A Monte Carlo investigation of the RF performance of partially-depleted SOI MOSFETs

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2006 Semicond. Sci. Technol. 21 273
(http://iopscience.iop.org/0268-1242/21/3/010)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 212.128.175.120
This content was downloaded on 05/09/2016 at 18:15

Please note that terms and conditions apply.

You may also be interested in:

A physically based investigation of the small-signal behaviour of bulk and FD SOI MOSFETs for microwave applications
R Rengel, D Pardo and M J Martín

Monte Carlo analysis of dynamic and noise performance of submicron MOSFETs at RF and microwave frequencies
R Rengel, J Mateos, D Pardo et al.

Numerical and experimental study of a 0.25 micron FD SOI MOSFET
R Rengel, J Mateos, D Pardo et al.

Towards the nanoscale: influence of scaling on the electronic transport and small-signal behaviour of MOSFETs
Raúl Rengel, Daniel Pardo and María Jesús Martín

Noise in strained Si MOSFETs for low-power applications
K Fobelets and J E Velázquez-Pérez

2D ensemble Monte Carlo modelling of bulk and FD SOI MOSFETs
R Rengel, D Pardo and M J Martín-Martínez
A Monte Carlo investigation of the RF performance of partially-depleted SOI MOSFETs

Raúl Rengel¹, María J Martín¹, Gilles Dambrine² and François Danneville²

¹ Departamento de Física Aplicada, Universidad de Salamanca, 37008 Salamanca, Spain
² Département Hyperfréquences et Semiconducteurs, IEMN, F-59655 Villeneuve d’Ascq, France

E-mail: raulr@usal.es

Received 26 September 2005, in final form 28 November 2005
Published 20 January 2006
Online at stacks.iop.org/SST/21/273

Abstract

A Monte Carlo investigation of the high-frequency performance of partially-depleted silicon-on-insulator MOSFETs is presented. From static current–voltage characteristics to the RF dynamic and noise performance, the simulator is able to properly reproduce the experimental measurements for the device under test with the need of only two model parameters, the rate of diffusive surface scattering mechanisms and the threshold voltage. Once the model is validated for a 90 nm effective gate length device, a forecast of the dynamic and noise performance of the device when the gate length is downscaled is carried out and the influence of scaling on internal transport quantities is discussed.

1. Introduction

Among the different alternatives proposed to eliminate or mitigate the problems appearing in sub-100 nm bulk silicon MOSFETs, silicon-on-insulator (SOI) technology has emerged as the most feasible technology due to its high level of compatibility and improved features offered by SOI-based MOSFET devices as compared to bulk counterparts [1]. In the field of high-frequency analogue applications, a physically based modelling of the dynamic and noise performance of the transistors becomes mandatory in order to achieve a complete development of SOI-based RF IC’s. To this end, significant effort has been focused on the high-frequency modelling of fully-depleted (FD) SOI transistors. However, in the short term, partially-depleted (PD) SOI MOSFETs represent the most solid alternative to replace bulk transistors in industrial production lines; so a detailed investigation of the RF properties of such devices is urgent. To carry out such a study, the ensemble Monte Carlo (MC) technique [2] is one of the best suited methodologies due to its stochastic and microscopic nature, which intrinsically incorporates important phenomena in scaled devices such as hot carrier effects or non-stationary transport. It can provide the noise behaviour of the transistors in a natural way without the necessity of defining the primary noise sources or making any pre-assumption about its physical origin, and it also allows us to determine the dynamic parameters and the high-frequency small-signal equivalent circuit [3, 4].

In this work, we have employed our in-house MC device simulator to analyse the static, dynamic and high-frequency noise behaviour of 90 nm effective gate length experimental PDSOI MOSFETs. To the authors’ knowledge, this is the first time that a Monte Carlo device simulator is applied to the investigation of the high-frequency performance of a PDSOI transistor. Together with the results for the experimental device, the MC simulator is employed to evaluate the consequences of reducing the effective gate length by performing additional simulations considering \( L_{\text{eff}} = 60 \) and \( 30 \) nm. Our objective therefore is not to perform a rigorous scaling, but to check the limits for an acceptable performance of the experimental structure considered, thus helping to guide the efforts of designers. This paper is organized as follows: in section 2, a brief description of the device under test, the simulated structure and the Monte Carlo procedure are presented. The main static characteristics, together with some internal quantities of interest and a discussion about the modifications of the charge transport conditions when \( L_{\text{eff}} \) is scaled are analysed in section 3. In section 4, the RF dynamic

0268-1242/06/030273+06$30.00 © 2006 IOP Publishing Ltd Printed in the UK
and noise performance of the device is examined, and finally the main conclusions of our work are drawn.

2. Simulated device and Monte Carlo procedure

The device under test is a body-contacted PDSOI MOSFET processed by ST microelectronics [5]. The effective gate length $L_{\text{eff}}$ is 90 nm, the active layer thickness is 150 nm and the gate oxide thickness is 2 nm; the buried oxide thickness is 400 nm and the overlap length is 15 nm. The device has 30 fingers each with 2 $\mu$m device width. The threshold voltage is 0.3 V. The 2D simulated structure (figure 1) reproduces the main features of the experimental device topology, the device width considered as the non-simulated dimension; hence, most of the results will be presented normalized per unit width. The doping of the $p$-body is $2 \times 10^{18}$ cm$^{-3}$ and the doping of the n$^+$ regions is $10^{20}$ cm$^{-3}$. While in the fabricated device the body contact is not placed all along the active layer below the gate, but at the lateral sides in the non-simulated dimension, in the 2D MC structure we have placed a small contact (short-circuited to the source) at the bottom of the active layer to mimic the performance of the body contact, thus allowing the flux of holes and avoiding undesired floating-body effects.

The numerical simulations have been carried out by means of an ensemble Monte Carlo device simulator self-consistently coupled to a Poisson solver. The main features of the simulator can be found elsewhere [6, 7] and references therein. In the present paper, an empirical combination of 20% of diffusive surface scatterings and 80% of specular reflections was considered in order to reproduce the experimental transconductance value. Impurity scattering, acoustic and optical phonons are also considered. Access resistances $R_S = 1.5$ $\Omega$ and $R_D = 1.3$ $\Omega$, provided by the experimental measurement extraction procedure, are incorporated in the simulation following the procedure indicated in [7].

3. Static characteristics

To evaluate the effects of scaling the gate length, we considered three different values in the simulations for $L_{\text{eff}}$: 90 nm (corresponding to the experimental device), 60 nm and 30 nm. All the rest of the topology parameters, as well as the source and drain access resistances, are kept the same, including the gate oxide $t_{ox}$ (scaling this parameter below 2 nm could produce undesired leakage phenomena). The output $I_D$–$V_{DS}$ characteristics for the three devices considered are shown in figures 2(a), (b) and (c), together with the values of $I_D$ as a function of the gate overdrive for $V_{DS} = 1.2$ V (the inset of figure 2(a)). An excellent agreement is found between the numerical MC results for $L_{\text{eff}} = 90$ nm and the experimental measurements, particularly in the saturation regime, which is the fundamental bias condition for analogue RF applications and the conditions for which the experimental high-frequency dynamic and noise measurements were carried out. When the gate length is reduced, an important increase of $I_D$ is observed in the output characteristic: at $V_{DS} = 1.2$ V and $V_{GS} = 0.8$ V, the values of $I_D$ are 0.31, 0.51 and 1.18 mA $\mu$m$^{-1}$ for $L_{\text{eff}}$ equal to 90, 60 and 30 nm, respectively. The dramatic increase of $I_D$ for $L_{\text{eff}} = 30$ nm is due to several reasons: first of all, the noticeable reduction in the threshold voltage, which is reduced by 0.1 V for $L_{\text{eff}} = 60$ nm and 0.5 V for the shortest device (at high $V_{DS}$); a much stronger short channel effect takes place in this case, as will be discussed afterwards. Moreover, the analysis of several internal quantities provided by the MC simulation also offers interesting information. As the device length is reduced, the electron concentration in the inversion layer (not shown in the graphs) is reduced due to the increased influence of the depletion region associated with the drain–substrate junction (especially at high $V_{DS}$). However, a relevant fact takes place concerning the electron velocity (figure 3). For all the values of $L_{\text{eff}}$ considered, a velocity overshoot phenomenon [8] takes place at the drain end of the channel, which is the signature of non-stationary effects. As the gate length is reduced, this velocity overshoot area becomes progressively more important: the percentages of the channel for which the electrons are under velocity overshoot conditions are 39%, 56% and 86% for $L_{\text{eff}}$ equal to 90, 60 and

Figure 1. Scheme of the simulated MC structure.

Figure 2. Output $I_D$–$V_{DS}$ characteristics for the three devices considered. The inset of figure (a) shows the transfer characteristics as a function of the gate overdrive for $V_{DS} = 1.2$ V.
considered: black lines correspond to the gate Si–SiO\(_2\) interface and found that while the peak value of the longitudinal electric field and in this case the injection of carriers in the active layer induces the well-known phenomenon of the barrier lowering, barrier hardly changes from 90 to 60 nm. However, in the case the active layer. In the 90 and 60 nm transistors, a high barrier is observed some tens of nanometres below, at the middle of

\[ V_{\text{GS}} = 0.8 \text{ V} \text{ and } V_{\text{DS}} = 1.2 \text{ V} \text{ for the three values of } L_{\text{eff}} \text{ considered.} \]

The value of saturation velocity in bulk silicon is included as a reference.

Immediately under the Si–SiO\(_2\) interface, it can be observed (black lines) and at the middle of the active layer (grey lines). For the three effective gate lengths 90, 60, and 30 nm. Since the carrier concentration decreases when decreasing \( L_{\text{eff}} \), the increase of drain current is attributed to the superior average velocity of carriers within the whole channel, which in fact are undergoing quasiballistic transport for the shortest device.

To further analyse this enhanced short channel effect, we examined the profile of the conduction band along the \( X \) dimension for two different vertical locations, just below the gate oxide (black lines) and at the middle of the active layer (grey lines). Immediately under the Si–SiO\(_2\) interface, it can be observed that at the source overlap region the band conduction is lowered around 0.1 eV, thus allowing the injection of electrons from the \( N^+ \) area to the channel. But the most remarkable situation is observed some tens of nanometres below, at the middle of the active layer. In the 90 and 60 nm transistors, a high barrier avoids the injection of electrons in that area: the height of the barrier hardly changes from 90 to 60 nm. However, in the case of the 30 nm transistor, the proximity of the drain to the source induces the well-known phenomenon of the barrier lowering, and in this case the injection of carriers in the active layer volume is favoured, thus additionally increasing current in the channel. In good accordance with this behaviour, it has been found that while the peak value of the longitudinal electric field near the drain hardly changes with \( L_{\text{eff}} \), the average value is clearly increased in the rest of the channel, thus producing higher electric fields within the whole channel for \( L_{\text{eff}} = 30 \text{ nm} \).

It is interesting to remark on an important feature of the 30 nm device. While for the largest devices a concentration of holes is maintained in the active layer, for the shortest transistor the depletion regions associated with drain and source wells yield to the elimination of majority carriers in the body; as a consequence, the device no longer operates as a partially-depleted transistor, but as a fully-depleted one, thus explaining the important reduction in the conduction band profile at the middle of the active layer and the significant reduction of the threshold voltage.

In conclusion, while the 90 nm effective gate length device still shows a usual picture for the main internal quantities of interest (although non-stationary phenomena are already relevant), and for \( L_{\text{eff}} = 60 \text{ nm} \) this condition is relatively maintained, for \( L_{\text{eff}} \) equal to few tens of nanometre the proximity of source and drain regions yields a rather different situation, with severe non-stationary conditions and high electric fields within the channel of the device, which is due to the aggressive scaling of the gate length, keeping some important topology parameters unchanged, like the body doping or the gate oxide thickness. This will have important consequences on the RF performance of the device, as we shall see in the next section.

### 4. High-frequency performance

The high-frequency performance of the device was carried out through the determination of the small signal equivalent circuit (SSEC) and the four noise parameters of the device. The SSEC extraction was made following a method originally developed for III–V FETs described in [9], but adapted to SOI MOSFETs by using a direct extraction to determine the series resistances [10]. The method used to extract the four noise parameters of the device is described in [11].

The Monte Carlo simulation of the dynamic performance of the transistor is made by considering the usual procedure [3]: from a stationary bias condition, small voltage steps are applied at terminals. Fourier analysis of the transient response of gate and drain currents gives the four complex admittance parameters \( Y \), and from them a non-quasi-static small-signal equivalent circuit representation is obtained [11]. For the study of noise, a typical two-port device configuration with two noise current generators \( S_{\text{IG}} \) and \( S_{\text{ID}} \), one at the input and the other at the output and both correlated, is considered [12]. The values of \( S_{\text{IG}} \), \( S_{\text{ID}} \) and the cross-correlation are directly obtained in the Monte Carlo simulation through the analysis of instantaneous current fluctuations at terminals in long-time simulations [6]. From the admittance parameters and the noise current sources, the main circuitual noise parameters and figures of merit (minimum noise figure \( \text{NF}_{\text{min}} \), equivalent noise resistance \( R_{\text{neq}} \), complex optimum reflection coefficient \( \Gamma_{\text{opt}} \) and associated gain \( G_{\text{smax}} \)) are calculated following an analogous procedure to that shown in [13] and also incorporating the contribution of parasitics, thus giving the extrinsic noise performance of the device.
transconductance is shown in figure 5 as a function of the gate overdrive for the three values of $L_{\text{eff}}$ under consideration for a fixed $V_GS$ equal to 1.2 V. Labels are the same as in figure 2.

Figure 5(a) shows the values of the gate-to-source capacitance $C_{GS}$ as a function of the gate overdrive. As can be observed, when reducing the effective gate length the values of $C_{GS}$ in strong inversion are also reduced; simultaneously, there is a manifest loss of gate control on $C_{GS}$, which shows a small variation with $V_GS-V_T$ in the case of $L_{\text{eff}} = 60$ nm and no dependence at all for $L_{\text{eff}} = 30$ nm. In this case, parasitic overlap capacitances (which value is estimated around $3 \times 10^{-10}$ F m$^{-1}$) dominate over the intrinsic channel charge capacitance. This loss of gate control is accompanied by a significant degradation of the profile of the electric field (as compared to the ideal shape for a well-tempered MOSFET) and a significant velocity overshoot in the entire channel, as previously discussed. As a result, the transistor turns into a drain-controlled device for extremely short gate lengths. The transconductance is shown in figure 5(b). For the 90 nm device, a very good agreement is obtained for the MC values for this figure of merit and the experimental measurements. As $L_{\text{eff}}$ is diminished, from the results shown in the graph $g_m$ is expected to be visibly improved, with maximum values around 1100 and 1500 S m$^{-1}$ for $L_{\text{eff}}$ equal to 60 and 30 nm, respectively. Taking into account the results for $g_m$ and the total gate capacitance (adding $C_{GS}$ and the gate-to-drain capacitance $C_{GD}$, which has been found to be barely dependent on $L_{\text{eff}}$ and $V_GS-V_T$ and with a value around $5 \times 10^{-10}$ F m$^{-1}$), the cut-off frequency $f_T$ (figure 5(c)) has been determined. The good agreement between the experimental measurements and the MC simulation is remarkable, with a maximum value of $f_T$ equal to 87 GHz for $V_GS = 0.75$ V and $V_GS = 1.2$ V. When the effective gate length is reduced to 60 nm, the maximum value of the intrinsic $f_T$ rises to 160 GHz, and if $L_{\text{eff}}$ is scaled to 30 nm, the maximum $f_T$ would be around 330 GHz. A dramatic improvement of this figure of merit is therefore to be expected when downsizing $L_{\text{eff}}$; however, from the results obtained for $C_{GS}$ and $g_m$ it can be concluded that the improvement in $f_T$ for $L_{\text{eff}} = 30$ nm is achieved mainly because of an undesired loss of gate control ($C_{GS}$ drops near a 70% for the smallest $L_{\text{eff}} = 90$ nm if considering a long channel analytical model.

Let us discuss now the main results obtained for the high-frequency noise performance of the PDSOI MOSFET under consideration. The results for the spectral density of drain current fluctuations $S_{ID}$ (which shows white noise behaviour in the frequency range under consideration) are shown in figure 6. First, let us focus on the results for $L_{\text{eff}} = 90$ nm (corresponding to the experimental device). Together with the values obtained with the MC simulation, the results predicted by the long-channel theory [12] are also shown (dashed line).

As can be observed, in the strong inversion significantly higher values than the theoretical prediction are obtained, especially as $V_GS$ is increased, which is the signature of excess noise at the drain. As the gate length is scaled, the values of $S_{ID}$ are progressively augmented, so under constant bias conditions it can be concluded that the device becomes noisier. However, this is not strictly relevant: what is important is how this noise quantity is combined with the dynamic performance. To analyse this, normalized noise parameters are usually employed [14, 15]. The values of the intrinsic drain noise parameter $P$, which is defined as

$$ P = \frac{S_{ID}}{4K_BT|V_{GS} - V_T|} $$

reach maximum values of 2.8, 5.2 and 5.8 for $L_{\text{eff}}$ 90 nm, 60 nm and 30 nm, respectively. Therefore, for the bias conditions considered (constant $V_GS$ and gate overdrive) the excess noise is progressively increased as $L_{\text{eff}}$ is scaled down, but the augmentation takes place mainly when passing from 90 nm to 60 nm gate length.

Finally, let us analyse the results for the main circuitual noise figures of merit. Figure 7 shows the results for the minimum noise figure at 6 GHz as a function of $V_GS-V_T$ for $V_GS = 1.2$ V. The inset shows the results for the associated gain $G_{\text{ass}}$ under those same bias conditions. The MC simulation is able to reproduce reasonably the experimental measurements for $L_{\text{eff}} = 90$ nm; the minimum value of $NF_{\text{min}}$ is near 0.6 dB at 6 GHz, which confirms the exceptional performance of the fabricated device to be used in high-frequency analogue designs. If the effective gate length is reduced to 60 nm, an improvement in $NF_{\text{min}}$ is to be expected at a low gate voltage, which is significant for low-noise low-voltage applications. For $L_{\text{eff}} = 30$ nm clearly lower values of $NF_{\text{min}}$ (which is
shifted towards higher values of $V_{GS}-V_T$ would be obtained according to the MC simulation. The main reason for this improvement is to be found in the important augmentation of the cut-off frequency of the device in this case (figure 5) and not in the intrinsic noise performance, since according to the study of noise sources, in fact the smaller is the device the noisier it becomes. Therefore, it is the dynamic performance of the transistor which provides an in principle ‘better’ noise performance from the point of view of $\text{NF}_{\text{min}}$ for shorter gate lengths. However, as already discussed, the improvement of the dynamic figures of merit (and consequently, the better $\text{NF}_{\text{min}}$) must be considered with caution due to the important loss of gate control over the channel charge for the bias conditions considered and the negative threshold voltage observed for $L_{\text{eff}}=30$ nm, which results in an inadequate switching device operation (the gate terminal shows difficulties to completely turn off the transistor). This undesired effect could be mitigated by an adequate scaling of other topology parameters or bias applied, or with satisfactory source/drain engineering. The detrimental performance due to the degradation of the electrical behaviour of the transistor is evidenced in the results for $G_{\text{ass}}$, which is improved for $L_{\text{eff}}=60$ nm but shows an important reduction for the shortest transistor (the inset of figure 7) and consequently a smaller gain in minimum noise operation conditions.

Although the MC results for $R_n$ and $\Gamma_{\text{opt}}$ show a certain quantitative discrepancy with the experimental measurements, the bias dependences are correctly reproduced, and qualitatively the tendencies when $L_{\text{eff}}$ is reduced can be extrapolated to give useful information. The equivalent noise resistance $R_n$ is not significantly changed when scaling $L_{\text{eff}}$ (figure 8(a)); in this case, the augmentation of the intrinsic noise source $S_{\text{ID}}$ is compensated by the increase in the transconductance of the device, and consequently only slight differences are observed. In contrast, the optimum reflection coefficient shows an increase of the module (which is closer to 1) and a reduction of the phase as the effective gate length is scaled down (figure 8(b)), thus showing more difficulties to build the matching network at the transistor input for minimum noise operation when reducing the gate length.

5. Conclusions

A Monte Carlo investigation of the static and high-frequency performance of body contacted PD SOI MOSFETs with sub-100 nm effective gate length has been presented for the first time. In general, the Monte Carlo simulator correctly reproduces the values and bias dependences for the experimental measurements for the main characteristics and figures of merit under saturation conditions, thus confirming the reliability of the simulator, which has also been used to predict the performance of the transistor when scaling down the gate length. An important non-stationary transport is observed at the drain end of the channel, with a significant velocity overshoot effect which becomes progressively more important as the dimensions are shrunk. The simulation results indicate that for an effective gate length of 60 nm, the main dynamic and noise figures of merit could be improved without a severe degradation of the device behaviour. However, for even more reduced values of $L_{\text{eff}}$ (30 nm), the partial depletion of the body no longer takes place and the device turns into a fully-depleted transistor. The electrical performance of the device is degraded due to an important loss of gate control when source and drain depletion regions interact with each other and the overlap parasitic capacitances dominate over the channel charge capacitance. Although the minimum noise figure would be improved by reducing the effective gate length, the worse short channel effect and gain of the transistor, and the non-proper operation of the device would require a more sophisticated scaling or a significant improvement of source/drain engineering to avoid the problems related to the degradation of the electrical performance. Then, the need for designing a ‘well-tempered’ structure becomes critical if a good high-frequency noise performance is desired for devices with gate lengths near or below 30 nm.

Acknowledgments

The authors wish to thank Christine Raynaud from CEA-LETI, in charge of SOI CMOS Process Integration in ST Microelectronics. This work was supported by the research
References