NAND flash memory technology utilizing fringing electric field

Jong-Ho Leea,⇑, Sang-Goo Jungb

aSchool of Electrical Engineering and Inter-University Semiconductor Research Center (ISRC), Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742, Republic of Korea
bSchool of Electrical Engineering and Computer Science, Kyungpook National University, Sankyuk-dong, Buk-gu, Daegu 702-701, Republic of Korea

ABSTRACT

In this review article, basic properties of NAND flash memory cell strings which consist of cells with virtual source/drain (S/D) (or without S/D) were discussed. The virtual S/D concept has advantages of better scalability, less cell fluctuation due to effectively longer channel length at the same technology node, and less program disturbance. The fringing electric field from the control-gate and/or the floating-gate is essential to induce the virtual S/D (charges) in the space region of the body between control-gates and becomes effective as cell size shrinks. A cell string consisting of planar channel silicon-oxide-nitride-oxide-silicon (SONOS) cells formed in bulk Si substrate needs to have a bit-line body doping of \( \frac{10^{17}}{cm^2} \) in the channel and a less doping in the space region to keep high bit-line read current. The floating gate (FG) flash memory cell string gives larger bit-line current compared to that of SONOS flash memory cell string at given similar body doping. Non-planar channel cells like arch and fin-type body structures were more effective to focus the fringing electric field on the space region. The virtual S/D concept is also useful in 3-dimensional (3-D) stacked NAND flash memory where thin film (or nanowire, nanotube) body is adopted.

1. Introduction

As IT technology including mobile devices advances, the flash memory market has shown growth at an extremely fast rate. The memory market needs continuously a mass storage device with high bit density. Traditionally, many researchers have studied non-volatile memory technology to achieve high-density [1–3]. NAND flash memory has been already known the best solution to meet high bit density at low cost until now. In shrinking conventional NAND flash memory cell with poly-Si floating-gate (FG) to 2X generation and beyond, reliability issues regarding threshold voltage \( V_{th} \) fluctuation due to random telegraph noise (RTN) [4] and disturbance during program and read, cycling [5], retention, cross-talk and short channel effect (SCE) [6,7] become serious. Alternative cell devices like fin-type silicon-oxide-nitride-oxide-silicon (SONOS) field-effect-transistor (FET) [5], planar channel (PC) SONOS device [8], nano-floating-gate-memory (NFGM) device [9,10], have been investigated and shown better scalability. But there were also reliability issues in these alternative devices when scaled down. Although NAND flash memory has still some issues, it has been widely adopted in successively created mobile devices due to its high bit density and continuous cost reduction through scaling-down of cell size and adoption of multi-level cell (MLC). At present, planar channel NAND flash memory with the FG structure is believed to a key technology for a mass storage device until around 2013 [11]. After that, the planar channel NAND flash memory would be saturated in cell size and 3-dimensional (3-D) stack structure can be very promising for mass production [11]. Thus, NAND flash memory will be very promising technology for future...
mass storage device. If we can provide an approach to improve the reliability and decrease $V_{th}$ fluctuation of NAND flash memory cells with the scaling-down of cell size, it will be very helpful.

We have introduced, for the first time, virtual source/drain (S/D) or without S/D for NAND cell string [12] to improve reliability and suppress short channel effect and $V_{th}$ fluctuation due to RTN. Here, the virtual S/D is sometimes called as junction-free [13–15] or inversion S/D [16], and means that there are no neural source/drain in the space between adjacent control-gates (CG) of a NAND cell string at equilibrium and charges between adjacent cells are induced by the fringing electric field from the side of the CGs whenever needed. This virtual S/D concept can be applied to various NAND cell strings which include cells having planar or non-planar (or 3-D) channel structure [13–21], and also even vertical NAND cell string in 3-D stacked NAND technology [22–27]. The virtual S/D concept has been demonstrated experimentally in planar channel NAND flash memory and reported very good performance [13–16]. In this work, we review and investigate the fringing field effect on the virtual S/D in NAND flash memory technologies which have planar channel structure, 3-D channel structure, and 3-D stacked (or vertical) structure. We investigate each structure of the memory in terms of a cell string, since the basic building block of the NAND flash memory is not a unit cell but a cell string which consists of 32 or 64 cells and selection devices.

2. NAND flash memory cells with planar channel structure

The planar channel structure of a NAND flash memory cell has a poor short-channel-effect (SCE) with scaling-down, but more importantly smaller cells size (-4F²) compared to that of 3-D channel structures. Therefore the planar channel structure is very attractive to achieve high-density memory capacity. To check the fringing electric effect in the NAND flash memory cells with the planar channel structure, we investigate NAND cell strings which consist of cells with two different charge storage nodes: nitride [8,13] and poly-Si FG [16].

We first go over planar channel SONOS flash memory in which cells in a string cell string have a layer of nitride as a storage node. Key advantages of SONOS cells are process compatibility with conventional complementary-metal-oxide-semiconductor (CMOS) process and smaller discharge of charges stored in a dielectric storage node than that of FG cells.

Fig. 1a shows a schematic cross-sectional view of planar channel NAND flash memory cell string which has 3 cells and two selection devices at both ends as an example. The selection devices are denoted by ground selection line (GSL) and string selection line (SSL). There is no S/D region in the space region between adjacent cells in the cell string. Neutral S/D regions are formed only in both ends of the cell string (only near the selection devices) in a non-overlapped manner in this example, but can be formed in an overlapped manner [12].

In Fig. 1a, the thicknesses of the tunneling oxide, nitride storage layer, and blocking oxide are 4 nm, 5 nm, and 6 nm, respectively, $L_g$ and $L_e$ represent a gate length and a length of space between adjacent cells, respectively, and generally $L_g = L_e$. $R_{pol}$, $N_{peak}$, and $N_{peak}$'s represent the projected range, peak concentration of S/D and the peak concentration of boron, respectively, when nMOS type cells are assumed. In Fig. 1b, the cell devices have non-overlapped S/D. Here the S/D regions can be fully depleted depending on the doping concentration, which leads to a virtual S/D cell string. $W_{SD}$ and $x_i$ stand for S/D width and S/D junction depth, respectively. If the $W_{SD}$ is 0 nm, then there is no S/D in a cell string. $R_{SD}$ represents the peak doping position of the S/D regions.

As $L_g$ ($\leq L_e$) decreases, the charge induction by the fringing electric field from the CGs becomes easy in the space region, which means $V_{th}$ of a cell decreases because the $V_{th}$ in the space region doped with a p-type impurity is higher than that of the channel and short channel effect becomes large. We first check drain-induced-barrier-lowering (DIBL) and subthreshold-swing (SS) of a cell in a cell string by changing $W_{SD}$ at a fixed $L_g = L_e = 30$ nm. Here, $x_i$ is 20 nm and $N_{peak}$ is a variable.

Fig. 2 shows the DIBL and SS behavior of the cell with $W_{SD}$. As mentioned above, there is no S/D when $W_{SD}$ is 0 nm. When $N_{peak}$ is less than $1 \times 10^{18}$ cm$^{-3}$ at $W_{SD} = 0$ nm, DIBL and SS show the smallest values for each $N_{peak}$. The S/D is overlapped with the CG for a $W_{SD}$ wider than 30 nm and non-overlapped for a $W_{SD}$ narrower than 30 nm. As $W_{SD}$ increases, $V_{th}$ decreases for various $N_{peak}$'s due to decreasing $V_{th}$ in the non-overlapped region. As shown in this figure, devices with the non-overlap structure show lower DIBL than that of the devices with overlapped structure. With increasing $W_{SD}$, SS's for $N_{peak}$'s lower than $2 \times 10^{18}$ cm$^{-3}$ increase significantly due to SCE, but the SS at $2 \times 10^{18}$ cm$^{-3}$ decreases due to $V_{th}$ reduction in the space region by the fringing field. The $V_{th}$ decrease in the space region means better channel...
controllability by the CG, resulting in SS reduction. Although not shown as a figure, DIBLs and SSs of the cell having S/D were relatively immune to the change of $x_j$ (0–20 nm). For $N_{\text{peak}}$ lower than $1 \times 10^{18}$ cm$^{-3}$, the $W_{\text{S/D}}$ needs to be small as possible if needed at a given $L_g = L_s = 30$ nm. If the S/D just meet the gate ($W_{\text{S/D}} = 30$ nm) or is overlapped with the gate ($W_{\text{S/D}} > 30$ nm), an $N_{\text{peak}}$ of $\sim 2 \times 10^{18}$ cm$^{-3}$ (or higher) can be reasonable. However, such high $N_{\text{peak}}$ can give high junction leakage which deteriorates program disturbance [16].

Fig. 3 shows $V_{\text{in}}$ variation with the change of $L_g$ as a parameter of $N_{\text{S/Dpeak}}$ at a fixed pitch ($L_g + L_s$). Here $N_{\text{peak}}$ is fixed at $5 \times 10^{17}$ cm$^{-3}$, which is generally adopted at industry. Since there may be a process-dependent tolerance in $L_g$, we assume $\pm 10\%$ of the tolerance in this work. The projected range ($R_{\text{Ch}}$) and the straggle ($\Delta R_{\text{Ch}}$) of $N_{\text{peak}}$ are fixed at 16.5 and 15.6 nm, respectively. 40 nm device without S/D has a $V_{\text{in}}$ variation of $\sim 50$ mV, which is much smaller than that ($\sim 130$ mV) at an $N_{\text{peak}}$ of $1 \times 10^{18}$ cm$^{-3}$. With decreasing $N_{\text{peak}}$, an inversion layer can be induced more easily at the space region. With the decrease of $L_g$, the $V_{\text{in}}$ variation is also reduced. At the $L_g$ of 20 nm, the variation is $\sim 10$ mV and quite similar to that of device with S/D. Thus the cell scheme without S/D gives better DIBL and SS with scaling-down while keeping similar $V_{\text{in}}$ variation with the $L_g$ variation. This figure provides important information to find a reasonable $N_{\text{peak}}$ value to fabricate a NAND string with immunity to the process variation.

It is very important to verify the read operation of an erased (or programmed) cell in a cell string utilizing the fringing field con-

Fig. 2. DIBL and SS with S/D width ($W_{\text{S/D}}$) as a parameter of $N_{\text{peak}}$. Peak S/D doping concentration is fixed at $3 \times 10^{18}$ cm$^{-3}$.

For reasonable sensing, the read current (bit-line current $= I_{\text{bitline}}$) needs to be high as possible. As an example, we check the $I_{\text{bitline}}$ for erased cells with $L_g = L_s = 30$ nm (a) and $L_g = L_s = 20$ nm (b) as a parameter of S/D doping concentration in the space region. To make the cells erased, positive charge densities of $4.27 \times 10^{12}$ cm$^{-2}$ and $3.68 \times 10^{12}$ cm$^{-2}$ are stored in the storage nodes of 30 nm and 20 nm cells, respectively, which gives $\sim 2$ V of $V_{\text{th}}$ shift. To check the $I_{\text{bitline}}$, 0 V is applied the CG of the erased cell (center cell) and a bit-line bias is applied to the drain of the selection device (SSL). As shown in Fig. 4a and b, the read current for w/o $N_{\text{S/Dpeak}}$ is the lowest in each figure, but much higher than 1 $\mu$A at a $V_{\text{bitline}}$ of 0.8 V. Considering that $I_{\text{bitline}}$ under read operation is generally $\sim 0.1$ mA, $I_{\text{bitline}}$ in these figures is reasonable with a margin which should be taken into account for unwanted charge trapping into the insulator on the space region. As the $N_{\text{S/Dpeak}}$ increases from 0 (w/o S/D) to $1 \times 10^{18}$ cm$^{-3}$, $I_{\text{bitline}}$ increases by 38% and 12%, respectively, in figures (a) and (b). These data say that the w/o S/D (or virtual S/D) concept guarantees sufficient $I_{\text{bitline}}$ and becomes more useful as $L_g (=L_s)$ scales down. If the $N_{\text{peak}}$ is increased to $1 \times 10^{18}$ cm$^{-3}$, the $I_{\text{bitline}}$ is decreased to $\sim 1.5$ mA (3 times smaller) at a $V_{\text{bitline}}$ of 0.8 V. In this case, it becomes a method to increase $I_{\text{bitline}}$ to adopt localized buried oxide (L-BOX) just underneath the space regions in a cell string [21]. It is also useful to use arch shape body to increase $I_{\text{bitline}}$ because the fringing

Fig. 3. $V_{\text{in}}$ variation versus $L_g$ as a parameter of $N_{\text{S/Dpeak}}$. Here, for the $L_g$s of 20, 30, and 40 nm, the $L_g$ variation is $\pm 10\%$. $N_{\text{peak}}$ is fixed at $5 \times 10^{17}$ cm$^{-3}$.

Fig. 4. Read current ($I_{\text{bitline}}$) during read operation of an erased cell in a cell string as a parameter of doping concentration of n-type S/D region in the space region. (a) $L_g = L_s = 30$ nm. (b) $L_g = L_s = 20$ nm. The channel width of a cell is fixed at 20 nm. The peak p-type channel doping concentration ($N_{\text{S/peak}}$) is fixed at $5 \times 10^{17}$ cm$^{-3}$. The $R_{\text{Ch}}$ and $\Delta R_{\text{Ch}}$ are 16.5 nm and 15.6 nm, respectively.
electric field is concentrated on the space region by the arch-shaped structure [19].

It is interesting to check the effect of non-uniform channel doping along the bit-line body. We can selectively increase channel doping and decrease the body doping in the space region. In Fig. 5, the \( N_{Cpeak} \) is 1 \( \times 10^{18} \) cm\(^{-3} \) and \( L_g \) (=\( L_e \)) of cells is 30 nm. For the virtual S/D case, the p-type doping in the space is 5 \( \times 10^{17} \) cm\(^{-3} \). When n-type S/D is formed, the \( N_{Cpeak} \) is 3 \( \times 10^{18} \) cm\(^{-3} \) and \( x_j \) is 10 nm.

According to our simulation, the \( I_{bitline} \) is -0.25 \( \mu A \) at a \( V_{bitline} \) of 0.8 V when the \( N_{Cpeak} \) is laterally uniform (1 \( \times 10^{18} \) cm\(^{-3} \)) as represented by triangle symbols. If we adopt the non-uniform body doping profile shown in the inset of Fig. 5 (peak p-type doping in the space region is \( \approx 5 \times 10^{17} \) cm\(^{-3} \)), the \( I_{bitline} \) is increased by more than 5 times (~1.4 \( \mu A \)). Thus it is very important to keep low body doping in the space region to enhance read current. The lower doping in the space region can be implemented by using counter-doping. If the S/D is formed in the space region with a peak concentration of 3 \( \times 10^{18} \) cm\(^{-3} \), \( I_{bitline} \) is increased to ~2.63 \( \mu A \) at a \( V_{bitline} \) of 0.8 V as represented by circle symbols.

A SONOS NAND flash cell strings utilizing the fringing electric field concept have been demonstrated experimentally and shown good performance [13–15].

As described above, we confirmed the fringing electric field plays very important role in inducing charge layer in a space region between adjacent cells in scaled NAND flash memory cell string which has SONOS type planar channel cells with sub-3X nm size, and makes possible the read operation of cells with virtual S/D (without S/D). The effect is getting better as the cell size shrinks. Now we investigate the fringing electric field effect in FG NAND flash memory cell string which consists of FG planar channel cells.

Fig. 6a shows 3-D schematic view showing a NAND flash memory cell string consisted of one planar channel FG cell and two selection devices. In this example, we reduced the number of cells because of a limited grid problem. GSL, SSL, and ‘selected’ in this figure represent selection devices for ground and string selection lines, and a read cell, respectively. In figure (b), shown is the cross-sectional view along the bit-line body. Since there is no S/D (virtual S/D) in the string, the space region is connected electrically by the charge layer induced from the fringing field from the CG and FG during read and program operations. The S/D regions formed in the left side of GSL and the right side of the SSL in figure (b) have a peak doping concentration of 1 \( \times 10^{18} \) cm\(^{-3} \) and a junction depth of 30 nm. In this figure, we adopted non-overlap S/D to the CG of the GSL and SSL devices to stress the importance of the fringing field effect, and the non-overlap length is 15 nm. As a method to implement such non-overlap S/D regions, we can utilize oxide spacers formed on sides of the CG and FG gates which fill up fully the gap on both space regions of the center cell before an ion implantation to form the S/D regions [12,13]. The channel has a Gaussian doping profile with a \( \mu_{CG} \) of 15 nm and a \( N_{Cpeak} \) of 2 \( \times 10^{18} \) cm\(^{-3} \), and the doping at the surface is 1 \( \times 10^{18} \) cm\(^{-3} \). The tunneling oxide thickness is 7 nm and the blocking insulator consists of a typical oxide/nitride/oxide (4 nm/4 nm/6 nm) stack.

As the \( L_g \) decreases from 40 nm to 20 nm, DIBL of the cell with S/D (the S/D just meets the edge of the CG) significantly increases to ~470 mV which is two times higher than that with the virtual S/D. At a \( V_{bitline} \) of 1 V, the SS of the 40 nm cell with S/D is 120 mV/dec which is lower than that (140 mV/dec) of the device with the virtual S/D. However, as the \( L_g \) decreases from 40 nm to 20 nm, the SS of cell with the virtual S/D becomes 210 mV/dec which is better than that (~215 mV/dec) of the cell with S/D at such high \( N_{Cpeak} \) (=2 \( \approx 2 \times 10^{18} \) cm\(^{-3} \)) because the increasing rate of DIBL of the device with S/D with decreasing \( L_g \) is faster. The SS for the virtual S/D is improved significantly when laterally uniform \( N_{Cpeak} \) is reduced to ~5 \( \times 10^{17} \) cm\(^{-3} \) or the p-type body doping in the space is selectively reduced to ~5 \( \times 10^{17} \) cm\(^{-3} \) or less. Now we check read current of the device with the virtual S/D.

In the read mode of the NAND flash memory string, selected cell (read cell) and selection devices are biased by 0 V and a \( V_{pass} \) of 4 V, respectively. The \( I_{bitline} \) is an important parameter in the NAND flash string to decide the program/erase state of a selected cell. The \( I_{bitline} \) of the erased cell needs to be ~1 \( \mu A \) by considering a margin, and this condition is satisfied at a \( V_{bitline} \) of ~0.2 V. By considering 32 cells string or beyond, a \( V_{bitline} \) larger than ~0.5 V would be needed. As the \( L_g \) decreases, \( I_{bitline} \) increases at the same
V_{\text{bitline}}$ because the charge layer in the space is induced more easily by the fringing electric field from adjacent FGs. With this structure, we could extend the scalability in conventional floating-gate NAND flash memory.

By comparing the read current ($I_{\text{bitline}}$) of FG and SONOS flash memory cells with the virtual S/D, we can understand that the FG cell gives larger current than that of SONOS cell at even higher $N_{\text{Cpeak}}$. The 30 nm SONOS cell with an $N_{\text{Cpeak}}$ of $1 \times 10^{18}$ cm$^{-3}$ gives $\sim$0.25 $\mu$A as represented by tri-angle symbols in Fig. 5. However, 30 nm FG cell gives $\sim$2.5 $\mu$A as represented by circle symbols in Fig. 7. It seems that the holes stored in the FG of the erased cell can effectively induce electrons in the space because the FG is conductive and the tunneling oxide is 7 nm thick. For 40 nm-node NAND flash string, inversion S/D cells were proved to have better current drivability and superior program disturbance [16].

3. NAND flash memory cells with 3-D channel structure

Now we discuss on 3-D channel (or non-planar) structure NAND flash memory cell which is known to be useful in suppressing short channel effect. Some of the 3-D structure cells are very effective in inducing charge layer (or virtual S/D) by the fringing electric field from the control-gate during read or program operation.

Fig. 8a and b shows schematically 2-D cross-sectional views of simplified concave channel, and convex channel SONOS flash memory cell strucrures, respectively. The channel doping is localized under the gate or is formed uniformly along the bit-line. The substrate is uniformly doped with p-type impurity with a concentration of $1 \times 10^{17}$ cm$^{-3}$ and n$^+$ poly-Si gate is applied. The peak channel doping ($N_{\text{Cpeak}}$) has a Gaussian profile with a peak concentration of $5 \times 10^{17}$ cm$^{-3}$ to suppress punch-through and $V_{\text{th}}$ control. There are three types of channel doping ($N_{\text{type}}$). The $N_{\text{type1}}$ represents that the Gaussian doping vertically is formed uniformly along the bit-line. The $N_{\text{type3}}$ represents channel doping formed only under the gate and the $N_{\text{type3}}$ stands for channel doping formed only under the space region between cells. $R_{\text{cch}}$ represents distance from the surface to the position of $N_{\text{Cpeak}}$. In the concave channel, $R_{\text{cch}}$ represents the distance from the recess bottom to the position of $N_{\text{Cpeak}}$. In the convex channel, $R_{\text{cch}}$ represents the distance from the top of convex channel to the position of $N_{\text{Cpeak}}$. Note no S/D is formed. The planar channel cells are control devices, and have the same types of doping mentioned above.

It was studied that which doping profile is reasonable to each transistor. The $N_{\text{type1}}$ (or $N_{\text{type2}}$) is effective in improving DIBL and SS of planar channel (control) device for a given $N_{\text{Cpeak}}$ of $5 \times 10^{17}$ cm$^{-3}$. If the $N_{\text{Cpeak}}$ is increased further, the $N_{\text{type2}}$ will be effective. The $N_{\text{type1}}$ is also reasonable for the concave channel device by considering DIBL and SS. In the convex channel transistor, the $N_{\text{type2}}$ is the most effective. Table 1 shows comparison of $V_{\text{th}}$, DIBL, and SS of 3 different transistors. The convex channel transistor shows the best performance in terms of DIBL and SS, because the channel structure concentrates effectively the electric field from the gate on the channel. However, since the concave channel transistor has opposite channel structure to the convex structure, it shows the worst performance. Note, among 3 different transistors, the control device shows the largest on-current because of the shortest channel length. Although the convex channel device shows the best DIBL and SS, it has the lowest on-current since the path of the fringing electric field from the gate is the longest due to the recess of the space region. The on-current of convex channel device is $\sim$1 $\mu$A at a $V_{\text{bitline}}$ of 1 V for 20 nm channel width, which may be insufficient if we consider unwanted charge trapping into the oxide on the space region. In this case, we can further reduce the p-type doping by a counter doping as an example so that the $V_{\text{th}}$ in the space region is lowered further and $I_{\text{bitline}}$ is increased.

Fig. 9 shows program characteristics of a cell in a cell string. Here, the channel structures of the program cells have convex and planar shapes. The planar channel structure is compared as control cell. The convex channel cell shows the larger $\Delta V_{\text{th}}$ than the control cell since the channel structure is more effective to focus the electric field from the gate at a given program bias of 12 V. During the program, the pass bias is 4.5 V. The $V_{\text{th}}$ of the convex channel cell increases until a program time of $\sim$1 ms, and then saturated, which comes from that the charge density starts to saturate $\sim$1 ms as shown in the inset. After $\sim$1 ms at a given program bias.

Table 1

<table>
<thead>
<tr>
<th>Channel Type</th>
<th>$V_{\text{th}}$ (V)</th>
<th>DIBL (mV/V)</th>
<th>SS (mV/dec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planar ($N_{\text{type1}}$)</td>
<td>0.685</td>
<td>421.9</td>
<td>173</td>
</tr>
<tr>
<td>Concave ($N_{\text{type1}}$)</td>
<td>2.245</td>
<td>599.05</td>
<td>195</td>
</tr>
<tr>
<td>Convex ($N_{\text{type2}}$)</td>
<td>0.267</td>
<td>359.05</td>
<td>158</td>
</tr>
</tbody>
</table>
of 12 V, the charges stored in the nitride layer hinders the charge injection from the channel by reducing FN (Fowler Nordheim) tunneling.

Fig. 10a shows 3-D schematic twin-FinFET SONOS NAND flash cell string which is composed of a cell and two select transistors (SSL and GSL). Fig. 10b shows cross-sectional schematics of along the bit-line direction. The cell device has twin-channel fin bodies formed by etching the body in the center of the fin body along the bit-line direction [28]. There are 3 transistors in a string where two devices at both edges of string are select transistors as in conventional NAND string. There is no S/D in the cell of the cell string. But S/D regions are formed on the left side of GSL and the right side of SSL. \(W_{ox}, W_{cfin}, H_{fin} \) and \(W_{fin} \) stand for the distance between fins in a cell (10 nm), the width with single fin (15 nm), height of side wall channel (50 nm) and fin width of conventional FinFET structure (40 nm), respectively. The fin body is doped uniformly with the p-type impurity of a concentration of \(2 \times 10^{17} \text{ cm}^{-3} \). Since the body is thin and short channel effect can be effectively suppressed as a result, we can reduce the fin body doping in this twin-fin flash memory cell. The induced layer between cells can be formed easily by the fringing field.

It is very important to verify the read operation of an erased cell in a NAND cell string utilizing fringing field concept. In case of multi-level cell (MLC), it is also important to check the read current in a programmed cell with the same manner. To check the read current operation, 0 V is applied the CG of the erased (\(V_{th} = \frac{1}{2} \text{ V} \)) cell (selected cell or center cell in the string) and a bit-line voltage is applied to the drain of the select device shown in Fig. 10. Here, \(V_{pass} \) is 4.5 V and is applied to the two select transistors: SSL and GSL. As shown in Fig. 11, the read current with erased state is higher than \(I_{bitline} > 0.15 \text{ V} \) for the cell device with a channel fin width of 15 nm, which is enough for reasonable read operation at a higher \(V_{bitline} \) (normally 0.8 V). With decreasing the \(L_g = L_s \), the \(I_{read} \) increases due to increased charge density induced by the fringing field.

4. 3-D stacked NAND flash memory cells

Recently, 3-D stacked NAND flash structures have been proposed to overcome scaling limit of the planar flash memory structure [22–27]. However some of these structures have a limitation in the number of vertical bit cell and large threshold voltage (\(V_{th} \)) distribution in a bit-line (BL) because the change of through-hole size becomes significant along the vertical bit-line as the number of bit cell increases [22–24]. Here, the body has nanotube shape. One cell size per layer is \(~65^2 \), which is relatively large, in [22–24]. Metal gate structure can give bad retention characteristics in high temperature operation due to high thermal expansion coefficient of the metal gate [24]. Thus we require 3-D stacked NAND
The diameter of the nanowires are different from layer to layer, resulting the same shape and diameter of the nanowires, the shape and enhance memory capacity [25]. But since it is very difficult to keep margin for other regions like body (and/or insulator) in the through-hole limited in size (diameter). Nanowires can be stacked vertically with an insulator between adjacent nanowires to enhance memory capacity [25]. But since it is very difficult to keep the same shape and diameter of the nanowires, the shape and diameter of the nanowires are different from layer to layer, resulting in \( V_{in} \) fluctuation.

We discuss briefly on two candidates to solve the problems mentioned above. We first consider CG stacked structure as shown by the cross-sectional view shown in the inset of Fig. 12. In this structure, we stack a layer of control-gate and a layer of insulator alternatively and form trenches to provide independent CG stacks. It is well known that the difference between the top and bottom widths of the trench is relatively small compared to that between the top and bottom diameters of the through-hole.

In the trench between CG stacks, gate stack, poly-Si body and backside oxide (BOX) are formed. S/D pads are formed on top of the CG stack. With increasing the number of CG in a stack, etched trench width is more consistent than the etched through-hole diameter. The body on the top of the CG stack and the bottom of the trench can be doped by \( n^+ \) dopants for reasonable conduction. The top \( n^+ \) body regions are used for S or D and the bottom \( n^+ \) body region is used for electrical conduction between adjacent vertical bodies. One CG can control two bodies on both sides of a gate stack so that we can achieve high density (~4P\(^2\)). The cell transistors have no S/D regions and operate by utilizing the fringing electric field. The \( T_{n}, T_{s}, L_{s}, T_{w} \) represent poly-silicon body thickness, trench width, length between adjacent CGs, and gate length, respectively. The shield region in the center of the trench is to remove any cross-talk between the bodies facing each other in the trench [26].

Fig. 12 shows \( \text{h}_{\text{biline}}-V_{\text{CC}} \) characteristics of the structure shown in the inset. If there is no \( n^+ \) region in the body in the bottom of the trench, we can apply a pass bias to the \( n^+ \) region formed in the substrate under the bottom of the trench to turn-on the channel of the body in the bottom of the trench. It is also possible to program the storage node on the \( n^+ \) region formed in the substrate by applying a program bias to the \( n^+ \) region. Then the channel of the cell formed in the bottom of the trench is always turned on. Here the pass bias is 4.5 V. The structure with \( L_{s} = 60 \text{ nm} \) and \( L_{s} = 40 \text{ nm} \) shows quite reasonable \( \text{h}_{\text{biline}}-V_{\text{CC}} \) characteristics.

\[ \text{h}_{\text{biline}} = 10 \text{ nm}, T_{\text{BOX}} = 20 \text{ nm} \]
\[ L_{s} = 60 \text{ nm}, L_{s} = 40 \text{ nm}, T_{n} = 10 \text{ nm} \]
\[ T_{\text{BOX}} = 20 \text{ nm} \]

Body doping in these bodies can be reduced to a value less than \( 5 \times 10^{-17} \text{ cm}^{-3} \) without serious short channel effect. Therefore, the virtual S/D concept can be effectively applied to 3-D stacked NAND flash memory.

Another example of not using through-hole is to stack bit-line bodies vertically with a layer of insulator between adjacent bodies and to form the channel on the sides of etched body [27]. To separate body stacks, trenches are formed between adjacent body stacks by etching process. In the trench, the CGs are standing vertically on the O/N/O gate stack formed on the sides of adjacent body stacks and shared by the body stacks so that memory density is increased. It was checked that the virtual S/D concept could be applied to this structure [27]. However, since the space in the body between adjacent CGs becomes narrow as going down to the bottom of the trench, it is recommended that the body doping in the space region is reduced by a counter-doping or neural S/D regions are formed.

As mentioned above, the body shapes of cells in 3-D stacked NAND flash memory are nanotube, nanowire, and thin film. The body doping in these bodies can be reduced to a value less than \( 5 \times 10^{-17} \text{ cm}^{-3} \) without serious short channel effect. Therefore, the virtual S/D concept can be effectively applied to 3-D stacked NAND flash memory.

5. Conclusions

We have investigated fundamental properties of NAND flash memory cell strings which have virtual S/D (also called as junction free or inversion S/D). It was confirmed that a NAND flash cell string consisting of nano-scale cells without S/D (or virtual S/D) can be worked reasonably by utilizing the fringing electric field from the control-gate (CG) and/or floating-gate (FG) to induce charge layer in the space region between adjacent cells. It was shown that the doping in the space region should be reduced as low as possible to increase read current (bit-line current). The cell structure (fin and arch body structures) which has body and CG structures to effectively focus the fringing electric field on the
space region could induce more easily the charges in the space region. For the planar and non-planar channel cells, the virtual S/D concept was proven to be very promising in sub-3X technology nodes because the charges can be induced more easily in decreasing space regions. It was also shown that the virtual S/D concept can be applied to 3-D stacked NAND flash memory technology.

Acknowledgment

This work was supported by the IT Research and Development Program of Ministry of Knowledge Economy/Korea Evaluation Institute of Industrial Technology under Grant 10035320 (Development of novel 3-D stacked devices and core materials for the next-generation Flash memory).

References