Flexible Nanoscale High-Performance FinFETs

Galo A. Torres Sevilla, † Mohamed T. Ghoneim, † Hossain Fahad, † Jhonathan P. Rojas, Aftab M. Hussain, and Muhammad Mustafa Hussain *

Integrated Nanotechnology Lab, King Abdullah University of Science and Technology, Thuwal, Makkah 23955-6900, Saudi Arabia. † Equally contributing authors.

ABSTRACT With the emergence of the Internet of Things (IoT), flexible high-performance nanoscale electronics are more desired. At the moment, FinFET is the most advanced transistor architecture used in the state-of-the-art microprocessors. Therefore, we show a soft-etch based substrate thinning process to transform silicon-on-insulator (SOI) based nanoscale FinFET into flexible FinFET and then conduct comprehensive electrical characterization under various bending conditions to understand its electrical performance. Our study shows that back-etch based substrate thinning process is gentler than traditional abrasive back-grinding process; it can attain ultra-flexibility and the electrical characteristics of the flexible nanoscale FinFET show no performance degradation compared to its rigid bulk counterpart indicating its readiness to be used for flexible high-performance electronics.

KEYWORDS: flexible silicon · FinFETs · high-performance flexible electronics

Since the first introduction of the microprocessor in the early 1970s, silicon electronics have become an essential part of our daily life. Its unparalleled performance and cost-yield advantage makes silicon the most highly used material in today’s digital world. With more than 90% of electronics based on silicon micro- and nanofabrication processes, silicon has become indispensable for our technology-centered lives. However, advances in ultra-mobile computation and implantable and wearable electronics are still hindered by silicon’s brittleness and its inherited lack of flexibility. For this reason, in recent years, many different approaches to make flexible electronics have been introduced using polymer-based substrates. 1–4 Although polymer based electronics have shown exciting progress in displays and sensors, their thermal instability and inherent low carrier mobility 5 hinder their potential for high performance and low stand-by power electronics used in today’s ultramobile computation and communication devices. As a different approach, extremely high mobilities (10^5 cm^2 V^-1 s^-1) have been demonstrated with 2D single walled carbon nanotubes and graphene. However, the difficulty of integrating them with commonly known fabrication techniques and their low integration density impede their potential to become candidates for future high performance electronics. Also, graphene’s nearly zero bandgap property and doping difficulties makes it an undesired material for logic electronics due to the lack of distinguishable on/off behavior. 7,8 However, graphene’s potential for radio frequency (RF) applications has already been demonstrated with promising results.9 To address all previous concerns, recently, many different transfer techniques have been demonstrated to combine the outstanding flexibility of polymer based substrates with the excellent electrical properties of silicon. 10–18 Transfer techniques are based on the creation of silicon nanoribbons from unorthodox (111) or expensive silicon-on-insulator substrate and then transferring them to a polymer based carrier substrate where the device is fabricated and tested. This approach shows promising results and advances in the area of unusual electronics. But still there is substantial area of improvement between transfer techniques and industry standard complementary metal oxide semiconductor (CMOS) processes. Ultra-large-scale-integration (ULSI) densities can only be achieved with today’s most advanced lithographic techniques which
involve critical alignment of different layers, making plastics an undesired substrate due to their nonuniformity. Also, etching and deposition techniques usually consist of using complex chemistry to remove undesired material from the substrate and in the case of plastic substrates, polymer cross contamination with the channel and gate stack is still an issue to be solved. On the other hand, alternate approaches to reduce silicon thickness have been recently developed using fracture of single crystalline substrates (spalling) at different depths in order to obtain thin flexible film.\(^{19,20}\) However, the use of costly substrates like ET/UTSOI (compared to regular SOI that we have used), processes like high energy ion implantation, inherited limited bendability (nickel thin film has inherent high residual stress therefore bending is limited compared to our bending radius of \(<0.5\) mm; additionally, global presence of highly stressful film potentially increases the risk of breaking the spalled substrate) and film opaqueness are some of the factors that limit their potential for high performance applications. One more technique has been developed which makes use of a double layer of monocrystalline silicon, while the first layer is kept intact for device fabrication, the bottom layer is anodic etched to introduce porosity allowing the top layer to be easily peeled off with stamp-transfer techniques.\(^{21,22}\) Although highly complex circuits have been achieved with this method, the complexity of the process makes it difficult to integrate with known silicon microfabrication techniques. Also, the high processing costs associated with processes such as silicon epitaxy as well as limited flexibility in the resulting film hinder its potential for low-cost/high-yield applications. Finally, commercially available thin silicon films have been introduced to overcome all previous problems. However, the higher cost in handling difficulties, fragility and processing complications make them undesired substrates for commonly known fabrication methods.

To eliminate all previous concerns, in recent years we have demonstrated a different method to transform conventional electronics into flexible ones without compromising cost, yield, performance and efficiency.\(^{23–27}\) Since the process creates deterministic network of trenches/holes, we followed a back-etch process from the back-side of the substrate to eliminate any holes/trenches in the peeled-off (thinned down) silicon fabric. In this work we demonstrate the world’s most advanced transistor architecture (FinFET) and material set (high-x/metal-gate) on a flexible monocrystalline silicon-based substrate. In addition to the advanced architecture and material set, we show a process to reduce silicon thickness in an easy and repeatable way in order to obtain high performance devices without comprising silicon’s cost/yield advantage. The back etch process is softer than traditional back grinding, which is abrasive in nature and can thin down a substrate to a limited thickness of \(50\) \(\mu\)m only. Additionally, the objective of our study is to acquire a comprehensive understanding of the electrical characteristics of a flexible nanoscale FinFET; therefore, absence of hole/trenches in the flexible substrate offers a more holistic perspective on the device operation. The fabricated devices show extremely high performance and reduced short channel effects due to the increased electrostatic control achieved with dual-gate FinFETS. From all the perspectives, we envision this process and fabrication technique as a step forward toward very large scale integration of complex and high-performance circuits on flexible substrates.

**FABRICATION**

The FinFETs fabrication process started with 8 in. SOI monitor grade wafers using state of the art gate first flow. First, the Fins were patterned using deep ultraviolet light (DUV) and resist trimming to obtain 20 nm features. Then, the top silicon layer was etched using anisotropic reactive ion etching (RIE). Next, we formed the gate stack with 10–20 nm of titanium nitride (TiN) as metal gate and 2–4 nm of hafnium dioxide (HfO\(_2\)) as dielectric. Then, the gate is patterned to obtain different gate lengths (250 nm to 10 \(\mu\)m). Source and drain were formed using ion implantation followed by NISi to form ohmic contact between the test pads and the gate, source and drain regions. Finally, activation anneal is performed on the fabricated wafers at 1000 °C for 10 s followed by forming gas anneal (FGA-\(\text{N}_2/\text{H}_2\) at 420 °C). Figure 1 shows the process to transform the fabricated devices into flexible ones. To process each die separately, the fabricated wafers were diced into 2.5 cm \(\times\) 3 cm pieces (Figure 1a). At this point, each die is spin-coated with thick photoresist (PR, 7 \(\mu\)m) to protect the on chip devices from the following thinning processes (Figure 1b). Then, the dies are turned upside-down and the back of the substrate is etched using anisotropic etching to reduce its thickness and achieve the required flexibility (Figure 1c). The back-etching process is divided into 4 different steps to make sure that the required thickness is achieved without under or overetching. The first step reduces the thickness of the substrate from 800 to 200 \(\mu\)m (\(\pm30\) \(\mu\)m) in a single etch step. Then, following the same methodology and changing the etching time, the substrate is etched 50 \(\mu\)m in each step until the required 50 \(\mu\)m in thickness is achieved (\(\pm10\) \(\mu\)m). The etch process is carried out at a temperature of –20 °C (Etch step, 7 s, 1300 \(\text{W}_{\text{ICP}}\), 30 \(\text{W}_{\text{RF}}\), 35 mTorr, 5 sccm \(\text{C}_4\text{F}_8\), 120 sccm \(\text{SF}_6\) deposition step, 5 s, 1300 \(\text{W}_{\text{ICP}}\), 5 \(\text{W}_{\text{RF}}\), 35 mTorr, 100 sccm \(\text{C}_4\text{F}_8\), 5 sccm \(\text{SF}_6\)). It is important to note that between each etching step, a control thickness measurement of the substrate is performed in order to prevent over etching and damage of the on-chip devices. All the measurements were done using two different profiling systems in order to confirm the
thickness of the substrate. The first system is a contact-based profiler DETAK-8, which makes use of microtip to measure the step size between two surfaces. The second system consists of an optical ZYGO profiler, which makes use of reflected light in order to measure the step size between two different surfaces. According to the readings, the difference between the measurements is less than $\pm 5 \mu m$. At this point, due to the reduced thickness of the substrate, the devices can be bent to a minimum-bending radius of 1.5 cm. Finally, the protecting PR is removed from the top of the substrate (Figure 1e) and the devices are tested at different bending conditions (compressive and tensile). Figure 2a shows the fabricated and processed FinFETs at 0.5 mm bending radius. To confirm the final thickness of the substrate, scanning electron microscopic (SEM) imaging of the sample was performed (Figure 2b). Finally, Figure 2c shows the resulting transmission electron microscopic (TEM) imaging of the patterned Fins on standard SOI wafer. For electrical characterization, we found P-MOS ($L_g = 250 \text{ nm}$) and N-MOS ($L_g = 1 \mu m$) has highest yield (out of nearly 200 devices yield rate was close to 75%).

RESULTS

To study the behavior of the fabricated FinFETs, $I-V$ characteristics were obtained in linear and saturation regions. The saturation currents at $V_{GS} = \pm 1.25 \text{ V}$ and $V_{DD} = \pm 1.25 \text{ V}$ for released P-MOS ($L_g = 250 \text{ nm}$) and N-MOS ($L_g = 1 \mu m$) FinFETs are 383 and 83 $\mu A/\mu m$, respectively (Figure 3a,b). Figure 3c,d depicts the output characteristics of flexible N- and P-MOS devices. We focused our study on P-MOS (shorter $L_g$) to comprehend a full analysis at different bending radii and stress direction (compressive and tensile). Figure 4a,b shows the transfer and output characteristics of released P-MOS at different bending radii. The electrical characterization started by extracting the threshold voltage of the devices ($V_t$) from the $I_D-V_G$ curve using linear extrapolation method. With the $I_D-V_G$ curves, the threshold voltage can be determined by

$$V_t = \frac{V_{GS} - V_{DG}}{2} \quad (1)$$

where $V_{SS}$ is the intercepting point between the linear extrapolation at maximum slope of the $I_D-V_G$ curve.
with the gate voltage axis. With eq 1, the obtained threshold voltage is \( \frac{-C_0}{0.556 \text{ V}} \) in the linear region and \( \frac{-C_0}{0.474 \text{ V}} \) in saturation (Figure 5a). To continue with the device characterization, mobility was extracted and calculated at low drain voltages by

\[
\mu_{\text{eff}} = \frac{L}{W} \frac{g_d}{C_{\text{ox}}(V_{\text{GS}} - V_t)}
\]

where \( L \) and \( W \) are the channel length and width of the transistor, \( V_{\text{GS}} \) and \( V_t \) are the gate to source voltage and threshold voltage, respectively, \( C_{\text{ox}} \) is the gate dielectric capacitance, and \( g_d \) is the drain conductance, which can be determined by

\[
g_d = \frac{\partial I_{\text{D}}}{\partial V_{\text{GS}}} \quad V_{\text{GS}} = \text{constant}
\]

where \( I_{\text{D}} \) is the drain current and \( V_{\text{GS}} \) is the gate to source voltage. \( I_{\text{D}} \) is calculated when extremely scaled dielectric thickness is used in the transistor fabrication causing gate leakage to oppose drain current and hence reducing the calculated mobility value. \( I_{\text{D}} \) is calculated by

\[
I_{\text{D}} = I_{D} + I_{G}
\]

where \( I_{D} \) is the drain current and \( I_{G} \) is the gate current. With eqs 2, 3, and 4, the calculated peak effective mobility was found to be \( 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1} \) (Figure 5b).

To continue the device characterization, the subthreshold swing was obtained from the \( I_{\text{D}} \) vs. \( V_{\text{G}} \) curve for released (63 mV/dec) and bent P-MOS (Figure 5c). Also, DIBL was extracted for P-MOS devices obtaining 68 mV/V in the case of released sample. Figure 5c also depicts the change in DIBL due to the induced stress during bending of the sample. The obtained value for \( I_{\text{on}}/I_{\text{off}} \) ratio in the case of released P-MOS devices is 5 decades and is kept almost unchanged at different bending radii (not shown). The obtained peak transconductance (\( g_{m} \)) of the device was found to be 0.33 mS.
Intrinsic gate delay was extracted using moving \( V_{DD} \) method and is shown in Figure 6a depicting no significant changes at different bending conditions. Finally, the gate leakage \( J_G \) was obtained to be 0.9 A/cm\(^2\) in the case of released P-MOS devices. Figure 6b shows the obtained results for the gate leakage extraction at different bending radii.

For the first time, we also characterize the performance at different applied stress while bending the samples at different bending radii with a minimum compressive radius of 3 cm (bent downward) and a minimum tensile radius of 1.5 cm (bent upward). The maximum change in saturation current is 7.6% for PMOS at bending radius of 1.5 cm (0.17% tensile nominal strain). The maximum change SS is 10.3% for PMOS at a bending radius of 5 cm (0.05% tensile nominal strain). The maximum change in DIBL is 3.8 times for PMOS at bending radius of 1.5 cm (0.17% tensile nominal strain). The maximum change in \( J_G \) is \( \sim 2 \) order of magnitude for PMOS at a bending radius of 1.5 cm (0.05% compressive nominal strain). The maximum change in \( V_T \) is 7.4% and 5.1% for PMOS, in linear and saturation mode, at 1.5 cm, and 5 cm bending radius, respectively. The maximum change in peak \( g_m \) is 36.4% for PMOS at a bending radius of 3 cm. The maximum change in peak mobility is 29.6% for PMOS at a bending radius of 1.5 cm and \( V_{GS} \) of 0.625 V. The maximum change in \( I_{on}/I_{off} \) ratio is 15.9% for PMOS at a bending radius of 5 cm. Table 1 shows a comparison of the past reports and this work. Finally, the obtained results concur with previously reported studies, which show that saturation current is normally reduced when stress is applied in the longitudinal direction of the transistor’s channel.\(^{29} \)

**DISCUSSION**

In this work, we have demonstrated a CMOS compatible method to transform state of the art rigid silicon-on-insulator (SOI) based FinFET devices into flexible ones without compromising performance, integration density and cost/yield advantage of silicon microfabrication processes. Our method has the advantage of allowing complete fabrication of devices prior to release, allowing us to keep standard processes...
without introducing any constraint in design and commonly know deposition, etching and lithography methods. Also, since the devices are completely silicon based and transfer to polymer substrates is not required, thermal budget constraints are kept intact when compared to industry standard processes. Our process then sets a major step toward true integration of state of the art ultramobile applications. The advantage offered by the superior electrical characteristic of inorganic substrates is here demonstrated that may be combined with standard industry processes to produce flexible devices. Also, the maturity of silicon microfabrication techniques makes the transformation of silicon semiconductor from rigid and brittle substrates to flexible ones the most logical step toward very large integration of flexible high performance electronics. We have also demonstrated that with the use of simple etching techniques and without introducing any additional photolithography step, we are able to control the final thickness of the substrate with an accuracy of ±10 μm in a 50 μm thick substrate. In this work, we have shown P- and N-MOS devices on a flexible monocrystalline silicon substrate parting from standard 8 in. SOI wafers. We have chosen to characterize P-MOS devices in depth due to its scaled gate length (250 nm). It is to be noted that both N- and P-MOS devices were fabricated on the same wafer and with the same stress (forward and downward direction) show no performance degradation (compared to the FinFETs on rigid-bulk-traditional SOI substrate). The study is a concrete step forward toward introduction of flexible high-performance electronics for the Internet of Things (IoT).

**CONCLUSION**

We have reported a flexible (0.5 mm bending radius) nanoscale FinFET on silicon-on-insulator using a back-etch based substrate thinning process (50 μm silicon fabric). The process proves to be gentler than the traditional abrasive back grinding process and also does not introduce network of holes/trenches, typical signature of our previously demonstrated trench-protect-release-reuse process. Comprehensive electrical characterizations using various bending radii (both upward and downward direction) show no performance degradation (compared to the FinFETs on rigid-bulk-traditional SOI substrate). The study is a concrete step forward toward introduction of flexible high-performance electronics for the Internet of Things (IoT).

**Acknowledgment:** The authors declare no competing financial interest.

**REFERENCES AND NOTES**


---

**TABLE 1. Comparison for Different Fabrication Methods of Flexible MOS Transistors**

<table>
<thead>
<tr>
<th>group</th>
<th>IBM (flex)</th>
<th>Intel (rigid)</th>
<th>this work (flex)</th>
<th>Rogers (flex)</th>
<th>Banerjee (flex)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arch.</td>
<td>Planar</td>
<td>Tennate Fin</td>
<td>Double-gate Fin</td>
<td>Planar</td>
<td>Planar</td>
</tr>
<tr>
<td>Lg (nm)</td>
<td>40</td>
<td>40</td>
<td>10000/(N/250(P)</td>
<td>20000</td>
<td>150</td>
</tr>
<tr>
<td>ION (mA/cm)</td>
<td>0.02(N)/0.5(P)</td>
<td>1.4(N)/1.1(P)</td>
<td>0.083(N)/0.383(P)</td>
<td>0.09625(N)</td>
<td>0.15(P)</td>
</tr>
<tr>
<td>Ion (nA/μm)</td>
<td>5</td>
<td>100</td>
<td>2.2</td>
<td>0.025</td>
<td>1</td>
</tr>
<tr>
<td>fO/Io</td>
<td>10^6</td>
<td>10^6</td>
<td>10^6</td>
<td>10^6</td>
<td>10^6</td>
</tr>
<tr>
<td>SS (mV/dec)</td>
<td>-</td>
<td>90</td>
<td>150(N)/63(P)</td>
<td>190</td>
<td>720(N)/81(P)</td>
</tr>
<tr>
<td>VN (V)</td>
<td>0.25</td>
<td>-</td>
<td>0.36(N)/0.556(P)</td>
<td>-0.72</td>
<td>-</td>
</tr>
<tr>
<td>μeff (cm^2/V·s)</td>
<td>-</td>
<td>400(N)/200(P)</td>
<td>102(P)</td>
<td>375</td>
<td>252(N)/51(P)</td>
</tr>
</tbody>
</table>