Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2664851 A1 2008/06/05

(21) 2 664 851

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2007/11/29
(87) Date publication PCT/PCT Publication Date: 2008/06/05
(85) Entrée phase nationale/National Entry: 2009/03/30
(86) N° demande PCT/PCT Application No.: CA 2007/002149
(87) N° publication PCT/PCT Publication No.: 2008/064480
(30) Priorité/Priority: 2006/11/30 (US11/565,170)

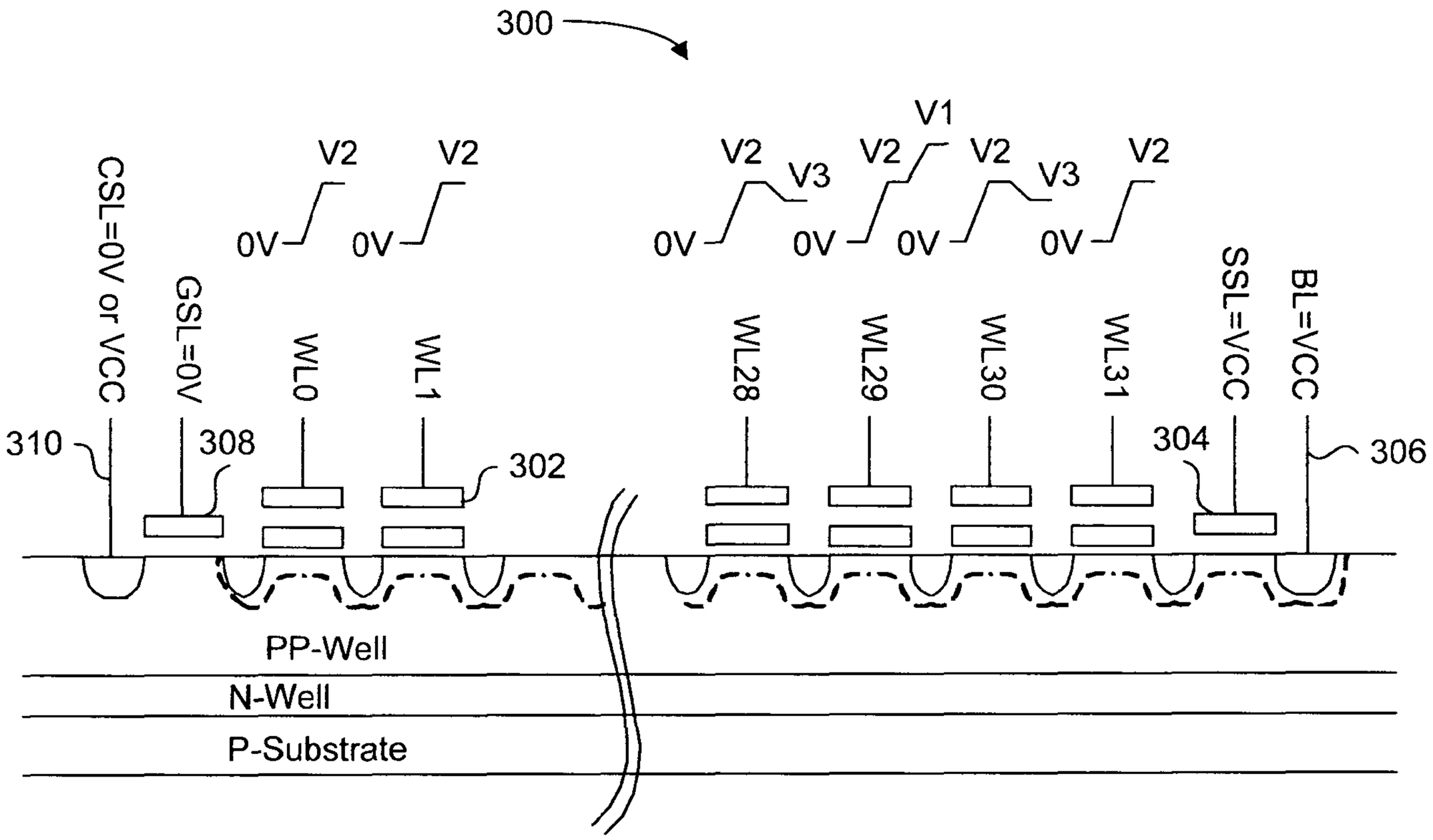
(51) Cl.Int./Int.Cl. *G11C 7/12* (2006.01),
G11C 16/24 (2006.01)

(71) Demandeur/Applicant:
MOSAID TECHNOLOGIES INCORPORATED, CA

(72) Inventeur/Inventor:
KIM, JIN-KI, CA

(74) Agent: SUMI, SHUJI

(54) Titre : PLAN D'INHIBITION DE PROGRAMME DE MEMOIRE FLASH
(54) Title: FLASH MEMORY PROGRAM INHIBIT SCHEME



(57) Abrégé/Abstract:
A method for minimizing program disturb in Flash memories. To reduce program disturb in a NAND Flash memory cell string where no programming from the erased state is desired, a local boosted channel inhibit scheme is used. In the local boosted channel inhibit scheme, the selected memory cell in a NAND string where no programming is desired, is decoupled from the other cells in the NAND string. This allows the channel of the decoupled cell to be locally boosted to a voltage level sufficient for inhibiting F-N tunneling when the corresponding wordline is raised to a programming voltage. Due to the high boosting efficiency, the pass voltage applied to the gates of the remaining memory cells in the NAND string can be reduced relative to prior art schemes, thereby minimizing program disturb while allowing for random page programming.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 June 2008 (05.06.2008)

PCT

(10) International Publication Number
WO 2008/064480 A1

(51) International Patent Classification:
G11C 7/12 (2006.01) *G11C 16/24* (2006.01)

(21) International Application Number:
PCT/CA2007/002149

(22) International Filing Date:
29 November 2007 (29.11.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/565,170 30 November 2006 (30.11.2006) US

(71) Applicant *(for all designated States except US)*: **MOSAID TECHNOLOGIES INCORPORATED** [CA/CA]; 11 Hines Road, Kanata, Ontario K2K 2X1 (CA).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **KIM, Jin-Ki** [CA/CA]; 46 Ironside Court, Ottawa, Ontario K2K 3H6 (CA).

(74) Agents: **HUNG, Shin** et al.; BORDEN LADNER GERVAIS LLP, World Exchange Plaza, 100 Queen Street, Suite 1100, Ottawa, Ontario K1P 1J9 (CA).

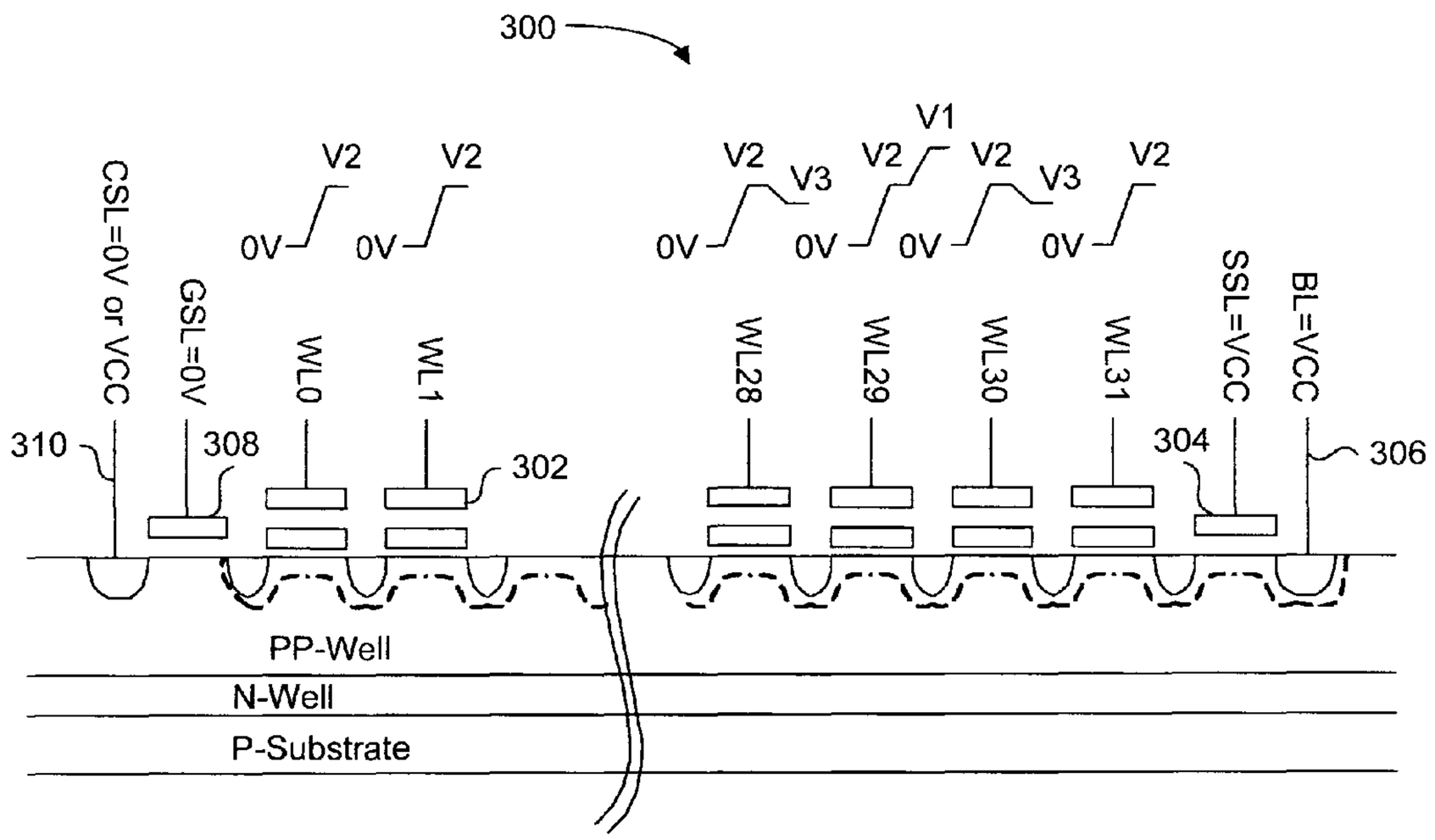
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN,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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:
— *with international search report*

(54) Title: FLASH MEMORY PROGRAM INHIBIT SCHEME



(57) Abstract: A method for minimizing program disturb in Flash memories. To reduce program disturb in a NAND Flash memory cell string where no programming from the erased state is desired, a local boosted channel inhibit scheme is used. In the local boosted channel inhibit scheme, the selected memory cell in a NAND string where no programming is desired, is decoupled from the other cells in the NAND string. This allows the channel of the decoupled cell to be locally boosted to a voltage level sufficient for inhibiting F-N tunneling when the corresponding wordline is raised to a programming voltage. Due to the high boosting efficiency, the pass voltage applied to the gates of the remaining memory cells in the NAND string can be reduced relative to prior art schemes, thereby minimizing program disturb while allowing for random page programming.

WO 2008/064480 A1

FLASH MEMORY PROGRAM INHIBIT SCHEME

FIELD OF THE INVENTION

The present invention relates generally to non-volatile memory. In particular, the present invention relates to non-volatile Flash memory programming schemes.

BACKGROUND OF THE INVENTION

Numerous types of consumer electronics products rely on some form of mass storage for retaining data or software for the execution of code by a microcontroller. Such consumer electronics are prolific, and include devices such as personal digital assistants (PDA's), portable music players, portable multimedia players (PMP's) and digital cameras. In PDA's, mass storage is required for storing applications and data, while portable music players and digital cameras require large amounts of mass storage for retaining music file data and/or image data. The mass storage solution for such portable electronics is preferably small in size, consumes minimal power, and has high storage density. This limits the selection to non-volatile forms of memory since volatile memories, such as static random access memory (SRAM) and dynamic random access memory (DRAM), require a constant application of power in order to retain data. As is known in the art, portable electronics rely on batteries that have a finite power supply. Therefore, non-volatile memories that retain data after power is removed are preferred.

While many consumer products use commodity Flash memory, Flash memory is indirectly used by consumers in products such as cell phones and devices with microprocessing functions. More specifically, the application specific integrated circuits (ASIC) commonly found in consumer electronics can have integrated Flash memory to enable firmware upgrades. Needless to say, Flash memory is versatile due to its optimal balance in

size, storage density, and speed, making it a preferred non-volatile mass storage solution for consumer electronics.

Figure 1 is a general block diagram of typical Flash memory of the prior art. Flash memory **10** includes logic circuitry for controlling various functions of the Flash circuits, registers for storing address and data, high voltage circuits for generating the required program and erase voltages, and core memory circuits for accessing the Flash memory array. The functions of the shown circuit blocks of Flash memory **10** should are well known in the art. Persons skilled in the art will understand that Flash memory **10** shown in Figure 1 represents one possible Flash memory configuration among many possible configurations.

A read operation is a relatively straight-forward access of data stored at a particular memory location of the memory array, called an address. Prior to a write operation to a specific block of the memory array, the specific block must first be erased with the application of high voltages. A write operation, more accurately called a program operation, requires the careful application of high voltages to a selected memory location, followed by a program verify operation to ensure that the data has been properly programmed. Furthermore, since high voltages are used, the Flash chip must be designed to be relatively tolerant to inadvertent programming of non-selected memory cells.

Figures 2a, 2b and 2c are illustrations of the NAND memory cell string used in memory cell array **28**. Figure 2a is a circuit schematic of two NAND memory cell strings. Figure 2b is a chip layout of the two NAND memory cell strings shown in Figure 2a. Figure 2c is a cross-sectional view of one NAND memory cell string shown in Figure 2b along line A-A'. Each NAND memory cell string includes **32** serially connected floating gate memory cells **50** each connected to respective wordlines WL0 to WL31, a string select transistor **52** connected between the bitline **54** and the first floating gate memory cell **50**, and a ground select transistor **56** connected between a common source line (CSL) **58** and the last floating gate memory cell **50**. The gate of string select transistor **52** receives a string select signal SSL,

while the gate of ground select transistor **56** receives a ground select signal GSL. The NAND memory cell strings share common wordlines, string select SSL, and ground select GSL signal lines. The construction and arrangement of the shown NAND memory string is well known in the art.

As previously mentioned, the NAND memory cell strings of the memory array are first erased, according to well known techniques in the art. Each block of NAND memory cell strings can be selectively erased, therefore one or more blocks can be simultaneously erased. When successfully erased, all erased floating gate memory cells **50** will have a negative threshold voltage. In effect, all erased memory cells **50** are set to a default logic state, such as a logic "1", for example. Programmed memory cells **50** will have their threshold voltages changed to a positive threshold voltage, thus representing the opposite "0" logic state.

Figure 3 shows a threshold voltage (Vt) distribution graph for erased memory cells and programmed memory cells. Due to process and voltage supply variations, the erased and programmed threshold voltages will be distributed within a voltage range. As shown in Figure 3, erased memory cells can have a negative threshold voltage between -3V to -1V, while programmed memory cells can have a positive threshold voltage between 1V and 3V. Generally, a cell is programmed by applying a high voltage to its gate while keeping its source and drain terminals grounded. The high electrical field causes electrons in the memory cell channel to cross the gate oxide and embed in the floating gate (known as Fowler-Nordheim (F-N) Tunneling), thereby increasing the effective threshold voltage of the memory cell.

A brief discussion of how data is read from the NAND memory cell string now follows with reference to Figures 2a to 2c. To read data one memory cell **50**, SSL and GSL are both set to a read voltage (Vread, typically higher than Vcc), of 5V for example, to turn on string select transistor **52** and ground select transistor **56**. All wordlines except the wordline to be accessed are set to a read voltage (Vread), of 5V for example, while the wordline to be

accessed is set to 0V. Therefore, all memory cells **50** having a threshold lower than 5V are turned on to couple the source and drain terminals of the selected memory cell having the 0V wordline, to the bitline **54** and the source line **58**. If the selected memory cell is in the erased state (having a negative threshold voltage), it will turn on, thereby coupling the bitline **54** to the source line **58**. On the other hand, if the selected memory cell in the programmed state (having a positive threshold voltage), it will not turn on, thereby isolating the bitline **54** from the source line **58**. The presence or absence of current is then detected by sense amplifiers.

The aforementioned reading scheme will reliably access memory cell data if the erased and programmed threshold voltages remain within their respective ranges. However, the described situation is ideal and the threshold voltages of erased and programmed memory cells can potentially shift when memory cells are programmed.

Programming is typically done by the page, meaning that all the memory cells **50** in the block connected to the same wordline are selected to be programmed with write data (logic "0") at the same time. The remaining memory cells are thus unselected during programming. Since the memory cells start in the erased state (logic "1") prior to programming, only the memory cells to be programmed with the logic "0" should be subjected to the high electric fields necessary to promote F-N tunneling. However, due to the physical connections of the memory array, all the memory cells along the same wordline receive the same high voltage programming level. As a result, there is a potential that erased memory cells will have their threshold voltages unintentionally shifted. This is called program disturb, which is well known in the Flash memory field.

Therefore a program inhibit scheme is used for preventing those memory cells where no change from the erased state is required, from being programmed to the logic "0" state. There are two known program inhibit schemes that can be used. The first is a basic inhibit scheme and the second is a self-boosted inhibit scheme. Table 1 summarizes the voltages applied to the relevant signal lines of Figure 2a for both schemes. It is assumed that all the

memory cells **50** connected to WL27 are erased to the logic "1" state, and BL0="0" and BL1="1" data is to be written to the memory cells **50** accessed by wordline WL27.

Table 1

	Basic inhibit	**Self-boosted inhibit**
BL0	0V (VSS)	0V (VSS)
BL1	VPI (8V for example)	VCC (2.5V for example)
WL27	VPGM (18V for example)	VPGM (18V for example)
WL0-WL26, WL28-31	VPASS (10V for example)	VPASS (10V for example)
SSL	VPASS (10V for example)	VCC (2.5V for example), then 0V
GSL	0V (VSS)	0V (VSS)

For the basic inhibit scheme, VPGM is set to a sufficiently high program voltage to cause F-N tunneling with drain voltage of 0V on the selected cell. VPASS is set sufficiently high to render unselected cell transistors in the selected string conductive regardless of their programmed state, and to pass VPI to the memory cell where no programming is desired. At the same time, VPASS should be insufficiently high to initiate F-N tunneling on unselected cells. VPI is an inhibit voltage on BL1 set sufficiently high to inhibit F-N tunneling on the selected cell connected to WL27, since the voltage difference between VPGM and VPI in the channel of the memory cell connected to WL27 is now too small.

There are some problems with the basic inhibit scheme. VPI is provided by an internal high voltage generator during program operations, and a large capacity charge pump is required to supply VPI to the highly capacitive bitlines. This leads to a drastic increase in the power consumption and the chip size, which are both highly undesirable. Page buffers connected to the bitlines must now be configured for high voltage operation in order to

provide VPI to the bitlines. High voltage transistors are larger than regular voltage transistors, which increase the page buffer size and consequently, the chip size. Programming speed is reduced due to the need to charge the highly capacitive bitlines to VPI, with the on-chip voltage generator which has a limited current supply.

The self-boosted inhibit scheme addresses the problems of the basic inhibit scheme. With the SSL transistors **52** turned on and the GSL transistors **56** turned off, 0V or ground voltage, is applied to bitline BL0 , while a high voltage such as VCC is applied to bitline BL1. A 0V bitline (as in BL0) ties the channel of the associated NAND string to ground. When the program voltage VPGM is applied to the gate of the selected cell, the large potential difference between gate and channel results in F-N tunneling of electrons onto the floating gate, thereby programming the cell. In program inhibited cells, BL1 initially precharges the associated channels towards the VCC voltage. When the voltage of WL27 rises to VPGM, and the remaining wordlines reach VPASS, the series capacitances through the control gate, floating gate, channel, and bulk are coupled and the channel potential is automatically boosted.

This boosting occurs when the coupled channel voltage rises to Vcc-Vth (where Vth is the threshold voltage of the SSL transistor). At this point, the SSL transistor **52** turns off and the channel becomes a floating node. It has been calculated that the floating channel voltage rises to approximately 80% of the gate voltage, which is sufficiently high to prevent F-N tunneling from occurring.

However, program disturb can still occur. In particular, if VPASS is set to be too high, erased cells coupled to BL0 can be unintentionally soft programmed to the logic “0” state due to the relatively high difference in voltage between VPASS and the 0V channel. On the other hand, if VPASS is set to be too low, the erased memory cell connected to WL27 and coupled to BL1 may not receive sufficient channel boosting to inhibit F-N tunneling.

Unfortunately, with continued scaling down due to advances in semiconductor fabrication technology, the power supply VCC is also scaled to a lower level. This is disadvantageous for the self-boosted inhibit scheme. This is because the NAND cell string channel is initially precharged to Vcc – Vth of the SSL transistor **52** before self-boosting by VPGM and VPASS, and the boosted voltage is largely affected by the precharge voltage Vcc-Vth. Furthermore, to achieve higher packing density, designers are increasing the number of memory cells in each NAND string. Therefore, relative to a NAND string having 16 memory cells **50**, the 32 memory cell NAND string shown in Figures 2a to 2c will endure twice as many program disturb cycles than a 16 memory cell NAND string.

Accordingly, the threshold of programmed and erased memory cells can be unintentionally shifted during programming operations. Figure 4 shows threshold voltage (Vt) distribution graph for program disturbed erased memory cells and programmed memory cells. The solid curves correspond to the threshold distribution originally shown in Figure 3, while the dashed curves show the shifted threshold distribution due to program disturb. This shifting can be due to accumulated number times a cell is disturbed, or a single program disturb event. This is very problematic as the shifted thresholds can affect read operations. As previously discussed for a read operation, the gate of a selected memory cell **50** is driven to 0V, while unselected memory cells receives a read voltage Vread at its gate, while unaccessed memory cells have their gates driven to a read voltage Vread. As shown in Figure 4, the 0V level may not turn on those cells having threshold voltages shifted above 0V. Furthermore, unselected cells having their thresholds shifted over the Vread voltage will remain off, thereby isolating the accessed memory cell from its corresponding bitline.

One technique for minimizing program disturb is to reduce the voltage level of VPASS. This introduces an operational constraint on the Flash memory, whereby the memory cells of the NAND string must be programmed in a sequential pattern, starting from the memory cell farthest from the bitline. In this scheme, all memory cells in the NAND string

between the memory cell being programmed (selected page) and the bitline contact must be in the erased state. Hence a lower VPASS voltage can be used for non-selected cells to ensure that the bitline voltage can be coupled to the selected memory cell in the NAND string. Unfortunately, no random page program operations can be executed, since programmed memory cells (having Vth higher than VPASS for example) between the selected memory cell and the bitline contact will impede the bitline voltage from reaching the selected memory cell. This random page program prohibition causes many restrictions and performance degradation in certain applications.

It is, therefore, desirable to provide a program inhibit scheme that minimizes program disturb in NAND Flash memory cells. More specifically, it is desirable to provide a program inhibit scheme where the VPASS voltage level can be reduced for minimizing program disturb in unselected memory cells where programming is undesired, while maximizing the boosted channel voltage of the selected memory cell where programming is undesired.

SUMMARY OF THE INVENTION

It is an object of the present invention to obviate or mitigate at least one disadvantage of previous Flash memory program disturb inhibit schemes. In particular, it is an object of the present invention to provide a NAND Flash memory program disturb inhibit scheme that minimizes pass disturbance of unselected memory cells during random programming operations.

In a first aspect, the present invention provides a method for minimizing program disturb in NAND string. The NAND string includes a selected memory cell, upper memory cells between the selected memory cell and a bitline, and lower memory cells between the selected memory cell and a sourceline, and a string select transistor for coupling the memory cells to the bitline. The method includes the steps of coupling a voltage level to the NAND string; precharging a channel under the selected memory cell and the upper memory cells; and

locally boosting the selected memory cell channel. The voltage level corresponds to program inhibit data from the bitline. The channel under the selected memory cell and the upper memory cells are precharged to a primary boosted voltage level after the voltage level is coupled to the channel of the NAND string. The selected memory cell channel is locally boosted to a secondary boosted voltage level after the channel is precharged, the secondary boosted voltage level being higher than the primary boosted voltage level.

According to an embodiment of the aspect, the voltage level corresponding to program inhibit data is coupled to the bitline in a first time period, the channel is precharged in a second time period following the first time period, the selected memory cell channel is locally boosted in a third time period following the second time period, and the step of locally boosting includes driving a selected wordline connected to the selected memory cell to a programming voltage level. In an aspect of the present embodiment, the step of coupling can include driving the string select transistor to a first voltage level in the first time period, followed by driving the string select transistor to a decoupling voltage level in the second time period. The decoupling voltage level can be lower than the first voltage level. In another aspect of the present embodiment, the step of precharging can include driving the selected wordline cell and upper wordlines connected to the upper memory cells to a first pass voltage level. The first pass voltage can have a value effective for minimizing program disturb in the upper memory cells and for maximizing the primary boosted voltage level. The first pass voltage can be less than 10 volts, but a 7 volt pass voltage can be used.

In a further aspect of the present embodiment, wherein the step of locally boosting includes electrically turning off the upper memory cell adjacent to the selected memory cell. The step of electrically turning off can include reducing the upper wordline adjacent to the selected wordline from the first pass voltage level to a decoupling voltage level while the selected wordline is driven to the programming voltage level. Alternately, the selected wordline can be driven to the programming voltage level after the upper memory cell adjacent

to the selected memory cell is turned off. More specifically, the selected wordline can be driven to the programming voltage at a delay time after the upper wordline begins to drop towards the decoupling voltage. Alternately, the step of electrically turning off can include raising the upper wordlines except the upper wordline adjacent to the selected wordline, from the first pass voltage level to a second pass voltage while the selected wordline is driven to the programming voltage level, and the upper wordline adjacent to the selected wordline can be maintained at the first pass voltage level.

In yet another aspect of the present embodiment, the step of precharging can include driving lower wordlines connected to the lower memory cells to a second pass voltage, where the second pass voltage is less than the first pass voltage. The step of locally boosting can include reducing the lower wordline adjacent to the selected wordline from the second pass voltage to an off voltage level while the selected wordline is driven to the programming voltage level. Alternately, the step of precharging can include driving lower wordlines connected to the lower memory cells except for the lower wordline adjacent to the selected wordline to a second pass voltage, the lower wordline adjacent to the selected wordline being maintained at an off voltage level during the first time period, the second time period and the third time period. In all the previous embodiments, the programming voltage level can be iteratively increased by predetermined voltage steps while maintaining the first pass voltage level and the second pass voltage level.

In a second aspect, the present invention provides a method for minimizing program disturb in a Flash memory NAND string. The NAND string can have a selected memory cell, upper memory cells between the selected memory cell and a bitline, lower memory cells between the selected memory cell and a source line, and a string select transistor for coupling the memory cells to the bitline. The method includes a) driving the string select transistor for coupling data voltage of the bitline to the NAND string in a first time period; b) driving the upper memory cells and the selected memory cell with a pass voltage level in a second time

period; c) driving the selected memory cell to a program voltage level in a third time period; and d) electrically turning off the upper memory cell adjacent to the selected memory cell in the third time period.

According to embodiments of the present aspect, the step of driving the upper memory cells includes driving the lower memory cells except the lower memory cell adjacent to the selected memory cell with a second pass voltage level in the second time period, where the second pass voltage level is lower than the pass voltage. The step of electrically turning off can include driving an upper memory cell adjacent to the selected memory cell from the pass voltage level to a decoupling voltage while the selected memory cell is driven to the programming voltage level in the third time period. The selected memory cell can be driven to the programming voltage level at a delay time after the upper memory cell adjacent to the selected memory cell starts to be driven to the decoupling voltage. Alternately, the step of electrically turning off can include driving the upper memory cells except an upper memory cell adjacent to the selected memory cell from the pass voltage to a second pass voltage in the third time period, where the second pass voltage is greater than the pass voltage. The program voltage level can be iteratively increased by predetermined voltage steps while maintaining the pass voltage level and the off voltage level.

According to further embodiments, the lower memory cell can be driven to 0V during the second time period, and the step of driving the upper memory cells can include driving the lower memory cells with a second pass voltage level in the second time period, the second pass voltage level being lower than the pass voltage. A lower memory cell adjacent to the selected memory cell can be driven with an off voltage level in the third time period.

Other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments of the invention in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the present invention will now be described, by way of example only, with reference to the attached Figures, wherein:

Fig. 1 is a block diagram of typical Flash memory;

Fig. 2a is a circuit schematic of two NAND memory cell strings;

Fig. 2b is plan view layout of the two NAND memory cell strings shown in Figure 2a;

Fig. 2c is a cross-sectional view of one NAND memory cell string shown in Figure 2b along line A-A';

Fig. 3 is a threshold voltage (Vt) distribution graph for erased memory cells and programmed memory cells;

Fig. 4 is a threshold voltage (Vt) distribution graph for erased memory cells and programmed memory cells after being program disturbed;

Fig. 5 is a flow chart illustrating a general program inhibit method according to an embodiment of the present invention;

Fig. 6 is a flow chart illustrating an embodiment of the general program inhibit method, according to an embodiment of the present invention;

Fig. 7a is a simplified cross sectional view of a NAND string showing applied gate voltages during the program inhibit method, according to an embodiment of the present invention;

Fig. 7b is a sequence diagram further illustrating the sequence of the voltages being applied to the NAND string of Figure 7a;

Fig. 8 is a sequence diagram illustrating the program inhibit sequence according to an embodiment of the present invention;

Fig. 9 is a sequence diagram illustrating the program inhibit sequence according to another embodiment of the present invention;

Fig. 10 is a sequence diagram illustrating the program inhibit sequence of Figure 7b modified to minimize channel leakage;

Fig. 11 is a sequence diagram illustrating the program inhibit sequence of Figure 7b modified to maximize the local secondary boosted voltage level;

Fig. 12 is a sequence diagram illustrating the program inhibit sequence according to another embodiment of the present invention.

Fig. 13 is a simplified cross sectional view of a NAND string showing applied gate voltages, according to an embodiment of the present invention;

Fig. 14 is a sequence diagram further illustrating the sequence of the voltages being applied to the NAND string of Figure 13;

Fig. 15 is a simplified cross sectional view of a NAND string showing applied gate voltages, according to an embodiment of the present invention; and,

Fig. 16 is a sequence diagram further illustrating the sequence of the voltages being applied to the NAND string of Figure 15.

DETAILED DESCRIPTION

Generally, the present invention provides a method and system for minimizing program disturb in Flash memories. To reduce program disturb in a NAND Flash memory cell string where no programming from the erased state is desired, a local boosted channel inhibit scheme is used. In the local boosted channel inhibit scheme, the selected memory cell in a NAND string where no programming is desired, is locally decoupled from the other cells in the NAND string. This allows the channel of the decoupled cell to be locally boosted to a voltage level sufficient for inhibiting F-N tunneling when the corresponding wordline is raised to a programming voltage. Due to the high boosting efficiency, the pass voltage applied to the gates of the remaining memory cells in the NAND string can be reduced relative to

prior art schemes, thereby minimizing program disturb while allowing for random page programming.

Figure 5 is a flow chart illustrating a general program inhibit method according to an embodiment of the present invention. It should be understood that while a selected memory cell of one NAND string is being programmed, a selected memory cell of another NAND string can be inhibited from being programmed. The bitline is biased with a program inhibit voltage at step **100**, which corresponds to a logic "1" state. Then at step **102**, the channel of the NAND string is precharged to a primary boosted voltage level. This can be done by driving all the wordlines to a pass voltage. At step **104**, the channel of a selected memory cell is locally boosted to a secondary boosted voltage level when the corresponding wordline is driven to a programming voltage. This can be done by decoupling, or isolating, the selected memory cell from the rest of the NAND string by controlling the wordline voltage of the memory cells immediately adjacent to the selected memory cell. For the NAND string where programming is to be inhibited, this corresponds to a localized boosting phase where the channel of the selected memory cell is boosted to a level sufficient to inhibit F-N tunneling. The sequence ends at step **106** where all the wordlines and the select lines, such as SSL are driven to 0V.

Figure 6 is a flow chart illustrating an embodiment of the program inhibit method shown in Figure 5. In the present discussion, upper wordlines refer to those wordlines and memory cells between the selected memory cell and the bitline, while lower wordlines refer to those wordlines and memory cells between the selected memory cell and the sourceline. The program inhibit method of Figure 6 begins at step **200**, where the bitlines are biased with an inhibit voltage or a program voltage. In the presently described example, it is presumed that the bitline is driven to an inhibit voltage of VCC. At step **202**, the selected wordline and the upper wordlines are driven to a first pass voltage. This first pass voltage is high enough to turn on each memory cell corresponding to an upper wordline, regardless of its programmed

or unprogrammed state. Concurrently with step **202**, the lower wordlines are driven to a second pass voltage at step **204**. According to embodiments of the present method, the first pass voltage and the second pass voltage can be identical to each other, or the second pass voltage can be lower than the first pass voltage. The selected wordline is then driven to a programming voltage at step **206**, followed by an isolation operation at step **208** for decoupling the selected memory cell from the unselected memory cells of the NAND string. As will be described later, there are different techniques for isolating the selected memory cell. According to another embodiment of the present invention, a memory cell connected to a lower wordline can be turned off for reducing charge leakage from the selected memory cell.

Figure 7a is a simplified cross sectional view of a NAND string **300**, similar to the one shown in Figure 2b. NAND string **300** includes memory cells **302** having gates connected to wordlines WL0 to WL31, a string select transistor **304** for coupling the NAND string **300** to a bitline **306**, and a ground select transistor **308** for coupling the NAND string **300** to a source line **310**. The NAND string **300** is formed within a PP-Well, which is itself formed within an N-Well of the P-substrate. The source/drain regions of the NAND string memory cell transistors are n+ diffusion regions.

Figure 7a includes annotations to show relative voltage levels applied to specific wordlines, according to an embodiment of the present invention. It is noted that the physical NAND string remains unchanged from the prior art. In the example shown in Figure 7a, it is assumed that the memory cell connected to WL29 is selected for programming, and the data to be programmed is the logic "1" state. Since the erased state is a logic "1", programming is to be inhibited. Figure 7b is a sequence diagram further illustrating the sequence of the voltages being applied in Figure 7a.

Time period T1 is an initial state, where the gates of all the memory cells **302**, signal SSL, and signal GSL are biased to 0V, while the bitline voltages for programming memory cells and for inhibiting programming of memory cells are applied. CSL can be biased to either

0V or VCC, but can be preferably biased to VCC in the present example to minimize leakage. In the present example, bitline **306** is biased to VCC. At time period T2, signal SSL is biased to VCC for coupling the bitline voltage to NAND string **300**. In NAND string **300** shown in Figure 7a, string select transistor **304** electrically turns off because its drain terminal connected to the bitline **306** and its gate are both biased to VCC. More specifically, string select transistor 304 will turn off once its source rises to about VCC minus a threshold voltage of the transistor. Time periods T1 and T2 correspond to step **100** in Figure 5. During time period T3, which corresponds to step **102** in Figure 5, all the wordlines (upper, lower and selected wordlines) are driven to a pass voltage V2, where V2 is selected to be higher than VCC. This has the effect of boosting the entire channel of NAND string **300** to the primary boosted voltage. At the beginning of time period T4, the selected wordline WL29 is raised to the programming voltage V1 while the adjacent wordlines WL28 and WL30 are reduced to a decoupling voltage V3. These correspond to steps **104** and **106** respectively in Figure 5. The voltage level of V1 is set to be higher than V2, and the voltage level of V3 is set to be lower than V1 and V2.

A further advantage of reducing the wordline voltage of WL28 and WL30 to V3 is to compensate for capacitive coupling between the control gate connected to WL29 and the floating gates coupled to WL28 and WL30. As WL29 is driven to V1, the capacitive coupling effect can increase the floating gate voltage, thereby increasing the probability of pass disturbance in the memory cells immediately adjacent to the selected memory cell. Therefore, the reduction of the wordline voltage to V3 will reduce the floating gate voltage.

In the embodiment shown in Figure 7a the voltage levels of V1, V2 and V3 should be selected according to the following criteria. V1 should be a voltage level high enough to promote F-N tunneling in a memory cell coupled to a bitline biased to 0V. The voltage level of V2 should be selected to satisfy the following three criteria. First, V2 should be high enough to ensure that the channels of unselected cells are made to be conductive. Second,

when the voltage V2 falls down to the voltage V3 for WL28 and WL30 during time period T4, the memory cells connected to WL28 and WL30 are turned off and remain turned-off during time period T4. This is the local decoupling effect that is achieved. Third, when the selected wordline WL29 rises to the program voltage V1 from voltage V2, the channel of the selected cell can be boosted to a desired voltage, namely to the secondary boosted voltage level. V3 should be a voltage level high enough to pass a 0V bitline voltage through unselected cells to the drain of a selected cell, while being sufficiently low to substantially turn off the channels of the memory cells (connected to WL28 and WL30) immediately adjacent the selected cell (connected to WL29).

A significant advantage of the present program inhibit scheme is that the voltage level of V2 can be reduced relative to the analogous VPASS voltage of the prior art to minimize program disturb in unselected memory cells, while being sufficiently high to pass 0V on a bitline through the unselected memory cells that have been previously programmed to the logic "0" state.

Now that the relative levels of voltages V1, V2 and V3 and the application sequence thereof have been described according to a generic embodiment of the invention, following is a description of example values which can be used. Table 2 lists example voltage values for the embodiment shown in Figure 7a, based on a particular process technology.

Table 2

Parameter	Voltage
VCC	2.5V
Vth_ssl (Threshold voltage of **304**)	0.8V
γ (Cell coupling ratio)	0.8
V1	18V
V2	7V

V3	5V
Vthc (worst case logic "0" programmed cell voltage threshold voltage)	3V

With these values, the primary boosted voltage level for the NAND string channel **300** and the secondary boosted voltage level for the selected memory cell where no programming is desired, can be calculated. The string select transistor **304** is electrically turned off when its source terminal reaches VCC –Vth, due to a gate voltage of VCC and a drain voltage (connected to bitline **306**) of VCC. Thus VCC-Vth_ssl is an initial channel voltage Vich (i.e. starting voltage) for channel boosting.

The primary boosted voltage level (Vprimary) of the channel is calculated according to equation (1) below:

(1) $Vprimary = Vich + \gamma * (Vunsel - Vthc - Vich)$,

where Vunsel is the gate voltage applied to an unselected memory cell, and Vich = VCC-Vth_ssl.

The secondary boosted voltage level Vsecondary of the selected memory cell is calculated according to equation (2) below:

(2) $Vsecondary = Vich + \gamma * (Vsel - Vthc - Vich)$,

where Vsel is the gate voltage applied to a selected memory cell, and Vich = VCC-Vth_ssl.

Vsecondary in equation (2) can be alternately expressed in equation (3) below:

(3) $Vsecondary = Vprimary + \gamma * (Vsel - Vunsel)$, provided $Vunsel > Vthc + Vich$ and $Vsel > Vunsel$

A description of the sequence diagram of Figure 7b follows with the application of the example values of Table 2. With bitline **306** set to 2.5V (VCC), signal SSL rising to 2.5V (VCC) during time period T2 will cause the channel voltage of NAND string **300** to rise to Vich of about 1.7V. When the wordlines all rise to 7V (V2) during time period T3, the channel of NAND string **300** is raised to the primary boosted voltage level of:

Vprimary = 1.7V + 0.8 * (7V – 3V – 1.7V) = 3.54V

During time period T4, the selected wordline rises to 18V (V1), while the two adjacent unselected wordlines WL28 and WL30 fall to 5V (V3). Since unselected cells, except the memory cells connected to WL28 and WL30, retain the boosted channel voltage of 3.54V in T3, the memory cells connected to WL28 and WL30 will be electrically turned off as the gate of the selected cell rises to 18V (V1). Thus the channel of the selected cell is decoupled from the rest of NAND string **300**, and the channel boosting on the selected cell is localized. Localized channel boosting is more efficient than the prior art channel boosting scheme, thereby providing a higher boosted channel voltage. The resultant secondary boosted voltage level is approximately:

Vsecondary = 3.54V + 0.8 (V1-V2) = 12.34V,

Hence, this secondary boosted voltage level is high enough to prevent the selected cell from being programmed. That is, the secondary boosted voltage is sufficiently high for the selected cell to retain its erased state under the V1 program voltage of 18V. In described embodiments, the ratio of the secondary boosted voltage level vs the program voltage V1 is at least about 70%.

The previously described program inhibit sequence applies for any selected memory cell having adjacent memory cells, since the adjacent memory cells can be turned off to decouple the selected memory cell from the rest of the NAND string. However, the NAND string **300** includes the end memory cells connected to wordlines WL0 and WL31 which do

not have a second adjacent memory cell. Following is a discussion of the program inhibit sequence when either end memory cell is to be inhibited from being programmed.

Figure 8 is a sequence diagram illustrating the program inhibit sequence when the memory cell connected to wordline WL31 is to be driven to the programming voltage V1. Immediately adjacent to the memory cell connected to WL31 is the memory cell connected to WL30, and the string select transistor **304** connected to signal SSL. As in the sequence of Figure 7b, bitline **306** is biased to VCC and SSL rises to VCC during time period T2. String select transistor **304** eventually turns off as the coupled channel voltage rises to VCC-Vth_ssl. During time period T3 all the wordlines rise to V2 to raise the channel of NAND string **300** to the primary boosted voltage level. Wordline WL31 rises to V1 and adjacent wordline WL30 falls to V3 during time period T4 to turn off its respective memory cell. Since string select transistor **304** is already turned off, the memory cell connected to WL31 is decoupled from the NAND string **300**, and its channel is locally boosted to the second boosted channel voltage level.

Figure 9 is a sequence diagram illustrating the program inhibit sequence when the memory cell connected to wordline WL0 is to be driven to the programming voltage V1. Immediately adjacent to the memory cell connected to WL0 is the memory cell connected to WL1, and the ground select transistor **308** connected to signal GSL. GSL is biased to 0V in the initial state during time period T1 while bitline **306** is biased to VCC. SSL rises to VCC during time period T2, and all the wordlines rise to V2 to raise the channel of NAND string **300** to the primary boosted voltage level during time period T3. Wordline WL0 rises to V1 and adjacent wordline WL1 falls to V3 during time period T4 to turn off its respective memory cell. Since ground select transistor **308** is already turned off, the memory cell connected to WL1 is decoupled from the NAND string **300**, and its channel is locally boosted to the second boosted channel voltage level.

As shown in Figures 7b to 9, the selected memory cell in a NAND string where no programming is desired can be decoupled from the rest of the NAND memory cells by controlling the sequence and voltages being applied to the unselected memory cells and the select transistors.

During time period T3 in Figures 7b to 9, the channel of the NAND string **300** can leak charge to the bitline through capacitive coupling between WL31 and string select transistor **304**. This has the effect of reducing the primary boosted voltage level. Therefore, according to an embodiment of the present invention, signal SSL can be reduced from VCC to a leakage minimization voltage V4. By example, for the NAND string **300** shown in Figure 7a, V4 can be 1V. By dropping SSL to V4, the channel of string select transistor **304** becomes more resistant to leaking charge to bitline **306**. V4 is can be selected to be at a level sufficient to pass a 0V level on a bitline to the selected memory cell to be programmed.

Figure 10 is a sequence diagram illustrating the program inhibit sequence of Figure 7b modified to minimize channel leakage. At the initial state during time period T1, the gates of all the memory cells **302**, signal SSL, and signal GSL are biased to 0V, while bitline voltage (VCC) is applied and CSL is biased to VCC. At time period T2, signal SSL is biased to VCC for coupling the bitline voltage to NAND string **300**. In NAND string **300** shown in Figure 7a, string select transistor **304** is turned off because its drain terminal connected to the bitline **306** and its gate are both biased to VCC. More specifically, string select transistor 304 will turn off once its channel rises to about VCC minus a threshold voltage of the transistor. During time period T3, all the wordlines are driven to a pass voltage V2 for boosting the entire channel of NAND string **300** to the primary boosted voltage. However, to minimize leakage of the primary boosted voltage to bitline **306**, SSL is reduced to voltage V4 as the wordlines are driven to V2. At the beginning of time period T4, the selected wordline WL29 is raised to the programming voltage V1 while the adjacent wordlines WL28 and WL30 are reduced to a decoupling voltage V3. The SSL voltage reduction to the leakage minimization voltage V4

can be used in the program inhibit sequences shown in Figures 8 and 9 with equal effectiveness.

The local secondary boosted voltage level of the selected memory cell where no programming is desired can be further increased, according to another embodiment of the present invention. In the previous program inhibit sequence embodiments shown in Figures 7a to 10, all the wordlines are driven to the V2 pass voltage during time period T3, after which, only the selected wordline is driven from V2 to the program voltage V1. The local secondary boosted voltage level can be maximized when the difference between V2 and V1 is maximized. Hence, according to an embodiment of the present invention, the voltage level of V2 for the selected wordline is reduced relative to the previously shown program inhibit sequences.

Figure 11 is a sequence diagram illustrating the program inhibit sequence of Figure 7b modified to maximize the local secondary boosted voltage level in the channel of the selected transistor where programming is to be inhibited. As shown during time period T3, all the wordlines are driven to the V2 pass voltage except for the selected wordline WL29, which is driven to a lower voltage of V5. Subsequently in time period T4, WL29 is driven to the programming voltage V1. For the NAND string **300** shown in Figure 7a, V5 can be set to a voltage level of 5V, for example. It is noted that V5 is selected to be sufficiently high to ensure conductivity of the channel of the selected cell during time period T3. Therefore, since the difference between V5 and V1 is greater than the difference between V2 and V1, a higher local secondary boosted voltage level can be obtained in the selected memory cell. This is referred to as the high differential local boosting scheme.

The following relational expressions summarize the relative constraints of the voltage levels used in the previously described embodiments of the invention.

(1) $V3 < V2 < V1$

(2) $0V < V4 < V3$

(3) 0V < V5 < V2

The actual values depend on memory cell and transistor geometry, and the process being used. V1 is a programming voltage sufficient for inducing F-N tunneling. V2 is a pass voltage i) high enough for making the memory cells conductive; ii) high enough such that a drop to V3 will turn off the memory cells it is connected to; iii) high enough such that an increase to V1 will locally boost the selected memory channel to the desired secondary boosted voltage level. V2 can be higher than VCC. V3 is high enough to pass a 0V bitline voltage, and should be at least as high as the worst case threshold voltage of a programmed memory cell. In Table 1 for example, if Vthc is 3 volts, then V3 should be at least 3 volts.

Various program inhibit sequences according to embodiments of the present invention have been shown in Figures 7b to 11. Embodiments have been shown for reducing leakage or for increasing the secondary boosted voltage level during programming operations. These embodiments can be combined together to attain all the benefits each individual scheme has to offer. As shown in the sequence diagram of Figure 12, the SSL signal can be reduced to the leakage minimization voltage V4 during time period T3 while the selected wordline WL29 is set to a reduced pass voltage of V5 during time period T3.

In the previously described embodiments, the SSL signal can be reduced to the leakage minimization voltage V4 to minimize charge leakage to bitline **306**, thereby maximizing the boosted channel voltage prior to localized boosting of the selected memory cell channel to the secondary boosted voltage level. During localized boosting, some charge can leak from the channel of the selected memory cell through its two immediately adjacent memory cells. This is caused by dropping the wordlines of the adjacent cells from V2 to V3, which reduces the channel voltage of their respective memory cells. Therefore, while still turned off, some of the charge will leak to the other memory cells connected to the lower wordlines.

According to a charge leakage reduction embodiment of the present invention, the lower wordlines are driven with voltages different than the upper wordlines for reducing charge leakage from the locally boosted channel of the selected memory cell. An embodiment of this scheme is shown in Figure 13.

Figure 13 shows the same NAND string 300 shown in Figure 7a, and includes annotations to show relative voltage levels applied to specific wordlines. The voltages applied to the selected wordline WL29, and upper wordlines WL30 and WL31 can be the same as previously shown and described for the embodiment of Figure 7a. In the present embodiment the lower wordlines WL0-WL28 are driven to a pass voltage of V6, and the immediately adjacent wordline WL28 is then driven down to a low voltage level. By example, the low voltage can be 0V, 0.1V, 0.2V.

Figure 14 is a sequence diagram illustrating the program inhibit sequence according to the currently described embodiment. The sequence during time period T1 and T2 is the same as previously described for Figure 7b. At time period T3, the upper wordlines WL30-WL31, and the selected wordline, are driven to the first pass voltage V2. The lower wordlines WL0-WL28 are driven to a second pass voltage V6. In another embodiment, second pass voltage V6 is less than first pass voltage V2, but higher than the threshold of a programmed memory cell. A primary boosted voltage will still result because charge is distributed, or shared, throughout the entire channel of the NAND memory cell string. At the beginning of time period T4, the selected wordline WL29 is raised to the programming voltage V1 while the upper adjacent wordline WL30 is reduced to the decoupling voltage V3. The lower adjacent wordline WL28 on the other hand, is reduced to an off voltage of 0V. The effect of the aforementioned sequence and voltage levels will now be discussed.

If the memory transistor connected to WL28 has been programmed with a positive threshold, then it will be electrically turned off when WL28 falls below the positive threshold, such as to 0V for example. On the other hand, if the memory cell is still erased (negative

threshold), then the off voltage of even 0V can turn on the memory cell. However, since the second pass voltage V6 was applied to the lower wordlines, the channels of the lower memory cells are boosted to the lower primary boosted voltage. Hence the source terminal (drain terminal is connected to selected memory cell) of the memory cell connected to WL28 is positive, thereby turning off the memory cell. By closing the leakage path from the boosted channel of the selected transistor to the lower memory cells, local boosting efficiency is further improved. Those skilled in the art will understand that the off voltage is not restricted to being 0V, and that any low voltage effective for minimizing charge leakage through the memory cell connected to WL28 can be used.

According to alternate embodiments, string select signal SSL can be reduced to V4 at the beginning of time period T4 as shown by the dashed line in the trace of SSL, and WL28 can be set to remain at the off voltage during time period T3, as shown by the dashed line in the trace of WL28. Advantages of keeping the adjacent lower wordline (such as WL28 in the described embodiment) at the off voltage include a reduction in voltage consumption and coupling capacitance. Voltage consumption is reduced since the wordline driver does not need to raise and lower the wordline. As wordline pitches are decreased to improve memory array packing density, capacitive coupling between adjacent wordlines becomes more pronounced as wordlines rise and fall. Since the adjacent lower wordline does not change between time period T3 and T4, capacitive coupling is reduced. The second pass voltage V6 can be a low voltage, such as a read voltage level used for Flash read operations, and is lower than V3.

Figure 15 is another embodiment of the present invention, combining some of the features described in the previous embodiments. Figure 15 shows the same NAND string **300** shown in Figure 13, and includes annotations to show relative voltage levels applied to specific wordlines. In particular, the present example uses the charge leakage reduction scheme of Figure 13, a variant of the high differential local boosting scheme of Figure 11, and a novel decoupling scheme. In the novel decoupling scheme, the first and second upper

adjacent memory cells of NAND string **300** are controlled using the previously established voltages for decoupling the selected memory cell from the NAND string **300**.

Figure 16 is a sequence diagram illustrating the program inhibit sequence according to the currently described embodiment. Figure 16 shows relative timing variations of WL29, and an alternate pass voltage of WL29, which will be described later. The sequence during time period T1 and T2 is the same as previously described for Figure 13. At time period T3, all the upper wordlines (WL30 and WL31 in the present example) and the selected wordline WL29, are driven to the decoupling voltage V3. In effect, V3 is being used in time period T3 as a first pass voltage. Meanwhile, the lower wordlines are driven to the second pass voltage V6, including adjacent lower wordline WL28 which remains at the default off voltage level (0V for example). Hence the channel of NAND string **300** is now effectively restricted to the memory cells connected to wordlines WL29-WL31, which is raised to the primary boosted voltage.

At the beginning of time period T4, the selected wordline WL29 is raised to the programming voltage V1 while the upper adjacent wordline WL30 remains at V3. At about the same time, all the upper wordlines except the upper adjacent wordline WL31 are driven up to the first pass voltage of V2. In this embodiment, the memory cell connected to WL30 is electrically turned off when WL29 is raised to V1 and when WL31 is raised to V2, and WL28 is reduced to the off voltage in time period T4. In an alternate embodiment, the string select signal SSL can be reduced to V4 at the beginning of time period T4 as shown by the dashed line in the trace of SSL. The following relational expressions summarize the relative constraints of the voltage levels used in the embodiments of Figures 13 to 16.

(1) $V3 < V2 < V1$

(2) $0V < V4 < V3$

(3) $0V < V6 \leq V3$

Therefore, there is virtually no pass disturbance experienced by the lower memory cells due to the low voltage level being applied to their gates.

In Figure 16, WL29 can be initially set to V3 during time period T3 to increase the voltage difference between V3 and V1, and to provide high differential local boosting similar to that achievable by the embodiment of Figure 11. Alternatively, local boosting can be maximized by maintaining WL29 at 0V during time period T3, and then driving WL29 directly to V1 during time period T4. This is shown by dashed line 400. It is noted that the memory cell connected to WL29 is an erased cell having a negative threshold voltage.

The previously shown sequence diagrams are intended to show the general timing relationships for the application of specific voltages to selected and unselected wordlines. However, the relative timing between driving the selected wordline from the first pass voltage (or 0V) to the programming voltage (V1) and control of the immediately adjacent wordlines can be further refined as follows. With reference to Figure 10 for example, the selected wordline WL29 can be driven to V1 at the same time the adjacent wordlines WL28 and WL30 are reduced to V3. In other words, the selected memory cell is decoupled from the rest of the NAND memory cells as the V1 programming voltage is applied. During the initial part of the transition of WL29 to V1, some off the charge resulting from the WL29 boost will leak away through adjacent memory cells before these cells have been fully turned off by voltage V3.

To maximize local boosting efficiency of the selected memory cell during time period T4, the selected wordline can be driven to V1 shortly after the selected memory cell is decoupled from the other memory cells. In the embodiment of Figure 16 for example, WL29 can remain at V3 until after WL28 reaches 0V, as shown in trace 402. More specifically as shown in Figure 16, WL29 can be driven to V1 after a delay time td after WL28 starts to drop to 0V. Delay time td can be any preset value. For the embodiment where WL29 remains at 0V during time period T3, WL29 can remain at 0V until after WL28 reaches 0V after delay time

td, as shown in trace 404. Those skilled in the art will understand that the relative timing between the selected wordline with the other wordlines can be configured as required to minimize charge leakage or to maximize local channel boosting.

In the previously described embodiments, V2 is a value optimized for minimizing program disturb in unselected memory cells, and for maximizing the primary boosted channel level. Those skilled in the art will understand that a practical V2 value will depend on various design and fabrication parameters of the memory cells, and can therefore be determined through simulation or modeling.

The presently described embodiments of the invention can be used in standard stepped programming schemes. According to an embodiment of the present invention, the selected memory cell program voltage V1 will be an initial programming voltage applied to the gate of the selected memory cell. Thereafter, the voltage level of V1 can be iteratively increased by predetermined voltage steps while the pass voltages and decoupling voltages applied to the other wordlines are kept at the same voltage levels. By fixing these voltage levels during the incremental programming steps, chip size and power consumption can be reduced relative to prior art programming schemes.

In the described embodiments, the off voltage applied to the adjacent lower wordline (ie. WL28) can be 0V, or any low voltage that is selected based on the chosen design and fabrication parameters for the memory cells, and the boosted channel voltage resulting from the application of the V6 pass voltage.

The wordline and control line (SSL and GSL) voltages and sequence of application can be controlled by wordline driver circuits, and associated wordline logic circuits, such as decoders. Those skilled in the art will understand that well known multi-voltage level wordline driver circuits can be easily modified to provide the voltage levels used in the illustrated program inhibit sequence embodiments. The decoding logic can ensure that for any

selected wordline, the immediately adjacent wordlines (or control lines) are controlled in the disclosed manner for decoupling the selected memory cell from the NAND string.

Therefore, the aforementioned wordline/control line control scheme is effective for programming selected memory cells from an erased state, while inhibiting programming in selected memory cells where the erase state is to be retained. Since a lower pass voltage is used relative to the prior art program inhibit schemes, program disturb in unselected memory cells is minimized. Since the selected memory cell can be decoupled and isolated from the other memory cells in the NAND string, its channel can be efficiently boosted to a level for inhibiting F-N tunneling and random page programming can be performed.

The previously described embodiments of the program inhibit scheme can be implemented for any NAND string based Flash memory. Those of skill in the art will understand that the specific voltage levels used in the scheme are unique to a particular process and NAND cell design. The voltage level transitions shown in Figures 7b to 16 are intended to provide a general sequence of events, and not intended to provide specific timing relationships between signals. Persons skilled in the art will appreciate that minor timing differences between signals can occur without departing from the scope of the invention.

The above-described embodiments of the present invention are intended to be examples only. Alterations, modifications and variations may be effected to the particular embodiments by those of skill in the art without departing from the scope of the invention, which is defined solely by the claims appended hereto.

What is claimed is:

1. A method for minimizing program disturb in NAND string having a selected memory cell, upper memory cells between the selected memory cell and a bitline, lower memory cells between the selected memory cell and a sourceline, and a string select transistor for coupling the memory cells to the bitline, comprising:

a) coupling a voltage level corresponding to program inhibit voltage from the bitline to the NAND string;

b) precharging channels of the selected memory cell and the upper memory cells to a primary boosted voltage level after the voltage level is coupled to the channel of the NAND string; and,

c) locally boosting the selected memory cell channel to a secondary boosted voltage level after the channel is precharged, the secondary boosted voltage level being higher than the primary boosted voltage level.

2. The method for minimizing program disturb of claim 1, wherein the voltage level corresponding to program inhibit voltage is coupled to the bitline in a first time period.

3. The method for minimizing program disturb of claim 2, wherein the channel is precharged in a second time period following the first time period.

4. The method for minimizing program disturb of claim 3, wherein the selected memory cell channel is locally boosted in a third time period following the second time period.

5. The method for minimizing program disturb of claim 4, wherein the step of locally boosting includes driving a selected wordline connected to the selected memory cell to a programming voltage level.

6. The method for minimizing program disturb of claim 5, wherein the step of coupling includes driving the string select transistor to a first voltage level in the first time period, followed by driving the string select transistor to a decoupling voltage level in the second time period, the decoupling voltage level being lower than the first voltage level.

7. The method for minimizing program disturb of claim 5, wherein the step of precharging includes driving the selected wordline and upper wordlines connected to the upper memory cells to a first pass voltage level.

8. The method for minimizing program disturb of claim 7, wherein the first pass voltage is less than 10 volts.

9. The method for minimizing program disturb of claim 8, wherein the first pass voltage is approximately 7 volts.

10. The method for minimizing program disturb of claim 7, wherein the first pass voltage has a value effective for minimizing program disturb in the upper memory cells and for maximizing the primary boosted voltage level.

11. The method for minimizing program disturb of claim 7, wherein the step of locally boosting includes electrically turning off an upper memory cell adjacent to the selected memory cell.

12. The method for minimizing program disturb of claim 11, wherein the step of driving the selected wordline includes driving the selected wordline to the programming voltage level after the upper memory cell adjacent to the selected memory cell is turned off.

13. The method for minimizing program disturb of claim 11, wherein the step of electrically turning off includes reducing an upper wordline connected to the upper memory cell from the first pass voltage level to a decoupling voltage level while the selected wordline is driven to the programming voltage level.

14. The method for minimizing program disturb of claim 11, wherein the step of electrically turning off includes reducing an upper wordline connected to the upper memory cell from the first pass voltage level to a decoupling voltage level, and driving the selected wordline to the programming voltage level at a predetermined delay time after the upper wordline begins to drop towards the decoupling voltage level.

15. The method for minimizing program disturb of claim 11, wherein the step of electrically turning off includes increasing the upper wordlines except the upper wordline adjacent to the selected wordline, from the first pass voltage level to a second pass voltage while the selected wordline is driven to the programming voltage level, the upper wordline adjacent to the selected wordline being maintained at the first pass voltage level.

16. The method for minimizing program disturb of claim 11, wherein the step of precharging includes driving lower wordlines connected to the lower memory cells to a second pass voltage, the second pass voltage being less than the first pass voltage.

17. The method for minimizing program disturb of claim 16, wherein the step of locally boosting includes reducing a lower wordline adjacent to the selected wordline from the second pass voltage to an off voltage level while the selected wordline is driven to the programming voltage level.

18. The method for minimizing program disturb of claim 11, wherein the step of precharging includes driving lower wordlines connected to the lower memory cells except for a lower wordline adjacent to the selected wordline to a second pass voltage, the lower wordline adjacent to the selected wordline being maintained at an off voltage level during the first time period, the second time period and the third time period.

19. The method for minimizing program disturb of claim 16, further including iteratively increasing the programming voltage level by predetermined voltage steps while maintaining the first pass voltage level and the second pass voltage level.

20. A method for minimizing program disturb in a Flash memory NAND string, the NAND string having a selected memory cell, upper memory cells between the selected memory cell and a bitline, lower memory cells between the selected memory cell and a source line, and a string select transistor for coupling the memory cells to the bitline, comprising:

a) driving the string select transistor for coupling data voltage of the bitline to the NAND string in a first time period;

b) driving the upper memory cells and the selected memory cell with a pass voltage level in a second time period;

c) driving the selected memory cell to a program voltage level in a third time period; and

d) electrically turning off the upper memory cell adjacent to the selected memory cell in the third time period.

21. The method for minimizing program disturb of claim 20, wherein the step of driving the upper memory cells includes driving the lower memory cells except a lower memory cell adjacent to the selected memory cell with a second pass voltage level in the second time period, the second pass voltage level being lower than the pass voltage.

22. The method for minimizing program disturb of claim 21, wherein the lower memory cell is driven to 0V during the second time period.

23. The method for minimizing program disturb of claim 20, wherein the step of driving the upper memory cells includes driving the lower memory cells with a second pass voltage level in the second time period, the second pass voltage level being lower than the pass voltage.

24. The method for minimizing program disturb of claim 23, further including driving a lower memory cell adjacent to the selected memory cell with an off voltage level in the third time period.

25. The method for minimizing program disturb of claim 20, wherein the step of electrically turning off includes driving an upper memory cell adjacent to the selected memory cell from the pass voltage level to a decoupling voltage while the selected memory cell is driven to the programming voltage level in the third time period.

26. The method for minimizing program disturb of claim 25, wherein the selected memory cell is driven to the programming voltage level at a delay time after the upper memory cell adjacent to the selected memory cell starts to be driven to the decoupling voltage.

27. The method for minimizing program disturb of claim 20, wherein the step of electrically turning off includes driving the upper memory cells except an upper memory cell adjacent to the selected memory cell from the pass voltage to a second pass voltage in the third time period, the second pass voltage being greater than the pass voltage.

28. The method for minimizing program disturb of claim 20, further including iteratively increasing the program voltage level by predetermined voltage steps while maintaining the pass voltage level and the off voltage level.

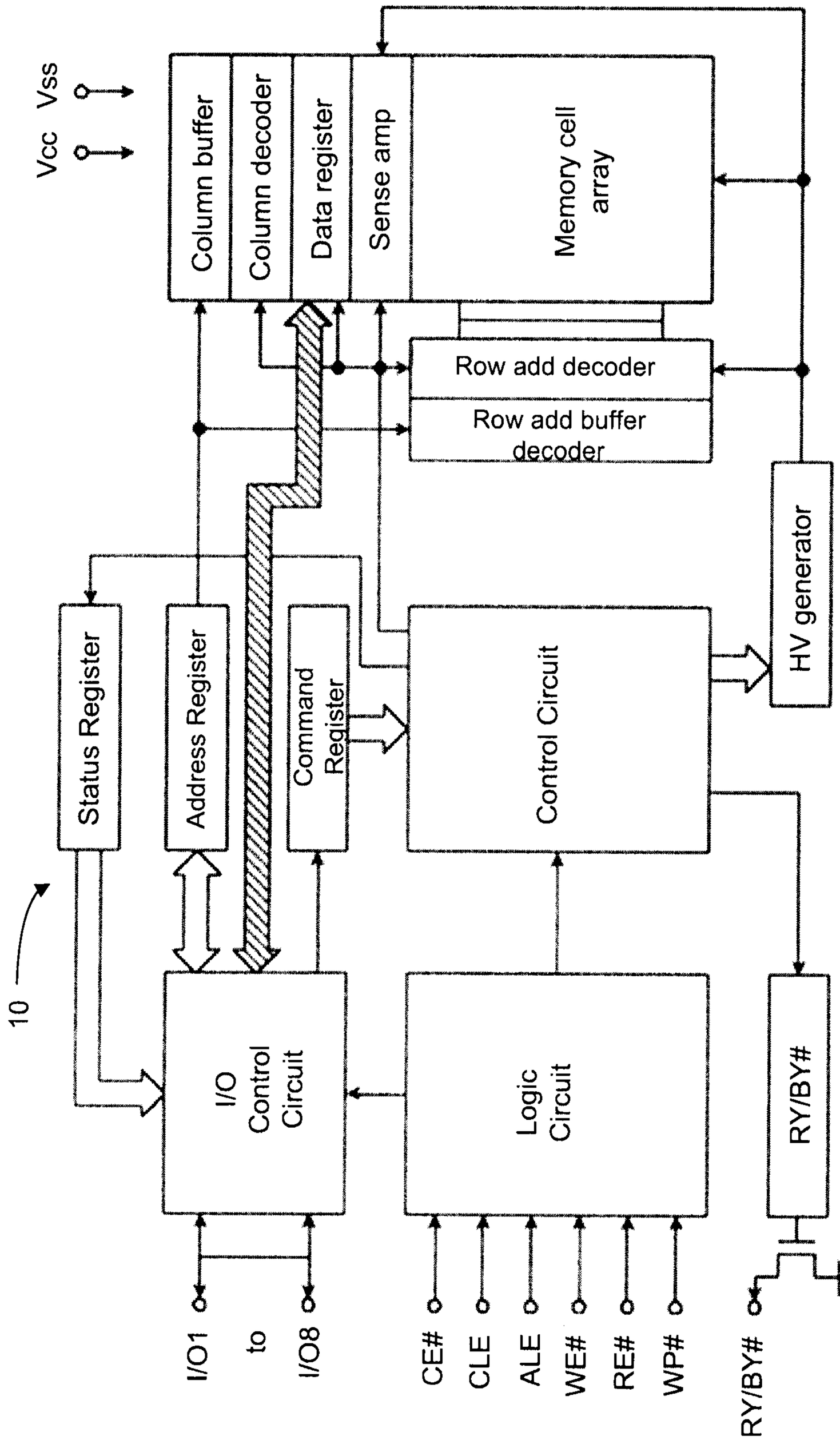

Vcc
Vss
Column buffer
Column decoder
Data register
Sense amp
Memory cell array
Row add decoder
Row add buffer decoder
HV generator
Status Register
Address Register
Command Register
Control Circuit
10
I/O Control Circuit
Logic Circuit
RY/BY#
I/O1
to
I/O8
CE#
CLE
ALE
WE#
RE#
WP#
RY/BY#

Figure 1

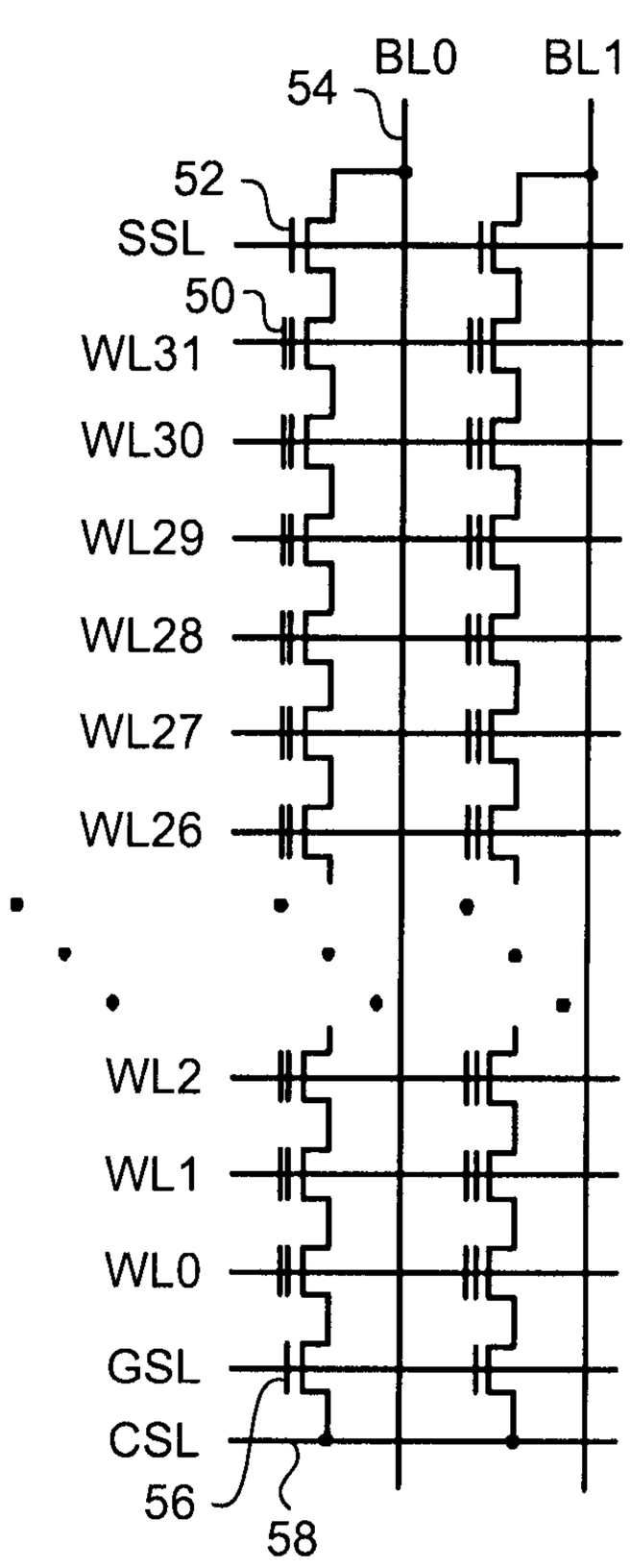


Figure 2a

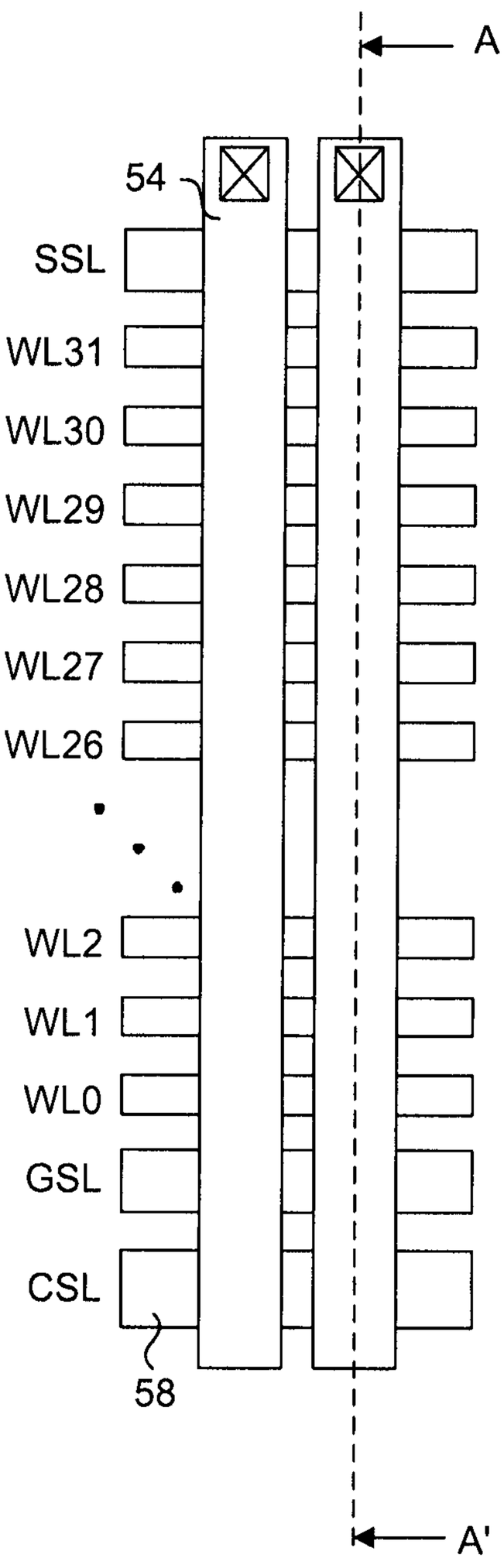


Figure 2b

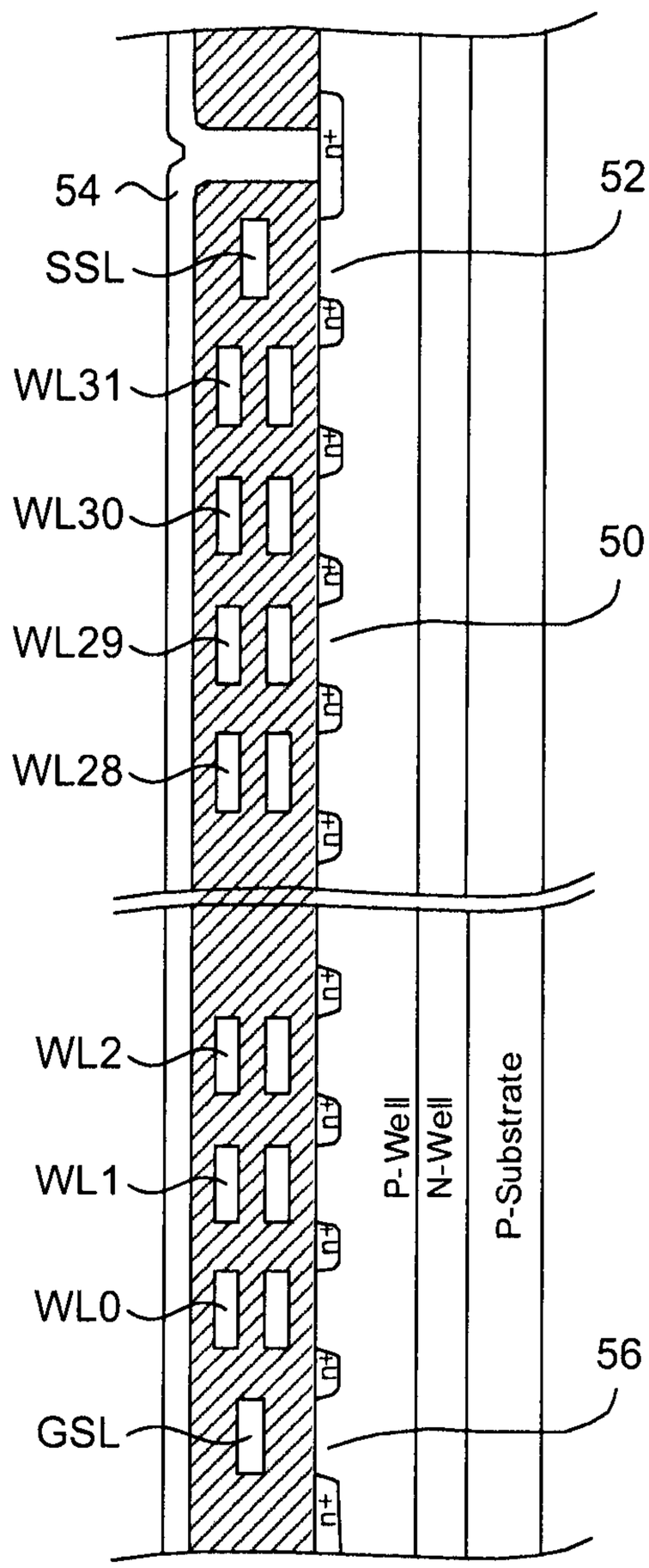

54
SSL
WL31
WL30
WL29
WL28
WL2
WL1
WL0
GSL
52
50
56
P-Well
N-Well
P-Substrate
n+

Figure 2c

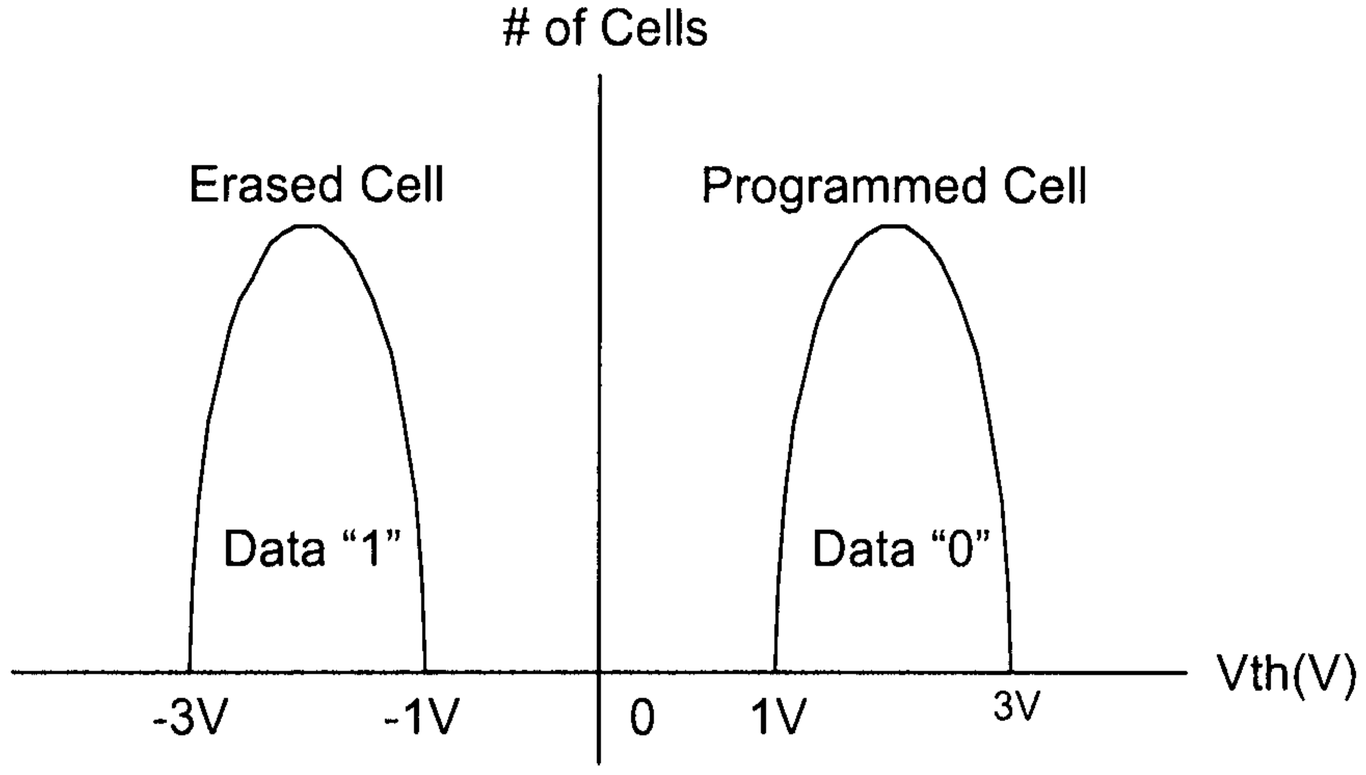


Figure 3

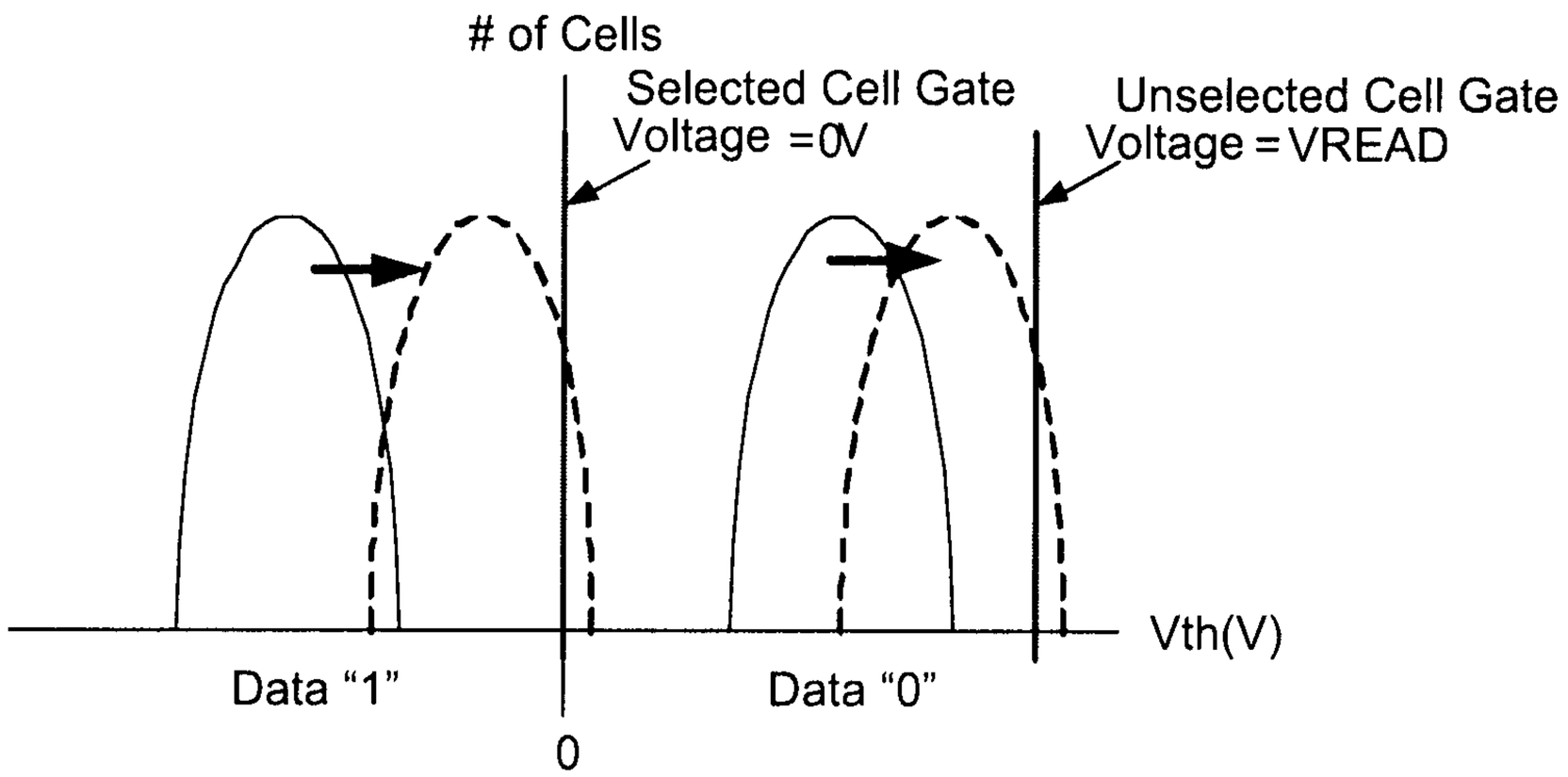


Figure 4

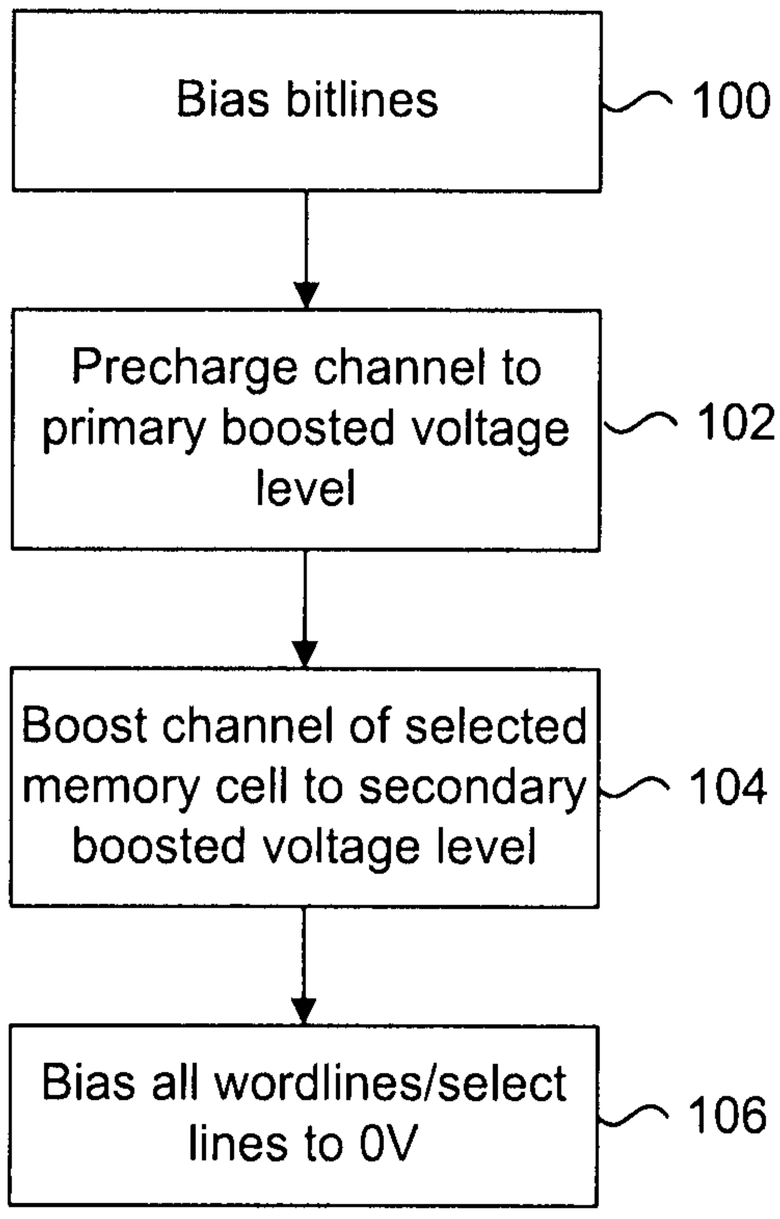

Bias bitlines
100
Precharge channel to primary boosted voltage level
102
Boost channel of selected memory cell to secondary boosted voltage level
104
Bias all wordlines/select lines to 0V
106

Figure 5

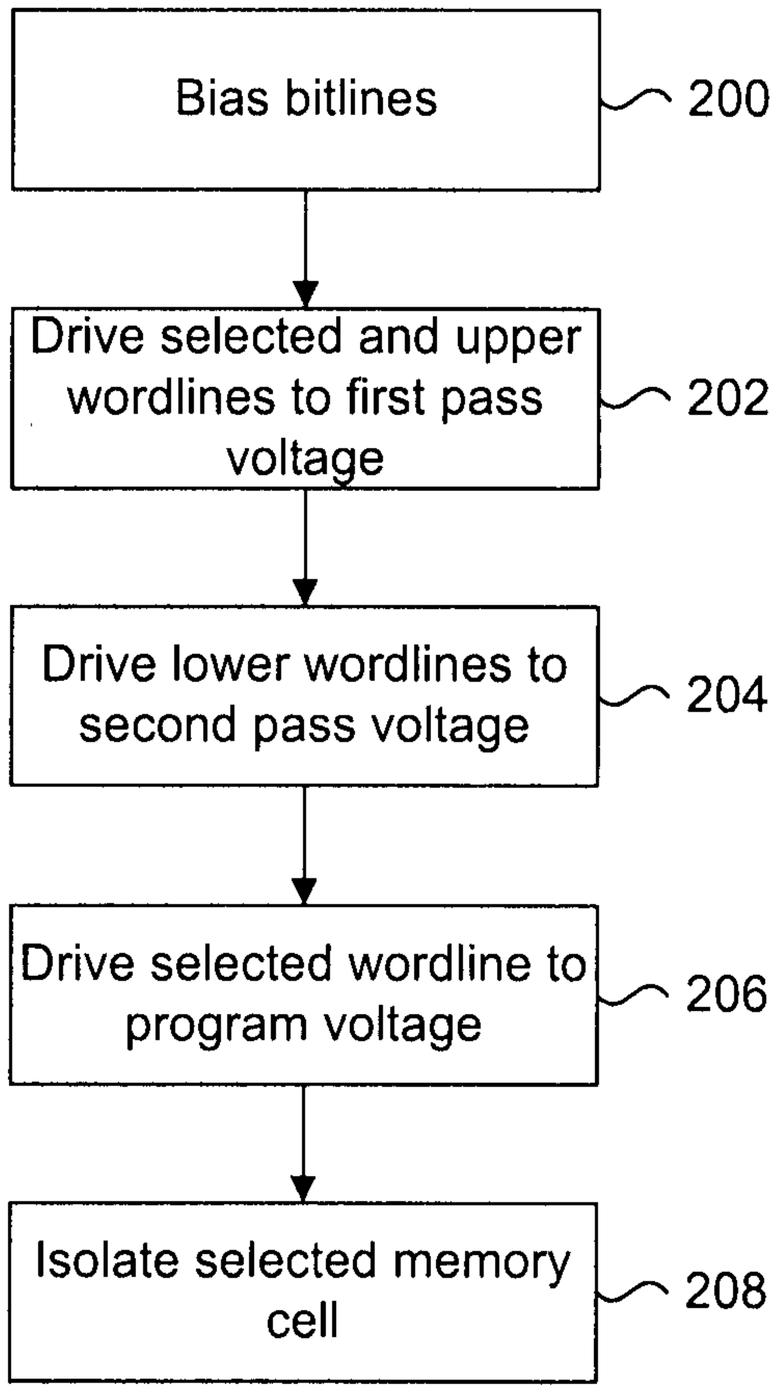

Bias bitlines
200
Drive selected and upper wordlines to first pass voltage
202
Drive lower wordlines to second pass voltage
204
Drive selected wordline to program voltage
206
Isolate selected memory cell
208

Figure 6

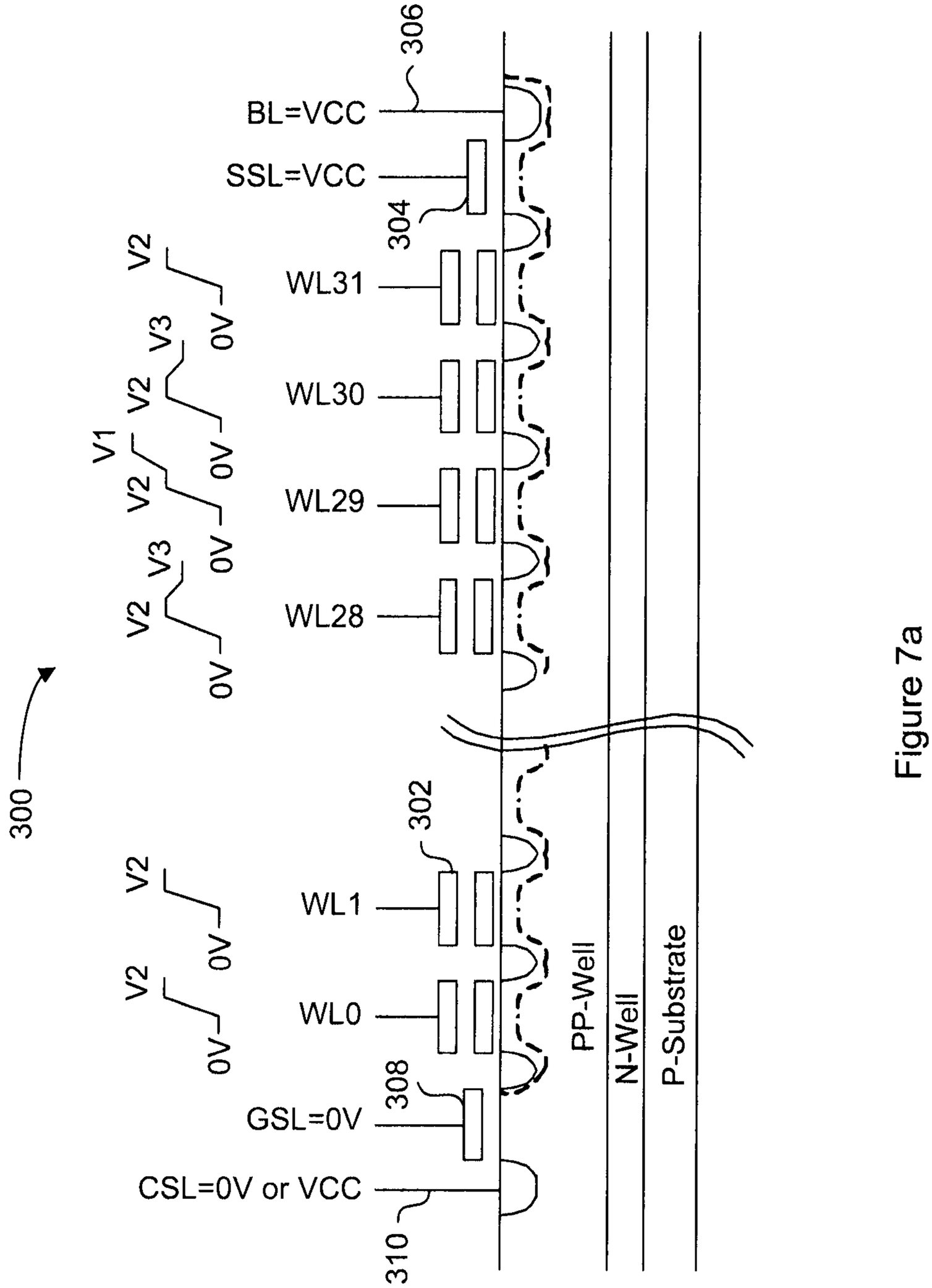

300
BL=VCC
SSL=VCC
WL31
WL30
WL29
WL28
WL1
WL0
GSL=0V
CSL=0V or VCC
V1
V2
V3
0V
302
304
308
310
PP-Well
N-Well
P-Substrate

Figure 7a

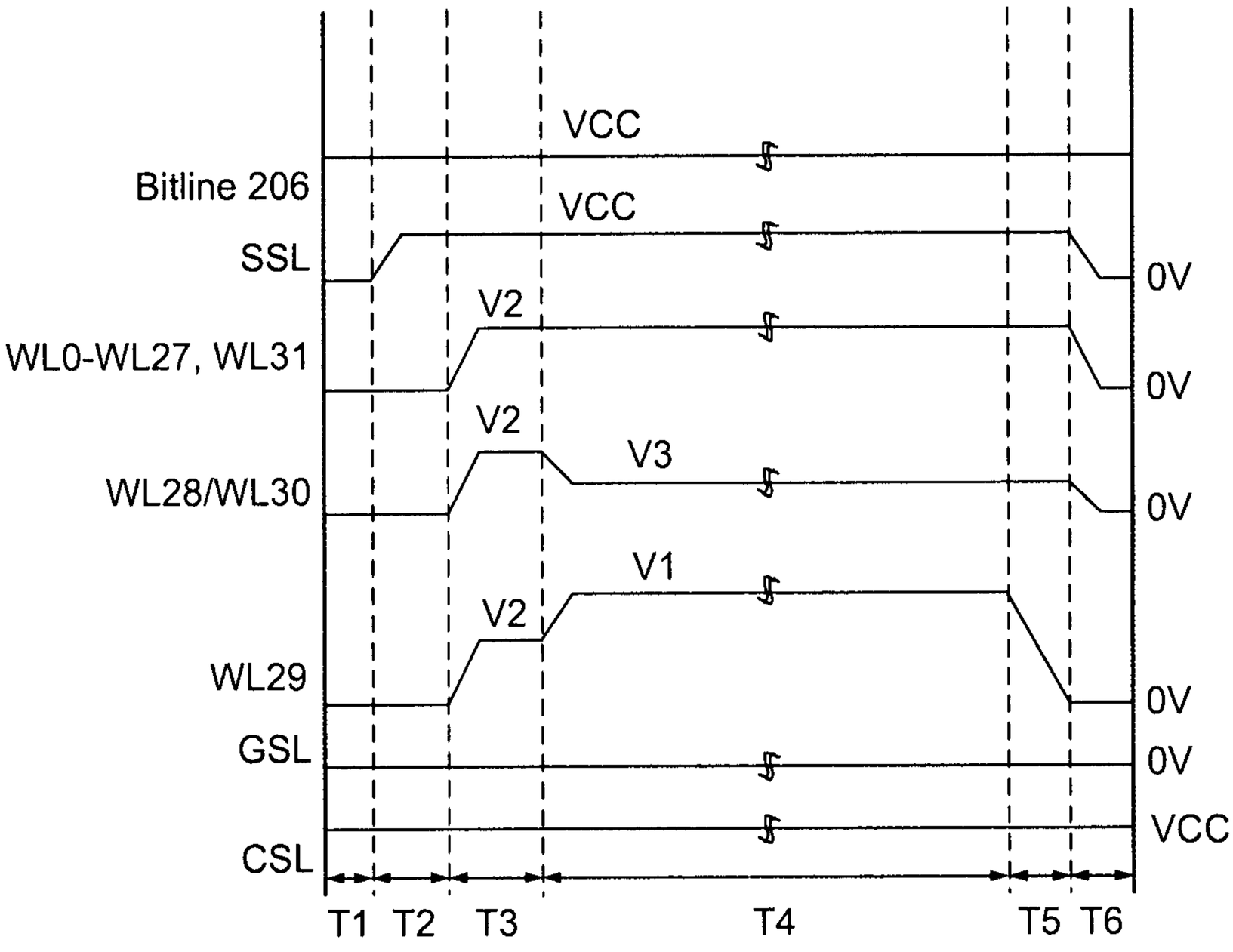

Bitline 206
SSL
WL0-WL27, WL31
WL28/WL30
WL29
GSL
CSL
VCC
VCC
V2
V2
V3
V1
V2
0V
0V
0V
0V
0V
VCC
T1
T2
T3
T4
T5
T6

Figure 7b

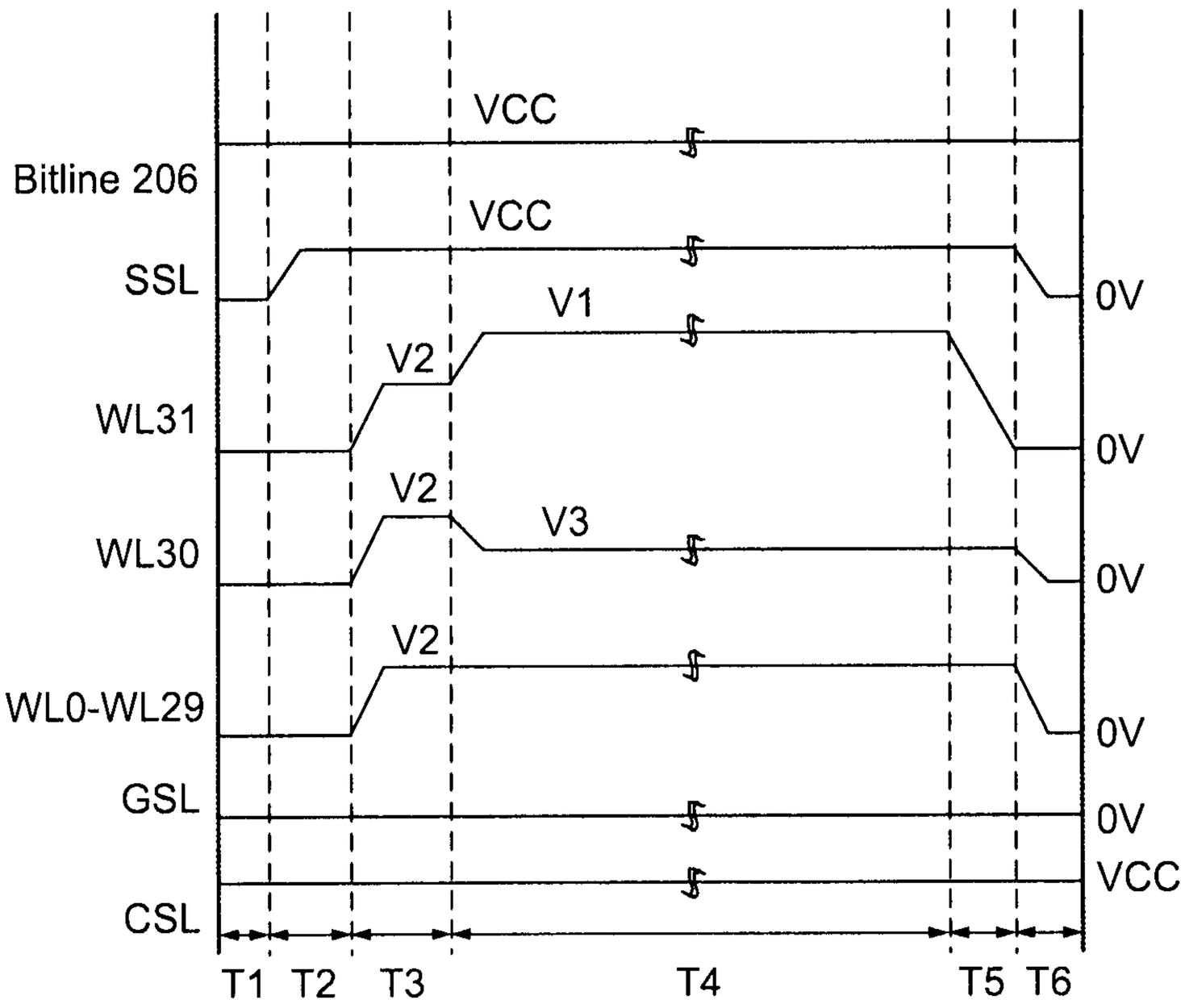

VCC
Bitline 206
VCC
SSL
V1
V2
WL31
V2
V3
WL30
V2
WL0-WL29
GSL
CSL
0V
0V
0V
0V
VCC
T1
T2
T3
T4
T5
T6

Figure 8

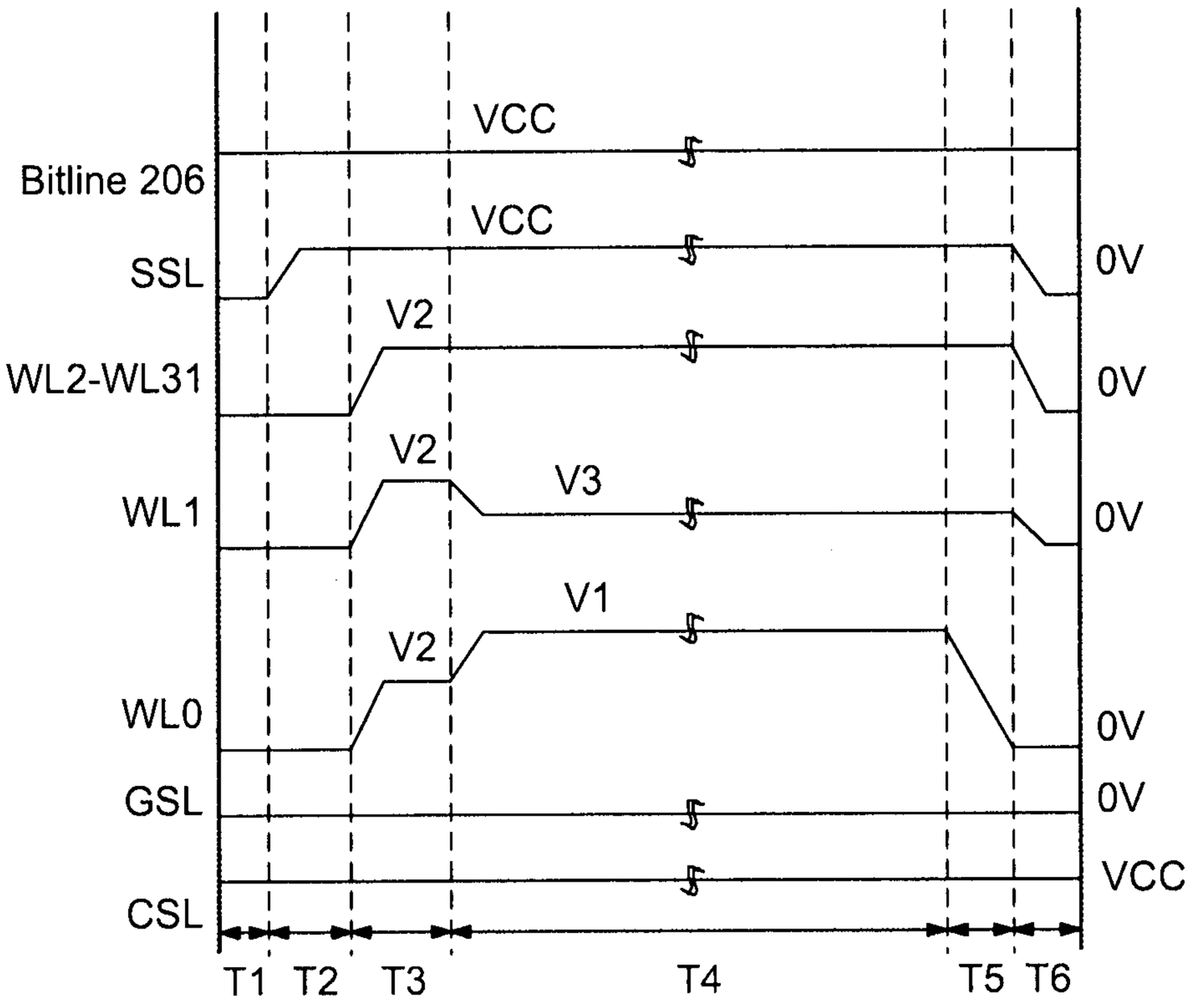

VCC
Bitline 206
VCC
SSL
0V
V2
WL2-WL31
0V
V2
V3
WL1
0V
V1
V2
WL0
0V
GSL
0V
VCC
CSL
T1
T2
T3
T4
T5
T6

Figure 9

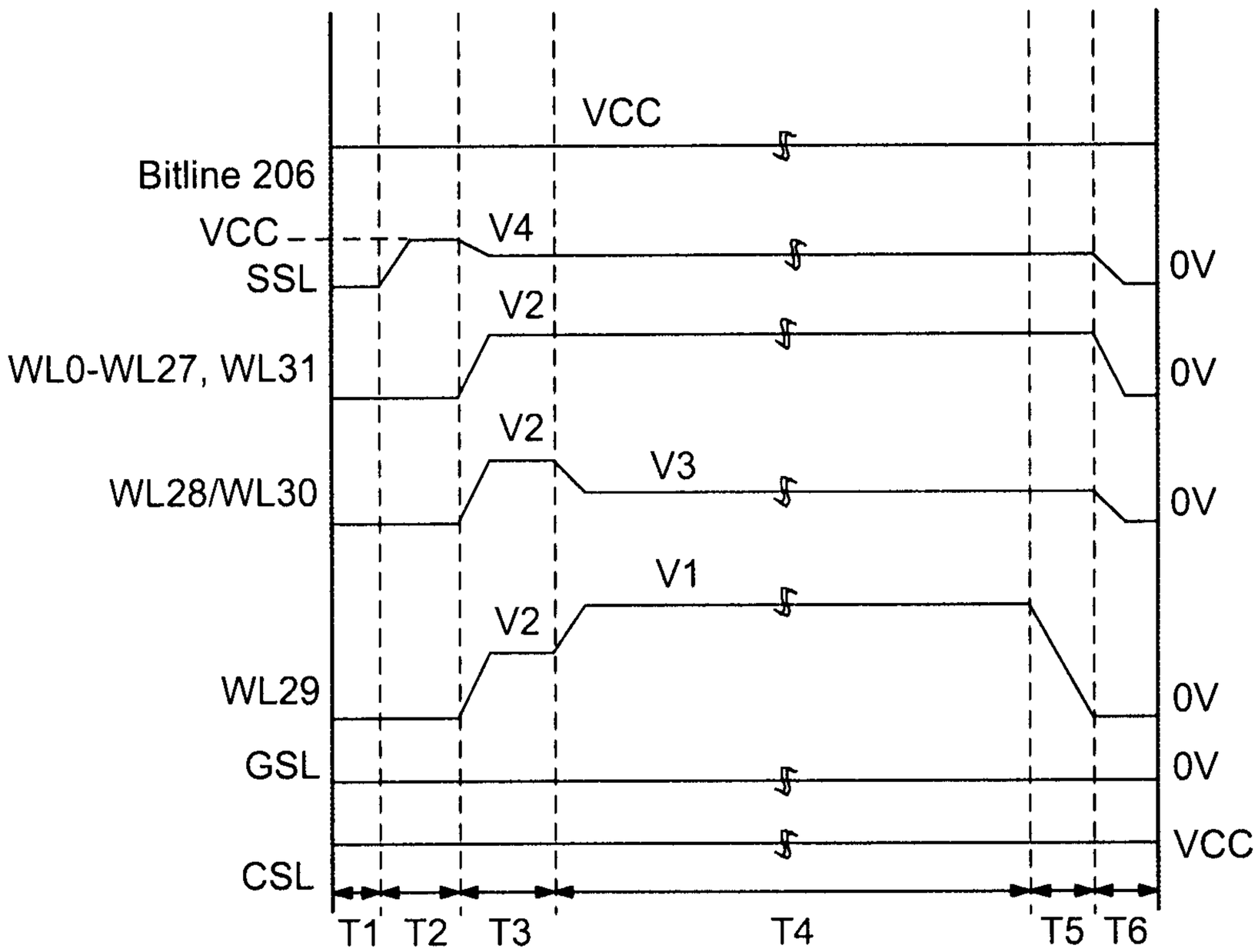

VCC
Bitline 206
VCC
V4
SSL
0V
V2
WL0-WL27, WL31
0V
V2
V3
WL28/WL30
0V
V1
V2
WL29
0V
GSL
0V
VCC
CSL
T1
T2
T3
T4
T5
T6

Figure 10

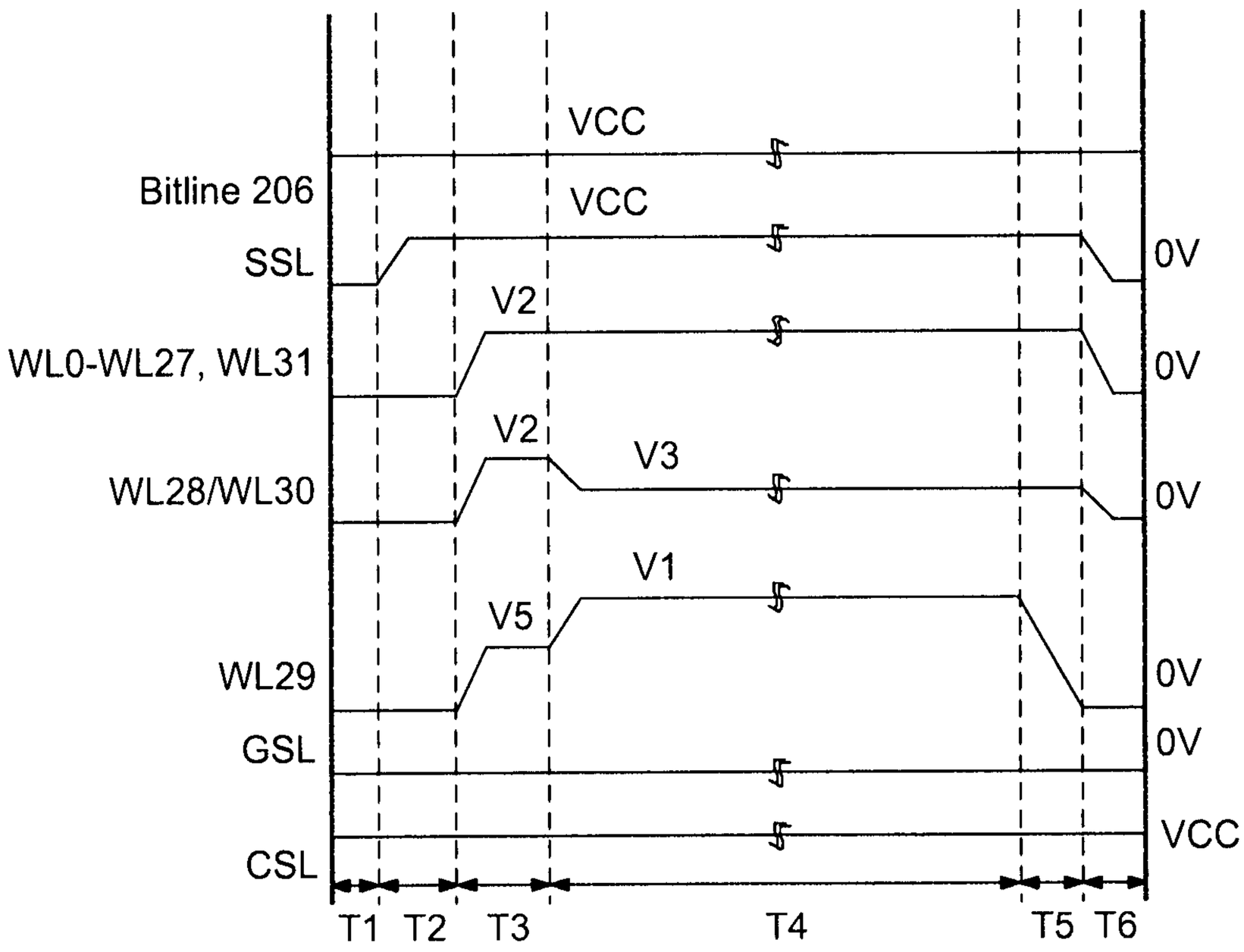

VCC
Bitline 206
VCC
SSL
0V
V2
WL0-WL27, WL31
0V
V2
V3
WL28/WL30
0V
V1
V5
WL29
0V
GSL
0V
VCC
CSL
T1
T2
T3
T4
T5
T6

Figure 11

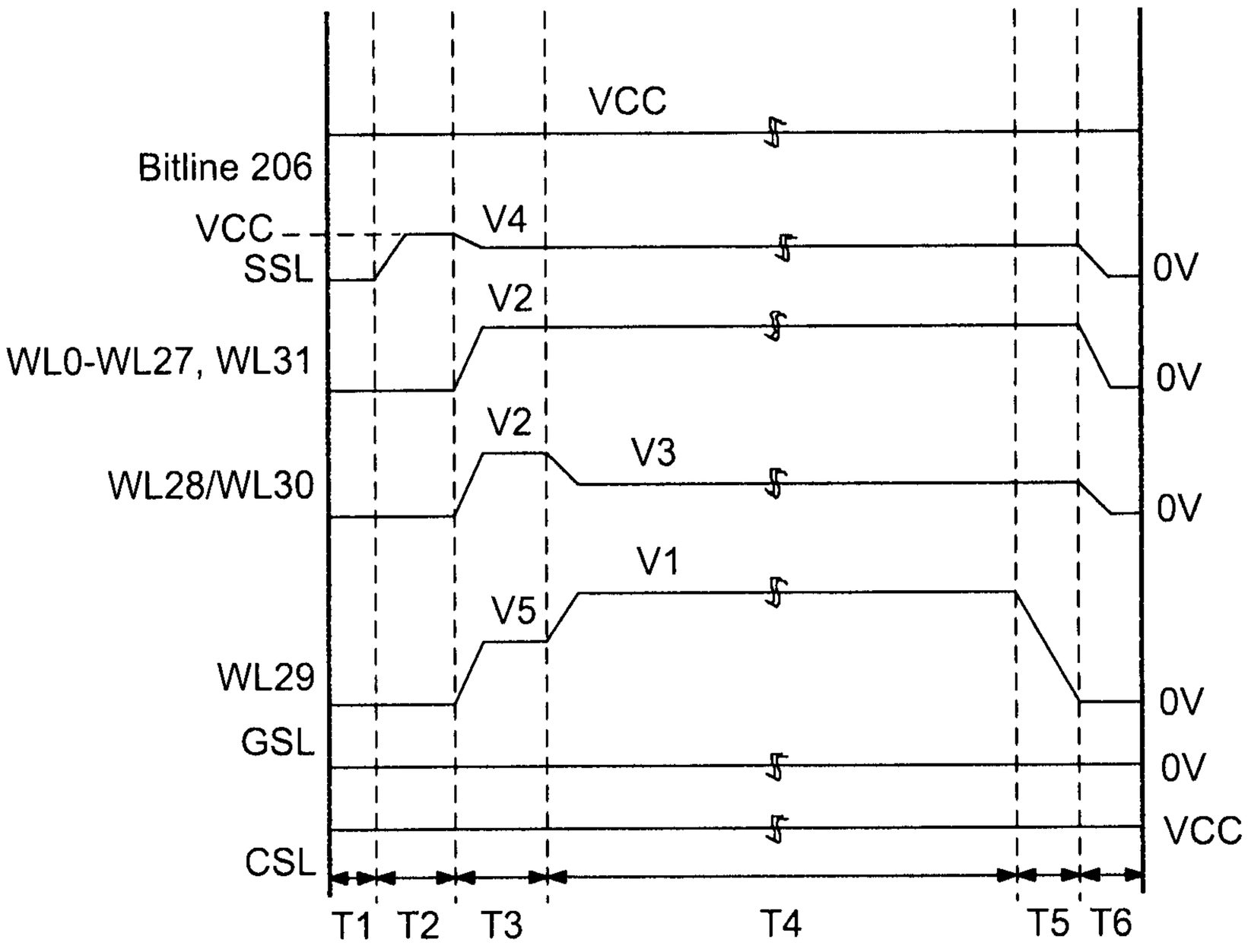

VCC
Bitline 206
VCC
SSL
V4
V2
WL0-WL27, WL31
V2
V3
WL28/WL30
V1
V5
WL29
GSL
CSL
0V
0V
0V
0V
0V
VCC
T1
T2
T3
T4
T5
T6

Figure 12

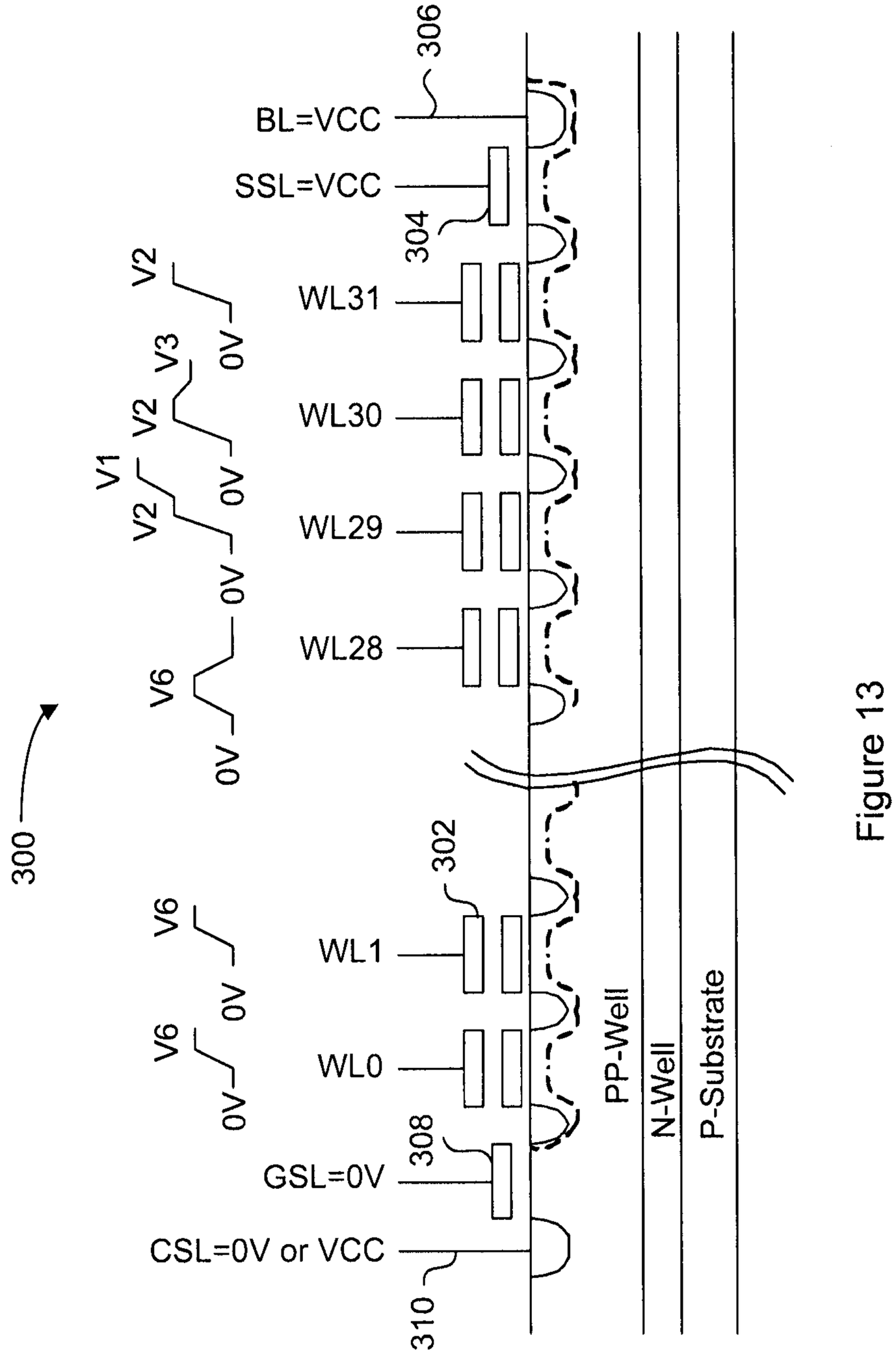

300
BL=VCC
SSL=VCC
WL31
WL30
WL29
WL28
WL1
WL0
GSL=0V
CSL=0V or VCC
V1
V2
V3
V6
0V
304
308
310
302
PP-Well
N-Well
P-Substrate

Figure 13

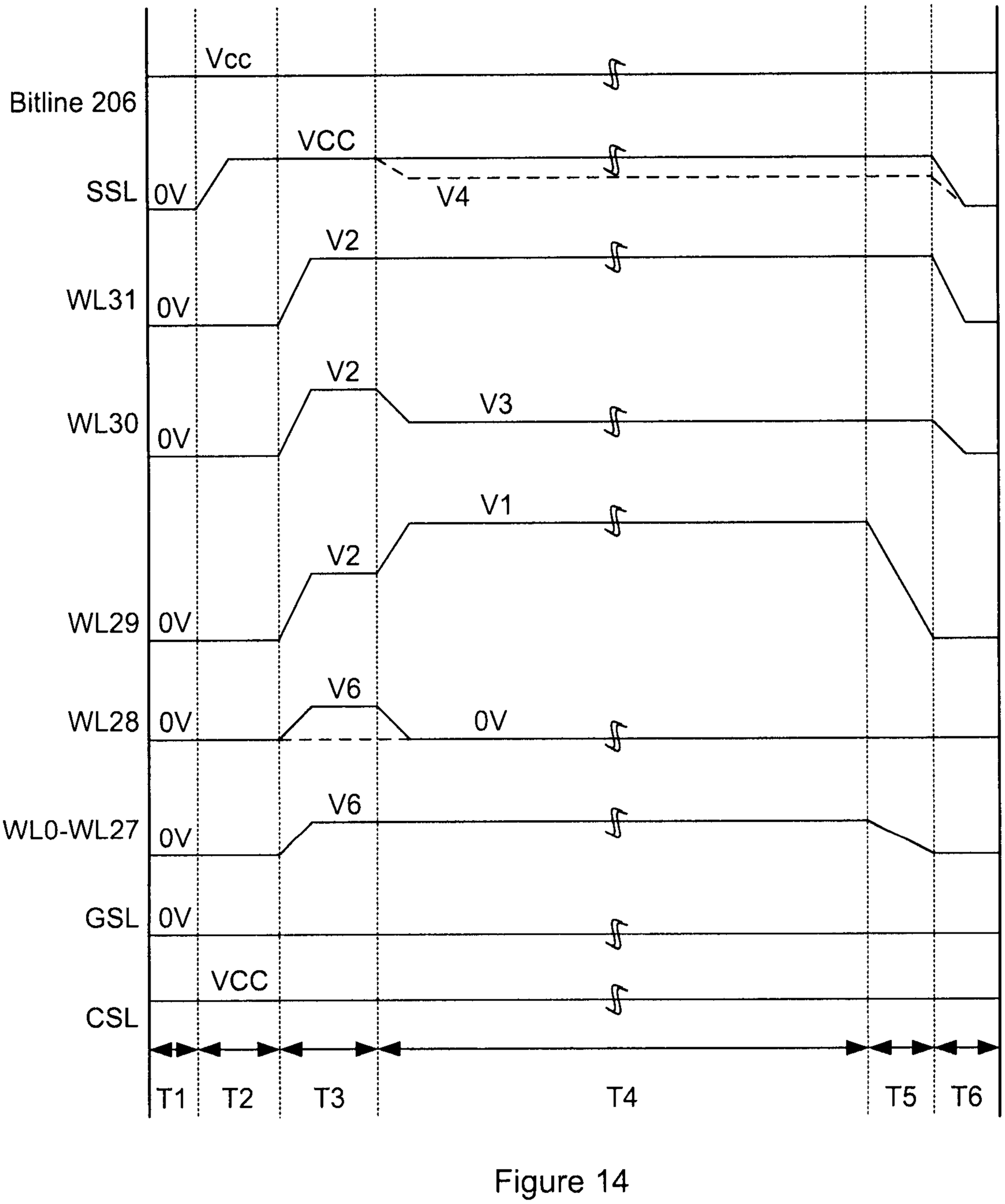

Vcc
Bitline 206
VCC
SSL
0V
V4
V2
WL31
0V
V2
V3
WL30
0V
V1
V2
WL29
0V
V6
WL28
0V
0V
V6
WL0-WL27
0V
GSL
0V
VCC
CSL
T1
T2
T3
T4
T5
T6

Figure 14

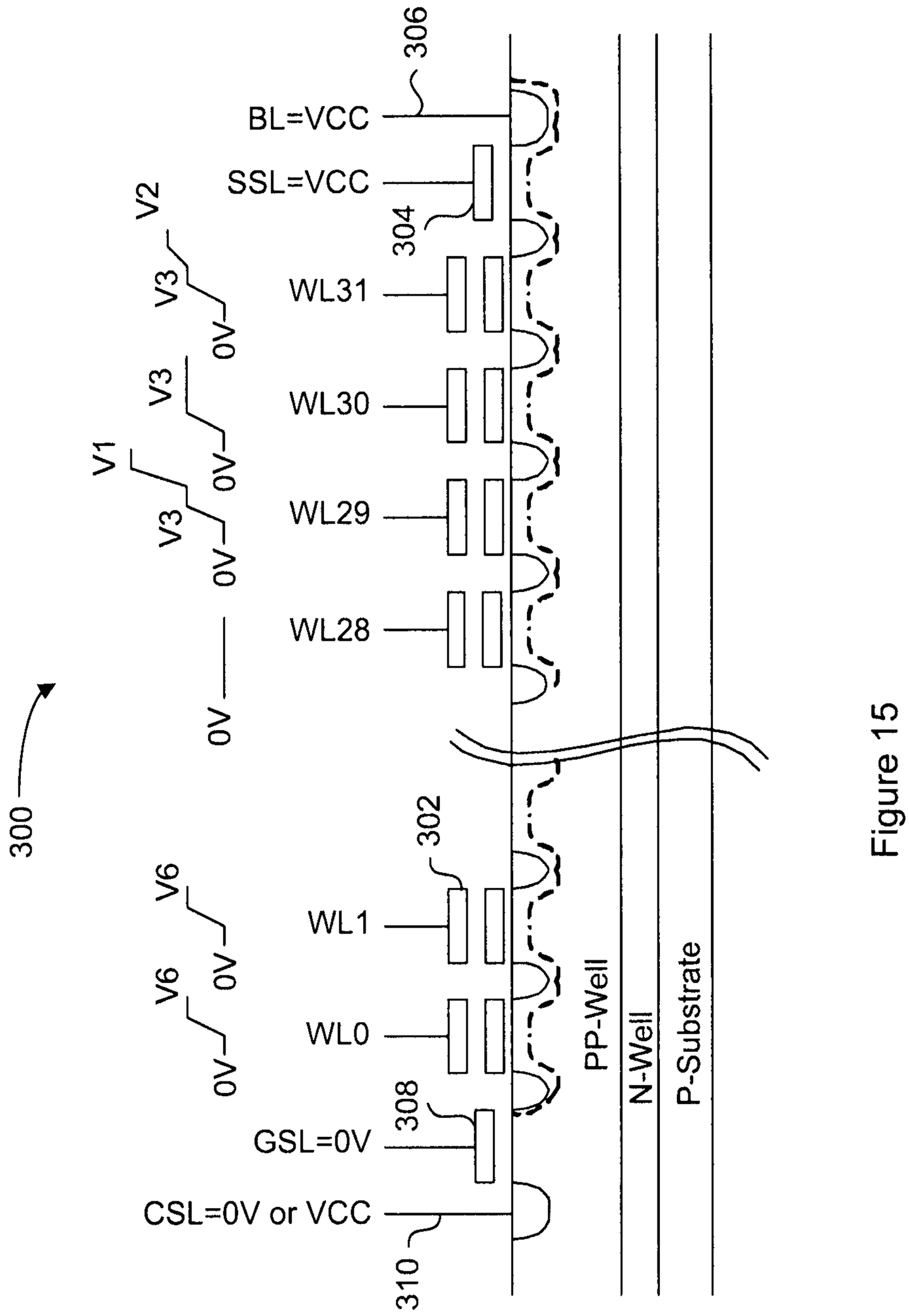

300
BL=VCC
SSL=VCC
WL31
WL30
WL29
WL28
WL1
WL0
GSL=0V
CSL=0V or VCC
V1
V2
V3
V6
0V
302
304
308
310
PP-Well
N-Well
P-Substrate

Figure 15

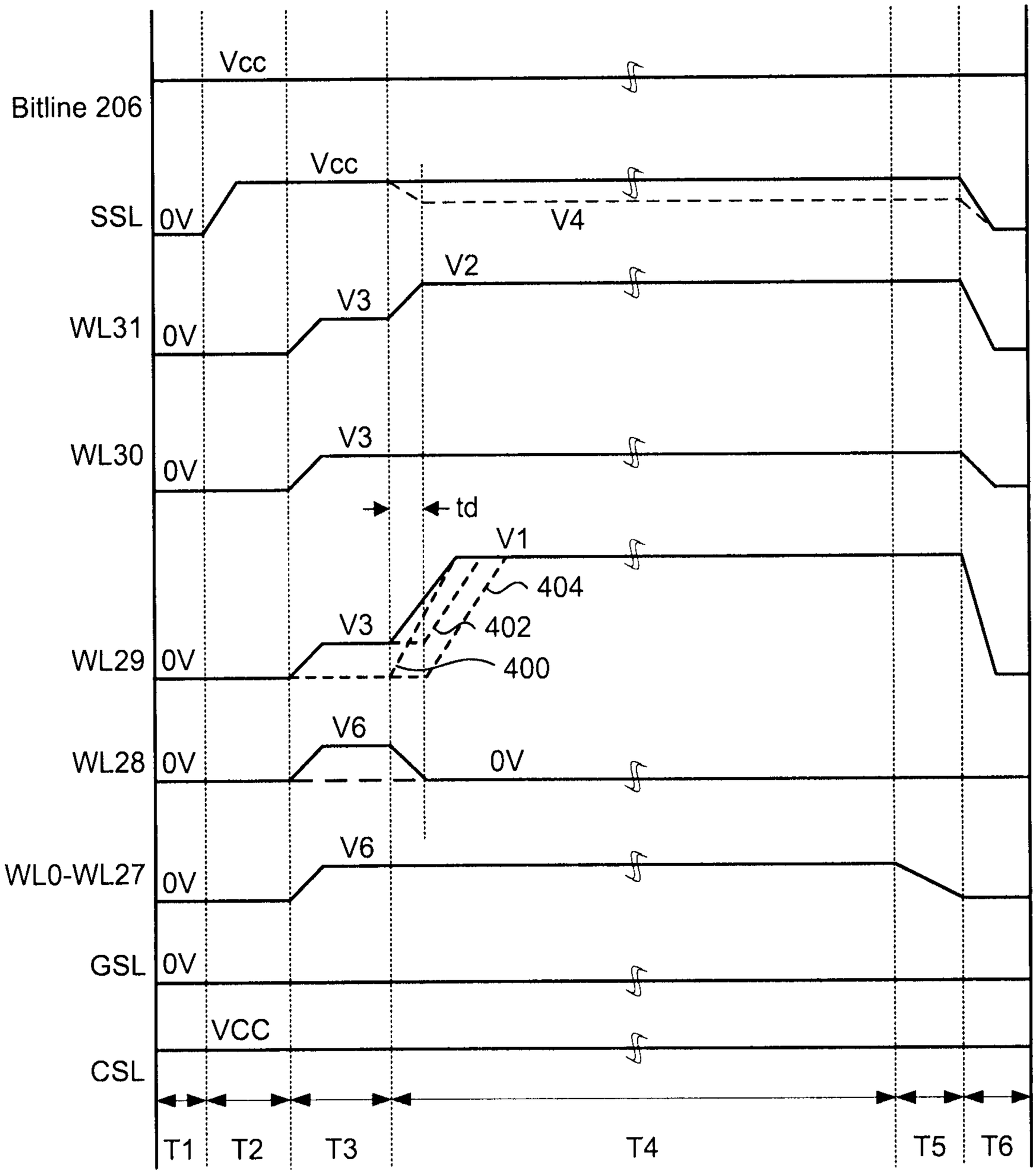
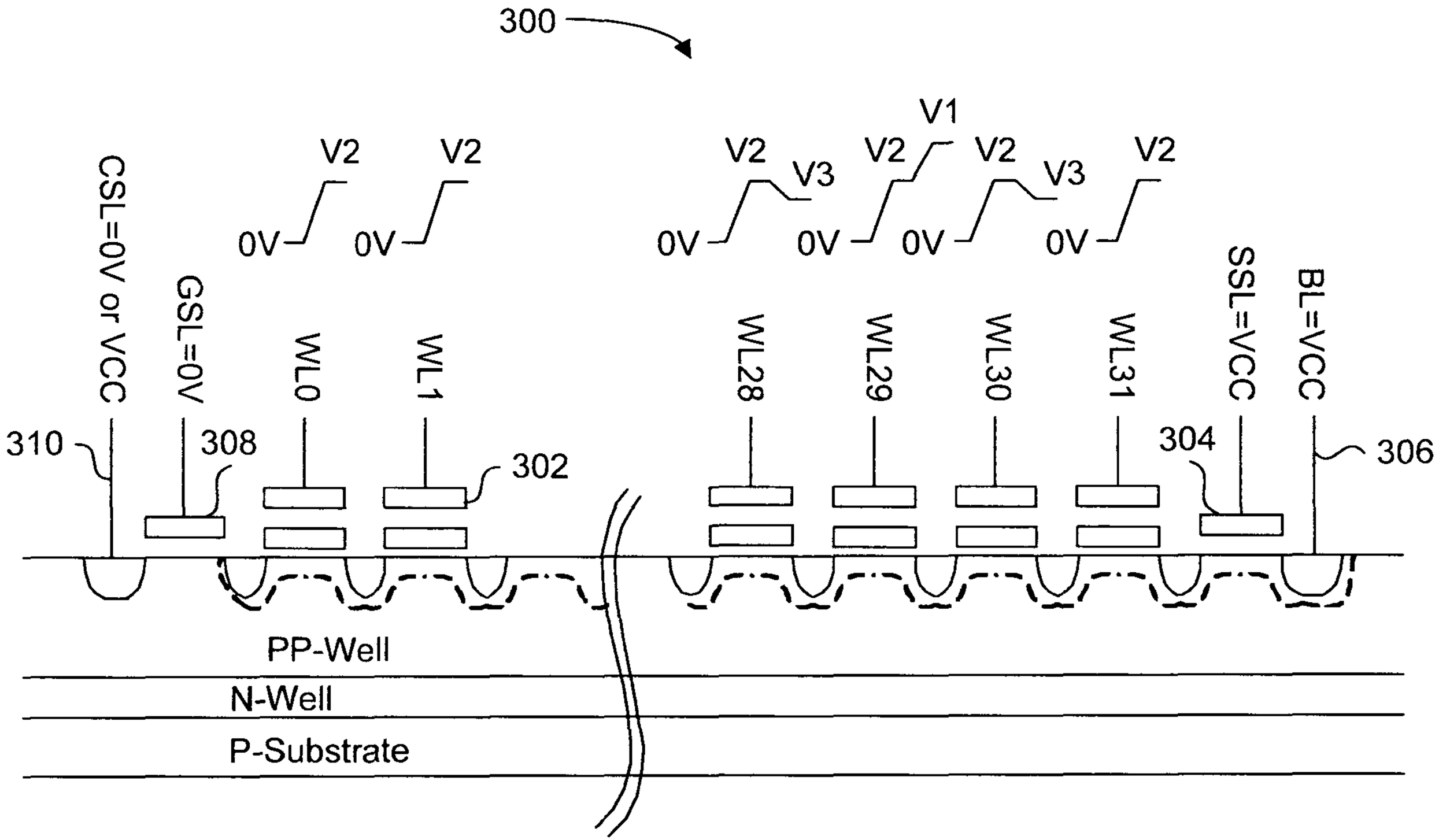

Bitline 206
Vcc
SSL
0V
Vcc
V4
WL31
0V
V3
V2
WL30
0V
V3
td
WL29
0V
V3
V1
404
402
400
WL28
0V
V6
0V
WL0-WL27
0V
V6
GSL
0V
CSL
VCC
T1
T2
T3
T4
T5
T6

Figure 16


300
CSL=0V or VCC
GSL=0V
WL0
WL1
WL28
WL29
WL30
WL31
SSL=VCC
BL=VCC
0V
V2
V3
V1
310
308
302
304
306
PP-Well
N-Well
P-Substrate