A novel 3D Multi-Subband Monte Carlo approach including material-dependent nonparabolic effects

L. Donetti, C. Medina-Bailon, J. L. Padilla, C. Sampedro, F. Gamiz
Nanoelectronics Research Group and CITIC, Universidad de Granada, Granada, Spain.

***Abstract*—Accurate modeling of carrier transport in nanodevices requires incorporating nonparabolic band effects, especially under strong quantum confinement. In this work, we present a novel implementation of nonparabolic corrections in a 3D multi-subband Monte Carlo simulator that can handle nonuniform electric fields and arbitrary confinement profiles. We derive the modified equations of motion accounting for nonseparable energy-momentum-position dependencies and validate our method in Si and InP nanowire transistors, showing that nonparabolicity has a material-dependent impact.**

***Index Terms*—Monte Carlo, multi-subband, nonparabolicity**

I. Introduction

Modeling the carrier band structure is a crucial aspect in the simulation of electronic devices. The effective mass approximation is widely employed, especially for near-equilibrium properties, and nonparabolicity can be added to expand its validity range [1], [2]. However, the standard formulation of nonparabolicity couples the kinetic energy in the transport and lateral directions. Therefore, when low-dimensional (1D or 2D) electron gases with lateral quantum confinement are considered, different approximations have been proposed [2]–[5]. Nonparabolic corrections are especially important when 2D confinement is included because it induces large kinetic energy in the confinement plane [5]. However, the approach in [5] is only valid if the lateral confinement is homogeneous in the transport direction. If the confining potential varies along the device, as in the Multi-Subband Ensemble Monte Carlo (MSEMC) approach [6], nonstandard equations of motion are obtained. In this paper, we derive such equations (Section II), propose an implementation in a 3D MSEMC simulator (Section III), study the impact of nonparabolicity on the device characteristics (Section IV) and state our conclusions (Section V).

II. Nonparabolic corrections for 2D confinement

The energy eigenvalues, E_μ^{p}, of the (parabolic) 2D Schrödinger equation can be written as

$$E_\mu^p = U_\mu + \gamma_\mu \tag{1}$$

This work was supported by Spanish "Secretaria de Estado de Telecomunicaciones e Infraestructura Digital, Ministerio de Transformación Digital" through project "+QCHIP TSI-069100-2023-0003" and by European Chips Joint Undertaking through HORIZON-JU-Chips-2023-RIA-CPL-2 (FAMES ID 101182279).

where U_μ and γ_μ denote the expectation values of the 2D potential $U(x,y)$ and kinetic energy, respectively, for the μ-th eigen-function:

$$\begin{aligned} U_\mu &= \langle U(x,y)\rangle \\ \gamma_\mu &= \left\langle \frac{\hbar^2 k_x^2}{2m_x} + \frac{\hbar^2 k_y^2}{2m_y} \right\rangle \end{aligned} \tag{2}$$

According to [5], nonparabolic correction to energy can be taken into account as:

$$E_\mu(k_z) = U_\mu + \frac{1}{2\alpha}\left(\sqrt{1+4\alpha\left(\gamma_\mu + \frac{\hbar^2 k_z^2}{2m_z}\right)} - 1\right) \tag{3}$$

where α is the nonparabolicity parameter. The subband minimum energy taking into account nonparabolicity can be computed as:

$$E_\mu^{\mathrm{np}} = E_\mu(0) = U_\mu + \frac{1}{2\alpha}\left(\sqrt{1+4\alpha\gamma_\mu} - 1\right) \tag{4}$$

In a simulator based on the mode-space approach [7], the Schrödinger equation is solved in several device cross sections located at positions z_i, providing values for $E_{\mu,i}^p$ and $U_{\mu,i}$. In each interval i, with $z_i < z < z_{i+1}$, we define the functions $E_\mu^{\mathrm{p}}(z)$ and $U_\mu(z)$ as linear interpolations of the respective values at the boundaries (see Fig. 1):

$$\begin{aligned} E_\mu^{\mathrm{p}}(z) &= E_{\mu,i}^{\mathrm{p}} + \frac{z-z_i}{z_{i+1}-z_i}F_{\mu,i} \text{ for } z_i < z < z_{i+1} \\ U_\mu(z) &= U_{\mu,i} + \frac{z-z_i}{z_{i+1}-z_i}V_{\mu,i} \text{ for } z_i < z < z_{i+1} \end{aligned} \tag{5}$$

where the slopes $F_{\mu,i}$ and $V_{\mu,i}$ are defined as:

$$F_{\mu,i} = \frac{E_{\mu,i+1}^{\mathrm{p}} - E_{\mu,i}^{\mathrm{p}}}{z_{i+1}-z_i}, \quad V_{\mu,i} = \frac{U_{\mu,i+1}-U_{\mu,i}}{z_{i+1}-z_i} \tag{6}$$

In this way, the energy of an electron at position z with wave-vector k_z can be computed as:

$$E_\mu(z,k_z) = U_\mu(z) + \frac{\sqrt{1+4\alpha\left(E_\mu^{\mathrm{p}}(z) - U_\mu(z) + \frac{\hbar^2 k_z^2}{2m_z}\right)} - 1}{2\alpha} \tag{7}$$

As a consequence, in $E_\mu(z,k_z)$ the dependence on z and k_z cannot be separated. Moreover, the subband profile $E_\mu^{\mathrm{np}}(z) = E_\mu(z,0)$ (see Figure 1) is not linear in z unless $F_{\mu,i} = V_{\mu,i}$.

Thanks to energy conservation, during a trajectory of an electron with total energy E we can relate the position z

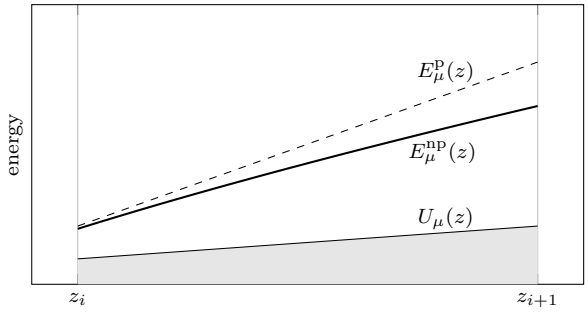

Fig. 1. Energy diagram including the linear interpolations of $E_\mu^{\mathrm{p}}(z)$ and $U_\mu(z)$ between positions z_i and z_{i+1} and the resulting nonlinear $E_\mu^{\mathrm{np}}(z)$.

and the wave-vector k_z: $E = E_\mu(z, k_z)$. In this way, we can compute the absolute value of k_z as a function of z:

$$k_z = \pm\frac{\sqrt{2m_z}}{\hbar}\sqrt{E - E_\mu^{\mathrm{p}}(z) + \alpha\big(E - U_\mu(z)\big)^2} \tag{8}$$

Then, employing (7) and energy conservation, the 1D equations of motion, *i. e.* the time evolution of electron position, z, and wave-vector, k_z, can be computed as:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \pm\sqrt{\frac{2}{m_z}}\frac{\sqrt{E - E_\mu^{\mathrm{p}}(z) + \alpha\big(E - U_\mu(z)\big)^2}}{1 + 2\alpha\big(E - U_\mu(z)\big)} \tag{9}$$

$$\frac{\mathrm{d}k_z}{\mathrm{d}t} = -\frac{1}{\hbar}\frac{F_{\mu,i} + 2\alpha V_{\mu,i}\big(E - U_\mu(z)\big)}{1 + 2\alpha\big(E - U_\mu(z)\big)} \tag{10}$$

where the sign in the first equation depends on the sign of k_z.

Equation (9) can be solved implicitly as:

$$t - t_0 = \pm\big(G(z) - G(z_0)\big) \tag{11}$$

where t_0 and z_0 are the initial time and position, respectively, and

$$G(z) = \sqrt{\frac{m_z}{2\alpha}}\frac{\big(V_{\mu,i} - F_{\mu,i}\big)\ln\big(g(z) - h(z)\big) - g(z)}{V_{\mu,i}^2} \tag{12}$$

$$g(z) = 2V_{\mu,i}\sqrt{\alpha}\sqrt{E - E_\mu^{\mathrm{p}}(z) + \alpha\big(E - U_\mu(z)\big)^2} \tag{13}$$

$$h(z) = F_{\mu,i} + 2\alpha V_{\mu,i}\big(E - U_\mu(z)\big) \tag{14}$$

In (12), the argument of the logarithm, $g(z) - h(z)$, could in principle be negative. However, it can be shown that its sign is preserved during a flight: a possible imaginary part of $G(z)$ cancels out with that of $G(z_0)$. As a consequence, in (12) we can safely employ the absolute value in the argument of the logarithm and consider $\ln\big|g(z) - h(z)\big|$.

III. Implementation in the MSEMC simulator

The previous equations are the building blocks for the implementation of the equations of motion in the Monte Carlo code. The simulation of the flight of an electron with total energy E and duration t_f is represented schematically in the flowchart of Fig. 2 and explained in the rest of this Section.

At the beginning of the flight, we locate the interval i corresponding to the position of the electron and set the direction of its motion, $s = \pm 1$, according to the sign of k_z. This value, s, represent the sign in (9) and (11) and we need to detect when the particle changes its direction of motion. This can happen during a flight only if the electron reaches the barrier defined by the subband profile $E_\mu^{\mathrm{np}}(z)$ (in this semi-classical approach we do not consider the possibility of tunneling). We also need to detect when the electron moves from one interval to an adjacent one, since in this case the values for the interpolations of $E_\mu^{\mathrm{p}}(z)$ and $U_\mu(z)$, including $F_{\mu,i}$ and $V_{\mu,i}$, change. Combining all these considerations, the first step is to check whether the electron can reach the barrier inside the current interval, as in the case depicted in red in Fig. 3. A first trivial condition that can be used to detect the possibility of "rebound" against the barrier is that the nonparabolic energy at the end of the interval (at $E_\mu^{\mathrm{np}}(z_{i+1})$ if $s = +1$ or $E_\mu^{\mathrm{np}}(z_i)$ if $s = -1$) is smaller than E. However, since $E_\mu^{\mathrm{np}}(z)$ is not linear, it could have a maximum inside the interval. This maximum exists and is located inside interval i if the two following conditions hold:

$$\begin{aligned} F_{\mu,i} &> -2\alpha V_{\mu,i}(E_\mu^{\mathrm{np}}(z_i) - U_{\mu,i}) \\ F_{\mu,i} &< -2\alpha V_{\mu,i}(E_\mu^{\mathrm{np}}(z_{i+1}) - U_{\mu,i+1}) \end{aligned} \tag{15}$$

and its value can be computed as:

$$\frac{1}{V_{\mu,i} - F_{\mu,i}}\left(E_{\mu,i}^{\mathrm{p}}V_{\mu,i} - U_{\mu,i}F_{\mu,i} + \frac{F_{\mu,i}^2}{4\alpha V_{\mu,i}}\right) \tag{16}$$

Therefore, a rebound is also possible if the electron energy, E, is lower than such maximum, if it exists.

By checking the previous conditions we detect whether a rebound is possible. In this case, we compute its position, z_r, by solving the equation $E_\mu^{\mathrm{np}}(z_r) = E$ (a quadratic equation in z_r). Then, we need to check if this happens during the requested flight time: we compute the time t_r needed to reach position z_r by employing (11). If $t_r < t_f$, we move the particle to z_r, invert the direction of its motion (by changing the sign s), and simulate the remaining flight with duration $t_f - t_r$.

If the electron energy, E, is larger than the maximum subband energy in the interval, it could reach the boundary position z_b (that is z_{i+1} if $s = +1$ or z_i if $s = -1$), as in the case depicted in blue in Fig. 3. To check if this happens before t_f, we compute t_b, the time needed to reach z_b, by employing (11). If $t_b < t_f$, we move the electron to the boundary of the interval and simulate the flight in the adjacent interval with duration $t_f - t_b$.

If the electron does not reach either the barrier or the boundary of the interval, we compute the final position z_f by solving numerically the (implicit) equation given by (11). This computation is performed numerically through the iteration of Newton's method until a given tolerance is reached. Once the position is known, the final value of the electron wave-vector, k_z is obtained employing energy conservation (8). Now, the sign is kept equal to the initial sign of k_z, since we know that the electron does not change its direction in the remaining time.

IV. Results

The 3D MSEMC simulator [6] with nonparabolic corrections has been employed to study cylindrical nanowires with

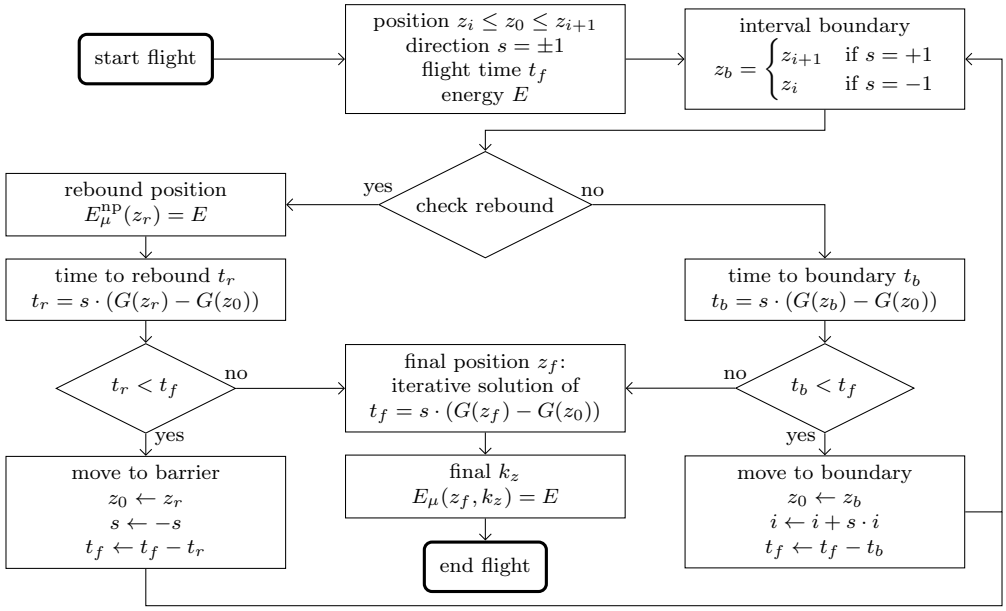

Fig. 2. Flow diagram of the numerical simulation of the dynamics of an electron inside a device slice between z_i and z_{i+1}, with flight time t_f.

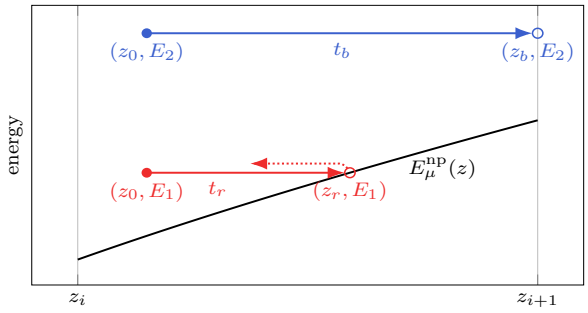

Fig. 3. Schematic motion of an electron with energy E_1 that would rebound against the subband profiles, and another one with energy E_2 which can reach the boundary of interval i.

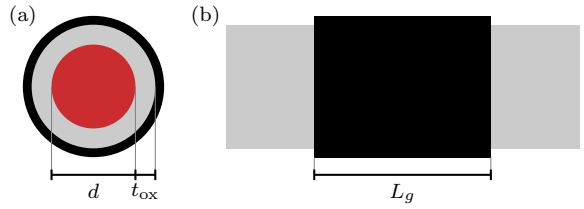

Fig. 4. (a) Cross section and (b) lateral view of the simulated devices.

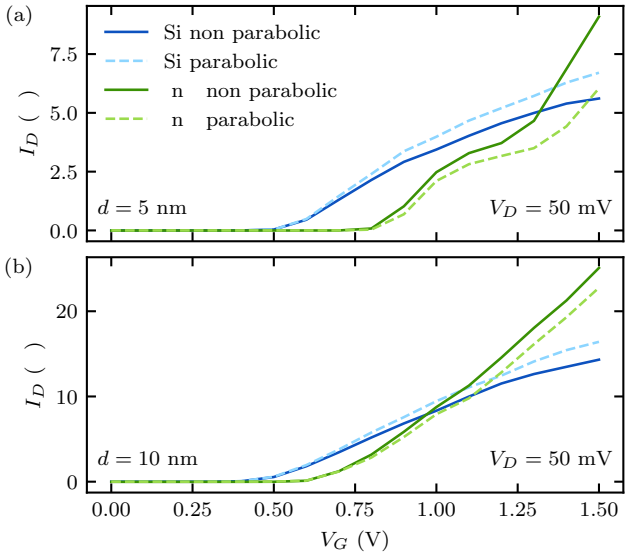

Fig. 5. I_D-V_G curves of transistors based on Si and InP nanowires with diameter (a) $d = 5$ and (b) $d = 10$.

Si and InP channel with diameter $d = 5$ and $d = 10$ (see Fig. 4). Other common parameters employed are: gate length $L_G = 15$ gate insulator thickness $t_{\rm ox} = 1.5$ (SiO_2 and Al_2O_3 for Si and InP, respectively), and mid-gap metal gate work function. For Si conduction band valleys, we employ $\alpha = 0.5$ [2]; for InP, we consider parameters from [8], with $\alpha = 0.61$ for the Γ valley and neglect nonparbolicity for the L valleys (for which [8] recommends a negative value in the longitudinal direction and a positive value in the transverse direction).

Fig. 5 compares the transfer characteristics in linear regime ($V_d = 0.05$) of these devices with and without taking into account nonparabolicity. It can be observed that the non-parabolic correction decreases the drain current, I_D, for Si devices but increases it for InP devices for all nanowire diameters and all considered values of V_G. For both materials, the variation induced by nonparabolic corrections increases for narrower nanowires, due to the larger in-plane kinetic energy induced by confinement. For the narrower ($d = 5$)

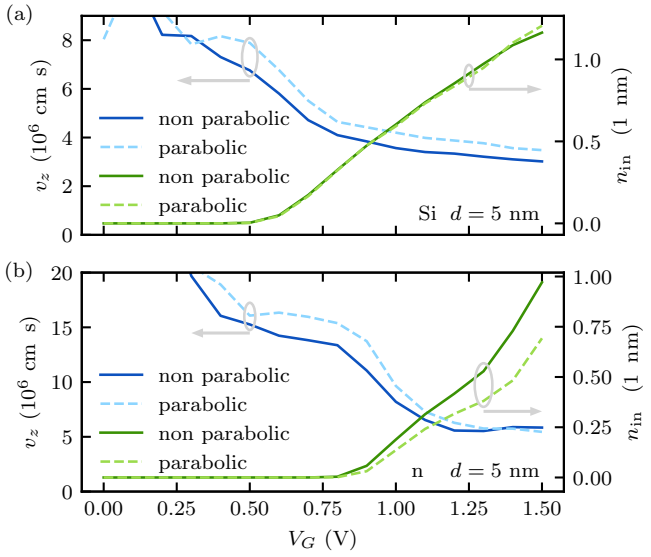

Fig. 6. Maximum velocity v_z and linear electron density n_{inv} at position z_p where velocity is maximum, as a function of V_G for (a) Si- and (b) InP-based transistors.

InP nanowire, a large increase of I_D can be observed around $V_G = 1.25$ due to the onset of conduction of the L valleys.

To investigate the cause of the different behavior in Si and InP nanowires, we compute the average electron velocity in the channel direction, v_z, and the linear inversion density n_{inv}. Fig. 6 shows both quantities computed for nanowires with $d = 5$ (for which the effect is stronger) at position z_p where v_z reaches its maximum. In the case of Si, the inversion density is essentially the same whether the nonparabolic correction is included or not, while the average velocity is lower if it is considered. This is reasonable, since nonparabolicity decreases the energy and the velocity for electrons with a given value of k_z. This last observation is also true in the case of InP nanowires, *i. e.* the average velocity is lower when nonparabolic correction are included. However, in this case, we also observe an increase of the inversion charge density, which compensates the velocity decrease and produces an increase of I_D.

V. Conclusions

We have presented a novel implementation of nonparabolic corrections within a Monte Carlo transport kernel, integrated into a multi-subband simulator for 3D nano-devices. This approach enables accurate modeling of carrier dynamics across a wide energy range, beyond the limitations of parabolic band approximations. Our results reveal that nonparabolicity significantly influences key device characteristics and we demonstrate that the magnitude and nature of its impact is material-dependent: it reduces current in Si devices due to decreased carrier velocity, while it enhances current in InP devices through increased inversion charge.

References

[1] C. Jacoboni and L. Reggiani, "The monte carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Rev. Mod. Phys.*, vol. 55, pp. 645–705, Jul 1983.

[2] M. V. Fischetti and S. E. Laux, "Monte carlo study of electron transport in silicon inversion layers," *Phys. Rev. B*, vol. 48, pp. 2244–2274, Jul 1993.

[3] J. A. López-Villanueva, F. Gámiz, I. Melchor, and J. A. Jiménez-Tejada, "Density of states of a two-dimensional electron gas including nonparabolicity," *Journal of Applied Physics*, vol. 75, no. 8, p. 4267–4269, Apr. 1994.

[4] C. Jungemann, A. Emunds, and W. Engl, "Simulation of linear and nonlinear electron transport in homogeneous silicon inversion layers," *Solid-State Electronics*, vol. 36, no. 11, pp. 1529–1540, Nov 1993.

[5] S. Jin, M. V. Fischetti, and T.-w. Tang, "Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity," *Journal of Applied Physics*, vol. 102, no. 8, p. 083715, Oct 2007.

[6] L. Donetti, C. Sampedro, F. Ruiz, A. Godoy, and F. Gamiz, "Multi-Subband Ensemble Monte Carlo simulations of scaled GAA MOSFETs," *Solid-State Electronics*, vol. 143, pp. 49–55, May 2018.

[7] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic, "Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches," *Journal of Applied Physics*, vol. 92, no. 7, p. 3730–3739, Oct. 2002.

[8] M. Fischetti, "Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. I. Homogeneous transport," *IEEE Transactions on Electron Devices*, vol. 38, no. 3, p. 634–649, Mar. 1991.