VLSI DESIGN
2001, Vol. 13, Nos. 1–4, pp. 103–109
Reprints available directly from the publisher
Photocopying permitted by license only

3D Modelling of Fluctuation Effects in Highly Scaled VLSI Devices

THOMAS D. LINTON JR.*, SHAOFENG YU and REAZ SHAHEED

Intel Technology CAD Division, 2200 Mission College Blvd., Santa Clara, CA

Fluctuation effects are becoming important in advanced VLSI devices because of their increasing impact on circuit performance and chip yields. Accurate modelling of these effects generally requires full 3D simulation, which is used here to analyse four of the primary such effects. Polysilicon line edge roughness causes excess device leakage, which can be reduced at the cost of decreased performance. Phase-shift mask defects can reduce current drive and increase capacitance. Random dopant fluctuation, which causes variation in threshold voltage, is evaluated for three technology generations and it is shown that proper tip scaling can reduce these variations. Finally, a study of alpha particle strikes evaluates the effectiveness of SOI in improving soft error reliability.

Keywords: Fluctuation; 3D; Device simulation; MOSFET

INTRODUCTION

Continued VLSI device scaling is increasing the importance of statistical fluctuation phenomena. This is because decreasing size increases the relative size of fluctuations at the same time that the larger number of devices on a chip raises the number of devices which are in the "tail" of the device distribution and which therefore show large deviations from the nominal characteristics.

This paper addresses four types of device variation that are becoming important: MOS gate line edge roughness (LER), phase-shift mask defects (MD), random dopant fluctuation (RDF), and soft error reliability (SER). All the studies were carried out in Intel's SONATA device simulation scripting environment using Intel's PRIME structure editor and the 3D device simulator DESSIS from Integrated Systems Engineering.

LINE EDGE ROUGHNESS

The lithographic process used to define polysilicon MOSFET gates yields rough gate sidewalls due to reticle roughness, polymer molecular weight variations, exposure shot noise, oxide sputtering variations, etch variations, and other variations. This roughness results in a variation of MOS channel length across the width of the device. Channel length variations are a problem because

*Corresponding author. Tel.: +1 408 765 0419, Fax: +1 408 765 9322, e-mail: tom.linton@intel.com

shortened parts of the channel will have excessive off-state leakage and lengthened parts will have degraded drive current. The study here quantifies the effect of LER on device performance using the method of [1].

LER Assessment Methodology

LER was modelled by a square wave shape of amplitude λ on the drain side of the device (Fig. 1). This simple structure permits variation of LER in a controlled manner which can easily be analyzed. In the device simulations, a half-period of the square wave is modelled; reflecting boundary conditions make the simulation act as one slice out of a periodic structure. The devices simulated had standard source/drain and tip dopings, no halo dopings, and flat wells. The well dopings were adjusted to achieve an off current of 3 nA/um. Two device sizes were analyzed: (1) LGATE = 100 nm (LMET = 60 nm) and (2) LGATE = 80 nm (LMET = 40 nm). In the simulations performed here, the width of the short half of the channel is kept the same as that of the long half, unless otherwise noted.

Effect of LER on performance is determined in the following way: (1) the nominal device without LER is simulated to get base on-current ION (VG = VD = VCC); (2) off-current IOFF (VG = 0, VD = 1.1*VCC) of the LER device is calculated from simulation; (3) the LER device average gate length is increased until its IOFF matches the leakage criterion of 3 nA/um; (4) ION is calculated for the adjusted device and compared against the base ION of step (1) to determine performance degradation.

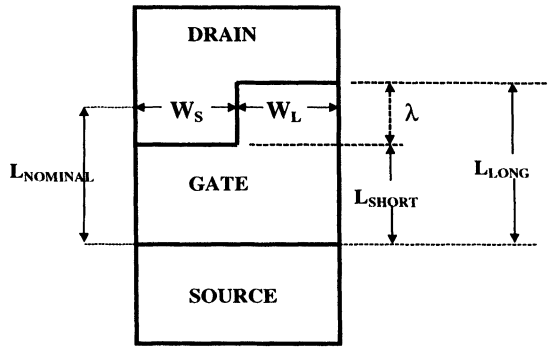

FIGURE 1 Square-wave model of LER with amplitude λ.

LER Performance Degradation

The dependence of IOFF on LER amplitude is shown in Figure 2 for a device of width (half-cycle of LER waveform) 100 nm. In this figure there are plots of normalized IOFF for the short half of the channel, for the long half of the channel, for their arithmetic average, and for the actual IOFF from the 3D simulations. Increasing λ decreases the length of the short half, and Figure 2 shows that IOFF current increases rapidly here, as expected. Increasing λ lengthens the long half and reduces its IOFF. Results of the 3D simulation in Figure 2 also show that total IOFF increases with λ, as one

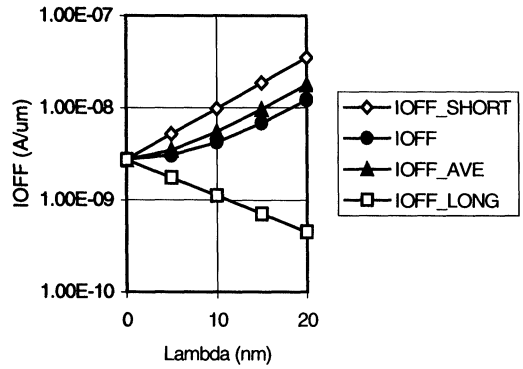

FIGURE 2 Off-current (IOFF) *versus* LER amplitude.

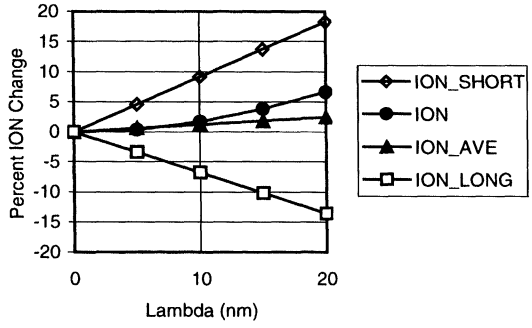

FIGURE 3 On-current (ION) *versus* LER amplitude.

TABLE I Performance degradation in adjusted LER devices

L_{GATE}	λ	$L_{NOMINAL}$ adjustment	I_{ON} change	C_{GATE} change	Performance = Δ (I_{ON}/C_{GATE})
100 nm	20 nm	7.5 nm	−3.6%	+7.5%	−10.3%
100 nm	15 nm	3.9 nm	−2.3%	+3.9%	−6.0%
100 nm	10 nm	2.4 nm	−1.5%	+2.4%	−3.8%
100 nm	5 nm	0.48 nm	−0.5%	+0.48%	−0.98%
80 nm	12 nm	3.5 nm	−5.4%	+4.4%	−9.4%

would expect, since the current through the short part of the device will increase faster than the current through the long half will decrease, as a consequence of the superlinear dependence of IOFF on inverse channel length. The average of the currents from the short and long parts of the channel is higher than the actual total current. This implies that the 3D interaction between the short and long parts of the channel is reducing IOFF from what would be predicted by a simple model with a short and long transistor in parallel. In other words, 3D effects are reducing the negative impact of LER from what a simple model would suggest.

The dependence of ION on λ, shown in Figure 3, is similar to that of IOFF except that the dependence is essentially linear instead of exponential. Also, the average ION is less than what 3D simulation predicts. Here the simple model of a short and long device in parallel underestimates the actual current. In addition, simulations of ION at different device widths show that LER perturbation effects extend as far as 50 nm in the width direction.

Analysis of LER Effects

Results of the performance analysis are shown in Table I. For LGATE = 80 nm, 12 nm LER requires a gate length increase of 3.5 nm, resulting in 5.4% ION degradation, 4.4% increase in gate capacitance, and net performance degradation of about 9.4%.

These results may be used to help set lithographic tolerances. They also show that, because of the superlinear dependence of IOFF on inverse gate length, LER "notches" are much worse than protrusions. Because 3D effects extend out to at least 50 nm, excess leakage will be particularly severe if two notches are closer than that distance. Finally, the high cost in capacitance due to gate lengthening suggests other approaches to mitigating LER, such as dopant level adjustment, should be explored.

Additional simulations were performed with LER positioned at the source instead of the drain edge of the gate. These produced essentially the same results as with LER at the drain edge. The effect of LER at both edges, both in and out of phase, is worthy of further study.

PHASE-SHIFT MASK DEFECTS

Phase-shift masks (PSMs) are increasingly being used to define minimum-size features, and with their use has come the need to develop tools for PSM inspection. What resolution is required in a PSM inspection tool? To help answer this question, a study was performed of the impact of PSM defect size on device performance.

Relevant PSM defects are additive; they cause excess material in the patterned layer to be left behind. PSM defects are therefore modelled with a square tab at the drain end of a MOSFET (Fig. 4). Here the interest is in a single isolated defect, unlike in the LER study, and therefore the effect of the reflecting boundary conditions in the device simulator is eliminated by using wide devices. Degradation of ION was determined as a function of tab size for devices with LGATE = 80 nm. IOFF was not considered because the tab only improves IOFF.

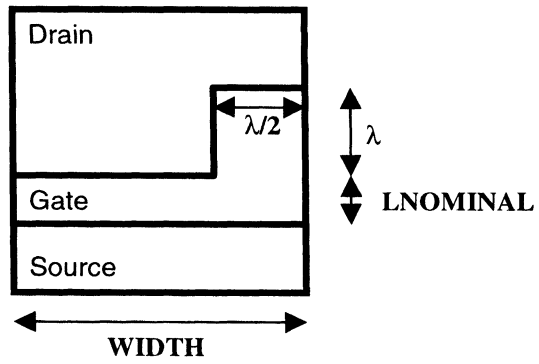

FIGURE 4 Model of phase-shift mask defect as a squart tab on the drain.

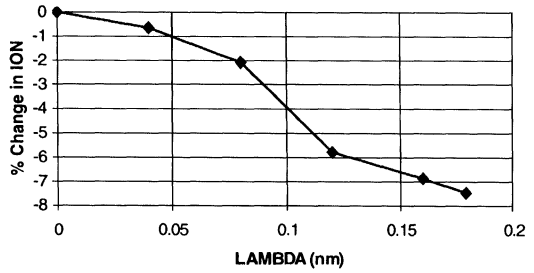

FIGURE 5 On-current (ION) degradation *versus* PSM defect size for a device of width 640 nm and length 80 nm.

Figure 5 shows ION degradation as a function of defect size for a device of width 640 nm and LGATE = 80 nm. A defect of size 80 nm by 80 nm reduces ION only by 2.1% and gate capacitance increases by about 13%. A narrower device, of width 250 nm, undergoes a reduction of ION by 5.4% for an 80 nm defect and a reduction of 10% for a 100 nm defect. Gate capacitance of the narrow device will increase by approximately 32% for the 80 nm defect and 50% for the 100 nm defect. 2D simulations approximating the channel as an L-shaped resistive region verified that even a large "tab" at the drain has only a small impact on resistance.

The conclusion is that defects must be large relative to channel length to significantly degrade device drive current. However gate capacitance, which loads the driving transistor, is increased significantly at these dimensions, especially for narrow devices.

RANDOM DOPANT FLUCTUATION

Random dopant fluctuation (RDF) is a variation in the number of placement of MOSFET dopant atoms due to the stochastic nature of ion implantation and diffusion [2]. It is an inherent variation that would be present even if the manufacturing process were perfectly controlled. The impact of RDF generally increases with scaling because smaller channels translate into a smaller number of dopant atoms, and therefore greater relative variation in their number. Doping levels generally increase with scaling but this increase in the density of atoms tends to be more than offset by the smaller channel size. RDF must be simulated in 3D because channel currents can flow under or around channel impurity atoms.

RDF is important because it causes variations in both IOFF and ION. IOFF variations can increase chip power or can impact circuit density and performance if less aggressive scaling is employed to meet leakage criteria. Variations in ION may degrade circuit speed by reducing drive currents of critical path transistors.

An atomistic simulation [3, 4] approach was used with full 3D single-carrier device simulation of 500 fluctuated devices for each device structure under study.

RDF Results and Analysis

Figure 6 shows the effect of RDF on threshold voltage for minimum length and width devices for three technology generations. Lengths and dopings are scaled appropriately except that all three technologies here use the same tip depth and concentration. The figure shows that variations get worse with technology scaling and for lower target threshold voltages.

Proper scaling of the tip doping profiles reduces sensitivity to RDF, as shown in Figure 7. Here, the variation in threshold voltage of the 0.1 um technology has been reduced to a level comparable to that of the previous generation technologies by making the tips shallower. When the tips are

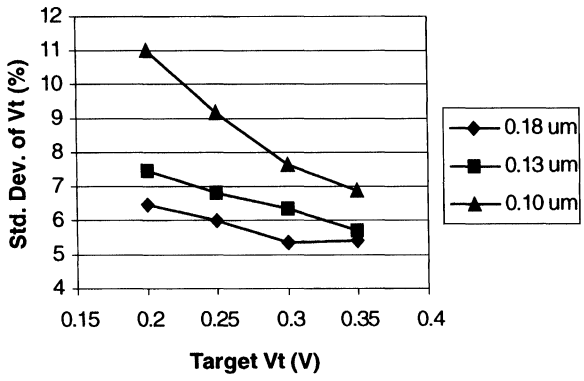

FIGURE 6 Relative V_t variation for 3 technology generations.

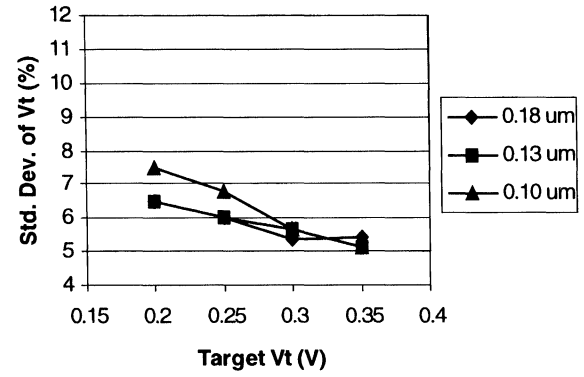

FIGURE 7 Relative V_t variation for 3 technology generations using scaled tips.

changed, the well doping and channel length are also re-optimized to achieve the leakage current requirements. With this change in the doping profiles, the number of atoms in the channel region is a bit lower for the shallow-tip device, but not enough to account for the improved resistance to RDF. The improvement in scaled tip devices can be explained by their reduced channel depletion depth [5]. This mechanism will be dealt with in a future publication.

It appears, then, that incorrectly scaled devices are likely to suffer from severe device characteristic variations due to RDF at the 0.1 um level and below. Proper design of the device structure, based on insights gleaned from 3D device simulation, can mitigate this problem.

SOFT ERROR RELIABILITY

Alpha particles, which come from packaging materials, and neutrons from background radiation generate a burst of undesired charge if they hit a silicon device [6, 7]. This charge disturbance can upset the logic state of a gate, resulting in soft error reliability (SER) problems. The effect of SER becomes more severe with scaling because smaller devices work with smaller amounts of charge at logic nodes and therefore a given amount of SER charge has a larger relative effect.

One device design that has been proposed to reduce SER effects is the silicon-on-insulator (SOI) MOSFET. This section therefore compares the sensitivity of SOI and bulk silicon devices to SER effects. The devices had gate lengths of 0.15 um, tips, halo dopings, and retrograde wells. The SOI devices had a silicon thickness of 0.2 um.

Alpha particles were injected with normal incidence to the silicon at a variety of locations along the length of the devices. Trench edges and the effects of a strike as a function of distance from the trench edge were not studied.

Alpha Strike Results

Figure 8 shows two distinctive differences between the response of SOI and bulk devices to alpha strikes. First, the peak current for the bulk device is significantly higher (about 40%) than for the SOI device. Second, the bulk device current quickly drops back to a low level within about

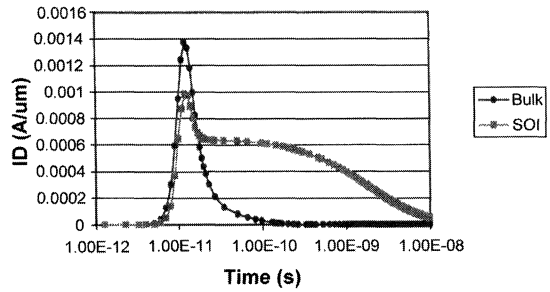

FIGURE 8 Drain current *versus* time in bulk and SOI devices.

40 pSec, whereas the SOI device has a long tail where the current stays at significant levels for a much longer time, on the order of 10 nSec. The total amount of charge collected by the drain is also much greater for the SOI device.

Figure 9 shows how SOI drain current changes with strike location. The peak current is greatest when the alpha particle strikes in the middle of the channel and decreases as the strike location moves toward the drain or source (see also Fig. 10). Also, there is a distinctive difference between strikes on the source side and drain side. Strikes near the drain show a high initial current spike whereas strikes near the source do not. Strikes near the drain also show a larger residual tail current.

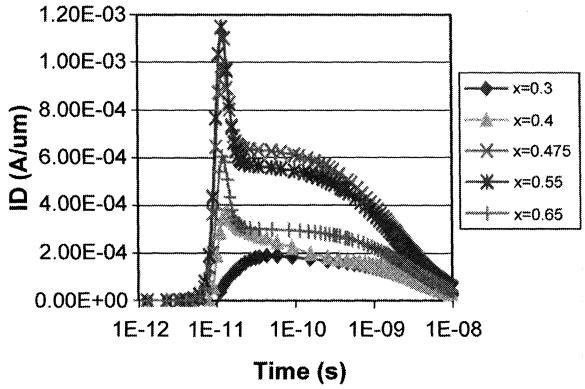

FIGURE 9 Drain current response in SOI as a function of strike location. The edges of the gate are at $x = 0.3$ um and $x = 0.55$ um. The source is at $x < 0.3$ um and the drain is at $x > 0.55$ um.

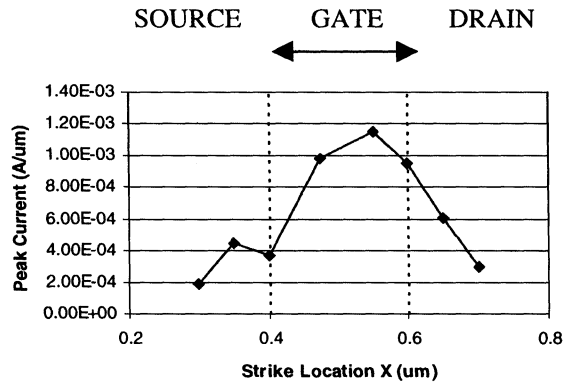

FIGURE 10 Peak drain current in SOI as a function of strike location.

SER Analysis

In a bulk MOSFET, an alpha strike generates charge that drifts under the influence of the source and drain depletion fields and is collected by the drain or substrate. Charge is collected relatively quickly and current soon returns to its original low value. In a partially-depleted SOI device, some of the charge is generated in the low field region near the insulator underneath the device. This charge takes a relatively long time to dissipate through recombination and drift processes. It induces the long tail of current in the SOI devices both through the slow collection process and through an alteration of the SOI threshold voltage.

Alpha particle-induced drain current in SOI devices depends strongly on the strike location. Simulation show that alpha particle strikes in the source or drain regions have essentially no effect because of the high doping in these regions. Strikes under the gate have the initial pulse and long tail for the reasons described above. As the strike location moves toward the drain, the response has the same shape, but its magnitude is less because more of the charge is induced in the drain, where it does not contribute to the current. As the strike location moves toward the source, the initial pulse of current disappears because none of the charge is generated in the high-field region near the drain; the primary effect is to generate charge in low-field regions where it is collected slowly. The peak current and total charge collected drop as the strike location moves toward the source because more and more of the charge is generated inside the source.

The effect of an alpha particle strike on an SOI channel can be more severe than that for a bulk device, but an SOI device is only sensitive to alpha particle strikes in the channel. Bulk devices can collect charge in isolation regions or under sources and drains; this charge also causes significant current. Therefore, for a given layout, a bulk

process may have a far larger area that is susceptible to alpha particle strikes and therefore the sensitivity of the circuit can be significantly larger. There are also circuit considerations that determine how important the initial pulse is compared with the tail and thereby alter the assessment of susceptibility to alpha strikes.

CONCLUSION

Four phenomena that cause variations in scaled VLSI devices have been examined: line edge roughness, mask defects, random dopant fluctuation, and alpha particle strikes. The studies show that each of these phenomena is significant but may be addressed through proper device design. Three important lessons may be drawn from these studies: the first is that fluctuation phenomena get stronger in scaled devices. RDF, for example, is changing from a second-order to first-order effect. The second lesson is that scaling increases the importance of 3D effects. Devices are becoming so small that perturbations in structure of charge affect device operation not just in the length direction but in the width direction as well. Whereas inherent active device behavior for wide devices may still be studied with 2D simulations, the effects studied here all require 3D simulation. The final lesson is a consequence of the previous one: 3D device simulation is becoming increasingly important. Even though the effects analyzed here are strong, they are extremely difficult and costly to measure experimentally. Accurate 3D simulation, however, provides a tool that can be used to understand and address them.

References

[1] Linton, T., Giles, M. and Packan, P., "The Impact of Line Edge Roughness on 100 nm Device Performance", 1999 Silicon Nanoelectronics Workshop, Kyoto, Japan.

[2] Keyes, R. W., "Effect of Randomness in the Distribution of Impurity Ions on FET Thresholds in Integrated Electronics", *Journal of Solid-State Circuits*, August, 1975, pp. 245–247.

[3] Wong, H.-S. and Taur, Y., "Three-Dimensional 'Atomistic' Simulation of Discrete Random Dopant Distribution Effects in Sub-0.1 um MOSFET's", *1993 International Electron Devices Meeting Technical Digest*, pp. 705–708.

[4] Asenov, A., "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 um MOSFET's: A 3-D 'atomistic' simulation study", *IEEE Transactions on Electron Devices*, **45**(12), December, 1998, 2505–2513.

[5] Vasileska, D., Gross, W. J. and Ferry, D. K., "Modeling of Deep-Submicrometer MOSFETs: Random Impurity Effects, Threshold Voltage Shifts and Gate Capacitance Attenuation", *Extended Abstracts of 1998 Sixth International Workshop on Computational Electronics*, Osaka, Japan.

[6] May, C. and Woods, M. J., "Alpha Particle-Induced Soft Errors in Dynamic Memories" *IEEE Trans. Electron Devices*, January, 1979, **ED-26**, 2–9.

[7] Srinivasan, G. R. (1996). "Modeling the Cosmic-Ray-Induced Soft-Error Rate in Integrated Circuits – An Overview", *IBM J. Res. Develop.*, **40**, 77–89.