Quantum Confinement Induced Performance Enhancement in Sub-5-nm Lithographic Si Nanowire Transistors

Krutarth Trivedi,† Hyungsang Yuk,‡ Herman Carlo Floresca,‡ Moon J. Kim,‡ and Walter Hu*†‡

†Department of Electrical Engineering and ‡Department of Material Science and Engineering, University of Texas at Dallas, Richardson, Texas 75080, United States

ABSTRACT: We demonstrate lithographically fabricated Si nanowire field effect transistors (FETs) with long Si nanowires of tiny cross sectional size (≈3–5 nm) exhibiting high performance without employing complementarily doped junctions or high channel doping. These nanowire FETs show high peak hole mobility (as high as over 1200 cm²/(V s)), current density, and drive current as well as low drain leakage current and high on/off ratio. Comparison of nanowire FETs with nanobelt FETs shows enhanced performance as a result of significant quantum confinement in these 3-5 nm wires. This study suggests simple (no additional doping) FETs using tiny top-down nanowires can deliver high performance for potential impact on both CMOS scaling and emerging applications such as biosensing.

KEYWORDS: Nanowire, nanobelt, quantum confinement, mobility, FETs

Following Feynman’s visionary glimpse into manufacturing of the nanorealm in 1960,† there has been sustained effort to reach atomic scales in both materials and devices. Now, low dimensional structures such as nanowires and quantum dots of a variety of materials are extensively investigated for their unique properties. Silicon nanowires, in particular, have generated much interest as one-dimensional structures, owing to relatively low material cost and maturity of Si processing and integration techniques. Si nanowire field effect transistors (NWFETs) are emerging as important devices, poised to play a significant role both in extending performance of scaled CMOS and in applications such as biosensing. A popular type of Si NWFET, known as a FinFETs, is typically fabricated by lithography in a top-gate configuration with highly doped junctions and a short channel (≈50–300 nm).6–7 Along this line, recent progress has demonstrated Si NWFETs with ≈10 nm or smaller nanowires using gate-all-around8,9 and multigate or stacked architecture.10 In contrast to devices for CMOS application, biosensing FET devices typically employ a back-gated configuration with longer nanowires. Sensitivity of nanowire biosensors is inversely related to nanowire width, and hence, low dimensional silicon nanowires could allow ultrahigh sensitivity.11 For long nanowire devices, bottom-up methods such as vapor—liquid—solid or solution—liquid—solid (VLS/SLS) growth12,13 have emerged as easily accessible routes to produce small nanowires (diameter down to a few nanometers).12,13 Si nanowires, made with such growth methods, have been used to assemble NWFETs with high performance.14 However, intrinsic limitations of bottom-up methods including substantial device integration challenges13 have made adoption of bottom-up techniques for practical manufacturable devices difficult. These issues can be mitigated with a top-down approach. However, lithographically fabricated back-gated NWFETs with long nanowires of dimensions comparable to grown nanowires have not been demonstrated.

Equally important is investigation of effects of varying degrees of channel confinement on device performance at normal operating conditions. As nanowire size decreases, quantum confinement causes nanowire properties to change from their bulk values. Simulations have shown band structure15,16 and mobility17–19 to be very sensitive to carrier and phonon confinement in sub-10-nm Si nanowires, even at room temperature. Numerous studies with Si NWFETs of sub-50-nm nanowires report significant confinement induced effects only at ultralow temperature or drain bias6,9,20,21 because increase in carrier energy, by increasing temperature or drain bias, greatly reduces effects of carrier confinement in NWFETs with relatively large nanowires. Confinement effects, such as theoretically expected enhancement of mobility,17,19,22 have not been observed under normal operating conditions which, for most applications, are room temperature and sufficient drain bias (millivolt range). Consequently, it is not clear if nanowire FETs offer performance benefits over larger geometries. Here, we demonstrate simple lithographic fabrication of high-performance Si nanowire P-MOSFETs without complementarily doped junctions or high channel doping. We study NWFETs with channel length on the order of micrometers and width less than 5 nm, similar to grown nanowires. Performance characteristics of sub-5-nm NWFETs are compared to wider nanobelt

Received: September 16, 2010
Revised: January 15, 2011
Published: March 04, 2011
FETs to investigate effects of increasing confinement (2D vs 1D confinement). High-resolution transmission electron microscopy (HR-TEM) is used to verify dimensions of nanowires and nanobelts in measured devices. Results show higher mobility and current density in all nanowire P-MOSFETs, indicating strong 2D quantum confinement induced enhancement of hole mobility. Our results present experimental evidence supporting theoretical calculations of enhanced hole mobility in sufficiently confined nanowires. Moreover, scaled performance of NWFETs shown here exceeds that of NWFETs reported in recent studies. This study shows the viability of long channel, sub-5-nm Si NWFETs as high-performance transistors for emerging applications such as nanowire biosensors.

Single nanowire and nanobelt Si P-MOSFETs were made on an ultrathin p-type (boron, $2 \times 10^{15}$ cm$^{-3}$) silicon-on-insulator (SOI) substrate with a 10 nm top Si layer (001) and 150 nm buried oxide (BOX) layer (LG Siltron, S. Korea). A schematic of typical devices is shown in Figure 1. Electron beam lithography (EBL) with hydrogen silsesquioxane (HSQ) resist and inductively coupled plasma etching were used to make nanowires ($\sim 12$ nm) and wider nanobelts ($\sim 175$ nm) in the [110] channel direction between Si pads separated by 3, 10, and 20 nm. The devices were oxidized for 10 min at 900 °C in O$_2$ to reduce the size of etched Si nanowires and nanobelts. After oxidation, the devices immediately underwent annealing in nitrogen for an improved Si–oxide interface. This process differs from stress limited oxidation, in which a significantly larger Si nanowire is oxidized for several hours, typically resulting in two Si cores of arbitrary smaller sizes, one of which is carefully etched away by plasma while the other is released from oxide by wet etching.

Unlike stress limited oxidation, which is time-consuming and difficult to control, our process, starting with small nanowires on ultrathin SOI, reliably produces nanowires with sub-5-nm cross sectional dimensions in a fraction of the time and without excessive processing. A back-gate configuration was implemented, as the original motivation for this study was to develop lithographically defined long channel Si NWFETs for biosensing. Pt (200 nm) and Al (250 nm) were deposited on Si source/drain pads and backside Si, respectively. Unlike conventional MOSFETs, no extra doping beyond background doping level is used in the source/drain regions or the nanowire/nanobelt channels. Low channel doping is also beneficial in biosensing, as sensitivity is expected to increase with decreasing dopant concentration. Finally, devices were annealed in forming gas at 400 °C to improve metal–semiconductor contacts and reduce fixed charge in grown oxide on the surface of the nanowires and nanobelts. This process allows for simple fabrication of nanowire transistors.
shown in Figure 2a, plasma etched Si nanowires do not show significant sidewall roughness in our process. Growing thermal oxide on plasma etched nanowires evidently reduces sidewall roughness and rounds sharp corners, producing a cross section that closely resembles grown nanowires. Therefore, in addition to reducing trap density at the Si–oxide interface, thermal oxidation plays a key role in reducing surface roughness scattering. As seen in Figure 1, nanowires, being confined in both (001) (top) and (−110) (sides), are 1D structures. Nanobelts, being confined only in (001) (top), are 2D structures. Oxidation resulted in nanowires of width between 3.4 and 5.7 nm and nanobelts of ∼169 nm. Both nanowires and nanobelts have a height of 4.3 nm. Variation in nanowire size originates from line edge roughness and possibly variations in resist thickness and beam focus over large distances during e-beam lithography.

All devices were measured at room temperature with a Keithly 4200 semiconductor characterization system and a shielded probe station with triax connectors to minimize noise. Before each measurement, all devices underwent gate voltage cycling (between −5 and +5 V) to minimize hysteresis effects. Transfer characteristics of typical P-MOSFETs with a nanowire or nanobelt are shown in Figure 3a. Both types of devices exhibit p-type transfer characteristics with $I_{ON}/I_{OFF}$ greater than $10^5$, SS less than 100 mV/decade, $I_{OFF}$ lower than the measurement limit ($\sim 10^{-15}$ A), and threshold voltage ($V_T$) of less than 2 V. NWFETs with channel length of 3 μm or less exhibit drain current oscillation. Drain current oscillation is a byproduct of quantum confinement and is typically observed only at ultralow temperature and drain bias. Theoretically, it may be possible to observe such oscillation at room temperature if the nanowire cross sectional area is sufficiently small. The oscillation observed in our NWFETs shows a strong dependence on channel length, with oscillation subsiding as channel length increases. As seen in Figure 3a, longer channel NWFETs do not exhibit drain current oscillation. Even for shorter NWFETs, such oscillation can be reduced by increasing drain bias.

Both nanowire and nanobelt devices behave similar to conventional enhancement mode P-MOSFETs operating in accumulation. Despite the lack of complementarily doped current blocking junctions between source/drain and channel, $I_{OFF}$ in both nanowire and nanobelt FETs is less than $10^{-14}$ A. A likely due to a barrier of more than 0.2 eV, given the high work function of Pt. These devices exhibit high drive current ($I_{ON}$), even with small geometry and no additional channel doping. Unlike in conventional MOSFETs with higher channel doping and larger geometry, appreciable accumulation of carriers in our devices is not probable, particularly in devices with sub-5-nm nanowires. Considering wafer doping density of $\sim 2 \times 10^{15}$ cm$^{-2}$, even long 20 μm length nanowires and nanobelts in our devices would statistically contain only $\sim 0.7$ and $\sim 29$ dopant atoms, respectively, indicating a severe shortage of carriers in the channel region. Therefore, unlike conventional MOSFETs in which carriers flow through thin highly conductive accumulation or inversion channels at the gate oxide–Si interface, holes in our NWFETs likely flow through fully depleted nanowires. Unlike in the tiny nanobelts and nanowires, holes can accumulate at the BOX/Si interface in the larger source/drain pads with applied back-gate bias. As back-gate bias is increased, the holes accumulated in the source/drain pads find it energetically favorable to flow through the channel. The exact mechanism by which the channel conductivity is modulated remains unclear and warrants further investigation.

**Figure 3.** (a) Transfer characteristics of typical single nanowire and nanobelt P-MOSFETs ($V_{DS} = 50$ mV). (b) Peak hole mobility ($\mu_{FE}$) of single nanowire and nanobelt devices. Nanowire height is 4.3 nm for all nanowires and nanobelts, while nanowire width ranges from 3.4 to 5.7 nm, with smaller nanowires exhibiting higher mobility. (c) Current density through typical nanowire and nanobelt P-MOSFETs.

without any additional doping beyond wafer background level, which is known to require increasingly complicated and costly techniques as CMOS devices continue to be scaled to smaller dimensions.

Certain representative devices (e.g., a nanowire and nanobelt device of each channel length) were cross sectioned with focused ion beam and imaged by HR-TEM after electrical characterization. Figure 2a shows an SEM image of an EBL-defined Si nanowire on BOX as well as a HR-TEM image of the same nanowire before oxidation. Panels b and c of Figure 2 show cross sectional HR-TEM images of an oxidized nanowire and nanobelt from two typical devices, respectively. The cross sections in panels b and c of Figure 2 show single crystalline Si with a smooth and well-defined interface with oxide. It has been suggested that wet etching Si nanowires produces superior Si NWFETs, with higher sub-threshold slope (SS) due to smoothness of sidewalls.28 As
where higher performance of our devices.

Table 1. Comparison of Recent Nanowire Field Effect Transistor Devices

<table>
<thead>
<tr>
<th>Si nanowire (this study)</th>
<th>Si nanowire</th>
<th>SiGe nanowire</th>
</tr>
</thead>
<tbody>
<tr>
<td>carrier type</td>
<td>P</td>
<td>p&lt;sup&gt;28&lt;/sup&gt;</td>
</tr>
<tr>
<td>gate type</td>
<td>back gate</td>
<td>GAA</td>
</tr>
<tr>
<td>channel width (nm)</td>
<td>3.4</td>
<td>30.5</td>
</tr>
<tr>
<td>channel height (nm)</td>
<td>4.3</td>
<td>10</td>
</tr>
<tr>
<td>channel length (μm)</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>V&lt;sub&gt;ds&lt;/sub&gt; (mV)</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>I&lt;sub&gt;on&lt;/sub&gt; (μA/μm)</td>
<td>2.91</td>
<td>0.49</td>
</tr>
<tr>
<td>I&lt;sub&gt;off&lt;/sub&gt; (μA/V)</td>
<td>582.3</td>
<td>9.8</td>
</tr>
<tr>
<td>L&lt;sub&gt;ox&lt;/sub&gt; (Å)</td>
<td>∼10&lt;sup&gt;-15&lt;/sup&gt;</td>
<td>∼10&lt;sup&gt;-15&lt;/sup&gt;</td>
</tr>
<tr>
<td>I&lt;sub&gt;on&lt;/sub&gt;/I&lt;sub&gt;off&lt;/sub&gt;</td>
<td>∼10&lt;sup&gt;-7&lt;/sup&gt;</td>
<td>∼10&lt;sup&gt;-6&lt;/sup&gt;</td>
</tr>
<tr>
<td>SS (mV/decade)</td>
<td>&lt;100</td>
<td>64</td>
</tr>
<tr>
<td>peak mobility (cm&lt;sup&gt;2&lt;/sup&gt;/(V·s))</td>
<td>932–1235</td>
<td>40</td>
</tr>
</tbody>
</table>

<sup>a</sup> Key: I<sub>on</sub><sup>1</sup>, on current normalized to width or diameter; I<sub>on</sub><sup>2</sup>, on current normalized to L/(WV<sub>ds</sub>); GAA, gate all around.

Table 1 compares characteristics of NWFETs published in recent studies to a typical NWFET in this study with a 10 μm long and 3.4 nm wide nanowire. A similar comparison is made for nanobell FETs in the Supporting Information (see Table S2 in Supporting Information). The drive current is usually normalized to width, and is shown as I<sub>on</sub><sup>1</sup>, but it neglects device length and drain bias (V<sub>ds</sub>). To compare devices more completely, I<sub>on</sub><sup>2</sup> shows drive current normalized to L/(WV<sub>ds</sub>), revealing performance of our NWFETs surpasses recently reported p-type Si NWFETs. We find it especially noteworthy that I<sub>on</sub><sup>2</sup> of our back-gated NWFETs (without a highly doped channel) is significantly higher than top-gated devices with thinner gate oxide as well as shorter and wider nanowires. In contrast to a recent study of junctionless nanowire transistors (∼30 nm width), in which high channel doping (∼10<sup>19</sup> cm<sup>−3</sup>) is said to be required to ensure sufficient drive current,<sup>28</sup> the NWFETs shown here operate with higher current (scaled for length) despite being fully depleted due to their tiny volume and low doping. Low channel doping is beneficial to mobility because low-field mobility is primarily bound by impurity scattering.<sup>29</sup> As drive current in MOSFETs is proportional to μ<sub>ox</sub><sup>W/L</sup>, higher drive current in our devices can only be attributed to higher mobility because W is shorter, C<sub>ox</sub> is lower, and L is longer in our back-gated NWFETs in comparison to top-gated devices. Therefore, we believe quantum confinement induced enhancement in hole mobility may account for higher performance of our devices.

In MOS device analysis, field effect mobility is a standard parameter extracted from transfer characteristics. For typical back-gated NWFTs, field effect mobility is calculated as (derived directly from standard MOSFET μ<sub>E</sub>—see Supporting Information)

\[ \mu_{FE} = \frac{L^2G_{m}}{CV_{ds}} \]

where L is gate length (cm), G<sub>m</sub> is transconductance (S), C is nanowire to back-gate capacitance (F), and V<sub>ds</sub> is drain–source voltage. For accurate calculation of μ<sub>FE</sub> the nanowire to back-gate capacitance must be precisely calculated or measured. As direct measurement of capacitance is challenging with a back-gated configuration on SOI, an analytical expression (wire-to-plate model<sup>30</sup>) is typically used for calculating capacitance in most back-gated NWFETs. The analytical expression only estimates capacitance and accuracy depends on how closely a device’s physical characteristics align with assumptions employed in developing the expression. To calculate actual capacitance values in our devices, finite element mesh (FEM) simulation (COMSOL—see Supporting Information) was performed with the specific geometrical parameters of devices of each size, as measured by HR-TEM. Figure 3b shows peak hole mobility for both nanowire and nanobell P-MOSFET devices of different channel lengths. Peak mobility occurs at low gate voltage, between ∼1.7 and 5 V (V<sub>bg</sub>). Mobility in both our nanowire and nanobell devices is higher than in conventional MOSFETs (∼40 cm<sup>2</sup>/(V·s)), with nanowire devices exhibiting much higher mobility, apparently due to higher degree of confinement. We measured the sizes of nanowires in devices with mobility of 1235 cm<sup>2</sup>/(V·s) (10 μm length) and ∼737 cm<sup>2</sup>/(V·s) (20 μm length) to be 3.4 and 4.3 nm, respectively. Interestingly, a sharp peak in mobility is expected (by theoretical simulation) for nanowire diameter of ∼3 nm.<sup>22</sup>

Mobility can be enhanced by multiple factors in NWFETs. For example, strain in the Si lattice can increase mobility, as demonstrated by well-known Si strain engineering techniques.<sup>31</sup> Long duration oxidation (hours) has been shown to produce only minor strain-induced performance enhancement in NWFETs.<sup>32</sup> In our process, oxidation time to grow thin oxide (∼5–6 nm) is short (10 min). Average d-spacings of oxidized and unoxidized Si nanowires as well as bulk substrate were measured (by HR-TEM) to be nearly identical (see Supporting Information), indicating no observable strain is introduced by oxidation. Moreover, the gate oxide—Si nanowire interface should not be affected by oxidation, as the gate oxide is the BOX in these back-gated devices. Consequently, we believe strain cannot account for significant enhancement of mobility in our nanowire FETs. Thus, enhancement of mobility in nanowire FETs must be, in large measure, a result of increased quantization of carrier energy distribution. As spatial confinement increases, the continuous band structure of Si becomes progressively discrete, leading to a 1D density of states (DOS), especially in the sub-5-nm nanowires. The highly discrete DOS may produce carriers with a more uniform energy distribution. Another shifting attribute of carriers in confined states is effective mass, which is directly related to
energy of carriers. In bulk Si, the valence band maximum is comprised of bands with vastly different dispersion relations, resulting in both heavy and light holes contributing to current in conventional P-MOSFETs. In nanowire P-MOSFETs, dispersion relations of the discrete subbands comprising the valence band is expected to be entirely different from the heavy and light hole bands in bulk Si. At higher gate field, current in our sub-5-nm nanowire devices may well be comprised of holes with similar effective mass, thereby reducing intersubband scattering. Moreover, the effective mass of holes in the first few subbands of 3 nm [110] nanowires is estimated to be between 2 and 4 times lower than the effective mass of heavy holes in bulk Si. At higher gate field, significant mobility oscillation is observed in nanowire devices (see Supporting Information), in accordance with expected increase in intersubband scattering as holes of increasingly varying effective masses (due to inclusion of more subbands) start to contribute to current flow. While this high gate bias phenomenon clearly indicates quantum confinement, it does not impede device performance because FET devices are rarely operated at high gate bias.

Typically, extracted mobility can have other contributing factors, such as processing and measurement conditions. However, both our nanowire and nanobelt devices are made at the same time, on the same substrates, and with the same process. Moreover, the same techniques are employed for extraction of relevant characteristics (capacitance, mobility). Therefore, processing and measurement conditions or data analysis techniques are not responsible for the mobility differences between nanowire and nanobelt devices. Higher mobility in nanowire devices is clearly a result of increased current conduction despite the tiny channel as evidenced by higher current density in nanowire devices, as shown in Figure 3c. A comparative assessment of our nanowire and nanobelt devices directly infers that higher hole mobility in sub-5-nm nanowire FETs compared with nanobelt FETs (with the same height of ~4.3 nm) can only be attributed to increased confinement (2D for nanowire vs 1D for nanobelt) of holes in sub-5-nm nanowires. We also observed a larger spread in hole mobility of nanowire devices, which may be attributed to multiple factors including discrete dopant fluctuation and also size variation of nanowires due to EBL over large areas. It is important to note that the observed spread is not simply due to change in calculated capacitance resulting from size variation of the geometry used in capacitance simulation; we show that simulated capacitance exhibits only minute variation with changes in nanowire size (see Supporting Information). Therefore, the spread in mobility of nanowire devices is a phenomenon caused by an intricate combination of varying quantum confinement in nanowires of slightly different widths and discrete dopant fluctuations, which is known to cause transconductance oscillations and threshold voltage variation in low-doped devices. This may well explain why nanobelt devices do not exhibit much spread in mobility, due to their larger size and, therefore, statistically less variation of dopant atoms. It may be possible to eliminate discrete dopant fluctuation in nanowire devices while retaining high mobility, by using intrinsic Si or multiple Si nanowires as channel.

In conclusion, we have developed a simple lithographic process, without requiring high doping, for making high-performance long channel NWFETs with sub-5-nm nanowires. A comparative assessment of nanowire and nanobelt FETs shows strong enhancement of performance characteristics such as mobility, drive current, and current density in NWFETs as a result of increased quantum confinement. Performance of these back-gated NWFETs surpasses that of recently reported Si nanowire devices despite the lack of complementarily doped junctions and highly doped channel or compound semiconductor materials. In addition to high performance, our approach facilitates practical and inexpensive fabrication of novel devices such as ultrasensitive biosensors.

**ASSOCIATED CONTENT**

# Supporting Information.

Simulation of back-gate capacitance, mobility oscillation, a table containing comparison of nanobelt FETs with ultrathin-body FETs, and interatomic spacing measurement comparison between oxidized nanowire (from a sub-5-nm nanowire FET) and the bulk substrate. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

Corresponding Author

*Electronic mail: walter.hu@utdallas.edu*

**ACKNOWLEDGMENT**

This research was supported by the National Science Foundation (ECSS #0955027), by KETI through the international collaboration program of COSAR, by the State of Texas Emerging Technology Fund (ETF), and by Texas Instruments Inc., Richardson, TX. The authors thank Professors Bruce Gnade and William Frensley of The University of Texas at Dallas for useful discussions on data interpretation.

**REFERENCES**

(2) Cui, Y.; Wei, Q. Q.; Park, H. K.; Lieber, C. M. Science 2001, 293 (5533), 1289–1292.