RC22711 (W0301-140) January 27, 2003
Physics

IBM Research Report

Six-band k*p Calculation of the Hole Mobility in Silicon Inversion Layers: Dependence on Surface-orientation, Strain, and Silicon Thickness

Massimo V. Fischetti, [+]hibin Ren, Paul M. Solomon, Min Yang, Kern Rim

IBM Research Division
Thomas J. Watson Research Center,
P. O. Box 218
Yorktown Heights, NY 10598

[+]IBM Semiconductor
Research and Development Center (SRDC)

Research Division
Almaden - Austin - Beijing - Delhi - Haifa - India - T. J. Watson - Tokyo - Zurich

Six-band k · p calculation of the hole mobility in silicon inversion layers: Dependence on surface orientation, strain, and silicon thickness

M. V. Fischetti, Z. Ren†, P. M. Solomon, M. Yang, and K. Rim
IBM Semiconductor Research and Development Center (SRDC)
IBM Research Division, Thomas J. Watson Research Center,
P. O. Box 218, Yorktown Heights, NY 10598, USA
† *Microelectronics Division, Hopewell Junction, NY 12533, USA*
(December 20, 2002)

A six-band $\mathbf{k}\cdot\mathbf{p}$ model has been used to study the mobility of holes in Si inversion layers for different crystal orientations, for both compressive or tensile strain applied to the channel, and for a varying thickness of the Si layer. Scattering assisted by phonons and surface roughness has been accounted for, also comparing a full anisotropic model to an approximated isotropic treatment of the matrix elements. Satisfactory qualitative (and in several cases also quantitative) agreement is found between experimental data and theoretical results for the density and temperature dependence of the mobility for (001) surfaces, as well as for the dependence of the mobility on surface orientation (for the (011) and (111) surfaces). Both compressive and tensile strain are found to enhance the mobility, while confinement effects result in a reduced hole mobility for a Si thickness ranging from 30 to 3 nm.

I. INTRODUCTION

Theoretical studies of the hole mobility in Si inversion layers have lagged substantially those dedicated to electron transport. Oberhuber and Vogl[1] speculate correctly that this may be attributed to the complicated nature of the valence bands – not amenable to a simple analytic description – and to the overwhelming importance of n-channel devices in the present Complementary Metal-Oxide-Semiconductor (CMOS) technology. Yet, hole transport in inversion layer constitutes an interesting subject on its own, both from a fundamental and from a technological viewpoint. The latter motivation is probably gaining importance, as devices are scaled towards their ultimate limit of (quasi)ballistic transport for which the performance disadvantage of p-channel devices may disappear[2].

The purpose of this paper is to reduce this lag, at least partially, by presenting results which are complementary to those of Oberhuber and Vogl[1]: The computationally expensive self-consistency between the Schrödinger and Poisson equations will be bypassed in order to analyze a much wider range of applications and account for scattering processes (with phonons and surface roughness) as rigorously as possible.

After a discussion of the procedure followed to calculate the subband structure in the triangular-well approximation (Sec. II), of the evaluation of the hole mobility (Sec. III A) and of the relaxation rates due to scattering with phonons and surface roughness (Sec. III B), results will be presented for relaxed Si with different surface orientations and direction of the applied field (Sec. IV B), for both compressive and tensile in-plane stress (Sec. IV C), and, finally, for decreasing thickness of the Si layer (Sec. IV D).

II. SUBBAND STRUCTURE

As mentioned above, the calculation of the subband structure of p-type inversion layers is complicated by the strong anisotropy and nonparabolicity of the heavy- and light-hole bands, caused by their energetic proximity. In Si the small value of the spin-orbit splitting at the symmetry point Γ causes additional complications. Thus, at least 6 bands should be considered. This is usually done in the framework of the $\mathbf{k}\cdot\mathbf{p}$ approximation[3–5], the recent work by Nakatsuji *et al.*[6], who employ self-consistent pseudopotentials, being a recent welcome exception. The resulting numerical complexity has forced drastic analytic approximations early on[7–9], while the first 6-six band self-consistent calculations are due to Bangert *et al.*[10,11], Landwehr[12], and Ohkawa and Uemura[13,14]. Extension of their approach to heterostructures has required a 'symmetrized' reformulation of this 6-band $\mathbf{k}\cdot\mathbf{p}$ approach in order to address the issue of the boundary conditions at hetero-interfaces. The resulting symmetrized $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian[15] (but see Ref. 16 for an alternative form) has been used in the context of strained-Si-on-$Si_{1-x}Ge_x$ structures[1,15,17,18] and quantum-dots[19]. Here, even when considering strained-Si on relaxed $Si_{1-x}Ge_x$, the presence of the SiGe substrate will be neglected, assuming it is sufficiently removed from the strained-Si/SiO_2 interface. Therefore, symmetrization issues can be ignored. Thus, following Refs. 10–14, we solve the wave equation:

$$[\,\widehat{\mathbf{H}}(\mathbf{K},k_z)+\mathbf{I}\,V(z)]\;\cdot\psi_{\mathbf{K}}(z)\;=\;E(\mathbf{K})\;\psi_{\mathbf{K}}(z)\;. \quad (1)$$

Here $\mathbf{I}$ is the 6×6 identity matrix, $\psi_{\mathbf{K}}(z)$ is a 6-vector (function also of the two-dimensional in-plane components $\mathbf{K}=(k_x,k_y)$ of the hole wavevector) subject to the boundary conditions $\psi_{\mathbf{K}}(0)=0$ and $\psi_{\mathbf{K}}(W)=0$ (where W is either the thickness of the Si layer or $W\to\infty$ for

'bulk' inversion layers), and the 6×6 Hamiltonian is given by:

$$\widehat{\mathbf{H}} \;=\; \widehat{\mathbf{H}}_{\mathbf{kp}} \;+\; \widehat{\mathbf{H}}_{strain} \;+\; \widehat{\mathbf{H}}_{so}\;, \qquad (2)$$

where the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian, $\widehat{\mathbf{H}}_{\mathbf{kp}}$, the spin-orbit Hamiltonian, $\widehat{\mathbf{H}}_{so}$, and the strain Hamiltonian, $\widehat{\mathbf{H}}_{strain}$, have been derived in Refs. 3–5 and 20, and their expressions are given in Ref. 21, 22 with the parameters given in Ref. 23. It must be understood that in the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian as given in Ref. 21, the appropriate rotations in $\mathbf{k}$-space must be performed when dealing with (001), (011) and (111) surfaces, while $k_z \rightarrow -id/dz$. Finally, for the case of strain on (001) surfaces, if a_{SiGe} is the lattice constant of the $Si_{1-x}Ge_x$ substrate and a_0 the lattice constant of (relaxed) Si, the only nonzero components of the strain tensor $\widehat{\mathbf{e}}$ are

$$e_{xx} \;=\; e_{yy} \;=\; \frac{a_{SiGe}}{a_0} \;-\; 1\;, \quad e_{zz} \;=\; -\frac{2c_{12}}{c_{11}}\, e_{xx}\;, \qquad (3)$$

where c_{11} and c_{12} are the Si elastic constants, also from Ref. 23.

Equation (1) is solved numerically on a z-mesh of N_z points in the interval $(0, z_{max})$ (where z_{max} is chosen to be either the finite thickness W of the Si layer or, for 'bulk' inversion layers, to be sufficiently large to approximate correctly the boundary condition $\psi_{\mathbf{K}}(z \rightarrow \infty) = 0$ with $\psi_{\mathbf{K}}(z_{max}) = 0$) $z_l = z_{min} + (l-1)\Delta$, yielding a $6N_z \times 6N_z$ eigenvalue problem of the tri-diagonal block form:

$$\begin{bmatrix} . & . & . & . & . & . & . \\ . & \widehat{\mathbf{D}}_- & \widehat{\mathbf{D}}_{l-1} & \widehat{\mathbf{D}}_+ & 0 & 0 & . \\ . & 0 & \widehat{\mathbf{D}}_- & \widehat{\mathbf{D}}_l & \widehat{\mathbf{D}}_+ & 0 & . \\ . & 0 & 0 & \widehat{\mathbf{D}}_- & \widehat{\mathbf{D}}_{l+1} & \widehat{\mathbf{D}}_+ & . \\ . & . & . & . & . & . & . \end{bmatrix} \begin{bmatrix} . \\ \psi_{l-1} \\ \psi_l \\ \psi_{l+1} \\ . \end{bmatrix}$$

$$= E(\mathbf{K}) \begin{bmatrix} . \\ \psi_{l-1} \\ \psi_l \\ \psi_{l+1} \\ . \end{bmatrix}\;, \qquad (4)$$

where each $\psi_l = \psi(z_l)$ is a 6-component column-vector $\psi_i(z_l)$, the index i running over the $\mathbf{k}\cdot\mathbf{p}$ basis, and $\widehat{\mathbf{D}}_l$, $\widehat{\mathbf{D}}_-$, and $\widehat{\mathbf{D}}_+ = \widehat{\mathbf{D}}_-^{\dagger}$ are 6×6 block-diagonal difference operators, functions of the in-plane wavevector $\mathbf{K}$. The 'band-structure' sign convention used in Eqns. (1) and (2), from Ref. 23, yields negative eigenvalues $E(\mathbf{K})$. Below the sign convention appropriate to analyze transport (positive kinetic energies $E(\mathbf{K})$) will be employed.

In principle, the potential $V(z)$ should result from three terms: An image-term, $V_{im}(z)$, which, in the limit of an infinitely thick Si layer, has the form $-e^2\tilde{\epsilon}/(16\pi z)$, where e is the magnitude of the electron charge, $\tilde{\epsilon} = (\epsilon_{Si} - \epsilon_{ox})/(\epsilon_{Si} + \epsilon_{ox})$, ϵ_{Si} being permittivity of Si, ϵ_{ox} of the SiO_2 insulator; an exchange and correlation potential, $V_{xc}(z)$ (Ref. 24); and, finally, the Hartree term, $V_H(z)$, solution of the self-consistent problem consisting of solving Eq. (1) (or its discretized form (4)) together with the Poisson equation

$$\frac{d^2}{dz^2} V_H(z) \;=\; -\frac{e^2}{\epsilon_{Si}}[\rho(z) - N_A(z)]\;, \qquad (5)$$

where $N_A(z)$ is the concentration of ionized acceptors in the depletion layer, and $\rho(z) = \sum_{\text{occupied } \mathbf{K}} \psi^{\dagger}_{\mathbf{K}}(z) \cdot \psi_{\mathbf{K}}(z)$ is the hole charge density. This last expression has been evaluated in the past either using an approximated subband-structure[25], or by expanding the vector-wavefunctions $\psi_{\mathbf{K}}(z)$ on a suitable basis of orthogonal[13] or Airy functions[10], or, recently, with a direct numerical evaluation[1]. In either case, one either oversimplifies the problem or must face a very expensive numerical effort. This is especially true when a very fine discretization of the two-dimensional $\mathbf{K}$-space is required to achieve a satisfactory evaluation of the hole mobility. As stated above, here emphasis is given to the width of the range of applications. Therefore, also motivated by its satisfactory validity in the context of n-type inversion layers at high substrate doping concentrations of current practical interest[26,27], the so-called 'triangular-well approximation' $V(z) \approx eF_s z$, where F_s is the surface field, will be embraced. One should keep in mind that the results presented below, while expected to be qualitatively correct, may be quantitatively inaccurate at either low or high hole sheet densities, $n_s = eF_s/\epsilon_{Si}$, depending on substrate doping.

The calculation of the hole mobility described in Sec. III below require the tabulation of the subband structure. This is accomplished following a procedure analogous to Ohkawa's and Uemura's[13]: Having selected a surface field, F_s, the energy E_0 of the ground-state subband is determined by solving Eq. (1) for $\mathbf{K} = \mathbf{0}$. A maximum energy E_{max} is selected in order to account correctly for the thermal occupation of the top-most subband at the highest temperature considered (443K). The range $[0, \pi]$ of the polar angle ϕ in the two-dimensional $\mathbf{K}$-space is divided into N_ϕ intervals[28] of width $\Delta\phi = \pi/(N_\phi - 1)$ and the energy range $[0, E_{max}]$ is divided into N_E intervals of width $\Delta E = (E_{max} - E_0)/(N_E - 1)$ and for each ϕ_m (with $m = 1, N_\phi$) and each $E_n = E_0 + \Delta E(n-1)$ (with $n = 1, N_E$), the corresponding wavevector $K_{m,n}$ is determined by inverting the $\mathbf{k}\cdot\mathbf{p}$ problem, Eq. (4). Because of the quadratic dependence in $\mathbf{K}$ of the total Hamiltonian $\widehat{\mathbf{H}}(\mathbf{K}, k_z)$, writing $\mathbf{K}$ as $(K\cos\phi, K\sin\phi)$, Eq. (1) can be recast in the form of an eigenvalue problem in K of rank twice as large as the rank of the original problem. Thus, setting $\phi = \phi_m$, $E = E_n$, and solving this new eigenvalue problem – discretized as in Eq. (4) – one obtains the desired 'equienergy lines' $K_{m,n} = K(E_n, \phi_m)$. Finally, the gradients $\nabla_{\mathbf{K}} E$ at the same points (E_n, ϕ_m) are also evaluated and stored.

Employing $N_z \approx 250$ points along the z-axis in the interval $[0, z_{max}]$, with $z_{max} = 4E_{max}/(eF_s)$ (or replacing z_{max} with the thickness of the Si layer, W,

when required), $N_E \approx 100$ energy points in the interval $[E_0, E_0 + E_{max}]$, with $E_{max} = 0.3$ eV, $N_\phi \approx 25$-to-45 angles in the interval $[0, \phi_{max}]$ (see footnote 28), the tabulation of the subband structure requires the solution of about one eigenvalue problem of rank 1500 (given by Eq. (1) and its discretized form Eq. (4)) to obtain the wavefunctions and subband energies and of an additional 2500-4500 eigenvalue problems of rank 3000 to obtain the equienergy lines $K_{m,n} = K(E_n, \phi_m)$. This information is stored in look-up tables which are used to evaluate the hole mobility. Despite the relative fine discretization employed (and the resulting significant size of these tables), some numerical noise is still noticeable in the results presented below.

Using these tabulated values, the numerical evaluation of the Fermi level, E_F for a given hole density n_s at temperature T is performed by using a combination of the bisection and Newton's methods to look for the root of the equation:

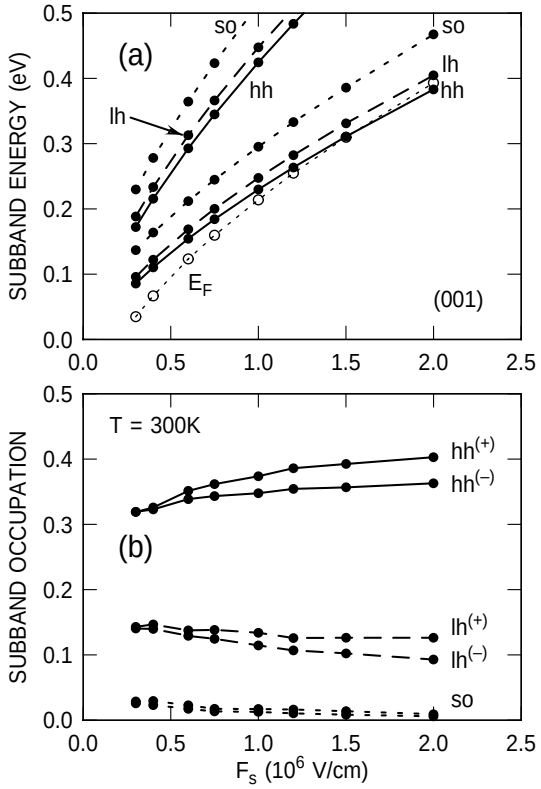

FIG. 1. Energy of the lowest-lying subbands (a) and their occupation at 300 K (b) for the (001) surface of relaxed Si. In (a) the subband energies are measured from the surface potential. The symbols are calculated values, the connecting lines only a guide to eye. Solid lines refer to the *hh*, heavy-hole-like subbands, the dashed line to the *lh*, light-hole-like subbands, while the dotted lines refer to the *so*, split-off-hole-like subbands. The thin dotted line shows the position of the Fermi energy at 300 K. Note in (b) that spin-degeneracy is lifted away from the center of the zone, so that the density of states and the occupation of given subband depends on spin variables.

$$n_s = \sum_\nu \int \frac{d\mathbf{K}}{(2\pi)^2} f^{(\nu)}(\mathbf{K}) = \sum_\nu \int_0^\infty dE \, \frac{\rho_\nu(E)}{1 + \exp\left(\frac{E + E_\nu^{(0)} - E_F}{k_B T}\right)} , \qquad (6)$$

where $f^{(\nu)}(\mathbf{K}) = \{1 + \exp[(E_\nu(\mathbf{K}) - E_F]/(k_B T)\}^{-1}$ is the Fermi-Dirac function in subband ν, $E_\nu^{(0)} = E_\nu(\mathbf{K} = \mathbf{0})$, and k_B is the Boltzmann constant. The density of states in subband ν at energy E is given by

$$\rho_\nu(E) = \theta[E - E_\nu^{(0)}] \, \frac{1}{(2\pi)^2} \int_0^{2\pi} d\phi \, \frac{K_\nu(E, \phi)}{\left|\frac{\partial E_\nu}{\partial K}\right|_{K_\nu(E,\phi)}} , \qquad (7)$$

where $\theta(x)$ is the step-function. The maximum number of subbands considered depends on surface field and lattice temperature. At a given temperature and surface field, the Fermi level is determined and the energy $E_{top} = E_\nu^{(0)}$ of the top-most subband with fractional occupation exceeding 10^{-4} is also determined. All subbands μ such that $E_\mu^{(0)} < E_{top} + \hbar\omega_{op} + 4k_B T$ are retained in order to account properly for absorption of optical phonons of energy $\hbar\omega_{op}$ for intervalley processes originating from the most energetic subband considered. A maximum of 36

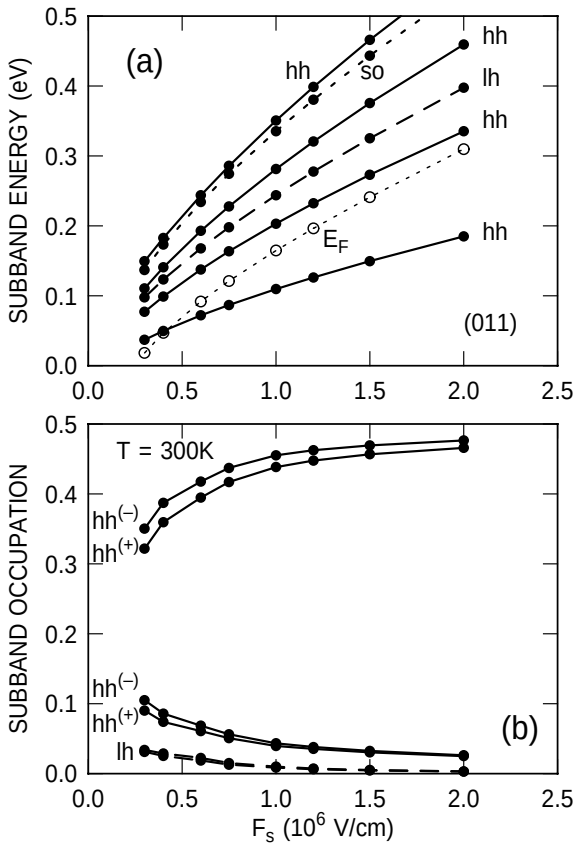

FIG. 2. Energy of the subbands (a) and their occupation at 300 K (b) for the (011) surface of relaxed Si, as in Fig. 1.

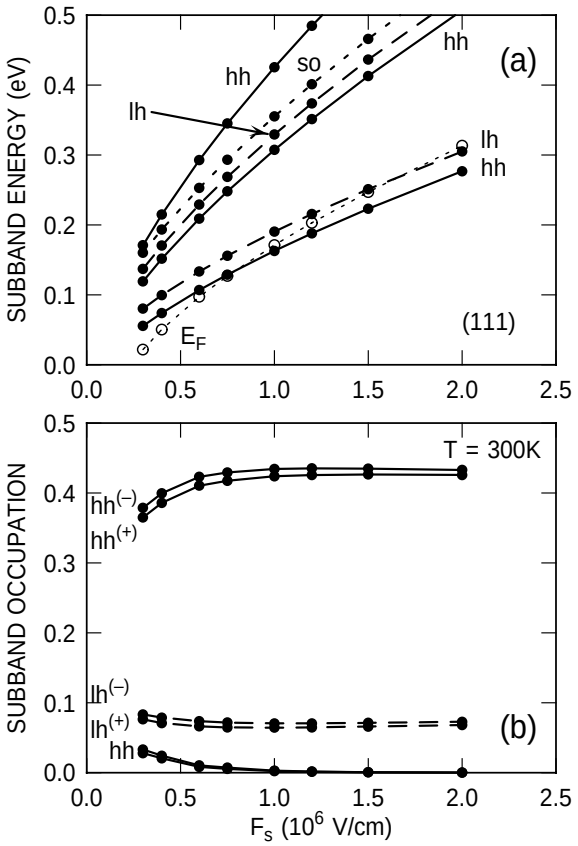

FIG. 3. Energy of the subbands (a) and their occupation at 300 K (b) for the (111) surface of relaxed Si, as in Fig. 1.

subbands has been used at $T = 443$ K and for $F_s = 3\times10^5$ V/cm. The number of subbands which must be included in the calculation grows very quickly as F_s decreases. This is what limits the magnitude of minimum confining field considered in this work.

It is useful to illustrate in some detail the subband energies and hole equienergy lines in a few cases of interest, since the behavior of the hole mobility shown in Sec. IV below can be understood only by considering the occupation of the subbands as a function of the confinement and orientation dependence of the conductivity mass. The top frames of Figs. 1-3 show the energies of the subbands (measured from the surface potential) as a function of the surface field F_s for three orientations of the Si surface. Note that the heavy-hole-like subband (here labeled *hh* for simplicity, while the label v_2 was used in Ref. 23) is always the ground state. The nature of each subband can be established by looking at the 'shape' of the equienergy lines in **K**-space and 'counting the nodes' of the wavefunctions. The first excited subband is a light-hole-like state (*lh*) for the (001) and (111) surfaces, it is a heavy-hole-like state for the (011) surface. This agrees with previous calculations for the (001) and (111) surfaces[13], but in the case of the (011) surface it disagrees with the calculations by Bangert *et al.*[10], who assign the first excited subband to a light-hole state. This is simply due to the 'early' $\mathbf{k}\cdot\mathbf{p}$ parameters used in Ref. 10, which reproduce only approximately the low-energy band-structure obtained from nonlocal empirical pseudopotentials[23] using form factors close to those determined recently[29,30]. The second excited subband is a split-off-hole state (*so*) for the (001) surface, a light-hole state for the (011) surface (also at odds with Ref. 10, as the nature of the first two excited subbands is reversed), a heavy-hole state for the (111) surface. Note in particular the high degree of degeneracy present even at high temperatures in the case of the (011) surface. As discussed below, this causes a

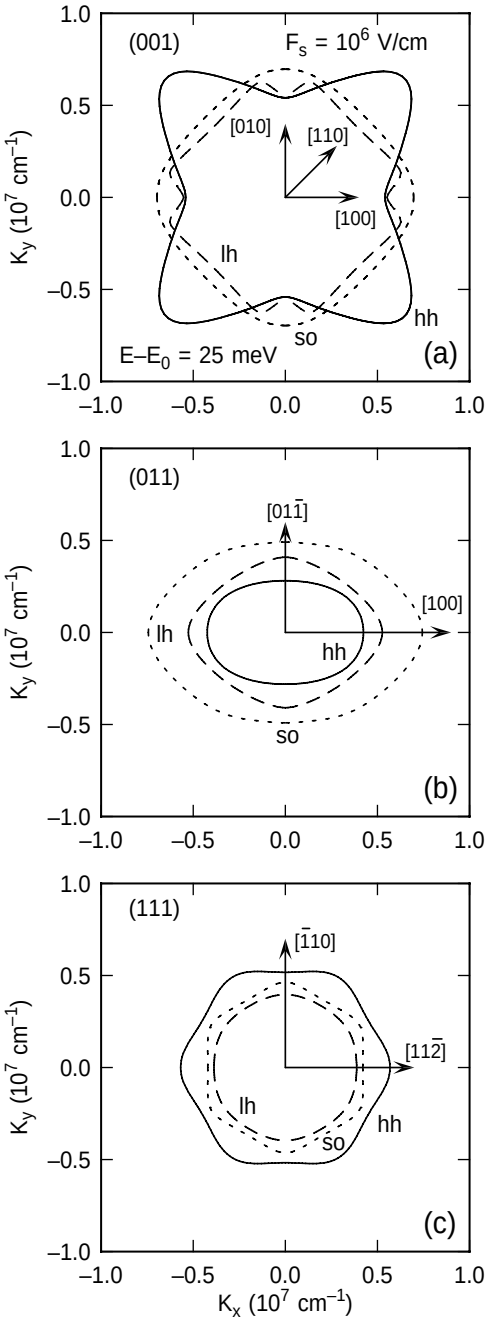

FIG. 4. Equienergy lines in the lowest-lying *hh*, *lh*, and *so* subbands for the (001) (a), (011) (b), and (111) (c) surfaces of relaxed Si. Only one spin state is plotted for clarity.

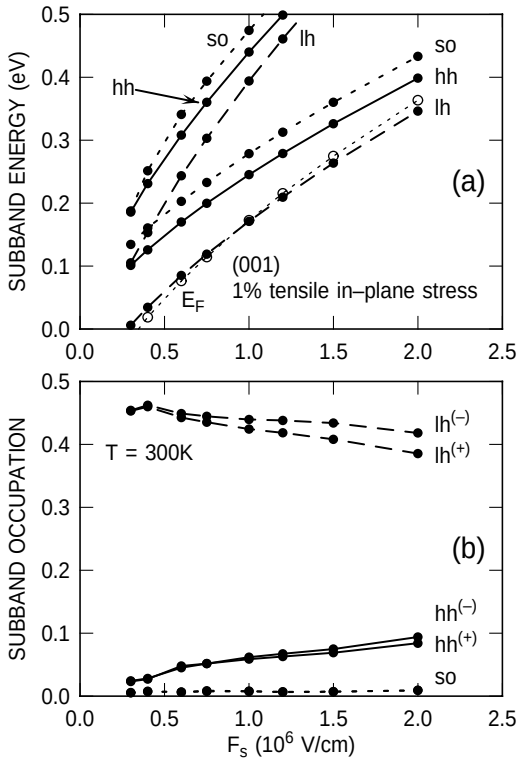

FIG. 5. Energy of the subbands (a) and their occupation at 300 K (b) for the (001) surface of Si with 1% of tensile in-plane stress applied on the (x, y)-plane (as for Si grown on $Si_{1-x}Ge_x$), as in Fig. 1.

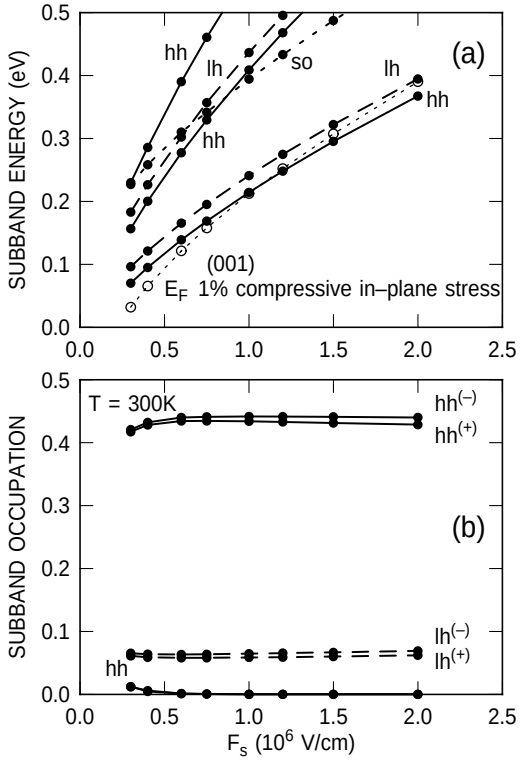

FIG. 6. As in Fig. 5, but for Si with 1% of compressive in-plane stress applied on the (x, y)-plane.

unique dependence on carrier density (especially at low temperatures) of the phonon-limited hole mobility. The bottom frames of these figures show the occupation of the subbands at 300 K. Since the spin-orbit interaction lifts the degeneracy of each band away from the Γ point, the density of states – and so the occupation – of the two subbands associated with each heavy-, light-, and split-off-band is different for different spin states. In the figures these spin states are arbitrarily labeled as $hh^{(\pm)}$, $lh^{(\pm)}$, and $so^{(\pm)}$. Figure 4 shows the equienergy lines at a kinetic energy of 25 meV above the subband bottom for the lowest-lying subbands of each type.

The same information is shown in Figs. 5, 6, and 7 for the (001) surface of Si subjected to 1% tensile or compressive strain in the (x, y)-plane. Note in Figs. 5(a) and 6(a) the crossing of subbands, since the effect of confinement is compensated by the effect of strain at small confining fields. Finally, Fig. 8, shows the variation of the subband energies as the thickness of the Si layer is reduced.

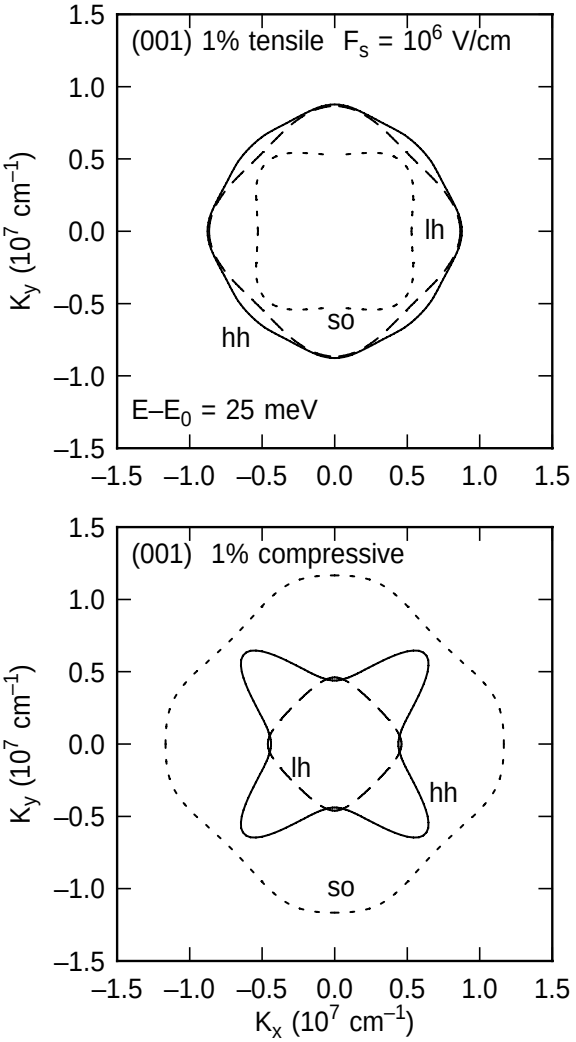

FIG. 7. As in Fig. 4, but for Si with 1% of tensile (a) or compressive (b) in-plane stress applied on the (x, y)-plane.

III. HOLE MOBILITY

A. Kubo-Greenwood formula

The hole mobility in inversion layer can be calculated from a linearization of the Boltzmann equation. Following the discussion in Refs. 23 and 31, the xx-component of the mobility tensor, μ_{ij}, can be written as:

$$\mu_{xx} = \frac{e}{4\hbar^2\pi^2 k_B T}\sum_\nu \frac{1}{n_\nu}\int_0^{2\pi} d\phi \int_{E_\nu^{(0)}}^{\infty} dE \frac{K_\nu(E,\phi)}{\left|\frac{\partial E_\nu}{\partial K}\right|_{K_\nu(E,\phi)}} \times \left(\frac{\partial E_\nu}{\partial K_x}\right)^2_{K_\nu(E,\phi)} \tau_x^{(\nu)}[K_\nu(E,\phi),\phi]\; f_0(E)[1-f_0(E)] \;, \quad (8)$$

where n_ν is the population of subband ν, $n_s = \sum_\mu n_\mu$, $\tau_x^{(\nu)}(K,\phi)$ is the (anisotropic) relaxation time for x-component of the momentum in subband ν, and $f_0(E) = \{1+\exp[(E-E_F)/(k_BT)]\}^{-1}$.

As discussed before[31], the major problem one has to confront originates from the fact that, even within first-order perturbation theory, the relaxation time can only be defined implicitly via the integral equation:

$$\frac{1}{\tau_x^{(\mu)}(\mathbf{K})} = \sum_\nu \int \frac{d\mathbf{K}'}{(2\pi)^2}\,|M_{\nu\mathbf{K}'\mu\mathbf{K}}|^2\,\delta[E_\mu(\mathbf{K}) - E_\nu(\mathbf{K}') \pm \hbar\omega] \times \left\{1 - \frac{v_x^{(\nu)}(\mathbf{K}')\tau_x^{(\nu)}(\mathbf{K}')f^{(\nu)}(\mathbf{K}')}{v_x^{(\mu)}(\mathbf{K})\tau_x^{(\mu)}(\mathbf{K})f^{(\nu)}(\mathbf{K})}\right\} \;, \quad (9)$$

where $v_x^{(\nu)}(\mathbf{K})$ is x-component of the hole group velocity at the point $\mathbf{K}$ and the matrix element $M_{\nu\mathbf{K}'\mu\mathbf{K}}$ associated to the scattering potential ϕ is given by:

$$M_{\nu\mathbf{K}'\mu\mathbf{K}} = \int \frac{d\mathbf{R}}{(2\pi)^2}\, e^{-i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{R}} \times \int_0^W dz\; \psi_{\mathbf{K}'}^{(\nu)\dagger}(z)\cdot\psi_{\mathbf{K}}^{(\mu)}(z)\;\phi(\mathbf{R},z) \;. \quad (10)$$

While self-consistent solutions of this problem may be sought for, it is convenient to simplify the problem. Note that for isotropic and elastic processes the term in curly brackets in the equation above reduces to unity, for isotropic bands and elastic (even if anisotropic) processes it simplifies to $= 1-\cos\theta_{\mathbf{K}',\mathbf{K}}$. By analogy, for general cases it is customary to approximate this troublesome term as:

$$\left\{1 - \frac{v_x^{(\nu)}(\mathbf{K}')f^{(\nu)}(\mathbf{K}')}{v_x^{(\mu)}(\mathbf{K})f^{(\nu)}(\mathbf{K})}\right\} \;, \quad (11)$$

assuming a sufficiently slow variation of $\tau_x^{(\mu)}(\mathbf{K})$ with $\mathbf{K}$. This converts the integral equation Eq. (9) to a conceptually simple evaluation of a (numerically demanding) multidimensional integral.

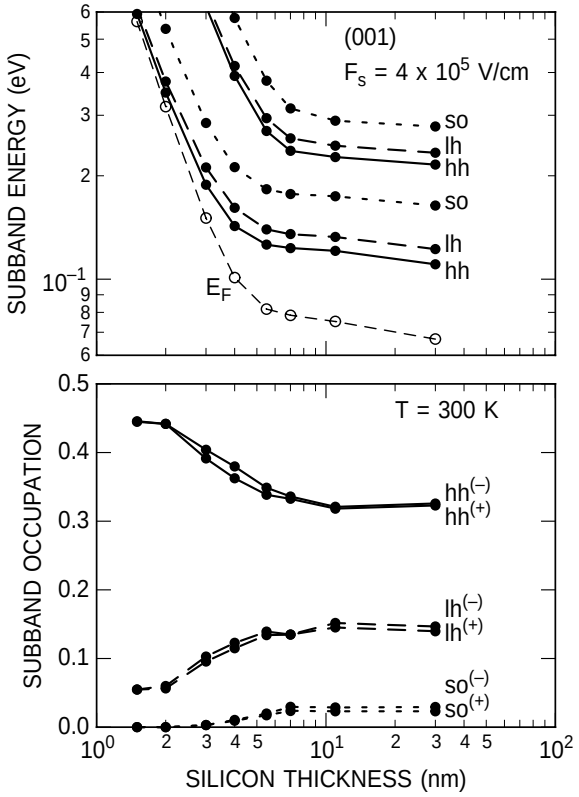

FIG. 8. Subband energies (a) and occupation (b) as in Fig. 1, but as a function the thickness of the Si layer.

B. Relaxation rates

The momentum relaxation rates caused by scattering with phonons and surface-roughness will be discussed in this section. Coulomb scattering with ionized impurities in the substrate or oxide/interface charge will be ignored.

Hole-phonon. The evaluation of the momentum relaxation time due to scattering with phonons is complicated not only by the anisotropy and nonparabolicity of the hole dispersion, but also by the anisotropy of the hole-phonon matrix elements. Usually, even sophisticated calculations in the context of inversion layers[1] or quantum wells[32,33] have relied on some isotropic approximation, similarly to what is done for bulk silicon[23,34–38]. In general, following Bir[20] in Hinckley's and Singh's reformulation[21,22], employing Fermi golden rule and the elastic, equipartition approximation (satisfactory at sufficiently high lattice temperatures), the scattering rate for emission or absorption of an acoustic phonon for a hole of wavevector $\mathbf{K} = (K\cos\beta, K\sin\beta)$ in subband μ is given by:

$$\frac{1}{\tau_x^{(\mu)}(K,\beta)} = \frac{k_BT}{\hbar\rho u_\eta^2}\sum_\nu \int_0^{2\pi}\frac{d\phi}{2\pi}\int_{-\infty}^{\infty}\frac{dq_z}{2\pi}\frac{K_\nu(E,\phi)}{\left|\frac{\partial E_\nu}{\partial K}\right|_{K_\nu(E,\phi)}}$$

$$\times \sum_{\alpha\beta} \left| \frac{q_\alpha}{q} e_\beta^{(\eta)}(\mathbf{q}) \sum_{i,j} D_{\alpha\beta}^{(ij)} \int_0^W dz\ \psi_{j\mathbf{K}'}^{(\nu)*}(z)\ e^{-iq_z z}\ \psi_{i\mathbf{K}}^{(\mu)}(z) \right|^2 \times \left[1 - \frac{v_x^{(\nu)}(\mathbf{K}') f^{(\nu)}(\mathbf{K}')}{v_x^{(\mu)}(\mathbf{K}) f^{(\mu)}(\mathbf{K})} \right] \theta[E_\mu(\mathbf{K}) - E_\nu^{(0)}] \ , \tag{12}$$

while for scattering with optical phonons (considering only one branch $\eta = op$ of frequency ω_{op})

$$\begin{aligned} \frac{1}{\tau_x^{(\mu)}(K,\beta)} &= \frac{3d_0^2}{2\rho a_0^2 \omega_{op}} \sum_\nu (n_{op} + 1/2 \pm 1/2) \\ &\times \int_0^{2\pi} \frac{d\phi}{2\pi} \int_{-\infty}^{\infty} \frac{dq_z}{2\pi} \frac{K_\nu(E,\phi)}{\left|\frac{\partial E_\nu}{\partial K}\right|_{K_\nu(E,\phi)}} \\ &\times \sum_\alpha \left| e_\alpha^{(op)}(\mathbf{q}) \sum_{i,j} A_\alpha^{(ij)} \int_0^W dz\ \psi_{j\mathbf{K}'}^{(\nu)*}(z)\ e^{-iq_z z}\ \psi_{i\mathbf{K}}^{(\mu)}(z) \right|^2 \\ &\times \left[1 - \frac{v_x^{(\nu)}(\mathbf{K}') f^{(\nu)}(\mathbf{K}')}{v_x^{(\mu)}(\mathbf{K}) f^{(\mu)}(\mathbf{K})} \right] \theta[E_\mu(\mathbf{K}) \pm \hbar\omega_{op} - E_\nu^{(0)}] \ , \end{aligned} \tag{13}$$

where $\mathbf{q} = (\mathbf{K} - \mathbf{K}', q_z)$ is the phonon wavevector, $\mathbf{K}' = K_\nu(E,\phi)(\cos\phi, \sin\phi)$ is the final hole wavevector, $\mathbf{D}_{\alpha\beta}$ and $\mathbf{A}_\alpha$ are 6×6 matrices (with the indices i, j running over x, y, z) describing the anisotropic deformation potential interaction between hole and phonons[21–23,38], $\mathbf{e}^{(\eta)}(\mathbf{q})$ is the polarization vector for phonons of branch η and wavevector $\mathbf{q}$, which is approximated by the expression given by Ehrenreich and Overhauser[39]. n_η is the Bose occupation factor, ω_η the frequency, and u_η the sound velocity of phonons of branch η, d_0 is the optical deformation potential, a_0 the Si lattice constant, and, finally, ρ the crystal density.

Surface roughness. It is probably fair to say that we are still unable to treat properly scattering with roughness at the Si/SiO_2 interface. The early formulation by Prange and Nee[40], Saitoh[41] and Ando[24,42] is still the best model available. Despite its semi-heuristic nature and the many approximations it relies upon, this model is often subject to additional approximations, mainly consisting in simplifying the treatment of dielectric screening[1,32,43] (almost invariably treated by replacing the dielectric matrix with a scalar[44]), by simplifying the matrix element[43], or even the wavefunctions[44]. When dealing with thin Si layers (as in silicon-on-insulator (SOI), double-gate (DG), or quantum-well (QW) structures), one has to deal with presence of the 'bottom' interface, problem studied early on by Gold[45] and Sakaki[46] in the context of Si and GaAs/AlAs quantum wells, respectively, and later by Gámiz *et al.*[47–49] in the context of n-type Si inversion layers and SOI structures. Recently, this problem has been reconsidered theoretically by Mou and Hong[50].

For reasons which will be briefly discussed shortly, following Eq. (23) below, here Ando's model will be followed extending it to account for the proximity of the bottom Si/SiO_2 interface, a distance W away from the 'top' Si/gate-insulator interface at $z = 0$. Accordingly, the momentum relaxation rate can be obtained from the first-order perturbation theory as:

$$\begin{aligned} &\frac{1}{\tau_x^{(\mu)}(K,\beta)} \\ &= \frac{2\pi}{\hbar} \sum_\nu \int \frac{d\mathbf{K}'}{4\pi^2} \left| V_{\nu\mathbf{K}'\mu\mathbf{K}}^{(SR)} \right|^2 \delta[E_\mu(\mathbf{K}) - E_\nu(\mathbf{K}')] \\ &\quad \times \left[1 - \frac{v_x^{(\nu)}(\mathbf{K}') f^{(\nu)}(\mathbf{K}')}{v_x^{(\mu)}(\mathbf{K}) f^{(\mu)}((\mathbf{K})} \right] \\ &= \frac{2\pi}{\hbar} \sum_\nu \frac{1}{4\pi^2} \int_0^{2\pi} d\phi\ \frac{K_\nu(E,\phi)}{\left|\frac{\partial E_\nu}{\partial K}\right|_{K_\nu(E,\phi)}} \left| V_{\nu\mathbf{K}'\mu\mathbf{K}}^{(SR)} \right|^2 \\ &\quad \times \left[1 - \frac{v_x^{(\nu)}(\mathbf{K}') f^{(\nu)}((\mathbf{K}')}{v_x^{(\mu)}(\mathbf{K}) f^{(\mu)}((\mathbf{K})} \right] \theta[E_\mu(\mathbf{K}) - E_\nu^{(0)}] \ , \end{aligned} \tag{14}$$

where the matrix element $V_{\nu\mathbf{K}'\mu\mathbf{K}}^{(SR)}$is given by[24]:

$$\begin{aligned} V_{\nu\mathbf{K}'\mu\mathbf{K}}^{(SR)} &= \int \frac{d\mathbf{R}}{2\pi}\ e^{-i(\mathbf{K}-\mathbf{K}')\cdot\mathbf{R}} \int_0^W dz \\ &\times \left\{ \psi_{\mathbf{K}'}^{(\nu)\dagger}[z - \Delta(\mathbf{R})] \cdot [H_0 + \Delta V(\mathbf{R}, z)]\ \psi_{\mathbf{K}}^{(\mu)}[z - \Delta(\mathbf{R})] \right. \\ &\quad \left. - \psi_{\mathbf{K}'}^{(\nu)\dagger}(z) \cdot H_0\ \psi_{\mathbf{K}}^{(\mu)}(z) \right\} \ . \end{aligned} \tag{15}$$

The roughness at the interface is usually described by the power spectrum, $|S(\mathbf{Q})|^2$, of the autocorrelation, $\Delta(\mathbf{R})$, of the roughness,

$$|\Delta(\mathbf{R})|^2 = \int \frac{d\mathbf{Q}}{2\pi}\ e^{-i\mathbf{Q}\cdot\mathbf{R}}\ |S(\mathbf{Q})|^2 \ . \tag{16}$$

Gaussian[24,40–42] and exponential[51] autocorrelation functions – characterized by a correlation length Λ and an average step-height Δ – have been assumed. Recently, 'intermediate' forms have been considered by Ishihara *et al.*[52] and by Leadley and coworkers[53]. Here we assume a form $|\Delta(\mathbf{R})|^2 \propto (2^{1/2}R/\Lambda)^n K_n(2^{1/2}R/\Lambda)$ (where $K_n(x)$ is a modified Bessel functions of the third kind), corresponding to a power spectrum

$$|S(\mathbf{Q})|^2 = \frac{\pi\Delta^2\Lambda^2}{(1 + Q^2\Lambda^2/2)^{1+n}} \ . \tag{17}$$

For $n = 1/2$ one recovers the exponential autocorrelation function[51], while here we follow the suggestion of Ref. 53 and choose $n = 2$.

According to Ando's model[24,41,42], the potential associated to the interfacial roughness arises from two contributions: The shift along the z-axis of the hole density

and the potential arising from the dipole moments induced by the 'steps' at the deformed interface(s). Thus, from Eq. (15), and employing the Green's function for Poisson equation appropriate to the geometry at hand, (see Eq. (32) below):

$$V^{(SR)}_{\nu\mathbf{K}'\mu\mathbf{K}} \;=\; S(\mathbf{K}-\mathbf{K}') \times \left\{ \Gamma_{\nu\mathbf{K}'\mu\mathbf{K}} \;+\; \frac{e^2}{\epsilon_s}\left[H^{(1)}_{\nu\mathbf{K}'\mu\mathbf{K}} \;+\; \tilde{\epsilon}\; H^{(2)}_{\nu\mathbf{K}'\mu\mathbf{K}} \right] \right\} . \quad (18)$$

In this expression the matrix element

$$\Gamma_{\nu\mathbf{K}'\mu\mathbf{K}} = \int_0^W dz \left\{ \psi^{(\nu)\dagger}_{\mathbf{K}'}(z)\cdot\psi^{(\mu)}_{\mathbf{K}}(z)\;\frac{dV(z)}{dz} \;+\; E^{(0)}_\nu \psi^{(\nu)\dagger}_{\mathbf{K}'}(z)\cdot\frac{d\psi^{(\mu)}_{\mathbf{K}}(z)}{dz} - E^{(0)}_\mu \frac{d\psi^{(\nu)\dagger}_{\mathbf{K}'}(z)}{dz}\cdot\psi^{(\mu)}_{\mathbf{K}}(z) \right\} , \quad (19)$$

results from the 'direct scattering at the steps'. As originally derived by Ando[24], it captures correctly *only* the effect of the confinement due to the field $(1/e)(dV(z)/dz)$, but fails to account for fluctuations of the additional confinement caused by the finite thickness of the layer, as already discussed by Gámiz *et al.*[47]. This stems from Eq. (15) which assumes a rigid shift of the wavefunction under the random translation $z \rightarrow z - \Delta(\mathbf{r})$. Therefore, when dealing with transport in thin Si layers, it is necessary to employ an alternative formulation[42], shown by Ando[24] to be equivalent to the original expression by Prange and Nee[40] in the bulk case. Within the context of the $\mathbf{k}\cdot\mathbf{p}$ approximation employed here, ignoring the dependence of the wavefunctions $\psi^{(\mu)}_{\mathbf{K}}$ on $\mathbf{K}$, and assuming that roughness affects only the 'top' Si/SiO_2 interface, this expression takes the form:

$$\Gamma^{(PN)}_{\nu\mathbf{K}'\mu\mathbf{K}} \;\approx\; \widehat{T}_{\mathbf{k}\cdot\mathbf{p}}\left(\frac{d\psi^{(\nu)}_{\mathbf{0}}(0)}{dz}, \frac{d\psi^{(\mu)}_{\mathbf{0}}(0)}{dz} \right) , \quad (20)$$

where, for (001) surfaces, $\widehat{T}_{\mathbf{k}\cdot\mathbf{p}}(\mathbf{a},\mathbf{b}) = -M(a^*_1 b_1 + a^*_2 b_2 + a^*_4 b_4 + a^*_5 b_5) - L(a^*_3 b_3 + a^*_6 b_6)$. The term

$$H^{(1)}_{\nu\mathbf{K}'\mu\mathbf{K}} \;=\; \int_0^W dz\; \psi^{(\nu)\dagger}_{\mathbf{K}'}(z)\cdot\psi^{(\mu)}_{\mathbf{K}}(z) \times \int_0^W dz'\; G_Q(z,z')\;\frac{\partial n_s(z')}{\partial z'} , \quad (21)$$

(where $G_Q(z,z')$ is given by Eq. (32) below) originates from the z-shift of the hole density, and the term

$$H^{(2)}_{\nu\mathbf{K}'\mu\mathbf{K}} \;=\; \int_0^W dz\; \psi^{(\nu)\dagger}_{\mathbf{K}'}(z)\cdot\psi^{(\mu)}_{\mathbf{K}}(z) \times \left\{ \frac{1}{1-\tilde{\epsilon}^2\; e^{-2QW}}(n_s+n_d)[\; e^{-Qz} - \tilde{\epsilon}\, e^{-Q(2W-z)}] + \frac{Q^2}{16\pi}\left[\sum_{n=0}^{\infty}\left(\frac{K_1[Q(2nW+z)]}{Q(2nW+z)} - \frac{\tilde{\epsilon}}{2}K_0[Q(2nW+z)] \right)\tilde{\epsilon}^{\;2n} - \sum_{n=1}^{\infty}\left(\frac{K_1[Q(2nW-z)]}{Q(2nW-z)} - \frac{\tilde{\epsilon}}{2}K_0[Q(2nW-z)] \right)\tilde{\epsilon}^{\;2n-1} \right] \right\} , \quad (22)$$

results from the potential associated to the interfacial dipoles. In these expressions n_d is the depletion charge (neglected here), $\tilde{\epsilon} = (\epsilon_s - \epsilon_{ox})/(\epsilon_s + \epsilon_{ox}) \approx 1/2$, $Q = |\mathbf{K}-\mathbf{K}'|$, and the sum over the index n – reflecting the potential due to all images located at $z' = \pm 2nW$ – results from having expanded the denominator of the Green's function, $(1-\tilde{\epsilon}^2 e^{-2QW})^{-1} = \sum_{n=0}^{\infty} \tilde{\epsilon}^{2n}\, e^{-2QnW}$ in order to perform the necessary integrations. Since scattering with surface roughness is important mainly at large carrier densities and the values of the Q-vectors of interest in calculating the mobility are of the order of the Fermi wavevector, K_F, it is sufficient to retain the leading terms in e^{-2QnW}, as $QW \approx K_F W >> 1$, and approximate this last expression as:

$$H^{(2)}_{\nu\mathbf{K}'\mu\mathbf{K}} \;\approx\; \int_0^W dz\; \psi^{(\nu)\dagger}_{\mathbf{K}'}(z)\cdot\psi^{(\mu)}_{\mathbf{K}}(z) \times \left\{ \frac{1}{1-\tilde{\epsilon}^2\; e^{-2QW}}(n_s+n_d)[\; e^{-Qz} - \tilde{\epsilon}\; e^{-Q(2W-z)}] + \frac{Q^2}{16\pi}\left[\frac{K_1(Qz)}{Qz} - \frac{\tilde{\epsilon}}{2}K_0(Qz) - \tilde{\epsilon}\left(\frac{K_1[Q(2W-z)]}{Q(2W-z)} - \frac{\tilde{\epsilon}}{2}K_0[Q(2W-z)] \right) \right] \right\} , \quad (23)$$

resulting now from the potential associated to the interfacial dipoles and their closest images.

Previous treatments of the effect of surface roughness in thin layers[45,46] have made use of an expression similar to Eq. (20), expressed in terms of the local energy-level fluctuations $\partial E^{(0)}_\nu/\partial W$, ignoring additional terms. Mou and Hong[50] have discussed the validity of employing the term $\partial E^{(0)}_\nu/\partial W$, and the importance of these additional terms, arising from the potential associated with the change of the carrier-density caused by the roughness. Note, however, that Ando's model already accounts for this effect in more general terms: Rather than looking directly at the change of carrier density $\partial n_s(\mathbf{R})/\partial W$ in terms of $\partial E^{(0)}_\nu/\partial W$, as done by Mou and Hong, the perturbation term $H^{(1)}_{\nu\mathbf{K}'\mu\mathbf{K}}$ given by Eq. (21) accounts for the same effect via the term $\partial n_s(z)/\partial z$, which reflects the shift of the carrier density with varying well-width. In the same limits considered by Mou and Hong, it can be shown that $H^{(1)}_{0\mathbf{K}'0\mathbf{K}} \sim e^2 S(K_F)/(\epsilon_s K_F^2 W^3)$ for $|\mathbf{K}-\mathbf{K}'| \sim K_F$, as long as the carrier sheet density is large enough so that $K_F W >> 1$. Thus, for small W, this term exhibits the same asymptotic behavior exhibited by the 'phenomenological' term involving local energy-level fluctuations employed in Refs. 45 and 46, or by the more rigorously-defined term of Mou and Hong[50] (the second term at the right-hand side of their Eqns. (2) or (3)).

Isotropic approximation. The evaluation of the relaxation rates is numerically a daunting task: Similarly to what is found in bulk Si, the anisotropy of the matrix elements in Eqns. (12) and (13) poses significant problems. In addition, in the case of the specific problem of holes in inversion layers, the dependence of the wavefunctions $\psi_{\mathbf{K}}(z)$ on the in-plane wavevector $\mathbf{K}$ requires the calculation of a large number of overlap integrals. Previous work[23,34–37] has shown that the hole mobility in bulk Si can be calculated quite accurately by approximating the anisotropic hole-phonon matrix element with appropriate angle-averaged quantities. In addition, it is seems plausible to ignore the $\mathbf{K}$-dependence of the wavefunctions and set $\psi_{\mathbf{K}}(z) \approx \psi_{\mathbf{0}}(z) = \psi(z)$. No argument can be given *a priori* supporting the validity of these approximations. Numerical results will be given below showing that relatively small errors are expected. Therefore, embracing these approximations and introducing the average acoustic deformation potential[23,34,36]:

$$\Xi_{eff}^2 \;=\; a^2 + \frac{c_l}{c_t}\left(b^2 + \frac{1}{2}d^2\right) \qquad (\Xi_{eff} \;\approx\; 7.12 \text{ eV}) \;, \tag{24}$$

where $c_l = (3c_{11} + 2c_{12} + 4c_{44})/5$ and $c_t = (c_{11} - c_{12} + 3c_{44})/5$, are average longitudinal and transverse elastic coefficients expressed in terms of the elastic constants c_{11}, c_{12}, and c_{44}, and the average optical deformation potential[23,35–37]:

$$(DK)_{op}^2 \;=\; \frac{c_l + 2c_t}{\rho \overline{v}_s{}^2}\left(\frac{d_0}{a_0}\right)^2 \qquad (\;(DK)_{op} \;\approx\; 13.24 \text{ eV}) \;, \tag{25}$$

where $\overline{v}_s{}^2 = (v_l^2 + 2v_t^2)/3$ is an average sound velocity, the relaxation times Eqns. (12) and (13) simplify, respectively, to the following expressions:

$$\frac{1}{\tau_x^{(\mu)}(K)} \;\approx\; \frac{2\pi\; k_B T\; \Xi_{eff}^2}{\hbar \rho u_l^2} \sum_{\nu} \;\mathcal{F}_{\mu\nu}\; \rho_\nu[E_\mu(\mathbf{K})] \;, \tag{26}$$

and

$$\frac{1}{\tau_x^{(\mu)}(K)} \;\approx\; \frac{\pi\; (DK)_{op}^2}{\rho \omega_{op}} \sum_{\nu} \mathcal{F}_{\mu\nu}\; \rho_\nu[E_\mu(\mathbf{K}) \mp \hbar\omega_{op}] \\ \frac{1 - f_0[E_\mu(\mathbf{K}) \mp \hbar\omega_{op}]}{1 - f_0[E_\mu(\mathbf{K})]} \left(n_{op} + \frac{1}{2} \pm \frac{1}{2}\right) \;, \tag{27}$$

where the 'form factor' $\mathcal{F}_{\mu\nu}$ is given by the 'usual' expression

$$\mathcal{F}_{\mu\nu} \;=\; \int_0^W dz\; |\psi_0^{(\mu)}(z) \cdot \psi_0^{(\nu)\dagger}(z)|^2 \;. \tag{28}$$

Similarly, the relaxation rate associated to scattering with interface roughness, Eq. (14), is approximated by:

$$\frac{1}{\tau_x^{(\mu)}(K)} \\ \approx\; \frac{2\pi}{\hbar} \sum_{\nu} \;\rho_\nu[E_\mu(\mathbf{K})] < |V^{(SR)}_{\nu\mathbf{K}'\mu\mathbf{K}}|^2 (1 - \cos\theta_{\mathbf{K}',\mathbf{K}}) > \;, \tag{29}$$

where the brackets $< ... >$ denote averaging over the polar angle ϕ, performed numerically.

Dielectric screening. Finally, dielectric screening is treated as described previously[26,31]: Assuming that the perturbing potential associated with the interfacial

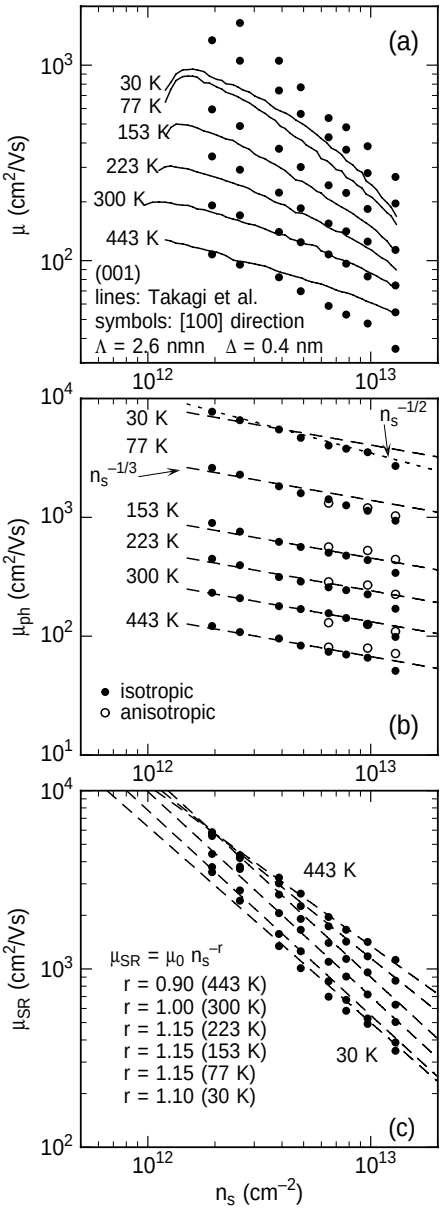

FIG. 9. (a): Calculated hole mobility for the (001) surface along the [100] direction as a function of hole sheet density at various lattice temperatures. The experimental data are by Takagi *et al.*[54]. (b): Calculated phonon-limited mobility in the isotropic approximation (circles) compared to the results obtained accounting for the full anisotropy of the hole-phonon matrix element (dots). The dashed and dotted lines show the approximate dependence of the phonon-limited mobility on carrier density. (c): Calculated mobility limited by surface roughness at various temperatures. The approximated power-law dependence is illustrated by the dashed lines.

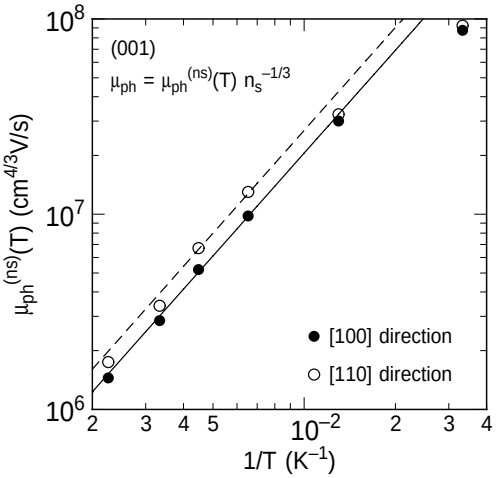

FIG. 10. Calculated temperature dependence of the phonon-limited mobility for the (001) surface and along both the [100] and [110] directions according to the approximated expression shown.

roughness is weak enough to leave the Bloch functions unaffected, denoting by $\varphi_{Q;\mu\mu}$ the matrix element of a general scattering potential, of which $V^{(SR)}_{\nu\mathbf{K}'\mu\mathbf{K}}$ is a particular example of interest here, the corresponding screened matrix elements, $\varphi^{(s)}_{Q;\mu\mu}$, are the solutions of the linear problem:

$$\varphi^{(s)}_{Q;\mu\mu} \;=\; \varphi_{Q;\mu\mu} \;-\; 2\sum_{\lambda}\; \beta_\lambda(Q)\; \mathcal{G}_{\mu\mu,\lambda\lambda}(Q)\; \varphi^{(s)}_{Q;\lambda\lambda}\;. \tag{30}$$

This simple formulation can be obtained assuming, as done above, that the wavefunctions $\psi_{\mathbf{K}}$ do not depend on $\mathbf{K}$, that the inter-subband polarization is weak enough to be neglected, and it is valid for static perturbations. In Eq. (30)

$$\mathcal{G}_{\mu\mu',\lambda\lambda'}(\mathbf{Q}) = \int\; dz\; \int\; dz'\; \psi_{\mathbf{0}}^{(\mu')\dagger}(z)\cdot\psi_{\mathbf{0}}^{(\mu)}(z)\; G_Q(z,z') \times\; \psi_{\mathbf{0}}^{(\lambda')\dagger}(z')\cdot\psi_{\mathbf{0}}^{(\lambda)}(z') \tag{31}$$

is the form-factor associated with the Green's function for the geometry of interest (*i.e.*, a Si layer from $z=0$ to $z=t$ sandwiched between two semi-infinite SiO_2 films):

$$G_Q(z,z') = \frac{1}{2Q(1-\tilde{\epsilon}^2\; e^{-2QW})}\left\{e^{-Q|z-z'|}+\tilde{\epsilon}e^{-Q|z+z'|} + \;\tilde{\epsilon}\; e^{-2QW}\left[e^{Q|z+z'|}+\tilde{\epsilon}\; e^{Q|z-z'|}\right]\right\}\;, \tag{32}$$

and

$$\beta_\lambda(Q) \;\approx\; \frac{e^2}{2\epsilon_{Si}}\frac{\partial n_\lambda}{\partial E_F}\; g_1(QL^{(\lambda)}_{th})\;, \tag{33}$$

is the static, wavelength dependent screening parameter. The function $g_1(x)$ is:

$$g_1(x) \;=\; \frac{2\pi^{1/2}}{x}\;\Phi\left(\frac{x}{4\pi^{1/2}}\right)\;, \tag{34}$$

where $\Phi(y)$ the 'plasma dispersion function'[26,31]. The thermal wavelength $L^{(\lambda)}_{th}$ in subband λ is approximated by an angular average $L^{(\lambda)}_{th} \approx 2\pi/K^{(\lambda)}_{th}$ with $K^{(\lambda)}_{th} \approx< K^{(\lambda)}(E^{(0}_\lambda + k_BT,\phi) >$.

As discussed in Ref. 26, scattering with phonons should be screened dynamically in a two-dimensional context. This results in very weak screening effects which shall be ignored altogether in this work. Finally, note that the numerical work required to evaluate the form-factors Eq. (31), to set-up the polarization matrix, and to solve the linear problem Eq. (30) is one of the heaviest computational step, surpassed only by the tabulation of the (sub)band structure outlined in Sec. II above.

IV. RESULTS

A. Relaxed Si, (001) surface, [100] direction

Before presenting and discussing results relative to strained or thin Si, it is necessary to present results for what may be viewed as a 'benchmark' case, namely, for the (001) surface of relaxed Si. Figure 9 shows the calculated hole mobility along the [100] direction. Comparison with the experimental data by Takagi and coworkers[54] (Fig. 9(a)) shows a satisfactory agreement at temperatures high enough for the elastic/equipartition approximation made to treat scattering with acoustic phonons

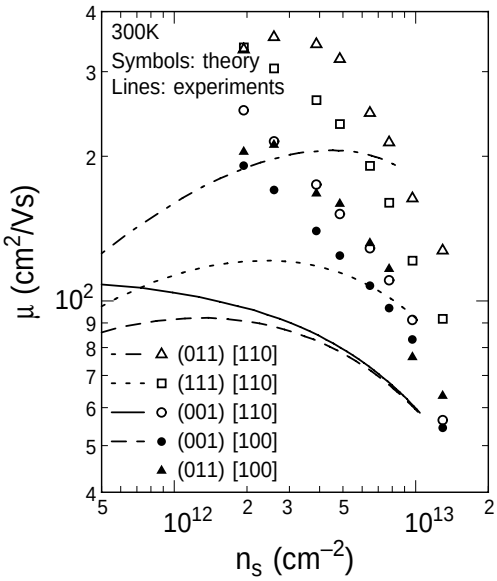

FIG. 11. Calculated total hole mobility at 300 K for the (001) surface along the [100] (dots) and [110] (circles) directions, for the (011) surface along the [100] (solid triangles) and [110] (open triangles) directions, and for the (111) surface along the [110] direction (open squares). Experimental data relative to the mobility along the [110] direction for the (001) (solid line), (011) (dot-dashed line), and (111) (dotted line) surfaces and along the [100] direction for the (001) (dashed line) are also shown.

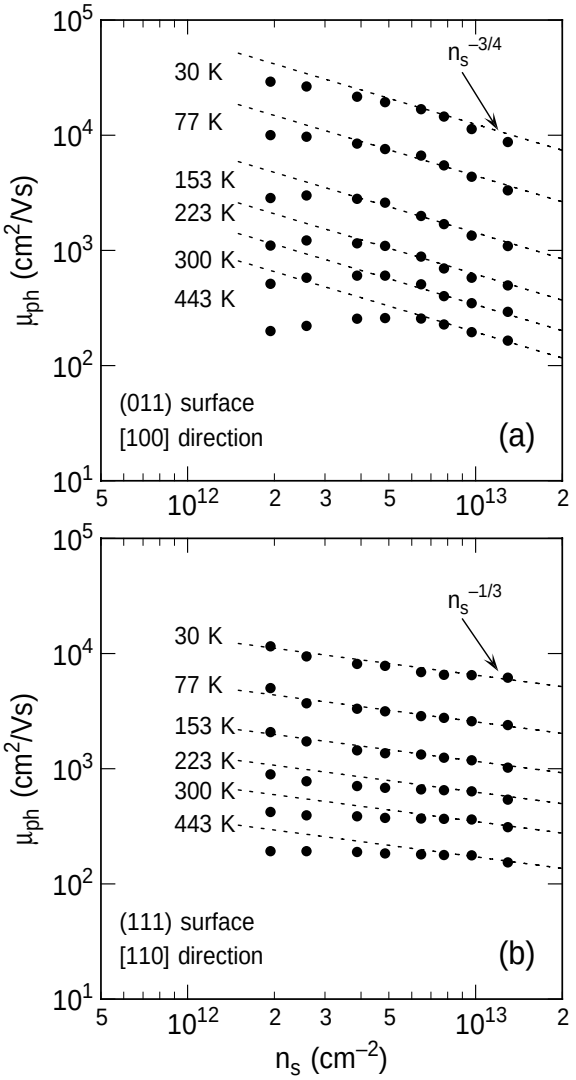

FIG. 12. Calculated phonon-limited mobility at various temperatures along the [110] direction for the (011) surface (a) and for the (111) surface (b).

to be valid. At low hole densities Coulomb scattering with dopants and/or interface charges, ignored here, may account for the difference between experimental data and theoretical results. At large densities, on the other side, the comparison is made difficult by the usual uncertainty surrounding the proper surface-roughness parameters Δ and Λ (here taken to be 0.4 and 2.6 nm, respectively). Furthermore, as illustrated in Fig. 9(b), the approximations made in simplifying Eqns. (12) and (13) to their isotropic forms, Eqns. (26) and (27), respectively, causes a slight underestimation of the phonon-limited mobility at large n_s. The 'anisotropic' results shown in Fig. 9(b) have been obtained by calculating the relaxation rates (12) and (13) with a quite laborious and time-consuming Monte Carlo sampling technique analogous to the one described in Ref. 23. In this figure note how the phonon-limited mobility, μ_{ph}, decreases with hole sheet density faster than the 'usual' $n_s^{-1/3}$ law. This results from the fact that the (doubly degenerate at $\mathbf{K} = \mathbf{0}$) lowest-lying *lh* subband exhibits a higher phonon-limited mobility because of its smaller conductivity mass (see Fig. 4(a)), but its population decreases with increasing confinement (see Fig. 1(b) and even more strongly at low T). Thus, a faster drop of μ_{ph} ($\sim n_s^{-1/2}$) results from the expected $n_s^{-1/3}$ drop augmented by a decreasing occupation of these 'fast' subbands. By fitting the calculated low-n_s behavior of μ_{ph} to an expression of the form $\mu_{ph}^{(n_s)}(T) n_s^{-1/3}$, Fig. 10 shows that the parameter $\mu_{ph}^{(n_s)}(T)$ exhibits the $T^{-1.75}$ dependence observed experimentally[54]. This is a nontrivial results, since the temperature dependence of the phonon-limited mobility depends strongly on the relative weight of scattering processes with acoustic and optical phonons. This is encouraging as it indicates that the values for Ξ_{eff} and $(DK)_{op}$ given from Ref. 23 and Eqns. (24) and (25) for bulk Si seem to apply also in *p*-type inversion layers. This consideration, as well as the results shown in Fig. 9(b) relative to the isotropic-anisotropic models, suggests that it is not unreasonable to assume the isotropic approximation.

An unanticipated result is shown in Fig. 9(c): The surface-roughness-limited mobility, μ_{SR}, exhibits a strong temperature dependence. This results from the contributions of two effects: Since the Fourier transform of the surface-roughness correlation function, $S(Q)$, decreases sharply with Q, in a non-degenerate situation at higher temperatures 'hotter' holes will have a shorter wavelength and will feel less the roughness. This results in an increasing mobility at higher temperatures. (The case of degenerate subbands, for which the Fermi level moves higher into the subband at lower temperatures, exhibits the opposite behavior, as discussed below for the (011) surface). Furthermore, at low temperatures the more 'confined' low-energy subbands are more populated and these exhibit a lower surface-roughness-limited mobility. The stronger dielectric screening at reduced temperatures helps in boosting μ_{SR} to grow at lower temperature, but not enough to overcome the former two effects. Note also how μ_{SR} decreases with increasing hole density as some small power of n_s, that is, as n_s^{-r}, with $r \sim 1$, quite a weak dependence when compared to the n_s^{-2}-to-$n_s^{-2.6}$ behavior observed for electrons[27,54]. This is caused by the stronger (with respect to electrons) screening properties of the 'slower', more massive holes. Note, however, as discussed in Ref. 27, that the particular choice of Λ affects this behavior. It should be remarked explicitly that neither the n_s-dependence of the phonon-limited mobility nor the temperature dependence of the surface-roughness limited mobility observed here for the (001) surface are general features: The case of (011) surface, in particular, exhibits an interestingly qualitatively different behavior.

Finally, we should note another unanticipated result: the dramatic failure of the well-accepted Matthiessen's rule, even more dramatic than in the case of *n*-type

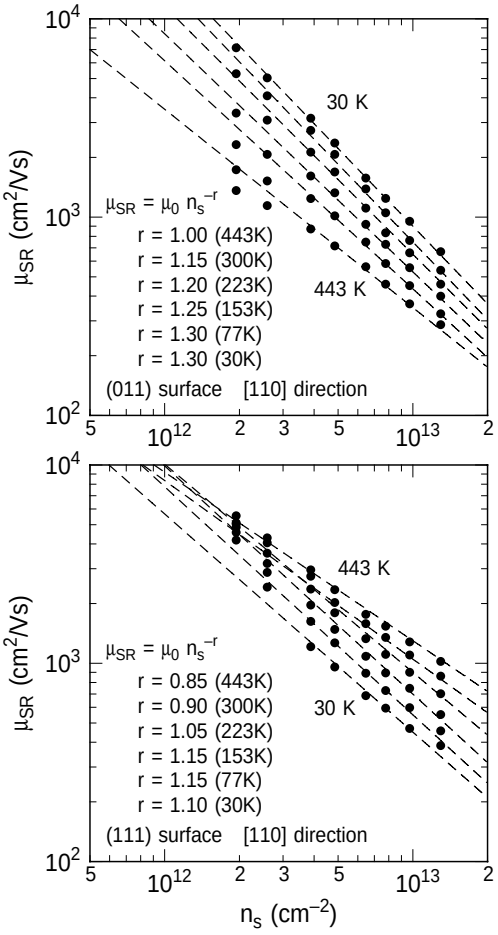

FIG. 13. Calculated surface-roughness-limited mobility at various temperatures along the [110] direction for the (011) surface (a) and for the (111) surface (b).

inversion layers[27]. Whenever the phonon- and surface-roughness-limited mobilities in each subband differ substantially (because, for example, of the competing effects of different effective masses, inter-subband processes, and a varying degree of confinement), Matthiessen's rule is approximately valid in each subband, but it overestimates dramatically the total mobility.

B. Dependence on crystal orientation

Figure 11 shows the calculated mobility for several surface orientations (and, for the (001) and (011) surfaces, also along two crystallographic directions) at 300 K. Experimental results are also shown (see Ref. 55 for experimental details). In general, the results of the calculations are in qualitative agreement with the experimental data reported here as well as those published previously[56–58] (Ref. 58 actually dealing with orientation dependence of the on-current of devices). Comparison with data obtained here is hampered by the large substrate doping (resulting in strong Coulomb scattering not accounted for theoretically) and by the (poorly-understood) role of heavily nitrided oxides, which usually depress the carrier mobility. Thus, only qualitative conclusions will be drawn here and in the following. Of particular interest are the following observations:

1. For the (001) surface, the calculated mobility along the [110] direction is somewhat larger than the mobility along the [100] direction. Indeed, despite the large mass of the hh-subband along the [110] direction, it is the much smaller conductivity mass of the lh-subband along this direction (see Fig. 4) which is responsible for this behavior. Experimentally, this behavior is not observed. Indeed, the difference between the experimental mobilities along the [100] and [110] directions shown in Fig. 11 is not significant, being smaller than sample-to-sample variations, while usually in the literature no difference is observed[56], and even the opposite behavior has been recently reported[59]. We can only speculate and blame slightly inaccurate $\mathbf{k}\cdot\mathbf{p}$-parameters resulting in an overestimation of the occupation of the lh-subband. Indeed, even a minor change of the relative occupations of the hh- and lh-subbands would affect significantly our result.
2. As shown in Fig. 12(a), at large hole sheet densities and for (011) surfaces, the phonon-limited mobility decreases much faster than $n_s^{-1/2}$, approaching an $n_s^{-3/4}$ behavior (shown by the dashed lines in the figure). At low n_s and high temperatures, on the other hand, it decreases with decreasing hole density. This behavior can be understood as follows: At small n_s, the small energetic separation between two lowest-lying hh bands (see Fig. 2) favors inter-subband transitions assisted by absorption of optical phonons, provided the lattice temperature is large enough to allow a significant phonon population. As the confinement increases, this process weakens and the phonon-limited mobility increases. However, recalling the high level of degeneracy already noted above, with the ground-state hh-subband being the only significantly occupied subband, as n_s increases, so does the Fermi energy. Carriers now feel a larger momentum relaxation rate, because of their higher energy and the strong nonparabolicity of the hh-subband. This behavior is peculiar to the (011) surface, because of its high degeneracy. Indeed the phonon-limited mobility for the (111) surface shows a very weak dependence on the hole density (see Fig. 12(b)), because of the close energetic proximity of the lowest-energy hh and lh subbands (see Fig. 3).
3. The temperature dependence of the surface-roughness-limited mobility for the (011) surface exhibits the opposite behavior seen for the (001) and (111) surfaces, namely, μ_{SR} increases with decreasing temperature, as illustrated in Fig. 13. Once more, the aforementioned high level of degeneracy is responsible for this behavior, since now the reduced value of the function $S(Q)$ at the larger

Fermi energies resulting from a reduces temperature and the enhanced dielectric screening dominate over the (very weak, in this case) subband-repopulation. The n_s-dependence is always of the form n_s^{-r} with $r \approx 1$ at high densities, growing more 'steeply' (roughly as n_s^{-2}) at lower confining fields.

C. Dependence on strain

Because of the promise of strained-Si in the present technology[60], several groups have investigated experimentally the enhancement of the hole mobility in Si grown on relaxed $Si_{1-x}Ge_x$ substrates[61–68]. Figure 14 shows the results of the calculation of the hole mobility at 300 K along the [100] and [110] directions for the (001) surface in strained Si layers with 1% of strain applied on the (x, y) plane, as in Ref. 23. Also the case of compressive strain – not yet investigated beyond the linear piezoresistive regime[70,69] – is shown for comparison. In agreement with the recent results by Nakatsuji and coworkers[6], a strong enhancement of the mobility is seen in the case of tensile strain. The fast drop of the mobility at large hole sheet densities is due to the fact that tensile strain enhances the *bulk* hole mobility more than compressive strain does, but the quantization effects caused by the confining potential tend to 'cancel' the effect of strain: Tensile strain, as shown in Fig. 1(b) of Ref. 23, increases the energetic separation between the *hh* and *lh* bands, while confinement tends to maintain a small energetic splitting between the *hh* and *lh* bands, while pushing the *so* (split-off) hole band to higher kinetic energies. Therefore, at large hole densities the hole mobility

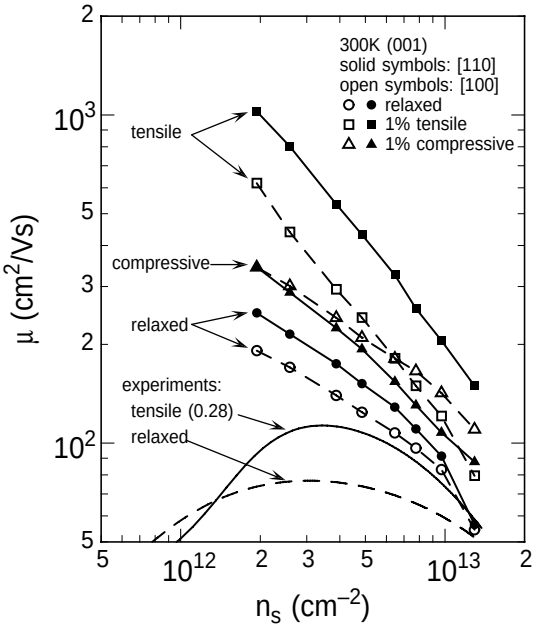

FIG. 14. Calculated total hole mobility at 300 K for the (001) surface along the [100] (open symbols) and [110] (solid symbols) directions, for relaxed Si (circles) and for Si under 1% in-plane strain, both compressive (triangles) and tensile (squares). Experimental data for Si-on-$Si_{0.72}Ge_{0.28}$ (solid line) and a relaxed 'control' device (dashed line) are also shown.

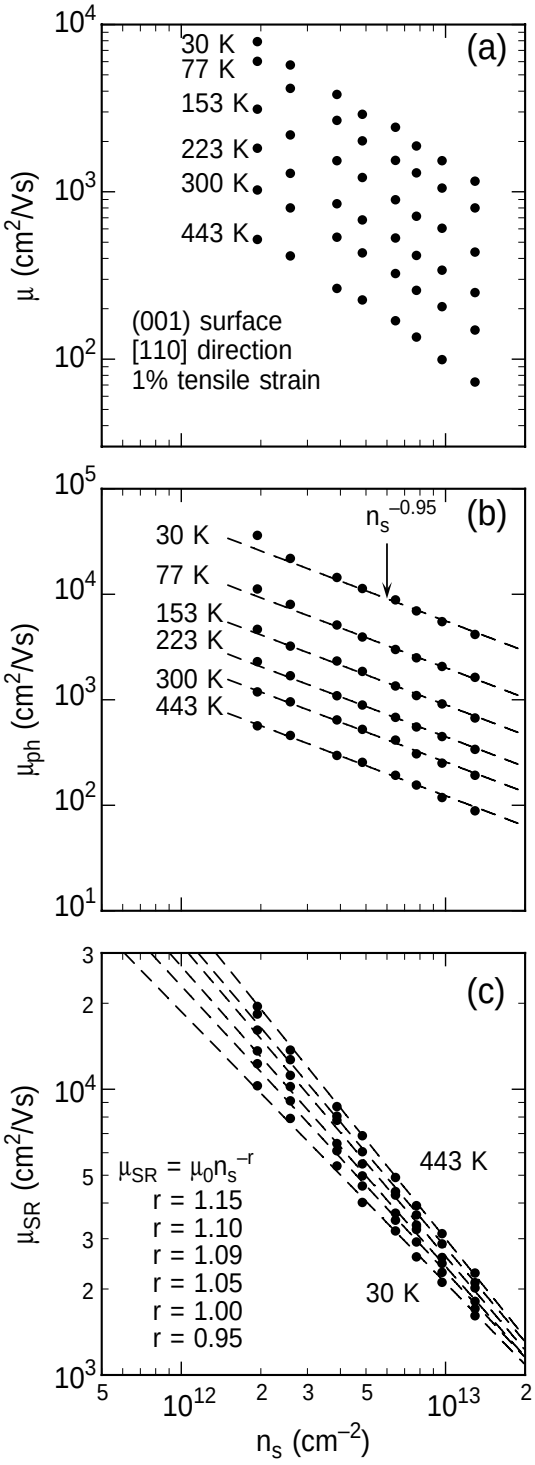

FIG. 15. Calculated total (a), phonon-limited (b), and surface-roughness-limited mobility (c) for Si under 1% tensile in-plane strain at various temperatures. The dashed lines in (b) and (c) illustrate qualitatively the approximate dependence on carrier density.

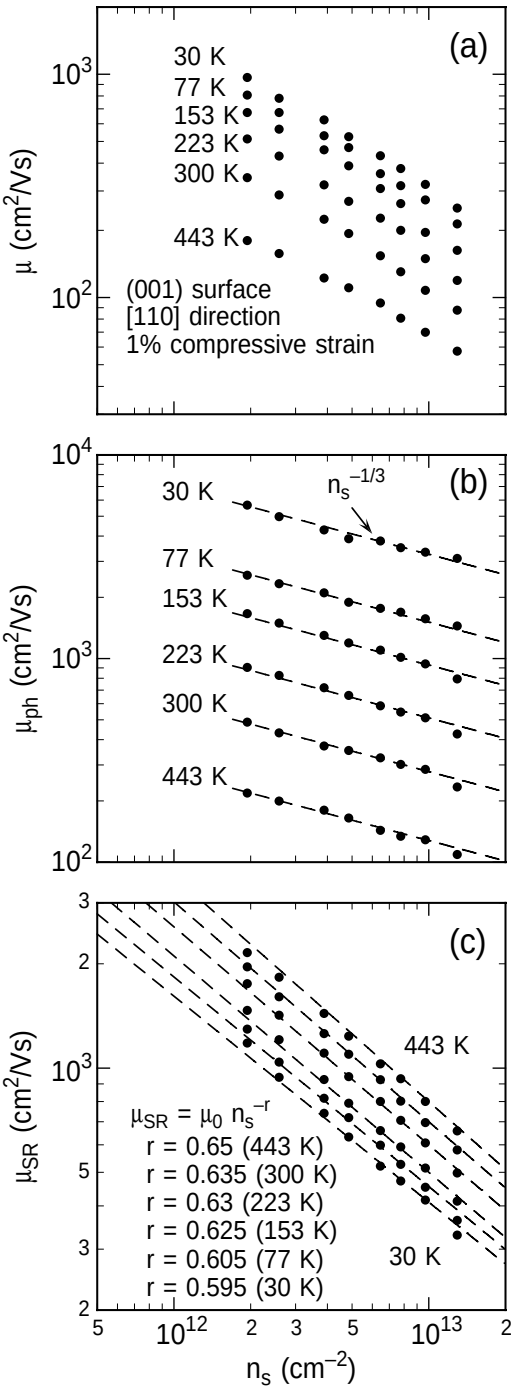

FIG. 16. As in Fig. 15, but for 1% compressive in-plane strain.

approaches the mobility of relaxed Si. The opposite happens when compressive strain is considered: Strain and confinement effects both work in maintaining a constant energetic splitting between the *hh* and *lh* bands and in pushing the *so* band to higher energies (see Fig. 6(a) which shows that the *lh* subband is actually the lowest-energy state), so the mobility enhancement caused by strain persists even at the largest hole density considered here. Eventually, as is the case for n-type inversion layers[27], at large-enough hole densities, only the effect of quantization remains and all curves in Fig. 14 expected to merge. Figures 15 and 16 show the temperature dependence of the total hole mobility (a), as well as of the separate phonon- and surface-roughness-limited mobilities, in (b) and (c) respectively. Note how strongly μ_{ph} decreases with increasing n_s (approximately as n_s^{-1}) in the case of tensile strain, as a result of this competition between stress and confinement.

A final remark is necessary: Here, as in Ref. 23, the application of stress of either sign (*i.e.*, tensile or compressive) is predicted to enhance the hole mobility. On the contrary, a linearization for small strain (in the so-called piezoresistive regime) predicts that the mobility should decrease for one type of strain, increase when the opposite stress is applied, the sign of the piezoresistance coefficients Π_{11} and Π_{12} determining the sign of the mobility-change. The main reason for this difference stems from the fact that for the small values of stresses which can be applied mechanically[7,8,69,71] (10 kPa to 200kPa, corresponding to a strain of the order of 0.01 to 0.1%) or induced by the field-oxidation[72,73], the mobility is affected mainly by (small, as they are approximately linear with strain) changes of the effective masses and the associated band/subband repopulation. On the contrary, at the much higher levels of strain present in strained-Si layers grown on SiGe substrates, band repopulation caused by large strain-induced energy shifts and suppression of inter-band scattering dominate the picture, as suggested by Tezuka *et al.*[61]. Large quantitative and even qualitative differences should be expected between the two regimes.

D. Dependence on Si thickness

Experiments performed on Si-on-insulator (SOI) structures have shown that the hole mobility decreases monotonically as the thickness of the Si layer is reduced[43,74,75] from $\approx$ 60 nm to $\approx$ 2.7 nm, behavior quite similar to what measured[74,75] and calculated[47,49,75,76] for electrons. For both n- and p-type channels the additional confinement caused by the proximity of the 'bottom' Si/SiO_2 interface is the origin of this behavior. Two main competing factors are at play: As the thickness W of the Si layer is reduced, the increasing separation of the subbands and the increasing scattering form factor ($\approx \int |\psi(z)|^4 dz \sim W^{-2}$ for intra-subband phonon-assisted processes, $\Gamma_{\nu\mathbf{K}'\mu\mathbf{K}} \sim W^{-3}$ for thin Si layers) result in an enhancement and in a reduction, respectively, of the carrier mobility.

For n-type channels, calculations have shown that as W is reduced to 3 or 4 nm, the enhancement of the form factor causes a reduction of the mobility.[47,49,75,76] As the Si thickness approaches 3-to-4 nm, the increasing energetic subband-spacing boosts the mobility. But once the Si thickness is reduced below approximately 3 nm,

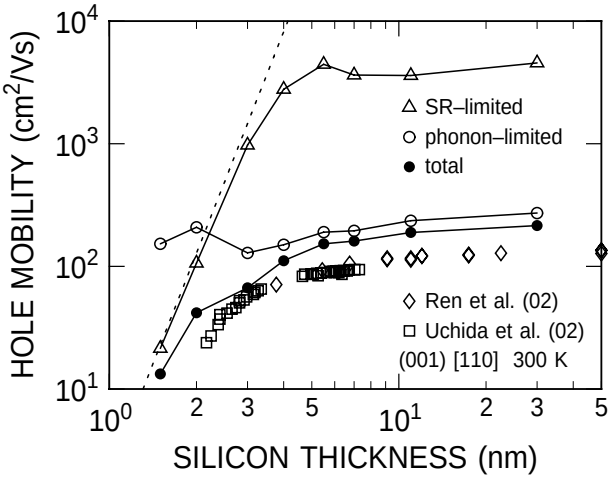

FIG. 17. Calculated total (solid circles), phonon-limited (open circle), and surface-roughness-limited (open triangles) hole mobility for the (001) surface along the [110] direction at 300 K and a surface field of 4×10^5 V/cm ($n_s \approx 2.6 \times 10^{12}$ cm^{-2}) as a function of the thickness of the Si layer. The symbols are calculated data, the lines are only a guide to the eye. Experimental data from Ren *et al.* (Ref. 43, open diamonds, for $n_s \approx 3 \times 10^{12}$ cm^{-2}) and Uchida *et al.* (Ref. 75, open squares, for $F_{eff} = 3 \times 10^5$ V/cm) are also shown. The dotted line shows the 'expected' W^{-6} dependence of the mobility in the regime in which scattering with surface-roughness dominates transport in thin layers.

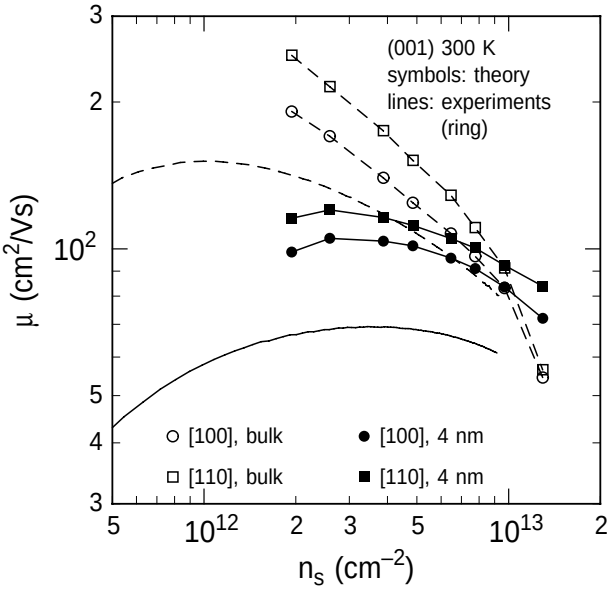

FIG. 18. Calculated hole mobility at 300 K for the (001) surface along the [100] (circles) and [110] (squares) directions, for bulk Si (open symbols, dashed lines) and for a Si layer 4 nm-thick (solid symbols, solid lines). Experimental data are shown by the solid line (4 nm-thick Si layer) and by the dashed line ('control' bulk device). Since experiments have been performed using ring-FETs, the measured mobility represents an angular average over all directions lying on the (001) plane. The solid and dashed lines connecting the calculated data are only a guide to the eye.

inter-subband processes are completely suppressed and the mobility now drops sharply as the form factor dominates once more. The expected mobility-peak at 3 nm, predicted in Refs. 47, 49, 76 has indeed been observed experimentally[75] in a narrow range of electron density.

For p-type channels the situation is qualitatively similar. As shown in Fig. 8(a), the energetic separation of the lowest-energy hh, lh, and so subbands does not change much down to $W \approx 7$ nm. In this regime, the mobility decreases monotonically, as shown in Fig. 17, mainly because of the enhancement of hole-phonon scattering, because of the enhancement of the form factors, Eq. (28). Note that here μ_{SR} has been calculated using Eq. (20), which assumes a perfectly smooth bottom interface, lacking any experimental guidance about the nature of the roughness at this interface and about the presence or lack of correlation between the roughness at two interfaces. However, the presence of this interface has been accounted for by dealing with the image forces it induces, via Eq. (22). As the thickness of the Si layer is reduced to approximately 5 nm, the increased separation between the hh and lh subbands causes a reduction of inter-subband surface-roughness scattering, so that μ_{SR} exhibits a peak. For thinner layers ($W < 4$ nm, in qualitative agreement with Gold's conclusion for n-channels[45]), inter-subband scattering is ineffective and the mechanism discussed by Gold[45], Sakaki[46], and Ren *et al.*[43] becomes dominant. As the energy separation is increased by an additional shift $\approx \hbar\omega_{op}$, also optical-phonon assisted inter-subband processes are reduced, so that the peak of the phonon-limited mobility is observed at a small Si thickness, ≈ 2 nm. It is interesting to analyze the density and temperature dependence of the hole mobility for a thin Si layer. Figs. 18 compares theoretical and experimental results for $t = 4$ nm. (Note that while the calculations have been performed for the mobility along the indicated direction, experimental results are obtained using 'ring-FETs', which provide a mobility averaged over all directions lying on the surface plane.) There is a qualitative agreement, especially regarding the 'flattening' of the μ *vs.* n_s curves at low densities. Indeed, at low n_s the form factor Eq. (28) is not controlled by the gate bias anymore, but by the limited thickness of the Si layer. Figure 19 illustrates a few additional interesting effects: In (b), note the aforementioned independence on n_s of the phonon-limited mobility. In (c), the behavior of the surface-roughness-limited mobility emphasizes the importance of dielectric screening: At low n_s the density has no effect on the confinement, so that the improved screening at larger hole sheet densities results in an enhancement of μ_{SR}, especially at low temperatures for which screening becomes stronger. Finally, note that in Fig. 17 no peak is observed for the total hole mobility, unlike what calculated and observed for n channels. However, such as peak seems to occur at larger hole densities, as seen in the crossover shown in Fig. 18.

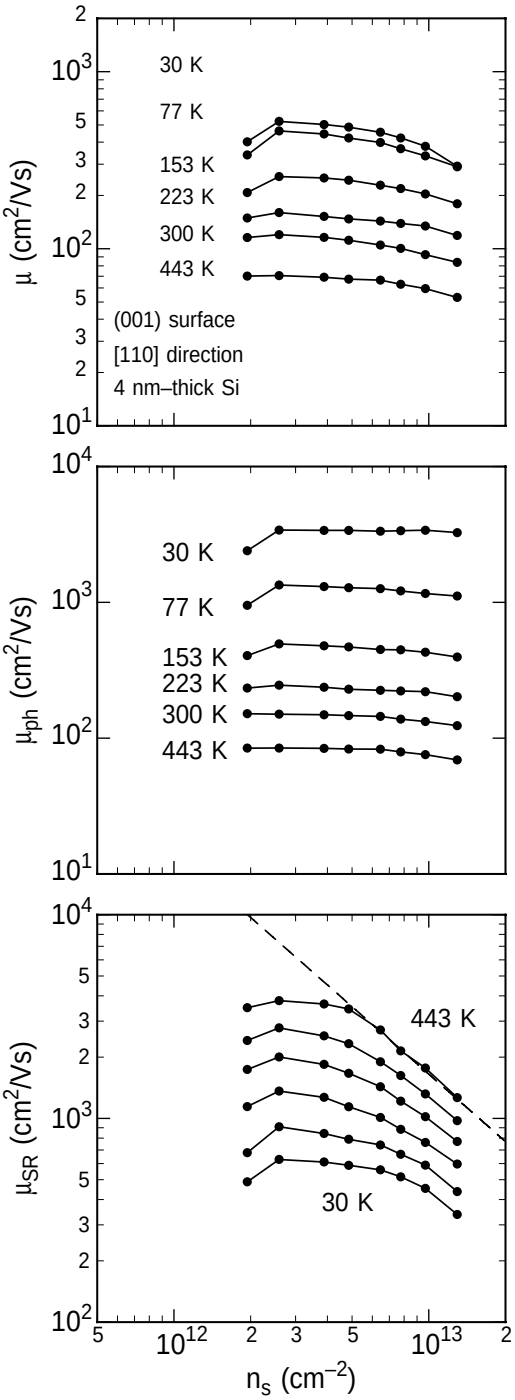

FIG. 19. Calculated total (a), phonon-limited (b), and surface-roughness-limited (c) hole mobility for the (001) surface along the [110] direction at various temperatures for a 4 nm-thick Si layer. The symbols are calculated data, the lines are only a guide to the eye. The dashed line in (c) shows the approximate $n_s^{-1.1}$ behavior of μ_{SR} at 443 K.

V. CONCLUSIONS

In conclusion, the mobility of holes in inversion layers can be calculated quite reliably using an isotropic approximation for the hole-phonon and surface-roughness matrix elements, but retaining the full anisotropy of the (sub)band structure. The 'isotropic' model predicts qualitatively (and quantitatively, to a satisfactory degree within the limitations of the non-self-consistent calculations performed here) the observed mobility for the (001), (011), and (111) surfaces, its temperature dependence, and its depression caused by the increase confinement in thin Si layers. Interestingly, the surface-roughness-limited mobility has been found to exhibit a strong temperature dependence caused mainly by a confinement-induced subband-repopulation (resulting in a higher μ_{SR} at higher temperatures for the (001) and (111) surfaces) or by screening and degeneracy effects (resulting in a lower μ_{SR} at higher temperatures for the (011) surface). The application of both compressive or tensile in-plane strain yields a significant enhancement of the mobility, resulting not only from the suppression of inter-subband scattering, as suggested by Tezuka *et al.*[61], but also from a strain-induced reduction of the conductivity mass, as suggested by Nakatsuji[6] and as seen in bulk Si (Ref. 23).

Despite this 'success' of the model, disagreement with available experimental data arises in a couple of cases:

1. For (001) surfaces, the hole mobility along the [110] direction is found to be slightly larger than the mobility along the [100] direction. Although the physical origin of this effect is quite obvious when considering the anisotropy of the subband-structure, it has never been observed experimentally.
2. The strain-induced enhancement of the mobility under compressive in-plane strain is predicted to be significantly larger than what seen experimentally. Paradoxically, this situation is completely opposite to the case of n-type channels, in which case the experimental data show an enhancement much larger that what predicted theoretically[27]. An increased roughness at the strained-Si/SiO_2 interface, as recently observed in n-channel strained-Si/SiGe MOSFETs[77], could explain the low mobility observed experimentally in strained-Si p-channels, at the price of worsening the disagreement between theory and some observations (see Ref. 27 for a discussion) for n-channels

ACKNOWLEDGMENTS

One of us (MVF) would like to thank Mr. Hiroshi Nakatsuji (Osaka University) and Dr. Shin-ichi Takagi (Toshiba) for having provided copies of their articles prior to publication (Refs. 6 and 75, respectively) and for interesting discussions.

[1] R. Oberhuber, G. Zandler, and P. Vogl, Phys. Rev. B **58**, 9941 (1998).
[2] M. V. Fischetti and S. E. Laux, IEEE Trans. Electron Devices, **ED-38**, 650 (1991).

[3] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
[4] G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. **98**, 368 (1955).
[5] J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).
[6] H. Nakatsuji, Y. Kamakura, and K. Taniguchi, International Electron Device Meeting Tech. Digest (IEEE, Piscataway, New Jersey, 2002), p. 727.
[7] D. Colman, R. T. Bate, and J. P. Mize, J. Appl. Phys. **39**, 1923 (1968).
[8] G. Dorda, H. Friedrich, and E. Preuss, J. Vac. Sci. Technol. **9**, 759 (1972).
[9] L. M. Falikov and N. Garcia, Solid State Commun. **17**, 473 (1975).
[10] E. Bangert, K. von Klitzing, and G. Landwehr, in *Proc. 12th Intern. Conf. Phys. Semicond. Stuttgart,* M. H. Pilkuhn ed. (Teubner, Stuttgart, 1974), p. 714.
[11] E. Bangert and G. Landwehr, Surf. Sci. **58**, 138 (1976).
[12] G. Landwehr, "Quantum transport in silicon inversion layers", in *Festkörperprobleme (Advances in Solid State Physics)*, H. J. Queisser ed. (Pergamon-Vieweg, Braunschweig, 1975), vol. XV, p. 49.
[13] F. J. Ohkawa and Y. Uemura, Prog. Theor. Phys. Suppl. **5**, 164 (1975).
[14] Y. Uemura, Surf. Sci. **58**, 138 (1976).
[15] C. Y. P. Chao and S. L. Chuang, Phys. Rev. B **46**, 4110 (1992).
[16] B. A. Foreman, Phys. Rev. B **48**, 4964 (1993).
[17] Y. X. Liu, D. Z. Ting, and T. C. McGill, Phys. Rev. B **54**, 5675 (1996).
[18] L.-H. Peng, J. Appl. Phys. **82**, 5711 (1997).
[19] E. P. Pokatilov, V. A. Fonoberov, V. M. Fomin, and J. T. Devreese, Phys. Rev. B **64**, 245328 (2001).
[20] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974).
[21] J. M. Hinckley and J. Singh, Phys. Rev. B **41**, 2912 (1990).
[22] J. M. Hinckley and J. Singh, J. Appl. Phys. **76**, 4192 (1994).
[23] M. V. Fischetti and S. E. Laux, J. Appl. Phys. **80**, 2234 (1996).
[24] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[25] C. Moglestue, J. Appl. Phys. **59**, 3175 (1986).
[26] M. V. Fischetti and S. E. Laux, Phys. Rev. B **48**, 2244 (1993).
[27] M. V. Fischetti, F. Gámiz, and W. Hänsch, J. Appl. Phys. **92**, 7320 (2002).
[28] Actually, thanks to the crystal symmetry, it is sufficient to consider the interval $[0, \phi_{max}]$, where $\phi_{max} = \pi/4$, $\pi/2$, and $\pi/3$ for the (001), (011), and (111) surfaces, respectively.
[29] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976).
[30] M. M. Rieger and P. Vogl, Phys. Rev. B **48**, 14276 (1993).
[31] M. V. Fischetti, J. Appl. Phys. **89**, 1232 (2001).
[32] B. Laikhtman and R. A. Kiehl, Phys. Rev. B **47**, 10515 (1993).
[33] Z. Ikonic, P. Harrison, and R. W. Kelsall, Phys. Rev. B **64**, 245311 (2001).
[34] P. Lawaetz, Phys. Rev. **174**, 867 (1968); **183**, 730 (1969).
[35] J. D. Wiley, Solid State Comm. **8**, 1865 (1970).
[36] M. Costato and L. Reggiani, Lett. Nuovo Cimento **4**. 848 (1970); Phys. Status Solidi (b) **58**, 471 (1973).
[37] K. Takeda, A. Taguchi, and M. Sakata, J. Phys. C: Solid State Phys. **16**, 2237 (1983).
[38] F. L. Madarasz and F. Szmulowicz, Phys. Rev. B **26**, 2101 (1981).
[39] H. Ehrenreich and A. W. Overhauser, Phys. Rev. **104**, 331 (1956).
[40] R. E. Prange and T.-W. Nee, Phys. Rev. **168**, 779 (1968).
[41] M. Saitoh, J. Phys. Soc. Japan **42**, 201 (1977).
[42] T. Ando, J. Phys. Soc. Japan **43**, 1616 (1977).
[43] Z. Ren, P. M. Solomon, T. Kanarsky, B. Doris, O. Dokumaci, P. Oldiges, R. A. Roy, E. C. Jones, M. Ieong, R. J. Miller, W. Hänsch, and H.-S. Wong, International Electron Device Meeting Tech. Digest (IEEE, Piscataway, New Jersey, 2002), p. 51.
[44] M. J. Kearney and A. I. Horrell, Semicond. Sci. Technol. **13**, 174 (1998).
[45] A. Gold, Solid State Commun. **60**, 531 (1986); Phys. Rev. B **35**, 723 (1987).
[46] H. Sakaki, T. Noda, K. Hirakawa, M. Tanaka, and T. Matsusue, Appl. Phys. Lett. **51**, 1934 (1987).
[47] F. Gámiz, J. B. Roldán, J. A. López-Villanueva, P. Cartujo-Cassinello, and J. E. Carceller, J. Appl. Phys. **86**, 6854 (1999).
[48] F. Gámiz, J. B. Roldán, P. Cartujo-Cassinello, J. A. López-Villanueva, and P. Cartujo, J. Appl. Phys. **89**, 1764 (2001)
[49] F. Gámiz and M. V. Fischetti, J. Appl. Phys. **89**, 5478 (2001).
[50] C.-Y. Mou and T.-M. Hong, Phys. Rev. B **61**, 12612 (2000).
[51] S. M. Goodnick, D. K. Ferry, C. W.Wilmsen, Z. Lilienal, D. Fathy, and O. L. Krivanek, Phys. Rev. B **32**(12), 8171-8186 (1985).
[52] T. Ishihara, K. Matsuzawa, M. Takayanagi, and S. Takagi, Jpn. J. Appl. Phys. **41**, 2353 (2001).
[53] D. R. Leadley, M. J. Kearney, A. I. Horrell, H. Fischer, L. Risch, E. H. C. Parker, and T. E. Whall, Semicond. Sci. Technol. **17**, 708 (2002).
[54] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, IEEE Trans. Electron Devices, **41**, 2357 (1994).
[55] M. Yang, E. Gusev, M. Ieong, O. Glushenkov, D. C. Boyd, K. K. Chan, P. M. Kozlowski, C. P. D'Emic, R. M. Sicina, P. C. Jamison, and A. I. Chou, submitted to IEEE Electron. Device Lett. (2002).
[56] T. Satô, Y. Takeishi, and H. Hara, Jap. J. Appl. Phys. **8**, 588 (1969).
[57] T. Satô, Y. Takeishi, H. Hara, and Y. Okamoto, Phys. Rev. B **4**, 1950 (1971).
[58] M. Kinugawa, M. Kakumu, T. Usami, and J. Matsunaga, International Electron Device Meeting Tech. Digest (IEEE, Piscataway, New Jersey, 1985), p. 581.
[59] T. Matsumoto, S. Maeda, H. Dang, T. Uchida, K. Ota, Y. Hirano, S. Sayama, T. Iwamatsu, T. Ipposhi, H. Oda, S. Maegawa, Y. Inoue, and T. Nishimura, International Electron Device Meeting Tech. Digest (IEEE, Piscataway, NJ, 2001), p. 663.
[60] F. Schäffler, Semicond. Sci. Technol. **12**, 1515 (1997).
[61] T. Tezuka, N. Sugiyama, S. Takagi, and A. Kurobe, Proc. 25th Int. Conf. Phys. Semicond., Osaka 2000 (Eds. N. Miura and T. Ando), p. 1753.
[62] T. Mizuno, N. Sugiyama, H. Satake, and S. Takagi, 2000

Symposium on VLSI Technology Digest, (IEEE, Piscataway, New Jersey, 2000), p. 210.
[63] T. Mizuno, S. Takagi, N. Sugiyama, H. Satake, A. Kurobe, and A. Toriumi, IEEE Electron Device Lett. **21**, 230 (2000).
[64] T. Mizuno, N. Sugiyama, A. Kurobe, and S. Takagi, IEEE Trans. Electron Devices, **48**, 1612 (2001).
[65] N. Sugii, D. Hisamoto, K. Washio, N. Yokoyama, and S. Kimura, International Electron Device Meeting Tech. Digest (IEEE, Piscataway, NJ, 2001), p. 33-4.1.1
[66] C. W. Leitz, M. T. Currie, M. L. Lee, Z.-Y. Cheng, D. A. Antoniadis, and E. A. Fitzgerald, J. Appl. Phys. **92**, 3745 (2002).
[67] K. Rim, J. Chu, K. A. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Ieong, and H. S. Wong, 2002 Symposium on VLSI Technology Digest, (IEEE, Piscataway, New Jersey, 2002), p. 98.
[68] L. Huang, J. O. Chu, S. A. Goma, C. P. D'Emic, S. J. Koester, D. F. Canaperi, P. M. Mooney, S. A. Cordes, J. L. Speidell, R. M. Anderson, and H. S. Wong, IEEE Trans. Electron Devices, **49**, 1566 (2002).
[69] A. T. Bradley, R. C. Jaeger, J. C. Suhling, and K. J. O'Connor, IEEE Trans. Electron Devices, **48**, 2009 (2001).
[70] P. Kleinmann, B. Semmache, M. Le Berre, and D. Barbier, Phys. Rev. B **57**, 8966 (1998).
[71] C. S. Smith, Phys. Rev. **94**, 42 (1954).
[72] C.-L. Huang, H. R. Soleimani, G. J. Grula, J. W. Sleight, A. Villani, H. Ali, and D. A. Antoniadis, IEEE Trans. Electron Devices **44**, 646 (1997).
[73] J.-W. Lee, H.-K. Kim, J.-W. Yang, W.-C. Lee, J.-H. Oh, M.-R. Oh, and Y.-H. Koh, IEEE Electron Device Lett. **20**, 176 (1999).
[74] D. Esseni, M. Mastrapasqua, G. K. Celler, C. Fiegna, L. Selmi, and E. Sangiorgi, IEEE Trans. Electron Devices, **48**, 2842 (2001).
[75] K. Uchida, H. Watanabe, A. Kinoshita, J. Koga, T. Numata, and S. Takagi, International Electron Device Meeting Tech. Digest (IEEE, Piscataway, New Jersey, 2002), p. 47.
[76] H. Iwata, J. Appl. Phys. **90**, 866 (2001).
[77] S. H. Olsen, A. G. O'Neill, D. J. Norris, A. G. Cullis, N. J. Woods, J. Zhang, K. Fobelets, and H. A. Kemhadjian, Semicond. Sci. Technol. **17**, 655 (2002).