Improved effective mass theory for silicon nanostructures

K. Nehari, M. Lannoo, F. Michelini, N. Cavassilas, M. Bescond,[a] and J. L. Autran
IM2NP, UMR CNRS 6242, Bât. IRPHE, 49 rue Joliot Curie, BP 146, 13384 Marseille Cedex 13, France

(Received 2 May 2008; accepted 12 August 2008; published online 4 September 2008)

Effective mass theory is known to fail for nanostructures of indirect band gap materials such as silicon. We show that this situation can be remedied by going beyond the conventional second order expansion in the wave vector $\boldsymbol{k}$. The method including fourth order terms is developed for the conduction band of silicon and applied to the ⟨100⟩-oriented wells, wires, and dots. The energy minima, their shift in k-space, as well as the variation in the effective mass with size are fully predicted from analytical equations containing only bulk parameters. This approach opens the way for accurate simulations of realistic quantum devices avoiding heavy calculations. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2978196]

With the recent advances in semiconductor technology, it has become possible to fabricate solid state devices integrating quantum nanostructures generally based on indirect gap semiconductors such as silicon. Further advances require theoretical tools that are at the same time simple and accurate to provide with reasonable means quantitative predictions for realistic structures. This is not yet the case of *ab initio* calculations which can only be applied to limited size systems.[1] One possible way of improving this situation is to use semiempirical techniques such as tight-binding,[2] pseudopotential,[3] or multiband $\boldsymbol{k}\cdot\boldsymbol{p}$ (Ref. 4) methods, which can handle much larger sizes. However even these techniques are not easily applicable to complex problems such as a full self-consistent three-dimensional calculation of the drain current.[5] The aim of this paper is thus to rediscuss the applicability of the much simpler single band effective mass approximation (EMA). This has proved to be extremely powerful for most direct gap semiconductors.[6] However it was shown to fail for indirect band gap semiconductors particularly for Si, the most interesting from application point of view.[7] A critical example is provided by the case of a Si nanowire. Figure 1 shows its band structure obtained from the tight-binding calculations of Nehari *et al.*[2] The lower subbands are parabolic, but it has been shown[8] that the corresponding effective mass and the position of the indirect minimum are both strongly dependent on the size of the wire cross section. Furthermore the energy shift of the minima is not correctly predicted by standard EMA. Both effects have been tentatively ascribed to the importance of nonparabolic terms but without any quantitative proof.[9,10]

In this paper, we demonstrate how the EMA failures for Si can be fully remedied by including third and fourth order terms in the expansion of the bulk conduction band in powers of the wave vector $\boldsymbol{k}$. This allows us to derive simple analytical equations that accurately describe the electronic properties in Si quantum nanostructures. Comparison with full numerical tight-binding calculations shows quantitative agreement for Si quantum wells, wires, and dots.

We consider in the following a higher order EMA (HEMA) for the conduction band of Si. Our treatment is applied to three different situations: a [100] quantum well, a [100] wire with ⟨010⟩ and ⟨001⟩ faces, and a parallepipedic quantum dot with ⟨100⟩ faces. The boundary conditions correspond to an infinite potential barrier. We consider a minimum of the bulk conduction band taken to be in the k_x direction. To take into account the nonparabolic terms, we write a fourth order expansion of the bulk conduction band around the minimum k_0,

$$\varepsilon(k) = \varepsilon_0 + a(k_x - k_0)^2 + b(k_x - k_0) + \alpha_2(k_x - k_0)^3 + \beta_3(k_x - k_0)^4 + c, \tag{1}$$

with

$$a = \frac{\hbar^2}{2m_l^*} + \beta_1(k_y^2 + k_z^2),$$

$$b = \alpha_1(k_y^2 + k_z^2),$$

$$c = \frac{\hbar^2}{2m_t^*}(k_y^2 + k_z^2) + \beta_2 k_y^2 k_z^2 + \beta_4(k_y^4 + k_z^4), \tag{2}$$

where m_l^* and m_t^* are the bulk longitudinal and transverse effective masses. The coefficients α_i and β_j in Eqs. (1) and (2) are determined from a fit to the bulk conduction band. To get an analytical solution, we first drop the fourth order term in Eq. (1) and search for the shift δk_0 of the minimum induced by the nonzero values of k_y and k_z. The solution giving the right analytical behavior in the limit $a_2 \to 0$ is

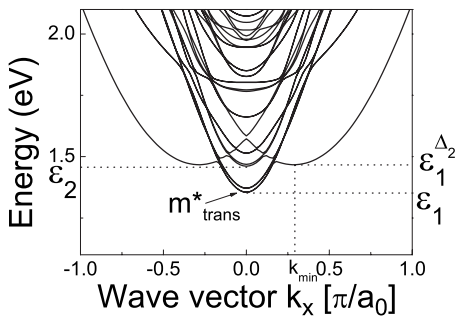

FIG. 1. Conduction band for a [100] infinite Si nanowire with main caracteristics. The transverse dimension of the squared cross-sectional nanowire is $D^* = 2.72$ nm.

[a]Electronic mail: marc.bescond@im2np.fr.

$$\delta k_0 = \frac{a}{3\alpha_2}\left(1 - \sqrt{1 - \frac{3b\alpha_2}{a^2}}\right). \tag{3}$$

From Eqs. (2) and (3), we note that δk_0 is at least of second order in k_y and k_z. We now rewrite expression (1) again to the fourth order with respect to the new minimum $k_0' = k_0 + \delta k_0$, where ε_0 is taken as the origin of energy,

$$\varepsilon(k_x) = (a\delta k_0^2 + b\delta k_0 + c) + (a + 3\alpha_2\delta k_0)(k_x - k_0')^2 + \alpha_2(k_x - k_0')^3 + \beta_3(k_x - k_0')^4. \tag{4}$$

At this step, in any [100] nanostructure, there are two possibilities: either k_x is a direction of propagation or it is a direction in which the system is confined.

If k_x is a direction of propagation: Then the previous treatment shows that the minimum of the corresponding band is shifted by δk_0 given by Eq. (3). From Eq. (4) the minima in energy and effective mass along this direction are given by

$$\varepsilon_{\min} = a\delta k_0^2 + b\delta k_0 + c,$$

$$\frac{\hbar^2}{2m^*} = a + 3\alpha_2\delta k_0. \tag{5}$$

Note that in these expressions the values of k_y and k_z depend on the case considered (quantum well or quantum wire). For the infinite quantum well, if k_y is the other direction of propagation, then we have $k_z = m\pi/D$, where m is a strictly positive integer and D is the width of the quantum well. For the quantum wire, $k_y = n\pi/D$ and $k_z = m\pi/D$, where n and m are both strictly positive integers. Note that this situation (k_x is a propagation direction) does not correspond to the lower subbands because of the values of $m_{l,t}$.

If k_x is a direction of confinement: Here we have to start from expression (4) in which we apply the change $k_x - k_0' \rightarrow -i\partial/\partial x$, which leads to the differential equation

$$\left\{-(a + 3\alpha_2\delta k_0)\frac{\partial^2\phi_x}{\partial x^2} + i\alpha_2\frac{\partial^3\phi_x}{\partial x^3} + \beta_3\frac{\partial^4\phi_x}{\partial x^4}\right\} = (\varepsilon - \varepsilon_{\min})\phi_x. \tag{6}$$

Keeping only the first term in Eq. (6) leads to the equation

$$-(a + 3\alpha_2\delta k_0)\frac{\partial^2\phi_x}{\partial x^2} = (\varepsilon - \varepsilon_{\min})\phi_x, \tag{7}$$

which gives

$$\phi_x = A\sin\left(\frac{l\pi x}{D}\right),$$

$$\varepsilon_l = \varepsilon_{\min} + (a + 3\alpha_2\delta k_0)\left(\frac{l\pi}{D}\right)^2, \tag{8}$$

where l is a strictly positive integer.

Now we treat the two corrective terms in Eq. (6). The term in $\partial^4/\partial x^4$ is diagonal in the basis $\{\phi_l\}$ and only gives an additional term to the energy equal to $\beta_3(l\pi/D)^4$. The term in $\partial^3/\partial x^3$ is nondiagonal and thus gives nothing in first order perturbation theory. We have thus to go to second order perturbation theory to get a contribution $\delta\varepsilon_1$. Finally, the energy is

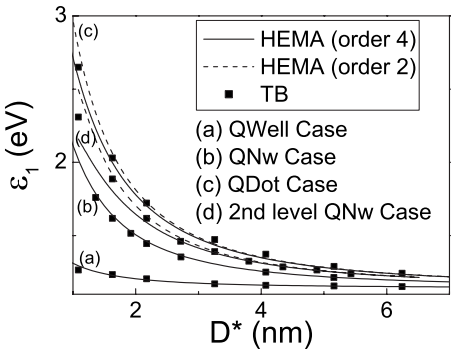

FIG. 2. First energy level predicted by the analytical model [Eq. (9)] in an infinite Si well (a), nanowire (b), and a Si quantum dot (c) with terms of the second (dotted line) and fourth (solid line) order compared with the tight-binding results in Ref. 11. Curve (d) shows the second energy level ε_2 in the nanowire (relevant for electron transport).

$$\varepsilon = \varepsilon_{\min} + (a + 3\alpha_2\delta k_0)\left(\frac{l\pi}{D}\right)^2 + \beta_3\left(\frac{l\pi}{D}\right)^4 + \delta\varepsilon_1, \tag{9}$$

where $\varepsilon_{\min}$ is given by Eq. (5).

It can be shown that for the ground state $l=1$, $\delta\varepsilon_1$ has the following expression:

$$\delta\varepsilon_1 = -1.333\frac{2m_l}{\hbar^2}\alpha_2^2\left(\frac{\pi}{D}\right)^4. \tag{10}$$

If y, z is a direction of propagation, we get the corresponding effective mass m^*_{trans} from the second order derivative of Eq. (9) with respect to $k_{y,z}$.

The explicit analytical HEMA expressions correspond to those derived above which, according to the case, are expanded either to second or fourth order in π/D. Its numerical values are compared in the following to full tight-binding results.[11] All the parameters discussed before are deduced from the bulk band structure of the same reference. This leads to α_1 and α_2 respectively equal to -3.46×10^{-48} and -8.47×10^{-49} J m^3, while β_1, β_2, and β_3 are given by -5.9×10^{-59}, -3.98×10^{-57}, and 1.12×10^{-58} J m^4.

In the following figures, D^* is taken to be the distance between the two limiting ⟨100⟩ Si planes. Our HEMA calculations correspond to vanishing boundary conditions and the width of the corresponding infinite potential well does not correspond to D^*. In fact, it is well known that, for example, in the simple case of tight-binding linear chain the vanishing of the wave function has to be taken on the first missing atom on each side (see Ref. 1, p. 55). Here we deal with the ⟨100⟩ plane distant from $a_0/4$ (where a_0 is the lattice parameter), the last two Si planes being connected to hydrogen atoms. These, in fact, play the role of extra Si atoms but with no dangling bonds. In the same spirit the vanishing of the wave function should then occur on the first missing planes. This means that our HEMA potential well should have a width D equal to $D^* + a_0$ for comparison with the results of tight binding.

We consider in the following the HEMA applied to Si conduction band in three different situations: (a) a [100] quantum well, (b) a [100] wire with ⟨010⟩ and ⟨001⟩ faces ($D_y = D_z = D$), and (c) a parallepipedic quantum dot with ⟨100⟩ faces. ε_1, m^*_{trans}, and $k_{\min}$ values (defined in Fig. 1) obtained for various dimensions by our model are compared to the full-band tight-binding results. Figure 2 shows the

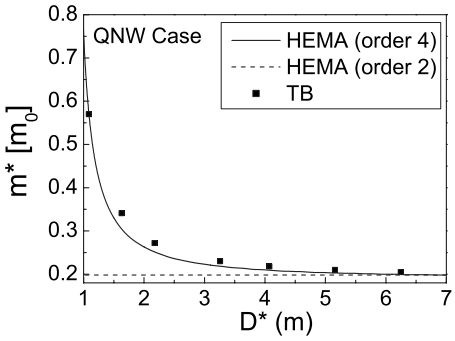

FIG. 3. Transport effective mass m^*_{trans} of Eq. (13) predicted by the analytical model in an infinite Si nanowire with term of the fourth order (solid line) compared with the tight-binding results in Ref. 11.

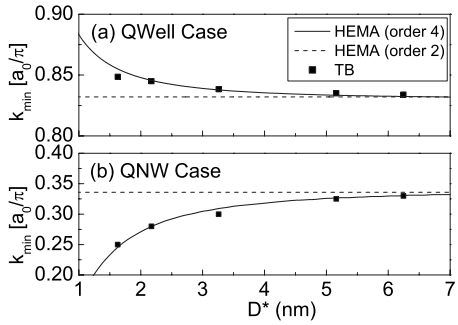

FIG. 4. Position of the minimum of the band k_{min} predicted by the analytical model [Eq. (3)] in an infinite Si nanowire and well with term of the fourth order (solid line) compared with the tight-binding results in Ref. 11 (square).

minimum of first subband ε_1 as well as the second level ε_2 in the case of nanowire. We show that HEMA and EMA (HEMA of the order of 2) curves are superposed in cases A and B, illustrating the fact that the "usual" effective mass approach remains very accurate to determine the energy minima. In the case of quantum dot, HEMA is more accurate than EMA. In fact, as expected, in this case higher order terms are more important than for one-dimensional (1D) and two-dimensional (2D) situations. Figure 3 shows a very good agreement between the analytical model and tight-binding results for the effective mass in wire, even for a transverse dimension down to 1.08 nm. Note that we did not show the variation in the effective mass in the quantum well because it is not significant. In Fig. 4, the position of k_{min} (defined in Fig. 1) is shown for various transverse dimensions. Once again the agreement between the HEMA model and tight-binding results is very satisfactory. In this figure, we can note that this position decreases with transverse dimensions in the case of nanowire.

We have shown that a single band effective mass Hamiltonian including third and fourth order terms in $\boldsymbol{k}$ components allows one to quantitatively describe the lower conduction subbands of the ⟨100⟩-oriented Si nanostructures. The use of perturbation theory leads to an analytical model based on well defined bulk parameters. This approach fully explains the origin of the minimum position for 1D and 2D structures as well as the evolution of the effective mass in Si wires. As the EMA in direct gap semiconductors, the presented HEMA would reveal itself to be extremely powerful for silicon based nanodevice investigations. We have to mention that our approach is analytical only for the ⟨100⟩-orientations for which the effective mass tensor is diagonal. Its extension to other orientations could be made by the methods introduced in Ref. 12 for simple EMA but, as in this case, would require numerical calculations.

[1] C. Delerue and M. Lannoo, *Nanostructures: Theory and Modelling* (Springer, Berlin, 2004).
[2] K. Nehari, N. Cavassilas, F. Michelini, M. Bescond, J. L. Autran, and M. Lannoo, Appl. Phys. Lett. **90**, 132112 (2007).
[3] L.-W. Wang and A. Zunger, J. Phys. Chem. **98**, 2158 (1994).
[4] S. Richard, F. Aniel, and G. Fishman, Phys. Rev. B **72**, 245316 (2005).
[5] M. Luisier, A. Schenk, and W. Fichtner, Appl. Phys. Lett. **90**, 102103 (2007).
[6] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Editions de Physique, Paris, 1990).
[7] J. Wang, A. Rahman, A. Ghosh, G. Klimeck, and M. Lundstrom, IEEE Trans. Electron Devices **52**, 7 (2005).
[8] Y. Zheng, C. Rivas, R. Lake, K. Alam, T. B. Boykin, and G. Klimeck, IEEE Trans. Electron Devices **52**, 1097 (2005).
[9] A. Rahman, G. Klimeck, T. B. Boykin, M. Lundstrom, Tech. Dig. - Int. Electron Devices Meet. **2004**, 139.
[10] E. Gnani, S. Reggiani, A. Gnudi, P. Parruccini, R. Colle, M. Rudan, and G. Baccarani, IEEE Trans. Electron Devices **54**, 2243 (2007).
[11] K. Nehari, N. Cavassilas, J. L. Autran, M. Bescond, D. Munteanu, and M. Lannoo, Solid-State Electron. **50**, 716 (2006).
[12] M. Bescond, N. Cavassilas, and M. Lannoo, Nanotechnology **18**, 255201 (2007).