Materials Science Forum
ISSN: 1662-9752, Vol. 1090, pp 171-178
doi:10.4028/p-84nz7l

Submitted: 2022-09-04
Revised: 2023-01-10
Accepted: 2023-01-12
Online: 2023-05-31

Complications of Charge Pumping Analysis for Silicon Carbide MOSFETs

Jan Lettens[1,2,3,a*], Fredrik Allerstam[4,b*], Jimmy Franchi[4,c*], Sofie Cambré[2,d*], Etienne Goovaerts[2,e*], Henk Vrielinck[3,f*], Peter Moens[1,g*]

[1]onsemi, Westerring 15, 9700 Oudenaarde, Belgium

[2]Physics Department, University of Antwerp, Universiteitsplein 1, B-2610, Antwerp, Belgium

[3]Ghent University, Krijgslaan 281-S1, 9000 Gent, Belgium

[4]onsemi, Isafjordsgatan 32C, 16440 Kista, Sweden

[a]jan.lettens@onsemi.com, [b]fredrik.allerstam@onsemi.com, [c]jimmy.franchi@onsemi.com, [d]sofie.cambre@uantwerpen.be, [e]etienne.goovaerts@uantwerpen.be, [f]Henk.Vrielinck@ugent.be, [g]peter.moens@onsemi.com

Keywords: defect characterization, trap modeling, charge pumping

Abstract. Unexpected behavior is observed when charge pumping is performed on silicon carbide MOSFETs with a thermally grown silicon dioxide gate dielectric. Supported by experimental evidence, two root causes are proposed: the trap density and the channel non-equilibrium. These are difficult to overcome experimentally due to limitations by oxide breakdown and doping variation along the channel, respectively. A correct interpretation then requires a 2D model.

Introduction

Charge pumping is a powerful trap characterization technique [1]. Part of its success originates from the existence of practical models that allow interpreting experiments in a fast and simple way [2]. Moreover, the technique is very sensitive, boasting single trap detectability [3] while requiring only commonly available equipment. Similarly, it does not require a special test structure; a simple lateral MOSFET is a good option. Since all suitable structures provide access to both types of carriers, *i.e.* electrons and holes, the method allows for studying most traps across the bandgap, in particular important for wide bandgap materials. Since the method can be applied to a lateral MOSFET, it is an excellent fit to study quasi in-situ mobility degradation under stress.

Despite this potential, charge pumping on silicon carbide has not yet produced many first insights into the traps at the interface and sometimes discrepancies are found when comparing with other techniques. In this paper, some assumptions are discussed that were previously made for silicon [2,4] but are sometimes invalid for silicon carbide. An example where these assumptions are invalid is when the dielectric is thermally grown silicon dioxide. Interpreting silicon carbide results with the silicon model can lead to wrong conclusions. A simple numerical alternative is presented, and its shortcomings are discussed.

Experimental

Sample preparation. Lateral n-channel MOSFETs have been produced with a gate width of 700 μm and various channel lengths L ranging from 1 to 400 μm. Their cross-section is pictured in Fig. 1. The gate oxide was grown with thermal oxidation in a dry ambient (O_2) to a thickness of 40 nm. All devices were made on highly nitrogen doped 4° off-axis, 4H-SiC substrates. On these, 10 μm silicon carbide was grown epitaxially with a nitrogen doping concentration of 10^{16} cm^{-3}. P-type doping for the well was done by aluminum implantation to a depth of 1 μm and a concentration of 10^{17} cm^{-3}.

Measurement method. Charge pumping was performed using a Keithley S4200 semiconductor characterization system equipped with a 4200-PG2 dual channel pulse generator module. The measurement setup is shown in Fig. 1. In such a charge pumping experiment, a voltage is applied to

the gate which alternates between a base voltage (V_B) and the sum of the base voltage and the pulse amplitude (V_A) at frequencies (f) between 100 Hz and 10 MHz. Unless explicitly mentioned, the rise time (t_r) and fall time (t_f) of the gate signal are 100 ns and the dutycycle is 50%. Currents are measured on all terminals of the MOSFET except on the gate, as indicated by the subscript mentioned in Fig. 1 (SD for source-drain, B for bulk and SUB for substrate). After the measurement, these currents (I) are converted to a density, with units of cm^{-2}, by dividing by frequency, gate area and the absolute value of the electron charge. Note that in the experiment, the source and drain are electrically shorted. All measurements were performed at 25°C.

Simulation method. Differential equations were solved numerically with an explicit, adaptive step-size Runge-Kutta method (5th order embedded Dormand-Prince). Since a trap cannot be filled above its maximum concentration, this was used as an extra, absolute error estimate. The charge pumping current was determined by counting charge carrier capture and subtracting charge carrier emission.

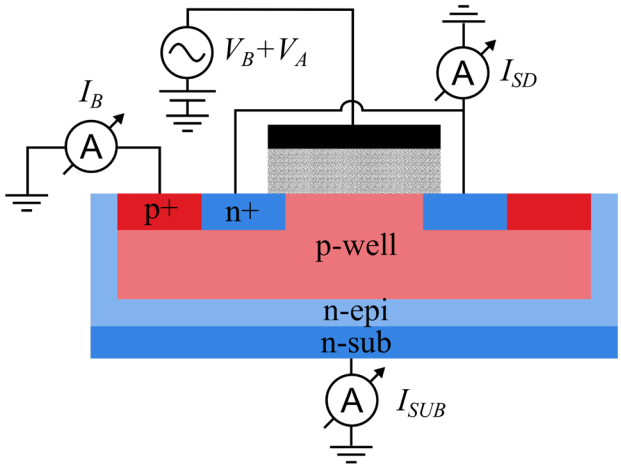

Fig. 1: Cross-section of the lateral MOSFET and schematics of the measurement setup. Currents are measured at the bulk (I_B), source/drain (I_{SD}) and substrate (I_{SUB}) contacts while supplying a DC and AC voltage to the gate (V_B+V_A).

Discussion

Theory. During charge pumping, the voltage applied on the MOSFET gate is repeatedly pulsed between a low voltage (V_B), to attract holes to the channel, and a high voltage (V_B+V_A), to attract electrons. When a trap captures an electron during one part of the cycle and a hole during the other, recombination takes place, which leads to a direct current from the bulk to source and drain, which is proportional to the frequency of the alternating voltage on the gate.

The phenomenon was discovered by Brugler and Jespers [1]. Following this early work, Kaden and Reimer [5] explained it in more detail using the theory of non-steady-state emission of carriers. Afterwards, Groeseneken *et al.* [2,4] developed a practical method to characterize traps. They calculated the energy range ΔE, in which traps contribute to the charge pumping current, as a function of experimental settings. This was done by finding the characteristic trap energies for electrons $E_{em,e}$ and holes $E_{em,h}$ that are unable to emit fast enough to remain in equilibrium with the channel under a changing gate voltage:

$$\Delta E = E_{em,e} - E_{em,h} = 2k_B T \ln\left[C \frac{V_A}{\sqrt{t_r t_f}\,(V_{T,CP} - V_{FB,CP})}\right]. \quad (1)$$

Where k_B is the Boltzmann constant, T the temperature, C a constant dependent on material properties with a unit of time, t_r and t_f the rise and fall time of the pulse respectively, V_A the pulse amplitude, $V_{T,CP}$ the charge pumping threshold voltage and $V_{FB,CP}$ the charge pumping flatband voltage. $V_{T,CP}$ and $V_{FB,CP}$ are dynamic quantities, which shift as traps are filled and emptied. For now, they can be approximated by the static threshold voltage V_T and flatband voltage V_{FB} as is done for silicon.

Trap density. The relation between the charge pumping current I_{CP} and the number of recombination events N is given by the well-known formula [1]:

$$I_{CP} = q f N. \tag{2}$$

Where q is the elementary charge and f the charge pumping frequency. Assuming recombination only happens via traps, Eq. 2 can be used as a lower bound for the number of traps in the energy range ΔE. Eq. 2 only permits to extract the total number of traps in the special case where charge can be supplied to all of them. To estimate the number of charges that can be supplied to a MOS structure, the upper bound $C_{ox}V_A/q$ can be used. Where C_{ox} is the parallel plate capacitance and V_A the pulse amplitude applied to the gate. For 40 nm SiO_2 and a pulse amplitude of 10 V this results in an areal density of approximately 5 x 10^{12} cm^{-2} elementary charges.

This number is much larger than trap densities typically encountered in silicon, which are of the order of 10^{11} cm^{-2} [2] or lower. However, it is smaller than trap densities reported on silicon carbide, which can easily reach 10^{13} cm^{-2} [6]. Consequently, this prevents simply extracting the total trap density from the observed charge pumping current when large trap densities are present.

This is experimentally reflected in the behavior of the charge pumping current for different pulse amplitudes, see Fig. 2, top left, where the density of charges pumped per cycle is presented in function of the base voltage. An increase in the density of charges pulsed (*i.e.* $C_{ox}\Delta V_A/q$) directly corresponds to a greater density of charges per cycle (N_B). While Eq. 1 predicts an increase of the charge pumping current for an increasing pulse amplitude, this cannot be a complete explanation since it predicts that, by lowering the rise and fall times, the effect of the increased pulse amplitude can be compensated for. However, in Fig. 3 a comparison between the rise and fall time behavior for two different pulse amplitudes is made. The peak density of charges per cycle pumped with a pulse amplitude of 20 V and 10^{-5} s rise and fall times is still higher than with a pulse amplitude of 10 V and 10^{-7} s rise and fall times. Eq. 1 predicts ΔE to decrease, if the denominator of the fraction in the logarithm is increased by a larger amount ($\sqrt{t_{r,2}t_{f,2}} = 100\sqrt{t_{r,1}t_{f,1}}$) than the numerator ($V_{A,2} = 2V_{A,1}$).

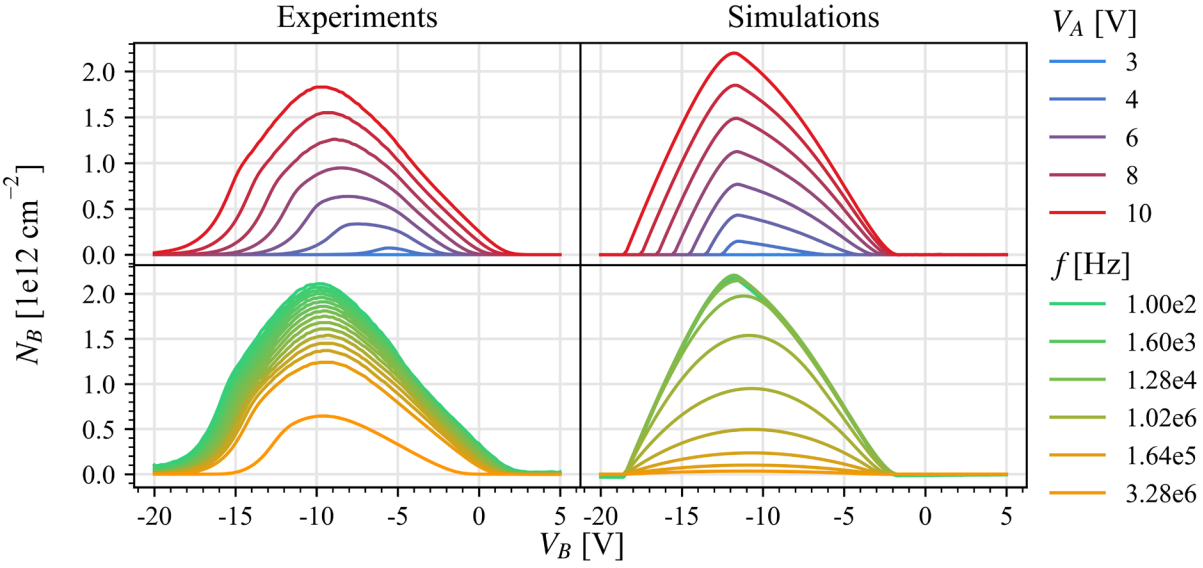

Fig. 2: The density of charges per cycle (N_B) versus base voltage (V_B) for different pulse amplitudes V_A (*top row*, f = 10 kHz) and different frequencies f (*bottom row*, V_A = 10 V). *(left)* Measurements on a MOSFET with L = 1 µm. *(right)* Simulations with one type of acceptor trap (N_T = 6 x 10^{12} cm^{-2}, $c_n = c_p$ = 10^{-7} cm^2s^{-1}, e_n = 5 x 10^5 s^{-1} and e_p = 1 s^{-1}), V_T = -5 V, and V_{FB} = -10 V. *(top left)* From a pulse amplitude of 4 V onwards, every volt increase of the pulse amplitude pumps ~3 x 10^{11} cm^{-2} more traps, close to $\Delta V_A C_{ox} \approx$ 5 x 10^{11} cm^{-2}. *(bottom left)* The number of recombination events per cycle decreases for increasing frequency. *(top and bottom right)* The simulations show qualitatively similar behavior as the measurements.

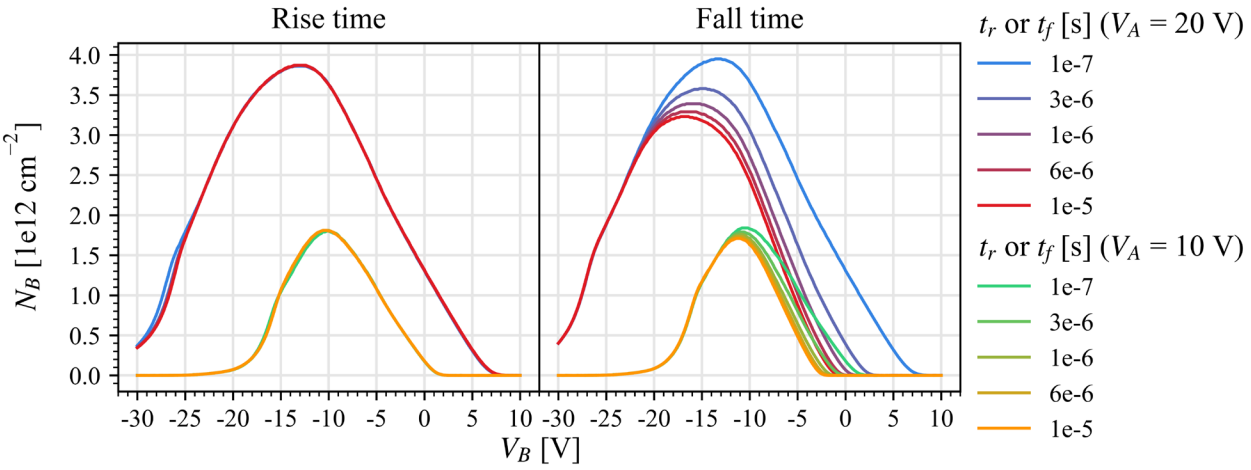

Fig. 3: The density of charges per cycle (N_B) versus base voltage (V_B) for different rise and fall times and two different pulse amplitudes, V_A = 10 V (green to orange) and 20 V (blue to red). L = 1 µm, f = 10 kHz. *(left)* The rise time hardly affects the peak. *(right)* The fall time does affect the peak. Less traps are pumped for the fastest fall time with 10 V pulse amplitude than for the slowest fall time with 20 V pulse amplitude, which is not predicted by Eq. 1.

In the derivation of Eq. 1 it was assumed that traps can be filled and emptied completely. This is a valid assumption for small trap densities where the dynamic range of the charge pumping threshold voltage $V_{T,CP}$ and the charge pumping flatband voltage $V_{FB,CP}$ is small compared to the pulse amplitude V_A. For large trap densities this assumption is not valid. This difficulty could be circumvented if it was possible to make $C_{ox}V_A/q \gg N_T$ by choosing an appropriate pulse amplitude or oxide thickness. But on silicon dioxide this is limited by oxide degradation and breakdown.

Capture rate. In Fig. 2 (bottom left) the density of charges per cycle versus base voltage is plotted for different frequencies. This density decreases as the frequency increases. Also this can be a consequence of the large trap density. Consider the rate equation:

$$\frac{dn_t}{dt} = c_n n_{SiC}(N_T - n_t) - e_n n_t \,, \tag{3}$$

which describes trap communication with the conduction band. N_T is the total empty trap density, n_t the occupied trap states, n_{SiC} the surface concentration of electrons in the silicon carbide, c_n the electron capture rate and e_n the electron emission rate.

For deep traps at moderate temperatures emission can be neglected. Assuming the capture rate c_n is constant, traps are initially empty and $N_T >> n_{SiC}$ then $N_T >> n_t$ and $N_T - n_t \approx N_T$, hence:

$$\frac{dn_t}{dt} \propto n_{SiC} \,. \tag{4}$$

Consequently, capture will quickly decrease as the density of electrons in the silicon carbide n_{SiC} decreases. Eq. 3 can be used to model the charge pumping current if the hole surface concentration p_{SiC}, hole capture rate c_p and emission rate e_p are also included:

$$\frac{dn_t}{dt} = c_n n_{SiC}(N_T - n_t) - e_n n_t \; - \; c_p n_t p_{SiC} + e_p(N_T - n_t). \tag{5}$$

Eq. 5 only describes the evolution of a single trap. In a more realistic case, different traps exist. This results in a system of equations, coupled through the surface concentration of charge carriers, n_{SiC} and p_{SiC}. Even with a single trap, Eq. 5 reproduces characteristic charge pumping features, such as the observed amplitude and frequency dependencies, which are presented in Fig. 2 on the right side.

It is certainly not excluded that such behavior is the reflection of traps which intrinsically charge slowly. However, it also seems linked with the trap density if the latter is sufficiently large.

Model dimensions. Given the success of Eq. 5, it is tempting to continue with a 0D spatial model [2,4]. It seems reasonable to ignore the spatial dimension from the interface to the gate (1D). For one, reports [6, 7] indicate that traps are near the interface, so the effect of their spatial distribution on the electrostatic potential will be small for thick oxides. Another incentive to consider this dimension is trap-to-trap charge transfer. From the experimental data it is not evident this needs to be considered. In that case, it is not interesting to add complexity if a simpler model suffices.

Problems occur when the spatial dimensions from the interface to the bulk and along the channel are considered (2D). The latter is important if trap behavior is not uniform along the channel dimension of the MOSFET or when the channel does not reach equilibrium on the experimental timescale. Nonuniform trap behavior can result from different variations along the channel (*e.g.,* oxide thickness, doping, trap density, trap characteristics, *etc.*). The channel can be out of equilibrium for two reasons: when charge cannot get in fast enough or when charge cannot get out fast enough.

Transport out. When electrons in the channel cannot escape towards the source or drain before holes accumulate, they will recombine and contribute to the charge pumping current. This has previously been defined as the 'geometric component' [1,8], since long transistors are primarily affected, see Fig. 4. Such a contribution to the recombination is regarded as undesirable, as it is not trap related. The typical strategy is to avoid it experimentally. For silicon carbide [9], this leads to constraints on the rise time and fall time, which severely limit the traps that can be characterized. It is clearly problematic when the time ranges for channel evacuation and trap emission overlap. Inspecting the geometric component more closely, it seems that the situation on silicon is again a special case, where problems can be avoided due to the combination of a high channel mobility and a small trap density.

Consider a high channel mobility with a large trap density. It was established that not all traps are characterized when insufficient carriers are supplied. For low base voltages, electrons will be the limiting carrier and for high base voltages, holes. This leads to the rising and falling edge of the charge pumping curve. At the rising edge, when electron capture slows down and hole emission starts to dominate, the emitted holes transport back to the bulk and do not recombine. At the falling edge, as hole capture slows down, traps continue to empty by electron emission. Some of these electrons will diffuse into the bulk and recombine. This is bulk recombination, but trap related.

A channel with low mobility is always problematic. Whenever the emission rate is much faster than the channel evacuation rate, traps will hide behind the geometrical component. As an extreme example, consider the case where transport is limited by a trap-assisted mechanism, such as by hopping transport or impurity band conduction. Charges will effectively move along the channel from trap to trap. These traps can never be characterized without measuring the 'geometric component'.

Very fast states exist at the nitrided silicon carbide silicon dioxide interface [10]. It is not obvious this is relevant for the previous example. The response time of these states was reported to be 5 ns. Coincidentally, an upper bound estimate for the transit time t_c is also 5 ns for a 1 μm long channel. This is based on the following equation [11], valid under gate overdrive and saturation:

$$t_c = \frac{L^2}{\mu(V_G - V_T)}. \tag{6}$$

In this equation, a mobility μ of 20 $cm^2\ V^{-1}\ s^{-1}$ was used, and 10 V gate overdrive V_G-V_T. This must be an upper bound: the gate is turned off, there is only self-induced drift and diffusion, and a typical peak field effect mobility for a nitrided sample was used. Based on this simple calculation it cannot be excluded that these fast states can be separated from the geometrical component.

The channel evacuation rate can be increased by making the transistor shorter, but there is a lower bound before channel nonuniformity needs to be considered. Using the method described in [12] the effective channel length is determined to be 0.3 μm shorter than the mask length. This reduction in

channel length must be due to doping variation. TEM cross-sections (not shown) confirm a uniform oxide thickness on the lowly doped silicon carbide. The oxide mask variation can also be ruled out, since the process is not self-aligned and there is 2 µm gate-to-source and gate-to-drain overlap. A reduction of 0.3 µm is large for a short channel transistor (1 µm), which makes it questionable to ignore the threshold voltage variation along the channel.

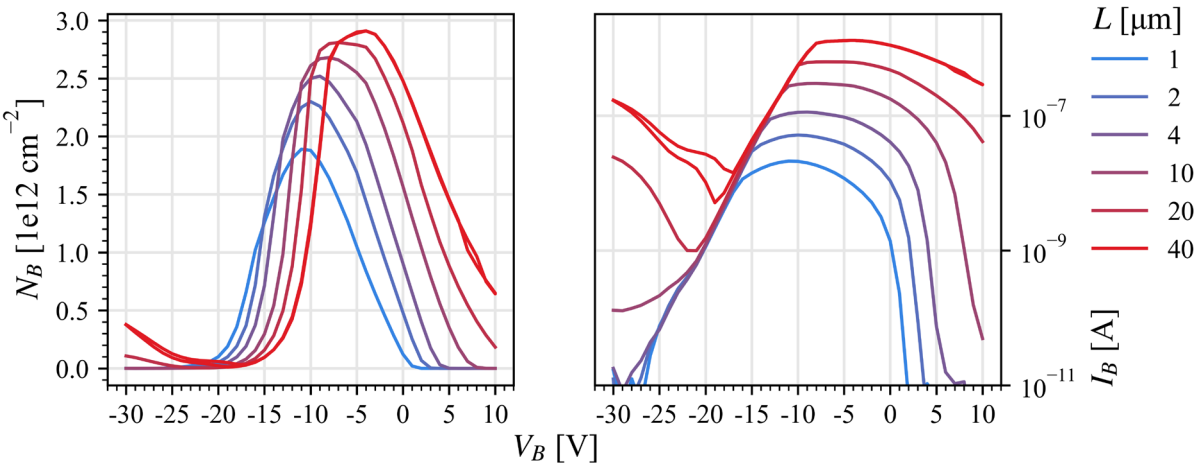

Fig. 4: *(left)* The density of charges per cycle (N_B) versus base voltage (V_B) for different gate lengths L, V_A = 10 V, f = 10 kHz. For longer transistors there is a substantial amount of current at high base voltages. *(right)* Charge pumping current for different gate lengths on a logarithmic scale. Same data as on the left. The charge pumping curves at low base voltages follow the same trend for all gate lengths. The abnormal behavior for longer transistors is explained later.

Transport in. Fig. 4 on the right shows the charge pumping current instead of the density of charges per cycle. The rising edge of the curves follow the same trend for all lengths. If this was because of a variation of the threshold voltage along the channel, then it should not continue for as long. Otherwise, the rising edges should only partly overlap, since each transistor shares this 0.3 µm long region.

Therefore, it is required to consider how charge enters the channel. Consider another extreme example: if the electron in Fig. 5 passes trap 1 and 2, which one will it occupy? When the traps have large capture rates, the electron will simply occupy the first trap it encounters. If the emission rate is low, there is no significant redistribution. Looking at the situation along the channel dimension, the electrons fill the traps they first encounter. This is consistent with our observations: electrons enter the channel in the same way for all transistor lengths, so the behavior of the smaller transistor is recognized as part of the larger transistor. As long as charge does not reach the center of the transistor, the charge pumping will be independent of channel length.

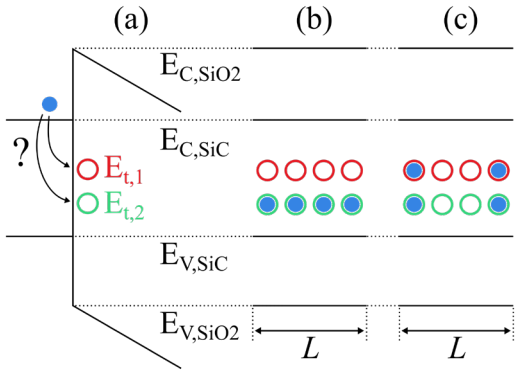

Fig. 5: $E_{C,SiC}$, $E_{V,SiC}$, $E_{C,SiO2}$ and $E_{V,SiO2}$ are the silicon carbide and silicon dioxide conduction and valence bands respectively. $E_{t,1}$ and $E_{t,2}$ are trap energies. (a) Which state will the electron occupy? In equilibrium, according to Fermi-Dirac statistics, the lowest available energy state, $E_{t,2}$ has the highest chance to be occupied. When the capture rate is large compared to the diffusion coefficient, the electron will most likely occupy the first trap encountered. If the barrier for re-emission is also large, then the channel might not find equilibrium on the timescale of the experiment. This makes situation (b) much less likely than situation (c) when for example a large concentration of deep donor trap states is present (which Fig. 3 indicates).

Deep depletion effects. The channel nonequilibrium causes another unexpected phenomenon, visible in the substrate current, Fig. 6 (bottom left). When charge cannot respond to a rising gate signal, the p-well continues depleting, which results in forward biasing of the p-well/n-epi junction. This signal is indeed sensitive to a rise time variation (Fig. 6, middle), not to a fall time variation (Fig. 6, right). This can be due to trapped holes or when electrons from the source/drain are too slow to reach the center of the transistor. This indicates large fields in the MOSFET under switching as a result of hole traps, low electron mobility or both. This can be important for degradation mechanisms. Also, field dependent emission has been reported [6], which then also demands 2D modeling to characterize.

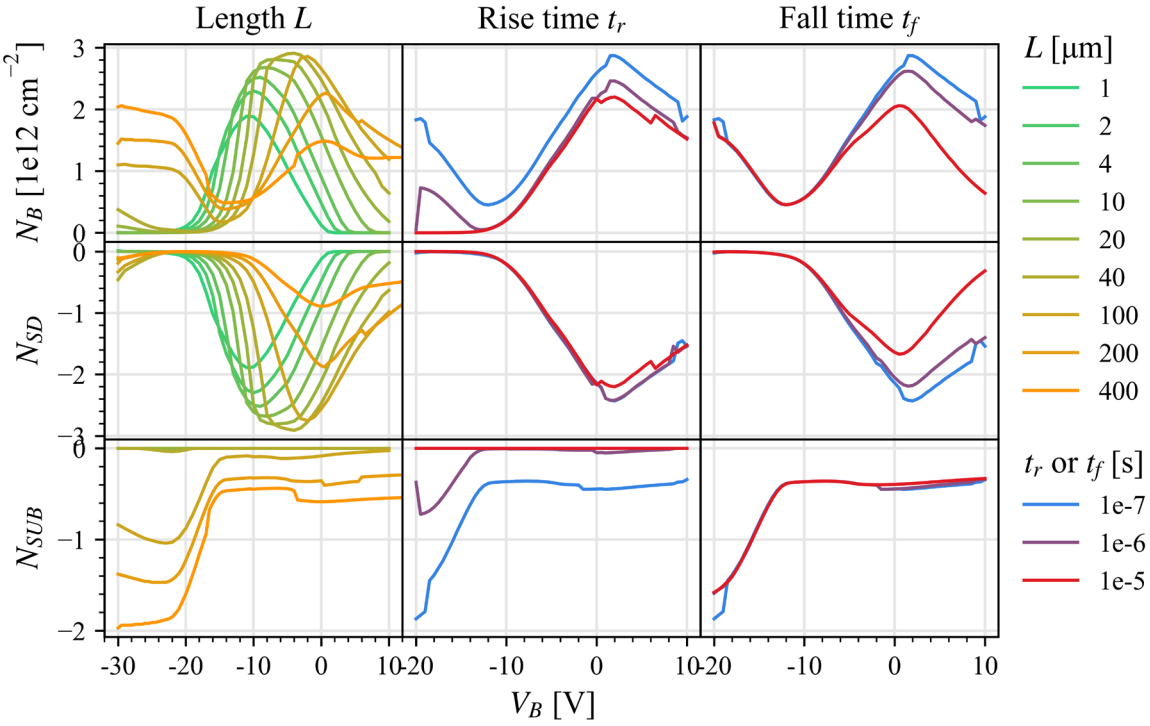

Fig. 6: Density of charges per cycle measured at the bulk (N_B), source/drain (N_{SD}) and substrate (N_{SUB}), V_A = 10V, f = 10 kHz. *(left)* Different lengths, *(middle)* rise times and *(right)* fall times.

Future analysis through modeling. It is possible to detect and characterize traps using the charge pumping technique. However, for accurate characterization, numerical modeling is necessary. The system of differential equations that needs to be solved has a wide range of time constants; easily ranging from milliseconds to picoseconds. Even worse, sometimes multiple charge pumping periods are necessary to reach equilibrium. This makes it a stiff differential equation. These are normally solved using implicit methods or explicit methods with adaptive step-size. Unfortunately, implicit methods lose some of their benefits since small time steps are necessary to accurately determine the charge pumping current during the very fast capture.

To avoid ambiguous interpretation a wide range of experimental conditions need to be fitted simultaneously. This is feasible for a 0D approximation. But such an approximation seems inaccurate, given the experimental observations. Modeling two dimensions will dramatically increase the computational requirements. There exist trap characterization techniques which avoid these problems. For example, by limiting the excitation signal to small variations like in conductance spectroscopy [10] or to low frequency (DC) as in thermal dielectric relaxation current [6,13].

Conclusion

The models for charge pumping developed on silicon contain at least two assumptions which may not always be valid for silicon carbide. One of them is that there is enough charge to completely fill and empty traps quickly. Incorporating the relatively large trap density into a 0D numerical model reproduces some of the anomalies observed in experiments. But there are more discrepancies evident

when results for transistors with different gate lengths are compared. For long transistors, the data suggests that the channel is no longer in equilibrium. Due to doping variation, there is a lower bound on the smallest channel length that can be used in the uniform channel approximation. It seems necessary to extend the model to 2D. The wide range of timescales make the numerical equations difficult to solve. Still, these results show that such advanced modelling is necessary to understand the atypical charge pumping characteristics of SiC MOSFET devices, and that interpretation solely based on simplified theory may be meaningless.

Acknowledgement

We gratefully acknowledge the financial support of the Flemish Government and Flanders Innovation & Entrepreneurship (VLAIO) through the Baekeland project number HBC.2019.2171.

References

[1] S. J. Brugler and P. G. A. Jespers, "Charge Pumping in MOS Devices", IEEE Trans. Electron Devices, vol. 16, no. 3, pp. 297–302, 1969.

[2] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. De Keersmaecker, "A Reliable Approach to Charge-Pumping Measurements in MOS Transistors", IEEE Trans. Electron Devices, vol. 31, no. 1, pp. 42–53, 1984.

[3] G. V. Groeseneken, I. De Wolf, R. Bellens, and H. E. Maes, "Observation of single interface traps in submicron MOSFET's by charge pumping", IEEE Trans. Electron Devices, vol. 43, no. 6, pp. 940–945, 1996.

[4] G. Van den bosch, P. Heremans, G. V. Groeseneken, and H. E. Maes, "Spectroscopic Charge Pumping: A New Procedure for Measuring Interface Trap Distributions on MOS Transistors," IEEE Trans. Electron Devices, vol. 38, no. 8, pp. 1820–1831, 1991.

[5] G. Kaden and H. Reimer, "Ladungstragereinfang in MOS-Strukturen mit externer Minoritatstragerquelle," Phys. Stat. Sol., vol. 32, pp. 183–194, 1975.

[6] T. E. Rudenko, I. N. Osiyuk, I. P. Tyagulski, H. Ö. Ólafsson, and E. Ö. Sveinbjörnsson, "Interface trap properties of thermally oxidized n-type 4H-SiC and 6H-SiC," Solid. State. Electron., vol. 49, no. 4, pp. 545–553, 2005.

[7] F. Allerstam, H. Ö. Ólafsson, G. Gudjónsson, D. Dochev, and E. Ö. Sveinbjörnsson, "A strong reduction in the density of near-interface traps at the SiO_2 / 4H-SiC interface by sodium enhanced oxidation," J. Appl. Phys., vol. 101, no. 124502, pp. 1–5, 2007.

[8] G. Van den bosch, G. Groeseneken, and H. E. Maes, "On the Geometric Component of Charge-Pumping Current in MOSFET's," IEEE Electron Device Lett., vol. 14, no. 3, pp. 107–109, 1993.

[9] D. Okamoto, H. Yano, T. Hatayama, Y. Uraoka, and T. Fuyuki, "Criteria for accurate measurement of charge-pumping current in 4H-SiC MOSFETs," Mater. Sci. Forum, vol. 600–603, pp. 747–750, 2009.

[10] H. Yoshioka, T. Nakamura, and T. Kimoto, "Generation of very fast states by nitridation of the SiO_2/SiC interface," J. Appl. Phys., vol. 112, no. 024520, pp. 1–7, 2012.

[11] J. B. Kuo, R. W. Dutton, and B. A. Wooley, "MOS Pass Transistor Turn-Off Transient Analysis," IEEE Trans. Electron Devices, vol. 33, no. 10, pp. 1545–1555, 1986.

[12] J. G. J. Chern, P. Chang, R. F. Motta, and N. Godinho, "A New Method To Determine MOSFET Channel Length," IEEE Electron Device Lett., vol. 1, no. 9, pp. 170–173, 1980.

[13] E. Pippel, J. Woltersdorf, H. Ö. Ólafsson, and E. Ö. Sveinbjörnsson, "Interfaces between 4H-SIC and SiO_2: Microstructure, nanochemistry, and near-interface traps," J. Appl. Phys., vol. 97, no. 3, pp. 1–8, 2005.