

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

MICRON TECHNOLOGY, INC.,
MICRON SEMICONDUCTOR PRODUCTS, INC., and
MICRON TECHNOLOGY TEXAS, LLC,
Petitioner

v.

BESANG INC.,
Patent Owner.

Case No. IPR2023-00900
U.S. Patent No. 7,378,702

**DECLARATION OF DR. JACK C. LEE IN SUPPORT
OF PETITIONER'S REPLY TO PATENT OWNER'S RESPONSE**

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. BACKGROUND AND QUALIFICATIONS	1
III. INFORMATION CONSIDERED	1
IV. RELEVANT LEGAL STANDARDS	5
V. SUMMARY OF OPINIONS	5
VI. CLAIM CONSTRUCTION	6
A. “stackable add-on layer”	6
B. “SOI pillar”	15
VII. GROUNDS 5-6: CLAIMS 13-17 ARE OBVIOUS OVER LEE IN VIEW OF WATANABE AND FURTHER IN VIEW OF ENDOH.....	22
A. A POSITA would have been motivated to combine the teachings of Lee and Watanabe and would have had a reasonable expectation of success.....	22
B. A POSITA would have been motivated to combine the teachings of Lee, Watanabe, and Endoh and would have had a reasonable expectation of success.....	33
VIII. GROUNDS 3-4: CLAIMS 13-17 ARE OBVIOUS OVER LEE IN VIEW OF ENDOH	40
IX. GROUNDS 1-2: CLAIMS 13-17 ARE OBVIOUS OVER WATANABE IN VIEW OF ENDOH	43
A. Watanabe discloses a nonvolatile memory cell	43
B. Watanabe in view of Endoh discloses an “SOI pillar”	45
C. Watanabe’s stacking method is enabled	47
X. CONCLUSION	47

LIST OF EXHIBITS

Ex-1001	U.S. Patent No. 7,378,702 (“the ’702 Patent”)
Ex-1002	Declaration of Dr. Jack C. Lee
Ex-1003	Curriculum Vitae of Dr. Jack C. Lee
Ex-1004	Prosecution History of U.S. Patent No. 7,378,702 (App. No. 10/934,270)
Ex-1005	U.S. Patent No. 5,091,762 (“Watanabe”)
Ex-1006	U.S. Patent Application Pub. No. US2002/0154556 (“Endoh”)
Ex-1007	U.S. Patent No. 6,881,994 (“Lee”)
Ex-1008	U.S. Patent No. 7,221,008 (“Matamis”)
Ex-1009	G. Groeseneken et al., <i>Basics of Nonvolatile Semiconductor Memory Devices</i> , in <i>Nonvolatile Semiconductor Memory Technology: A Comprehensive Guide to Understanding and Using NVSM Devices</i> (William D. Brown & Joe Brewer eds., 1997)
Ex-1010	Philip M. Sailer et al., <i>Creating 3D Circuits Using Transferred Films</i> , IEEE Circuits & Devices (1997)
Ex-1011	U.S. Patent No. 5,808,938 (“Tran”)
Ex-1012	U.S. Patent No. 6,982,905 (“Nguyen”)
Ex-1013	U.S. Patent No. 6,807,098 (“Jeong”)
Ex-1014	Kee-Ho Yu et al., <i>Real-Time Microvision System with Three-Dimensional Integration Structure</i> , IEEE Int’l Conference on Multisensor Fusion and Integration for Intelligent Sys. (1996)
Ex-1015	U.S. Patent No. 5,374,564 (“Bruel”)
Ex-1016	Horacio Mendez, <i>Silicon-on-Insulator: SOI Technology and Ecosystem – Emerging SOI Applications</i> , SOI Indus. Consortium (2009)
Ex-1017	U.S. District Courts – Combined Civil and Criminal Federal Court Management Statistics (Dec. 31, 2022), available at: https://www.uscourts.gov/statistics-reports/federal-court-management-statistics-december-2022
Ex-1018	T. Mikolajick et al., <i>The Future of Charge Trapping Memories</i> , 2007 IEEE Int’l Symposium on VLSI Tech., Sys. & Applications (2007)
Ex-1019	Complaint, <i>BeSang Inc. v. Micron Tech., Inc.</i> , No. 2:23-cv-00028 (E.D. Tex. Jan. 23, 2023)
Ex-1020	Jim Handy, <i>3D NAND: Benefits of Charge Traps over Floating Gates</i> , The Memory Guy Blog (Nov. 22, 2013)
Ex-1021	E. Bertagnolli et al., <i>ROS: An Extremely High Density MASK ROM Technology Based on Vertical Transistor Cells</i> , IEEE (1996)

Ex-1022	Tetsuo Endoh et al., <i>Novel Ultrahigh-Density Flash memory with A Stacked-Surrounding Gate Transistor (S-SGT) Structured Cell</i> , 50 IEEE Transactions on Electron Devices 945 (2003)
Ex-1023	F. Nemati & J.D. Plummer, <i>A Novel High Density, Low Voltage SRAM Cell with A Vertical NDR Device</i> , IEEE (1998)
Ex-1024	K.W. Lee et al., <i>Three-Dimensional Shared Memory Fabricated Using Wafer Stacking Technology</i> , IEEE (2000)
Ex-1025	<i>NAND vs. NOR Flash Memory</i> , Toshiba America Electronic Components, Inc. (2006)
Ex-1026	G.K. Celler & Sorin Crisloveanu, <i>Frontiers of Silicon-on-Insulator</i> , 93 J. Applied Physics 4955 (2003)
Ex-1027	Timeline, Soitec, available at: https://www.soitec.com/en/company/soitec-activity/history (last visited Apr. 25, 2023)
Ex-1028	G. Shahidi et al., <i>Fabrication of CMOS on Ultrathin SOI Obtained by Epitaxial Lateral Overgrowth and Chemical-Mechanical Polishing</i> , IEEE (1991)
Ex-1029	MicroChemicals, Further Processing of Silicon Wafers: SiO_2 , Si_3N_4 , Metallization, Epitaxy and Dicing, available at: https://www.microchemicals.com/technical_information/silicon_wafer_sio2_si3n4_metallization_epitaxy_soi.pdf (last visited Apr. 28, 2023)
Ex-1030	Hiroshi Sakuraba et al., <i>New Three-Dimensional High-Density Stacked-Surrounding Gate Transistor (S-SGT) Flash Memory Architecture Using Self-Aligned Interconnection Fabrication Technology Without Photolithography Process for Tera-Bits and Beyond</i> , 43 Japanese J. Applied Physics 2217 (2004)
Ex-1031	Letter dated September 21, 2023 to counsel for Patent Owner from John Kappos
Ex-1032	<i>BeSang, Inc. v. Micron Technology Inc. et al.</i> , 2:23-CV-00028-JRG-RSP, D.I. 95 (E.D. Tex. Dec. 20, 2023), Order Granting Stipulation and Joint Motion to Stay Pending <i>Inter Partes</i> Review (“Order Granting Stay”)
Ex-1033	Y. Beilliard et al., <i>Advances Toward Reliable High Density Cu–Cu Interconnects by Cu-SiO₂ Direct Hybrid Bonding</i> , 2014 International 3D Systems Integration Conference (3DIC) (2014)
Ex-1034	Perrine Batude et al., <i>3-D Sequential Integration: A Key Enabling Technology for Heterogeneous Co-Integration of New Function With CMOS</i> , 2 IEEE J. Emerging and Selected

	Topics in Circuits and Systems 714-22 (2012)
Ex-1035	Rafael Reif et al., <i>Technology and Applications of Three-Dimensional Integration</i> , 2004 Electrochemical Society Proceedings Volume 261-76 (2004)
Ex-1036	Dr. Lee's Deposition Errata Sheet
Ex-1037	SEM Image
Ex-1038	Two-level Stack of Watanabe's Fig. 7
Ex-1039	Four-level Stack of Watanabe's Fig. 7
Ex-1040	<i>Certain Semiconductor Chips and Prods. Containing Same</i> , ITC Inv. No. 337-TA-753, 2012 WL 927056 (Mar. 2, 2012)
Ex-1041	Declaration of Dr. Jack C. Lee in Support of Petitioner's Reply to Patent Owner's Response
Ex-1042	Email thread from Sang-Yun Lee, dated Nov. 8-26, 2007
Ex-1043	Sang-Yun Lee & Junil Park, <i>Architecture of 3D Memory Cell Array on 3D IC</i> , IEEE (2012)
Ex-1044	Email thread from Scott DeBoer (Micron), dated Oct. 10, 2017
Ex-1045	Email from Scott DeBoer (Micron), dated July 19, 2016
Ex-1046	Sang-Yun Lee & Junil Park, <i>History and Trend of 3-Dimensional Integrated Circuit</i>
Ex-1047	International Technology Roadmap for Semiconductors: Front End Processes (2003)
Ex-1048	U.S. Patent No. 6,951,798
Ex-1049	Gerald Lucovsky et al., <i>Integration of Plasma-Assisted and Rapid Thermal Processing for Low-Thermal Budget Preparation of Ultra-Thin Dielectrics for Stacked-Gate Device Structures</i> , 33 Jpn. J. Appl. Phys. Vol. 7061-70 (1994)
Ex-1050	J. Wong et al., <i>Rapid Thermal Annealing of Deposited SiO₂ Films</i> , 35 Mat. Res. Soc. Symp. Proc. 514-20 (1985)
Ex-1051	Dr. Steven A. Przybylski Deposition Transcript (Apr. 5, 2024)
Ex-1052	D.E. Mercer et al., <i>Will Thermal Budget Really Matter in the Future</i> , 2001-9 Electrochemical Soc. Proceedings 247 (2000)
Ex-1053	Jürgen Regner, <i>An Analytical Approach to Quantify the Thermal Budget in Consideration of Consecutive Thermal Process Steps</i> , 10 th IEEE Int'l Conference on Advanced Thermal Processing of Semiconductors – RTP 15-20 (2002)
Ex-1054	R. L. Smith & S. D. Collins, <i>A Wafer-to-wafer Alignment Technique</i> , 20 Sensors & Actuators 315-16 (1989)
Ex-1055	Frank Niklaus et al., <i>A Method to Maintain Wafer Alignment Precision During Adhesive Wafer Bonding</i> , 107 Sensors & Actuators

	273-78 (2003)
Ex-1056	Updated Guidance for Making a Proper Determination of Obviousness, 89 Fed. Reg. 14,449-453 (Feb. 27, 2024), https://www.federalregister.gov/documents/2024/02/27/2024-03967/updated-guidance-for-making-a-proper-determination-of-obviousness
Ex-1057	Roberto Bez et al., <i>Introduction to Flash Memory</i> , 91 Proc. IEEE 489-502 (2003)
Ex-1058	Paolo Pavan, <i>Flash Memory Cells—An Overview</i> , 85 Proc. IEEE 1248 (1997)
Ex-1059	Keith Thompson, <i>Electromagnetic Fast-Firing for the Rapid Thermal Processing of Silicon</i> (Feb. 25, 2003) (P.D. dissertation, University of Wisconsin-Madison)
Ex-1060	Keith Thompson et al., <i>Direct Silicon-Silicon Bonding by Electromagnetic Induction Heating</i> , 11 J. Microelectromechanical Sys. 285-92 (2002)
Ex-1061	U.S. Patent No. 5,519,236
Ex-1062	U.S. Patent No. 5,932,920

1. I, Jack C. Lee, declare as follows:

I. INTRODUCTION

2. I provide the present declaration to supplement the opinions provided in my initial declaration (Ex-1002) provided on May 3, 2023, for this proceeding (“my initial declaration”). My opinions herein further confirm the opinions and conclusions in my initial declaration and deposition of February 6, 2024, and respond to the assertions in Dr. Przybylski’s declaration (Ex-2029) and Patent Owner’s Response.

3. I continue to be compensated at my ordinary and customary consulting rate for my work. As before, my compensation is in no way contingent on the nature of my findings, the presentation of my findings in testimony, or the outcome of this or any other proceeding. I have no other interest in this proceeding.

II. BACKGROUND AND QUALIFICATIONS

4. My qualifications are set forth in my initial declaration, provided on May 3, 2023, which I incorporate herein by reference. The only change is that in September 2023, I was promoted to the Professor Emeritus level.

III. INFORMATION CONSIDERED

5. In addition to the materials and information listed in my initial declaration, and in addition to my knowledge and experience, I reviewed the following information in forming my opinions and conclusions stated herein:

a. Paper 6: Patent Owner’s Preliminary Response

- b. Paper 7: Petitioner's Preliminary Reply In Support of *Inter Partes* Review
- c. Paper 8: Patent Owner's Preliminary Sur-Reply
- d. Paper 9: Board's Institution Decision
- e. Paper 18: Patent Owner's Response to Petition for *Inter Partes* Review
- f. Ex-2022: Executive Summary, ITRS (2003)
- g. Ex-2028: Dr. Jack C. Lee Deposition Transcript (Feb. 6, 2024)
- h. Ex-2029: Declaration of Steven A. Przybylski
- i. Ex-1033: Y. Beilliard et al., *Advances Toward Reliable High Density Cu–Cu Interconnects by Cu-SiO₂ Direct Hybrid Bonding*, 2014 International 3D Systems Integration Conference (3DIC) (2014)
- j. Ex-1034: Perrine Batude et al., *3-D Sequential Integration: A Key Enabling Technology for Heterogeneous Co-Integration of New Function With CMOS*, 2 IEEE J. Emerging and Selected Topics in Circuits and Systems 714-22 (2012)

- k. Ex-1035: Rafael Reif et al., *Technology and Applications of Three-Dimensional Integration*, 2004 Electrochemical Society Proceedings Volume 261-76 (2004)
- l. Ex-1047: International Technology Roadmap for Semiconductors: Front End Processes (2003)
- m. Ex-1048: U.S. Patent No. 6,951,798
- n. Ex-1049: Gerald Lucovsky et al., *Integration of Plasma-Assisted and Rapid Thermal Processing for Low-Thermal Budget Preparation of Ultra-Thin Dielectrics for Stacked-Gate Device Structures*, 33 Jpn. J. Appl. Phys. Vol. 7061-70 (1994)
- o. Ex-1050: J. Wong et al., *Rapid Thermal Annealing of Deposited SiO₂ Films*, 35 Mat. Res. Soc. Symp. Proc. 514-20 (1985)
- p. Ex-1051: Dr. Steven A. Przybylski Deposition Transcript (Apr. 5, 2024)
- q. Ex-1052: D.E. Mercer et al., *Will Thermal Budget Really Matter in the Future*, 2001-9 Electrochemical Soc. Proceedings 247 (2000)
- r. Ex-1053: Jürgen Regner, *An Analytical Approach to Quantify the Thermal Budget in Consideration of Consecutive Thermal Process*

Steps, 10th IEEE Int'l Conference on Advanced Thermal Processing of Semiconductors – RTP 15-20 (2002)

s. Ex-1054: R. L. Smith & S. D. Collins, *A Wafer-to-wafer Alignment Technique*, 20 Sensors & Actuators 315-16 (1989)

t. Ex-1055: Frank Niklaus et al., *A Method to Maintain Wafer Alignment Precision During Adhesive Wafer Bonding*, 107 Sensors & Actuators 273-78 (2003)

u. Ex-1057: Roberto Bez et al., *Introduction to Flash Memory*, 91 Proc. IEEE 489-502 (2003)

v. Ex-1058: Paolo Pavan, *Flash Memory Cells—An Overview*, 85 Proc. IEEE 1248 (1997)

w. Ex-1059: Keith Thompson, *Electromagnetic Fast-Firing for the Rapid Thermal Processing of Silicon* (Feb. 25, 2003) (Ph.D. dissertation, University of Wisconsin-Madison)

x. Ex-1060: Keith Thompson et al., *Direct Silicon-Silicon Bonding by Electromagnetic Induction Heating*, 11 J. Microelectromechanical Sys. 285-92 (2002)

y. Ex-1061: U.S. Patent No. 5,519,236

z. Ex-1062: U.S. Patent No. 5,932,920

6. I have also relied on my own education, training, experience, and knowledge in the relevant fields.

7. I may rely on any of these materials, experiences, and knowledge, in addition to the evidence specifically cited as supportive examples in particular sections of this declaration, as additional support for my opinions.

8. I reserve the right to supplement or amend this declaration if additional facts and information that affect my opinions become available.

IV. RELEVANT LEGAL STANDARDS

9. I provided my understanding of the legal standards in my initial declaration. I rely on those legal standards herein where applicable.

10. In addition to the legal standards identified in my initial declaration, I have been informed and understand that patent drawings should not be presumed to be drawn to scale without an explicit statement that they are to scale or are tied to a quantitative value.

V. SUMMARY OF OPINIONS

11. For the reasons set forth here, in my deposition of February 6, 2024, and in my initial declaration, it is my opinion that under any interpretation of the claims, Claims 13-17 of the '702 Patent are obvious under the grounds set forth in my initial declaration. Ex-1002, Section X.

VI. CLAIM CONSTRUCTION

A. “stackable add-on layer”

12. I agree in substance with the Board’s preliminary construction of “stackable add-on layer,” but I proposed a slightly modified construction of “stackable add-on layer” during my deposition: “a ~~layer~~ structure ~~that is~~ formed separately from the substrate ~~and then~~ suitable to be bonded ~~as whole~~ to the dielectric layer positioned above the substrate.” See Paper 9 at 9 (differences marked); Ex-2028, 12:18-16:8, 56:12-59:11, 88:8-92:4, 115:1-118:6, 119:18-133:11. I further agree with the Board that “stackable add-on layer” does not include “memory cells that are formed by sequential deposition on or above a substrate without being added to a separately formed substrate.” Ex-1002, 13 n.2.

13. The intrinsic record supports my and the Board’s construction. The plain language of Claim 13 requires “a stackable add-on layer having a plurality of vertically oriented semiconductor memory cells.” Ex-1001, 12:29-30. The ’702 Patent’s specification states, “non-volatile memory cells, and similar structures may be formed in a stackable add-on layer for use in *conjunction* with the *separately fabricated* substrate.” *Id.* 2:4-9 (emphasis added). “The plurality of vertically oriented semiconductor memory cells can be added to the *separately fabricated* substrate as a thin layer including several doped semiconductor regions,” or “[a]lternatively, the plurality of vertically oriented semiconductor memory cells may

be completely fabricated prior to attachment.” *Id.*, Abstract (emphasis added); *see id.* 1:47-50. The ’702 Patent consistently states that “a stackable add-on layer” is to be formed separately from the substrate.

14. Further, as the Board noted, the ’702 Patent’s definition of “SOI” refers to “a relatively thin, single crystal portion of a semiconductor wafer that can be cleaved and bonded to another *previously fabricated* wafer, or similar type of substrate, such that a three dimensional stack is formed from the SOI layer and the previously fabricated wafer or similar type of substrate.” *Id.* 4:47-53 (emphasis added). “[T]he SOI layer may be thought of as an attachment layer, or stackable add-on structure which is *suitable for bonding* to a semiconductor substrate already containing devices and/or interconnections.” *Id.* 4:53-57 (emphasis added); *id.* 5:11-17 (defining “3D-IC” to include a semiconductor substrate and a stackable add-on layer “stacked and bonded to each other”). Thus, the ’702 Patent states that “a stackable add-on layer” is formed separately from the substrate and then stacked on and bonded to the substrate. Because my construction follows the ’702 specification closely, that construction is strongly supported. *See, e.g.*, Ex-1001, Abstract, 1:47-50, 2:4-9, 4:45-5:7, 5:11-17, 12:29-30.

15. The Board also asked the parties “to explain how a ‘stackable add-on layer’ is structurally distinguishable from a layer that is sequentially deposited.” Paper 9, 9 n.5. During my deposition, I identified three structural distinctions

between a “stackable add-on layer” and a sequentially-deposited structure.

16. First, a memory array formed through the stacking method often has “single-crystal silicon in the upper layer[s].” Ex-2028, 141:20-22. Because each layer in the stacking method is processed separately, one can start with monocrystalline silicon for each layer, in which “the speed of the transistor and, therefore, the speed of the circuit might be higher.” *Id.* 141:22-142:2; *see* Ex-1016, Ex-1028, Ex-1029. By contrast, the sequential deposition method typically deposits polycrystalline silicon by CVD in upper transistors. *See* Ex-1034 at 715, Fig. 5. Compared to monocrystalline silicon, transistors formed on polycrystalline or amorphous silicon would exhibit lower speed and, thus, the speed of the circuit might also be lower. Although polycrystalline silicon can be converted to monocrystalline silicon by heating to high temperatures, repeatedly processing the upper transistors at high temperatures will cause “dopant diffusion and interfacial oxide growth, which is detrimental to the characteristics of the transistors.” Ex-2028, 122:19-128:14.

17. For example, Figure 5 below from Batude illustrates different methods of forming “a crystalline semiconductor layer above processed transistors,” including a recrystallization technique using a seed window and wafer bonding. Ex-1034 at 715.

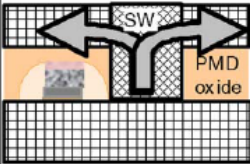
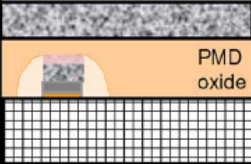
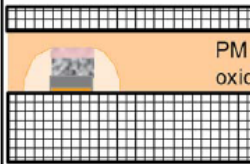
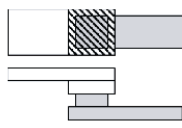
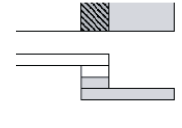
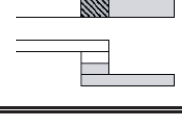
	Seed window (SW)	Poly-Si	Wafer bonding
Description			
Density	limited due to SW	Same than bottom level	Same than bottom level
Crystalline quality	Defect in SW region with controlled location	Random defects location	Perfect quality ~SOI supply quality
Thickness control	10s nm range	nm range	Å range
layer orientation	same orientation	random orientation for top substrate	different orientation possible
References	[5] [6] [7] [8] [9] [10]	[3] [11]	[12] [13] [14] [15]

Fig. 5. Description, benchmark, and main references for top active realization techniques.

Id., Fig. 5. Notably, the crystalline quality of the layers formed through the recrystallization technique using a seed window and sequential deposition of polysilicon result in defects, whereas the wafer bonding method results in “[p]erfect quality ~SOI supply quality.” *Id.* In addition, the wafer bonding method “does not require a seed window[] (SW) required in recrystallization techniques..., allowing hence a higher integration density” and “the bonding operation benefits from pristine crystalline quality and accurate thickness control.” *Id.* at 715. Thus, the stacking method enables one to use single-crystal silicon in the upper layers of the memory array, whereas sequential deposition results in upper layers of polycrystalline or amorphous semiconductor. *Id.*

18. Second, “the stacking method requires larger contact area” between the interconnects of the stacked layers (as compared with the sequential deposition method), Ex-2028, 142:3-7, 110:8-111:14, because there is a risk of misalignment using the stacking method. Ex-1034 at 714 (“Among 3-D technologies, 3-D sequential integration offers the highest alignment capability and the smallest contact size.”). Misalignment of the cells would lead to larger interconnect resistance and possibly render the device inoperable. Ex-2029 ¶80. A larger interconnect resistance would also result in lower speed for the circuit. Thus, a POSITA designing a memory array using the stacking method would use larger contact areas, sometimes referred to as landing pads. See Ex-1033 at 0003. For example, Table I below illustrates “the bonding pads of DCP1 located on the top wafer are larger than the ones on the bottom wafer. This design allows maintaining a constant contact area despite the bonding misalignment, which helps to better estimate the contact resistivity.” *Id.* at 0003, Table I.

Table I. Detailed characteristics of the three daisy chain structures.

Daisy chain	Interco. number	Pad width top / bottom (μm)	Pitch x / y (μm)	Top view and cross section
DCP1	7068	5 / 3	15 / 15	
DCP2	11 056	3 / 3	12 / 12	
DCP3	30 160	3 / 3	7 / 7	

Id., Table I. Further, Table III illustrates that the resistivity of DCP1, with larger contact areas, was far below the resistivity of the structures with smaller contact areas.

Table III. Electrical characterization results conducted on daisy chains DCP1, DCP2 and DCP3 after annealing at 200 and 400 °C for 2 h.

	DCP1	DCP2	DCP3
200 °C - α			
Functional yield (%)	99	96	96
Mean R_{exp} (Ω)	888.02	1873.86	3264.80
Std. deviation (%)	1.33	1.97	2.67
Contact resistivity	154 m Ω . μm^2		
400 °C - β			
Functional yield (%)	99	99	98
Mean R_{exp} (Ω)	873.48	1857.50	3227.54
Std. deviation (%)	0.99	0.84	1.37
Contact resistivity	123 m Ω . μm^2		

Id., Table III, 0006.

19. This structural difference is also illustrated by comparing Watanabe's Figs. 5E (sequential deposition) and 6D (stacking) below:

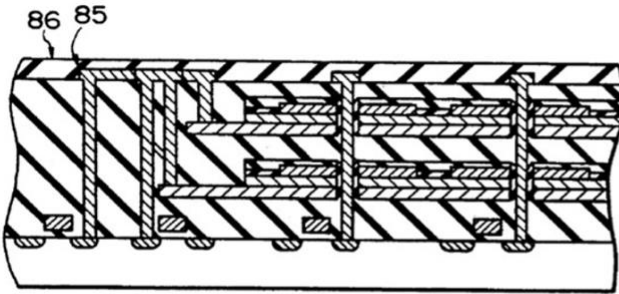


FIG. 5E

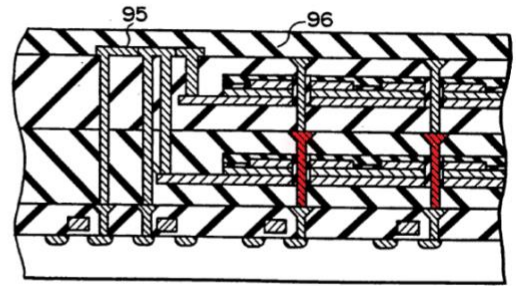


FIG. 6D

Ex-1005, Figs. 5E, 6D (showing enlarged contact areas as funnel-shaped landing pads annotated in red). Dr. Przybylski admitted that “[t]he landing pads of the stacking method...have no comparable structures in the figures associated with the lamination [*i.e.*, sequential deposition] method of Watanabe.” Ex-1051, 87:20-88:24, 85:10-86:19. Thus, the stacking method typically requires structures with larger contact areas and in which “the chip size might be larger and, therefore, the cost per memory cell might be higher and the...overall yield might be degraded.” Ex-2028, 142:3-7, 119:18-121:15; *see also* Ex-1033 at 0003, Table I.

20. Third, “there is an interfacial quality concern with the stacking method and not with sequential deposition method.” Ex-2028, 142:8-11. For example, when forming a three-dimensional integrated circuit through the stacking method, the interfaces between stacked layers are more prone to defects. *See id.* 132:3-12. Voids caused by air bubbles in the interface can cause surface roughness that leads to adhesion issues, resulting in weak bonding. *Id.* And “[i]f bonding strength is not good,” “over time, or even after fabrications, the layer might...start peeling off and that affects your yield and that affects your reliability.” *Id.* 131:7-12; *see* Ex-1033,

Figs. 1, 9; Ex-1035 at 268-71.

21. A POSITA would have recognized that interface defects occur using a stacking method because this was a well-known issue in the art. As shown in Figure 1 below, copper dishing causes gaps at the metal-to-metal interface (Figure 1(c)) that would exist unless the stack goes through a high-temperature annealing step (Figure 1(d)). See Ex-1033, Fig. 1, 0002-03.

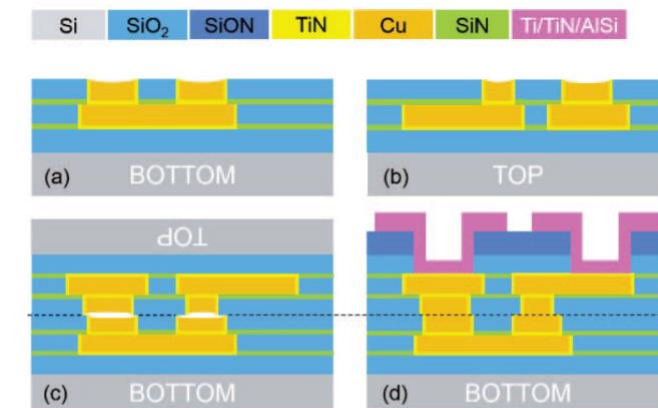


Figure 1. Test vehicle fabrication process: (a, b) successive damascene processes optimized for multilayer Cu-SiO₂ direct hybrid bonding, (c) bonding at room temperature, atmospheric pressure and ambient air, (d) annealing, Si removal and redistribution layer (RDL) realization.

Id., Fig. 1. However, even after a high-temperature annealing step, the bonding interface typically exhibits “circular voids mostly due to surface roughness and oxide diffusion (red arrows) and slit [voids] due to surface imperfections (blue arrows).” *Id.*, Fig. 9, 0005-06; Ex-1060, 290-91, Fig. 10.

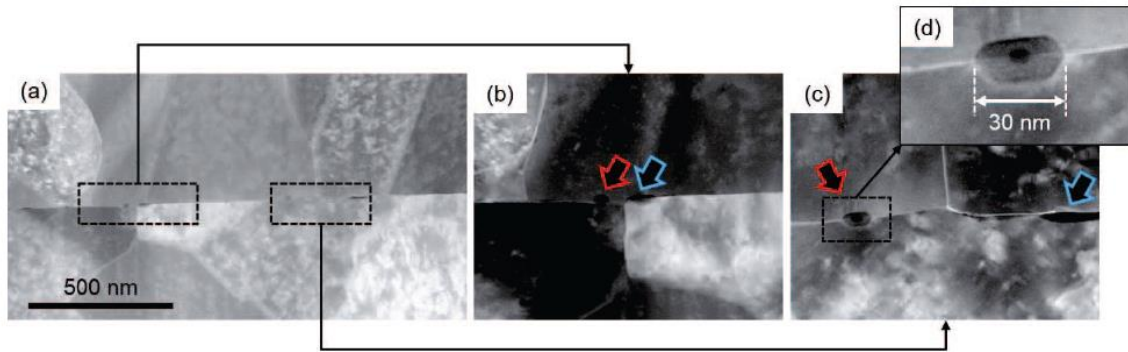


Figure 9. TEM characterization of the bonding interface of a DCP3 connection annealed at 400 °C for 2 h; (b)(c) Close up views of circular voids mostly due to surface roughness and oxide diffusion (red arrows) and slit ones due to surface imperfections (blue arrows). Copper grain growths at triple points are also visible; (d) Typical interfacial void.

Id., Fig. 9. Further, “[d]uring wafer bonding, trapped air bubbles can cause undesired interface voids and this will affect the interface quality of bonded wafer pair.” Ex-1035 at 270. “It is evident that voids form in this wafer pair. The voids form possibly due to out-gassing from the undensified PECVD oxide during post-bonding anneal.” *Id.* Thus, the physical differences in the layer interfaces present another structural distinction between the two approaches.

22. I discussed each of these structural differences and the literature that illustrates these differences during my deposition. Ex-2028, 110:8-111:14, 115:2-132:12, 141:2-144:9.

23. While PO states that it “disagrees with the Board’s preliminary construction,” PO does not offer an alternative. Paper 18 at 13. PO also does not address the construction I advanced during my deposition, Ex-2028, 12:18-16:8, which is consistent with the Board’s constructions. Dr. Przybylski also did not offer any opinion on this term. *See generally* Ex-2029; Ex-1051, 30:10-14 (“I was not asked to form an opinion on the term ‘stackable add-on layer.’”), 83:3-84:18.

B. “SOI pillar”

24. The ’702 Patent defines “SOI layer” “to refer to a relatively thin, *single crystal portion of a semiconductor wafer* that can be cleaved and bonded to *another previously fabricated wafer, or similar type of substrate.*” Ex-1001, 4:47-53 (emphasis added). “The *transferred*, unpatterned, SOI layer has only a vertically oriented stack of layers.” *Id.* 6:13-25 (emphasis added). “[A] FLD can be directly contacted by metal electrodes at top, bottom, and intermediate regions of the individual *doped stack structures (i.e., pillar structures).*” *Id.* 6:26-29 (emphasis added). Thus, the ’702 Patent’s specification defines “SOI pillar” to mean “a doped stack structure formed from a single-crystal portion of a semiconductor wafer transferred from one wafer to another previously fabricated wafer or similar type of substrate.” Ex-2028, 136:11-18.

25. I identified support for this construction during my deposition. *Id.* 79:20-80:17, 137:15-140:17. As explained in my testimony, “nowhere in the ’702 patent when they describe what a pillar is, mention[s] the dimension, whether it’s vertical dimension, horizontal dimension, whether vertical dimension is wider—larger than the horizontal dimensions or vice versa.” *Id.* 137:22-138:6. “So, therefore, if a device...is of broad horizontal structure, as long as it is a doped stack structure formed from a single-crystal portion of a semiconductor wafer transferred from one wafer to another previously fabricated wafer, or similar type of substrate,

whether it's broad or narrow...it will still be an SOI pillar.” *Id.* 138:7-15. This is because, by 2004, the horizontal dimension of a memory cell was harder to control during processing than the vertical dimension, which was smaller and easier to control. *See id.* 139:2-140:17 (“So, therefore, in general, the horizontal dimension...back in the 2003, 2004 timeframe are much larger than the vertical dimension.”).

26. Dr. Przybylski also conceded, “In general, the vertically oriented collections of doped silicon...Ns, Ps, or N+, P+, constitute SOI pillars.” Ex-1051, 64:3-7. PO’s argument that a “pillar” must be “taller than it is wide,” is not supported by record evidence. Taking a term from the patent specification out of context, PO argues that “the aspect ratio of SOI pillar 124” is “‘relatively small’ given its larger height relative to its width.” Paper 18 at 15 (quoting Ex-1001, 8:5-10). But a POSITA would have found “relatively small” to be unclear as to the numerical value. Moreover, Dr. Przybylski’s statement implies that he defines “aspect ratio” as width/height—otherwise the aspect ratio would not be “relatively small.” But this definition and Dr. Przybylski’s argument conflict with the correct meaning of “aspect ratio” in the semiconductor field, which is defined as height/width. Ex-1057 at 501 (defining “cell aspect ratio” as “the y pitch over the x pitch”), Fig. 29; *see* Ex-1061, 1:20-21 (defining “the aspect ratio (height/width) of the silicon column”); *see also* Ex-1062, 1:59 (same). Under the correct meaning of “aspect ratio,” PO’s

construction that calls for a pillar to be “taller than it is wide” requires an aspect ratio of greater than one, but this construction conflicts with the ’702 specification, which expresses a preference for “relatively small” aspect ratios. Ex-1001, 8:5-10. The ’702 Patent repeatedly seeks to reduce the aspect ratio, recognizing that it is undesirable to have a taller-than-wide structure. For example, the ’702 Patent explains that “[o]ne drawback of the structure in FIG. 20 is its high aspect ratio, which may cause toppling during processing.” Ex-1001, 9:19-23. As an alternative, the ’702 Patent describes a design in which “two stackable add-on layers...are vertically stacked as shown in FIG. 22” to resolve “the problem of high aspect ratio.” *Id.*, 9:20-23. As shown in the patent figure, the structure of Figure 22 is formed by separately forming and bonding “two stackable add-on layers” labeled as 101 and 102, each having its own “pillar” (i.e., doped stack structure, typically labeled as 124 throughout the patent). *See id.*, Fig. 22. Following the ’702 Patent’s expressed preference for “relatively small” aspect ratios, a POSITA would have understood how to minimize the height-to-width ratio of each stacked pillar. I am not aware of any teaching in the ’702 Patent’s specification or file history that requires or supports a height-to-width ratio being greater than one, and the ’702 specification actually conflicts with PO’s construction of “SOI pillar” insofar as it teaches the benefits of structures having a “relatively small” height to width.

27. I disagree with PO’s attempt to rely on the proportions shown in the

patent figures. I have been informed that patent drawings should not be presumed to be drawn to scale without an explicit statement that they are to scale or are tied to a quantitative value. As such, I understand that the figures in the '702 Patent, Lee, and Watanabe should not be presumed to be drawn to scale because none of these patents include any explicit statement that the figures are drawn to scale or provide quantitative dimensions. Ex-2028, 139:22-140:3; Ex-1051, 54:9-56:3; Ex-2029 ¶80. As I explained during my deposition, even though the figures “might [be] draw[n] in such a way that it looks like [the] vertical dimension is larger than [the] horizontal dimension...the actual reality is that it’s reversed. The vertical dimension is smaller than [the] horizontal dimension.” Ex-2028, 140:3-9. “Perhaps one reason why they do this is so that they can write n^+ , p , n^+ within the space that is allowed.” *Id.* 140:10-12. Thus, a POSITA reviewing the figures in the '702 Patent would not have understood the figures to be drawn to scale, but would have instead understood that the vertical dimensions are likely to be smaller than the horizontal dimensions.

28. PO’s construction is further called into question because Dr. Przybylski refused to provide any meaningful bounds for a “pillar” during his deposition—even when asked whether a single doped silicon region (*e.g.*, n^+ or p) of a memory cell could constitute an SOI pillar (it plainly cannot). Ex-1051, 64:3-67:7, 70:24-71:8, 71:25-72:7, 93:17-96:11. Thus, the '702 specification does not provide quantitative guidance to require an “SOI pillar” to be taller than it is wide.

29. PO cannot identify any intrinsic support for its construction, instead relying on common dictionary definitions outside the field of the relevant technology. Based on my experience with the relevant technology, a POSITA would not have consulted a common dictionary to understand the term “SOI pillar.” “SOI” is a specialized technical term not typically defined by common dictionaries, and the patent specification describes both “SOI” and “pillar” in the context of the relevant technology. Moreover, as explained in my deposition testimony, a POSITA around 2004—when the state-of-art process technology was at the 90-nanometer node—would have understood that the processing limitations in the horizontal direction had dimensions far greater than that in the vertical direction, resulting in a structure wider than it is tall. Ex-2028, 139:2-140:17; *see also* Ex-1047, Table 74a; Ex-2022, Table B; *see also* Ex-1034 at 715-16, Fig. 5.

30. To increase integration density, a POSITA would have been motivated to make transistors as small as possible by minimizing the area in the x-y plane and reducing the height in the z-dimension. *See* Ex-1022, 945-46, Fig. 2; *see also* Ex-1001, 3:51-61; Ex-1007, 11:18-23; Ex-1005, 1:49-52. As technology evolves, the optimization of pillar height becomes crucial for addressing various challenges associated with 3D NAND fabrication and operation. In general, the height of a pillar is related to the “channel length” of the metal-oxide-semiconductor field-effect transistor (MOSFET)—*i.e.*, the distance between the n⁺ (heavily-doped n-type)

source region and the n⁺ drain region. And the channel length of a MOSFET is the most critical dimension of the memory device's function. For example, a POSITA would have understood that *increasing* the channel length of a MOSFET would lead to slower programming speed, in part, because the electrons (electrical signals) have a further distance to travel. Increasing the channel length would also lead to decreased performance and slower programming speed because a larger channel length would produce lower internal electric field and, thereby, less energetic electrons (so-called hot electrons) would be injected into the floating gate or the charge-trap layer of the memory cell. A larger channel length would also decrease charge retention while increasing the opportunity for trapped electrons to leak away; increase parasitic capacitance, which slows reading speeds; and increase resistance, which increases the power required for reading and writing data.

31. In addition, taller pillars are more difficult to fabricate. The complexity and precision required to create structures with a higher height-to-width aspect ratio increase with the stack height. In general, taller pillars exhibit more issues with process variability, including layer misalignment and nonuniformity across layers. Taller pillars also use more material per cell and are more susceptible to mechanical stress and deformation, especially during thermal cycles of processing. *See* Ex-1001, 9:19-23 (“One drawback of the structure in FIG. 20 is its high aspect ratio, which may cause toppling during processing.”). Moreover, taller pillars require longer

processing time, *i.e.*, longer deposition times and longer etching times, which leads to lower manufacturing throughput and higher manufacturing costs. Thus, by increasing the pillar height, the structural stability of the pillars degrades, which leads to decreased device reliability, operational life, manufacturing yield, and cost efficiency.

32. Conversely, *reducing* the channel length would lower the parasitic capacitance and energy consumption per circuit function (or per memory function) while increasing charge retention, processing speed, and processing efficiency. Shorter pillars also enable designers to increase the number of stacks within the same footprint, which increases the overall memory storage density and lowers the cost per memory bit. Indeed, reducing the height of memory pillars in 3D NAND arrays not only tackles technical and physical limitations but also enhances performance metrics such as speed, power efficiency, and reliability—all of which are critical for meeting demands for increased scalability, better cost efficiency, and higher manufacturing yield.

33. Since at least 2004, I have taught courses at The University of Texas at Austin and short courses to industrial partners on semiconductor memory design and the effects of dimensional scaling (*e.g.*, reducing channel length and oxide thickness). By 2004, a POSITA would have understood that there are numerous benefits to making the memory cell shorter than it is wide and that it is

disadvantageous to make the memory cell taller than it is wide—including slowing operating speeds, consuming more power, worsening programming efficiency, and increasing overall manufacturing costs. *See* Ex-1001, 9:19-23 (explaining that a “high aspect ratio...may cause toppling during processing”). A POSITA would not, therefore, have understood “pillar” to require an aspect ratio of greater than one, as such a dimensional requirement would result in an inferior device and would run counter to common process technology parameters. PO has not shown that its construction of “SOI pillar” is supported by intrinsic or extrinsic evidence.

VII. GROUNDS 5-6: CLAIMS 13-17 ARE OBVIOUS OVER LEE IN VIEW OF WATANABE AND FURTHER IN VIEW OF ENDOH

A. A POSITA would have been motivated to combine the teachings of Lee and Watanabe and would have had a reasonable expectation of success

34. I disagree with PO’s argument that one would not apply Watanabe’s stacking method to Lee’s structure. Watanabe contemplates both stacking and sequential deposition¹ methods and explains that both approaches may be used to arrange memory cells “three-dimensionally, thereby providing a high integration density.” Ex-1005, 1:49-52. PO tries to distance Watanabe from Lee by arguing that

¹ PO does not dispute that Lee and Watanabe each disclose all elements of at least Claim 13 if the Board were to construe “stackable add-on layer” to encompass sequential deposition.

Lee expresses a preference for “monolithic” structures by comparing three-dimensional monolithic and two-dimensional non-monolithic structures. Ex-1007, 11:24-31. However, as explained in my deposition testimony, “A [POSITA] would understand that [a] three-dimensional array, whether you do it monolithically or non-monolithically, in other words, whether you use stacking method or sequential deposition method, would improve the device density.” Ex-2028, 62:20-63:4. Thus, a POSITA would have understood that Lee’s benefits derived from 3-D integration would be obtained if Lee’s memory structures were formed using Watanabe’s alternative stacking method. As discussed above and in my initial declaration, the stacking method results in three-dimensional memory devices with higher performance and higher speed.

35. Further, PO alleges two obstacles for which there were well-known solutions in the art within the skill of a POSITA by 2004. I disagree with PO’s arguments. First, PO argues that Watanabe’s high-temperature bonding would damage Lee’s memory transistors, Paper 18 at 24-27, but that is not so in view of the knowledge of a POSITA. When developing a semiconductor process, it is important to consider the effect that each process step may have on the structures that were previously fabricated. For example, field-effect transistors (FETs) are typically formed directly on the substrate surface. Then, subsequent structures are formed above the FETs, such as dielectric layers, metal interconnect layers, or, in

the case of a 3D device, one or more additional FET layers. If these subsequent process steps are performed at high temperatures, they may adversely impact the FETs already formed on the substrate surface, such as causing additional (and unintended) dopant diffusion or interfacial oxide growth. Thus, care must be taken such that these subsequent process steps do not impose excessive thermal effects on the FETs already formed in the device. The process of managing these thermal effects is referred to as “thermal budgeting.”

36. Generally, the thermal budget (*TB*) of a process is defined as the product of process temperature (*T*) and process time (*t*). See Ex-1052 at 247-48; Ex-2029 ¶71 (Dr. Przybylski acknowledged that thermal budget is a measure of both temperature and time.). The *TB* defines the maximum temperature-time a device can tolerate before the thermal effects would deteriorate its performance beyond acceptable metrics. As shown in the equation below, the *TB* associated with a high-temperature process at a shorter time would be the same as that of a low-temperature process for a longer period of time.

$$TB = T \cdot t,$$

where the process temperature (*T*) is in the unit of *Kelvin* and the process time (*t*) is in the unit of *seconds*.

37. For a multi-step process, *e.g.*, *n* number of process steps in a process sequence, thermal budget is the cumulative sum of the product of process

temperature (T) and process time (t) of each step. See Ex-1053 at 15-18.

$$TB = \sum_1^n T \cdot t = (T_1 \cdot t_1) + (T_2 \cdot t_2) + \cdots + (T_n \cdot t_n).$$

38. Although thermal budget TB as defined above has been considered a rough, first-order approximation of the thermal effects of semiconductor process step(s), the above equations have been used regularly and are widely accepted in the literature and in the field of semiconductor and integrated circuits today. The values provide a useful metric for quantifying the thermal effects of process step(s).

39. At a fixed process temperature in a certain process step, a longer process time would impose higher thermal effects on the previously fabricated structures than a shorter process time. Similarly, for a fixed process time duration, a higher process temperature would have higher thermal effects on the previously fabricated structures than a lower process temperature. It is important to note that high-temperature processes, *e.g.*, 900-1000°C, for a short time may have less thermal effect than low-temperature processes, *e.g.*, 650°C, for a longer time. This is because dopant diffusion or interfacial oxide growth not only depends on the process temperature but also the length of time maintained at that temperature. Ex-2028, 125:6-14; *see also* Ex-1052 at 247. Notably, the '702 Patent is silent as to the duration of its "low temperature processing." See *generally* Ex-1001. Thus, the '702 Patent does not teach a POSITA how to manage the thermal budget.

40. But it was known to POSITAs in 2004—the priority date of the '702

Patent—that thermal budget takes account of process temperature and process time. It was also known to POSITAs in 2004 that bonding a structure (containing FETs) formed separately from the substrate to a dielectric layer can be achieved at 900-1000°C (as described in Watanabe) *without* exceeding the thermal budget—and, therefore avoiding damage to the lower memory cells—if the process *time* is carefully controlled. See Ex-1053 at 15, 17-18, Table 1.

41. Several techniques were published and known to POSITAs well before the 2004 priority date of the '702 Patent. For example, a POSITA would have recognized that the electromagnetic induction heating (EMIH) bonding technique makes use of “rapid heating by directly coupling electromagnetic radiation into the silicon wafer.” Ex-1059 at i. “Minimizing...diffusion requires that the wafer is (1) rapidly ramped to the target anneal temperature, typically between 950°C and 1050°C; (2) soaked for the desired amount of time; and (3) cooled back below 500°C as rapidly as possible.” *Id.* at 10. Because EMIH heats the bonding surface rapidly for a very short period of time, as compared to furnace heating, EMIH offers “a method of bonding a stack of silicon wafers...in a shorter time period than prior methods” to bond wafers at a high temperature without exceeding the thermal budget. Ex-1048, 1:26-29, 1:40-43, 1:64-67.

42. Indeed, “[b]ecause the temperatures can be rapidly ramped to the bonding temperature of interest, the bonds can be set in a matter of minutes,” which

“compares favorably to the hour-long anneals necessary using more conventional heating techniques.” Ex-1059 at 11. “[A]ny time spent in the 700°C - 900°C range enhances deactivation and diffusion without contributing to activation,” and “[m]inimizing the time spent at 1000°C limits both ‘kick-out’ and traditional diffusion processes.” *Id.* at 9-10; see Ex-1049 (“[H]igher temperature (900°C), but low-thermal budget, rapid thermal annealing...promote[s] physical and/or chemical relaxations that minimize both interfacial and bulk dielectric defects, thereby improving device performance.”); see also Ex-1050 (concluding that short processing time is a more efficient process than furnace annealing and rapid thermal annealing can control shallow junction diffusion). EMIH can be applied to bond various semiconductor layers, including silicon-silicon and silicon-silicon oxide. *Id.* at 11. It has been demonstrated that a full-strength, void-free bond is achieved in five minutes at 1000°C, all without exceeding the thermal budget or degrading the performance of the FETs. *Id.* at ii. Because a POSITA would have understood that EMIH bonds stacked wafers at 900-1000°C while limiting heating time to avoid undesirable dopant diffusion, a POSITA would have known how to perform Watanabe’s stacking method without exceeding thermal budget.

43. PO’s argument also reads Watanabe in such a way as to render it unworkable. For example, PO argues that “repeated baking” is required but that is not so. Paper 18 at 25. A POSITA would have understood that multiple stacked

layers in Watanabe can be bonded during a single step. For example, Watanabe's Figure 6C shows two previously fabricated planar members stacked and bonded together with the substrate and peripheral circuit, and describes bonding all three structures during a single bonding step, not separate bonding steps per layer. *See* Ex-1005, 6:6-7. Thus, Watanabe's stacking method does not require that each stackable add-on layer is separately bonded and avoids "repeated baking."

44. In addition, a POSITA would have known about several methods of bonding, including EMIH, that allow all levels of the memory array to be bonded at once. Using EMIH, "four pairs of stacked wafers were bonded *simultaneously* in 5 min, demonstrating the potential for multiwafer bonds and high-throughput batch processing." Ex-1060 at 285 (emphasis added); *see* Ex-1059 at 29 ("[T]he same temperature is achieved for simultaneously heating several wafers or a single wafer."); *see also* Ex-1060 at 292 ("The low thermal budget and potential for batch processing also makes it suitable for high volume manufacturing."). Thus, a POSITA would have known how to bond the layers of Watanabe's stacking method without cumulatively exceeding the thermal budget.

45. Second, PO argues that Watanabe does not disclose how to "precisely align" its layers and contends that Watanabe results in a "facially unworkable technique." Paper 18 at 21. But PO again disregards the skill and knowledge of a POSITA. A POSITA would have known how to precisely align the planar members

using Watanabe's stacking method because techniques for wafer alignment were well-known and discussed extensively in the art.

46. There are many known techniques for wafer-to-wafer alignment for fabricating a three-dimensional semiconductor memory structure using stackable add-on layers. For example, two common wafer-to-wafer alignment techniques are mechanical alignment and optical alignment. In mechanical alignment, physical structures—such as the silicon wafers' main flats (or grooves)—are used to align the wafers to each other. Ex-1054 at 315-16, Fig. 1; *see also* Ex-1055 at 276-77. Typically, a set-up involving an alignment pin, as shown in Figure 1 below, is used to accomplish the task of alignment.

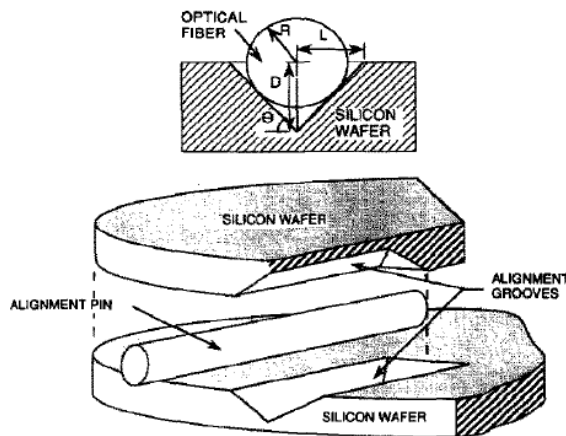


Fig. 1. Diagrammatic representation of the wafer-to-wafer alignment in partial cross-section. A cross-section of the important dimensions of the alignment groove geometry for a {100} silicon wafer (see eqn. (1)) is shown at the top of the Figure.

Ex-1054, Fig. 1. Optical alignment uses well-defined patterns called alignment marks, located on the frontside or backside of the wafers to be stacked. Ex-1055 at 274, Fig. 1. Typically, because silicon wafers are not transparent, wafer-to-wafer

alignment is performed indirectly, *i.e.*, a digitized image of the first loaded wafer is stored and the second wafer is aligned to the image reference of the first wafer, as shown in Figure 1 below. *See id.* at 274-75.

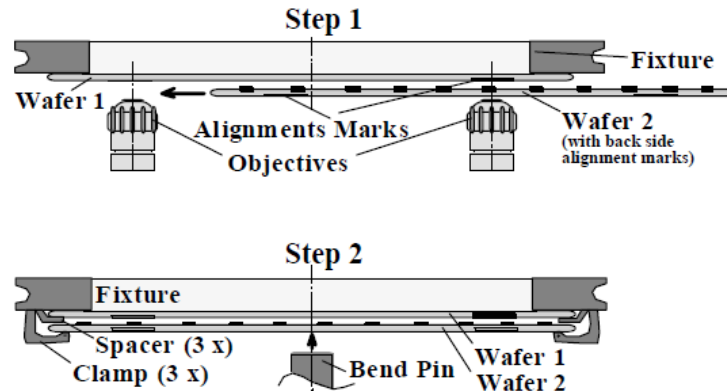


Fig. 1. Schematic of conventional wafer bond alignment [15]. Step 1: alignment using back side alignment marks and a digitised image; Step 2: aligned wafers on transport fixture before moving the fixture with the clamped wafers to the bond chamber.

Id., Fig. 1. The alignment accuracy using optical alignment is generally better than the accuracy obtained from mechanical alignment. *Id.* at 274 (providing “an effective alignment accuracy of equal or better than 5 μm ”); Ex-1054 at 315 (providing “high-accuracy (5 μm) wafer-to-wafer alignment”). Thus, a POSITA would have understood how to precisely align the planar members of Watanabe’s stacking method when combining Lee and Watanabe because the teachings of optical and mechanical alignment were well known by 2004.

47. By my review, the weaknesses PO argues for the prior art are also present in the ’702 Patent. For example, the ’702 Patent fails to acknowledge that thermal budget is defined as the product of process temperature and process time,

and is silent on the duration of its “low temperature processing.” *See generally* Ex-1001. The ’702 Patent also never mentions how to bond its stackable add-on layers. *See generally id.* Instead, the ’702 Patent relies on what a POSITA would have understood about thermal budgeting and wafer bonding techniques based on his or her knowledge and training.

48. Similarly, the ’702 Patent’s alleged teaching regarding wafer alignment merely identifies several known methods, such as using “a wafer alignment mark, or...a bump-type alignment structure” or “a simple notch alignment,” without any teaching of how to use these techniques. Ex-1001, 6:13-23. Again, the ’702 Patent relies on what a POSITA would have known of these common alignment techniques based on his or her knowledge and training. PO contends that not describing how to precisely align Watanabe’s wafers is “a fundamental flaw,” Paper 18 at 21, while inconsistently arguing that the world would know how to build the ’702 Patent’s alleged invention with nothing more than a mention of known techniques. Ex-1051, 111:19-116:16. Yet, Dr. Przybylski was unable to identify “when those techniques [came] into the common knowledge of a POSITA in the field” and had not even “considered whether [the ’702 Patent’s] description is sufficient for a POSITA to perform the necessary alignment.” Ex-1051, 111:24-114:7.

49. Moreover, the ’702 Patent’s alignment disclosure does more to emphasize the difference between the stacking and sequential deposition methods

than it does to explain how to properly align the wafers of the '702 Patent. For example, the '702 Patent's mention of alignment contrasts alignment required for conventional "photolithographic processes," *i.e.*, processes that use sequential deposition, with those for "combining the SOI layer with a semiconductor substrate" through the stacking method. Ex-1001, 6:13-23.

50. The '702 Patent also assumes that a POSITA would know how to apply its teachings to multiple types of memory devices—the same proposition PO rejects for Watanabe. The '702 Patent applies its teachings to multiple types of memory devices (*i.e.*, DRAM, SRAM, and nonvolatile memory, each claimed separately), without disclosing different circuit design techniques, by relying on what was known in the art. *See, e.g.*, Ex-1001, Abstract, 2:4-9, 9:24-35 (explaining that "a vertical memory cell...could work as a DRAM and/or a nonvolatile memory device"), 10:29-32 ("Since the body of SOI pillar is usually floating, the nonvolatile could work as a DRAM device."). PO's criticisms of the prior art ignore the level of skill in the art by 2004, including well-known teachings on which the '702 Patent relies. Because by 2004 a POSITA would have known how to overcome PO's alleged obstacles—thermal budget and alignment—using well-known techniques, PO has not identified any reason why the teachings of Lee and Watanabe cannot be combined.

B. A POSITA would have been motivated to combine the teachings of Lee, Watanabe, and Endoh and would have had a reasonable expectation of success

51. A POSITA would have known how to construct Lee's and Watanabe's memory array using Endoh's NAND architecture. Flash memory products have existed since at least the late 1980s. Ex-1025 at 2; Ex-1057 at 490. In flash memories, a single cell can be electrically programmed and a large number of cells, generally called a "block," can be electrically erased at the same time. Ex-1025; *see also* Ex-1058 at 0001. As described in my initial declaration, the two most common types of flash memory array architectures are NAND and NOR architectures. Ex-1002, VIII.H, VIII.I; *see also* Ex-1057 at 489 ("[T]oday two [types of Flash cells and architectures] can be considered as industry standard: the common ground NOR Flash...and the NAND Flash.... Most of these considerations are also valid for the NAND cell, since it is based on the same concept of floating-gate MOS transistor.").

52. In NOR architecture, "the floating gate memory cells of the memory array are arranged in a matrix. The gates of each floating gate memory cell of the array matrix are coupled by rows to word select lines and their drains are coupled to column bit lines." Ex-1012, 1:29-36. In "NAND array architecture ... the memory cells of the array are coupled together in series, source to drain, between a source line and a column bit line." Ex-1012, 1:47-49. Because the differences between the NOR and NAND architecture reside in the circuitry that couples the memory cells,

the memory cells themselves can be used in either architecture—*i.e.*, the same memory cell can be implemented in NAND or NOR architecture by simply changing the circuitry. Thus, one can drop in any type of nonvolatile semiconductor memory cells, whether they are of floating-gate device type or charge-trapping device type, to either the NAND architecture or the NOR architecture arrangement.

53. In my opening declaration, I defined the level of skill of a POSITA as someone having “a bachelor’s degree in electrical engineering, computer engineering, or a related field, and at least two years of experience in the research, design, or development of semiconductor memory device structures, or the equivalent, with additional education substituting for experience and vice versa.” Ex-1002 ¶34. PO and PO’s expert Dr. Przybylski adopted my definition. Paper 18 at 12; Ex-2029 ¶37. Under my definition, a POSITA would have learned about NAND memory design in his or her classroom education towards an electrical engineering or computer engineering degree, such as through common textbooks. *See, e.g.*, Ex-1009 at 0041. In fact, as discussed above, I taught semiconductor memory design and how dimensional scaling affects memory function, *e.g.*, programming efficiency, in courses at The University of Texas at Austin since at least 2004. A POSITA would have also gained knowledge about NAND and NOR memory designs through his or her two years of experience in the research, design, or development of semiconductor memory device structures because both NAND

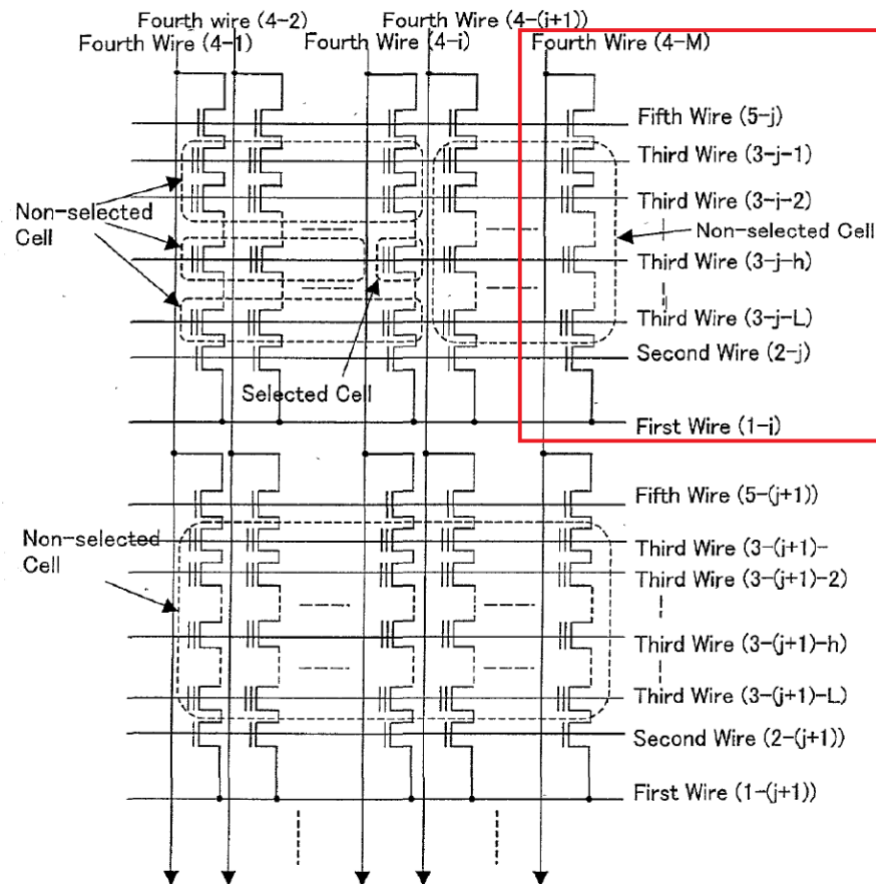
and NOR flash memory were already on the market and in common use by the early 2000s.

54. In addition, a POSITA would have known generally how to implement memory cells within NOR or NAND architecture because the differences between NOR and NAND architectures reside in the circuitry that couples the memory cells. *Id.* Neither the '702 Patent nor the prior art the Petition cites are directed to a particular circuit design—they apply their teachings to multiple types of memory devices and architectures. For example, the '702 Patent relies on the knowledge of a POSITA in its attempt to claim stackable add-on layers for DRAM, SRAM, and nonvolatile memory cells, all without specifying how a POSITA would modify the architecture. Ex-1001, 2:4-9. In fact, the '702 Patent states generally, “Since the body of SOI pillar is usually floating, the nonvolatile could work as a DRAM device.” *Id.* 10:31-32; *see also id.* 3:11-15, 3:18-22, 3:25-29, 9:24-35, 9:36-50 (applying the same teachings to DRAM and nonvolatile memory devices). Thus, a POSITA would have understood that the memory cells of the '702 Patent were interchangeable.

55. Similarly, a POSITA would have known how to implement Lee and Watanabe's memory cells in a NAND architecture by using well-known circuit design teachings. As Dr. Przybylski conceded, Endoh and Nguyen (Ex-1012) both disclose “memory cell transistors in a NAND flash memory array...connected

serially.” Ex-2029 ¶138. For example, Endoh’s Figure 65 illustrates a NAND array with four floating gate transistors (within annotated red box) connected in series forming a chain between the bit line (Fourth Wire) and ground (First Wire) through two select gate transistors (above and below the four floating gate transistors). Although Figure 65 illustrates four floating gate transistors, a POSITA would have understood that in an actual NAND memory array, there could be more than four floating gate transistors between the two select gate transistors.

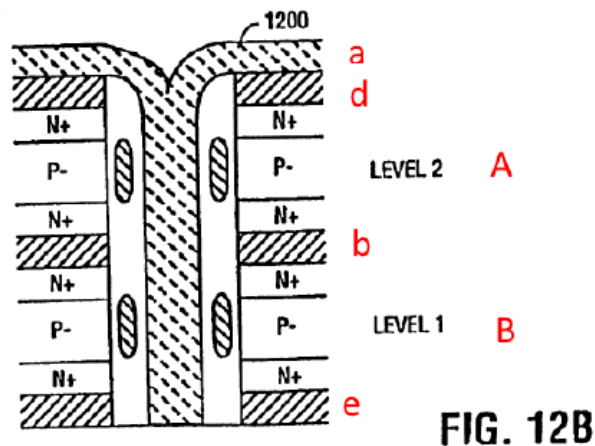
Fig. 65



Ex-1006, Fig. 65; *see also* Ex-1058 at 0020-21, Fig. 40. Each of the four floating gate transistors is controlled by a control gate (within each of the floating gate transistors); *see* Ex-1007, Fig. 2, 15:41-54. Because NAND architecture was well-known in the art and a POSITA would have understood that the memory cells were interchangeable, there would not be a “cascading set of modifications” to overcome, and no incompatible data flows, bit lines, sensing amplifiers, read and write protocols, or control circuitry. Rather, a POSITA would have begun with the teachings of NAND architecture, such as those from Endoh, and would have implemented the nonvolatile memory cells of Watanabe’s or Lee’s array in series within that architecture.

56. During my deposition, PO’s counsel questioned me on how to “convert” Lee’s Figure 12B into a NAND architecture. Counsel’s questions missed the point in framing the issue as “convert[ing]” Lee’s Figure 12B into a NAND architecture. Ex-2028, 94:13-17. Instead, the question is whether the memory cells shown in Lee’s Figure 12B can be implemented in a NAND architecture. Implementing Lee’s memory cells in the NAND or NOR architecture can be accomplished by starting with NAND or NOR circuitry and then using Lee’s Figure 12B memory cells within that circuitry. This task was within the technical ability of a skilled artisan in 2004. *See* Ex-1006 (Endoh), Ex-1022 (Endoh Article), Ex-1025 (Toshiba), Ex-1012 (Nguyen); *see also* Ex-1057, Ex-1058.

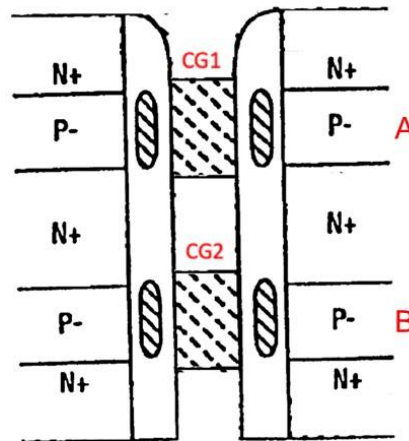
57. But even if asked to convert Figure 12B of the Lee Patent (Ex-1007), it is clear that a POSITA could have arranged the memory cells shown in Figure 12B within NAND flash memory array architecture. Shown below is Fig. 12B as annotated in Dr. Przybylski's Declaration. Ex-2029, 66.



Ex-2029, 66 (annotating Lee's Figure 12B). And during my deposition, PO's counsel asked whether layers d, b, and e are conductors or insulators. The Lee Patent does not explicitly describe the material of these layers. I testified that they are insulators because if they were not, conductor 1200 (annotated a), which connects to the gate of transistor A, would be electrically connected to the drain of transistor A, which would short the gate to the drain. In both NOR and NAND flash memory arrays, each transistor has a source, drain, and a gate, but in no case is the gate of any transistor shorted to its drain because this would render the transistor inoperable as a memory cell.

58. Assuming layers d, b, and e in Fig. 12B are insulators, then in order to

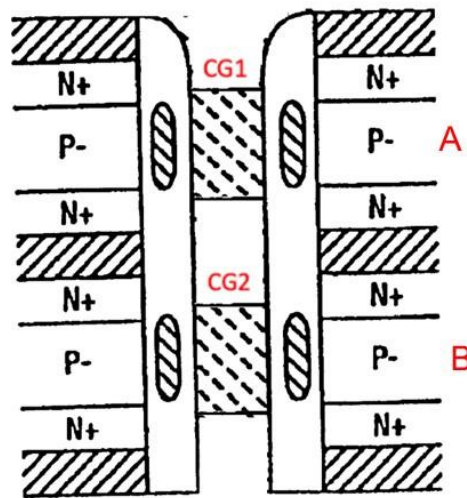
implement these memory cells in NAND architecture, we first need to remove the insulators d, b and e, as I discussed during my deposition. See Ex-2028, 94:13-96:20. Now, the two floating gate memory transistors (A and B) are connected in series. The next step is to incorporate individual control gates (CG1 and CG2) as shown in the annotated figure below.



This modification results in a NAND flash memory array architecture with two serial-connected memory cells (A and B), as shown below. This modification would have been well within the capability of a POSITA, and would allow for implementation of Lee's memory cells in NAND architecture.

59. Dr. Przybylski and BeSang contend that layers d, b, and e in Fig. 12B of the Lee Patent shown above are conductors. Ex-2029 ¶134; Paper 18 at 44. If that were true, then it would be even easier to arrange the memory cells of Lee's Figure 12B in a NAND architecture because the floating gate transistors A and B are already connected in series, similar to the structure discussed above and as required for

NAND architecture. To implement these memory cells in NAND architecture, a POSITA would install individual control gates (CG1 and CG2), as shown in the annotated figure below. This modification would result in a NAND flash memory array architecture with two serial-connected memory cells (A and B), each of which is controlled by a control gate (CG1 and CG2).



This modification would have also been well within the capability of a POSITA, and would allow for implementation of Lee's memory cells in NAND architecture. Thus, irrespective of whether layers d, b, and e in Lee's Fig. 12B are insulators or conductors, implementing Lee's memory cells in a NAND architecture is an exercise within the technical ability of a skilled artisan as of the 2004 priority date of the '702 Patent.

VIII. GROUNDS 3-4: CLAIMS 13-17 ARE OBVIOUS OVER LEE IN VIEW OF ENDOH

60. PO appears to adopt the Board's preliminary construction throughout

its Response in trying to distinguish Lee. And PO does not dispute that Lee alone or in view of Endoh discloses all elements of at least Claim 13 if the Board were to construe “stackable add-on layer” to encompass the sequential deposition method.

61. PO’s arguments amount to misinterpretation of the Petition’s mapping for Grounds 3 and 4, a failure to appreciate the level of skill in the art regarding NAND architecture, and an inconsistent construction of “SOI pillar.” First, as stated in the Petition and clarified in Petitioner’s Preliminary Reply, the Petition does not select or combine separate embodiments to arrive at the claimed subject matter.² The

² “Lee’s Figure 35A is a high-level schematic that generally illustrates how the charge storage devices of the invention may be construed, i.e., using multiple device layers deposited above a monocrystalline substrate and separated by insulating layers. Ex. 1007, 12:16–32. Figure 1A depicts an individual memory device that would be constructed in the manner shown in Figure 35. *Id.* at 12:50–56.” Paper 9 at 24–25. Figure 35’s “array of charge storage logic or memory devices 3101 [i.e., stackable add-on layers] are thus arranged as a three dimensional monolithic array [of] thin film transistors [TFTs] or diodes in amorphous or polysilicon layers,” where the “thin film transistors (TFTs) and diodes [are] arranged in a pillar configuration (i.e., the vertical direction with

Board correctly recognized this. Paper 9 at 25. Second, as discussed above, a POSITA would have known how to implement Lee's memory cell array within NAND architecture.

62. Third, my previous identification of the memory cells in Lee's Figure 12B as separate pillars does not mean that multiple memory cells cannot form a single pillar. The '702 specification in fact allows for multiple memory cells within a single pillar. *See, e.g.*, Ex-1001, Figures 4, 20, 27 (all showing multiple memory cells on structure 124, which is described as "SOI *pillar* [singular] 124" (emphasis added)), 8:7-10, 9:14-17, 9:51-55. PO's construction does not define whether a single or multiple memory cells constitute a "pillar," and Dr. Przybylski did not provide any bounds for a "pillar" during his deposition—testifying that "'taller than it is wide'...is the only indication of what constitutes a pillar." Ex-1051, 66:25-67:1. Thus, applying PO's logic, the two memory cells of Lee's Figure 12B may form a single pillar consistent with PO's construction in the same way that multiple memory cells of the '702 Patent's Figures 4, 20, and 27 form a single pillar.

63. The memory cells of Lee's Figure 12B may also form a single pillar consistent with my construction of "SOI pillar" because the '702 Patent defines a

respect to the substrate, where the length of the device is perpendicular to the substrate)...." Ex-1007, 12:20-23, 13:33-36.

pillar as a “doped stack structure.” Ex-1001, 6:28-29. “[T]he pillars in Figure 12B are just the n+, p-, n+ dope stack structures,” Ex-2028, 79:15-19, and a POSITA would have understood that two memory cells could together form a single pillar. Thus, under either PO’s or my construction, Lee alone or in view of Endoh renders the challenged claims obvious.

IX. GROUNDS 1-2: CLAIMS 13-17 ARE OBVIOUS OVER WATANABE IN VIEW OF ENDOH

A. Watanabe discloses a nonvolatile memory cell

64. Ground 1 does not combine Watanabe’s DRAM and EPROM embodiments. Watanabe discloses two “method[s] used to manufacture a semiconductor memory device,” having “memory cells arrayed three-dimensionally.” Ex-1005, 5:8-6:19, 1:49-52. Watanabe explains that “the present invention provides a 3-dimensional memory cell array in which memory cells are respectively located at the cross points of a plurality of planar members (21, 22) arranged in parallel.... By virtue of such a 3-dimensional memory cell array, the integration density of the semiconductor memory device can be greatly increased.” Ex-1005, 1:65-2:5. Watanabe further states, “FIG. 1 is a rough schematic illustration of a semiconductor memory device...made up of two conductive planes, or planar members, 21 and 22.” *Id.* 2:44-48. “In this instance [Figure 1], there are 8 cross points. The memory cell of the present invention is applicable to any type of memory device i.e., a SRAM, a PROM, or a DRAM....” *Id.* 2:50-55 (emphasis added); see,

e.g., Paper 1 at 9, 25. Thus, Watanabe explicitly states that the memory cells depicted at the cross points of Figure 1 (*i.e.*, 27-34), a high-level schematic of Watanabe's 3D structure, are applicable to any type of memory, including the EPROM nonvolatile memory cell depicted in Figure 7.

65. Although the memory cells in Figures 5 and 6 illustrating the manufacturing methods are "depicted as incorporated in a DRAM," Ex-1005, 2:53-55, a POSITA would have understood that Watanabe's teachings regarding the manufacturing methods apply equally to any type of memory cell disclosed in the invention. Petitioner does not attempt to combine or alter embodiments of different memory types shown in Figures 2A and 5-7. Rather, Petitioner is using these figures to illustrate how Watanabe's "stacking" and "laminating" (*i.e.*, sequential deposition) methods apply to the EPROM memory cell structure shown in Figure 7. For example, a POSITA would have understood from Watanabe that one could replace the contents of Figure 2A's cell 50 (shown as a dashed-line box) with the nonvolatile memory cell shown in Figure 7 because the photolithography steps necessary to make an EPROM memory cell were well known and within the skill of a POSITA. *See* Ex-1005, 3:37-41, 2:51-55, Figs. 2A, 7.

66. The '702 Patent takes a similar approach in its description of volatile and nonvolatile memory devices. For example, "FIG. 23 shows a vertical memory cell with a vertical MOSFET in a stackable add-on layer, where the vertical memory

has a floating gate 123k and a control gate 124, and could work as a DRAM and/or a nonvolatile memory device.” Ex-1001, 9:24-28. Thus, a POSITA would have understood that Watanabe’s “stacking” and “laminating” (*i.e.*, sequential deposition) methods apply equally to the EPROM nonvolatile memory device depicted in Figure 7 and would have understood how to manufacture the resulting device.

B. Watanabe in view of Endoh discloses an “SOI pillar”

67. “SOI pillar” should be construed as “a doped stack structure formed from a single-crystal portion of a semiconductor wafer transferred from one wafer to another previously fabricated wafer or similar type of substrate.” Because Watanabe discloses a vertical structure with “[a]t least two stack structures (two layers) [*i.e.*, two planar members] each consisting of layers 41 to 46”—*i.e.*, a doped stack structure—formed over an insulator, a POSITA would have understood that Watanabe discloses a SOI pillar structure or renders it obvious. *See* Ex-1005, 3:12-13, Figs. 1, 2A (showing individual doped stack structures forming vertical pillars above insulator 49). Further, Watanabe discloses that the “n+ type Si layer 45 is layered on an insulating film 49” and that “[w]hen a vapor phase epitaxial method”—commonly used to form a SOI structure—“is used for forming these layers, these may be successively formed by appropriately changing doping gases.” Ex-1005, 3:33-37, Figs. 2A-2B. Endoh also discloses “a semiconductor memory device” in which “a bulk substrate but also a SOI substrate, a multi-layered SOI

substrate and the like may be used,” and where “island-like semiconductor layers are formed in the form of *pillars* by selective epitaxial silicon growth.” Ex-1006 ¶¶329, 1206 (emphasis added). Thus, Watanabe in view of Endoh discloses or renders obvious a doped stack structure formed from a single-crystal portion of a semiconductor wafer transferred from one wafer to another previously fabricated wafer or similar type of substrate.

68. In contrast, PO’s construction of “SOI pillar” is inconsistent with the state of the art at the time. As discussed above, the figures of Watanabe, Lee, and the ’702 Patent were not drawn to scale. Ex-2028, 139:22-140:3; Ex-1051, 54:9-56:3; Ex-2029 ¶80. And regardless of scale, PO compares, inexplicably, the alleged “SOI pillar 124” of the ’702 Patent’s Figure 1—comprising a source, channel, and drain—with only drain 110 of Watanabe’s Figure 7. Paper 18 at 60. *But see* Paper 1 at 29 (explaining that Figure 7’s vertical transistor—comprising source 108, channel 109, and drain 110—could be formed on a SOI pillar).

69. Moreover, Watanabe discloses a “SOI pillar” even under PO’s construction, which does not limit “SOI pillar” to a single doped silicon region or even a single memory cell. Thus, even if a pillar is construed as taller than it is wide, Watanabe discloses a “pillar” because it describes stacking multiple layers as shown in Figure 6C, which will result in a structure that is taller than it is wide. Ex-1005, 3:12-15, 6:3-5, Fig. 6C. Watanabe alone or in view of Endoh discloses a “pillar”

under either my or PO's construction.

C. Watanabe's stacking method is enabled

70. Despite Watanabe's bonding process at 900-1000°C, a POSITA would have understood how to manage thermal budget. For example, a POSITA would have understood that EMIH offers "a method of bonding a stack of silicon wafers...in a shorter time period than prior methods" to manage the thermal budget while bonding the wafers at a high temperature. Ex-1048, 1:26-29, 1:40-43, 1:64-67; *see also* Ex-1049; Ex-1050. Thus, a POSITA would have known how to perform Watanabe's stacking method without exceeding thermal budget.

X. CONCLUSION

71. I declare that all statements made herein of my knowledge are true, that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

72. I declare under penalty of perjury that the foregoing is true and correct.

Respectfully submitted,

May 3, 2024



Dr. Jack C. Lee