Ph.D. DISSERTATION

Layer Selection by Multi-level Permutation in Three-Dimensional Stacked NAND Flash Memory

3 차원 적층 낸드 플래시 메모리에서의 멀티 레벨 순열에 의한 층 선택 방법

BY

SANG-HO LEE

August 2018

DEPARTMENT OF ELECTRICAL AND
COMPUTER ENGINEERING
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Layer Selection by Multi-level Permutation in Three-Dimensional Stacked NAND Flash Memory

3차원 적층 낸드 플래시 메모리에서의 멀티 레벨 순열에 의한 층 선택 방법

지도교수 박 병 국

이 논문을 공학박사 학위논문으로 제출함

2018 년 8 월

서울대학교 대학원

전기컴퓨터공학부

이 상 호

이상호의 공학박사 학위논문을 인준함

2018 년 8 월

위 원 장 : 이 종 호 (인)

부위원장 : 박 병 국 (인)

위 원 : 송 윤 흡 (인)

위 원 : 조 성 재 (인)

위 원 : 윤 장 근 (인)

Abstract

Layer selection by multi-level permutation (LSMP) method for the layer decoding of 3D stacked NAND flash memory is proposed, proved mathematically, and verified by the TCAD simulations and the measurements of the fabricated 3D stacked NAND flash memory.

LSMP method provides the most efficient way to arrange the V_{th}'s of the SSTs for the layer selection as compared to VG-NAND, LASER, and LSM. By using this method, the number of selectable layers can be increased or the minimum required numbers of SSLs and V_{th} states can be reduced. In 3D stacked NAND flash memory architectures with layer selection methods by V_{th}'s of SSTs and SSL biases, as the number of SSLs increases, the number of WLs decreases in a limited area, which degrades the memory density. Since SSL often has much longer gate length to suppress leakage current, the array overhead (including SSL/GSL, BL pad, contact, and CSL area) can become significant. As the

number of V_{th} states of SSTs is increased, V_{th} margin is narrowed due to the limited memory window of SST and the layer selection may become unstable. Therefore, V_{th} arrangement of SSTs should be performed efficiently so that the maximum number of layers can be distinguished by the minimum number of SSLs and V_{th} states. How to arrange V_{th} states of SSTs for the unique selection is permuting grouped V_{th} multisets whose sums of V_{th} state numbers (l) are constant where available V_{th} states (V_i) are considered as $V_0 < V_1 < \ldots < V_{k-1}$ and i is state number. The SSL bias set for the layer selection can be determined by choosing the SSL bias voltage, V_{bi}, for the arranged V_{th} of SST, V_i, such that $V_i < V_{bi} < V_{i+1}$ ($0 \leq i \leq k$-1). The condition for the maximum number of the selectable layers is that l is equal to the integer which is closest to $n(k - 1)/2$. Furthermore, mathematical proofs for LSMP are made including the unique selection by the equal sum permutations and the recursion formula of the maximum number of selectable layers.

To verify the LSMP method in 3D stacked NAND flash memory, the TCAD simulation and the measurement of the fabricated device are performed.

The V_{th} states of SSTs are arranged by LSMP method. When the SSL bias set for layer selection is applied, the transfer characteristic and electron concentration of each stacked layer are monitored and each layer can be uniquely selectable.

Keywords: 3-D stacked NAND flash memory, NAND flash memory, layer selection by multi-level permutation, layer selection by multi-level operation, selection by thresholds.

Student number: 2012-20836

Table of Contents

Chapter 1
Introduction

1.1 Three-Dimensional Stacked NAND Flash Memory

In recent years, the demand for low cost and high density NAND flash memory has been increasing due to advances in mobile electronic devices such as smartphones, tablet PCs and laptop computers, which adopt NAND flash memory as their main data storage device. Fig. 1.1 shows the compound annual growth rate (CAGR) of the NAND flash application. As the scaling down of the NAND flash memory is accelerated, the short channel effect and

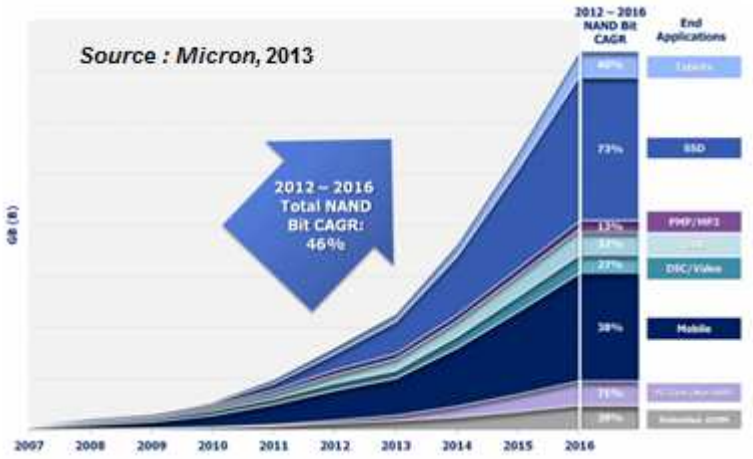


Fig. 1.1 Compound annual growth rate of the NAND flash applications.

reliability problems become more serious [1-5] and more advanced and high cost fabrication technologies are required for the further scaling down [6-8]. Fig. 1.2(a) shows the gap between the required photo lithography and NAND flash memory technology nodes. As shown in fig. 1.2(b), there are some concerns on whether it is possible to maintain the current scaling trend because of available lithography tools and their cost. Various photo lithography technologies and compensation technologies such as optical proximity correction (OPC), phase shift masking (PSM), immersion lithography, and double patterning technology (DPT) have been introduced to

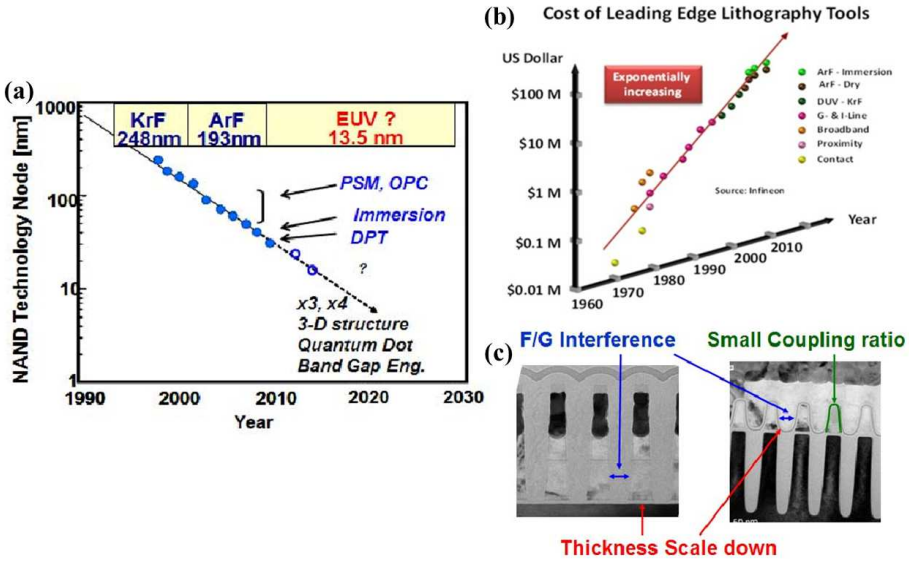


Fig. 1.2 (a) The gap between required photolithography and NAND flash memory technology nodes. (b) Cost of lithography tools (c) Major obstacles in scaling down NAND flash memory [10].

overcome the limitations of photo lithography tools [6-10]. However, beyond the 20 nm technology node, the gap will be wider. Therefore, we need new breakthroughs for the further reduction of the per-bit manufacturing cost. Fig. 1.2(c) shows the major obstacles to scaling down NAND flash memory. These physical limitations also block further scaling down [10-12]. As the feature

size becomes smaller, the volume of the floating gate decreases. And a small number of stored electrons causes the V_{th} shift of the programmed cell in small feature size structures. This implies that the loss of stored charge due to poor retention characteristics has a negative effect on the V_{th} shift of the cell [9, 10]. Moreover, the further scaling down causes short channel effects in planar NAND flash. As the gate length of the transistor decreases, the doping concentration of the substrate must be increased to maintain effective channel length and to prevent source / drain punch through. However, in NAND flash, the high doping concentration of the substrate degrades the boosted potential of the channel in the unselected bit lines under program-inhibit condition, thereby narrowing the V_{pass} bias operating window [10].

In spite of these problems, it is still required to maintain the current trends of memory capacity increase and bit cost reduction. To maintain the current trends, the scaling down of the device dimension has been in progress and multi-level operation of cells has been developed. Fig. 1.3 shows the

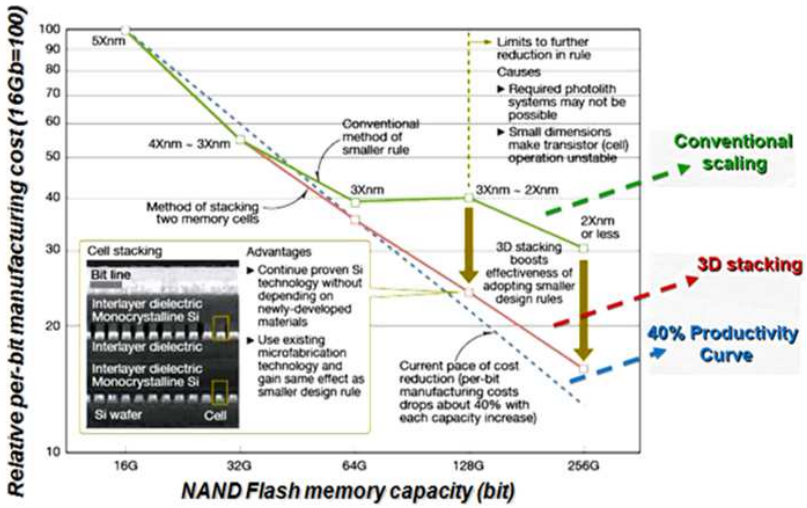


Fig. 1.3. Relative per-bit manufacturing cost as a function of NAND flash memory capacity [13].

relative per-bit manufacturing cost as a function of NAND flash memory capacity. The per-bit manufacturing cost is increasing faster than the shrink rate [13]. The scaling down of the device can achieve both per-bit manufacturing cost reduction and memory capacity increase. However, with 2D-based technology, it is difficult to maintain the current scaling down trend beyond the 2X nm node.

To overcome these limitations, various 3D stacked NAND flash memory

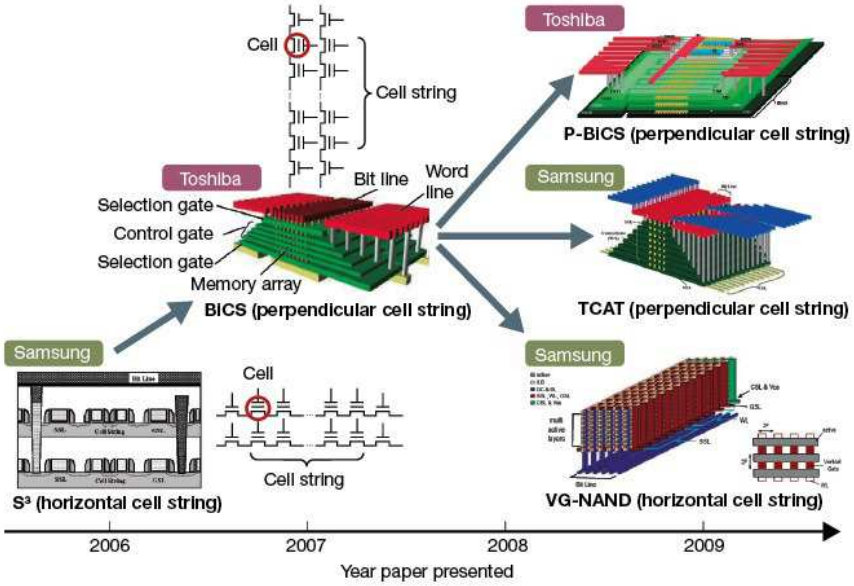


Fig. 1.4. History of 3D stacked NAND flash memory technology [39].

architectures have been proposed and reported [13-38]. Fig. 1.4 shows the history of 3D stacked NAND flash memory [39]. In 2006, the vertically stacked NAND flash memory technology was proposed, but it had a disadvantage that the process cost increased in proportion to the number of layers [13, 14].

In 2007, Toshiba proposed a Bit-Cost Scalable (BiCS) 3D cell technology [16]. The BiCS structure alternately stacks the dielectric films and the poly-Si electrodes, then punches channel holes through the entire stack

from top to bottom and fills the holes with ONO films, poly-Si, filler oxide, and so on. By using this punch and plug process, even if a large number of layers are stacked, the process cost does not increase much [16-19].

In 2009, Toshiba proposed a Pipe-shaped Bit Cost Scalable (P-BiCS) flash memory with improved BiCS technology [20]. P-BiCS structure consists of NAND strings folded like a U-shape, and a 32-Gbit NAND flash memory is implemented with 60-nm process technology using 16-layer stacking and multi-level-cell operation [20, 21].

Samsung Electronics also proposed Terabit Cell Array Transistor (TCAT) in 2009 [22]. Most of the issues of BiCS structure were solved in the TCAT structure. By the gate replacement process, the damascened metal gate SONOS type cell is implemented in vertical NAND flash string. And conventional bulk erase operation can be achieved because the channel poly plug is connected to Si substrate.

1.2 Gate Stack Type NAND Flash Memory

Several 3D stacked NAND flash memory architectures have been proposed [15-38]. The 3D stacked NAND flash memory architectures can be divided into two categories: gate stack type [15-26], in which channel current flows along a vertically formed channel and gates are shared horizontally by all strings, and channel stack type [27-38], in which the current flows along the horizontally stacked channel and the gates are vertically shared by all the strings. SONOS and MANOS devices using charge trapping layer as data storage are preferred as 3D stacked NAND flash memory architectures because of their easy process integration and small cell size. Also, floating gate type 3D NAND flash has been proposed [25, 26].

Fig. 1.5 shows the schematic diagram and representative structure of the gate stack type NAND flash memory [20, 22]. The current flows through a

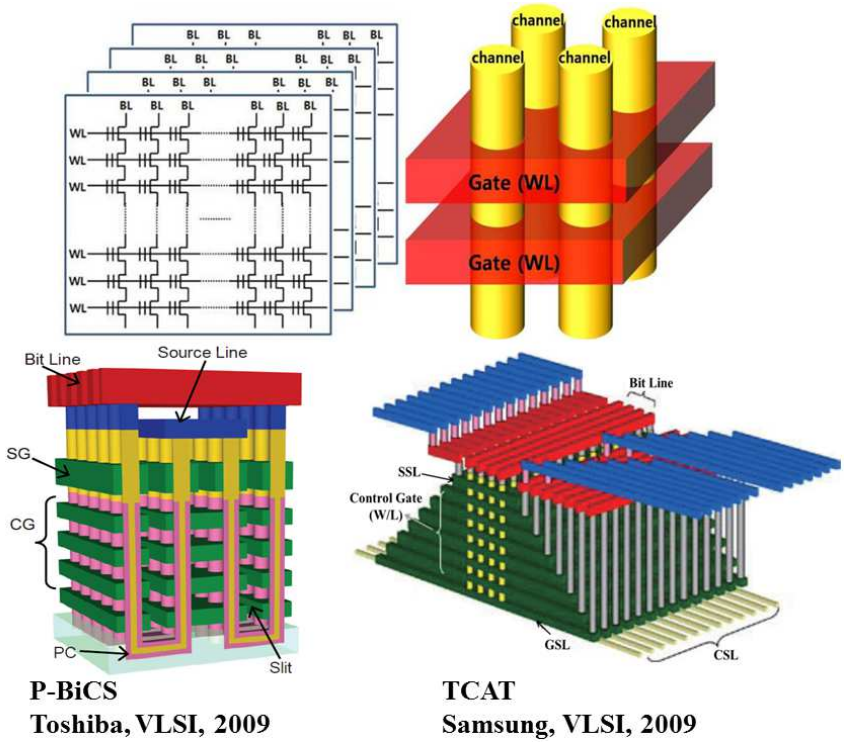


Fig. 1.5. Schematic diagram and representative architectures of gate stack type NAND flash memory [20, 22].

vertically formed nanowire poly-Si channel and the cell structure is a gate-all-around structure. The memory capacity can be increased by adding more control gate layer stacks, while the number of critical lithography steps remains constant because whole stack of control gates and dielectric films is punched with one lithography step.

After the proposal of the BiCS architecture in 2007, several advanced gate stack type 3D NAND flash memory architectures have been introduced. As shown in Fig. 1.5, Pipe-shaped BiCS (P-BiCS) of Toshiba and TCAT of Samsung are two representative structures.

BiCS flash memory array is shown in fig. 1.6 [16]. A string of cells is on the plugs located vertically in the holes punched through the whole stack of gate plates. Each plate acts as a control gate except for the top and bottom plates, which take the roles of the upper and lower select gates (USG/LSG). The cell size of BiCS memory is $6F^2$ per layer. A single cell is accessed by the control gate on the string, which is selected by a BL and a select gate (SG). The bottom of the memory plug is connected to source diffusion formed on the silicon substrate. For the erase operation of BiCS flash memory, the hole current generated by gate-induced drain leakage (GIDL) near the LSG is used. However, BiCS flash memory has several critical issues to be resolved as follows [20, 22]; (1) SiN-based tunnel film is implemented to minimize the

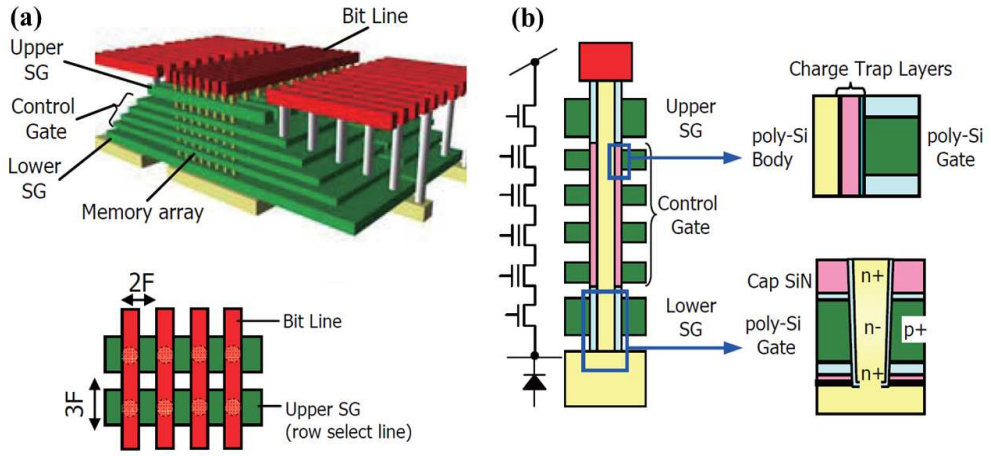


Fig. 1.6. BiCS NAND flash memory by Toshiba. (a) Bird's eye view, top down view and (b) schematic diagram of one string [16].

damage by diluted-HF treatment for the "gate-first" process; however, the program/erase (P/E) window which is vulnerable to read disturb and the data retention is insufficient to operate it as a multi-level cell (MLC). (2) The high-resistance source line is hardly clamped on its initial voltage owing to the large total current from a huge number of NAND strings during a read operation. (3) The LSG transistor is placed upon the heavily doped source diffusion and suffers from many thermal process steps for the memory-film formation; therefore, the diffusion profile is not controllable. (4) It is almost

impossible to implement metal gate structure for BiCS flash because it is very difficult to etch metal/oxide multilayer simultaneously. We cannot utilize various advantages of metal gate SONOS cell structure, for example, faster erase speed, wider V_{th} margin, and better retention characteristics [40]. (5) Another concern is GIDL erase of BiCS flash. An extensive circuit change may be necessary to apply negative bias on word line (WL) during erase operation. Area penalty and limited erase voltage are expected.

To overcome these issues, P-BiCS flash memory was proposed as shown in fig. 1.7 [20]. Two adjacent NAND strings are connected at the bottoms by a so-called pipe-connection (PC), which is gated by the bottom electrode. One of the terminals for the U-shaped pipe is connected to the BL, and the other is bound by the source line (SL). The SL consists of the meshed wiring of the third-level metal and is accessed by the first- and second-level metal as in conventional planar technology; therefore, the resistance of the SL is sufficiently low. Both the SG transistors are placed below the SL and the BL.

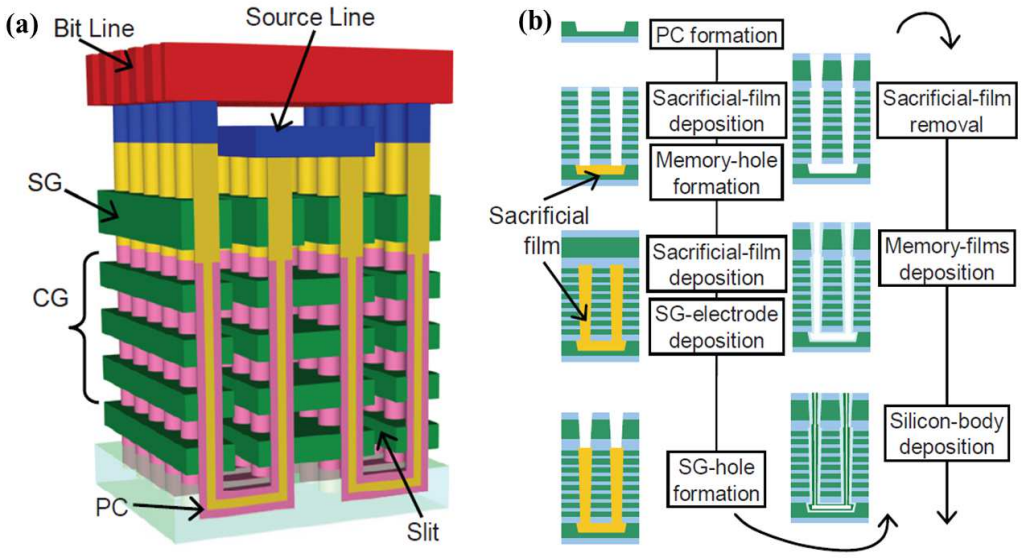


Fig. 1.7. P-BiCS NAND flash memory by Toshiba. (a) Bird's eye view and (b) process flow [20].

The control gates (CGs) are isolated by the slit and face each other in a couple-of-combs pattern. The memory stack consists of a blocking layer, a charge trap layer, and an oxide-based layer as a tunnel film. The both of the P/E window and speed are dramatically improved by the P-BiCS memory films.

Samsung Electronics proposed vertical NAND flash memory cell array using TCAT technology in 2009 [22]. Distinctive structural differences of

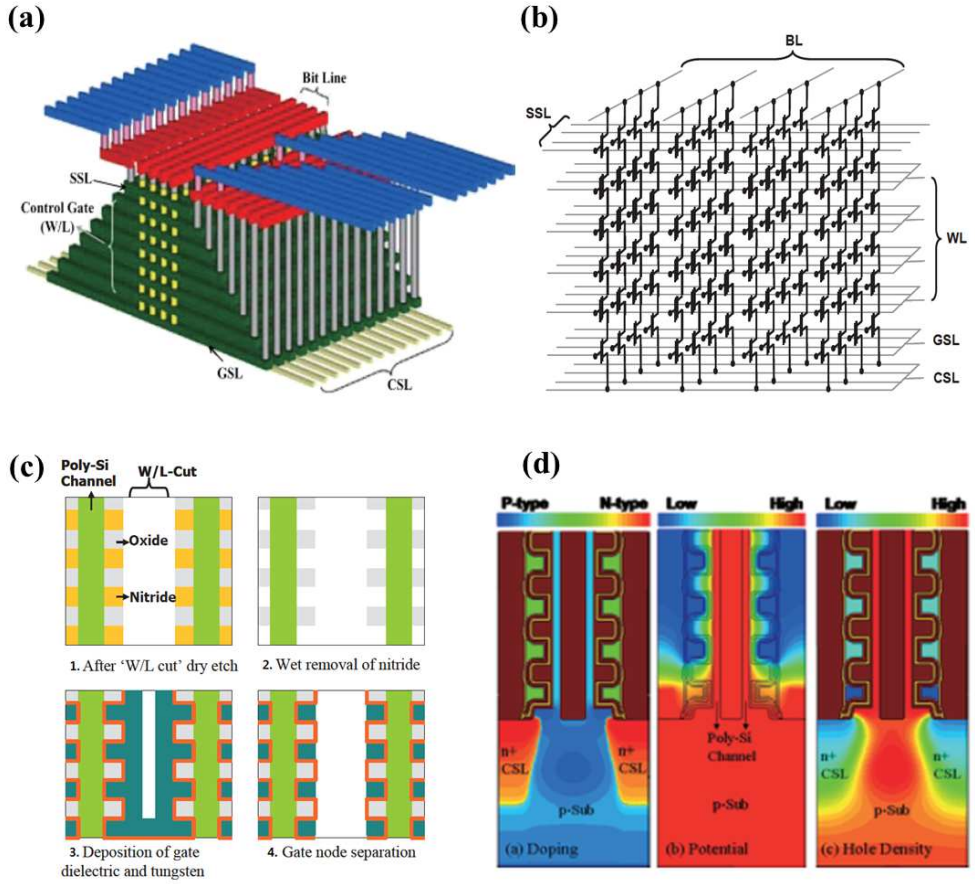


Fig. 1.8. TCAT flash memory by Samsung Electronics. (a) Bird's eye view (b) Equivalent circuit (c) Gate replacement process flow (d) Simulated profiles of doping, potential, and hole density during bulk-erase operation of TCAT flash [22].

TCAT with BiCS are oxide/nitride multilayer stack, line-type 'W/L cut' etched through the whole stack between each row array of channel poly plug, line-type common source line (CSL) formed by an implant through the 'word line

(WL) cut', and replaced metal gate lines for each row of poly plug. The most unique process is 'gate replacement' to achieve the MONOS structure. Fig. 1.8(c) shows a detailed view. After the WL cut dry etch and wet removal of the sacrificial nitride layer, gate dielectric layers and gate metal are deposited in the conventional order. It is not 'gate-first' process as for BiCS. The cell string has NAND cell transistors with a string select line (SSL) transistor at the top and the ground select line (GSL) transistor at the bottom. The channel poly plug in TCAT flash structure is connected to Si substrate. Therefore, conventional bulk erase operation can be achieved as displayed in the simulated profiles of Fig. 1.8(d) and the NAND string can be operated without any major peripheral circuit change.

1.3 Channel Stack Type NAND Flash Memory

Several channel stack type NAND flash memory architectures have been proposed [27-38]. Fig. 1.9 shows a schematic diagram and a representative structure of channel stack type NAND flash memory [27, 35]. The channel current flows through a horizontally formed Si nanowire channel and the cell structure is a double gate or gate-all-around structure. The memory capacity can be increased by adding the number of the stacked Si channel layers, while the number of critical lithography steps is constant since the entire stacked channel layer can be dry etched with one lithography step. As shown in fig. 1.9, Vertical Gate (VG) NAND architecture and Channel STacked ARray (CSTAR) architecture are representative structures. Channel stacked NAND flash memory has the advantages of good pitch scalability and small cell size. However, decoding innovations are required because it is more difficult to

decode BLs than gate stack type.

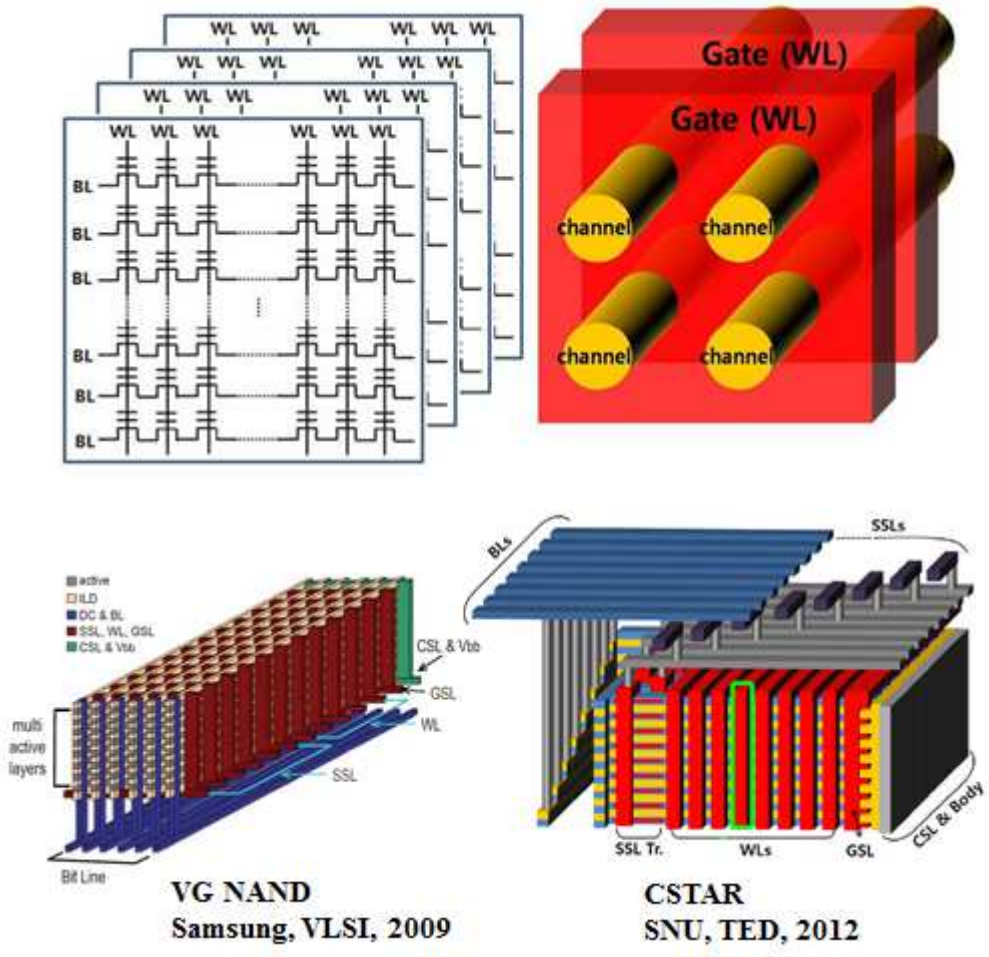


Fig. 1.9. Schematic diagram and representative architectures of channel stack type NAND flash memory [27, 35]

Samsung Electronics introduced VG-NAND in 2009 [27]. Fig. 1.10(a) and (b) show a bird's eye view of the VG-NAND flash memory and its equivalent circuit. VG-NAND flash memory includes WLs, BLs, CSLs, horizontal active strings with a pattern, VGs (for SSLs, WLs, and GSLs), charge trap layers between the active and the vertical gate, vertical plugs of DC, and source and active body (V_{bb}). WLs and BLs are formed at the beginning of fabrication before the cell array is constructed, making interconnect between WLs and BLs and the decoder easier. The source and active body (V_{bb}) are electrically tied to the CSL for enabling the body erase operation. A positive bias is applied to the CSL during erase operation. The array schematic of each layer is identical to that of the planar NAND flash memory except for the SSL. VG-NAND requires 6 SSLs for 8 active layers and 8 SSLs for 16 active layers. The reason for the multiple SSLs is to enable selection of data from a chosen layer out of multilayers since VG-NAND uses a common BL and a common WL between multi-active layers. The cell size of

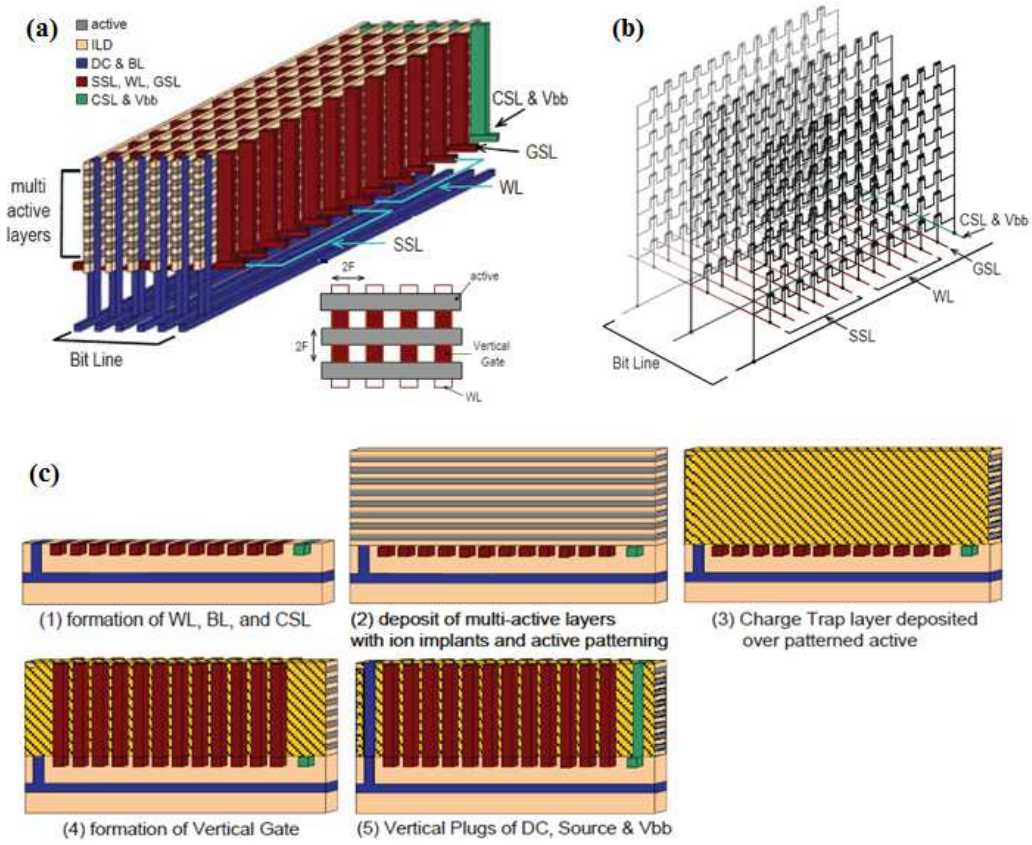
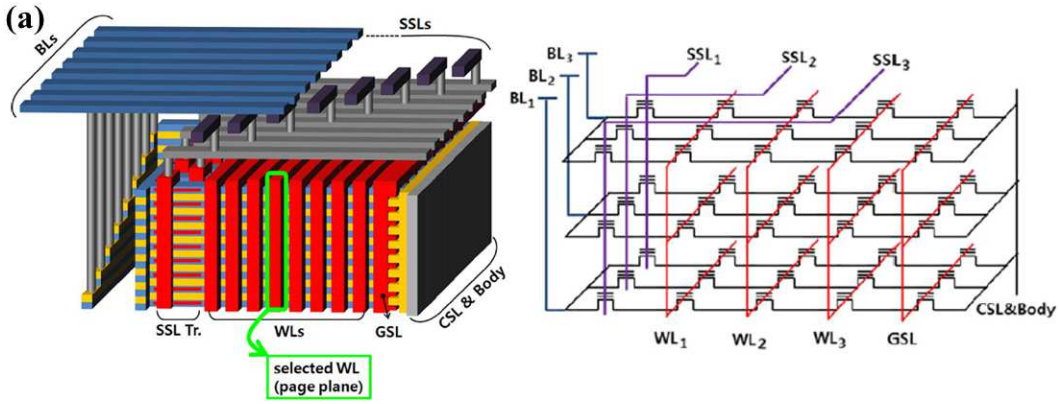


Fig. 1.10. Multiple layered VG-NAND. (a) Bird's eye view, (b) equivalent circuit and (c) process flow [27].

the VG-NAND is $4F^2$ per layer, as shown in Fig. 1.10(a). Fig. 1.10(c) explains integration sequence of VG-NAND and the integration scheme is based on simple patterning and plugging. A BL having n^+ poly-Si is fabricated first and then an n^+ poly-Si WL is formed on top of it. Multiple active layers with p-type poly-Si are formed with n-type ion implants for SSL layer selection and

alternating interlayer dielectrics are inserted between active layers. Then patterning is carried on the multi-active layers and charge trap layers (ONO) are deposited over the patterned actives. Consecutively VG is formed and connected to the WL.

CSTAR architecture was introduced in 2012 [35]. Fig. 1.11 shows the proposed CSTAR structure for the 3D NAND flash architecture. This unit structure is defined as “building.” The specific descriptions of each part of the building are as follows. The 3D stacked NAND flash memory, unlike conventional two-dimensional (2D) planar NAND flash memory, needs one more address that is determined by the SSLs. In conventional 2D NAND, the WL selects a “page (line),” whereas the WL of the 3D stacked NAND selects a “page plane.” Then, the column and row in the selected page plane are determined by the BL and SSL, respectively. The BLs are formed on the top floor of a building. In addition, BLs are perpendicular to other lines (WLs and SSLs). In a full array, all BLs of each block are connected with those of other



(b)

	PROGRAM	READ	ERASE
selected WL	Vpgm	0 V	0 V
unselected WL	Vpass	Vread	float
selected SSL	Vcc	Vread	float
unselected SSL	0 V	0 V	float
selected BL	0 V	Vcc	float
unselected BL	Vcc	0 V	float
GSL	0 V	Vcc	float
CSL&Body	0 V	0 V	Vers

Fig. 1.11. Channel stacked NAND flash memory. (a) Bird's eye view of the unit building structure and equivalent circuit. (b) Operation scheme [35].

blocks to the page buffer. Consequently, there must not be any overlapping with other metal lines of contact holes in the BL region. The WLs and SSLs are parallel with different levels, and they will be connected with each decoder at the opposite end site.

The key features of CSTAR are that it uses a single-crystalline silicon

channel and a gate-all-around structure using the Si/SiGe epitaxial growth process [33-35]. VG-NAND has a poly-Si channel and a double gate structure. Thus, CSTAR has the merit of more cell current drivability, short channel effect immunity, and more reliable memory operation.

Channel stack type 3D NAND flash memory with layer selection by V_{th}'s of SSTs and SSL biases such as VG NAND, LASER, and LSM [27, 37, 38] holds many advantages over the conventional CSTAR array. Fig. 1.12 shows a bird's eye view and a top view (layout) of the conventional CSTAR array and the CSTAR with LSM. In the case of the conventional CSTAR array, SSLs must be of island type for channel separation in a selected array, because the BL is common in a layer. So, minimum BL pitch is constrained by the lithography and dry etch margin to form the island-type SSLs. Moreover, according to the operation scheme, the SSL bias induces interference between the neighboring channel and the SSL. Additionally, the conventional CSTAR has complex wiring. Too many contacts are needed for the island-type SSL

and the number of metal layers is increased to form different levels of WLs and SSLs.

However, in the case of CSTAR with layer selection by V_{th}'s of SSTs and SSL biases, the BL pitch can be scaled down compared with conventional CSTAR, because it enables line type SSLs. Scaling down the BL pitch offers an area advantage of increased cell density. The CSTAR array is also free of interference between neighboring channels and SSLs owing to the shielding of the gate-all-around structure. In addition, it is not necessary to use different levels to form the WLs and SSLs, so wiring complexity can be reduced.

It is more difficult to decode a stacked layer and cell in a conventional channel stacked NAND flash memory. This decoding difficulty leads to various problems such as wiring complexity, limitation of BL pitch scaling, low BL throughput, and restriction on building expandability [38]. However, channel stacked NAND flash memory with layer selection by V_{th}'s of SSTs

and SSL biases has no island-type. Channels are formed in the horizontal direction, and SSLs are formed in the other horizontal direction to be parallel to WLs like a planar type NAND flash array. Layer selection can be performed by the electrical initialization and the SSL bias. By virtue of the above mentioned features, this layer selection method solves the problems of conventional CSTAR.

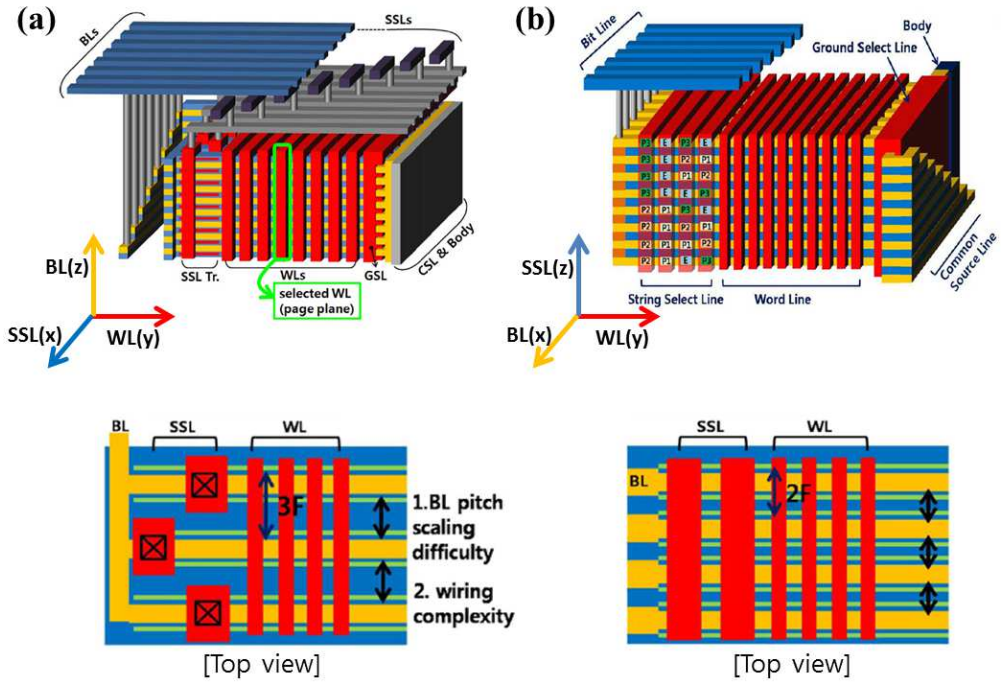
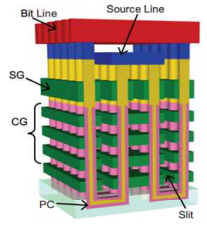
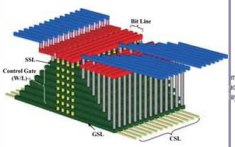
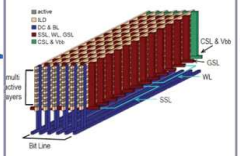
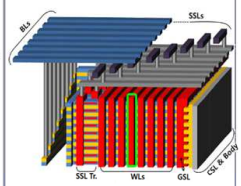


Fig. 1.13. Bird's eye view and top view (layout) of (a) conventional CSTAR array and (b) CSTAR with LSM [38]

1.4 Comparison between Gate Stack Type and Channel Stack Type NAND Flash Memory

Table 1.1 shows various 3D stacked NAND flash memory architectures [20, 22, 27, 35]. In this section, we discuss the advantages and disadvantages of gate stack type NAND flash memory and channel stack type NAND flash memory.

Table 1.1 Various 3D NAND flash memory architecture [20, 22, 27, 35].

	Gate Stack Type		Channel Stack Type	
	P-BiCS	**TCAT**	**VG NAND**	**CSTAR**
Structure				
Cell Structure	Macaroni & GAA Floating body	Macaroni & GAA	Double Gate	Gate-All-Around
Channel	Poly-Si channel	Poly-Si channel	Poly-Si channel	Single crystalline Si channel
Key Issue	▪ Low read current ▪ Reliability ▪ Pitch scaling	▪ Low read current ▪ Reliability ▪ Pitch scaling	▪ Number of litho. and implantation ▪ a lot of SSLs	▪ Layer decoding ▪ Wiring complexity ▪ Layer separation

The gate stack type NAND flash memory uses a poly-Si channel. The poly-Si channel has lower electron mobility and worse current drivability than single-crystalline Si channel due to grain boundary traps. Thus, gate-all-around and macaroni poly-Si channels were applied to reduce the effect of grain boundary trap [17, 19]. In the gate stack type, the bit density can be increased by stacking more control gate layers. However, as the number of stacked gate layers increases, the overall resistance of the channel string becomes higher and the read current decreases. As the diameter of the channel hole decreases, the channel width of the cell device decreases and the read current decreases. Due to the low current drivability of the poly-Si channel, the number of stacked control gate layers and cell size scaling down can be limited. The cell size is also limited by the minimal ONO thickness the macaroni channel thickness, and lithography constraints such as the overlay between the drill hole and SSL, and the minimal SSL/WL cut space [41]. The use of poly-Si channels also has instability problems due to grain boundary

traps [42-45]. Memory operation may become unstable due to slow charge trapping / detrapping rate at the grain boundaries as the gate bias changes.

The channel stack type NAND flash memory architectures are free from these problems. Since the length of the channel string is kept constant irrespective of the number of stacks, the read current does not deteriorate even if the number of stacks is increased. In the channel stack type, a single-crystalline Si channel can be used using the Si / SiGe epitaxial growth process, thus, it has no instability problems due to grain boundary trap [33-38]. For channel stack type, the limiting factor is the finite ONO thickness that restricts the BL pitch [41]. So, in general, the channel stack type has a smaller minimum cell size than the gate stack type. However, it is more difficult to decode stacked layers in channel stacked NAND flash memory architectures.

Chapter 2

Layer Selection Methods by V_{th}'s of SSTs and SSL biases

2.1 Concept of Layer Selection

In the channel stack type NAND flash memory with layer selection by V_{th}'s of SSTs and SSL biases such as VG-NAND, LASER, and LSM [27, 37, 38], BL is common for all layers and is separated in a selected layer as an example of Fig. 2.1. Fig.2.1 shows the bird's eye view and the equivalent circuit of layer selection by erase (LASER) as an example. SSLs are formed parallel to the WLs as in a planar type NAND flash array. When a layer is

selected by SSTs and SSL biases, the array schematic of the selected layer is identical to that of the planar NAND flash memory except for the SSL. CSL and body can be formed differently in each structure according to V_{th} initialization method of SST.

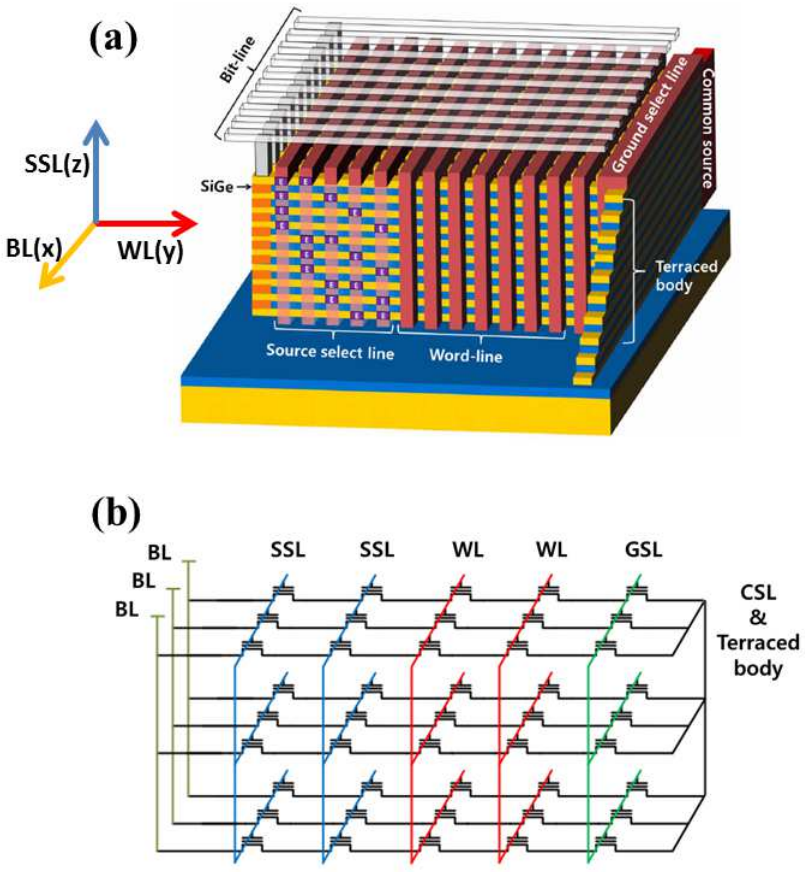


Fig.2.1. (a) Bird's eye view and (b) equivalent circuit of LASER [37].

Basic description of layer selection method by V_{th}'s of SSTs and SSL biases is shown in Fig.2.2. When SSL biases are applied to SSTs whose V_{th}'s are properly arranged, it is determined whether or not the channels of each layer are connected to the BLs by the ON and OFF states of the SSTs of each layer. The V_{th} arrangement of SSTs in a layer must be all different from the other layers. As shown in Fig. 2.2(a), when 2 V and 0 V are applied to the 1st SSL and the 2nd SSL, respectively, all string select transistors whose V_{th}'s are 1 V and -1 V of the 1st layer are turned on and each channel of the 1st layer is connected to each BL, while the other unselected layers (the 2nd layer in this example) are electrically disconnected to BLs, because the other layers have at least one turned-off SST(s). In Fig. 2.2(a), the 2nd SST of the 2nd layer is turned off because the threshold voltage of SST is lower than the applied SSL bias.

On the other hand, when 0 V and 2 V are applied to the 1st SSL and the 2nd SSL, respectively, all string select transistors whose V_{th}'s are -1 V and 1 V

of the 2nd layer are turned on and each channel of the 2nd layer is connected to each BL, while the other unselected layers (the 1st layer in this example) are electrically disconnected to BLs, because the other layers have at least one turned-off SST(s) as shown in Fig. 2.2(b).

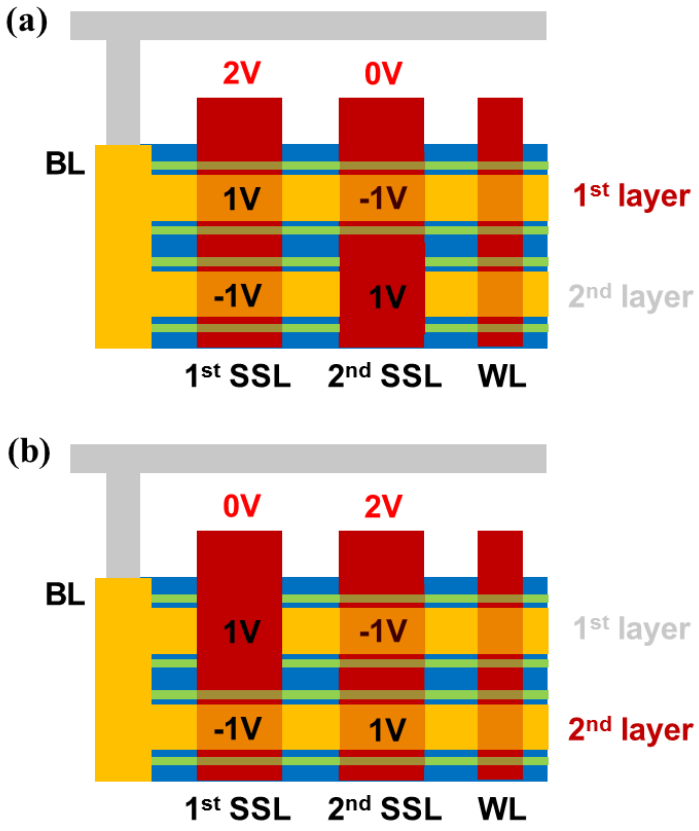


Fig. 2.2. Basic description of layer selection method by V_{th}'s of SSTs and SSL biases. (a) 1st layer selection (b) 2nd layer selection

2.2 Conventional Layer Selection Methods

Vertical Gate NAND (VG-NAND), LAyer Selection by ERase (LASER), and Layer Selection by Multi-level operation (LSM) have been proposed as layer selection methods by V_{th}'s of SSTs and SSL biases for the channel stack type NAND flash memory. The V_{th} arrangement strategy is different for each method, and the body and CSL are formed differently according to the V_{th} initialization method.

2.2.1. Vertical Gate NAND (VG-NAND)

In the VG-NAND [27] structure as shown in Fig.2.3(a), multi-active layers are deposited and then SSTs are formed as depletion transistors or enhancement transistors by n-type ion implants. A transistor with shade

(depletion transistor) always turns on regardless of applied voltage on SSL while one without shade (enhancement transistor) only turns on under proper voltage application. The operation table of SSL for a layer selection and the schematics of SSL including depletion and enhancement transistors are shown in Fig.2.3(b). The number of selectable layers S is $2^{n/2}$, where n is the number of SSLs. The V_{th} arrangement strategy of VG-NAND is shown in Fig.2.3(c). The number in each cell where a SSL and a layer intersect means the V_{th} of SST. V_{th}'s of the 1st SSL are arranged in ascending order and V_{th}'s of the 2nd SSL are arranged in descending order. And this V_{th} arrangement can be expanded for more SSLs. Due to the pair SSL operation, the number of SSLs should be an even number.

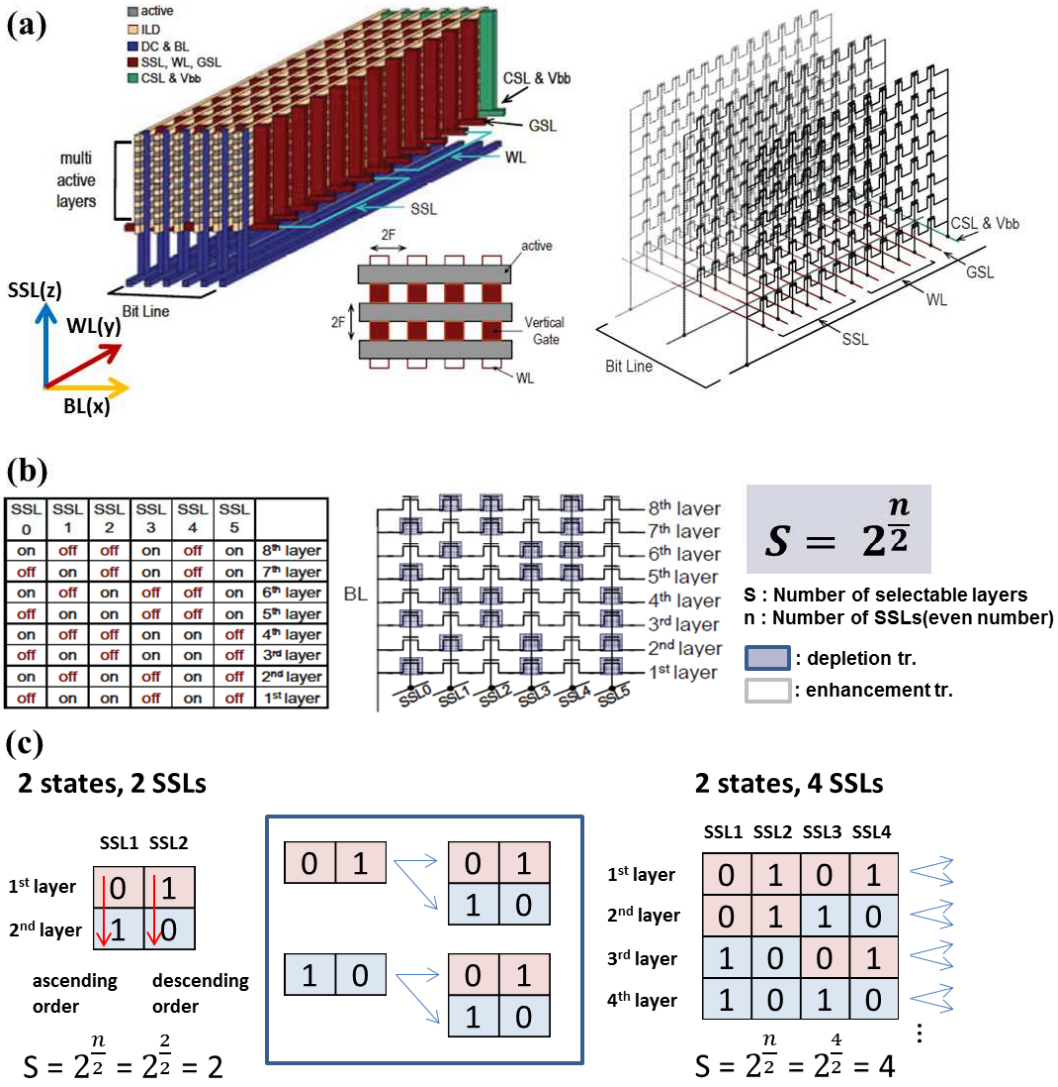


Fig. 2.3. (a) Bird's eye view and equivalent circuit of VG-NAND (b) The operation table of SSL for a layer selection and the schematics of SSL (c) V_{th} arrangement strategy of VG-NAND [27].

2.2.2. LAyer Selection by ERase (LASER)

In the LASER [37], to distinguish the stacked layers, some SSTs in a layer are electrically initialized (erased) by the terraced body as shown in Fig2.4(a). For the initial erase operation, bodies of each stacked layer should be contacted. However, with conventional architecture, which has a common source line (CSL) next to a ground select line (GSL), there is no space to contact the body for the erase operation. In the LASER architecture, terraced body regions are directly contacted for the layer-by-layer erase operation. On the other hand, the CSL region is required for the read operation. The GSL covers over the stacked BLs to make current from the SSL to the channels.

An operation example of the LASER method is shown in Fig. 2.4(b). As shown in Fig. 2.4(b), the arrangement of initially erased cells in a layer is all different for all stacked layers. The number of selectable layer is *nCr*, where the number of erased cells in a layer is *r*, the number of SSLs is *n*, and the

number of initial-state (fresh) cells is $n-r$. To get the maximum number of stacked BLs while fixing n to save the area consumed by them, $n = 2r$ (when n is even) or $n = 2r \pm 1$ (when n is odd) as examples of Fig.2.4(c).

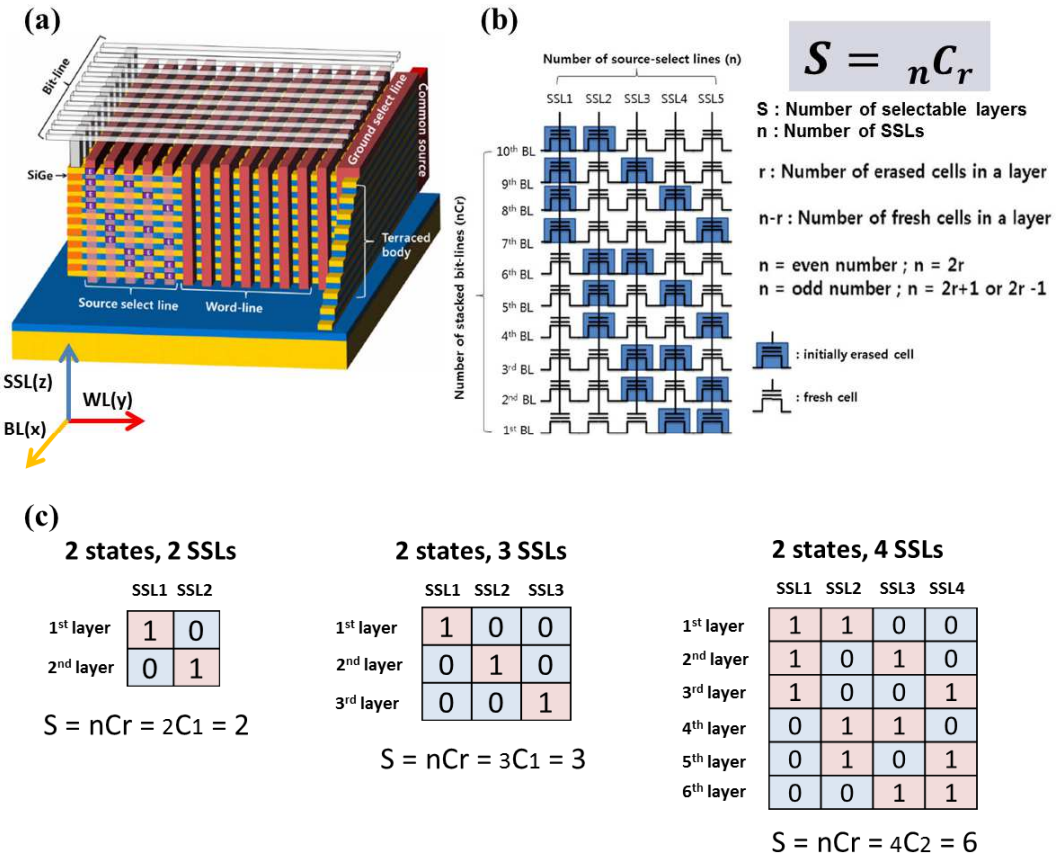


Fig. 2.4. (a) Bird's eye view of LASER structure. (b) The operation example of LASER. (c) V_{th} arrangement strategy of LASER [37].

2.2.3. Layer Selection by Multi-level operation (LSM)

For layer selection in the LSM [38], SSTs have tunnel oxide / charge trapping layer / blocking layer for initial program / erase operations and CSLs are formed separately for each layer as shown in Fig. 2.5(a). V_{th}'s of SSTs are electrically initialized by the body and the stair-like CSLs. Thus, SSTs can have multi-level V_{th} states. Fig. 2.5(b) shows the generalized LSM operation. The number of selectable layers is $k^{n/2}$ where k is the number of multi-level states of SSTs and n is the number of SSLs. The examples of the V_{th} arrangement strategy of LSM when k=3 are shown in Fig.2.5(c). V_{th}'s of the 1st SSL are arranged in ascending order and V_{th}'s of the 2nd SSL are arranged in descending order. And this V_{th} arrangement can be expanded for more SSLs. Due to the pair SSL operation, the number of SSLs should be an even number as in the case of VG-NAND.

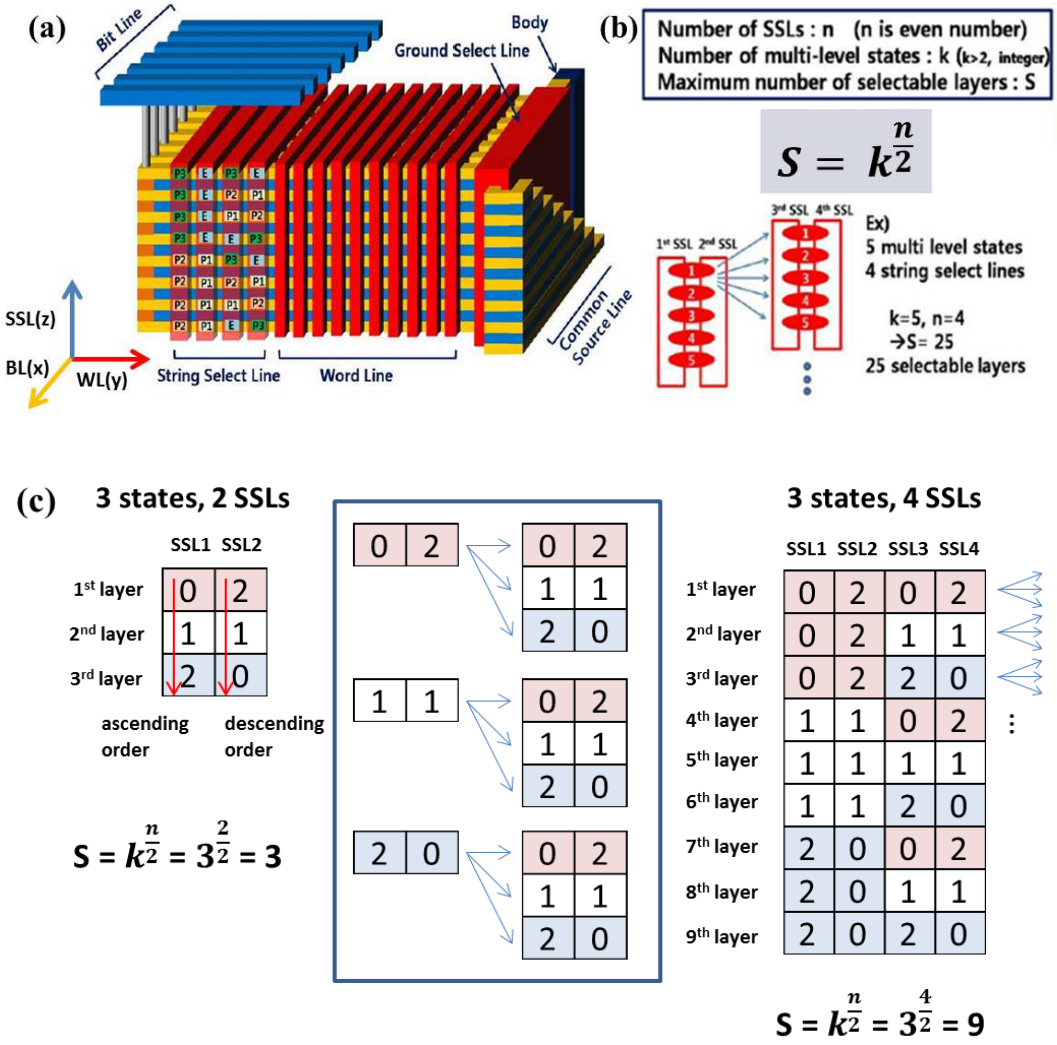


Fig. 2.5(a) Bird's eye view of LSM structure (b) Generalized LSM operation (c) The examples of the V_{th} arrangement strategy of LSM when k=3 [38].

2.3 Operation Scheme of Channel Stacked NAND Flash Memory with LSM

2.3.1. SST Initialization to Enable LSM

To perform LSM, stacked SSTs should be initialized to have specific threshold voltages. There are two different methods to initialize the stacked SSTs. Fig. 2.6 and Fig. 2.7 show the two different methods to initialize V_{th}'s of SSTs in the channel stacked NAND structure [38, 46].

One method is using separated CSL for self-boosting of the channel of the unselected layer as shown in Fig. 2.6 [38]. In the LSM structure, the array has a separated CSL on each layer. This separated CSL enables the self-boosting of the channel of the unselected layer. In the selected layer to be programmed, 0V is applied to CSL. On the contrary, in the unselected layer, V_{cc} is applied to CSL. To maintain the self-boosting program inhibition in the

channel of unselected layers, the opposite side of CSL must be cut off. There are two methods to block the BL bias. One is adding a dummy SSL between the BL and the 1st SSL, and the other is using the 1st SSL to cut off the BL voltage. The dummy SSL method is simple, but has a disadvantage of areal loss due to the additional SSL.

The other method is using one erase operation [46]. After the SSTs in all stacked layers are blind-programmed as shown in Fig. 2.7(c), the SSTs of the selected layer can be set to targeted V_{th}'s simultaneously during the erase operation by adjusting the time for making each SST gate floated state, as shown in the erase pulse of Fig. 2.7(a). The simulated result of Fig. 2.7(b) shows that the SSTs have different gate-to-channel potential (V_{GC}) values according to the gate floated time since each SST gate potential is coupled to the channel potential rising from the moment each SST gate is floated. To prevent the erase disturbance, V_{inhibit} is applied to the CSLs of the unselected layers. The figure shows the cross-sectional images of the proposed SSTs, with

color coding indicating electrostatic potential contours. The result shows that each SST has a different gate potential according to the gate-floated time and the SST gate potential induces a soft program in the unselected layer. The simulation results in LSM structure show that the SSTs successfully achieve multi-level states by these V_{GC} differences, as shown in Fig. 2.7(d).

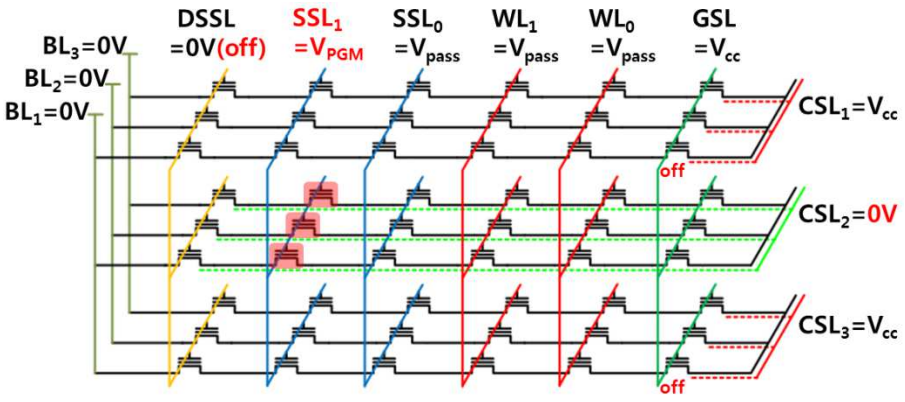


Fig. 2.6. Method to initialize V_{th}'s of SSTs using the separated CSLs.

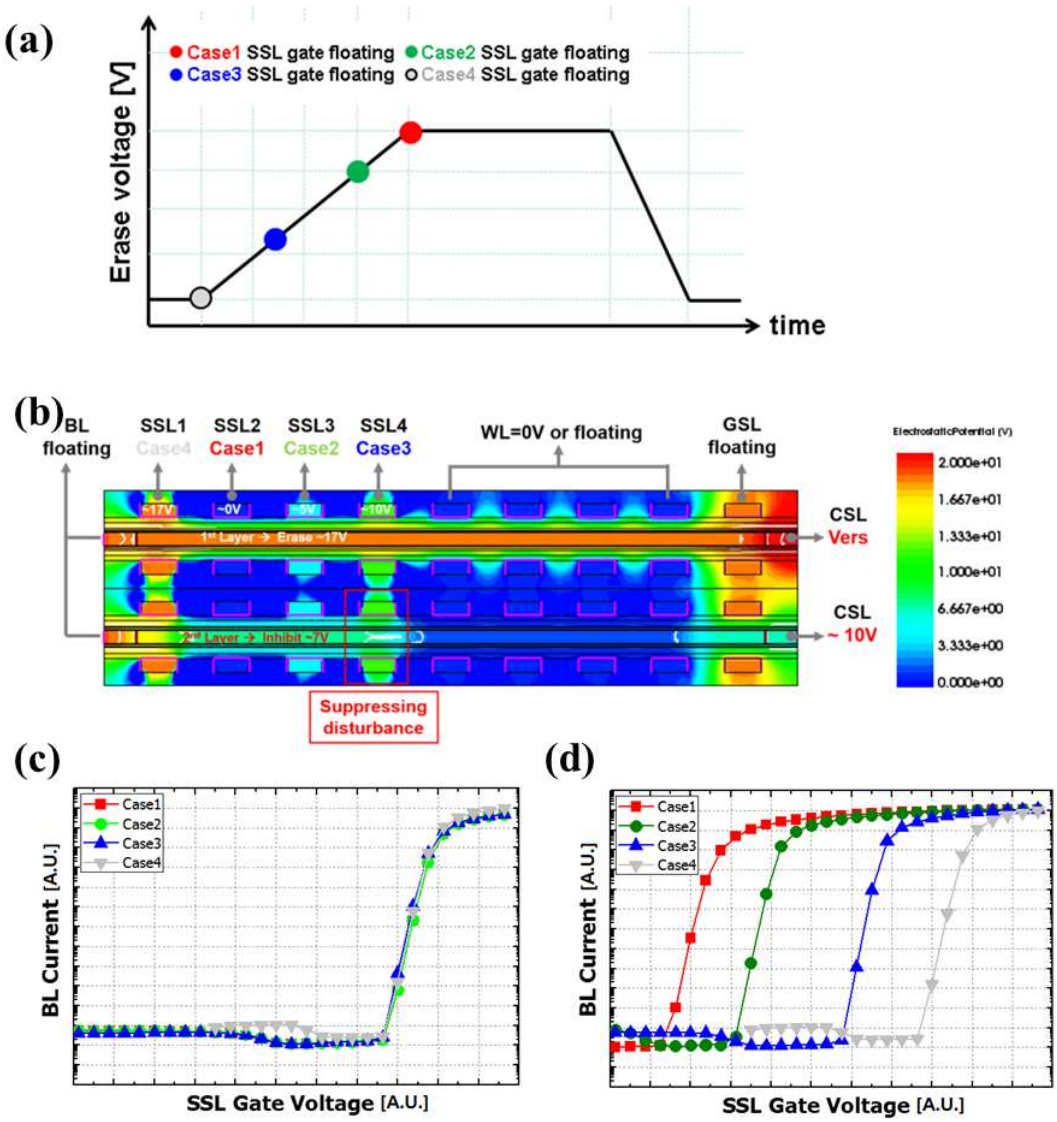


Fig. 2.7. (a) Erase pulse that shows the floated time of each SST gate. (b) Simulation result of the SST initializing method by using one erase operation. (c) SSTs in all stacked layers are blind-programmed. (d) Transfer characteristics of the SSTs that are set to the targeted V_{th}'s by one erase operation [46].

2.3.2. Read / Program / Erase Operation with LSM

For the channel stacked NAND flash memory array with LSM, read, program and erase operation schemes are similar to those of conventional planar NAND flash memory except for layer selection [48]. Fig. 2.8 shows the equivalent circuit diagram for program operation in the channel stacked NAND flash memory with LSM. Layer selection is performed by the LSM method. In the selected layer, BLs and WLs perform program operations similar to those of conventional NAND flash memory.

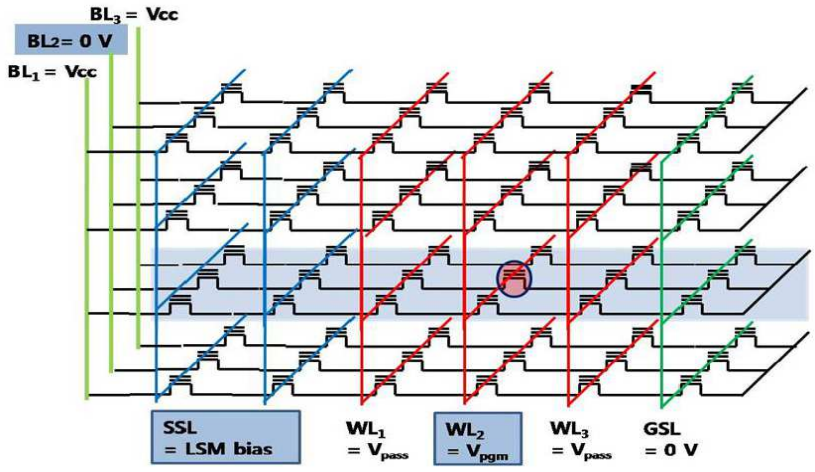


Fig. 2.8. Equivalent circuit diagram for program operation in LSM [38].

Figs. 2.9(a) and (b) show two different cases for program inhibition in the channel stacked NAND flash memory array. One is unselected BL program inhibition in the selected layer, and the other is unselected layer program inhibition in the selected BL.

In the case of unselected BL program inhibition in the selected layer, when V_{cc} is applied to the BL and SSLs of the layer is under the selection bias, self-boosting program inhibition occurs. Otherwise, in the case of unselected layer program inhibition in the selected BL, the BL is applied with 0 V. However, SSTs of the unselected layer are turned off, and then self-boosting program inhibition also occurs. For maintaining the self-boosting in the inhibit channel, ground select transistors must be turned off as in conventional NAND flash memory operation. In the case of a selected string, as 0 V is applied to the BL and SSTs are turned on, a 0 V bias of the BL is applied to the channel of the selected string. Therefore, it is possible to program the selected cell in the selected string, while the unselected cells are under the self-boosting

program inhibition, as shown in Fig. 2.9.

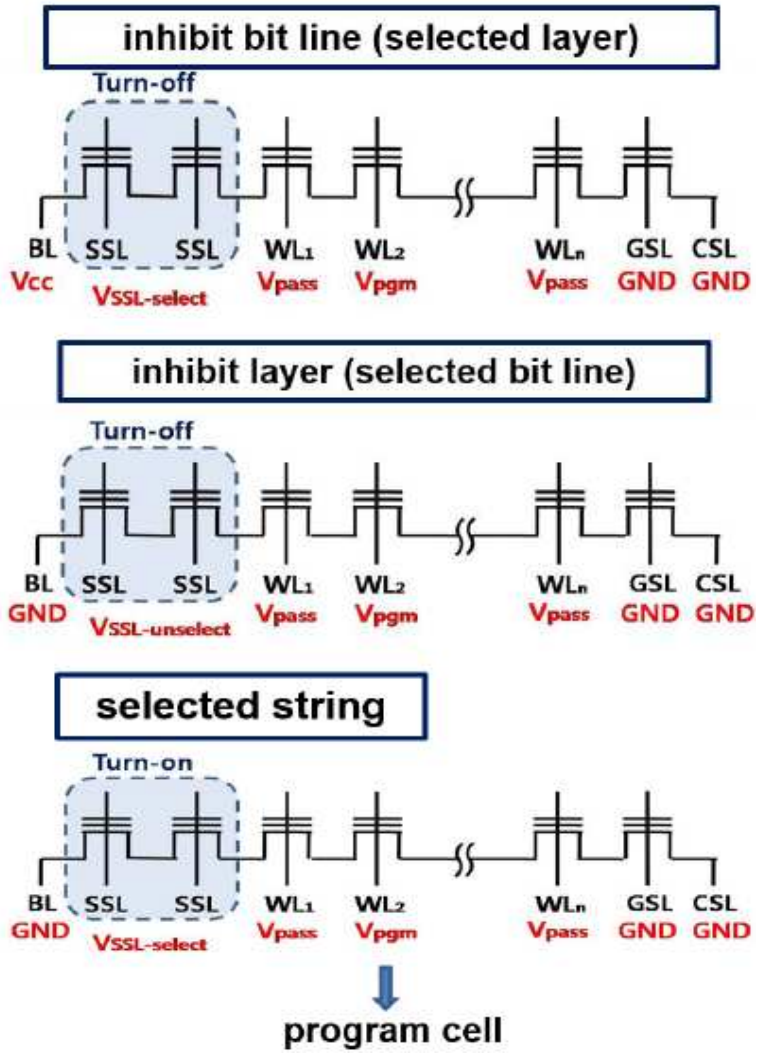


Fig. 2.8. (a) Unselected BL program inhibition in the selected layer, (b) unselected layer program inhibition in the selected BL, and (c) program operation in the LSM array [48].

The erase operation in the channel stacked NAND flash memory array with LSM is the same as that in the conventional NAND flash memory array, because the body and GSL are designed to connect the channel to the body. All the word lines are set to 0 V, and a high erase voltage is applied to the channel through the body contact to induce Fowler–Nordheim tunneling so that the erase operation is possible. All the CSL, SSL, and BL voltages are floated.

Chapter 3

Layer Selection by Multi-level Permutation Method

3.1 Introduction

In the LSM architecture of Fig. 3.1(a), all SSTs that share the same layer and SSL must be set to the same V_{th} for layer selection. Once V_{th}'s of SSTs are properly set, when viewed from the side as shown in Fig. 3.1(b), the V_{th} arrangement can be tabulated.

In 3D stacked NAND flash memory architectures with layer selection methods by V_{th}'s of SSTs, as the number of SSLs increases, the number of

WLs decreases in a limited area, which degrades the memory density. Since SSL often has much longer gate length to suppress leakage current, the array overhead (including SSL/GSL, BL pad, contact, and CSL area) can become significant [41]. As the number of V_{th} states of SSTs is increased, V_{th} margin is narrowed due to the limited memory window of SST and the layer selection may become unstable. Therefore, V_{th} arrangement of SSTs should be performed so that the maximum number of layers can be distinguished by the limited number of SSLs and V_{th} states.

To generalize how to arrange V_{th}'s of SSTs for layer selection, we first need to figure out under what conditions a unique selection is possible. And then, we have to figure out the condition that leads to the maximum number of layer selection cases where unique selection is possible. Then, the maximum number of layers can be decoded under the given number of SSLs and V_{th} states. A V_{th} table shown in Table 3.1 can be filled with V_{th} arrangement obtained by LSMP method in this chapter. After V_{th}'s of SSTs are set and

appropriate SSL biases are applied, each layer can be selected uniquely.

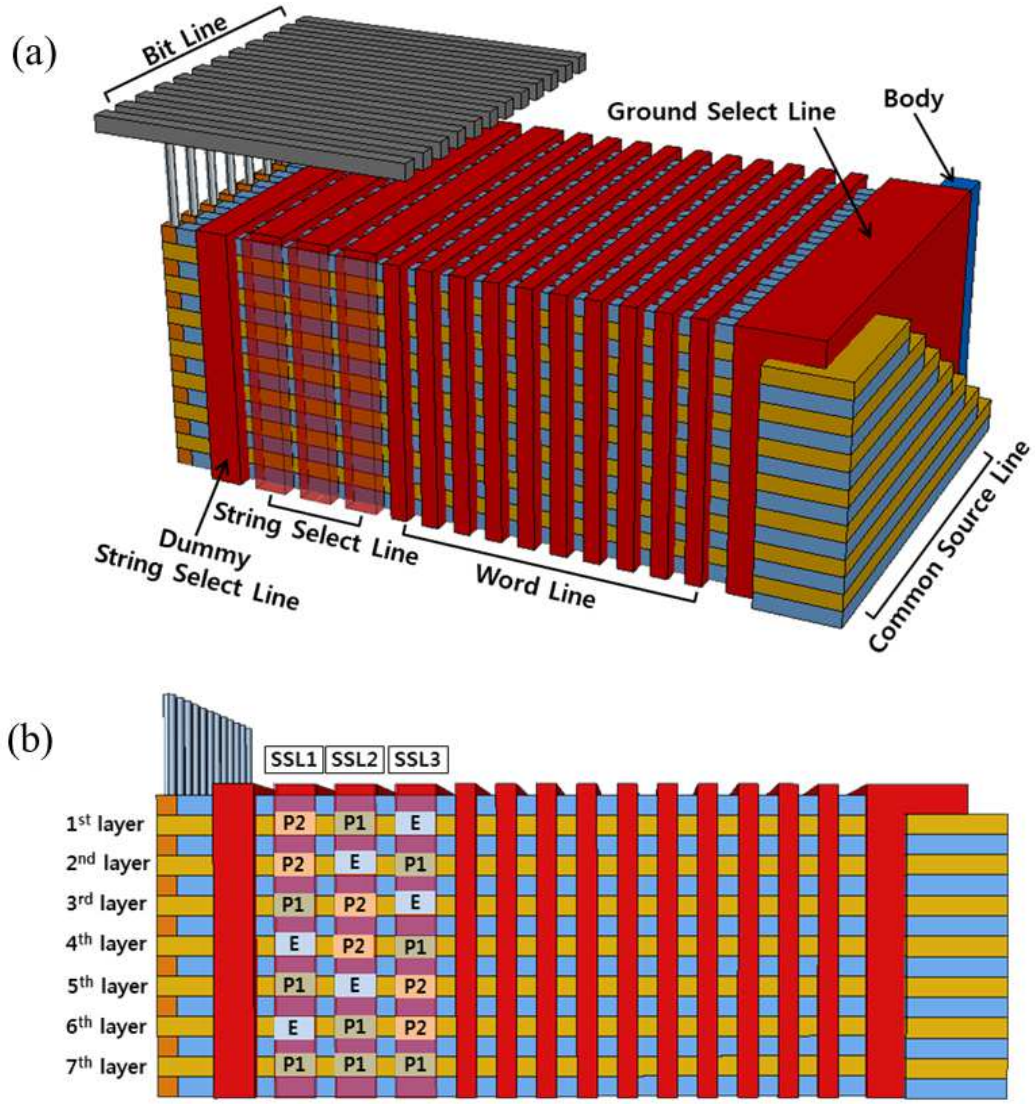


Fig. 3.1. (a) Bird's eye view of channel stacked NAND flash memory with LSM. (b) Side view of the 3D NAND array architecture with LSMP method.

Table 3.1. V_{th} arrangement table for string select transistors. Each cell of the table is filled with V_{th} state of string select transistor for layer selection.

V_{th} table	1st SSL	2nd SSL	⋯	nth SSL
1st Layer				
2nd Layer				
3rd Layer				
⋮				

3.2 How to Arrange V_{th} states of SSTs for Unique Selection

The unique selection means that when one layer is selected in the 3D stacked NAND flash memory, the other layers are inhibited. The method of V_{th} arrangement for the unique selection can be explained based on the multinomial theorem. Let the number of SSLs be n and the number of V_{th} states of SST be k. Then, we can consider a set $\boldsymbol{V} = \{V_0, V_1, V_2, \dots, V_{k-1}\}$ whose elements are the multi-level V_{th} states such that $V_0 < V_1 < \dots < V_{k-1}$. The number of possible arrangements of V_{th} states is k^n because this is the same as n-tuples of k elements where repetition is allowed. However, since not all of these cases enable the unique selection, we have to find specific conditions for the unique selection.

Now, let us define a multinomial formula of threshold voltages. For any positive integer k and n, the multinomial formula tells us how a sum with k terms expands when raised to an arbitrary power n :

$$(\boldsymbol{V_0} + \boldsymbol{V_1} + \cdots + \boldsymbol{V_{k-1}})^{\boldsymbol{n}} = \sum_{\boldsymbol{m_0}+\boldsymbol{m_1}+\cdots+\boldsymbol{m_{k-1}}=\boldsymbol{n}} \frac{\boldsymbol{n}!}{\boldsymbol{m_0}!\,\boldsymbol{m_1}!\cdots\boldsymbol{m_{k-1}}!} \prod_{\boldsymbol{i}=\boldsymbol{0}}^{\boldsymbol{k-1}} \boldsymbol{V_i}^{\boldsymbol{m_i}}. \quad (3.1)$$

When the multinomial formula is expanded, each term can be expressed as follow:

$$\frac{\boldsymbol{n}!}{\boldsymbol{m_0}!\,\boldsymbol{m_1}!\cdots\boldsymbol{m_{k-1}}!} \boldsymbol{V_0}^{\boldsymbol{m_0}} \boldsymbol{V_1}^{\boldsymbol{m_1}} \cdots \boldsymbol{V_{k-1}}^{\boldsymbol{m_{k-1}}}$$

where $n = m_0 + m_1 + \cdots + m_{k-1}$. The sum is taken over all combinations of nonnegative integer indices m_0 through m_{k-1} such that the sum of all m_i is *n*. That is, for each term in the expansion, the exponents of the V_i must add up to *n*.

The value of *l* can be defined which is determined by following equations:

$$\boldsymbol{l = \sum_{i=0}^{k-1} i \cdot m_i \quad = (0 \times m_0) + (1 \times m_1) + \cdots + \{(k-1) \times m_{k-1}\}} \qquad (3.2)$$

where $0 \leq l \leq n(k-1)$ and l is a nonnegative integer. l is equal to the sum of V_{th} state numbers (V_{th} indices) of each term in the expansion and i is V_{th} state number. The terms of Eq. (3.1) with the same l can be grouped. The unique selection can be achieved by arranging V_i's of the grouped terms as permutations. For example, when n=3 and k=3, the third power of the trinomial ($V_0 + V_1 + V_2$) is given by

$$\begin{aligned}
(\boldsymbol{V_0 + V_1 + V_2})^{\boldsymbol{3}} = {} & \left(\frac{\boldsymbol{3!}}{\boldsymbol{3!\,0!\,0!}} \boldsymbol{V_0{}^3 V_1{}^0 V_2{}^0}\right)_{\boldsymbol{l=0}} + \left(\frac{\boldsymbol{3!}}{\boldsymbol{2!\,1!\,0!}} \boldsymbol{V_0{}^2 V_1{}^1 V_2{}^0}\right)_{\boldsymbol{l=1}} \\
& + \left(\frac{\boldsymbol{3!}}{\boldsymbol{2!\,0!\,1!}} \boldsymbol{V_0{}^2 V_1{}^0 V_2{}^1} + \frac{\boldsymbol{3!}}{\boldsymbol{2!\,1!\,0!}} \boldsymbol{V_0{}^1 V_1{}^2 V_2{}^0}\right)_{\boldsymbol{l=2}} \\
& + \left(\frac{\boldsymbol{3!}}{\boldsymbol{1!\,1!\,1!}} \boldsymbol{V_0{}^1 V_1{}^1 V_2{}^1} + \frac{\boldsymbol{3!}}{\boldsymbol{0!\,3!\,0!}} \boldsymbol{V_0{}^0 V_1{}^3 V_2{}^0}\right)_{\boldsymbol{l=3}} \\
& + \left(\frac{\boldsymbol{3!}}{\boldsymbol{0!\,2!\,1!}} \boldsymbol{V_0{}^0 V_1{}^2 V_2{}^1} + \frac{\boldsymbol{3!}}{\boldsymbol{1!\,0!\,2!}} \boldsymbol{V_0{}^1 V_1{}^0 V_2{}^2}\right)_{\boldsymbol{l=4}} \\
& + \left(\frac{\boldsymbol{3!}}{\boldsymbol{0!\,1!\,2!}} \boldsymbol{V_0{}^0 V_1{}^1 V_2{}^2}\right)_{\boldsymbol{l=5}} + \left(\frac{\boldsymbol{3!}}{\boldsymbol{0!\,0!\,3!}} \boldsymbol{V_0{}^0 V_1{}^0 V_2{}^3}\right)_{\boldsymbol{l=6}}
\end{aligned}$$

$$= \left(\boldsymbol{V_0}^{3}\boldsymbol{V_1}^{0}\boldsymbol{V_2}^{0}\right)_{\boldsymbol{l=0}} + \left(\boldsymbol{3V_0}^{2}\boldsymbol{V_1}^{1}\boldsymbol{V_2}^{0}\right)_{\boldsymbol{l=1}}$$

$$+\left(\boldsymbol{3V_0}^{2}\boldsymbol{V_1}^{0}\boldsymbol{V_2}^{1} + \boldsymbol{3V_0}^{1}\boldsymbol{V_1}^{2}\boldsymbol{V_2}^{0}\right)_{\boldsymbol{l=2}}$$

$$+\left(\boldsymbol{6V_0}^{1}\boldsymbol{V_1}^{1}\boldsymbol{V_2}^{1} + \boldsymbol{V_0}^{0}\boldsymbol{V_1}^{3}\boldsymbol{V_2}^{0}\right)_{\boldsymbol{l=3}}$$

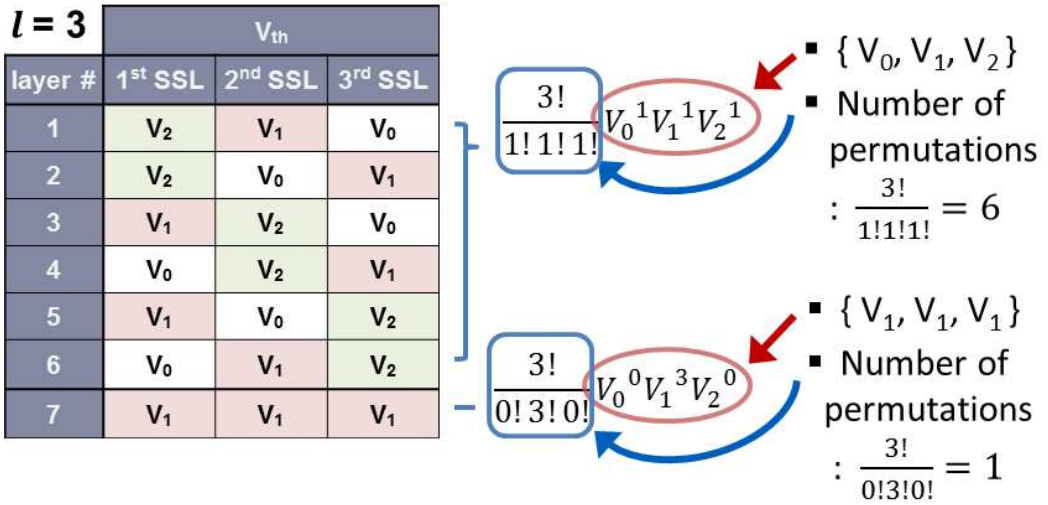
$$+\left(\boldsymbol{3V_0}^{0}\boldsymbol{V_1}^{2}\boldsymbol{V_2}^{1} + \boldsymbol{3V_0}^{1}\boldsymbol{V_1}^{0}\boldsymbol{V_2}^{2}\right)_{\boldsymbol{l=4}}$$

$$+\left(\boldsymbol{3V_0}^{0}\boldsymbol{V_1}^{1}\boldsymbol{V_2}^{2}\right)_{\boldsymbol{l=5}} + \left(\boldsymbol{V_0}^{0}\boldsymbol{V_1}^{0}\boldsymbol{V_2}^{3}\right)_{\boldsymbol{l=6}}$$

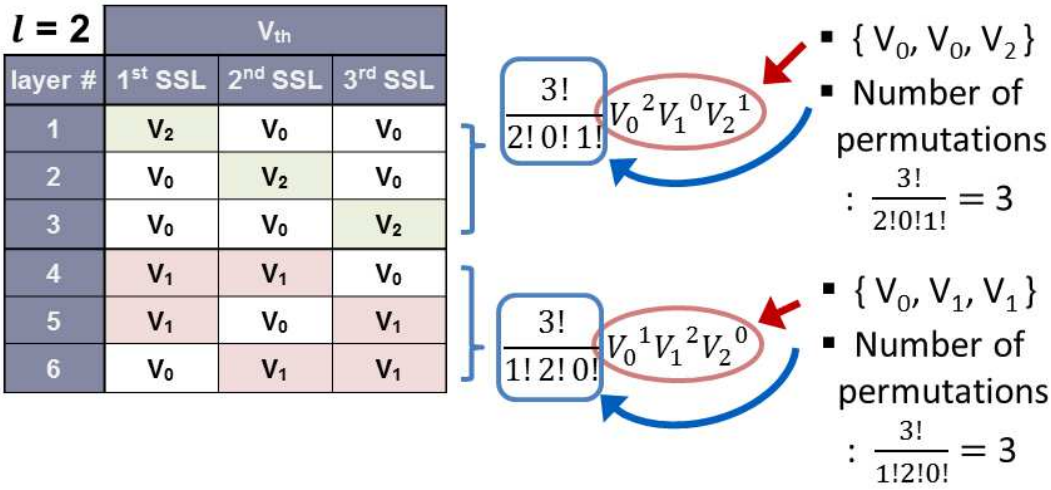
where $0 \leq l \leq 6$. One of seven groups can be selected to make a set of permutations for the unique selection. The sum of each coefficient in one group equals the number of selectable layers. When l = 3, the maximum number of layer selection is possible because the sum of the coefficients in the group is the largest 7. As shown in Fig. 3.2(a), the arrangements of V_{th}'s of SSTs are the permutations of {V_0, V_1, V_2} and {V_1, V_1, V_1} which are V_i's of the terms in the selected group. The coefficients of the terms denote the number of permutations that can be arranged using each term. The SSL bias

set for the layer selection can be determined by choosing the SSL bias voltage, V_{bi}, for the arranged V_{th} of SST, V_i, such that $V_i < V_{bi} < V_{i+1}$ ($0 \leq i \leq k$-1).

The V_{th} arrangement for the unique selection can be made for each l. For example, when l = 2, 6 layers can be uniquely selected as shown in Fig. 3.2(b). There are three permutations of the set {V_0, V_0, V_2}, namely: (V_0, V_0, V_2), (V_0, V_2, V_0), and (V_2, V_0, V_0). Also, there are permutations of the set {V_0, V_1, V_1}, namely: (V_0, V_1, V_1), (V_1, V_0, V_1), and (V_1, V_1, V_0). These are all possible orderings of these two 3-tuples.



(a)



(b)

Fig. 3.2. Possible V_{th} arrangement tables when n = 3, k = 3. (a) l = 3. (b) l = 2.

3.3 Condition for Maximum Number of Selectable Layers

The number of selectable layers is defined as $P_k(l, n)$ where k is the number of V_{th} states, l is the equal sum of V_{th} state numbers (indices) in the permutation, and n is the number of SSLs. For example, $P_k(l, n) = P_3(3, 3) = 7$ and $P_k(l, n) = P_3(2, 3) = 6$. $P_k(l, n)$ corresponding to each k, l, and n condition can be tabulated as shown in Table 3.2. These tables have some properties and patterns of numbers. One of the patterns of numbers is very similar to Pascal's triangle as shown in Fig. 3.3. Especially when $k = 2$, the pattern is the same as Pascal's triangle. This triangle can be expanded to all k values for simple calculation of $P_k(l, n)$. As shown in Fig. 3.3, each entry of each subsequent row is constructed by adding adjacent k numbers in the above row, treating blank entries as 0. This relation in the Table 3.2 can be defined using the recurrence:

$$\boldsymbol{P_k(l,n)} = \begin{cases} \sum_{\boldsymbol{m=0}}^{\boldsymbol{k-1}} \boldsymbol{P_k(l-m,n-1)} & \boldsymbol{0 \le l \le n(k-1)} \\ \boldsymbol{0} & \boldsymbol{else} \end{cases} . \qquad (3.3)$$

Using the recursion formula above, tables of $P_k(l, n)$ shown in Table 3.2 can be created without calculating permutations of V_{th}'s one by one. l-values, which maximize $P_k(l, n)$ at given k and n, are marked in red. According to Table 3.2, the condition for the maximum number of the selectable layers is that l is equal to the integer which is closest to the median in the range of l, where $0 \le l \le n(k-1)$. The condition for the maximum number of the selectable layers in the all cases is given by

$$\boldsymbol{l = \left\lfloor \frac{n}{2}(k-1) \right\rfloor}. \qquad (3.4)$$

where the bracket is notation of the floor function. The floor function, $\lfloor x \rfloor$,

maps x to the greatest integer less than or equal to x. In the cases that n is odd and k is even, the condition for the maximum selection is given by

$$\boldsymbol{l = \left\lfloor \frac{n}{2}(k-1) \right\rfloor \text{ or } \left\lceil \frac{n}{2}(k-1) \right\rceil} \tag{3.5}$$

where the ceiling function, $\lceil x \rceil$, maps x to the least integer greater than or equal to x. There are two integers closest to the median of the range of l.

Now, it is possible to complete Table 3.1 of V_{th}'s of SSTs more simply without expanding the multinomial formula, Eq.(3.1), by calculating the l-value that can maximize the number of selectable layers. Let us define a set I of V_{th} state numbers (indices) as $I = \{0, 1, \ldots, k\text{-}1\}$ for easy calculation. Each V_{th} state number refers to the V_{th} of that number as shown in Table 3.3. By using Eq.(3.4), the l-value which maximizes the number of selectable layers can be calculated for given n and k. Multisets of cardinality n whose sums of

the elements are equal to l can be made by taking elements n times from the set I with repetitions. All permutations of the multisets which have the equal sum l can be tabulated in Table 3.1. Finally, the state numbers of the tabulated permutations have to be replaced with the corresponding V_{th} states of the set V.

Table 3.2. $P_k(l, n)$ tables corresponding to each k, l, and n condition when (a) $k = 2$, (b) $k = 3$, and (c) $k = 4$.

(a) $k = 2$

n \ l	0	1	2	3	4	5
1	1	1	0	0	0	0
2	1	2	1	0	0	0
3	1	3	3	1	0	0
4	1	4	6	4	1	0
5	1	5	10	10	5	1

(b) $k = 3$

n \ l	0	1	2	3	4	5	6	7	8	9	10
1	1	1	1	0	0	0	0	0	0	0	0
2	1	2	3	2	1	0	0	0	0	0	0
3	1	3	6	7	6	3	1	0	0	0	0
4	1	4	10	16	19	16	10	4	1	0	0
5	1	5	15	30	45	51	45	30	15	5	1

(c) $k = 4$

n \ l	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	1	2	3	4	3	2	1	0	0	0	0	0	0	0	0	0
3	1	3	6	10	12	12	10	6	3	1	0	0	0	0	0	0
4	1	4	10	20	31	40	44	40	31	20	10	4	1	0	0	0
5	1	5	15	35	65	101	135	155	155	135	101	65	35	15	5	1

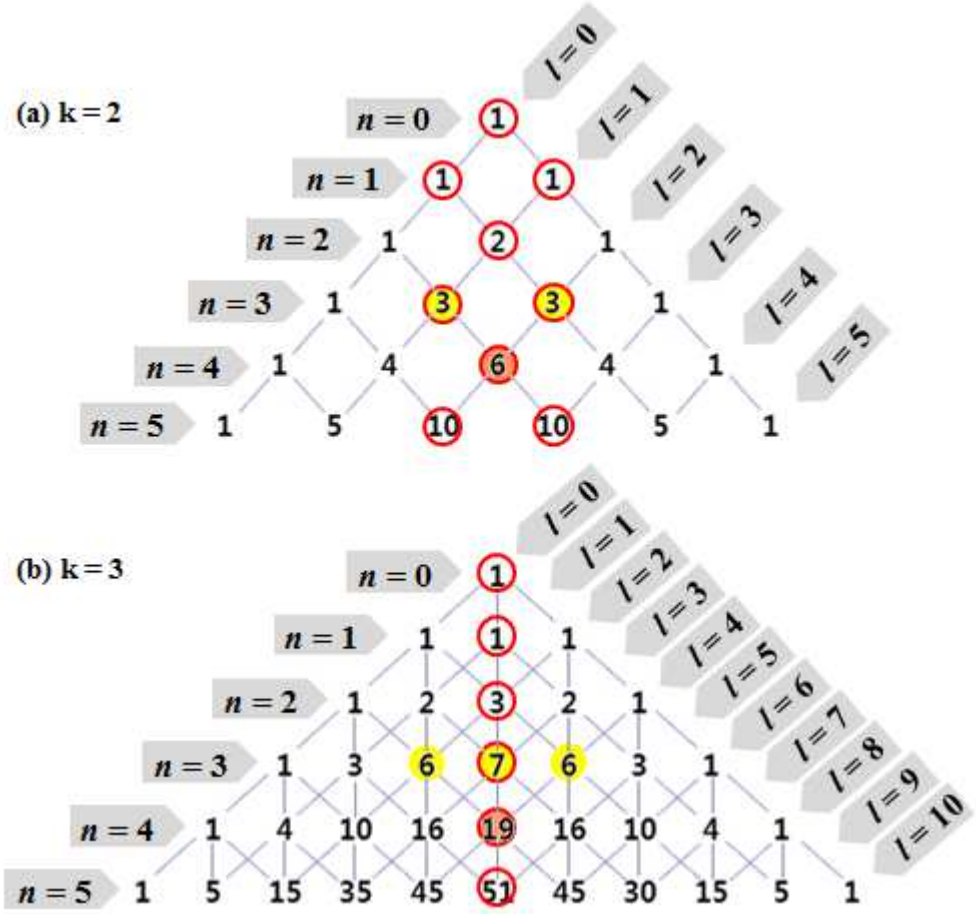


Fig. 3.3. Patterns of numbers are very similar to Pascal's triangle. (a) When k = 2, the pattern is exactly the same as Pascal's triangle. (b) Expanded triangle for k = 3. Each entry of each subsequent row is constructed by adding adjacent k numbers in the above row, treating blank entries as 0.

Table 3.3. V_{th} states and V_{th} state numbers (indices).

State	1st	2nd	3rd	…	k^{th}
V_{th}	V_0	V_1	V_2	…	V_{k-1}
I	0	1	2	…	k-1

For example, let us complete Table 3.1 for 3 SSLs and 4 V_{th} states. A set I of state number is defined as I = {0, 1, 2, 3}. The equal sum l is $\left\lfloor \frac{n}{2}(k-1) \right\rfloor = \left\lfloor \frac{3}{2}(4-1) \right\rfloor = \lfloor 4.5 \rfloor = 4$. The multisets of cardinality 3 taken from the set I with equal sum l = 4 are {3, 1, 0}, {2, 2, 0}, and {2, 1, 1}. Now, we have to find all permutations of these multisets. There are six permutations of the multiset {3, 1, 0} : namely, (3, 1, 0), (3, 0, 1), (1, 3, 0), (0, 3, 1), (1, 0, 3), and (0, 1, 3). There are 3 permutations of the multiset {2, 2, 0}: namely, (2, 2, 0), (2, 0, 2), and (0, 2, 2). There are also 3 permutations of the multiset {2, 1, 1}: namely, (2, 1, 1), (1, 2, 1), and (1, 1, 2). Thus, the number of the possible permutations of all the multisets is 12 as shown in Table 3.4. Also, the state numbers of permutations have to be replaced by the real V_{th} states of the set V.

The permutations of V_{th} states are summarized as "V_{th} of SST" in Table 3.5. Also, the corresponding SSL bias voltages for each layer selection are tabulated as "SSL bias for selection" in Table 3.5. When V_{b3}, V_{b1}, and V_{b0} are applied to the 1st SSL, 2nd SSL, and 3rd SSL, respectively, all the SSTs of the 1st layer are turned on and the channel of the 1st layer is connected to BL. On the other hand, the other layers are inhibited because they have one or more turned-off SST(s).

In this example, since n (=3) is odd and k (=4) is even, there is one more l-value for the maximum unique selection. The other equal sum l is $\left\lceil \frac{n}{2}(k-1) \right\rceil = \left\lceil \frac{3}{2}(4-1) \right\rceil = \lceil 4.5 \rceil = 5$. The multisets of cardinality 3 taken from the set I with equal sum l = 5 are {3, 2, 0}, {3, 1, 1}, and {2, 2, 1}. There are six permutations of the multiset {3, 2, 0} : namely, (3, 2, 0), (3, 0, 2), (2, 3, 0), (0, 3, 2), (2, 0, 3), and (0, 2, 3). There are 3 permutations of the multiset {3, 1, 1}: namely, (3, 1, 1), (1, 3, 1), and (1, 1, 3). There are also 3 permutations of the multiset {2, 2, 1}: namely, (2, 2, 1), (2, 1, 2), and (1, 2, 2). With the same

procedure as the case of *l*=4, the number of the possible permutations of all the multisets is 12 as shown in Table 3.6.

Table 3.7 shows the multisets whose elements are state numbers to maximize the layer selection by LSMP method. The condition is $l = \left\lfloor \frac{n}{2}(k-1) \right\rfloor$ where *n* is the number of SSLs and *k* is the number of V_{th} states.

Table 3.4. All possible permutations of multisets when *n* = 3, *k* = 4, and *l* = 4.

	V_{th} state number		
layer #	1st SSL	2nd SSL	3rd SSL
{ 3, 1, 0 } 1	3	1	0
2	3	0	1
3	1	3	0
4	0	3	1
5	1	0	3
6	0	1	3
{ 2, 2, 0 } 7	2	2	0
8	2	0	2
9	0	2	2
{ 2, 1, 1 } 10	2	1	1
11	1	2	1
12	1	1	2

Table 3.5. The state numbers of permutations of Table 3.4 have to be replaced by V_{th} states of the set V. The permutations of V_{th} states are summarized as "V_{th} of SST". Also, the corresponding SSL bias voltages for each layer selection are tabulated as "SSL bias for selection". The SSL bias sets for the layer selection are determined by choosing the SSL bias voltage, V_{bi}, for the arranged V_{th} of SST, V_i, such that $V_i < V_{bi} < V_{i+1}$ ($0 \leq i \leq k$-1).

	V_{th} of SST			SSL bias for selection		
layer #	1st SSL	2nd SSL	3rd SSL	1st SSL	2nd SSL	3rd SSL
1	V_3	V_1	V_0	V_{b3}	V_{b1}	V_{b0}
2	V_3	V_0	V_1	V_{b3}	V_{b0}	V_{b1}
3	V_1	V_3	V_0	V_{b1}	V_{b3}	V_{b0}
4	V_0	V_3	V_1	V_{b0}	V_{b3}	V_{b1}
5	V_1	V_0	V_3	V_{b1}	V_{b0}	V_{b3}
6	V_0	V_1	V_3	V_{b0}	V_{b1}	V_{b3}
7	V_2	V_2	V_0	V_{b2}	V_{b2}	V_{b0}
8	V_2	V_0	V_2	V_{b2}	V_{b0}	V_{b2}
9	V_0	V_2	V_2	V_{b0}	V_{b2}	V_{b2}
10	V_2	V_1	V_1	V_{b2}	V_{b1}	V_{b1}
11	V_1	V_2	V_1	V_{b1}	V_{b2}	V_{b1}
12	V_1	V_1	V_2	V_{b1}	V_{b1}	V_{b2}

Table 3.6. All possible permutations of multisets when n = 3, k = 4, and l = 5.

		V_{th} state number		
	layer #	1st SSL	2nd SSL	3rd SSL
{ 3, 2, 0 }	1	3	2	0
	2	3	0	2
	3	2	3	0
	4	0	3	2
	5	2	0	3
	6	0	2	3
{ 3, 1, 1 }	7	3	1	1
	8	1	3	1
	9	1	1	3
{ 2, 2, 1 }	10	2	2	1
	11	2	1	2
	12	1	2	2

Table 3.7. Multisets whose elements are state numbers to maximize the layer selection by LSMP method. The condition is $l = \left\lfloor \frac{n}{2}(k-1) \right\rfloor$ where n is the number of SSLs and k is the number of V_{th} states.

	2 states	3 states	4 states		5 states		
2 SSLs	{1,0}	{1,1} {2,0}	{2,1}	{3,0}	{2,2}	{3,1}	{4,0}
3 SSLs	{1,0,0}	{1,1,1} {2,1,0}	{2,1,1} {2,2,0}	{3,1,0}	{2,2,2} {3,3,0}	{3,2,1} {4,2,0}	{4,1,1}
4 SSLs	{1,1,0,0}	{1,1,1,1} {2,2,0,0} {2,1,1,0}	{2,2,2,0} {2,2,1,1} {3,3,0,0}	{3,2,1,0} {3,1,1,1}	{2,2,2,2} {3,3,2,0} {3,3,1,1}	{3,2,2,1} {4,4,0,0} {4,3,1,0}	{4,2,2,0} {4,2,1,1}
5 SSLs	{1,1,0,0,0}	{1,1,1,1,1} {2,2,1,0,0} {2,1,1,1,0}	{2,2,2,1,0} {2,2,1,1,1} {3,3,1,0,0}	{3,2,2,0,0} {3,2,1,1,0} {3,1,1,1,1}	{2,2,2,2,2} {3,3,3,1,0} {3,3,2,2,0} {3,3,2,1,1}	{3,2,2,2,1} {4,4,2,0,0} {4,4,1,1,0} {4,3,3,0,0}	{4,3,2,1,0} {4,3,1,1,1} {4,2,2,2,0} {4,2,2,1,1}
6 SSLs	{1,1,1,0,0,0}	{1,1,1,1,1,1} {2,2,2,0,0,0} {2,2,1,1,0,0} {2,1,1,1,1,0}	{2,2,2,2,1,0} {2,2,2,1,1,1} {3,3,3,0,0,0} {3,3,2,1,0,0}	{3,3,1,1,1,0} {3,2,2,2,0,0} {3,2,2,1,1,0} {3,2,1,1,1,1}	{2,2,2,2,2,2} {3,3,3,3,0,0} {3,3,3,2,1,0} {3,3,3,1,1,1} {3,3,2,2,2,0} {3,3,2,2,1,1}	{3,2,2,2,2,1} {4,4,4,0,0,0} {4,4,3,1,0,0} {4,4,2,2,0,0} {4,4,2,1,1,0} {4,4,1,1,1,1}	{4,3,3,2,0,0} {4,3,3,1,1,0} {4,3,2,2,1,0} {4,3,2,1,1,1} {4,2,2,2,2,0} {4,2,2,2,1,1}

3.4 Comparison with Conventional Layer Selection Methods

Table 3.8 shows the maximum number of selectable layers by VG NAND, LASER, LSM, and LSMP under a certain number of SSLs and states. LSMP method allows more layers to be selected compared to the VG NAND, LASER and LSM methods. In the LSMP method, the number of SSLs does not have to be an even number in contrast to VG NAND and LSM. Therefore, LSMP provides more flexible options than LSM and increases the maximum number of selectable layers. Let's assume a 48-layer-stacked 3D NAND array. There are 48 layers to be distinguished by SSLs. VG NAND requires 2 states and 12 SSLs to decode 48 layers. In the LASER method, 2 states and 8 SSLs are required. In the LSM method, 4 states and 6 SSLs are needed (3 states and 8 SSLs also can be used). In the LSMP method, only 3 states and 5 SSLs are required to distinguish 48 layers. Thanks to the reduced number of states, the stacked layers can be more stably selected by widening the V_{th} margin of SST

in the limited V_{th} window. Also, the minimized number of SSLs implies that the space occupied by SSLs is reduced and the memory density can be increased.

When decoding many layers by V_{th} arrangements of SSTs and SSL biases, using more number of V_{th} states will reduce the number of the required SSLs, but will narrow the V_{th} margin of SST. If the V_{th} margin is narrowed, layer selection failure can occur due to the V_{th} shift of SST during retention as shown in Fig. 3.4(a). Especially, SSTs with high V_{th}'s show severe V_{th} shift. The inhibition failure may occur because the leakage current flows through the unselected layers as shown in Fig. 3.4(b). If the number of V_{th} arrangements is maximized by LSMP method, all stacked layers can be decoded with a small number of V_{th} states. Thus, V_{th} margin can be widened and stable layer selection operation can be performed. As a result, the leakage current of the inhibited layers can be more stably suppressed as shown in Fig. 3.5(a) and (b).

Table 3.8. The maximum number of selectable layers by VG NAND, LASER, LSM, and LSMP under a certain number of SSLs and states.

number of SSLs	Maximum Number of Selectable Layers									
	VG NAND	LASER	LSM				LSMP			
	2 states	2 states	2 states	3 states	4 states	5 states	2 states	3 states	4 states	5 states
2	2	2	2	3	4	5	2	3	4	5
3	-	3	-	-	-	-	3	7	12	19
4	4	6	4	9	16	25	6	19	44	85
5	-	10	-	-	-	-	10	51	155	381
6	8	20	8	27	64	125	20	141	580	1751
7	-	35	-	-	-	-	35	393	2128	8135
8	16	70	16	81	256	625	70	1107	8092	38165

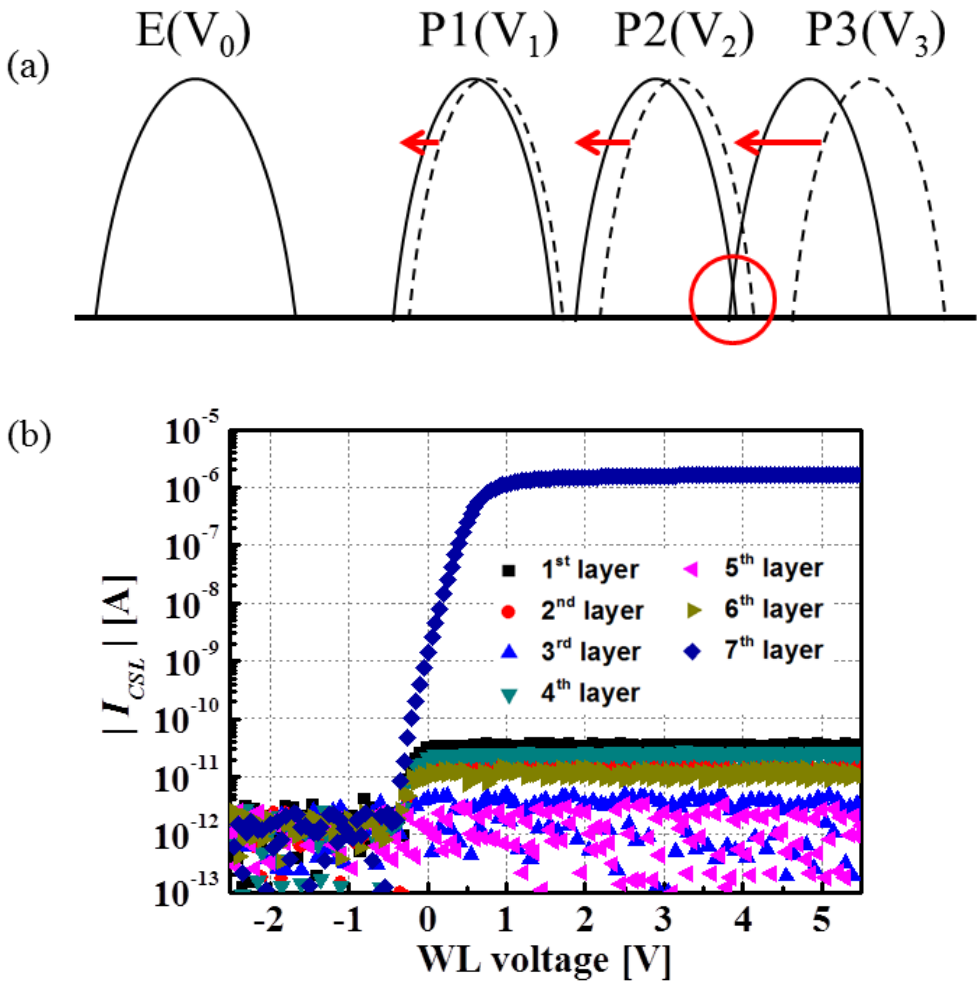


Fig. 3.4. (a) Narrow V_{th} margin of SST due to increased number of V_{th} states. Due to V_{th} shift during retention, an overlap of V_{th} distributions is produced. (b) The inhibition failure may occur because the leakage current flows through the unselected layers.

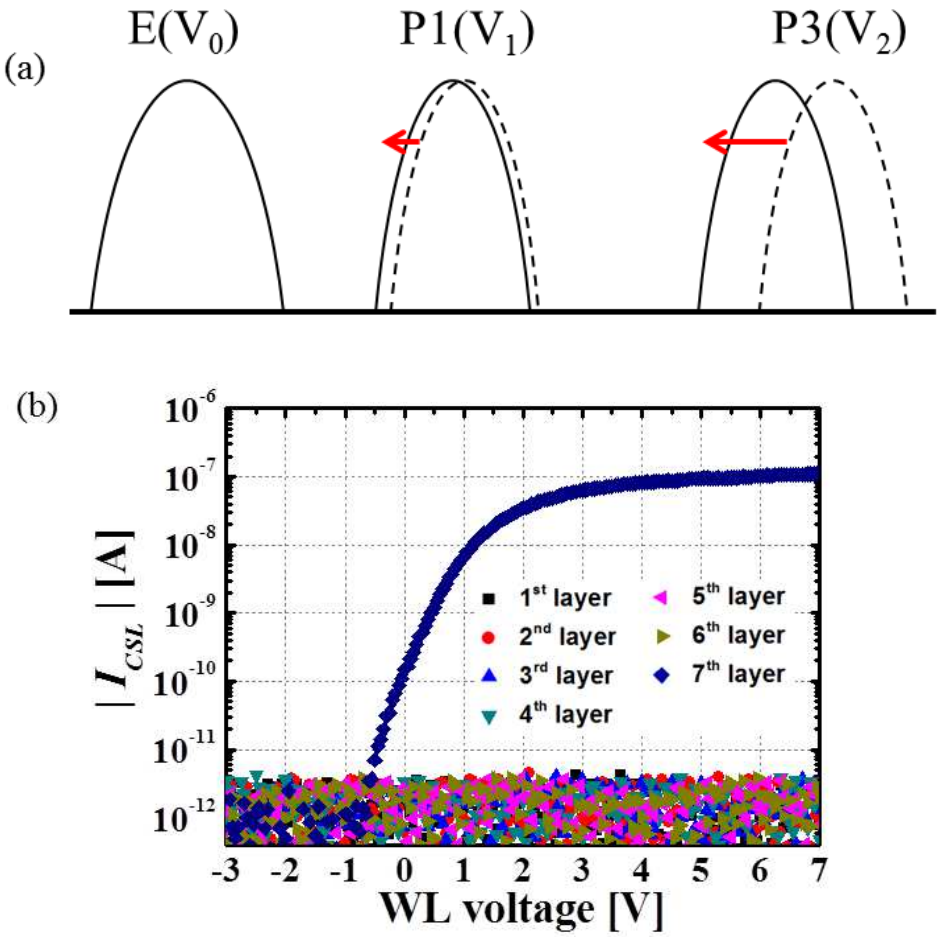


Fig. 3.5. (a) Widened V_{th} margin due to the reduced number of the V_{th} states. (b) Suppressed leakage current of the inhibited layers due to widened V_{th} margin. As a result, stable layer selection operation can be performed.

3.5 Mathematical proofs of LSMP method

Definition 1. Let $V = \{V_0, V_1, \ldots V_{k-1}\}$ ($k \geq 2$) be a set of real numbers that satisfy $V_0 < V_1 < \ldots < V_{k-1}$. i and j are integers such that $0 \leq i \leq k\text{-}1$ and $0 \leq j \leq k\text{-}1$. Then, a threshold function, $T_i(j)$, is defined as: 1 if $V_j \leq V_i$, and 0 if $V_j > V_i$.

Definition 2. Consider a permutation of $V_0, V_1, \ldots, V_{k-1}$ taken n (≥ 2) at a time with repetitions. Let the repetition number of V_i's in this permutation be m_i. Then,

$$\sum_{i=0}^{k-1} m_i = n\,, \quad 0 \leq m_i \leq n\,.$$

Now, we define a set S_l of equal sum permutations with sum of arranged state numbers (l), as a set of all possible permutations for which

$$\sum_{i=0}^{k-1} i \cdot m_i = l\,, \quad 0 \leq l \leq n(k-1)$$

where i is V_{th} state number.

Definition 3. A permutation is uniquely selectable by thresholds, if we can find a set of threshold functions such that all the values of threshold function for the elements of that permutation are 1.

Definition 4. The number of elements of S_l is defined as $P_k(l, n)$, where k is the number of elements of V and n is the number taken at a time.

$$\boldsymbol{P_k(l,n)} = \begin{cases} \boldsymbol{positive\ integer} & \boldsymbol{0 \le l \le n(k-1)} \\ \boldsymbol{0} & \boldsymbol{else} \end{cases}$$

$P_k(l, n)$ is the sum of $\frac{n!}{m_0!m_1!\cdots m_{k-1}!}$ for all possible combinations of m_i's that satisfy the equation below:

$$\sum_{i=0}^{k-1} \boldsymbol{i \cdot m_i = l}\,.$$

Theorem 1. In a set of equal sum permutations, all permutations are uniquely selectable by thresholds.

Proof : Let the given permutation be $V_{p1}V_{p2}\ldots V_{pn}$. If we take the threshold functions as T_{p1}, T_{p2}, …, T_{pn}, $T_{p1} = T_{p2} = \ldots = T_{pn}$=1. Thus, this permutation is selected. Now, let us assume there is another permutation that is also selected by the same set of threshold functions. Then, at least one of the values of V_{pi}'s should be smaller than the value of the given permutation at the same position. This is contradictory to the definition of equal-sum permutation. Thus, any permutation is uniquely selectable by thresholds.

Theorem 2. If any elements of other sets of equal sum permutations are added to the given set of equal-sum permutations, the permutations in the given set cannot be uniquely selectable by thresholds.

Proof : Let the sum of the given set be *l*. If the sum for the other set is *l'*

(< l), then we can always find a permutation in the given set such that all the elements of the given set are larger than or equal to those of the other set. If the sum for the other set is l" (> l), then we can always find a permutation in the other set such that all the elements of the given set are smaller than or equal to those of the other set. Thus, the resulting set cannot be uniquely selectable by thresholds.

Theorem 3. $P_k(l, n)$ is the number of cases where the integer l is divided into n non-negative integers less than or equal to k-1. This is the same as if one has l identical objects to be placed into n boxes, $P_k(l, n)$, and the number of the objects placed into one box cannot be over k-1. This is a kind of the partition problem or combination with repetitions.

Proof : Suppose one has l identical objects placed to be n boxes. The number of the objects placed into one box cannot be over k-1. In the example

below l = 3, k = 4, and n = 4. Instead of starting to place objects into boxes, one starts by placing the l objects and n-1 partitions on a line as shown in Fig. 3.6.

Fig. 3.6. l objects and (n-1) partitions are placed on a line.

One may place multiple partitions between two objects, as well as placing partitions before the first object or after the last object. The objects for the first box will be taken from the left, followed by the objects for the second box, and so forth. The number of the partitioned objects cannot be over k-1. Thus, for example, the arrangements of (0, 2, 1, 0) or (0, 3, 0, 0) may be represented by Fig. 3.7.

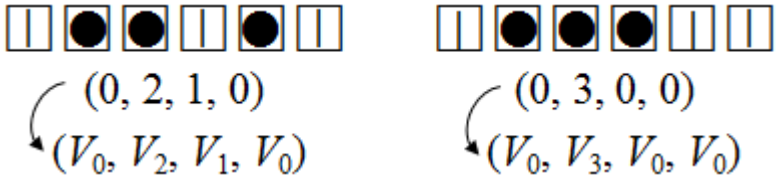


Fig. 3.7. The arrangements of (0, 2, 1, 0) or (0, 3, 0, 0) using partition number theory. The arranged integers (state numbers) can be replaced with the corresponding V_{th} states.

The desired arrangements are consist of $l+n-1$ objects (l objects and $n-1$ partitions). Choosing the positions for the objects leaves exactly $n-1$ spots left for the $n-1$ partitions. That is, choosing the positions for the objects determines the entire arrangement. However, these arrangements cannot be counted by the combination, ${}_{l+n-1}C_l$ because there is the upper limit of the number the partitioned objects, k-1. Also, this problem can be considered as a kind of combination with repetitions as shown in Fig .3.8.

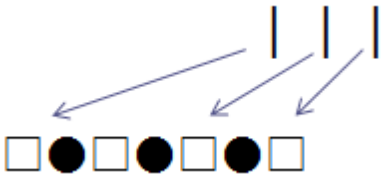


Fig. 3.8. The arrangement method by using combination with repetitions. The number of the partitioned objects cannot be over *k*-1.

One may place objects on a line first. And then, multiple partitions are placed between two objects, as well as placing partitions before the first object or after the last object with repetitions. The number of the partitioned objects cannot be over *k*-1.

Theorem 4. If $n = 1$,

$$\boldsymbol{P_k(l,n) = P_k(l,1) = \begin{cases} 1 & 0 \le l \le n(k-1) \\ 0 & else \end{cases}}$$

Proof : Suppose one has *l* identical objects to be placed into a single box,

$P_k(l,\ 1)$. The number of the objects placed into one box cannot be over k-1. The only way to put l objects into one box is to put all the given l objects into the box as the example below:

$$P_k(0,\ 1) \rightarrow (0), \quad P_k(1,\ 1) \rightarrow (1), \quad \cdots \quad , \quad P_k(k-1,\ 1) \rightarrow (k-1).$$

Theorem 5. If $n \geq 2$,

$$\boldsymbol{P_k(l,n)} = \begin{cases} \displaystyle\sum_{\boldsymbol{r=0}}^{\boldsymbol{k-1}} \boldsymbol{P_k(l-r,n-1)} & \boldsymbol{0 \leq l \leq n(k-1)} \\ \boldsymbol{0} & \boldsymbol{else} \end{cases}$$

Proof : Suppose one has l identical objects to be placed into n boxes, $P_k(l, n)$. The number of the objects placed into one box cannot be over k-1. As shown in Fig. 3.9, this is the same as the sum of the number of cases that l-r objects are placed into n-1 boxes after r objects are placed into first box where $0 \leq r \leq k$-1.

$$P_3(3,3) = \sum_{r=0}^{2} P_3(3-r, 2) = 7$$

$P_3(3, 2) = 2$ $(2, 1)$ $(1, 2)$ $\xrightarrow{r=0}$ $(0, 2, 1)$ $(0, 1, 2)$

$P_3(2, 2) = 3$ $(2, 0)$ $(0, 2)$ $(1, 1)$ $\xrightarrow{r=1}$ $(1, 2, 0)$ $(1, 0, 2)$ $(1, 1, 1)$

$P_3(1, 2) = 2$ $(1, 0)$ $(0, 1)$ $\xrightarrow{r=2}$ $(2, 1, 0)$ $(2, 0, 1)$

k=3 $V=\{V_0, V_1, V_2\}$

	$P_k(l, n)$				
n \ l	0	1	2	3	4
1	1	1	1	0	0
2	1	2	3	2	1
3	1	3	6	7	6
4	1	4	10	16	19

Fig. 3.9. r objects are placed into first box where $0 \le r \le k\text{-}1$. And then, the number of cases that the rest, l-r, are put in n-1 boxes is equal to $P_k(l\text{-}r, n\text{-}1)$. The number of cases has to be counted for all r. After arranging, the integers can be replaced by V_{th} state of the set V.

Also, it can be calculated as follows:

$$\left(\sum_{r=0}^{k-1} q^r\right)^n = \sum_{l=0}^{n(k-1)} P_k(l,n) q^l$$

$$\left(\sum_{r=0}^{k-1} q^r\right)^{n-1} = \sum_{l=0}^{(n-1)(k-1)} P_k(l,n-1) q^l$$

$$\sum_{l=0}^{n(k-1)} P_k(l,n) q^l = \left\{\sum_{l=0}^{(n-1)(k-1)} P_k(l,n-1) q^l\right\}\left(\sum_{r=0}^{k-1} q^r\right)$$

$$\Rightarrow P_k(l,n) = \sum_{r}^{k-1} P_k(l-r, n-1), \qquad P_k(l,n) = 0 \quad \forall\, l < 0, l > n(k-1)$$

Theorem 6. $P_k(0, n) = P_k(n(k\text{-}1), n) = 1$.

Proof : If $l = 0$, $m_0 = n$ and $m_i = 0$ for all $i \neq 0$. Thus, $P_k(0, n) = \frac{n!}{n!} = 1$.

If $l = n(k\text{-}1)$, $m_{k-1} = n$ and $m_i = 0$ for all $i \neq k\text{-}1$. Thus, $P_k(n(k\text{-}1), n) = \frac{n!}{n!} = 1$

Theorem 7. $P_k\,(n(k\text{-}1)\text{-}l, n) = P_k(l, n)$.

Proof :

$$\boldsymbol{n(k-1)-l=(k-1)\sum_{i=0}^{k-1} m_i-\sum_{i=0}^{k-1} i\cdot m_i}$$

$$\boldsymbol{=\sum_{i=0}^{k-1}(k-1-i)\cdot m_i}$$

$$\boldsymbol{=\sum_{i=0}^{k-1} i\cdot m_{k-1-i}}$$

That is, if we substitute the repetition numbers (m_i's) of elements of S_l with those (m_{k-1-i}'s) of $S_{n(k-1)-l}$, we can obtain P_k**(**$n(k\text{-}1)\text{-}l$, n**)**. Since $\frac{n!}{m_{k-1}!\cdots m_0!} = \frac{n!}{m_0!\cdots m_{k-1}!}$,

$$\boldsymbol{P_k(n(k-1)-l,n)} = \sum \frac{\boldsymbol{n!}}{\boldsymbol{m_{k-1}!\cdots m_0!}} = \sum \frac{\boldsymbol{n!}}{\boldsymbol{m_0!\cdots m_{k-1}!}} = \boldsymbol{P_k(l,n)}.$$

Theorem 8. $P_2(l, n) = {}_nC_l$. <Combination>

Proof : If $k = 2$, the set V contains only two elements (V_0 and V_1). Since $m_0 = n - l$ and $m_1 = l$, $P_2(l, n) = \frac{n!}{(n-l)!\, l!} = {}_nC_l$.

Theorem 9. If $0 \le l < k$ and $l \le n(k\text{-}1)/2$, $P_k(l, n) = {}_{l+n-1}C_l$.

Proof : $P_k(l, n)$ is the same as the number of permutations for the case of l identical objects and $(n\text{-}1)$ partitions. Therefore, $P_k(l, n) = \frac{(l+n-1)!}{l!\,(n-1)!} = {}_{l+n-1}C_l$

Theorem 10. If $k \le l < 2k$ and $l \le n(k\text{-}1)/2$,

$$\boldsymbol{P_k(l,n)} = {}_{\boldsymbol{n+l-1}}\boldsymbol{C_l} - {}_{\boldsymbol{n}}\boldsymbol{C_1}\sum_{\boldsymbol{i=0}}^{\boldsymbol{l-k}} \boldsymbol{P_k(i,n-1)}$$

$$= {}_{n+l-1}C_l - {}_nC_1 \sum_{i=0}^{l-k} {}_{n+i-2}C_i$$

Proof : If $l \geq k$, the simple equation in **Theorem. 9** overestimates the number of permutations, since the elements we can use are limited to those with indices from 0 to k-1 (V_0, V_1, … , V_{k-1}). Thus, we have to subtract the number of permutations containing elements (V_k , … , V_l) with indices $k \sim l$ from ${}_{n+1-1}C_l$. If $l < 2k$, any permutation can contain no more than one element with indices from k to l. Otherwise, the sum of indices would be larger than l. The number of permutations that contain the element with index (l-i) is ${}_nC_1 P_k(i, n\text{-}1)$, since the number of positions V_{l-i} can have is ${}_nC_1$ and the numbers of ways we can put i identical objects into n-1 positions is $P_k(i, n\text{-}1)$.

Lemma. If $n \leq 4$, $P_k(l, n)$ can be evaluated by **Theorem 9** and **Theorem 10** for all k's.

Proof : According to **Theorem 7**, we can evaluate all $P_k(l, n)$'s if we can evaluate $P_k(l, n)$ for $0 \le l \le n(k\text{-}1)/2$. If $n \le 4$, $l \le 2(k\text{-}1) < 2k$. Thus, we can evaluate all $P_k(l, n)$'s if $n \le 4$.

Theorem 11. If $n \le 4$ and $l \le n(k\text{-}1)/2$, $P_k(l, n) > P_k(l\text{-}1, n)$.

Proof : $l \le 2(k\text{-}1) < 2k$

If $1 \le l < k$,

$$\boldsymbol{P_k(l,n) = {}_{n+l-1}C_l}$$

$$\boldsymbol{P_k(l,n) - P_k(l-1,n) = {}_{n+l-1}C_l - {}_{n+l-2}C_{l-1} = \frac{(n+l-2)!}{(n-2)!\ l!} = {}_{n+l-2}C_l > 0}$$

If $k \le l < 2k$,

$$\boldsymbol{P_k(l,n) = {}_{n+l-1}C_l - {}_nC_1 \sum_{i=0}^{l-k} {}_{n+l-2}C_i}$$

$$\boldsymbol{P_k(l,n) - P_k(l-1,n) = {}_{n+l-2}C_l - n \cdot {}_{n+l-k-2}C_{l-k}}$$

$l \le n(k\text{-}1)/2 \rightarrow 2l/n + 1 \le k \rightarrow 2l/n < l \rightarrow 2 < n$

1) $n = 3$: $\quad 2k \le 2l \le 3(k\text{-}1) \rightarrow 3k\text{-}2l \ge 3,\ k \ge 3,\ l \ge 3$

$$\boldsymbol{P_k(l,3) - P_k(l-1,3) = {}_{l+1}C_l - 3\,{}_{l-k+1}C_{l-k}}$$

$$\boldsymbol{= (l+1) - 3(l-k+1)}$$

$$\boldsymbol{= 3k - 2l - 2 > 0}$$

2) $n = 4$: $\quad l \le 2(k\text{-}1) \rightarrow l\text{-}k+2 \le k$

For $l = k$,

$$\boldsymbol{P_k(l,4) - P_k(l-1,4) = {}_{k+2}C_k - 4\ {}_1C_0 = (k+2)(k+1) - 4 > 0}$$

For $l = k+1$,

$$\boldsymbol{P_k(l,4) - P_k(l-1,4) = {}_{k+3}C_{k+1} - 4\ {}_2C_1 = (k+2)(k+1) - 8 > 0}$$

For $l > k$,

$$\boldsymbol{P_k(l,4) - P_k(l-1,4) = {}_{l+2}C_l - 4\,{}_{l-k+2}C_{l-k}}$$

$$\boldsymbol{= \frac{(l+2)(l+1)}{2} - 2(l-k+2)(l-k+1)}$$

Since this formula decreases monotonically as a function of l, it has its minimum at $l = 2k - 2$.

$$\boldsymbol{P_k(2k-2,4) - P_k(2k-3,4) = {}_{2k}C_2 - 4\ {}_kC_2}$$

$$\boldsymbol{= k(2k-1) - 2k(k-1)}$$

$$\boldsymbol{= k > 0}$$

Lemma. If $n \leq 4$, $P_k(l, n)$ is maximum when $l = \left\lfloor \frac{n(k-1)}{2} \right\rfloor$. Here, $\lfloor x \rfloor$ is the maximum integer $\leq x$.

Proof : According to **Theorem 7** and **Theorem 11**, this is obvious.

Comments on the case of $n > 4$: For $n > 4$, we can obtain an analytical form of $P_k(l,n)$ using recursion and considering repetition. But, the complexity of calculation increases too fast to be practical.

Chapter 4

Verification of LSMP method

4.1 Structure of Simulated Device

In order to verify LSMP operation, TCAD simulation is performed. Fig. 4.1 shows the equivalent circuit diagram of channel stacked NAND flash memory with layer selection by multi-level V_{th}'s of SSTs and SSL biases. For the simple simulation, channel stack NAND array structure with 1 BL is simulated. Fig. 4.2 shows the structure of the simulated device. BL is common for all layers and each SST has charge trapping layer to enable layer selection.

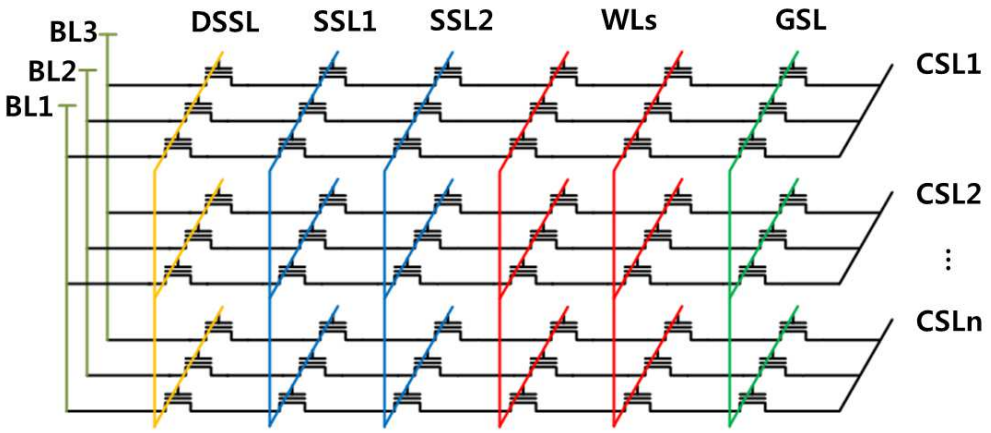


Fig. 4.1. Equivalent circuit diagram of channel stacked NAND flash memory with layer selection by multi-level V_{th}'s of SSTs and SSL biases.

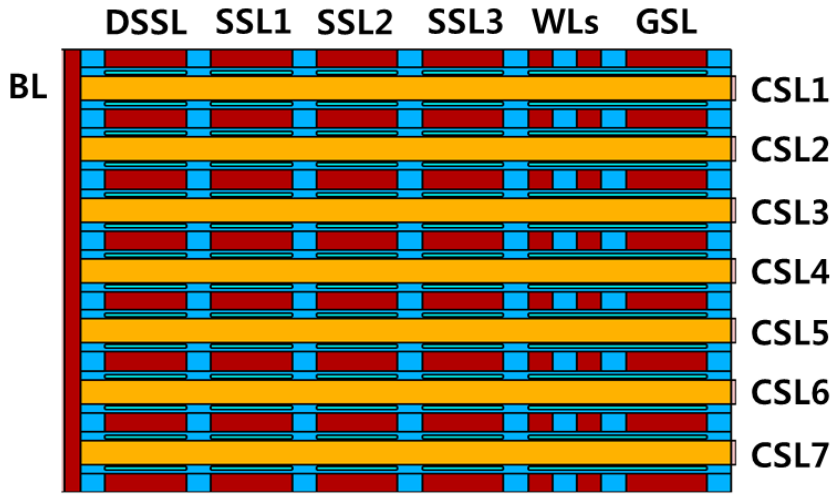


Fig. 4.2. Structure of simulated channel stack NAND array structure with 1 BL.

4.2 Simulation for Verification

7 stacked layers can be decoded when 3 SSLs and 3 V_{th} states are available by LSMP method. The arranged V_{th} states and the corresponding SSL bias voltages for each layer selection are summarized as "V_{th} of SST" and "SSL bias for selection" in Table 4.1. I_d-V_g curves of SSTs which have three different V_{th} states are demonstrated in Fig. 4.3 where V_0 = -1V, V_1 = 1V, and V_2 = 3V. Each V_{th} state is set as the SSL voltage at I_{BL} = 100nA. The SSL bias for stable layer selection, V_{bi}, should make SSTs whose V_{th} state are lower than or equal to V_i be in strong inversion region and should make SSTs whose V_{th} state are larger than V_i be in off-state where $V_i \leq V_{bi} \leq V_{i+1}$. In this simulation, V_{b0} = 0V, V_{b1} = 2V, and V_{b2} = 4V. Fig. 4.4 shows the $|I_{CSL}|$-V_{WL} curves of all 7 layers when the first layer is selected by LSMP method. When SSL bias set for the first layer selection is applied, only the channel of the first layer is connected to BL while the other layers are inhibited. Since each BL is common to all layers, BL is connected to the channel of the selected layer by

the layer selection operation. Thus, in order to verify LSMP method, the current flowing to the common source line, which is separated for each layer, has to be read. Fig. 4.5 shows the contour images of electron concentration of the stacked channels under LSMP bias condition. Each layer can be uniquely selectable. SSL bias set for each layer selection is applied to SSLs as shown in table. Unselected layers have at least one or more turned-off SST(s) and are all inhibited. Only the selected layer has all turned-on SSTs and the layer can be connected to bit-line selectively.

Table 4.1. The arranged V_{th} states of SSTs and SSL bias for each layer selection used in simulation for 7 layers.

Layer #	V_{th} of SST [V]			SSL bias for selection [V]		
	1st SSL	2nd SSL	3rd SSL	1st SSL	2nd SSL	3rd SSL
1	3	1	-1	4	2	0
2	3	-1	1	4	0	2
3	1	3	-1	2	4	0
4	-1	3	1	0	4	2
5	1	-1	3	2	0	4
6	-1	1	3	0	2	4
7	1	1	1	2	2	2

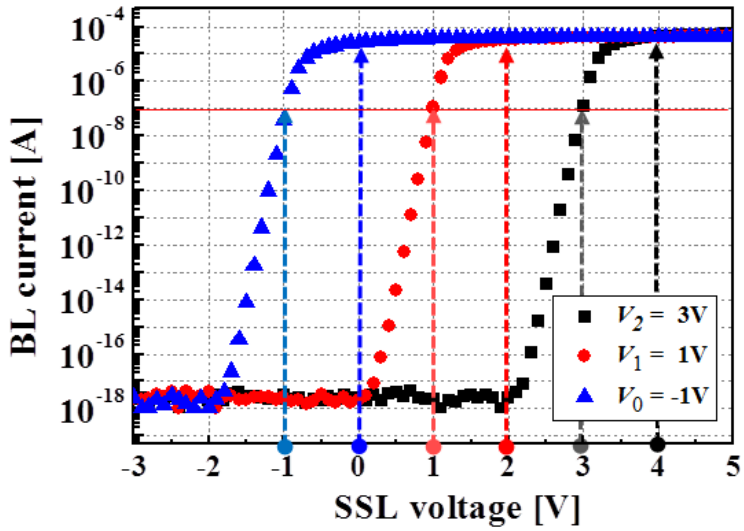


Fig. 4.3. I_d-V_g curves of SSTs which have three different V_{th} states.

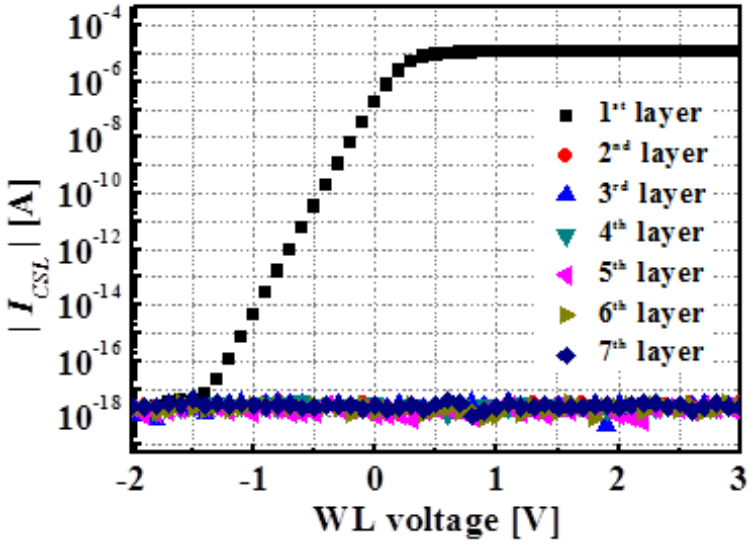
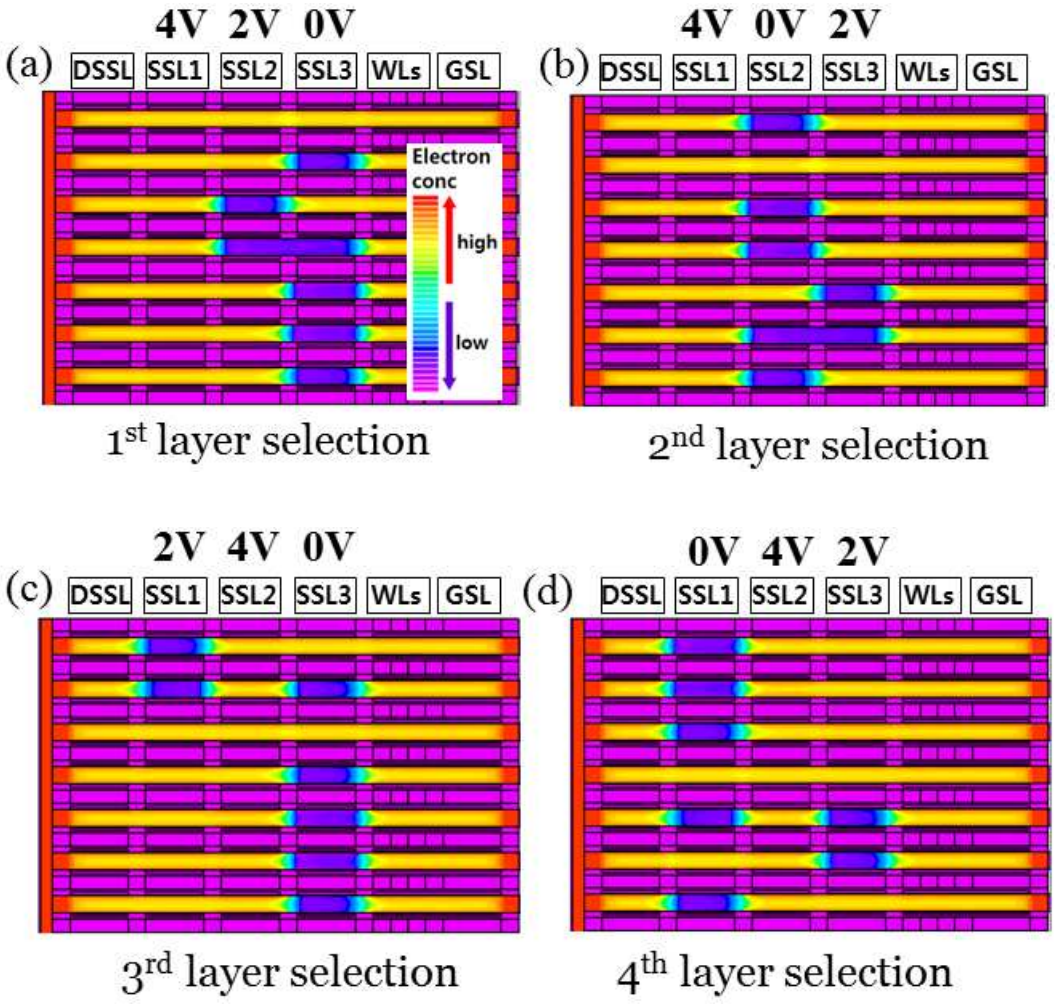


Fig. 4.4. Transfer characteristics of the cells in 7 layers when the first layer is selected by LSMP method where SSL bias set is (4V, 2V, 0V).


4V 2V 0V
(a)
DSSL
SSL1
SSL2
SSL3
WLs
GSL
Electron conc
high
low
1st layer selection
4V 0V 2V
(b)
DSSL
SSL1
SSL2
SSL3
WLs
GSL
2nd layer selection
2V 4V 0V
(c)
DSSL
SSL1
SSL2
SSL3
WLs
GSL
3rd layer selection
0V 4V 2V
(d)
DSSL
SSL1
SSL2
SSL3
WLs
GSL
4th layer selection

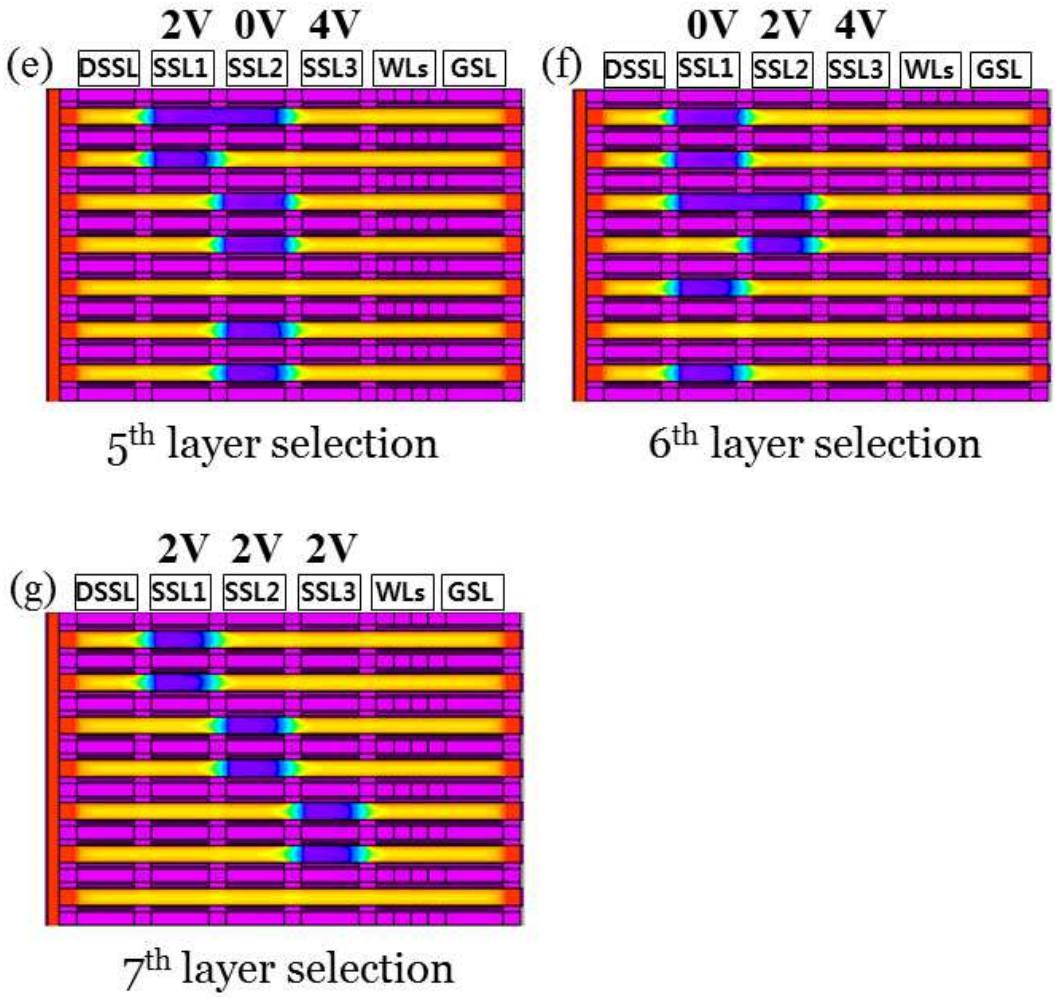


Fig. 4.5. Contour images of electron concentration of the stacked channels under LSMP bias condition. (a) 1st layer selection (b) 2nd layer selection (c) 3rd layer selection (d) 4th layer selection (e) 5th layer selection (f) 6th layer selection (g) 7th layer selection.

4.3 Equivalent Circuit of Measured Device

In order to verify LSMP operation, the gate stack type 3D NAND flash memory fabricated for research purposes by Samsung Electronics is measured. The fabricated memory has macaroni poly-Si body and gate dielectric stacks including SiN charge-trapping layer. Its cell array structure is similar to TCAT [22]. If TCAT structure is rotated by 90°, it can be used for the verification of the LSMP operation as shown in Fig. 4.6.(a) and (b). BLs, CSL, GSL, and SSLs of the TCAT structure can be used as CSLs, BL, DSSL, and GSL of the pseudo-LSM structure with 1 BL. Due to CSL of TCAT is common for all strings, the pseudo-LSM has 1 BL. Every string in a layer is considered as one string and the same GSL bias is applied to the separated SSLs of TCAT in the pseudo-LSM. TCAT structure with 7 BLs is used as the equivalent channel stacked NAND array structure with 1 BL and 7 stacked layers as shown in Fig.4.7. The bias scheme for the verification of the layer selection is also shown.

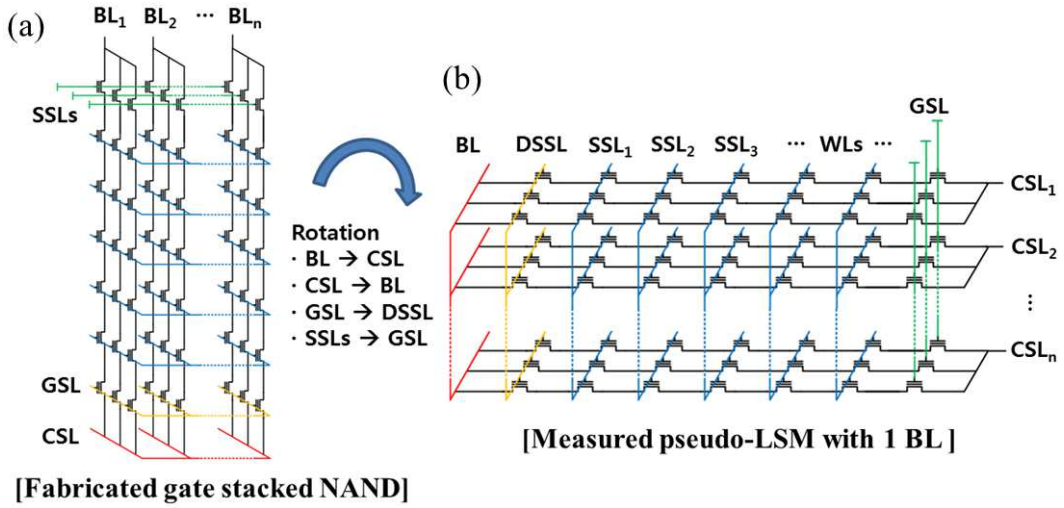


Fig. 4.6. (a) Equivalent circuit of fabricated gate-stacked type NAND flash memory. (b) Equivalent circuit of measured pseudo-LSM.

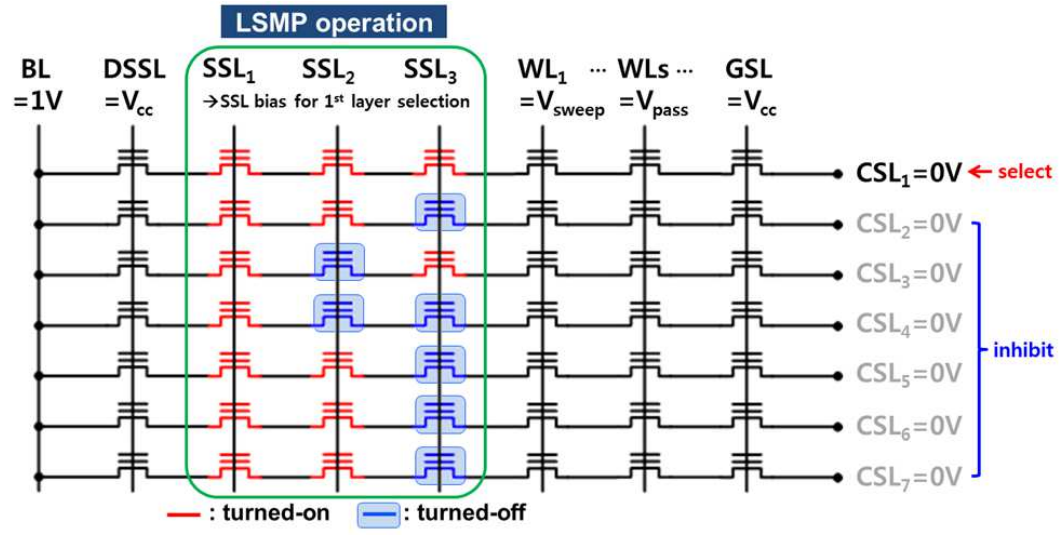


Fig. 4.7. Equivalent circuit of measured device. 1 BL, 3 SSLs, and 7 stacked layers are available.

4.4 Measurement for Verification

The V_{th} states setting of SSTs are carried out to decode 7 stacked layers as shown in Table 4.2. There are two methods for the V_{th} setting of SSTs. One is that all SSTs are initially erased and then programmed one by one to the targeted V_{th} values [38]. The other is that all SSTs on each layer are set to the targeted V_{th} states simultaneously by one erase operation [46], [47]. Both methods can be used for initial V_{th} setting. To verify the LSMP method, initial V_{th} setting of SST is performed by the separated CSL. For the accurate V_{th} setting of SST, incremental step pulse programming (ISPP) is used. This programming scheme is that the gate voltage of program pulse V_{pp} is increased by a constant value $\Delta V_{pp} = V_{step}$ after each program step [49]. Fig. 4.8(a) shows constant V_{th} shift of SST during ISPP. After the initial V_{th} setting, layer selection can be performed by SSL biases. Three V_{th} levels of SSTs used in the device measurement are demonstrated in Fig.4.8(b). Each V_{th} of SST is set as the SSL voltage at I_{BL} = 100nA. The V_{th} states of SSTs and SSL biases should

be determined by considering the available memory windows and subthreshold swings of the SSTs. The use of poly-Si channel may result in narrowing the V_{th} margin between states since poly-Si channel has poor subthreshold behavior, low current drivability, and low on/off current ratio as compared with single-crystalline Si channel. If the LSMP operation is applied to single-crystalline Si channel, the V_{th} margin can be improved and more stable operation can be performed. Fig. 4.9 shows the common source line current ($|I_{CSL}|$) as a function of selected WL voltage under LSMP bias condition for each layer selection. For example, when 7V, 4V, and 1V are applied to the 1st SSL, 2nd SSL, and 3rd SSL respectively, all SSTs of 1st layer which have 6V, 3V, and 0V threshold voltages are turned on and 1st layer is selected due to V_{b2} (7V) > V_2 (6V) > V_{b1} (4V) > V_1 (3V) > V_{b0} (1V) > V_0 (0V). Unselected layers have at least one or more turned-off SST(s) and are all inhibited. Only the selected layer has all turned-on SSTs and is connected to bit-line.

Table 4.2. The arranged V_{th} states of SSTs and SSL bias for each layer selection used in device measurement for 7 layers.

Layer #	V_{th} of SST [V]			SSL bias for selection [V]		
	1st SSL	2nd SSL	3rd SSL	1st SSL	2nd SSL	3rd SSL
1	6	3	0	7	4	1
2	6	0	3	7	1	4
3	3	6	0	4	7	1
4	0	6	3	1	7	4
5	3	0	6	4	1	7
6	0	3	6	1	4	7
7	3	3	3	4	4	4

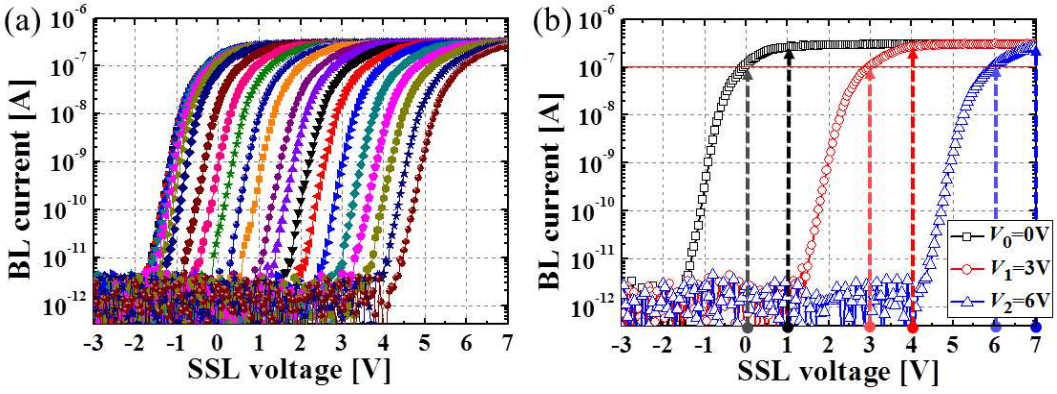
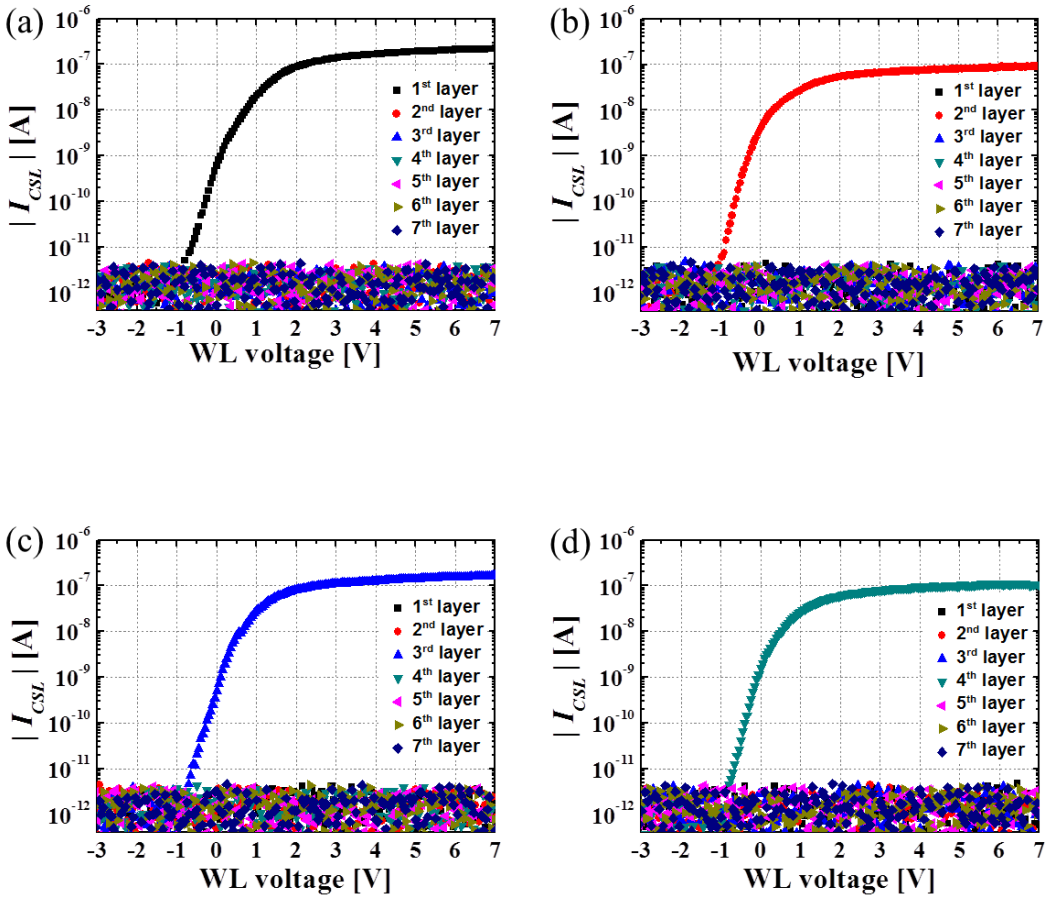


Fig. 4.8. (a) Constant V_{th} shift of SST during incremental step pulse programming. (b) 3 different V_{th} levels of SST of the fabricated NAND flash memory are used to distinguish 7 stacked layers.


(a)
(b)
(c)
(d)
| I_{CSL} | [A]
WL voltage [V]
1st layer
2nd layer
3rd layer
4th layer
5th layer
6th layer
7th layer

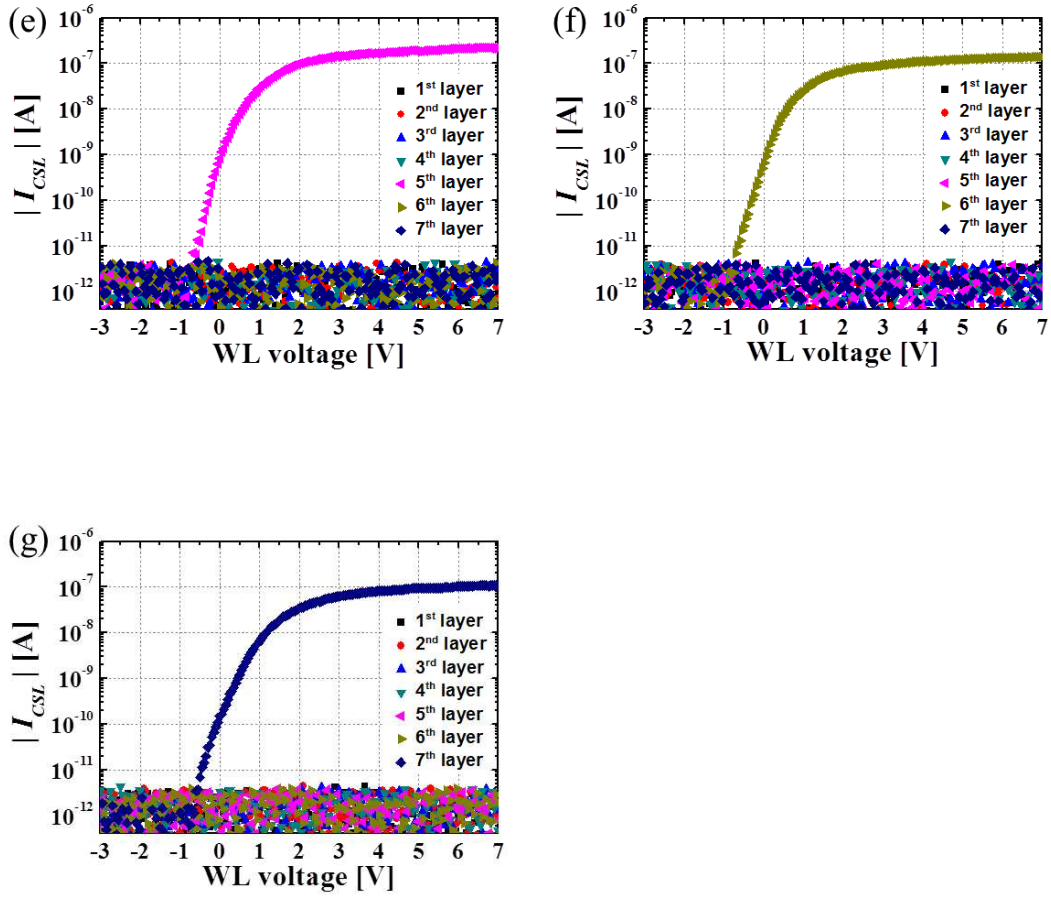


Fig. 4.9. Common source line current ($|I_{CSL}|$) as a function of WL voltage under LSMP bias condition for each layer selection. (a) 1st layer selection (b) 2nd layer selection (c) 3rd layer selection (d) 4th layer selection (e) 5th layer selection (f) 6th layer selection (g) 7th layer selection.

Chapter 5

Application to Gate Stack Type NAND Flash Memory

Channel stack type NAND flash memory has many advantages. In the channel stack type, a single-crystalline Si channel can be used which has better performance with a uniform and stable characteristic by the Si/SiGe epitaxial growth process [33-38]. The read current drivability of the channel stack type is independent of the number of stacks because the channel length is constant. On the other hand, gate stack type has an issue that the read current drivability degrades as the number of stacked layers increases due to an inherent vertical channel structure. Also, channel stack type normally has a

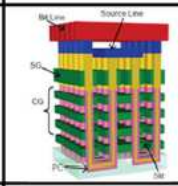
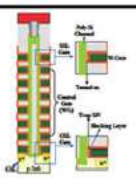
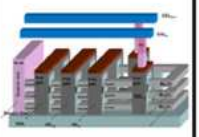
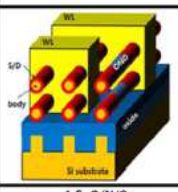
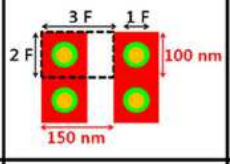
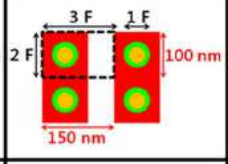
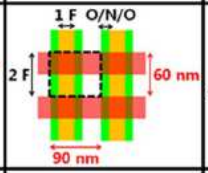
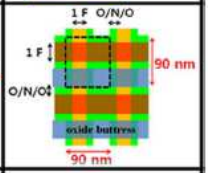
smaller minimal feature size than gate stack type. On the other hand, channel stack type NAND flash memory is difficult to commercialize, and difficult to decode the stacked layer since the bit-line (BL) is placed horizontally and parallel to each other.

The layer selection method by threshold voltage arrangement of string select transistors (SSTs) and string select line (SSL) biases is originally proposed for the channel stack type NAND flash memory. In the case of the conventional CSTAR structure [35], SSLs have to be formed as island-type for the channel selection in a selected layer, because the BL is common in a layer. Thus, BL pitch scaling is difficult due to constraints of lithography and dry etch margin for the island-type SSL formation. Moreover, many contacts are needed for the island-type SSLs and the number of metal layers is increased to form different levels of word-lines (WLs) and SSLs. Those issues can be solved by applying the layer selection method by threshold voltage arrangement of SSTs and SSL biases [27, 37, 38]. Instead of forming island-

type SSLs, this layer selection method can lower the process difficulty and easily achieve BL pitch scaling by forming SSLs like WLs.

On the other hand, in the case of gate stack type NAND flash memory, SSLs can be simply formed on the top of the channel hole to take a role of row selection. In this case, the fabrication is easy, but space is wasted due to the SSL cut. Table 5.1 shows the minimum unit cell size and feature size of various 3D NAND architectures [35]. As shown in Fig. 5.1 [41], at the top of the gate stack type NAND flash memory, SSLs and BLs are perpendicularly crossed each other, and select a row and a column of the channel hole array. SSLs must be separated as the number of rows for decoding and SSL cut spaces are required. Thus, the estimated minimal unit cell size in x, y direction is $6F^2$. By applying LSMP to gate stack type NAND flash memory and forming SSLs like WLs, the SSL cut process can be omitted and minimal unit cell size can be reduced by $4F^2$. The memory density can also be increased accordingly.

Table 5.1. Unit cell size and possible minimal feature size of various 3D NAND architectures [35].

	Gate Stack Type		Channel Stack Type	
	P-BiCS	TCAT	VG NAND	STAR
Structure				
Unit Cell Size				
Device Structure	Macaroni & GAA	Macaroni & GAA	Double Gate	Gate-All-Around
Channel	Macaroni poly-Si channel	Macaroni poly-Si channel	poly-Si channel	Single-Crystalline channel
Limitation of the number of stacks	Low read current	Low read current	No degradation of read current	No degradation of read current

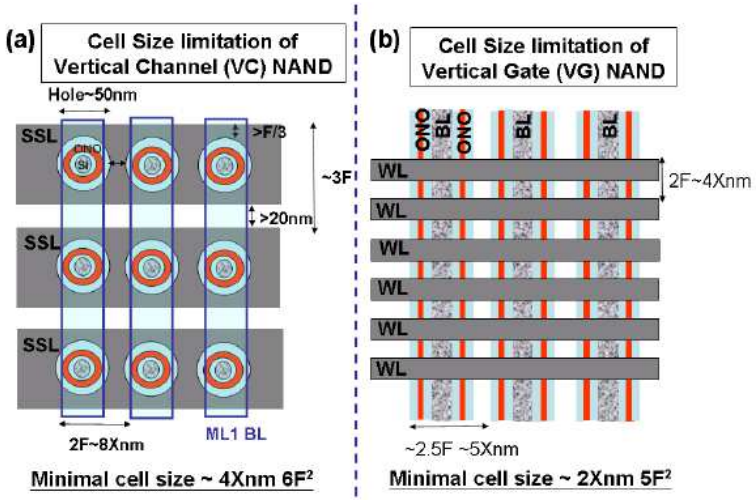


Fig. 5.1. Layout illustration of (a) gate stack type (vertical channel), and (b) channel stack type (vertical gate) NAND [41].

As shown in Fig. 5.2, when the equivalent circuit of CSTAR structure which is channel stack type is rotated 90 degrees, it is the same as the equivalent circuit of TCAT structure which is gate stack type. Likewise, if the equivalent circuit of the LSM structure is rotated 90 degrees and processed as a gate stack type, LSMP method can be applied. In order to apply LSMP method to the gate stack type NAND flash memory, the bottom common source lines must be separately formed. The row of the channel hole array is selected by LSMP operation of SSLs, the column is selected by BLs, and the z-direction layer is selected by WLs.

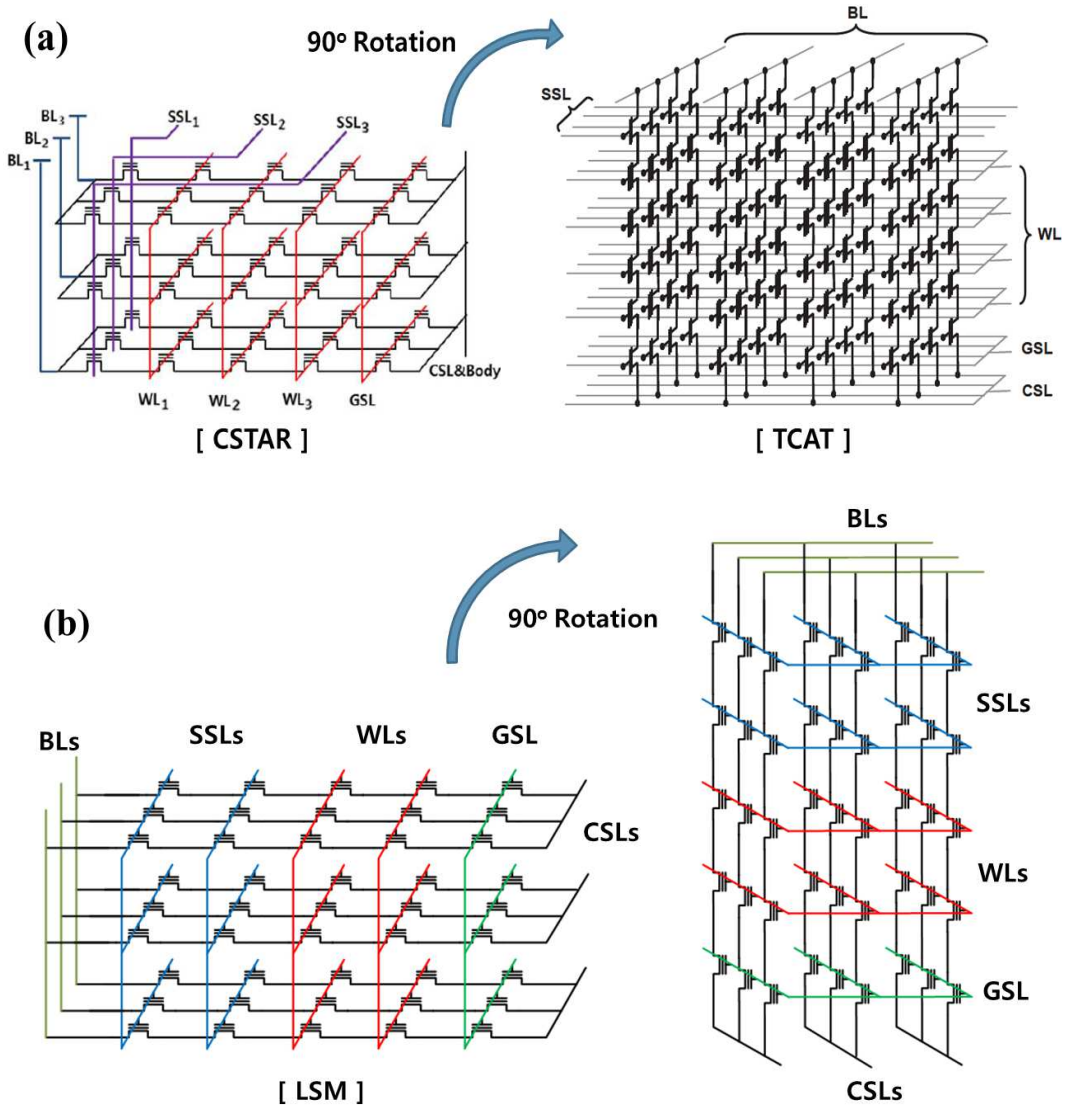


Fig. 5.2. (a) Equivalent circuit of CSTAR rotated 90 degrees is the same as equivalent circuit of TCAT [22, 35]. (b) Equivalent circuit of LSM structure rotated 90 degrees can be processed as gate stack type.

Fig. 5.3 shows the process flow for the gate stack type NAND flash memory with LSMP operation. STI structure and separated line-type bottom common source lines are formed. Control gates to be used for ground select line, word-line, and string select line are deposited. SSLs are stacked like WLs without any special process. Holes for vertical channels are punched through and blocking layer, charge trapping layer, and tunnel oxide films are deposited. The bottom of dielectric films are removed by RIE and poly-Si channel and filler oxide are deposited. Edges of control gates are processed into stair-like structure. And bit-lines are formed and vertical plugs are connected to control gates. The body is omitted because the process difficulty increases to form it. The erase operation uses hole generation by GIDL at dummy string select line and ground select line edge like BiCS [16].

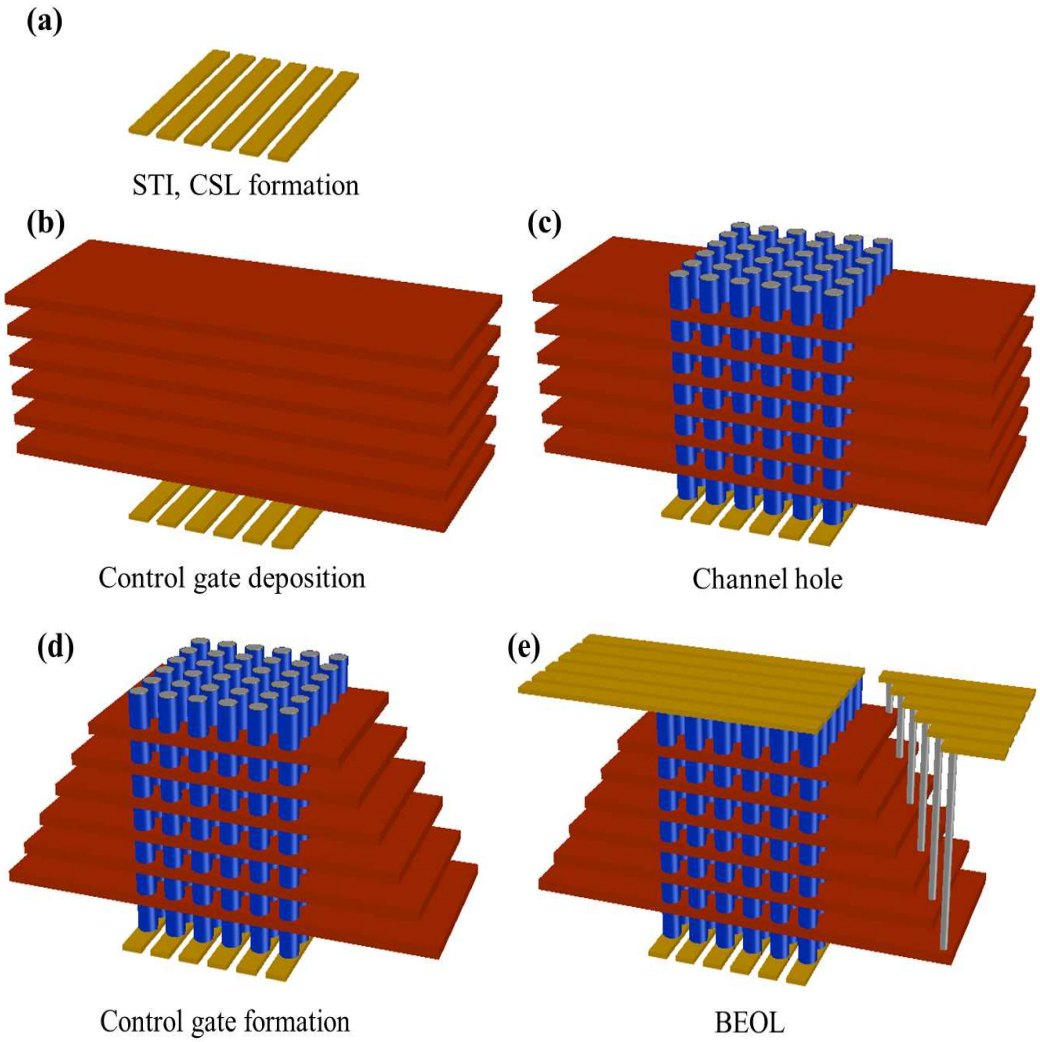


Fig. 5.3. Fabrication sequence of gate stack type NAND flash memory with LSMP method.

Chapter 6

Conclusions

In this dissertation, layer selection by multi-level permutation (LSMP) method for the layer decoding of 3D stacked NAND flash memory is proposed, proved mathematically, and verified by the TCAD simulations and the measurements of the fabricated 3D stacked NAND flash memory.

LSMP method provides the most efficient way to arrange the V_{th}'s of the SSTs for the layer selection as compared to VG-NAND, LASER, and LSM. By using this method, the number of selectable layers can be increased or the minimum required numbers of SSLs and V_{th} states can be reduced. In 3D

stacked NAND flash memory architectures with layer selection methods by V_{th}'s of SSTs and SSL biases, as the number of SSLs increases, the number of WLs decreases in a limited area, which degrades the memory density. Since SSL often has much longer gate length to suppress leakage current, the array overhead (including SSL/GSL, BL pad, contact, and CSL area) can become significant. As the number of V_{th} states of SSTs is increased, V_{th} margin is narrowed due to the limited memory window of SST and the layer selection may become unstable. Therefore, V_{th} arrangement of SSTs should be performed efficiently so that the maximum number of layers can be distinguished by the minimum number of SSLs and V_{th} states. How to arrange V_{th} states of SSTs for the unique selection is permuting grouped V_{th} multisets whose sums of V_{th} state numbers (l) are constant where available V_{th} states (V_i) are considered as $V_0 < V_1 < \ldots < V_{k-1}$ and i is state number. The SSL bias set for the layer selection can be determined by choosing the SSL bias voltage, V_{bi}, for the arranged V_{th} of SST, V_i, such that $V_i < V_{bi} < V_{i+1}$ ($0 \leq i \leq k$-1). The

condition for the maximum number of the selectable layers is that l is equal to the integer which is closest to $n(k - 1)/2$. Furthermore, mathematical proofs for LSMP are made including the unique selection by the equal sum permutations and the recursion formula of the maximum number of selectable layers.

To verify the LSMP method in 3D stacked NAND flash memory, the TCAD simulation and the measurement of the fabricated device are performed. The V_{th} states of SSTs are arranged by LSMP method. When the SSL bias set for layer selection is applied, the transfer characteristic and electron concentration of each stacked layer are monitored and each layer can be uniquely selectable.

Bibliography

[1] J.-D. Lee, S.-H. Hur, and J.-D. Choi, “Effects of Floating-Gate Interference on NAND Flash Memory Cell Operation,” *IEEE Electron Device Letters*, Vol. 23, No. 5, pp. 264-266 May. 2002.

[2] E. Vianello, F. Driussi, A. Arreghini, P. Palestri, D. Esseni, L. Selmi, N. Akil, M. J. van Duuren, and D. S. Golubovi’c, “Experimental and Simulation Analysis of Program/Retention Transients in Silicon Nitride-Based NVM Cells,” *IEEE Transactions on Electron Devices*, Vol. 56, no. 9, pp. 1980-1990, Sep. 2009.

[3] Y.W. Chang, G.W. Wu, P.C. Chen, C.H. Chen, I.C. Yang, C.Y. Chin, I.J. Huang, W.J. Tsai, T.C. Lu, W.P. Lu, K.C. Chen, and C.Y. Lu, “A new interference phenomenon in sub-60nm nitride-based Flash memory,” *Proceedings of NVSMW*, pp. 81-82, 26-30 Aug. 2007.

[4] C.-H. Lee, J.Choi, Y. Park, C. Kang, , “B.-I. Choi, H. Kim, H. Oh, and W.-S. Lee, “Highly scalable NAND Flash memory with robust immunity to program disturbance using symmetric inversion-type source and drain structure,” *Symposium on VLSI technology*, pp. 118-119, 17-19 June 2008.

[5] E.S. Cho, H.J. Kim, B.T. Kim, J.H. Song, J.H. Song, D.H. Song, J.-H. Choi, KD. Suh, and C. Chung, “Optimal Cell Design for Enhancing Reliability Characteristics for sub 30 nm NAND Flash Memory,” in *IEEE IRPS Tech*. pp.611-614, May. 2010.

[6] J.N. Wileu, T.-Y. Fu, T. Tanaka, S. Takeuchi, S. Aoyama, J. Miyazaki, and Y. Watakabe, “Phase-shift mask pattern accuracy requirements and inspection technology,” *Proc. SPIE 1464, Integrated Circuit Metrology, Inspection, and Process Control V,* 346, Jul. 1991.

[7] M. Garza, J.V. Jensen, N.K. Eib, and K.K. Chao., “Optical Proximity Correction Method and Apparatus,” U. S. Patent No. 5,723,233, Mar. 3, 1998.

[8] Soichi Owa, Hiroyuki Nagasaka, “Immersion lithography; its potential performance and issues,” *Proc. SPIE 5040, Optical Microlithography XVI, 724,* Jun. 2003.

[9] C.-M. Lim, S.-M. Kim, Y.-S. Hwang, J-S. Choi, K.-D. Ban, S.-Y. Cho, J.-K. Jung, E.-K. Kang, H.-Y. Lim, H.-S. Kim, and S.-C. Moon, “Positive and negative tone double patterning lithography for 50 nm flash memory.” *Proc. SPIE 6154, Optical Microlithography XIX, 615410,* Mar. 2006.

[10] Yohwan Koh, “NAND Flash Scaling Beyond 20 nm,” *Memory Workshop, 2009. IMW’09. IEEE International,* pp.1-3, 10-14 May 2009.

[11] J. E. Brewer, M. Gill, “Nonvolatile Memory Technologies with Emphasis on Flash,” *IEEE Press,* pp. 2, 2008.

[12] Kirk. Prall, “Scaling Non-Volatile Memory Below 30 nm,” *NVSMW Tech Dig.*, pp.5-10, Aug. 2007.

[13] S.-M. Jung, J. Jang, W. Cho, H. Cho, J. Jeong, Y. Chang, J. Kim, Y. Rah, Y. Son, J. Park, M.-S. Song, K.-H. Kim, J.-S. Lim, and K. Kim, “Three dimensionally stacked NAND Flash memory technology using stacking single crystal Si layers on ILD and TANOS structure for beyond 30 nm node,” *IEDM tech. digest,* pp. 1-4, 11-13 Dec. 2006.

[14] E.-K. Lai, H.-T. Lue, Y.-H. Hsiao, J.-Y. Hsieh, C.-P. Lu, S.-Y. Wang, L.-W. Yang, T. Yang, K.-C. Chen, J. Gong, K.-Y. Hsieh, R. Liu, and C.-Y. Lu, “A Multi-Layer Stackable Thin-Film Transistor (TFT) NAND-Type Flash Memory,” *IEEE International Electron Devices Meeting*, pp. 1-4, 11-13 Dec. 2006.

[15] T. Endoh, K. Kinoshita, T. Tanigami, Y. Wada, K. Sato, K. Yamada, T. Yokoyama, N. Takeuchi, K. Tanaka, N. Awaya, K. Sakiyama, and F. Masuoka, “Novel ultrahigh-density flash memory with a stacked-surrounding gate transistor (S-SGT) structured cell,” *IEEE Transactions on Electron Devices*, vol. 50, no. 4, pp. 945-961, April 2003.

[16] H. Tanaka, M. Kido, K. Yahashi, M. Oomura, R. Katsumata, M. Kido, Y. Fukuzumi, M. Sato, Y. Nagata, Y. Matsuoka, Y. Iwata, H. Aochi, and A. Nitayama, “Bit Cost Scalable technology with punch and plug process for ultra high density Flash memory,” *Symposium on VLSI technology*, pp. 14-15, 2007.

[17] Y. Fukuzumi, R. Katsumata, M. Kito, M. Kido, M. Sato, H. Tanaka, Y. Nagata, Y. Matsuoka, Y. Iwata, H. Aochi and A. Nitayama, "Optimal integration and characteristics of vertical array devices for ultra-high density, bit-cost scalable flash memory." in *IEDM Tech. Dig.*, 2007, pp.449-452.

[18] Y. Komori, M. Kido, M. Kito, R. Katsumata, Y. Fukuzumi, H. Tanaka, Y. Nagata, M. Ishiduki, H. Aochi, and A. Nitayama, “Disturbless flash memory due to high boost efficiency on BiCS structure and optimal memory film stack for ultra high density storage device,” in *IEDM Tech. Dig.*, pp. 1-4, Dec. 15~17, 2008.

[19] A. Nitayama, and H. Aochi, "Bit Cost Scalable (BiCS) flash technology for future ultra high density storage devices," in *VLSI Symp. Tech. Dig.*, 2010, pp.130-131.

[20] R. Katsumata, M. Kito, Y. Fukuzumi, M. Kido, H. Tanaka, Y. Komori, M. Ishiduki, J. Matsunami, T. Fujiwara, Y. Nagata, L. Zhang, Y. Iwata, R. Kirisawa, H. Aochi, and A. Nitayama, “Pipe-shaped BiCS flash memory with 16 stacked layers and multi-level-cell operation for ultra high density storage devices,” *Symposium on VLSI Technology*, pp. 136-137, 16-18 June 2009.

[21] M. Ishiduki, Y. Fukuzumi, R. Katsumata, M. Kito, M. Kido, H. Tanaka, Y. Komori, Y. Nagata, T. Fujiwara, T. Maeda, Y. Mikajiri, S. Oota, M. Honda,Y. Iwata, R. Kirisawa, H. Aochi and A. Nitayama, "Optimal device structure for pipe-shaped BiCS flash memory for ultra high density storage device with excellent performance and reliability," in *IEDM Tech. Dig.*, 2009, pp.625-628.

[22] J. Jang, H.-S. Kim, W. Cho, H. Cho, J. Kim, S.I. Shim, Younggoan, J.-H. Jeong, B.-K. Son, D.W. Kim, Kihyun, J.-J. Shim, J.S. Lim, K.-H. Kim, S.Y. Yi, J.-Y. Lim, D. Chung, H.-C. Moon, S. Hwang, J.-W. Lee, Y.-H. Son, U.-I. Chung, and W.-S. Lee, “Vertical cell array using TCAT (Terabit Cell Array Transistors) technology for ultra high density NAND Flash memory,” *Symposium on VLSI technology*, pp. 192-193, 2009.

[23] J. Kim, A. J. Hong, M. Ogawa, S. Ma, E. B. Song, Y.-S. Lin, J. Han, U.-I. Chung and K. L .Wang, "Novel 3-D structure for ultra high density flash memory with VRAT (Vertical-Recess-Array-Transistor) and PIPE (Planarized Integration on the same PlanE)," in *VLSI Symp.*

Tech. Dig., 2008, pp.122-123.

[24] J. Kim, A. J. Hong, S. M. Kim, E. B. Song, J. H. Park, J. Han, S. Choi, D. Jang, J.T. Moon, and K. L .Wang, “Novel Vertical-Stacked-Array - Transistor (VSAT) for ultra-high-density and cost-effective NAND Flash memory devices and SSD (Solid State Drive),” in *VLSI Symp. Tech. Dig.*, 2009, pp.186-187.

[25] S. Whang, K. Lee, D. Shin, B. Kim, M. Kim, J. Bin, J. Han, S. Kim, B. Lee, Y. Jung, S. Cho, C. Shin, H. Yoo, S. Choi, K. Hong, S. Aritome, S. Park, and S. Hong, “Novel 3-Dimensional Dual Control-Gate with Surrounding Floating-Gate (DC-SF) NAND Flash Cell for 1 Tb File Storage Application,” in *IEDM Tech. Dig.*, pp. 29.7~4, Dec. 6~8, 2010.

[26] M.-S. Seo, B.-H. Lee, S.-K. Park, and T. Endoh, "Novel Concept of the Three-Dimensional Vertical FG nand Flash Memory Using the Separated-Sidewall Control Gate," *IEEE Transactions on Electron Devices*, Vol.59, No.8, pp. 2078-2084, Aug. 2012.

[27] W. Kim, S. Choi, J. Sung, T. Lee, C. Park, H. Ko, J. Jung, I. Yoo, and Y. Park, “Multi-layered vertical gate NAND flash overcoming stacking limit for terabit density storage,” in *VLSI Symp. Tech. Dig.*, pp. 188-189, 2009.

[28] H.-T. Lue, T.-H. Hsu, Y.-H. H, S.P. Hong, M.T. Wu, F.H. Hsu, N.Z. Lien, S.-Y. Wang, J.-Y. Hsieh, L.-W. Yang, T. Yang, K.-C. Chen, K.-Y. Hsieh, and C.-Y. Lu, “A highly scalable 8-layer 3D vertical-gate (VG) TFT NAND Flash using junction-free buried channel BE-SONOS device,” in *VLSI Symp. Tech. Dig.*, pp. 131-132, 2010.

[29] C.H. Hung, H.T. Lue, K.P. Chang, C.P. Chen, Y.H. Hsiao, S.H. Chen,

Y.H. Shih, K.Y. Hsieh, M. Yang, J. Lee, S.Y. Wang, T. Yang, K.C. Chen, and C.Y. Lu, “A Highly Scalable Vertical Gate (VG) 3D NAND Flash with Robust Program Disturb Immunity Using a Novel PN Diode Decoding Structure,” in *VLSI Symp. Tech. Dig.*, 2011, pp.68-69.

[30] C.P. Chen, H.T. Lue, K.P. Chang, Y.H. Hsiao, C.C. Hsieh, S.H. Chen, Y.H. Shih, K.Y. Hsieh, T. Yang, K.C. Chen, and C.Y. Lu, “A Highly Pitch Scalable 3D Vertical Gate (VG) NAND Flash Decoded by a Novel Self-Aligned Independently Controlled Double Gate (IDG) String Select Transistor (SSL),” in *VLSI Symp. Tech. Dig.*, 2012, pp. 91-92.

[31] K.P. Chang, H.T. Lue, C.P. Chen, C.F. Chen, Y.R. Chen, Y.H. Hsiao, C.C. Hsieh, S.H. Chen, Y.H. Shih, T. Yang, K.C. Chen, K.Y. Hsieh, C.H. Hung, and C.Y. Lu, “An Efficient Memory Architecture for 3D Vertical Gate (3DVG) NAND Flash Using Plural Island-Gate SSL Decoding and Study of Its Program Inhibit Characteristics”, *IEEE International Memory Workshop*, 2012, pp. 25-28.

[32] H.-T. Lue, T.-H. Hsu, C.-J. Wu, W.-C. Chen, T.-H. Yeh, K.-P. Chang, C.-C. Hsieh, P.-Y. Du, Y.-H. Hsiao, Y.-W. Jiang, G.-R. Lee, R. Lo, Y.-R. Su, C. Huang, S.-C. Lai, L.-Y. Liang, C.-F. Chen, M.-F. Hung, C.-W. Hu, C.-J. Chiu, and C.-Y. Lu, "A novel double-density, single-gate vertical channel (SGVC) 3D NAND Flash that is tolerant to deep vertical etching CD variation and possesses robust read-disturb immunity." in *IEDM Tech. Dig.*, 2015, pp.44-47.

[33] A. Hubert, E. Nowak, K. Tachi, V. Maffini-Alvaro, C. Vizioz, C. Arvet, J.-P. Colonna, J.-M. Hartmann, V. Loup, L. Baud, S. Pauliac, V. Delaye, C. Carabasse, G. Molas, G. Ghibaudo, B. De Salvo, O.

Faynot and T. Ernst, “A stacked SONOS technology, up to 4 levels and 6 nm crystalline nanowires, with gate-all-around or independent gates (Φ-Flash), suitable for full 3-D integration,” in *IEDM Tech. Dig.*, 2009, pp. 637-640.

[34] J.-G. Yun, G. Kim, J.-E. Lee, Y. Kim, W.B. Shim, J.-H. Lee, H. Shin, B.-G. Park, and J.D. Lee, “Single-Crystalline Si STacked ARray (STAR) NAND Flash Memory,” *IEEE Transactions on Electron Devices*, vol. 58, no. 4, pp. 1006-1014, Mar. 2011.

[35] Y. Kim, J.-G. Yun, S.H. Park, W. Kim, J.Y. Seo, M. Kang, K.-C. Ryoo, J.-H. Oh, J.-H. Lee, H. Shin, and B.-G. Park, “Three-Dimensional NAND Flash Architecture Design Based on Single-Crystalline STacked ARray,” *IEEE Transactions on Electron Devices*, vol. 59, no. 1, pp. 34-45, Jan. 2012.

[36] S. H. Park, Y. Kim, W. Kim, J. Y. Seo, H. Kim, and B.-G. Park, "Novel three dimensional (3D) NAND flash memory array having tied bit-line and ground select transistor (TiGer)," *IEICE Transactions on Electronics*, Vol. E95-C, No. 5, pp. 837-841, May. 2012

[37] J.-G. Yun, S.H. Park, and B.-G. Park, “LAyer Selection by ERase (LASER) With and Etch-Through-Spacer Technique in a Bit-Line Stacked 3-D NAND Flash Memory Array,” *IEEE Transactions on Electron Devices*, vol. 58, no. 7, pp. 1892-1897, Jul. 2011.

[38] W. Kim, J.Y. Seo, Y. Kim, S.H. Park, S.H. Lee, M.H. Baek, J.-H. Lee, and B.-G. Park, “Channel-stacked NAND flash memory with layer selection by multi-level operation (LSM),” in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Dec. 2013, pp. 3.8.1–3.8.4.

[39] Masahide Kimura, “3D cells make terabit NAND flash possible,”

Nikkei Electronics Asia, Sept 17, 2009.

[40] C. H. Lee, K. I. Choi, M. K. Cho, Y. H. Song, K. C. Park, and K. Kim, “A Novel SONOS Structure of SiO2/SiN/Al2O3 with TaN Metal Gate gor Multi-Giga Bit Flash Memories,” in *IEDM Tech. Dig.*, 2003, pp.26.5.1-26.5.4.

[41] H.-T. Lue, S.-H. Chen, Y.-H. Shih, K.-Y. Hsieh, and C.-Y. Lu, “Overview of 3D NAND Flash and Progress of Vertical Gate (VG) Architecture,” *in ICSICT Tech. Dig.*, Oct. 2012, pp. 1-4.

[42] H. Ikeda, “Characterization of switching transient behaviors in polycrystalline-silicon thin-film transistors,” *Jpn. J. Appl. Phys.*, Vol. 43, No. 2, pp. 477-484, Feb. 2004.

[43] B. Kim, S.-H. Lim, D. W. Kim, T. Nakanishi, S. Yang, J.-Y. Ahn, H. Choi, K. Hwang, Y. Ko, and C.-J. Kang, “Investigation of Ultra Thin Polycrystalline Silicon Channel for Vertical NAND Flash,” in *Proc. IEEE Int. Rel. Phys. Symp.*, pp. 126–129, Apr. 2011.

[44] H.-J. Kang, M.-K. Jeong, S.-M. Joe, J.-H. Seo, S.-K. Park, S. H. Jin, B.-G. Park and J.-H. Lee, “Effect of Traps on Transient Bit-line Current Behavior in Word-line Stacked NAND Flash Memory with Poly-Si Body,” in *VLSI Symp. Tech. Dig.*, 2014, pp. 28–29.

[45] Wen-Jer Tsai, W.L. Lin, C.C. Cheng, S.H. Ku, Y.L. Chou, Lenvis Liu, S.W. Hwang, T.C. Lu, K.C. Chen, Tahui Wang, and Chih-Yuan Lu, “Polycrystalline-Silicon Channel Trap Induced Transient Read Instability in a 3D NAND Flash Cell String,” in *IEDM Tech. Dig.*, 2016, pp. 288-291.

[46] D.W. Kwon, W. Kim, D.-B. Kim, S.-H. Lee, J. Y. Seo, M. H. Baek, J.-H. Park, E. Choi, G. S. Cho, S.-K. Park, and B.-G. Park, "Multi-Level

Threshold Voltage Setting Method of String Select Transistors for Layer Selection in Channel Stacked NAND Flash Memory." *IEEE Electron Device Letters*, Vol.36, No.12, pp. 1318-1320, Oct. 2015.

[47] D.W. Kwon, W. Kim, D.-B. Kim, S.-H. Lee, J. Y. Seo, M. H. Baek, J.-H. Park, E. Choi, G. S. Cho, S.-K. Park, J.-H. Lee, and B.-G. Park, "Analysis on Program Disturbance in Channel-Stacked NAND Flash Memory with Layer Selection by Multilevel Operation." *IEEE Transactions on Electron Devices*, Vol.63, No.3, pp. 1041-1046, Mar. 2016.

[48] W. Kim, “Channel-stacked NAND flash memory with layer selection by multi-level operation,” Ph.D. Dissertation, Seoul National University, Aug. 2013.

[49] K.-D. Suh B.-H. Suh, Y.-H. Lim, J.-K. Kim, Y.-J. Choi, Y.-N. Koh, S.S. Lee, S.-C. Kwon, B.-S. Choi, J.-S. Yum, J.-H. Choi, J.-R. Kim, H.-K. Lim, ”A 3.3 V 32 Mb NAND flash memory with incremental step pulse programming scheme,” in *ISSCC Tech. Dig.*, Feb. 1995, pp. 128-129.

초 록

본 논문에서는 3차원 적층 낸드 플래시 메모리에서의 멀티 레벨 순열을 이용한 층 선택 방법을 제안하고 수학적으로 증명하였으며, TCAD 시뮬레이션과 실제 공정된 3차원 적층 낸드 플래시 메모리소자에 적용하여 동작을 검증 하였다.

제안된 방법은 기존에 제안되었던 방법들과 비교했을 때 가장 효율적으로 스트링 셀렉트 트렌지스터의 문턱전압을 배열하는 방식을 제공한다. 이 방법을 통하여, 같은 수의 스트링 셀렉트 라인과 문턱전압상태 조건에서는 가장 많은 층을 선택할 수 있고, 적층된 수가 주어졌을 때는 가장 적은 수의 스트링 셀렉트 라인과 문턱전압상태로 모든 층을 선택할 수 있다. 스트링 셀렉트 트랜지스터의 문턱전압배열과 스트링 셀렉트 라인의 바이어스로 층 선택 동작을 하는 3차원 적층 낸드 플래시 메모리 구조에서는 스트링 셀렉트 라인의 수가 늘어날수록, 워드라인을 배치할 공간이 줄어들기 때문에 메모리 밀도가 열화 된다. 또한, 스트링 셀렉트 라인은 보통 누설전류를 억제하기 위하여 게이트 길이가 길기 때문에 어레이 오버헤드가 심각한 문제로 대두될 수 있다. 한편, 스트링 셀렉트 트렌지스터의 문턱전압상태가 많아지게 되면 메모리 윈도우가

한정되어있기 때문에 문턱전압 마진이 줄어들게 되고 결국 층 선택 동작이 불안정해질 수 있다. 그러므로 , 스트링 셀렉트 트렌지스터의 문턱전압 배열은 적은 수의 스트링 셀렉트 라인과 문턱전압상태를 쓰고도 최대한 많은 층을 선택할 수 있도록 효율적으로 시행되어야 한다. 본 논문에서는 유일 선택을 할 수 있는 문턱전압의 배열조건과 그 중에서 최대의 층 선택이 가능한 조건을 찾아 배열방법을 제안하고 수학적으로 증명하였다. 그리고 이를 확인하기 위해 TCAD 시뮬레이션과 제작된 소자의 측정으로 기존에 제안되었던 방법들 보다 같은 조건에서 더 많은 층의 선택이 가능함을 검증 하였다.

주요어 : 3차원 적층 낸드 플래시 메모리, 낸드 플래시 메모리, 멀티 레벨 순열에 의한 층 선택, 멀티 레벨 동작을 통한 층 선택, 문턱 전압에 의한 선택 방법.

학번 : 2012-20836