

Latchup prevention by using guard ring structures in a 0.8 μm bulk CMOS process

Felipe Coyotl Mixcoatl¹, Alfonso Torres Jacome
Instituto Nacional de Astrofísica, Óptica y Electrónica
 Luis Enrique Erro #1, Tonantzintla, Puebla, México, C. P. 72840
 Tel: 266-31-00, ¹Ext: 1123, Cub: 1113, ²Ext: 1418, 1425

Latchup is a parasitic effect in CMOS technology, this is a PNP parasitic structure formed by at least two coupled bipolar transistors. When a transitory voltage/current overshoot/undershoot at a input/output node occurs, PNP structure can be turned on and a low impedance path between V_{DD} and V_{SS} can be formed. This low impedance state can produce either a momentary or a permanent loss of circuit functioning. In modern processes there are several techniques to reduce latchup in CMOS, some of them are the use of an epi layer, retrograde wells, SOI, deep trenches, etc. to spoil or decouple the parasitic bipolar transistors. Although such techniques can solve the latchup problem, they increase the cost of production. Thus, the bulk CMOS processes are still used. In this work an analysis of the latchup phenomena with the use of guard ring structures in order to prevent the latchup effect is made. Several structures have been analyzed with SSUPREM4 and S-PISCES with and without guard rings. The results show an increase in the holding voltage when guard ring structures are used. These structures will be fabricated in a 0.8 μm twin-well silicided bulk CMOS process in the Microelectronics Laboratory at INAOE.

Keywords: Latchup; Avalanching; Punchthrough; Guard ring structures

1. Introduction

CMOS technology was initially proposed in 1962, however, this was not applied until the 1980's decade for his low velocity and larger area consumption. At same time, CMOS requires more complicated fabrication process than NMOS technology and special careful in the latchup prevention, a destructive, parasitic, and intrinsic effect to CMOS. To understand the latchup phenomena, it is necessary to know the triggering mechanisms and the techniques to either minimize or eliminate this effect. In this work a summary of basic triggering mechanisms and some techniques to minimize latchup are presented. Then a design technique to decouple bipolar transistors by using guard ring structures in a twin-well bulk CMOS process in SSUPREM4 and S-PISCES is presented.

2. Problem description

In a CMOS circuit at least one parasitic PNP and at least one parasitic NPN bipolar transistors, normally in blocking state, with a very high impedance, are present. The formation of the main two parasitic bipolar transistors (one lateral and another vertical) connected in positive feedback loop, is shown in Figure 1a. In Figure 1b a concentrated simple model of the PNP parasitic structure with associated well and substrate resistances is shown too.

Latchup phenomena can be described as follow (throughout the work we consider a twin-well process over a lightly doped p-type substrate, although in this case P-well is analogous to P-substrate). In presence of some either current or voltage overshoot/undershoot in a I/O pad, some electrons(holes) can be injected to the P-well(N-well). Un-recombined electrons(holes) are collected in the well(substrate) by the reverse biased P-well/N-well

junction to form a majority carriers current. If this current is large enough, it can produce a ohmic drop in a parasitic emitter/base junction by R_W (R_S) which can turn on the VPNP (LNPN) transistor. Then, the collector current of VPNP (LNPN) injected to the base of LNPN (VPNP) can form a positive feedback loop which can generate high currents and burn up or cause either a momentary or a permanent loss of circuit functioning. There are several triggering latchup modes and the down scaling produce better parasitic bipolar transistors [1, 2, 3, 4].

The I-V characteristic of high and low impedance states are shown in Figure 2. Two key points are identified, (V_S , I_S) and (V_H , I_H), I_S marks the transition from high impedance region to the negative differential resistance region and I_H marks the transition from negative differential resistance region to the low impedance region. Usually, I_S and V_S are known as triggering (or switching) current and voltage, respectively, whereas I_H and V_H are known as holding current and voltage, respectively. In most cases $V_H > V_{DD}$ implies a fairly robust CMOS technology against latchup. If this relation is made sure, do not worry how many and how long can be the transient pulses of either triggering current or voltage, the biasing voltage can not support latchup condition when the excitation disappear. On the other hand, if $V_H < V_{DD}$, it is possible that the low impedance state can be maintained after that the excitation is removed.

In general terms, to induce latchup several conditions should be satisfied:

1. Bipolar loop gain should be larger than one [2].

$$\beta_{\text{NPN}}\beta_{\text{PNP}} > 1 + \frac{(\beta_{\text{NPN}} + 1)(I_{\text{Rs}} + I_{\text{Rw}}\beta_{\text{PNP}})}{I_{\text{DD}} - I_{\text{Rs}}}$$

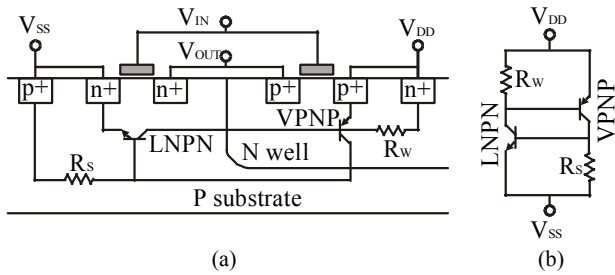


Figure 1. Latchup in CMOS. (a) A CMOS inverter with two parasitic bipolar transistors, (b) First order model with concentrated elements.

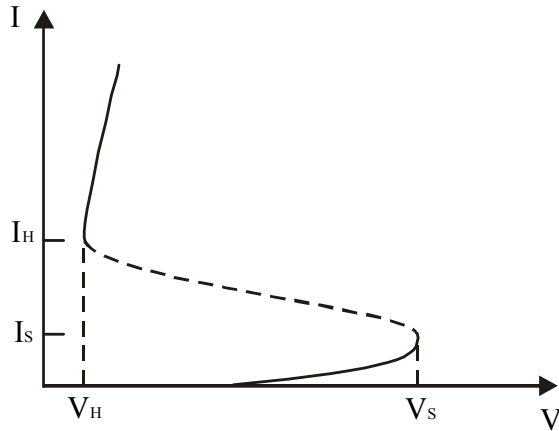


Figure 2. I-V typical PNP curve.

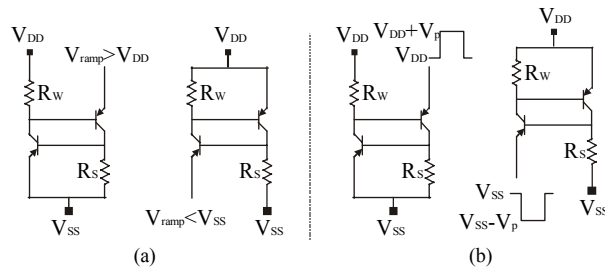


Figure 3. Two characterization modes, (a) V_{ramp} larger/smaller than V_{DD} , (b) transitory overshoot/undershoot pulses.

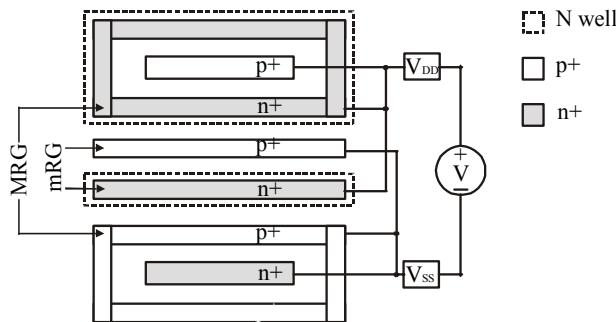


Figure 4. Guard ring structures for majority carriers (MGR), and for minority carriers (mGR).

2. In order to the lateral and vertical bipolar transistors can be turned on and carry a current high enough to reach the critical commutation value, a high enough bias power should exist.
3. Power supply voltage and associated circuitry should be capable of supply both a current at least equal to I_S , such that the PNP structure can move since the high impedance state to the negative differential resistance, and a current at least equal to I_H to reach the low impedance state.

3. Triggering modes

There are a lot of modes to induce latchup, some of them are summed up here in a few words. For a better comprehension refer to the excellent works in latchup phenomena by Troutman [1], and Deferm [3].

3.1 Overshoot/Undershoot

When voltage in a p^+ I/O node is higher than a fixed N-well voltage, holes are injected to the N-well. Then they are collected by the reverse biased P-substrate/N-well junction and produce a ohmic drop in the P-substrate. If ohmic drop is high enough, it can forward bias the n^+ /P-substrate junction and turn on the LNP transistor. If no protection techniques have been taken, the parasitic SCR would be turned on.

3.2 Avalanching

P-substrate/N-well junction should be inversely biased, if this voltage increases considerably, an avalanching current across R_S and R_W can be present. For a process without either epitaxial layer or buried layers, conduction is on the surface.

3.3 Punchthrough

If a n^+ diffusion is near to N-well, an increase in inverse bias of P-substrate/N-well junction can join the depletion regions with voltages smaller than avalanching voltages. As the N-well voltage increases, the depletion regions can spread into closely spaced n^+ diffusions, and a punchthrough current across N-well can turn on the VPNP transistor.

3.4 Parasitic field devices

In field thick oxide regions exist parasitic MOS transistors formed by interconnection lines over field thick oxide. If voltage in the lines and trapped charges is high enough to invert a field region, a parasitic channel across two regions with the same polarity can be formed (for example a parasitic channel across n^+ and N-well and a parasitic channel across p^+ and P-substrate), which can inject current across R_W or R_S and cause voltage drops that forward bias a parasitic emitter/base junction.

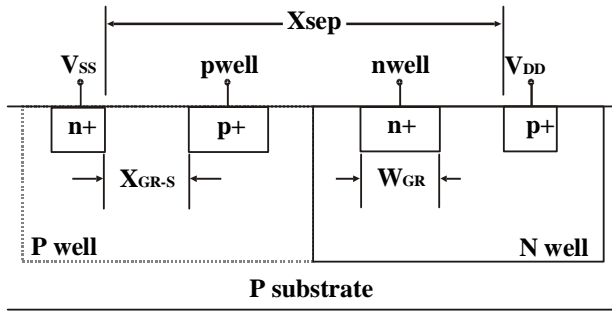


Figure 5. The three variables used in the analysis of latchup.

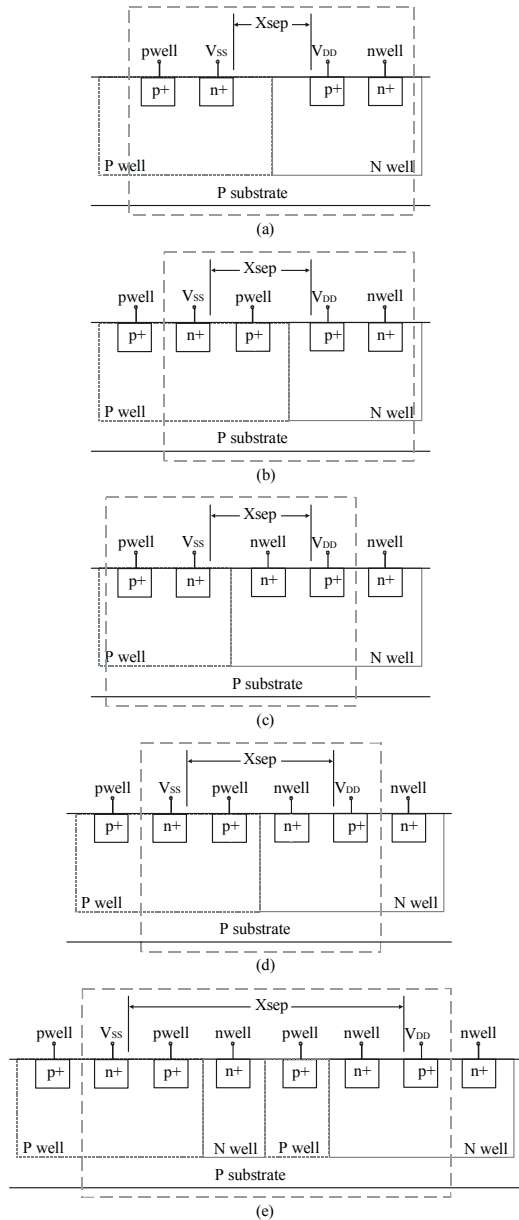


Figure 6. Test structures for simulation in SSUPREM4 and S-PISCES (a) without guard rings, (b) with P-MRG, (c) with N-MRG, (d) with P-well and N-well MRG, and (e) with both P-and-N MRG and floating (not biased) P-and-N MRGs.

There are some other triggering modes, some of them are radiation sources which generate hole-electron pairs and a current flow in a inversely biased junction, such as P-substrate/N-well [1], but they are out of the scope of this work.

4. Characterization

One of the transitory triggering modes is the power supply voltage ramp at a constant rate. The resultant displacement current flow across R_W and R_S resistors and junction capacitance in N-well/P-substrate can cause latchup if the change rate of the ramp is over a critical value even though the final voltage is smaller than the triggering voltage.

A second dynamic triggering mechanism is transitory undershoot/overshoot. Transitory pulses are directly applied to n^+ and p^+ diffusions. The result is a momentarily forward biased n^+ /P-substrate (p^+ /N-well) junction which creates a current flow that turn on the another bipolar transistor and so latchup can be induced. The direct bias pulse width in the emitter/base junction is approximated to an asymptotic limit and increases quickly for narrow pulses. As the pulse width decreases the high required to induce latchup increases, whereas when the pulse width increases a quasi-static condition corresponding to forward biasing of emitter/base junction is reached. Relation between pulse high and width changes with spacing between n^+ and p^+ [4]. Both characterization modes, for voltage ramp and overshoot/undershoot pulses are shown in Figure 3.

5. Preventions Methods

In reference [1] the latchup prevention methods are referenced and two general strategies to avoid latchup are mentioned, bipolar spoiling and bipolar decoupling. At least one of these are followed in latchup protection techniques. In the first the fabrication process is modified by adding or by changing the order in the steps of fabrication process in order to reduce the bipolar gain ($\beta_{PNP} = 10-100$, $\beta_{NPN} = 2-5$), for instance, gold doping, neutron irradiation, a base retarding field, and Schottky barrier source/drain. In the second one bipolar transistor is prevented of turning on the another bipolar transistor by using either layout or processing techniques. Layout examples include butted contacts and majority and minority carrier guard rings. Process examples include a lightly doped epitaxial layer on a highly doped substrate, retrograde wells and trench isolation. In trying to spoil the bipolar transistors a lot of damages and spoiling are induced in MOS devices. So the most efforts are devoted to decoupling bipolar transistors. Here a division in layout techniques and process techniques related to bipolar decoupling is made.

5.1. Layout Techniques

Guard ring structures. Guard rings are diffusions which decouple the parasitic bipolar transistors. There are two guard rings structure types, guard rings for minority carriers and guard rings for majority carriers connected to V_{DD} and V_{SS} metal lines, as shown the Figure 4. The guard rings for minority carriers are used to collect minority carriers before they are collected in an inversely biased P-substrate/N-well junction and flow as majority carriers. The guard rings for majority carriers decouple the bipolar transistors minimizing voltage drops created for majority carriers current. However, those are of the same diffusion-type that P-substrate or N-well and serve to reduce the sheet resistance.

Multiple well/substrate contacts. Majority carriers current flow across well and substrate whose brings voltage drops that can turn on a bipolar transistor. So it is necessary minimize the resistance to reduce voltage drop. This can be done putting multiple well/substrate contacts each 5-10 transistors or each 25-100 μm .

Butted source contacts. Butting two diffusions n^+ and p^+ in just one contact reduce the emitter/base resistance. Those contacts are used to fix both the substrate and well potentials in parasitic emitter vicinity and to reduce carriers injection in the substrate reducing lateral bipolar gain. Butting substrate and well to its corresponding source contacts in a typical CMOS inverter increases I_H and V_H values.

5.2. Process Techniques

There are several techniques to reduce or eliminating latchup by increasing either the substrate or the well conductivity and reduce R_S or R_W . For example, lightly doped epitaxial layers over highly doped substrates, buried layers and/or retrograde wells. Other techniques decouple bipolar transistors making the current flow path longer across V_{SS} to V_{DD} , by instance trench isolation and silicon on isolating. However, these techniques imply a higher cost of production and require more expensive equipment. For example, to do either buried layers or retrograde wells a high energy implanter is required.

Therefore, to obtain a cheap and competitive fabrication process, bulk CMOS can be used with guard ring structures as protection against latchup. Additionally, process improvements will make this a good competitor to more complicated and expensive technologies.

6. Simulation results

Considering the high cost of both epitaxial layers and retrograde wells, a latchup free bulk CMOS process can be obtained with guard rings. As the dimensions in CMOS structures are down scaling, the V_{SS} to V_{DD} separation (X_{sep}) should decrease, but when (X_{sep}) increases the effectiveness of CMOS devices decreases. Otherwise, when (X_{sep}) is scaled, the parasitic bipolar transistors effectiveness is

increased and a larger separation between V_{SS} to V_{DD} is needed. Then if (X_{sep}) increases, the layout area consumption will be larger and some avoiding latchup technique of section IV should be taken. In this section the simulation results of a simple twin-well bulk CMOS process with and without guard rings in SSUPREM4 and S-PISCES are presented. Applying a positive voltage ramp in electrode V_{DD} larger than 5 V, the holding voltage show an increase when guard ring are used and X_{sep} increases. The analysis of this effect was attacked since three defined variables for five structures shown in Figures 5 and 6.

1. The increase in V_{SS} to V_{DD} separation (X_{sep}).
2. The distance of $pwell$ and $nwell$ to V_{SS} and V_{DD} , respectively (X_{GR-S}).
3. The guard ring diffusion width (W_{RG}).

All five structures were simulated in SSUPREM4 and S-PISCES by using different arrays as is shown in Figure 6. The bias of the structures is as follow: V_{SS} and $pwell$ are biased to 0 V, $nwell$ is polarized to 5 V, and a voltage ramp larger than 5 V is applied in V_{DD} in order to forward bias the p^+/N -well junction. If conditions of section II are accomplished at least one parasitic bipolar transistor can be turned on and a latchup condition can be reached.

As was seen in section V-A, majority carriers guard ring (MRG) reduces the sheet resistance in either the well or the substrate because R_S and R_W are low. Then, if the V_{SS} to V_{DD} separation increases, the holding voltage increases too as is shown in Figure 7. As the Figure 7 shows, the use of majority carriers ring guard in N-well is more effective than guard ring in P-well by his higher doping concentration and his lower resistance. The use of both majority carriers N and P guard ring structures rise V_H faster than the use of only either P or N guard ring structure. And the use of not biased minority carriers guard ring structures in Figure 6e do not show an improvement (as one can expect) when a comparison is made with the results of Figure 6d, however, the same slope is observed. The explanation of this effect is related to the non biasing of the electrodes. If these electrodes were biased an improvement should be expected.

In order to found a optimal V_{SS} to $pwell$ separations, two simulations on Figure 6d were made. As $pwell$ and $nwell$ go away of V_{SS} and V_{DD} , respectively, the V_H decreases. In Figure 8 the decreasing of V_H when X_{GR-S} is increased is shown. So the optimal X_{GR-S} tends to be 0 μm as in butted contacts.

The last variable in this latchup analysis is the guard ring diffusion width (W_{RG}). When W_{RG} increases, the effectiveness in carrier collection is better, in fact, as deepest is the ring guard diffusion, it is better in carrier collection. Four simulations were made with W_{RG} as variable, they are shown in Figure 9 and were based on Figure 6d. As the Figure 9 show, with $W_{RG} = 10 \mu\text{m}$ and $X_{sep} = 28 \mu\text{m}$, the majority carrier collection efficiency improves the holding voltage, rising it to 3.35 V. Therefore, with a correct combination of X_{sep} , W_{RG} and an

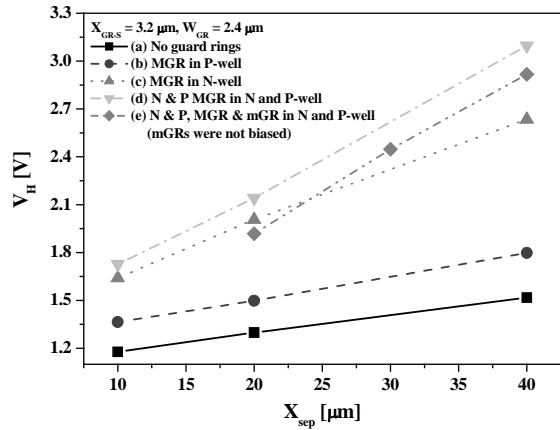


Figure 7. V_H vs. X_{sep} for the five guard ring structures shown in Figure 6.

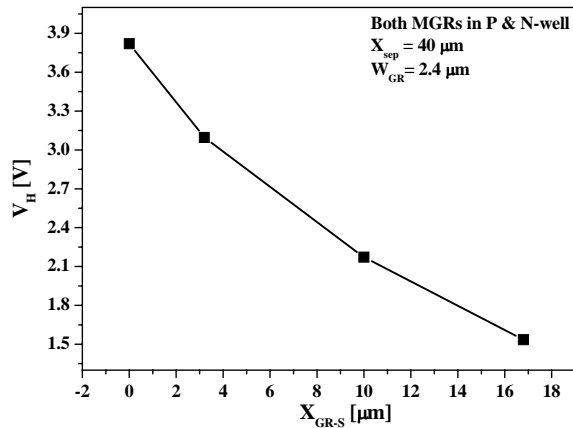


Figure 8. V_H vs. X_{GR-S} for Figure 5.

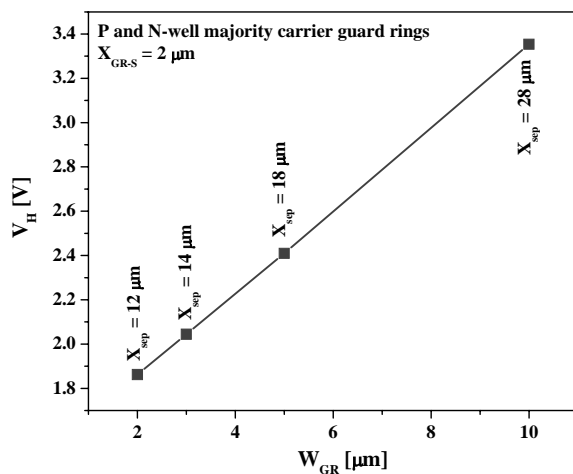


Figure 9. V_H vs. X_{sep} and W_{GR} for Figure 5.

appropriate design, a latchup free bulk CMOS process can be obtained. If a linear regression is made to the curves in Figures 7a-e, and 9, show a linear relation of V_H with X_{sep} and W_{GR} of the form

$$V_H = a + bX$$

where X can be changed by X_{sep} or W_{GR}

The Figure 8 shows a square relation of V_H with X_{GR-S} having the form

$$V_H = a + bX_{GR-S} + cX_{GR-S}^2$$

Some authors consider necessary mandatory to put majority carriers ring guards in P and N-well in the I/O circuits and minority carrier guard rings between those and the internal circuits. So the internal circuit are protected against triggering minority carrier current not collected by the majority guard rings [5].

7. Conclusion

In searching a low cost, latchup free, and simple CMOS process, a structure without guard ring structures requires V_{SS} to V_{DD} spacing as large as 350 μm to obtain $V_H=5$ V. Nevertheless, with a higher area consumption, the guard rings are an efficient technique to reduce this parasitic effect. For a structure without guard rings, V_H is as low as 1.2 V, and for structures with guard rings V_H can be as high as one want with the higher area consumption as tradeoff. However, the area layout could be larger in a latchup free bulk CMOS process if guard ring structures were not used. Therefore, with the base of the last simulation results in V_H values and the polynomial regressions, some predictions can be done to put the CMOS structures in a "safe space" maintaining V_H larger than 5 V.

1. The best structures in reducing latchup are shown in Figures 6d and 6e.
2. The increase in X_{sep} improves the resistance to latchup, with a double majority guard ring structure with $W_{GR} = 2.4 \mu\text{m}$, the minimum X_{sep} is 82 μm .
3. The optimal X_{GR-S} is 0 μm .
4. In a double majority guard ring structure, the minimal W_{RG} is 20 μm with a $X_{sep} = 38 \mu\text{m}$.
5. If mGR electrodes were connected to V_{DD} (n - mGR) and V_{SS} (p - mGR) in Figure 6e, a faster increase in V_H would be expected.

Acknowledgements

Authors thanks to CONACyT by the support with doctoral grant number 129162 and to the INAOE and SMCSyV by the support in the assistance to the XXIII National Congress of SMCSyV.

References

- [1] R. R. Troutman, Kluwer Academic Publishers, USA, 1986.
- [2] C. Y. Chen, Ch. 8, Prentice Hall, USA, 1990.
- [3] L. Deferm, IMEC, Belgium, 1989.
- [4] M. J. Hargrove, S. Voldman, R. Gauthier, J. Brown, K. Duncan, and W. Craig, IEEE 98CH36173. 36th Annual Reliability Physics Symposium, Reno, Nevada, 269 (1998).
- [5] M.-D. Ker, and W.-Y. Lo, IEEE Transactions on semiconductor manufacturing, **16-2**, 319 (2003).

