The Effect of Compact Modelling Strategy on SNM and Read Current variability in Modern SRAM

P. Asenov[1], F. Adamu-Lema[1], S. Roy[1], C. Millar[1,2], A. Asenov[1,2]
School on Engineering
University of Glasgow[1]
Glasgow, Scotland
(P.Asenov) 0405152a@student.gla.ac.uk

G. Roy[2], U. Kovac[2], D. Reid[2]
Gold Standard Simulations Ltd. [2]
Glasgow, Scotland
G.Roy@GoldStandardSimulations.com

***Abstract*— It has been shown that sub 100nm SRAM is particularly sensitive to stochastic device variability. In this paper we consider two correlated figures of merit for SRAM, Static Noise Margin (SNM) and Read Current. For the purposes of this paper 1,000 3D atomistic simulations of microscopically different 25nm P and N bulk MOSFETs were performed, and statistical compact models were then extracted for each device. Using these models simulations are performed to calculate the SNM and Read Current distributions of SRAM cells constructed using devices from the device ensemble. Variability in device performance has been then introduced via Gaussian or skewed Gaussian threshold voltages (V_t) and by using values of V_t extracted directly from the individual device compact models and the results of these simulations are then compared to the baseline simulations using fully extracted models. The results clearly demonstrate the errors that can be introduced in the estimation of SNM and Read Current distribution of a 6T SRAM cell when statistical device variability is not correctly modelled.**

Keywords-SNM, RandomSpice, Statistical Variability, Compact Model Extraction

I. INTRODUCTION

Modern silicon design is complicated by the presence of statistical MOSFET variability, a problem which has been highlighted in both industry[1] and academia[2]. Exacerbated by the continuous drive toward reducing device sizes, stochastic statistical variability results from the discreteness of charge and granularity of matter. Dominant stochastic variability sources include random discrete dopants (RDD)[3], line edge roughness (LER)[4] and polysilicon (PSG)[5] and metal gate (MGG) granularity. Statistical variability introduces random differences in the behaviour of technologically ideal devices. Recently, it has been demonstrated that in contemporary CMOS technology the level of statistical variability is overtaking that of process variability[6].

A significant portion of the chip area in modern SoC applications is occupied by SRAM, and its dependence on minimal width transistors leaves SRAM acutely sensitive to random statistical variability. The huge number of SRAM cells in modern memory arrays provides a strong motivation to simulate SRAM performance with statistical confidence greater than 5 sigma. Two traditional figures of merit, used to determine the functionality and yield of an SRAM design, are the static noise margin (SNM)[7] and read current. SNM provides a measure of the stability of a cell and read current is a measure of the limiting readability of the cell given a stored 0. In a successful SRAM design there is a trade-off between cell stability (SNM) and read current which can only be accurately assessed in the presence of variability via simulation.

It has been shown that in sub 100nm technologies statistical variability has a significant impact on SRAM SNM, yield and performance[8], and it is clear that this must be factored into the SRAM design process. Techniques have been proposed to model the effects of variability on SNM and read current, however these are generally limited to including only the variability of the threshold voltage [9][10]. In order to determine the impact of variability over the full current-voltage characteristics of devices e.g. including the impact on on-current, off-current and transconductance, we compare statistical SRAM SNM and read current simulations performed with statistically extracted BSIM4 models, which accurately capture the complete range of variable device behaviour, and compare these to simulations where only threshold voltage variability in various guises is considered.

II. METHODOLOGY

A. Compact Model Extraction

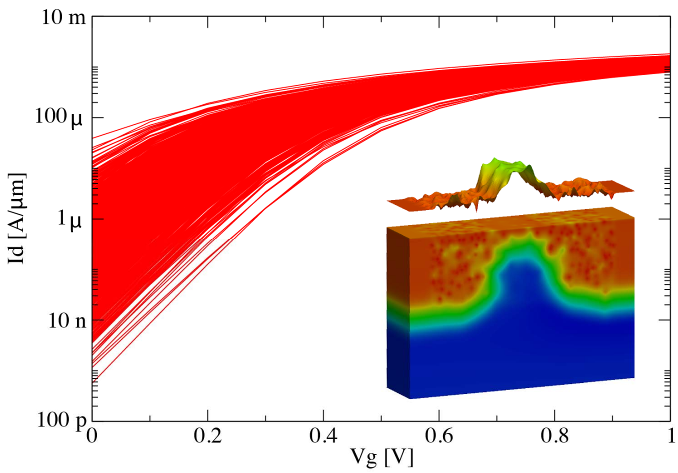

Figure 1. Id-Vg transfer charactesistics for NMOS at high drain condition

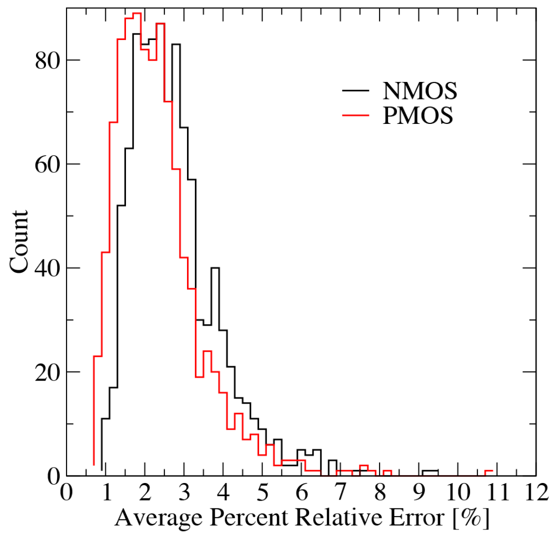

Figure 2. Average Percentage Relative Error for NMOS and PMOS 1000 model extraction

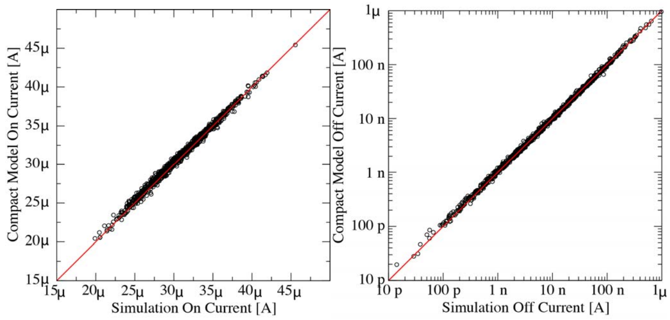

Figure 3. On current (left) and off current (right) comparison between device simulation and extracted compact model

Using the atomistic 3D simulator Garand[11] the transfer characteristics (I_d-V_g) of an ensemble of 1000 p- and n-, 25nm channel length devices was simulated[12] including the effects of variability due to RDD, LER and MGG. An example of the simulated device and the corresponding statistical transfer characteristics is shown in figure 1. For each device in the ensemble a compact model is then extracted as described by Cheng *et al.*[13] using the Mystic statistical compact model extraction tool[11]. This is a two-stage process. First a nominal compact model based on a 'uniform' device - containing no sources of variability - is extracted. Then, based on a detailed sensitivity analysis, a subset of the standard BSIM4 parameters is chosen (in this case 6 parameters) which are then used to model the impact of variability on individual instances of the device. It is important to accurately capture the characteristics of each device in the statistical compact models in order ensure the accuracy of baseline circuit simulations against which the accuracy of approximate methods can be judged.

Figure 2 shows the distribution, for both the extracted PMOS and NMOS compact models, of the average percentage relative error over all I_d-V_g curves when compared to the simulated device characteristics. It is important to ensure that the baseline statistical compact model ensemble accurately reproduces the distribution of important figures of merit for the simulated device ensemble. Figures 3(left) and 3(right) show a comparison of the on and off current distribution obtained from statistical compact modelling and 3D Drift Diffusion simulation for the NMOS device at high drain bias. Accurate 1-to-1 correlation between device on/off current and extracted compact model on/off current is denoted by the straight line at 45 degrees. For the device ensemble, the average error, calculated over multiple I_d-V_g characteristics at different drain biases, introduced by compact model extraction is ~2.6% for the NMOS and ~2.4% for PMOS devices. An additional compact model is extracted from the ensemble average device behaviour to serve as the baseline device model for simulations where only V_t variation is considered. This corrects for the fact that, due to V_t lowering from the effect of RDD, the behaviour of an idealised uniform device obtained from TCAD is not the same as the average behaviour of an ensemble of variable devices, and allows a more accurate comparison of the two simulation methods.

B. SNM/Read Current Simulation

The simulation configuration and calculation method for SNM can be seen in figure 4. In the following results 50,000 SNM and 10,000 read current values were obtained using the average performance compact model and threshold variability is introduced by varying the BSIM4 VTH0 parameter. VTH0 values for individual devices are generated from Gaussian and skewed Gaussian distributions and by using the values of VTH0 extracted directly from the statistical compact model ensemble. The baseline set of simulations, performed using the full compact models extracted from the device ensembles, accurately capture the effect of variability over the whole range of device operation. These simulations are considered to be the "gold standard" against which a comparison of the errors introduced by VTH0-only methodologies can be tested.

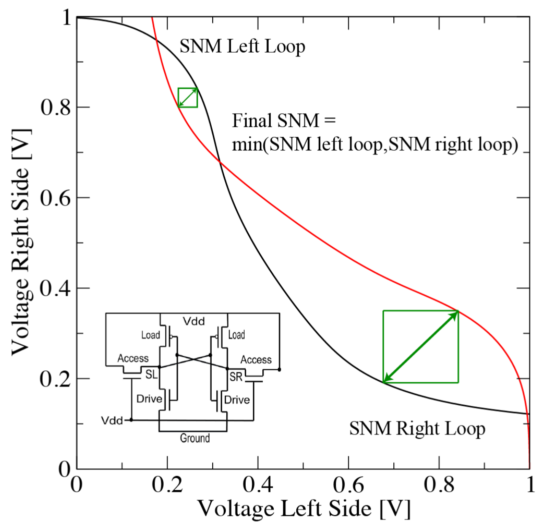

Figure 4. SNM Simulation Setup and Calculation

In order to reduce simulation time, the simulations were performed in parallel on a High Performance Compute (HPC) cluster, facilitated by using the Gold Standard Simulations Push-button cluster technology. All circuit simulation was performed using the RandomSpice[14] statistical circuit simulation toolkit.

III. Results

A. Gaussian and Skew-Gaussian V_{th}

Initial simulations were performed where variability is introduced via Gaussian distributed threshold voltage variability, where the mean and standard deviation of the distribution are the same as that obtained from full compact model extraction. In order to assess the impact of V_t skew on read current, further simulations were performed with skewed Gaussian distribution where the skew of the distribution is varied from -0.5 to 0.5 and the same mean and standard deviation are used.

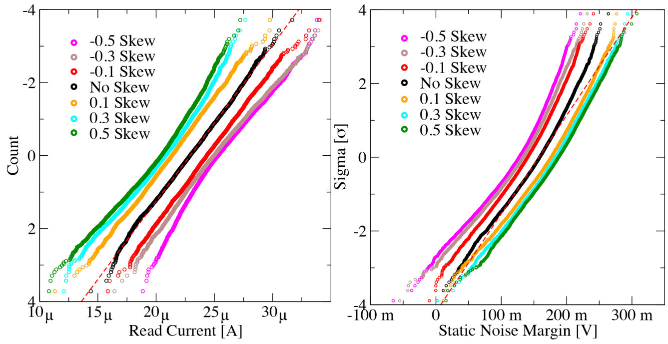

Figure 5. Read Current Distribution (left) and SNM distribution (right) at different amounts of threshold voltage skew

TABLE I. Simulated Read Current Distribution Moments

	-0.5	-0.3	-0.1	0.0	0.1	0.3	0.5
Mean	25.50	25.05	24.39	22.90	21.30	20.67	20.20
Std Dev	2.345	2.493	2.335	2.370	2.343	2.370	2.379
Skew	0.38	0.23	0.04	0.05	0.09	0.35	0.46

TABLE II. Simulated SNM Distribution Moments

	-0.5	-0.3	-0.1	0.0	0.1	0.3	0.5
Mean	0.122	0.129	0.137	0.156	0.173	0.180	0.185
Std Dev	0.037	0.037	0.035	0.035	0.037	0.036	0.036
Skew	-0.70	-0.64	-0.48	-0.47	-0.47	-0.41	-0.38

The resulting SNM and Read Current distributions are shown in Figures 5 (left) and 5 (right). It is apparent that V_t skew causes a shift in the mean and skewness of the final SNM and Read Current distributions, the this is more clearly shown in tables 1 and 2. Gaussian V_t variability produces a Gaussian distribution of read current, This indicates that simulation methodologies based on the assumption of Gaussian distributions of these figures of merit could introduce large errors especially in the tails of the distribution if the true distribution of device V_t is non-normal. It is interesting to note that the distribution of the SNM is always negatively skewed, even in the case where the underlying V_t distribution is positively skewed. This is due to the Min operation performed during the SNM calculation. This negative skew of static noise margin is important as it leads to the conclusion that in order to estimate SRAM yield with respect to SNM, we must accurately determine higher moments of the SNM distribution and not simply the mean and standard deviation.

As expected the effect on SNM and Read Current is anti-correlated as these two figures of merit define opposing effects - SNM is a measure of the ability to hold data, and Read Current is a measure of the ability to read data

B. Extracted V_{th} and Full Compact Model

Simulations were then performed using fully extracted compact models, and models where the average device mode is modified by using the values of VTH0 obtained from the extracted models. This provides an accurate distribution of V_t and allows the assessment of the impact of variability in other device figures of merit, since the fully extracted models incorporate the impact of variability on V_t, DIBL, on-current and off-current.

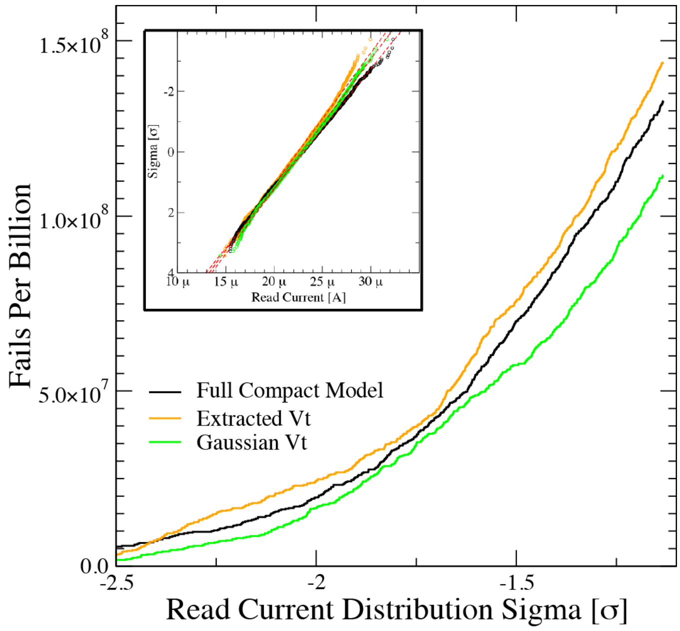

Figure 6. Read Current Distribution with different compact model variability injection approaches

The resulting SNM and read current distributions are shown in figures 6 (Read Current) and 7 (Static Noise Margin). SNM and read current simulations with threshold voltage variability capture the mean of the resultant distribution well, but badly underestimate the standard deviation and the skew or the resultant distribution when compared to the more accurate full compact model simulations, as illustrated by the data in table 3.

TABLE III. Simulated Read Current Distribution Moments

	Full Models	Extracted V_t	Gaussian V_t
Mean	22.98	22.48	22.90
Std Dev	2.58	2.31	2.37
Skewness	-0.0188	0.139	0.0494

It is worrying to note that, using the extracted distribution in both read current and SNM simulations, introducing variability via V_t alone does not accurately capture the correct distribution of the output values. This effect is clearly visible when one considers yield. Figure 7 shows the fail count (measured in parts per billion (PPB)) for a simulated SRAM cell when one considers its SNM. Using a 3.5 sigma cutoff as an example, it is clear that Gaussian V_t simulations significantly overestimate yield, predicting only ~880,000 PPB failures, while extracted V_t simulations underestimate yield predicting ~2,260,000 PPB failures, compared to the full compact model simulations which predict ~1,440,000 PPB failures.

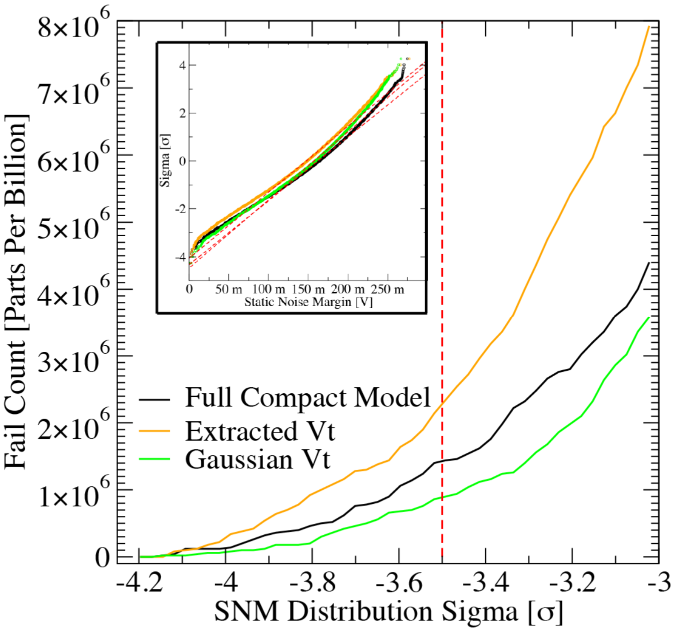

Figure 7. SNM Fail count predictions with different injection approaches and full distribution QQ plots (inset)

This data leads to the conclusion that any Monte-Carlo or statistically enhanced SRAM simulation methodologies used to determine SRAM performance at 4σ or greater, will produce inaccurate results if only variability in V_t is considered. Even in the case where a more accurate non-Gaussian distribution of V_t is used to model this effect the result of simulations do not provide realistic values. These problems are most pertinent when attempting to estimate the yield of SRAM, and when one attempts to optimize a design based on figures of merit. For these evaluations we rely on accurate distributions of the figures of merit of the SRAM cell. If the shape, and therefore, tails of these distributions are incorrect, large errors can be introduced leading to inefficient optimization, over/under design and overall design failure.

IV. CONCLUSIONS

In modern SRAM design where it is necessary to simulate parameter distributions to ±5 sigma in order to detect possibly fatal fail states and determine the correct yield of a design, simulations based on V_t variability alone are not sufficiently accurate. V_t only simulations underestimate the complexity of the SRAM figures of merit and could lead to under/overdesign. More advanced simulation methods, which accurately reproduce the statistical device behavior, including accurate statistical compact model extraction strategies, such as those described in this paper, are required.

REFERENCES

[1] S.R. Nassif, "Design for variability in DSM technologies" , Quality Electronic Design, 2000. ISQED 2000. Proceedings. p.451 - 454

[2] Y Cao, P Gupta, AB Kahng, D Sylvester, "Design Sensitivities to Variability, Extrapolations and Assessments in Nanometer VLSI" Proc. ASiC/SOC, 2002

[3] A. Asenov, G. Slavcheva, A. R. Brown, J. H. Davies and S. Saini, "Increase in the random dopant induced threshold fluctuations and lowering in sub-100nm MOSFETs due to quantum effects: A 3-D density-gradient simulation study," IEEE Trans. Electron Devices, vol. 48, no.4, 2001, pp. 271–350

[4] A. Asenov, S. Kaya, A. R. Brown, "Intrinsic parameter fluctuations in decananometer MOSFETs introduced by gate line edge roughness," IEEE Trans. Electron Devices, vol. 50, no.5, 2003, pp. 1254–1260

[5] A. R. Brown, G. Roy and A. Asenov, "Impact of Fermi level pinning at polysilicon gate grain boundaries on nano-MOSFET variability: A 3D simulation study," in Proc. 36th ESSDERC, 2006, pp.451-454

[6] H. Aikawa, T. Sanuki, A. Sakata, E. Morifuji, H. Yoshimura, T. Asami, H. Otani and H. Oyamatsu "Compact Model for Layout Dependent Variability" IEDM09 P699-702

[7] E. Seevink, F. J. List, J. Lohstron, "Static-Noise Margin Analysis of MOS SRAM Cells", IEEE Journal of Solid-State Circuits, Vol 22. No. 5, Oct 1987

[8] R. Herald, P. Wang, "Variability in Sub-100nm SRAM Designs", in Proc. Int. Conf. Comput.-Aided Des., 2004, pp. 347–352

[9] K. Agarwal, S. Nassif, "The Impact of Random Device Variation on SRAM Cell Stability in Sub-90-nm CMOS Technologies", IEEE Trans. On VLSI Systems, Vol. 16, No. 1

[10] B. H. Calhoun, A. P. Chandrakasan, "Static Noise Margin Variation for Sub-threshold SRAM in 65-nm CMOS", IEEE Journal of Solid-State Circuits, Vol. 41, No. 7, July 2006

[11] url=www.goldstandardsimulations.com

[12] A. Asenov, "Simulation of Statistical Variability in Nano MOSFETs" in 2007 Symposium on VLSI Technology, p. 86 – 87

[13] B. Cheng, D. Dideban, N. Moezi, C. Millar, G. Roy, X. Wang, S. Roy, A. Asenov, "Statistical-Variability Compact-Modeling Strategies for BSIM4 and PSP", IEEE Design & Test of Computers, Vol. 27, Issue 2, March-April 2010

[14] P. Asenov, N. A. Kamsani, D. Reid, C. Millar, S. Roy, A. Asenov, "Combining process and statistical variability in the evaluation of the effectiveness of corners in digital circuit parametric yield analysis", in Proc. ESSDERC 2010, 14-16 Sept 2010, p. 130-133