A Silicon Nanowire Ferroelectric Field-Effect Transistor

Violetta Sessi,* Maik Simon, Halid Mulaosmanovic, Darius Pohl, Markus Loeffler, Tom Mauersberger, Franz P. G. Fengler, Terence Mittmann, Claudia Richter, Stefan Slesazeck, Thomas Mikolajick, and Walter M. Weber

The design and characterization of a Schottky-type ferroelectric field-effect transistor based on a nominally undoped silicon nanowire are reported. The nanowire transistor is fabricated by top-down technology starting from a silicon-on insulator wafer. A thin ferroelectric HfO$_{2.33}$Zr$_{0.62}$O$_2$ layer is integrated via a gate-first approach. Abrupt Schottky source/drain contacts to the undoped silicon are provided by NiSi$_2$ formation. Two distinct nonvolatile transistor states (programmed and erased) are observed in correspondence to negative and positive polarization in the ferroelectric layer, delivering a memory window of $\approx 1.5$ V and, differently to conventional ferroelectric field effect transistors, yielding an occurrence difference of up to 30%. These results are interpreted as a combination of effects, arising from the proximity of the ferroelectric layer to both the channel and the Schottky-junction regions. The threshold voltage shift, due to a polarization field acting on the channel, adds up to a polarization field-driven tuning of the current injection through the Schottky-source junction. This provides a strategy for manufacturing Schottky-type nanoscale transistors with the add-on nonvolatile option, following a complementary metal-oxide-semiconductor compatible process. In particular, the device concept is of great interest for achieving nonvolatile polarity modification in reconfigurable field-effect transistors.

One route to add new functionality to metal-oxide-semiconductor field-effect transistors (MOSFETs) is the integration of switching materials to the transistors. In this way, memory and computation are merged into a single building block. Such embedded memory concepts are expected to alleviate the bottleneck between processor speed and memory access time, with an overall improvement of the computing performances. A prominent candidate for embedded nonvolatile memory is the ferroelectric field-effect transistor (FeFET), which can operate as a nonvolatile memory cell but also serves to the realization of added functionalities like standard logic operations, memory-in-logic, or true random number generators. Moreover, the particular switching properties of FeFETs, such as multi-level and accumulative switching, make them appealing to realize synapses and neurons for neuromorphic computing. Between numerous existing ferroelectric gate-oxide materials, hafnium oxide (HfO$_2$) in its metastable orthorhombic phase has gained increasing attention in the device community in recent years. Different from conventional ferroelectrics like Pb(Zr$_{0.52}$Ti$_{0.48}$)$_2$O$_3$, which show several integration and scaling issues, the material HfO$_2$ is already in use as a high-k dielectric in highly scaled CMOS manufacturing. When doped with suitable dopants (e.g., Si, Al, Zr, Y) and/or a particular fabrication procedure is applied, the orthorhombic ferroelectric HfO$_2$ phase can be stabilized. Very importantly for applications, HfO$_2$-based ferroelectrics show stable ferroelectricity.

The end of classical scaling has become evident as manufacturing and design scaling reaches physical and economical limits. Already since many years, Moore’s law is complemented in semiconductor industry by the so-called “More-than-Moore” approach. Here, a cost-effective integration of many addi-

---

Dr. V. Sessi,[1] T. Mauersberger, Prof. T. Mikolajick
Technische Universität Dresden
Institut für Halbleiter- und Mikrosystemtechnik
Professur für Nanoelektronische Materialien
Nöthnitzer Str. 64a 01187, Dresden, Germany
E-mail: violetta.sessi@namlab.com

[1] Present address: Namlab gGmbH, Nöthnitzer Str. 64a 01187, Dresden, Germany
© 2020 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

DOI: 10.1002/aelm.201901244
with high coercive fields in the 1 MV cm\(^{-1}\) range for comparatively thin films of 5–30 nm thickness.\(^{[14]}\) This allows downsizing of devices and reduction of the necessary gate-stack height. Compatibility with miniaturized MOSFETs has been established for laterally scaled FeFETs having a channel length and width of 30 and 80 nm, respectively.\(^{[15]}\) Recently, large FeFET memory arrays based on Si:HfO\(_2\) have been successfully demonstrated.\(^{[4,16]}\) Among several dopants, Zr-doped films reach the largest compound concentration window and lowest annealing temperatures for the stabilization of the ferroelectric phase.\(^{[17]}\)

So far, most of the work on HfO\(_2\)-based FeFETs focused on planar geometry and doped silicon substrates. However, interesting devices adding functionality in a “More-than-Moore” way exploit the special electronic and structural properties of nonplanar nanostructures. One prominent example is reconfigurable field effect transistors (RFETs), which take advantage of the geometry and of the nanoscale size of silicon nanowires (SiNWs).\(^{[18]}\) RFETs can be electrostatically programmed at runtime as either p- or n-type transistors and are interesting as logic elements, i.e., building blocks of circuits expressing more functions than CMOS electronics counterparts in a single layout. The feature of polarity control is also valuable in conjunction with tunneling FETs,\(^{[19,20]}\) where the ambipolar behavior could be suppressed electrostatically.\(^{[19,20]}\)

The polarity control is usually achieved by application of an additional static voltage at a dedicated gate electrode, the so-called program-gate, which selects the carrier-type to be injected in the transistor. Such operation is volatile, and the information would be lost upon powering off the computer. An interesting development toward energy-efficient reconfiguration encompasses the addition of a memory-conserving layer within the FET gate-stack. Recently, a nonvolatile memory function integrated into the computing device was achieved,\(^{[21]}\) with a charge trapping layer added to a SiNW-RFET. However, this solution requires high switching voltages. In ref. [22], a nonvolatile tuning of the junction transmissibility was demonstrated for SiNW-FETs, by means of a HfO\(_2\)-based ferroelectric layer. Due to the employed top-gate geometry, the standard FeFET behavior was not shown. In both works, bottom-up grown SiNWs were used limiting the scalability of the devices.

Here, we present the fabrication and transfer characteristics of a top-down fabricated silicon nanowire ferroelectric field effect transistor (SiNW-FeFET). A gate-first approach was employed to integrate a ferroelectric Hf\(_{0.38}\)Zr\(_{0.62}\)O\(_2\) layer with nominal thickness of 5.5 nm as gate oxide. The SiNW-FeFET shows two distinct nonvolatile transistor states (programmed and erased) in correspondence to negative and positive polarization in the ferroelectric layer, separated by a memory window of \(\approx 1.5 \text{ V}\). Additional to this typical FeFET behavior, an on-current difference of about 30% between the two states is also observed. These results are interpreted as a combination of effects, arising from the proximity of the ferroelectric layer to both the channel and the Schottky-junction regions. The threshold voltage shift, due to a polarization field acting on the channel, adds up to a polarization field-driven tuning of the current injection through the Schottky-source junction. Potentially, in such devices the junctions’ transmissibility for a specific charge carrier can be made more transparent or more opaque simply by a voltage pulse, enabling nonvolatile programming of the transistor polarity.

The SiNW-FeFETs were produced by a gate-first approach as illustrated in Figure 1, where further details are given in the Experimental Section below.

SiNWs were fabricated by a top-down process, based on commercially available silicon-on-insulator (SOI) substrates with initial Si thickness \(t_{\text{Si}} = 20 \text{ nm}\) and buried oxide thickness \(t_{\text{SiO}_2} = 100 \text{ nm}\). Hydrogen silsesquioxane (HSQ), e-beam lithography

Figure 1. Side view: top-down fabrication of silicon nanowire (SiNW) FETs with ferroelectric Hf\(_{0.38}\)Zr\(_{0.62}\)O\(_2\) (HZO) gate-oxide, using a gate-first approach. Details can be found in the Experimental Section. a) SiNW structuring from an SOI substrate. b) Thermal oxide growth. c) Atomic layer deposition of HZO. d) Top gate-stack structuring. e) HZO chemical etching. f) Ni source and drain contact formation. g) Nickel-silicide creation. h) TEM cross-section image and i) color-coded EDS elemental mapping of the final SiNW-FeFET. The NWs have a trapezoidal cross-section with height \(\approx 19 \text{ nm}\) and width between 14 and 24 nm.
and reactive ion etching were used to structure the SiNWs (Figure 1a), followed by rapid thermal annealing (RTA) at 850 °C in O₂ atmosphere to grow a thermal oxide shell (Figure 1b). Next, a HfₓZr₁₋ₓO₂ (HZO) layer was deposited by atomic layer deposition (ALD; Figure 1c) at 250 °C. A Pt top gate was formed by e-beam lithography, metal evaporation, and lift-off (Figure 1d). Openings through HZO for source and drain contacts were self-aligned using the Pt top gate as a hard mask via wet chemical etching (Figure 1e). A further e-beam lithography was employed for deposition of the source/drain Ni contact metal (Figure 1f). Finally, RTA at 450 °C (Figure 1g) was performed in order to simultaneously i) promote Ni intrusion in Si and ii) crystallize the HZO in its orthorhombic phase. Lateral intrusion of Ni in the Si nanowire results in the formation of NiSi₂–Si interfaces, as known from transmission electron microscopy (TEM) analysis of structures fabricated under similar conditions. Nickel-disilicide was chosen as the contact compound given its electrical, structural, and interfacial properties, e.g., low formation temperature, sufficiently low contact resistance, sharp and atomically flat interfaces, allowing adequately high currents and a sub-micrometer control of the lateral encroachment. The TEM cross-section in Figure 1h shows the SiNW device with silicon core and SiO₂/HZO shell. The Si core has shrunk slightly to a thickness of t₅ = 18.8 nm due to the oxidation process. Furthermore, the SiNW channel section has a trapezoidal shape with width between 14 and 24 nm. The gate oxide is wrapped around the three sides with a thickness of ≈4.3 nm and geometry similar to a trigate FinFET. The thickness of the HZO is estimated to be ≈5.5 nm with composition of Hf₀.₃₈Zr₀.₆₂O₂. An inhomogeneity in the HZO layer thickness is observed in the cross-section of Figure 1h, it is unclear whether this is the result of processing or of the preparation of the TEM lamella.

In Figure 2, the transfer characteristics of the SiNW-FeFET are shown. In contrast to other works that show hysteretic

![Figure 2](image-url)
effects in ferroelectric devices, a pulsed programming and erasing of the device was used to explore the shift of the $I-V$ curve instead of sweeping up to high voltages in both directions. This was done to avoid detrimental effects caused by the slow ramping during the DC $I-V$ curves.[36] The measurements were performed using the terminals sketched in Figure 2a. The top-gate stack covered both the intrinsic Si channel and the metal-semiconductor (NiSi$_2$-Si) Schottky junctions (SJs) and was connected to a pulse measurement unit (PMU). The latter was employed for both recording the transfer characteristics and applying the program (PRG)/erase (ERS) pulses. The source terminal was fixed at ground potential, while the drain terminal was connected to a second PMU. During pulses both source and drain terminals were grounded.

Schottky-type SiNW devices based on intrinsic Si are ambipolar.[27] However, the n-current is comparably small if no optimization is performed[28,29] and falling below the detection limits of the PMU setup. This is the reason why the transfer characteristics of Figure 2c only show p-type currents for the chosen $V_{TG}$ range.

The effect of PRG and ERS pulses was evaluated with the simple pulse sequence sketched in Figure 2b. Pulse amplitudes $V_p = +9$ V and $V_N = -9$ V were employed for PRG and ERS, respectively. The pulse duration in both cases was $\Delta = 4$ ms. For reproducibility, the sequence was repeated five times. In Figure 2c, the drain current $I_D$ recorded in the programmed (after PRG pulse) and erased (after ERS pulse) state are plotted versus top gate voltage $V_{TG}$.

The transfer curves for the programmed and erased state appear shifted with respect to each other, with a memory window $MW = 1$ V. In particular, the shift is in the positive (negative) voltage direction for the transfer curve recorded after the ERS (PRG) pulse. This is coherent to what has been reported for n-type FeFETs in ref. [30] and in general is expected for FeFETs. In contrast to charge trapping that shifts the transfer curve always in the direction of voltage pulse, for ferroelectric gate oxide the transfer curve shifts always to the opposite direction.[26,31] We propose a qualitative interpretation based on the band diagrams sketched in Figure 2e, for the transistor in the on-state. A voltage pulse generates a polarization $P_{FE}$ within the ferroelectric HZO layer due to alignment of the internal dipoles, which remains after the pulse. This polarization directly above the Si channel acts as an additional electric field $E_{FE}$. A PRG pulse (positive) results in an additional positive $E_{FE}$, i.e., for a p-type FET the device at zero applied external voltage is more "off" compared to the case without $P_{FE}$. As a consequence, similar to what is reported for MOSFET-type FeFETS, the transfer characteristics after a PRG pulse is shifted toward more negative voltages by $E_{FE}$. The behavior reverses for the ERS pulse, i.e., $E_{FE}$ is negative and therefore the FET appears more "on" at zero applied external voltage.

An additional effect of $E_{FE}$ which is not observed in MOSFET-type FeFETs is evidenced in the linear plot of Figure 2d, where the average $I_D$ is plotted versus overdrive voltage. The latter is calculated as the difference between the applied voltage and the threshold voltage $V_{TH}$. Here, $V_{TH}$ is defined as the $V_{TG}$ at which $I_D = I_{TH} = 10^{-8}$ A. When comparing the transfer characteristics after PRG and ERS pulses in Figure 2d, a slightly different $I_D$ for the same overdrive is found. In particular, an almost constant relative current variation between ERS and PRG states $\Delta I_D^{\text{ON}} = (I_D^{\text{ERS}} - I_D^{\text{PRG}})/I_D^{\text{ERS}} \approx 30\%$ is observed in the overdrive range below $\approx -1.2$ V, approximately corresponding to the transistor on-state. This is not the case for MOSFET-type p-FeFET[30] and n-FeFET[9] devices, where the transfer characteristics after PRG and ERS pulses are only "horizontally" shifted. This effect may be attributed to the particular carrier injection mechanism in Schottky-barrier FETs (SB-FETs) that involves thermal emission but also tunneling, with the latter being dominant in the transistor's on-state.[28] The band-bending due to $E_{FE}$ described above in the specific case also modifies the SB width at the source and drain contacts. In particular, the SB width for hole injection (i.e., into the valence band $E_v$) is enlarged by a positive $E_{FE}$ (programmed state, Figure 2e, left) and reduced by a negative $E_{FE}$ (erased state, Figure 2e, right). Consequently, the junction transmissibility via tunneling is reduced or enhanced, respectively. As for SB-FETs the tunneling current is the dominant contribution to the on-current, the latter is very sensitive to modifications of the potential landscape around the SJs. In particular, the reported $\Delta I_D^{\text{ON}}$ of $\approx 30\%$ could be due to spatial inhomogeneity of $E_{FE}$, i.e., stronger $E_{FE}$ in proximity of the SJs, resulting in a sizeable modification of the tunneling barrier. A spatial inhomogeneity of $E_{FE}$ could have electrostatic reasons, in connection to the intrinsic character of the SiNW channel. This interpretation is coherent to what has been reported in ref. [22], where a strong on-current difference between PRG and ERS states was achieved by a remanent $E_{FE}$ localized at one single SJ (at the source contact). Note that when an SB-FET is used, the additional modulation of the SB width increases the memory window compared to the one that can be observed in a classical MOSFET. Therefore, SB-FETS can be beneficial to realize FeFET memory cells.

We have further characterized the SiNW-FeFET by varying the pulse intensity ($V_p$ and $V_N$), duration ($\Delta$), and wait time ($W$) between pulse and read, according to the sequences in Figure 3a,b. The results are displayed in Figure 3c,e for the PRG pulse, and in Figure 3d,f for the ERS pulse. In order to preset the devices always to the same reference state, the PRG (ERS) pulse sequence was preceded by a PRG-ERS (ERS-PRG) reset series, as in Figure 3a,b. For both PRG and ERS pulses, voltage pulse amplitudes above 6 V are required to achieve a detectable shift of the transfer characteristics. These large values result from a relatively thick SiO$_2$ layer in the gate-stack. In fact, due to its significantly lower dielectric constant and to a comparable physical thickness with respect to the ferroelectric layer, the resulting capacitive voltage divider in the gate-stack is such that a large portion of the gate voltage drops over the SiO$_2$ layer. Nevertheless, to increase the switching efficiency, it is desirable to increase the voltage drop over the HZO layer. This can be achieved by reducing the SiO$_2$ thickness and/or increasing the HZO thickness. In fact, as recently demonstrated for planar FeFETs with 20 nm thick HfO$_2$, this not only enhances the switching efficiency but also increases the memory window and the resilience against parasitic charge trapping.[32] Charge trapping is in fact known to affect HfO$_2$-based ferroelectric devices[34] and we expect it to influence also the memory window in our SiNW-FeFET device.
Above 6 V the shift in the transfer characteristics increases gradually, as can be seen in Figure 3e,f. Here, the extracted $V_{TH}$ values at $I_{TH} = 10^{-8}$ A are plotted versus $V_P$ ($V_N$), for several values of the pulse duration $\Delta t$. $V_{TH}$ is extracted from (c) as the $V_{TG}$ corresponding to drain current $I_D = I_{TH} = 10^{-8}$ A. f) Threshold voltage $V_{TH}$ plotted versus ERS pulse amplitude $V_N$, for several values of the pulse duration $\Delta t$.

In conclusion, we have shown the manufacturing and working mechanisms of a silicon nanowire FeFET. A
ferroelectric Hf$_{0.38}$Zr$_{0.62}$O$_2$ layer was integrated via a gate-first approach in the top gate stack of a silicon nanowire transistor fabricated by top-down technology. Two distinct transistor states (programmed and erased) were observed in correspondence to negative and positive polarization in the ferroelectric layer, with a memory window of $\sim$1.5 V. As the top gate also covers the SJs between intrinsic Si and NiSi$_2$ segments, associated with the ferroelectric polarization, also a non-volatile modulation of the SJ transmissibility was observed, with a relative current variation $\Delta I_{\text{ON}}^{\text{N}} \sim$ 30%. Consequently, the memory window is increased in the SiNW Schottky-FET compared to a conventional FeFET because of the additional effect of the modulation of the tunneling barrier transmissibility of charge carriers. The reported proof of principle of a nonvolatile SiNW FET with CMOS compatible materials can be exploited in multiple-gate, nonvolatile transistor architectures with the aim to test “More-than-Moore” emerging concepts merging memory and logic such as nonvolatile circuit programing.

Experimental Section

SiNWs were fabricated by top-down technology, starting from commercially available SOI wafers with initial Si thickness $t_{\text{Si}} = 20$ nm and buried oxide thickness $t_{\text{SOI}} = 100$ nm. E-beam lithography was used to pattern a hard-mask of HSQ that was spin-coated onto the SOI wafer. Subsequently, the not-masked silicon was removed by reactive ion etching utilizing SF$_6$, O$_2$, and CHF$_3$ gases and moderate radio frequency forward power of 50 W, leaving a top-Si mesa structure. After HSQ removal by hydrofluoric acid (HF), a thermal SiO$_2$ shell of $\approx$5 nm was grown above the SiNWs by RTA in O$_2$ atmosphere at 850 °C for 10 min. In order to passivate Si dangling bonds, annealing in N$_2$ atmosphere at 850 °C for 5 min and forming gas anneal (N$_2$:H$_2 = 9:1$) at 450 °C for 5 min were also performed.

Next, a Hf$_{0.5}$Zr$_{0.5}$O$_2$ layer with nominal thickness $t_{\text{HZO}} = 25$ nm and nominal composition Hf$_{0.5}$Zr$_{0.5}$O$_2$ was deposited by ALD. Tetrakis(ethylmethylamino)zirconium (TEMAZr) were employed as precursors, while H$_2$O was used as oxidant source. The HZO films were deposited by 100 supercycles of alternating TEMAHzf and TEMAZr-based processes with a cycle ratio of 1:1. The deposition temperature was 250 °C. The metal-gate stack for the SiNW FeFETs was fabricated by e-beam lithography, deposition of Ti/Pt (10 nm/20 nm) by an e-beam evaporator, and lift-off. The sample was then immersed in HF with concentration 2.5%, for 7 min in order to form openings through the HZO for the source and drain contacts. The different HF etching rate for the metals composing the gate stack (low for Pt, high for Ti) resulted in a selective under-etching of the Ti layer and thinning of the HZO layer underneath. The element-specific TEM in Figure 1i confirmed that no substantial Ti was present in the gate stack. From local energy-dispersive X-ray spectroscopy (EDS) of the HZO layer, the atomic ratio of Hf to Zr was found to be 38:62 (see the Supporting Information). The discrepancy with the expected 50:50 ratio could be attributed to a slightly higher etching rate of diluted HF for HfO$_2$ with respect to ZrO$_2$.

The polarization versus voltage curves of metal-insulator-metal structures fabricated using the same ALD process as for the SiNW FeFETs are reported in the Supporting Information, for the as-deposited (25 nm) as well as for a 5 nm thick HZO film. Note however that in this reference system an alternative metal electrode (TiN) was used due to processing constrictions.

The source and drain contacts were fabricated by e-beam lithography, a short dip in NH$_4$F-buffered HF, Ni sputter deposition (42 nm), and lift-off. RTA for silicidation and HZO crystallization was performed in a single RTA experiment in forming gas atmosphere (10% H$_2$ in N$_2$), at 450 °C for 50 s.

In order to investigate the structure of the SiNW-FeFET via TEM, a cross-section lamella was prepared via focused ion beam technique. The cross-section was analyzed utilizing a Jeol JEM F200 operated at 200 kV equipped with a dual window-less silicon drift detector.

Up to 7000 cycles were applied to the SiNW FeFETs for device screening before recording the transfer characteristics. Additionally, 100 wake-up cycles were run in concomitance with the PRG/ERS routine.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors acknowledge the financial support from CfAED (EXC 1056) and DFG ReproNanoWE4853/1-3. The authors are indebted to A. Tahn for the preparation of the TEM lamella.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

FeFETs, ferroelectrics, HZO, multigate FETs, nonvolatile memory, Schottky-barrier FETs, silicon nanowires, SOI substrates

Received: November 11, 2019
Revised: January 24, 2020
Published online: March 4, 2020
