require careful control.

Another method uses tetraethyl orthosilicate, $Si(OC_2H_5)_4$, abbreviated TEOS.[18] By decomposing the vaporized liquid TEOS, silicon dioxide can also be deposited in an LPCVD reactor at temperatures between 650 and 750°C. The reaction is written as

$$Si(OC_2H_5)_4(l) \rightarrow SiO_2(s) + \text{by-products}(g) \tag{5.3}$$

Micron Ex. 1023, p. 243
Micron v. YMTC
IPR2024-00909

where the by-products are organic and organosilicon compounds. LPCVD TEOS is often used to deposit the spacers beside the polysilicon gates. TEOS depositions have very good uniformity and step coverage, but the high-temperature process limits the application of LPCVD TEOS on aluminum interconnects.

Silicon dioxide can also be deposited by LPCVD at about 900°C. The reaction is expressed as

$$SiCl_2H_2(g) + 2N_2O(g) \rightarrow SiO_2(s) + 2N_2(g) + 2HCl(g) \qquad (5.4)$$

The advantage of this process is film uniformity. Like LPCVD TEOS, this process can deposit silicon dioxide over polysilicon, but its disadvantage is the chlorine-containing damage on the polysilicon.

Silicon dioxides can be doped by feeding dopant hydrides such as phosphine, arsine, and diborane into the CVD reactors. Such dopant compounds as halides or organic compounds can also be utilized.[20–23]

Recently, atmospheric-pressure and low-temperature CVD processes using TEOS and ozone (O_3) have been proposed,[24] as shown in Fig. 8. This CVD technology enables the formation of oxide films with high conformality and low viscosity under low deposition temperatures. The step angle depends on the ozone concentration, as shown in Fig. 9. In addition, the shrinkage of the oxide film during annealing is also a function of the ozone concentration, as shown in Fig. 10. Furthermore, the deposition temperature will determine the selectivity ratio of the O_3–TEOS films deposited on the SiO_2 to the Si substrate. This selectivity ratio will decrease with increasing deposition temperature, as shown in Fig. 11. Because of their porosity, O_3–TEOS CVD oxides are often accompanied by plasma-assisted oxides to execute the planarization technology in ULSI processing.

High-pressure techniques have been studied by many groups. The major advantages of high-pressure techniques stem mainly from their ability to use lower processing temperatures, which lead to reduced impurity diffusion and fewer substrate

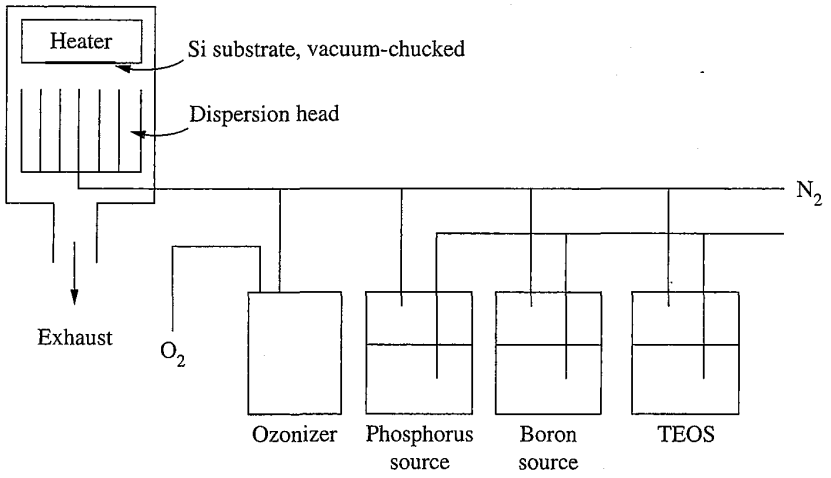


FIGURE 8
Experimental apparatus for the O_3–TEOS CVD system.

Micron Ex. 1023, p. 244
Micron v. YMTC
IPR2024-00909

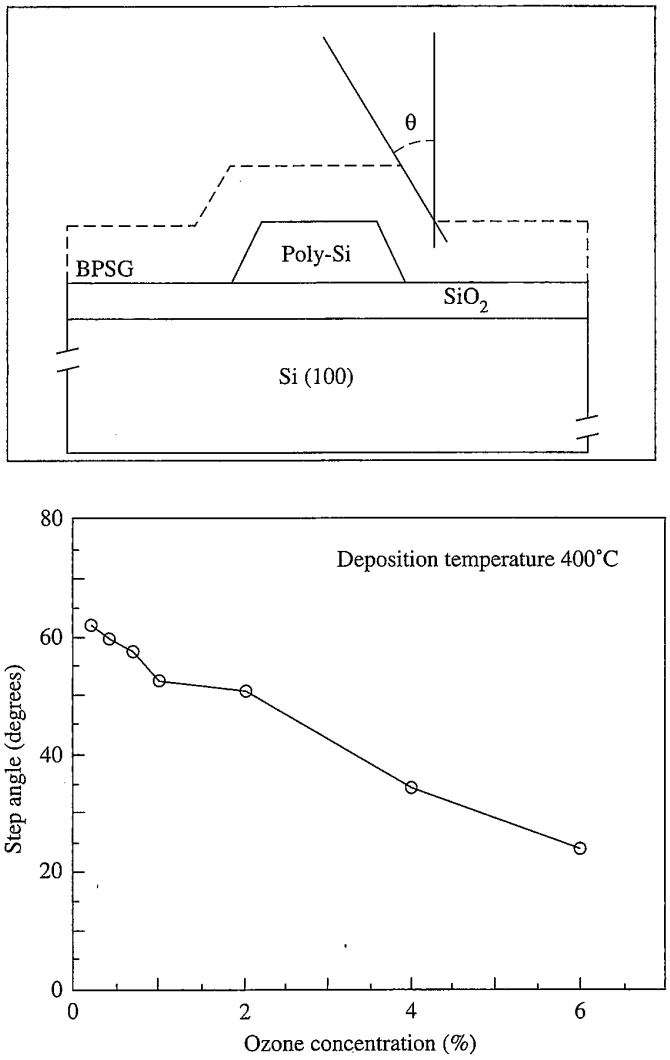


FIGURE 9
(*a*) Definition of step angle. (*b*) The step angle of the O_3–TEOS film as a function of the ozone concentration. (*Courtesy of SAMCO Company, Japan.*)

defects. However, the cost of facilities and the risk of danger discourage interest in high-pressure techniques.

5.3.2 Deposition Variables

Pressure, deposition temperature, reactant concentration, and dopants will affect the quality of silicon dioxide. For CVD the deposition rate of the silicon dioxide increases with deposition temperature. The silane–oxygen reaction needs less decomposition energy and, therefore, lower activation energy than TEOS decomposition. As the phosphorus dopants are added, the activation energies for these two cases

Micron Ex. 1023, p. 245
Micron v. YMTC
IPR2024-00909

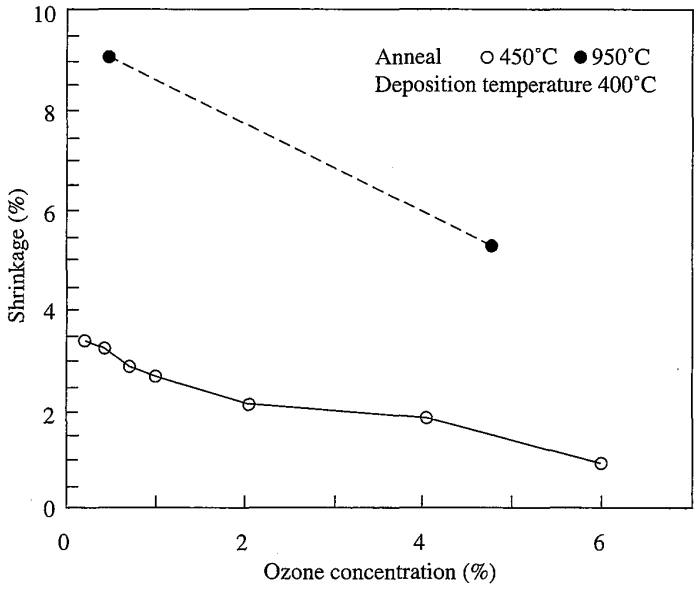


FIGURE 10
Dependence of the shrinkage of the O_3–TEOS film on ozone concentration during annealing. (*Courtesy of SAMCO Company, Japan.*)

will decrease. By introducing trimethylphosphite (TMOP) into the silane–oxygen or TEOS, phosphorus-doped oxides can also be formed. Figure 12 shows the dependence of deposition rate and the phosphorus content (mole % P_2O_5) on TMOP flow rate with a total TEOS/N_2 flow rate of 3.5 standard liters per minute (slm). Ozone concentration is another parameter that influences film characteristics, as shown in Fig. 13. The surface-reaction-rate-limited mode and the surface catalysis by phosphorus and ozone can be used to explain the differences of the activation energies.

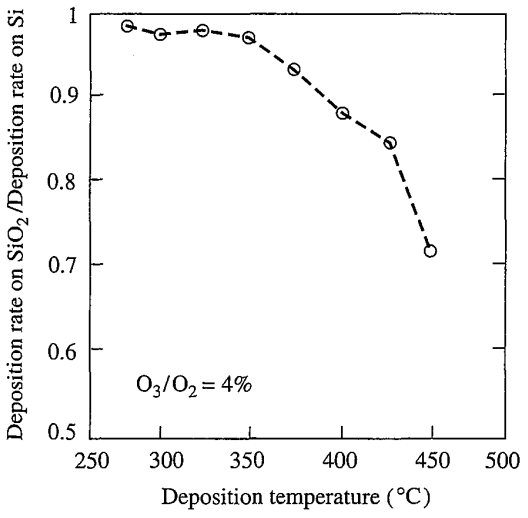


FIGURE 11
The selectivity ratio of the O_3-TEOS films deposited on the SiO_2 to the Si substrate as a function of the deposition temperature. (*Courtesy of SAMCO Company, Japan.*)

Micron Ex. 1023, p. 246
Micron v. YMTC
IPR2024-00909

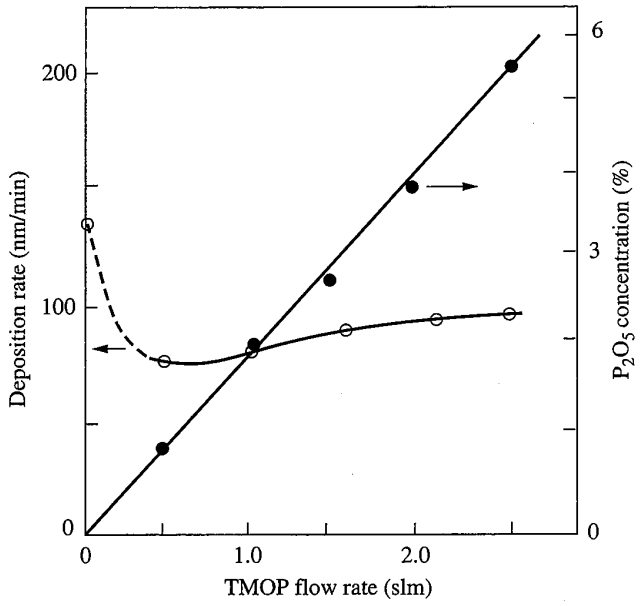


FIGURE 12
Dependence of deposition rate and P_2O_5 concentration of PSG films on TMOP/N_2 flow rate, deposited at 2.4% ozone and 400°C. (*After Fujino et al., Ref. 24.*)

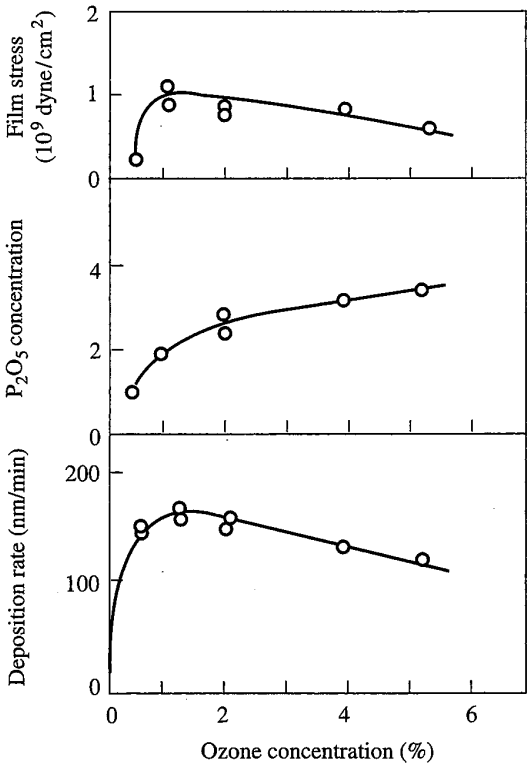


FIGURE 13
Dependence of deposition rate, P_2O_5 concentration, and film stress of PSG films on ozone concentration, deposited at 400°C. (*After Fujino et al., Ref. 24.*)

Micron Ex. 1023, p. 247
Micron v. YMTC
IPR2024-00909

As for the case of silicon dioxide deposited with dichlorosilane and nitrous oxide, a high deposition temperature (900°C) results in a mass-transport-limited reaction, exhibiting a strong pressure dependence. For the dichlorosilane–nitrous oxide reaction this doping is impossible, because the phosphorus oxides are volatile at the deposition temperature of 900°C.[25] Inert gases often make the gas flow over the wafer surfaces uniform and improve film uniformity. On the other hand, by introducing diborane, boron tricholoride, or trimethylborate into the silane–oxygen or TEOS reactions, boron-doped oxides can be obtained.

5.3.3 Step Coverage and Reflow

Figure 14 shows three types of step coverage often found for the deposited dioxides in a trench. A completely uniform or *conformal* step coverage means that the film thickness along the walls and at the bottom of the step is constant, as shown in Fig.

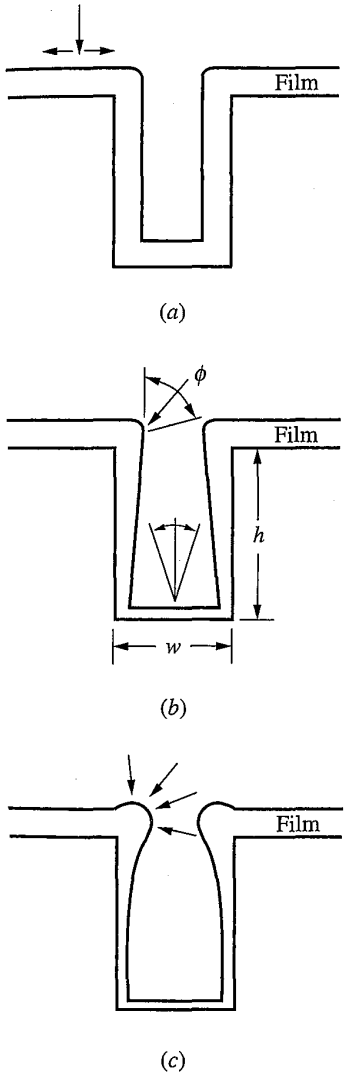


FIGURE 14
Step coverage of deposited films. (*a*) Uniform coverage resulting from rapid surface migration. (*b*) Nonconformal step coverage for long mean free path and no surface migration. (*c*) Nonconformal step coverage for short mean free path and no surface migration.

Micron Ex. 1023, p. 248
Micron v. YMTC
IPR2024-00909

14*a*. As the reactants or reactive intermediates adsorb on the surface and then rapidly migrate along the surface before reaction, the resulting films will have a uniform surface concentration on the substrate and a constant thickness.

On the other hand, the deposition rate is proportional to the arrival angle of the gas molecules if the absorbed reactants predominantly do not migrate. The arrival angle ϕ is expressed as

$$\phi = \arctan\left(\frac{w}{h}\right) \tag{5.5}$$

where w is the width of the via hole and h is the height from the top surface. Thus, the film thickness will gradually decrease along the step wall and even form an opening at the bottom, as shown in Fig. 14*b*. Furthermore, if the gas atoms have only a short mean free path, the arrival angle at the top of the step is 270°. Consequently, a thick cusp at the top and a thin tail at the bottom of the step will be formed, as shown in Fig. 14*c*; it will be unfavorable for subsequent metal deposition.

In general, doped oxides used for the reflow process contain 6 to 9 wt % phosphorus.[20] Silicon dioxide (glass) with higher phosphorus concentration will facilitate the reflow smoothing but is detrimental to the metallization, because of aluminum corrosion. After the doped silicon dioxide is deposited, a subsequent heating is necessary until the oxide softens and flows, as shown in Fig. 15. Therefore, the procedure is named *reflow*. In addition to the phosphorus concentration, the reflowing morphology of the doped silicon dioxide will also be determined by the heating temperature, heating time, heating rate, and heating ambient. In gen-

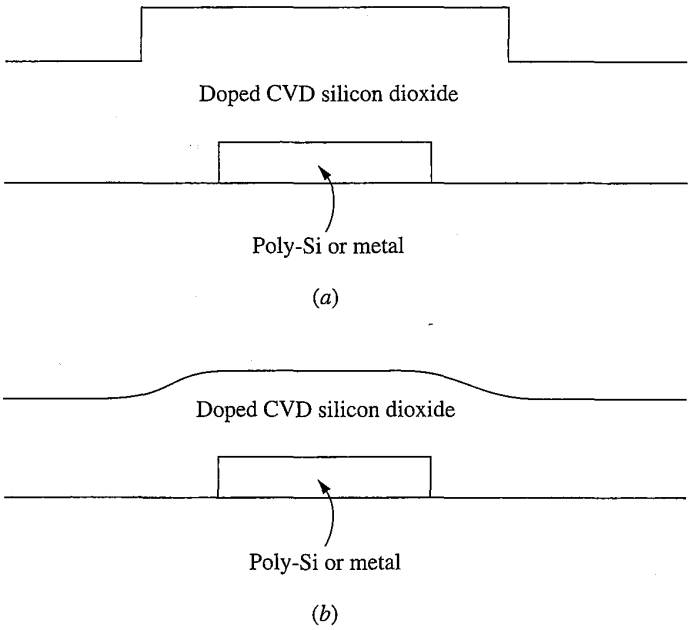


FIGURE 15
Schematic diagram of the reflow process. Polysilicon or metal step covered with doped CVD silicon dioxide (*a*) before and (*b*) after reflow process.

Micron Ex. 1023, p. 249
Micron v. YMTC
IPR2024-00909

eral, P-glass reflow processes need a heating temperature as high as 950 to 1100°C with a 15 to 30-min cycle. The temperatures will be determined by the phosphorus concentration.

Sometimes, boron dopants are added to the phosphorus-doped silicon dioxide to further reduce the softening temperature by decreasing the glass viscosity.[21] Silane-based BPSG films are deposited at low temperatures (400 to 450°C) and are then immediately densified at ~800°C for one hour.[22,23] The purpose of this step is to completely stabilize the BPSG films, which would otherwise be prone to blistering during subsequent processing. The higher the boron concentration in such films, the lower the reflow temperature; however, BPSG films with high boron concentrations (i.e., greater than 5%) are not stable. Typical dopant concentrations of boron are 1 to 4 wt % for the oxide films with 4 to 6 wt % phosphorus. On a weight percent basis, boron is 1.5 times more effective in the reflow process than phosphorus.

Furnace annealing times coupled with high temperatures often result in significant dopant diffusion, leading to remarkable junction movement, autodoping, and outdiffusion. In addition, high-temperature reflow for a long time also possibly causes further oxidation of the contact hole or substrate. These phenomena must be avoided in the fine-geometry ULSI devices. In order to prevent these problems, and also to prevent leach-out of phosphorus caused by steam reflow, a two-layer film is sometimes used for the PMD; that is, a thin (120-nm-thick) silicon nitride layer is deposited before the BPSG film. The nitride prevents dopants in the BPSG from diffusing into the polysilicon or substrate device regions during the flow and reflow thermal cycles. In advanced ULSI technology, an alternative method, utilizing rapid thermal annealing (RTA) at 1000 to 1150°C for 10 to 60 seconds, has been performed to reflow the BPSG without attacking the underlying devices, as shown in Fig. 16. Obviously, the step angle, reflecting the smoothing topography, will increase with rapid thermal annealing temperature and time.

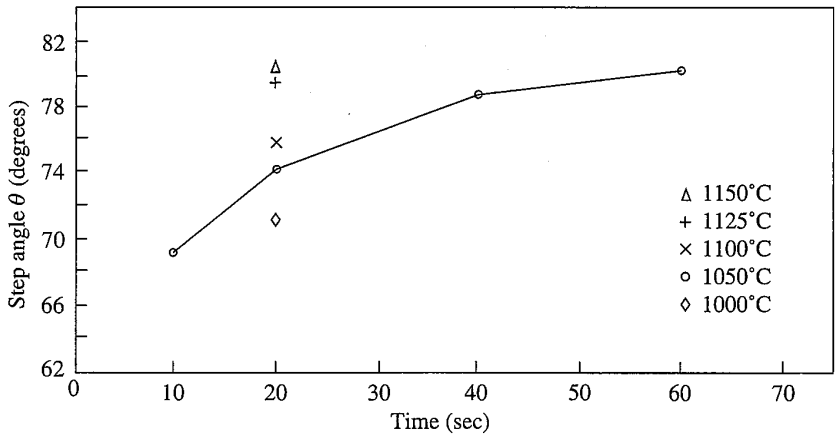


FIGURE 16
Step angles of LPCVD BPSG with concentrations of 3.6 wt % B and 3.9 wt % P, reflowed in the RTA system at different temperatures for various times. (*Courtesy of Peak Systems, Inc., California.*)

Micron Ex. 1023, p. 250
Micron v. YMTC
IPR2024-00909

The step coverage of deposited oxides can be further improved by *planarization* or *etch-back* techniques. An abrupt step in phosphorus-doped silicon dioxide (P-glass) is often covered with a sacrificial organic coating, called SOG (spin-on-glass) because the organic material usually has a smooth surface. Then the sample is plasma-etched to remove the organic material and the P-glass at equal rates. Since plasma etching preserves the original smooth surface of the coating, the P-glass is left with the same smooth contour. The planarization process reduces step heights independent of the phosphorus concentration in P-glass and requires temperatures less than 200°C. The disadvantages are the increased thickness of the original doped oxide film and the need for additional processing steps. Another planarization process uses PSG films accompanied by chemical/mechanical polishing. In this planarization technique, stacked O_3–TEOS/plasma oxide films are used to cover the polysilicon or metal layers[24] and are subsequently polished by chemical/mechanical methods to obtain a smooth surface. These planarizations will be further discussed in Section 5.8.

5.3.4 Properties of Silicon Oxide, Phosphosilicate Glass (PSG), Borophosphosilicate Glass (BPSG), and Tetraethyl Orthosilicate (TEOS) Films

In order to improve the device characteristics and to use the deposited dielectric and polysilicon films in various applications, the properties of these films are important. As an example, low-dielectric-constant films will be used for intermetal dielectric, whereas high-dielectric-constant films will be used for MOS capacitors. Hence, dielectric constant is one important property of the deposited films. Deposited oxide films commonly have an amorphous structure. By varying deposition temperatures or conditions, it is possible to achieve oxide films with different composition and density. In general, the thermal CVD schemes produce films with a thermodynamic-equilibrium composition. A subsequent densification heating of the CVD oxides at 600 to 900°C is necessary, because higher density of the deposited films is desirable. On the other hand, higher deposition temperatures improve the step coverage. Silicon dioxides deposited from TEOS and dichlorosilane will attain better conformal characteristics. To avoid deterioration during subsequent processing, the deposited films used in integrated circuits generally need high thermal stability. Low deposition temperature and impurities will be harmful to thermal stability. Other parameters, such as refractive index, film thickness, film stress, dielectric strength, and leakage current, all significantly determine the film properties and even the applications of these deposited oxide films. Refractive index and film thickness are often measured by the ellipsometer and the surface profiler. Silicon dioxide has a refractive index of about 1.46, depending on the process. Dielectric strength and leakage current are tested with the semiconductor parameter analyzer. A silicon-rich oxide film will have higher leakage current and is often used as the tunneling gate oxide in flash memory and EEPROM (electrically erasable programmable read-only memory) devices. Stoichiometric SiO_2 is used as the gate oxide because of its lower leakage current. For use in ULSI devices the reliability of the silicon dioxide is also very important. Hence, the time to failure and charge to breakdown under constant-voltage or constant-current stress are analyzed to determine the oxide quality. The stress

Micron Ex. 1023, p. 251
Micron v. YMTC
IPR2024-00909

in silicon dioxide can change the film quality. Tensile stress in silicon dioxide may induce cracks in the films. Stress in silicon dioxide depends on the deposition rate, deposition temperature, post-annealing cycle, dopant concentration, film porosity, and water content. It can be measured with a Michelson interferometer or optical reflectance meter. Hence, the processing of silicon dioxide should be carefully controlled.

5.4 LPCVD SILICON NITRIDES

Silicon nitride (Si_3N_4) is commonly used in passivation layers on integrated devices because of its ability to protect against the diffusion of impurities and water, which would make the devices unstable. The local oxidation of silicon (LOCOS) process also uses the silicon nitride as the mask. The patterned silicon nitride will prevent the underlying silicon from oxidation and leave the exposed silicon to be oxidized. Silicon nitride is also used as the dielectric for DRAM MOS capacitors when it combines with silicon dioxide. Silicon nitride can be chemical-vapor-deposited by reacting silane and ammonia at atmospheric pressure and temperatures ranging from 700 to 800°C:

$$3SiH_4(g) + 4NH_3(g) \rightarrow Si_3N_4(s) + 12H_2(g) \qquad \text{(APCVD)} \qquad (5.6)$$

or by reacting dichlorosilane and ammonia at reduced pressure and temperature between 700 and 800°C:

$$3SiCl_2H_2(g) + 4NH_3(g) \rightarrow Si_3N_4(s) + 6H_2(g) \qquad \text{(LPCVD)} \qquad (5.7)$$

Better uniformity and higher wafer throughput can be achieved by the low-pressure process.

5.4.1 Deposition Variables

Total pressure, reactant concentrations, deposition temperature, and temperature gradients in the furnace will affect the properties of silicon nitride. For example, increasing the total pressure and the partial pressure of dichlorosilane will increase the deposition rate, but increasing the ammonia will reduce the deposition rate. The deposition rates will increase with deposition temperature.

5.4.2 Properties of Silicon Nitride Films

As deposited, silicon nitride is in an amorphous phase and often contains a large amount of hydrogen, depending on the deposition temperature and the ratio of ammonia to dichlorosilane.[26,29] High ammonia ratio and low deposition temperature will increase the hydrogen content. These hydrogen atoms can passivate the dangling bonds of silicon and can sometimes be used for passivation of polysilicon devices. The refractive index of the deposited silicon nitride is about 2.01, but it changes with

Micron Ex. 1023, p. 252
Micron v. YMTC
IPR2024-00909

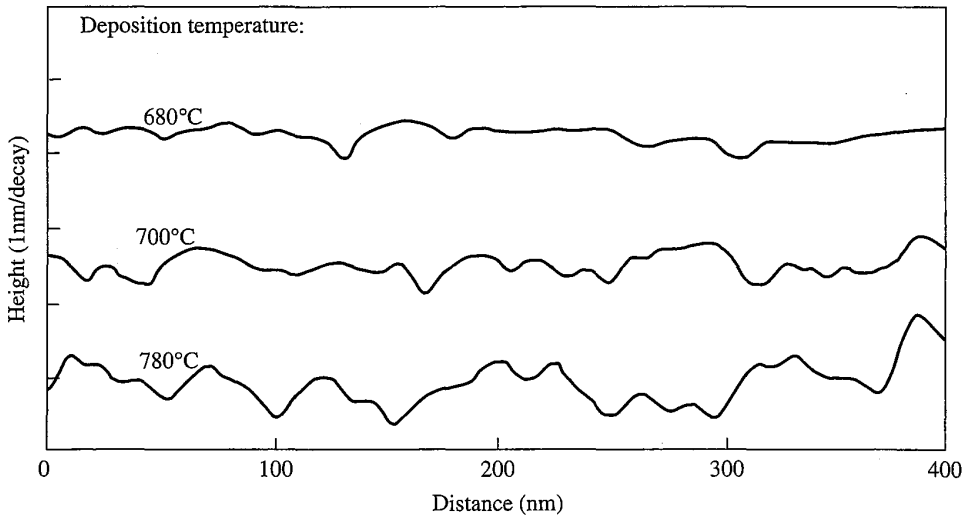


FIGURE 17
AFM surface profiles of Si_3N_4 films deposited at various deposition temperatures. (*After Tanaka et al., Ref. 28.*)

nitride composition. Silicon nitride films that contain oxygen are *silicon oxynitride* and have lower refractive indices. The composition of silicon oxynitride can vary from silicon dioxide, with a refractive index of 1.46, to silicon nitride, with an index of 2.01. On the other hand, silicon nitride has a refractive index that increases with increasing amounts of excess silicon. Because of high tensile stress, silicon nitride thicker than 200 nm may crack. The resistivity of silicon nitride at room temperature depends on film stoichiometry, deposition temperature, and the amount of impurities in the film.

In recent ULSI devices, thin CVD nitrides are extensively used in the dielectrics of the storage capacitors for dynamic random-access memories (DRAMs) and as interpoly dielectrics for nonvolatile memory devices. Because of their high dielectric constant and bulk-limited current, thin nitrides, oxidized in a pyrogenic ambient, have been used for high-density DRAMs. In hopes of shrinking the dimension of the dielectric, thin nitrides grown at different temperatures (680 to 780°C) were analyzed by atomic force microscopy (AFM) as shown in Fig. 17. Obviously, higher deposition temperatures result in rougher nitride surfaces. Hence, the ability of resisting oxidation increases with decreasing deposition temperature.[28]

5.5 LPCVD POLYSILICON FILMS

The low-pressure chemical-vapor deposition (LPCVD) method is also widely used in integrated-circuit manufacturing for the deposition of polycrystalline silicon films. Polysilicon films are still used as the gate electrodes in modern ULSI MOS devices. Furthermore, a metal or metal silicide, such as tungsten or tantalum silicide, is often deposited over the polysilicon gate to form a *polycide* or *salicide* structure and reduce

Micron Ex. 1023, p. 253
Micron v. YMTC
IPR2024-00909

the electrical resistivity.[29,30] Polysilicon can also be used for resistors, conductors, and diffusion sources. Other applications of poly-Si films, such as thin-film transistors in ULSI SRAM, stacked MOS capacitors in ULSI DRAM,[31,32] photovoltaic conversion, thermal and mechanical sensors, and large-area LCDs,[33,34] have been demonstrated. The polysilicon is deposited by dissociating silane at a temperature between 575 and 650°C. Either pure silane or 20 to 30% silane, diluted in nitrogen, can be fed into the LPCVD system at a pressure of 0.2 to 1.0 torr to fabricate the polysilicon films. The chemical reaction is expressed as

$$SiH_4(g) \rightarrow Si(s) + 2H_2(g) \tag{5.8}$$

For practical use a deposition rate of about 10 to 20 nm/min is required.

5.5.1 Deposition Variables

The properties of LPCVD polysilicon films are determined by deposition pressure, silane concentration, deposition temperature, and dopant content. In a low-pressure CVD reactor, deposition pressure can be changed by varying the pumping speed at a constant inlet gas flow or by varying the gas flow but keeping the pumping speed constant. To obtain better deposition reproducibility, the former is used. Figure 18 shows the results of the polysilicon depositions for different silane partial pressures.[35] At first, silane is adsorbed on the surface. The adsorbed silane then decomposes to form SiH_2 and adsorbed hydrogen. The SiH_2 further decomposes to form polysilicon and more adsorbed hydrogen. Finally, the adsorbed hydrogen

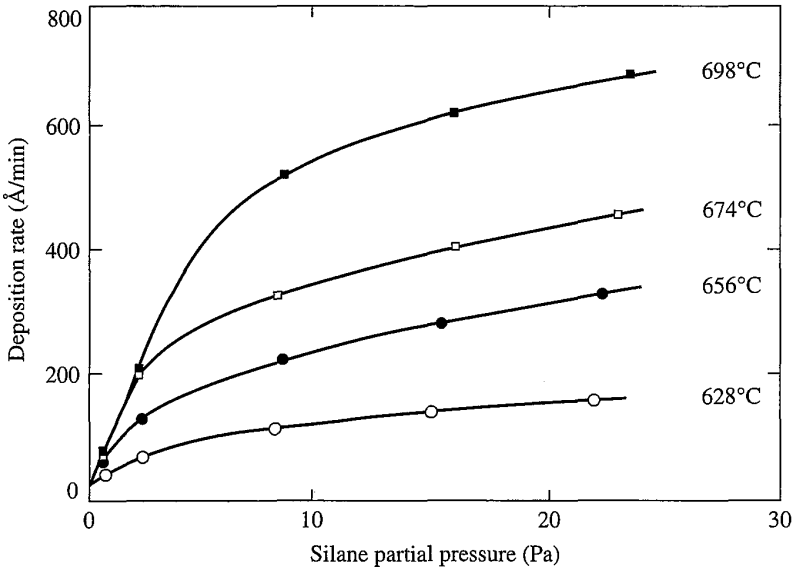


FIGURE 18
Deposition rate as a function of silane partial pressure for the polysilicon films deposited at different deposition temperatures. (*After Claassen et al., Ref. 35.*)

Micron Ex. 1023, p. 254
Micron v. YMTC
IPR2024-00909

leaves, and additional silane is adsorbed to continue this reaction. The reaction sequence is written as

$$SiH_4(g) \longleftrightarrow SiH_4(ad) \tag{5.9}$$

$$SiH_4(ad) \longleftrightarrow SiH_2(ad) + H_2(ad) \tag{5.10}$$

$$SiH_2(ad) \longleftrightarrow Si + H_2(ad) \tag{5.11}$$

$$2H_2(ad) \longleftrightarrow 2H_2(g) \tag{5.12}$$

If the last step, desorption of the hydrogen, is slow as compared to the other steps, the overall reaction rate is given by

$$\text{Rate} = K_1 P_S^{1/2}(1 + K_2 P_S^{1/2}) \tag{5.13}$$

where P_S is the partial pressure of silane and K_1 and K_2 are derived from the equilibrium constants for the individual reaction steps of Eqs. (5.10) and (5.11). Hence, the deposition rate is correlated to the square root of the silane partial pressure.

At high temperatures, the surface reaction rate becomes higher than the arrival rate of the unreacted silane to the surface. When this occurs, the deposition rate no longer increases with deposition temperature, and the reaction is mass-transport-limited. Deposition occurring in the mass-transport-limited regime depends on the reactant concentration, reactor geometry, and gas flow. When the surface reaction rate is smaller than the arrival rate of the reactant, as with low deposition temperature, the deposition is surface-reaction-limited and the critical variables are reactant concentration and deposition temperature. Depositions that are surface-reaction-limited have excellent thickness uniformity and step coverage. The deposition rate increases with deposition temperature. The temperature dependence is usually exponential and follows the Arrhenius equation

$$R = A \exp(-E_a/kT) \tag{5.14}$$

where R is the deposition rate, A is the frequency factor, E_a is the activation energy in eV, T is the absolute temperature in K, and k is the Boltzmann constant. The activation energies for polysilicon deposition are found to be about 1.7 eV. The Arrhenius equation is useful for finding the deposition rate at different deposition temperatures. For example, the deposition rates of 10 nm/min and 2.5 nm/min can be obtained directly from Eq. (5.14) for the deposition temperatures of 600°C and 550°C, respectively, if the deposition rate is 18.75 nm/min at 625°C and the activation energy is 1.7 eV. For ULSI technology, polysilicon films are generally deposited at the temperatures ranging from 500 to 650°C. Higher temperatures result in a rough, loosely adhering deposit and poor uniformity. For temperatures much lower than 550°C, the deposition rate is too low. However, an amorphous structure will be achieved for deposition temperatures below 575°C. A subsequent annealing can be used to transform the amorphous silicon films into polycrystalline silicon with even larger grain sizes. These large-grain polysilicon films have been used for many applications.

Polysilicon can be doped in situ by adding phosphine, arsine, or diborane to the reaction gases. Here an inert gas is often accompanied to improve the film uniformity. The addition of diborane will significantly enhance the deposition rates of polysilicon films. However, the addition of phosphine or arsine seriously decreases

Micron Ex. 1023, p. 255
Micron v. YMTC
IPR2024-00909

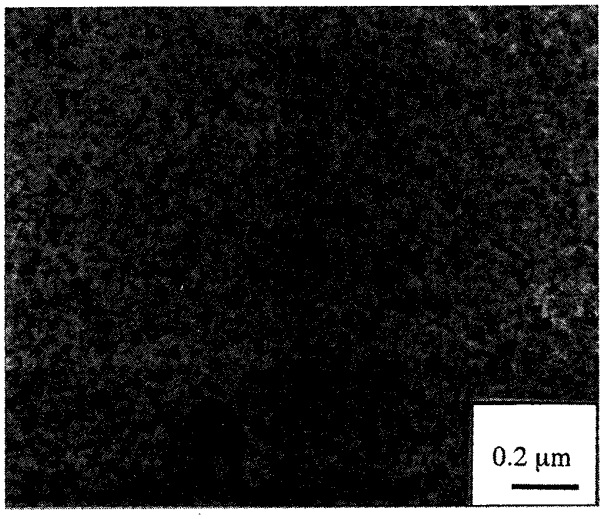
the deposition rates, because the phosphine or arsine is strongly adsorbed on the silicon substrate surface, inhibiting the dissociative chemisorption of SiH_4. However, diborane forms borane radicals, BH_3, that catalyze gas-phase reactions and increase the deposition rate.

5.5.2 Structures

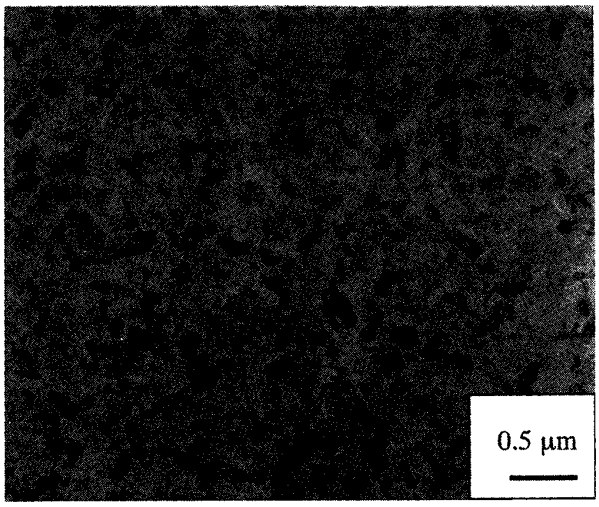
The structure of deposited polysilicon will be determined by the deposition temperature, deposition pressure, postdeposition annealing, dopant, and impurity contents. For deposition temperatures below 600°C the deposited silicon films will exhibit a microcrystalline or amorphous structure. For the microcrystalline phase, only grains with a size of tens to hundreds of angstroms (Å) will be detected using transmission electron microscopy (TEM). For the amorphous silicon, only the diffraction pattern of diffuse rings will be found in the TEM. Silicon films deposited above 600°C will possess the polycrystalline structure. The typical polysilicon deposited at 625°C generally has a grain size of about 0.1 μm with columnar grain structure. Note that the transition temperature, here 600°C, from the amorphous or microcrystalline to polycrystalline silicon films is not a constant but depends on the deposition pressure, deposition rate, and impurity or dopant incorporation. For example, the transition temperature will be lowered from 625°C to 575°C as the total pressure decreases from 50 Pa to 15 Pa. In general, the transition temperatures range between 550°C and 625°C. Polysilicon deposited at 600 to 650°C will have a {110} preferred orientation. For higher deposition temperatures, {100} orientation will increase. The texture will be affected by the deposition temperature, dopant, substrate material, and impurities.

After the silicon films are deposited, a subsequent annealing is required. The subsequent heating cycle can be used to crystallize the amorphous film and to increase the polysilicon grain size. The annealing processes can also be used for oxidation or dopant activation of the deposited silicon films. The polysilicon film as deposited has columnar grains, with grain size between 0.03 and 0.1 μm exhibiting a larger size on the top surface and a smaller size on the bottom interface. The deposited polysilicon grains do not significantly grow after a subsequent high-temperature annealing, as shown in Fig. 19*a*. On the other hand, the deposited amorphous silicon films or microcrystalline films will grow into the polycrystalline structure with a grain size even larger than those grown from a polycrystalline phase.[36] With a longer annealing time, the amorphous silicon can achieve a grain size larger than 0.2 μm, as shown in Fig. 19*b*. The technique has been popularly used in the thin-film transistors (TFTs) of SRAMs or LCDs addressed by TFTs to improve the electrical characteristics of the poly-Si TFTs. Other annealing schemes such as laser heating[37,38] can increase the polygrains from the amorphous type to a grain size as large as 10 μm, as shown in Fig. 19*c*. The electrical properties of TFTs fabricated with these polysilicon films are shown in Fig. 20. Higher I_{on}/I_{off} current ratio and field effect mobility have been achieved for the TFTs with larger grain sizes obtained by laser annealing (LA) as compared with those by furnace annealing (FA).[37,38] Recrystallization temperatures of the amorphous silicon films will also be affected by the dopants, impurities, and the heating methods. Oxygen, nitrogen,

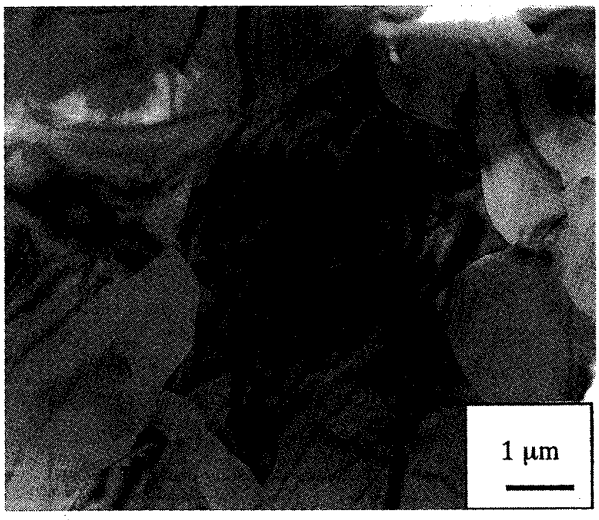
Micron Ex. 1023, p. 256
Micron v. YMTC
IPR2024-00909



(*a*)



(*b*)



(*c*)

FIGURE 19
(*a*) Brightfield (BF) TEM showing the as-deposited poly-Si films furnace-annealed at 600°C for 48 h. (*b*) BF showing the amorphous-Si films furnace-annealed at 600°C for 48 h. (*c*) BF showing the amorphous-Si specimens CO_2-laser-annealed at 22.5 W. (*After Tsai and Cheng, Ref. 36.*)

Micron Ex. 1023, p. 257
Micron v. YMTC
IPR2024-00909

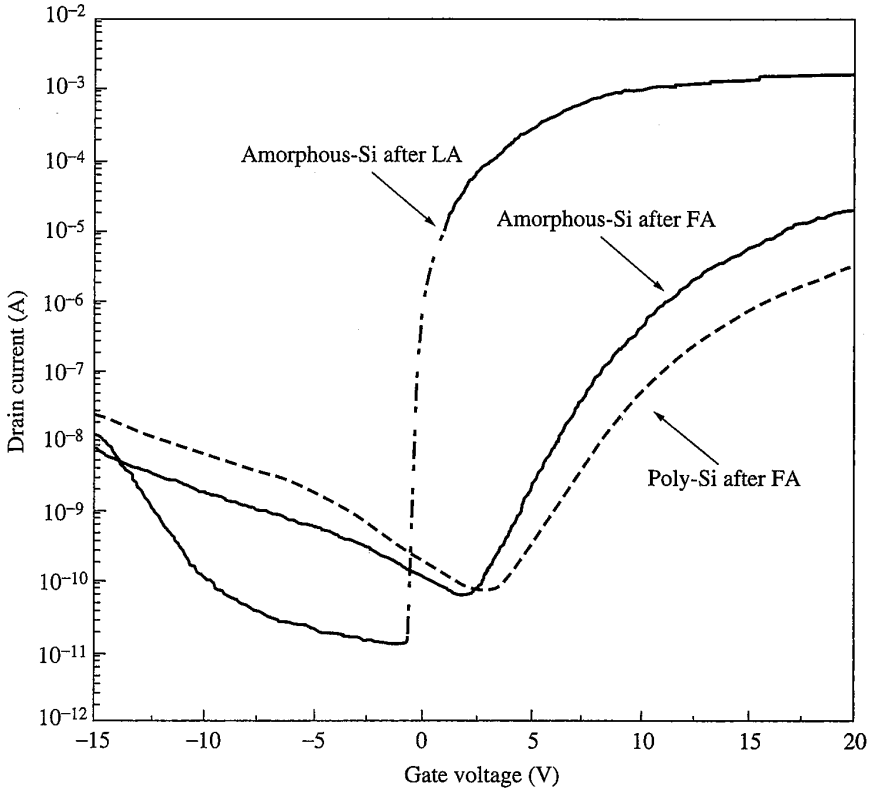


FIGURE 20
Transfer characteristics of TFT with the various recrystallized films shown in Fig. 19. $W/L = 25\ \mu m/10\ \mu m$ and $V_d = 5$ V. (*After Tsai and Cheng, Ref. 36.*)

and carbon will stabilize the amorphous films and raise the crystallization temperatures. Furthermore, the dopants, impurities, and heating methods will influence the preferred orientations. Table 3 lists the preferred orientations of the amorphous and polysilicon films after furnace annealing and laser annealing.

5.5.3 Doping Polysilicon

In general, implantation, diffusion, and in-situ doping are used to dope polysilicon films in ULSI technology.[39–41]

For dopant-implanted polysilicon films the resistivities are influenced by implant energy, implant species, implant dose, annealing temperature, and annealing time. Polysilicon films implanted with a low dose will lead to n-type grains alternating with intrinsic ones, because of the dopant traps in the grain boundaries. Polysilicon resistors can therefore be fabricated by this technique. Once these traps are saturated with dopants, the resistivity rapidly decreases. Diffusion of dopants results in low resistivities of polysilicon films. Because the dopants become segregated in the grain boundaries, the dopant concentrations in polysilicon films often exceed

Micron Ex. 1023, p. 258
Micron v. YMTC
IPR2024-00909

TABLE 3
Orientation of amorphous and polysilicon films

Samples	X-ray texture (%) (111)	(110)	(113)	(100)	(133)
As-deposited poly-Si	1.9	92.6	5.5	0	0
As-deposited amorphous-Si	—	—	—	—	—
Furnace-annealed poly-Si	3.4	85.5	6.8	2.6	1.7
Furnace-annealed amorphous-Si	30.2	25.8	27.3	9.1	7.6
Laser-annealed poly-Si (in solid-phase regime)	2.1	73.2	8.1	15.4	1.2
Laser-annealed amorphous-Si (in solid-phase regime)	11.5	14.2	26.3	38.2	9.8
Laser-annealed poly-Si (in liquid-phase regime)	14.7	13.6	8.2	58.8	4.7
Laser-annealed amorphous-Si (in liquid-phase regime)	4	2.4	4	79.4	10.2

the solid solubility limit. Polysilicon films doped in situ with phosphine, arsine, or diborane have resistivities that depend on the deposition temperature, dopant concentration, and annealing temperature. Increasing the deposition temperature will facilitate the decrease in the resistivities of polysilicon films; however, the resistivity still remains high. For in-situ boron doping, polysilicon can obtain a lower resistivity at a much lower deposition temperature than can be obtained for in-situ phosphorus doping.

For applications in the active region of TFTs, polysilicon films need to be grown to have larger grain size. Hence, an amorphous initial phase should be deposited and subsequently annealed for a long time or with a laser source. The device characteristics of the TFTs with polysilicon active layers can be improved by passivating the dangling bonds at the grain boundaries with dopants or hydrogen impurities, as mentioned in Section 5.5.4. For polysilicon films to be used as resistors, phosphorus implantation with appropriate low dose or double implantations of boron and phosphorus ions with an almost equal dose are utilized to increase the barriers at the grain boundaries. For polysilicon films to be used as conductors, such as gate electrodes or interconnections, high-dose doping will be necessary. In the future, low-temperature processes such as implantation and in-situ doping look promising. Implantation will offer precise control, and in-situ doping will have the benefit of a low-temperature process.

5.5.4 Properties of Polysilicon

Because polysilicon films have various electrical characteristics that depend on the doping processes, deposition conditions, post-annealing conditions, and passivation effects, they have been widely used in ULSI devices and other thin-film microelectronics. As mentioned above, the resistivity of polysilicon used in conductors is a main consideration. For such applications the doping concentration must be raised as

Micron Ex. 1023, p. 259
Micron v. YMTC
IPR2024-00909

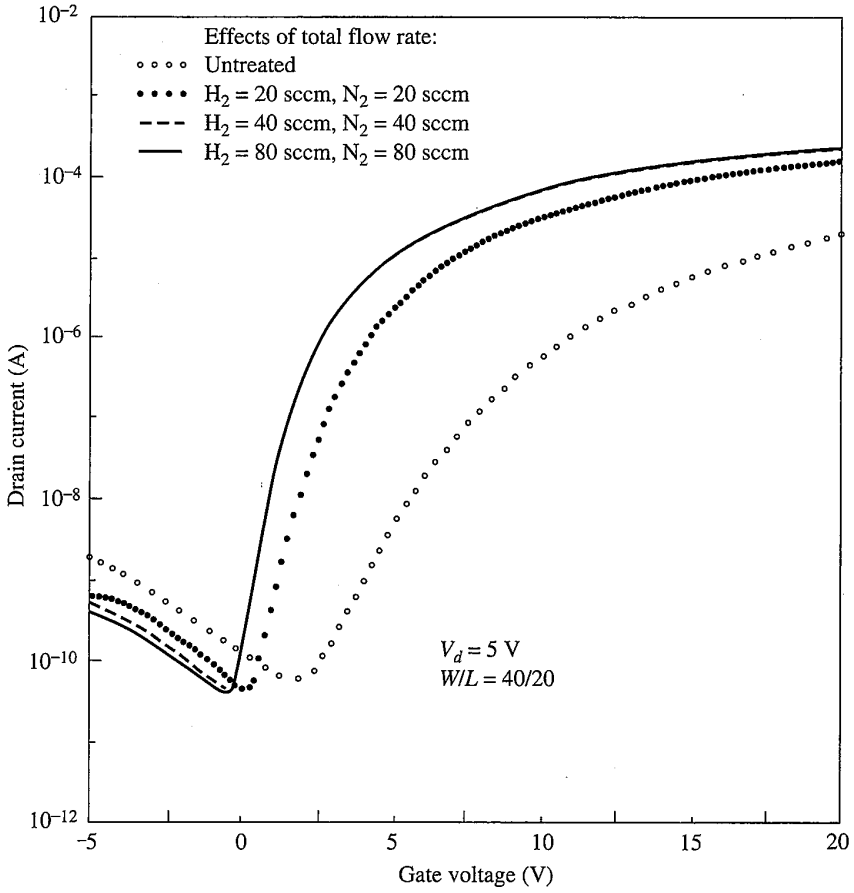


FIGURE 21
Characteristics of a TFT passivated with N_2–H_2 plasma for different flow rates. (*After Tsai and Cheng, Ref. 36.*)

high as possible. If polysilicon films are used as resistors, the doping concentration and the grain structure and silicon phase must be carefully controlled to increase the sheet resistance significantly. As for the use of polysilicon as a semiconductor, the deposition film phase, grain size, and passivation effects by hydrogen or nitrogen, etc., become very important. The previous section has described the electrical characteristics of TFTs fabricated with various polysilicon films. The electrical properties of TFTs are mainly determined by the grain sizes of these polysilicon films.

The passivation effects of polysilicon films on the electrical characteristics of TFTs are shown in Fig. 21. The on currents exhibit an increase for all samples passivated with hydrogen, nitrogen-diluted hydrogen, and nitrogen. Hydrogen has been reported to be capable of passivating the dangling bonds of polysilicon and improving the electrical properties of polysilicon.[42] The role of nitrogen in passivation is attributed to three effects. First, the added nitrogen would cause the formation of hydrogen radicals, which contain more hydrogen atoms to passivate the dangling bonds. Secondly, the added nitrogen suppresses the accumulation of hydrogen on the field oxide, which could induce surface leakage current. Thirdly, the nitrogen plasma

Micron Ex. 1023, p. 260
Micron v. YMTC
IPR2024-00909

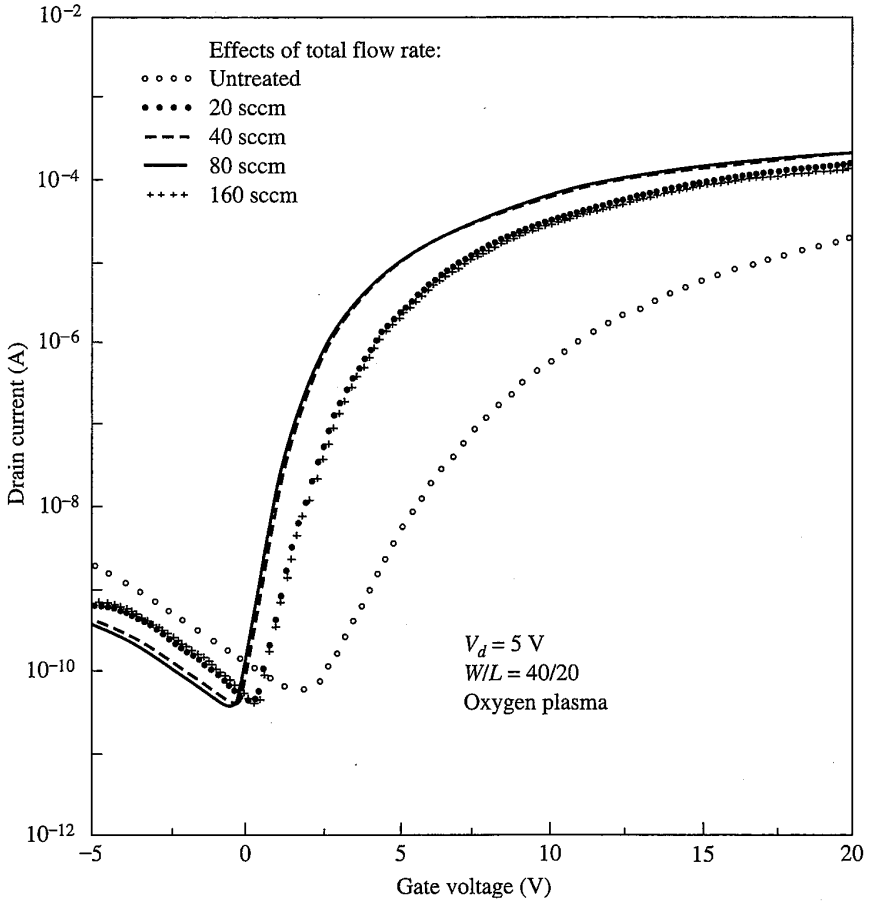


FIGURE 22
Characteristics of a TFT passivated with O_2 plasma for different flow rates. (*After Tsai and Cheng, Ref. 36.*)

could effectively passivate the poly-Si TFT by means of the three valence electrons of nitrogen. Hence, the TFT using nitrogen passivation also exhibits the lowest leakage current. Oxygen passivation effects on the polysilicon have also been found, as shown in Fig. 22. The passivation effects of various grain structures on polysilicon films are listed in Table 4. After hydrogen passivation, the laser-annealed samples still possessed better performance than the furnace-annealed specimens, because of the lower density of grain-boundary traps.

Polysilicon films are also utilized as dopant sources. As ULSI devices are scaled down, junction depths also shrink. With conventional implantation and activation, as well as the silicidation process, the control of shallow junction depth is difficult. A novel scheme, *implantation through polysilicon* (ITP), has been proposed to improve this phenomenon by using polysilicon as the diffusion source. By implanting dopant into thin poly-Si films, good junctions can be formed without these problems.[43,44] Figure 23 shows the leakage current J_r as a function of annealing temperature for the 1500Å-thick polysilicon films on silicon substrate with BF_2^+ implanted at various energies to a dose of 5×10^{15} cm^{-2}. Poly-Si serves as the implantation buffer layer

Micron Ex. 1023, p. 261
Micron v. YMTC
IPR2024-00909

TABLE 4

Effects of passivation on performance of TFTs with various grain structures

Samples	Before hydrogen passivation						After hydrogen passivation					
	$I_{d,min}$(pA)	S(V/decade)	V_t(volt)	μ_{eff}(cm^2/v · cm)	I_{on}/I_{off}	N_t(1/cm^2)	$I_{d,min}$(pA)	S(V/decade)	V_t(volt)	μ_{eff}(cm^2/v · cm)	I_{on}/I_{off}	N_t(1/cm^2)
AP†	124	2.76	18.5	3.36	3.1E4	1.85E12	119	0.85	4.0	16.2	1.2E6	7.23E11
HA‡	58	1.94	13.2	7.23	2.9E5	1.67E12	52	0.76	3.2	22.7	3.9E6	6.78E11
LA§	65	1.76	11.3	8.7	4.1E5	1.48E12	47	0.62	2.0	37	5.6E6	5.57E11
LL¶	10	0.198	0.49	162	1.0E8	5.08E11	8	0.178	-0.42	188	1.2E8	4.9E11
LS††	66	2.14	14.9	5.4	1.3E5	1.65E12	65	0.74	2.82	23.2	2.7E6	6.33E11

The channel width and channel length were 25μm and 10μm, respectively.

†AP: As-deposited poly-Si film.

‡HA: Amorphous-Si film after high-temperature grain growth.

§LA: Amorphous-Si film after low-temperature annealing.

¶LL: Amorphous-Si film after laser annealing under liquid-phase scheme.

††LS: Amorphous-Si film after laser annealing under solid-phase regime.

Micron Ex. 1023, p. 262
Micron v. YMTC
IPR2024-00909

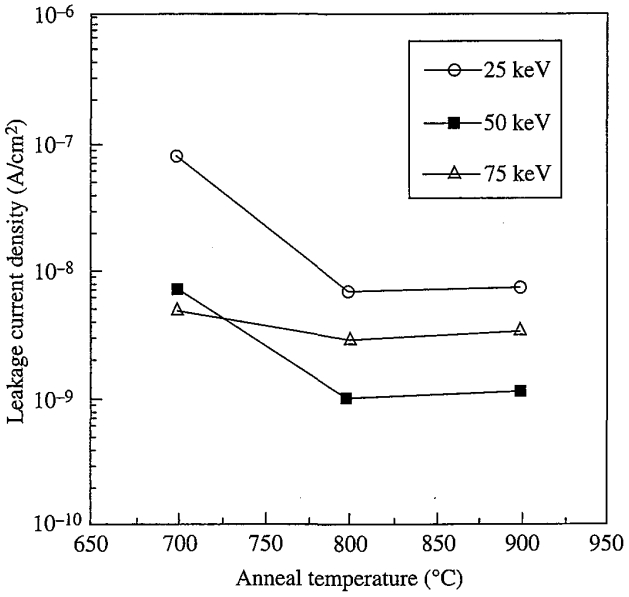


FIGURE 23
Dependence of leakage current density on annealing temperature for diodes fabricated with polysilicon films on silicon substrate samples BF_2^+-implanted at various energies to a dose of 5×10^{15} cm^{-2}. (*After Juang et al., Ref. 43.*)

as well as the doping source here. Excellent shallow p^+n junctions, which show a leakage of 1 nA/cm^2 and a junction depth of about 0.05 μm, have been formed by implanting BF_2^+ ions into thin poly-Si films and subsequent annealing.[43]

5.5.5 Oxygen-Doped Polysilicon

Oxygen-doped polysilicon, which has very high resistivity, is known as *semi-insulating polysilicon* (SIPOS).[45] SIPOS is deposited at temperatures ranging from 600 to 700°C. Silane mixed with a small quantity of nitrous oxide is fed into the reactor; the reactions are expressed as

$$SiH_4(g) \rightarrow Si(s) + 2H_2(g) \tag{5.15}$$

$$SiH_4(g) + xN_2O(g) \rightarrow SiO_x(s) + 2H_2(g) + xN_2(g) \tag{5.16}$$

SIPOS will have multiple phases consisting of crystalline silicon, amorphous silicon, silicon dioxide, and silicon monoxide. The resulting properties of SIPOS will be influenced by the amount of nitrous oxide, the deposition temperature, and the post-annealing condition. For practical use, SIPOS films generally have oxygen concentration of about 30 at%.

5.5.6 Oxidation of Polysilicon

Recently, polycrystalline silicon has become an important material in silicon technology. In most integrated circuits, *polyoxides* grown on polysilicon films have been widely applied in SRAMs, EPROMs,[46] EEPROMs, flash memories, and thin-film transistors. Polyoxides are often used to isolate the devices and are generally fabricated by annealing the polysilicon films in dry or wet oxygen ambient at 800

Micron Ex. 1023, p. 263
Micron v. YMTC
IPR2024-00909

to 1000°C. The polyoxide thickness depends on the ambient, annealing temperature, and annealing time. Furthermore, polysilicon oxidation must be achieved in a short duration if the polyoxides are to be used as gate dielectrics.

For polysilicon films by LPCVD at around 625°C, the oxidation mechanism is predominantly surface-reaction-limited. The oxidation rates depend on the orientations of silicon grains, which cause the roughness of the interface between the polyoxide and the polysilicon layer. At higher oxidation temperatures the oxidation mechanism is predominantly diffusion-controlled, resulting in a smoother polyoxide/polysilicon interface.[47,48] The high electrical conductivity found for the polyoxides is ascribed to the enhanced local electric field. On the other hand, the polyoxides of amorphous silicon films deposited at 560°C and subsequently annealed at 600°C to form the polycrystalline structures are reported to have breakdown fields higher than those of the polysilicon films as deposited at 620°C. Both polysilicon films deposited at 625°C and amorphous films deposited at 550°C with subsequent annealing at 600°C for 24 hours to form polysilicon films are used to grow polyoxides at 850°C in wet or dry oxygen. For the recrystallized amorphous silicon specimens, the polyoxide/polysilicon interface is very flat, as shown in Fig. 24. On the other hand, the polyoxide/polysilicon interface is rough for the recrystallized as-deposited polysilicon samples, as shown in Fig. 25. Cross-sectional HRTEM (high-resolution transmission electron microscope) observation shows that intergranular oxidation occurs for the polysilicon specimens after wet oxidation at 850°C for 60

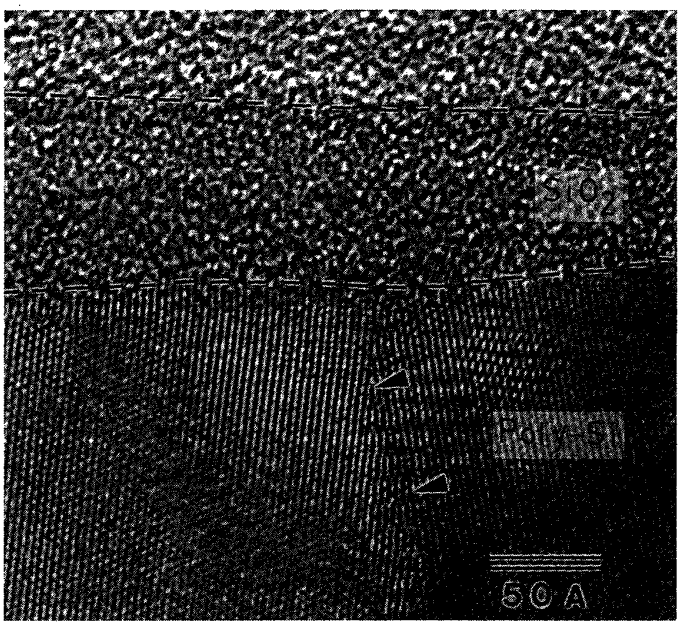


FIGURE 24
Cross-sectional high resolution TEM (HRTEM) for wet polyoxide grown on the amorphous specimen at 850°C for 150 s in wet oxygen ambient, showing the smooth polyoxide/polysilicon interface on the low-angle grain boundary. (*After Wang et al., Ref. 48.*)

Micron Ex. 1023, p. 264
Micron v. YMTC
IPR2024-00909

FIGURE 25
Cross-sectional HRTEM image of polysilicon samples after dry oxidation at 850°C for 18 min, indicating the annealed as-deposited polyoxide thickness fluctuation on the high-angle grain boundary, marked by the arrows. (*After Wang et al., Ref. 48.*)

Micron Ex. 1023, p. 265
Micron v. YMTC
IPR2024-00909

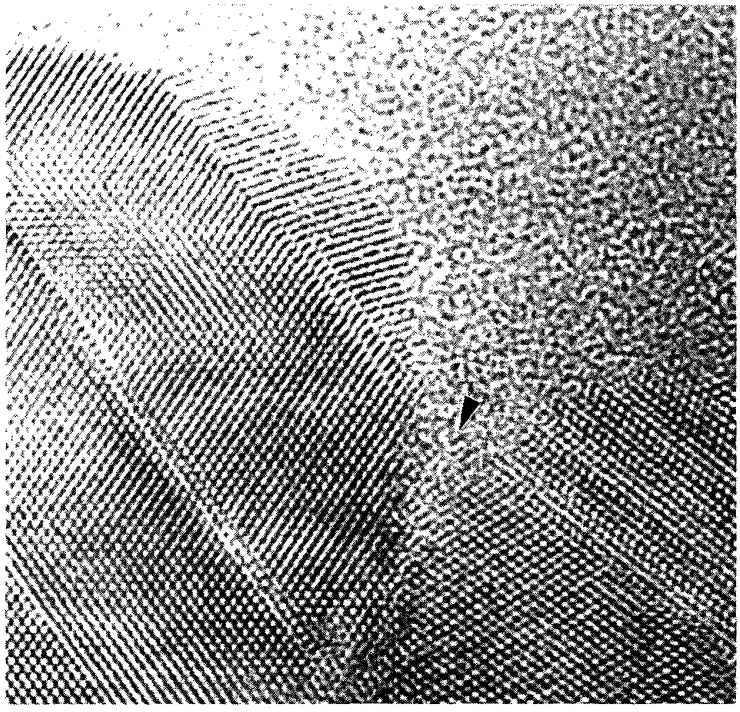
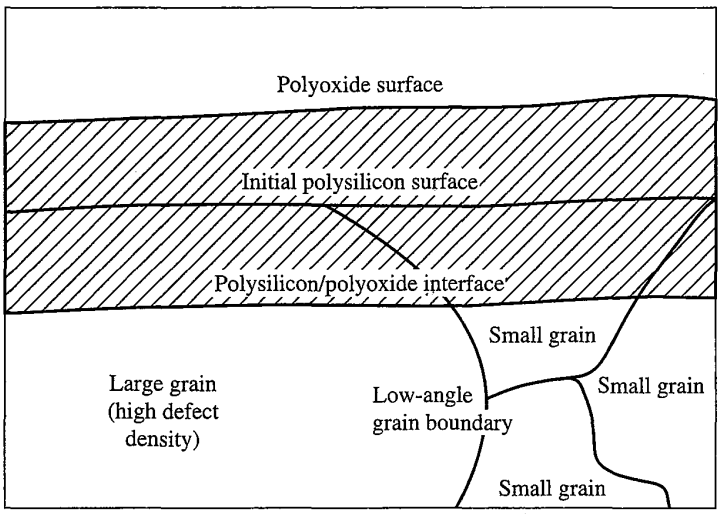


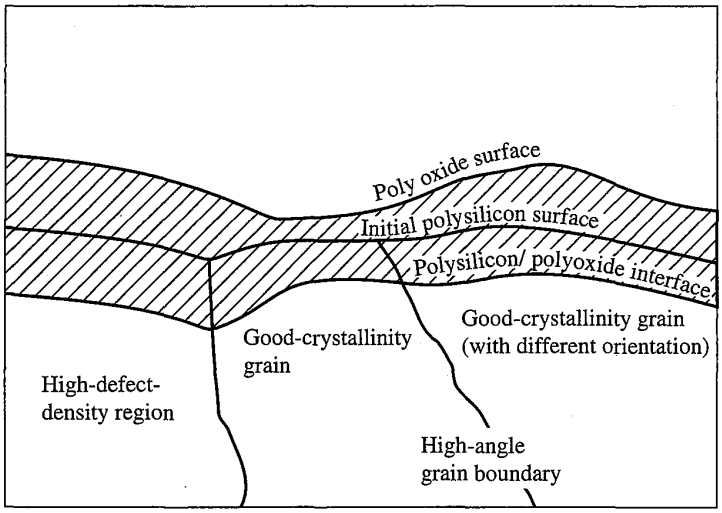
FIGURE 26
Cross-sectional HRTEM image showing the intergranular oxidation phenomenon, marked by arrows, for the polysilicon recrystallized as-deposited specimens wet-oxidized at 850°C for 60 min. (*After Wang et al., Ref. 48.*)

min, as shown in Fig. 26. The oxidation phenomenon of recrystallized amorphous samples is illustrated in the schematic diagram in Fig. 27*a*. Due to the poor crystallinity and low-angle grain boundaries of the recrystallized amorphous films, defect density is not much higher at grain boundaries than in the other parts of polysilicon films. Hence, the oxidation rates are similar for regions of highly damaged large grains, grain boundaries, and small or microcrystalline grains. On the other hand, the recrystallized as-deposited polycrystalline specimens possess good crystallinity and high-angle boundaries. The oxidation rates at the highly damaged region and grain boundaries are very different from those at the grains with good crystallinity. This results in the polyoxide/polysilicon interface roughness. The oxidation phenomenon of the as-deposited polysilicon specimens is illustrated in Fig. 27*b*. Figure 28 shows the Weibull plots of the time-zero-dielectric-breakdown (TZDB) characteristics for the polyoxides grown on the annealed as-deposited polysilicon films and the annealed amorphous silicon films. A wide range of breakdown events, which distribute from 1 to 7 V for the as-deposited polysilicon specimens and a narrow distribution for the recrystallized amorphous samples, is observed. Much better electrical characteristics have been achieved by using polyoxides on the recrystallized amorphous silicon films.[48]

Micron Ex. 1023, p. 266
Micron v. YMTC
IPR2024-00909



(*a*)



(*b*)

FIGURE 27
(*a*) Schematic diagram illustrating the oxidation phenomenon of the recrystallized amorphous samples. (*b*) Schematic diagram depicting the oxidation phenomenon of the as-deposited polysilicon samples.

5.6 PLASMA-ASSISTED DEPOSITIONS

A CVD method is categorized not only by the pressure regime but also by its energy input. As mentioned in Section 5.2.2, PECVD uses a radio frequency (rf) power

Micron Ex. 1023, p. 267
Micron v. YMTC
IPR2024-00909

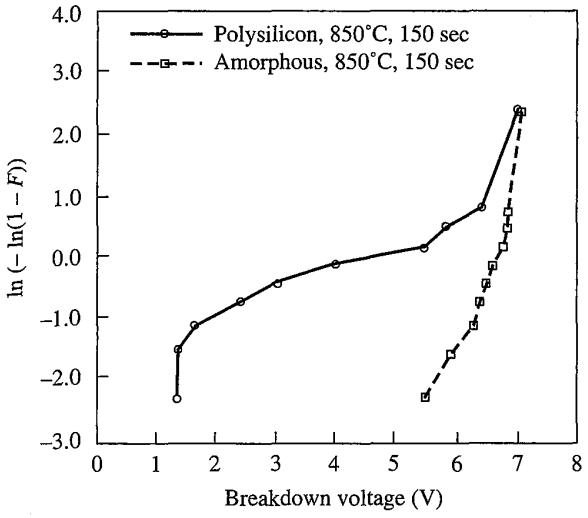


FIGURE 28
Weibull plots of the TZDB characteristics for the polyoxides grown on both recrystallized amorphous-Si and poly-Si samples. (*After Wang et al., Ref. 48.*)

to generate glow discharge to transfer the energy into the reactant gases, allowing the deposition on the substrate at a lower temperature than in the APCVD or the LPCVD process.[9,49–55] Hence, low deposition temperatures are a major advantage of PECVD. Desirable properties such as good adhesion, low pinhole density, good step coverage, adequate electrical properties, and compatibility with fine-line pattern transfer processes have made these PECVD films useful in ULSI circuits. Owing to good protection properties of scratch resistance, sodium diffusion, and moisture protection, plasma-deposited silicon nitride is commonly used as the final passivation layers of devices. Plasma-deposited nitride and oxide can be used as insulators between the metal layers, especially since the underlying metals have low melting points.[56] PECVD also provides deposition of amorphous silicon films, although these films have not yet been used in the ULSI circuits. The PECVD amorphous silicon has been widely applied in TFT LCDs.

5.6.1 Deposition Methods

The plasma is generated by the application of an rf field to a low-pressure gas, thereby creating free electrons within the discharge region. The electrons gain energy from the electric field. When they collide with gas molecules, the reactant gases (e.g., silane and nitrogen- or oxygen-containing species) dissociate and ionize. The energetic species are then adsorbed on the film surface. These radicals tend to have high sticking coefficients and also appear to migrate easily along the surface after adsorption. These two factors can lead to excellent film conformality. Upon being adsorbed on the substrate, the radicals are subjected to ion and electron bombardment and rearrangements and react with other adsorbed species to form new bonds and the resulting film. Atom rearrangement includes diffusion of the adsorbed atoms onto stable sites and concurrent desorption of reaction product. The desorption rates depend on the substrate temperatures; higher temperatures produce films with fewer entrapped

Micron Ex. 1023, p. 268
Micron v. YMTC
IPR2024-00909

by-products. Note that homogeneous gas-phase nucleation should be avoided to reduce particulate contamination.

Three types of PECVD reactors are available:

1. The parallel plate type
2. The horizontal tube type
3. The single wafer type

As compared with rf sputtering systems, the key difference in designing the reactor and electrode configuration in PECVD systems is that the potentials of both the powered and the grounded electrodes, relative to the plasma, must be approximately equal in PECVD systems. The walls of parallel plate reactors are made of either quartz-, ceramic-, or aluminum oxide–coated steel in order to place them at a floating potential with respect to the plasma. This can minimize bombardment and sputtering of the wall and reduce contamination on the growing films.

Parallel-plate PECVD reactors

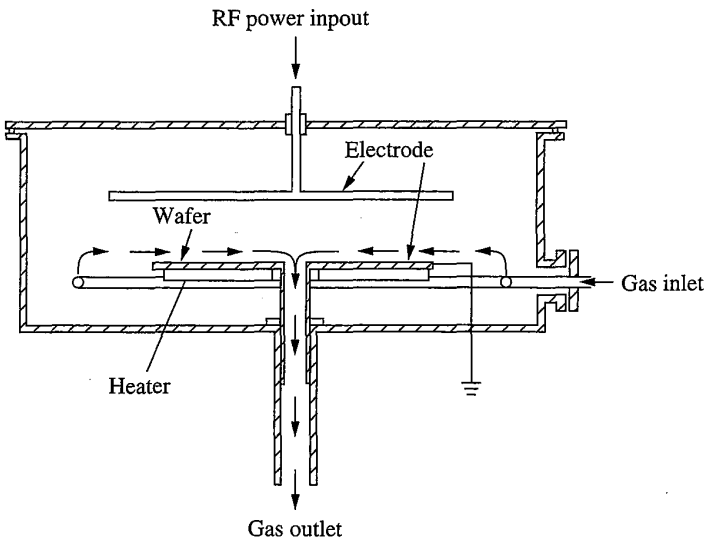
The radial parallel plate reactor was developed by Reinberg in the early 1970s, and the first commercial model was offered in 1976. A schematic diagram of this reactor type is shown in Fig. 29*a*. The reaction chamber is a short, vertically oriented cylinder, typically constructed of aluminum-coated stainless steel. The rf power (at frequencies of 380 kHz to 13.56 MHz), which establishes the plasma, is applied to the upper electrode, and the wafers reside on the bottom, grounded electrode, which can be rotated, for improved uniformity, at substrate temperatures up to 400°C. The electrode spacing is typically 5 to 10 cm, and such systems operate in the pressure range of 0.1 to 5 torr. The parallel plate system, however, suffers from low throughput for large-diameter wafers. In addition, particulates flaking off from the walls or the upper electrode can fall onto the horizontally positioned wafers.

Horizontal-tube PECVD reactors

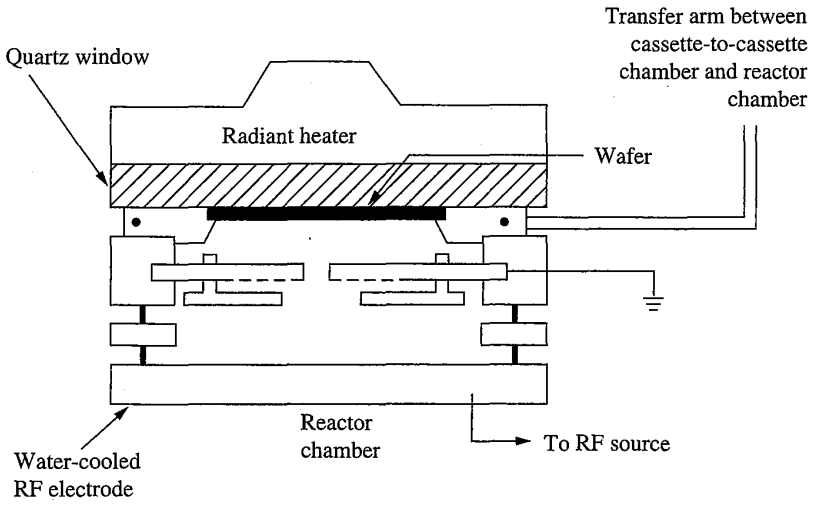
A new PECVD reactor design, the horizontal tube reactor, is shown in Fig. 29*a*. This design allows the throughput of PECVD reactors to be greatly increased. The reactor resembles a "hot-wall" LPCVD system. It consists of a long horizontal quartz tube that is radiantly heated. Gas is fed into one end and flows linearly to the other. Special long rectangular graphite plates serve both as the electrodes needed to establish a plasma and as the holders of the wafers. The electrode configuration is also designed to provide a uniform plasma environment for each wafer to ensure film uniformity. These vertically oriented graphite electrodes are stacked parallel to one another, side by side, with alternating plates serving as power and ground for the rf voltage. A plasma is formed in the space between each pair of plates.

The use of several long slabs allows an increased number of wafers (up to 120 pieces of 100 mm in diameter) to be loaded into the reactor at one time. Since the wafers are held vertically, most particles do not fall on the wafer surfaces. The entire graphite assembly is withdrawn from the reactor and then reinserted for processing. Care must be taken, however, to prevent particulate contamination when the wafers are loaded and unloaded.

Micron Ex. 1023, p. 269
Micron v. YMTC
IPR2024-00909



(*a*)



(*b*)

FIGURE 29
Schematic diagrams of plasma deposition reactors: (*a*) parallel plate type, and (*b*) single-wafer type.

Single-wafer PECVD reactors

The most recently introduced PECVD reactor is the single-wafer design shown in Fig.29*b*. The reactor, which is load-locked, offers cassette-to-cassette operation and provides rapid radiant heating of each wafer as well as allowing in-situ monitoring of the film deposition. Wafers larger than 200 mm in diameter can be adopted. Such reactors are being designed to allow the fabrication of large wafers with low

Micron Ex. 1023, p. 270
Micron v. YMTC
IPR2024-00909

particulate contamination. Multichamber PECVD, with or without a reactive ion etching (RIE) system, has also been developed and used to deposit the oxide, nitride, and silicon film free of air-exposure contamination.[57]

5.6.2 Deposition Variables

Plasma-assisted CVD requires the control and optimization of rf power density, frequency, and duty cycle in addition to the conditions similar to those of an LPCVD process, which possesses the parameters of gas composition, flow rate, deposition temperature, and total pressure. Like the LPCVD process at low temperatures, the PECVD process is surface-reaction-limited, and adequate substrate temperature control is thus necessary to ensure film thickness uniformity.

By reacting silane and oxygen or nitrous oxide in plasma, silicon dioxide films can be deposited. The reactions are expressed as

$$SiH_4(g) + O_2(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{5.17}$$

$$SiH_4(g) + 4N_2O(g) \rightarrow SiO_2(s) + 4N_2(g) + 2H_2O(g) \tag{5.18}$$

Furthermore, silicon nitride films are deposited by reacting silane and ammonia or nitrogen in plasma, with the reactions

$$SiH_4(g) + NH_3(g) \rightarrow SiN : H(s) + 3H_2(g) \tag{5.19}$$

$$SiH_4(g) + N_2(g) \rightarrow 2SiN : H(s) + 3H_2(g) \tag{5.20}$$

Amorphous silicon is prepared mainly by the glow discharge decomposition of silane, with the reaction expressed as

$$SiH_4(g) \rightarrow Si(s) + 2H_2(g) \tag{5.21}$$

Processing parameters such as deposition rate will be affected by such deposition variables as total pressure, reactant partial pressure, discharge frequency and power, electrode materials, gas species, reactor geometry, pumping speed, electrode spacing, and deposition temperature. The higher the deposition temperature and the rf power, the higher the deposition rate is. Furthermore, the properties of all these PECVD films will be influenced by these deposition conditions. These properties will be discussed in the following section.

5.6.3 Properties of Plasma-Deposited Films

The fact that the radicals formed in the plasma discharge are highly reactive causes some stoichiometric problems, because the reactions are so varied and complicated. Moreover, the by-products and incidental species are incorporated into the reactant films, especially hydrogen, nitrogen, and oxygen. Excessive incorporation of these contaminants may lead to threshold voltage shifts in MOS circuits that use these PECVD films.

The refractive indices of plasma-deposited silicon nitride are distributed in a large range (1.8 to 2.6). The resistivity and breakdown field of the PECVD silicon

Micron Ex. 1023, p. 271
Micron v. YMTC
IPR2024-00909

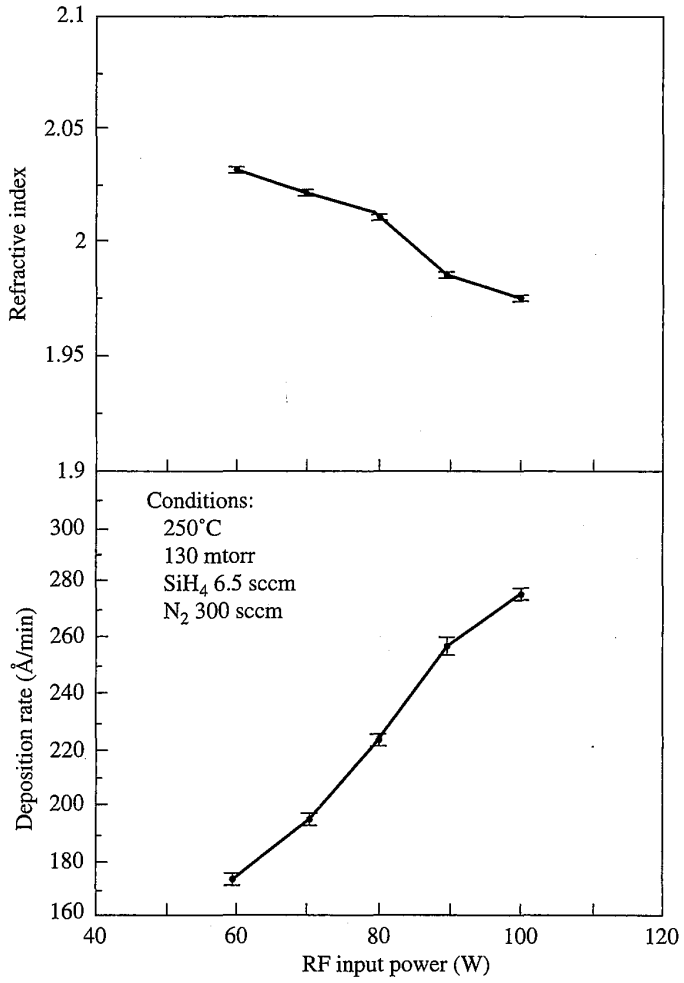


FIGURE 30
Refractive index and deposition rate as functions of rf power for plasma nitride films.

nitride also vary with different deposition conditions (10^5 to 10^{21} ohm-cm and 3 to 9 MV/cm, respectively). Figure 30 shows the refractive index and deposition rate as functions of rf power for plasma nitride films. For silicon-rich nitride films, which lower rf power tends to form, refractive indices will be high and resistivities will be low. Low rf powers also exhibit lower deposition rates. In addition, the dielectric breakdown field, dielectric constant, and band gap of plasma nitrides also vary with the compositions.

Plasma silicon nitrides often contain high hydrogen concentrations of between 10 and 35 at% hydrogen. The relative concentrations of Si–H and N–H vary widely and cause a large variation in film properties. Because high stress will cause cracking of the nitride film, minimizing stress in PECVD silicon nitride is most important. Crack-free films with low tensile stress can be deposited by precisely controlling the process parameters. Stress depends on the local bonding around individual atoms and is affected by deposition variables such as rf frequency and deposition temper-

Micron Ex. 1023, p. 272
Micron v. YMTC
IPR2024-00909

TABLE 5
Properties of silicon nitride films

	Deposition	
	Plasma	**LPCVD**
Temperature(°C)	250–350	700–800
Composition	SiN_xH_y	$Si_3N_4(H)$
Stress (10^9 dyne/cm)	2C-5T	10T
Dielectric constant	6–9	6–7
Dielectric strength (10^6 V/cm)	5	10
Resistivity (Ω-cm)	10^6–10^{15}	10^{16}
Refractive index	1.8–2.5	2.01
Density (g/cm^3)	2.4–2.8	2.9–3.1
Energy gap (eV)	4–5	5
Si/N ratio	0.8–1.2	0.75
Atom % H	20–25	4–8

ature. Low rf frequency facilitates the plasma nitride films with compressive stress. The general properties of plasma-deposited silicon nitride are listed in Table 5. Since the film composition is a variable, the properties of plasma nitride vary over large ranges.

In contrast, the properties of plasma oxide are nearly constant, because the oxide composition is nearly stoichiometric. The refractive index and deposition rate of plasma silicon oxide as functions of deposition temperature are shown in Fig. 31. High deposition temperatures attain higher deposition rates. However, the refractive indices of these plasma oxide films are almost a constant, of about 1.45. Table 6 lists the properties of PECVD silicon oxide films. Note that the plasma oxides become conformal if the reaction of SiH_4 with O_2 or N_2O is replaced by that of TEOS with O_2 or N_2O. This feature makes PECVD TEOS very useful in ULSI technology. Combining the isotropic flow characteristics of the APCVD or slightly-reduced-pressure CVD O_3–TEOS films with the PECVD TEOS provides the planarization scheme for ULSI technology, especially for refilling the aluminum interconnects, which require low-temperature processing. Here, PECVD TEOS films can provide moisture protection for the APCVD O_3–TEOS in addition to the conformality feature. The processing conditions of the PECVD TEOS are a TEOS flow rate of 7 sccm, an O_3 to TEOS ratio of 13 : 1, a total pressure of 0.6 torr, an rf power of 0.5 W/cm, and substrate temperature of 350°C. The resulting deposition rate of PECVD TEOS is 250 Å/min.[58]

Some new dielectric films prepared by PECVD have been applied in ULSI memory devices (e.g., MOS capacitors of DRAMs), including Ta_2O_5,[59] TiO_2, $LiNbO_3$, and PLZT (lead lanthanum zirconate titanate, (Pb, La)(Zr, Ti)O_3).[57] Note that the use of Ta_2O_5 films has been announced by some microelectronics companies to increase the storage capacitance in 256-Mbit DRAM devices by a factor of about 4.[60] Conventional MOS capacitors are fabricated with the stacked oxide/nitride/oxide (ONO) structure prepared by LPCVD. In the ONO structure the tensile stress in nitride film is offset by the compressive stress of the oxide. However, all of the *oxynitrides,* or stacked oxide/nitride/oxide films, will not provide enough

Micron Ex. 1023, p. 273
Micron v. YMTC
IPR2024-00909

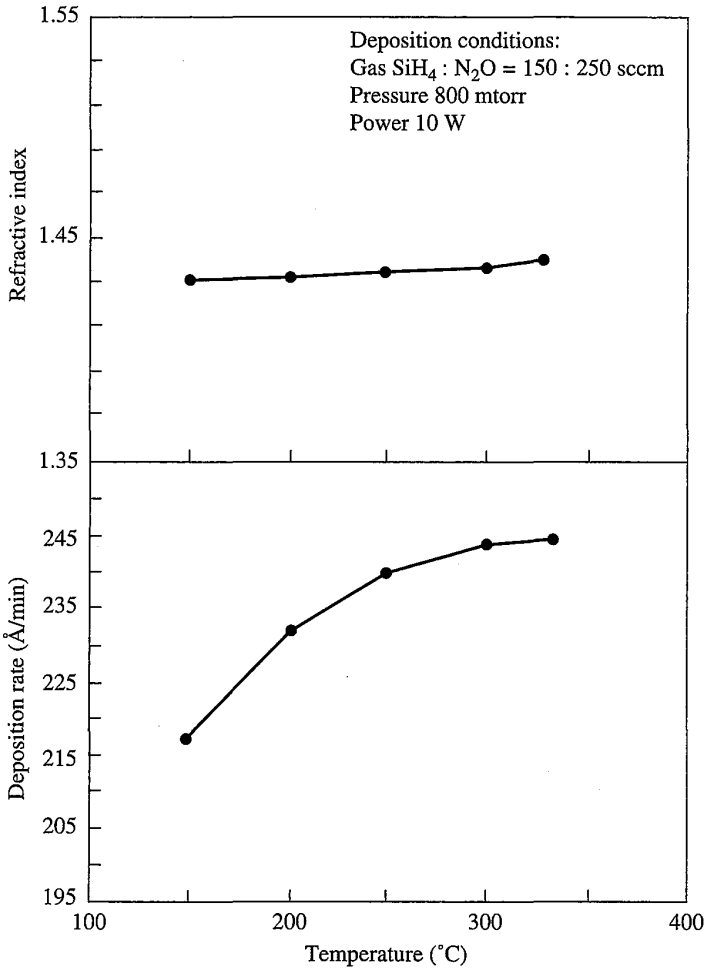


FIGURE 31
Refractive index and deposition rate of plasma silicon oxide as functions of deposition temperature.

capacitances for future ULSI memory devices. Hence, Ta_2O_5 films are fabricated in CVD from solid sources such as $TaCl_5$, or related halogens and liquid materials, such as $Ta(OCH_3)_5$, $Ta(OC_2H_5)_5$, and related alkoxides. In addition to PECVD, Ta_2O_5 can be prepared by LPCVD, metal-organic CVD (MOCVD), and photon-induced CVD. Photon-induced CVD and MOCVD will be described in Section 5.7.

5.7 OTHER DEPOSITION METHODS

5.7.1 Photon-Induced Chemical Vapor Deposition

Photon-induced chemical vapor deposition (PHCVD, photo CVD) is a technique that may fulfill the need for an extremely-low-temperature deposition process without the

Micron Ex. 1023, p. 274
Micron v. YMTC
IPR2024-00909

TABLE 6
Properties of silicon oxide films

Deposition	**Plasma** $SiH_4 + O_2$(or N_2O)	**Plasma** TEOS + O_2	**APCVD** TEOS + O_3	**LPCVD** $SiH_4 + O_2$	**LPCVD** TEOS + O_2	**LPCVD** $SiCl_2H_2 + N_2O$	**Thermal**
Temperature(°C)	250	400	400	450	700	900	1000
Step coverage	Nonconformal	Conformal	Isotropic Flow	Nonconformal	Conformal	Conformal	Conformal
Composition	$SiO_{1.9}$(H)	SiO_x	SiO_2(−OH)	SiO_2(H)	SiO_2(−OH)	SiO_2(−Cl)	SiO_2
Stress (10^9 dyne/cm^2)	3C-3T	C-T	T	3T	1C	3C	3C
Dielectric constant	4.9	—	—	4.3	4.0	—	3.9
Dielectric strength (10^6 V/cm)	3–6	6	3–6	8	10	10	11
Refractive index	1.45	1.45	—	1.44	1.46	1.46	1.46
Density (g/cm^3)	2.3	—	1–2	2.1	2.2	2.2	2.2
Etch rate, nm/min (100 : 1 H_2O : HF)	40	164 20 : 1 H_2O : HF	—	6	3	3	2.5
Comments	High density Cusp problem and microparticle contamination	High density Limited surface coverage (combinations with SOG or O_3-TEOS)	Excellent surface coverage Self-planarization Low density High moisture affinity Stress-induced crack over 2mm thick	High density Low temperature	High density High temperature High moisture affinity Backing-induced crack	High density High temperature No doping capability	Excellent quality High temperature

Micron Ex. 1023, p. 275
Micron v. YMTC
IPR2024-00909

problems of PECVD.[61] PHCVD uses high-energy, high-intensity photons to heat the substrate surface or to dissociate and excite reactant species in the gas phase. In the case of substrate surface heating the reactant gases are transparent to the photons, and the energy for gas-phase reactions is completely eliminated. On the other hand, in the case of reactant gas excitation the energy of photons can be chosen for energy-efficient transfer either to the reactant molecules themselves or to a catalytic intermediary such as mercury vapor. This technique enables deposition at extremely low substrate temperatures (e.g., room temperature). Unlike PECVD, PHCVD possesses the advantage of deposited films with a low defect density. Photo CVD has been used to deposit SiN_x film with the underlying substrate free of ion-bombardment damage that can be induced by PECVD. In addition, PHCVD films also show good step coverage. However, they may suffer from low density and molecular contamination as a result of the low deposition temperature.

There are three classes of PHCVD reactors, depending on the energy source: (1) UV lamp, (2) laser,[62] and (3) IR lamp. UV lamp reactors generally use mercury vapor for the energy transfer between photons and the reactant gases. UV radiation at the wavelength of 253.7 nm is efficiently absorbed by mercury atoms, which then transfer energy to the reactant species. Deposition rate for UV PHCVD reactors are typically much slower than those for other low-temperature techniques. The process data and film quality of Si_3N_4, amorphous silicon, and SiO_2 are listed in Table 7. The reactor chamber and the PHCVD system are schematically shown in Figs. 32*a* and *b*. The chemical reactions of PHCVD of a silicon film are expressed as

$$Hg(g) + h\nu\,(253.7\text{ nm}) \rightarrow Hg^*(g) \tag{5.22}$$

$$Hg^*(g) + SiH_4(g) \rightarrow Hg(g) + SiH_4^*(g) \tag{5.23}$$

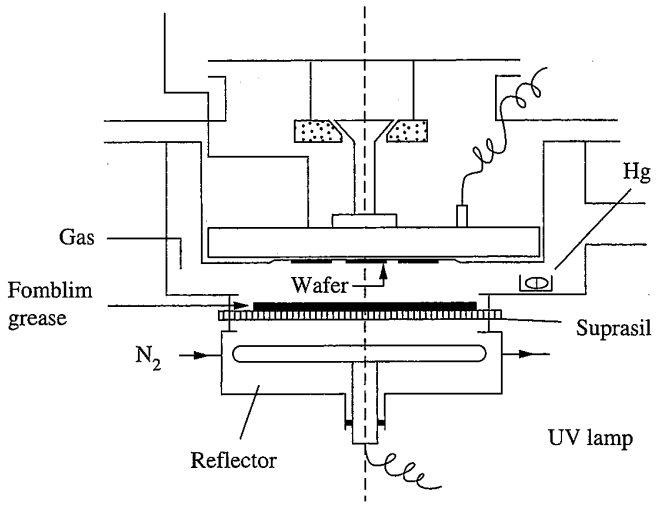
$$SiH_4^*(g) \rightarrow Si(s) + 2H_2(g), \tag{5.24}$$

where Hg^* is the excitation state of Hg and $h\nu$ (253.7 nm) indicates the UV light. As for the deposition of the oxide film, its chemical reactions are written as

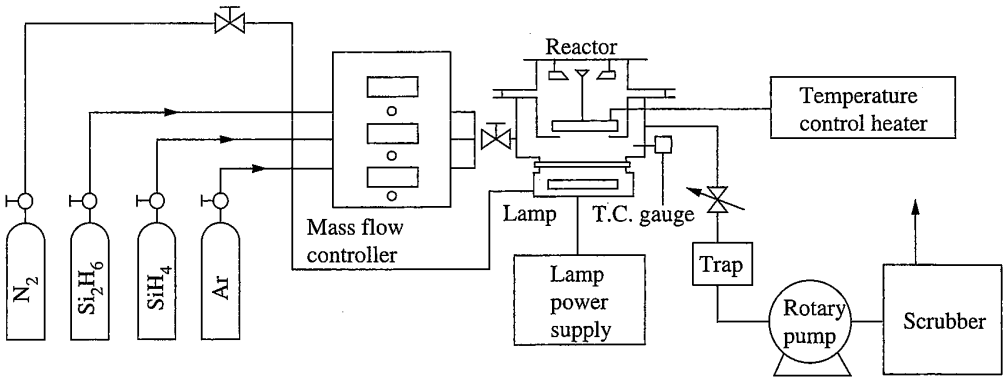
TABLE 7
Properties of films deposited by PHCVD

	Film type		
	Si_3N_4	α-Si	SiO_2
Thickness uniformity (%)	±5	±5	±5
Deposition rate (nm/min)	7–10	6	12
Deposition temperature (°C)	50–250	50–250	50–200
Deposition pressure (torr)	0.3–1.5	0.2–1.0	0.3–1.0
Dielectric constant	5.5	—	4.0
Refractive index	1.8–2.0	3.3–3.4	1.45–1.50
Density (g/cm^3)	1.8–2.5	2.2–2.3	2.1
Hydrogen content (at %)	10–15	8–20	—
Etch rate (nm/min) 20 : 1 H_2O : HF	5–10	—	10–15

Micron Ex. 1023, p. 276
Micron v. YMTC
IPR2024-00909



(a)



(b)

FIGURE 32
Schematic configuration of (*a*) the photo CVD reactor chamber, (*b*) the photo CVD system.

$$Hg(g) + h\nu\,(253.7\text{ nm}) \rightarrow Hg^*(g) \tag{5.25}$$

$$Hg^*(g) + N_2O(g) \rightarrow Hg(g) + N_2(g) + O(g) \tag{5.26}$$

$$SiH_4(g) + 2O(g) \rightarrow SiO_2(s) + 2H_2(g) \tag{5.27}$$

Similarly, Si_3N_4 films can be deposited by reacting SiH_4 with NH_3. Figure 33 shows the deposition rate of silicon dioxide deposited by PHCVD as a function of substrate temperature. Higher substrate temperatures will facilitate higher deposition rates, and the activation energy of films deposited by PHCVD is lower than that in the thermal CVD process.

Micron Ex. 1023, p. 277
Micron v. YMTC
IPR2024-00909

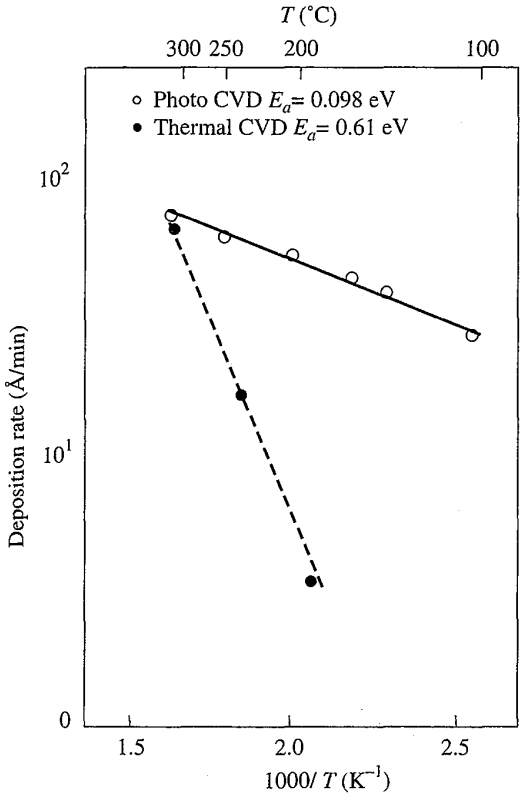


FIGURE 33
The deposition rates of silicon dioxide by thermal CVD (dashed line) and photo CVD (solid line) as a function of substrate temperature. (*After Peters et al., Ref. 61.*)

Laser PHCVD reactors offer the advantages of frequency tunability and a high-intensity light source. Tunability is useful to target the specific photon energies for particular dissociation reactions, enabling more flexible control of the depositions. The energy from the high-intensity laser also increases the reaction rate. Laser PHCVD opens the possibility of CVD writing, i.e., selective deposition by spatial control of the laser. Nevertheless, the deposition rates of the current PHCVD processes are still too low to allow them to be adopted for production applications.

On the other hand, an IR lamp has often been used in deposition of silicon dioxide, silicon film, and silicon nitride. Due to the rapid heating rate of the IR lamp, the reactors with an IR lamp heating source are also called rapid thermal processing (RTP) systems. If the reactions for formation of SiO_2, Si_3N_4, and polysilicon by chemical vapor deposition (CVD) are performed using the IR lamp, this technology is named as RTPCVD. The resulting film properties are all similar to those for the LPCVD or APCVD processes. However, the short processing duration can reduce autodoping of CVD epitaxial silicon or polysilicon films.

5.7.2 Electron Cyclotron Resonance Chemical Vapor Deposition

The microwave electron cyclotron resonance (ECR) CVD method is expected to realize low-temperature, low-ion-damage, and high-deposition-rate processes.[63–65] Silicon nitride films formed by ECR CVD without intentional heating of substrates can show good electrical insulating and protective properties as passivation films.

Micron Ex. 1023, p. 278
Micron v. YMTC
IPR2024-00909

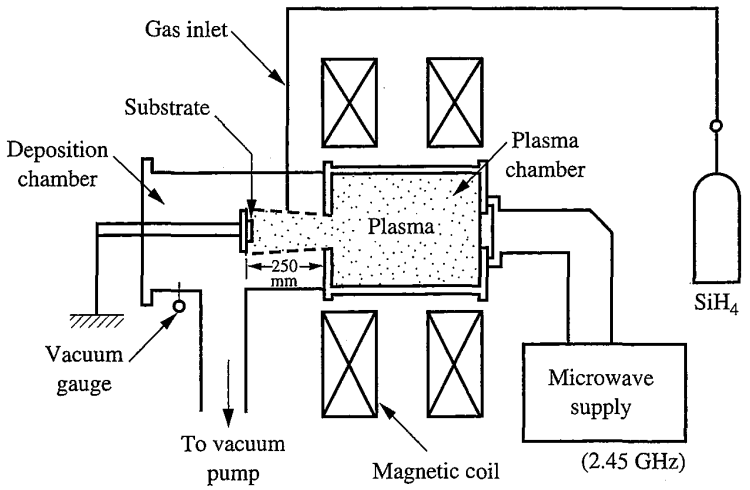


FIGURE 34
Schematic diagram of the microwave ECR CVD system.

Figure 34 illustrates a schematic configuration of the microwave ECR plasma CVD system. Microwave power is supplied to a plasma chamber through a rectangular waveguide. The magnetic field for the ECR plasma excitation is applied perpendicularly to the substrate surface. The plasma stream is introduced into the deposition chamber in the direction of the substrate by the magnetic field. The distance from the substrate to the plasma chamber is 250 mm. The gases are introduced into the deposition chamber. For ECR CVD the deposition variables include the microwave power, total pressure, gas species, and flow rate. These conditions are very similar to those for PECVD, except that the rf power is replaced by the microwave power. However, the ECR source generally creates a large number of electrons to generate a high density of plasma. Hence, the total pressure for the ECR CVD is much less than that for PECVD. The required power is also much lower than that for PECVD. Consequently, the induced plasma possess a lower ion energy, and the resulting film will achieve a higher quality due to low defect density. Typical deposition conditions are listed in Table 8. Figure 35 shows the effect of silane gas flow and the total pressure on the deposition rate of silicon oxide. Higher total pressures and flow rates result in higher deposition rates. The chemical reactions of SiO_2,

TABLE 8
ECR deposition conditions

Back pressure (torr)	2×10^{-6}
Gas flow rate (sccm)	4–30
Pressure (torr)	1×10^{-4} - 2×10^{-3}
Microwave frequency (GHz)	2.45
Microwave power (W)	200
Magnetic field (gauss)	875
Substratte temperature	Without intentional heating

Micron Ex. 1023, p. 279
Micron v. YMTC
IPR2024-00909

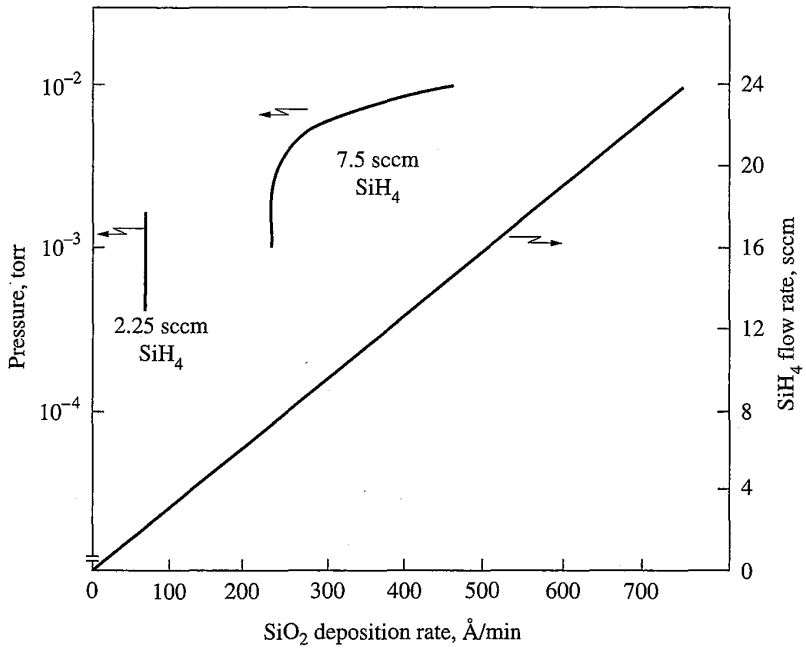


FIGURE 35
The effect of silane gas flow and the total system pressure on the deposition rate of ECR CVD silicon oxide film. (*After Popov and Waldron, Ref. 65.*)

Si_3N_4, and amorphous silicon (α-Si) films in ECR CVD are all the same as those in PECVD.

5.7.3 Ultrahigh-Vacuum Chemical Vapor Deposition

The ultrahigh-vacuum chemical vapor deposition (UHVCVD) apparatus is shown schematically in Fig. 36.[66] Prior to its initial use, the entire UHV section of the apparatus was subjected to a thorough bakeout as well as hydrogen plasma scouring until an ultimate vacuum in the 10^{-9} torr range was achieved. The UHVCVD system features a load-lock and growth chamber with a base pressure maintained at about 10^{-8} torr. Such an ultraclean environment ensures an undoped film of minimum impurity incorporation. In-situ mass spectrometry shows hydrogen to be in the greatest abundance, with water vapor and oxygen present at less than 10^{-10} torr partial pressure. For as-grown films, the oxygen and carbon concentrations were below the detection limit of secondary ion mass spectroscopy, which are about 10^{17} and 10^{18} cm^{-3} for O and C, respectively.[67] The low pressure also prevents the formation of a porous Si structure,[68] which will dramatically degrade the device performance.

Recently, many researchers have reported the growth of poly-Si and poly-$Si_{1-x}Ge_x$ at reduced pressure (e.g., below 10 mtorr) in UHVCVD systems. It was found that lower pressures led to larger grain size, lower transition temperature, or both. Moreover, the use of poly-$Si_{1-x}Ge_x$ instead of poly-Si would lower the process

Micron Ex. 1023, p. 280
Micron v. YMTC
IPR2024-00909

FIGURE 36
UHVCVD system schematic. (*After Meyerson, Ref. 66.*)

Micron Ex. 1023, p. 281
Micron v. YMTC
IPR2024-00909

temperature as well as the process thermal budget. At a reduced-pressure condition, the gas flow is in a molecular mode, and the surface reaction plays a dominant role during growth. Liehr et al.[69] have shown by the desorption rate of surface-bonded hydrogen that the growth of poly-Si films was dominant. This is very similar to the growth mechanism observed for low-temperature (below 600°C) Si epitaxy from a hydride source (e.g., SiH_4).[70] Figure 37 shows the deposition rates of poly-Si films at four different growth temperatures, namely 600, 575, 550, and 525°C, respectively. The SiH_4 flow rate was kept at 20 sccm. It seems that the deposition rates are essentially the same for epitaxial and poly growth at temperatures below 600°C. These results indicate an identical deposition mode between these two cases. Figure 38 shows that the effect of GeH_4 flow rate on the deposition rate for poly films was smaller than epi-layer growth at a higher GeH_4 flow. This indicates that the addition of GeH_4 would change the deposition mode remarkably.[71] Poly-$Si_{1-x}Ge_x$ is potentially more advantageous than poly-Si as a gate electrode for the metal-oxide-semiconductor field-effect transistor (MOSFET).[72] With the former, the threshold voltage can be adjusted simply by varying the Ge content inside the film, which is much easier than the conventional channel implantation method.

The UHVCVD system could deposit in-situ-doped poly-Si films by using SiH_4 and 1% B_2H_6 in H_2. Figure 39 shows the resistivity as a function of the B_2H_6 flow rate for poly-Si films; a very low resistivity (about 4 mΩ-cm) was obtained.

Hence, the use of poly films deposited by a UHVCVD system can significantly lower the deposition temperature and the process thermal budget, which is required for deep submicron Si technology and some other applications demanding a low

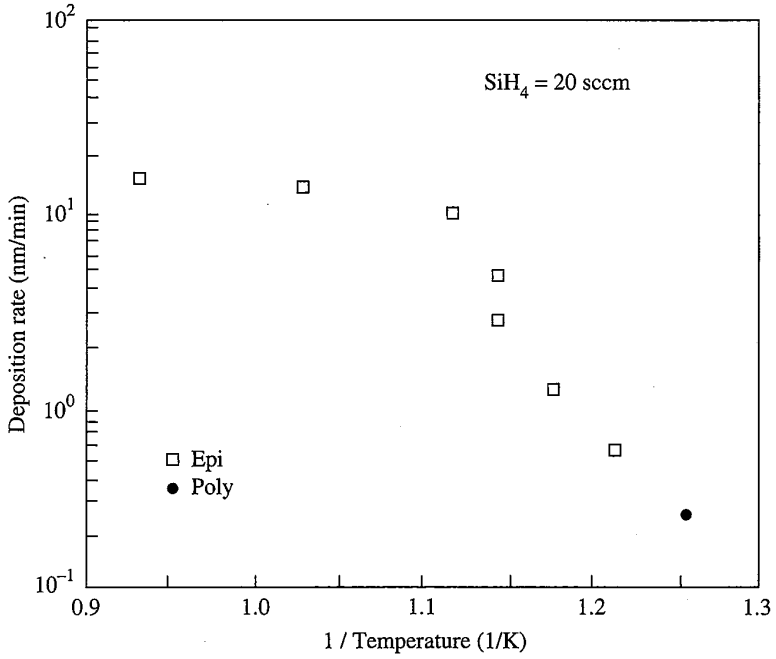


FIGURE 37
An Arrhenius plot of deposition rate for the poly-Si and epitaxial Si films. (*After Lin, Chang, and Lin, Ref. 71.*)

Micron Ex. 1023, p. 282
Micron v. YMTC
IPR2024-00909

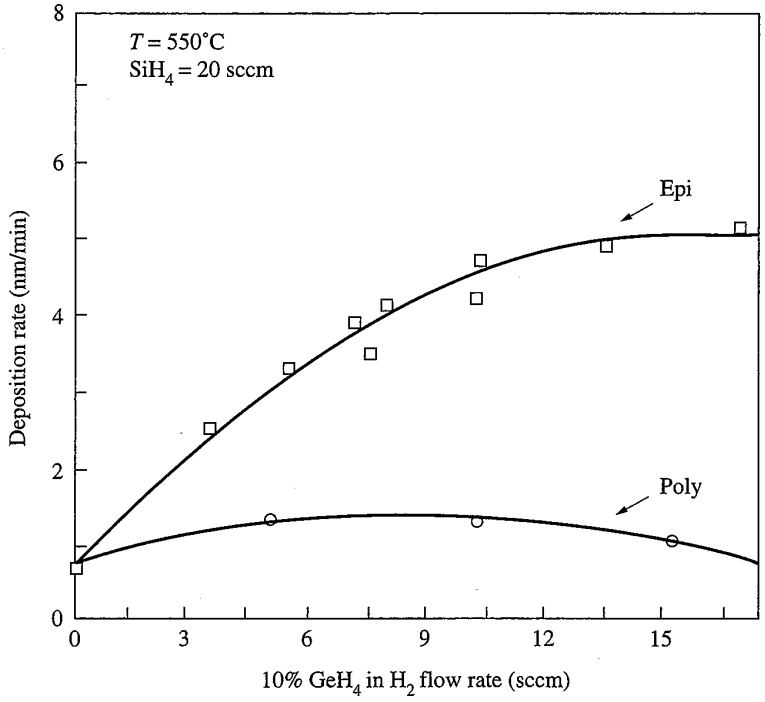


FIGURE 38
Deposition rate for poly and epitaxial films as a function of GeH_4 flow. (*After Lin, Chang, and Lin, Ref. 71.*)

processing temperature, such as thin-film transistor (TFT) fabrication for active-matrix liquid crystal displays (AMLCDs)[73,74] and photoconductors.[75]

5.7.4 Other CVD Techniques

Other CVD techniques, such as metal-organic CVD (MOCVD), laser ablation, and hybrid-excitation CVD, are used to deposit dielectric and silicon films.[76] Like the ferroelectric dielectrics, $LiNbO_3$ and PLZT, they can be prepared by MOCVD or laser ablation. MOCVD uses a metal-organic compound vapor that is decomposed or reacts with other reactants to form the desired film. The excitation is generally thermal but can also be plasma-assisted. Like the PECVD of Ta_2O_5 mentioned in Section 5.6.3, MOCVD of Ta_2O_5 adopts the liquid metal-organic compounds $Ta(OCH_3)_5$ and $Ta(OC_2H_5)_5$. However, MOCVD uses thermal energy to dissociate these organic compounds, rather than the rf power used in PECVD. The chemical reactions can be expressed as

$$Ta(OC_2H_5)_5(l) \rightarrow Ta(s) + OC_2H_5(g) \quad (5.28)$$

$$2O_2(g) \rightarrow 2O(g) + O_2(g) \quad (5.29)$$

$$Ta(s) + O(g) \rightarrow Ta_2O_5(s) \quad (5.30)$$

$$OC_2H_5(g) + O_2(g) \rightarrow CO_2(g), CO(g), H_2O(g), C_2H_4(g), CH_4(g) \quad (5.31)$$

Micron Ex. 1023, p. 283
Micron v. YMTC
IPR2024-00909

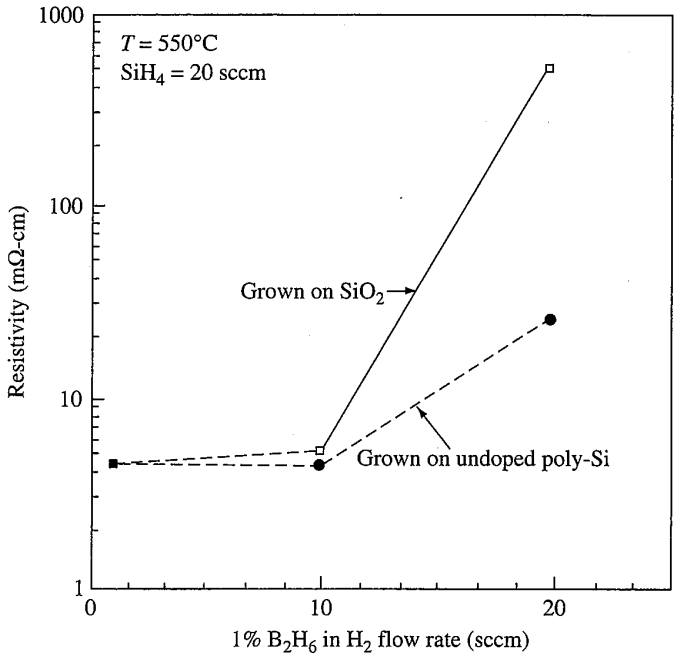


FIGURE 39
Effect of B_2H_6 (1% B_2H_6 in hydrogen) flow rate on the resistivity of poly-Si films grown on either SiO_2 or undoped poly-Si substrate. (*After Lin et al., Ref. 73.*)

Laser ablation deposits the dielectric films by directly scanning the laser beam on the dielectric target. The advantage is the low-temperature process, but the disadvantage is the uniformity control and lack of large-wafer capability.

The hybrid-excitation idea has been used in depositing some dielectric films. The hybrid-excitation techniques consist of plasma-assisted direct photolysis and photo-enhanced CVD methods. The environment contamination caused by the sensitizer encountered in photo CVD will be prevented. A schematic diagram of the hybrid-excitation CVD apparatus is shown in Fig. 40. The plasma of reactive gas is generated in a glow discharge tube and then diffuses into a process chamber. Therefore, the samples will also be free from plasma damages.

5.8 APPLICATIONS OF DEPOSITED POLYSILICON, SILICON OXIDE, AND SILICON NITRIDE FILMS

Polycrystalline silicon, also called polysilicon or poly-Si, thin films have many important applications in integrated circuit (IC) technology. The most widespread application of poly-Si is in metal-oxide-semiconductor (MOS) ICs. Heavily doped poly-Si films have been widely used as the gate electrodes and interconnections in MOS circuits. Poly-Si is utilized because of its compatibility with subsequent high-temperature processing, its excellent interface with thermal oxide (low interface

Micron Ex. 1023, p. 284
Micron v. YMTC
IPR2024-00909

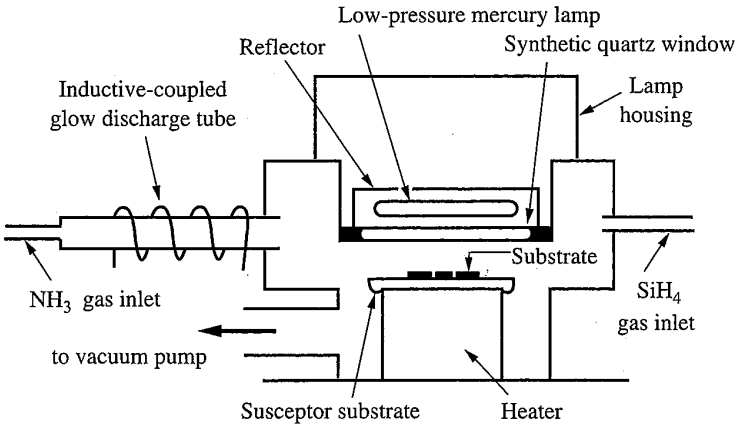


FIGURE 40
Schematic configuration of the hybrid-excitation CVD system.

state density), its higher stability than Al gate material, its ability to be deposited conformally over steep topography, and its capability of reacting with the overlaying metal to form *silicide* structures. Silicide films formed onto the poly-Si, known as *polycides,* on the gate dielectrics have been fabricated to lower the sheet resistance of poly-Si gates. This is especially important for ULSI devices. Moreover, self-aligned silicides (*salicides*) have also been used in ULSI circuits.[77] Through the reaction of metal with the underlying polysilicon and silicon substrate, the self-aligned silicide contacts on the source, drain, and gate will be simultaneously formed, avoiding any misalignment due to photolithography. Emitter structures with heavily doped poly-Si films are also utilized in advanced bipolar devices. Selective growth of poly-Si is studied as a planarization technique for filling the contact hole. Lightly doped poly-Si films are used as high-resistance resistors in static random access memories (SRAMs), and oxygen-doped poly-Si (semi-insulating poly-Si or SIPOS) can be used to form a barrier to the carrier injection. In addition, poly-Si films have been extensively used in ULSI DRAMs, SRAMs, and flash memory devices.[78] For 1-Mbit (and even higher packing density) DRAMs, high capacitances of the MOS capacitors are still needed. However, the decreasing area of the ULSI devices will require a three-dimensional capacitor structure. A layer of *rugged* polysilicon (hemispherical grains of polysilicon) is developed to increase the surface area. The hemispherical grains of polysilicon have been grown by selection of the deposition temperature and the subsequent annealing of the LPCVD system.[79]

Silicon dioxide has several uses: to serve as a mask against implant or diffusion of dopants into silicon, as a gate oxide and capacitor dielectric in MOS devices, and as a tunneling oxide in (EEPROMs). It can also be used to provide surface passivation, to isolate one device from another, to act as a component in MOS structure, and to provide electrical isolation of multilevel metallization systems. Several techniques for forming the oxide layers have been developed, such as wet or dry thermal oxidation (including rapid thermal techniques), wet anodization, chemical vapor deposition (CVD), PECVD, PHCVD, ECR CVD, and plasma anodization or oxidation. When the interface between the oxide and the silicon requires a low-trap

Micron Ex. 1023, p. 285
Micron v. YMTC
IPR2024-00909

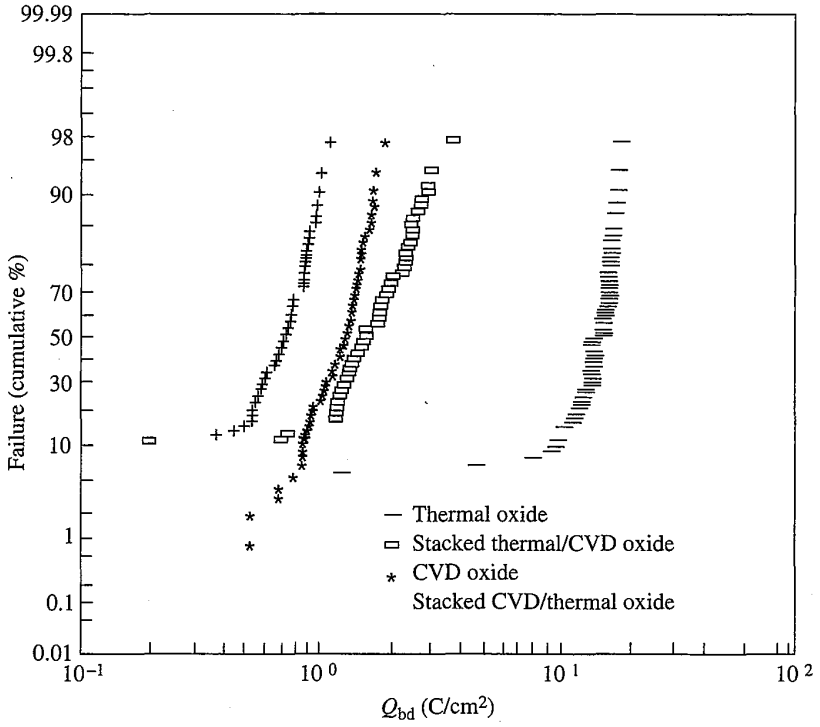


FIGURE 41
The time-dependent dielectric-breakdown charge Q_{bd} in coulomb/cm^2 by step current stress: thermal oxide (−); CVD oxide (*); stacked CVD/thermal oxide (+); stacked thermal/CVD oxide (o). (*After Tseng et al., Ref. 80.*)

density, thermal oxidation has been the preferred technique. In the past, CVD oxides have seldom been used in ULSI technology as the gate dielectrics, because an extrusion phenomenon of the CVD oxides forms at the n^+ polysilicon/CVD oxide interface. The resulting electrical properties of CVD and thermal oxides are compared in Fig. 41.[80] However, CVD oxides are particularly useful as the gate dielectrics on the polysilicon active channel. For example, CVD oxides can offer better characteristics than thermal polyoxides for the poly-Si TFTs in 4-Mbit and higher-density SRAMs because polyoxides generally have the problem of interface roughness. Owing to the improvement of the characteristics by the deposition schemes, CVD oxides will become more important in future ULSI technology, which requires multilayered structures. CVD oxides with low and high dielectric constants will be used in interlayer isolations and MOS capacitors, respectively.[81]

Silicon nitride films are amorphous insulating materials and have three main applications in ULSI technology:

1. As final passivation and mechanical protective layers for integrated circuits, especially for parts encapsulated in plastic packages
2. As a mask for selective oxidation of silicon
3. As a gate dielectric material in MOS devices

Micron Ex. 1023, p. 286
Micron v. YMTC
IPR2024-00909

Silicon nitride also has a high dielectric constant, about 6–9 versus about 4.2 for CVD oxide, making it less attractive for the interlevel insulation because of the resultant higher parasitic capacitances between conductive layers. However, silicon nitrides are especially valuable in applications of ULSI circuits as the oxide/nitride/oxide (ONO) dielectrics. In ULSI technology, nitride films deposited by LPCVD are often accompanied by bottom thermal oxides and oxide overlayers in the MOS capacitors of 1-Mbit DRAMs, because the higher dielectric constant—7.6 for silicon nitride relative to 3.9 for silicon dioxide—can offer a higher capacitance for nitride under the constant film thickness. Especially for the DRAMs above 4 Mbit, the required oxide film thickness has approached the fabrication limit of thermal oxide. The use of silicon nitride resolves the problem, although numerous other techniques, such as rugged polysilicon, hemispherical-grained polysilicon, Ta_2O_5 film, and other ferroelectric films, have been developed and reported. Silicon nitride is highly suitable as a passivation layer because of its following properties:

1. It can be deposited by PECVD under low temperatures, which will not degrade the underlying aluminum interconnects.
2. It behaves as a nearly impervious barrier to the diffusion of moisture and sodium.
3. It can be prepared by PECVD to have a low compressive stress, which allows it to be subjected to severe environmental stress with less likelihood of delimitation or cracking.
4. Its coverage on underlying metals is conformal.
5. It can be deposited with acceptably low pinhole densities.

A notable application of deposited dielectric films is in planarization for ULSI technology. Multilevel metallization becomes an essential technology for ULSI devices because it determines the packing density as well as device performance and affects yield and reliability. For deep submicron applications the metallization schemes are becoming more complex and dense. Therefore, the intermetal planarization technology faces serious challenge.

The intermetal dielectric application has the following requirements:

1. Good step coverage properties and void-free filling of high-aspect-ratio gaps on the underlying metal lines
2. Good planarization capability at low temperature ($\leq$ 400°C), so as not to damage the underlying interconnect

The sandwich scheme of the silane-based or TEOS-based PECVD combined with SOG (spin-on-glass) is widely utilized for the intermetal dielectric. Figure 42 indicates the simple structure. Prior to SOG coating, PE-oxide or PE-TEOS is deposited as the underlayer to prevent SOG from attaching directly to the metal. After SOG annealing, one subsequent PECVD capping layer is also performed. The main concerns are high hydrogen concentration and moisture absorption of SOG films. Furthermore, because of concern about *poisoned* vias[82] (i.e., significant increase of via contact resistance caused by SOG outgassing) and related reliability issues, the SOG etchback process was well developed to remove SOG on the metal.

However, for subsequent gap filling of sub-half-micron geometry and for higher-temperature metallization processes such as tungsten plug (W-plug), the gap-filling

Micron Ex. 1023, p. 287
Micron v. YMTC
IPR2024-00909

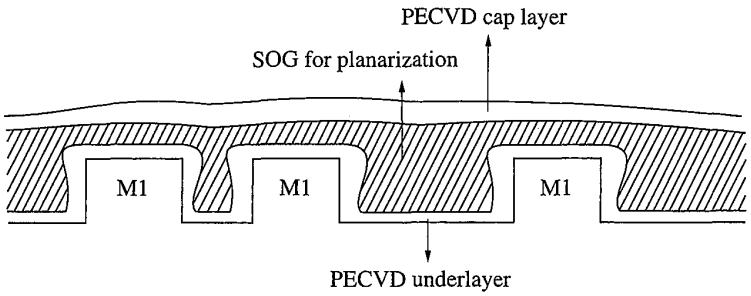


FIGURE 42
Schematic diagram of sandwiched intermetal dielectric scheme with SOG (spin-on-glass).

capability and thermal stability of SOG film suffers severe challenge. O_3–TEOS deposition process becomes another alternative. The chemistry is thermal reaction of O_2/O_3/TEOS. Basically, higher O_3/TEOS gas ratios achieve better film quality and gap-filling capability. However, there are still some problems in the O_3–TEOS process. The most important concerns are the serious moisture absorption of O_3–TEOS film and the associated impacts on reliability and device performance. So far, much effort has been expended on ways to eliminate the moisture absorption or strengthen the film characteristics by various treatments[83] and to integrate O_3–TEOS related modules.

Recently, electron cyclotron resonance (ECR) CVD technique has become more mature and has been shown to fill high-aspect-ratio gaps without voids and seams. Moreover, ECR CVD oxide possesses excellent film characteristics, including low dielectric constant (~4.07), high breakdown voltage (> 7 MV/cm), and low moisture absorption. Therefore, it meets the requirements of devices in which capacitive coupling between metal lines causes propagation delays and crosstalk. The deposition chemistry includes the reaction of SiH_4 with O_2 as well as in-situ Ar sputtering, which can achieve good gap filling and oxide quality.[84]

For planarization, the etchback process for SOG or photo resist is already widely used for production. The structure is shown in Fig. 43. One new approach, CMP (chemical-mechanical polishing) technology, expresses very good planarization performance and has become a very hot research field. The factors that affect chemical-mechanical polishing are classified into two groups. Chemistry-related parameters include slurry type, pH value of slurry, solid content in slurry, slurry flow, as well as process temperature. Mechanical-related factors include polish pressure, back pressure, platen speed, and pad type. Currently, this process is widely studied for metal and dielectric polishing. The material and pattern sensitivities are also reported. With this planarization process, the step height of wafer topography can be reduced to within 1000 Å. The schematic diagram is shown in Fig. 44.

In summary, there are two major topics for deep submicron planarization. One is the filling of gaps with high aspect ratios. The other is to achieve actual global planarization, at distances exceeding one micron. Various approaches are being developed in parallel, and extensive room remains for investigation.

Micron Ex. 1023, p. 288
Micron v. YMTC
IPR2024-00909

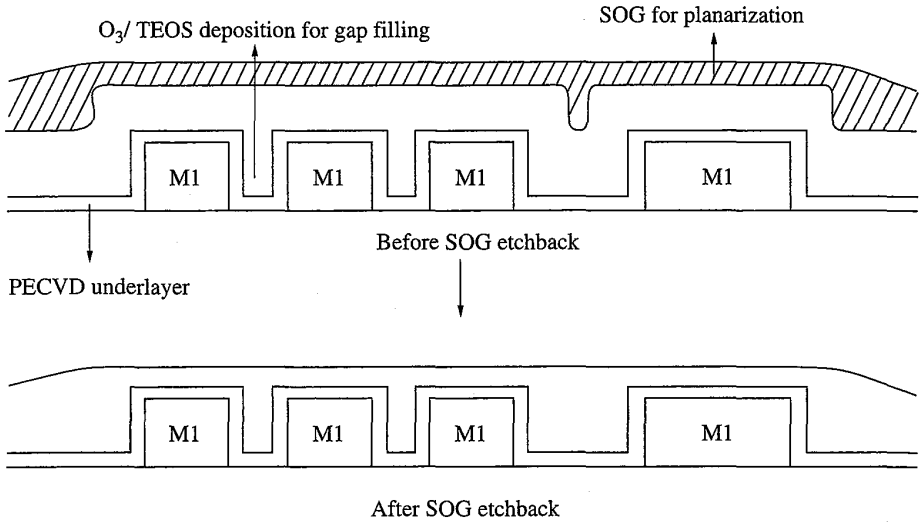


FIGURE 43
Planarization scheme with SOG etchback.

5.9 SUMMARY AND FUTURE TRENDS

To fabricate ULSI circuits, various dielectric and polysilicon film depositions are necessary. The major issues include small-geometry fabrication, low-temperature processing, particles or contamination, step coverage, planarization, selective deposition, throughput, uniformity, film properties, and large-wafer capability. ULSI devices with small dimensions require small channel length, shallow junctions, and small contacts. All of them will need precise lithography, fine pattern transfer with etching, and minimal diffusion under low-temperature processing. For precise lithography and anisotropic etching, conformal step coverage becomes more and

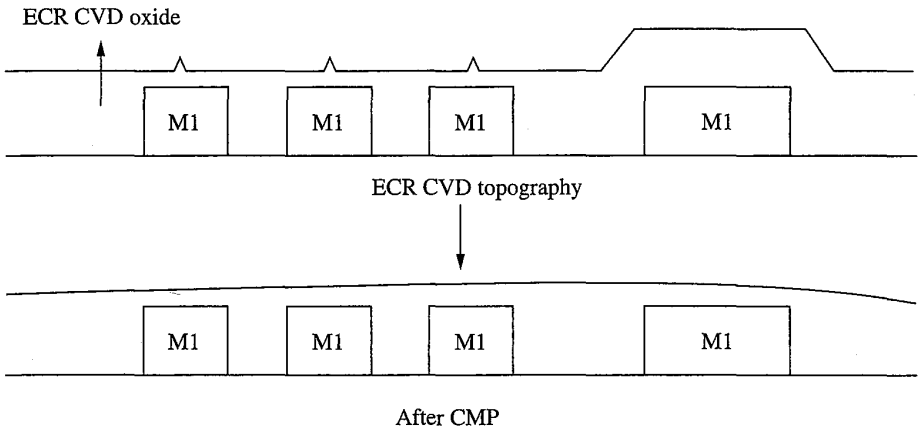


FIGURE 44
Planarization scheme with CMP.

Micron Ex. 1023, p. 289
Micron v. YMTC
IPR2024-00909

more important. Low-pressure deposition of dielectric and polysilicon films will offer this advantage. In contrast, physical vapor deposition and atmospheric pressure deposition will have nonconformal step coverage. In addition, throughput and uniformity will also be challenged with physical vapor deposition, especially for future large wafers in ULSI technology (200-mm wafers have been introduced in newly established ULSI factories).

Moreover, reducing particles and contaminations also requires a low temperature and cold-wall deposition. Defects formed during the processing will seriously lower the ULSI device yield, because these devices are scaled down and more sensitive to particles and impurity contamination. Plasma-assisted and photo-assisted depositions have supplied the benefits of lowering the deposition temperatures. However, further study on these deposition schemes will be needed to improve the properties of the deposited films. Liquid-phase deposition of silicon dioxide via the decomposition of H_2SiF_6 supersaturated with silica at reaction temperatures below 100°C has been reported. Because the conventional holder will produce particles through the motion of the holder's rollers, the cantilever-type holder and vertical-type CVD furnace have been used to decrease particle contamination, as shown in Fig. 45.

High throughput is needed to reduce cost. High deposition rates are therefore required, since the single-wafer process will likely be adopted in future ULSI technology. Deposition through generation of high-density reacting radicals is therefore emphasized. Magnetic fields have been utilized to create the helicon source or to form the electron cyclotron resonance (ECR) for the radiofrequency (rf) or microwave systems, correspondingly. Cassette-to-cassette and multichamber concepts are also favorable for increasing throughput and decreasing contamination, as shown in Fig. 46.

Various deposited films with specific properties are required for ULSI devices. As an example, dielectric films with high dielectric constants have been investigated for the MOS capacitors in ULSI DRAMs above 1 Mbit. Oxide/nitride/oxide, ferroelectric films, and stacked or trench MOS capacitors need new deposition techniques for the dielectric and polysilicon films to satisfy various requirements for the film properties. In order to increase the capacitor area, polysilicon layers can also be deposited as semispherical-grained structures by tuning the deposition temperatures. Furthermore, flash memories or SRAMs above 1 Mbit need better characteristics of polysilicon films and dielectrics on the polysilicon layers. Polysilicon with large and small grain sizes and CVD dielectrics are utilized in these ULSI devices.

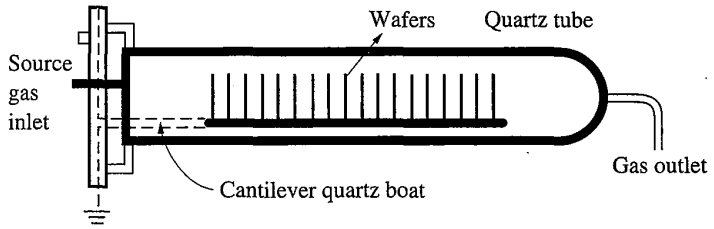


FIGURE 45
A cantilever-type LPCVD system.

Micron Ex. 1023, p. 290
Micron v. YMTC
IPR2024-00909

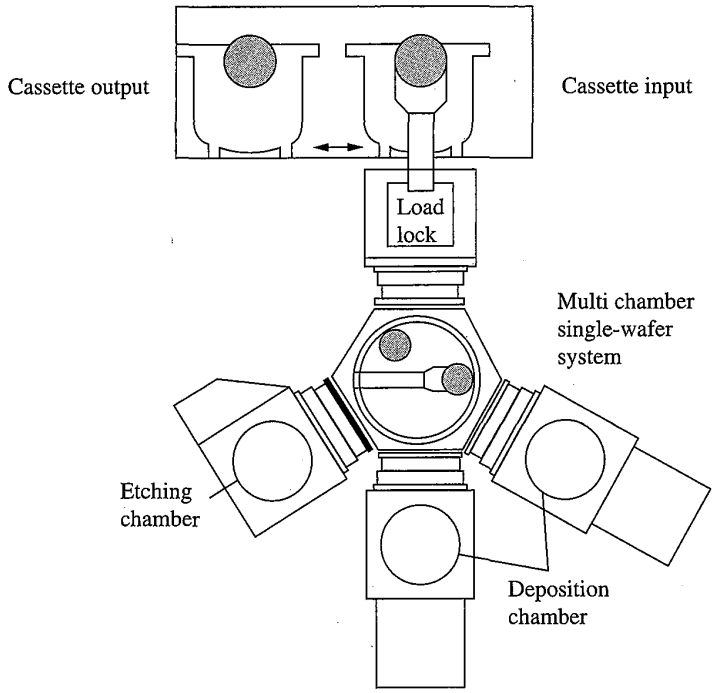


FIGURE 46
A cassette-to-cassette multichamber deposition system.

The research and development of these dielectric and polysilicon films in ULSI applications can have important impacts in other fields, such as optoelectronics, displays, microsensors, and bioelectronics. In the future these applications will need new and novel deposition techniques for these films with better properties.

REFERENCES

1. W. Kern and G. L. Schnable, "Low Pressure Chemical Vapor Deposition for VLSI Processing—A Review," *IEEE Trans. Electron Dev.* **ED-26,** 647 (1979).
2. P. Singer, "Techniques of Low Pressure CVD," *Semic. Intl.,* 72 (May, 1984).
3. M. de Fraiteur and J. Goldman, "Pressure Control in LPCVD System," *Semic. Intl.,* 250 (May, 1984).
4. A. Learn, "Modeling the Reaction of Low Pressure Chemical Vapor Deposition of Silicon Dioxide," *J. Electrochem. Soc.* **132,** 390 (1985).
5. B. Gorowitz, T. B. Gorczyca, R. J. Saia, "Application of PECVD in VLSI," *Sol. State Technol.,* 197 (June, 1985).
6. A. A. Chernov, "Growth Kinetics and Capture of Impurities during Gas Phase Crystallization," *J. Cryst. Growth,* **42,** 55 (Dec., 1977).
7. W. Kern and V. Ban, "Chemical Vapor Deposition of Inorganic Thin Films," *Thin Film Processes,* J. L. Vossen and W. Kern, Eds., Academic, New York, 1978, pp. 257–331.
8. M. Hammond, "Introduction to Chemical Vapor Deposition," *Sol. State Technol.,* 61 (Dec., 1979).

Micron Ex. 1023, p. 291
Micron v. YMTC
IPR2024-00909

9. A. C. Adams, "Plasma-Assisted Deposition of Dielectric Films," in R. Reif and G. R. Srinivasan, Eds., *Reduced Temperature Processing for VLSI,* The Electrochemical Society, Pennington, NJ, 1986, p. 111.
10. A. Weiss, "PECVD: Silicon Nitride and Beyond," *Semic. Intel.* **6(7),** 88 (July, 1983).
11. W. L. Johnson, "Design of Plasma Deposition Reactors," *Sol. State Technol.,* 191 (April, 1983).
12. M. L. Hammond, "Safety in Chemical Vapor Deposition," *Sol. State Technol.,* **23,** 104 (1980).
13. C. Cobianu and C. Pavelescu, "A Theoretical Study of the Low-Temperature Chemical Vapor Deposition of SiO_2 Films," *J. Electrochem. Soc.,* **127,** 2254 (1980).
14. C. Cobianu and C. Pavelescu, "Silane Oxidation Study: Analysis of Data for SiO_2 Films Deposited by Low Temperature Chemical Vapor Deposition," *Thin Sol. Films,* **117,** 211 (1984).
15. A. J. Learn, "Phosphorus Incorporation Effects in Silicon Dioxide Films at Reduced Pressure," *J. Electrochem. Soc.,* **132,** 405 (1985).
16. A. C. Adams and C. D. Capio, "The Deposition of Silicon Dioxide Films at Reduced Pressure," *J. Electrochem. Soc.,* **126,** 1402 (1979).
17. R. M. Levin and C. D. Adams, "Low Pressure Deposition of Phosphosilicate Glass Films," *J. Electrochem. Soc.,* **129,** 1588 (1982).
18. H. Huppertz and W. L. Engl, "Modeling of Low Pressure Deposition of SiO_2 by Decomposition of TEOS," *IEEE Trans. Electron Dev.,* **ED-26,** 658 (1979).
19. K. Watanabe, T. Tanigaki, and S. Wakayama, "The Properties of LPCVD SiO_2 Film Deposited by SiH_2Cl_2 and N_2O Mixtures," *J. Electrochem. Soc.,* **128,** 2630 (1981).
20. A. C. Adams and C. D. Capio, "Planarization of Phosphorus-Doped Silicon Dioxide," *J. Electrochem. Soc.,* **128,** 423 (1981).
21. R. A. Levy and K. Nassau, "Reflow Mechanisms of Contact Bias in VLSI Processing," *J. Electrochem. Soc.,* **133,** 1417 (1986).
22. W. Kern and R. K. Smeltzer, "Borophosphosilicate Glasses for Integrated Circuits," *Sol. State Technol.,* **28,** 171 (1985).
23. T. Foster, G. Hoeye, and J. Goldman, "A Low Pressure BPSG Deposition Process," *J. Electrochem. Soc.,* **132,** 505 (1985).
24. K. Fujino, Y. Nishimoto, N. Tokumasu, and K. Meada, "Doped Silicon Oxide Deposition by Atmospheric Pressure and Low Temperature Chemical Vapor Deposition Using Tetraethoxysilane and Ozone," *J. Electrochem. Soc.,* **138,** 3019 (1991).
25. J. D. C-Sokol, C. J. Giunta and R. Gordon, "A Kinetic Study of the Atmospheric Pressure CVD Reaction of Silane and Nitrous Oxide," *J. Electrochem. Soc.,* **136,** 2993 (1989).
26. T. Makno, "Composition and Structure Control by Source Gas Ratio in LPCVD Silicon Nitride Deposited with Different NH_3/SiH_2Cl_2 Gas Ratios," *J. Electrochem. Soc.,* **132,** 3001 (1985).
27. J. V. Dalton and J. Drobek, "Structure and Sodium Migration in Silicon Nitride Films," *J. Electrochem. Soc.,* **115,** 865 (1968).
28. H. Tanaka, H. Uchida, N. Hirashida, and T. Ajioka, "The Effect of Surface Roughness of Si_3N_4 Films on TDDB Characteristics of ONO Films," *International Reliability Physics Symposium,* IEEE, New York, 1992, p. 31.
29. T. Kamins, *Polycrystalline Silicon for Integrated Circuit Application,* Kluwer Academic Publishers, Boston, 1988.
30. S. Wolf, R. N. Taauber, *Silicon Processing for the VLSI Era, vol. 1,* Chap. 6, Lattice Press, 1986, p. 161.
31. M. Ohkura, K. Kusukawa, and H. Sunami, "Beam-Induced Seeded Lateral Epitaxy with Suppressed Impurity Diffusion for a Three-Dimensional DRAM Cell Fabrication," *IEEE Trans. Electron Dev.,* **ED-36,** 333 (1989).

Micron Ex. 1023, p. 292
Micron v. YMTC
IPR2024-00909

32. T. Yamanaka, T. Hashimoto, N. Hashimoto, T. Nishida, A. Shimizu, K. Ishibashi, Y. Sakai, K. Shimohigashi, and E. Takeda, "A 25 m^2, New Poly-Si PMOS Load (PPL) SRAM Cell Having Excellent Soft Error Immunity," *Tech. Dig. IEDM,* 48 (1988).
33. A. Mimura, N. Konishi, K. Ono, J.-I. Ohwada, Y. Hosokawa, Y. A. Ono, T. Suzuki, K. Niyata, and H. Kawakami, "High Performance Low-Temperature Poly-Si n-Channel TFT's for LCD," *IEEE Trans. Electron Dev.,* **ED-36,** 351 (1989).
34. M. Yuki, K. Masuino, and M. Kunigita, "A Full-Color LCD Addressed by Poly-Si TFT's Fabricated below 450°C," *IEEE Trans. Electron Dev.,* **ED-36,** 1934 (1989).
35. W. A. P. Claassen, J. Bloem, W. G. J. N. Valkenburg, and C. H. J. Van den Brekel, "The Deposition of Silicon from Silane in a Low-Pressure Hot-Wall System," *J. Cryst. Growth,* **57,** 259 (1982).
36. M. J. Tsai and H. C. Cheng, "Hydrogen Passivation of Thin Film Transistors with Different Film Structures," *Tech. Rep. IEICE,* **93,** 25 (1993).
37. A. Nakamura, F. Emoto, E. Fujii, A. Yamamoto, Y. Uemoto, "Analysis of Solid Phase Crystallization in Amorphized Polycrystalline Si Films," *J. Appl. Phys.,* **66,** 4248 (1989).
38. S. D. Brotherton, D. J. McCulloch, J. B. Clegg, and J. P. Gowers, "Excimer-Laser-Annealed Poly-Si Thin-Film Transistors," *IEEE Trans. Electron Dev.,* **ED-40,** 407 (1993).
39. R. E. Jones, Jr. and S. P. Wesolowski, "Electrical, Thermoelectric, and Optical Properties of Strongly Degenerate Polycrystalline Silicon Films," *J. Appl. Phys.,* **56,** 1701 (1984).
40. H. Kurokawa, "P-Doped Polysilicon Film Growth Technology," *J. Electrochem. Soc.,* **129,** 2620 (1982).
41. B. S. Meyerson and W. Olbricht, "Phosphorus-Doped Polycrystalline Silicon via LPCVD. 1. Process Characterization," *J. Electrochem. Soc.,* **131,** 2361 (1984).
42. I. W. Wu, T. Y. Huang, W. B. Jackson, A. G. Lewis, and A. Chiang, "Passivation Kinetics of Two Types of Defects in Polysilicon TFT by Plasma Hydrogenation," *IEEE Electron Dev. Lett.,* **EDL-12,** 181 (1991).
43. M. H. Juang, C. T. Lin, S. T. Jan, and H. C. Cheng, "The Process Limitation for Forming Ti Silicided Shallow Junction by BF_2^+ Implantation into Thin Polycrystalline Si Films and Subsequent Ti Silicidation," *Appl. Phys. Lett.,* **63,** 1267 (1993).
44. C. Y. Lu, J. J. Sung, R. Liu, N. S. Tsai, R. Singh, S. J. Hellenius, and H. C. Kirsch, "Process Limitation and Device Design Trade-offs of Self-Aligned $TiSi_2$ Junction Formation in Submicrometer CMOS Devices," *IEEE Trans. Electron Dev.,* **ED-38,** 246 (1991).
45. W. R. Knolle and H. R. Maxwell, Jr., "A Model of SIPOS Deposition Based on Infrared Spectroscopic Analysis," *J. Electrochem. Soc.,* **127,** 2254 (1980).
46. S. D. Chu and A. J. Steckl, "The Effect of Trench-Gate-Oxide Structure EPROM Device Operation," *IEEE Electron Dev. Lett.,* **EDL-9,** 284 (1988).
47. B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," *J. Appl. Phys.,* **36,** 3770 (1965).
48. P. W. Wang, H. P. Su, M. J. Tsai, G. Hong, M. S. Feng, and H. C. Cheng, "A New Portrayal of Oxidation of Undoped Polycrystalline Silicon Films for Short Duration," *Jpn. J. Appl. Phys.,* **33,** 429 (1994).
49. H. Dun, P. Pan, F. R. White, and R. W. Douse, "Mechanisms of Plasma-Enhanced Silicon Nitride Deposition Using SiH_4/N_2 Mixture," *J. Electrochem. Soc.,* **128,** 1555 (1981).
50. A. C. Adams, F. B. Alexander, C. D. Capio, and T. E. Smith, "Characterization of Plasma-Deposited Silicon Dioxide," *J. Electrochem. Soc.,* **128,** 1545 (1981).
51. S. Veprek, "Plasma-Induced and Plasma-Assisted Chemical Vapor Deposition," *Thin Sol. Films,* **130,** 135 (1985).
52. M. J. Rand and J. F. Roberts, "Silicon Oxynitride Films from the NO-NH_3-SiH_4 Reaction," *J. Electrochem. Soc.,* **120,** 446 (1973).

Micron Ex. 1023, p. 293
Micron v. YMTC
IPR2024-00909

53. W. A. P. Claassen, H. A. J. Th. V. d. Pol, A. H. Goemans, and A. E. T. Kuiper, "Characterization of Silicon-Oxynitride Films Deposited by Plasma-Enhanced CVD," *J. Electrochem. Soc.,* **133,** 1458 (1986).
54. A. C. Carlson, T. H. Tom Wu, and H. B. K. Liang, "Deposition Method to Control Plasma-Enhanced Chemical Vapor Deposition Tetraethylortho Silicate Oxide Charge," *J. Electrochem. Soc.,* **140,** 774 (1993).
55. R. S. Rosler, "The Evolution of Commercial Plasma Enhanced CVD Systems," *Sol. State Technol.,* 67 (June, 1991).
56. I. Avigal, "Inter-metal Dielectric and Passivation-Related Properties of Plasma BPSG," *Sol. State Technol.,* **26,** 217 (1983).
57. A. Madan, P. Rava, R. E. I. Schropp, and B. v. Roedern, "A New Modular Multichamber Plasma Enhanced Chemical Vapor Deposition System," *Appl. Surf. Sci.,* **70/71,** 716 (1993).
58. A. C. Adams, "Dielectric and Polysilicon Film Deposition," in S. M. Sze, Ed., *VLSI Technology,* McGraw-Hill Book Company, New York, 1983.
59. P. A. Murawala, M. Sawai, T. Tatsuta, O. Tsuji, "Structural and Electrical Properties of Ta_2O_5 Grown by the Plasma-Enhanced Liquid Source CVD Using Penta Ethoxy Tantalum Source," *Jpn. J. Appl. Phys.,* **32,** pt. 1, 368 (1993).
60. G. Q. Lo, D. L. Kwong, P. C. Fazan, V. K. Mathews, and N. Sandler, "Highly Reliable, High-C DRAM Storage Capacitors with CVD Ta_2O_5 Film on Rugged Polysilicon," *IEEE Electron Dev. Lett.,* **EDL-14,** 216 (1993).
61. J. W. Peters, F. L. Gebhart, and T. C. Hall, "Low-Temperature Photo-CVD Silicon Nitride—Properties and Applications," *Sol. State Tech.,* **23,** 121 (1980).
62. R. Solaki, C. Moore, and G. Collins, "Laser Induced CVD," *Sol. State Technol.,* 220 (June, 1985).
63. M. Kitagawa, S. I. Ishihara, K. Setsune, Y. Manabe, and T. Hirao, "Low Temperature Preparation of Hydrogenated Amorphous Silicon by Microwave Electron-Cyclotron-Resonance Plasma CVD," *Jpn. J. Appl. Phys.,* **26,** L231 (1987).
64. M. Kitagawa, K. Setsune, Y. Manabe, and T. Hirao, "Preparation of Doped Hydrogenated Amorphous Silicon by Microwave Electron-Cyclotron-Resonance Discharge Deposition," *J. Appl. Phys.,* **61,** 2048 (1987).
65. O. A. Popov and H. Waldron, "Electron Cyclotron Resonance Plasma Stream Source for Plasma Enhanced Chemical Vapor Deposition," *J. Vac. Sci. Technol. A,* **7,** 914 (1989).
66. B. S. Meyerson, "Low-Temperature Silicon Epitaxy by Ultrahigh Vacuum/Chemical Vapor Deposition," *Appl. Phys. Lett.,* **48,** 797 (1986).
67. H. C. Lin, H. Y. Lin, C. Y. Chang, T. F. Lei, P. J. Wang, and C. Y. Chao, "Growth of Undoped Polycrystalline Si by an Ultrahigh Vacuum Chemical Vapor Deposition System," *Appl. Phys. Lett.,* **63,** 1351 (1993).
68. M. Miyasaka, T. Nakazawa, I. Yudasaka, and H. Ohshima, "TFT and Physical Properties of Polycrystalline Silicon Prepared by Very Low Pressure Chemical Vapor Deposition (VLPCVD)," *Jpn. J. Appl. Phys.,* **30,** 3733 (1991).
69. M. Liehr, S. S. Dana, and M. Anderle, "Nucleation and Growth of Silicon on SiO_2 during SiH_4 Low-Pressure Chemical Vapor Deposition as Studied by Hydrogen Desorption Titration," *J. Vac. Sci. Tech. A,* **10,** 869 (1992).
70. S. M. Gates and S. K. Kalkarn, "Kinetics of Surface-Reactions in Very Low Pressure Chemical Vapor Deposition of Si from SiH_4," *Appl. Phys. Lett.,* **58,** 2963 (1991).
71. H. C. Lin, C. Y. Chang, W. H. Chen, W. C. Tsai, T. C. Chang, T. G. Jung, and H. Y. Lin, "Effects of SiH_4, GeH_4, and B_2H_6 on the Nucleation and Deposition of Polycrystalline $Si_{1-x}Ge_x$ Films," *J. Electrochem. Soc.,* **141,** 2559 (1994).
72. T. J. King and K. C. Saraswat, "Low-Temperature (Less Than or Equal to 550°C) Fabrication of Poly-Si Thin-Film Transistors," *IEEE Electron Dev. Lett.,* **EDL-13,** 309 (1992).

Micron Ex. 1023, p. 294
Micron v. YMTC
IPR2024-00909

73. H. C. Lin, H. Y. Lin, C. Y. Chang, T. F. Lei, P. J. Wang, R. C. Deng, J. Lin, and C. Y. Chao, "Deposition and Device Application of In-Situ Boron-Doped Polycrystalline SiGe Films Grown at Low Temperature," *J. Appl. Phys.*, **74,** 5395 (1993).
74. T. J. King, J. R. Pfiester, J. D. Shott, J. P. McVittie, and K. C. Saraswat, "A Polycrystalline-$Si_{1-x}Ge_x$-Gate CMOS Technology," *IEEE Proc. IEDM,* 253 (1990).
75. A. Hai, J. D. Morse, and R. W. Dutton, "1-GHz Integrated Poly-Si-and-SiGe Photoconductors with BiCMOS Compatibility," *IEEE Proc. IEDM,* 41 (1991).
76. S. Yamamoto and M. Migitaka, "Hybrid-Excitation Chemical Vapor Deposition of Silicon Nitride on (100)Si—The Film and Interface Properties," *Jpn. J. Appl. Phys.*, **31,** Pt. 1, 348 (1992).
77. D. Peters, "Implanted-Silicide Polysilicon Gate for VLSI Transistors," *IEEE Trans. Electron Dev.,* **ED-33,** 1391 (1986).
78. H. Sunami, "Cell Structure for Future DRAMs," *IEEE Proc. IEDM,* 694 (1985).
79. N. Matsuo and H. Ogawa, "Nucleation and Growth Mechanism of Hemispherical Grain Polycrystalline Silicon," *Appl. Phys. Lett.,* **60(21),** 2607 (1992).
80. K. S. Tseng, H. C. Cheng, C. C. Chang, C. G. Lou, and F. H. Yang, "Effects of n^+-Polysilicon/SiO_2 Interface on the Electrical Characteristics of MOS Capacitors," Unpublished.
81. H. H. Tseng, P. J. Tobin, J. D. Hayden, K. M. Chang, "Advantages of CVD Stacked Gate Oxide for Robust 0.5 μm Transistors," *IEEE Proc. IEDM,* 75 (1991).
82. N. Rutherford, M. Camenzind, and A. Belic, "Outgassing and Oxidative Damage in Non-Etchback Siloxane SOG Process," *Proc. 10th VMIC,* 141 (1993).
83. K. Fujino, Y. Nishimoto, N. Tokumasu, S. Fisher, and K. Maeda, "Plasma Post-Treatment for Tetraethoxysilane/O_3 Atmospheric Pressure CVD," *Proc. 10th VMIC,* 96 (1993).
84. C. Y. Chang, J. P. McVittie, J. Li, K. C. Saraswat, S. E. Lassig, and J. Dong, "Profile Simulation of Plasma Enhanced and ECR Oxide Deposition with Sputtering," *IEEE Proc. IEDM,* 853 (1993).

PROBLEMS

1. In order to increase the capacitance of MOS capacitors in ULSI DRAMs, Ta_2O_5 has been plasma-deposited to replace the oxide/nitride/oxide dielectric. The dielectric constants of SiO_2, nitride, and Ta_2O_5 are about 3.9, 7.6, and 25, respectively. What is the capacitance ratio for the capacitors with the Ta_2O_5 and oxide/nitride/oxide dielectrics for the same dielectric thickness, provided that the oxide/nitride/oxide has thickness ratio 1 : 1 for the oxide to the nitride?

2. For the CVD of phosphosilicate glass, 20% silane is mixed in 80% nitrogen at a reaction temperature of 500°C. The flow rate is 600 sccm (standard cubic centimeter per minute). The reactor is loaded with 10 wafers, 10 cm in diameter. The growth rate and film density are 600 Å/min and $2 \times 10^{22}/cm^3$, respectively. With what efficiency is the silane transformed into silicon dioxide film? (*Hint:* A flow of one sccm is defined as a flux of one cm^3 of gas, as measured at 273 K and 760 torr, per minute.)

3. Phosphosilicate glass consists of SiO_2 and P_2O_5. Derive the relationships of weight percent (wt %), atom percent (at%), and mole percent (mole %), for the dopant concentration.

Micron Ex. 1023, p. 295
Micron v. YMTC
IPR2024-00909

4. (*a*) What factors will affect the step coverage of the deposited oxide films?

   (*b*) What techniques have been used to improve the step coverage of the deposited oxide films?

5. The LOCOS process often causes a "bird's beak" phenomenon, i.e., field-oxide encroachment under the silicon nitride mask. A decrease of the stress between CVD nitride and Si substrate can reduce this effect. What methods can you suggest to attain this purpose?

6. If LPCVD polysilicon deposition has an activation energy of 1.65 eV and a deposition rate of 8 nm/min at 600°C, what is the deposition rate at 620°C?

7. In a PECVD system, the chamber pressure is 1 torr, the total gas flow rate is 100 sccm, the temperature is 27°C, and the volume of the chamber is 1000 cm^3. What is the average residence time for a gas molecule in the PECVD reactor?

8. For an ECR CVD system the magnetic field is 875 gauss. What is the frequency of the microwave power to generate electron cyclotron resonance?

Micron Ex. 1023, p. 296
Micron v. YMTC
IPR2024-00909

CHAPTER 6

Lithography

K. Nakamura

6.1 INTRODUCTION

Lithography is a kind of art made by impressing, in turn, several flat embossed slabs, each covered with greasy ink of a particular color, onto a piece of paper. The various colors or levels must be accurately aligned with respect to one another within some registration tolerance. Many "originals" can be made from the same slabs as long as the quality remains adequately high.

Several methods can be used to make ULSI circuit patterns on wafers, as shown in Fig. 1*a*. The most common process is to make the master photomask using an electron beam exposure system and replicating its image by optical printers, as shown in Fig. 1*b*. The exposing radiation is transmitted through the "clear" part of a mask. The opaque part of the circuit pattern blocks some of the radiation. The resist, which is sensitive to the radiation and has resistance to the etching, is coated on the wafer surface. The mask is aligned within the required tolerance on the wafer; then radiation is applied through the mask, the resist image is developed, and the layer underneath the resist is etched.

Therefore, lithography for integrated circuit manufacturing is analogous to the lithography of the art world. The slabs correspond to masks for the various circuit levels. The press corresponds to the exposure system, which not only exposes each level but also aligns it to a completed level. The ink may be compared to either the exposing radiation or the radiation-sensitive resist; the paper can represent the wafer into which the pattern will be etched, using the resist as a stencil.

Lithography is the key technology in semiconductor manufacturing, because it is used repeatedly in a process sequence that depends on the device design. It determines the device dimensions, which affect not only the device's quality but also its product amount and manufacturing cost.

Micron Ex. 1023, p. 297
Micron v. YMTC
IPR2024-00909

FIGURE 1
Device lithography generalization. (*a*) Lithographic process for ULSI. (*b*) Optical replication process.

Micron Ex. 1023, p. 298
Micron v. YMTC
IPR2024-00909

6.2 OPTICAL LITHOGRAPHY

Optical lithography comprises the formation of images with visible or ultraviolet radiation in a photoresist using proximity or projection printing. Two methods are available to make masks for optical lithography: electron beam exposure and laser beam scanning. These are described in the next section. In the 1970s, the major technology was the combination of a negative resist and proximity printing. At present the most common method is the combination of a positive resist and a stepper. However, proximity printing is still used because of its convenience and low cost.

Table 1 lists examples of commercially available optical printers[1] designed to manufacture ULSI circuits. These machines are classified into three groups: proximity, reflective projection, and refractive projection. The key parameters such as numerical aperture (NA), depth of focus (DOF), resolution (usable linewidth), overlay accuracy, and throughput are also listed.

The most advanced ULSI has a minimum feature size of 0.3 μm to 0.4 μm, which has been reported by several organizations in 64-Mbit dynamic random access memories (DRAMs). It is believed that the linewidth limit of optical lithography lies near 0.2 μm using phase-shifting masks combined with a high numerical aperture projector and a short-wavelength light source. The performances of the machines listed in Table 1 are very close to the resolution limit of optical lithography.

6.2.1 Contact and Proximity Printing

Contact and proximity printings are relatively simple because they do not have any means of image formation between masks and wafers. A typical contact or proximity mask aligner consists of a light source, a condenser, a filter, a mirror, a shutter, the wafer stage, and the alignment microscope. In contact printing, a photomask is pressed against the resist-covered wafer, with pressures typically in the range of 0.05 atm to 0.3 atm, and is exposed by light with a wavelength near 400 nm. Very high resolution of less than 0.5 μm linewidth is possible, but because of spatial nonuniformity of the contact, resolution may vary considerably across the wafer. To provide better contact over the whole wafer, a thin (0.2 mm) flexible mask has been used; 0.2-μm space patterns have been formed by using 3-μm-thick PMMA (poly-methyl methacrylate) resist and 200 to 260 nm radiation.[2] Quartz or Al_2O_3 mask substrates must be used to pass these shorter wavelengths, since the usual borosilicate glass strongly absorbs wavelengths less than 300 nm.

Contact printing produces defects in both the mask and the wafer while the two are in contact and as they are separated from each other, so that the mask, whether thick or thin, may have to be replaced after a short period of use. Nevertheless, contact printing is still widely used. It is the most convenient way to get high resolution. Strictly speaking, *contact* printing is never actually practiced, because of the difficulty of uniform contact.[3] When a typical mask and wafer are brought into hard contact, the nonuniform gap can be as large as 15 μm because of unevenness of the surfaces. Besides that, the thickness of the resist between the mask and the wafer cannot be neglected for features of less than 1 μm.

Micron Ex. 1023, p. 299
Micron v. YMTC
IPR2024-00909

TABLE 1
Commercially available optical printers[1]

Category	Model	Maker	NA†	DOF† (μm)	Field size (mm)	Usable linewidth (μm)	Overlay accuracy (μm)	Throughput/6″-hr
Contact/proximity	PLA600	Canon			160	1	0.5	120
	3HRP	JBA				0.25	0.5	45/10″
	Q-6000	Quintel				0.25		60
	MA150M	Karl Suss				0.5	0.5	60
Catadioptric/4×	Micrascan IIi	SVG	0.5	1.2	22 × 32.5	0.5	0.07	51/8″
/1×	Micralign 700	SVG	0.167	6	150 × 135	1.0–1.2	0.25	100
/1×	2000gh	Ultratech	0.24–0.40	2–3	18 × 18–39 × 11	0.8–1.0		78
I-line stepper	PAS5500/100	ASM	0.6	1	22 × 27.6	0.4	0.07	88
	XLS7500/29	GCA	0.55	1.1	21 × 21	0.5	0.09	70
	NSR2005i9C	Nikon			21.2 × 22.8/18 × 25.2	0.45	0.1	47/8″
	FPA2500i3	Canon	0.6	1	20 × 20	0.4	0.12	62
KrF stepper	PAS5500/90	ASM	0.5	1	21 × 27.7	0.35	0.07	80
	XLS7800/31	GCA	0.53	1.2	22 × 22	0.35	0.07	72
	ISR2005EXBA	Nikon			21.2 × 21.2/16 × 25.2	0.4	0.1	34/8″
	FPA3000EX1	Canon	0.45	1	20 × 20	0.3	0.1	66

†NA - numerical aperture
DOF - depth of focus

Micron Ex. 1023, p. 300
Micron v. YMTC
IPR2024-00909

Masks used for proximity printing have the advantage of longer life because there is no contact between the mask and the wafer. Typical separations between mask and wafer are in the range of 20 to 50 μm. Resolution is not so good as in contact printing or projection printing. Commercial machines can select the mode automatically and switch to either contact or proximity printing.

Figure 2*a* shows proximity printing schematically with a slit of width W, illuminated by a monochromatic source, separated from a parallel image plane (wafer) by a gap g. We assume that g and W are larger than the wavelength λ of the imaging light and that $\lambda \ll g < W^2/\lambda$—the region of Fresnel diffraction. In this case, the diffraction that forms the image of the slit is a function only of the particular combination of λ, W, and g, which we call the parameter Q, where[4]

$$Q = W\sqrt{2/g\lambda} \tag{6.1}$$

This parameter can be considered as a normalized slit width. The lower limit of the nearfield diffraction region is at $Q = \sqrt{2}$ from Eq. (6.1), or $g = (W^2/\lambda)(2/Q^2) < W^2/\lambda$. Thus, the resolution, W in Eq. (6.1), becomes better at smaller gaps and shorter wavelengths. The resist image heights, which for practical purposes are converted to resist thickness and mask-wafer gap, have been measured as a function of linewidth, wavelength, and intensity tolerance.[3] The optimum image heights of 0.25-μm and 0.5-μm patterns are 0.9 μm and 3.2 μm, respectively, in the case of a deep UV (wavelength less than about 200 nm) light source with ±5% intensity tolerance.

Another parameter is the divergence of the illuminating beam as a result of the size of the light source. One effect is to smooth out the undulations in the image intensity profile, and the other is to produce a greater linewidth variation as the mask-to-wafer distance varies. The apparent source size of the illumination system must

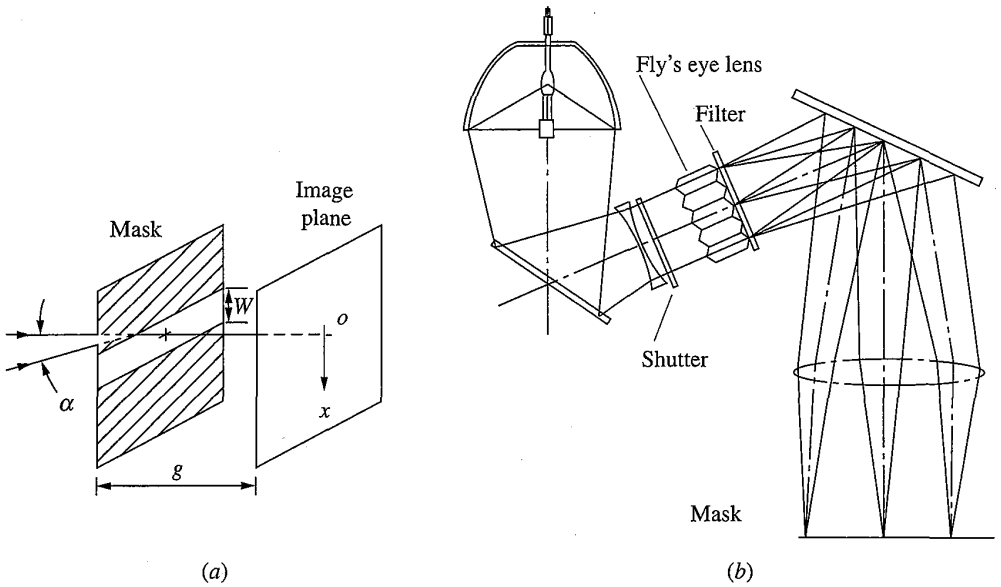


FIGURE 2
Schematic illustration of a proximity printer and illumination system. (*a*) Proximity printing schematic. (*b*) Illumination system of the Canon PLA600.

Micron Ex. 1023, p. 301
Micron v. YMTC
IPR2024-00909

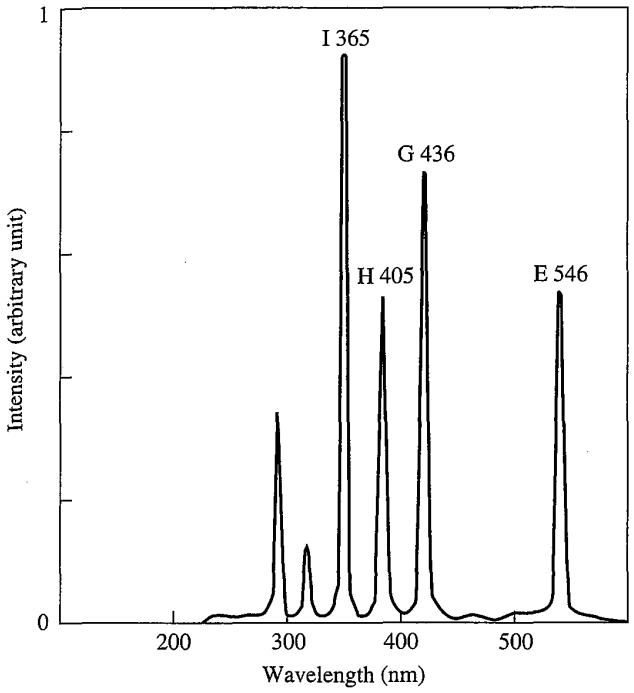


FIGURE 3
Spectrum of a high-pressure mercury lamp.

be large enough to give a value of α (Fig. 2) that allows the smallest features to be printed, typically, a few degrees. The mercury arc lamp used as the source[5] is too small to yield the required α. The illumination is telecentric or normally incident at the mask to prevent runout (magnification) error similar to that in x-ray lithography. The optical system must also minimize nonuniformity of the intensity across the field. The illumination system illustrated in Fig. 2*b* attains ±3% uniformity across a 6-inch wafer, as well as a large enough α and normal incidence. With a Hg arc source, the strong lines at 436 nm, 405 nm, and 365 nm provide the exposure flux shown in Fig. 3. The same printer is available with a Xe-Hg source for enhanced output in the 200–300 nm spectral region.

Commercially available machines are listed in the first group in Table 1.

6.2.2 Projection Printing

Projection printing offers higher resolution than proximity printing. It has a larger separation between the mask and the wafer because of its image formation system. According to Rayleigh's criterion, the resolution W of the optical systems and the depth of focus (DOF) are given in the following equations, where the numerical aperture NA = $n \sin \alpha$, n is the refractive index = unity for air, 2α is the solid angle of the ray reaching the imaging point from the objective lens, and λ is the wavelength:[6]

$$W = 0.6\lambda/\text{NA} \tag{6.2}$$

$$\text{DOF} = \pm\lambda/2(\text{NA})^2 \tag{6.3}$$

Micron Ex. 1023, p. 302
Micron v. YMTC
IPR2024-00909

Modern projection printers employ diffraction-limited optics, which means that the design and fabrication of the optical elements are not a source of optical problems. Accordingly, the image characteristics are dominated by diffraction effects associated with the finite apertures in the condenser and with projection optics rather than by aberrations. When monochromatic light from a very small point is imaged by a diffraction-limited lens, the image consists of diffraction rings of light surrounding a central bright spot called the *Airy disc*.[6] The diameter of the pattern is $1.2\lambda/\text{NA}$.

It is very useful to apply a modulation transfer function (MTF) to characterize the resolution capability of projection printers and proximity printers as well.[7] We consider the mask pattern as consisting of a periodic grid (line and space) with equal linewidths b. Light intensities in the center of dark lines and of bright lines are assigned as $I_{\min}$ and $I_{\max}$, respectively. Modulation is defined as

$$M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{6.4}$$

The MTF is the ratio of the modulation in the image plane (wafer) to that in the object plane (mask). It is a function of the spatial frequency ($\nu = 1/2b$) of the mask and the NA. The shape of the MTF, related to the spatial frequency curve, depends on the coherence of the illumination system.[7] The degree of coherence is defined by $\sigma = (\text{NA})_c/(\text{NA})_o$, where $(\text{NA})_c$ is the numerical aperture of the condenser and $(\text{NA})_o$ is the numerical aperture of the objective as shown in Fig. 4. In the coherent system, Fig. 4*a*, an object at point A is illuminated by only a narrow angle. Hence, all the light diffracted by A is coherent in amplitude at the image plane. In Fig. 4*b*, A is illuminated by rays from all portions of the extended incoherent source. Each ray is diffracted by the object A and forms an image in the image plane.

In Fig. 5, the solid lines show a plane wave front [$\sigma = (\text{NA})_c/(\text{NA})_o \rightarrow 0$] incident normal to the mask, which contains a grating pattern with b width lines and spaces (spatial frequency $\nu = 1/2b$).[8] The undiffracted component of light emerging from the mask contains no information about ν. This information is contained only in the diffracted light. The direction of the first diffraction peak is given by the grating formula $2b \sin\theta = \lambda$, so that $\nu = \sin\theta/\lambda$. If the light diffracted to direction θ is to reach the image plane, $\theta \approx \alpha$ where $\text{NA} = \sin\alpha$. Therefore, the highest grating frequency that can be imaged by an optical system with coherent illumination becomes the result of Eq. (6.5).

For incoherent illumination (shown by dotted lines in Fig. 5), incident at a general angle i, the direction of the first maximum is given by $2b(\sin i + \sin\theta) = \lambda$. For both the undiffracted beam and the first diffraction peak to reach the image plane, i and θ both $\leq \alpha$. Therefore, $2b \geq \lambda/2\sin\alpha$ and $\nu_{\max} = 2b = 2\text{NA}/\lambda$. Consequently, $\nu_{\max}$ (incoherent) $= 2\nu_{\max}$ (coherent). In modern printers the illumination is intermediate between the coherent and the incoherent limit.

The curve for MTF ($H(\nu)$) vs. spatial frequency[9] is shown in Fig. 6 and is described as[5]

$$\nu_{\max} = \text{NA}/\lambda \qquad \text{(coherent)} \tag{6.5}$$

Micron Ex. 1023, p. 303
Micron v. YMTC
IPR2024-00909

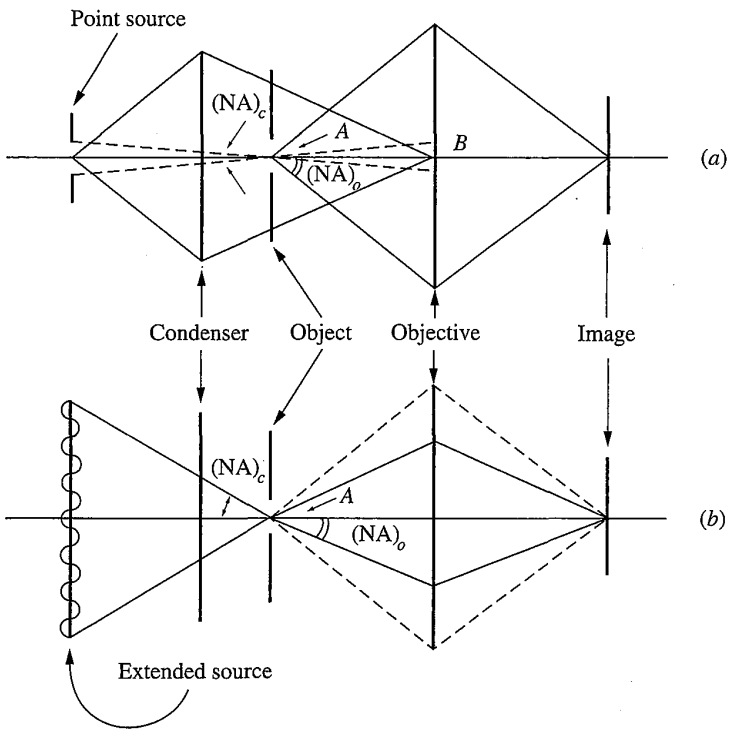


FIGURE 4
Coherent and incoherent illumination of projection printers. (*After King and Goldrick, Ref. 7.*) (*a*) Coherent system. (*b*) Incoherent system.

$$H(\nu) = \frac{2}{\pi}[\cos^{-1}(\nu/\nu_{\max}) - \nu/\nu_{\max}\sqrt{1-(\nu/\nu_{\max})^2}] \qquad \text{(incoherent)} \quad (6.6)$$

where $\nu_{\max} = 2\text{NA}/\lambda$

Figure 6 shows the modulation of the image intensity versus the spatial frequency for different values of coherence σ with a 0.3 NA lens. Note that the MTF of a fully coherent optical system has a spatial frequency cutoff that is half that of an incoherent system as explained in Eq. (6.5) and Eq. (6.6). Partially coherent illumination has a higher MTF in the region of ν larger than $\frac{1}{2}\nu_0$. The useful range, MTF > 0.6, is extended to higher spatial frequencies; edge gradients in the image become steeper, and the image is somewhat less sensitive to focusing.

The focus error is a simple but important aberration. The error is a displacement of best focus away from its intended position. In Fig. 6, the dashed curve shows the effect of a displacement of the image plane or wafer from the focal plane by one Rayleigh unit $w = \lambda/2(\text{NA})^2$, corresponding to a phase error of $\pi/2$ at the edge of the pupil.

It was once believed that a high NA is always better. However, in the submicrometer region there is an optimum NA if a resolution requirement and the imaging wavelength are given.[10] If the NA is too low, the resolution cannot be achieved, but if the NA is too high, the DOF, which is inversely proportional to $(\text{NA})^2$, becomes

Micron Ex. 1023, p. 304
Micron v. YMTC
IPR2024-00909

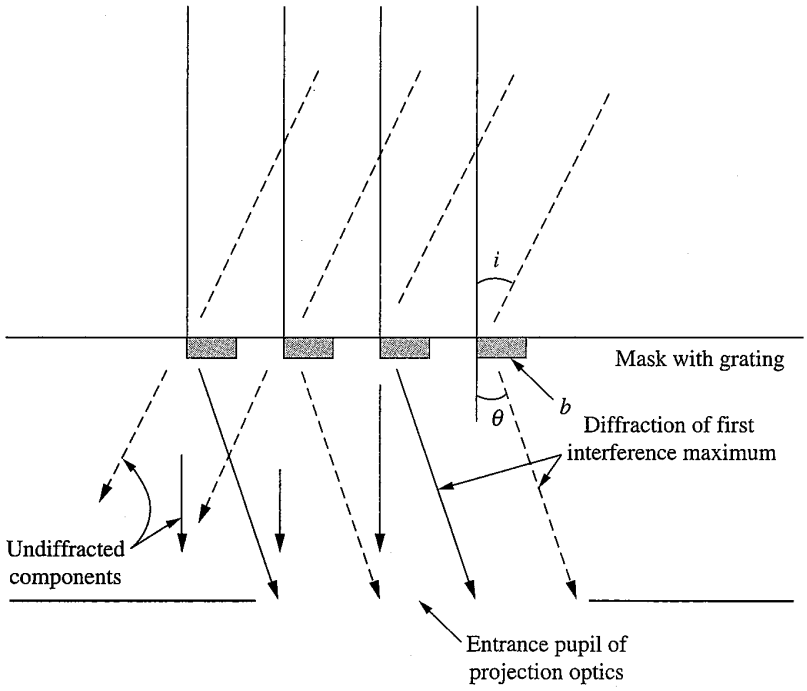


FIGURE 5
Diffraction of coherent and incoherent light by a grating pattern. (*After Cuthbert, Ref. 8.*)

unacceptable. There is an optimum NA at which the DOF is maximum. The normalization of the resolution W is given by

$$k_1 = W\frac{\text{NA}}{\lambda} \tag{6.7}$$

where k_1 is usually known as Rayleigh's coefficient for resolution but is now redefined as the normalized resolution. The normalized DOF is given by

$$k_2 = \Delta Z\frac{(\text{NA})^2}{\lambda} \tag{6.8}$$

where ΔZ is the physical axial displacement from the focal plane, and k_2, which is usually known as Rayleigh's coefficient for DOF, is now redefined as the normalized DOF.

With normalized resolution and defocus, the lithographic imaging behavior of a given feature and illumination can be universally plotted in the exposure-defocus (E-D) space in the form of constant linewidth contours. The exposure dosage required to keep the image linewidth constant at each defocal plane is evaluated to form contours. This set of contours, which is the E-D tree of the particular feature and illumination condition, is universally applicable for all W-λ-NA combinations leading to an identical k_1. The DOF is defined as the total amount of defocus allowed without violating a given linewidth tolerance. Typically, a $\pm 10\%$ tolerance is

Micron Ex. 1023, p. 305
Micron v. YMTC
IPR2024-00909

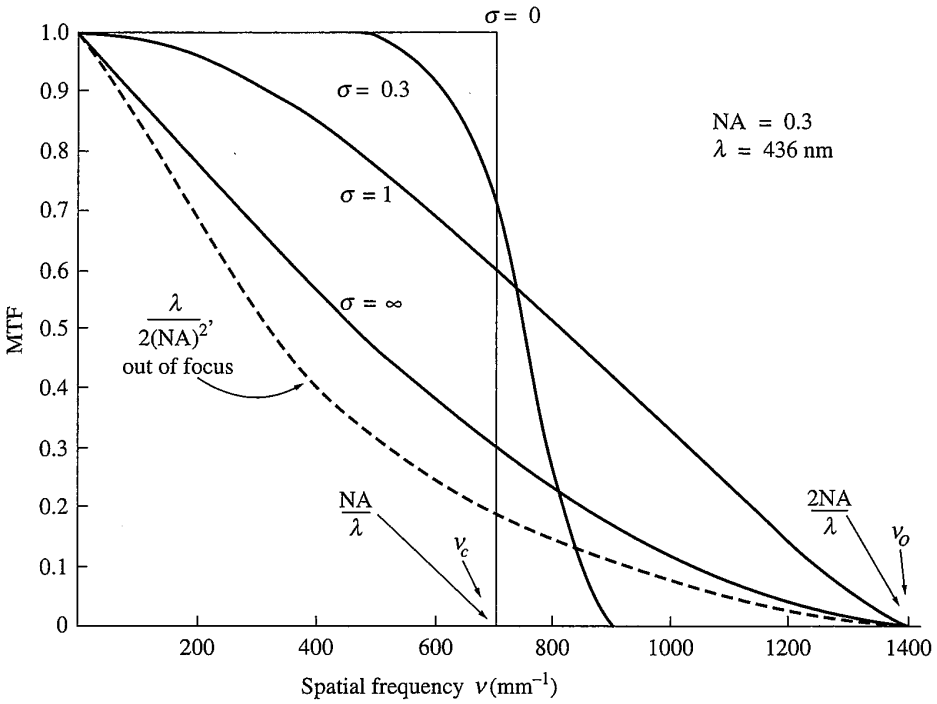


FIGURE 6
Comparison of MTFs for different optical conditions. (*After Lacombat et al., Ref. 9.*)

adopted. The DOF is the height of the largest window whose width is the exposure budget (process margin) that accommodates the tolerances that can affect the actual exposure dosage. Figure 7*a* shows the DOF determined by 10% and 30% exposure budgets, respectively. Figure 7*b* shows two E-D windows. One is bound by the common zone of three kinds of feature (line and spaces, isolated line openings, isolated spaces), and the other by five kinds of feature (including holes and islands).

The DOF equation is produced from Eq. (6.7) and Eq. (6.8) as

$$\text{DOF} = \frac{k_2}{k_1^2}\frac{W^2}{\lambda} \tag{6.9}$$

where λ/W = NA coefficient
W^2/λ = DOF coefficient
k_2/k_1^2 = DOF, normalized to W and λ

Figure 8*a* and 8*b* shows k_2 and k_2/k_1^2, respectively, as a function of k_1, and Fig. 8*c* shows W^2/λ and λ/W as a function of W. The normalized DOF and k_1 can be converted to physical DOF and NA by multiplying by the DOF coefficient and the NA coefficient, respectively, which are plotted in Fig. 8*c*.

The optimum NA for individual feature shapes is much lower than expected, as a result of Fig. 8*c*. The results lead to the problems that occur in reducing the usable k_1 factor for manufacturing. The optimum k_1 for the single-layer resist process (10% linewidth budget estimated) ranges from 0.57 to 0.87 depending on the feature shape.

Micron Ex. 1023, p. 306
Micron v. YMTC
IPR2024-00909

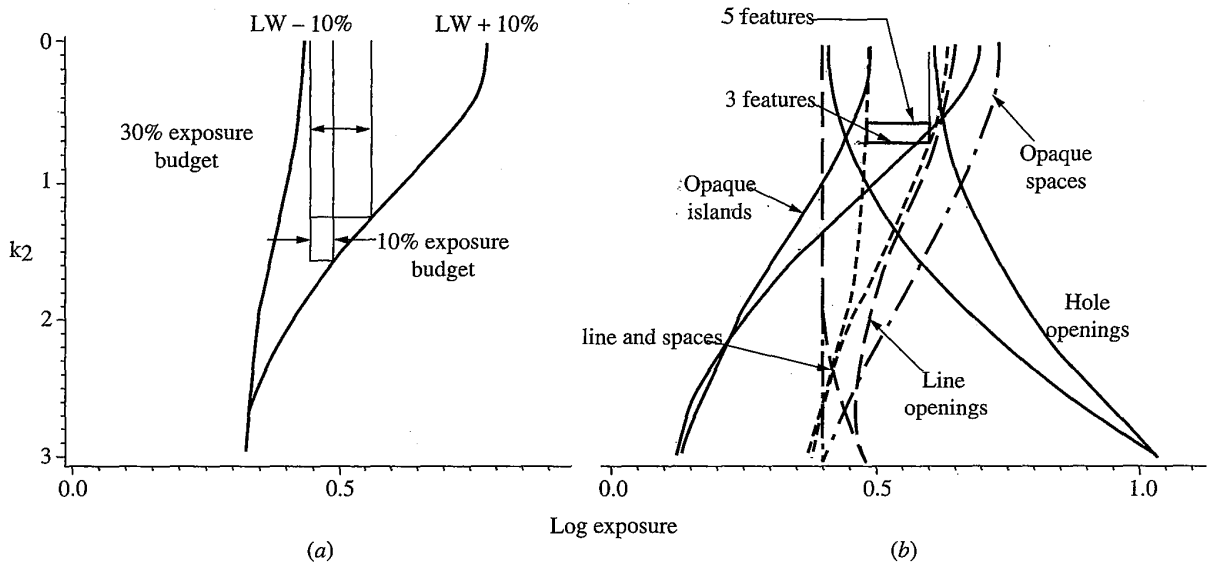
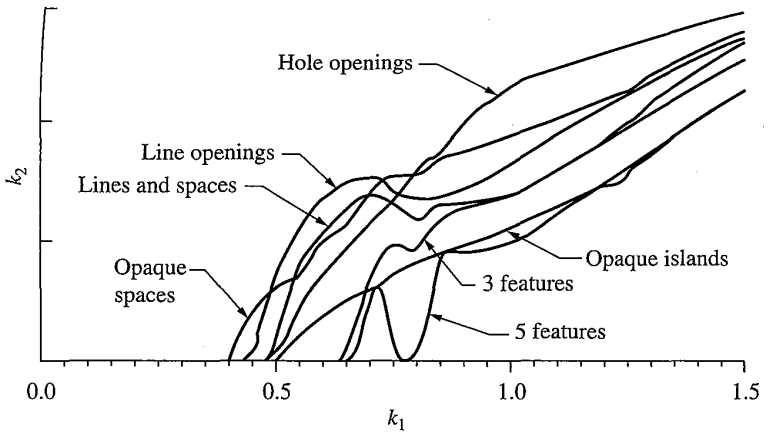
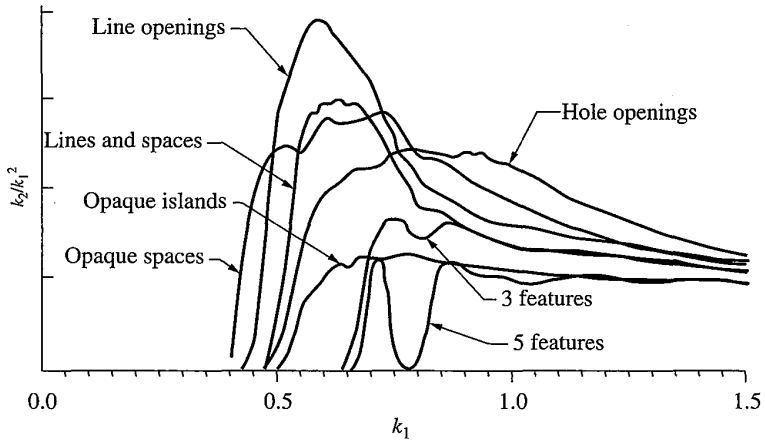


FIGURE 7
E-D windows. (*After Lin, Ref. 10.*) (*a*) The 10% and 30% E-D window bounded by ±10% E-D branches of line and space pairs at $k_1 = 0.8$, $\sigma = 0.4$. (*b*) The 3-feature and 5-feature E-D windows at $k_1 = 0.7$, $\sigma = 0.52$.

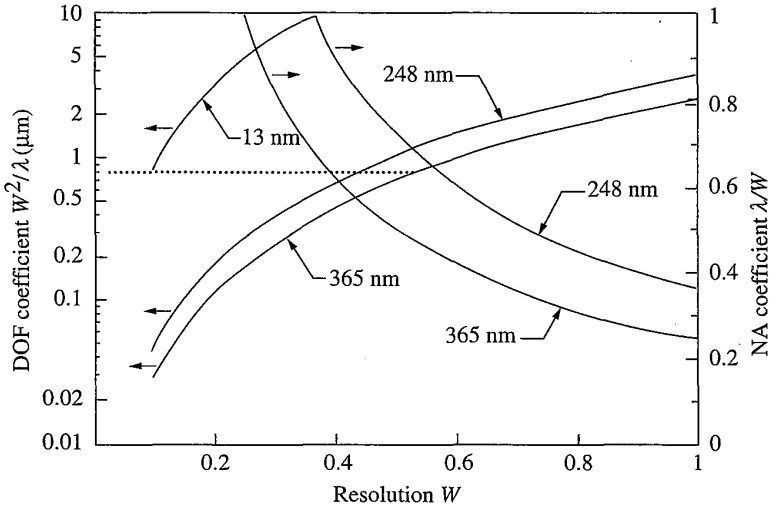
Micron Ex. 1023, p. 307
Micron v. YMTC
IPR2024-00909



(*a*)



(*b*)



(*c*)

FIGURE 8
NA optimization parameters for optical projection printers. (*After Lin, Ref. 10.*) (*a*) k_2 as a function of k_1 at 10% budget. (*b*) k_2/k_1^2 as a function of k_1 at 10% budget. (*c*) The DOF coefficient and NA coefficient as a function of W. The dotted line shows the similar difficulty of achieving 0.1μm by X-ray and 0.5μm by I-line.

Micron Ex. 1023, p. 308
Micron v. YMTC
IPR2024-00909

For a multilayer resist process (30% budget), k_1 ranges from 0.42 to 0.7, although it is generally defined as 0.6 as shown in Eq. (6.2).

Reflection and catadioptric projection

A significant advantage of the reflective or the catadioptric system[11] (combining reflecting and refracting components) is the larger spectral bandwidth. Reflection projection systems generally have higher throughput with less standing-wave interference than refraction systems because of their polychromatic characteristics. Catadioptric systems usually require more than just one optical axis and thus can lead to great difficulty in aligning the optical elements. Several systems of this type[12] have been developed as shown in the second category of Table 1 and in Fig. 9.

The Micralign®, made by Perkin Elmer, consists of only three optical components, as shown in Fig. 9*a*. It has a bandwidth of 400 nm. The aberrations vary as a function of the distance from the center and can be zero for a narrow zone about the center. Only a small zone of the spherical mirror is used, limited by the slit, providing nearly diffraction-limited imaging. The mask and the wafer are swept in unison through an arc to form an image of the whole mask. Figure 9*b* is the catadioptric system, also by Perkin Elmer, which is still widely used as a 1× full-scan system. The throughput is very high, as shown in Table 1, but the disadvantages of these systems are the low resolution and the difficulty of alignment during scanning; hence, the overlay accuracy is rather poor, as listed in Table 1.

Figure 9*c* is a catadioptric 1× stepper made by Ultratech (Table 1). Only the wafer has a step-and-repeat movement, and a die-by-die alignment is done. This system also has a very simple optical configuration, featuring a variable NA that depends on object field size—a low NA corresponds to a large field size. The practical resolution is 0.7 μm.

Figure 9*d* shows a 4× catadioptric reduction projector (Table 1) developed by Perkin Elmer that has 0.5-μm resolution. This machine also has a *ring field* similar to Fig. 9*a* and 9*b* with field-by-field alignment and a focus adjustment. The image field is larger than in I-line steppers.

Refraction projection

In refraction projection, the image of the mask is projected on the wafer through a high-resolution lens. A refraction projection system consists of a light source, a condenser, a heat-removing filter, beam-orienting mirrors, a shutter, and the mask and wafer stages. The refraction projection system is very similar to the proximity printers except for the imaging lens between the mask and the wafer. The optical stepper has become the most common lithography tool for ULSI production. The resolution is 0.4 μm using the I-line of the Hg lamp with the Kohler method, which images the exposure source through a condenser lens in an entrance pupil of the projection lens. However, the DOF is reduced to 1 μm because of the stepper's large NA and short wavelength. The features of the stepper are high resolution, and overlay accuracy as shown in Table 1.

The last group in Table 1 is an excimer laser stepper[13] whose wavelength is 248 nm. It can obtain the highest resolution of the equipment listed in Table 1 based on the excimer's shortest wavelength. The system configuration is almost the same as an I-line stepper, but the light source and the lens system are different. In the deep

Micron Ex. 1023, p. 309
Micron v. YMTC
IPR2024-00909

FIGURE 9
Examples of reflection projection printers. (*After Lin, Ref. 12.*)

Micron Ex. 1023, p. 310
Micron v. YMTC
IPR2024-00909

UV region, the materials of the optical components are limited to fused silica and several crystalline fluorides because of UV transmittance characteristics. The lenses are made entirely of fused silica and therefore are chromatic. The source bandwidth must be narrow to achieve sufficient resolution. The focus and the magnification differences due to wavelength difference are 0.15 μm/pm and 0.3 ppm/pm, respectively. The current performance of the KrF laser source, such as a 2.2-pm bandwidth and a ±0.25-pm wavelength stability, is considered adequate to meet the requirement for practical use. A 0.35-μm resolution within the 0.03-μm critical dimension spread is reported.[13] The use of the excimer laser is another possible way to fabricate ULSI patterns without resists by using the damaging effect of a short-wavelength light.

6.2.3 Enhancement

Many efforts have been made to solve optical lithography problems arising from wavelength or DOF. Among these, the phase-shifting mask and the off-axis illumination technique seem to be useful. The basic concept of the phase-shifting mask[14] is shown in Fig. 10. At the transmission mask, the electric field $\mathscr{E}$ has the same phase at every aperture (clear area) in Fig. 10*a*. Diffraction and the limited resolution of the optical system spread the electric field $\mathscr{E}$ at the wafer, as shown by the dotted

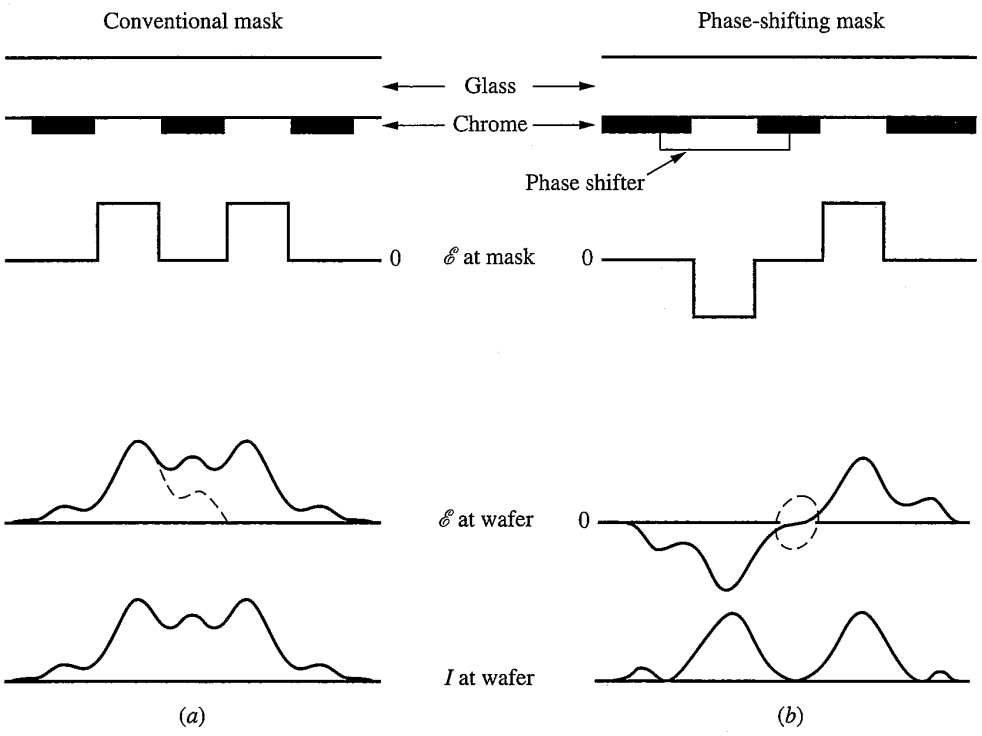


FIGURE 10
The principle of phase-shift technology. (*After Levenson et al., Ref. 14.* © 1992 IEEE.) (*a*) Conventional technology. (*b*) Phase-shifting technology.

Micron Ex. 1023, p. 311
Micron v. YMTC
IPR2024-00909

line. Interference between waves diffracted by the adjacent apertures enhances the field between them. The intensity I is proportional to the square of the electric field.

The phase-shifting layer that covers adjacent apertures reverses the sign of the electric field as shown in Fig. 10*b*. The intensity at the mask is unchanged. The electric field of these images at the wafer, shown by the dotted line, can be canceled. Consequently, images that are projected close to one another can be separated completely. A 180° phase change occurs when a transparent layer of thickness $d = \lambda/2(n-1)$, where n is the refraction index and λ is the wavelength, covers one aperture as shown in Fig. 10*b*.

The Fourier analysis[14] of phase-shifting and conventional transmission masks is explained in Fig. 11. The conventional resolution power of a projection system for transmission objects is the critical frequency $\nu_c = \text{NA}/\lambda$, which is related to the wavelength λ and numerical aperture NA and assumes coherent illumination. With incoherent illumination some spatial modulation is transmitted through such a system up to a frequency of $2\nu_c$; but for coherent illumination the modulation transmission function is unity up to ν_c and zero for $\nu > \nu_c$, as shown in Fig. 6. For the purpose of calculation it is convenient to define for the mask a transmission function that has sinusoidal intensity, as illustrated at the top of Fig. 11. In each case,

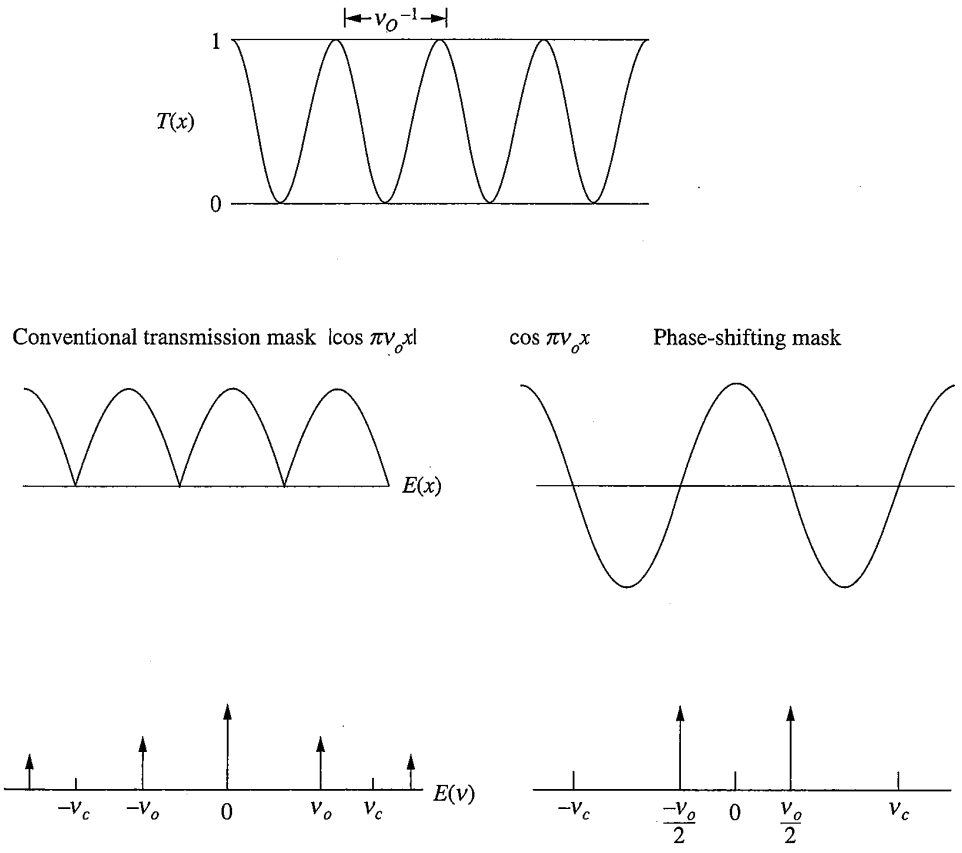


FIGURE 11
Fourier analysis of the optics of transmission and phase-shifting masks. (*After Levenson et al., Ref. 14.* © 1992 IEEE.)

Micron Ex. 1023, p. 312
Micron v. YMTC
IPR2024-00909

the intensity transmitted by the mask is assumed to be

$$T(x) = \tfrac{1}{2}(1 + \cos 2\pi\nu_0 x) \tag{6.10}$$

The electric field at the mask plane can have two possible forms when the illumination is normally incident and coherent. A transmission mask would produce an electric field profile $\mathscr{E}$ that is proportional to $|\cos \pi\nu_0 x|$, as in the left side of Fig. 11, whereas a phase-shifting mask would produce the electric field on the right, or $\mathscr{E} \propto \cos \pi\nu_0 x$. Both profiles can be Fourier-analyzed as shown at the bottom of Fig. 11. The transmission mask yields a Fourier spectrum of the form

$$\begin{aligned}\mathscr{E}_T(\nu) = \frac{4}{\pi}\Bigg\{ &\frac{1}{2}\delta(\nu) + \frac{1}{6}[\delta(\nu - \nu_0) + \delta(\nu + \nu_0)] \\ &- \frac{1}{30}[\delta(\nu - 2\nu_0) + \delta(\nu + 2\nu_0)] + \cdots\Bigg\}\end{aligned} \tag{6.11}$$

where $\delta(\nu)$ is the Dirac delta function and higher harmonics have been ignored. The phase-shifting mask projects only two Fourier components:

$$\mathscr{E}_\phi(\nu) = \tfrac{1}{2}[\delta(\nu - \nu_0/2) + \delta(\nu + \nu_0/2)] \tag{6.12}$$

The same information, in the sense of an optical intensity, is transmitted in each case, but the phase-shifting mask requires less spatial bandwidth than the transmission mask.

The optical system through which these Fourier components must propagate can be modeled as having a response function equal to unity for $\nu \leq \nu_c$ and zero for $\nu > \nu_c$, where ν_c is the cutoff frequency. Thus, the Fourier components of the electric field at the image plane are (except for a scale factor that can be set to unity) identical to those at the mask plane for $\nu < \nu_c$ and zero otherwise. From this treatment, the critical frequency for the transmission of modulation produced by a phase-shifting mask is $\nu_0 = 2\nu_c$, twice that of a transmission mask.

It is expected that an I-line stepper can resolve features down to 0.2 μm using a phase-shifting mask without reducing the DOF. The problem with this technology is the restriction on the pattern layout. It is very difficult to apply this technology to random patterns, although it is very useful in a periodic layout. Several ideas have been proposed[15] to solve the problem, as shown in Fig. 12. The halftone method, shown in Fig. 12*a*, has advantages such as not requiring extra data and ease of processing, but the resolution is not so good as the alternative method of Fig. 12*c*.

Phase-shifting masks are made by using an electron-beam (e-beam) direct-writing method. After the first layer is etched, the second layer is exposed on a newly coated resist, referenced to the alignment marks made in the first layer. In Fig. 12*b* or 12*c*, two kinds of patterns are necessary for each wafer layer. The problems of phase-shift technology are the design of the shifter pattern and the complexity of the mask process. Other key technologies required to make phase-shifting masks are improved inspection and repair technologies for the shifter patterns and the opaque patterns.

Off-axis illumination[16] is another useful method to enhance the resolution and DOF. The basic principle is to tilt the illumination to an angle that passes through both the undiffracted light and the first-order diffracted light symmetrically as shown

Micron Ex. 1023, p. 313
Micron v. YMTC
IPR2024-00909

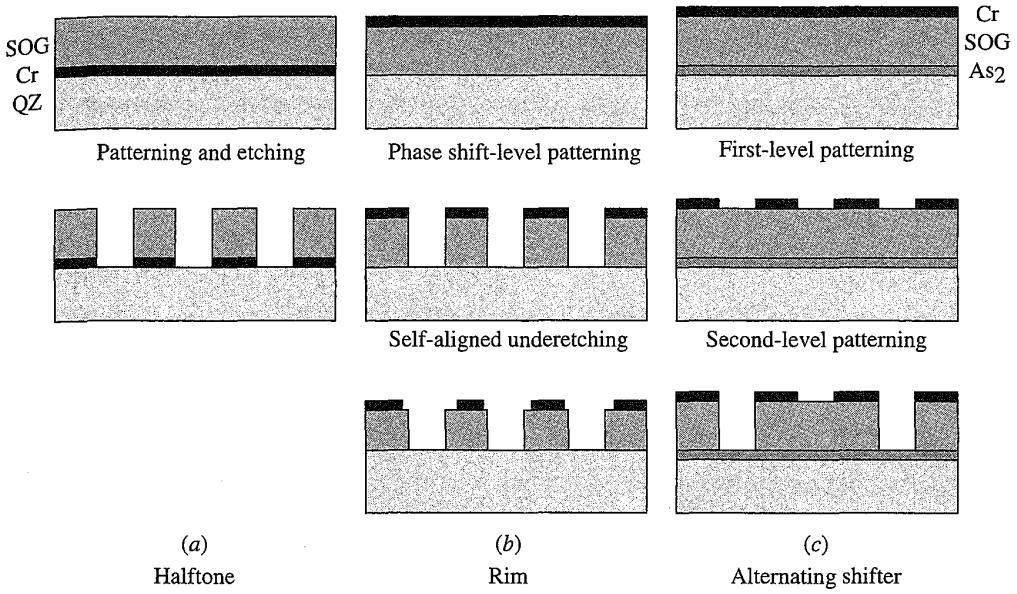


FIGURE 12
Examples of phase-shifting masks. (*After Ronse et al., Ref. 15.*) (*a*) Halftone. (*b*) Rim. (*c*) Alternating shifter.

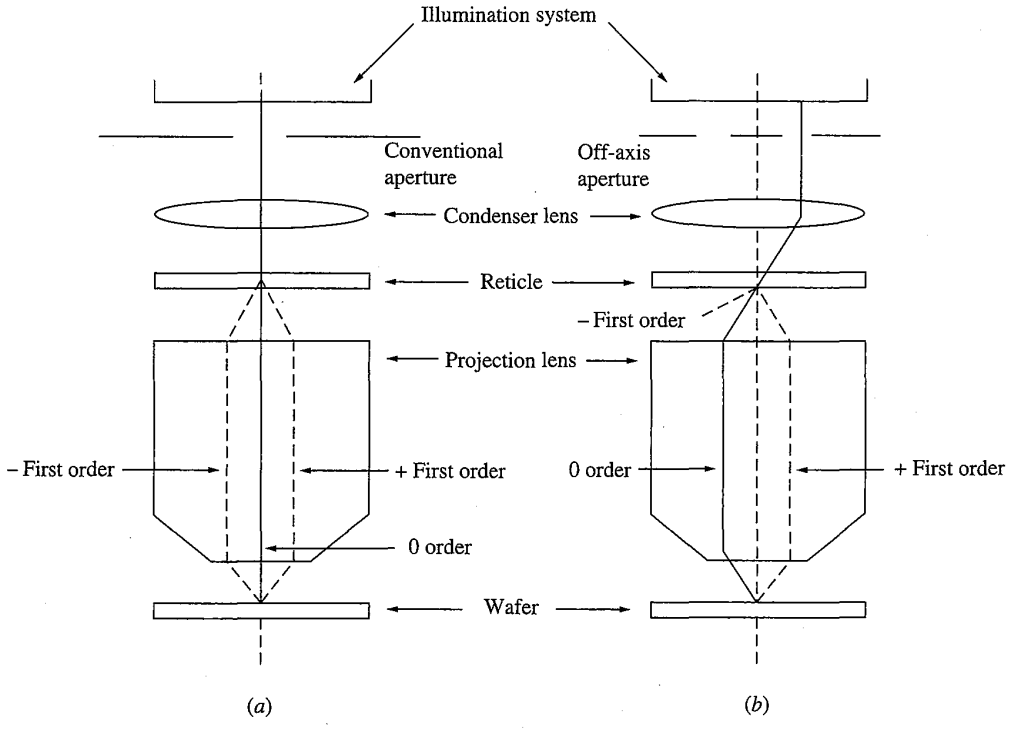


FIGURE 13
The off-axis illumination principle. (*After Shiraishi et al., Ref. 16.*) (*a*) Conventional illumination. (*b*) Off-axis illumination.

Micron Ex. 1023, p. 314
Micron v. YMTC
IPR2024-00909

in Fig. 13*b*, compared with the conventional method in Fig. 13*a*. The light beams transmitted through the off-axis aperture illuminate the reticle pattern with a specific incident angle ϕ, defined as $f \sin\phi = x$, where x is the distance between the optical axis and the transmissive portion of the aperture and f is the focal length of the condenser. The light beams are diffracted by the reticle pattern. The −first-order or higher-order diffracted light will not enter the projection lens because the pitch of the pattern is so fine and the diffraction angle ($\sin\theta$) is larger than the numerical aperture of the projection lens, as explained in Fig. 5. As a result, only zero- and first-order diffracted light will interfere on the wafer surface and contribute to the image formation. It is reported[16] that this technique has a line-and-space resolution of 0.27 μm and a 2.8-μm DOF for a 0.35-μm line-and-space pattern with a 1.18-μm resist thickness. It is necessary to adjust the incident angle depending on pattern spatial frequencies and to increase the exposure time to compensate for the loss of illumination due to tilting.

6.2.4 Overlay Accuracy

To this point we have emphasized resolution, but overlay accuracy is another fundamental technology component.[12] Overlay error is generated mainly in two process steps—the alignment and the exposure. Many types of errors are included in the exposure process, such as the relative placement error between the mask and the wafer, distortion or magnification error of the lens, and mechanical instability of the exposure system caused by vibration, temperature drift, or atmospheric pressure drift.

Alignment methods are classified into two types: off-axis and TTL (through the lens). Off-axis offers the possibility of using nonactinic rays and of broad-band and high-NA viewing, with the flexibility of brightfield, darkfield, or even phase-contrast viewing. It is also easy to upgrade the off-axis alignment system because of its total separation from the imaging optics. However, off-axis requires high mechanical stability and precision as well as accurate means to refer the positions of the mask, the wafer, and the alignment microscopes to each other.

TTL offers the inherent advantage of a direct mask-to-wafer reference and thus relaxes the requirement for high mechanical stability and precision. However, some accuracy of the direct reference is lost if nonactinic alignment is used or if the wafer has to be moved into the exposure position after alignment. Both of these process steps require compensation.

Brightfield alignment suffers from multiple interferences of the wafer film stacks in case of TTL narrow-bandwidth illumination. The alignment signal is sensitive to variation in film thickness, on the order of 10 nm. This problem is reduced by using broad-band illumination. Alternatively, darkfield alignment removes the dependence on film thickness but does not completely nullify the effect of uneven resist coverage. Also, darkfield alignment has the potential problem of signal degradation on a grainy substrate. The solution to this problem is using the slit beam and the periodic alignment marks. One way to solve the aforementioned problems is to remove the resist over the alignment marks in the positive resist process, but this is impossible in the case of multilayer resist and negative resist processes.

An alignment system[17] is illustrated in Fig. 14. Off-axis global alignment measures the relative position of the wafer with a He-Ne laser. Die-by-die alignment is

Micron Ex. 1023, p. 315
Micron v. YMTC
IPR2024-00909

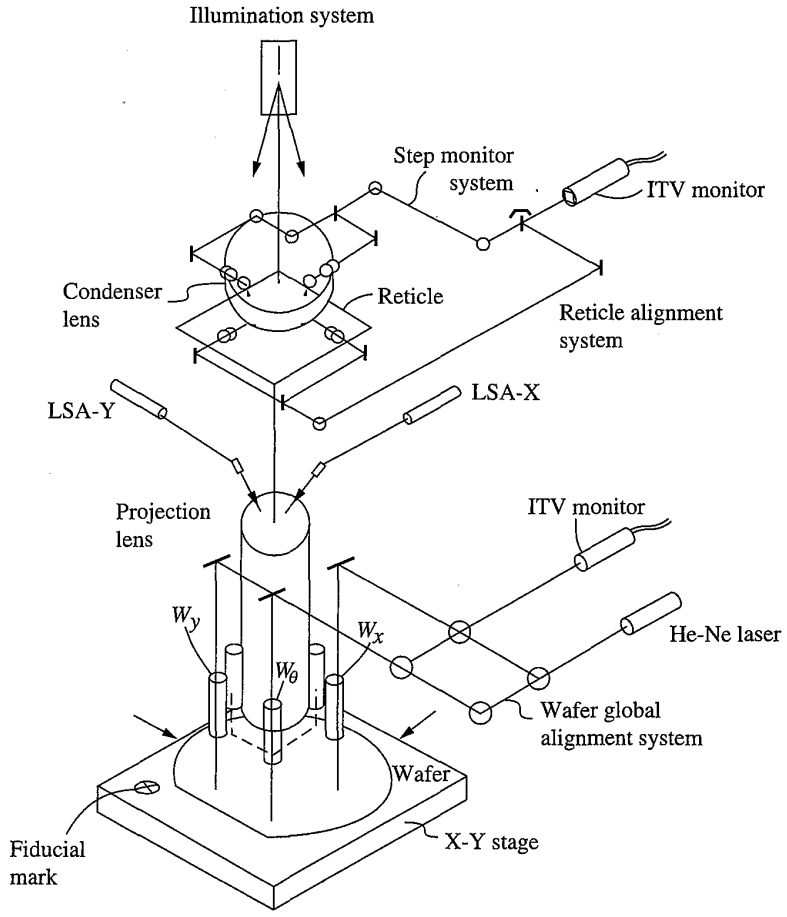


FIGURE 14
Wafer alignment system of an optical stepper. (*After Shiotake and Yoshida, Ref. 17.*)

more accurate because it uses an extra laser source (LSA-X, LSA-Y) through the projection optical path, i.e., the TTL method. Step-and-repeat exposure offers level-to-level registration, which is independent of wafer size, by separate alignment of each exposure field. In the reduction projection system, the mask pattern dimensions are larger than in systems imaged with unity magnification, and therefore it is more convenient to adjust the overlay position.

The basic sources of overlay error are listed in Table 2. The problem of overlay is common to both optical and x-ray lithography. X-ray lithography and e-beam direct write are considered distortionless. The registration error, defined as the ability to locate the alignment marks, is assumed to be identical for all optical alignment systems. Resist-induced alignment error is considered to be systematically removable. Wafer inplane distortion is caused by various hot processes. The major part of the magnification error of reduction systems is eliminated by fine tuning, but not easily, owing to the inherent symmetry of 1× optical systems. Coupled with the linear part of wafer inplane distortion, the magnification error of 1× optical systems can be as large as 0.15 μm. Similarly, x-ray systems have an error of 0.1 μm as a result

Micron Ex. 1023, p. 316
Micron v. YMTC
IPR2024-00909

TABLE 2
Sources of overlay errors[12]

Items	5× optics (μm)	1× optics (μm)	1× x-ray (μm)	E-beam direct write (μm)
1. Lens distortion	0.07	0.06	0	0
2. Registration error	0.05	0.05	0.05	0.05
3. Wafer inplane distortion	0.05	0.05	0.05	0.05
4. Magnification error	0.01	0.1	0.1	0.01
5. Mask-to-mask placement error	0.05	0.05	0.05	0.05
6. Total contributions (RSS)	0.1	0.15	0.13	0.09
7. Total without lens distortion	0.07	0.13	0.13	0.09
8. Total without magnification error	0.1	0.11	0.09	0.09
9. Total without distortion or magnification error	0.07	0.09	0.09	0.09

of wafer inplane distortion and mask heating during exposure. The mask-to-mask placement error reflects the accuracy of e-beam mask making. For x-ray systems, the placement error of the membrane mask can be worse than for the optical mask, but it is not included here. In the case of e-beam direct write, the mask error is simply the composite placement error of the e-beam between two masking levels. In the case of 5× lithography, the mask-to-mask placement error is divided by 5 before being included in the RSS (square root of the sum of the squares) total of the contributions.

6.2.5 Optical Resists

Photoresist is the general term for polymers that can create patterns by the use of solvents after irradiation. The development of the resist is based on a chemical reaction and depends on the solubility difference between irradiated and unirradiated areas. Photoresists are of two types: negative, which on exposure to light become less soluble in a developer solution, and positive, which become more soluble.

Negative resists generally consist of a chemically inert polyisoprene rubber, which is the film-forming component, and a photoactive agent. The photoactive agent releases nitrogen gas on exposure to light, and the radicals generated react with the double bonds to form cross-links between rubber molecules, making the rubber less soluble in an organic developer solvent. The reactive species formed during the exposure can react with oxygen and be rendered ineffective for cross-linking, so an inert atmosphere is used. The developer solvent dissolves the unexposed resist. The exposed resist has low molecular weight, so it swells as the uncrosslinked molecules are dissolved. The swelling distorts the pattern features and limits resolution to 2 to 3 times the initial film thickness. For VLSI/ULSI applications, the use of negative resists has been supplanted by positive resists because negative resists have a resolution limit of about 2 μm, although they have many advantages such as resistance to etching and good adhesion to the substrate.

Positive resists have two components: a resin and a photoactive compound dissolved in a solvent. The photoactive compound is a dissolution inhibitor. When the photoactive compound is destroyed by exposure to light, the resin becomes more

Micron Ex. 1023, p. 317
Micron v. YMTC
IPR2024-00909

soluble in an aqueous developer solution. The unexposed regions have high molecular weight and so do not swell much in the developer solution; therefore higher resolution, suitable for ULSI, is possible with positive resists. The development process of projection-printed images in positive resists has been modeled theoretically;[18] it is an isotropic etching or removal process.

One of the fundamental properties of a resist, which determines resolution, is the contrast.[19] Contrast, in the case of negative resist, is related to the rate of formation of a cross-linked network and may be measured by exposing pads of known area to varying radiation doses. D_g^i is the minimum dosage required to form the first insoluble film. Thereafter, the film thickness, and consequently the thickness remaining after development, increases with increasing dose until, ultimately, the thickness of the developed pad is not detectably different from the original film thickness (in practice it is slightly less owing to volume contraction during cross-linking). This process is shown schematically in Fig. 15*a*. The resist contrast γ_n is then defined as

$$\gamma_n = 1/(\log D_g^0 - \log D_g^i) \tag{6.13}$$

For a positive resist, the film thickness of the irradiated region that remains after development decreases with increasing dose until eventually a dose D_c is reached, which results in complete removal of the film on development, as shown in Fig. 15*b*. The contrast of positive resist γ_p is defined as

$$\gamma_p = 1/(\log D_c - \log D_0) \tag{6.14}$$

When photoresist films are exposed using monochromatic radiation, standing waves are formed in the resist.[20] These are caused by coherent interference effects due to a reflecting substrate. Coherent interference results in periodic intensity distributions in the direction perpendicular to the plane of the resist with a period $\lambda_r/2$, where λ_r is the wavelength in the resist. The refractive index of most resists is about 1.6, leading to an optical mismatch at the air/resist interface. Figure 16 shows

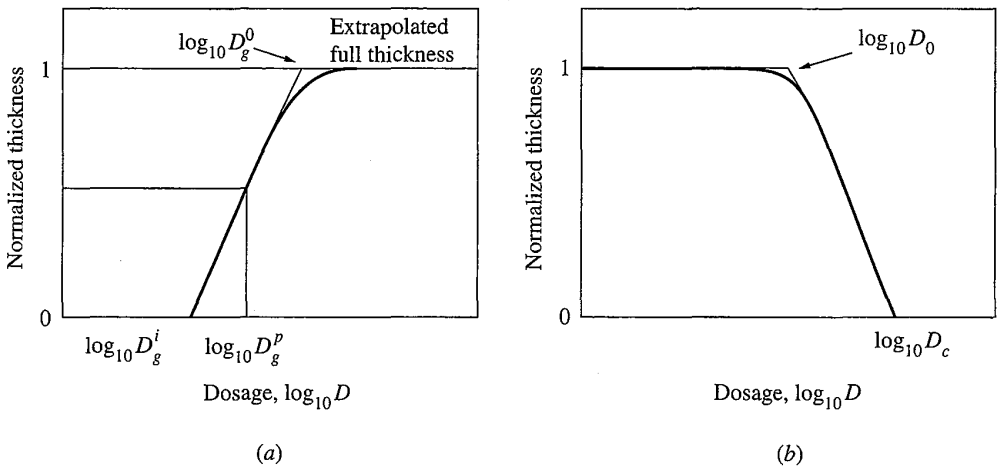


FIGURE 15
Sensitivity curves of negative and positive resists. (*After Thompson, Ref. 19.*) (*a*) Negative resist. (*b*) Positive resist.

Micron Ex. 1023, p. 318
Micron v. YMTC
IPR2024-00909

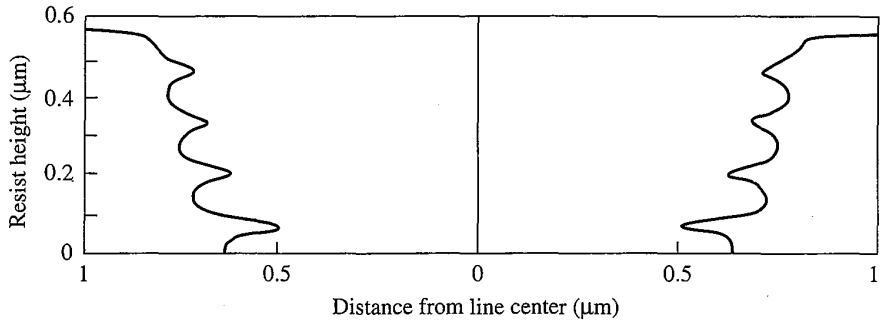


FIGURE 16
Simulated standing wave edge profile for a nominal 1-μm line in AZ1350 photoresist developed for 85 sec in 1 : 1 AZ developer : water. (*After Dill et al., Ref. 20.* © 1975 IEEE.) Substrate: Si + SiO_2, 60 nm; resist thickness: 583.6 nm; exposure wavelength: 435.8 nm (G-line); exposure energy: 57 mJ/cm^2

the standing wave. This variation in peak intensity with resist thickness becomes less with decreasing substrate reflectivity and increasing absorption by the resist. Since resist thickness varies at a step in the substrate topography, the resulting difference in effective exposure leads to size variations in the resist image.

Commercially available resists are listed in Table 3. The negative RD2000N-type resists, which have a diazide compound as a reactor with polyvinylphenol resin, have a resolution similar to a positive resist and do not swell on development. Photoresists are being developed for exposure at shorter wavelengths where higher resolution is possible. A few such deep-UV resists are PMMA, sensitive for $\lambda < 250$ nm; polybutene sulfone, sensitive for $\lambda < 200$ nm; and Microposit MP-2400, sensitive for $\lambda = 250$ nm. At these shorter wavelengths, the radiation quantum is large enough to produce scission of the molecular chain.

The other very important resists for deep UV are the chemically amplified resists, which exhibit high photo-speed, excellent resolution, and process tolerance.[21] The chemically amplified resist consists of a polymer host and a generator of acid in the presence of light. In the postexposure bake the photogenerated acid catalyzes thermal reactions that alter the solubility of the exposed region. Generally, resolution and sensitivity are affected by postexposure bake conditions. Some chemically amplified resists are applicable to electron lithography.

Other important properties of resists are adhesion to the substrate and resistance to wet and dry etch processing. In general, the commercially available optical resists are compatible with such processes.

Some inorganic materials are also effective as resists. Fundamental functions are based on the Ag-doping effect, which is induced by photoirradiation or electron-beam irradiation on a stacked-layer system of Ag and Se-Ge chalcogenide-glass film. When doped with Ag, the chalcogenide-glass films become almost insoluble in an alkaline solution. The light source, such as a Hg lamp, should have a wavelength that is shorter than the absorption edge of Se-Ge. This method is expected to have high resolution, with applications in dry processing.[22,23]

Micron Ex. 1023, p. 319
Micron v. YMTC
IPR2024-00909

TABLE 3
Commercial optical resists

Name	Maker	Type	Chemical structure	Remarks
OMR83	Tokyo-oka	Negative	Isoprene rubber, azide compound	1× aligner
CIR	Nippon Synthetic Rubber	"	"	"
RD-2000N	Hitachi Chemical	"	Phenol resin, azide compound	Deep UV
MES-U	Nippon Synthetic Rubber	"	Chloropolystyrene	"
THMR-iN200	Tokyo-oka	"	Novolac resin, azide	I-line
TDUR-N7	"	"	Phenol resin	Deep UV
Microposit 1350J	Shipley	Positive	Novolac resin, O-kyniadiazide	G-line
AZ1350J	Hoechst	"	"	"
OFPR800	Tokyo-oka	"	"	"
Microposit2400	Shipley	"	"	Deep UV
ODUR-120	Tokyo-oka	"	PMMA	"
ODUR1000	"	"	PMIPK	I-line
THMR-IP3000	"	"	Novolac resin, azide	"
TSCR-80I	"	"	"	"
PFI28	Sumitomo Chemical	"	"	"
PFRIX060	Nippon Synthetic Rubber	"	"	"
PFRIX061	"	"	"	"
FHI-3950	Fuji Hunt	"	"	"
AZ7500	Hoechst	"	"	"
NPR-A18SH7	Nagase	"	"	"

6.2.6 Process Technologies

Increasing the NA and decreasing the wavelength are logical approaches to improving resolution. However, if these techniques are used, DOF decreases, based on Eq. (6.3). In the actual production process, many possible causes of interference to focusing arise. Some occur during the wafer process, such as surface topography and warping. Others are due to the exposure process, such as vibration, illumination nonuniformity, and aberrations. One solution to these problems is the combination of planarization and surface imaging. The thick bottom planarization layer reduces optical reflection and pattern-width degradation due to the surface topography beneath. The thin top layer enhances resolution, pattern-width control, and process latitude.

Surface-imaging resist processes are summarized in Fig. 17. The first item is treatment through a development cycle to enhance the contrast of the resist.[24] In this cycle the wafer is dipped in the developer, rinsed, and dried. This sequence is repeated until the resist film of the exposed area is completely dissolved. This interrupted development process promotes the formation of a passivation layer on the

Micron Ex. 1023, p. 320
Micron v. YMTC
IPR2024-00909

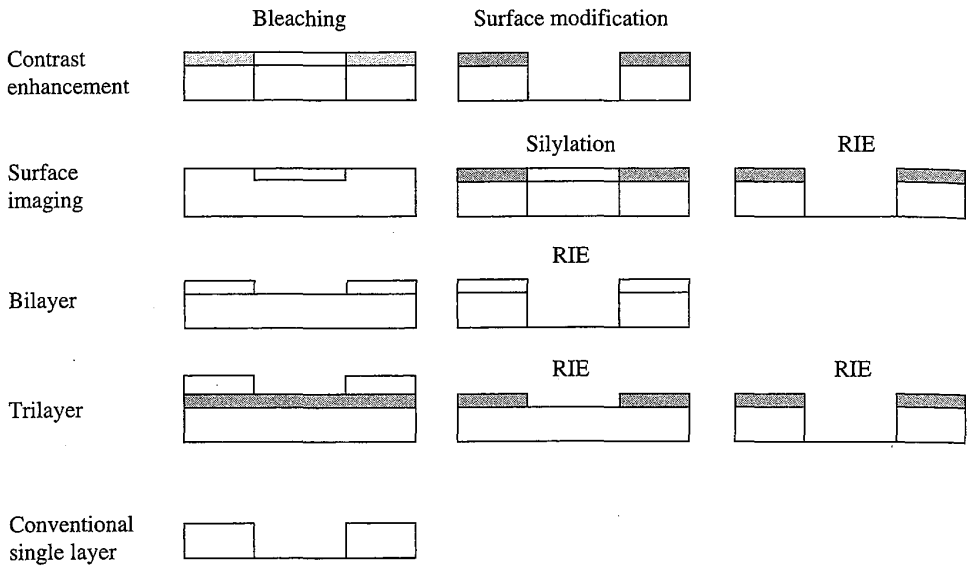


FIGURE 17
Advanced resist processes.

unexposed surface. As a result, the side walls become vertical and sensitivity is improved 30 to 50% in AZ resist. This technology is also used in electron lithography.

The second process in Fig. 17 is silylation.[25] A single layer of PLASMASK® (UCB from Belgium) resist is spincoated onto the substrate, with a thickness from 1.5 to 2.5 μm, and prebaked to a self-planarizing layer. After exposure, the wafers are treated with a gas-phase silylating agent (hexamethyldisilazane) at elevated temperature to reduce the silylation time. The resist material of the exposed areas selectively bonds chemically with silicon to a depth of 100–200 nm, and remains stable for a long time. The wafers are developed in oxygen plasma. During this treatment, the silicon is converted into silicon dioxide, which forms a thin protective mask that stops the etching of these exposed areas. It is reported that 0.35-μm features can be achieved using a 248-nm source.[26]

In the bilayer method, the image of the top layer formed by exposure and development is transferred to the bottom planarization layer by reactive ion etching (RIE). Resolution better than 0.5 μm is achieved by using a resist containing tungsten, which has a high resistance to etching, as the top layer.[26] There are some disadvantages in the bilayer method, such as intermixing between the resist layers and the formation of cracks while different materials are coated and baked.

The trilayer resist structure typically consists of an inorganic interlayer that is 0.1 to 0.2 μm thick, sandwiched between a top organic imaging layer that is 0.3 to 0.4 μm thick and a bottom organic planarization layer that is 1 to 2 μm thick. The trilayer method is more complex than the bilayer but is more flexible because of the intermediate layer.[27,28] Because of the intermediate barrier, the processes for the top layer and for the bottom layer are independently selectable, so many combinations of processes can be achieved by the proper choice of the intermediate layer materials and the top layer resists. The thickness of the top layer can be minimized to get high

Micron Ex. 1023, p. 321
Micron v. YMTC
IPR2024-00909

resolution with enough etching duration for the intermediate layer (usually oxide). This method is also used in electron lithography.

6.3 ELECTRON LITHOGRAPHY

Electron lithography has the possibility of higher resolution than optical lithography because of the small wavelength (less than 1 angstrom) of the 10–50 keV electrons. Resolution in electron lithography systems is not limited by diffraction but by electron scattering in the target materials including the resist and by the various aberrations of the electron optics. Scanning electron beam pattern generators have been under development for more than 20 years and were derived from the scanning electron microscope.[29,30] Because of the serial nature of the pattern writing, throughput is much lower than for optical systems. However, a wide variety of applications is available in the pattern-generating function for electron lithography, such as mask fabrication for optical or x-ray lithography, direct writing on the wafers, and direct reaction with some materials on the substrate. Electron lithography is classified into two types—scanning and projection—and the scanning type can be either raster or vector scanning.

The exposure scheme is illustrated in Fig. 18. The ULSI pattern is composed by a computer-aided design (CAD) system. The output format from the CAD system is converted into the internal format of the individual exposure systems. The electron

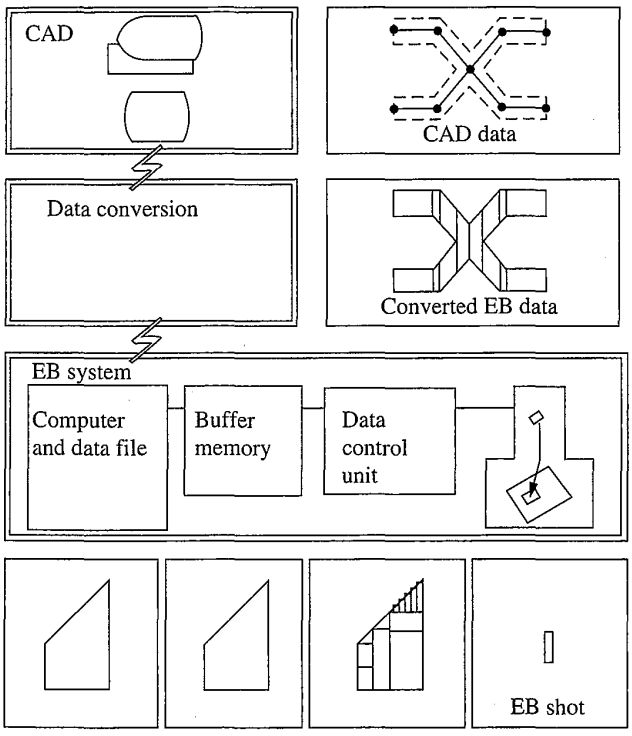


FIGURE 18
Flow of data in electron beam exposure systems.

Micron Ex. 1023, p. 322
Micron v. YMTC
IPR2024-00909

exposure machine decomposes the data into simple elements (trapezoids or rectangles, depending on the machines) to control the electron beam irradiation. Electron beam exposure machines are bigger and more complicated than the optical printers because of their data-handling function.

Although electron projection and proximity-printing methods have been developed for full image exposure and are expected to be one of the alternatives to optical technology after an optical resolution limit is reached, these techniques have not become practical yet. Among many kinds of electron beam exposure machines, the MEBES® from Perkin Elmer has proven to be the best photomask pattern generator. Some special products, such as microwave transistors, have for many years been manufactured by direct writing. In some low-volume production devices, direct writing has also been applied in the development stages of ULSIs and in personalization of application-specific integrated circuits (ASICs) to minimize the mask-related cost and development time. Commercially available electron-beam machines are listed in Table 4. Generally, Gaussian spot vector machines are used for research and development of small-geometry devices, and the variable-shaped or raster spot machines for production. About 450 machines have been installed and used in the world. The last entry of Table 4 is included because it functions much like an electron-beam exposure machine but uses an optical system.

6.3.1 Electron Optics

Figure 19 shows the basic electron optical system. Magnetic lenses form a demagnified image of the source d_o on the image plane d_i. The position of an electron beam, irradiated on the substrate, is controlled by deflectors. The on-off control of an electron beam is carried out by a blanking plate with an aperture underneath.

The characteristics of some cathode materials are shown in Table 5. The one used most commonly is the LaB_6 single crystal, which has a long life, high stability, and high current capability compared to the conventional tungsten hairpin cathode. The thermal field emission (TFE) source consists of a tungsten tip with a radius of 0.5 to 1 μm heated sufficiently (about 1900 K) to emit high current with high stability.

Emission current density from the cathode J_c is given by

$$J_c = AT^2 \exp(-E_w/kT) \tag{6.15}$$

where A is the Richardson constant, k the Boltzmann constant, and E_w the work function of the cathode material. The electrons are accelerated by the voltage V (10 to 50 kV) and focused by the gun at a point near the anode called the *crossover,* which has a diameter d_o of 10 to 100 μm. The configuration of the electron gun is illustrated schematically in Fig. 20*a*. The maximum value of the brightness is

$$\beta = J_c eV/\pi kT \tag{6.16}$$

The lenses in the electron optical column (Fig. 19) are magnetic, and their structure is illustrated in Fig. 20*b*. If a parallel beam of radius r_o enters the field **B** of the lens, the electrons experience a force that causes those not on the axis to revolve

Micron Ex. 1023, p. 323
Micron v. YMTC
IPR2024-00909

TABLE 4
Commercially available electron-beam exposure systems and a laser-beam pattern generator

Model	Maker	Type	Optics	Acceleration voltage (kV)	Beam size (μm)	Resolution (μm)	Overlay accuracy (μm)
MEBES4	ETEC	Raster/continuous	TFE, Gaussian	10	0.05	0.2	0.4
AEBLE150	"	Vector/continuous	LaB_6, variable shape	20	~2	0.5	0.15
EBPG4	Leica	Vector/step-repeat	LaB_6, Gaussian	50	0.01~	0.03	0.08
EBML300	"	"	"	50	0.025	0.1	"
JBX7000MV	JEOL	"	LaB_6, variable shape	20	~4	0.5	0.05
JBX8000DX	"	"	"	20	"	0.2	
JBX-5DII	"	"	LaB_6, Gaussian	25/50	0.008~	<0.1	0.06
HL700M/D	Hitachi	"	LaB_6, variable shape	30	~4	0.2	0.06
HL700F	"	"	TFE, Gaussian	30	0.03	0.03	0.1
CORE2564	ETEC	Raster/continuous	364 nm laser	Optical system		0.6	0.05

Micron Ex. 1023, p. 324
Micron v. YMTC
IPR2024-00909

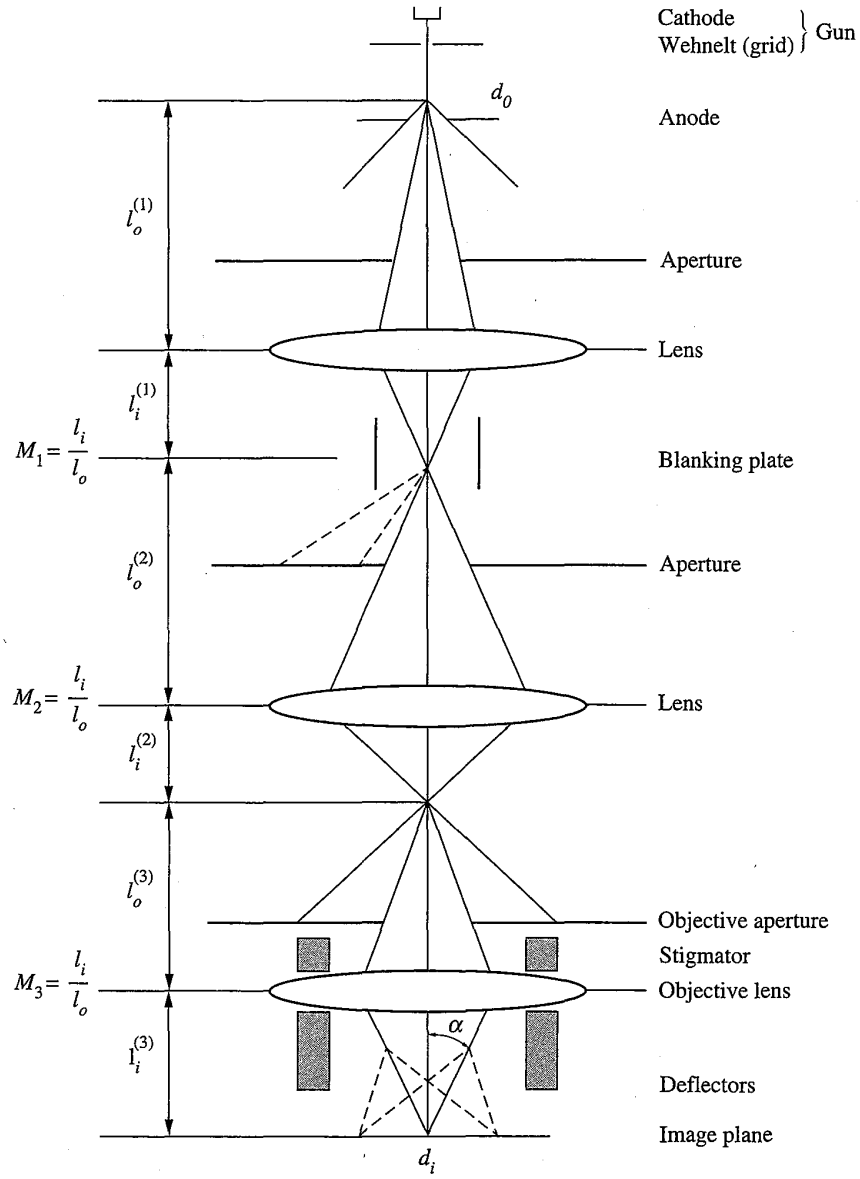


FIGURE 19
Basic electron optical column.

about the axis and turn toward it. For a thin lens, the electron path beyond the lens is given by

$$\frac{d_r}{d_z} \approx -\frac{r_o e}{8mV}\int_{-\infty}^{+\infty} B_z^2 dz \qquad \frac{d_r}{d_z} = \frac{-r_o}{f} \tag{6.17}$$

Micron Ex. 1023, p. 325
Micron v. YMTC
IPR2024-00909

TABLE 5
Characteristics of cathodes used in electron scanning systems

Materials	Operating temperature (K)	Energy dispersion (eV)	Current density (μA/sr)	Source size (Å)	Brightness A/cm^{2}sr	Drift (%/hr)	Life (hr)
W hairpin	2700	4.5	10^4	50×10^4	10^5	< 1	100
LaB_6	1800	2.7	10^5	10×10^4	10^6	< 1	~5000
W(310)-FE	Room temp.	0.3	50	30	10^8	30	~5000
ZrW-TFE	1900	3.1	10^3	100	10^9	< 1	~5000

where r_o = beam radius
e = electronic charge
m = mass of electron
B_z = axial component of **B**
f = focal length
V = acceleration voltage

Many rules of light optics apply here also, such as

$$1/l_o + 1/l_i = 1/f \qquad M = l_i/l_o \tag{6.18}$$

where l_o = object distance
l_i = image distance
M = magnification

The current density J and current I on the image plane are

$$J \leq \pi\beta\alpha^2 \qquad \text{and} \qquad I = J(\pi d_i^2/4) \tag{6.19}$$

where β = gun brightness
d_i = diameter of the electron beam on the image plane

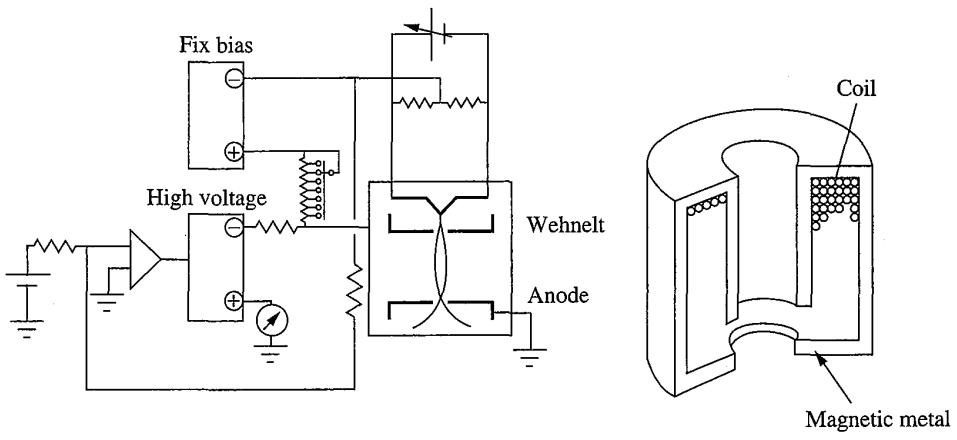


FIGURE 20
Configuration of the electron gun and magnetic lens. (*a*) Configuration of electron gun and its power supply. (*b*) Magnetic lens structure.

Micron Ex. 1023, p. 326
Micron v. YMTC
IPR2024-00909

Spot sizes of d_i that are of interest range from 0.01 to 0.1 μm. From Section 6.2.2 the diameter d of the central spot of the Airy pattern is $1.2\lambda/\alpha$. NA($\approx \sin\alpha$, where 2α is the solid angle of the ray reaching the imaging point from the objective lens) is the numerical aperture. For 15-keV electrons, the wavelength $\lambda = 0.1$Å. Taking $\alpha = 10^{-2}$ radians, we have a diffraction spot width

$$d_{\text{diff}} = 1.2\lambda/\alpha = 10^{-3}\mu\text{m} \ll d \tag{6.20}$$

Thus, diffraction can be ignored. However, aberrations in the final lens and in the deflection system will increase the spot size and can change its shape as well. The actual spot size in the presence of aberrations is the square root of the sum of the squares of the contributions of the independent aberrations. The aberrations are of two types: undeflected beam aberrations and aberrations that are functions of the distance r, in the image plane from the axis, to the position of the deflected beam. The aberrations of the first type are

$$\begin{aligned} \text{Spherical aberration: } d_s &= \tfrac{1}{2}C_s\alpha^3 \\ \text{Chromatic aberration: } d_c &= C_c(\Delta V/V)\alpha \end{aligned} \tag{6.21}$$

Spherical aberration is a result of focusing rays passing through the lens at different distances from the axis. The spread of electron energies, $e\Delta V$, is usually a few eV. C_s and C_c are constants determined by the lens parameters. Astigmatism, a third type of aberration, results from the breaking of cylindrical symmetry in the column and can be removed by introducing compensating fields with *stigmator* coils as shown in Fig. 19. Deflection aberrations, such as coma and field curvature, are proportional to $r^2\alpha$ or $r\alpha^2$. Another source of spot broadening is the mutual Coulomb repulsion of the electrons as they traverse the column. For total column length L (cm) and beam current I (A) this contribution is given approximately by

$$d_{\text{ee}} \approx [LI/(\alpha V^{3/2})] \times 10^8\mu\text{m} \tag{6.22}$$

This spreading becomes important only at large currents when small spot sizes or sharp edges of large spots are required. For example, if $I = 500$ nA, $L = 70$ cm, $V = 20$ kV, and $\alpha = 10^{-2}$ rad, then Eq. (6.22) gives $d_{\text{ee}} = 0.12$ μm.

Note that the distortion, which causes the positioning error of the electron beam landing on the substrate, depends on the deflection distance from the center. The actual amount of error depends on the condition of the electron optics, such as the manufacturing accuracy of the electron optical parts or deflection range, but an approximate value is on the order of 10^{-3} on the deflection edge. An electron lithography machine has many correction functions that differ from those in optical systems because of its sequential exposure method and the real-time response characteristic of the electron beam. These functions include corrections for distortion, beam drift, stage position, beam size related to deflection distance or substrate height, and beam unblanking time due to beam current drift.

6.3.2 Raster Scan and Vector Scan

The EBES® machine, developed by AT&T Bell Laboratories, and the commercially available MEBES® system from ETEC (Hayward, CA) use beam deflection in one

Micron Ex. 1023, p. 327
Micron v. YMTC
IPR2024-00909

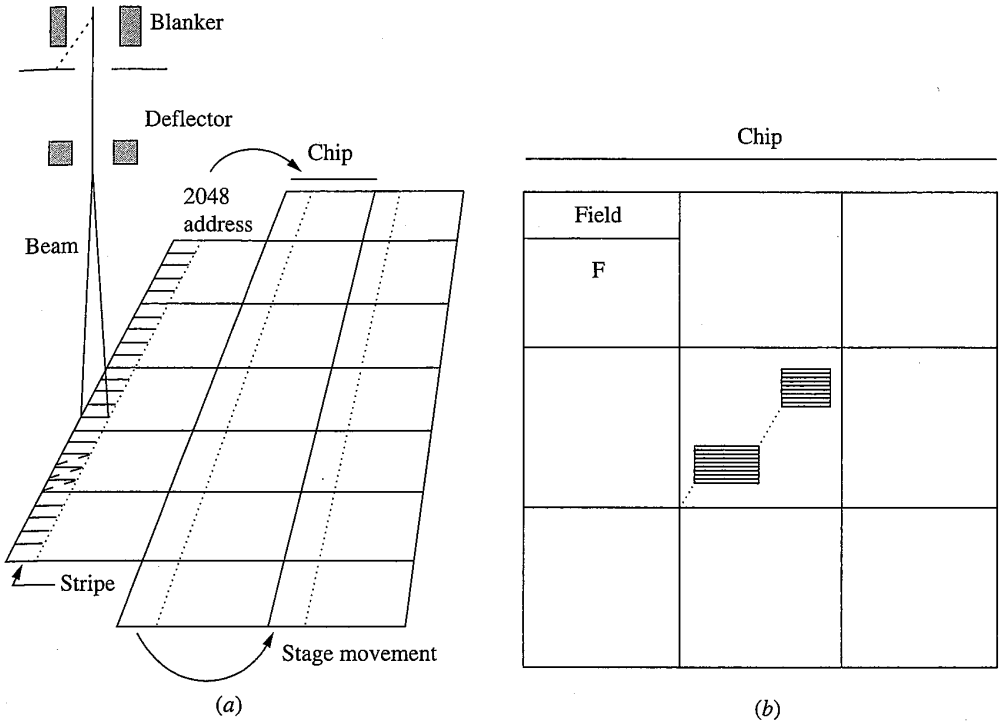


FIGURE 21
The principles of raster and vector scan. (*a*) Raster scan. (*b*) Vector scan.

dimension (mainly).[31] The writing scheme is shown in Fig. 21*a*. The stage moves continuously in a direction perpendicular to the beam scanning direction. The electron beam is moved repetitively in one direction while it is slowly deflected in the perpendicular direction, corresponding to the stage movement. The pattern data are decomposed into a number of stripes parallel to the stage movement, and one stripe is written on all chips of the same type before the next stripe is begun. The stripe is decomposed into many scan lines whose widths are basically similar to the feature addressing unit. The electron beam is irradiated on the substrate by controlling the blanker that turns the beam on at the exposure starting points and turns it off at the end point during rastering. The performance of MEBES4 is given in Table 4.

In vector scan the beam is directed sequentially to the parts of the chip pattern to be exposed.[32] The pattern is decomposed into a number of elements (rectangles, triangles, etc.) and each is exposed by the writing beam. Many vector scan machines expose in a step-and-repeat fashion. Figure 21*b* shows an exposure field of dimensions $F \times F$, which is the electron beam deflection area. The dotted lines indicate where the beam is turned off (deflected by the blanking plates), similar to blanking in a raster scan. After all elements in one field are exposed, the stage is stepped to the next field and the exposure process is repeated. The stage need not provide highly precise positioning. Stage position is monitored by a laser interferometer, and small differences from the desired stage location are compensated by small offsets of the beam. If a chip is larger than an exposure field, as in Fig. 21*b*, several fields can be used to expose the chip. The scan field must be as large as allowed by deflection

Micron Ex. 1023, p. 328
Micron v. YMTC
IPR2024-00909

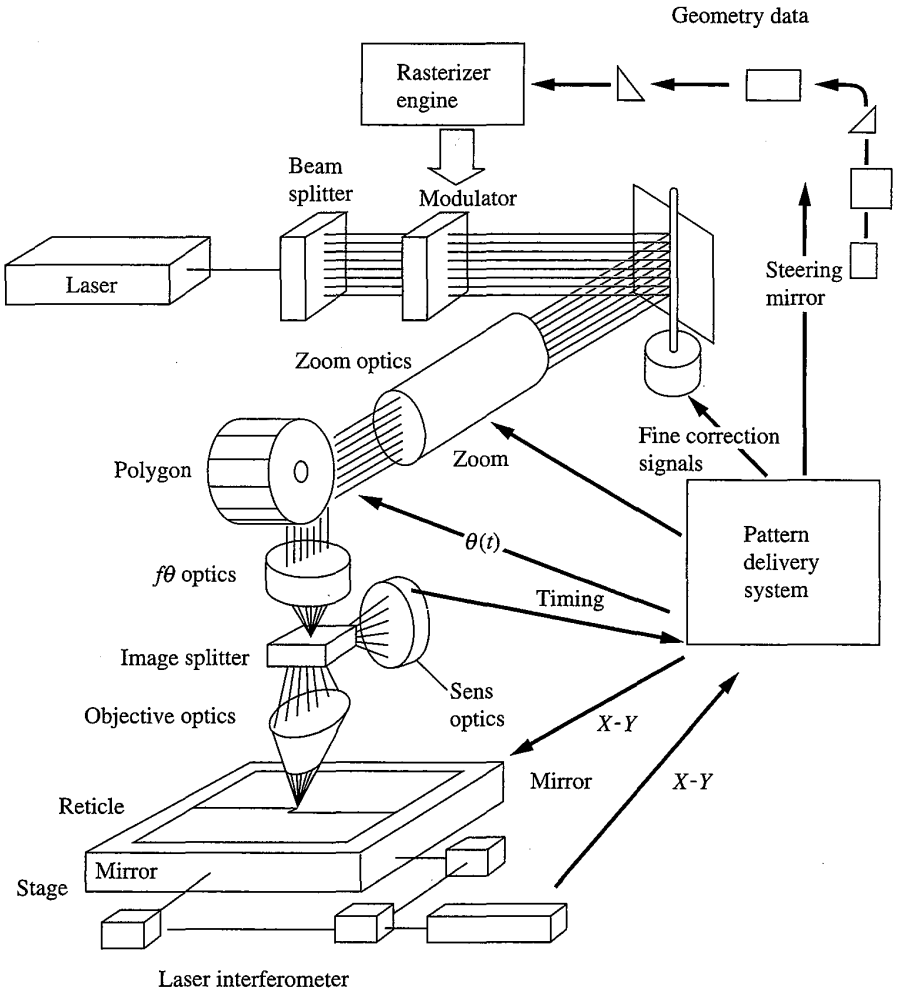


FIGURE 22
System configuration of a laser reticle generator. (*After Burns and Schoeffel, Ref. 34.*)

positioning accuracy because stage-stepping time affects throughput. There are many vector-type machines commercially available, as shown in Table 4.

The EBES4 machine developed by AT&T Bell Laboratories has a continuously moving stage, three hierarchical deflection systems, and a high-brightness gun that uses a thermal field emission cathode in order to manufacture deep-submicron-geometry (less than about 0.5 μm) masks or wafers within an acceptable exposure time.[33]

The CORE® laser reticle writer, the last entry in Table 4, is an optical pattern generator.[34] The writing concept is very similar to the MEBES® electron beam machine. A narrow laser beam is scanned very accurately by the combination of a polygon mirror, a piezo-controlled steering mirror, and a high-frequency beam modulator (on-off controller) on the continuous-movement mask stage. The system configuration is illustrated in Fig. 22. This machine can use the same process as with wafer fabrication for mask making because it is an optical method. It also has the capability of direct writing and phase-shifting mask making.

Micron Ex. 1023, p. 329
Micron v. YMTC
IPR2024-00909

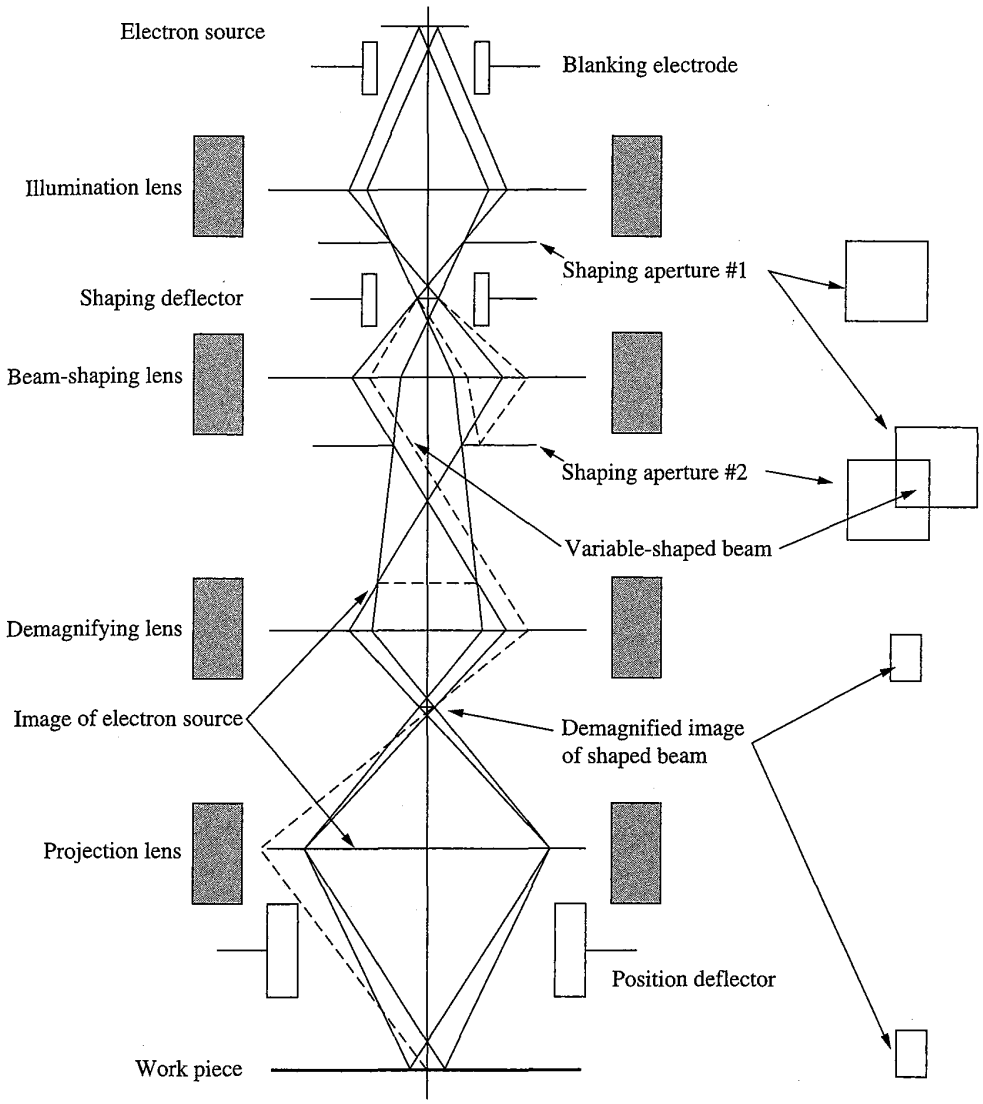


FIGURE 23
Variable-shaped-beam column of JEOL model JBX-6A3.

6.3.3 Variable Shaping

For vector scan machines the exposure time of a specified area is proportional to the number of electron beam irradiations if the current density and the resist sensitivity are fixed. The *shaped-beam* method offers a way to project many image points in parallel and to achieve faster exposure by reducing the irradiation numbers.[35,36] In the JEOL JBX-6A3 machine shown schematically in Fig. 23, the shape of the electron beam is rectangular, and the size is variable. The image of the first square aperture is shifted in two dimensions to cover various portions of the second aperture, which is the actual object of the image projected on the mask or wafer. The minimum and

Micron Ex. 1023, p. 330
Micron v. YMTC
IPR2024-00909

maximum rectangle widths are 0.2 μm and 6 μm, respectively. Maximum current density is 2A/cm^2 with the LaB_6 cathode. Electrostatic beam deflection is employed.

Another interesting approach is the combination of a variable shaped beam with continuous stage movement technology to minimize stage movement overhead. The first commercially available machine is AEBLE-150 from ETEC.[37] It has static and magnetic two-stage deflectors. The magnetic deflector decides the static deflection positions, and the static deflections control the electron-beam exposure positions corresponding to the stage positions. The stage is continuously moved like a raster machine, and the position, monitored by laser interferometer, is fed back to both deflectors. Machines developed by NTT[38] and Toshiba[39] use a similar concept.

6.3.4 Electron Proximity/Projection Printing

An electron proximity printing system[40] is illustrated in Fig. 24*a*. It needs step-and-repeat stage movement because of its small mask size and thus differs from an optical proximity printer. The distance between the mask and the wafer is 0.6 mm. The beam, with a diameter of approximately 1 mm, is scanned over the full chip-size mask, which is made on a silicon membrane. To permit electrons to pass through the transparent parts of the mask physical holes must be present, unlike an x-ray mask. Creating isolated opaque patterns such as the interiors of doughnut shapes is a problem because they have no attachment to the rest of the mask. A solution to this problem is to use a set of two complementary (half) masks. Registration is by reference to alignment marks on each chip. Resolution is reported as 0.3 μm.

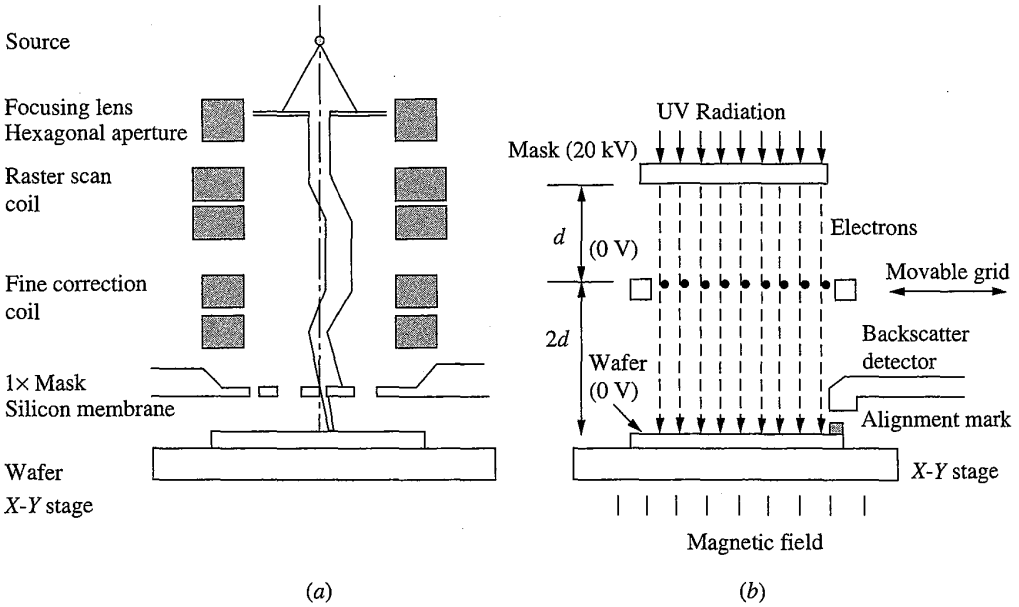


FIGURE 24
The principles of electron proximity and 1× electron projection. (*a*) Electron proximity. (*After Meissner et al., Ref. 40.*) (*b*) The 1× electron projection. (*After Ward et al., Ref. 41.*)

Micron Ex. 1023, p. 331
Micron v. YMTC
IPR2024-00909

Overlay error is less than 0.05 μm (1 σ). Estimated exposure time is around 30 sec per 5-inch wafer with 10 $\mu C/cm^2$ sensitivity. The major disadvantages of this system are the need for two masks for each pattern and the heating problem associated with the stencil mask.

Electron projection systems[41] are another method of achieving high resolution over a large field with high throughput as shown in Fig. 24*b*. In a 1 : 1 projection system, parallel electric and magnetic fields image electrons onto the wafer. The "mask" is quartz, patterned with chrome and covered with CsI on the side facing the wafer. Photoelectrons are generated on the mask/cathode by backside UV illumination. The advantages of the system include a stable mask, good resolution, fast-repeat exposure with low-sensitivity resists, large field, and fast alignment. Proximity effects can be compensated for by undersizing or oversizing the features on the mask or by increasing the acceleration energy to 50 keV or more. Apparently neither method is entirely satisfactory for production use. Another problem is that the cathode has an unacceptably short life, only 50 exposures, before the CsI must be replaced by fresh material.

Recently, another projection method, a modification of the variable-shaping method, has been proposed by several organizations.[42] Several different unit patterns, such as the cell patterns of a memory chip, are located in the second shaping aperture of the variable-shaping column as shown in Fig. 25. The deflection range of the electron-beam shaping is larger than with the variable-shaping method in order to select the different types of cells. Hence, redeflection is necessary to eliminate the beam shape deformation passing through the field far from the center of the lenses. The reduction ratio of the cell patterns is the same as that of the variable rectangle. Repetitive patterns are exposed by using the cell aperture, which is designed and manufactured in advance, that is specific to the device layout. Random patterns or areas that connect the repeated patterns are exposed by a conventional variable rectangular beam. Cell apertures are manufactured by a method similar to the one used to make the electron proximity mask, but the pattern size of the cell aperture is some ten times as large as the actual wafer geometry. This technique is very flexible because it can use either a conventional variable-shaping function or preprogrammed shaping. The problems of this method are the formulation of the flexible algorithms that combine the data preparation of the ULSI patterns with exposure control, the adjustment of electron-beam optics, and the life of the cell aperture.

6.3.5 Electron Resists

Electron exposure of resists occurs through bond breaking (positive resist) or the formation of bonds between polymer chains (negative resist). The incident electrons have energies far greater than the bond energies in the resist molecules, so the exposure energies can cause direct molecular reactions. Both bond scission and bond formation occur simultaneously. Which one predominates determines whether the resist is positive or negative.

In a negative resist, electron beam–induced crosslinks between molecules make the polymer less soluble in the developer solution. One crosslink per molecule is sufficient to make the polymer insoluble. Resist sensitivity increases with increasing

Micron Ex. 1023, p. 332
Micron v. YMTC
IPR2024-00909

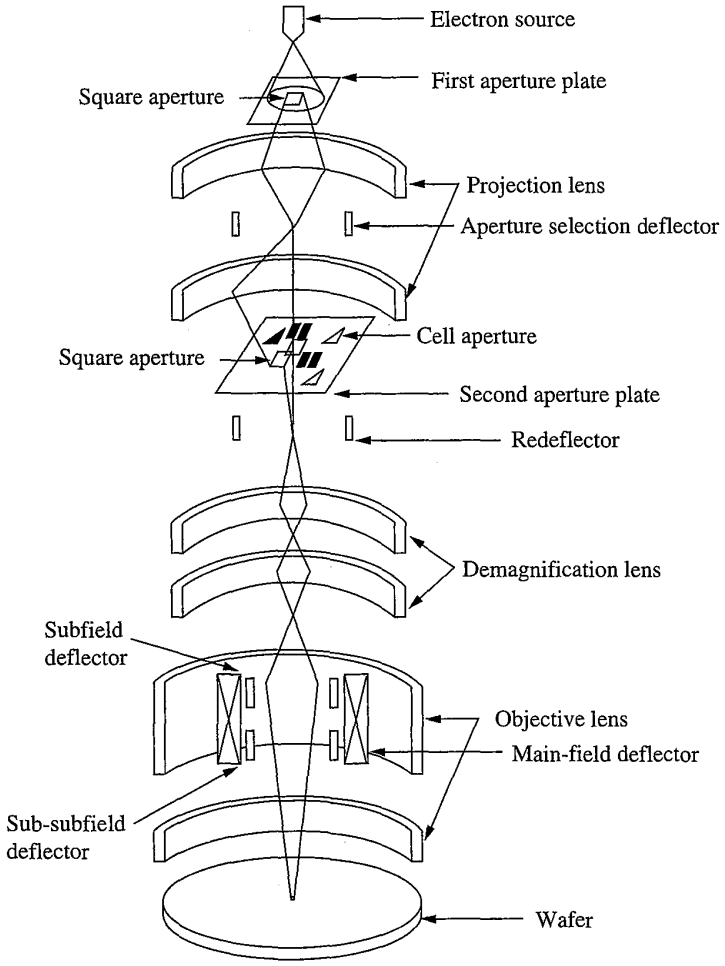


FIGURE 25
The electron optics of the cell projection method. (*After Sakitani et al., Ref. 42.*)

molecular weight. If the molecules are larger, then fewer crosslinks are required per unit volume for insolubility. The polymer molecules in the unexposed resist will have a distribution of lengths or molecular weights, and thus a distribution of sensitivities to radiation. The narrower the distribution, the higher the contrast γ—similar to an optical resist, as shown in Fig. 15. The exposure dose has units of charge deposited by the beam per unit area, or C/cm^2.

In a positive resist the scission process predominates; the exposure leads to lower molecular weights and greater solubility. Again, high molecular weight and narrow distribution are advantageous.

Two major factors limit resist resolution: swelling of the resist in the developer (more severe for negative resists) and electron scattering. Swelling of a negative resist, whether optical, electron, or x-ray, has two deleterious effects. First, two adjacent lines of resist may swell enough that they touch. Second, this expansion and contraction weakens the adhesion of very small resist features to the substrate and

Micron Ex. 1023, p. 333
Micron v. YMTC
IPR2024-00909

can cause small undulations in narrow lines. Both problems become less severe as resist thickness is reduced.

Recently developed chemically amplified resists have good resolution, sensitivity, and dry-etching resistance capabilities without swelling.[43] The basic process is explained in Section 6.2.5. It is reported to have a 0.2-μm resolution with a 4.0-$\mu C/cm^2$ sensitivity at 20 kV. Dry-etching performance is as good as that of conventional diazoquinone-based positive resists.

A few of the readily available electron resists are listed in Table 6. As the wavelength of optical lithography becomes shorter, several optical resists also have sensitivity to an electron beam. The resist PMMA has the highest resolution known. Note that the values for sensitivity and resolution are approximate. Because faster electrons penetrate more deeply, more current is required at higher voltages. A resist is about one-half as sensitive for 20-keV electrons as it is for 10-keV electrons.

When electrons are incident on a resist or other material, they enter the material and lose energy by a collision process known as scattering, producing secondary electrons, x-rays, and, finally, heat. This fundamental process limits the resolution of electron resists to an extent that depends on resist thickness, beam energy, and substrate composition. The envelope of the electron cloud in the material is shaped like an onion bulb; the cloud is pulled closer to the surface as beam voltage decreases. At higher beam voltage the electrons penetrate farther before being scattered over larger lateral distances. These phenomena are shown in Fig. 26, from the results of Monte Carlo simulations of the penetrating paths of the individual electrons in the target materials.[44] The scattering range of electrons in the Si substrate is about 1 μm for both distance and depth at 10 keV, and 2 μm for distance and 3.5 μm for depth at 20 keV.

For an electron beam of zero width incident at position r on the resist-covered substrate, the distribution of energy deposited in the resist at depth z is closely approximated by

$$f(r, z) = a_1 \exp[-r^2/\beta_f^2(z)] + a_2 \exp[-r^2/\beta_b^2(z)] \tag{6.23}$$

where a_1, a_2 = ratios of forward- and backscattering of the total electrons, respectively
$\beta_f(z)$ = width for the forward-scattered beam in the resist
$\beta_b(z)$ = backscattering from the substrate

and $\beta_b(z) \gg \beta_f(z)$, $f(r, z)$ is the point spread function. Generally, the value of z of interest is that corresponding to the resist/substrate interface z_i. For a 25-keV electron beam penetrating a 0.5-μm thick resist substrate, for example, $\beta_f(z_i) = 0.06$ μm, $\beta_b(z_i) = 2.6$ μm, and $a_2/a_1 = 2.7 \times 10^{-4}$. As separation between lines decreases, the backscattered electrons contribute a greater dose between the lines, where the dose should be zero. This feature is somewhat similar to the reduced modulation in an optical image at higher spatial frequencies.

An exposed pattern element adjacent to another element receives exposure not only from the incident electron beam but also from scattered electrons from the adjacent element. This is called the *proximity effect* and is, of course, more pronounced the smaller the space between pattern elements. For example, an isolated 0.5-μm line requires 20–30% more exposure than 0.5-μm lines separated from each other

Micron Ex. 1023, p. 334
Micron v. YMTC
IPR2024-00909

TABLE 6
Electron-beam resists

Name	Maker	Type	Polymer	Sensitivity ($\mu C/cm^2$) at 20 keV	Contrast	Resolution (μm)
CMS-EXS	Toyo Soda	Negative	Chloromethyl polystyrene	2	1.5	0.75
CMS-EXR	"	"	"	20	3	0.3
SEL-N	Somar	"	PGMA[†]–maleic methylate	0.4	0.9	1
RE4000N	Hitachi Chemical	"	Iodinated polystyrene	4	1.3	1
GMCIA	ATT/Mead	"	GMA[‡]–3-chlorostyrene	7	1.7	0.5
SAL 601	Shipley	"	Novolac	5	5	0.2
PBS	Mead Tech	Positive	Polybutene SO_2	2	1.7	0.5
PMMA	KTI Chemical	"	PMMA	100	2	<0.1
EBR-9	Toray	"	Poly-2,2,2-ethyl a-acrylate	1.5	3	0.5
FBM-120	Daikin Ind.	"	Fluoroallyl methacrylates	2	5	1.5
AZ 2400	Shipley	"	Novolac	200	2	0.5
RE 5000P	Hitachi Chemical	"	PMPS[§]-Novolac	4	1	0.8
ZEP 520	NihonZeon	"	a-Chloromethacrylate, a-Methylstyrene	5–20	2.5	<0.1

[†]Polyglycyl methacrylate
[‡]Glycyl methacrylate
[§]Polymethylpentane sulfone

Micron Ex. 1023, p. 335
Micron v. YMTC
IPR2024-00909

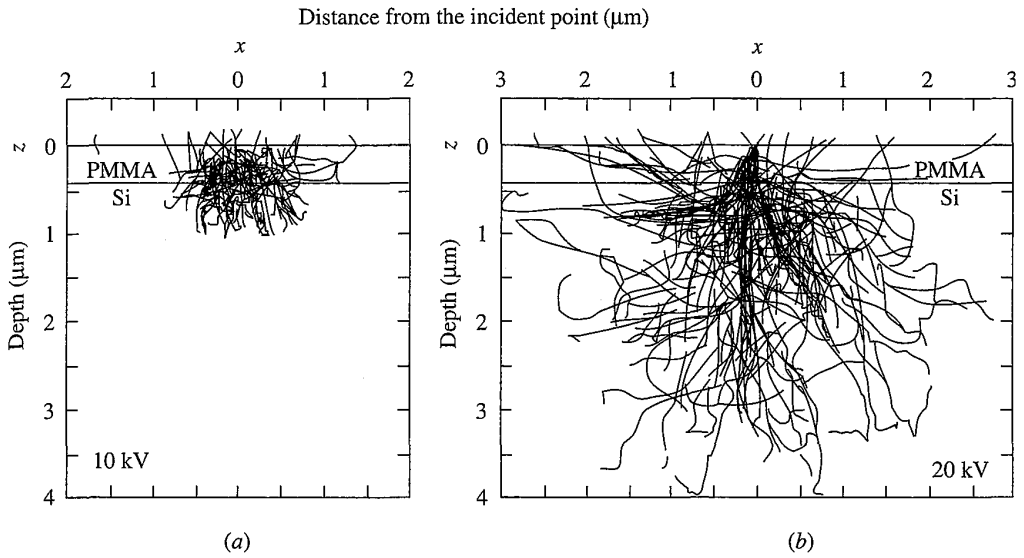


FIGURE 26
Simulated electron trajectories for a beam incident at the origin. (*After Kyser and Viswanathan, Ref. 44.*) The trajectories of 100 primary electrons are shown projected onto the x-z plane. (*a*) A 10-kV point source. (*b*) A 20-kV point source.

by 0.5 μm. Thus, as pattern density increases, it becomes necessary to adjust the exposure for various classes of elements or, in the extreme case, for different parts of an element.[45] This process is carried out through data preparation from the CAD output into the machine-oriented format in Fig. 18. In general, it takes a very long time, even for advanced computers, to apply perfect correction in ULSI patterns.

6.3.6 Electron-Beam Applications

Mask fabrication

Electron beam exposure machines have a wide variety of applications because of their pattern-generating function. Applications include direct writing and mask writing for the other lithographies such as the optical stepper, x-ray, ion, or electron projection. By the beginning of the 1970s the 10× reticle pattern was exposed by optical pattern generators. The rectangles are generated by a pair of L-shaped mechanical masks, and then the reduced images of these rectangles are projected on the blanks. A large, dense chip requires 20 hours or more of optical pattern generator time, but only two hours or less of electron beam pattern generator time. The first widespread use of electron beam pattern generations was in photomask making.

Table 7 lists specifications for the advanced masks required for optical steppers. In this table pattern placement errors are displacements of patterns from the desired locations on the mask. One of the types of error components that degrade these mask specifications is the relative-position error between the stage and the exposure position, which occurs because of mechanical slippage of the mask due to stage movement or thermal expansion. Fused-silica mask substrates, with their low

Micron Ex. 1023, p. 336
Micron v. YMTC
IPR2024-00909

TABLE 7
Specifications of masks for optical steppers

Minimum feature size (μm)	Line size on mask (μm)	Pattern placement error (μm)	Critical dimension uniformity (μm)	Defect size limit (μm)	Applied DRAMs
0.8	4	0.2	0.15	1.5	1 M
0.65	3.25	0.2	0.15	1	4 M
0.5	2.5	0.12	0.1	0.74	16 M
0.35	1.75	0.09	0.07	0.5	64 M
0.25	1.0*	0.05*	0.04*	0.25*	256 M

*Values on 4× mask.

thermal expansion (coefficient of linear expansion is 0.4×10^{-6}/°C), can reduce thermal contributions that result in relative-placement errors among several masks. Another type of placement error is caused by the bending of the mask. To minimize this error, the use of 0.25-inch thick blanks is beneficial compared to the conventional 0.09-inch ones. Commercially available machines have metrology capabilities to measure exposure performance of the masks.

Several factors also contribute to decrease the critical dimension (CD) uniformity, that is, the exposed pattern width uniformity. If the chemical process for the resist development is perfect, the factors that affect the CD are the dose and beam profile. Usually dosage is controlled by beam irradiation time (beam-on time with unblanking) and depends on the beam current. Beam blanking and unblanking control has a very quick response—within several nanoseconds. The problem is the heating effect of electron bombardment, which is more important in direct writing because of rather thick resists.[46,47]

Another important parameter of mask performance is the defect density. Usually, the etched pattern is inspected by an optical method and compared with the original data base used for the electron exposure. Mask defects are either opaque spots in areas that should be transparent or pinholes where the chromium layer should be continuous. The allowable defect size, which depends on the layout rule of the devices, is shown in Table 7. The mask must be perfect within that size. Consequently, inspection and repair technology is very important not only to ensure the device performance but also to increase the mask production yield. The minimum thickness of the resist is set by the need to avoid pinholes and by its resistance to etching. The resist thickness of the photomask is approximately 0.2 to 0.4 μm. In the case of thin resists, it is easy to control the CD by using uniform process conditions, such as minimizing swelling, proximity effect, heating effect, and charging of electrons.

Mask making can be more difficult than direct writing in spite of its 5× larger geometry. Absolute position accuracy and pattern-integrity perfection on the completely clean blanks, without any references such as alignment marks on wafers, are required during a rather long exposure time. The reduction ratio of the optical stepper will become 4× in the case of 256-Mbit DRAM production instead of the conventional 5× because of the large chip size and the limit on field size of the stepper as shown in Table 1. The required specifications of masks for 256-Mbit DRAM

Micron Ex. 1023, p. 337
Micron v. YMTC
IPR2024-00909

become more severe than the conventional ones because of changing the reduction ratio.

Direct writing

Direct writing means making patterns directly on the resist-coated substrate. The advantages of direct writing are that no mask is needed and that it produces a high-resolution pattern. Because it is maskless, direct writing can be used in many applications. For example, quick design modification; shortened manufacturing time from design to device test, usually called QTAT for quick turnaround time; and the capacity to put different types of pattern layouts on one wafer (one wafer module) are advantages of maskless applications. In ASIC manufacturing, the basic design and process are common from chip to chip; only the wiring and through-hole layers are different based on the logical functions of the devices.[48] Hence, the common layers can be manufactured using a stepper, and the personalization layers can be exposed by electron beam. Direct writing is very efficient not only because it is economical but also because of its higher production speed. If the mask-and-stepper process is used, it needs almost one full day from receiving the exposure data to making a mask without defect. In the direct writing process, the exposed pattern data are transferred directly to an electron beam machine, without the use of external components such as magnetic tape, through a network from the CAD system connected to the electron exposure machines. The trilayer resist process has been applied to actual production to get higher throughput, to achieve better resolution, and to minimize radiation damage.[48]

The first 64-Mbit DRAM was made by a mix-and-match method using the electron beam and the stepper.[49] Depending on pattern density, several newly developed technologies were used, such as the positive resist OEBR200 (Tokyo-oka) and the negative resist SAL601 (Shipley). To control CD uniformity in 0.3-μm geometry, a two-step development process (contrast enhancement), a conductive material coating to eliminate charge accumulation, and a new proximity-correction algorithm (to classify the pattern density area instead of only the pattern width and gap) are employed.

Another direct writing application, with long experience, is GaAs field-effect transistor (FET) production. A short gate length and low resistance are very important for this device to have high performance. A T-shaped gate is one way to accomplish this. A double exposure with different exposure conditions is another.[50] Two types of PMMA, with different sensitivities (molecular weights), are coated on the wafer. The top resist has a higher sensitivity and lower molecular weight compared to the bottom one. After the top layer has been opened, the second layer is exposed by referring to the alignment mark as shown in Fig. 27. It is reported[50] that the gate contact width is 0.15 μm.

As the pattern size becomes smaller and smaller, it becomes more difficult to make the pattern on the resist and transfer it to the wafers. There have been many approaches to making patterns directly on the wafer without a resist, especially in applications that need a geometry of less than 0.1 μm.[51,52] For direct electron beam deposition, an electron beam reacts with the source gas directly. The gas is introduced to the specimen chamber, and a Au pattern is deposited. Fifty-nm lines and spaces are reported.[51]

Micron Ex. 1023, p. 338
Micron v. YMTC
IPR2024-00909

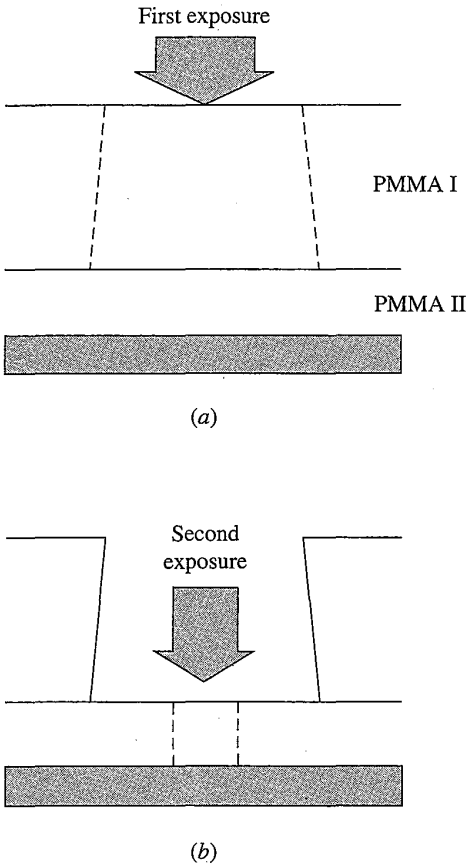


FIGURE 27
Two-step exposure principle. (*After Samoto et al., Ref. 50.*) (*a*) First exposure on high-sensitivity resist. (*b*) Second exposure on low-sensitivity resist.

6.4 X-RAY LITHOGRAPHY

Since x-ray lithography was proposed in 1972, development of the technology has been pursued in many laboratories.[53,54] In the previous section we noted that diffraction effects and resolution are improved by reducing the wavelength. If the wavelength is reduced further than deep UV, all optical materials become opaque because of fundamental absorption, but transmission increases again in the x-ray region. In x-ray lithography, an x-ray source illuminates a mask that casts shadows on a resist-covered wafer. There are several advantages in x-ray lithography in addition to short wavelength. Some contaminants, such as light organic materials, do not print as a defect; and the depth of focus is larger than that of optical printers. The essential technology components of this process are (1) a mask consisting of a device pattern made of x-ray-absorbing materials on transmitting material, (2) an x-ray source, and (3) an x-ray resist.

The x-ray absorption of several materials is shown in Fig. 28. The absorption coefficient of an elemental material of density ρ and atomic number Z is proportional to $\rho Z^4 \lambda^3$ over a wide range of wavelengths. The proportionality constant decreases in a step function fashion at the *absorption edge,* a wavelength that corresponds to

Micron Ex. 1023, p. 339
Micron v. YMTC
IPR2024-00909

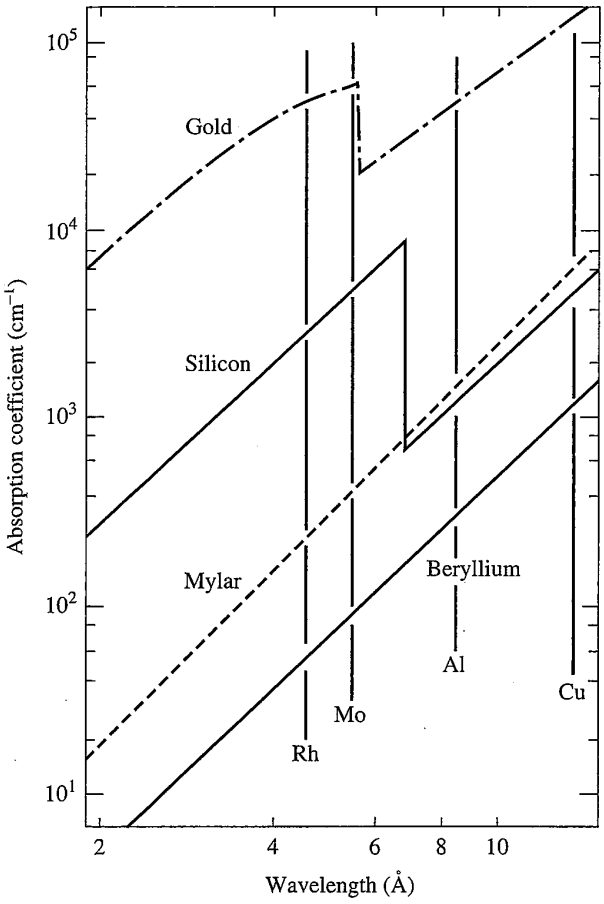


FIGURE 28
Typical x-ray absorption coefficients with some lines characteristic of commonly used materials.

the ionization energies of inner electrons of the K, L, and other shells. Note in Fig. 28 that the big differences in absorption coefficients for different materials are observed at the same wavelengths that are utilized for absorbing materials to make patterns and the transmitting substrate of the x-ray mask.

The performance required of x-ray lithography, as the alternative to optical lithography, has become more exacting as optical technology has continued to improve. The resolution and placement accuracy should soon surpass 0.2 μm and 0.03 μm, respectively.

6.4.1 Proximity Printing

Because the x-ray wavelength is short enough, simple geometrical considerations can be used to relate the image formation and the wafer without having to consider diffraction, as shown in Fig. 29*a*. The opaque parts of the mask cast shadows onto the wafer below. The edge of the shadow is not absolutely sharp because of the finite extent of the x-ray source at distance D from the mask. If the gap between mask and wafer is g, this blur δ is given by

$$\delta = Sg/D \tag{6.24}$$

Micron Ex. 1023, p. 340
Micron v. YMTC
IPR2024-00909

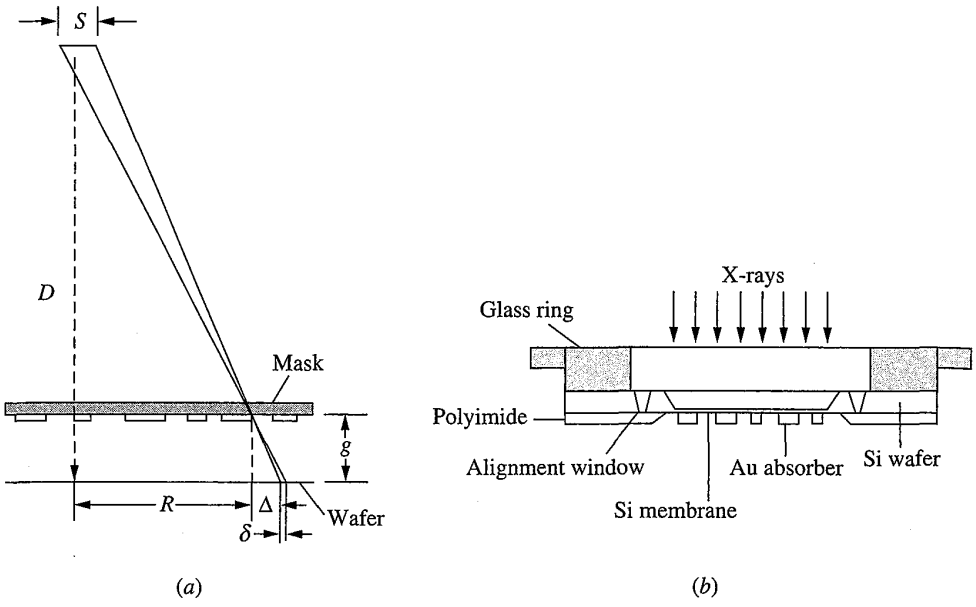


FIGURE 29
X-ray shadowing errors and x-ray mask structure. (*a*) X-ray proximity printing consideration. (*b*) X-ray mask structure example. (*After Fleming et al., Ref. 54.*)

The angle of incidence of the x-ray on the wafer varies from 90° at the center of the wafer to $\tan^{-1}(D/R)$ at the edge of the exposure field of radius R. The shadows are slightly longer at the edge by the amount

$$\Delta = g(R/D) \tag{6.25}$$

This small magnification is generally of no concern. In the special case where it may be undesirable, it can be compensated for when the mask is patterned. For multilevel devices the magnification must have the same value for each level, or at least its variation must be within the registration tolerance. This implies stringent control of the gap g. Wafer warping in processing can be nearly eliminated with a proper vacuum chuck. It is not necessary that the gap have the same value at all points on the wafer, only that the spatial variations be the same, within close tolerance, for all levels. The step-and-repeat motion is indispensable to the narrow beam line of x-ray lithography.

6.4.2 X-Ray Mask

An example of a typical mask structure is shown in Fig. 29*b*. An x-ray mask consists of a transmissive membrane substrate and an absorber patterned on it. The ratio of metal thickness to substrate thickness is greater than for a photomask because no materials are available that are fully transparent or fully absorbent, unlike the combination of glass and chromium in optical lithography. These thicknesses are determined by the transmission of the materials for the x-ray wavelength of interest. X-ray masks are made by electron lithography and use the technologies of mask making and direct

Micron Ex. 1023, p. 341
Micron v. YMTC
IPR2024-00909

writing. The pattern on the x-ray mask must be as perfect as a photomask, but the dimension is the same as on the wafer because of proximity printing. For features down to a 0.1 to 0.2-μm geometry it is harder to check the pattern integrity and to repair defects in x-ray than in optical lithography.[54]

Of the heavy metals with large ρZ^4 values, gold was once widely used, because it is relatively easily patterned by liftoff or electroplating. Tungsten and tantalum are used now, because they are easily etched by dry etching. The thicknesses of gold necessary for absorption of 90% of the incident x-ray flux are 0.7 μm, 0.5 μm, 0.2 μm, and 0.08 μm for x-ray wavelengths 4.4 Å (Pd_L), 8.3 Å (Al_K), 13.3 Å (Cu_L), or 44.8 Å (C_K), respectively. In general, the metal is considerably thicker than the chromium layer (0.1 μm) on a photomask. Methods for high-resolution patterning of the gold include electroplating and ion milling. Electroplating produces excellent definition with vertical walls but requires a vertical wall primary pattern in a resist that has a thickness equal to that of the metal to be plated.[55] More often, a subtractive process has been employed in which a thinner resist layer is used to pattern a thin layer of a refractive metal; the refractive metal serves as a mask for ion-milling the underlying gold. With this method, walls that depart from the vertical by 20° or less can be formed. The minimum linewidth attainable by ion milling 0.5-μm-thick gold is ~ 0.4 μm. For higher resolution, and where gold thickness can be reduced, longer wavelengths such as the 12 Å Cu_L radiation may be used. Lines as small as 0.16 μm have been replicated with this type of radiation.

The membrane forming the mask substrate should be as transparent as possible to the x-rays, smooth, flat, dimensionally stable, reasonably rugged, and transparent to visible light if an optical registration scheme is used. Materials that have been used include polymers such as polyimide and polyethylene terephthalate, silicon, SiC, Si_3N_4, Al_2O_3, and a Si_3N_4–SiO_2–Si_3N_4 sandwich structure. Although different mask substrates are appropriate for different portions of the soft x-ray spectrum, there is not yet general agreement on the best material for any particular wavelength.

The major questions remaining about x-ray masks concern their dimensional stability, minimum attainable defect densities, and ease of handling. The stress applied on the thin membrane during processing and formation of the absorber structure causes distortions. Absorber-induced distortion is not noticeable because it is evaluated by comparing the measured fiducial marks before and after electroplating. Resist films on the membrane produce tensile stress on the membrane. In the case of a multilayer process, RIE heating causes membrane distortion with a maximum error of 0.1 μm.[55] Dimensional stability can be degraded by radiation damage produced by x-ray flux, which also makes the mask substrate optically opaque. Pattern placement and CD accuracies are reported as 0.06 μm (3 σ) and are repeatable.[54]

6.4.3 X-Ray Sources

X-rays are produced by the interaction of incident electrons and a target material. The maximum x-ray energy is the energy E of the incident electrons. If E is greater than the excitation energy E_c of the characteristic lines of the atoms of the material, the x-ray spectrum will contain these lines.

Micron Ex. 1023, p. 342
Micron v. YMTC
IPR2024-00909

X-ray generation by electron bombardment is a very inefficient process; most of the input power is converted into heat in the target. The x-ray flux is generally limited by the heat dissipation in the target. With electrons focused to a spot 1 mm in diameter on an aluminum target on a water-cooled stem, 400 to 500 W is a typical upper limit for the input power. The x-ray power produced is only about 10 mW, and this power is distributed over a hemisphere. The x-ray power is proportional to $(E - E_c)^{1.63}$.

Another x-ray source, which is capable of an order of magnitude greater flux, is plasma discharge.[56] There are several versions, but all function by heating a plasma to a temperature high enough to produce x-ray radiation. The radiation consists of strong lines superimposed on a weak continuum. The source is pulsed at a low rate. In one embodiment, the source size is 2 mm. The repetition rate is 3 Hz. Special problems with such a source are reliability and contamination produced in the plasma chamber.

Typical x-ray sources[57] are listed in Table 8. Although the electron bombardment and plasma sources are included,[58,59] their resolution and registration are limited because of their x-ray trajectory divergences and δ values shown in Table 8.

6.4.4 Synchrotron Radiation

We have considered many kinds of lithography equipment in this chapter. The synchrotron is the biggest and the most expensive; including its support facilities, it is more than two orders higher than any other lithography equipment. Some experiments have been reported in which the radiation from electron synchrotrons and storage rings was used for x-ray lithography. The small angular divergence of the radiation simplifies mask–wafer registration, and the high intensity of the radiation leads to short exposure times. A single storage ring could provide radiation to a large number of exposure stations.

In synchrotrons and storage rings, high-energy electrons are forced into closed curved paths, or orbits, by magnetic fields. An electron moving through a perpendicular magnetic field has an acceleration directed toward the center of the orbit and emits radiation converted from the kinetic energy. For the high-energy electrons of interest, which have velocities very nearly equal to that of light, the radiation is emitted in a narrow cone in the forward direction of motion of the electron. An observer looking along a tangent to the orbit sees a bright spot. The radiation is very different from that of a point source because of the narrow beam from each electron. The radiation from a circular divergence of the radiation in the vertical direction is $\Psi \approx (1957\,E)^{-1}$, where E is the electron energy in GeV and Ψ is in radians. Thus, for a 1 GeV machine the vertical divergence is only 0.5 mradians. The synchrotron is the only equipment that has a small divergence angle and high brightness.

High-energy electrons are provided to the storage ring by a small synchrotron or a small linear accelerator. The ring is operated, briefly, as a synchrotron to boost the electron energy to the final value. Then the electrons may circulate for several hours in a stable orbit. The loss due to the power radiated as synchrotron radiation is compensated by one or more acceleration cavities around the ring. Nevertheless,

Micron Ex. 1023, p. 343
Micron v. YMTC
IPR2024-00909

TABLE 8
The characteristics of x-ray sources[57]

Source	Divergence (mrad)	Distance source/mask, D (m)	Source diameter, S (mm)	Proximity gap, g (μm)	Runout, Δ (nm)	Blurring, δ (nm)	Spectrum (nm)	Total power on mask (mW/cm^2)
Synchotron	5	4	0.5	50	250	< 10	Broadband, 0.4–2	> 100
Electron bombardment	50	0.4	3	50	250	400	Line, 0.44–1.3	> 1
Plasma focus	50	0.4	0.2	50	250	< 35	Line, 0.7–1.5	> 10
Laser plasma	50	0.4	0.1	50	250	15	Line, 0.9–1.7	> 1

Micron Ex. 1023, p. 344
Micron v. YMTC
IPR2024-00909

the current slowly decays because electrons collide with residual gas molecules or the walls.

The peak of the power spectrum of the synchrotron radiation occurs at wavelength λ_p. This is related to the electron energy E (in GeV) and magnet bending radius R (in meters) by

$$\lambda_p = 2.35R/E^3 \tag{6.26}$$

A 0.83 GeV machine with $R = 2.1$ m would have a power spectrum with a peak at $\lambda_p = 8.4$ Å.

Oxford Instrument's compact synchrotron Helios® was installed at IBM, East Fishkill, NY. Figure 30 shows the total facility arrangement.[60] The size of Helios is about 2 m × 4 m. It needs many other components—the linear accelerator, beam lines, RF controllers, magnet controllers, pumping units, and a radiation protection building. It has a 0.9-nm wavelength, 0.4-mm source size, and a 45-hr beam life. The x-ray stepper made by Karl-suss has a 0.25-μm resolution and 0.1 μm overlay accuracy.[61] Synchrotron radiation actually has been used in 0.35-μm static RAM (SRAM) manufacturing. The critical layers (isolation, gate, contact hole, and metal 1) were fabricated using the synchrotron on the completed layers made by optical *i*-line or KrF steppers. The CD control of eight lots is reported[62] as 30–50 nm (3 σ).

6.4.5 X-Ray Projection

The use of soft x-ray projection lithography has been demonstrated.[63] The key components of this system are x-ray mirrors, which consist of alternating layers of high-index and low-index materials deposited on a smooth substrate. The choice of materials is determined by the x-ray wavelength, and the individual layer thickness is determined by the wavelength and the incidence angle. The normal incidence reflectivity is better than 50% at $\lambda = 4.5$ nm. A laser-produced plasma source is adopted for the x-ray source. The system configuration is illustrated schematically in Fig. 31. The demagnification ratio is 1 : 5. The advantages, compared to proximity methods, are greater latitude with respect to placement of the mask pattern and less radiation-induced damage of the mask. Theoretical resolution is 8.9×10^3 line-pairs per millimeter, or a linewidth of 56 nm, based on $1/\lambda f^{\#}$, where $\lambda = 4.5$ nm and $f^{\#}$ is 125. The image field size is as small as 5 mm in diameter. The exposure time is estimated roughly as 15 minutes for a 6-inch wafer.

6.4.6 X-Ray Resists

An electron resist is also an x-ray resist, since an x-ray resist is exposed largely by photoelectrons produced during x-ray absorption. The energies of these photoelectrons are much smaller (0.3 keV to 3 keV) than the 10 keV to 50 keV energies used in electron lithography, making proximity effects negligible with x-rays and promising higher ultimate resolution.

On traversing a path of length z in a resist or any other material, an x-ray flux is attenuated by the factor $\exp(-\alpha z)$. The 1-μm thick resist film absorbs about 10%

Micron Ex. 1023, p. 345
Micron v. YMTC
IPR2024-00909

1. Concrete shielding
2. Linear accelerator
3. Control room
4. RF room
5. DC power room
6. Research area
7. Exposure station
8. Cryogenics
9. Electron storage ring
10. X-ray steppers
11. Wafer processing area

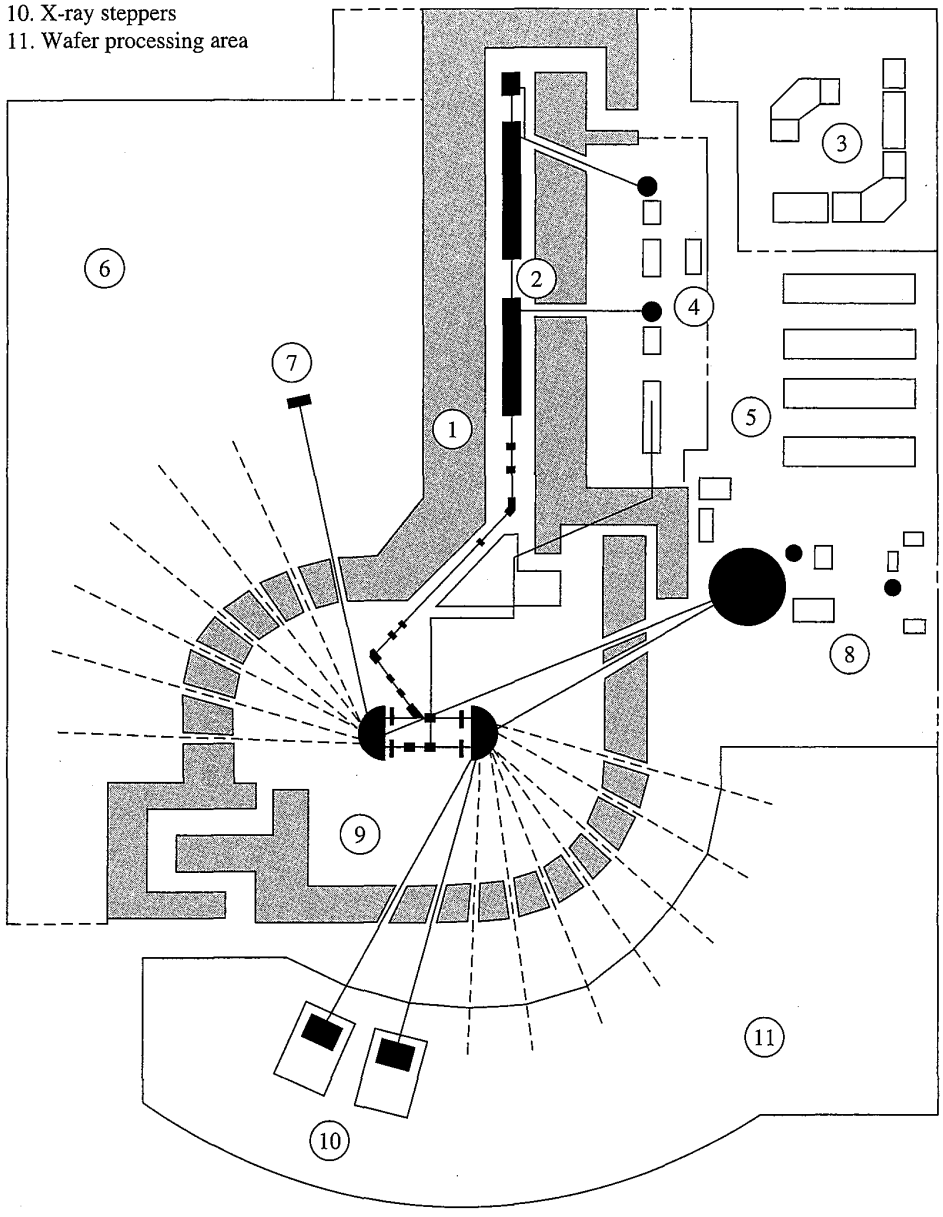


FIGURE 30
Layout of IBM's synchrotron facility. (*After Archie et al., Ref. 60.*)

Micron Ex. 1023, p. 346
Micron v. YMTC
IPR2024-00909

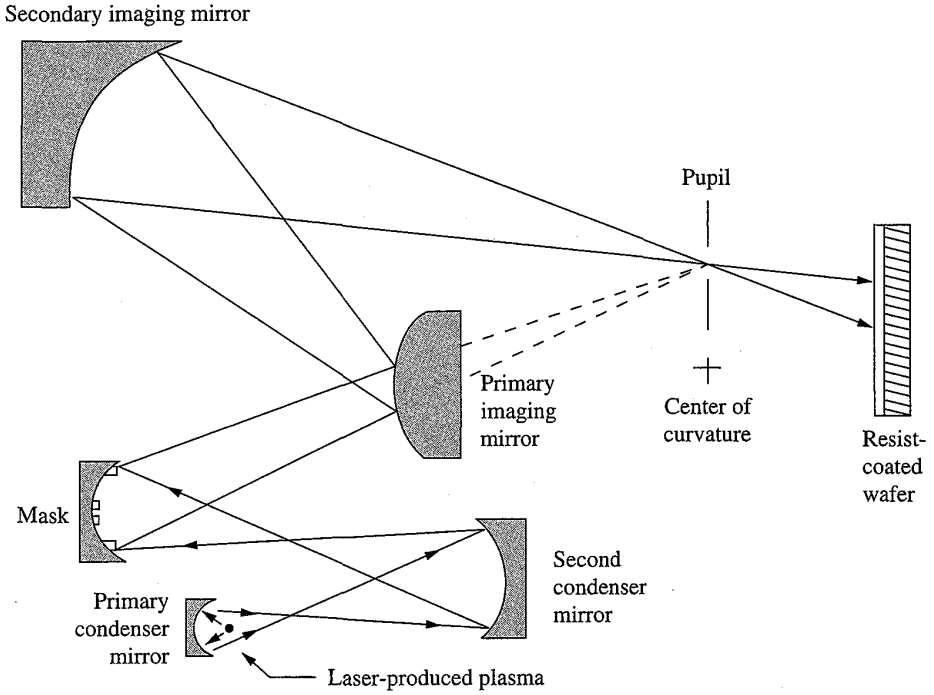


FIGURE 31
Schematic illustration of an x-ray projection system. (*After Hawryluk and Seppala, Ref. 63.*)

of the incident x-ray flux at the characteristic x-ray $Al_K\alpha$ ($\lambda = 8.3$Å), because most resists consist of H, C, and O only with a density $\rho \approx 1$ g/cm^3. This small absorption has the advantage of providing uniform exposure throughout the resist thickness z and the disadvantage of reduced sensitivity. As in the optical case, x-ray resist sensitivity is generally quoted in terms of incident dose Q (J/cm^2) required for exposure; sometimes absorbed dose αQ (J/cm^3) is used.

For the shorter x-ray wavelengths the λ^3 dependence of the absorption coefficient leads to low sensitivity. This can be offset by incorporating heavier elements in resist components to increase absorption (sensitivity). In x-ray lithography, we expect a special sensitivity enhancement due to the relationships between resist materials and x-ray absorption edges. For the characteristic x-ray $Pd_L\alpha$ radiation, $\lambda = 4.37$ Å, the DCOPA resist (a mixture of poly (3.3-dichloro-1-propyl acrylate) and PGMA/ZA) incorporates chlorine. The Cl_K absorption edge at 4.40 Å provides higher absorption, and chlorine is a chemically reactive species. Thus, two effects contribute to higher sensitivity.

6.5 ION LITHOGRAPHY

When an ion beam is used to expose a resist, higher resolution is possible than with an electron beam because of less scattering. In addition, resists are more sensitive to ions than to electrons. There is also the possibility of a resistless wafer process.

Micron Ex. 1023, p. 347
Micron v. YMTC
IPR2024-00909

However, the most important application is the repair of masks for optical or x-ray lithography, a task for which commercial systems are available.

The sputtering yield increases with beam energy if the beam energy is larger than some small threshold value. There is an energy limit beyond which the yield decreases because the ions penetrate more deeply and fewer surface atoms receive enough energy to leave the surface. For example, the peak in the sputtering curve for Ar^+ ions incident on Cu occurs at 23 keV. For ion implantation, energies from 30 keV to 500 keV are used, and the dose ranges up to 10^{15} ions/cm^2 (or 1.6×10^{-4}C/cm^2 for monovalent ions). This represents a much larger dose than that used for resist exposure.

The sensitivity of PMMA resist has been measured for 30 keV, 60 keV, and 200 keV He^+ ions and for 100 keV and 150 keV Ar^+ ions. The required dose is nearly two orders of magnitude less than with 20 keV electrons.[64] The perpendicular straggle of the penetrating ion path and the range of low-energy secondary electrons produced are less than the range of backscattered electrons in electron lithography. The ion energies for exposing a resist depend on the ion species. If the ion must penetrate 2500 Å of resist, then a proton would need 14 keV; and a Au ion, 600 keV for the projected range to be 3000 Å.

Ion lithography systems are of two types: a scanning focused-beam system and a mask-beam system. The problems of ion optics for scanning ion systems are more serious than for electron optics. One problem is the ion source, and another is the beam-forming system. In the two practical types of field ionization source, ions are produced in the strong electric field near a pointed tungsten tip; the source of the ionized material is a gas or a liquefied metal surrounding the tip. The largest current densities obtained in the focused image of such a source are 1.5 A/cm^2 for Ga^+ in a 0.1 μm spot and 15 mA/cm^2 for H^+ in a 0.65 μm spot.[65] This value is two orders less than the conventional electron beam system. There are no bright sources for such useful implant species as B and P. Electrostatic lenses rather than magnetic lenses must be used to focus ion beams. If a magnetic lens were used to focus an ion beam, the field would have to be huge since the required field is proportional to $(mV/f)^{1/2}$ from Eq. (6.17), where m is the particle mass. Different isotopes would be focused at different points. Electrostatic optical systems generally have higher aberrations, necessitating a small aperture α (beam solid angle similar to optical or electron beams) and small scan fields.

A prototype scanning system[65] in which a beam of 57 keV Ga^+ ions is focused to a 0.1-μm diameter spot with current density 1.5 A/cm^2, is illustrated in Fig. 32. Spot size is limited by chromatic aberration of the electrostatic lens and the large 14-eV energy spread of the source.

Mask-based systems are of two types. One type is the 5×-reduction projection step-and-repeat system, which projects 60–100 keV light ions such as H_2^+ through a stencil mask.[66] The resolution and overlay accuracy are 0.15 μm and ±0.05 μm. The major disadvantage is the fragile foil mask, which can treat re-entrant patterns only by double exposure with two masks, similar to the technique used in the electron proximity method.

The second type is a step-and-repeat proximity printer in which 300-keV protons are projected through the 0.5-μm thick portion of an all-silicon mask.[67] The mask is aligned so that the ions travel along the channels in the [100] direction. Projected

Micron Ex. 1023, p. 348
Micron v. YMTC
IPR2024-00909

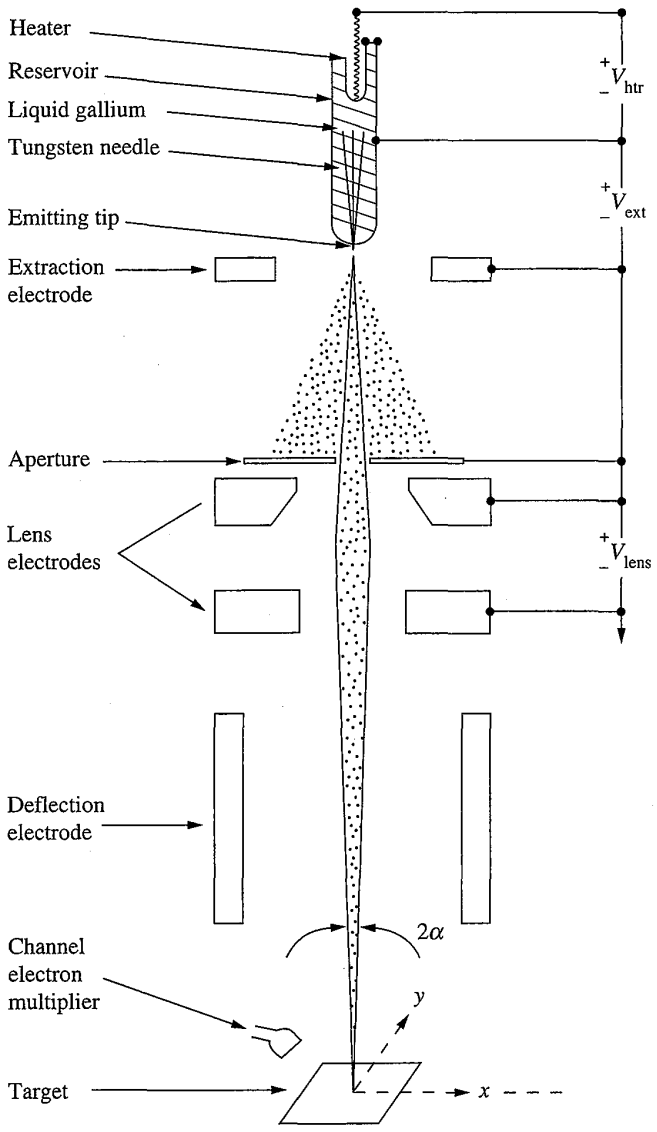


FIGURE 32
The optics of an ion lithography system. (*After Seliger et al., Ref. 65.*)

throughput is sixty 100-mm wafers/hr, with an overlay error of ~ 0.1μm—barely adequate for 0.5-μm lithography. Resolution is set by ions scattering as they emerge from the channels. Edge resolution is 0.1 μm for the beam-scattering angle of 0.3°. The main disadvantage is the fragile mask, with a 0.5 μm membrane. However, if a large fraction of the mask consists of the thicker silicon absorber, then the mask may be sufficiently rugged. The highest resolution in resist is obtained with a stencil mask and a beam of protons with an energy in the range of 40 to 80 keV. Lines as fine as 400 Å have been printed with this proximity printer.

Micron Ex. 1023, p. 349
Micron v. YMTC
IPR2024-00909

6.6 SUMMARY AND FUTURE TRENDS

In general, DRAMs have been used as the indicator of progress in ULSI technology. The most advanced DRAM currently in mass production is the 4-Mbit type with 0.8-μm geometry. The 64 Mbit will start production at the end of this century, with a 0.35-μm feature size. If we consider the resolution, optical lithography systems have a capability very close to the required performance as listed in Table 1. Optical lithography is also better than the other methods from the overlay consideration as described in Table 2. Hence, it is likely that conventional optical lithography will be applied to 64-Mbit DRAM production with the combination of several sophisticated technologies such as the phase-shifting mask and off-axis illumination. Other reasons for the continued use of optical lithography are the wide variety of tools and technologies that already exist, its convenience, and its reasonable cost. Some commercially available resists can resolve down to 0.2 μm, so optical lithography still has a margin over next-generation device manufacturing.[68]

In addition, the DOF has become very tight because the wavelength has been reduced and the NA increased, as discussed in Section 6.2.2. One solution to the problem of tight DOF is the multilayer resist process, which has other problems because of its complex process steps.

The performance of various lithography methods based on throughput and resolution are summarized in Fig. 33. The borderline of each technology is very fuzzy. It will be difficult for another technology to replace optical lithography until 0.2-μm resolution is required. This chapter mainly discussed resolution. However, other key

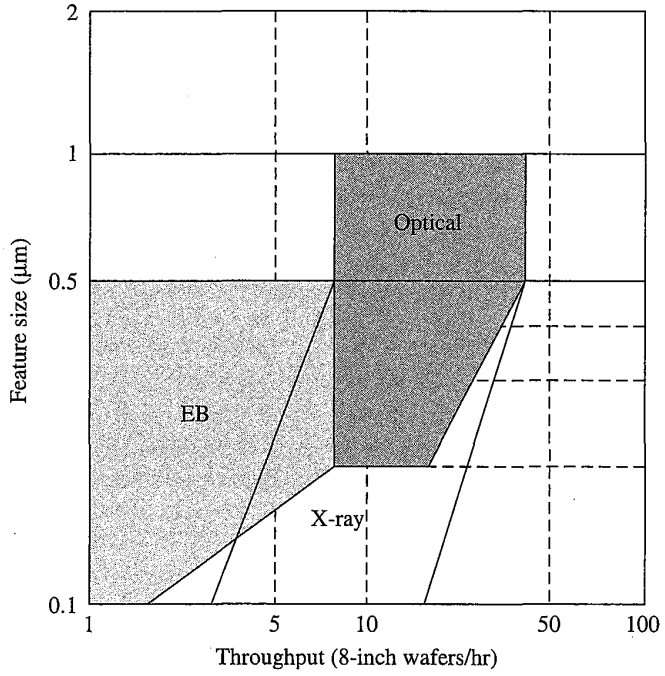


FIGURE 33
Resolution and throughput in the submicron region.

Micron Ex. 1023, p. 350
Micron v. YMTC
IPR2024-00909

issues are the pattern-width uniformity and the overlay accuracy. If the tolerances of these items are defined as 10% of the feature size, they should be less than 0.02 μm. It is very difficult to guarantee these accuracies because the accuracy of present measurement tools cannot meet this requirement.

Pattern-geometry accuracy is degraded in each step of the device manufacturing process. In general, the patterning sequences are (1) electron beam irradiation, (2) resist patterning on the mask, (3) chromium patterning on the mask, (4) stepper exposure, (5) resist patterning on the wafer, and (6) pattern transfer to a specified layer on the wafer. To minimize error, two methods are available. One is to skip some manufacturing steps, for example, using a direct reaction on the wafer by electron beam with gas molecules. The second is a statistical approach, which is used because of the difficulty in analyzing the cause of errors in the intermediate steps. This method has recently been applied to mask writing to improve the pattern-placement accuracy by multiple exposures. The more active approach is the evaluation and repair of function units of the final product, such as memory or logic cells, instead of inspecting and repairing in the intermediate fabrication steps such as masks.

It is very difficult to predict the trend in lithography beyond 0.2 μm or the architecture or structure of future ULSI. Optical technology is considered almost impossible to use for devices that have less than 0.2-μm geometry because of its resolution limit. We have only two choices: electron beam direct writing or an x-ray technology. However, it is very difficult to make perfect x-ray masks, either 1× transmission or 5× reflection, because of the complicated process steps including inspection and repair. In conclusion, the only available choice is electron beam direct writing. The throughput of this method is just a few 8-inch wafers per hour, as shown in Fig. 33, if the cell projection method is applied to production. For the mass production application, we must minimize the cost and footprint (required floor area) of the machine and develop high-sensitivity resists that are easy to process. It is considered possible that all of these requirements can be fulfilled.

REFERENCES

1. E. G. Cromer, Jr., "Mask Aligners and Steppers for Precision Microlithography," *Sol. State Technol.,* **36,** 23 (Apr. 1993).
2. B. J. Lin, "Deep UV Lithography," *J. Vac. Sci. Technol.,* **12,** 1317 (1975).
3. B. J. Lin, "Optical Methods for Fine Line Lithography," *Fine Line Lithography,* Roger Newman, Ed., North-Holland, Amsterdam, Ch. 2, 1980.
4. J. G. Skinner, "Some Relative Merits of Contact, Near-Contact, and Projection-Printing," *Proceedings of Microelectronics Seminar, Kodak Interface,* **73,** 53 (1974).
5. R. K. Watts, "Advanced Lithography," *Very Large Scale Integration, Fundamentals and Applications,* D. F. Barbe, Ed., Springer, New York, Ch. 2, 1982.
6. M. Born and E. Wolf, *Principles of Optics,* 6th ed., Pergamon Press, New York, 1980.
7. M. C. King and M. R. Goldrick, "Optical MTF Evaluation Techniques for Microelectronic Printers," *Sol. State Technol.,* **20,** 37 (Mar. 1977).
8. J. D. Cuthbert, "Optical Projection Printing," *Sol. State Technol.,* **20,** 59 (Aug. 1977).
9. M. Lacombat, J. Massin, G. M. Dubroeucq, and M. Brevignon, "Laser Projection Printing," *Sol. State Technol.,* **23,** 115 (Aug. 1980).

Micron Ex. 1023, p. 351
Micron v. YMTC
IPR2024-00909

10. B. J. Lin, "The Optimum Numerical Aperture for Optical Projection Microlithography," *Proc. SPIE,* **1463,** 42 (1991).
11. W. J. Smith, "Modern Optical Engineering," *The Design of Optical Systems,* 2nd ed., McGraw-Hill, New York, Ch. 13, 1991.
12. B. J. Lin, "The Path to Subhalf-Micrometer Optical Lithography," *Proc. SPIE,* **922,** 256 (1988).
13. R. Unger, C. Sparkes, P. A. DiSessa, and D. J. Elliott, "Design and Performance of a Production-Worthy Excimer-Laser-Based Stepper," *Proc. SPIE,* **1674,** 708 (1992).
14. M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, "Improving Resolution in Photolithography with a Phase-Shifting Mask," *IEEE Trans. Electron. Dev.,* **ED-29,** 1828 (1982).
15. K. Ronse, R. Jonckheere, C. Juffermans, and L. Van den Hove, "Comparison of Various Phase Shift Strategies and Application to 0.35 μm ASIC Designs," *Proc. SPIE,* **1927,** 2 (1993).
16. N. Shiraishi, S. Hirukawa, Y. Takeuchi, and N. Magome, "New Imaging Technique for 64M-DRAM," *Proc. SPIE,* **1674,** 741 (1992).
17. N. Shiotake and S. Yoshida, "Recent Advances of Optical Step-and-Repeat System," *Proc. SPIE,* **537,** 168 (1985).
18. F. H. Dill, W. P. Hornberger, P. S. Hauge, and J. M. Shaw, "Characterization of Positive Photoresist," *IEEE Trans. Electron. Dev.,* **ED-22,** 445 (1975).
19. L. F. Thompson, "Design of Polymer Resist for Electron Lithography," *Solid State Technol.,* **17,** 27 (1974).
20. F. H. Dill, A. R. Neureuther, J. A. Tuttle, and E. J. Walker, "Modeling Projection Printing of Positive Photoresists," *IEEE Trans. Electron. Dev.,* **ED-22,** 456 (1975).
21. H. Ito and C. G. Willson, "Applications of Photoinitiators to the Design of Resists for Semiconductor Manufacturing," *Polymer in Electronics,* ACS Symposium Series **242,** T. Davidson, ed., American Chemical Society, 11, 1984.
22. A. Yoshikawa, O. Ochi, and Y. Mizushima, "Dry Development of Se-Ge Inorganic Photoresist," *Appl. Phys. Lett.,* **36,** 107 (1980).
23. K. L. Tai, R. G. Vadimisky, C. T. Kemmerer, J. S. Wagner, V. E. Lamberti, and A. G. Timko, "Submicron Optical Lithography Using an Inorganic Resist/Polymer Bilevel Scheme," *J. Vac. Sci. Technol.,* **17,** 1169 (1980).
24. K. G. Chiong, K. Petrillo, F. J. Hohn, and A. D. Wilson, "Contrast and Sensitivity Enhancement of Resists for High-Resolution Lithography," *J. Vac. Sci. Technol.,* **B6,** 2238 (1988).
25. B. Roland, R. Lombaerts, C. Jakus, and F. Coopmans, "The Mechanism of the DESIRE Process," *Proc. SPIE,* **771,** 69 (1987).
26. G. R. Misium, M. Tipton, and C. M. Garza, "Surface Imaging Lithography at 248 nm," *J. Vac. Sci. Technol.,* **B8,** 1749 (1990).
27. J. M. Moran and D. Maydan, "High Resolution, Steep Profile Resist Patterns," *J. Vac. Sci. Technol.,* **16,** 1620 (1979).
28. L. E. Stillwagon, A. Kornblit, and G. N. Taylor, "Thin Titanium Dioxide Films as Interlayers in Trilayer-Resist Structures," *J. Vac. Sci. Technol.,* **B6,** 2229 (1988).
29. R. F. M. Thorney and T. Sun, "Electron Beam Exposure of Photoresists," *J. Electrochem. Soc.,* **112,** 1151 (1965).
30. H. C. Pfeiffer, "Recent Advances in Electron-Beam Lithography for the High-Volume Production of VLSI Devices," *IEEE Trans. Electron. Dev.,* **ED-26,** 663 (1979).
31. D. R. Herriott, R. J. Collier, D. S. Alles, and J. W. Stafford, "EBES: A Practical Electron Lithographic System," *IEEE Trans. Electron. Dev.,* **ED-22,** 385 (1975).
32. A. J. Speth, A. D. Wilson, A. Kern, and T. H. P. Chang, "Electron-Beam Lithography Using Vector-Scan Techniques," *J. Vac. Sci. Technol.,* **12,** 1235 (1975).

Micron Ex. 1023, p. 352
Micron v. YMTC
IPR2024-00909

33. D. S. Alles, C. J. Biddick, J. H. Bruning, J. T. Clemens, R. J. Collier, E. A. Gere, L. R. Harriott, F. Leone, R. Liu, T. J. Mulrooney, R. J. Nielsen, N. Paras, R. M. Richman, C. M. Rose, D. P. Rosenfeld, D. E. A. Smith, and M. G. R. Thomson, "EBES4: A New Electron-Beam Exposure System," *J. Vac. Sci. Technol.,* **B5,** 47 (1987).
34. G. A. Burns and J. A. Schoeffel, "Performance Evaluation of the ATEQ CORE-2000 Scanning Laser Reticle Writer," *Proc. SPIE,* **772,** 269 (1987).
35. E. Goto, T. Soma, and M. Idesawa, "Design of a Variable-Aperture Projection and Scanning System for Electron Beam," *J. Vac. Sci. Technol.,* **15,** 883 (1978).
36. H. C. Pfeiffer, "Variable Spot Shaping for Electron-Beam Lithography," *J. Vac. Sci. Technol.,* **15,** 887 (1978).
37. A. M. Carroll and J. L. Freyer, "Measuring Performance of the AEBLE™150 Direct-Write e-Beam Lithography Equipment," *Proc. SPIE,* **537,** 25 (1985).
38. M. Fujinami, N. Shimazu, T. Hosokawa, and A. Shibayama, "EB60: An Advanced Direct Wafer Exposure Electron-Beam Lithography System for High-Throughput, High-Precision, Submicron Pattern Writing," *J. Vac. Sci. Technol.,* **B5,** 61 (1987).
39. T. Takigawa, H. Wada, Y. Ogawa, R. Yoshikawa, I. Mori, and T. Abe, "Advanced e-Beam Lithography," *J. Vac. Sci. Technol.,* **B9,** 2981 (1991).
40. K. Meissner, W. Haug, S. Silverman, and S. Sonchik, "Electron-Beam Proximity Printing of Half-Micron Devices," *J. Vac. Sci. Technol.,* **B7,** 1443 (1989).
41. R. Ward, A. R. Franklin, I. H. Lewin, P. A. Gould, and M. J. Plummer, "A 1 : 1 Electronstepper," *J. Vac. Sci. Technol.,* **B4,** 89 (1986).
42. Y. Sakitani, H. Yoda, H. Todokoro, Y. Shibata, T. Yamazaki, K. Ohbitsu, N. Saitou, S. Moriyama, S. Okazaki, G. Matsuoka, F. Murai, and M. Okumura, "Electron-Beam Cell-Projection Lithography System," *J. Vac. Sci. Technol.,* **B10,** 2759 (1992).
43. H. Y. Liu, M. P. deGrandpre, and W. E. Feely, "Characterization of a High-Resolution Novolak Based Negative Electron-Beam Resist with 4 $\mu C/cm^2$ Sensitivity," *J. Vac. Sci. Technol.,* **B6,** 379 (1988).
44. D. F. Kyser and N. S. Viswanathan, "Monte Carlo Simulation of Spatially Distributed Beams in Electron-Beam Lithography," *J. Vac. Sci. Technol.,* **12,** 1305 (1975).
45. M. Parikh, "Self-Consistent Proximity Effect Correction Technique for Resist Exposure (SPECTRE)," *J. Vac. Sci. Technol.,* **15,** 931 (1978).
46. F. Murai, S. Okazaki, N. Saito, and M. Dan, "The Effect of Acceleration Voltage on Linewidth Control with a Variable-Shaped Electron Beam System," *J. Vac. Sci. Technol.,* **B5,** 105 (1987).
47. E. Kratschmer and T. R. Groves, "Resist Heating Effects in 25 and 50 kV e-Beam Lithography on Glass Masks," *J. Vac. Sci. Technol.,* **B8,** 1898 (1990).
48. M. Fujita, K. Shiozawa, T. Kase, H. Hayakawa, F. Mizuno, R. Haruta, F. Murai, and S. Okazaki, "Application and Evaluation of Direct-Write Electron Beam for ASIC's," *IEEE J. Solid State Circuits,* **23,** 514 (1988).
49. F. Murai, Y. Nakayama, I. Sakama, T. Kaga, Y. Nakagome, Y. Kawamoto, and S. Okazaki, "Electron Beam Direct Writing Technology for 64-Mb DRAM LSIs," *J. Appl. Phys.,* **29,** 2590 (1990).
50. N. Samoto, Y. Makino, K. Onda, E. Mizuki, and T. Itoh, "A Novel Electron-Beam Exposure Technique for 0.1 μm T-Shaped Gate Fabrication," *J. Vac. Sci. Technol.,* **B8,** 1335 (1990).
51. K. L. Lee and M. Hatzakis, "Direct Electron-Beam Patterning for Nanolithography," *J. Vac. Sci. Technol.,* **B7,** 1941 (1989).
52. E. M. Clausen, J. P. Harbison, C. C. Chang, H. G. Craighead, and L. T. Florez, "Electron Beam Induced Modification of GaAs Surfaces for Maskless Thermal Cl_2 Etching," *J. Vac. Sci. Technol.,* **B8,** 1830 (1990).

Micron Ex. 1023, p. 353
Micron v. YMTC
IPR2024-00909

53. D. L. Spears and H. I. Smith, "High-Resolution Pattern Replication Using Soft X Rays," *Electron. Lett.,* **8,** 102 (1972).
54. D. Fleming, J. R. Maldonado, and M. Neisser, "Prospect for X-Ray Lithography," *J. Vac. Sci. Technol.,* **B10,** 2511 (1992).
55. R. Viswanathan, R. E. Acosta, D. Seeger, H. Voelker, A. Wilson, I. Babich, J. Maldonado, J. Warlaumont, O. Vladimirsky, F. Hohn, D. Crockatt, and R. Fair, "Fully Scaled 0.5 μm Metal-Oxide Semiconductor Circuits by Synchrotron X-Ray Lithography: Mask Fabrication and Characterization," *J. Vac. Sci. Technol.,* **B6,** 2196 (1988).
56. I. Okada, Y. Saitoh, S. Itabashi, and H. Yoshihara, "A Plasma X-Ray Source for X-Ray Lithography," *J. Vac. Sci. Technol.,* **B4,** 243 (1986).
57. A. Heuberger, "X-Ray Lithography," *Sol. State Technol.,* **29,** 93 (Feb. 1986).
58. B. S. Fay and W. T. Novac, "Advanced X-Ray Alignment System," *Proc. SPIE,* **632,** 146 (1986).
59. R. B. McIntosh, G. P. Hughes, J. L. Kreuzer, and G. R. Conti, "X-Ray Step-and-Repeat Lithography System for Submicron VLSI," *Proc. SPIE,* **632,** 156 (1986).
60. C. N. Archie, J. I. Granlund, R. W. Hill, K. W. Kukkonen, J. A. Leavey, L. G. Lesoine, J. M. Oberschmidt, A. E. Palumbo, C. Wasik, M. Q. Barton, J. P. Silverman, J. M. Warlaumont, A. D. Wilson, R. J. Anderson, N. C. Crosland, A. R. Jorden, V. C. Kempson, J. Schouten, A. I. C. Smith, M. C. Townsend, J. Uythoven, M. C. Wilson, M. N. Wilson, D. E. Andrews, R. Palmer, R. Webber, and A. J. Weger, "Installation and Early Operating Experience with the Helios Compact Synchrotron X-Ray Source," *J. Vac. Sci. Technol.,* **B10,** 3224 (1992).
61. A. C. Chen, C. J. Progler, F. F. Couch, T. A. Gunther, H. Fair, and K. A. Cooper, "First X-Ray Stepper in IBM Advanced Lithography Facility," *J. Vac. Sci. Technol.,* **B10,** 2628 (1992).
62. J. Conway, C. Alcorn, D. Patel, J. Ricker, R. Yandow, L. Hsia, and A. Flamholz, "Fabrication of High-Density SRAM Chips Using Mix-and-Match X-Ray Lithography," *Proc. SPIE,* **1924,** 309 (1993).
63. A. M. Hawryluk and L. G. Seppala, "Soft X-Ray Projection Lithography Using an X-Ray Reduction Camera," *J. Vac. Sci. Technol.,* **B6,** 2162 (1988).
64. M. Komuro, N. Atoda, and H. Kawakatsu, "Ion Beam Exposure of Resist Materials," *J. Electrochem. Soc.,* **126,** 483 (1979).
65. R. L. Seliger, J. W. Ward, V. Wang, and R. L. Kubena, "A High-Intensity Scanning Ion Probe with Submicrometer Spot Size," *Appl. Phys. Lett.,* **34,** 310 (1979).
66. G. Stengel, G. Bosch, A. Chalupka, J. Fegel, R. Fischer, G. Lammer, H. Loscher, L. Malek, R. Nowak, C. Traher, P. Wolf, P. Mauger, A. Shimkunas, S. Sen, and J. C. Wolfe, "Ion Projector Wafer Exposure Results at 5× Ion-Optical Reduction Obtained with Nickel and Silicon Stencil Masks," *J. Vac. Sci. Technol.,* **B10,** 2824 (1992).
67. J. L. Bartelt, "Masked Ion Beam Lithography: An Emerging Technology," *Sol. State Technol.,* **29,** 215 (May 1986).
68. S. Okazaki, "Resolution Limits of Optical Lithography," *J. Vac. Sci. Technol.,* **B9,** 2829 (1991).

PROBLEMS

1. A proximity printer operates with a 10-μm mask-to-wafer gap and a wavelength of 436 nm. Another printer uses a 30-μm gap with wavelength 365 nm. Which offers higher resolution?

Micron Ex. 1023, p. 354
Micron v. YMTC
IPR2024-00909

2. Why must the reticle (mask) used in a wafer stepper be completely free of defects? Why can some defects be tolerated in systems exposing the entire wafer at once?

3. Derive Eq. (6.6),

$$H(\nu) = \frac{2}{\pi}[\cos^{-1}(\nu/\nu_m) - (\nu/\nu_m)\sqrt{1-(\nu/\nu_m)^2}]$$

using a figure and simple geometric considerations.

4. Which projector has a better resolution: a coherent machine or a partially coherent one with a coherency factor 1? The MTF required is better than 0.5.

5. Calculate the wavelength of a 20-keV electron beam.

6. Calculate the current density on the image plane of a LaB_6 cathode at 20 kV. The work function, Richardson constant, and solid angle are 2.4 eV, 40 $A/cm^2 \cdot K^2$, and 10^{-2} rad, respectively.

7. Calculate the total exposure time of a variable-shape electron beam system under the following conditions: resist sensitivity, 2 $\mu C/cm^2$; current density, 10 A/cm^2; field size, 3 mm; chip size, 5 mm $\times$ 10 mm; stage speed, 0.2 s/step; deflection settling, 0.1 μs; shot number per chip, 1×10^8; chip number in a wafer, 100.

8. Compare the chip cost between a stepper and electron direct-writing under the following conditions:

   (*a*) Machine costs of the stepper and the e-beam are S and E, respectively.

   (*b*) Throughput of stepper and e-beam are α and β per hour.

   (*c*) Reticle unit price is C, and n pieces are necessary.

   (*d*) Chip yields, uptimes, process speeds, running costs, depreciation years, and other conditions are the same.

9. Calculate the actual cost of the chip in the previous question if S = \$2M; E = \$3M; α, β = 50, 10/hr; C = \$3K; n = 5000/y; yields are 80%; running cost = 0; uptimes = 5000 hr/y; 100 chips/wafer; and 5-year depreciation.

10. Suppose that an x-ray resist must see a mask modulation greater than or equal to 0.6 to form useful resist images. What is the minimum gold thickness required on an x-ray mask to satisfy this requirement if the exposure wavelength is 4 Å?

Micron Ex. 1023, p. 355
Micron v. YMTC
IPR2024-00909

CHAPTER 7

Etching

Y. J. T. Lii

7.1 INTRODUCTION

Devices are built from a number of different layer materials deposited sequentially. The lithography techniques described in Chapter 6 are used to replicate circuit and device features, and the desired pattern is transferred by means of etching. As we enter the era of ULSI manufacturing, the etching process grows more important for fabricating features with sub-half-micrometer dimensions. Dry etching is synonymous with plasma-assisted etching, which encompasses several techniques that use plasmas in the form of low-pressure gas discharges. Integrated circuit fabrication processes that use reactive plasmas are commonplace in today's semiconductor production lines because of their potential for very-high-accuracy transfer of resist patterns, i.e., anisotropic etching, as shown in Fig. 1. By contrast, wet chemical etching results in isotropic etching, where both vertical and lateral etch rates are comparable.

In the early 1970s, photo lithography with Novolak-based resists was used to pattern devices with dimensions less than three micrometers. Oxygen plasmas were explored for stripping of resist. The need to maintain dimensional tolerance during etching became a critical requirement that led to numerous advances in plasma processing and improved reactors during the 1980s. Today, there is a broad base of empirical knowledge and some qualitative understanding of etch mechanisms.[1] These insights and systematic experiments led to the process developments needed to meet the requirements of today's submicrometer device manufacturing.

Etch techniques consist of dry and wet etch methods. Dry etch methods include plasma etching, reactive ion etching, sputtering, ion beam etching, and reactive ion beam etching. For dry etch methods, this chapter concentrates only on plasma and reactive ion etching because they are the most popular methods used in semiconductor industry. This chapter starts with a discussion of the fundamentals of

Micron Ex. 1023, p. 356
Micron v. YMTC
IPR2024-00909

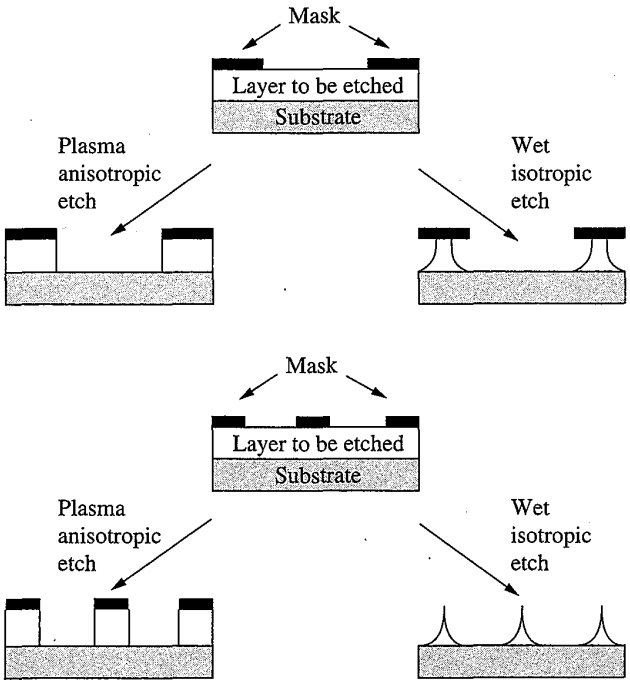


FIGURE 1
Illustration of the difference between plasma (anisotropic) and wet (isotropic) etching.

low-pressure discharge. Next, fluorocarbon plasma chemistry is introduced to review the means for achieving etching directionality and selectivity in reactive etching using glow discharges. Specific plasma processes used to etch different materials important to semiconductor processing are reviewed. Plasma diagnostics, process control, and plasma-induced damage are introduced briefly. In addition, relevant trends in new, magnetically enhanced reactive ion etching systems, in high-density plasma systems, and in integrated etching clustering are discussed. Finally, alternative wet etching methods used in the semiconductor industry are briefly presented.

7.2 LOW-PRESSURE GAS DISCHARGE

Plasma is an ionized gas with approximately equal amounts of positively charged particles (positive ions) and negatively charged particles (usually electrons and a small amount of negative ions). The plasma useful to ULSI processing is a weakly ionized plasma, called a "glow discharge," containing a significant density of neutral particles—more than 90% in most etchers. Although plasma is neutral in a macroscopic sense, it behaves quite differently from a molecular gas, because it consists of charged particles that can be influenced by applied electric and magnetic fields.

Plasma etching systems are extremely complex, and many details of the physical and chemical interactions, both within the plasma and on the surface exposed to the plasma, are not yet understood. The interaction of plasma with surfaces will be discussed in order to describe some of the mechanisms thought to be important in etching. A comprehensive review of this subject can be found in Ref. 2.

Micron Ex. 1023, p. 357
Micron v. YMTC
IPR2024-00909

7.2.1 AC Plasma Discharge

In plasma etching, a plasma is created by applying a voltage across two electrodes, between which a gas is confined at low pressure. Simple dc power can be used to generate plasma; however, insulating materials cannot be used to cover electrode surfaces, because the insulating materials stop the current flow between the two electrodes and cause surface charging. To deal with the surface charging problem, an ac discharge is used so that the positive charge accumulated during one half-cycle can be neutralized by electron bombardment during the next half-cycle. In plasma etching, an rf field is normally used to generate the gas discharge. One reason for doing so is that the electrodes do not have to be made of a conducting material. The other reason is that electrons can pick up sufficient energy during field oscillation to cause more ionization by electron–neutral atom collisions. As a result, the plasma can be generated at pressures as low as 10^{-3} torr.

The free electrons in the gas are accelerated by the electric field and gain energy. Energy is released when free electrons collide with gas atoms; the gas atoms are then ionized, and more electrons are freed. Electron collisions with neutrals in plasma cause dissociation, ionization, and light emission. These collisions generate atoms, molecules, molecular fragments, and ions. The ions and electrons generated from these collisions raise the density of the plasma. Additional electrons are generated by secondary emission from energetic, positive ions colliding with the electrode surfaces. However, some collisions on the electrode surfaces lower plasma density, such as an ion-ion recombination and an electron-ion recombination. The process continues until equilibrium is reached, that is, until the number of ion-electron pairs created is the same as the number of pairs lost at the electrode surface and the chamber wall. At this point, the plasma is self-sustaining. The density for the plasmas of interest ranges from 10^9 to 10^{12} cm^{-3}. The density of gas molecules at 1 torr is about 10^{16} cm^{-3}, so these discharges are weakly ionized. Therefore, at low pressure, plasma is a partially ionized gas containing electrons, ions, and neutral atoms and/or molecules.

7.2.2 Properties of RF Discharge

When a plasma comes in contact with an electrode or surface, a dark space or *sheath* is formed next to it. The potentials of the rf discharge near the electrodes are important in determining the energies of ions incident on surfaces in the plasma.[3] The potentials near the electrodes, as shown in Fig. 2, result from currents of different particles, specifically electrons and ions. A light particle such as an electron will accelerate more than a heavier particle such as a positive ion in a given electric field. A large number of electrons rush toward electrodes after an electric field is applied. The depletion of electrons from the plasma and a negative charge buildup on the electrodes are inhibited by the development of a negative charge barrier that repels electrons and attracts positive ions. In Fig. 2, V_c is the potential at the rf-powered electrode measured with respect to ground, and V_p is the plasma potential with respect to ground. The potentials across the ion sheaths are $V_p - V_c$ at the rf-powered electrode

Micron Ex. 1023, p. 358
Micron v. YMTC
IPR2024-00909

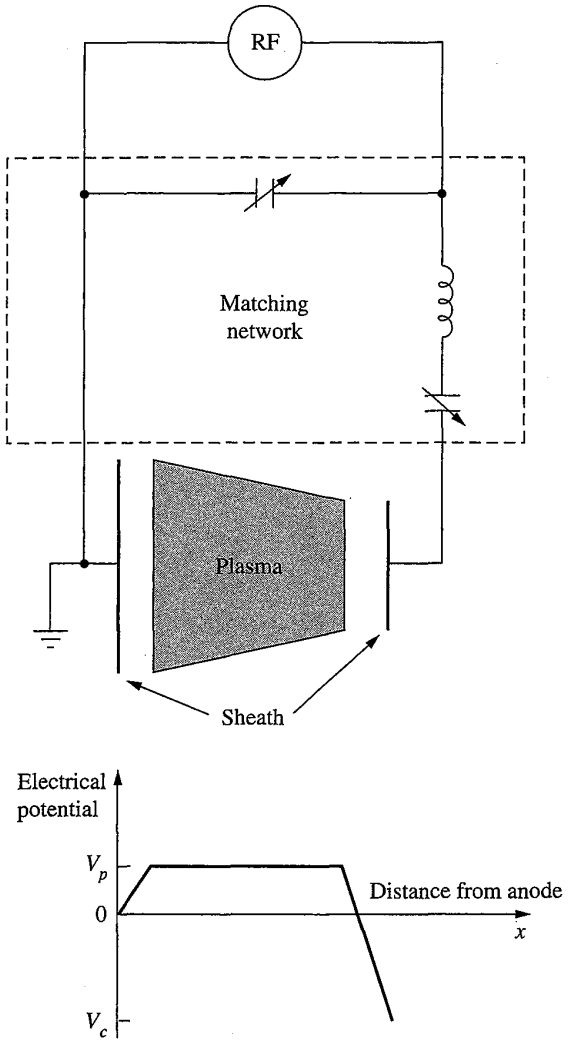


FIGURE 2
Representation of a generic, reactive-ion plasma etcher. RF power is supplied to the electrode supporting the wafer by a generator and matching network. The potential is shown as a function of position in the discharge for the case in which the area of the powered electrode is much less than the area of all grounded surfaces.

and V_p at the grounded surface. The potential of the electrode or surface with respect to that of the plasma determines the maximum energy of the ions bombarding that electrode or surface.

To a first approximation, the sheaths are treated as capacitors. The current density of the positive ions across sheaths is uniform and is equal at both electrodes. The ratio, R, of the area of the rf-powered electrode to the area of all grounded surfaces in contact with the plasma is a key parameter in determining how the applied voltage is distributed among the plasma sheaths. The potential $V_p - V_c$ increases as R decreases. A classic theory of plasma characteristics can be expressed as[2]

$$\frac{V_p - V_c}{V_p} = \left(\frac{A_a}{A_c}\right)^4 \tag{7.1}$$

where A_a and A_c are the grounded surface and rf-powered electrode areas, respectively. The sheath voltage near the small rf-powered electrode is larger than the

Micron Ex. 1023, p. 359
Micron v. YMTC
IPR2024-00909

sheath voltage near the grounded surface. In a typical diode reactive-ion etcher, $V_p - V_c$ can be in the several-hundred-volt range when V_p is less than 100 volts.

7.2.3 Matching Network

A matching network is usually located between the rf power supply and the plasma, as shown in Fig. 2. The purpose of this network is to increase the power dissipation in the discharge and to protect the rf power supply. An rf power supply normally has a purely resistive output with a usual value of 50 ohms. However, an rf plasma normally has a partly capacitive impedance. Therefore, a load equal to the impedance of the rf power-supply output can be simulated by combining the plasma impedance with a variable matching network load. Most plasma etching systems lose power in the matching network; but with careful system design, the power loss can be minimized.

7.2.4 Glow Discharge Chemical Phenomena

Electron energies achieved at low pressure can be so high that they dissociate molecules into free radicals, ions, and new electrons, thus cascading the electrical breakdown of the gas and igniting or sustaining the plasma. At rf frequencies higher than 1 MHz, the ions, because of their low mobility, do not respond to the rf field instantaneously. Electrons undergo primarily inelastic electron-molecule collisions, which supply fresh species to the plasma. The steady-state constitution of a glow discharge is governed by the equilibrium of the rates of production and loss of the various species.

Two main reactions take place in a plasma system:

1. Homogeneous reactions in the bulk plasma, caused mainly by electron impact with neutral species
2. Heterogeneous reactions at the surface contact with the plasma, initiated by ions, electrons, and neutrals

Some examples of the most important types of electron–neutral particle collisions are summarized here:

Simple ionization

$$Ar + e^- \rightarrow Ar^+ + 2e^- \tag{7.2}$$

$$O_2 + e^- \rightarrow O_2^+ + 2e^- \tag{7.3}$$

In molecular gases, ionization may be concurrent with fragmentation; in this case, dissociative ionization is said to occur.

Dissociation

$$CF_4 + e^- \rightarrow CF_3^+ + F + 2e^- \tag{7.4}$$

$$CF_4 + e^- \rightarrow CF_3^+ + F^- + e^- \tag{7.5}$$

Light emission

$$O_2 + e^- \rightarrow O_2^* + e^- \tag{7.6}$$

$$O_2^* \rightarrow O_2 + h\nu \tag{7.7}$$

Micron Ex. 1023, p. 360
Micron v. YMTC
IPR2024-00909

In the plasma, each formation step balances various loss processes. The equilibrium between formation and loss determines the steady concentrations of species in a discharge. For the charged species in a plasma, these formation and loss processes may be grouped into a few categories as follows:

Electron-ion recombination $O_2^+ + e^- \rightarrow 2O$ (7.8)

Ion-ion recombination $CF_3^+ + F^- \rightarrow CF_4$ (7.9)

Dissociative attachment $e^- + CF_4 \rightarrow CF_3 + F^-$ (7.10)

Atoms and radicals can be lost either by homogeneous reactions or by heterogeneous reactions. The type of loss reaction that will dominate for any species depends on many factors such as pressure, the type of surface, and the ratio of surface area to the volume of discharge.

7.3 ETCH MECHANISMS, SELECTIVITY, AND PROFILE CONTROL

Much of the early work in plasma etching focused on the discovery and optimization of fluorocarbon chemistry.[4] Although this initial work provided many useful industrial etching processes, an understanding of the associated etch mechanisms was typically lacking. Plasma etching is a process in which a solid film is removed by a chemical reaction with ground-state or excited-state neutral species, and plasma etching is often enhanced or induced by energetic ions generated in a gaseous discharge. Some important processing issues discussed in this section include basic etch mechanisms, etch selectivity, and profile control.

7.3.1 Etch Mechanisms

Unlike ion sputtering, which can be compared to sandblasting on an atomic scale, plasma etching methods are based on chemical reactions between the solid to be etched and an active species that arrives on the surface of the solid from the plasma. The microscopic processes of plasma etching of a silicon wafer are shown schematically in Fig. 3.[5] Electrons are accelerated by an applied rf field, and the energetic electrons collide with gas atoms and molecules. Some atoms and molecules become electronically excited, and light emission (glow) is produced from their deexcitation. Other atoms and molecules are dissociated or ionized to form radicals, atoms, and ions by colliding with high-energy electrons. The active species are then transported to the wafer surface where they are adsorbed. The adsorbed species can react with the wafer to form etch products or desorb from the wafer surface without reaction. Etch products then desorb into a gas phase if they are volatile. A brief summary of processes that occur during plasma etching is that a molecular gas, usually relatively inert, is used to generate reactive species, such as F or Cl radicals. By choosing proper chemistry, the radicals react with the wafer surface to form "volatile" products. The chemical compound formed must be volatile with a high vapor pressure

Micron Ex. 1023, p. 361
Micron v. YMTC
IPR2024-00909

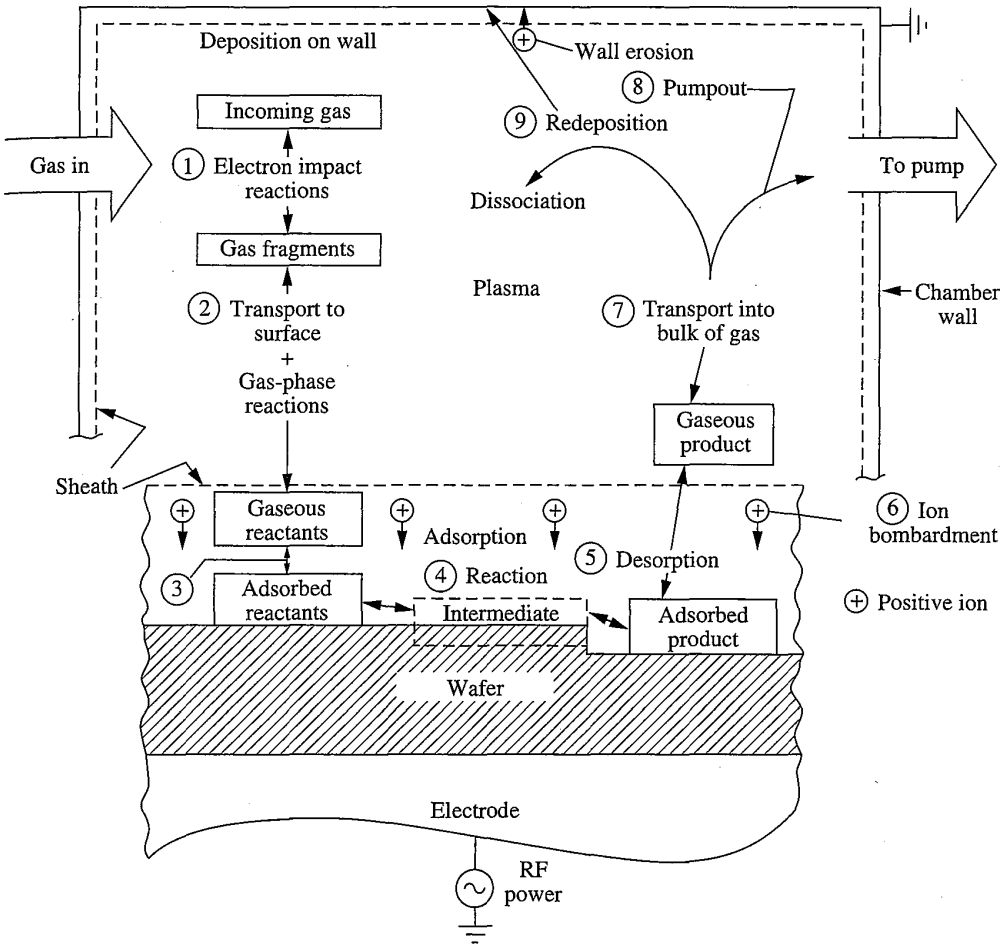


FIGURE 3
Schematic view of microscopic processes that occur during etching of a silicon wafer. (*After Oehrlein and Rembetski, Ref. 5.*)

and stable with a high bonding energy so that it is not dissociated in the plasma and, consequently, can be pumped out of the reactor.

Plasma etching comprises chemical, ion-sputter, and ion-enhanced plasma etching. All these methods are based on the generation of plasma by an rf discharge in a gas at low pressure. But two basic methods can be distinguished—physical methods and chemical methods. The former include sputter etching and the latter include pure chemical etching. There is also a hybrid technique, which uses a physical method to enhance chemical etching. In chemical etching, neutral reactive species generated by the plasma interact with the material's surface to form volatile products. In physical etching, positive ions bombard the surface at high speed; small amounts of negative ions formed in the plasma cannot reach the wafer surface and therefore play no direct role in plasma etching. Chemical and physical etch mechanisms have different basic characteristics. Chemical etching exhibits a high etch rate and good selectivity, produces low ion-bombardment-induced damage, and yields isotropic profiles. Physical etching can yield anisotropic profiles, but it is associated with low

Micron Ex. 1023, p. 362
Micron v. YMTC
IPR2024-00909

etch selectivity and high bombardment-induced damage. Combinations of chemical and physical etching give anisotropic etch profiles, reasonably good selectivity, and moderate bombardment-induced damage.

7.3.2 Fluorocarbon Chemistry

In the early 1980s, molecular gases containing one or more fluorine atoms were used for plasma etching of silicon and silicon compounds. These gases were selected because they are often nontoxic, and fragments produced in a plasma react with silicon or silicon compounds to form volatile compounds at room temperature. For example, CF_4 is extremely stable but dissociates into F atoms and fluorinated fragments (CF_x) in a plasma. Atomic F is the active etchant for Si and SiO_2 by the formation of the volatile SiF_4 and O_2.

$$Si + 4F \rightarrow SiF_4 \tag{7.11}$$

$$SiO_2 + 4F \rightarrow SiF_4 + O_2 \tag{7.12}$$

Energetic electrons colliding with CF_4 are responsible for the generation of radicals, atoms, and ions. However, these radicals and atoms can react with other species, such as gas molecules, to alter the overall chemistry. For example, the etch rates of Si increase dramatically, as shown in Fig. 4,[6] if O_2 is added to CF_4. A maximum

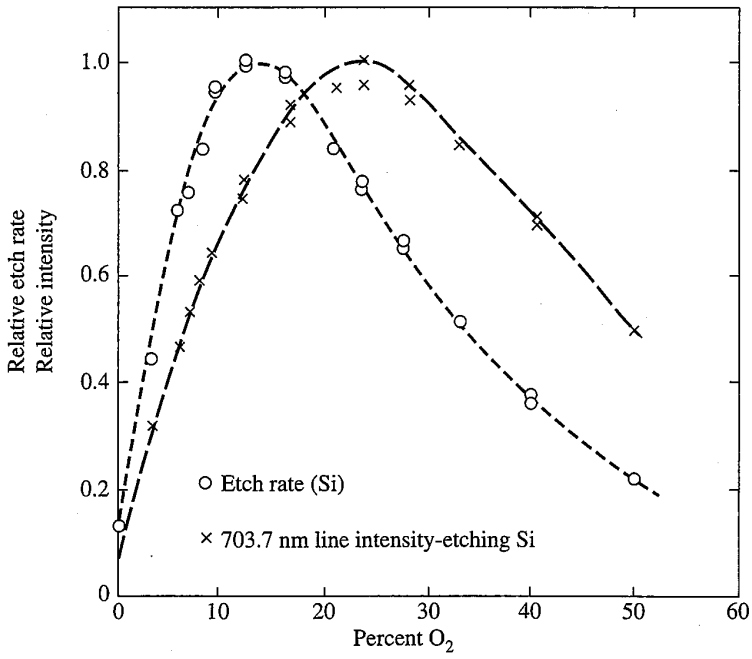


FIGURE 4
Dependence of both the normalized silicon etch rate and the normalized intensity of the emission from electronically excited F atoms on oxygen concentration in the CF_4/O_2 mixture. (*After Mogab, Adams, and Flamm, Ref. 6.*)

Micron Ex. 1023, p. 363
Micron v. YMTC
IPR2024-00909

etch rate occurs with a CF_4/O_2 mixture that has 12% O_2. A similar effect is observed for SiO_2, where the maximum etch rate is reached at 20% O_2. These effects can be explained by the decreasing recombination rate of F and CF_x ($x \leq 3$). Adding O_2 to the plasma leads to the formation of COF_2, CO, and CO_2, which decreases the concentration of CF_x, thereby decreasing the recombination rate of F and CF_x. This results in a higher F concentration and, therefore, higher Si and SiO_2 etch rates. F concentration increases until the effect of O_2 dilution by CF_4 becomes dominant. This point is reached more quickly with Si etching because oxygen chemsorbs on the Si surface more readily than on the SiO_2 surface thereby blocking the reaction of F with Si.

When H_2 is added to the CF_4 plasma, quite different etch rate changes are observed. Figure 5 shows the etch rates of Si and SiO_2 in a CF_4/H_2 plasma as a function of H_2 concentration.[7] The SiO_2 etch rate is relatively insensitive to the H_2 concentration, but the Si etch rate decreases significantly as the H_2 concentration increases. When the H_2 concentration reaches 40%, the silicon etch stops. A generally accepted explanation of these effects is that two main mechanisms result from H_2 addition. First, H_2 reacts with F to form HF. This reaction consumes the F concentration and decreases the Si etch rate.

$$H + F \rightarrow HF \tag{7.13}$$

Second, the reaction of CF_x with SiO_2 forms etch products SiF_4, CO, CO_2, COF_2, etc. However, the reaction of CF_4 with Si forms a carbon or hydrocarbon polymer on the silicon surface, which blocks silicon etching by F.

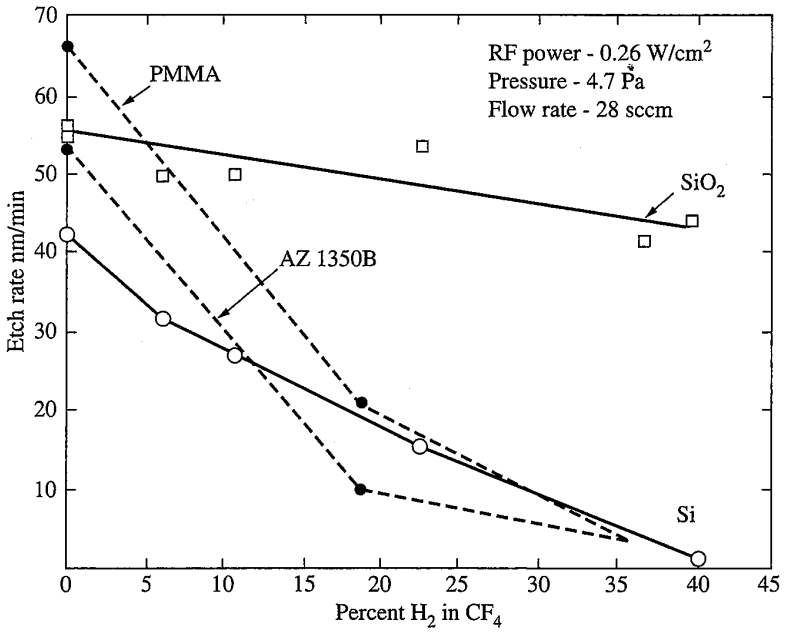


FIGURE 5
Etch rates of silicon (solid line, round points), silicon dioxide (solid line, square points), and PMMA and AZ1350B resists (dashed lines) vs. percentage of H_2 in CF_4. (*After Ephrath, Ref. 7.*)

Micron Ex. 1023, p. 364
Micron v. YMTC
IPR2024-00909

$$CF_4 + SiO_2 \rightarrow SiF_4 + (CO, CO_2, COF_2) \tag{7.14}$$

$$CF_4 + Si \rightarrow SiF_x + (C, CF_y, CH_xF_y) \tag{7.15}$$

Similarly, CHF_3 can be used to replace the CF_4/H_2 mixture for selective silicon dioxide–to-silicon etch. A mixture of CHF_3/CF_4 is often used in industry for selective silicon dioxide–to-silicon etch, because etch selectivity can be controlled by changing the ratio of CHF_3 to CF_4 concentrations.

7.3.3 Anisotropy

Etch profiles are extremely important for submicrometer semiconductor fabrication, and careful control of the surface etch processes is required to ensure directional etching. Etch mechanisms can be understood by analyzing surface etching processes. Plasma consists of neutrals, ions, and electrons, and the effects of each species on etch results are not easily identified. Much of the fundamental understanding of the surface processes that occur during plasma etching comes from well-controlled plasma experiments in which beams of ions, electrons, and neutrals are directed, either together or alternately, at well-characterized surfaces, using apparatus such as that shown in Fig. 6.[8] Etch products are usually analyzed by mass spectrometry. The surface compositions are diagnosed by in-situ or ex-situ x-ray photoelectron spectroscopy, Auger spectroscopy, and secondary ion mass spectrometry. Without radiation or any external energy, the probability of a reaction of Si with F is approximately 10^{-2}, which could be enhanced by ion-beam bombardment. Figure 7 shows the etch rate of silicon with inert Ar ions and with reactive flux to the silicon surface. This etch rate is larger than the sum of the individual etch rates for ion sputtering and neutral etching because chemical etching is enhanced by ion bombardment.[9] Without experiments like this, it is difficult to identify the dominant etch mechanism.

The precise anisotropic etch mechanisms for any given material system are not fully understood, but it is clear that the ions striking the surface must themselves be anisotropic, that is, traveling mainly in a direction perpendicular to the wafer surface. Ions are oriented and accelerated in a sheath. Two proposed mechanisms for anisotropic plasma etching are as follows:

1. Perpendicular ion bombardment creates a damaged surface that is then more reactive toward neutral etchants.
2. Ions help to desorb etch-inhibiting species, such as etch products, from the surface.

In both mechanisms, ion transport must be perpendicular to the surface so that only the etch rate of the bottom surface is enhanced. An illustration of these anisotropic mechanisms is shown in Fig. 8.[4]

Ion–neutral particle collisions from the plasma sheath result in a fraction of the ions hitting the sidewall, and some lateral etching may occur. The number of ion–neutral collisions in the sheath is directly proportional to the sheath thickness and inversely proportional to the ion mean free path. Because the ion mean free path is

Micron Ex. 1023, p. 365
Micron v. YMTC
IPR2024-00909

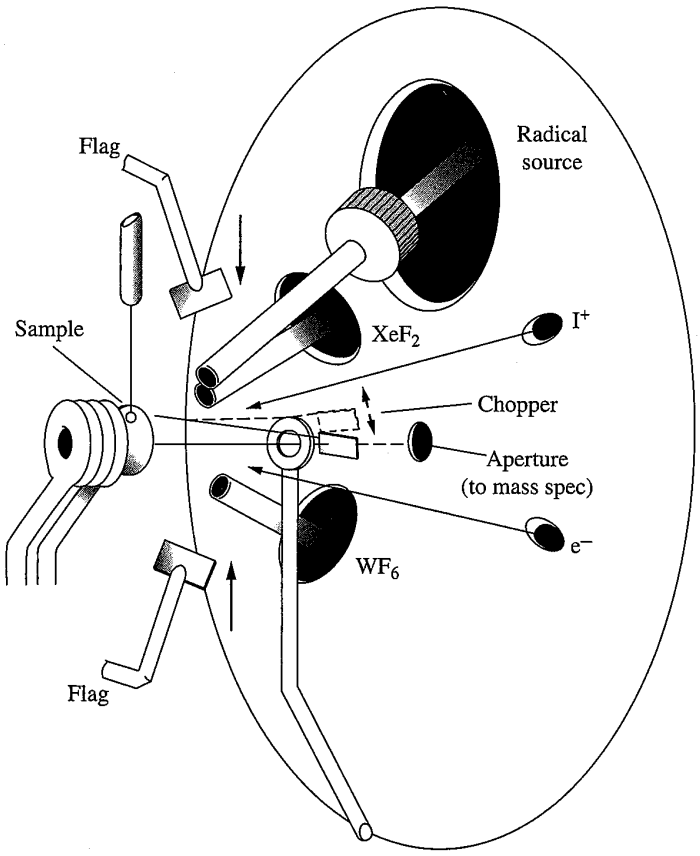


FIGURE 6
Beam-surface experimental apparatus used for the simulation of plasma etching. Such experiments are important for understanding mechanisms of surface-plasma interactions. (*After Winters and Plumb, Ref. 8.*)

usually proportional to pressure, lowering the pressure reduces ion–neutral collisions and enhances anisotropic etching.

In mechanism (2), the formation of sidewall films during etching plays a critical role in the development of the anisotropic etch profile; the films protect the sidewalls from etching by nonperpendicular incoming ions. In other cases, such as during silicon trench etching, thick deposits, often in the upper area of the trench, lead to progressive narrowing of the trench opening and a round bottom.[10] Various mechanisms contribute to sidewall film growth. Three cases are important in the formation of sidewall films:[11,12]

1. Redeposition of etch products or sputtered material from the bottom of the etched structure by adsorption and recombination
2. Deposition from glow discharge
3. Redeposition of sputtered mask material

A schematic description of sidewall deposition phenomena during silicon trench etching by HBr is shown in Fig. 9. Sidewall passivation is governed by the

Micron Ex. 1023, p. 366
Micron v. YMTC
IPR2024-00909

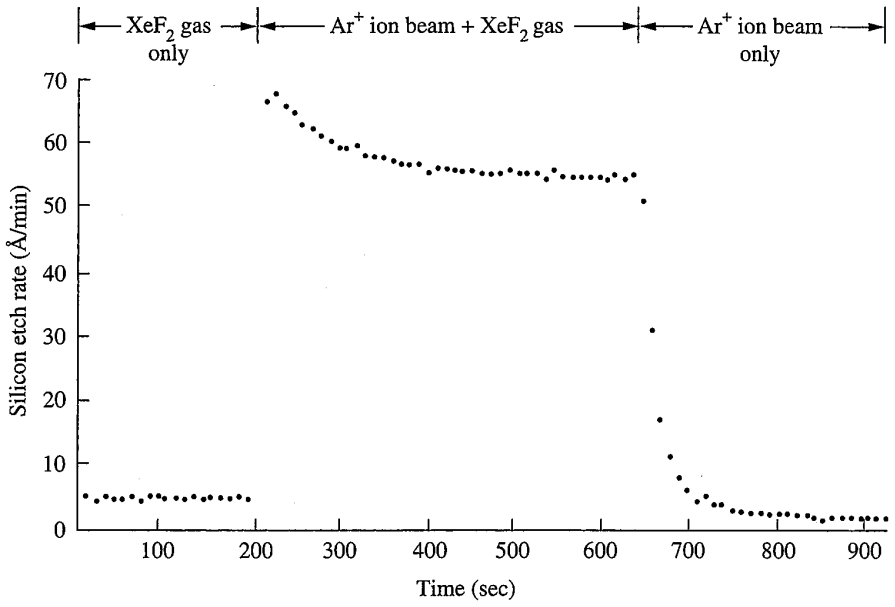


FIGURE 7
Etch rate of silicon as a function of time. This graph illustrates ion-enhanced chemical etching of silicon by an Ar^+ ion beam and XeF_2 flux. (*After Coburn and Winters, Ref. 9.*)

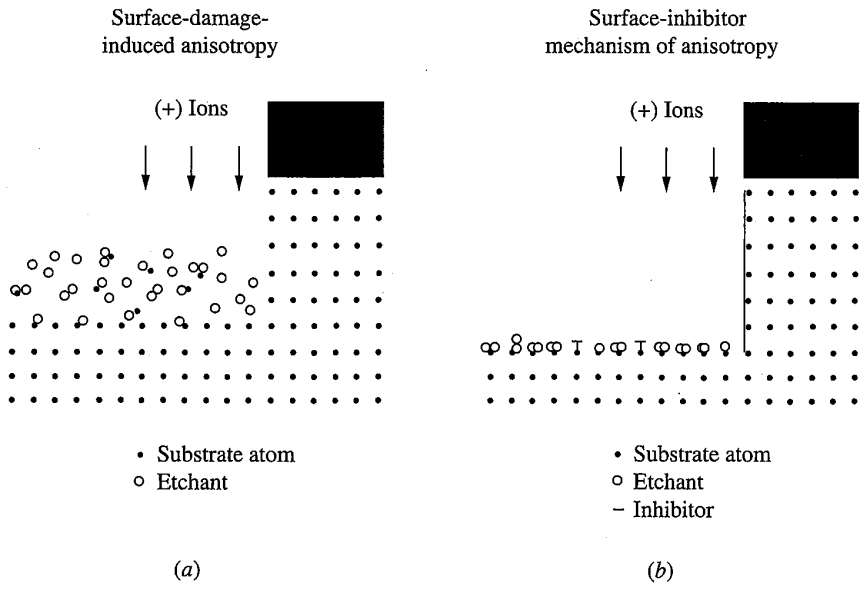


FIGURE 8
Schematic of two proposed mechanisms for ion-enhanced surface etching. Ion-enhanced surface etching results in anisotropic etch profiles. (*After Flamm and Donnelly, Ref. 4.*)

Micron Ex. 1023, p. 367
Micron v. YMTC
IPR2024-00909

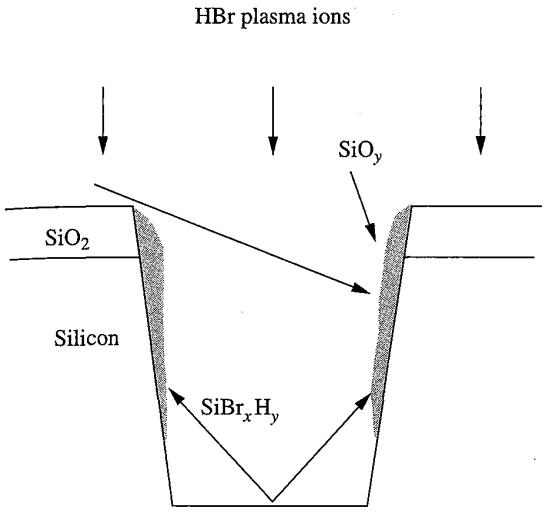


FIGURE 9
Schematic of sidewall formation mechanisms during silicon trench etching by HBr plasma.

equilibrium between deposition and ionic sputtering on vertical surfaces. Sidewall film thickness control becomes critical if tapered sidewalls and dimensional loss is not desired. For example, if the passivation layer becomes too thick during polysilicon gate etching, the gate dimensions increase and problems result.

7.3.4 Selectivity

Selectivity is defined as one film etching faster than another film under the same etching conditions. A higher etch rate ratio (ERR) between different layers is the crucial advantage of reactive ion etching over physical sputtering. The etch rate difference between two different materials is due to different surface etch mechanisms, such as adsorption, reaction, and desorption. During etching, the selectivities with respect to the masking material and the underlying layer require careful control. The required selectivity is defined according to a specified percentage of overetch and film thickness. The gate oxide thickness is in the range of 10 nm for today's 0.8-micrometer, complementary metal-oxide-silicon (CMOS) technologies, and an etch selectivity of polysilicon to thin gate oxide of more than 10 : 1 is required. In the case of a contact hole etch, a large amount of overetch time is required because of the different thicknesses of the oxide layers over the polysilicon gate and over the source and drain. Therefore, a high etch selectivity between silicon and oxide is required for oxide contact hole etch. Thinner masks, such as photoresist, are preferred for optical lithography. A thinner mask film can be used in the etching if a high ERR between the etched film and the mask is obtainable.

Selectivity usually results from two main mechanisms. First, the gas chemistry chosen should be able to generate thermodynamically favorable reactions with the primary film. For example, chlorine-based plasma is usually employed for a polysilicon–oxide etch, and the selectivity arises from the fact that Cl atoms etch SiO_2 very slowly. Etching SiO_2 with chlorine is not a thermodynamically favored reaction. Second, gas chemistry is selected to produce a volatile etch product with

Micron Ex. 1023, p. 368
Micron v. YMTC
IPR2024-00909

the primary film. For example, to etch SiO_2 over aluminum, a fluorine-based chemistry is used because of the high vapor pressure of SiF_4 from the SiO_2 etching. A similar principle has also been applied to SiO_2-over-silicon etching, and the selectivity is usually achieved by forming a passivation layer on silicon. Fluorocarbon polymer forms are used to etch oxide selectively with respect to silicon. Carbon reacts with oxygen during oxide etch, and a fluorocarbon polymer film forms on the silicon surface and provides a good selectivity.

7.3.5 Microscopic Uniformity

Some problems in achieving microscopic uniformity occur because etching rates and profiles depend on feature size and pattern density. Microscopic uniformity problems can be grouped into two categories—aspect ratio–dependent etching (ARDE) and pattern-dependent etching, known as microloading.[13] The cause of the problems is ion and neutral transport.

A typical trench is shown in Fig. 10 with an aspect ratio of 8. Significant ARDE is seen, and trenches with a large aspect ratio etch more slowly than trenches with a small aspect ratio,[14] as shown in Fig. 11. This phenomenon became serious when the

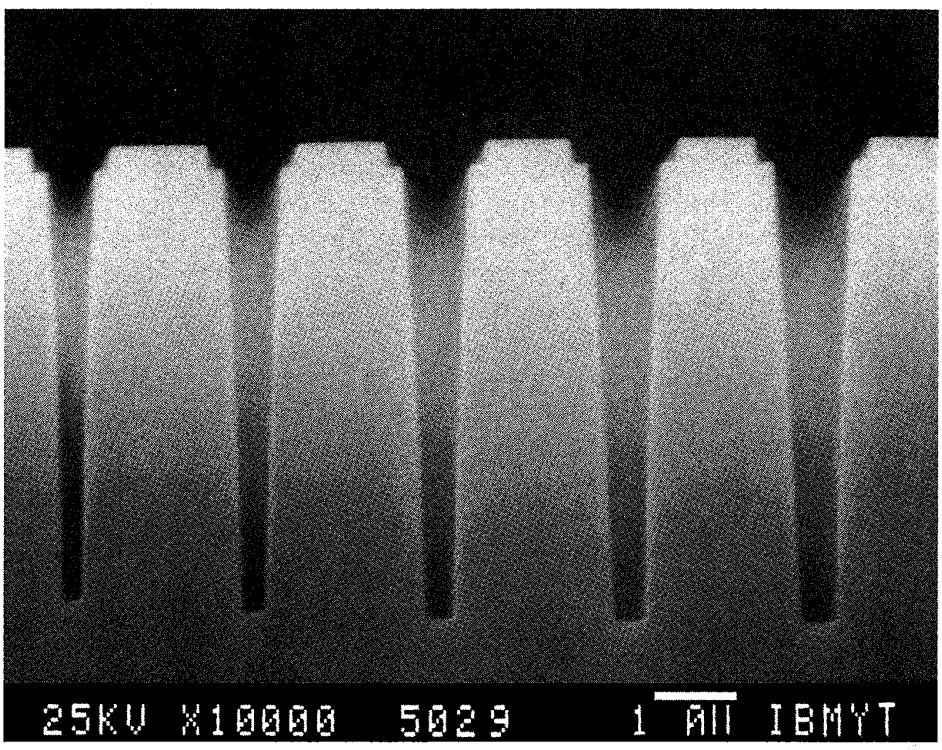


FIGURE 10
Effect of aspect ratio–dependent etching on a silicon trench etch. The etching rate decreases as the trench dimensions are reduced.

Micron Ex. 1023, p. 369
Micron v. YMTC
IPR2024-00909

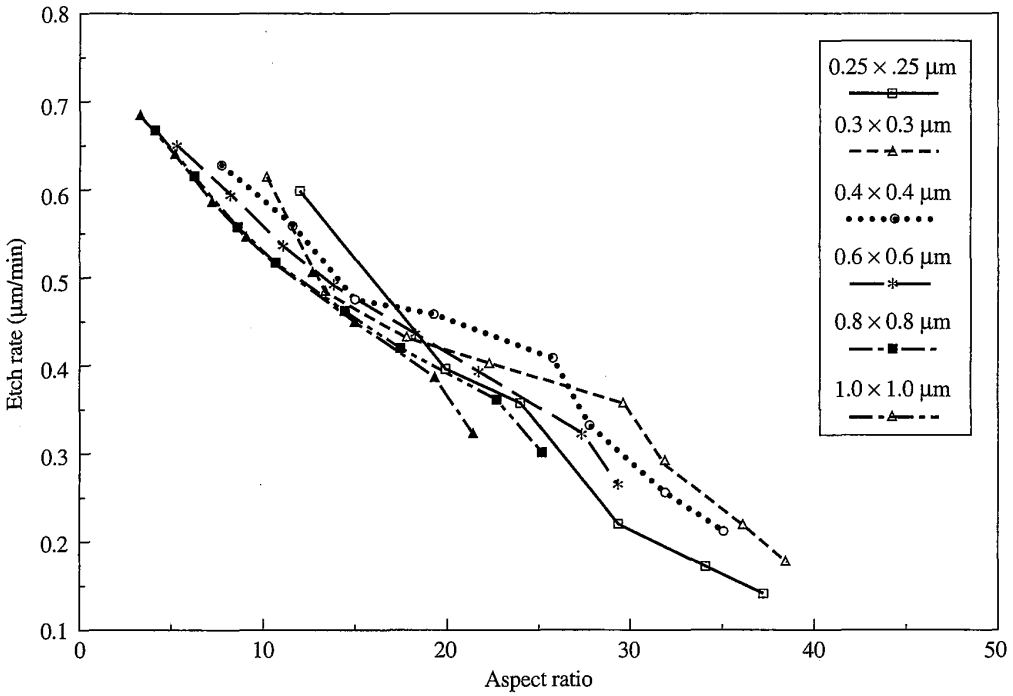


FIGURE 11
Dependence of average silicon trench etch rate on aspect ratio.

era of submicrometer etching began in recent years. Ion bombardment, electron bombardment, reactive neutral species, product desorption, and redeposition all appear to be important in determining the relative etch rates of trenches. Four different mechanisms are the main contributors to ARDE: limited Knudsen transport of neutrals, ion shadowing from ion–neutral collisions in the plasma sheath, neutral shadowing caused by geometrical effects, and ion scattering from insulator charging.[15,16,17]

The term *microloading* usually refers to the dependence of the etching rate on pattern density for identical features.[18] Microloading results from depletion of reactants because the wafer has a local, higher-density unmasked area.[14]

7.4
REACTIVE PLASMA ETCHING TECHNIQUES AND EQUIPMENT

Plasma reactor technology in the electronics industry has changed dramatically since the first application of plasma processing to photoresist stripping. A reactor for plasma etching contains a vacuum chamber, pump system, power supply generators, pressure sensors, gas flow control units, and an end-point detector. Plasma etch tools may be categorized according to the etch mechanism each of them employs. Table 1 shows the similarities and differences in the types of etch equipment that are commercially available. A comparison of pressure operating ranges and ion energies

Micron Ex. 1023, p. 370
Micron v. YMTC
IPR2024-00909

the primary film. For example, to etch SiO_2 over aluminum, a fluorine-based chemistry is used because of the high vapor pressure of SiF_4 from the SiO_2 etching. A similar principle has also been applied to SiO_2-over-silicon etching, and the selectivity is usually achieved by forming a passivation layer on silicon. Fluorocarbon polymer forms are used to etch oxide selectively with respect to silicon. Carbon reacts with oxygen during oxide etch, and a fluorocarbon polymer film forms on the silicon surface and provides a good selectivity.

7.3.5 Microscopic Uniformity

Some problems in achieving microscopic uniformity occur because etching rates and profiles depend on feature size and pattern density. Microscopic uniformity problems can be grouped into two categories—aspect ratio–dependent etching (ARDE) and pattern-dependent etching, known as microloading.[13] The cause of the problems is ion and neutral transport.

A typical trench is shown in Fig. 10 with an aspect ratio of 8. Significant ARDE is seen, and trenches with a large aspect ratio etch more slowly than trenches with a small aspect ratio,[14] as shown in Fig. 11. This phenomenon became serious when the

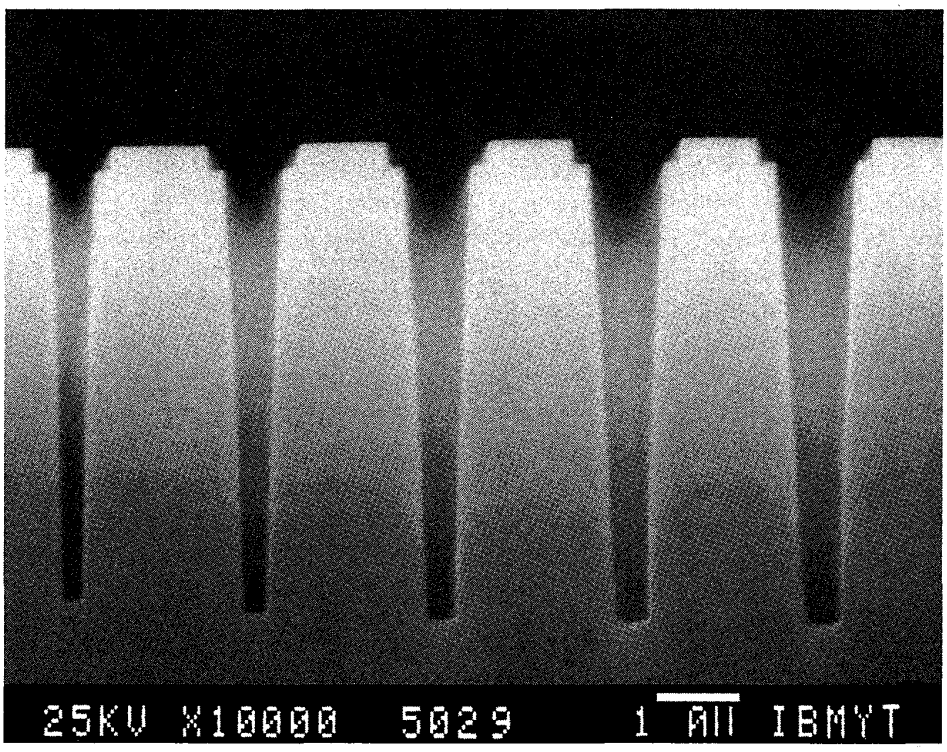


FIGURE 10
Effect of aspect ratio–dependent etching on a silicon trench etch. The etching rate decreases as the trench dimensions are reduced.

Micron Ex. 1023, p. 371
Micron v. YMTC
IPR2024-00909

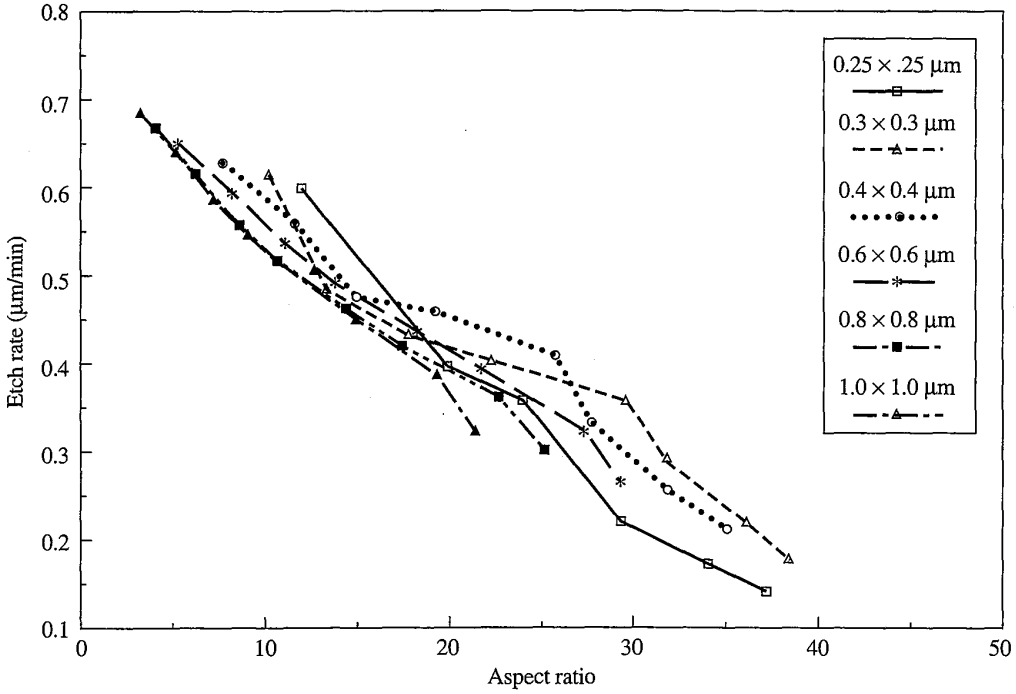


FIGURE 11
Dependence of average silicon trench etch rate on aspect ratio.

era of submicrometer etching began in recent years. Ion bombardment, electron bombardment, reactive neutral species, product desorption, and redeposition all appear to be important in determining the relative etch rates of trenches. Four different mechanisms are the main contributors to ARDE: limited Knudsen transport of neutrals, ion shadowing from ion–neutral collisions in the plasma sheath, neutral shadowing caused by geometrical effects, and ion scattering from insulator charging.[15,16,17]

The term *microloading* usually refers to the dependence of the etching rate on pattern density for identical features.[18] Microloading results from depletion of reactants because the wafer has a local, higher-density unmasked area.[14]

7.4
REACTIVE PLASMA ETCHING TECHNIQUES AND EQUIPMENT

Plasma reactor technology in the electronics industry has changed dramatically since the first application of plasma processing to photoresist stripping. A reactor for plasma etching contains a vacuum chamber, pump system, power supply generators, pressure sensors, gas flow control units, and an end-point detector. Plasma etch tools may be categorized according to the etch mechanism each of them employs. Table 1 shows the similarities and differences in the types of etch equipment that are commercially available. A comparison of pressure operating ranges and ion energies

Micron Ex. 1023, p. 372
Micron v. YMTC
IPR2024-00909

TABLE 1
Etch mechanisms, pressure ranges, and electrode arrangements of plasma reactors

Etch tool configuration	Etch mechanism	Pressure range (torr)	Electrode arrangement
Barrel etching	Chemical	0.1 ~ 10	Electrodes outside barrel
Downstream plasma etching	Chemical	0.1 ~ 10	Electrodes located in another chamber
Reactive ion etching	Chemical and physical	0.01 ~ 1	Planar diode, triode, cylindrical diode
Magnetic enhanced RIE	Chemical and physical	0.01 ~ 1	Planar diode with magnetic field parallel to the cathode
Magnetic confinement triode RIE	Chemical and physical	0.001 ~ 0.1	Planar diode or triode with magnetic field confinement
Electron cyclotron resonance etching	Chemical and physical	0.001 ~ 0.1	External microwave excitation with biased electrodes
ICP or helicon plasma etching	Chemical and physical	0.001 ~ 0.1	External RF excitation with biased electrodes

for different types of reactors is shown in Fig. 12. Each etch tool is designed empirically and uses a particular combination of pressure, electrode configuration and type, and source frequency to control the two primary etch mechanisms—chemical and physical. High etch rates and tool automation are emphasized for most etchers used in manufacturing.

7.4.1 Barrel Reactor

A barrel reactor usually consists of a cylindrical vacuum chamber with a pair of rf electrodes concentrically located inside. The cross-sectional and side views are shown in Fig. 13. The wafers are placed upright in a boat along the axis of the chamber so that they lie concentrically within the reactor. The process gas pressure is usually between 0.1 to 5 torr, depending on the type of material to be etched. A large batch of wafers can be processed in each run. In the barrel reactor the active radicals are formed between the two electrodes and diffuse through the holes in the electrode to the wafers. The resultant etch profile is isotropic, because there is not much ion bombardment on the wafer surface. The barrel reactor is suitable for less critical process steps, such as the removal of photoresist. Reactants transport to the wafer surface through gas-phase diffusion from the edge of the wafer to the center of the wafer. Therefore, etch uniformities within a wafer and among wafers are very poor, and a large amount of overetch is usually needed to solve this etch uniformity problem.

Micron Ex. 1023, p. 373
Micron v. YMTC
IPR2024-00909

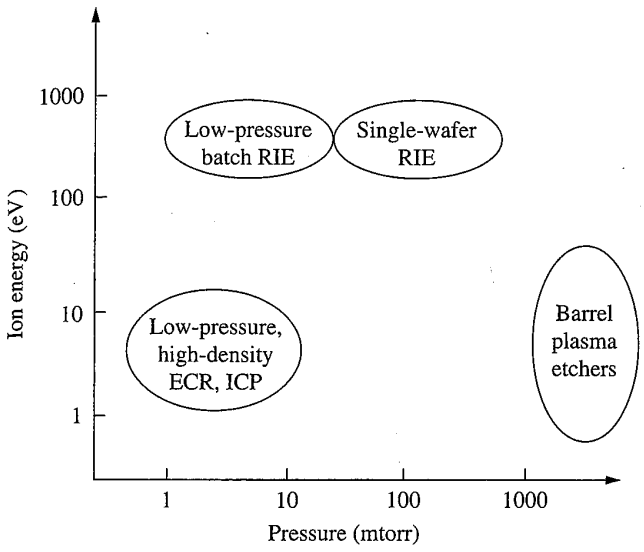


FIGURE 12
Comparison of ion energy and operating pressure ranges for different types of plasma reactors.

7.4.2 Downstream Plasma Etchers

To completely eliminate ion bombardment and photon radiation from the barrel reactor, wafers can be placed in a downstream process chamber. For example, in photoresist stripping, the elimination of ion bombardment is desirable to avoid ion-induced damage or wafer charging. Microwave or rf energy can be used to excite the plasma in the upstream chamber. In addition to its application for photoresist stripping, some single-wafer downstream etchers are used for isotropic plasma etching of silicon, silicon nitride, and silicon dioxide.

7.4.3 Reactive Ion Etchers (RIE)

In the late 1970s, scientists from AT&T Bell Laboratories invented the hexode RIE reactor, which found extensive use in the microelectronic industry. A batch of wafers is mounted on a central, hexagonal-shaped cathode powered by rf, and the

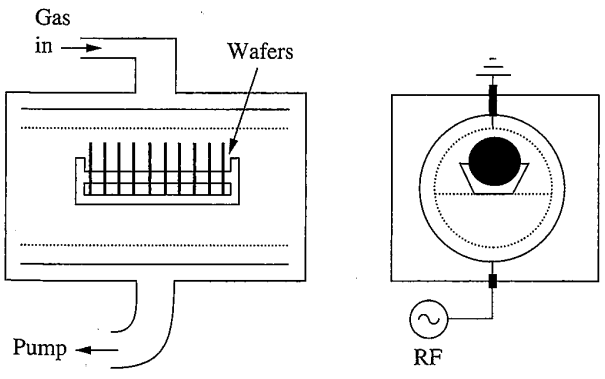


FIGURE 13
Schematic of a barrel plasma reactor showing (*a*) side view and (*b*) cross section.

Micron Ex. 1023, p. 374
Micron v. YMTC
IPR2024-00909

surrounding bell jar is grounded. As device dimensions shrink, etch uniformity across the wafers becomes a problem for batch RIE systems, and these reactors are not easily automated. Plasma processing is changing from a batch operation to a single-wafer operation, and the parallel-plate diode system has become popular. In this etcher, the wafer is placed on an rf capacitive-coupled bottom electrode, which is considerably smaller than the grounded part of the system. This and the relatively low working pressure (< 500 mtorr) mean that the wafers are subjected to a heavy bombardment of energetic ions from the plasma as a result of the large, negative self-bias at the wafer surface. Anisotropic etching is obtained because ion-enhanced chemical etching has a higher etching rate in the direction perpendicular to the wafer surface than in the direction parallel to the wafer surface.

The etch selectivity of this system is relatively low compared to barrel etch systems because of strong physical sputtering. However, selectivity can be improved by choosing the proper etch chemistry, for example, by polymerizing the silicon surface with fluorocarbon polymers to obtain selectivity of SiO_2 over silicon. Alternatively, a triode-configuration RIE etch, as shown in Fig. 14, can separate plasma generation from ion transport. Ion energy is controlled through a separate bias on the wafer electrode, thereby minimizing the loss of selectivity and the ion-bombardment-induced damage observed in most traditional RIE systems.

7.4.4 Magnetic-Enhanced and Confinement Plasma Etchers

Magnetic-enhanced reactive ion plasma etchers (MERIEs) use a group of permanent magnets located behind the etching wafer or pairs of direct-current electric coils to generate a magnetic field parallel to the wafer surface, as shown in Fig. 15. The magnetic field is perpendicular to the electric field (because of the cathode dc bias) and confines electrons to a circular trajectory near the cathode. Electron confinement reduces the loss of electrons to the wall of the system and increases the frequency of

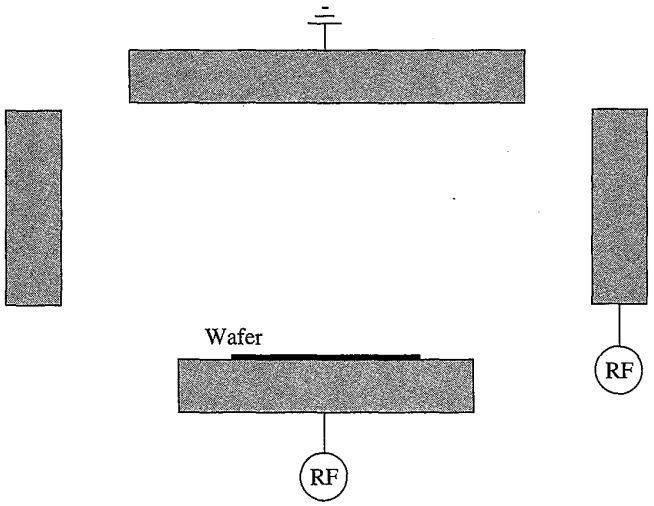


FIGURE 14
Schematic of a triode reactive ion etch reactor. The ion energy is separately controlled by a bias voltage on the bottom electrode.

Micron Ex. 1023, p. 375
Micron v. YMTC
IPR2024-00909

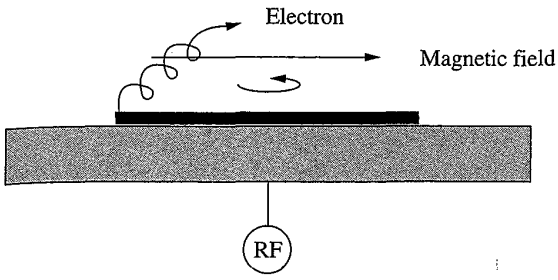


FIGURE 15
Magnetic-enhanced reactive ion etch (MERIE) reactor with a time-average magnetic field applied parallel to the cathode.

electron–neutral collisions. The associated higher frequency of collisions increases ion density, and electron confinement reduces the mobility of electrons toward the cathode, thereby reducing the self-bias voltage,[19] as shown in Fig. 16.[10] To achieve the required time-average plasma uniformity across the wafer surface, the field direction is rotated electrically with a period that is short with respect to the processing time. Magnetic fields also modify the ion-bombardment energies and can be used as another parameter for process optimization.

Uniformity of the plasma and radical fluxes to the wafer are critical issues. Magnetic confinement plasma etchers use permanent-magnet pole pieces and arrange them alternately around the process chamber to create a magnetic-field-free region around the wafer, as shown in Fig. 17. A high density of ions is achieved through the reflection of electrons back into the plasma from the surface magnetic-field bucket. This greatly increases the effective path length of electrons, and higher ion density is generated from the higher frequency of electron–neutral collisions.

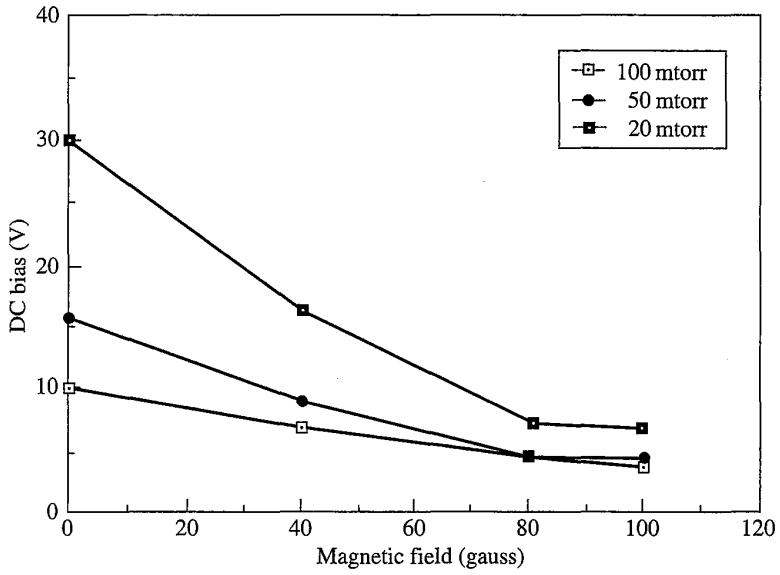


FIGURE 16
Effect of magnetic field on self-bias voltage. (*After Lii, Ng, and Danner, Ref. 10.*)

Micron Ex. 1023, p. 376
Micron v. YMTC
IPR2024-00909

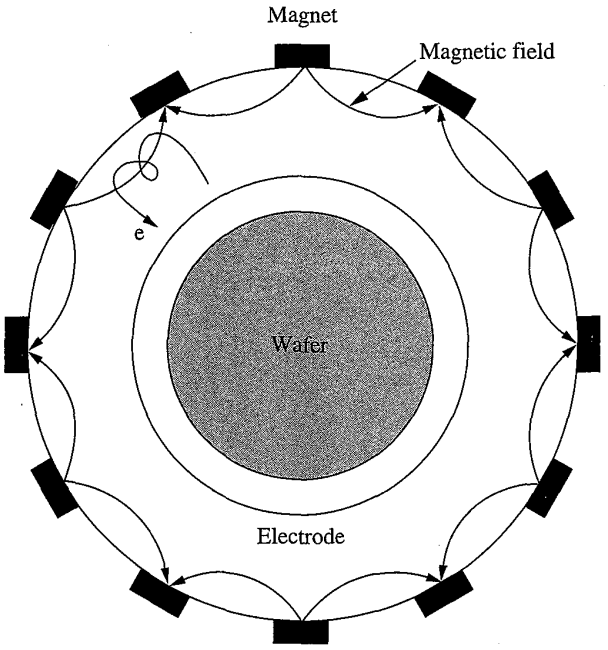


FIGURE 17
Schematic of magnetic-confinement reactive ion etch reactor with a multipolar magnetic bucket surrounding the reactor chamber.

7.4.5 Electron Cyclotron Resonance Plasma Etchers (ECR)

Most parallel-plate plasma etchers, except triode RIE, do not provide the ability to independently control plasma parameters, such as electron energy, ion energy, plasma density, and reactant density. As a result, ion-bombardment-induced damage becomes serious as device dimensions shrink.

ECR etching uses microwave excitation in the presence of a magnetic field to generate a high-density discharge. The Lorentz force causes the electrons to circulate around the magnetic field lines in circular orbits, with a characteristic cyclotron frequency of

$$\omega_{ce} = \frac{eB}{m_e} \tag{7.16}$$

where e is the electron charge, B is the magnetic field, and m_e is the electron mass. When this frequency equals the applied microwave frequency, a resonance coupling occurs between the electron energy and the applied electric field, which results in a high degree of dissociation and ionization (10^{-2} for ECR compared to 10^{-6} for RIE).[20] With a microwave frequency of 2.45 GHz, the required magnetic field is 875 gauss. Figure 18 shows one of the possible ECR plasma etching configurations.[21] Microwave power is coupled via a waveguide through a dielectric window into the ECR source region. The magnetic field, supplied from magnetic coils, decreases as a function of distance from the coils.

Because the gradient in the magnetic field decreases, the electrons are accelerated away from the plasma source, creating a negative potential in the direction of the wafer. Ions diffuse by ambipolar diffusion out of the source region into the the wafer process chamber. The wafer is rf- or dc-biased to control the energy of the ions to achieve the desired etch anisotropy. Plasma uniformity degrades

Micron Ex. 1023, p. 377
Micron v. YMTC
IPR2024-00909

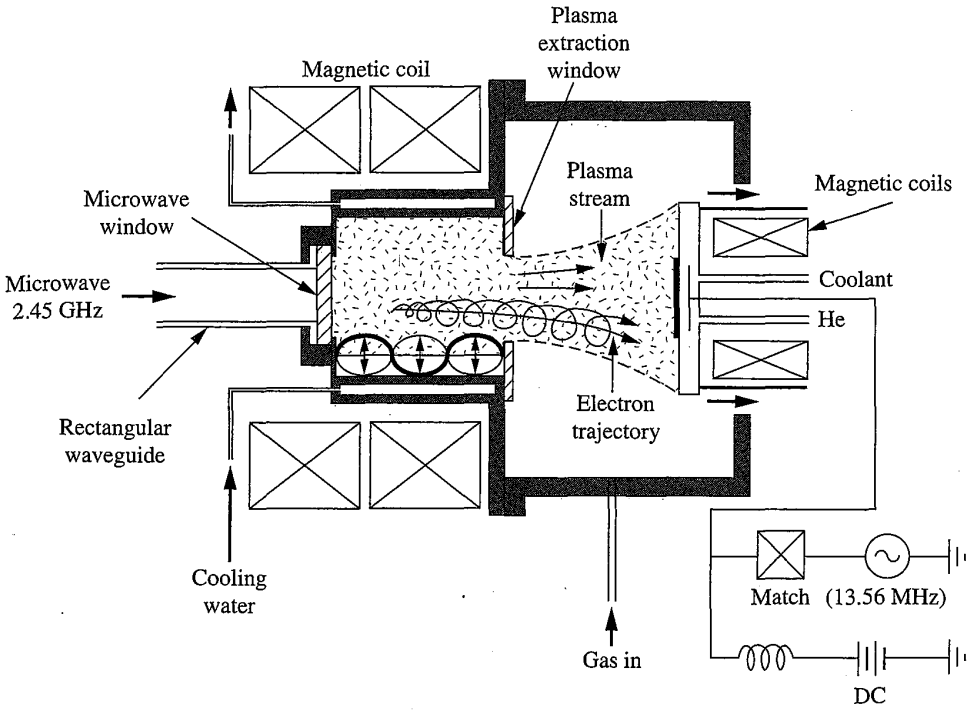


FIGURE 18
Schematic of electron cyclotron resonance etch reactor with a mirror magnetic field around the wafer to confine electron trajectory. (*After Chen et al., Ref. 21.*)

because of the ambipolar diffusion and mirror magnetic field,[21] but a multipole magnetic bucket[22] can be used to improve plasma uniformity. Etch uniformity can also be improved by putting the wafer in an ECR source region surrounded by an optimized magnetic field.[23]

7.4.6 Inductively Coupled and Helicon Wave RF Plasma Etchers

As feature sizes for ULSI continue to decrease, the limits of the conventional rf capacitive-coupled parallel system are being approached. Most ECR plasma reactors are suitable for etching ULSI. However, they are not popular in manufacturing because of their complexity. Other types of high-density plasma sources, such as inductively coupled plasma (ICP) sources or helicon plasma sources, may become the main plasma sources for future ULSI processing. A comparison of recent plasma sources is shown in Table 2.[24] The PMT MORI®source is a helicon wave plasma source. The Lam TCP source is a transformer-coupled plasma source.

An inductively coupled plasma source, shown in Fig. 19, generates high-density, low-pressure plasma that is decoupled from the wafer, and it allows independent control of ion flux and ion energy.[25] Plasma is generated by a flat spiral coil that is separated from the plasma by a dielectric plate on the top of the reactor. The wafer is located several skin depths away from the coil, so it is not affected by the electromagnetic field generated by the coil. There is little plasma density loss because plasma

Micron Ex. 1023, p. 378
Micron v. YMTC
IPR2024-00909

TABLE 2
Comparison of different commercially available plasma sources[24]

	Plasma source comparison						
	PMT MORI	**Lam TCP**	**Hitachi ECR**	**Drytech Hol. Anode**	**Prototech Helical Res.**	**Lucas Labs Helicon**	**Tegal HRe**
Source freq. (MHz)	13.56	13.56	2450	13.56	13.56	13.56	13.56
Sub. bias freq. (MHz)	13.56	13.56	2/13.56	13.56	13.56	13.56	< 500 kHz
Source *B* field (gauss)	50–120	—	800	—	—	50–400	†
Ion density (cm^3 at wafer)	3×10^{12}	$0.5–2 \times 10^{12}$	1×10^{11}	1×10^{12}	1.00×10^{12}	3×10^{11}	3×10^{11}
Ion potential (eV)	20+	20+	20+	20+	20+	20+	20+
Electron temp. (eV)	3.5	3.5–6	4	3–5	4	4	4–6
Pressure range (mtorr)	0.5–10	1–25	0.4–10	10–180	0.05–3000	1–10	1–20

†No magnetic field at wafer.

Micron Ex. 1023, p. 379
Micron v. YMTC
IPR2024-00909

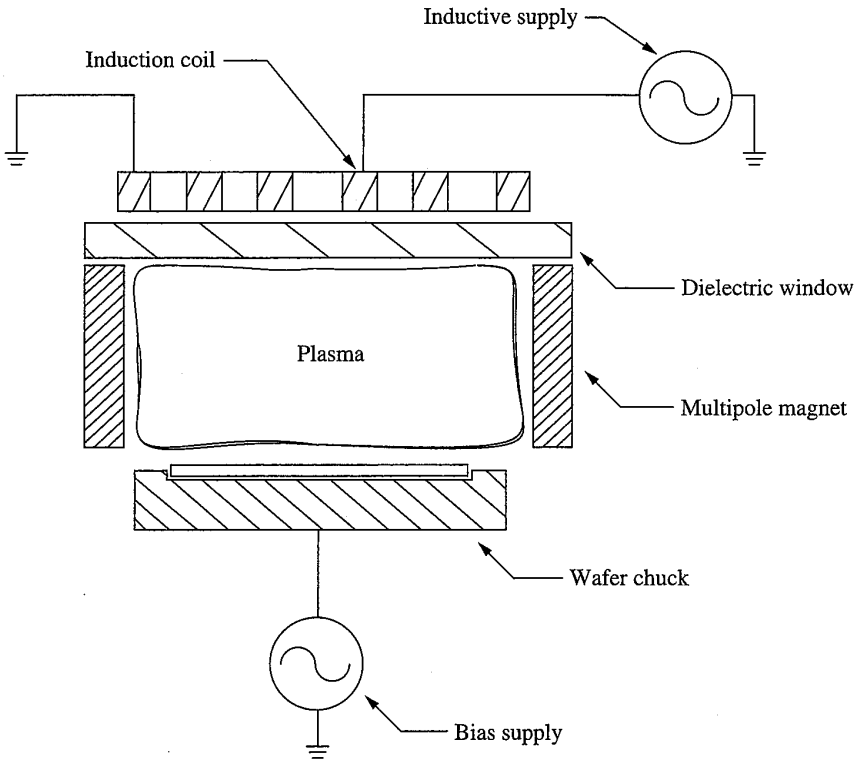


FIGURE 19
Illustration of an inductively coupled plasma reactor. (*After Keller, Forster, and Barnes, Ref. 25.*)

is generated only a few mean free paths away from the wafer surface. Therefore, a high-density plasma and high etch rates are achieved.

A helicon plasma source can also be used to generate a high-density ($> 10^{11}/cm^3$) discharge. A transverse electromagnetic radio-frequency wave (13.56 MHz), excited by a double-loop or single-loop antenna located outside a quartz source tube, is coupled with a steady longitudinal magnetic field B_0 (approximately 100 gauss) generated by a solenoid coil, as shown in Fig. 20.[26] The resonance condition (propagation of a helicon wave) depends on the magnitude of the longitudinal field and the dimension of the reactor. If the wavelength of the helicon wave is the same as the antenna length, the coupling will be resonant. High-density plasma then diffuses into the wafer chamber. In addition, the wafer can be biased separately with a second rf generator.

7.4.7 Clustered Plasma Processing

Microelectronic device wafers are processed in cleanrooms to minimize exposure to ambient particulate contamination. As device dimensions shrink, particulate contamination becomes a more serious problem. To minimize particulate contamination and preserve the integrity of thin film interfaces, clustered plasma tools use a wafer

Micron Ex. 1023, p. 380
Micron v. YMTC
IPR2024-00909

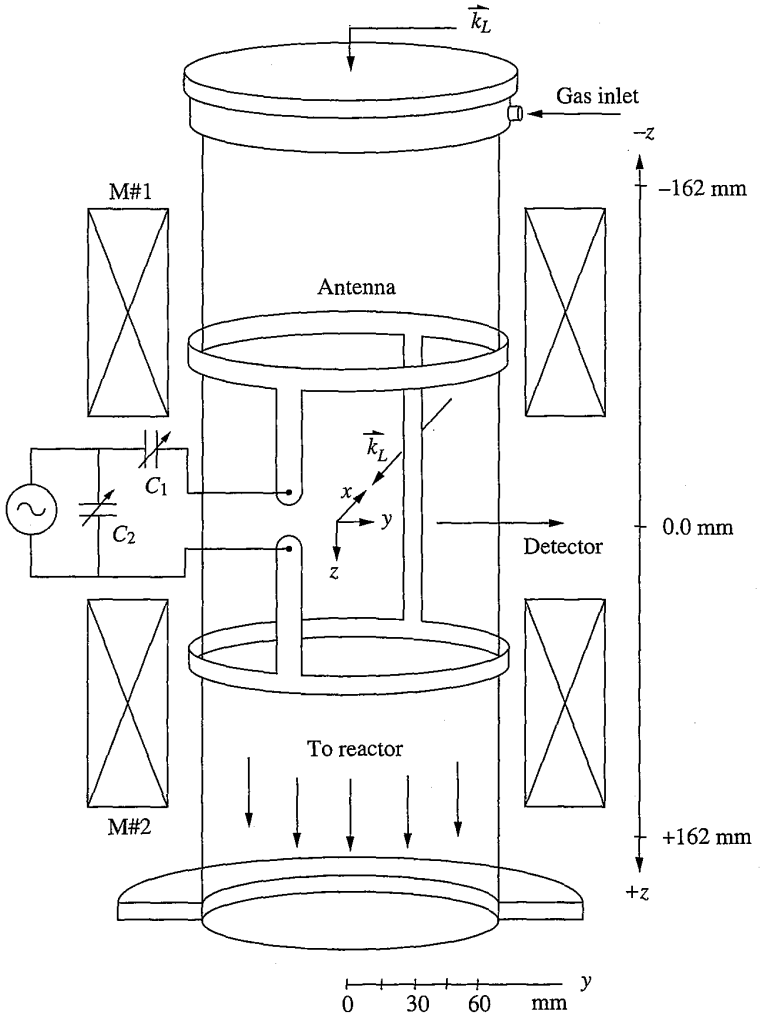


FIGURE 20
Schematic of a helicon wave plasma reactor with a double-loop antenna. (*After Giapis et al., Ref. 26.*)

handler to pass wafers from one process chamber to another in a vacuum environment, as shown in Fig. 21. The wafer can be transferred from one process chamber to another without being exposed to and contaminated by the ambient.[27]

Clustered plasma processing tools can also increase throughput. For example, integration of sequential process steps into multichamber in-situ processing modules, multilayer CMOS gate level etching, dielectric layer deposition with etchback planarization, and metal etching with resist strip are usually accomplished with clustered processing. However, tool reliability becomes very important as tools become even more sophisticated. If they are reliable, clustered tools can provide an economic advantage of high chip yield. The yield is key to reducing manufacturing costs, and clustered tools should provide better yield mainly because the wafer is exposed to less ambient contamination and is handled less.

Micron Ex. 1023, p. 381
Micron v. YMTC
IPR2024-00909

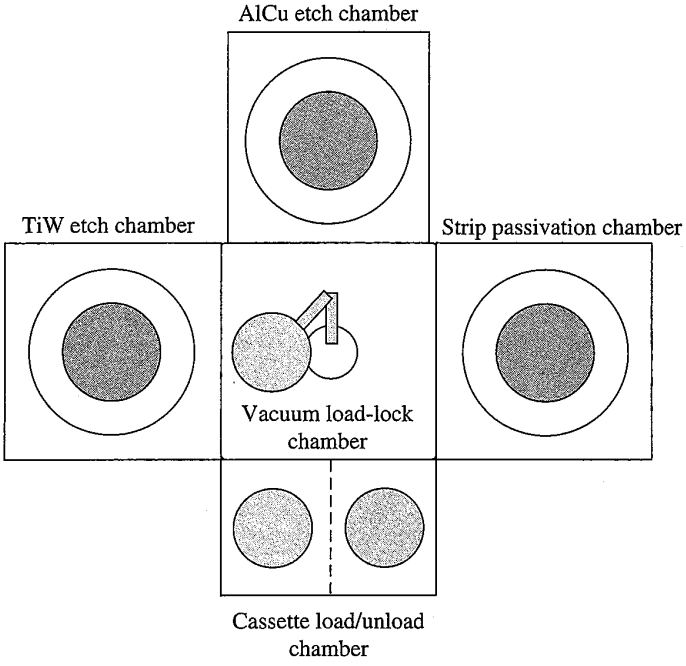


FIGURE 21
Cluster reactive ion etch tool for multilayer metal (TiW/AlCu/TiW) interconnect etching.

7.5 PLASMA PROCESSING PROCESSES

The chemical and physical complexity of plasma-surface interactions makes computer-based modeling and plasma simulation inadequate for developing plasma reactors. As a result, the detailed descriptions required to guide the transfer of processes from one reactor to another or to scale processes from a small to a large reactor are not available. Therefore, the plasma processes for fabricating microelectronic devices have been developed largely by time-consuming, costly, empirical methods.

Excellent critical dimension control and minimized profile microloading are the two most important issues for ongoing research. Extensive research in plasma etching explores ways to optimize the performance of etching processes, characterize and avoid contamination and damage effects, identify the important reactive species and measure their densities, and, in general, improve our limited understanding of plasma etching. See Refs. 28 and 29 on various aspects of plasma etching. Other studies concentrate on the limitations of conventional rf plasma. This research led to the development of enhanced plasma source configurations and processes that are still being examined.

Etch chemistries also play a critical role in the performance of etch processes. Table 3 lists some conventional and new etch chemistries for different etch processes.[30] A truly useful plasma process must have a robust process window.

Micron Ex. 1023, p. 382
Micron v. YMTC
IPR2024-00909

TABLE 3
Etch chemistries of different etch processes[30]

Material being etched	Conventional chemistry	New chemistry	Benefits
PolySi	Cl_2 or BCl_3/CCl_4 / CF_4 / $CHCl_3$ / CHF_3 } sidewall passivating gases	$SiCl_4/Cl_2$	No carbon contamination
		BCl_3/Cl_2	
		$HBr/Cl_2/O_2$	Increased selectivity to SiO_2 and resist
		HBr/O_2	No carbon contamination
		Br_2/SF_6	
		SF_6	Higher etch rate
		CF_4	
Al	Cl_2	$SiCl_4/Cl_2$	Better profile control
	BCl_3 + sidewall-passivating gases	BCl_3/Cl_2	No carbon contamination
	$SiCl_4$	HBr/Cl_2	
Al-Si (1%); Cu (0.5%)	Same as Al	$BCl_3/Cl_2 + N_2$	N_2 accelerates Cu etch rate
Al-Cu (2%)	$BCl_3/Cl_2/CHF_3$	$BCl_3/Cl_2 + N_2 + Al$	Additional aluminum helps etch copper
W	$SF_6/Cl_2/CCl_4$	SF_6 only	No carbon contamination
		NF_3/Cl_2	Etch stop over TiW and TiN
			No carbon contamination
TiW	$SF_6/Cl_2/O_2$	SF_6 only	
WSi_2, $TiSi_2$, $CoSi_2$	CCl_2F_2	CCl_2F_2/NF_3	Controlled etch profile
		CF_4/Cl_2	No carbon contamination
Single-crystal Si	Cl_2 or BCl_3 + sidewall-passivating gases	CF_3Br	Higher-selectivity trench etch
		HBr/NF_3	
SiO_2 (BPSG)	CCl_2F_2	CCl_2F_2	
	CF_4	CHF_3/CF_4	CFC alternatives
	C_2F_6	CHF_3/O_2	
	C_3F_8	CH_3CHF_2	
Si_3N_4	CCl_2F_2	CF_4/O_2	CFC alternatives
	CHF_3	CF_4/H_2	
		CHF_3	
		CH_3CHF_2	

Micron Ex. 1023, p. 383
Micron v. YMTC
IPR2024-00909

Developing an etch process usually means optimizing many characteristics, such as etch rate, selectivity, profile control, endpoint control, damage, etc., by adjusting a large number of process parameters. Because we do not fully understand etch processes, systematic experiments are required to study the large number of process parameters. The most common method is to use the design experiment and analyze the results with a surface response methodology. In this section, several of the most important etch processes for ULSI fabrication are introduced.

7.5.1 Silicon Trench Etching

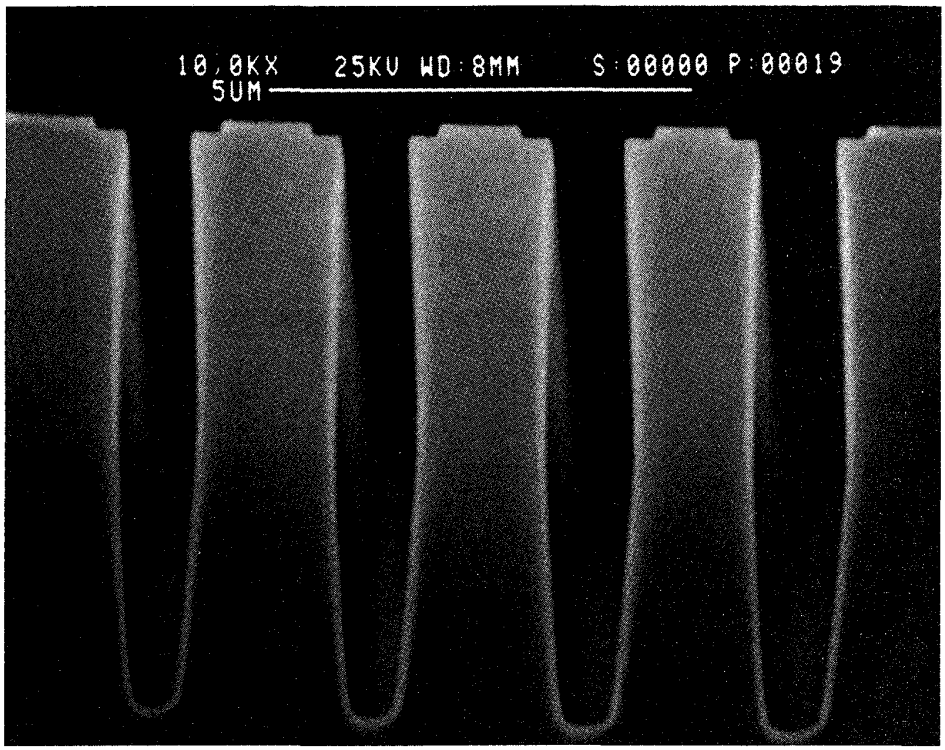
As feature sizes in ULSI decrease, a corresponding decrease is needed in the area occupied by both the isolation between circuit elements and the storage capacitor of the DRAM memory cell. This area can be reduced by etching trenches into the silicon substrate and filling them with suitable dielectric or conductive materials. Silicon trenches can also be used as alignment marks for optical or electron beam lithography. Deep silicon trenches, usually with a trench depth larger than 5 μm, are used mainly for forming trench capacitors. Shallow silicon trenches, usually with a trench depth less than 1 μm, are often used for isolation between device elements.

To guarantee a void-free trench refill, anisotropic etching with a high etch selectivity to mask material is desired for submicrometer-deep silicon trench etching. This precludes the use of fluorocarbon chemistry, which tends to undercut the mask material.[18] Fluorine chemistry has a high silicon etch rate, but it has low etch selectivity to the oxide mask and to the isotropic trench profile. However, it is often used for etching shallow silicon trenches with trench depths less than 1 μm, because it satisfies the requirement of etch selectivity of silicon to the photoresist mask. In this section, the deep-trench etching process is discussed because of its stringent requirements of etch rate, anisotropy, and selectivity.

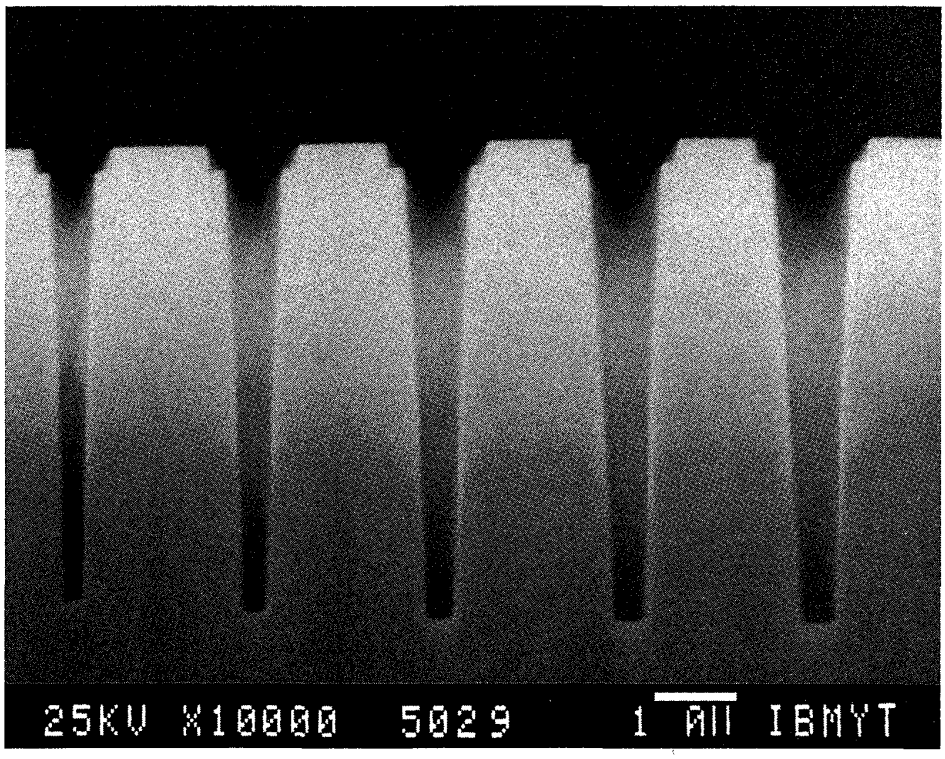
Chlorine-based and bromine-based chemistries have a high silicon etch rate and high etch selectivity to the silicon dioxide mask. However, chlorine-based chemistries tend to give slightly isotropic trench profiles,[31,32] especially in the ion-implanted phosphorus-doped region. Carbon-containing gas has been employed to provide sidewall passivation to avoid undercutting from fast chlorine-silicon reactions.[33] The sidewall passivation layer, however, causes contamination and reduces the etch selectivity of silicon to a silicon dioxide mask. A simple carbon-free etch chemistry leading to minimum thickness of the sidewall passivation is desired to guarantee high accuracy in the pattern transfer. As a general rule, the reactivity of the halogens follows the order F > Cl > Br. Recently, bromine-based chemistries have become popular for etching silicon with anisotropy and extremely high selectivity to the oxide mask.[34]

Sidewall passivation protects the sidewall from lateral etch attack and local ion-enhanced trenching. The desired trench shapes with smooth vertical sidewalls are obtained by controlling the balance between the etch and the inhibition by sidewall passivation. The sidewall shape depends significantly on the wafer temperature. Increasing temperature results in less deposition of sidewall passivation and more lateral etching. Fig. 22[10] refers to the effects of backside cooling under different He pressure. Greater pressure of He results in a cooler wafer temperature.

Micron Ex. 1023, p. 384
Micron v. YMTC
IPR2024-00909



(a)



(b)

FIGURE 22
Effect of wafer backside cooling on a silicon trench profile: (*a*) backside He pressure 0.5 torr (*b*) backside He pressure 6.0 torr. (*After Lii, Ng, and Danner, Ref. 10.*)

Micron Ex. 1023, p. 385
Micron v. YMTC
IPR2024-00909

Aspect ratio–dependent etching (ARDE) is often observed in submicrometer-deep silicon trench etching caused by limited ion and neutral transport within the trench. Trenches with large aspect ratios etch more slowly than trenches with small aspect ratios. Ion scattering results from ion–neutral collisions in the plasma sheath, and the electrical charging on the masks causes ARDE. Some neutrals are transported to the bottom of the trench by Knudsen diffusion, also contributing to ARDE. Low gas pressure reduces the ARDE effect, and chlorine-based chemistry shows less ARDE than fluorine-based chemistry during deep trench etching, because ion-assisted etching is dominant in chlorine-based chemistry.[35]

7.5.2 Polysilicon and Polycide Gate Etching

Low-pressure chemical vapor deposited (LPCVD) polysilicon is usually used as a gate material for MOS devices because of its superior interface property with thin gate oxide at high temperature. Metal silicide over polysilicon has also been used for the MOS gate because of its high electrical conductivity. Anisotropic etching and high etch selectivity between polysilicon and the gate oxide are the most important requirements for gate etching. Achieving high selectivity and etch anisotropy at the same time is difficult for most ion-enhanced etching processes. Therefore, multistep processing is used in which different etch steps in the process are optimized for etch anisotropy and selectivity.

Most chlorine-based and bromine-based chemistries can be used for gate etching to achieve the required etch anisotropy and selectivity. The main etchant today is bromine-based chemistry, because bromine gives higher etch selectivity of polysilicon to gate oxide than chlorine does. Figure 23 shows a scanning electron microscope (SEM) cross section of a polysilicon gate etched by HBr plasma[36] that has a 0.1 μm gate width and a 45 Å gate oxide thickness. The thin native oxide on top of the polysilicon should be removed by fluorine-based plasma before polysilicon etching, i.e., initiation, to avoid any micromasking formation during polysilicon etching. To further improve the selectivity of polysilicon to gate oxide requires removing carbon-containing gases and carbon-containing materials from the etch process and the etch reactor, because carbon tends to reduce the selectivity of polysilicon to gate oxide,[37] as shown in Fig. 24. It is suspected that carbon reacts with the oxygen atoms of silicon dioxide. An enhanced etch attack at the edge of the gate, a phenomenon called "trenching," is shown in Fig. 25. Trenching results in broken gate oxide and is usually observed at low gas pressure. The trenching problem can be solved by increasing etch selectivity, such as adding bromine- and oxygen-containing gases to the etch process.

7.5.3 Oxide and Nitride Etching

An oxide etching process is often used to open a contact window down to a silicon device or metal conductor; therefore, the selectivity of SiO_2 to Si must be high. It is well known that selectivity is achieved by increasing the ratio of CF_x to F atoms

Micron Ex. 1023, p. 386
Micron v. YMTC
IPR2024-00909

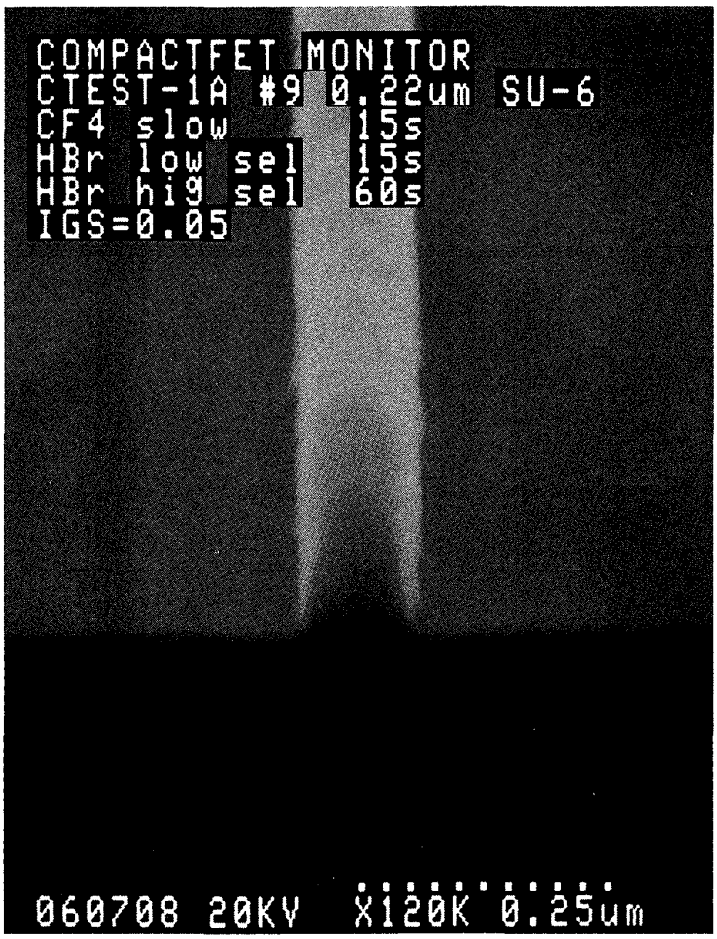


FIGURE 23
SEM cross section of HBr-plasma-etched polysilicon gate with a 0.1 μm gate width and a 45 Å gate oxide thickness. (*After Lii et al., Ref. 36.*)

in the plasma.[7] CHF_3/CF_4 chemistry is a common oxide etchant because it provides high selectivity to Si from carbon polymer deposition on the silicon surface. Therefore, a further silicon surface cleaning must be added to remove the polymer passivation layer after oxide etching to avoid high contact resistance between metal and silicon.

LPCVD or plasma-enhanced CVD silicon nitride are often used as a sidewall around the polysilicon gate in MOS devices for source and drain extension or used as a final passivation layer. Sidewall formation for a polysilicon gate requires anisotropic etching and high selectivity of silicon nitride to silicon dioxide. Si_3N_4 can be etched by both fluorine atoms and CF_x-containing plasma. CF_4 and CF_4/O_2 are widely used to etch silicon nitride but are not selective to silicon or silicon dioxide.[38] Si_3N_4 can be etched anisotropically in an RIE mode with high selectivity to Si and SiO_2 by CHF_3/O_2, CH_2F_2, or CH_3F.[39,40] Details of why the etch mechanism results in high selectivity are not clear. Fluorine-deficient fluorocarbon polymer is thought to contribute to this high selectivity.

Micron Ex. 1023, p. 387
Micron v. YMTC
IPR2024-00909

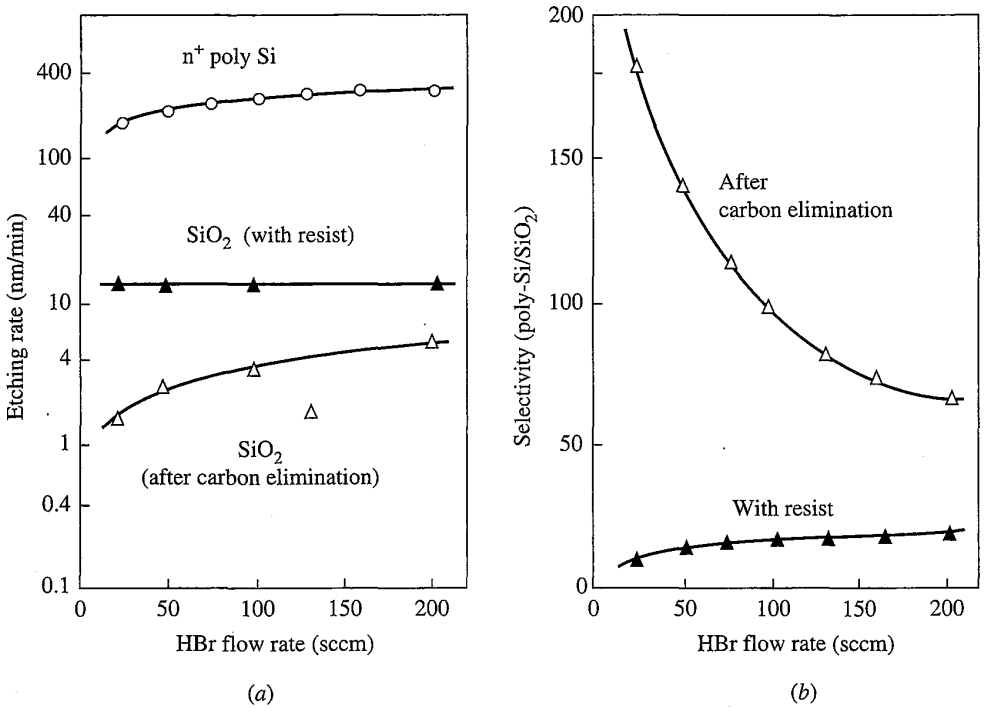


FIGURE 24
Dependence of (*a*) the polysilicon etch rate and (*b*) Si/SiO_2 selectivity on HBr flow rate, carbon reducing Si/SiO_2 selectivity. (*After Nakamura, Iizuka, and Yano, Ref. 37.*)

7.5.4 Aluminum and Tungsten Metal Etching

Etching of a metallization layer is a very important step in semiconductor device fabrication. Aluminum and tungsten are the most popular materials used for interconnection, and anisotropic etching of these materials is usually required. An aluminum etch with fluorine-based chemistry does not work, because of the very low vapor pressure of the etch product AlF_3. Chlorine-based chemistry has been used for Al etching most frequently, and bromine-based chemistry has been investigated recently.[41]

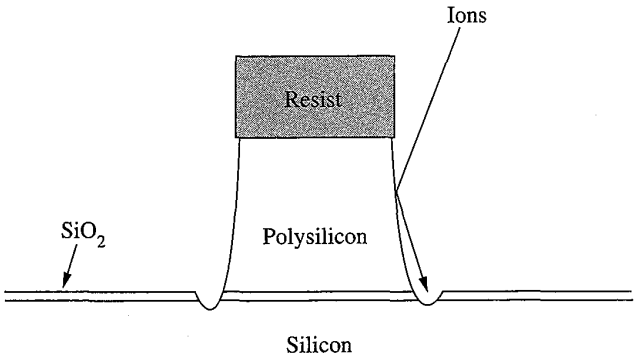


FIGURE 25
Schematic of the trenching mechanism during polysilicon gate etching.

Micron Ex. 1023, p. 388
Micron v. YMTC
IPR2024-00909

Cl_2 spontaneously etches clean aluminum. Chlorine has a very high chemical etch rate with aluminum and tends to produce an undercut during aluminum etching. Carbon-containing gas is added to form sidewall passivation during aluminum etching to obtain etch directionality. The Al_2O_3 layer, which is always present on air-exposed aluminum, is etched very slowly in chlorine-based gases. Energetic ion bombardment, for example, sputtering, is required to remove Al_2O_3 before aluminum etching. In the semiconductor industry, BCl_3 plasma is often used to remove Al_2O_3 on the aluminum surface.[42] BCl_x also appears to recombine with chlorine at the wafer surface to ensure an anisotropic etch profile during Al etching with a Cl_2/BCl_3 mixture. A small percentage of silicon and copper are usually added to enhance the electromigration resistance of aluminum conductors, but Cu causes micromasking during aluminum etching with chlorine-based plasma because it generates low-volatility etch products. Micromasking can be suppressed by using stronger ion bombardment during etching and addition of aluminum chloride gas.[43] More recently, a barrier metal layer, such as TiW and TiN, has been used to prevent Si migration into Al. TiW and TiN etch readily in chlorine-based plasma.

Corrosion of the aluminum pattern exposure to the ambient is another problem in aluminum etching. Residual chlorine on both the Al sidewall and the photoresist tends to react with water to form HCl, which etches aluminum. In-situ exposure of the wafer to a CF_4 discharge to exchange Cl with F and then to an oxygen discharge to remove the resist, followed by immediate immersion in deionized (DI) water, is successful in eliminating Al corrosion.[44] High-temperature heating of the wafer to evaporate chlorine or hydrolyzing chlorine with hydrogen-containing gases can be used to remove chlorine trapped in the resist and the sidewall.

LPCVD tungsten has been widely used for contact hole filling and first-level metalization because of its excellent deposition conformability. Both fluorine- and chlorine-based chemistry etch W and form volatile etch products.[45,46] Fluorine-based chemistry also etches the silicon dioxide layer; thus it is very difficult to obtain selectivity between W and SiO_2. Chlorine-based chemistries etch tungsten with selectivity to SiO_2. Two tungsten etch processes, blanket W etchback to form a W plug and W pattern etch, are discussed briefly in this section.

To form a W plug, blanket LPCVD W is deposited on top of a TiN barrier layer as shown in Fig. 26. Reducing loading effects during the etchback of W is very important to avoid a W recess in the plug. A two-step process is used for W etchback. In the first step, 90% of the W is etched at a high etch rate and excellent uniformity to achieve high throughput. In the second step, the etch rate is reduced to remove the

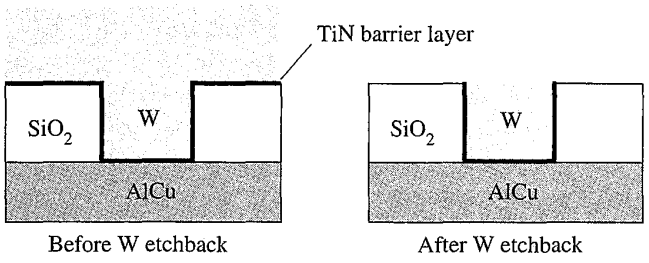


FIGURE 26
Formation of tungsten plug in a contact hole by depositing blanket LPCVD W and then using RIE etchback.

Micron Ex. 1023, p. 389
Micron v. YMTC
IPR2024-00909

remaining W with an etchant that has a high W-to-TiN-barrier-layer etch selectivity and a minimum loading effect.[47] The loading effect is reduced by decreasing the gas pressure and wafer temperature because of the low chemical etch rate of F with W at these process conditions.

For W interconnect etching, the selectivity between W and the photoresist mask is very low for both fluorine- and chlorine-based chemistries. N_2 is added to the SF_6/Cl_2 mixture to get a slight selectivity increase.[48] Inorganic material, such as SiO_2, can be used as a hard mask during W etching with chlorine-based chemistry because of high etch selectivity of W to the hard mask.

7.5.5 Organic Materials Etching

Oxygen plasmas have been used with a barrel reactor for resist etching since the 1970s. Oxygen atoms are the primary species responsible for etching polymers at a pressure of 1–3 torr. Oxygen atoms initiate polymer etching by extracting hydrogen atoms from a polymer chain to form radicals. These radicals then react with oxygen molecules to form volatile products under high wafer-temperature conditions or under ion bombardment.

Recently, multilayer lithography has become increasingly important for ULSI processing because it provides high resolution that is independent of the underlying materials and substrate topography. The trilayer scheme is a very popular multilayer lithography method. It uses a conventional resist to pattern an intermediate layer, such as silicon dioxide or nitride, that acts as a mask during the subsequent planarization of the organic layer by oxygen RIE. A high degree of anisotropic and selective etching is required in these applications. This implies that both chemical and physical processes are needed, because purely physical etching is not selective, whereas purely chemical etching is not anisotropic. Reactive ion etching contributes both chemical and physical processes. To achieve anisotropic etching, sidewall passivation is required to suppress chemical and physical etching on the sidewalls of the feature.[49] A low wafer temperature and extra carbon- or chlorine-containing gases increase the deposition rate of sidewall films and improve etch anisotropy.[50]

7.5.6 Planarization Etch

With the advent of three or more levels of interconnection, more than 50% of the cost of chip manufacturing will be in the fabrication of electrical interconnects. Between these interconnection levels are inorganic insulators or polymer films, such as doped silicon dioxide glass or polyimide. Because the lithographic exposure achieves maximum resolution over only a limited depth of focus, these films must be flat, or planarized, over the optical exposure area to construct higher levels of interconnection. If the temperature budget does not allow planarization of the oxide layer by means of heat flow, planarization is sometimes accomplished by a repetitive sequence of plasma etching and deposition. An organic material with good flow properties is

Micron Ex. 1023, p. 390
Micron v. YMTC
IPR2024-00909

deposited on top of the oxide layer, and the wafer is etched in a plasma that removes the organic material and the oxide at an approximately equal rate, leaving a uniform oxide layer.

Mixtures of fluorine-based chemistry, such as NF_3, CF_4, and CHF_3, with oxygen are suitable for a nonselective silicon dioxide/organic material etch. Good etch uniformity and a high etch rate are the most important requirements for a planarization etch. The planarization etching step also reduces surface topography because different etch rates are used at different surface angles.

7.6 DIAGNOSTICS, END POINT CONTROL, AND DAMAGE

Dry etching differs from traditional wet chemical etching, and it usually does not have enough etch selectivity to the layer underneath. The plasma reactor must be equipped with a monitor that indicates when the etching process is to be stopped; an end point detection system is needed to avoid etching the underlying layer. Several different types of end-point systems are available. Most of them detect the change in the amount of active reactants or etch products in the discharge. Two of the most popular methods are introduced here. Some etching systems use a combination of these methods for end point detection.

7.6.1 Optical Emission Spectroscopy

Optical emission spectroscopy is a method of analyzing the light emitted by excited atoms and molecules in the gas discharge. The concentration of excited residual products or excited reactants during the etching process is studied by equipping a detector with a filter that lets light of a specific wavelength pass through to be detected, as shown in Fig. 27. Photons passing through a tunable monochromator are detected by a photomultiplier. The presence of neutrals and ions can be determined by correlating these spectra with previously determined spectral series. However, absolute concentrations cannot be obtained unless a complex actinometry technique is used. The emission signal derived from the primary etchant or by-product begins to rise or fall at the end of the etch cycle.

7.6.2 Laser Interferometry

Laser interferometry uses laser light directed toward the wafer surface. If the films on the wafer are transparent, the laser beams reflected from the top and bottom of the etched layer will interfere with each other. As the thickness of the etched layer is reduced, the degree of interference will change, either constructively or destructively. The time between light maxima and minima can be used to determine the etching rate. At the end of the etching process, the interference stops and the signal flattens out. The period of the oscillation is related to the change in film thickness, Δd where

Micron Ex. 1023, p. 391
Micron v. YMTC
IPR2024-00909

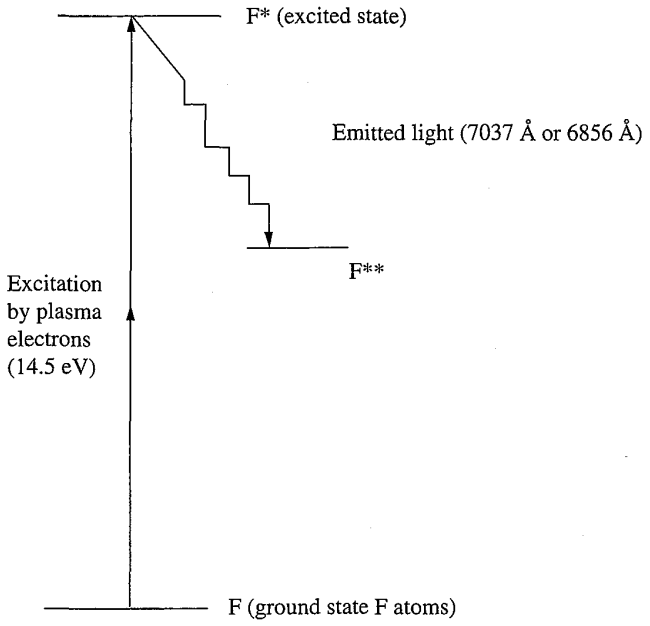


FIGURE 27
Mechanism of optical emission spectroscopy used for end-point detection during plasma etching.

$$\Delta d = \lambda/2n \tag{7.17}$$

λ is the wavelength of the laser light, and n is the refractive index of the etching layer. For example, Δd for polysilicon is 800 Å measured by using a helium-neon laser source for which $\lambda = 6328$ Å. Laser interferometry not only detects the end point but also determines the in-situ etch rate. The disadvantages of this method are that information is obtained from only a limited part of the wafer surface and the output signal is very weak if the etched surface is rough.

7.6.3 Plasma-Induced Damage

As stated above, plasma consists of energized ions, electrons, and excited molecules. When the excited particles recombine, they give off photons having an energy of a few eV. In addition, ion and electron bombardment may all contribute to damage mechanisms. Plasma-induced damage can take many forms, such as a trapped interface charge, material defects migration into bulk materials, and contamination caused by the deposition of etch products on material surfaces.

Etch damage can change the characteristics of sensitive components, such as Schottky diodes. For example, their rectifying capability can be reduced considerably. During oxide contact hole etching, heavy polymer deposition often causes high contact resistance. Sometimes the low-volatility etch products formed during plasma etching, such as silicon bromide from polysilicon etching with HBr, or $CuCl_x$ from AlCu etching with chlorine, are not completely removed after etching and cause device reliability problems. For example, corrosion can occur on dry-etched aluminum lines. Therefore, it is important that the impurities are removed and the damaged area restored before processing is continued. To reduce possible damage from contamination, wet etching is usually used to rinse wafers after plasma etching.

Micron Ex. 1023, p. 392
Micron v. YMTC
IPR2024-00909

Charge buildup can occur in reactive ion etching because of severe ion bombardment of the wafer surface, and nonuniform plasma stimulates the accumulation of charges on the gate electrode. These stored charges can cause an electric breakdown of the thin gate oxide.[51] In addition, nonuniform electron and ion currents that do not balance locally cause charge buildup and oxide damage.[52] Therefore, it is increasingly important to find new processes and equipment that do not contaminate or cause damage in the underlying layers.

7.6.4 Plasma-Process-Induced Particle Contamination

Plasma processes can generate a large number of particles on the wafer surface during etching. These particles are created by chemical or mechanical means during plasma exposure. Laser light scattering shows that particles are trapped near the plasma/sheath interface during plasma etching.[53] Negatively charged particles are suspended at the plasma/sheath interface because of the electrical potential difference between the sheath boundary and the electrode/wafer surface. These particles range in size from less than a quarter of a micrometer to tens of micrometers at a density of over 10^7 cm^{-3}, and they can fall on the wafer surface when the plasma is turned off. Fluorine-containing gas mixtures produce fewer particles during silicon etching than chlorine- or bromine-containing gases do because fluorine generates etch products that have a high vapor pressure. Also, particle traps are formed by a change of electrode material or topography, which can increase the localized plasma potential. The tool design, especially the electrode, significantly influences the cleanliness of a plasma tool.

Particle contamination control includes the proper operation of plasma shutdown, appropriate process chemistry, and good tool design. Particles suspended near the plasma/sheath boundary can be removed before shutting the plasma process off by using the proper changes in rf power, gas flow, and magnetic field near the end of processing. Particles constitute the etch species in the plasma, so choosing the correct plasma chemistry reduces particle generation, and proper tool design can purge particles to the pump during processing.

7.7 WET CHEMICAL ETCHING

Wet etching is still the practical alternative for a high-throughput, flexible production process. With properly selected chemicals, etch reactions with the primary film are thermodynamically favored over reactions with the other films.[54] The etch-rate ratio usually approaches infinity. Wet etches are especially suitable for blanket etches of polysilicon, oxide, nitride, and metal. Recently, there has been a resurgence of wet etching for certain process steps because plasma etching fails to provide the required etch selectivity, damage-free interface, and particle-contamination-free wafers. Using robotic handling and ultrapure chemicals has improved particle control and process consistency. These improvements have revived wet etching processes for ULSI processing.

Micron Ex. 1023, p. 393
Micron v. YMTC
IPR2024-00909

7.7.1 Wet Etching Mechanisms and Techniques

Etching of a solid in a solution is a heterogeneous process. Consecutive steps are thought to take place in the wet etch process. These steps are the diffusion of the reacting ions or molecules from solution, the adsorption of the reacting ions or molecules on the surface, the formation of a surface complex, the dissociation of this complex into reaction products, the desorption of the final reaction products, and the diffusion of the etch products into solution. The key criterion is that the etch products must be soluble in the etchant solution.

Most wet etch reactions involve oxidation-reduction. A chemical reaction is considered to be the transition of a system with an initial energy E_r to a final energy E_p. Heat of reaction ΔH is equal to $E_r - E_p$. A heat of reaction less than zero indicates that the reaction is favorable thermodynamically.

Two basic wet etching techniques are immersion etching and spray etching. The choice of one or the other depends on the material to be etched, the etchants, and the economic considerations. Immersion etching is the simplest technique. The masked or unmasked wafer is submerged in the etch solution, and mechanical agitation is usually required to ensure etch uniformity and a consistent etch rate. Spray etching requires less volume of chemicals and is faster than immersion etching. Good process control and etch uniformity are easily obtained from spray etching because fresh etchant is constantly supplied to the wafer surface while the etch products are continuously removed. Fully enclosed systems that eliminate the particle problems incurred when wafers are drawn out of a liquid/air interface are important for future ULSI processing.

7.7.2 Polysilicon and Silicon Nitride Etching

Most etchants may be classified as redox etchants. For silicon etching, H_2O_2 and HNO_3 are used to oxidize the silicon, and HF and HCl are used to dissolve the oxides. Selective isotropic chemical etching of polysilicon is the most widely used etching process for patterning deposited polysilicon films on the oxide. Concentrated KOH at 80 to 100°C is often used for selective polysilicon to oxide etching because of its low SiO_2 etch rate. Dopant concentrations and temperature may affect the etch rate of polysilicon.

Silicon nitride films are etchable at room temperature in concentrated HF or buffered HF and in H_3PO_4 at 150 to 200°C. The etch rate is strongly affected by the oxygen concentration in the nitride films. Selective etching of nitride to oxide is done with 85% H_3PO_4 at 180°C. The etch rate is typically 100 Å/min for CVD nitride, but only a few angstroms per minute for thermal oxide film. Plasma-enhanced CVD nitride has a much higher etch rate than CVD nitride. Sputter-deposited nitride films also have a different etch rate than CVD nitride films.

7.7.3 Silicon Dioxide Etching

SiO_2 can be etched by aqueous fluoride solutions, usually HF with or without the addition of NH_4F. The exact chemical mechanisms are extremely complex. The

Micron Ex. 1023, p. 394
Micron v. YMTC
IPR2024-00909

addition of NH_4F to HF controls the pH value and prevents depletion of the fluoride ions, thus maintaining stable etching performance. Silicon dioxide can also be etched in vapor phase HF. Vapor phase HF oxide etch technology looks promising for submicrometer feature etching because process can be tightly controlled and a desired hydrogen surface passivation can be precisely selected. The main reactions between oxide and HF vapor are

$$SiO_2 + 6HF \rightarrow H_2SiF_6 + H_2O \tag{18}$$

$$H_2SiF_6 + H_2O \rightarrow SiF_4 + 2HF \tag{19}$$

The etch rate of wet oxide etching depends on etchant concentration, agitation, and temperature. In addition, density, porosity, microstructure, and the purity of the silicon dioxide also affect the etch rate. CVD nitride etches are much slower than SiO_2 in HF solution, so an oxide-to-nitride selective wet etch is possible with HF solution.

7.8 SUMMARY AND FUTURE TRENDS

The etch selectivity and etch uniformity within the wafer become very important as film thickness is reduced, device dimensions shrink, and wafer size increases. Half-micrometer processes stretch conventional dry etching approaches to the limit. The first problem observed at quarter-micrometer generation is that conventional RIE will not achieve the clean, vertical sidewall etches demanded for that generation. The second problem is achieving uniformity across a wafer, because etch rates will vary with wafer surface topography. A related problem is known as microloading, in which varying degrees of etch behavior are triggered by alternating dense and sparse etching patterns. The final problem is plasma-induced damage. The concern is that electrostatic charging during the plasma etching process will damage the gate insulators. Process engineers are also concerned about driving unwanted ions and other contaminants into the silicon during etching.

High-density plasma sources with a lower electric potential difference between the plasma and the wafer will, therefore, become necessary to minimize the effects of ion-bombardment-induced damage. The solution lies largely in the high-density plasma sources in which plasma generation is wholly or partly separated from the processing region. The independent control of ion energy and plasma density promises a wider window for process optimization. Such approaches generate considerably denser plasma and avoid creating damaging, high-energy particles. The newer systems offer the necessary selectivity to the resist materials that protect the unetched portions of the wafer as well as to the underlying films that should not be etched. They also help prevent unwanted wafer depositions by using simple gas chemistry. In addition, the combination of a high-density plasma source and a short-gas-residence-time reactor design, i.e., a reactor equipped with high pumping capacity, is important to obtain a high etch rate.[55] Though such solutions appear viable, they remain to be proved in

Micron Ex. 1023, p. 395
Micron v. YMTC
IPR2024-00909

production. Etching will likely pose a major challenge at quarter-micrometer generation.

It is well known that many plasma processes drift from desired conditions and, therefore, require continual correction. Typically, the process drifts to a marginal performance before corrective measures are taken, leading to lost throughput and yield. Future etching systems will need a much higher degree of control and a self-diagnosis capability, so that high process stability can be maintained. Active feedback controls are needed to improve process reliability and reduce process maintenance. Currently, controls are used in plasma etching only for end point detection, with the process completed at a specified time after the end point. With automatic feedback control based on real-time detection of the plasma and wafer state, the optimal etch results could be maintained. Computer modeling of various aspects of plasma etching processes can increase our understanding of etch mechanisms and may enhance system design and process development. Clustered processing will become more popular and will be used to maintain the integrity of sensitive interfaces between thin films.

REFERENCES

1. R. A. Morgan, *Plasma Etching in Semiconductor Fabrication,* Elsevier, New York, 1985.
2. B. Chapman, *Glow Discharge Processes*, Wiley, New York, 1981.
3. J. L. Vossen, "Glow Discharge Phenomena in Plasma Etching and Deposition," *J. Electrochem. Soc.,* **126,** 319 (1979).
4. D. L. Flamm and V. M. Donnelly, "The Design of Plasma Etchants," *Plasma Chem. Plasma Process.,* **1,** 317 (1981).
5. G. S. Oehrlein and J. F. Rembetski, "Plasma-Based Dry Etching Techniques in the Silicon Integrated Circuit Technology," *IBM J. Res. Devel.,* **36,** 140 (1992).
6. C. J. Mogab, A. C. Adams, and D. L. Flamm, "Plasma Etching of Si and SiO_2: the Effect of Oxygen Additions to CF_4 Plasmas," *J. Appl. Phys.,* **49,** 3796 (1978).
7. L. M. Ephrath, "Selective Etching of Silicon Dioxide Using Reactive Ion Etching with CF_4–H_2," *J. Electrochem. Soc.,* **126,** 1419 (1979).
8. H. F. Winters and I. C. Plumb, "Etching Reactions for Silicon with F Atoms: Product Distributions and Ion Enhancement Mechanisms," *J. Vac. Sci. Technol.,* **B9,** 197 (1991).
9. J. W. Coburn and H. F. Winters, "Ion and Electron Assisted Gas-Surface Chemistry—An Important Effect in Plasma Etching," *J. Appl. Phys.,* **50,** 3189 (1979).
10. Y. T. Lii, H. Ng, and D. A. Danner, "Magnetic Enhanced RIE of Silicon Deep Trench," *Proceedings of the 8th Symposium on Plasma Processing,* The Electrochemical Society, Pennington, NJ, 1990, p. 462.
11. Y. T. Lii and J. Jorne, "Redeposition during Deep Trench Etching," *J. Electrochem. Soc.,* **137,** 2837 (1990).
12. H. Hubner, "Calculations of Deposition and Redeposition in Plasma Etch Processes," *J. Electrochem. Soc.,* **139,** 3302 (1992).
13. R. A. Gottscho and C. W. Jurgensen, "Microscopic Uniformity in Plasma Etching," *J. Vac. Sci. Technol.* **B10,** 2133 (1992).

Micron Ex. 1023, p. 396
Micron v. YMTC
IPR2024-00909

14. Y. T. Lii and T. V. Rajeevakumar, "Quarter Micron Deep Trench Etch for ULSI," *Proceedings of the 9th Symposium on Plasma Processing,* The Electrochemical Society, Pennington, NJ, 1992, p. 328.
15. H. C. Jones, R. Bennett, and J. Singh, "Size Dependent Etching of Small Shapes," *Proceeding of the 8th Symposium on Plasma Processing,* The Electrochemical Society, Pennington, NJ, 1990, p. 45.
16. D. J. Econnomou and R. C. Alkire, "Effect of Potential Field on Ion Reflection and Shape Evolution of Trenches during Plasma Assisted Etching," *J. Electrochem. Soc.,* **135,** 941 (1988).
17. J. W. Coburn and H. F. Winters, "Conductance Consideration in the Reactive Ion Etching of High Aspect Ratio Features," *Appl. Phys. Lett.,* **55,** 2730 (1989).
18. C. J. Mogab and H. J. Levinstein, "Anisotropic Plasma Etching of Polysilicon," *J. Vac. Sci. Technol.,* **17,** 721 (1980).
19. G. Y. Yeom and M. J. Kushner, "Magnetic Field Effects on Cylindrical Magnetron Reactive Ion Etching of Si/SiO_2 in CF_4/H_2 Plasma," *J. Vac. Sci. Technol.,* **A7,** 987 (1989).
20. K. Suzuki, S. Okudaira, N. Sakudo, and I. Kanomata, "Microwave Plasma Etching," *Jpn. J. Appl. Phys.,* **16,** 1979 (1977).
21. C. Chen, M. Chang, C. Yang, and T. Ebata, "Electron Cyclotron Resonance Microwave Plasma Etching of Submicron Polysilicon Gate Structures," *Proceeding of the 8th Symposium on Plasma Processing,* The Electrochemical Society, Pennington, NJ, 1990, p. 368.
22. D. Dane, P. Gadgil, T. D. Mantei, M. A. Carlson, and M. E. Weber, "Etching of Polysilicon in a High-density Electron Cyclotron Resonance Plasma with Collimated Magnetic Field," *J. Vac. Sci. Technol.,* **B10,** 1312 (1992).
23. A. Hall, "Electron Cyclotron Resonance Etching of AlSiCu," *Sol. State Technol.,* **34(5),** 107 (1991).
24. P. Singer, "Meeting Oxide, Poly and Metal Etch Requirements," *Semic. Int.,* 50 (April, 1993).
25. J. H. Keller, J. C. Forster, and M. S. Barnes, "Novel Radio-Frequency Induction Plasma Processing Techniques," *J. Vac. Sci. Technol.,* **A11,** 2487 (1993).
26. K. P. Giapis, N. Sadeghi, J. Margat, R. Gottscho, and T. C. Lee, "Limits to Ion Energy Control in High Density Discharge: Measurement of Absolute Metastable Ion Concentration," *J. Appl. Phys.,* **73,** 7188 (1993).
27. A. S. Bergendahl, D. V. Horak, P. E. Bakeman, and D. J. Miller, "Cluster Tools 2. 16Mb DRAM Processing," *Semic. Intl.,* **13(10),** 94 (1990).
28. J. W. Coburn, "Plasma-assisted Etching," *Plasma Chem. and Plasma Process.* **2,** 1 (1982).
29. D. B. Graves, "Plasma Processing in Microelectronics Manufacturing," *AIChE J.,* **35,** 1 (1989).
30. L. Peters, "Plasma Etch Chemistry: The Untold Story," *Semic. Intl.,* **15(6),** 66 (1992).
31. H. B. Pogge, J. A. Bondur, and P. J. Burkhardt, "Reactive Ion Etching of Silicon with Cl_2/Ar," *J. Electrochem. Soc.,* **130,** 1592 (1983).
32. H. Crazzolara and N. Gellrich, "Profile Control Possibilities for a Trench Etch Process Based on Chlorine Chemistry," *J. Electrochem. Soc.,* **137,** 708 (1990).
33. G. Wohl, A. Weisheit, I. Flohr, and M. Bottcher, "Trench Etching Using a $CBrF_3$ Plasma and Its Study by Optical Emission Spectroscopy," *Vacuum,* **42,** 905 (1991).
34. L. Y. Tsou, "Highly Selective Reactive Ion Etching of Polysilicon with Hydrogen Bromide," *J. Electrochem. Soc.,* **136,** 3003 (1989).
35. M. Sato, S. Kato, and Y. Arita, "Effect of Gas Species on the Depth Reduction in Silicon Deep-Submicron Trench Reactive Ion Etching," *Jpn. J. Appl. Phys.,* **30,** 1549 (1991).

Micron Ex. 1023, p. 397
Micron v. YMTC
IPR2024-00909

36. Y. T. Lii, C. M. Reeves, D. A. Danner, P. J. Coane, and L. K. Wang, "MRIE 0.1 μm Polysilicon Lines by Using HBr," *Proceedings of the 9th Symposium on Plasma Processing,* The Electrochemical Society, Pennington, NJ, 1992, p. 334.
37. M. Nakamura, K. Iizuka, and H. Yano, "Very High Selective n^+ Poly-Si RIE with Carbon Elimination," *Jpn. J. Appl. Phys.,* **28,** 2142 (1989).
38. J. W. Coburn, "Plasma-Assisted Etching," *Plasma Chem. and Plasma Process.,* **2,** 1 (1982).
39. H. Stocker, "Selective Reactive Ion Etching of Silicon Nitride on Oxide in a Multifacet ("HEX") Plasma Etching Machine," *J. Vac. Sci. Technol.,* **A7,** 1145 (1989).
40. Y. Kawamoto, T. Kure, N. Hashimoto, and T. Takaichi, "Highly Selective Dry Etching Si_3N_4," Electrochem. Soc. Meeting, 84-2, abs. 395 (1984).
41. K. Fujino and T. Oku, "Dry Etching of Al Alloy Films Using HBr Mixed Gases," *J. Electrochem. Soc.,* **139,** 2585 (1992).
42. D. W. Hess, "Plasma Etching of Aluminum," *Sol. State Technol.,* **24,** 189 (1981).
43. M. Sato and Y. Arita, "Al-Cu Alloy Etching Using In-Reactor Aluminum Chloride Formation in Static Magnetron Triode Reactor Ion Etching," *Jpn. J. Appl. Phys.,* **32,** 3013 (1993).
44. P. Riley, S. Peng, and L. Fang, "Plasma Etching of Aluminum for ULSI Circuits," *Sol. State Technol.,* **36(2),** 47 (1993).
45. D. S. Fischl and D. W. Hess, "Plasma Enhanced Etching of Tungsten and Tungsten Silicide in Chlorine Containing Discharges," *J. Electrochem. Soc.,* **134,** 2265 (1987).
46. N. Mutsukura and G. Turban, "Reactive Ion Etching of Tungsten in SF_6–N_2 Plasma," *J. Electrochem. Soc.,* **137,** 225 (1990).
47. I. Miller, R. Frazier, and M. Su, "Controlling Tungsten Etchback on Submicron Devices," *Microelectron. Manuf. Technol.,* 28 (Jan., 1992).
48. R. Rossen, "Magnetically Enhanced Patterned Tungsten Etching," *Microelectron. Manuf. Technol.,* 17 (March, 1991).
49. W. Pilz, J. Janes, K. P. Muller, and J. Pelka, "Oxygen Reactive Ion Etching of Polymers—Profile Evolution and Process Mechanisms," *SPIE, Vol. 1392, Advanced Techniques for Integrated Circuit Processing,* Society of Photo Instrumentation Engineers, 1990.
50. T. Kure, H. Kawakami, S. Tachi, and H. Enami, "Low Temperature Etching for Deep Submicron Trilayer Resist," *Jpn. J. Appl. Phys.,* **30,** 1562 (1991).
51. Y. Kawamoto, "MOS Gate Insulator Breakdown Caused by Exposure to Plasma," *Proceedings of the 7th Symposium on Dry Process,* Institute of Electrical Engineers, Tokyo, 1985, p. 132.
52. C. T. Gabriel and J. P. McVittie, "How Plasma Etching Damages Thin Gate Oxides," *Sol. State Technol.,* 81 (June, 1992).
53. G. S. Selwyn, J. Singh, and R. S. Bennett, "In Situ Laser Diagnostic Studies of Plasma-Generated Particulate," *J. Vac. Sci. Technol.,* **A7,** 2758 (1989).
54. J. L. Vossen and W. Kern, *Thin Film Processes,* Chapter V-1, Academic Press, New York, 1978, p. 401.
55. K. Tsujimoto, T. Kumihashi, and S. Tachi, "Novel Short-Gas-Residence-Time Electron Cyclotron Resonance Plasma Etching," *Appl. Phys. Lett.,* **63,** 1915 (1993).

PROBLEMS

1. Assume that the sheaths are pure capacitors in the plasma model and that positive ions of mass m_i come from the plasma and traverse the sheaths without colliding. The flux i_i is

Micron Ex. 1023, p. 398
Micron v. YMTC
IPR2024-00909

$$i_i = \frac{KV^{3/2}}{m_i^{1/2}D^2}$$

Derive the equation.

2. What are the major distinctions between traditional reactive-ion etching and high-density plasma etching (ECR, ICP, etc.)? Compare the advantages and limitations of these techniques.

3. Explain why end point detection techniques, such as optical emission and laser interferometry, can detect etch uniformity across a wafer surface?

4. The main etching reaction of chlorine plasma with Si is

$$Si + 4Cl \rightarrow SiCl_4$$

If 90% of input chlorine is consumed by silicon etching, the gas pressure of the etcher is 10 mtorr. In order to achieve a 1 μm/min silicon etch rate for an eight-inch wafer, what is the minimum pumping capacity required?

5. A multiple-step etch process is required for etching a polysilicon gate with thin gate oxide. How do you design an etch process that has no micromasking, has an anisotropic etch profile, and is selective to thin gate oxide?

6. Explain how the mechanisms of using extra Al in the RIE chamber can reduce residue formation during AlCu etching with Cl_2 plasma?

7. What are the major distinctions between bromine-based chemistry and chlorine-based chemistry for Al etching? Compare the advantages and limitations of these chemistries.

8. Based on your understanding of this chapter, briefly summarize the requirements of etching ULSI.

Micron Ex. 1023, p. 399
Micron v. YMTC
IPR2024-00909