This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Analysis of ESD Behavior of Stacked nMOSFET RF Switches in Bulk Technology

Matteo Rigato, Clément Fleury, Benedikt Schwarz, Markus Mergens, Sergey Bychikhin, Werner Simbürger, and Dionyz Pogany

***Abstract*— The operation of stacked MOSFET circuit for RF switch application under electrostatic discharge (ESD) conditions is studied by transmission line pulse (TLP) and transient interferometric mapping (TIM) techniques combined with circuit simulation. TLP pulses with 100–840 ns durations were applied to the device composed of 16 stacked multifinger MOSFET blocks with gate and drain resistors, fabricated in a bulk technology. ESD discharge paths related to MOSFET channel and to open-base breakdown have been identified to have dominant role in explaining the complex voltage and current waveforms. The overall circuit behavior during TLP pulses is analyzed taking into account calibrated breakdown measurements on test structures. In order to explain the heat dissipation in MOSFET blocks measured by TIM, additional discharge paths related to block-to-block coupling due to parasitic bipolar action have been modeled and discussed. Besides the analysis of the device behavior, we have also investigated peculiar optical TIM signal related to anisotropic reflectivity and phase response due to dense multifinger block structure.**

***Index Terms*— Breakdown, electrostatic discharge (ESD), optical testing, RF CMOS switch, stacked MOSFETs, transmission line pulse (TLP) technique.**

I. Introduction

RF SWITCHES based on CMOS technology are a cost-effective alternative to GaAs switches [1] in RF front end for mobile applications [2]. However, advanced CMOS technologies present low breakdown voltage (2–3 V). This requires designing the switch devices in a stack configuration [1], [3], i.e., connecting several transistors in series. Using stacked devices is a common practice to enhance the voltage capability of ESD protection devices [9], [10]. However, their gate biasing circuit topology differs to that used in switches. Current leakage in RF switches can be prevented using silicon-on-insulator (SOI) technology with the drawback of significantly higher cost and increased OFF-capacitance C_{OFF} compared to bulk-CMOS-technology. During the manufacturing and operation, such a CMOS switch has to be robust against different disturbing pulses, such as electrostatic discharge (ESD). Such pulses have typically nanosecond rising edge and hundred nanoseconds duration [4]. The introduction of a parallel protection device is not allowed due to RF performance degradation, so the device has to be self-protected. Up to now, the ESD behavior of stacked nMOS RF switches has only been studied in [5] and [6] without giving physical insight into the switch operation. The main issue on RF stacked switches concerns the voltage imbalance among its transistors [7], [8].

We have recently reported on transmission line pulse (TLP) measurements (pulsewidth dependence), transient interferometric mapping (TIM), and failure analysis of multifinger CMOS test structures which are basic building blocks of such RF switch [11].

In this paper, we analyze stacked switch structures, fabricated in a 0.13-μm bulk-CMOS-technology, under ESD-like pulses. Besides understanding overall transient behavior, we investigate the role of device breakdown and parasitic bipolar transistor (BJT) action. Transient electrical analysis using TLP measurements is combined with TIM and correlated with results of compact modeling. In particular, TIM is used to probe dissipated energy distribution in the stacked transistors. Besides device aspects, we have discovered and analyzed that TIM response in such multifinger devices differs from the TIM response of samples with homogeneously reflecting surfaces [12]. This is relevant for laser beam probing of IC chips in general.

This paper is organized as follows. Section II introduces studied structures and used methods. Section III presents calibration measurements on test structures for the purpose of breakdown source determination and demonstration of parasitic BJT action. Section IV with the Appendix presents enhanced TIM signal calibration procedure. Section V shows transient TLP and TIM results on stacked devices. Section VI provides the circuit simulation used to interpret the experiments and gives insight into parasitic BJT action between the stacked nMOS blocks, followed by the conclusion given in Section VII.

II. Device Structure and Experiment

The layout and circuit schematics of the studied stacked RF nMOS device are given in Fig. 1(a)–(c). The stacked

Manuscript received October 23, 2017; revised December 7, 2017; accepted December 28, 2017. The work of B. Schwarz was supported by Austrian Science Fund within the Project NanoPlas under Grant P28914-N27. The review of this paper was arranged by Editor E. Rosenbaum. *(Corresponding author: Matteo Rigato.)*

M. Rigato, C. Fleury, B. Schwarz, and D. Pogany are with the Institute for Solid State Electronics, TU Wien, 1040 Vienna, Austria (e-mail: matteo.rigato@tuwien.ac.at).

M. Mergens is with QPX Gmbh, 85614 Kirchseeon, Germany.

S. Bychikhin is with Alten Gmbh, 80687 Munich, Germany.

W. Simbürger is with Infineon Technologies AG, 85579 Neubiberg, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2018.2789941

0018-9383 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

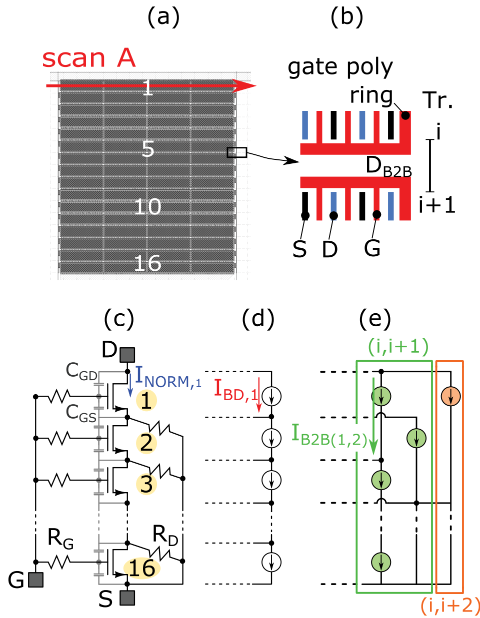

Fig. 1. (a) Layout of the 16-block device where the TIM scan line A on first transistor is indicated. (b) Simplified cropped layout of the region between two consecutive multifinger blocks; the B2B distance (D_{B2B}), gate (G), drain (D), and source (S) fingers are indicated. (c) Equivalent circuit of a basic stacked MOSFET configuration. (d) Additional breakdown current sources at individual transistors, and (e) parasitic next neighbor coupling (i, $i+1$) current sources and further coupling among (i, $i+2$) transistors.

switch consists of 16 MOSFET blocks which have the same layout and all design parameters. This device, here referred as "switch" or "stack," mimics functionality of an RF switch and is fabricated on low-doped Si substrate. The gates of all MOSFET blocks are connected to a common gate terminal via polysilicon resistors (R_G = 400 kΩ) which decouple the gate from RF signal. Under RF operation the gate-to-source voltage V_{GS} is +1.5 V for ON-state condition and −1.5 V for OFF-state condition; the drain-to-source voltage V_{DS} < 2 V and the substrate is biased at −3 V in order to reduce the substrate junction capacitance. R_G determines also the switching RC time constant between the conductive and nonconductive switch states [3]. Each block consists of four multifinger MOSFET subblocks [Fig. 1(a)]. All blocks are surrounded by a polysilicon gate ring which connects all gate fingers together. The width of one finger is 12.5 μm, and the total gate width of a single block is 6.96 mm. The drain resistors (R_D) of 400 kΩ connect the drains to source terminal except the transistor no.1 [see Fig. 1(c)]. The function of R_D is to prevent charge storage during switching OFF the whole switch. Fig. 1(d) and (e) provide model implementations discussed in Section VI.

The switch is provided with separated gate, bulk, drain, and source pads. In particular, the p^+-silicon ring contact for bulk is placed more than 200 μm from the device to represent the same layout conditions like in a product. The long distance between stacked device and bulk contact is because in RF operation the substrate is biased at −3 V in order to deplete it and to reduce the C_{OFF}. Devices with two different block-to-block distances D_{B2B} = 5.68 and 1.84 μm have been analyzed [see Fig. 1(b) for D_{B2B} definition].

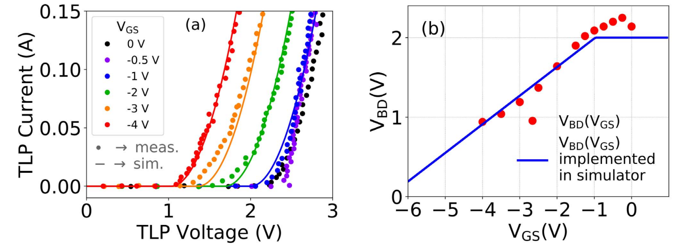

Fig. 2. (a) Measured (dotted lines) and simulated (solid lines) TLP-IVs for 100-ns pulsewidth on a single multifinger transistor (W = 1.95 mm) for different V_{GS} bias. (b) Drain–source breakdown voltage (V_{BD}) (red points) extracted from TLP-IVs in (a) at 1 mA. The blue line shows empirical $V_{BD}(V_{GS})$ function for which the TLP IV results in (a) fits the best.

Additional test structures have been investigated for calibration purposes: 1) multifinger MOSFETs with one single polysilicon ring with total width W = 1.95 mm (150 fingers with 13-μm finger width W_F), similar as in [11] and 2) transistors with finger number N_F = 1, 2, and 4 having W_F = 10 μm have also been analyzed. The device test structures are provided with gate, bulk, drain, and source separated pads. P^+ bulk ring is also placed far away from the device active region as for the stacked transistors. All studied structures have gate length L_G = 0.13 μm and gate oxide thickness of 2.2 nm. The one finger transistor has been studied also for multiple L_G from 0.1 to 10 μm.

A 50-Ω TLP pulser from HPPI has been used for ESD-like stress of the devices [13]. TLP pulses with 1-ns rise time and several pulse widths (PWs) like 100, 450, 650, and 840 ns were used for device stressing. The switch has been probed applying positive pulses at drain of block no.1 versus grounded source of block 16 using GGB Picoprobe probes. The positive stress represents the worst case because the transistors tend to go to breakdown leading to high voltage on them. (The negative stress would cause the MOSFETs to operate in normal operation with positive V_{GS} causing lower voltage.) The gate pad has been grounded by a flexible pitch ground fixture clamp from HPPI. Since we found that grounding of bulk contact has negligible effect to ESD robustness and device behavior, we present results with floating bulk unless explicitly stated otherwise.

The heat energy dissipation in transistors 1–16 during the TLP stress pulse is probed by TIM [14]. In TIM, a probe infrared laser beam enters the chip from the polished backside and scans the device laterally while stressing the device with pulses of 1-Hz repetition frequency. The space and time resolution are 1.5 μm and 3 ns, respectively. The measurements are performed at wafer level. As the optical response of the studied samples differs from the TIM phase signal response of samples with flat reflecting surface, we have performed signal calibration in Section IV.

III. Single Transistor Characterization

The measured TLP–IV characteristics of a single block multifinger MOSFET (W = 1.95 mm) with gate bias in the range 0—(−4 V) are shown by symbols in Fig. 2(a).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

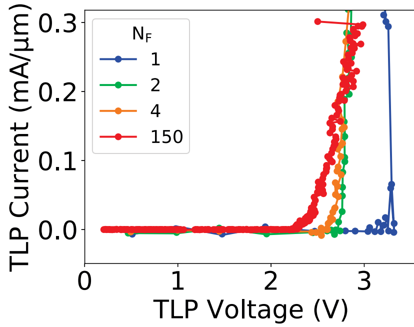

Fig. 3. Normalized TLP-IVs for 100-ns pulsewidth in grounded gate configuration on transistor with different number of fingers (N_F).

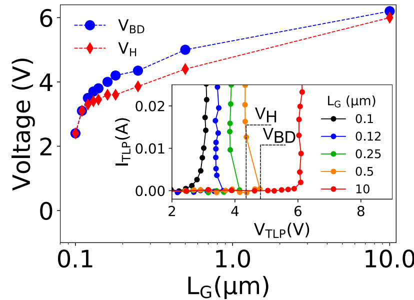

Fig. 4. Breakdown voltage (V_{BD}) and holding voltage (V_H) as a function of L_G measured by TLP on single finger transistor in grounded gate configuration for 100-ns pulsewidth. Selected TLP–IV curves are shown in the inset. The definition of V_{BD} and V_H is shown in one curve.

The negative gate bias was chosen since negative V_{GS} occurs during transients in the switch, see Section VI. For $V_{GS} = 0$, -0.5, -1 V the IV characteristics are nearly identical, while for $V_{GS} < -1$ V they shift toward lower drain breakdown voltage V_{BD}. The extracted V_{BD} decreases linearly for $V_{GS} < -1$ V, see Fig. 2(b) (red dotted points).

Fig. 3 shows the TLP–IV characteristics of small MOSFETs as a function of N_F. A snapback-like characteristics can be seen for $N_F = 1$, while for higher N_F the snapback is not visible. The device voltage decreases as the number of gate fingers increases. Another set of TLP measurements has been done on single finger transistors with L_G variation from 0.1 μm (no snapback) to 10 μm (snapback behavior), see inset of Fig. 4. The TLP breakdown (V_{BD}) and holding voltage (V_H) are plotted in Fig. 4.

Since the substrate contact in these devices has negligible role, we consider our devices as BJTs with open base where a nonsnapback behavior is usually observed [15]. The situation is similar to floating body avalanche breakdown in SOI MOSFETS [16]–[19]. Part of the emitter current, i.e., electrons injected from the source which do not recombine in the base are multiplied in the high field drain region, providing thus the collector current. A snapback-like IV is observed only for long channel devices where intrinsic substrate resistance and current dependence of current gain can play a role [17]. The decrease of V_{BD} with N_F can be explained via finger to finger coupling [20]: the electrons injected from source regions laterally diffuse and are multiplied not just in the drain region of the same MOSFET finger but also in the drain regions

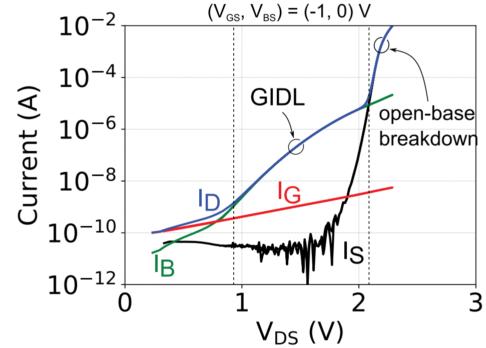

Fig. 5. Four terminal OFF-state dc I–V characteristic for $V_{GS} = -1$ V and $V_{BS} = 0$ V. I_D, I_S, I_B, and I_G stands for drain, source, bulk, and gate currents.

of the neighboring fingers which decreases V_{BD}. For large multifinger blocks ($N_F = 150$), eventually a lower boundary value of V_{BD} is reached [see Fig. 3]. All these results indicate that the studied MOSFETs exhibit parasitic BJT action where coupling via substrate plays a role.

Finally to elucidate the nature of the breakdown voltage decrease at negative V_{GS}, a dc characterization in four terminal configuration has been performed. The selected measurement has been carried out for $V_{GS} = -1$ V and $V_{BS} = 0$ V (see Fig. 5). In the region $V_{DS} < V_{BD} = 2.1$ V, the drain current I_D coincides with the bulk current I_B. The involved physical phenomenon is the gate-induced drain leakage, which is typically observed in transistors with thin gate oxides [21], [22]. This leakage mechanism is due to band-to-band tunneling taking place in the overlap region of the gate oxide with the drain n-well for negative gate bias. For further increase of V_{DS}, the device enters into the open-base avalanche breakdown mode [15].

IV. TIM Calibration

Initial TIM measurements on a single multifinger block [11] revealed that the experimentally observed phase shift values are lower than the theoretically predicted value for the given areal thermal energy density [12]. Furthermore, the overall sample reflectivity was at the limit of the acceptable TIM signal to noise ratio. Such effects can be related to light scattering at the gate edges together with multiple interference effects and can depend on light polarization (see the Appendix). In the previous work, where devices had well-separated fingers, edge scattering effects led to isolated optical artifacts which were disregarded [23]. In the present structures with dense finger spacing, edge scattering effects have to be considered as an integral part of the measured TIM response. As an example, for the actual finger pitch of 0.56 μm and the beam spot diameter of 1.5 μm, the optical signal comes from around three fingers. In order to investigate the optical response in more detail, we have modified our TIM setup [14] by introducing a half-wave plate (HWP) [24] to allow the rotation of the polarization of the incident probing beam against the axis of the multifinger device [see Fig. 6(a) and (b)]. Notice that, the rotation of HWP by angle θ against its fast optical axis rotates the polarization by 2θ. The sample orientation angle α

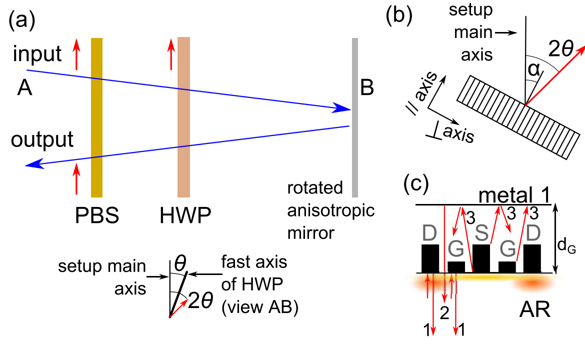

Fig. 6. (a) Modified TIM setup with possibility of polarization rotation of incident beam: HWP, polarizing beam splitter (PBS), sample (B). The beam polarization is indicated by the red arrow. Elements like microscope objective are disregarded since they do not affect the polarization. (b) Orientation of the sample (angle α) and beam polarization (angle 2θ) against the main setup axis. (c) Schematic explanation of interference between different parts of the same probing beams; 1-light reflected on drain–source and gate metallizations; 2-light reflected on metal 1; 3-scattered light; AR-active region.

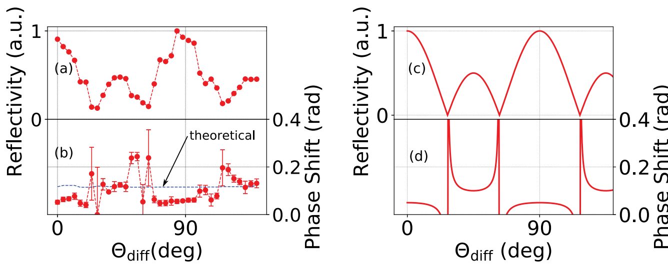

Fig. 7. (a) Measured reflectivity and (b) phase as a function of $\theta_{diff} = \theta - \alpha$, where θ is the HWP angle and α is the sample angle [see Fig. 6(b)]. The phase shift was measured in the center of a single multifinger block ($W = 1.95$ mm) at time 200 ns for $V_{GS} = 0$ V, $I = 0.2$ A. (c) Simulated reflectivity and (d) phase shift; parameters used in simulation: $r_{//} = 1$, $r_{\perp} = 0.5$, $\Delta\varphi_{//} = 0.1$ rad, $\Delta\varphi_{\perp} = 0.05$ rad taken from experiments, and $\Delta\varphi_{//,0} = 0$, $\Delta\varphi_{\perp,0} = \pi$ from fitting.

is defined by the position of the device pads and probes and cannot be varied during TIM measurements [see Fig. 6(b)].

The measured reflectivity signal as a function of the difference angle $\theta_{diff} = \theta - \alpha$ exhibits a 90° period as expected [see Fig. 7(a)]. The interesting feature, which is not present in samples with flat reflecting surface, is the nonequal reflectivity at the peaks distant by 45° [i.e., for mutually perpendicular polarization states, see, e.g., peaks at $\theta_{diff} = 0°$ and $\theta_{diff} = 45°$ in Fig. 7(a)]. This indicates an anisotropic sample reflectivity. The effect is similar to a wire polarizer [24], where the multifinger drain–source metallization and gate are considered to be analogous to the wires. The reflectivity is high for the light polarization parallel to the wire (i.e., $\theta_{diff} = 0°$ and $\theta_{diff} = 90°$; the electrons can move along the wire, thus responding efficiently to the incident field), while the reflectivity is low for the perpendicular polarization (i.e., $\theta_{diff} = 45°$ and $\theta_{diff} = 135°$; i.e., negligible response).

In the phase response as a function of θ_{diff}, shown in Fig. 7(b), one can remark that the ideal "flat surface" phase response is related to low reflectivity from Fig. 7(a), while a lower response (factor 0.5) is related to the high reflectivity. Thus, the phase response also shows the strong anisotropy due to the same scattering/interference effects. Surprisingly, for the device operated in the normal mode, the phase response oscillates between the "flat surface" response and a low value (factor 0.18) for a 90°-rotated polarization. This difference is attributed to different distributions of heat sources in the breakdown and normal modes, which modify the scattering light path and thus the overall optical response. The optical simulation can reproduce both the reflectivity and phase features, see Fig. 7(c) and (d) and the Appendix. This confirms that in contrast to structures with a flat reflecting surface, the TIM signal in such multifinger devices is influenced by scattering effects resulting in lower phase response. Since we needed to optimize signal to noise ratio we worked at the point of maximum reflectivity, which resulted, however, to the lower phase. Considering the superposition of phase signals coming from the response in normal and breakdown modes, the total TIM phase signal in the ith block, used in simulations, can be expressed as [12]

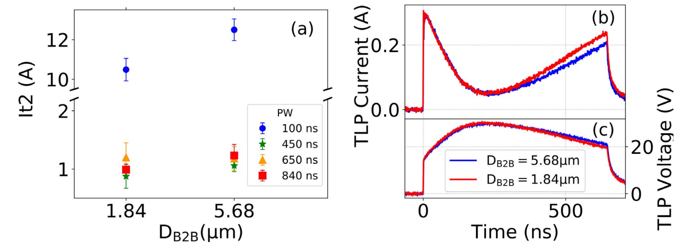

Fig. 8. (a) I_{t2} in the switch device as a function of B2B distance (D_{B2B}) for different PWs. Notice the vertical scale interruption. (b) Comparison of measured TLP current and (c) voltage waveforms for the same TLP charging voltage (V_{charge} = 30 V) for switches with different D_{B2B}; PW = 650 ns.

$$\Delta\phi_{total,i}(t) = 4\pi/(\lambda \cdot c_v) \cdot dn/dT \cdot [0.18 \cdot E_{DISS,NORM,i}(t) + 0.5 \cdot E_{BREAK,i}(t)] \quad (1)$$

with

$$E_{DISS,NORM,i}(t) = \frac{1}{A}\int_0^t V_{DS,i}(\tau) \cdot I_{NORM,i}(\tau) d\tau \quad (2)$$

$$E_{DISS,BREAK,i}(t) = \frac{1}{A}\int_0^t V_{DS,i}(\tau) \cdot I_{BREAK,i}(\tau) d\tau \quad (3)$$

where the volume specific heat $c_v = 1.63 \cdot 10^6$ JK^{-1}m^{-3} [12], $\lambda = 1.3$ μm is the laser wavelength, $dn/dT = 1.9 \times 10^{-4}$ K^{-1} the thermo-optical coefficient of silicon [12], $E_{DISS,NORM,i}$ ($E_{DISS,BREAK,i}$) is the thermal energy dissipated in the normal (breakdown) mode, $I_{NORM,i}$ the normal mode component of the current (i.e., formed channel in MOSFET) in the ith transistor, $I_{BREAK,i}$ the breakdown current component of the ith transistor, $V_{DS,i}$ the voltage on the ith transistor block and A the effective area of the block. This signal calibration is essential for correct interpretation of measured TIM response.

V. Stacked Transistor Analysis During ESD Stress

Analysis of stacked devices under TLP measurements shown in Fig. 8(a) indicates that the failure current I_{t2} for

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

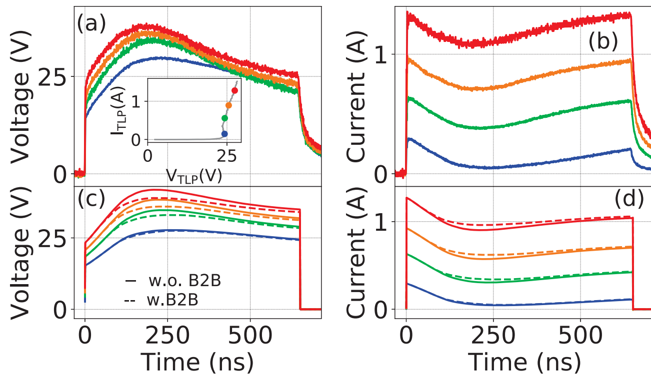

Fig. 9. (a) and (b) Measured and (c) and (d) simulated TLP voltage and current waveforms for PW = 650 ns of the switch device for different increasing charging voltages. The curves marked in green are related to the bias conditions used in TIM experiments in Figs. 10(a) and 11. The TLP–IV curve at 650 ns is shown in the inset of (a). The simulations considering the breakdown source I_{BD} only (i.e., without I_{B2B}) are indicated in (c) and (d) by solid lines. The dashed curves refer to simulations with additional B2B current sources $I_{B2B,(i,i+1)}$.

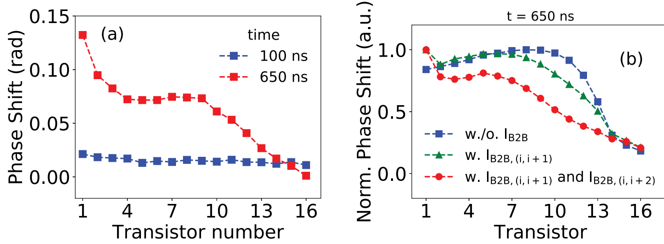

Fig. 10. (a) Measured phase shift distribution at two time instants along all blocks 1–16 in a switch with long D_{B2B} for $I = 0.5$ A and PW = 650 ns [bias condition are indicated by green lines in Fig. 9(a) and (b)]. The displayed value in the block arises from averaging over the finger width. The 15 pulses were applied at each scanning position. (b) Simulated phase shift distribution along the blocks for the same bias conditions as in (a): squares for the model with breakdown sources $I_{BD,i}$, triangles for the model with additional B2B current sources $I_{B2B(i,i+1)}$ and dots for the model with additional B2B sources $I_{B2B(i,i+2)}$.

PW = 100 ns is higher than 10 A while it is near 1 A for PW ≥ 450 ns. The devices with the short D_{B2B} exhibit lower I_{t2}. Thanks to $I_{t2} > 10$ A for PW = 100 ns, both devices can reach equivalent human body model (HBM) voltage higher than 15 kV. Charge device model (CDM) testing also showed very good performances of at least 1 kV.

Selected measured TLP voltage and current waveforms for a long D_{B2B} device for different charging voltages are depicted in Fig. 9(a) and (b). The corresponding working points in a 650-ns TLP–IV curve are indicated in the inset of Fig. 9(a). The voltage increases and the current decreases in the first 150 ns, whereas for $t > 200$ ns the voltage decreases and the current increases. Similar waveforms can be seen also for short D_{B2B} device. Fig. 8(b) and (c) compares TLP waveforms for the same TLP charging voltage for the device with short and long D_{B2B}. We notice that the short D_{B2B} switch shows steeper current increase (and related voltage decrease) for $t >$ 200 ns.

The complex shape of the voltage and current waveforms is a combined effect of a time-dependent impedance of the stack (discussed in Section VI) and pulser loadline of 50 Ω.

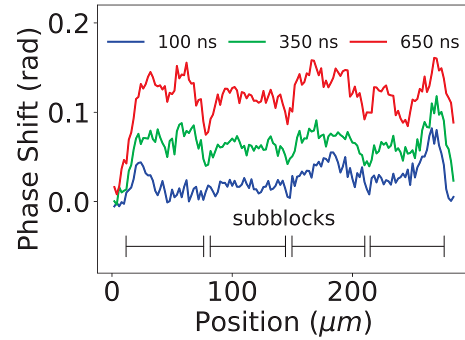

Fig. 11. Phase shift distributions at three time instants along first transistor of the switch on a device with long D_{B2B}. The bias conditions are the same as in Fig. 10(a). The scan path is marked in Fig. 1(a) as "scan A."

This indicates a complex voltage distribution over the transistors in the stack. Therefore, we have probed the device by TIM. Fig. 10(a) shows the phase distribution across 16 blocks of the stacked device for two time instants of a 650-ns long pulse. The device stress condition corresponds to green curves in Fig. 9(a) and (b). At 100 ns, nearly homogenous power dissipation is observed among the blocks [Fig. 10(a)]. However at longer time instants, the phase shift rises in block no.1 and increases also in central blocks (no. 6–9), but remains low in the bottom transistors no.14–16. Fig. 11 shows a phase distribution along the block no.1 [see path in Fig. 1(a) indicated as "scan A"], showing a homogeneous power distribution. Small dips can be recognized which are related to separation between the four subblocks. Similar homogeneous power distribution has been found also in other blocks.

VI. Circuit Simulation and Interpretations

1) Simulations With Breakdown Model: The MOSFET current in each block in normal operation, $I_{NORM,i}$, [Fig. 1(c)] has been modeled by BSIM model [25] under isothermal conditions. It includes drain to source and gate to drain overlap capacitors. External forward and reverse diodes between gate–source and gate–drain have been incorporated to model the gate leakage current that leads to better simulation of the gate voltage evolution. [They are not drawn in Fig. 1(c).]

The breakdown current of a single block transistor I_{BD} has been modeled by voltage controlled current source based on calibration measurements derived from Fig. 2(a). So I_{BD} is both function of V_{DS} and V_{GS}. The incorporation of these sources in the simulated circuit is indicated in Fig. 1(d). An efficient fitting of the results of Fig. 2(a) (see solid lines) is obtained by the following function:

$$I_{BD}(V_{GS}, V_{DS}) = \begin{cases} 0, & \text{for } V_{DS} < V_{BD} \\ 0.18 \text{ A/V}^2 \cdot (V_{DS} - V_{BD})^2, & \text{for } V_{DS} \geq V_{BD} \end{cases} \quad (4)$$

with

$$V_{BD}(V_{GS}) = \begin{cases} 2 \text{ V}, & \text{for } -1 < V_{GS} < V_{th} \\ 0.36 \cdot V_{GS} + 2.35 \text{ V}, & \text{for } V_{GS} \leq -1 \text{ V}. \end{cases} \quad (5)$$

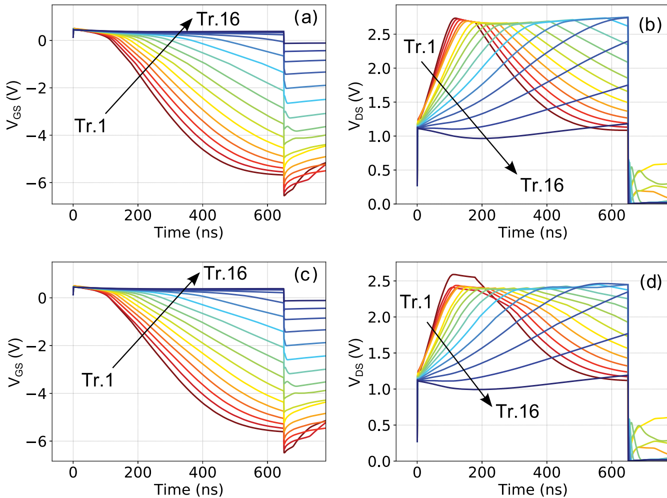

Fig. 12. Simulated transients of (a) and (c) $V_{GS,i}$ and (b) and (d) $V_{DS,i}$ at transistors 1–16 for the model with (a) and (b) the breakdown sources $I_{BD,i}$ and for the model with (c) and (d) additional B2B current sources $I_{B2B,(i,i+1)}$. The bias condition (V_{charge} = 50 V) is the same as for the green lines in Fig. 9 (i.e., TIM measurement condition).

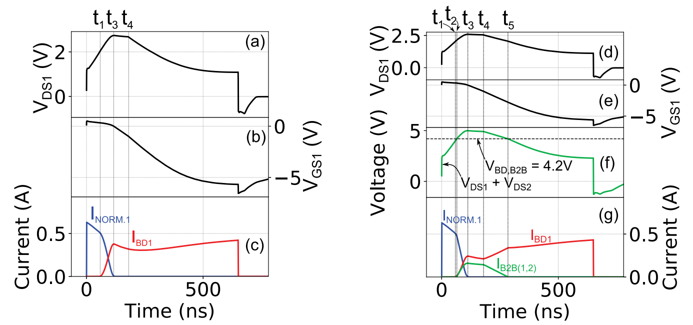

Fig. 13. Simulated current contributions [I_{NORM1}, I_{BD1}, $I_{B2B(1,2)}$] and voltage transients (V_{DS1}, V_{GS1}, V_{DS1} + V_{DS2}) in transistor 1 for the model with (a)–(c) breakdown sources $I_{BD,i}$ and for the model with (d)–(g) additional block-to-block current sources $I_{B2B(i,i+1)}$, parameter $K = 0.2$ A/V^2.

The fitting curves related to (4) and (5) are given with solid lines in Fig. 2(a) and (b), respectively. The TLP pulse has been simulated by an ideal square voltage pulse source (rising edge of 1 ns) with a 50-Ω resistor in series.

Solid lines in Fig. 9(c) and (d) shows simulated voltage and current waveforms which correspond to bias conditions in experiments of Fig. 9(a) and (b), respectively. The breakdown model already reproduces the main features of the experimental current and voltage waveforms. Fig. 12(a) and (b) shows the corresponding extracted $V_{GS,i}$ and $V_{DS,i}$ transients at each transistor block for V_{charge} = 50 V, respectively. [It corresponds to the TLP condition of the green curves in Fig. 9(a) and (b)]. During the pulse rising edge the C_{GS} and C_{GD} capacitors are charged [i.e., gate coupling effect [16], see Fig.1(c)], causing the similar initial V_{GS} at each MOSFET block as shown in Fig. 12(a). Since initial V_{GS} of ∼0.6 V is higher than the threshold voltage V_{th} = 0.3 V, each MOSFET is conducting initially in normal mode (saturation regime) [see Fig. 12(b)]. After the end of the pulse rising edge, the capacitors discharge via R_G resistors. As a consequence, the normal MOSFET current in each transistor and consequently the total current decreases which is in accordance to experiments [see curves for t < 150 ns in Fig. 9(a)]. As a consequence, the voltage $V_{DS,i}$ on each transistor increases, see waveforms for t < 150 ns in Fig. 12(b). Eventually, when the voltage $V_{DS,i}$ on transistor becomes higher than V_{BD} a parallel conductive path due to transistor avalanche breakdown (i.e., breakdown source) is activated. This is seen as an increase of the current [Fig. 9(a)] and related decrease in voltage [Fig. 9(b)] of the stack for t > 200 ns. Since the breakdown voltage V_{BD} depends on V_{GS} [see Fig. 2(b)] and V_{GS} decreases during the discharge [Fig. 12(a)], the $V_{DS,i}$ of each transistor starts to decrease [Fig. 12(b)] when the $V_{GS,i}$ becomes negative. A plateau in V_{DS}(t) in Fig. 12(b) occurs in the interval when the breakdown voltage is V_{GS}-independent in the interval −1 V < V_{GS} < V_{th} [see Fig. 2(b)].

Furthermore, since the top transistor (no.1) is on highest potential, $V_{GS,1}$ becomes negative earlier than on other transistors, as shown in Fig. 12(a). Consequently, there is a progressive delay in $V_{GS,i}$(t) decrease in Fig. 12(a) and $V_{DS,i}$(t) increase in Fig. 12(b) as going from blocks 1 to 16. The $V_{DS,i}$ in blocks no.14–16 even do not reach the V_{BD}, explaining low TIM signal in these blocks in Fig. 10(a).

The current sharing between the MOSFET normal and breakdown modes is explained in detail for transistor 1 in Fig. 13(a)–(c). For better visualization, $V_{GS,1}$(t) and $V_{DS,1}$(t) transients from Fig. 12(a) and (b) are also shown in Fig. 13(a) and (b). At t_1, the breakdown condition occurs and $I_{NORM,1}$ decreases, whereas $I_{BD,1}$ starts to increase. When V_{GS1} = 0 V at t_3, $I_{BD,1}$ takes over the whole current. When V_{GS1} = −1 V at t_4, V_{DS1} start to decrease due to the V_{BD} decrease for V_{GS} < −1 V.

The normalized simulated phase distribution for t = 650 ns are given in Fig. 10(b), see the blue lines with squares. The TIM signal is calculated using (1)–(4) taking $I_{BREAK,i}$ = $I_{BD,i}$. For t = 650 ns the simulated distribution reaches maximum at block 12, due to activity of the breakdown sources. However, the simulated distribution differs from the shape found in the experiments [Fig. 10(a)]. We suggest that this discrepancy could be due to negligence of parasitic BJT action between transistor blocks.

2) Modeling Parasitic BJT Action Between Blocks: The increase of the current for t > 200 ns for the short D_{B2B} switch in Fig. 8(b), where I_{NORM} is vanishing, indicates that an additional breakdown path or parasitic BJT action exists between the blocks. Looking at Fig. 1(b), such a discharge path can exist between the bottom side of the source of the $(i+1)$th block and the top side of drain of the ith block. The snapback/open-base breakdown behavior in Figs. 3 and 4 and N_F dependence of voltage in Fig. 3 also indicates BJT action.

For modeling of the additional discharge current between neighboring blocks, so-called block-to-block (B2B) current source $I_{B2B(i,i+1)}$ [see also schematics in Fig. 1(e)], we consider that open-base avalanche breakdown occurs between them when the voltage between the drain of the $(i+1)$th transistor and source of the ith transistor (i.e., $V_{DS,i} + V_{DS,i+1}$) overcome certain voltage $V_{BD,B2B}$. The latter value is chosen higher than $2 \cdot V_{BD}(V_{GS} = 0)$. The $I_{B2B(i,i+1)}$ current source

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

is modeled as a stepwise linear function

$$I_{\mathrm{B2B}(i,i+1)}(V_{DS,i}, V_{DS,i+1}) = \begin{cases} 0, & \text{for } V_{\mathrm{DS},i} + V_{\mathrm{DS},i+1} < V_{\mathrm{BD,B2B}} \\ K \cdot (V_{\mathrm{DS},i} + V_{\mathrm{DS},i+1} - V_{\mathrm{BD,B2B}}), & \text{for } V_{\mathrm{DS},i} + V_{\mathrm{DS},i+1} \geq V_{\mathrm{BD,B2B}} \end{cases} \quad (6)$$

where K is the slope and $i = 1, \ldots 15$. Parameters $V_{\mathrm{BD,B2B}}$ and K are considered as fitting parameters.

Dashed lines in Fig. 9(c) and (d) shows the simulated $I(t)$ and $V(t)$ waveforms taking into account the B2B current sources for the same bias conditions as shown in Fig. 9(a) and (b). The inclusion of I_{B2B} results in increased current, in agreement with results of Fig. 8(b), where higher coupling leads to higher current. Fig. 12(c) and (d) shows the transient $V_{\mathrm{GS},i}(t)$ and $V_{\mathrm{DS},i}(t)$ for the conditions of the green line in Fig. 9(a) and (b). Fig. 13(d)–(g) shows the voltage and current components related to first transistor. $I_{\mathrm{B2B}(1,2)}$ is nonzero only in case when $V_{\mathrm{DS},1} + V_{\mathrm{DS},2} > V_{\mathrm{BD,B2B}} = 4.2$ V, i.e., in the interval $t_2 < t < t_5$ [see Fig. 13(d) and (f)]. Moreover, with I_{B2B} sources, simulated phase distribution indicated by triangles at $t = 650$ ns [Fig. 10(b)] matches better the experimental distribution of Fig. 10(a). In particular a peak on the first transistor appears because of the faster decrease in $V_{\mathrm{GS},1}$ [Figs. 12(c) and 13(e)] and consequent faster rise in $V_{\mathrm{DS},1}$ compared to other blocks [Figs. 12(d) and 13(d)]. Here in the simulated energy dissipation, $I_{\mathrm{BREAK},i}$ in (3) was modeled as $I_{\mathrm{BREAK},i} = I_{\mathrm{BD},i} + \Sigma I_{\mathrm{B2B}(i,i+1)}$, where the symbol Σ indicates summation over nearest neighbors of the ith site.

The above results show that considering the B2B coupling between neighboring fingers can qualitatively explain the TIM results. We have also performed simulation including additional B2B sources between the second nearest neighbors [see Fig. 1(e)] which shows some better agreement between TIM measurements and simulations [see red circles in Fig. 10(b)]. Although the simulated voltage and current waveforms show qualitative agreement to experiments the existing differences can be attributed to model simplifications. Unlike the breakdown calibration in Section III, we did not have suitable test structures which could be used for calibration of interblock coupling. Thus, the used implementation of B2B current sources might still be an oversimplification, since $I_{\mathrm{B2B}(i,i+1)}$ might also depend on gate voltages, similar to the case of I_{BD} [see (4) and (5)]. Furthermore, the self-heating effect has not been included, but it can influence both the breakdown voltage and the triggering voltage of a parasitic BJT which can modify exact transient behavior. For more detailed analysis, 3-D mixed mode TCAD simulation would be necessary.

VII. Conclusion

ESD behavior of an RF switch, composed of stacked MOSFET blocks, has been analyzed. The observed complex current and voltage waveforms are fingerprints of dynamical changes of the bias conditions along the blocks during the TLP pulse. For times $t < 150$ ns, the current first decreases due to progressing deactivation of initially dominant MOSFET channel current. The channel current dominance in this time scale explains the high HBM and CDM robustness of this device. For $t > 200$ ns, the current increases due to the activation of open-base avalanche breakdown in the majority of the transistors. The drain-to-source voltage on single MOSFETs and reaching the condition for breakdown is strongly dependent on gate-to-source bias at each block. In particular, the decrease of the breakdown voltage with negative gate bias is crucial to understand the results. The electrical and TIM measurements have also indicated that the simulation model needs to include additional current paths due to parasitic B2B bipolar coupling. To protect the device against pulses of more than 200 ns [4], an external shunt inductance is used in switch products.

Apart from understanding device/circuit behavior, we have explained a peculiar TIM response related to anisotropic reflectivity and phase. It is inherent to the studied multi-finger device structure where scattering and multiple beam interference effects plays important role. The lower phase signal should be taken into account in quantitative optical probing [12] and failure analysis.

Appendix

The modified part of the existing heterodyne TIM setup starting with a polarizing beam splitter (PBS), HWP and sample (B) is shown in Fig. 6(a). The PBS is part of the Faraday insulator which separates the incident and reflected beams (see [26, Fig. 1]). The polarization state of the beam is treated using Jones matrix approach [24]. Each optical element is represented as a 2×2 matrix which transforms the initial light polarization state at the element input to another state at its output. The incident beam, here linearly polarized in the direction of the so-called setup main axis, is represented as the Jones vector (E_0; 0), where E_0 is the electric field amplitude. By reflecting on the sample, passing via HWP again, and passing again via PBS the returning beam is again linearly polarized with field amplitude $E_{\mathrm{ret}} = \tilde{r} E_0$, where $\tilde{r}$ is the complex number containing information on reflectivity and phase

$$\begin{pmatrix} E_{\mathrm{ret}}(\theta) \\ 0 \end{pmatrix} = S_{\mathrm{pol}} \times S_{\mathrm{HWP}}(-\theta) \times S_{\mathrm{sample}} \times S_{\mathrm{HWP}}(\theta) \times \begin{pmatrix} E_0 \\ 0 \end{pmatrix} \quad (\mathrm{A1})$$

with

$$S_{\mathrm{pol}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \; S_{\mathrm{HWP}}(\theta) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

$$S_{\mathrm{sample}} = \begin{pmatrix} r_{\parallel} \exp i(\varphi_{\parallel 0} + \Delta\varphi_{\parallel}) & 0 \\ 0 & -r_{\perp} \exp i(\varphi_{\perp,0} + \Delta\varphi_{\perp}) \end{pmatrix}$$

where S_{pol}, S_{HWP}, and S_{sample} are matrices related to PBS, HWP, and sample, respectively. The anisotropic sample reflectivity coefficients in the polarization direction parallel and perpendicular to the gate width are denoted by $r_{//}$ and $r_{\perp}$, respectively [see Fig. 6(b)]. The anisotropic phase response (i.e., the useful response to thermal signal) in the "//" and "⊥" direction of the sample are denoted by $\Delta\varphi_{//}$ and $\Delta\varphi_{\perp}$, respectively. The terms $\varphi_{//,0}$ and $\phi_{\perp,0}$ denote constant phase terms. The matrices related to sample rotation by angle α were

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

omitted in (A1) because the sample rotation just shifts the θ-axis. The total reflectivity r and phase response $\Delta\varphi$ of the sample are thus equal to

$$r(\theta) = |\tilde{r}| \tag{A2}$$

$$\Delta\phi(\theta) = \arg[\tilde{r}(\Delta\varphi_{//}, \Delta\varphi_{\perp})] - \arg[\tilde{r}(\Delta\varphi_{//}=0, \Delta\varphi_{\perp}=0)] \tag{A3}$$

where the second, subtracted term in (A3), represents a reference phase, so that $\Delta\phi = 0$ for $\Delta\varphi_{||} = \Delta\varphi_{\perp} = 0$. Fig. 7(c) and (d) represents simulated curves using (A2) and (A3), respectively. The variation of the reflectivity with θ [Fig. 7(a)] is due to the fact that the returning beam is generally elliptically polarized and only one polarization passes by PBS. The simulated phase jumps [Fig. 7(b)] at positions with low reflectivity are correlated with increased noise in experiments.

Finally, let us discuss the possible physical reasons for the phase anisotropy. In a first approximation, we consider that the used multifinger structure in its cross section can be considered as a grating in Fig. 6(c). When the probing beam approaches from the backside, one part of the beam reflects partially on the drain–source metal or gate (see returning beams "1") and the second part reflects on the top interconnect metallization (e.g., metal 1), see beams "2." The reflected beams, such as 1 and 2, interfere and provide response which depends not just on the optical path difference due to the refractive index change in the active region (AR), but also on the step height d_G [Fig. 6(c)]. The effect is similar to multiple beam interference effects as in a Fabry–Perot resonator [24], observed by TIM, e.g., in an SOI structure [27]. As a result, the measured phase response is not directly proportional to the heat energy input and higher or lower phase shifts can be observed [27]. This picture describes the main contributions, but neglects near-field effects, e.g., edge scattering (see beams "3"). The scattering can be influenced by the refractive index profile in the vicinity of scatterers, which could be the reason why the TIM response for the normal (i.e., heat dissipation in the channel) and breakdown (i.e., heat dissipation at the drain side) operational modes is different. Numerical simulations of light propagation or multiphysics modeling of the actual device structure would give a rigorous description that includes all possible effects and will be part of our follow up work.

Acknowledgment

The authors would like to thank B. Seyboldt from Keysight Technologies Germany, Böblingen, Germany, for his strong support with Advanced Design System used for circuit simulations. They would also like to thank K.-E. Möbus from Infineon Technologies AG, Neubiberg, Germany, for fruitful discussions on the BSIM simulation model; and G. Reider and J. Darmo from TU Wien, Wien, Austria for discussion of light scattering and diffraction problems.

References

[1] M. B. Shifrin, P. J. Katzin, and Y. Ayasli, "Monolithic FET structures for high-power control component applications," *IEEE Trans. Microw. Theory Techn.*, vol. 37, no. 12, pp. 2134–2141, Dec. 1989, doi: 10.1109/22.44132.

[2] P. Hindle, "The state of RF/microwave switch devices," *Microw. J.*, vol. 53, no. 11, pp. 20–36, 2010.

[3] A. Tombak, M. S. Carroll, D. C. Kerr, J.-B. Pierres, and E. Spears, "Design of high-order switches for multimode applications on a silicon-on-insulator technology," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 10, pp. 3639–3649, Oct. 2013, doi: 10.1109/TMTT.2013.2277989.

[4] *IEC 61000-4-2 System ESD Specification, 2.0*, Int. Electrotech. Commiss., Geneva, Switzerland, 2008.

[5] Y.-Y. Chen, T.-Y. Lee, E. Lawrence, and J. Woods, "ESD considerations for SOI switch design," in *Proc. IEEE Int. SOI Conf.*, Oct. 2011, pp. 1–2, doi: 10.1109/SOI.2011.6081695.

[6] X. S. Wang *et al.*, "Concurrent design analysis of high-linearity SP10T switch with 8.5 kV ESD protection," *IEEE J. Solid-State Circuits*, vol. 49, no. 9, pp. 1927–1941, Sep. 2014, doi: 10.1109/JSSC.2014.2331956.

[7] A. Joseph *et al.*, "Power handling capability of an SOI RF switch," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2013, pp. 385–388, doi: 10.1109/RFIC.2013.6569611.

[8] Y. Zhu, O. Klimashov, A. Roy, G. Blin, D. Whitefield, and D. Bartle, "High voltage SOI stacked switch with varying periphery FETs," in *Proc. Asia–Pacific Microw. Conf. (APMC)*, vol. 3. Dec. 2015, pp. 1–3, doi: 10.1109/APMC.2015.7413448.

[9] W. R. Anderson and D. B. Krakauer, "ESD protection for mixed-voltage I/O using NMOS transistors stacked in a cascode configuration," in *Proc. Electr. Overstress/Electrostatic Discharge Symp.*, Oct. 1998, pp. 54–62, doi: 10.1109/EOSESD.1998.737022.

[10] J. W. Miller, M. G. Khazhinsky, and J. C. Weldon, "Engineering the cascoded NMOS output buffer for maximum V/sub t1/," in *Proc. Electr. Overstress/Electrostatic Discharge Symp.*, Sep. 2000, pp. 308–317.

[11] M. Rigato, C. Fleury, M. Heer, M. Capriotti, W. Simbürger, and D. Pogany, "ESD characterization of multi-finger RF nMOSFET transistors by TLP and transient interferometric mapping technique," *Microelectron. Rel.*, vol. 55, nos. 9–10, pp. 1471–1475, 2015, doi: https://doi.org/10.1016/j.microrel.2015.06.019.

[12] D. Pogany *et al.*, "Quantitative internal thermal energy mapping of semiconductor devices under short current stress using backside laser interferometry," *IEEE Trans. Electron Devices*, vol. 49, no. 11, pp. 2070–2079, Nov. 2002, doi: 10.1109/TED.2002.804724.

[13] W. Simbürger, D. Johnsson, and M. Stecher, "High current TLP characterisation: An effective tool for the development of semiconductor devices and ESD protection solutions," in *Proc. ARMMS, RF Microw. Soc.*, 2012, pp. 1–9. [Online]. Available: http://www.armms.org/media/uploads/03_armms_nov12_wsimbuerger.pdf

[14] M. Litzenberger, C. Furbock, S. Bychikhin, D. Pogany, and E. Gornik, "Scanning heterodyne interferometer setup for the time-resolved thermal and free-carrier mapping in semiconductor devices," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 6, pp. 2438–2445, Dec. 2005, doi: 10.1109/TIM.2005.858121.

[15] B. El-Kareh and R. J. Bombard, *Introduction to VLSI Silicon Devices: Physics, technology and characterization*. Norwell, MA, USA: Kluwer, 1986.

[16] A. O. Adan and K. Higashi, "OFF-State leakage current mechanisms in bulkSi and SOI MOSFETs and their impact on CMOS ULSIs standby current," *IEEE Trans. Electron Devices*, vol. 48, no. 9, pp. 2050–2057, Sep. 2001, doi: 10.1109/16.944195.

[17] P. Smeys and J. P. Colinge, "Analysis of drain breakdown voltage in enhancement-mode SOI MOSFETs," *Solid-State Electron.*, vol. 36, no. 4, pp. 569–573, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/003811019390268U

[18] J. Chen *et al.*, "An accurate model of thin film SOI-MOSFET breakdown voltage," in *IEDM Tech. Dig.*, Dec. 1991, pp. 671–674, doi: 10.1109/IEDM.1991.235333.

[19] H. Yu, J.-S. Lyu, S.-W. Kang, and C. Kim, "A physical model of floating body thin film silicon-on-insulator nMOSFET with parasitic bipolar transistor," *IEEE Trans. Electron Devices*, vol. 41, no. 5, pp. 726–733, May 1994, doi: 10.1109/16.285024.

[20] M. D. Ker and J. H. Chen, "Self-substrate-triggered technique to enhance turn-on uniformity of multi-finger ESD protection devices," *IEEE J. Solid-State Circuits*, vol. 41, no. 11, pp. 2601–2609, Nov. 2006, doi: 10.1109/JSSC.2006.883331.

[21] J. Li, H. Li, R. Barnes, and E. Rosenbaum, "Comprehensive study of drain breakdown in MOSFETs," *IEEE Trans. Electron Devices*, vol. 52, no. 6, pp. 1180–1186, Jun. 2005, doi: 10.1109/TED.2005.848858.

[22] J. Liu, H. Fan, J. Li, L. Jiang, and B. Zhang, "The gate-bias influence for ESD characteristic of NMOS," in *Proc. IEEE 8th Int. Conf. ASIC*, Oct. 2009, pp. 1047–1050, doi: 10.1109/ASICON.2009.5351505.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[23] M. Heer *et al.*, "Analysis of triggering behaviour of high voltage CMOS LDMOS clamps and SCRs during ESD induced latch-up," *Microelectron. Rel.*, vol. 46, nos. 9–11, pp. 1591–1596, 2006, doi: https://doi.org/10.1016/j.microrel.2006.07.040.

[24] H. Eugene, *Optics*. Reading, MA, USA: Addison-Wesley, 2002.

[25] Y. S. Chauhan *et al.*, "BSIM compact MOSFET models for SPICE simulation," in *Proc. 20th Int. Conf. Mixed Design Integr. Circuits Syst. (MIXDES)*, Jun. 2013, pp. 23–28.

[26] M. Goldstein, G. Sölkner, and E. Gornik, "Heterodyn interferometer for the detection of electric and thermal signals in integrated circuits through the substrate," *Rev. Sci. Instrum.*, vol. 64, no. 10, pp. 3009–3013, 1993, doi: 10.1063/1.1144348.

[27] S. Bychikhin *et al.*, "Transient interferometric mapping of smart power SOI ESD protection devices under TLP and vf-TLP stress," *Microelectron. Rel.*, vol. 44, pp. 1687–1692, Sep./Nov. 2004, doi: http://www.sciencedirect.com/science/article/pii/-S0026271404002896.

Matteo Rigato received the B.S. and M.S. degrees in electronic engineering from the University of Padua, Padua, Italy, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Institute for Solid-State Electronics, Vienna University of Technology, Vienna, Austria.

His current research interests include semiconductor device characterization and modeling, with a particular focus on ESD domain.

Clément Fleury, photograph and biography not available at the time of publication.

Benedikt Schwarz, photograph and biography not available at the time of publication.

Markus Mergens, photograph and biography not available at the time of publication.

Sergey Bychikhin, photograph and biography not available at the time of publication.

Werner Simbürger, photograph and biography not available at the time of publication.

Dionyz Pogany, photograph and biography not available at the time of publication.