

UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE PATENT TRIAL AND APPEAL BOARD

MICRON TECHNOLOGY, INC.,
Petitioner,

v.

YANGTZE MEMORY TECHNOLOGIES COMPANY, LTD.,
Patent Owner.

Case No. IPR2024-00795
U.S. Patent No. 11,501,822

REPLY DECLARATION OF DAVID KUAN-YU LIU, PH.D.

I, David Kuan-Yu, Liu, Ph.D. declare as follows:

I. INTRODUCTION

1. I have been retained by Micron Technology, Inc. (“Micron” or “Petitioner”) as an independent expert consultant in this proceeding before the United States Patent and Trademark Office (“PTO”). I am not an employee of Micron or any affiliate or subsidiary of Micron.

2. I have been asked to reply to certain of the opinions set forth in the January 3, 2025 declaration of Steven Przybylski, Ph.D submitted by Patent Owner in an *inter partes* review proceeding relating to U.S. Patent No. 11,501,822, which I refer to herein as the ’822 patent.

3. I am being compensated at my current standard consulting rate, plus reimbursement for any reasonable expenses. My compensation is based solely on the amount of time that I devote to activity related to this case and 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 financial interest in this proceeding.

4. The prior declaration I submitted in connection with this proceeding, which I understand has been designated as Ex. 1003, sets forth among other topics my educational background and professional experience (Ex. 1003, ¶¶ 5-14 and associated Exhibit 1017), my understanding of the law (*id.*, ¶¶ 19-34), an overview

of pertinent technical background (*id.*, ¶¶ 40-73), an overview of the '822 patent and its prosecution history (*id.*, ¶¶ 74-103), the knowledge and characteristics of a person having ordinary skill in the art (*id.*, ¶¶ 35-39), and certain claim construction opinions (*id.*, ¶¶ 104-107).

5. In forming the opinions set forth in this declaration, I have considered and relied upon my education, knowledge of the relevant field, knowledge of scientific and engineering principles, my work with those of ordinary skill at the time of the alleged invention, and my general experience in the field of art.

6. I have also reviewed and considered the materials listed in my prior declaration (*id.*, ¶ 18), the Board's Institution Decision (Paper 9), Patent Owner's Response (Paper 12), the Przybylski Declarations (Exs. 2001 and 2011), and the following materials to the extent relevant to my opinions:

Exhibit	Description
2001	July 17, 2024, Declaration of Dr. Steven Przybylski
2002	U.S. Pat. No. 11,081,164
2004	U.S. Pat. No. 10,700,157 ("Chen")
2005	U.S. Pub. 2011/0194351 ("Kim")
2021	Transcript of the December 17, 2024, Deposition of Dr. David Kuan-Yu Liu
2022	Hambley, "Electrical Engineering Principles and Applications, Fourth Edition" (2008)

1022	Spinelli et al., “Reliability of NAND Flash Memories: Planar Cells and Emerging Issues in 3D Devices,” Computers 2017, 6, 16 (2017)
------	---

II. RESPONSE TO CERTAIN OF PATENT OWNER’S ARGUMENTS

7. The paragraphs below respond to certain of the arguments and assertions made by Patent Owner and Dr. Przybylski in this proceeding. In these paragraphs, I address only a fraction of the opinions and assertions Dr. Przybylski set forth in his declaration. The fact that I do not respond to any particular opinion expressed by Dr. Przybylski should not be considered to indicate agreement with that opinion.

A. Ground 1: Obviousness Over Lee

8. According to Dr. Przybylski, “Lee’s pre-bias voltages do not qualify as ‘word line pre-pulse signal[s]’ because they are not formed as pulses.” Ex. 2011, ¶ 79. Rather than being supplied as a pulse, Dr. Przybylski posits that Lee’s word lines would be supplied with “flat bias voltages that enable the voltage level in the memory string to reach a single steady state value before being followed immediately by programming voltages.” *Id.* Dr. Przybylski goes on to opine that not pulsing Lee’s pre-charge voltages would somehow improve “programming time” and a “uniform level of charge.” *Id.*, ¶¶ 79-82. Dr. Przybylski also takes the

position that there is purportedly no “reason” why use of pulsed programming renders the use of pulsed pre-charge voltages. *Id.*, ¶ 87.

9. I disagree with these opinions. Dr. Pryzbylski’s analysis is erroneous and fails to account for (1) Lee’s teachings and (2) the realities of 3D NAND memory operation.

10. First, Dr. Pryzbylski is wrong when he argues that Lee teaches a memory in which bias voltages are applied as a “flat” value that is “immediately” followed by a programming voltage.

11. 3D NAND is unable to operate in this way. Word line potential does not instantaneously change from one value to another. Instead, there is necessarily a period of transition, ramp up, or ramp down as the potential moves from an original value to another value. The different voltages are typically derived from output of charge pumps. Any new voltage level needs to be supplied by the charge pump to the word line driver, which will take time. Put another way, there will be a continuous potential gradient between the starting potential and the new potential rather than an “immediate” shift from one potential to another as Dr. Pryzbylski argues.

12. This operation is reflected throughout the prior art. For example, in just the references at issue in this proceeding, Zhao’s Figure 13 shows word line potential gradually increasing when voltage is applied and gradually decaying when voltage

is removed. *E.g.*, Zhao, Fig. 13. Yang's Figures 8B and 9B show the same thing: rather than potential immediately changing from one value to another, there is a period of transition. *E.g.*, Yang, Figs. 8B, 9B. This delay is caused by the RC component of the driving circuit, with the resistance being the resistance of the word line driver circuit and capacitance being the word line / gate layer to the substrate / channel of the memory array. The manner in which voltage is applied to word lines in NAND is discussed in greater detail below. During this operation, the word line effectively acts as a single capacitor plate. The voltage on this plate/word line is the same along its entire length throughout operation (it just takes a short period of time for the voltage to reach the intended level as dictated by, for instance, the charge pump supplying the voltage).

13. Because word line potential does not immediately change, if Lee's pre-charge voltages were not pulsed and were instead applied in the way Dr. Pryzbylski argues, word line potential would not "immediately" transition from the pre-charge to the programming value.

14. Instead, there would be a period of transition that would cause pre-charging and programming to effectively blend together: pre-charging would continue into the period where programming is meant to occur. And programming would extend back into the period when pre-charging is meant to occur. This would in turn cause pre-charging to continue for some indeterminate period while

programming voltages are being applied. And it would cause programming to begin earlier than intended.

15. This is not how a POSITA would have understood Lee to operate.

16. To begin, Lee plainly teaches that its pre-charging and programming occur in separate steps. These steps are not meant to overlap and be blended together as Dr. Pryzbylski argues.

17. Figure 7 of Lee exemplifies this. There, Lee’s “precharging” occurs in an earlier step S205, while a “program” operation is performed in separate later step S207. Lee, Fig. 7.

18. This is also repeated throughout Lee. Lee repeatedly references that its system employs separate and distinct “precharging” and “program” steps. *E.g., id.*, 11:13-22. Indeed, Lee even explains that “program” and “pass” voltages are not applied to Lee’s word lines until “**after** ... the precharge operation is **completed**.” *Id.*, 9:7-14. This is plainly referring to pulsed operation: pre-charge is not “completed” until the word lines are permitted to return to their original value (*i.e.*, ground). The various steps also—contrary to Dr. Pryzbylski’s arguments—employ different voltages. This includes not only program and pass voltages, but different pre-bias voltages applied to different groups of word lines. *See id.*, 6:29-31; 9:41-44.

19. Moreover, Lee explains that its pre-charge voltages are applied only “**during** a precharge operation” or “precharging step.” *Id.*, 7:58-65 (emphasis added); *see also id.*, 8:15-17; 8:25-29; 8:32-47; 8:53-64 (all referencing application of pre-charge voltages “during” specific steps or periods of operation).

20. Lee never says that these steps can (or even should) be blended together in the way Dr. Pryzbylski blends them. It certainly never says that programming voltages should be applied while pre-charging is still occurring. That, however, is how Dr. Pryzbylski is now interpreting Lee. By arguing that a POSITA would not understand word line voltage to go back to 0V between pre-charging and programming, Dr. Pryzbylski has blurred the distinction between Lee’s pre-charging and programming steps and has effectively construed Lee as employing a single, continuous operation during these two periods instead of two separate operations as Lee itself teaches.

21. Dr. Pryzbylski’s interpretation of Lee is also contrary to how a POSITA would have understood a NAND memory—like that taught by Lee—to operate for additional reasons. As shown in Figure 1, Lee employs a “voltage generation unit 150” to generate the voltages its system needs (including V_{pgm}, V_{pass}, and the pre-bias voltages) and an “address decoder” to ensure that those voltages are supplied to the correct word lines. *See also* Lee, 6:19-39.

22. These components were well-known in the art at the time the '822 patent was filed.

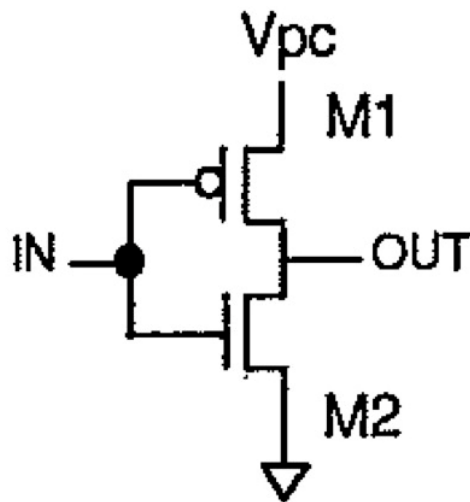
23. NAND memory employs a voltage generation unit because it is typically supplied with only a single low voltage power supply (often referred as " V_{DD} "). Required high voltages are not externally supplied, instead they are internally generated. The voltage generation unit is responsible for internally generating—using the low voltage power supply—the different higher voltages required for memory operation. A charge pump is used for this purpose. This is discussed repeatedly in the literature. *See, e.g.*, Ex. 1008, 259 (“The voltages that are required for Program and Erase are generated internally starting from the V_{CC} value available. ... [I]t is from this low value of power supply that the internal voltages for Program and Erase must be generated. The increment of the voltage is obtained by using charge pump circuits”); Ex. 1012, 189 (“Both program and erase operations require voltages higher than V_{DD} High voltages are generated by a charge pump and filtered by a proper voltage regulator.”)

24. A charge pump works by storing charge—with the appropriate voltage—in a capacitor. The capacitor is then able to supply a required voltage to signal lines. But, it can only do so for short periods of time and without the flow of current (the charge pump simply stores charge; it is not a continuous source of power). *See, e.g.*, Ex. 1008, 311 (“The principle of charge pumping is based on the

storing of charge with the proper sign at the desired node. ... [T]he element used to transfer charge from one node to another is the capacitor, because this component has the property of keeping the voltage drop across its plates unchanged as far as small time intervals are considered”).

25. Because they operate by accumulating charge on capacitors, charge pumps also require time to prepare for the next voltage that needs to be supplied. For instance, in a typical programming flow, an initial “voltage setup” step occurs, where the “charge pumps are enabled in order to raise the voltage from V_{cc} to the program voltage.” Ex. 1008, 354. Then, the prepared voltage is supplied to the word lines in the form a “pulse.” *Id.* “At the end of the pulse, the charge pumps get ready to switch to the voltage needed in the verify phase.” *Id.*

26. Address decoders (also referred to as “row decoders” or “word line decoders”) like that referenced in Lee were also well known. The most basic of these is a simple circuit like that shown below:



Ex. 1008, 269. A logical input signal supplied by control circuitry determines whether the output (which will be connected to a word line) either receives a voltage generated by the charge pump (indicated by “Vpc” in the figure above) through transistor M1, or is pulled to ground (indicated by the triangle) through transistor M2. As can be seen, whenever a voltage from the charge pump is not being supplied to the word lines, the word lines are grounded and thus will return to 0 V.

27. A further example of this is shown below. As depicted, whether voltage is supplied to word lines during pre-charge is “driven by the ENABLE signal.” Ex. 1012, 191-192.

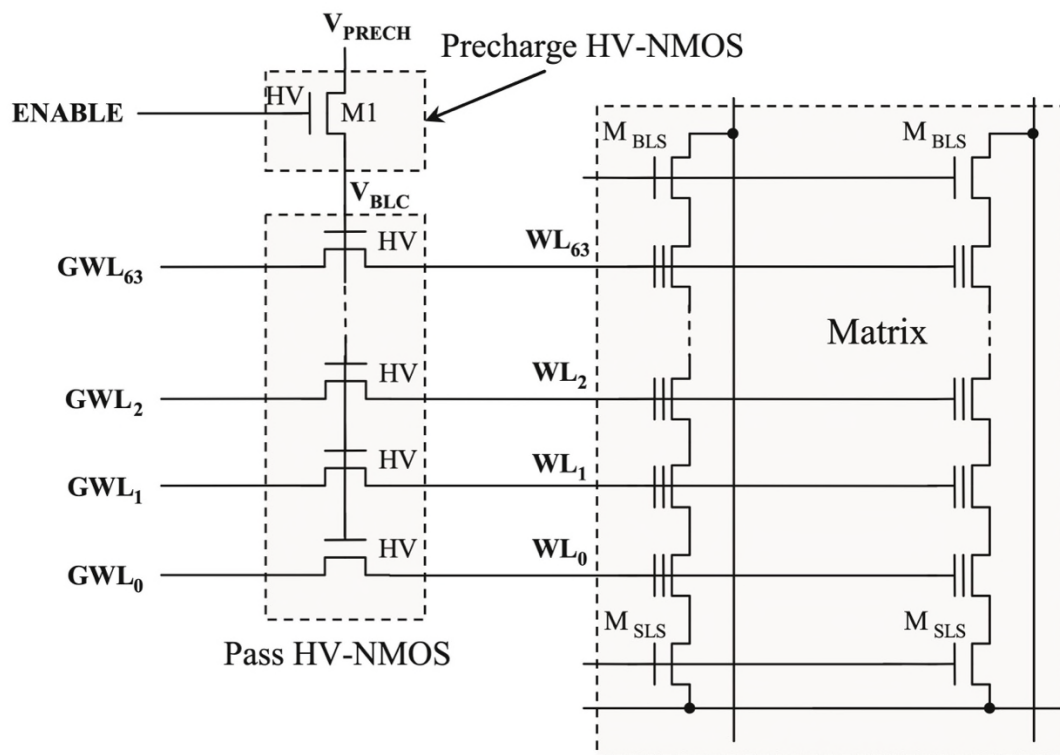


Fig. 6.41 All-NMOS wordline driver

Id. High voltage supplied by the charge pump will make its way to the word lines only when the ENABLE signal is active. When the ENABLE signal is not active, the word lines will be pulled back to ground.

28. NAND's use of these components—including a charge pump to generate high voltages and a row decoder to supply those voltages to word lines—has a number of important effects. In particular, it means that a POSITA would have understood that all required voltages are supplied in the form of pulses. Again, the charge pump (1) requires time to develop a charge, (2) only supply charge for relatively short periods of time, and (3) requires time to transition from one voltage to another. Moreover, when a voltage is not being supplied (i.e., when the charge

pump is preparing itself to supply another, different voltage), (4) the word lines are pulled to ground by the row decoder.

29. Thus, applying a flat pre-charge voltage that is “immediately” followed by a programming voltage would be a very unusual, atypical, and difficult to implement way of operating a NAND memory. Rather than employing the type of circuitry a NAND memory typically employs (like a charge pump and row decoder), it would require unique, unusual, and special circuitry nowhere discussed in Lee. Instead, as I have explained, Lee teaches standard NAND voltage generation and supply circuitry that will (or at the very least obviously) supply voltages in the form of pulses.

30. Moreover, it bears noting that all voltages were routinely supplied to signal lines in NAND are supplied in the form of pulses due to the origination of these voltages in the components I discuss above. This includes programming voltages applied to selected word lines (*see, e.g.*, Ex. 1012, 89 (discussing program pulses), pass voltages applied to unselected word lines (*see, e.g., id.*, 89 (discussing both program and pass voltage pulses), and erase voltages (*see id.*, 94, 181 (discussing erase pulses)). A POSITA would see no reason why Lee’s pre-bias voltages would not be applied in the same way.

31. Second, Dr. Pryzbylski is simply wrong when he argues that applying Lee’s precharge voltages as “flat” values immediately followed by programming

voltages would somehow improve uniformity and speed operation. The opposite is true. Dr. Pryzbylski's proposed operation would destroy uniformity.

32. As I explain above, word line potential is unable to instantaneously shift. Because of this, if Lee is operated in the way Dr. Pryzbylski argues—and programming or pass voltages are applied immediately after pre-charging without first allowing potential to return to 0 V—there will be some indeterminate amount of extra unintended pre-charging disrupting the desired pre-charge level. And this amount of indeterminate, extra programming will differ from word line to word line because different pre-charge and programming voltages are applied.

33. In my opinion, a POSITA would have readily understood that this would **decrease** word line pre-charge uniformity. Rather than adopting the precise values Lee intends, by immediately applying programming voltages rather than allowing potential to return to 0 V, pre-charge would begin to drift away from the intended value as the word line potential transitions from the pre-charge voltage to the programming voltage.

34. A POSITA would have understood that this unwanted drifting effect is avoided—and the desired uniformity maintained—by separating the pre-charge and programming periods in time.

35. This is why Lee employs separate in time pre-charge and programming steps: by separating the steps in time, pre-charging can be controlled and the level

of pre-charge is prevented from drifting away from the controlled value before programming occurs.

36. Next, in my opinion, a POSITA also would have understood that pulsed voltages are a necessity given how higher voltages are generated in NAND (via a charge pump that generates voltage, whose application to word lines is controlled by an on/off row decoder as discussed above). Dr. Pryzbylski's argument that use of pulses somehow slows operation fails to account for this. The operation he proposes—immediately switching from one voltage to another without returning to ground in between—does not allow adequate time for voltages to be developed by the charge pump. To accomplish what Dr. Pryzbylski proposes would either (1) make the NAND more costly or difficult to manufacture (by requiring more components, such as additional charge pumps to supply voltages one after another without a charging period between) or (2) complicate operation (thus reducing the risk of failure).

37. Third, a POSITA would have recognized that Lee includes only a single “voltage generation unit 150” that functions to output all the voltages its system requires, including the pre-charge and programming voltages. Lee, 7:13-19; *see also id.*, 6:19-21, 6:28-36, Fig. 1.

38. Because a single unit is employed, a POSITA would plainly have understood that because Lee's pre-charging and programming voltages are generated by the same unit, they would have the same form.

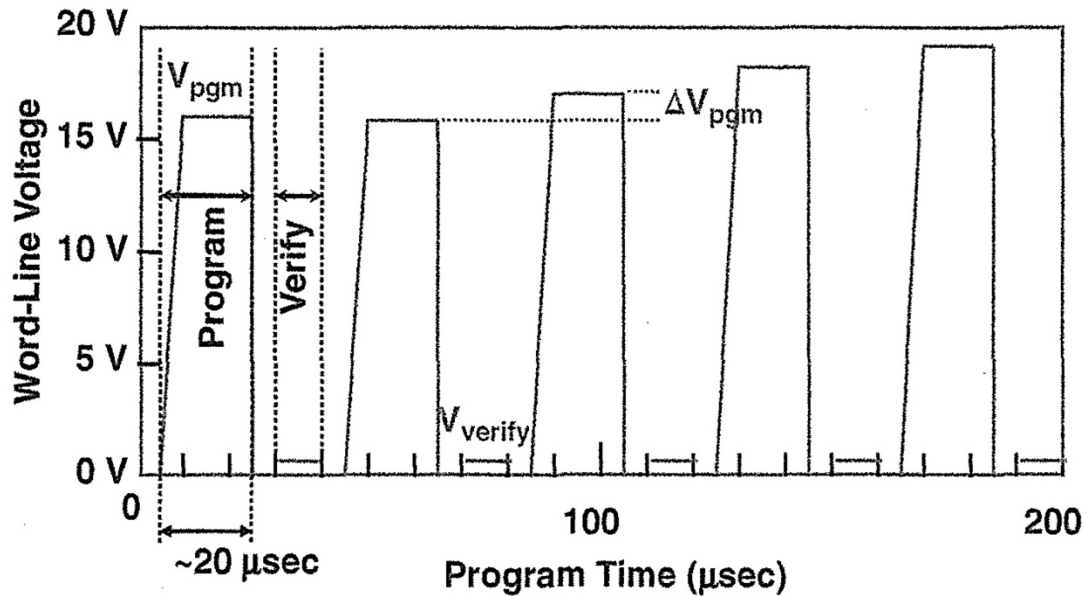
39. Dr. Pryzbylski does not appear to dispute that programming voltages were routinely supplied as pulses when the '822 patent was filed. Because both the programming and pre-charge pulses are coming from the same unit in Lee, a POSITA would have had every reason to believe that they are applied in the same general form: *i.e.*, as pulses.

40. Next, a POSITA would have understood that application of programming voltages as pulses requires the use of pre-charge voltages in the form of pulses.

41. "[I]ncremental step pulse programming" (or ISPP) was a "scheme ... used in NAND flash memory products for a long time, more than 20 years, due to fast programming speed, tight V_t distribution, and excellent reliability." Ex. 1006, 116-118; *see also* Ex. 1012, 82 (cell programming is "typically accomplished by" ISPP); Ex. 1008, 354.

42. An example of ISPP is shown below:

Incremental Step Pulse Programming (ISPP)



Ex. 1006, 46.

43. As shown, during ISPP, the word line voltage starts at 0 V. *See id.* Then, a series of voltage pulses with specific durations and magnitudes are applied to program the memory cell allowing it to reach the desired threshold voltage representing stored data. *See id.*

44. Additionally, when programming voltages are applied to selected word lines in the form of pulses as shown above, the pass voltages applied to unselected word lines must also be in the form of pulses with the same pulse duration. *See, e.g.,* Ex. 1012, 89. In other words, use of ISPP to supply a programming voltage requires the use of matching pass voltages. If the pass and program voltages were not supplied in the same pulsed form, then there would be prolonged voltage differences between the selected and unselected word lines (the selected lines would be at V_{prog}

while the unselected lines receive an un-pulsed pass voltage; all while the channel of the selected string is held at 0V from the bitline). This creates a substantial risk of unwanted electron movement and program disturb that NAND memory seeks to avoid, as the intended applied Vpass voltage is optimized to reduce the disturb on both the selected wordline of the unselected string (disturb voltage is “Vprog-Vpass”), and the unselected wordlines of the selected string (disturb voltage is “Vpass-0V”). Ensuring that the selected and unselected word lines are pulsed in the same way (with both transitioning between a program/pass voltage and ground at the same time) avoids this.

45. Lee applies pre-charge voltages to all its word lines (this includes both the selected and unselected lines) during the same pre-charge period.

46. In my opinion, a POSITA would have recognized that if pre-charging were not pulsed in Lee, and programming voltages were instead applied “immediately” after the pre-charge voltage as Dr. Pryzbylsk argues, then the later programming procedure would be altered.

47. The ISPP programming pulse applied to the selected word line after pre-charging would no longer have the intended magnitude and duration. Instead, it would blend into and be combined with the pre-charge voltage (which never returned to zero) resulting in an indeterminate and undesired increase in programming.

48. This would in turn defeat the purpose of ISPP: precise control of the desired threshold voltage would no longer be achieved.

49. A POSITA would consider this non-sensical and would never operate memory in this manner. When ISPP is used—as Dr. Przybylski admits was known and readily done—then pre-charge voltages would also obviously be pulsed to avoid creating programming issues. As a final note, there is no meaningful difference between the purpose of pre-charging in the ‘822 patent—*i.e.*, purportedly to “enhance” “the channel potential gradient” (Ex. 1001, 4:54-59)—and Lee or any other reference discussed in this proceeding. The prior art—just like the ‘822 patent—seeks to address charge distribution issues in the channel during pre-charging. While references like Lee and Zhao do address different portions of the channel, the end goal is the same: to ensure that the channel is charged in a way that avoids program disturb.

B. Ground 2: Obviousness Over Zhao

50. According to Dr. Przybylski, a POSITA would not have “replace[d]” Zhao’s “Vss with some kind of pulsed signal during the time period.” Ex. 2011, ¶ 102. Instead, according to Dr. Przybylski, the voltage of word lines WLn-x would be changed to a non-zero voltage at all times, not just during time period t1-t3. *Id.*, ¶ 106.

51. This is not what Zhao teaches. Dr. Przybylski’s interpretation of the reference is unreasonable and ignores how it is intended to operate.

52. According to Zhao, all of its lines—including every word line—start “at 0 volts” (or “Vss”) “[a]t time t0.” Zhao, 22:34-35. As I explain above, this is consistent with the generation of voltages in this context, because a voltage generator, e.g., charge pump, must be prepared and connected to the respective word line.

53. Then, during the pre-charge period t1-t3, different voltages can be applied to and received by different lines for purposes of pre-charging. *See id.*, 22:39-23:12.

54. At “time t3,” the word lines that received a non-zero voltage are again “lowered to ground/Vss” and the “pre-charge process has completed.” *Id.*, 23:8-11.

55. Zhao does explain that in one embodiment, the “source side unselected word lines WLn-x” can “**remain** at Vss (e.g., 0 V)” during pre-charge period t1-t3. Zhao, 23:7-9.

56. So here, Zhao is discussing operation in which no signal or voltage is applied to these word lines and they remain at their starting potential of 0 V.

57. Zhao, however, provides an alternative: “unselected source side word lines” WLn-x can instead “**receive ... one or more voltages** that are too low to pre-

charge and/or too low to turn on memory cells in the highest data state (e.g., data state S7).” *Id.*, 19:22-26 (emphasis added).

58. Per Zhao, this received voltage is applied to word lines WLn-x during pre-charging step 650. *See id.*, 19:9-26.

59. Given all this, a POSITA would have understood that Zhao teaches two different ways to handle word lines WLn-x: (1) the word lines can remain at the original potential of 0 V or (2) a variety of non-zero voltage can be applied during pre-charging.

60. A POSITA would not, however, have understood that the “one or more voltages that are too low to pre-charge” would be applied to word lines WLn-x at all times as Dr. Przybylski argues.

61. There are multiple reasons for this.

62. First, per Zhao, all of its word lines are at 0 V at t0 and return to 0 V at t3. Zhao includes no embodiments where the starting/ending voltages in its system are changed. That, however, is what Dr. Przybylski is proposing.

63. Second, Zhao explains that the “one or more voltages that are too low to pre-charge” are “received” by WLn-x. Vss or ground is not a “received” voltage. It is simply the ground state that the word lines return to when no voltage is applied. A POSITA would plainly understand that the “one or more voltages that are too low to pre-charge” are applied voltages. Further, these voltages are plainly applied by

the same components in Zhao, during the same time period, and in the same way as Zhao's other pre-charge voltages.

C. Ground 3: Obviousness Over Zhao and Lee

64. Dr. Przybylski also argues that a POSITA would not have been motivated to apply Lee's teachings to Zhao. He advances a number of reasons for this: (1) there is supposedly a difference between "charging operations performed via the common source line" and "charging from the bit line," (2) the "resistive effects" Lee seeks to address are supposedly "not present ... in Zhao's system," and (3) differences between Zhao and Lee's "approaches for mitigating program disturb" somehow weigh against the combination. *See* Ex. 2011, ¶ 112.

65. Once again, all of this is wrong. In fact, Dr. Przybylski does not appear to understand how precharging occurs in 3D NAND.

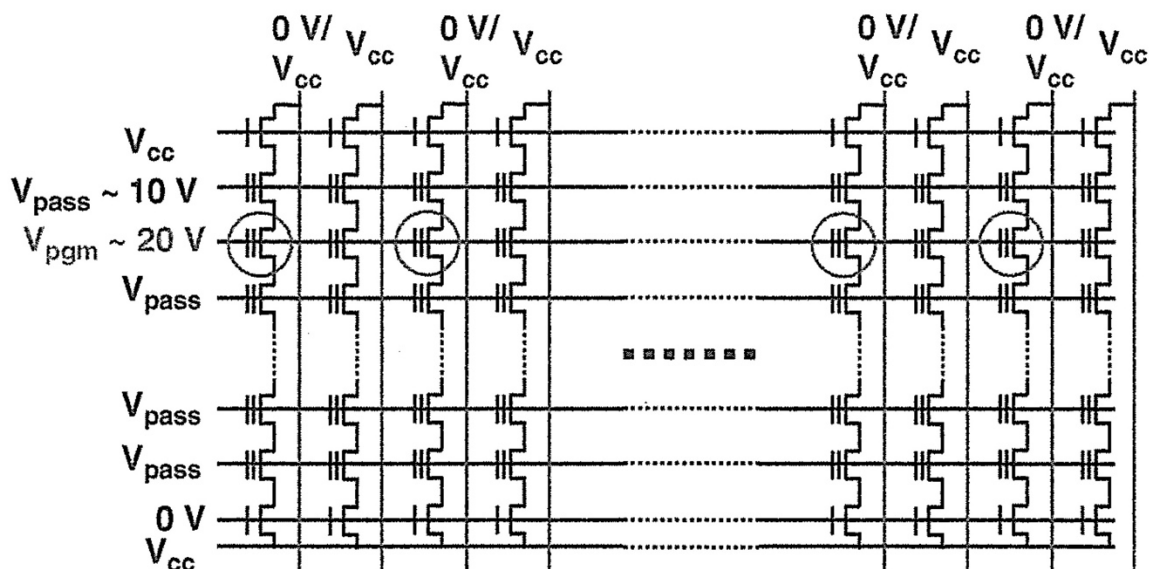
66. Precharging is meant to prepare a NAND memory string's channel for later programming. *See, e.g.,* Yang, 2:23-24 (discussing "pre-charging NAND strings" before a "program operation").

67. During both programming and the preceding pre-charging step, voltages are applied but current is not flowing through the memory string channel. This is because one of the select gates is closed during these periods.

68. Thus, Dr. Przybylski is wrong when he characterizes Lee as entailing current flow. Current is not flowing in either Lee or Zhao during pre-charging.

69. The only period during NAND operation when channel current flows between bitline and the source is during a read. During this phase of operation, only the low power supply voltage is applied across the channel of the string. The charge pumps are not employed, so there is no risk to this componentry. Channel current between bitline and the source does not flow during any other period of operation. One reason for this is that during other phases of operation, including program and erase, high voltages are supplied from the charge pumps. Flowing current through the channel during the precharging and programming operations would cause the charge pumps supplying the bitline voltage to collapse, adversely impacting memory operation.

70. This is reflected throughout the literature. It was well known that during programming, only one of the channel's two select gates is open. An example of this is shown below. As depicted, Vcc is applied to the drain select transistors to turn them on. However, 0 V is applied to the source select transistor, preventing current flow through the channel:



E.g., Ex. 1006, 45; *see also* Ex. 1022, 27 (Figs. 31 and 32 similarly show that only the drain select transistor is on during programming; the source select transistor is off due to application of a ground voltage to its gate).

71. The purpose of applying voltages during pre-charging is to purge and redistribute electrons already in the channel that may later cause program disturb. This is explained in both Lee (*e.g.*, Lee, 1:52-56 (discussing the mitigation of “program disturb”)) and Zhao (Zhao, 17:4-12 (noting that “program disturb” can be caused by “electrons trapped in the channel”). Indeed, Dr. Przybylski noted this himself, explaining that in Zhao pre-charging “allow[s] electrons to flow freely” for purposes of “eliminat[ing]” unwanted “voltage differential.” Ex. 2001, ¶ 46.

72. During pre-charging, a voltage is applied to both the channel and the word lines.

73. The overall amount of channel pre-charging (*i.e.*, total electron movement and purging) is controlled by the voltage applied to the channel itself.

74. The relative amount of charging along the channel (*i.e.*, the electron distribution) is controlled by the application of different voltages to the word lines intersecting the channel.

75. A greater potential difference between the voltage applied to a particular word line and the voltage applied to the channel results in more pre-charging in the portion of the channel associated with that word line.

76. In an ideal case, the potential difference between the word line and the channel (gate to channel) would be the same, especially for the memory cells that are not programmed yet. This same potential difference would translate into a uniform boosted channel potential to prevent program disturb.

77. Again, the purpose of the voltage applied to the channel is simply to create a potential difference that facilitates electron redistribution. Because of this, the channel voltage can be applied to either end of the channel. All that is required is that there be electrical connection between the point where the voltage is applied to the channel and the portion of the channel targeted for pre-charging so as to allow the channel potential to properly develop.

78. Yang provides an example of this. In some embodiments, Yang applies a voltage to the “drain end of the NAND string,” while in others it applies voltage

to the “source end.” Yang, 17:16-23; Figs. 8C, 9C, 10C. The goal in all cases is the same: to develop potential differences between the channel and word lines to facilitate electron redistribution.

79. Given all this, in my opinion a POSITA would not have considered Lee and Zhao to be different at all. Instead, both references are compatible and employ complementary solutions to the same problem.

80. More particularly, both Lee and Zhao address program disturb. *See, e.g.*, Lee, 1:52-56; 2:18-25.

81. Further, like all prior art teaching channel pre-charging, both Lee and Zhao create voltage differentials between a NAND channel and word lines to control channel charge distribution. *See, e.g.*, Lee, 1:62-2:32; Zhao, 19:9-56.

82. The only real difference between Lee and Zhao is the portion of the channel they address.

83. In particular, Lee tailors potentials to ensure uniform pre-charging in the portion of the channel between the selected word line and the bit line when programming from source to drain. *See* Lee, 8:36-51.

84. Lee explains that when programming a NAND channel starting at the source side and moving towards the drain (where the bit line is located), a “resistance component between the source select transistor and the drain select transistor”

operates to relatively reduce the “precharge level” of memory cells closer to the channel’s source side. *Id.*, 8:25-32.

85. To avoid this and ensure development of a uniform pre-charge, Lee applies different pre-charge voltages to different groups of word lines. Word lines closer to the selected word line receive lower voltages. Word lines closer to the bit line and further from the selected word line receive higher voltages. This results in a uniform distribution of channel potential gradient, and a more uniform potential difference with the channel among these memory cells.

86. In contrast, Zhao addresses pre-charging issues in the portion of the channel **crossing** the selected word line.

87. According to Zhao, its pre-charging scheme is meant to prevent “electrons trapped in the channel [from] being accelerated from one side of a selected word line to another side of the selected word line and redirected into the selected word line” during programming. Zhao, 2:14-18.

88. To accomplish this, Zhao applies word line voltages around the selected word line in a way that is intended to “open the channel from one side of the selected word line to the other side” so electrons can move across the selected word line to “purg[e]” the “already programmed side of the channel.” *Id.*, 2:18-30.

89. Rather than teaching incompatible approaches, in my opinion a POSITA would have recognized that Zhao and Lee discuss pre-charging issues that

occur in different parts of the channel. These issues can of course occur at the same time: electrons may be trapped on the already programmed side of the channel (causing program disturb) at the same time the length of the unprogrammed side impacts precharging (by impacting uniformity).

90. Because of this, a POSITA would have considered Zhao and Lee to together present a more comprehensive solution to addressing program disturb.

91. Lee addresses the not-yet-programmed side of the channel by ensuring that precharging occurs uniformly despite channel resistance effects.

92. Zhao addresses the charge imbalances that can exist between the already programmed and not-yet-programmed portion of the channel.

93. Next, a POSITA would have recognized that Lee's teachings are directly applicable to Zhao. Further, a POSITA would have applied Lee's different word line voltages to Zhao in the same way Lee itself teaches.

94. Lee applies a "precharge voltage" to its "common source line." Lee, 2:30-32.

95. A POSITA would have understood that one purpose for doing so would be to set the source to a known voltage so that it can later be connected to the channel via the source select transistors at the appropriate time. While Lee does note that voltage is applied to the source during pre-charging, implying that the channel voltage may be applied from the source side, the purpose of such a voltage is only

to create potential differences between the channel and the word lines. Current is not flowing through Lee's channel. In fact, an additional effect of this voltage on the common source line is to further ensure that the string would be shut-off since a higher source voltage would translate into a more robust off state in a memory cell. A POSITA would have understood this as a further proof, (although a POSITA would have already known), that there is no channel current flow in the memory string during Lee's precharge operation,

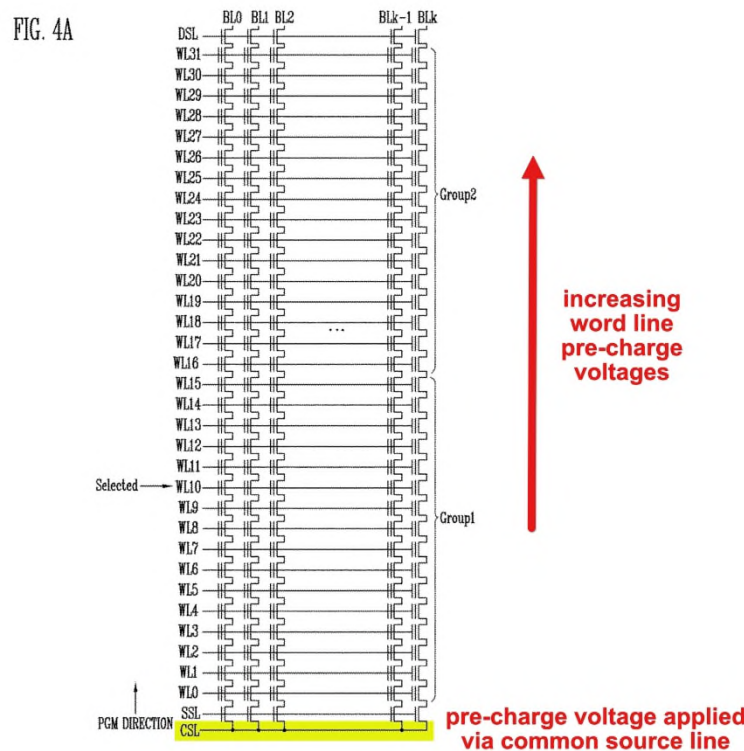
96. Thus, it does not matter whether voltage is applied to the channel from the source end or the drain (bit line) end of Lee's channel. The same effect would be achieved—and the same word line voltages would be used—in either case.

97. This should have been apparent to Dr. Przybylski both because (1) this is how NAND precharging occurs (as I discuss above) and (2) Lee itself shows that it operates this way (as I discuss in the paragraphs below).

98. The fact that Lee has nothing to do with current flow—and instead only concerns itself with the development of potential differences—is exemplified by Lee's own examples.

99. Figure 4A of Lee depicts an embodiment in which Lee's memory array is programmed starting from the source side and moving in the direction of the drain (or bit line). *See* Lee, 8:18-20.

100. When programming in this direction, a “precharge voltage” is applied to the common source line. *Id.*, 4:65-67. At the same time, pre-charge voltages are also applied to the word lines. These voltages increase with distance from selected word line in the direction of programming. *Id.*, 8:32-47. This is shown in the below annotated version of Figure 4A:

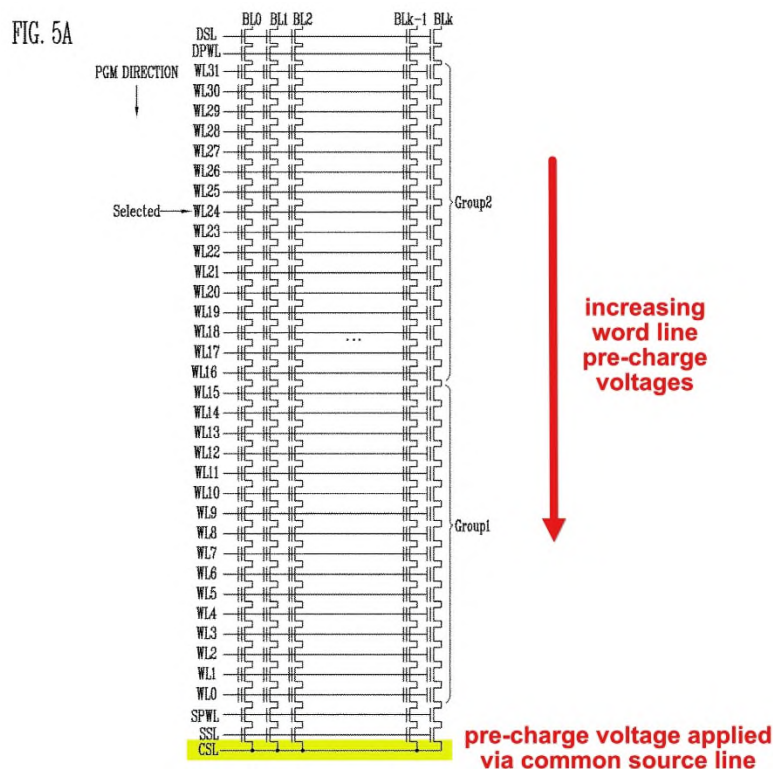


Id., Fig. 4A (annotated).

101. In Figure 5A, Lee’s memory array is programmed in the opposite direction: starting from the drain side (or the bit line) and moving towards the source. *See id.*, 10:16-21.

102. Once again, Lee applies voltage to the common source line (and not the bit line) during pre-charging. See *id.*, 1:30-37, 1:52-58, 8:65-67, 11:34-38, 11:45-61, Fig. 8.

103. But, the word lines now receive opposite voltages. So, unlike in Figure 4A, now lower voltages are applied closer to the drain (or bit line) and higher voltages are applied closer to the source. *Id.*, 10:28-50. The voltages still increase in the direction of programming. This is shown in Figure 5A below:



Id., Fig. 5A (annotated).

104. This confirms that Lee is not operating in the way Dr. Przybylski assumes it is.

105. Lee's word line voltages are not compensating for some sort of "current flow" through Lee's channel. If they were, then both Figure 4A and Figure 5B would use the same word line pre-charge voltages because a voltage is applied to the channel from the source side in both figures.

106. Instead, these figures use opposite word line voltages (such that the lowest voltages are always applied closest to the selected word line and the highest voltages are always applied furthest from the selected word line regardless of how close/far these word lines are to/from the source of the channel potential).

107. This is because in Lee—as in all pre-charging—it is the potential differences between the channel and the word lines that matters. There is no current flow.

108. I recognize that Lee does state that its various pre-charge voltages are meant to "offset" a "resistance component between the source select transistor and the drain select transistor." Lee, 8:25-29, 8:47-51.

109. But Dr. Przybylski is wrong when he argues that this is meant to be a reference to current flow. No current is applied to Lee's channel during pre-charging.

110. I explained this to Patent Owner during my deposition. Here, Lee is referring the fact that the threshold voltages of transistors connected in series along a channel vary in the direction of programming. *See* Ex. 2021, 83:12-84:2

111. Lee simply adjusts the voltages applied to each word line during pre-charging to account for these transistor threshold voltage differences, so that there is a gradient in the channel potential to facilitate uniform precharging, for eventual uniform channel boosting. *See* Lee, 8:25-29, 8:47-51. Moreover, Lee is plainly tailoring its applied word line voltages to not only account for any channel resistance, but also to prepare the array's cells using the minimum necessary voltage.

112. In my opinion, a POSITA would have recognized that Zhao will experience the very same problem.

113. In particular, Zhao's memory also includes transistors connected in series along a channel. *See generally* Zhao, Fig. 4G.

114. Because of this, when programming the memory from the source to the drain (the same direction of programming shown in Figure 4A of Lee), the threshold voltages of Zhao's transistors will vary in the same way discussed in Lee.

115. But, because Zhao applies the same pre-charge voltages to all its not-yet-programmed word lines (*see, e.g.,* Zhao, Fig. 13), these variations in threshold voltage will result in inadequate pre-charging in the portion of the channel closest to the selected word line.

116. This is what necessitates use of Lee: different voltages must be applied to Zhao's not-yet-programmed word lines (which will have varying threshold

voltages due to their location in series long the channel) to achieve uniform pre-charging.

D. Ground 4: Obviousness Over Yang

117. In addressing the Yang reference, Dr. Przybylski appears to do little more than perform a side by side comparison of the voltages shown in Figure 8 of Yang.

118. This, in my opinion, was improper. A POSITA would plainly have read the entirety of Yang. And this would have led a POSITA to conclude that Yang employs three different pre-charge voltages.

119. First, Yang separately refers to three pre-charge voltages: (1) an “over-drive voltage” (Vpre_OD), (2) a “bypass voltage” (Vbypass), and (3) a “dummy pre-charge voltage” (Vpre_dmy). Yang, 22:4-37; 24:11-33.

120. These three voltages are also separately depicted in Figure 8B. *See id.*, Fig. 8B.

121. Further, reviewing Yang’s discussion of these voltages confirms that they are different.

122. “[O]verdrive voltage” (Vpre_OD): is “above all the data states” of the memory’s memory cells. *Id.*, 20:53-55.

123. “[D]ummy pre-charge voltage” (Vpre_dmy): this has “a magnitude such that the dummy memory cells will operate as a pass gate (e.g., conduct a

current). ... An example of Vpre_dmy is 5V, assuming that the” memory cell “threshold voltage is less than 5V.” *Id.*, 24:18-27.

124. “[B]ypass voltage” (Vbypass): this is at least Vev, but no greater than Vv7....” *Id.*, 20:64-66. Vv7 is the memory’s highest programmed state threshold voltage.

125. So, Vpre_OD is highest. Vpre_dmy is intermediate (it is intended to exceed the programmed threshold voltage). Vbypass is lowest (it is not 0 V, but does not exceed the programmed threshold voltage).

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements were made with the knowledge that willfully false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1101 of Title 18 of the United States Code.

Date: March 28, 2025

Location: Fremont, California

Signature: 
David Kuan-Yu Liu, Ph.D.