Grain-Orientation Induced Quantum Confinement Variation in FinFETs and Multi-Gate Ultra-Thin Body CMOS Devices and Implications for Digital Design

Seid Hadi Rasouli, Student Member, IEEE, Kazuhiko Endo, Member, IEEE, Jone F. Chen, Member, IEEE, Navab Singh, Senior Member, IEEE, and Kaustav Banerjee, Senior Member, IEEE

Abstract—This paper identifies and investigates a new source of random threshold voltage variation, which is referred to as Grain-Orientation-induced Quantum Confinement (GOQC) in emerging ultra-thin-body metal-gate complementary metal-oxide-semiconductor (CMOS) devices including FinFET, tri-gate, and nanowire field-effect transistors. Due to the dependence of the work function of the metal gates on their grain orientations, different parts of the gate in multigate CMOS devices can have different work functions, resulting in a high electric field in the channel (body) of these devices and, hence, in electrical confinement of the carriers. GOQC effect is shown to be the dominant source of the quantum threshold voltage variation in all emerging ultra-thin multi-gate devices including FinFETs. It is also highlighted for the first time that such variations can have significant implications for the performance and reliability of minimum-sized digital circuits such as static random-access memory cells.

Index Terms—FinFET, grain orientation, intrinsic variability, nanowire-FET, quantum confinement, threshold voltage fluctuation, tri-gate FET, work-function variation.

I. INTRODUCTION

M

etal HAS become the primary gate material in advanced complementary metal–oxide–semiconductor (CMOS) technologies due to the incompatibility of polysilicon with high-κ dielectric materials [1], [2]. Introduction of the metal gate also results in lower gate resistance and eliminates the poly-depletion layer, thereby increasing the ON current of the transistors [3], [4]. While high-κ/metal-gate planar technologies can suppress short-channel effects, beyond the 22-nm technology node, a single gate cannot effectively control the channel. Therefore, multigate CMOS devices (including FinFETs, tri-gate field-effect transistors (FETs), and nanowire FETs) have been proposed to substitute bulk metal–oxide–semiconductor field-effect transistors (MOSFETs) for ultimate scaling [5], [6]. On the other hand, grain orientation (GO)-induced work-function variation (WFV) has been recently identified [7], [8] and experimentally verified [8]–[10] as a new source of variation in metal-gate devices, which affects the characteristics of these devices. While all previous work have focused on the classical effects of the WFV, in this paper, we investigate and report the quantum effects of the GO-induced WFV in multigate devices.

A. Metal-Gate Devices: GO-Induced WFV

It is known that the work function (WF) for a solid material depends on its “surface density,” which refers to the number of metal atoms per unit area on the surface of the metal. Surface density, on the other hand, varies for different orientations. Hence, WF depends on GO, which has been validated by experimental measurements [11].

As shown in Fig. 1, a hypothetical metal gate is composed of several grains, where the WF of each grain depends on the GO, which itself cannot be controlled in the fabrication process. Therefore, the WF of the entire gate (which depends on the WF of the individual grains on the gate) varies from one device to another device, which is referred to as GO-induced WFV and has been identified [7], [8] and experimentally verified [8]–[10] as the new intrinsic source of threshold voltage variation.
Fig. 2. (a) In a generic multigate device, the channel is surrounded by the gate. (b) For thick-channel devices (with thickness in the X- or Y-direction of about 30 nm), the energy subbands are very close together. $E_0 - E_2$ are energy subbands in the unprimed valley, and $E'_0 - E'_2$ are energy subbands in the primed valley. (c) In ultrathin devices (with thickness in the X- or Y-direction of less than 10 nm), energy subbands become quantized, which is referred to as SC. (d) Due to random GOs (which results in WFV), four hypothetical gates with different WFs can be assumed for the nanowire gate, resulting in a transverse electric field inside the channel. (e) High transverse electric field ($E_{SC}$) results in energy subband quantization, which is referred to as EC.

B. Quantum Confinement in Multigate CMOS Devices

There are two types of quantum confinement in ultra-thin body multi-gate CMOS devices, i.e., structural confinement (SC) and electrical confinement (EC), as explained here.

Fig. 2(a) shows a cross section of the general multigate device (nanowire FET), where the channel is surrounded by the gate. The cross section of the multigate device can be nonrectangular, for instance, cylindrical, but in this paper, a rectangular channel cross section will be considered since it is easier to explain the quantum effects. In a nanowire FET [Fig. 2(a)], when channel thickness (either in the X- or Y-direction) is around 30 nm, energy bands are very close together and are placed at the edge of the conduction band [Fig. 2(b)]. However, if the channel thickness drops below 10 nm (in either the X- or Y-direction), energy bands become quantized [Fig. 2(c)], and the minimum energy of subbands increases. Since the lowest subband energy level ($E_0$) is higher due to quantum-mechanical (QM) effects, more gate voltage is required to reach the same inversion carrier concentration as in the classical case [12], [13]. This is referred to as SC since it arises from the physical structure of multigate devices. Due to SC, quantum threshold voltage ($V_{T}^{QM}$) is higher than the classical threshold voltage ($V_{T}^{CL}$). The difference between these two voltages ($\Delta V_{T}^{QM}$) defines the QM effects. SC is important in ultra-thin-body single-gate silicon-on-insulator (SOI) and multigate devices.

Moreover, it is also well known that a high transverse electric field (electric field from the gate to the channel) can also cause energy subband quantization, which is referred to as EC. In the literature, several attempts have been made to model the EC in double-gate devices but focused on the strong-inversion effects [14], [15]. SC has also been discussed in [16]–[20], without considering EC. In [13] and [21], a physical model is developed for quantum effects on the subthreshold electrostatics (including threshold voltage) of the generic (asymmetric) undoped double-gate MOSFETs. In [13] and [21], however, it is assumed that, in a symmetric double-gate device [such as 3-terminal FinFET (3T-FinFET)], the transverse electric field is zero; hence, the carriers are only subjected to SC, whereas this is not the case for metal-gate multigate devices (due to GO-induced WFV) as explained here.

In order to explain the GO-induced quantum effects in multigate CMOS devices, consider a cross section of a nanowire FET [Fig. 2(d)]. For simplicity, it is assumed that each side of the gate is composed of only one grain. Randomness of the GO implies that the work function of the different sides of the gate can be different, which results in a high transverse electric field in the channel. In this paper, we show that this GO-induced transverse electric field causes carrier quantum confinement (EC) [Fig. 2(e)], which is referred to as GO-induced quantum confinement (GOQC).

It is worth noting that the source of this transverse electric field is a random phenomenon (GO); therefore, its quantum effect on the threshold voltage should also be considered as a random variable. Our simulation results clearly indicate the significant impacts of this new source of variation on the total threshold voltage variation (due to both classical and quantum sources of variations). In this paper, we mostly focus on FinFET devices, as an example of multigate CMOS devices, to explore the GOQC effect.

The rest of this paper is organized as follows: In Section II, we experimentally verify that WFV is the main source of variation in FinFET devices and explain our simulation methodology. In Section III, GOQC is modeled, and the probability of occurrence of GOQC for different gate materials is calculated. The effect of GOQC on four-terminal (4T) FinFETs, trigate FETs, and nanowire FETs is also analyzed. Section IV presents the implications of this new source of variation for digital design. Finally, concluding remarks are made in Section V.

II. GO-INDUCED WFV

A. How Important Is WFV?

In multigate CMOS devices, the channel is near intrinsic; therefore, the impact of random dopant fluctuation on the threshold voltage ($V_T$) variation is significantly reduced (therefore, its effect on threshold voltage variation and GOQC is neglected in this work). In order to compare the effect of other sources of variation including fluctuation in gate length ($\Delta L_g$), fin thickness ($\Delta T_{si}$), oxide thickness ($\Delta T_{ox}$), and WFV in metal-gate FinFET (as an example of the multigate devices), we have fabricated several FinFETs with different lengths and fin
Fig. 3. (a) Scanning-electron-microscopy plane view and (b) cross-sectional TEM view of the fabricated 4T-FinFET.

Fig. 4. (a) Threshold voltage variations due to \( \Delta L_g \), \( \Delta T_{si} \), and \( \Delta t_{ox} \). These values are small, indicating that the dominant \( V_T \) variation source is WFV. (b) Measured \( I_{ON} \) variance. The dominant source of \( I_{ON} \) variation is \( V_T \) variation (and, hence, WFV).

thicknesses [Fig. 3(a) and (b)]. Fig. 4(a) shows the effect of different sources of variations on the total threshold voltage variation. Furthermore, in order to separate the effect of threshold voltage variation and parasitic resistance on the ON current of the FinFET, the variation of the ON current is measured at \( V_G = 1 \) V and at an overdrive gate voltage \( (V_G = V_T + 0.65 \text{ V}) \) [Fig. 4(b)]. Note that the effect of threshold voltage variation is canceled out in the overdrive case \( (V_G = V_T + 0.65 \text{ V}) \). Hence, it can be concluded from Fig. 4(a) and (b), that GO-induced threshold voltage variation is the dominant source of the variations in FinFET devices.

**B. How Can We Consider the Effects of GO in Our Simulations?**

As shown in Fig. 5(a), in our simulations, using a 3-D device simulator in Silvaco (ATLAS) \[22\], the gate area is divided into several segments. The area of each segment is determined based on the average size of the grains in a specific gate material (as an example, 17 nm for MoN). Then, the work function of each segment is assigned randomly based on the probabilities of the GOs and corresponding work function (Table I). As an example, Fig. 5(b) shows the \( I_D-V_{GS} \) current curves for a p-type MoN metal-gate FinFET device. In this simulation, the fin thickness, effective oxide thickness (EOT), channel length, and fin height are assumed to be 10 nm, 1 nm, 30 nm, and 30 nm, respectively. Since the average size of the MoN grain is approximately 17 nm, each side of the gate is composed of three grains, where different combinations of the grains (with different orientations) result in threshold voltage variation and different \( I_D-V_{GS} \) current curves.

In order to explore the impact of the GO on the transverse electric field \( (E_{xc}) \) in the channel, we use the aforementioned device (except that the fin thickness is reduced to 4 nm) and consider two different combinations of the grains. Assuming that \( V_{GS} = 0 \) V, when all grains have \( \langle 110 \rangle \) orientation \( (WF = 5.0 \text{ eV}) \), the transverse electric field is negligible. In case 2, it is assumed that one grain in the front gate has \( \langle 112 \rangle \) orientation \( (WF = 4.4 \text{ eV}) \), and other grains have \( \langle 110 \rangle \)
orientation (WF = 5.0 eV) (V GS = 0 V) [Fig. 6(a)]. In this case, the GO-induced transverse electric field [Fig. 6(b) and (c)] significantly affects both subthreshold regime characteristics (such as threshold voltage and I OFF) and inversion regime characteristics (such as I ON).

III. MODELING AND SIMULATION OF QUANTUM CONFINEMENT VARIATION

As detailed in [13], for an n-type FinFET, the transverse electric field is given by

$$E_{\text{xc}} = \frac{(V_{\text{FG}} - V_{\text{BG}}) - (V_{\text{FBH}} - V_{\text{FFB}})}{\frac{1}{t_{\text{ox}}} (t_{\text{oxh}} + t_{\text{oxi}}) + T_{\text{si}}}.$$  (1)

The parameters of (1) are defined in Table II. Without considering WFV, in a symmetric 3T-FinFET, $V_{\text{FG}} = V_{\text{BG}}$ and $V_{\text{FBH}} = V_{\text{FFB}}$; hence, $E_{\text{xc}}$ is zero, and only SC is important. However, in the presence of the WFV, $V_{\text{FBH}} \neq V_{\text{FFB}}$, and $E_{\text{xc}}$ is significant and affects the QM channel potential as given by [13, 14]

$$\Delta \phi_{\text{QM}} = \frac{E_0}{q} - \frac{kT}{q} \ln \left( \frac{gm_d}{\pi \hbar^2 N_c} \frac{qE_{\text{xc}}}{1 - \exp(-qE_{\text{xc}}T_{\text{si}}/kT)} \right).$$  (2)

and QM inversion charge density ($Q_{\text{QM}}$) is given by [13]

$$Q_{\text{QM}} = -\frac{q n_t kT}{N_c} \times \exp \left( \frac{q \phi_{\text{QM}}}{kT} \right) \times \left[ \sum_{j=0} \frac{gm_d}{\pi \hbar^2} \exp \left( -\frac{E_j}{kT} \right) + \frac{g'm_d}{\pi \hbar^2} \sum_{j=0} \exp \left( -\frac{E_j'}{kT} \right) \right]$$  (3)

where $h$ is the reduced Planck's constant; $N_c$ is the 3-D effective density of states (DOS) for the conduction band; $E_0$ is the ground-state subband energy in unprimed valley; $E_j$ ($E'_j$) is the energy separation between the $j$th subband in the unprimed (primed) valley and the bottom of the conduction band at the front surface; and $g$ ($g'$) and $m_d$ ($m'_d$) are the degeneracy and DOS effective masses in the unprimed (primed) valley, respectively. $n_t$ is the intrinsic carrier density, $T$ is temperature, and $k$ is Boltzmann constant. $\gamma$ and $E_0$ are given by [13]

$$\gamma = 1 + \sum_{j=1} \exp \left( \frac{E_0 - E_j}{kT} \right) + \frac{g'm_d}{gm_d} \sum_{j=0} \exp \left( \frac{E_0 - E_j'}{kT} \right)$$  (4)

$$E_0 = \frac{\hbar^2}{2m_x} \left( \frac{\pi}{T_{\text{si}}} \right)^2 + b_0^2 \left( 3 - \frac{4}{3} \left( \frac{1}{[b_0T_{\text{si}}/\pi]^2 + 1} \right) \right)$$  (5)

$$b_0 = \left( \frac{3 m_x q E_{\text{xc}}}{2 \hbar^2} \right)^{\frac{1}{2}}$$  (6)

where $m_x$ is the electron effective mass in the confinement direction.

Fig. 7 shows the effect of GOQC and SC on the energy subbands along the fin thickness of a FinFET with $T_{\text{si}} = 5$ nm. As it can be observed, the separation between the minimum of the conduction band and the first subbands ($E_0$) increases due to GOQC, resulting in higher threshold voltage. Assuming a nonideal subthreshold swing $S$ for FinFET, the change in threshold voltage due to quantum effects is given by [12]

$$\Delta V_T^{\text{QM}} = \frac{S}{S_0} \Delta \phi_{\text{QM}}$$  (7)

where $S_0$ is the ideal subthreshold swing (60 mV/dec). Figs. 8 and 9 show the effect of the GOQC in n- and p-type FinFETs for different WFs of the front gate and back gate ($\Delta WF$).

It is worth noting that the variation in quantum confinement effect due to GO is an exclusive source of variation in multigate devices. In bulk devices, the body thickness is large (which can be treated as a thick $T_{\text{si}}$), and in SOI devices, back-gate oxide ($t_{\text{ox}}$) is thick. In both cases, as can be inferred from (1), the resulting transverse electric field is low. However, in a FinFET (as an example of multigate devices), due to thin channel and oxide thicknesses, GO results in high transverse electric field.

In our simulations, the drain current is derived in both classical (Figs. 6 and 10) and quantum regimes (Fig. 10), assuming identical mobility models (Darwish CVT model and Shockley-Read-Hall recombination [22]). In the CVT model, the transverse-field-, doping-, and temperature-dependent parts of the mobility are given by three components that are combined using Matthiessen’s rule [22]. In the quantum regime, the drift-diffusion mode space (DDMS) method is employed to
Fig. 7. Effect of GOQC on the subband energy levels. $E_C$, $E_1$, and $E_2$ represent the conduction band and the first and second subbands, respectively. SC represents the effect of structural confinement on the subbands ($WF_{FG} = WF_{BG} = 4.6$ eV). GOQC shows the effect of GO-induced EC ($WF_{FG} = 4.7$ eV and $WF_{BG} = 5.1$ eV). $T_{Si}$ is assumed to be 5 nm.

Fig. 8. Comparison between threshold voltage variation induced by structural confinement and GOQC for an n-type FinFET ($T_{Si} = 5$ nm, $L_{ch} = 40$ nm, EOT = 1 nm). The $\Delta WF$ of 200 meV represents the difference between the WFs of the grains in TiN, and the $\Delta WF$ of 800 meV corresponds to the difference between the WFs of the grains in TaN with $\langle 100 \rangle$ and $\langle 220 \rangle$ orientations.

Fig. 9. Comparison between threshold voltage variation induced by structural confinement and GOQC for a p-type FinFET ($T_{Si} = 5$ nm, $L_{ch} = 40$ nm, EOT = 1 nm). The $\Delta WF$ of 600 meV represents the difference between the WFs of the grains in MoN. For WN, the $\Delta WF$ can be up to 1100 meV as shown in Table I.

determine the device characteristics. A self-consistent coupled Schrodinger-Poisson solver has been used in these simulations. Additionally, Von Neumann Boundary Conditions for potential has been employed at the contacts [22]. Since the DDMS and classical approaches use identical transport and mobility models, the difference between the currents from the two approaches is only due to the quantum effect inside the channel. Fig. 11 shows the quantum effect on the threshold voltage for different $T_{Si}$’s without considering WFV, which is inversely proportional to the square of $T_{Si}$ (as expected). Moreover, the GOQC effect has been explored in ballistic transport regime using a fully quantum nonequilibrium Green’s function (NEGF) approach. Exchange-correlation effects (effect of the local carrier density) are also considered in our simulations. Fig. 10 reveals that [in both drift diffusion (using DDMS) and ballistic transport (using NEGF) regimes] quantum effect is increased due to GOQC. In these simulations, a 3T-FinFET with channel length of 34 nm and fin thickness of 6 nm has been used. Cross-sectional view of such a device is shown in Fig. 12, where the average grain size is assumed to be 17 nm. For the case where GOQC is neglected, all grains have a WF of 4.5 eV. For the case where GOQC is considered, it is assumed that grains 1 and 2 have WFs of 4.7 eV, whereas grains 3 and 4 have WFs of 4.5 eV ($T_{Si} = 6$ nm).

Since we extract the threshold voltage from the drain current, we need to assure that the change in the threshold voltage is from EC and not from other sources of variation. For example, one may argue that, by changing the WF of one grain, the transverse electric field below that grain is different, and since the mobility is field dependent, it also affects the drain current apart from the GOQC. In order to separate the effect of work-function-induced variation in mobility, in one set of our simulations, we employ a mobility model that is independent of the transverse electric field. Comparison between the results of two sets of simulations...
where $P_i$ is the probability of the $i$th type of grains. It is obvious that the probability of non-occurrence of the GOQC effect ($Pr(GOQC = 0)$) is given by

$$Pr(GOQC \neq 0 \text{ between grains 1 and 3}) = \sum_{i=1}^{n} P_i \times (1 - P_i) = 0.6 \times 0.4 + 0.4 \times 0.6 = 0.48$$  \hspace{1cm} (8)$$

$$Pr(GOQC = 0 \text{ between grains 1 and 3}) = 1 - \sum_{i=1}^{n} P_i \times (1 - P_i) = 1 - (0.6 \times 0.4 + 0.4 \times 0.6) = 0.52.$$  \hspace{1cm} (9)$$

As mentioned earlier, orientation of the grain is not controllable and should be considered as random variable. Therefore, as explained below, one should calculate the probability of GOQC occurrence in metal-gate devices.

### A. Probability of Occurrence of GOQC

Table I summarizes the WFs of the grains with different orientations in different materials. (TiN and TaN are used for n-type devices, whereas WN and MoN are used for p-type devices.) As mentioned earlier, in order to observe the GOQC effect in the channel, grains on different sides of the gate should have different work functions. As an example, consider a cross section of the multigate MoN device, as shown in Fig. 12. For simplicity, we assume that the metal gate is composed of two grains. GOQC happens inside the channel between grains 1 and 3, if they have different WFs; hence, its probability ($Pr$) is given by

$$Pr(GOQC \neq 0 \text{ between grains 2 and 4}) = 1 - \sum_{i=1}^{n} P_i \times (1 - P_i)$$

$$= 1 - (0.6 \times 0.4 + 0.4 \times 0.6) = 0.52.$$  \hspace{1cm} (10)$$

Therefore, the probability of not observing the GOQC effect in the entire channel is given by

$$Pr(GOQC = 0 \text{ in the entire channel}) = Pr(GOQC = 0 \text{ between grains 1 and 3}) \times Pr(GOQC = 0 \text{ between grains 2 and 4})$$

$$= \left(1 - \sum_{i=1}^{n} P_i \times (1 - P_i)\right)^2.$$  \hspace{1cm} (11)$$

As a result, the probability of GOQC occurrence in the channel is derived from

$$Pr(GOQC \neq 0 \text{ inside the channel}) = 1 - Pr(GOQC = 0 \text{ in the entire channel})$$

$$= 1 - \left(1 - \sum_{i=1}^{n} P_i \times (1 - P_i)\right)^2.$$  \hspace{1cm} (12)$$

Similarly, in a FinFET device, if the gate is composed of $N$ grains, we have $N/2$ grains in the front gate, and the probability of GOQC occurrence is given by

$$Pr(GOQC \neq 0 \text{ in a FinFET}) = 1 - Pr(GOQC = 0 \text{ in FinFET})$$

$$= 1 - \left(1 - \sum_{i=1}^{n} P_i \times (1 - P_i)\right)^{N/2}.$$  \hspace{1cm} (13)
From (13), it is found that the probability of GOQC occurrence for the metal with higher $n$ (different types of the grains) is greater than that of the metals with lower number of different types of the grains. For a given gate area, one should first calculate the number of the grains in the gate (based on average size of the grains) and then, based on the probability of the different GOs, calculate the probability of GOQC occurrence [using (13)]. As an example, for a 3T-FinFET with a channel length of 22 nm and a fin height of 66 nm, the number of grains in the gate (considering average sizes of the grains) is estimated to be 6, 16, 28, and 54 for TiN, MoN, WN, and TaN, respectively. Hence, probabilities of GOQC occurrence for these metal-gate devices are approximately 0.86, 0.99, 1, and 1, respectively. Note that a probability of 1 implies that GOQC always happens in WN and TaN gates. Fig. 13(a)–(d) shows the GOQC induced threshold voltage variation for the aforementioned metal-gate devices. As it can be observed, while the probabilities of GOQC occurrence in WN and TaN metal gates are the same, the GOQC effects in these devices are different (due to different $\Delta WF$ among the grains in these devices). Moreover, TiN and MoN are preferable gate materials for n- and p-type devices, respectively (due to the lower GOQC effect).

It is worth noting that (13) only gives the probability of GOQC happening while the effect of GOQC on the threshold voltage is given by (1)–(7). As an example, consider a TiN and a MoN metal-gate FinFET device, where there are two types of grains in each device (Table I). If it is assumed that there are identical number of grains on the gates, the probability of GOQC occurrence in two devices is the same and equal to

$$P_{\text{GOQC}} = 1 - \left(1 - \sum_{i=1}^{2} P_i \times (1 - P_i)\right)^{N_1/2}$$

where $p_1 = 0.4$, $p_2 = 0.6$, and $N_1$ is the number of grains on the gates of the two devices. While the probabilities of GOQC happening for the two devices are the same, the GOQC effect on the threshold voltage is more severe in MoN devices. The reason is that the difference between the WFs of the grains with different orientations in MoN devices is 600 meV, which is greater than that in TiN devices (200 meV).

Note that, in such small FinFET devices (in the sub-22-nm technology node), in order to consider the effects of different GOs, each device can be considered as a parallel combination of local channels (from source to drain). The threshold voltage of each local channel is calculated using (1)–(7). Subsequently, the threshold voltage of the device is analytically calculated from threshold voltages of the local channels (similar to the problem of calculating the threshold voltage of a multifin FinFET from the threshold voltages of individual fins [33], [34]. Fig. 14 shows the effect of the fin thickness on the GOQC effect. As it can be observed, thicker fins result in lower effect of the GOQC on the threshold voltage.

However, in FinFET technology, there is a design space for the appropriate fin thickness. The upper bound of this design space is determined by the allowed leakage current for a specific design. The leakage current of FinFET significantly increases for the thicker fins. Therefore, it is recommended that the $V_{th}$ should be less than $L_{\text{channel}}/2$. Moreover, when high-$k$ materials are employed, thinner fins are required to reduce the fringing-induced barrier lowering effect [28]–[31]. On the other hand, with very thin fins, the maximum height of the fin should also be reduced (to maintain the mechanical stability), which itself reduces the ON current of single-fin devices.

**B. GOQC Effect in 4T-FinFET, Nanowire FET, and Tri-Gate FET**

1) 4T-FinFET: In a 4T-FinFET [35], the front gate and back gate are separated from each other and can be independently biased. Therefore, as it can be inferred from (1), apart from the WVF, different biases of the front gate and back gate can also be the source of the transverse electric field. As an example, in a TiN metal gate, assuming an EOT of 1 nm and a $V_{FG}$ of 0 V and $V_{BG} = -0.5$ V, the maximum GOQC effect is 70 mV (corresponding to $\Delta WF = 700$ meV). As a result, in a 4T-FinFET, the GOQC effect is more severe than that in a 3T-FinFET.

2) Nanowire FETs: In a nanowire FET, four hypothetical gates can be assumed around the channel [Fig. 2(d)]. In a nanowire FET, GOQC happens when one of the hypothetical gates has a different WF, compared with those of the other gates. Therefore, the probability of GOQC occurrence in a nanowire FET is different from that in a 3T-FinFET (for the same gate material). The reason is that, in the 3T-FinFET, GOQC happens when two grains in two sides of the channel have different WFs. In a nanowire FET, however [see Fig. 2(d)], GOQC happens when one of the four hypothetical gates has a different WF, compared with those of the other hypothetical gates. As a result, generally, the probability of GOQC occurrence in a nanowire FET is higher than that in a 3T-FinFET (for identical channel length and approximately same gate area).

3) Tri-Gate FET: In a tri-gate FET [Fig. 15(a)], the channel is surrounded by three hypothetical gates (two side gates and a top gate). In the case where two side gates have different WFs, the GOQC effect is similar to the case in a 3T-FinFET, where the front gate and back gate have different WFs. However, the GOQC effect due to the different WF of the top gate requires specific attention. Since there is a thick oxide layer buried...
underestimation of the mean value and standard deviation of the delay by 4% and 12%, compared with the case, where GOQC is considered.

**IV. IMPLICATIONS FOR DIGITAL DESIGN**

In this section, we study the effect of GOQC induced threshold voltage fluctuations on the performance and reliability parameters of the digital circuits such as an inverter and a six-transistor static random-access memory (SRAM) cell. It is assumed that TiN is used as the gate material of the n-type devices and MoN is employed for the p-type devices. In all figures, “WFV” represents the case where classical effects of the WFV, along with SC variation, are considered. “WFV+GOQC” represents the case where the GOQC effect, along with the effects of other sources of variations, is considered. Fig. 16 shows the delay of an inverter, where neglecting the GOQC effect results in underestimation of the inverter delay (n-channel metal–oxide–semiconductor (MOS) and p-channel MOS are assumed to have identical (minimum) sizes with a channel length of 22 nm and a channel height of 66 nm).

Similarly, access time ($T_A$) (the time required to read the data from the SRAM cell) and write time ($T_W$) (the time required to write the data into the SRAM cell) can also be

\[
RNM = RNM_{NOM} + k_{R(AL)}\Delta V_{T(AL)} + k_{R(AR)}\Delta V_{T(AR)} + k_{R(PR)}\Delta V_{T(PR)} + k_{R(NL)}\Delta V_{T(NL)} + k_{R(NR)}\Delta V_{T(NR)}
\]

where $RNM_{NOM}$ is the nominal value of the $RNM$, and $k_R$ are the sensitivity coefficients of $RNM$ to the threshold voltages of the different transistors (Table III). Fig. 18 shows the $RNM$ probability density function, where neglecting the GOQC effect results in underestimation of the read failure probability.

**A. Effects of GOQC on the 3T-FinFET SRAM Characteristics**

The schematic of a six-transistor SRAM cell is shown in Fig. 17. NL and NR are pull-down (n-type) devices, PL and PR are pull-up (p-type) devices, and AL and AR are access devices (n-type). $Q_L$ and $Q_B$ represent the stored charge and its reciprocal value, respectively. BL and BLB represent the bitline and its reciprocal value, respectively, whereas WL represents the wordline voltage. In this paper, we investigate the effect of the GOQC on the reliability parameters of the SRAM cell, including read noise margin ($RNM$), access time ($T_A$), and write time ($T_W$). Monte Carlo simulations were performed with 1000 samples.

$RNM$: An analytical model is developed in [36], which considers the impact of random device variation on the $RNM$ of the SRAM cell, as given by
TABLE III
SENSITIVITY COEFFICIENTS OF THE STABILITY METRICS WITH RESPECT TO THE \( V_T \) VARIATIONS IN A SIX-TRANSISTOR 3T FinFET SRAM CELL, WHICH ARE EXPRESSED IN V\(^{-1}\) UNITS

<table>
<thead>
<tr>
<th></th>
<th>AL</th>
<th>NL</th>
<th>PL</th>
<th>AR</th>
<th>NR</th>
<th>PR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNMM</td>
<td>-0.95</td>
<td>1.07</td>
<td>-0.06</td>
<td>0.2</td>
<td>-2.52</td>
<td>0.69</td>
</tr>
<tr>
<td>Normalized (1/( T_A ))</td>
<td>0.63</td>
<td>-0.13</td>
<td>0.0005</td>
<td>0.001</td>
<td>0.03</td>
<td>0.0005</td>
</tr>
<tr>
<td>Normalized (1/( T_W ))</td>
<td>-0.27</td>
<td>0.32</td>
<td>-0.09</td>
<td>-0.65</td>
<td>-0.38</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Fig. 18. Neglecting the GOQC effect results in underestimation of the read failure probability.

Fig. 19. GOQC results in higher access time and, therefore, higher access failure.

expressed based on their nominal values and sensitivity coefficients to the threshold voltages of the different transistors (which are given in Table III). As shown in Figs. 19 and 20, GOQC increases the access time \( (T_A) \) and reduces the write delay \( (T_W) \) of the SRAM cell. The Write delay reduces since GOQC makes p-type devices weaker. (n-Type devices are less affected by GOQC [see Fig. 13(a) and (b)]).

V. CONCLUSION

In this paper, we have investigated a new intrinsic source of random threshold voltage variation in emerging ultra-thin-body metal-gate devices including FinFETs, tri-gate FETs, and nanowire FETs. We have shown that the randomness of the GO in metal gate devices causes high transverse electric fields in the channel, resulting in quantum confinement of the carriers, which is referred to as GOQC. GOQC is an exclusive source of threshold voltage variation in ultra-thin-body multigate devices. In bulk devices, due to thick body, and in SOI devices, due to thick back-gate oxide, the transverse electric field and, hence, the GOQC effect is negligible. GOQC is shown to be the main source of quantum confinement and quantum threshold voltage variation in ultra-thin-body multi-gate CMOS devices. GOQC effect is more severe for 4T-FinFETs compared to 3T-FinFETs and it is projected to be the worst in nanowire-FET devices.

The GOQC effect for different gate materials has been investigated. It has been shown that TiN and MoN are preferable gate materials for n- and p-type devices, respectively (due to lower GOQC effect). Our simulation results clearly indicate the significant effect of the GOQC on the characteristics of multigate CMOS devices and circuits. Therefore, the GOQC effect should be considered in transistor design as well as reliability and performance modeling of multigate CMOS devices and circuits.

REFERENCES


Seid Hadi Rasouli (S’07) received both the B.S. and M.S. degrees in electrical engineering from the University of Tehran, Tehran, Iran, in 2001 and 2004, respectively. He is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of California, Santa Barbara (UCSB), Santa Barbara, From 2004 to 2006, he was a Research Assistant and a VLSI Laboratory Instructor with the University of Tehran. He is with the Nanoelectronics Research Laboratory, with Prof. K. Banerjee, USCB. His current research interests include technology-circuit interactions in the design of low-power and robust digital integrated circuits using emerging technologies.

Kazuhiko Endo (M’99) received the Ph.D. degree in electrical engineering from Waseda University, Tokyo, Japan, in 1999.

From 1993 to 2003, he was with Silicon Systems Research Laboratories, NEC Corporation, Otsu, Japan, where he worked on research and development of multilevel interconnects and high-k gate stack technologies for ultra-large-scale integration. From August 1999 to August 2000, he was a Visiting Scholar with the Center for Integrated Systems, Stanford University, Stanford, CA. He is currently a Senior Researcher with the Silicon Nanoscale Devices Group, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan. His research interests include nanometer-scale manufacturing for aggressively scaled multigate devices in advanced very large scale integration technologies. Dr. Endo is a member of the IEEE Electron Devices Society and the Japan Society of Applied Physics. He was the recipient of the Best Paper Award at the 2013 Advanced Metallization Conference and at the 1998 Meeting of Japan Society of Applied Physics.

Jone F. Chen (S’93–M’98) received the B.S. degree from the National Cheng Kung University, Tainan, Taiwan, in 1990 and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1995 and 1998, respectively, all in electrical engineering. Since 1999, he has been with the Institute of Microelectronics, Department of Electrical Engineering, National Cheng Kung University, where he is currently a Professor. He is also currently a Visiting Professor with the Nanoelectronics Research Laboratory, Department of Electrical and Computer Engineering, University of California, Santa Barbara. His research interests include the reliability of deep-submicrometer and high-voltage metal–oxide–semiconductor devices.
Navab Singh (M’06–SM’09) received the M.Tech. degree in solid-state materials from Indian Institute of Technology Delhi, New Delhi, India, in 1995 and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2008.

After working for five years in the semiconductor industry in the area of lithography process technology, in July 2001, he joined the Institute of Microelectronics (IME), Agency for Science, Technology and Research (A∗STAR), Singapore. He is currently the Head with the Department of Postgraduate Research, IME, A∗STAR. He is the author or coauthor of more than 150 technical papers in archival journals and refereed conferences. His research interests include semiconductor nanowire technology and devices for green electronics and clean energy including storage, embedded memory devices, and nanoelectromechanical switch relays. His nanowire device papers have been selected for preconference publicity by conferences including the IEEE International Electron Devices Meeting and the International Conference on Solid-State Devices and Materials.

Dr. Singh was a recipient of the IEEE Electron Devices Society George E. Smith Award in 2007; the Singapore National Technology Award in 2008 for his outstanding contributions to the research and development of nanowire technology platform, enabling the realization of ultimately scaled complementary metal–oxide–semiconductor integrated circuits and a new class of electronic biosensors; and the A∗STAR TALENT award 2010 for leading, educating, and nurturing talents.

Kaustav Banerjee (S’92–M’99–SM’03) received the Ph.D. degree in electrical engineering and computer sciences from the University of California, Berkeley, in 1999.

In July 2002, he joined the faculty of the Department of Electrical and Computer Engineering, University of California, Santa Barbara (UCSB), where he has been a Full Professor since 2007. At UCSB, he directs the Nanoelectronics Research Laboratory and is an affiliated Faculty with the California NanoSystems Institute and the Institute for Energy Efficiency. Prior to joining UCSB, he was a Research Associate with the Center for Integrated Systems, Stanford University, Stanford, CA, from 1999 to 2001. He has also held a number of summer/Visiting positions at Texas Instruments, Dallas, from 1993 to 1997; École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, in 2001; and the Circuits Research Laboratories, Intel Corporation, Hillsboro, OR, in 2002. In December 2010, he held a Visiting Professorship at the Tokyo Institute of Technology, Tokyo, Japan. His research is chronicled in more than 220 journal articles, international conference proceeding papers, and book chapters, including more than 50 invited papers. He is also a Co-Editor of the book Emerging Nanoelectronics: Life With and After CMOS (Springer-Verlag, 2004). His technical ideas and innovations have seen widespread proliferation in both the industry and academia, as exemplified by his h-index of 36 as of April 2011. He has made a number of seminal contributions in the area of nanoscale integrated circuit (IC) interconnects and innovative interconnect solutions, including 3-D ICs and carbon-based interconnects, which have helped shape the semiconductor industry’s R&D efforts in those areas. His research interests include nanometer-scale issues in very large scale integration (VLSI), as well as circuits and systems issues in emerging nanoelectronics. He is currently involved in exploring the physics, technology, and applications of carbon nanomaterials for next-generation electronics, as well as energy harvesting and storage.

Dr. Banerjee has delivered more than 100 keynote/panel speeches, tutorials, and invited talks at major international forums and academic/research institutes around the world. He has served on the technical and organizational committees of several leading IEEE and Association for Computing Machinery conferences, including the IEEE International Electron Devices Meeting, International Reliability Physics Symposium, Design Automation Conference (DAC), International Conference on Computer-Aided Design, International Symposium on Quality Electronic Design, and International Conference on Simulation of Semiconductor Processes and Devices. From 2005 to 2008, he served as a member of the Nanotechnology Committee of the IEEE Electron Devices Society (EDS). He currently serves on the IEEE/EDS VLSI Technology and Circuits Committee. He was the recipient of the Best Paper Award at the Design Automation Conference in 2001, the Association of Computing Machinery Special Interest Group on Design Automation Outstanding New Faculty Award in 2004, the IEEE Micro Top Picks Award in 2006, and an IBM Faculty Award in 2008. He has been a Distinguished Lecturer of the IEEE Electron Devices Society since 2008.