ESD CIRCUIT DESIGN AND MEASUREMENT TECHNIQUES

BY

ALEX AYLING

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Adviser:

Professor Elyse Rosenbaum

ABSTRACT

Part 1 of this thesis presents a method to measure sub-nanosecond reverse recovery in wafer-level test structures. The setup uses a transmission line pulse generator with a time domain through connection to measure the device under test current. The setup is then used to measure reverse recovery in a 65 nm CMOS ESD diode, and it is found that a quasi-static compact model does not accurately describe the observed transient. A non-quasi-static charge control model is used to accurately simulate both the reverse recovery and forward bias behavior.

Part 2 of this thesis reports the design and fabrication of an active feedback based high-voltage tolerant power clamp with optimally biased positive and negative feedback to bypass the trade-off between ESD performance and mis-trigger immunity. The circuit was fabricated in 28 nm CMOS, and characterization results show a 70% improvement in failure current over previous designs while maintaining mis-trigger immunity.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Electrostatic discharge (ESD) is a major source of failures in integrated circuits (ICs). ESD damage occurs when an IC is brought into contact with an object that is charged relative to that IC, causing current to flow between the two as the charge is redistributed. There are two dominant models for component level ESD, the human body model (HBM) and charged device model (CDM). HBM, as its name suggests, represents a charged human touching an IC and causing a discharge. The Charged Device Model (CDM), in contrast, represents the scenario where the IC is charged and then grounded, typically by a mechanical device during manufacturing or handling. The current injected into a chip during HBM testing is on the order of several amperes, and the current has a rise-time of approximately ~10 ns and a duration of ~1 μs. In CDM testing the current can be several tens of amperes, has a rise-time of 100 ps and a duration of 1 ns. The purpose of ESD protection is to safely shunt the current away from sensitive devices on-chip so it does not damage them. A significant challenge is to design the protection devices and components such that they satisfy the ESD requirements while causing as little degradation as possible to the normal operation of the chip.

The thesis is split into two parts discussing two issues encountered during the design of ESD protection devices and components. The first is the need for accurate models. ESD current is much larger than what is encountered during normal operation, and devices will behave much differently than predicted by models extracted under nominal conditions. Therefore, ESD models must be developed and extracted manually for higher accuracy. The models are only as good as the measurements used in their fitting procedure, however, and the first part of this thesis will prove that existing ESD diode models fail to predict the reverse recovery transient by introducing a new, more accurate measurement setup for sub-nanosecond reverse recovery transients.

The second part of the thesis discusses the tradeoffs encountered during the design of ESD protection components by example of an active rail clamp. The active rail clamp must have a low on-resistance to minimize the voltage drop across sensitive devices during ESD; however, it should not take too much chip area. This trade-off can be mitigated by introducing more instability into the clamp's design through the use of positive feedback, but then the reliability may be compromised. These tradeoffs can be balanced by careful circuit design techniques.

PART I

CHAPTER 2: REVERSE RECOVERY BACKGROUND

Reserve recovery is the non-zero "turn-off time" of a diode that is driven quickly from a forward to reverse bias. Consider the setup shown in Fig. 1, where an initial forward bias is applied to the device and then switched instantaneously to a reverse bias. The current through the device will not immediately drop to zero, but rather will decay to zero after some time (assuming the reverse bias voltage is less than the avalanche breakdown voltage). A representative example of the current waveform is shown in Fig. 1 [1]. The non-zero current occurs because the minority carriers injected into the quasi-neutral regions during the forward bias must be removed before the device can enter reverse bias. How long this takes depends on the speed at which the carriers are removed, and initial charge present in the quasi-neutral regions.
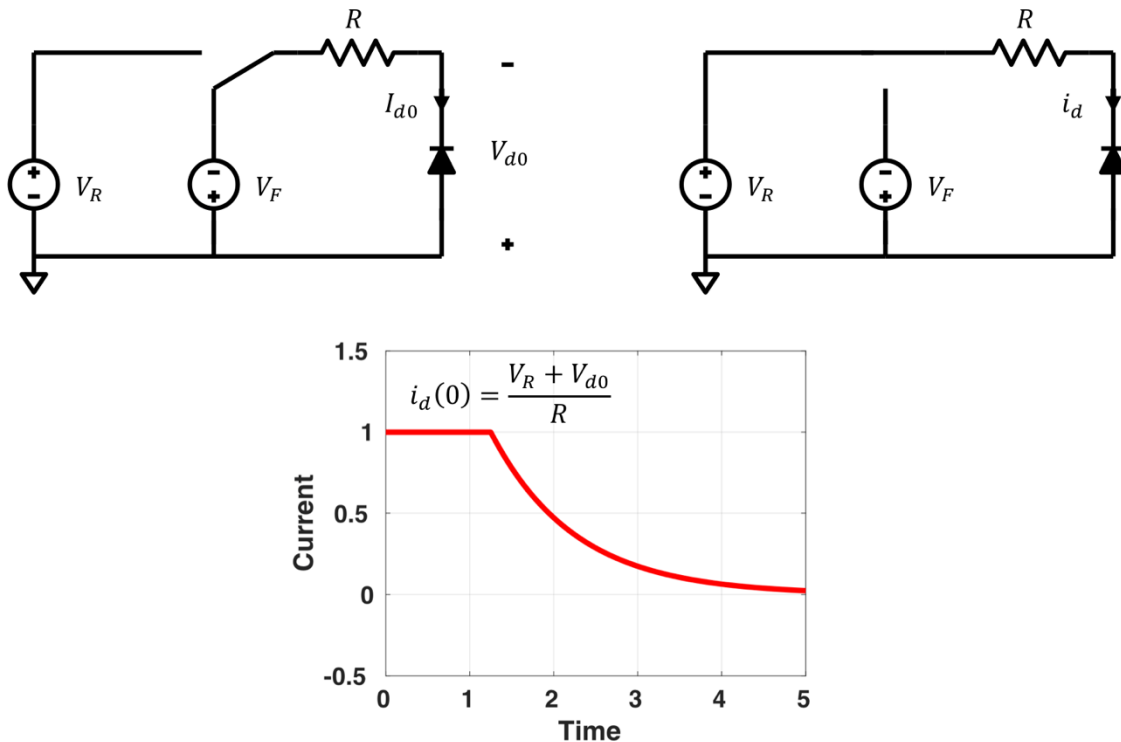
Figure 1: Simplified reverse recovery measurement setup: initial forward bias (top left), switch to reverse bias (top right) and example waveform (bottom).

The reverse recovery is extremely important in power electronics circuits such as rectifiers, switching converters etc. Reverse recovery is not typically associated, however, with ESD protection. This is likely because most ESD testing is unipolar, meaning that the injected current is either positive or negative, but not both. Reverse recovery can only occur when there are bipolar currents. There are some ESD tests that have bipolar currents such as machine model (MM) testing, and unexpected failures have been reported due to the reverse recovery during MM testing [2]-[4]. Additionally, system-level ESD testing can produce bipolar currents that may cause reverse recovery [5]-[7]. Therefore, the notion that reverse recovery is something insignificant for ESD protection is false. To understand how reverse recovery affects an ESD protection circuit, it must first be characterized with measurement. Fig. 1 does not provide any details on how a practical reverse recovery measurement is done, and chapter 3 will discuss the particular challenges for implementing a measurement setup for ESD devices.

CHAPTER 3: TIME DOMAIN THROUGH REVERSE RECOVERY MEASUREMENT

Most reverse recovery measurements are intended for large power devices, with relatively slow transients, and are based on the "clamped inductive load test" [8]. The clamped inductive load test has a minimum rise-time of approximately 5-10 ns for high power switching from forward to reverse bias [9]. ESD devices, in contrast, will have recovery times less than a nanosecond due to their small size. Reverse recovery can only be resolved if the device is switched faster than the expected recovery time [9], and therefore a sub-nanosecond reverse recovery of an ESD device is not measurable with this setup.

A transmission line-based approach, first proposed in [10] and adopted in [11]-[13], satisfies the ultra-fast switching requirement using a dedicated voltage pulse generator. In particular, [11] proposed a "100-Ω Time Domain Through (TDT)" setup using TLP as the pulse generator. TLP can supply high power pulses at rise-times as fast as 100 ps, making it a good fit for a sub-nanosecond reverse recovery measurement (see Appendix A for more information on TLP). The TDT measurement relative to time domain reflectometry (TDR) has the potential for higher fidelity measurements of the current transient; however, [11] did not present wafer-level measurements validating the setup. It will be shown that parasitic inductance introduced from wafer-level probing will compromise the ability to resolve sub-nanosecond transients.

A schematic of the proposed setup is shown in Fig. 2. An ordinary TLP setup uses a single transmission line where the device under test (DUT) is terminated as a one-port. In contrast, in the TDT setup the DUT appears as a series two-port between two transmission lines. The primary advantage of the TDT setup is that the DUT current can be measured directly at the rightmost transmission line by terminating it with a 50 Ω high-bandwidth oscilloscope port. TDR, in contrast, requires that both the incident and reflected voltage pulse be measured and

then aligned computationally, which introduces additional error and noise in the measurement [14]. These differences are expanded upon further in Appendix B.
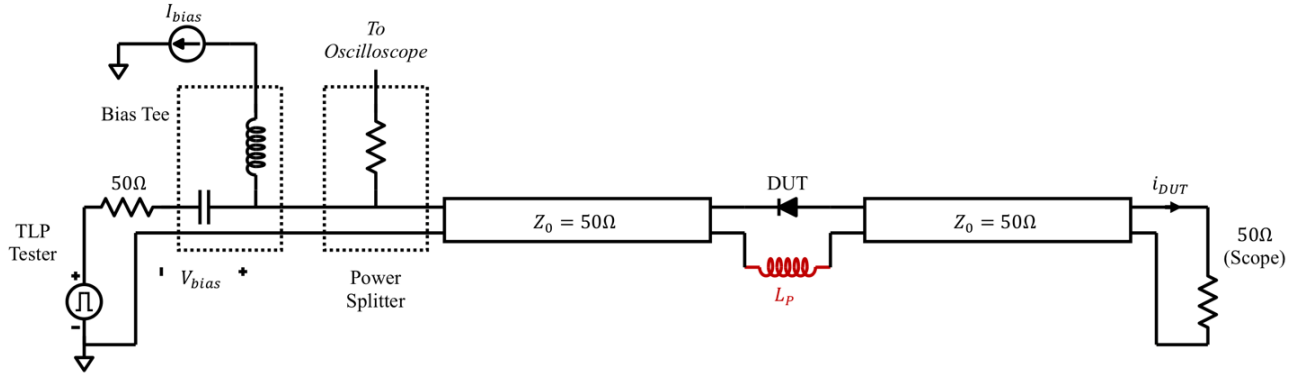

Figure 2: 100 Ω TDT measurement setup. L_p represents the inductance of the short between the transmission lines' shields.

The basic reverse recovery measurement procedure is as follows. For additional insight on obtaining high fidelity measurements, see Appendix E.

1. Provide a forward bias current with the DC source (I_{bias}) through the inductor of the bias tee.

2. Once the forward bias has been established, pulse the device into reverse bias using a TLP tester (this work uses an HPPI 3010-C TLP tester). The pulse has a 100 ps rise-time, an example of which is shown in Fig. 3.

3. Measure the device current, $i_{dut}(t)$, which flows through the rightmost transmission line into the oscilloscope port. The device current induces a voltage, $v_{meas}(t)$, across the input impedance. The current is given by

$$i_{dut}(t) = \frac{v_{meas}(t)}{50}.$$ (1)

4. The voltage can be measured at the left port of the DUT using TDR, and the right port voltage is $v_{meas}(t)$. Correctly aligning and then subtracting these in time will produce

6

the voltage across the DUT (Appendix B). If TDR is used for either the current or voltage, it must be subtracted by I_{bias} or V_{bias}, respectively (Appendix C).
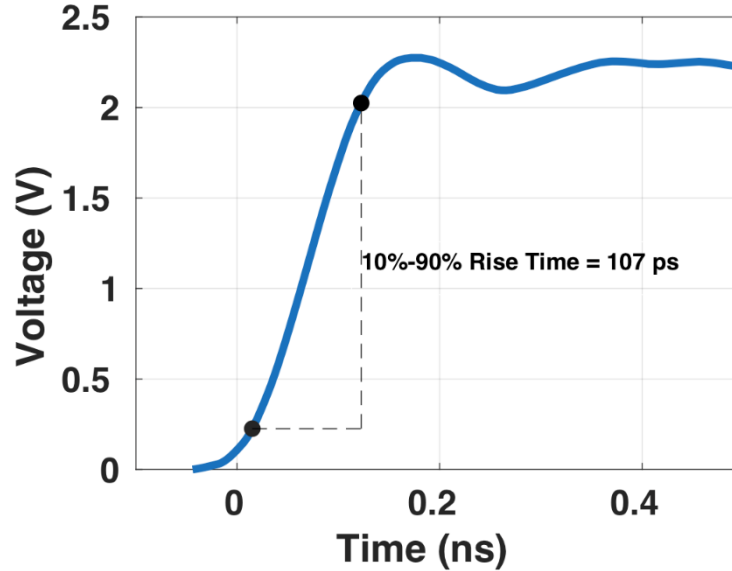

Figure 3: Example pulse from TLP tester with a rise time of 100 ps into a 50 Ω load.

The forward bias current, because it is DC, must be limited so as to not cause damage to the device. How much current that can be handled depends on the technology and size of the DUT. For the 90 nm and 65 nm devices that will be used in this work, a forward bias greater than approximately 50 mA shows significant self-heating and outright fails above 100 mA. To avoid self-heating, the forward current was restricted to no more than 40 mA.

A typical two-terminal test structure for TLP testing with Kelvin probing is shown in Fig. 4 with a diode as the example DUT. All devices in this work have that four-pad configuration. It is desirable to use the same four-pad test structure, and the wafer-level probes associated with it, for the reverse recovery measurement so that no additional test structures are required for characterizing the device. The probes, however, introduce a break in the transmission lines' shields, and for the TDT measurement to work, those two shields must be shorted together.

7

Invariably, this short has some inductance associated with it, denoted as L_p in Fig. 2, that is dependent on the probing method, three of which are compared.
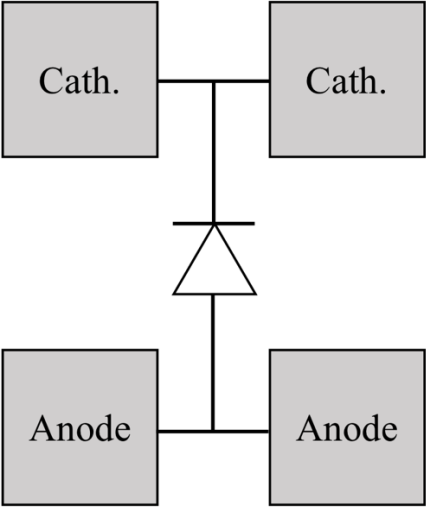
Figure 4: Four-pad TLP test structure of a two-terminal device.

1) Shield Strap

The shields of the two coaxial probes are shorted together by a strap located off the die, a diagram and picture of which are shown in Figs. 5 and 6. The strap is made of braided copper and has a diameter of 0.1 cm and a length of 1 cm.
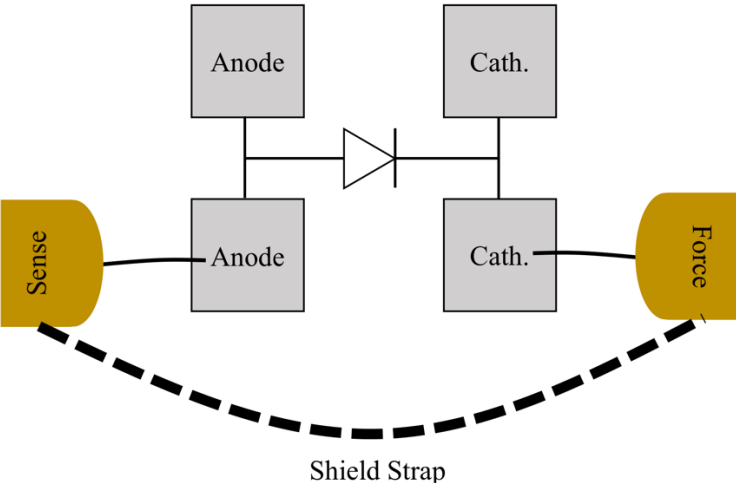
Figure 5: Shield strap probing technique.
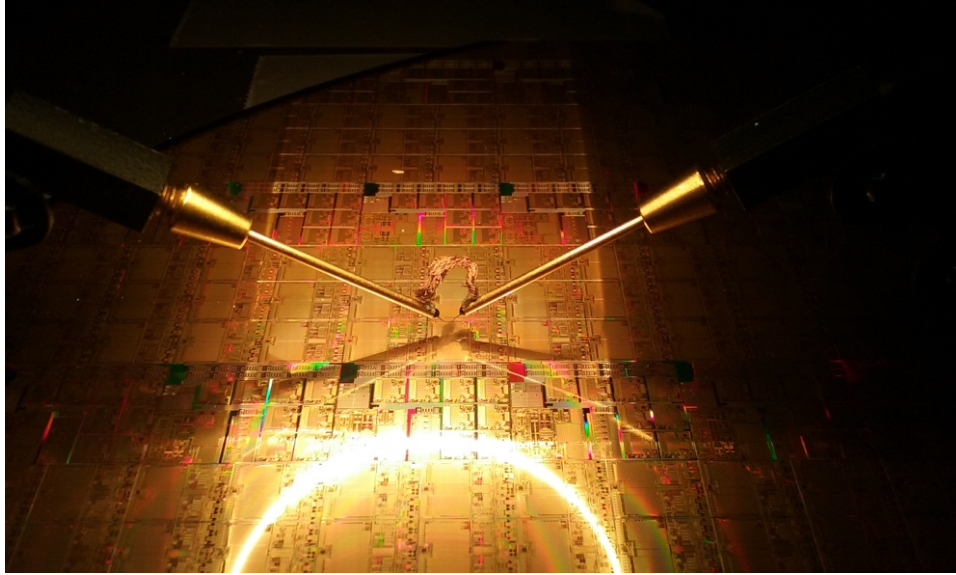
Figure 6: Shield strap example where strap is located above the left and right probes.

2) Die-short

The short is implemented on the die rather than off as shown in Fig. 7. The length of the on-die short is an order of magnitude less than the off-die strap (150 μm vs. 1 cm), greatly reducing the parasitic inductance. Preferably, the short should be implemented in layout before the test structure is fabricated; however, if this is not possible a well-placed piece of copper foil will work.

Figure 7: Die-short probing technique.

9

3) GSSG RF Probe

The two signal needles of a ground-signal-signal-ground (GSSG) RF probe are placed on the two terminals of the device (Fig. 8). The ground probes are left to rest on the 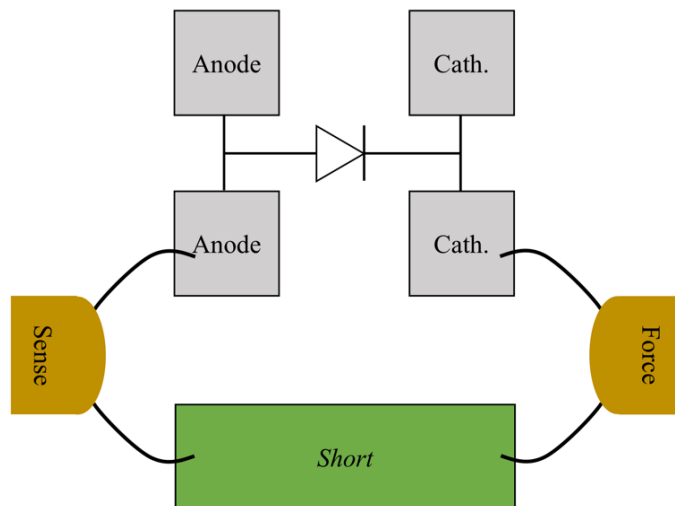passivation layer. The GSSG probe shorts the two signal lines' shields internally, so no external connection is required.
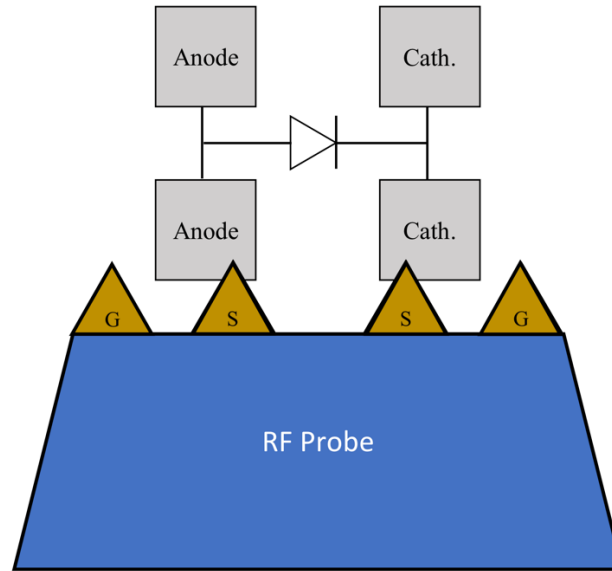
Figure 8: GSSG probing technique.

To estimate the effect of the inductance on the current waveform, a simulation of the TDT setup is performed in ADS [15]. The simulation schematic is shown in Fig. 9. The DUT is replaced with a 100 fF capacitor to mimic the behavior of a reverse biased diode, and the bias is omitted to satisfy the DC solution of the capacitor. The voltage pulse has a 100 ps rise-time and a maximum value of 10 V. Two inductor values are simulated, $L_p = 6$ nH and $L_p = 0$ nH. The former represents the shield strap probing method based on the dimensions of the short, and the latter the ideal scenario where there is no parasitic inductance. The resulting current waveforms for these two cases are shown in Fig. 10.
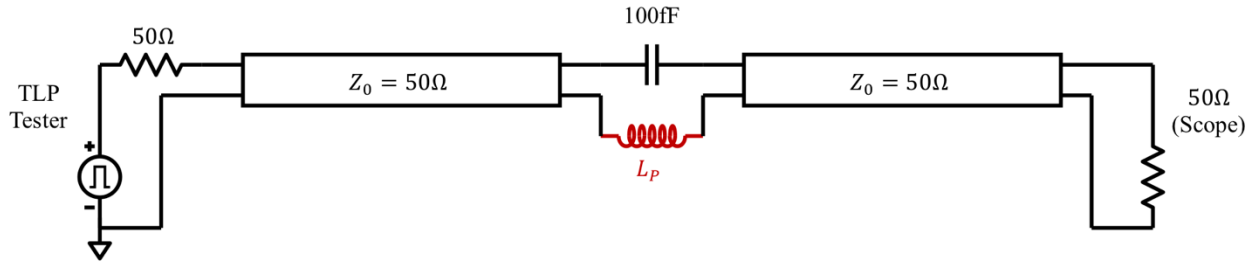
Figure 9: Schematic for simulating the current measured at the rightmost 50 Ω termination for different values of L_p.
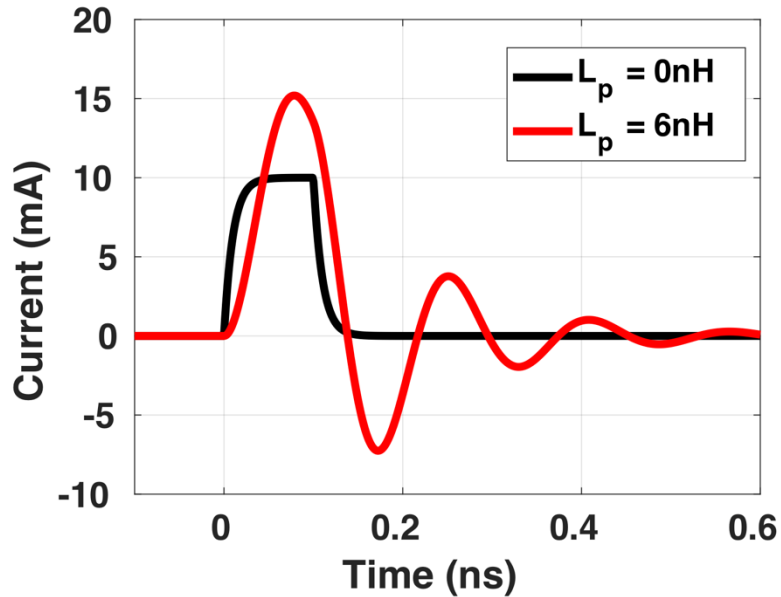
Figure 10: 100 Ω TDT simulated current with 100 fF DUT and $L_p = 0$ nH, 6 nH.

The simulation shows that an inductance of 6 nH causes the current waveform to oscillate which compromises the measured signal. To confirm this in measurement, the reverse recovery transient of a 90 nm P-well diode was measured with the three probing techniques outlined previously. The current waveforms for these measurements are shown in Fig. 11. The results confirm that the shield strap probing technique produces oscillations in the measured current, whereas the die-short and GSSG probing methods eliminate them. Additionally, a higher peak current is observed in the GSSG probe measurement because it has a higher bandwidth than the coaxial probes used in the die-short technique (40 GHz vs. 7 GHz). In the former case, the

11

overall bandwidth is limited by the 12 GHz oscilloscope. Clearly, the GSSG probe produces the best results, and therefore is the method that is used in the rest of this work.

Now that a measurement technique has been validated, it will be used to determine whether ESD device models, with an emphasis on diodes, are capable of predicting the measured reverse recovery characteristics.
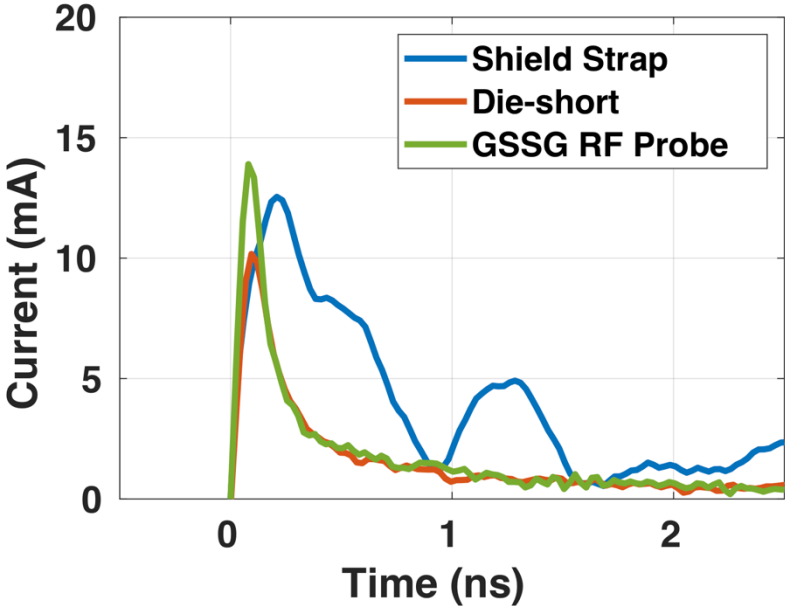

Figure 11: Reverse recovery of a 90 nm P-well diode using the three probing techniques.

CHAPTER 4: ESD DIODE MODEL REVERSE RECOVERY VALIDATION

ESD devices operate at high current with fast rise-times, and typically models (if they exist) offered by a foundry will not be well characterized in those regions of operation. Therefore, dedicated ESD models are needed for use in simulations. A popular, physics-based model for ESD diodes was proposed in [16], and its model equations have been adopted in other diode [17], PNP [18] and SCR [19]-[20] models. Henceforth the model in [16] is referred to as the "standard model". The procedure for fitting the standard model involves measuring a variety of TLP I-V curves and vf-TLP transient waveforms [16]. In particular, the latter is extremely important for characterizing transient overshoot voltage due to forward recovery during a fast rise-time event, such as CDM. Reverse recovery, however, is not typically measured for the standard model, and there have been no comprehensive studies on its ability to predict a reverse recovery transient. The goal of this chapter is to use the TDT measurement setup to determine how well the standard model predicts reverse recovery after going through its fitting procedure for a 65 nm P-well diode.

As stated before, the model is typically fitted to the TLP I-V curve and the transient overshoot voltage during vf-TLP. This procedure was carried out for the 65 nm P-well diode, where the TLP I-V measurement and simulation fit are shown in Fig. 12. Good correspondence is seen between measurement and simulation. Additionally, the measured and simulated voltage overshoot during forward recovery is shown for two different waveforms in Figs. 13 and 14. The model is able to well-predict the transient voltage overshoot.
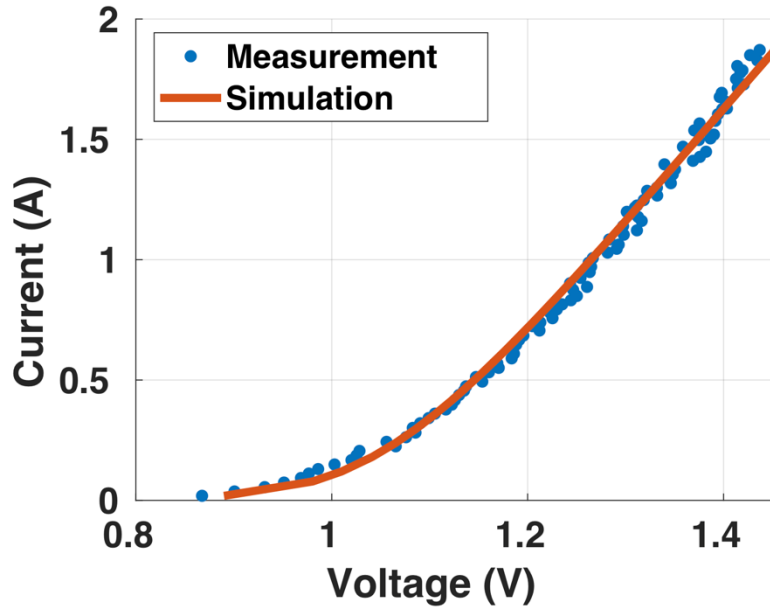
Figure 12: Standard model TLP I-V of 65 nm P-well diode. 100 ps rise-time, 5 ns pulse width.
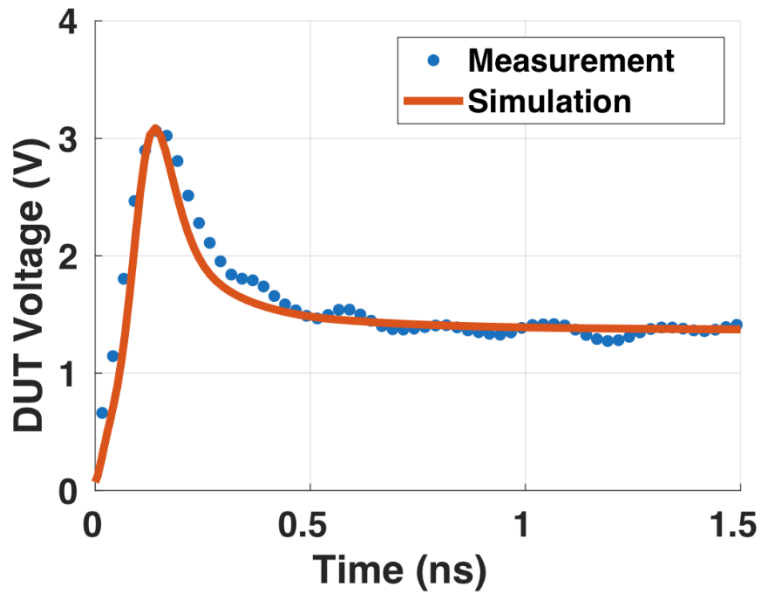

Figure 13: Standard model forward recovery voltage overshoot of 65 nm P-well diode. 1.4 A current injection, 100 ps rise-time.
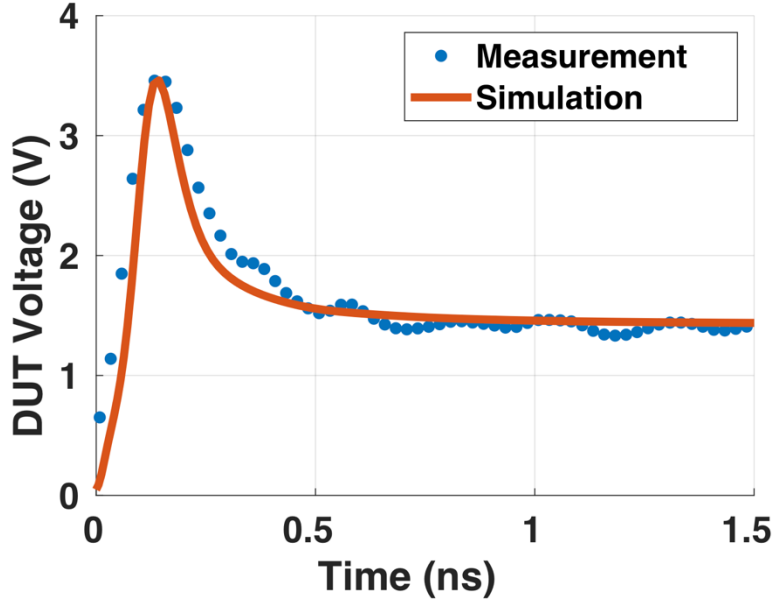
Figure 14: Standard model forward recovery voltage overshoot of 65 nm P-well diode. 1.7 A current injection, 100 ps rise-time.

Using the fitted model, reverse recovery was simulated and then compared with measurement, an example of which is shown in Fig. 15. Relative to the measurement, the simulation has a longer duration, higher peak current and a steeper decay. Fig. 16 and Fig. 17 show the recovery time (defined in Fig. 15) and the peak current vs. the reverse voltage for different forward currents, respectively. Inaccuracy is present across the bias conditions.

The simulated reverse recovery transient is dependent on how the charge in the quasi-neutral regions (Q_{diff}) of the diode is defined. The standard model uses the following formulation:

$$Q_{diff} = \tau I_d, \tag{2}$$

where τ is a fitting parameter related to the transit time and I_d is the conduction current. The transit time also controls the conductivity of a modulated series well resistor which predicts the forward recovery voltage overshoot, in addition to the velocity saturation.

15

$$R_{sm} = \frac{R_{sm0}}{1 + \dfrac{Q_{diff}}{Q_0}} \tag{3}$$

$$i_M = \frac{V_M}{R_{sm}\left(1 + \dfrac{V_M}{V_{SAT}}\right)} . \tag{4}$$

In (3) R_{sm} is the resistance of the well resistor and Q_0 is a fitting parameter. In (4) i_M is the total

current through the device, V_M is the voltage across the well resistor and V_{SAT} is a fitting

parameter.

The simulated peak current during reverse recovery has a linear relationship with the

reverse bias voltage (Fig. 17), which is not the case in measurement, and highlights the issue

with the standard model. The peak current will be a linear function of the reverse bias voltage if

very little charge is swept out of the device during the rising edge of the pulse, or equivalently

$$\tau \gg t_r, \tag{5}$$

where t_r is the rise-time of the pulse. If (5) is valid, the initial voltage across the diode, V_{d0}, will

not change significantly after the pulse has risen to its final value. Therefore, the peak current is

given by

$$I_{peak} = \frac{V_R + |V_{d0}|}{2Z_0}, \tag{6}$$

where V_R is the maximum value of the reverse pulse. Equation (6) is a similar to the classic

expression given in the chapter 2, Fig. 1. A transit time of 2.2 ns was needed to optimize the

forward recovery model. The rise-time is 100 ps, so (5) is satisfied and the model predicts the

peak current in accordance with (6). The poor agreement between measurement and simulation

confirms that this single time-constant model, which well replicates the forward transients, cannot accurately predict the reverse recovery transient.
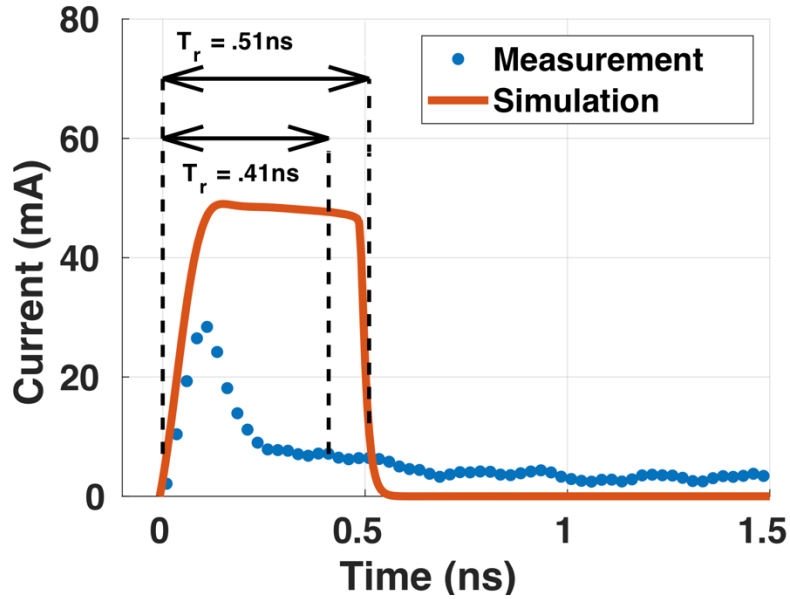
Figure 15: Standard model reverse recovery transient of a 65 nm P-well diode. The reverse bias voltage is 4 V and the current density, J_f, is 2.5×10^{-5} A/µm (normalized to the junction perimeter). The recovery time, T_r, is defined as the time for the current to decay to 25% of its peaks value.
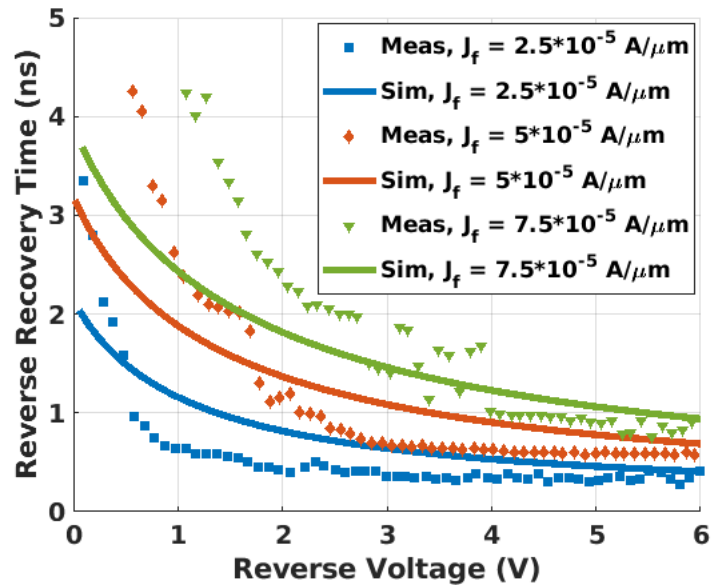

Figure 16: Standard model recovery time vs. reverse voltage for a 65 nm P-well diode biased at three current densities.
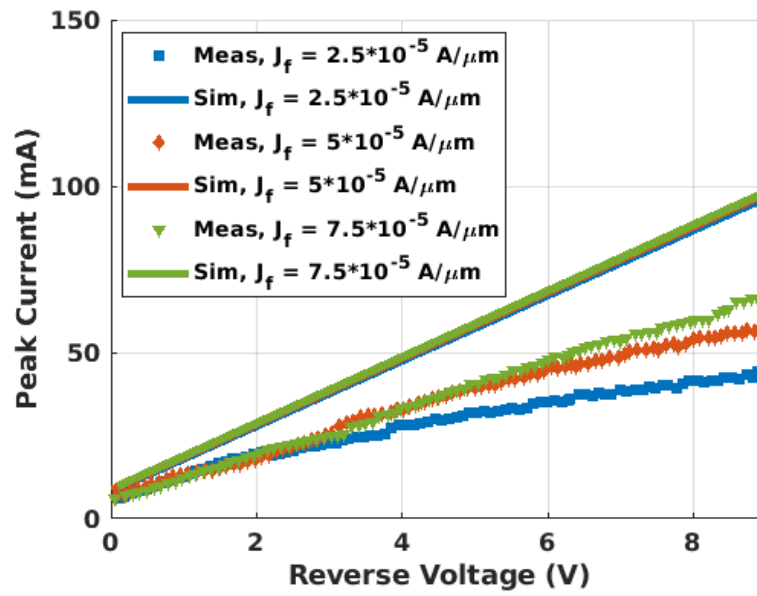
17

Figure 17: Standard model peak current vs. reverse bias voltage for a 65 nm P-well diode at three current densities.

CHAPTER 5: NON-QUASI-STATIC DIODE MODEL

As was shown in chapter 4, the standard model is incapable of accurately predicting the reverse recovery transient, and this is due to the quasi-static assumption in (2). This is a known issue acknowledged by Manouvrier et al. in [16], where they stated that the quasi-static diffusion charge model was unlikely to capture the reverse recovery transient, and chapter 4 has proven that conclusively. Other works, such as [21], have demonstrated through TCAD simulations and measurement that the quasi-static assumption is inaccurate for ESD devices, although [21] only shows that during forward recovery. To capture both forward and reverse recovery, the diode model must be changed from a quasi-static to non-quasi-static (NQS). For more information on the NQS model, see [22] and [23]. For the purposes of this work, only the results of the NQS will be discussed.

The forward recovery overshoot voltage of the diode was simulated using the NQS model in Fig. 18, where the measurement data is the same as that used for the standard model. The NQS model gets a good match between measurement and simulation. In addition, the reverse recovery was simulated under three different bias conditions shown in Figs. 19-21. The NQS model, with the same model parameters as the forward recovery, can well capture the initial reverse recovery transient. There is a slight discrepancy in the decaying edge of the current transient, especially at higher voltage bias, and this is due to the electrons stored deep in the P-well of the diode recombining at a much slower rate than the initial sweeping action near the junction [22]-[23]. It is not critically important to model this for purposes of ESD simulation, however, as peak current is the most crucial element.

As a future study, the NQS model can be tested with an oscillatory stimulus such as OTLP [24]. Appendix D outlines a simplified measurement setup for simulating the OTLP response of any arbitrary DUT.
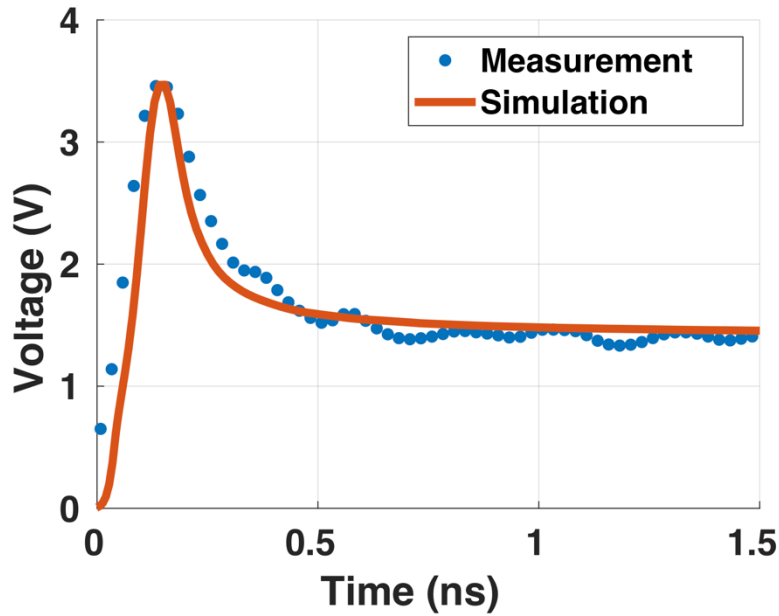
Figure 18: NQS model simulated vs. measured forward recovery for 65 nm P-well diode under a 1.7 A current injection.

Figure 19: NQS model simulated vs measured reverse recovery for 65 nm P-well diode with a forward current density of $J_f = 2.5 \times 10^{-5}$ A/μm and reverse bias voltage of 4 V.
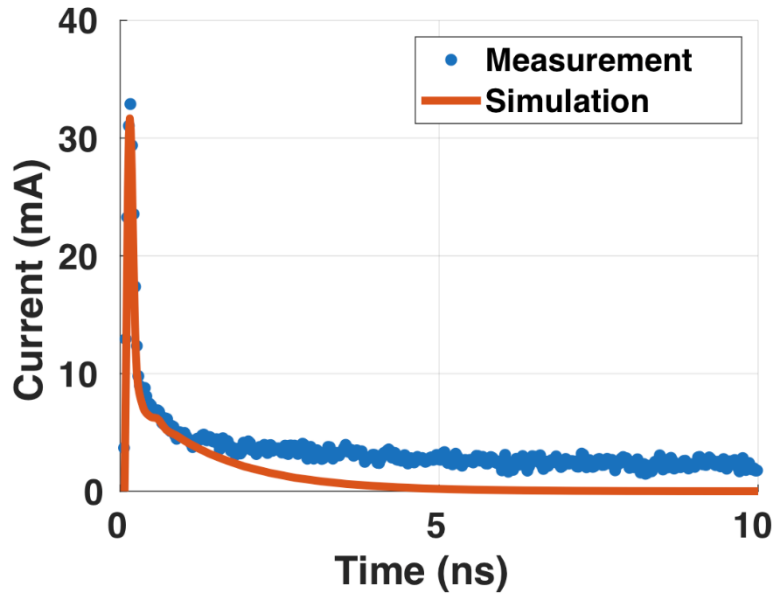
Figure 20: NQS model simulated vs measured reverse recovery for 65 nm P-well diode with a forward current density of $J_f = 2.5 \times 10^{-5}$ A/µm and reverse bias voltage of 6 V.

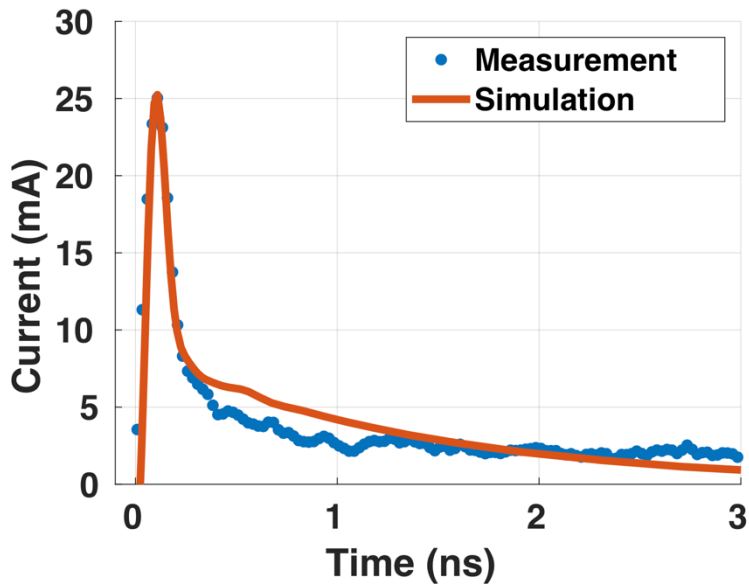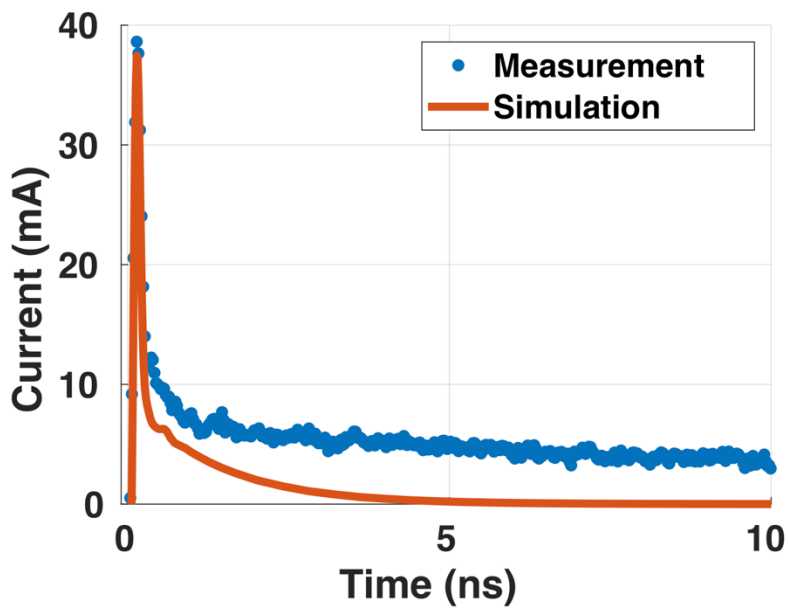

Figure 21: NQS model simulated vs measured reverse recovery for 65 nm P-well diode with a forward current density of $J_f = 2.5 \times 10^{-5}$ A/µm and reverse bias voltage of 8 V.

21

CHAPTER 6: CONCLUSION

A sub-nanosecond reverse recovery measurement technique was developed and used to evaluate ESD compact models' accuracy. The results of this work demonstrate that a compact model may fail to predict transients it was not directly fitted to, and point toward a fundamental weakness in existing modeling methodology, which relies on TLP and vf-TLP. Optimizing model equations to a narrow set of stimuli limits the predictive scope to only simulations with similar excitations.

By expanding the set of stimuli, in this case the inclusion of reverse recovery, not only is there more data available for fitting, but physical errors in the model can be more readily identified. Correcting those physical errors, in this work the quasi-static assumption, brings the model closer to physical reality and therefore is better able to aid a designer when simulating ESD devices.

PART II

CHAPTER 7: ACTIVE RAIL CLAMP BACKGROUND

A basic dual-diode ESD protection scheme is shown in Fig. 22. There must exist a safe discharge path between any two pads for the chip to be protected. The component responsible for the discharge from V_{DD} to V_{SS} is called the power clamp.
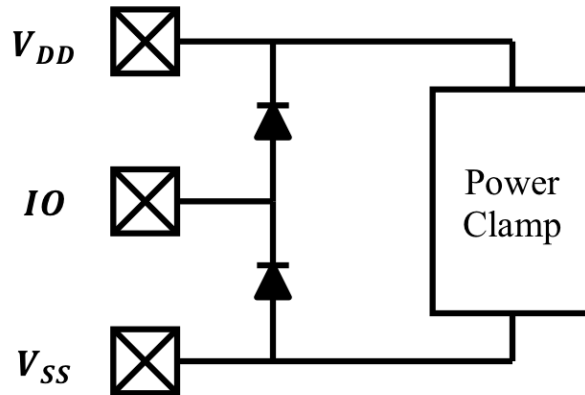
Figure 22: Dual-diode ESD protection scheme.

An active rail clamp [25] is a popular power clamp implementation, a block diagram of which is shown in Fig. 23. The detector block is responsible for triggering the clamp during an ESD event and is typically made up of an RC network. If an ESD event is detected, the bias block, typically made up of inverters, will bias the clamp to start discharging the power rail. The clamp is the device where most of the ESD current flows and is typically an NMOS transistor.
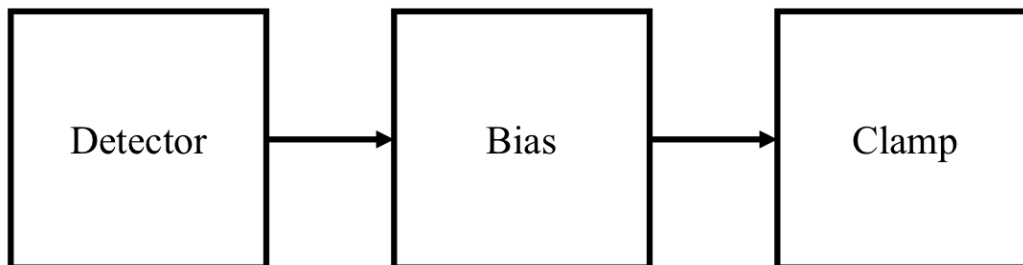
Figure 23: Active rail clamp block diagram.

During a discharge, the active clamp's bias is maintained by the detector, and has a natural time-out based on the time-constant of the RC circuit. The time-constant must be larger than the duration of the ESD stress, so it is designed to be greater than 1 μs to protect against an HBM event. This causes the detector circuit to be large in area. One method for reducing the detector area is to use "active feedback" to extend the bias beyond the RC time-constant [26]. The general idea is shown in the block diagram of Fig. 24. The primary signal path of the detector, bias and clamp blocks are the same as for the traditional active clamp. Positive feedback is added to "latch" the clamp into its discharge state after the detection of an ESD event. The detector need only be large enough to initially sustain the bias until the positive feedback activates. This can lead to significant area reductions in the footprint of the circuit [26].

Rather than having a time-out from the detector, the active feedback clamp will turn off when there is no longer a sufficient Vdd voltage, which naturally occurs once the ESD has been removed. This presents a danger during normal operation, however, because if the clamp were to "mis-trigger" and start conducting current due to some noise on the power rail, there would be no way of turning it off. This would cause errors or damage to the chip. Several subsequent designs [27]-[29] have added negative feedback elements to compensate for this danger. The negative feedback acts as another time-out mechanism that shuts down the clamp after a specified time.

A tradeoff exists in designing the relative strength of the positive and negative feedback. If the negative feedback is much stronger than the positive feedback, then it will turn off the clamp before the ESD stress has been fully dissipated. If it is too weak, then the positive feedback will too easily mis-trigger. This tradeoff was investigated with a 28 nm CMOS high-voltage tolerant active feedback clamp, and several circuit design techniques are used to decouple the mis-trigger immunity from the ESD performance.
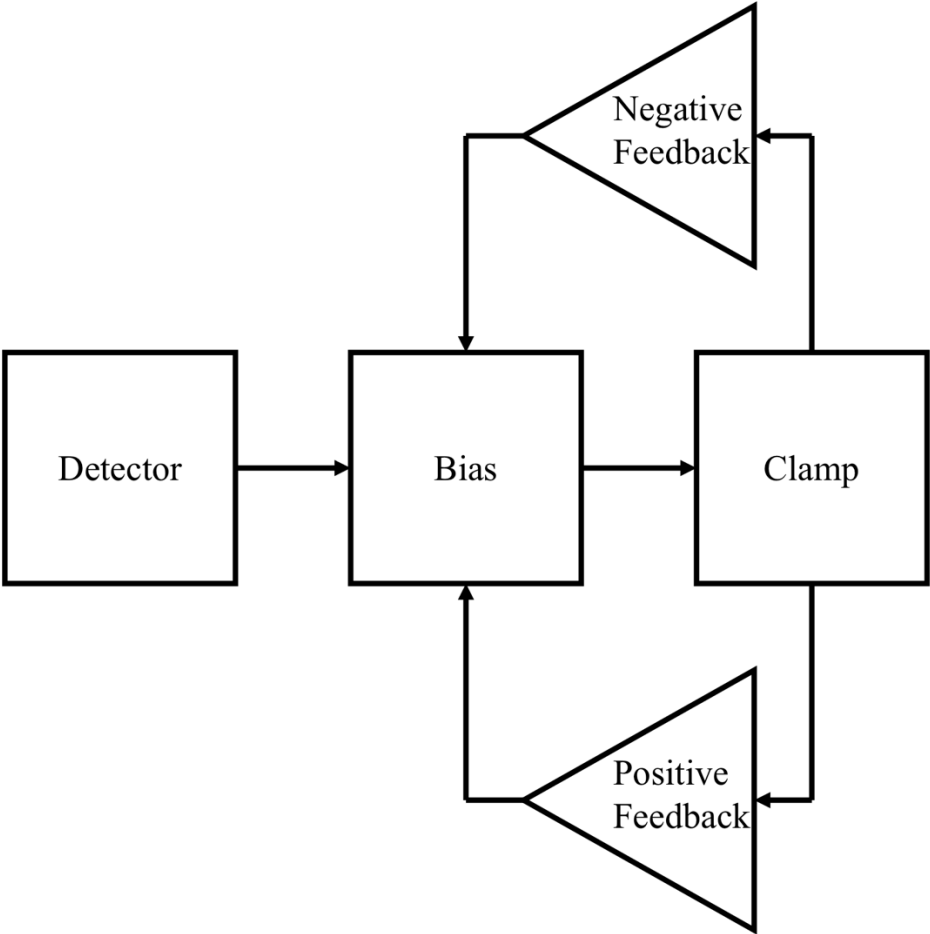
Figure 24: Active feedback power clamp block diagram.

CHAPTER 8: HIGH-VOLTAGE RAIL CLAMP WITH ACTIVE FEEDBACK

A critical design requirement of the clamp is that it must tolerate a DC voltage larger than the gate-oxide breakdown of a single I/O device. Specifically, the I/O transistors are rated for 1.8 V at DC. The operating Vdd is 3.3 V, so a cascoded topology is required to avoid damage to the devices. This specification is in addition to the other requirements of any clamp such as low leakage, compact area, small on-resistance etc. Two clamp designs are investigated. The first, based on [30], is referred to as the "original design" (Fig. 25) and will be used as a benchmark for the "new design" (Fig. 26) where changes to feedback have been made. Both designs are created in a 28 nm CMOS technology.

The original design detects a transient signal on Vdd with the RC detector composed of C2 and PMOS diode-connected transistors M9 and M10 which act as resistors. The detector time constant is approximately 30 ns. Node RC2 will rise instantaneously in response to an ESD transient and forward bias the base-emitter junctions of Q1, Q2 and Q3 which will buffer the signal on VG1 and VG2. The lateral BJTs are a standard foundry offering whose cross section is shown in [30].

VG1 and VG2 bias the large width MOSFETs M13 and M14 that are responsible for discharging the ESD current. VG1 and VG2 will also bias the inverter comprised of M15, M16 and R3, causing VFB1 and VFB2 to be pulled towards Vss. VFB1 and VFB2 are applied to the gates of M11 and M12, which hold RC2 at Vdd. The loop formed by M11, M12 and the inverter is the positive feedback loop.
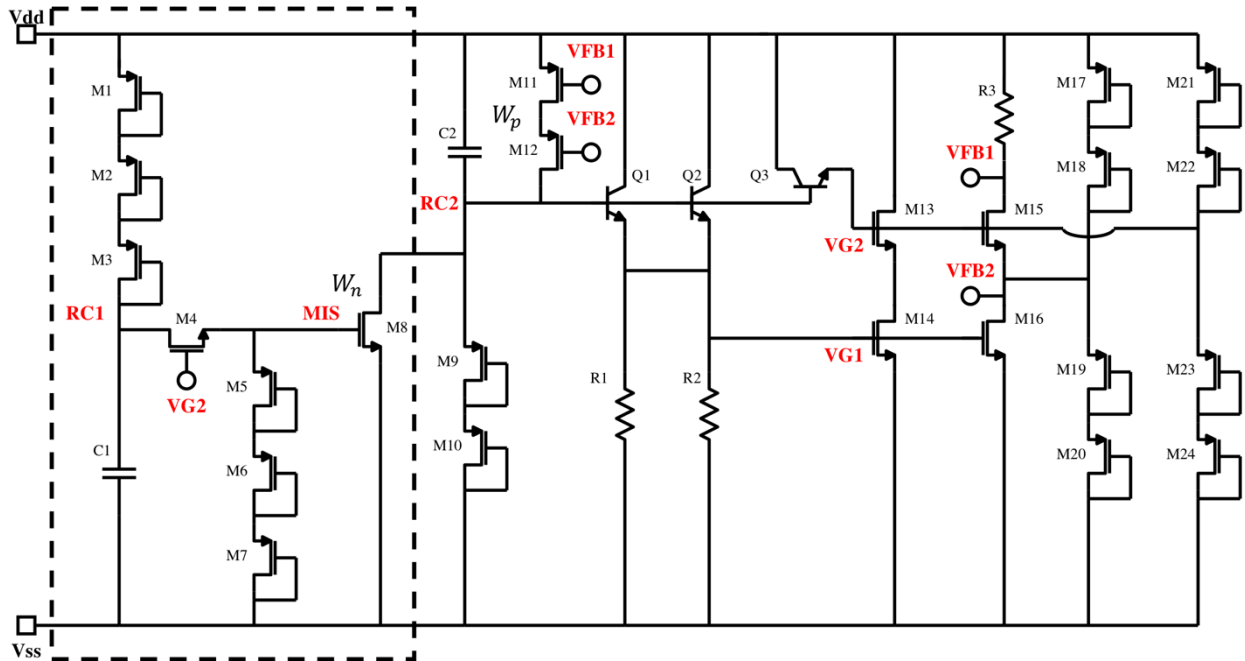
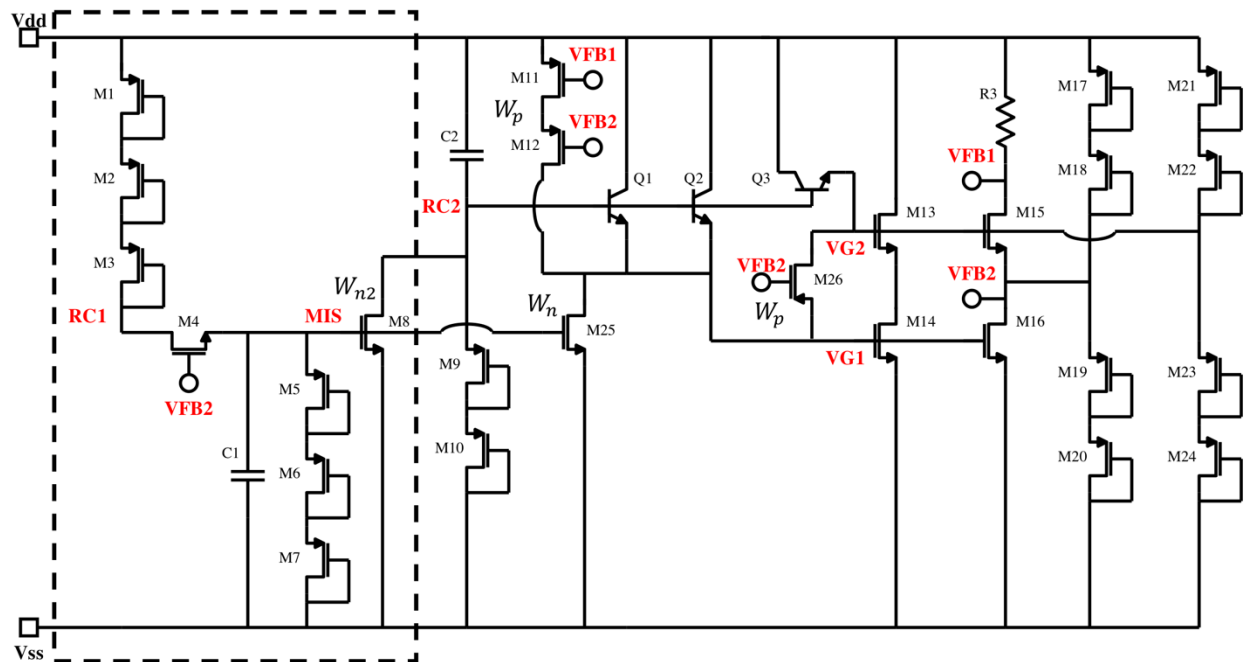Figure 25: Schematic of the original design, based on [30].

Figure 26: Schematic of the new design.

The dashed sub-circuit on the left side of Fig. 25 (and Fig. 26), in contrast, prevents the clamp from being "mis-triggered". A mis-trigger is defined as the clamp conducting a sustained current outside of an ESD event due to the triggering of the positive feedback. A mis-trigger can

27

occur either from transients on the Vdd rail during normal operation or when Vdd is initially energized by a voltage ramp. Transistor M8 is responsible for limiting the voltage on VG2 and VG1 during those scenarios to block the active feedback from triggering. The mis-trigger mechanisms are discussed further in chapter 9.

The time-constant at RC1 and the width of M8 must be carefully optimized to balance the performance and mis-trigger immunity. That trade-off can be significantly relaxed by decoupling the mis-trigger protection from the rest of the clamp during an ESD event. This is done in the new design by biasing switch M4 with VFB2 (VG1 in the original design) and placing C1 at the source side of M4 (Fig. 26). In response to ESD, the active feedback will trigger faster than the mis-trigger protection. Therefore, the active feedback will force M4 off before a significant voltage develops on RC1, which keeps M8 from turning on and eliminates its influence on the rest of the circuit for the duration of the discharge.

The new design additionally optimizes the biasing of M13 and M14. In [31] it was shown that the optimal gate basing during ESD operation of the cascoded NMOS transistors in a high-voltage tolerant design is at the Vdd rail. The BJTs in the original design introduce an offset in their buffering, biasing the VG2 and VG1 nodes at a reduced voltage. To remedy this, in the new design the active feedback is placed at the gates of M13 and M14 rather than at node RC2. M26 shorts VG2 and VG1 during ESD stress and isolates them during normal operation. The shift in the placement of active feedback requires an additional shunting transistor, M25, at node VG1. The resistors R1 and R2 are removed in the new design because M25 carries out the same function, mainly shunting node VG1 during normal operation. It will be shown in chapter 9, however, that the resistors do have some additional benefits when it comes to power-up mis-trigger immunity.

In the new design, capacitor C1 is placed at node MIS so that if the positive feedback

engages and turns off M4 during normal operation, a positive gate bias will be maintained on M8

and M25 due to the presence of C1 and the very slow discharge of MIS through the three high-

resistance PMOS diodes. This allows those transistors to continue limiting the voltage on VG1

and VG2.

The maximum length, minimum width transistors M17-M24 make up the voltage

dividers used purely for biasing the cascoded transistors during normal operation to avoid gate-

oxide breakdown. Additionally, the leakage current remains the same in the original and new

designs as it is primarily dependent on the M14, M15 and the diode voltage dividers, which do

not change.

The width W_n is defined as the M8 width in the original design and M25 in the new

design. The M8 width in the new design is defined as W_{n2}. These are important parameters

relative to the M11, M12 PMOS transistors widths (W_p) in optimizing the design and will be

discussed further in chapter 9.

To better understand the operation of the circuit, a TLP SPICE simulation was done [32].

Fig. 27 shows the simulated voltage waveforms for a 1.5 A current injection for the original

design. The triggering detector initiates the active feedback which pulls RC2 toward Vdd, but

RC1 is also charged which causes M8 to shunt the RC2 voltage towards Vss. This can be seen in

the downward decay of the RC2 waveform in Fig. 27. The decaying RC2 is buffered onto VG1

through the BJTs, increasing the clamping voltage until the bias is too small to sustain the

discharge and the clamp stops conducting.

Fig. 28 shows the same current injection applied to the new design. The active feedback

pulls VG1 and VG2 to Vdd, significantly lowering the voltage across the clamp relative to the

original design for the same amount of current. Additionally, because M4 is biased with VFB2 node MIS is isolated from RC1 which keeps M8 and M25 in subthreshold. Thus, in the new design VG2 and VG1 remain strongly biased for the duration of the stress.

To highlight the trade-offs in the two designs, a thorough optimization via SPICE simulations is done in chapter 9.
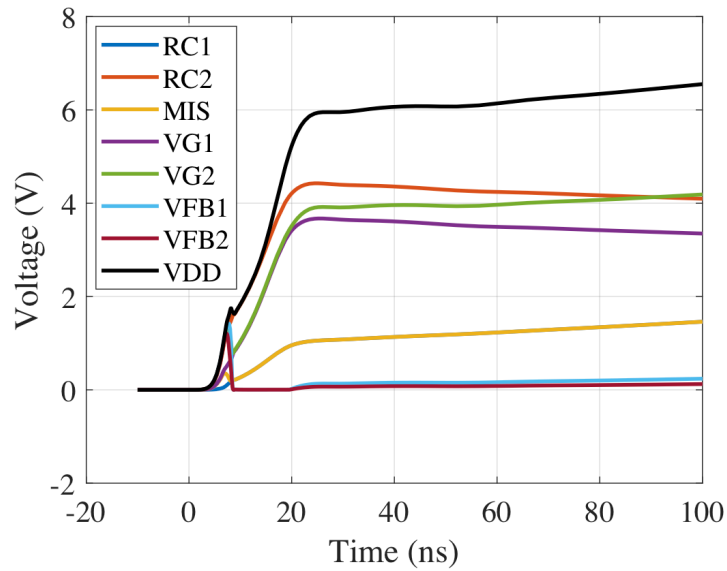

Figure 27: Original design simulated voltage waveforms with 1.5 A TLP current injection. 100 ns pulse-width, 10 ns rise-time.
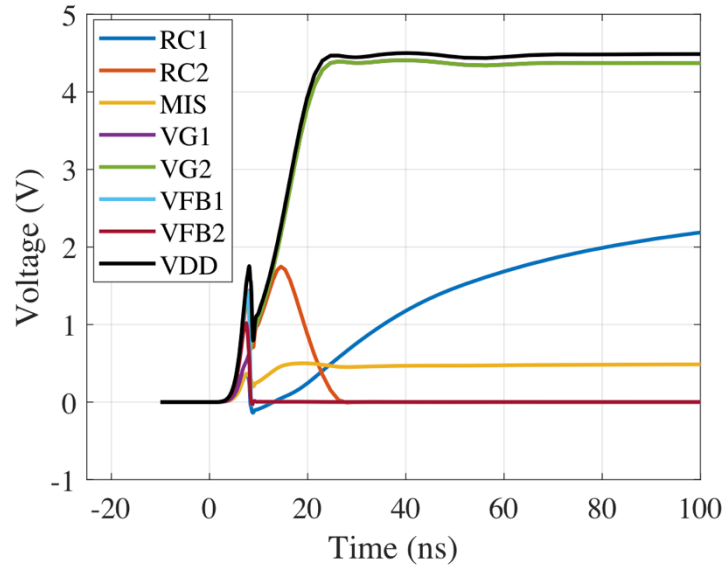
Figure 28: New design simulated voltage waveforms with 1.5 A TLP current injection. 100 ns pulse-width, 10 ns rise-time.

CHAPTER 9: DESIGN OPTIMIZATION

The key design parameters in terms of ESD performance and mis-trigger immunity are the size of capacitor C1 and the relative widths of the PMOS pull-up transistors (W_p) and NMOS pull-down transistors (W_n) as shown in Fig. 25 and Fig. 26. The parameter W_p is sized such that the active feedback pull-up transistors are capable of quickly biasing M13 and M14 during ESD operation. Given that W_p sizing, the magnitude of W_n is normalized to W_p and represents the relative strength of the positive and negative feedback. The ratio W_n/W_p, is the parameter to be optimized in simulation. In the new design, there is an additional pull-down transistor whose width is denoted as W_{n2} in Fig. 26 which will also be normalized to W_p and swept in simulation.

TLP

The first simulation done is TLP as explained in Appendix A. TLP is useful for extracting the on-resistance of the clamp and the expected clamping voltage for a given current injection. The critical phenomenon for these clamp designs, especially the original, is the tendency of the on-resistance to increase due to the influence of the mis-trigger protection. A transient plot of this process was shown for the original design in Fig. 27 in chapter 8.

Three TLP I-V curves of the original design with W_n/W_p = 0.1, 1 and 2 were simulated and are shown in Fig. 29. The pulse width is 100 ns and the rise-time 10 ns. It should be emphasized that the models used in the simulations do not include the effects of self-heating, and therefore the increase in on-resistance is primarily due to the influence of the mis-trigger protection. This increase in on-resistance is evident in the "folding" of the TLP curve for higher current injections, which gets worse as W_n/W_p increases. All of the curves in Fig. 29 have the same on-resistance at lower current injections, and this minimum resistance is defined as R_{on0}. The current injection at which the on-resistance becomes $2R_{on0}$ is defined as the "folding

current", an example of which is shown in the blue curve in Fig. 29. Clearly, a higher folding current is desirable in terms of lower clamping voltage at higher ESD currents. Because the widths of M13 and M14 are the same in the new design, it will have approximately the same on-resistance. Therefore, the same folding current definition is used for TLP simulations of the new design.

A plot of the folding current versus W_n/W_p for the original and new designs is shown in Fig. 30. Three values of C1 capacitance are simulated, where C_0 is the nominal capacitance in the original design (approximately 400 fF). The results show that as W_n/W_p increases, the folding current decreases. This intuitively makes sense, as the negative feedback becomes stronger and more easily quenches the active feedback. Additionally, as C1 increases the folding current increases, and this is because the voltage that develops on node MIS is reduced at the sampling point of the transient simulations, so the pull-down transistors are biased at a lower voltage and their drive strength is reduced. Those same trends are seen in the new design but with much reduced sensitivity due to the M4 switch blocking the charging of node MIS during ESD operation. This not only improves the performance but, as will be discussed later, allows more flexibility in meeting other design requirements.

There is an additional parameter that can be swept in the new design, W_{n2}/W_p, and simulations of the folding current are shown in Fig. 31 for different W_{n2}/W_p. Increasing W_{n2}/W_p, in contrast to W_n/W_p, does not reduce the folding current. In fact, the folding current actually increases with W_{n2}/W_p in Fig. 31. This is because M8 in the new design does not directly contend with the active feedback and therefore cannot degrade the VG2 and VG1 bias. It can, however, add capacitance to node MIS, which has the same effect as simply increasing the size of C1.
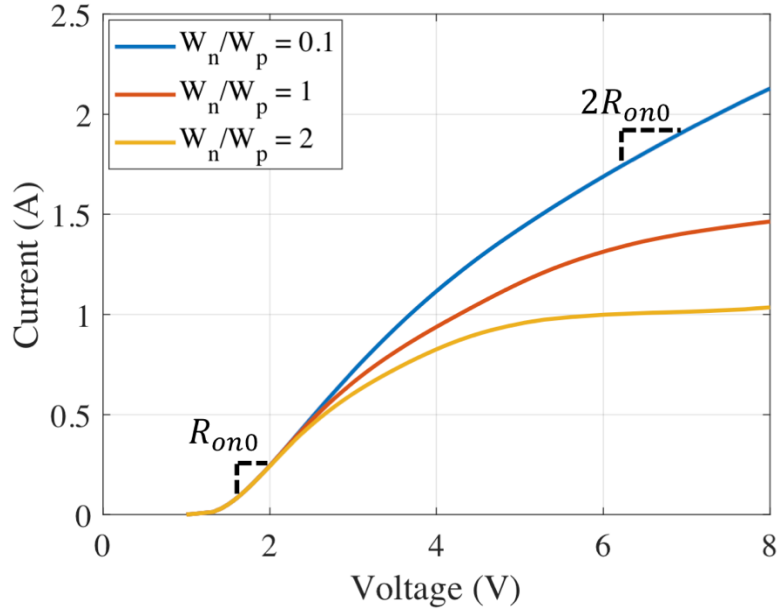
33

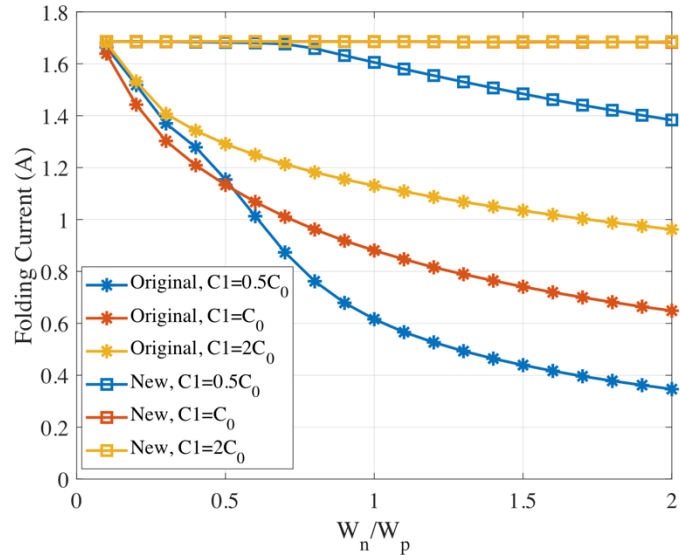Figure 29: TLP I-V curve of original design with W_n/W_p = 0.1, 1 and 2 and C1 = C_0.

Figure 30: Folding current of original and new design (W_{n2}/W_p = 1) versus W_n/W_p with C1 = 0.5C_0, C_0 and 2C_0.
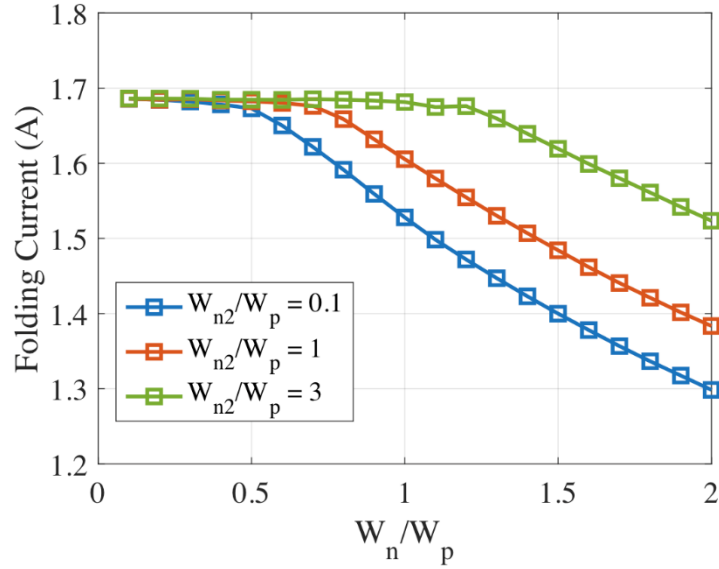
Figure 31: Folding current of new design versus W_n/W_p with $W_{n2}/W_p = 0.1$, 1 and 3 - C1 = $0.5C_0$

Power-on TLP

When the clamp is operating at nominal Vdd, it must not trigger due to any fluctuations in the Vdd voltage. To characterize this situation both in simulation and later in measurement, a power-on TLP approach is used [33], the schematic of which is displayed in Fig. 32. The setup is similar to the TDT measurement in Part I, where a bias tee is used to supply some DC voltage through a large inductor in addition TLP pulses through a large capacitor. The DC voltage bias will be nominal Vdd, or 3.3 V. A transient simulation of this schematic for the original design with $W_n/W_p = 0.5$ was done with pulse voltages of 4.0 V and 4.7 and a rise-time of 100 ps (Fig. 33).
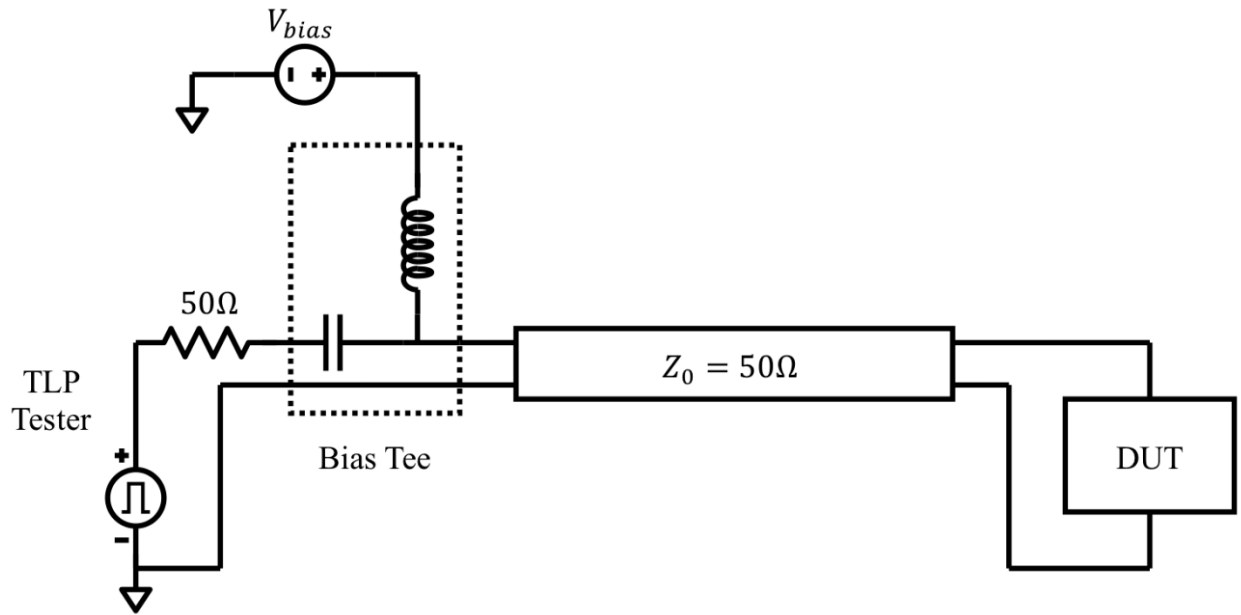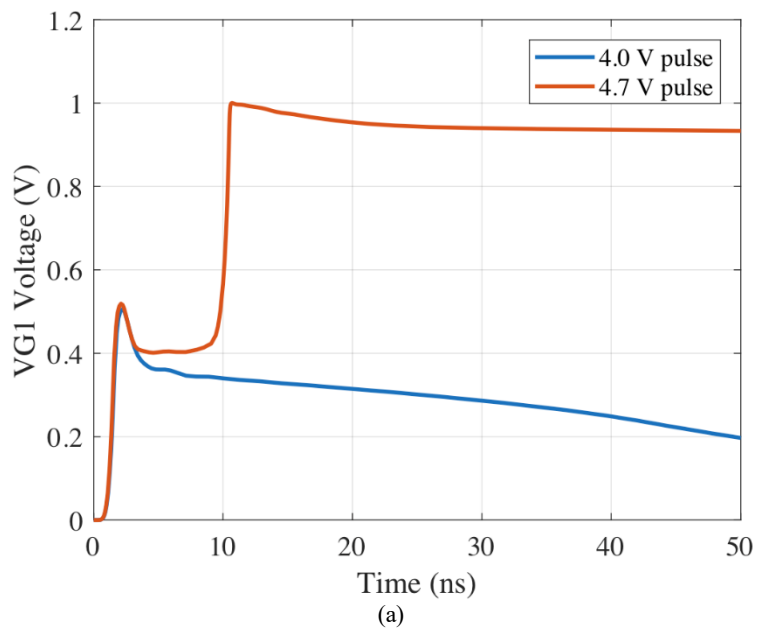
Figure 32: Power-on TLP schematic.
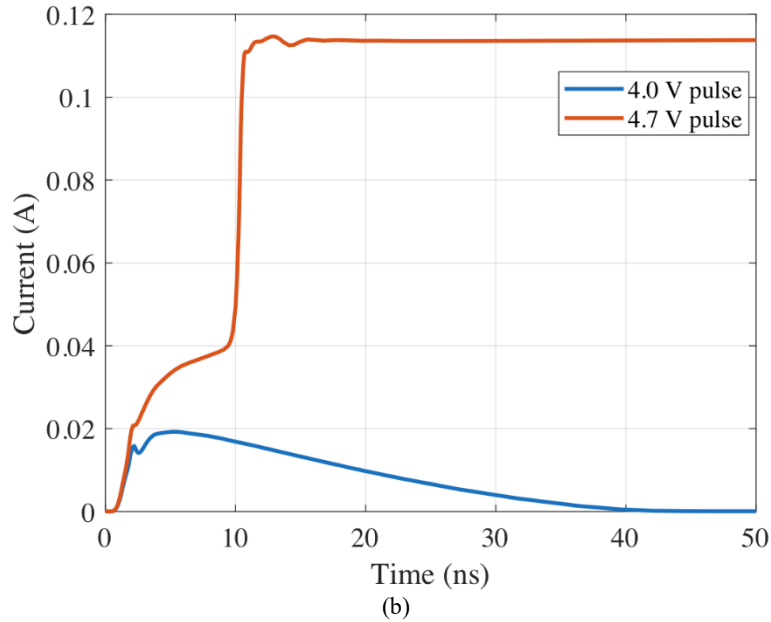

(a)

Figure 33 (cont.)

Figure 33: Power-on TLP simulation of VG1 voltage (a) and DUT current (b) for a 4.0 V and 4.7 V pulse, rise-time = 100 ps.

In the 4.0 V case, a voltage develops on VG1 but eventually decays back to zero, leading to a similar decaying transient current. In the 4.7 V case, the active feedback overpowers the mis-trigger protection and is latched to a high voltage that causes a sustained current to flow through the clamp, corresponding to a mis-trigger. A TLP I-V curve can be constructed from those transient simulations, and two examples are shown in Fig. 34 with $W_n/W_p = 0.5$ and $W_n/W_p = 1.5$.

For the $W_n/W_p = 0.5$ case, the I-V curves exhibit snapback behavior corresponding to the situation discussed in the 4.7 V power-on TLP pulse of Fig. 33. A "trigger voltage" (V_{t1}) can be defined at the onset of snapback for this type of I-V curve. A higher trigger voltage is desirable as it indicates a larger voltage excursion on the Vdd rail is required to force a mis-trigger. In the $W_n/W_p = 1.5$ case, however, there is no snapback. This is because the active feedback is incapable of overpowering the mis-trigger protection, and thus the clamp current always decays back to zero after a sufficient time. What is being measured in the I-V curve is this decaying current. For purposes of comparison, a trigger voltage will be defined for the non-snapback

37

curves as the voltage at which the current equals 10,000x the standby leakage current, or approximately 10 mA. These definitions are displayed on the I-V curves in Fig. 34.
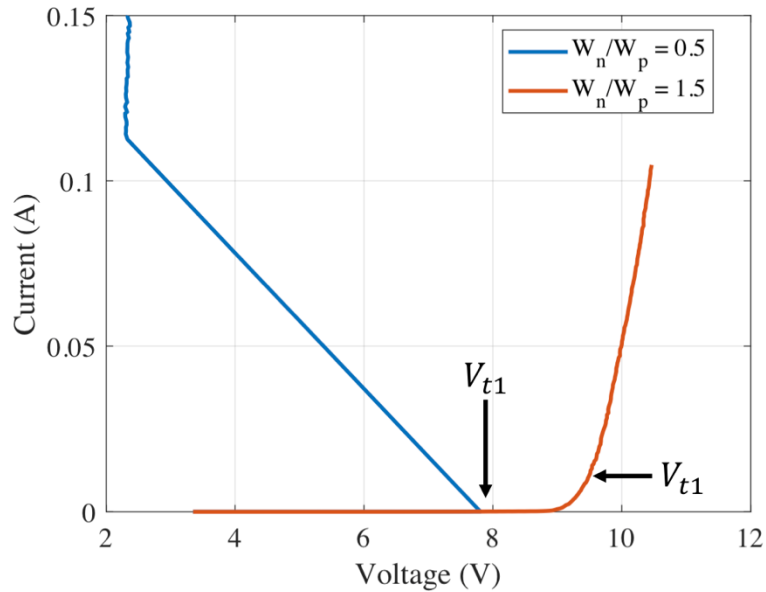
Figure 34: Original design power-on TLP I-V curves – W_n/W_p = 0.5, 1.5.

The same parametric sweeps of the folding current will be done for the trigger voltage during power-on TLP, but first the rise-time dependency should be discussed. The voltage coupled onto VG1 should increase as the rise-time is reduced, and this can be seen in Fig. 35 where the trigger voltage of the original design versus W_n/W_p is plotted for rise-times of 10 ns, 1 ns and 100 ps. Indeed, the trigger voltage is larger for the 10 ns rise-time as would be expected, but there is negligible difference between 1 ns and 100 ps. This indicates that continuing to decrease the rise-time beyond a certain point produces minimal reductions in the trigger voltage, and thus a 1 ns rise-time is effectively equivalent to 100 ps. For the subsequent simulations, a 100 ps rise-time is used.

Figure 35: Original design power-on TLP trigger voltage for rise-time = 100 ps, 1 ns and 10 ns - W_n/W_p = 1 and $C1 = C_0$.
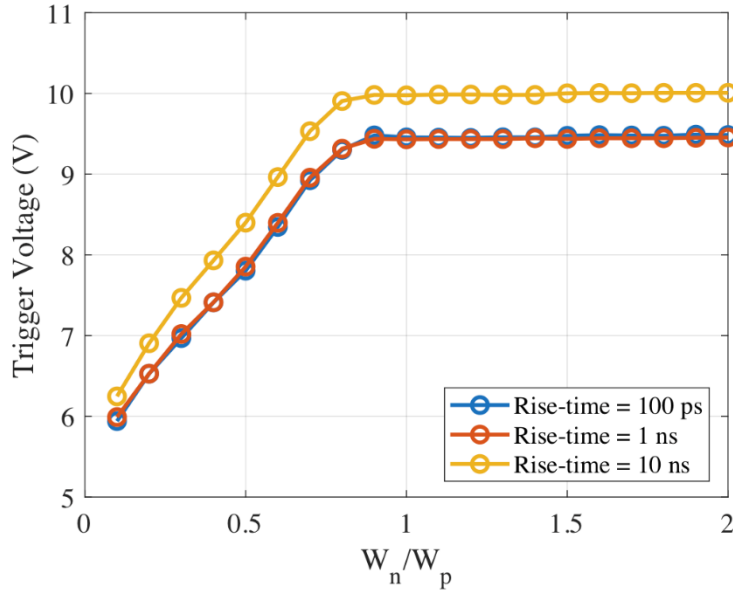
Parametric plots of the trigger voltage versus W_n/W_p can now be constructed for the original and new designs, as shown in Fig. 36. Two things are notable. The first is that there is little dependency on the value of C1. This can be understood by looking at the transient plot of Fig. 33. The mis-trigger occurs within 10 ns of the applied pulse, which is not enough time for the voltage across C1 to change enough to improve or degrade the mis-trigger immunity regardless of slight increases or decreases in the capacitance. Secondly, the trigger voltage of the original and new designs improves as W_n/W_p increases but saturates beyond W_n/W_p= 1 in the original design and W_n/W_p = 1.5 in the new design. Past those points, the clamp ceases to experience strong snapback and thus the trigger voltage does not improve with a further increase in W_n/W_p. This saturation occurs at a slightly higher voltage in the new design, but both are well above any realistic noise, even in extreme cases, expected during nominal operation.
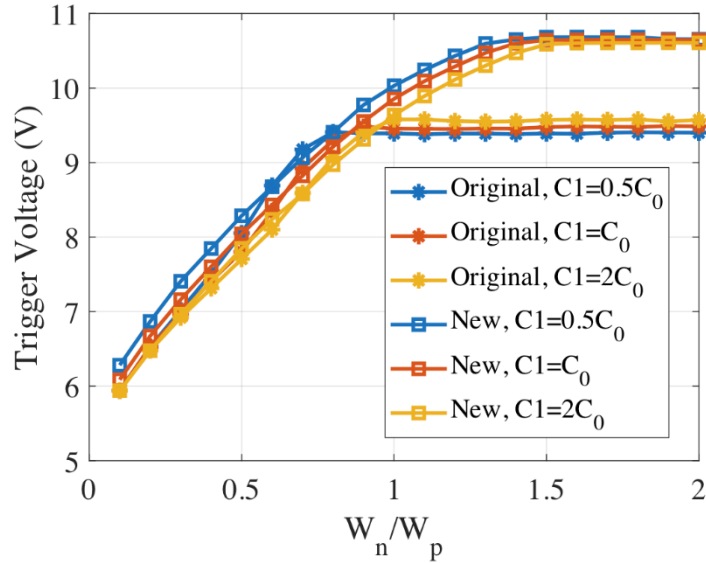
Figure 36: Power-on TLP trigger voltage of original and new design ($W_{n2}/W_p = 1$) versus W_n/W_p with C1 = $0.5C_0$, C_0 and $2C_0$.

In general, any improvement in the trigger voltage is in direct opposition to the folding current, and hence presents a trade-off between ESD performance and mis-trigger immunity. This is where the modifications in the new design become important. The folding current sensitivity to W_n/W_p is greatly reduced and thus a larger W_n/W_p can be used for more robust mis-trigger immunity without compromising ESD performance.

Finally, as was done for the TLP simulations, the parameter W_{n2}/W_p can be varied as well, a plot of which is shown in Fig 37. The trigger voltage has essentially no dependency on W_{n2}/W_p, and this is because it is not in direct contention with the active feedback like W_n/W_p is. This makes W_{n2} an interesting design variable because no tradeoff exists for it. As will be discussed in the next section, this property is attractive for improving the power-up mis-trigger immunity, where W_{n2} plays a more prominent role.
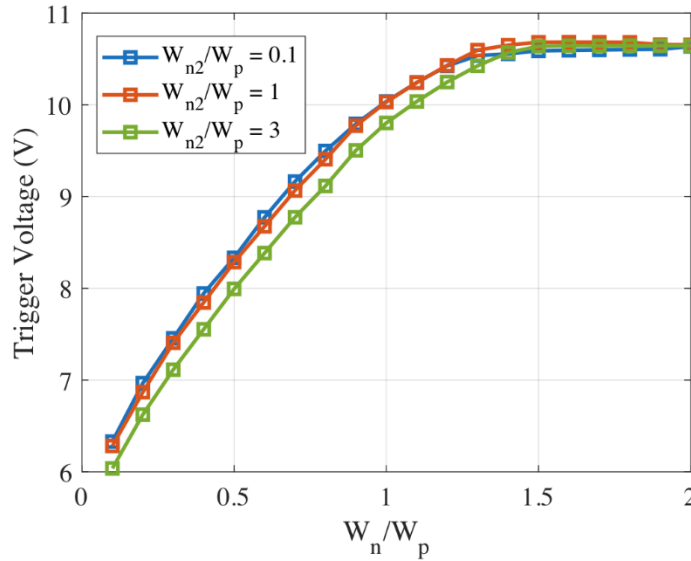
Figure 37: Power-on TLP trigger voltage of new design versus W_n/W_p with $W_{n2}/W_p = 0.1$, 1 and 3 - $C1 = 0.5C_0$.

Power-up Transient

There is another potential mis-trigger hazard during the power-up voltage ramp on Vdd, where the clamp must be given a sufficiently slow rise-time such that it does not start conducting current. This is defined as the power-up mis-trigger. There exists a rise-time (10%-90%) threshold such that any Vdd ramp above this value will not cause a power-up mis-trigger. Conversely, any rise-time including and below the threshold will always cause a power-up mis-trigger. Therefore, the rise-time threshold can be thought of as the maximum Vdd rise-time that forces a mis-trigger.

To simulate the power-up mis-trigger the schematic in Fig. 38 is used. A 50 Ω pulse generator supplies a variable rise-time voltage ramp to the clamp, where its terminal current and voltage are monitored. At the rise-time threshold, the active feedback triggers, causing a sustained current to flow through the clamp that collapses the Vdd rail. This can be seen from an example simulation of the original design in Fig. 39, where the voltage ramp suddenly drops as the active feedback engages and forces the clamp to conduct current.

It may be alarming to see the 50 Ω series impedance in the schematic upon first glance, as the actual impedance of the power delivery network (PDN) would be much less in any practical design. The 50 Ω impedance is used in the schematic because that same setup will be used for measurement, where the pulse generator will have a 50 Ω impedance. The impedance can be safely neglected, however, because the impedance of the clamp up to and including the exact point in time when it mis-triggers is much larger than the 50 Ω impedance. This can be seen in the plot of the current in Fig. 39, where essentially zero current flows through the device and all of the voltage from the source develops across the terminals. Thus, the only discrepancy occurs after the mis-trigger, where the current will be limited to 40 mA by the series impedance, whereas with a more realistic PDN impedance the clamp would conduct a huge amount of current and likely be damaged. Regardless, the rise-time threshold extracted from this setup will be the same.
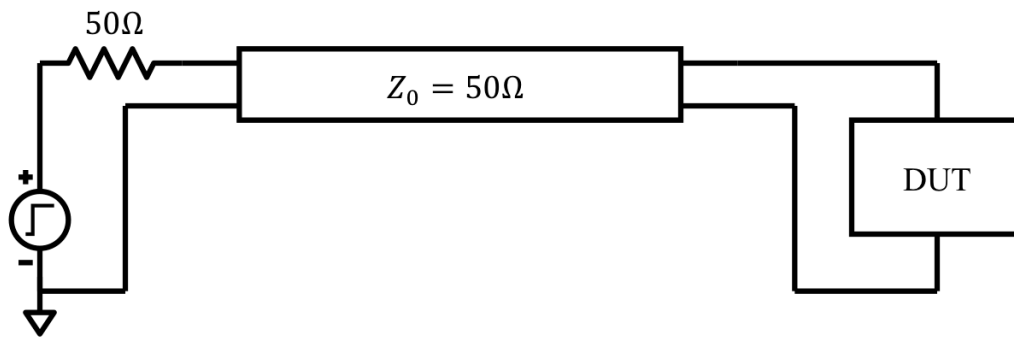

Figure 38: Power-up voltage ramp measurement schematic.
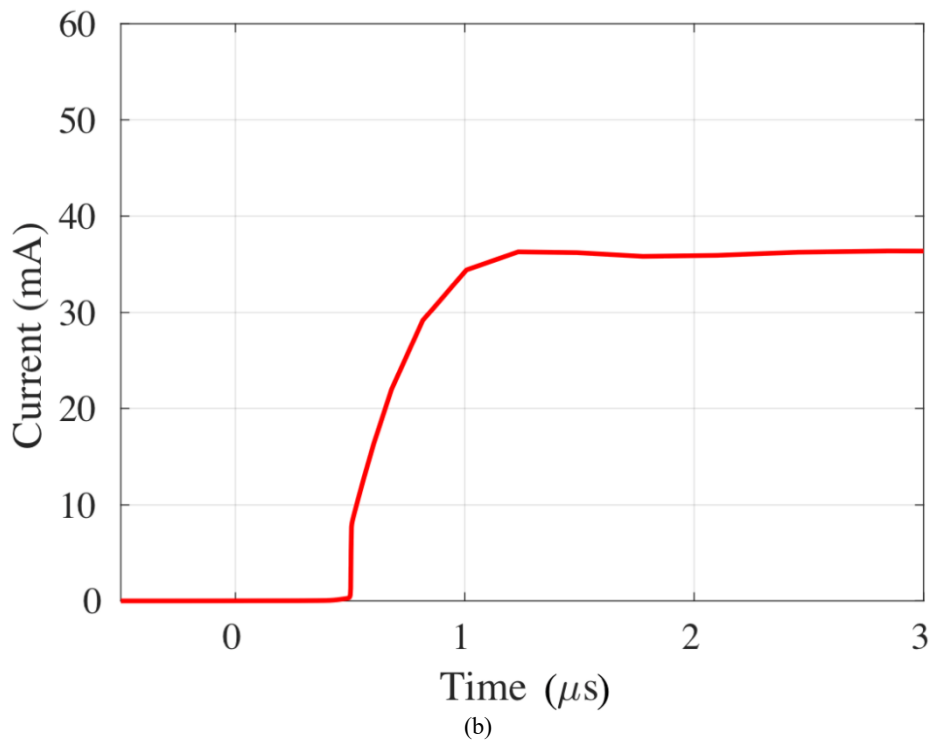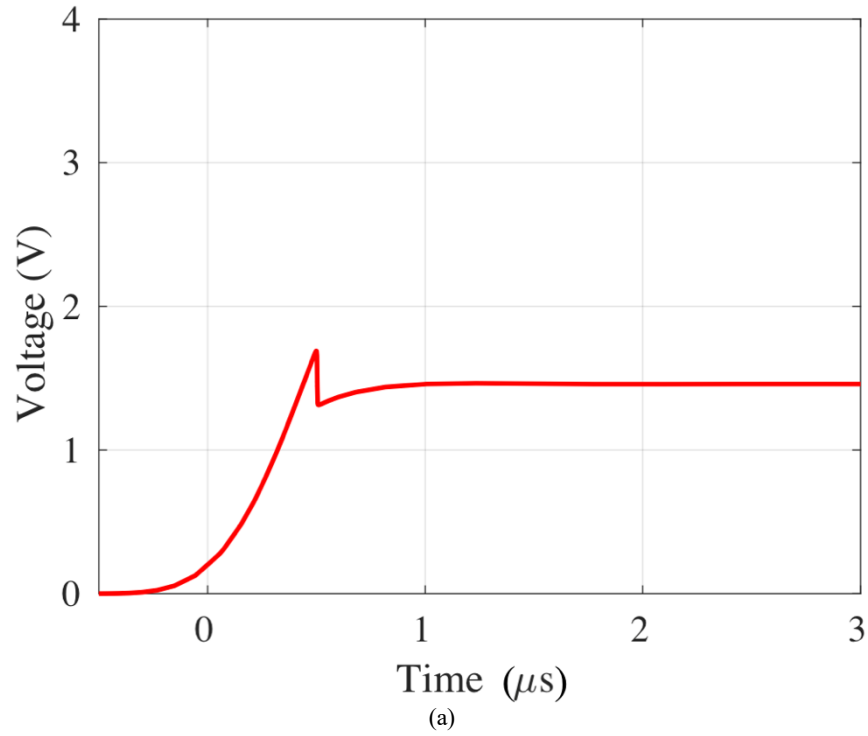
(a)

(b)

Figure 39: Power-up mis-trigger voltage (a) and current (b) for original design with $W_n/W_p = 1$ and $C1 = C_0$.

The rise-time threshold, as was the case for the folding current and trigger voltage, depends on C1 and W_n/W_p. A plot of the rise-time threshold versus W_n/W_p for the original and

new designs is shown in Fig. 40 and Fig. 41. In general, the rise-time threshold decreases as W_n/W_p increases, which is expected as the mis-trigger protection is stronger at resisting the active feedback. The rise-time threshold, in contrast with the trigger voltage, has a strong dependency on the C1 capacitance. This is due to the fact that the bias on node MIS must be charged up during the voltage ramp. Node MIS determines the strength of the mis-trigger protection, and therefore as C1 capacitance increases it accumulates less voltage for a faster rise-time and weakens the mis-trigger protection.

The new design's rise-time threshold has the undesirable property of much greater sensitivity to changes in the parameters than the original design. This is also the case for W_{n2}/W_p, where increasing that parameter reduces the rise-time threshold (Fig. 42). For the case of $C1 = 0.5C_0$ and $W_{n2}/W_p = 3$, the rise-time threshold of the new design is comparable to that of the original design, while trigger voltage is similar and the folding current much improved. The question is why the sensitivity is so much greater in the new design. As was alluded to before, the answer is the removal of the poly resistors R1 and R2.
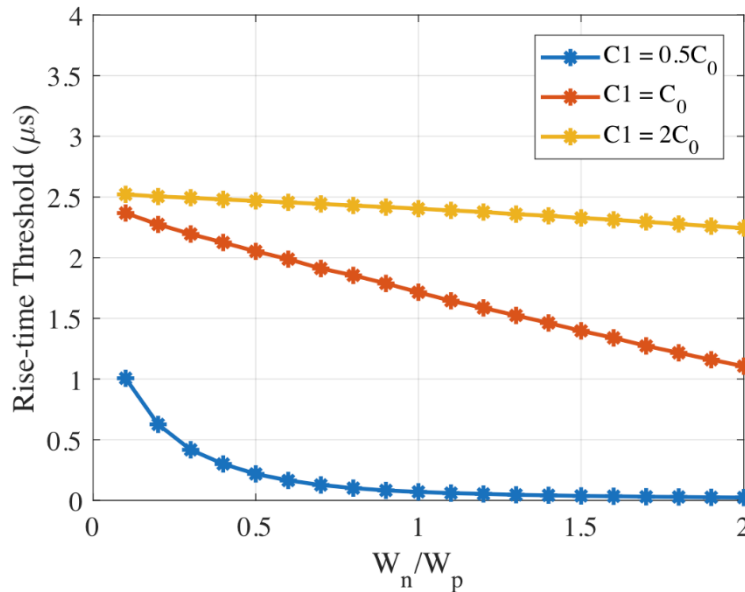
Figure 40: Power-up rise-time threshold of original design versus W_n/W_p with $C1 = 0.5C_0$, C_0 and $2C_0$.

Figure 41: Power-up rise-time threshold of new design ($W_{n2}/W_p = 1$) versus W_n/W_p with C1 = $0.5C_0$, C_0 and $2C_0$.

Figure 42: Power-up rise-time threshold of new design versus W_n/W_p with W_{n2}/W_p = 0.1, 1 and 3 - C1 = $0.5C_0$.

During the crucial early stages of the Vdd ramp, M25 is in subthreshold and therefore its impedance is quite large, so a significant VG1 voltage is developed for extremely small currents, which degrades the mis-trigger immunity. Because M25 is in subthreshold, its impedance is highly dependent on its gate bias, transistor width and process variations which cause dramatic

changes in the rise-time threshold. When the resistors are included back into the design, they clamp the voltage on VG1 during this subthreshold stage. The resistors will not affect the clamp during an ESD event or nominal Vdd operation, as either the active feedback or the mis-trigger protection has a low impedance compared to the resistors (50 kΩ) and therefore they are negligible. This can be seen by plotting the folding current and trigger voltage with and without the resistors added into the new design (Fig. 43, Fig. 44). No significant difference is discernable. During the power-up transient, however, VG1 is clamped by the resistors while M25 is in subthreshold, improving the performance and reducing the sensitivity to C1 and W_n/W_p (Fig. 45). In fact, W_n/W_p only affects the rise-time threshold as a capacitor would, where increasing it actually increases the threshold rather than decreasing.

Sweeping W_{n2}/W_p for the new design with resistors shows that it will reduce the rise-time threshold (Fig. 46). This is an attractive property, as W_{n2} is essentially a free design parameter that can be used to meet the rise-time threshold target without compromising any ESD performance or power-on mis-trigger immunity. Unfortunately, the R1 and R2 resistors were not included in the fabricated test structures that will be discussed in chapter 10, as it was originally thought they were unnecessary due to their minimal effect on ESD performance and power-on mis-trigger immunity. In a future design, the addition of the resistors would likely lead to even greater improvements over the original design.

Regardless of whether the resistors are included, from the simulation sweeps it is clear that the changes made to the new design decouple the ESD performance and mis-trigger immunity, allowing significantly improved TLP I-V with similar and or better mis-trigger immunity compared to the original design. This will be validated with measurement results in chapter 10.
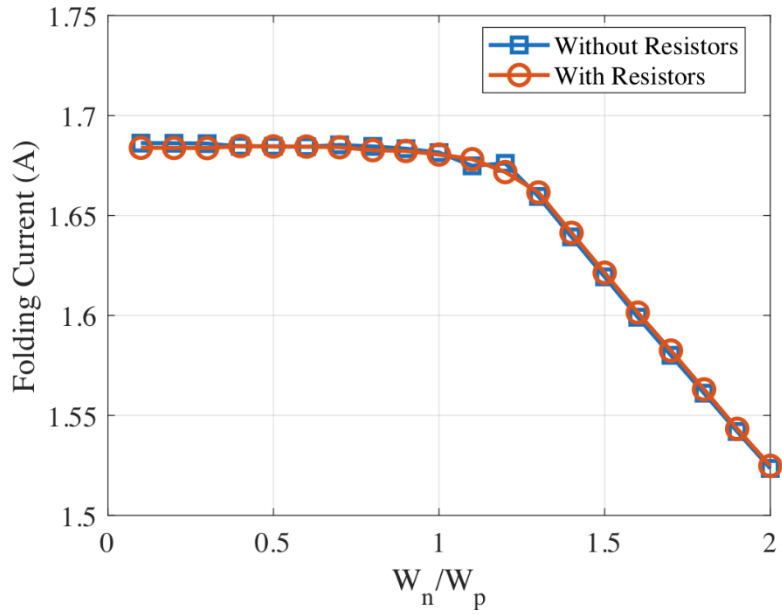
Figure 43: New design folding current with and without R1, R2 – W_{n2}/W_p = 3 and C1 = $0.5C_0$.

Figure 44: New design trigger voltage with and without R1, R2 – W_{n2}/W_p = 3 and C1 = $0.5C_0$.
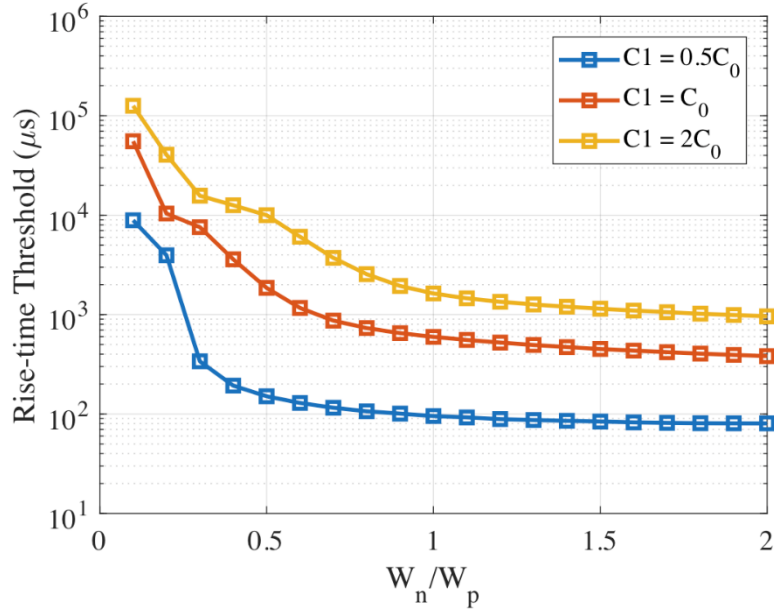
Figure 45: Power-up rise-time threshold of new design ($W_{n2}/W_p = 1$) with R1, R2 versus W_n/W_p with C1 = $0.5C_0$, C_0 and $2C_0$.
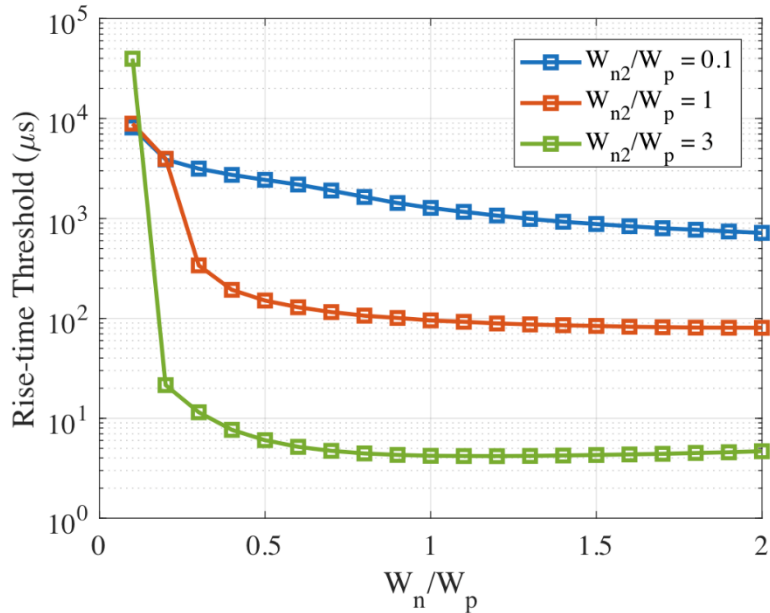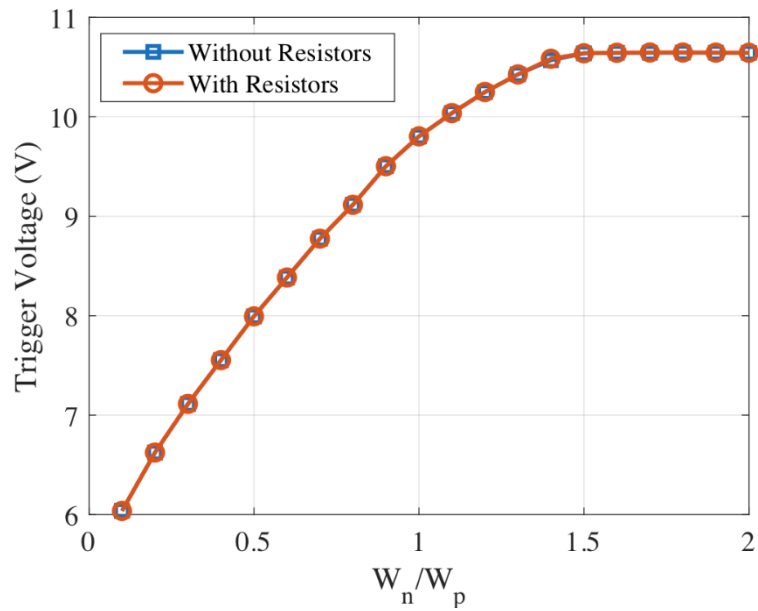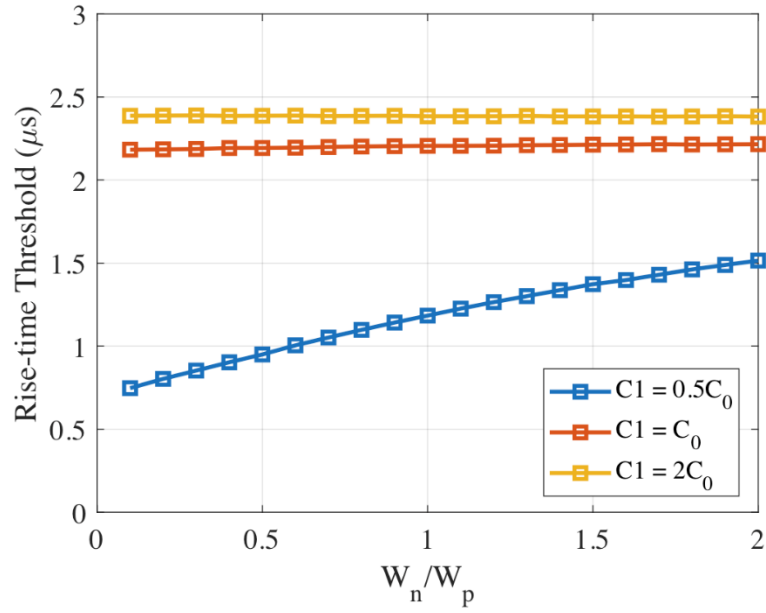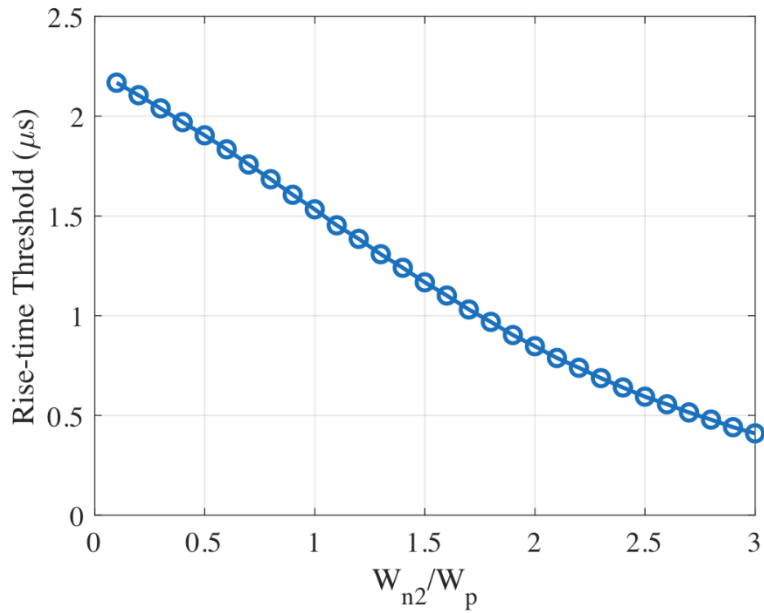

Figure 46: Power-up rise-time threshold of new design with R1, R2 versus W_{n2}/W_p - $W_n/W_p = 3$ and C1 = $0.5C_0$.

CHAPTER 10: MEASUREMENT RESULTS

To verify the performance of the clamps, the two designs were fabricated in 28 nm CMOS and tested with DC, TLP, power-on TLP and power-up voltage ramp measurements. The original design has a C1 capacitance of C_0 and $W_n/W_p = 1$ as a compromise between ESD performance and mis-trigger immunity. The new design has a C1 capacitance of $0.5C_0$, $W_n/W_p = 1$ and $W_{n2}/W_p = 3$. Other than these parameter differences and the schematic changes in the new design, the two are the same. In particular, the M13 and M14 widths are the same and so is the layout area.

A simple DC sweep of the voltage using a source measurement unit (SMU) verifies that the clamp will operate at the high voltage with no breakdown of the devices. The DC sweep results at 25 °C are shown in Fig. 47. The results confirm that the clamps do not break upon reaching 3.3 V (no sudden increase in leakage current, indicating a failure). It is almost always desirable to have low leakage current, but the maximum tolerable is application dependent. For this design, it was desired to have a leakage current below 1 μA [30]. Finally, the curves of Fig. 47 agree with the expectation that the original and new design would have approximately the same leakage current.

Next, the TLP I-V curves were measured and the results are shown in Fig. 48 with accompanying leakage current measurement after each pulse. In the original design, the I-V curve flattens at approximately 0.8 A. The flattening of the I-V curve due to the mis-trigger protection circuit in the original design leads to a large increase in the clamping voltage which causes a parasitic device, likely the NPN BJTs of M13 and M14, to conduct the ESD current beyond that point. In contrast, the new design, with its optimal biasing, has a much higher folding current, lower clamping voltage and avoids the parasitic operating mode. The snapback-

like behavior in the new design I-V curve is expected, and it is a result of the triggering

mechanism in the original and the new design being the same, but the active feedback being

placed in different locations. Essentially, when the active feedback engages in the original design

it is simply sustaining the initial bias coupled onto the RC detector. In the new design, however,

the active feedback will increase the VG2 and VG1 bias over the initial coupling once triggered

that will shift the I-V curve to a lower voltage. What can also be seen from Fig. 48 is that the

new design not only has improved on-resistance and lower clamping voltage, but the biasing also

gives it a 70% larger failure current normalized to M13 & M14 width ($It_2/\mu m$) shown in Table 1.
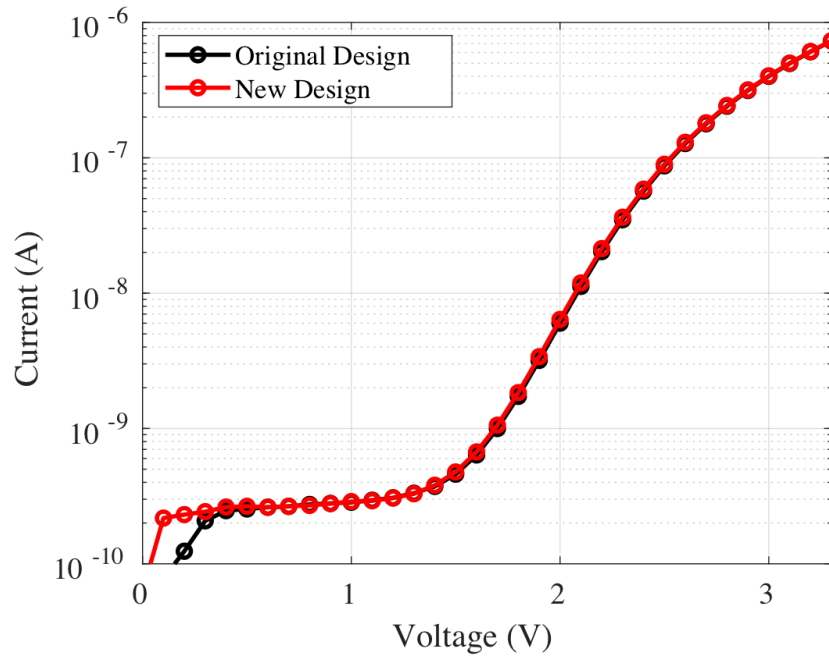
Figure 47: DC leakage measurement of original and new designs at 25 °C.

To see whether the same advantages are present during faster discharges, 2.5 ns pulse-

width 100 ps rise-time vf-TLP testing was done (Fig. 49). The results confirm that the new

design vf-TLP I-V curve has a lower clamping voltage than the original design. The same

parasitic device turns on in both designs once the clamping voltage reaches approximately 8 V.

A transient plot of the clamp voltage during a vf-TLP test is shown in Fig. 50. The initial

"overshoot" voltage is the same in both designs as the triggering mechanism is the same, but after approximately 1 ns the active feedback in the new design brings it to a lower voltage. A plot of the current injection versus the overshoot voltage is shown in Fig. 51 to confirm over the entire vf-TLP I-V range that the two are the same. Thus, the new design has greatly improved ESD performance both in the HBM and CDM time domains as shown through TLP and vf-TLP testing.

	It$_2$ (mA/μm)	R$_{on}$ (Ωcm)
Original Design	0.69	354
New Design	1.18	331

Table 1: Failure current and on-resistance normalized to MOSFET width from TLP I-V (Fig. 48).
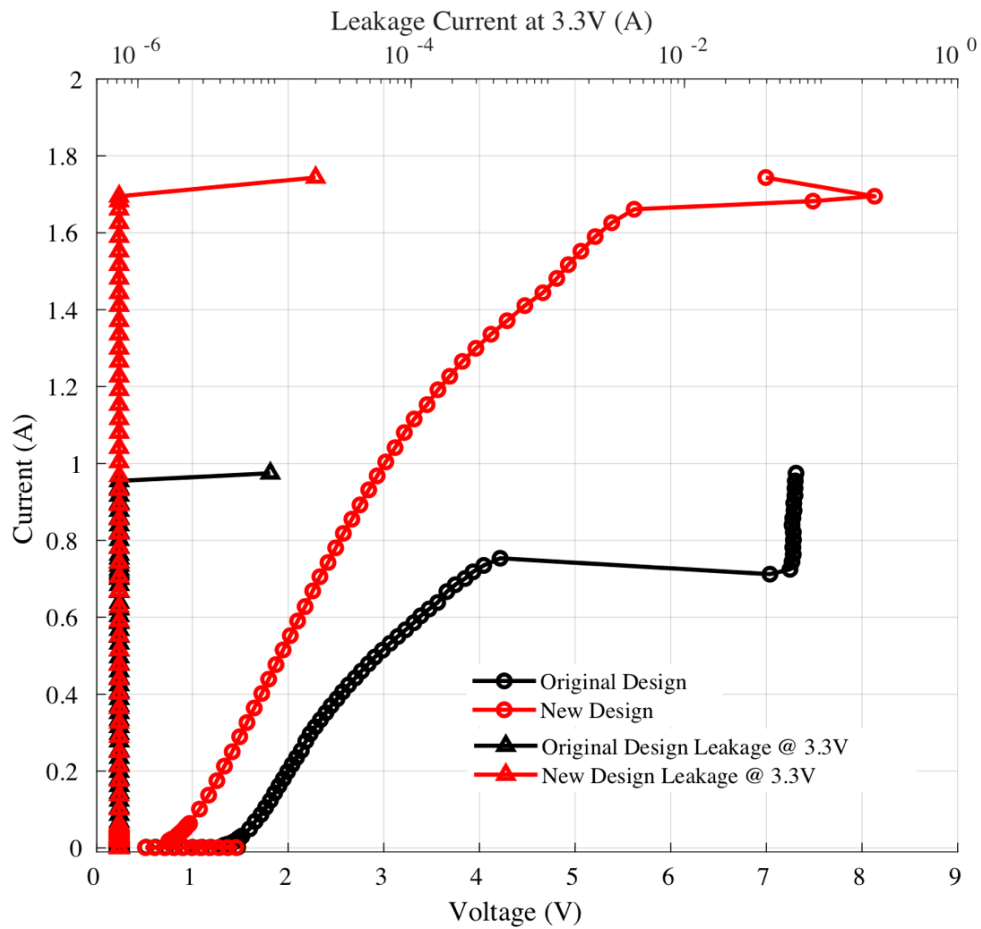

Figure 48: TLP I-V with leakage measurement at 3.3 V - 100 ns pulse-width, 10 ns rise-time.
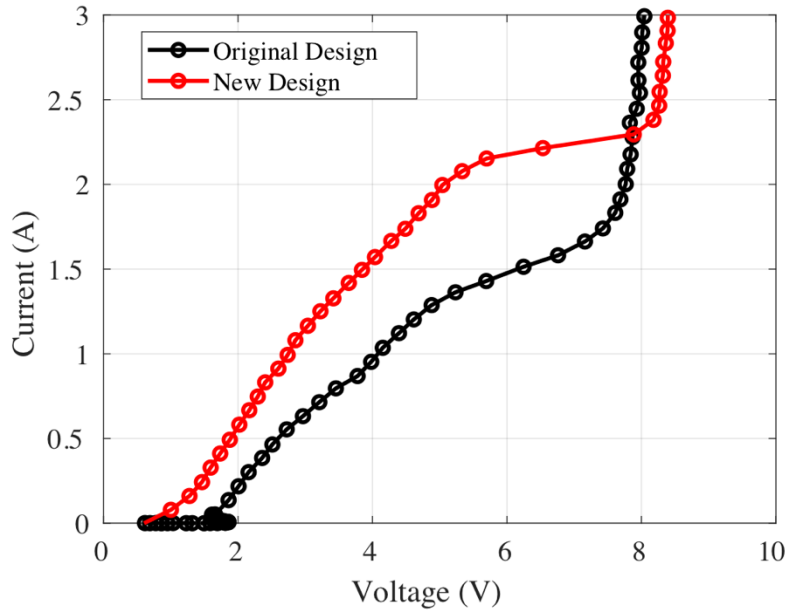
Figure 49: vf-TLP I-V - 2.5 ns pulse width, 100 ps rise-time.
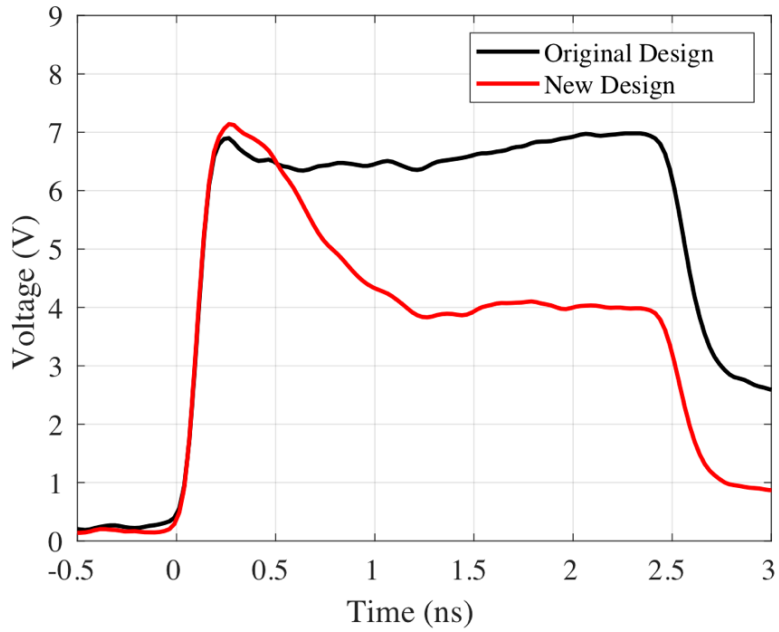
Figure 50: vf-TLP voltage transient for a 1.6 A current injection – 100 ps rise-time, 2.5 ns pulse width.
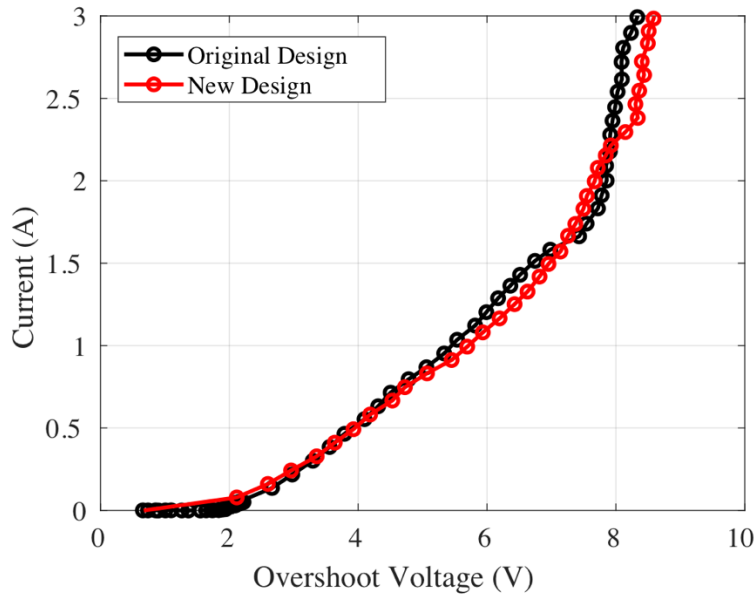
Figure 51: vf-TLP current versus overshoot voltage - 2.5 ns pulse width, 100 ps rise-time.

There is one final note on the TLP testing. Upon initially measuring the I-V of the clamps, a discrepancy was noticed in the trigger voltage when measurement was done with and without a leakage test. It was hypothesized that the reason for this was node MIS was charged during the leakage test and not given sufficient time to discharge before the next pulse was applied. This makes the clamp more difficult to trigger, and thus an increase in the trigger voltage is seen. The simplest way of eliminating this problem is to increase the time between TLP pulses to allow node MIS to fully discharge, and when done this eliminated the increase in trigger voltage (Fig. 52). The takeaway from this is to emphasize the importance of guaranteeing that there is no "memory" from the previous TLP pulses that influence the subsequent measurement.

To verify the clamps will remain on during a long HBM discharge, the HPPI-3010C TLP waveform generator was used to force a pseudo-HBM current. A waveform of the current and voltage for a >2 kV pseudo-HBM discharge is shown in Figs. 53 and 54. A 2x version of the

original and new designs was used in this measurement to replicate a typical scenario at an I/O

pad, where at least two would be used. Both designs are capable of handling the discharge, but

the new design has the added advantage of lowered clamping voltage as was seen in TLP testing.
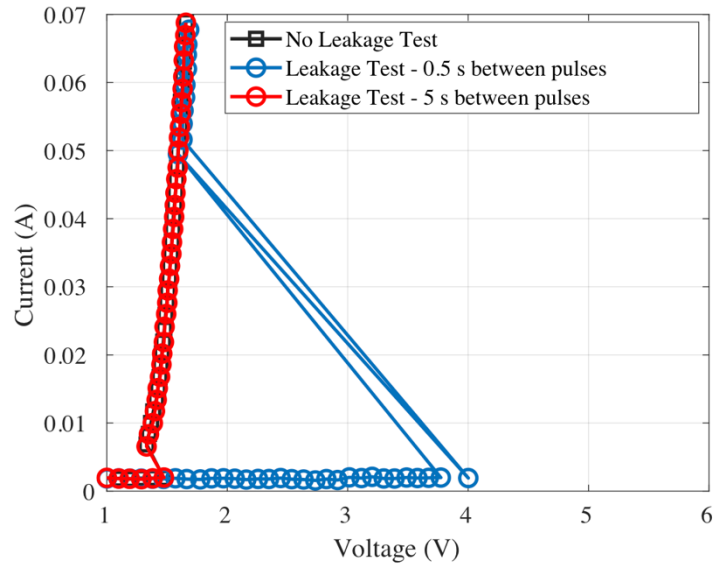
Figure 52: Original design TLP I-V without leakage test and with leakage test.
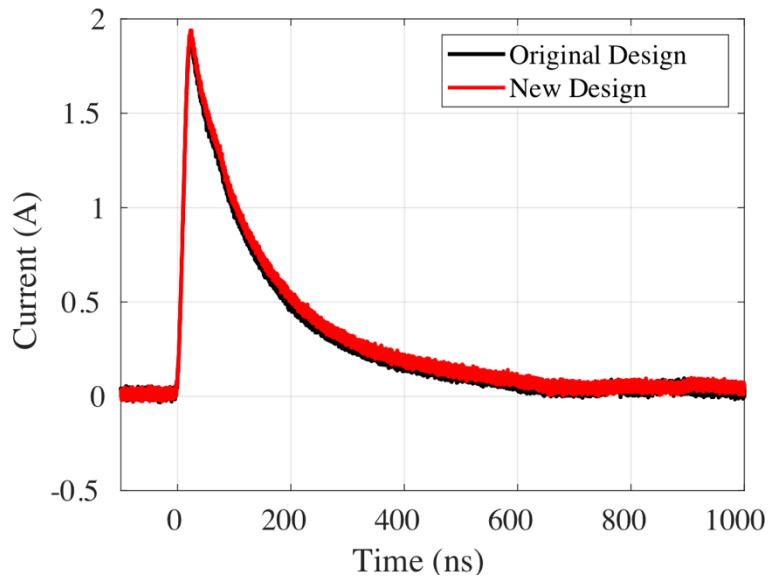
Figure 53: Pseudo-HBM current transient injection for 2x original design and 2x new design.
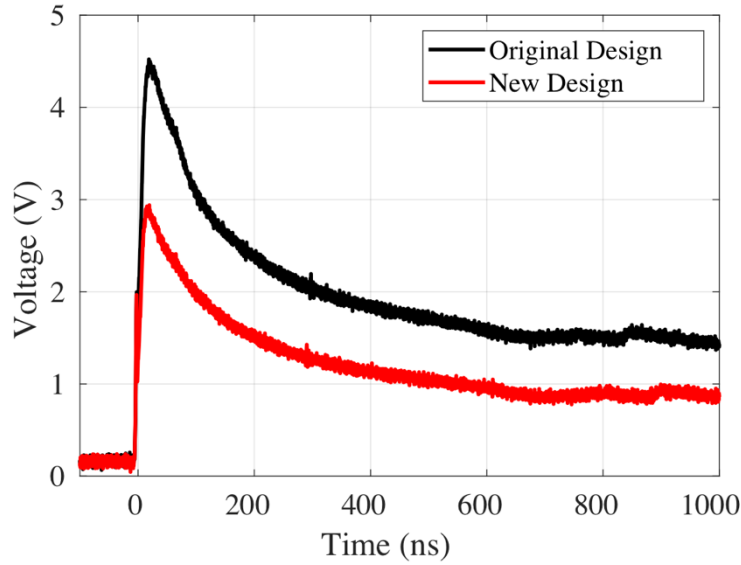
Figure 54: Pseudo-HBM voltage transient for 2x original design and 2x new design.

To determine the mis-trigger immunity due to Vdd-noise, the power-on TLP setup from the previous chapter 9 (Fig. 32) is used, with more details on the measurement in [33]. The power-on TLP I-V curves with 3.3 V DC bias are shown in Fig. 55. Neither clamp conducts below 8 V, demonstrating that both designs are resilient to well over 1x Vdd in noise.
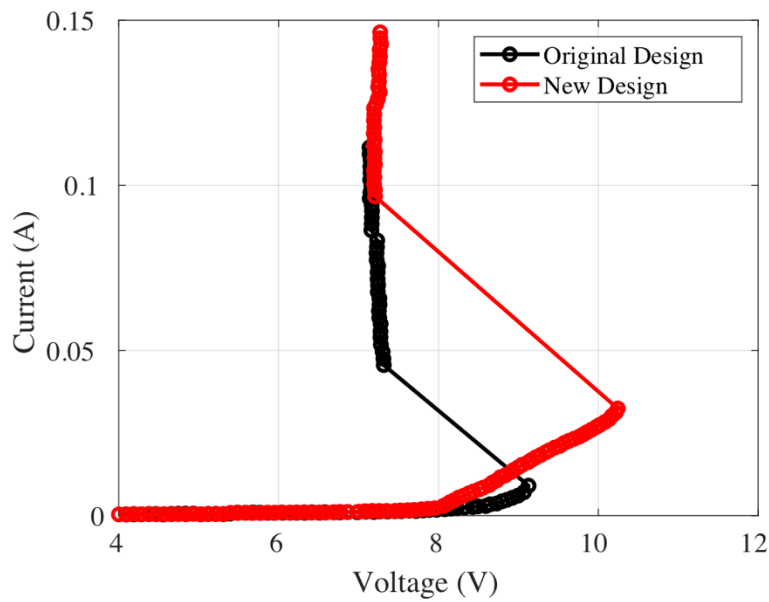
Figure 55: Power-on TLP I-V for original design and new design. 100 ns pulse-width, 1 ns rise-time.

Finally, the power-up rise-time threshold was measured using the schematic in Fig. 38, with the voltage ramp being supplied by a pulse generator with a variable rise-time. More information on the measurement setup is given in [34]. To extract the rise-time threshold, the clamp is given a 3.3 V voltage ramp with a steadily decreasing rise-time in increments of 50 ns until a sustained current is seen flowing through the clamp. Examples of a successful voltage ramp and a mis-trigger for the original design are shown in Figs. 56 and 57 along with simulations of the schematics with parasitic R and C extracted from layout. Good correspondence is seen between the measurement and simulation.

The rise-time threshold is dependent on temperature and process variation, and this was captured in measurement by varying the temperature with a hot chuck and measuring 20 different clamp test structures for the original and new designs. Additionally, simulations at the process corners were done to compare the match of measurement and simulation. These results are shown in Table 2 (original design) and Table 3 (new design).

The minimum rise-time threshold in simulation is associated with fast devices and the maximum slow devices. This intuitively makes sense based on the discussion of the mis-triggering mechanism in chapters 8 and 9. The clamp mis-triggers when node VG1 rises above the inverter threshold voltage, triggering the positive feedback and causing the clamp to conduct. M8 in the original design and M8, M25 in the new design are responsible for shunting that node to Vss and are more effective if the devices are skewed fast because they conduct more current. The opposite is true for slow devices. It should be noted that the process corners are highly pessimistic or optimistic estimates. The minimum rise-time threshold simulation is significantly faster than seen in measurement, and the maximum is significantly slower for both designs.

The new design has a much greater sensitivity to temperature than the original, with the rise-time threshold decreasing significantly with increasing temperature. This is again due to the removal of the resistors on node VG1. In the original design, the resistors are made of P+ poly and have a moderate positive temperature coefficient. The resistors will shunt any voltage that develops on node VG1 with only a moderate temperature dependency. This is in contrast to the new design, where the shunting is done purely by M25. The transistor leakage current increases exponentially with temperature when it is in subthreshold during the beginning stage of the voltage ramp, where several hundreds of mV develop on VG1 for relatively small currents. This is significant enough to cause a mis-trigger event. This also explains the dramatic decrease in the rise-time threshold at 110 °C, because the shunting capabilities are much improved with the temperature increase. The rise-time threshold can not only be improved by adding the resistors back into the new design but can also be made more stable over the operating temperature, with basically no effect on ESD performance or power-on mis-trigger immunity. Nonetheless, even without the resistors the new design has a worst-case rise-time threshold of 1.5 μs, which is only slightly larger than 1 μs in the original design.

(a)

(b)
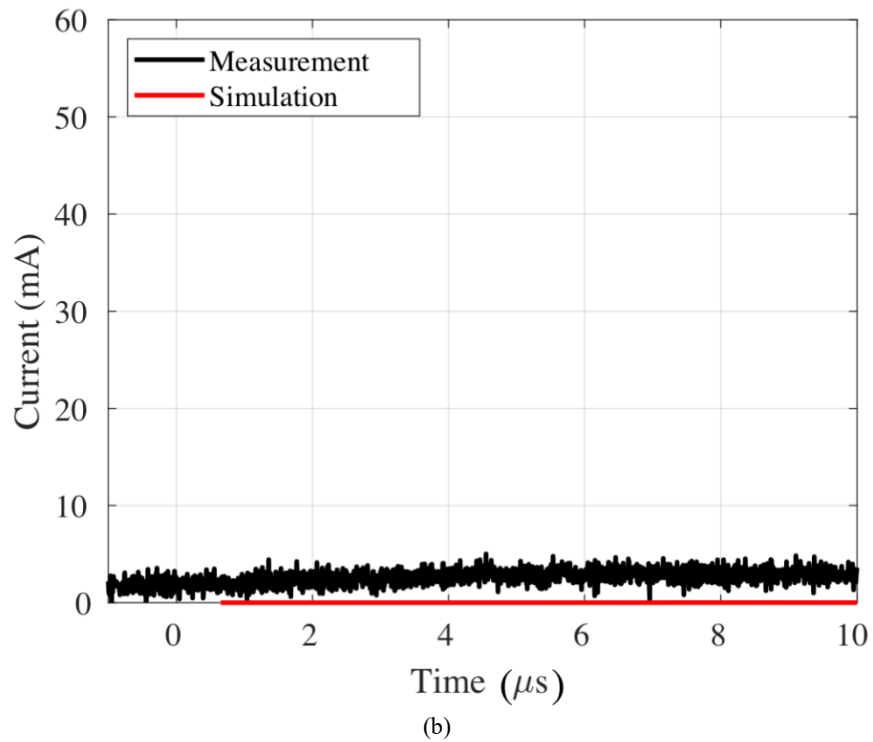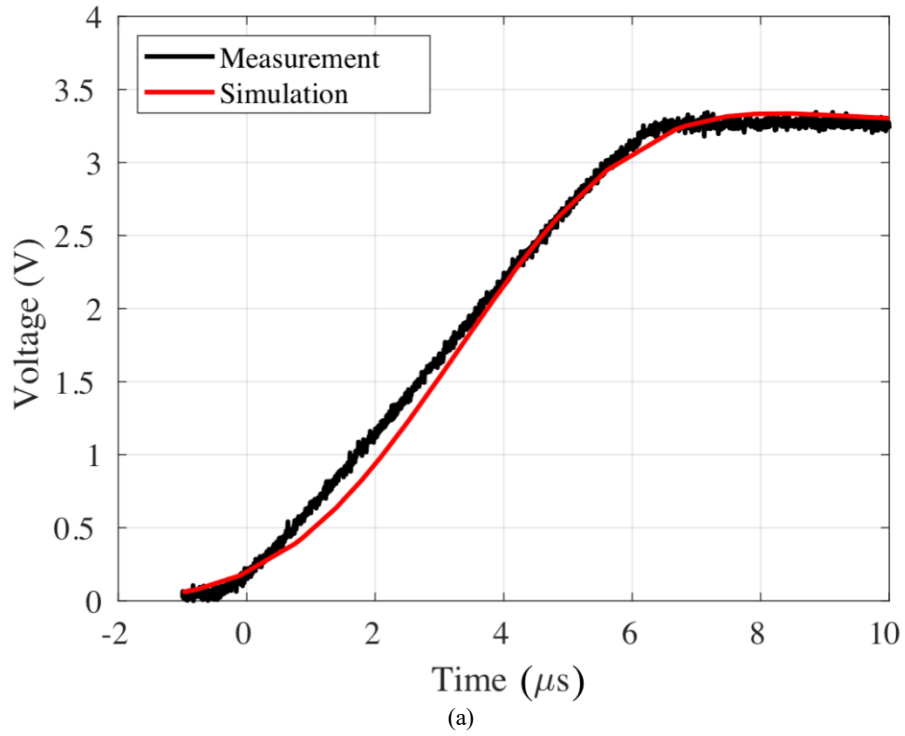
Figure 56: Original design 5 μs voltage ramp. (a) DUT voltage measurement and simulation. (b) DUT current measurement and simulation.

(a)

(b)
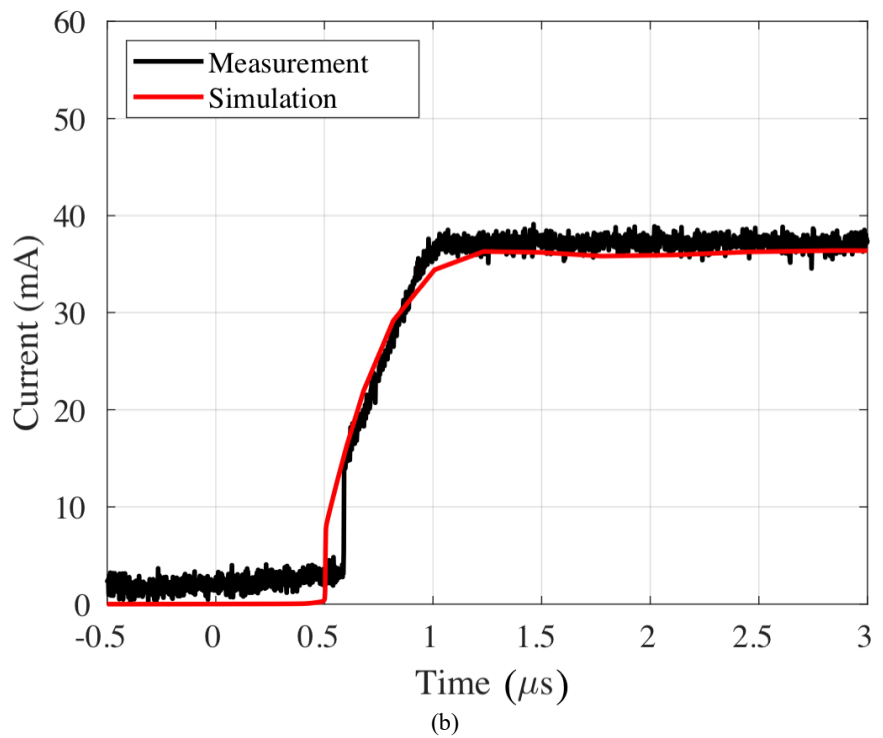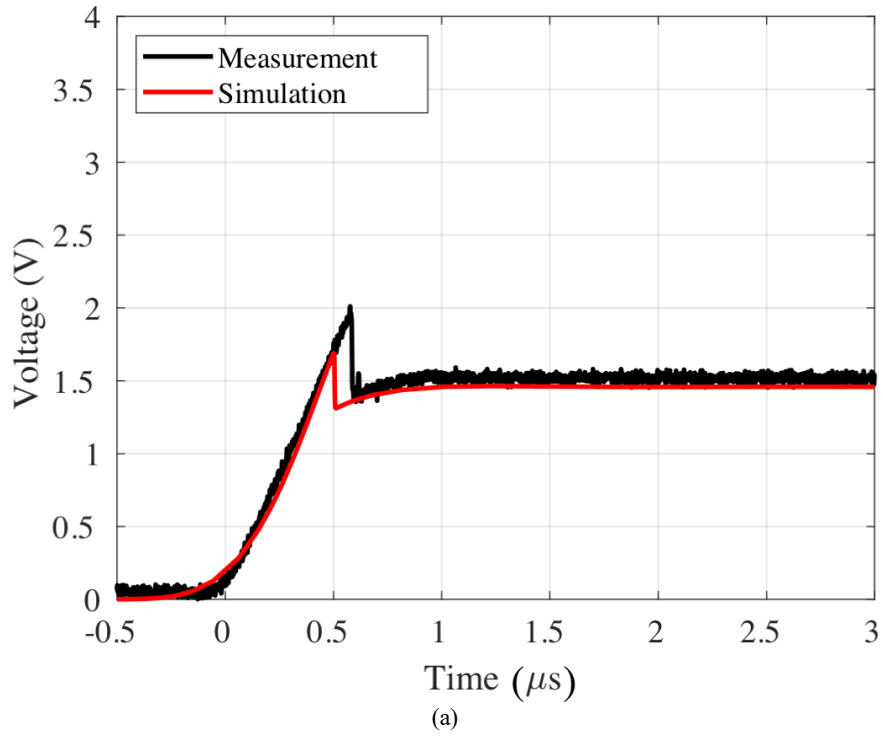
Figure 57: Original design 850 ns voltage ramp. (a) DUT voltage measurement and simulation. (b) DUT current measurement and simulation.

	Average Rise-time Threshold (25 °C | 110 °C)		Minimum Rise-time Threshold (25 °C | 110 °C)		Maximum Rise-time Threshold (25 °C | 110 °C)	
Measurement	850 ns	800 ns	700 ns	700 ns	1 μs	950 ns
Simulation	850 ns	500 ns	50 ns	50 ns	1.8 μs	2.2 μs

Table 2: Original design rise-time threshold at 25 °C and 110 °C extracted for 20 devices. Simulation average corresponds to a typical device, minimum a fast device skew and maximum a slow device skew.

	Average Rise-time Threshold (25 °C | 110 °C)		Minimum Rise-time Threshold (25 °C | 110 °C)		Maximum Rise-time Threshold (25 °C | 110 °C)	
Measurement	1.3 μs	300 ns	1.1 μs	300 ns	1.5 μs	350 ns
Simulation	700 ns	200 ns	100 ns	350 ns	30 μs	1.6 μs

Table 3: New design rise-time threshold at 25 °C and 110 °C extracted for 20 devices. Simulation average corresponds to a typical device, minimum a fast device skew and maximum a slow device skew.

CHAPTER 11: CONCLUSION

ESD performance and mis-trigger immunity in an active feedback rail clamp were improved by carefully decoupling the positive and negative feedback loops during ESD operation. Additionally, it was reiterated that optimally biasing the cascoded transistors in a high-voltage tolerant design significantly reduces the clamping voltage, thereby improving ESD performance. The design shows a 70% improvement in $It_2/\mu m$ and significant reduction in R_{on}.

APPENDIX A: TLP BACKGROUND

Transmission line pulsing, or TLP, has been the standard method of characterizing ESD devices and components since its introduction [35]. TLP is used to replicate the approximate magnitude, rise-time and duration of the current transient during an ESD event. To understand the need for TLP, consider the simplified 50 Ω pulse generator schematic shown in Fig. 58. ESD components have a very low impedance, and therefore the current injected into the DUT is well approximated with the short circuit current.

$$I_{DUT} \approx I_{sc} = {V_{max}}/{50}. \tag{7}$$
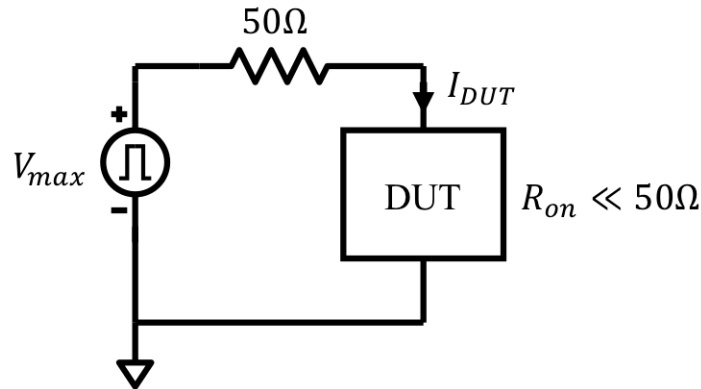
Figure 58: Simplified pulse generator schematic.

Using (7), a 10 A current injection would require a maximum pulse voltage, V_{max}, of 500 V. Most semiconductor-based pulse generators capable of handling those voltages have slow rise-times, and vice versa. This is the niche that TLP fills, as it can generate both high power and fast rise-time pulses.

A basic schematic of a TLP tester is shown in Fig. 59 (a)-(c). The tester consists of a high-voltage source in series with a large resistance, two transmission line cables and a switch. The transmission line on the left, often called the charged cable, is first charged to a fixed pre-

charge voltage, V_{pre}. Once the voltage has been established on the charged cable, the switch is flipped, connecting the two transmission lines together. Subsequently, a forward-going $V_{pre}/2$ pulse propagates into the right-most transmission line, and a reverse going $-V_{pre}/2$ pulse into the charged cable. These pulses are generated to satisfy the new boundary condition once the switch has flipped; i.e., the charged cable looks into a 50 Ω impedance rather than an open.

The forward-going pulse will hit the DUT and inject approximately $V_{pre}/50$ Amps of current into it assuming its impedance is small. The reverse-going pulse will hit the open end of the charge cable and fully reflect back, eliminating the remaining voltage on the charged cable and eventually cancelling out the voltage applied at the DUT (Fig. 41(c)). This whole process produces a pulse that lasts for approximately as long as it takes the reverse-going pulse to traverse the charged cable twice. Mathematically, this can be expressed as

$$Pulse\ Width = 2 * {}^{l}/_{v},\tag{8}$$

where l is the length of the charged cable and v is the propagation velocity of the transmission line. Before the reflection from the charged cable eliminates the pulse, the DUT sees the same Thevenin equivalent circuit shown in Fig. 58 but with $V_{max} = V_{pre}$.

Measuring the average voltage and current across the DUT after the pulse settles, an I-V curve of the DUT can be generated by injecting successively increasing currents until the DUT fails. See Simburger et al. [11] for more information on the different ways of measuring the voltage and current.
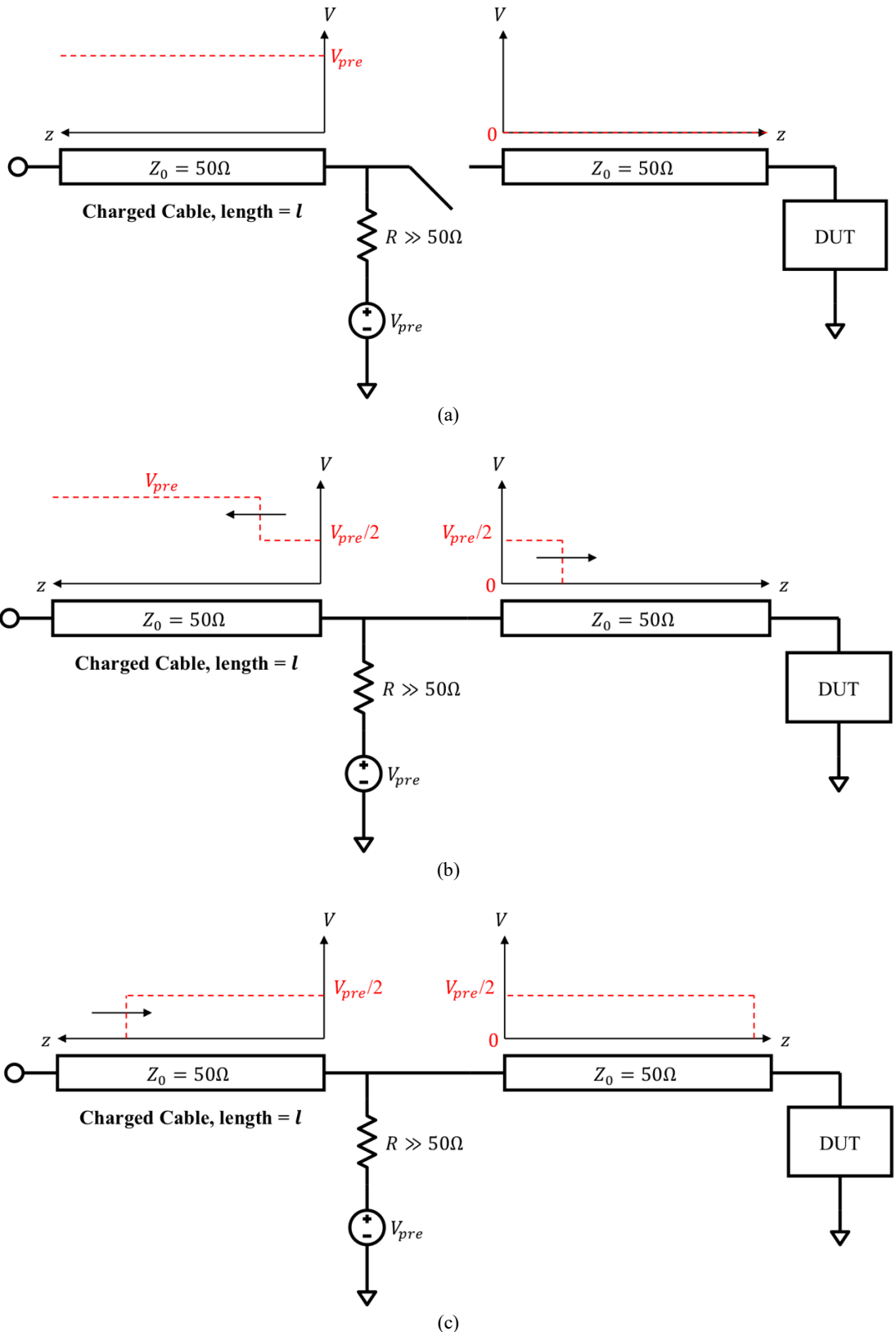
Figure 59 (a)-(c): Simplified TLP tester schematic (a) before the switch is closed, (b) a short time after the switch is closed, (c) and after the reverse going pulse is reflected at the open end of the charge cable.

APPENDIX B: TDR VS. TDT

There are two methods that can be used to measure the transient current during reverse recovery with the setup from chapter 3, time domain reflectometry (TDR) and time domain through (TDT). TDT has already been explained in detail, but to understand TDR, consider the simplified schematic in Fig. 60.
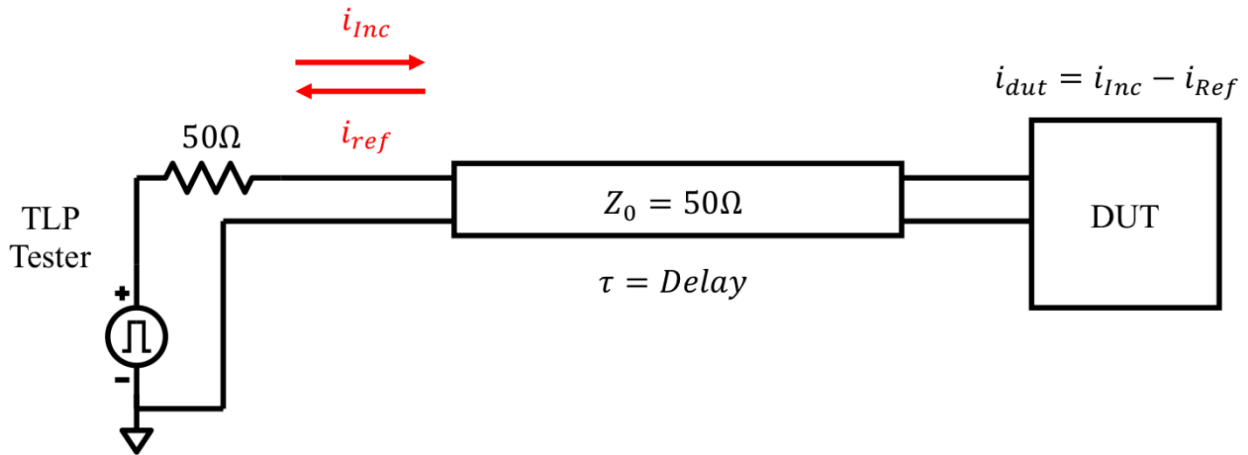
Figure 60: TDR measurement setup schematic.

For high-speed measurements, it is difficult to measure the current at the DUT because a sensor cannot be placed directly at its terminals. Suppose a current sensor can be placed no closer than 0.1 m (10 cm) to the DUT. This is a reasonable value given the physical limitations of a typical magnetic current sensor; i.e., their size does not allow them to be placed any closer to the wafer. A coaxial cable with a Teflon dielectric has a delay of 5 ns/m, and therefore the delay associated with the 0.1 meters between the sensor and the DUT is 0.5 ns. For a slowly varying transient relative to this delay, there is no visible discrepancy and the current measured at the sensor will be the current at the DUT. Consider, however, a 100 ps rise-time pulse with a pulse width of 1 ns. Replace the DUT in Fig. 60 with a 100 Ω resistor and simulate the current 0.1 m away from the load on the transmission line in SPICE. The current simulated away from the DUT and at the DUT is compared in Fig. 61.
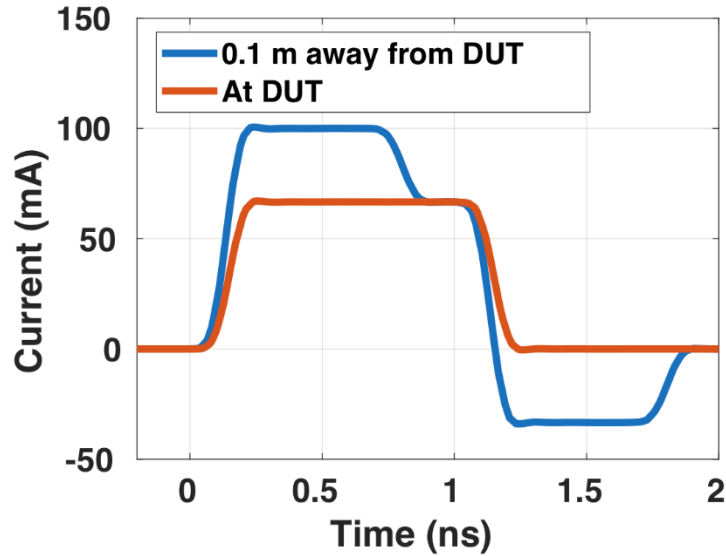
Figure 61: 10V, 100 ps rise-time, 1 ns pulse-width into a 100 Ω load. Current is simulated 0.1 m away from the DUT on the transmission line and at the DUT.

The current measured away from the DUT is distorted because the incident and reflected waveforms from the DUT are not perfectly aligned in time with each other like they are directly at the DUT. There exists a small time period around t = 1 ns where the current measurement is correct, but it is very brief and only occurs because the waveform is a pulse and the transients have settled. All of the information at the rising edge, in contrast, is lost in this measurement. The previous analysis also assumes the current sensor has the bandwidth to measure a 100 ps rise-time. A rough estimate of the bandwidth needed to measure a particular rise-time is given by the following equation:

$$BW \times t_r = 0.35, \tag{9}$$

where BW is the bandwidth and t_r is the rise-time [36]. Note that (9) is technically only valid for a single-pole RC or Gaussian frequency response, but it is nonetheless a reasonable approximation. The bandwidth required to detect a 100 ps rise-time is ∼ 3.5 GHz. The current sensor available in the HPPI-3010C TLP tester used in this work only has a bandwidth of

approximately 2 GHz, so it cannot be used to measure that transient. TDR measurements are used to bypass those problems.

In Fig. 60, the boundary condition at the load requires that the current at the DUT be equal to the incident current waveform minus the reflected, or

$$i_{dut} = i_{inc} - i_{ref}.$$ (10)

The incident waveform is the forward-travelling signal from the source that hits the DUT, whereas the reflected waveform is the reverse travelling signal from the DUT that travels back to the source. If twice the delay of the transmission line, 2τ, is greater than the duration of the applied pulse, the incident and reflected waveforms measured at the source will be completely separate from each other. Therefore, these two signals can be measured independently using a high-bandwidth power splitter at the source, and then combined computationally during post-processing to determine the current at the DUT.

$$i_{dut} = i_{inc} - i_{ref} = \frac{v_{inc}}{Z_0} - \frac{v_{ref}}{Z_0}$$ (11)

So, rather than attempting to place a sensor close to the DUT, TDR places it **farther** from the DUT, which is much easier to do. The current and voltage on the transmission line in (11) are proportional by the characteristic impedance, Z_0. Note that it is also possible to measure the voltage using TDR with the equation

$$v_{dut} = v_{inc} + v_{ref}.$$ (12)

Particularly important to TDR is properly calibrating the alignment of the incident and reflected waveforms. This is done by placing a known DUT, and therefore a known reflection,

67

and then shifting the reflected signal and subtracting it from the incident according to (11) until the result gives the expected current response at the DUT. The simplest DUT to use is an open, which ideally has a current of zero and a reflection coefficient of 1 for any arbitrary signal. This process is shown graphically in Fig. 62. It is important to remember that doing this shifting and subtracting means that the current measurement is only accurate for the brief time period when the incident and reflected waveforms are aligned, i.e. anything before or after that period is incorrect.

A perfect open, unfortunately, does not exist in reality, and will have some nonzero capacitance. The same alignment calibration is done with a 100 fF capacitor as the DUT rather than a perfect open. The DUT now exhibits some current at the rising edge of the pulse due to the capacitance. The key takeaway is that it is now incorrect to try and align the waveforms such that there is zero current at the DUT, as was done for the perfect open, but rather it should be aligned to replicate the response shown in blue curve of Fig. 63. If there is a slight error in the alignment, it will manifest as error in the resulting current measurement. This is shown for two cases in Fig. 63, one where the alignment mismatch is +10 ps and the other -10 ps. Given that a 40 Gigasample/second oscilloscope will have a sampling time of 25 ps, it is not unreasonable to expect an offset of this magnitude if the waveforms are not aligned perfectly.

The problem during calibration is that if the exact response of the DUT is not known a priori, as is typically the case, it is impossible to determine which of the three traces in Fig. 63 is correct. This can also be seen in measurement of an open structure shown in Fig. 64. It should be emphasized that the error introduced by the offset is negligible after the rising edge of the pulse, and therefore this is not a great concern if the purpose of the current measurement is to get an I-V curve from vf-TLP. It is a major problem if a high-fidelity transient waveform of the current is

desired, as is the case for a reverse recovery measurement. The three different calibrations from Fig. 65 are used to measure the current of a diode with TDR using the measurement setup from chapter 3, and then is compared with the TDT current measurement. These results are shown in Fig. 65.

Just as was the case for the open calibration, the alignment mismatch manifests itself as error in the reverse recovery current measurement. The TDR measurement has a negative current value after the initial peak due to offset error, reflection losses and the oscilloscope ranging that can be calibrated. The TDT has a slight increase in current after the initial peak due to the small inductance of the test structure. The main distinction between TDR and TDT is that TDT does not require a careful alignment calibration and therefore will always give the exact current at the DUT, assuming there is sufficient bandwidth and sampling rate at the oscilloscope. The TDT connection can also be made as small as possible, whereas the TDR transmission line must be sufficiently long to separate the incident and reflected waveforms. A longer transmission line has more loss, which introduces more error.

A final advantage of TDT over TDR is lower noise. A histogram of the peak current for the same measurement in Fig. 65 repeated 100 times is shown in Fig. 66 for TDR (best alignment) and Fig. 67 for TDT. The results show that the TDR measurement has quadruple the standard deviation in its peak current relative to TDT. There are two reasons for this, the first being that the incident and reflected waveforms are measured independently of each other for TDR, and therefore each has noise independent of each other. So, when the two are subtracted to determine the current at the DUT, the noise is effectively doubled. The second reason is the oscilloscope voltage range needed to capture the incident and reflected waveforms is much larger than the actual DUT signal after post-processing. This means a larger mV/div is required to

measure it with TDR than if it were measured directly with TDT, where voltage range only depends on the actual signal at the DUT. The internal noise associated with an oscilloscope trace, as specified by the manufacturer, increases as the voltage divisions increase. The increased noise is much more pronounced on the relatively small TDR signal and contributes to the larger peak current spread.
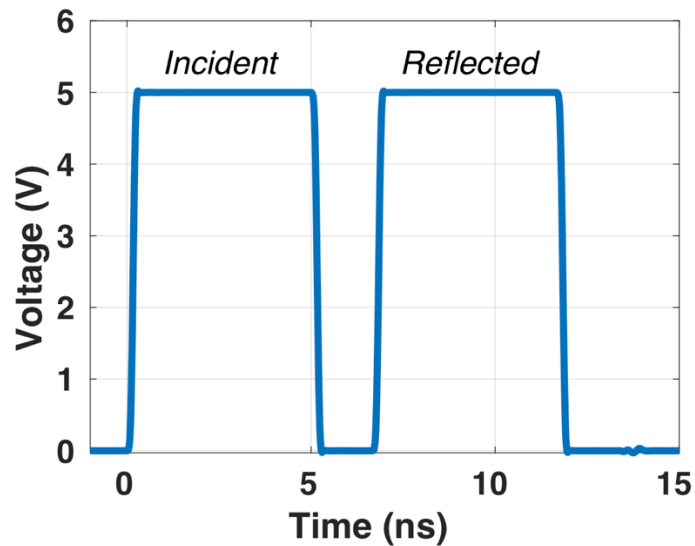
(a)
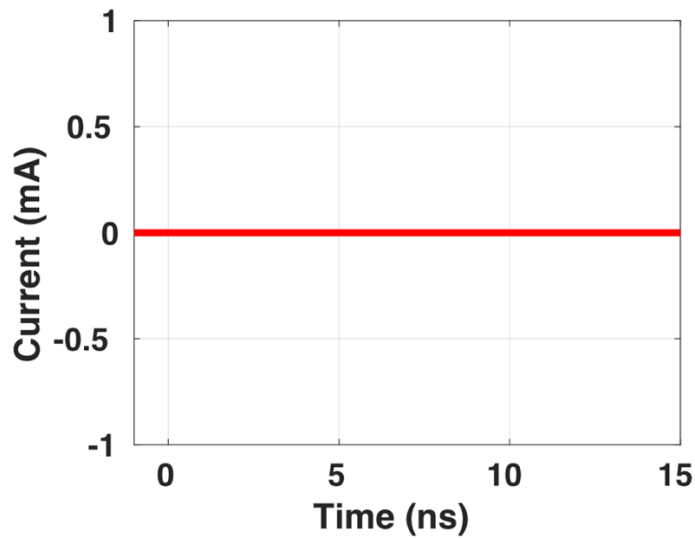
(b)

Figure 62: TDR simulation of (a) voltage at the source and (b) current at the DUT after alignment of the incident and reflected waveforms for an open calibration.
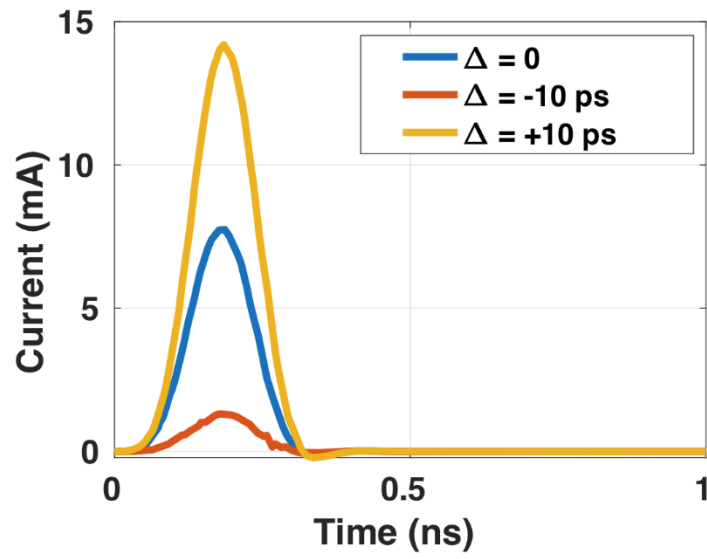
Figure 63: TDR simulated calibration for a 100 fF DUT. The Δ represents an offset in the alignment of the incident and reflected waveforms.
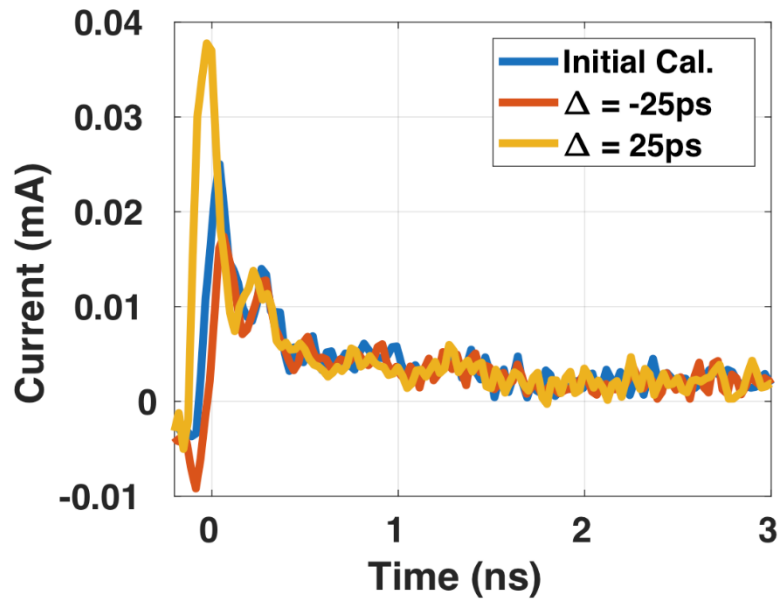
Figure 64: TDR open calibration measurement.
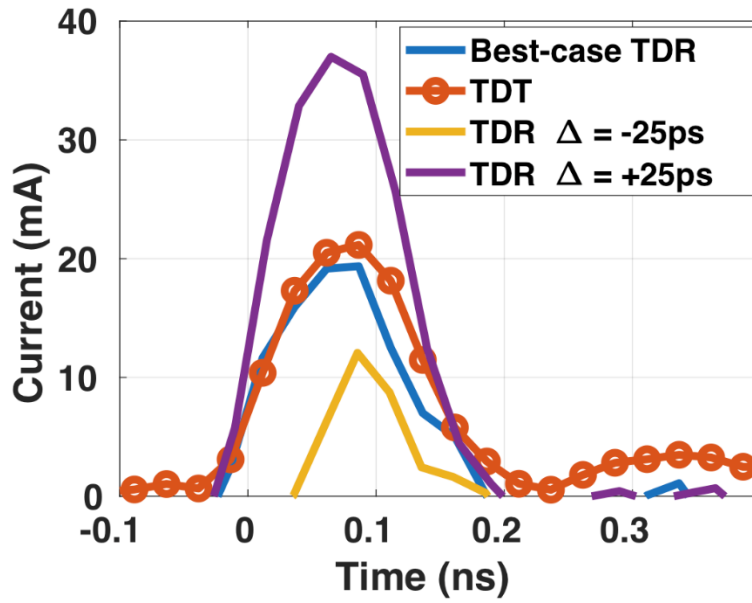
Figure 65: Reverse recovery current measurement of 65 nm P-well diode with TDR and TDT. The initial bias was set to zero for simplicity. Reverse pulse has a 100 ps rise-time and a maximum value of 8 V.

Figure 66: TDR peak current histogram of reverse recovery measurement for 100 samples.
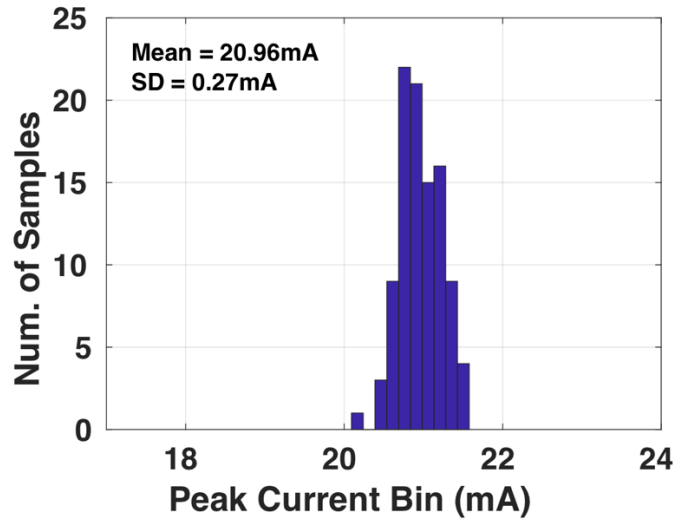
Figure 67: TDT peak current histogram of reverse recovery measurement for 100 samples.

The foregoing discussion proves that a TDT measurement will produce more accurate measurements of the transient DUT current in comparison with TDR. It is important to note that this does not mean TDT should always be preferred over TDR; in fact, there is one key disadvantage. When using TDT, the DUT is connected as a series two-port, so a single Kelvin probe cannot be used to measure the DUT voltage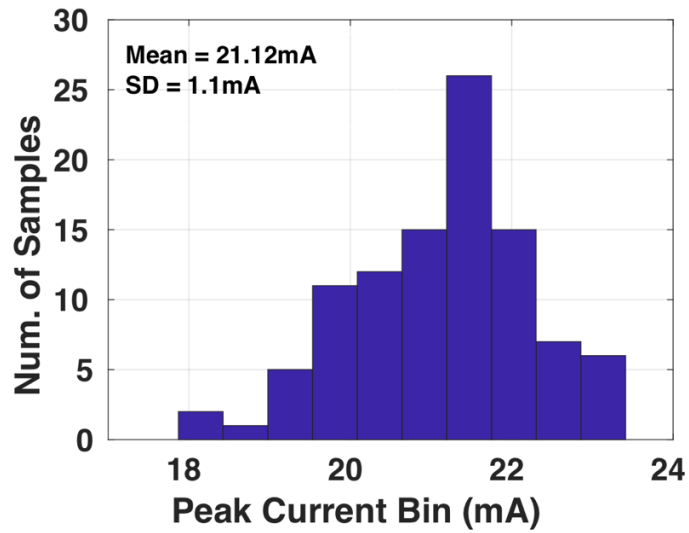 (using a single Kelvin probing requires one of the terminals be connected to ground). Kelvin probing must be done to get an accurate measurement of the voltage transient directly at the DUT, but two would be needed for the TDT setup, one placed at the anode and the other at the cathode. This is impractical to probe on the wafer-level test structures.

The best way of thinking about when to use TDR and when to use TDT is to decide whether a good transient measurement of the voltage or current is needed. TDR will give a high-accuracy transient measurement of the voltage, and TDT the current. For example, in vf-TLP measurements typically only the value of the current after the transients have settled is desired, which does not require ultra-high fidelity. The voltage, especially the overshoot voltage,

73

however, is an extremely important measurement during vf-TLP that requires Kelvin probing to accurately measure. The opposite is the case during reverse recovery measurement, where the current transient, but not the voltage, must be measured with extreme accuracy.

APPENDIX C: TDR REVERSE RECOVERY DC OFFSET DERIVATION

When measuring current or voltage using TDR with a bias tee, as was done in chapter 3 for reverse recovery, the measurement must be offset by the DC bias. Note that this is only the case when the incident and reflected signals are manually aligned. If the current is sampled when the incident and reflected signals physically overlap in time (TDRO), as is done for 100 ns TLP, this is not necessary. The purpose of this appendix is to prove that this DC offset is valid for any arbitrary load. Before beginning the derivation, the offset can be intuitively interpreted as "counting" the DC voltage and current at the source twice when the incident and reflected waveforms are measured, leading to a redundancy when the two are combined that must be corrected with an offset.

The schematic used for the derivation is shown in Fig. 68. The derivation will be done purely in the time-domain, and therefore does not assume the load is linear [37]. To conduct the analysis, the transmission line is replaced with a time-domain two-port (Fig. 69) derived in [37]. The time domain two-port is equivalent to writing the general solution of the telegrapher's equations in the time domain, and therefore is mathematically complete. The only assumptions made in the analysis are that the transmission line is lossless, and that the inductor and capacitor of the bias tee are large enough to be approximated as current and voltage sources for the duration of the transient. Those assumptions are reasonable as long as the transmission line and bias tee are operated within specifications.

Fig. 70 shows the schematic of the setup immediately before the voltage, $v_{pulse}(t)$, is applied. In accordance with the requirements of TDR it is assumed that the duration of the pulse width, PW, is less than the time it takes for the reflection from the DUT to arrive at the source.

Fig. 71 shows the connections at the source during the time period $0 \le t < PW$ when the incident wave is entering the transmission line. During this period, the reflection has yet to arrive back at the source, and therefore the time-dependent two-port sources retain their DC values. The voltage entering the leftmost transmission line port during this time is the incident pulse, or $v_{inc}(t)$. To simplify the math, the (t) is omitted from all further time-dependent voltages and currents without loss of generality. The incident voltage can be solved by writing Kirchhoff's current law at that node, and then the incident current can be found by solving for the current entering the port.
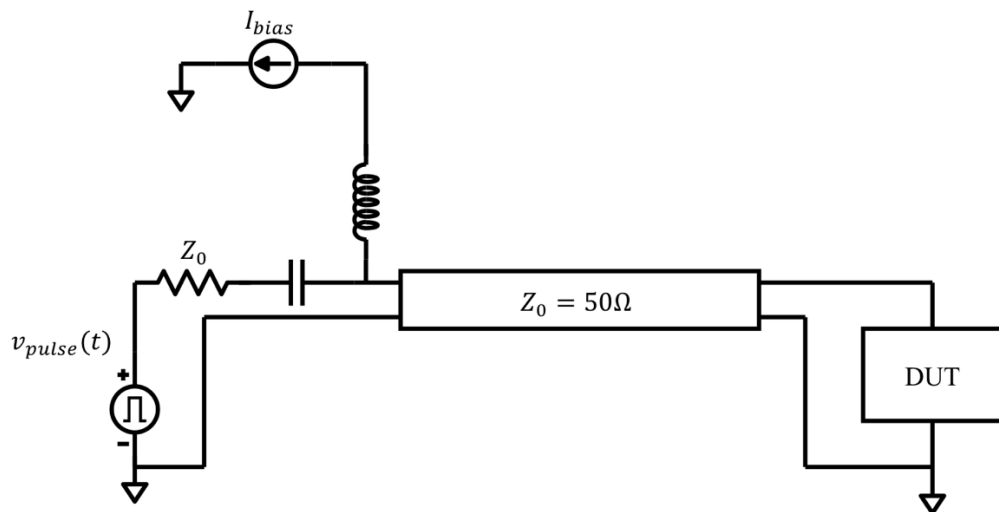
Figure 68: TDR with bias tee schematic.
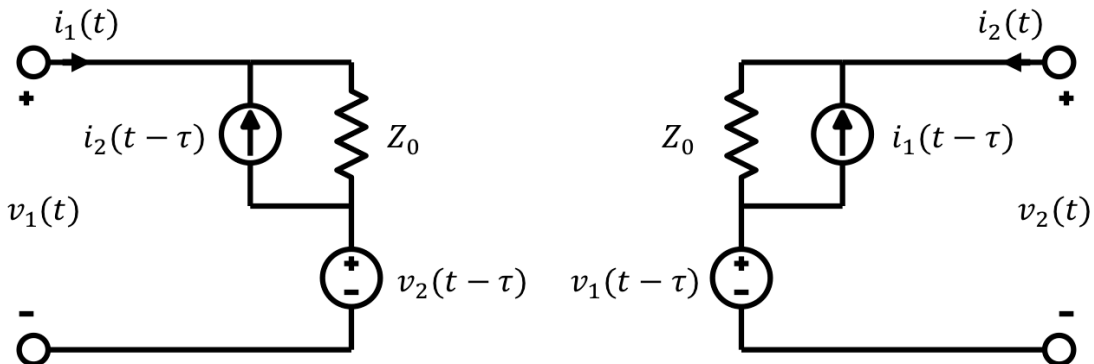
Figure 69: Lossless transmission line time domain two-port, τ is the delay of the line.

76
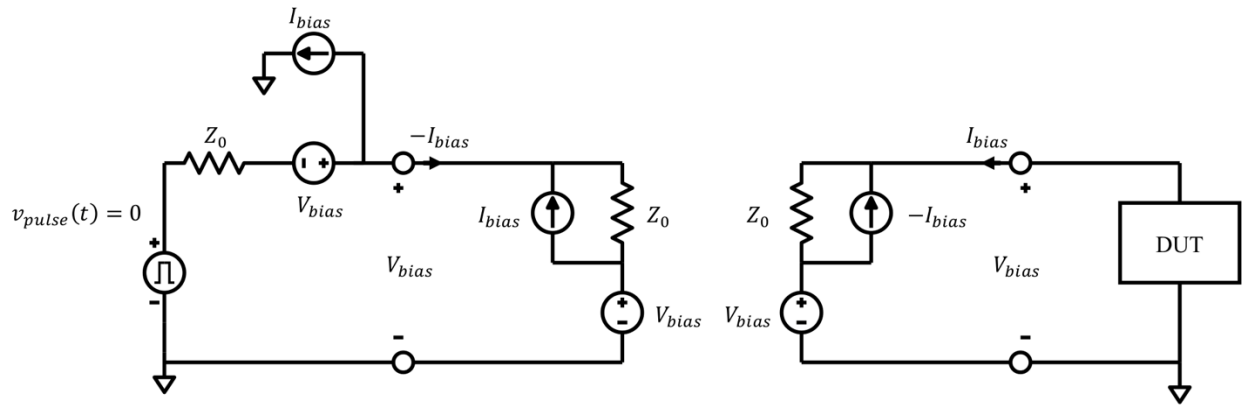
Figure 70: TDR with bias tee schematic where the transmission line has been replaced with its two-port equivalent. The node voltages and currents are valid just before the pulse is applied.
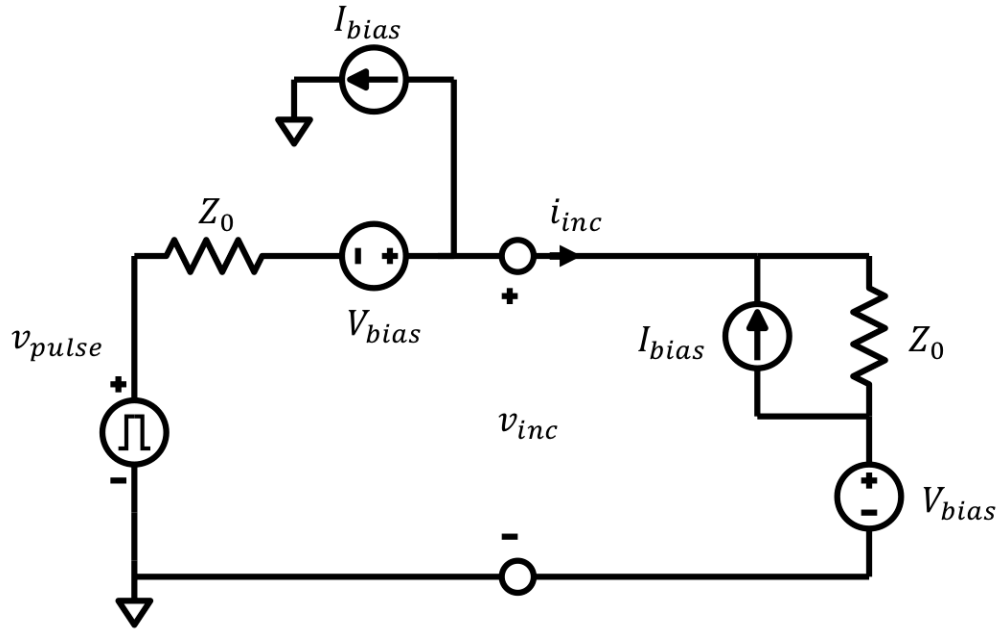
Figure 71: TDR with bias tee schematic at the leftmost port when $0 \leq t < PW$.

$$\frac{v_{inc} - V_{bias}}{Z_0} + \frac{v_{inc} - V_{bias} - v_{pulse}}{Z_0} + I_{bias} - I_{bias} = 0$$

$$v_{inc} = \frac{v_{pulse}}{2} + V_{bias} \tag{13}$$

$$i_{inc} = \frac{v_{pulse}}{2Z_0} - I_{bias}. \tag{14}$$

Note that $v_{inc} \neq Z_0 i_{inc}$, which is due to the voltage and current contributions from the bias tee offsetting them by V_{bias} and $-I_{bias}$ respectively.

The second time period of interest is when the incident pulse hits the DUT, or $\tau \leq t < \tau + PW$, where τ is the delay of the transmission line. A voltage v_{DUT} and current i_{DUT} are forced into the load, and this is shown in Fig. 72. i_{dut} can be written in terms of v_{inc}, i_{inc} and v_{DUT} by writing KCL at the v_{DUT} node.
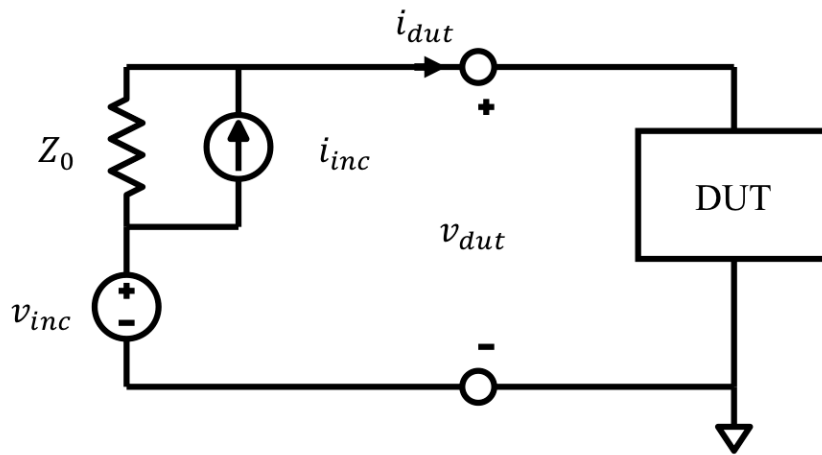
Figure 72: TDR with bias tee schematic at the rightmost port when $\tau \leq t < \tau + PW$.

$$i_{dut} = i_{inc} - \frac{v_{dut} - v_{inc}}{Z_0}. \tag{15}$$

Finally, the voltage at the leftmost port must be solved again when the reflection from the DUT arrives. This occurs during the time period $2\tau < t < 2\tau + PW$, and the schematic of this is shown in Fig. 73. The pulse source is no longer active at this point, so it is represented as a short.

The reflected voltage pulse, v_{ref}, is solved in the same manner as v_{inc}. The results are shown in (16).

$$v_{ref} = \frac{v_{dut} + V_{bias}}{2} - \frac{Z_0(I_{bias} + i_{dut})}{2}. \tag{16}$$

Now that both v_{inc} and v_{ref} are known, and the current that is measured with TDR is given by (17).

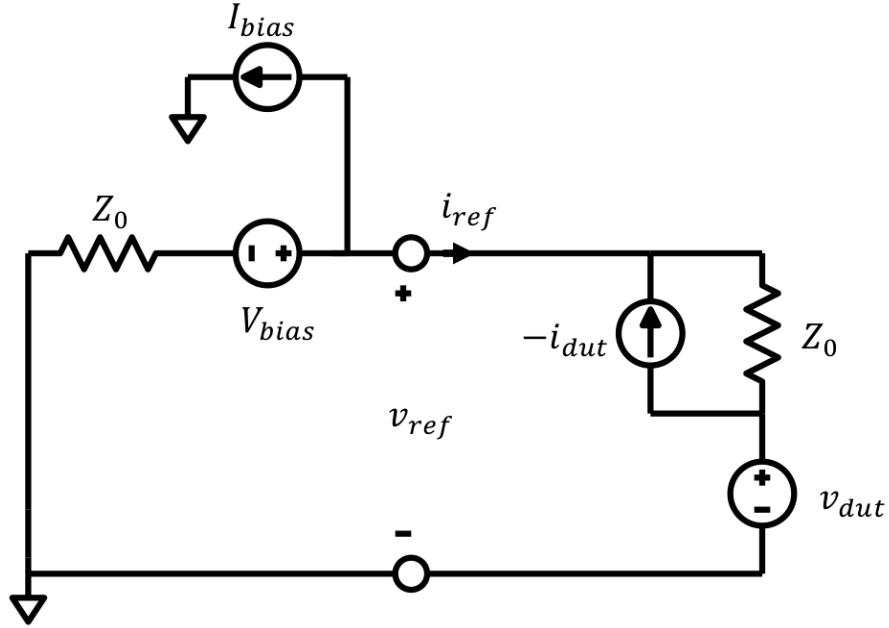Figure 73: TDR with bias tee schematic at the leftmost port when $2\tau \le t < 2\tau + PW$.

$$i_{TDR} = \frac{v_{inc} - v_{ref}}{Z_0}. \tag{17}$$

Use (13) and (16) in (17) to replace v_{inc} and v_{ref}.

$$i_{TDR} = \frac{v_{pulse}}{2Z_0} + \frac{V_{bias}}{Z_0} - \frac{v_{dut}}{2Z_0} - \frac{V_{bias}}{2Z_0} + \frac{I_{bias}}{2} + \frac{i_{dut}}{2}. \tag{18}$$

Plug (15) into (18) for $\frac{v_{dut}}{2Z_0}$.

$$i_{TDR} = \frac{v_{pulse}}{2Z_0} + \frac{V_{bias}}{Z_0} - \frac{i_{inc}}{2} - \frac{v_{inc}}{2Z_0} + \frac{i_{dut}}{2} - \frac{V_{bias}}{2Z_0} + \frac{I_{bias}}{2} + \frac{i_{dut}}{2}. \qquad (19)$$

Finally, use (13) and (14) for v_{inc} and i_{inc} in (19).

$$i_{TDR} = \frac{v_{pulse}}{2Z_0} + \frac{V_{bias}}{Z_0} - \frac{v_{pulse}}{4Z_0} + \frac{I_{bias}}{2} - \frac{v_{pulse}}{4Z_0} - \frac{V_{bias}}{2Z_0} - \frac{V_{bias}}{2Z_0} + \frac{I_{bias}}{2} + i_{dut}.$$

Simplifying the right-hand-side yields

$$i_{TDR} = i_{dut} + I_{bias}. \qquad (20)$$

Therefore, the current measured using TDR is the DUT current offset by DC current applied from the bias tee. It is simple to show that a voltage measurement using TDR gives a similar result to the current.

$$v_{TDR} = v_{dut} + V_{bias}. \qquad (21)$$

As a quick verification, the reverse recovery of an ESD diode is simulated using the model from chapter 4. The applied DC current is 10 mA. The DUT current is simulated directly at the DUT with TDT and using TDR computationally. The results are shown in Fig. 74. The TDR waveform is offset by the bias current, 10 mA, exactly as predicted. The TDT current waveform subtracted by the TDR waveform gives -10 mA, also predicted by (21).
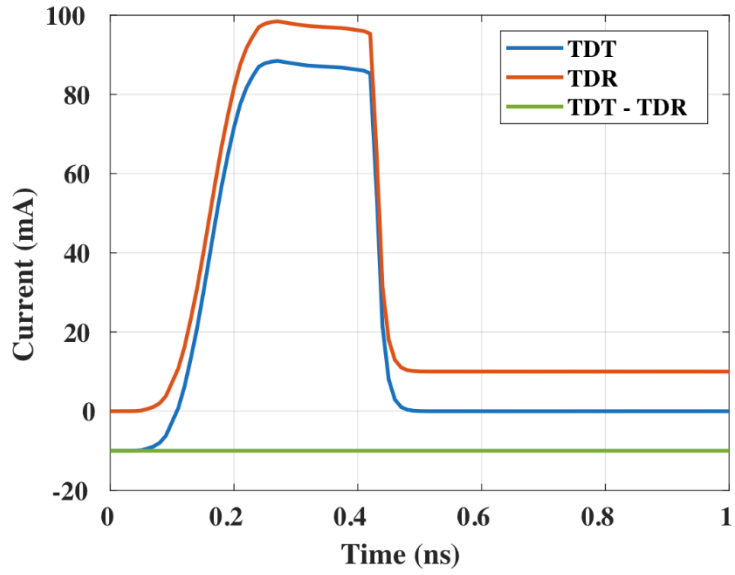
Figure 74: TDT vs. TDR current for the same reverse recovery simulation of the standard model from chapter 4. The TDR measurement is offset by the DC current, which is 10 mA.

APPENDIX D: OSCILLATORY TLP EQUIVALENT CIRCUIT DERIVATION

The purpose of this appendix is to develop an equivalent circuit model of an oscillatory TLP (OTLP) [24] tester that only requires measurement of the incident waveform to simulate the response of any arbitrary load, linear or nonlinear. This is especially useful if an exact model of the LC network is not known or well characterized. The proposed equivalent circuit is shown in Fig. 75.

The schematic implies that as long as the waveform, with the correct characteristic impedance, that hits the DUT is exactly replicated, then the response at the DUT will be the same as if the full schematic were used. To prove this rigorously, start with the full OTLP schematic [24] (Fig. 76).
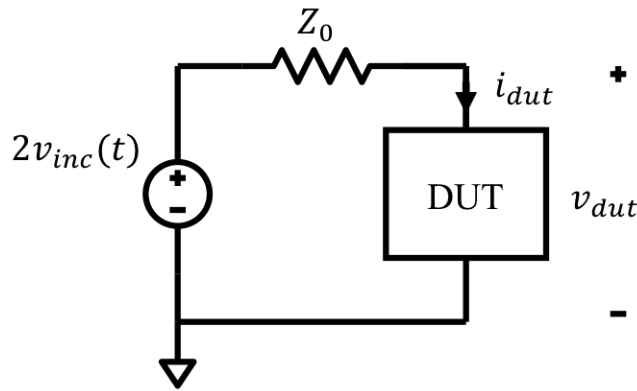
Figure 75: OTLP equivalent circuit, v_{inc} is the incident voltage.
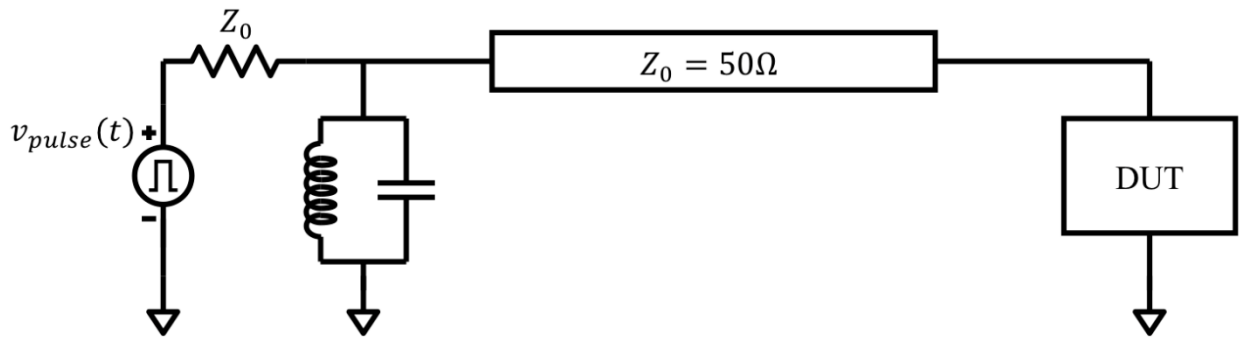
Figure 76: OTLP measurement setup schematic.

82

To conduct the analysis, the transmission line is replaced with its two-port circuit from Appendix C (Fig. 69). During the time period $0 \le t < \tau$ (Fig. 77), the dependent sources of the leftmost and rightmost transmission line ports retain their initial DC value, and thus looking into leftmost port only a Z_0 impedance is seen. The voltage and current at this port are defined as v_{inc} and i_{inc}, where the relationship between the two is

$$i_{inc} = \frac{v_{inc}}{Z_0}.$$

(22)
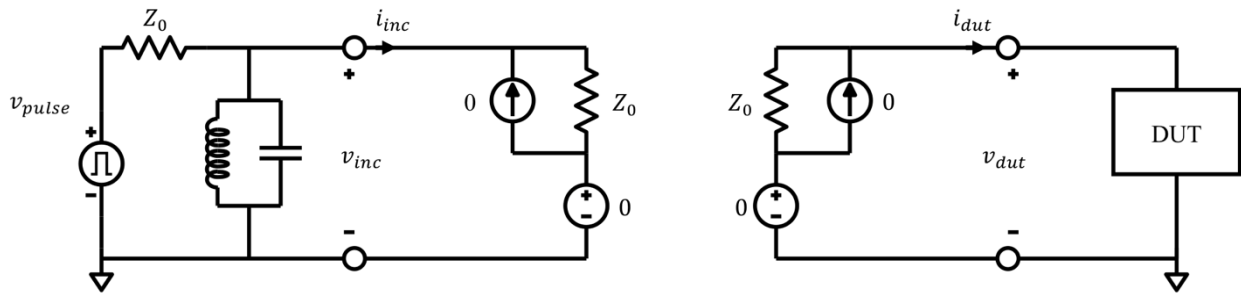
Figure 77: OTLP schematic with transmission line two-port, $0 \le t < \tau$.

The measurement window at the DUT is between $\tau \le t < 3\tau$, where τ is the delay of the transmission line. This window guarantees that no reflections from the source mismatch interfere with the response at the DUT. During this time period, the incident waveform hits the DUT and forces a voltage and current. The schematic of this event at the rightmost transmission line port is shown in Fig. 78. Solve for the DUT current (i_{dut}) in terms of v_{inc}, i_{inc}, and v_{dut}.

$$i_{dut} = i_{inc} - \frac{v_{dut} - v_{inc}}{Z_0}$$

$$i_{dut} = 2\frac{v_{inc}}{Z_0} - \frac{v_{dut}}{Z_0}.$$

(23)
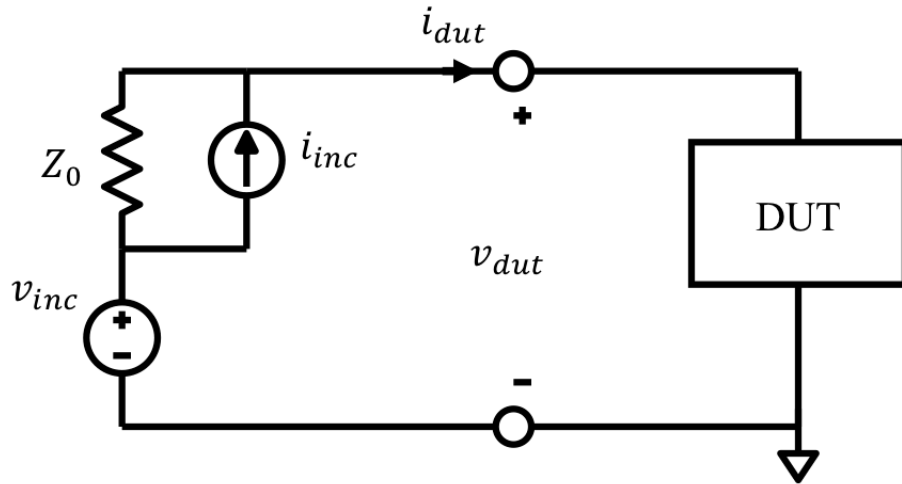
Figure 78: OTLP schematic at the DUT from $\tau \leq t < 3\tau$.

Looking closely at (23), it is exactly equivalent to the circuit in Fig. 75. Therefore, the OTLP response of any arbitrary DUT from time $\tau \leq t < 3\tau$ can be simulated with this circuit and a piecewise linear voltage source to represent the incident waveform. The waveform for this simulation can be measured by placing a matched Z_0 load as the DUT and recording the resulting voltage and current.

APPENDIX E: TDT MEASUREMENT TRICKS

There are a few nuances that can affect the accuracy of a reverse recovery measurement utilizing the TDT setup. For ease of use, each discussion is split into the following sections.

- Oscilloscope Sampling Rate

- Oscilloscope Range

- Oscilloscope Attenuation

- Oscilloscope Offset Error

- Transmission Lines

Oscilloscope Sampling Rate

The two most important specifications of an oscilloscope are the bandwidth and sampling rate. They are two related but altogether different metrics. The bandwidth of the oscilloscope is its frequency response, i.e. the frequency at which there is significant attenuation of the signal. Obviously, a higher bandwidth than the expected signal is required to get a good measurement.

The sampling rate is related to how the resulting signal at the oscilloscope port is discretized by the analog to digital converter (ADC). The Nyquist sampling theorem establishes the lowest sampling rate for resolving a particular signal, but it is almost always best to use the maximum sampling rate available. One quick method of determining if a reverse recovery transient, or any peaking type signal, is under-sampled is to look at the peak value. If the peak changes significantly when the same measurement is done several times, it is likely that the signal is under-sampled. Each time the signal is measured, it is measuring a slightly different peak dependent on the overall jitter of the system. A good example of this can be seen in [21], Fig. 2(a). The higher the sampling rate, the better the measurement.

Oscilloscope Range

The noise associated with the oscilloscope is dependent on the mV/div setting and is usually specified by the manufacturer as some percentage of the full voltage range. For the reverse recovery measurement, it is not necessary to capture the entire waveform during the acquisition. This is especially the case for the DC portion, as this is a known quantity. Significant reduction in noise can be realized by "zooming" into the interesting parts of the current transient, mainly the peak and decaying edge, and ignoring the rest of the transient.

Oscilloscope Attenuation

High-bandwidth oscilloscopes have strict limits on the maximum voltage that can be applied to its ports, and often are around $\pm$ 5 V. To measure a voltage outside of this range, a resistive attenuator must be used. The attenuator is placed in series with the oscilloscope, and it takes the incident waveform and scales it by a factor of $\alpha = 10^{-dB/20}$, where dB is the attenuation given by the manufacturer. The current measured at the oscilloscope with TDT will therefore be

$$i_{dut}(t) = \alpha \frac{v_{meas}(t)}{50}. \tag{24}$$

High-bandwidth attenuators with cutoff frequencies above 20 GHz are readily available and will not significantly degrade the overall bandwidth of the system. Placing an attenuator is not without cost, however, as it will increase the overall noise floor of the system. Use as little attenuation as needed for a given measurement.

<u>Oscilloscope Offset Error</u>

The DC measurement of an oscilloscope trace has error, usually defined is "DC error" by the manufacturer, that may shift the settling point of the reverse recovery measurement. This can be seen by measuring the current well after the recovery has occurred, as shown in Fig. 79. The trace does not settle out to zero as expected, and given that the reverse bias voltage is well below the avalanche breakdown voltage, it must be related to the DC error. To confirm this, the average current over the window in Fig. 79 is measured over a range of reverse bias voltages for two settings. The first auto-ranges the waveform to fit into the minimum possible mv/div. The second setting uses a larger, constant mv/div to capture all possible transients. The results are shown in Fig. 80.

There is a discontinuity in the I-V curve of the measurement with auto-ranging, and this is because the scope is increasing its mV/div at the measurement point which is increasing the DC error. The fixed range measurement has no discontinuity because there is no change in the range of the scope. This shows that the settling level is simply a product of the oscilloscope's error and can be calibrated by subtracting the entire waveform by the settling current. Something similar can be done for regular TLP using an offset correction software available in commercial TLP testers such as the HPPI-3010C. That correction will measure the oscilloscope trace before the pulse is applied and shift the entire waveform such that it is settled at zero, eliminating kinks similar to that seen in Fig. 80. An important caveat for the reverse recovery offset correction is that it is only valid if the diode is operating below the avalanche breakdown voltage, where it is known that the diode does not conduct significant current.
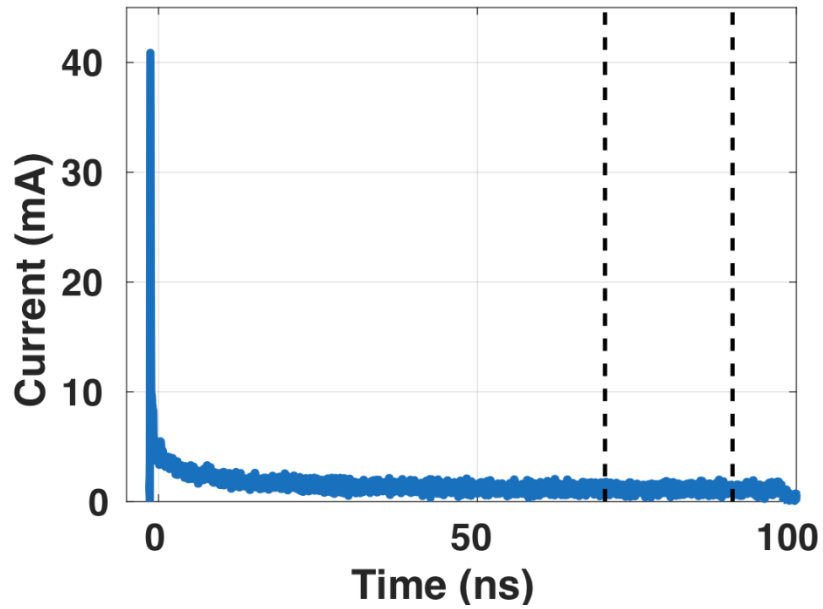
Figure 79: Reverse recovery current measurement of a diode with a 100 ns pulse width.
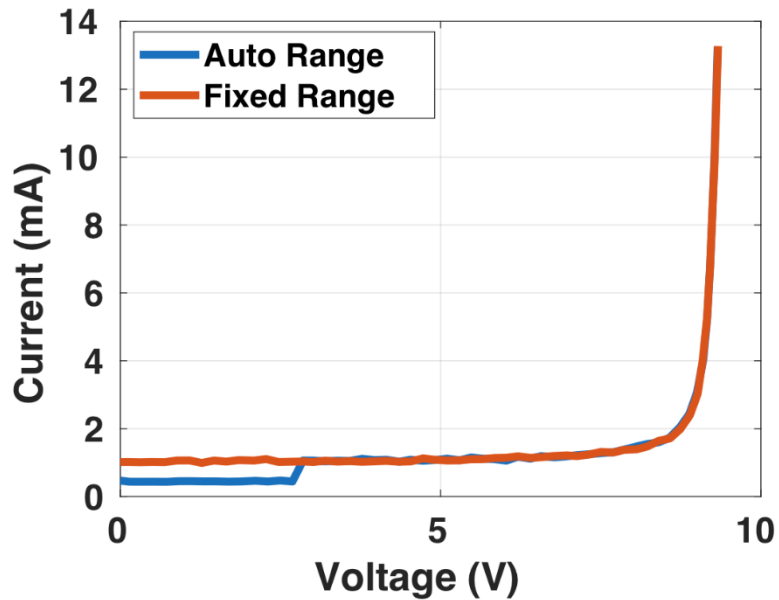
Figure 80: Settling current vs. reverse voltage.

Transmission Lines

Not all transmission lines are created equal. Manufacturers will give S-parameter data of the cables that they offer, including the insertion-loss (S_{21}). The insertion-loss is related to how much of the signal reaches the other side of the transmission line, and ideally it should have a

magnitude of 0 dB. Real transmission lines have losses, however, so the insertion-loss will increase as the frequency increases. Therefore, the first decision when choosing a cable is to look at the insertion-loss and determine which cable has a magnitude closest to 0 dB over the frequency range of interest.

Additionally, the loss can further be minimized by choosing the smallest cable length possible for the TDT connection. A smaller cable will have lower loss than a longer one, and since there is no minimum length requirement on the TDT transmission line like there is for TDR, it should be as short as physically possible.

REFERENCES

[1]   R. Muller, T. Kamins and M. Chan, "Currents in pn junctions," in *Device Electronics for Integrated Circuits,* third ed. U.S: John Wiley & Sons, 2003, ch. 5, pp. 258-259.

[2]   J. Whitfield *et al.*, "ESD MM failures resulting from transient reverse currents," *2006 IEEE International Reliability Physics Symposium Proceedings*, San Jose, CA, 2006, pp. 136-139.

[3]   T. Hirano, M. Hase, T. Ogura, S. Tanaka and S. Fujiwara, "Anomalous ESD failures in NLDMOS during reverse recovery," *Electrical Overstress/Electrostatic Discharge Symposium Proceedings 2010,* Reno, NV, 2010, pp. 1-6.

[4]   Z. Lin, G. Sun and F. Lin, "A MM (machine model) failure case study: Parasitic NPN turn-on induced by reverse recovery current," *2017 IEEE 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, Chengdu, 2017, pp. 1-3.

[5]   M. D. Ker and S. F. Hsu, "Evaluation on board-level noise filter networks to suppress transient-induced latchup in CMOS ICs under system-level ESD test," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 48, no. 1, pp. 161-171, Feb. 2006.

[6]   A. Patnaik *et al.*, "Systematic evaluation of soft failures in system-level ESD transient events," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 60, no. 5, pp. 1263-1269, Oct. 2018.

[7]   M. Park *et al.*, "Measurement and analysis of statistical IC operation errors in a memory module due to system-level ESD noise," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 61, no. 1, pp. 29-39, Feb. 2019.

[8]   J. Catt, "An improved method for ultra fast recovery diode testing," *Proceedings of 1994 IEEE Applied Power Electronics Conference and Exposition - ASPEC'94*, Orlando, FL, USA, 1994, pp. 473-479 vol.1.

[9]   D. L. Mauch *et al.*, "Ultrafast reverse recovery time measurement for wide-bandgap diodes," in *IEEE Transactions on Power Electronics*, vol. 32, no. 12, pp. 9333-9341, Dec. 2017.

[10] R. H. Dean and C. J. Nuese, "A refined step-recovery technique for measuring minority carrier lifetimes and related parameters in asymmetric p-n junction diodes," in *IEEE Transactions on Electron Devices*, vol. 18, no. 3, pp. 151-158, March 1971.

[11] W. Simburger, D. Johnsson and M. Stecher, "High current TLP characterisation: an effective tool for the development of semiconductor devices and ESD protection solutions," *ARMMS RF & Microwave Society,* 2012.

[12] M. Sauter, W. Simburger, D. Johnsson, M. Stecher, "On-wafer measurement of the reverse-recovery time of integrated diodes by Transmission-Line-Pulsing (TLP)," in *Proc. 2011 Microelectronics Reliability Conf.*, pp. 1309-1314.

[13] K. Rajagopal *et al.*, "Investigation of reverse recovery effects on the SOA of integrated high frequency power transistors," *2016 IEEE International Reliability Physics Symposium (IRPS),* Pasadena, CA, 2016, pp. 6A-3-1-6A-3-9.

[14] H. A. Gieser and H. Wolf, "Survey on very fast TLP and ultra fast repetitive pulsing for characterization in the CDM-domain," *2007 IEEE International Reliability Physics Symposium Proceedings. 45th Annual*, Phoenix, AZ, 2007, pp. 324-333.

[15] *Advanced Design System.* (2015). Keysight.

[16] J. Manouvrier *et al.*, "A physics-based compact model for ESD protection diodes under very fast transients," *EOS/ESD 2008 - 2008 30th Electrical Overstress/Electrostatic Discharge Symposium*, Tucson, AZ, 2008, pp. 67-75.

[17] H. Li, M. Miao, Y. Zhou, J. A. Salcedo, J. Hajjar and K. B. Sundaram, "Modeling and simulation of comprehensive diode behavior under electrostatic discharge stresses," in *IEEE Transactions on Device and Materials Reliability*, vol. 19, no. 1, pp. 90-96, March 2019.

[18] X. Han and S. Zhao, "A comprehensive physical model for PNP based ESD protection devices in SOI technology," *2019 41st Annual EOS/ESD Symposium (EOS/ESD)*, Riverside, CA, USA, 2019, pp. 1-7.

[19] A. Romanescu, H. Beckirch-Ros, P. Fonteneau, C. Legrand, P. Ferrari and J. Arnould, "Scalable modeling studies on the SCR ESD protection device," *EOS/ESD Symposium Proceedings*, Anaheim, CA, 2011, pp. 1-8.

[20] R. Mertens and E. Rosenbaum, "A physics-based compact model for SCR devices used in ESD protection circuits," *2013 IEEE International Reliability Physics Symposium (IRPS)*, Anaheim, CA, 2013, pp. 2B.2.1-2B.2.7.

[21] Z. Pan, D. Schroeder, S. Holland and W. H. Krautschneider, "Understanding and modeling of diode voltage overshoots during fast transient ESD events," in *IEEE Transactions on Electron Devices*, vol. 61, no. 8, pp. 2682-2689, Aug. 2014.

[22] A. Ayling, S. Huang and E. Rosenbaum, "Sub-nanosecond reverse recovery measurement for ESD devices" in *International Reliability Physics Symposium (IRPS) 2020,* April 2020.

[23] S. Huang, "Model for sub-nanosecond switching transients of ESD diodes", B.S. Thesis, University of Illinois – Urbana Champaign, May 2020.

[24] J. P. Di Sarro and E. Rosenbaum, "Oscillatory transmission line pulsing for characterization of device transient response," in *IEEE Electron Device Letters*, vol. 30, no. 2, pp. 168-170, Feb. 2009.

[25] R. Merrill and E. Issaq, "ESD design methodology," in *Proc. Elect. Overstress / Electrostatic Discharge Symposium*, pp. 233-237, 1993.

[26] J. C. Smith and G. Boselli, "A MOSFET power supply clamp with feedback enhanced triggering for ESD protection in advanced CMOS technologies," *2003 Electrical Overstress/Electrostatic Discharge Symposium*, Las Vegas, NV, 2003, pp. 1-9.

[27] J. Li, R. Gauthier and E. Rosenbaum, "A compact, timed-shutoff, MOSFET-based power clamp for on-chip ESD protection," *2004 Electrical Overstress/Electrostatic Discharge Symposium*, Grapevine, TX, 2004, pp. 1-7.

[28] H. Sarbishaei, O. Semenov and M. Sachdev, "A new flip-flop-based transient power supply clamp for ESD protection," in *IEEE Transactions on Device and Material Reliability*, vol. 8, no. 2, pp. 358-367.

[29] R. Venkatasubramanian, K. Oertle and S. Ozev, "Rail clamp with dynamic time-constant adjustment," in *IEEE Journal of Solid-State Circuits*, vol. 51, no. 5, pp. 1313-1324, May 2016.

[30] S. Parthasarathy, J. A. Salcedo, S. Herrera and J. Hajjar, "ESD protection clamp with active feedback and mis-trigger immunity in 28nm CMOS process," *IEEE International Reliability Physics Symposium*, Monterey, CA, 2015, EL.3.1-EL.3.5.

[31] T. J. Maloney and W. Kan, "Stacked PMOS clamps for high voltage power supply protection," *Electrical Overstress/Electrostatic Discharge Symposium Proceedings. 1999 (IEEE Cat. No.99TH8396)*, Orlando, FL, USA, 1999, pp. 70-77.

[32] R. Mertens and E. Rosenbaum, "Separating SCR and trigger circuit related overshoot in SCR-based ESD protection circuits," *2013 35th Electrical Overstress/Electrostatic Discharge Symposium*, Las Vegas, NV, 2013, pp. 1-8.

[33] S. Ruth *et al.*, "Investigation of product burn-in failures due to powered NPN bipolar latching of active MOSFET rail clamps," *2013 35th Electrical Overstress/Electrostatic Discharge Symposium*, Las Vegas, NV, 2014, pp. 1-10.

[34] A. Gerdemann, J. W. Miller, M. Stockinger, N. Herr, A. Dobbin and R. Ricklefs, "When good trigger circuits go bad: A case history," *EOS/ESD Symposium Proceedings*, Anaheim, CA, 2011, pp. 1-6.

[35] T. Maloney and N Khurana, "Transmission line pulsing techniques for circuit Modeling of ESD phenomena", *1985 EOS/ESD Symposium Proceedings*, pp. 49-54.

[36] J. Schutt-Aine. (2020). Timing and signaling [PowerPoint slides]. Available: http://emlab.uiuc.edu/ece546/Lect_22.pdf

[37] H. W. Dommel, "Digital computer solution of electromagnetic transients in single-and multiphase networks," in *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-88, no. 4, pp. 388-399, April 1969.