System-Level IEC ESD Failures in High-Voltage DeNMOS-SCR: Physical Insights and Design Guidelines

N. K. Kranthi, James Di Sarro, Rajkumar Sankaralingam, Gianluca Boselli, *Senior Member, IEEE*, and Mayank Shrivastava, *Senior Member, IEEE*

***Abstract*— A unique failure mechanism for International Electrotechnical Commission (IEC) stress through a common-mode (CM) choke is investigated. The presence of a CM choke in the stress path was found to change the current waveform shape that the electrostatic discharge (ESD) protection device experiences on-chip. Minor variations in the stress current waveform shape for specific IEC stress levels are found to cause an unexpected window failure in drain-extended nMOS silicon controlled rectifier (DeNMOS-SCR). The 3-D technology computer-aided (TCAD) simulations are used to understand the device behavior and failure under the peculiar two-pulse-shaped IEC current waveform attributed to the presence of a CM choke. DeNMOS-SCR failure sensitivity to different components of the unique pulse shape is studied in detail. A novel device architecture is proposed to increase the DeNMOS-SCR robustness against the peculiar two pulse stimuli. The proposed DeNMOS-SCR was found to eliminate the window failures against system-level IEC stress through a CM choke in communication pins in automotive ICs. The proposed concept is universal and can be extended to all high-voltage DeNMOS-SCRs. A detailed physical insight is provided for the operation of the engineered structure.**

***Index Terms*— Current filaments, drain-extended nMOS [laterally double diffused MOS (LDMOS)], electrostatic discharge (ESD), International Electrotechnical Commission (IEC), system-level ESD.**

I. Introduction

THE electrostatic discharge (ESD) protection design is particularly challenging in automotive applications because product requirements often dictate qualification for a variety of stress models in addition to human body model (HBM) and charge device model (CDM). For example, the communication pins in automotive environments, such as a controlled area network (CAN) and a local interconnection network (LIN), must be designed to protect against system-level stresses, e.g., International Electrotechnical Commission (IEC) 61000-4-2. High-voltage laterally double diffused MOS (LDMOS)-/drain-extended nMOS (DeNMOS) devices cannot be used at the high-voltage pins as their failure current per unit area is small, resulting in unacceptably large cell size and capacitance [1]. The low failure current levels in high-voltage LDMOS/DeNMOS devices are attributed to space charge modulation (SCM)-induced filament formation at the onset voltage snapback [1]–[4]. Though previous works in [5], [6] tried to improve the LDMOS/DeNMOS ESD robustness, device survival beyond snap-back cannot be guaranteed. High-holding voltage ESD solutions, such as p-n-p's and n-p-n's, also exhibit a relatively low failure current per unit area, making on-chip system-level protection a significant challenge because of the large ESD current requirements.

In contrast to high-holding voltage devices, snapback devices that operate in conductivity modulated mode can sustain very high ESD current per unit area [7], [8]. Thus, one of the most area-efficient solutions for on-chip system-level ESD protection is the high-voltage silicon controlled rectifier (SCR) device in an LDMOS/DeMOS process [9], [10]. However, there are multiple challenges in using the HV-SCR as a protection device. For example, the problem of power scalability of such devices for long-duration discharges was highlighted in [11]–[13]. In addition, HV-SCRs are found to be vulnerable for IEC system-level failures when stressed with a common-mode (CM) choke in the path [14], [15]. Altered pulse rise time with CM choke in the stress path was attributed to high-current saturation of the choke [14]. Increased pulse rise time was found to cause multifinger turn-on issues at specific IEC stress levels. Mutual ballasting technique is proposed to mitigate multifinger turn-on problems. The high current saturation of choke was also demonstrated in [16] with transmission line pulse (TLP) stressing the choke alone. Though the mutual ballasting proposed in [14] improves the device IEC performance with uniform turn-on across all fingers, however, it is found that problem still persists with very narrow and sporadic failure window. Previous work [14] has overlooked the change in the shape of the waveform with choke and tried to address the problem simply as an increased rise time effect. In this work, which is an extension of [15],

Manuscript received April 24, 2021; revised June 21, 2021; accepted July 19, 2021. Date of publication August 9, 2021; date of current version August 23, 2021. The review of this article was arranged by Editor C. Duvvury. *(Corresponding author: N. K. Kranthi.)*

N. K. Kranthi was with the Department of Electronic Systems Engineering, Indian Institute of Science, Bengaluru 560012, India. He is now with the Analog Technology Development, Texas Instruments, Bangalore 560093, India (e-mail: kranthinagothu@gmail.com).

James Di Sarro, Rajkumar Sankaralingam, and Gianluca Boselli are with the Analog Technology Development, Texas Instruments, Dallas, TX 75243 USA.

Mayank Shrivastava is with the Department of Electronic Systems Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail:mayank@iisc.ac.in).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TED.2021.3100810.

Digital Object Identifier 10.1109/TED.2021.3100810

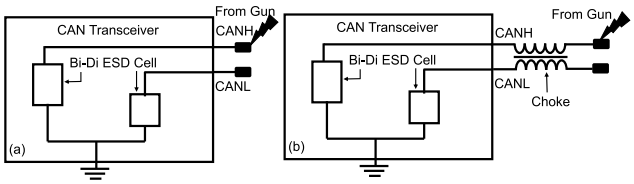

Fig. 1. Schematic representation of (a) direct air discharge test on CAN pins and (b) IEC stress through CM choke.

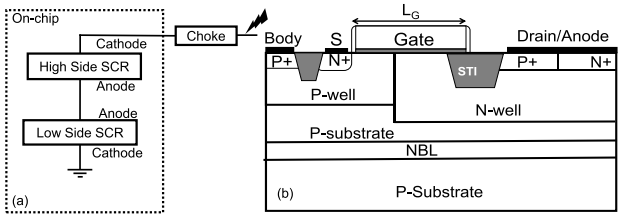

Fig. 2. (a) Representation of how a single-ended SCR is connected in a bidirectional configuration to withstand both positive and negative ESD discharges. (b) Cross-sectional view of a DeNMOS SCR.

we study the behavior of the HV-SCR when stressed by the unusual stimulus caused by choke saturation. In the extended work, a device-level solution to the IEC choke problem is proposed and experimentally validated.

The article is arranged as follows. Section II presents the problem in detail. The approach used to demonstrate the problem using 3-D technology computer-aided (TCAD) simulation is highlighted. In Section III, physics related to the root cause of the problem is discussed. Section IV discusses few key components of the unique pulse shape and its impact on the device failure. Finally, a novel modified DeNMOS-SCR design is disclosed in Section V to increases device robustness during the IEC test through a choke. The proposed design shows no window failures at any stress levels and is found to be robust against the different process and IEC test setup variations. The proposed design does not require any additional process changes.

II. Choke Saturation Problems Under High Current ESD Stress

The CAN pins in the automotive ICs presented in this work are protected using a high-voltage bidirectional DeNMOS-based SCR device at both CAN high and low pins, as depicted in Fig. 1(a). Two MOS-SCR devices are connected in a back-to-back configuration, as depicted in Fig. 2(a). During positive/negative IEC strike, the top MOS-SCR will be in the diode/SCR mode, and the bottom MOS-SCR is in the SCR/diode mode. The CM choke (off-chip element) is added to the CAN pins to reject the CM noise at the differential communication pins, as depicted in Fig. 1(b). The CM choke is a transformer that presents an inductive load to the CAN pins. When the communication pins are tested for system-level ESD protection, they are required to qualify for two scenarios: direct stress, in which the ESD gun touches one of the CAN pins, and an indirect test, in which CAN pins are stressed through a CM choke. Fig. 1 clearly depicts both the direct and indirect test setup. However, CM chokes are not designed to handle the large currents that a system-level ESD event produces [16]. The core saturation effect at high

TABLE I
IEC Results Summarizing the Pass/Fail Information at Every Stress Level for Both With and Without Choke Cases. Sporadic Device Failures Are Seen At −3-kV Stress Level When CM Choke Is Present in the Stress Path

IEC Stress Level (KV)	Through Choke(pass/fail)	Direct(pass/fail)
-1	Pass	Pass
-2	Pass	Pass
-3	Fail	Pass
-4	Pass	Pass
-5	Pass	Pass

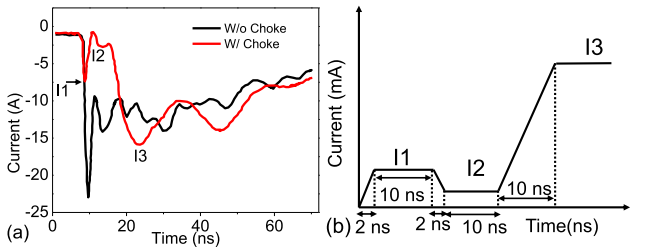

Fig. 3. (a) Measured IEC current waveforms (at the Gun tip) for a −5-kV precharge level for both direct contact stress and stress through the choke. (b) Idealized two peak current waveform for 3-D TCAD studies.

currents in CM inductive chokes has previously been found to cause unexpectedly low system-level ESD failure levels [14], [16]. Table I presents the IEC test results for the direct and indirect tests. For the direct test, the automotive IC was qualified for all stress levels. However, with a CM choke in the ESD stress path, the IC passed low and high-stress levels but failed at −3 kV. When the same experiment was repeated with much smaller step sizes (200 V) between −3 and −4 kV, failures were observed in a very fine window of stress levels. It is worth mentioning that product qualification tests use a 100-μH choke that adheres to the IEC standard for the CAN bus (IEC-62228).

The IEC current waveform measure at the gun tip for both direct and indirect IEC tests is shown in Fig. 3. ESD stress through choke exhibits a two-pulse current waveform, as depicted in Fig. 3(a). Unlike the traditional IEC 61000-4-2 contact discharge current waveform, there exists a first peak with small current (I1), followed by reduction in peak to a residual current (I2) for a certain time duration, followed by a larger amplitude second current peak (I3). It is worth highlighting here that the magnitudes of I1 and I2 vary depending on the precharge stress level and various system conditions. This variation of first peak current (I1) and residual current magnitude (I2) is the source of window failures, as they, in turn, impact the SCR turn-on behavior (in the next sections). The 3-D TCAD simulations are used to understand the device behavior under such a peculiar stimulus. The idealized current waveform used for 3-D TCAD investigations is shown in Fig. 3(b).

III. Understanding the Root Cause of Failure

An ideal current source (defined pulse shape) is used as the stress stimulus during electrothermal TCAD simulations

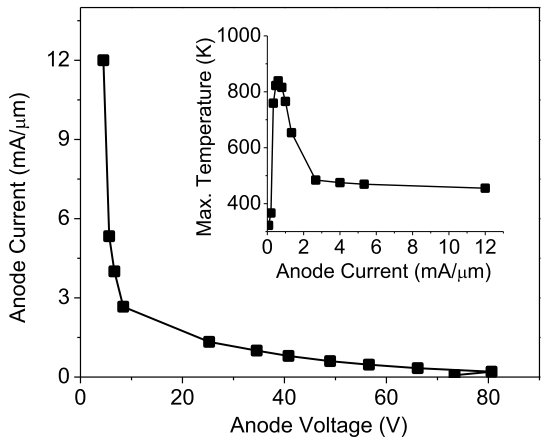

Fig. 4. 3-D TCAD simulated TLP *I–V* characteristics of DeNMOS-SCR stressed with 100-ns duration pulses. Device is observed to heat up (higher lattice temp.) at lower current values, whereas lower temperature is observed at higher current values.

with proper electrical and thermal contacts. The 3-D TCAD simulation approach is borrowed from our previous works [2], [4]. Various physical models are included to capture high field and high current effects, such as avalanche action, carrier recombination (Shockley-Read-Hall (SRH) and auger), and electric field-dependent mobility degradation models. All the IEC through choke failures are located (failure analysis (FA) data not presented) in the device, which is acting like SCR in bidirectional configuration, not in diode component. Hence, we limit our TCAD studies to understand SCR turn-on in response to peculiar two pulse shape waveform. Cathode and gate are shorted (to study without MOS-current) with a negative current pulse [stimulus shape shown in Fig. 3(b)] applied at the cathode.

A. Understanding From TLP Characteristics

Although TLP itself cannot capture the observed "window" failures under IEC through choke, it is worth analyzing the simulated TLP characteristics. Fig. 4 depicts the 100-ns TLP I–V characteristics extracted from 3-D TCAD simulations. The inset in Fig. 4 shows the extracted maximum temperature for each stress (current) value. It can be observed that the lattice temperature is highest at stimulus current levels near the SCR's holding current and then decreases significantly as stimulus current increases. This higher lattice temperature was previously observed to cause failure near snapback for long-duration pulse discharges (PW $>$ 100 ns) measured with a high-impedance load line TLP system [14]. The LDMOS/DeNMOS SCR operation at different injected current levels was explained in detail [14]. At low-current levels (prior to trigger), avalanche action at HV n-well and p-well junction generates electron–hole pairs. Generated electrons are collected at the n-well contact and holes at the p-well contact. The conduction of avalanche-generated holes is through p-well turn-on parasitic n-p-n, which floods the n-well with excess electrons. When the electron concentration exceeds background n-well doping, SCM-induced LDMOS filamentation is observed, followed by localized p-n-p turn-on and deep

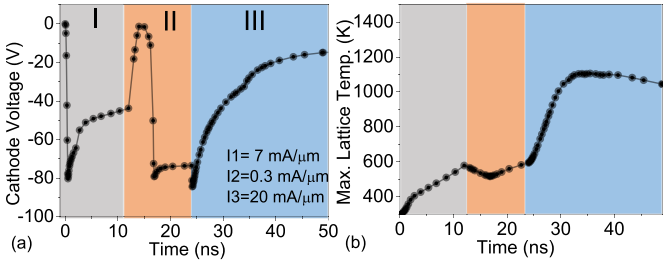

Fig. 5. 3-D TCAD extracted device response for the two pulse stimuli. (a) Transient cathode voltage. (b) Maximum lattice temperature as a function of stress time.

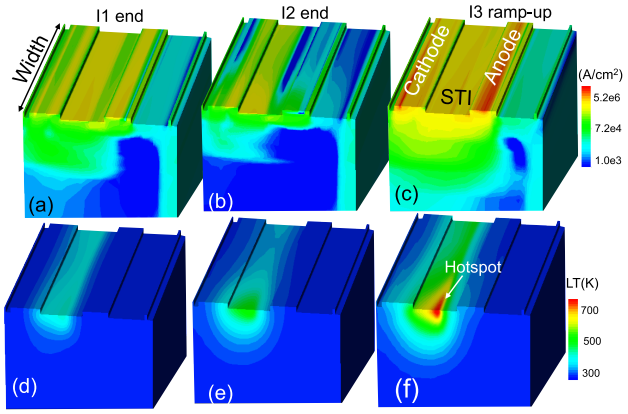

Fig. 6. (a)–(c) Conduction current density and (d)–(f) lattice temperature extracted at end of I1 (a) and (d), end of I2 (b) and (e), and during ramp up of I3 (c) and (f). The stimulus currents are same as in Fig. 5.

snap-back. When the device is designed with strong SCR action, the filament will not cause localized damage during the lower current levels, and safe snapback is achieved for longer pulsewidths. The observation that nonuniform conduction near the holding point (at a lower current level) leads to increased lattice heating is instructive for understanding the failure observed during IEC stress through a CM choke.

B. Root Cause for Failure Under Choke/Two Pulse Stimuli

The DeNMOS-SCR was simulated using 3-D TCAD and stressed with the pulse shown in Fig. 3(b) that is an idealized version of the current waveform measured for an IEC 61000-4-2 discharge through the CM choke. Fig. 5 shows device voltage response and maximum lattice temperature inside the DeNMOS SCR during the stimulus. While I1 is applied, voltage snapback occurs as the SCR turns on. It is also evident that snapback in I1 is not deep (region-I in Fig. 5). Current density in I1 indicates that the SCR cannot turn on uniformly (see Fig. 6), and lattice heating in I1 is also nonuniform (see Fig. 6). As the stimulus transitions to the I2 region, ramp down of current causes reduction (absolute number) in the cathode voltage (see Fig. 5), until the carrier concentrations in the n-well and p-well decrease significantly, and the SCR cannot maintain the low impedance state, after which a significant increase in cathode voltage is observed. This is where a low-current filament is formed (see Fig. 6). As the current begins to increase during the rising edge of I3,

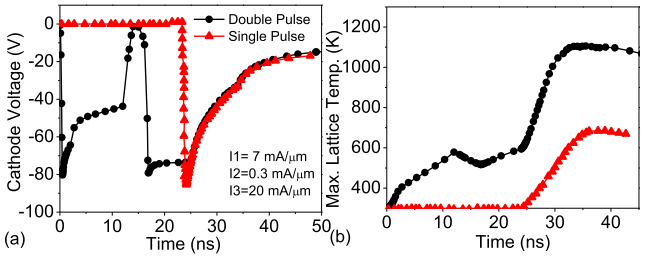

Fig. 7. 3-D TCAD extracted device response for stimulus with and without first pulse. (a) Transient cathode voltage and (b) maximum lattice temperature as a function of stress time. When there is no first pulse, device survives, as seen from direct contact IEC measurements.

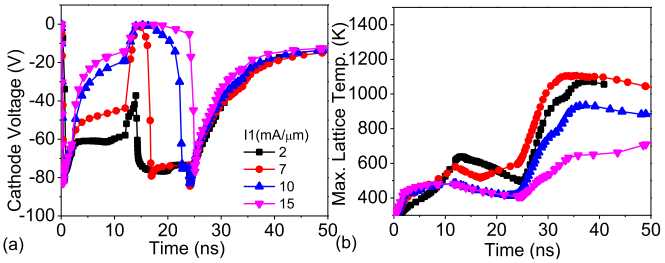

Fig. 8. 3-D TCAD extracted (a) cathode voltage and (b) maximum lattice temperature inside the device plotted as function of time for different current values in the first pulse (I1). In these studies, I2 (0.3 mA/μm) and I3 (20 μm) are kept constant. The peak temperature in during I3 found to have a maximum value in a window of currents in I1, beyond which it decreases drastically.

this filament intensifies, and an increase in current density inside the filament causes a localized hotspot to form (see Fig. 6), resulting in a peak in the lattice temperature. It is this peak during I3 that causes device failure under two pulse stimuli when a choke is present in the IEC discharge path. It is observed that the device will not see such a high lattice temperature peak during ramp-up of I3 (for the same current value in I3) when I1 and I2 current pulses are not present, as depicted in Fig. 7. Hence, this failure is particular to the shape of the stimulus that the SCR experiences under choke saturation.

IV. Failure Sensitivity for Different Stimulus Components

In this section, we investigate the effect of different current components (I1, I2) of the unique pulse shape and their influence on the device failure. For all these studies, I3 is kept constant at 20 mA/μm.

A. Sensitivity of Device Failure for First Pulse Magnitude (I1)

The sensitivity of device failure (peak temperature during ramp up of I3) is studied against the magnitude of current that is injected into SCR during I1. Several observations can be made from Fig. 8, which shows cathode voltage and maximum lattice temperature plotted for different current values in I1 (I2 and I3 are kept unchanged for these studies).

1) The increase in current in I1 causes deeper voltage snapback during the I1 pulse.

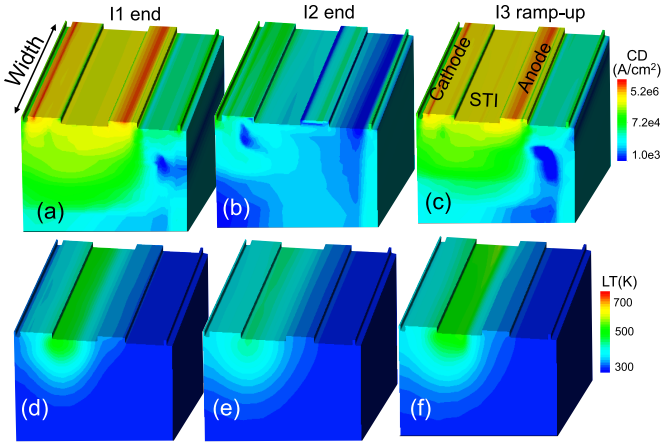

Fig. 9. (a)–(c) Conduction current density and (d)–(f) lattice temperature extracted at (a) and (d) end of I1 and (b) and (e) end of I2, and (c) and (d) during the ramp up of I3. Injecting larger current through SCR in I1 (15 μm in this case) causes more uniform SCR turn on in I1, which eventually causes more uniform turn-off in I2 and eliminates the filament-induced SCR failure in larger IEC current pulse I3.

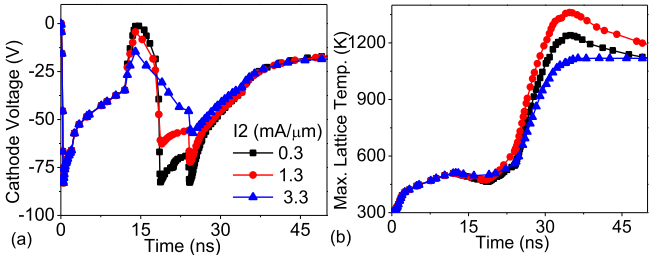

Fig. 10. (a) Cathode voltage and (b) maximum lattice temperature plotted as a function stress time for different current levels in I2 region of the stimulus. The I1 for this experiment is chosen to be 7 mA/μm. The peak temperature first increases with I2 and then decreases with further increases once I2 exceeds 1.3 mA/μm.

2) As I1 increases, the SCR takes a longer time to turn off, accompanied by a "snap up" in its cathode voltage.
3) Maximum lattice temperature during I3 first increases and then decreases to lower values with an increase in injected current in I1.

The deeper snapback at higher I1 values is attributed to stronger and more uniform SCR turn-on, as shown in Fig. 9. Larger current density in I1 implies that carriers take a longer time to recombine when the current at the cathode is ramped down to I2. This was also found to cause a more uniform current distribution during I2 (see Fig. 9), which mitigates filament formation during the ramp-up of I3. Even if there is some nonuniform conduction during I3, the distribution of carriers across the device in I2 causes faster filament spreading and lower lattice temperature.

B. Sensitivity of Device Failure for Second Pulse Magnitude (I2)

The risk of choke-induced failures is also sensitive to the residual current magnitude (I2) that is in between the two current peaks (I1 and I3). When I1 is assumed to be a lower value (around 7 mA/μm in this case), the sensitivity to I2 is amplified, as shown in Fig. 10. Furthermore, there is a window

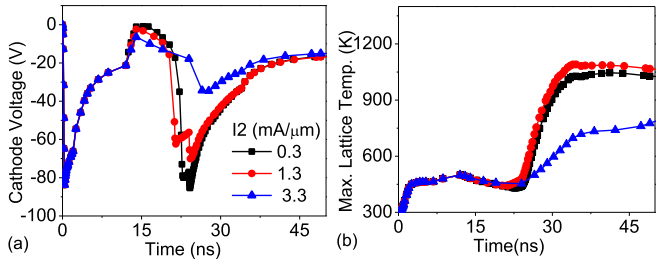

Fig. 11. (a) Cathode voltage and (b) maximum lattice temperature plotted as a function stress time for different current levels in I2. The I1 for this experiment is 10 mA/mum. The larger current in I1 reduces the temperature swing or sensitivity for changes in I2.

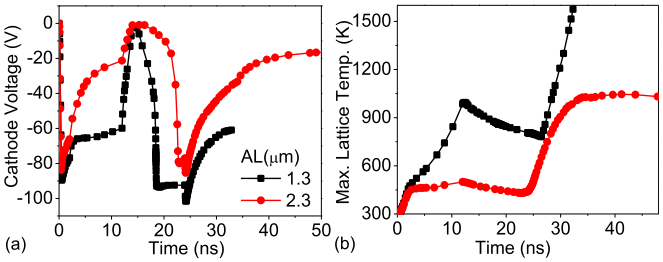

Fig. 12. (a) Cathode voltage and (b) maximum lattice temperature as a function stress time for different SCR designs when they are stressed with the same stimulus (two pulse stimuli). The deeper snapback in I1 and small temperature peak in I3 are observed for device with strong SCR (larger P+ anode length AL) action compared to the weaker SCR device (smaller AL).

observed in I2 for which the peak temperature in I2 is a maximum. At lower I2 values, current density inside filament is reduced, and the peak does not occur during I2, resulting in uniform turn-on in the I3 region. Furthermore, the swing in peak temperature as a function of I2 depends on the magnitude of I1. At higher I1, the device conducts more uniformly and turns off more uniformly in I2. In this case, the magnitude of I2 does not strongly affect the peak value, as shown in Fig. 11, as the filament is not strong in I2. Also, a dramatic reduction in peak temperature is observed at higher values of I2 when I1 is larger. This is further attributed to a strongly turned on SCR in I1; in this case, the SCR will not substantially turn off in I2 and will continue to conduct in SCR mode before it experiences the I3 pulse.

C. Two Pulse Stimuli Behavior for Different SCR Strengths

The behavior of two MOS-SCR designs with different SCR strengths was studied to understand the two-pulse device behavior. Fig. 12 presents a comparison of simulated voltage and peak temperature for two SCRs during the two-pulse TCAD simulation. SCR strength in the designs is tweaked by changing the P+ anode length (AL) (emitter area for the inherent p-n-p). The stronger SCR device has a deeper snapback in the I1 region when stressed with the same stimulus current. The deeper snapback is attributed to strong SCR action and more uniform current conduction in I1 (smaller temperature is also observed in I1 region). The stronger SCR design also takes a long time to transition into the filament mode during I2

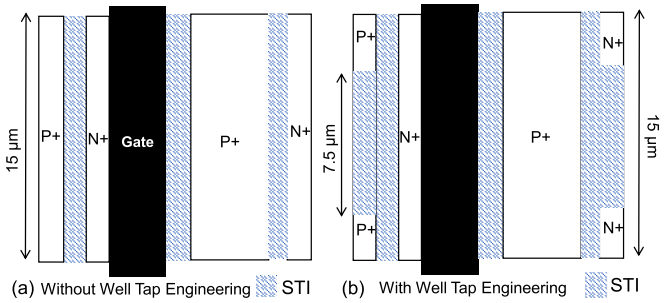

Fig. 13. Layout view of (a) reference DeNMOS-SCR and (b) well tap engineered DeNMOS-SCR. Two well taps are present at both the ends of the fingers for both n- and p-wells.

(where cathode voltage reaches to higher value). This implies that the stronger SCR design conducts more uniformly in the I3 region and reduces the peak temperature during I3. Hence, efforts to mitigate the SCR IEC-through-choke sensitivity should be focused on increasing the strength of SCR action based on the results presented in Fig. 12.

V. Device Engineering: Selective Well-Tap Placement

The device engineering is focused on filament spreading at low-current levels in I1. From Sections II and III, it is evident that the deep snapback during I1 for lower current injected values is critical. This causes more uniform conduction in I1 and hence, uniform turn-off during the residual current I2 before large second peak (I3) hits device under test (DUT). Various design parameters are investigated to solve filament damage in the DeNMOS-SCR, at different injection current values in the first pulse (data are not shown). The parameters include the P+ AL, the N+ cathode length, the silicide blocking on the P+ anode, and the N+ cathodes. However, most designs did not show significant gains to mitigate the failure under two pulse stimuli. However, engineering the well-taps was found to be crucial for DeNMOS SCR filament spreading. Different well-tap engineered structures, the physics of well tap engineering, and each structure's effectiveness are compared in this work.

A. Well-Taps at the End of the Finger

Fig. 13 depicts the top layout view of DeNMOS SCR without and with well-tap engineering. The N+ and P+ islands closest to the gate are cathode and anode, respectively. Anodes and cathodes run across the full finger width. In reference DeNMOS-SCR [see Fig. 13(a)], p-well pickup (P-tap) is isolated from N+ cathode using a small shallow trench isolation (STI). Similarly, in n-well, N+ Pickup(N-tap) is isolated from P+ anode via STI. However, in a well-tap engineered structure [see Fig. 13(b)], both N+ and P+ tap islands are not run across full finger width. STI is used to block N+ and P+ taps in the middle of each finger. The engineered design has two well-taps in each well, at both ends of the finger. Furthermore, the well-tap coverage ratio is defined as the effective tap width/finger width. The layout shown has a 50% coverage ratio. It is worth mentioning

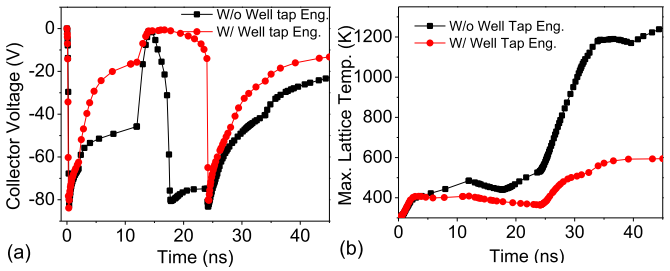

Fig. 14. (a) Cathode voltage response and (b) maximum lattice temperature inside the 3-D TCAD simulated device for device without and with well tap engineered. Taps are placed only at the edge of fingers with 50% coverage ratio. The same stimulus (I1 = 7 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm) is used for comparison.

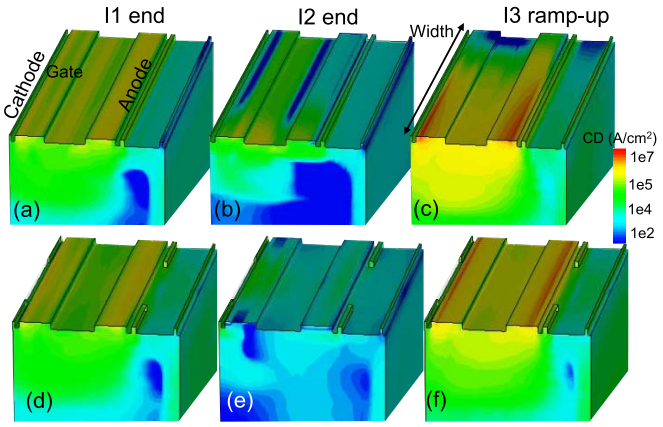

Fig. 15. (a)–(c) 3-D TCAD simulated conduction current density (A/cm^2) extracted for reference structure and (d)–(f) well-tap engineered structure. The contours are extracted at (a) and (d) end of first pulse I1 and (b) and (e) end of residual current I2, and (c) and (f) second pulse peak I3 at the same stimulus (I1 = 7 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm).

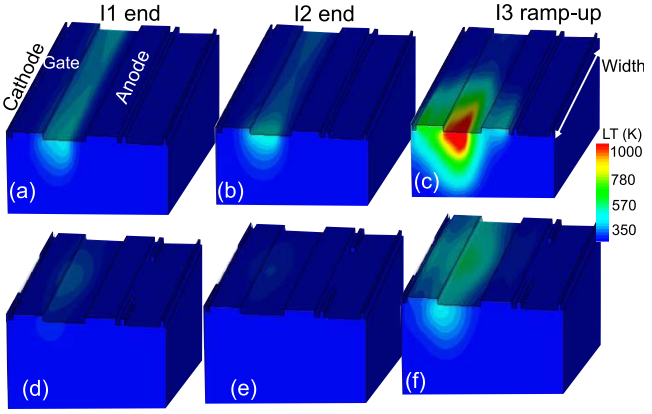

Fig. 16. (a)–(c) 3-D TCAD simulated lattice temperature (K) extracted for reference structure and (d)–(f) well-tap engineered structure. The contours are extracted at (a) and (d) end of first pulse I1 and (b) and (e) end of residual current I2, and (c) and (f) second pulse peak I3 at the same stimulus (I1 = 7 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm).

that the proposed well-tap engineering does not require any additional mask or process changes, hence cost-effective.

Fig. 14 depicts the simulated cathode voltage response and maximum lattice temperature plotted for reference structure (without tap engineering) and well-tap engineered construction with taps at ends. Comparison is made at the same current levels in I1, I2, and I3 of two pulse stimuli. The following observations can be made.

1) DeNMOS-SCR shows a deeper voltage snapback during I1 in engineered structure at the same current.
2) The time to snap-up in engineered structure is longer than reference structure (during I2) and, hence, stays in filament mode for a shorter time before the larger second pulse (I3) hits the DUT.
3) Lattice temperature significantly decreases during larger second pulse I3 in engineered structure.

During I1 in well-tap engineered DeNMOS-SCR, deeper snapback is attributed to strong SCR action because of increased well resistances (both for n-well and p-well). By blocking well-taps along finger-width, avalanche-generated carriers (electrons in n-well and holes in p-well) have to travel longer to get collected at the well-taps finally. This increase in well resistance requires smaller well currents to sustain SCR action. Hence, at the same current in I1, the well-tap structure shows deeper snapback. Stronger SCR action in the first pulse also implies that SCR conduction becomes much more uniform across the finger width. This, in turn, causes uniform turn-off during residual current I2. Figs. 15 and 16 compare the current density and lattice temperature, respectively, for reference structure and well-tapped structure. Contours are extracted at the end of I1 and I2, and I3 for the same currents. Stronger SCR action causes lower heating in I1, which does not cause strong filament during I2. This reduced filament strength in I2 helps in conducting the larger current pulse I2 uniformly. It is worth mentioning that filament formation is inevitable even in the well-tapped structure during I3. However, well tap (WT) engineering helps to spread the filament much faster, and hence, the current does not get localized. Current spreading/filament spreading is a function of electron and hole density inside the n-well and p-well. By reducing the effective area of well-taps (which are meant to collect electrons in n-well and holes in p-well), an effective electron and hole population increases. Higher base resistance and increased electron–hole populations inside wells make the filament spread effectively along device width. This reduces the strength of the hotspot during I3. The term "strong hotspot" refers to highly localized heating, which translates to high lattice temperature. This way, a well-tap engineered structure with taps has shown improved current distribution during the peculiar two-pulse stimuli and can mitigate the choke-induced window failures.

B. Split Well-Taps

Furthermore, in well-tap engineering, different well-tap placement configurations are investigated. A spilt configuration of wider well-taps into small pieces across full device width is investigated by keeping the same coverage ratio (effective tap width/total device width). Fig. 17 shows the layout view of spilt tap structure under test, along with the reference structure. Instead of two long taps at the end (3.75 μm wide each), we have chosen different splits

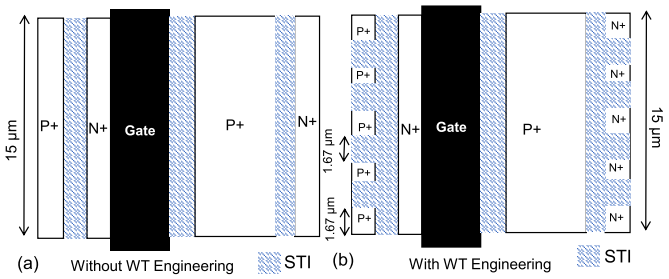

Fig. 17. Top layout view of (a) reference DeNMOS-SCR and (b) well-tap engineered DeNMOS-SCR with spilt taps.

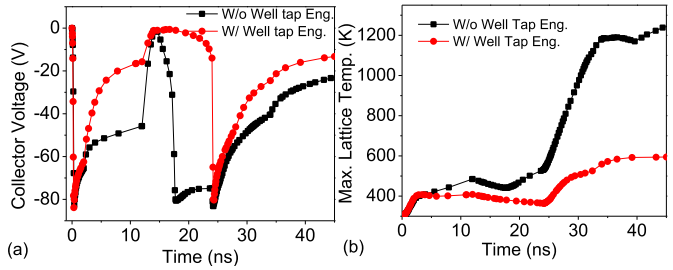

Fig. 18. (a) Cathode voltage response and (b) maximum lattice temperature inside the 3-D TCAD simulated device without and with well-tap engineering (split tap eng.). The same stimulus (I1 = 7 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm) is used for comparison.

(each with 1.67 μm) and studied response under the two-pulse stimulus. The tap size is selected to ensure that a 50% coverage ratio exists. The structure also will have two taps at both edges.

Fig. 18 depicts cathode voltage response and maximum lattice temperature for split-tap structure and reference structure plotted for same stimulus values (I1, I2, and I3). Similar to the well-tap design in Fig. 14, a deeper snapback is present in the split well-tap engineered design during the first pulse (I1). Deeper voltage snapback is attributed to stronger SCR action present in the well-tap structure, increasing n-p-n and p-n-p's effective base resistances. This strong SCR action results in a uniform turn-off of the device during residual current I2 between the IEC pulses (I1 and I3). This is also evident in the voltage response in Fig. 18(a), where well-tap structure takes a longer time to snap-up or to completely turn off compared to the reference structure. The current density distribution, as depicted in Fig. 19(b) and (e), extracted at the end of I2 shows a uniform turn-off of split well-tap engineered structure compared to reference design. The reference structure shows a filament-like behavior and stronger hotspot in Fig. 20(b) and (e). The stronger hotspot intensifies during the larger second pulse (I3) in the reference device. However, the lack of strong filament during I2 causes a more uniform turn-on of SCR in I3 in the split well-tap structure. Strong hotspot formation presented in reference structure is not seen in well-tap engineered structure, during second peak current (I3), as depicted in Fig. 20(c) and (f).

C. Well-Taps at Ends Versus Split Well-Tap Structure

The difference between both the well-tap placements discussed above is explored. Device response for the same

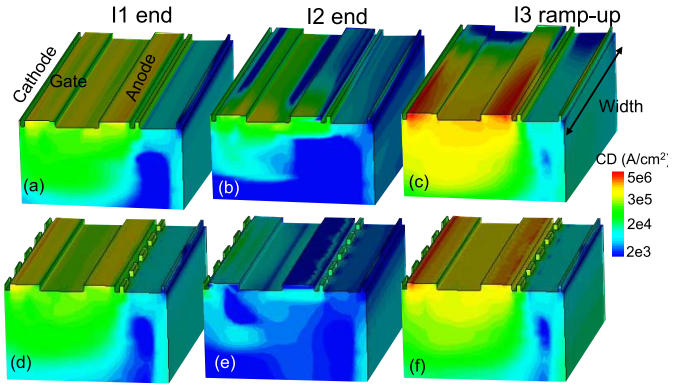

Fig. 19. (a)–(c) 3-D TCAD simulated conduction current density (A/cm^2) extracted for reference structure and (d)–(f) well-tap engineered (split taps) structure. The contours are extracted at (a) and (d) end of first pulse I1, (b) and (e) end of residual current I2, and (c) and (f) second pulse peak I3. Used stimulus values are I1 = 7 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm.

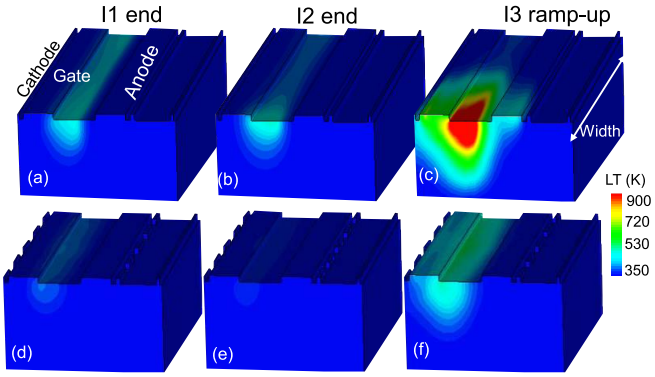

Fig. 20. (a)–(c) 3-D TCAD simulated lattice temperature (K) extracted for reference structure and (d)–(f) well-tap engineered (split taps) structure. The contours are extracted at (a) and (d) end of first pulse I1 and (b) and (e) end of residual current I2, and (c) and (f) second pulse peak I3. Both structures are simulated under same stimulus of I1 = 7 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm.

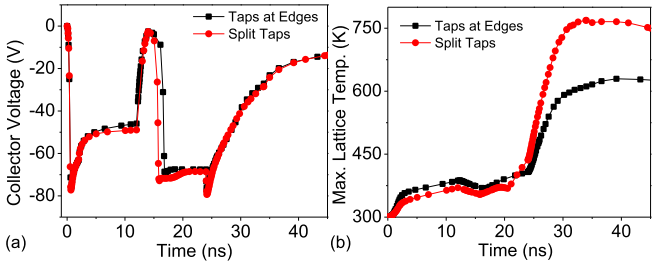

Fig. 21. Comparison of (a) cathode voltage response and (b) maximum lattice temperature for two different well-tap engineering designs at the same stimulus (I1 = 2 mA/μm, I2 = 0.3 mA/μm, and I1 = 20 mA/μm) and the same coverage ratio (50%).

stimulus is plotted in Fig. 21. From Figs. 14 and 18, difference between both well-tap configurations looks negligible. However, it is worth mentioning that the current value I1 in the stimulus is 7 mA/μm. As both device structures effectively spread current filament at this current in I1 (on the higher side), the difference is almost negligible.

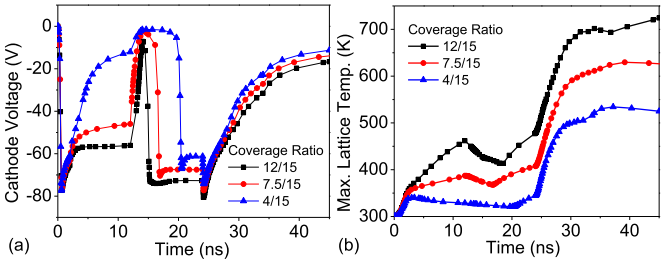

Fig. 22. Comparison of (a) cathode voltage response and (b) maximum lattice temperature extracted for with different coverage ratios (well-taps at the finger edges) under the same stimulus (I1 = 2 mA/μm, I2 = 0.3 mA/μm, and I3 = 20 mA/μm).

However, when we investigated a range of injected values in I1, the difference can be found at lower current levels in I1. In Fig. 21, 2 mA/um is used in I1, and other values of stimulus are kept unchanged (I2 and I3). Though the cathode voltage response difference is minimal, the lattice temperature difference during the second pulse is significant (around 150 K). This difference is attributed to the location of filament formation in both structures. In design, where taps are present only at the end of fingers, the favorable location for filament formation is in the middle of the finger (maximum base resistance can be experienced at this location). When filament location is in the middle, more silicon volume is available for heat conduction; well-tap engineering further helps spread the current. However, in a split well-tap structure, filament location is still found to be at one of the edges. The device is more symmetric, and the difference in base resistance is uniform across device width. This makes the hotspot stronger, and spreading becomes less effective than a structure where two well-taps are present only at the end of the finger.

D. Tap Coverage Ratio

The tap coverage ratio is found to be one of the critical design parameters in well-tap engineered structures. Fig. 22 depicts the cathode voltage and the maximum lattice temperature response plotted for different well-tap coverage ratios, stressed with the same stimulus (the same I1, I2, and I3). The well-tap configuration used for these investigations is depicted in Fig. 13. The following observations can be made.

1) Reducing well-tap coverage gives deeper snapback during the first pulse and lower lattice temperature during I1.
2) Device with the lowest coverage ratio takes a longer time to snap-up during residual current conduction in I2.
3) The lower well-tap coverage device shows reduced lattice temperature during the second pulse and, hence, is projected to show more robustness for the two-pulse stimulus.

Reducing coverage ratio means that the efficiency of carrier collection from n-well and p-well is reduced. More electrons and holes population in n-well and p-well aids the filament spreading; a weaker hotspot is present in the design with least coverage ratio. Smaller well-tap coverage also means that n-p-n and p-n-p bases become more resistive, and hence, SCR conduction also becomes stronger. Deep snapback has been seen during I1 and I3. In summary, well-tap engineering, with a proper coverage ratio, will help to solve current nonuniformity issues during the IEC choke stress conditions. The experimental evaluation of well-tap structures is presented in Table II.

TABLE II
IEC RESULTS SUMMARIZED FOR SELECTIVE WELL-TAP ENGINEERED DeNMOS SCR UNDER DIFFERENT STRESS TYPES. MULTIPLE DEVICE AND PROCESS SPLITS ARE TESTED, ALL OF WHICH FOUND TO PASS THE TEST. MEASUREMENT SYSTEM'S VARIATION ARE ALSO CAPTURED BY TESTING ON DIFFERENT IEC MEASUREMENTS SYSTEMS

IEC Stress Level (KV)	Through Choke(pass/fail)	Direct(pass/fail)
-1	Pass	Pass
-2	Pass	Pass
-3	Pass	Pass
-4	Pass	Pass
-5	Pass	Pass

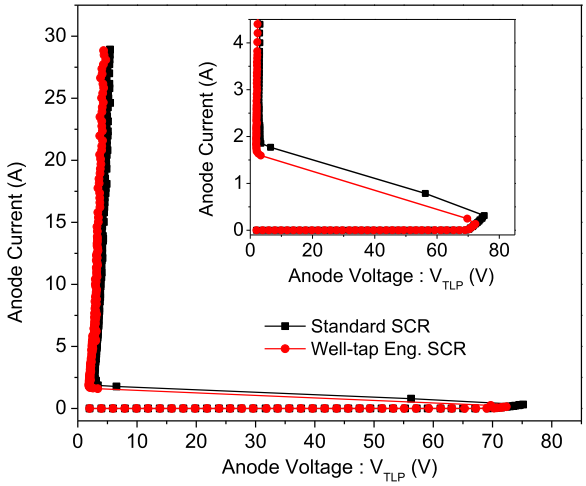

Fig. 23. Measured TLP *I–V* characteristics of standard and well-tap engineered DeNMOS-SCR. Inset depicts the trigger and holding states more clearly.

Finally, in Fig. 23, the measured TLP I–V characteristics of standard DeNMOS-SCR with 100% tap coverage and with well-tap engineered DeNMOS-SCR (50% coverage) are compared. Last data point in the I–V is the failure point. Zoomed-in view near the trigger and holding state is shown in the inset of Fig. 23. The following observations can be made from the TLP plot: 1) the trigger voltage decreases from 75 to 72 V with well-tap engineering (50% coverage); 2) the current required to trigger the SCR reduces as the n-well and p-well become more resistive in the engineered design; and 3) the uniform SCR turn-on across the finger also contributes to reduced holding current with deeper snapback. It is worth highlighting that these measurements are done with a standard 50-Ω line TLP setup; hence, the holding current depicted here is not the actual holding current of the SCR. Furthermore, the maximum operating voltage at CAN pin, in this case, is 50 V. Well-tap engineered DeNMOS-SCR with 50% coverage ratio triggers at 72 V and still has enough margin to avoid any mistriggering during functional operation.

VI. Conclusion

A choke-induced ESD failure in a system-level automotive ESD environment is revisited. Change in current pulse waveform shape that occurs within a narrow range of IEC stress levels is found to be the concern. A weak current pulse (first pulse) prior to the large IEC current peak (second pulse) was observed to cause filamentation inside the device, which eventually causes device failure during the larger second pulse. Higher first pulse current can mitigate the device failure during a large second current pulse; this is attributed to uniform turn-on of the SCR during the first pulse, which, in turn, causes a weak filament inside the device and uniform conduction during the larger IEC second pulse. The residual current between first and second pulses is also found to significantly impact the failure. Higher residual current results in a more uniform turn-off of the SCR going from the first pulse to the larger IEC current pulse, mitigating the filament formation during the larger second pulse. The proposed novel design with well-taps presents only at the edges, found to causes stronger SCR action. The reduced efficiency of taps to collect electrons and holes in the n-well and p-well causes SCR to spread the filament across its width. The increased base resistance with well-tap engineering also provides stronger and uniform SCR action, deeper snapback. The improved thermal distribution inside the well-tap engineered device for the two-pulse stimuli is observed compared with the reference design. Finally, the coverage ratio in the well-tap designs was a critical parameter as it helps to make wells more resistive and, in turn, turn-on SCR more uniformly during the first and second IEC pulses.

Acknowledgment

The authors would like to thank Yang Xiu, Texas Instruments (TI), Dallas, USA, for her help with test chip design and measurements. They would also like to thank Hans Kunz, TI, Dallas, USA for extensive IEC measurement support and the Analog ESD Team, TI, Dallas, USA for providing their feedback during the technical discussions.

References

[1] M. Shrivastava and H. Gossner, "A review on the ESD robustness of drain-extended MOS devices," *IEEE Trans. Electron Devices*, vol. 12, no. 4, pp. 615–625, Dec. 2012, doi: 10.1109/TDMR.2012.2220358.

[2] M. Shrivastava, H. Gossner, M. S. Baghini, and V. R. Rao, "Part II: On the three-dimensional filamentation and failure modeling of STI type DeNMOS device under various ESD conditions," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 2243–2250, Sep. 2010, doi: 10.1109/TED.2010.2055278.

[3] G. Boselli, V. Vassilev, and C. Duvvury, "Drain extended nMOS high current behavior and ESD protection strategy for HV applications in sub-100 nm CMOS technologies," in *Proc. Int. Rel. Phys. Symp. Proc.*, May 2007, pp. 342–347, doi: 10.1109/RELPHY.2007.369913.

[4] M. Shrivastava, C. C. Russ, H. Gossner, S. Bychikhin, D. Pogany, and E. Gornik, "ESD robust DeMOS devices in advanced CMOS technologies," in *Proc. EOS/ESD Symp. Process.*, Sep. 2011, pp. 1–10.

[5] K.-H. Kim, W.-J. Choi, and Y.-J. Seo, "Effects of background doping concentration on ESD protection properties of high voltage operation extended drain N-type MOSFET device," in *Proc. IEEE Int. Rel. Phys. Symp. Proceedings. 45th Annu.*, Apr. 2007, pp. 334–341, doi: 10.1109/RELPHY.2007.369912.

[6] J. Lee, H. Su, C. Chan, D. Yang, J. F. Chen, and K. Wu, "The influence of the layout on the ESD performance of HV-LDMOS," in *Proc. 22nd Int. Symp. Power Semiconductor Devices IC's (ISPSD)*, Jun. 2010, pp. 303–306.

[7] V. A. Vashchenko and A. Shibkov, *ESD Design for Analog Circuits*. USA: Springer, 2010.

[8] E. A. Amerasekera and C. Duvvury, *ESD in Silicon Integrated Circuits*. Chichester, U.K.: Wiley, 2002.

[9] C. Duvvury, J. Rodriguez, C. Jones, and M. Smayling, "Device integration for ESD robustness of high voltage power MOSFETs," in *IEDM Tech. Dig.*, Dec. 1994, pp. 407–410, doi: 10.1109/IEDM.1994.383381.

[10] S. Pendharkar, R. Teggatz, J. Devore, J. Carpenter, T. Efland, and C. Tsai, "SCR-LDMOS. A novel LDMOS device with ESD robustness," in *12th Int. Symp. Power Semiconductor Devices Process.*, May 2000, pp. 341–344, doi: 10.1109/ISPSD.2000.856839.

[11] G. Boselli, A. Salman, J. Brodsky, and H. Kunz, "The relevance of long-duration TLP stress on system level ESD design," in *Proc. Electr. Overstress/Electrostatic Discharge Symp. Proc.*, Oct. 2010, pp. 1–10.

[12] N. Karmel Kranthi, B. Sampath Kumar, A. Salman, G. Boselli, and M. Shrivastava, "Physical insights into the low current ESD failure of LDMOS-SCR and its implication on power scalability," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Mar. 2019, pp. 1–5.

[13] N. K. Kranthi, B. S. Kumar, A. Salman, G. Boselli, and M. Shrivastava, "Performance and reliability co-design of LDMOS-SCR for self-protected high voltage applications on-chip," in *Proc. 31st Int. Symp. Power Semiconductor Devices ICs (ISPSD)*, May 2019, pp. 407–410, doi: 10.1109/ISPSD.2019.8757641.

[14] A. A. Salman, F. Farbiz, A. Concannon, H. Edwards, and G. Boselli, "Mutual ballasting: A novel technique for improved inductive system level IEC ESD stress performance for automotive applications," in *Proc. 35th Electr. Overstress/Electrostatic Discharge Symp.*, Sep. 2013, pp. 1–7.

[15] N. K. Kranthi, J. di Sarro, R. Sankaralingam, G. Boselli, and M. Shrivastava, "Insights into the system-level IEC ESD failure in high voltage DeNMOS-SCR for automotive applications," in *Proc. 42nd Annu. EOS/ESD Symp. (EOS/ESD)*, Sep. 2020, pp. 1–7.

[16] M. Ammer, S. Miropolskiy, A. Rupp, F. Z. Nieden, M. Sauter, and L. Maurer, "Characterizing and modelling common mode inductors at high current levels for system ESD simulations," in *Proc. 41st Annu. EOS/ESD Symp. (EOS/ESD)*, Sep. 2019, pp. 1–7, doi: 10.23919/EOS/ESD.2019.8870005.