Process and Layout Dependent Substrate Resistance Modeling for Deep Sub-Micron ESD Protection Devices

Xin Y. Zhang, Kaustav Banerjee, Ajith Amerasekera*, Vikas Gupta*,
Zhiping Yu, and Robert W. Dutton
Department of Electrical Engineering, Stanford University, Stanford, CA 94305
*Silicon Technology Development, Texas Instruments Inc. Dallas, TX 75265
Tel: (650)723-1482 Email: xyz@gloworm.stanford.edu

ABSTRACT

This paper demonstrates a new methodology for bringing accurate substrate resistance modeling into circuit level ESD simulation. The impact of layout and process variations on the effective substrate resistance of deep sub-micron ESD devices is analyzed and modeled using a quasi mixed-mode approach. The substrate resistance simulated by this method shows good agreement with the values extracted from experimental data. This technique can be employed to simulate turn-on characteristics of ESD protection devices and determine the impact of process and layout variations on their reliability before fabrication of the actual devices.

I. INTRODUCTION

During electro-static discharge (ESD), the magnitude of substrate resistance determines the on/off state of the parasitic bipolar transistor (n^+-p-n^+ or p^+-n-p^+), which provides protection by forming a current path from the drain to the source and substrate. More importantly, the interactions of different circuit elements through the common substrate can have a significant impact on the circuit's ESD performance. Therefore, accurate modeling of the substrate resistance to capture the effects of layout and process is essential for performing accurate circuit level ESD simulation. Moreover, the fact that the substrate resistance becomes conductivity modulated due to the injection of minority carriers into the base after the turn-on of the parasitic BJT also needs to be modeled in order to simulate the substrate current correctly [1-2].

The effects of conductivity modulation can be seen from experimental data, which shows that the substrate current continues to increase after snapback; hence, to maintain a constant base voltage of the parasitic BJT, the substrate resistance must decrease [1]. Instead of explicitly modeling the dynamic substrate resistance, the standard approach is to model the substrate potential as a current controlled voltage source [2-4],

$$V_{sub} = R_{sub} \cdot I_{sub} - R_d \cdot (I_d - I_{ds}) \qquad (1)$$

as shown in Fig. 1, where the reduction of the substrate resistance due to conductivity modulation has been implicitly modeled. I_{sub} is the substrate current, I_d is the total drain current,

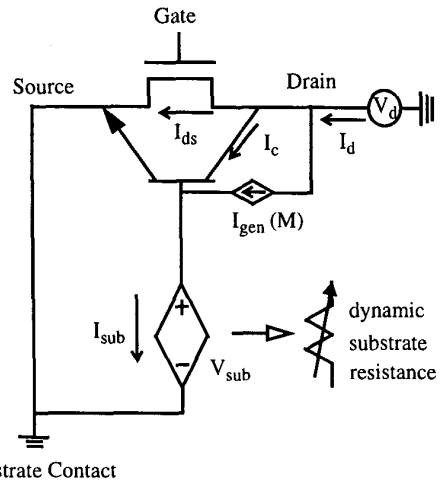

Fig. 1 Compact model for simulating MOSFET breakdown uses a current controlled voltage source to represent the dynamic substrate resistance during ESD stress.

and I_{ds} is the MOSFET's surface current under normal operating condition. R_{sub} and R_d are circuit model parameters that can be extracted from experimental data as illustrated in Fig. 2 [2]. R_{sub} is the substrate resistance at the on-set of the snapback, and it is extracted using the y-intercept,

$$R_{sub} = \frac{V_{beon}}{I_{sub0}} \approx \frac{0.8}{I_{sub0}} \qquad (2)$$

where V_{beon} is the turn-on voltage (~0.8V) of the base and emitter junction of the parasitic BJT and I_{sub0} is the y-intercept. R_d models the conductivity modulation by relating the substrate current (I_{sub}) to the total drain current (I_d), and it is extracted using the slope,

$$R_d = \frac{\Delta I_{sub}}{\Delta I_d} \cdot R_{sub} \qquad (3)$$

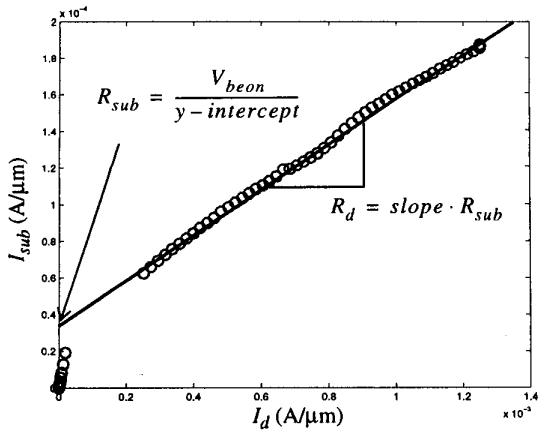

Fig. 2. The straight line fitting ($I_{sub} = slope \cdot I_d + y - intercept$) is a good approximation to extract R_{sub} and R_d parameters from I_{sub} vs. I_d plot, which is obtained at V_g=0V.

where $\frac{\Delta I_{sub}}{\Delta I_d}$ is the slope. The $R_d \cdot (I_d - I_{ds})$ term is negligible until after the snapback, after which, it begins to offset the $R_{sub} \cdot I_{sub}$ term to keep the substrate potential constant, emulating minority carrier injection into the substrate. This substrate potential model can estimate the reduction of the substrate resistance due to conductivity modulation; however, once extracted, the R_{sub} and R_d are fixed parameters that fail to predict the effects of layout and process variations, and they must be extracted again from experimental data for a different layout or process.

Other approaches employ substrate resistance networks to simulate the effects of different layouts [5-6]. Without accounting for conductivity modulation, they overestimate the substrate resistance and underestimate the substrate current, which could lead to inaccurate simulation results for the second breakdown as shown in Fig. 3. Hence, it is desirable to formulate a methodology that can account for conductivity modulation, process, and layout variations in substrate resistance modeling to allow for improved circuit simulation capabilities.

The aim of this work is to extend the capability of the current controlled substrate potential model (V_{sub} model) using the quasi-mixed-mode (qmm) method so that the R_{sub} and R_d parameters can be simulated for different layouts and processes based on a few calibrated devices, instead of extracting them from experimental data. This allows the accurate modeling of the substrate resistance due to the effects of layout and processes, and enables the circuit designers to simulate and design an effective protection device based on layout and identify critical current path during the ESD stress. In this paper, the qmm method is calibrated against a single, fixed-geometry device with one doping profile. The calibrated doping profile is used as a basis to simulate R_{sub} and R_d parameters of other

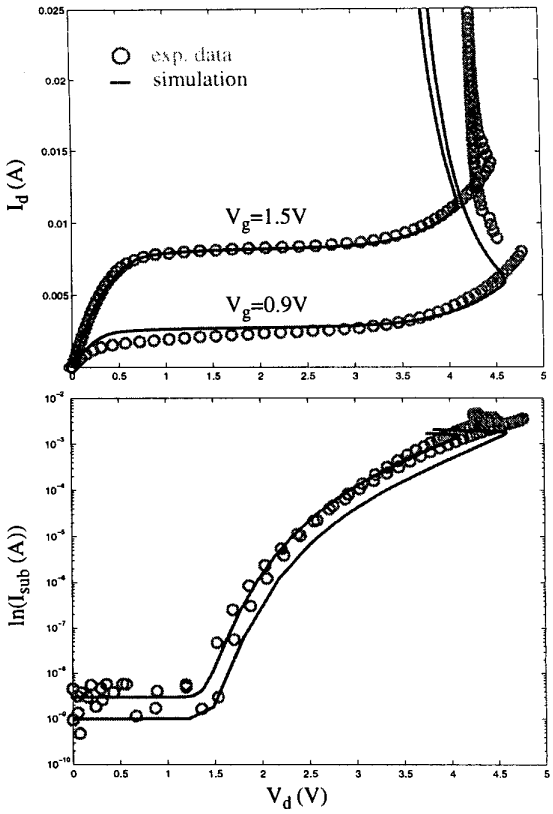

Fig. 3 The ESD I-V curve (solid line) is obtained using a constant substrate resistance ($V_{sub} = R_{sub} \cdot I_{sub}$) without considering conductivity modulation. The discrepancy between the simulated I_{sub} results and the experimental data (circles) is significant.

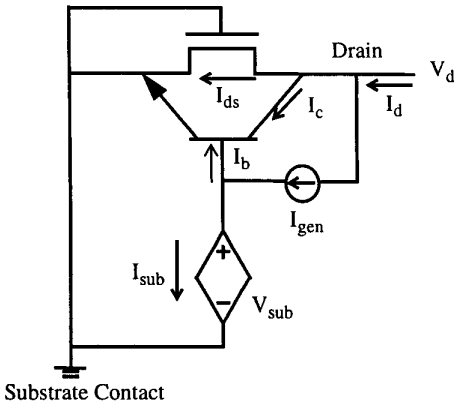

Fig. 4 This is the schematic representation of hole injection (I_{gen}) into the structure for the device simulation part of the qmm. It is equivalent to sweep I_{gen} while fixing V_d.

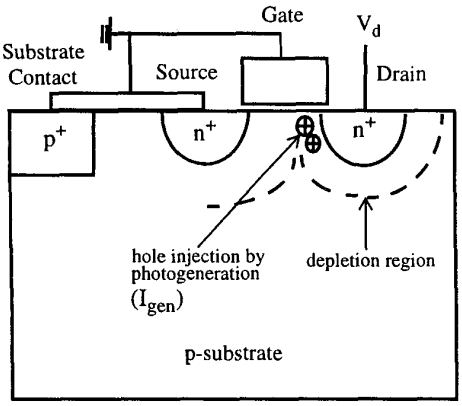

Fig. 5 With the boundary conditions established, the full device simulation can be greatly simplified by using photogeneration function to replace hole generation by impact ionization. This corresponds to steps 2 and 3 of Fig. 6.

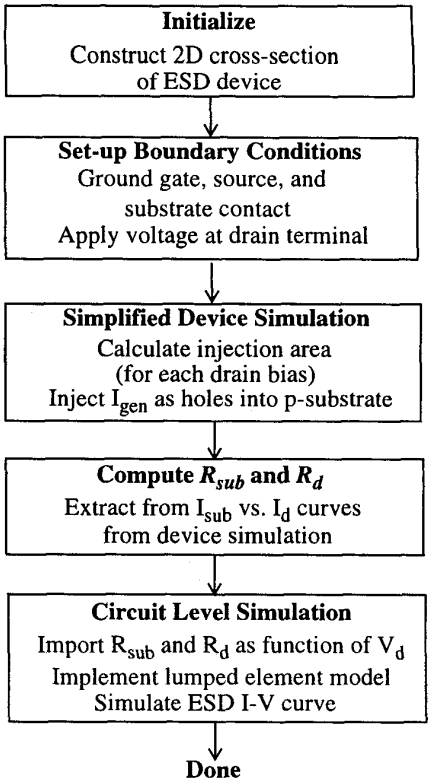

Figure 6. The flow diagram illustrates the system level set-up of the quasi-mixed-mode model.

devices with different layouts (i.e. channel length and source to substrate contact spacing variations) and p-well doping. The simulated R_{sub} and R_d are compared to the extracted R_{sub} and R_d, showing the accuracy of the qmm method.

The qmm methodology is briefly described in section two; in section three, the qmm approach is applied to model the substrate resistance of on-chip ESD devices with layout and process variations. The extracted substrate resistance parameters are discussed along with their significance in section four, and followed by the conclusions in section five.

II. QUASI-MIXED-MODE APPROACH

The quasi-mixed-mode model is a marriage between device and circuit simulation. It differs from the traditional mixed-mode (device/circuit) simulation by not using a fully coupled matrix approach. The qmm utilizes the circuit model and device simulation to model lumped and distributed elements respectively [7].

For a given technology, the process dependent parameters do not vary once extracted, so the physical effects can be modeled as lumped elements; hence, the impact ionization model (M) parameters are implemented directly in the compact model along with the parameters that govern normal MOSFET operation. On the other hand, the substrate resistance parameters tend to depend on layout; therefore, it is better suited to use distributed element modeling. The device simulator computes the substrate resistance based on layout. The compact model takes the simulated substrate resistance, and simulates the resulting ESD I-V curve.

The qmm was developed with the purpose of modeling substrate resistance of the protection device, and it is much faster, robust and easier to calibrate compared to the full device simulation. In addition, the substrate resistance parameters are able to account for layout and process variations which extends the capabilities of the circuit model described in the introduction. Hence, the qmm can be used as an effective tool in designing the optimal ESD devices without building and testing them on silicon.

The speed improvement, robustness, and ease of calibration are possible because the device simulation part of the qmm does not simulate with impact ionization model. The qmm breaks the feedback loop caused by impact ionization: instead of stressing the drain terminal as in normal simulation (Fig. 1) or measurement set-up, the drain voltage is fixed while sweeping the I_{gen} until the parasitic BJT is on as shown in Fig 4. This is equivalent to the snapback region, where the drain voltage is roughly constant, while the currents change quite dramatically. This process of sweeping I_{gen} can be accomplished by proper placement of the boundary conditions in device simulation [7].

Fig. 5 illustrates the placement of boundary conditions on the ESD devices. The gate, source, and substrate contacts are all tied to the ground, and the drain terminal is biased to establish the corresponding electric field and depletion area. The holes are injected into the depletion region that has the highest electric field around drain junction, much like the mechanism of hole generation by impact ionization. The hole injection is achieved using photogeneration function. For the I_{gen} sweep, it can be shown that during snapback when the

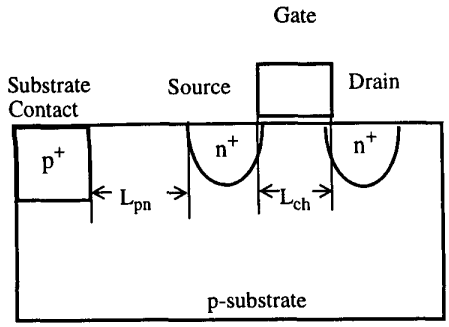

Fig. 7 ESD devices have two different types of layouts: changing L_{ch} (channel length) with fixed L_{pn} (source to substrate contact space), and changing L_{pn} with fixed L_{ch}.

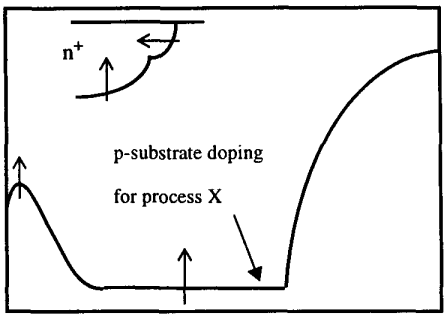

Fig. 8 The LDD and S/D junction depth and lateral diffusion ratio along with p-substrate doping are scaled in the direction of the arrows by the ratio of $\frac{P-well_{doseY}}{P-well_{doseX}}$.

parasitic BJT is on at every V_d with V_g=0,

$$I_{gen} = I_b + I_{sub} \quad (4)$$

$$I_d = I_c + I_{gen} \quad (5)$$

where I_c, I_{gen}, I_d, I_b, *and* I_{sub} are as shown in Fig. 4 and 5.

The information flow of the quasi-mixed-mode model is described in Figure 6. To begin with, a 2D cross section of the ESD device is constructed using the device simulator. Then boundary conditions are imposed on the device that allow the holes to be injected into the silicon substrate as described earlier. The placement of boundary conditions and the execution of device simulations are automated by using computer scripts. After running device simulations, a set of substrate current (I_{sub}) vs. drain current (I_d) curves under different drain bias are obtained. The values of substrate resistance parameters (R_{sub} and R_d) can be extracted from these curves as a function of drain bias (V_d) according to eq. (1) and (2), and imported into the compact model as [7],

	Process X		Process Y		
Devices	L_{ch}	L_{pn}	L_{ch}	L_{pn}	Devices
A	0.21	2.5	0.21	2.5	G
B	0.21	10	0.21	10	H
C	0.18	2.3	0.18	2.3	I
D	0.21	2.3	0.21	2.3	J
E	0.25	2.3	0.25	2.3	K
F	0.30	2.3	0.30	2.3	L

Table 1. Devices A-L have the same layout dimensions, aside from the different dimensions listed. All the dimensions listed in the Table 1 are in μm. Process X differs from process Y only in p-well dose. Devices A-F are fabricated using process X, and devices G-L are fabricated using process Y.

$$V_{sub} = R_{sub}(V_d) \cdot I_{sub} - R_d(V_d) \cdot (I_d - I_{ds}) \quad (6)$$

for circuit level simulation. The compact model shown in Fig. 1 is implemented inside the circuit simulator. The circuit parameters for normal MOSFET operation and impact ionization are already extracted from experimental data according to previously published research [2-4,8].

III. CALIBRATION AND SIMULATION OF SUBSTRATE RESISTANCE

In this work, the qmm is applied to model the substrate resistances of ESD devices fabricated by state-of-the-art CMOS technology with two different p-well dopings. Process X has a lower p-well doping than process Y. These ESD devices are all 20μm wide with varying gate lengths (L_{ch}) or source to substrate contact spacings (L_{pn}), as illustrated in Fig. 7. The exact layout dimensions along with the process information of each device are listed in Table 1.

An analytical doping profile for process X is tuned until the simulated R_{sub} and R_d parameters fit the experimental parameters of device A. No additional change is made to the doping profile and model coefficients after this calibration. The calibrated doping profile was used to predict the substrate resistance of devices (B-F) fabricated under the same process X. To simulate the effect of p-well doping variation on R_{sub} and R_d parameters, the doping profile for process Y was generated by simply scaling the doping profile for process X according to the ratio of the two p-well doses,

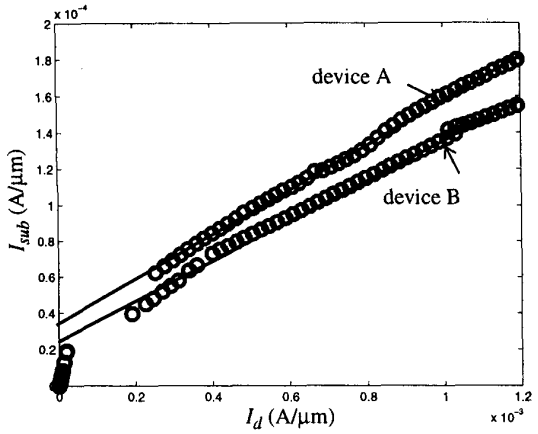

Fig. 9 Experimental I_{sub} vs. I_d curves obtained with gate grounded for devices A&B showing the impact of increasing the distance of substrate contact to source contact (L_{pn}) on the magnitude of R_{sub} and R_d parameters.

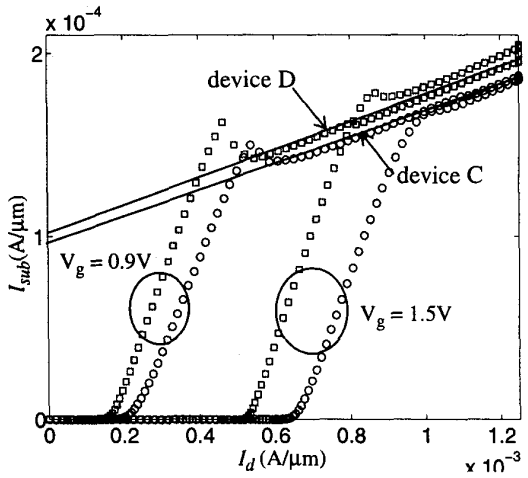

Fig. 11 Experimental I_{sub} vs. I_d curves for devices C (plotted in circles, L_{ch}=0.18μm)) and D (plotted in squares, L_{ch}=0.21μm) show the impact of increasing channel length (L_{ch}) on the magnitude of R_{sub} and R_d parameters.

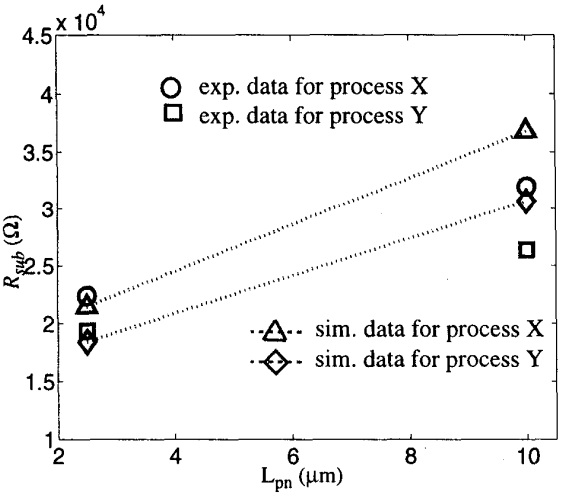

Fig. 10 The resistance values plotted in circles and squares are extracted from experimental data of devices A&B and G&H fabricated using process X and Y, and the resistance values plotted in triangle and diamond shapes with dotted lines are extracted from the simulation results of A&B and G&H. The error between the simulated and experimental data is ~15%.

as shown in Fig. 8.

The device simulation part of qmm was performed using MEDICI, and the circuit simulation part was done using HSPICE. After performing the qmm simulation as described in the previous section, the R_{sub} and R_d parameters were extracted from the simulation results. The experimental R_{sub} and R_d parameters for devices A&B were extracted based on eqs. (2) and (3) as shown in Fig. 9, and the R_{sub} and R_d parameters of devices G&H were extracted using the same method [2-4,7]. The predicted R_{sub} values obtained using qmm method are plotted against the extracted R_{sub} as shown in Fig. 10.

The experimental R_{sub} and R_d parameters for devices C&D were extracted as shown in Fig. 11, and similarly the R_{sub} and R_d parameters of devices E&F and I-L were extracted using the same method. The predicted R_{sub} and R_d values obtained using qmm method are plotted against the extracted R_{sub} and R_d as shown in Figs. 12 and 13 for process X and Y respectively.

IV. EFFECTS OF LAYOUT AND PROCESS VARIATIONS

From the extracted experimental values in Fig. 10, it can be observed that as the distance from the source to substrate contact (L_{pn}) increases, the substrate resistance (R_{sub}) becomes larger. This is due to an increase in the effective substrate area. However, the slope ($\Delta I_{sub}/\Delta I_d$) remains the same as shown in Fig. 9 because the property of the parasitic BJT has not been altered by changing the L_{pn}.

In addition to R_{sub} increase due to L_{pn}, the simulation results also captured the fact that R_{sub} decreases as the p-well doping increases for process Y. The error is ~15% for R_{sub} values of devices B and H of which 3% is propagated by the calibration error from device A.

The current flow contours are plotted for both devices A and B in Fig. 14. It indicates that as the substrate contact

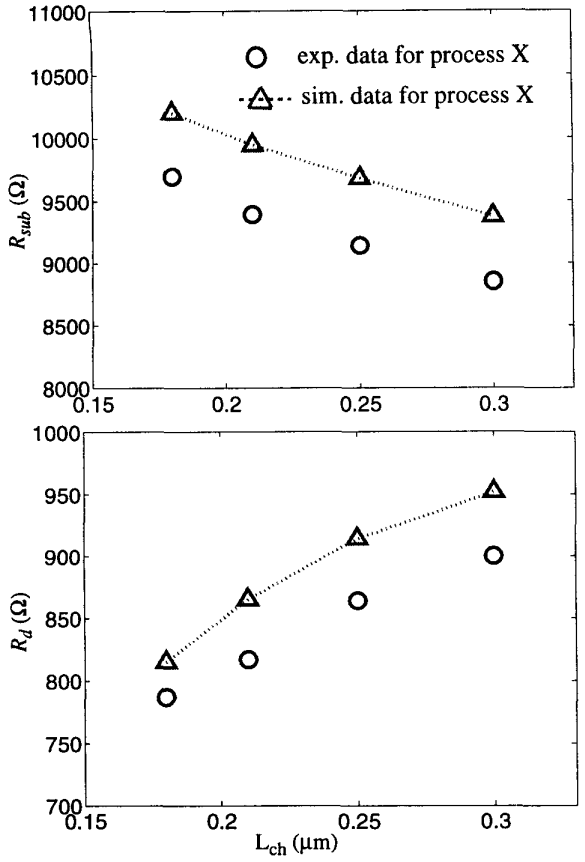

Fig. 12 The resistance values plotted in circles are extracted from experimental data of devices C-F fabricated using process X, and the resistance values plotted in triangles are extracted from the quasi-mixed-mode simulation results. The maximum error between the measured and experimental data does not exceed 6%.

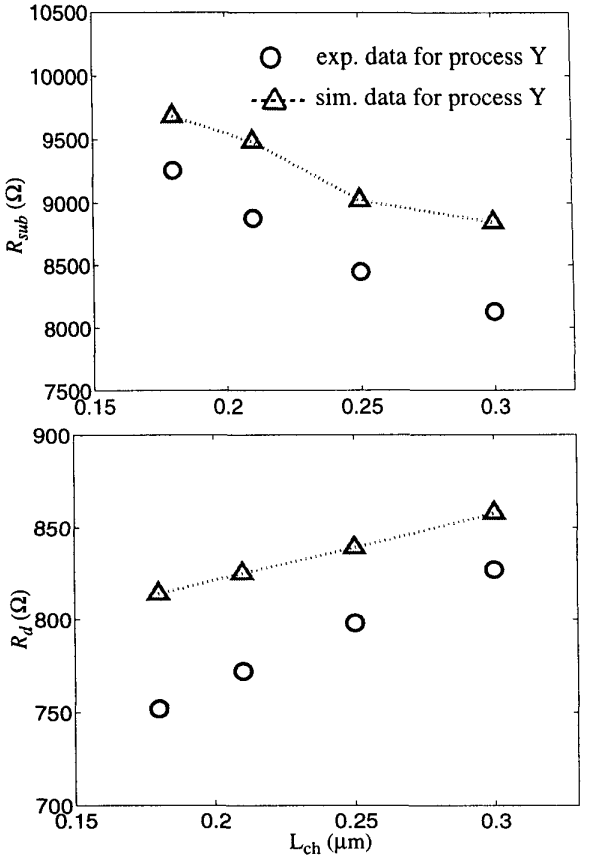

Fig. 13 The resistance values plotted in circles are extracted from experimental data of devices I-L fabricated using process Y, and the resistance values plotted in triangles are extracted from the quasi-mixed-mode simulation results. The maximum error between the measured and experimental data does not exceed 9%.

moves further away from the NMOS (from 2.5μm to 10μm), the current flow path also becomes more spread out, and more current flows deeper through the bulk of the substrate (P+ substrate), such as devices B and H. And this could also explain the larger percentage of the simulation error for devices B and H (15%) compared to devices A and G (5%) because the initial calibration done for device A did not accurately calibrate the doping of P+ substrate, since I_{sub} of device A did not flow as deep through the substrate. It is well known that the current flows along the least resistive path, and in this case, the current path is mainly determined by the location of the substrate contact; hence, the spreading resistance determines the R_{sub} values. The qmm approach was able to take this nonlinear effect due to layout into account when modeling the substrate.

In this case, two experimental data points of R_{sub} (Fig. 10) per process is not enough to find the influence of L_{pn} on R_{sub} for design analysis; therefore, additional structures are simulated using the qmm method. The simulated R_{sub} parameters are plotted against corresponding L_{pn}s as shown in Fig. 15. From the plot, it is clear that as the substrate contact is moved further from the NMOS (larger L_{pn} values), R_{sub} values do not increase as rapidly since most of the substrate current is flowing through P+ part of the substrate. The precise analytical and physical relationship between L_{pn} and R_{sub} still needs to be formulated with the aid of more simulations and experiments.

According to Figs. 12 and 13, the experimental data and the simulation results for both processes demonstrate that R_{sub} decreases and R_d increases as channel length (L_{ch}) increases, which shows that the simulation results are in good agreement with experimental data. As the p-well doping increases for process Y, the magnitude of simulated R_{sub} values in Fig. 13 also decreased compared to that of Fig. 12 (process X). The percentage errors for devices C-F and I-L, which all have different L_{ch}, are less than 9%. The error could be caused by the inaccuracies in the 2-D doping profile.

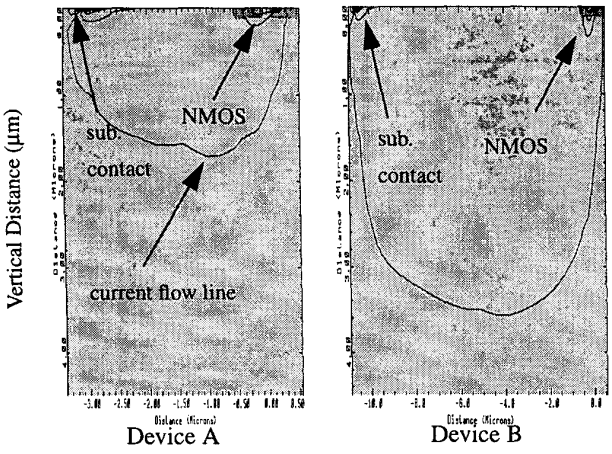

Fig. 14 The current flowlines are plotted for devices A&B. The distance from NMOS to substrate contact is 2.5μm for device A and 10μm for device B. Most of the current still flows near the surface for both devices, the remaining current spreads out much deeper (3.6 μm vs. 1.7 μm) for device B than A.

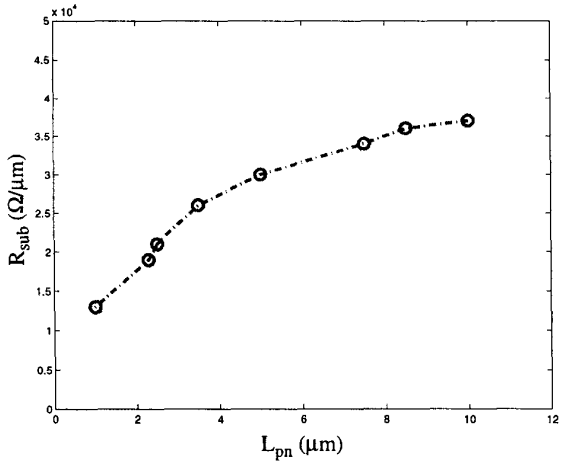

Fig. 15 R_{sub} values between 1-10μm L_{pn}s are obtained using the qmm method. As L_{pn} increases beyond 4μm, there is a decrease in the slope of the curve, showing a reduction in incremental R_{sub} value.

As illustrated in Fig. 2 and eq. (2), R_{sub} is the substrate resistance at the turn-on of the parasitic BJT (i.e. $I_c = I_b = 0$). At that point, eq. (4) reduces to,

$$I_{gen} = I_{sub} \tag{7}$$

and the injected hole current (I_{gen}) flows from the drain junction towards the substrate contact. And the substrate spreading resistance (R_{sub}) can be estimated by [9],

$$R_{sub}/\rho = \frac{1}{2(W - L_{ch})} \ln\left(\frac{W(L_{ch} + 2T)}{L_{ch}(W + 2T)}\right) + \frac{X_j + X_d/2}{W(L')} \tag{8}$$

$$L' \propto L_{ch} \tag{9}$$

where ρ is the substrate resistivity, T is the substrate depth, X_j is the junction depth, X_d is the depletion width at the drain junction, w is the channel width, and L_{ch} is the channel length. The exact relation between L' and L_{ch} in eq. (9) depends on the hole current distribution inside the substrate, and it can be obtained by empirically fitting calculated R_{sub}s (eq. (8)) to the experimental values [9]. Eqs. (8) and (9) demonstrate that the magnitude of R_{sub} decreases with channel length.

On the other hand, R_d (slope) increases because the β of the parasitic BJT decreases more rapidly as the drain current increases for shorter channel length as shown in Fig. 16. It also can be shown that during snapback [3,8,10],

$$\beta \cdot (M - 1) \geq 1 \tag{10}$$

$$I_{gen} = (M - 1) \cdot (I_c + I_{ds}) \tag{11}$$

$$\beta = \frac{I_c}{I_b} \tag{12}$$

As β decreases due to high current injection for both the short and long channel devices such as devices C and F, M increases for both as governed by eq. (10) to keep the parasitic BJT on; therefore, I_{gen} increases, but $I_{genC} < I_{genF}$ because $\beta_C > \beta_F$ at each I_d according to Fig. 16 [3]. I_b also increases as β decreases, but the rate of increase for I_{bF} is less than that of I_{bC} because the rate of decrease of β_F is much less than that of β_C; hence, the slope of the I_{sub} vs. I_d curve decreases as the sharp drop in β requires more I_b (eq. (12)) from I_{gen} for shorter channel device.

The beta degradation in both devices C and F is caused by high current injection after the parasitic BJT turns on. During this process, the injection of electrons into base is sufficient to cause a significant increase in the hole concentration in the base, thereby reducing the collector current (I_c). The base charge under high current injection can be described by [10],

$$\int_0^{x_B} p(x)dx = \int_0^{x_B} \left[N_a(x) + n'(x)\right]dx \tag{13}$$

$$x_B \propto L_{ch} \tag{14}$$

where $\int_0^{x_B} N_a(x)dx$ is the gummel number (the initial built in base charge caused by the processing of the transistor), x_B is the width of the quasi-neutral region of the base, and $n'(x)$ is the injected electron concentration. Based on eqs. (13) and (14), device F with larger L_{ch}, has a larger gummel number than device C, resulting in a lower β as shown in Fig. 16, but for incremental β change (β degradation) during high current injection, the device with a larger gummel number is less affected by the injected electrons than the device with a smaller gummel number; hence, the β degrades at a slower rate for longer channel device.

The magnitude of R_{sub} and R_d are very sensitive to the process and layout changes as can be seen from Figs 10, 12, and 13, and that reflects in the shapes of ESD I-V curves. For example, in order to quantitatively predict the effects of increasing L_{ch} from 0.18μm (device C) to 0.30μm (device F) on the ESD I-V curve without fabricating device F, R_{sub} and R_d from device C had to be used. Using the qmm method, R_{sub} and R_d for device F can be predicted resulting in a more accurate ESD I-V curve as shown in Fig. 17.

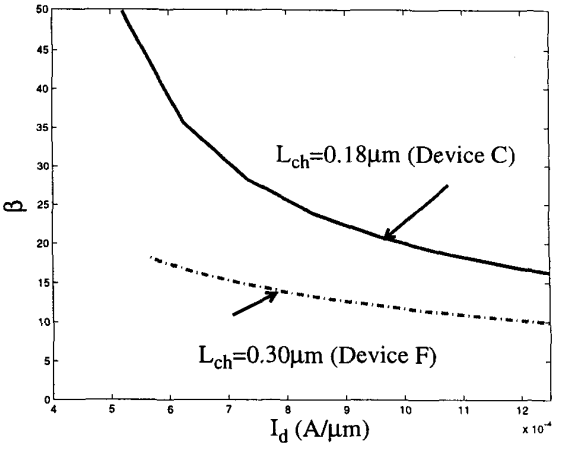

Fig. 16 β for devices C and F are extracted from the qmm simulation results. β of device C decreases at a much faster rate than device F; hence, the percent increase in I_b is larger as the I_d increases for device C.

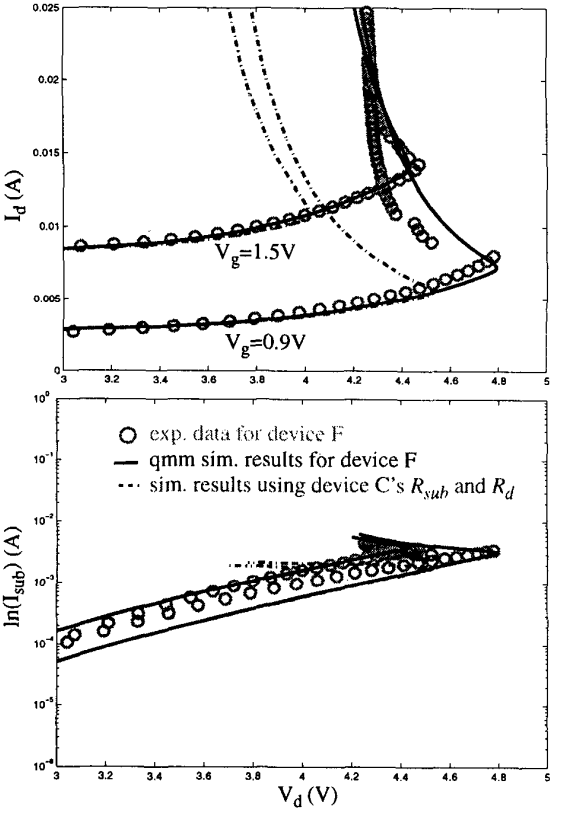

Fig. 17 The ESD I-V curve of device F (solid line) is simulated using the values of R_{sub} and R_d obtained from the qmm method, and the simulation results matches the experimental data of device F (circles). The dashed line shows discrepancy between the simulated curve and the experimental data using the R_{sub} and R_d of device C. The discrepancy will get larger as the layout becomes more different.

V. CONCLUSIONS

In conclusion, we have demonstrated a quasi-mixed-mode approach to accurately model the substrate resistance parameters, R_{sub} *and* R_d, across different layouts and processes for deep sub-micron ESD devices. This methodology can be used to evaluate different layout and technology options for optimizing the performance of ESD protection devices before fabrication. It can be extended to model the common silicon substrate of the protection circuit to capture the substrate coupling effects between different circuit elements and also define the turn-on process for multi-finger devices.

ACKNOWLEDGEMENT

The authors would like to thank Sridhar Ramaswamy and Charvaka Duvvury from TI for giving valuable suggestions and comments. Finally the authors gratefully acknowledge the support from SRC and TI.

REFERENCES

[1] T. Skotnicki, G. Merckel, and A. Merrachi, "New Physical Model of Multiplication-Induced Breakdown in MOSFETs", *Solid State Electronics*, vol. 34, p.1297-1307, 1991.

[2] S. Ramaswamy, A. Amerasekera, and M. Chang, "A

Unified Substrate Current Model for Weak and Strong Impact Ionization in sub-0.25μm NMOS Devices", *International Electron Device Meeting*, p.885-888, 1997.

[3] A. Amerasekera, V. Gupta, K. Vasanth, and S. Ramaswamy, "Analysis of Snapback Behavior on the ESD Capability of Sub-0.20μm NMOS", *Proc. of 37th IEEE International Reliability Physics Symposium*, p.159-166, 1999.

[4] X. Zhang, Z. Yu, S. Beebe, and R. Dutton, "Substrate Resistance Model for Simulating MOSFET Breakdown in ESD Protection", *Techcon*, 1998.

[5] T. Li, T. Tsai, E. Rosenbaum, and S. Kang, "Substrate Resistance Modeling and Circuit-Level Simulation of Parasitic Device Coupling Effects for CMOS I/O Circuits under ESD Stress", *Proc. of 20th EOS/ESD Symposium*, 1998.

[6] R. Troutman and M. Hargrove, "Transmission Line Modeling of Substrate Resistances and CMOS Latchup", *IEEE Transactions on Electron Devices*, p.945-954, 1986.

[7] X. Zhang, Z. Yu, and R. Dutton, "A Quasi-Mixed-Mode MOSFET Model for Simulation and Prediction of Substrate Resistance", *Proc. of International Conference on Simulation of Semiconductor Processes and Devices*, p.211-4, 1999.

[8] A. Amerasekera, S. Ramaswamy, M. Chang, and C. Duvvury, "Modeling MOS Snapback and Parasitic Bipolar Action for Circuit-Level ESD and High Current Simulations", *Proc. of 34th IEEE International Reliability Physics Symposium*, p.318-326, 1996.

[9] F. Hsu, P. Ko, S. Tam, C. Hu, and R. Muller, "An Analytical Breakdown Model for Short-Channel MOSFET's", *IEEE Transactions on Electron Devices*, vol. ED-29, p.1735-1740, 1982.

[10] R. Muller and T. Kamins, *Device Electronics for Integrated Circuits*, p.323, Willey, New York, 1986.