

Measurement of Carrier Mobility in Silicon Nanowires

Okı Gunawan,* Lidija Sekaric,* Amlan Majumdar, Michael Rooks,
Joerg Appenzeller, Jeffrey W. Sleight, Supratik Guha, and Wilfried Haensch

IBM Research Division, T. J. Watson Research Center,
Yorktown Heights, New York 10598

Received October 14, 2007; Revised Manuscript Received January 18, 2008

ABSTRACT

We report the first direct capacitance measurements of silicon nanowires (SiNWs) and the consequent determination of field carrier mobilities in undoped-channel SiNW field-effect transistors (FETs) at room temperature. We employ a two-FET method for accurate extraction of the intrinsic channel resistance and intrinsic channel capacitance of the SiNWs. The devices used in this study were fabricated using a top-down method to create SiNW FETs with up to 1000 wires in parallel for increasing the raw capacitance while maintaining excellent control on device dimensions and series resistance. We found that, compared with the universal mobility curves for bulk silicon, the electron and hole mobilities in nanowires are comparable to those of the surface orientation that offers a lower mobility.

Research on one-dimensional (1D) semiconducting structures such as nanowires (NWs)^{1–8} and carbon nanotubes (CNTs)^{9–15} is motivated, among other reasons, by device scaling trends, in particular, the need for improving the gate control over the transistor channel in short-channel field-effect transistors (FETs). One-dimensional structures offer the opportunity for employing the ultimate three-dimensional (3D) electrostatic control, the wrap-around gate. This lessens the need for gate dielectric thickness scaling. A 1D channel device is also a natural evolution of the channel width and thickness scaling in planar devices.¹⁶ Of particular interest are silicon nanowires (SiNWs) because of the potential ease of processing and large-scale integration when compared with other materials. Despite the widespread interest in 1D semiconducting nanostructures, there have been few successful attempts at experimentally determining the intrinsic transport properties of semiconductor NWs and CNTs. Most reported values of carrier mobility in NWs and CNTs have been estimated using measured channel resistance and calculated values of channel capacitance rather than directly measuring the channel capacitance.^{3–8,11–15} What makes direct determination of channel mobility in NWs or CNTs difficult is both the high series resistance and the ultralow capacitance typically in the sub-fF regime. Recently, successful measurement of NW and CNT capacitance has been reported, although the determination is at low temperatures (77 K in ref 17 and 150 K in ref 18) and therefore of limited relevance to room temperature device operation. In this paper, we report the first direct capacitance measurements of SiNWs and the

extraction of carrier mobility in undoped-channel SiNW FETs at room temperature.

NW fabrication can be divided into two classes: top-down and bottom-up fabrication. In the latter, NWs are grown on separate substrates, harvested, and then transferred to the target substrate. The top-down approach employs conventional CMOS fabrication techniques to build the NWs and thus has perfect control over the dimensions and location of the NWs on the target substrate. Although the bottom-up approach potentially offers advantages in material combination and choices, it lacks the positioning accuracy required for complex electrical circuits. Compatibility with conventional complementary metal-oxide-semiconductor (CMOS) processes and realistic circuit requirements were the deciding factor in choosing top down fabricated wires for this study.

Top-down fabrication also enables us to introduce the necessary elements to perform direct mobility measurements by (i) fabricating a large number of wires in parallel (up to 1000) for increased capacitance while maintaining excellent dimensional control and by (ii) introducing arbitrary changes in the FET gate length, accompanied by a tight control of the series resistance via precise doping and control of extension (the ungated regions) and contact geometry. We employ a two-FET method to accurately determine the channel resistance and gate capacitance necessary for mobility extraction.

The SiNWs described here and shown in Figure 1 were fabricated on 25 nm silicon-on-insulator (SOI) substrates using wafer-scale conventional processing used for planar CMOS devices. The wires are defined first using electron

* Corresponding authors. E-mail: ogunawa@us.ibm.com (O.G.), lidija@us.ibm.com (L.S.).

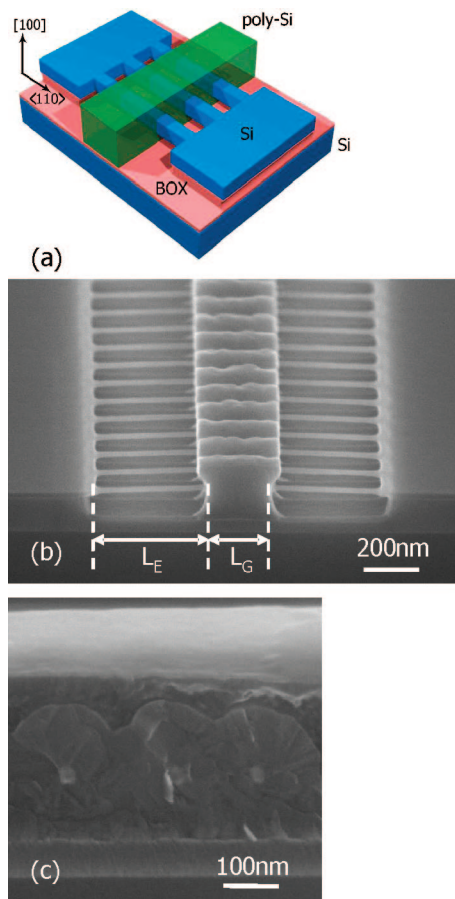


Figure 1. (a) Schematic of the final Si nanowire array FET structure on (100) SOI wafers with the wires oriented along the $\langle 110 \rangle$ direction. (b) A side-view scanning electron micrograph (SEM) of the final structure after gate formation. The gate length $L_G = 250$ nm and the extension region length $L_E = 375$ nm for this structure. (c) A cross-sectional SEM of the Si nanowires with wrap-around poly Si gate.

beam lithography in a negative resist. The wire widths for two sets of devices are 25 and 70 nm, and the wire lengths are 0.5 and 1 μm . Wire array structures ranging from 6 to 1000 SiNWs are connected to standard pad sets for electrical probing. Using resist as a mask, the exposed SOI layer is removed by reactive ion etching, followed by the immersion in HF for removal of the underlying oxide in order to suspend the SiNWs. After the undercut, 5 nm of thermal SiO_2 is grown around the wires, and then polycrystalline Si is deposited as the gate electrode. The gate oxidation step further thins down the wires to a thickness of 20 nm. Using resist as a mask, the gate is defined by selectively etching away the polycrystalline Si. The gate length is in the 300–800 nm range. Following a thin spacer deposition, the gate and SOI source/drain (S/D) contacts are doped by ion implantation. We use As or P for n-type FETs (NFETs), and B or BF_2 for p-type FETs (PFETs). Subsequently, the dopants are activated by rapid thermal anneal (RTA). Source/drain extension regions are formed by the diffusion of the dopants from the source/drain contact regions during RTA. The channel is protected by the gate material and remains undoped after RTA, which was confirmed using secondary

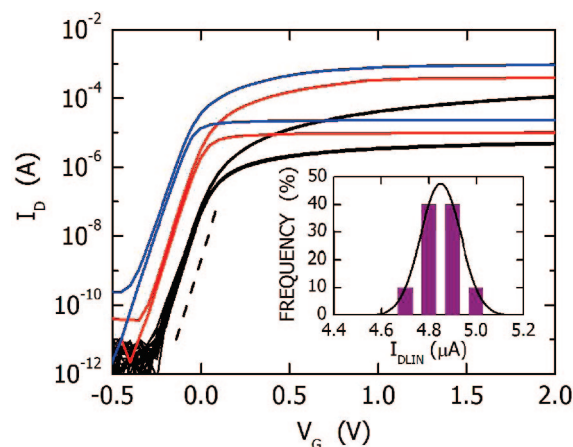


Figure 2. Drain current I_D vs gate voltage V_G of SiNW array NFETs with 9 (black), 100 (red), and 1000 (blue) wires. The nanowires in all three arrays have gate length $L_G = 300$ nm, height $H = 20$ nm, and width $W = 20$ nm, and were implanted with As to form source/drain contacts. For each device, I_D is measured at drain-to-source voltage $V_{DS} = 50$ mV (lower curve) and 2 V (upper curve). The dashed gray line denotes a subthreshold slope of 60 mV/decade. Inset: Histogram of linear on-current I_{DLIN} of 10 nominally identical 9-wire FETs with NW dimensions mentioned above. I_{DLIN} is defined as I_D at $V_G = 2$ V and $V_{DS} = 50$ mV. The solid line is a Gaussian fit that indicates a mean of 4.85 μA and net width of 0.17 μA ($\sim 4\%$).

ion mass spectroscopy (SIMS) performed on blanket wafers with equivalent films that went through the same implantation and RTA. Finally, we form nickel–platinum silicide contacts to the source/drain and gate electrodes. The wafers are also subjected to forming gas anneal after silicidation to improve the quality of the gate oxide. The device yield is 100% (of 420 devices per wafer), showing the robustness of the process.

SiNW FET electrical current–voltage measurements show the expected FET behavior. Figure 2 shows the postsilicidation drain current I_D versus gate voltage V_G characteristics of SiNW NFETs, measured at drain-to-source bias $V_{DS} = 50$ mV and 2 V. The SiNW FETs have 9, 100, and 1000 wires in parallel. The SiNWs in all three arrays have gate length $L_G = 300$ nm, wire height $H = 20$ nm, and width $W = 20$ nm. The deep S/D implant was As. We observe excellent long-channel device characteristics with a subthreshold slope of ~ 60 mV/decade, which is the theoretical limit at room temperature. This indicates that the gate oxide is of high quality and with a negligible interface state density (which might give rise to interface state capacitance and therefore affects capacitive determination of channel charge). It is also clear that the subthreshold current for $V_G < -0.1$ V scales with the number of wires. This is expected because, in the sub-threshold regime, the channel resistance far exceeds the series resistance, R_{EXT} , which is defined as the net additional resistance offered by the device excluding the intrinsic channel resistance. On the other hand, we find that the on-state current for $V_G > 0.5$ V is severely limited by R_{EXT} and, as a result, while the current increases with the number of wires, it does not scale proportionately. In the inset of Figure 2, we plot a histogram of the linear on-current

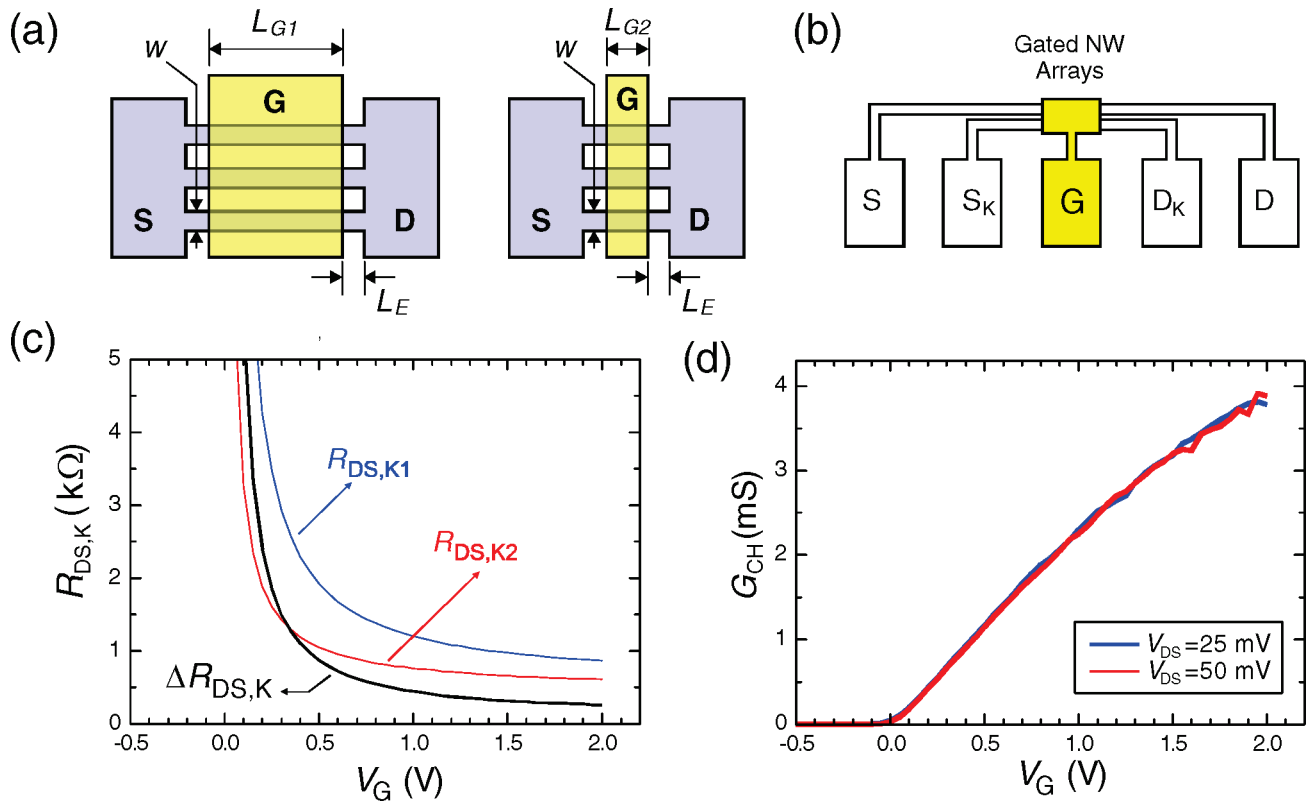


Figure 3. (a) Schematic of devices used for two-FET method for extracting mobility. The two FETs have different gate lengths L_{G1} and L_{G2} but identical number of wires N , wire width W , wire height H , and length of extension region L_E . (b) Layout of individual FETs employed for mobility extraction. Each of the FETs are laid out in the Kelvin configuration with drain D, drain voltage probe D_K , source S, source voltage probe S_K , and gate G. (c) Kelvin resistance, $R_{DS,K1}$ (blue) and $R_{DS,K2}$ (red), of two NFETs with $L_{G1} = 0.8 \mu\text{m}$ and $L_{G2} = 0.3 \mu\text{m}$, respectively, vs gate voltage V_G . Each FET has $N = 100$, $W = 20 \text{ nm}$, $H = 20 \text{ nm}$, and $L_E = 0.1 \mu\text{m}$, and was implanted with As to form source/drain contacts. The difference in the Kelvin resistances $\Delta R_{DS,K} = R_{DS,K1} - R_{DS,K2}$ (black) represents the intrinsic channel resistance of a $\Delta L_G = L_{G1} - L_{G2} = 0.5 \mu\text{m}$ long channel. (d) Intrinsic channel conductance $G_{CH} = 1/\Delta R_{DS,K}$ vs V_G obtained from the two-FET method applied to the devices shown in (c). G_{CH} was extracted at drain bias $V_{DS} = 25$ and 50 mV . The overlapping of the two curves in (d) indicates that the measurements are done in the linear ohmic regime.

I_{DLIN} , which is defined as I_D at $V_G = 2 \text{ V}$ and $V_{DS} = 50 \text{ mV}$, of 10 nominally identical 9-wire FETs with $L_G = 300 \text{ nm}$, $H = 20 \text{ nm}$, and $W = 20 \text{ nm}$. The histogram shows that the I_{DLIN} distribution has a spread of $\sim 4\%$. Since I_{DLIN} is limited by R_{EXT} , the $\sim 4\%$ spread reflects of the width of the R_{EXT} distribution.

The accurate extraction of carrier mobility for any system depends on both precise measurement of intrinsic channel resistance (or conductance) and channel charge. Channel charge can be inferred from the Hall resistance or by measuring the channel capacitance.^{19,20} Since the Hall resistance is measured in the presence of a perpendicular magnetic field, the method cannot be applied to nonplanar device geometries. Therefore, capacitive determination of charge is the only known method for nonplanar FETs, including NW FETs. Since individual NWs have ultralow capacitance, a large number of NWs must be placed in parallel to increase the device capacitance into the fF range, where it is measurable with good accuracy. The measurement of intrinsic channel resistance of NWs is also challenging. For two-dimensional, planar FETs or even ultrathin body SOI FETs, the intrinsic channel resistance can be determined using long-channel Kelvin FETs with voltage probes under-

neath the gate.²¹ However, for a NW array FET, placing voltage probes underneath the gate is complicated.

In light of the above-mentioned difficulties with conventional methods, we accurately measure intrinsic channel resistance and capacitance using a two-FET method, which provides a simple approach for eliminating series resistance and parasitic capacitance. Figure 3 summarizes our approach. We use two FETs with identical width W , number of wires N , and extension length L_E but different gate lengths, L_{G1} and L_{G2} (Figure 3a). The FET layout is in a four-point Kelvin configuration²² but with the voltage probes outside of the gated NW region (Figure 3b). For the SiNW devices, we have used for mobility extraction, $N = 100$, $L_E = 0.1 \mu\text{m}$, $L_{G1} = 0.8 \mu\text{m}$, and $L_{G2} = 0.3 \mu\text{m}$. The difference of resistance and capacitance of these two FETs is attributed to an intrinsic FET with channel length $\Delta L_G = L_{G1} - L_{G2} = 0.5 \mu\text{m}$. For this two-FET method to be valid, the following conditions must be satisfied: the two FETs must have the same (i) series resistance R_{EXT} , (ii) electrical width $W_E = 2N \times (W + H)$, and (iii) mobility μ that is independent of channel length. Deviations from the above conditions would lead to error in the extracted mobility. We have earlier deduced a 4% spread in the width of the R_{EXT} distribution of nominally identical FETs. Taking the worst-case scenario, a simple analysis

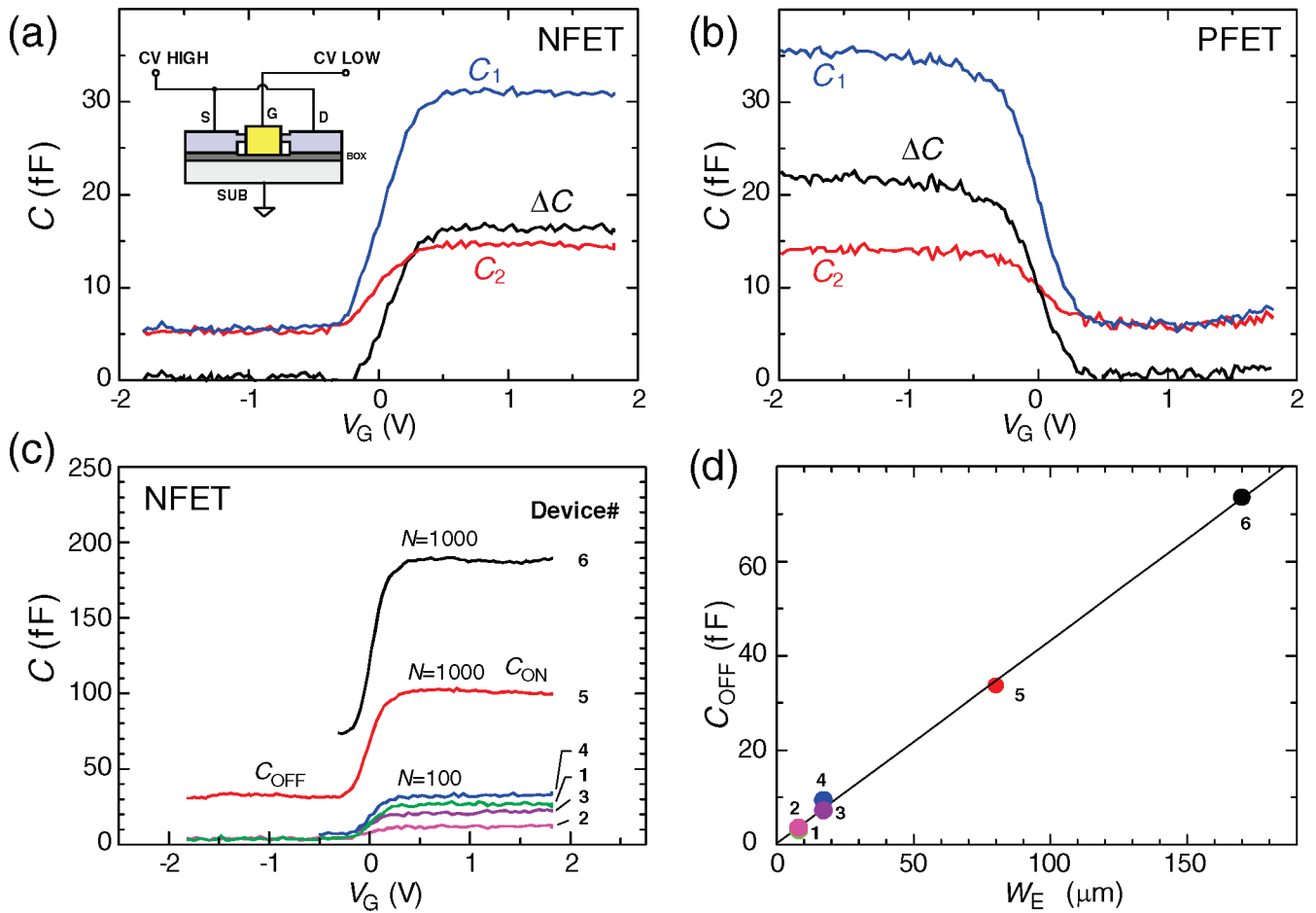


Figure 4. (a) Capacitance, C_1 (blue) and C_2 (red), of two Si nanowire (SiNW) NFETs with $L_{G1} = 0.8 \mu\text{m}$ and $L_{G2} = 0.3 \mu\text{m}$, respectively, vs gate voltage V_G . Each FET has number of wires $N = 100$, wire width $W = 20 \text{ nm}$, wire height $H = 20 \text{ nm}$, and extension length $L_E = 0.1 \mu\text{m}$, and was implanted with As to form source/drain contacts. The capacitance difference $\Delta C = C_1 - C_2$ (black) represents the intrinsic channel capacitance of a $\Delta L_G = L_{G1} - L_{G2} = 0.5 \mu\text{m}$ long channel. Inset: Split capacitance–voltage (C – V) measurement configuration used for measuring channel capacitance. (b) C – V data for PFETs similar to (a). (c) C – V curves of SiNW arrays with $L_G/W/N = 800 \text{ nm}/20 \text{ nm}/100$, $300 \text{ nm}/20 \text{ nm}/100$, $500 \text{ nm}/65 \text{ nm}/100$, $300 \text{ nm}/20 \text{ nm}/1000$, and $300 \text{ nm}/65 \text{ nm}/1000$ for devices labeled 1 through 6. (d) Off-state capacitance C_{OFF} vs electrical width $W_E = 2N \times (W + H)$ for the devices shown (c). A straight-line fit to the data indicates a nearly zero intercept, indicating that all the wires are functional and that our estimate of the wire widths W from cross-sectional scanning electron microscope (SEM) images is correct.

shows that this would lead to an error of $\sim 5\%$ in the mobility. We discuss unequal electrical width as a source of error later in this paper.

Figure 3c shows the Kelvin probe resistance of the two FETs measured by applying $V_{\text{DS}} = 50 \text{ mV}$. The difference of the resistance of the two FETs (and therefore the intrinsic channel resistance), denoted by $\Delta R_{\text{DS,K}}$ and attributed to $\Delta L_G = 0.5 \mu\text{m}$, is also shown in Figure 3b. From these measurements, we deduce a series resistance $R_{\text{EXT}} \sim 500 \Omega$, which is significant and must be carefully subtracted out for accurate mobility extraction. This R_{EXT} is deduced from the measurement of $R_{\text{DS,K}}$ of the two FETs by solving the corresponding linear equations assuming a constant R_{EXT} and intrinsic channel resistance that scales with the gate length. Finally, the intrinsic channel conductance G_{CH} is obtained from $G_{\text{CH}} = 1/\Delta R_{\text{DS,K}}$. Figure 3d shows G_{CH} versus V_G obtained from the two-FET method at applied drain bias V_{DS} in the 25–50 mV range. It is clear that the channel conductance is measured in the linear ohmic regime for both FETs.

In order to determine the inversion charge, capacitance–voltage (C – V) is measured in the split- C – V configuration under dark conditions.²⁰ The C – V measurement configuration is detailed in the inset of Figure 4a. An Andeen–Hagerling 2700A ultra precision capacitance bridge is used in a room-temperature probe station with 200 mV excitation signal at a frequency of 1 kHz. This excitation signal is optimized to reduce the noise without causing significant deviations in the capacitance data compared to the measurement with smaller excitation signals.

Figure 4a shows the measured C – V curves of the two FETs with 0.3 and $0.8 \mu\text{m}$ gate lengths for NFETs. The on-state capacitance C_{ON} , defined at $V_G > 1 \text{ V}$, is different for the two FETs because of the difference in gate lengths. The difference of the two C – V curves, denoted by ΔC , is also shown in Figure 4a and represents the intrinsic channel capacitance from $\Delta L_G = 0.5 \mu\text{m}$. We obtain the inversion charge, Q_{INV} , using $Q_{\text{INV}} = \int \Delta C_{\text{CH}} dV_G$. In Figure 4a, we observe that the off-state capacitance C_{OFF} , defined at $V_G < -0.2 \text{ V}$, is practically the same for both FETs. Direct overlap,

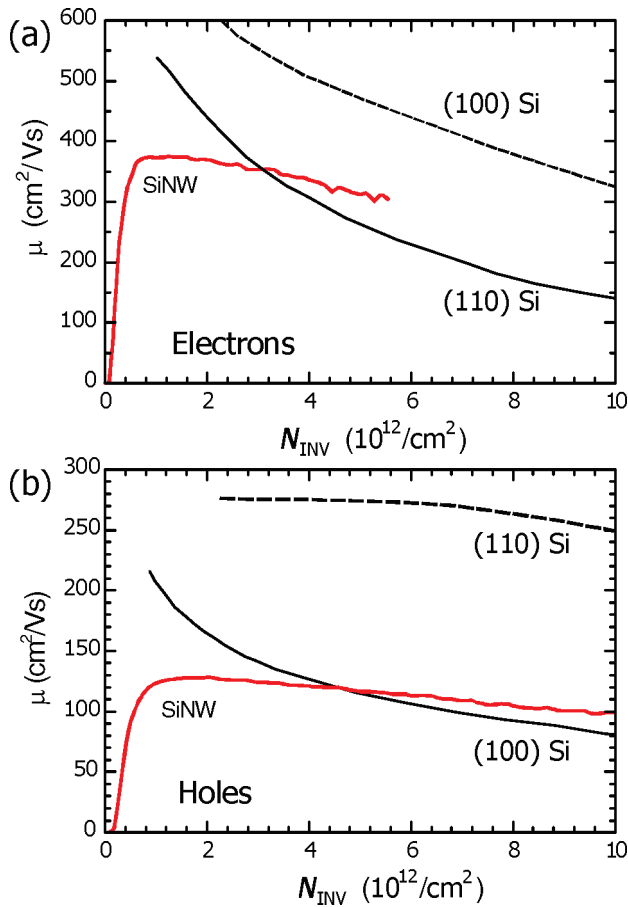


Figure 5. Comparison of (a) electron and (b) hole mobility μ vs inversion carrier density N_{INV} in Si nanowire FETs with Si universal mobility data.^{23–25} The height H and width W of the Si nanowires are both 20 nm. Since our SiNW FETs have undoped channels, we convert effective vertical field E_{EFF} into N_{INV} by setting $N_{\text{DEP}} = 0$ in $E_{\text{EFF}} = (q/\epsilon_{\text{Si}})(N_{\text{DEP}} + \eta N_{\text{INV}})$. We use $\eta = 1/2$ for (100) electrons and $1/3$ for (110) electrons, (100) holes, and (110) holes.

inner fringe, and outer fringe capacitances contribute to C_{OFF} and are proportional to the electrical width W_{E} . The same value of C_{OFF} of the two FETs is an indication that these two FETs have indeed identical electrical widths and, therefore, satisfy condition ii for the validity of the two-FET method. Similar observation is obtained for PFETs as shown in Figure 4b. We plot the C – V curves measured on NW arrays with $N = 100$ and 1000 , and $W = 20$ and 65 nm in Figure 4c. It is clear that NW arrays with larger NW width or larger number of NWs have larger C_{OFF} . In Figure 4d, we plot C_{OFF} versus W_{E} for the devices shown in Figure 4c. A straight-line fit to the data indicates a nearly zero intercept, showing that all of the wires in the array are functional and that our estimate of the wire widths W from cross-sectional scanning electron microscope (SEM) images is correct.

By having deduced the intrinsic channel conductance G_{CH} and inversion charge Q_{INV} , the carrier mobility μ is obtained from the relation $\mu = G_{\text{CH}} \Delta L_{\text{G}}^2 / Q_{\text{INV}}$. A plot of electron and hole mobility versus inversion charge density is shown in Figure 5, with the inversion charge density obtained from $N_{\text{INV}} = Q_{\text{INV}} / (W_{\text{E}} \Delta L_{\text{G}})$. The data is compared against the electron and hole universal mobility curves for both (100)

Si and (110) Si in Figure 5.^{23–25} Since these SiNW FETs have undoped channels, N_{INV} is found from effective vertical field E_{EFF} (at the semiconductor–oxide interface) by setting $N_{\text{DEP}} = 0$ in $E_{\text{EFF}} = q/\epsilon_{\text{Si}} \times (N_{\text{DEP}} + \eta N_{\text{INV}})$, where q is the charge of an electron, ϵ_{Si} is the permittivity of Si, and N_{DEP} is the areal density of the depletion charge. We use η , that represents the fractional contribution of N_{INV} to E_{EFF} , to be $1/2$ for (100) electrons and $1/3$ for (110) electrons, (100) holes, and (110) holes. The NWs are bound by a combination of roughly a comparable admixture of (100) and (110) Si surfaces. Assuming that the devices operate in the surface inversion regime at high carrier densities, the effective carrier mobilities in the SiNWs may be expected to lie between the (100) and the (110) mobility curves on planar Si inversion layers. Instead, the data indicate that the SiNW carrier mobilities are comparable to those of the lower mobility surface in planar Si inversion layers, that is, (110) Si for electrons and (100) Si for holes. This may be explained by noting that the carriers predominantly appear to “feel” the lower mobility surface potential, when two competing surfaces are presented to the carrier at distances of the order of the inversion layer depth. It is also difficult to resolve the effects of strain across the wire direction²⁶ which can be complicated in wire geometry.

In conclusion, we have performed the first direct capacitance measurements of SiNWs at room temperature and demonstrated a method for determining accurate electron and hole mobilities in undoped-channel SiNW FETs. A two-FET method is employed to measure the intrinsic channel resistance and capacitance using SiNW array FETs fabricated via a top-down method. Using this method, we have compared the measured carrier mobilities with the universal mobility curves for bulk silicon and have shown that the NW carrier mobilities appear to be closer to those of the surface orientation that offers a lower mobility for the specific carrier type.

Acknowledgment. We thank Jim Bucchignano and the staff of the Microelectronics Research Laboratory for technical assistance with device fabrication.

References

- (1) Appenzeller, J.; Knoch, J.; Tutuc, E.; Reuter, M.; Guha, S. *IEDM Tech. Dig.* **2006**, 53, 555–558.
- (2) Yeo, K. H.; Suk, S. D.; Li, M.; Yeo, Y.; Cho, K. H.; Hong, K.-H.; Yun, S. K.; Lee, M. S.; Cho, N.; Lee, K.; Hwang, D.; Park, B.; Kim, D.-W.; Ryu, B.-I. *IEDM Tech. Dig.* **2006**, 53, 539–542.
- (3) Lauhon, L. J.; Gudiksen, M. S.; Wang, D.; Lieber, C. M. *Nature* **2002**, 420, 57–61.
- (4) Zhang, D.; Li, C.; Han, S.; Liu, X.; Tang, T.; Jin, W.; Zhou, C. *Appl. Phys. Lett.* **2003**, 82, 112–114.
- (5) Greytak, A. B.; Lauhon, L. J.; Gudiksen, M. S.; Lieber, C. M. *Appl. Phys. Lett.* **2004**, 84, 4176–4178.
- (6) Wernersson, L. E.; Bryllert, T.; Lind, E.; Samuelson, L. *IEDM Tech. Dig.* **2005**, 52, 265–268.
- (7) Singh, N.; Agarwal, A.; Bera, L. K.; Liow, T. Y.; Yang, R.; Rustagi, S. C.; Tung, C. H.; Kumar, R.; Lo, G. Q.; Balasubramanian, N.; Kwong, D.-L. *IEEE Elec. Dev. Lett.* **2006**, 27 (5), 383–386.
- (8) Motayed, A.; Vaudin, M.; Davydov, A. V.; Melngailis, J.; He, M.; Mohammad, S. N. *Appl. Phys. Lett.* **2007**, 90, 043104 1–3.
- (9) Martel, R.; Schmidt, T.; Shea, H. R.; Hertel, T.; Avouris, Ph. *Appl. Phys. Lett.* **1998**, 73, 2447–2449.
- (10) Tans, S. J.; Verschueren, A. R. M.; Dekker, C. *Nature* **1998**, 393, 49–52.

- (11) Javey, A.; Kim, H.; Brink, M.; Wang, Q.; Ural, A.; Guo, J.; McIntyre, P.; McEuen, P.; Lundstrom, M.; Dai, H. *Nat. Mater.* **2002**, *1*, 241–246.
- (12) Javey, A.; Guo, J.; Wang, Q.; Lundstrom, M.; Dai, H. *Nature* **2003**, *424*, 654–657.
- (13) Durkop, T.; Getty, S. A.; Cobas, E.; Fuhrer, M. S. *Nano Lett.* **2004**, *4*, 35–39.
- (14) Li, S.; Yu, Z.; Rutherglen, C.; Burke, P. J. *Nano Lett.* **2004**, *4*, 2003–2007.
- (15) Zhou, X.; Park, J.-Y.; Huang, S.; Liu, J.; McEuen, P. L. *Phys. Rev. Lett.* **2005**, *95*, 146805 1–4.
- (16) Haensch, W.; Nowak, E. J.; Dennard, R. H.; Solomon, P. M.; Bryant, A.; Dokumaci, O. H.; Kumar, A.; Wang, X.; Johnson, J. B.; Fischetti, M. V. *IBM J. Res. & Dev.* **2006**, *53*, 339–361.
- (17) Ilani, S.; Donev, L. A. K.; Kindermann, A.; McEuen, P. L. *Nat. Phys.* **2006**, *2*, 687–691.
- (18) Tu, R.; Zhang, L.; Nishi, Y.; Dai, H. *Nano Lett.* **2007**, *6*, 1561–1565.
- (19) Fowler, A. B.; Fang, F.; Hochberg, F. *IBM Journal* **1964**, *8*, 427–429.
- (20) Sodini, C. G.; Ekstedt, T. W.; Moll, J. L. *Solid-State Electron.* **1982**, *25*, 833–841.
- (21) Uchida, K.; Takagi, S. *Appl. Phys. Lett.* **2003**, *82*, 2916–2918.
- (22) Schroder, D. K. *Semiconductor Material and Device Characterization*; John Wiley & Sons: New York, 1990.
- (23) Takagi, S.; Toriumi, A.; Iwase, M.; Tango, H. *IEEE. Trans. Elec. Dev.* **1994**, *41*, 2357–2362.
- (24) Takagi, S.; Toriumi, A.; Iwase, M.; Tango, H. *IEEE. Trans. Elec. Dev.* **1994**, *41*, 2363–2368.
- (25) Irie, H.; Kita, K.; Kyuno, K.; Toriumi, A. *IEDM Tech. Dig.* **2004**, *51*, 225–228.
- (26) He, R.; Yang, P. *Nat. Nanotechnol.* **2006**, *1*, 42–46.

NL072646W