Transformation: nanotechnology—challenges in transistor design and future technologies

B. Ullmann, T. Grasser

The transistor is one of the key inventions of modern society which have paved the way for innovative technologies being so ubiquitous in everyday life. Due to their economic importance and the pressure to keep production costs low while simultaneously increasing efficiency, transistors have been scaled to the deca-nanometer regime during the last decades. This excessive downscaling of device dimensions has now reached several physical limits, whereby transistor design and fabrication have been challenged numerous times. Although some of these challenges have been overcome by the use of new materials and cleverly designed device geometries, a lot of additional effort must be put into the research of innovative nanotechnologies in order to open up for new opportunities. In this work, we summarize recent developments in transistor technology and give a short overview of possible future technologies.

Keywords: transistor; MOSFET; reliability; nanotechnology

Transformation: Nanotechnologie – Herausforderungen beim Design von Transistoren und Technologien der Zukunft.

Der Transistor ist eine der wichtigsten Erfindungen der modernen Gesellschaft, die den Weg für innovative und im Alltag allgegenwärtige Technologien geebnet hat. Aufgrund ihrer wirtschaftlichen Bedeutung und des Drucks, die Produktionskosten niedrig zu halten und gleichzeitig die Effizienz zu steigern, wurden die Dimensionen der Transistoren in den letzten Jahrzehnten skaliert und haben kürzlich den 14-nm-Bereich erreicht. Diese Skalierung der Abmessungen hat inzwischen mehrere physikalische Grenzen erreicht, wobei Design und Herstellung von Transistoren mehrfach vor Herausforderungen standen. Obwohl einige dieser Herausforderungen durch den Einsatz neuer Materialien und clever gestalteter Geometrien überwunden wurden, muss weiter in die Forschung investiert werden, um mittels Nanotechnologie neue Möglichkeiten zu erschließen. In dieser Arbeit fassen die Autoren die jüngsten Entwicklungen in der Transistortechnologie zusammen und geben einen kurzen Überblick über mögliche Technologien der Zukunft.

Schlüsselwörter: Transistor; Zuverlässigkeit; Nanotechnologie

CrossMark

1. Introduction

More than 200 years ago the origin of electricity was analyzed for the first time and throughout this process the way to more revolutionary technological inventions has always been enabled by an enhanced insight into the physical process of electric charge manipulation. Although many inventions and patents were considered fundamental research and initially ignored by industry, they have eventually changed our way of living and as such are of enormous economic importance. A rather prominent example in this regard is the transistor, which is one of the key inventions of modern society arguably comparable to the domestication of fire and the invention of the wheel.

As a result of basic research on the physics of solids, transistors were able to replace vacuum tubes in the 1950s, which led in the following to the development of the integrated circuit and the microprocessor, which are at the heart of modern electronics. Thus, transistors paved the way for a new generation of powerful and efficient electronic devices with a seemingly unlimited number of applications in everyday life. As a consequence of the continuous improvement of their performance, modern technologies have enabled numerous innovative ways of global networking by connecting things and people, optimizing work flows as well as saving valuable resources.

The economic importance of these new technologies and the pressure to keep production costs low have been the driving forces for the development and improvement of the transistor. Moore predicted in 1965 an exponential relationship between device complexity (number of transistors per area unit) and time, by stating that "the complexity will double annually." This prediction has in the meantime been revised to a doubling every two years [1], resulting in over 5 billion transistors being processed on a single chip today. As a consequence of the complexity increase, transistors have been downscaled to the deca-nanometer regime during the past decades, which has resulted in severe reliability issues, power loss and instabilities as the physical limits are approached. Thus, new materials and geometries are required to overcome these challenges.

The new approaches in transistor design have led to the transformation from the field of Microelectronics to the field of Nanoelectronics, where Nanotechnology naturally plays a major role. Here we give a short overview of the consequences that have resulted from

Ullmann, Bianka, Institute for Microelectronics, Vienna University of Technology, Gusshausstrasse 27-29, 1040 Vienna, Austria (E-mail: bianka.ullmann@tuwien.ac.at);
Grasser, Tibor, Institute for Microelectronics, Vienna University of Technology, Gusshausstrasse 27-29, 1040 Vienna, Austria (E-mail: tibor.grasser@tuwien.ac.at)
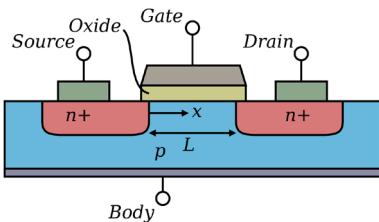
Fig. 1. Lateral planar *n*-type MOSFET: Two highly *n*-doped source and drain regions separated by a *p*-doped body region (e.g. Si) and an insulating layer (e.g. silicon oxynitride SiON) separating the gate contact from the body [2]

downscaling and recent developments in transistor technology together with their challenges.

2. The transistor in digital circuits

In digital circuits, the metal-oxide-semiconductor field-effect transistor (MOSFET) is the most important transistor technology due to the achievable short switching times and the nearly loss-less control at low frequencies. In a very simplified way, a transistor can be compared to a switch realized by modifying the conductivity properties of semiconductors (e.g., Si). For a more detailed explanation, Fig. 1 shows the cross section of an *n*-type MOSFET containing two highly *n*-doped regions, source and drain, separated by a *p*-doped body region. An insulating layer (e.g. silicon oxynitride SiON) is sandwiched between the gate and the body, separating them from each other and thus preventing a current flow. Given that a positive supply voltage (V_{DD}) is applied between the drain and the source (V_D), the voltage applied between the gate and the bulk (V_G) controls the current flow between the drain and the source I_D. If $V_G = 0$ V, the MOSFET is in its off-state and current flow is inhibited because of the reverse biased $p{-}n$ junction. By applying a low positive V_G the majority carriers are forced away from the oxide/body interface and therefore a depletion layer near the interface forms which is populated by minority carriers. As soon as V_G exceeds a certain threshold voltage V_{th}, the concentration of minority carriers is high enough to form a thin inversion layer near the interface. As a consequence, the minority charge carriers flow freely between the source and drain and the gate voltage controls the carrier concentration in the channel and thus the resistance. Further increase of V_G results in an increase of I_D until I_D reaches the saturation region which corresponds to the on-state of the MOSFET.

Unfortunately, any real device differs from such an ideal switch. The real switching characteristics, as shown in Fig. 2, are typically characterized by four parameters determined by materials, doping and geometry: the off-current, the subthreshold swing SS (reciprocal value of the slope S in a log-lin plot), V_{th} and the on-current. The off-current ($I_D \neq 0$ A at $V_G = 0$ V) is caused by leakage currents between source and drain and is inevitable. Furthermore, the switching process between off- and on-state shows switching dynamics characterized by V_{th} and SS which are limited to certain minimum values, and the on-current is the current flowing through the channel when the MOSFET is in its on-state.

The parameters mentioned in the previous paragraph have to meet certain requirements in digital circuits in order to ensure a correct interaction with other circuit components and a low power consumption of the circuit. A correct interaction assumes that the off-current and the on-current of a MOSFET correspond to output voltages fulfilling the limits for the digital levels low and high, respectively. In the CMOS technology, the digital low level is defined as a
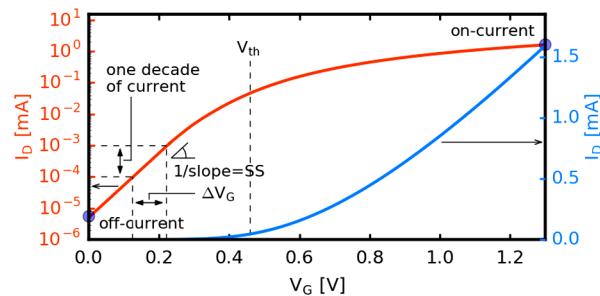

Fig. 2. Typical I_D–V_G characteristics of an nMOSFET: Drain current I_D plotted against gate voltage V_G on a log-lin (*red, left scale*) and a lin–lin (*blue, right scale*) scale. Off-current (current flowing when MOSFET is switched off), subthreshold slope (slope of the subthreshold region in a log-lin plot), threshold voltage (V_G where inversion layer is formed) and on-current (current flowing in the on-state) are the most important parameters characterizing the I_D–V_G curve (Color figure online.)
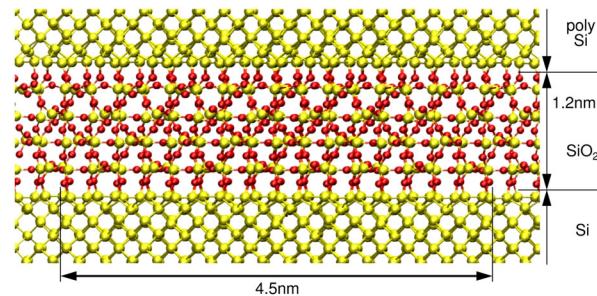

Fig. 3. Schematic atomic structure of a nanoscale MOSFET: polysilicon gate contact on SiO_2 insulator (with $t_{ox} = 1.2$ nm) on a silicon body [3]

voltage between 0 V and $1/3V_{DD}$ and the digital high level is defined as a voltage between $2/3V_{DD}$ and V_{DD}. Meeting the requirement of clearly distinguishable output voltage levels for the two digital states results in limitations for the ratio between on- and off-current which has to be made as large as possible. This can be achieved by maximizing V_{DD} while simultaneously keeping SS and V_{th} as low as possible. Besides increasing the ratio, a high V_{DD} would also ensure an on-current which is high enough to drive subsequent digital stages, and low SS and V_{th} would enhance the switching dynamics. However, increasing V_{DD} conflicts with the requirement of low power consumption since $P \propto V_{DD}^2$ and, as mentioned in the previous paragraph, SS and V_{th} are either fundamental limits or limited due to materials, doping and geometry of the MOSFET.

The consideration of these aspects in the fabrication process provide fundamental challenges for the design of MOSFETs in general. Furthermore, excessive scaling of the MOSFET geometries has led to further limitations as discussed in the next section.

3. Scaling trend of MOSFETs and challenges

While decades ago transistor structures were processed in the micrometer range, modern MOSFET structures have been scaled down to 22 nm in 2008 and to 14 nm in the current generation of transistors. In addition to the scaling of the width (W) and length (L) of the transistors, the oxide thickness (t_{ox}) has also been scaled down, reaching values less than 2 nm. Considering SiO_2 as the insulating material, 2 nm in fact contain a rather small number layers of atoms as shown in Fig. 3.
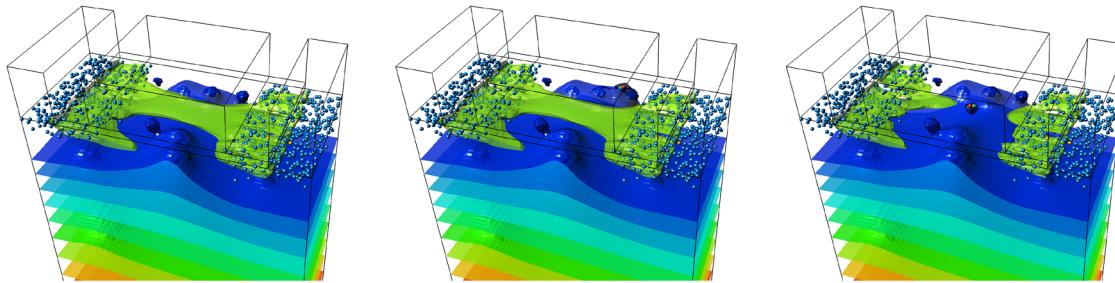
Fig. 4. A single percolation path formed by random discrete dopants (current flow shown in the uppermost layer) and contours of constant potential in a pMOSFET: *Left:* **current flow without a disturbance due to oxide defects.** *Center:* **reduced current flow when a defect is located beside the percolation path.** *Right:* **disturbance of the current flow when a trap is located directly in the center of the percolation path [4]**

The scaling of the geometry results in several design and fabrication challenges, for example, a short transistor length (less than 100 nm) leads to short-channel effects, e.g., drain-induced barrier lowering, which affects the MOSFET performance. Additionally, channel leakage currents are more pronounced, therefore, the off-current increases significantly. Moreover, considering that $C_{ox} \propto A/t_{ox}$, where $A = L \times W$ is the area, t_{ox} has to be scaled in the same manner as A. Otherwise, C_{ox} would increase significantly which would have a great impact on V_{th} and the switching dynamics. However, t_{ox} is limited to a certain minimum value because of quantum effects like tunneling, leading to a dramatic increase of leakage currents if the oxide thickness is further reduced. With a t_{ox} below this value, a loss-less control cannot be ensured any longer. Furthermore, the supply voltages cannot be scaled in the same manner as the device geometry, otherwise, the ratio between the on- and off-current would deteriorate. Last but not least, the electric field in the oxide E_{ox} has increased considerably due to the indirect proportionality to t_{ox}. As a consequence, degradation effects depending on E_{ox}, e.g., the bias temperature instability have a greater impact than in devices with thicker oxides.

Another consequence of downscaling is a higher variability due to the variance of parameters between MOSFETs processed in the same manner. The variance is higher for nanoscale devices than for large devices in the micrometer range because nanoscale devices contain, in contrast to large devices, only a countable number of discrete dopants. The slightest deviations of their number or position influence the non-uniformly current flow over the width, the so-called percolation path (see Fig. 4 left) [4, 5]. Furthermore, the relative change of device dimensions due to fabrication variability increases with scaling. As a result of both, the variance of discrete dopants and dimension deviations, even MOSFETs of the same technology and processed in the same manner show a significant variance in their characteristics, like V_{th}, and behave differently in digital circuits.

Even more dramatic, in contrast to Fig. 3, which shows an ideal arrangement of atoms making up a transistor, in reality the atoms are not perfectly arranged. On the one hand, if the oxide material is amorphous, structural defects related to dangling bonds inevitably occur at the bulk/oxide interface and on the other hand, the oxide contains intrinsic defects. Some of these defects have the ability to capture and emit single charge carriers from the conducting channel, disrupting the electrostatics and characteristics of the device. Due to the comparatively large amount of charge carriers in the channel, in large devices, a single capture or emission event has a small impact on the MOSFET parameters. In stark contrast, in nanoscale MOSFETs, such events affect device performance severely [6–8]. Depending on the position of the defect, the percolation path is disturbed as can be seen in the center and right panels of Fig. 4.
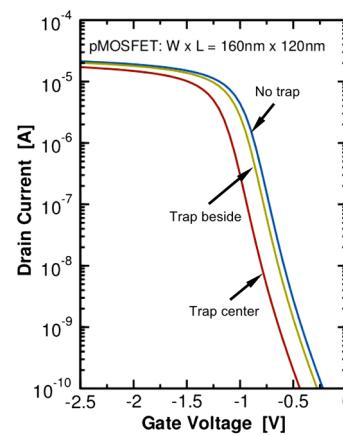

Fig. 5. Schematic $I_D–V_G$ characteristics of a pMOSFET for the three cases shown in Fig. 4. The more the current flow is disturbed by a trap, the more V_{th} shifts, the subthreshold slope decreases and the on-current reduces

This causes a V_{th} shift, a degradation of the subthreshold slope and a reduction of the on-current as summarized in Fig. 5. Even one single active defect may shift V_{th} by a detrimental value, thus, change the transistors behavior and dynamics in digital circuits dramatically. Such a shift of transistor characteristics can endanger the correct interaction with other components, which makes the circuit less reliable and more likely to fail.

In order to understand the root cause of degradation mechanisms, a large effort has been put into studying single defects related to the failure of devices. Figure 6 shows the recently proposed non-radiative multi-phonon model for one particular defect configuration in the oxide including its interplay with hydrogen [10]. The latter has to be considered since hydrogen is the most abundant element, very reactive and everywhere in the MOSFET due to processing steps like forming gas annealing. The interplay between hydrogen, charge carriers and the particular atomic configuration of the defect can be summarized as a number of clearly discernible configurations (states) [6, 9]. Depending on the position of its energy level within the band gap and its particular properties, an active defect can capture or emit charge carriers which can be described by transitions between four states. However, due to movement of or reaction with hydrogen, the defect can also be deactivated or transformed to a precursor state, both of which being neutral and not contributing to MOSFET parameter changes under operating conditions.

The deep understanding of degradation mechanisms in conventional MOSFETs is one basis for new approaches in transistor design.
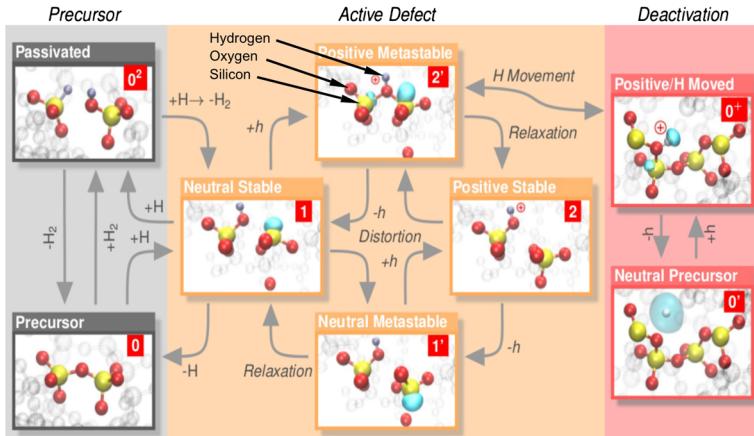
Fig. 6. Non-radiative multi-phonon model including hydrogen exchange describes the various states of a single oxide defect. The shown defect is a hydroxyl-E' center, which is a hydrogen related defect in the amorphous oxide (a hydrogen atom is attached to a bridging oxygen atom) [10]
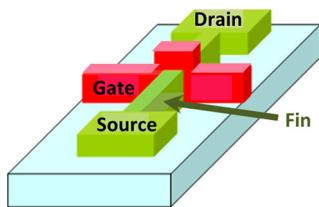

Fig. 7. Non-planar double-gate field-effect transistor (FinFET): The gate is wrapped around the channel which allows for a better electrostatic control over the channel [14]

Fig. 8. Tunnel field-effect transistor (TFET): At a sufficient gate bias a tunnel current flows across the device [15]

For example, by using different materials and designs than in CMOS technologies the root cause for degradation mechanisms like interface defects can be avoided and a profound development of future electronic circuits can meet the challenges posed by limits due to fundamental physics of solids.

4. Transformation to nanotechnology

Recently, researchers in industry and academia have put an effort in finding new geometries and materials that will help in overcoming the challenges due to downscaling. Already in mass production, for example, non-planar multi-gate field-effect transistors (FinFETs) have a cleverly designed gate structure shown in Fig. 7; in contrast to planar devices, the gate is wrapped around the channel instead of being processed on top of it. This allows for a better electrostatic control over the channel, a reduction of the channel leakage current and a decrease of the switching time, making them also appropriate for use in memory cells. Additionally, high-k (high relative permittivity κ, also ϵ_r) materials, e.g., hafnium oxide (HfO$_2$) are used instead of SiO$_2$, which only has a relatively small permittivity and thus has to be very thin for good electrostatic control over the channel [11, 12]. Due to the fact that $C_{ox} \propto \kappa/t_{ox}$, an increase of κ allows an increase of t_{ox} without changing the capacitance. At the same time $E_{ox} \propto 1/t_{ox}$, and thus an increase of t_{ox} makes challenges associated with high E_{ox} less severe.

Nevertheless, even with these improvements, the fundamental physics of solids will limit downscaling of chip structures again soon since ultra-scaled FinFETs suffer from the same detrimental mechanisms like planar MOSFETs [13]. Innovative ideas for future solutions have been suggested, for example the tunnel field-effect transistor (TFET). This transistor technology is based on the quantum mechan-
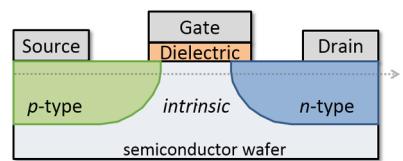
ical tunneling effect. As shown in Fig. 8, the source and drain terminals are doped of opposite type. At a sufficient gate bias, band-to-band tunneling occurs and electrons from the valence band of the p-type region tunnel into the conduction band of the intrinsic region which results in a current flow across the device. Taking advantage of this effect opens up new possibilities for transistor design. In principle, transistors using band-to-band tunneling for injecting charge carriers into the channel would reach better subthreshold swings—less than 63 mV/decade of current—leading to significant power savings and allowing for further downscaling. However, the experimental realization of a field-effect transistor with $SS < 63$ mV/decade of current has not been demonstrated so far.

A highly inspiring idea for future transistor technologies is based on the revolutionary discovery of the field-effect in 2D materials as shown in Figs. 9 and 10. Depending on their chemical composition, such 2D layers can have metallic, semiconducting or insulating properties, as has been demonstrated during the last years. One important class of materials are the transistor metal dichalcogenides (TMD). There the molecules are formed by transition metal atoms, e.g., Mo and of chalcogen atoms, e.g., S bound covalently to, e.g., MoS$_2$ [17]. Hundreds of transition metal atoms and chalcogen atoms combinations are available and each of them shows different properties. One remarkable advantage of TMDs is that transistors can be built by "sticking" together layers with different properties without creating structural defects related to dangling bonds at the interfaces as they occur between crystalline and amorphous materials in conventional CMOS transistors. The reason for this is that adjacent layers of TMDs are weakly held together by van der Waals forces. In contrast to conventional CMOS transistors, where structural defects at the interface strongly affect the performance, degradation caused by such defects does not play a role here. Therefore, this technology might be one innovation towards more reliable transistors even in the several nanometer regimes.
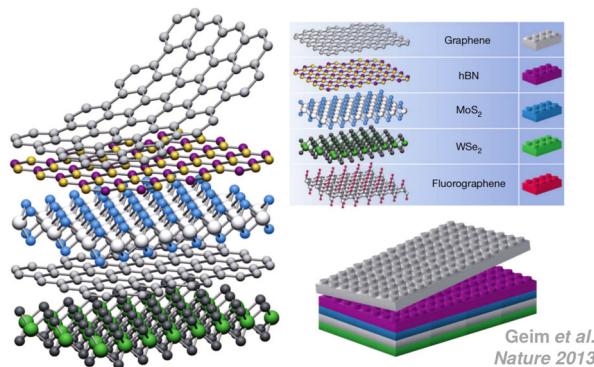
Geim *et al.*
Nature 2013

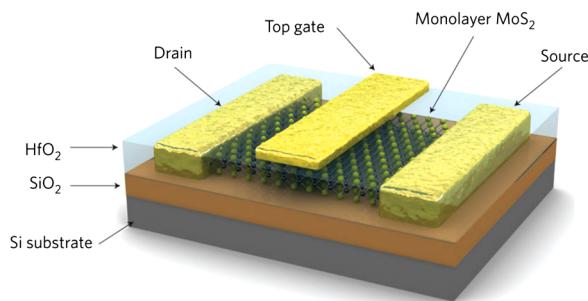**Fig. 9. Combination of 2D materials as transistor design technology [16]**

Fig. 10. Field-effect transistor with a channel layer of MoS$_2$, a transition metal atom layer sandwiched between two layers of chalcogen atoms. Layers hold together by van der Waals interactions [17]

Although approaches for constructing improved transistors have been made, technologies which can enable scaling transistor dimensions to the sub deca-nanometer regime and further are not yet ready for mass production. Reliability issues, power loss, temperature instability and many other detrimental effects have to be eliminated first. Therefore, further research in this field is inevitable and will open the doors towards new approaches for the improvement of digital circuits. These improvements will in a next step enable new applications which are important for many aspects of our lives, such as healthcare, the automotive sector, and security.

5. Conclusions

The pressure on industry to keep the production costs for high-efficiency electronic devices low has led to an excessive downscaling of the most important part in digital circuits, the MOSFET. As a consequence, the limits of physics have been reached several times during the past two decades which has resulted in severe reliability issues, power loss, and instabilities. Although innovative approaches have been made recently in order to meet some of these issues, an

holistic solution which will enable the further increase of complexity in digital circuits has not been introduced yet. However, ideas for new transistor technologies based on other materials and designs than used in CMOS technology have already been suggested and might one day open the door to future applications. With further research, these new technologies can guarantee affordable electronic devices in the future.

References

1. Moore, G. E. (1998): Cramming more components onto integrated circuits. Proc. IEEE, 86(1), 82–85.
2. https://www.wikipedia.org/, Field-effect transistor, 21 September 2017.
3. Schanovsky, F., Grasser, T. (2013): Bias temperature instability for devices and circuits (pp. 379–408). New York: Springer.
4. Bina, M., Triebl, O., Schwarz, B., Karner, M., Kaczer, B., Grasser, T. (2012): Simulation of reliability on nanoscale devices. In Proc. of the international conference on simulation of semiconductor processes and devices, SISPAD (pp. 109–112).
5. Gerrer, L., Hussin, R., Amoroso, S., Franco, J., Weckx, P., Simicic, N., Horiguchi, N., Kaczer, B., Grasser, T., Asenov, A. (2015): Experimental evidences and simulations of trap generation along a percolation path. In Proc. of the 2015 45th European solid state device research conference (pp. 226–229).
6. Grasser, T. (2012): Stochastic charge trapping in oxides: from random telegraph noise to bias temperature instabilities. Microelectron. Reliab., 52, 39–70.
7. Ghetti, A., Compagnoni, C. M., Spinelli, A. S., Visconti, A. (2009): Comprehensive analysis of random telegraph noise instability and its scaling in deca-nanometer flash memories. IEEE Trans. Electron Devices, 56(8), 1746–1752.
8. Franco, J., Kaczer, B., Toledano-Luque, M., Roussel, P. J., Mitard, L. A. R. J., Witters, L., Chiarella, T., Togo, M., Horiguchi, N., Groeseneken, G. (2012): Impact of single charged gate oxide defects on the performance and scaling of nanoscaled FETs. In Proc. international reliability physics symposium, IRPS (pp. 1–6).
9. Wimmer, Y., El-Sayed, A.-M., Gös, W., Grasser, T., Shluger, A. (2016): Role of hydrogen in volatile behaviour of defects in SiO$_2$-based electronic devices. Proc. R. Soc., Math. Phys. Eng. Sci., 472, 1–23.
10. Grasser, T., Waltl, M., Gös, W., Wimmer, Y., El-Sayed, A., Shluger, A., Kaczer, B. (2015): On the volatility of oxide defects: activation, deactivation, and transformation. In Proc. international reliability physics symposium, IRPS (pp. 5A.3.1–5A.3.8).
11. Wilk, G. D., Wallace, R. M., Anthony, J. M. (2001): High-K gate dielectrics: current status and materials properties considerations. J. Appl. Phys., 89, 5243.
12. Robertson, J. (2001): High dielectric constant oxides. Eur. Phys. J. Appl. Phys., 28, 265–291.
13. Rzepa, G., Waltl, M., Goes, W., Kaczer, B., Grasser, T. (2015): Microscopic oxide defects causing BTI, RTN, and SILC on high-k FinFETs. In Proc. of the international conference on simulation of semiconductor processes and devices (SISPAD) (pp. 144–147).
14. https://www.wikipedia.org/, Multigate device, 1 October 2017.
15. https://www.wikipedia.org/, Tunnel field-effect transistor, 15 June 2017.
16. Geim, A. K., Grigorieva, I. V. (2013): Van der Waals heterostructures. Nature, 499, 419–425.
17. Radisavljevic, B., Radenovic, A., Brivio, J., Giacometti, V., Kis, A. (2011): Single-layer MoS2 transistors. Nat. Nanotechnol., 6, 147–150.

Authors

Bianka Ullmann
received her diploma degree in physics from the TU Wien (Vienna University of Technology) in 2012. She has been working as a project assistant at the TU Wien since 2013 and joined the Institute for Microelectronics in January 2014. Her research focus lies on the characterization of single oxide defects in MOSFETs especially after mixed hot carrier and bias temperature instability stress conditions. Currently, she is working on her Ph.D. thesis.

Tibor Grasser
is an IEEE Fellow and currently head of the Institute for Microelectronics at TU Wien. He has edited various books, e.g. on the bias temperature instability (Springer) and hot carrier degradation (Springer), is a distinguished lecturer of the IEEE EDS, has been involved in outstanding conferences such as IEDM, IRPS, SISPAD, ESSDERC, and IIRW, is a recipient of the Best and Outstanding Paper Awards at IRPS (2008, 2010, 2012, and 2014), IPFA (2013 and 2014), ESREF (2008) and the IEEE EDS Paul Rappaport Award (2011). He currently serves as an Associate Editor for Microelectronics Reliability (Elsevier).