Ultrathin silicon-on-insulator vertical tunneling transistor

A. Zaslavsky and C. Aydin
Division of Engineering and Department of Physics, Brown University, Providence, Rhode Island 02912

S. Luryi
Department of Electrical Engineering, SUNY—Stony Brook, Stony Brook, New York 11794

S. Cristoloveanu
IMEP-ENSERG, 23 rue des Martyrs, Grenoble, France

D. Mariolle, D. Fraboulet, and S. Deleonibus
CEA-DRT, LETI/DTS, CEA, 17 rue des Martyrs, Grenoble, France

(Received 29 April 2003; accepted 13 June 2003)

We have fabricated silicon-on-insulator (SOI) transistors with an ultrathin Si channel of ~5 nm, tunneling gate oxide of ~1 nm, and 100 nm gate length. In addition to good transistor characteristics, these same devices exhibit additional functionality at low temperature. The drain current $I_D$ exhibits steps near the turn-on threshold voltage as a function of the backgate $V_{BG}$ bias on the substrate. When operated as a gate-controlled tunneling device, with source shorted to drain and $I_G$ originating from tunneling from the gate to the channel, we observe structure in the $I_G(V_{BG})$ due to resonant tunneling into the quantized channel subbands. In the future, as SOI device fabrication improves and the buried oxide thickness is reduced, these quantum effects will become stronger and appear at lower $V_{BG}$, offering the prospect of ultralarge scale integration-compatible devices with standard transistor operation or quantum functionality depending on the electrode biasing.


Silicon-on-insulator (SOI) transistors built in thin fully depleted Si channels on top of an insulating buried oxide are predicted by the various technology roadmaps to take over from bulk Si complementary metal–oxide–semiconductor (CMOS) devices over the next few years, leading to an ongoing debate about an appropriate double-gate SOI transistor architecture. At the same time, the ongoing miniaturization of SOI devices, with available Si channel and gate insulator thicknesses dropping to the nanoscale, is opening the door to quantum effect devices based on tunneling and/or charge quantization fabricated in and integrable with mainstream CMOS. This is significant because it appears increasingly unlikely that any incompatible quantum effect architecture will make inroads against the rapidly evolving CMOS technology. To date, most of the work in CMOS-compatible tunneling transistors has focused on quantum dots replacing the usual channel, wherein the tunneling into discrete levels combined with Coulomb charging produces a sharply peaked $I_D(V_G)$ characteristic. Here we report on a device, produced by a conventional technological process, that combines standard $I_D(V_G)$ curves under ordinary transistor biasing and a backgate controlled tunneling current $I_G(V_{BG})$ when operated in the quantum capacitance mode.

One of the most widely studied quantum effect devices is the resonant tunneling (RT) structure, where strongly nonlinear current–voltage ($I(V)$) characteristics and negative differential resistance (NDR) arise due to carrier tunneling selection rules into a reduced dimensionality density of states [a quantum well (QW) or dot confined by tunneling barriers]. The difficulty with Si-based RT structures has been the absence of sufficiently high heteroepitaxial barriers. Technologically compatible strained Si/SiGe pseudomorphic RT structures operate with tunneling barriers in the 0.2–0.3 eV range, leading to low-temperature operation and much less pronounced NDR—suitable for spectroscopy of confined states in SiGe QWs and dots, but problematic for devices. Other Si-based resonant tunneling structures involve exotic materials, such as CaF$_2$. In this letter, we report on a tunneling structure with SiO$_2$ barriers built in the standard SOI transistor geometry with $L_G=0.1$ μm gate length, shown in Fig. 1(a). The key difference is that the Si channel is thinned down to ~5 nm, to foster quantization in the channel, and the gate oxide is reduced to ~1 nm to create a tunneling barrier. Despite these stringent criteria, these devices exhibit good transistor $I_D(V_G)$ and $I_D(V_{BG})$ curves at both room and cryogenic temperatures [$V_{BG}$ refers to biasing the substrate under the buried oxide, see Fig. 1(a)]. At low temperature we observe two quantum effects. First, the drain current $I_D(V_{BG})$ at small $V_D$ exhibits clear steps near the threshold, corresponding to channel subbands becoming available for charge transport, indicating sufficient uniformity of the Si channel over the entire active region under the gate. Second, when the device is operated in a purely tunneling mode—source shorted to drain, with drain current due only to tunneling from the gate—we observe structure in the $I_D(V_{BG})$ due to the changing alignment of the quantized channel subbands with the occupied states in the gate, as predicted by the quantum capacitance mechanism and first observed in III–V structures.

The devices were fabricated on an 8 in. silicon line at LETI-CEA on standard UNIBOND SOI substrates (400 nm of buried oxide) using an existing CMOS mask set. The active region Si was thinned to 50 nm using repeated sacrificial oxidation and removal and then locally thinned to ~5 nm in the gate region. The thermal gate oxide was kept as thin as possible by densifying the native oxide, resulting in ~1 nm

*Electronic mail: alexander_zaslavsky@brown.edu
SiO₂. After in situ doped poly-Si gate material (P-doped to \(\sim 10^{19} \text{ cm}^{-3}\) range), the device followed standard transistor processing with low-energy (3 keV) As source/drain implants. The dimensions of the transistors reported here were gate length \(L_G=0.1 \mu\text{m}\) and width of 10 \(\mu\text{m}\). The fabrication sequence was simulated on Silvaco software, using the actual implantation and activation anneal parameters. The resulting cross-sectional view of the device is illustrated in Fig. 1(a).

Standard transistor \(I_D(V_G, V_{BG}=0)\) characteristics for \(V_D=0.1\) V are shown in Fig. 1(b) for \(T=300, 77,\) and 4.2 K. Because of \(n^+\)-poly gate material, the threshold \(V_T<0\), but otherwise the room-temperature characteristics show good subthreshold slope and acceptable drain-induced barrier lowering (compare \(V_D=0.1, 0.6\) V at \(T=300\) K). Detailed performance analysis of these devices as standard transistors will be published elsewhere.\(^{17}\) 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 ultrathin gate oxide. Given the SOI geometry, the transistor can also be turned on using the substrate backgate voltage \(V_{BG}\). Figure 1(c) shows the room-temperature \(I_D(V_{BG}, V_G=-0.8\) V) curve at \(V_D=0.1\) V.

At \(T=4.2\) K, below the threshold \(V_G<\sim -0.1\) V and at small \(V_D=1\) mV, \(I_D\) exhibits clear current steps as the transistor is turned on by \(V_{BG}\), as shown in Fig. 2. These steps, which persist in a weaker fashion to \(T=77\) K, correspond to quantized subbands in the Si channel being pulled down below the source Fermi level. As is clear from Fig. 2, changing front-gate \(V_G\) shifts the \(I_D(V_{BG})\) curve along the \(V_{BG}\) axis:

\[\Delta V_G=12.5\text{ mV}, \ \text{the corresponding } \Delta V_{BG} \approx -0.6\text{ V due to ratio of the buried oxide thickness to the combination of top oxide and Si channel thicknesses—see inset of Fig. 2. The results of Fig. 2 are quite similar to the Si quantum dot transistors,}^{4-6}\text{ except that in dots the } I_D \text{ exhibits sharp peaks followed by NDR regions corresponding to tunneling into discrete states, whereas here we have tunneling into effectively two-dimensional (2D) subbands } E_n \text{ in the channel. Since these 2D subbands contain higher-energy states corresponding to in-plane motion, the NDR is weakened by impurity and phonon scattering-assisted tunneling into these states.}^{18}\text{ Also, inhomogeneities in the Si channel thickness and the SiO}_2/Si interface are certain to broaden } E_n \text{ and, hence, the } I_D(V_{BG}) \text{ steps.}

Figure 3 illustrates the quantum capacitance mode of operation of this device. The gate is grounded and acts as the

---

**FIG. 1.** (a) Cross-sectional view of the device: the Si channel thickness under the gate is \(\sim 5\) nm thick, the front gate oxide is \(\sim 1\) nm. Grayscale contours show the \(10^{19}, 10^{18},\) and \(10^{17} \text{ cm}^{-3}\) doping contours in the source and drain extensions, the channel under the gate is essentially undoped. (b) Standard \(I_D(V_G)\) transfer characteristics at \(T=300, 77,\) and 4.2 K for \(V_D=0.1\) V (and 0.6 V at \(T=300\) K, dashed line), \(V_{BG}=0\). (c) Backgate \(I_D(V_{BG})\) transistor characteristics for \(V_D=0.1\) V, \(V_G=-0.8\) V at \(T=300\) K.

---

**FIG. 2.** Drain current \(I_D(V_{BG})\) characteristic at \(T=4.2\) K and \(V_D=1\) mV for several values of \(V_G\). The current steps correspond to the population of quantized subbands \(E_n\) in the channel under the gate. Inset of Fig. 2 plots the \(V_{BG}\) position of the first current step vs \(V_G\).

---

**FIG. 3.** (a) Schematic vertical band diagram through the device midpoint under bias, indicating the electron tunneling into the quantized Si channel (only the lowest subband \(E_1\) is 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 \(E_1\) with poly-Si gate \(E_F\). (b) Tunneling \(I_G(V_{BG})\) characteristic at \(T=4.2\) K for \(V_S=V_D=0.2\) V (arrow marks the transconductance \(g=\partial I_G/\partial V_{BG}\) minimum), together with smoothed \(g(V_{BG})\) for various \(V_S=V_D=0.2-0.35\) V.
“emitter,” the source and drain electrodes are shorted and biased to a $V_D$, and $I_G$ is modulated by the substrate voltage $V_{BG}$. The schematic vertical band diagram through the midpoint of the device and the tunneling current is shown in Fig. 3(a), with the $\sim 1$ nm gate oxide acting as the thin first barrier, the buried oxide acting as an impenetrable second barrier, and the thinned Si channel acting as the QW containing 2D subbands $E_n$ (only $E_1$ is shown). Electrons tunnel from the $n^-$-poly-Si gate into $E_1$ and are extracted laterally via the source and drain contacts. The tunneling $I_G$ is the only current component in this biasing mode (no measurable substrate leakage is observed for any $V_{BG}$). As in all RT structures, this tunneling $I_G$ depends on the alignment of $E_1$ with the occupied states in the gate—see Fig. 3(a). Three-terminal operation is achieved via $V_{BG}$, which induces an electric field and alters the alignment between the channel and the gate. In principle, once $E_1$ is lowered below the bottom of the occupied states in the gate electrode, $I_G$ is cut off by the energy and transverse momentum conservation. This should lead to a negative transconductance, $g = dI_G/dV_{BG} < 0$, but the effect is weakened by energy or transverse momentum nonconserving tunneling, as well as broadening of $E_1$ due to Si channel nonuniformity and the possible inhomogeneous broadening of the emitter states in the gate electrode due to quantum-size energy shifts in small poly-Si grains.

The first observation of this phenomenon in a Si-based device is shown in Fig. 3(b), where we plot the $I_G(V_{BG})$ and the transconductance $dI_G/dV_{BG}$ for $V_G=0$ and $V_D$ ranging 0.2–0.35 V ($V_S = V_D$). As $V_{BG}$ is increased from 0, the backgate transconductance $g = dI_G/dV_{BG}$ for any given value of $V_D$ first increases and then drops, with one or more (at higher $V_D$) 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 tunneling $I_G$. The eventual drop in $g$ at large $V_{BG}$ [that is, the near saturation of $I_G$ as $V_{BG}$ exceeds $\sim 15$ V at $V_D = 0.2$ V, see Fig. 3(b)], corresponds to a large carrier density being established in the Si channel. The tunneling oxide emitter barrier is then completely screened from the electric field produced by $V_{BG}$. 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. The actual alignment of the 2D subbands with the gate is a complex electrostatic problem that will be the subject of future study.

In estimating the potential impact of such devices, it is worth noting that an analogous RT structure was originally fabricated in a III–V heterostructure by Morkoç and co-workers. There, both barriers were AlGaAs, with a much thicker second barrier ensuring the isolation between the GaAs QW and the substrate. The main technical difficulty in the III–V implementation was making good contact to the QW without leakage to the substrate. This problem is absent in SOI devices, where the buried oxide is essentially impenetrable. For the current generation of UNIBOND substrates, the required $V_{BG}$ to shift quantized subbands in the Si channel runs to $>10$ V because of the 400-nm-thick buried oxide, but much thinner buried oxides will become available as SOI transistors are scaled down.

To summarize, we have fabricated proof-of-concept SOI resonant tunneling transistors, that combine standard transistor $I_D(V_G)$ transfer characteristics at large $V_D$ with backgate control of the tunneling current $I_G(V_{BG})$ as the source and drain are shorted and the front gate is used as the emitter of the resulting RT structure. For now, the features in the tunneling transconductance are pronounced at $T=4.2$ K, but with improved fabrication techniques the operating temperature will increase. As a result, our tunneling transistors offer, at least in principle, the prospect of enhancing silicon integrated circuits with SOI quantum tunneling devices featuring functionally useful nonlinear and NDR characteristics.

The work at Brown and SUNY-Stony Brook has been supported by the AFOSR MURI program (Award No. F49620-00-1-0331, managed by T. Steiner). The device fabrication at LETI-CEA was carried out in the CEA-LETI/CPMA collaboration framework, with PLATO organization teams and tools. A.Z. is grateful for a productive sabbatical stay in the Micro-electronics Department at LETI-CEA and visits to the IMEP-ENSERG laboratory in Grenoble. The authors acknowledge use of the Microelectronics Central Facility at Brown, supported by the NSF MRSEC (Award No. DMR-0079964).

---

1 The latest publicly released version of the ITRS roadmap is available on the http://public.itrs.net web site.


