SiGe quantum dot single-hole transistor fabricated by atomic force microscope nanolithography and silicon epitaxial-regrowth

Xiang-Zheng Bo, L. P. Rokhinson,[a] N. Yao, D. C. Tsui, and J. C. Sturm[b]
Princeton Institute for the Science and Technology of Materials (PRISM), Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08540

A SiGe quantum dot single-hole transistor passivated by silicon epitaxial regrowth with extremely stable Coulomb blockade oscillations has been demonstrated. The quantum dot was defined by atomic force microscopy nanopatterning technique and subsequently passivated by the epitaxial regrowth of silicon. Such passivation of the dot avoids any potential defect states on the dot associated with the Si/SiO_2 interface. Coulomb blockade oscillations controlled by side planar gating at ~ 0.3 K are reproducible, in sharp contrast with the noisy and irreproducible *I-V* characteristics of unpassivated SiGe quantum dot devices. An additional top gate was used to further tune the Coulomb blockade oscillations, enabling a shift in side-gate voltage of up to three periods. © *2006 American Institute of Physics.* [DOI: 10.1063/1.2358398]

INTRODUCTION

Si-based quantum dot (QD) single-electron (or hole) transistors (SETs or SHTs) are of growing interest for their usefulness in single-electron memories,[1] logic operations,[2] and quantum computing.[3] Si-based QDs have been usually defined by the trench isolation, using electron beam lithography and reactive ion etching (RIE) to directly pattern the silicon layers. Defined QD devices with Si metal-oxide-semiconductor field-effect-transistor (MOSFET) structure have been demonstrated, exhibiting periodic Coulomb blockade oscillations at temperatures greater than 4 K, even at room temperature.[1,2,4–6] Memory and logic operations have been realized at room temperature.[1,2] However, most Si-based devices show parasitic multidot transport characteristics when temperatures go down to tens of millikelvins, which unfortunately is the temperature range at which silicon exhibits long coherent times, as required for quantum computing.[3] It has been found that the parasitic extra QDs resulted from the inevitable charge traps at the Si/SiO_2 interface on the surface of the dot.[6] These traps can interfere with the electrical operation of single electron devices either by randomly altering the potential landscape in the vicinity of the lithographically defined QD or by acting as additional parasitic QDs,[6–9] making it unlikely for QD devices with SiO_2 passivation to be used for the study of quantum computing.

Band gap engineering and strain effects in Si/SiGe heterojunctions can lead to superior carrier transport mobility than in Si devices, which has motivated the study of SiGe single-electron/hole transistors.[10–14] Coulomb blockade devices in thin SiGe layers on insulating layers[10] or on silicon substrates[13] [in which the carriers (holes) are confined to the SiGe layer at low temperatures due to valence band offset] have been demonstrated. In these devices, the SiGe was pat-

terned by etching and left unpassivated. Coulomb blockade oscillations were observed at low temperatures but varied from scan to scan. The irreproducibility of the current-voltage characteristics was thought to be due to the trapping/detrapping of carriers at the dot surfaces.

The main contribution of this work is threefold. First, we demonstrate the three-dimensional epitaxial regrowth of silicon over a SiGe QD to electrically passivate the QD surface without the use of any SiO_2. By surrounding the SiGe QD with epitaxial silicon to create a defect-free Si/SiGe hetero-interface, parasitic surface traps could be eliminated entirely in principle. Second, SiGe QD devices (single-hole transistors) reported in this paper with this epitaxial silicon passivation approach show very reproducible and stable Coulomb blockade oscillations at 0.3 K, with extremely narrow linewidths. Third, in addition the primary side gate, an optional top gate for modulating the QD potential has been added, which is then used to tune the *I-V* characteristics of the main side gate.

DEVICE STRUCTURE AND FABRICATION BY ATOMIC FORCE MICROSCOPE (AFM) NANOLITHOGRAPHY AND SILICON EPITAXIAL REGROWTH

The schematic device structure is shown in Fig. 1(a). It consists of a strained doped SiGe layer, which is patterned into regions of source, drain, quantum dot, and two side gates. The entire structure is then passivated by epitaxial silicon. The device fabrication began with a multilayer growth consisting of 10-nm-thick $p^+ - Si_{0.7}Ge_{0.3}$ and 2-nm-thick intrinsic Si cap on a (100) silicon substrate grown by rapid thermal chemical vapor deposition, with growth temperatures of 625 °C for SiGe and 700 °C for Si. The top and bottom 1 nm of the SiGe were not intentionally doped to help keep the boron away from the actual Si/SiGe interface. At temperatures below 50 K, the only conducting layer is the $p^+ - Si_{0.7}Ge_{0.3}$, as carriers in the lightly doped Si substrate and Si cap freeze out. The carrier density in the $p^+ - Si_{0.7}Ge_{0.3}$ layer is $\sim 3 \times 10^{12}$ cm^{-2}, obtained by Hall

[a]Department of Physics, and Electrical Engineering and Computer Science, Purdue University, West Lafayette, IN 47907.
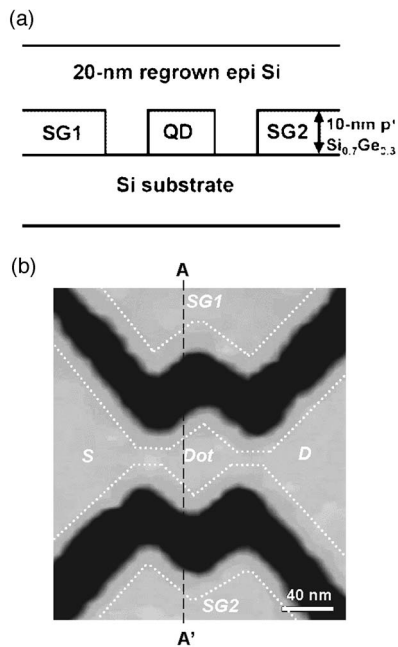[b]Electronic mail: sturm@princeton.edu

FIG. 1. (a) Schematic device structure cross section [along A–A′ in (b)] of *p*-type strained SiGe quantum dot (QD) with electrical potential modulated by SiGe side gates from the same SiGe layer. The dot is passivated by epitaxial silicon. (b) Top-view AFM image of patterned device structure, showing source (S), drain (D), dot, and side gates (SG1 and SG2). The dark area is where $p^+-Si_{0.7}Ge_{0.3}$ layer (light color) was etched. The dotted lines show the estimated actual SiGe dimensions accounting for the wet-etching undercut.

measurement at $T=4.2$ K. The conducting SiGe layer was then patterned by AFM local anodic oxidation (LAO) and selective wet etching [Fig. 2(a)].[13] AFM LAO was performed at room temperature in tapping mode using a Digital Instruments Dimension 3100 AFM. The humidity during the LAO was kept ~70%, and the tip bias voltage was ~−20 V.

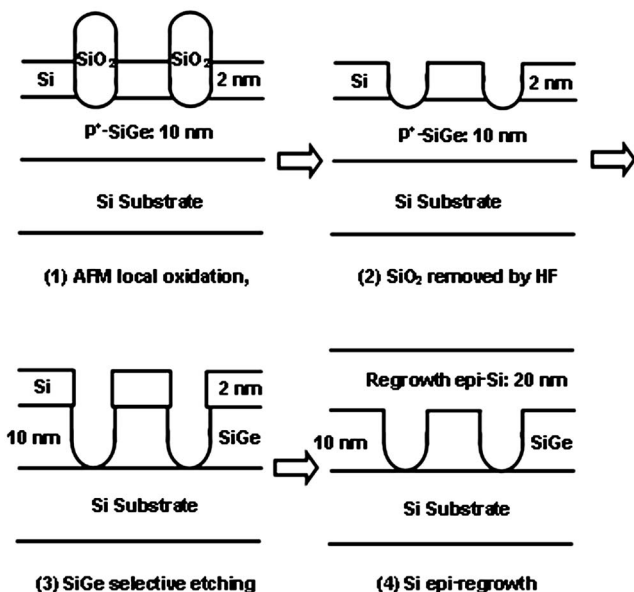The 2-nm-thick Si cap was first locally oxidized by

FIG. 2. Nanopatterning process of Si/SiGe heterostructures: (1) AFM LAO on top Si cap, (2) selective removal of SiO_2 by diluted HF, (3) selective wet etching of SiGe, and (4) Si epitaxial regrowth.

AFM, and then the pattern was transferred to the Si cap by removing SiO_2 with diluted hydrofluoric acid, and was finally transferred to the $p^+-Si_{0.7}Ge_{0.3}$ layer by a selective wet etch[13] to remove SiGe (Fig. 2).[15] The SiGe was not directly patterned by AFM oxidation because the maximum oxide thickness achieved by this process is only ~4 nm, which consumes only ~2 nm of SiGe. The etched trenches defined the p^+-SiGe layer into three regions. Figure 1(b) is an AFM scan, showing this structure immediately after the SiGe etching. In Fig. 1(b), the top and bottom regions are two planar side gates (SG1 and SG2), which are electrically isolated from the central transport region and used to modulate the potential of the central dot. The central region consists of the current path of the source, dot, and drain. The narrow gaps between the source/drain and the dot create potential barriers to isolate the dot from the source and drain. The linewidth, dot diameter, and gap width as drawn by AFM were about 40, 70, and 40 nm, respectively, but the undercut of the isotropic wet etch under the Si cap into the SiGe was estimated to change these to 70, 40, and 5–10 nm, respectively. The final SiGe edges are estimated by dotted lines in Fig. 1(b). The number of holes in the dot is thus estimated to be ~40. Assuming a gap of 10 nm, a one-dimensional (1D) calculation of quantum confinement in the SiGe well between the Si barriers gives ~8 meV. This is the potential barrier for holes between the source/drain (S/D) and the dot, and is the primary factor restricting the device to low temperature operation.

After the QD was patterned, 20 nm epitaxial silicon was regrown on the top and edges of the SiGe and the exposed Si substrate. To remove the native oxide and other impurities on the SiGe and Si surfaces before regrowth, the sample was first cleaned by $H_2SO_4:H_2O_2(1:1)$ for 15 min and subsequently by dilute HF for 1 min before loading into the epitaxial reactor. *In situ* hydrogen baking was performed at a temperature of 800 °C, hydrogen gas flow rate of 3 lpm, and pressure of 10 Torr for 2 min. The subsequent silicon regrowth was done at 700 °C for 20 min using gases of dichlorosilane and hydrogen.[16] Here the relatively low cleaning temperature (800 °C) was chosen in order to suppress boron diffusion from $p^+-Si_{0.7}Ge_{0.3}$ into Si layers. Assuming a boron diffusion coefficient in SiGe of 2×10^{-17} cm^2 s^{-1},[17] the boron diffusion length at 800 °C for 2 min is less than 1 nm, so we expect that the heavy doping in the $p^+-Si_{0.7}Ge_{0.3}$ layer be kept from entering into the silicon substrate. Secondary ion mass spectroscopy analysis and photoluminescence spectra indicate that the cleanliness of the Si/SiGe interface on an unpatterned SiGe surface with the above procedure is comparable to such an interface created without any growth interruption and external processing.[18] Ohmic metal contacts for electrical measurements were achieved by 300-nm-thick aluminum evaporation, lift-off, and forming gas anneal at $T=450$ °C.

After the SiGe layer was patterned before Si epitaxial regrowth, the gap between adjacent doped SiGe regions was insulating (current less than ~100 pA; lower current measurements were limited by instrumentation) at $T=4.2$ K up to applied voltages of 8–10 V before breakdown occurs. After regrowth, the gate "insulator" between the gate and the dot
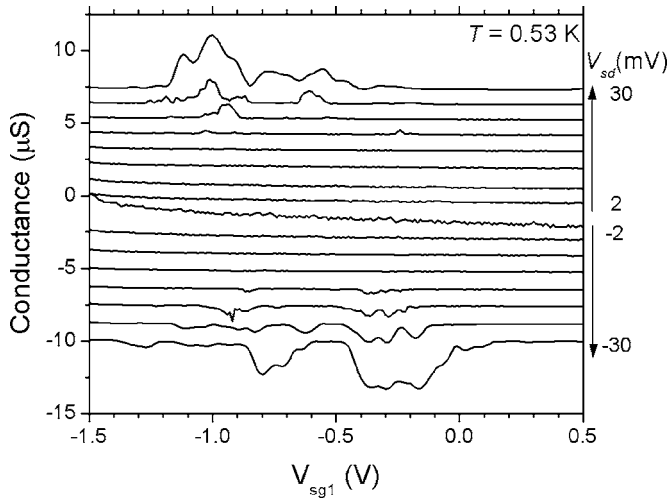
FIG. 3. Dot conductance vs gate voltage for devices without Si regrowth passivation at 0.53 K. Conductance curves corresponding to the source-drain bias from −30 to +30 mV with a step of 4 mV are offset for clarity. One side gate is scanned and the other is held at zero voltage.

region is the regrown Si. The structure insulates because of the valence band offset between the doped strained SiGe and the unstrained Si ($\sim$0.25 eV for $Si_{0.7}Ge_{0.3}/Si$). After regrowth, the breakdown voltage is reduced to $\sim$2–3 V, still sufficient for the gate voltage to modulate the dot potentials.

SIDE-GATE SIGE QUANTUM DOT SINGLE-HOLE TRANSISTOR CHARACTERISTICS

The combination of energy confinement in all three dimensions in the dot leads to an energy spectrum quantized in the dot. The energy difference between empty energy levels and filled energy levels includes Coulomb charging energy e^2/C and quantum energy level spacing $\Delta\varepsilon$. Due to the energy difference, the electrons or holes cannot tunnel from the source to the dot through the narrow opening, where there is a potential barrier due to quantum confinement, and then through another barrier to the drain. This is called the Coulomb blockade effect. Only when empty levels of the dot are aligned by the gate voltage to the source/drain Fermi energies, is the current able to flow.

Without Si epitaxial regrowth, the unpassivated QD devices behaved poorly with broad peaks ($>$0.1 V in gate voltage) at 0.53 K, and the peak positions are irreproducible from scan to scan (Fig. 3). Furthermore, the current peaks are suppressed for a drain-source bias $|V_{ds}|$ of $|V_{ds}|$ less than 18 mV. Such a suppression is characteristic of multiple dot transport.[19] We attribute both the extra multiple QDs and irreproducible conductance oscillations to the trap states at the surface of SiGe, whose occupation fluctuates as a function of time.

After the device is passivated by silicon regrowth, much better electrical characteristics are achieved. The typical conductance oscillations versus the two side-gate voltages (V_{sg1} and V_{sg2}, measured with respect to the source, which is grounded) are observed at $T=0.3$ K and source-drain bias $V_{ds}=100$ μV (Fig. 4).[15] The conductance peaks are extremely narrow with a full width at half magnitude (FWHM) of $\sim$12 mV, corresponding to energy level broadening
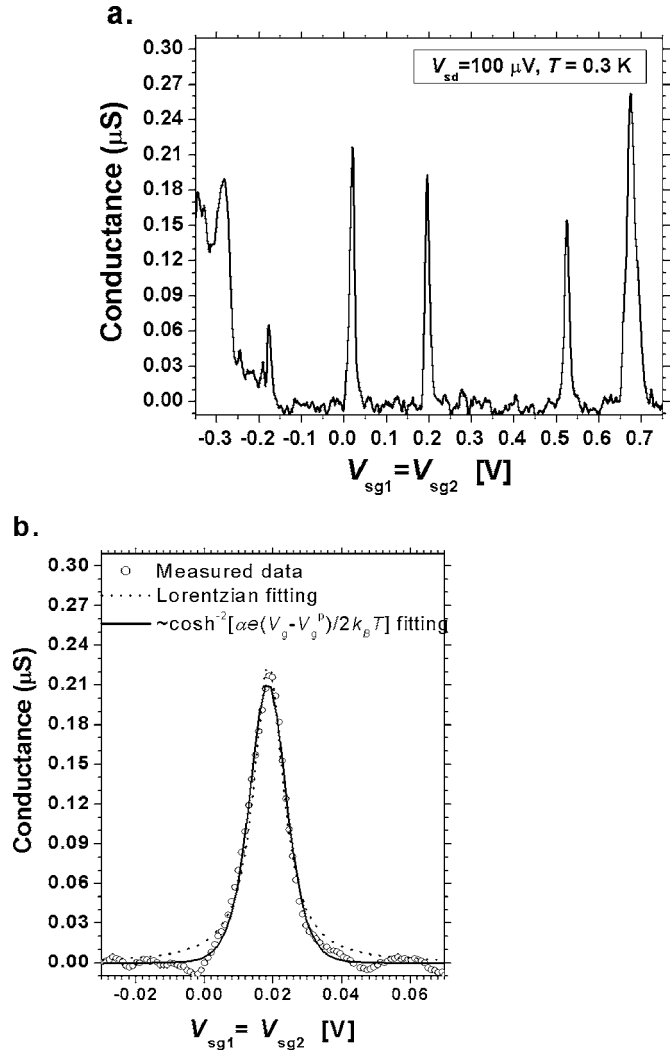

a.

b.

FIG. 4. (a) Dot conductance vs side-gate voltage at $T=0.3$ K for devices with epitaxial regrowth passivation. Drain-source bias is $\sim$0.1 mV and both side gates are connected together. (b) Conductance peak (scattered circles) near $V_g \sim 0.02$ V of (a) fitted to Lorentzian (dotted line) and a thermally broadened Coulomb blockade resonance (solid line).

$\sim$0.68 meV.[20] Within a single cooldown, the scans are reproducible over multiple scans in either direction for a period of many hours, in sharp contrast with the unpassivated SiGe (Ref. 13) and Si (Refs. 7–9) devices. The conductance peaks are better fitted by a Fermi-Dirac function $G \propto \cosh^{-2}[\alpha e(V_g - V_g^{peak})/2k_BT]$ than by a Lorentzian function [Fig. 4(b)], where V_g^{peak} is the gate voltage at which the conductance peak appears, and parameter $\alpha \equiv C_g/C_{total}$ relates the change of gate voltage to the energy level shifts in the dot relative to the source/drain Fermi energies.[21] Thus, the dot is in a weak coupling regime and the peak width is dominated by thermal broadening.

Coulomb blockade oscillations in the QD at different temperatures (0.36, 0.45, 0.72) all show conductance peaks at the same gate voltages [Fig. 5(a)]. The stability of the device allowed us to measure the differential conductance $g=\partial I_{ds}/\partial V_{ds}$ of the QD versus both gate and drain voltages at 0.3 K. The Coulomb blockade region with zero conductance shows the expected diamond shape [dark color in Fig. 5(b)].[15] From these diamonds we can measure the energy
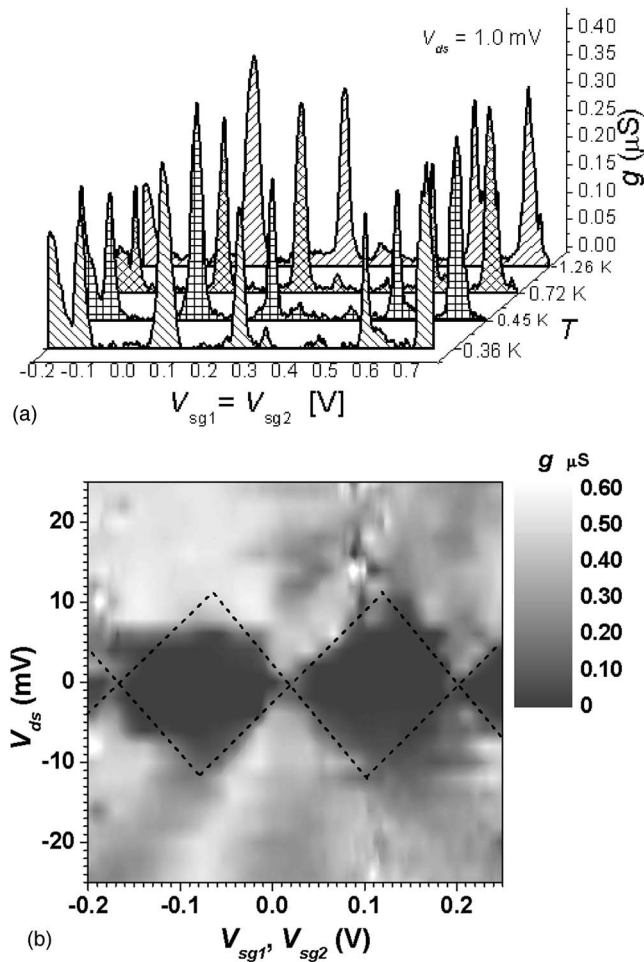
FIG. 5. (a) Conductance vs gate voltage (side gates connected together) at drain-source bias of 1.0 meV at 0.36, 0.45, 0.72, and 1.26 K for a passivated dot. (b) Differential conductance $\partial I/\partial V_{ds}$ on a linear gray scale as a function of side gate voltage (gates connected together) and drain-source bias. The dotted lines outlining the zero-conductance diamond are estimated by eye.

conversion factor $\alpha = 0.057$.[21] Assuming that the smallest peak spacing in 4(a) corresponds to filling of the same energy level with two holes of opposite spin, we estimate a charging energy of ~ 10 meV, consistent with electrostatic estimates for a 40 nm dot.[22] The single particle energy spacing can be deduced from larger peak spacings by subtracting the charging energy and is ~ 8 meV.

TOP GATE TUNING OF COULOMB BLOCKADE OSCILLATION

In practice, the current nanopatterning technique cannot identically process all quantum dot devices to exhibit oscillation peaks at exactly the same gate voltages. Slight differences in their fabricated dimensions will shift the oscillation peaks. Because of the multi-peaked nature of current in QD devices [compared to the monotonic nature of conventional field effect transistor (FET) curves], slight variations in the location of peaks can make circuit design more complex than for conventional MOSFET circuits. To control the position of the oscillation peaks when scanning the side-gate voltage, a top gate was added to the passivated SiGe QD device.
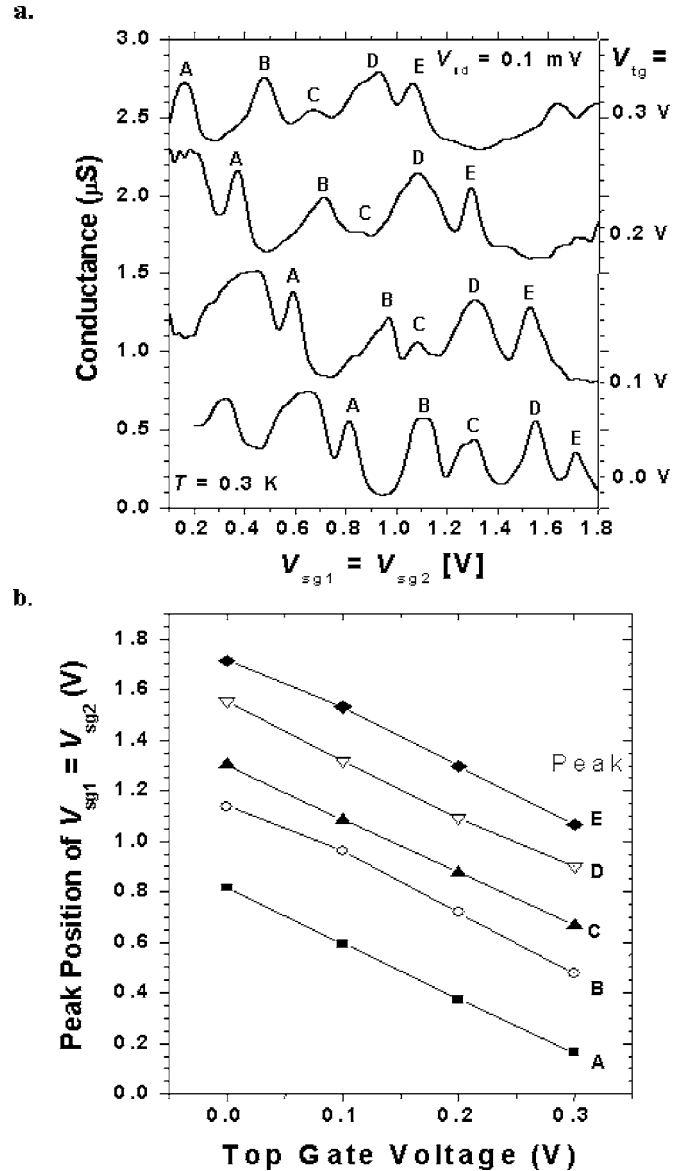

FIG. 6. (a) Conductance oscillations with side-gate voltage at different top gate voltages. The two side gates were held equal. The drain-source voltage was 0.1 mV and the temperature was 0.3 K. The individual curves are vertically offset for clarity. (b) Side-gate voltage peak positions for peaks A, B, C, D, and E vs top-gate voltage.

After the silicon epitaxial regrowth, tetraethylorthosilicate (TEOS) oxide was deposited at 575 °C and pressure of 600 mtorr in a low pressure chemical vapor deposition chamber. The oxide thickness is ~ 38 nm. Holes were opened in the oxide by conventional lithography and etching to allow subsequent aluminum, deposited by thermal evaporation, to contact the source, drain, and side gates. The aluminum was patterned by lift-off and subjected to a forming gas anneal. A second aluminum evaporation, again patterned by lift-off, was then done to create the top gate, which covered the whole dot. It is important to note that holes do not come into direct contact with the gate oxide (since they are confined in the SiGe, which is separated from the oxide by the epitaxial silicon), keeping the carriers in the device away from the oxide traps.

The desired use of the top gate to shift the oscillation

positions of the side gate is shown in Fig. 6. The dot conductance exhibits Coulomb blockade oscillations as a function of the double side-gate voltages, with the top gate voltage varying from 0.0 to 0.3 V at a step of 0.1 V. Positive top gate voltages tend to deplete holes in the dot as do the positive side-gate voltages. Therefore, the individual oscillation peaks labeled by A, B, C, D, and E shift to more negative side-gate voltages at each step increase of the top gate voltage to compensate for the effect of the top gate voltage. The voltage shift in side gate divided by the step of applied top-gate voltage is ∼2.2 [Fig. 6(b)]. Thus we infer that the top gate capacitance to the dot is twice that of the two side gates in parallel. Thus, the ability of a top gate to shift the side-gate characteristics to a desired peak position is clearly demonstrated. As an example, the side-gate-voltage position of peak B with zero volts of top gate voltage is tuned to that of peak E by applying 0.3 V to the top gate. Thus the effective tuning of the side-gate characteristics is up to three periods of oscillations [Fig. 6(b)].

Note that the oscillation peaks in Fig. 6(a) are broad and not perfectly periodic. Furthermore, the peak curve is well fitted with a Lorentzian function rather than a thermal broadening function $\cosh^{-2}[\alpha(V_g - V_g^{\text{peak}})e/2k_BT]$, which fit the oscillation peaks of the device without a top gate in Fig. 4. This indicates that the potential barriers between the source/drain and dot in the top gate device are smaller than in the device of Fig. 4, therefore leading to strong source-dot coupling. This may also explain why the conductance of the oscillation valley does not reach zero.[21] This lower confinement potential is probably due to a wider "narrow gap" region in the SiGe between source/drain and the dot, and is not related to the top gate itself. A 20 nm gap would lower the barrier to 2 meV (from ∼8 meV at 10 nm). In practice, the width of the narrow gap region was a very difficult parameter to control.

CONCLUSION

In summary, a SiGe single-hole transistor was fabricated by AFM nanolithography and silicon epitaxial regrowth. Very reproducible Coulomb blockade oscillations were observed, confirming a low number of parasitic states on the dot surfaces, which reflects the high quality interface between SiGe and regrown Si. The Coulomb blockade oscillations may be tuned by an additional top gate, enabling circuit flexibility. This work shows that epitaxial passivation of dot surfaces is an attractive approach for developing repeatable and stable silicon-based quantum dot devices.

[1]L. J. Guo, E. Leobandung, and S. Y. Chou, Science **275**, 649 (1997).
[2]Y. Takahashi, A. Fujiwara, Y. Ono, and H. Inokawa, Mater. Res. Soc. Symp. Proc. **686**, A7 (2002).
[3]M. Friesen, P. Rugheimer, D. E. Savage, M. G. Lagally, D. W. van der Weide, R. Joynt, and M. A. Eriksson, Phys. Rev. B **67**, 121301 (2003).
[4]H. Ishikuro and T. Hiramoto, Appl. Phys. Lett. **71**, 3691 (1997).
[5]M. Saitoh and T. Hiramoto, Appl. Phys. Lett. **84**, 3172 (2004).
[6]L. P. Rokhinson, L. J. Guo, S. Y. Chou, and D. C. Tsui, Appl. Phys. Lett. **76**, 1591 (2000).
[7]D. Ali and H. Ahmed, Appl. Phys. Lett. **64**, 2119 (1994).
[8]R. A. Smith and H. Ahmed, J. Appl. Phys. **81**, 2699 (1997).
[9]A. T. Tilke, F. C. Simmel, R. H. Blick, H. Lorenz, and J. P. Kotthaus, Prog. Quantum Electron. **25**, 97 (2001).
[10]D. J. Paul, J. R. A. Cleaver, and H. Ahmed, Appl. Phys. Lett. **63**, 631 (1993).
[11]S. Kanjanachuchai, J. M. Bonar, and H. Ahmed, Semicond. Sci. Technol. **14**, 1065 (1999).
[12]U. Dötsch, U. Gennser, C. David, G. Dehlinger, D. Grützmacher, T. Heinzel, S. Lüscher, and K. Ensslin, Appl. Phys. Lett. **78**, 341 (2001).
[13]X.-Z. Bo, L. P. Rokhinson, H. Yin, D. C. Tsui, and J. C. Sturm, Appl. Phys. Lett. **81**, 3263 (2002).
[14]L. J. Klein *et al.*, Appl. Phys. Lett. **84**, 4047 (2004).
[15]X. Z. Bo, L. P. Rokhinson, D. C. Tsui, and J. C. Sturm, Conference Digest 61st Device Research Conference, 2003, pp. 129–130.
[16]J. C. Sturm, P. V. Schwartz, E. J. Prinz, and H. Manoharan, J. Vac. Sci. Technol. B **9**, 2011 (1991).
[17]M. S. Carroll, L. D. Lanzerotti, and J. C. Sturm, Mater. Res. Soc. Symp. Proc. **527**, 417 (1998).
[18]M. S. Carroll, J. C. Sturm, and M. Yang, J. Electrochem. Soc. **147**, 4652 (2000).
[19]I. M. Ruzin, V. Chandrasekhar, E. I. Levin, and L. I. Glazman, Phys. Rev. B **45**, 13469 (1992).
[20]From the diamond slope in Fig. 5(b), we know $\alpha \equiv C_g/C_{\text{total}} = 0.057$.
[21]L. P. Kouwenhoven, C. M. Marcus, P. L. McEuen, S. Tarucha, R. M. Westervelt, and N. S. Wingreen, in *Proceedings of the Advanced Study Institute on Mesoscopic Electron Transport*, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schön (Kluwer, Dordrecht 1997).
[22]Assuming a dot shape of a flat disk with radius r, the Coulomb charging energy is $E_c = e^2/C = e^2/8\varepsilon r$, and the quantum energy level spacing is roughly $\Delta\varepsilon = \hbar^2/m^*r^2$, where ε is the dielectric constant, m^* is the effective mass, and $\hbar$ is Planck's constant. E_c is greater than $\Delta\varepsilon$ for a dot radius over 1 nm.