Calibrated Mobility Corrections for Drift Diffusion Simulation of Strained MOSFET Devices.

D. Villanueva[1], A. Dray[2], S. Orain[1], V. Fiori[2], C. Ortolland[1], E. Fuchs[2], F. Salvetti[1] and A. Juge[2]
[1]Philips Semiconductors, 860 rue Jean Monnet 38920 Crolles, France,+33 (0)4 38 92 26 49, davy.villanueva@philipscrolles.st.com
[2]STMicroelectronics, 850 rue Jean Monnet 38926 Crolles, France

Abstract – In this paper, simple corrections to traditional Drift-Diffusion mobility models are derived from strained full-band Monte Carlo simulations in order to correctly account for strain effects. The validity of this approach is benchmarked upon Monte Carlo MOSFET simulations as well as experimental data featuring tensile nitride capping layers.

I. INTRODUCTION

Process induced stress has become one of the key performance boosters of modern MOSFET devices. Industrially employed techniques such as the use of tensile nitride capping layers or compressive source-drain regions [1] lead to highly non homogeneous stress field within devices. From the TCAD device simulation point of view, handling non constant stress fields remains a problematic issue. State-of-the-art full-band Monte Carlo simulators are up till now bound to constant strain fields throughout the device to be simulated: this is satisfactory for strained silicon films onto $Si_{1-x}Ge_x$ but does not help in case of process induced strains. It is therefore mandatory to have a tool which allows to handle any strain field. In this paper a simple and original Drift-Diffusion (DD) based procedure is proposed to correctly assess the impact of any non homogeneous stress field on device performance. This approach consists in deriving corrections for each mobility component of the Mathiessen's rule from full band Monte Carlo (MC) simulations. The validity of the proposed corrections is first benchmarked against Monte Carlo simulations of $Si_{1-x}Ge_x$ strained devices for various gate lengths. The methodology is then applied to process induced strain fields as obtained from mechanical simulations. The results are compared to two sets of experimental data.

II. MODELING METHODOLOGY

A. State of the art and general approach

Several options have been already proposed in the literature to account for the modification of carrier mobility under stress [2][3][4]. The first one is the so-called piezo-resistive approach which relates the strained to the unstrained mobility as given by Eq.1. This approach has the disadvantage to be applied to the overall carrier mobility, leading thus to various sets of piezo coefficients depending on the doping and electric field. Furthermore, it used to relate mobility to stress while actually only strain matters as it directly gives information on the lattice parameters, which are the relevant quantities with respect to transport properties. The second alternative [3] limits the change in mobility to that of the effective conduction mass, omitting thereby the impact of the reduced inter-valley scatterings [5]. Although the approach [4] has not been applied to non homogeneous stress fields, it allows to treat independently each mobility component and their associated scattering mechanisms. In this work mobilities are also dealt separately and the applied corrections are based either on a piezo-resistive like approach or on the change in the effective conduction mass.

B. Low field bulk mobility

The strained bulk electron mobility is treated according to Eq.1. Following Eq.2, the α_{11} and α_{12} coefficients may be extracted from strained and unstrained MC transport simulations along the two directions x and y (e.g. [100] and [010]). The band structure of the strained MC simulations corresponds to that of a [001] oriented silicon layer grown on a relaxed $Si_{0.9}Ge_{0.1}$ buffer. The strain components are given in Eq.3. As a result, the bulk mobility is known for any diagonal strain tensor.

$$\mu_{strained} = \mu\left(\bar{1} - \bar{\mathrm{A}} \cdot \bar{\varepsilon}\right)$$

$$[\mathrm{A}] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{12} & 0 & 0 & 0 \\ \alpha_{12} & \alpha_{11} & \alpha_{12} & 0 & 0 & 0 \\ \alpha_{12} & \alpha_{12} & \alpha_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_{44} \end{bmatrix}$$

Eq. 1: Strained mobility with A being the ε related piezo-resistive matrix.

$$\frac{\mu_x}{\mu_0} = 1 - \alpha_{11}\varepsilon_{xx} - \alpha_{12}(\varepsilon_{yy} + \varepsilon_{zz}) = R_x$$
$$\frac{\mu_y}{\mu_0} = 1 - \alpha_{11}\varepsilon_{yy} - \alpha_{12}(\varepsilon_{xx} + \varepsilon_{zz}) = R_y$$

Eq. 2: Coefficients extraction based on mobility ratios.

$$\varepsilon_{xx} = \varepsilon_{yy} = \left(a_{si} + 0.2x_{Ge}(1 - x_{Ge}) + (a_{Ge} - a_{si})x_{Ge} - a_{si}\right)/a_{si}$$
$$= 0.37\%$$

$$\varepsilon_{zz} = \varepsilon_{xx} \cdot \frac{-2\upsilon}{1-\upsilon} = -0.29\%$$

Eq. 3: Strain components for a $Si_{0.9}Ge_{0.1}$ buffer layer.

C. Coulomb mobility

In the case of the ionized impurity limited mobility, the corrections to be brought are less obvious. On the one hand the scatterings due to ionized impurities can be considered as being intra-valley, which implies that the scattering rate is not modified by the strain induced reduction of the available density of states. On the other hand, the effective conduction mass depends on strain. Thus, assuming that the mobility may be written as in Eq.4, the high doping strained mobility is obtained by changing only the effective conduction mass as depicted in Figure1. In this way, the doping limited mobility appears to be in good agreement with MC data.

$$\mu = q\frac{\langle \tau \rangle}{m^*}$$

Eq. 4: mobility equation with $<\tau>$ being the average scattering rate and m^* the effective conduction mass.

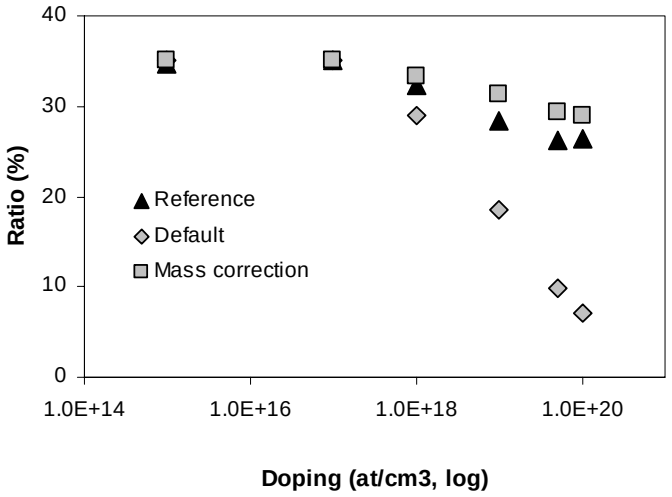

Figure1: strained/unstrained mobility ratio as a function of the doping level as obtained from MC simulations (reference) and DD with (Mass correction) or without (default) changing the high doping mobility term (the intrinsic mobility being always modified).

D. Surface mobility

Figure2 depicts the mobility enhancement for both the low field bulk mobility and the surface interaction limited mobility. For comparison purpose, the results are given for 10% and 20% Ge contents. It appears that for both Ge contents, the enhancement of the surface mobility is higher than that of the bulk mobility. Therefore, unlike [4] the strain dependency of the surface mobility is defined by another set of piezo coefficients. These are derived from surface mobility MC simulations for two directions as defined by Eq2.

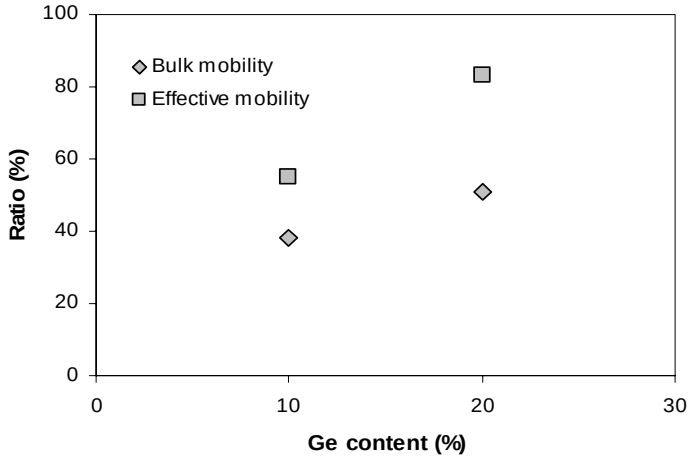

Figure2: Enhancement of the low filed bulk mobility and the surface interaction limited mobility for two Ge contents.

E. Saturation velocity

The last important aspect of the mobility to consider in MOSFET devices is the high longitudinal field behavior. Although it is clear that the saturation velocity is merely left unchanged by strain, this latter has been shown to impact the quasi-ballistic transport [6]. Unfortunately, the saturation velocity is often used as a fitting parameter for velocity overshoots in DD nano-scale MOSFET's simulations. However these conflicting views may not be an issue with respect to the impact of strain: indeed it has been demonstrated that the velocity at the drain end has only a marginal impact on the on-state current of nano-scale devices [7]. Thus the calibrated saturation velocity used to reproduce unstrained device current may be used for strained devices as well.

III. RESULTS & DISCUSSIONS

A. Comparison with Monte Carlo simulations

In order to validate the previously developed corrections, the saturation current enhancement ratio induced by the $Si_{0.9}Ge_{0.1}$ buffer is plotted as a function of the gate length both for the DD and the MC approaches in the case of an nMOSFET (Figure3). The DD surface and high field mobilities have been fit on the unstrained 65nm device simulation in order to account for nanoscale transport features. As a result the saturation velocity is set well above the default 1e7cm/s value. Figure3 shows that the enhancement ratio decreases with the gate length. This can be mainly attributed to the contribution of the non-linear regime in the overall transport [6]. At small gate length, the DD ratio lines up pretty well upon the MC trend line. This supports the fact that setting a strain dependent saturation velocity is not necessary to obtain the right current enhancement for small devices. The mismatch for long devices is mostly due to the fitted saturation velocity. Indeed, the saturation mobility being adjusted for the 65nm device to a higher value, it is no longer suited for long devices: the non-linear transport contribution is underestimated and this leads to a too high enhancement ratio.

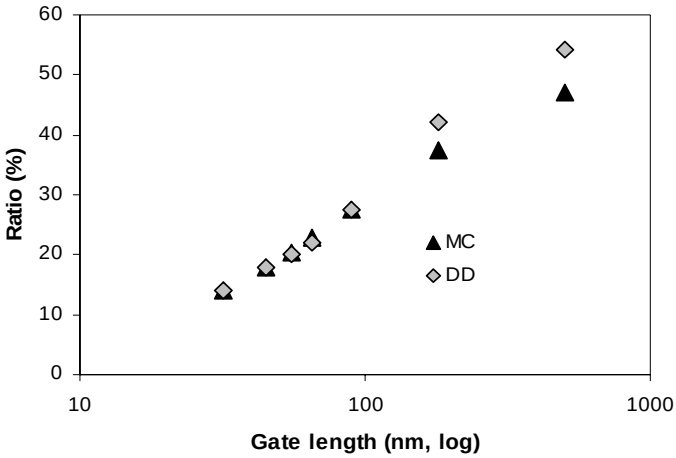

Figure3: Comparison of the current enhancement ratio vs gate length for DD and MC simulations.

B. Simulation of Strained CESL nMOSFET

B.1 Mechanical data

The comparison is extended to experimental data: the current enhancement ratio is reported as a function of the gate length and the CESL process conditions. The strain fields induced by tensile CESL's are computed by means of mechanical simulations. Figure4 shows a 3D view of the simulated CESL strained MOSFET. The resulting strain field for a 65nm device turns out to be non homogeneous as depicted in Figure5.

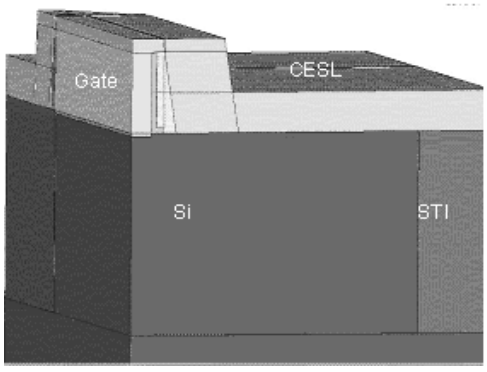

Figure4: 3D view of CESL strained MOSFET.

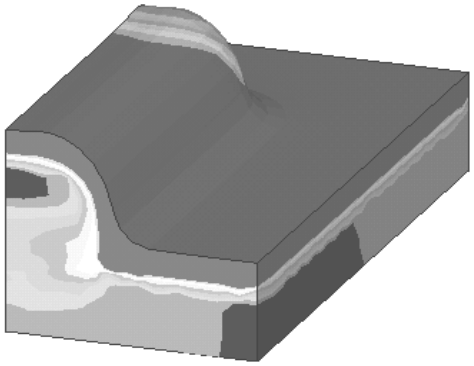

Figure5: Strain field within a 65nm CESL strained MOSFET: high strain at the gate edge.

B.2 Device description

Investigating the impact of strain on device characteristics requires that standard device characteristics such access resistance or short channel effects are correctly accounted for. Indeed, the strain enhancement ratio depends on quantities such as normal fields or doping levels. Therefore, it is mandatory to have a good representation of the doping profiles within the device. With that respect, doping profiles have been calibrated upon stress "free" experimental data.

B.3 Electrical simulation setup.

The output of these process and mechanical simulations is then transferred to the device simulation tool. Electrical simulations are performed with either averaged or as-simulated non homogeneous strain fields. The constant strain is obtained by averaging each strain component over the channel region. In the case of the non constant strain field, the device simulation is made of several 2D devices mounted in parallel in order to account for the gradients in the third direction. This allows to spare computationally prohibitive 3D device simulations. However, it is necessary to check whether device slices can indeed be treated separately. Therefore a 2D resistor has been simulated assuming the L×W surface strain field; the resistor standing for a MOSFET in the on-state regime. The simulated current flow lines turned out to be nearly parallel. This suggests that each slice can be treated independently.

B.4 Results.

Figure6 compares the current enhancement ratios obtained by 2D and pseudo 3D simulations with experimental data (750MPa intrinsic stress - 50nm thick CESL). A good agreement is achieved with both methods. Even though the pseudo 3D simulations lead to more accurate results than the constant strain field approach, the latter does capture most of the experimental trends. As observed with MC simulations, the enhancement ratio is too high for long devices. This can also be related to the fact that the saturation velocity is kept constant whatever the gate length.

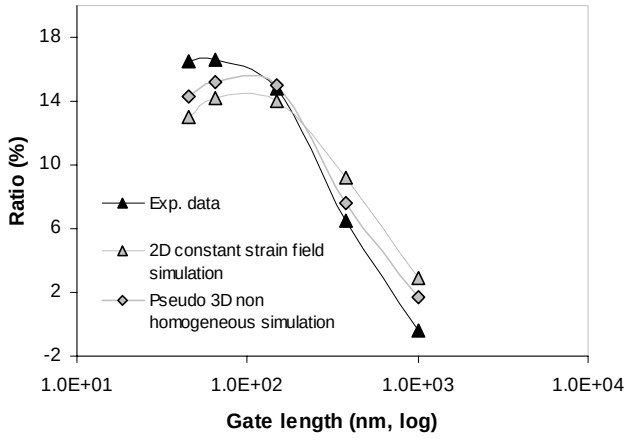

Figure6: Comparison of the current enhancement ratio as a function of the gate length with experimental data for a 50nm – 750MPa CESL.

In Figure7, homogeneous strain field results are plotted for a 30nm–700MPa CESL. The experimental results are correctly reproduced as well.

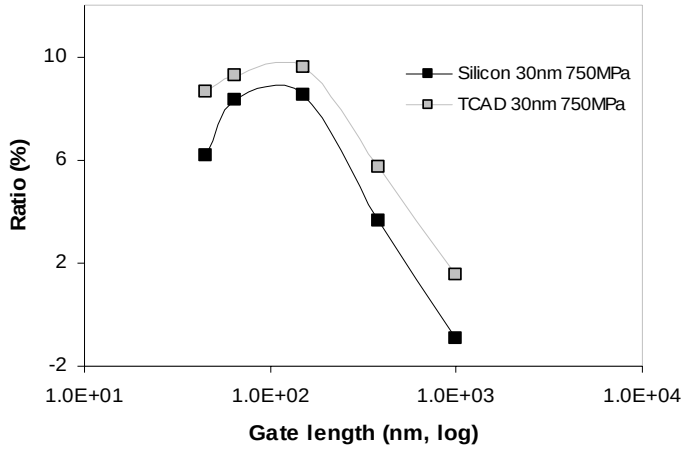

Figure7: Comparison of the current enhancement ratio as a function of the gate length with experimental data for a 30nm – 700MPa CESL.

IV. CONCLUSION

In this work, a procedure aiming at correcting the DD mobility components has been presented. Based on strained full band Monte Carlo simulations, calibrated strain dependent corrections have been added to existing mobility models. First this methodology has been benchmarked against full band MC simulations of $Si_{1-x}Ge_x$ strained devices. A very good agreement has been achieved over a broad range of gate lengths, supporting thereby the relevance of the developed corrections. Moreover it confirms that the saturation velocity, which is often used as a fitting parameter for overshoot effects within a DD framework, does not need to be set strain dependent. Last, comparisons with devices strained by tensile contact etch stop layers have been carried out as a function of the gate length. The simulated 3D strain fields have been handled using two averaging techniques. A good agreement has been achieved whatever the gate length. Interestingly using averaged constant strain fields allows to correctly reproduce experimental data. This result gives insight on the relevance of using strained MC simulations, which are bound to constant strain fields, to assess the impact of process induced strain.

ACKNOWLEDGMENTS

This work was partly funded by the European NANOCMOS project.

REFERENCES

[1] : T. Ghani, et. al, IEDM Technical Digest, pp. 978-980, 2003.
[2] : Z. Wang, *Modelisation de la piezoresistivite*, Ph.D. thesis, Lille 1994.
[3] : T. Guillaume, et. al, ESSDERC Proceedings, pp. 383-386, 2004.
[4] : H. M. Nayfeh, et. al, IEDM Technical Digest, pp. 475-478, 2003
[5]: S. Takagi, J. Appl. Phys., vol. 80, pp. 1567-1577, 1996.
[6]: F. M. Bufler et.al. IEEE T.E.D., vol. 50, pp. 278-284, 2003.
[7]: M.Lundstrom, *Fundamentals of carrier transport*, 2nd edition, Cambridge University Press, 2000.