

Nonclassical devices in SOI: Genuine or copyright from III–V

S. Luryi ^{a,*}, A. Zaslavsky ^b

^a *Department of Electrical and Computer Engineering, SUNY-Stony Brook, Stony Brook, NY 11794-2350, USA*

^b *Division of Engineering and Department of Physics, Brown University, Providence, RI 02912, USA*

The review of this paper was arranged by Raphaël Clerc, Olivier Faynot and Nelly Kernevez

Abstract

The combination of semiconductor-on-insulator (SOI) substrates with ultrathin Si (or Ge) channel and gate insulator layers opens new opportunities for nonclassical CMOS-compatible devices and possibly optical sources. Unlike their III–V counterparts, which often came first, SOI-based devices have the crucial advantage of potential integrability with dominant silicon technology. We discuss the examples of lateral and vertical tunneling transistors, as well as a tunneling-based SOI intersubband laser. None of these devices has progressed beyond either proof-of-concept demonstrations or, in the case of the intersubband laser, a purely theoretical concept. Still, the unique characteristics deriving from quantum mechanical tunneling make such devices an interesting playground for innovative device research, especially as standard Si CMOS heads towards the rapidly approaching end of scaling.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

Semiconductor-on-insulator (SOI) transistors built in thin fully-depleted Si channels on top of an insulating buried oxide [1] are an attractive alternative to bulk Si CMOS in ULSI technology [2] due to their superior scaling properties, enhanced performance, and improved device isolation. In mainstream technology, it is clear that one needs both new materials, such as high- κ dielectrics and possibly non-Si channel layers, and new device layouts, such as fin-FETs, to continue scaling down to ultimate CMOS – the main controversy regards the most appropriate device technology that can meet the stringent reliability and yield requirements. At the same time, the continuing miniaturization of SOI devices, with available channel and gate insulator thickness dropping to the nanoscale is opening the door to quantum effect and hot-electron devices. Many such devices, based on quantum tunneling, hot-electron injection or charge quantization, have been demonstrated in III–V heterostructures over the past two decades, driven

by the bandgap engineering capabilities and nearly perfect heterointerfaces of modern epitaxy [3]. Yet the mainstream semiconductor technology continues to be dominated by Si CMOS. Today, it appears hardly conceivable that any stand-alone III–V quantum effect or hot-electron architecture will make inroads against Si ULSI [4], and the same is true for other contender devices, like carbon nanotube or molecular electronics [5]. At the same time, the end of CMOS scaling, visible within a couple of decades due to absolute physical limitations, endows the search for ULSI-compatible innovative devices that can be integrated with standard CMOS with a particular urgency.

In this paper, we will present several device structures that appear suitable for exploitation in the SOI world: lateral and vertical tunneling transistors (TTs), as well as an intersubband laser based on vertical tunneling. All of these devices can and have been envisaged and, in several cases, demonstrated in III–V heterostructures. However, SOI implementations offer the crucial CMOS compatibility, which history indicates to be more important than stand-alone performance. We emphasize that while some of these SOI devices have been demonstrated at a rudimentary proof-of-concept level, others exist only as proposals, and

* Corresponding author. Tel.: +1 631 632 8420; fax: +1 631 632 8494.
E-mail address: serge.luryi@stonybrook.edu (S. Luryi).

none is likely to have immediate technological impact. All require some fabrication advances beyond the current state of the art, whether in terms of material integration or extreme control over layer thickness and composition. At the same time, these devices do not abandon CMOS compatibility or room temperature operation, making them somewhat less exotic than the much researched single-electron devices, like gated quantum dots built in SOI that show strongly nonlinear conductance due to single-charge tunneling [6–8]. What we hope to convey is that novel devices and functionalities can still be grafted onto the rapidly growing SOI technology – an opportunity that device physicists should not miss.

2. Interband tunneling transistors in SOI

A device class that was originally proposed and realized in III–V materials but appears eminently suitable for SOI implementation is the tunneling transistor (TT), where the current transport occurs via quantum-mechanical tunneling. This tunneling can either be interband, with carriers tunneling between conduction and valence bands, as in the Esaki diode [9], or resonant, with carriers tunneling into a reduced-dimensionality subband, as in a resonant tunneling diode (RTD). In both cases, the tunneling of carriers requires an available final state on the other side of the barrier. The hypothetical advantage of TT over standard FETs is twofold. First, unlike thermionic injection over a barrier, tunneling through a barrier can, at least in principle, beat the 60 mV/decade limit on the subthreshold current of an FET, although this holds over a relatively narrow range of V_G and V_D [10]. Second, TTs can produce extremely nonlinear current–voltage characteristics, including regions of negative differential resistance (NDR), that may add logic functionality and hence reduce the area required for a logic gate or a memory cell. A number of TT-based devices and small demonstration circuits, typically involving the bandgap engineering available in epitaxial III–V heterostructures, were investigated over the past decades [3]. In this paper, we will restrict our focus to devices that can be fabricated within silicon-based SOI, only mentioning the direct III–V antecedents when necessary to illustrate the operating mechanism.

First, we consider the lateral interband TT where the drain current flows via interband tunneling modulated by a gate electrode. The original idea for controlling interband breakdown via a gate can be traced back to the 1960s, when Shockley and Hooper proposed a gated avalanche pn junction [11]. However, an avalanche drift region occupies too much area, so a transistor-like layout is preferred for ULSI compatibility. The simplest lateral interband TT layout requires a standard MOSFET with the source and drain of opposite doping, as illustrated in Fig. 1a. This type of structure, first demonstrated in GaAs [12] and subsequently in bulk Si [13–16], uses the gate voltage V_G to influence the electric field F (and hence the interband tunneling current) between the inversion channel and the source or

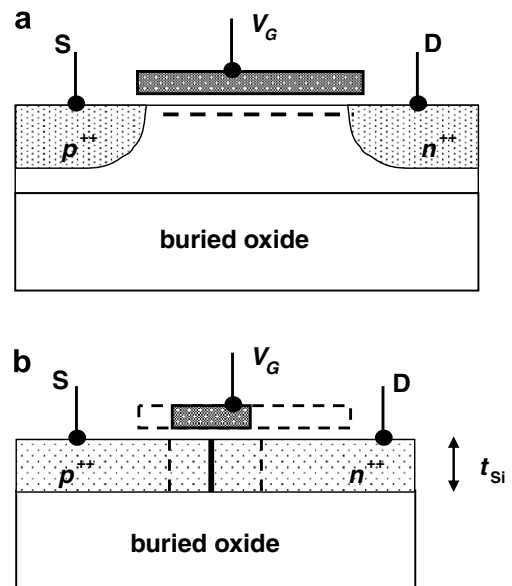


Fig. 1. (a) Schematic layout of a lateral interband n -channel TT with a counterdoped source in partially depleted SOI: dashed line shows the electron inversion channel under the gate, body is tied to source (similar bulk Si devices have been integrated in advanced CMOS [22]). (b) Schematic layout of an interband TT without an inversion channel implemented in SOI, bold line marks the pn junction while the dashed lines indicate the corresponding lateral depletion region. Ideally, the gate overlaps the depletion region only to minimize capacitance (hatched); in the first fabricated prototypes [18], the gate overlapped source and drain space charge regions (dashed).

drain. The effective cross-sectional area of the channel-drain pn junction in Fig. 1a is kept small by the physical depth of the inversion channel. The channel-drain or channel-source junction can be either forward or reverse-biased, since as in the Esaki diode a heavily-doped pn junction passes current in reverse bias as well [17].

An alternative implementation without a channel altogether is based on a heavily-doped lateral pn junction in an ultra-thin Si channel on SOI, with a standard gate oxide and a gate electrode overlapping the junction depletion region – a schematic illustration is shown in Fig. 1b. Here V_G controls the source–drain interband tunneling current, while the cross-sectional area of the pn junction is set by the actual thickness t_{Si} of the Si [18]. In both variants of Fig. 1, the electric field F that drives the tunneling current depends on both V_D and V_G . A serious hindrance to the technological insertion of interband TTs is the lack of an adequate quantitative model for interband tunneling in indirect materials like Si and Ge [19–21]. Most of the available expressions contain adjustable parameters for the electron–phonon coupling constant and predict a sharply increasing current with the electric field F . Their quantitative accuracy is open to question, especially since the calibrations on planar heavily-doped pn junctions [10,22], where maximum F is normal to the junction, are difficult to apply directly to the three-terminal TTs of Fig. 1, where the magnitude and direction of maximum F is self-consistently determined by V_G and V_D . Since interband tunneling

between the heavily-doped contact region and the body is a limiting process to conventional CMOS FETs and is gaining relevance in standard technology, we expect further gains in the theoretical understanding of indirect interband tunneling in the future.

Fabrication and CMOS process flow integration of counterdoped TT analogous to Fig. 1a has been pursued by a German university–industry collaboration [16,23,24], with results on a bulk CMOS process at the 65 nm node recently reported by Nirschl et al. [24]. The measured TT performance appears slightly better than a comparable MOSFET as regards to static leakage in the off state, but the on-state current is also lower, likely due to the added tunneling resistance of the source–channel pn junction. Simulations for this type of device predict an improved subthreshold swing for some range of V_D and V_G biasing [21], although as mentioned previously the quantitative validity of empirical interband tunneling formulae used in industrial simulators is open to question.

The $I_D(V_D, V_G)$ characteristics of a TT without an inversion channel as in Fig. 1b is shown in Fig. 2. Ideally, an ultra-narrow gate would overlap only the depletion region of the lateral pn junction, minimizing gate capacitance and leakage. However, in the first proof-of-concept devices a standard CMOS process with the $L_G = 0.35 \mu\text{m}$ gate overlapping the junction as well as regions of source and drain was employed (see Fig. 1b). The channel thickness t_{Si} was $\sim 40 \text{ nm}$ and the pn junction was formed by double implan-

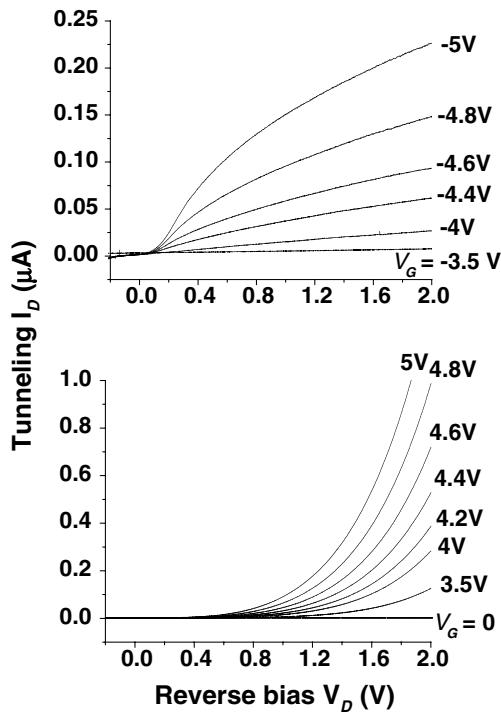


Fig. 2. Reverse-bias $I_D(V_D, V_G)$ curves for interband TT without a channel at $T = 300 \text{ K}$ as a function of V_G of both polarities (double-implanted pn junction doping of 6×10^{19} : 10^{20} cm^{-3} , $L_G = 0.35 \mu\text{m}$, $W = 10 \mu\text{m}$, Si channel thickness $t_{\text{Si}} \sim 40 \text{ nm}$, gate oxide $\sim 5 \text{ nm}$).

tation. As expected, V_G of either polarity adds to the drain-voltage controlled lateral F in the junction, operated in reverse bias, and hence modulates the drain current. However, due to insufficient doping and sharpness of the pn junction, the current drive is unacceptably low for these $W = 10 \mu\text{m}$ devices. Estimates of $F_{\text{MAX}}(V_D, V_G)$ in such devices indicate that a truly abrupt junction (even at lower doping levels) would lead to much higher current drive, with the same caveat that only empirical tunneling current expressions are currently available [10]. Another possibility is the use of a Ge-on-insulator channel, where the interband tunneling barrier would be much lower because of the smaller bandgap – the development of thin epitaxial GeOI layers on Si substrates is particularly promising in this regard [25].

The TT without an inversion channel has no intrinsic current saturation mechanism: as long as F_{MAX} is due to both V_D and V_G , I_D should increase with V_D , limited only by extrinsic factors, such as series resistance in the contacts. If we consider such a device operated in reverse drain bias for high-speed analog amplification, the limitation on its high-frequency operation arises essentially from the capacitance, since in reverse bias there is no minority carrier storage. However, given the absence of the inversion channel, the gate capacitance in the ideal TT of Fig. 1b, where the gate overlaps the source–drain depletion only, is minimal. As for the source–drain capacitance, it depends on the doping level of the pn junction. The trade-off between source–drain capacitance (increasing with pn junction doping) and the tunnel resistance (decreasing with pn junction doping) will determine the proper design point. Obviously, the double implantation procedure employed to produce the devices of Fig. 2 is not ideal, with the source region undergoing two high-dose implantations and the junction sharpness suffering from lateral diffusion during activation. Epitaxial regrowth of the source region would be better, although the resulting nonplanar TT layout would complicate the gate level lithography.

Finally, it should be emphasized that with source and drain contacts of opposite polarity, the no channel variant of lateral TT of Fig. 1b is not affected by short channel effects and does not obey the scaling limitations of a MOSFET. This may become important for future end-of-the-roadmap ULSI devices. Conversely, the counter-doped FET variant of Fig. 1a appears to scale similarly to advanced CMOS, but with lower leakage and hence less static power consumption [24].

3. Vertical resonant tunneling transistor in SOI

Turning to resonant TT, the idea of employing the sharp $I(V)$ nonlinearities and NDR arising due to tunneling into a reduced dimensionality quantum well to perform device functions was extensively explored in the 1980s and 1990s [3]. However, Si-compatible double-barrier epitaxial heterostructures with sufficiently tall barriers and monolayer control over barrier and quantum well thickness have proved elusive. The much studied Si/SiGe heterostructure

is limited by the lattice mismatch considerations to barriers on the order of 0.2 eV [26], insufficient for strong quantization, at least at room temperature.

However, now that the Si channel thicknesses t_{Si} in SOI can be reduced to a few nanometers, one can envisage a vertical TT with lateral extraction of the operating tunnel gate current I_G . The principle of operation, first proposed [27] and demonstrated [28] in epitaxial GaAs/AlGaAs (with a thick AlGaAs second barrier in place of the buried oxide), is illustrated in Fig. 3a. Carriers tunnel from the gate electrode into the quantum well, where two-dimensional (2D) subbands E_N arise from size quantization. Elastic tunneling through an ultrathin gate oxide into subbands E_N conserves energy and transverse momentum k_{\perp} [29], so the tunneling current depends on the alignment between E_N and the occupied states in the gate. This alignment, in turn, is controlled by the voltage on the back-gate electrode V_{BG} , which induces an electric field that changes the 2D confining potential and alters the alignment between the channel and the gate [3,27]. In principle, once E_1 is lowered below the bottom of the occupied states in the gate electrode, the tunneling current is cut off by the energy and

transverse momentum conservation, leading to NDR in the $I_G(V_{\text{BG}})$ characteristic and hence a negative transconductance, $g \equiv \partial I_G / \partial V_{\text{BG}} < 0$. The tunneling I_G is extracted laterally, via a separate contact to the quantum well. If the V_{BG} is increased further, the process is repeated, with the tunneling current dropping due to E and k_{\perp} conservation once E_2 is lowered below the occupied states in the gate – this is the actual alignment illustrated in Fig. 3a. As a result, the $I_G(V_{\text{BG}})$ curve should contain repeated regions of both positive and negative transconductance, provided the quantization in the Si channel is sufficiently strong and the channel thickness t_{Si} is sufficiently uniform.

First proof-of-concept SOI resonant TT devices were reported [30], with $t_{\text{Si}} \sim 5$ nm under the gate produced by sacrificial oxidation and an ultrathin densified native oxide of ~ 1 nm thickness serving as the tunnel barrier. The gate was *in situ* doped poly-Si gate material (P-doped to $\sim 10^{19}$ cm $^{-3}$ range), the device followed standard transistor processing using e-beam gate level lithography ($L_G = 0.1$ μm), nitride spacers, and low-energy source/drain implants. Despite the ultrathin Si channel and gate oxide, these devices successfully operated as standard transistors, with good subthreshold slope and acceptable drain-induced barrier lowering [30,31]. Once the channel is depleted with $V_G < V_T$, an even more negative V_G results in a slowly increasing and relatively temperature-insensitive tunneling current through the ultra-thin gate oxide. At low $T = 4.2$ K, this tunneling current can be modulated by V_{BG} , as shown in Fig. 3b. For this measurement, source and drain electrodes are shorted together and biased to a fixed voltage V_D with respect to the grounded gate electrode. The tunneling current I_G is then modulated by V_{BG} and the resulting transconductance g first increases and then drops, with one or more clear minima in-between. The initial increase of g with V_{BG} corresponds to the V_{BG} -induced lowering of E_1 with respect to the gate, leading to a higher I_G . The eventual drop in g at large V_{BG} ($V_{\text{BG}} > 15$ V, see Fig. 3b), corresponds to a large carrier density being established in the Si channel, which screens the V_{BG} -induced electric field. The minima in g at intermediate V_{BG} corresponds to the E_1 subband going out of alignment with the occupied states in the gate (at larger $V_S = V_D$ the next E_2 subband gives rise to a second feature). In these early, proof-of-concept devices, g does not really go negative and the V_{BG} required to modulate I_G is unacceptably large. However, a truly uniform (and thinner) Si channel would lead to much sharper features at noncryogenic temperatures, while a thinner buried oxide would (an expected development in SOI technology) would reduce the necessary V_{BG} . Finally, a major problem with the proof-of-concept implementation of Fig. 3b is the ultrathin SiO $_2$ tunneling barrier that keeps the tunneling I_G to impractically low values because of its $\Delta E_C > 3$ eV conduction band offset. This very high barrier suppresses tunneling in standard MOSFETs, which is a plus because there the tunneling electrons mean an undesirable gate leakage. However, there is currently an enormous research effort into

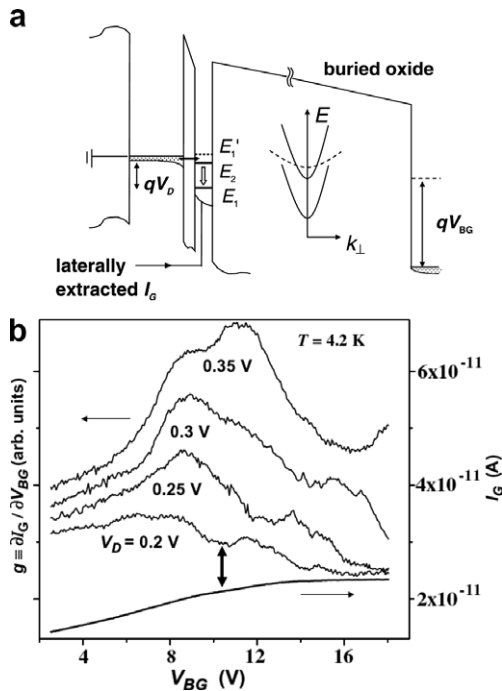


Fig. 3. (a) Schematic vertical band diagram through the VTT under bias, indicating the electron tunneling into the quantized Si channel (the lowest three subbands E_1 , E_2 and E_1' are shown) and their subsequent lateral extraction via the (shorted) source and drain contacts. Control of the tunneling current I_G via V_{BG} is due to the penetration of the electric field through the Si channel, which alters the alignment of E_N with poly-Si gate. Inset shows the in-plane dispersions of the lowest subbands, the E_1' subband (dashed) is due to the light electron mass normal to the quantum well and is therefore heavy in the plane of the QW. The open arrow shows the radiative transition of the proposed intersubband laser. (b) Tunneling $I_G(V_{\text{BG}})$ characteristic at $T = 4.2$ K for $V_S = V_D = 0.2$ V (arrow marks the transconductance $g \equiv \partial I_G / \partial V_{\text{BG}}$ minimum), together with smoothed $g(V_{\text{BG}})$ for various $V_S = V_D = 0.2 - 0.35$ V. At $V_D = 0.35$ V a second g minimum appears.

alternative high- κ dielectrics, in order to enhance the gate-to-channel capacitance (and hence the MOSFET transconductance). A number of metal oxides that have been investigated for high- κ applications, such as zirconium silicates and barium oxides, have relatively low ΔE_C , in the 0.5–0.8 eV range [32]. While such low barriers are problematic for standard MOSFETs, they might prove ideal for vertical TTs. Evidently, the high- κ dielectric barriers would require both smooth interfaces and low defect densities to keep interface-roughness-assisted and defect-assisted tunneling components acceptably small and prevent these contributions to the valley current from swamping the resonant current peak [3]. Another important valley current component is due to phonon-assisted inelastic tunneling, where phonon emission (or absorption) relaxes the energy and transverse momentum k_\perp conservation. The phonon-emission mechanism is very strong in polar III–V heterostructures, such as GaAs/AlGaAs RTDs [33]. In Si-based TTs phonon-assisted tunneling should be weaker, but should certainly be taken into account (thus, in the proposed intersubband laser, below, the energy separation of the subbands in the Si quantum well is set below the optical phonon energy).

4. Intersubband laser in SOI

Provided a sufficiently uniform, smooth, and more transmissive tunneling barrier becomes available, a slightly modified vertical TT structure could enable an intersubband laser compatible with dominant SOI technology. The fundamental operating principle involves preferentially filling a higher-energy subband E_2 in a quantum well via tunneling and extracting the carriers from a lower subband E_1 , producing stimulated emission at $\hbar\omega = (E_2 - E_1)$. Population inversion requires that once carriers relax radiatively to E_1 they should leave the active region faster than E_2 is filled. An intersubband source involving resonant tunneling for both filling E_2 and emptying E_1 in a double-well structure was originally proposed for GaAs/AlGaAs [34]; and the same mechanism underpins the rapidly developing of III–V quantum cascade lasers (QCLs) [35]. Since the radiative transition is intersubband, no direct bandgap is required and a Si quantum well can also be used for the gain medium. However, directly translating III–V QCL design into Si-based materials is hampered, as always, by the relatively low barriers and strain in Si/SiGe heterostructures [36].

The proposed SOI-based intersubband laser is based on the vertical TT structure of Fig. 3a. The added modifications are that the poly-Si gate thickness should be kept down to ~ 20 nm and the substrate should be undoped, with only a thin, heavily doped layer near the buried oxide interface to which V_{BG} can be applied. This minimization of doped region is necessary to reduce free carrier absorption, as discussed later.

Exactly as in the vertical TT, electrons tunnel from the gate electrode into the E_2 subband in the Si QW and are

extracted laterally at the contacts. The in-plane dispersions of the relevant subbands in the Si QW are shown schematically in the inset of Fig. 3a. Doubly-degenerate subbands E_1 and E_2 correspond to the heavy $m^* = 0.98m_0$ electron mass in the quantization direction and to the light $m^* = 0.2m_0$ in the plane, whereas the fourfold-degenerate subband E'_1 lies above E_2 and is heavy in the plane. Coupling between E_2 and E'_1 reduces the in-plane diffusivity in the E_2 subband.

Once injected into the E_2 subband, electrons diffuse laterally towards the contacts with a diffusion constant D_2 . They will relax to the lowest E_1 subband with a characteristic lifetime τ and there also diffuse towards the contacts, with a different diffusion coefficient D_1 . As long as the energy separation $(E_2 - E_1)$ lies below the Si optical phonon energy, nonradiative relaxation to E_1 is suppressed. Assuming uniform injection from the gate, the diffusion equation in the QW can be solved, leading to carrier densities n_2 and n_1 for the upper and lower subband, respectively [37]. Defining $L_D \equiv (D_2\tau)^{1/2}$ as the diffusion length in the upper subband with respect to relaxation to E_1 and setting $L_G = 2L$ symmetrically about the $x = 0$ midpoint of the gate, one obtains in the limit of $L_D > L$:

$$\frac{n_2(x)}{n_1(x)} = \frac{D_1}{D_2} \left(\frac{12L_D^2}{5L^2 - x^2} - 1 \right) \quad (1)$$

Since x varies from zero to L , Eq. (1) predicts population inversion even for equal diffusion coefficients $D_1 = D_2$, because the electrons do not have time to relax from E_2 to E_1 during their diffusion towards the contacts. The ratio D_1/D_2 increases the inversion further, as the in-plane diffusivity D_2 is expected to be markedly smaller because of the hybridization of the E_2 subband with the E'_1 subband that lies closest in energy and has a heavy in-plane mass (as well as a large density of states) – see Fig. 3a. Given that optical phonon scattering is turned off in the proposed structure, as $(E_2 - E_1) < \hbar\omega_{LO}$, the main nonradiative relaxation processes are acoustic-phonon assisted or the weaker multi-electron Auger-like processes. In a similarly narrow GaAs QW, the acoustic phonon intersubband scattering time at $T = 77$ K was calculated to be in the several tens of ps range, implying L_D of several μm [38].

Several other necessary conditions are needed of an effective infrared source, using a typical vertical III–V QCL design for comparison. Given the intersubband nature of the $E_2 \rightarrow E_1$ transition, the optical matrix element in our Si QW is, in principle, quite similar to that in a GaAs QW. In order to avoid competing nonradiative relaxation via phonon emission, the Si QW in Fig. 3a should not be too narrow, e.g., if $t_{\text{Si}} = 6$ nm we have $(E_2 - E_1) \sim 30$ meV $< \hbar\omega_{LO}$ in Si. Optical mode confinement for far infrared wavelengths has been a major issue in III–V QCL designs because of the relatively small change in refractive index available in the epitaxial III–V heterostructures. On the other hand, the active region in epitaxial III–V QCLs can be repeated many times to bring the total active thickness

closer to λ , increasing the gain to counter the additional optical losses due to incomplete confinement. In our device, the structure is based on a single QW and cannot easily be repeated, due to the SOI substrate fabrication technology. As a result, the optical confinement Γ at $\lambda = 40 \mu\text{m}$ is very weak for the TM mode, despite the large refractive index discontinuity between Si and SiO_2 [39]:

$$\Gamma = 2\pi^2(d/\lambda)^2(\epsilon_{\text{SiO}_2}/\epsilon_{\text{Si}})(\epsilon_{\text{Si}} - \epsilon_{\text{SiO}_2}) \sim 10^{-3} \quad (2)$$

where $d \sim 25 \text{ nm}$ is the total thickness of the active structure (thin 20 nm poly-Si gate, tunneling dielectric, and $t_{\text{Si}} = 6 \text{ nm}$ SOI channel). Consequently, there is no significant difference between the optical field intensity in the Si QW gain region and the doped poly-Si gate and substrate regions, which are the regions of optical loss due to free-carrier absorption. Estimating the energy broadening in the Si QW to be $\sim 5 \text{ meV}$, corresponding to monolayer fluctuations in t_{Si} and assuming strong population inversion ($n_2 - n_1 \approx n_2$ as predicted by Eq. (1) for $L_D < L$, we obtain a gain $\alpha \sim 350(n_2/10^{11}) \text{ cm}^{-1}$ [37]. This gain must overcome the optical losses due to free-carrier absorption in the poly-Si emitter, in the narrow doped region under the buried oxide, and in the lateral contacts, as well as possible absorption in the SiO_2 cladding layers [40]. Free-carrier absorption in heavily doped semiconductors α_L is given by $\alpha_L \sim N\lambda^2$, where N is the carrier concentration. In n-Si, absorption in the infrared was measured by Spitzer and Fan [41]: at $T = 77 \text{ K}$ the absorption for Si doped $N_D \sim 3 \times 10^{17} \text{ cm}^{-3}$ is roughly $\alpha_L \sim 150 \text{ cm}^{-1}$ when extrapolated to $\lambda = 40 \mu\text{m}$. Since the total thickness of the doped regions in our device is 40 nm whereas the quantum well $t_{\text{Si}} = 6 \text{ nm}$, to overcome the free-carrier absorption losses we need to get the sheet density n_2 in the upper E_2 subband to reach at least $3 \times 10^{11} \text{ cm}^{-2}$. The tunneling current densities measured in the vertical TT of Fig. 3b were far too small to achieve such a charge density in the well, but an alternative dielectric barrier with a smaller barrier height [32,42] could be used to increase the current density. As a result, operation at $\lambda = 40 \mu\text{m}$ may be feasible.

The remaining question regarding the feasibility of SOI-based unipolar laser concerns the uniformity of the Si QW on top of the BOX layer. The possibility of thinning the Si channel by means of repeated sacrificial oxidation is a relatively recent development [43], and the near-perfect control of layer thickness by epitaxy, crucial to existing III–V QCL designs, is not yet attainable. As a result, devices based on the existence of sharp 2D quantization in the QW suffer from the subband broadening due to QW thickness nonuniformity. In our device, given a reasonable carrier density in the well, the lateral diffusion of these electrons to the contact regions will not be affected by the local potential dips and hills due to thickness fluctuations. At the same time, recent reports of photoluminescence from Si QW down to 2 nm thickness sandwiched between SiO_2 barriers on SOI indicate that thickness nonuniformity is dropping towards a single monolayer [44]. It

is this monolayer nonuniformity that we assumed to estimate the available gain.

5. Conclusions

In this paper, we have briefly presented several nonclassical device prototypes, ranging from tunneling transistors to a proposed intersubband laser, that had their beginnings in III–V heterostructures but can possibly be implemented in SOI-based Si technology. More generally, the ultrathin Si layers becoming available in SOI, in combination with Si-based heterostructures, ultrathin oxides and other Si-compatible dielectrics, beckon device physicists and designers to transfer the entire arsenal of bandgap-engineered, size and charge quantized device ideas to the world of silicon technology. While it is uncertain that any nonclassical devices will have genuine technological impact in the CMOS-dominated semiconductor industry, various end-of-the-roadmap scenarios do envision the integration of alternative devices with CMOS in the not-too-distant future.

Acknowledgements

We are grateful to S. Cristoloveanu, M. Mastrapasqua, S. Deleonibus, R. Beresford, and R. Soref for useful discussions. The experimental work at Brown and Stony Brook has been supported by AFOSR and NSF, and has benefited from a long-term collaboration with the Microelectronics Department of LETI–CEA and ENSERG in Grenoble, France.

References

- [1] Celler GK, Cristoloveanu S. Frontiers of silicon-on-insulator. *J Appl Phys* 2003;93:4955–78.
- [2] The latest publicly released version of the ITRS roadmap is available on the <<http://public.itrs.net>> web site.
- [3] Luryi S, Zaslavsky A. Quantum effect and hot electron devices. In: Sze SM, editor. *Modern semiconductor device physics*. New York: Wiley-Interscience; 1998. p. 253–341.
- [4] Solomon PM. Strategies at the end of CMOS scaling. In: Luryi S, Xu JM, Zaslavsky A, editors. *Future trends in microelectronics: the nano millennium*. New York: Wiley-Interscience; 2002. p. 28–42.
- [5] See Luryi S, Xu JM, Zaslavsky A, editors. *Future trends in microelectronics: the nano, the giga, and the ultra*. New York: Wiley-Interscience; 2004.
- [6] Likharev KK. Electronics below 10 nm. In: Greer J et al., editors. *Nano and giga challenges in microelectronics*. Amsterdam: Elsevier; 2003. p. 27–68.
- [7] Leobandung E, Guo LJ, Wang Y, Chou SY. Observation of quantum effects and Coulomb-blockade in silicon quantum-dot transistors at temperatures over 100 K. *Appl Phys Lett* 1995;67:938–40; Leobandung E, Guo LJ, Chou SY. Single hole quantum-dot transistors in silicon. *Appl Phys Lett* 1995;67:2338–40.
- [8] Saitoh M, Saito T, Inukai T, Hiramoto T. Transport spectroscopy of the ultrasmall silicon quantum dot in a single-electron transistor. *Appl Phys Lett* 2001;79:2025–7.
- [9] Esaki L. New phenomenon in narrow Ge p–n junction. *Phys Rev B* 1958;109:603–4.

- [10] Zhang Q, Zhao W, Seabaugh A. Low-subthreshold-swing tunnel transistors. *IEEE Electron Dev Lett* 2006;27:297–300.
- [11] Shockley W, Hooper WW. The surface controlled avalanche transistor. In: Proceedings of WESCON 1964. p. 12.1; since this reference is generally unavailable, it can be consulted at <<http://www.ece.sunysb.edu/~serge/Shockley-Hooper-1964.pdf>>.
- [12] Uemura T, Baba T. First observation of negative differential resistance in surface tunnel transistors. *Jpn J Appl Phys* 1994;33:L207–L210;
Chun YJ, Uemura T, Baba T. Self-aligned surface tunnel transistors fabricated by a regrowth technique. *Jpn J Appl Phys* 1999;38:L1163–L1165.
- [13] Reddick WM, Amaratunga GAJ. Silicon surface tunneling transistor. *Appl Phys Lett* 1995;67:494–6.
- [14] Koga J, Toriumi A. Three-terminal silicon surface junction tunneling device for room temperature operation. *IEEE Electron Dev Lett* 1999;20:529–31.
- [15] Hansch W, Fink C, Schulze J, Eisele I. A vertical MOS-gated Esaki tunneling transistor in silicon. *Thin Solid Films* 2000;369:387–9;
Bhuwalka KK, Sedlmaier S, Ludstek AK, Schulze J, Eisele I. Vertical tunnel field-effect transistor. *IEEE Trans Electron Dev* 2004;51:279–282.
- [16] Wang P-F, Hilsenbeck K, Nirschl T, Oswald M, Stepper C, Weis M, et al. Complementary tunneling transistor for low power application. *Solid State Electron* 2004;48:2281–6.
- [17] Sze SM. *Physics of semiconductor devices*. 2nd ed. New York: Wiley; 1981. p. 513–39.
- [18] Aydin C, Zaslavsky A, Luryi S, Mariolle D, Fraboulet D, Deleonibus S. Lateral interband tunneling transistor in silicon-on-insulator. *Appl Phys Lett* 2004;84:1780–2.
- [19] Kane EO. Zener tunneling in semiconductors. *J Phys Chem Solids* 1959;12:181–8;
Kane EO, Blount EI. In: Burstein E, Lundqvist S, editors. *Tunneling phenomena in solids*. New York: Plenum; 1969. p. 79–91, The numerical device simulators ATLAS and MEDICI employ expressions based on Kane's model with adjustable parameters in the exponential.
- [20] Schenk A. Rigorous theory and simplified model of the band-to-band tunneling in silicon. *Solid State Electron* 1993;36:19–34;
Schenk A. On the origin of tunneling currents in scaled silicon devices. *IEICE Trans Electron* 1994;E77-C:148–54.
- [21] Tanaka S. An indirect interband tunneling formulation for an arbitrary electric field direction in semiconductors. *Solid State Electron* 1995;38:1017–23.
- [22] Stork JMC, Isaac RD. Tunneling in base-emitter junctions. *IEEE Trans Electron Dev* 1983;ED30:1527–34.
- [23] Wang P-F, Nirschl T, Schmitt-Landsiedel D, Hansch W. Simulation of the Esaki-tunneling FET. *Solid State Electron* 2003;47:1187–92.
- [24] Nirschl T, Henzler S, Fischer J, et al. Scaling properties of the tunneling field effect transistor (TFET): device and circuit. *Solid State Electron* 2006;50:44–51.
- [25] Preisler EJ, Guha S, Perkins BR, Kazazis D, Zaslavsky A. Ultrathin all-epitaxial Ge-on-insulator devices. *Appl Phys Lett* 2005;86:223504.
- [26] Schaffler F. High-mobility Si and Ge structures. *Semicond Sci Technol* 1997;12:1515–49, contains a fairly complete review.
- [27] Luryi S. Quantum capacitance devices. *Appl Phys Lett* 1988;52:501–503.
- [28] Beltram F, Capasso F, Luryi S, Chu SNG, Cho AY, Sivco DL. Negative transconductance via gating of the quantum well subbands in a resonant tunneling transistor. *Appl Phys Lett* 1988;53:219–21.
- [29] Luryi S. Frequency limit of double-barrier resonant-tunneling oscillators. *Appl Phys Lett* 1985;47:490–2.
- [30] Zaslavsky A, Aydin C, Luryi S, Cristoloveanu S, Mariolle D, Fraboulet D, et al. Ultrathin silicon-on-insulator vertical tunneling transistor. *Appl Phys Lett* 2003;83:1653–5.
- [31] Ohata A, Pr et J, Cristoloveanu S, Zaslavsky A. Correct biasing rules for virtual double-gate mode operation in SOI-MOSFETs. *IEEE Trans Electron Dev* 2005;52:124–6.
- [32] Peacock PW, Robertson J. Band offsets and Schottky barrier heights of high dielectric constant oxides. *J Appl Phys* 2002;92:4712–4721.
- [33] Goldman VJ, Tsui DC, Cunningham JE. Evidence for LO-phonon-emissions-assisted tunneling in double-barrier heterostructures. *Phys Rev B* 1987;36:7635–7.
- [34] Kastalsky A, Goldman VJ, Abeles JH. Possibility of infrared laser in resonant-tunneling structure. *Appl Phys Lett* 1991;59:2636–8.
- [35] A discussion of cascade lasers is available in chapter 3 of Luryi S, Xu JM, Zaslavsky A, editors. *Future trends in microelectronics: the nano, the giga, and the ultra*. New York: Wiley; 2004. p. 283–390.
- [36] Dehlinger G, Diehl L, Gennser U, et al. Intersubband photoluminescence from silicon-based quantum cascade structures. *Science* 2000;290:2277–80;
Diehl L, Mentese S, M ller E, et al. Electroluminescence from strain-compensated Si_{0.2}Ge_{0.8}/Si quantum-cascade structures based on a bound-to-continuum transition. *Appl Phys Lett* 2002;81:4700–2.
- [37] Luryi S, Zaslavsky A. On the possibility of an intersubband laser in silicon-on-insulator. In: Iwai H, Nishi Y, Shur MS, Wong H, editors. *Frontiers in electronics: materials to systems (Proc. WOFE-2004)*. Singapore: World Scientific; 2006. p. 411–20. Also published in *Intern. J. High Speed Electronics Syst.* 2006; 16:411–420.
- [38] Smet JH, Fonstad CG, Hu Q. Intrawell and interwell intersubband transitions in multiple quantum wells for far-infrared sources. *J Appl Phys* 1996;79:9305–20.
- [39] Casey Jr HC, Panish MB. *Heterostructure lasers. Part A: fundamental principles*. New York: Academic Press; 1978. p. 54–7.
- [40] Philipp HR. The infrared optical properties of SiO₂ and SiO₂ layers on silicon. *J Appl Phys* 1979;50:1053–7.
- [41] Spitzer W, Fan HY. Infrared absorption in n-type silicon. *Phys Rev* 1957;108:268–71, The samples measured in that study were quite compensated, with carrier density $n < N_D$. Since free-carrier absorption is inversely proportional to mobility, a modern clean Si sample would exhibit less absorption for the same free-carrier density.
- [42] Luryi S, Zaslavsky A. Blue sky in SOI: new opportunities for quantum and hot-electron devices. *Solid State Electron* 2004;48:877–886, discusses the use of low-barrier height dielectrics for tunneling barriers in SOI quantum devices.
- [43] Suzuki E, Ishii K, Kanemaru S, et al. Highly suppressed short-channel effects in ultrathin SOI n-MOSFETs. *IEEE Trans Electron Dev* 2000;47:354–9;
Esseni D, Mastrapasqua M, Celler GK, et al. Low field mobility of ultrathin SOI n- and p-MOSFETs: measurements and implications on the performance of ultra-short MOSFETs. *Tech Digest IEDM* 2000.
- [44] Green MA. Silicon-on-insulator (SOI): a path for integration of silicon light emitters into future microelectronic chips? In: Luryi S, Xu JM, Zaslavsky A, editors. *Future trends in microelectronics: the nano, the giga, and the ultra*. New York: Wiley; 2004. p. 243–252.