UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
16/402,202	05/02/2019	Wenyu Hua	10018-01-0045-US

CONFIRMATION NO. 7288

153988
Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

PUBLICATION NOTICE

OC000000119529995

Title:THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

Publication No.US-2020-0273873-A1
Publication Date:08/27/2020

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Public Records Division. The Public Records Division can be reached by telephone at (571) 272-3150 or (800) 972-6382, by facsimile at (571) 273-3250, by mail addressed to the United States Patent and Trademark Office, Public Records Division, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently https://portal.uspto.gov/pair/PublicPair. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Managment, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

Micron Ex. 1002, p. 518
Micron v. YMTC
IPR2024-00789

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

153988 7590 10/26/2020
Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

EXAMINER
GHEYAS, SYED I

ART UNIT	PAPER NUMBER
2812	

DATE MAILED: 10/26/2020

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/402,202	05/02/2019	Wenyu Hua	10018-01-0045-US	7288

TITLE OF INVENTION: THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1200	$0.00	$0.00	$1200	01/26/2021

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Maintenance fees are due in utility patents issuing on applications filed on or after Dec. 12, 1980. It is patentee's responsibility to ensure timely payment of maintenance fees when due. More information is available at www.uspto.gov/PatentMaintenanceFees.

PTOL-85 (Rev. 02/11)

Micron Ex. 1002, p. 519
Micron v. YMTC
IPR2024-00789

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), by mail or fax, or via EFS-Web.

By mail, send to: Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

By fax, send to: (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

153988 7590 10/26/2020
Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being transmitted to the USPTO via EFS-Web or by facsimile to (571) 273-2885, on the date below.

(Typed or printed name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/402,202	05/02/2019	Wenyu Hua	10018-01-0045-US	7288

TITLE OF INVENTION: THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1200	$0.00	$0.00	$1200	01/26/2021

EXAMINER	ART UNIT	CLASS-SUBCLASS
GHEYAS, SYED I	2812	257-314000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-09 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1_____
2_____
3_____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document must have been previously recorded, or filed for recordation, as set forth in 37 CFR 3.11 and 37 CFR 3.81(a). Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. Fees submitted: ❑Issue Fee ❑Publication Fee (if required) ❑Advance Order - # of Copies _____

4b. Method of Payment: *(Please first reapply any previously paid fee shown above)*

❑ Electronic Payment via EFS-Web ❑ Enclosed check ❑ Non-electronic payment by credit card (Attach form PTO-2038)

❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment to Deposit Account No. _____

5. **Change in Entity Status** (from status indicated above)

❑ Applicant certifying micro entity status. See 37 CFR 1.29

❑ Applicant asserting small entity status. See 37 CFR 1.27

❑ Applicant changing to regular undiscounted fee status.

NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

Micron Ex. 1002, p. 520
Micron v. YMTC
IPR2024-00789

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/402,202	05/02/2019	Wenyu Hua	10018-01-0045-US	7288

153988 7590 10/26/2020
Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

EXAMINER
GHEYAS, SYED I

ART UNIT	PAPER NUMBER
2812	

DATE MAILED: 10/26/2020

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

PTOL-85 (Rev. 02/11)

Micron Ex. 1002, p. 521
Micron v. YMTC
IPR2024-00789

OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No. 16/402,202	Applicant(s) Hua et al.	
	Examiner SYED I GHEYAS	Art Unit 2812	AIA (FITF) Status Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☑ This communication is responsive to 08/18/2020.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☑ The allowed claim(s) is/are 1-14 and 21-26 . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information , please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐ All b) ☐ Some *c) ☐ None of the:

   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____ .

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.
5. ☐ Examiner's Amendment/Comment
6. ☑ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/SYED I GHEYAS/
Primary Examiner, Art Unit 2812

Micron Ex. 1002, p. 523
Micron v. YMTC
IPR2024-00789

Notice of Pre-AIA or AIA Status

1. The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

Allowable Subject Matter

Claims ***1-14 & 21-26*** are allowed.

The following is an examiner's statement of reasons for allowance:

Regarding Claim 1: The prior art of record to the examiner's knowledge does not teach or render obvious the instant invention, particularly characterized by a three-dimensional (3D) memory device, comprising: *wherein each of the dummy source structures comprises a conductive contact.*

The most relevant prior art reference due to **Jeong et al. (Pub. No.: US 2016/0293622 A1)** substantially discloses a three-dimensional (3D) memory device, comprising:

a substrate (Par 0055, Fig. 3 - substrate 100);

a memory stack comprising interleaved conductive layers and dielectric layers on the substrate (Par. 0057-0058, Fig. 3 - interleaved conductive layers 130, 135 & 138 and dielectric layers 110);

a staircase structure on one side of the memory stack (Par. 0057-0058, Fig. 3 – staircase structure (stepped structure) shown in region R2);

a staircase contact in the staircase structure (Par. 0068, Figs. 1-3 – staircase contact (cell contact plug) 150c shown in region R2); and

a plurality of dummy source structures each extending vertically through the staircase

structure, the plurality of dummy source structures surrounding the staircase contact (Par. 0062-0065 & 0082-0085, Figs. 1-3 – dummy pillars DP1 & DP2 shown in region R2 could be considered as the dummy source; dummy pillars of this prior art, like the dummy source of the instant application, function as a mechanical support for conductive layers).

Additionally, the prior arts made of record and not relied upon are considered pertinent to applicant's disclosure. See form PTO-892.

However, none of these prior art references indicated above or the prior arts made of record in form PTO-892, disclose all the limitations of claim 1 (the individual limitations may be found in a plurality of prior arts but there is no motivation to combine). Because no reference alone teaches all the limitations, nor is there any motivation to combine the prior arts to construct all the limitations of this independent claim, claim 1 is deemed patentable over the prior arts.

Regarding Claims 2-7 & 21-26: these claims are allowed because of their dependency status from claim 1.

Regarding Claim 8: The prior art of record to the examiner's knowledge does not teach or render obvious the instant invention, particularly characterized by a three-dimensional (3D) memory device, comprising: *wherein each of the dummy source structures comprises a conductive contact.*

The most relevant prior art reference due to **Jeong et al. (Pub. No.: US 2016/0293622 A1)** substantially discloses a three-dimensional (3D) memory device, comprising:

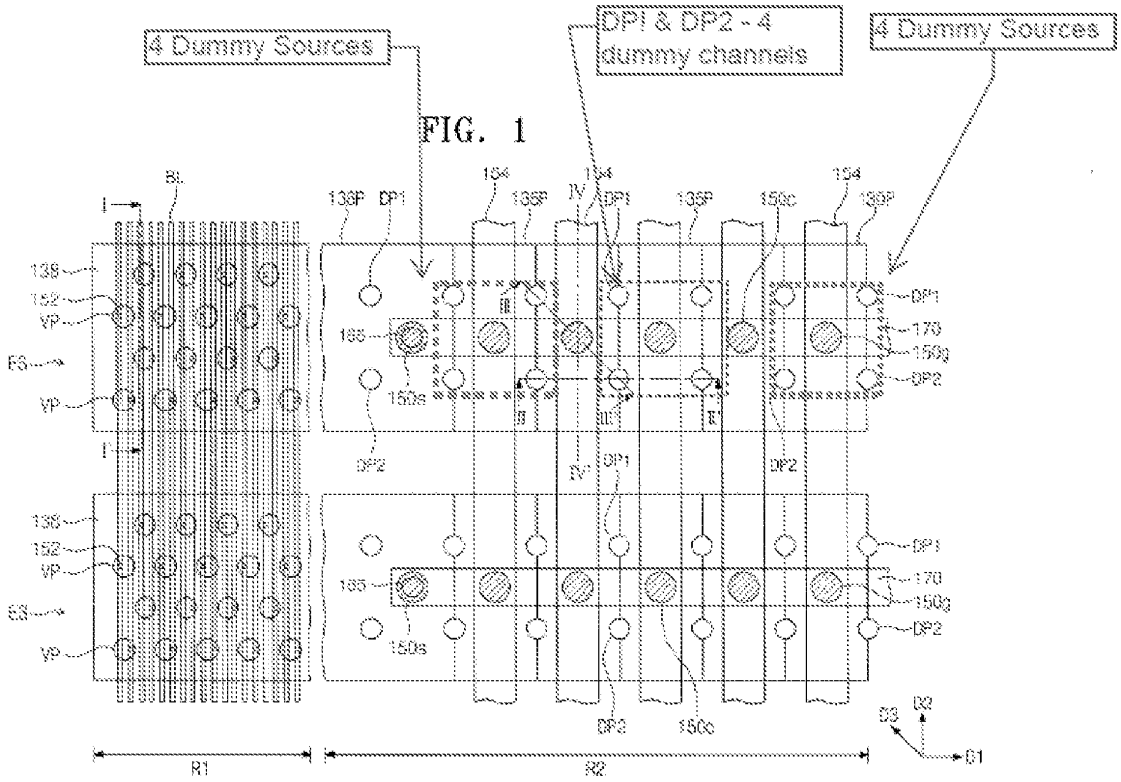
a substrate (Par 0055, Fig. 3 - substrate 100);

a memory stack comprising interleaved conductive layers and dielectric layers on the substrate (Par. 0057-0058, Fig. 3 - interleaved conductive layers 130, 135 & 138 and dielectric layers 110);

a staircase structure on one side of the memory stack (Par. 0057-0058, Fig. 3 – staircase structure (stepped structure) shown in region R2);

an array of dummy channel structures each extending vertically through the staircase structure (Par. 0062-0065 & 0082-0085; see annotated Fig. 1 below – some of the dummy pillars DP1 and DP2 could be considered as the dummy channel structures); and

a plurality of dummy source structures each extending vertically through the staircase stack, wherein a subset of the dummy channel structures is surrounded by the plurality of dummy source structures (Par. 0062-0065 & 0082-0085, Figs. 1-3 and annotated Fig. – some of the dummy pillars DP1 & DP2 shown in region R2 could be considered as the dummy source; dummy pillars of this prior art, like the dummy source of the instant application, function as a mechanical support; it is important to note that under BRI we are considering some of the dummy pillars DP1 and DP2 as dummy channels and others as the dummy sources; a subset of the dummy channel structures is at least partially surrounded by the plurality of dummy source structures).



Additionally, the prior arts made of record and not relied upon are considered pertinent to applicant's disclosure. See form PTO-892.

However, none of these prior art references indicated above or the prior arts made of record in form PTO-892, disclose all the limitations of claim 8 (the individual limitations may be found in a plurality of prior arts but there is no motivation to combine). Because no reference alone teaches all the limitations, nor is there any motivation to combine the prior arts to construct all the limitations of this independent claim, claim 8 is deemed patentable over the prior arts.

Regarding Claims 9-14: these claims are allowed because of their dependency status from claim 8.

Conclusion

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SYED I GHEYAS whose telephone number is (571)272-0592. The examiner can normally be reached on M-F 8:30 AM - 5:30 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Charles Garber can be reached on 571-272-2194. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

10/17/2020
/SYED I GHEYAS/
Primary Examiner, Art Unit 2812

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/402,202	Hua et al.
	Examiner	**Art Unit**
	SYED I GHEYAS	2812

CPC - Searched*		
Symbol	**Date**	**Examiner**
H01L27/11582, H01L23/535, H01L29/0847, H01L29/1037	06/03/2020	Syed Gheyas

CPC Combination Sets - Searched*		
Symbol	**Date**	**Examiner**

US Classification - Searched*			
Class	**Subclass**	**Date**	**Examiner**

* See search history printout included with this form or the SEARCH NOTES box below to determine the scope of the search.

Search Notes		
Search Notes	**Date**	**Examiner**
PE2E search	06/03/2020 & 10/17/2020	Syed Gheyas
NPL search using Google.com & IP.com	06/03/2020	Syed Gheyas

Interference Search			
US Class/CPC Symbol	**US Subclass/CPC Group**	**Date**	**Examiner**
	see PE2E search	10/17/2020	Syed Gheyas

/SYED I GHEYAS/ Primary Examiner, Art Unit 2812	

Micron Ex. 1002, p. 530
Micron v. YMTC
IPR2024-00789

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/402,202	Hua et al.
	Examiner	**Art Unit**
	SYED I GHEYAS	2812

CPC						
Symbol					**Type**	**Version**
H01L	/	27	/	11582	F	2013-01-01
H01L	/	23	/	535	I	2013-01-01
H01L	/	29	/	0847	I	2013-01-01
H01L	/	29	/	1037	I	2013-01-01

CPC Combination Sets								
Symbol					**Type**	**Set**	**Ranking**	**Version**
	/		/					

NONE (Assistant Examiner)	(Date)	**Total Claims Allowed:** 20	
/SYED I GHEYAS/ Primary Examiner, Art Unit 2812 (Primary Examiner)	17 October 2020 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 1D

U.S. Patent and Trademark Office

Part of Paper No.: 20201017

Micron Ex. 1002, p. 531
Micron v. YMTC
IPR2024-00789

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/402,202	Hua et al.
	Examiner	**Art Unit**
	SYED I GHEYAS	2812

INTERNATIONAL CLASSIFICATION				
CLAIMED				
H01L	/	27	/	11582

NON-CLAIMED				
	/		/	

US ORIGINAL CLASSIFICATION	
CLASS	**SUBCLASS**

CROSS REFERENCES(S)						
CLASS	**SUBCLASS (ONE SUBCLASS PER BLOCK)**					

NONE (Assistant Examiner) (Date)		**Total Claims Allowed:** 20	
/SYED I GHEYAS/ Primary Examiner, Art Unit 2812 (Primary Examiner)	17 October 2020 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 1D

U.S. Patent and Trademark Office Part of Paper No.: 20201017

Micron Ex. 1002, p. 532
Micron v. YMTC
IPR2024-00789

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/402,202	Hua et al.
	Examiner	**Art Unit**
	SYED I GHEYAS	2812

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIMS

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1	16	10		19										
2	2	17	11		20										
3	3	18	12	8	21										
4	4	19	13	9	22										
5	5	20	14	10	23										
6	6		15	11	24										
7	7		16	12	25										
14	8		17	13	26										
15	9		18												

NONE (Assistant Examiner)	(Date)	**Total Claims Allowed:** 20	
/SYED I GHEYAS/ Primary Examiner, Art Unit 2812 (Primary Examiner)	17 October 2020 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 1D

U.S. Patent and Trademark Office Part of Paper No.: 20201017

Micron Ex. 1002, p. 533
Micron v. YMTC
IPR2024-00789

Bibliographic Data

Application No: 16/402,202

Foreign Priority claimed: ○ Yes ◉ No

35 USC 119 (a-d) conditions met: ☐ Yes ☑ No ☐ Met After Allowance

Verified and Acknowledged: /SYED I GHEYAS/ (Examiner's Signature) [ ] (Initials)

Title: THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

FILING or 371(c) DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
05/02/2019 **RULE**	257	2812	10018-01-0045-US

APPLICANTS

YANGTZE MEMORY TECHNOLOGIES CO., LTD., Wuhan, CHINA

INVENTORS

Wenyu Hua, Wuhan, CHINA

Fandong Liu, Wuhan, CHINA

Zhiliang Xia, Wuhan, CHINA

CONTINUING DATA

This application is a CON of PCT/CN2019/076186 02/26/2019

FOREIGN APPLICATIONS

IF REQUIRED, FOREIGN LICENSE GRANTED**

05/16/2019

STATE OR COUNTRY

CHINA

ADDRESS

Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102
UNITED STATES

FILING FEE RECEIVED

$1,720

PE2E SEARCH - Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	British Equivalents	Time Stamp
L1	0	"16402202"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 03:41 PM
L2	9	("8822285"|"20190043883").PN.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 03:44 PM
L3	354	(Wenyu near2 Hua).in. (Fandong near2 Liu).in. (Zhiliang near2 Xia).in.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:14 PM
L4	1469	(YANGTZE near2 MEMORY).as.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:15 PM
L5	14	L4 and L3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:16 PM
L6	11806	(YANGTZE).as.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:16 PM
L7	14	L6 and L3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:17 PM
L8	62816	((three-dimension$2) or (3D)) near5 (memory)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:23 PM
L9	746	L6 and L8	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/26 05:28 PM
L10	1470	(((three-dimension$2) or (3D)) near5 (memory)) and (staircase)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:48 AM
L11	1481	(((three-dimension$2) or (3D)) near5 (memory)) and (staircase or stair adj	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT;	OR	ON	ON	2020/05/27 09:49 AM

Micron Ex. 1002, p. 535
Micron v. YMTC
IPR2024-00789

		case)	IBM_TDB				
L12	11806	(YANGTZE).as.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:50 AM
L13	68	L12 and L11	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:50 AM
L14	104	(((three-dimension$2) or (3D)) near5 (memory)) and (staircase or stair adj case) and @ad<"20100101"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:51 AM
L15	43	(((three-dimension$2) or (3D)) near5 (memory)) and (staircase or stair adj case) and @ad<"20050101"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:53 AM
L16	3	(((three-dimension$2) or (3D)) near5 (memory)).ab. and (staircase or stair adj case) and @ad<"20050101"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:54 AM
L17	10	(((three-dimension$2) or (3D)) near5 (memory)).ab. and (staircase or stair adj case) and @ad<"20100101"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:55 AM
L18	61	(((three-dimension$2) or (3D)) near5 (memory)).ab. and (staircase or stair adj case) and (interleav$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:56 AM
L19	332	(((three-dimension$2) or (3D)) near5 (memory)).ab. and (staircase or stair adj case) and (interleav$3 or alternat$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/05/27 09:56 AM
L20	53524	(H01L27/11582 or H01L23/535 or H01L29/0847 or H01L29/1037).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 08:49 AM
L21	29263	(H01L29/0847 or H01L29/1037).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 08:50 AM

Micron Ex. 1002, p. 536
Micron v. YMTC
IPR2024-00789

L22	25616	(H01L29/0847).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 08:50 AM
L23	12013	H01L27/11524.cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:21 PM
L24	3922	H01L27/11529.cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:22 PM
L25	3922	H01L27/11529.cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:22 PM
L26	15267	(H01L27/11526-11548).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:24 PM
L27	9886	(H01L27/11526).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:25 PM
L28	2242	(H01L27/11548).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:25 PM
L29	12690	(H01L27/11551-11556).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:26 PM
L30	4003	(H01L27/11551).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:26 PM
L31	10148	(H01L27/11556).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:26 PM
L32	14400	L23 or L24	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:27 PM
L33	4649	L32 and L29	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:27 PM

Micron Ex. 1002, p. 537
Micron v. YMTC
IPR2024-00789

L34	1840	L33 and L26	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:29 PM
L35	4724	dummy near2 source	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:29 PM
L36	4832	(dummy or nonfunctional) near2 source	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:30 PM
L37	5020	(dummy or nonfunctional or non adj functional) near2 source	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:30 PM
L38	5060	(dummy or nonfunction$3 or non adj function$3) near2 source	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:31 PM
L39	23	L34 and L38	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:31 PM
L40	2	"20180337192".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 02:54 PM
L41	2	"20160293622".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 03:19 PM
L42	4	"20130161182"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 03:32 PM
L43	4	"20130161821"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 03:33 PM
L44	315	((dummy or nonfunction$3 or non adj function$3) near2 (source)) and ((dummy or nonfunction$3 or non adj function$3) near2 (channel))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:13 PM
L45	12013	H01L27/11524.cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO;	OR	ON	ON	2020/06/03 06:21 PM

Micron Ex. 1002, p. 538
Micron v. YMTC
IPR2024-00789

			JPO; DERWENT; IBM_TDB				
L46	3922	H01L27/11529.cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L47	15267	(H01L27/11526-11548).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L48	12690	(H01L27/11551-11556).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L49	14400	L45 or L46	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L50	4649	L49 and L48	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L51	1840	L50 and L47	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L52	10	L51 and L44	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/03 06:21 PM
L53	14413	(H01L27/11524 or H01L27/11529).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:26 PM
L54	12702	(H01L27/11551-11556).cpc.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:36 PM
L55	4656	L54 and L53	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:36 PM
L56	317	((dummy or nonfunction$3 or non adj function$3) near2 (source)) and ((dummy or nonfunction$3 or non adj function$3) near2 (channel))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:37 PM

Micron Ex. 1002, p. 539
Micron v. YMTC
IPR2024-00789

L57	11	L55 and L56	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:37 PM
L58	237	L54 and L53 and (stair)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:39 PM
L59	507	L54 and L53 and (stair or staircase)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:39 PM
L60	507	L54 and L53 and (stair or staircase or stair adj case)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:39 PM
L61	507	L54 and L53 and (stair or staircase or stair-case)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:40 PM
L62	507	L54 and L53 and (stair or staircase or stair adj case)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:40 PM
L63	1908	L54 and L53 and (stair or staircase or stair adj case or step)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:41 PM
L64	10	L63 and L56	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 12:41 PM
L65	85	(NAND) and ((dummy or nonfunction$3 or non adj function$3) near2 (source)) and ((dummy or nonfunction$3 or non adj function$3) near2 (channel))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 02:30 PM
L66	80	L65 NOT L64	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 02:31 PM
L67	4	"20180026047"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/06/04 03:10 PM
L68	3	(NAND) AND ((dummy OR nonfunction$3 OR	US-PGPUB; USPAT; USOCR; FPRS; EPO;	OR	ON	ON	2020/10/17 03:23 PM

Micron Ex. 1002, p. 540
Micron v. YMTC
IPR2024-00789

		non ADJ function$3) near2 (source)) AND ((dummy OR nonfunction$3 OR non ADJ function$3) near2 (metal OR conductor))	JPO; DERWENT; IBM_TDB				
L69	32	(NAND) AND (((dummy OR nonfunction$3 OR non ADJ function$3) near2 (source OR pillar)) WITH (metal OR conductor))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	ON	2020/10/17 03:46 PM

PE2E SEARCH - Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	British Equivalents	Time Stamp
N1	36	(((dummy OR nonfunction$3 OR non ADJ function$3) near2 (source OR pillar)) WITH (metal OR conductor)).clm.	US-PGPUB; USPAT	OR	ON	ON	2020/10/17 07:28 PM
N2	12	((((dummy OR nonfunction$3 OR non ADJ function$3) near2 (source OR pillar)) WITH (metal OR conductor)) AND (memory)).clm.	US-PGPUB; USPAT	OR	ON	ON	2020/10/17 07:28 PM

Micron Ex. 1002, p. 541
Micron v. YMTC
IPR2024-00789

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), by mail or fax, or via EFS-Web.

By mail, send to: Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

By fax, send to: (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

153988 7590 10/26/2020
Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being transmitted to the USPTO via EFS-Web or by facsimile to (571) 273-2885, on the date below.

Anna J. O'Connor	(Typed or printed name)
/Anna J. O'Connor/	(Signature)
October 27, 2020 E-FILED	(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/402,202	05/02/2019	Wenyu Hua	10018-01-0045-US	7288

TITLE OF INVENTION: THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1200	$0.00	$0.00	$1200	01/26/2021

EXAMINER	ART UNIT	CLASS-SUBCLASS
GHEYAS, SYED I	2812	257-314000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-09 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 BAYES PLLC
2
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document must have been previously recorded, or filed for recordation, as set forth in 37 CFR 3.11 and 37 CFR 3.81(a). Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE	(B) RESIDENCE: (CITY and STATE OR COUNTRY)
YANGTZE MEMORY TECHNOLOGIES CO., LTD.	WUHAN, CHINA

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ☒ Corporation or other private group entity ❑ Government

4a. Fees submitted: ☒ Issue Fee ❑ Publication Fee (if required) ❑ Advance Order - # of Copies _____

4b. Method of Payment: *(Please first reapply any previously paid fee shown above)*

☒ Electronic Payment via EFS-Web ❑ Enclosed check ❑ Non-electronic payment by credit card (Attach form PTO-2038)

☒ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment to Deposit Account No. 60-2227

5. **Change in Entity Status** (from status indicated above)

❑ Applicant certifying micro entity status. See 37 CFR 1.29

❑ Applicant asserting small entity status. See 37 CFR 1.27

❑ Applicant changing to regular undiscounted fee status.

NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date 10/27/2020

Typed or printed name Zhiwei Zou Registration No. 66,041

PTOL-85 Part B (08-18) Approved for use through 01/31/2020 OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Micron Ex. 1002, p. 542
Micron v. YMTC
IPR2024-00789

Electronic Patent Application Fee Transmittal

Application Number:	16402202
Filing Date:	02-May-2019
Title of Invention:	THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME
First Named Inventor/Applicant Name:	Wenyu Hua
Filer:	Zhiwei Zou/Anna O'Connor
Attorney Docket Number:	10018-01-0045-US

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
UTILITY APPL ISSUE FEE	1501	1	1200	1200

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**1200**

Electronic Acknowledgement Receipt

EFS ID:	40964678
Application Number:	16402202
International Application Number:	
Confirmation Number:	7288
Title of Invention:	THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME
First Named Inventor/Applicant Name:	Wenyu Hua
Customer Number:	153988
Filer:	Zhiwei Zou/Anna O'Connor
Filer Authorized By:	Zhiwei Zou
Attorney Docket Number:	10018-01-0045-US
Receipt Date:	27-OCT-2020
Filing Date:	02-MAY-2019
Time Stamp:	21:40:53
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$1200
RAM confirmation Number	E20200QL41152265
Deposit Account	602227
Authorized User	Anna O'Connor

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.16 (National application filing, search, and examination fees)

37 CFR 1.17 (Patent application and reexamination processing fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Issue Fee Payment (PTO-85B)	10018-01-0045-US_IFTransmittal.pdf	290642 afe637f2a3281a465c41524ae207bc42c2e80eae	no	1
Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	30367 d048b930f9d53323fad265eeffbc574cd3cffc74	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	321009		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.
National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.
New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Not for submission under 37 CFR 1.99)

Application Number	16402202
Filing Date	2019-05-02
First Named Inventor	Wenyu HUA
Art Unit	2812
Examiner Name	GHEYAS, SYED I.
Attorney Docket Number	10018-01-0045-US

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	108886040	CN	A	2018-11-23	SANDISK TECHNOLOGIES LLC		
	2	104157654	CN	A	2014-11-19	INST OF MICROELECTRONICS CAS		
	3	2018071143	WO	A2	2018-04-19	MONOLITHIC 3D INC.		

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	16402202
	Filing Date	2019-05-02
	First Named Inventor	Wenyu HUA
	Art Unit	2812
	Examiner Name	GHEYAS, SYED I.
	Attorney Docket Number	10018-01-0045-US

	4	108695336	CN	A	2018-10-23	SAMSUNG ELECTRONICS CO LTD		

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1		

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	16402202
	Filing Date	2019-05-02
	First Named Inventor	Wenyu HUA
	Art Unit	2812
	Examiner Name	GHEYAS, SYED I.
	Attorney Docket Number	10018-01-0045-US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☒ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

☒ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Zhiwei Zou/	Date (YYYY-MM-DD)	2020-10-30
Name/Print	Zhiwei Zou	Registration Number	66,041

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Micron Ex. 1002, p. 549
Micron v. YMTC
IPR2024-00789

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Espacenet

CN108886040A Three-dimensional memory device with short-free source select gate contact via structure and method of making same

Applicants: SANDISK TECHNOLOGIES LLC

Inventors: ARIYOSHI JUNICHI

Classifications:

IPC ***H01L27/11565; H01L27/1157; H01L27/11573; H01L27/11582;***

CPC **H01L23/5226 (US); H01L27/11524 (US); H01L27/11556 (EP,US); H01L27/11565 (EP,US); H01L27/1157 (US); H01L27/11573 (US); H01L27/11575 (EP); H01L27/11582 (EP,US);**

Priorities: US2017063478W 2017-11-28; US201715632703A 2017-06-26; US201762469618P 2017-03-10

Application: CN108886040A·2017-11-28

Publication: CN108886040A·2018-11-23

Published as: CN108886040A;**US10256245B2; US2018261613A1; WO2018164743A1**

Three-dimensional memory device with short-free source select gate contact via structure and method of making same

Abstract

Electrical short caused by misalignment of source select level contact via structure and support pillar structures can be prevented by modifying the pattern of the support pillar structures such thatthe support pillar structures are omitted from the area in which source select gate contact via structures are formed. The insulating layer at the level overlying the source select level electricallyconductive layer can have a sufficient thickness to prevent deformation during formation of the backside recesses. A minimum lateral separation distance between the source select level contact via structure and the support pillar structures is greater than any minimum lateral separation distance between the word line level contact via structures and the support pillar structures.

Micron Ex. 1002, p. 551
Micron v. YMTC
IPR2024-00789

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 108886040 A
(43)申请公布日 2018.11.23

(21)申请号 201780002980.0

(22)申请日 2017.11.28

(30)优先权数据
62/469,618 2017.03.10 US
15/632,703 2017.06.26 US

(85)PCT国际申请进入国家阶段日
2018.03.01

(86)PCT国际申请的申请数据
PCT/US2017/063478 2017.11.28

(87)PCT国际申请的公布数据
WO2018/164743 EN 2018.09.13

(71)申请人 桑迪士克科技有限责任公司
地址 美国得克萨斯州

(72)发明人 有吉润一

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 邱军

(51)Int.Cl.
H01L 27/11565(2006.01)
H01L 27/1157(2006.01)
H01L 27/11573(2006.01)
H01L 27/11582(2006.01)

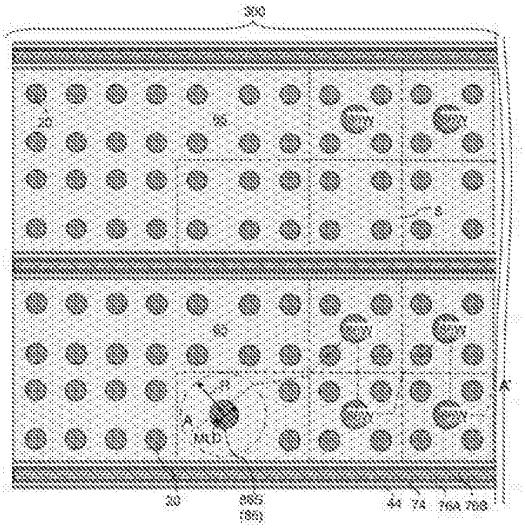
权利要求书4页 说明书21页 附图30页

(54)发明名称

具有无短路源极选择栅极接触通孔结构的三维存储器器件及其制造方法

(57)摘要

通过改变支撑柱结构的图案，使得从其中形成源极选择栅极接触通孔结构的区域中省略支撑柱结构，可以防止由源极选择级接触通孔结构和支撑柱结构的未对准导致的电短路。源极选择级导电层上面的级处的绝缘层可以具有足够的厚度，以在形成背侧凹陷期间防止变形。源极选择级接触通孔结构与支撑柱结构之间的最小横向分离距离大于字线级接触通孔结构与支撑柱结构之间的任何最小横向分离距离。



CN 108886040 A

Micron Ex. 1002, p. 552
Micron v. YMTC
IPR2024-00789

1.一种三维存储器器件,包括:

位于基板之上的绝缘层和导电层的交替堆叠体,其中所述导电层包括源极选择级导电层和位于所述源极选择级导电层之上的字线级导电层;

存储器堆叠体结构,所述存储器堆叠体结构延伸穿过所述交替堆叠体,其中所述存储器堆叠体结构中的每一个包括存储器膜和垂直半导体沟道,所述垂直半导体沟道接触所述存储器膜的内侧壁;

梯台区域,在所述梯台区域中,除所述交替堆叠体内的最顶部导电层之外的每个导电层比所述交替堆叠体内的上面的导电层横向地延伸得更远;

倒退阶梯式电介质材料部分,所述倒退阶梯式电介质材料部分在所述梯台区域上面;

支撑柱结构,所述支撑柱结构位于所述梯台区域中,并且延伸穿过所述倒退阶梯式电介质材料部分和所述交替堆叠体;

源极选择级接触通孔结构,所述源极选择级接触通孔结构延伸穿过所述倒退阶梯式电介质材料部分,并且接触所述源极选择级导电层;以及

字线级接触通孔结构,所述字线级接触通孔结构延伸穿过所述倒退阶梯式电介质材料部分,并且接触所述字线级导电层中的相应的一个;

其中所述源极选择级接触通孔结构与所述支撑柱结构之间的最小横向分离距离大于所述字线级接触通孔结构与所述支撑柱结构之间的任何最小横向分离距离。

2.根据权利要求1所述的三维存储器器件,其中:

所述源极选择级导电层包括源极选择栅电极;并且

与穿过所述源极选择栅极接触通孔结构的几何中心的垂直轴线相距所述源极选择栅极接触通孔结构的最大横向尺寸的两倍内的整个体积不含任何支撑柱结构。

3.根据权利要求1所述的三维存储器器件,其中

阶梯式表面在所述梯台区域中从所述交替堆叠体的最底层连续延伸到所述交替堆叠体的最顶层;并且

所述支撑柱结构穿过所述阶梯式表面的每个水平表面。

4.根据权利要求3所述的三维存储器器件,其中所述字线级接触通孔结构中的每一个具有相应的最大横向尺寸,并且与所述支撑柱结构中的最接近的一个横向地间隔开小于所述相应的最大横向尺寸的横向距离。

5.根据权利要求1所述的三维存储器器件,其中所述源极选择级接触通孔结构与所述源极选择级导电层的最接近的侧壁横向地间隔开大于所述源极选择栅极接触通孔结构的最大横向尺寸的横向分离距离。

6.根据权利要求1所述的三维存储器器件,其中所述支撑柱结构中的每一个包括基座沟道部分,所述基座沟道部分接触所述基板中的半导体材料层。

7.根据权利要求6所述的三维存储器器件,其中所述支撑柱结构中的每一个包括管状电介质间隔体,所述管状电介质间隔体位于所述基座沟道部分与所述源极选择级导电层之间。

8.根据权利要求6所述的三维存储器器件,其中所述支撑柱结构中的每一个包括存储器膜、位于所述存储器膜内并且接触所述基座沟道部分的垂直半导体沟道、以及位于所述垂直半导体沟道的顶端部处的漏极区域。

Micron Ex. 1002, p. 553
Micron v. YMTC
IPR2024-00789

9.根据权利要求8所述的三维存储器器件，其中所述支撑柱结构中的一个穿过所述源极选择级导电层和所述倒退阶梯式电介质材料部分，包含与所述倒退阶梯式电介质材料部分接触的相应的基座沟道部分，并且不穿过任何所述字线级导电层。

10.根据权利要求8所述的三维存储器器件，其中：

每个存储器堆叠体结构在相应的基座沟道部分上面；

每个垂直半导体沟道的上端部与相应的漏极区域的接触；

接触存储器堆叠体结构的每个漏极区域与相应的漏极接触通孔结构接触，所述漏极接触通孔结构嵌入在接触级电介质层内，所述接触级电介质层在所述存储器堆叠体结构上面；并且

所述支撑柱结构中的每一个的整个顶表面与所述接触级电介质层的底表面接触。

11.根据权利要求1所述的三维存储器器件，其中：

所述三维存储器器件包括单片三维NAND存储器器件；

所述字线级导电层包括或电连接到所述单片三维NAND存储器器件的相应的字线；

所述基板包括硅基板；

所述单片三维NAND存储器器件包括所述硅基板之上的单片三维NAND串的阵列；

所述单片三维NAND串的阵列的第一器件级中的至少一个存储器单元位于所述单片三维NAND串的阵列的第二器件级中的另一存储器单元之上；

所述硅基板含有集成电路，所述集成电路包括位于其上的所述存储器器件的驱动器电路；

所述字线级导电层包括具有条形的多个控制栅电极，所述条形实质上平行于所述基板的顶表面延伸，所述多个控制栅电极至少包括位于所述第一器件级中的第一控制栅电极和位于所述第二器件级中的第二控制栅电极；并且所述单片三维NAND串的阵列包括：

多个半导体沟道，其中所述多个半导体沟道中的每一个的至少一个端部部分实质上垂直于所述基板的顶表面延伸，以及

多个电荷储存元件，每个电荷储存元件位于所述多个半导体沟道中的相应的一个附近。

12.一种形成三维存储器器件的方法，包括：

在基板之上形成绝缘层和牺牲材料层的交替堆叠体；

形成梯台区域，在所述梯台区域中，所述交替堆叠体内的除最顶部牺牲材料层之外的每个牺牲材料层比所述交替堆叠体内的上面的牺牲材料层横向地延伸得更远；

在所述梯台区域之上形成倒退阶梯式电介质材料部分；

形成延伸穿过所述交替堆叠体的存储器堆叠体结构和支撑柱结构；

用导电层取代所述牺牲材料层；以及

穿过所述倒退阶梯式电介质材料部分并直接在所述导电层中的相应的一个的顶表面上形成接触通孔结构，

其中：

所述导电层包括源极选择级导电层和在所述源极选择级导电层上面的字线级导电层；

源极选择级接触通孔结构接触所述源极选择级导电层；

字线级接触通孔结构接触所述字线级导电层中的相应的一个；并且

Micron Ex. 1002, p. 554
Micron v. YMTC
IPR2024-00789

所述源极选择级接触通孔结构与所述支撑柱结构之间的最小横向分离距离大于所述字线级接触通孔结构与所述支撑柱结构之间的任何最小横向分离距离。

13.根据权利要求12所述的方法，还包括：

在存储器阵列区域中穿过所述交替堆叠体内的每层形成存储器开口；和

在所述梯台区域中穿过所述倒退阶梯式电介质材料部分形成支撑开口；以及

在所述存储器开口和所述支撑开口中的每一个内形成存储器膜和垂直半导体沟道，其中所述存储器堆叠体结构和所述支撑柱结构分别形成在所述存储器开口中和所述支撑开口中。

14.根据权利要求13所述的方法，还包括在所述存储器开口和所述支撑开口中的每一个的底部部分处形成基座沟道部分，其中每个存储器膜和每个垂直半导体沟道直接形成在相应的基座沟道部分上。

15.根据权利要求14所述的方法，其中通过选择性半导体沉积工艺来形成每个存储器开口中的所述基座沟道部分，所述选择性半导体沉积工艺从所述基板中的半导体材料层生长半导体材料且不从电介质表面生长半导体材料。

16.根据权利要求14所述的方法，其中：

每个基座沟道部分的顶表面形成在包含最底部牺牲材料层的顶表面的水平平面上方；并且

在穿过所述最底部牺牲材料层且不穿过任何其他牺牲材料层的支撑开口中形成的基座沟道部分接触所述倒退阶梯式电介质材料部分的侧壁。

17.根据权利要求12所述的方法，其中：

所述源极选择级导电层包括源极选择栅电极；并且

与穿过所述源极选择栅极接触通孔结构的几何中心的垂直轴线相距所述源极选择栅极接触通孔结构的最大横向尺寸的两倍内的整个体积不含任何支撑柱结构。

18.根据权利要求12所述的方法，其中：

形成所述梯台区域包括形成阶梯式表面，所述阶梯式表面从所述交替堆叠体的最底层连续延伸到所述交替堆叠体的最顶层；

穿过所述阶梯式表面的水平表面形成所述支撑柱结构的子集；并且

所述字线级接触通孔结构中的每一个形成有相应的最大横向尺寸，并且与所述支撑柱结构中的最近的一个横向地间隔开小于所述相应的最大横向尺寸的横向距离。

19.根据权利要求12所述的方法，其中所述源极选择级接触通孔结构形成在与所述源极选择级导电层的最接近的侧壁横向地间隔开大于所述源极选择栅极接触通孔结构的最大横向尺寸的横向分离距离的位置处。

20.根据权利要求12所述的方法，其中：

所述三维存储器器件包括单片三维NAND存储器器件；

所述字线级导电层包括或电连接到所述单片三维NAND存储器器件的相应的字线；

所述基板包括硅基板；

所述单片三维NAND存储器器件包括所述硅基板之上的单片三维NAND串的阵列；

所述单片三维NAND串的阵列的第一器件级中的至少一个存储器单元位于所述单片三维NAND串的阵列的第二器件级中的另一存储器单元之上；

Micron Ex. 1002, p. 555
Micron v. YMTC
IPR2024-00789

所述硅基板含有集成电路，所述集成电路包括位于其上的所述存储器器件的驱动器电路；

所述字线级导电层包括具有条形的多个控制栅电极，所述条形实质上平行于所述基板的顶表面延伸，所述多个控制栅电极至少包括位于所述第一器件级中的第一控制栅电极和位于所述第二器件级中的第二控制栅电极；并且

所述单片三维NAND串的阵列包括：

多个半导体沟道，其中所述多个半导体沟道中的每一个的至少一个端部部分实质上垂直于所述基板的顶表面延伸，以及

多个电荷储存元件，每个电荷储存元件位于所述多个半导体沟道中的相应的一个附近。

Micron Ex. 1002, p. 556
Micron v. YMTC
IPR2024-00789

具有无短路源极选择栅极接触通孔结构的三维存储器器件及其制造方法

[0001] 相关申请

[0002] 本申请要求于2017年3月10日提交的申请号为62/469,618的美国临时专利申请和于2017年6月26日提交的申请号为15/632,703的美国非临时申请的优先权,其全部内容通过引用并入本文。

技术领域

[0003] 本公开总体上涉及半导体器件的领域,并且特别地,涉及采用无短路源极选择栅极接触通孔结构的三维存储器器件及其制造方法。

背景技术

[0004] 在T.Endoh等人在IEDM Proc.(2001)33-36的题为“Novel Ultra High Density Memory With A Stacked-Surrounding Gate Transistor(S-SGT)Structured Cell”的文章中公开了每单元具有一个位的三维垂直NAND串。

发明内容

[0005] 根据本公开的方面,提供了一种三维存储器器件,其包括:位于基板之上的绝缘层和导电层的交替堆叠体,其中导电层包括源极选择级导电层和位于源极选择级导电层之上的字线级导电层;存储器堆叠体结构,存储器堆叠体结构延伸穿过交替堆叠体,其中存储器堆叠体结构中的每一个包括存储器膜和垂直半导体沟道,并且垂直半导体沟道接触存储器膜的内侧壁;梯台区域,在梯台区域中,除交替堆叠体内的最顶部导电层之外的每个导电层比交替堆叠体内的上面的导电层横向地延伸得更远;倒退阶梯式电介质材料部分,倒退阶梯式电介质材料部分在梯台区域上面;支撑柱结构,支撑柱结构位于梯台区域中,并且延伸穿过倒退阶梯式电介质材料部分和交替堆叠体;源极选择级接触通孔结构,源极选择级接触通孔结构延伸穿过倒退阶梯式电介质材料部分,并且接触源极选择级导电层;以及字线级接触通孔结构,字线级接触通孔结构延伸穿过倒退阶梯式电介质材料部分,并且接触字线级导电层中的相应的一个。源极选择级接触通孔结构与支撑柱结构之间的最小横向分离距离大于字线级接触通孔结构与支撑柱结构之间的任何最小横向分离距离。

[0006] 根据本公开的另一方面,提供了一种形成三维存储器器件的方法,其包括以下步骤:形成梯台区域,在梯台区域中,交替堆叠体内的除最顶部牺牲材料层之外的每个牺牲材料层比交替堆叠体内的上面的牺牲材料层横向地延伸得更远;在基板之上形成绝缘层和牺牲材料层的交替堆叠体;形成延伸穿过交替堆叠体的存储器堆叠体结构和支撑柱结构;在梯台区域之上形成倒退阶梯式电介质材料部分;用导电层取代牺牲材料层;以及穿过倒退阶梯式电介质材料部分并直接在导电层中的相应的一个的顶表面上形成接触通孔结构。导电层包括源极选择级导电层和在源极选择级导电层上面的字线级导电层。源极选择级接触通孔结构接触源极选择级导电层。字线级接触通孔结构接触字线级导电层中的相应的一

Micron Ex. 1002, p. 557
Micron v. YMTC
IPR2024-00789

个。源极选择级接触通孔结构与支撑柱结构之间的最小横向分离距离大于字线级接触通孔结构与支撑柱结构之间的任何最小横向分离距离。

附图说明

[0007] 图1是根据本公开的实施例的示例性结构在形成至少一个外围器件、半导体材料层和栅极电介质层之后的示意性垂直截面图。

[0008] 图2是根据本公开的实施例的示例性结构在形成绝缘层和牺牲材料层的交替堆叠体之后的示意性垂直截面图。

[0009] 图3是根据本公开的实施例的示例性结构在形成阶梯式梯台和倒退阶梯式电介质材料部分之后的示意性垂直截面图。

[0010] 图4A是根据本公开的实施例的示例性结构在形成存储器开口和支撑开口之后的曲折(zig-zag)示意性垂直截面图。

[0011] 图4B是图4A的示例性结构的俯视图。垂直平面A-A'是图4A的截面的平面。

[0012] 图4C是图4A和图4B的示例性结构沿着图4A的水平平面C-C'的水平截面图。

[0013] 图4D是图4A-4C的示例性结构沿着图4A和图4C的垂直平面D-D'的垂直截面图。

[0014] 图5A-5H是根据本公开的实施例的直到第二半导体沟道层的沉积的工艺步骤的示例性结构内的存储器开口的顺序示意性垂直截面图。

[0015] 图6是根据本公开的实施例的示例性结构在形成存储器堆叠体结构和支撑柱结构之后的示意性垂直截面图。

[0016] 图7A是根据本公开的实施例的示例性结构在形成背侧沟槽之后的曲折示意性垂直截面图。

[0017] 图7B是图7A的示例性结构的部分透视俯视图。垂直平面A-A'是图7A的示意性垂直截面图的平面。

[0018] 图7C是图7A和图7B的示例性结构沿着图7A的水平平面C-C'的水平截面图。

[0019] 图8是根据本公开的实施例的示例性结构在形成背侧凹陷之后的曲折示意性垂直截面图。

[0020] 图9A-9D是根据本公开的实施例的示例性结构的区域在形成导电层期间的顺序垂直截面图。

[0021] 图10是示例性结构在图9D的工艺步骤处的曲折示意性垂直截面图。

[0022] 图11是根据本公开的实施例的示例性结构在从背侧沟槽内移除所沉积的导电材料之后的曲折示意性垂直截面图。

[0023] 图12A是根据本公开的实施例的示例性结构在形成绝缘间隔体和背侧接触结构之后的曲折示意性垂直截面图。

[0024] 图12B是图12A的示例性结构的区域的放大图。

[0025] 图13A是根据本公开的实施例的示例性结构在形成附加接触通孔结构之后的曲折示意性垂直截面图。

[0026] 图13B是图13A的示例性结构的俯视图。垂直平面A-A'是图13A的示意性垂直截面图的平面。

[0027] 图13C是图13A和图13B的示例性结构沿着图13A的水平平面C-C'的水平截面图。图

Micron Ex. 1002, p. 558
Micron v. YMTC
IPR2024-00789

13D是在形成附加接触通孔结构之后的示例性结构沿着图13E中的平面D-D'的另一曲折示意性垂直截面图。图13E是与图13C相同的视图。图示了曲折垂直截面平面D-D'。

[0028] 图14A是本公开的替代实施例的替代结构沿着图14B中所示的平面B-B'的曲折示意性垂直截面图。图14B是图14A的替代实施例结构的俯视图。垂直平面A-A'是图14A的示意性垂直截面图的平面。

[0029] 图15是现有技术结构的曲折示意性垂直截面图。

具体实施方式

[0030] 如上所述,本公开涉及包含多级存储器阵列的垂直的堆叠体的三维存储器器件以及其制造方法,以下描述了其各方面。本公开的实施例可以用于形成包含多级存储器结构的各种结构,其非限制性示例包含半导体器件,诸如包括多个NAND存储器串的三维单片存储器阵列器件。

[0031] 附图未按比例绘制。除非明确地描述或清楚地指示了不存在元件的重复,否则在示出了元件的单个实例的情况下,可以重复元件的多个实例。诸如"第一"、"第二"和"第三"的序数仅用于标识相似的元件,并且在本公开的说明书和权利要求书中可能采用不同的序数。相同的附图标记指代相同的元件或者相似的元件。除非另外指明,否则具有相同附图标记的元件假定为具有相同的组成。如本文所使用的,位于第二元件"上"的第一元件可以位于第二元件的表面的外侧上或者位于第二元件的内侧上。如本文所使用的,如果在第一元件的表面和第二元件的表面之间存在物理接触,则第一元件"直接"位于第二元件"上"。

[0032] 如本文所使用的,"层"是指包含具有厚度的区域的材料部分。层可以在下面的或上面的结构的整体之上延伸,或者可以具有比下面的或上面的结构的范围更小的范围。此外,层可以是均质或者非均质的连续结构的区域,层的厚度小于连续结构的厚度。例如,层可以位于任一对水平平面之间,水平平面在连续结构的顶表面和底表面之间或者在连续结构的顶表面和底表面处。层可以水平地、垂直地和/或沿着锥形表面延伸。基板可以是层,可以在其中包含一个或多个层,或者可以在其上、在其上方和/或在其下方具有一个或多个层。

[0033] 单片三维存储器阵列是其中在诸如半导体晶片的单个基板上方形成多个存储器级而没有介入基板的阵列。术语"单片"意味着阵列中的每级的层直接沉积在阵列中的每个下面的级的层上。相比之下,二维阵列可以分开形成,并且然后封装在一起以形成非单片存储器器件。例如,如题为"Three-dimensional Structure Memory"、专利号为5,915,167中的美国专利所描述的,已经通过在分开的基板上形成存储器级并且垂直地堆叠存储器级而构造了非单片堆叠存储器。在接合前可以将基板减薄或者从存储器级移除,但是由于存储器级最初形成在分开的基板上,因此这样的存储器不是真正的单片三维存储器阵列。本公开的各种三维存储器器件包含单片三维NAND串存储器器件,并且可以采用本文所描述的各种实施例来制造。

[0034] 参考图1,示出了根据本公开的实施例的示例性结构,其可以用于例如制造含有垂直NAND存储器器件的器件结构。示例性结构包含基板(9,10),其可以是半导体基板。基板可以包含基板半导体层9和可选的半导体材料层10。基板半导体层9可以是半导体晶片或半导体材料层,并且可以包含至少一种单质半导体材料(例如,单晶硅晶片或层)、至少一种III-

Micron Ex. 1002, p. 559
Micron v. YMTC
IPR2024-00789

V族化合物半导体材料、至少一种II-VI族化合物半导体材料、至少一种有机半导体材料、或者本领域中已知的其他半导体材料。基板可以具有主表面7，其可以是例如基板半导体层9的最顶部表面。主表面7可以是半导体表面。在一个实施例中，主表面7可以是单晶半导体表面，诸如单晶半导体表面。

[0035] 如本文所使用的，“半导材料”是指具有从1.0×10^{-6}S/cm到1.0×10^{5}S/cm的范围内的电导率的材料。如本文所使用的，“半导体材料”是指在其中没有电掺杂剂的情况下具有从1.0×10^{-6}S/cm到1.0×10^{5}S/cm的范围内的电导率且一经用电掺杂剂适当地掺杂能够产生具有从1.0S/cm到1.0×10^{5}S/cm的范围内的电导率的掺杂材料的材料。如本文所使用的，“电掺杂剂”是指将空穴添加到能带结构内的价带的p型掺杂剂，或者是将电子添加到能带结构内的导带的n型掺杂剂。如本文所使用的，“导电材料”是指具有大于1.0×10^{5}S/cm的电导率的材料。如本文所使用的，“绝缘体材料”或“电介质材料”是指具有小于1.0×10^{-6}S/cm的电导率的材料。如本文所使用的，“重掺杂半导体材料”是指掺杂有足够高原子浓度的电掺杂剂而变为导电材料(即具有大于1.0×10^{5}S/cm的电导率)的半导体材料。“掺杂半导体材料”可以是重掺杂半导体材料，或者可以是包含提供从1.0×10^{-6}S/cm到1.0×10^{5}S/cm的范围内的电导率的浓度的电掺杂剂(即，p型掺杂剂和/或n型掺杂剂)的半导体材料。“本征半导体材料”是指未掺杂有电掺杂剂的半导体材料。因此，半导体材料可以是半导的或者导电的，并且可以是本征半导体材料或者掺杂半导体材料。根据在其中的电掺杂剂的原子浓度，掺杂半导体材料可以是半导的或者导电的。如本文所使用的，“金属材料”是指其中包含至少一种金属元素的导电材料。电导率的所有测量是在标准条件下进行的。

[0036] 用于外围电路的至少一个半导体器件700(未示出)可以形成在半导体材料层9的部分上。至少一个半导体器件可以包含例如场效应晶体管。例如，可以通过蚀刻基板半导体层9的部分并在其中沉积电介质材料来形成至少一个浅沟槽隔离结构120。栅极电介质层、至少一个栅极导体层和栅极帽电介质层可以被形成在基板半导体层9之上，并且可以随后被图案化，以形成至少一个栅极结构(150，152，154，158)，其中的每一个可以包含栅极电介质150、栅电极(152，154)和栅极帽电介质158。栅电极(152，154)可以包含第一栅电极部分152和第二栅电极部分154的堆叠体。可以通过沉积和各向异性地蚀刻电介质衬垫来在至少一个栅极结构(150，152，154，158)周围形成至少一个栅极间隔体156。可以例如通过采用至少一个栅极结构(150，152，154，158)作为掩模结构来引入电掺杂剂，来将有源区130形成在基板半导体层9的上部部分中。可以按需采用附加的掩膜。有源区130可以包含场效应晶体管的源极区域和漏极区域。可以可选地形成第一电介质衬垫161和第二电介质衬垫162。第一和第二电介质衬垫(161，162)中的每一个可以包含硅氧化物层、硅氮化物层、和/或电介质金属氧化物层。如本文所使用的，硅氧化物包含二氧化硅以及非化学计量的硅氧化物，非化学计量的硅氧化物对每个硅原子具有多于或少于两个的氧原子。二氧化硅是优选的。在说明性示例中，第一电介质衬垫161可以是硅氧化物层，并且第二电介质衬垫162可以是硅氮化物层。外围电路的至少一个半导体器件可以含有将随后形成的存储器器件的驱动器电路，存储器器件可以包含至少一个NAND器件。

[0037] 诸如硅氧化物的电介质材料可以被沉积在至少一个半导体器件之上，并且可以随后被平坦化以形成平坦化电介质层170。在一个实施例中，平坦化电介质层170的平坦化的顶表面可以与电介质衬垫(161，162)的顶表面共面。随后，平坦化电介质层170和电介质衬

Micron Ex. 1002, p. 560
Micron v. YMTC
IPR2024-00789

垫(161,162)可以被从区域移除,以物理暴露基板半导体层9的顶表面。如本文所使用的,如果表面与真空或者气相材料(诸如空气)物理接触,则表面是“物理暴露”的。

[0038] 如果可选的半导体材料层10存在,则在形成至少一个半导体器件700之前或者之后,可以通过由选择性外延来沉积单晶半导体材料来而在基板半导体层9的顶表面上形成可选的半导体材料层10。沉积的半导体材料可以相同于或者可以不同于基板半导体层9的半导体材料。基板半导体层9如上所述。半导体材料层10的单晶半导体材料可以与基板半导体层9的单晶结构外延对准。可以例如通过化学机械平坦化(CMP)来移除沉积的半导体材料的位于平坦化电介质层170的顶表面的上方的部分。在这种情况下,半导体材料层10可以具有与平坦化电介质层170的顶表面共面的顶表面。

[0039] 至少一个半导体器件700的区域(即,区)在本文中称为外围器件区域200。其中随后形成存储器阵列的区域在本文中称为存储器阵列区域100。可以在存储器阵列区域100和外围器件区域200之间提供用于随后形成导电层的阶梯式梯台的接触区域300。可选地,栅极电介质层12可以形成在半导体材料层10和平坦化电介质层170的上方。栅极电介质层12可以是例如硅氧化物层。栅极电介质层12的厚度可以在从3nm到30nm的范围内,虽然还可以采用更小和更大的厚度。

[0040] 参考图2,交替的多个第一材料层(其可以是绝缘层32)和第二材料层(其可以是牺牲材料层42)的堆叠体形成在基板的顶表面之上,其可以是例如在栅极电介质层12的顶表面上。如本文所使用的,“材料层”是指在其整体通体包含材料的层。如本文所使用的,交替的多个第一元件和第二元件是指第一元件的实例和第二元件的实例在其中交替的结构。第一元件的不是交替的多元件的端部元件的每个实例在两侧由第二元件的两个实例邻接,且第二元件不是交替的多元件的端部元件的每个实例在两端由第一元件的两个实例邻接。第一元件在其间可以具有相同的厚度,或者可以具有不同的厚度。第二元件在其间可以具有相同的厚度,或者可以具有不同的厚度。交替的多个第一材料层和第二材料层可以以第一材料层的实例或者以第二材料层的实例开始,并且可以以第一材料层的实例或者以第二材料层的实例终止。在一个实施例中,第一元件的实例和第二元件的实例可以形成在交替的多元件内周期性地重复的单元。

[0041] 每个第一材料层包含第一材料,并且每个第二材料层包含不同于第一材料的第二材料。在一个实施例中,每个第一材料层可以是绝缘层32,并且每个第二材料层可以是牺牲材料层。在这种情况下,堆叠体可以包含交替的多个绝缘层32和牺牲材料层42,并且构成包括绝缘层32和牺牲材料层42的交替层的原型堆叠体。如本文所使用的,“原型”结构或者“处理中”的结构是指瞬态结构,该瞬态结构随后在其中至少一个组件的形状或成分上被修改。

[0042] 交替的多元件的堆叠体在本文中称为交替堆叠体(32,42)。在一个实施例中,交替堆叠体(32,42)可以包含由第一材料构成的绝缘层32以及由不同于绝缘层32的材料的第二材料构成的牺牲材料层42。绝缘层32的第一材料可以是至少一种绝缘材料。就此而言,每个绝缘层32可以是绝缘材料层。可以用于绝缘层32的绝缘材料包含但不限于硅氧化物(包含掺杂和未掺杂的硅酸盐玻璃)、硅氮化物、硅氮氧化物、有机硅酸盐玻璃(OSG)、旋涂电介质材料、通常称为高介电常数(高k)电介质氧化物(例如铝氧化物、铪氧化物等)的电介质金属氧化物及其硅酸盐、电介质金属氮氧化物及其硅酸盐,以及有机绝缘材料。在一个实施例中,绝缘层32的第一材料可以是硅氧化物。

Micron Ex. 1002, p. 561
Micron v. YMTC
IPR2024-00789

[0043] 牺牲材料层42的第二材料是可以对绝缘层32的第一材料有选择性地被移除的牺牲材料。如本文所使用的，如果移除工艺移除第一材料的速率至少两倍于第二材料的移除的速率，则第一材料的移除是“对”第二材料“有选择性”的。第一材料的移除的速率与第二材料的移除的速率的比率在本文中称为第一材料的移除工艺对于第二材料的“选择度”。

[0044] 牺牲材料层42可以包括绝缘材料、半导体材料或者导电材料。牺牲材料层42的第二材料可以随后被用导电电极取代，导电电极可以例如起到垂直NAND器件的控制栅电极的功能。第二材料的非限制性示例包含硅氮化物、非晶半导体材料(诸如非晶硅)和多晶半导体材料(诸如多晶硅)。在一个实施例中，牺牲材料层42可以是包括硅氮化物的或者包含硅和锗中的至少一个的半导体材料的间隔体材料层。

[0045] 在一个实施例中，绝缘层32可以包含硅氧化物，并且牺牲材料层可以包含硅氮化物牺牲材料层。可以例如由化学气相沉积(CVD)来沉积绝缘层32的第一材料。例如，如果硅氧化物用于绝缘层32，则原硅酸四乙酯(tetraethyl orthosilicate，TEOS)可以用作CVD工艺的前驱体材料。可以例如由CVD或者原子层沉积(ALD)来形成牺牲材料层42的第二材料。

[0046] 可以适当地图案化牺牲材料层42，使得通过牺牲材料层42的取代随后形成的导电材料部分可以起到导电电极(诸如随后形成的单片三维NAND串存储器器件的控制栅电极)的功能。牺牲材料层42可以包括具有实质上平行于基板的主表面7延伸的条形的部分。

[0047] 绝缘层32和牺牲材料层42的厚度可以在从20nm到50nm的范围内，虽然也可以对每个绝缘层32和对每个牺牲材料层42采用更小或更大的厚度。绝缘层32和牺牲材料层(例如，控制栅电极或者牺牲材料层)42的对的重复的数目可以在从2到1024的范围内，并且典型地是从8到256，虽然还可以采用更大数目的重复。堆叠体中顶部的和底部栅电极可以起到选择栅电极的功能。在一个实施例中，交替堆叠体(32，42)中的每个牺牲材料层42可以具有在每个相应的牺牲材料层42内实质上不变的均匀的厚度。

[0048] 尽管采用其中间隔体材料层是随后被导电层取代的牺牲材料层42的实施例来描述本公开，但是本文明确地涵盖了其中牺牲材料层形成为导电层的实施例。在这种情况下，用导电层取代间隔体材料层的步骤可以省略。

[0049] 可选地，绝缘帽层70可以形成在交替堆叠体(32，42)之上。绝缘帽层70包含不同于牺牲材料层42的材料的电介质材料。在一个实施例中，绝缘帽层70可以包含可以用于如上所述的绝缘层32的电介质材料。绝缘帽层70可以具有比绝缘层32中的每一个更大的厚度。绝缘帽层70可以例如由化学气相沉积来沉积。在一个实施例中，绝缘帽层70可以是硅氧化物层。

[0050] 参考图3，阶梯式腔可以形成在接触区域300内，接触区域300位于存储器阵列区域100和含有外围电路的至少一个半导体器件的外围器件区域200之间。阶梯式腔可以具有各种阶梯式表面，使得阶梯式腔的水平截面形状作为距基板10的顶表面的垂直距离的函数逐步改变。在一个实施例中，可以通过重复进行工艺步骤的集合来形成阶梯式腔。工艺步骤的集合可以包含，例如第一类型的蚀刻工艺和第二类型的蚀刻工艺，第一类型的蚀刻工艺将腔的深度垂直地增加一个或多个级，第二类型的蚀刻工艺横向地扩大在随后的第一类型的蚀刻工艺中将垂直地蚀刻的面积。如本文所使用的，包含交替的多元件的结构的“级”定义为结构内的第一材料层和第二材料层的对的相对位置。

[0051] 在形成阶梯式腔之后，交替堆叠体(32，42)的外围部分在形成阶梯式腔之后可以

Micron Ex. 1002, p. 562
Micron v. YMTC
IPR2024-00789

具有阶梯式表面。如本文所使用的,“阶梯式表面”是指包含至少两个水平表面和至少两个垂直表面的表面的集合,使得每个水平表面邻接于从水平表面的第一边缘向上延伸的第一垂直表面,并且邻接于从水平表面的第二边缘向下延伸的第二垂直表面。“阶梯式腔”是指具有阶梯式表面的腔。

[0052] 梯台区域通过图案化交替堆叠体(32,42)形成。交替堆叠体(32,42)内除了最顶部的牺牲材料层42之外的每个牺牲材料层42比交替堆叠体(32,42)内的上面的(多个)牺牲材料层42横向地延伸得更远。梯台区域包含交替堆叠体(32,42)的阶梯式表面,阶梯式表面从交替堆叠体(32,42)内的最底部层连续地延伸到交替堆叠体(32,42)内的最顶部层。

[0053] 可以通过在阶梯式腔中沉积电介质材料而在阶梯式腔中形成倒退阶梯式电介质材料部分65(即绝缘填充材料部分)。例如,诸如硅氧化物的电介质材料可以沉积在阶梯式腔中。可以由化学机械平坦化(CMP)来从绝缘帽层70的顶表面上方移除所沉积的电介质材料的超出部分。填充阶梯式腔的所沉积的电介质材料的剩余部分构成倒退阶梯式电介质材料部分65。如本文所使用的,“倒退阶梯式”元件是指具有阶梯式表面和作为距元件所在的基板的顶表面的垂直距离的函数而单调地增加的水平截面积的元件。如果硅氧化物用于倒退阶梯式电介质材料部分65,则倒退阶梯式电介质材料部分65的硅氧化物可以掺杂有或者可以不掺杂有诸如B、P和/或F的掺杂剂。

[0054] 可选地,可以穿过绝缘帽层70和牺牲材料层42的位于漏极选择级处的子集来形成漏极选择级隔离结构72。可以通过形成漏极选择级隔离沟槽并且用诸如硅氧化物的电介质材料填充漏极选择级隔离沟槽来形成漏极选择级隔离结构72。可以从绝缘帽层70的顶表面的上方移除电介质材料的超出部分。

[0055] 参考图4A-4D,至少包含光刻胶层的光刻材料堆叠体(未示出)可以被形成在绝缘帽层70和倒退阶梯式电介质材料部分65之上,并且可以被光刻地图案化以在其中形成开口。开口包含形成在存储器阵列区域100之上的开口的第一子集和形成在接触区域300之上的开口的第二子集。可以由将图案化的光刻材料堆叠体用作蚀刻掩模的至少一个各向异性蚀刻,来穿过绝缘帽层70或倒退阶梯式电介质材料部分65并穿过交替堆叠体(32,42)来转移光刻材料堆叠体中的图案。刻蚀交替堆叠体(32,42)的在图案化的光刻材料堆叠体的开口下面的部分,以形成存储器开口49和支撑开口19。如本文所使用的,“存储器开口”是指在其中随后形成诸如存储器堆叠体结构的存储器元件的结构。如本文所使用的,“支撑开口”是指在其中随后形成机械支撑其他元件的支撑结构(诸如支撑柱结构)的结构。穿过存储器阵列区域100中的绝缘帽层70和交替堆叠体(32,42)的整体来形成存储器开口49。穿过接触区域300中的倒退阶梯式电介质材料部分65和交替堆叠体(32,42)的在阶梯式表面下面的部分来形成支撑开口19。

[0056] 存储器开口49在存储器阵列区域100中延伸穿过交替堆叠体(32,42)的每层。支撑开口19延伸穿过交替堆叠体(32,42)内的层的子集。可以在梯台区域中穿过倒退阶梯式电介质材料部分65,形成支撑开口19的主要(predominant)子集(多于50%)或全部。用于蚀刻穿过交替堆叠体(32,42)的材料的各项异性蚀刻工艺的化学过程可以交替,以最优化交替堆叠体(32,42)中的第一和第二材料的蚀刻。各向异性蚀刻可以是例如一系列反应离子蚀刻。存储器开口49和支撑开口19的侧壁可以实质上是垂直的,或者可以是锥形的。可以例如通过灰化,随后移除图案化的光刻材料堆叠体。

Micron Ex. 1002, p. 563
Micron v. YMTC
IPR2024-00789

[0057] 可以穿过栅极电介质层12来形成存储器开口49和支撑开口19，使得存储器开口49和支撑开口19从交替堆叠体(32,42)的顶表面至少延伸到包含半导体材料层10的最顶部表面的水平平面。在一个实施例中，在半导体材料层10的顶表面物理暴露在每个存储器开口49和每个支撑开口19的底部处之后，可以可选地进行到半导体材料层10中的过蚀刻。过蚀刻可以在移除光刻材料堆叠体之前或者之后进行。换言之，半导体材料层10的凹陷的表面可以从半导体材料层10的未凹陷的顶表面垂直地偏移凹陷深度。凹陷深度可以例如在从1nm到50nm的范围内，虽然还可以采用更小和更大的凹陷深度。过蚀刻是可选的，并且可以省略。如果不进行过蚀刻，则存储器开口49和支撑开口19的底表面可以与半导体材料层10的最顶部表面共面。存储器开口49和支撑开口19中的每一个可以包含实质上垂直于基板的最顶部表面延伸的侧壁(或多个侧壁)。存储器开口49的二维阵列可以形成在存储器阵列区域100中。支撑开口19的二维阵列可以形成在接触区域300中。基板半导体层9和半导体材料层10共同地构成基板(9,10)，其可以是半导体基板。替代地，半导体材料层10可以省略，并且存储器开口49和支撑开口19可以延伸到基板半导体层9的顶表面。

[0058] 交替堆叠体(32,42)的阶梯式表面包含邻接于水平表面(即，阶梯的顶部)H的垂直阶梯S。每个垂直阶梯S的顶部边缘(即，阶梯的侧面)可以连接到上水平表面的边缘(即，阶梯的顶部)，并且每个垂直阶梯S的底边缘可以连接到下水平表面的边缘(即，下面的阶梯的顶部或台阶(staircase)的底表面)。在一个示例中，垂直阶梯S可以垂直于第一水平方向hd1(例如，字线方向)，交替堆叠体(32,42)的层的横向范围沿着第一水平方向hd1变化。在一个示例中，每个阶梯可以包含一个绝缘层32和一个牺牲材料层42。

[0059] 在另一可选示例中，垂直阶梯S的一个子集可以垂直于第一水平方向hd1，并且垂直阶梯S的另一子集可以垂直于第二水平方向hd2(例如，位线方向)，第二水平方向hd2垂直于第一水平方向hd1。在此实施例中，可以将第一阶梯式表面提供为与第二阶梯式表面相邻，第一阶梯式表面在每个垂直阶梯S处沿着第一水平方向改变两个绝缘层32和两个牺牲材料层42的高度，第二阶梯式表面沿着第一水平方向hd1改变两个绝缘层32和两个牺牲材料层42的高度。第一阶梯式表面的水平表面的级可以从第二阶梯式表面的水平表面的级垂直偏移一对绝缘层32和牺牲材料层42。

[0060] 在图4C和图4D中所示的示例中，每个垂直阶梯S与两个在第二水平方向hd2上相邻的垂直阶梯相差两层(例如，一个绝缘层32和一个牺牲材料层42)。这形成两个阶梯在第二水平方向上偏移两层的重复图案，如图4D所示。在图4C中所示的此示例中，水平表面(即，阶梯的顶部)H2在两层(例如，一个绝缘层32和一个牺牲材料层42)上面，水平表面(即，阶梯的顶部)H4各自在四层(例如，两个绝缘层32和两个牺牲材料层42)上面，水平表面H6各自在六层上面，水平表面H8各自在八层上面，水平表面H10各自在十层上面，以此类推。在此示例中，每个垂直阶梯S与两个在第一水平方向hd1上相邻的垂直阶梯相差四层(例如，两个绝缘层32和两个牺牲材料层42)。然而，还可以采用其他阶梯配置。

[0061] 总体上，可以沿着第二水平方向hd2提供N(其中N是大于零的整数)个阶梯式表面，使得N个阶梯式表面中的每一个在垂直于第一水平方向hd1的每个垂直阶梯S处沿着第一水平方向hd1改变N个绝缘层32和N个牺牲材料层42的高度。N个阶梯式表面的汇总为牺牲材料层42的每级提供水平表面H。例如，如果N＝3，则在第二水平方向hd2上可以存在三个垂直阶梯S的重复图案，其中相邻垂直阶梯S之间的每个阶梯高度在第一水平方向hd1上相差三个

Micron Ex. 1002, p. 564
Micron v. YMTC
IPR2024-00789

绝缘层32和三个牺牲材料层42(总共六层)。

[0062] 最底部牺牲材料层42对应于源极选择栅极级。在一个示例中,最底部牺牲材料层42的水平顶表面的接触倒退阶梯式电介质材料部分65的接触区域可以大于除最底部牺牲材料层42之外的牺牲材料层42与倒退阶梯式电介质材料部分65之间的任何接触区域。

[0063] 在一个示例中,最底部牺牲材料层42的水平顶表面的接触区域的区域可以不含穿过其的任何支撑开口19或存储器开口49。在一个示例中,最底部牺牲材料层42的水平顶表面的不含穿过其的任何支撑开口49或存储器开口49的接触区域的区域可以包含与垂直于第一水平方向的垂直阶梯S相接(border)且与两个垂直阶梯S的平行于第一水平方向hd1的区段相接的区域。在此情况下,最底部牺牲材料层42的水平顶表面的不含穿过其的任何支撑开口19的接触区域的区域可以包含最底部牺牲材料层42的水平顶表面的接触倒退阶梯式电介质材料部分65的整个接触区域的多于50%。

[0064] 图5A-5H示出了存储器开口49中的结构改变,其是图4A和图4B的示例性结构中的存储器开口49中的一个。在其他存储器开口49中的每一个中并在每个支撑开口19中同时发生相同的结构改变。

[0065] 参考图5A,示出了图4A和图4B的示例性器件结构中的存储器开口49。存储器开口49延伸穿过绝缘帽层70、交替堆叠体(32,42)、栅极电介质层12,并且可选地延伸到半导体材料层10的上部部分中。在该工艺步骤中,每个支撑开口19可以延伸穿过倒退阶梯式电介质材料部分65、交替堆叠体(32,42)中的层的子集、栅极电介质层12,并且可选地穿过半导体材料层10的上部部分。每个存储器开口的底表面相对于半导体材料层10的顶表面的凹陷深度可以在从0nm到30nm的范围内,虽然还可以采用更大的凹陷深度。可选地,可以例如通过各向同性蚀刻来使牺牲材料层42部分地横向凹陷,以形成横向凹陷(未示出)。

[0066] 参考图5B,可以例如通过选择性半导体沉积工艺来在每个存储器开口49和每个支撑开口19的底部部分处形成可选的基座沟道部分(例如,外延基座)11。选择性半导体沉积工艺从诸如基板(9,10)中的半导体材料层10的表面的半导体表面生长半导体材料,而不从电介质表面生长半导体材料。如果半导体材料层10是单晶的,则基座沟道部分11可以包含单晶半导体材料。替代地,基座沟道部分11可以包含多晶半导体材料。

[0067] 在半导体材料层10是单晶的且采用选择性外延工艺作为选择性半导体沉积工艺的情况下,每个基座沟道部分11可以包括与半导体材料层10的单晶半导体材料外延对准的单晶半导体材料。在一个示例中,基座沟道部分11可以掺杂有与半导体材料层10相同导电型的电掺杂剂。

[0068] 在一个示例中,每个基座沟道部分11的顶表面可以形成在包含最底部牺牲材料层42的顶表面的水平平面上方,并且在包含从底部第二个牺牲材料层42(即,最接近于最底部牺牲材料层42的牺牲材料层42)的底表面的水平平面之下。在此情况下,可以随后通过用导电材料层取代位于包含基座沟道部分11的顶表面的水平平面之下的牺牲材料层42来形成源极选择栅电极,如下面将关于图13A-13E描述的。在另一实施例中,可以用导电层取代若干下部牺牲材料层,以形成若干源极选择栅电极,如下面将关于图14A-14B描述的。支撑开口19的主要(predominant)子集穿过多于两个牺牲材料层42。在这样的支撑开口19中形成的基座沟道部分11不物理接触倒退阶梯式电介质材料部分。一些支撑开口19穿过最底部牺牲材料层42,并且不穿过任何其他牺牲材料层42。在这样的支撑开口19中形成的基座沟道

Micron Ex. 1002, p. 565
Micron v. YMTC
IPR2024-00789

部分11物理接触倒退阶梯式电介质材料部分65的侧壁。

[0069] 基座沟道部分11可以是晶体管沟道的部分，该晶体管沟道的部分在将随后在基板(9,10)中形成的源极区域和将随后在存储器开口49的上部部分中后续形成的漏极区域之间延伸。存储器腔49'存在于存储器开口49的在基座沟道部分11上方的未填充部分中。在一个实施例中，基座沟道部分11可以包括单晶硅。在一个实施例中，基座沟道部分11可以具有第一导电型的掺杂，其与基座沟道部分所接触的半导体材料层10的导电型相同。如果半导体材料层10不存在，则基座沟道部分11可以直接形成在基板半导体层9上，基板半导体层9可以具有第一导电型的掺杂。

[0070] 参考图5C，包含阻挡电介质层52、电荷储存层54、隧穿电介质层56，以及可选的第一半导体沟道层601的层的堆叠体可以随后沉积在存储器开口49中。阻挡电介质层52可以包含单个电介质材料层或者多个电介质材料层的堆叠体。在一个实施例中，阻挡电介质层可以包含基本由电介质金属氧化物构成的电介质金属氧化物层。如本文所使用的，电介质金属氧化物是指包含至少一种金属元素和至少氧的电介质材料。电介质金属氧化物可以基本由至少一种金属元素和氧构成，或者可以基本由至少一种金属元素、氧和诸如氮的至少一种非金属元素构成。在一个实施例中，阻挡电介质层52可以包含具有大于7.9的介电常数(即具有大于硅氮化物的介电常数的介电常数)的电介质金属氧化物。

[0071] 电介质金属氧化物的非限制性示例包含铝氧化物(Al_2O_3)、铪氧化物(HfO_2)、镧氧化物(LaO_2)、钇氧化物(Y_2O_3)、钽氧化物(Ta_2O_5)、其硅酸盐、其氮掺杂化合物、其合金及其堆叠体。可以通过例如学气相沉积(CVD)、原子层沉积(ALD)、脉冲激光沉积(PLD)、液体源雾化化学沉积或其组合来沉积电介质金属氧化物层。电介质金属氧化物层的厚度可以在从1nm到20nm的范围内，虽然还可以采用更小和更大的厚度。电介质金属氧化物层可以随后起到电介质材料部分的功能，电介质材料部分阻挡储存的电荷泄露到控制栅电极。在一个实施例中，阻挡电介质层52包含铝氧化物。在一个实施例中，阻挡电介质层52可以包含具有不同材料成分的多个电介质金属氧化物层。

[0072] 替代地或者附加地，阻挡电介质层52可以包含电介质半导体化合物，诸如硅氧化物、硅氧氮化物、硅氮化物或其组合。在一个实施例中，阻挡电介质层52可以包含硅氧化物。在这种情况下，可以通过保形沉积方法(诸如低压化学气相沉积、原子层沉积或者其组合)来形成阻挡电介质层52的电介质半导体化合物。电介质半导体化合物的厚度可以在从1nm到20nm的范围内，虽然还可以采用更小和更大的厚度。替代地，可以省略阻挡电介质层52，并且可以在要随后形成的存储器膜的表面上形成背侧凹陷之后形成背侧阻挡电介质层。

[0073] 随后，可以形成电荷储存层54。在一个实施例中，电荷储存层54可以是电荷捕获材料的连续层或者图案化的离散部分，电荷捕获材料包含可以是例如硅氮化物的电介质电荷捕获材料。替代地，电荷储存层54可以包含诸如通过例如在横向凹陷内被形成到牺牲材料层42中而被图案化为多个电隔离的部分(例如，浮置栅极)的掺杂多晶硅或金属材料的导电材料的连续层或者图案化的离散部分。在一个实施例中，电荷储存层54包含硅氮化物层。在一个实施例中，牺牲材料层42和绝缘层32可以具有垂直重合的侧壁，并且电荷储存层54可以形成为单个连续层。

[0074] 在另一个实施例中，牺牲材料层42可以相对于绝缘层32的侧壁横向地凹陷，并且可以采用沉积工艺和各项异性蚀刻工艺的组合来将电荷储存层54形成为垂直地间隔开的

Micron Ex. 1002, p. 566
Micron v. YMTC
IPR2024-00789

多个存储器材料部分。尽管采用其中电荷储存层54是单个连续层的实施例来描述本公开，但是本文明确涵盖其中用垂直地间隔开的多个存储器材料部分(其可以是电荷捕获材料部分或者电隔离的导电材料部分)取代电荷储存层54的实施例。

[0075] 电荷储存层54可以形成为均质成分的单个电荷储存层，或者可以包含多个电荷储存层的堆叠体。多个电荷储存层(如果采用)可以包括多个间隔开的浮置栅极材料层，浮置栅极材料层含有导电材料(例如，诸如钨、钼、钽、钛、铂、钌及其合金的金属，或诸如钨硅化物、钼硅化物、钽硅化物、钛硅化物、镍硅化物、钴硅化物或其组合的金属硅化物)和/或半导体材料(例如，包含至少一种单质半导体元素或至少一种化合物半导体材料的多晶或非晶半导体材料)。替代地或附加地，电荷储存层54可以包括绝缘电荷捕获材料，诸如一个或多个硅氮化物段。替代地，电荷储存层54可以包括导电纳米颗粒(诸如金属纳米颗粒)，其可以是例如钌纳米颗粒。可以例如由化学气相沉积(CVD)、原子层沉积(ALD)、物理气相沉积(PVD)、或用于在其中储存电荷的任何适当的沉积技术来形成电荷储存层54。电荷储存层54的厚度可以在从2nm到30nm的范围内，虽然还可以采用更小和更大的厚度。

[0076] 隧穿电介质层56包含电介质材料，可以在适当的电偏置条件下穿过该电介质材料进行电荷隧穿。可以根据要形成的单片三维NAND串存储器器件的操作的模式，通过热载流子注入或者由福勒-诺得海姆(Fowler-Nordheim)隧穿诱导的电荷转移来进行电荷隧穿。隧穿电介质层56可以包含硅氧化物、硅氮化物、硅氮氧化物、电介质金属氧化物(诸如铝氧化物和铪氧化物)、电介质金属氮氧化物、电介质金属硅酸盐、其合金、和/或其组合。在一个实施例中，隧穿电介质层56可以包含第一硅氧化物层、硅氮氧化物层和第二硅氧化物层的堆叠体，其通常已知为ONO堆叠体。在一个实施例中，隧穿电介质层56可以包含实质上不含碳的硅氧化物层或者实质上不含碳的硅氮氧化物层。隧穿电介质层56的厚度可以在从2nm到20nm的范围内，虽然还可以采用更小和更大的厚度

[0077] 可选的第一半导体沟道层601包含半导体材料，诸如至少一种单质半导体材料、至少一种III-V族化合物半导体材料、至少一种II-VI族化合物半导体材料、至少一种有机半导体材料或者在本领域已知的其他半导体材料。在一个实施例中，第一半导体沟道层601包含非晶硅或者多晶硅。可以通过保形沉积方法(诸如低压化学气相沉积(LPCVD))形成第一半导体沟道层601。第一半导体沟道层601的厚度可以在从2nm到10nm的范围内，虽然还可以采用更小和更大的厚度。存储器腔49’形成在每个存储器开口49的未填充沉积的材料层(52,54,56,601)的体积中。

[0078] 参考图5D，采用至少一种各向异性蚀刻工艺，顺序地各向异性地蚀刻可选的第一半导体沟道层601、隧穿电介质层56、电荷储存层54、阻挡电介质层52。可以通过至少一个各向异性蚀刻工艺移除第一半导体沟道层601、隧穿电介质层56、电荷储存层54以及阻挡电介质层52的位于绝缘帽层70的顶表面的上方的部分。另外，可以移除第一半导体沟道层601、隧穿电介质层56、电荷储存层54以及阻挡电介质层52的在每个存储器腔49’的底部处的水平部分，以在其剩余部分中形成开口。可以通过采用相应的蚀刻化学过程的相应的各项异性蚀刻工艺来蚀刻第一半导体沟道层601、隧穿电介质层56、电荷储存层54以及阻挡电介质层52中的每一个，蚀刻化学对于各种材料层可以相同或者可以不相同。

[0079] 第一半导体沟道层601的每个剩余部分可以具有管状配置。电荷储存层54可以包括电荷捕获材料或者浮置栅极材料。在一个实施例中，每个电荷储存层54可以包含一经编

Micron Ex. 1002, p. 567
Micron v. YMTC
IPR2024-00789

程即储存电荷的电荷储存区域的垂直堆叠体。在一个实施例中，电荷储存层54可以是其中相邻于牺牲材料层42的每个部分构成电荷储存区域的电荷储存层。

[0080] 可以穿过第一半导体沟道层601、隧穿电介质层56、电荷储存层54和阻挡电介质层52来在开口的下面物理暴露基座沟道部分11的表面(或者在不采用基座沟道部分11的情况下的半导体材料层10的表面)。可选地，在每个存储器腔49'的底部处物理暴露的半导体表面可以垂直地凹陷，使得在存储器腔49'的下面凹陷的半导体表面从基座沟道部分11的(或者在不采用基座沟道部分11的情况下的半导体材料层10的)最顶部的表面垂直地偏移凹陷距离。隧穿电介质层56位于电荷储存层54之上。存储器开口49中的阻挡电介质层52、电荷储存层54和隧穿电介质层56的集合构成存储器膜50，其包含多个电荷储存区域(如被实现为电荷储存层54)，该多个电荷储存区域由阻挡电介质层52和隧穿电介质层56与围绕的材料相绝缘。在一个实施例中，第一半导体沟道层601、隧穿电介质层56、电荷储存层54和阻挡电介质层52可以具有垂直重合的侧壁。

[0081] 参考图5E，可以直接在基座沟道部分11的半导体表面上(或者如果省略基座沟道部分11，则直接在半导体材料层10的半导体表面上)，并且直接在第一半导体沟道层601上沉积第二半导体沟道层602。第二半导体沟道层602包含半导体材料，诸如至少一种单质半导体材料、至少一种III-V族化合物半导体材料、至少一种II-VI族化合物半导体材料、至少一种有机半导体材料或者本领域中已知的其他半导体材料。在一个实施例中，第二半导体沟道层602包含非晶硅或者多晶硅。可以通过保形沉积方法(诸如低压化学气相沉积(LPCVD))形成第二半导体沟道层602。第二半导体沟道层602的厚度可以在从2nm到10nm的范围内，虽然还可以采用更小和更大的厚度。第二半导体沟道层602可以部分填充每个存储器开口中的存储器腔49'，或者可以完全填充每个存储器开口中的腔。第一半导体沟道层601和第二半导体沟道层602的材料全体地称为半导体沟道材料。换言之，半导体沟道材料是在第一半导体沟道层601和第二半导体沟道层602中的所有半导体材料的集合。

[0082] 参考图5F，在每个存储器开口中的存储器腔49'未被第二半导体沟道层602完全填充的情况下，可以在存储器腔49'中沉积电介质芯层62L，以填充每个存储器开口内的存储器腔49'的任何剩余部分。电介质芯层62L包含诸如硅氧化物或者有机硅酸盐玻璃的电介质材料。可以由保形沉积方法(诸如低压化学气相沉积(LPCVD))或者由自平坦化沉积工艺(诸如旋涂)来沉积电介质芯层62L。

[0083] 参考图5G，可以例如通过凹陷蚀刻从绝缘帽层70的顶表面的上方移除电介质芯层62L的水平部分。电介质芯层62L的每个剩余部分构成电介质芯62。另外，第二半导体沟道层602的位于绝缘帽层70的顶表面的上方的水平部分可以通过平坦化工艺移除，平坦化工艺可以采用凹陷蚀刻或者化学机械平坦化(CMP)。第二半导体沟道层602的每个剩余部分可以整个位于存储器开口49内或者整个位于支撑开口19内。

[0084] 第一半导体沟道层601和第二半导体沟道层602的每个接邻的对可以全体地形成垂直半导体沟道60，当包含垂直半导体沟道60的垂直NAND器件导通时，电流可以流动穿过该垂直半导体沟道60。隧穿电介质层56被电荷储存层54围绕，并且横向地围绕垂直半导体沟道60的部分。阻挡电介质层52、电荷储存层54和隧穿电介质层56的每个相邻的集合全体地构成存储器膜50，其能够以宏观保留时间储存电荷。在一些实施例中，在该步骤处，阻挡电介质层52可以不存在于存储器膜50中，并且可以在形成背侧凹陷之后随后形成阻挡电介

Micron Ex. 1002, p. 568
Micron v. YMTC
IPR2024-00789

质层。如本文所使用的，宏观保留时间是指适合于永久性存储器器件的存储器器件的操作的保留时间，诸如超过24小时的保留时间。

[0085] 参考图5H，例如通过凹陷蚀刻到位于绝缘帽层70的顶表面和绝缘帽层70的底表面之间的深度，使每个电介质芯62的顶表面进一步在每个存储器开口内凹陷。可以通过在电介质芯62的上方在每个凹陷区域内沉积掺杂半导体材料来形成漏极区域63。漏极区域63可以具有与第一导导电型相反的第二导电型的掺杂。例如，如果第一导电型是p型，则第二导电型是n型，反之亦然。漏极区域63中的掺杂剂浓度可以在从$5.0\times10^{19}/cm^3$到$2.0\times10^{21}/cm^3$的范围内，虽然还可以采用更小和更大的掺杂剂浓度。掺杂半导体材料可以是例如掺杂多晶硅。可以例如由化学机械平坦化(CMP)或者凹陷蚀刻来从绝缘帽层70的顶表面的上方移除沉积的半导体材料的超出部分，以形成漏极区域63。

[0086] 存储器开口49内的存储器膜50和垂直半导体沟道60(其是垂直半导体沟道)的每个组合构成存储器堆叠体结构55。存储器堆叠体结构55是半导体沟道、隧穿电介质层、多个存储器元件(如实现为电荷储存层54的部分)和可选的阻挡电介质层52的组合。存储器开口49内的基座沟道部分11(如果存在)、存储器堆叠体结构55、电介质芯62和漏极区域63的每个组合在本文中称为存储器开口填充结构(11,55,62,63)。每个支撑开口19内的基座沟道部分11(如果存在)、存储器膜50、垂直半导体沟道60、电介质芯62和漏极区域63的每个组合填充相应的支撑开口19，并且构成支撑柱结构。

[0087] 参考图6，示出了在存储器开口49和支撑开口19内分别形成存储器开口填充结构58和支撑柱结构20之后的示例性结构。在图4A和图4B的结构的每个存储器开口49内可以形成存储器开口填充结构58的实例。在图4A和图4B的结构的每个支撑开口19内可以形成支撑柱结构20的实例。

[0088] 存储器开口填充结构58中的每一个可以包括相应的基座沟道部分11、相应的存储器堆叠体结构55(即，包括相应的存储器膜50和相应的垂直半导体沟道60)、以及相应的漏极区域63。支撑柱结构20中的每一个可以包括相应的基座沟道部分11、相应的存储器膜50、相应的垂直半导体沟道60、以及位于相应的垂直半导体沟道60的顶端部处的相应的虚设漏极区域63。支撑柱结构20的每个部件是电力上无源的且浮置的虚设部件。换言之，支撑柱结构20的每个基座沟道部分11是虚设基座沟道部分，支撑柱结构20的每个存储器膜50是虚设存储器膜，支撑柱结构20的每个垂直半导体沟道60是虚设垂直半导体沟道，并且支撑柱结构20的每个漏极区域63是虚设漏极区域，其不电连接到位线。

[0089] 可以在梯台区域中穿过阶梯式表面的水平表面形成支撑柱结构20的子集或全部。支撑柱结构20中的一个或多个可以穿过最底部牺牲材料层42与倒退阶梯式电介质材料部分65之间的水平界面，并且可以包含接触倒退阶梯式电介质材料部分65的相应的基座沟道部分11。

[0090] 在存储器开口49和支撑开口19中的每一个中，存储器膜50和垂直半导体沟道60直接形成在相应的基座沟道部分11上。每个存储器堆叠体结构55包含垂直半导体沟道60和存储器膜50，垂直半导体沟道60可以包括多个半导体沟道层(601,602)。存储器膜50可以包括横向围绕垂直半导体沟道60的隧穿电介质层56，以及横向围绕隧穿电介质层56的电荷储存区域(如实现为存储器材料层54)和可选的阻挡电介质层52的垂直堆叠体。尽管采用为存储器堆叠体结构示出的配置来描述本公开，但是本公开的方法可以应用于替代的存储器堆叠

Micron Ex. 1002, p. 569
Micron v. YMTC
IPR2024-00789

体结构，替代的堆叠体结构包含用于存储器膜50和/或用于垂直半导体沟道60的不同层堆叠体或结构。

[0091] 参考图7A-7C，接触级电介质层73可以形成在绝缘层32和牺牲材料层42的交替堆叠体(32,42)之上，并且形成在存储器堆叠体结构55和支撑柱结构20之上。接触级电介质层73包含与牺牲材料层42的电介质材料不同的电介质材料。例如，接触级电介质层73可以包含硅氧化物。接触级电介质层73可以具有在从50nm到500nm的范围内的厚度，虽然还可以采用更小和更大的厚度。

[0092] 光刻胶层(未示出)可以被施加在接触级电介质层73之上，并且可以被光刻地图案化，以在存储器堆叠体结构55的群集(cluster)之间的区域中形成开口。可以采用各向异性蚀刻来穿过接触级电介质层73、交替堆叠体(32,42)和/或倒退阶梯式电介质材料部分65来转移光刻胶层中的图案，以形成背侧沟槽79，背侧沟槽79从接触级电介质层73的顶表面至少垂直地延伸到基板(9,10)的顶表面，并且横向地延伸穿过存储器阵列区域100和接触区域300。在一个实施例中，背侧沟槽79可以包含源极接触开口，在源极接触开口中可以随后形成源极接触通孔结构。可以例如通过灰化来移除光刻胶层。

[0093] 参考图8和图9A，可以例如采用蚀刻工艺将相对于绝缘层32的第一材料选择性地蚀刻牺牲材料层42的第二材料的蚀刻剂引入到背侧沟槽79中。图9A示出了图8的示例性结构的区域。背侧凹陷43形成在从其中移除牺牲材料层42的体积中。牺牲材料层42的第二材料的移除可以对绝缘层32的第一材料、倒退阶梯式电介质材料部分65的材料、半导体材料层10的半导体材料、和存储器膜50的最外层的材料是有选择性的。在一个实施例中，牺牲材料层42可以包含硅氮化物，并且绝缘层32和倒退阶梯式电介质材料部分65的材料可以从硅氧化物和电介质金属氧化物中选择。

[0094] 对第一材料和存储器膜50的最外层有选择性地移除第二材料的蚀刻工艺可以是采用湿法蚀刻溶液的湿法蚀刻工艺，或者可以是气相(干法)蚀刻工艺，气相蚀刻工艺中蚀刻剂以气相被引入到背侧沟槽79中。例如，如果牺牲材料层42包含硅氮化物，则蚀刻工艺可以是湿法蚀刻工艺，其中示例性结构浸入到包含磷酸的湿法蚀刻槽内，磷酸相对于硅氧化物、硅和本领域中采用的各种其他材料有选择性地蚀刻硅氮化物。支撑柱结构20、倒退阶梯式电介质材料部分65和存储器堆叠体结构55提供结构的支撑，而背侧凹陷43存在于先前被牺牲材料层42所占据的体积内。

[0095] 每个背侧凹陷43可以是横向延伸的腔，该腔具有大于腔的垂直范围的横向尺寸。换言之，每个背侧凹陷43的横向尺寸可以大于背侧凹陷43的高度。多个背侧凹陷43可以形成在从其中移除牺牲材料层42的第二材料的体积内。与背侧凹陷43相比，在其中形成存储器堆叠体结构55的存储器开口在本文中称为前侧开口或前侧腔。在一个实施例中，存储器阵列区域100包括具有设置在基板(9,10)上方的多个器件级的单片三维NAND串的阵列。在这种情况下，每个背侧凹陷43可以限定用于接收单片三维NAND串的阵列的相应的字线的空间。

[0096] 多个背侧凹陷43中的每一个可以实质上平行于基板(9,10)的顶表面延伸。可以由下面的绝缘层32的顶表面和上面的绝缘层32的底表面垂直地界定背侧凹陷43。在一个实施例中，每个背侧凹陷43可以通体具有均匀的高度。

[0097] 可以通过将半导体材料热转化和/或等离子体转化为电介质材料，而将可选的基

Micron Ex. 1002, p. 570
Micron v. YMTC
IPR2024-00789

座沟道部分11和半导体材料层10的物理暴露表面部分转化为电介质材料部分。例如，热转化和/或等离子体转化可以用于将每个基座沟道部分11的表面部分转化成管状电介质间隔体116，并且用于将半导体材料层10的每个物理暴露的表面部分转化成平坦电介质部分616。在一个实施例中，每个管状电介质间隔体116可以是与环面(torus)拓扑同胚的，即大体上为环形。如本文所使用的，如果元件的形状可以在不破坏孔或者形成新孔的情况下连续拉伸为环面的形状，则元件与环面拓扑同胚。电介质间隔体116包含电介质材料，电介质材料包含与基座沟道部分11相同的半导体元素，并且附加地包含诸如氧和/或氮的至少一种非金属元素，使得管状电介质间隔体116的材料是电介质材料。在一个实施例中，管状电介质间隔体116可以包含基座沟道部分11的半导体材料的电介质氧化物、电介质氮化物或电介质氮氧化物。同样地，每个平坦电介质部分616包含电介质材料，该电介质材料包含与半导体材料层相同的半导体元素，并且附加地包含诸如氧和/或氮的至少一种非金属元素，使得平坦电介质部分616的材料是电介质材料。在一个实施例中，平坦电介质部分616可以包含半导体材料层10的半导体材料的电介质氧化物、电介质氮化物或电介质氮氧化物。

[0098] 参考图9B，可以可选地形成背侧阻挡电介质层44。背侧阻挡电介质层44(如果存在)包括电介质材料，该电介质材料起到随后在背侧凹陷43中形成的控制栅极的控制栅极电介质的功能。在阻挡电介质层52存在于每个存储器开口内的情况下，背侧阻挡电介质层44是可选的。在省略阻挡电介质层52的情况下，背侧阻挡电介质层44是存在的。

[0099] 背侧阻挡电介质层44可以形成在背侧凹陷43中和背侧沟槽79的侧壁上。背侧阻挡电介质层44可以直接形成在背侧凹陷43内的绝缘层32的水平表面和存储器堆叠体结构55的侧壁上。如果形成背侧阻挡电介质层44，则在形成背侧阻挡电介质层44之前形成管状电介质间隔体116和平坦电介质部分616是可选的。在一个实施例中，可以通过诸如原子层沉积(ALD)的保形沉积工艺来形成背侧阻挡电介质层44。背侧阻挡电介质层44可以基本由铝氧化物构成。背侧阻挡电介质层44的厚度可以在从1nm到15nm的范围内，诸如从2到6nm，虽然还可以采用更小和更大的厚度。

[0100] 背侧阻挡电介质层44的电介质材料可以是诸如铝氧化物的电介质金属氧化物，至少一种过渡金属元素的电介质氧化物，至少一种镧系元素的电介质氧化物，铝、至少一种过渡金属元素和/或至少一种镧系元素的组合的电介质氧化物。替代地或者附加地，背侧阻挡电介质层44可以包含硅氧化物层。可以通过保形沉积方法(诸如化学气相沉积或原子层沉积)来沉积背侧阻挡电介质层44。背侧阻挡电介质层44形成在背侧沟槽79的侧壁、绝缘层32的水平表面和侧壁、存储器堆叠体结构55的侧壁表面的物理暴露于背侧凹陷43的部分，以及平坦电介质部分616的顶表面上。背侧腔79'存在于每个背侧沟槽79的未填充背侧阻挡电介质层44的部分内。

[0101] 参考图9C，金属屏障层46A可以沉积在背侧凹陷43中。金属屏障层46A包含导电金属材料，其可以起到要随后沉积的金属填充材料的扩散屏障层和/或粘合促进层的作用。金属屏障层46A可以包含诸如TiN、TaN、WN或其堆叠体的导电金属氮化物材料，或者可以包含诸如TiC、TaC、WC或其堆叠体的导电金属碳化物材料。在一个实施例中，可以通过保形沉积工艺(诸如化学气相沉积(CVD)或原子层沉积)来沉积金属屏障层46A。金属屏障层46A的厚度可以在从2nm到8nm的范围内，诸如从3nm到6nm，虽然还可以采用更小和更大的厚度。在一个实施例中，金属屏障层46A可以基本由诸如TiN的导电金属氮化物构成。

Micron Ex. 1002, p. 571
Micron v. YMTC
IPR2024-00789

[0102] 参考图9D和图10，金属填充材料沉积在多个背侧凹陷43中，沉积在至少一个背侧沟槽79的侧壁上，并且沉积在接触级电介质层73的顶表面之上，以形成金属填充材料层46B。金属填充材料可以通过保形沉积方法来沉积，保形沉积方法可以是例如化学气相沉积(CVD)、原子层沉积(ALD)、无电镀、电镀或其组合。在一个实施例中，金属填充材料层46B可以基本由至少一种单质金属构成。可以例如从钨、钴、钌、钛和钽中选择金属填充材料层46B的至少一种单质金属。在一个实施例中，金属填充材料层46B可以基本由单个单质金属构成。在一个实施例中，可以采用诸如WF_6的含氟前驱体气体来沉积金属填充材料层46B。在一个实施例中，金属填充材料层46B可以是钨层，钨层包含作为杂质的残留水平的氟原子。金属填充材料层46B由金属屏障层46A与绝缘层32和存储器堆叠体结构55间隔，金属屏障层46A是阻挡氟原子穿过其进扩散的金属屏障层。

[0103] 多个导电层46可以形成在多个背侧凹陷43中，并且连续金属材料层46L可以形成在每个背侧沟槽79的侧壁上并且形成在接触级电介质层73之上。每个导电层46包含金属屏障层46A的部分和金属填充材料层46B的部分，该金属屏障层46A的部分和金属填充材料层46B的部分位于电介质材料层的垂直相邻的对之间，该电介质材料层的垂直相邻的对可以是绝缘层32的对、最底部的绝缘层和栅极电介质层12的对、或者最顶部的绝缘层和绝缘帽层70的对。连续金属材料层46L包含金属屏障层46A的连续部分和金属填充材料层46B的连续部分，该金属屏障层46A的连续部分和金属填充材料层46B的连续部分位于背侧沟槽79中或者位于接触级电介质层73的上方。

[0104] 导电层46包含最底部导电层和在最底部导电层上面的导电层46。最底部导电层位于源极选择级，并且因此，在本文中称为源极选择级导电层46S。源极选择级导电层46S包括源极选择栅电极，其选择位于背侧沟槽的相邻对之间的垂直半导体沟道60的组，用于在三维存储器器件的操作期间的存取。在最底部导电层上面的导电层46包含可选的一个或多个附加的(多个)源极选择栅电极46SA(图14A中所示)、包含或电连接(短路)到三维存储器器件的字线的字线级导电层46W、以及位于源极选择栅电极与字线之间的可选的虚设字线46WD(如图14A中所示)。如果采用漏极选择级隔离结构72，则包含最顶部导电层的导电层46中的至少一个可以为漏极选择级导电层。漏极选择级导电层沿着第一水平方向相互由漏极选择级隔离结构72横向隔离开，该漏极选择级隔离结构72沿着第一水平方向横向延伸。

[0105] 可以用导电层46取代每个牺牲材料层42。背侧腔79'存在于每个背侧沟槽79的未填充背侧阻挡电介质层44和连续金属材料层46L的部分中。每个管状电介质间隔体116横向地围绕相应的基座沟道部分11。导电层46中的最底部的一个是源极选择级导电层46S。一经形成导电层46，源极选择级导电层46S横向地围绕每个管状电介质间隔体116。支撑柱结构20中的每一个包括管状电介质间隔体116，管状电介质间隔体116位于基座沟道部分11与源极选择级导电层之间。

[0106] 参考图11，例如通过各向同性湿法蚀刻、各向异性干法蚀刻或其组合，从每个背侧沟槽79并从接触级电介质层73的上方回蚀刻连续导电材料层46L的沉积的金属材料。沉积的金属材料在背侧凹陷43中的每个剩余部分构成导电层46。每个导电层46可以是导电线结构。因此，可以用导电层46取代牺牲材料层42。

[0107] 每个字线级导电层46W可以起到位于相同级处的多个控制栅电极和将位于相同级处的多个控制栅电极电互连(即电短路)的字线的组合的功能。字线级导电层46W内的多个

Micron Ex. 1002, p. 572
Micron v. YMTC
IPR2024-00789

控制栅电极可以是包含存储器堆叠体结构55的垂直存储器器件的控制栅电极。换言之，每个字线级导电层46W可以是字线，字线起到多个垂直存储器器件的公共控制栅电极的功能。

[0108] 在一个实施例中，连续导电材料层46L的移除可以对背侧阻挡电介质层44的材料是有选择性的。在这种情况下，背侧阻挡电介质层44的水平部分可以存在于每个背侧沟槽79的底部处。栅极电介质层12可以通过背侧阻挡电介质层44的水平部分而与背侧沟槽79间隔。

[0109] 在另一个实施例中，连续导电材料层46L的移除可以对背侧阻挡电介质层44的材料是没有选择性的，或者可以不采用背侧阻挡电介质层44。在这种情况下，根据在移除连续导电材料层46L期间栅极电介质层12是未被移除还是被部分地移除，栅极电介质层12的顶表面和/或侧壁表面可以物理暴露在背侧沟槽79的底部处。可以在移除连续导电材料层46L期间共同地移除平坦电介质部分616。背侧腔79'存在于每个背侧沟槽79内。

[0110] 参考图12A和图12B，通过保形沉积工艺，可以在至少一个背侧沟槽79中并在接触级电介质层73之上形成绝缘材料层。示例性的保形沉积工艺包含但不限于化学气相沉积和原子层沉积。绝缘材料层包含诸如硅氧化物、氮氧化物、电介质金属氧化物、有机硅酸盐玻璃或其组合的绝缘材料。在一个实施例中，绝缘材料层可以包含硅氧化物。可以通过低压化学气相沉积(LPCVD)或者原子层沉积(ALD)来形成绝缘材料层。绝缘材料层的厚度可以在从1.5nm到60nm的范围内，虽然还可以采用更小和更大的厚度。

[0111] 如果存在背侧阻挡电介质层44，则绝缘材料层可以直接形成在背侧阻挡电介质层44的表面上，并且直接形成在导电层46的侧壁上。如果不采用背侧阻挡电介质层44，则绝缘材料层可以直接形成在绝缘层32的侧壁上，并且直接形成在导电层46的侧壁上。

[0112] 进行各向异性蚀刻，以从接触级电介质层73的上方以及每个背侧沟槽79的底部处移除绝缘材料层的水平部分。绝缘材料层的每个剩余部分构成绝缘间隔体74。背侧腔79'存在于被每个背侧沟槽74围绕的体积内。

[0113] 各项异性蚀刻工艺可以在蚀刻化学过程改变或者不改变的情况下继续，以移除可选的背侧阻挡电介质层44和平坦电介质部分616的在穿过绝缘间隔体74的开口下面的部分。可以在每个背侧沟槽79的底部处物理暴露半导体材料层10的顶表面。

[0114] 可以通过将电掺杂剂注入到半导体材料层10的物理暴露的表面部分，在半导体材料层10的在每个背侧腔79'的下方的表面部分处形成源极区域61。每个源极区域61形成在基板(9,10)的在穿过绝缘间隔体74的相应的开口处下面的表面部分中。由于在注入工艺期间所注入的掺杂剂原子的散乱(straggle)和在随后的激活退火工艺期间所注入的掺杂剂原子的横向扩散，每个源极区域61可以具有比穿过绝缘间隔体74的开口的横向范围更大的横向范围。

[0115] 半导体材料层10的在源极区域61和多个基座沟道部分11之间延伸的上部部分构成多个场效应晶体管的水平半导体沟道59。水平半导体沟道59通过相应的基座沟道部分11连接到多个垂直半导体沟道60。水平半导体沟道59接触源极区域61和多个基座沟道部分11。一经形成交替堆叠体(32,46)内的导电层46而提供的最底部的导电层46可以包括场效应晶体管的选择栅电极。每个源极区域61形成在半导体基板(9,10)的上部部分中。半导体沟道(59,11,60)在每个源极区域61和漏极区域63的相应的集合之间延伸。半导体沟道(59,11,60)包含存储器堆叠体结构55的垂直半导体沟道60。

Micron Ex. 1002, p. 573
Micron v. YMTC
IPR2024-00789

[0116] 背侧接触通孔结构76可以形成在每个背侧腔79'内。每个接触通孔结构76可以填充相应的腔79'。可以通过在背侧沟槽79的剩余未填充的体积(即,背侧腔79')中沉积至少一种导电材料来形成接触通孔结构76。例如,至少一种导电材料可以包含导电衬垫76A和导电填充材料部分76B。导电衬垫76A可以包含导电金属衬垫,诸如TiN、TaN、WN、TiC、TaC、WC、其合金,或者其堆叠体。导电衬垫76A的厚度可以在从3nm到30nm的范围内,虽然还可以采用更小和更大的厚度。导电填充材料部分76B可以包含金属或者金属合金。例如,导电填充材料部分76B可以包含W、Cu、Al、Co、Ru、Ni、其合金,或者其堆叠体。

[0117] 可以采用在交替堆叠体(32,46)上面的接触级电介质层73作为停止层来平坦化至少一种导电材料。如果采用化学机械平坦化(CMP)工艺,则接触级电介质层73可以用作CMP停止层。至少一种导电材料在背侧沟槽79中的每个剩余的连续部分构成背侧接触通孔结构76。

[0118] 背侧接触通孔结构76延伸穿过交替堆叠体(32,46),并且接触源极区域61的顶表面。如果采用背侧阻挡电介质层44,则背侧接触通孔结构76可以接触背侧阻挡电介质层44的侧壁。

[0119] 参考图13A-13E,可以穿过接触级电介质层73并可选地穿过倒退阶梯式电介质材料部分65来形成附加接触通孔结构(88,86,8P)。例如,可以穿过每个漏极区域63上的接触级电介质层73来形成漏极接触通孔结构88。可以穿过接触级电介质层73并且穿过倒退阶梯式电介质材料部分65,在导电层46上形成导电线接触通孔结构86。可以穿过倒退阶梯式电介质材料部分65直接在外围器件的相应节点上形成外围器件接触通孔结构8P。导电线接触通孔结构86可以包含接触源极选择级导电层46S的源极选择级接触通孔结构86S、接触字线级导电层46W的中的相应的一个的字线级接触通孔结构86W、以及接触相应的漏极选择级导电层46D的漏极选择级接触通孔结构86D。

[0120] 图14A-14B图示了本公开的替代实施例的接触区域300的下部部分。在此实施例中,交替堆叠体还可以包含附加的源极选择栅电极46SA和可选的虚设字线46WD,以及相应的接触通孔结构86SA和86WD。虚设字线46WD不控制存储器单元电荷储存状态,而字线46W起到NAND器件的控制栅电极的作用,并且控制存储器单元电荷储存状态。如图13C、图13E和14B所示,除源极选择级接触通孔结构86S之外的接触通孔结构86(例如,86SA,86W,86WD)中的每一个可以被支撑柱结构20的相应的集合横向围绕。在一个示例中,接触通孔结构86中的每一个可以形成在由四个支撑柱结构20横向限制的空间内,支撑柱结构20穿过相同水平平面中的开口,该水平平面是导电层46的水平顶表面与倒退阶梯式电介质材料部分65之间的界面。然而,省略了源极选择级接触通孔结构86S周围的四个支撑柱结构20。

[0121] 支撑柱结构20起到在存在背侧凹陷43的同时防止薄绝缘层32的弯曲和塌陷的作用。在直接在最底部背侧凹陷43上面的绝缘层(例如,倒退阶梯式电介质材料部分65和/或下部绝缘层32)具有足够的厚度的情况下,由于这样的厚绝缘层提供的机械强度,在最底部背侧凹陷43上面的此绝缘层的弯曲可以是不显著的。在说明性示例中,在最底部背侧凹陷43上面的绝缘层的厚度可以在从60nm至600nm的范围内,并且阶梯式表面中的每个阶梯的横向尺寸可以在从200nm至4000nm的范围内。在此情况下,在最底部背侧凹陷43上面的绝缘层的机械强度可以足以防止塌陷,即使在源极选择栅极接触通孔结构86S的区域内没有支撑柱结构20。

Micron Ex. 1002, p. 574
Micron v. YMTC
IPR2024-00789

[0122] 在一个示例中，源极选择级接触通孔结构86S与支撑柱结构20之间的最小横向分离距离可以大于字线级接触通孔结构86W与支撑柱结构20之间的任何最小横向分离距离。在一个示例中，在与垂直轴线VA相距源极选择栅极接触通孔结构86S的最大横向尺寸MLD的三倍内(诸如2.5倍内、诸如两倍内(即，二倍)、诸如一倍半内)的整个体积不含任何支撑柱结构20，垂直轴线VA穿过源极选择栅极接触通孔结构86S的几何中心GC。支撑柱结构20位于其余的字线栅极接触通孔结构(86W，86WD)和漏极选择栅极接触通孔结构周围(即，在最大横向尺寸二倍内)，如图14B中所示。

[0123] 在一个示例中，源极选择级接触通孔结构86S可以形成在与源极选择级导电层46S的最接近的侧壁横向地间隔开横向分离距离的位置处，分离距离大于源极选择栅极接触通孔结构46S的最大横向尺寸MLD。如本文中所用的，元件的“最大横向尺寸”是指元件在水平截面图的整个汇总中可以测量的最大尺寸。在一些情况下，源极选择级导电层46S的最大横向尺寸MLS可以是源极选择级导电层46S的直径。在一些其他情况下，源极选择级导电层46S的最大横向尺寸MLS可以是在源极选择级导电层46S的水平截面形状之中可以找到的最大椭圆的主轴线。源极选择级导电层46S可以具有垂直侧壁，或锥形侧壁。

[0124] 例如，居中在穿过源极选择接触通孔结构86S的几何中心GC的垂直轴线VA处且相切地接触源极选择栅极接触通孔结构46S的最接近的侧壁的圆柱的半径可以具有半径R，半径R是源极选择栅极接触通孔结构86S的最大横向尺寸MLD的至少一倍半。在一个示例中，字线级接触通孔结构86W中的每一个可以形成有相应的最大横向尺寸，并且可以与支撑柱结构20中的最近的一个横向地间隔开横向距离，横向距离小于相应的最大横向尺寸。在一个示例中，源极选择栅极接触通孔结构86S的最大横向尺寸MLD可以与字线级接触通孔结构86W的最大横向尺寸相同。

[0125] 本公开的示例性结构包含三维存储器器件。三维存储器器件包括：位于基板(9，10)之上的绝缘层32和导电层46的交替堆叠体(32，46)，其中导电层46包括源极选择级导电层46S和位于源极选择级导电层46S之上的字线级导电层46W；延伸穿过交替堆叠体(32，46)的存储器堆叠体结构55，其中存储器堆叠体结构55中的每一个包括存储器膜50和接触存储器膜50的内侧壁的垂直半导体沟道60；梯台区域，交替堆叠体(32，46)内除最顶部导电层46之外的每个导电层46在梯台区域中比交替堆叠体(32，46)内的上面的导电层46横向地延伸得更远；在梯台区域上面的倒退阶梯式电介质材料部分65；支撑柱结构20，支撑柱结构20位于梯台区域中，并且延伸穿过倒退阶梯式电介质材料部分65和交替堆叠体(32，46)；源极选择级接触通孔结构86S，源极选择级接触通孔结构86S延伸穿过倒退阶梯式电介质材料部分65并且接触源极选择级导电层46S；以及字线级接触通孔结构86W，字线级接触通孔结构86W延伸穿过倒退阶梯式电介质材料部分65并且接触字线级导电层46W中的相应的一个；其中源极选择级接触通孔结构86S与支撑柱结构20之间的最小横向分离距离大于字线级接触通孔结构86W与支撑柱结构20之间的任何最小横向分离距离。

[0126] 在一个示例中，源极选择级导电层46S包括源极选择栅电极，与穿过源极选择栅极接触通孔结构86S的几何中心GC的垂直轴线VA相距源极选择栅极接触通孔结构86S的最大横向尺寸MLD两倍内的整个体积可以不含任何支撑柱结构20，并且可以由包含交替堆叠体(32，46)的最顶层的顶表面的水平平面与包含源极选择级导电层46S的顶表面的水平平面之间的倒退阶梯式电介质材料部分65的电介质材料构成。

Micron Ex. 1002, p. 575
Micron v. YMTC
IPR2024-00789

[0127] 在一个示例中，阶梯式表面在梯台区域中从交替堆叠体(32,46)的最底层连续延伸到交替堆叠体(32,46)的最顶层，并且支撑柱结构20穿过阶梯式表面的每个水平表面。在一个示例中，字线级接触通孔结构46W中的每一个具有相应的最大横向尺寸，并且与支撑柱结构20中的最近的一个横向地间隔开小于相应的最大横向尺寸的横向距离。例如，每个字线级接触通孔结构46W的最大直径可以大于字线级接触通孔结构46W与支撑柱结构20之间的最小横向距离。在一个示例中，源极选择级接触通孔结构86S可以与源极选择级导电层46S的最接近的侧壁横向地间隔开大于源极选择栅极接触通孔结构86S的最大横向尺寸的横向分离距离。

[0128] 在一个示例中，支撑柱结构20中的每一个包括基座沟道部分11，基座沟道部分11接触基板(9,10)中的半导体材料层10。在一个示例中，支撑柱结构20中的每一个包括位于基座沟道部分11与源极选择级导电层46S之间的管状电介质间隔体116。

[0129] 在一个示例中，支撑柱结构20中的每一个包括存储器膜50、位于存储器膜50内且接触基座沟道部分11的垂直半导体沟道60、以及位于垂直半导体沟道60的顶端部处的漏极区域63。在一个示例中，支撑柱结构20中的一个穿过源极选择级导电层46S和倒退阶梯式电介质材料部分65，并且包含接触倒退阶梯式电介质材料部分65的相应的基座沟道部分11。在此情况下，这样的支撑柱结构20不穿过除源极选择级导电层46S之外的任何其他导电层46。

[0130] 在实施例中，每个存储器堆叠体结构55在相应的基座沟道部分11上面并与之接触，每个垂直半导体沟道60的上端部与相应的漏极区域63接触，并且接触存储器堆叠体结构55的每个漏极区域63与相应的漏极接触通孔结构88接触，相应的漏极接触通孔结构88嵌入在接触级电介质层73内，接触级电介质层73在存储器阵列区域(例如，存储器平面)100中在存储器堆叠体结构55上面。支撑柱结构20中的每一个的整个顶表面可以在接触区域300中与接触级电介质层73的底表面接触。

[0131] 在一个实施例中，三维存储器器件包括垂直NAND存储器器件。导电层46可以包括或可以电连接到单片三维NAND存储器器件的相应字线。基板(9,10)可以包括硅基板。垂直NAND存储器器件可以包括硅基板之上的单片三维NAND串的阵列。单片三维NAND串的阵列的第一器件级中的至少一个存储器单元(如实现为电荷储存层54的在字线级导电层46W的级处的部分)可以位于单片三维NAND串的阵列的第二器件级中的另一个存储器单元(如实现为电荷储存层54的在另一个字线级导电层46W的级处的另一个部分)之上。硅基板可以含有集成电路，集成电路包括位于硅基板上的存储器器件的驱动器电路。字线级导电层46W可以包括多个控制栅电极，控制栅电极具有实质上平行于基板(9,10)的顶表面(例如在背侧沟槽79的对之间)延伸的条形。多个控制栅电极至少包括位于第一器件级中的第一控制栅电极和位于第二器件级中的第二控制栅电极。单片三维NAND串的阵列可以包括：多个半导体沟道(59,11,60)，其中多个半导体沟道(59,11,60)中的每一个的至少一个端部部分60实质上垂直于基板(9,10)的顶表面延伸；以及多个电荷储存元件(如实现为电荷捕获材料部分)。每个电荷储存元件可以位于相邻于多个半导体沟道(59,11,60)中的相应的一个。

[0132] 如图15所示，在现有技术三维存储器器件的制造工艺中，可能发生支撑柱结构20与导电层接触通孔结构86之间的覆盖变化。此外，基座沟道部分11的顶表面形成在包含源极选择级导电层46S的顶表面的水平平面上方，因为源极选择级导电层46S需要在其整个厚

Micron Ex. 1002, p. 576
Micron v. YMTC
IPR2024-00789

度之上提供电偏置。位于仅延伸穿过源极选择级导电层46S而不延伸穿过任何字线级导电层46W的每个支撑柱结构20内的基座沟道部分11接触倒退阶梯式电介质材料部分65。这样的基座沟道部分11有电短路到源极选择级接触通孔结构86S的倾向，如图15所示。

[0133] 在本公开的实施例中，通过从居中在每个源极选择级接触通孔结构86S的几何中心处的预定体积消除支撑柱结构20，降低或消除了源极选择级接触通孔结构86S与基座沟道部分11之间的电短路的可能性。因此，可以通过改变支撑柱结构20的图案，使得从其中形成源极选择栅极接触通孔结构86的区域省略支撑柱结构20，而防止由源极选择级接触通孔结构86S与支撑柱结构20的未对准所导致的电短路。在源极选择级导电层46S上面的级处的倒退阶梯式电介质材料部分65具有足够的厚度，以在形成背侧凹陷43期间防止变形，使得在源极选择级接触通孔结构86S的区域中不需要支撑柱结构20。

[0134] 通过包含源极选择级接触通孔结构86S和支撑柱结构20的各种元件的几何配置，本公开的方法和结构可以提供三维存储器器件的更高产率和提高的可靠性。

[0135] 尽管前述涉及特定的优选实施例，但是应当理解，本公开不限于此。对于本领域的普通技术人员而言，可以对所公开的实施例进行各种修改，并且此类修改意图在本公开的范围内。在本公开中示出采用特定结构和/或配置的实施例的情况下，应该理解，本公开可以采用功能上等同的任何其他兼容结构和/或配置来实践，只要此类取代不是明确禁止的，或对于本领域的普通技术人员而言是不可能的。在本文中所引用的所有出版物、专利申请和专利均通过引用以其整体并入本文。

Micron Ex. 1002, p. 577
Micron v. YMTC
IPR2024-00789

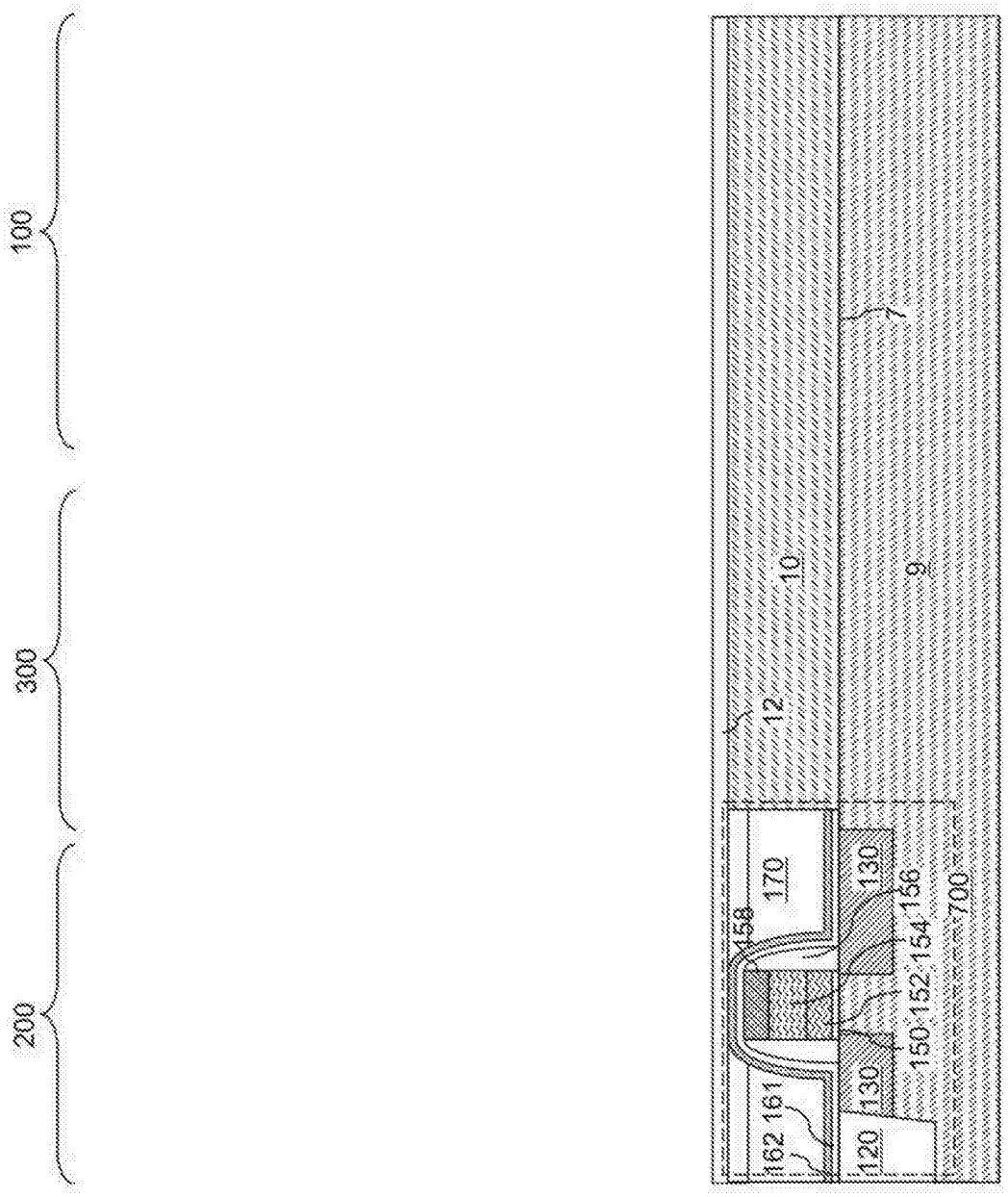


图1

Micron Ex. 1002, p. 578
Micron v. YMTC
IPR2024-00789

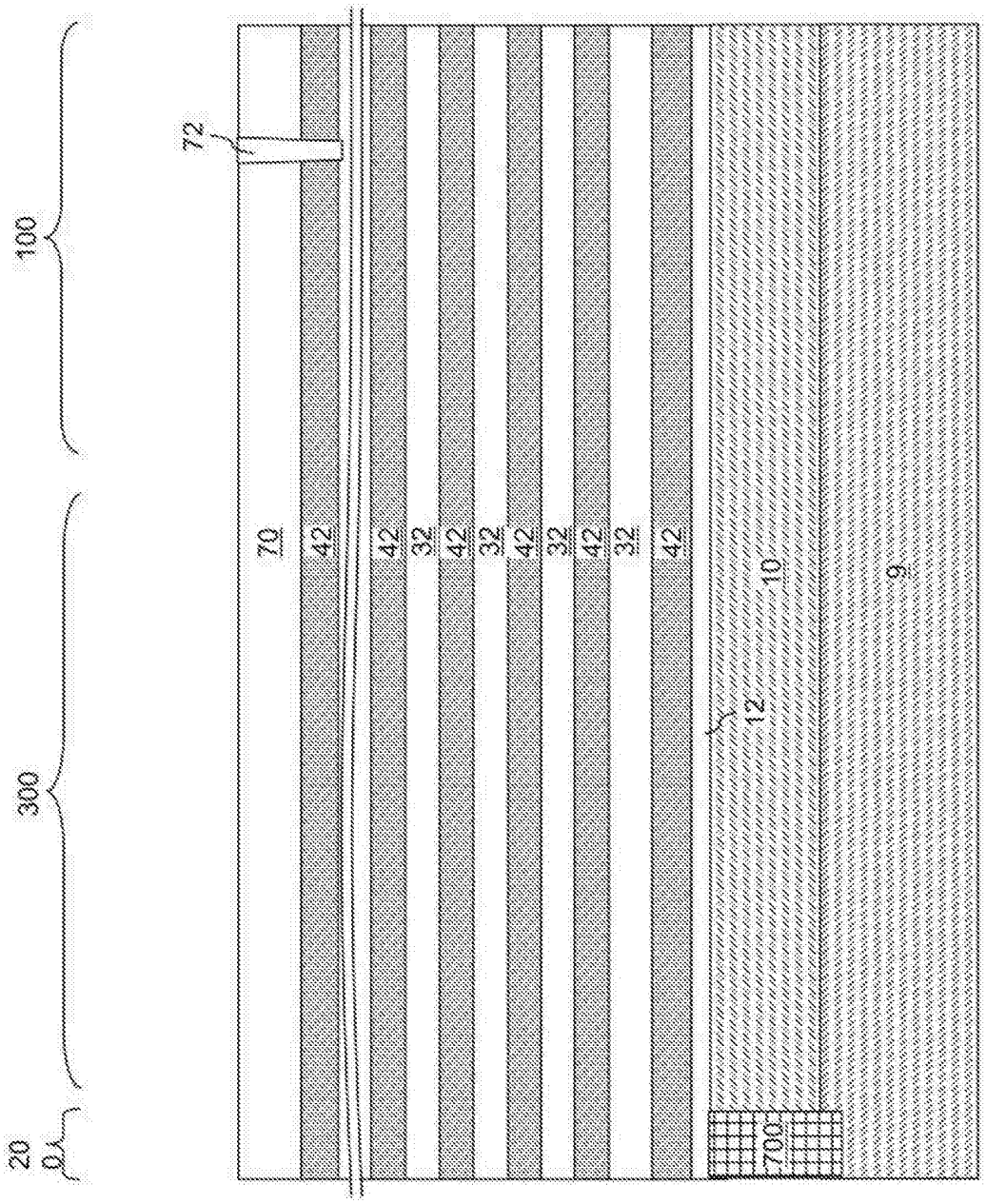


图2

Micron Ex. 1002, p. 579
Micron v. YMTC
IPR2024-00789

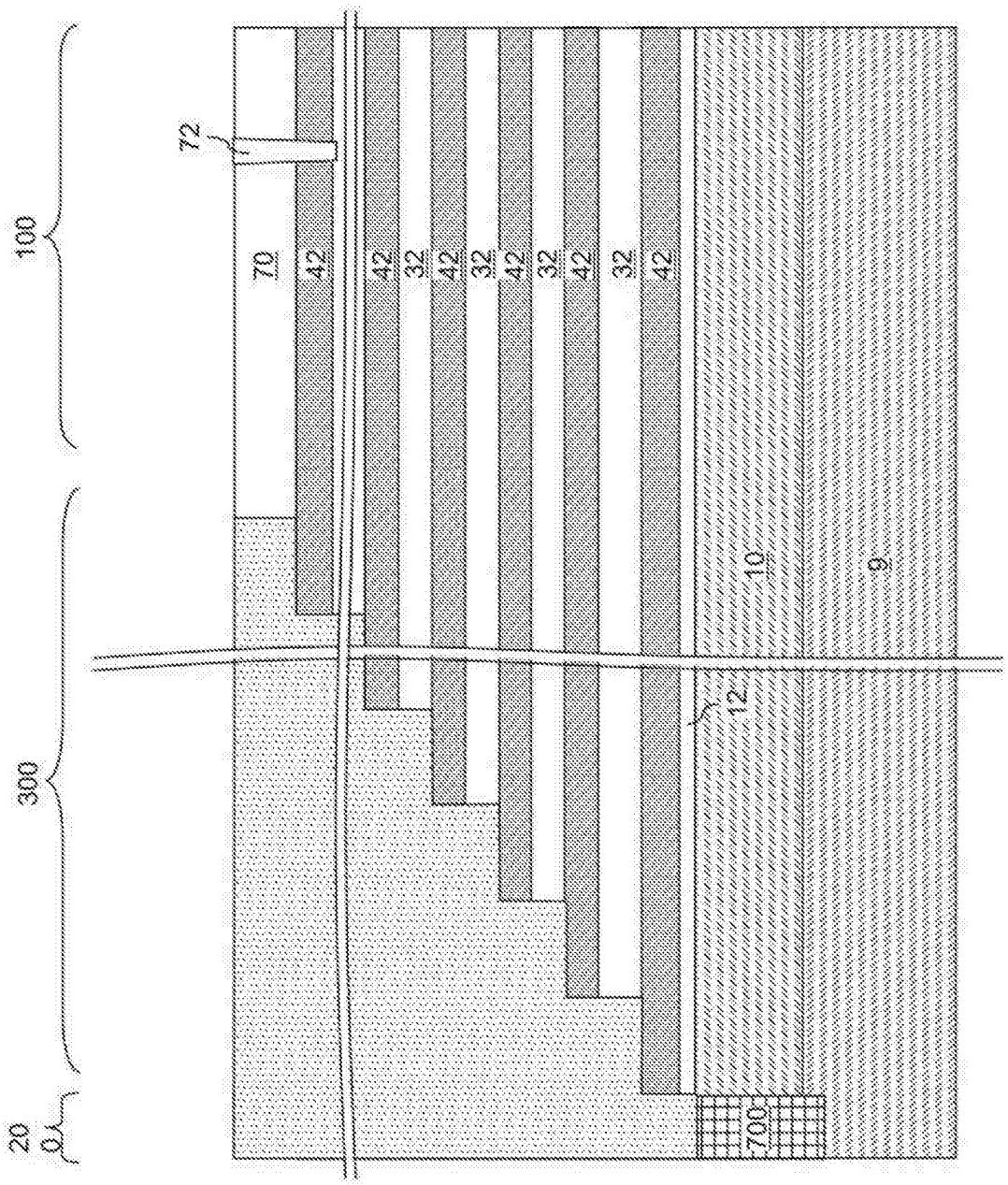


图3

Micron Ex. 1002, p. 580
Micron v. YMTC
IPR2024-00789

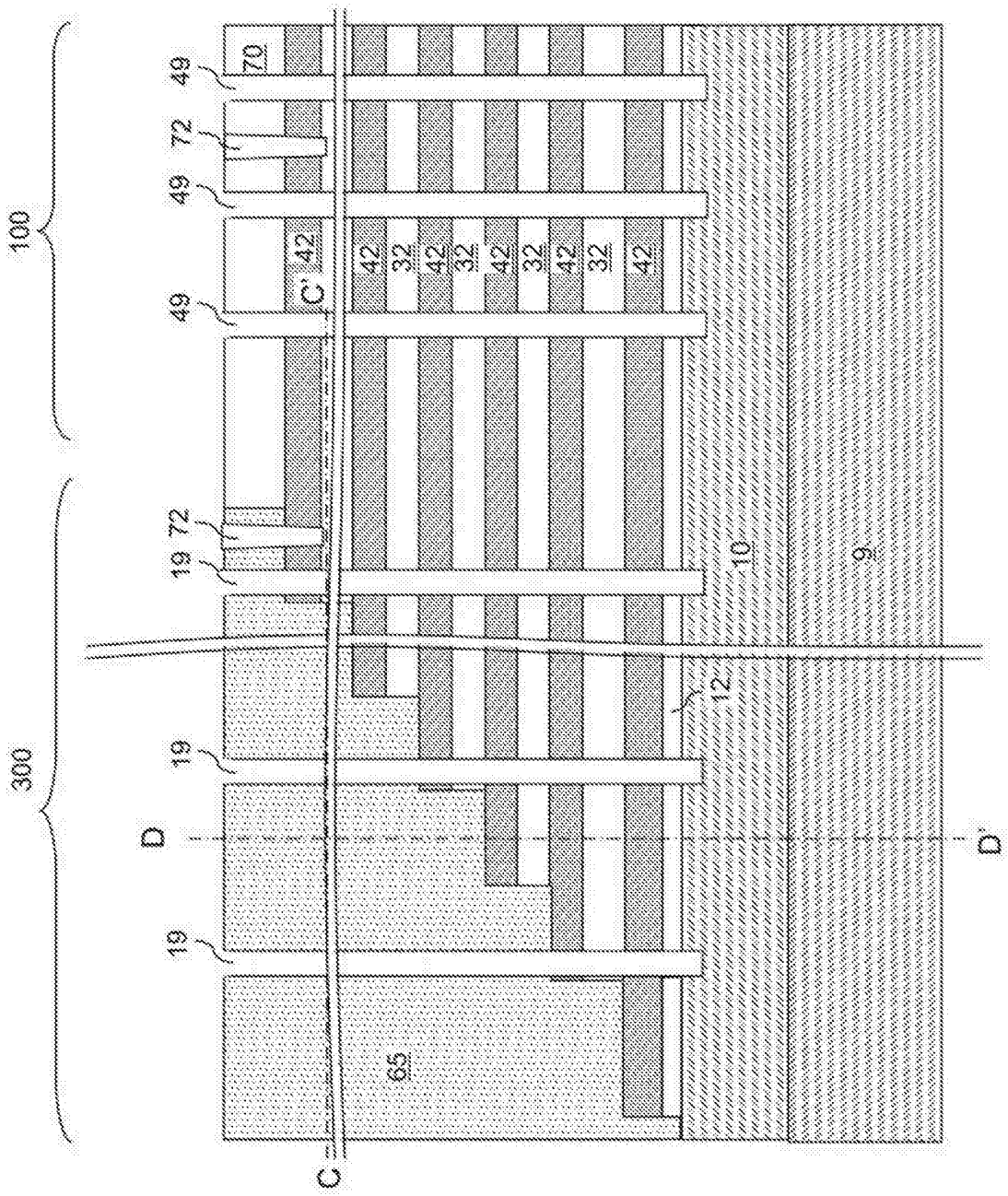


图4A

Micron Ex. 1002, p. 581
Micron v. YMTC
IPR2024-00789

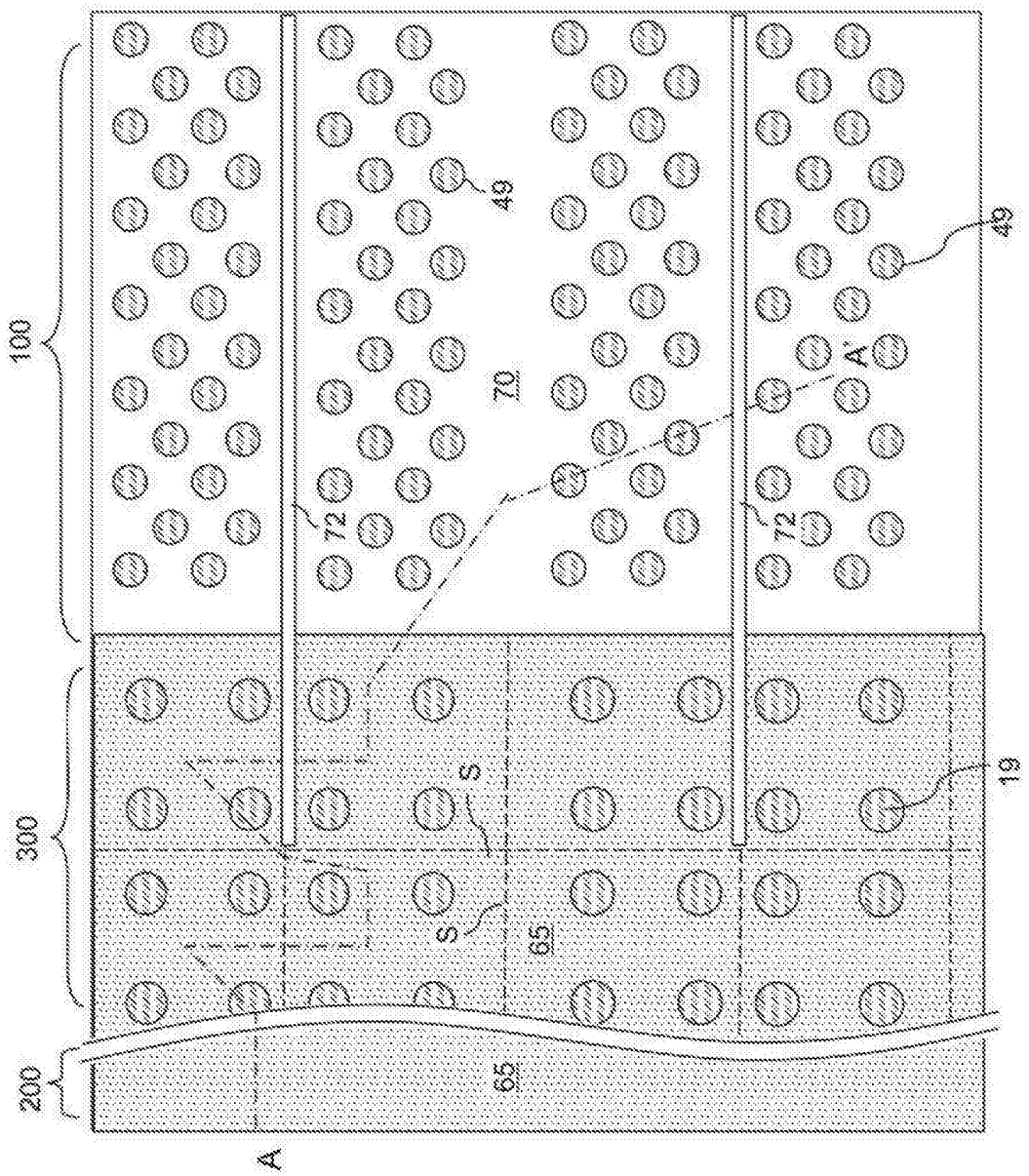


图4B

Micron Ex. 1002, p. 582
Micron v. YMTC
IPR2024-00789

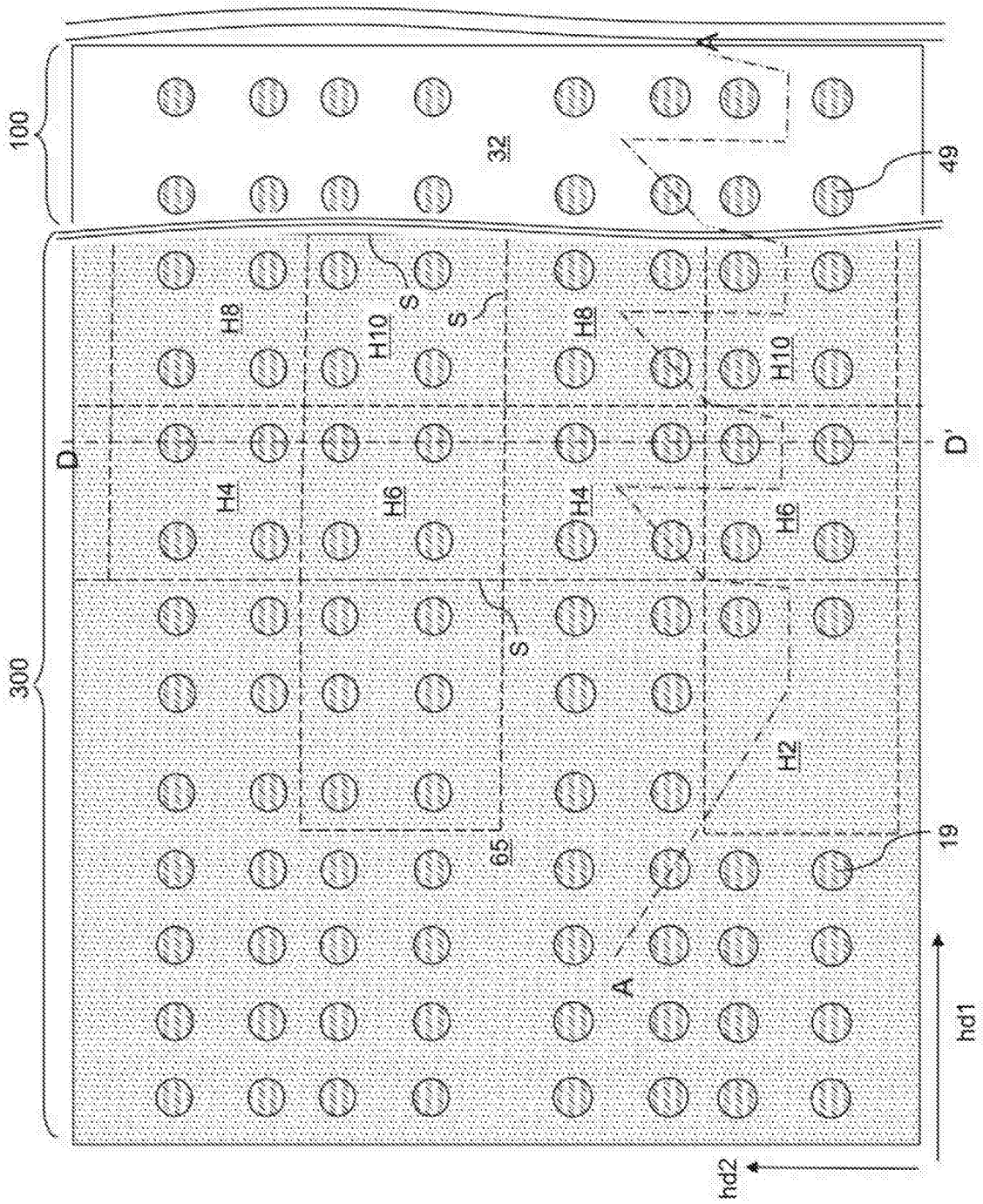


图4C

Micron Ex. 1002, p. 583
Micron v. YMTC
IPR2024-00789

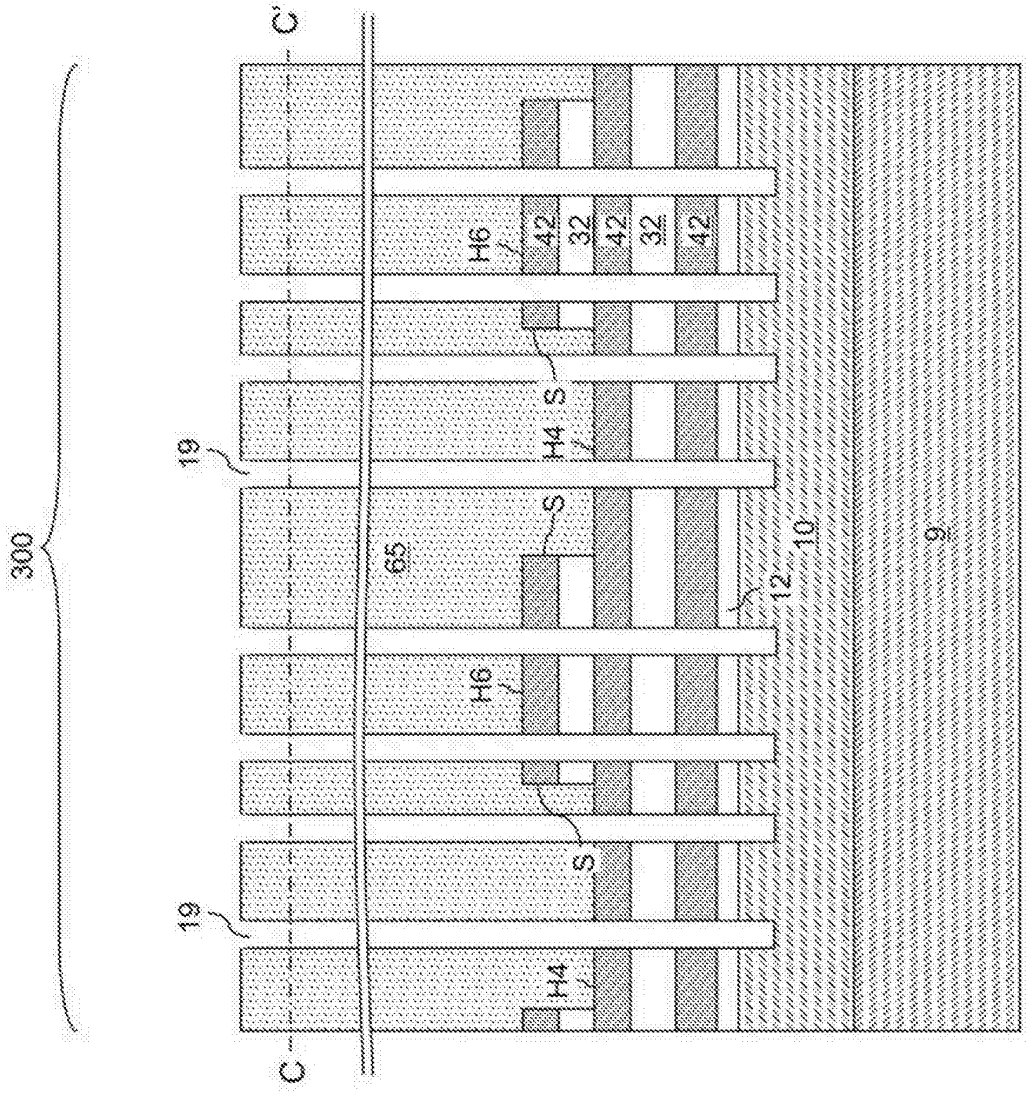


图4D

Micron Ex. 1002, p. 584
Micron v. YMTC
IPR2024-00789

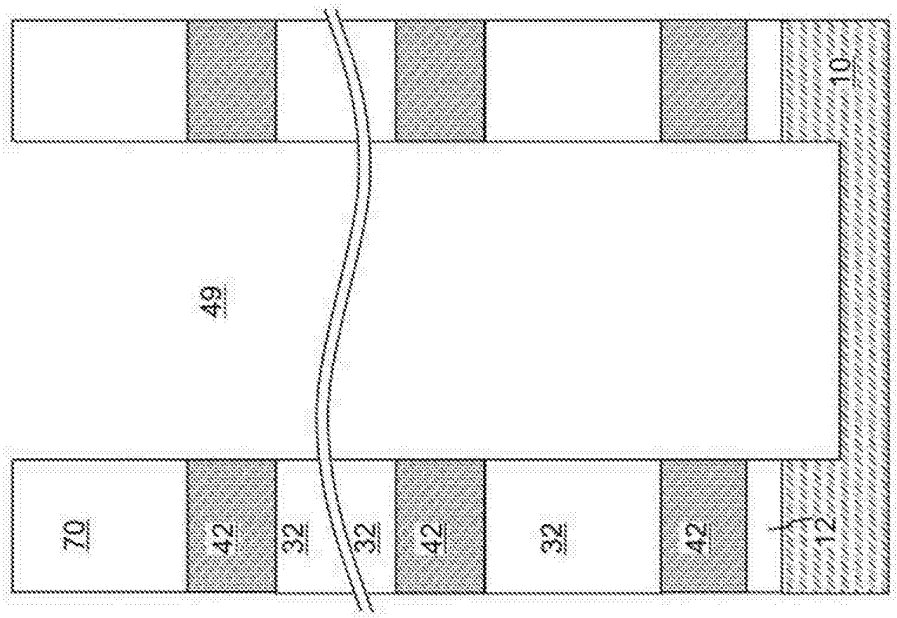


图5A

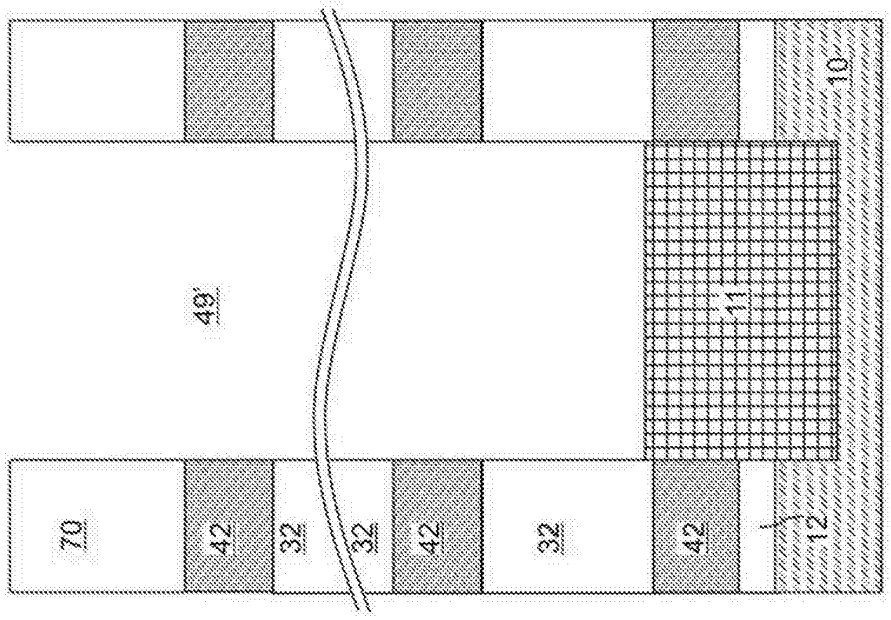


图5B

Micron Ex. 1002, p. 585
Micron v. YMTC
IPR2024-00789

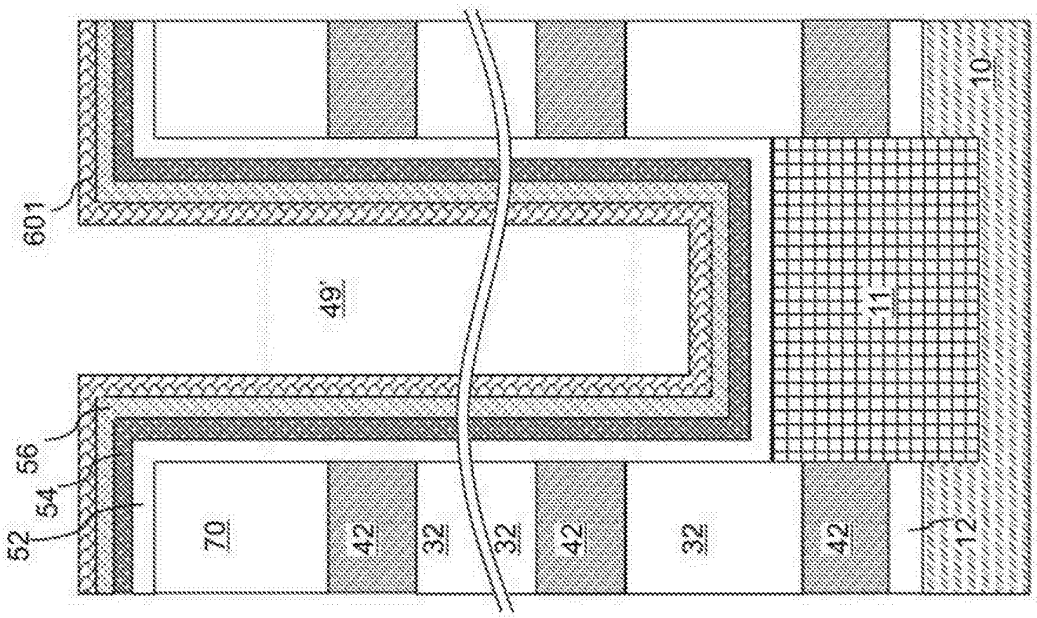


图5C

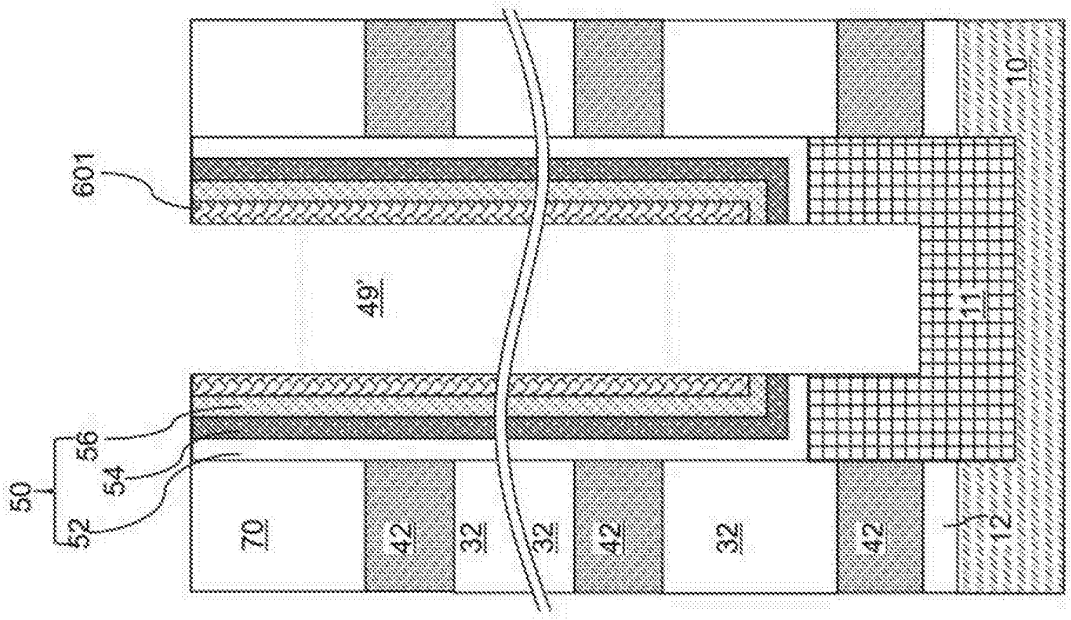


图5D

Micron Ex. 1002, p. 586
Micron v. YMTC
IPR2024-00789

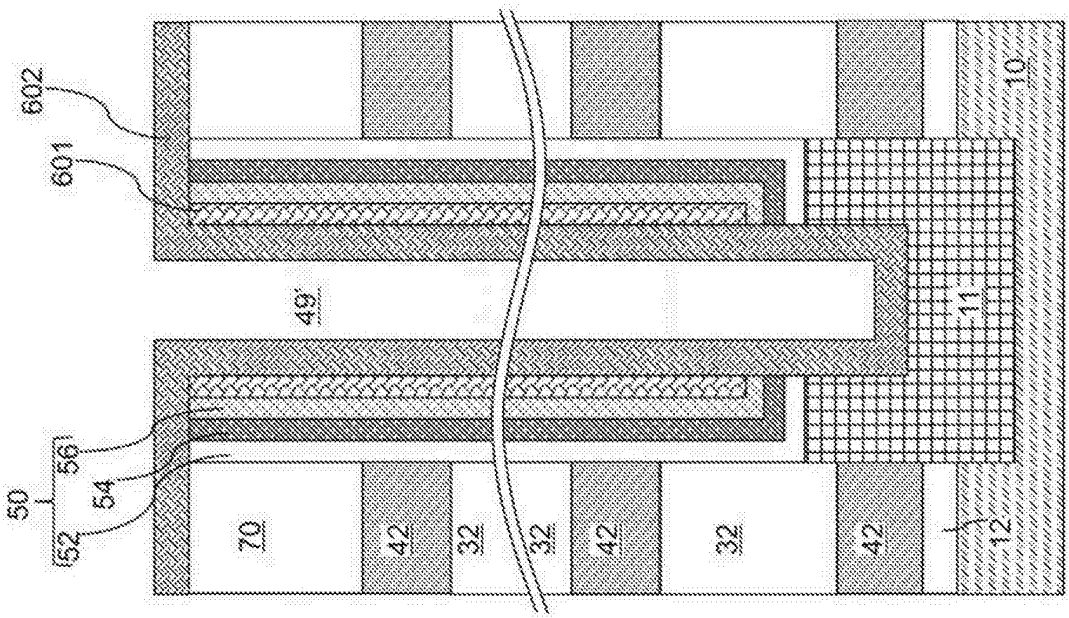


图5E

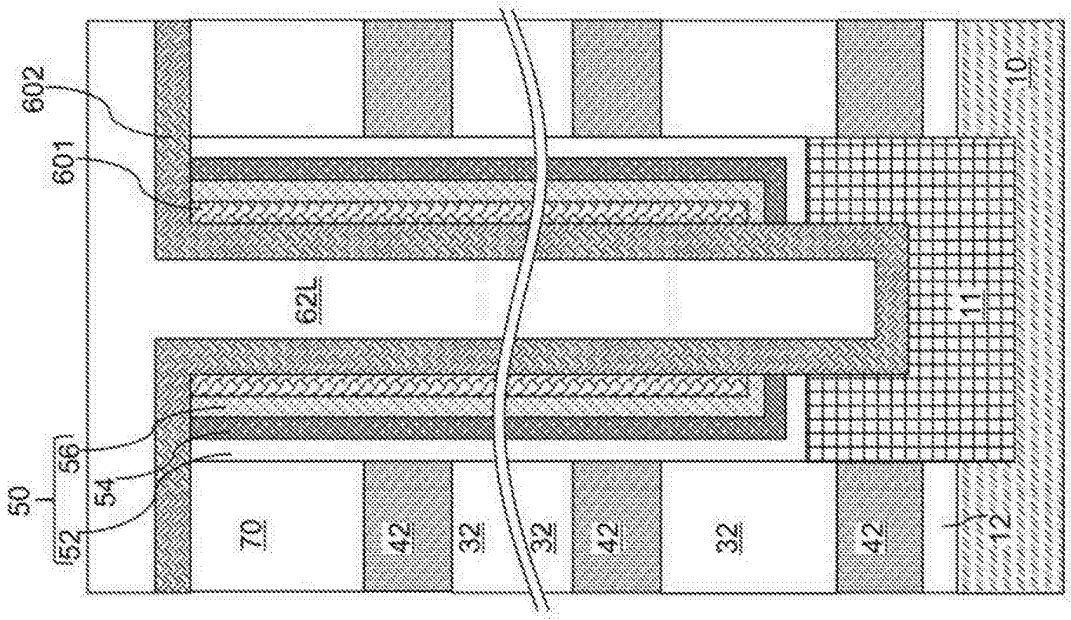


图5F

Micron Ex. 1002, p. 587
Micron v. YMTC
IPR2024-00789

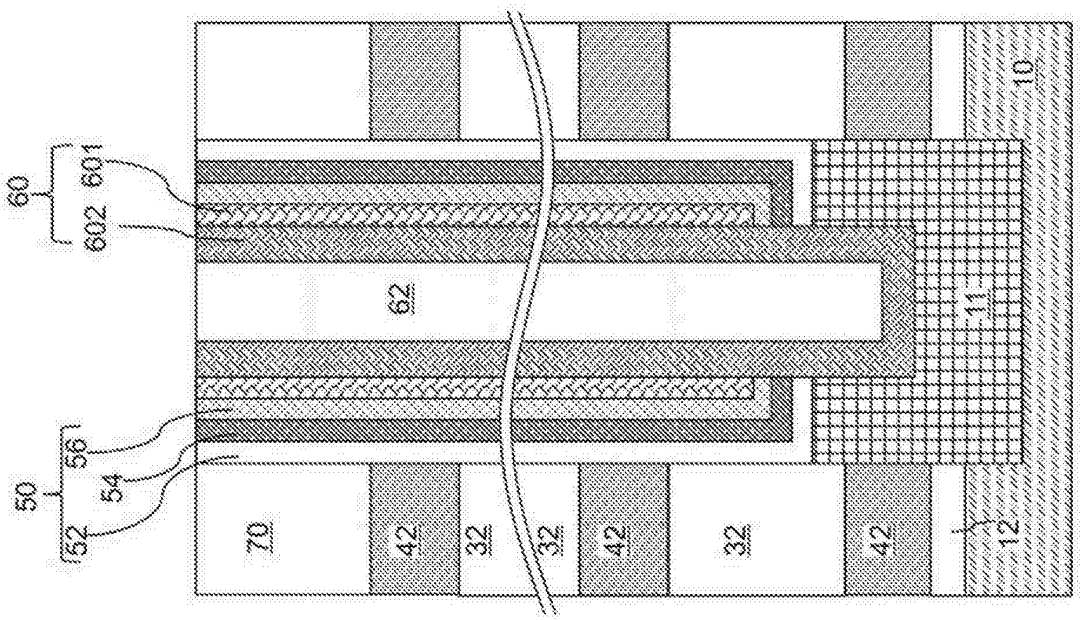


图5G

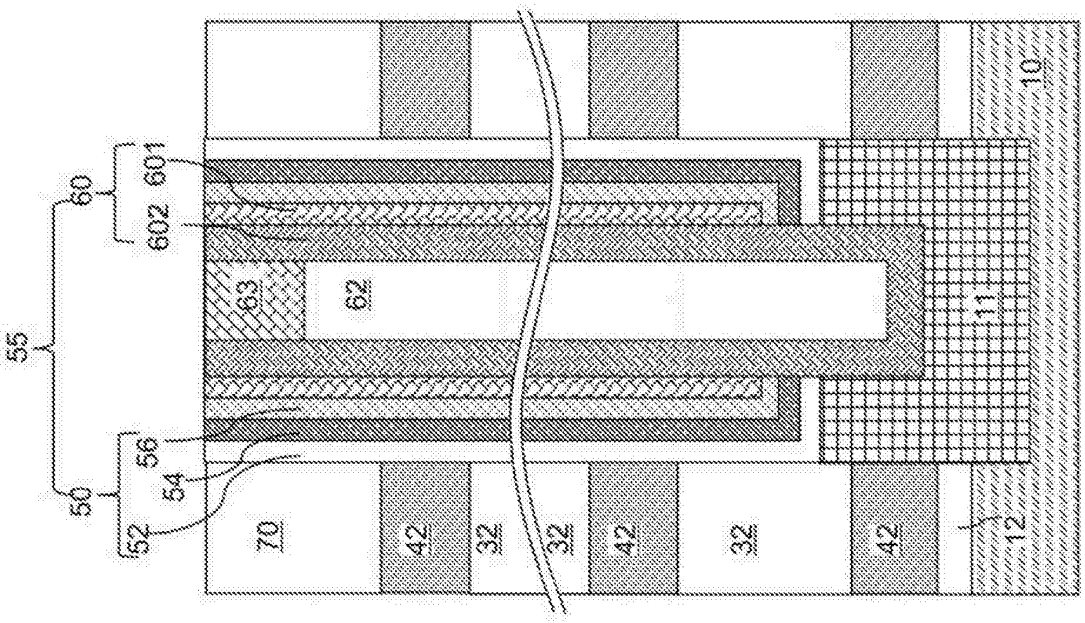


图5H

Micron Ex. 1002, p. 588
Micron v. YMTC
IPR2024-00789

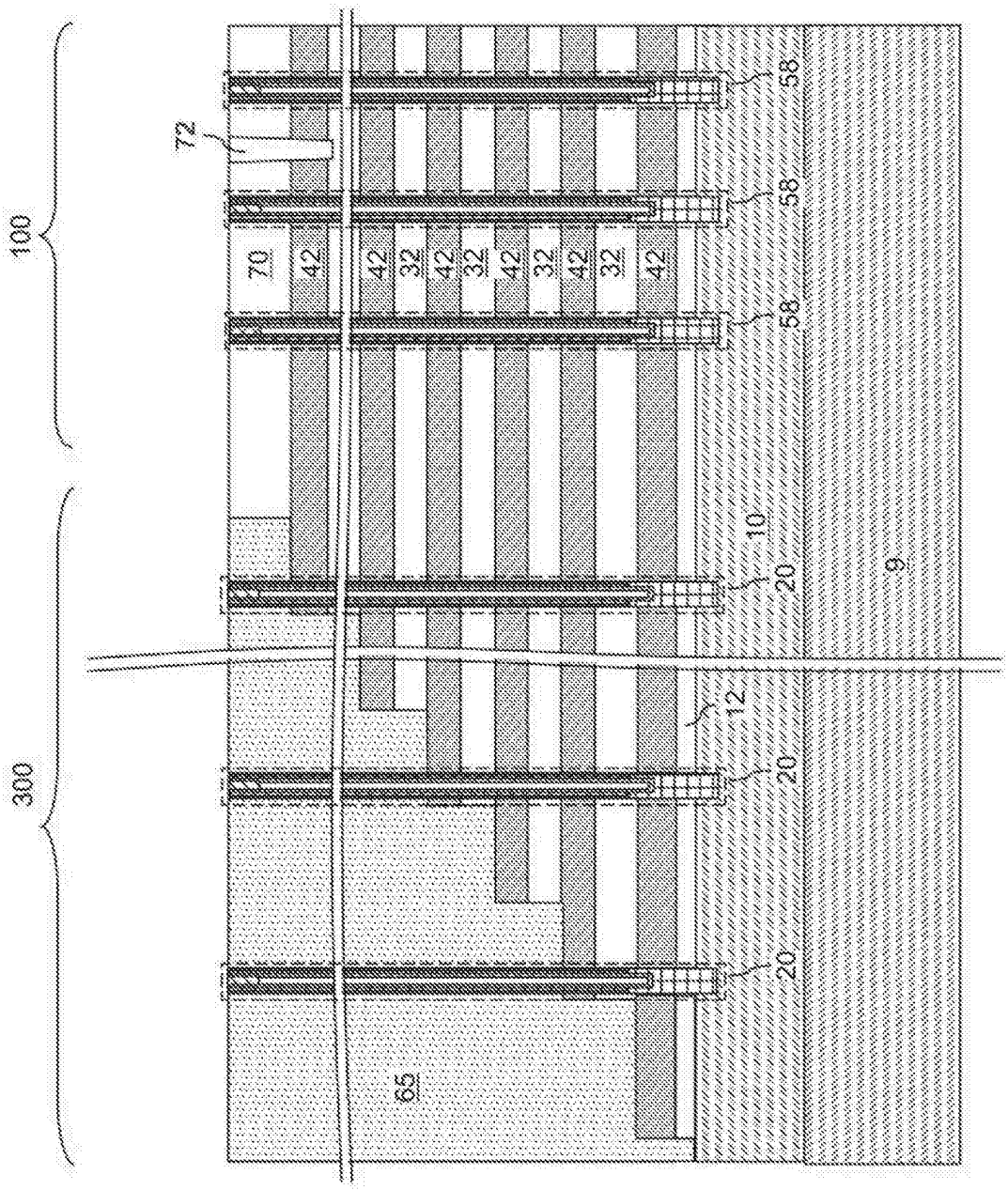


图6

Micron Ex. 1002, p. 589
Micron v. YMTC
IPR2024-00789

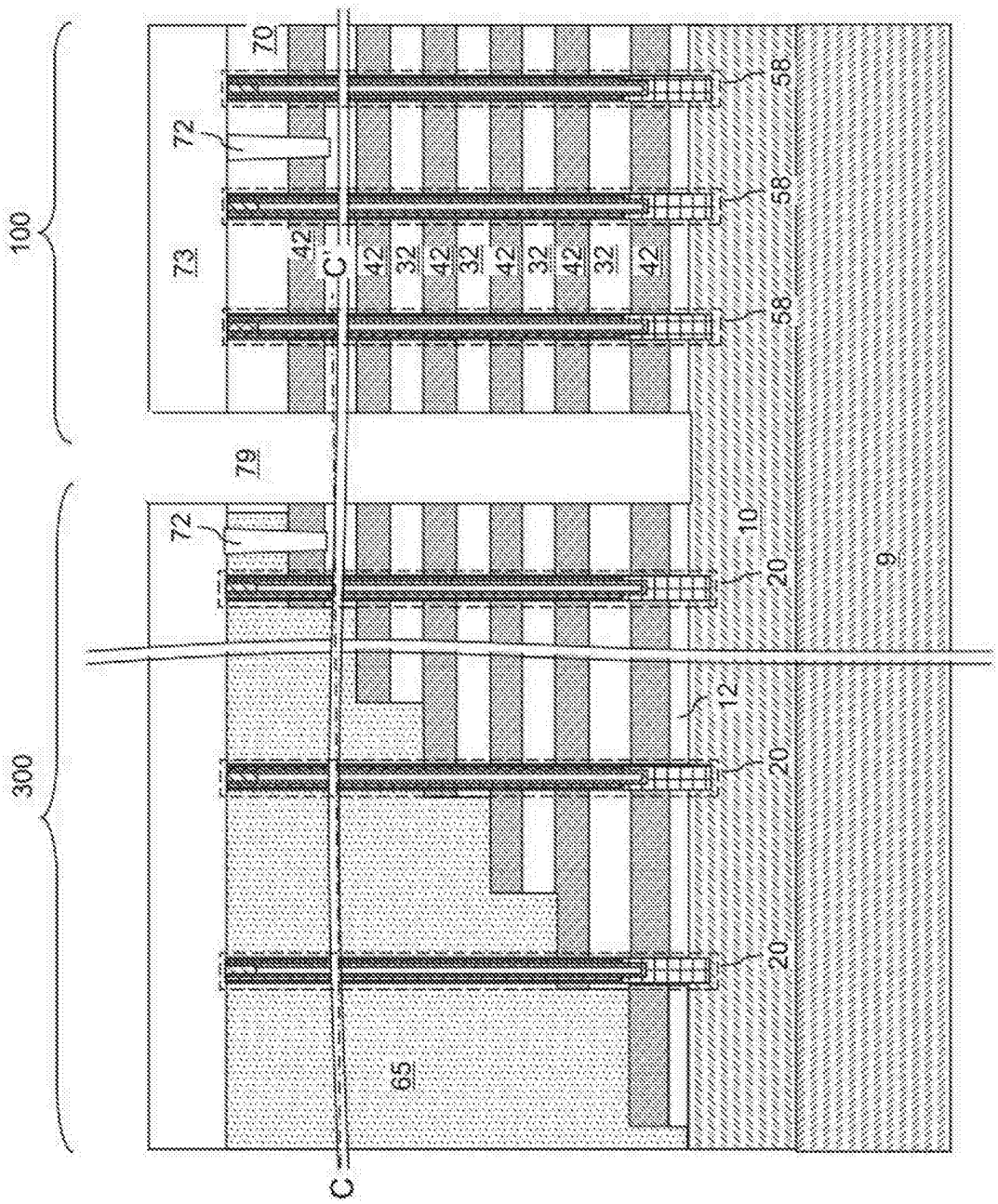


图7A

Micron Ex. 1002, p. 590
Micron v. YMTC
IPR2024-00789

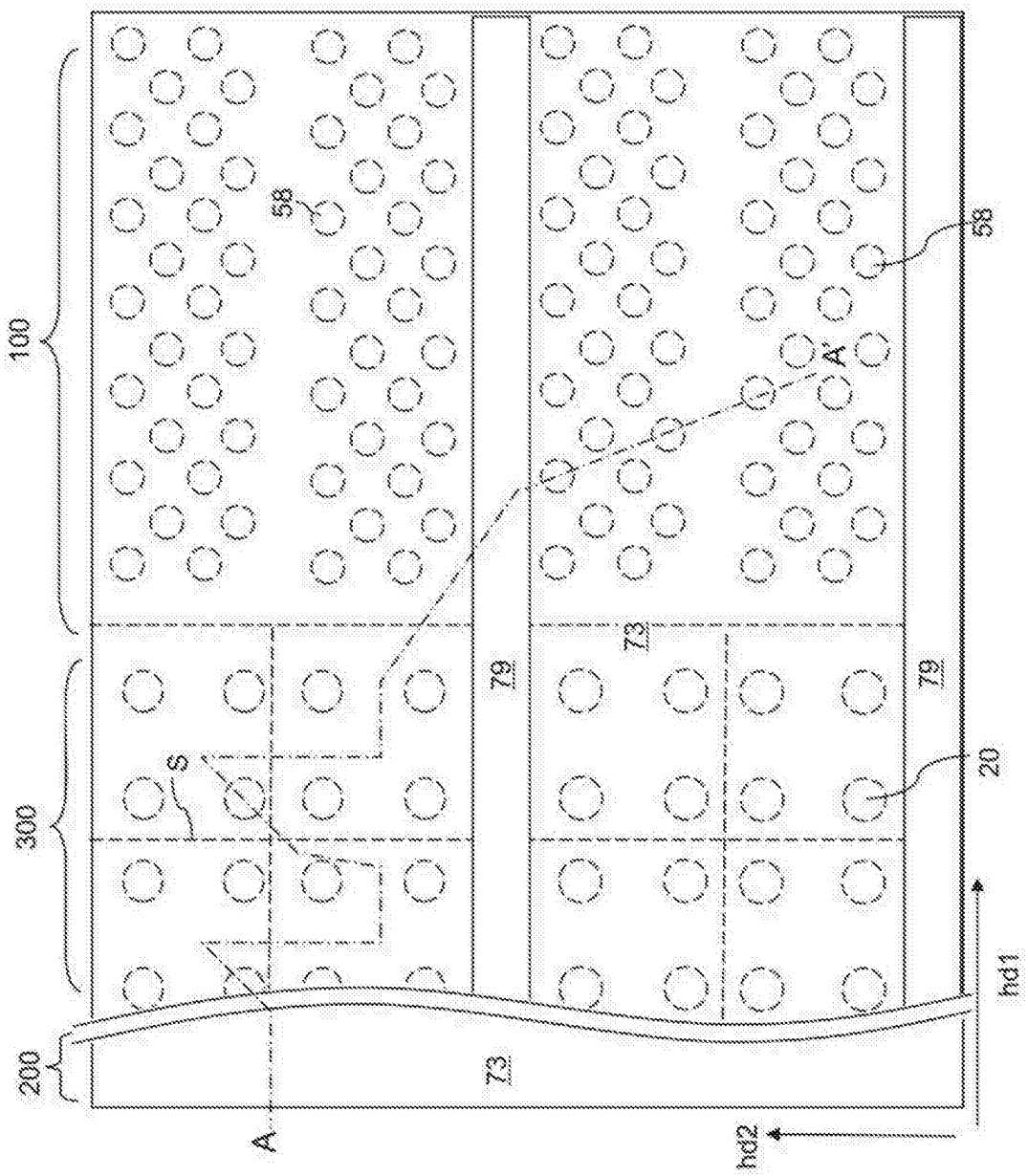


图7B

Micron Ex. 1002, p. 591
Micron v. YMTC
IPR2024-00789

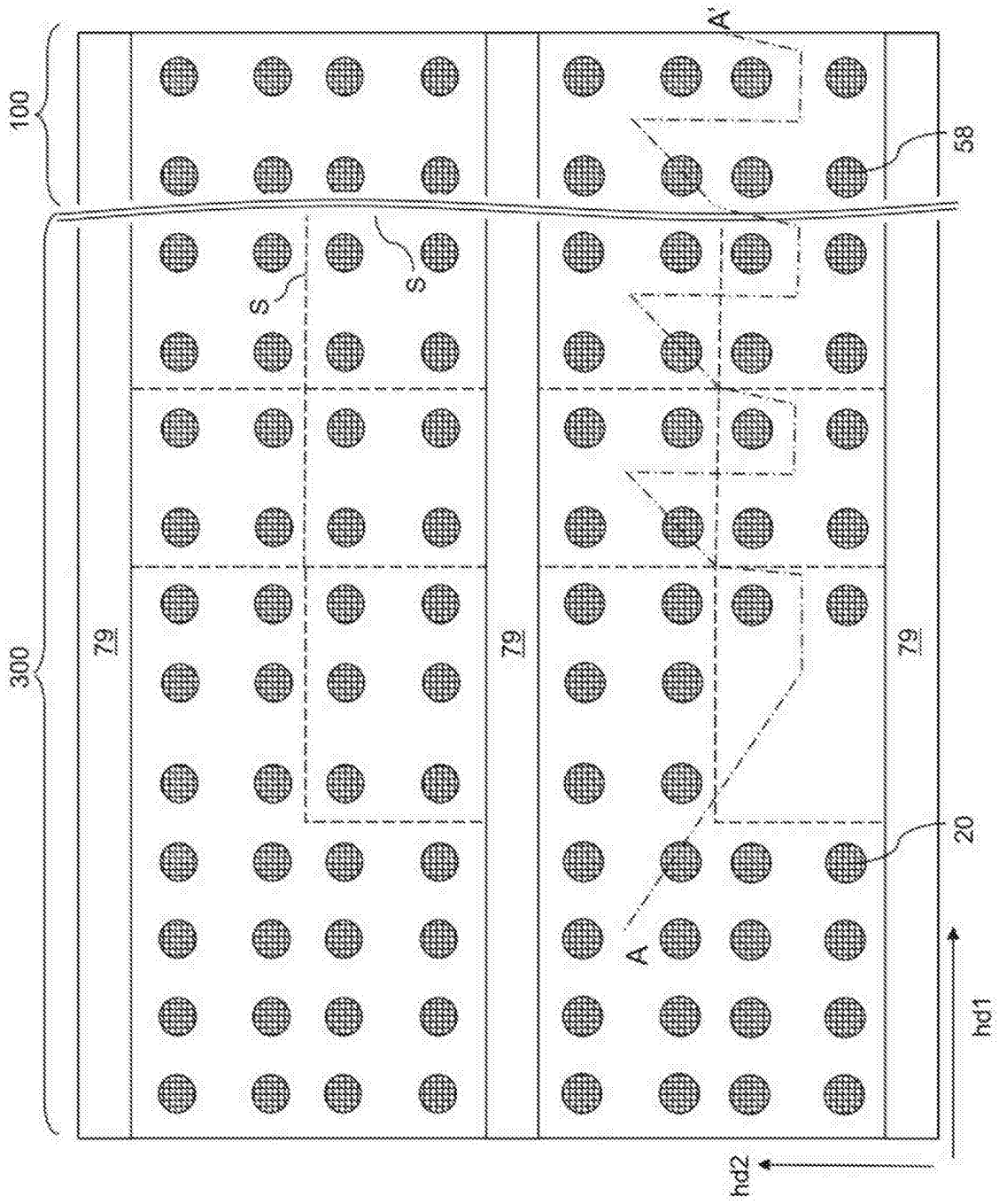


图7C

Micron Ex. 1002, p. 592
Micron v. YMTC
IPR2024-00789

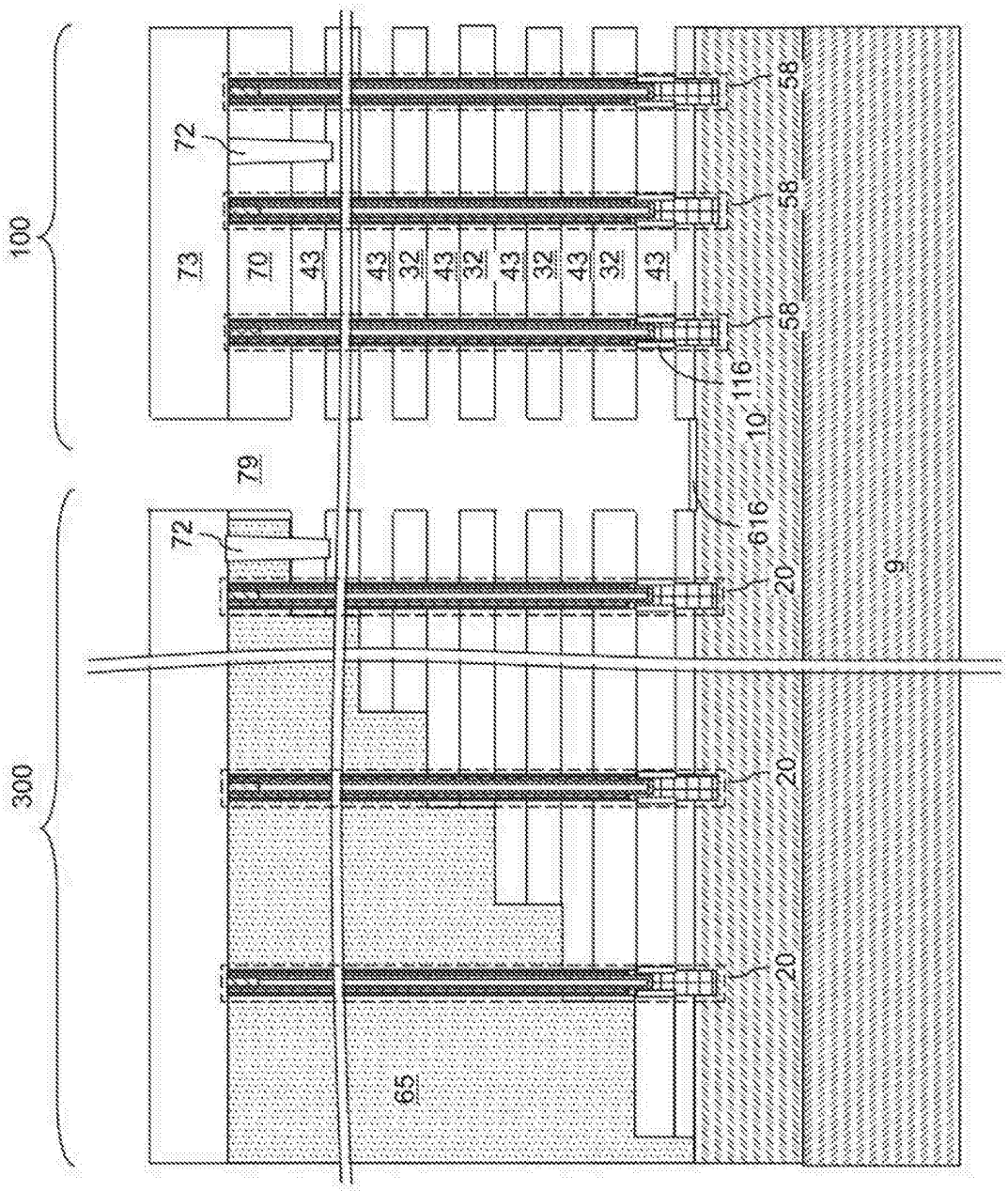


图8

Micron Ex. 1002, p. 593
Micron v. YMTC
IPR2024-00789

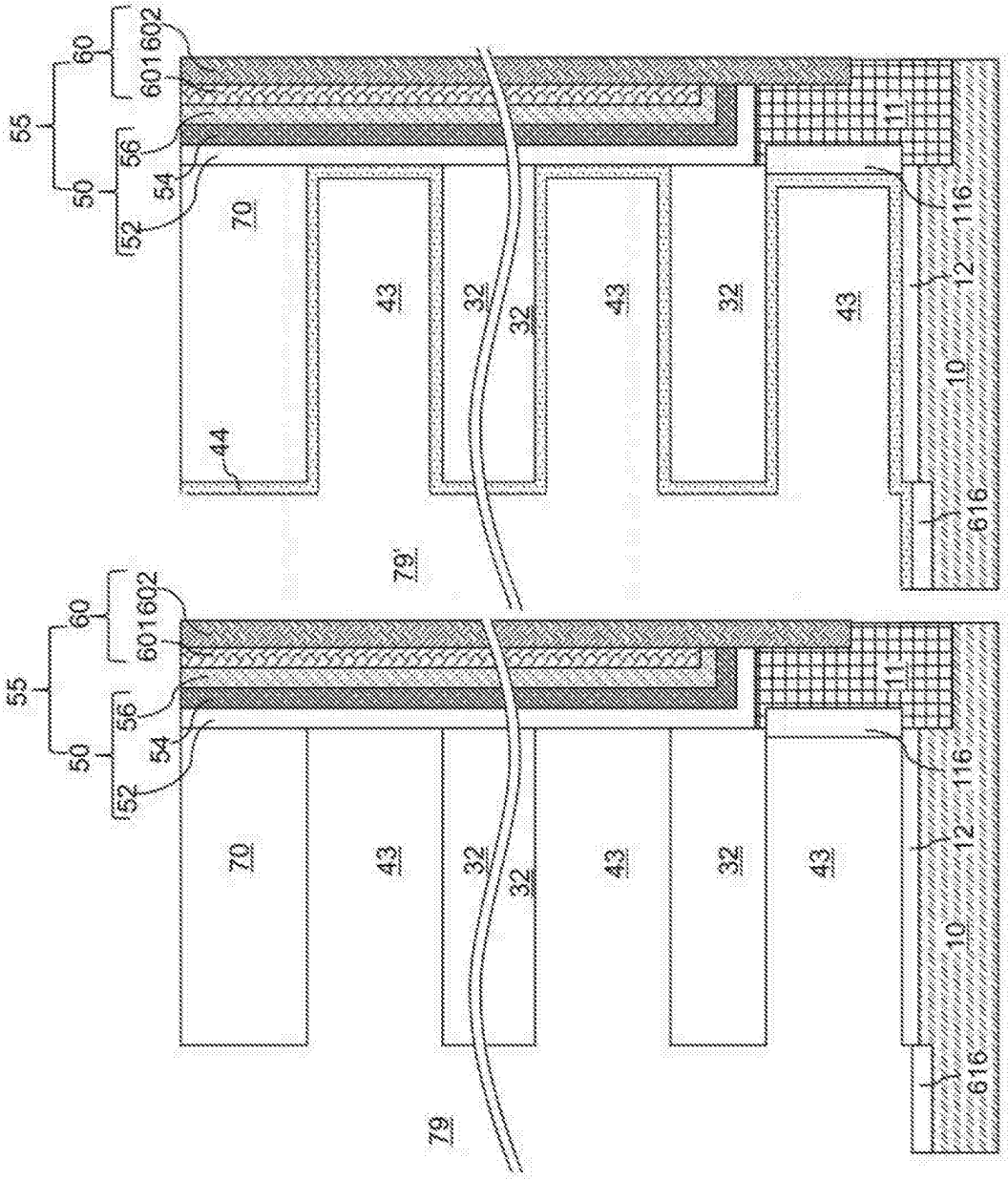


图 9B

图 9A

Micron Ex. 1002, p. 594
Micron v. YMTC
IPR2024-00789

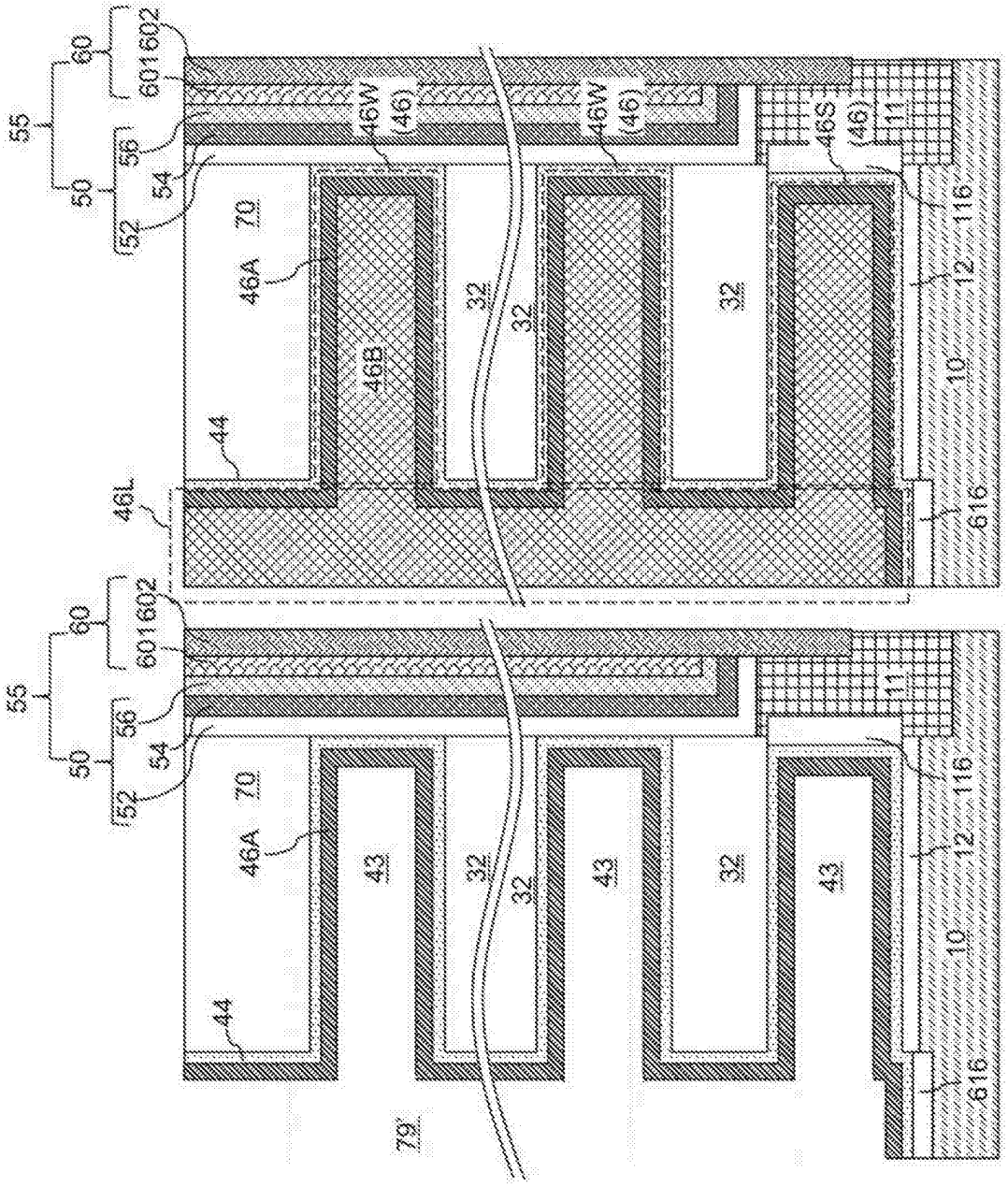


图 9C

图 9D

Micron Ex. 1002, p. 595
Micron v. YMTC
IPR2024-00789

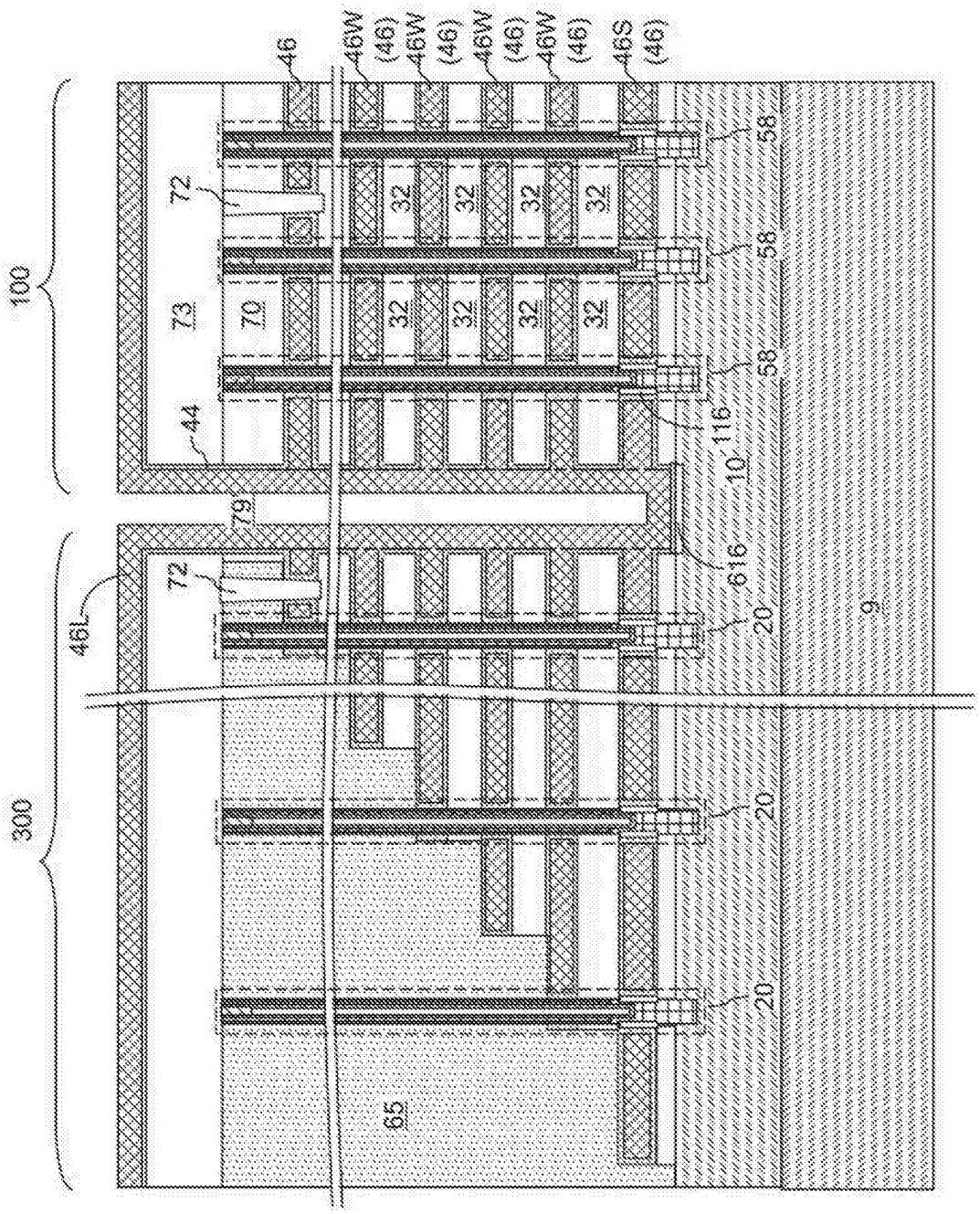


图10

Micron Ex. 1002, p. 596
Micron v. YMTC
IPR2024-00789

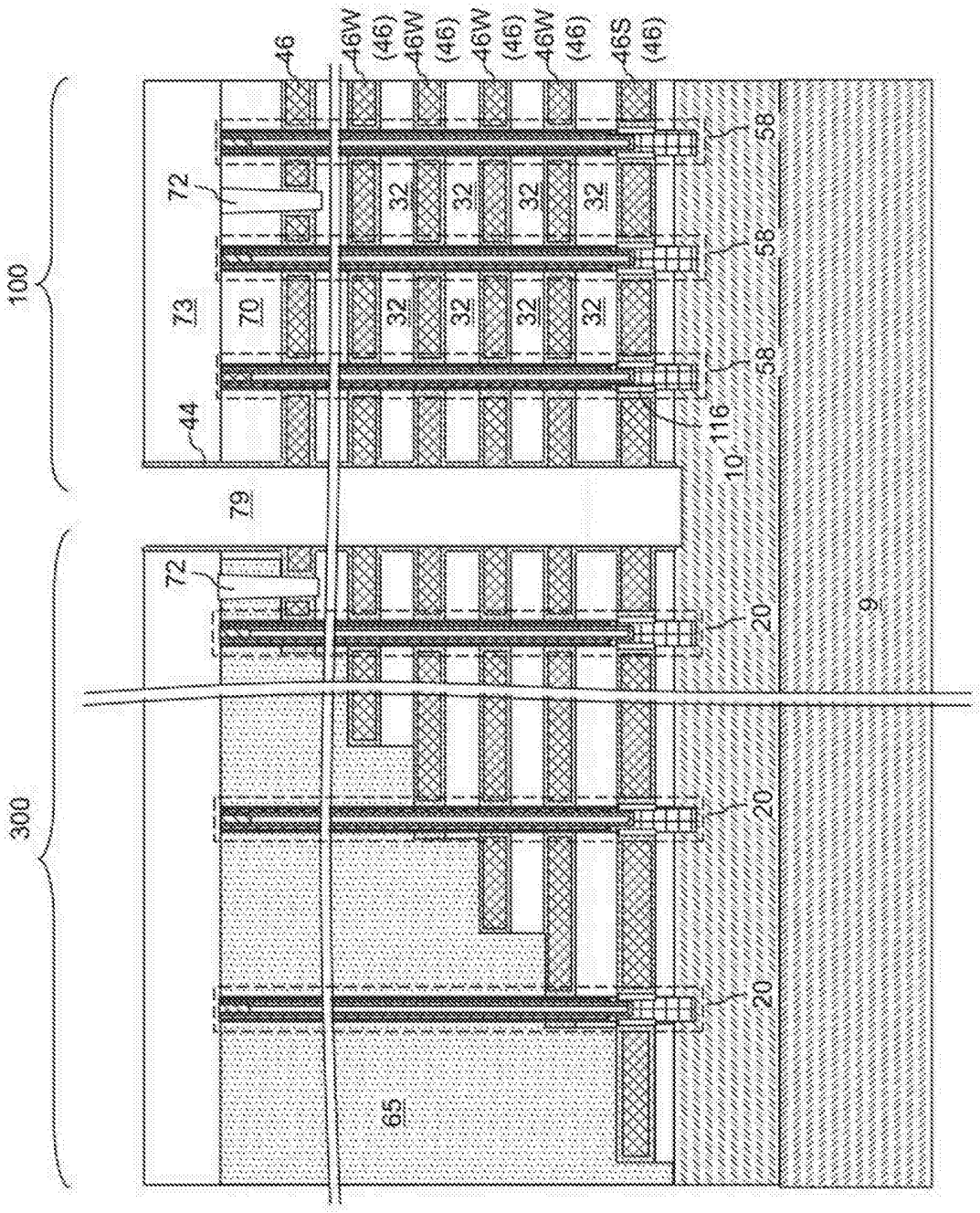


图11

Micron Ex. 1002, p. 597
Micron v. YMTC
IPR2024-00789

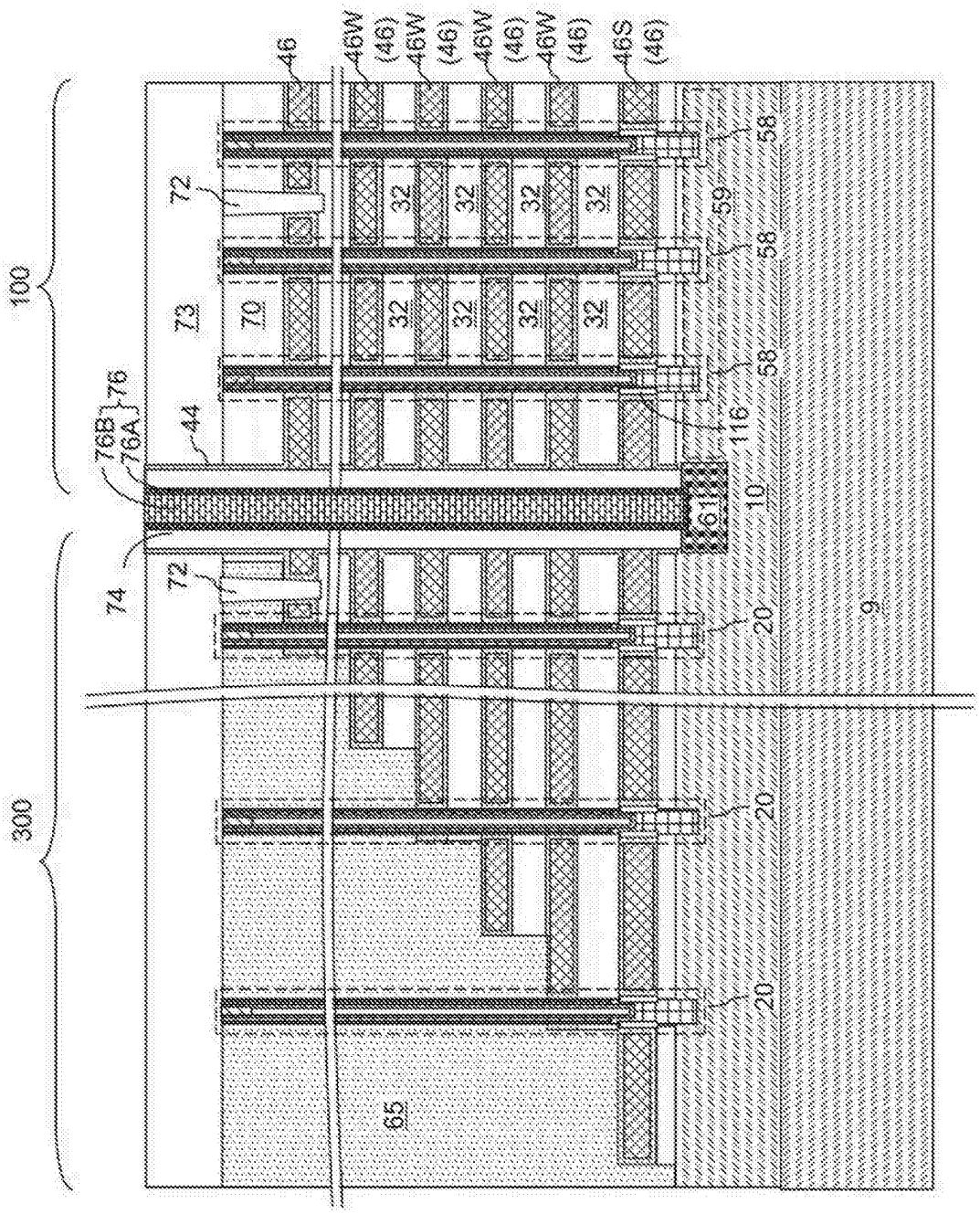


图12A

Micron Ex. 1002, p. 598
Micron v. YMTC
IPR2024-00789

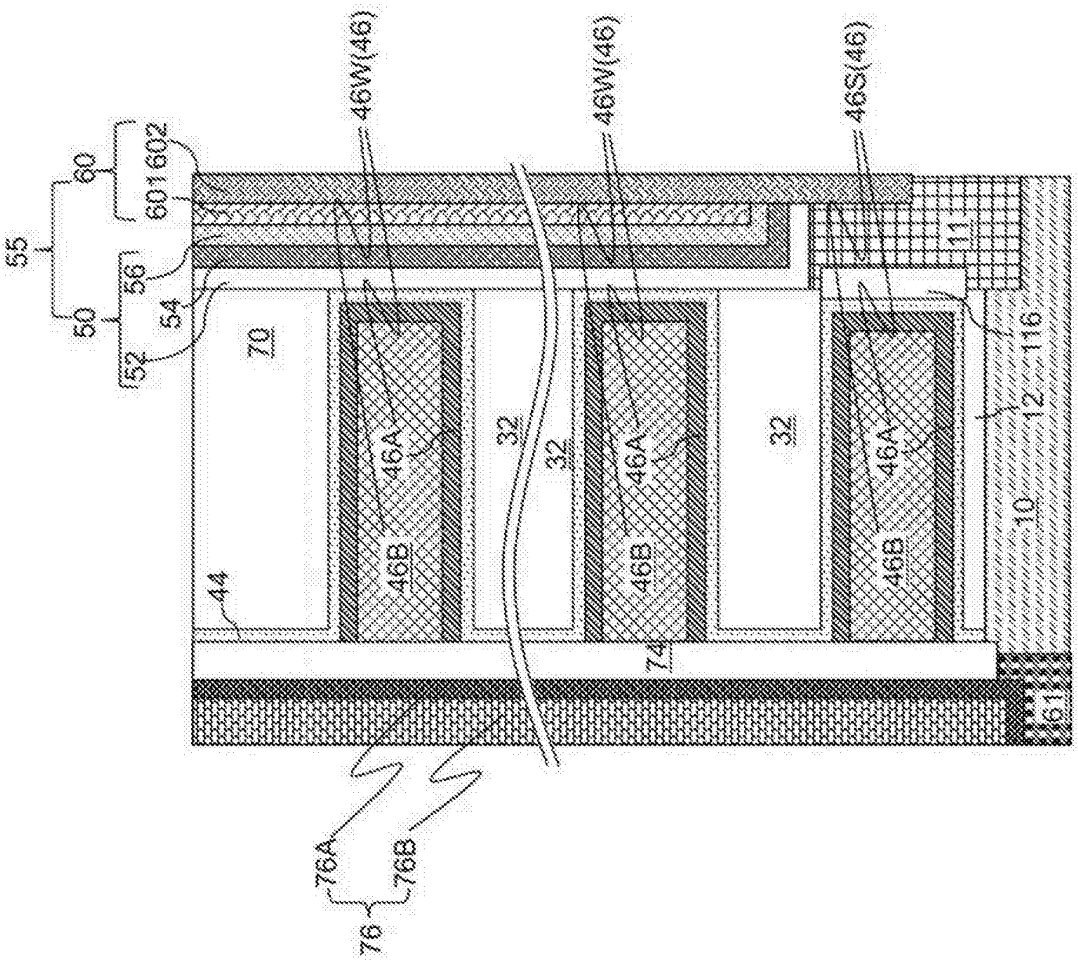


图12B

Micron Ex. 1002, p. 599
Micron v. YMTC
IPR2024-00789

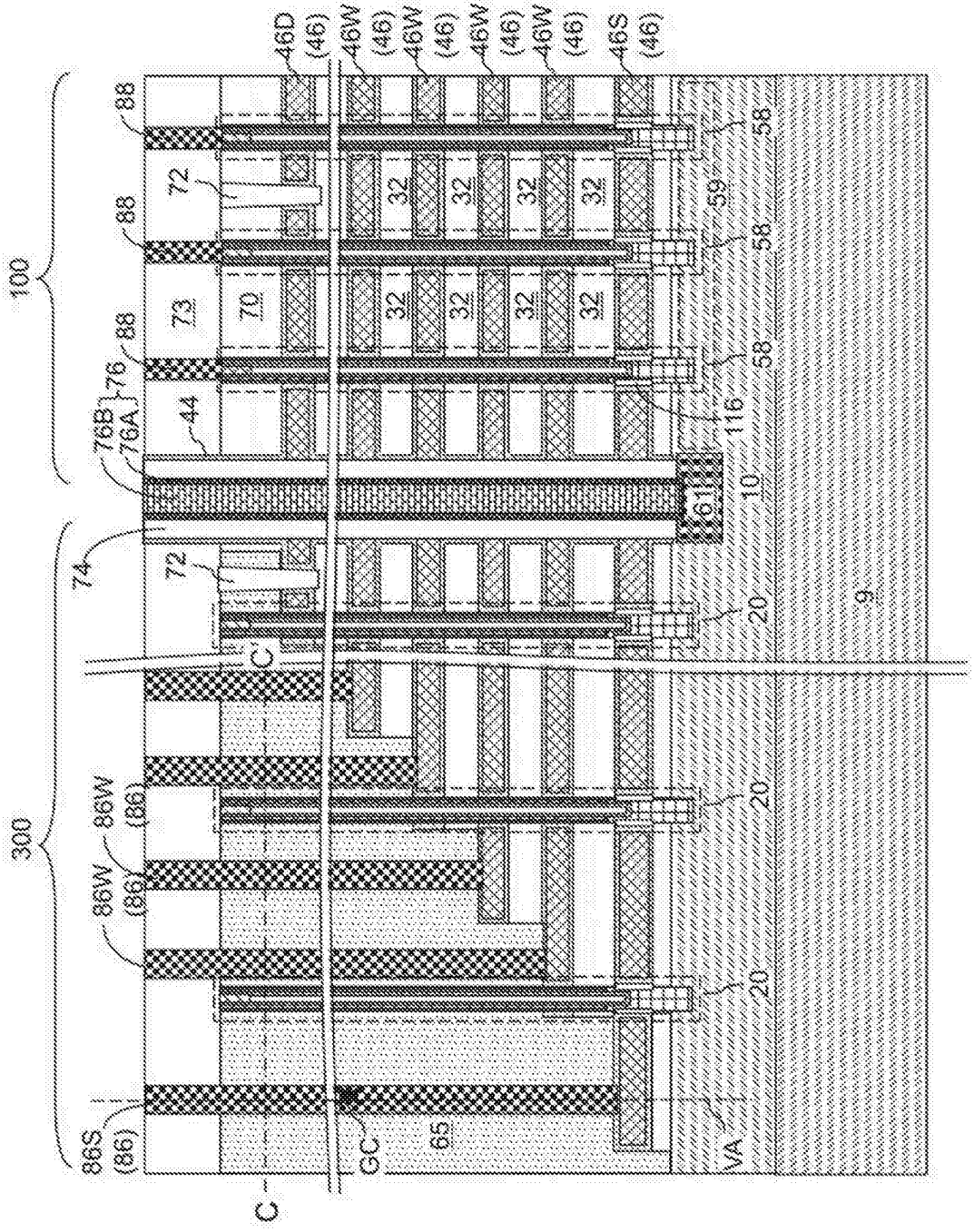


图13A

Micron Ex. 1002, p. 600
Micron v. YMTC
IPR2024-00789

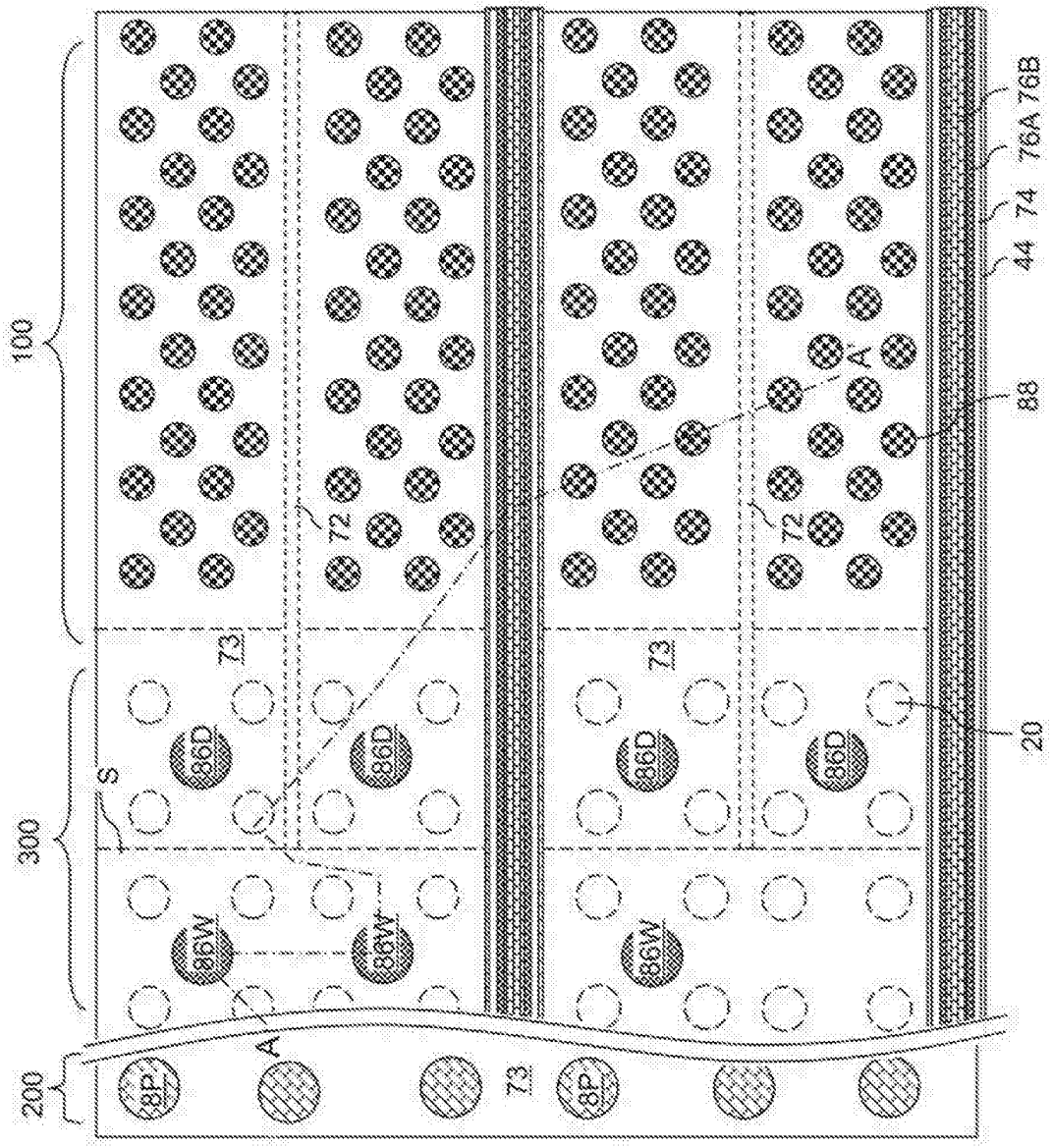


图13B

Micron Ex. 1002, p. 601
Micron v. YMTC
IPR2024-00789

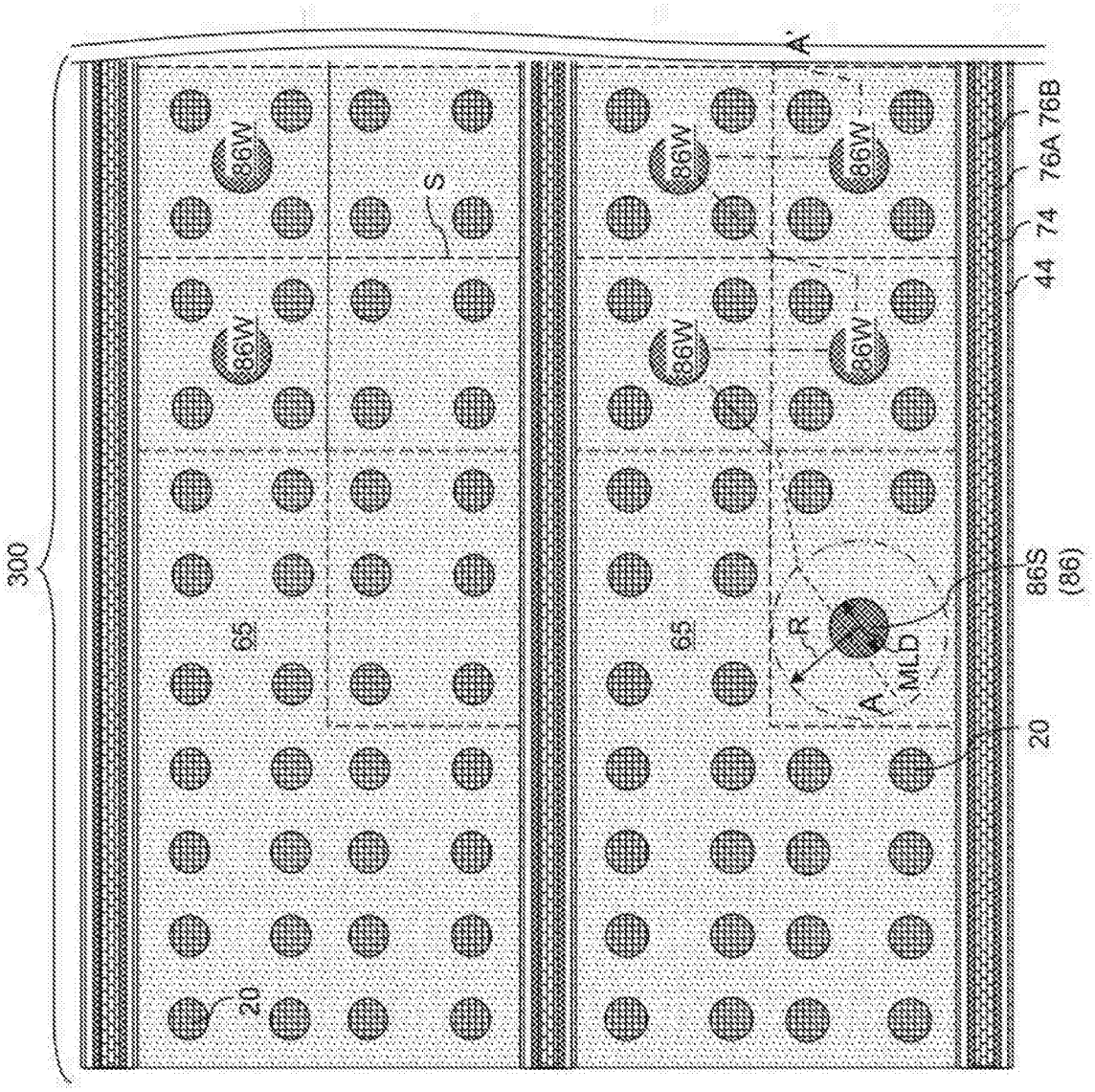


图13C

Micron Ex. 1002, p. 602
Micron v. YMTC
IPR2024-00789

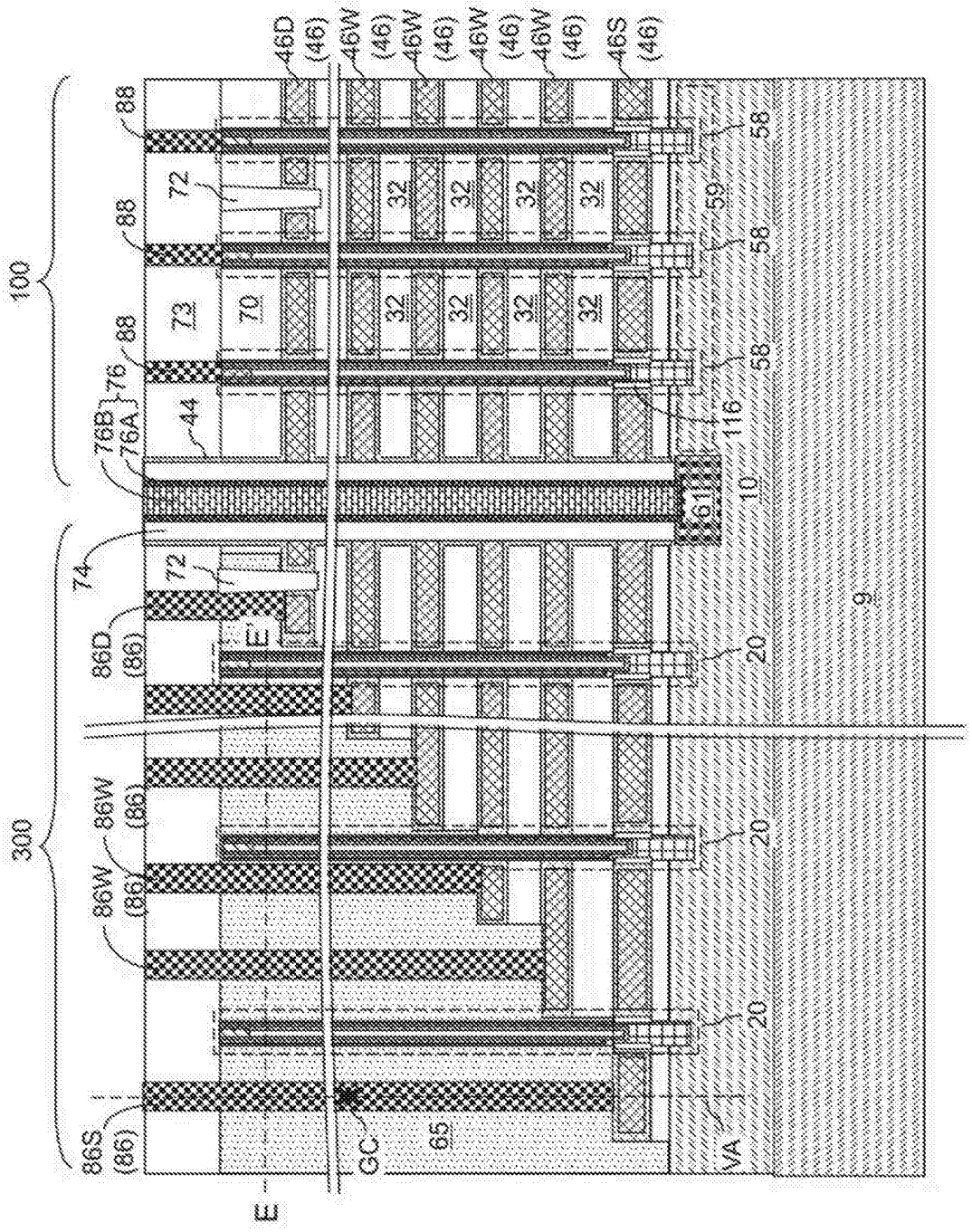


图13D

Micron Ex. 1002, p. 603
Micron v. YMTC
IPR2024-00789

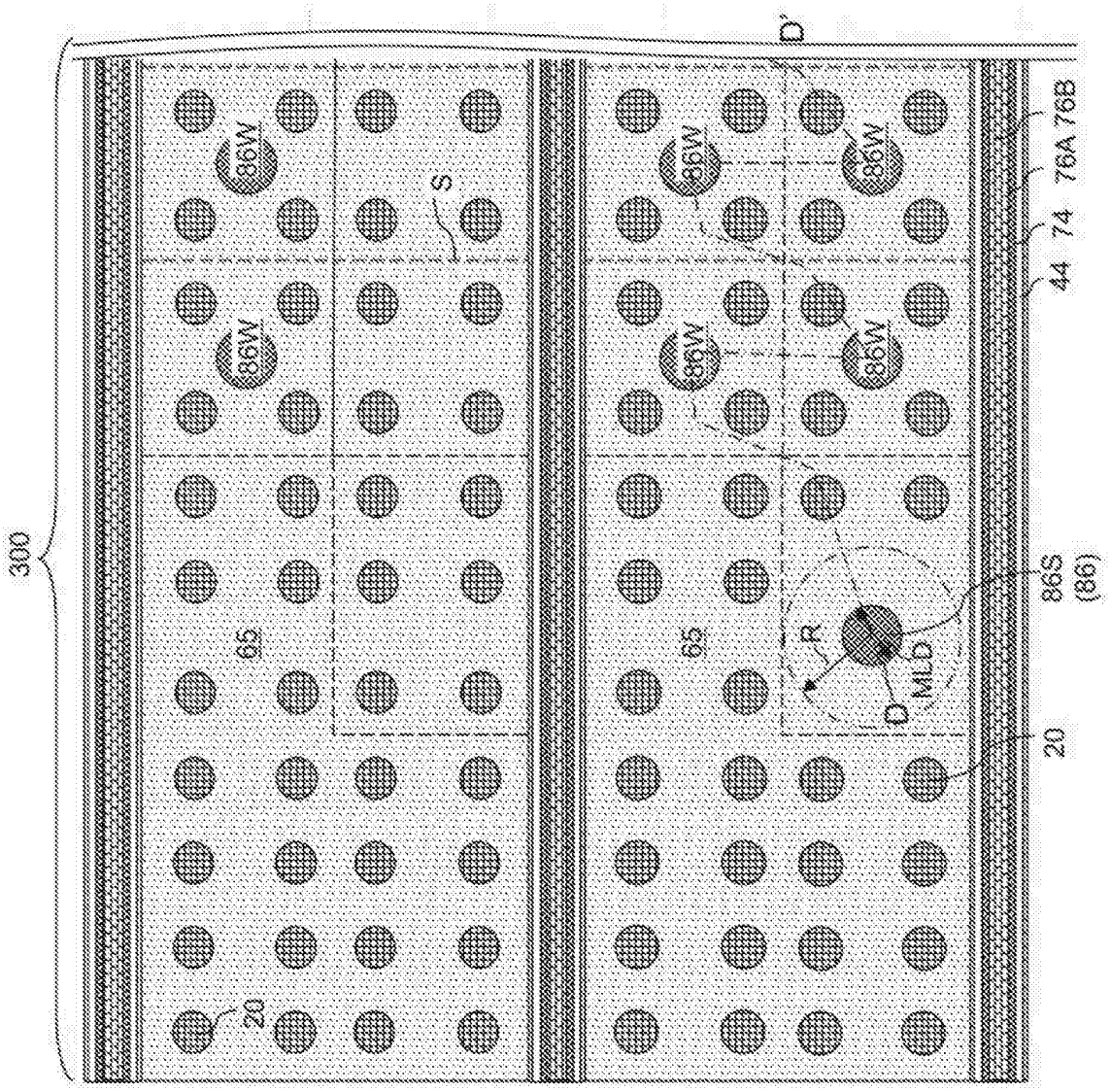


图13E

Micron Ex. 1002, p. 604
Micron v. YMTC
IPR2024-00789

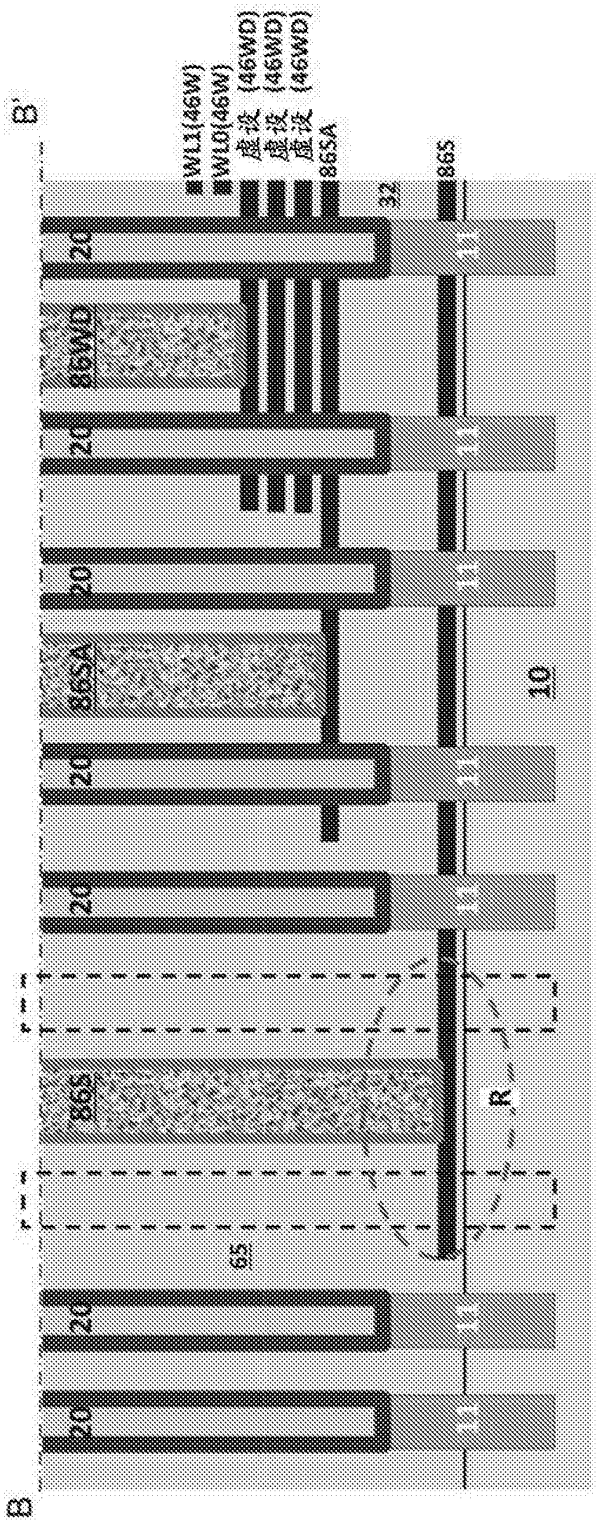


图14A

Micron Ex. 1002, p. 605
Micron v. YMTC
IPR2024-00789

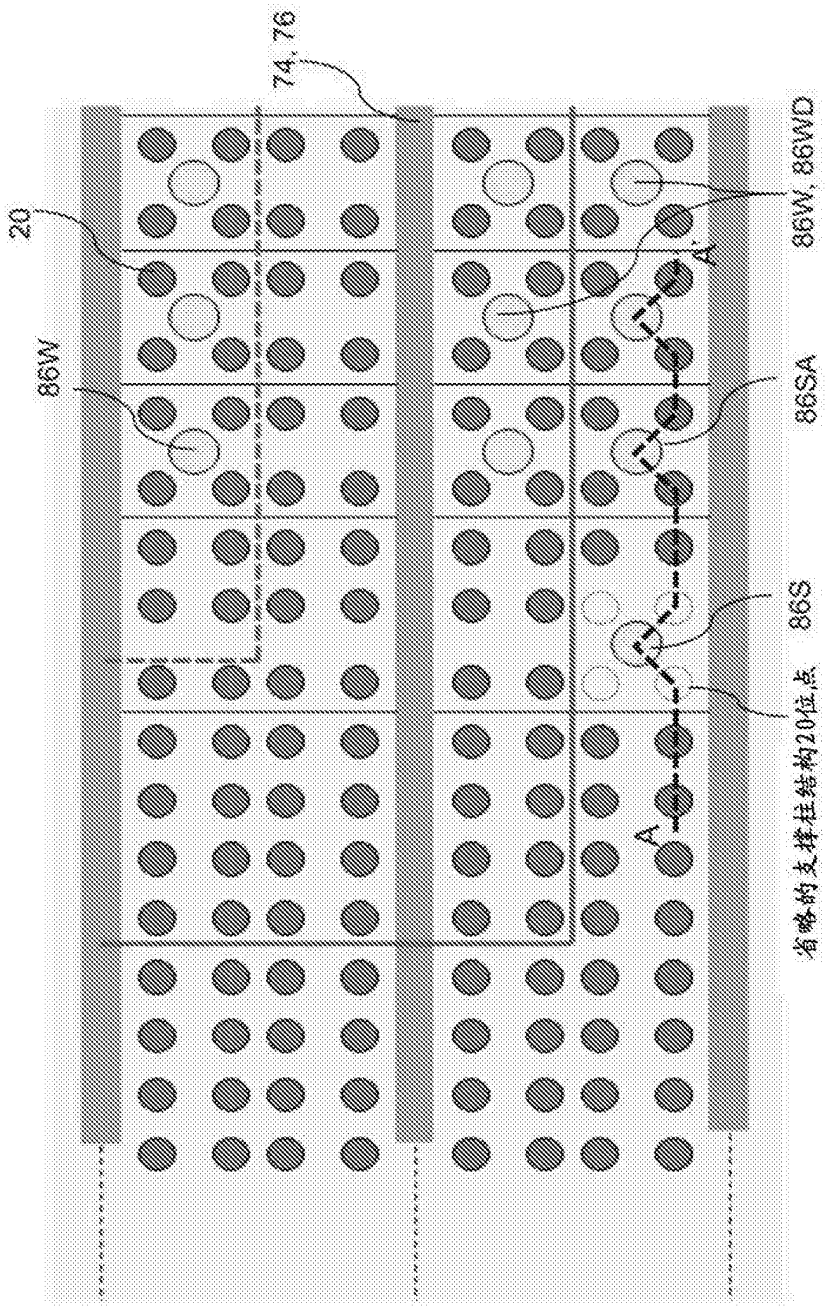


图14B

Micron Ex. 1002, p. 606
Micron v. YMTC
IPR2024-00789

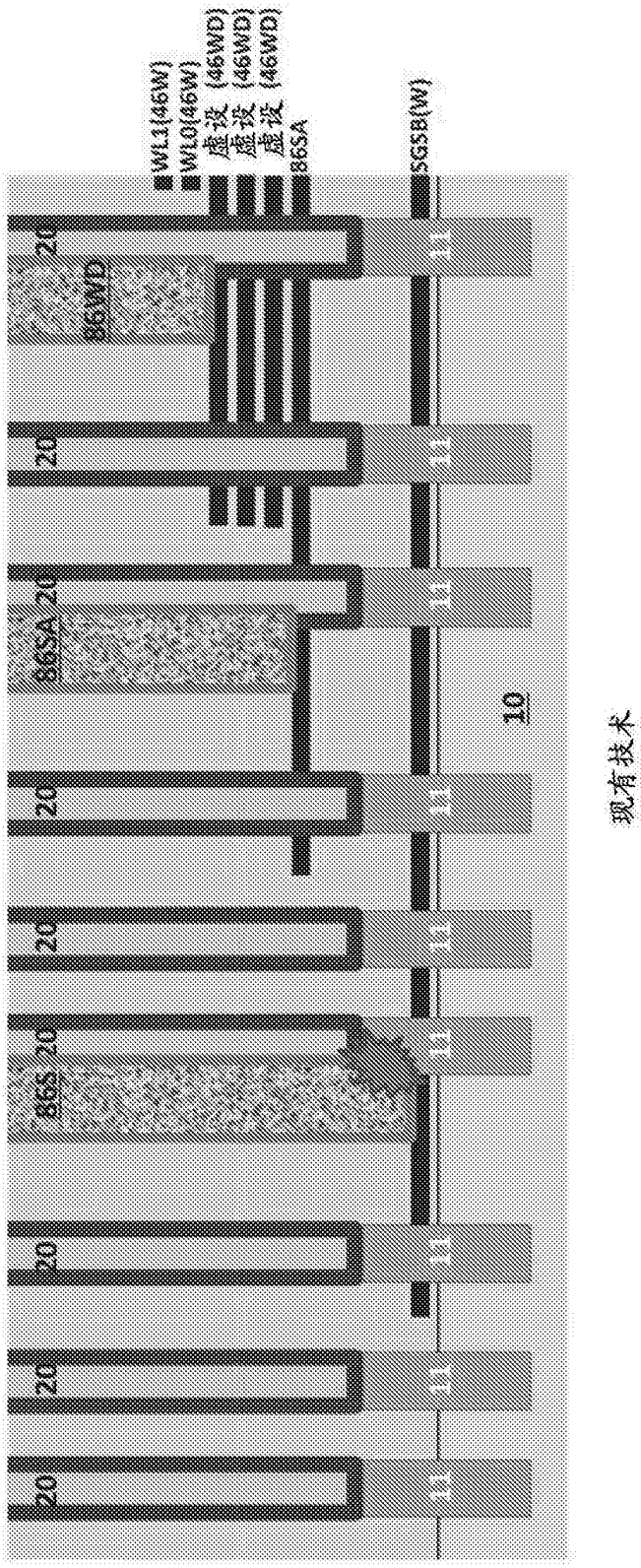


图15

Micron Ex. 1002, p. 607
Micron v. YMTC
IPR2024-00789

Espacenet

CN104157654A Three-dimensional memory and manufacturing method thereof

Applicants: INST OF MICROELECTRONICS CAS

Inventors: HUO ZONGLIANG,JIN LEI,LIU MING

Classifications:
IPC ***H01L21/8247; H01L27/115;***

CPC **H01L27/11578 (US); H01L27/11582 (EP,US); H01L28/00 (EP,US);**

Priorities: CN201410404550A 2014-08-15

Application: CN104157654A·2014-08-15

Publication: CN104157654A·2014-11-19

Published as: CN104157654A;**CN104157654B; US10475807B2; US2017236836A1; WO2016023260A1**

Three-dimensional memory and manufacturing method thereof

Abstract

The invention discloses a three-dimensional memory manufacturing method. The method comprises steps: a stacking structure of a first material layer and a second material layer is formed on a substrate; the stacking structure is etched to expose the substrate, and multiple vertical first open holes are formed; a filling layer is formed in each first open hole; the stacking structure is etched to expose the substrate round each first open hole, and multiple vertical second open holes are formed, a vertical channel layer and a drain are formed in each second open hole; filling layers are selectively etched and removed to expose the first open holes again; part of the second material layer is laterally etched or the second material layer is fully removed, and a groove is left; a gate stacking structure is formed in the groove; and a common source is formed on/in the substrate of the bottom part of each first open hole. According to the three-dimensional memory manufacturing method of the invention, a word line deep groove of a TCAT three-dimensional device is replaced by deep hole etching to achieve the same functions, the integration density is improved, the etching technology of the stacking structure is simplified, and control performance of the metal gate is remained.

Micron Ex. 1002, p. 608
Micron v. YMTC
IPR2024-00789

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104157654 A
(43)申请公布日 2014.11.19

(21)申请号 201410404550.X

(22)申请日 2014.08.15

(71)申请人 中国科学院微电子研究所
地址 100029 北京市朝阳区北土城西路 3#

(72)发明人 霍宗亮 刘明 靳磊

(74)专利代理机构 北京蓝智辉煌知识产权代理事务所（普通合伙）11345
代理人 陈红

(51)Int.Cl.
H01L 27/115(2006.01)
H01L 21/8247(2006.01)

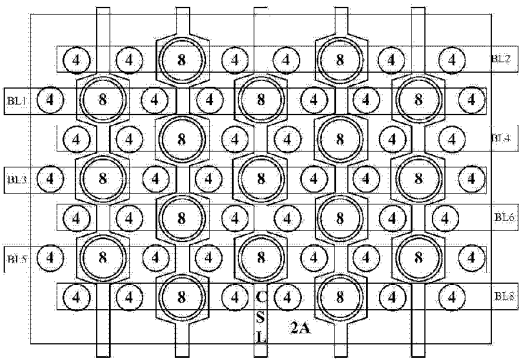
权利要求书2页 说明书7页 附图6页

(54)发明名称

三维存储器及其制造方法

(57)摘要

一种三维存储器制造方法，包括步骤：在衬底上形成第一材料层与第二材料层的堆叠结构；刻蚀堆叠结构露出衬底，形成垂直的多个第一开孔；在每个第一开孔中形成填充层；在每个第一开孔周围，刻蚀堆叠结构露出衬底，形成垂直的多个第二开孔；在每个第二开孔中形成垂直的沟道层和漏极；选择性刻蚀去除填充层，重新露出第一开孔；侧向刻蚀部分或者完全去除第二材料层，留下凹槽；在凹槽中形成栅极堆叠结构；在每个第一开孔底部的衬底上和/或中形成共源极。依照本发明的三维存储器制造方法，将 TCAT 三维器件的字线深槽替换为深孔刻蚀来完成相同的功能，提高集成密度，简化堆叠结构的刻蚀工艺，保留了金属栅控制性能。



CN 104157654 A

Micron Ex. 1002, p. 609
Micron v. YMTC
IPR2024-00789

1. 一种三维存储器制造方法，包括步骤：

在衬底上形成第一材料层与第二材料层的堆叠结构；

刻蚀堆叠结构露出衬底，形成垂直的多个第一开孔；

在每个第一开孔中形成填充层；

在每个第一开孔周围，刻蚀堆叠结构露出衬底，形成垂直的多个第二开孔；

在每个第二开孔中形成垂直的沟道层和漏极；

选择性刻蚀去除填充层，重新露出第一开孔；

侧向刻蚀部分或者完全去除第二材料层，留下凹槽；

在凹槽中形成栅极堆叠结构；

在每个第一开孔底部的衬底上和/或中形成共源极。

2. 如权利要求1所述的三维存储器制造方法，其中，第一材料层、第二材料层、填充层三者之间具有各自不同的刻蚀选择性。

3. 如权利要求2所述的三维存储器制造方法，其中，第一材料层、第二材料层、填充层材料选自氧化硅、氮化硅、氮氧化硅、非晶硅、非晶锗、DLC、非晶碳的任意一种及其组合。

4. 如权利要求1所述的三维存储器制造方法，其中，第一开孔的尺寸大于或等于第二开孔的尺寸。

5. 如权利要求1所述的三维存储器制造方法，其中，沟道层为中心包括绝缘体的中空结构。

6. 如权利要求1所述的三维存储器制造方法，其中，形成共源极之后进一步包括，在每个第一开孔侧壁形成绝缘层，在绝缘层侧壁以及每个第一开孔底部形成接触共源区的接触层。

7. 如权利要求6所述的三维存储器制造方法，其中，形成接触层之时或者之后进一步包括，控制沉积工艺参数或者回刻使得接触层的顶面低于栅极堆叠结构最顶层的底面，并且采用绝缘层回填。

8. 如权利要求6所述的三维存储器制造方法，其中，形成接触层之后进一步包括，刻蚀去除部分第一材料层、栅极堆叠结构形成第三开孔，在第三开孔中沉积绝缘材料形成与栅极堆叠结构最顶层之间的隔离绝缘区。

9. 如权利要求8所述的三维存储器制造方法，其中，形成隔离绝缘区之后进一步包括，在器件上形成层间介质层，刻蚀层间介质层形成第四开孔直至露出接触层，填充金属形成共源线引线。

10. 如权利要求9所述的三维存储器制造方法，其中，形成共源线接触塞之后进一步包括，在器件上形成第二层间介质层，刻蚀第二层间介质层形成第五开孔直至暴露沟道区，填充金属形成位线接触。

11. 如权利要求1所述的三维存储器制造方法，其中，刻蚀形成第一开孔之前进一步包括，在阵列区域周围的字线接触区域刻蚀堆叠结构形成台阶，依次暴露每个第一材料层和第二材料层的端部。

12. 一种三维存储器制造方法，包括步骤：

在衬底上形成第一材料层与第二材料层的堆叠结构；

刻蚀堆叠结构露出衬底，同时形成垂直的多个第一开孔以及在每个第一开孔周围的多

Micron Ex. 1002, p. 610
Micron v. YMTC
IPR2024-00789

个第二开孔；

在每个第一开孔中形成填充层；

在每个第二开孔中形成垂直的沟道层和漏极；

选择性刻蚀去除填充层，重新露出第一开孔；

侧向刻蚀部分或者完全去除第二材料层，留下凹槽；

在凹槽中形成栅极堆叠结构；

在每个第一开孔底部的衬底上和／或中形成共源极。

13. 一种三维存储器，包括：

在衬底上垂直分布的多个共源极引线，与衬底中和／或上的共源极接触；

在每个共源极引线周围垂直于衬底分布的多个沟道层，每个沟道层顶部具有漏极；

在每个沟道层侧壁上的多个绝缘隔离层以及在每个绝缘隔离层之间的多个栅极堆叠结构。

Micron Ex. 1002, p. 611
Micron v. YMTC
IPR2024-00789

三维存储器及其制造方法

技术领域

[0001] 本发明涉及一种半导体器件及其制造方法，特别是涉及一种高密度三维存储器及其制造方法。

背景技术

[0002] 为了改善存储器件的密度，业界已经广泛致力于研发减小二维布置的存储器单元的尺寸的方法。随着二维（2D）存储器件的存储器单元尺寸持续缩减，信号冲突和干扰会显著增大，以至于难以执行多电平单元（MLC）操作。为了克服 2D 存储器件的限制，业界已经研发了具有三维（3D）结构的存储器件，通过将存储器单元三维地布置在衬底之上来提高集成密度。

[0003] 业界目前一种常用的 3D 存储器件结构是太比特单元阵列晶体管（TCAT）。具体地，可以首先在衬底上沉积多层叠层结构（例如氧化物和氮化物交替的多个 ONO 结构）；通过各向异性的刻蚀工艺对衬底上多层叠层结构刻蚀而形成沿着存储器单元字线（WL）延伸方向分布、垂直于衬底表面的多个沟道通孔（可直达衬底表面或者具有一定过刻蚀）；在沟道通孔中沉积多晶硅等材料形成柱状沟道；沿着 WL 方向刻蚀多层叠层结构形成直达衬底的沟槽，露出包围在柱状沟道周围的多层叠层；针对叠层中相邻层之间的刻蚀选择性，选择刻蚀选择比较高的腐蚀液湿法去除叠层中的第二类型材料，在柱状沟道周围留下横向分布的第一类型材料构成的突起结构；在沟槽中突起结构的侧壁沉积高 k 介质材料的栅极介质层以及金属材质的栅极导电层构成的栅极堆叠；刻蚀叠层结构形成源漏接触并完成后端制造工艺。此时，叠层结构在柱状沟道侧壁留下的一部分突起形成了栅电极之间的隔离层，而留下的栅极堆叠夹设在多个隔离层之间作为控制电极。当向栅极施加电压时，栅极的边缘电场会使得例如多晶硅材料的柱状沟道侧壁上感应形成源漏区，由此构成多个串并联的闪存单元构成的门阵列而记录所存储的逻辑状态。其中，为了将单元区多个串并联 MOSFET 信号引出，在柱状沟道顶部沉积填充多晶硅材料形成漏区，并形成与漏区电连接的金属接触塞以进一步电连接至上方的位线（bit-line，BL）。此外，在多个垂直柱状沟道之间衬底中形成带有金属硅化物接触的共用源区。在单元导通状态下，电流从共用源区流向周围的垂直沟道区，并在控制栅极（与字线 WL 相连）施加的控制电压作用下向上穿过垂直沟道中感应生成的多个源漏区，通过沟道顶部的漏区而进一步流向上方的位线。

[0004] 该 TCAT 器件结构具有体擦除（改变控制栅极可以引起感应源漏区以及浮栅极中电势变化，能整体擦除）、金属栅极（能较方便通过控制金属材料控制功函数从而调节晶体管阈值）等诸多优点。但是另一方面，由于除了顶部的选择晶体管（USG，位于存储晶体管单元串上方）之外，其余栅极与字线（WL）连接都是通过刻蚀孔进行共享链接，并且采用后栅工艺刻蚀去除伪栅极形成栅极开口并沉积金属栅极，这种极高深宽比（AR 例如通常大于 40:1 乃至 100:1）的深接触孔以及栅极开口将由于沉积多层薄膜而使得宽度增大，进而使得 TCAT 的存储单元密度无法有效进一步减小。同时，深槽的刻蚀和深孔沟道的刻蚀都是对多层堆栈的刻蚀，工艺复杂度很高，深孔和深槽的形状差异使得刻蚀的工艺也要有所变化。

Micron Ex. 1002, p. 612
Micron v. YMTC
IPR2024-00789

[0005] 与 TCAT 技术对应的，另一种常用的器件结构例如是采用位成本可缩减 (BiCS) 的 NAND 结构，通过将存储器单元三维地布置在衬底之上来提高集成密度、其中沟道层垂直竖立在衬底上，栅极分为下层的选择栅极、中层的控制栅极以及上层的选择栅极三部分，通过将栅极信号分布在三组栅电极中以减小信号之间的串扰。具体地，上层和下层的器件用作选择晶体管——栅极高度 / 厚度较大的垂直 MOSFET，栅极介质层为常规的单层高 k 材料；中层的器件用作存储单元串，栅极高度 / 厚度较小，栅极介质层为隧穿层、存储层、阻挡层的堆叠结构。基于 BiCS 的 NAND 结构器件的具体制造工艺一般包括，在硅衬底上沉积下层选择栅电极层，刻蚀下层选择栅电极层形成直达衬底的孔槽以沉积沟道层的下部分以及下层栅电极的引出接触，在上方沉积控制栅极层，刻蚀控制栅极层形成作为存储器单元区域的中间沟道区以及中层控制栅电极的引出接触，刻蚀形成控制栅极，按照字线、位线划分需要将整个器件分割为多个区域，在之上沉积上层选择 栅极并刻蚀、沉积形成上部沟道以及上层引出接触，之后采用后续工艺完成器件的制造。在 BiCS 结构中，除了最上方的选择晶体管 USG 之外，下方的所有栅电极都可以是平板状，相对于 TCAT 而言可以避免深沟槽、深孔接触工艺，有利于提高存储器密度。在这种工艺过程中，最为关键的刻蚀步骤仅在于对于中间层存储器沟道区和引出接触的光刻，这直接决定了整个器件的集成度以及信号抗干扰能力。然而，BiCS 结构虽然通过存储阵列与选择晶体管堆叠放置而分别利用控制栅极阈值，并且通过分层连接栅极避免了过大深宽比的接触孔或栅极开口的复杂工艺，但是只能通过多晶硅材质的栅极诱导漏极泄漏电流 (GIDL) 进行擦除，无法进行体擦除，读写效率较低。

发明内容

[0006] 由上所述，本发明的目的在于克服上述技术困难，一种能以较低成本实现接触互连并且保持金属栅控制性能的新型三维存储器结构的制造方法。

[0007] 为此，本发明一方面提供了一种三维存储器制造方法，包括步骤：在衬底上形成第一材料层与第二材料层的堆叠结构；刻蚀堆叠结构露出衬底，形成垂直的多个第一开孔；在每个第一开孔中形成填充层；在每个第一开孔周围，刻蚀堆叠结构露出衬底，形成垂直的多个第二开孔；在每个第二开孔中形成垂直的沟道层和漏极；选择性刻蚀去除填充层，重新露出第一开孔；侧向刻蚀部分或者完全去除第二材料层，留下凹槽；在凹槽中形成栅极堆叠结构；在每个第一开孔底部的衬底上和 / 或中形成共源极。

[0008] 其中，第一材料层、第二材料层、填充层三者之间具有各自不同的刻蚀选择性。

[0009] 其中，第一材料层、第二材料层、填充层材料选自氧化硅、氮化硅、氮氧化硅、非晶硅、非晶锗、DLC、非晶碳的任意一种及其组合。

[0010] 其中，第一开孔的尺寸大于或等于第二开孔的尺寸。

[0011] 其中，沟道层为中心包括绝缘体的中空结构。

[0012] 其中，形成共源极之后进一步包括，在每个第一开孔侧壁形成绝缘层，在绝缘层侧壁以及每个第一开孔底部形成接触共源区的接触层。

[0013] 其中，形成接触层之时或者之后进一步包括，控制沉积工艺参数或者回刻使得接触层的顶面低于栅极堆叠结构最顶层的底面，并且采用绝缘层回填。

[0014] 其中，形成接触层之后进一步包括，刻蚀去除部分第一材料层、栅极堆叠结构形成第三开孔，在第三开孔中沉积绝缘材料形成与栅极堆叠结构最顶层之间的隔离绝缘区。

Micron Ex. 1002, p. 613
Micron v. YMTC
IPR2024-00789

[0015] 其中，形成隔离绝缘区之后进一步包括，在器件上形成层间介质层，刻蚀层间介质层形成第四开孔直至露出接触层，填充金属形成共源线引线。

[0016] 其中，形成共源线接触塞之后进一步包括，在器件上形成第二层间介质层，刻蚀第二层间介质层形成第五开孔直至暴露沟道区，填充金属形成位线接触。

[0017] 其中，刻蚀形成第一开孔之前进一步包括，在阵列区域周围的字线接触区域刻蚀堆叠结构形成台阶，依次暴露每个第一材料层和第二材料层的端部。

[0018] 本发明另一方面提供了一种三维存储器制造方法，包括步骤：在衬底上形成第一材料层与第二材料层的堆叠结构；刻蚀堆叠结构露出衬底，同时形成垂直的多个第一开孔以及在每个第一开孔周围的多个第二开孔；在每个第一开孔中形成填充层；在每个第二开孔中形成垂直的沟道层和漏极；选择性刻蚀去除填充层，重新露出第一开孔；侧向刻蚀部分或者完全去除第二材料层，留下凹槽；在凹槽中形成栅极堆叠结构；在每个第一开孔底部的衬底上和／或中形成共源极。

[0019] 本发明又一方面提供了一种三维存储器，包括：在衬底上垂直分布的多个共源极引线，与衬底中和／或上的共源极接触；在每个共源极引线周围垂直于衬底分布的多个沟道层，每个沟道层顶部具有漏极；在每个沟道层侧壁上的多个绝缘隔离层以及在每个绝缘隔离层之间的多个栅极堆叠结构。

[0020] 依照本发明的三维存储器制造方法，将 TCAT 三维器件的字线深槽替换为深孔刻蚀来完成相同的功能，提高集成密度，简化堆叠结构的刻蚀工艺，保留了金属栅控制性能。

附图说明

[0021] 以下参照附图来详细说明本发明的技术方案，其中：

[0022] 图 1 至图 10 为依照本发明的三维半导体存储器件制造方法的各个步骤的剖视图和／或顶视图。

具体实施方式

[0023] 以下参照附图并结合示意性的实施例来详细说明本发明技术方案的特征及其技术效果，公开了能以较低成本实现接触互连并且保持金属栅控制性能的新型三维存储器结构的制造方法。需要指出的是，类似的附图标记表示类似的结构，本申请中所用的术语“第一”、“第二”、“上”、“下”等等可用于修饰各种器件结构或制造工序。这些修饰除非特别说明并非暗示所修饰器件结构或制造工序的空间、次序或层级关系。

[0024] 如图 1 所示，在衬底 1 上交替形成第一材料层 2A 与第二材料层 2B 的堆叠结构 2。衬底 1 材质可以包括体硅 (bulk Si)、体锗 (bulkGe)、绝缘体上硅 (SOI)、绝缘体上锗 (GeOI) 或者是其他化合物半导体衬底，例如 SiGe、SiC、GaN、GaAs、InP 等等，以及这些物质的组合。为了与现有的 IC 制造工艺兼容，衬底 1 优选地为含硅材质的衬底，例如 Si、SOI、SiGe、Si:C 等。堆叠结构 2 的选自以下材料的组合并且至少包括一种绝缘介质：如氧化硅、氮化硅、非晶碳、类金刚石无定形碳 (DLC)、氧化锗、氧化铝、等及其组合。第一材料层 2A 具有第一刻蚀选择性，第二材料层 2B 具有第二刻蚀选择性并且不同于第一刻蚀选择性（例如两种材料之间的刻蚀选择比大于 5:1 并优选大于 10:1）。在本发明一个优选实施例中，叠层结构 2A/2B 均为非导电材料，层 2A/ 层 2B 的组合例如氧化硅与氮化硅的组合、氧化硅与（未掺杂）多晶

Micron Ex. 1002, p. 614
Micron v. YMTC
IPR2024-00789

硅或非晶硅的组合、氧化硅或氮化硅与非晶碳的组合等等。在本发明另一优选实施例中，层 2A 与层 2B 在湿法腐蚀条件或者在氧等离子干法刻蚀条件下具有较大的刻蚀选择比（例如大于 5:1）。层 2A、层 2B 的沉积方法包括 PECVD、LPCVD、HDPCVD、MOCVD、MBE、ALD、热氧化、蒸发、溅射等各种工艺。在本发明一个最优实施例中，层 2A 为二氧化硅，层 2B 为氮化硅。如图 1 所示，最低部的层 2B 将用于未来最底层选择栅电极形成，因此标注为 LSG，最顶部的层 2B 将用于未来最顶层选择栅电极形成，因此标注为 USG，中部具有从 WL0 ～ WLn 的多个层 2B 以用于 n+1 条位线。图 1 中仅示出了 n ＝ 2 的情形，但是实际上 n 可以依照需要 为大于等于 1 的任何正整数。

[0025] 任选的，如图 2 所包括的剖视图 2A 和顶视图 2B 所示，刻蚀堆叠结构 2A/2B，依次露出一部分层 2A/2B 的侧面端部以形成台阶状的字线焊垫区域 (WL-Pad Region,WLPR)，中心区域为阵列区域 (Array Region，AR)。其中，在 WLPR 区域，依次暴露了顶层 USG 上方的层 2A:USG、用于 WL2 的层 2B 上方的层 2A:WL2、用于 WL1 的层 2B 上方的层 2A:WL1、用于 WL0 的层 2B 上方的层 2A:WL0、用于 LSG 的层 2B 上方的层 2A:LSG、以及保护衬底 1 的层 2A。以下图 3 至图 10 的步骤将针对图 2 中的中心的阵列区域进行，换言之也即各个顶视图仅示出了区域 2A:AR 中的一部分。

[0026] 在现有技术的 TCAT 结构制造方法中，一般是刻蚀深孔并沉积多晶硅等沟道材料形成沟道后，通过执行深槽刻蚀暴露出作为层 2B 的伪栅 (SiN)，之后移去层 2B 的伪栅并完成栅堆栈介质层和金属电极沉积等步骤完成存储单元的制备；随后通过侧墙、注入和硅化来完成共源区形成；接着通过深槽的氧化物填充、CMP，共源区 CSL 接触孔刻蚀和填充，金属线连线来完成 CSL 的连线。如背景技术部分所述，这种后形成的深槽往往难以兼顾深宽比以及金属填充率，容易导致器件失效。

[0027] 与之对照的，本发明的一个实例不用深槽而用如图 3 所示的深孔刻蚀，这样工艺和沟道形成工艺相同，在完成存储单元的制备之后，直接通过侧墙、硅化和金属的深孔填充完成共源区的接触，然后引线完成 CSL 连线。相对而言，多层介质刻蚀工艺简单，存储阵列的密度提高，共源区的接触孔引出工艺更加简单，原有的二氧化硅填充深槽并平坦化、刻蚀接触孔、填充接触孔的方式直接被深孔的侧墙绝缘层和金属接触孔填充所替代

[0028] 如包括剖视图 3A、顶视图 3B 的图 3 所示，在阵列区域刻蚀堆叠结构 2 直至暴露衬底 1（该区域将形成未来的共源区），形成作为伪栅极开孔的第一开孔（未示出）并在伪栅极开孔中形成填充层 3。在稍后图 5 所示的步骤中，第一开孔将作为形成共源区 1S 以及侧向刻蚀去除部分层 2B 所用的入口区域。优选地，采用 RIE 或等离子干法刻蚀各向异性刻蚀层 2A/ 层 2B 的堆叠结构 2，形成露出衬底 1 以及衬底 1 上交替堆叠的层 2A/ 层 2B 的侧壁的第一开孔（未示出）。刻蚀气体例如针对二氧化硅和氮化硅等材质的碳氟基刻蚀气体，并且通过增加碳氟比而在侧壁形成由含 C 聚合物形成的临时保护侧壁，最终获得较好的垂直侧壁。在本发明一个优选实施例中刻蚀气体优选 C_3F_6、 C_4F_8 等含 C 量比较高的气体并进一步优选通过增加氧化性气体如 O_2、CO 等控制侧壁形貌。平行于衬底 1 表面切得的孔槽的截面形状可以为矩形、方形、菱形、圆形、半圆形、椭圆形、三角形、五边形、五角形、六边形、八边形等等各种几何形状。填充层 3 沉积方法包括 PECVD、HDPCVD、MOCVD、MBE、ALD、蒸发、溅射等，材质优选为与堆叠结构 2 的层 2A、层 2B 均具有高选择性的材料，例如层 3、层 2A、层 2B 三者之间每两个之间的刻蚀选择比均大于等于 5 :1。在本发明一个优选实施例中，层 2A 为

Micron Ex. 1002, p. 615
Micron v. YMTC
IPR2024-00789

氧化硅，层 2B 为氮化硅，填充层为非晶硅、非晶锗、非晶碳、DLC 等，反之亦然。

[0029] 接着，如包括剖视图 4A、顶视图 4B 的图 4 所示，在每个第一开孔周围刻蚀形成多个第二开孔并形成垂直的沟道区 4。与刻蚀形成第一开孔的工艺类似，RIE 或等离子干法刻蚀各向异性刻蚀层 2A/ 层 2B 的堆叠结构 2，在第一开孔的周围形成多个露出衬底 1 以及衬底 1 上交替堆叠的层 2A/ 层 2B 的侧壁的第二开孔（未示出）。刻蚀气体例如针对二氧化硅和氮化硅等材质的碳氟基刻蚀气体，并且通过增加碳氟比而在侧壁形成由含 C 聚合物形成的临时保护侧壁，最终获得较好的垂直侧壁。在本发明一个优选实施例中刻蚀气体优选 C_3F_6、C_4F_8 等含 C 量比较高的气体并进一步优选通过增加氧化性气体如 O_2、CO 等控制侧壁形貌。如图 4A、4B 所示，用于暴露共源区的第一开孔的尺寸（例如直径）要大于或等于用于形成沟道区的第二开孔尺寸，例如两者尺寸（直径或者多边形的最大跨距的比值）比大于 1..5 并优选大于等于 2。在本发明一个实施例中，每一个第一开孔周围具有六个第二开孔，以便于提高稍后侧向刻蚀层 2B 的效率以及均匀性。在本发明其他实施例中，每一个第一开孔周围可以具有的第二开孔的数目为 2、3、4、5、6、7、8、9 乃至更多。

[0030] 此后，在第二开孔中形成多个垂直的沟道层 4。沟道层 4 的材质可以包括单晶硅、单晶锗、SiGe、Si:C、SiGe:C、SiGe:H 等半导体材料，沉积工艺包括 LPCVD、PECVD、HDPCVD、MOCVD、MBE、ALD 等。在本发明一个实施例中，沟道层 4 的沉积方式为局部填充第二开孔的侧壁而形成为具有空气隙的中空柱形。在本发明其他实施例中，选择沟道层 4 的沉积方式以完全或者局部填充第二开孔，形成实心柱、空心环、或者空心环内填充绝缘层（未示出）的核心 - 外壳结构。沟道层 4 的水平截面的形状与第二开孔类似并且优选地共形，可以为实心的矩形、方形、菱形、圆形、半圆形、椭圆形、三角形、五边形、五角形、六边形、八边形等等各种几何形状，或者为上述几何形状演 化得到的空心的环状、桶状结构（并且其内部可以填充绝缘层）。优选地，对于空心的柱状沟道层 4 结构，可以进一步在沟道层 4A 内侧填充绝缘隔离层 4B，例如通过 LPCVD、PECVD、HDPCVD 等工艺形成例如氧化硅材质的层 4B，用于支撑、绝缘并隔离沟道层 4A。此后，在沟道层 4 顶部沉积漏区接触 4D。优选地，采用与沟道层 4 材质相同或者相近（例如与 Si 相近的材质 SiGe、SiC 等，以便微调晶格常数而提高载流子迁移率，从而控制单元器件的驱动性能）的材质沉积在第二开口的顶部而形成存储器件单元晶体管的漏区 4D，并且可以进一步形成硅化物（未示出）以降低接触电阻。

[0031] 如剖视图 5A 和顶视图 5B 所示，选择性刻蚀去除填充层 3，重新露出第一开孔 2T，利用第一开孔 2T 侧向刻蚀去除堆叠结构中的第二材料层（伪栅极层）2B。首先，采用垂直的各向异性刻蚀工艺，例如采用 RIE 或等离子干法刻蚀工艺（进一步提高碳氟比以刻蚀非晶硅、非晶锗，或者采用氧等离子干法刻蚀去除非晶碳、DLC 材质的填充层 3)，或者针对填充层 3 的材质而选用基本或者完全不腐蚀层 2A、层 2B 的腐蚀液来湿法腐蚀去除层 3（例如针对非晶硅材质的层 3，采用 TMAH 腐蚀液）。垂直的各向异性刻蚀工艺完全去除了填充层 3，重新露出了较宽的垂直的第一开孔 2T。随后，采用各向同性干法刻蚀工艺，横向刻蚀去除层 2B，在层 2A 之间留下了侧向凹槽 2R。例如减小碳氟比以横向刻蚀氮化硅的层 2B、或者采用热磷酸腐蚀氮化硅的层 2B。备选地，当层 2A 为氮化硅、层 2B 为氧化硅时，可以采用 HF 基腐蚀液腐蚀层 2B。

[0032] 如附图 6 所示，在第一开孔 2T 底部形成共源区 1S，在凹槽 2R 中形成栅极堆叠结构 5A/5B。可以通过离子注入掺杂、以及优选地进一步在表面形成金属硅化物（未示出）而形

Micron Ex. 1002, p. 616
Micron v. YMTC
IPR2024-00789

成源区 1S。金属硅化物例如 $NiSi_{2-y}$、$Ni_{1-x}Pt_xSi_{2-y}$、$CoSi_{2-y}$ 或 $Ni_{1-x}Co_xSi_{2-y}$，其中 x 均大于 0 小于 1，y 均大于等于 0 小于 1。栅极堆叠结构 5 包括栅极绝缘层 5A 和栅极导电层 5B。层 5A 的沉积方法包括 PECVD、HDPCVD、MOCVD、MBE、ALD、蒸发、溅射等。图中未示出的是，层 5A 优选地进一步包括多个子层，例如隧穿层、存储层、阻挡层。其中隧穿层包括 SiO_2 或高 k 材料，其中高 k 材料包括但不限于氮化物（例如 SiN、SiON、AlN、TiN）、金属氧化物（主要为副族和镧系金属元素氧化物，例如 MgO、Al_2O_3、Ta_2O_5、TiO_2、ZnO、ZrO_2、HfO_2、CeO_2、Y_2O_3、La_2O_3）、氮氧化物（如 HfSiON）、钙钛矿相氧化物（例如 $PbZr_xTi_{1-x}O_3$(PZT)、$Ba_xSr_{1-x}TiO_3$(BST)）等，隧穿层可以是上述材料的单层结构或多层堆叠结构。存储层是具有电荷俘获能力的介质材料，例如 SiN、HfO、ZrO 等及其组合，同样可以是上述材料的单层结构或多层堆叠结构。阻挡层可以是氧化硅、氧化铝、氧化铪等介质材料的单层结构或多层堆叠结构。在本发明一个实施例中，层 5A 例如是氧化硅、氮化硅、氧化硅组成的 ONO 结构。

[0033] 栅极导电层 5B 可以是多晶硅、多晶锗硅、或金属，其中金属可包括 Co、Ni、Cu、Al、Pd、Pt、Ru、Re、Mo、Ta、Ti、Hf、Zr、W、Ir、Eu、Nd、Er、La 等金属单质、或这些金属的合金以及这些金属的氮化物，栅极导电层 5B 中还可掺杂有 C、F、N、O、B、P、As 等元素以调节功函数。栅极绝缘层 5A 与栅极导电层 5B 之间还优选通过 PVD、CVD、ALD 等常规方法形成氮化物的阻挡层（未示出），阻挡层材质为 M_xN_y、$M_xSi_yN_z$、$M_xAl_yN_z$、$M_aAl_xSi_yN_z$，其中 M 为 Ta、Ti、Hf、Zr、Mo、W 或其它元素。同样地，层 5B 可以是单层结构也可以是多层堆叠结构。

[0034] 进一步地，如图 6 所示，在第一开孔 2T 侧壁形成较薄的绝缘层 6A 并在绝缘层 6A 中填充金属形成源极接触 6B。绝缘层 6A 可以采用热氧化、化学氧化、PECVD、HDPCVD 等工艺形成，材质可以为氧化硅、氮化硅或高 K 材料。层 6A 优选仅覆盖第一开孔 2T 侧壁，底部堆积的较薄部分可以采用湿法腐蚀工艺或者各向同性干法刻蚀工艺去除，类似于侧墙技术。随后，采用 MOCVD、MBE、ALD、蒸发、溅射等工艺形成低电阻的共源极接触线 6B，也称作 CSL。优选地，采用 CMP 或者回刻方法，或者调整沉积金属工艺参数，降低层 6B 的高度使其顶部低于最顶部的栅极堆叠结构 5A/5B 的底面（顶部的该金属栅极结构即用于顶部选择晶体管的顶部选择栅线 USG），接着再回填绝缘层 6A 直至与最顶层的层 2A 齐平。

[0035] 此后，优选地，如图 7A 剖视图和图 7B 顶视图所示，刻蚀去除 CSL 层 6B 顶部的绝缘层 6A，露出下方的层 6B。在该刻蚀步骤中，刻蚀形成的第三开孔 6T 尺寸要大于第一开孔 2T 的尺寸（例如圆形的直径或多边形的最大跨距），去除了一部分 USG 层级的栅极堆叠结构 5A/5B。如图 7B 所示，共源极接触线 CSL 图形连接了多个第三开孔 6T，并且可以进一步延伸出存储单元区而连接至外部电路形成互连线。

[0036] 如图 8 所示，在第三开孔 6T 中填充同样的绝缘层 6A，形成了顶层选择栅极线的隔离绝缘区。

[0037] 如图 9 所示，在整个器件之上形成层间介质层（ILD）7。例如采用喷涂、旋涂、CVD 沉积等工艺形成低 k 材料的 ILD 7，低 k 材料包括但不限于有机低 k 材料（例如含芳基或者多元环的有机聚合物）、无机低 k 材料（例如无定形碳氮薄膜、多晶硼氮薄膜、氟硅玻璃、BSG、PSG、BPSG）、多孔低 k 材料（例如二硅三氧烷（SSQ）基多孔低 k 材料、多孔二氧化硅、多孔 SiOCH、掺 C 二氧化硅、掺 F 多孔无定形碳、多孔金刚石、多孔有机聚合物）。随后刻蚀 ILD 7 直至暴露 CSL 层 6B，形成第四开孔（未示出）并沉积金属材料形成接触塞 8，连接了阵列的共源区 1S。

Micron Ex. 1002, p. 617
Micron v. YMTC
IPR2024-00789

[0038] 如图 10A 和图 10B 所示，完整位线连接。在第一 ILD 7 之上采用类似工艺和材质沉积第二 ILD 9 并刻蚀形成暴露沟道区 4D 的接触孔，沉积金属形成位线连接 10。

[0039] 值得注意的是，图 1 至图 10 仅示出了根据本发明一个实施例的工艺方法，本发明还可以采用其他工序完成三维存储器件的制造。例如，在本发明一个优选的备选实施例中，图 3B 和图 4B 所示的用于共源区的第一开孔与用于沟道区的第二开孔可以通过一次光刻 / 刻蚀构图实现，也即图 3 和图 4 所示工艺可以合并，只要在每个第一开孔周边存在多个第二开孔即可。在本发明另一备选实施例中，在执行完成图 6 所示的栅极堆叠结构填充之后，在图 7A、图 7B 以及之后所示的连线阶段按照图 7B、图 10B 所示的版图进行刻蚀，完成各层栅电极隔离（中心包围了金属层 6A 的隔离层 6B 图形）并形成共源极连线 CSL 图形。具体地，在刻蚀孔完成栅堆栈沉积以后，然后介质填充平坦化整个平面，然后按照图 10B 所示 CSL 标注的区域进行物理刻蚀，类似于 TCAT 工艺中 WL-Cut 的步骤，沉积薄的介质层侧墙，然后沉积金属等导电层形成一个共源区的导电面。好处在于，前一个实例中除了 USG 隔离其他的 WL 没有隔离，这个方法可以完全像 TCAT 一样完成各层 WL 的隔离。

[0040] 在本发明另一备选实施例中，可以在图 5 移除作为伪栅极的层 2B 工序之中，与图 5 所示相邻两个第一开孔 2T 直线连线之间仅有一个沟道层 4 不同，如图 10C 所示，可以在两列（沿图 5B 中上下方向排列）第一开孔 2T 之间具有多个沟道层 4，由此增大了第一开孔 2T 之间的距离，并且在刻蚀去除层 2B 时仅部分去除而在凹槽 2R 中保留部分伪栅极层 2B，利用伪栅极层 2B 的自对准完成了控制栅的自动隔离。换言之，相邻的两列沟道层 4 之间的栅极堆叠结构 5A/5B 的侧面通过残留的第二材料层 2B 而绝缘隔离，增强了器件的绝缘性能。

[0041] 依照本发明的三维存储器制造方法，将 TCAT 三维器件的字线深槽替换为深孔刻蚀来完成相同的功能，提高集成密度，简化堆叠结构的刻蚀工艺，保留了金属栅控制性能。

[0042] 尽管已参照一个或多个示例性实施例说明本发明，本领域技术人员可以知晓无需脱离本发明范围而对器件结构或方法流程做出各种合适的改变和等价方式。此外，由所公开的教导可做出许多可能适于特定情形或材料的修改而不脱离本发明范围。因此，本发明的目的不在于限定在作为用于实现本发明的最佳实施方式而公开的特定实施例，而所公开的器件结构及其制造方法将包括落入本发明范围内的所有实施例。

Micron Ex. 1002, p. 618
Micron v. YMTC
IPR2024-00789

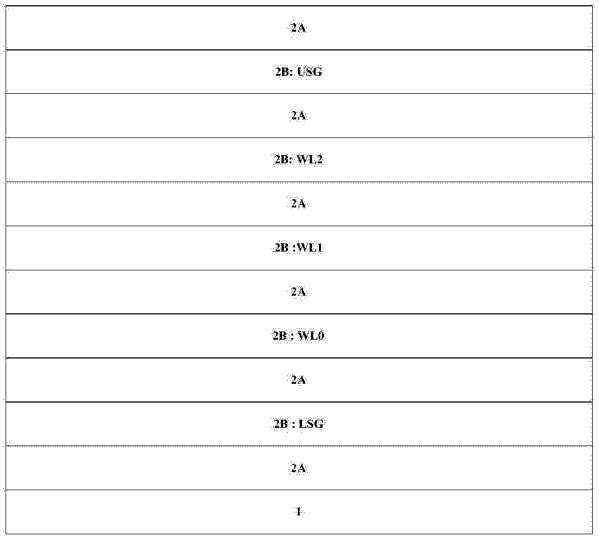


图 1

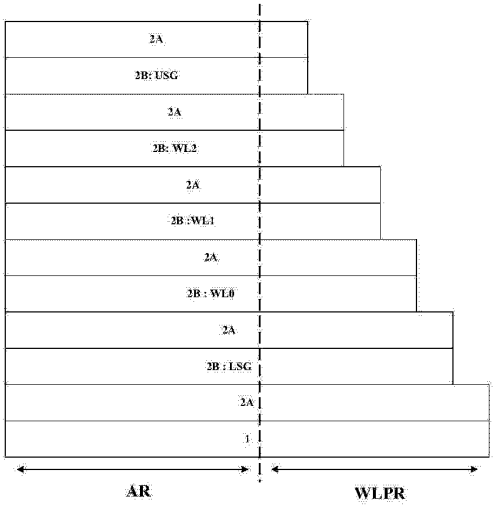


图 2A

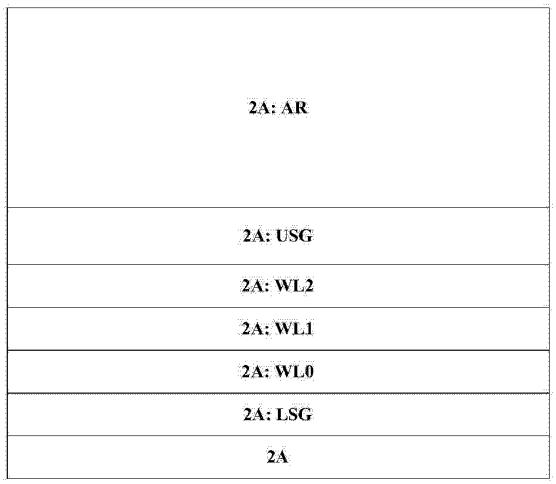


图 2B

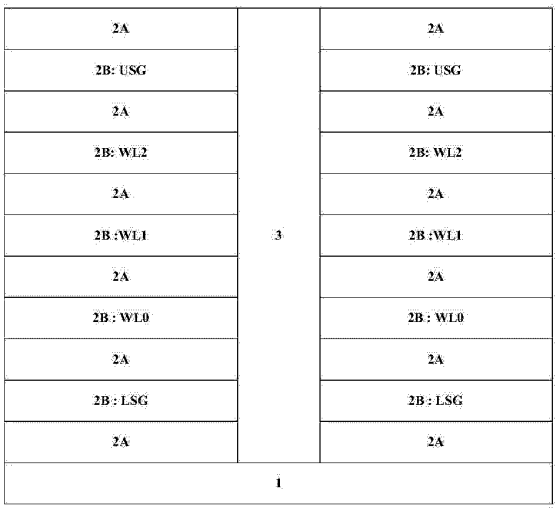


图 3A

Micron Ex. 1002, p. 619
Micron v. YMTC
IPR2024-00789

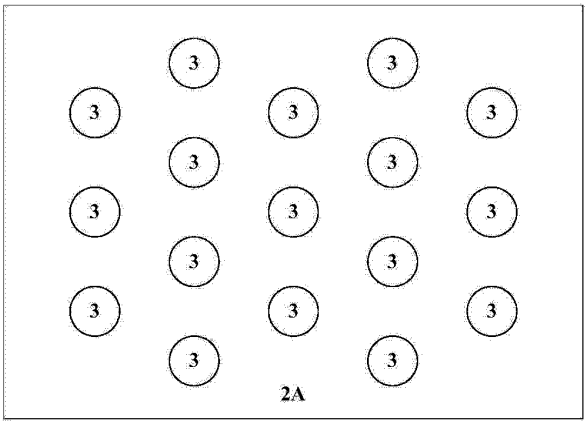


图 3B

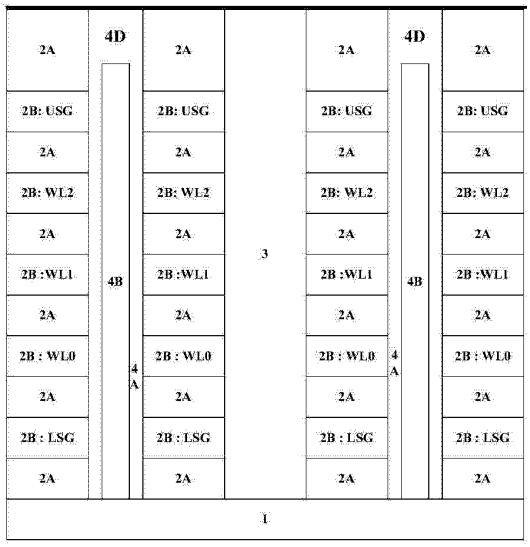


图 4A

Micron Ex. 1002, p. 620
Micron v. YMTC
IPR2024-00789

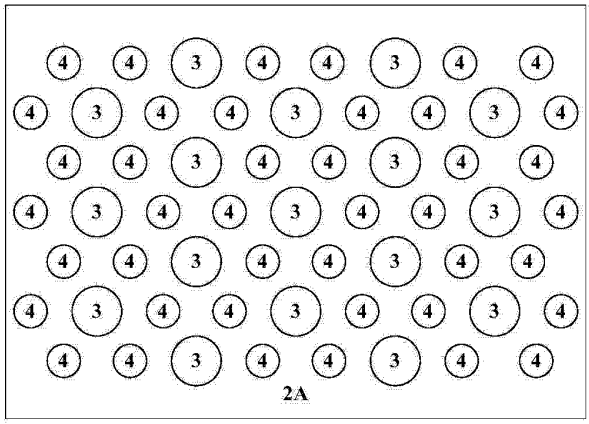


图 4B

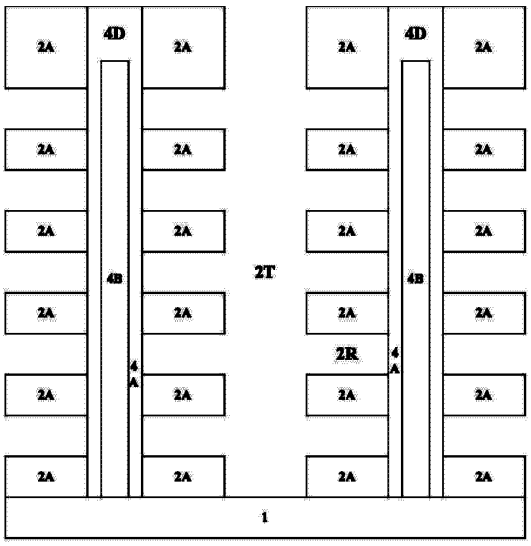


图 5A

Micron Ex. 1002, p. 621
Micron v. YMTC
IPR2024-00789

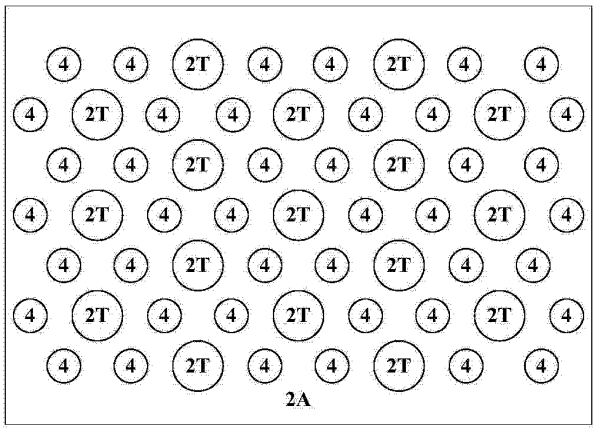


图 5B

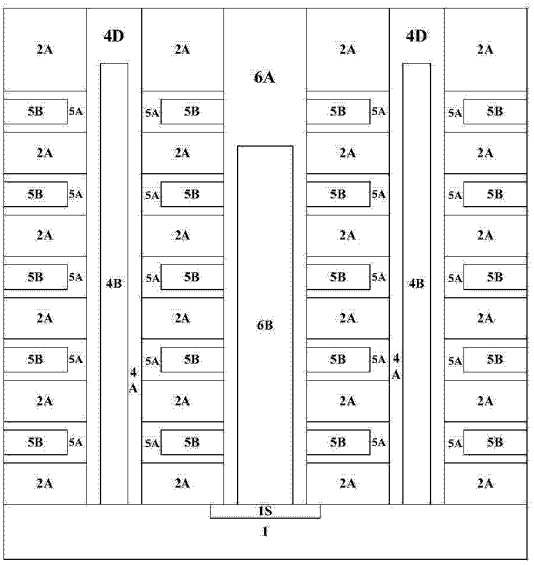


图 6

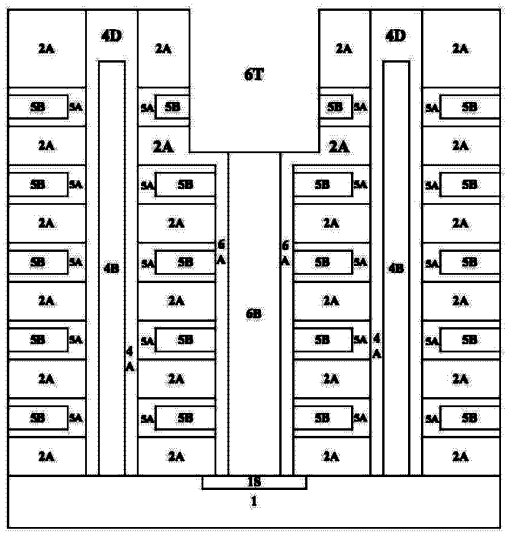


图 7A

Micron Ex. 1002, p. 622
Micron v. YMTC
IPR2024-00789

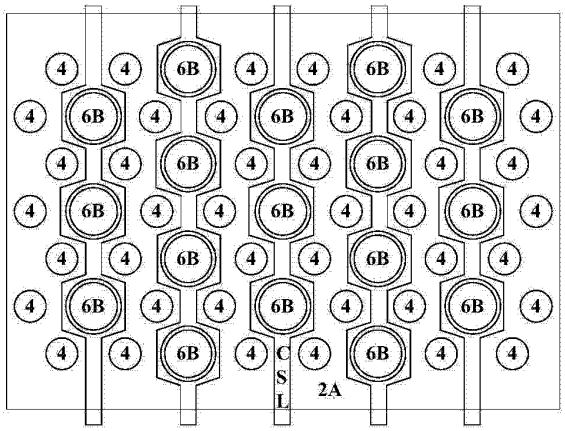


图 7B

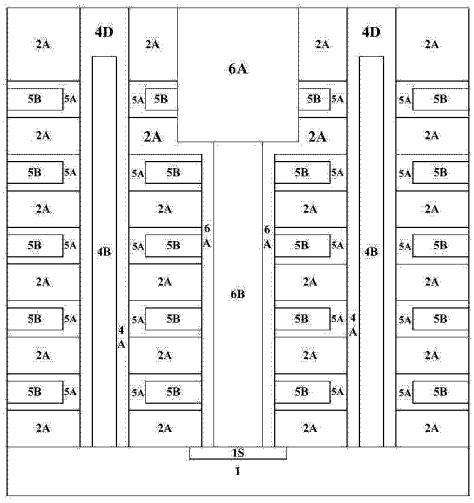


图 8

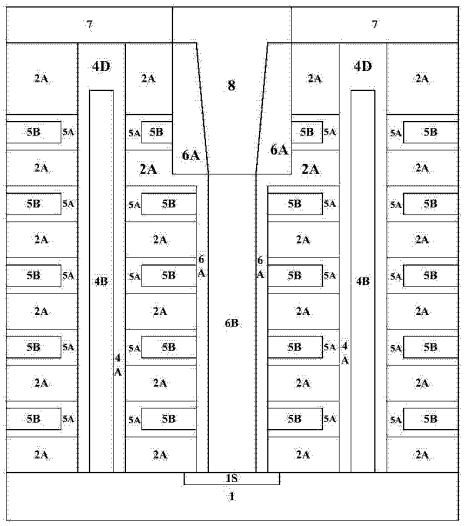


图 9

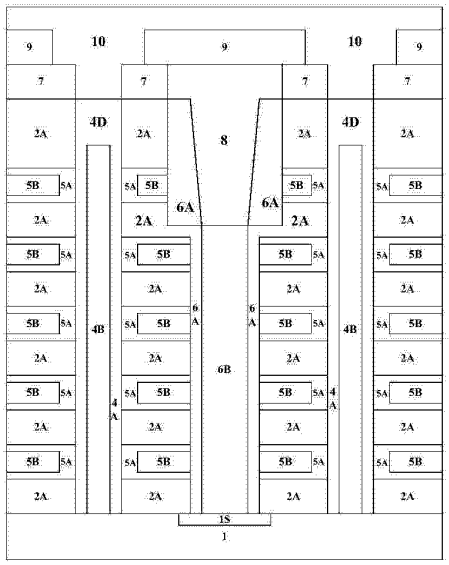


图 10A

Micron Ex. 1002, p. 623
Micron v. YMTC
IPR2024-00789

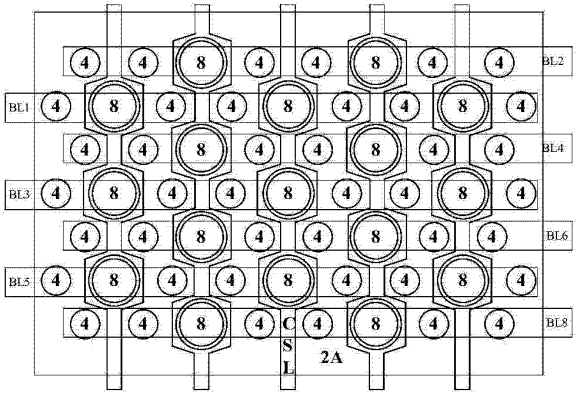


图 10B

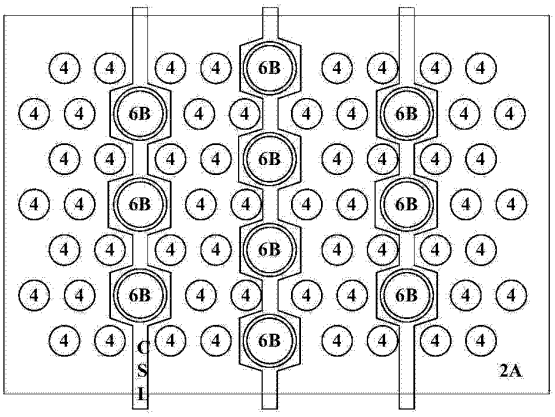


图 10C

Micron Ex. 1002, p. 624
Micron v. YMTC
IPR2024-00789

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 April 2018 (19.04.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/071143 A2

(51) International Patent Classification:
H01L 27/06 (2006.01) *H01L 27/108* (2006.01)
H01L 21/8238 (2006.01) *H01L 27/11* (2006.01)
H01L 27/11551 (2017.01)

(21) International Application Number:
PCT/US2017/052359

(22) International Filing Date:
19 September 2017 (19.09.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/406,376	10 October 2016 (10.10.2016)	US
62/432,575	11 December 2016 (11.12.2016)	US
62/440,720	30 December 2016 (30.12.2016)	US
62/457,838	11 February 2017 (11.02.2017)	US
62/460,989	20 February 2017 (20.02.2017)	US
62/471,963	16 March 2017 (16.03.2017)	US
62/480,529	02 April 2017 (02.04.2017)	US
62/484,398	12 April 2017 (12.04.2017)	US
62/488,821	23 April 2017 (23.04.2017)	US
62/517,152	08 June 2017 (08.06.2017)	US
62/530,173	08 July 2017 (08.07.2017)	US
62/535,265	21 July 2017 (21.07.2017)	US
62/549,952	24 August 2017 (24.08.2017)	US

(71) Applicant: **MONOLITHIC 3D INC.** [US/US]; 3555 Woodford Drive, San Jose, California 95124 (US).

(72) Inventors: **OR-BACH, Zvi**; 3555 Woodford Drive, San Jose, California 95124 (US). **HAN, Jin-Woo**; 2131 Samaritan Drive, San Jose, California 96124 (US). **CRONQUIST, Brian**; 6315 Cessna Court, San Jose, California 95123 (US). **LUSKY, Eli**; Aaron Katzir Avenue, 5200200 Ramat-Gan (IL).

(54) Title: 3D SEMICONDUCTOR DEVICE AND STRUCTURE

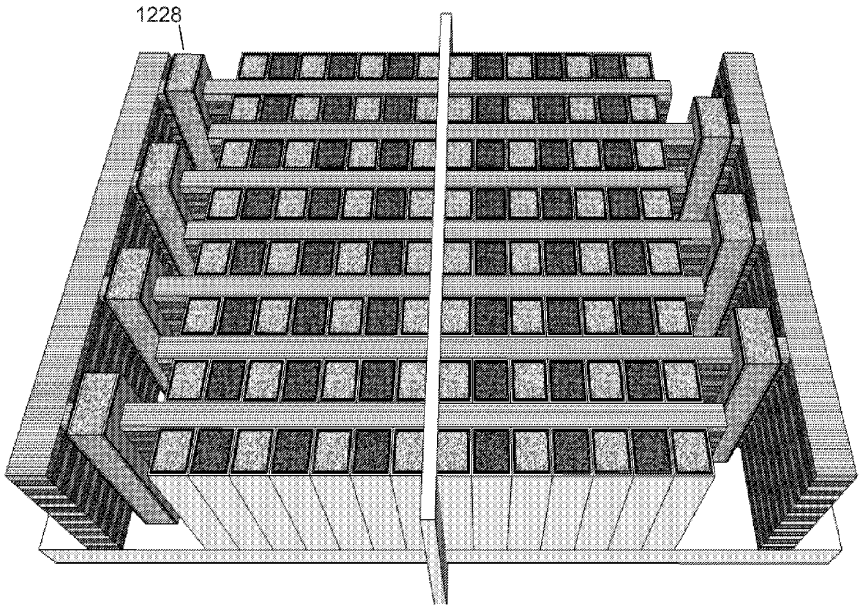


Fig. 12D

(57) Abstract: A 3D device, the device comprising: a first stratum comprising a first bit-cell array, the first bit-cell array includes three independent first rows; a second stratum including a second bit-cell array, the second bitcell array includes three independent second rows, where the second stratum overlays the first stratum; and at least three vertical bitlines each connected to respective three horizontal first bitlines and three horizontal second bitlines, where the three horizontal first bitlines include control of the first bit-cell array, where the three horizontal second bitlines include control of the second bit-cell array, and where each of the three vertical bitlines could be used to control a different one of the three independent first rows, or control a different one of the three independent second rows

WO 2018/071143 A2

(74) **Agent: TRAN, Bao**; 3555 Woodford Drive, San Jose, California 95124 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

3D SEMICONDUCTOR DEVICE AND STRUCTURE

BACKGROUND OF THE INVENTION

1. Field of the Invention

[0001] This application relates to the general field of Integrated Circuit (IC) devices and fabrication methods, and more particularly to multilayer or Three Dimensional Integrated Memory Circuit (3D-Memory) and Three Dimensional Integrated Logic Circuit (3D-Logic) devices and fabrication methods.

2. Discussion of Background Art

[0002] Over the past 40 years, there has been a dramatic increase in functionality and performance of Integrated Circuits (ICs). This has largely been due to the phenomenon of "scaling"; i.e., component sizes such as lateral and vertical dimensions within ICs have been reduced ("scaled") with every successive generation of technology. There are two main classes of components in Complementary Metal Oxide Semiconductor (CMOS) ICs, namely transistors and wires. With "scaling", transistor performance and density typically improve and this has contributed to the previously-mentioned increases in IC performance and functionality. However, wires (interconnects) that connect together transistors degrade in performance with "scaling". The situation today is that wires dominate the performance, functionality and power consumption of ICs.

[0003] 3D stacking of semiconductor devices or chips is one avenue to tackle the wire issues. By arranging transistors in 3 dimensions instead of 2 dimensions (as was the case in the 1990s), the transistors in ICs can be placed closer to each other. This reduces wire lengths and keeps wiring delay low.

[0004] There are many techniques to construct 3D stacked integrated circuits or chips including:

- Through-silicon via (TSV) technology: Multiple layers of dice are constructed separately. Following this, they can be bonded to each other and connected to each other with through-silicon vias (TSVs).
- Monolithic 3D technology: With this approach, multiple layers of transistors and wires can be monolithically constructed. Some monolithic 3D and 3DIC approaches are described in U.S. Patents 8,273,610, 8,557,632, 8,298,875, 8,642,416, 8,362,482, 8,378,715, 8,379,458, 8,450,804, 8,574,929, 8,581,349, 8,642,416, 8,687,399, 8,742,476, 8,674,470, 8,803,206, 8,902,663, 8,994,404, 9,021,414, 9,023,688, 9,030,858, 9,117,749, 9,219,005; U.S. patent publication 2011/0092030; and pending U.S. Patent Applications, 62/077,280, 62/042,229, 13/803,437, 61/932,617, 14/607,077, 14/642,724, 62/139,636, 62/149,651, 62/198,126, 62/239,931, 62/246,054, 62/307,568, 62/297,857, 15/095,187, 15/150,395, 15/173,686, 62/383,463, 15/243,941, PCT/US2016/052726,15/494,525 (WO2017053329), 62/406,376, 62/432,575, 62/440,720, 62/457,838, 62/460,989, 62/471,962, 62/480,529, 62/484,398, 62/488,821, 62/501,136, 62/517,152, 62/523,760, 62/530,173, 62/535,265, 62/539,054, 62/549,952, 15/333,138 and 15/344,562. The entire contents of the foregoing patents, publications, and applications are incorporated herein by reference.
- Electro-Optics: There is also work done for integrated monolithic 3D including layers of different crystals, such as U.S. Patents 8,283,215, US 8,163,581, 8,753,913, 8,823,122, 9,197,804; and U.S. Patent Application No. 14/461,539. The entire contents of the foregoing patents, publications, and applications are incorporated herein by reference.

[0005] Additionally the 3D technology according to some embodiments of the invention may enable some very innovative IC devices alternatives with reduced development costs, novel and simpler process flows, increased yield, and other illustrative benefits.

SUMMARY

[0006] The invention relates to multilayer or Three Dimensional Integrated Circuit (3D IC) devices and fabrication methods. Important aspects of 3D IC are technologies that allow layer transfer. These technologies include technologies that support reuse of the donor wafer, and technologies that support fabrication of active devices on the transferred layer to be transferred with it.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Various embodiments of the invention will be understood and appreciated more fully from the following detailed description, taken in conjunction with the drawings in which:

[0008] Fig. 1A is an example illustration of selective etch by Applied Material etcher;

[0009] Fig. 1B is an example illustration of alternating Si and SiGe layers;

[00010] Figs. 2A-2I are example illustrations of an alternative system process flow;

[00011] Figs. 3A-3D are example illustrations of the structure transfer of an alternative system process flow;

[00012] Figs. 4A-4H are example illustrations of an alternative donor wafer process flow;

[00013] Figs. 5A-5D are example illustrations of an additional alternative donor wafer process flow;

[00014] Figs. 6A-6D are example illustrations of structures and methods to thermally isolate stratums in a 3D IC;

[00015] Fig. 7 is an example illustration of a logic unit redundancy scheme;

[00016] Fig. 8 is an example illustration of an alternative logic redundancy scheme;

[00017] Figs. 9A-9B are example illustrations of a staircase contact scheme;

[00018] Figs. 10A-10J are example illustrations the formation and structure of a staircase connectivity scheme;

[00019] Figs. 10K-10N are example illustrations of process simulations of the formation and structure of a staircase connectivity scheme;

[00020] Figs. 11A-11D are example illustrations of a ''Lego'' scheme with 3DNOR;

[00021] Fig. 11E is an example illustration of an isolation structure to keep the logic heat away from the memory array;

[00022] Figs. 11F-11K are example illustrations of alignment techniques and structures;

[00023] Figs. 12A-12J are example illustrations of array access and staircase schemes;

[00024] Figs. 13A-13D are example illustrations of write and erase schemes for the 3D NOR array;

[00025] Fig. 14A is an example illustration of I(V) curves of a silicided nanowire formed with microwave annealing technology;

[00026] Fig. 14B is an illustration of I(V) curves demonstrating the performance of a self-referenced differential sense amplifier when applied to a programmed and erased memory cell;

[00027] Figs. 15A-15E are example illustrations of a 'DS-SB' 3D NOR structure and cell formation;

[00028] Fig. 15F is an example illustration of a programming scheme for 3D NOR with a staircase;

[00029] Fig. 15G is an example illustration of an alternative to a staircase;

[00030] Figs. 16A-16D are example illustrations of writing a ferroelectric memory cell;

[00031] Figs. 17A-17D are example illustrations of two bits stored in one facet of a memory cell;

[00032] Figs. 18A-18D are example illustrations of writing bits in a small section of a ridge of the 3D-NOR fabric;

[00033] Figs. 18E-18F are example illustrations of programming multiple bits in a cell of the 3S-NOR fabric;

[00034] Fig. 19A is an example illustration of a 3D system;

[00035] Fig. 19B is an example illustration of customization of a 3D system;

[00036] Figs. 20A- 20F are example illustrations of the formation of multiple stratum integrated into a 3D system via 3D integration with minimum per strata processing;

[00037] Fig. 20G is an example illustration of a connectivity technique adapted to die level operation;

[00038] Fig. 20H is an example illustration of a through strata via structure;

[00039] Figs. 21A- 21H are example illustrations of a control line arrangement for memory integration of the 3D system of Figs.19 and 20;

[00040] Figs. 22A-22B are example illustrations of a stratum select connectivity scheme;

[00041] Figs. 22C-22E are example illustrations of wordline and bitline select schemes;

[00042] Fig. 22F is a section of a partition an array of memory units;

[00043] Figs. 23A-23B are example illustrations of two-layer select and select schemes;

[00044] Fig. 24A is an example illustration of an alternative 3D computer system;

[00045] Fig. 24B is an example illustration of a generic 3D memory structure "G3DM";

[00046] Fig. 24C is an example illustration of 3D structure with active thermal cooling;

[00047] Figs. 25A-25D are an example illustrations of an alternative flow and structure for 3D stacking without changing the main processing of the memory/logic;

[00048] Fig. 25E is an example illustration of forming alignment marks for the method and structure of Figs 25A-25D;

[00049] Figs. 25F and 25G are example illustrations of using a lithographically defined doping process to simplify 3D stacking;

[00050] Figs. 25H-25J are example illustrations of over the array connectivity structures;

[00051] Figs. 26A-26D are example illustrations of using a punch and plug process scheme;

[00052] Fig. 26E is an example illustration of the single hole punch process to construct elements which may be needed for the 3D NOR fabric;

[00053] Fig. 26F is an example illustration of holes/vias etched or punched together/simultaneously;

[00054] Fig. 26G is an example illustration of holes/vias etched or punched in two or more independent etch steps;

[00055] Fig. 26H is an example illustration of global strata select for stacking of 3D memory structures;

[00056] Fig. 27 is an example illustration of a memory unit refresh operation flow;

[00057] Fig. 28 is an example illustration of an alternative per layer access by sidewall strapping though one-time-programmable anti-fuses; and

[00058] Figs. 29A and 29B are example illustrations of a pass-through path add-on structure and an alternative structure.

DETAILED DESCRIPTION

[00059] An embodiment of the invention is now described with reference to the drawing figures. Persons of ordinary skill in the art will appreciate that the description and figures illustrate rather than limit the invention and that in general the figures are not drawn to scale for clarity of presentation. Such skilled persons will also realize that many more embodiments are possible by applying the inventive principles contained herein and that such embodiments fall within the scope of the invention which is not to be limited except by any appended claims.

[00060] Some drawing figures may describe process flows for building devices. The process flows, which may be a sequence of steps for building a device, may have many structures, numerals and labels that may be common between two or more adjacent steps. In such cases, some labels, numerals and structures used for a certain step's figure may have been described in the previous steps' figures.

[00061] The use of layer transfer in the construction of a 3D IC based system could enable heterogeneous integration where each of strata may include one or more of MEMS sensor, image sensor, CMOS SoC, volatile memory such as DRAM and SRAM, persistent memory, and non-volatile memory such as flash and OTP. Such could include adding memory control circuits, also known as peripheral circuits, on top or below a memory array. The memory strata may contain only memory cells but not control logic, thus the control logic may be included on a separate strata. Alternatively, the memory strata may contain memory cells and simple control logic where the control logic on that strata may include at least one of decoder, buffer memory, sense amplifier. The circuits may include the charge pumps and high voltage transistors, which could be made on a strata using silicon transistors or other transistor types (such as SiGe, Ge, CNT, etc.) using a manufacturing process line that is different than the low voltage control circuit manufacturing process line. The analog circuits, such as for the sense amplifiers, and other sensitive linear circuits could also be processed independently and be transferred over to the 3D fabric. Such 3D construction could include "Smart Alignment" techniques presented in this invention or leverage the repeating nature of the memory array to reduce the impact of the wafer bonder misalignments on the effectiveness of the integration.

[00062] In patents such as, for example, U.S. Patent Application No. 15/173,395, layer transfer techniques called ELTRAN (epitaxial layer transfer) are presented and may be part of the formation process of a 3DIC. The ELTRAN technique utilizes an

epitaxial process or processes over porous layers. Alternatively other epitaxial based structures could be formed to support layer transfer techniques by leveraging the etch selectivity of these epitaxial layers, such as the very high etch selectivity of SiGe vs. Silicon, and variations such as Silicon (single crystal or poly or amorphous), SiGe (mix of silicon and Germanium), P doped silicon, N doped silicon, etc. Alternately, these layer(s) could be combined with types of detachment processes, such as 'cold splitting,' for example the Siltectra stress polymer and low temperature shock treatment, to provide a thin layer transfer process.

[00063] Recently it become a very attractive concept for processing gate all around horizontal transistors and has become the target flow for next generation devices such as the 5 nm technology node. Some of the work in respect to selective etching of SiGe vs. silicon has been presented in a paper by Jang-Gn Yun *et al.* titled: "Single-Crystalline Si Stacked Array (STAR) NAND Flash Memory" published in IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 58, NO. 4, APRIL 2011, and a more recent work by K. Wostyn *et al.* titled "Selective Etch of Si and SiGe for Gate All-Around Device Architecture" published in ECS Transactions, 69 (8) 147-152 (2015), and by V. Destefanis *et al.* titled: "HCl Selective Etching of Si1-xGex versus Si for Silicon On Nothing and Multi Gate Devices" published in ECS Transactions, 16 (10) 427-438 (2008), all of the forgoing incorporated herein by reference. Since the SiGe over Si substrate process is becoming mature, this facilitates using a SiGe layer as a sacrificial layer for production worthy 3D layer transfer. Fig. 1A illustrates the high etch selectivity of SiGe vs. Silicon, which, in this example, could be formed using, for example, the Applied Material Selectra etch system. Alternatively, the selective etch may be made using a wet chemical etch. Fig. 1B illustrates a putative retrograde Ge composition with stack thickness as explained thoroughly later.

[00064] An exemplary layer transfer process could include the steps A-K, referencing the illustrations Fig. 2A-2I and Fig. 3A-3D:

[00065] A. As illustrated in Figs. 2A and 2B, epi (the term 'epi' herein means epitaxial, as often used in the art) layer 204 such as SiGe may be formed on a donor wafer or reusable donor wafer - base substrate 202, for example, by an epitaxy processes. The donor wafer may contain a strata layer over the sacrificial layer, also called herein a 'cut-layer', on the base substrate 202 where the strata layer is subsequently transferred to a receptor wafer. The epitaxy process may utilize but not is limited to vapor-phase epitaxy (VPE), a modification of chemical vapor deposition, molecular-beam and liquid-phase epitaxy (MBE and LPE). If desired so as to at least increase etch selectivity, dopant may be incorporated during the epitaxial growth process by adding impurities to the source gas and/or reaction chamber. The type of epitaxy may be homoepitaxy with the same material grown on the base substrate 202. In the homoepitaxy, the doping type and concentration of the epi layer 204 may be substantially different from those of the base substrate 202 and the subsequently formed silicon layer 206, which could provide sufficient etch selectivity. Alternatively, another type of epitaxy may be heteroepitaxy with different material grown on the substrate. Such examples include SiGe on Si. Epi layer 204 may be formed with a thickness of about 20nm, 50nm, 100nm, or about 200nm, or about 500nm, or about 1000nm or about 2000nm, depending on process integration, etching throughput, stiction resistance, and other process and device architecture engineering decisions and tradeoffs. The base wafer 202 thickness could be similar to the industry standard for these types of processing such as about 775 microns used for most current silicon fabs. The base wafer 202 may include sizes of about 2 inch, about 4 inch, about 8 inch, or about 12 inch in diameter or about 16 inch in the future (these wafer diameter sizes are often known with the corresponding mm sizes: 50 mm, 100 mm, 150 mm, 200 mm, 300 mm, 450 mm). Heterogeneously grown epi layer 204 may include various materials, for example SiGe, in anticipation of a sacrificial etch later in the process flow. The content of Ge in the SiGe may be designed per the selectivity desired and in consideration of the stress, for example, such as about 20% Ge, or about 10% Ge, or about 25% Ge, etc. With respect to the etch selectivity, the epi layer 204 could be favorably removed from the base substrate 202 and the subsequently formed silicon layer 206. The level of stress needs to be controlled to not cause dislocations. The stress can be controlled by at least growth rate, temperature, and film thickness. Predefined trenches on the silicon layer 206 and partially or fully across the epi layer 204 (not shown) in the designated dicing streets (or other non-circuit areas, for example, such as the streets between projection fields) could be used to provide some release of the potential stress. These trenches could have a width and a depth approximately corresponding to the thickness of the SiGe layer. Other techniques could be used to reduce stress, associated with epitaxial of SiGe on or under a silicon layer, such as incorporating smaller atoms in the SiGe layer such as boron or carbon; these and similar techniques are known in the art and could be used with combination with the techniques herein to support various forms of layer transfer or other applications herein.

[00066] B. As illustrated in Fig. 2C, silicon layer 206 as an active layer for the active devices to be formed may be epitaxially grown on top of epi layer 204. The silicon layer 206 may also be single crystalline Ge, SiGe or Si:C depending on the applications. However, we herein will use silicon layer 206 as the active layer unless otherwise specified. The thickness of silicon layer 206 may include a thickness of about 10nm, or about 20nm, or about 50nm, or about 100nm, or about 200nm, or about 500nm, or about 1,000nm or about 2,000nm, as desired for the electronic circuits, depending on engineering, materials and scientific device considerations. In many formations, silicon layer 206 may be considered to consist of mono-crystalline or single crystal silicon.

[00067] C. As illustrated in Fig. 2D, desired circuits 212 may be processed, including n-type and p-type transistors and other devices, such as inductors, capacitors, resistors, optoelectronic devices, gas sensors, etc., and may include a processed contact layer. For example, the desired circuit 212 may be processed to include metal 0 layer or metal 4 layer. This could be done using conventional processing including the appropriate high temperature processes (~600-900 °C) such as gate oxidation, dopant activation, contact silicidation, and so on. Types of transistors and circuits may include, for example, DRAM, NAND, or RRAM, RCAT, continuous array and FPGA structures, gate array, memory blocks, logic blocks, CMOS p-type and n-type transistors, MOSFET transistors, junction-less transistors, JFET, replacement gate transistors, thin-side-up transistors, double gate transistors, horizontally oriented transistors, finfet transistors, fully depleted thin-body transistor, JLRCAT, DSS Schottky transistors, and/or trench MOSFET transistors.

[00068] D. As illustrated in Fig. 2E, first set of holes 214 may be formed (by, for example, a conventional mask and etch sequence of processing) and may extend through the top silicon layer 206. The bottom of the etched trench may reside inside the SiGe epi layer 204 and may not touch the surface of base substrate 202 in order to reclaim the base substrate 202. Alternatively, the bottom of the etched trench may be located below the bottom of SiGe epi layer 214. These holes could then be filled with oxide or other material that would remain and play as a supporter for the desired circuits during future silicon and SiGe etches. Filled first set of holes 214 may serve as posts to hold the top layer designated to be transferred in later steps. First set of holes 214 may be located in the dicing streets area or some field oxides such as shallow trench isolation, and may be designed to be weak enough to be torn out, according to design and process integration engineering considerations.

[00069] E. As illustrated in Fig. 2F, one or more interconnection layers 216 may be added. This is optional, depending on engineering and design considerations. Interconnection layers 216 may include wiring, contacts and vias, and may include materials such as, for example, copper, aluminum, tungsten, titanium, tantalum, cobalt metals and/or silicides of the metals. Interconnection layers 216 may be covered with isolation layer 222 (as illustrated in Fig. 2G) with materials such as SiO_2, carbon containing oxides, and so on. Isolation layer 222 may be planarized, for example, with CMP or other forms of planarization in anticipation of future wafer to wafer bonding.

[00070] F. As illustrated in Fig. 2H, second set holes 224 may be formed to expose portions of sacrificial SiGe epi layer 204, thus allowing a substantially full etch of the sacrificial SiGe layer (the former SiGe epi layer 204). These holes could be made at un-used locations or in locations designated for future Through-Layer-Via (TLVs). These holes could also be made in locations designated for shallow trench isolation ("STI"). These holes may be opened through the top layers such as isolation layer 222 all the way into the epi layer 204. The formation of the holes for the sacrificial layer etches may include steps to protect the top silicon layer 206 and its holes side wall such as sidewall spacer by: 1. Use isotropic deposition techniques such as ALD to deposit isolation material such as silicon oxide in the open holes covering the side wall and the bottom of the holes. 2. Then may use anisotropic etching to open only the bottom of the holes to direct access to the sacrificial layers while protecting the side walls.

[00071] G. As illustrated in Fig. 2I, selectively etch the remaining sacrificial epi layer 204 creating void 226 underneath desired circuit layer 212. The design of posts and allocation of sacrificial etch holes could be designed so after the sacrificial layer etching is completed, the suspended circuit layer be remained substantially flat for layer transfer.

[00072] H. As illustrated in Fig. 3C, the structure such as illustrated in Fig. 3A (the structure from Fig. 2I) may be flipped and bonded onto a target wafer 302 illustrated in Fig. 3B. The target wafer 302 may be fully processed wafer including metallization, for example, an arbitrarily designed SoC or generic circuit. Alternatively, the target wafer 302 may be specially fabricated wafer which could be the underlying base in the 3D structure. For example, see the incorporated references listed as part of the Background of the

Invention section of this specification. This results in bonded structure 390 such as illustrated in Fig. 3C. The bonding could be oxide to oxide bonding which could be followed with top wafer interconnection through a TLV (Thru Layer Via) process, or metal to metal bonding, or hybrid bonding (oxide to oxide and metal to metal bonding). The bonding process could be made so it results in sufficiently a strong adhesion between bonding surfaces of donor wafer and the target wafer 302. The target wafer 302 may include transistors of one or more types in one or more layers, metallization such as, for example, copper or aluminum in one or more layers, interconnections to and between layers above and below, and interconnections within the layer. The transistors may be of various types that may be different from layer to layer or within the same layer. The transistors may be in various organized patterns. The transistors may be in various pattern repeats or bands. The transistors may be in multiple layers involved in the transfer layer. The transistors may be, for example, junction-less transistors or recessed channel array transistors. Target wafer 302 may further comprise semiconductor devices such as resistors and capacitors and inductors, one or more programmable interconnects, memory structures and devices, sensors, radio frequency devices, or optical interconnect with associated transceivers. Target wafer 302 may further include isolation layers, such as, for example, silicon and/or carbon containing oxides and/or low-k dielectrics and/or polymers, which may facilitate oxide to oxide wafer or substrate bonding. Target wafer 302 may also be a base substrate to build the 3DIC stricture onto.

[00073] I. The donor structure, substantially donor wafer - base substrate 202, may be detached from bonded structure 390 leaving intermediate 3D IC structure 399 as is illustrated in Fig. 3D. The donor wafer, substantially reusable donor wafer - base substrate 202, could be reclaimed for processing to prepare for reuse, perhaps as a seed wafer of subsequent epitaxial growth as explained in Fig. 2. Intermediate 3D IC structure 399 may include target wafer 302, desired circuits 212, surface 213, first set of holes 214, and interconnection layers 216. Alternatively, the substrate 202 could be grind and etched back without being reusable for future processing.

[00074] J. The top surface 213 of intermediate 3D IC structure 399 may be cleaned and prepared for interconnections. Optionally cover with isolation.

[00075] K. TLVs may be formed for interconnection from the top to the bottom strata to form the 3DIC layer interconnects if necessary.

[00076] The donor wafer 'tearing off' detach could be assisted by known techniques such as, for example, water-jet, wedge, laser cutting, etched assisted tearing off and mechanical twist and pull.

[00077] Alternatively, additional interconnection layers and other processing could be added in between step 'G' and 'H' above. So, the structure illustrated in Fig. 2I could be further processed before being flipped and bonded to the target wafer 302. This add-on process could include adding additional metal layers or any other structure including additional transistor layers using similar techniques such as layer transfer.

[00078] The sacrificial layer removal holes 224 process could include side wall oxide deposition to further protect the side walls from the etch process designed to remove the sacrificial layers. These holes could be later sealed by a second step of, for example, oxide deposition. Such two steps oxide fill could be visible under proper magnification or other imaging techniques.

[00079] These layer transfer techniques could allow many of the benefits associated with monolithic 3D technologies including avoiding thermal budgets associated with forming one circuit strata affecting another circuit stratum, enabling mixing of technology nodes, mixing circuit substrate types, crystal structure, orientation and many other advantages associated with heterogeneous integration without process temperature restrictions described herein and in the incorporated art.

[00080] The use of SiGe for epitaxial based 'cut layer' instead of porous silicon or porous SiGe 'cut layer' could be adapted to many of the flows presented in at least U.S. Applications 14/642,724, 15/095,187, and 15/173,686, all the forgoing are incorporated herein by reference. It does add some complexity related to the holding posts formation and the holes to etch the SiGe thoroughly prior to performing the layer transfer. For applications in which two layers of active silicon, and isolation layer in between, is desired, the in-between SiGe could be removed after the transfer and replaced with isolation material.

[00081] Use of SiGe as a sacrificial layer for transferring a single crystal structure of one crystal on top of another structure has been presented in U.S. patent application 2015/0137187, incorporated herein by reference. Many studies of SiGe etch selectivity in respect

to silicon have been done and published such as: In a work by T.Salvetat *et al* titled "Comparison between three Si1-xGex versus Si selective etching processes" presented at 214th ECS Meeting; and by M. Stoffel titled "SiGe wet chemical etchants with high compositional selectivity and low strain sensitivity" published in Semicond. Sci. Technol. 23 (2008) 085021; by V. Destefanis *et a*l titled "High pressure in situ HCl etching of Si1−xGex versus Si for advanced devices" published in Semicond. Sci. Technol. 23 (2008) 105019; by T. K. Carns *et al* titled "Chemical Etching of Si,Ge in HF:H202:CH3COOH" published in J. Electrochem. Soc., Vol. 142, No. 4, April 1995; and by Marius Orlowski *et al* titled "Si, SiGe, Ge, and III-V Semiconductor Nanomembranes and Nanowires Enabled by SiGe Epitaxy" published at ECS Transactions, 33 (6) 777-789 (2010), all of the forgoing are incorporated herein by reference.

[00082] Another alternative is to skip steps related to Fig. 2D-2I and use the SiGe layer 204 as an etch stop, and after transfer grind and etch back the donor substrate 202, using the SiGe as an etch stop. And if desired, remove the SiGe layer 204 with an additional selective etch step, etching SiGe and very little or negligibly thin backside of the top silicon 206. The base substrate 202 could be about 775 microns thick while the SiGe 204 could be ten nm or less, or few tens of nm or even few 100 nm. For example, a 3D technique of flip bond and etch back of an SOI donor such as presented in at least U.S. patents 6,821,826, 7,723,207 and 7,312,487, all the forgoing are incorporated herein by reference. As an alternative to the use of SOI wafers, the base substrate 202 would not be reused but rather be ground and etched away from its wafer backside. The back grind and etch back could use wet etching and the SiGe layer 204 could be designed to be very resistive to the silicon wet etching. The SiGe could be designed to have multiple layers including one that might have high Ge content, for example, such as over about 20% or over about 40% or over about 80%, followed by other layers with low Ge content such as less than about 20% or even less than about 10% to reduce stress so to support the silicon layer 206. Fig. 1B illustrates an example of having the alternating Si and SiGe layers being comprised by multiple sub-layers with a varying content of Ge in the SiGe. Depositing or epitaxially growing the SiGe interface layer to the silicon with a smaller amount of Ge content decreases the stress due to the lattice mismatch. Then gradually increasing the Ge content to the desired level and then after the desired thickness has been grown gradually reducing the Ge content toward the next level of silicon could be used to reduce stress from both sides of the Silicon-SiGe-Silicon structure. Reducing the stress could help reduce the risk for formation of point defects and dislocations, and could help the engineering of the SiGe thickness as needed for the application. The alternative of use of SiGe as an etch stop later is further discussed herein later.

[00083] Alternatively, the 'cut' process could be integrated with Siltectra's 'Cold Split' technology as has been detailed in at least U.S. Patents 8,440,129 and 8,877,077, and U.S. applications 20160064283, 20160086839, all of which are incorporated herein by reference. These techniques would allow reuse/recycling of the donor wafer (base substrate 202A middle location inside SiGe or the interface between Si and SiGe could be used to provide the "Pre-Defined Break Initiation Point" as an alternative to the Siltectra use of laser or in addition to it. The Siltectra's 'Cold Split' could reduce the need for the undercut etch and posts formation processing while providing reuse of the base substrate 202. For this technique, a multilevel SiGe could be designed to support the 'cut' on the one hand but also to reduce damage to the device layer on the other. This could be accomplished by increasing the Ge content in the interface with the base substrate 202 to have high Ge content such as over about 20% or over about 40% or even over about 80% and then on the side interfacing with device layer 206 forming a low Ge content such as less than about 20% or even less than about 10% to reduce stress to the silicon circuit layer 206. Alternatively, a few atomic leyers thick Ge rich SiGe layer or even a pure Ge layer maybe used as a predefined break layer.

[00084] Once the base substrate 202 is removed, a selective etch could be used to remove the SiGe residues. Additional thinning processes such as etch and/or CMP could be used to further thin the back side of the device layer 206. Connection layers could be added included vias aligned to the target wafer 302 using "Smart Alignment" and similar 3D integration techniques discussed here and the incorporated by reference art.

[00085] This use of Cold Split could be used to form SOI wafers and could be less expensive to manufacture when compared to the current ion-cut methods.

[00086] A variation of flow in respect to Fig. 2A-2I, is to have the first set of holes 214 posts formed as part of the substrate process prior to desired circuits 212 processing. Accordingly, the flow in respect to Fig. 2E could be done to the donor wafer illustrated in Fig.

2C. These posts could be positioned at the dicing streets such as those between reticle projections, so they would not interfere with the future desired circuits 212. Alternatively post formation may be done to the donor wafer illustrated in Fig. 2B, which could then be processed with an epitaxial process which could be thick enough to fill-in the space on top of these posts allowing the following circuit processing without concern for the posts' locations. The number of posts per wafer could be 1-2, 4-10, 10-50 or over 100 first set of holes 214 posts per reticle. The diameter of these posts could be the size of vias for the designated process, or 50-100, 100-200, 200-400 nm or even larger. The material in these posts could be material compatible with these semiconductor processes and may be designed to be very selective with respect to the SiGe etch such as Silicon-Nitride, or any of the materials used for contacts such tungsten, or their combination, or, for example, copper, aluminum, titanium, tantalum, cobalt metals and/or silicide of the metals.

[00087] Another alternative is to use a similar flow to form a donor substrate which could support layer transfer as an alternative to an ELTRAN based donor wafer. This embodiment offers silicon on nothing structures anchored by post structures on the silicon ends. Then, the process follows using wafers with silicon on nothing. An exemplary donor wafer construction flow could include the steps A-F, referencing the illustrations in Fig. 4A-4H:

A. As illustrated in Figs. 4A and 4B, on a reusable donor wafer - base substrate 402 and epi layer 404 as a sacrificial layer may be formed, for example, by epitaxy processes. Epi layer 404 may be formed with a thickness of about 100 nm, or about 200 nm, or about 500 nm, or about 1000 nm or about 2000 nm, depending on process integration and architecture engineering decisions and tradeoffs. Epi layer 404 may include various materials, for example SiGe, in anticipation of a sacrificial etch later in the process flow. The content of Ge in the SiGe may be designed per the selectivity desired and in consideration of the stress. Predefined trenches (not shown) in the designated dicing streets (or other non-circuit areas, for example, such as the streets between projection fields) could be used to release the potential stress. These trenches could have a width and a depth approximately corresponding to the thickness of the epi layer 404 SiGe layer.

B. As illustrated in Fig. 4C, first epi silicon layer 406 as an active device layer may be epitaxially grown on top of epi layer 404. The thickness of silicon layer 406 may include a thickness of about 10nm, or about 20nm, or about 50nm, or about 100nm, or about 200nm, or about 500nm, or about 1000nm or about 2000nm, as desired for the electronic circuits, depending on engineering, materials and scientific device considerations. In many formations, first epi silicon layer 406 may be considered to consist of mono-crystalline or single crystal silicon.

C. As illustrated in Fig. 4D, first set of holes 414 may be formed to eventually be anchors of silicon on nothing structures (by, for example, a conventional mask and etch sequence of processing) and may extend through the first epi silicon layer 406 and the SiGe epi layer 404. These holes could then be filled with silicon nitride or oxide or other material that would be remained to future silicon and SiGe selective etches. Filled first set of holes 414 may serve as posts to hold the top layer designated to be transferred in later steps. First set of holes 414 may be located in the dicing streets area and may be designed to be weak enough to be torn out at the 'cut' step, by methods, for example, such as mechanical tear-off, edge and twist tear-off, water jet, according to design and process integration engineering considerations.

D. As illustrated in Fig. 4E, form multiple second holes 424 to expose the SiGe layer 404 and resultantly allow a full etch of the sacrificial SiGe layer (the former SiGe epi layer 404). These holes need to be open through the first epi silicon layer 406 all the way into the SiGe epi layer 404.

E. As illustrated in Fig. 4F, selectively etch the sacrificial SiGe layer creating void 426 underneath the first epi silicon layer 406. The suspended first epi silicon layer 407 may remain substantially flat for good bonding process.

F. As is illustrated in Fig. 4G perform additional epitaxial processing adding second epi silicon 432 (incorporated suspended first epi silicon layer 407 in drawing) and sealing second holes 424. The material and doping type of second epi silicon 432 may be the same or different from those of the first epi silicon 406 depending on the applications. The additional epitaxial width could exceed the second holes 424 radius to ease the holes sealing. A smoothing technique such as chemical mechanical polishing and H_2 annealing could be used to improve the top layer surface.

[00088] The donor wafer illustrated in Fig. 4G could then be used to process circuits 212, similar manner such as illustrated in Fig.2D and could form some interconnections 216 similar manner as illustrated in Fig. 2F, and then be flipped and bond on top of a

target wafer in a similar flow as is illustrated in Fig. 3A-3D, and many flows presented in the incorporated by reference material, such as U.S. 15/173,686 in respect to ELTRAN base donor wafer.

[00089] The silicon epitaxial layer 206/406 could be constructed from two layers such as first layer doped n+ followed by p- doped layer. Such double layer construction could allow smoothing of the surface 213 of the transferred layer after the transfer. Selective etch could etch the doped n+ layer leaving a smooth p- doped layer. Alternatively, the silicon epitaxial layer 206/406 could be made with three layers as is illustrated for example in Fig. 4H. First a p- layer 442, followed by n+ layer 444 and finally the upper most p- layer 446. The upper layer 446 could be used for the transistor layer, underneath it layer 444 which could support back-bias as suggested by Zeno Semiconductors and as published such as in papers by Jin-Woo Han *et al* titled "A Novel Bi-stable 1-Transistor SRAM for High Density Embedded Applications" published at IEDM 2015, and "A CMOS-Compatible Boosted Transistor Having >2x Drive Current and Low Leakage Current" published at ESSDERC 2016, incorporated herein by reference. And the bottom layer 442, as a sacrificial layer, to support smoothing, post transfers by selective etch as discussed above.

[00090] In U.S. patent applications 15/095,187 and 15/173,686 incorporated herein by reference, ELTRAN base layer transfer techniques are shown being adapted to support die to wafer 3D IC construction. Some of the die-to-wafer flows suggest transfer of dies having a relatively greater thickness such as 6 microns or even 20 microns and further thinning these dies after being bonded to the target wafer. Such die thinning could leverage a multi layers die structure. As an example, a multilayer such as is illustrated in Fig. 4H could be used. For such the bottom layer 442 could be silicon on top of it SiGe layer 444 and at the top silicon device layer 446. So the multilayered 440 could be 'cut' and bonded at a die level onto the target wafer, then a selective etch from the top could be used to first remove the silicon layer 442 and then thin the die all the way to the device layer by selectively etching the SiGe layer 444. These extra layers 442 and 444 could have a thickness of about 1 micron, 1-3 microns, 3-6 microns or even higher. The device layer 446 itself could have more sub-layers such as n+ and p- to support the back-bias scheme as discussed before. The use of multilayers, such as SiGe, allows having flexibility so a layer thickness is first set to support transfer at the wafer or die level providing the mechanical strength required for the handling and transfer, to be followed by selective etch trimming the thickness of the devices to support the electrical function and to allow forming via and other connection for following steps. These could be engineered by the artisan in the art. The 'cut' techniques could include selective under-etch or grinding and etch back as could be engineered for specific applications.

[00091] All these multilayer structures could be formed during the epitaxial growth by adding materials as gases to the epitaxial growth chamber as well known in the art.

[00092] These variations could be used for donor wafer substrate formation as discussed in reference to Fig. 4A-4H, or for transferable device layer as discussed in reference to Fig. 2A-2I.

[00093] As been stated before, a buried SiGe layer could be used as an etch stop layer. Use of buried SiGe as an etch stop layer to transfer a crystalline layer on top of another wafer structure has been presented in U.S. Patents 6,521,041, 6,689,211, 6,940,089, 7,348,259 and U.S. patent applications 2014/0342523, and in combination with ion cut in U.S. patent applications 2007/0023066, 2008/018959 all of the forgoing are incorporated herein by reference.

[00094] An additional alternative is to combine the porous formation technology of the ELTRAN based wafer transfer with the epitaxial ease of formation of silicon -SiGe technology base layer transfer presented herein. In U.S. patents 5,685,946 and 5,757,024 and in paper by Mondiali, V., *et al.* "Micro and nanofabrication of SiGe/Ge bridges and membranes by wet-anisotropic etching." *Microelectronic Engineering* 141 (2015): 256-260, all incorporated herein by reference, SiGe is shown to be stain etched forming a porous layer with about 100 to 1 selectivity with respect to silicon. Using this selectivity could allow forming a 'cuttable substrate' from the structure of Fig. 2C or Fig. 4C without the need of the posts formation of Fig. 2E or Fig. 4D. This concept could be applied to substantially all of the transfer flows presented herein including die level and wafer level transfer. The buried SiGe layer would become mostly porous Ge or porous SiGe layer and would withstand the following process temperature and other related process steps. For the 'cut' step the techniques presented in respect to the ELTRAN process could be used, such as mechanical tear-off, edge and twist tear-off, water jet, and extremely selective etch (being porous and Ge vs. silicon). This would simplify the layer transfer process and would allow substrate reuse, both cost saving opportunities.

[00095] The stain etch of a buried (Si or Ge or) SiGe layer converting it to porous layer could also be used when thermal isolation is required. Porous layers function well as a thermal isolation layer and oxidizing it could further add mechanical strength and further decrease its thermal mobility. Accordingly such a layer process could be useful in forming thermal isolation between the stratum of the 3D structure. So, for example, using SiGe as an etch stop could be followed by, instead of etching away the SiGe after grinding and etching of the silicon substrate, stain etching the SiGe converting it to a thermal isolation layer.

[00096] Converting the buried SiGe layer to porous layer by stain etching as presented above could leverage the STI etching step to use it as access to the buried SiGe layer or could include dedicated holes 224, 424, etching step. These access holes could be designed to provide access for effective conversion of the full buried SiGe layer underneath the die to SiGe. Such full buried SiGe conversion could be engineered based on the height of the buried SiGe layer the percent of Ge and other engineering aspects. As presented in U.S. patents 5,685,946 and 5,757,024 the side spread of the stain etching could extend to over 1 micron from the access holes. The engineering aspect of such full SiGe conversion could include Electronic Design Automation ("EDA") to support the design process to place these holes throughout the die surface to provide sufficient access to full SiGe stain etching. Such EDA support could include adapting the macro-cell library to include access for very large structures, providing a holes adder utility to add holes in 'white' spaces area that do not need holes or STI for the active circuits, and adding modules to the Design Rule Checker ("DRC") utilities. Similar type of EDA enhancements to support process modules, is common practice in the industry.

[00097] The layer transfer process could include two steps. First step could be performed at the frontend of the line of the process optionally as part of the STI process in which the SiGe layer is stain-etched converting it to a substantially porous layer. The second step could be performed just as before the layer transfer. In this second step, the porous-SiGe layer is selectively etched to make it ready for "cut" – detach. At that point the porous-SiGe could be selectively etched with extremely high selectivity. As discussed before, etching porous layer is about 5 orders of magnitude faster than etching the respective same material in full solid form. In addition the porous-SiGe is mostly Ge which could be extremely selective etched in respect to silicon. Accordingly the porous-SiGe could be etched with many orders of magnitude selectivity vs. silicon or other elements of the active circuits. In this case, the top silicon sidewall protection process described in Fig. 2H and Fig. 2I may be skipped. For the second step of weakening the porous-SiGe the required deep holes access could be designed with very low area loss. An example for such low cost access could be using the dice lanes. Deep etch of the dice lanes could help the layer transfer and detaching process as 'divide and conquer'. Additionally, the second step of weakening the porous-SiGe etch could be design to leave only small porous regions in the center of dies, keeping the dies in place for the bonding but make it easy to then detach the wafer leaving the active die bonded and the substrate detached and ready to be refreshed for reuse. An additional aspect of these porous-SiGe and related layer transfer techniques herein is an improved bonder machine which could include a detaching module. Such a detaching module could be a simple twist and pull apart, tearing off the substrate for reuse.

[00098] Additionally, a substrate 502 similar to one illustrated in Fig. 2C could be used in a conventional semiconductor fabrication process to process device layer contacts and potentially some connectivity 516 as illustrated in Fig. 5A. The SiGe layer 518 could be made as a buffer layer with multiple gradient Ge or as a uniform SiGe with Ge content of 15-20% or 20-30% or even higher. The circuit layer could be covered with oxide layer 517 and then planarized and made ready for being bonded to a target wafer 504 shown in Fig. 5B. The target wafer 504 may also be processed with the desired level of metallization. An oxide to oxide bond could be used to bond it to the target wafer 504 as illustrated in Fig. 5C. Other types of bonding could be used; such as metal to metal or hybrid bonding. Then a grind and etch could be used to remove the silicon substrate 502 from the bonded structure 522 leveraging the SiGe layer 528 as an etch stop resulting with the structure 524 as is illustrated in Fig. 5D.

[00099] Alternatively, the substrate 502 could be made with perforations similar to as been described in U.S. patent 8,273,610, incorporated herein by reference, in respect to at least Fig. 184 and Fig. 185. The SiGe layer could be etched or stain etched through these perforations allowing the detachment of the carrier wafer 502 for reuse.

[000100] Additional advantage of the techniques described herein is having the transferred circuit being an SOI circuit with its active silicon thickness to be fully depleted channel. The single crystalline silicon layer such as 530 could be made thin enough and being

bonded over oxide and covered with oxide effectively could provide the SOI functionality and if made thinner such as 10 nm provides FD SOI functionality.

[000101] Fig. 6A is a 'cuttable' wafer carrying circuits 610 such as memory control (periphery) circuits. Fig. 6D illustrates transferring the circuits 610 of Fig. 6A on top of the structure 604 of Fig. 6B, transferring the substrate 601 and partially the 'cut-layer' (SiGe) 603, and then removing base substrate 601 and partially the 'cut-layer' 603. An additional inventive embodiment is to optionally form a thermal isolation 608 on top of the memory matrix of target wafer 604 as is illustrated in Fig. 6B. In U.S patent 9,023,688, incorporated herein by reference, in at least Figs. 1-5 and associated specification sections, various thermal isolation layer techniques are presented to allow high temperature processing of the upper layer with minimal effect of the underlying circuit. These techniques could also be used to allow thermal isolation between one stratum and the other strata. Such could help isolate the operating temperature of one stratum so that it would not affect the operation of the other stratum. For example, the target wafer 604 could be primarily memory circuits while the transfer circuit 610 could be logic circuits that consume higher power during operation. The logic circuit might be generating heat and operate at a higher temperature that the memory circuit underneath; for example, a temperature difference of greater than about 20°C, greater than about 40°C, greater than about 60°C, greater than about 100°C. Accordingly forming a thermal isolation layer 608 in-between might help isolate temperature across these strata. Such isolation layer techniques could also include having layer 608 act as a sacrificial layer that could be etched under the bonding layer 606 in similar way to what have been described in respect to the SiGe layer 204 in reference to Fig. 2D to Fig. 2I. Etching away layer 608 results as is illustrated in Fig. 6C, could further increase the thermal isolation by providing an air gap isolation 630 between stratums. The etch technique could be such which forms a first porous layer which could be fully oxidized providing buried air pockets and thus almost an air gap level of thermal isolation. Using these techniques could allow forming a thermal isolation layer 630 between two strata having a thickness of few tens of nm or few hundreds of nm or even few microns, and having a low thermal conductivity such as less than about 1 W/mK or less than about 0.4 W/mK or less than about 0.1 W/mK or even less than about 0.05 W/mK. Alternatively, the isolation layer, for example, thermal isolation 608, 630 in between the target wafer 604 circuits and the transfer circuits 610 could comprise aerogel or high porosity dielectric as detailed in U.S patent 9,023,688 and could be made extra thick to further increase the thermal isolation, such as, for example, greater than about 100 nm, greater than about 200 nm, greater than about 400 nm, or greater than about 1 micron thick. In some cases, a thickness of 1-10 microns may be indicated, depending on design and engineering tradeoffs. Additionally, bonding layer 606 could include a heat spreader structure to reduce hot spots and further help protect the lower stratum from the upper stratum operating heat or vice versa. Fig. 6D illustrates a two strata 3D circuit 620 using techniques described herein. Furthermore, the through layer via ("TLV") used to connect the upper stratum to the lower one could utilize titanium to further reduce the heat conductivity between these strata. These TLV could use extra thick isolation to reduce thermal conductivity to the surrounding areas. Additionally, some of those TLV may be dedicated thermal TLVs wherein their main function is to serve as heat delivery channels to the heat sink and/or outside surface of the device. These even more important in respect to the TLV being used to conduct the heat from the operating circuit to the device external surface which is part of the device heat removal structure. Such extra thick lateral isolation could be greater than about 100 nm, greater than about 200 nm, greater than about 300 nm, greater than about 500 nm or more than a micron thick.

[000102] In a 3D system such as is illustrated in respect to at least Fig. 6D or Fig. 11D-11E, may include both a memory array and a logic circuit, the thermal isolation 1157, to keep the logic heat away from the memory array, could be placed in between the logic circuit 1156 (or logic layer) and the memory control circuit 1155 (or memory control circuit layer) as is illustrated in Fig. 11E. In general many vias, such as Thru Layer Vias (TLVs), may be connecting between the memory control circuit 1155 and the memory array 1131. These vias could connect to substantially every word-line, bit-line, source-line, and so forth. On the other hand the logic circuit 1156 could connect to the memory control circuit 1155 using address lines which represent far less connections (as an example - 10 address line could control, with proper decoder circuits, 1024 word-lines). Reducing the number of TLVs through the thermal isolation layer 1157 helps reduce the thermal connectivity through it and greatly improves the effective thermal isolation.

[000103] Formation of multiple levels of arrays of transistors or other transistor formations in the structures described herein may be described at least by the terms 'multilevel device' or 'multilevel semiconductor device.' Some examples of multilevel device may include memory device such as DRAM, SRAM, and Flash memory and image sensors such as CCD and CIS.

[000104] 3D devices could include redundancy for defect recovery in addition to redundancy techniques know for 2D devices. Such 3D devices could include one-time-programmable memory for at least packaged level memory repair. Such redundancy techniques and structure has been presented in U.S. patent 8,395,191, incorporated herein by reference, in respect to at least Figs 41, 86, 87, 114-132.

[000105] Additional variations of redundancy and repair techniques could be integrated within the 3D SRAM. 3D DRAM or 3D NOR fabric as detailed in PCT/US patent application 16/52726 and U.S. application 15/333,138, incorporated herein by reference. Hereinafter, the use of 3D NOR fabric in any of embodiment in this invention may be 3D SRAM fabric or 3D DRAM fabric unless otherwise specified.

[000106] Further inventive aspects follow. Fig. 7 illustrates a logic function of general 'look up table' and more specifically 3D architected look up table LUT-U 702 being processed together on top of logic function LUT-B 704, both sharing the same control gates and each may act as redundancy for the other, where -U and -B respectively denote upper stratum and bottom stratum, or upper portion and the bottom portion of the 3D structure. Fig. 7 illustrates that the LUT-U is included in the upper strata and the LUT is included in the bottom strata where two strata are constructed the mean explained in this invention. Placement of the stratum may vary due to engineering and design considerations. A programmable interconnect (I/C) structure 701 may include first I/C structure 706 and second I/C structure 707 which allow connecting either one of the logic function outputs, LU-Out 742 or LB-Out 744, to the primary output 724. Such programmable interconnect may be SRAM or latch. Alternatively, such programmable interconnect may be antifuse based OTP or other non-volatile memory such as RRAM, EEPROM, or flash, depending on the applications. The programmable interconnect may be implemented in the fabric layers in between the upper and the bottom portion and could be considered as a third strata or comprise a portion of a third stratum. Self-test could be used to choose which of the outputs to connect using the programmable connectivity structure first I/C structure 706 and second I/C structure 707. Redundancy could also be used for the interconnects such as is illustrated third I/C structure 714, fourth I/C structure 716, and fifth I/C structure 717, which would connect the LUT outputs to the secondary output 726. Another alternative is to have redundancy of the support circuits 712 on the bottom of the 3D NOR fabric as well.

[000107] Fig. 8 illustrates a redundancy scheme wherein a middle LUT-M 803 look up table functions to repair faults in either the right LUT-R 802 or the left LUT-L 804. The LUT-M, LUT-R, and LUT-L may be included at the same strata. Alternatively, LUT-M, LUT-R, and LUT-L may be placed in different strata of the 3D chip. The right LUT-R 802 and the middle LUT-M 803 could share the same control gates, and the middle LUT-M 803 and the left LUT-L 804 could also share the same gates disposed in the valleys of the corresponding ridge as is illustrated. Programmable interconnects structure 806 , 807, 808, 814, 816, 817, 818, 827, 828 allow connecting to selected LUT to the proper first primary output 824 or second primary output 826 to function as two complementing LUT functions as discussed in respect to Fig. 23 of PCT/US patent application 16/52726.

[000108] These redundancy schemes have the advantage that they provide the repair locally keeping the overall circuit functionality consistent with or without activating the redundancy function. These are especially important for logic circuit operation.

[000109] In these forms of redundancy read or write could be performed in parallel to two adjacent ridges once both ridge selects are enabled to accelerate the redundancy operation.

[000110] Fig. 9A and Fig. 9B are corresponding to the 3D memory fabric presented in Fig. 14B and Fig. 14C of PCT/US patent application 16/52726. They illustrate the designated 3D memory fabric such as 3D NOR structure for staircase per layer connection 902 and the staircase contact holes 906. An alternative technique for staircase formation is illustrated in reference to Fig. 10A-10K. This technique leverages selective etch rates between the layers forming the 3D multilayer substrate, such as in the case of 3D NOR are the source/drain (S/D) layers, that could include N+ type silicon, and the channel layers, that could include P type SiGe layers. Alternatively, the channel may be realized in the Si layers wherein the S/D substantially resides in N+ SiGe layers/regions. The selective etch could be designed to etch the S/D layer faster than the channel layer, for example, at a rate such as twice as fast or even

higher than 2x such as (2-3)x, (3-4)x or even higher. This etch process is designed to be an isotropic etch, wet or dry etch could be used with proper consideration for sizes. However, this does not preclude the use of other etches, for example, one that is primarily but not substantially isotropic, and so forth.

[000111] Multiple etch processes have been developed providing selectivity between the silicon etch rate and the SiGe etch rate including variations related to the proportional content of Ge in the SiGe. Such work was reported by V. Loup *et al* in work titled "Silicon and SiGe alloys wet etching using TMAH chemistry, published at Abstract #2101, 224th ECS Meeting; by Borel, S., *et al.* "Isotropic etching of SiGe alloys with high selectivity to similar materials." Microelectronic engineering 73 (2004): 301-305; and by Stephan BOREL titled "Control of Selectivity between SiGe and Si in Isotropic Etching Processes" published at Japanese Journal of Applied Physics Vol. 43, No. 6B, 2004, pp. 3964–3966, all incorporated herein by reference.

[000112] Fig. 10A illustrates the starting structure showing S/D layer that is to be converted into a structure for staircase per layer access region, having a top layer 1002 of mask, and multilayer ridge of S/D layers 1004 and channel layers in-between the S/D layers. The channel layer herein the per layer access region is a sacrificial layer that to be replaced by dielectric. The mask layer 1002 could be made to have a side etch similar to the S/D etch rate. The opening in the mask 1002 could be similar to the depth of the S/D layer or few times larger. The width could be similar to the width of the ridge.

[000113] Fig. 10B illustrates the structure after etching through to top most S/D layer through the opening in the mask.

[000114] Fig. 10C illustrates the structure after etching the top most channel layer through the opening in the mask and through the top most S/D layer. It could be seen the top most S/D layer could be etched about twice the thickness (Z direction) of the channel layer to the sides (X-Y direction) as the etching is isotropic and the etch rate of the S/D layer is twice higher than the etch rate of the channel layer.

[000115] Fig. 10D illustrates the structure after etching the second S/D layer through the top most channel layer. Note: These figures are not accurate and serve for illustrating the concept. Fig. 10K-Fig. 10N resulted from simulation and resemble the expected reality.

[000116] Fig. 10E illustrates the structure after etching the second channel layer through the formed opening.

[000117] Fig. 10F illustrates the structure after etching the third S/D layer through the formed opening.

[000118] Fig. 10G illustrates the structure after etching the third channel layer through the formed opening.

[000119] Fig. 10H illustrates the structure after etching the forth S/D layer through the formed opening.

[000120] Fig. 10I illustrates the structure after etching the forth channel layer through the formed opening.

[000121] Repeat the n-th S/D layer and n-th channel layer etch until the etch process reaches the desired bottom most S/D and channel layer.

[000122] Fig. 10J illustrates the structure after selective etch of the channel material in the structure and replacing the channel layer with isolation material. This process could accordingly form the desired staircase for per layer connection using one lithography step and multiple etch steps. Thus, top layer 1002 becomes etched top layer1003, and the various multilayer ridge of S/D layers 1004 become the etched version 1005. Fig. 10K-10N are process simulation charts illustrating such a stair-case formation process. Fig. 10K may be the starting point having a resist processed to form a lithographic defined hole for the center of the stair-case structure. Other lithographic shapes may be formed due to engineering and design considerations. Herein, the SiGe was assumed to be the channel layer and the Si was assumed to be S/D layer. Fig. 10L may be the first step of a multilayer etch using an isotropic etch having a SiGe:Si selectivity of 1:4. Fig. 10M illustrates the structure after the etch process has reached the base silicon going through five Si/SiGe layer pairs. The staircase is now formed. Fig. 10N illustrates the structure after removal of the photo resist and an optional cleaning etch with 100:1 SiGe:Si selectivity. The specific etch selectivity may be adjusted based on engineering and design considerations. Many variations of this concept could be considered by an artisan in semiconductor process including changing the etch process between these layers or even layer replacement techniques such as replacing the channel layers. Controlling the isotropic etch of the S/D strips in the ridge allows forming the desired staircase. Many options of materials which have different etch rates are well known in the art. An example for such could be found in a paper by Kirt R. Williams *et al*, titled "Etch Rates for Micromachining Processing" published at JOURNAL OF MICROELECTROMECHANICAL SYSTEMS, VOL. 12, NO. 6, DECEMBER 2003 incorporated herein by reference. The staircase formation techniques presented in respect to Fig. 10A-10N could be adapted by an

artisan in semiconductor processing to many of the 3D memory structures including structures that use multilayers of poly over oxide or tungsten over oxide.

[000123] Fig. 11A-Fig. 11D correspond to Fig. 35A –Fig. 35D of PCT/US patent application 16/52726. It is used to illustrates 3D system which includes 3D-NOR fabric 1130 over 'cut layer' 1113 over substrate 1110, and brought over from another substrate 1150 and cut layer 1143 periphery circuits 1154, and brought over additional processing logic 1156 enabling such as "Processing-in-Memory", or "co-processor in memory" or "function in memory". Such memory centric architecture may be useful compared to the conventional processor centric architecture for future machine learning, neural network, approximate computation and so forth. As discussed there many system options could leverage such 3D system 'Lego'. A generic memory array could be diced using predefined potential dice lines or etch defined dice line to fit specific overlay logic, the overlay logic itself could be broken to more than one stratum, and could include dedicated stratum for I/O, cut from another generic wafer. Additional'Lego' options could be available from the 3D NOR fabric. One such option is segmenting and allocating the number of layers into the 3D NOR fabric and the amount of memory and programmable logic accordingly. Fig. 11D illustrates one type of 3D NOR fabric 1130 with control circuit 1154 and processing logic on top 1156. Alternatively another 3D system could be construct with similar "Lego" parts such as the processing logic 1156 with modified control circuit 1155 and a 3D NOR fabric with much fewer number of layers 1131. Another type of system variation using similar 'Lego' parts is to modify within the 3D NOR fabric where the thickness of the tunneling oxide is processed differently for sets of words line groups and accordingly creating variations with the memory type in the 3D system between retention time and access time. Another type of structure variation that could be applied is memory usage and the corresponding write and read techniques such as multilevel and mirror-bit tradeoffs between access time and density. Another type of modification that could be formed in the 3D system is the allocation of the fabric to memory versus to programmable logic. These variations could form completely different end 3D system with lower investment in new masks and higher leverage of the volume of generic product produced in the process line.

[000124] An additional inventive embodiment for such a 3D system as is illustrated in at least Fig. 11D, is to add additional layer on top to provide an electro optics circuit which could allow use of light, such as, for example, fiber-optics or photonic components, to communicate with other systems and for other systems to communicate to the 3D system illustrated. Alternatively, the top layer of the 3D system may be image sensors, hyperspectral sensor, or time-of-flight sensor. Such a 3D system could comprise a memory fabric such as 3D NOR fabric 1130 which could be allocated as sub memory sections, some for high speed and some for high density, and a compute circuits within control circuit 1154 that could have many cores and control circuits designed for communications and control of the underling memory fabric (such as 3D NOR fabric 1130), and communication layer within processing logic 1156 that could be made with material optimized for those tasks, such as those optimized for RF.. Such a 3D computer could be made to support very efficient programming as all its internal routing are far shorter compared with current Printed Circuit Board (PCB) technique of computer system integration or even the2.5D/3D initiatives.

[000125] Such heterogeneous 3D integration allow the use of one type of fabrication facility for one of the strata, for example, a memory oriented fabrication facility to produce the memory array 1130, and very different facility for a different stratum, for example, such as a logic orientated fabrication facility producing the memory control circuit 1154, thus allowing an increased flexibility in the design of the overall system including use of much more advance fabrication lines for some of the stratum.

[000126] Use of the alignment technique we call 'Smart Alignment' allows connection between the upper strata and the lower strata with vias (Through Layer Via – TLV) that are as small as the thickness of the layer and the process capabilities allow. Such is useful for connecting memory control circuits in one stratum to memory control lines such as bit-lines and word-lines on the other stratum.

[000127] Fig. 11F illustrates the "smart Alignment' technique. The target wafer 1164, such as illustrated in Fig. 11B, could have its alignment mark 1162 and a connection strip target 1160 along direction 'X' 1190, having a length longer than the worst-case misalignment in 'X' direction of the wafer bonding (and subsequent release, planarization, and cleaning process). The transferred layer or wafer 1170, such as for example control circuit 1154 illustrated in Fig. 11A-11C, may have its own transferred layer alignment mark 1166. The intersection between the horizontal connection strip target 1160 and the designated upper stratum vertical connection strip 1172, which could be aligned to the transferred layer alignment mark 1166, is now known and will be processed to be

locating the TLV_{21} 1168. The processed vertical connection strip 1172 length should be designed to be longer than the worst-case misalignment in 'Y' direction of the wafer bonding. Accordingly, the via 1168 could be aligned, to target wafer alignment mark 1162 in the Y direction and to the transferred layer alignment mark 1166 in the X direction. Once the via layers are formed an upper layer metal mask aligned to the upper layer alignment mark 1166 could be used to form connection between the transferred upper stratum and the target stratum underneath. Fig. 11G illustrates connecting the upper stratum strips 1178 to memory control lines 1176 in the target wafer using the 'smart alignment' technique individually connecting to each control lines even for bonding misalignment which far exceed the memory control lines pitch.

[000128] Yet the target wafer 1164 in most cases of memory array would have at least two set of control lines one in X direction and one in Y direction. To allow effective connectivity the 'Smart Alignment' technique could be enhanced to have two sets of TLV. One TLV_{21} aligned to target wafer's alignment mark 1162 in Y direction and to the transferred layer's alignment mark 1166 in X direction. And the other TLV_{12} aligned to target wafer's alignment mark 1162 in X direction and to the transferred layer's alignment mark 1166 in Y direction. This may require two step of lithography. Fig. 11H illustrates such two connection sets. Bit-lines in X direction 1180 connected to upper stratum by strips in Y direction 1182 while word-lines in Y direction 1184 connected to upper stratum by strips in X direction 1186.

[000129] An additional inventive embodiment relates to monolithic 3D by layer transfer whereby a unique structure may be formed by replacing silicon with high quality oxide prior to the layer transfer at the time that high temperatures processes are acceptable. For example, the silicon in the zone 1179 that is being designated for TLVs may be etched and filled in with high quality oxide (or a lower quality oxide deposition followed by a high temperature anneal) that would have leakage current of less than one picoamp per micron at a device power supply voltage of 1.5 and at a measurement temperature of 25°C. Thus, as well, the TLVs would not require any insulative lining to pass thru the TLV transiting layer, which could be islands/mesas of silicon in a sea of oxide, or vice versa.

[000130] An alternative to, two lithography steps with two via masks, could be the smart use of direct write eBeam in which the eBeam alignment could be managed to provide proper placement for the TLV_{12} and TLV_{21}.

[000131] In some applications, it could be desired to transfer stratum including interconnection performing what could be called parallel integration instead of sequential integration. Bonding layer or die in such case could utilize hybrid bonding forming bonding and direct metal to metal connection in the process. In general, such hybrid bonding utilizes connection pads that are large enough to accommodate the bonding misalignment which in advanced bonder is approaching 100nm worst case misalignment. Yet some memory stratum might use control line pitches which could not accommodate the bonder misalignment. An alternative for such cases could be use of bonding oxide that could be made to conduct by electrical signal, using what is known as One Time Programmable – "OTP" or Resistive RAM technologies. In such case one stratum could have some control signal and power signal connected using the hybrid bonding while the memory control lines could be connected by programming.

[000132] Fig. 11I illustrates a section of memory control lines in one stratum. Fig. 11J illustrates connection segment on the other stratum. For the case in which the memory stratum is the one under the connection structure of Fig. 11J includes connection strips in Y direction 1192 that are formed long enough to cover the bonding misalignment in Y direction, connected by via 1193 to the X direction connection wires 1194 on the upper layer. Fig. 11K illustrates these connection structures after the bonding. For misalignment of less than three memory control lines pitch, three independently controlled programming signal VP1, VP2, VP3 could be used to form the connection between the memory stratum and the logic stratum. These could be provided using properly arranged and connected diodes. Other arrangements could be designed and engineered.

[000133] Fig. 12A and Fig. 12B correspond to Fig. 13E and Fig. 14D of PCT/US patent application 16/52726. Fig. 12A illustrates an optional alternating allocation of ridge select transistors1202, 1203 formed in series of the S/D lines on both ridge ends. The ridge select transistor 1202, 1203 can allow selectively the specific S/D lines. Fig. 12B illustrates the 'Y' direction per S/D layer staircase structure 1204. Fig. 12C illustrates X-Y cut (top view) of such alternating ridge select 1213 (RS1, RS3, RS5), 1222 (RS2, RS4) with 'Y" oriented common staircase access on both sides 1214, 1224. The ridge such as 1216 may be associated with multiple word line controls (w1, w2, …) which could be consider first gates w-1, and second gates w-2 and controlling channel from both sides of the ridge, odd side w-1o, w-2o, and even side w-1e, w-2e. Fig. 12D illustrates similar structure with a perspective 3D illustration of Fig.

12C except that the select transistors are staggered on left and right end of ridges. Such arrangement of staggered select gates on left and right end of ridges allows sufficient space for the per ridge select gate (without affecting the adjacent ridge). Fig. 12E illustrates a variation on Fig. 12C in which the side of ridge with no ridge select 1217 is directly connected to the respective 'Y' direction staircase structure 1214, yet isolated from the ridge select of its adjacent ridge 1233. The isolation layer 1233 may be a relatively thick oxide in order to minimize electrostatic coupling from the gate and the ridge across the isolation layer 1233. So, the select gate would only control a ridge on its relatively thin oxide side. Such arrangement of select gate enables equivalent function of Fig 12C. The variation of Fig. 12E would allow access to each cell from both ridge ends, as there are both still connected to the respective staircase access 1214, 1224.

[000134] In consideration of the 'smart alignment' as was previously discussed in respect to Fig. 11H, the 3D NOR fabric as is illustrated in Fig. 12B could be modified to allow easier connection to upper stratum accommodating bonding misalignment in X and Y direction. Fig. 12F illustrates such accommodation in modifying the structure of Fig. 12E. Two global wordlines (WL) are assigned for one local WL pitch; one for the odd side and another for the even side of a single ridge. The global word-lines WL1 1236 and WL2 1238 are oriented in the Y direction and correspond to Y direction word lines 1184. Alternatively, no global WL is used but each vertical local WL may be directly accessed by a 3D stacked peripheral circuit on its top. The ridge select transistor 1228 could be modified as illustrated in Fig. 12F to be extended in the X direction 1234 to accommodate X direction misalignment. Thus the ridge select transistor 1228 may be a long channel transistor, which improves the leakage behavior for an unselected ridge. The per level via 1205 of Fig. 12B could be extended in the Y direction 1232 to accommodate Y direction misalignment. Fig. 12G is a 3D perspective illustration with extended gate length ridge select transistor 1244 corresponding to 1234. Fig 12H illustrates the addition of connection vias, and Fig. 12I illustrates adding the word lines 1246 and the extended per S/D layer connections 1242. Fig. 12J illustrates side view of the extended per S/D layer connection. Precise wafer bonders available in the industry are proven and capable of bonding misalignments of less than 100 nm. Accordingly the structures to support such a bonder could be made with connection pad of about 100 nm which could be used for ridge select pad 1234 to be about 100 nm in the X direction and the per layer pad 1232 to be about 100 nm in the Y direction.

[000135] Placing the staircase along the Y direction vertically aligned to the ridge direction reduces the area overhead associated with per layer access. When combined with the 3D integration of memory control circuits transferred and added on top and/or under the memory matrix, this supports an array built from many micro-arrays each with its own memory control circuits. Such a micro array or unit could have X direction and/or Y direction size of few tens of microns, or few hundreds of microns. This arrangement reduces the capacitance and resistance of the memory control lines thus allowing lower power and higher speeds for the memory device.

[000136] Additional variation for the 3D NOR fabric is to use SiGe with 5% Ge content for the S/D layers and SiGe with 20% or higher content for the channel layers to achieve higher growth rate for the epitaxial growth of the multilayer structure which could be used for cost reduction or alternatively reduce epitaxial process temperature with lower effect on the growth rate.

[000137] Additional alternative for the 3D NOR fabric is to utilize charge trapping also as a variable memory function for brain type functionality in a similar fashion to what have been proposed in PCT/US2016/052726 for the RAM part of the fabric.

[000138] Fig. 18 of PCT/US patent application 16/52726, presents a programming table for the 3D NOR memory. It utilizes what is called Hot-Electron injection. Fig. 13 of that same application depicts table for Flash memory programming. It presents Fowler-Nordheim tunneling ("FN") which is a common flash programming and erasing technique and could allow lower energy per bit writing and erasing. Use of FN programming for NOR type flash memory is presented in a paper by Masayoshi Ohkawa *et al* titled "A 98 mm^2 Die Size 3.3-V 64-Mb Flash Memory with FN-NOR Type Four-Level Cell" published at IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 31, NO. 11, NOVEMBER 1996, incorporated herein by reference. The use of FN tunneling as a programming mechanism may save write energy when compared to the less efficient hot-carrier programming. The FN tunneling can be bit-specifically conducted for the presented 3D-NOR architecture. In an unselected ridge, the ridge-select gates are turned off so that the S/D lines become floating so to inhibit writing (programming) of those ridge memory cells. In a selected ridge, the ridge-select gates are turned on so that the S/D voltages for writing or reading can be applied through it. For the selected cell in the selected ridge, a high enough programming voltage is applied to the selected wordline and a pair of S/Ds could be connected to a lower

voltage, such as grounded, so that there is sufficient electric field across the selected S/D line and the selected wordline to effect charge transfer into the charge trap layer or into the floating gate. For the unselected cell in the selected ridge, a disabling high enough positive voltage could be applied to all unselected S/D lines to reduce the electric field across the unselected memory cells, and thus minimize program disturb. Note, under typical FN voltages wherein the WL voltage for writing is assumed to be 20V, the inhibit voltage for the non-selected cell S/D could be ~10V. The half-selected cells are those cells that are sharing one of the selected S/D lines and thus may suffer from disturb. The S/D lines could be designed to withstand such voltage differential to avoid source-to-drain breakdown or punch-through from selected S/D line to the adjacent unselected S/D line Alternatively, the method is suitable for lower PGM (programming) voltages (assuming ~ 5V), possibly adequate for thinner O/N/O. Mirror bit techniques might not be useable with FN tunneling or direct tunneling programming and some other disadvantages could be associated with such FN or direct tunneling programming.

[000139] Despite the term "FN tunneling" solely appearing throughout here, it should be understood that it is not intended to differentiate FN tunneling from direct tunneling, rather there is also contribution due to direct tunneling as well as FN tunneling. Unless otherwise specified, the sole use of "FN tunneling" herein contains the meaning of direct tunneling.

[000140] Fig. 13A illustrates a small section of the 3D NOR structure which could assist in presenting alternative writing and erasing schemes of memory cells within the 3D NOR memory. Illustrating two columns, each with its local word-line, WL1 1304, WL2 1306, with 8 S/D lines 1311-1318, and Ridge Select gate 1302.

[000141] The entire ridge could be Fowler-Nordheim ("FN") erased by selecting the ridge and either by grounding all S/D lines and channel while powering all the word-lines with negative erase voltage –Ve, or by positively biasing the channel while grounding all S/D lines and WL's. Optionally, positively biasing S/D lines and grounding WL and Channel lines can be carried out as well. As discussed herein before, specific voltages could be set for specific structures. For example, -5 volts could be used for FN erase for -Ve, and +5 volts for Ve. For use of the channel access for any programming or erasing explained, the body contact structure needs to be disposed, which will be explained later herein.

[000142] Alternatively, the entire ridge could be hot-hole erased, by selecting the ridge and grounding one S/D line from S/D pair and positively biasing the other S/D line while powering substantially all the wordlines with a negative erase voltage and optionally grounding substantially all channel lines. In this case, the hot-hole erase favorably occurs near the positive biased S/D region. Therefore, such hot-hole erase may be used to one-side ridge erase in a mirror-bit memory system. In order to completely erase both the source side and the drain side storage nodes using hot carriers, the erase may be accomplished by two steps; one-side erasing followed by opposite-side erasing by swapping grounded S/D line and positively based S/D line.

[000143] A specific column within the ridge could be FN or direct tunneling erased by having all the other word-lines at the same voltage as the S/D lines while the selected column word-line would be at negative differential Ve. For example, such as grounding all the S/D lines (and channel 'body' if such is used in the structure) and all the non-selected word-lines and driving the selected column word-line with an erase voltage such as -5 volt in some specific structures.

[000144] Fig. 13B illustrates FN erase condition for one cell 1320. That cell word-line could be set to -5v while the other column word-lines are kept at ground. The S/D line and channel 'body' of the selected cell could be grounded while all other S/D and channel lines are left floating. In such FN single cell erase the adjacent cells (the half selected cell) with the column 1322 might be disturbed and partially erased. All the other cells 1324 should be unaffected. The partially erased cell could be refreshed to their original state by per cell write steps. Alternatively, the use module could be utilized only unaffected cells such as using only the odd level cells. This would reduce the memory density while simplifying the operation of FN per cell erase.

[000145] The inverse of FN erase could be used for FN write (programming) so instead of negative differential using a positive differential such as +5v instead of -5v of Fig. 13B. Accordingly the above discussion would apply to FN writing.

[000146] Many variations known for Flash memories could be applied and adapted to the 3D NOR fabric. Such could be reversing the gate stack order by having the tunneling oxide between the gate and the charge trap layer. Such could be used for FN write and erase from the word-line to the charge trap or the floating gate. Other variations could be engineered by memory artisan using the techniques known in the flash memory art.

[000147] In general FN is known to be few orders of magnitude more efficient in power then hot electron techniques, and accordingly preferred for many applications. Yet, the additional benefit of using Schottky Barrier techniques is its highly efficient hot electron programming, which could be effective enough for many of these applications. Additionally the use of SiGe for the channel with N+ silicon for the S/D, could provide an additional enhancement to secondary hot electron injection as presented in a paper by Kencke, D. L., *et al.* "Enhanced secondary electron injection in novel SiGe flash memory devices," Electron Devices Meeting, 2000. IEDM'00. Technical Digest. International. IEEE, 2000. This mechanism combines both the secondary electron injection with the smaller bandgap to generate higher impact ionization rates and subsequent electron injection probability. The proposed mechanisms may significantly reduce maximum voltages required for the program operation applied to WL, S/D and channel lines.

[000148] The 3D memory structures shown in Fig. 13C and 13D may include floating-body devices, so the body potential may be strongly capacitively coupled with the S/D line voltage. So, it is assumed that when both S/D line voltages are positively raised up, the body potential is also raised up accordingly. As a result, no sharp energy band bending near P-N junction is made to cause the band-to-band tunneling. It is assumed that the minimum voltage for FN tunneling erasing across the gate and the S/D line was -5V as shown in Fig. 13B. Therefore, the voltage difference of -4 V across the gate and the S/D lines shown in Fig. 13C and Fig. 13D most likely will not cause the FN tunneling erasing. However, the techniques illustrated in Fig. 13C and 13D are intended for hot-hole injection caused by a potential gradient induced by asymmetric voltage for one grounded S/D line and another partially positively biased S/D line along with a partially negatively biased gate.

[000149] Fig. 13C illustrates Gate Induce Drain Leakage ("GIDL") erase for one cell 1330 within a selected ridge 1332 wherein channel lines are floating. The selected cell 1330 Source line 1352 could be grounded (0 volt), its Drain line 1353 could be powered with positive voltage such as 2v, and its word-line 1334 is connected to negative erase voltage such as -2v. The differential voltage between the positive Drain line and the negative word-line (-4v) should be below the voltage inducing FN erase. The specific voltage could be set for specific device as previous discussed. The non-selected word lines 1336 could be grounded in order to disable GIDL current. The S/D lines above the selected cell 1351 could be grounded and the S/D lines below the selected cell 1354-1358 could be powered to positive 2v. The non-selected cells 1344 could therefore see their Source and Drain at same potential and accordingly no current is induced to the channel and accordingly non creation of hot-holes. And having the voltage difference across a gate and S/D line below the threshold level of FN tunneling should keep the trap charge with no change. In the selected cell the differential voltage between its Source and Drain could induce leakage under the negative gate bias, such leakage could form hot-holes, and the negative field of the word-line 1334 will pull these holes into the charge trap (or floating gate) layer to erase its stored electrons charge.

[000150] Using positive voltage (+2v) for the selected word-line 1334 of Fig. 13C could allow writing in the charge trap (or floating gate) layer of the selected cell 1330. Having a positive gate would actually open the selected cell transistor strongly increasing the transistor current and accordingly the formation of Hot-Electron to be pulled in by the positive word-line voltage. Using the up side flip similar to Fig. 13D could change the location of trap charge. Note, the above mentioned voltages are extremely small for typical NVM device and may be obtained if extremely thin ONO thickness is selected, such as bottom-oxide of 1 nm, nitride thickness of 2 nm and top-oxide such as 2 nm. Further, for CHE programming the drain to source voltage must be typically larger than ~3V. Other mechanisms such as SB injection may require lower drain voltage (absolute value), such as -2V.

[000151] An alternative writing and erasing could be performed by having the non-selected S/D lines floating. In order to use such a scheme, the floating body 3D structure is preferred as the body region can be also floated according to the floated S/D.

[000152] As stated the specific voltages for write erase of selected cell and not affecting of non-selected cells could be tuned for specific devices. These tuning could include the presented write and erase techniques and variations of those as known in the flash memory art including mixing techniques such as FN and Hot-Electron/Hot-Holes.. Many variations are known in the art and could be adapted for the 3D NOR memory. Such as those presented in U.S. patent 7,113,431, 7,220,634, 7,590,005, 8,183,616 and applications 2006/0125121, 2009/0086548, 2011/0095353 and 2012/0231613, incorporated herein by reference. And papers such as by Lei SUN et al. titled "Characteristics of Band-to-Band Tunneling Hot Hole Injection for Erasing Operation in Charge-Trapping Memory" published at Japanese Journal of Applied Physics Vol. 45, No. 4B, 2006, pp. 3179–3184, by Alvaro Padilla et al titled "Enhanced Endurance of Dual-bit SONOS NVM Cells using the GIDL Read Method" published at 2008 Symposium on VLSI Technology, by

Kyoung-Rok Han et al titled "5-bit/cell Characteristics using mixed program/erase mechanism in recessed channel non-volatile memory cells" published at Current Applied Physics 10 (2010) e2–e4, by LIU LiFang et al, titled "A 1G-cell floating-gate NOR flash memory in 65 nm technology with 100 ns random access time" published at SCIENCE CHINA, Information Sciences April 2015, Vol. 58, by Yu Wang et al titled "A 65-nm 1-Gb NOR floating-gate flash memory with less than 50-ns access time" published at Chin. Sci. Bull. (2014) 59(29–30):3935–3942, by Ken Uchida et al titled "Enhancement of hot-electron generation rate in Schottky source metal–oxide–semiconductor field-effect transistors" published at Applied Physics Letters 76, 3992 (2000); by Kyeong-Rok Kim titled "Design of NOR flash memory cells, with high speed programming by utilizing an asymmetric Silicide(TiSi2) Drain" published at ICASIC07, by E. J. Prinz et al titled "90nm SONOS Flash EEPROM Utilizing Hot Electron Injection Programming and 2-Sided Hot Hole Injection Erase" published at NVM Workshop 2003, by Li-Jung Liu et al titled "Performance enhancement in p-channel charge-trapping flash memory devices with Si/Ge super-lattice channel and band-to-band tunneling induced hot-electron injection" published at Thin Solid Films 533 (2013) 1–4, by Choi, Sung-Jin, *et al.* "A novel TFT with a laterally engineered bandgap for of 3D logic and flash memory," published at VLSI Technology (VLSIT), 2010 Symposium, IEEE, 2010, and by Yu-Hsien Lin et al, titled "Band-to-Band Hot Hole Erase Mechanism of p-Channel Junctionless Silicon Nanowire Nonvolatile Memory" published at IEEE TRANSACTIONS ON NANOTECHNOLOGY, VOL. 15, NO. 1, JANUARY 2016, all incorporated herein by reference.

[000153] The silicidation process suggested in PCT/US patent application 16/52726 and U.S. application 15/333,138 could be done by known technique in the art. These could include two annealing steps. First to form the metal penetration into the silicon followed, then removal of the excess metal by dry or wet etch, followed by second anneal to activate the silicide metal. These processes could include microwave annealing which was demonstrated to allow reduced temperature. Such process has been presented in papers such as by Xiangbiao Zhou et al titled "Schottky Barrier Height Tuning via Nickel Silicide as Diffusion Source Dopant Segregation Scheme with Microwave Annealing" published at 15th International Workshop on Junction Technology (IWJT), by Shih, Tzu-Lang, and Wen-Hsi Lee. "High Dopant Activation and Diffusion Suppression of Phosphorus in Ge Crystal with High-Temperature Implantation By Two-Step Microwave Annealing." ECS Transactions 72.4 (2016): 219-225, by Chun-Hsing Shih et al titled "Metallic Schottky barrier source/drain nanowire transistors using low temperature microwave annealed nickel, ytterbium, and titanium silicidation" published at MSSP 16, by Sounak K. Ray et al titled "Enhanced charge storage characteristics of nickel nanocrystals embedded flash memory structures" published at Journal of Experimental Nanoscience, 2013 Vol. 8, No. 3, 389–395, by Chaochao Fu et al titled "Schottky Barrier Height Tuning via the Dopant Segregation Technique through Low-Temperature Microwave Annealing" published at Materials 2016, 9, by Jian Deng et al titled "A modified scheme to tune the Schottky Barrier Height of NiSi by means of dopant segregation technique" published at Vacuum 99 (2014) 225e227,by Y.-J. Lee *et al.* titled "Record-Thin 10.5nm Ni Silicide Film for 2012-2021 by Two-step Low Temperature Microwave Anneal" published at IEDM 11; Y.-J. Lee *et al.* titled "Full Low Temperature Microwave Processed Ge CMOS Achieving Diffusion-Less Junction and Ultrathin 7.5nm Ni Mono-Germanide" published at IEDM 12; by Y.-J. Lee titled "A Novel Junctionless FinFET Structure with Sub-5nm Shell Doping Profile by Molecular Monolayer Doping and Microwave Annealing" published at IEDM 14; by Y.-J. Lee *et al.* titled "Low-Temperature Microwave Annealing Processes for Future IC Fabrication—A Review" published in IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 61, NO. 3, MARCH 2014; by T. Yamaguchi *et al.* titled "Low-Resistive and Homogenous NiPt-Silicide Formation using Ultra-Low Temperature Annealing with Microwave System for 22nm-node CMOS and beyond," published at IEDM 2010; by Ming-Kun Huang *et al.* titled "Dopant segregated Schottky barrier nanowire transistors using low-temperature microwave annealed ytterbium silicide" published in Japanese Journal of Applied Physics 53; and by Ming-Kun Huang *et al.*, titled "Dopant segregated Schottky barrier nanowire transistors using low-temperature microwave annealed ytterbium silicide" published in Japanese Journal of Applied Physics 53, 116501 (2014), all of the forgoing are incorporated herein by reference. Equipment for Microwave Annealing has been offered by DSG technologies of California. An additional advantage for use of Microwave Annealing and ALD is the compatibility of these processes with 3D structures. For example, the laser annealing might deliver non-uniform energy along the depth of the multiple stacks of semiconductor layers, which may result in non-uniformity. Microwaves could be designed to penetrate inside the semiconductor so that the annealing result may be uniform across the multiple stacks of semiconductor layers. In addition, the metal ALD may further enhance the metal silicidation, as the metal ALD can enable a deposit of a precise amount of metal to be used for the

silicidation. Consequently, no residual metal unreacted with the silicon may left and thereby no subsequent residual metal removal may be necessary. In some embodiments, the exact amount of the layered metal may be fully consumed during the silicidation process, which could save the subsequent removal process of non-reacted metal. Such could be also effective for additional applications such as FinFET and nano-wires providing additional advantage to reduce Schottky Barrier variations.

[000154] Fig. 14A presents a drain current versus gate voltage (Id-Vg) curve taken from Fig. 8 of the paper by Ming-Kun Huanget et al. It illustrates the advantage of microwave annealing of a silicided nano-wire transistor. The chart illustrates the advantage of Dopant Segregation for Schottky Barrier in respect to reducing ambipolarity without sacrificing the drive current. Yet in some alternatives a 3D NOR memory structure could be made with transistors having some level of ambipolarity. In such cases the sense amplifier could be made to support improved memory function. Assuming that the chart related to Schottky Barrier 1402 represent the memory cell of such 3D NOR. Assuming that the VT shift effected for '0' state is 1v while '1' state is as plotted in Fig. 14A. The sense amplifier for read circuit could be made to senses the differential drain current for gate voltage of 2v and 0v. For state '1' such differential current could be about 10^{-6} A, while for state '0' the differential current could be about 10^{-7} -8*10^{-8} ~ 2*10^{-8} A. Accordingly such differential sense amplifier could enable range of memory transistors to effectively function.

[000155] Fig. 14B is an Id-Vg curve taken from Figure 11. (b) of a paper by Liu, Yongxun, *et al.*, "Comparative Study of Charge Trapping Type SOI-FinFET Flash Memories with Different Blocking Layer Materials," Journal of Low Power Electronics and Applications 4.2 (2014): 153-167, incorporated herein by reference. It illustrates that in some memory applications a self-reference differential sense amplifier could be an effective technology for sensing the memory information. For example, if the Id differential read is determined by two step gate biasing between Vg_2=4v and Vg_1=0.5v, then in the erase state the differential Id is relatively large compared to the programmed state. Accordingly such self-reference differential sense amplifier could be a useful memory sensing technique for memory structure presented herein and elsewhere. Setting the gate voltages for such detection technology could be so the memory transistor threshold voltage (Vt) is in the voltages range (Vg1 <Vt< Vg2) for one memory state and out of it (Vg1 >Vt or Vt>Vg2) for the other. Such self-reference differential sensing could be effective in reducing the effect of absolute Id current value that is not related to the change in that cell wordline voltage.

[000156] An additional inventive embodiment is to reduce the ambipolarity effect by using a relatively reduced drain to source voltage (V_{DS}) such as 0.5-1 V during the read operation (for example, as compared to state of the art Mirror-bit technology of 1.4-1.6 V). Lowering drain voltage can suppress the band-to-band tunneling leakage causing ambipolarity. Thus, effectively reducing the ambipolar current of the unselected bitline cell during the reading of the selected wordline cell. The reduced V_{DS} may still distinguish between two physical bits per cell such as mirror bit scheme in reverse and forward read operations when a small Vt shift between the programmed and the erased states such as about 1 volt may be implemented and as a narrow charge distribution is obtained using SB or DSSB injection compared to Channel Hot Electron injection in doped np junction. Furthermore, the reduced V_{DS} improves immunity to read disturb and extends the retention time and endurance of the memory transistor.

[000157] The Dopant Segregation Schottky Barrier ("DS-SB") formation process suggested in PCT/US patent application 16/52726 in respect to at least Fig. 8E and 41E, could be modified to allow broader device control. Fig. 15A illustrates a Y-Z cut-view of two level 3D NOR memory (resembling Fig. 2 of 16/52726). On top of a section of a carrier wafer 1500, illustrated are side view of two ridges, having two layer of memory having channels 1502, 1504 each between two S/D regions 1501, 1503 and 1506. A hard mask portion on top 1508 that could have been used for the etch mask forming these ridges. The S/D regions 1501, 1503 and 1506 could be made of N+ doped silicon and the channel 1502, 1504 could be made with P doped SiGe, or many other variations including replacing the NPN with PNP SiGe with Silicon etc. For the following process alternatives, the materials of the channel 1502, 1504 could be selectively or non-selectively isotropic etched in respect to the S/D 1501, 1503, 1506 in order to create selective indentation. Fig. 15B illustrates the structure after a selective isotropic etch of the channel regions 1505. For the case in which the S/D is N+ silicon and the channels are P type SiGe the selectivity could be applied to about 100:1 as is illustrated in Fig. 1. Fig. 15C illustrates the structure after deposition of protective isolation 1510 such as silicon dioxide followed by directional etch using the hard mask 1511 to remove protective isolation on S/D side walls while leaving the protective isolation 1510 only at the region of the etched away SiGe. The structure is now ready for the extra S/D diffusion doping and S/D silicidation. The protective isolation could be later replaced with

O/N/O (Tunneling Oxide/trapping Nitride/isolation Oxide) or be already design to be used as at least part of the O/N/O structure. In this step, the process temperature may be determined with consideration that doesn't substantially impact on the silicide and avoid junction spike.

[000158] Adding the metallic material for the silicidation could be done by known in the art deposition techniques, for example, such as ALD. Examples may be found in a paper by Hyungjun Kim, titled "Atomic layer deposition of transition metals for silicide contact formation: Growth characteristics and silicidation" published at Microelectronic Engineering 106 (2013) 69–75; in a paper by Viljami Pore *et al*. titled: "Nickel Silicide for Source-Drain Contacts from ALD NiO Films" published at the Interconnect Technology Conference and 2015 IEEE Materials for Advanced Metallization Conference (IITC/MAM), 2015 IEEE International; and by Jinho Kim *et al.* titled "Characteristics of Nickel Thin Film and Formation of Nickel Silicide by Remote Plasma Atomic Layer Deposition using Ni(iPr–DAD)2": published at Journal of the Korean Physical Society, March 2015, Volume 66, Issue 5, pp 821–827; and by Kinoshita, A., *et al.* "Solution for high-performance Schottky-source/drain MOSFETs: Schottky barrier height engineering with dopant segregation technique," published at VLSI Technology, 2004. Digest of Technical Papers. 2004 Symposium on. IEEE, 2004, all of which are incorporated herein by reference.

[000159] An embodiment of the invention is to form full metallic source/drains . Such could include tuning the Schottky Barrier height or Fermi level de-pinning by a very thin deposition of isolation, for example, such as silicon oxide or high-k dielectric, prior to the deposition of metal. Such could be achieved by first fully etching away the S/D regions and then filling in, the isolation using deposition techniques such as ALD, followed then by metal deposition and completed by removing the excess metal using etching, for example, anisotropic etching. Such techniques are known in the art as Metal Isolation Silicon ("MIS"). Such has been presented by Connelly, Daniel, *et al*. in a paper titled "A new route to zero-barrier metal source/drain MOSFETs," published in IEEE transactions on nanotechnology 3.1 (2004): 98-104; by Demaurex, Bénédicte in a dissertation titled "Passivating contacts for homojunction solar cells using a-Si: H/c-Si hetero-interfaces." at ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, 2014; by Chiu, Fu-Chien titled "A review on conduction mechanisms in dielectric films," published in Advances in Materials Science and Engineering 2014 (2014); and B. E. Cossa, *et al.* in a paper titled "CMOS band-edge schottky barrier heights using dielectric-dipole mitigated (DDM) metal/Si for source/drain contact resistance reduction" VLSI Technology 2009, pp. 104-105; all of the forgoing are incorporated herein by reference. The full metal source/drain to channel junction can be formed on one or more well-defined crystallographic orientation surfaces of the semiconductor channel as referenced in U.S. patent application publication 2010/0065887, incorporated herein by reference. Such approaches may be used to control the effective Schottky Barrier height. The metal source/drain can include a single type of metal such as tungsten, cobalt, platinum, nickel, or their silicide. Alternatively, the metal source/drain may include a stack of multiple metals in order to form a desired metal work-function and thereby a specific effective Schottky Barrier height. Alternatively, the stack of multiple metals may be used where the first metal contacts the semiconductor channel is thin, but predominately defines the effective Schottky barrier height and the subsequent metal(s) may be chosen for process convenience. For example, such as is disclosed in U.S. patent application publication 2011/0008953, incorporated herein by reference. For the memory application as presented herein the Schottky Barrier could be tuned to about 0.1 - 0.5 eV. Alternatively, the multilayer substrate, such as is illustrated in Fig. 3A of PCT/US16/52726, could be made with undoped or P doped silicon designated to become the memory channel while the S/D layers could be first be made of sacrificial SiGe to be replaced with metal and function as the S/D of the memory structure, effectively exchanging the role of the silicon and SiGe in the structure. A variation of such a sacrificial SiGe flow could include doping the bottom most and the upper most of the silicon strips after the removal of the SiGe strips. Such doping could use techniques such as solid phase or gas based diffusion or monolayer doping (MLD), as previously discussed. Such could be used to form N+ regions in the silicon forming NPN type vertical transistors, and by adding metal into the grooves left by the removal of the SiGe the S/D lines could be completed reducing the bit lines resistivity, and could include forming DSSB vertical memory cells.

[000160] Fig. 15D illustrates an alternative in which the structure of Fig. 15B is formed via a two-step etch; selective SiGe etch followed by non-selective isotropic etch, etching both the S/D and the channel at about equal rate, which create notch near Si and SiGe corners. The results should have round corner unlike what is in Fig. 15D illustration. The channel region 1516 are narrowed further but the S/D regions are also etch forming an S/D neck 1514 regions for the space etch from the side that was open by the prior channel

etch, and a thinner S/D 1512. The height of the neck 1514 in Z direction could be controlled by the etch depth of this second etch step. Fig 15E illustrates the structure of Fig. 15D after similar formation of protective isolation 1520. Since the protective isolation 1520 is extended over the junction near the S/D 1521, the silicidation junction direct short with the channel may be substantially avoided while the silicidation proceeds along the S/D sides. The protective isolation could be later replaced with O/N/O (Tunneling Oxide/trapping Nitride/isolation Oxide) or be already design to be used as at least part of the O/N/O structure. In this step, the process temperature may be determined with consideration that doesn't substantially impact on the silicide and avoid junction spike. The top most S/D 1521 could be used as mask for excess protective isolation directional etch removal. Alternatively, the hard mask 1508 could be trimmed to guarantee the S/D side wall exposure for the following silicidation. This second etch step allow forming a neck to the S/D providing more control for the DS-SB formation.

[000161] An additional inventive embodiment is an additional alternative for the 3D NOR formation process, wherein it use a multilayer in which the S/D layers are kept un-doped prior to the subsequent silicidation process. Such undoped S/D layer and S/D doping last process can prevent the channel autodoping problem during a multilayer epitaxial growth process. For the SB type that would be fine as the S/D are defined by the silicidation process. For other types of memory cells and for DSSB the moderate N concentration such as order of 10^{16} /cm3 or higher concentration of N+ dopant such as higher than 10^{20} /cm^3 could be selectively added to the S/D region. which could be silicided after the ridge formation. In these cases, the channel regions could first be protected by techniques similar to those in reference to Fig. 15C-15E herein. Solid phase or gas base diffusion could be used to dope the vertically arranged multiple layers though the exposed S/D regions. An example of these types of doping techniques are presented in papers by Ajay Kumar Kambham *et al.* titled "Three dimensional doping and diffusion in nano scaled devices as studied by atom probe tomography" published in Nanotechnology 24 (2013) 275705 (7pp); by Thomas E. Seidel titled "Atomic Layer Deposition of Dopants for Recoil Implantation in finFET Sidewalls" published at Ion Implantation Technology (IIT), 2014 20th International Conference on Ion Implantation; and by U.S. patent U.S. 5,891,776, and by D. Raj titled "Plasma Doping of High Aspect Ratio Structures" published at Ion Implantation Technology (IIT), 2014 20th International Conference on Ion Implantation; all of the forgoing incorporated herein by reference. A similar technique is also called Mono Layer Doping (MLD) as presented by Ye, Liang, *et al.* "Doping of semiconductors by molecular monolayers: monolayer formation, dopant diffusion and applications." Materials science in semiconductor processing 57 (2017): 166-172, all are incorporated herein by reference. These techniques could be used with substantially every 3D memory herein. For example, MLD techniques could be used on a dedicated region designated for S/D or on the outer side of a region designated as channel. It could also be used for DSSB formation.

[000162] The epitaxial process forming the multilayer base structure of the 3D NOR formation process could utilize an alternative technique to reduce the probability of dopant transition from the future S/D regions to the channel regions. In a paper by Robert J. Mears *et a.l* titled "Punch-Through Stop Doping Profile Control via Interstitial Trapping by Oxygen-Insertion Silicon Channel" presented at EDTM 2017, incorporated herein by reference, an ultra thin layer of oxide integrated in the epitaxial process is suggested to keep dopant from drifting or diffusing by heat. A mono layer or even less than monolayer of oxide is an effective barrier to dopant pass through. Accordingly, for the multilayer epitaxial process, such oxide blocking could be integrated to support in-situ doping of the S/D regions with blocking oxide at the interface of the S/D layer to the channel layer. Such a method should be carefully considered as it may degrade SB current performance if SB technology is to be implemented.

[000163] As note previously, the role of S/D and Channel may be replaced wherein Si serves as the channel and SiGe as the S/D region. Further and in accordance with 15A-E, SiGe may be completely selectively etched wherein top and bottom planes of silicon layers may be doped by various methods such as Molecular Monolayer Doping, thus forming both channel and S/D region within the Silicon crystalline layer.

[000164] A known concern with memory arrays is various types of disturbs. Some of those are related to parasitic capacitance and similar forms of signal coupling due to the relatively long and close proximity of parallel memory control lines such as bit lines and word lines. These concerns are part of the engineering challenge of any memory device and could be included in the engineering of a 3D NOR memory structure. Some of the alternative techniques in the following could be adapted for such memory engineering.

[000165] A body contact as discussed in respect to Fig. 42A-42E of PCT/US16/52726 could be engineered either at the edge of a memory unit or multiple times along the ridge. The connection lines (4248 of PCT/US16/52726) could be called body-lines and could be connected to ground, or to a 'body control' which could be connected to a specific voltage control as part of the memory control circuit and logic. Such 'body control' could be set to a positive voltage to assist the memory erase step. In some embodiments, a positive body voltage erase may be accomplished with the ground voltage to the selected word lines, eliminating the need of a negative voltage for all operations, thus as a result the periphery area (memory control circuits) used for a negative voltage generator can be saved. In some applications it could be combined with higher doping in the center of the channel, for example, such as indicated in Fig. 1B as $Si_{0.8}SiGe_{0.2}$, to improve the body horizontal conductivity.

[000166] An additional embodiment is to have within a selected ridge all of the unselected S/D lines left floating (e.g. Fig. 13B), which may be connected together, for example, through a multiplexer in the periphery circuits. By connecting together all the S/D lines left floating, they will form a much larger capacitive load thus significantly reducing the capacitive coupling to the two active S/D lines.

[000167] An additional embodiment, such as when using in FN or direct tunneling, is to sequence the writing sequence to reduce cross talk by first activating the selected Source and Drain (the two adjacent S/D lines), then assert the selected word-line to turn on the vertical transistor of the selected bit cell. Then after the cross-talk ripple to adjacent S/D lines has subsided, move the wordline bias high enough so that the writing process will be effective only in the selected memory cell, thus reducing the disturb effect.

[000168] An additional embodiment is a 3D NOR structure alternative: to replace the full ridge shared body contact with a per layer body contact using a stair-case to allow selective access control for each body layer. Fig. 15F illustrates such structure in which the body between S/D1 and S/D2 could be controlled by signal B1, the body between S/D2 and S/D3 could be controlled by signal B2, and so forth. In such a 3D NOR structure, an alternative writing technique could be achieved by using one word-line such as WL1 and one body layer of a selected ridge, such as B2, to select a specific cell 1540. Using FN tunneling or direct tunneling by having the voltage difference between B2 and WL1, for example using a positive voltage for the selected WL and zero or a negative voltage for the selected channel 'body' line will pull charge into the related charge trapping region. All S/D lines may be left floating. For the erasing operation using a positive body contact voltage, all S/D lines could be left floating while the selected WL is grounded and the remainder of the WL's floating. Alternatively, the aforementioned configuration of sharing voltages between selected WL and Channel lines may benefit by requiring a smaller voltage range to perform program and erase operations. Such maybe accomplished by negative half of erasing voltage to the gate and positive half of erasing voltage to the body to create a full erase voltage across the gate and the body while all S/D lines could be left floating. Furthermore, inhibit of non-selected cells is obtained naturally. Alternatively, to avoid the need of negative voltages and associated array size penalty, only positive voltages may be utilized. Programming operation using only positive word line voltages is accomplished by grounding the channel line and all other unselected channel lines are left floating. The unselected S/D lines could be biased by half voltage of the programming word line voltage so that the voltage difference between the selected word line and unselected cells is low enough to avoid any undesired FN tunneling. Alternatively, for programming operations using a positive word line voltage, the unselected S/D lines could be left floating or alternatively biased by half voltage of the programming word line voltage so that the voltage difference between the selected word line and unselected S/D line is low enough to avoid any undesired FN tunneling from the S/D lines. The selected body voltage to the selected channel B2 should be selected per ridge to avoid writing in the non-selected ridge. An additional embodiment in respect to such form of writing is to steer the charges to be close to one of the S/D lines such that at least two charge locations could be formed to increase the memory density. Such could be done for example by having S/D3 at ground while all other S/D lines are floating, preferably shorted together. The electric field between the negative body B2 and the grounded S/D3 during program operation wherein also a positive voltage is applied on WL's could pull the electrons toward the S/D3 side for writing on that side of the charge trap region and for the other side replace roles with S/D2. Accordingly, the storage locations per facet could be made effective with such a writing technique. Additional variations could include a blend by combining FN type writing with some level of hot electron by proper voltage control of S/D2 and S/D3. An additional alternative for charge steering could include modulating the steering S/D line by a wavelet function similar to what has been presented in at least U.S. application 15/333,138 with respect to at least Fig. 27 to Fig. 32.

[000169] Fig. 15G is an Y-Z cut view illustrating an optional alternative to a staircase for forming a per layer 'body' contact presented in respect to Fig. 15F herein with reference to B1-B7. It illustrates a pillar electrode 1558 to form a programmable connection per layer through programmable isolation 1556 such as a one-time programmable antifuse, for example, formed with silicon oxide or other resistive switching material that could be electrically programmed to form a conductive connection from the pillar electrode 1558 to the designated body such as body 1564. The side view illustrates S/D lines that could be referred as S/Dn+2 1552, S/Dn+1 1560, S/Dn 1568 and S/Dn-1 1570. The designated channels to form a programmable connection to, are illustrated as 1562, 1564, 1566 for which 1564 could be considered as the 'body' for channel 'n'. During the formation of the structure, the S/D protective isolations 1550, 1554 could be processed in a similar way to the process used to form channel protection 1510. The structure illustrated in Fig. 15G could be designed to form enough vertical pillars 1558 to allow at least one pillar per body. Then the per layer contacts to the S/D lines could be used to program the connection between the vertical pillar electrode 1558 to the designated body. Such programming could be formed by pulsing +Vpp to the vertical pillar 1558 and –Vpp/2 to S/Dn and S/Dn+1 resulting in forming programmable link 1580 between the vertical pillar electrode 1558 and the Channel n body 1564. The adjacent S/Dn-1 and S/Dn+2 could be pulled-up as additional protection. The vertical connection illustrated in Fig. 15G could also be made oriented in Y direction parallel to the connection structure for the S/D lines as is illustrated in Fig. 12B.

[000170] An additional alternative when the memory structure includes body access is to use the programming method known as Channel Initiated Secondary ELectrons injection ("CHISEL") which could allow lower writing voltages. For example, the Source line could be held at Vs=0v the Drain at Vd= 2 to 3v and the 'body; at Vb= -2 to -3v. Such programming techniques has been detailed in a paper by Mahapatra, Souvik, S. Shukuri, and Jeff Bude. "CHISEL flash EEPROM. I. Performance and scaling," *IEEE Transactions on Electron Devices* 49.7 (2002): 1296-1301; by Mahapatra, Souvik, S. Shukuri, and Jeff Bude. "CHISEL flash EEPROM. I. Performance and scaling," *IEEE Transactions on Electron Devices* 49.7 (2002): 1296-1301; by Driussi, Francesco, David Esseni, and Luca Selmi. "Performance, degradation monitors, and reliability of the CHISEL injection regime." *IEEE Transactions on Device and Materials Reliability* 4.3 (2004): 327-334; by Nair, Deleep R., et al. "Explanation of P/E cycling impact on drain disturb in flash EEPROMs under CHE and CHISEL programming operation." *IEEE Transactions on Electron Devices* 52.4 (2005): 534-540; and by Stefanutti, Walter, et al. "Monte Carlo simulation of substrate enhanced electron injection in split-gate memory cells." *IEEE Transactions on Electron Devices* 53.1 (2006): 89-96, all of the forgoing are incorporated herein by reference.

[000171] The memory structure herein was presented as a charge trap memory. When the target application could make tradeoffs between write speed and retention time, the charge trap layer may be tuned accordingly. As illustrated in IEEE ELECTRON DEVICE LETTERS, 16, 11, p. 491, 1995 by H. Clement Wann and Chenming Hu, "High-Endurance Ultra-Thin Tunnel Oxide in MONOS Device Structure for Dynamic Memory Application", while thinning down the bottom-oxide thickness improves program speed, retention time decreases significantly. An alternative approach may be considered wherein the bottom-oxide may be replaced with low-trapping nitride such as oxinitride as published in Masayuki Terai, Koji Watanabe, and Shinji Fujieda , "Effect of Nitrogen Profile and Fluorine Incorporation on Negative-Bias Temperature Instability of Ultrathin Plasma-Nitrided SiON MOSFETs", IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 54, NO. 7, JULY 2007 or JVD nitride as published in IEEE ELECTRON DEVICE LETTERS, VOL. 21, NO. 11, pp. 540 2000, by Yee Chia Yeo, Qiang Lu, Wen Chin Lee, Tsu-Jae King, Chenming Hu, Xiewen Wang, Xin Guo, and T. P. Ma, "Direct Tunneling Gate Leakage Current in Transistors with Ultrathin Silicon Nitride Gate Dielectric". For a similar retention time, such a layer may provide better control of ONO formation as the required thickness is significantly larger, directly related to the ratio of dielectric constants, 7/3.9. Such an advantage may be realized with very a thin nitride layer of about 1 nm which may be production worthy compared ~0.6 nm silicon-oxide which may not feasible to realize in a production environment. An additional advantage is a faster FN erase speed thanks to the significantly smaller offset in valence band with respect to the silicon. Such a method may therefore enable an erase operation and limit erase voltages to lower or similar values as the program voltage. Program speed may also be improved significantly thanks to the smaller band offset if a hot electron injection mechanism is utilized. Such a method may be Channel Hot Electron Injection (CHEI) or Channel Hot Injection Secondary Electrons (CHISEL) in doped np junctions or alternatively in Schottky Barrier or DSSB transistors hot electrons are generated next to the source wherein injection also takes place. Further improvement in program and erase speed or voltages may be brought about by thinning

down to about 2 nm the trapping nitride layer which may be formed either by LPCVD or ALD methods with a potentially significant program window of 1-2V as referred in IEEE ELECTRON DEVICE LETTERS, VOL. 32, NO. 11, p.1501, 2011 by G. Van den bosch, G. S. Kar, P. Blomme, A. Arreghini, A. Cacciato, L. Breuil, A. De Keersgieter, V. Paraschiv, C. Vrancken, B. Douhard, O. Richard, S. Van Aerde, I. Debusschere, and J. Van Houdt, "Highly Scaled Vertical Cylindrical SONOS Cell With Bilayer Polysilicon Channel for 3-D NAND Flash Memory". Other forms of such multilayers could be engineered to fit the specific design objective of retention vs. write time.

[000172] An additional option is to use the low trapping nitride, such as oxynitride and Jet Vapor Nitride mentioned above, as a replacement to the high doping charge trapping nitride layer, thus realizing a layer which accommodates injected charges in the conduction band of the low trapping nitride. Such a method would be an alternative approach to the common industry standard of floating gate polysilicon.

[000173] The memory structure herein was presented as a charge trap memory. Recently good progress has been made in respect to ferroelectric type memories, especially in respect to utilizing Hafnium Oxide and Zirconium Oxide based materials (HfO, ZrO, HfZrO, ZrSio, …). These memories are referenced as FRAM and could provide higher write and erase speeds compared to charge trap based memories. At the current time a development effort is being pursued by the industry to develop such FRAM technology to be commercially viable. The technologies in development for such FRAM include advanced ferroelectric deposition techniques such as ALD. Adapting FRAM to the 3D-NOR fabrics presented herein could be a good option. These could include also a mix, in which some of the regions are processed with O/N/O and some of the other regions with FRAM materials such as HfO_2 and silicon doped HfO2 (Si:HfO2) or Zr doped HFO_2 based materials. For example, the locations where the O/N/O layer used to be formed can be replaced with a stack of dielectric to function as a ferroelectric memory transistor. Alternatively, such locations may instead be formed with a stack of a charge trapping layer and a ferroelectric layer. Such a mix could be attractive for many systems as it provides multiple memory technologies, low power, and time for transferring data between these memory technologies as was discussed herein. Using FRAM within the 3D-NOR memory could include many of the variations referred herein in respect to charge trap memories. These could include multi-level cells in which multi-bits are coded in via different writing voltages or different write times. These could also include multi-bit locations similar to mirror bit and multi-gates steering of bit location such as discussed in respect to at least Fig. 10E of PCT/US16/52726 and Fig. 15A to Fig. 23 of US 15/333,138. Similar to charge trap, FRAM is an electric field based memory and is an isolative material, and accordingly could support these multi-bit coding techniques to support higher memory density. As such, the stored data states can be differentiated by either or both the degree and the location of polarization of the ferroelectric. The use and methods of constructing an FRAM memory is covered in many recent papers and their teaching could be applied to incorporate such ferroelectric dielectric for the 3D NOR. Such papers as by J. Muller *et al.* titled "Ferroelectric Hafnium Oxide Based Materials and Devices: Assessment of Current Status and Future Prospects" published in ECS Journal of Solid State Science and Technology, 4(5) N30-N35 (2015); by Patrick D. Lomenzo *et al.* titled "Annealing behavior of ferroelectric Si-doped HfO2 thin films" published at Thin Solid Films 615 (2016) 139–144; by Uwe Schroeder *et al.* titled "Chapter 3, Nonvolatile Field-Effect Transistors Using Ferroelectric Doped HfO2 Films" published by Springer Science & Business Media Dordrecht 2016, B.-E. Park et al. (eds.), Ferroelectric-Gate Field Effect Transistor Memories, Topics in Applied Physics; by U. Schroeder *et al.* titled "Impact of field cycling on HfO2 based non-volatile memory devices" presented at ESSDERC16; by Shinji Migita titled " Thickness-Independent Behavior of Coercive Field in HfO2-based Ferroelectrics" presented at EDTM 2017; by Cheng, Chun-Hu, *et al.* titled "Low-leakage-current DRAM-like memory using a one-transistor ferroelectric MOSFET with a Hf-based gate dielectric" published in IEEE Electron Device Letters 35.1 (2014): 138-140; and by Zhen Fan titled "Ferroelectric HfO2-based materials for next-generation ferroelectric memories" published at JOURNAL OF ADVANCED DIELECTRICS Vol. 6, No. 2 (2016); all of the forgoing are incorporated herein by reference.

[000174] Some of the ferroelectric materials could act also as a charge trapping materials. These two could impair each other as discussed in a paper by Yurchuk, Ekaterina, *et al.* titled "Charge-Trapping Phenomena in HfO 2-Based FeFET-Type Nonvolatile Memories," published in IEEE Transactions on Electron Devices 63.9 (2016): 3501-3507; the forgoing is incorporated herein by reference. On the other hand the memory cell could be engineered to have these two enhancing each other, so the ferroelectric and

charge trap could be combined for enhanced memory functionality, such as presented byYu-Chien Chiu *et al.* titled "Low Power 1T DRAM/NVM Versatile Memory Featuring Steep Sub-60-mV/decade Operation, Fast 20-ns Speed, and Robust 85oC-Extrapolated 1016 Endurance" presented at 2015 Symposium on VLSI Technology; Chiu, Yu-Chien, *et al.* "On the variability of threshold voltage window in gate-injection versatile memories with Sub-60mV/dec subthreshold swing and 10^{12}-cycling endurance" Reliability Physics Symposium (IRPS), 2016 IEEE International, IEEE, 2016; and by Chia-Chi Fan *et al.* titled "Impact of Ferroelectric Domain Switching in Nonvolatile Charge-Trapping Memory" presented at EDTM 2017; and in U.S. patent application publication 2016/0308070; all of the forgoing are incorporated herein by reference. Such enhancement is accomplished when the electric field generated by the charge trapping and the polarization of ferroelectric are oriented to enhance each other.

[000175] Writing a ferroelectric memory cell techniques are provided, for example, in U.S. patent 6,667,244, incorporated herein by reference. The following ferroelectric writing example for the 3D-NOR structure could be also adapted to charge trap programming using FN by adapting voltage levels and to the above structure for which ferroelectric cells are designed to be enhanced by charge trapping.

[000176] An example of writing conditions in the following assumes that the 4 V difference across the gate and the S/D is engineered to be sufficient to polarize the ferroelectric while the voltage difference below half of it (2V) is not sufficient to disturb the states of the ferroelectric.

[000177] Fig. 16A illustrates a schematic of a single cell written to the '0' logic state. Fig. 16B illustrates such single cell structure in a simplified cross section.

[000178] Fig. 16C illustrates a schematic of a single cell written to the '1' logic state. Fig. 16D illustrates such single cell structure in a simplified cross section.

[000179] The specific voltages of these structures are for illustration only and are to be defined specifically for a specific memory cell as part of such memory cell engineering. In a case where the selected word-line (the gate) is shared with other cells in the array the unselected cell could have their bit-lines (Source-lines and Drain-lines – S/D lines) at ground (zero volts – 0V) or in some alternatives floating. Likewise, in a case where the selected bit-line (the source and the drain – S/D lines) is shared with other cells in the array, the unselected cell could have their word-lines (the gate) are at ground or in some alternative floating.

[000180] Fig. 17A-17D illustrates an alternative in which two bits are stored in one facet of the memory cell by directing the electric field only to the source side of the memory transistor, or alternatively (not illustrated) to the drain side. This could resemble the mirror –bit concept in charge trapping cells. This writing method could be adapted for the 3D NOR fabric. The channel could be floating in the cases outlined above.

[000181] Fig. 18A illustrates a small section of a ridge of the 3D-NOR fabric. These writing schemes are driven by an electric field between the channel, driven from the S/D lines, and the word-lines. Note that the symbol 'x' in the figures represents that no writing is to be effected at the related ferroelectric zone. For such writing techniques, the odd level cells are to be used while the even level memory cells could be ignored. This is since in this writing technique both sides of the active S/D are affected in the regions close to the active word-line. Fig. 18A illustrates writing 'zero' to a memory cell. Fig. 18B illustrates writing 'one' to a half of a memory cell. Fig. 18C illustrates writing 'one' to a group of memory cells sharing the same bit-lines (S/D lines). Fig. 18D illustrates writing 'zero' to a group of memory cells having common word-line (gate line). These write modes illustrations are indicative of the options available for writing of a ferroelectric memory option within 3D NOR structure. These could be combined with multilevel programming techniques to increase storage capacity. Such multilevel storage is presented in a paper by Mulaosmanovic, Halid, *et al.* titled "Switching kinetics in nanoscale hafnium oxide based ferroelectric field effect transistors," published at ACS Applied Materials & Interfaces (2017), incorporated herein by reference.

[000182] These could also include multi-bit locations similar to mirror bit and multi-gates steering of bit location such as is discussed in respect to at least Fig. 10E of PCT/US16/52726 and Fig. 15A to Fig. 23 of US 15/333,138 for further increases of memory capacity. The multi storage locations writing, is similar to that of charge trap: the side gates could be used to modify the electric field directing the effect of the major gate to the selected location of the related facet. Also the techniques presented for transferring memory, from and to high speed cells, and high density cells, could be used in respect to ferroelectric based memory cells.

[000183] In hot-electron writing techniques the side gate steering is targeting the channel region in which the hot electrons are being formed along the channel width direction perpendicular to the source-drain direction. In FN and even more so for FN in which the gate is the source of the charge, or in ferroelectric writing techniques, the steering could also be directed to the electric field formed in the O/N/O and/or the ferroelectric regions. Accordingly it would be enough if the side gates are positioned to affect the electric field between the primary gate and the semiconductor region.

[000184] Fig. 18E and Fig. 18F are example illustrations how this effect might be arranged and performed. Fig. 18E is an X-Z cut-view of a small ridge region of a 3D NOR. It shows a section of two 1st gates 1821, 1823 and two 2nd gates 1822, 1824. It also shows two S/D regions 1801, 1803 and the channel 1802 in between them. Fig. 18F is an X-Y cut-view at the active region of the ridge 'marked plane' 1805 of Fig. 18E. It shows the side view of the channel 1802 marked channel 1832, the respective gates 1821, 1822, 1823 and the O/N/O or ferroelectric or combination 1834. It also illustrates two memory sites that could be formed by these steering techniques. By assigning the role of main gate to 1822 and pulling the electric field by side gate 1823 and pushing the electric field by side gate 1821 the write location could be made to 1828. And, for example, by pulling the electric field by side gate 1821 and pushing the electric field by side gate 1823, the write location could be made to be 1826. To avoid disturb to the memory related to the side gate, the side gate voltage could be set to be below the level that could disturb the side memory. The specific voltages could be set for a specific memory structure and all numbers presented in here are for illustration only.

[000185] The writing techniques presented with respect to Fig. 15F, could also be used for ferroelectric based memory and the presented charge-trap/ferroelectric structures presented herein. In such case at least one of the channel facets storage layer (O/N/O) could incorporate ferroelectric material such as HfO_2 as presented herein before. The writing techniques presented in respect to Fig. 18A-18D could be adapted accordingly, so that the writing and erasing electric field is formed between the selected word-line and the selected 'body' while the S/D lines are left floating, or used for electric field steering as has been described herein.

[000186] An embodiment of the invention is to use such memory structures or a portion of such for Synapse-like functions. Such use of memory has been suggested in the art for RRAM and PCM cross bar but could be applied to charge trap or other memory types presented herein. Published work for RRAM and PCM cross bar has been by Chen, C-YM *et al.* titled "A solid-state electronic linear adaptive neuron with electrically alterable synapses" published at Neural Networks, 1991, 1991 IEEE International Joint Conference on. IEEE, 1991; by Lee, Myoung-Sun, *et al.* titled "Implementation of Short-Term Plasticity and Long-Term Potentiation in a Synapse Using Si-Based Type of Charge-Trap Memory" published at IEEE Transactions on Electron Devices 62.2 (2015): 569-573; and by Kornijcuk, Vladimir, *et al.* titled "Leaky integrate-and-fire neuron circuit based on floating-gate integrator" published at Frontiers in neuroscience 10 (2016); all of the forgoing are incorporated herein by reference.

[000187] In many of the memory structure herein the writing technique could include a reduced gate voltage in conjunction with use of negative voltages on the S/D to reduce the overall power requirements of the device. Such technique could also take advantage of the heterogeneous integration of memory control circuit layer(s) disposed over a memory matrix.

[000188] An additional inventive embodiment is an alternative for a 3D device-system; such a system as is illustrated in Fig. 11E. It is to construct the system as array(s) of memory units, such as illustrated in Fig. 12B, and with a corresponding processor cores on top of it or underneath. For many compute tasks the program code loaded on a processor core could be set according to the content loaded in its corresponding memory unit on top or under it. Thus processing of data could be accomplished with a short distance data transfer of less than about 1 mm or less than about 100 microns or even less than about 20 microns, as opposed to PCB (Printed Circuit Board) based computers in which the data from the memory being fed though over 20 mm wires using the PCB. The memory unit and the processor core could be rectangular in shape with an area of less than about 4 mm^2, or less than about 1 mm^2. Alternatively, the data transfer into the 3D device-system could be set such that data is placed in correspondence with the designated processor to process it. An additional alternative is the use programmable logic as part of the processing logic. With the use of programmable logic, or FPGA, the logic configuration could be adapted to the data stored in the corresponding 3D memory to further enhance processing efficiency while reducing data movement power and delay. Fig.19A illustrates an X-Z cut view of a 3D system in which a first strata of memory units 1902, 1904, 1906, 1908, is overlaid by a second strata of memory units 1912, 1914, 1916, 1918, is overlaid by a third strata of memory control circuits 1922,1924,1926,1928, and is overlaid by fourth strata of processor core units 1932, 1934, 1936, 1938. In

some embodiments, the first strata of memory units may be high density non-volatile memory such as Flash memory or RRAM. In some embodiments, the second strata of memory unit may be high-speed memory such as DRAM or static memory such as SRAM. For example, such computer system could be tasked to search for a person which meets a specific criterion. Accordingly the fourth strata cores could be set with the search in parallel in which core 1932 is access using memory control 1922 for that person in the data base being hold at memory cores 1912 and 1902. In another case only one core at a time will perform the search while the other cores perform other tasks, yet the search is perform by a core 1932 in the data bank in 1912, 1902, and then the search task is assigned to core 1934 which will run it for the data in 1914, 1904 and so forth.

[000189] An additional advantage in such a 3D memory system relates to the potential defects in semiconductor manufacturing. For example, the structure illustrated in Fig. 19A could be processed at full reticle level with the expectation that some of the cores or the memory units could be defective and would not be activated on the end product, which could be designed to function with only 80% of the units functional. As such, the third strata of memory control circuits may have an on-chip testing function (not shown) in order to assess the functionality of the memory layers underneath and allocate those memory blocks into the enabled and the disabled blocks, and update the data routing path(s) accordingly. An important advantage of this 3D structure is the ability to effectively support a very fine grain of unit based construction. Such units could be designed and engineered to be sized to less than about 1 mm^2 or less than about 0.2 mm^2 or even less than about 0.05 mm^2. While the 3D system size could be larger than 100 mm^2 or larger than 600 mm^2, or larger than 2,000 mm^2 or even larger than 10,000 mm^2.

[000190] Moreover, a mix of redundancy techniques could be used. As such the multi-core multi-unit 3D system of Fig. 19 could have a system control function 1940 which controls the overall 3D system and could be constructed with two strata: one strata provides redundancy to the other at the logic cone level as been presented in U.S. 8,994,404, incorporated herein by reference, in respect to at least Figs. 24A-44B.

[000191] The system control function 1940 could include input output channels to other systems, or to a communication channel such as the internet or to wireless systems such as G4, G5. This could include such as fiber optic channel, free space optical channel, wireless channel and other forms of communication channels. The Monolithic 3D technology presented herein enables heterogenous integration to enable those forms of communication.

[000192] The 3D architecture also could be useful to enable common manufacturing of a modular system that could be customized to specific needs by techniques presented herein, such as the use of each of a continuous structure as presented U.S. 8,994,404 as related to at least Figs. 11A-12E, Figs. 14-17, and Figs. 22A-23D. Fig. 19B illustrates such customization. The upper portion 1952 is a magnification of section of a structure such as is illustrated in Fig. 19A, and it overlay a substrate 1954. The generic wafer could be then customized by dicing it to the desired end chip size. The dicing 1956 could be done by many of the known techniques including conventional dicing saw, or plasma etching also called plasma dicing or laser assisted dicing. The dicing could be done at designated potential dice lines 1958. These potential dice lines could include various restrictions (for example, design rule restrictions) and support for potential future dicing including guard rings and avoiding active regions or metal lines through them. Alternatively, the dicing could be done by advanced dicing techniques such as laser assisted or plasma assisted dicing. And could be supported by additional techniques to seal and provide side wall protection to support good functionality and reliability of the end device.

[000193] An additional inventive embodiment is an additional aspect of a 3D computer system, such as is related to Fig. 19A, where there is the need to integrate multiple memory stratums to achieve a larger memory bank. In one alternative, multiple stratums could be integrated via 3D integration with minimum per strata processing, and then a memory control stratum could be added to control each and every memory cell in the strata underneath. These memories could be structured for such integrations. These memories could be volatile memory such as DRAM, Non Volatile memory such as 3D NOR, or 3D NAND or even a mix of such. In the following description of such integration, it is assumed that the memory are constructed in a same size memory unit array and each such unit is controlled by the same pitch of memory control lines so when one memory wafer is bonded on top of another memory wafer these memory units and their control lines (wordlines and bitlines) are precisely overlaid to each other, This overlaying is subject to the wafer to wafer or die to wafer misalignment precision of the bonding equipment. The integration technique leverages copper to copper, hybrid or 'fusion' bonding in which the bonding process also functions as an electrical connection process between these

wafer/strata. Such precise bonding is presented in a paper by Kurz, Florian, et al. "High Precision Low Temperature Direct Wafer Bonding Technology for Wafer-Level 3D ICs Manufacturing." *ECS Transactions* 75.9 (2016): 345-353, incorporated herein by reference. Utilizing precise bonder and thin layer transfer to construct 3D generic memory structure and integrating it with a logic wafer to form high performance 3D compute system is presented in a paper by Zvi Or-Bach titled "A 1,000x Improvement in Computer Systems by Bridging the Processor Memory Gap" IEEE-S3S 2017, incorporated herein by reference. Fig. 20A- 20F illustrates preparation of a wafer for such connectivity using wafer bonding, including construction of the TLV to allow multiple stratum integration, leveraging the 3D layer transfer techniques presented herein or other thin layer transfer techniques.

[000194] Fig. 20A illustrates a Y-Z 2000 cut view of section of a wafer, a base wafer 2002 including a SiGe "cut layer" 2001 and memory circuits 2003.

[000195] Fig. 20B illustrates the structure of Fig. 20A after etching hole region 2004 substantially all the way to expose the cut layer 2001.

[000196] Fig. 20C illustrates the structure of Fig. 20B after forming top metal landing pad 2006. The isolation layer to protect the silicon sidewall of the memory circuit 2003 (not drawn). This top landing pad 2006 may be drawn larger in the X and the Y direction than the pitch of a TLV as determined by the bonding process, in order to accommodate the wafer bonding alignment tolerance of the TLV. Top landing pad 2006 may be in-plane with the silicon layer of the memory circuit 2003.

[000197] Fig. 20D illustrates the structure of Fig. 20C after covering the structure with isolation layer 2008.

[000198] Fig. 20E illustrates the structure of Fig. 20D after adding interconnection layer 2010 including at least one via 2014 to the landing pad 2006. This could be a memory control line. The interconnection layer 2010 may be wordline, bitline and/or sourceline of the memory circuit 2003.

[000199] Fig. 20F illustrates the structure after adding the bottom connection pad 2012 connected to the interconnection layer 2010.

[000200] Such prepared stratum may be bonded onto another target wafer and once the cut is performed the target strata is ready to have additional stratum bonded and connected onto it.

[000201] Fig. 20G illustrates an alternative to adapted such technique to die level operation as was presented in reference to Fig. 4H herein. As an example, a multilayer such as is illustrated in Fig. 4H could be used. For such the bottom layer 2042 could be silicon, and SiGe layer 2044 on top of the silicon layer 2042, and top silicon device layer 2046. Then using a similar flow to the one in reference to Fig. 20A-20F herein, bottom pads 2022 and top pads 2044 could be formed. Thus the multilayered structure 2040 could be 'cut' and diced out and bonded at a die level onto another target wafer (not drawn), then a selective etch from the top could be used to first remove the silicon layer 2042 and then thin the die all the way to the device layer, for example, by selectively etching the SiGe layer 2044. These support layers 2042 and 2044 could have a thickness of about 1 micron, 1-3 microns, 3-6 microns or even higher. The device layer 2046 itself could include sub-layers such as n+ and p- to support the back-bias scheme as discussed before herein and in incorporated references. By having the inter-stratum connectivity structure 2022, 2024 pre-built, the stacking process could provide both mechanical bonding and through silicon connectivity – hybrid bonding – thus simplifying the 3D system formation which could includes wafer level stacking and die level stacking yet having thin stratum in the stack. These stratum could be at a thickness of about 10, 20, 40, 100, 200, 400 nm or about one or a few microns.

[000202] Fig. 21A- 21C illustrates a small region of a memory control line X-Y 2100 top view. In this integration technique, a layer selected (not shown) could be used to allow multiple stratum control lines to be connected in parallel yet by enabling the layer select only the selected stratum could be accessed.

[000203] Fig. 21A illustrates top metal landing pads 2102 for control lines 2104 such as bit line, word line, or source line that are sized to the maximum bonding misalignment margin 2101 to guarantee that the bottom connection pad 2108 of the following stratum to be bonded will land on the top landing pad of the prior stratum. In some of the advanced wafer bonders the bonding misalignment is less than 100 nm (three sigma).

[000204] Fig. 21B illustrates the structure with the added control lines 2104 and their connection 2106 to the top landing pads. In many cases, the control lines pitch is denser than the expected worst-case misalignment and accordingly the landing pads are placed on multiple rows as illustrated. These control lines could be the bit-lines or the word-lines. The connectivity structure of Fig. 21A-21C

assumes a control lines pitch of about 80 nm. Using an advanced lithography process, control lines pitch could be further pushed to even below 30 nm. The connectivity approach illustrated in Fig. 21A-21C could be adjusted accordingly.

[000205] Fig. 21C illustrates the structure after adding the bottom connection pads 2108. The use of the term bottom and top connecting pads is just for the ease of explanation and being part of layer transfer process top and bottom could be flipped.

[000206] Fig. 21D illustrates the X-Z cut-view or the Y-Z cut-view 2120 of the top landing pad during the wafer processing. Having a carrier substrate 2110, the area designated for intra stratum connectivity could be first processed to etch the silicon 2118 all the way to expose the SiGe layer 2111 using selective etch with the SiGe as an etch stop. Silicon 2118 may include a bottom layer silicon layer 2122 (dark black line) which later after flip bonding and 'cut' could become the top layer. The etch process could be wet or dry as the region being etched is relatively large, for example, of about 1 micron by 200 microns, or about 0.5 microns by 300 microns. Then the region could be filled with an electrically isolative material, for example, such as oxide 2113. Then the top landing pads may be patterned and then filled in with a conductive material, for example, such as copper, forming top landing pads 2112. The size of these pads 2112 could be large enough to assure the desired electrical and physical contact after the following wafer bonding step. For example, such as about 100 nm by 100 nm, or about 200 nm by 200 nm, or about 220 nm by 220 nm, or about 180 nm by 180 nm, or about 250 nm by 250 nm, or about 180 nm by 220 nm depending on production, design and other engineering considerations, especially the wafer bonder alignment capability. Then vias 2116 to these landing pads 2112 may be formed and additional isolation 2113 could be added if needed. These vertical connectivity elements pads 2112 and vias 2116 could be called Through Layer Via (TLV) or nano-TSV. The processing of these nano-TSVs could take place after the wafer has completed the high temperature process, often called Front End of the line ('FEOL"), which includes forming all the transistors, their isolation, and contacts in the active silicon 2118. The metal pattern could include a 'bonder alignment marks' 2119 to support the following face-to-face precise alignment. These 'bonder alignment marks' could be placed at the die level or even at the reticle level as they are part of the full wafer alignment process.

[000207] Fig. 21E illustrates the structure after completing the interconnect layers – the Back-End of The Line ("BEOL"). The memory array interconnect 2130, the nano-TSV which includes the future bottom pads 2138, the landing pads 2132, and the vias 2136 connecting them, and the bit-lines or the word-lines 2134. An oxide 2131 could be added to cover the array interconnect 2130. The process could be designed so the top surfaces of the future bottom pads 2138 are exposed to support the future metal to metal or hybrid bonding. Such preparation could include a slight height adjustment to ensure connectivity between the stack stratum.

[000208] Fig. 21F illustrates the resulting structure after having the structure of Fig. 21E, first structure 2144, flipped and bonded (metal to metal or hybrid bonding) to a second structure, base structure 2142 (may be similar to first structure 2144). The base structure 2142 may not need full nano-TSV or SiGe layer, for example, if it is supposed to be the uppermost stratum of the 3D chip; however, having a unified memory wafer could be preferred and having the ability to connect controls from both sides could be desired.

[000209] Fig. 21G illustrates the structure after removal of the base silicon 2146 of the top wafer first structure 2144. This could be done with a conventional grind followed by etch leveraging the SiGe layer 2148 as an etch stop.

[000210] Fig. 21H illustrates the structure after removal of the SiGe cut-layer 2148, using selective etch to etch mostly SiGe and not silicon. The exposed silicon layer 2122 could be oxidized to support subsequent hybrid bonding of additional structures, for example, such as first structure 2144 on top to form a three stratum stack or as many as needed. Alternatively, the hybrid bonding could be made of silicon to oxide and metal to metal. A few processes could be used to convert the now top layer silicon 2122 to oxide. Such as simple etch and deposition with potential CMP step to expose the pads, or to oxidize the top silicon surface 2122 using low temperature techniques such as presented in a paper by H. Kakiuchi *et al* titled "Formation of silicon dioxide layers at low temperatures (150–400 ^{o}C) by atmospheric pressure plasma oxidation of silicon" published at Science and Technology of Advanced Materials 8 (2007) 137–141; and by Masaki Hirayama *et al* titled "Low-Temperature Growth of High-Integrity Silicon Oxide Films by Oxygen Radical Generated in High-Density Krypton Plasma" published at IEDM 99. Alternatively the boding could be made between the top silicon 2122 and oxide 2131 of the added stratum such as presented in a paper by R. Do Black *et al* titled "Silicon and

silicon dioxide thermal bonding for silicon-on-insulator applications" published at J. Appl. Phys. 63 (8l, 15 April 1988, all of which are incorporated herein by reference.

[000211] The process for removing the base silicon 2146 and the SiGe cut-layer 2148 could include use of grinding and selective etch as previously discussed. First selectively etch silicon using the SiGe layer 2148 for an etch stop, and then etching selectively the SiGe using the silicon 2122 and the pads 2132 as an etch stop. Alternatively the SiGe layer 2148 could be pre etched or mostly etched similar to the process in reference to Fig. 2H to Fig. 3D and Fig. 4E to Fig. 4H herein. The 3D integration between memory control circuits and the bitline/wordline of the memory array could utilize the concept illustrated in Fig. 21A-21C, as an alternative to the 'smart-alignment' technique such as in reference to Fig. 11F-11H, or the programmable technique of Fig. 11I-11K.

[000212] Fig. 22A-22B are X-Z 2200 cut view of the stratum select connectivity. It supports a generic stratum design which could be integrated in any stack numbers of 3D integration and allow a top select of each stratum.

[000213] Fig. 22A illustrates one stratum section which is designed to support up to four stratum integrations (for example) with top access to select each of the stratum in the stack by the top access – SL1, SL2, SL3, SL4.

[000214] Fig. 22B illustrates a stack of four stratum 2211, 2212, 2213, 2214 which are stacked so that SL1 could be used to select stratum 2211 and so forth to SL4 to select the top stratum 2214.

[000215] Fig. 22C illustrates a conventional DRAM block diagram. In the 3D computer system presented in herein the memory array could be in one strata while the control circuits, for example, such as Row Decoder, Sense Amps, Column Decoder, and Data In/Out Buffers, are placed on upper most overlaying (or lower most underlying) strata. Such memory multiple array stratums could be combined by techniques such as has been described herein to form a larger memory 3D array. For example, the per layer select to be added per unit array as is illustrated in Fig. 22D for the bitlines and as is illustrated in Fig. 22E for the wordlines. At the edge of the unit array the per layer bitlines –L-BLj could be selected by the control line SLi by activating select transistor 2222, which its output G-BLj is one of the control line 2104 illustrated in Fig. 21B to connected together as General Bit Line-j, as presented in respect to Fig. 21A-21C herein. Herein the symbol i indicates the number of the layers in the stack, and the symbol j te the count of the control lines. Similarly, at the other edge of the unit array the per layer wordlines –L-WLj could be selected by the control line SLi and its inversion NSLi by select transistor 2224, which its output G-WLj is one of the control lines 2104 illustrated in Fig. 21B to connected together as General Word Line-j, with additional transistor 2225 to deactivate the unselected wordlines (as gate signal is preferred not to be left floating).

[000216] As an additional embodiment, the per layer select circuits could be made to either the bit lines (Fig. 22D) or the wordlines (Fig. 22E) and the select for the wordlines could be made with primarily N type transistors for which both the SLi and the NSLj signals could be routed from the control stratum. Fig. 22F illustrates a section of such partition to array of units illustrating a section of 3x3 units 2231-2239 along X-Y direction 2230. Each unit may be an array of bit-cells having wordlines 2242 traveling in the X direction and bit-lines 2243 traveling in the Y direction. The memory unit array (2231-2239) size could be about 200 microns x 200 microns while the gap between units could be about one micron to allow for the vertical connecting pads 2246 of Fig. 21A-21C, for the layer select 2244, 2268 of Fig. 22E for the wordlines and of Fig. 22D for the bitlines. In the corners between units the 'layer-select' vertical connection structure 2247 of Fig. 22A could be placed. It should be noted that Fig. 22F like many other figures herein is not to scale and the unit size (~200 micron x 200 micron) are not drawn in proportion to the size of gap (~1 micron) between units, and so on.

[000217] An additional embodiment is to have two layers select circuits for each control line as is illustrated in Fig. 23A. One select sector 2314 controlling the bitlines or the wordlines 2312 coming out from memory array unit 'n' 2304 before the connections pads 2316, and one select sector 2318 after it, in between the adjacent memory array unit n+1 2306. The select signals pads 2317 coming from the memory control circuits could include two signals SLn and SLn+1. This way the vertical memory control lines from the memory control circuits connected to the horizontal memory control lines (Word-line or Bit-lines) via pad structure 2316 could drive each of the adjacent memory array units 2304, 2306. This connectivity structure enables many use options, for example, including redundancy used to overcome defects, or multiple memory access options from single units to multiple units. A 3D computer system could leverage this flexible connectivity to blend between high speed local access with multiple processor cores, each accessing a

local memory in parallel operation, combined with global memory access in which multiple units are functioning as a larger memory array. The hybrid bonding of stratum in the 3D stack allows connecting not just the active signals but also support signals such as ground, power and feed-through 2308 as needed for such stratum within the stack and below and on top of it. The space between the memory units 2311 could be designed to accommodate the landing pads 2316 and the layer select transistor 2314, 2318.

[000218] Fig. 23B illustrates a block diagram for the generation of per layer select signal SLi. For example, a case in which 8 stratum of memory array are needed, the lower address bits A_0-A_2 could be decoded 2324 to eight layer select - SL'_0-SL'_7. In such a 3D system it could be desired to use one extra 9th stratum in which one stratum is used as redundancy to replace a defective stratum. The operation could include first a testing cycle to check if any memory cell in a unit had a defect for which the redundancy could be used. In general, the big memory may be segmented to many small units such as multiple thousands of units each about 200 micron by 200 micron. The likelihood of two defects in two overlaying units is extremely low so repair at the unit level could enable extremely high system yield. For memory unit by redundancy unit replacement, the redundancy stratum may be segmented into multiple units according to the unit size of the memory stratum. For such the 8 signals SL'_0-SL'_7 could be input to the repair control unit 2326. The repair control unit will allocate the 8 layer select to the 8 good stratum out of potential 9-th stratum leaving one stratum always unselected, by generating the proper 9-th layer select signal SL_0-SL_8. Fig. 23B illustrates an optional a sub unit repair control by having some additional address lines A_{10}-A_{11} control signals to the repair control unit 2326, so to allow different distribution of the layer select lines to the 9-th potential stratum to each quarter of the memory array. The repair control unit 2326 could be constructed to be programmable so after testing the arrays it could be programmed to avoid use of a defective sub array or array region. Using the enhanced access scheme of Fig. 23A, memory stratum of adjacent unit(s) could be used as a replacement of defective unit within a stratum if needed, thereby giving a larger range of recovery options.

[000219] Persons in the memory art could adapt these techniques in many variations to engineer 3D Computer system with the desired memory size with consideration to process yield. Such could include, having first the logic stratum then the memory control and then overlaying the memory stack, or having the memory stack first as illustrated in Fig. 11E. The 3D memory array 1131 could be a monolithic memory array or stacked memories array such as been presented in respect to Fig. 20A- Fig. 23, or both above and below. In such a 3D computer system it could be desired to include a thermal isolation layer such as layer 1157 of Fig. 11E to isolate the relatively high operating temperature of the logic layer from the memory structure. Redundancy techniques could be also used for the memory control circuits and the processing logic. The redundancy technology utilized could include techniques presented herein or in the incorporated by reference art, leveraging the unit modularity aspect of the 3D computer system and the 3D integration in which the repair could be provided in a very close proximity overlaying the replaced part and preserving the full system functionality.

[000220] In the 3D Memory stack presented herein, the unit partition could be symmetrical in which the length of the wordlines within a unit is similar to the length of the bitlines, or the unit partition could be very asymmetrical. These control line length and accordingly the size of the respective unit size in X direction or in Y direction could be about 50, 100, 200, 400 micron or even one or few millimeters. The number of connections associated with these control lines is order of magnitudes larger than the number of vertical connections associated with the access control, the per layer select (SLi). In some applications the control could be broken into a few banks, each with its own select line allowing more control flexibility to individual memory banks within the unit. Such could allow better granularity for redundancy use or parallel access to the unit memory array. These banks could be allocated horizontally (X, Y) or vertically (Z). Such could also be used for parallel access from logic overlaying and or logic underlying the array. Such could also allow for sections of the memory array to be mapped for global access across multiple units. Such variation and the support control logic to support them are known in the art and could be designed by an artisan in computer architecture and memory controls.

[000221] Fig. 24A illustrates an alternative 3D computer system utilizing the technologies presented herein. The base 2410 is a carrier substrate which is also provides cooling to the main multi cores computing stratum 2424, through a first thermal isolation layer 2426 the computer stratum is connected to the multi-unit memory control stratum 2428, which controls the multi-unit memory array strata 2430. Overlaying the memory strata is a second memory control stratum 2432 which provides second access to the same memory strata 2430. Through a second thermal isolation layer 2434 a second computing stratum 2436 could be connected to the second memory control stratum 2432. The second computing stratum 2436 could communicate with external devices utilizing a

communication stratum 2438. The communication stratum 2438 could utilize wired, wireless, optical or other communication channels to communicate with external devices. An upper heat removal apparatus could overlay the communication stratum 2438.

[000222] An additional alternative is to integrate in such 3D computing structure active cooling. Such active cooling work was recently supported by DARPA and the report on these techniques is presented in a paper by Chainer, Timothy J., *et al.* "Improving Data Center Energy Efficiency With Advanced Thermal Management." *IEEE Transactions on Components, Packaging and Manufacturing Technology* (2017), incorporated herein by reference. Such active cooling could be incorporated in addition or as replacement of the thermal isolations 2426, 2434. Fig. 24C illustrated a 3D structure with active thermal cooling supporting feed through of electrical interconnects 2472 and thermal vias 2474.

[000223] Herein the term layer transfer or layer cut could be applied to use of SiGe as a cut layer either as sacrificial layer with far different etch rate vs. silicon as presented such as in reference to Fig. 2A-Fig. 3D, or as an etch stop layer for back grinding and silicon etch scheme. Either one of these techniques could be used for the 3D system presented herein.

[000224] Additionally, alternative structures to SiGe could be used for the formation of the 'cut layer'. In some embodiments, the 'cut layer' may also function as an etch stop layer or sacrificial layer which could be selectively removed. Such alternatives have been detailed in PCT/US patent application 16/52726 and U.S. application 15/333,138, incorporated herein by reference. For example, one may use a highly doped layer of N+ or P+ or porous layers. A unique advantage of a doped layer used as 'cut layer' is the ability to make it at the processing fab as part of the conventional process flow via conventional processes such as ion implantation or in-situ doped epitaxial growth. Another aspect is the ease to make the 'cut layer' selectively using patterning which opens up more options; for example, instead of a full layer or to allow change in the layer thickness in different location across the wafer. The use of a doped layer as a 'cut layer' could be combined with other functions, such as a back bias connection for transistors or other devices. The choice of 'cut' between undercut and lift off or grind and etch back could be in consideration of the type of etch and its selectivity in respect to choice of the 'cut layer' structure.

[000225] Many other variations of 3D system could be constructed utilizing techniques presented herein or in incorporated by references. In some applications, the peripherals circuit could be placed on more than one stratum. This could be used for memory partitioning to small units such that the area of a unit is too small to fit all the require memory control on a single stratum. For example, the upper most stratum may be control logic to control about the upper half of memory stratums while the lower most stratum may be control logic to control about the lower half of the memory stratums.

[000226] Another approach that could leverage such monolithic 3D technology is multiple port access to the memory array. This could also include non-symmetrical multiport access, such as one access port could access single unit, while another access port could access multiple units. This multiport non-symmetrical access could be achieved by controlling the access to the segments of word-lines and/or bit-lines. The access from the top and the access from the bottom could be independent, yet synchronized. In such, for example, the wordlines and bitlines could be accessed by per unit memory control from the top control stratum, while the bottom control stratum provides access to the same wordlines and bitlines with multiple units control, providing one memory port access per unit from the top, while the bottom control stratum could provide access to a block of memory that could include multiple units.

[000227] The 3D memory architectures herein constructed with arrays of memory units each comprising memory strata in which every stratum has at least one select controlled from the overlaying and/or underlying memory control stratum, and the multiple options opened up by such architecture including yield repair, local and global access, is applicable to many memory technologies, including volatile and non-volatile. These architectures benefits are applicable to many of the 3D integration techniques presented herein including epitaxial based with shared lithography and layer stacking with grind and etch-back. A technologist in the art of memory systems could engineer a specific system leveraging the techniques presented herein.

[000228] Fig. 24B illustrates a generic 3D memory structure "G3DM" which could be constructed according to the techniques presented herein. Such a 3D memory could include a controller to manage the memory including self test and advanced refresh techniques. The 3D memory could include at least one or two stratum of memory control circuits, first memory control stratum 2448 and second memory control stratum 2452 and the 3D memory stack 2450. The 3D memory stack 2450 structure could be constructed and may include an array of memory units each with its own memory control structure as a tile of the 3D structure, it could also

include 3D memory array structures such as 3D NOR or 3D NOR-P disclosed herein or elsewhere. It could be provided as a wafer ready to have additional customer specific circuits, for example, such as control and encryption, which could be similar to those presented in Fig. 24A 2436, 2438. And it could be constructed on top of a 'cut-layer' so it could be cut over other structures. These additional integration steps could be done at a die level after dicing, or at a wafer level to be diced afterwards. Dicing afterward could be performed by: conventional saw dicing, laser assisted dicing, or etch assisted dicing. The structure could be useful to support more than one device size as previously presented forming a continuous structure which could be tiled to specific device size(s) near the final phase of the processing, thus allowing stocking of generic wafers, etc. The external surfaces such as 2454 could include the pads for the additional custom circuits to interface with. Alternatively, such external surfaces with the pads may be used for subsequent conventional chip packaging. The decoding circuits could be part of the generic 3D memory structure "G3DM" such as part of the memory control circuit 2448 or 2452, and accordingly the number of wires per such memory unit (about 200μ x 200μ) connectivity with the customer specific circuits could be at the range of 30-100. Such connectivity could be readily achieved with today's face to face bonding capabilities. The G3DM could incorporate self test to invoke redundancy memory stratum per unit at product release and also during normal operation to extend operations with self-repair capabilities. The G3DM could also includes wireless test and report capabilities as discussed such as in U.S. patent 9,142,553 in reference to at least Fig. 24A-C and Fig. 48-Fig. 50. The system level memory structure herein could be used for many types of memory technologies and products. A very common memory technology is DRAM for which additional enhancements could be integrated in such high granularity memory structures. DRAM is known to require refresh with the common refresh cycle of about every 60ms. The refresh rate is known to be driven by the worst case of relatively few memory cells that exhibit high leakage. Recent works have suggested adaptive refresh to reduce the refresh energy by adapting the refresh to those sections that require higher refresh rate while reducing the refresh rate for most of the device's cells. Such as presented in papers by Ahn, Jin-Hong, *et al.* titled "Adaptive self refresh scheme for battery operated high-density mobile DRAM applications," published in Solid-State Circuits Conference, 2006. ASSCC 2006. IEEE Asian. IEEE, 2006; and by Mukundan, Janani, *et al.* titled "Understanding and mitigating refresh overheads in high-density DDR4 DRAM systems," published at ACM SIGARCH Computer Architecture News. Vol. 41. No. 3. ACM, 2013, all incorporated herein by reference. The high granularity of this 3D structure with arrays of relatively small sized units could enable deploying such techniques at the unit level, or even at a layer of a unit, so the refresh rate could be reduced to units that either have not be written into yet, or that exhibit lower leakage. In addition, for a unit in which all the memory layers are good, the choice of the unit to be left unused could be based on refresh needs.. The partition to units with the associated reduction of the word-lines and bit—lines length could itself reduce leakage and accordingly the required lower refresh rate. Alternatively these techniques could enable the reduction of the DRAM capacitor size for some application which could enable significant overall memory cell size reductions.

[000229] Memory centric applications such as intelligent systems or search applications could be implemented as a memory focused processing system utilizing such 3D systems as is illustrated in Fig. 24A. In such a system, a new approach could be used, instead of the conventional processor centric approach in which data is transferred to and from the central processing unit, transfer the process to where the data is. As an example, the dashed border 2462 of Fig. 24A may represent a database stored in the memory 2430 associated with people in city A while the dashed line 2464 indicates data bank stored in the array 2430 associated with people in city B. And if a search is needed in respect to city A the program performing the search could be transferred to the processing unit in the logic layer 2436 located in the area marked by 2462, while the program code related to a search in city B could be transferred to the processing core located in the region marked by 2464. Memory centric system operations could leverage the 3D computer system illustrated herein as a new compute paradigm. The program or code itself could also be stored in the memory matrix 2430. Additional option is to run parallel processing on the memory stored in the memory matrix 2430 converting it from one form to another form. There are many form of data transfer such as from amplitude domain to frequency domain as often called Fourier Transform. Another type of transform is to form multiple feature plans with one or very few bits from a many bit per data point representation, which is useful for technique for brain inspired algorithm.

[000230] Another alternative for constructing a 3D memory stacked structure is to reduce changes in the memory wafer processing and compensate by adding process steps to the stacking process. Fig. 25A illustrates structure 2500 as alternative in respect to the

structure of Fig. 21E. A wafer substrate 2501 with cut-layer 2502, memory semiconductor structure 2504, memory interconnection structure 2506 and oxide layer 2508 may be formed. The in-between units control line 2510, through memory via 2536 and connecting pads 2538 are formed similar to those of Fig. 21E. The difference is that the in-between silicon 2505 is not etched away but rather shallow trench isolation 2512 and bottom dummy contacts 2514 are added, using conventional memory processing. The bottom dummy contact 2514 may be formed at the same step as the source and drain region of memory bit cell formation. The bottom dummy contact may be an n+ diffusion region. The through memory vias 2536 are connected to the respective bottom contacts 2514. Fig. 25B illustrates this alternative in respect to Fig. 21H after having a structure such as 2500 flipped and bonded to similar structure 2508, and have it substrate and cut layer removed, thus shown as processed bonded layer 2506. Then, vias 2582 could be opened by removing silicon 2505 to expose the bottom side of the bottom contacts 2514 as is illustrated in Fig. 25C. If the STI 2512 of the standard process are not deep enough, then the via formation process could include the STI locations to etch through to assure full isolations of the vertical connection from the substrate and each other, thus forming full etch structures 2583. Alternatively, the silicon at the regions could be first be etched out and replaced with isolation material. Then add landing pads 2522 on top of bottom contact 2514 as is illustrated in Fig. 25D, to prepare the structure for the following layer.

[000231] An additional step could be added which is forming alignment marks for this stacking process. The bonding alignment mark could be included in the metal layer as the bonder could see these alignment marks from the top view of the wafer. Fig. 25E illustrates the structure of Fig. 25A with alignment marks 2532 utilizing the STI process. These marks could be used once the wafer has been flipped and the substrate and the cut layer have been removed. Other alternatives could be the use of an ion implant process or leveraging the contact process for the alignment marks.

[000232] Additional steps that could be taken in the memory fab to help the following stacking process could include using a lithographically defined doping process. Fig. 25F illustrates an optional use of N+ doping for extending the contacts 2514 into the silicon with conductive N+ silicon 2534 reducing the need of forming the metal connections illustrated in Fig. 25C-25D 2502, 2522. Such N+ doping can be accomplished by adding extra high dose deep ion implantation in the memory wafer fab or even attainable using a part of a standard process step. The depth of N+ layer may be the substantially close to the body thickness of the stratum to be transferred, so that N+ regions play as nano-TSVs. Fig. 25G illustrates an optional use of N type silicon as the cut-layer 2540, and leaving the regions 2536 for the nano-TSV as P-. Once flipped and bonded, the process of substrate removal using the N as etch stop could allow etching these P- regions 2536 exposing the contacts 2514. Filling with conductive metal and then removing the excess using processes such as CMP would make the stack ready for the next stratum. As discussed previously, the selective etch of P silicon to N silicon could be an anodizing process which would first etch the P silicon to become porous and then the porous silicon could be selectively etched away. Utilizing these techniques the memory array wafer could be processed using a standard memory process or such with simple changes, and then stacked using simple processes that could be just bond, grind and etch, or with some additional steps as presented herein.

[000233] As a general note, the use of top pad and bottom pads herein are exchangeable as with the use of layer transfer techniques. These structures could be flipped for specific applications using the presented technology and structures herein. In some cases there might be a need to flip the layer before bonding it to the target wafer. A carrier wafer, such as presented in at least US patent 8,273,610, incorporated herein by reference, could be used to support such flipping. The carrier wafer could also leverage techniques presented herein in respect to the term “cut layer”, and could be designed to be grind and etched out, or to be reused having it ‘refurbished’ and used again. Additional techniques for such a carrier wafer, could be to form a porous layer at the top of a carrier, such as presented in respect to the ELTRAN process, without the need for the epitaxial step but rather just use it with silicon top or add oxide for the bonding. Another option is to use a wafer with thick oxide and/or a nitride cover of a few microns and optionally add grooves at the dice lane or between lithographic projection fields. Then detach the carrier wafer by a though-side etch leveraging the very high selectivity of etch rates between silicon and oxide or nitride. An additional alternative is to implant ions such as a combination of helium and hydrogen and then use low temperature (~400 ^{o}C) ion-cut for detach. An example for a need of flipping is in a case when the desired landing pads 2006 are in a range of about 200x200nm^2 or about 400x400nm^2 while the designated location for these pads might be desired for operating silicon. In such case a via smaller than 100x100 nm^2 through the transferred silicon film

(strata) could be used and the landing pads could be constructed over the carrier wafer, once the layer was transferred onto the carrier wafer.

[000234] An additional enhancement could be by adding to through-strata-via that we could also call a through-layer-via 'TLV" such as illustrated in Fig. 20H including top landing pad 2054, a via or chain of vias 2050, and a bottom connection pad 2052. Such connection path might serve as path through without connecting to other elements within the strata it is passing through. It could allow signal paths through such as connecting signals between first memory control stratum 2448, of Fig. 24B, and second memory control stratum 2452 without connecting to any element within the 3D memory stack 2450. Furthermore, a plurality of dummy vias not connecting any pad may be included to improve process uniformity, serve as mechanical support, or in some application when it is desired facilitate heat dissipation, such as thermal pipes and paths to an external surface of the device.

[000235] An example for such a feed-through TLV is illustrated in Fig. 24A. A base wafer carrying, for example, processors and other circuits 2410, 2424, could be sourced from conventional 2D fabrication process and facility. It could include connecting pads on its upper surface. Then a wafer carrying multilayer memory strata 2430 with its memory control circuits 2428 could be bonded on top of the sourced wafer forming the connection between the standard flow wafer and the stack memory and control circuits. Then using the feed-through TLV connections could be made with the upper stratum 2436 which could include the I/O circuits to interface the 3D system with external devices. These could also include wired connections such as, for example, pads, balls or pins, or wireless systems such as electromagnetic, optical, etc. This heterogeneous integration supports the use of different device crystalline material, RF, Analog and other forms of heterogeneous integration. Such could include magnetic films using technique such as presented in US patents 9,337,251, 9,357,650, 9,357,651, 9,647,053, 9,679,958, incorporated herein by reference. These ferromagnetic films could be added to construct an on-chip inductor using standard metallization layer(s) with high quality factor for voltage regulators or RF transceiver/receiver to improve the 3D system effectiveness and capability.

[000236] A standard wafer fabrication technology or baseline technology could be established for the memory per unit pin out position and function. That standard wafer fabrication technology or baseline technology could also be used for the custom logic design so it could integrate the generic memory wafer presented herein , for example, by bonding. Each standard wafer could include alignment marks for the custom logic top layer to help align the generic memory wafer during the bonding process. The standard wafer could include processing cores compatible with the size of the memory unit such as, for example, about 200 μm by 200 μm, a street width between units such as, for example, about 1 μm. The signals to be connected in-between such as: 40 pins for address, 16 pins for data, 10 pins for control (such as read and write) and 4 pins for pass through paths. Some of these pins could be defined in the industry standard as expansion options or to allow more than one memory type or architecture. With about 100 pins per unit, the area for each pin could be about 20 μm by 20 μm, which allows the use of most wafer bonders available currently in the industry. Additionally, the generic memory and control stack could be designed to be about 50 μm thick so it could be shipped, handled, and bonded by industry standard processes and machines. Such could become also a standard for which the memory stack could include a path-through the interface layer with the proper thickness so the total stack would be about 50 μm thick. For example, for a 16 memory layer stack of 1 μm each and control stratum with I/O stratum of 2 μm, the stack thickness could be about 18 μm, and a path-through layer of about 32 μm could be bonded on top to bring the overall stack thickness to about 50 μm thick, compatible with the current industry capability. The pass through paths could be built using technology such as TSV to pass, for example, the approximately 100 signals from the generic memory to the custom 2D processor device, to the processor device such as based on planar, SOI, FinFET, or gate-all-around technology.

[000237] Fig. 29A illustrates such 'pass-through add-on structure' 2902 constructed within a silicon wafer substrate, having base substrate 2910, designated cutline 2908 that could be formed by cut-line technique presented herein or just by timing the grinding and etch once bonded as it design to be many microns thick. The 'pass-through add-on structure' could include pre-built isolated TSV 2904 to function as the conductive feed through. An enhancement of such 'pass-through add-on structure' is to have it function also as a thermal isolation between the heat generating processor and the memory stack, for example, as is illustrated in Fig. 29B. The enhanced 'pass-through add-on structure' 2912 could have some of the silicon 2906 be replaced by etch, deposition and planarization with thermal isolating material 2917 such as silicon oxide. With feed-through TSV 2914 and the remainder of the silicon substrate

2916 similar to those illustrated in Fig. 29A. Those 'pass-through' structures could be reserved for the arbitrary purpose that can be determined later by the system designer. The strata containing 'pass-through' structures may be a generic product such as memory, sensor, power or communication chip. Alternatively, use of such feed-throughs could be, from the processor logic in the logic layer 2424, to the I/O layer 2438 to connect to an I/O or analog function being placed on the 2438 layer, or to interconnect memory control layer being placed on top 2432 and under 2428 the memory structure 2430. Such could be used to coordinate use of the redundancy control function between the top control 2432 and under control 2428. An additional use could be to coordinate dual control to support multiport access to the same memory matrix.

[000238] The technique of using a precise bonder with staggered pads on word-lines or bit-lines presented in respect to Fig. 21A-21C, could be used to add memory control logic over 3D memory such as 3D NAND or 3D NOR. This could be an alternative to the technique presented here in respect to Fig. 11F-11K and Fig. 12F-12J. An advantage of a staggered pads approach could be the parallel processing of the memory and memory control which then could be integrated into a 3D device by bonding.

[000239] The stacking techniques presented in respect to Fig. 20A-Fig. 25G, and Fig. 29A-29B could be extended to 3D memory such as 3D NOR. Such could include stacking multiple strata of 3D NOR each having multiple layers of memory. In such case, the per layer select transistors of Fig. 22D-Fig. 22E could become per memory strata select. This per layer select transistors could be implemented within a control stratum on top or below the multilayer memory strata or alternatively without such control stratum but rather in the memory structure such as by doubling the ridge select transistors to have to serial transistors instead of the one ridge select transistor (such as 1213, 1222, 1332). In such case, one transistor could function as the ridge select while the other could be globally controlled by the per layer select signal SLi, effectively providing strata select to support simple vertical stacking. This could allow use of the same memory control circuits to control multiple strata with relatively simple additional control and z-decoding circuit to control the strata select signals. Such stacking could include periodical stratum of rebuffering and redrive to support the relevant control lines. Fig. 26H illustrates an X-Y 2690 top view of such added transistor for forming Strata-Select SS1 2692, SS2 2694. Fig. 26H is a portion of Fig. 12E with the added transistors 2692, 2694, 2696 to function as global strata select for such stacking of 3D memory structures. In such case the stacking flow presented in respect to Fig. 21A-24B could be adapted for stacking of 3D memory strata such as 3D NOR. The 3D memory strata could be designed with units such as presented in respect to Fig. 22F. These units could be structured similar to the 3D memory units such as presented in respect to Fig. 12D. Each such unit could be designed to have the wordline access both from the top and the bottom and each of those word lines could include stacking pads arranged such as presented in respect to Fig. 21B. The bit-lines could also be structured to be access from the top and bottom and to have landing pad structures similar to Fig. 21B. And similar pad structures could be added to the ridge selects. For the Strata Select, the pad structure could be designed similar to the structure of Fig. 22A.

[000240] Additional option is to stack different memory type. Clearly stacking could include many type of stratum, yet the unique aspect of the stacking technique in here is to form vertical connection of the word-lines and the bit-lines while having per strata select line such as SLi of Fig. 22D-22E. These word-lines and the bit-lines could control plurality of memory type including volatile and non-volatile. Sharing these memory control lines could allow efficient form of memory structure and in some cases, could even allow direct data transfer from one memory type to another memory while keeping some of these control line unchanged reducing power for such data transfer and reducing time for such transfer.

[000241] An additional step that could be included in the preparation of 3D memory structures for stacking is adding bonder alignment marks. Precise bonders generally need alignment marks to align wafer to wafer These alignment marks could be incorporated in the proper location on the top and/or bottom layers of the 3D memory wafer structure using a proper mask of the 3D memory structure such as in a nonfunctioning zone over the wafer such as in-between memory structures.

[000242] The technology for precise wafer bonding is being enhanced, recently demonstrating the improvement of wafer to wafer alignment tolerance from 200 nm three sigma to 100 nm three sigma, and efforts are underway to develop bonding precision to 50 nm. These works have been reported by papers such as by Peng, Lan, *et al.* "W2W permanent stacking for 3D system integration." *Electronics Packaging Technology Conference (EPTC), 2014 IEEE 16th*. IEEE, 2014; by Sakuma, Katsuyuki, *et al.* "Bonding technologies for chip level and wafer level 3D integration." *Electronic Components and Technology Conference (ECTC),*

2014 IEEE 64th. IEEE, 2014; by Sugaya, Isao, *et al.* "Precision wafer bonding process for future cost-effective 3DICs." *Advanced Semiconductor Manufacturing Conference (ASMC), 2015 26th Annual SEMI*. IEEE, 2015; and by Kurz, Florian, *et al.* "High Precision Low Temperature Direct Wafer Bonding Technology for Wafer-Level 3D ICs Manufacturing." *ECS Transactions* 75.9 (2016): 345-353, all of the forgoing are incorporated herein by reference. Herein we presented at least two stacking flows, one using 'Smart-Alignment' techniques as presented in reference to at least Fig. 11F-Fig. 12J, and one utilizing 'Hybrid' and/or 'Fusion' as presented in respect to at least Fig. 20A-Fig.23B. The fusion bonding techniques enable a simple stacking operation as it could be done without the need to form vias in the transferred layer thereby reducing the need for lithography, metal deposition, and processing per stack layer at the stacking fab. But for such fusion bonding based stacking, the landing pads as illustrated in Fig. 21A are required. In some applications, it might be effective to mix these stacking techniques for the formation of 3D structure such as, for example, the one illustrated in Fig. 24B. An example for such could be to use fusion bonding for the stacking of the memory stratum as related to Fig. 21A-Fig. 23A, and use 'Smart-Alignment' for connecting the memory control strata 2428 (providing the decoding, sense amplifiers and other memory control). Precise bonders could align the wafers being stacked but could not overcome die level misalignment within these wafers. It is relatively easier to control the die level alignment for wafers processed in the same processing fab line using the same stepper, or matched steppers. The memory stratum could be processed in the same line as they are providing the same circuits. The control stratum would be processed in a different process and might be harder to achieve the same die to die misalignment as the memory stratum. By using 'Smart-Alignment' for it, the landing pads could be exchanged to landing strips as illustrated in Fig. 11G and could allow a larger overall misalignment between the control stratum and the memory stratum. Such hybrid stacking could keep a tight space between units of the memory stratum with landing pads, as an example, 200 nm x 200 nm, while the connection of the control stratum to the memory stack could use landing strips of 300 nm length (and less than 80 nm width) to accommodate the additional die level misalignment between the control stratum and the memory stratum.

[000243] An additional option to accommodate large total stacking misalignment is to build the relatively larger landing pads and pins over the memory unit. This way the space 2311 between the memory units could be kept relatively small while the landing pads could be made large enough to accommodate the total error which could include the bonder alignment error or/and the die placement errors. Such over the array pads construction could add costs associated with the processing of such over the array pads and additional per stacking layer costs in the stacking fabs to build these extra layers of landing pins. Fig. 25H-25J could illustrate such over the array connectivity structures. Fig. 25H is lower portion 2560 is similar to Fig. 25A. The upper portion 2561 illustrates the landing pads 2550 constructed over the memory bit cell array 2556. These landing pads are connected to the respective bit-line or word-line by via 2552, connection wire 2554, and via 2562. These additional metal layers are added on top of the isolation covering the array 2558 and are with the additional isolation material 2557. It might be desired to have two versions of over the array pads: one with landing pads 2550 and one with landing pins 2570 as illustrated in the lower part of Fig. 25I. Fig. 25I illustrates bonding a such wafer with large landing pads 2550 on top of another such wafer with landing pins 2570. Fig. 25J illustrates the structure after removal of the top substrate and SiGe layer and constructing landing pins 2572 or alternatively large landing pads (not drawn) for the following stacking step. Artisans in the art could mix and match this alternative stacking option to the specific 3D system design.

[000244] The 3D memory stack herein enables stacking multilayers of memory stratum in which the vertical connectivity is at the word-line and bit-line level. Such 3D stacking enables use of a memory control for multiple memory stratums thus reducing cost in addition to benefits in performance and power reduction. Yet vertical connectivity at the word-line/bit-line level could be a technology challenge as the high pitch of these memory control lines may prevent the use of techniques such as 'smart-alignment' as there might not be enough room to run TLVs through. For such cases the Hybrid/Fusion bonding techniques presented herein, at least in respect to Fig. 21A-Fig. 25J, are an effective technique to provide word-line/bit-line level 3D connectivity.

[000245] The 3D memory stacking presented herein could be modified to accommodate technology limitations or cost objectives. Such modification could include connecting only the bitlines at the unit level while connecting the wordlines at a far courser granularity or vice versa (connecting the wordlines at the unit level and the bitline at the bank on multiple unit level). Other modifications could include staggering the layer select transistors position of Fig. 22D or Fig. 22E, to accommodate the high metal

pitch of the respected control lines. Another modification could be to position the connecting pads for/on odd control lines in one side of the unit and for/on even control lines on the other side of the unit thereby resembling the structure illustrated in Fig. 12C-12E.

[000246] An additional alternative to a form buried cuttable layer is to replace the buried SiGe with oxide, nitride or other layers. This could be done following a step of isotropic etch as been described such as in reference to Fig. 2I. This could use deposition techniques such as ALD or other conformal deposition techniques to refill the etched-out space with a proper dielectric material. This could be done while forming and keeping posts to hold the top silicon layer or in two steps; first portion and then additional etch and replace the remaining portion. Following the replacement an epitaxial process could be used to seal the entrance holes 224. This could be done for the whole wafer or at selective sections for specific applications. This could be done as part of forming a generic substrate or per a specific application. The replacement could be done for specific circuit considerations, for example, such as substrate capacitance, or back bias, or back gate. This could also be done to support any of the 3D integration flows such as presented herein for which the replacement material could be later etched with a much higher selectivity. For example the selectivity of dry etch of nitride vs. silicon could be set to be 2,000:1, which is much higher than the selectivity of SiGe vs. silicon. This could also be done in combination with bonding to another substrate and detaching with a finishing process such as residue etch CMP and epitaxial. This flow could also be done as an alternative process for the formation of SOI substrates, and may have lower cost to manufacture compared to the current methods. This could also be done using a stain etching to convert the SiGe layer to porous layer as previously presented herein.

[000247] An additional alternative for a 'cut-layer' is to use a single atom layer of Graphene as presented in a paper by Kim, Jeehwan, *et al.*, "Principle of direct van der Waals epitaxy of single-crystalline films on epitaxial graphene." Nature communications 5 (2014); and Yunjo Kim, *et al.*, "Remote epitaxy through graphene enables two-dimensional material-based layer transfer" published at NATURE | VOL 544 | 20 APRIL 2017, incorporated herein by reference. It was discovered that a single atom layer of graphene being placed on a single crystal substrate could allow a single crystal epitaxial growth on top having the base crystal orientation and quality. Yet the layer grown on top of the graphene could be pulled off as the graphene layer has "weak van der Waals interactions, and which also allows facile layer release from 2D surfaces". Enabling "the grown single-crystalline films are rapidly released from the graphene-coated substrate and perform as well as conventionally prepared films". Accordingly, such a single atom graphene layer could serve as alternative to the porous layers described herein or combined with such or other forms of 'cut-layer' presented herein. The base substrate could be reused after 'cutting' off the functional layer. A graphene cut could be used in a similar way to the original concept of porous layer for formation of SOI wafer as was named ELTRAN by Cannon.

[000248] Many mix and match of these cutting techniques could be utilized for different product formation and related flows. One such mix could be used for die to wafer 3D integrations as discussed in U.S. patent applications 15/095,187 and 15/173,686 and herein. So, the cut of the 6 microns thick die could use the graphene as the cut layer but then the following a step of thinning the layer after being bonded to the target wafer, and could leverage the SiGe etch selectivity for etch and controlled thinning to below 1 micron to allow a simple process with nano-TSV through (with less than 400nm via diameter) the thinned die.

[000249] Additional technique that could be used for a 2D material such as Graphene as a 'cut layer' is an oxide type post that could be etched out prior to the layer transfer step. As the substrate with a cut layer being built in could go through the full front end of theline processing and some back of the line processing before the layer transfer, it could be desired to add in such posts to keep the stability of the structure for the various processing steps prior to the transfer step. Using a modified STI step, holes could be etched all the way through the graphene into the underlaying substrate, and filled with oxide. These holes could be made in the dice lanes. Then as one of the last step before performing the layer transfer operation these oxide posts could be etched away releasing their hold. Additionally, these in the dice lanes could be extended to a full dice lane etch so that in the layer transfer step each die may be peeled off independently from the other dies.

[000250] The release process could include a polymer or other material such as nickel to help form a stress which together with temperature, such as liquid nitrogen or less than 400 °C degree spike heating, could help the detach and release of the re-useable substrate from the 3D structure comprising the target wafer and the bonded transferred layer. An alternative technique could include the use of pulling 5-30 micron thin layers off reusable wafers using a technique called controlled spalling such as presented in papers by Shahrjerdi, Davood, and Stephen W. Bedell. "Extremely flexible nanoscale ultrathin body silicon integrated circuits on

plastic." Nano letters 13.1 (2012): 315-320; and by Bedell, Stephen W., et al. "Layer transfer by controlled spalling." Journal of Physics D: Applied Physics 46.15 (2013): 152002; and U.S. patents 9,698,039, 9,704,736 and 9,713,250, incorporated herein by reference. A reusable "cuttable" substrate could be constructed using the following steps; 1) Form the thin layer with desired silicon thickness on top of a SiGe etch stop layer using an epitaxial process over a donor wafer. 2) Form a reusable carrier by growing 3-10 micron thick oxide (or nitride) over a silicon wafer. 3) By using controlled spalling pull out 5-10 micron layer off the silicon over SiGe (top Si/buried SiGe/fractured bulk Si stack) from the donor wafer and bond it on top of the reusable carrier, thus forming the reusable "cuttable" substrate. Optionally, the fractured surface portion of bulk Si maybe treated to be planarized for better bonding to the resuable carrier. The reusable "cuttable" substrate could now be processed with building the desired circuits on top of it. Than it could be bonded on top of a target wafer. Then using selective oxide or nitride etch from the side of the wafer, the bulk of the reusable "cuttable" substrate could be detached leaving over the target wafer the circuits and the layers previously being bonded to the 3-10 micron thick oxide. Than using SiGe as an etch stop the 5-10 micron silicon could selectively etched followed by a SiGe etch. In the process the edge of the wafer could include protection of the interconnect layers to protect them from the side oxide detaching etch.

[000251] The 'cut-layer' technology presented herein could also be used for applications requiring a very thin device. An example of such application is integrating a semiconductor device in a contact lens or in application requiring a very flexible circuit layer. In these applications the ability to use a standard semiconductor fabrication process following by thinning the device thickness to a few microns or hundreds of nano-meters or even less, could be key enabling technology.

[000252] An additional inventive embodiment for a 3D memory constructed of arrays of relatively small memory units, with the memory control circuits on top or under of such memory units relate to the ability to perform per unit refresh and other techniques to extend memory effectiveness. This could be applied for DRAM type memory as presented herein before and also for non-volatile memory such as charge trap, floating gate and ferro-electric based memory. These memory units could have an X direction and/or Y direction size of a few tens of microns, or a few hundreds of microns. For example, some of the general concerns with memory structures relate to disturb and other forms of losing memory fidelity. These could impact the level of memory density utilization. With such a 3D memory system as illustrated in Fig. 19A, the control system could copy the contents of a memory unit to cache storage, erase the unit and re-write the content to restore memory fidelity. These refresh cycles could be performed based on time or activity of that memory unit. These refresh operations could be performed at a time there is no active use of the memory so auto-maintenance could be performed. Fig. 27 illustrates a block diagram for such a refresh operation flow. Such a refresh could extend the effectiveness of the memory by enabling a greater number of bit site locations with the memory cell and a greater number of storage levels within such storage sites.

[000253] An alternative to form the 3D NOR fabric illustrated in Fig. 3A-3B and Fig. 4A-Fig. 10D of PCT/US16/52726 is to use the "punch and plug" technique, commonly used in current 3D NAND formation. Fig. 26A illustrates a top XY 2600 view of holes 2608, 2609 formed in a multilayer structure such as is illustrated in Fig. 3A of PCT/US16/52726. The multilayer structure may be the stack of multiple single crystalline N+/P/N+ semiconductors layers. The multilayer structure may further include silicon for N+ layer (designated for the source and drain S/D) and SiGe for P layer (designated for the channels). Alternatively, the SiGe region may serve as S/D while Si may serve as the channel material. The memory transistor may be constructed for single crystalline vertical channels and single crystalline horizontal bit lines as presented in reference to Fig. 4A-Fig. 10D of PCT/US16/52726. Herein, the term "punch" represents a deep etch process to make a hole deep through multiple stack of layers while "plug" denotes a deposition process that either substantially fills the punch hole or partially fills it, such as a layer or layers deposition on the inner sidewall of the punch hole. The holes are usually punched substantially all the way through the multilayer structure. The holes could be formed in horizontal rows 2602, 2604, 2606 which could function similar to the valleys 308 in Fig. 3B of PCT/US16/52726. After the vertical anisotropic etch forming the holes, an isotropic etch could be used to extend the holes to slightly overlap as is illustrated in Fig. 26B. Alternatively the holes could be formed directly as illustrated in Fig. 26B. Then using deposition, such as ALD, the O/N/O layers 2628, 2629 could be deposited in the holes, similar to as is done in 3D NAND, as illustrated in Fig. 26C. The gap opened by the overlaps 2618, 2619 between the neighboring punch holes, could be designed so, the deposition of the O/N/O layers would close it, by having the proper O/N/O thickness, forming rows of isolated holes, 2622, 2624, 2626, with isolated ridge 2623, 2625 which function similar to the

ridges 309 in Fig. 3B of PCT/US16/52726. Such auto-shutting of the gaps enables for the subsequent gate formation to be self-aligned, reducing a lithography step for the gate patterning. Then the inside of the holes could be filled with gate material 2630 which could be polysilicon, tungsten or alternative gate material or combination of such. The O/N/O stack thickness is usually more than 10 nm and less than 30 nm, and accordingly the gap between the holes 2619, 2629 could be designed to be less than 20 nm to assure isolation between the gates of adjacent holes. Fig. 26D illustrates having the gates along the same column in Y direction connected by global wordlines 2632, 2634. The holes of even rows 2604 could be phased from the holes on odd rows 2602, 2606, to simplify the wordlines connections, to enable individual selection of each storage facet as an alternative to the connectivity in Fig. 5A or 5B or 5E or 8A of PCT/US16/52726. Most of the variation and enhancement presented in the PCT/US16/52726 could be adapted and implemented with such "punch and plug" process. For example, the indentation presented in respect to Fig. 10B of PCT/US16/52726, could implemented by having 'odd' holes 2608 larger than 'even' holes 2609. Furthermore, the formed ridges 2623, 2625 could be sliced with slits ('valleys') to expose the sidewall of multilayer structure and selectively expose the bitlines – S/D region for silicidation and so forth as explained in reference to Fig. 8 of PCT/US16/52726. An artisan in memory technology would be able to adapt many of the technology alternatives presented in PCT/US16/52726 in respect to a straight lines ridge in valleys to the punch holes formation of ridge and valleys as illustrated in reference to Fig. 26B.

[000254] An additional advantage of the punch holes technique illustrated in Fig. 26A-26C is the ability to use the adjacent wordline to steer storage location. So, for example, when writing using WLn the adjacent wordline WLn+1 could be used to pull the charge toward it or to push charge away resembling the concept presented in PCT/US16/52726 in respect to Fig. 10B and Fig. 10E. Accordingly using both WLn-1 and WLn+1 could extend the storage capacity when writing using WLn.

[000255] An alternative process could include sealing some holes such as the 'even' holes 2509, then plugging the 'odd' holes 2508 with O/N/O and gates, then remove the sealing and optionally isotropic etching the unplugged 'even' holes 2509 . Such selective sealing process is often used for the in-situ sealing process to create a vacuum cavity in MEMS technology by using very low step coverage deposition process or very non-conformal deposition process, causing voids. In order to protect the multilayer structure by some residual sidewall deposition of the sealing material, a dummy mask pattern on the very top of the multilayer structure may be incorporated. In this approach, the holes to be sealed first have a substantially smaller diameter than the holes to be plugged first. Holes sealing could be done by processing such as presented in U.S. patent application 12/979,592, incorporated herein by reference.

[000256] Fig. 26E is a top view X-Y illustration to show the use of the single hole punch process to construct the various elements which may be needed for the 3D NOR fabric. The region cut is illustrated by the dash line 2660 and the in picture cut 2662 is to indicate that the structure could include many more memory cells in the X direction. The structure resembles the structure of Fig. 12D. The ledger could be read as follows, the un-punched multilayer structure 2640 forming the in bitlines ridge select transistor 2642 and in-silicon layer transistor 2644 used as part of the per layer programming (PE1-9 of Fig. 28), 1st gate pillar 2646, 2nd gate pillar 2648, control gate 2650 of the in-silicon transistor for ridge select and PE gate, a vertical pillar for ground 2652 for all layers (2800 of Fig. 28), programmable gate 2654 for the PE transistors, isolation pillars 2656, contact pillars 2658 for per layer contacts (L1-L4, 2812-2818), optional feed-through pillars 2639 which could be used to transfer signals from the upper side of the fabric to the bottom side. The processing of these holes could be done for each function while the other holes are sealed or by other alternative techniques presented herein. Processing these punch holes together saves processing cost because in 3D memory structures the holes punch process through the multi-layers structure is a slow and expensive process. In many cases the holes diameter is about 100nm or smaller yet the multilayer structure could be of a few microns thick.

[000257] An additional advantage of simultaneous holes punching is having these holes self-aligned in the vertical direction enabling dense structures as illustrated in Fig. 26A-26E. Fig. 26F and Fig. 26G are side cut views along X-Z direction 2670 to illustrates the difference between holes etch together/simultaneously (Fig. 26F) vs. holes punch in two independent etch steps (Fig. 26G). The self-aligned holes 2674, 2676, 2678 in a multilayer structure 2672 could be used for different functions as illustrated in Fig. 26E. Fig. 26G illustrates holes etched into a multilayer structure 2682 in which holes 2684, 2688 are etched in one step while holes 2686 in another step. With 'Single Punch' the sidewall vertical scallop pattern of the etched hole is substantially the same as the sidewall vertical

scallop pattern of the other holes etched at the same time, being processed at the same process into a uniform structure. The peak and valley pattern of the scallop is continuous along the X-Y direction between these simultaneously etched holes.

[000258] A known challenge in 3D memory formation relates to the etch aspect ratio for the holes punching process. At the current state of etch technology it is about 1:60 which imply that for multilayer substrates with 3 microns thickness the smallest consistently attainable hole would be about 50 nm diameter. To keep the holes diameter small for thicker multilayer substrates the following techniques could be applied. A layer transfer technique could be used to punch holes from both sides of the multilayer substrate enabling about double the multilayer thickness. And successive holes punching, followed by epitaxial growth of the substrate, could be applied to construct a thin hole in a thick multilayer structure.

[000259] An additional alternative that could be combined with many of the 3D-NOR structures presented herein and in at least PCT/US16/52726 could be to use the ridge split process to replace the SiGe portion with metal to function as the S/D and having the Silicon portion function as the channel. For example, in respect to Fig. 41E of PCT/US16/52726, instead of a partial SiGe etch perform a full SiGe etch and then either use a thin oxide first or just perform conformal deposition of metal to be followed by an etch removing the metal from the side walls and leaving it as replacement of the SiGe regions, thus functioning as S/D. For such alternative, the silicon region should be kept un-doped or P- doped to function as the 3D NOR channel regions. At the edge of the ridge the proper adjustment should be made in forming the staircase access. These adjustments could include replacement of region of the metallic S/D with p type silicon for the formation of the ridge select. Alternatively an epitaxial step of N+ type silicon could be used to form the S/D. An additional step of etching the ridge split again could be used to clean the slit from the epitaxial over growth and reduce the risk of leakage between S/D lines. Additionally the epitaxial step could be engineered to only partially fill the space formed by the SiGe removal, followed by adding metal as described above, combining both techniques to replace the SiGe forming the S/D. An additional aspect that could be integrated in the 3D NOR formation is to use a high work-function metal for the wordlines to reduce punch-through leakage risks. Persons in the memory art could adapt this alternative to the various relevant structures of the 3D NOR memories presented herein.

[000260] The 3D integration technique such as presented in respect to Fig. 1 to Fig. 6D herein could be used to support many derivatives of memory products by a mix and match of memory control circuits on top or below the memory array. These could allow multiple control circuits, some that use multi-bits per cell and some that do not, and also the many techniques of multi bit presented herein. The memory array could also be made with a generic size which could then be customized for a specific memory product by properly designing the memory control circuit and sizing the end product by the placement and setting of the dicing lines.

[000261] An additional inventive embodiment to enhance 3D memories is to use an alternative per layer access by sidewall strapping through one-time-programmable anti-fuseto the method and structure presented in reference to Fig. 43 of PCT/US2016/052726, or by similar multiple programmable connections such as is used for RRAM and Bridge-RAM. Fig. 28 is a modified structure to support this alternative method. L1, L2, L3, L4.,.(2812, 2814, 2816, 2818) are the vertical contact pillars to serve as per layer access. Each one of these contact pillars could be programmed to connect to a Bit-Line (BL1-BL4) by breaking a thin oxide (OTP), which could be considered an anti-fuse between it and the Bit-Line (2841, 2842, 2843, 2844). The Bitlines and/or pillars could be engineered to provide a more stable anti-fuse providing an ohmic connection when linked; for example, by include a high concentration of atomically large lattice atoms, such as Arsenic in silicon, which would be incorporated into the link when fused. The pillar maybe formed using polysilicon with the same doping polarity of BL to provide an ohmic anti-fuse. The connection region could include the horizontal 'programming-enable-transistors' PE1-PE9 controlled by gate pillar 2803 having a contact 2832 PE, which provides controlled connection to the vertical grounding pillar 2801 with ground contact 2800 GND. The horizontal transistors PE1-PE9 are embedded in the Bitline using a similar technique presented herein for the ridge select transistors. These transistors comprise charge trap so they could be programmed to be disconnected. The programming could be performed by the memory control circuit 2850 using the Vpp Gen to generate programming voltage and 'P Enable' to enable the programming, as follows:

1. Initially, all PEs (PE1-PE9) are erased (to low Vt) to pass ground potential to their respective BLs (BL1-BL9).
2. Vpp is set to a high enough voltage to break the 'anti-fuse' made of the thin oxide (OTP), than P1 is activated presenting the programming voltage on L1. Then 'P Enable' activates the gate PE opening up all the horizontal programming

enabling transistors (PE1-PE9), connecting all the Bitlines to ground. One of the anti-fuses would break which would connect L1 to one of the Bitlines (randomly). Let say (for example) BL1 is connected to L1. Then the programming voltage drops as the current through the activated anti-fuse pull the voltage on L1 down enough so no more anti-fuses would break. A soak algorithm could also be initiated to make the anti-fused link more stable.

3. Then Vpp and PE could be set to program PE transistor (to high Vt) connected to the randomly fused BL in step 1 to stop passing ground voltage. In this case, PE1 transistor is programmed to high Vt. Now, only the rest of the PEs excluding PE1 could be passing ground potential when PE is activated.
4. After P1 is disabled, P2 is enabled and the cycle repeated creating connection between L2 and random BL.
5. Repeat step 2 to 4 for rest of BLs.

The selection of which Bitline would be fused first could be guided by changing the vertical pillar ground connection 2801 from metallic pillar to resistive pillar-like poly silicon which could give advantage to the upper Bitline closest to the GND to fuse first. This concept of random (or guided) selective fusing by parallel access to multiple anti-fuses is been implemented for random number generator as is detailed in a paper by Chuang, K-H., et al., "Physically unclonable function using CMOS breakdown position," International Reliability Physics Symposium (IRPS), 2017 IEEE International, IEEE, 2017, incorporated herein by reference. During this programming of per layer connection, all the relevant ridge selects could be disabled to reduce the risk of sneak paths.

The per layer connection technique presented in respect to Fig. 28 could be combined with the per layer connection technique presented in respect to Fig. 43 of PCT/US2016/052726 to various alternatives of mix and matched by an artisan in the art.

[000262] It will also be appreciated by persons of ordinary skill in the art that the invention is not limited to what has been particularly shown and described hereinabove. For example, the use of SiGe as the designated sacrificial layer or etch stop layer could be replaced by compatible material or combination of other material including additive materials to SiGe like carbon or various doping materials such as boron or other variations. And for example, drawings or illustrations may not show n or p wells for clarity in illustration. Further, any transferred layer or donor substrate or wafer preparation illustrated or discussed herein may include one or more undoped regions or layers of semiconductor material. Further, transferred layer or layers may have regions of STI or other transistor elements within it or on it when transferred. Rather, the scope of the invention includes combinations and sub-combinations of the various features described hereinabove as well as modifications and variations which would occur to such skilled persons upon reading the foregoing description. Thus, the invention is to be limited only by appended claims (if any).

We claim:

1. A 3D device, the device comprising:
a first stratum comprising a first bit-cell array, said first bitcell array comprises three independent first rows;
a second stratum comprising a second bit-cell array, said second bitcell array comprises three independent second rows,
wherein said second stratum overlays said first stratum; and
at least three vertical bitlines each connected to respective three horizontal first bitlines and three horizontal second bitlines,
wherein said three horizontal first bitlines comprise control of said first bit-cell array,
wherein said three horizontal second bitlines comprise control of said second bit-cell array, and
wherein each of said three vertical bitlines could be used to control a different one of said three independent first rows, or control a different one of said three independent second rows.

2. The 3D device according to claim 1,
wherein said three horizontal first bitlines are spaced apart by less than 200 nm.

3. The 3D device according to claim 1,
wherein said three horizontal first bitlines each comprise a pass transistor to control access to the respective rows of said first bit-cell array.

4. The 3D device according to claim 1,
wherein said first stratum is bonded to said second stratum by fusion bonding or by hybrid bonding.

5. The 3D device according to claim 1, further comprising:
a third stratum overlaying said second stratum,
wherein said third stratum comprises at least three bitline decoders, one for each of said three vertical bitlines.

6. A 3D device, the device comprising:
a first stratum comprising a first bit-cell array, said first bitcell array comprises three independent first rows;
a second stratum comprising a second bit-cell array, said second bitcell array comprises three independent second rows,
wherein said second stratum overlays said first stratum; and
at least three vertical wordlines each connected to respective three horizontal first wordlines and three horizontal second wordlines,
wherein said three horizontal first wordlines comprise control of said first bit-cell array,
wherein said three horizontal second wordlines comprise control of said second bit-cell array, and
wherein each of said three vertical wordlines could be used to control a different one of said three independent first rows, or control a different one of said three independent second rows.

7. A 3D device, the device comprising:
a first stratum of array of units, said array of units comprising a first unit comprising a first bit-cell array and a second unit comprising a second bit-cell array,
wherein said first bit-cell array comprises a first and a second row of bit-cells controlled by a first and a second bit-line respectively, and
wherein said second bit-cell array comprises a third and a fourth row of bit-cells controlled by a third and a fourth bit-line respectively; and

a first transistor controlling the connectivity of said first bit-line to said third bit-line; and

a second transistor controlling the connectivity of said second bit-line to said fourth bit-line.

8. The 3D device according to claim 7,

wherein said first bit-line and said second bit-line are spaced apart by less than 200 nm.

9. The 3D device according to claim 7, further comprising:

a third transistor controlling the connectivity of said first bit-line to said third bit-line.

10. A 3D device, the device comprising:

a first stratum of array of units, said array of units comprising a first unit comprising a first bit-cell array and a second unit comprising a second bit-cell array,

wherein said first bit-cell array comprises a first and a second row of bit-cells controlled by a first and a second word-line respectively, and

wherein said second bit-cell array comprises a third and a fourth row of bit-cells controlled by a third and a fourth word-line respectively; and

a first transistor controlling the connectivity of said first word-line to said third word-line; and

a second transistor controlling the connectivity of said second word-line to said fourth word-line.

11. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrays, overlaid by a third stratum,

wherein said third stratum comprises a plurality of bitline decoders to control said first bit-cell array and said second bit-cell array.

12. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrasy, overlaid by a third stratum,

wherein said third stratum comprises a plurality of wordline decoders to control said first bit-cell array and said second bit-cell array.

13. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrays, overlaid by a third stratum,

wherein said third stratum comprises a plurality of n type and p type transistors, and

wherein said first stratum and second stratum comprise only n type transistors.

14. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrays, overlaid by a third stratum,

wherein said third stratum comprises a first stratum select contact and a second stratum select contact and a bit-line contact,

wherein said first stratum comprises a first bit-line controlling at least one of said first bit-cell arrays, and

wherein said second stratum comprises a second bit-line, overlaying said first bit-line, controlling at least one of said second bit-cell arrays, and

wherein said first stratum select contact could be used to connect said bit-line contact to said first bit-line, and said second stratum select contact could be used to connect said bit-line contact to said second bit-line.

.

15. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrays, overlaid by a third stratum,

wherein said third stratum comprises a first stratum select contact and a second stratum select contact and a word-line contact,

wherein said first stratum comprises a first word-line controlling at least one of said first bit-cell arrays, and

wherein said second stratum comprises a second word-line, overlaying said first word-line, controlling at least one of said second bit-cell arrays, and

wherein said first stratum select contact could be used to connect said word-line contact to said first word-line, and said second stratum select contact could be used to connect said word-line contact to said second word-line.

16. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrays, overlaid by a third stratum of third bit-cell arrays, overlaid by a fourth stratum,

wherein said first stratum comprises a first word-line controlling at least one of said first bit-cell arrays, and

wherein said second stratum comprises a second word-line, overlaying said first word-line, controlling at least one of said second bit-cell arrays,

wherein said third stratum comprises a third word-line, overlaying said second word-line, controlling at least one of said second bit-cell arrays, and

wherein said fourth stratum comprises control circuit to enable use of said third word-line as an alternative to said first word-line or said second word-line.

17. A 3D device, the device comprising:

a first stratum of first bit-cell arrays, overlaid by a second stratum of second bit-cell arrays, overlaid by a third stratum of third bit-cell arrays overlaid by a fourth stratum,

wherein said first stratum comprises a first bit-line controlling at least one of said first bit-cell arrays, and

wherein said second stratum comprises a second bit-line, overlaying said first bit-line, controlling at least one of said second bit-cell arrays,

wherein said third stratum comprises a third bit-line, overlaying said second bit-line, controlling at least one of said second bit-cell arrays, and

wherein said fourth stratum comprises control circuits to enable use of said third bit-line as an alternative to said first bit-line or said second bit-line.

18. A 3D device, the device comprising:

a first stratum comprising a first bit-cell array, said first bitcell array comprises three independent first rows and three horizontal first bitlines,

wherein said three horizontal first bitlines are BLsub1-1, BLsub1-2, and BLsub1-3;

a second stratum comprising a second bit-cell array, said second bitcell array comprises three independent second rows and three horizontal second bitlines,

wherein said second stratum overlays said first stratum, and

wherein said three horizontal second bitlines are BLsub2-1, BLsub2-2, and BLsub2-3;

at least three vertical bitlines, vBLsub1, vBLsub2, and vBLsub3,

wherein vBLsub1 is selectively connectable to BLsub1-1 or BLsub 2-1,

wherein vBLsub2 is selectively connectable to BLsub1-2 or BLsub 2-2,

wherein vBLsub1 is selectively connectable to BLsub1-3 or BLsub 2-3,

wherein said three horizontal first bitlines comprise control of said first bit-cell array,

wherein said three horizontal second bitlines comprise control of said second bit-cell array, and

wherein each of said three vertical bitlines could be used to control a different one of said three independent first rows, or control a different one of said three independent second rows.

Fig. 1B

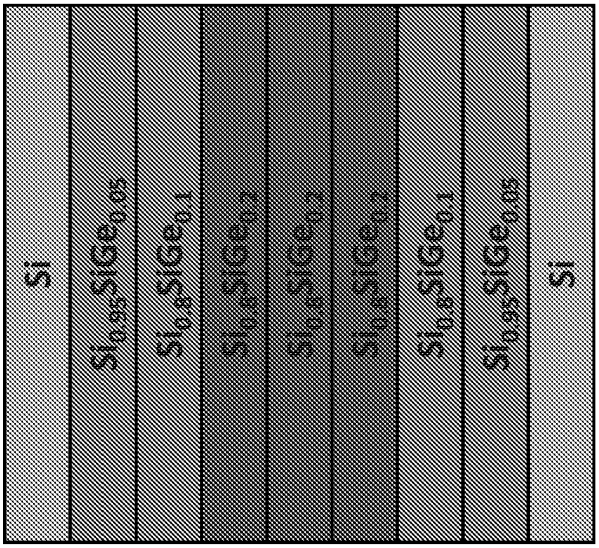
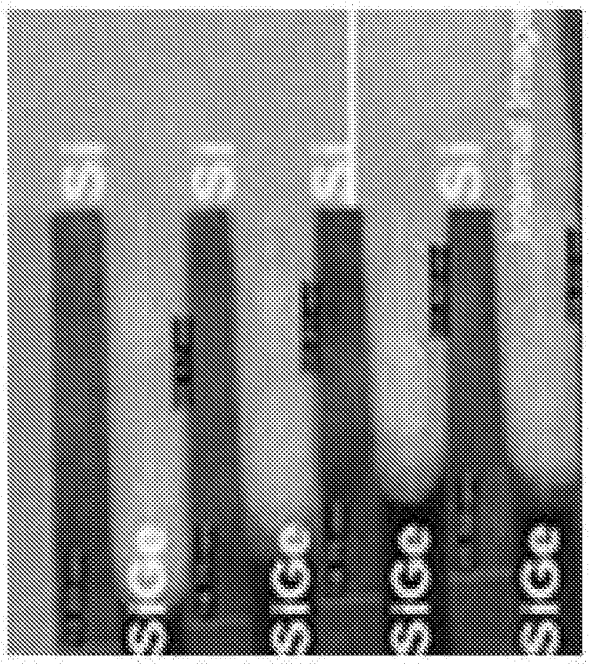

Si
$Si_{0.95}SiGe_{0.05}$
$Si_{0.8}SiGe_{0.1}$
$Si_{0.8}SiGe_{0.2}$
$Si_{0.8}SiGe_{0.2}$
$Si_{0.8}SiGe_{0.2}$
$Si_{0.8}SiGe_{0.1}$
$Si_{0.95}SiGe_{0.05}$
Si

Fig. 1A


SiGe
SiGe
SiGe
SiGe

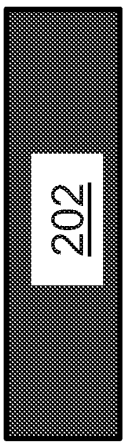


Fig. 2A

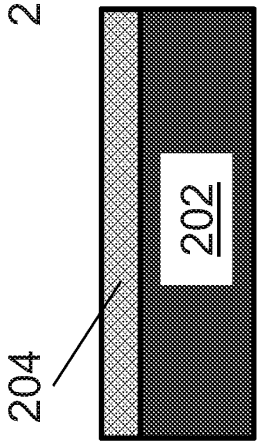


Fig. 2B

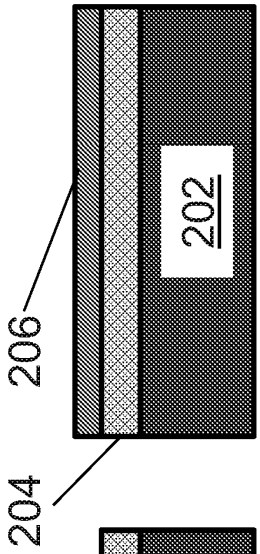


Fig. 2C

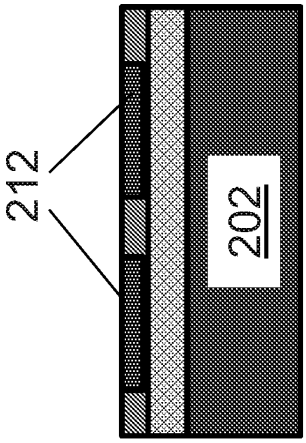


Fig. 2D

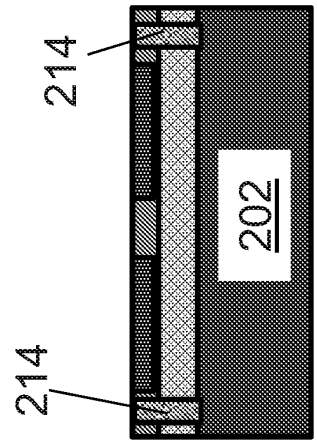


Fig. 2E

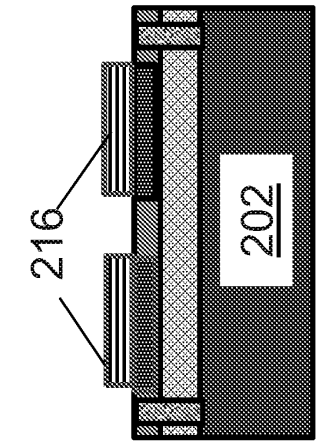


Fig. 2F

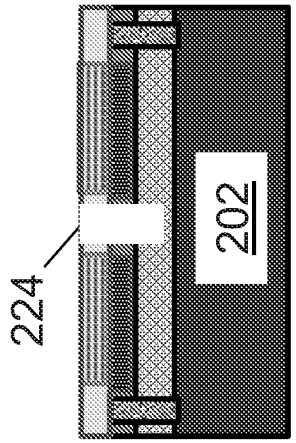


Fig. 2H

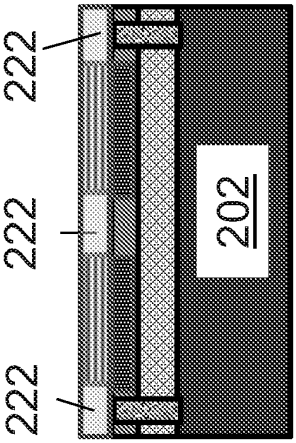


Fig. 2G

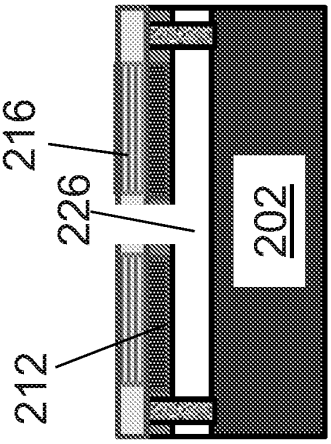


Fig. 2I

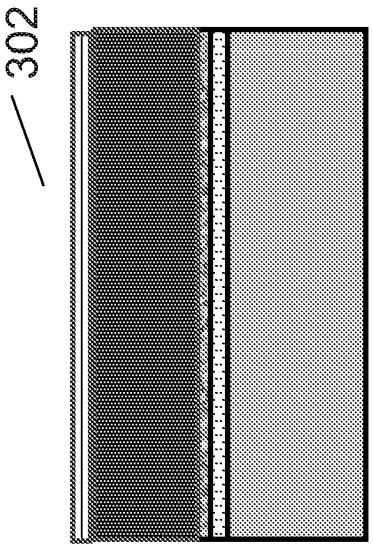


Fig. 3B

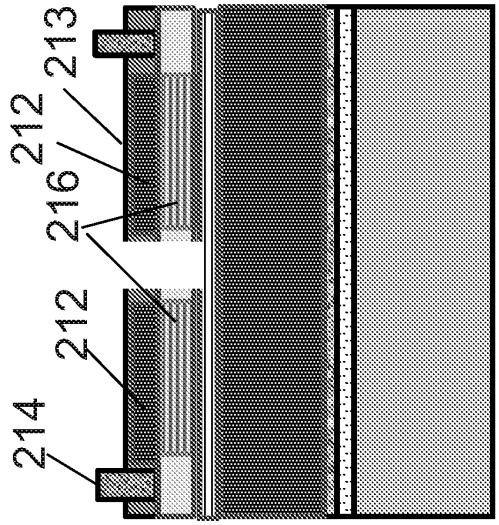


Fig. 3D

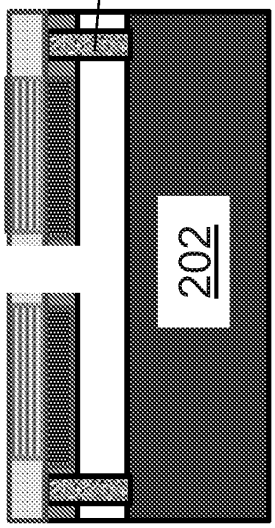


Fig. 3A

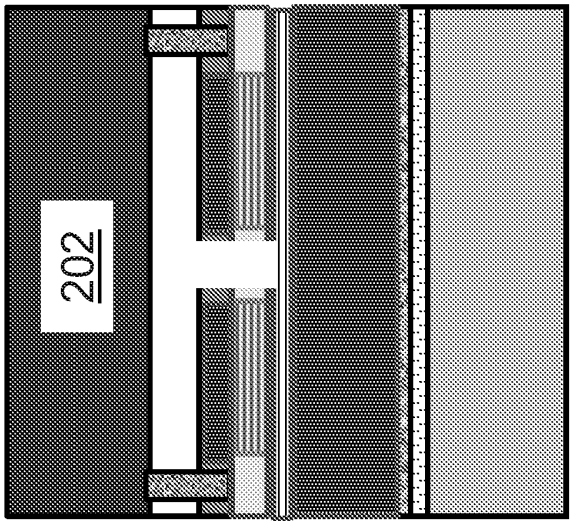


Fig. 3C

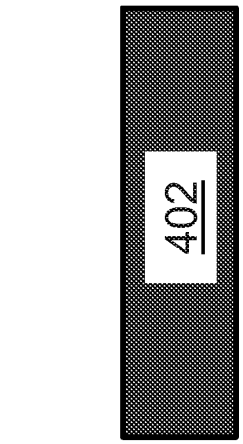


Fig. 4A

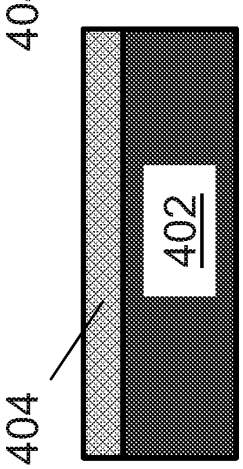


Fig. 4B

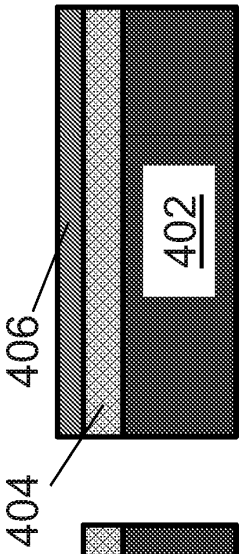


Fig. 4C

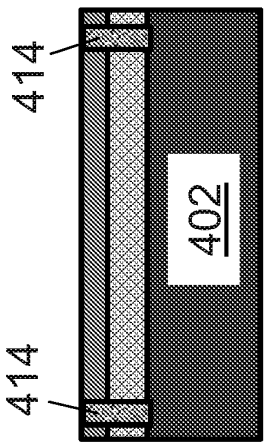


Fig. 4D

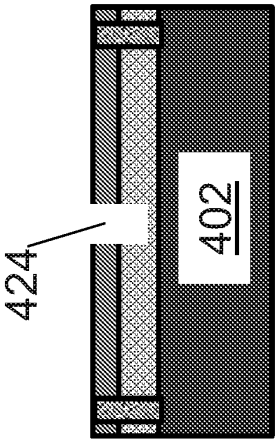


Fig. 4E

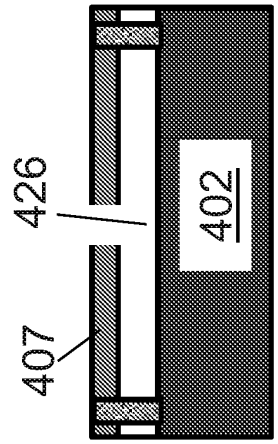


Fig. 4F

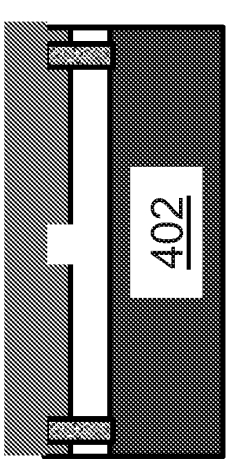


Fig. 4G

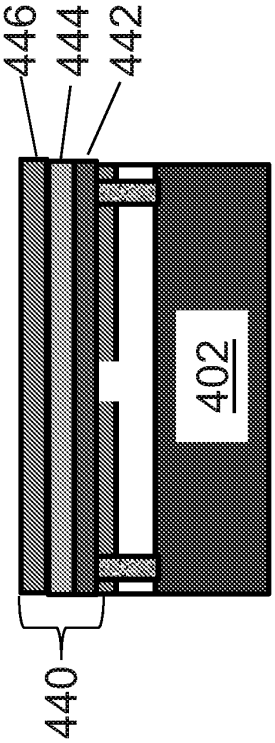


Fig. 4H

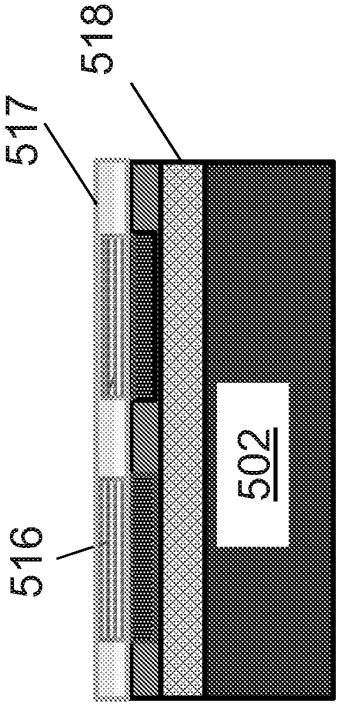


Fig. 5A

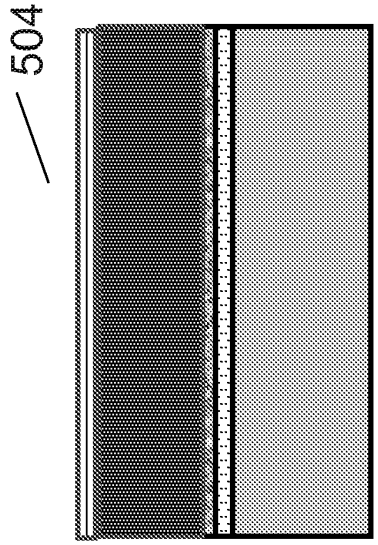


Fig. 5B

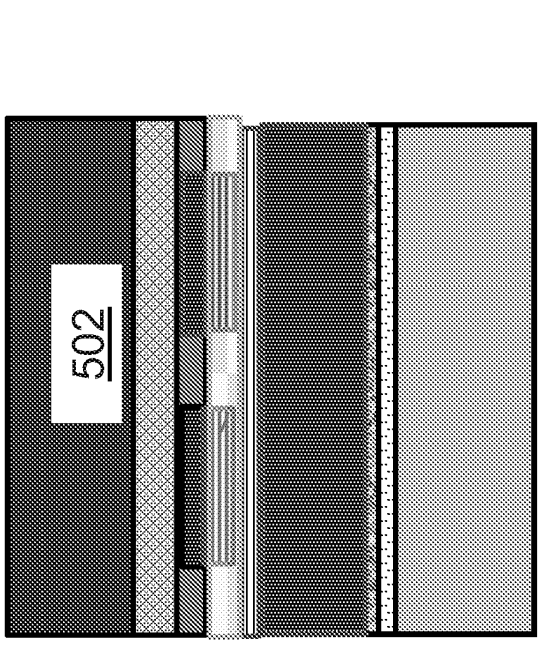


Fig. 5C

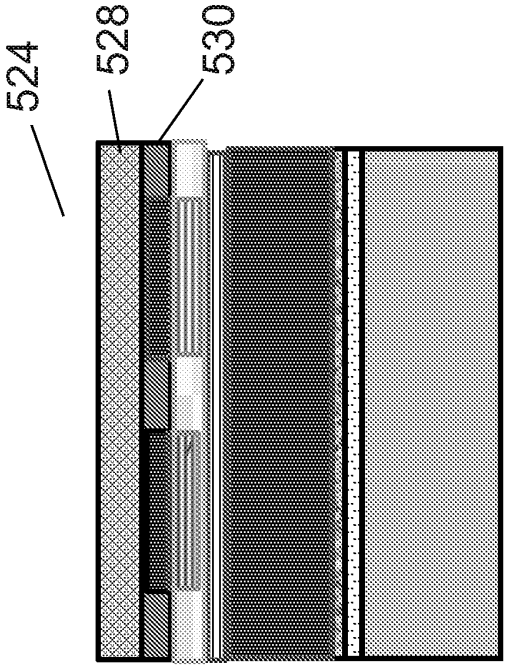


Fig. 5D

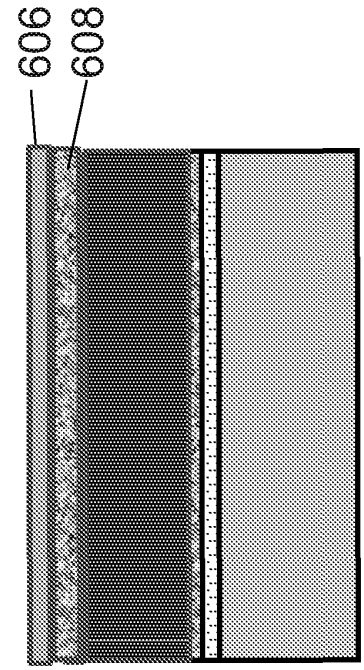


Fig. 6B

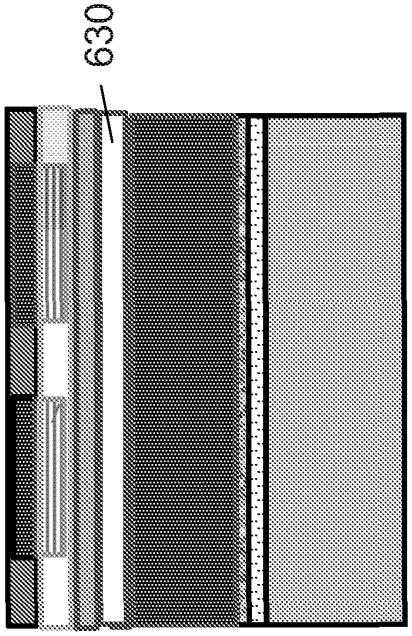


Fig. 6D

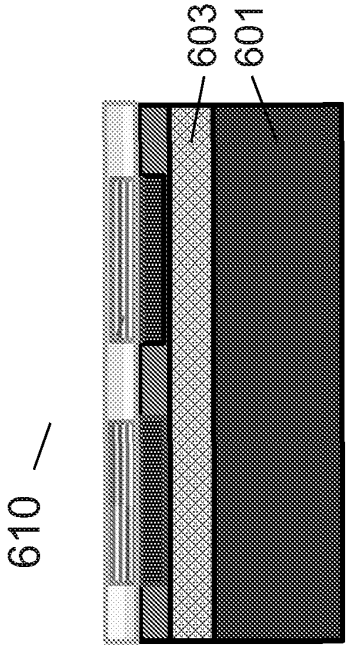


Fig. 6A

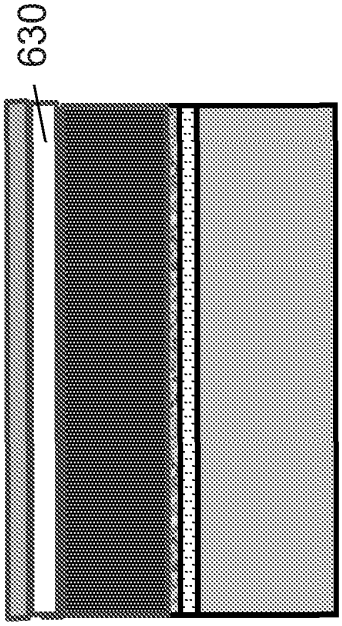


Fig. 6C

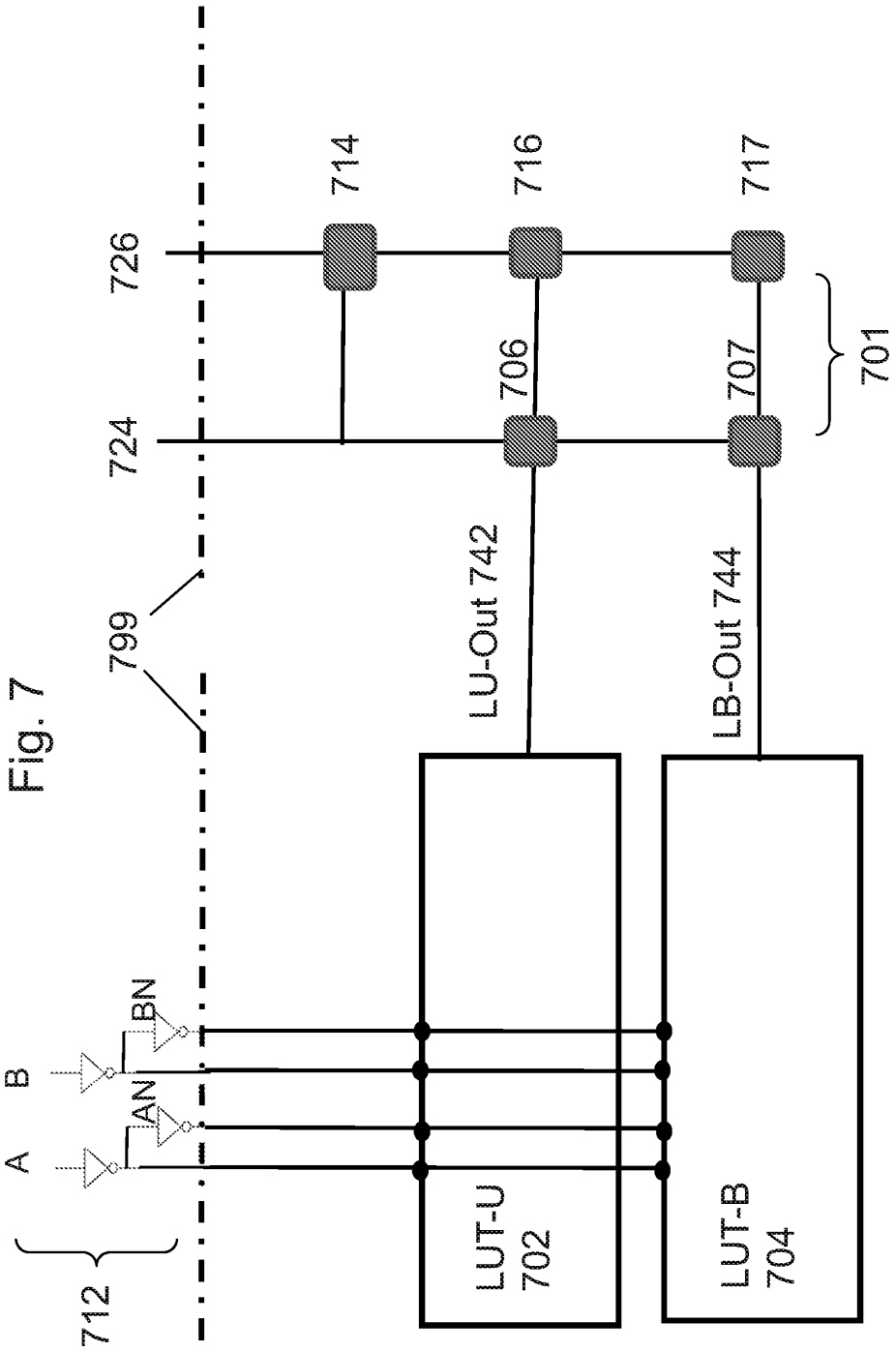

712
A
B
AN
BN
799
726
724
714
716
717
706
707
701
LU-Out 742
LB-Out 744
LUT-U
702
LUT-B
704

Fig. 7

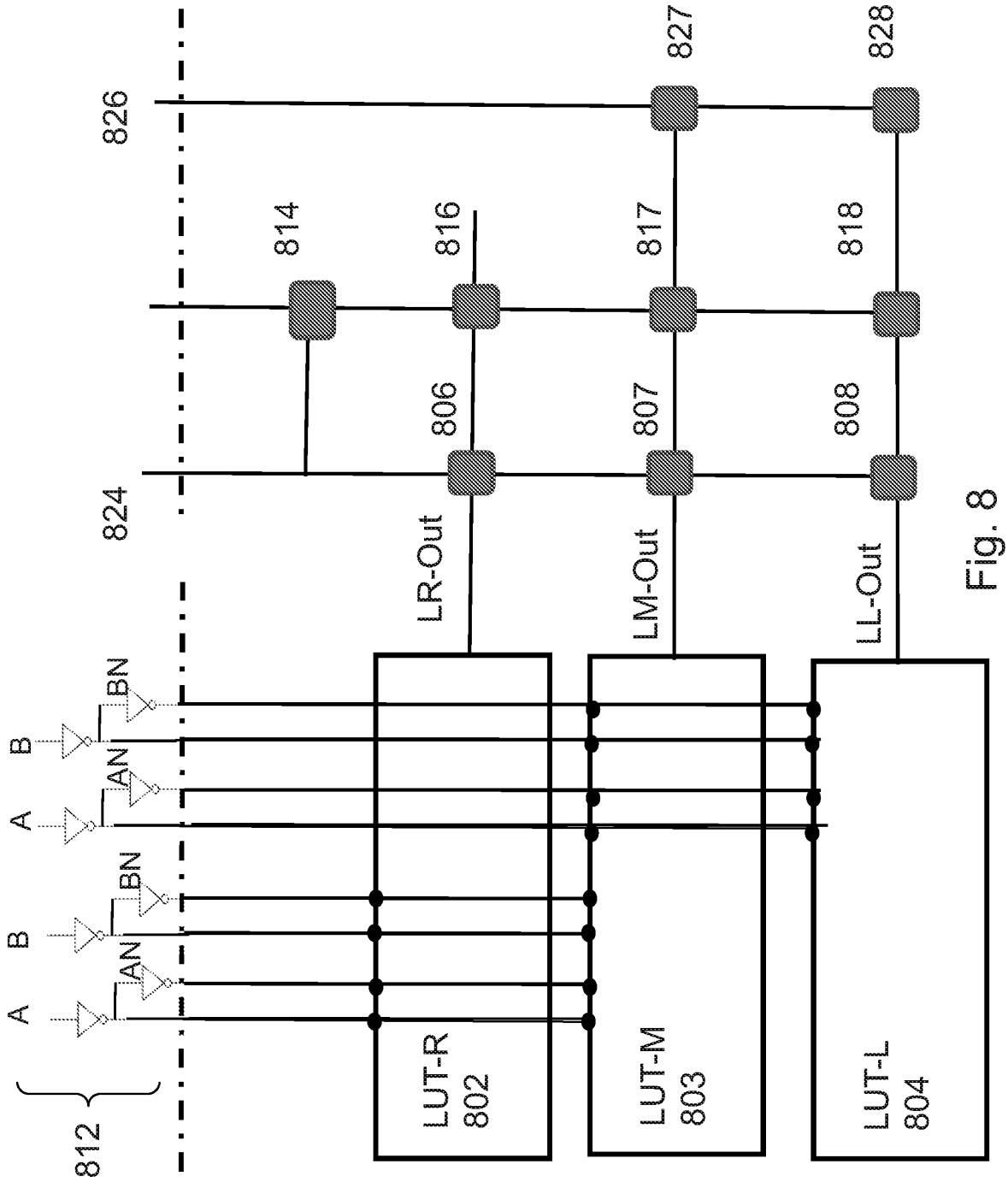
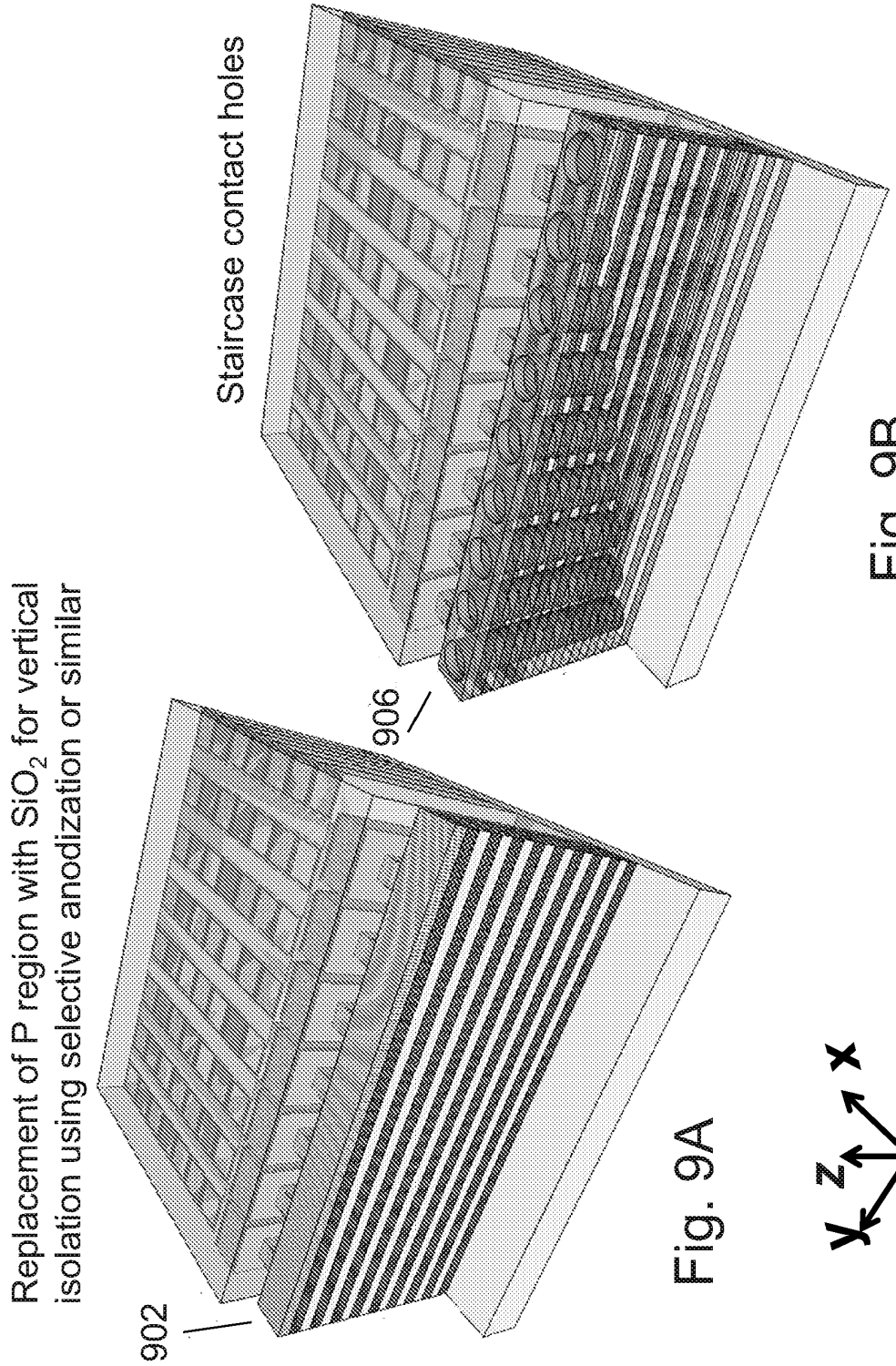

826
827
828
814
816
817
818
806
807
808
824
LR-Out
LM-Out
LL-Out
BN
B
AN
A
BN
B
AN
A
812
LUT-R
802
LUT-M
803
LUT-L
804

Fig. 8


Replacement of P region with SiO_2 for vertical isolation using selective anodization or similar
902
906
Staircase contact holes
Fig. 9A
Fig. 9B
x
z
y

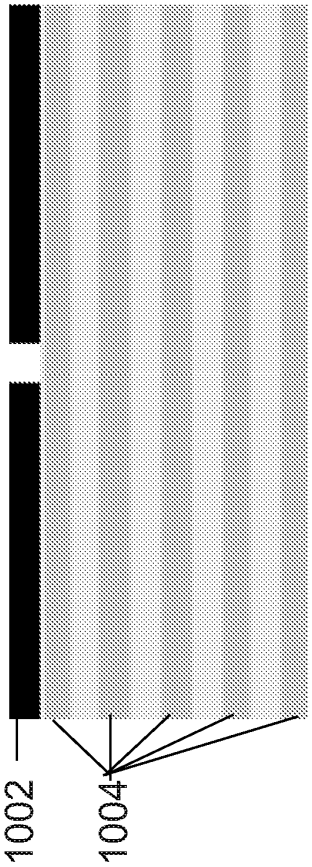


Fig. 10A

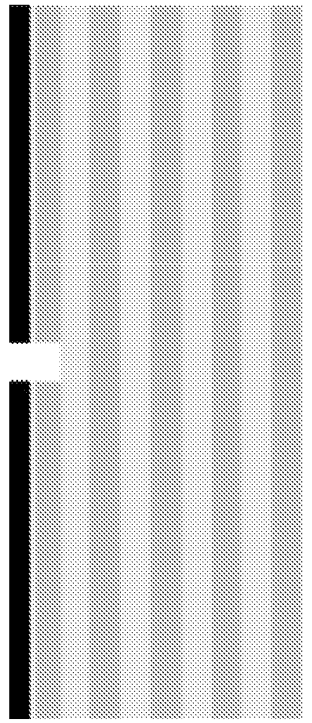


Fig. 10B

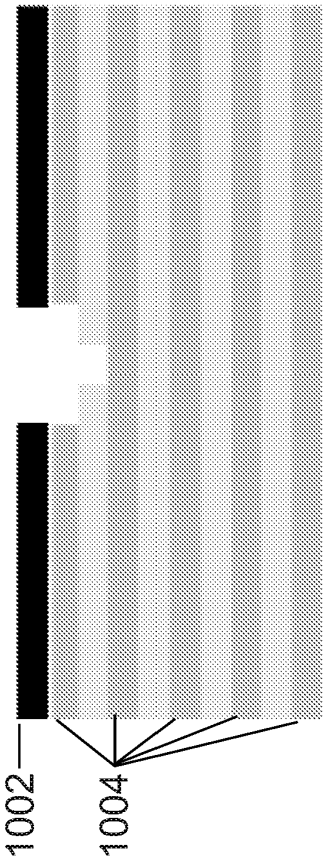


Fig. 10C

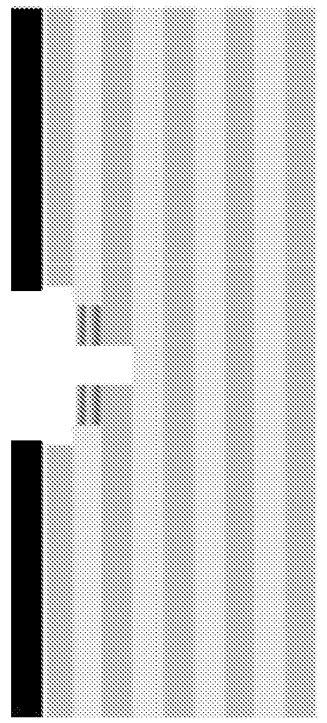


Fig. 10D

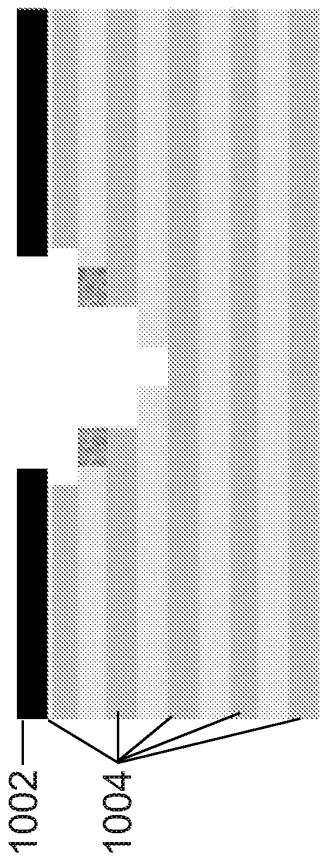


Fig. 10E

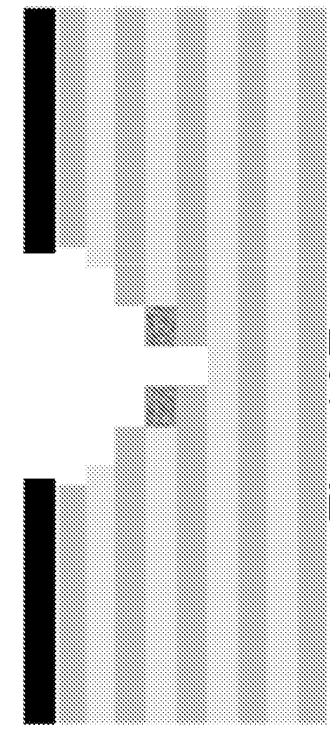


Fig. 10F

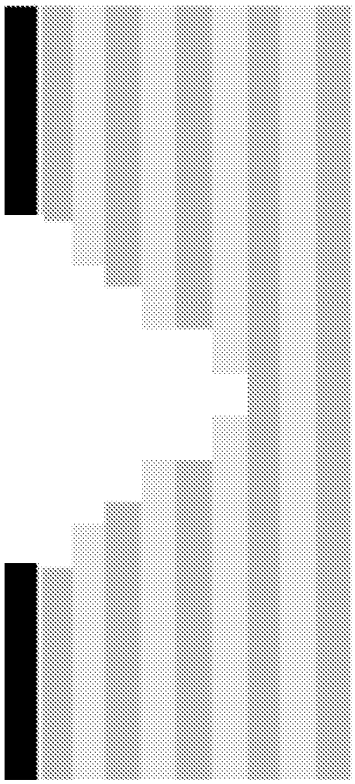


Fig. 10G

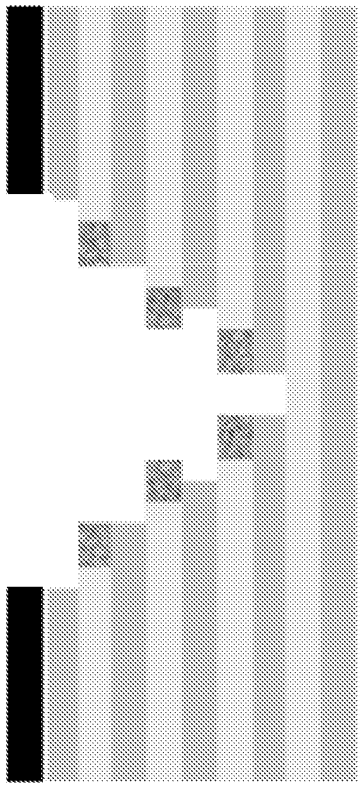


Fig. 10H

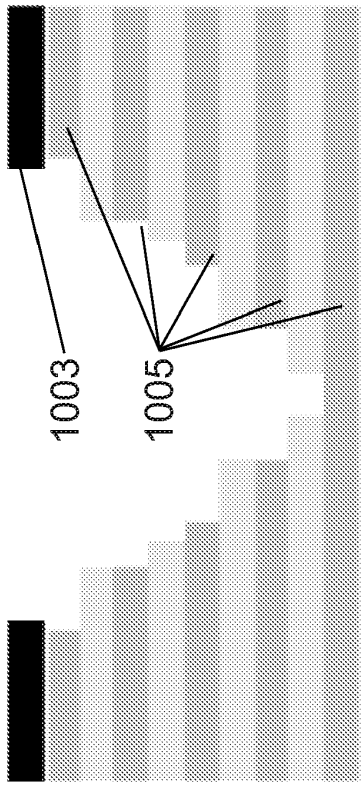


Fig. 10I

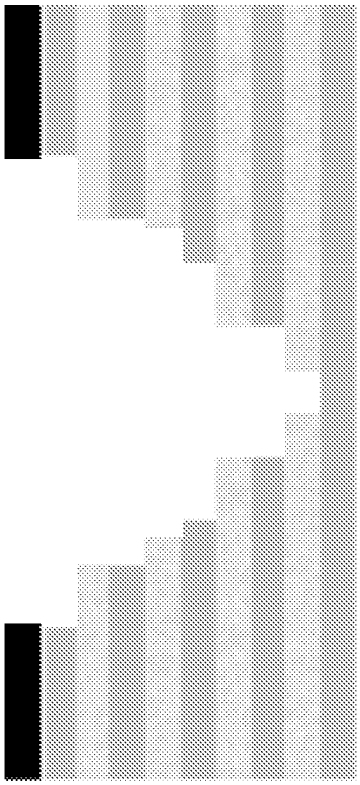


Fig. 10J

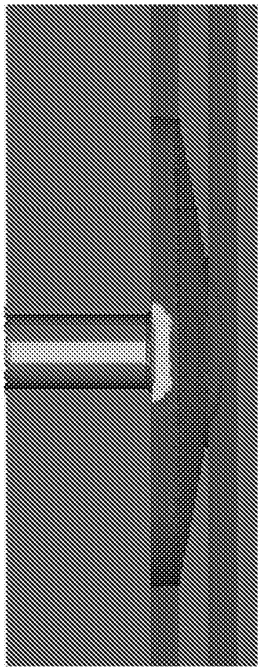


Fig. 10L

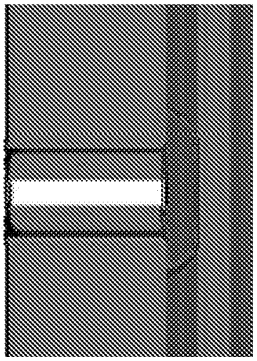


Fig. 10K

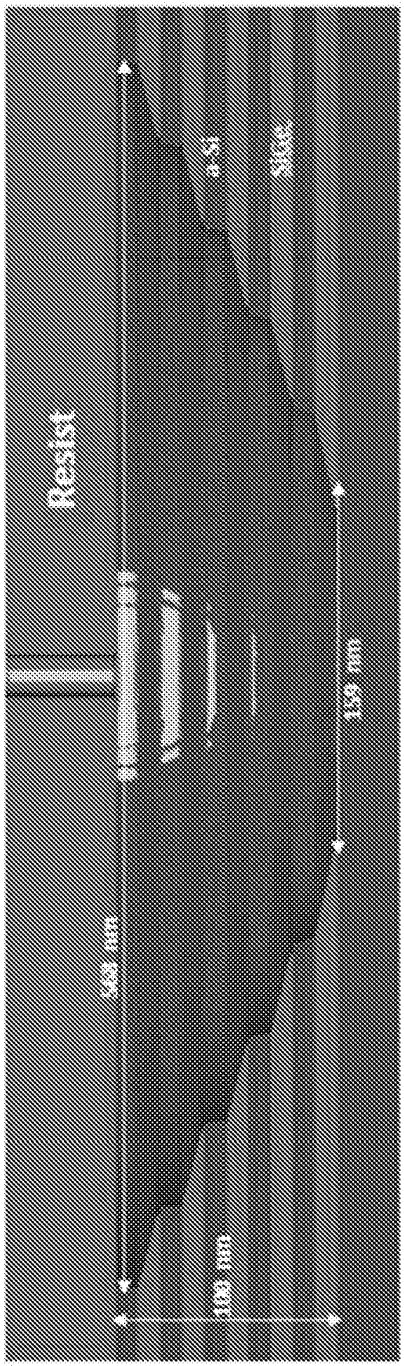


Fig. 10M

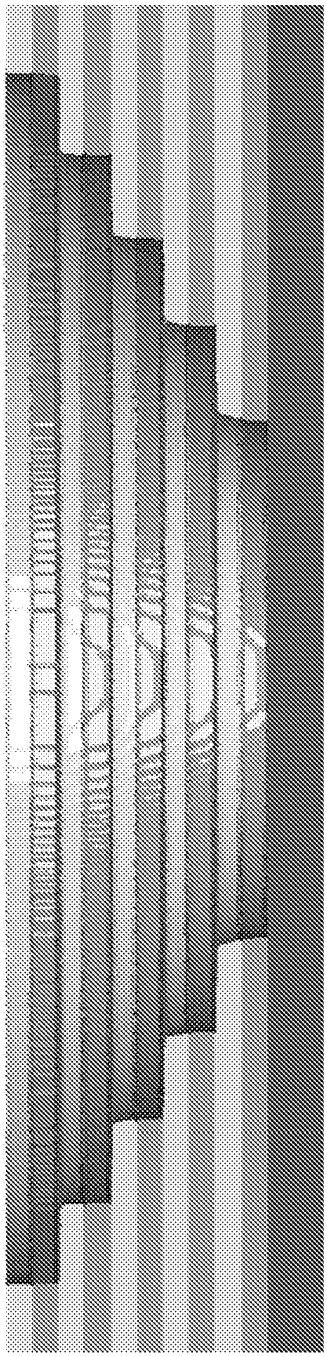


Fig. 10N

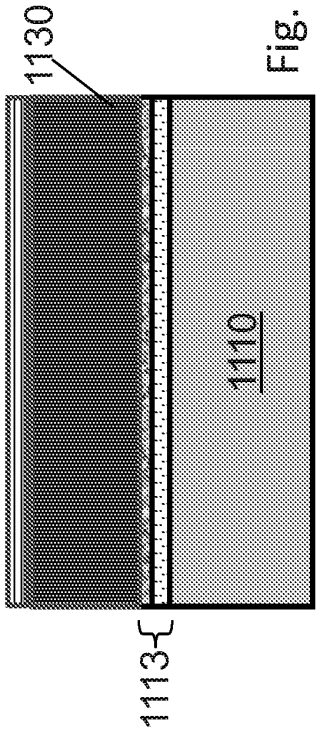


Fig. 11B

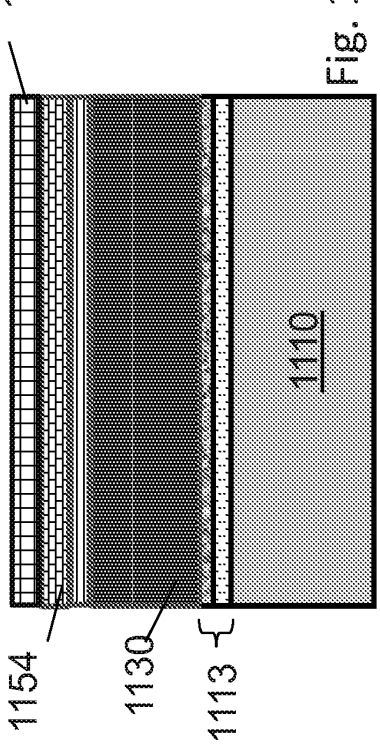


Fig. 11D

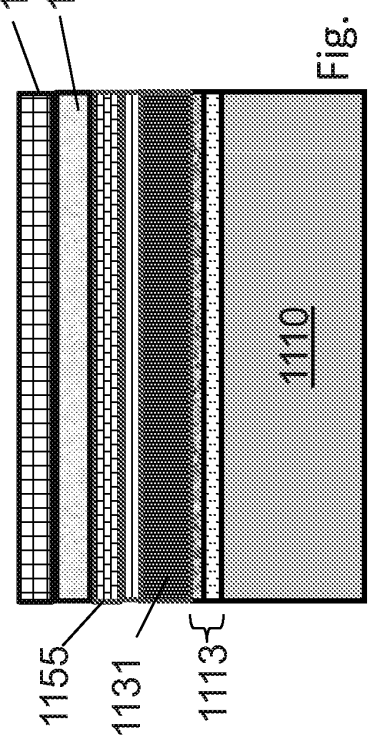


Fig. 11E

Side Cut-Views Y-Z Plane

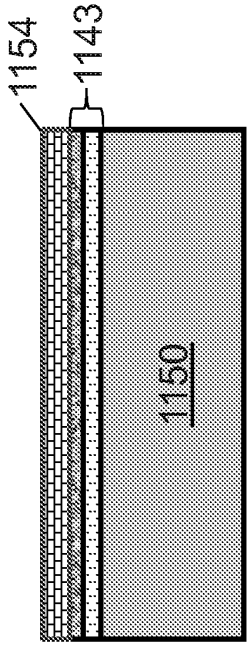


Fig. 11A

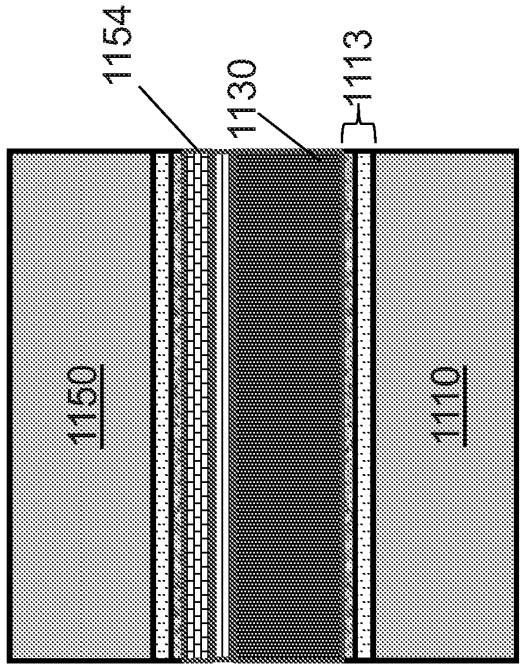


Fig. 11C

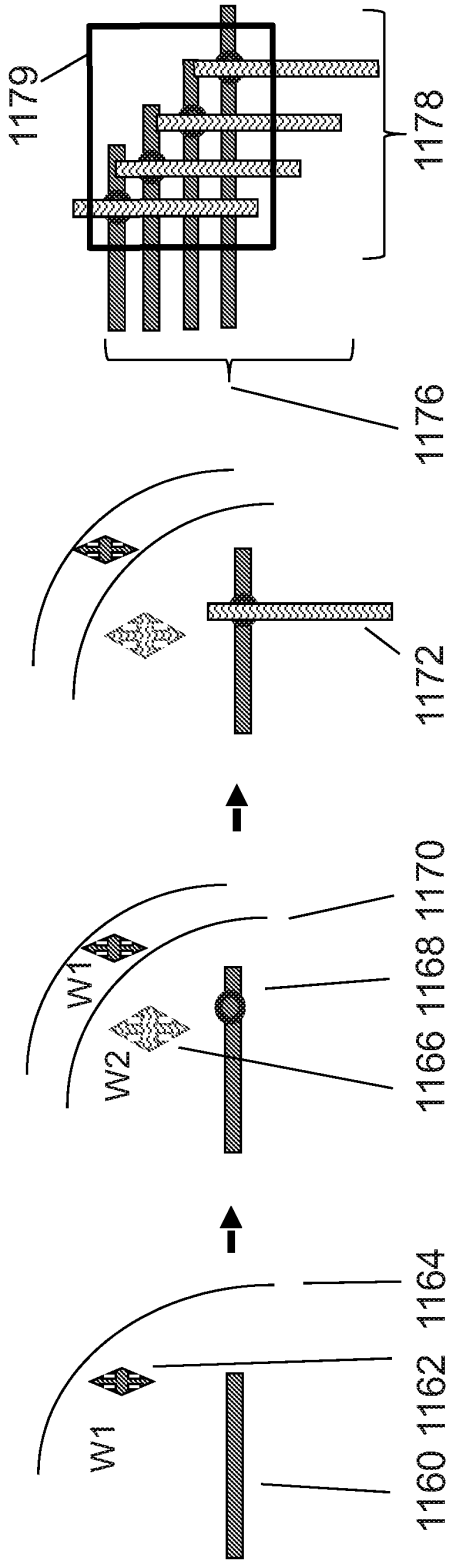


Fig. 11G

Fig. 11F

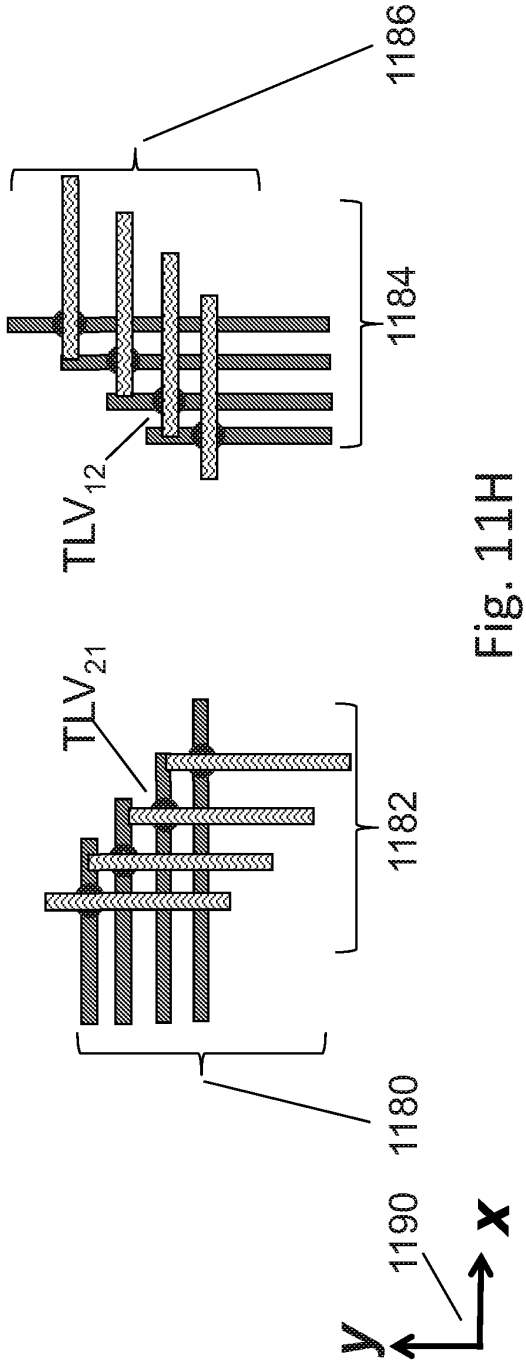


Fig. 11H

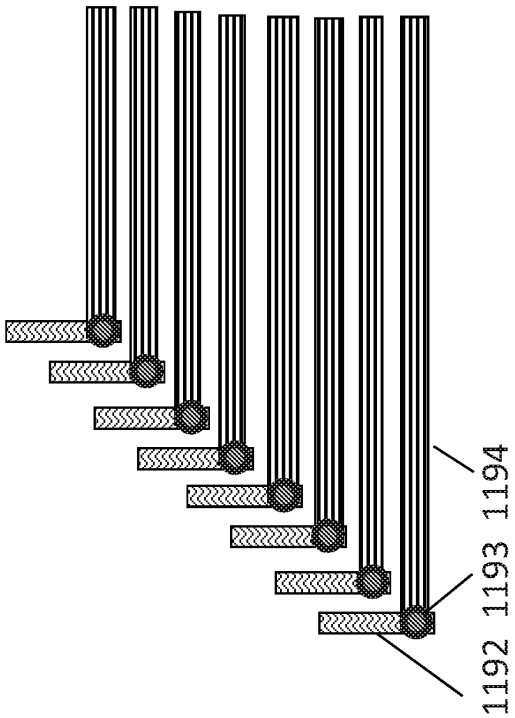


Fig. 11J

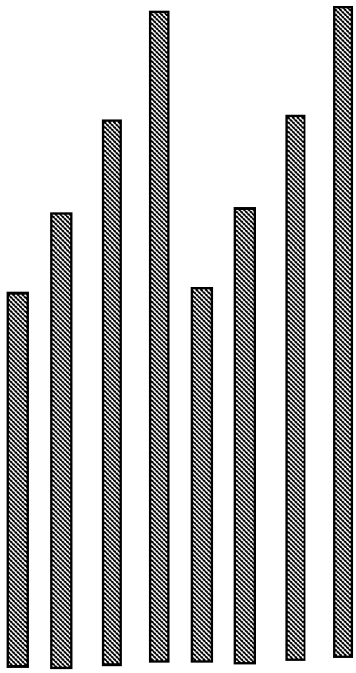


Fig. 11I

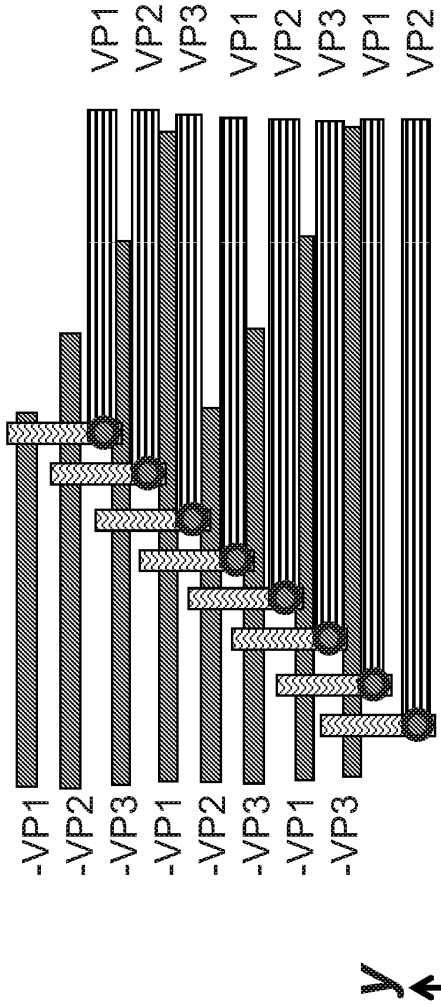


Fig. 11K

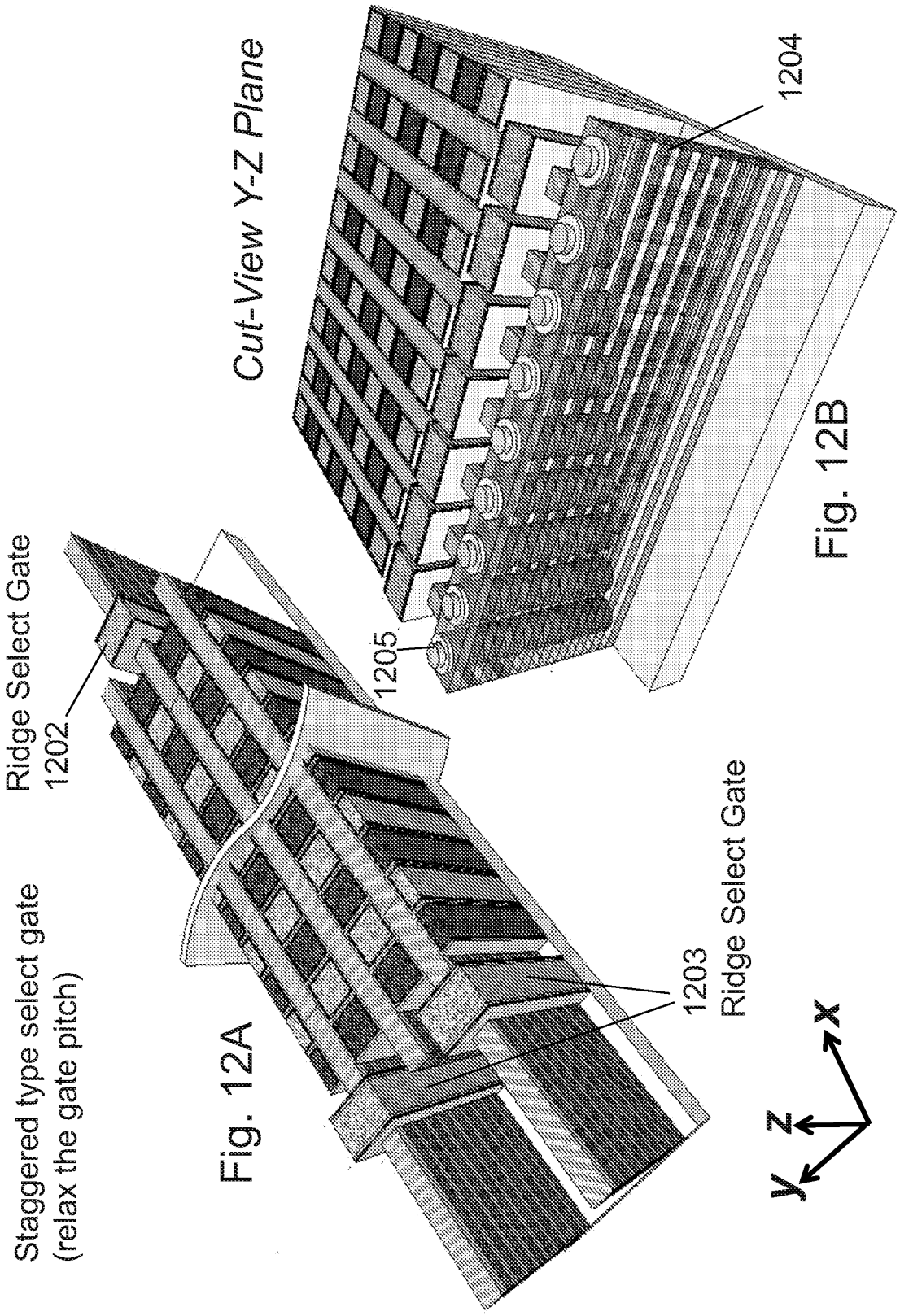


Fig. 12A

Fig. 12B

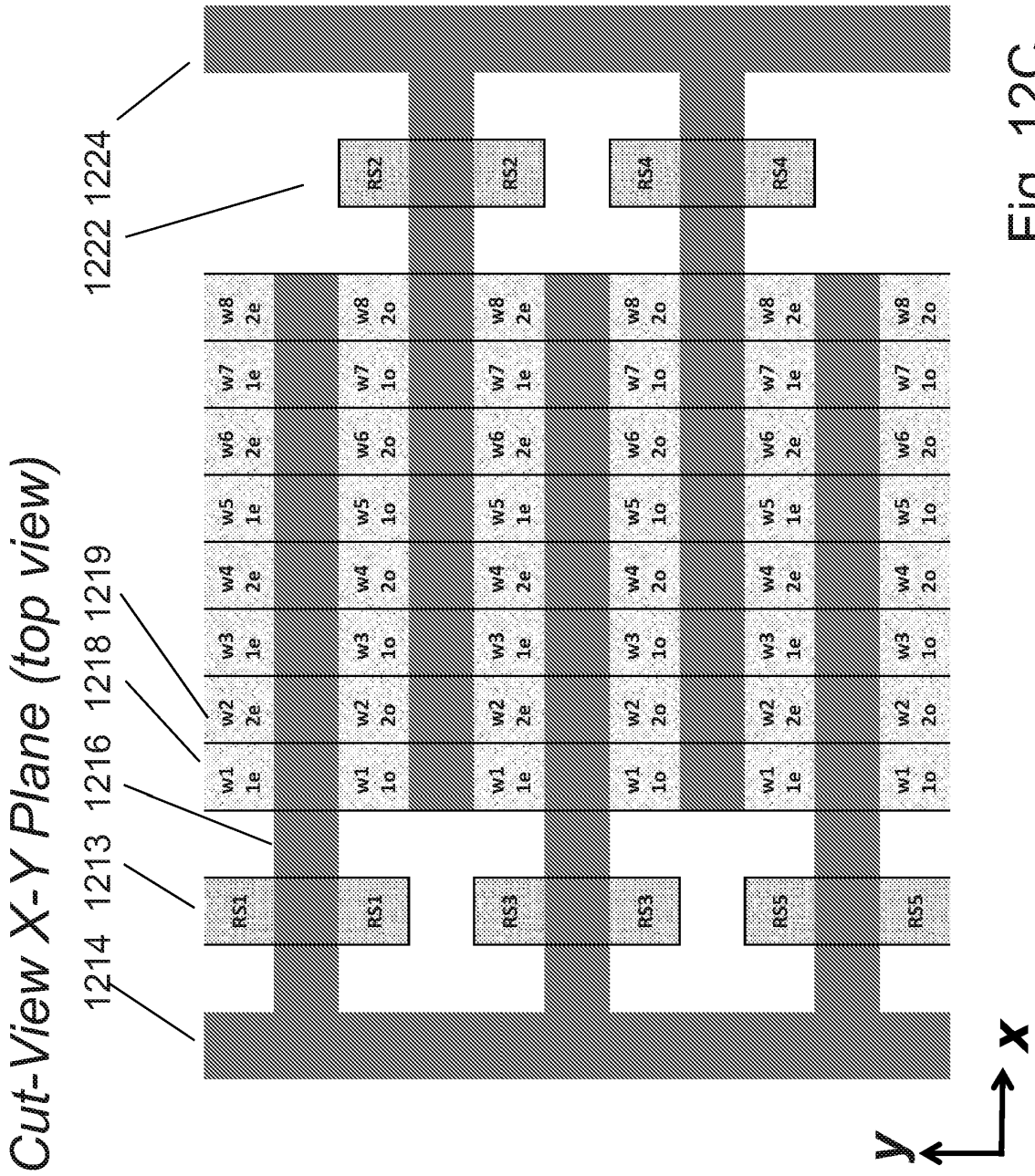

Cut-View X-Y Plane (top view)
1214
1213
1216
1218
1219
1222
1224
RS1
RS1
RS3
RS3
RS5
RS5
RS2
RS2
RS4
RS4
w1 1e
w1 1o
w2 2e
w2 2o
w3 1e
w3 1o
w4 2e
w4 2o
w5 1e
w5 1o
w6 2e
w6 2o
w7 1e
w7 1o
w8 2e
w8 2o
y
x

Fig. 12C

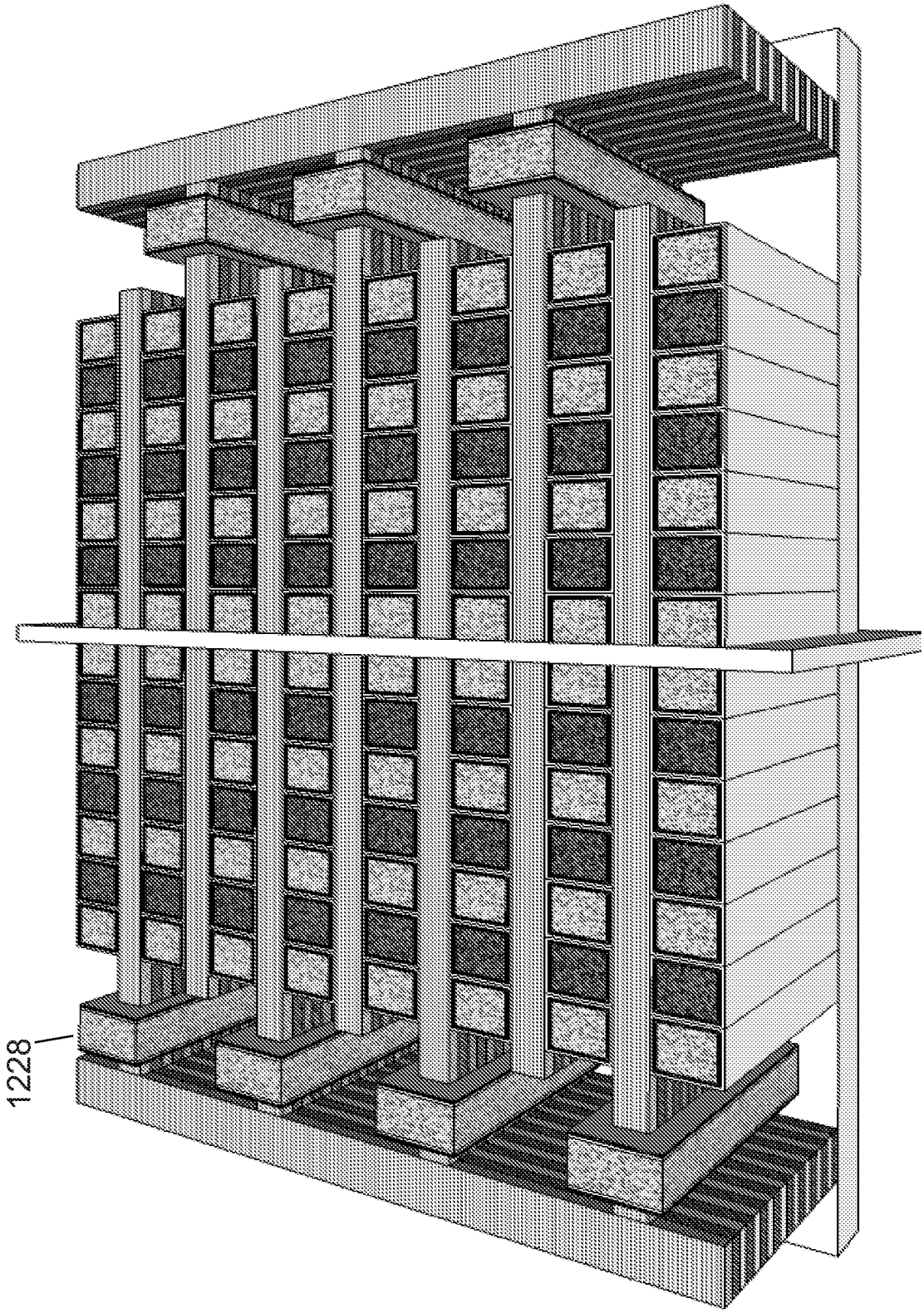


Fig. 12D

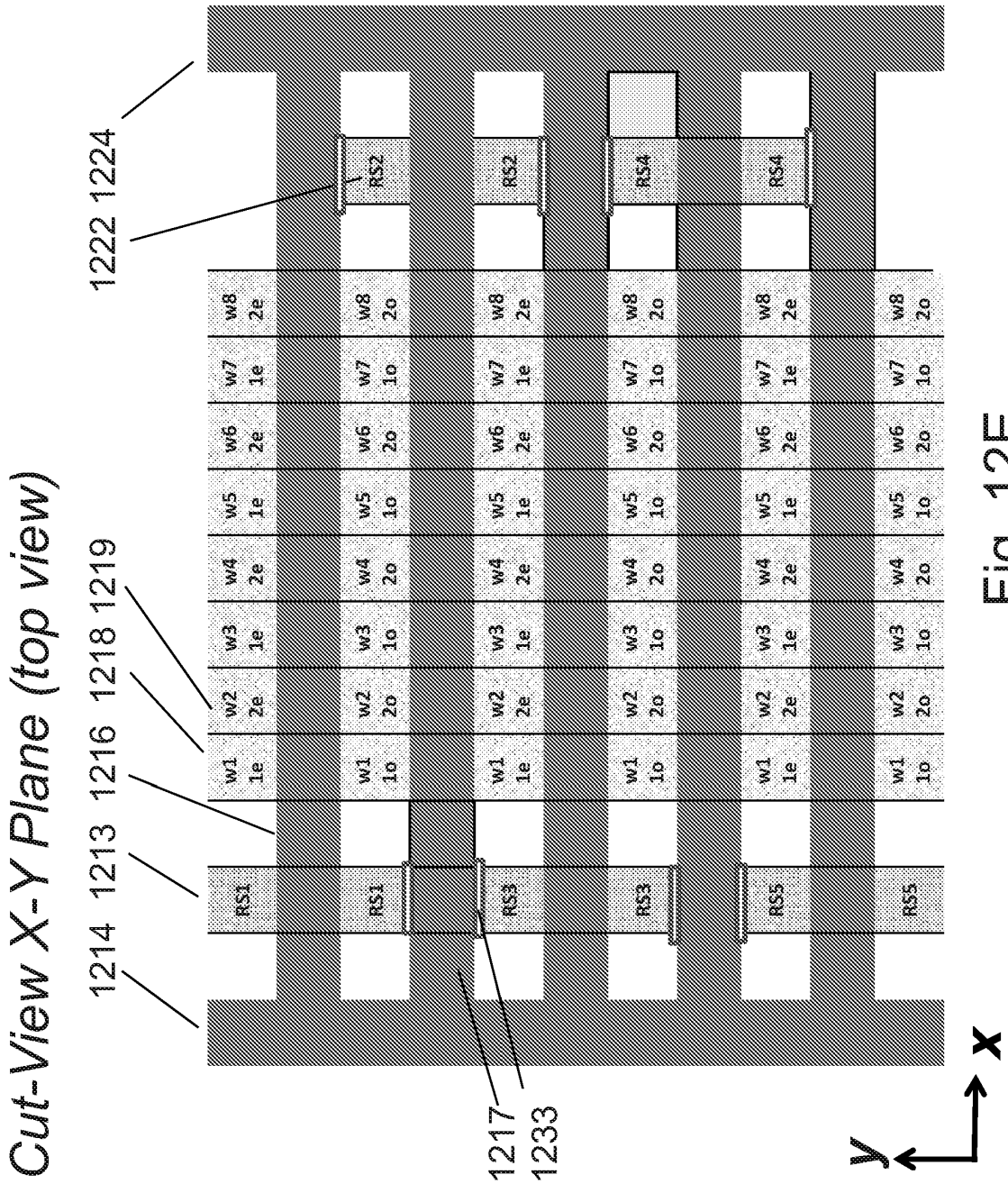

Cut-View X-Y Plane (top view)
1214
1213
1216
1218
1219
1222
1224
1217
1233
RS1
RS1
RS3
RS3
RS5
RS5
RS2
RS2
RS4
RS4
w1 1e
w1 1o
w2 2e
w2 2o
w3 1e
w3 1o
w4 2e
w4 2o
w5 1e
w5 1o
w6 2e
w6 2o
w7 1e
w7 1o
w8 2e
w8 2o
x
y

Fig. 12E

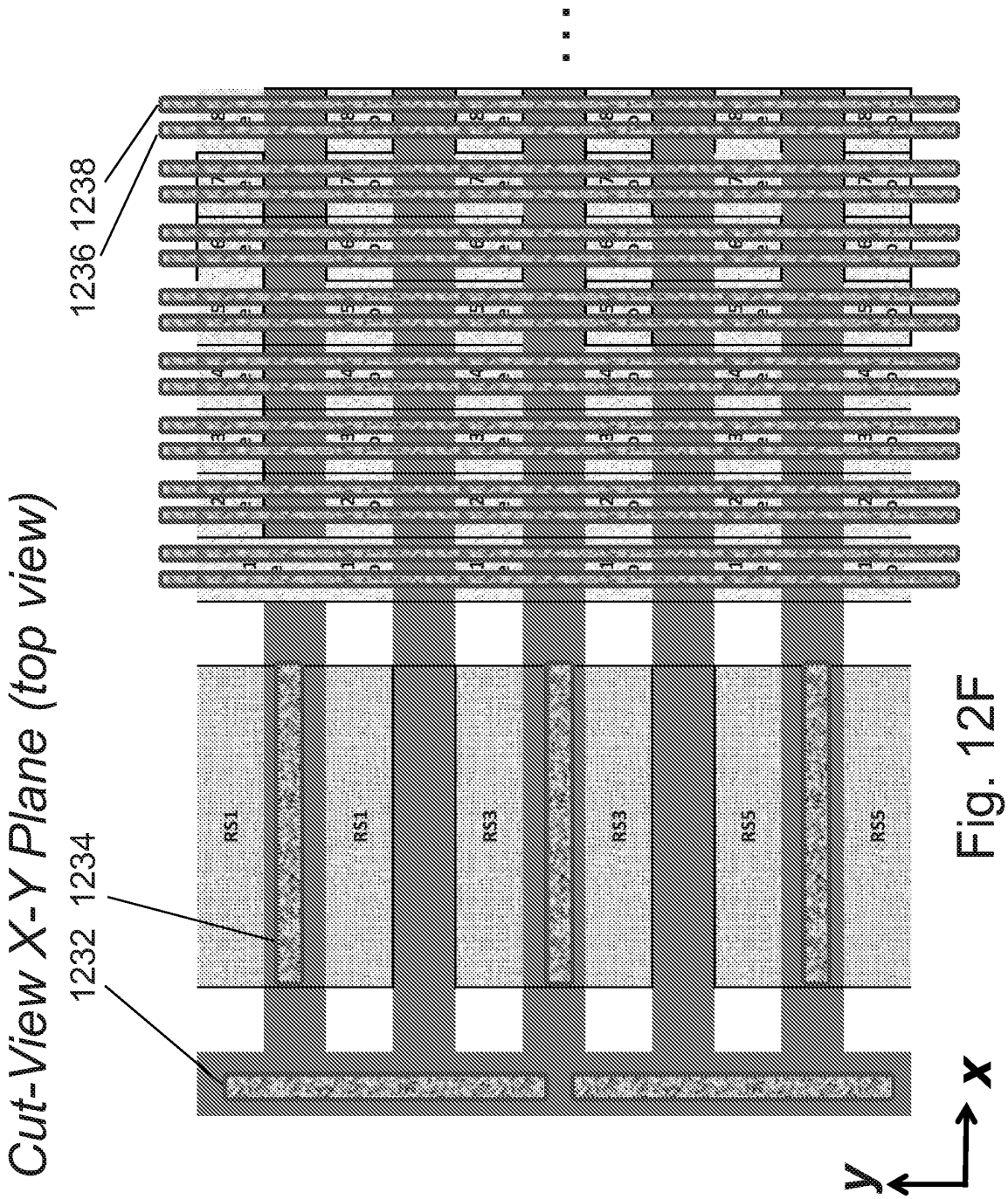

Cut-View X-Y Plane (top view)
1232
1234
1236
1238
RS1
RS1
RS3
RS3
RS5
RS5
x
y

Fig. 12F

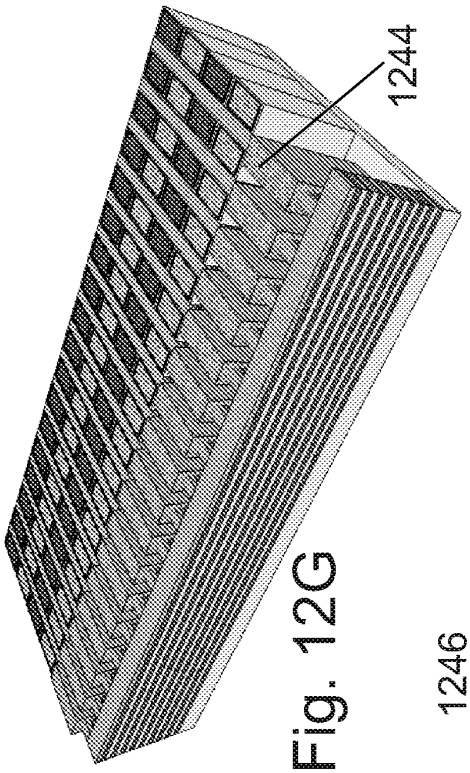


Fig. 12G

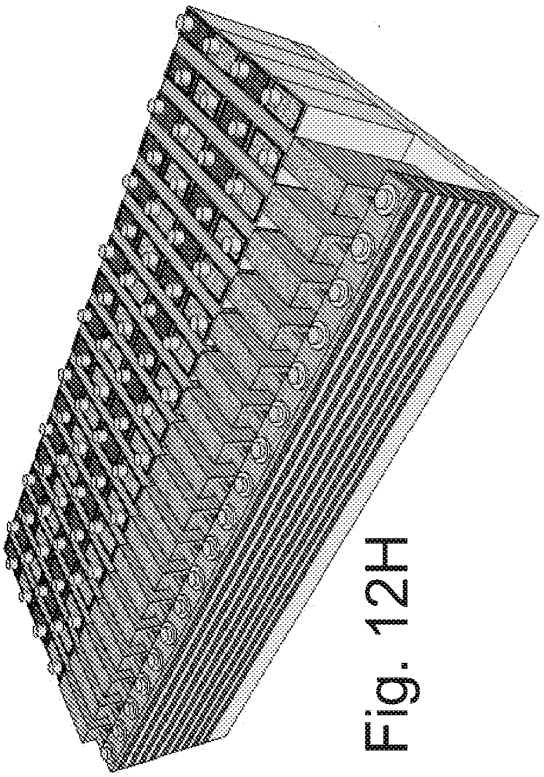


Fig. 12H

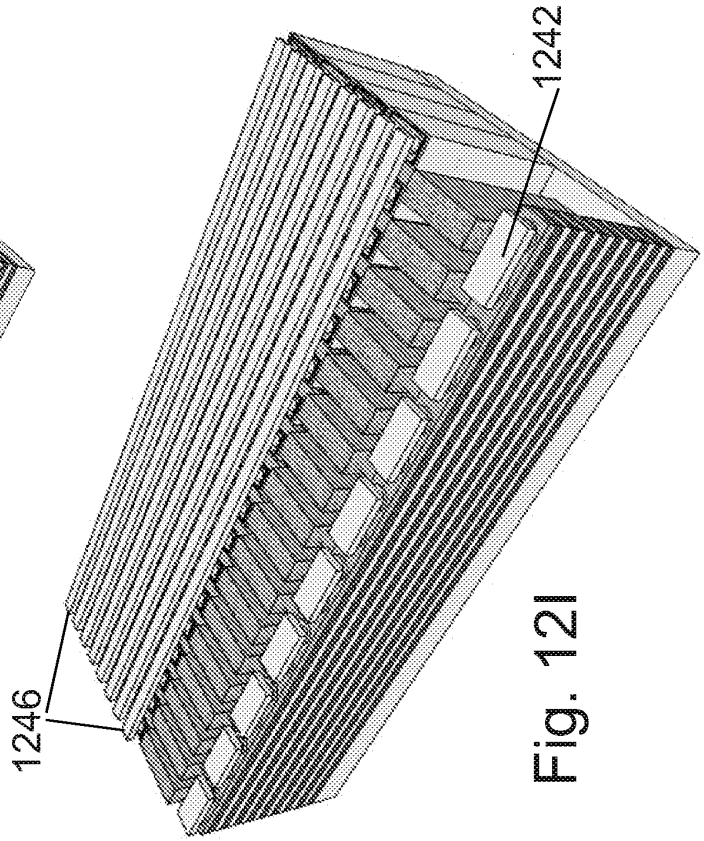


Fig. 12I

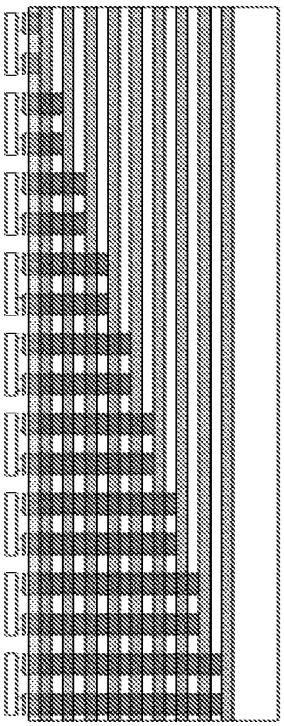


Fig. 12J

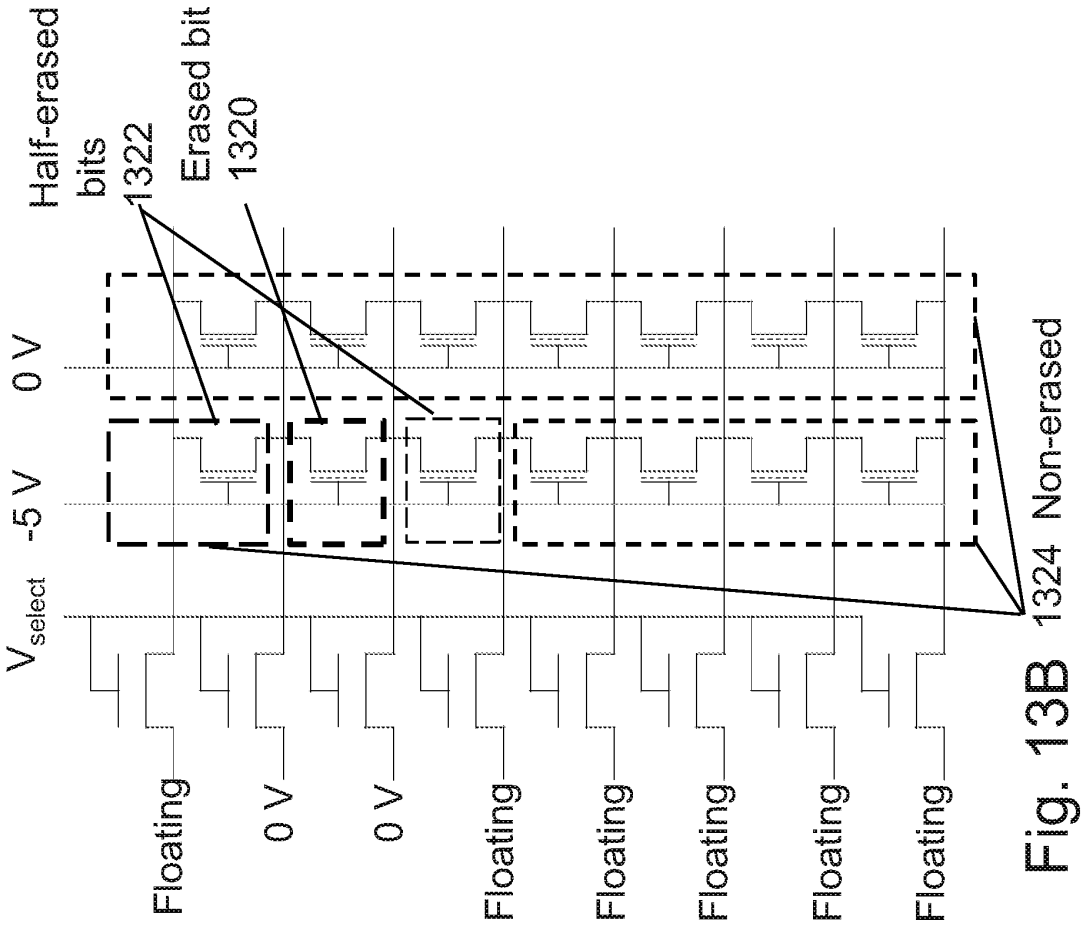


Fig. 13B

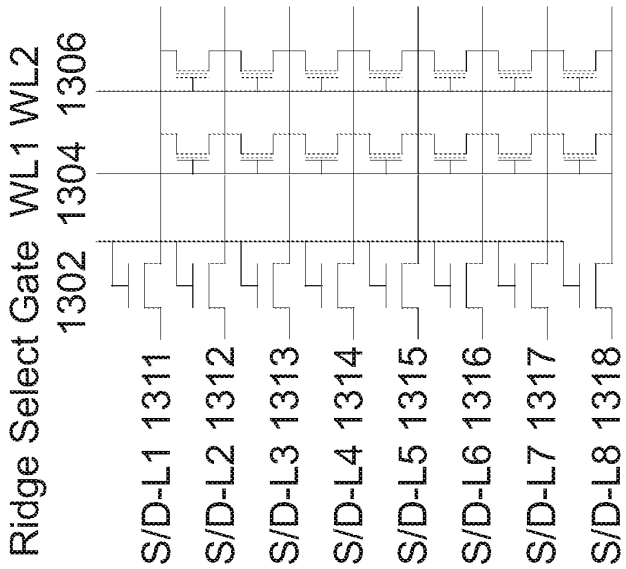


Fig. 13A

Per bit Erase - Gate Induced Drain Leakage ("GIDL")

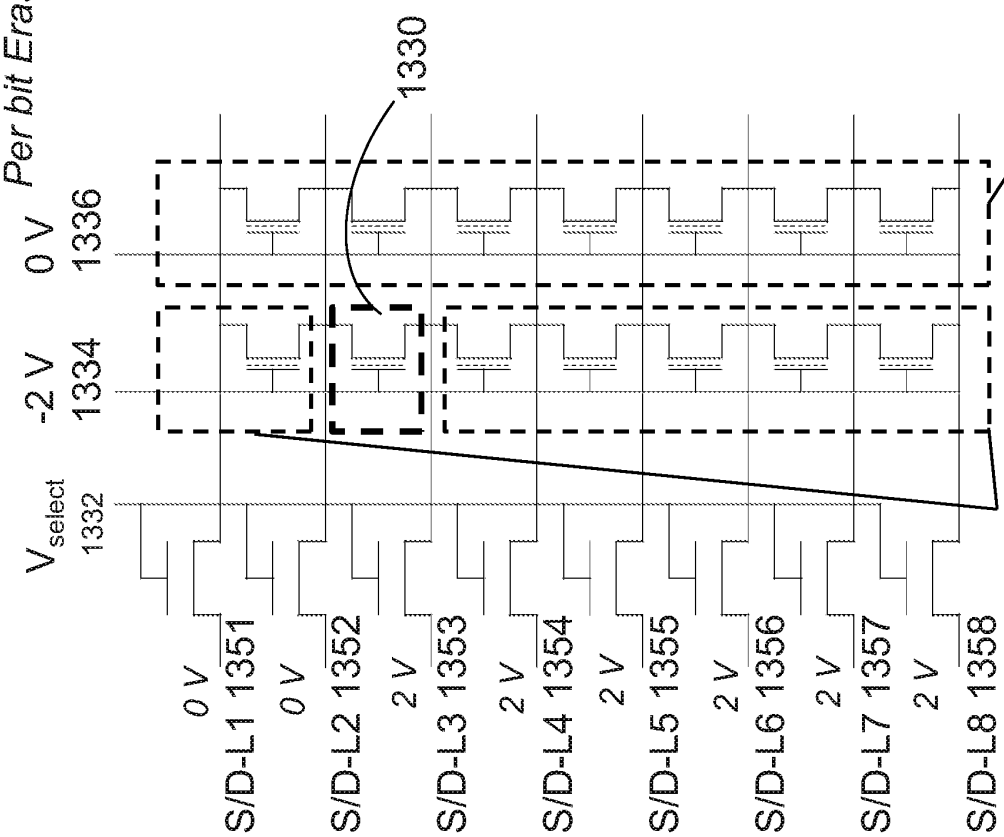


Fig. 13C Source side injection

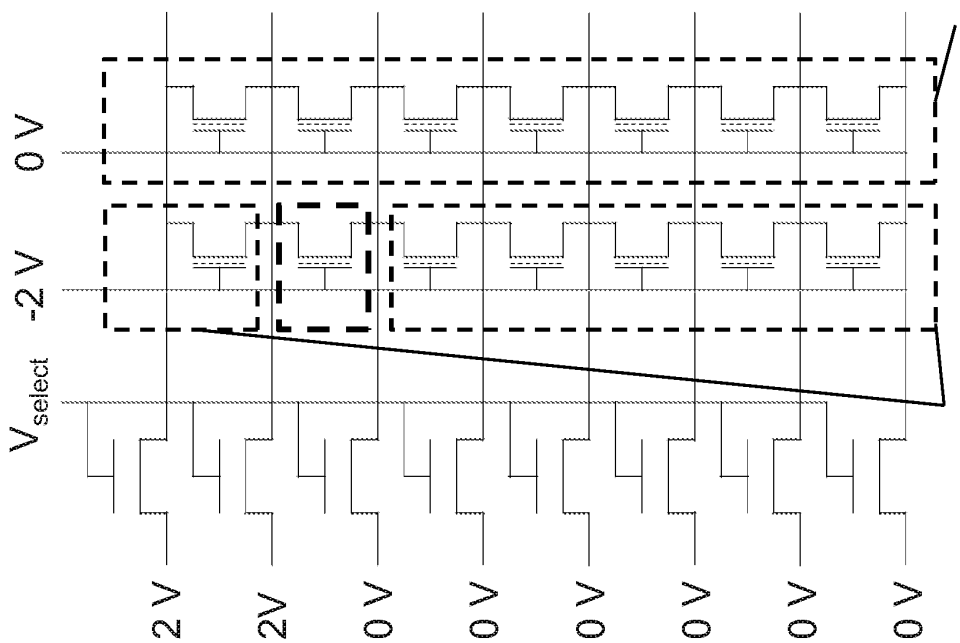


Fig. 13D Drain side injection

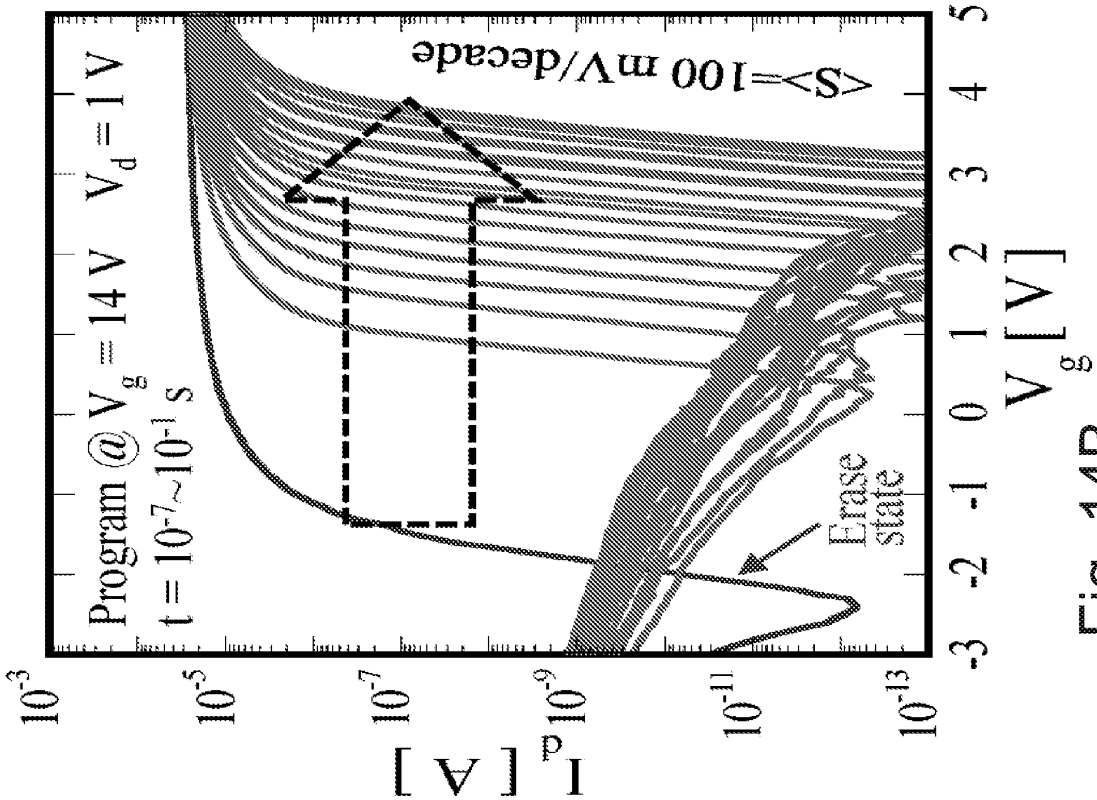


Fig. 14B

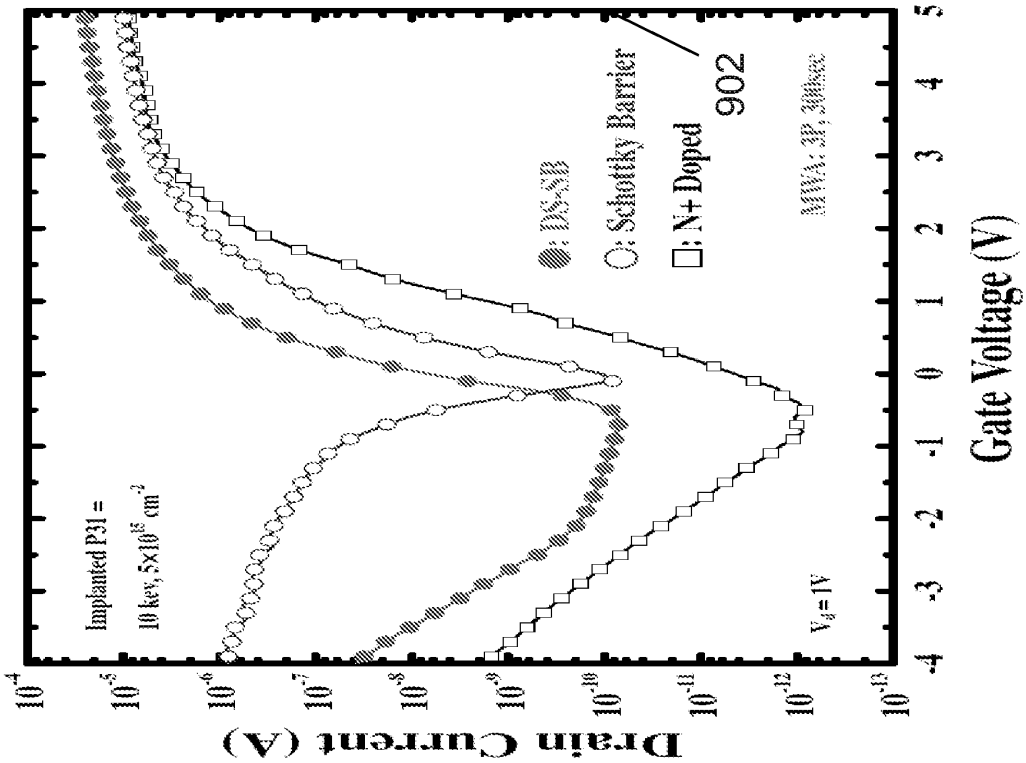


Fig. 14A

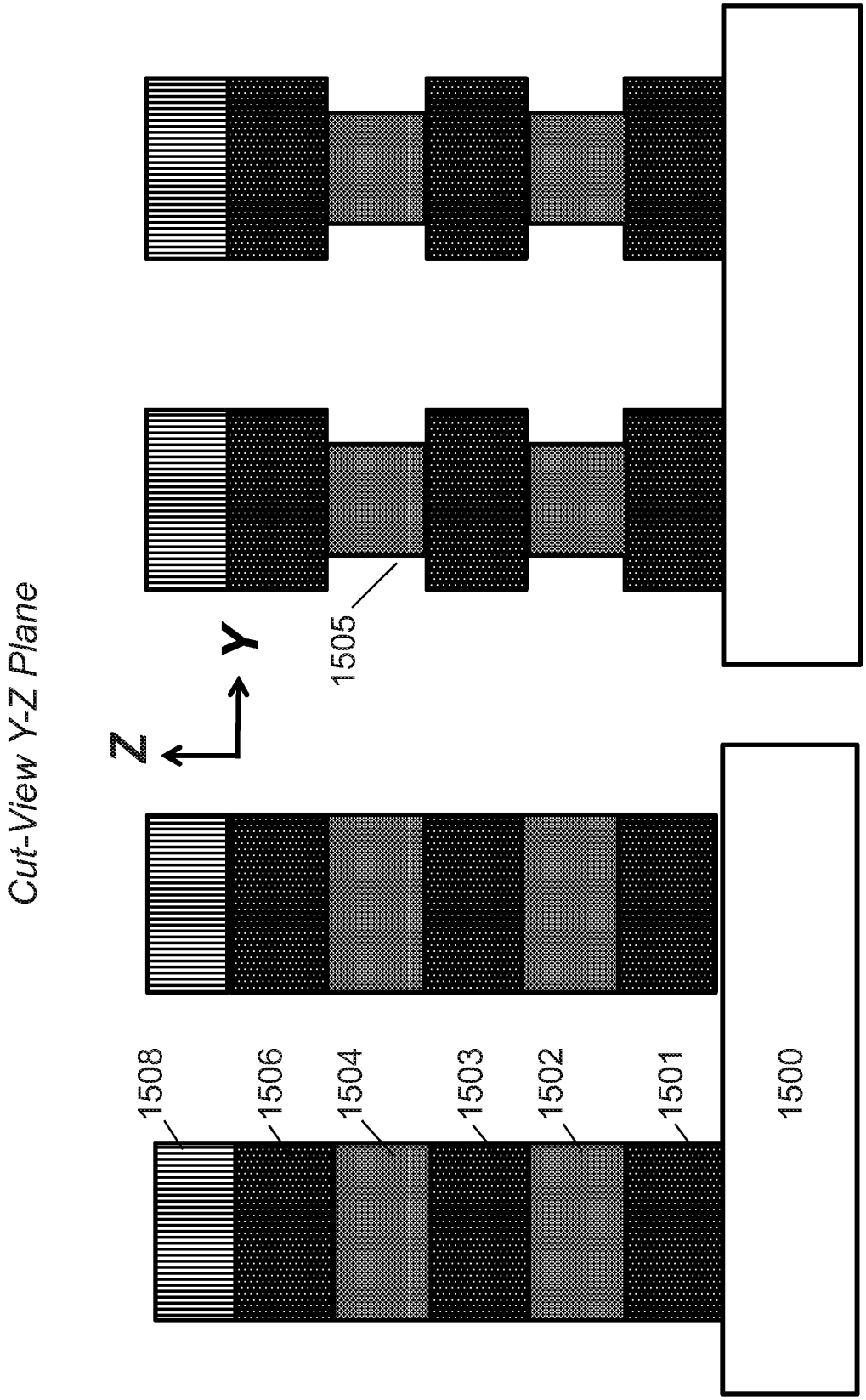


Fig. 15B

Fig. 15A

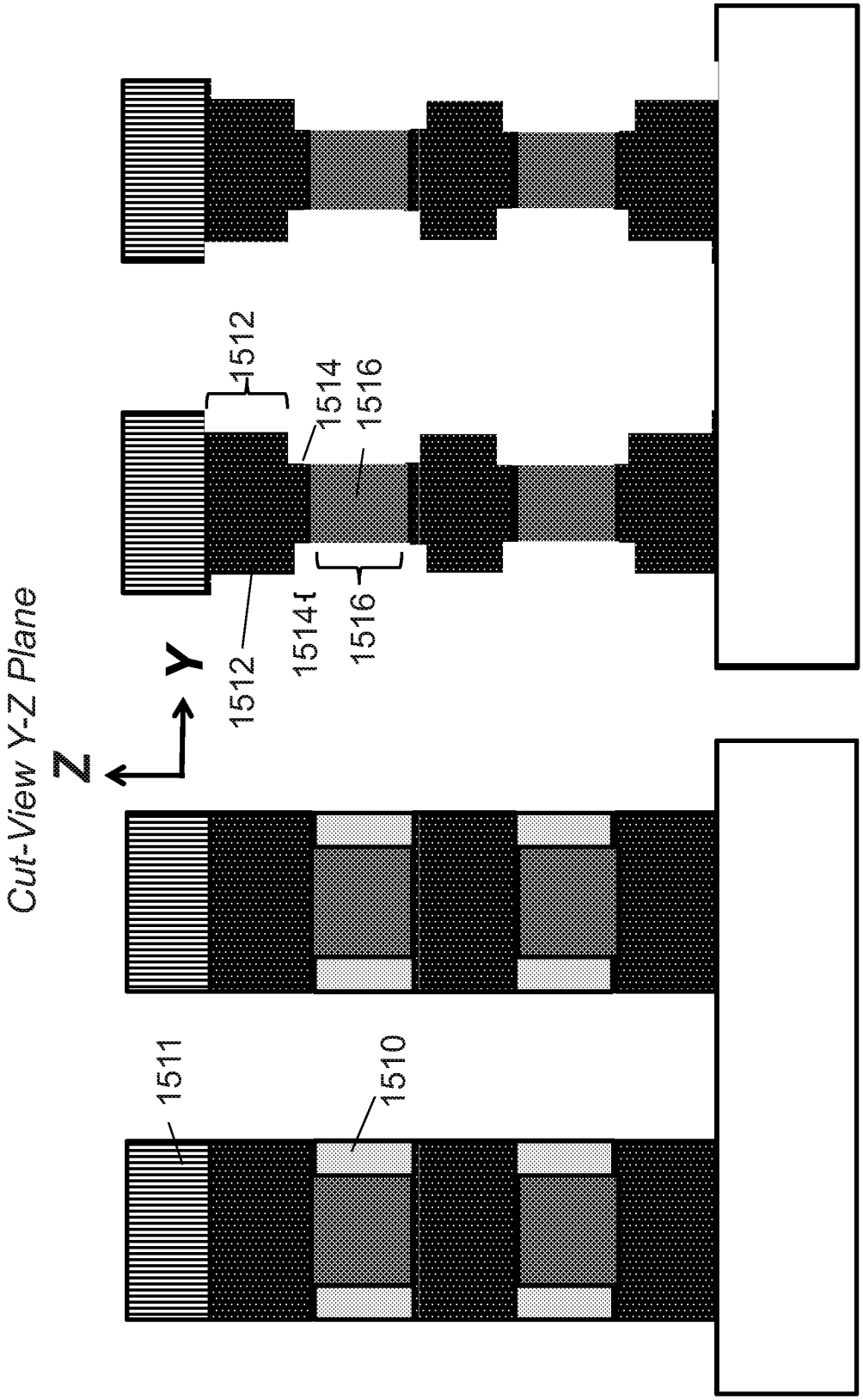


Fig. 15C

Fig. 15D

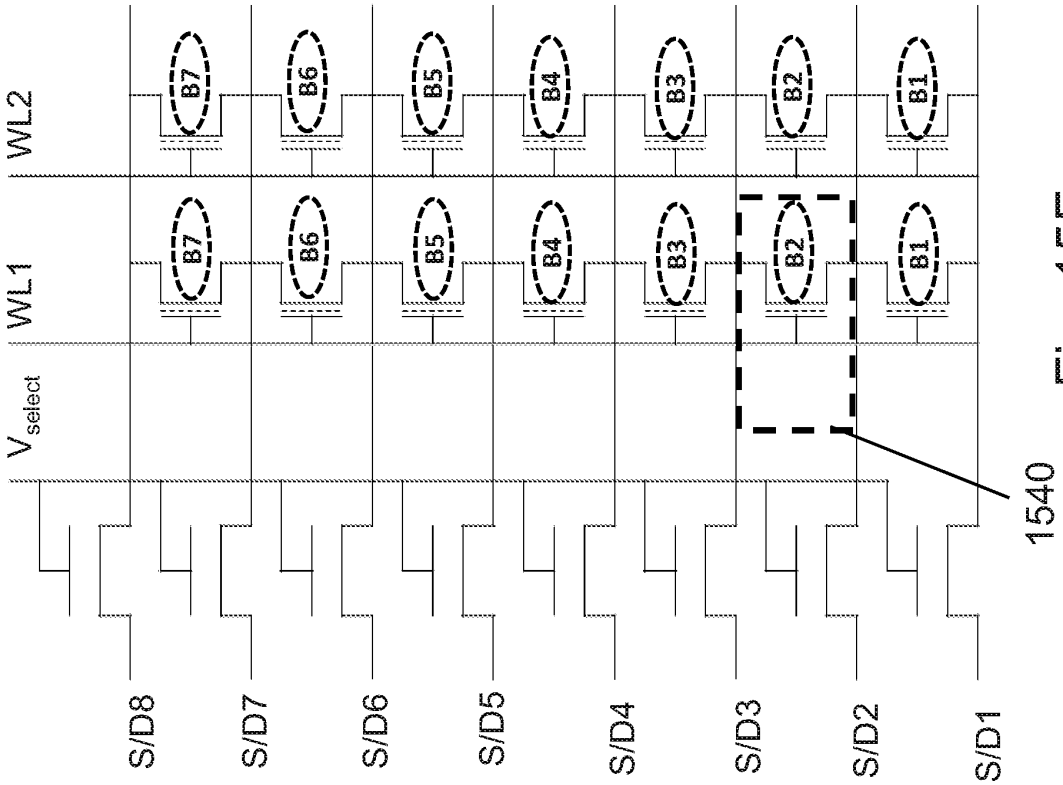


Fig. 15F

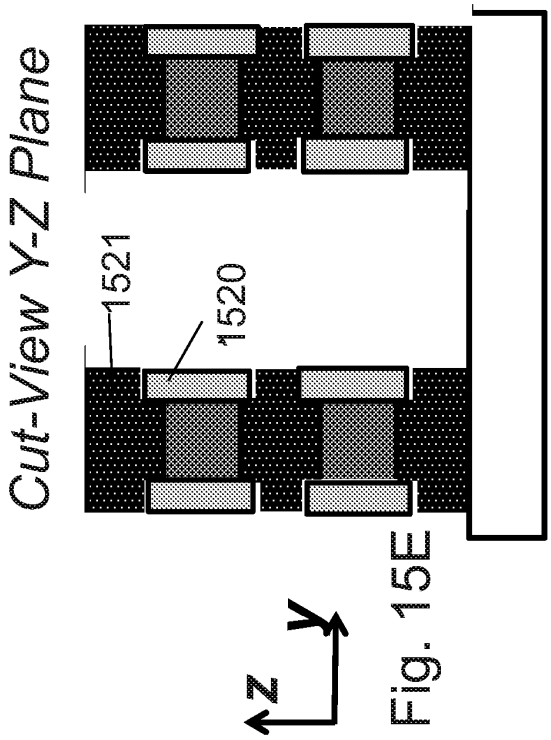


Fig. 15E

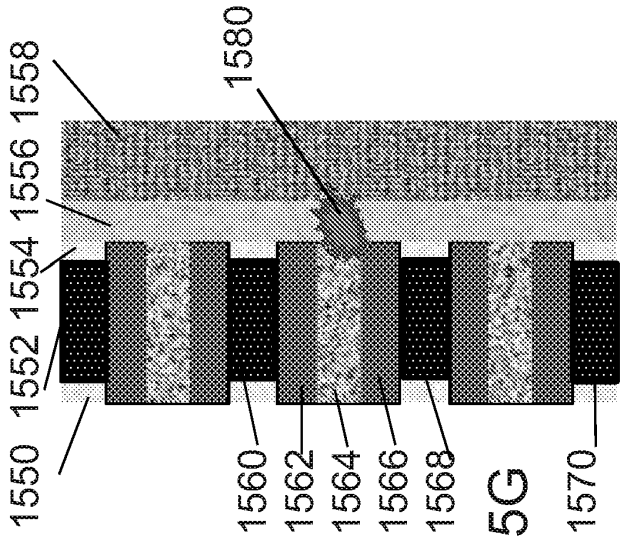


Fig. 15G

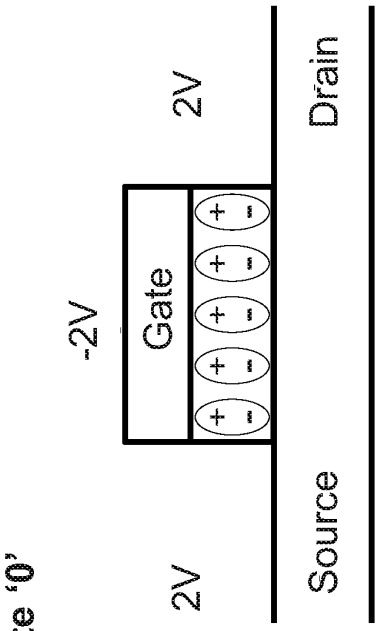


Fig. 16B

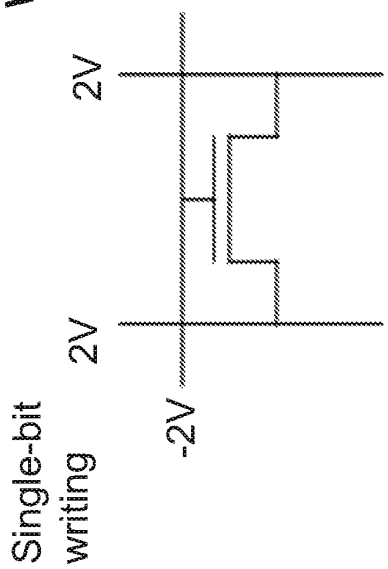


Fig. 16A

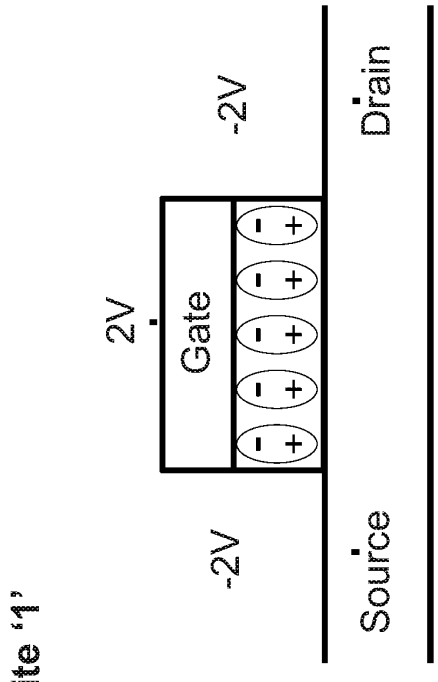


Fig. 16D

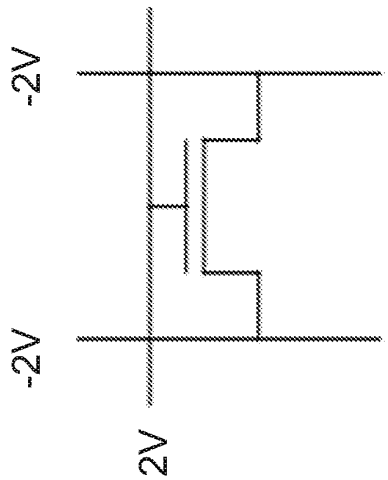
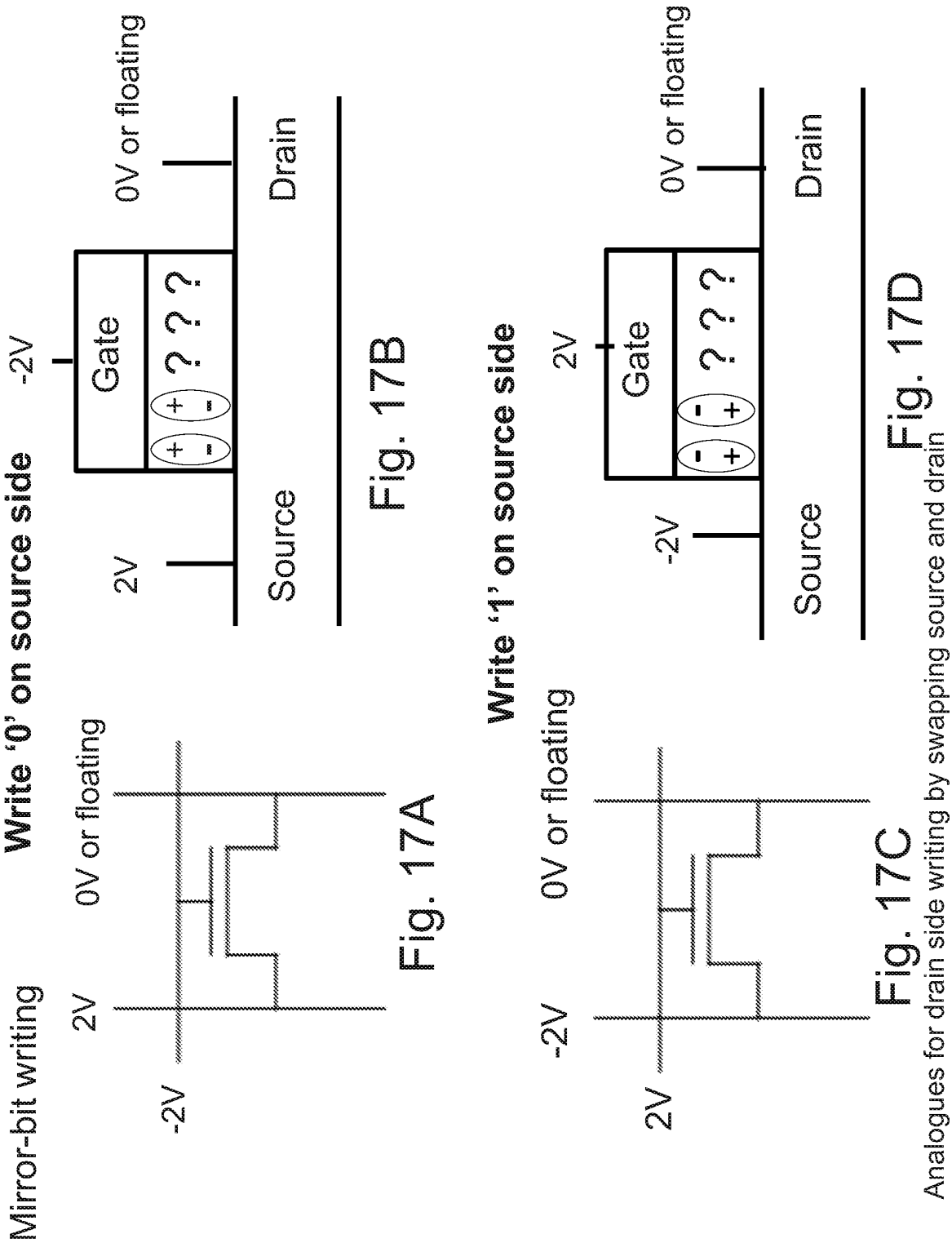
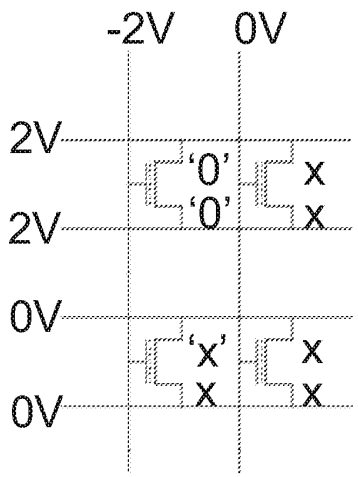


Fig. 16C



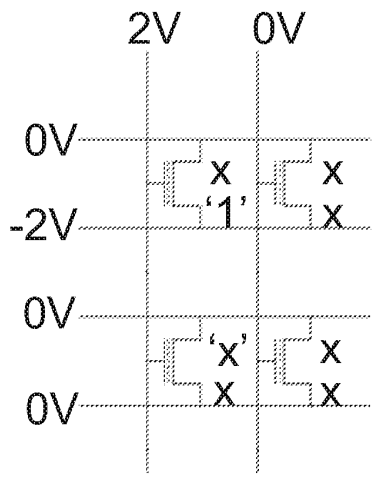
Analogues for drain side writing by swapping source and drain

Array with independent SD lines



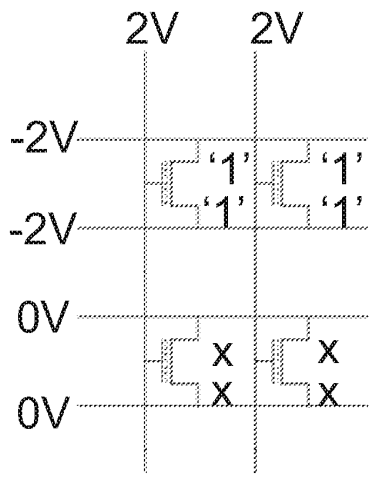
Bit specific write '0'

Fig. 18A



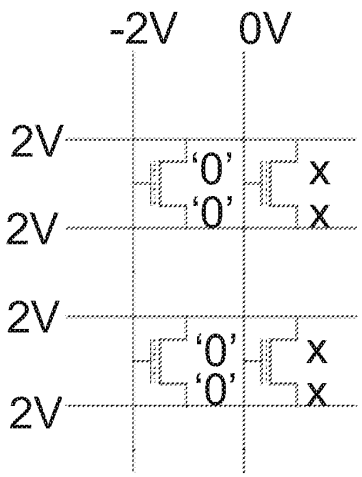
Bit specific in 2 bit per cell scheme write '1'

Fig. 18B



Entire SDL write

Fig. 18C



Entire WL write

Fig. 18D

For example: Apply 2V or –2V for selected WL and BL.
Apply 0V for unselected WL and BL (BL could be floating).
Symbol 'x' signifies non-affected bit.

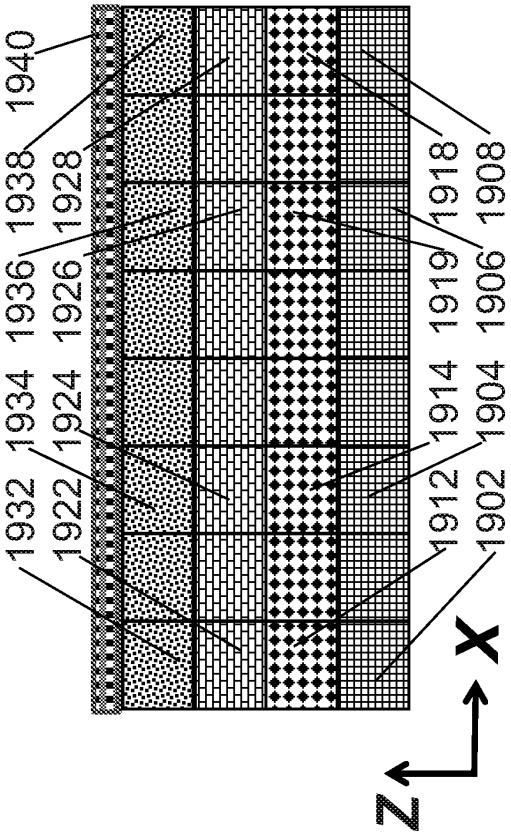


Fig. 19A

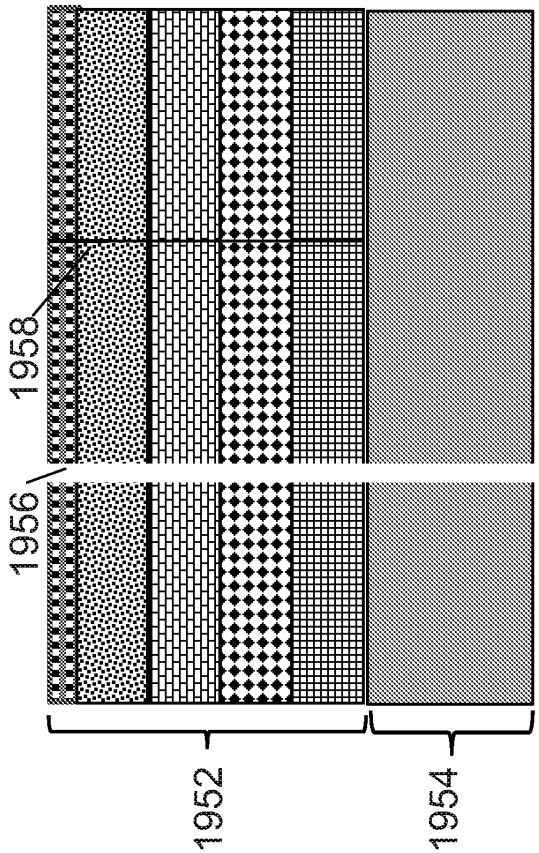


Fig. 19B

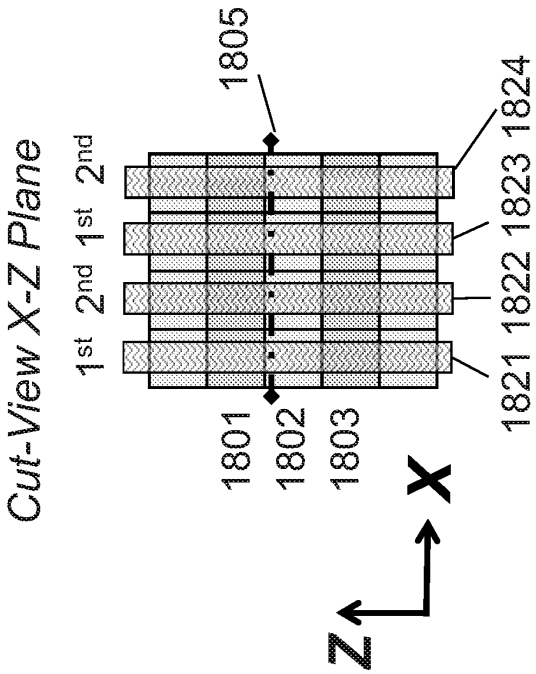


Fig. 18E

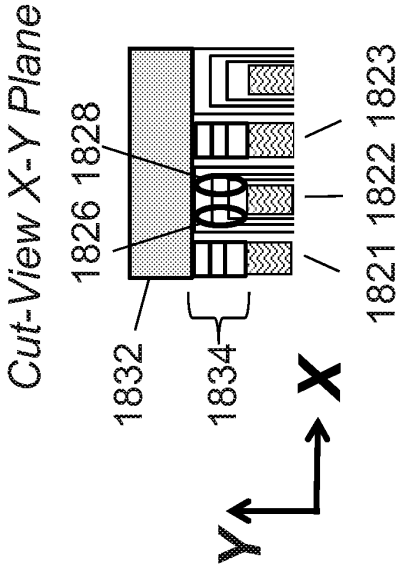
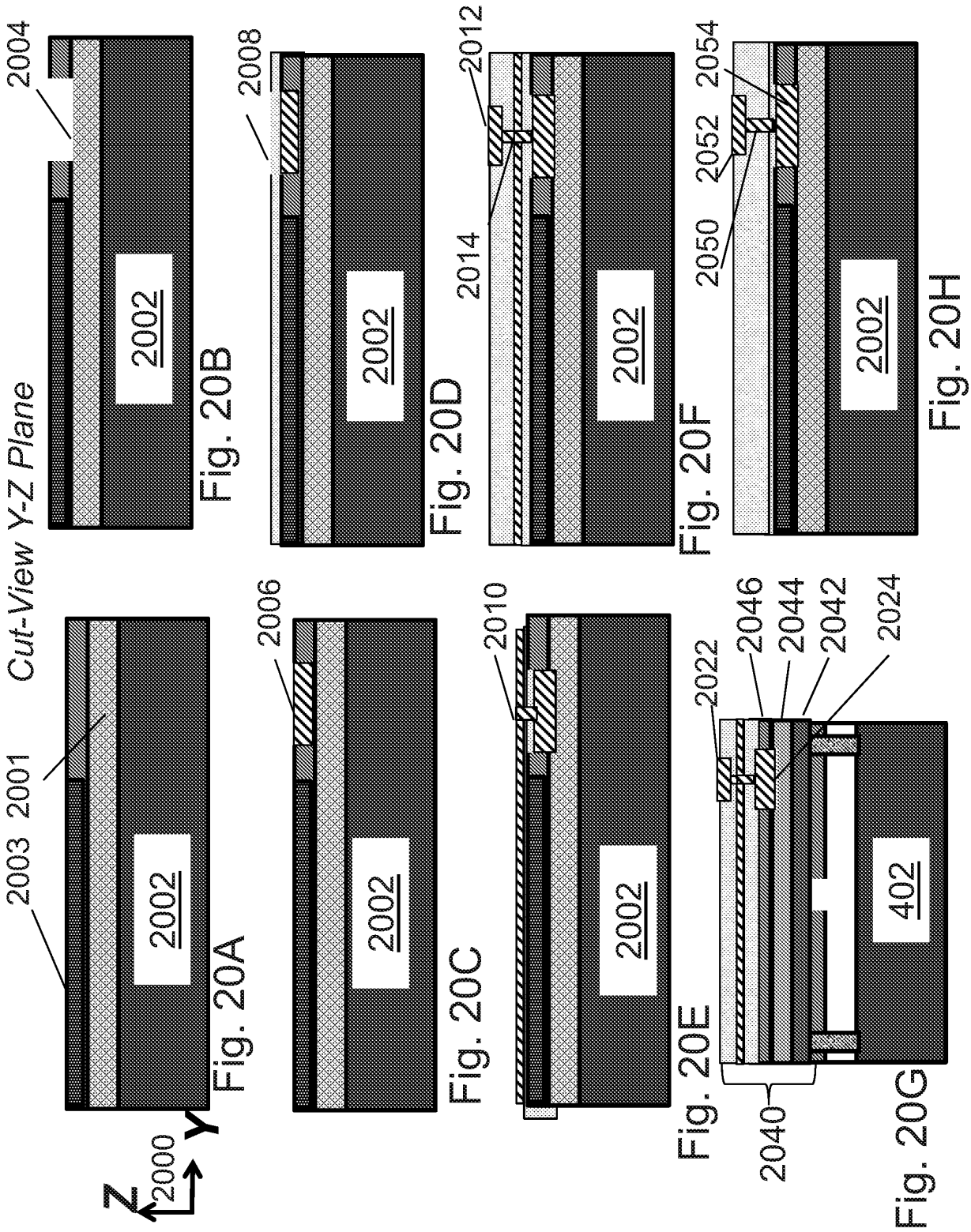


Fig. 18F


Cut-View Y-Z Plane
2003
2001
2002
Fig. 20A
2004
2002
Fig. 20B
2006
2002
Fig. 20C
2008
2002
Fig. 20D
2010
2002
Fig. 20E
2012
2014
2002
Fig. 20F
2022
2046
2044
2042
2024
402
2040
Fig. 20G
2054
2052
2050
2002
Fig. 20H
Z
Y
2000

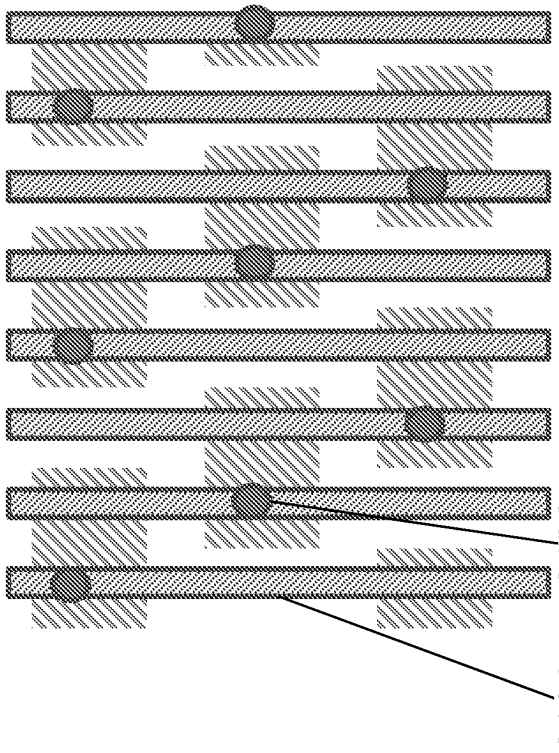


Fig. 21B

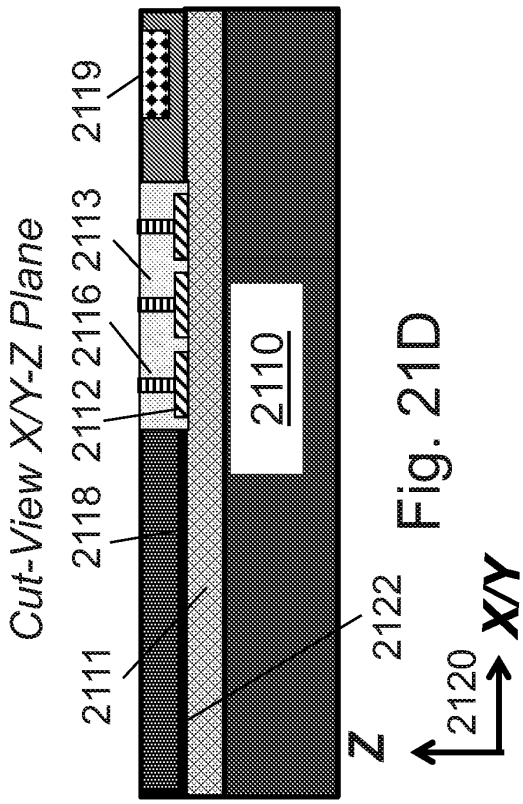


Fig. 21D

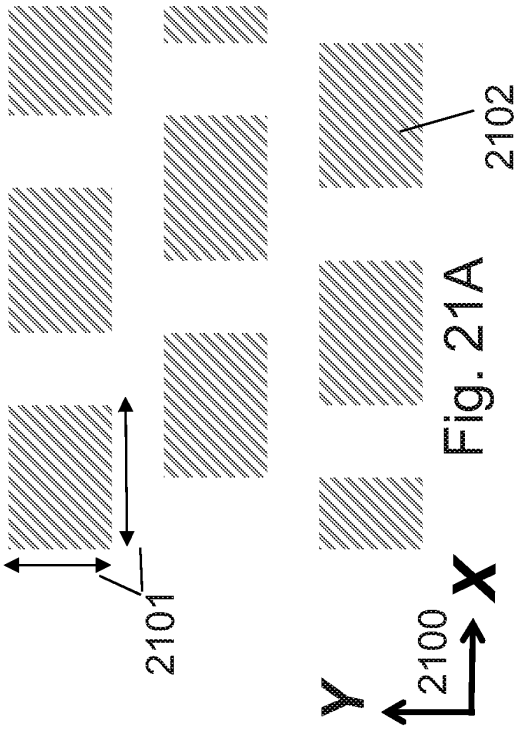


Fig. 21A

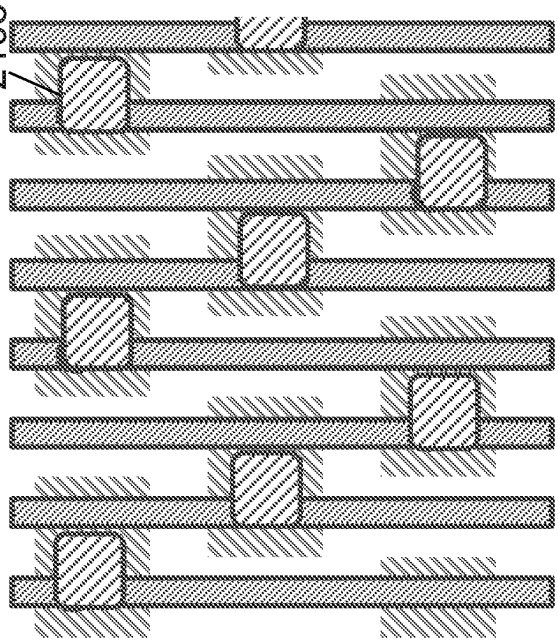


Fig. 21C

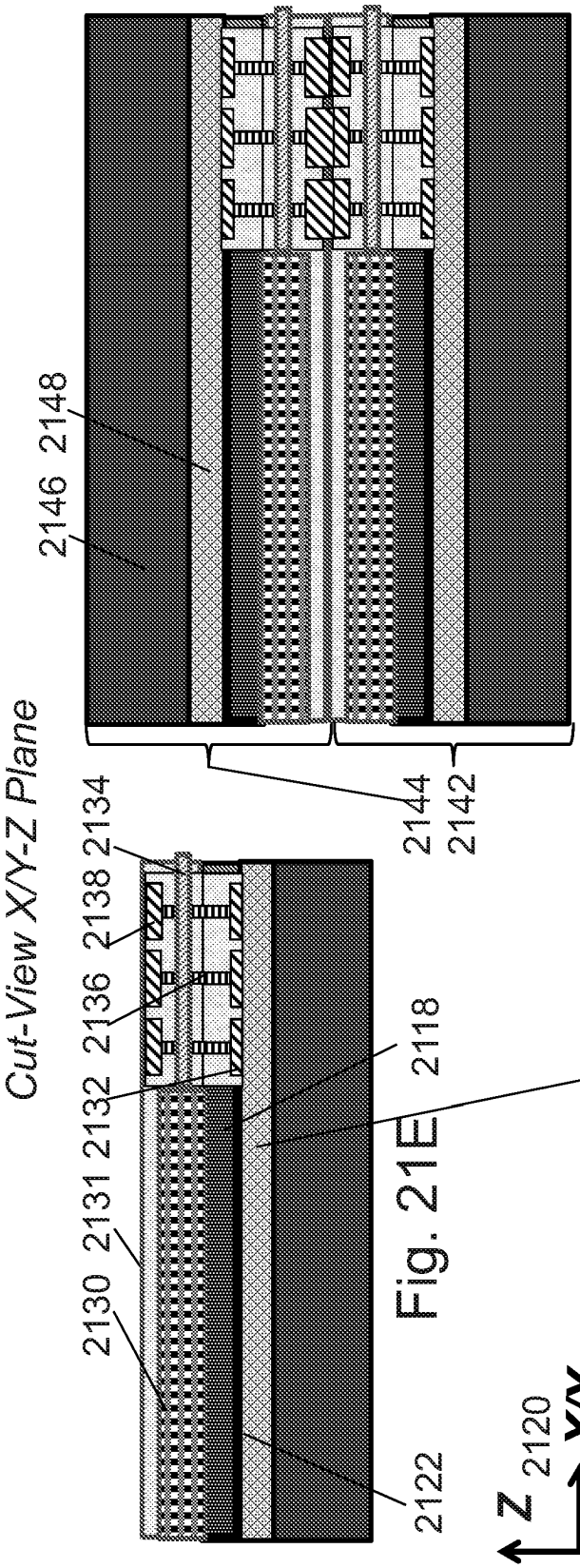


Fig. 21F

Fig. 21E

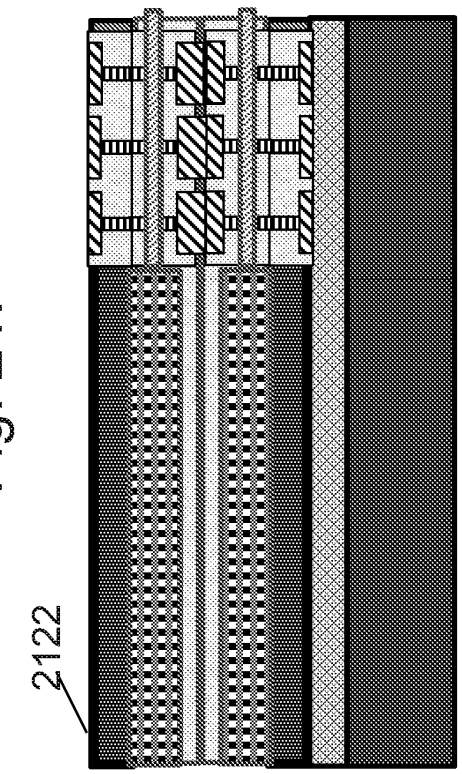


Fig. 21H

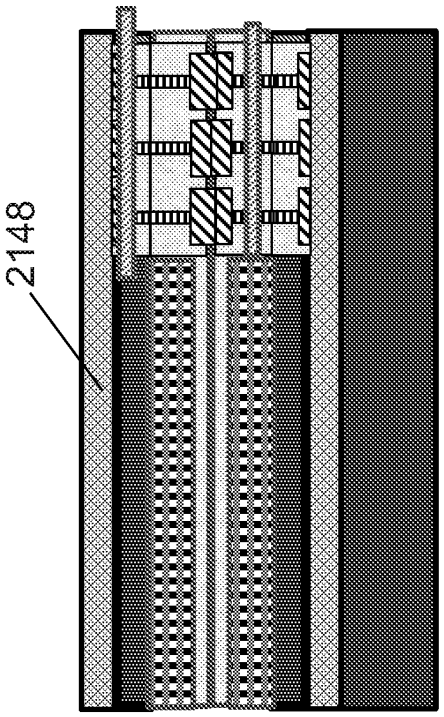


Fig. 21G

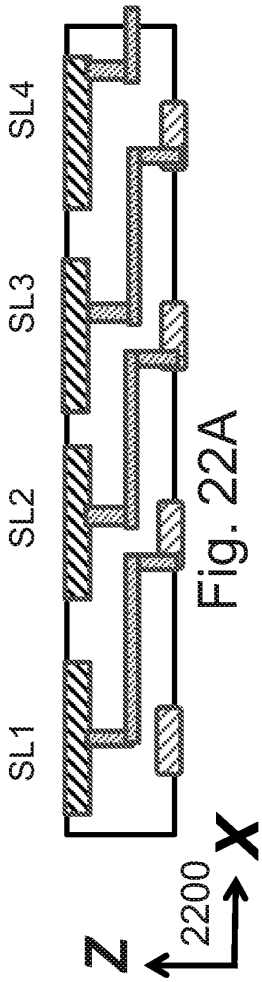


Fig. 22A

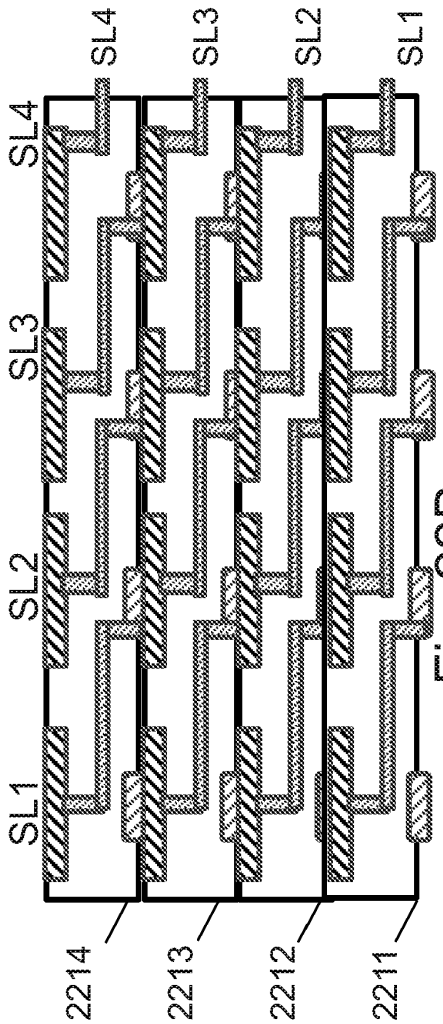


Fig. 22B

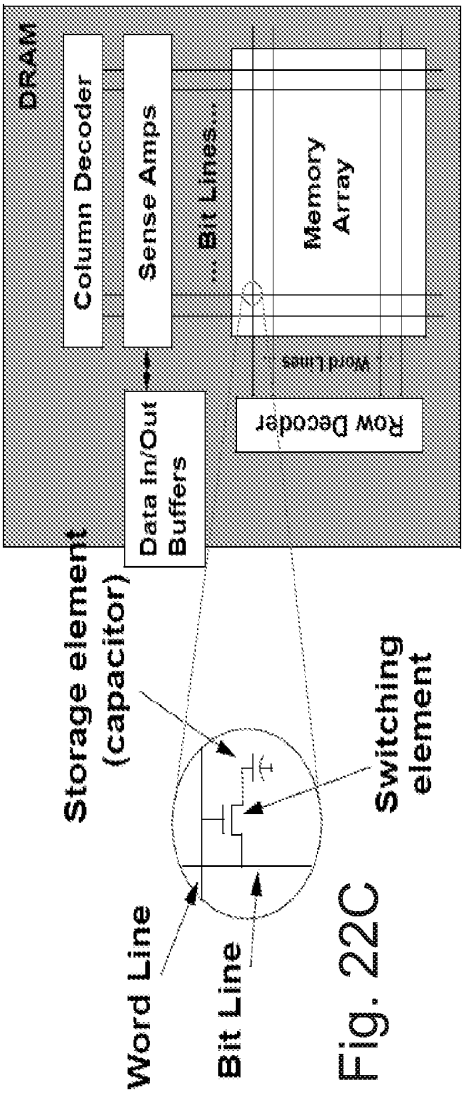


Fig. 22C

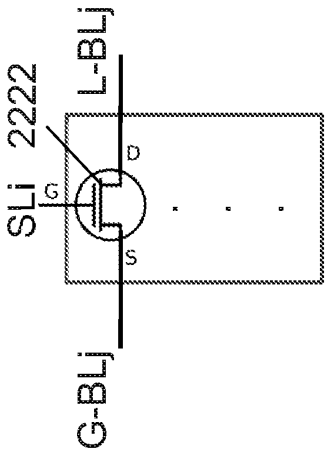


Fig. 22D

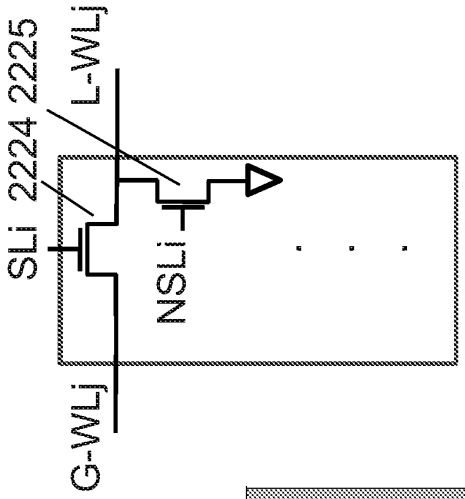


Fig. 22E

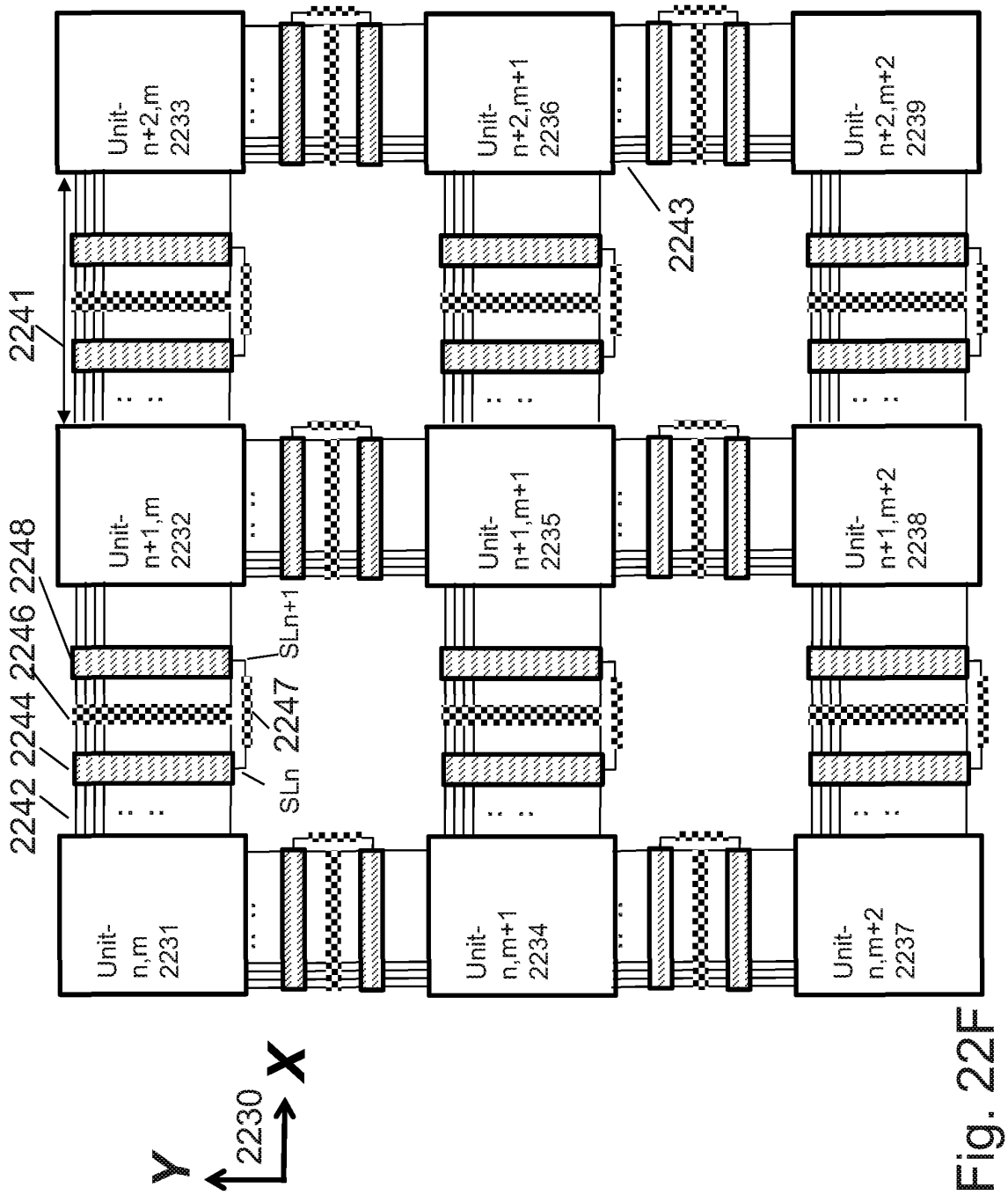

Unit-
n+2,m
2233
Unit-
n+2,m+1
2236
Unit-
n+2,m+2
2239
2243
2241
Unit-
n+1,m
2232
Unit-
n+1,m+1
2235
Unit-
n+1,m+2
2238
2248
2246
2244
2242
SLn+1
2247
SLn
Unit-
n,m
2231
Unit-
n,m+1
2234
Unit-
n,m+2
2237
Y
2230
X

Fig. 22F

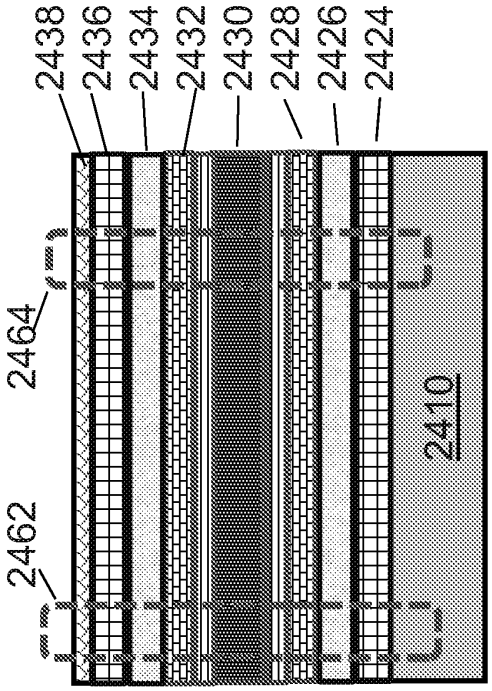


Fig. 24A

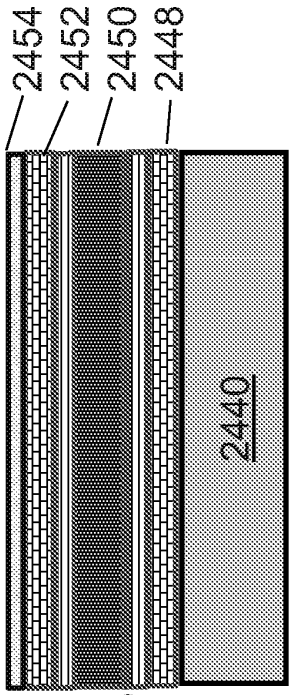


Fig. 24B

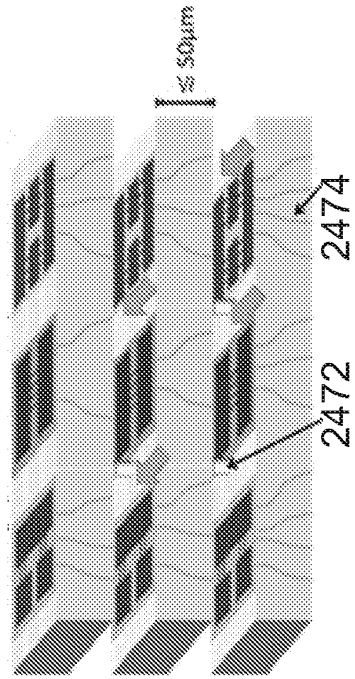


Fig. 24C

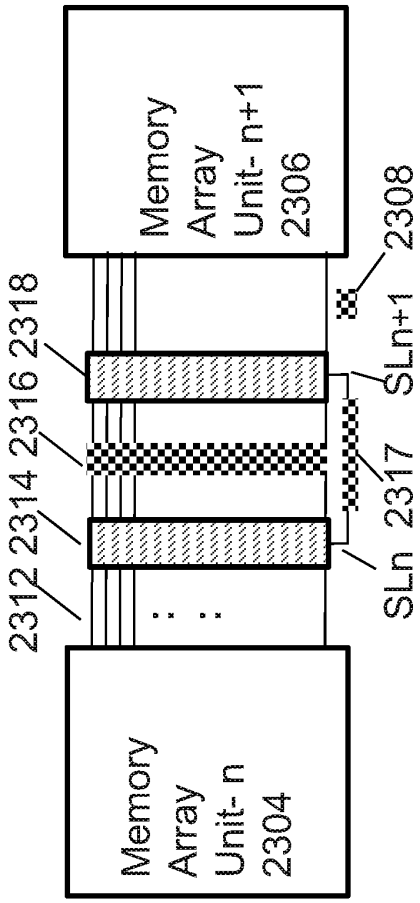


Fig. 23A

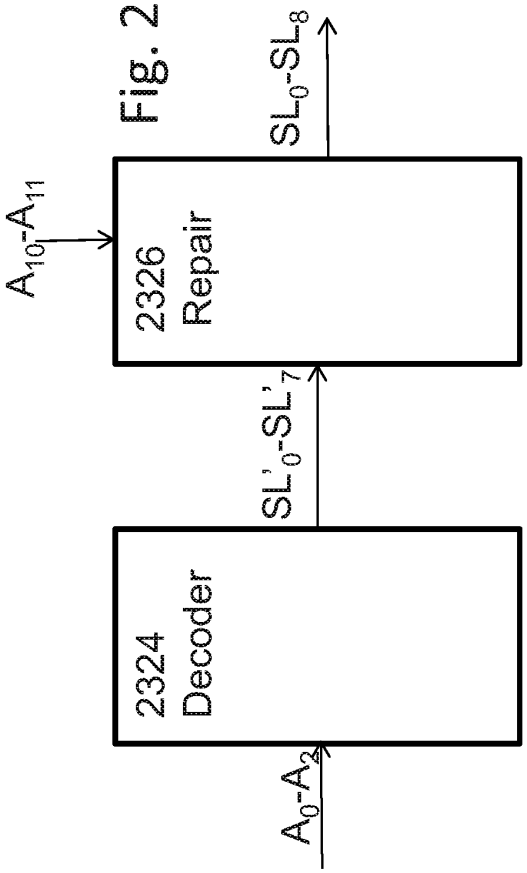


Fig. 23B

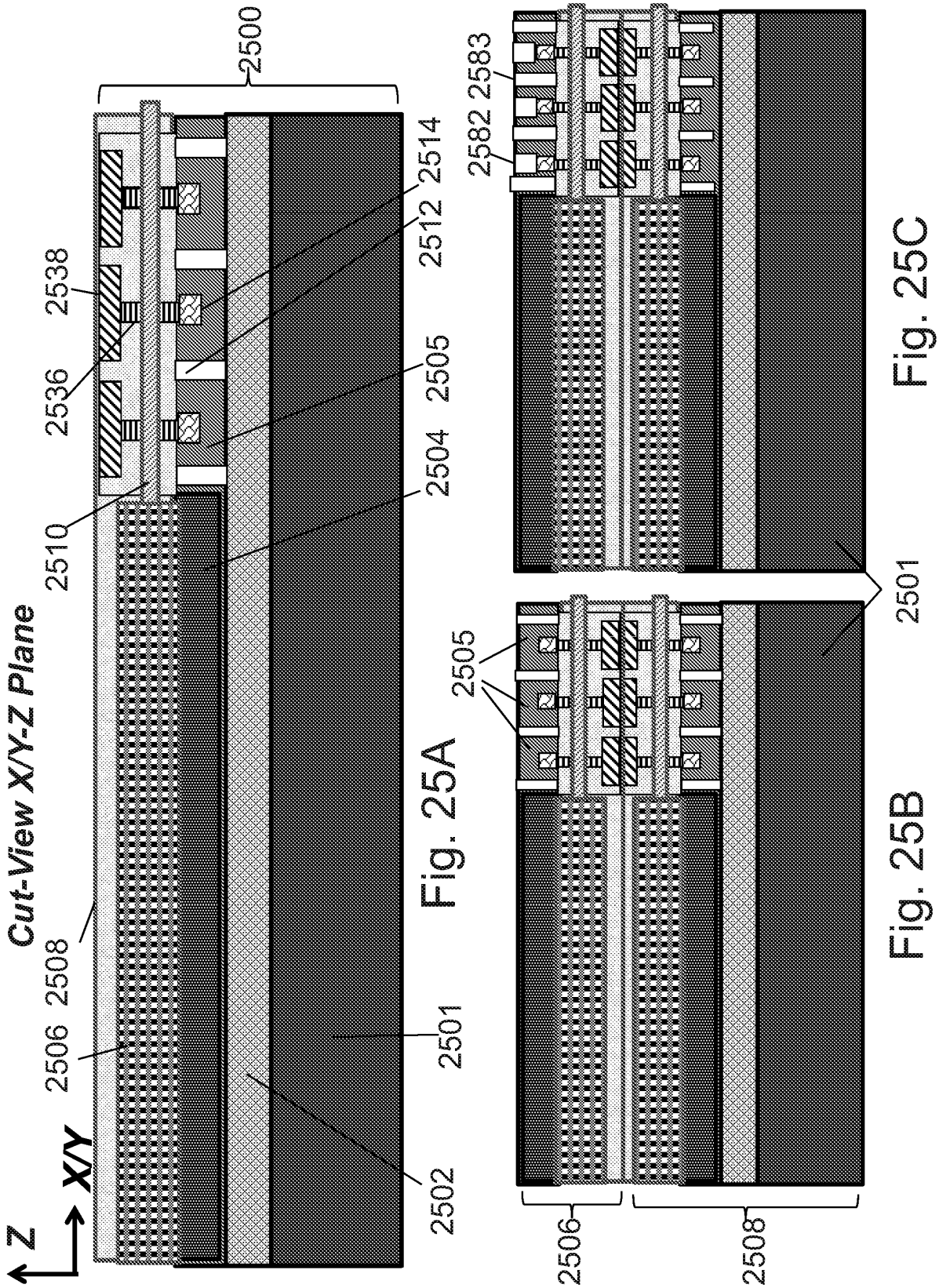


Fig. 25A

Fig. 25B

Fig. 25C

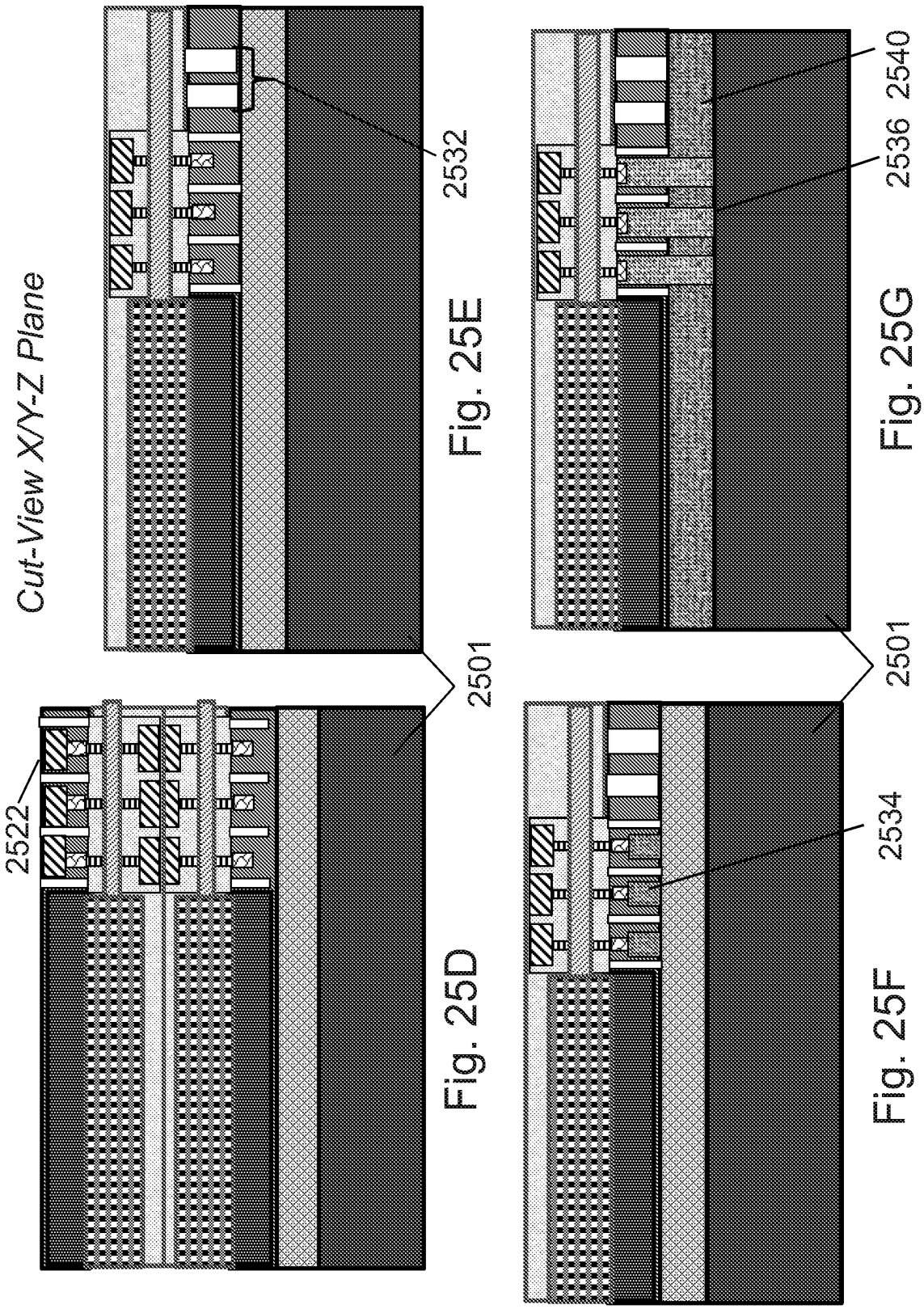


Fig. 25D

Fig. 25E

Fig. 25F

Fig. 25G

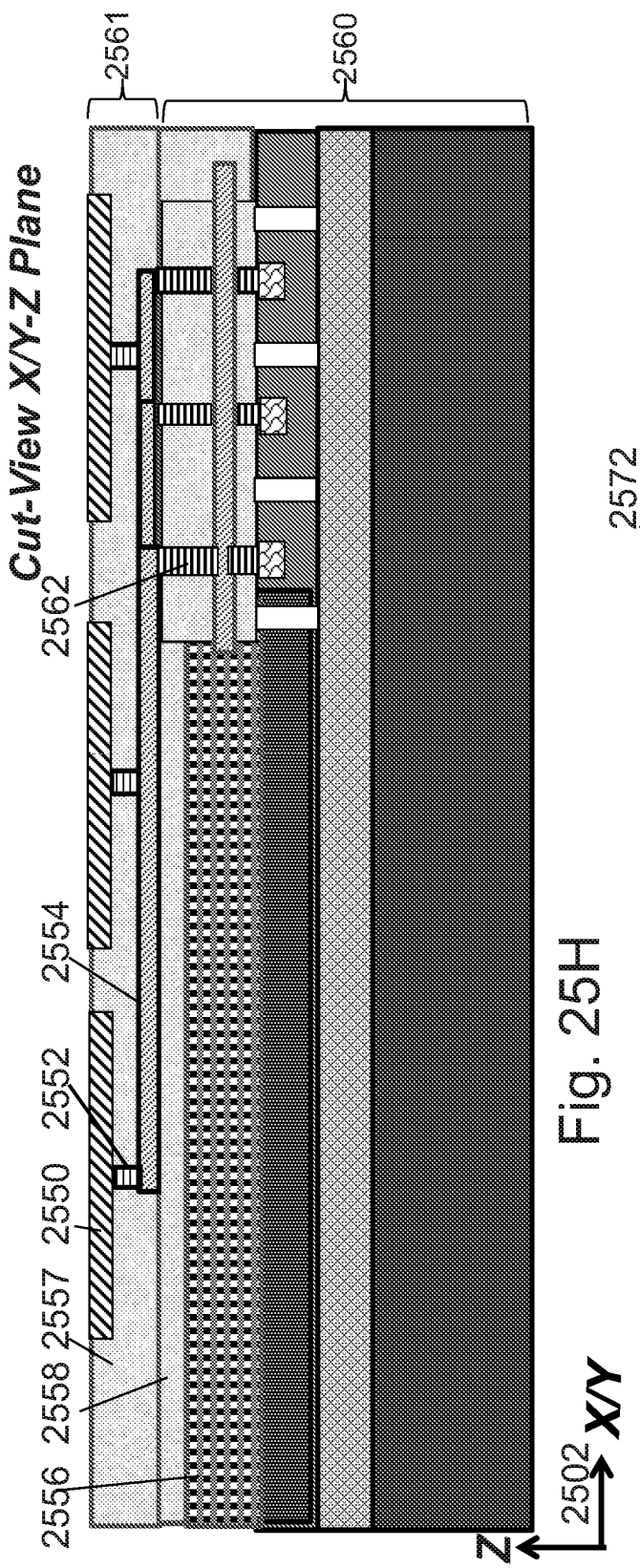


Fig. 25H

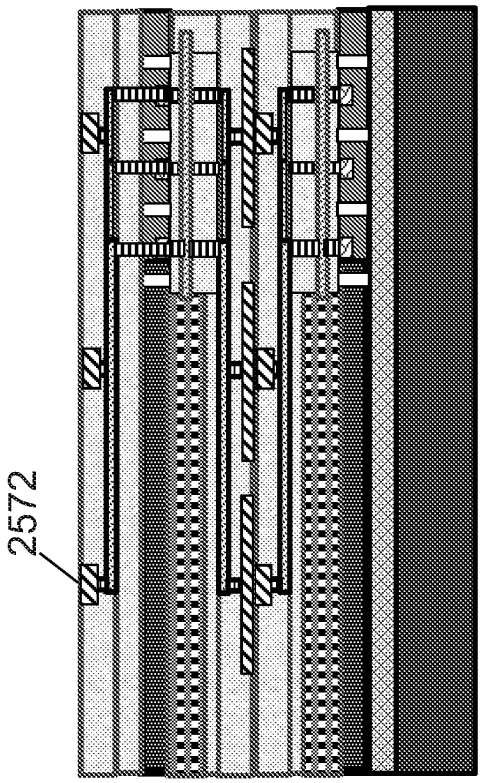


Fig. 25J

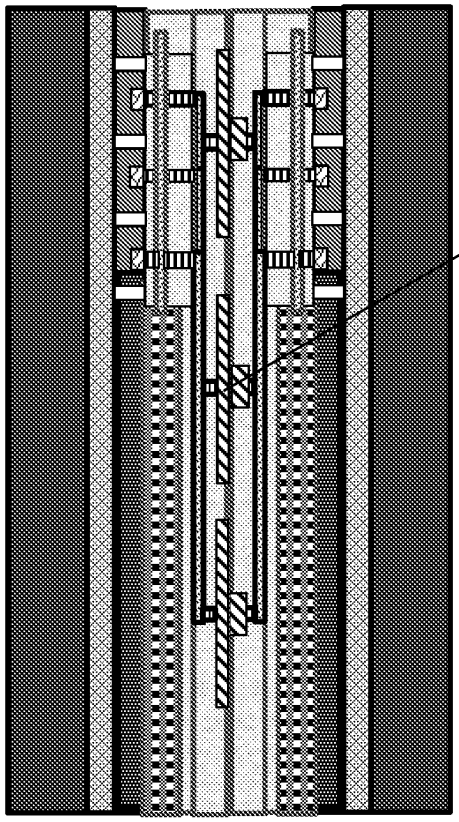


Fig. 25I

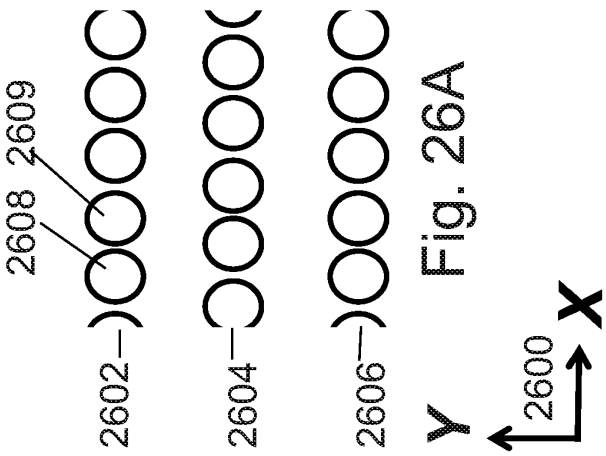


Fig. 26A

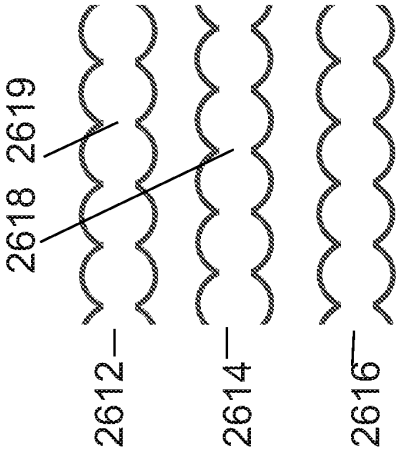


Fig. 26B

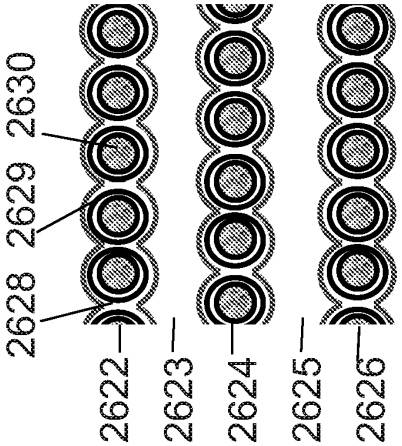


Fig. 26C

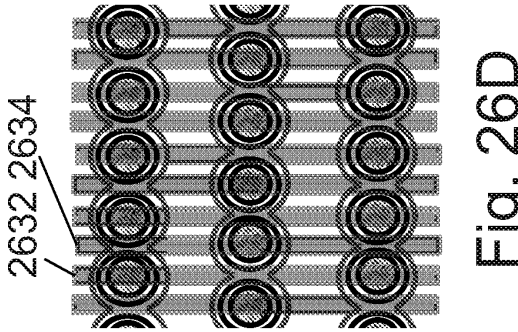


Fig. 26D

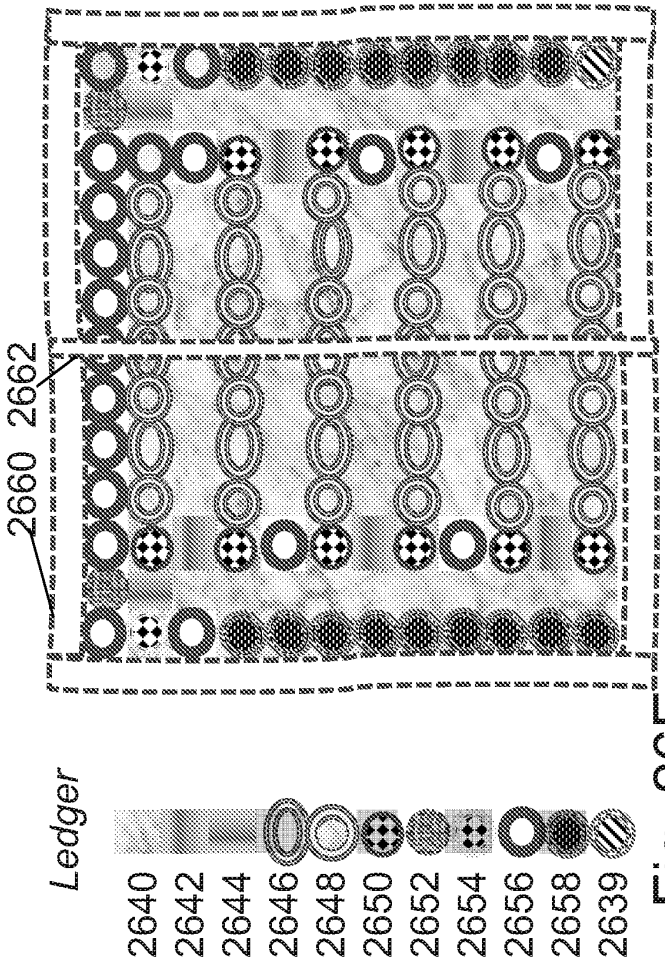


Fig. 26E

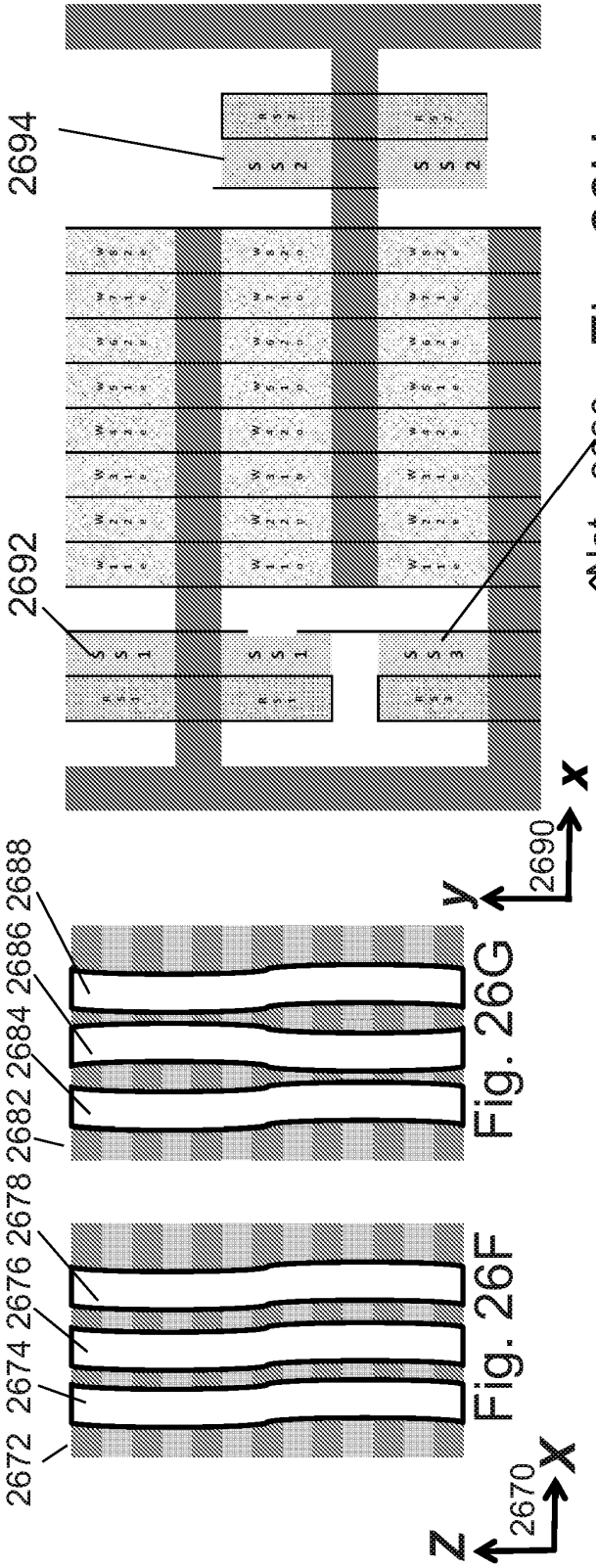


Fig. 26H

Fig. 26G

Fig. 26F

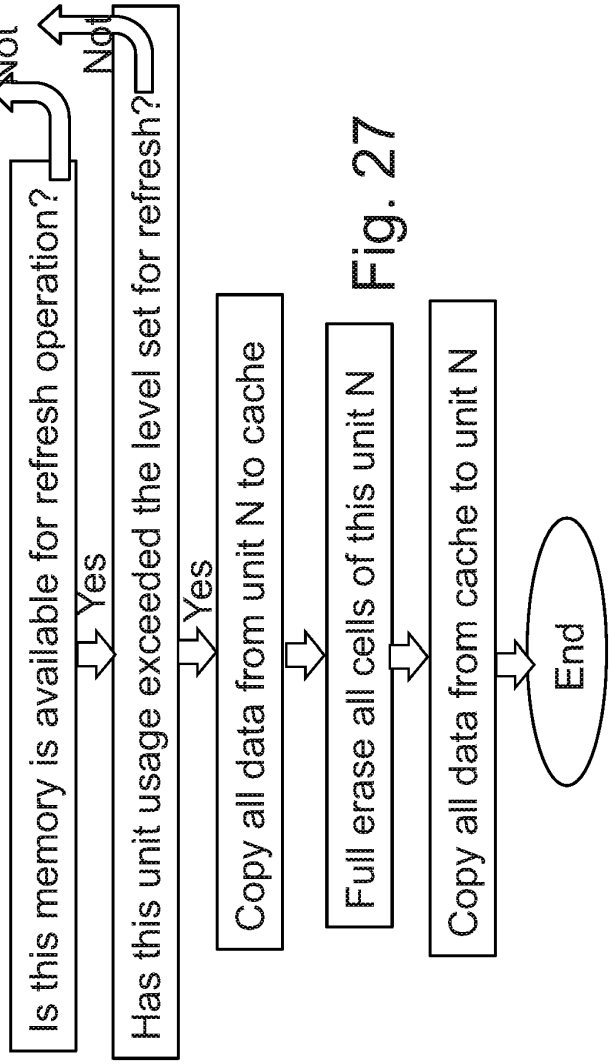


Fig. 27

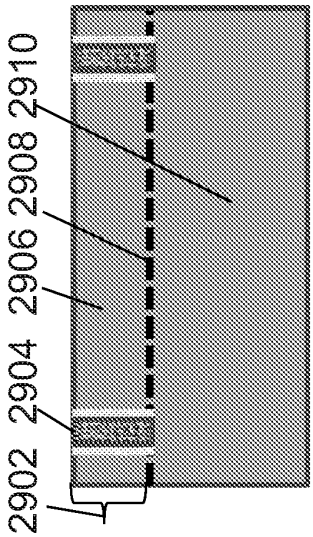


Fig. 29A

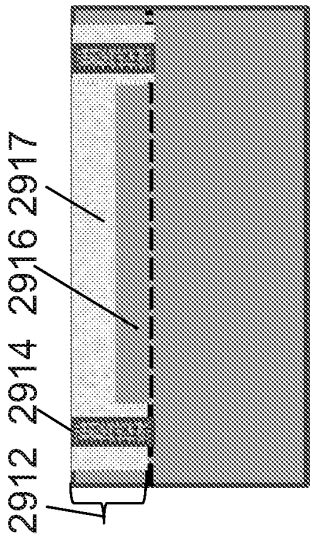


Fig. 29B

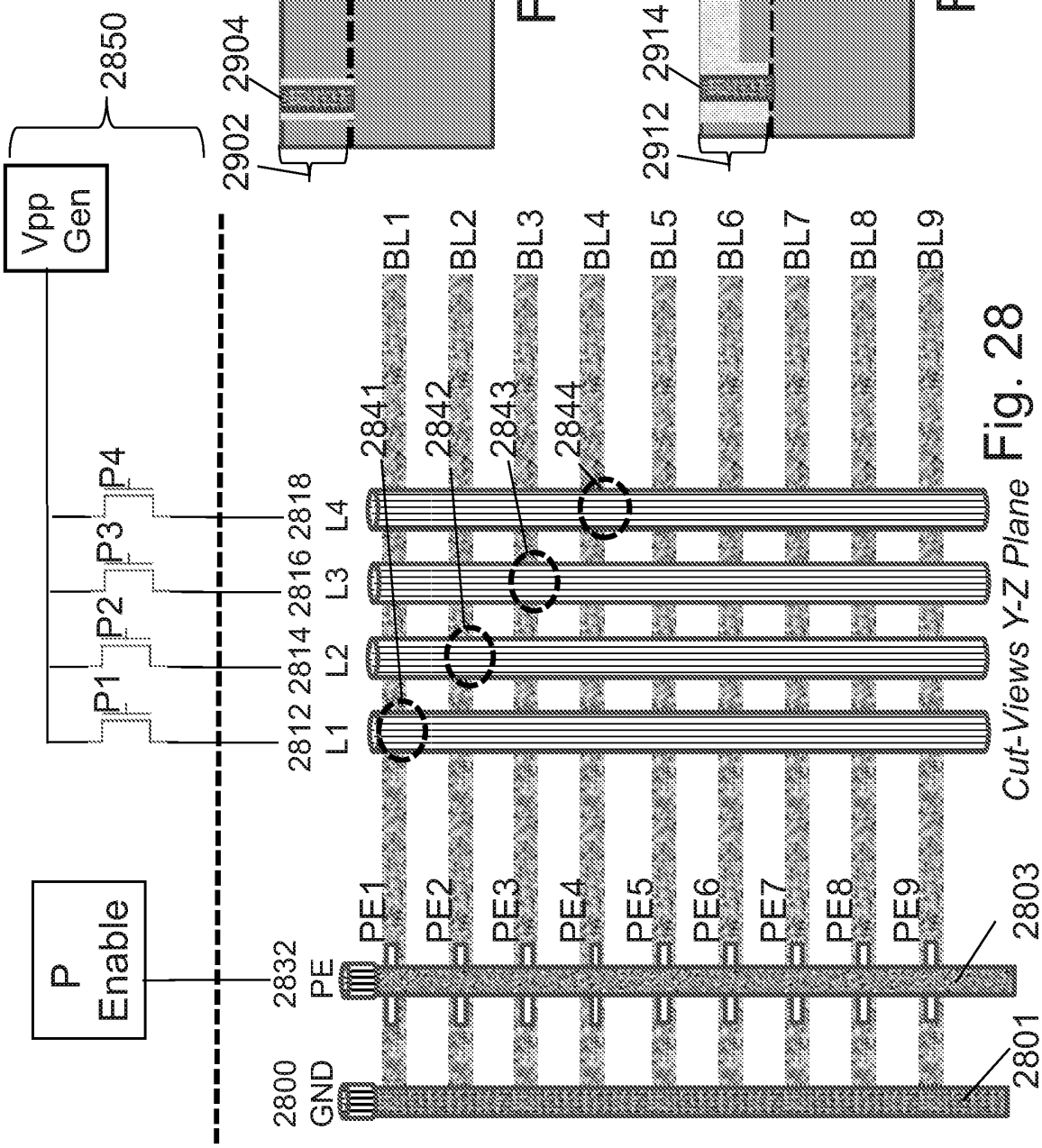


Fig. 28

Espacenet

CN108695336A THREE-DIMENSIONAL SEMICONDUCTOR MEMORY DEVICE AND METHOD OF FABRICATING THE SAME

Applicants: SAMSUNG ELECTRONICS CO LTD

Inventors: HWANG JEE-HOON,LEE JU-YEON,LEE TAE-HEE

Classifications:

IPC ***H01L27/11529; H01L27/11531; H01L27/11551; H01L27/11573; H01L27/11578;***

CPC **H01L21/76816 (EP); H01L23/485 (EP); H01L23/5283 (US); H01L27/11526 (US); H01L27/11529 (CN); H01L27/11531 (CN); H01L27/11551 (CN); H01L27/11556 (US); H01L27/11565 (EP); H01L27/11573 (EP,CN,US); H01L27/11575 (EP); H01L27/11578 (CN); H01L27/11582 (EP,US);** H01L2221/1063 (EP);

Priorities: KR20170045114A 2017-04-07

Application: CN108695336A·2018-04-08

Publication: CN108695336A·2018-10-23

Published as: CN108695336A;**JP2018182319A; KR20180114262A; US2018294225A1**

THREE-DIMENSIONAL SEMICONDUCTOR MEMORY DEVICE AND METHOD OF FABRICATING THE SAME

Abstract

Provided are a three-dimensional semiconductor memory device and a method of fabricating the same. The device may include a substrate including a peripheral circuit region and a cell array region, anelectrode structure including a plurality of electrodes vertically stacked on the cell array region of the substrate, a peripheral logic circuit provided on the peripheral circuit region of the substrate, the peripheral logic circuit including a first impurity region doped with first impurities, a peripheral contact plug connected to the first impurity region, and a second impurity region betweenthe first impurity region and the peripheral contact plug, with the second impurity region including second impurities different from the first impurities. The peripheral contact plug includes a lowerportion contacting the second impurity region, and an upper portion continuously extending from the lower portion, a lower width of each of the lower and upper portions is less than an upper width thereof, and the upper width of the lower portion is greater than the lower width of the upper portion.

Micron Ex. 1002, p. 718
Micron v. YMTC
IPR2024-00789

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 108695336 A
(43)申请公布日 2018.10.23

(21)申请号 201810305417.7

(22)申请日 2018.04.08

(30)优先权数据
10-2017-0045114 2017.04.07 KR

(71)申请人 三星电子株式会社
地址 韩国京畿道

(72)发明人 李太熙 李朱嬿 黄志勋

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 张波

(51)Int.Cl.
H01L 27/11551(2017.01)
H01L 27/11529(2017.01)
H01L 27/11531(2017.01)
H01L 27/11573(2017.01)
H01L 27/11578(2017.01)

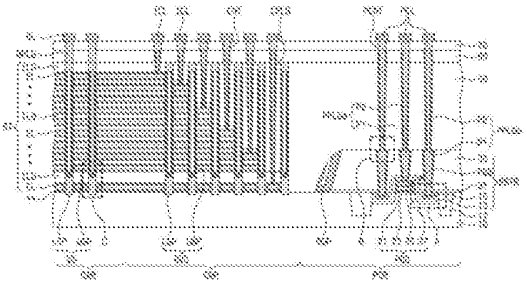
权利要求书3页 说明书14页 附图34页

(54)发明名称

三维半导体存储器件及制造其的方法

(57)摘要

提供了一种三维半导体存储器件及制造其的方法。该器件可以包括:衬底,其包括外围电路区域和单元阵列区域;电极结构,其包括垂直地堆叠在衬底的单元阵列区域上的多个电极;外围逻辑电路,其被提供在衬底的外围电路区域上,外围逻辑电路包括掺杂有第一杂质的第一杂质区域;外围接触插塞,其连接到第一杂质区域;以及第二杂质区域,其在第一杂质区域与外围接触插塞之间,第二杂质区域包括与第一杂质不同的第二杂质。外围接触插塞包括接触第二杂质区域的下部和从下部连续延伸的上部,下部和上部的每个的下宽度小于其上宽度,并且下部的上宽度大于上部的下宽度。



CN 108695336 A

1.一种三维半导体存储器件,包括:

衬底,其包括外围电路区域和单元阵列区域;

电极结构,其包括垂直地堆叠在所述衬底的所述单元阵列区域上的多个电极;

外围逻辑电路,其被提供在所述衬底的所述外围电路区域上,所述外围逻辑电路包括掺杂有第一杂质的第一杂质区域;

外围接触插塞,其连接到所述第一杂质区域;以及

第二杂质区域,其在所述第一杂质区域与所述外围接触插塞之间,所述第二杂质区域包括与所述第一杂质不同的第二杂质,

其中所述外围接触插塞包括接触所述第二杂质区域的下部以及从所述下部连续延伸的上部,

所述下部和所述上部的每个的下宽度小于其上宽度,以及

所述下部的所述上宽度大于所述上部的所述下宽度。

2.根据权利要求1所述的三维半导体存储器件,其中所述外围接触插塞具有与所述第一杂质区域间隔开的底表面。

3.根据权利要求1所述的三维半导体存储器件,其中所述外围接触插塞具有位于所述衬底的顶表面之下的底表面。

4.根据权利要求1所述的三维半导体存储器件,其中所述第二杂质区域具有比所述第一杂质区域的深度更小的深度。

5.根据权利要求1所述的三维半导体存储器件,其中所述第二杂质区域的宽度小于所述第一杂质区域的宽度。

6.根据权利要求1所述的三维半导体存储器件,其中所述第二杂质区域包含碳(C)、氮(N)和氟(F)中的至少一种。

7.根据权利要求1所述的三维半导体存储器件,其中所述外围接触插塞的所述下部的垂直长度小于其所述上部的垂直长度。

8.根据权利要求1所述的三维半导体存储器件,其中所述外围接触插塞的所述上部的顶表面在所述电极中的最上面的电极的顶表面之上。

9.根据权利要求1所述的三维半导体存储器件,还包括:

下绝缘层,其覆盖所述外围逻辑电路;以及

上绝缘层,其覆盖所述电极结构和所述下绝缘层,

其中所述上绝缘层的底表面部分地覆盖所述外围接触插塞的所述下部的顶表面。

10.根据权利要求9所述的三维半导体存储器件,还包括:

单元接触插塞,其穿透所述上绝缘层并且分别连接到所述电极结构的所述电极的端部,

其中所述外围接触插塞具有与所述单元接触插塞的顶表面共平面的顶表面。

11.根据权利要求1所述的三维半导体存储器件,还包括:

多个垂直结构,其穿透所述电极结构并且连接到所述衬底,

其中所述垂直结构的每个包括穿透所述电极结构的下部并连接到所述衬底的下半导体图案以及穿透所述电极结构的上部并连接到所述下半导体图案的上半导体图案,以及

所述下半导体图案的顶表面位于所述外围接触插塞的所述下部的顶表面之下。

Micron Ex. 1002, p. 720
Micron v. YMTC
IPR2024-00789

12.一种三维半导体存储器件，包括：

衬底，其包括外围电路区域和单元阵列区域；

电极结构，其包括垂直地堆叠在所述衬底的所述单元阵列区域上的多个电极；

外围逻辑电路，其被提供在所述衬底的所述外围电路区域上，所述外围逻辑电路包括外围栅极堆叠和源极/漏极杂质区域，所述外围栅极堆叠包括第一侧和第二侧，所述源极/漏极区域在所述外围栅极堆叠的两侧；以及

外围接触插塞，其分别连接到所述源极/漏极杂质区域，

其中所述源极/漏极杂质区域的每个包括，

第一杂质区域，其掺杂有第一杂质，以及

第二杂质区域，其包括与所述第一杂质不同的第二杂质，

其中所述外围接触插塞与所述源极/漏极杂质区域的所述第二杂质区域接触。

13.根据权利要求12所述的三维半导体存储器件，其中所述外围接触插塞的每个包括：

下部，其与所述第二杂质区域接触；以及

上部，其从所述下部连续地延伸，

其中所述下部和所述上部的每个具有比上宽度更小的下宽度，以及

所述下部的所述上宽度大于所述上部的所述下宽度。

14.根据权利要求13所述的三维半导体存储器件，其中所述外围接触插塞的所述下部和所述上部连续地延伸并且其间没有界面。

15.根据权利要求12所述的三维半导体存储器件，其中所述源极/漏极杂质区域的所述第二杂质区域用所述第一杂质和所述第二杂质共掺杂。

16.根据权利要求12所述的三维半导体存储器件，其中所述第二杂质区域具有比所述第一杂质区域的深度更小的深度。

17.根据权利要求12所述的三维半导体存储器件，其中所述第二杂质区域的宽度小于所述第一杂质区域的宽度。

18.根据权利要求13所述的三维半导体存储器件，还包括：

多个垂直结构，其穿透所述电极结构并且连接到所述衬底，

其中所述垂直结构的每个包括穿透所述电极结构的下部并连接到所述衬底的下半导体图案以及穿透所述电极结构的上部并连接到所述下半导体图案的上半导体图案，以及

所述下半导体图案的顶表面位于所述外围接触插塞的所述下部的顶表面之下。

19.一种制造三维半导体存储器件的方法，包括：

提供包括外围电路区域和单元阵列区域的衬底；

在所述衬底的所述外围电路区域上形成外围结构，所述外围结构包括外围栅极堆叠、在所述外围栅极堆叠的两侧的源极/漏极杂质区域、以及覆盖所述外围栅极堆叠和所述源极/漏极杂质区域的下绝缘层；

形成穿透所述下绝缘层并且分别与所述源极/漏极杂质区域接触的牺牲插塞；

形成与所述外围结构间隔开的电极结构，所述电极结构包括垂直地堆叠在所述衬底的所述单元阵列区域上的电极；

形成覆盖所述电极结构、所述外围结构和所述牺牲插塞的上绝缘层；

形成穿透所述上绝缘层并且分别暴露所述牺牲插塞的上接触孔；

Micron Ex. 1002, p. 721
Micron v. YMTC
IPR2024-00789

去除由所述上接触孔暴露的所述牺牲插塞;以及

在下接触孔和所述上接触孔中形成外围接触插塞,其中所述下接触孔穿透所述下绝缘层,所述外围接触插塞分别电连接到所述源极/漏极杂质区域。

20.根据权利要求19所述的方法,其中所述源极/漏极杂质区域的形成包括:

用第一杂质掺杂所述衬底以形成第一杂质区域;以及

用不同于所述第一杂质的第二杂质掺杂所述第一杂质区域以形成第二杂质区域,

其中所述牺牲插塞形成为与所述源极/漏极杂质区域的所述第二杂质区域接触。

21.根据权利要求20所述的方法,其中所述牺牲插塞的去除被执行,暴露所述源极/漏极杂质区域的所述第二杂质区域。

22.根据权利要求20所述的方法,其中所述外围结构的形成包括在所述下绝缘层中形成下接触孔,所述下接触孔的每个暴露所述第一杂质区域,

所述第二杂质区域在所述下接触孔的形成之后被形成。

23.根据权利要求20所述的方法,其中所述外围结构的形成包括在所述第二杂质区域的形成之后,在所述下绝缘层中形成下接触孔,所述下接触孔的每个暴露所述第二杂质区域。

24.根据权利要求19所述的方法,其中所述下接触孔和所述上接触孔的每个形成为具有在向下方向上减小的宽度。

25.根据权利要求19所述的方法,还包括:

形成穿过所述电极结构连接到所述衬底的垂直结构。

Micron Ex. 1002, p. 722
Micron v. YMTC
IPR2024-00789

三维半导体存储器件及制造其的方法

技术领域

[0001] 本公开涉及高度集成的三维半导体存储器件及制造其的方法。

背景技术

[0002] 期望半导体器件的更高集成度以满足消费者对性能和价格的要求。在半导体器件的情况下，因为其集成度是决定产品价格的重要因素，所以特别期望增大的集成度。在常规二维或平面半导体器件的情况下，因为其集成度主要由单位存储单元所占据的面积决定，所以集成度受到精细图案形成技术水平极大地影响。然而，增加图案精细度所需的昂贵的工艺设备为二维或平面半导体器件设定了集成复杂度的实际限制。为了克服这样的限制，近来已经提出了包括三维布置的存储单元的三维半导体存储器件。

发明内容

[0003] 发明构思的一些实施方式提供了具有改善的集成密度的三维半导体存储器件及制造其的方法。

[0004] 根据发明构思的一些实施方式，一种三维半导体存储器件可以包括：衬底，其包括外围电路区域和单元阵列区域；电极结构，其包括垂直地堆叠在衬底的单元阵列区域上的多个电极；外围逻辑电路，其被提供在衬底的外围电路区域上，外围逻辑电路包括掺杂有第一杂质的第一杂质区域；外围接触插塞，其连接到第一杂质区域；以及第二杂质区域，其在第一杂质区域与外围接触插塞之间，第二杂质区域包括与第一杂质不同的第二杂质。外围接触插塞包括接触第二杂质区域的下部和从下部连续延伸的上部，下部和上部的每个的下宽度小于其上宽度，并且下部的上宽度大于上部的下宽度。

[0005] 根据发明构思的一些实施方式，一种三维半导体存储器件可以包括：衬底，其包括外围电路区域和单元阵列区域；电极结构，其包括垂直地堆叠在衬底的单元阵列区域上的多个电极；外围逻辑电路，其被提供在衬底的外围电路区域上，外围逻辑电路包括外围栅极堆叠和源极/漏极杂质区域，外围栅极堆叠包括第一侧和第二侧，源极/漏极区域在外围栅极堆叠的两侧；以及外围接触插塞，其分别连接到源极/漏极杂质区域。源极/漏极杂质区域的每个包括掺杂有第一杂质的第一杂质区域以及包括与第一杂质不同的第二杂质的第二杂质区域。外围接触插塞与源极/漏极杂质区域的第二杂质区域接触。

[0006] 根据发明构思的一些实施方式，一种制造三维半导体存储器件的方法可以包括：提供包括外围电路区域和单元阵列区域的衬底；在衬底的外围电路区域上形成外围结构，外围结构包括外围栅极堆叠、在外围栅极堆叠的两侧的源极/漏极杂质区域、以及覆盖外围栅极堆叠和源极/漏极杂质区域的下绝缘层；形成穿透下绝缘层并分别与源极/漏极杂质区域接触的牺牲插塞；形成与外围结构间隔开的电极结构，电极结构包括垂直地堆叠在衬底的单元阵列区域上的电极；形成覆盖电极结构、外围结构和牺牲插塞的上绝缘层；形成穿透上绝缘层并分别暴露牺牲插塞的上接触孔；去除由上接触孔暴露的牺牲插塞；以及在下接触孔和上接触孔中形成外围接触插塞，外围接触插塞分别电连接到源极/漏极杂质区域。

Micron Ex. 1002, p. 723
Micron v. YMTC
IPR2024-00789

附图说明

[0007] 示例实施方式将由以下结合附图的简要描述被更清楚地理解。附图表示如在此所述的非限制性的示例实施方式。

[0008] 发明构思可以涉及三维布置的存储单元。例如，发明构思可以涉及如美国专利第7,679,133号、第8,553,466号、第8,654,587号、第8,559,235号以及美国专利公开第2011/0233648号中公开的三维布置的存储单元，其每个的全部内容通过引用在此合并。

[0009] 图1是根据本发明构思的一些实施方式的三维半导体存储器件的俯视图。

[0010] 图2A至图2J是沿图1的线I-I'截取的剖视图，以示出根据本发明构思的一些实施方式的制造三维半导体存储器件的方法。

[0011] 图3A至图3C是示出根据本发明构思的一些实施方式的三维半导体存储器件的一部分(例如图2J的部分“A”)的放大剖视图。

[0012] 图4A和图4B是示出根据本发明构思的一些实施方式的三维半导体存储器件的一部分(例如图2J的部分“B”)的放大剖视图。

[0013] 图5是示出根据本发明构思的一些实施方式的三维半导体存储器件的一部分(例如图2J的部分“C”)的放大剖视图。

[0014] 图6A和图6B是示出根据本发明构思的各种实施方式的三维半导体存储器件的外围电路区域的俯视图。

[0015] 图7A至图7H是示出根据本发明构思的各种实施方式的制造三维半导体存储器件的方法的剖视图。

[0016] 图8A至图8I是示出根据本发明构思的各种实施方式的制造三维半导体存储器件的方法的剖视图。

[0017] 图9是根据本发明构思的各种实施方式的三维半导体存储器件的剖视图。

[0018] 应注意，这些附图旨在示出某些示例实施方式中利用的方法、结构和/或材料的一般特性，并且补充下面提供的书面描述。然而，这些附图不是按比例绘制的，并且可以不精确地反映任何给定实施方式的精确结构或性能特征，并且不应被解释为限定或限制由示例实施方式涵盖的值或性质的范围。例如，为了清楚，分子、层、区域和/或结构元件的相对厚度和位置可以被减小或夸大。相似或相同的附图标记在各种各样的附图中的使用旨在表示相似或相同的元件或特征的存在。

具体实施方式

[0019] 现在将参照其中示出示例实施方式的附图更全面地描述本发明构思的示例实施方式。

[0020] 图1是根据本发明构思的一些实施方式的三维半导体存储器件的俯视图。图2A至图2J是沿图1的线I-I'截取的剖视图，以示出根据本发明构思的一些实施方式的制造三维半导体存储器件的方法。

[0021] 参照图1和图2A，衬底10可以包括单元阵列区域CAR、连接区域CNR和外围电路区域PCR。连接区域CNR可以位于单元阵列区域CAR与外围电路区域PCR之间。

[0022] 衬底10可以由具有半导体性质的材料(例如硅晶片)、绝缘材料(例如玻璃衬底)、

Micron Ex. 1002, p. 724
Micron v. YMTC
IPR2024-00789

或者用绝缘材料覆盖的半导体或导体材料形成，或者包括具有半导体性质的材料(例如硅晶片)、绝缘材料(例如玻璃衬底)、或者用绝缘材料覆盖的半导体或导体材料。例如，衬底10可以是或者可以包括具有第一导电性的硅晶片。

[0023] 用于向存储单元写入数据或从存储单元读取数据的外围逻辑电路可以是例如可以形成在衬底10的外围电路区域PCR上。外围逻辑电路可以包括行解码器和列解码器、页缓冲器和/或控制电路。例如，外围逻辑电路可以包括电连接到存储单元的NMOS和PMOS晶体管、电阻器和电容器。

[0024] 例如，器件隔离层12可以形成在衬底10的外围电路区域PCR中以限定有源区域ACT。外围栅极堆叠PGS可以在衬底10的外围电路区域PCR上形成为交叉有源区域ACT。

[0025] 外围栅极堆叠PGS可以包括顺序地堆叠在衬底10上的外围栅极绝缘层21、掺杂多晶硅层23、栅极金属层25和硬掩模层27。外围栅极绝缘层21可以是或者可以包括硅氧化物层。硅氧化物层可以通过热氧化工艺形成。硅氧化物层可以通过原位水汽产生工艺被形成。间隔物可以形成在外围栅极堆叠PGS的两个侧表面上，并且源极/漏极杂质区域13可以通过在外围栅极堆叠PGS的两侧用第一杂质掺杂有源区域ACT而形成。第一杂质可以是从周期表的III族中选择的元素，例如硼。第一杂质可以是从周期表的V族中选择的元素，例如磷或砷。源极/漏极杂质区域13可以用离子注入工艺形成。例如，源极/漏极杂质区域13可以用高电流离子注入工艺形成。例如，源极/漏极杂质区域13可以用等离子体辅助掺杂工艺形成。然而，发明构思不限于此。

[0026] 在外围逻辑电路的形成之后，蚀刻停止层31和外围绝缘层33可以在衬底10上形成，例如顺序地形成。蚀刻停止层31可以被沉积以共形地覆盖外围逻辑电路。外围绝缘层33可以包括多个绝缘层(例如包括硅氧化物层、硅氮化物层、硅氮氧化物层和低k电介质层中的至少一个)。

[0027] 参照图1和图2B，第一下接触孔33a可以被形成，暴露外围栅极堆叠PGS，并且第二下接触孔33b可以被形成，分别暴露源极/漏极杂质区域13。

[0028] 第一下接触孔33a和第二下接触孔33b可以通过在外围绝缘层33上形成掩模图案(未示出)然后蚀刻例如各向异性地蚀刻外围绝缘层33而形成。第一下接触孔33a可以形成为暴露外围栅极堆叠PGS的栅极金属层25，第二下接触孔33b可以形成为部分地暴露源极/漏极杂质区域13。

[0029] 第一下接触孔33a和第二下接触孔33b的每个可以具有比上宽度更小的下宽度。当第二下接触孔33b被形成时，衬底10的顶表面可以通过各向异性蚀刻工艺而被凹入。

[0030] 在一些实施方式中，在第一下接触孔33a和第二下接触孔33b的形成之后，虚设杂质区域15可以通过用第二杂质掺杂源极/漏极杂质区域13而形成。第二杂质可以是与第一杂质的种类不同的种类的杂质，并且可以包括例如碳(C)、氮(N)或氟(F)。

[0031] 虚设杂质区域15可以用第二杂质和第一杂质共掺杂。虚设杂质区域15中的第二杂质的浓度可以小于源极/漏极杂质区域13中的第一杂质的浓度。例如，虚设杂质区域15中的第二杂质的浓度可以小一个数量级、或几个数量级。

[0032] 虚设杂质区域15的形成可以包括将第二杂质注入到衬底10的由第二下接触孔33b暴露的部分中。如图3A中所示，虚设杂质区域15可以形成为具有比源极/漏极杂质区域13的深度更浅的深度。换言之，虚设杂质区域15可以形成在第二下接触孔33b周围。

Micron Ex. 1002, p. 725
Micron v. YMTC
IPR2024-00789

[0033] 在某些实施方式中，如图3B中所示，虚设杂质区域15可以在源极/漏极杂质区域13的形成之后并且在外围绝缘层33的形成之前被形成。虚设杂质区域15可以通过使用离子注入工艺将第一杂质注入到衬底10中、形成源极/漏极杂质区域13、然后使用与用于源极/漏极杂质区域13的离子注入掩模相同的离子注入掩模将第二杂质浅注入到源极/漏极杂质区域13中而被形成。第二杂质可以用中等电流或低电流注入被注入到虚设杂质区域15中；然而，发明构思不限于此。

[0034] 或者，虚设杂质区域15可以与源极/漏极杂质区域13一起被形成。例如，当源极/漏极杂质区域13形成时，第一杂质和第二杂质可以被一起注入到衬底10中。

[0035] 参照图1和图2C，牺牲插塞层41和牺牲间隙填充层43可以在具有第一下接触孔33a和第二下接触孔33b的外围绝缘层33上顺序地形成至均匀的厚度。

[0036] 牺牲插塞层41可以由相对于外围绝缘层33具有蚀刻选择性的材料形成，牺牲间隙填充层43可以由相对于牺牲插塞层41具有蚀刻选择性的材料形成。牺牲插塞层41可以由例如多晶硅层、硅层、锗层或硅锗层形成，或者包括例如多晶硅层、硅层、锗层或硅锗层。牺牲间隙填充层43可以由例如硅氧化物层形成，或者包括例如硅氧化物层。

[0037] 牺牲插塞层41可以使用具有良好台阶覆盖性质的沉积工艺(例如使用化学气相沉积(CVD)或原子层沉积(ALD)工艺)被沉积。这里，牺牲插塞层41的厚度可以比第一下接触孔33a和第二下接触孔33b的上宽度的大约一半更小。因此，牺牲插塞层41可以均匀地覆盖第一下接触孔33a和第二下接触孔33b的内表面，并且可以部分地填充第一下接触孔33a和第二下接触孔33b。在一些实施方式中，牺牲插塞层41可以与虚设杂质区域15直接接触。在虚设杂质区域15不被形成的情况下，牺牲插塞层41可以与源极/漏极杂质区域13直接接触。

[0038] 牺牲间隙填充层43可以在牺牲插塞层41上沉积。在一些实施方式中，牺牲间隙填充层43可以形成为填充用牺牲插塞层41覆盖的第一下接触孔33a和第二下接触孔33b。

[0039] 参照图1和图2D，在牺牲插塞层41和牺牲间隙填充层43的形成之后，平坦化工艺可以被执行，暴露外围绝缘层33的顶表面。因此，牺牲插塞42和牺牲间隙填充图案44可以形成在第一下接触孔33a和第二下接触孔33b的每个中。

[0040] 此后，外围绝缘层33可以被图案化以形成暴露衬底10的单元阵列区域CAR和连接区域CNR的外围绝缘图案35。作为外围绝缘图案35的形成的结果，外围逻辑结构PSTR可以形成在衬底10的外围电路区域PCR上，并且这里，外围逻辑结构PSTR可以包括外围栅极堆叠PGS、源极/漏极杂质区域13和外围绝缘图案35。

[0041] 参照图1和图2E，在外围绝缘图案35的形成之后，模制结构110可以形成在衬底10的单元阵列区域CAR和连接区域CNR上。模制结构110可以包括堆叠例如交替地堆叠在衬底10上的牺牲层SL和绝缘层ILD。

[0042] 在模制结构110中，牺牲层SL可以由能相对于绝缘层ILD以高蚀刻选择性被蚀刻的材料形成，或者包括能相对于绝缘层ILD以高蚀刻选择性被蚀刻的材料。作为示例，牺牲层SL可以由与绝缘层ILD不同的绝缘材料形成。例如，牺牲层SL可以由硅氮化物层形成，绝缘层ILD可以由硅氧化物层形成。牺牲层SL可以具有基本相同的厚度，并且绝缘层ILD中的至少一个可以具有与其它绝缘层ILD不同的厚度。

[0043] 例如，模制结构110的形成可以包括在衬底10上形成其中交替地堆叠牺牲层SL和绝缘层ILD的层状结构，并且对层状结构执行修整工艺。这里，修整工艺可以包括以下步骤：

Micron Ex. 1002, p. 726
Micron v. YMTC
IPR2024-00789

在单元阵列区域CAR和连接区域CNR上形成覆盖层状结构的掩模图案(未示出)、使用掩模图案作为蚀刻掩模蚀刻层状结构、蚀刻掩模图案以减小掩模图案的平面面积、以及去除掩模图案。在去除掩模图案的步骤之前,蚀刻层状结构和掩模图案的步骤可以被重复若干次。

[0044] 作为修整工艺的结果,模制结构110可以从单元阵列区域CAR延伸到连接区域CNR,并且可以在连接区域CNR上具有阶梯结构。例如,模制结构110可以具有其高度在朝向外围电路区域PCR的方向上以台阶状方式减小的阶梯结构。模制结构110可以具有比外围结构的垂直高度更大的垂直高度。例如,模制结构110的垂直高度可以大于或等于外围结构的高度的大约2倍。

[0045] 此外,在形成模制结构110的修整工艺期间,虚设间隔物DSP可以形成在外围绝缘图案35的侧表面上。虚设间隔物DSP可以是或者可以包括牺牲层SL和绝缘层ILD的未被各向异性蚀刻工艺蚀刻的剩余部分。

[0046] 参照图1和图2F,上平坦化绝缘层50可以形成为覆盖提供有模制结构110的衬底10。上平坦化绝缘层50可以延伸为不仅覆盖模制结构110而且覆盖外围结构,并且可以具有基本上平坦的顶表面。上平坦化绝缘层50可以由相对于牺牲层SL具有蚀刻选择性的材料形成。

[0047] 在上平坦化绝缘层50的形成之后,垂直结构VS可以形成在单元阵列区域CAR上,穿透模制结构110。当在俯视图中被观察时,垂直结构VS可以在特定方向上或者以Z字形形状布置。

[0048] 此外,在垂直结构VS的形成期间,虚设垂直结构DVS可以在连接区域CNR上形成为部分地穿透模制结构110。虚设垂直结构DVS可以具有与垂直结构VS基本相同的结构,并且可以形成为穿透牺牲层SL的端部。

[0049] 垂直结构VS和虚设垂直结构DVS的形成可以包括形成穿透模制结构110并暴露衬底10的垂直孔、以及在垂直孔的每个中形成下半导体图案LSP和上半导体图案USP。这里,用于下半导体图案LSP和上半导体图案USP的半导体材料可以具有彼此不同的晶体结构。在一些实施方式中,下半导体图案LSP可以具有柱形状,并且下半导体图案LSP的顶表面可以位于外围电路区域PCR上的牺牲插塞42的顶表面之下。此外,导电垫D可以形成在上半导体图案USP的每个的顶部中。导电垫D可以是掺杂有杂质的杂质区域,或者可以由导电材料形成。将参照图5更详细地描述垂直结构VS和虚设垂直结构DVS。

[0050] 参照图1和图2G,第一层间绝缘层60可以形成在上平坦化绝缘层50上并且可以覆盖垂直结构VS和虚设垂直结构DVS的顶表面。

[0051] 在第一层间绝缘层60的形成之后,替换工艺可以用电极EL替换牺牲层SL被执行。作为替换工艺的结果,包括交替地堆叠在衬底10上的电极EL和绝缘层ILD的电极结构ST可以被形成。电极结构ST可以在连接区域CNR上具有阶梯结构。

[0052] 替换工艺可以包括在单元阵列区域CAR和连接区域CNR上形成穿透第一层间绝缘层60、上平坦化绝缘层50和模制结构110并暴露衬底10的沟槽。替换工艺可以包括去除由沟槽暴露的牺牲层SL以在绝缘层ILD之间形成栅极区域、以及分别在栅极区域中形成电极EL。

[0053] 这里,沟槽可以在第一方向D1上延伸并且可以在交叉第一方向D1的第二方向D2上彼此间隔开。在一些实施方式中,沟槽可以具有至少两个不同的长度,并且可以形成为允许模制结构110在俯视图中具有例如基本上“H”形的结构。沟槽可以与垂直结构VS间隔开,并

Micron Ex. 1002, p. 727
Micron v. YMTC
IPR2024-00789

且可以形成为暴露牺牲层SL和绝缘层ILD的侧表面。

[0054] 栅极区域的形成可以包括使用选择为相对于上平坦化绝缘层50、绝缘层ILD、垂直结构VS和衬底10具有蚀刻选择性的蚀刻配方来各向同性地蚀刻牺牲层SL。

[0055] 电极EL的形成可以包括在具有栅极区域的模制结构110上顺序地沉积阻挡金属层和金属层、然后各向异性地蚀刻沉积在沟槽的内表面上的阻挡金属层和金属层。阻挡金属层可以由金属氮化物层(例如TiN、TaN或WN)形成。金属层可以由金属材料(例如W、Al、Ti、Ta、Co或Cu)形成。

[0056] 在一些实施方式中,在电极EL的形成之前,水平绝缘图案HP可以被形成以共形地覆盖栅极区域的内表面,如图5中所示。水平绝缘图案HP可以用作NAND闪速存储晶体管的数据存储层的一部分。水平绝缘图案HP可以是或者可以包括高k电介质材料(例如铝氧化物和铪氧化物)中的一种。此外,在水平绝缘图案HP的形成之前,热氧化物层可以形成在下半导体图案LSP的侧表面上。

[0057] 此外,公共源极区域CSR可以形成在由沟槽暴露的衬底10中。公共源极区域CSR可以在第一方向D1上延伸并彼此平行,并且可以在第二方向D2上彼此间隔开。公共源极区域CSR可以通过用与衬底10的类型不同的类型的杂质掺杂衬底10而形成。公共源极区域CSR可以包含n型杂质(例如砷(As)或磷(P))。

[0058] 接着,参照图1和图2G,在电极结构ST的形成之后,第一层间绝缘层60和上平坦化绝缘层50可以被图案化以在连接区域CNR上形成单元接触孔50c,并且在外围电路区域PCR上形成上接触孔50a和50b。

[0059] 单元接触孔50c以及上接触孔50a和50b的形成可以包括在第一层间绝缘层60上形成掩模图案(未示出)、以及各向异性地蚀刻第一层间绝缘层60和上平坦化绝缘层50。在一些实施方式中,单元接触孔50c以及上接触孔50a和50b可以同时形成,但发明构思不限于此。在某些实施方式中,单元接触孔50c中的一些可以被形成,然后,上接触孔50a和50b可以与单元接触孔50c中的其它单元接触孔一起形成。在某些实施方式中,上接触孔50a和50b可以与接触孔50c中的一些一起形成,然后接触孔50c中的其它接触孔可以被形成。

[0060] 在一些实施方式中,单元接触孔50c可以分别形成为在连接区域CNR上暴露电极EL的端部。单元接触孔50c可以具有彼此不同的垂直长度,其中垂直长度是在垂直于衬底10的顶表面的方向上测量的长度。

[0061] 在一些实施方式中,当上接触孔50a和50b被形成时,上平坦化绝缘层50可以包括与牺牲间隙填充图案44相同的材料,并且可以与牺牲间隙填充图案44一起被蚀刻。在这种情况下,上接触孔50a和50b可以暴露牺牲插塞42。

[0062] 在一些实施方式中,上接触孔50a和50b可以具有小于单元接触孔50c的垂直深度的最大值的垂直深度。此外,上接触孔50a和50b的垂直深度可以大于下接触孔(例如参见图2B的33a和33b)的垂直深度。例如,上接触孔50a和50b的高宽比可以大于下接触孔(例如参见图2B的33a和33b)的高宽比。即使当各向异性蚀刻工艺被用于形成上接触孔50a和50b时,上接触孔50a和50b的每个可以具有比下宽度更大的上宽度。上接触孔50a和50b的下宽度可以小于下接触孔33a和33b的上宽度。此外,上接触孔50a和50b的每个的下宽度可以大于牺牲间隙填充图案44的上宽度。

[0063] 参照图1和图2H,由上接触孔50a和50b暴露的牺牲插塞42可以被去除,暴露下接触

Micron Ex. 1002, p. 728
Micron v. YMTC
IPR2024-00789

孔33a和33b的内表面。在牺牲插塞42被去除的情况下，下接触孔33a和33b可以分别被连接到上接触孔50a和50b。

[0064] 在一些实施方式中，牺牲插塞42的去除可以包括使用选择为相对于外围绝缘图案35具有蚀刻选择性的蚀刻配方各向异性或各向同性地蚀刻牺牲插塞42。

[0065] 作为牺牲插塞42的去除的结果，外围栅极堆叠PGS的金属层可以通过第一下接触孔33a被暴露，并且虚设杂质区域15可以通过第二下接触孔33b被暴露。虚设杂质区域15可以防止源极/漏极杂质区域13被用于去除牺牲插塞42的蚀刻气体或蚀刻剂损坏，或者帮助减小源极/漏极杂质区域13被用于去除牺牲插塞42的蚀刻气体或蚀刻剂损坏的可能性。例如，牺牲插塞42可以由多晶硅层形成，或者包括多晶硅层，并且虚设杂质区域15可以抑制并防止衬底10中的源极/漏极杂质区域13在蚀刻牺牲插塞42时被不必要地蚀刻，或者帮助减小衬底10中的源极/漏极杂质区域13在蚀刻牺牲插塞42时被不必要地蚀刻的可能性。

[0066] 参照图1和图2I，单元接触插塞CPLG可以形成在单元接触孔50c中和在连接区域CNR上，并且外围接触插塞PPLGa和PPLGb可以形成在下接触孔33a、33b和上接触孔50a、50b中以及在外围电路区域PCR上。

[0067] 单元接触插塞CPLG以及外围接触插塞PPLGa和PPLGb的形成可以包括在上接触孔50a、50b和下接触孔33a、33b以及单元接触孔50c中顺序地沉积阻挡金属层和金属层、然后执行平坦化工艺暴露第一层间绝缘层60的顶表面。这里，阻挡金属层可以由至少一种金属氮化物(例如TiN、TaN或WN)形成，或者包括至少一种金属氮化物(例如TiN、TaN或WN)。金属层可以由至少一种金属材料(例如W、Al、Ti、Ta、Co或Cu)形成，或者包括至少一种金属材料(例如W、Al、Ti、Ta、Co或Cu)。在一些示例实施方式中，单元接触插塞CPLG以及外围接触插塞PPLGa和PPLGb可以同时形成，并且在这种情况下，单元接触插塞CPLG可以具有与外围接触插塞PPLGa和PPLGb的顶表面基本共平面的顶表面。此外，因为上接触孔50a、50b和下接触孔33a、33b用金属层一次填充，所以外围接触插塞PPLGa和PPLGb的每个可以穿透第一层间绝缘层60、上平坦化绝缘层50和外围绝缘图案35而没有界面。

[0068] 单元接触插塞CPLG可以形成为穿透第一层间绝缘层60和上平坦化绝缘层50，并且可以分别联接到电极EL的端部。单元接触插塞CPLG可以形成为具有在朝向单元阵列区域CAR的方向上减小的垂直长度。

[0069] 外围接触插塞PPLGa和PPLGb可以包括与外围栅极堆叠PGS的栅极金属层25接触的第一外围接触插塞PPLGa、以及电连接到源极/漏极杂质区域13的第二外围接触插塞PPLGb。这里，第二外围接触插塞PPLGb的底表面可以位于衬底10的顶表面之下并且可以与虚设杂质区域15接触。

[0070] 第一外围接触插塞PPLGa和第二外围接触插塞PPLGb可以或者可以被提供为连续地穿透第一层间绝缘层60、上平坦化绝缘层50和外围绝缘图案35，并且可以在外围绝缘图案35与上平坦化绝缘层50之间的界面附近具有变化的宽度。例如，第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个可以包括穿透外围绝缘图案35的下部P1以及从下部P1连续地延伸的穿透上平坦化绝缘层50和第一层间绝缘层60的上部P2。当从衬底10的顶表面垂直地测量时，上部P2的垂直长度可以大于下部P1的垂直长度。下部P1和上部P2的每个可以在向下方向上具有减小的宽度，并且下部P1的上宽度可以大于上部P2的下宽度。

[0071] 参照图1和图2J，第二层间绝缘层70可以形成在覆盖单元接触插塞CPLG以及第一

Micron Ex. 1002, p. 729
Micron v. YMTC
IPR2024-00789

外围接触插塞PPLGa和第二外围接触插塞PPLGb的第一层间绝缘层60上。

[0072] 位线接触插塞BPLG、连接接触插塞CNT和外围连接接触插塞PCNT可以形成在第二层间绝缘层70中。

[0073] 位线接触插塞BPLG可以被提供为在单元阵列区域CAR上穿透第二层间绝缘层70，并且可以分别联接到垂直结构VS。连接接触插塞CNT可以在连接区域CNR上穿透第二层间绝缘层70，并且可以分别联接到单元接触插塞CPLG。外围连接接触插塞PCNT可以被提供为在外围电路区域PCR上穿透第二层间绝缘层70，并且可以分别联接到第一外围接触插塞PPLGa和第二外围接触插塞PPLGb。

[0074] 接着，位线BL、互连线ICL和外围连接线PCL可以形成在第二层间绝缘层70上。位线BL可以在单元阵列区域CAR上在第二方向D2上延伸并且可以连接到位线接触插塞BPLG。互连线ICL可以被提供在连接区域CNR上并且可以连接到连接接触插塞CNT。外围连接线PCL可以被提供在外围电路区域PCR上并且可以连接到外围连接接触插塞PCNT。

[0075] 图3A至图3C是示出根据本发明构思的一些实施方式的三维半导体存储器件的一部分(例如图2J的部分“A”)的放大剖视图。

[0076] 参照图3A，虚设杂质区域15可以局部地形成在源极/漏极杂质区域13中。例如，虚设杂质区域15可以具有比源极/漏极杂质区域13的宽度更小的宽度。第二外围接触插塞PPLGb可以具有与虚设杂质区域15直接接触的底表面，并且与源极/漏极杂质区域13间隔开。虚设杂质区域15可以包围第二外围接触插塞PPLGb的插入到衬底10中的部分。

[0077] 参照图3B，虚设杂质区域15可以形成在源极/漏极杂质区域13的上部区域中。例如，虚设杂质区域15可以具有比源极/漏极杂质区域13的深度更小的深度。第二外围接触插塞PPLGb可以具有位于虚设杂质区域15中并且与源极/漏极杂质区域13的底表面间隔开的底表面。

[0078] 在图3C的实施方式中，虚设杂质区域15可以被省略。在这种情况下，第二外围接触插塞PPLGb可以与源极/漏极杂质区域13直接接触。

[0079] 在某些实施方式中，第二外围接触插塞PPLGb可以与源极/漏极杂质区域13直接接触，并且这里，源极/漏极杂质区域13可以用第一杂质(例如硼(B)和磷(P)中的一种)和第二杂质(例如碳(C)、氮(N)和氟(F)中的至少一种)共掺杂。

[0080] 图4A和图4B是示出根据本发明构思的一些实施方式的三维半导体存储器件的一部分(例如图2J的部分“B”)的放大剖视图。

[0081] 参照图4A和图4B，第二外围接触插塞PPLGb的下部P1和上部P2的每个可以具有在向下方向上减小的宽度。例如，在第二外围接触插塞PPLGb中，下部P1的上宽度W1可以大于上部P2的下宽度W2。因此，第二外围接触插塞PPLGb可以在外围绝缘图案35与上平坦化绝缘层50之间的水平处具有转折点。此外，上平坦化绝缘层50可以具有直接覆盖第二外围接触插塞PPLGb的下部P1的顶表面的一部分的底表面。

[0082] 如图4A中所示，当在俯视图中被观察时，第二外围接触插塞PPLGb的上部P2可以与下部P1对准。在某些实施方式中，如图4B中所示，第二外围接触插塞PPLGb的上部P2可以从下部P1偏移。换言之，第二外围接触插塞PPLGb的上部P2可以从下部P1连续地延伸，但是上部P2的中心可以与下部P1的中心不对准。

[0083] 图5是示出根据本发明构思的一些实施方式的三维半导体存储器件的一部分(例

Micron Ex. 1002, p. 730
Micron v. YMTC
IPR2024-00789

如图2J的部分“C”)的放大剖视图。

[0084] 参照图5,垂直结构VS的每个可以包括下半导体图案LSP和上半导体图案USP。

[0085] 下半导体图案LSP可以通过其中由垂直孔暴露的衬底10用作籽晶层的选择性外延生长(SEG)工艺被形成。因此,下半导体图案LSP可以是填充垂直孔的下部区域的柱形结构。下半导体图案LSP可以形成为具有单晶或多晶结构,但本发明构思可以不限于此。下半导体图案LSP可以由例如碳纳米结构、有机半导体材料和/或化合物半导体材料形成。

[0086] 上半导体图案USP可以形成在提供有下半导体图案LSP的垂直孔中。上半导体图案USP可以与下半导体图案LSP接触。

[0087] 详细地,如图5中所示,上半导体图案USP可以包括第一半导体图案SP1和第二半导体图案SP2。第一半导体图案SP1可以联接到下半导体图案LSP,并且可以具有拥有闭合的底部和敞开的顶部的管形状或通心粉形状。第一半导体图案SP1的内部空间可以用绝缘间隙填充图案VI填充。此外,第一半导体图案SP1可以与第二半导体图案SP2的内表面和下半导体图案LSP的顶表面接触。例如,第一半导体图案SP1可以允许第二半导体图案SP2电连接到下半导体图案LSP。

[0088] 上半导体图案USP可以由掺杂或本征半导体材料(例如硅(Si)、锗(Ge)或其化合物)中的至少一种形成,或者包括掺杂或本征半导体材料(例如硅(Si)、锗(Ge)或其化合物)中的至少一种。此外,上半导体图案USP可以具有单晶、非晶和多晶结构中的一种。

[0089] 在一些实施方式中,在上半导体图案USP的形成之前,垂直绝缘图案VP可以形成在垂直孔的每个中,如图5中所示。换言之,垂直绝缘图案VP可以从电极EL与垂直结构VS之间的区域垂直地延伸到绝缘层ILD与垂直结构VS之间的其它区域。此外,水平绝缘图案HP可以从电极EL与阻挡绝缘层BLK之间的区域朝电极EL的顶表面或底表面水平地延伸到其它区域。

[0090] 垂直绝缘图案VP可以包括一个或更多个层。在一些实施方式中,垂直绝缘图案VP可以用作NAND闪速存储器件的存储元件,并且可以包括隧道绝缘层TIL、电荷存储层CIL和阻挡绝缘层BLK中的至少一些。例如,电荷存储层CIL可以是或者可以包括俘获绝缘层、浮置栅电极或具有导电纳米点的绝缘层。详细地,电荷存储层CIL可以包括硅氮化物层、硅氮氧化物层、富硅氮化物层、纳米晶体硅层和层叠俘获层中的至少一种。隧道绝缘层TIL可以由其带隙大于电荷存储层CIL的带隙的材料中的至少一种形成,并且阻挡绝缘层BLK可以由高k电介质材料(例如铝氧化物和铪氧化物)形成。

[0091] 在某些实施方式中,垂直绝缘图案VP可以用作相变存储器件或可变电阻存储器的存储元件,并且可以包括相变层或可变电阻层。

[0092] 图6A和图6B是示出根据本发明构思的各种实施方式的三维半导体存储器件的外围电路区域的俯视图。

[0093] 为了简要描述,先前参照图2A至图2J描述的元件可以由相似或相同的附图标记标识,而不重复其重叠描述。

[0094] 参照图6A和图6B,如上所述,外围电路区域PCR上的第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个可以包括下部P1和上部P2。

[0095] 在图6A的实施方式中,在第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个中,下部P1的顶表面可以具有其纵轴平行于特定方向的椭圆形状。因此,在第一外围接触

Micron Ex. 1002, p. 731
Micron v. YMTC
IPR2024-00789

插塞PPLGa和第二外围接触插塞PPLGb的每个中，可以防止下部P1与上部P2不对准，或者减小下部P1与上部P2不对准的可能性。类似地，可以防止第一外围接触插塞PPLGa和第二外围接触插塞PPLGb在外围绝缘图案35与上平坦化绝缘层50之间被切割，或者减小第一外围接触插塞PPLGa和第二外围接触插塞PPLGb在外围绝缘图案35与上平坦化绝缘层50之间被切割的可能性。

[0096] 在第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个中，下部P1的顶表面可以具有多边形形状，如图6B中所示。例如，在第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个中，下部P1的顶表面可以包括在第一方向上延伸的第一部分和在第二方向上延伸的第二部分。

[0097] 图7A至图7H是示出根据本发明构思的各种实施方式的制造三维半导体存储器件的方法的剖视图。

[0098] 为了简要描述，先前参照图2A至图2J描述的元件或步骤可以由相似或相同的附图标记标识，而不重复其重叠描述。

[0099] 参照图7A，缓冲绝缘层11和下牺牲层可以形成例如顺序地形成为覆盖其上形成外围栅极堆叠PGS和源极/漏极杂质区域13的衬底10。

[0100] 下牺牲层可以共形地形成为覆盖衬底10的顶表面和外围栅极堆叠PGS。下牺牲层可以由相对于缓冲绝缘层11具有蚀刻选择性的材料形成。缓冲绝缘层11可以从下牺牲层与衬底10之间的区域延伸，并且可以在下牺牲层与外围栅极堆叠PGS之间。缓冲绝缘层11可以是或者可以包括硅氧化物层并且可以通过使用热氧化工艺或沉积工艺而形成。

[0101] 在一些实施方式中，下牺牲层可以在外围电路区域PCR与连接区域CNR之间被部分地蚀刻。因此，下牺牲图案LSL可以形成在单元阵列区域CAR和连接区域CNR上，并且虚设牺牲图案DSL可以形成在外围电路区域PCR上。

[0102] 接着，下平坦化绝缘层20可以形成在衬底10上。下平坦化绝缘层20可以被沉积为在下牺牲图案LSL和虚设牺牲图案DSL上具有均匀的厚度。下平坦化绝缘层20可以具有通过平坦化工艺被平坦化的顶表面，并且可以与虚设牺牲图案DSL的最高顶表面基本上共平面。

[0103] 参照图7B，下接触孔33a和33b可以被形成，穿透下平坦化绝缘层20和虚设牺牲图案DSL。下接触孔33a和33b可以包括分别形成为暴露外围栅极堆叠PGS的第一下接触孔33a以及形成为暴露源极/漏极杂质区域13的第二下接触孔33b。

[0104] 在下接触孔33a和33b的形成之后，虚设杂质区域15可以通过用第二杂质(例如碳(C)、氮(N)和氟(F)中的至少一种)掺杂源极/漏极杂质区域13而形成，如先前参照图2B所述。

[0105] 参照图7C，牺牲插塞42和牺牲间隙填充图案44可以形成在下接触孔33a和33b的每个中。

[0106] 在一些实施方式中，牺牲插塞42和牺牲间隙填充图案44的顶表面可以与下平坦化绝缘层20的顶表面基本上共平面。牺牲插塞42可以由相对于下平坦化绝缘层20和虚设牺牲图案DSL具有蚀刻选择性的材料形成。

[0107] 参照图7D，其中交替地堆叠牺牲层SL和绝缘层ILD的模制结构110可以形成在下平坦化绝缘层20上。

[0108] 模制结构110可以在连接区域CNR上的下平坦化绝缘层20上具有阶梯结构。当模制

Micron Ex. 1002, p. 732
Micron v. YMTC
IPR2024-00789

结构110被形成时，外围电路区域PCR上的虚设牺牲图案DSL可以被暴露，或者外围电路区域PCR上的牺牲插塞42的顶表面可以被暴露。

[0109] 参照图7E，在模制结构110的形成之后，上平坦化绝缘层50可以形成在衬底10上。上平坦化绝缘层50可以从单元阵列区域CAR延伸到外围电路区域PCR，并且可以具有基本上平坦的顶表面。

[0110] 在上平坦化绝缘层50的形成之后，垂直结构VS和虚设垂直结构DVS可以被形成，如先前参照图2F所述。在这样的实施方式中，垂直结构VS和虚设垂直结构DVS可以形成为穿透模制结构110、下平坦化绝缘层20、下牺牲图案LSL和缓冲绝缘层11，并且可以连接到衬底10。

[0111] 参照图7F，电极结构ST可以通过用电极EL代替下牺牲图案LSL和牺牲层SL而形成。如先前参照图2G所述，用电极EL替换下牺牲图案LSL和牺牲层SL可以包括形成沟槽、去除由沟槽暴露的下牺牲图案LSL和牺牲层SL以形成栅极区域、以及分别在栅极区域中形成电极EL。

[0112] 在电极结构ST的形成之后，第一层间绝缘层60和上平坦化绝缘层50可以被图案化，以形成连接区域CNR上的单元接触孔50c以及外围电路区域PCR上的上接触孔50a和50b。

[0113] 当上接触孔50a和50b形成在上平坦化绝缘层50中时，牺牲间隙填充图案44可以被去除，暴露牺牲插塞42。单元接触孔50c可以形成为暴露电极EL在连接区域CNR上的端部，并且暴露电极EL中的最下面的一个的单元接触孔可以形成为穿透第一层间绝缘层60、上平坦化绝缘层50和下平坦化绝缘层20。

[0114] 接着，由上接触孔50a和50b暴露的牺牲插塞42可以被去除，以形成暴露虚设杂质区域15的下接触孔33a和33b。这里，牺牲插塞42可以使用选择为相对于下平坦化绝缘层20和虚设牺牲图案DSL具有蚀刻选择性的蚀刻配方被去除。作为牺牲插塞42的去除的结果，衬底10可以被下接触孔33a和33b暴露，并且虚设杂质区域15可以防止由下接触孔33a和33b暴露的衬底10在去除牺牲插塞42的蚀刻工艺期间被蚀刻或损坏，或者减小由下接触孔33a和33b暴露的衬底10在去除牺牲插塞42的蚀刻工艺期间被蚀刻或损坏的可能性。

[0115] 参照图7G，单元接触插塞CPLG可以形成在单元接触孔50c中和在连接区域CNR上，并且第一外围接触插塞PPLGa和第二外围接触插塞PPLGb可以形成在下接触孔33a、33b和上接触孔50a、50b中以及在外围电路区域PCR上。

[0116] 如参照图2I所述，第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个可以包括形成为穿透下平坦化绝缘层20的下部P1、以及从下部P1连续地延伸并形成为穿透上平坦化绝缘层50和第一层间绝缘层60的上部P2。

[0117] 在一些实施方式中，第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个的下部P1可以具有位于电极EL中的最下面一个与第二最下面的一个之间的顶表面。第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的顶表面可以位于比下半导体图案LSP的水平更高的水平处。

[0118] 参照图7H，第二层间绝缘层70可以形成在覆盖单元接触插塞CPLG以及第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的顶表面的第一层间绝缘层60上。

[0119] 如上所述，位线接触插塞BPLG、连接接触插塞CNT和外围连接接触插塞PCNT可以形成在第二层间绝缘层70中。此外，位线BL、互连线ICL和外围连接线PCL可以形成在第二层间

Micron Ex. 1002, p. 733
Micron v. YMTC
IPR2024-00789

绝缘层70上。

[0120] 图8A至图8I是示出根据本发明构思的各种实施方式的制造三维半导体存储器件的方法的剖视图。

[0121] 为了简要描述,先前参照图2A至图2J描述的元件或步骤可以由相似或相同的附图标记标识,而不重复其重叠描述。

[0122] 在图8A至图8H的实施方式中,虚设杂质区域15和牺牲插塞42可以在模制结构110的一部分(例如下模制结构110a)的形成之后被形成。

[0123] 参照图8A,外围逻辑结构PSTR可以形成在衬底10的外围电路区域PCR上。这里,外围逻辑结构PSTR可以包括外围栅极堆叠PGS、源极/漏极杂质区域13和外围绝缘图案35。外围绝缘图案35可以形成为覆盖外围栅极堆叠PGS和源极/漏极杂质区域13,并暴露衬底10的单元阵列区域CAR和连接区域CNR。

[0124] 参照图8B,下模制结构110a可以形成在衬底10的单元阵列区域CAR和连接区域CNR上。

[0125] 下模制结构110a可以包括垂直地和交替地堆叠在衬底10上的下牺牲层SLa和下绝缘层ILDa。下模制结构110a的形成可以包括在提供有外围逻辑结构PSTR的衬底10上交替地堆叠下牺牲层SLa和下绝缘层ILDa、然后对下牺牲层SLa和下绝缘层ILDa执行修整工艺。因此,下模制结构110a可以在连接区域CNR上具有阶梯结构。

[0126] 此外,在下模制结构110a的形成期间,虚设间隔物DSP可以形成在外围绝缘图案35的侧表面上。虚设间隔物DSP可以是或者可以包括下牺牲层SLa和下绝缘层ILDa的未被各向异性蚀刻工艺蚀刻的剩余部分。

[0127] 参照图8C,下平坦化绝缘层20可以形成在衬底10上。下平坦化绝缘层20可以形成为填充下模制结构110a与外围逻辑结构PSTR之间的间隙区域并具有基本上平坦的顶表面。

[0128] 下平坦化绝缘层20可以通过形成覆盖衬底10的绝缘层并对绝缘层执行平面化工艺而形成。下平坦化绝缘层20可以形成为覆盖外围逻辑结构PSTR。

[0129] 接着,第一下接触孔33a和第二下接触孔33b可以形成为穿透下平坦化绝缘层20和外围绝缘图案35。如上所述,在第一下接触孔33a和第二下接触孔33b的形成之后,虚设杂质区域15可以形成在源极/漏极杂质区域13中。

[0130] 参照图8D,在虚设杂质区域15的形成之后,牺牲插塞42和牺牲间隙填充图案44可以形成在第一下接触孔33a和第二下接触孔33b的每个中。

[0131] 参照图8E,上模制结构110b可以形成在下模制结构110a上。

[0132] 上模制结构110b可以包括垂直地和交替地堆叠在下模制结构110a上的上牺牲层SLb和上绝缘层ILDb。上模制结构110b可以通过在衬底10上交替地堆叠上牺牲层SLb和上绝缘层ILDb并且对上牺牲层SLb和上绝缘层ILDb执行修整工艺而形成。上模制结构110b可以形成为在连接区域CNR上具有阶梯结构。

[0133] 当上模制结构110b被形成时,外围电路区域PCR上的牺牲插塞42和牺牲间隙填充图案44的顶表面可以被暴露。

[0134] 参照图8F,上平坦化绝缘层50可以形成在外围逻辑结构PSTR和下平坦化绝缘层20上,覆盖上模制结构110b。

[0135] 在上平坦化绝缘层50的形成之后,垂直结构VS可以形成为穿透下模制结构110a和

Micron Ex. 1002, p. 734
Micron v. YMTC
IPR2024-00789

上模制结构110b。在一些实施方式中，垂直结构VS的形成可以包括形成穿透下模制结构110a和上模制结构110b并暴露衬底10的垂直孔以及在垂直孔的每个中形成与衬底10接触的垂直半导体图案。此外，垂直绝缘图案VP可以在垂直半导体图案的形成之前形成在垂直孔的每个中。在该实施方式中，图2F中所示的下半导体图案LSP的形成在垂直结构VS的形成期间可以被省略。

[0136] 如上所述，在垂直结构VS的形成期间，虚设垂直结构DVS可以在连接区域CNR上形成为穿透下模制结构110a和上模制结构110b。

[0137] 参照图8G，在垂直结构VS和虚设垂直结构DVS的形成之后，第一层间绝缘层60可以被形成。第一层间绝缘层60可以形成在上平坦化绝缘层50上，覆盖垂直结构VS和虚设垂直结构DVS的顶表面。

[0138] 接着，下牺牲层SLa和上牺牲层SLb可以用电极EL替换。结果，其中电极EL垂直地堆叠在衬底10上的电极结构ST可以被形成。

[0139] 如上所述，在电极结构ST的形成之后，单元接触孔50c以及上接触孔50a和50b可以形成为穿透第一层间绝缘层60和上平坦化绝缘层50。单元接触孔50c可以形成为分别暴露电极EL的端部，并且上接触孔50a和50b可以分别形成为暴露牺牲插塞42。

[0140] 接着，由上接触孔50a和50b暴露的牺牲插塞42可以被去除，暴露下接触孔33a和33b的内表面以及虚设杂质区域15。

[0141] 参照图8H，单元接触插塞CPLG可以形成在单元接触孔50c中和在连接区域CNR上，并且第一外围接触插塞PPLGa和第二外围接触插塞PPLGb可以形成在下接触孔33a、33b和上接触孔50a、50b中以及在外围电路区域PCR上。如上所述，第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个可以包括下部P1和上部P2。在一些实施方式中，第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的每个的下部P1的垂直长度可以取决于外围绝缘图案35的厚度和/或下平坦化绝缘层20的厚度而改变。

[0142] 参照图8I，第二层间绝缘层70可以形成在覆盖单元接触插塞CPLG以及第一外围接触插塞PPLGa和第二外围接触插塞PPLGb的顶表面的第一层间绝缘层60上。

[0143] 如上所述，位线接触插塞BPLG、连接接触插塞CNT和外围连接接触插塞PCNT可以形成在第二层间绝缘层70中。此外，位线BL、互连线ICL和外围连接线PCL可以形成在第二层间绝缘层70上。

[0144] 图9是根据发明构思的各种实施方式的三维半导体存储器件的剖视图。

[0145] 参照图9，第一电极结构ST1和第二电极结构ST2可以在衬底10的单元阵列区域上被提供为彼此间隔开。第一电极结构ST1和第二电极结构ST2的每个可以包括交替地堆叠在衬底10上的电极EL和绝缘层ILD。在一些实施方式中，第一电极结构ST1和第二电极结构ST2可以在一方向上延伸，并且绝缘间隙填充层120可以被提供在第一电极结构ST1与第二电极结构ST2之间。

[0146] 沟道结构CHS可以包括穿透第一电极结构ST1的第一垂直半导体柱VSP1和穿透第二电极结构ST2的第二垂直半导体柱VSP2、以及将第一垂直半导体柱VSP1和第二垂直半导体柱VSP2彼此连接的水平半导体图案HSP。

[0147] 第一垂直半导体柱VSP1和第二垂直半导体柱VSP2可以被提供在形成为穿透第一电极结构ST1和第二电极结构ST2的垂直孔中。第一垂直半导体柱VSP1和第二垂直半导体柱

Micron Ex. 1002, p. 735
Micron v. YMTC
IPR2024-00789

VSP2的每个可以包括提供在其最高水平处的导电垫D。第一垂直半导体柱VSP1可以连接到位线BL,第二垂直半导体柱VSP2可以连接到公共源极线CSL。

[0148] 水平半导体图案HSP可以被提供在形成于衬底10中的水平凹陷区域中。水平半导体图案HSP可以从第一电极结构ST1下面的区域水平地延伸到第二电极结构ST2下面的另一区域,并且可以将第一垂直半导体柱VSP1和第二垂直半导体柱VSP2彼此连接。

[0149] 外围接触插塞可以通过接触孔连接到MOS晶体管。根据发明构思的一些实施方式,接触孔的上部区域和下部区域可以分开形成,因而当堆叠在单元阵列区域上的电极的数量增加时,可以增大用于形成接触孔的工艺中的工艺余量。

[0150] 牺牲插塞用于形成外围接触插塞,并且这可以使得能够用导电材料一次填充接触孔的上部区域和下部区域。

[0151] 虚设杂质区域可以形成在源极/漏极杂质区域上。因此,可以防止源极/漏极杂质区域在去除牺牲插塞时被损坏,或者减小源极/漏极杂质区域在去除牺牲插塞时被损坏的可能性。

[0152] 虽然已经具体示出和描述了发明构思的示例实施方式,但是本领域普通技术人员将理解,可以在其中进行形式和细节上的变化而不背离所附权利要求的精神和范围。

[0153] 本申请要求享有2017年4月7日在韩国知识产权局提交的韩国专利申请第10-2017-0045114号的优先权,其全部内容通过引用在此合并。

Micron Ex. 1002, p. 736
Micron v. YMTC
IPR2024-00789

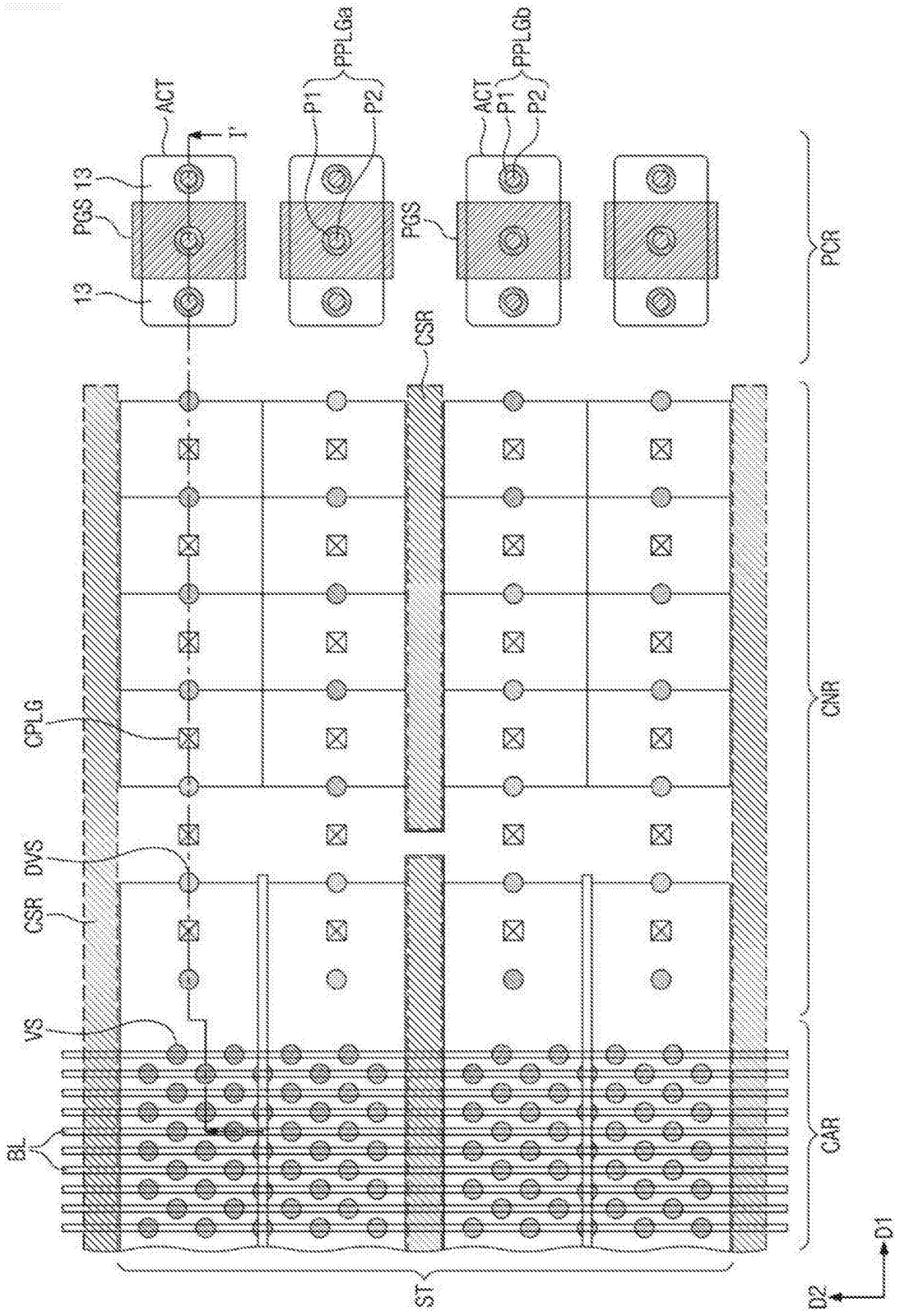


图1

Micron Ex. 1002, p. 737
Micron v. YMTC
IPR2024-00789

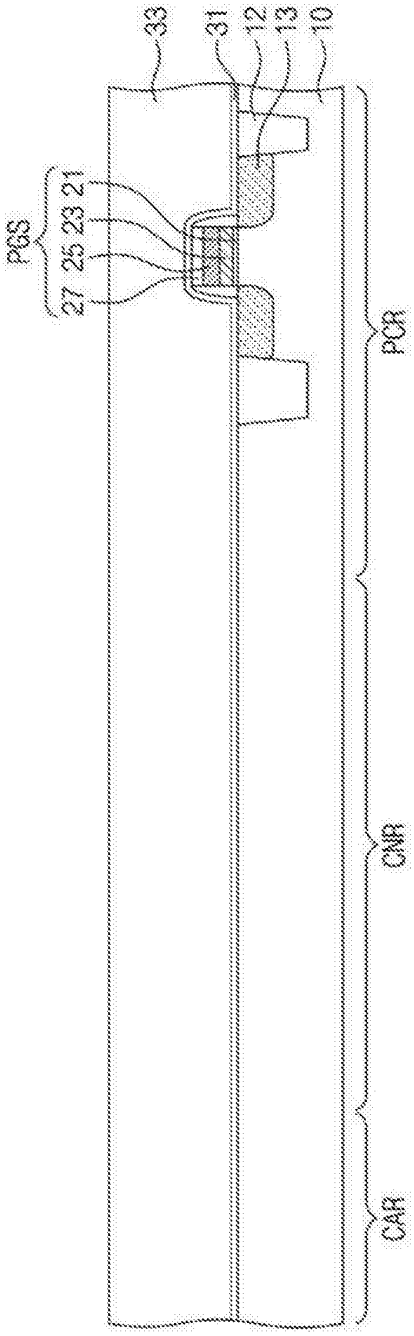


图2A

Micron Ex. 1002, p. 738
Micron v. YMTC
IPR2024-00789

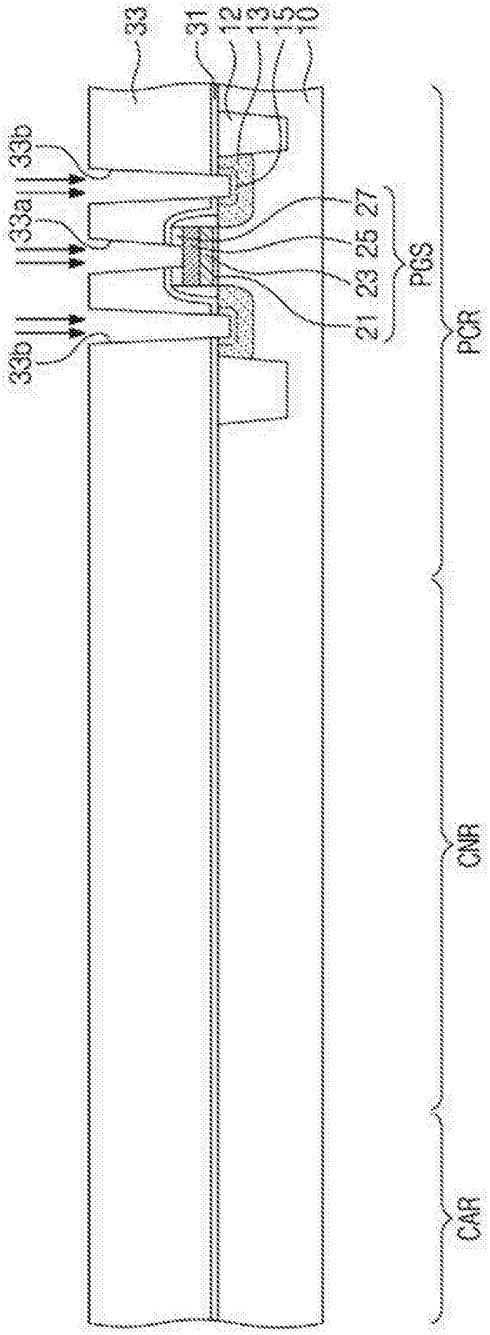


图2B

Micron Ex. 1002, p. 739
Micron v. YMTC
IPR2024-00789

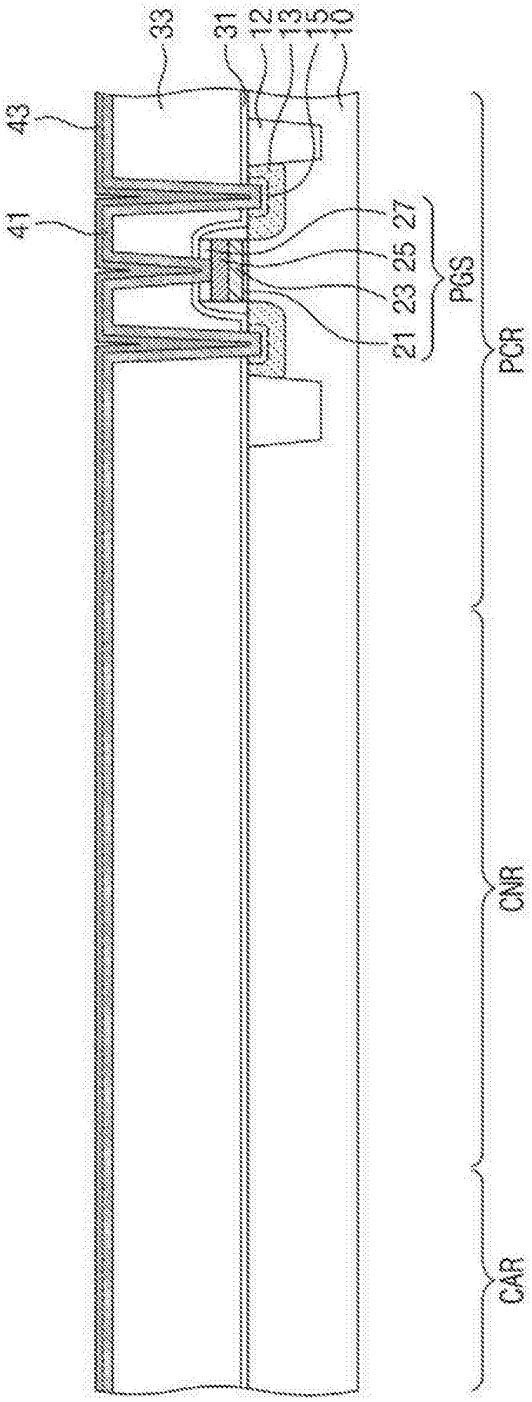


图2C

Micron Ex. 1002, p. 740
Micron v. YMTC
IPR2024-00789

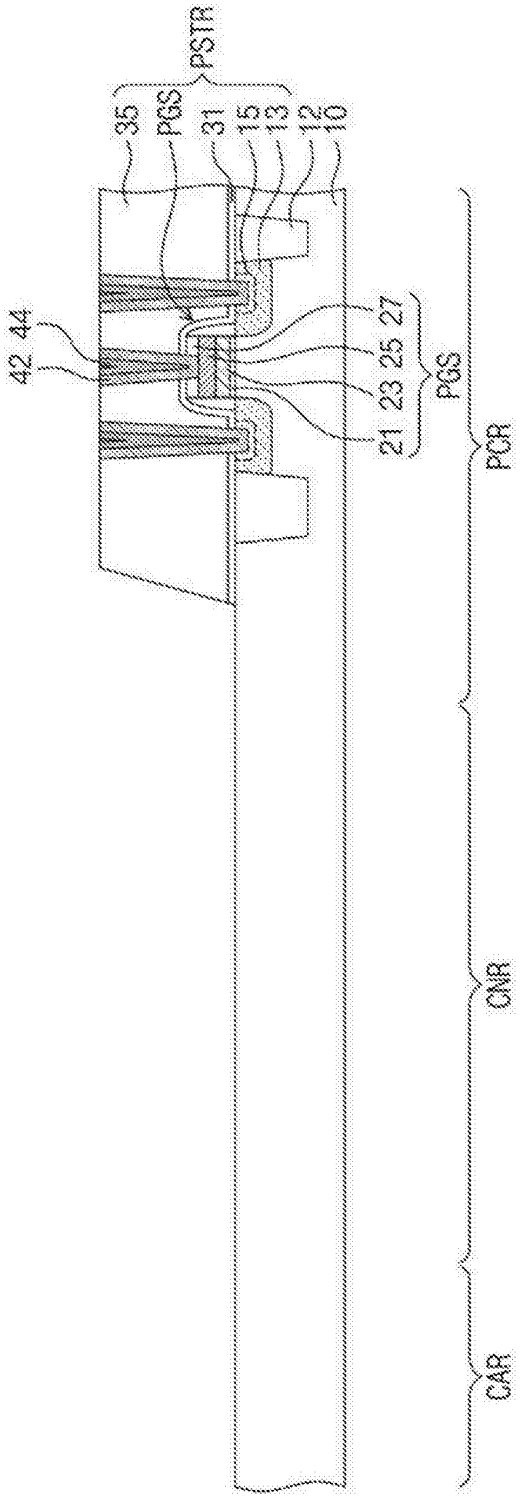


图2D

Micron Ex. 1002, p. 741
Micron v. YMTC
IPR2024-00789

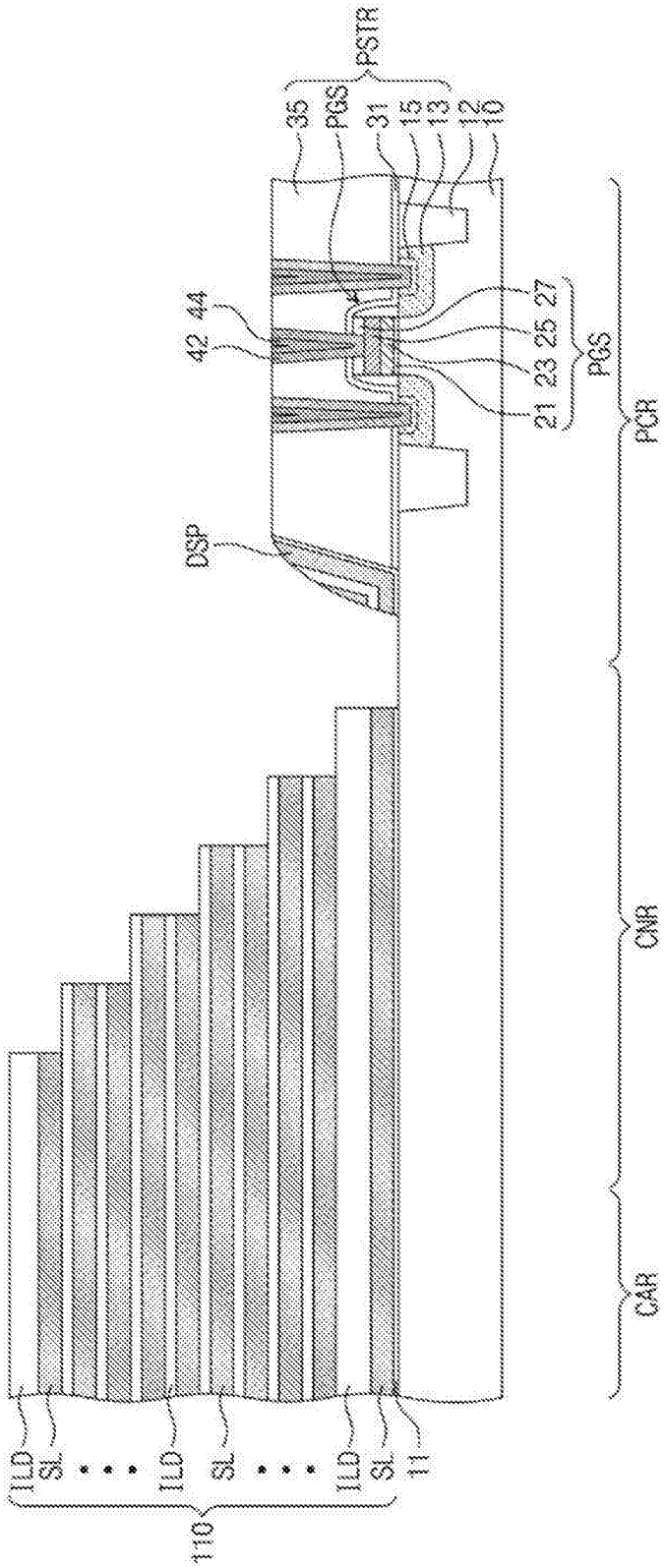


图2E

Micron Ex. 1002, p. 742
Micron v. YMTC
IPR2024-00789

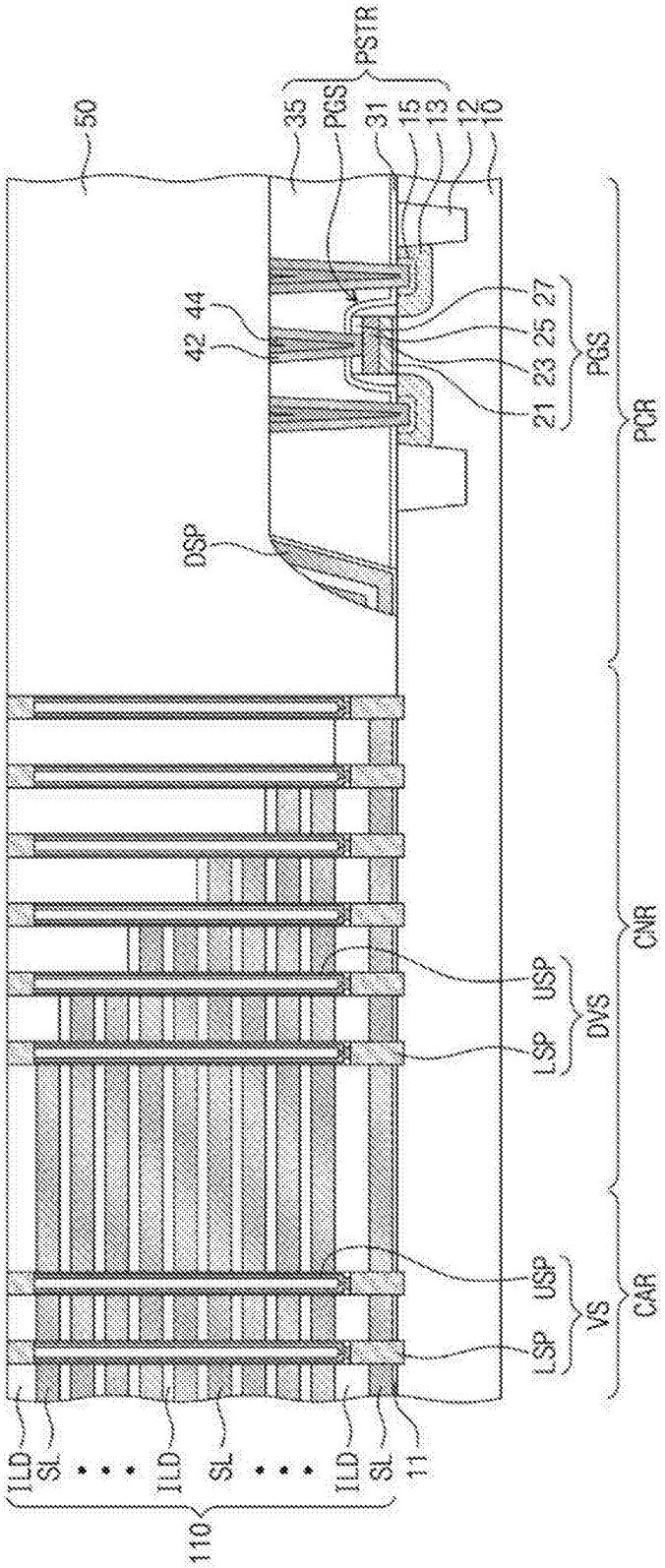


图2F

Micron Ex. 1002, p. 743
Micron v. YMTC
IPR2024-00789

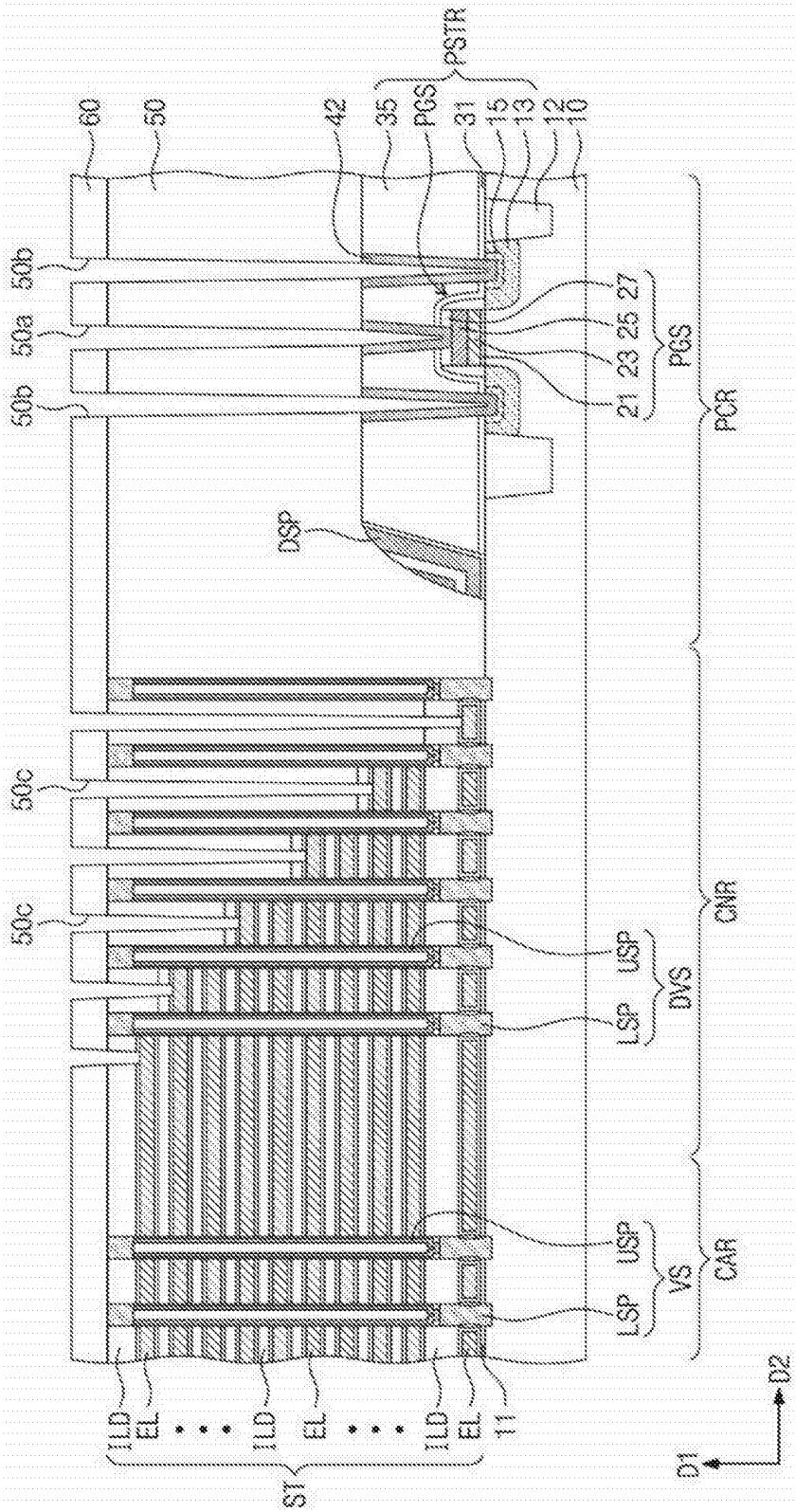


图2G

Micron Ex. 1002, p. 744
Micron v. YMTC
IPR2024-00789

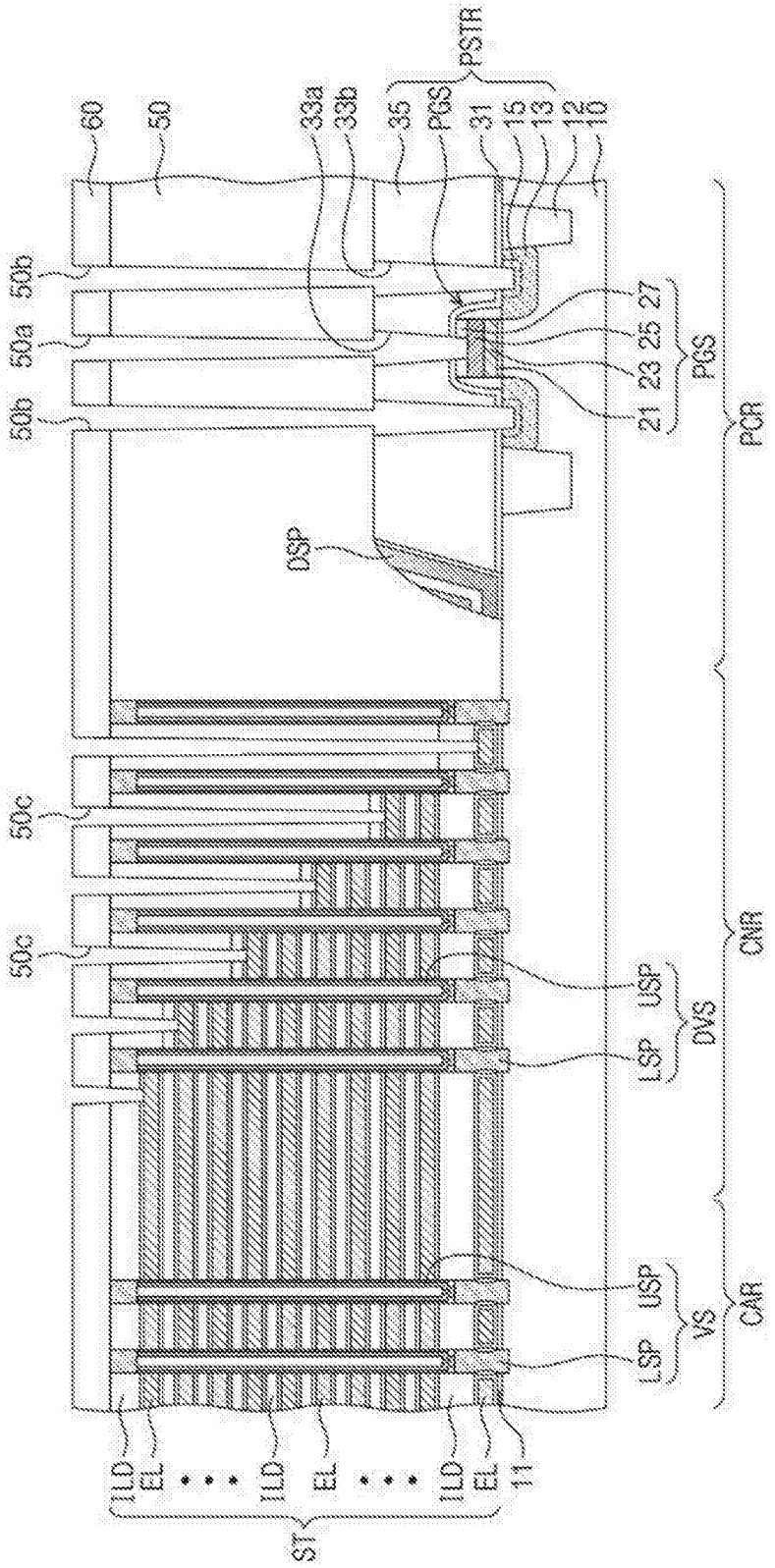


图2H

Micron Ex. 1002, p. 745
Micron v. YMTC
IPR2024-00789

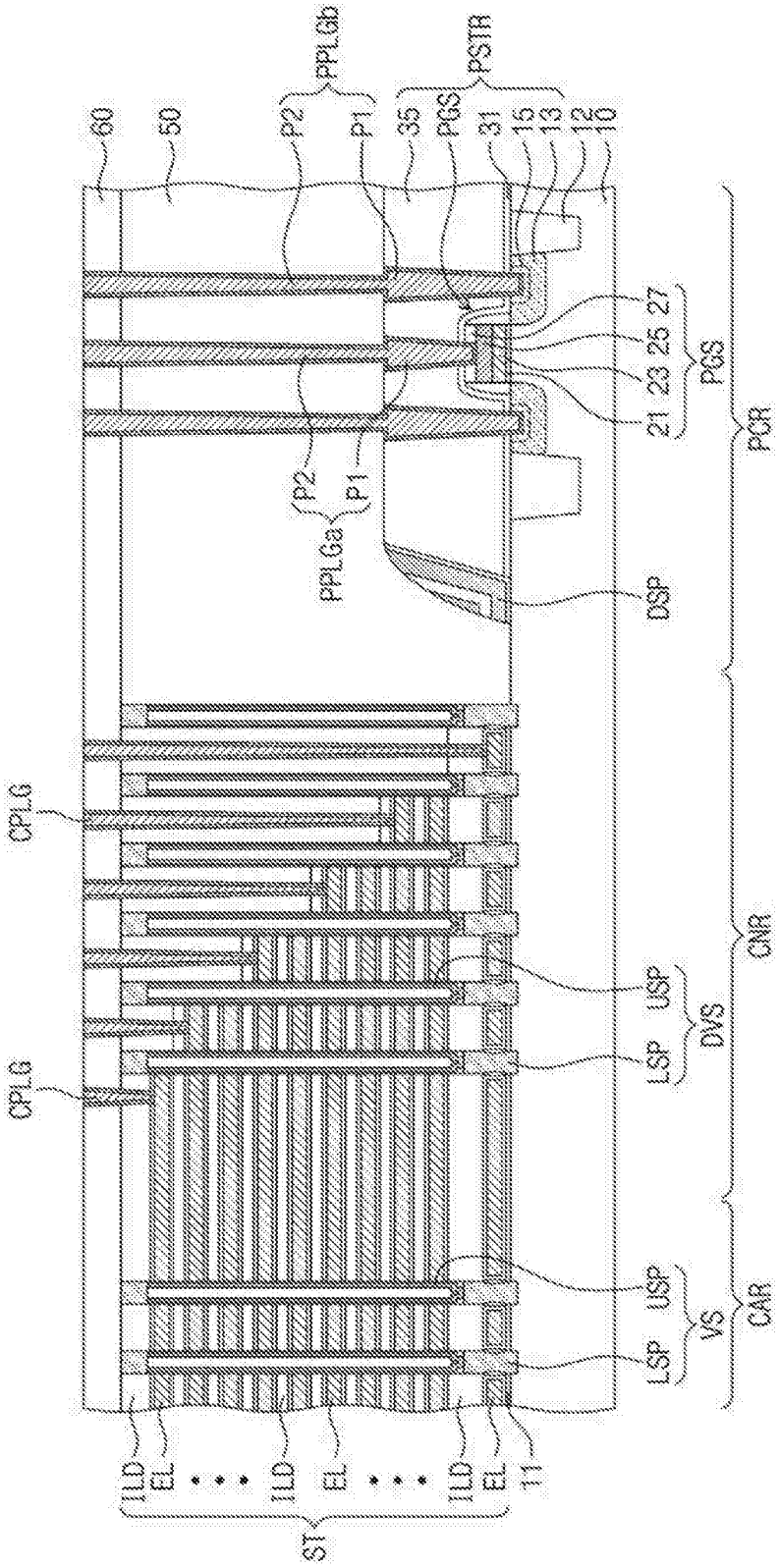


图2I

Micron Ex. 1002, p. 746
Micron v. YMTC
IPR2024-00789

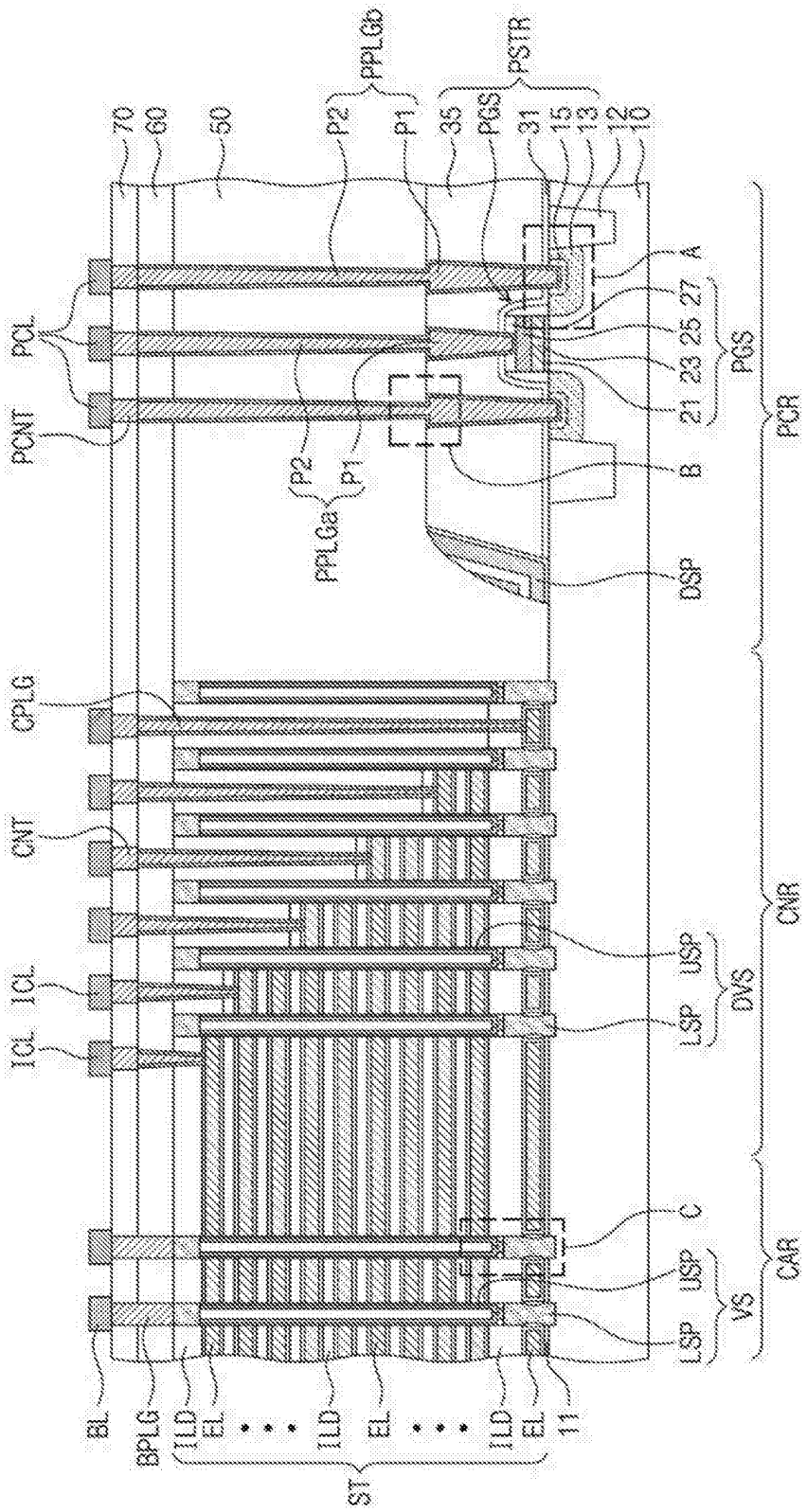


图2J

Micron Ex. 1002, p. 747
Micron v. YMTC
IPR2024-00789

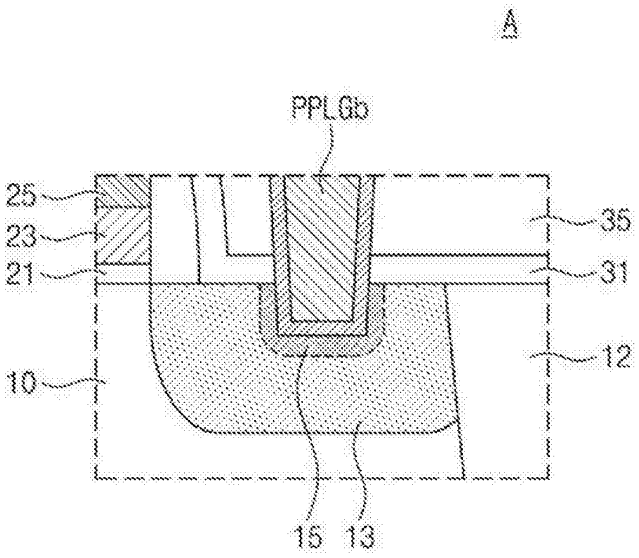


图3A

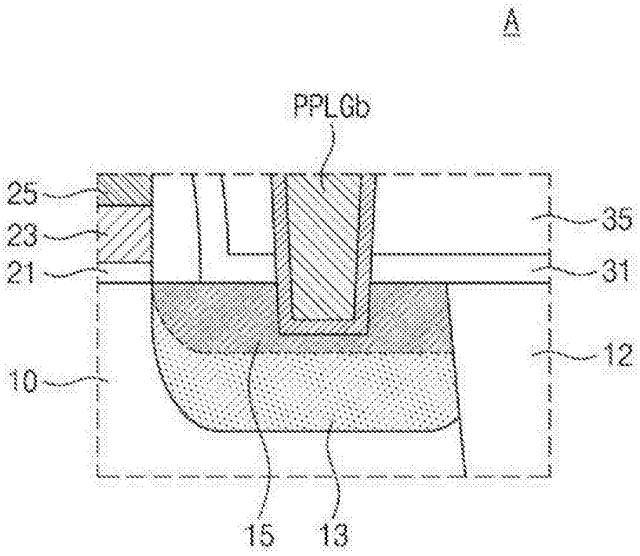


图3B

Micron Ex. 1002, p. 748
Micron v. YMTC
IPR2024-00789

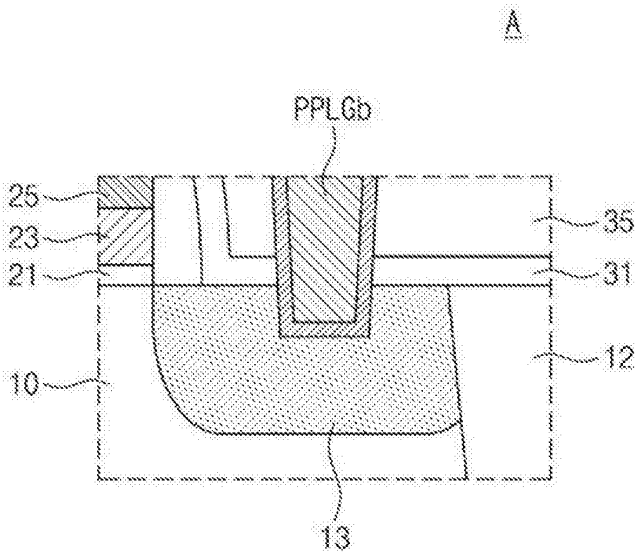


图3C

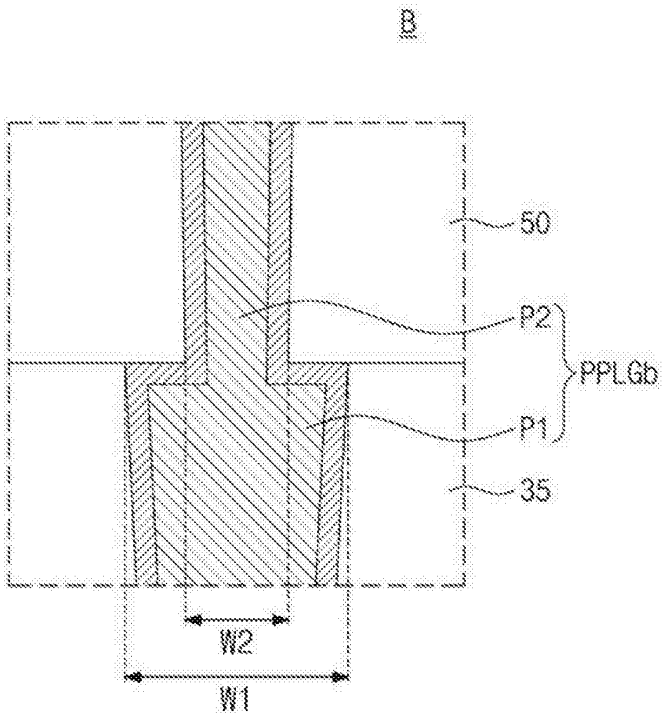


图4A

Micron Ex. 1002, p. 749
Micron v. YMTC
IPR2024-00789

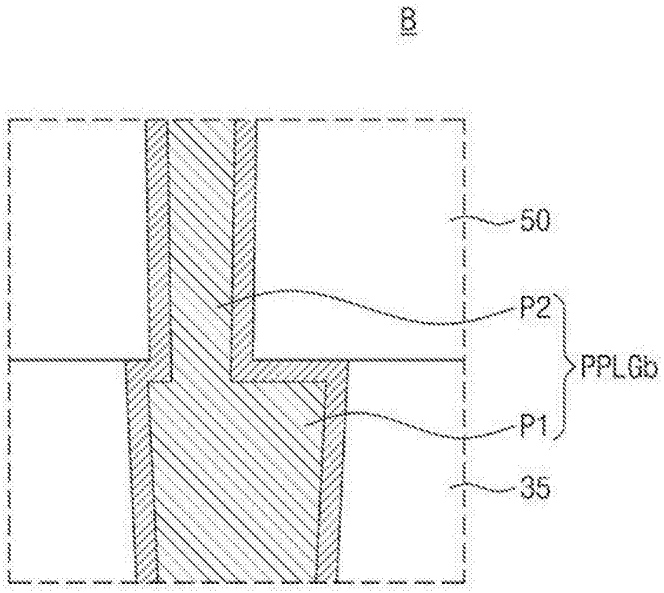


图4B

Micron Ex. 1002, p. 750
Micron v. YMTC
IPR2024-00789

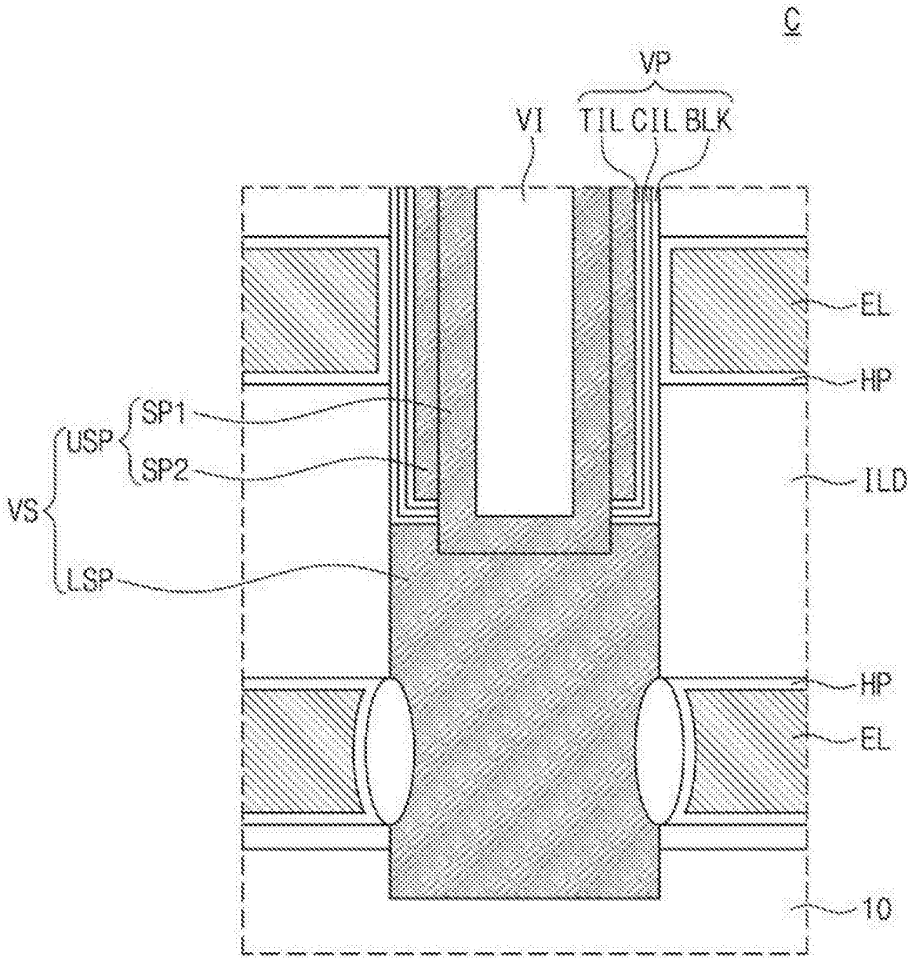


图5

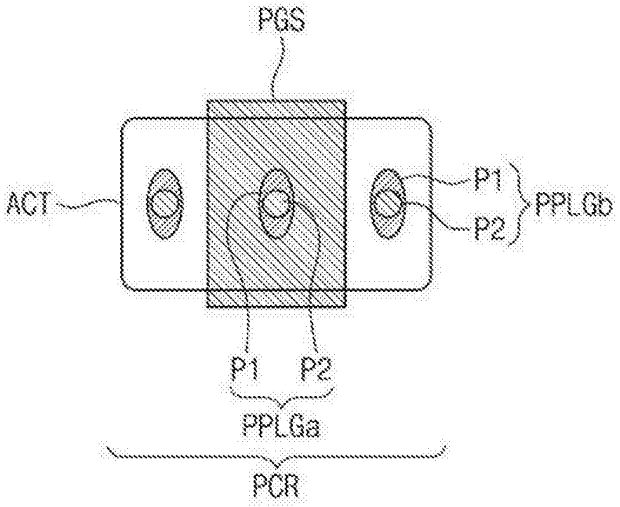


图6A

Micron Ex. 1002, p. 751
Micron v. YMTC
IPR2024-00789

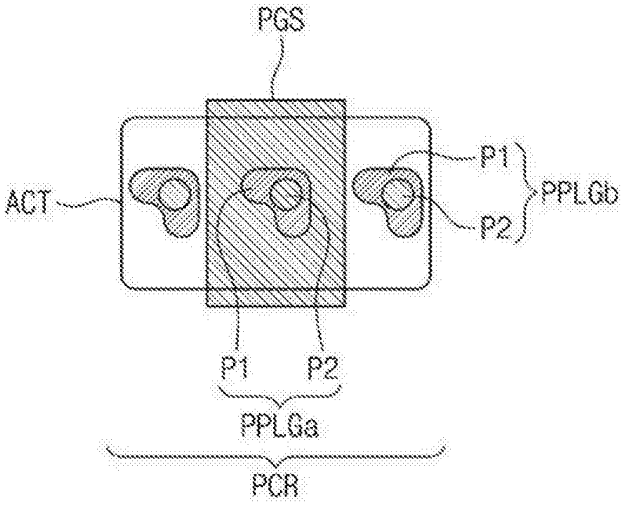


图6B

Micron Ex. 1002, p. 752
Micron v. YMTC
IPR2024-00789

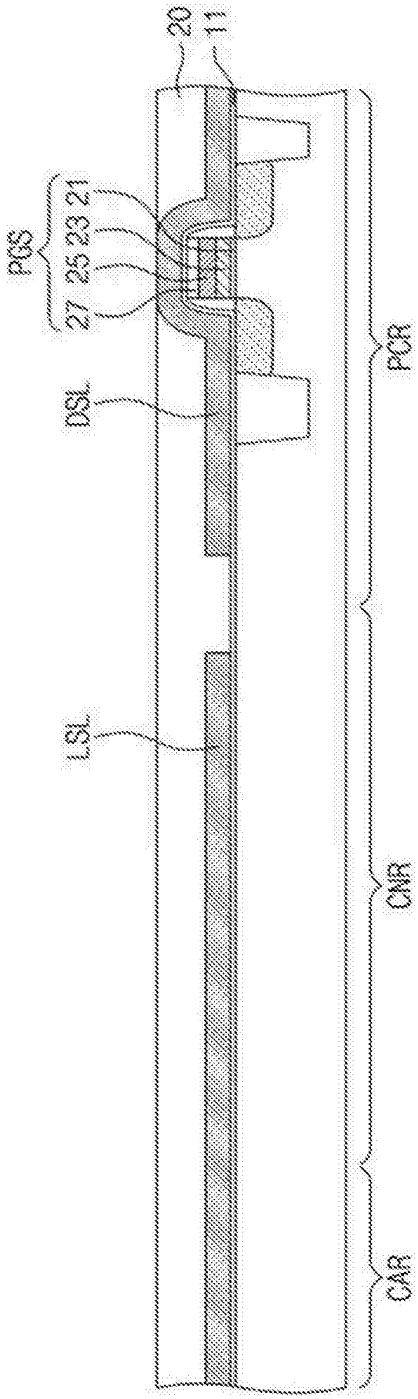


图7A

Micron Ex. 1002, p. 753
Micron v. YMTC
IPR2024-00789

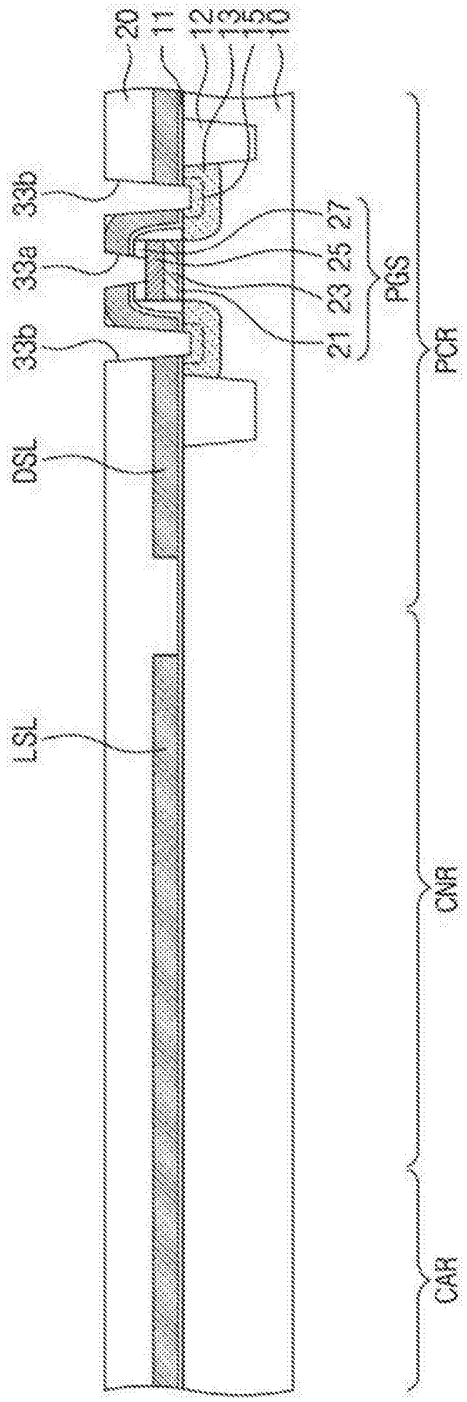


图7B

Micron Ex. 1002, p. 754
Micron v. YMTC
IPR2024-00789

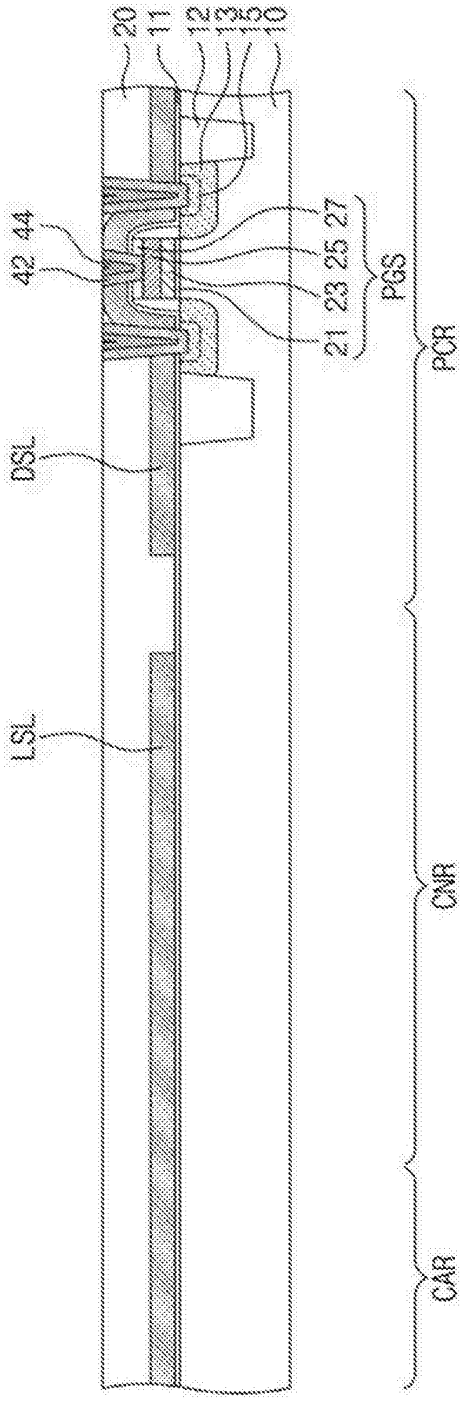


图7C

Micron Ex. 1002, p. 755
Micron v. YMTC
IPR2024-00789

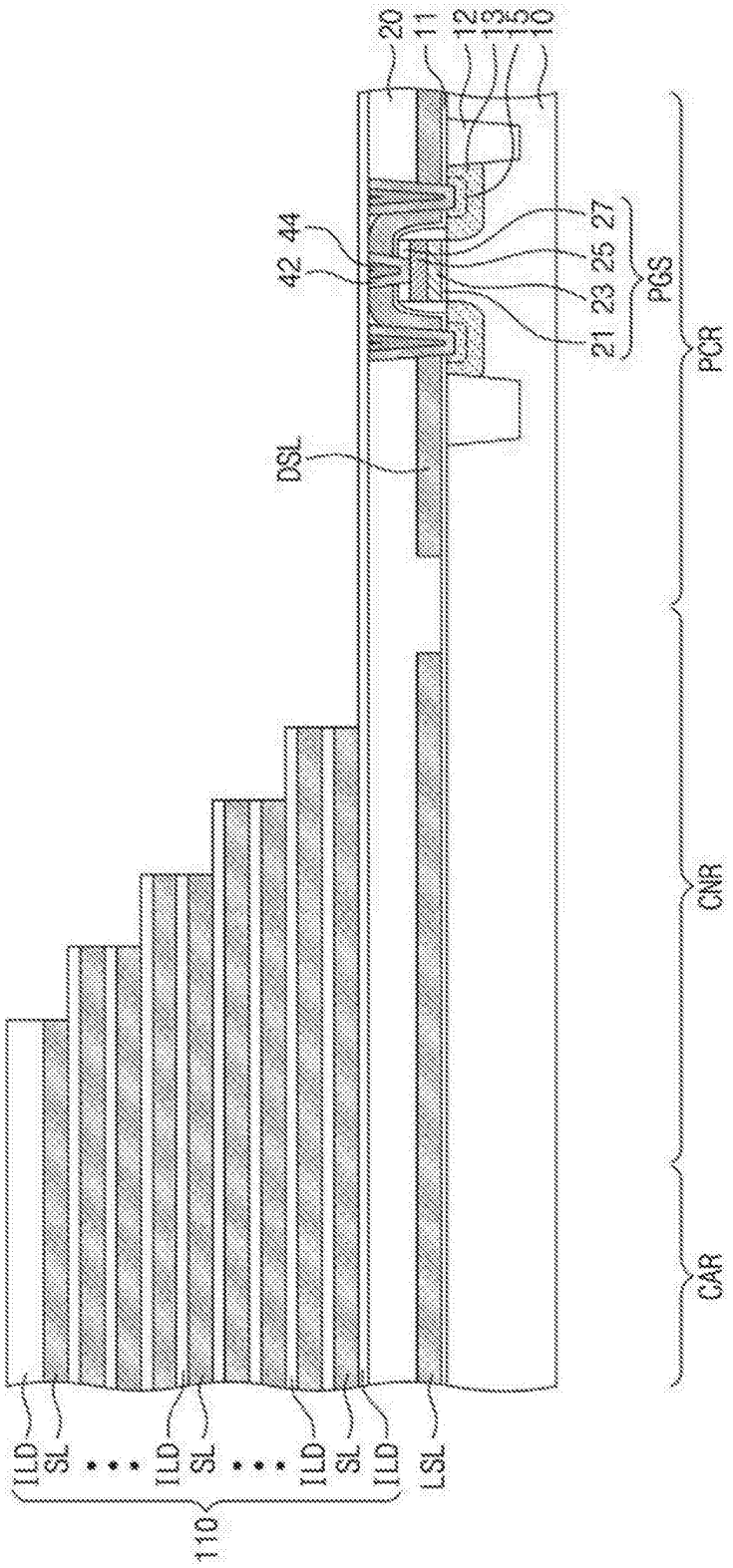


图7D

Micron Ex. 1002, p. 756
Micron v. YMTC
IPR2024-00789

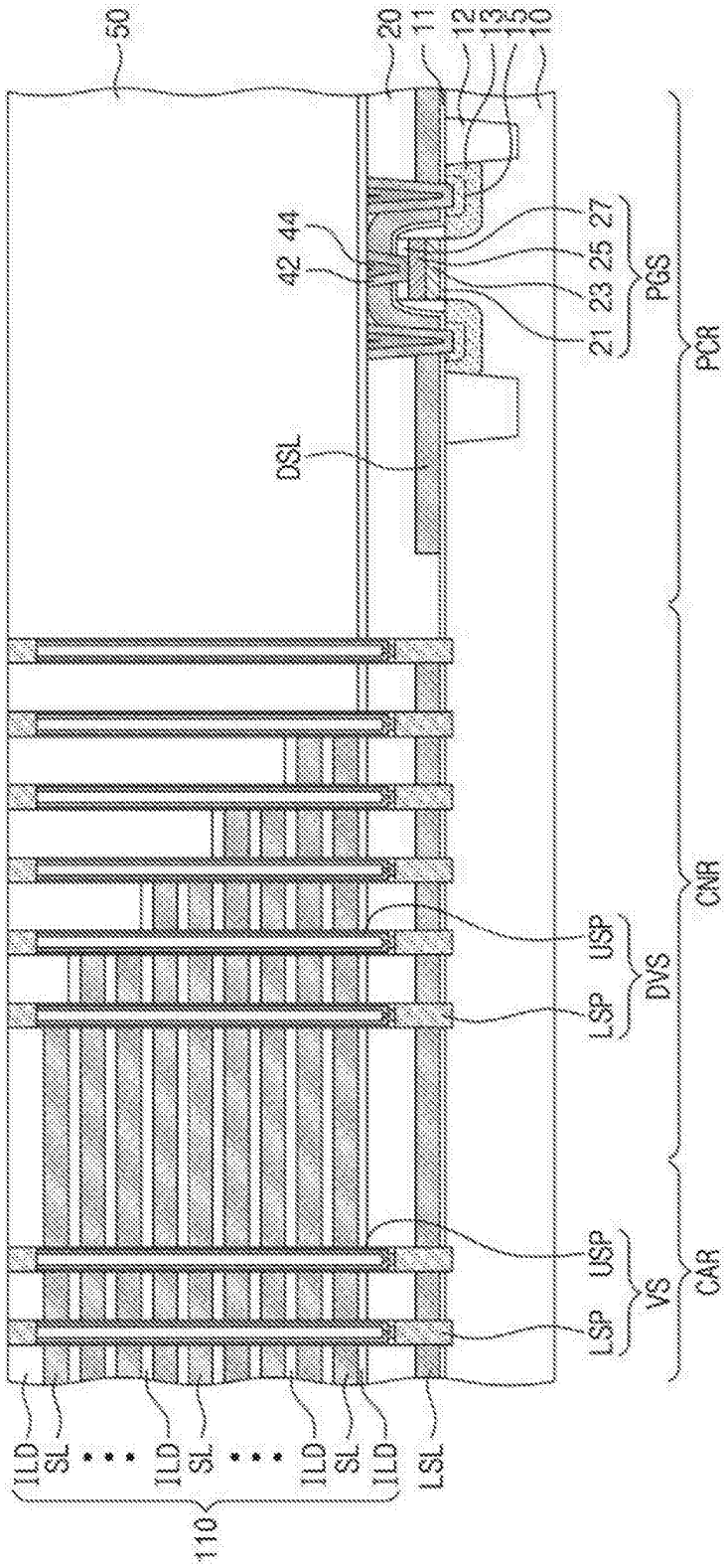


图7E

Micron Ex. 1002, p. 757
Micron v. YMTC
IPR2024-00789

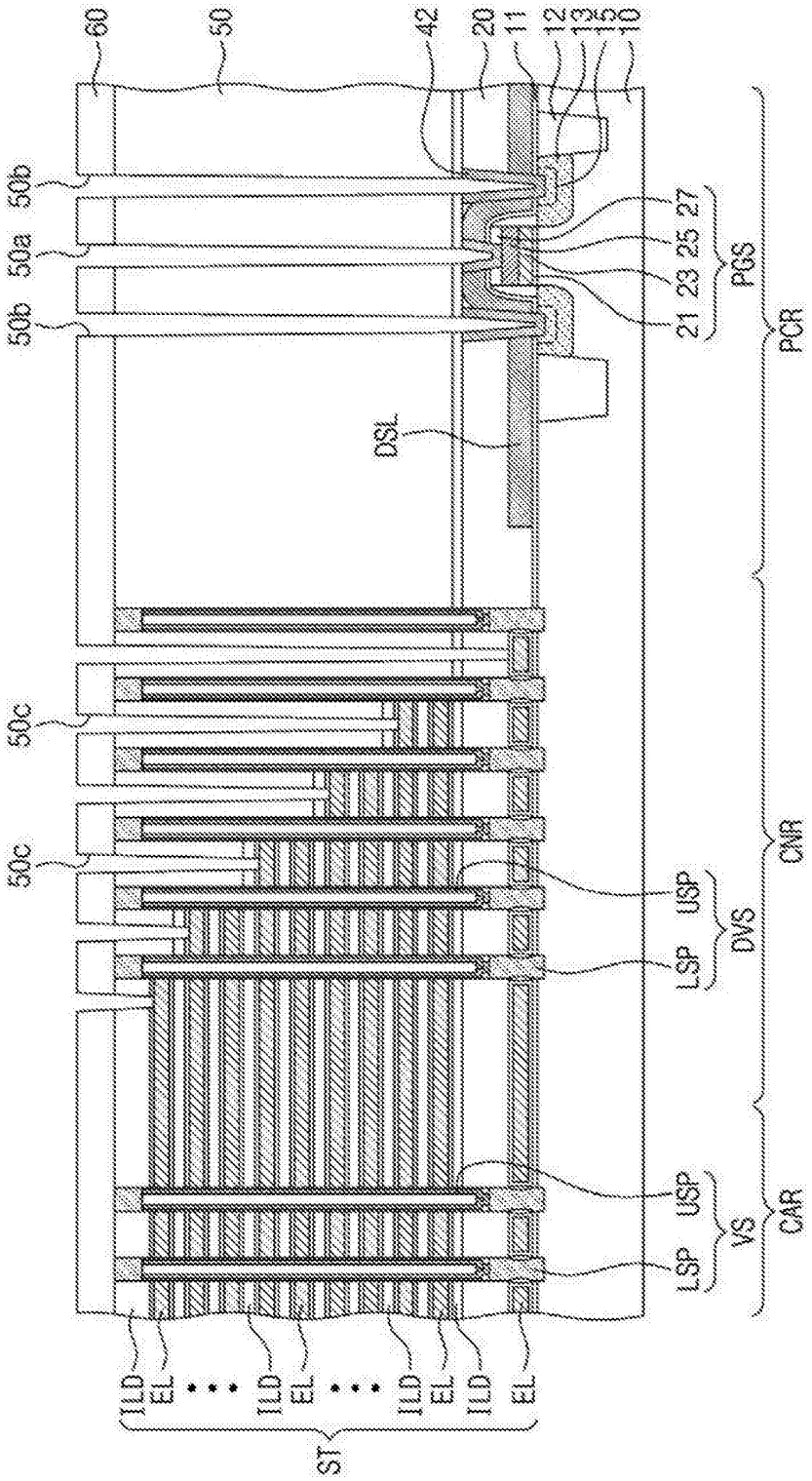


图7F

Micron Ex. 1002, p. 758
Micron v. YMTC
IPR2024-00789

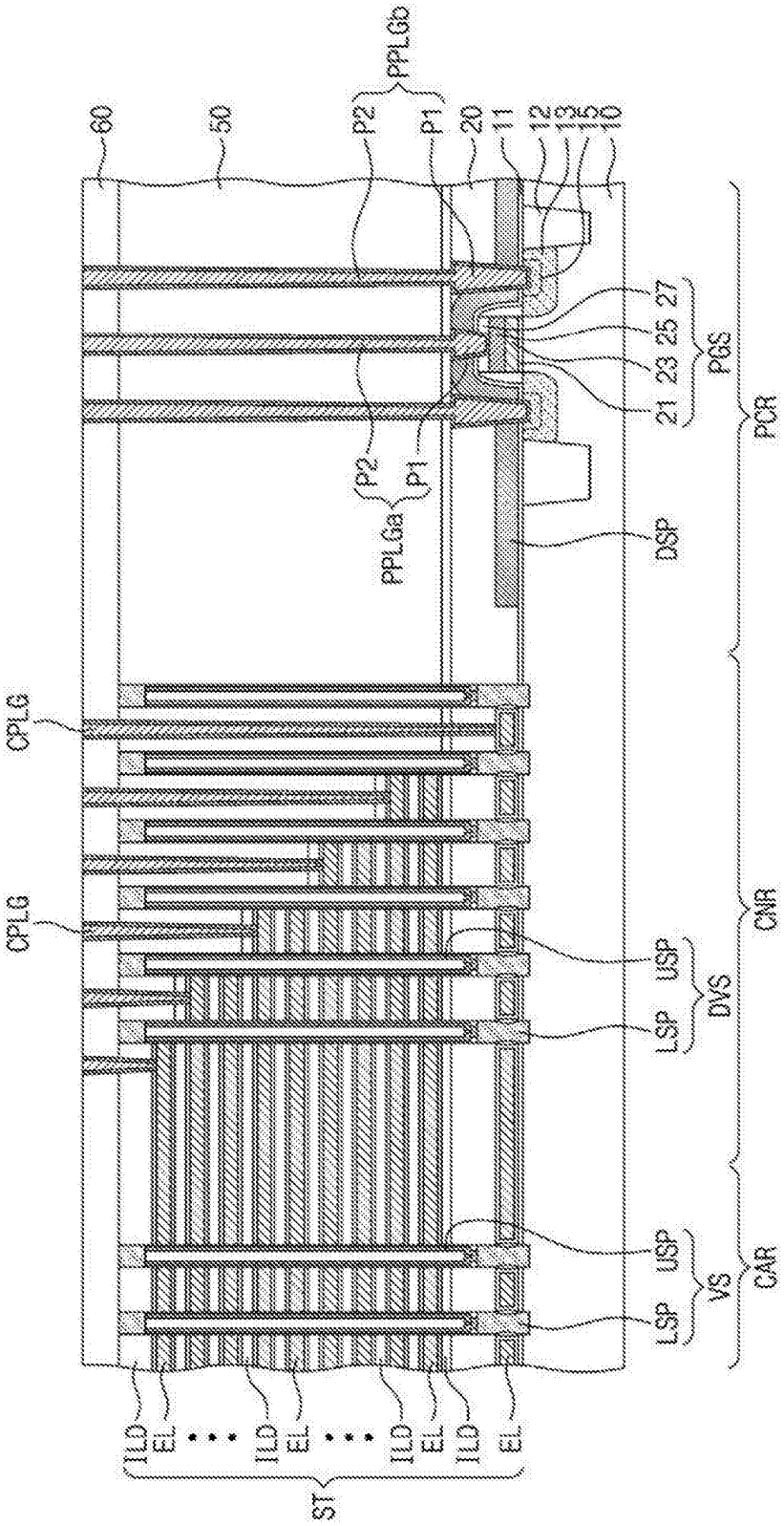


图7G

Micron Ex. 1002, p. 759
Micron v. YMTC
IPR2024-00789

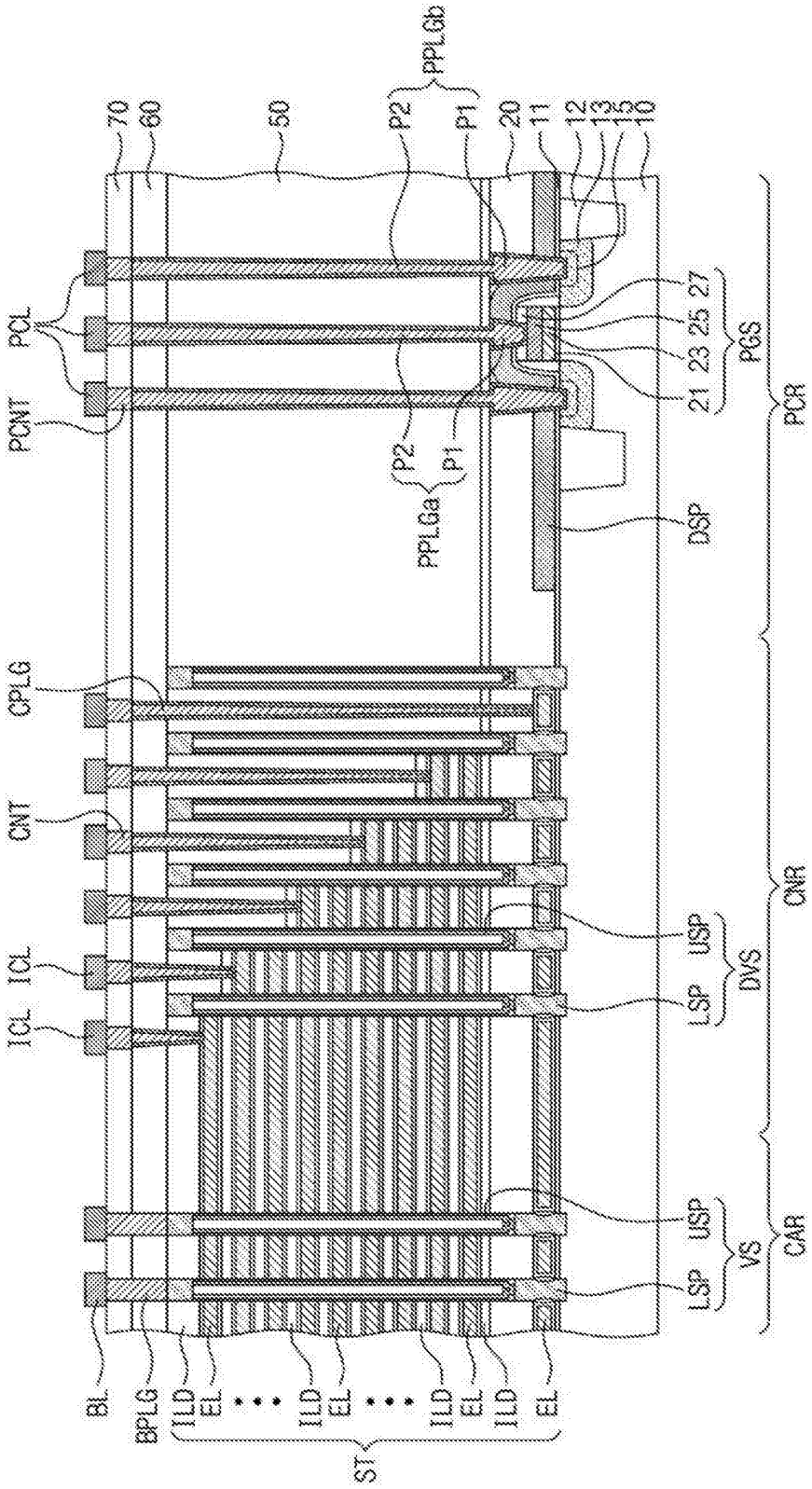


图7H

Micron Ex. 1002, p. 760
Micron v. YMTC
IPR2024-00789

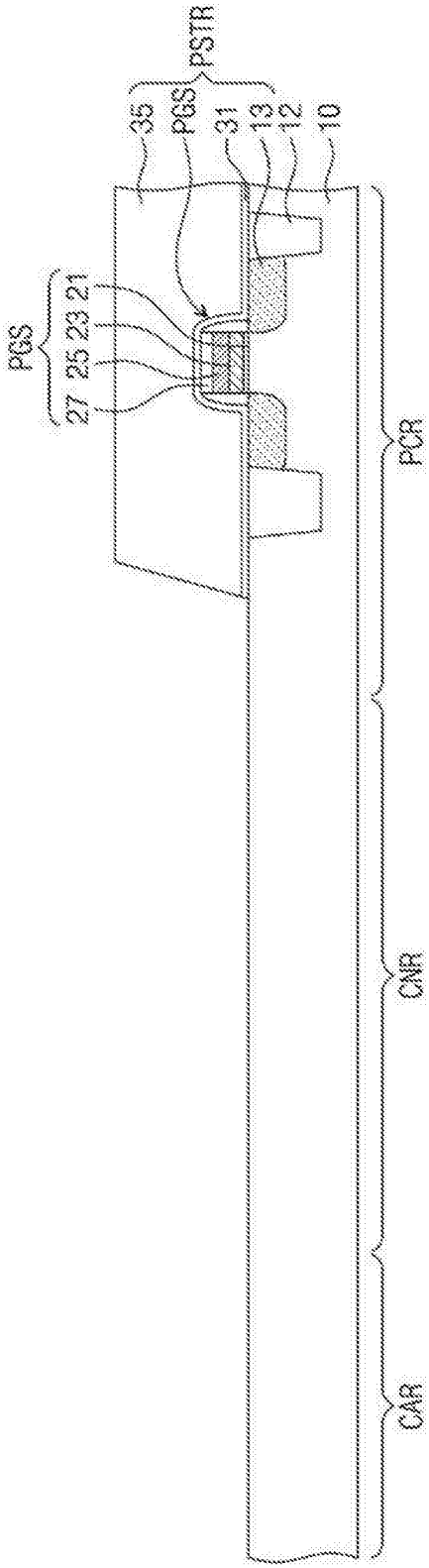


图8A

Micron Ex. 1002, p. 761
Micron v. YMTC
IPR2024-00789

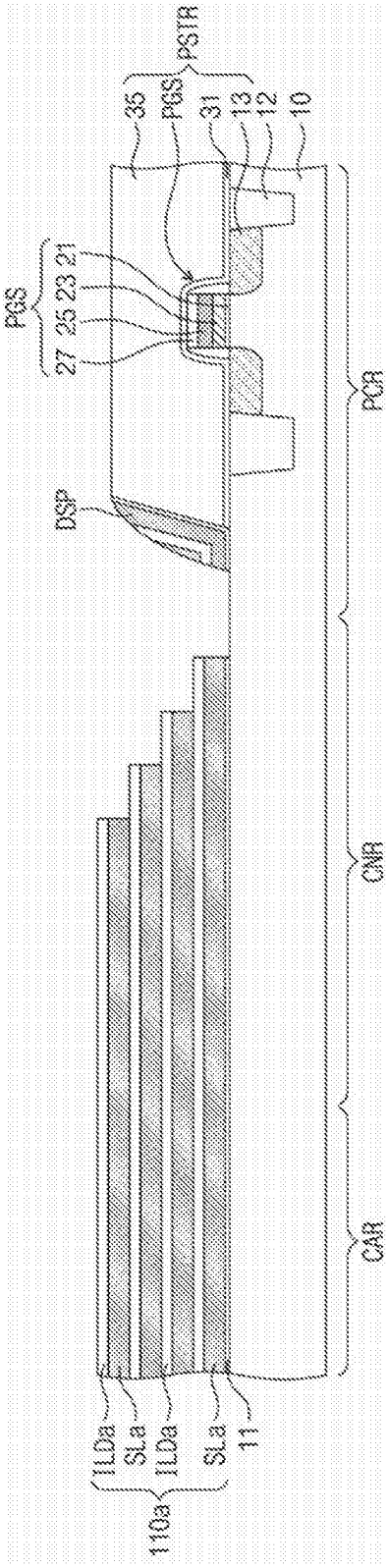


图8B

Micron Ex. 1002, p. 762
Micron v. YMTC
IPR2024-00789

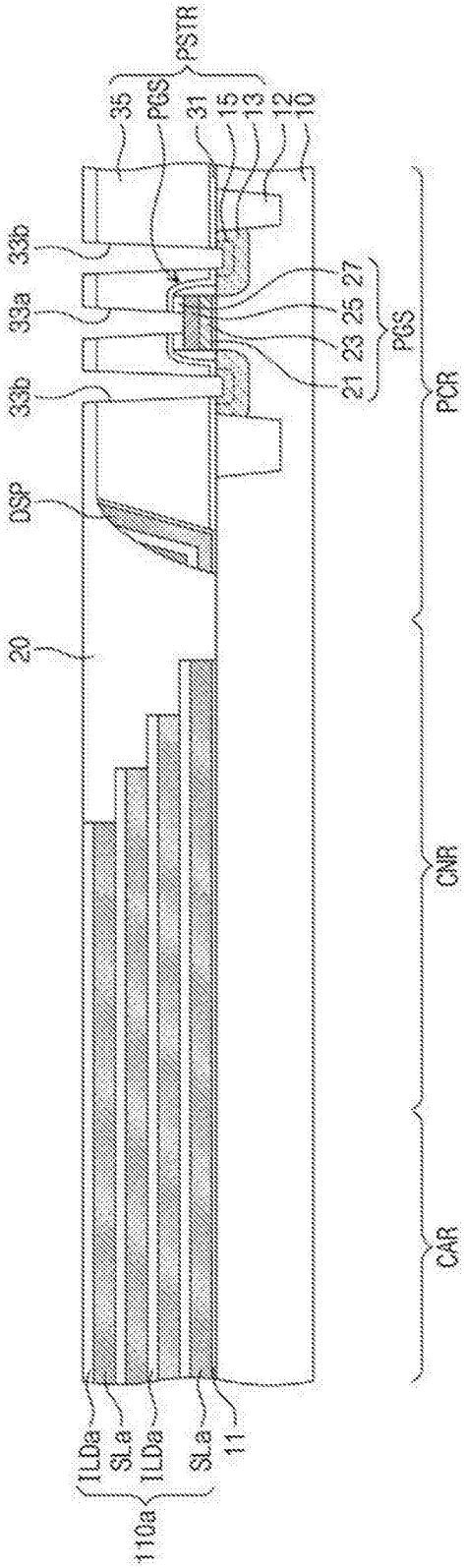


图8C

Micron Ex. 1002, p. 763
Micron v. YMTC
IPR2024-00789

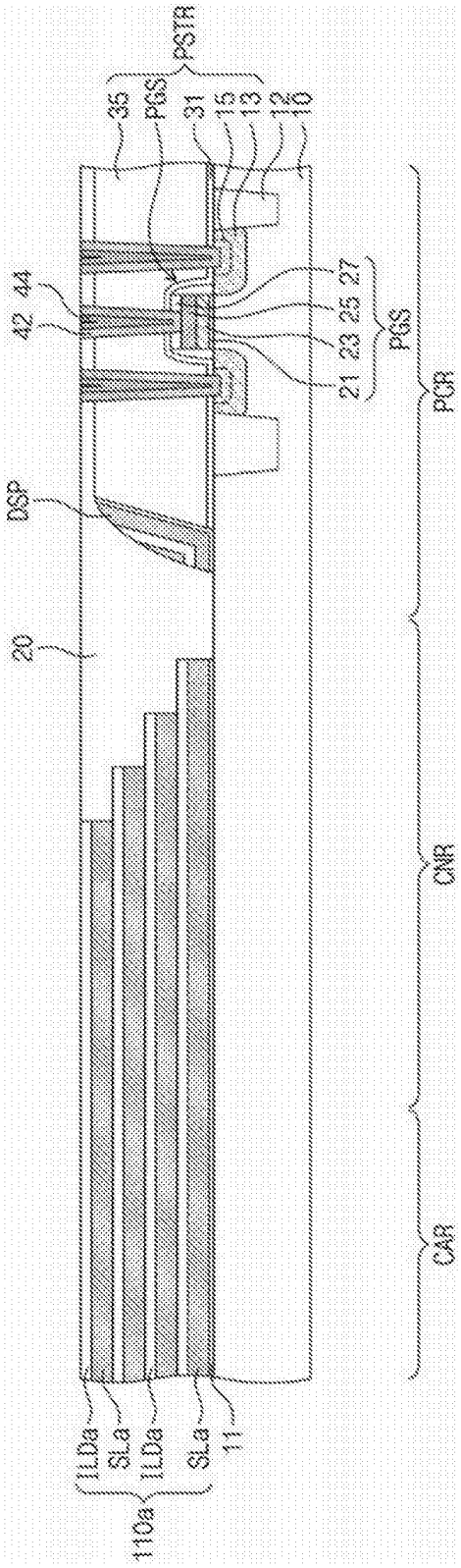


图8D

Micron Ex. 1002, p. 764
Micron v. YMTC
IPR2024-00789

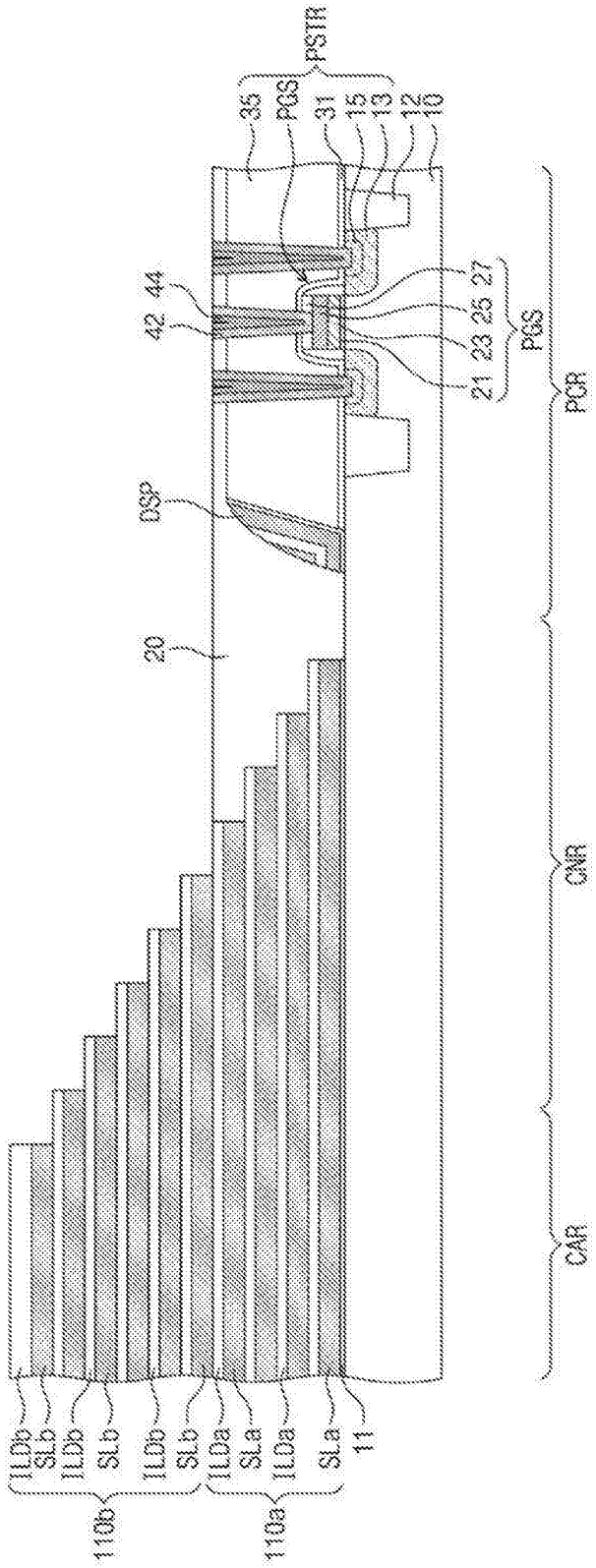


图8E

Micron Ex. 1002, p. 765
Micron v. YMTC
IPR2024-00789

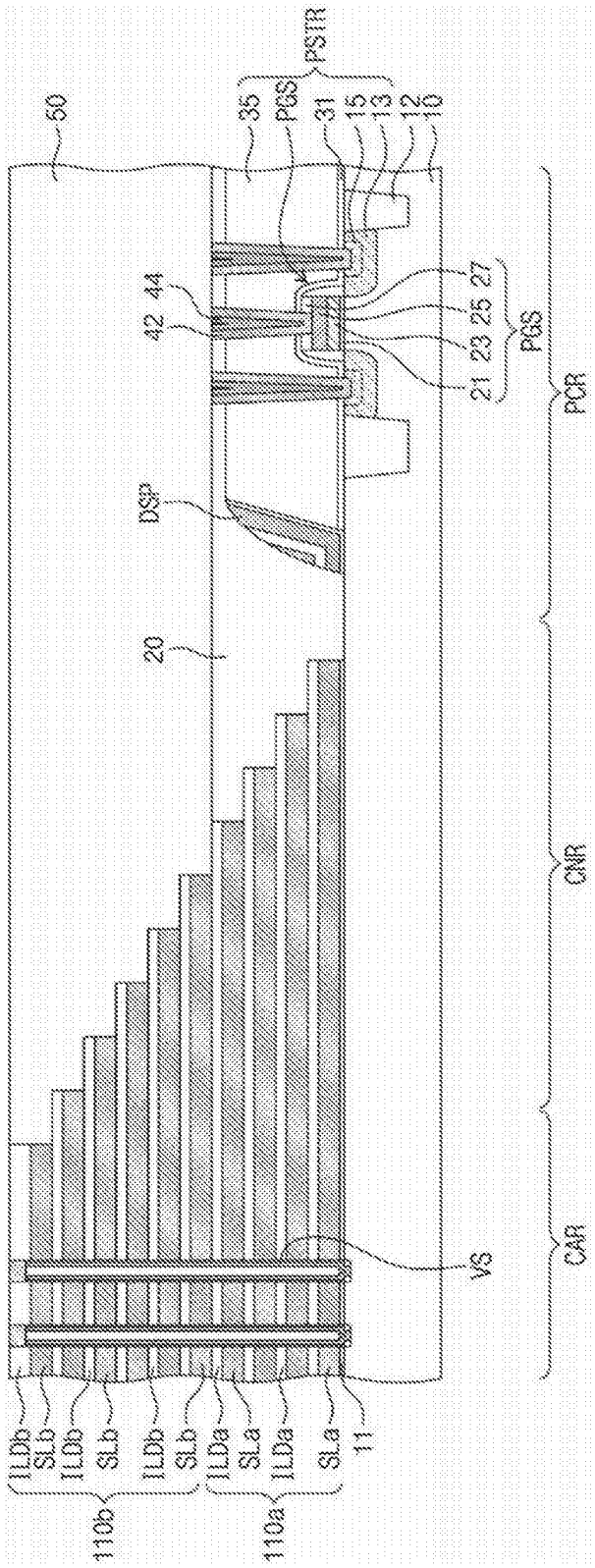


图8F

Micron Ex. 1002, p. 766
Micron v. YMTC
IPR2024-00789

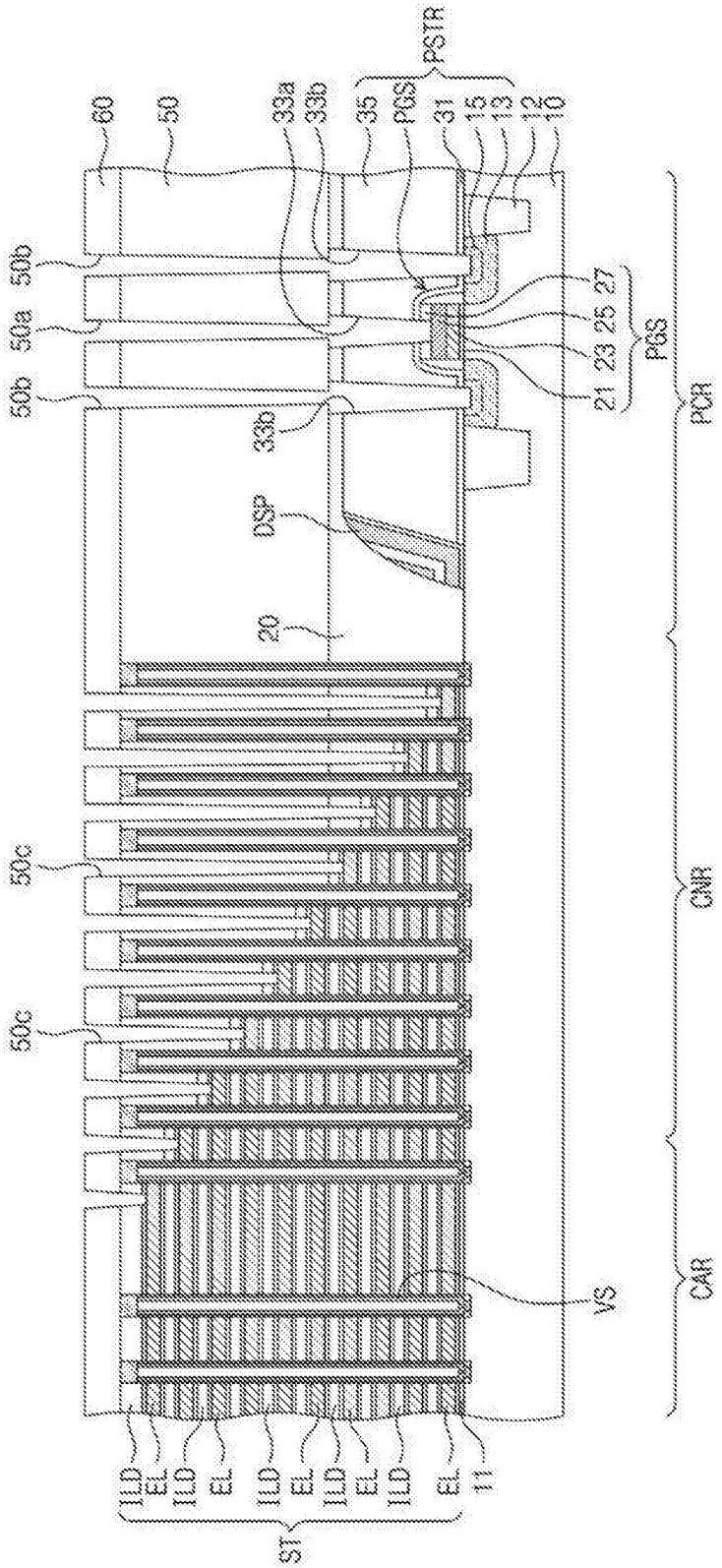


图8G

Micron Ex. 1002, p. 767
Micron v. YMTC
IPR2024-00789

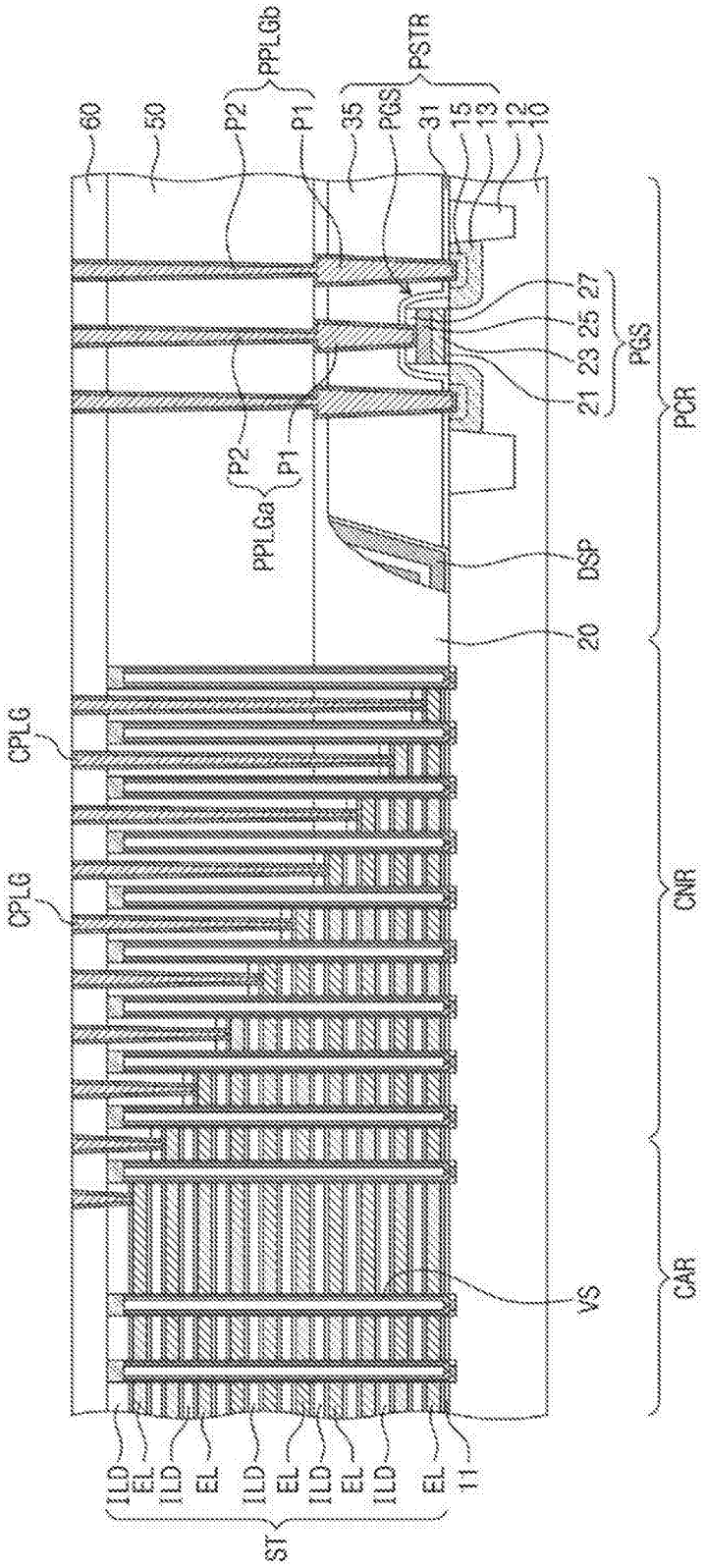


图8H

Micron Ex. 1002, p. 768
Micron v. YMTC
IPR2024-00789

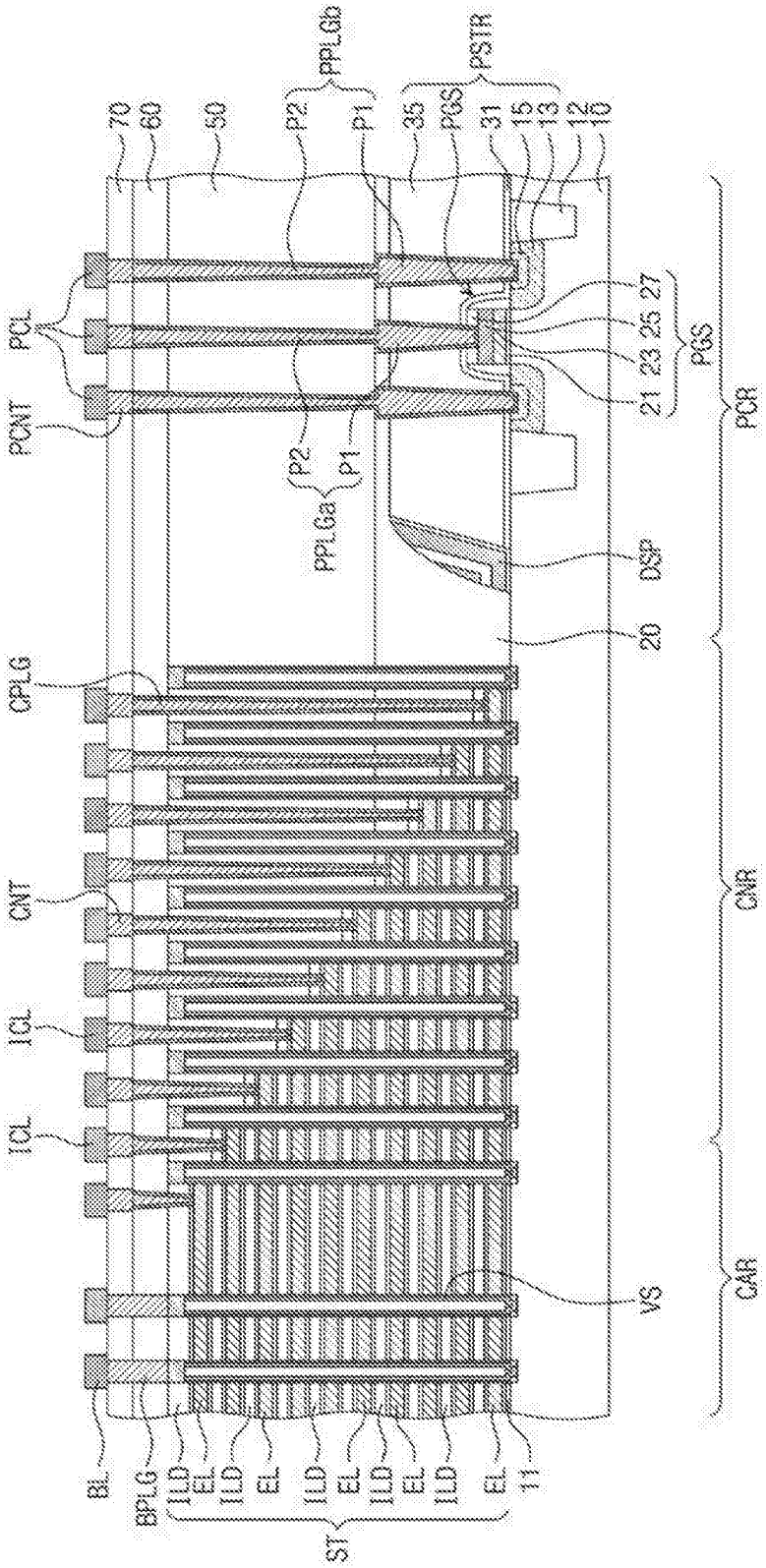


图8I

Micron Ex. 1002, p. 769
Micron v. YMTC
IPR2024-00789

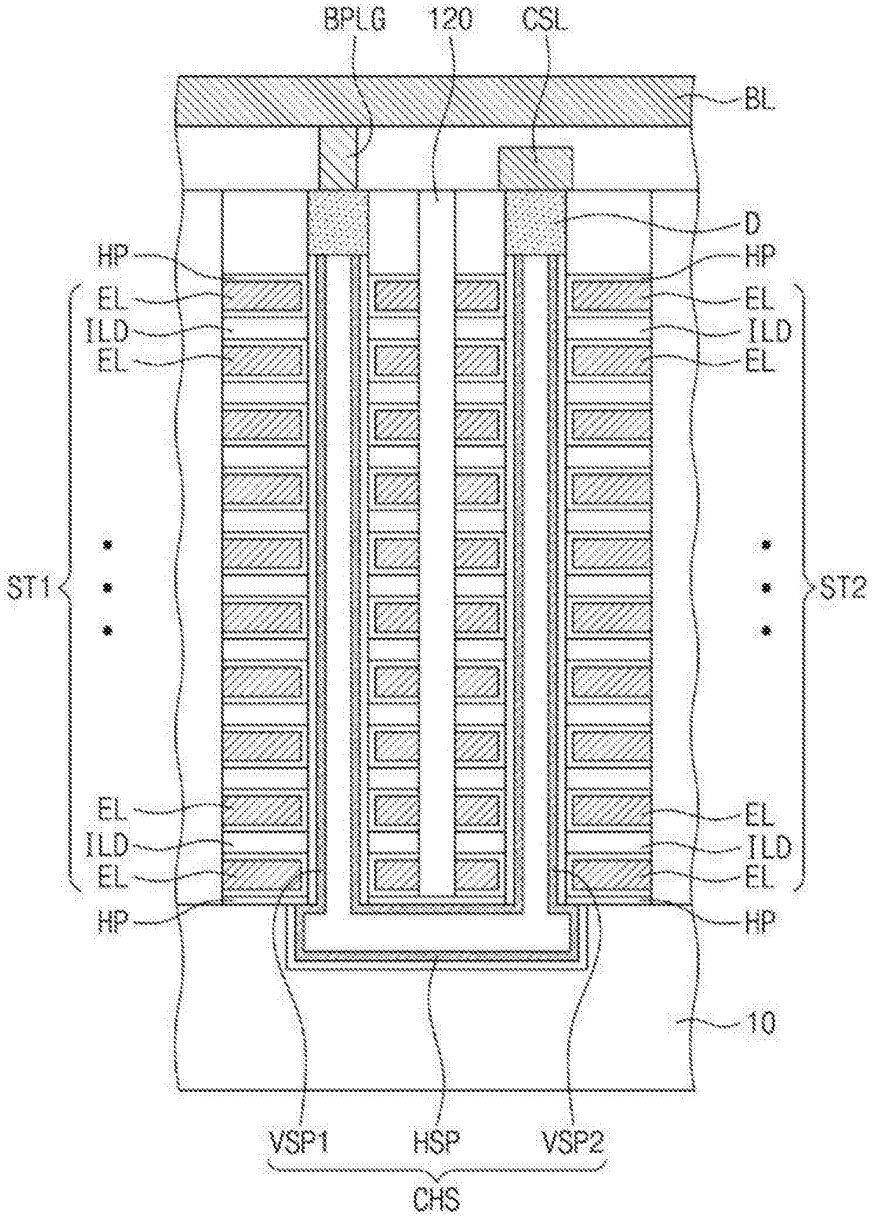


图9

Micron Ex. 1002, p. 770
Micron v. YMTC
IPR2024-00789

Doc Code: PET.AUTO Document Description: Petition automatically granted by EFS-Web	PTO/SB/140 U.S. Patent and Trademark Office Department of Commerce

Electronic Petition Request	**PETITION TO WITHDRAW AN APPLICATION FROM ISSUE AFTER PAYMENT OF THE ISSUE FEE UNDER 37 CFR 1.313(c)**
Application Number	16402202
Filing Date	02-May-2019
First Named Inventor	Wenyu Hua
Art Unit	2812
Examiner Name	SYED GHEYAS
Attorney Docket Number	10018-01-0045-US
Title	THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

An application may be withdrawn from issue for further action upon petition by the applicant. To request that the Office withdraw an application from issue, applicant must file a petition under this section including the fee set forth in § 1.17(h) and a showing of good and sufficient reasons why withdrawal of the application from issue is necessary.

APPLICANT HEREBY PETITIONS TO WITHDRAW THIS APPLICATION FROM ISSUE UNDER 37 CFR 1.313(c).

A grantable petition requires the following items:
(1) Petition fee; and
(2) One of the following reasons:
(a) Unpatentability of one or more claims, which must be accompanied by an unequivocal statement that one or more claims are unpatentable, an amendment to such claim or claims, and an explanation as to how the amendment causes such claim or claims to be patentable;
(b) Consideration of a request for continued examination in compliance with § 1.114 (for a utility or plant application only); or
(c) Express abandonment of the application. Such express abandonment may be in favor of a continuing application, but not a CPA under 37 CFR 1.53(d).

Petition Fee

- ○ Small Entity
- ○ Micro Entity
- ◉ Regular Undiscounted

Reason for withdrawal from issue

○ One or more claims are unpatentable ◉ Consideration of a request for continued examination (RCE) (List of Required Documents and Fees) ○ Applicant hereby expressly abandons the instant application (any attorney/agent signing for this reason must have power of attorney pursuant to 37 CFR 1.32(b)).	
RCE request,submission, and fee. ☐ I certify, in accordance with 37 CFR 1.4(d)(4) that : The RCE request ,submission, and fee have already been filed in the above-identified application on ☒ Are attached.	
THIS PORTION MUST BE COMPLETED BY THE SIGNATORY OR SIGNATORIES I certify, in accordance with 37 CFR 1.4(d)(4) that I am: ◉ An attorney or agent registered to practice before the Patent and Trademark Office who has been given power of attorney in this application. ○ An attorney or agent registered to practice before the Patent and Trademark Office, acting in a representative capacity. ○ A sole inventor ○ A joint inventor; I certify that I am authorized to sign this submission on behalf of all of the inventors as evidenced by the power of attorney in the application ○ A joint inventor; all of whom are signing this e-petition	
Signature	/Zhiwei Zou/
Name	Zhiwei Zou
Registration Number	66041

Electronic Patent Application Fee Transmittal

Application Number:	16402202
Filing Date:	02-May-2019
Title of Invention:	THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME
First Named Inventor/Applicant Name:	Wenyu Hua
Filer:	Zhiwei Zou/Anna O'Connor
Attorney Docket Number:	10018-01-0045-US

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
PETITION FEE- 37 CFR 1.17(H) (GROUP III)	1464	1	140	140
RCE- 1ST REQUEST	1801	1	1360	1360
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**1500**

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

Decision Date : October 30, 2020

In re Application of :

Wenyu Hua

DECISION ON PETITION

UNDER CFR 1.313(c)(2)

Application No : 16402202

Filed : 02-May-2019

Attorney Docket No : 10018-01-0045-US

This is an electronic decision on the petition under 37 CFR 1.313(c)(2), filed October 30, 2020 , to withdraw the above-identified application from issue after payment of the issue fee.

The petition is **GRANTED.**

The above-identified application is withdrawn from issue for consideration of a submission under 37 CFR 1.114 (request for continued examination). See 37 CFR 1.313(c)(2).

Petitioner is advised that the issue fee paid in this application cannot be refunded. If, however, this application is again allowed, petitioner may request that it be applied towards the issue fee required by the new Notice of Allowance.

Telephone inquiries concerning this decision should be directed to the Patent Electronic Business Center (EBC) at 866-217-9197.

This application file is being referred to Technology Center AU 2812 for processing of the request for continuing examination under 37 CFR 1.114 .

Office of Petitions

Electronic Acknowledgement Receipt

EFS ID:	41003286
Application Number:	16402202
International Application Number:	
Confirmation Number:	7288
Title of Invention:	THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME
First Named Inventor/Applicant Name:	Wenyu Hua
Customer Number:	153988
Filer:	Zhiwei Zou/Anna O'Connor
Filer Authorized By:	Zhiwei Zou
Attorney Docket Number:	10018-01-0045-US
Receipt Date:	30-OCT-2020
Filing Date:	02-MAY-2019
Time Stamp:	23:01:30
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$1500
RAM confirmation Number	E20200TN01234819
Deposit Account	602227
Authorized User	Anna O'Connor

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.21 (Miscellaneous fees and charges)

37 CFR 1.20 (Post Issuance fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Quick Path Information Disclosure Statement	10018-01-0045-US_RequestQPIDS.pdf	207510 193d6a94be0ce9ec7869475ae848ae815a6a3ff5	no	1
Warnings:					
Information:					
2	Request for Continued Examination (RCE)	10018-01-0045-US_RCE.pdf	697606 45a1729538edc4eed0861bbd6bc7e43fbc0407c3	no	3
Warnings:					
Information:					
3	Information Disclosure Statement (IDS) Form (SB08)	10018-01-0045-US_IDS4_SB082.pdf	612227 97e202d544648d8087fecb8ab4be4afa4673ff97	no	4
Warnings:					
Information:					
A U.S. Patent Number Citation or a U.S. Publication Number Citation is required in the Information Disclosure Statement (IDS) form for autoloading of data into USPTO systems. You may remove the form to add the required data in order to correct the Informational Message if you are citing U.S. References. If you chose not to include U.S. References, the image of the form will be processed and be made available within the Image File Wrapper (IFW) system. However, no data will be extracted from this form. Any additional data such as Foreign Patent Documents or Non Patent Literature will be manually reviewed and keyed into USPTO systems.					
4	Foreign Reference	CN108886040_Bibl.pdf	16561670 9d78133f8ec8ab56cb0994f6457161e96d5ee64d	no	57
Warnings:					
Information:					
5	Foreign Reference	CN104157654_Bibl.pdf	3478924 b3cc7f357208bac87a756a9762b34715354575f0	no	17
Warnings:					
Information:					

6	Foreign Reference	WO2018071143A2.pdf	22383153 e74f5cd1cd7e7077e020010de8aa6a89772becc0	no	93
Warnings:					
Information:					
7	Foreign Reference	CN108695336_Bibl.pdf	10674009 de0ff4db6e5e4994292312526d1d87c3de9a4e5e	no	53
Warnings:					
Information:					
8	Petition automatically granted by EFS	petition-request.pdf	31574 ddb4f8fdefcc58f81cb45aa614b788543105a9eb	no	2
Warnings:					
Information:					
9	Fee Worksheet (SB06)	fee-info.pdf	32145 7bed84c816594fe83961d4b817ad4367f5abb4ee	no	2
Warnings:					
Information:					
			Total Files Size (in bytes):	54678818	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.
<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.
<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Doc Code: QPIDS.REQ
Document Description: Quick Path Information Disclosure Statement

PTO/SB/09 (08-14)

CERTIFICATION AND REQUEST FOR CONSIDERATION OF AN INFORMATION DISCLOSURE STATEMENT FILED AFTER PAYMENT OF THE ISSUE FEE UNDER THE QPIDS PILOT PROGRAM

Non-Provisional Application Number: 16402202	Filing Date: May 2, 2019
First Named Inventor: Wenyu HUA	Title of Invention: THREE-DIMENSIONAL MEMORY DEVICE AND METHOD FOR FORMING THE SAME

THE UNDERSIGNED HEREBY CERTIFIES AND REQUESTS THE FOLLOWING FOR THE ABOVE-IDENTIFIED APPLICATION.

1. Consideration is requested of the information disclosure statement (IDS) submitted herewith, which is being filed after payment of the issue fee.
2. Check the box next to the appropriate selection:

   ☑ Each item of information contained in the IDS was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the IDS. See 37 CFR 1.97(e)(1).

   OR

   ☐ No item of information contained in the IDS was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the IDS was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the IDS. See 37 CFR 1.97(e)(2).
3. Please charge the IDS fee set forth in 37 CFR 1.17(p) to Deposit Account No. 60-2227.
4. A Petition to Withdraw from Issue After Payment of the Issue Fee (37 CFR 1.313(c)(2)), including the petition fee set forth in 37 CFR 1.17(h), is submitted herewith as a **Web-based ePetition**.
   WARNING: Do not submit the petition as a follow-on paper via EFS-Web. Submit the petition as a Web-based ePetition by signing on to EFS-Web as a registered user, selecting the radio button next to "Existing application/patent," and then selecting the radio button next to "ePetition (for automatic processing and immediate grant, if all petitions requirements are met)." Failure to use the Web-based ePetition interface will result in automatic entry of the RCE.
5. A request for continued examination (RCE) under 37 CFR 1.114 and the RCE fee under 37 CFR 1.17(e) are submitted herewith.
6. The RCE will be treated as a "conditional" RCE. In the event the examiner determines that any item of information contained in the IDS necessitates the reopening of prosecution in the application, the undersigned understands that (i) the RCE will be processed and treated as an RCE under 37 CFR 1.114 and therefore (ii) the IDS fee under 37 CFR 1.17(p) will be returned in accordance with 37 CFR 1.97(b)(4). In the event that no item of information in the IDS necessitates reopening prosecution, the undersigned understands that the RCE will not be processed and the RCE fee under 37 CFR 1.17(e) will be returned.
7. This certification and request is being filed as a **Web-based ePetition** and is not accompanied by an amendment to the application. Inclusion of an amendment will result in automatic entry of the RCE.

Signature /Zhiwei Zou/	Date 10/30/2020
Name (Print/Typed) Zhiwei Zou	Practitioner Registration Number (If applicable) 66,041

Note: *Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required in accordance with 37 CFR 1.33 and 11.18. Please see 37 CFR 1.4(d) for the form of the signature. If necessary, submit multiple forms for more than one signature, see below.**

☑ *Total of 1 forms are submitted.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

United States Patent and Trademark Office

Office of the Chief Financial Officer

Document Code:WFEE

User :75320

Sale Accounting Date:11/02/2020

Sale Item Reference Number	Effective Date
16402202	10/30/2020

Document Number	Fee Code	Fee Code Description	Amount Paid	Payment Method
I2020A2C12391491	1806	SUBMISSION- INFORMATION DISCLOSURE STMT	$260.00	Deposit Account

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/402,202	05/02/2019	Wenyu Hua	10018-01-0045-US	7288

153988 7590 11/09/2020
Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

EXAMINER
GHEYAS, SYED I

ART UNIT	PAPER NUMBER
2812	

NOTIFICATION DATE	DELIVERY MODE
11/09/2020	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

USPTO@dockettrak.com
eofficeaction@appcoll.com
uspto@bayes.law

PTOL-90A (Rev. 04/07)

Corrected ***Notice of Allowability***	**Application No.** 16/402,202	**Applicant(s)** Hua et al.	
	Examiner SYED I GHEYAS	**Art Unit** 2812	**AIA (FITF) Status** Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☑ This communication is responsive to 10/30/2020.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☑ The allowed claim(s) is/are 1-14 and 21-26 . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information , please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐All b) ☐ Some *c) ☐ None of the:

   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____ .

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☑ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 10/30/2020.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.
5. ☐ Examiner's Amendment/Comment
6. ☑ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/SYED I GHEYAS/
Primary Examiner, Art Unit 2812

Micron Ex. 1002, p. 782
Micron v. YMTC
IPR2024-00789

Notice of Pre-AIA or AIA Status

1. The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

Information Disclosure Statement

1. The information disclosure statement (IDS) submitted on October 30, 2020 was filed after the mailing date of the Notice of Allowance on 10/26/2020. The submission is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

Allowable Subject Matter

Claims ***1-14 & 21-26*** are allowed.

The following is an examiner's statement of reasons for allowance: The reasons for allowance have not changed since the previous Office Action (Notice of Allowance, mailed on 10/26/2020).

Conclusion

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to SYED GHEYAS whose telephone number is (571)272-0592. The examiner can normally be reached on M to F: 7:30 AM to 5 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Charles Garber can be reached on 571-272-2194. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

11/04/2020
/SYED I GHEYAS/
Primary Examiner, Art Unit 2812

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Not for submission under 37 CFR 1.99)

Application Number	16402202
Filing Date	2019-05-02
First Named Inventor	Wenyu HUA
Art Unit	2812
Examiner Name	GHEYAS, SYED I.
Attorney Docket Number	10018-01-0045-US

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	108886040	CN	A	2018-11-23	SANDISK TECHNOLOGIES LLC		
	2	104157654	CN	A	2014-11-19	INST OF MICROELECTRONICS CAS		
	3	2018071143	WO	A2	2018-04-19	MONOLITHIC 3D INC.		

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /S.I.G/

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16402202
	Filing Date	2019-05-02
	First Named Inventor	Wenyu HUA
	Art Unit	2812
	Examiner Name	GHEYAS, SYED I.
	Attorney Docket Number	10018-01-0045-US

	4	108695336	CN	A	2018-10-23	SAMSUNG ELECTRONICS CO LTD		

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1		

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/SYED I GHEYAS/	Date Considered	11/03/2020

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /S.I.G/

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16402202
	Filing Date	2019-05-02
	First Named Inventor	Wenyu HUA
	Art Unit	2812
	Examiner Name	GHEYAS, SYED I.
	Attorney Docket Number	10018-01-0045-US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

X That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

X The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Zhiwei Zou/	Date (YYYY-MM-DD)	2020-10-30
Name/Print	Zhiwei Zou	Registration Number	66,041

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /S.I.G/

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /S.I.G/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/402,202	12/08/2020	10861872	10018-01-0045-US	7288

153988 7590 11/18/2020

Bayes PLLC
1765 Greensboro Station Place
Suite 900
McLean, VA 22102

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 0 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

Wenyu Hua, Wuhan, CHINA;
YANGTZE MEMORY TECHNOLOGIES CO., LTD., Wuhan, CHINA;
Fandong Liu, Wuhan, CHINA;
Zhiliang Xia, Wuhan, CHINA;

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage and facilitate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

IR103 (Rev. 10/09)