

Influence of source-drain tunneling on the subthreshold behavior of sub-10nm double-gate MOSFETs

M. Städele

Infineon Technologies AG, Corporate Research ND

Otto-Hahn-Ring 6, D-81730 Munich, Germany

email: martin.staedele@infineon.com

Abstract

A microscopic tight-binding formalism is used to analyze source-drain tunneling and subthreshold currents in very short (< 12 nm) double-gate transistors. Tunneling is found to enhance source-drain currents substantially, especially at low temperatures. We quantify and discuss the corresponding degradation of subthreshold swings and compare with results obtained with the WKB approximation. The energy dependence of the effective tunnel mass is predicted and identified as an additional factor that deteriorates the subthreshold behavior.

1. Introduction

Tunneling of electrons from the source of a double-gate transistor through the channel into the drain region (see Fig. 1a) is expected to be an important inherent physical limitation to continued downscaling of very short devices with gate lengths on the order of 10 nm and less. Transport on these length scales is essentially ballistic [1]. Indications of source-drain tunneling are believed to be seen experimentally in ultrasmall devices [2] at low temperatures. Very recently, this mechanism has been included in ballistic device simulations that employ a single-band effective-mass [3-5] and/or the WKB approximation [4-6].

In this paper, we study source-drain tunneling currents in double-gate MOSFETs shorter than 12 nm, using a

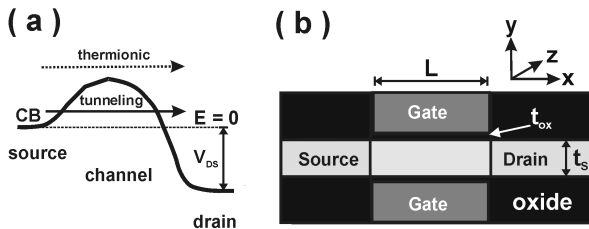


Fig. 1. (a) Profile of the minimum of the lowest conduction bands (CB) of Si along the transport direction in an n-MOSFET under negative gate bias. Electrons whose energy is higher (smaller) than the potential barrier in the channel contribute to the thermionic (tunnel) current. (b) Schematic view of the double-gate transistor structure studied here.

ballistic method that is based on scattering theory and embedded in an atomistic tight-binding (TB) formalism [7]. Compared with other ballistic formulations, this framework allows one to include full band structure effects, and, in principle, to assess the impact of microscopic features such as localized defects or the Si-SiO₂ interface structure on the device performance. As a starting point for more extended calculations [8], we utilize a simplified version of the TB scheme to analyze the degradation of the subthreshold characteristics of ultrashort double-gate transistors as gate lengths are decreased and the temperature is reduced. The TB results are compared with data obtained with the WKB approximation. Also, the energy dependence of the tunneling mass in Si and its influence on subthreshold swings is investigated. Other phenomena that may additionally deteriorate the subthreshold behavior such as band-to-band tunneling, gradual doping profiles, or gate leakage currents, are beyond the scope of this paper and will not be discussed.

2. Method

The double-gate structures which were considered here are characterized by a thin undoped Si channel (with thickness t_{Si}) and two identical gates with lengths L , separated from the channel by SiO₂ gate oxides with thicknesses t_{ox} (see Fig. 1b). The simulations were performed with $t_{Si} = 2$ and 4 nm, L between 2 and 12 nm, and $t_{ox} = 1$ and 2 nm. A metallic gate material with a workfunction of 4.66 eV and abruptly decaying dopant profiles in source and drain (with carrier concentrations of 10^{20} cm⁻³) were assumed. A relatively large drain-source voltage $V_{DS} = 1$ V has been used to accentuate the tunneling effects.

Tunneling currents are calculated using the formula $I = I_0 \int D(E) f(E) dE$, where E is the electron energy (measured from the bottom of the conduction bands in Si), $f(E)$ denotes the Fermi distribution function of the electrons in the source contact, I_0 is a constant, and $D(E) \sim \int dk_z t(E, k_z)$ represents an integral over electron transmission coefficients $t(E, k_z)$. These coefficients depend on the one-dimensional wavevector k_z parallel to the Si-SiO₂ interfaces (z direction, compare Fig. 1b for the axis

orientation) and can be obtained, including size quantization effects, by solving the 3D Schrödinger equation in a basis of localized valence orbitals [7] of the atoms that form the channel. To obtain the results presented in this paper, we have employed this formalism, but replaced $t(E, k_z)$ with bulk transmission coefficients $t(E, k_y, k_z)$ that were calculated for a set of slices along the channel (x) direction, and subsequently integrated over k_y and y . This approximation neglects size quantization effects, which lead to sizeable threshold voltage shifts for $t_{Si} < 10$ nm [9] but do not yield a subband profile that is substantially different from the conduction band profile. Therefore, we believe that the present approach preserves most of the essential physical aspects of source-drain tunneling. The sp^3 tight-binding parametrization of Ref. [10] was used, and the potential profiles were non-selfconsistently obtained from a commercial drift-diffusion simulator [11]. They represent a good approximation to fully quantum mechanical potentials in the subthreshold region, since the carrier densities in the channel are small [12]. For the WKB calculations, we have assumed that only the electrons in the conduction band ellipsoids with the transverse mass $m_t = 0.20 m_0$ in x direction contribute significantly to the tunneling current. These states are termed ‘ k_0 electrons’ in the following, since they have nonzero wavevector components in the yz plane. For the conduction band minima which contain states with the longitudinal mass ($m_l = 0.91 m_0$) in tunneling direction (we term them ‘ $k = 0$ electrons’), these components are zero.

3. Results

Energy dependence of the tunneling mass in bulk Si.

First, we briefly discuss the energy dependence of the tunneling mass m_{tunn} in Si, which is generally ignored in effective-mass based approaches but included in the full-band TB scheme. Since source-drain tunneling currents depend exponentially on m_{tunn} , a decrease of m_{tunn} from its value at the conduction band edge ($m_{tunn} = m_t = 0.20 m_0$) has the potential to increase the subthreshold currents substantially. To quantify these effects, we have calculated the complex bands of Si along the x direction through the conduction band minima at k_0 (see Fig. 2). Note that at a given energy, the transmission coefficients are proportional to $\exp[-\text{Im}(k) \delta x]$, where $\text{Im}(k)$ is the smallest imaginary part of the allowed wavevectors in the Si band gap (indicated by the arrows in Fig. 2) and δx the tunneling distance. Interestingly, we find that these wavevectors belong to *two different* bands which cross at $E \approx -1$ eV (this holds qualitatively also for the $k = 0$ electrons). Figure 3 depicts the energy dependence of the tunneling mass extracted from the complex bands (Fig. 2) via the relation $\text{Im}(k)^2 / m_{tunn}(\epsilon) = 2\epsilon\hbar^{-2}$; here $\epsilon = -E$ denotes the (positive) energy measured from the conduction band minimum. The mass varies linearly between the conduction band minimum and the kink at E

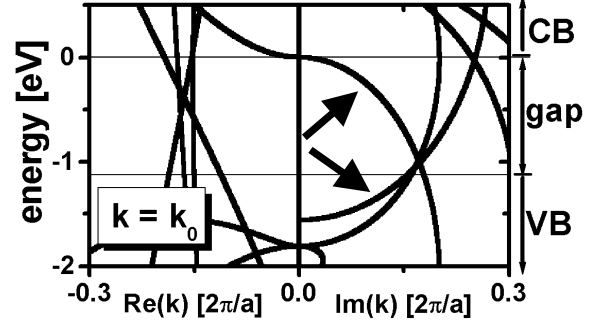


Fig. 2. Complex bands of Si along the tunneling direction through $k = k_0$ (see text). The conduction and valence band regions are denoted as ‘CB’ and ‘VB’, respectively. The arrows indicate the imaginary bands that determine the transmission coefficients of the tunneling states; $a = 0.357$ nm is the lattice constant of Si.

≈ -1 eV, which corresponds to the band crossing mentioned above.

The variation of the tunnel mass of the k_0 states is relatively small: it decreases by about 15 % between the conduction band minimum and the center of the gap. This change can be cast into the form $m_{tunn}(\epsilon) = m_{tunn}(0)(1 - \alpha_{el} \epsilon)$ with $m_{tunn}(0) = 0.200 m_0$ and $\alpha_{el} = 0.050 \text{ eV}^{-1}$. The inset of Fig. 3 illustrates exemplarily the change of the inverse subthreshold slopes S upon a similar (up to 15 %) increase of the (constant) tunnel mass, as predicted by WKB calculations for a representative device with $L = 6$ nm. S increases by about 8 % as the tunnel mass is decreased from $0.20 m_0$ to $0.17 m_0$.

It is interesting to note that for $E < -0.65$ eV, the mass of the $k = 0$ electrons (also shown in Fig. 3) is *smaller* than the mass of the k_0 electrons. This feature cannot be included in a single-band effective-mass model.

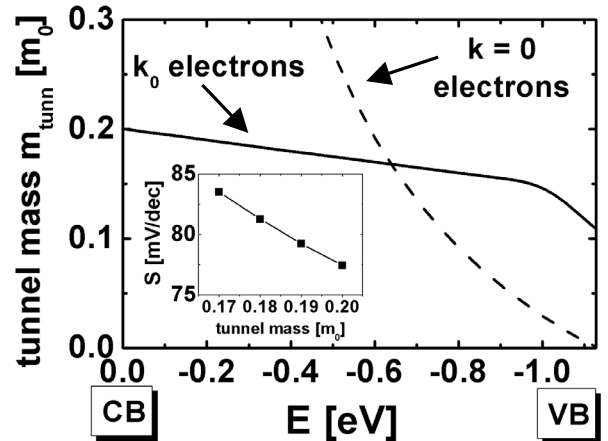


Fig. 3. Energy dependence of the effective tunnel masses in the band gap of Si, for electrons with wavevectors k_0 (solid line) and $k = 0$ (dashed line, see text). The energetic position of the bottom (top) of the conduction (valence) bands is indicated by ‘CB’ (‘VB’). The inset shows exemplarily the influence of the WKB tunnel mass on the inverse subthreshold swing of a typical device with a gate length of 6 nm.

Influence of source-drain tunneling on transfer characteristics at 300 K. We turn now to the discussion of the subthreshold characteristics, comparing TB results for double-gate devices with L ranging from 2 to 10 nm and with $t_{\text{Si}} = 2$ nm and $t_{\text{ox}} = 1$ nm (see Fig. 4). Data with and without the source-drain tunneling contribution are shown. The most salient feature is the strong decrease of the currents if L is increased: long gates control the channel more effectively and yield less drain-induced barrier lowering; this emphasizes the important role of electrostatic channel control. All transistors with $L \geq 4$ nm can be turned off fairly well because of the very small thicknesses of the channel and the gate oxide. Importantly, tunneling enhances the source-drain currents by a factor of about 1.5 ($L = 10$ nm) to 200 ($L = 4$ nm), degrading the subthreshold slopes significantly for the shorter devices. We have checked that a reduction of V_{DS} from 1 to 0.6 V reduces the currents due to the smaller drain-induced barrier lowering, but only slightly; the subthreshold slopes remain almost unaffected for the device parameters chosen here.

The corresponding WKB values for the thermionic and tunneling currents (not shown) are close to the TB data (typically within a factor of 3). This may be somewhat surprising, given the number of approximations in the WKB scheme such as the simple integral formula for the transmission coefficients [6], the assumption of a constant tunnel mass in the Si band gap, and a single band.

To gain a better understanding of the tunneling-induced current enhancement, we show in Fig. 5 the quantity $D(E) f(E, T)$ for $V_G = -0.4$ V and $V_{\text{DS}} = 1$ V. This product represents the energy-resolved contributions to the total ballistic current [to obtain the data in the figure, $D(E)$ has been approximated by the relevant contributions of the \mathbf{k}_0 electrons at the minima of the conduction bands of Si]. The arrows indicate the energies E_{barrier} of the maxima of the corresponding potential

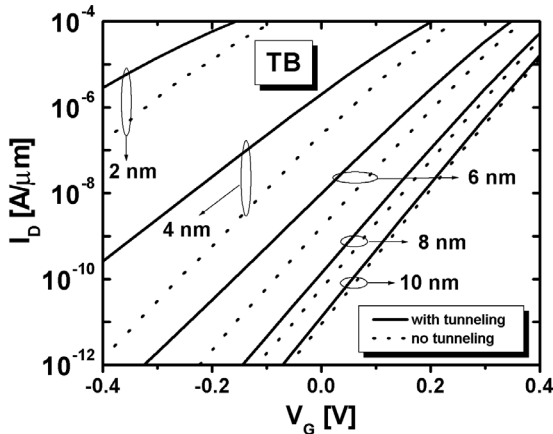


Fig. 4. Comparison of subthreshold drain currents I_D at $V_{\text{DS}} = 1$ V in double-gate transistors with $t_{\text{Si}} = 2$ nm and $t_{\text{ox}} = 1$ nm, calculated with the tight-binding (TB) formalism. Solid (dotted) lines refer to data including (excluding) the source-drain tunneling contribution. The numbers denote the corresponding gate lengths L .

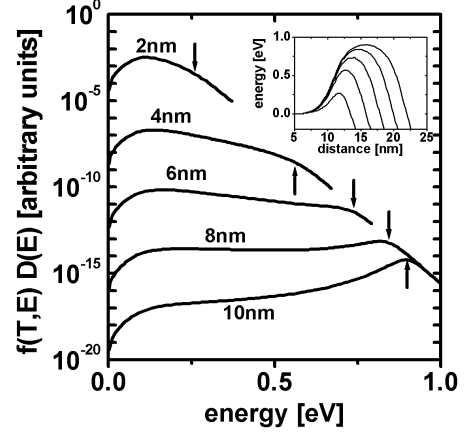


Fig. 5. The product $f(E, T) D(E)$ of the Fermi function $f(E, T)$ and the tight-binding transmission $D(E)$ for gate lengths from 2 to 10 nm at $T = 300$ K. The heights of the potential barriers are indicated by the arrows; $V_G = -0.4$ V, $V_{\text{DS}} = 1$ V. The inset shows the corresponding potential barriers in the center of the channel.

barriers (shown in the inset); note that the contributions for $E < E_{\text{barrier}}$ ($E > E_{\text{barrier}}$) yield the tunnel (thermionic) current. In the $L = 10$ nm transistor, most of the tunnel current stems from electrons with energies close to the top of the barrier, whereas the contributions of the cold electrons in the source contact are strongly suppressed. This changes as L is decreased: at a gate length of $L = 8$ nm, all the electrons in the source reservoir with $E < E_{\text{barrier}}$ contribute almost equally to the current; at even smaller L , the cold electrons at source dominate the tunnel current completely since the potential barriers become considerably lower and thinner. We note that a qualitatively similar effect occurs in tunneling through ultrathin gate oxides when the thickness is reduced to about 4 nm [13].

These considerations may provide a qualitative explanation for the small differences between the TB and the WKB results. In case that the tunnel current consists of electrons with high energies, both the WKB and TB transmission coefficients are close to unity and, thus, almost identical. If, on the other hand, cold electrons dominate the currents, the barriers are short and the transmission is relatively high. However, possible systematic differences between WKB and TB transmission coefficients are enhanced if the transmission is small, i.e., for long and high barriers.

To illustrate how the subthreshold behavior depends on the thicknesses of the gate oxide and the Si channel, Fig. 6 shows the influence of the parameters L , t_{ox} , and t_{Si} on the inverse subthreshold slope S . The tunneling effects (= differences between full and empty symbols) are more conspicuous for the thinner oxide (1 nm, left panel), since it allows for better gate control and thus, higher barriers that suppress the thermionic current effectively. We have checked that drain-induced barrier lowering is rather small for the combination of the thinnest layers ($t_{\text{ox}} = 1$ nm and $t_{\text{Si}} = 2$ nm); this is also supported by the excellent S values of the pure thermionic current.

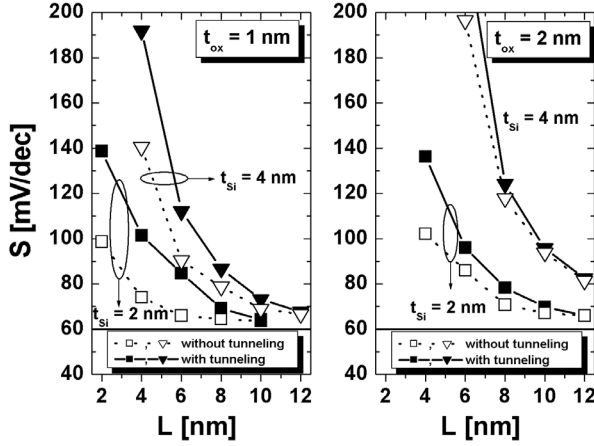


Fig. 6. Gate length dependence of inverse subthreshold slopes S for $t_{ox} = 1$ and 2 nm and $t_{Si} = 2$ and 4 nm. Full (empty) symbols include (do not include) the tunneling contribution to the subthreshold currents.

This implies that a smaller V_{DS} does not lead to considerably lower tunneling currents. A thicker gate oxide (2 nm) enhances the drain-induced barrier lowering, and the resulting drastic increase of the thermionic current increases S strongly and masks the tunneling contribution.

Subthreshold behavior at lower temperatures. At small L , tunneling of electrons with low energies becomes increasingly important as the temperature is reduced. Below a critical temperature T_c , the subthreshold slope converges quickly to a constant value [2]. This is shown in Fig. 7 for the structures with $t_{Si} = 2$ nm and $t_{ox} = 1$ nm (Figs. 4 and 5 refer also to these devices). As expected from the 300 K results in Fig. 5, S decreases considerably for $L = 8$ and 10 nm as the temperature is decreased, since the current ceases to be dominated by electrons with high energies. From Fig. 7, we find $T_c \sim 250$ and 180 K for the 8 and 10 nm devices, respectively. This trend (the smaller L , the higher T_c) is qualitatively consistent with the results presented in Ref. [2]. If tunneling currents are dominated by cold electrons (i.e., at lower T and/or for smaller L), the inverse subthreshold slopes calculated with the TB method are found to be $\sim 10\%$ larger than their WKB counterparts. As mentioned above, only in these cases appear the differences between the two methods; they can be partly explained by the reduced tunnel mass in the TB formalism that leads to a very similar degradation of the subthreshold slopes (see inset of Fig. 3), and also partly by the approximate WKB formula [6] for $D(E)$ itself.

4. Summary and Conclusions

Using an atomistic tight-binding method and a WKB scheme, we have studied the influence of source-drain tunneling on the subthreshold behavior of very small double-gate transistors. The WKB currents are found to be fairly close to the TB results; the inverse subthreshold swings differ by less than about 10% , which is related to

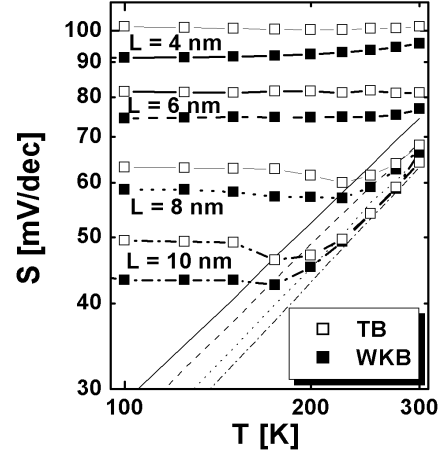


Fig. 7. Temperature dependence of inverse subthreshold slopes S for various gate lengths L , $t_{Si} = 2$ nm, $t_{ox} = 1$ nm, and $V_{DS} = 1$ V, calculated with the tight-binding (TB) scheme and the WKB method. The thin lines indicate the thermionic WKB contribution only.

the reduction of the tunnel mass in the band gap of Si. Unlike the thermionic current, which can be controlled by choosing sufficiently thin oxide and channel layers, the tunneling contribution will limit the minimally achievable subthreshold slopes for gate lengths below $\sim 6 - 8$ nm and result in rather small I_{on}/I_{off} ratios. Ultimately, this intrinsic tunneling component cannot be eliminated by electrostatics, but may be reduced, for instance, by changing the tunneling mass via a different crystal orientation.

We believe that the present work represents a first step towards the atomistic simulation of small MOSFETs.

I am indebted to F. Hofmann, L. Risch, W. Rösner, and M. Specht for enlightening discussions.

5. References

- [1] K. Natori, J. Appl. Phys. **76**, 4879 (1994).
- [2] H. Kawaura and T. Sakamoto, IEICE Trans. Electron. **84-C**, 1037 (2001).
- [3] Z. Ren et al., IEDM Tech. Dig. 2000, p. 107.
- [4] Y. Naveh and K. Likharev, IEEE Electron Device Lett. **21**, 242 (2000).
- [5] A. Poncet and B. Vergnet, Proc. ULIS 2002, p. 123.
- [6] C. B. Duke, in *Tunneling in Solids*, Solid State Physics Vol. 10, ed. By F. Seitz et al., (Academic, New York, 1969).
- [7] M. Städele, B. Tuttle, and K. Hess, J. Appl. Phys. **89**, 348 (2001), and references therein.
- [8] M. Städele, to be published.
- [9] Size quantization effects were included a posteriori by shifting the gate voltage by a constant that was estimated with standard 1D Schrödinger-Poisson simulations. For $t_{Si} = 2$ nm and $t_{ox} = 1$ nm, for example, this shift amounts to 96 mV.
- [10] G. Grosso and C. Piermarocchi, Phys. Rev. B **51**, 16772 (1995).
- [11] ATLASTM device simulation software, Silvaco International Inc., Santa Clara, 2000.
- [12] A. Svizhenko, private communication.
- [13] M. Städele, B. Fischer, B. Tuttle, and K. Hess, to appear in Solid-State Electron., 2002.