

- random access memory," *IEEE Trans. Nanotechnol.*, vol. 2, no. 4, pp. 335–340, Dec. 2003.
- [5] B. Eitan, P. Pavan, I. Bloom, E. Aloni, A. Frommer, and D. Finzi, "NROM: A novel localized trapping, 2-bit nonvolatile memory cell," *IEEE Electron Device Lett.*, vol. 21, no. 11, pp. 543–545, Nov. 2000.
- [6] H. I. Hanafi, S. Tiwari, and I. Khan, "Fast and long retention-time nano-crystal memory," *IEEE Trans. Electron Devices*, vol. 43, no. 9, pp. 1553–1558, Sep. 1996.
- [7] Y. Shi, K. Saito, H. Ishikuro, and T. Hiramoto, "Effects of traps on charge storage characteristics in metal-oxide-semiconductor memory structures based on silicon nanocrystals," *J. Appl. Phys.*, vol. 84, no. 4, pp. 2358–2360, Aug. 1998.
- [8] Y. Shi, K. Saito, H. Ishikuro, and T. Hiramoto, "Effects of interface traps on charge retention characteristics in silicon-quantum-dot-based metal-oxide-semiconductor diodes," *Jpn. J. Appl. Phys.*, vol. 38, no. 1B, pp. 425–428, Jan. 1999.
- [9] M. Zacharias, J. Heitmann, R. Scholz, U. Kahler, M. Schmidt, and J. Blasing, "Size-controlled highly luminescent silicon nanocrystals: A SiO/SiO₂ superlattice approach," *Appl. Phys. Lett.*, vol. 80, no. 4, pp. 661–663, Jan. 2002.
- [10] K.-J. Lee, T.-D. Kang, H. Lee, S. H. Hong, S.-H. Choi, T.-Y. Seong, K. J. Kim, and D. W. Moon, "Optical properties of SiO₂/nanocrystalline Si multilayers studied using spectroscopic ellipsometry," *Thin Solid Films*, vol. 476, no. 1, pp. 196–200, Apr. 2005.
- [11] K. J. Kim, D. W. Moon, S.-H. Hong, S.-H. Choi, M.-S. Yang, J.-H. Jhe, and J. H. Shin, "In situ characterization of stoichiometry for the buried SiO_x layers in SiO_x/SiO₂ superlattices and the effect on the photoluminescence property," *Thin Solid Films*, vol. 478, no. 1/2, pp. 21–24, May 2005.
- [12] B. De Salvo *et al.*, "How far will Silicon nanocrystals push the scaling limits of NVM technologies?" in *IEDM Tech. Dig.*, 2003, pp. 597–600.
- [13] R. A. Rao *et al.*, "Silicon nanocrystal based memory devices for NVM and DRAM applications," *Solid State Electron.*, vol. 48, no. 9, pp. 1463–1473, Sep. 2004.
- [14] H. I. Hanafi, S. Tiwari, and I. Khan, "Fast and long retention-time nanocrystal memory," *IEEE Trans. Electron Devices*, vol. 43, no. 9, pp. 1553–1558, Sep. 1996.
- [15] Y.-C. King, T.-J. King, and C. Hu, "MOS memory using germanium nanocrystals formed by thermal oxidation of Si_{1-x}Ge_x," in *IEDM Tech. Dig.*, 1998, pp. 115–118.
- [16] C. Y. Ng, T. P. Chen, M. Yang, J. B. Yang, L. Ding, C. M. Li, A. Du, and A. Trigg, "Impact of programming mechanisms on the performance and reliability of nonvolatile memory devices based on Si nanocrystals," *IEEE Trans. Electron Devices*, vol. 53, no. 4, pp. 663–667, Apr. 2006.

Strain Engineering to Improve Data Retention Time in Nonvolatile Memory

R. Arghavani, N. Derhacopian, V. Banthia, M. Balseanu, N. Ingle, H. M'Saad, S. Venkataraman, E. Yieh, Z. Yuan, L.-Q. Xia, Z. Krivokapic, U. Aghoram, K. MacWilliams, and S. E. Thompson

Abstract—Experimental data show that tensile stress improves and compressive stress degrades retention time for nonvolatile memory (NVM) devices. External mechanical tensile stress and compressive stress are introduced into the NVM floating-gate and nitride trap based memories via four-point wafer bending. The enhanced retention time under tensile stress results from stress-altered changes in the SiO₂/Si barrier height and out-of-plane conductivity mass for floating-gate memories and from changes in the trap activation energy in nitride based memories.

Index Terms—Compressive stress, data retention, nonvolatile memory (NVM), read current, strain, tensile stress.

I. INTRODUCTION

Conventional CMOS logic device scaling beyond 90-nm node requires local strain engineering to increase electron and hole mobilities and enhance transistor performance [1]–[4]. In addition to increasing mobility, which has some advantages in nonvolatile memory (NVM), strain also alters the SiO₂/Si barrier height, the conductivity mass perpendicular to the SiO₂ interface, and electron trap energy levels, all of which will alter the retention time.

The charge retention characteristics of an NVM cell are set by the leakage of stored electrons under the specified environmental and operational conditions including power-off mode. For the NVM NAND cell, electrons are generally stored on a floating poly-Si gate, while for NOR, electrons are stored on a floating gate or in nitride traps. With NVM technology in the nanoscale regime, many process steps from cell isolation formation to packaging introduce both intentional and unintentional strain in the floating gate, insulating layers, and channel. There have been early reports on compressive stress being important to minimize since it degrades the retention time [5], [6]. In this brief, we investigate the effect of tensile and compressive stress on floating-gate and nitride trap based NVM retention time. This has applications for intentionally engineering "good" types of strain and removing "bad" types from an NVM process flow.

II. EXPERIMENTAL SETUP

A flexure-based four-point bending setup is used in applying large (up to 600 MPa) tensile and compressive stress, as shown in Fig. 1(a). A flexure is a long beam with a notch on one end, which gives the beam only one degree of freedom and suppresses the others [7]. The setup has a system of eight such flexure beams, which provide uniaxial upward displacement to the bottom rods in a traditional four-point

Manuscript received May 25, 2006; revised October 17, 2006. The review of this brief was arranged by Editor S. Kimura.

R. Arghavani, V. Banthia, M. Balseanu, N. Ingle, H. M'Saad, S. Venkataraman, E. Yieh, Z. Yuan, L.-Q. Xia, and K. MacWilliams are with Applied Materials, Santa Clara, CA 95054 USA.

N. Derhacopian is with Adesto Technologies, Belmont, CA 94002 USA.

Z. Krivokapic is with Advanced Micro Devices, Sunnyvale, CA 94088 USA.

S. E. Thompson is with the College of Engineering, University of Florida, Gainesville, FL 32611 USA.

Color versions of one or more of the figures in this brief are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TED.2006.888827

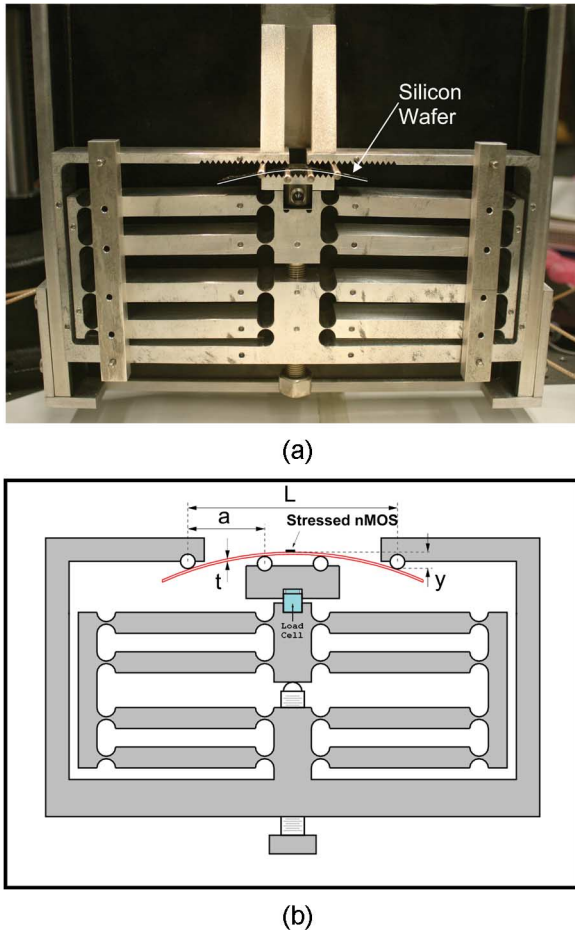


Fig. 1. Flexure-based wafer bending jig: (a) photograph and (b) schematic diagram.

bending setup, as shown in the schematic in Fig. 1(b). This eliminates error due to the uneven rotation of two screws present in traditional four-point bending setups and is essential for achieving high stress. The surface stress along the channel direction is calculated from the relationship $\sigma = Eyt/\{2a(L/2 - 2a/3)\}$, where E is the Young's Modulus ($E = 168$ GPa for $\langle 110 \rangle$ channel), y is the sample vertical displacement, t is the total thickness of the sample, L is the length of the sample between the two outer rods, and a is the distance between the inner and outer rods, as shown in the schematic in Fig. 1(b). The jig was calibrated with a load cell under the mounting platform and a strain gauge mounted on to the surface of the wafer.

III. EXPERIMENTAL DATA: STRAIN EFFECT ON DATA RETENTION

To understand the effect of strain on retention time, we perform two experiments. Both experiments apply longitudinal to the channel tensile and compressive stress on industrial samples. The first experiment measures the change in the stress-altered leakage current for thick SiO_2 tunnel oxides. The second experiment directly measures the stress-altered retention of store electrons in nitride traps.

In the first experiment, the change in strain-altered electron tunneling current is quantified by introducing mechanical compressive and tensile stress into the gate, oxide, and channel of a $1\text{-}\mu\text{m}$ n-MOSFET with a SiO_2 oxide thickness of 55 \AA , which is typically used in NVM. The data are taken at room temperature on ten samples with a repeatability of within 5%. The channel direction of the n-MOSFET is $\langle 110 \rangle$, and a mechanical stress of as large as 600 MPa is introduced.

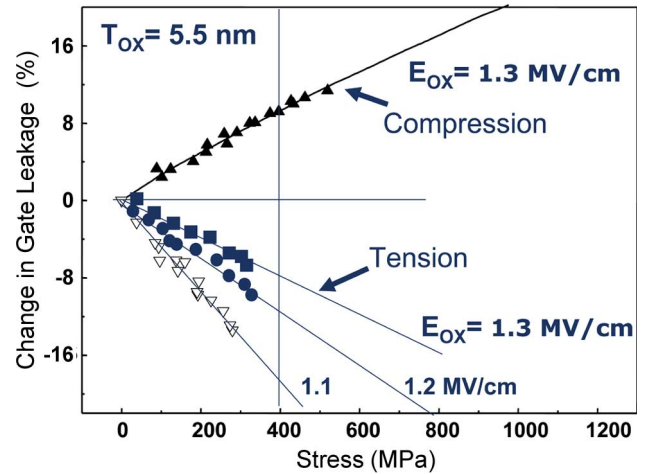


Fig. 2. Strain-induced change in gate tunneling current. Note that uniaxial tensile stress reduces the electron tunneling current, while compressive stress increases the current.

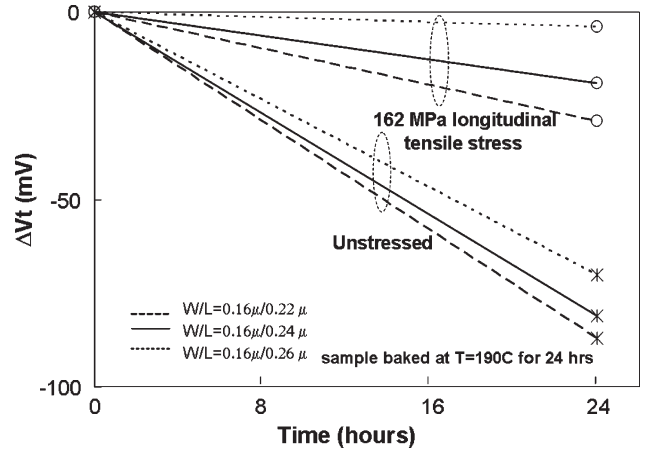


Fig. 3. NVM cell data retention (190°C for 24 h) with and without externally applied mechanical 162-MPa tensile stress. Note the improved retention time for devices under tensile stress.

The results are plotted in Fig. 2, and we observe that the SiO_2 tunnel leakage current increases under compressive stress and decreases under tensile stress. We also note a strong electric field dependence of the leakage, in which the stress-altered change in the leakage current is much larger at low vertical fields (retention condition) than at high electric fields (program condition).

In the second experiment, since it is difficult to directly measure the leakage current from nitride traps, we measure the change in retention time for the trapped electrons. Electrons are stored on the nitride based NVM cells of various sizes. All of the cell programming is done at room temperature using conventional programming conditions without any externally applied mechanical stress. Next, the samples are baked at 190°C for 24 h with and without externally applied tensile or compressive mechanical stress. For each experiment, the stress and unstressed samples are baked at the same time, and all devices come from the same wafer on neighboring dies. Figs. 3 and 4 plot the threshold voltage shift after the bake condition. For all device sizes, tensile stress is observed to significantly improve the retention of the trapped electrons, while the opposite is observed for compressive stress. The magnitude of stress used in the experiment of a few hundred megapascals is typical of stress intentionally/unintentionally induced on a Si channel of NVM-Flash cell via process steps.

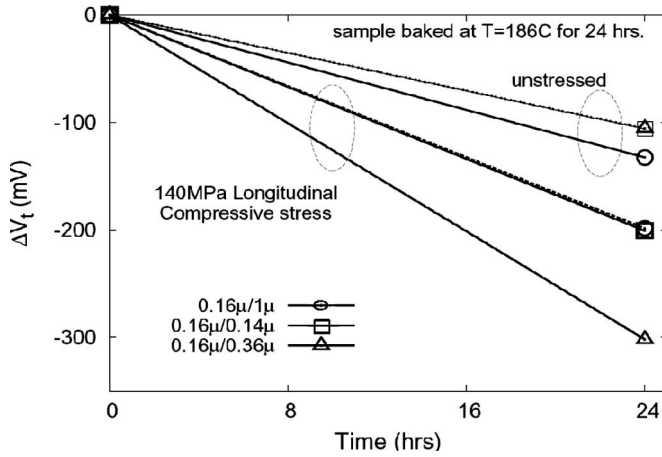


Fig. 4. NVM cell data retention (190 °C for 24 h) with and without externally applied mechanical 140-MPa compressive stress. Note the degraded retention time for devices under compressive stress.

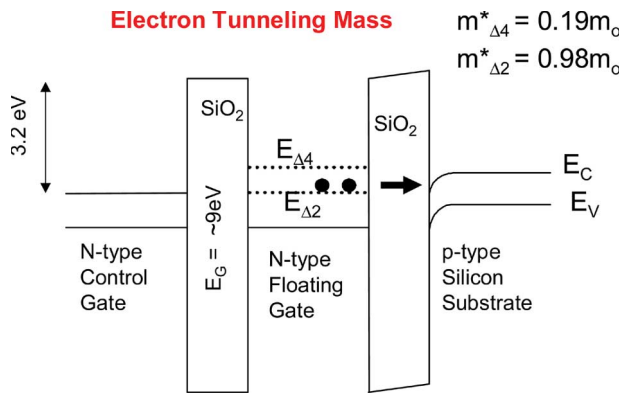


Fig. 5. Strain-induced splitting between Δ_2 and Δ_4 , resulting in an increased average conductivity effective mass in the direction of tunneling. The mass is increased due to increased electron population in Δ_2 valleys.

IV. DISCUSSION

Severe degradation in the retention time has been reported for mechanical stress, leading to the suggestion that process-induced mechanical stress needs to be eliminated [6]. However, in this brief, we see that it is compressive stress that degrades retention time, while tensile stress offers improvement. Improved retention time under tensile stress results from stress-altered changes in the SiO_2/Si barrier height and out-of-plane conductivity mass in floating-gate memories and from changes in the trap activation energy in nitride trap based memories. We will discuss each of these effects in more detail.

Strain alters tunnel oxide leakage currents by changing the SiO_2/Si barrier height and the electron conductivity effective mass. The barrier height is altered via strain-induced changes in the Si and SiO_2 electron affinity (see Fig. 5). The electron conductivity effective mass changes due to the valley repopulation and slight band warping for certain types of strain. Based on how tensile stress splits the conduction band, we should expect a reduction (improvement) in the magnitude of the tunneling current. This is understood by first discussing electron transport in unstressed silicon. For unstressed Si, the conduction band comprises six degenerate valleys. The degeneracy reflects the cubic symmetry of the silicon lattice. The effective mass for any direction is inversely proportional to the curvature of the energy dispersion relationship in that direction. Consequently, the effective mass of each ellipsoid is anisotropic, with the transverse mass (perpendicular to the axis) given by $m_t = 0.19m_0$ being significantly smaller than the

longitudinal mass (parallel to the axis) given by $m_l = 0.98m_0$, where m_0 is the free electron mass. For unstressed bulk silicon, the total electron conductivity mass m^* is obtained by adding the contributions of the six degenerate valleys and is given by

$$m^* = \left[\frac{1}{6} \left(\frac{2}{m_l} \right) + \left(\frac{4}{m_t} \right) \right]^{-1}.$$

Assuming a crystal orientation where $\langle 001 \rangle$ is the floating gate to the channel direction, in-plane tensile stress reduces the tunneling current by removing the degeneracy between the four in-plane valleys Δ_4 and the two out-of-plane valleys Δ_2 by splitting them in energy. Tensile stress lowers the energy of the Δ_2 valleys and results in preferential occupation of these states. The tunneling current in these states is significantly reduced due to m^* transforming to $m_l = 0.98m_0$ after enough band splitting to fully populate the Δ_2 valleys. Furthermore, the Δ_2 splitting increases the SiO_2/Si barrier height, which will also decrease the tunneling current. Which effect dominates depends on the magnitude of the oxide field [8]–[11] (Fig. 5).

Similarly, strain is known to alter the activation energy of SiO_2 electron traps and should be expected to alter the retention time. For nitride-based memories, the primary electron loss mechanism at elevated temperatures is the thermal detrapping of the stored electrons. The electron detrapping is proportional to $\exp(-Ea/kT)$, where Ea is the electron trap activation energy. Mechanical stress will shift the activation energy. Consistent with our measured data, compressive stress is observed to decrease the activation energy by ~ 40 meV for 100 MPa. Though no data are reported for tensile stress, it should have the opposite effect as that for compressive stress and should increase the activation energy.

V. CONCLUSION

We have experimentally measured using four-point mechanical wafer bending, the effect of tensile and compressive stress on floating-gate and nitride based NVM cells. We have shown experimentally that compressive stress degrades and tensile stress improves the retention time. Historically, process-induced stress in NVM process technologies has been eliminated due to deleterious effects. In this brief, we show that tensile stress is desirable. Recently, there have been several techniques to introduce tensile stress in logic technologies, which can also be used in NVM [12].

ACKNOWLEDGMENT

The authors would like to thank F. Moghadam, the Sr. VP of the Thin Films Group, Applied Materials, for his continuous support of this program.

REFERENCES

- [1] S. E. Thompson, G. Sun, K. Wu, J. Lim, and T. Nishida, "Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs," in *IEDM Tech. Dig.*, 2004, pp. 221–224.
- [2] H. S. Yang, R. Malik, *et al.*, "Dual stress liner for high performance sub-45 nm gate length SOI CMOS manufacturing," in *IEDM Tech. Dig.*, 2004, pp. 1075–1078.
- [3] P. Bai, C. Auth, *et al.*, "A 65 nm logic technology featuring 35 nm gate lengths, enhanced channel strain, 8 Cu interconnect layers, low- κ ILD and $0.57 \mu\text{m}^2$ SRAM cell," in *IEDM Tech. Dig.*, 2004, pp. 657–660.
- [4] R. Arghavani, Z. Yuan, N. Ingle, K.-B. Jung, M. Seamons, S. Venkataraman, V. Banthia, K. Lilja, P. Leon, G. Karunasiri, S. Yoon, and A. Mascarenhas, "Stress management in sub-90-nm transistor architecture," *IEEE Trans. Electron Devices*, vol. 51, no. 10, pp. 1740–1744, Oct. 2004.

- [5] A. Hamada and E. Takeda, "Hot-electron trapping activation energy in PMOSFET's under mechanical stress," *IEEE Electron Device Lett.*, vol. 15, no. 1, pp. 31–32, Jan. 1994.
- [6] Y. M. Park, J. S. Lee, M. Kim, M. K. Choi, T. K. Kim, J. I. Han, D. W. Kwon, W. K. Lee, Y. H. Song, and K. D. Suh, "The mechanical stress effects on data retention reliability of NOR Flash memory," in *IEDM Tech. Dig.*, 2001, pp. 711–715.
- [7] S. T. Smith, "Flexures: Elements of Elastic Mechanisms," CRC Press.
- [8] A. Gehring and S. Selberherr, "Modeling of tunneling current and gate dielectric reliability for nonvolatile memory devices," *IEEE Trans. Device Mater. Rel.*, vol. 4, no. 3, pp. 306–319, Sep. 2004.
- [9] A. Kolodny, S. Nieh, B. Eitan, and J. Shappir, "Analysis and modeling of floating gate EEPROM cells," *IEEE Trans. Electron Devices*, vol. ED-33, no. 6, pp. 835–844, Jun. 1986.
- [10] W. Zhao, A. Seabaugh, V. Adams, D. Jovanovic, and B. Winstead, "Opposing dependence of the electron and hole gate currents in SOI MOSFETs under uniaxial strain," *IEEE Electron Device Lett.*, vol. 26, no. 6, pp. 410–412, Jun. 2005.
- [11] X. Yang, J. Lim, G. Sun, K. Wu, T. Nishida, S. E. Thompson, *et al.*, "Strain-induced changes in the gate tunneling currents in p-channel MOSFETs," *IEEE Electron Device Lett.*, submitted for publication.
- [12] A. Al-Bayati, L. Washington, L.-Q. Xia, M. Balseanu, Z. Yuan, M. Kawaguchi, F. Nouri, and R. Arghavani, *Stress-Tunable Films Enhance Transistor Speed*, 26th ed. London, England: SEMICONDUCTOR FABTECH, 2005, pp. 84–88.

Performance Investigation of 50-nm Insulated-Shallow-Extension Gate-Stack (ISEGaS) MOSFET for Mixed Mode Applications

Ravneet Kaur, Rishu Chaujar, Manoj Saxena, and R. S. Gupta

Abstract—An extended study of electrical characteristics of 50-nm single-material-gate insulated-shallow-extension-gate-stack (ISEGaS) MOSFET is performed using ATLAS-2D. Incorporation of dual-material-gate architecture leads to the suppression of short channel effects along with the improvement in device intrinsic gain ($g_m \times R_{out}$), voltage gain (g_m/I_{DS}), and I_{on}/I_{off} ratio, thereby opening a new era of ISEGaS MOSFETs for mixed mode applications.

Index Terms—ATLAS-2D, dual material gate (DMG), energy-balance-transport (EBT), gate stack.

I. INTRODUCTION

Shortcomings of implanted pockets of high and low doping near the source/drain junctions lead to a new concept of dielectric pocket (DP) proposed by Jurczak *et al.* [1]. The buried dielectric spacer substituting the implanted pockets isolates vertical sidewalls of the deep junctions from the channel except the topmost region which constitutes the inversion layer. Thus, it can provide a means to curb the avalanche breakdown by preventing bulk punch through without reducing the depth of highly doped junction and without increasing the channel doping [1].

Manuscript received June 2, 2006; revised September 15, 2006. This work was supported by the University Grants Commission (UGC). The review of this brief was arranged by Editor T. Skotnicki.

R. Kaur, R. Chaujar, and R. S. Gupta are with the Semiconductor Devices Research Laboratory, University of Delhi, Dhaura Kuan, New Delhi 110021, India (e-mail: ravneetsawhney13@rediffmail.com; rishuchaujar@rediffmail.com; rsgu@bol.net.in).

M. Saxena is with the Department of Electronics, Deen Dayal Upadhyaya College, University of Delhi, Karampura, New Delhi 110015, India (e-mail: saxena_manoj77@yahoo.co.in).

Digital Object Identifier 10.1109/TED.2006.888722

In this brief, the electrical characteristics of 50-nm insulated-shallow-extension (ISE) gate-stack MOSFET are investigated by extensive simulation studies using ATLAS-2D [2]. Further, for minimizing short channel effects (SCEs), dual material gate (DMG) [3] with more advanced high- κ gate stack, as foreseen in the International Technology Roadmap for Semiconductors Road Map [4], has been incorporated onto the ISE architecture. L_c and L_s are the lengths of two laterally contacting gate materials of different work function— L_c serving as a control gate, having metal workfunction $q\Phi_{M1}$ and L_s as a screening gate, having workfunction $q\Phi_{M2}$, in DMG architecture. Work function difference of the two metal gates causes an abrupt change in the conduction band energy at the silicon surface providing an electric field peak in the channel [5] giving greater acceleration to the charge carriers and hence improves the carrier injection into the channel. For the feasibility of DMG architecture, several integration schemes have been suggested such as fully silicided metal gate [6]; metal wet etch process [7]; metal interdiffusion process [8], [9]; patterning and implanting a workfunction modifying species into one of the electrodes [10], and chemical mechanical polishing [11]. Further, combining the potential benefits of DP together with electric field tuned DMG and gate-stack architecture, optimum device structure is presented as an attractive design in view of system-on-chip realization, where digital mixed-signal base band and RF transceiver block are integrated.

II. RESULTS AND DISCUSSION

All device simulations are performed using energy-balance-transport model with relaxation time of 0.4 ps [12], [13], inversion layer Lombardi CVT mobility model [12], [13], Shockley Read-Hall and Auger recombination model. The simulated device described by the schematic cross section shown in Fig. 1 (inset) uses gate stack of a 2-nm-thick high- κ layer and a 1-nm-thin interfacial SiO_2 layer.

A. Surface Potential, Electric Field, and Electron Velocity

The penetration of drain electric field into the source region lowers the source-channel diffusion barrier resulting in poor drain-induced-barrier-lowering (DIBL). Fig. 1 shows step potential profile along the channel for DMG architecture that results in drain bias screening as compared to bulk and single material gate (SMG)-ISE. Upon decreasing Φ_{M2} (i.e., $\Delta q\Phi_M (= q\Phi_{M1} - q\Phi_{M2})$ increases) and increasing ϵ_{ox1} , step potential increases and the single electric field peak underneath the gate region, as shown in Fig. 2, increases resulting in improved gate transport efficiency, i.e., the carrier transport efficiency due to the effect of electric field distribution in the channel such that electron near the source are accelerated rapidly as a result of which average electron velocity in the channel increases [14], [15]. The electric field peak originates in the channel due to the work function difference between the two metal gates and hence, gate transport efficiency is the carrier transport efficiency due to the effect of gate electrode workfunction tuning. In Fig. 1, $L = 75$ nm curve has been plotted to illustrate the effect of shortening of screening metal gate length (L_s). It has been seen that with the reduction in the L_s , the potential step reduces and hence, deteriorates gate control. In scaled devices, electric field peak near the drain end causes carrier heating that result in the hot electron effect [3] thereby degrading the hot carrier reliability of the device. Hence, DMG-ISE architecture exhibits lowest field and reduced carrier temperature at the drain end as shown in Figs. 2 (inset 1) and 3; and thus improves the hot carrier reliability of the device. In SMG-ISE and bulk, occurrence of electric field peak