Compact Modeling of Stress Effects in Scaled CMOS

Chi-Chao Wang, Wei Zhao, Frank Liu*, Min Chen, and Yu Cao
Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287
*IBM Austin Research Laboratory, Austin, TX 78758
Email: cwang73@asu.edu

***Abstract* — Strained Si is implemented into the standard CMOS process to enhance carrier transport properties since the 90nm technology node. However, due to the non-uniform stress distribution in the channel, the enhancement of carrier mobility and threshold voltage strongly depend on layout parameters, such as channel length (L) and source/drain diffusion length (L_{sd}). In this work, a compact model that physically captures these behaviors is developed for circuit simulation with strained CMOS technology.**

Keywords* — *Layout Dependence, Stress Effect, Compact Modeling, Mobility, Threshold Voltage

I. INTRODUCTION

Strain technology, which alters band structure and reduces effective mass and scattering rate, is essential to elevate carrier mobility for continual scaling. The exact amount of mobility enhancement depends on both the applied stress level during the fabrication (for example, determined by the Ge composition for eSiGe technology) and circuit layout parameters, such as transistor length and source/drain size [1], because of the non-uniform stress distribution in the channel region. Such non-uniformity results in pronounced shifts in transistor and circuit performance. To capture such a systematic effect, traditional efforts resort to TCAD simulation to extract the stress level from the entire layout and analyze performance enhancement [2]. This approach is usually expensive in computation. Therefore, it is necessary to develop a more effective modeling approach that is able to extract the stress effect for each device and embed it into standard model parameters for circuit simulation.

The layout-dependent stress effect is first observed and reported from STI stress [3]. Layout-dependent models regarding STI stress are proposed on the basis of the experimental observation that the changes of drive current and threshold voltage follow the trend of the length of oxide definition area (LOD) [4] [5]. Moreover, a modeling approach of equivalent stress level is proposed to account for the mobility enhancement with an assumption that the mobility enhancement is proportional to the applied stress [6]. However, this approach also results in empirical fitting.

In this work, a new general modeling approach is proposed to capture the layout-dependent stress effect. This model is derived from the first principle and physically captures the impact of circuit layout on transistor performance such that model scalability is guaranteed for future technology generations.

II. COMPACT STRESS MODELING

A. *Bathtub Curve of Stress Distribution*

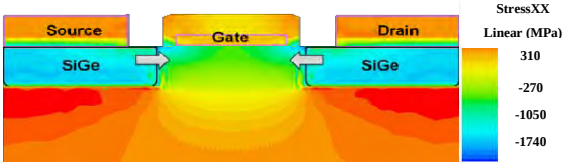
As investigated in [7], the stress magnitude in Si substrate decays sharply from the edge of the channel to the center, and becomes less dependent on the distance when the location is far from the origin of the applied stress. Figure 1 (a) shows the simulated stress contour of a PMOS with SiGe stressors in source/drain area, illustrating that the stress is imposed from the source/drain area and results in the non-uniform distribution [8]. In Fig. 1 (b), the simulation shows the stress profiles for devices with different channel lengths: *the shorter the channel length, the higher overall stress level*. Although the stress magnitude is different, the stress profile is similar and behaves like a bathtub curve. Based on this observation, without losing the generality, a linear piecewise approximation is proposed to capture the stress profile as Eqs. (1)-(3),

$$Y_1 = \sigma_P - dx \tag{1}$$

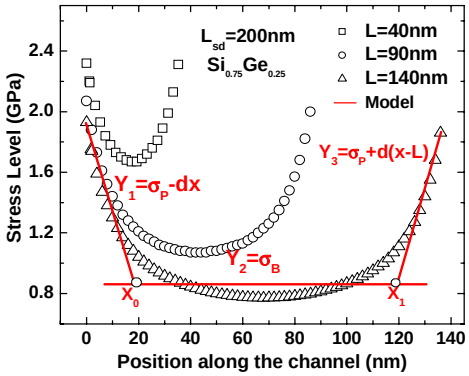
$$Y_2 = \sigma_B \tag{2}$$

$$Y_3 = \sigma_P + d(x - L) \tag{3}$$

where σ_P and σ_B denote the peak and bottom stress levels in the channel, respectively, and d represents the slope. Moreover,



(a) Stress contour of a MOSFET with eSiGe in S/D area.



(b) Bathtub curve approximation for stress distribution.

Figure 1. Modeling of stress distribution.

Y_1 and Y_3 intercept with Y_2 at points of x_0 and x_1, respectively. x_0 and x_1 are expressed by Eqs. (4) and (5).

$$x_0 = \frac{\sigma_P - \sigma_B}{d} \tag{4}$$

$$x_1 = L - \frac{\sigma_P - \sigma_B}{d} \tag{5}$$

The stress distribution in the channel is sensitive to some layout parameters and thus results in additional variations in device performance [1]. In Fig. 2, as S/D diffusion length (L_{sd}) increases, σ_P and σ_B become higher due to the increased amount of stressor material, and finally reaches the saturation state when L_{sd} is larger than the critical length [1]. Meanwhile, the stress for the device with smaller channel length is higher than that with larger channel length. To account for the stress dependence on L and L_{sd}, σ_P is modeled as Eq. (6), where σ_m is the saturation stress level and A and *m* are fitting parameters accounting for the dependence on L_{sd} and the stress decreasing rate over distance from neighboring transistors.

$$\sigma_P = \left(1 + \frac{m}{L} + \frac{m}{L + L_{sd}} + \frac{m}{2L + L_{sd}}\right) \cdot \frac{L_{sd}}{A + L_{sd}} \cdot \sigma_m \tag{6}$$

Each term in the parenthesis represents the contribution by a diffusion region, depending on their separation distance to the channel. The first two terms in Eq. (6) account for the contribution from the source/drain area of the target transistor, while the rest two terms represent the stress sources from the nearest neighboring transistors. Moreover, Eq. (6) assumes that all diffusion regions in the neighboring transistors have the same size L_{sd}. If they are different, the exact value should be used to replace the corresponding L_{sd}. On the other hand, as channel length becomes shorter, σ_B grows up and to the limit of σ_P when channel length reaches zero. This channel length dependence can be modeled by Eq. (7) with a fitting parameter C. In Fig. 2, the models show good agreement with TCAD simulation results.

$$\sigma_B = \frac{C}{C + L} \cdot \sigma_p \tag{7}$$

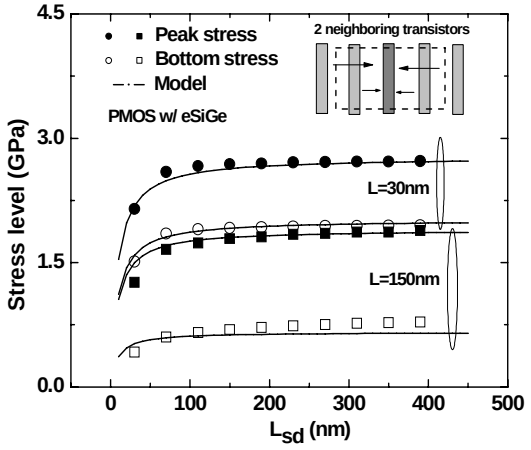


Figure 2. Layout dependence is modeled through σ_P and σ_B in Eqs. (6) and (7).

B. Equivalent Mobility (μ_e) in the Channel

When the stress is applied, the band structure is altered and further changes the symmetrical ellipsoids of constant energy of silicon. This results in carrier redistribution and reduce carrier effective mass and scattering rate; then carrier mobility is enhanced. Based on strain-induced band splitting, the model regarding strain-induced mobility change is physically modeled with the form as Eq. (8),

$$\frac{\mu}{\mu_0} = 1 + B \cdot \left[\exp\left(\frac{\Delta E}{kT}\right) - 1\right] \tag{8}$$

where the coefficient, B, is a physical constant [9] [10]. ΔE denotes the strain-induced energy splitting of conduction band or valence band and can be calculated by the deformation potential theory [11], which indicates the applied stress level is linearly proportional to energy splitting. Therefore, energy splitting is modeled by Eq. (9).

$$\Delta E = P \cdot \sigma \tag{9}$$

where *P* can be calculated by deformation potential constants, and it is also temperature dependent because the temperature variation alters the bandgap and further affects the energy band splitting. Therefore, the temperature-dependent behavior can be modeled as Eq. (10),

$$P(T) = P_0 \cdot \left(\frac{T}{T_0}\right)^{\alpha} \tag{10}$$

where P_0 denotes its value at room temperature (T_0) and α is a fitting parameter accounting for the temperature dependence. Moreover, since the stress level in the channel is not a constant, the enhancement in mobility is also non-uniformly distributed. Based on the principle of current continuity, the non-uniform mobility can be modeled as an equivalent mobility, μ_e, by using Eq. (11), [12]

$$\frac{\mu_o}{\mu_e} = \frac{1}{L}\int_0^L \frac{\mu_o}{\mu} dx \tag{11}$$

where μ_0 denotes the unstrained mobility. Therefore, an analytical solution for mobility can be derived as function of channel length and S/D diffusion length to bridge the layout parameters to mobility variation, as expressed in Eq. (12).

$$\frac{\mu_0}{\mu_e} = \frac{2kT}{dPL(B-1)} \cdot \left\{ \frac{-dPx_0}{kT} + \ln\left[\frac{1 + B\left(\exp\left(\frac{P\sigma_P}{kT}\right) - 1\right)}{1 + B\left(\exp\left(\frac{P\sigma_P - dPx_0}{kT}\right) - 1\right)} \right] \right\} + \frac{L - 2x_0}{L \cdot \left[1 - B + B\exp\left(\frac{P\sigma_B}{kT}\right)\right]} \tag{12}$$

In Fig. 3, TCAD simulation shows the amount of mobility enhancement is strongly dependent on channel length and S/D diffusion length; as channel length becomes shorter, the stress effect becomes more effective. Meanwhile, with a larger L_{sd}, the mobility enhancement is stronger. This dependence is captured by the new model well.

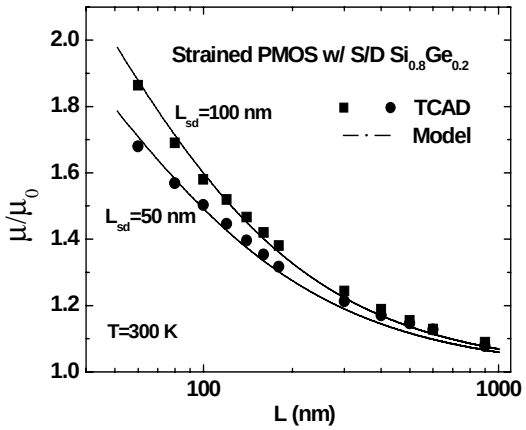


Figure 3. Layout dependence of channel mobility.

On the other hand, to assess the temperature effect, the device is operated at different temperatures from 100K to 800K. In Fig. 4, TCAD simulation shows that at 100K, the mobility enhancement increases 31% more than that at room temperature for the device with both L and L_{sd} at 100nm. On the contrary, as temperature increases more than the room temperature, the mobility enhancement declines. The sensitivity of mobility enhancement to temperature is higher at lower temperatures. This behavior can be explained by Eq. (8), where the temperature term is in the exponential function, so that the change is more dramatic under low temperatures. Moreover, the device with longer channel length is less sensitive to the temperature variation. The proposed model demonstrates the good agreement with TCAD simulation.

C. Strain Induced Threshold Voltage Shift

In addition to strain-induced mobility variation, strain-induced threshold voltage shift is also observed in the strained devices. The change in threshold voltage is attributed to strain-induced variation of energy bandgap, electron affinity, and density of states (DOS), where the effect of density of states (DOS) can be ignored due to its insignificant impact [13]**.** Based on the deformation potential theory [11], the strain-induced change in bandgap and electron affinity is proportional to the applied stress magnitude, so the threshold voltage change is modeled by Eq. (13),

$$\Delta V_{th}(\sigma_B) = VTH\_STR \cdot \sigma_B \tag{13}$$

where *VTH_STR* is a fitting parameter to capture the linear relationship between threshold voltage shift and the applied stress magnitude. Note that the bottom stress level (σ_B) is chosen for threshold voltage shift because the barrier peak between source and substrate is controlled by σ_B, as shown the stress bathtub curve in Fig. 5. The simulated valence bands of unstrained/strained PMOSFETs indicate that V_{bi} is lowered by stress effect and thus it becomes easier for holes to pass through the channel. This lowering valence band confirms the strain-induced threshold shift. In Fig. 6, TCAD simulation based on eSiGe technology shows that the strain-induced V_{th} shift has strong dependence on the channel length and source/drain diffusion length; the strain induced ΔV_{th} increases as L decreases, and meanwhile the lower ΔV_{th} is observed for a smaller L_{sd} due to less S/D stressors. This dependence is captured well by Eq. (13).

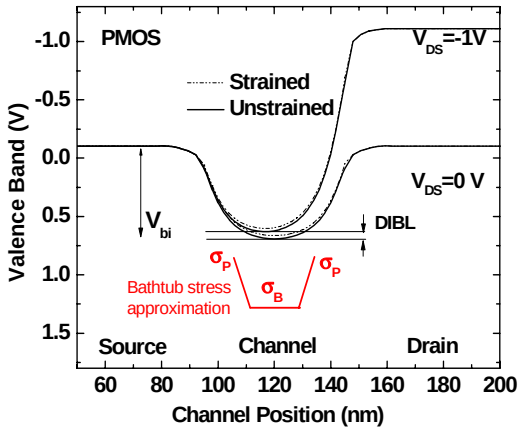


Figure 5. Simulated Valence band diagram for unstrained/ strained PMOS under different drain biases.

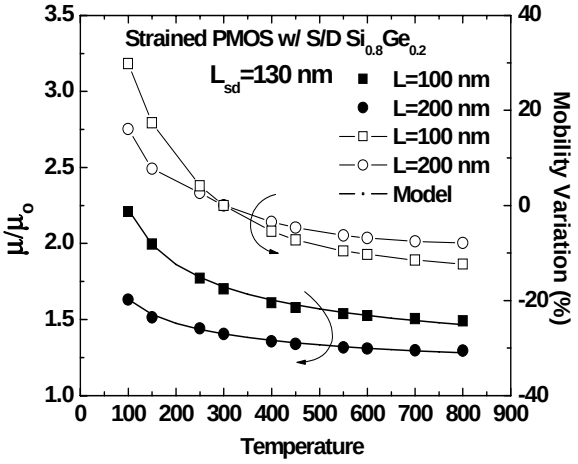


Figure 4. Temperature dependence of mobility.

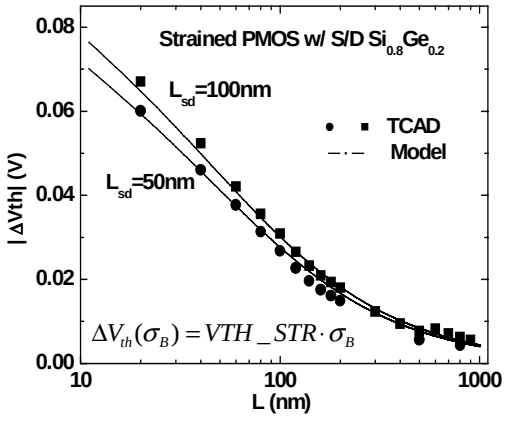


Figure 6. Strained induced ΔV_{th} as a function of channel length (L) and S/D diffusion length (L_{sd}).

D. Other Secondary Effects

The effect of drain induced barrier lowering (DIBL) becomes more important as channel keeps scaling. To the first order, the DIBL-induced threshold voltage shift can be modeled as Eq. (14) [14],

$$\Delta V_{th}(DIBL) \approx \left[3(V_{bi} - 2\phi_B) + V_{ds}\right] e^{-L/l} \quad (14)$$

where l denotes the characteristic length for DIBL effect. V_{bi} is the energy barrier between source and substrate and Φ_B is the bulk potential with expressions in Eqs. (15)-(16),

$$V_{bi}(\sigma_p) = \frac{kT}{q} \ln\left(\frac{N_{SOURCE} N_{SUB}}{n_i^2}\right) \quad (15)$$

$$\phi_B = \frac{kT}{q} \ln\left(\frac{N_{SUB}}{n_i}\right) \quad (16)$$

where N_{SOURCE} and N_{SUB} are the doping concentrations in source and substrate. V_{bi} and Φ_B are important factors to DIBL effect. Moreover, applying stress affects the intrinsic carrier density (n_i), which is an exponential function of bandgap, and further makes impact on V_{bi} and Φ_B. However, in Eq. (14), the term of (V_{bi}-2Φ_B) is independent of n_i, indicating the impacts of stress on V_{bi} and Φ_B cancel each other for DIBL effect. On the other hand, stress alters Φ_B and further influences the depletion depth (X_{dep}), shown in Eq. (17), which further affects DIBL and subthreshold swing through Eqs. (18) and (19), respectively.

$$X_{dep} = \sqrt{\frac{2\varepsilon_{si}(2\phi_B - V_{BS})}{qN_{SUB}}} \quad (17)$$

$$l = \sqrt{3T_{ox}X_{dep}} \propto (E_g)^{\frac{1}{4}} \quad (18)$$

$$S.S. = 60mV \cdot \left(1 + \frac{C_{dep}}{C_{ox}}\right) \quad (19)$$

Note that the characteristic length (l) in DIBL effect is the major term to be impacted by stress and is proportional to $E_g^{1/4}$, making DIBL relatively insensitive to the stress effect. In Fig. 7, our model captures strain-induced ΔV_{th} in strained devices operated under different drain biases as well.

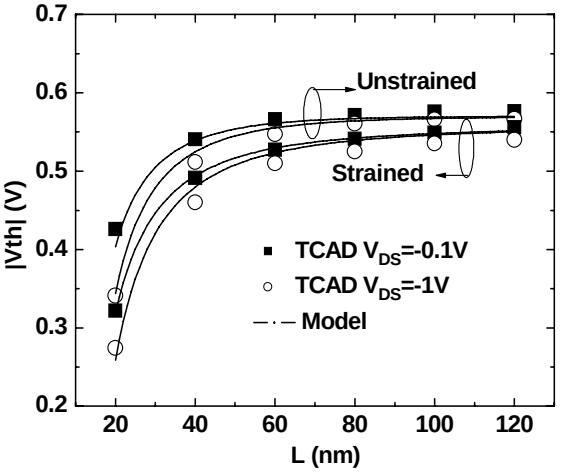


Figure 7. Threshold voltage for strained/unstrained devices under different drain biases.

III. Conclusion

With the scaling of device dimension, the strain-induced mobility and threshold voltage variation becomes more pronounced. Therefore, it is essential to develop compact models of the layout dependent stress effect for circuit analysis and optimization. In this work, the solution that bridges layout parameters to device electrical characteristics is proposed for simulation.

IV. Acknowledgement

The authors would like to acknowledge the support of the Focus Center for Materials, Structures, and Devices (MSD), as well as Center for Circuits & System Solutions (C2S2), two of five research centers funded under the Focus Center Research Program, a Semiconductor Research Corporation Program.

References

[1] G. Eneman, et al., "Scalability of the Si1-xGex Source/Drain technology for the 45-nm technology node and beyond," TED, vol. 53, no. 7, pp.1647-1656, Jul. 2006.

[2] V. Moroz, et al., "The impact of layout on stress-enhanced transistor performance," pp. 143-146, SISPAD 2005.

[3] G. Scott, et al., "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," in *IEDM Tech. Dig.*, 1999, pp. 827-830.

[4] R.A. Bianchi, et al., "Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance," *IEDM*, pp.117-120, 2002.

[5] K-W Su, et al., "A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics," *Custom Integrated Circuits Conference*, pp. 245-248, 2003.

[6] M. V. Dunga, et al., "Modeling advanced FET technology in a compact model," *TED*, vol. 53, No. 9, pp.1971-1978, Sept. 2006.

[7] C. E. Murray, "Mechanics of edge effects in anisotropic thin film/substrate systems," *Journal of Applied Physics*, vol. 100, 103532, 2006.

[8] Taurus-Tsuprem4, Manual, Oct. 2005. V. X-2005. 10.

[9] Sentaurus Device, Manual, June 2005. Version Y-2006. 06.

[10] J.L. Egley, et al., "Strain effects on device characteristics: implementation in drift-diffusion simulators," *Solid-State Electronics*, 36(12), pp. 1653-1664, 1993.

[11] J.-S. Lim, et al., "Comparison of threshold-voltage shifts for uniaxial and biaxial tensile-stressed n-MOSFETs," *EDL*, vol. 25, no. 11, pp. 731-733, Nov. 2004.

[12] F. Payet, et al., "Nonuniform mobility-enhancement techniques and their impact on device performance," *TED*, vol. 55, no. 4, pp. 1050-1057, April 2008

[13] W. Zhang, et al., "On the threshold voltage of strained-Si–Si1-xGex MOSFETs," *TED*, Vol. 52, no. 2, pp.263-268, Feb. 2005.

[14] Z.-H. Liu, et al., "Threshold voltage model for deep-submicrometer MOSFET's," *TED*, vol. 40, no. 1, pp. 86-95, Jan. 1993.