

回路シミュレータを用いた集積回路のばらつきと信頼性の評価

Evaluation of Variations and Reliabilities of VLSIs by Using Circuit Simulators

小林 和淑 Kazutoshi KOBAYASHI

京都工芸繊維大学 Kyoto Institute of Technology
〒 606-8585 京都市左京区松ヶ崎 Matsugasaki, Kyoto, 606-8585 Japan

1 はじめに

本稿では、ばらつきと信頼性を回路シミュレータを用いて評価する手法について概説する。ばらつきは主にモンテカルロ法で評価を行うが、集積回路の微細化により特に SRAM の不良解析が困難になってきている。ここでは、不良の境界付近のみでシミュレーションを行う手法の紹介も行う。信頼性においては、一時故障(ソフトウェア)と経年劣化(BTI)の評価法を説明する。ソフトウェアは中性子などの粒子線が集積回路に飛び込んで発生するエラーであり、FF やメモリの値が反転する。BTI (Bias Temperature Instability) は、MOS トランジスタのゲート酸化膜の欠陥により発生する劣化現象である。その双方の回路レベルの検証法について述べる。

2 ばらつき評価

ばらつき評価においては、すべてのトランジスタの性能を一律に変動させるワーストケース解析や、SPICE に標準で組み込まれているモンテカルロ (MC) 法を用いて実施することが多い。ロジック回路はそのトランジスタ数などから、クリティカルパス解析を行う場合には、 $\pm 3\sigma$ (σ は標準偏差) までの解析で済むことが多い。一方、ASIC や CPU 内の SRAM は数百 kbit から数 Mbit の規模である。含まれるトランジスタ数を考慮すれば、通常は $\pm 6\sigma$ までの解析を行わなければならない。ロジック回路も回路全体を対象とする場合にはより大きな σ に対して解析を行わなければならない。すべてのトランジスタが $\pm 6\sigma$ まで変動する確率は極めて低く、ワーストケース解析ではばらつきによる性能変動を悲観的に見積もってしまう。しかし、MC 法を素直に適用すると 10^{10} 回程度のシミュレーションを実行しなければならず現実的ではない。ここでは、MC 法と、MC 法を SRAM の SNM (Static Noise Margin) 解析に応用するための逐次重点的サンプリング法 [1] について解説を行う。

2.1 ばらつき

集積回路の微細化と大規模化ともない、トランジスタ特性などが大きくばらつく。トランジスタの飽和領域の電流(オン電流)は倍半分でばらつき、リーク電流(オフ電流)は桁違いにばらつく。ばらつきの主要な要因として、不純物ばらつき (Random Dopant Fluctuation, RDF), ゲート幅のばらつき (Line Edge Roughness, LER), ゲート酸化膜厚のばらつき (Oxide Thickness Fluctuation, OTF) が挙げられる。平本らは、多数のトランジスタを集積した DMA-TEG [2] の測定結果より、しきい値電圧 V_{th} のばらつきにおいて RDF が支配的であることを実証した [3]。RDF による V_{th} のばらつきは図 1 で表される正規分布 (Gaussian Distribution)

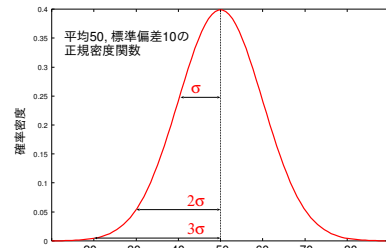


図 1 $\mu = 50, \sigma = 10$ の正規分布

となる。正規分布において重要なパラメータは平均値 μ と標準偏差 σ である。標本値 X_i に対する σ は次式により求められ、分布がどの程度広いかを表す尺度である。

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2} \quad (1)$$

正規分布の各範囲に含まれる割合は次のとおりとなる。

$\mu \pm \sigma$ この範囲に全体の 68.27%が含まれる。

$\mu \pm 2\sigma$ この範囲に全体の 95.45%が含まれる。

$\mu \pm 3\sigma$ この範囲に全体の 99.73%が含まれる。

SRAM のように、同一構造の回路ブロックを 100 万個を超えて集積する。 $\pm 6\sigma$ のばらつきまで考えて設計を行う理由はここにある。

2.2 モンテカルロ法

回路シミュレーションにおけるモンテカルロ法とは、乱数により生成された各種パラメータを用いて繰返しシミュレーションを行うことを指す。HSPICE では、`gauss`, `agauss` 関数を用いることで、正規分布している各種パラメータをばらつかせたシミュレーションを繰返し実行することができる。`gauss` 関数は次の通り用いる [4]。

```
.param p=gauss(nom_val, rel_var, sigma)
```

上記の例では、パラメータ `p` が、`nom_val` から $\pm \text{rel_var}$ までばらつく。`sigma=3` であれば、 $|\text{nom_val} - \text{rel_var}|$ が 3σ の正規分布でばらつく。

2.3 逐次重点的サンプリング法

佐藤らの提案する逐次重点的サンプリング法 (Sequential Importance Sampling, SIS) では、不良となる境界部分のみで MC 法を行うことにより、少ない回数で $\pm 6\sigma$ のばらつきによる不良推定が可能とする手法である。重点的サンプリングとは、図 2 に示すように、平均値をあらかじめ不良領域付近にシフトさせることによりシミュレーション回数を大幅に減らす手法である。本来サンプリングを行う分布 $f(x)$ とシフト後の代替分布 $g(x)$ について、重み関数 $w(x) = f(x)/g(x)$ を用いることで分布が異なる

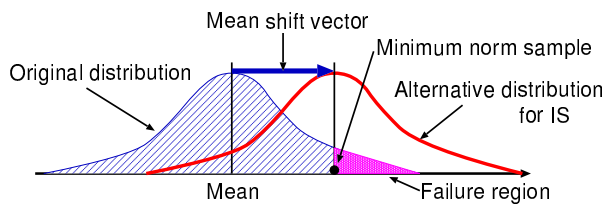


図 2 重点的サンプリングにおける平均値移動 [1]

ることによる起こりやすさをキャンセルできる。重点的サンプリング (IS) では、ある事象 A が起きる確率 p_{IS} を $g(x)$ からのサンプルを用いて次の通りに表す。

$$p_{IS} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N w(\tilde{X}_i) I(\tilde{X}_i \in A) \quad (2)$$

SIS ではさらに、 $g(x)$ を固定分布で構成する代わりに逐次的に更新することで、重点的サンプリングの効率化と計算の安定化を図る。

[1] では、SIS を SRAM の読み出しノイズマージン (RNM) の解析に応用し、従来の MC 法と比べて 2.5×10^6 倍の高速化を達成した。

3 信頼性評価

ここでは、信頼性のうち、ソフトエラーと BTI (Bias Temperature Instability) の SPICE による評価方法について解説する。

3.1 ソフトエラー

ソフトエラーとは、中性子、 α 粒子 (α 線, He の原子核) などのいわゆる放射線が LSI 内部に侵入し、FF や SRAM などの記憶素子の値が反転することである。電源を入れなおしたり、正しい値を再度書きこむことにより元に戻るため、永久故障を表すハードエラーに対して、ソフトエラーと呼ぶ。その発生要因は、粒子の侵入により電子正孔対が生成され、そのキャリアが OFF 状態のトランジスタに蓄えられている電荷を打ち消すことにより発生する SET (Single Event Transient) パルスと呼ばれるエラーパルスである。空乏層による電界によりドレインに引き寄せられる (ドリフト現象による) キャリアと拡散現象によるキャリアが SET パルスの要因となる。

3.1.1 ソフトエラーの要因

ソフトエラーは宇宙線によって引き起こされる。地上ではソフトエラーを引き起こす宇宙線は次の主に 3 種類である (図 3) [5]。

高エネルギー中性子 高エネルギー中性子は宇宙に飛来している高エネルギーのイオンが大気圏に入り、他の原子と反応することで発生する。Si 基板の Si 原子と衝突して Al や Mg の荷電イオンが生じる。中性子のエネルギーが 1MeV から 10MeV 以上のものを高エネルギー中性子と呼ぶ。

熱中性子 層間絶縁膜として使われている BPSG (borophosphosilicate glass) 膜に存在する Boron の同位体である ^{10}B はエネルギーの低い熱中性子と衝突しやすい性質を有する。衝突すると、 α 粒子と Li が生じる。熱中性子のエネルギーは 1eV と、高エ

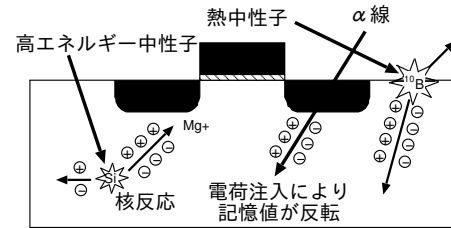


図 3 ソフトエラーの要因

ネルギー中性子と比べて小さいが、BPSG 膜が使われていた頃はソフトエラーの要因の約 75% を占めていた。しかし CMP (化学機械研磨) に移行した後は、約 1/100 と激減した。微細化が進み、コンタクト材料としてタングステンプラグが広く使われている。このタングステンプラグにも ^{10}B が含まれており、再び熱中性子がソフトエラーの要因となっている。 α 粒子 LSI やパッケージの材料に含まれる放射性不純物が α 崩壊することによって α 粒子が生じる。 α 粒子がドレインに突入した時、通過領域に電子正孔対が生成される。

ミューオン 素粒子の一つであるミューオンも微細化に伴う臨界電荷量の減少に伴い、ソフトエラーを引き起こす要因として注目されている [6]。

高エネルギー中性子は高度に依存して増加するため、山頂や飛行機中ではソフトエラー率は上昇する。一方、宇宙空間では高エネルギーの陽子線や重粒子線が主要な要因となる。地上では組み合わせ回路で生じるソフトエラーによるパルス幅は 100ps 程度だが、宇宙空間では 1ns にも達する [7]。

α 線は紙 1 枚で遮蔽できる。中性子は水や水を多く含むコンクリートで遮蔽することができるため、ビル内では中性子起因のソフトエラーは少なくなる。ミューオンは、物質とほとんど相互作用をしないため、遮蔽が困難である。

3.1.2 ソフトエラーの分類

ソフトエラーはラッチなどで起こる SEU (Single Event Upset) と、組み合わせ回路部分で起こる SET (Single Event Transient) に分類される。SEU はラッチや SRAM などの記憶素子で起きるソフトエラーである。これらは保持している値が反転するとエラーとなる。組み合わせ回路で起きる SET は、フリップフロップに取り込まれて初めてエラーとなる。

3.1.3 中性子起因の誘起電荷による電流

中性子起因の誘起電荷による電流として、まず空乏層における電界によって電子が引き寄せられるドリフト電流が発生する。その後、拡散による電流や、発生した電荷によって電界が生じファネリングによる電流が流れる。現在、中性子起因の誘起電荷による電流のモデルは 2 種類提案されている。[8, 9, 10] では Double Exponential Model (DEM), [11, 12, 13, 14] では式 (3) の Single Exponential Model (SEM) が提案されている。

$$I(t) = \frac{2Q}{T\sqrt{\pi}} \sqrt{\frac{t}{T}} \exp\left(-\frac{t}{T}\right) \quad (3)$$

T はプロセスによって決まる時定数であり、今後減少していく。式 (3) を $t=0$ から ∞ まで積分すると Q となる。 Q はドレインに集められる電荷の総量を示しており、荷電粒子により生成される電荷の総量に等しい。SEM においては、他のパラメータ値に関係なく Q により電荷の総量を設定することが可能であるが、DEM では、様々なパラメータの組合せで電荷総量を決めなければならない。回路シミュレーションにおいては、式 (3) を用いてあらかじめ電流値を計算し、その値を PWL (PieceWise Linear, 折線) 波形の電流源として保存しておく。計算時には総電荷量は固定だが、回路シミュレーション時には、電流制御電流源にて、電荷量の調整を行う。

3.1.4 SER (Soft Error Rate) の見積もり

単位時間あたりに発生するソフトエラーの数をソフトエラー率 (Soft Error Rate, SER) と呼び、FIT (10^9 時間あたりに故障が発生する回数の期待値) を単位とするのが一般的である。概算ではあるが、一般的な SRAM のエラー率は 1000FIT/Mbit である。これは、1Mbit の SRAM が 114 年に 1 回反転することを意味する。

SER を N_{SER} とすると式 (4) で表される [11, 13, 15].

$$N_{SER} = F \cdot CS = F \cdot A \cdot K \cdot \exp\left(-\frac{Q_{crit}}{Q_s}\right) \quad (4)$$

F は中性子量 (Neutron Flux) である。ニューヨークの海面レベルの中性子数を用いることが多く、約 20 neutron/cm²/h である。 CS は Cross Section (衝突断面積) と呼ばれ、単位時間の 1 粒子あたりのソフトエラーの数である。

ソフトエラーは、トランジスタのドレイン付近に粒子が突入することで発生する。従って A は、バルクプロセスにおいてはソフトエラーを起こす可能性のあるノードに接続された MOS トランジスタのドレイン面積の和となる。SOI プロセスでは基板に誘起された電荷は BOX 層によりドレインには到達しない。SOI でのソフトエラーはゲート直下に粒子が突入し、SOI 層で発生する電荷により寄生バイポーラトランジスタが ON することによるため、電流源のみでソフトエラーを見積もることは難しい [16, 17]。 Q_{crit} は臨界電荷量 (critical charge) と呼ばれ、保持データを反転させるのに必要な電荷量である。 Q_s は電荷収集効率 (charge collection efficiency) と呼ばれ、プロセスによって決まる量であり、微細化により減少していくと予想されている [15].

3.2 SPICE による SER の見積もり

図 4 は、中性子衝突などにより生成された電子正孔対による電流を模擬した電流源を付加したシミュレーション回路である。図 5 の通り総電荷量 1C を放出する電流源 I_c を接続したノード is を定義する。 is を接続した 1V の電圧源 Vis を接続した電流制御電流源 $F0$ を定義する。 $F0$ は電荷量 Q をパラメータとして持つため、これを変化させて、SET パルスの変動を見ることが出来る。SRAM などのメモリの場合は Q を変化させ、反転が起きる電荷量を求め、これを Q_{crit} とする。

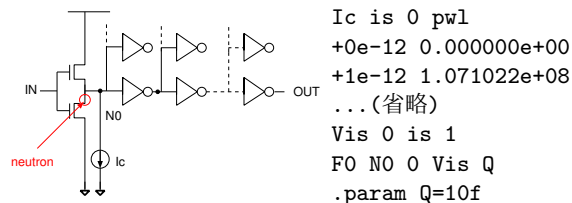


図 4 組合せ回路の評価

図 5 電流源付加方法

4 BTI 解析

BTI とは、ゲート酸化膜中の欠陥 (Oxide Trap) や、ゲート酸化膜とシリコンとの界面の欠陥 (Interface Trap) により引き起こされる経年劣化を指す。BTI 現象が問題視されていた当初は界面欠陥のみをモデル化した Reaction Diffusion (R-D) モデル [18] が提案された。その後、酸化膜中の欠陥をモデル化した Atomistic Trap-based モデル [19] も提案された。当初は両者の間で激しい論争 [20] を巻き起こしたが、現在では、回復する劣化は酸化膜中の欠陥が関係し、回復しない劣化は界面欠陥が関係しているという両者を合わせた説が有力となっている。ここでは、両モデルの概説を行い、回路シミュレータを用いた BTI 解析法について述べる。

4.1 BTI モデル

BTI の回復可能な成分の発生原理は RTN (Random Telegraph Noise) の重畳であるという節が有力である。RTN とはゲート酸化膜中の欠陥にキャリアが捉えられ、閾値電圧が動的かつ離散的にばらつく現象である。ゲート酸化膜にある複数の欠陥はそれぞれに異なった特性を持っており、時間の経過と共にキャリアを捕獲、放出する。キャリアが捕獲されている時に閾値電圧は高い状態に、放出されている時に低い状態にあり、それぞれの状態の平均持続時間を放出時間 (Emission Time: τ_e)、捕獲時間 (Capture Time: τ_c) という。これらは 10^{-9} s から 10^9 s 程度の広い範囲に対数等分布しているとされる [21]。欠陥によってはキャリアを一度捕獲したら半永久的に捕獲した状態となり、閾値電圧に対して恒常的な影響を及ぼすものもある。これらの時定数はゲート電圧依存性を持ち、ゲート・チャネル間にストレス電圧が印加された場合に τ_e は長くなり、 τ_c は短くなることが多いとされる。そのため、閾値電圧はストレス状態においては高くなりやすく、ストレス電圧を開放したりラックス状態においては低くなりやすいということが出来るので、BTI の基本的な特性と一致する。

ある一つの欠陥にキャリアが捕獲されたことで生じる閾値電圧の変動量 (μ) はゲート面積に反比例し、指数分布に従う。時定数 τ_e , τ_c と μ の相関については、有無を含めて現在も議論が行われている。

4.2 Atomistic Trap-based Model を用いた BTI 閾値電圧変動量の算出

MOSFET のゲート酸化膜にある欠陥の特性及び状態によって閾値電圧の変動量が決まる。ある閾値電圧 (V_{th}) を持つ MOSFET に N 個の欠陥がある場合には、閾値電

圧の変動量 (ΔV_{th}) は以下の式 (5) によって求められる.

$$\Delta V_{th}(t) = \sum_{j=1}^N k_j(t) \cdot \mu_j \quad (5)$$

ここで t は時間を表し, j は欠陥のインデックスであり 1 から N の値をとる. 欠陥の状態 k は捕獲状態で 1、放出状態で 0 である. k は各欠陥の捕獲確率 P_C によって決定され, 長期間の NBTI 劣化における P_C は式 (6) によって求められる [22].

$$P_C(t) = \frac{\tau_e^*}{\tau_c^* + \tau_e^*} \left[1 - \exp \left\{ - \left(\frac{1}{\tau_e^*} + \frac{1}{\tau_c^*} \right) t \right\} \right] \quad (6)$$

$$\frac{1}{\tau_c^*} = \frac{DF}{\tau_{ch}} + \frac{1-DF}{\tau_{cl}} \quad (7)$$

$$\frac{1}{\tau_e^*} = \frac{DF}{\tau_{eh}} + \frac{1-DF}{\tau_{el}} \quad (8)$$

ここで DF , τ_e^* , τ_c^* はそれぞれデューティファクタ, 実効放出時間, 実効捕獲時間である. ゲート入力信号の周波数 f , 1 周期中の信号が High の時間を t_H とした時 $DF = f \times t_H$ となり, ゲート電圧が DC の場合には $DF = 1(100\%)$ である. ゲート電圧が High の時の τ_e , τ_c をそれぞれ τ_{eh} , τ_{ch} , ゲート電圧が Low の時の τ_e , τ_c をそれぞれ τ_{el} , τ_{cl} としている. 各時定数は τ_{el} は 10^{-9} s から 10^9 s に対数等分布, τ_{ch} , τ_{eh} , τ_{cl} は τ_{el} と相関を持つと仮定している [21, 23]. 式 (6) は近似式であるが, 時定数が $1/f$ より十分大きい時にはよい近似となる.

4.3 BTI 劣化を考慮した回路解析手法

BTI 劣化を考慮した回路解析を行うためには, BTI による閾値電圧変動量を含んだ BTI-Aware Netlist が必要となる [24]. シミュレーションの対象となる回路の Netlist に含まれる各 MOSFET について, それぞれ時間 t に対応した閾値電圧変動量の予測値を与えることで, 各 MOSFET の劣化を再現することが出来る. 閾値電圧変動量を算出する BTI Model は, 計算のために MOSFET のデバイスパラメータとシミュレーションの条件を要する. 図 6 のように Netlist と BTI Model から BTI-Aware Netlist を作成し, これに対してシミュレーションを行うことで解析結果を得る.

HSPICE 向けの BSIM4 を用いた具体的な BTI-Aware Netlist の記述例を図 8 に示す. BSIM4 ではパラメータ DELVTO を設定することで個々の MOSFET の閾値電圧のパラメータ V_{th0} を変更可能であり, その素子の閾値電圧は設定した値の分シフトする [25]. この例では単一の NMOS, PMOS それぞれに dv_{thn} , dv_{thp} の値を設定している. 前述の通り劣化時間 t 毎に BTI-Aware Netlist が必要となるため, 例では $t1 \sim t3$ の 3 つを用意している.

4.4 解析例

条件として $N = 800$, $V_{th} = 500\text{mV}$, $DF = 50\%$ を設定した場合の閾値電圧変動量の平均と対数関数によるフィッティング結果を図 7 に示す. 横軸は対数目盛で劣化時間, 縦軸は閾値電圧の変動量である. 閾値電圧変動量の平均値は 10^8 s 近辺で飽和傾向を示しており, この領域においては対数関数からずれているといえる. この

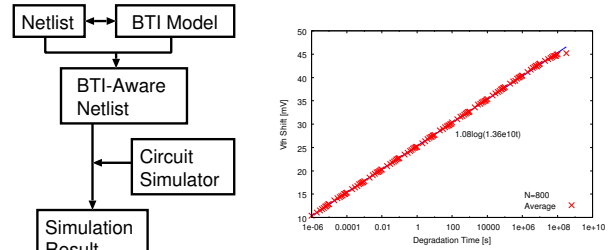


図 6 解析手法

図 7 閾値変動の計算結果

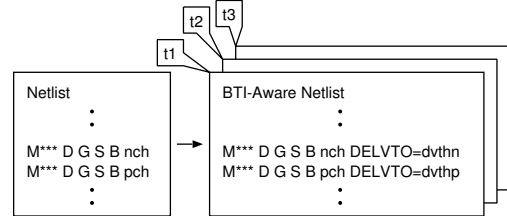


図 8 BTI-Aware Netlist の記述例.

ような飽和傾向は τ_{el} が最大 10^9 s で分布していて, 各時定数の最大値が 10^7 s から 10^{11} s になっていることに起因すると考えられる.

5 まとめ

本稿では, ばらつきと信頼性をどのように SPICE を用いて評価するかを述べた. 両者ともに半導体物理に関連する評価項目のため, 回路レベルでの評価が難しい. しかし, うまく使えば, デバイスシミュレーションに頼らなくとも, 簡便かつ高速に評価を行うことが可能となる.

謝辞

本稿執筆にあたり, 京都大学の小野寺秀俊教授, 佐藤高史教授, 東京大学の平本俊郎教授より資料の提供やアドバイスをいただいた. EDA ツールについては東大 VDEC を通じ, シノプシス社の協力で利用している.

参考文献

[1] ICCAD, pp. 703–708, 2010. [2] Trans. Semi. Manu., vol. 17, pp. 155–165, 2004, [3] Trans. ED., vol. 9, pp. 2073–2080, 2009. [4] “HSPICE Simulation and Analysis User Guide,” Synopsys, 2004.. [5] 日経エレクトロニクス, no. 903, pp. 145–155, 2005. [6] IRPS, pp. 3C.3.1–3C.3.6, 2011. [7] 日経エレクトロニクス, no. 903, pp. 63–70, 2005. [8] Trans. Nucl. Sci., vol. 29, pp. 2024–2031, 1982. [9] “Fault-tolerance Techniques for SRAM-based FPGAs,” Springer, 2006. [10] Trans. Nucl. Sci., vol. 51, pp. 3285–3290, 2004. [11] JSSC, vol. 35, pp. 1422–1429, 2000. [12] Trans. Nucl. Sci., vol. 47, pp. 2586–2594, 2000. [13] “Modeling the effect of technology trends on the soft error rate of combinational logic”, ISDSN, pp. 389–398, 2002. [14] DTC, vol. 19, pp. 54–68, 2002. [15] IEDM, pp. 21.5.1–21.5.4, 2003. [16] IRPS, SE.5.1–SE.5.4, 2013. [17] IEICE Trans. on Elec., E96-C, pp. 511–517, 2013. [18] IEDM, pp. 14.4.1–14.4.4, 2003. [19] IRPS., pp. 26–32, 2010. [20] Trans. on Dev. & Mat. Rel., vol. 14, pp. 182–193, 2014. [21] IRPS, pp. 7–15, 2010. [22] IRPS, pp. 4A.2.1–4A.2.8, 2011. [23] VLSI Tech., pp. 148–149, 2011. [24] IPSJ TSLDM, vol. 5, pp. 143–149, 2012. [25] “BSIM4. 3.0 MOSFET Model User; s Manual”, UCB