LIVERPOOL JOHN MOORES UNIVERSITY

LJMU Research Online

Liu, X, Ren, P, Chen, H, Ji, Z, Liu, J, Wang, R, Zhang, JF and Huang, R

Equiprobability-based Local Response Surface Method for High-Sigma Yield Estimation with Both High Accuracy and Efficiency

http://researchonline.ljmu.ac.uk/id/eprint/17292/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Liu, X, Ren, P, Chen, H, Ji, Z, Liu, J, Wang, R, Zhang, JF and Huang, R (2022) Equiprobability-based Local Response Surface Method for High-Sigma Yield Estimation with Both High Accuracy and Efficiency. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 42 (4), pp.

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/

This article has been accepted for publication in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. This is the author's version which has not been fully edi
content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3193875

Equiprobability-based Local Response Surface Method for High-Sigma Yield Estimation with Both High Accuracy and Efficiency

Xiang Liu, Pengpeng Ren, *Member, IEEE*, Haibao Chen, Zhigang Ji, *Member, IEEE* ,Junhua Liu, Runsheng Wang, Jianfu Zhang, and Ru Huang, *Fellow, IEEE*

***Abstract*— With the ever-increasing transistor density and memory capability in integrated circuits, the high-sigma yield estimation has become a growing concern. This work presents an equiprobability-based local response surface method (ELRS) that can perform high-sigma yield estimation with both high accuracy and efficiency. Demonstrating with 6T-SRAM, the proposed method exhibits more than 10 times improvement in accuracy when comparing with the state-of-the-art while maintaining the efficiency to the best record in literature.**

***Index Terms*—Monte Carlo, response surface, high sigma, rare events.**

I. Introduction

Increasing demand for data generation, storage, and intelligence generation from data is driving advances in memory technology with high capacity and speed without taking more power or volume, which urges for a high yield of memory cells [1]. **Fig. 1(a)** shows the evolution of the storage capability for different memory technologies in the past 10 years. To ensure sufficient quality for mass production at a low cost, the yield prediction with high sigma becomes essential. Taking one 256MB SRAM as an example, if we target a yield of 99.9%, the corresponding failure rate must be controlled below 4.6×10^{-13}. This requires the prediction estimation of the bit cell yield over 7-sigma.

The Monte Carlo (MC) method, which is considered as the golden standard for yield estimation, however, suffers the efficiency issue due to the low probability in sampling the points in the failure region at high sigma [2]. Extensive simulation runs are required and this leads to high computation time and exorbitantly long design cycles [3]. In recent years, extensive efforts have been made to improve the efficiency of high-sigma yield prediction.

One major route to tackle the issue is based on the importance sampling [4-15]. The aim is to construct the distorted probability density function (PDF) which can generate sample data in the failure region with high probability. The optimal distorted PDF is usually circuit-specific and difficult to construct in practice [5], therefore, most existing approaches tried to generate samples in regions where failure events most likely happen. The minimized normalization importance sampling (MN-IS) searched the most probable failure points (MPFP) to construct optimal shift vectors (OSV). The spherical importance sampling (S-IS) [7] and hypersphere sampling (HS) [8] utilized the spherical sampling to find the MPFP. The gradient importance sampling (G-IS) [12] used a gradient-based approach to find the MPFP. However, because

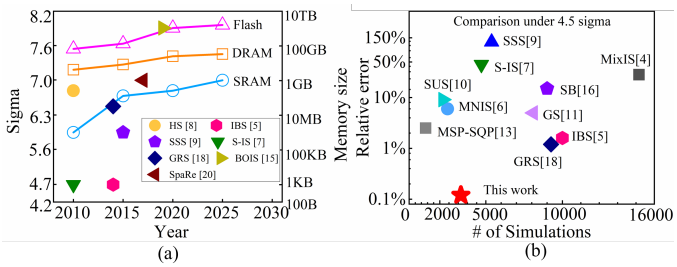

Fig. 1. (a) Trend of the capacity/yield requirements for different memory technologies (hollow points) and the capability of existing methods for high-sigma yield prediction (solid points). (b) Comparison of yield prediction results in efficiency and accuracy between the state-of-the-art and the proposed methods. 6T-SRAM is used for the yield estimation and the estimation is compared under 4.5 sigma. The relative error is defined as the percentage of the difference for yield prediction between the high-sigma prediction methods in literature and the standard MC method.

of the difficulties in obtaining optimized shift vectors, the confidence interval of these methods is quite wide, leading to significant inaccuracies. Such error gets even worse when the failure region has a spatial extent, or when multiple failure regions exist. Constructing a mixture of multiple mean-shift distributions by shifting the mean vectors to various failure regions has been proposed [13, 14] to improve the accuracy. However, these methods become inefficient with higher dimensions due to the complexity in identifying the failure boundary.

Another major route to improving the efficiency of the high-sigma yield prediction is through classification. For example, the Statistical blockade method (SB) [16], which is now widely used in industry, constructs a classifier and applies it to filter the likely-to-fail samples. Only the samples that passed the filter will be used for the circuit simulation. Therefore, the efficiency can be improved by reducing the number of simulation runs. The rigorous mathematics of extreme value theory is used to build sound models of these tail distributions. To reduce the error during the classification for better accuracy, a safety margin is usually applied [16]. More recently, this is further improved with the recursive SB [17] that constructs the conditional and SVM-based nonlinear classifiers. SB forms the foundation of well-accepted industry solutions like Solido or HSPICE HSMC. However, over 100% error can occur when compared to the standard MC [16]. This is because the training of such accurate classifiers is difficult, especially in high dimensions with which the complexity grows exponentially.

At the moment, there lack of methods that can perform high-sigma yield prediction with both satisfactory efficiency and accuracy. As shown in **Fig. 1(b)**, several state-of-the-art methods are compared for both efficiency and accuracy using

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: LIVERPOOL JOHN MOORES UNIVERSITY. Downloaded on July 28,2022 at 11:11:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. This is the author's version which has not been fully edi
content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3193875

the data published in the literature. The comparison is made at 4.5-sigma because of the data availability. Most methods exhibit an error higher than 10%. It is expected that such error can be even larger when predicting at a higher sigma region. Such high error in the yield prediction can misguide the circuit designer in practice [9].

The methods based on the response surface have been considered as one effective solution for accurate yield estimation. Rather than searching for the failure regions, the response-surface-based methods explore the entire input space and construct the equipotential surface, in which the failure probability density can be estimated through areal integration. For example, Weckx et al [18] proposed the global response surface method (GRS) and estimated the failure rate of SRAM with good accuracy. However, this is achieved by sacrificing efficiency due to the time-consuming response surface construction.

This work proposed an equiprobability-based local response surface method (ELRS) to tackle to inefficiency issue of the response surface method while maintaining its advantage in the prediction accuracy for high-sigma yield prediction. By constructing the SRAM circuit with a similar dimension, we show in Fig. 1(b) that the error of the proposed ELRS method can be reduced to only 0.1%, which is 10 times lower than the best record that can be achieved by the state-of-the-art method. Moreover, such highly accurate prediction is achieved with similar simulation runs to the best record in literature as well.

II. Background

A brief overview of problem formulation and response surface will be given first.

A. Problem formulation

Regarding a yield analysis problem, each process parameter is considered one dimension of the variation space. The performances of any circuit can be described with (1). Wherein, $\vec{x} = \{x_1, x_2, \dots, x_N\}$ represents the random process parameters with N dimensions and y represents the circuit performance index under $\vec{x}$. Taking SRAM as an example, $\vec{x}$ can be the threshold voltages for transistors and y is the corresponding static noise margin (SNM). Therefore, the task for the yield estimation is to find the cumulative failure rate, P_f, when y exceeds the pre-set failure criteria, y_0.

$$y(\vec{x}) = g(x_1, x_2, \dots, x_N) \quad (1)$$

Without loss of generality, we suppose these random variables are mutually independent. Thus, the PDF of $\vec{x}$ is

$$q(\vec{x}) = q(x_1, x_2, \dots, x_N) = \prod_{i=1}^{N} q(x_i) \quad (2)$$

According to the results of circuit simulations, the indicator function $I(\vec{x})$ is

$$I(\vec{x}) = \begin{cases} 0, \ pass & (y(\vec{x}) > y_0) \\ 1, \ failure & (y(\vec{x}) \le y_0) \end{cases} \quad (3)$$

Then the probability P_{fail} can be calculated as

$$P_{fail} = \int_{\Omega} q(\vec{x}) d\vec{x} \quad (4)$$

where Ω is the total failure region, including the cases with multiple failure regions. In most cases, the boundary of failure region Ω is unknown and (4) cannot be calculated analytically. Sampling based methods must be applied here to estimate the failure probability. For example, MC method is regarded as ground truth, which estimates the failure rate with a large amount of samples $\vec{x_i}, i = 1,2,..,M$, drawn from the PDF $q(\vec{x})$ of random variables

$$P_{MC} = \frac{1}{M} \sum_{i=1}^{M} I(\vec{x_i}) \quad (5)$$

where P_{MC} is the estimator, M is the number of samples, and $\vec{x_i}$ is the ith sample. The variance of the estimator P_{MC} in (5) can be approximated as

$$V_{MC} = P_{MC} \cdot (1 - P_{MC})/M \quad (6)$$

When MC is applied to estimate P_{fail} that is extremely small, most random samples drawn from the PDF $q(\vec{x})$ do not fail into the failure region Ω. Note that each MC sample is created by running an expensive transistor-level simulation. It, in turn, implies that MC can be extremely expensive for rare event estimation.

B. Response surface

Other than MC, the failure rate can be estimated accurately by constructing the response surface with the input parameters. The Worst Case Distance (WCD) [19] use perturbation analysis to construct an analytical linear response surface model of the input to output mapping. For highly non-linear response surfaces and non-normal input distributions, this approach however falls short of correctly assessing the true yield. To handle these issues, GRS[18] construct a non-analytical response surfaces. By simulating the points that are sampled in an equidistant discrete input space or a non-equidistant discrete input space, the response surface can be constructed. The error of the response surface methods originates from the lower and upper bounds of the input space since it is not possible to traverse all possible input parameter values. Therefore, a large number of points are required for the non-analytic response surface construction to cover most of the input space. This leads to the inefficiency issue.

III. Equiprobability-based Local Response Surface Technique

To improve the efficiency of the response surface method, we propose a dual-step equiprobability sampling method to identify multiple failure regions in an effective way. Therefore, the response surface can be constructed locally for the failure rate estimation. The proposed method is called the equiprobability-based local response surface technique (ELRS) hereafter. We will show that ELRS improves efficiency by reducing the number of points for the response surface construction. In the following, for ease of illustration, a case with two input dimensions is adopted: as shown in **Fig. 2**, two input dimensions (x_1&x_2) form a coordinate plane with four quadrants and the vertical y axis represents the performance index of the circuit.

A. Failure region identification

Since the probability distribution of each input parameter is known, the points with the same joint probability in the input plane can be easily achieved, which forms the equiprobability

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: LIVERPOOL JOHN MOORES UNIVERSITY. Downloaded on July 28,2022 at 11:11:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. This is the author's version which has not been fully edit
content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3193875

curve, as depicted with the dashed curves in **Fig. 2(a)**. Its shape

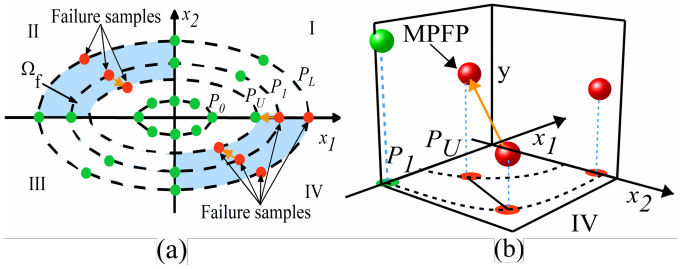

Fig. 2. The flow for response surface construction of the proposed method (2-D example). The failure points and the pass potins are illustrated in red and green cycles. (a) The critical failure regions (blue regions) can be bounded by the lower boundary P_L and the upper boundary P_U. (b) Search the MPFP along the maximum gradient between the failure point with probability P_1 and the good ones (use quadrant IV for example).

of the ellipse is only for illustration purposes.

The procedure starts by randomly choosing a probability, P_0, with which the equiprobability curve can be determined. Then the equiprobability sampling is taken within the quadrants as well as on the axis. For each sample, the circuit performance index, y, is evaluated. If y does not reach the failure criteria, a new equiprobability sampling is performed with the probability P_1, which is 10 times smaller than P_0. This procedure iterates until the first failure point (i.e. y exceeds the criteria, y_0) is found. With the known first failure point and its neighboring pass points, the MPFP can be determined by searching $y = y_0$ along the maximum gradient between the failure point and the passing points, As shown in **Fig. 2(b)**.

The gradient-based searching method works as follows:

- The target function is defined as $L(\theta) = y_0 - g(\overrightarrow{x_1} - \theta)$.
- While$(step > step_{min})$
  1) Extract the gradient of $L(\theta_{n_{iter}-1})$ by computing independently the derivative with respect to the $\overrightarrow{x_i}$ of each transistor with its normalized $\vec{x}$.
  2) Find the vector v of the norm step with the steepest positive slope based on the gradient $\frac{\Delta L(\theta_{n_{iter}-1})}{\Delta \vec{x}}$.
  3) Simulate the new sample $L(\theta_{n_{iter}}) = L(\theta + v + \overrightarrow{x_1})$
  4) If $(L(\theta_{n_{iter}}) > 0)$
     Set $\theta = \theta + v$
     Else if $(L(\theta_{n_{iter}}) < 0)$
     Set step = step/2 and go back to 2)
  5) $n_{iter} = n_{iter} + 1$

Once the highest probability P_U for the failure to occur is found. The upper boundary of the failure region can be defined by P_U. In principle, all the regions outside of the upper boundary form the failure region. However, in practice, we define the lower boundary of the failure region with the probability, $P_L = 0.001 * P_U$, as to be explained in Section II.C.

Bounded by the two equiprobability curves with the probability of P_L and P_U, the confined region within the input space formed the failure region Ω_f. Since the failure region can vary in different quadrants, the above procedure can be performed independently in each quadrant, and thus all the failure regions in the input space can be found.

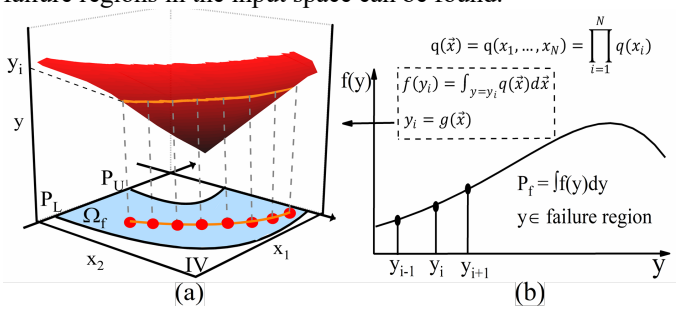

Fig. 3. (a) The PDF $f(y_i)$ for a given performance index y_i, can be evaluated by integrating $q(\vec{x})$ over all the possible samples for y_i. (b) The failure rate P_f can be numerically calculated by a cumulative sum of integration of $f(y)$.

B. Failure rate estimation with equipotential surface sampling

In the identified failure region, the response surface can be constructed by mapping the inputs to the output with SPICE simulations. **Fig. 3(a)** illustrates the response surface for the circuit performance y, as the function of input parameters$\vec{x}$. The samples with the same performance are chosen that forms the equipotential curve in the input plane. Circuit failures can occur with high probability in the failure region, and therefore much fewer points are required. This is the key to efficiency while maintaining good estimation accuracy.

The corresponding PDF of the circuit performance, $f(y)$, can be calculated by propagating the PDF of the input parameters in the equipotential curve as shown in **Fig. 3(b)**. By performing integrals along the equipotential surface $y = g(\vec{x})$, the probability density $f(y)$, can be evaluated, as described in (7):

$$f(y) = \int_{y=g(\vec{x})} q(\vec{x}) d\vec{x} \quad (7)$$

For a two-dimensional problem, (7) performs line integrals and for higher dimensional space, multiple integrals need be calculated by successive integrals based on Fubini's theorem [21]. By integration of the PDF, the cumulative distribution function (CDF) of the output parameter can be obtained, which is the failure rate, P_f, as described in (8).

$$P_f(y_t) = \int \int q(\vec{x}) d\vec{x} dy, y_t \in \Omega_f \quad (8)$$

C. Evaluation for Sources of Error

Since the response surface has no closed form solution, it cannot be constructed in a finite number of simulation runs. Truncation of the response surface is inevitable, which can introduce error as described by (9).

$$err = \int \int q(\vec{x}) d\vec{x} dy \ , \vec{x} \in \Omega - \Omega_f \quad (9)$$

Where Ω and Ω_f are the entire and the failure regions in the input space respectively.

In the proposed ELRS method, the failure region is truncated by defining the lower boundary, P_L. This error is reduced by selecting P_L which equals $0.001*P_U$, making it possible to reach deep into the tail. Moreover, in order to minimize the possibility of missing any failure points, P_U is chosen 1.1 times higher than the calculated value. The accuracy achieved in Section III suggests that the accuracy loss can be suppressed.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: LIVERPOOL JOHN MOORES UNIVERSITY. Downloaded on July 28,2022 at 11:11:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. This is the author's version which has not been fully edit
content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3193875

IV. Experiments and Results

In this section, we use the standard 6T SRAM circuit to verify the accuracy and efficiency of our proposed ELRS method. **Fig. 4(a)** depicts the structure of SRAM we designed

TABLE I
EFFICIENCY AND ACCURACY COMPARISON OF THE PROPOSED METHOD AND STATE-OF-THE-ART METHODS AT 4.5 SIGMA CONDITION.

Method	PoI	# of simulations	Dimension	Error
IBS[5]	RNM	100000	6	1.65
MNIS[6]	RNM	2506	6	6%
SUS[10]	RNM	2211	6	9.1%
GS[11]	RNM	8100	6	5%
MSP-SQP[13]	RNM	1078	6	2.5%
GRS[18]	RNM	9261	6	1.2%
This Work	RNM	3375	6	0.12%

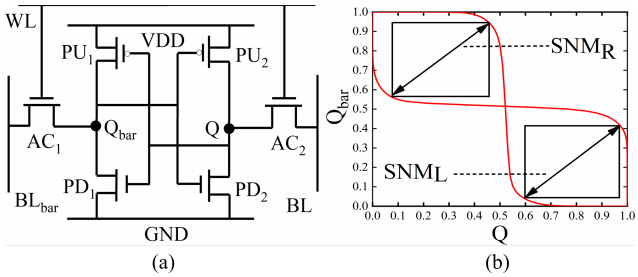

Fig. 4. (a) The 6T-SRAM cell used for simulation. (b) SNM margins are obtained by analyzing the butterfly curve during hold mode, where the word lines are set at 0V and the bitlines at VDD.

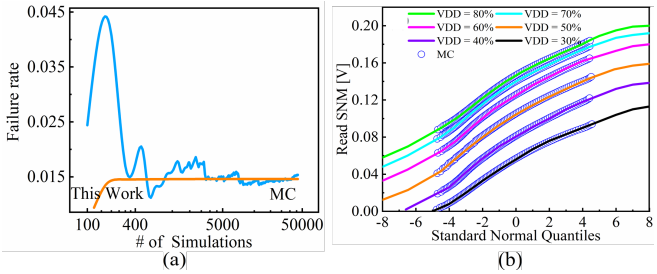

Fig. 5. Comparison of the failure rate estimation between the standard MC method and the proposed ELRS method for the 6T-SRAM circuit (a) under the fixed VDD of 0.5V at low-sigma and (b) under various VDD and sigma. The proposed ELRS method can easily reach into the tail.

with the 7-nm process node [22]. The SRAM cell consists of 4 core transistors (PU_1, PU_2, PD_1, PD_2) and two access transistors (AC_1, AC_2). The Read static noise margin (RNM) represents the maximum noise on the storage node Q, that can be tolerated during its read operation. For ease of comparison, this is used as the circuit performance index for the yield estimation. As shown in **Fig.4(b)**, RNM can be defined as the minimum of the two Seevinck squares (SNM_L and SNM_R) that can fit inside the eyes of the butterfly curve. With the random variation of the threshold voltages of these transistors, the RNM reduces and causes the read errors.

For the yield prediction with ELRS method, the input parameters are the threshold voltages V_{th} of all the six transistors, which exhibit device-to-device process variation and are modelled with the independent Gaussian distributions.

Fig. 5(a) compared the failure rate estimation at 2.2-sigma between the standard MC method and the proposed ELRS

TABLE II
EFFICIENCY AND ACCURACY COMPARISON OF THE PROPOSED METHOD AND GLOBAL RESPONSE SURFACE (GRS) METHOD [18] AT 7 SIGMA CONDITION.

Scenario	Method	Total runs (Failure samples)	Probability	Accuracy improvement
SNM	GRS[18]	225(8)	6.24E-12	11×
	This Work	197(27)	7.20E-11	
RNM	GRS[18]	9261(232)	2.6E-14	61×
	This Work	4856(2240)	1.6E-12	

TABLE III
EFFICIENCY AND ACCURACY COMPARISON OF THE PROPOSED METHOD AND STATISTICAL BLOCKADE (SB) METHOD [16] AT 4.526 SIGMA CONDITION.

Scenario	Standard MC (1M sims)	Method	Total runs	Probability	Relative Error
Write time	3.005E-6	SB[16]	5379	6.745E-6	124%
		This Work	5133	3.001E-6	0.13%

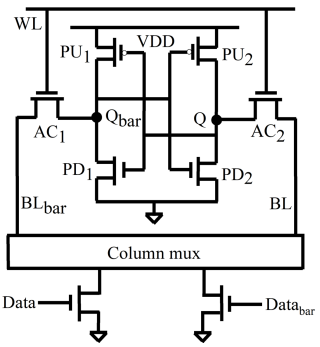

Fig. 6. The circuit for extracting the write time: 6-T SRAM cell with column mux and write drivers are used which is the same as used in ref. 16. The input dimension is 9 including the Vth on 8 transistors and the global oxide thickness (t_{ox}) variation, which also kept the same as ref. 16.

method. It is clear that the ELRS method converges quickly after around 400 simulation runs, which is almost 10 times faster than the standard MC method. In addition, the converged failure rate from the ELRS method also agrees very well with the MC, indicating its good accuracy.

Fig. 5(b) further evaluated the capability of the ELRS method for the high-sigma yield estimation. The failure rates of the circuits under different operating voltage (VDD) are estimated up to 8-sigma using the ELRS method. When comparing with the result from MC which is carried out up to 4.5-sigma, the good agreement is again obtained.

The comparison is further made with state-of-the-art methods. For a fair comparison, the same circuit topology, input dimensions and the performance index are used. The comparison is made at 4.5-sigma where we can find most of the data in the published literature. As shown in **Table I**, the error of the proposed method is only 0.12%, which is almost 10 times more accurate than the best one (i.e. GRS) in these state-of-the-art methods. In addition, ELRS use only 3375

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: LIVERPOOL JOHN MOORES UNIVERSITY. Downloaded on July 28,2022 at 11:11:18 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. This is the author's version which has not been fully edi
content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3193875

simulation runs, which is also among the lowest number of simulation, which confirms its good efficiency.

We further compared our ELRS method with the standard response-surface-based method, GRS, at 7 sigma. Both the read SNM (RNM) and the hold SNM (SNM) are used as the performance indices which is the same as the ones used in ref. 12. As shown in **Table II**, when running a similar number of simulations, the proposed ELRS method can find 11x and 61x more failure events when comparing with the GRS method, suggesting ELRS exhibits better accuracy. Such accuracy improvement is because the sampling in GRS spreads throughout the entire region, while the sampling in ELRS is only in the failure region, and therefore more points can cover the failure region with a similar amount of sampling.

Finally, the comparison is made between the ELRS method and the statistical blockade method (SB) [16]. The latter is widely adopted in commercial tools such as Solido or HSPICE HSMC. Again, for a fair comparison, we constructed the same circuit (**Fig. 6**) and used the same write time as the circuit performance index as ref. 16: a 6-T SRAM cell, with bit-lines connected to a column multiplexor and a non-restoring driver. There are 9 input dimensions, including the threshold voltage of all eight transistors and one global gate oxide thickness variation. The distributions of all 9 variations are set the same as the distributions in ref. 16. The error is defined by comparing the predicted yield using these two methods with the yield from the MC method. As shown in **Table III**, the error of the proposed ELRS method is only 0.13%, which is almost 1000 times more accurate than SB. In terms of efficiency, ELRS use about 5133 simulation runs, which is also comparable to the SB method. Therefore, ELRS can be an good candidate for future high-sigma yield prediction.

V. Conclusions

In this letter, a novel equiprobability-based local response surface method is proposed for the high-sigma yield estimation. Rather than sampling in the entire input space, the proposed method relies on equiprobability surface sampling to quickly find the failure regions, on which the response surface can be directly constructed for yield estimation. Therefore, the proposed method can perform high-sigma yield estimation with both high accuracy and efficiency. Demonstrated on 6T-SRAM, we show that the proposed method exhibits more than 10 times improvement in accuracy when comparing with the state-of-the-art while maintaining the efficiency to the best record in literature.

References

[1] D. Sylvester, K. Agarwal, and S. Shah, “Variability in nanometer CMOS: Impact, analysis, and minimization,” Integr., VLSI J., vol. 41, no. 3, pp. 319–339, 2008, doi: 10.1016/j.vlsi.2015.12.002.

[2] J. Yao, Z. Ye and Y. Wang, "An Efficient SRAM Yield Analysis and Optimization Method With Adaptive Online Surrogate Modeling," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 23, no. 7, pp. 1245-1253, July 2015, doi: 10.1109/TVLSI.2014.2336851.

[3] Ji, Z., Chen, H. & Li, X. Design for reliability with the advanced integrated circuit (IC) technology: challenges and opportunities. Sci. China Inf. Sci. 62, 226401 (2019).

[4] R. Kanj, R. Joshi and S. Nassif, "Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events," in Proc. DAC, jul. 2006, pp. 69-72 , doi: 10.1145/1146909.1146930.

[5] J. Yao, Z. Ye and Y. Wang, "Importance Boundary Sampling for SRAM Yield Analysis With Multiple Failure Regions," IEEE Trans. Comput-Aided Des. Integr. Circuits Syst., vol. 33, no. 3, pp. 384-396, March 2014, doi: 10.1109/TCAD.2013.2292504.

[6] L. Dolecek, M. Qazi, D. Shah and A. Chandrakasan, "Breaking the simulation barrier: SRAM evaluation through norm minimization," in Proc. Int. Conf. Comput. Aided Design, 2008, pp.322-329, doi:10.1109/ICCAD.2008.4681593.

[7] M. Qazi, M. Tikekar, L. Dolecek, D. Shah and A. Chandrakasan, "Loop flattening & spherical sampling: Highly efficient model reduction techniques for SRAM yield analysis," in Proc. DATE, 2010, doi: 10.1109/DATE.2010.5456940.

[8] T. Date, S. Hagiwara, K. Masu and T. Sato, "Robust importance sampling for efficient SRAM yield analysis," in Proc. ISQED, Mar. 2010, doi: 10.1109/ISQED.2010.5450410.

[9] S. Sun, X. Li, H. Liu, K. Luo and B. Gu, "Fast Statistical Analysis of Rare Circuit Failure Events via Scaled-Sigma Sampling for High-Dimensional Variation Space," TCAD, 2015, doi: 10.1109/TCAD.2015.2404895.

[10] S. Sun and X. Li, "Fast statistical analysis of rare circuit failure events via subset simulation in high-dimensional variation space," in Proc. ICCAD, Nov. 2014, pp. 324-331, doi: 10.1109/ICCAD.2014.7001370.

[11] S. Sun, Y. Feng, C. Dong and X. Li, "Efficient SRAM failure rate prediction via Gibbs sampling," TCAD 2012, doi: 10.1109/TCAD.2012.2209884.

[12] T. Haine et al., “Gradient importance sampling: An efficient statisticalex traction methodology of high-sigma sram dynamic characteristics,” in DATE, 2018, doi: 10.23919/DATE.2018.8342002.

[13] M. Wang, C. Yan, X. Li, D. Zhou and X. Zeng, "High-Dimensional and Multiple-Failure-Region Importance Sampling for SRAM Yield Analysis," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 25, no. 3, pp. 806-819, March 2017, doi: 10.1109/TVLSI.2016.2601606.

[14] Wu, Wei, Srinivas Bodapati, and Lei He. "Hyperspherical clustering and sampling for rare event analysis with multiple failure region coverage."P roceedings of the 2016 on International Symposium on Physical Design. ACM, 2016, doi: 10.1145/ACM.2872334.2872360.

[15] D. D. Weller, M. Hefenbrock, M. S. Golanbari, M. Beigl and M. B. Tahoori, "Bayesian Optimized Importance Sampling for High Sigma Failure Rate Estimation," in Proc. DATE, 2019, pp. 1667-1672, doi: 10.23919/DATE.2019.8714879.

[16] A. Singhee and R. A. Rutenbar, "Statistical Blockade: A Novel Method for Very Fast Monte Carlo Simulation of Rare Circuit Events, and its Application," in Proc. Conf. Design, Autom. Test Eur,. 2007, pp. 1-6, doi: 10.1109/DATE.2007.364490.

[17] Singhee, Amith, et al. "Recursive statistical blockade: An enhanced technique for rare event simulation with application to SRAM circuit design." 21st International Conference on VLSI Design (VLSID 2008). IEEE, 2008, doi: 10.1109/VLSI.2008.54.

[18] P. Weckx et al., "Non-Monte-Carlo methodology for high-sigmasimulations of circuits under workload-dependent BTI degradation—Application to 6T SRAM," in Proc. IRPS, 2014, pp. 5D.2.1-5D.2.6, doi: 10.1109/IRPS.2014.6860671.

[19] K. J. Antreich, H. E. Graeb and C. U. Wieser, "Circuit analysis and optimization driven by worst-case distances," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 13, no. 1, pp. 57-71, Jan. 1994, doi: 10.1109/43.273749.

[20] M. Malik, R. V. Joshi, R. Kanj, S. Sun, H. Homayoun and T. Li, "Sparse Regression Driven Mixture Importance Sampling for Memory Design," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 26, no. 1, pp. 63-72, Jan. 2018, doi: 10.1109/TVLSI.2017.2753139.

[21] Ponnusamy S . Foundations of mathematical analysis[J]. Springer Basel Ag, 2010:71-113.

[22] V. Vashishtha, M. Vangala, P. Sharma and L. T. Clark, "Robust 7-nm SRAM design on a predictive PDK," Porc. ISCAS, 2017, pp. 1-4, doi: 10.1109/ISCAS.2017.8050316.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: LIVERPOOL JOHN MOORES UNIVERSITY. Downloaded on July 28,2022 at 11:11:18 UTC from IEEE Xplore. Restrictions apply.