Received 8 October 2023; accepted 15 October 2023. Date of publication 20 October 2023; date of current version 3 November 2023.
The review of this article was arranged by Editor C. Bulucea.

Digital Object Identifier 10.1109/JEDS.2023.3326173

Analytical Modeling of Threshold Voltage and Subthreshold Slope for 3-D NAND Flash Memory With a Non-Uniform Doping Profile

AMIT KUMAR (Student Member, IEEE), AND SHUBHAM SAHAY (Senior Member, IEEE)

Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

CORRESPONDING AUTHOR: S. SAHAY (e-mail: ssahay@iitk.ac.in)

This work was supported in part by the Startup Research Grant (SRG) from Science and Engineering Research Board (SERB) under Grant SERB/EE/2021358.

ABSTRACT The emergence of data-driven technologies including Internet of Things (IoT), artificial intelligence (AI), and cloud computing has led to a surge in data generation and mining. The 3D NAND Flash memory has emerged as a promising technology for handling the big data owing to its ultra-high density, ultra-low cost per bit, fast random access, and multi-level programming capability per cell. However, the conventional punch and plug process used to fabricate the 3D NAND flash memory leads to an inherent tapering of the polysilicon channel of the stacked flash cells. The channel thickness of the flash cells increases monotonically from bottom of the string (common source line, CSL) to the top of the string (bit-line, BL). This results in a non-uniform intrinsic threshold voltage of the flash cells located at different word-line (WL) layers along the string. A non-uniform vertical (Gaussian) channel doping profile was recently proposed to mitigate the impact of channel tapering and realize a uniform threshold voltage distribution. However, an analytical model which may provide a physical insight and can be used for design exploration of 3D NAND flash memory with uniform intrinsic threshold voltage is still elusive. To this end, in this work, for the first time, we formulate an analytical model for the intrinsic threshold voltage and the subthreshold slope of 3D NAND flash cells with a vertical gaussian doping profile in the channel region. The results obtained using the analytical model are in close agreement with the TCAD simulations validating the accuracy of the developed model. We also utilize the analytical model to provide physical insights and necessary design guidelines for further optimizing the performance.

INDEX TERMS 3D NAND, flash memory, Macaroni body.

I. INTRODUCTION

The overwhelming need for high-dimensional data storage in this era of Internet of Things (IoT), artificial intelligence (AI), cloud computing and big data has propelled the development of planar NAND flash memories [1], [2], [3]. However, the incessant scaling of planar NAND flash to increase the bit-density is limited by the neighboring cell coupling disturbances and crosstalk which decrease the retention and restrict the multibit capability. Therefore, several 3D NAND architectures based on vertical stacking of flash cells such as BiCS [4], [5], TCAT [6], VRAT [7], VNAND, ZNAND [8], [9] etc. were proposed recently. 3D NAND flash memory has emerged as the most promising candidate to handle the data explosion with applications including SSDs, USB drives and cloud storage.

3D NAND flash memory utilizes the gate-all-around Macaroni body structure owing to its low footprint, higher immunity against short-channel effects as compared to the planar NAND flash memory [5], [10] and improved gate control which enhances the subthreshold characteristics. In a Macaroni body architecture, a hole is etched in the polysilicon channel and filled with a dielectric to restrict the thickness of the polysilicon region below the depletion region width. This leads to a significant reduction in the threshold voltage fluctuations due to the polysilicon grain boundaries and associated traps within the channel region [5].

Moreover, vertical stacking of more than 192 flash cells and word line (WL) layers in a string has already been demonstrated in commercial 3D NAND flash memories [11], [12]. However, there is an urgent need to further

increase the number of stacked layers to realize ultra-high density and cope-up with the data storage needs in this era of big data [13]. As the number of stacked layers increases, etching a narrow hole in such a large mold height to realize the string of flash cells becomes a technological challenge [14]. Furthermore, etching of such a large stack with aspect ratio greater than 40:1 even with the most advanced dry reactive ion etching (DRIE) technique leads to a substantial tapering of the hollow polysilicon channel region. This tapering is expected to further increase in the next generation 3D NAND flash memories with > 232 stacked layers owing to the significantly increased aspect ratio. The tapered shape of the channel leads to an increased channel thickness for flash cells located close to the bit-select line (BSL) and a smaller channel thickness for the cells located near the source-select line (SSL). Therefore, the conventional punch and plug fabrication process leads to an inherent non-uniformity in intrinsic threshold voltage for the flash cells located at different WL layers in the string. This necessitates the use of complex error-correction circuitry (ECC) [14] and WL layer-dependent programming schemes [15] which lead to an additional area and energy overhead.

To mitigate the impact of the inherent tapering of the channel region and the consequent WL layer-dependent threshold voltage (programming characteristics), the semiconductor IDMs are exploring several methodologies including blocking oxide compensation, non-uniform vertical (Gaussian) doping profile in the channel region, etc. [14]. A vertical non-uniform (Gaussian) doping profile in the polysilicon channel region of the 3D NAND flash string with Macaroni body enables the realization of a uniform intrinsic threshold voltage for flash cells located at different WL layers in the string [14] and appears lucrative alternative to the conventional 3D NAND flash memory for incessant stacking of WL layers without introducing complex WL-layer dependent program/erase scheme or ECC. Although a behavioral compact model [16] and an analytical model [17] have been proposed for 3D NAND flash cells with uniform doping profile, an analytical model which provides a physical insight and the necessary design guidelines for 3D NAND Macaroni body flash cell with vertical Gaussian doping profile was elusive. Therefore, we formulated the analytical model for the inner potential and the surface potential of a 3D NAND flash memory with a Gaussian doping profile in the channel region in [23]. However, the stored data is encoded in the threshold voltage of the flash cells in a 3D NAND flash memory and the subthreshold swing (SS) is an important parameter to assess the switching characteristics [24], [25]. Therefore, an analytical model for the intrinsic threshold voltage and sub-threshold swing becomes crucial for performance estimation and design space exploration [20], [21] of a Macaroni body 3D NAND flash memory with a Gaussian doping profile in the channel region. To this end, in this work, for the first time, we solve

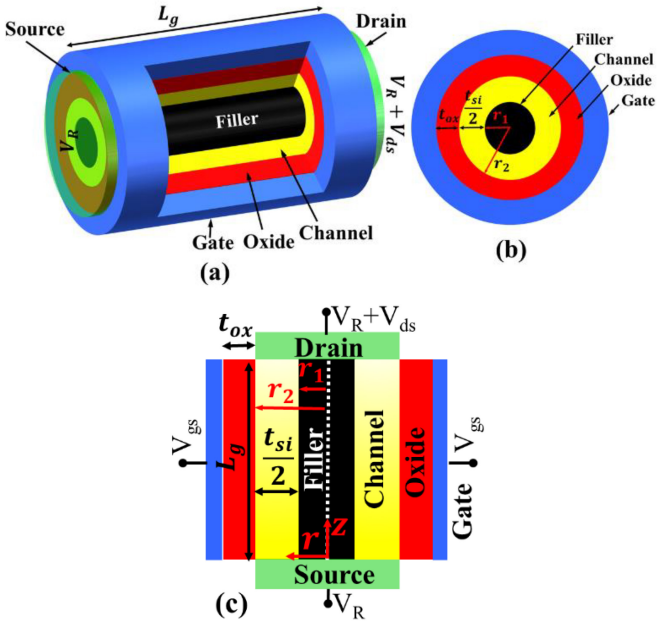

FIGURE 1. (a) 3D view and 2D cross-sectional view along (b) radial direction (*r*) (c) axial (along the channel length) of the Macaroni body 3D NAND flash cell with a vertical Gaussian doping profile and uniform threshold voltage across cells at different WLs.

the 3D Poisson's equation in cylindrical coordinates with appropriate boundary conditions and derive an analytical expression for the threshold voltage and subthreshold slope of the Macaroni body 3D NAND flash cell with a vertical Gaussian doping profile in the channel region. Although the TCAD simulations may provide similar physical insights as the analytical models, they are computationally intensive and require huge amounts of processing power and memory resources. The analytical models provide a faster means of design exploration across different geometric and design parameters and an enhanced speed of circuit simulations since they consume only a fraction of time and computational resources as compared to the TCAD simulations.

II. DEVICE STRUCTURE AND MODEL FORMULATION

The Macaroni body 3D NAND flash cell with a non-uniform (vertical Gaussian) doping profile in the channel region is shown in Fig. 1 and the structural parameters are reported in Table 1. Although the proposal of a 3D NAND flash with a non-uniform doping is given for a structure where the hole is fully filled with polySi [14], a Macaroni body leads to the realization of an ultra-thin poly-silicon channel region which improves the electrostatic integrity and leads to a significant reduction in the number of grain boundaries and associated traps in the channel region resulting in a tighter distribution of the string current and threshold voltage across different flash cells. Considering the enhanced robustness of the 3D NAND flash memory with a Macaroni body against the variation due to polySi grain boundaries and the associated traps, we have utilized this structure in our study. Moreover, the commercial 3D NAND flash memories also employ the

TABLE 1. Structural parameters of the Macaroni body 3D NAND flash cell with a vertical Gaussian doping profile.

Parameters	Values
Peak channel doping, $(N_{D,peak})$	10^{14} - 10^{18} cm^{-3} [17]
Gate length (L_g)	40 nm - 160 nm [17]
Oxide thickness (t_{ox})	6 nm, 12 nm [17]
Gate work function (ϕ_m)	4.6 eV [17]
Outer radius (r_2)	17.5 nm,19.5 nm, 21.5 nm, 23.5 nm [17]
Inner radius (r_1)	13.5 nm [17]

Macaroni body structure and suffer from the problem of etch taper due to the conventional punch and plug fabrication process. Therefore, a non-uniform doping profile is expected to be highly beneficial even for the 3D NAND flash with a Macaroni body. For realizing a uniform intrinsic threshold voltage in 3D NAND flash cells at different WLs along the string, the polysilicon channel region is heavily doped towards the CSL (source end) and lightly doped close to the BL (drain end) utilizing a multi-implant process [14]. As shown in Fig. 2, we have used a vertical Gaussian distribution for the analytical treatment of the non-uniform doping in the channel region ($N_D(z) = N_D exp(\frac{-kz^2}{2\sigma^2})$) and calibrated the parameters (mean and variance) to reproduce the optimal doping profile reported in [14] for the 3D NAND flash memory with a uniform threshold voltage distribution across the flash cells located at different word lines along the string. The doping concentration is gradually decreased from SSL to BSL along the z-direction (Fig. 1) [14]. As shown in Fig. 2, the value of N_D varies from $3 \times 10^{18}\text{cm}^{-3}$ at the source end (SSL) to 10^{15}cm^{-3} at the drain end (BSL). However, the doping remains constant at a particular value of z i.e., the doping concentration is uniform along the radial direction. Furthermore, for simplicity and ease of handling analytical expressions, following the prior modelling works on 3D NAND flash memory [16], [17], we have ignored the grain boundaries and their associated traps and replaced the ONO stack with a single oxide layer with effective oxide thickness of the ONO layer.

The analytical model is derived by simplifying the 3D Poisson's equation (equation (1)) to a second-order differential equation (equation (2)) utilizing Young's parabolic potential approximation [18] (equation (3)) and boundary conditions pertaining to the continuity of electrical flux [23].

$$\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\left(\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial\theta}\left(\frac{\partial\psi}{\partial\theta}\right) + \frac{\partial^2\psi}{\partial z^2} = \frac{-qN_D(z)}{\epsilon_{Si}} \; for \; r_1 \le r \le r_2 \tag{1}$$

where r, θ, and z are the cylindrical coordinate axes along the radial, angular, and axial (along the channel length) directions, ψ is the channel potential, q is the electron charge, ϵ_{si} is the dielectric constant of silicon and $N_D(z)$ is the donor concentration in the channel region of the

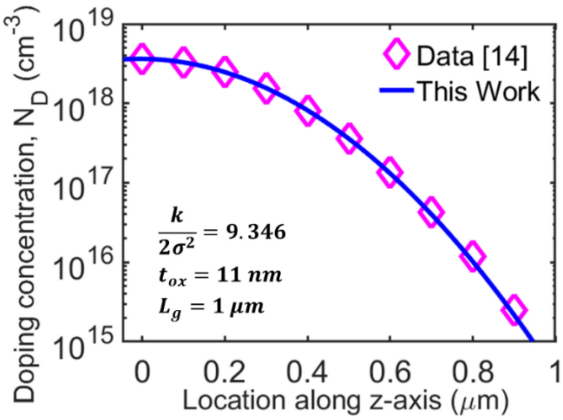

FIGURE 2. Parameters used for modeling the non-uniform doping profile in [14].

Macaroni body cell. Furthermore, the structure shown in Fig. 1 is symmetrical along the axial direction, results in ψ independent of θ i.e., $\frac{\partial\psi(r,\theta,z)}{\partial\theta} = 0$. Equation (1) will reduce to a 2D Poisson's equation given by [23] as:

$$\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\left(\frac{\partial\psi}{\partial r}\right) + \frac{\partial^2\psi}{\partial z^2} = \frac{-qN_D}{\epsilon_{Si}} exp\left(\frac{-kz^2}{2\sigma^2}\right) \tag{2}$$

Now, Young's parabolic potential approximation [18] can be utilized to simplify the 2D electrostatic potential distribution along the radial direction inside the channel region as:

$$\psi(r,z) = C_1(z) + r.C_2(z) + r^2.C_3(z) \tag{3}$$

where $C_1(z)$, $C_2(z)$ and $C_3(z)$ are arbitrary coefficients which can be obtained by using following boundary conditions [23]:

a) Electric field lines inside the channel terminate at the silicon body-core filler dielectric interface, i.e., the electric field at $r = r_1$ is zero: $\frac{\partial\psi(r_1,z)}{\partial r} = 0$.
b) Electric flux at the interface between the silicon film and the gate oxide must be continuous at $r = r_2$: $\epsilon_{si}\frac{\partial\psi(r_2,z)}{\partial r} = C_{ox}(V_{gs} - V_{fb} - \psi_s)$.

where, t_{Si} is the channel thickness, V_{fb} is the flat-band voltage, ψ_s is the surface potential ($\psi_s = \psi(r = r_2, z)$) and C_{ox} is the effective gate oxide capacitance per unit area of cylindrical Macaroni body cell given by [23]: $C_{ox} = \frac{\epsilon_{ox}}{r_2 \ln(1+\frac{t_{ox}}{r_2})}$, ϵ_{ox} is the dielectric constant of oxide, t_{ox} is the gate oxide thickness.

Moreover, we define the potential at the channel-core filler interface ($r = r_1$) as the inner potential, $\psi_o = \psi(r = r_1, z)$.

Now, utilizing above boundary conditions and assumption, equation (2) can be expressed as:

$$\frac{d^2\psi_0(z)}{dz^2} - \frac{\psi_0(z)}{\lambda^2} = \frac{-(V_{gs} - V_{fb})}{\lambda^2} - \frac{qN_D}{\epsilon_{Si}} exp\left(\frac{-kz^2}{2\sigma^2}\right) \tag{4}$$

where V_{fb} is the flat-band voltage and λ is characteristic length of the Macaroni body 3D NAND flash cell expressed as $\lambda = \sqrt{\frac{4\epsilon_{Si}t_{Si}+C_{ox}t_{Si}^2}{8C_{ox}}}$ [23].

The inner potential (ψ_0 at the channel/core-filler interface in Figure 1(c)) obtained by solving the differential equation utilizing the principle of superposition where the solution consists of a complementary function and a particular integral [23] as:

$$\psi_0(z) = \left[\frac{(V_R - K_1)\sinh\left(\frac{L_g - z}{\lambda}\right) + (V_R + V_{ds} - K_2)\sinh\left(\frac{z}{\lambda}\right)}{\sinh\left(\frac{L_g}{\lambda}\right)}\right] + \left\{(V_{gs} - V_{fb}) + \frac{qN_D\lambda^2}{\epsilon_{Si}}exp\left(\frac{-kz^2}{2\sigma^2}\right)\right\} \quad (5)$$

K_1 and K_2 are constants that can be expressed as [23]:

$$K_1 = (V_{gs} - V_{fb}) + \left(\frac{qN_D\lambda^2}{\epsilon_{Si}}\right) \quad (6)$$

$$K_2 = (V_{gs} - V_{fb}) + \left(\frac{qN_D\lambda^2}{\epsilon_{Si}}\right)exp\left(\frac{-kL_g^2}{2\sigma^2}\right) \quad (7)$$

The surface potential (ψ_S at channel/gate oxide interface in Fig. 1(c)) can be obtained as [23]:

$$\psi_s(z) = \psi_0'(z) + (V_{gs} - V_{fb} - \psi_0'(z))\left(\frac{t_{Si}^2}{8\lambda^2}\right) \quad (8)$$

where $\psi_0'(z)$ is inner potential obtained by solving the differential equation utilizing the principle of superposition where the solution consists of a complementary function and a particular integral (PI).

$$\psi_0'(z) = C_1'exp\left(\frac{z}{\lambda}\right) + C_2'exp\left(-\frac{z}{\lambda}\right) + PI \quad (9)$$

where C_1' and C_2' are constants with PI, can be expressed as [23]:

$$PI = (V_{gs} - V_{fb}) + \lambda^2\left(\frac{qN_D}{\epsilon_{Si}}\right)exp\left(\frac{-kz^2}{2\sigma^2}\right) \quad (10)$$

Furthermore, to estimate the C_1' and C_2' the following boundary conditions can be used [23]:

a) The surface potential at the source terminal $\psi_s(z = 0) = V_R$
b) The surface potential at the drain end $\psi_s(z = L_g) = V_R + V_{ds}$.

Utilizing the above boundary conditions, the analytical model for surface potential can be expressed as [23]:

$$\psi_s(z) = (V_{gs} - V_{fb}) + K_3exp\left(\frac{-kz^2}{2\sigma^2}\right) + \left[\frac{K_4\sinh\left(\frac{z}{\lambda}\right) - K_5\sinh\left(\frac{L_g - z}{\lambda}\right)}{\sinh\left(\frac{L_g}{\lambda}\right)}\right] \quad (11)$$

where K_3, K_4 and K_5 are constants that can be expressed as:

$$K_3 = \frac{qN_D\lambda^2}{\epsilon_{Si}}\left(1 - \frac{t_{Si}^2}{8\lambda^2}\right) \quad (12)$$

$$K_4 = V_R + V_{ds} - K_2 + \lambda^2\left(\frac{qN_D}{\epsilon_{Si}}\right)exp\left(\frac{-kL_g^2}{2\sigma^2}\right)\left(\frac{t_{Si}^2}{8\lambda^2}\right) \quad (13)$$

$$K_5 = K_1 - V_R - \lambda^2\left(\frac{qN_D}{\epsilon_{Si}}\right)\left(\frac{t_{Si}^2}{8\lambda^2}\right) \quad (14)$$

III. MODEL DESCRIPTION

A. THRESHOLD VOLTAGE, V_T

Since the stored data is encoded in the threshold voltage of the flash cells in a 3D NAND flash memory, an analytical model for the intrinsic threshold voltage becomes crucial for performance estimation and design space exploration [20], [21]. The threshold voltage for macaroni body 3D NAND flash cell is defined as the gate voltage at which the neutral region at the center of the channel disappears due to the merging of depletion region widths contributed by the gate electrode [22]. The threshold voltage can be obtained by finding the critical position (z_m), at which the inner potential ($\psi_0(z)$) of the channel is minimum and equating it to the reference potential V_R as [17]:

$$\psi_0(z_m) = \psi_{0,min} = V_R = 0 \quad (15)$$

$$\left[\frac{-K_1\sinh\left(\frac{L_g - z_m}{\lambda}\right) + (V_{ds} - K_2)\sinh\left(\frac{z_m}{\lambda}\right)}{\sinh\left(\frac{L_g}{\lambda}\right)}\right] + \left\{(V_T - V_{fb}) + \frac{qN_D\lambda^2}{\epsilon_{Si}}exp\left(\frac{-kz_m^2}{2\sigma^2}\right)\right\} = 0 \quad (16)$$

$$(V_{ds} - K_2)\sinh\left(\frac{z_m}{\lambda}\right) - K_1\sinh\left(\frac{L_g - z_m}{\lambda}\right) + \sinh\left(\frac{L_g}{\lambda}\right)(V_T - V_{fb}) + \sinh\left(\frac{L_g}{\lambda}\right)\left\{\frac{qN_D\lambda^2}{\epsilon_{Si}}exp\left(\frac{-kz_m^2}{2\sigma^2}\right)\right\} = 0 \quad (17)$$

Substituting the value of K_1 and K_2 in equation (17) and replacing V_{gs} with V_T, we get:

$$(V_T - V_{fb})\left\{\sinh\left(\frac{L_g}{\lambda}\right) - \sinh\left(\frac{L_g - z_m}{\lambda}\right) - \sinh\left(\frac{z_m}{\lambda}\right)\right\} = \left(\frac{qN_D\lambda^2}{\epsilon_{Si}}\right) \times \sinh\left(\frac{z_m}{\lambda}\right)exp\left(\frac{-kL_g^2}{2\sigma^2}\right) - \sinh\left(\frac{L_g}{\lambda}\right)\left\{\frac{qN_D\lambda^2}{\epsilon_{Si}}exp\left(\frac{-kz_m^2}{2\sigma^2}\right)\right\} - V_{ds}\sinh\left(\frac{z_m}{\lambda}\right) + \left(\frac{qN_D\lambda^2}{\epsilon_{Si}}\right)\sinh\left(\frac{L_g - z_m}{\lambda}\right) \quad (18)$$

$$V_T = V_{fb} + \frac{A_4\left[A_2 + A_3exp\left(\frac{-kL_g^2}{2\sigma^2}\right) - A_1exp\left(\frac{-kz_m^2}{2\sigma^2}\right)\right] - A_3V_{ds}}{A_1 - A_2 - A_3} \quad (19)$$

where

$$A_1 = \sinh\left(\frac{L_g}{\lambda}\right) \tag{20}$$

$$A_2 = \sinh\left(\frac{L_g - z_m}{\lambda}\right) \tag{21}$$

$$A_3 = \sinh\left(\frac{z_m}{\lambda}\right) \tag{22}$$

$$A_4 = \frac{qN_D\lambda^2}{\epsilon_{Si}} \tag{23}$$

B. SUBTHRESHOLD SLOPE, SS

The subthreshold swing (SS), an important parameter to assess the switching characteristics [24], [25], is defined as the variation in the gate-to-source voltage required to obtain an order of magnitude (decade) change in the drain-to-source current in the sub-threshold regime of operation. Mathematically, SS can be expressed as [26]:

$$SS = \left(\frac{\partial \log_{10} I_{ds}}{\partial V_{gs}}\right)^{-1} \tag{24}$$

The subthreshold current mainly flows through the channel/filler interface where the channel potential is minimum $\psi_0(z = z_m)$. Therefore, the subthreshold current is proportional to $exp(\frac{\psi_0(z=z_m)}{V_t})$ [26] and SS can be expressed as:

$$SS = \left(\frac{\ln(10)V_t}{\frac{\partial \psi_0(z=z_m)}{\partial V_{gs}}}\right) \tag{25}$$

where V_t is thermal voltage have value $\frac{KT}{q}$.

Differentiating equation (5) with respect to V_{gs} at $z = z_m$ and substituting the obtained expression of $\frac{\partial \psi_0(z=z_m)}{\partial V_{gs}}$ in equation (25), the SS can be expressed as:

$$SS = 2.3V_t\left[1 - \frac{\sinh\left(\frac{z_m}{\lambda}\right) + \sinh\left(\frac{L_g - z_m}{\lambda}\right)}{\sinh\left(\frac{L_g}{\lambda}\right)}\right]^{-1} \tag{26}$$

where z_m is a critical position of inner potential at channel interface which can be obtained by putting $\frac{\partial \psi_0}{\partial z} = 0$ [17], and can be expressed as:

$$z_m = \frac{\lambda}{2} ln\left[\frac{exp\left(\frac{L_g}{\lambda}\right) - \left(\frac{K_2 - V_{ds}}{K_1}\right)}{\left(\frac{K_2 - V_{ds}}{K_1}\right) - exp\left(-\frac{L_g}{\lambda}\right)}\right] \tag{27}$$

IV. RESULT AND DISCUSSION

It may be noted that experimental data for 3D NAND flash memory string/cells from the commercial semiconductor integrated device manufacturers (IDMs) are rarely reported in the literature. Therefore, considering the unavailability of extensive experimental data for 3D NAND flash cells with different structural parameters in the literature, we have followed a two-fold approach to ensure accuracy in validation of our developed analytical model. First, as shown in Fig. 3, we have calibrated our TCAD simulation set up by reproducing the experimental data for the average string current characteristics of a 3D NAND flash array with 10 cells [19]. Then, we utilized the experimentally calibrated TCAD simulation setup for exploring the characteristics of the 3D NAND flash memory array with different structural parameters and validating the results of the developed analytical model. To validate the efficacy of the developed analytical model for the intrinsic threshold voltage and SS, we have compared the results obtained from the model against the TCAD simulations for different design parameters including gate length, channel thickness, oxide thickness, drain (BL) voltage and peak doping concentration. Physical models were used for considering doping and carrier scattering-induced mobility degradation (Philips unified mobility model [27]), carrier recombination (SRH and Auger [27]), and band gap narrowing (Slotboom model [27]).

Fig. 4 shows the variation in the intrinsic threshold voltage with the channel length for different oxide thicknesses at $V_{ds} = 0.6\ V$ and $V_{ds} = 1\ V$. The threshold voltage (measured using the constant current method [19]) reduces significantly with gate length scaling due to the increased short channel effects (such as threshold voltage roll-off). However, the reduction in the threshold voltage is more prominent for a higher oxide thickness and a larger drain (BL) voltage owing to enhanced short-channel effects and reduced electrostatic integrity.

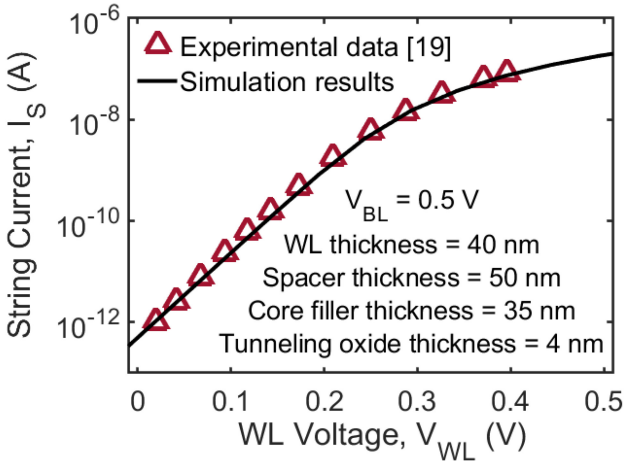

FIGURE 3. Calibration of the TCAD framework by reproducing the experimental characteristics of [19].

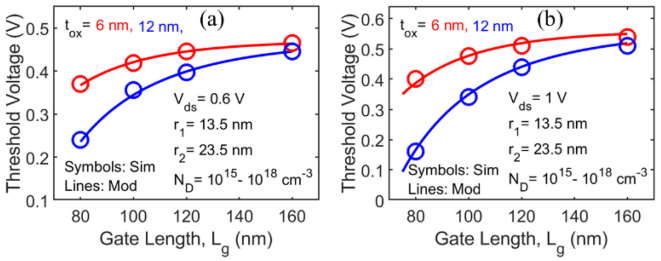

FIGURE 4. Variation in the intrinsic threshold voltage with the gate length (L_g) for different oxide thickness (t_{ox}) at (a) $V_{ds} = 0.6$, (b) $V_{ds} = 1$.

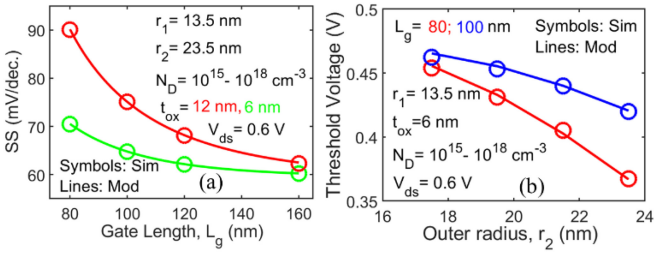

FIGURE 5. (a) The variation in subthreshold slope with the gate length (L_g) for different oxide thickness (t_{ox}), (b) Variation in the intrinsic threshold voltage with the outer radius (r_2) for different gate length (L_g).

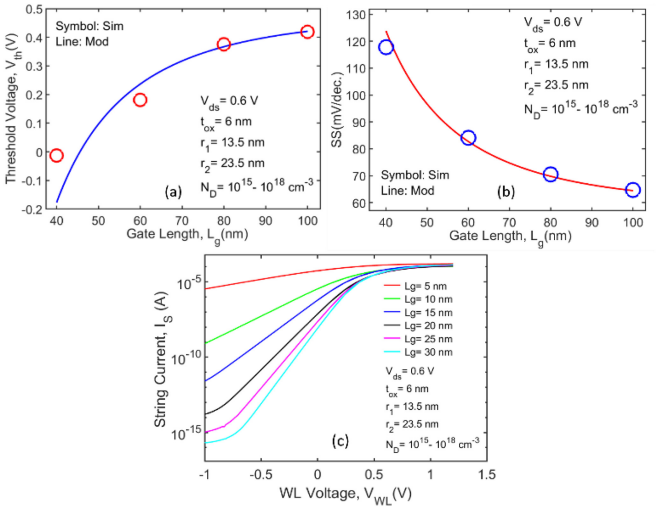

FIGURE 6. The variation in (a) threshold voltage and (b) subthreshold slope with the gate length (L_g) and (c) the simulated string current characteristics of the 3D NAND flash memory with a non-uniform doping profile for ultra-short gate lengths.

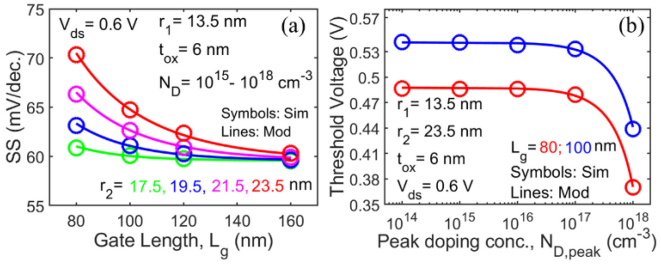

FIGURE 7. (a) The variation in subthreshold slope with the gate length (L_g) for different outer radii (r_2), (b) Variation in the intrinsic threshold voltage with peak doping concentration ($N_{D,peak}$) for different gate length (L_g).

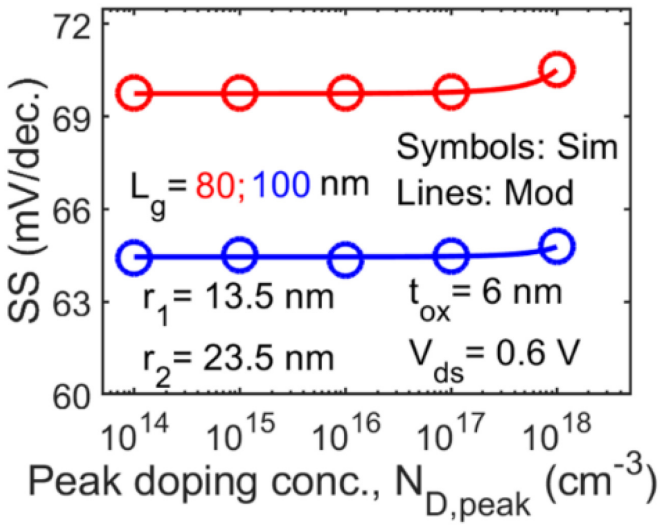

FIGURE 8. The variation in subthreshold slope with peak doping concentration ($N_{D,peak}$) for different gate length (L_g).

The variation in the intrinsic threshold voltage with the channel thickness is analyzed by changing the outer radius (r_2) while keeping a fixed inner radius (r_1) for two different gate lengths (L_g) as shown in Fig. 5(b). For a fixed gate length, threshold voltage increases significantly with scaling down the outer radius at fixed inner radius $r_1 = 13.5\ nm$. Furthermore, the degradation in threshold voltage is more for smaller gate lengths due to dominance of short-channel effects.

Fig. 5(a) shows the variation in the subthreshold slope (SS) with gate length for different oxide thicknesses. As the oxide thickness increases, the subthreshold slope also increases, which results in a poor gate control. Although a thin oxide may appear lucrative while designing 3D NAND flash cell owing to the enhanced immunity against the short-channel effects, an ultra-thin oxide also exhibits a large gate tunnelling leakage current. The gate leakage may induce retention failure while decreasing the memory window. Therefore, to eliminate the gate leakage current and achieve better gate controllability, a high-κ oxide material should be used [28]. Moreover, as shown in Fig. 6, the threshold voltage reduces significantly while the sub-threshold slope increases considerably due to the dominant short-channel effects such as threshold voltage roll-off and drain-induced barrier lowering (DIBL). Furthermore, to evaluate the scaling limits, the simulation results for the string current characteristics of the 3D NAND flash memory with non-uniform doping profile are shown in Fig. 6(c) for ultra-short gate lengths till $L_g = 5$ nm. The ON-state to OFF-state current ratio of the 3D NAND flash cell also degrades significantly when the gate length is scaled aggressively. Therefore, the short-channel effects limit the incessant gate length scaling and restrict the minimum mold height which dictates the aspect ratio during the punch and plug fabrication process for the 3D NAND flash memory with non-uniform doping profile in the channel region.

Fig. 7(a) illustrates the variation in the subthreshold slope with the gate length (L_g) for different outer radii (r_2). The increase in the subthreshold slope is more prominent for a larger channel thickness owing to the poor electrostatic integrity and a lower immunity against short-channel effects [29], [30]. Therefore, an ultra-thin Macaroni channel, a larger channel length, a thin equivalent oxide thickness

(EOT) and a lower BL voltage should be used to extract optimal performance from the 3D NAND flash cell.

Fig. 7(b) and Fig. 8 show the variation in the intrinsic threshold voltage and sub-threshold slope, respectively, with the peak doping concentration ($N_{D,peak}$) of the non-uniform Gaussian doping profile in the channel region for different gate lengths (L_g). The relative insensitivity of the threshold voltage to the peak doping concentration till $N_{D,peak} = 10^{17}\text{cm}^{-3}$ indicates that the performance is dominated by the geometric factors [31]. However, the performance degrades as the peak doping concentration is increased further. Therefore, a peak doping concentration $N_{D,peak} = 10^{17}\text{cm}^{-3}$ must be chosen for obtaining optimal performance from the 3D NAND flash cells with a vertical Gaussian doping profile in the channel.

V. CONCLUSION

We develop an analytical model for the intrinsic threshold voltage and SS of Macaroni body 3D NAND flash memory with a uniform threshold voltage distribution across the cells located at different WL layers and evaluate the impact of different design parameters. The strong agreement between the analytical model and TCAD simulations validates the accuracy of the developed model.

REFERENCES

[1] S. Okamoto et al., "Application driven SCM and NAND flash hybrid SSD design for data-centric computing system," in *Proc. IEEE Int. Memory Workshop (IMW)*, 2015, pp. 1–4, doi: 10.1109/IMW.2015.7150277.

[2] S. Liang et al., "An empirical study of quad-level cell (QLC) NAND flash SSDs for big data applications," in *Proc. IEEE Int. Conf. Big Data*, 2019, pp. 3676–3685, doi: 10.1109/BigData47090.2019.9006406.

[3] C.-W. Yoon, "The fundamentals of NAND flash memory: Technology for tomorrow's fourth industrial revolution," *IEEE Solid State Circuits Mag.*, vol. 14, no. 2, pp. 56–65, Mar. 2022, doi: 10.1109/MSSC.2022.3166466.

[4] H. Tanaka et al., "Bit cost scalable technology with punch and plug process for ultra high density flash memory," in *Proc. IEEE Symp. VLSI Technol.*, 2007, pp. 14–15, doi: 10.1109/VLSIT.2007.4339708.

[5] Y. Fukuzumi et al., "Optimal integration and characteristics of vertical array devices for ultra-high density, bit-cost scalable flash memory," in *Proc. IEEE Int. Electron Devices Meeting*, 2007, pp. 449–452, doi: 10.1109/IEDM.2007.4418970.

[6] J. Jang et al., "Vertical cell array using TCAT(terabit cell array transistor) technology for ultra high density NAND flash memory," in *Proc. Symp. VLSI Tech.*, 2009, pp. 192–193.

[7] J. Kim et al., "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 *Proc. Symp. VLSI Tech.*, pp. 122–123, doi: 10.1109/VLSIT.2008.4588587.

[8] R. Micheloni, *3D Flash Memories*. Dordrecht, The Netherlands: Springer, 2016.

[9] Y. Nishi and B. Magyari-Kope, *Advances in Non-Volatile Memory and Storage Technology*. Sawston, U.K.: Woodhead Publ., 2019.

[10] K. T. Kim, S. W. An, H. S. Jung, K.-H. Yoo, and T. W. Kim, "The effects of taper-angle on the electrical characteristics of vertical NAND flash memories," *IEEE Electron Device Lett.*, vol. 38, no. 10, pp. 1375–1378, Oct. 2017, doi: 10.1109/LED.2017.2747631.

[11] A. Khakifirooz et al., "A 1.67Tb, 5b/cell flash memory fabricated in 192-layer floating gate 3D-NAND technology and featuring a 23.3 Gb/mm^{2}bit density," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, 2023, pp. 27–29, doi: 10.1109/ISSCC42615.2023.10067616.

[12] M. Kim et al., "A 1Tb 3b/cell 8th-generation 3D-NAND flash memory with 164 MB/s write throughput and a 2.4 Gb/s interface," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, vol. 65, 2022, pp. 136–137, doi: 10.1109/ISSCC42614.2022.9731640.

[13] J.-W. Park et al., "A 176-stacked 512Gb 3b/cell 3D-NAND flash with 10.8 Gb/mm^{2}density with a peripheral circuit under cell array architecture," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, vol. 64. 2021, pp. 422–423. doi: 10.1109/ISSCC42613.2021.9365809.

[14] U. M. Bhatt, S. K. Manhas, A. Kumar, M. Pakala, and E. Yieh, "Mitigating the impact of channel tapering in vertical channel 3-D NAND," *IEEE Trans. Electron Devices*, vol. 67, no. 3, pp. 929–936, Mar. 2020, doi: 10.1109/TED.2020.2967869.

[15] S. Sahay, M. Klachko, and D. Strukov, "Hardware security primitive exploiting intrinsic variability in analog behavior of 3D NAND flash memory array," *IEEE Trans. Electron Devices*, vol. 66, no. 5, pp. 2158–2164, May 2019, doi: 10.1109/TED.2019.2903786.

[16] S. Sahay and D. Strukov, "A behavioral compact model for static characteristics of 3D NAND flash memory," *IEEE Electron Device Lett.*, vol. 40, no. 4, pp. 558–561, Apr. 2019, doi: 10.1109/LED.2019.2901211.

[17] Q. Nguyen-Gia, M. Kang, J. Jeon, and H. Shin, "Models of threshold voltage and subthreshold slope for macaroni channel MOSFET," *IEEE Electron Device Lett.*, vol. 41, no. 7, pp. 973–976, Jul. 2020, doi: 10.1109/LED.2020.2995642.

[18] K. K. Young, "Short-channel effect in fully depleted SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 36, no. 2, pp. 399–402, Feb. 1989, doi: 10.1109/16.19942.

[19] D. Resnati et al., "Characterization and modeling of temperature effects in 3D NAND flash arrays—part I: Polysilicon-induced variability," *IEEE Trans. Electron Devices*, vol. 65, no. 8, pp. 3199–3206, Aug. 2018, doi: 10.1109/TED.2018.2838524.

[20] R. Bez, E. Camerlenghi, A. Modelli, and A. Visconti, "Introduction to flash memory," *Proc. IEEE*, vol. 91, no. 4, pp. 489–502, Apr. 2003, doi: 10.1109/JPROC.2003.811702.

[21] W. Wang, T. Xie, and D. Zhou, "Understanding the impact of threshold voltage on MLC flash memory performance and reliability," in *Proc. 28th ACM Int. Conf. Supercomput.*, 2014, pp. 201–210, doi: 10.1145/2597652.2597681.

[22] S. Sahay and M. J. Kumar, *Junctionless Field-Effect Transistors: Design, Modeling, and Simulation*. New York, NY, USA: Wiley, 2019.

[23] A. Kumar, R. Kumar, and S. Sahay, "Analytical modeling of 3D NAND flash cell with a Gaussian doping profile," *IEEE Access*, vol. 10, pp. 85854–85863, 2022, doi: 10.1109/ACCESS.2022.3198398.

[24] D. Gola, B. Singh, and P. K. Tiwari, "Subthreshold modeling of tri-gate junctionless transistors with variable channel edges and substrate Bias effects," *IEEE Trans. Electron Devices*, vol. 65, no. 5, pp. 1663–1671, May 2018, doi: 10.1109/TED.2018.2809865.

[25] G. Hu, P. Xiang, Z. Ding, R. Liu, L. Wang, and T.-A. Tang, "Analytical models for electric potential, threshold voltage, and subthreshold swing of junctionless surrounding-gate transistors," *IEEE Trans. Electron Devices*, vol. 61, no. 3, pp. 688–695, Mar. 2014, doi: 10.1109/TED.2013.2297378.

[26] B. Singh, D. Gola, K. Singh, E. Goel, S. Kumar, and S. Jit, "Analytical modeling of subthreshold characteristics of ion-implanted symmetric double gate junctionless field effect transistors," *Mater. Sci. Semicond. Proc.*, vol. 58, pp. 82–88, Feb. 2017, doi: 10.1016/j.mssp.2016.10.051.

[27] *Sentaurus Device User Guide*, Synopsys, Mountain View, CA, USA, 2010.

[28] E. Gusev et al., "Ultrathin high-k gate stacks for advanced CMOS devices," in *Inter. Elec. Dev. Meeting Techn. Dig.*, 2001, p. 20, doi: 10.1109/IEDM.2001.979537.

[29] C. Jiang, R. Liang, J. Wang, and J. Xu, "A two-dimensionalanalytical model for short channel junctionless double-gate MOSFETs," *AIP Adv.*, vol. 5, no. 5, 2015, Art. no. 57122, doi: 10.1063/1.4921086.

[30] C. Li, Y. Zhuang, S. Di, and R. Han, "Subthreshold behavior models for nanoscale short-channel junctionless cylindrical surrounding-gate MOSFETs," *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3655–3662, Nov. 2013, doi: 10.1109/TED.2013.2281395.

[31] F. Liu, L. Zhang, J. Zhang, J. He, and M. Chan, "Effects of body doping on threshold voltage and channel potential of symmetric DG MOSFETs with continuous solution from accumulation to strong-inversion regions," *Semicond. Sci. Technol.*, vol. 24, no. 8, 2009, Art. no. 85005, doi: 10.1088/0268-1242/24/8/085005.

AMIT KUMAR (Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from the Haldia Institute of Technology, India, in 2016, and the M.Tech. degree in VLSI and embedded system from the Indian Institute of Technology Patna, India, in 2020. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology Kanpur. His research interests include modeling and simulation of 3-D NAND flash and other emerging memory devices.

SHUBHAM SAHAY (Senior Member, IEEE) received the B.Tech. degree (Hons.) in electronics engineering with four gold medals and several cash prizes from the Indian Institute of Technology (BHU) Varanasi in 2014, and the Ph.D. degree in electrical engineering from Indian Institute of Technology (IIT) Delhi in 2018. He is an Assistant Professor with the Department of Electrical Engineering, IIT Kanpur, India. Prior to joining IIT Kanpur, he worked as a Postdoctoral Research Scholar with the University of California at Santa Barbara, Santa Barbara, from 2018 to 2020. He has authored a book on "*Junctionless Field Effect Transistors: Design, Modeling and Simulation*" which is published by the (Wiley-IEEE Press). He has also published several peer-reviewed articles on topics, including semiconductor device design and modeling, neuromorphic computing, and hardware security primitives utilizing emerging non-volatile memories. His current research interests include hardware neuromorphic computing platforms, hardware security primitives, novel device architectures for sub-10 nm regime, analytical and compact modeling of semiconductor devices and non-volatile memories, and spintronics. He is the recipient of the IEEE EDS Early Career Award in 2023 and the Young Associate 2023 of Indian National Academy of Engineering. He is an Editor of *IETE Technical Review* and a Review Editor for Frontiers in Neuroscience: Neuromorphic Engineering and Frontiers in Electronics: Integrated Circuits and VLSI. He is the Chair of IEEE EDS UP Section. He also appeared in the list of Golden Reviewers for IEEE TRANSACTIONS ON ELECTRON DEVICES and IEEE ELECTRON DEVICE LETTERS from 2016 to 2020. For more information, see http://home.iitk.ac.in/ ssahay.