SCIENCE CHINA
Information Sciences

• **Supplementary File** •

A Comprehensive Study of Device Variability of Sub-5nm Nanosheet Transistors and Interplay with Quantum Confinement Variation

Haowen Luo[1], Ruihan Li[1], Xiangshui Miao[1,2,3,4] & Xingsheng Wang[1,2,3,4*]

[1]*School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan* 430074, *China;*
[2]*Hubei Yangtze Memory Laboratories, Wuhan* 430205, *China;*
[3]*School of Integrated Circuits, Huazhong University of Science and Technology, Wuhan* 430074, *China;*
[4]*Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan* 430074, *China*

Appendix A Device structure description

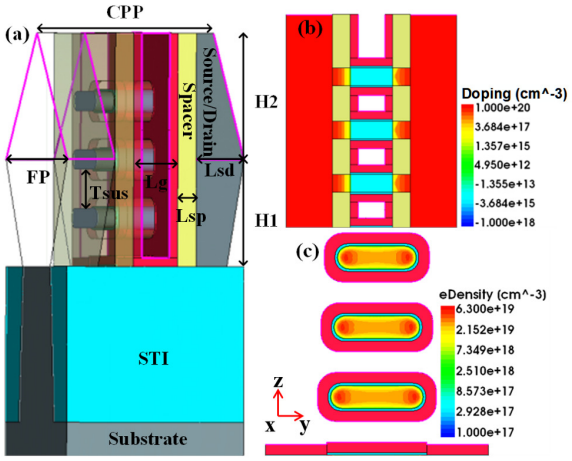

Figure A1 (a) The schematic view of the 3D Nanosheet Transistor in this study. (b) The cross view of doping concentration distribution along the channel. (c) The electron density distribution of the cross section perpendicular to the channel at V_g=0.7V and low drain voltage by TCAD simulation.

The structure of the simulated NST, consisted of three horizontally stacked nanosheets, is schematically illustrated in Figure A1(a). The key design parameters of three stacked sub-5nm technology NST referring to experimental data [1] and [2] are summarized in Table A1. The nominal NST channel is featured with a physical gate length (L_g) of 12nm and a nanosheet thickness (T_{sh}) of 5nm. The width (W_{sh}) difference between the top/middle/bottom nanosheets comes from the realistic trapezoid silicon fin. The gate metal sheet thickness variation reported in [3] could significantly alter the transfer characteristics of NSTs, but due to the irrelevance among the separated vertical sheet to sheet spacing (T_{space}) for top/middle, middle/bottom sheets and bottom sheet/parasitic channel, we assume the same T_{space} of 10nm [1] for 3 stacks in the nominal NST for easily defining the T_{space} in Sentaurus Structure Editor (SDE) during the calibration. The high-k metal gate stack is with equivalent oxide thickness (T_{ox}) of 0.9nm. The formation of epitaxial source/drain region comes from the different growth rates on the top and side surfaces due to the difference of crystal direction [4], whose side width is limited by Fin Pitch (FP). The constant doping ($10^{20} cm^{-3}$) in the epitaxial source/drain, the Gaussian doping profile in the extension region and un-doped ($10^{15} cm^{-3}$) channel are assumed (as shown in Figure A1(b)). The source/drain contact resistivity referring to the typical value [5] is assumed to be $2.5 \times 10^{-9} \Omega.cm^2$. Figure A1(c) shows the electron density distribution of three horizontal nanosheet channel cross-sections at V_g=0.7V and low drain voltage by TCAD simulation, which will be discussed in detail in the following. In this work, all simulations are carried out with 3D Sentaurus TCAD tools [6],'atomistic' simulator GARAND [7] and 2D PS simulator [8].

Appendix B Device calibration process description

Firstly, the 3D horizontally stacked n-channel NST based on the sub-5nm device parameters is established in the 3D Sentaurus TCAD simulation platform. The mobility model with modified high field saturation velocity, doping dependence and thin layer mobility models is utilized for channel carrier transport. Meanwhile, Shockley-Read-Hall Recombination and Avalanche Generation models also are included.

Table A1 PARAMETERS OF THE SIMULATED NST

Device Parameters	Value
Gate Length (L_g)	12nm
Top/Middle/Bottom Nanosheet Width (W_{sh})	17/18.5/20nm
Nanosheet Thickness (T_{sh})	5nm
Inner Spacer Length (L_{sp})	5nm
Source/Drain Length (L_{sd})	13nm
Vertical Sheet to Sheet Spacing (T_{sus})	10nm
Fin Pitch (FP)	48nm
Contacted Poly Pitch (CPP)	48nm
Equivalent Oxide Thickness (T_{ox})	0.9nm
S/D Epitaxy N-Type Doping Concentration	1e20cm^{-3}
Channel Doping Concentration (undoped)	1e15cm^{-3}
Substrate P-Type Doping Concentration	1e18cm^{-3}
Contact Resistivity (ρ_c)	$2.5 \times 10^{-9}\Omega.cm^2$
H_1	27.5nm
H_2	32.5nm
V_{dd}	0.7V

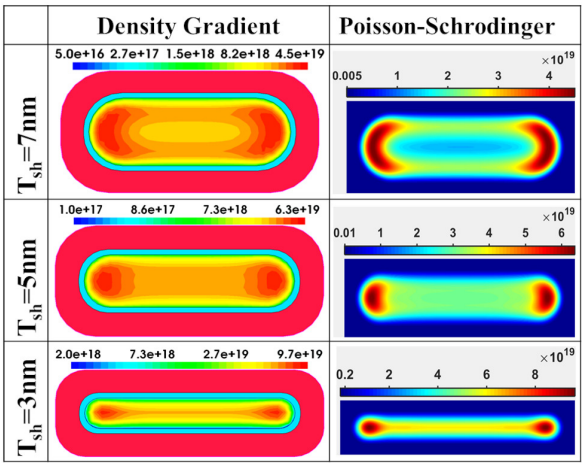

Figure B1 2D cross sectional charge distribution in different thick (T_{sh}=3nm, 5nm, 7nm) nanosheets (W_{sh}=20nm) compared between TCAD simulation based on density gradient quantum corrections and PS simulation at V_g=0.7V and low drain voltage.

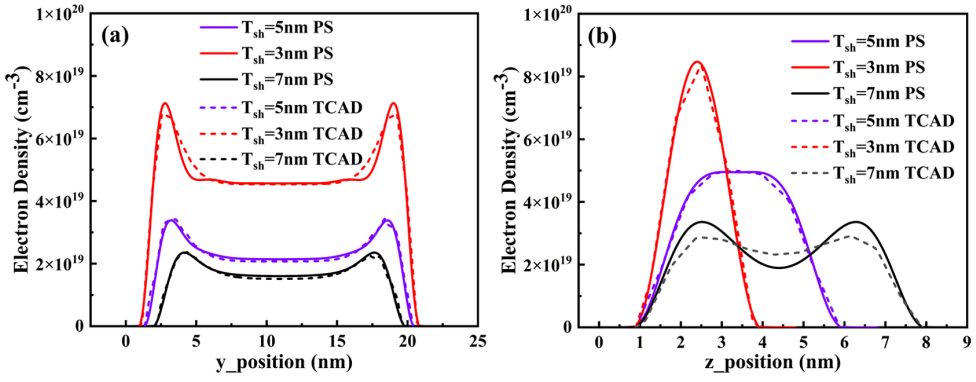

Figure B2 1D charge distributions along the horizontal direction in the up side (a) and along the vertical direction in the left corner (b) for different thick (T_{sh}=3nm, 5nm, 7nm) nanosheets (W_{sh}=20nm) compared between TCAD simulation and PS simulation at V_g=0.7V and low drain voltage.

Then, the quantum confinement effect plays a critical role in the channel carrier distribution, which will seriously affect the electrostatic performance of the NST, so it is also considered in the model. The direct solutions of Schrodinger equation or Non-Equilibrium Green's Functions (NEGF) are computationally heavy, so the density gradient (DG) quantum correction solution is

* Corresponding author (email: xswang@hust.edu.cn)

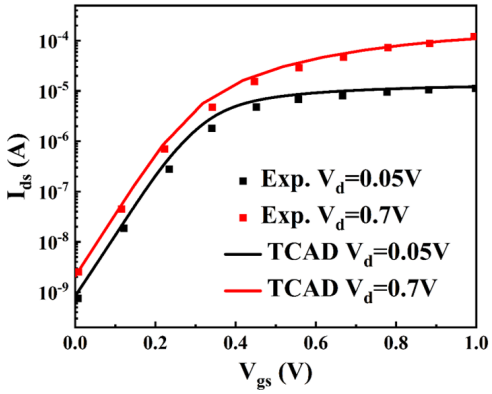

Figure B3 I_{ds}-V_{gs} characteristics of three stacked nanosheet transistors compared between TCAD simulation and the experimental data [1].

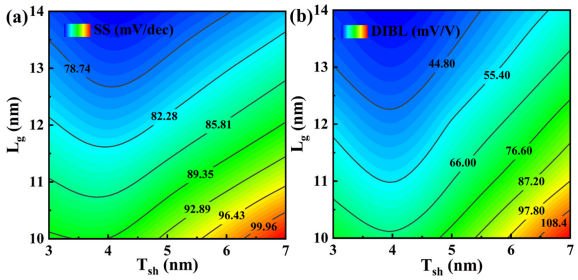

Figure C1 2D maps of (a) SS and (b) DIBL across the whole DoE space by TCAD simulation.

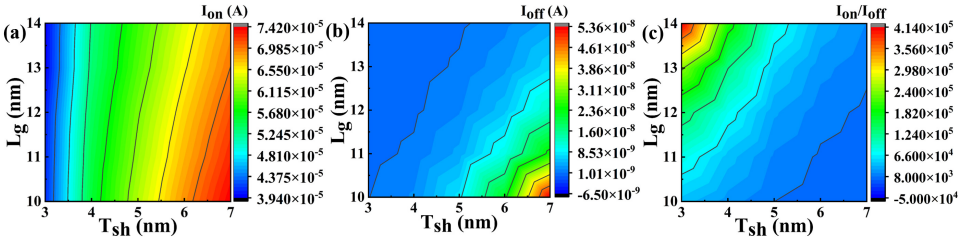

Figure C2 2D maps of (a) I_{on}, (b) I_{off} and (c) I_{on}/I_{off} across the whole DoE space by TCAD simulation.

widely used [9], but the DG formalism introduces an additional quantum correction term and is only an approximation to the Schrodinger equation. Hence, to improve the computational efficiency and accurately capture the quantum confinement effect, drift diffusion (DD) model coupled with the DG quantum correction in TCAD simulation has to be strictly calibrated against 2D PS simulation [10]. The DG quantum correction is calibrated here by fine-tuning model parameter γ [11] to the Poisson-Schrodinger charge distribution. Figure B1 shows the electron density distribution comparison in different thick NST channel cross-sections (W_{sh}=20nm) between TCAD and 2D PS simulations at V_g=0.7V and low drain voltage. Furthermore, 1D charge distributions along the horizontal direction ('y' direction) in the up side and along the vertical direction ('z' direction) in the left corner for different thick nanosheets compared between TCAD and 2D PS simulations at V_g=0.7V and low drain voltage are presented in Figure B2(a) and (b) respectively. It states more clearly that the DG quantum correction solution is accurately calibrated by 2D PS simulation in a wide range of NST channel thickness.

Finally, the device I_{ds}-V_{gs} characteristics by TCAD simulation are compared with the IBM experimental data [1] presented in Figure B3, illustrating an excellent agreement. The well-calibrated nominal NST will provide a baseline for the following research and analysis in this paper.

Appendix C Process induced global variation description

Since the difficulty and uncertainty of critical dimension (CD) control during the NST manufacturing including multilayer channel epitaxy, gate patterning and channel release [1], the key electrical performance parameters of NSTs with different nanosheet thickness and gate length are explored, including SS, DIBL, I_{on}, I_{off} and on-state current to off-state current ratio (I_{on}/I_{off}). The results will provide very useful references of NST's channel structure size selection for device designers to improve the device performance further. To cover the process variation, the 3σ process-induced long-range variations of ΔL_g=$\pm 2nm$ and ΔT_{sh}=$\pm 2nm$ are assumed to form a $5 \times 5 = 25$ node Cartesian product space. Each node represents a possible and specific case happening in the NST manufacturing.

Firstly, the 2D maps of SS and DIBL across the whole Design of Experiment (DoE) space by TCAD simulation are presented in Figure C1(a) and (b), respectively. Apart from that the SS and DIBL gradually reduces with the gate length increasing, it is also

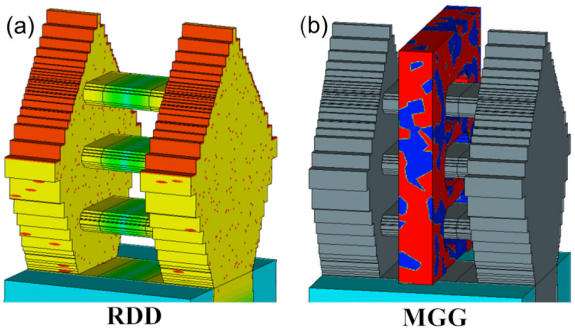

Figure D1 The 3D NST structure by 'atomistic' simulator GARAND with RDD generated in source/drain and extension regions (a) and MGG with average TiN grain-diameter of 5nm (b).

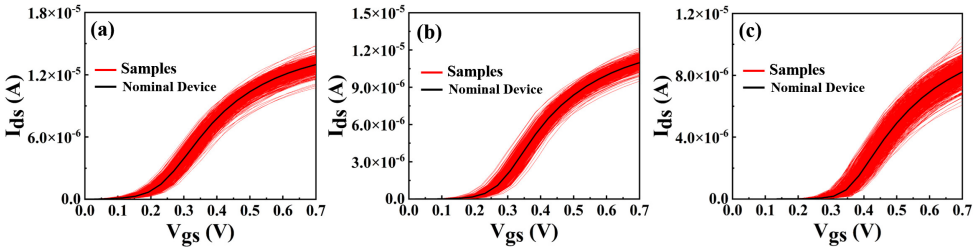

Figure D2 The full electrical transfer characteristics simulation of NSTs with the 10nm gate length and 7nm thick channel (a), 12nm and 5nm (b), 14nm and 3nm (c) including RDD, SER with RMS=0.45nm and Λ=20nm and MGG with average TiN metal grain diameter of 5nm at V_d=0.05V.

clearly found that the key figures of merit are the best for about 4nm thick NSTs and if the nanosheet thickness reduces further, the device performance parameters will drop sharply and even be unacceptable up to 3nm. The reason can be explained from the NST channel cross section charge affected by the strong quantum confinement effect induced by the thin channels and as shown in Figure B2(b), the peak of electron density in vertical direction is a 'flat top' peak for 5nm thick nanosheets but becomes a curved peak completely for 3nm thick nanosheets. At the moment, the electrons are firmly confined to the vertical center of the nanosheet so that the very strong quantum confinement leads to a sharp drop in device performance.

Then, we present the 2D maps of on-state and off-state currents across the whole DoE space by TCAD simulation in Figure C2(a) and (b), respectively. It is clearly seen from Figure C2(a) that the on-state current value is larger for the NSTs with short gate length and large nanosheet thickness due to the channel cross-section area increasing. In addition, for NSTs of less than 4nm thick, the on-state current value is very small so that it only has an insignificant increase with the gate length reduction. Meanwhile, from Figure C2(b), we can find that the off-state current also is very small for the NSTs with long gate length and small nanosheet thickness, which has the better gate control over the channel. At last, we calculate the on-state current to off-state current ratio across the whole DoE space presented in Figure C2(c). It clearly illustrates that the I_{on}/I_{off} value is larger for the nanosheets with long gate length and small nanosheet thickness, which has a higher ability to regulate the conduction current.

According to the global distributions of process parameters provided by measurements and last process generation, the global process variation induced performance parameter distributions can be obtained through the obtained response of device figures of merit to the long-range process variation in Figure C2.

Appendix D Statistical variability results discussion

In this section, the statistical variability, introduced by the three traditional local variability sources including RDD, nanosheet edge roughness (SER) and MGG, which is combined with quantum confinement variation effect is presented and analyzed comprehensively by TCAD simulation. It should be noted that different from global process variation, statistical variability derives from the intrinsic random nature of devices and demonstrates the local and short-range variation properties. Firstly, the 3D structure schematic view of NSTs with RDD and MGG by using the 'atomistic' drift-diffusion simulator GARAND are presented in Figure D1. Due to the un-doped channel, RDD is mainly generated in the source, drain and extension regions (as shown in Figure D1(a)). For SER [2], we only consider the top and bottom surface roughness of three stacked nanosheets modeled by using the 2-D autocorrelation function (ACF) [12] parameterized with correlation length (Λ) of 20nm and root mean square (RMS) of 0.45nm [2]. For simplicity, we assume that there is no correlation between the edge roughness lines along x- and y- direction. TiN gate metal grains with an average diameter of 5nm are assumed and MGG is modeled with two possible work functions spanning 0.2eV occupying 40% probability for $< 111 >$ grain and 60% probability for $< 200 >$ grain of occurrence [13] (as shown in Figure D1(b)). The Statistical Impedance Field Method (sIFM) is carried out to investigate the variation of the device characteristics.

To count and evaluate the impact of the statistical variability combined with quantum confinement variation effect on the threshold voltage (V_T) of NSTs, our method is to add the statistical variability sources to three cases of the DoE space, which are with the gate length of 10nm and nanosheet thickness of 7nm (fast corner), 12nm and 5nm (typical corner) and 14nm and 3nm (slow corner) according to the difference of the gate driven current I_{on}. At each node, ensembles of 1000 sample devices are used to simulate and obtain the statistical distribution results of device parameters. Figure D2 shows the I_{ds}-V_{gs} characteristics of 1000 samples simulated with the combination of three statistical variability sources at three nodes of the DoE space at V_d=0.05V. The threshold voltages extracted from the transfer characteristic curves of 1000 microscopically different samples with individual statistical variability source and their combination at the three nodes of DoE space are shown in Figure D3. It is clearly seen

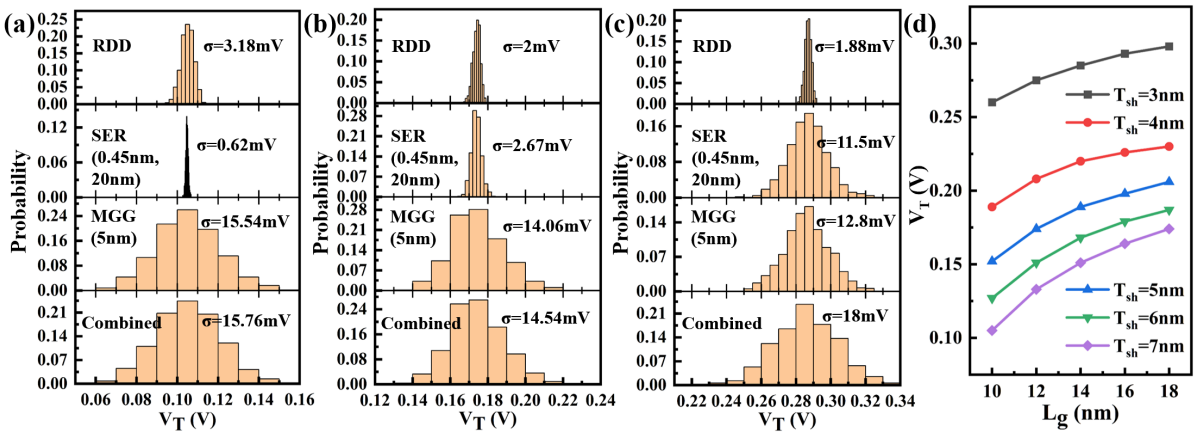

Figure D3 The threshold voltage (V_T) distribution of NSTs with different local variability sources at three nodes of the DoE space, which are with the 10nm gate length and 7nm thick channel (a), 12nm and 5nm (b), 14nm and 3nm (c) at V_g=0.7V. (d) The gate length dependence of threshold voltage (V_T) for different thick NST.

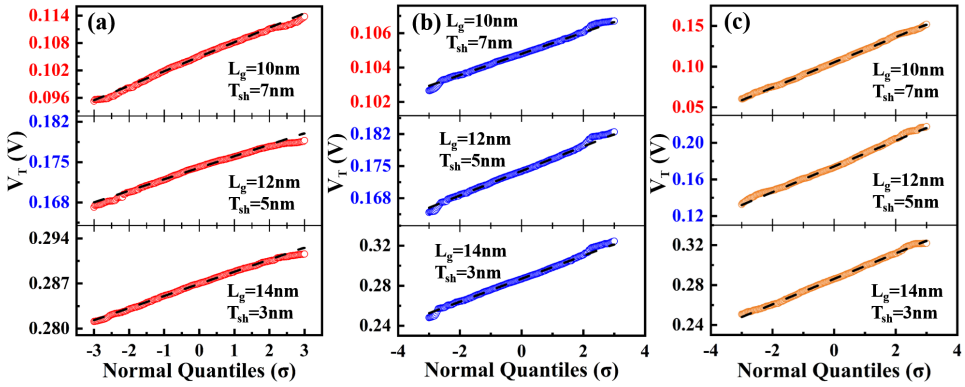

Figure D4 The Q-Q test on threshold voltage (V_T) distribution due to RDD generated in source/drain and extension regions (a), SER parametrized with RMS (0.45nm) and correlation length (20nm) (b) and MGG with average TiN grain-diameter (5nm) (c).

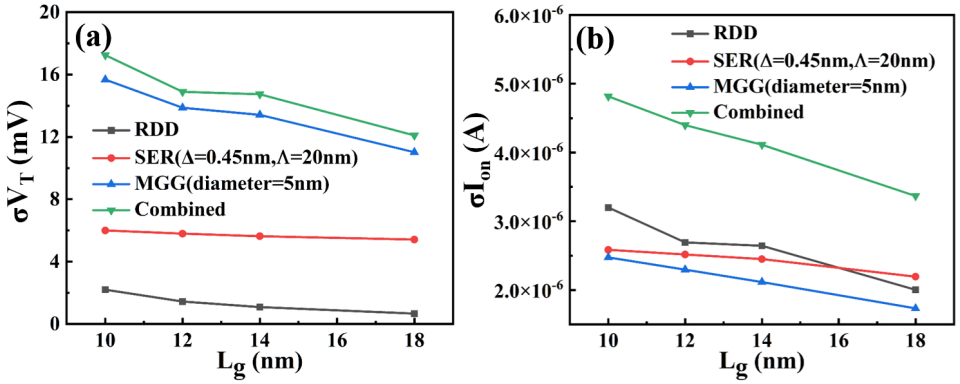

Figure D5 The gate length dependence of σV_T (a) and σI_{on} (b) for the 4m thick NSTs with different statistical variability sources.

that the standard deviation of threshold voltage can reach up to 15.76mV, 14.54mV and 18mV for the samples with L_g/T_{sh} of 10nm/7nm, 12nm/5nm and 14nm/3nm with three statistical variability sources combined, respectively.

Although RDD effects are suppressed due to the un-doped channel, the statistical variability induced by RDD combined with the strong quantum confinement effect, especially in extension region still results in unneglectable threshold voltage fluctuation. For SER, we can see clearly that the standard deviation of threshold voltage for the NST with L_g/T_{sh} of 14nm/3nm is much larger than the other two cases, which illustrates that the 14nm gate length and 3nm thick NST has the most serious threshold voltage variation because of extremely thin nanosheet and strong sensitivity to quantum confinement. As the threshold voltage dependent on the gate length of different thick NST is shown in Figure D3(d), it is illustrated clearly that with thinner nanosheet thickness, the gate length has smaller influence on V_T and the change of threshold voltage is more sensitive to nanosheet thickness variation

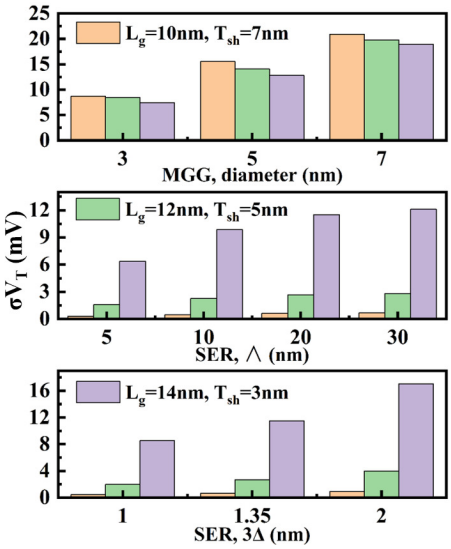

Figure D6 σV_T dependence of V_T variability respectively on MGG average grain diameter (top graph), SER correlation length (Λ) (middle graph) and SER root mean square (RMS, Δ) (bottom graph) for the NSTs at three nodes of DoE space.

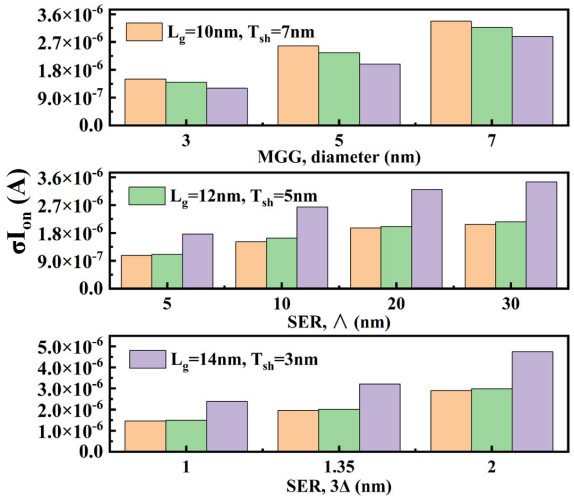

Figure D7 σI_{on} dependence of I_{on} variability respectively on MGG average grain diameter (top graph), SER correlation length (Λ) (middle graph) and SER root mean square (RMS, Δ) (bottom graph) for the NSTs at three nodes of DoE space.

compared with the gate length variation. Meanwhile, the SER-induced variability also results in the large overall threshold voltage variation of 3nm thick NSTs. As we move from Figure D3(a) to (c), it is found that the MGG-induced V_T variability is suppressed due to the increase in the gate area and for the fixed grain size of 5nm, the averaging effect [14] will be weakened. Compared to the SER-induced large V_T variability for 3nm thick NSTs, the strong quantum confinement for thin NSTs does not add a considerable contribution to the MGG-induced V_T variability.

Correspondingly, the normal Q-Q test on the threshold voltage distribution due to RDD in Figure D4(a) shows that it produces a skewed threshold voltage distribution. For the threshold voltage distribution due to SER, Figure D4(b) shows that it also produced a skewed distribution with a slightly raised right tail due to the asymmetrical sensitivity of threshold voltage to the changes in nanosheet thickness (as shown in Figure D3(d)). In more detail, Figure D3(d) shows that with T_{sh} gradually reduced, the equal thickness reduction results in a larger increment in V_T due to the non-linear channel thickness-dependent quantum confinement effect for the thin NSTs. Due to the small average TiN grain-diameter of 5nm, the Q-Q plots for V_T distribution due to MGG are shown as approximate Gaussian distribution without the flat tail for the three cases of NSTs in Figure D4(c). MGG renders itself still the dominant statistical variability source compared with RDD and SER but it is worth noting that SER brings the severe threshold voltage fluctuation (43.93%) for the NSTs with L_g/T_{sh} of 14nm/3nm because of extremely thin nanosheet and strong sensitivity to QCV.

For further checking the effect of statistical variability sources on the key performance parameters of 4nm thick NSTs as the best performing device mentioned in Appendix C, The gate length dependence of σV_T and σI_{on} for 4nm thick NSTs are shown in Figure D5(a) and (b), respectively. Due to the low RMS of 0.45nm for the 4nm thick nominal NST and the range of gate length varying close to the correlation length of 20nm, Figure D5(a) shows that the SER-induced σV_T slowly decreases with the gate length gradually increasing. Figure D5(b) makes it clear that the gate length dependence of σI_{on} has different distribution features, which can be seen that RDD results in the largest I_{on} variation due to the short channel effect.

In the following, the dependence of σV_T and σI_{on} on MGG average grain diameter, SER correlation length (Λ) and SER root

mean square (RMS, $\triangle$) for the NSTs at three corner nodes of DoE space is presented in Figure D6 and Figure D7, respectively. It is found that the MGG-induced V_T and I_{on} variability can be suppressed with the average TiN metal grain size reduction. The SER with larger Λ and RMS all brings obviously more V_T and I_{on} fluctuations, but SER-induced variability is more sensitive to RMS. Meanwhile it seems that the σV_T and σI_{on} are linearly proportional to the RMS of SER but are saturated for large Λ.

References

1 N. Loubet et al. Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET. In: Proceedings of Symposium on VLSI Technology, Kyoto, 2017. T230-T231.
2 Amita, A. Gorad and U. Ganguly. Analytical Estimation of LER-Like Variability in GAA Nano-Sheet Transistors. In: Proceedings of International Symposium on VLSI Technology, Systems and Application (VLSI-TSA), Hsinchu, Taiwan, 2019. 1-2.
3 P. Harsha Vardhan, et al. Threshold Voltage Variability in Nanosheet GAA Transistors. IEEE Transactions on Electron Devices, 2019, 66: 4433-4438.
4 A. Kaneko et al. Sidewall transfer process and selective gate sidewall spacer formation technology for sub-15nm finfet with elevated source/drain extension. In: Proceedings of IEEE International Electron Devices Meeting. IEDM Technical Digest., Washington, DC, 2005. 844-847.
5 N. Breil et al. Highly-selective superconformal CVD Ti silicide process enabling area-enhanced contacts for next-generation CMOS architectures. In: Proceedings of Symposium on VLSI Technology, Kyoto, 2017. T216-T217.
6 Sentaurus user manual, Synopsys, P-2019.03.
7 Garand user manual, Synopsys, P-2019.03.
8 M. Rother, 2D Schrodinger Poisson solver AQUILA (https://www.mathworks.com/matlabcentral/fileexchange/3344-2d-schroedinger-poisson-solver-aquila), MATLAB Central File Exchange. Retrieved January 12, 2020.
9 Brown, A.R., Watling, J.R., Roy, G. et al. Use of density gradient quantum corrections in the simulation of statistical variability in MOSFETs. J Comput Electron, 2010, 9: 187-196.
10 A. Asenov, G. Slavcheva, A. R. Brown, J. H. Davies and S. Saini. Quantum mechanical enhancement of the random dopant induced threshold voltage fluctuations and lowering in sub-0.1micron MOSFETs. In: Proceedings of International Electron Devices Meeting. Technical Digest (Cat. No.99CH36318), 1999. 535-538.
11 Ancona, M.G. Density-gradient theory: a macroscopic approach to quantum confinement and tunneling in semiconductor devices. J Comput Electron, 2011, 10: 65-97.
12 J. Park, H. Lee, S. Oh and C. Shin. Design for Variation-Immunity in Sub-10-nm Stacked-Nanowire FETs to Suppress LER-induced Random Variations. IEEE Transactions on Electron Devices, 2016, 63: 5048-5054.
13 X. Wang et al. Statistical Threshold-Voltage Variability in Scaled Decananometer Bulk HKMG MOSFETs: A Full-Scale 3-D Simulation Scaling Study. IEEE Transactions on Electron Devices,2011,58: 2293-2301.
14 A. Sudarsanan, et al. Superior Work Function Variability Performance of Horizontally Stacked Nanosheet FETs for Sub-7-nm Technology and Beyond. In: Proceedings of 4th IEEE Electron Devices Technology & Manufacturing Conference (EDTM), Penang, Malaysia, 2020. 1-4.