phys. stat. sol. (c) **5**, No. 1, 123–126 (2008) / **DOI** 10.1002/pssc.200776517

physica status solidi c
www.pss-c.com
*current topics in solid state physics*

# Onset of quasi-ballistic transport and mobility degradation in ultra scaled MOSFETs: a Monte Carlo study

**M. J. Martín[*,1], R. Rengel[1], E. Pascual[1], J. Łusakowski[2,3], W. Knap[2], and T. González[1]**

[1] Departamento de Física Aplicada, Universidad de Salamanca, Pza de la Merced s/n, 37008 Salamanca, Spain
[2] GES CNRS-UMR5650, Université Montpellier II, 34095 Montpellier, France
[3] Institute of Experimental Physics, Hoza 69, 00-681 Warsaw, Poland

Received 6 July 2007, revised 4 September 2007, accepted 3 October 2007
Published online 15 November 2007

**PACS** 02.70.Uu, 73.23.Ad, 85.30.Tv

[*] Corresponding author: e-mail mjmm@usal.es, Phone: +34 923 294 436, Fax: +34 923 294 584

A Monte Carlo investigation of quasi-ballistic transport in scaled MOSFETs is presented. Taking into account a realistic velocity and electric field profile in ultra-scaled transistors, the carrier mobility is calculated, observing a degradation of this macroscopic parameter strongly related to the peculiarities of the electrostatic profiles found for extremely short gate lengths. Simultaneously, the onset of ballistic phenomena within the channel is detected. A detailed investigation of the movement of carriers inside the channel allows providing numerous microscopic quantities related to electronic transport, such as mean free paths, transit times or average number of scattering mechanisms. Results evidence the existence of a significant proportion of ballistic carriers as the gate length is scaled down to few tens of nanometre.



**1 Introduction** The aggressive scaling of Silicon MOSFETs has been the driving force of microelectronics along the last decades, providing a spectacular improvement of the performance of the transistors. However, as the dimensions of the devices reach the nanometre scale, short channel effects, non-equilibrium transport and hot carrier phenomena become relevant and may affect the transistor behaviour as compared to long channel devices [1]. In particular, when the gate length ($L_g$) of the transistors becomes of the order of the carrier mean free path ($\lambda$), transport is expected to become strongly quasi-ballistic. As a consequence, conductivity becomes a highly non-equilibrium process, and mobility ($\mu$) is severely affected. Recent studies applying the magnetorresistance method [2] have shown a significant decrease of the electron mobility within MOSFET channels as the gate length is scaled down.

The determination of mobility in quasi-ballistic devices is not an obvious task, which requires the use of advanced models. Several papers have dealt with the study of mobility in the inversion layers of MOSFETs by means of ensemble Monte Carlo device simulators, e.g. [3–6]; in those papers, mobility is calculated under a constant and spatially uniform longitudinal electric field (which is the case of long channel devices) and is discussed as a function of the effective electric field in the direction perpendicular to the gate contact.

In this work we have performed a Monte Carlo investigation of transport in scaled MOSFET devices; the two-dimensional (2D) solution of the real electric field profile obtained by solving the Poisson equation in the whole channel is taken into account in the mobility calculations, thus providing a realistic consideration of the features of small size devices. Moreover, values of important microscopic quantities related to transport are also extracted, such as mean free paths, transit times, average number of scatterings within the channel, etc. This permits to identify the influence of the onset of quasi-ballistic phenomena on the degradation of the channel mobility for extremely short transistors.

The paper is organized as follows. In Section 2, the main features of the Monte Carlo simulator are presented, together with details about the simulated structures and the procedure followed to obtain the electron mobility. The

main results of our work are shown and discussed in Section 3. Finally, the main conclusions are presented.

**2 Monte Carlo simulation** A 2D bipolar ensemble Monte Carlo simulator, self-consistently coupled to a Poisson solver, has been employed for the numerical calculations. The simulator has been already successfully exploited for the study of different Silicon devices, such as BJTs, HBTs and MOSFETs [7, 8]. Details about the conduction and valence bands can be found in Ref. [7]. The scattering mechanisms considered are acoustic and nonpolar optical phonons and ionized impurities scattering, together with impact ionization and surface roughness scattering, as shown in [8] and references therein. Impurity scattering is treated via the Brooks Herring approximation [9]. Coulomb screening of the impurity charge and Pauli's exclusion principle are also taken into account [9, 10]. The devices under test correspond to n-channel bulk Silicon MOSFETs. The simulated structure reproduces the main features of real fabricated devices, including the presence of pockets adjacent to the source and drain regions incorporated to minimize short channel effects. The substrate doping is $5\times10^{18}$ cm$^{-3}$, and the p$^+$ pockets are $10^{19}$ cm$^{-3}$ doped. The depth of the source and drain areas is 25 nm, while the pocket depth is 15 nm. The oxide thickness is 1.8 nm. Several values of $L_g$ have been considered in the simulations, from 250 nm down to 20 nm. To achieve a fair comparison between the scaled transistors, identical inversion charge density ($n_s$) conditions are analyzed, with different values of $n_s$ raging from $5\times10^{11}$ cm$^{-2}$ to $2\times10^{12}$ cm$^{-2}$. The drain-to-source voltage is kept equal to 80 mV, which corresponds to triode bias conditions. In order to solve Poisson equation, an adaptive mesh is considered in the simulation, with the vertical size of the cells varying from 1.25 to 20 Å depending on the local doping and electron concentration. The time step considered is 0.2 fs. The number of simulated particles varies from 35000 to 70000 electrons and 8000 to 40000 holes depending on $L_g$ and the bias conditions. The total simulation time is 100 ps.

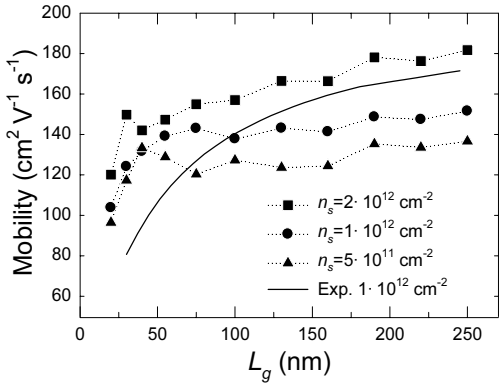


**Figure 1** Mobility *vs*. $L_g$ for several values of $n_s$. Tendency line for the experimental data at $n_s = 1\times10^{12}$ cm$^{-2}$ is also shown.

To determine the microscopic quantities related to charge transport like the mean free path or the transit time, a counting region is considered [11], which in this case corresponds to the entire channel. Each electron accessing the channel is flagged and all the relevant parameters related to its movement are recorded until it reaches the drain. Such study allows us to obtain not only average quantities, but also distribution functions. It must also be pointed out that the fact of entering the drain is considered as a scattering mechanism to avoid unphysical values of the mean free path.

In order to determine the mobility within the channel, a linear dependence of the average velocity on the average electric field within the channel is considered. I.e., a local relationship among the two quantities is not postulated, but a linear dependence between the average values of both, considering the channel as a whole. This definition is equivalent to the mobility experimentally determined through the magneto-resistance method, for which degradation has been observed in ultra-scaled devices [2]. The average velocity and electric field are determined according to the following expressions:

$$\langle v_x \rangle = \frac{\iint \bar{v}_x(x,y)\bar{n}(x,y)dxdy}{\iint \bar{n}(x,y)dxdy} \tag{1}$$

$$\langle E_x \rangle = \frac{\iint \bar{E}_x(x,y)\bar{n}(x,y)dxdy}{\iint \bar{n}(x,y)dxdy} \tag{2}$$

where $\bar{n}(x,y)$ is the average electron concentration at each location. Integrals extend over the whole channel area. In this way, the electron mobility is evaluated as:

$$\mu = \langle v_x \rangle / \langle E_x \rangle \tag{3}$$

**3 Results and discussion** Figure 1 shows the results obtained for the mobility as a function of $L_g$ for several values of the inversion charge density. As it can be observed, the mobility gets augmented with $n_s$, which indicates that in this range of inversion layer concentrations Coulomb screening of ionized impurities is highly effective, while surface scattering is not playing a dominant role [3]. When the gate length is reduced, a slight reduction in the values of $\mu$ is initially observed; however, when the devices are scaled below 50 nm $L_g$, an important drop in the mobility values takes place. This result is in good qualitative agreement with the experimental data reported in [2, 12], thus confirming an overall degradation of the channel mobility as devices are scaled into the nanometre range.

As described in Section 2, the mobility is obtained by assuming a linear relationship between the average velocity and electric field in the whole channel. However, from the information provided by the Monte Carlo simulator it is also possible to analyze the profiles of such quantities

 

within the channel. Figure 2 shows the results for the average electron velocity and average electric field (both properly weighted by the local concentration along the vertical axis) for $L_g$ = 20, 30, 75 and 160 nm, as a function of the longitudinal dimension normalized by the corresponding gate length, for $n_s = 1\times10^{12}$ cm$^{-2}$. For relatively long channels two velocity peaks are obtained at both ends of the channel, related to the electric field profile generated by the presence of the junctions between the channel substrate and the $p^+$ pockets. In the central part of the channel, reduced and practically constant values of the electric field are obtained, and consequently the velocity is also lower in that area. However, as $L_g$ is shrunk, pockets are much closer and even enter into contact. Simultaneously, source and drain junctions are extremely close, which significantly modifies the electric field profile, now showing a linear dependence in the whole channel. Consequently, the average velocity of carriers is significantly augmented inside the channel even under low drain bias. Nevertheless, when accounting for the total average value in the whole channel, the $<v_x>$ to $<E_x>$ ratio is lower than what would be expected from the direct scaling of $L_g$.

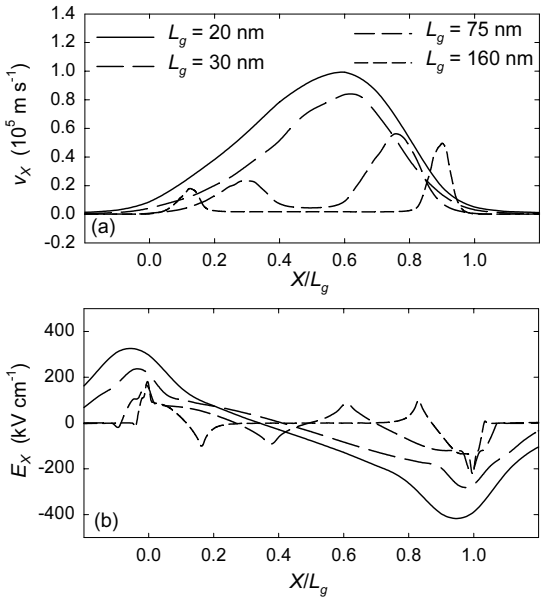


**Figure 2** Profiles of (a) average velocity and (b) electric field. The $X$ dimension is normalized by $L_g$ in order to ease the comparison. The values of $V_{GS}$ range from 0.59 V for $L_g$ = 160 nm to 0.22 V for $L_g$ = 20 nm in order to keep a constant $n_s = 1\times10^{12}$ cm$^{-2}$.

In order to evaluate the microscopic features of transport, we have also carried out the calculation of some relevant parameters as described in Section 2. Figure 3 shows the values obtained for the average transit time of carriers crossing the channel, $t_t$, together with the average time between scattering events, $\tau$. For long channel devices, as the density of inversion carriers is augmented, the transit time decreases due to the previously mentioned screening of charged impurities and shows a linear dependence on $L_g$, while the average time between scatterings slightly increases. As the gate length is reduced, the transit time shortens, while $\tau$ keeps almost constant; this is due to the fact that transport is still mainly diffusive, so $\tau$ is set up by the bulk material and pocket features, and the inversion charge. However, when the gate length is reduced below 50 nm the transit time is not dependent on $n_s$, decreases more pronouncedly with $L_g$, and tends to approach the values of $\tau$, thus indicating, in a first step, the onset of a quasi-ballistic regime of transport.

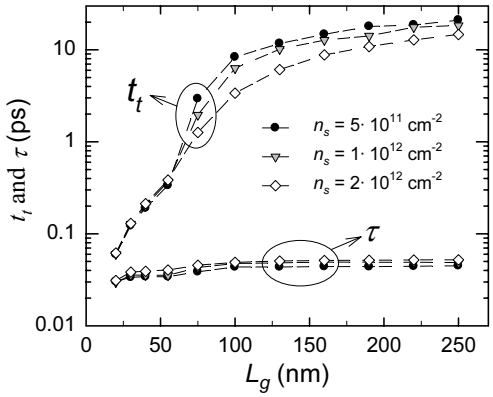


**Figure 3** Transit time and average time between scattering events as a function of $L_g$ for several values of $n_s$.

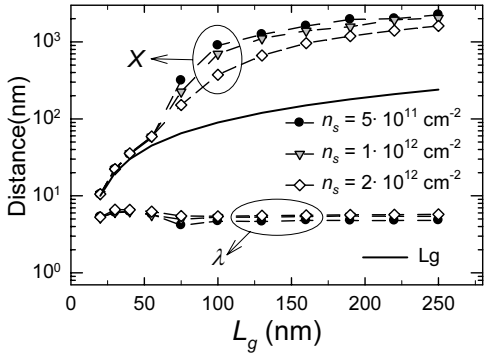


**Figure 4** Average distance travelled by electrons along the longitudinal direction and mean free path as a function of $L_g$ for several values of $n_s$. Solid line represents $L_g$ as a reference.

Figure 4 shows the results for the average distance travelled along the longitudinal dimension ($X$) as a function of $L_g$ for three different values of $n_s$. The value of $L_g$ is also plotted in the graph as a reference. As it can be observed, for long channels the distance travelled along the longitudinal dimension (which comprises both forward and backward movement of carriers in their way across the channel) is almost one order of magnitude larger than $L_g$, which is the signature of the diffusive character of transport in that case. The effect of impurity screening yields lower values of $X$ at higher $n_s$. As $L_g$ is scaled down, transport tends to get more unidirectional and the travelled distance approaches $L_g$, being of the order of the mean free path and not depending on $n_s$ for very short devices.

In order to make this clearer, in Fig. 5 we show the average number of scattering mechanisms for $n_s$ = 1×10$^{12}$ cm$^{-2}$ as a function of $L_g$, including the total number of events, interactions with acoustic and optical phonons, surface scattering phenomena and ionized impurity scattering. In the case of impurity scattering, since the model considered is anisotropic, an equivalent number of isotropic scattering mechanisms is determined by weighting the number of scatterings by the appropriate angular dispersion [9]. As it can be observed, the total number of scatterings significantly decreases with $L_g$, reaching values close to 1 for $L_g$ = 20 nm, which is the signature of a strong quasi-ballistic regime. It is also noticeable the relative increasing weight of surface scattering as the gate length is reduced. When the inversion carrier density is augmented (not shown in the graphs), surface scattering becomes dominant in the whole range of $L_g$ considered; however, the total number of scatterings keeps close to 1 in the case of $L_g$ = 20 nm for all the values of $n_s$ studied.

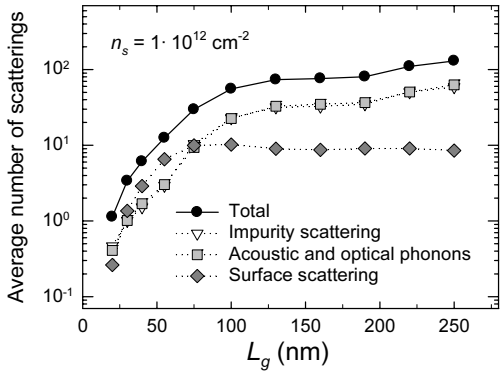


**Figure 5** Average number of scatterings as a function of $L_g$ for $n_s$ = 1×10$^{12}$ cm$^{-2}$.

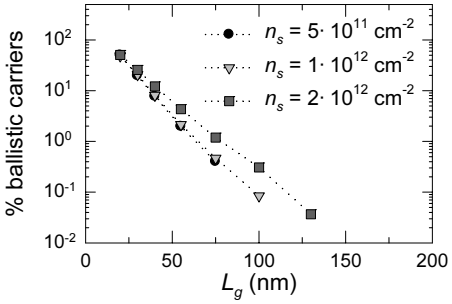


**Figure 6** Percentage of ballistic carriers over the total number of carriers crossing the channel as a function of $L_g$ for several values of $n_s$.

From the data provided by the Monte Carlo simulation, it is also possible to analyze the distribution function of the number of scattering mechanisms. In this way, we can determine the average number of carriers that are able to cross the channel from source to drain without suffering any scattering mechanism at all, this is, the number of purely ballistic carriers. Figure 6 shows the percentage of ballistic carriers as a function of $L_g$ for several values of the inversion charge density. As it can be observed, the percentage of ballistic carriers tends to increase exponentially as the gate length is reduced, being slightly higher for high values of $n_s$, as a consequence of the Coulomb screening of ionized impurities. Again, for ultra-short transistors screening does not affect the results significantly. In the case of the shortest $L_g$ considered, 20 nm, 50% of the total number of carriers crossing the channel corresponds to purely ballistic electrons. Therefore, it can be concluded that as the gate length enters the nanometre range, ballistic transport becomes dominant.

**4 Conclusion** We have presented a Monte Carlo investigation of mobility and quasi-ballistic transport in ultra-scaled Silicon MOSFET devices. The results evidence an important degradation of the channel mobility for gate lengths below 50 nm, which is accompanied by the modification of the electrostatic conditions under constant $n_s$. The appearance of quasi-ballistic charge transport takes place also for this gate length range. The study of the transit time and average time between scattering, together with the total travelled distance and the mean free path, have evidenced the onset of such transport regime. Impurity scattering has been found to play a decreasingly important role as the length of the devices is reduced. Half of the total carriers crossing the channel correspond to purely ballistic electrons for gate length equal to 20 nm.

**Acknowledgements** This work has been supported by research projects METAMOS (IST-016677) financed by the European Commission, SA010A07 from the Consejería de Educación de la Junta de Castilla y León in Spain and CNRS GDR and GDR-E projects "Semiconductor sources and detectors of THz frequencies" in France.

## References

[1] W. Haensch et al., IBM J. Res. Dev. **50**, 339 (2006).
[2] J. Łusakowski et al., Appl. Phys. Lett. **87**, 053507 (2005).
[3] F. Gámiz, J. A. Lopez-Villaneuva, and J. Banqueri, IEEE Trans. Electron Dev. **42**, 258 (1995).
[4] C. Jungemann, N. Subba, J. S. Goo, C. Riccobene, Q. Xiang, and B. Meinerzhagen, Solid-State Electron. **48**, 1417 (2004).
[5] D. Esseni and E. Sangiorgi, Solid-State Electron. **48**, 927 (2004).
[6] V. Aubry-Fortuna, P. Dollfus, and S. Galdin-Retailleau, Solid-State Electron. **49**, 1329 (2005).
[7] M. J. Martín-Martínez, S. Perez, D. Pardo, and T. González, J. Appl. Phys. **90**, 1582 (2001).
[8] R. Rengel, M. J. Martín, T. González, J. Mateos, D. Pardo, G. Dambrine, J.-P. Raskin, and F. Danneville, IEEE Trans. Electron Dev. **53**, 523 (2006).
[9] C. Jacoboni and P. Lugli, The Monte Carlo Method for Semiconductor Device Simulation (Springer, Vienna, 1989).
[10] J. Mateos et al., IEEE Trans. Electron Dev. **47**, 250 (2000).
[11] J. S. Martin, A. Bournel, and P. Dollfus, IEEE Trans. Electron Dev. **51**, 1148 (2004).
[12] J. Łusakowski et al., J. Appl. Phys. **101**, 114511 (2007).

 