TCAD methodology for ESD robustness prediction of smart power ESD devices

Christophe Salamero, Nicolas Nolhier, Amaury Gendron, Marise Bafleur, Patrice Besse, Michel Zecri

TCAD Methodology for ESD Robustness Prediction of Smart Power ESD Devices

Christophe Salaméro, Nicolas Nolhier, Amaury Gendron, Marise Bafleur, *Senior Member, IEEE*, Patrice Besse, and Michel Zécri

***Abstract*—This paper presents a new method to predict the electrostatic-discharge (ESD) protection robustness of a device with technology-in-computer-aided-design (TCAD) simulations. Tested on different devices and two Smart Power technologies, the results are validated through electrical measurement and failure analysis. Failure current is always predicted with a good accuracy compared to technology spreading. In addition, the methodology provides a significant simulation time speedup compared to classical methods based on a temperature criterion.**

***Index Terms*—Electrostatic discharge (ESD), predictive simulation, robustness, Smart Power technology, technology in computer aided design (TCAD) simulation.**

I. Introduction

PREDICTIVE technology in computer aided design (TCAD) has been shown to be a powerful tool for the development of low-voltage CMOS and VLSI memory technology development [1], [2] as well as for the development of mixed-mode and complex technologies such as the Smart Power ones [3]. There is a strong economical incentive for extrapolative quantitatively accurate TCAD capabilities, even if only possible under some restrictions. For a semiconductor company, performing circuit and system development concurrently with the experimental process development considerably shorten the time to qualification and market of a product.

The qualification of an integrated circuit also includes its electrostatic-discharge (ESD) robustness. In a Smart Power technology, the design of ESD protection structures is quite challenging, since it requires various clamping voltages and current capabilities according to the specifications of each I/O pin. Clamping voltages should be above the power supply one, which on some pins can be higher than 40 V. Moreover, the tough environment of applications such as the automotive one requires a strong ESD robustness (4- to 10-kV HBM), i.e., ESD currents ranging from 2.6 to 6.5 A.

Despite the continuous research efforts in the field of integrated ESD protections, ESD failures are still a major cause of both circuit redesigns and field returns (as high as 30% in the automotive application). Since about a decade, TCAD tools have been extensively used for the optimization of ESD protection structures. Applying the same predictive methodology to these devices should also have a strong impact on the time to market. To be efficient, it has to be carried out in parallel with the technology development and requires an accurate prediction of the maximum ESD current that the protection can withstand. The success of predictive simulations critically depends upon the quality of the physical models coded in the TCAD tool and their calibration to silicon data. ESD failure is generally a thermal mechanism inducing silicon and interconnection melting or oxide breakdown, which means high local temperatures: 1683 K [4] for silicon or 933 K [4] and 1356 K [4] for aluminum and copper metallization, respectively. During an ESD stress, the local temperature increase within the protection device can be rather high, and many authors [5], [6] used the rising of the silicon melting temperature as a failure criterion. Unfortunately, the physical models are only valid up to a temperature of 600 K and 800 K for the best cases [7], making such criterion very questionable.

In this paper, we present a predictive methodology based on TCAD simulation results obtained within the temperature validity range of the physical models. This method allows then, predicting with a good accuracy, the failure current I_{T2} value under a transmission-line-pulsing (TLP) stress. Instead of using the temperature value as a criterion, temperature related parameters, i.e., impact ionization and thermally generated carriers, are monitored until maximum allowed temperature for the models. Their behavior is extrapolated until their respective current contribution becomes equal: This condition is considered as the initiation of thermal breakdown and is used as the failure criterion. The methodology is validated on two Smart Power technologies with 0.35- and 0.25-μm minimum feature size.

II. Devices Under Test

A. First Technology

We first applied the method for I_{T2} prediction on three different ESD protection devices (device1, device2, and device3) from a junction-isolated Smart Power technology (0.35-μm SMARTMOS technology). Each ESD element results from a standard high-voltage n-p-n bipolar transistor (Fig. 1) that is designed with specific ESD design and used as a grounded-base bipolar ESD protection.

The main layout differences between the three ESD protection structures are the base doping, the spacing between

Manuscript received December 2, 2005; revised April 6, 2006. This work was supported in part by Freescale Semiconductors.

C. Salaméro, N. Nolhier, and M. Bafleur are with the Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS-CNRS), Université de Toulouse, 31077 Toulouse Cedex 4, France (e-mail: csalamer@laas.fr; nolhier@laas.fr; marise@laas.fr).

A. Gendron and P. Besse are with Freescale Semiconductors, 31023 Toulouse, France (e-mail: pbes01@freescale.com; agen01@freescale.com).

M. Zécri is with Freescale Semiconductors, 38920 Crolles, France (e-mail: michel.zecri@freescalecrolles.st.com).

Digital Object Identifier 10.1109/TDMR.2006.881483

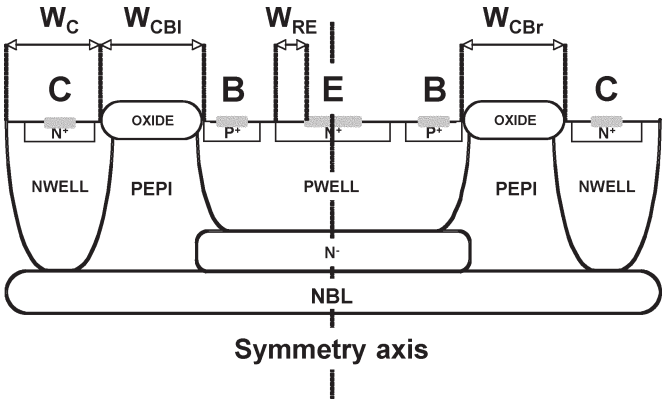

Fig. 1. 2-D cross section of standard n-p-n bipolar transistor.

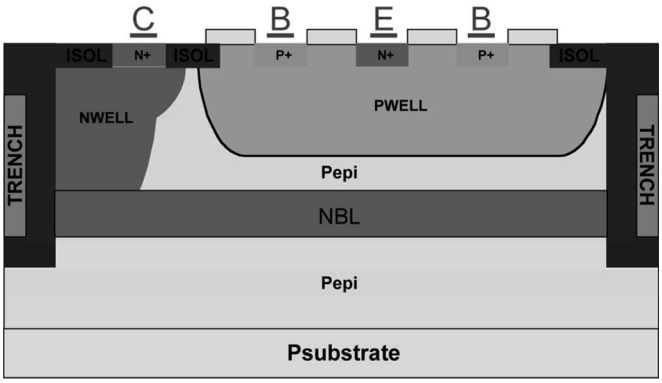

Fig. 2. Device4 technological cross section.

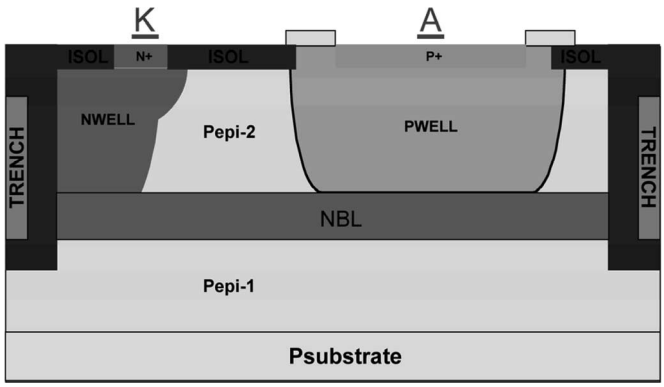

Fig. 3. Device5 technological cross section.

collector and base diffusion (W_{CB1} and W_{CBr} are equal for the standard n-p-n bipolar transistor), the collector diffusion width W_{C}, and emitter diffusion width W_{RE}, where the latter defines the intrinsic emitter ballast resistor.

B. Second Technology

The second technology is a more advanced Smart Power technology (0.25-μm SMARTMOS technology). Its doping profiles are similar to the previous technology ones, but in this case, isolation between components is obtained through deep oxide trenches (Fig. 2). Two ESD devices were studied: one n-p-n bipolar transistor (device4: Fig. 2) and one p-n diode (device5: Fig. 3).

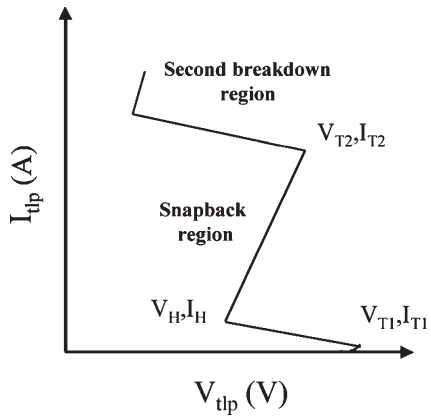

Fig. 4. Typical TLP characteristic.

III. Method

A. Experimental Failure Criterion

TCAD is widely used in power devices to simulate electrical characteristics such as breakdown voltage (BV) or ON-resistance to optimize performance [8]. Also, it is very useful for ESD protection design to adjust triggering V_{T1}, I_{T1}, and holding V_{H}, I_{H} points when the device exhibits a snapback behavior. Fig. 4 shows a typical quasi-static I–V curve from a TLP characterization [9]. However, simulating the failure point V_{T2}, I_{T2} is much more difficult.

The experimental determination of this point consists in monitoring the device leakage current between each increasing stress current pulses. When the leakage current is higher than a critical value, the protection is defined as failed. This method cannot be applied on simulation, since a defect creation such as silicon melting or oxide breakdown is not modeled. It is then necessary to develop a novel methodology based on TLP simulation with a new criterion allowing the I_{T2} failure current value estimation. For this aim, an accurate simulated TLP curve must be obtained compared to an experimental behavior, and it requires a meticulous calibration step.

In the following sections, all simulation results were obtained from TCAD-ISE tools release 7.0 [10].

B. Calibration

1) Device Profile Definition: There are two different ways of describing doping profiles for device simulation. The first method uses an analytical distribution, and the other one involves a process simulation.

In this paper, each device is described from a calibrated process simulation to obtain the best accuracy for the doping profiles. For instance and as depicted in Fig. 5, device1 is defined from a process simulation. Device1 has the same base doping as the original bipolar transistor but does not keep its symmetry. Indeed, the distance W_{CBl} on the left side of the device is reduced (compared to the W_{CBr} distance on the right side) to force the static breakdown to this side. It is localized at the NWELL/PEPI-P^{+} cylindrical junction (Fig. 5). As its BV occurs on a lateral junction, an accurate description of the lateral diffusion of the simulated doping profile is required. For

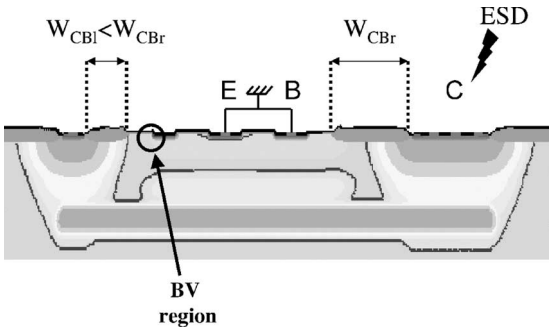

Fig. 5. Net-doping distribution on device1 obtained from a process simulation.

TABLE I
MEASUREMENT AND SIMULATED BV AND V_{T1} FOR DEVICE1 DOPING PROFILES ARE OBTAINED BY TWO DIFFERENT WAYS

	BV (V)	V_{T1} (V)
DC/TLP measurement	42	42.5
TCAD with analytical profiles	52	53
TCAD with process simulated profiles	43.8	45

the same device, we have also analytically described the doping profile. The lateral extension of the diffusions is then defined by the user who adjusts the value through the choice of a classical lateral-diffusion coefficient from a Gaussian function. In this case, the simulated BV value is different, and it occurs at a different location, namely at the PWELL/N^- planar junction. The results of this comparison are summarized in Table I for both the dc BV and the triggering voltage V_{T1} of the device1 under a TLP stress.

The comparison with electrical measurement data allows concluding that the doping-profile description of this device from the process simulation results in the best accuracy for both BV and V_{T1}. Moreover, the use of an analytical description of the lateral diffusion results in a wrong BV location and, therefore, a 20% error for the BV simulation. In this case, this means that the optimization of the structure cannot be correctly achieved by using analytical profiles. It is preferable to describe the doping profiles of the studied devices from a calibrated process simulation, which provides a better accuracy for the BV location, its voltage value, and the simulated response characteristic to a TLP stress.

2) Choice of Physical Models: The choice of physical models is important to improve the accuracy of the simulation results. This is not an easy task for two reasons. First, physical models have to be chosen according to physical phenomena occurring within the device during an ESD stress. Second, an additional difficulty is to select the best one among several models describing the same physical phenomenon.

When an ESD protection structure turns ON, two main physical phenomena are involved: avalanche breakdown and snapback effect with parasitic bipolar triggering. It has to be noticed that the second mechanism is associated with high current effects. Indeed, at high current density, in such graded collector bipolar device, the base–collector junction is not located anymore at the metallurgical junction. It is pushed toward a higher doped region, here, the NBL buried layer, where an effective base–collector junction can be defined according to the current level [11].

TABLE II
SIMULATED BV COMPUTED FROM DIFFERENT AVALANCHE MODELS COMPARED WITH MEASUREMENT

	BV (V)
Measurement	42
vanOverstraeten/de Man	43.8
Okuto/Crowell	44.1
Lackner	44.8
University of Bologna	45

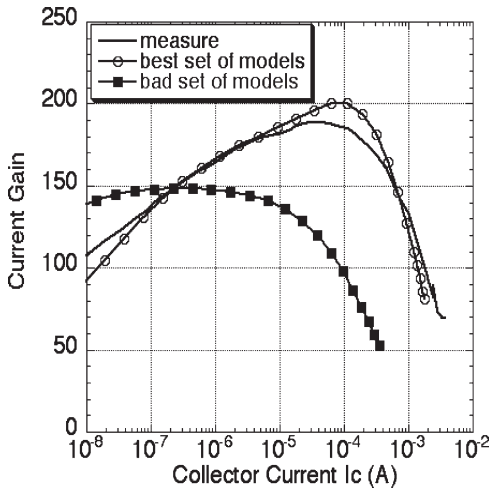

Fig. 6. Measurement and β simulation results according to the choice of physical models set.

To study the impact of the models on the electrical simulation results, we performed simulations on device1 (Fig. 5). We computed the base–collector BV with emitter and base grounded but also the current gain β.

Concerning the avalanche phenomenon, DESSIS-ISE [10] device simulator proposes a choice of four models for the ionization coefficients. The simulation results are summarized in Table II. Compared with measurement, Overstraeten and Man's [12] model offers the best fitting.

The current gain β is strongly influenced by the choice of physical models and corresponding model parameters, e.g., for mobility, intrinsic density, recombination models, and others. To illustrate their impact, Fig. 6 shows the β simulation results (obtained from Gummel plot) according to the selection of the best set of physical models/parameters in comparison to an example for a poor simulator calibration.

In addition to selecting the "right" model, we also calibrate model's parameters. The lifetime of electrons and holes is the most important parameter to adjust [5]. Indeed, the carrier-lifetime parameter is strongly process dependent, thus requiring to calibrate it for each technology under investigation. A bad calibration of physical models/parameters could induce future wrong TLP simulation and, as a result, wrong V_{T1} or V_H values. Therefore, these could lead to an overheating of the studied device and to the prediction of a premature failure.

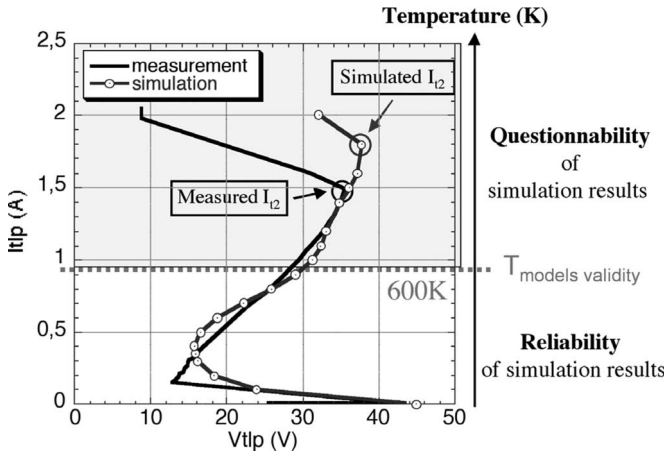

Fig. 7. Measured and simulated $I(V)$ TLP characteristic for device1.

In the device-simulator DESSIS, the following models are chosen.

1) For the effective intrinsic-density calculation, the band-gap narrowing is taken into account by the Slotboom model [13].
2) PhuMob for mobility model [13].
3) Van Overstraeten/deMan [12] impact ionization model for the avalanche process.

C. Predictive Methodology for I_{T2}

1) Former Approach: Actually, few people worked on the prediction of the failure current I_{T2} of an ESD protection device through the TCAD simulation.

In the literature, Esmark [5] is the major author who made investigations on a methodology to predict the failure-current value. This author applies TCAD-ISE tools for all physical two-dimensional (2-D) device simulations. As he observed two different failure modes, he proposed two criteria to extract I_{T2} value. The first method consists of running successive TLP simulations until thermal second breakdown I_{T2} occurrence. However, as some devices fail without exhibiting a visible second snapback on measured characteristics, a second criterion based on the temperature value is needed. Therefore, the criteria are summarized by

$$I_{\mathrm{T2simulated}} := \{T_{\max} > T_{\mathrm{crit}} \text{ locally, occurrence of second breakdown}\}$$

where $T_{\mathrm{crit}} = 1693$ K is the temperature of silicon melting point.

Unfortunately, whatever criterion is used, the temperature in the simulated device reaches a value that is well above the temperature validity of the physical models. As a consequence, the confidence that can be placed in the corresponding simulation results is questionable (Fig. 7).

We have applied Esmark's criteria to the five ESD protection devices under study. For device1, it resulted in 20% error compared to the measurement (Table III). This 2-D simulation result was obtained after seven days of computation time using a SUN workstation with a 1-GHz processor. This device is fine meshed with 15 000 grid points, and 24 (I_{tlp}, V_{tlp}) points were computed to reach the second snapback on the simulated TLP characteristic.

TABLE III
I_{T2} SIMULATED AND MEASURED VALUE FOR DEVICE1

I_{T2} (A)	Device1
Measure	1,5
Esmark's Criteria	1,8
Error (%)	20

2) Novel Approach: To cope with the temperature limitation of the simulation, we propose an I_{T2} prediction methodology based on the analysis of simulation results performed within the temperature validity range. The entire paper is only based on 2-D simulations for two major reasons. First, even if three-dimensional (3-D) simulations are known as the better way to obtain accurate triggering and holding voltage [5], they are too CPU time consuming to compute TLP curves. Second, for all the simulated devices elementary ESD design rules, such as ballasting, are applied. Therefore, these devices are assumed to have a uniform current density along the device width in the snapback region, within the model's validity temperature range.

The proposed method first consists of obtaining the most accurate TLP simulation results at low current level. This is achieved through a careful calibration of both the structure description (doping profiles and geometry) and the electrical simulation as described in previous sections.

To simulate TLP stresses, i.e., square stress pulses on a device under test (DUT), transient electrothermal simulations are performed in a mixed-mode configuration, taking account parasitic elements of our experimental TLP bench. A time-dependent current source generates the rectangular pulses of 100-ns duration and 2-ns rise time corresponding to the experimental conditions.

Next, successive TLP simulations are carried out using the temperature-dependent equation set. For each simulation, we localize the hottest spot in the device where a maximum temperature value T_{MAX} is reached. Simulations are stopped when the temperature in the device reaches 600 K at the end of the TLP current pulse. For each T_{MAX} point ($\leq$ 600 K), the values of both impact ionization G_{i} and Shockley Read Hall recombination rate R_{SRH} are extracted and plotted in the same graph as a function of TLP current I_{TLP}. The evolution of these two parameters is then exponentially extrapolated into the critical temperature regime ($>$ 600 K) until they intersect as demonstrated in Fig. 10. The I_{TLP} current value at the intersection where R_{SRH} begins to exceed G_{i} is defined as the simulated failure current I_{T2}. The underlying physics for this I_{T2} simulation methodology is explained in the following.

Second breakdown has been extensively studied [6], [14]. The onset of second breakdown is indicated by a voltage decrease resulting from a localized thermal run away induced by a hot spot. The related local power generation induces enough thermally generated carriers to sustain the current instead of the electrically generated ones. In the case of the n-p-n bipolar under study (device1), it means that the thermally generated current I_{th} becomes the dominant component of the base

current $I_{\rm b}$ that was initially provided by the avalanche generated current $I_{\rm av}$ [6]. The occurrence of the second breakdown $I_{\rm T2}$ can then be defined when $I_{\rm th} > I_{\rm av}$.

The thermally generated current $I_{\rm th}$ is proportional to the intrinsic concentration $n_{\rm i}$, since it is exponentially temperature dependent. In the same way, the electrically generated current $I_{\rm av}$ is proportional to the concentration $n_{\rm av}$ of generated carriers during the avalanche process.

As a result, a new formulation of the occurrence of second breakdown can be written: $n_{\rm i} > n_{\rm av}$. Effectively, the temperature in the device locally rises up, the intrinsic concentration $n_{\rm i}$ due to thermal carrier generation increases until it exceeds the carrier concentration generated by an impact ionization. Thus, $n_{\rm i}$ becomes dominant in the current flow. Concurrently, the voltage in the device strongly decreases, and the current crowds into an extremely localized area leading to current filamentation. In turn, the temperature further increases, resulting in the reduction of electrically generated carriers issued from the impact ionization and the increase of thermally generated carriers. This behavior is called a thermal run away or second breakdown. It corresponds to the point where thermal carrier generation exceeds electrical avalanche generation.

The evolution of the SRH recombination rate $R_{\rm SRH}$ can be used to monitor the thermally generated carrier evolution. Indeed, it represents the indirect recombination via deep impurities. $R_{\rm SRH}$ is an image of the carrier density with respect to its equilibrium value. The plasma density in equilibrium is given by the effective intrinsic carrier density $n_{\rm i,eff}$. The corresponding model of $R_{\rm SRH}$ is described by the following equation:

$$R_{\rm SRH} \propto \frac{np - n_{\rm i,eff}^2}{\tau_p(n + n_1) + \tau_n(p + p_1)}$$

with $n_1 = n_{\rm i,eff} e^{E_{\rm trap}/kT}$ and $p_1 = n_{\rm i,eff} e^{-E_{\rm trap}/kT}$.

$E_{\rm trap}$ is the difference between the defect level and the intrinsic level.

τ_p and τ_n are the minority carrier lifetimes and are modeled as

$$\tau_c = \tau_{\rm dop} \frac{f(T)}{1 + g_c(F)}$$

with $c = e$ for electrons or $c = h$ for holes.

$\tau_{\rm dop}$ is the doping dependence of the SRH lifetimes modeled with the Scharfetter relation [15]. f and g_c are two temperature and electric-field dependent functions, respectively.

$R_{\rm SRH}$ is positive when carrier recombination is larger than carrier generation. It becomes negative when thermal generation exceeds recombination. When $R_{\rm SRH}$ is negative, it represents the $n_{\rm i,eff}$ behavior, because $n_{\rm i,eff}$ is a monotonically increasing function of the lattice temperature. As previously stated, the thermally generated current $I_{\rm th}$ is proportional to the intrinsic concentration $n_{\rm i}$. Therefore, $R_{\rm SRH}$ represents indirectly the thermally generated current $I_{\rm th}$.

In the same way, impact ionization $G_{\rm i}$ indirectly represents $I_{\rm av}$ since it is proportional to $n_{\rm av}$.

We can finally reformulate the condition of the second-breakdown occurrence as $R_{\rm SRH} > G_{\rm i}$.

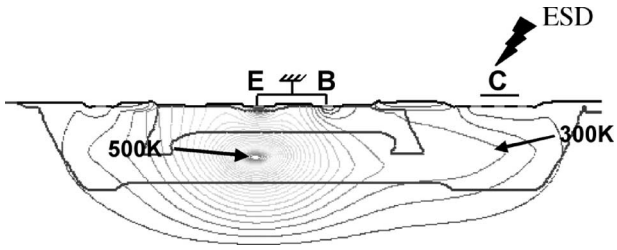

Fig. 8. Localization of the temperature maximum in the device1 at low current $I_{\rm TLP}$.

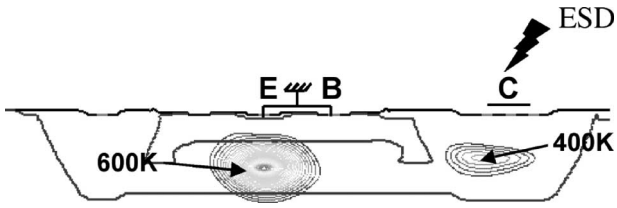

Fig. 9. Localization of the temperature maximum in the device1 when $I_{\rm TLP}$ increases.

The onset of fatal breakdown can then be tracked by monitoring now the evolution of both impact ionization $G_{\rm i}$ and the SRH recombination rate.

Right after the first snapback and at low current/temperature levels, the carriers are mainly electrically generated through impact ionization at low $n_{\rm i,eff}$ value. As a result, $R_{\rm SRH}$ is negligible compared to $G_{\rm i}$. Note that, $G_{\rm i}$ carrier generation is always greater than $|R_{\rm SRH}|$ in the temperature range below 600 K (Fig. 10).

At larger currents, the higher device temperature results in an increase of $n_{\rm i,eff}$ and, consequently, of $|R_{\rm SRH}|$.

It can be noticed that the $|R_{\rm SRH}|$ can only be extracted from a certain temperature level. Its evolution is fitted by a single exponential function. $G_{\rm i}$ evolution is described by two exponential functions over the temperature range (Fig. 10), but only the second one is used to apply our methodology.

$G_{\rm i}$ and $|R_{\rm SRH}|$ parameters exponentially increase with the TLP current until the temperature in the device reaches the maximum temperature for physical models validity (600 K). Above this value, the actual evolution of these parameters with regard to the temperature is not known. For our predictive method, we assume that the models keep the same exponential behavior. The simulation computation is stopped at 600 K, and $G_{\rm i}$ and $|R_{\rm SRH}|$ evolution are exponentially extrapolated until they intersect at the point where $R > G$. The intersection point defines the simulated failure current $I_{\rm T2}$.

3) Application to Device1: For this device, we first observe at low current level a one single hot spot localized below the emitter (Fig. 8).

At a higher current level, a second hot spot occurs in another region under the collector contact (Fig. 9).

We then chose to apply the methodology to the two hot spots (Figs. 10 and 11). On both figures, the 600-K model validity limit is indicated. For the hot spot on the collector side, the intersection between $G_{\rm i}$ and $|R_{\rm SRH}|$ occurs for $I_{\rm T2} =$ 1.52 A (Fig. 8).

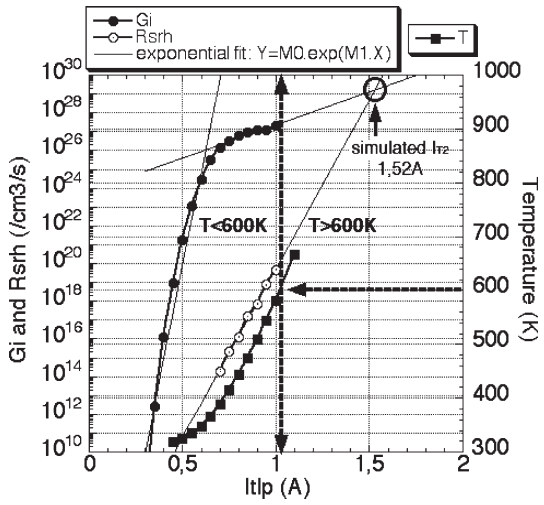

Fig. 10. Evolution of $G_{\rm i}$ and $|R_{\rm SRH}|$ parameters on the collector side.

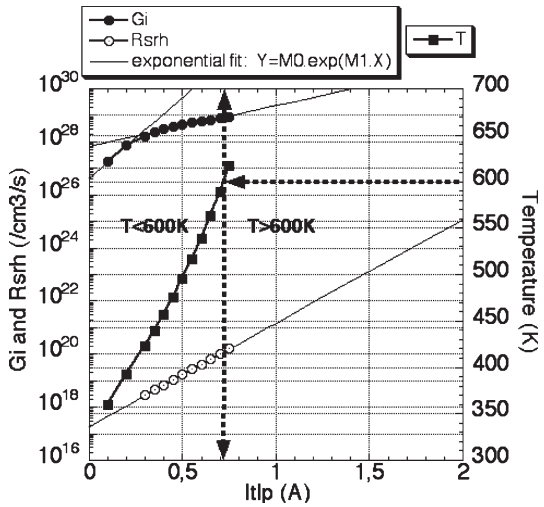

Fig. 11. Evolution of $G_{\rm i}$ and $|R_{\rm SRH}|$ parameters on the emitter side.

The failure occurs much earlier than for the hot spot on the emitter side where the $I_{\rm T2}$ intersection point is not visible within the range of the graph (Fig. 11). The predicted $I_{\rm T2}$ failure current extracted from the hottest spot (collector side) accurately correlates with the measurement, thus indicating that the failure of the device originates in the collector region. To validate this result, we performed a deprocessing of the failed device. It confirmed that the device destruction is initiated at the collector side (Fig. 12).

4) Justification of Two Hot-Spot Occurrences: When the current densities become important, we can observe in the area under the emitter (Fig. 13) that the maximum electric field, which is initially located in the silicon depth at the PWELL/N^- metallurgical junction, is shifted toward the effective junction N^-/NBL by Kirk effect [16].

As TLP current increases, an intrinsic zone of carriers occurs into the lowly doped diffusion region: PWELL and N^- (Fig. 14). Electron and hole carrier densities are equal and above the PWELL and N^- doping concentrations. The carrier and doping profiles can be observed at the Fig. 14 ob-

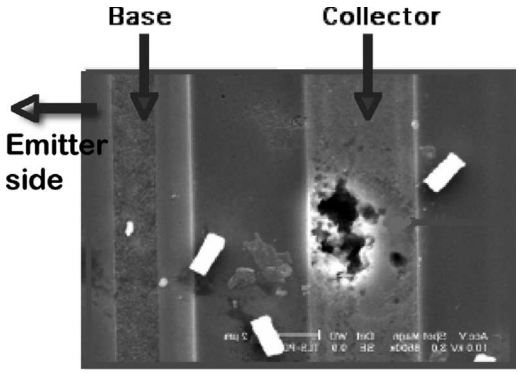

Fig. 12. SEM top view of the failure after delayering (device1).

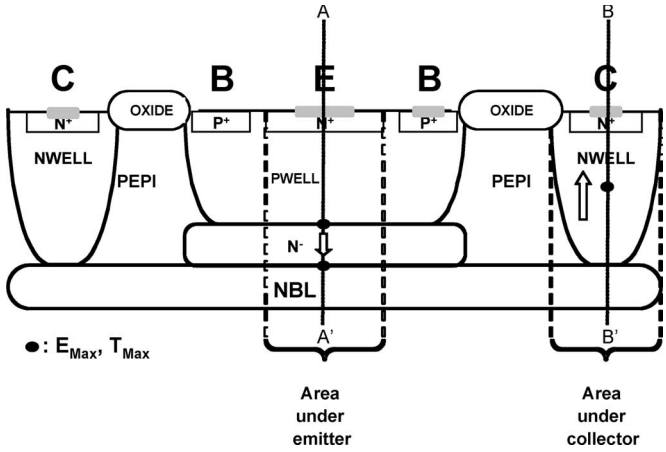

Fig. 13. Maximum electric field and temperature localization into device1 as and when current increases.

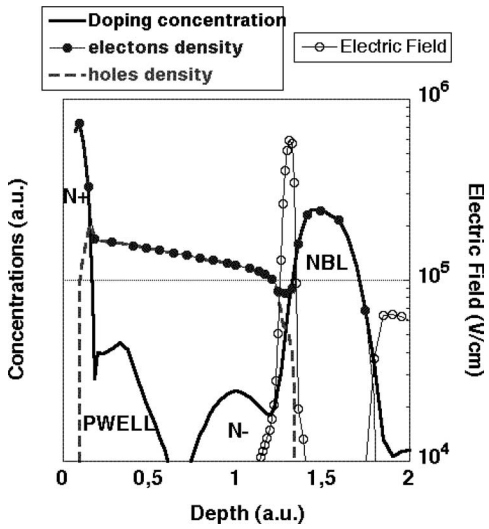

Fig. 14. Distribution of carriers and electric field maximum along AA′ cutline for a TLP current value $I_{\rm TLP} = 0.8$ A.

tained through along AA′ cutline (Fig. 13). When TLP current increases, the intrinsic zone spreads into the silicon depth where doping concentrations are the lowest. This intrinsic-zone extension is finally stopped by the buried layer doping (NBL), which is highly doped. For this reason and despite additional TLP current increase, the maximum electric field stays located at

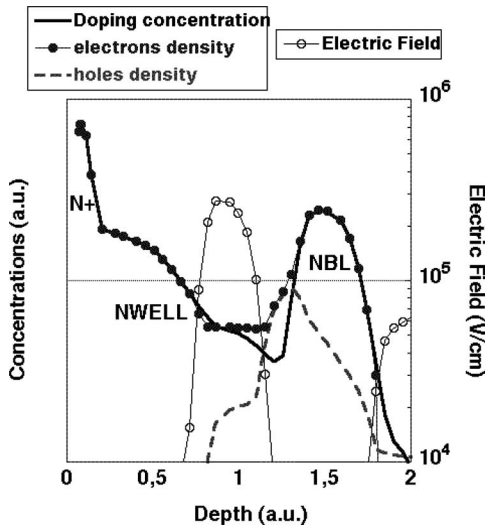

Fig. 15. Distribution of carriers and electric field maximum along BB′ cutline for a TLP current value $I_{\mathrm{TLP}} = 0.8$ A.

the effective junction N^-/NBL location. As the temperature is maximum at the location of concurrent maximum electric field and current, we can conclude that during TLP stress, the hot spot in the area under the emitter is always pushed toward the silicon-depth direction. The temperature is dissipated through the silicon volume that is more favorable for a best ESD robustness of the device.

As seen in the preview section, a second hot spot occurs for higher TLP current than the first one (Fig. 9). This hot spot is located in the area under the collector (BB′ cutline). Effectively, a high electric field builds up in the NWELL collector diffusion induced by the strongly doping gradient of this latter [17], [18]. With a TLP current increase, an intrinsic zone of carriers occurs and spreads toward the silicon surface (Fig. 15). The concurrent high electric field and large current in the collector area results in a local heating. This hot spot follows the maximum electric field and moves toward the silicon surface where temperature is more difficult to dissipate. Compared to the temperature evolution in the area below the emitter, this situation in the collector side is much more likely to evolve into a premature failure of the device. This failure scenario is validated by the failure analysis and confirms the results of the proposed predictive methodology.

5) Results Synthesis for the First Technology Devices: For the three devices, the comparison between I_{T2} issued from measurement and I_{T2} predicted with the simulation from Esmark's criteria and from the proposed predictive method is summarized in Table IV. We can notice for device2 that after a certain current level, there is a computation convergence problem noticed NA. In this case, it was impossible to apply Esmark's criteria because the simulation stopped before the occurrence of the second snapback or the melting-point temperature.

The I_{T2} value predictions obtained with the novel predictive methodology are more accurate for each studied device.

Another great advantage of our method is the computation time. Effectively, our predictive method can obtain I_{T2} value twice faster. The reason is that the simulation is stopped quite

TABLE IV
SYNTHESIS RESULTS FOR FIRST TECHNOLOGY DEVICES

Smart Power technology	0.35µm		
Device	1	2	3
Dimension (mm*mm)	25*80	40*100	50*100
I_{T2} measurement	1.5 A	3.1 A	3.8 A
I_{T2} conventional criteria simulated	1.8 A	NA*	4.3 A
I_{T2} with new method	1.52 A	3.3 A	3.5 A
Predict error with conventional criterion	20 %	NA*	14 %
Predict error with new method	2 %	11 %	8 %
CPU* Time for conventional criterion	7 days	NA*	5 days
CPU* Time for new method	4 days	11 days	3 days

* NA: Non Applicable

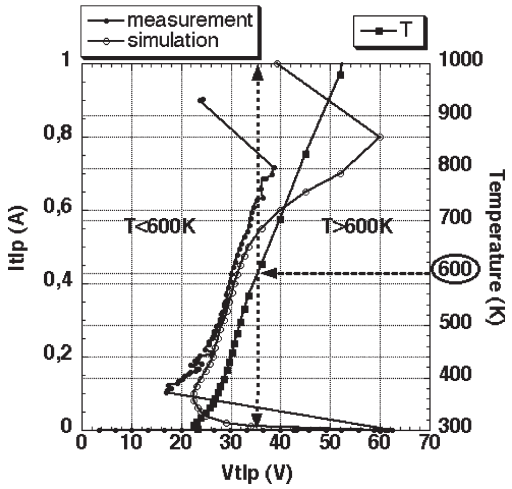

Fig. 16. TLP measurement and simulation (device4).

early before the failure current I_{T2} of the ESD structure occurs in opposition to Esmark.

A deprocessing was realized on device2 and device3, but the defect was so extent that it was impossible to determine the localization responsible of the failure initiation. We just observe a melted filament between collector and emitter.

6) Application to the Second Technology Devices: To demonstrate the validity and the generalization of the methodology, we applied it to a more advanced 0.25-μm Smart Power technology. Device4 (Fig. 2) is also a vertical n-p-n bipolar transistor used as ESD protection, and device5 (Fig. 3) is a p-n diode. In addition to dimension shrinking, this technology offers new features like deep trench isolation techniques. The doping profiles are also quite similar to the first technology ones.

Device4 is a preliminary design of this device, which was available on a test mask, and ESD robustness has not been optimized yet. For devices of this second technology, simulations were performed with a faster computer (3.6-GHz processor) than for the previous ones. Figs. 16 and 17 show the simulated

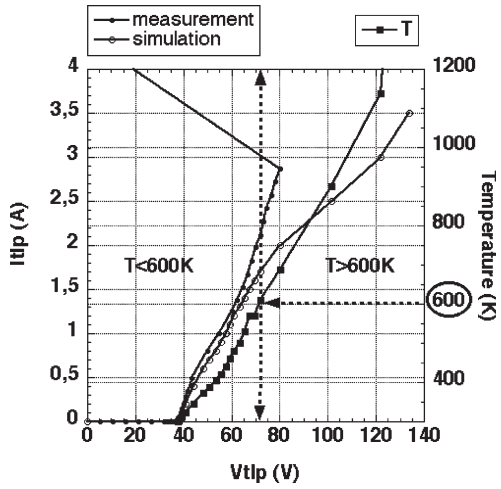

Fig. 17. TLP measurement and simulation (device5).

TABLE V
SYNTHESIS RESULTS FOR SECOND TECHNOLOGY DEVICES

Smart Power technology	0.25µm	
Device	4	5
Dimension (mm*mm)	35*60	40*200
I_{T2} measurement	0.72 A	2.8 A
I_{T2} conventional criteria simulated	0.8A	NA*
I_{T2} with new method	0.9 A	3.5 A
Predict error with conventional criterion	10 %	NA*
Predict error with new method	20 %	25 %
CPU* Time for conventional criterion	38 h	NA*
CPU* Time for new method	19 h	14h

* NA: Non Applicable

and measured TLP curves for bipolar transistor and diode, respectively. Up to 600 K, there is a good agreement between measurement and simulation. But for higher temperature, the simulated ON-resistance strongly increases, whereas the measured one keeps its linear behavior. This discrepancy is probably related to the physical model limits.

It has to be noticed that with the classical method, simulations are not able to provide a reliable result for both devices. Device4 simulation luckily provides a good I_{T2} prediction but exhibits an increase of the ON-resistance that is not seen in the real device. For device5 case, there is no snapback, and the simulator diverges before reaching the silicon melting point. TLP simulations are CPU time consuming. Indeed, each point of the TLP curves requires to run a 100-ns transient electrothermal simulation. The whole I–V curve could be obtained after several computation hours. These results emphasize the efficiency of the new method with respect to simulation time. As we only need the beginning of the TLP curve to predict current failure, the new method is twice faster (Table V).

IV. CONCLUSION

The conventional method to predict the ESD failure is not reliable, because the chosen failure criterion is a temperature that is outside the temperature validity range of physical models. Even if sometimes the current failure correlates experiment (device3), in many cases, there could be large errors or convergence problems.

The new proposed method always provides a result with an accuracy similar to the conventional method. Another benefit of this method is the simulation speedup with a computing time divided by two.

The proposed methodology is based on performing successive TLP simulations up to 600 K, which is the temperature limit for the validity of the physical models. For the prediction of the failure current, the main steps are the following:

1) localization of the hottest spot in the device for each TLP pulse;
2) extraction of the maximum values of impact ionization $G_{\rm i}$ and SRH recombination $R_{\rm SRH}$ rate at the hot spot for each TLP pulse;
3) exponential extrapolation of $G_{\rm i}(I_{\rm TLP})$ and $|R_{\rm SRH}|\ (I_{\rm TLP})$ curves until they intersect: corresponding TLP current value $I_{\rm TLP}$ defined as the failure-current value $I_{\rm T2}$.

It has to be noticed that the proposed method is aimed at detecting the occurrence of a thermal breakdown into a protection device, which represents the main cause of ESD failures.

REFERENCES

[1] M. R. Pinto, C. S. Rafferty, R. K. Smith, and J. Bude, "ULSI technology development by predictive simulation," in *IEDM Tech. Dig.*, 1993, pp. 701–704.
[2] H. Masuda *et al.*, "TCAD strategy for predictive VLSI memory development," in *IEDM Tech. Dig.*, 1994, pp. 153–156.
[3] M. Bafleur, T. Dinh, H. Park, R. Thoma, T. Zirkle, and A. Wild, "Concurrent process, device and integrated circuit development by predictive engineering for Smart Power technologies," in *Proc. Int. Conf. Model. Simul. Microsyst.*, 1998, pp. 414–419.
[4] *Handbook Chemistry and Physics*, CRC press, 56th Ed., 1975, pp. B1–B2.
[5] K. Esmark, "Device simulation of ESD protection elements," in *Microelectronics*, vol. 128. Konstanz, Germany: Hartung-Gorre Verlag Konstanz, 2002.
[6] A. Amerasekara and J. Seitchik, "Electrothermal behavior of deep submicron nMOS transistors under high current snapback (ESD/EOS)," in *IEDM Tech. Dig.*, 1994, pp. 446–449.
[7] S. Reggiani *et al.*, "Experimental extraction of the electron impact-ionization at large operating temperature," in *IEDM Tech. Dig.*, 2004, pp. 407–410.
[8] G. Charitat, "Modélisation et réalisation de composants planar haute-tension," *Doctorat d'Etat*, 1990, Toulouse, France.
[9] T. J. Maloney and N. Khurana, "Transmission line pulsing techniques for circuit modeling of ESD phenomena," in *Proc. EOS/ESD Symp.*, 1985, pp. 49–54.
[10] *Process and Device Simulation With ISE Tools Release 9.5*, 2004. [Online]. Available: http//:www.synopsys.com
[11] D. Trémouilles, G. Bertrand, M. Bafleur, N. Nolhier, and L. Lescouzères, "Design guidelines to achieve a very high ESD robustness in a self-biased NPN," in *Proc. EOS/ESD Symp.*, 2002, pp. 281–288.
[12] R. V. Overstraeten and H. D. Man, "Measurement of the ionization rates in diffused Silicon p-n junction," *Solid States Electron.*, vol. 13, no. 5, pp. 583–608, May 1970.
[13] J. W. Slotboom *et al.*, "Unified apparent band-gap narrowing in n- and p- type Silicon," *Solid State Electron.*, vol. 35, no. 2, pp. 125–129, 1992.
[14] H. A. Schafft and J. C. French, "A survey of second breakdown," *IEEE Trans. Electron Devices*, vol. 13, no. 8/9, pp. 613–618, Aug./Sep. 1996.

[15] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. New York: Springer-Verlag, 1984.
[16] C. T. Kirk, "A theory of transistor cutoff frequency (f_t) at high current densities," *IRE Trans. Electron Devices*, vol. ED-9, no. 2, pp. 164–174, Mar. 1962.
[17] G. Boselli, "On high injection mechanisms in semiconductor devices under ESD conditions," M.S. thesis, University of Twente, Enschede, The Netherlands, 2001.
[18] G. Notermans, "On the use of N-Well Resistors for uniform triggering of ESD protection elements," in *Proc. EOS/ESD Symp.*, 1997, pp. 221–229.

Christophe Salaméro was born in Béziers, France, in 1974. He received the M.S. degree in microelectronics engineering from the University of Paul Sabatier, Toulouse, France, in 2000. Since March 2002, he has been working toward the Ph.D. degree in cooperation with the Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS-CNRS) and Freescale Semiconductor, Toulouse.

His research activity focuses on the electrostatic discharge (ESD) robustness prediction of semiconductor devices through physical simulation.

Nicolas Nolhier was born in Pau, France, in 1966. He received the Ph.D. degree in electronics from the National Institute of Applied Sciences (INSA), Toulouse, France in 1992.

Since 1993, he has been a Lecturer with the University of Paul Sabatier, Toulouse, and has joined the Laboratoire d'Architecture et d'Analyse des Systèmes (LAAS-CNRS). Motivated by the rapid thermal processing control and modeling, his interests then moved to power-device design and simulations. In 1996, he initiated, in LAAS, the study of electrostatic discharge (ESD) protections using 2-D electothermal simulations and, investigated in TLP experiments. He is also active in power MOSFET modeling for RF applications.

Amaury Gendron was born in Rennes, France, in 1980. He received the degree from the Ecole Superieure d'electricite and received the M.S. degree in electronics from the University of Paris 6, Paris, France in 2003. Since October 2003, he has been working toward the Ph.D. degree in the R&D group of Freescale in cooperation with the Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS-CNRS).

His research activities focus on the development of reduced snapback ESD protection in Smart Power technologies.

Marise Bafleur (M'95–SM'01) received the Engineer degree in solid state physics from the Institut National des Sciences Appliquées of Toulouse, France, in 1979, and the Ph.D. and the State Doctorate degrees from Paul Sabatier University, Toulouse, in 1982, and 1987, respectively.

During her Ph.D. studies, she worked on defect characterization of molecular beam epitaxy grown GaAs layers. Since 1983, she has been with the Centre National de la Recherche Scientifique (C.N.R.S.) at the Laboratoire d'Analyse et d'Architecture des Systèmes (L.A.A.S.) in Toulouse. From 1983 to 1987, she devoted her research activities to CMOS digital circuit design and modeling of its dynamic behavior. From 1987 to 1994, she oriented her activity toward Smart Power IC's design and technology. From 1991 to 1994, she was a Project Leader in the PROMETHEUS/PROCHIP European program dedicated to the introduction of Smart Power in the automotive environment in cooperation with Renault and Peugeot automotive manufacturers. Then, she spent three years with Motorola in Phoenix (USA), first, as a low power Designer and, second, working on TCAD predictive engineering. Since the end of 1997, she has been with L.A.A.S. Research Institute leading a research activity in the field of electrostatic discharge (ESD) in CMOS and Smart Power technologies.

Patrice Besse was born in Poitiers, France, in 1975. He received the Ph.D. degree in electronics from the University of Paul Sabatier, Toulouse, France, in January 2004.

From September 2000 to November 2003, he developed new electrostatic discharge (ESD) protections for BiCMOS technologies in the R&D group of Motorola and in cooperation with the Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS-CNRS). Since April 2004, he has joined the Sensor & Analog design division of Freescale, Toulouse, where he is the Head of the ESD protections for automotive and power over ethernet applications. He is involved in physical simulations to improve the robustness of integrated circuits during ESD, especially for high-operating voltage and very high current stress level.

Michel Zécri received the M.Sc. degree in electrical engineering from the University of Aix-Marseille, France, in 1997 and the Ph.D. degree from the National Polytechnic Institute of Toulouse, France, in 2000.

In 2000, he joined Freescale Semiconductor, Toulouse, France, where he worked in the field of electrostatic discharge (ESD) protection for Smart Power technology. Since 2004, he has been with the Crolles Alliance (STMicroelectronics, Philips, Freescale) design services for Freescale Semiconductor Inc., where he manages a group responsible for developing IOs and ESD kits design in advanced CMOS technologies.