Cryogenic Electron Mobility and Subthreshold Slope of Oxygen-Inserted (OI) Si Channel nMOSFETs

Hiu Yung Wong
M-PAC Lab.
San Jose State University
San Jose, CA, USA
hiuyung.wong@sjsu.edu

Hideki Takeuchi
Atomera Inc.
Los Gatos, CA, USA
htakeuchi@atomera.com

Robert J. Mears
Atomera Inc.
Los Gatos, CA, USA
rjmears@atomera.com

***Abstract*— Oxygen-inserted (OI) Si channel nMOSFETs, which allow the formation and maintenance of an undoped epitaxial layer underneath the gate dielectric, are known to provide higher ON-state current and have higher inversion-layer carrier mobility compared to control silicon devices. However, it is not clear if the enhancement in mobility is due solely to reduced Coulomb scattering in the undoped layer. In this study, cryogenic measurements down to 4.2K are performed to deconvolve the contribution of phonon, surface roughness, and Coulomb scattering mobilities to total mobility. It is found that besides having reduced Coulomb scattering, due to the undoped epitaxial layer, the OI samples also exhibit 53% higher surface roughness mobility than the baseline samples. Notwithstanding the improved OI sample mobility, a saturation of the subthreshold slope at cryogenic temperature is still observed, as in the baseline and other technologies. In this paper, OI-Si nMOSFET mobility equations and empirical subthreshold slope saturation equation are developed which may be used for compact modeling from T = 330K to 4.2K.**

Keywords*— *Surface Roughness Mobility, Oxygen-insertion, Compact Modeling, Cryogenic Electronics, Subthreshold-Slope, TCAD

I. Introduction

Oxygen-inserted (OI) Si channel Metal-Oxide-Semiconductor Field Effect Transistors (MOSFET) are so-called because they include the insertion of partial mono-layers of oxygen atoms during epitaxy of the uppermost region of the silicon substrate, near the gate dielectric. Oxygen-insertion has been shown to improve various critical figures of merit in MOSFETs (Fig. 1) [1]-[6]. The inserted oxygen atoms are bonded covalently on Si-Si bonds while maintaining the overall silicon crystalline structure. Due to the high electronegativity of oxygen atoms and resulting electronic transfer with adjacent silicon atoms, the formation enthalpy of substitutional dopants and silicon point defects is modified, resulting in the trapping of defects and interstitials [1]. This can suppress oxidation-enhanced diffusion, enabling precise substrate doping profile control [2] and maintaining an undoped epitaxial layer above the OI layer even after subsequent high thermal-budget gate oxidation steps and activation anneals. As a result, the inversion layer electron mobility can be improved substantially due to reduced Coulomb scattering [3][4].

It is also known that OI-MOSFETs can enable improved gate dielectric quality metrics. For example, lower Negative Bias Temperature Instability (NBTI) [5] and reduced gate leakage current in both poly-Si/SiON [3] and High-κ-Metal-Gate stacks [6] have been observed. It is believed that these

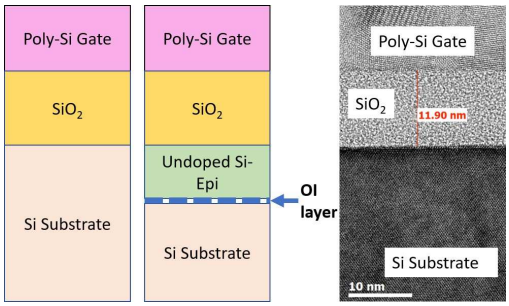

Fig. 1. Illustrations of the gate/channel stack of the baseline "STD" device (left) and the OI "MST" device (middle). The right figure shows a TEM picture of a "MST" device. Note that the OI layers are difficult to see in the TEM because they are partial monolayers and the silicon crystal lattice is maintained.

improvements are due, at least in part, to an exchange of oxygen atoms between the OI layer and the gate oxide during annealing. The exchange of oxygen between the OI layer and the gate oxide is evident from a recently reported isotope labeling experiment [1]. Hence, mobility improvement can also result from a reduction in surface roughness scattering.

Therefore, in order to understand better the physics of OI devices, and in particular their improved mobility, it is necessary to deconvolve the contributions of Coulomb scattering, surface roughness scattering, and phonon scattering. This requires performing mobility extraction down to 4.2K so that the phonon scattering component can be isolated.

Cryogenic electronics have gained significant attention in recent years due to their important role in quantum computing control [7], deep space exploration electronics [8], and cryogenic computing hardware for high-speed servers [9]. Therefore, the modeling of cryogenic MOSFETs becomes more important which includes the modeling of the carrier mobility [10]-[13], leakage mechanism [14], and the subthreshold slope (S.S.) [12][14][15]. It should be noted that the S.S. of a MOSFET is expected to reduce linearly with the temperature by following the Boltzmann statistics, kT/q, where k, T, and q are the Boltzmann constant, device temperature, and elementary charge, respectively. However, at T < 50K, it is found that S.S. saturates in the range of 10mV/dec to 20mV/dec. This may be explained by the band-tail states [10] or dielectric/channel interface traps [12][16], which is supported by an increase in 1/ f noise at cryogenic temperature [17]. Regardless of the cause of the abnormal S.S., it is important to investigate the S.S. saturation behavior,

as oxygen exchange is a major mechanism in OI-MOSFETs. Therefore, in this paper, we deconvolve the contribution of various scattering mechanisms to the improvement of electron mobility in OI-MOSFETs and compare the cryogenic S.S. saturation behavior to regular MOSFETs. OI-MOSFET mobility equations and empirical S.S. slope saturation equation are developed which may be used for compact modeling from T = 330K to 4.2K

II. Experiment

Two wafers (dubbed as "STD" and "MST6nm") were fabricated using a 0.18 µm commercial CMOS technology which is commonly used for power IC (Fig. 1) [18][19]. A standard wafer, "STD", was fabricated using the baseline process. A non-standard wafer, "MST6nm", was fabricated using the same process as "STD" except that the starting substrate included an epitaxial layer containing four oxygen partial monolayers followed by a 13 nm undoped Si epitaxial layer, to target final Si cap of 6nm to taking into account 7nm Si loss by pre-clean and oxidation processes. The well implants were also adjusted to have a similar maximum depletion width, and thus similar minimum depletion capacitance during the voltage-capacitance (CV) measurements. The effective substrate doping is about 1.9×10^{17} cm^{-3}. The gate oxide thickness is about 12.5nm as confirmed in the transmission microscopy measurement (TEM) in Fig. 1. Large area transistors (W = 120 µm and L = 120 µm) were fabricated for accurate CV extraction and to avoid statistical variations and minimize the effect of contact and lightly doped drain (LLD) resistance as is typical in accurate mobility extraction methods [20].

The transistors (W/L = 120 µm / 120 µm) in the two wafers were measured from 300 K to 4.2 K using a Lakeshore TTPX probe station system with four arms. Liquid Helium was used to cool down the sample with the temperature controlled by a Lakeshore Model 336 temperature controller. The temperatures were monitored by three sensors at the sample, probe arm, and radiation shield, respectively. The temperature reported below refers to the reading at the sample. To ensure accurate and repeatable measurement, each 4.2 K measurement was taken after the probe was in contact with the test structure for more than 5 minutes. For the 4.2 K measurement, the reading of the sample and probe arm temperatures were 4.2 K and 10 K, respectively.

Both gate voltage-drain current (I_DV_G) and CV measurements were performed to extract the electron mobility and the S.S. For the I_DV_G measurement, the drain voltage, V_D, was set to 50 mV, and the gate voltage, V_G, was swept from 0 V to 6 V using Keysight E5262A IV Analyzer.

For the CV measurement, both full-CV and split-CV were measured at 100kHz with 25 mV small signal amplitude and long integration. V_G was swept from -6 V to 6 V. In the full-CV measurement, the drain, source, and body were grounded and it was performed for a sanity check to ensure the measurement was normal at various temperatures and for the verification of the maximum depletion width. For the split-CV measurement, it was the same as the full-CV measurement except the body was allowed to float to enable an accurate extraction of the inversion charge for mobility calculation.

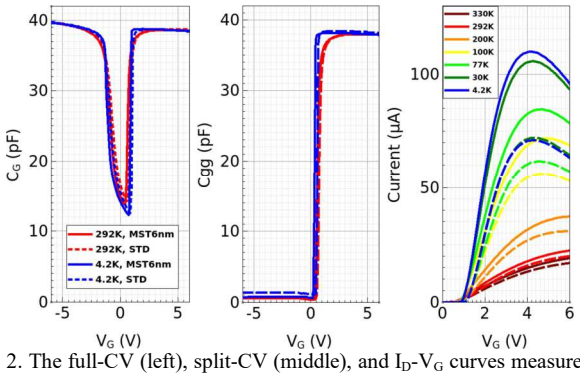

Fig. 2. The full-CV (left), split-CV (middle), and I_D-V_G curves measured at various temperatures. Dashed lines are "STD" and solid lines are "MST6nm". For clarity, only 292 K and 4.2 K are shown in the CV curves. V_D = 0.05 mV in the I_D-V_G measurement.

III. Mobility Modeling

Fig. 2 shows the CV measurement results at various temperatures. It can be seen that there is not much difference between the "STD" and "MST6nm" and between 300K and 4.2K. Fig. 2 also shows the I_D-V_G curves at V_D = 50mV. It is found that the current of "MST6nm" is much larger than that of "STD" at 200 K and below because phonon scattering, which is expected to be similar in both devices, becomes insignificant and the impurity scattering dominates at low temperatures. Since impurity scattering becomes stronger at lower temperatures, the peak current in "MST6nm" (V_G = 4V), which has an undoped epitaxial layer, is higher than that in "STD" by as much as 50% at 4.2 K.

Moreover, the I_D-V_G curves exhibit negative differential resistance (NDR) at high V_G and low temperatures (≤ 100 K). This is not due to gate leakage as the gate current is measured to be of the order of 10 pA (the resolution limit of the IV Analyzer). This is also not due to the dopant freeze-out effect under the spacer region because the LDD resistance is negligible compared to the channel resistance of the measured long-channel device (L = 120µm). It is also noted that the dopant freeze-out effect should be less at high V_G as the electric field under the spacer is higher at higher V_G which should mitigate the dopant freeze-out effect through field-enhanced ionization [11]. Therefore, the NDR can be explained purely due to the increase in surface roughness scattering at the cryogenic temperature under high V_G which is revealed when the phonon scattering diminishes [20].

Next, the contribution of impurity scattering and surface roughness scattering to the overall improvement of the "MST6nm" device was deconvolved. The electron mobility, µ, was extracted, following the standard split-CV method to find the inversion charge, Q_I, the depletion charge Q_D, and the vertical effective electric field, $E_\perp$ [20].

Q_I at a given V_G is found by integrating the split-CV curve from 0 V to V_G, $Q_I = \int_0^{V_G} C\,dV$, after offsetting by the minimum capacitance in the split-CV curve to account for parasitic capacitance.

Q_D is found by using the estimated effective uniform substrate doping, $N_A = 1.9\times10^{17}$ cm^{-3}, at various temperatures using the equation, $Q_D = \sqrt{2\varepsilon_s q N_A \Phi}$, where ε_s, q, and Φ are the Silicon permittivity, fundamental charge, and surface potential, respectively [21]. $\Phi = 2kT/q\ \ln(N_A/n_i)$, where n_i is the Si intrinsic carrier concentration [21]. At 4.2 K, since n_i is too low and causes numerical overflow, Φ is set to 1.18 eV.

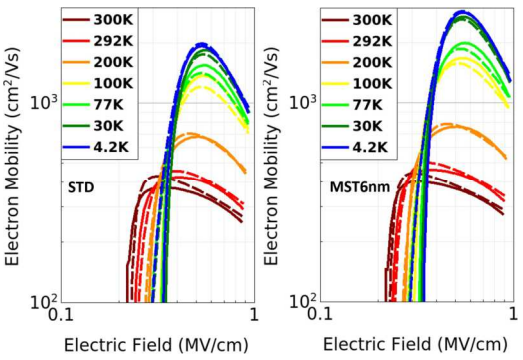

Fig. 3. Electron mobility as a function of temperature and $E_\perp$ for the "STD" (left) and "MST6nm" (right) devices. Solid lines are the mobilities extracted using the I_D-V_G and split-CV curves. Dashed lines are fitting curves using Eqs. (1)-(4).

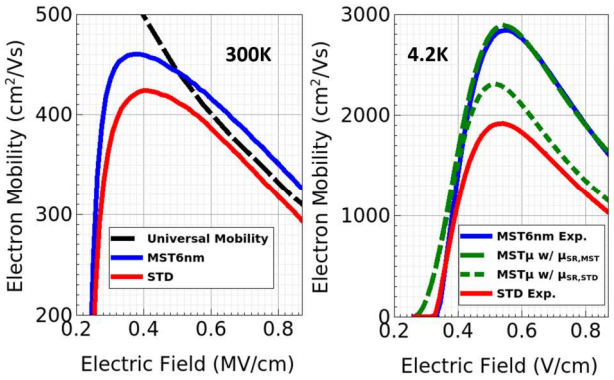

Fig. 4. Mobility analysis. Left: Experimental 300 K mobilities of "STD" and "MST6nm" compared to universal mobility [20]. Right: Experimental 4.2 K mobilities of "STD" and "MST6nm" compared to fitted mobility for MST6nm, μ, using either $\mu_{sr,MST}$ or $\mu_{sr,STD}$.

The mobility, μ, is found by using the drain current equation, $I_{DW} = Q_I \mu V_D/L$, where I_{DW} is the drain current per unit width. This equation is valid for small V_D such as in this study (50 mV).

$E_\perp$ is found by using the equation $E_\perp = (Q_I/2 + Q_D)/\varepsilon_s$ at each temperature and V_G [20].

μ at different temperatures and V_G is then plotted against the $E_\perp$ and shown in Fig. 3. Note that by extracting μ and $E_\perp$, the differences in the CV curves and the threshold voltages between the "STD" and the "MST6nm" are taken into account automatically.

All the measured mobility data were then found *consistently* deconvolved into Coulomb, phonon, and surface roughness scattering components using the following simple equations through Matthiesen's rule ($\frac{1}{\mu} = \frac{1}{\mu_{ph}} + \frac{1}{\mu_{sr}} + \frac{1}{\mu_C}$), as shown in Fig. 3:

$$\mu_{ph} = 8.62 \times 10^6 T^{-1.75} E_\perp^{-0.3} \quad (1)$$

$$\mu_{sr,STD} = 850 E_\perp^{-2} \quad (2)$$

$$\mu_{sr,MST} = 1300 E_\perp^{-2} \quad (3)$$

$$\mu_C = A Q_I^2 \quad (4)$$

where $\mu_{sr,STD}$ and $\mu_{sr,MST}$ are the μ_{sr} of the "STD" and "MST6nm", respectively. The Coulomb scattering model, Eq. (4), is an empirical model to capture the screening effect by the inversion electrons, where A is a temperature-dependent but field-independent pre-exponent. The phonon scattering model, Eq. (1), uses the same power exponents to T and E as those in [20][22] with the pre-exponent factor modified to fit the experimental data for a wider temperature range in this study. The power exponent of -2 in surface roughness scattering, Eqs. (2) and (3), was theoretically derived from 2DEG surface roughness scattering calculation [22].

The "STD" and the "MST6nm" share the same equation for μ_{ph} and μ_C, revealing that the higher mobility at the medium electric fields for "MST6nm" is due to reduced impurity in the epitaxial layer, but not due to different equation forms nor different physical mechanisms. Furthermore, the fact that "MST6nm" and "STD" have the same $E_\perp$-dependency indicates that inversion electrons are confined to the SiO_2/Si interface by the triangular potential in the same manner for the two samples. Hence, increased high-field mobility of "MST6nm" at low temperatures is attributed to a 53% larger μ_{sr} pre-exponent factor compared to "STD".

To clearly show the improvement of μ_{sr} in MST, Fig. 4 plots the experimental total mobility for "STD" and "MST6nm" compared to universal mobility at 300K [20]. At high $E_\perp$ (> 0.6 MV/cm), the mobility of "MST6nm" is 10% higher than that of "STD" and exceeds the universal mobility curve. Fig. 4 also shows the experimental mobilities at 4.2K on a linear scale. The proposed model fits the "MST6nm" experimental curve very well. It is also shown that if the $\mu_{sr,MST}$ is replaced by $\mu_{sr,STD}$ in Matthiesen's rule, the total mobility (MSTμ) is substantially lower and at high field, it approaches the "STD" experimental values. Therefore, the experimental data shows that high-field surface roughness scattering of NMOSFET on the Si channel can be improved. It should be noted that surface roughness scattering becomes more significant for more advanced nodes, where inversion carriers are exposed to higher vertical E -fields since V_{DD} and V_T do not scale as Dennard's scaling rule projects [23].

IV. Subthreshold Slope Modeling

The S.S. extracted between $I_D = 10^{-8}$A and $I_D = 10^{-9}$A is plotted in Fig. 5. It can be seen that despite the better μ_{sr} in "MST6nm" and the fact that oxygen exchange between the

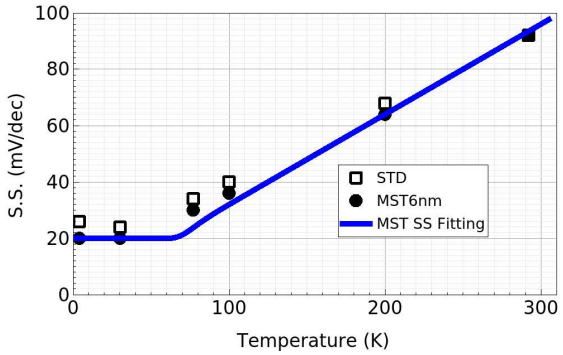

Fig. 5. Subthreshold slope (S.S.) of the "STD" and the "MST6nm" at $V_D = 50mV$ as a function of temperature. MST fitting curve based on Eq. (5) is also plotted.

gate oxide and the OI occurs during the gate oxidation process, it still exhibits the S.S. saturation effect as in the "STD" and other technologies. However, the "MST6nm" has almost 25% lower S.S. than the "STD" at 4K.

To model this, the following equation is derived and plotted in Fig. 5, which fits the data well,

$$SS = nkT\frac{\tanh\alpha(T-T_C)+1}{2} + S_0\frac{\tanh\alpha(T_C-T)+1}{2} \quad (5)$$

where n=1.6 (non-ideality factor), T_C =65K (saturation temperature), S_0=20mV (saturation subthreshold slope), and α=0.1 (smoothing factor).

V. Conclusions

Temperature-dependent Coulumb, surface roughness and phonon scattering mobility equations, and sub-threshold slope saturation equation are developed from 330K to 4.2K for MST devices which can be used for compact modeling. It also reveals that the surface roughness mobility is improved by 53% in MST.

Acknowledgment

Part of this work was supported by the National Science Foundation under Grant 2046220 and AMDT Endowment from the College of Engineering, San José State University.

References

[1] H. Takeuchi, R. J. Mears, M. Hytha, D. J. Connelly, P. E. Nicollian, and H.-Y. Wong, "Remote control of doping profile, silicon interface, and gate dielectric reliability via oxygen insertion into silicon channel," The 2022 International Meeting for Future of Electron Devices (IMFEDK 2022).

[2] H. Takeuchi, R. J. Mears, R. J. Stephenson, M. Hytha, D. Connelly, P. Fastenko, R. Burton, N. W. Cody, D. Weeks, D. Choutov, N. Agrawal, and S. Datta, "Punch-through stop doping profile control via interstitial trap- ping by oxygen-insertion silicon channel," IEEE Journal of the Electron Devices Society 6, 481–486 (2018).

[3] N. Xu, N. Damrongplasit, H. Takeuchi, R. J. Stephenson, N. W. Cody, A. Yiptong, X. Huang, M. Hytha, R. J. Mears, and T.-J. K. Liu, "Mosfet performance and scalability enhancement by insertion of oxygen layers," in 2012 International Electron Devices Meeting (2012) pp. 6.4.1–6.4.4.

[4] N. Xu, H. Takeuchi, M. Hytha, N. W. Cody, R. J. Stephenson, B. Kwak, S. Y. Cha, R. J. Mears, and T.-J. King Liu, "Electron mobility enhancement in (100) oxygen-inserted silicon channel," Applied Physics Letters 107, 123502 (2015), https://doi.org/10.1063/1.4931431.

[5] A. Marshall, S. Nimmalapudi, W. K. Loh, J. Krick, L. Hutter, H. Takeuchi, R. Stephenson, R. J. Mears, and S. Ikeda, "The impact of oxygen insertion technology on sram yield performance," in 2017 IEEE Electron Devices Technology and Manufacturing Conference (EDTM) (2017) pp. 191–192.

[6] J. Smith, H. Takeuchi, R. Stephenson, Y. Chen, M. Hytha, R. J. Mears, and S. Datta, "Experimental investigation of n-channel oxygen-inserted (oi) silicon channel mosfets with high-k/metal gate stack," in 2018 76th Device Research Conference (DRC) (2018) pp. 1–2.

[7] E. Charbon, F. Sebastiano, A. Vladimirescu, H. Homulle, S. Visser, L. Song, and R. M. Incandela, "Cryo-cmos for quantum computing," in 2016 IEEE International Electron Devices Meeting (IEDM) (2016) pp. 13.5.1–13.5.4.

[8] R. L. Patterson, A. Hammoud, J. E. Dickman, S. Gerber, M. Elbuluk, and E. Overton, "Electronics for deep space cryogenic applications," in Pro- ceedings of the 5th European Workshop on Low Temperature Electronics (2002) pp. 207–210.

[9] G.-H. Lee, D. Min, I. Byun, and J. Kim, "Cryogenic computer architecture modeling with memory-side case studies," in Proceedings of the 46th Inter- national Symposium on Computer Architecture, ISCA '19 (Association for Computing Machinery, New York, NY, USA, 2019) p. 774–787.

[10] A. Beckers, F. Jazaeri, and C. Enz, "Cryogenic mos transistor model," IEEE Transactions on Electron Devices 65, 3617–3625 (2018).

[11] H. Y. Wong, "Calibrated si mobility and incomplete ionization models with field dependent ionization energy for cryogenic simulations," in 2020 In- ternational Conference on Simulation of Semiconductor Processes and De- vices (SISPAD) (2020) pp. 193–196.

[12] P. Dhillon, N. C. Dao, P. H. W. Leong, and H. Y. Wong, "Tcad modeling of cryogenic nmosfet on-state current and subthreshold slope," in 2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (2021) pp. 255–258.

[13] P. Dhillon and H. Y. Wong, "A wide temperature range unified undoped bulk silicon electron and hole mobility model," IEEE Transactions on Elec- tron Devices 69, 1979–1983 (2022).

[14] T. Jiao and H. Y. Wong, "Robust cryogenic ab-initio quantum transport simulation for lg = 10 nm nanowire," Solid-State Electronics 197, 108440 (2022).

[15] A. Beckers, F. Jazaeri, and C. Enz, "Revised theoretical limit of subthresh- old swing in field-effect transistors," (2018).

[16] A. Beckers, F. Jazaeri, and C. Enz, "Characterization and modeling of 28-nm bulk cmos technology down to 4.2 k," IEEE Journal of the Electron Devices Society 6, 1007–1018 (2018).

[17] H. Oka, T. Matsukawa, K. Kato, S. Iizuka, W. Mizubayashi, K. Endo, T. Yasuda, and T. Mori, "Toward long-coherence-time si spin qubit: The origin of low-frequency noise in cryo-cmos," in 2020 IEEE Symposium on VLSI Technology (2020) pp. 1–2.

[18] J. -H. Lee et al., "ESD Robust Fully Salicided 5-V Integrated Power MOSFET in Submicron CMOS," in IEEE Electron Device Letters, vol. 38, no. 5, pp. 623-625, May 2017.

[19] Y.-A. Chen, C. Hong, S. Li, A. Raol, R. Burton, M. Duane, L. N. Hutter, H. Takeuchi, and R. J. Mears, "Source engineering on oxygen-inserted si channel for gate length scaling of low-voltage switch devices," in 2022 IEEE 34th International Symposium on Power Semiconductor Devices and ICs (ISPSD) (2022) pp. 197–200.

[20] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in si mosfet's: Part i-effects of substrate impu- rity concentration," IEEE Transactions on Electron Devices 41, 2357–2362 (1994).

[21] "Metal-insulator-semiconductor capacitors," in Physics of Semiconductor Devices (John Wiley Sons, Ltd, 2006) pp. 197–240, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470068328.ch4.

[22] S. Takagi, et al., "On the universality of inversion layer mobility in Si MOSFET's: Part II-effects of surface orientation," in IEEE Transactions on Electron Devices, vol. 41, no. 12, pp. 2363-2368.

[23] R. H. Dennard, F. H. Gaensslen, H. -N. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," in IEEE Journal of Solid-State Circuits, vol. 9, no. 5, pp. 256-268, Oct. 1974, doi: 10.1109/JSSC.1974.1050511.