# Statistical Diagnosis of Rare Event Failure Region by Smart Sample Selection and In-Spec Parameter Guidance

*Abstract*—**In this paper, we propose a new statistical diagnosis method for rare failure event region determination and failure probability estimation. Our method consists of two new design efficiency improvement techniques to mitigate the high computational cost and the lack of design guidance of the existing iterative blockage based method. First, the new approach employs a smart sample selection scheme, which can consider the effectiveness of samples and well-coverage for the parameter space. As a result, it can reduce an additional simulation costs by pruning less effective samples while keeping the accuracy of failure estimation. Second, the new approach identifies the failure regions in terms of parameters to provide a good design guideline for in-spec circuits. Applying variance based feature selection plays an important role in finding the dominant parameters. A quasi-random sampling with dominant parameters is then applied to determine in-spec boundaries of those parameters. In addition, we also provide complete formula for the probability determinations of failure regions in the iterative failure region searching framework. We demonstrate the advantage of our proposed method using two test benches: 6T-SRAM reading failure diagnosis with 27 process parameters, charge pump operation failure diagnosis in a PLL circuit with 81 process parameters. Experimental results show that the new method is 159x faster than traditional Monte Carlo method with only 0.0025% error. Our approach also provides the precise in-spec guidance of diverse parameters with 0.02% estimation error. Furthermore, the new method reduces the simulation cost by 2x than a recently published approach with same accuracy level.**

## I. Introduction

As the CMOS technology scaling continues, the performance uncertainties related to the process variation have become a major concern for IC development [1]. Many IC components such as SRAM bit-cells need to be tremendously robust as they are duplicated in millions [2]. Such modules require accurate statistical failure analysis in rare event region. However, the traditional Monte Carlo (MC) based statistical analysis method faces a challenge as it may require a million times of simulations [3]. Various statistical techniques have been developed to overcome this problem in the literature [3]–[11].

One way is by means of importance sampling (IS) [4], which consists of two steps. First, it shifts the mean value of the initial performance distribution and place it on interested failure region. The standard deviation based on the shifted mean is recalculated by considering samples only placed on the failure region. The new probability density function (PDF) is generated based on updated mean and standard deviation so that more samples in the failure region can be drawn. Work in [11] applied the mixture of IS for cross-validation of multiple failure regions due to disjointed process parameters. However, these approaches can only estimate single performance metric. Multiple important samplings are required to estimate more metrics. Also, it is difficult to calculate the failure probability for a generated distribution by IS.

The statistical blockade (SB) is another effective approach for improving the performance of MC [3]. The idea of this approach is using a threshold bound to separate an interested failure region from the whole distribution so that it can *block* some unnecessary sampling and simulation for efficiency improvement. This method builds a supervised learning model with the threshold bound and initial simulation data, which is named as ”classifier”, to recognize failure samples. Thus, later samples that tend to be placed in the failure region can be captured without simulation. This approach was improved by using the recursive statistical blockade (RSB) scheme to locate the rare event failure region in an iterative way [5]. This method can improve the accuracy of the classifier iteratively by increasing the number of samples in the failure region of interest. However, this method can incur the significant extra cost as it needs more samples for the simulation. Recently, Wu *et al.* [7] applied a nonlinear SVM classifier to model nonlinear and multiple disjoint failure regions of circuits. This method used the generalized Pareto distribution (GPD) fitting for tail of distribution to model failure probability in each iteration. However, this method cannot investigate further failure region without rerunning whole algorithm. The reason is that the pruned parameters depending on the initial samples cannot remained as important ones in the failure region, which keeps changing. Also, the previous approaches cannot provide the design guideline in terms of the design parameters to explicitly avoid performance fails, which are important for improving the yield of circuit.

In this paper, we propose a new statistical failure region diagnosis method. The new method is based on recursive statistical blockade [5] to gradually locate the failure regions. But the proposed method consists of two new techniques. First, it introduces the smart sample selection scheme, which can consider an effectiveness of samples and well-coverage for parameters space so as to further reduce the simulation cost. The selection process can reduce number of samples significantly while it still keeps sufficient representative samples for accurate failure probability analysis. Second, the new approach can generate safe boundaries of parameters to satisfy the design specification. The new method first applies variance based feature selection to find the dominant parameters. A quasi-random sampling with dominant parameters is then used to quickly determine in-spec boundaries of those parameters. In addition, we also provide complete formulas to determine the probabilities of failure regions in the iterative failure region searching framework.

The presented method has been tested on 6T-SRAM single-bit cell and circuit pump circuit in phase-locked loop (PLL) with each 27 and 81 dimension of process parameters. Ex-

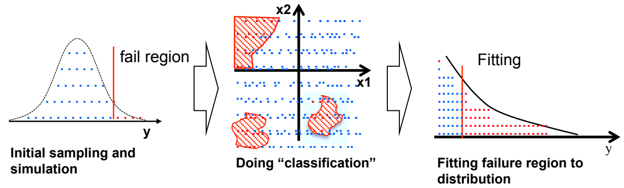
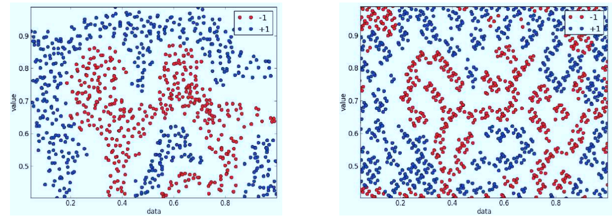



Fig. 1. General flow of statistical blockade approach



(a) Reference (b) Classification result

Fig. 2. The accuracy of classification by GRBF

perimental results show that the new method is 63-159x faster than traditional Monte Carlo method with only 0.0025-0.005% error. Our approach also provides the precise in-spec guidance of process parameters with 0.02-1.2% estimation error. Furthermore, the new method reduces the simulation cost by 1.7-2x than recently published approach with the same level of accuracy.

The reminder of this paper is organized as follows. We review some major techniques of the blockade-based method for failure analysis in section 2. In Section 3, we explain the proposed iterative diagnosis algorithm with an overall flow and detailed descriptions of the smart sampling and in-spec guidance techniques. We also introduce the mathematical framework to calculate the probability of iteratively updated failure region. Section 4 shows the experimental results for verifying the accuracy and efficiency of the proposed method. Section 5 provides conclusions of this paper.

## II. Background

We first briefly review the concept of the statistical blockade (SB) approach for fast estimation of properties of rare event region. A general framework of SB is shown in Fig.1. This method starts with drawing initial samples with uniform or normal distribution to capture a crude shape of performance distribution by circuit simulation. The classifier can be built by training with initial simulation data. Once we obtain the classifier, samples that tend to fall into the failure criteria can be identified without actual simulation. With these filtered samples, SB calculates the probability of failure region by fitting samples in proper distribution model. Thus, ”classification” and ”failure probability calculation” are both key steps in the SB method. The rest of this section describes these two steps.

### A. Classification

The classification filters samples as likely-to-fail samples for the circuit simulation. Building a classifier needs a training step with initial samples to render real shapes of the failure region. The classifier can then reduce the number of samples for real simulation because it can estimate mathematically whether a sample makes failure or not. However, the classifier is not capable of fully replacing the simulator due to its accuracy. So, a marginal filtering approach is used to improve the accuracy of classification [3], [5]–[7]. Some methods use relaxed threshold bounds instead of a real failure criterion to capture more samples to minimize classification error. Meanwhile, it is not sufficient to use the simple and linear classifier due to the nonlinearity of the failure region [3], [5], [6]. The gaussian radial basis function kernel (GRBF) is the suitable supervised learning model for nonlinear classification [7], [12]. GRBF can make curved boundaries to recognize patterns scattered with nonlinear and disjointed manner based on radial distances between samples. An example of GRBF classification with two input parameters is shown in Fig. 2. Even solution spaces are neither uniformly nor relevantly formed, GRBF can still recognize accurate boundaries of two different spaces.

### B. Failure probability calculation

The failure samples should be fitted to a particular distribution form in order to calculate the probability of the failure region. Suppose that the simulation result can be fitted to a normal distribution in form of its PDF. If there is a threshold $t$ which separates a tail region from the whole distribution, the conditional cumulative density function (CDF) of the tail distribution can be written as follows:

$$P_t(t_c) = P(Y \geq t_c \mid Y \geq t) = \frac{P(t_c) - P(t)}{1 - P(t)} \tag{1}$$

where $P_t(t_c)$ means the failure probability decided by $t_c$. Once we have a suitable fitting model for CDF of the failure region with a failure bound $t_c$, the failure probability with given values can be calculated as:

$$P(Y \geq t_c) = [1 - P(Y \leq t)] \cdot P_t(t_c) \tag{2}$$

In the several generalized extreme value distributions, GPD is one of the most accurate model to describe tail distribution corresponding to failure region [13]. With the location parameter $\mu$, the scale parameter $\sigma$ and the shape parameter $\xi$, CDF of the failure region can be formulated by GPD fitting.

$$\begin{aligned} &P_t(x) = G_{(\xi,\mu,\sigma)}(x) \\ &= \begin{cases} 1 - \left(1 + \dfrac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0 \\ 1 - e^{\frac{-(x-\mu)}{\sigma}} & \text{for } \xi = 0 \end{cases} \end{aligned} \tag{3}$$

The location parameter $\mu$ means a starting point of GPD and it is corresponded to the threshold $t$ of the tail distribution. Consequently, the failure probability with given threshold $t$ and failure bound $t_c$ can be computed as follows.

$$P(Y > t_c) = [1 - P(Y \leq t)] \cdot G_{(\xi,t,\sigma)}(t_c) \tag{4}$$

To approximate the rest of parameters for GPD fitting, we use the maximum likelihood estimation [14].

## III. Proposed new failure region diagnosis method

The overall flow of the proposed method, which is called ”IFRD” (Iterative Failure Region Diagnosis), is illustrated in Fig.3. Our algorithm starts with given data, such as process variations and some parameters for failure region detrmination. The failure criteria $t_c$ denotes the reference value of failure and the percentile bound $p$ to calculate threshold in each $i_{th}$ iteration. The first step is to perform initial MC sampling and simulation to capture overall circuit performance metrics. After this, the relaxed threshold $t_i$ can be obtained to separate a failure region from main PDF and the probability of this region is $P(Y > t_i) = p$. The classifier can then be modeled with the simulation result of the initial samples. In the classification step, the GRBF nonlinear classifier is used for accurate sample filtering. With the simulation result and the classifier, the new method can calculate the in-spec conditions of process parameters to achieve targeted yield in $i_{th}$ iteration. At the same time, the algorithm generates $m^i * n$ MC samples, which will be filtered by the classifier $C_i$ to likely-to-fail samples based on $t_i$. Then, the smart sample selection can be employed to further reduce the number of samples for actual simulation. After the simulation, the failure probabilities $P(Y > t_i)$ are updated by GPD fitting. Our approach iterates the above whole procedure with the updated threshold bound $t_i$ by percentile bound $p$, and the increased number of MC samples to calculate the failure probability $P(Y > t_i)$. Finally, it finishes when the threshold bound meets the given failure criterion $t_c$.

The rest of this section explains in detail the two major contributions of the proposed method: (1) Smart sample selection (2) In-spec guidance of process parameters. We also present the iterative formulation of failure probabilities.

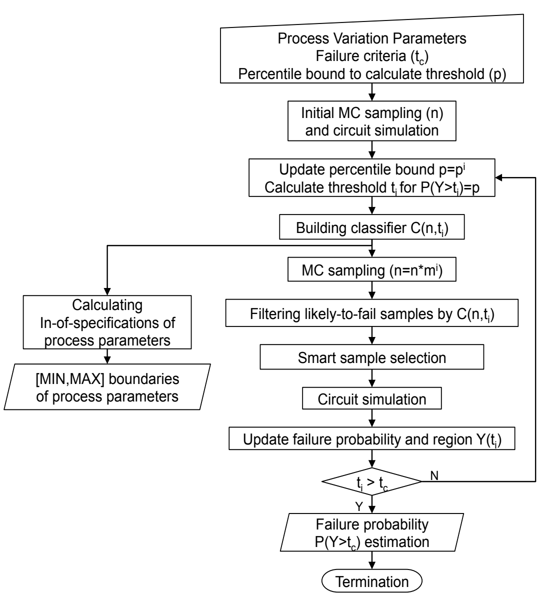



Fig. 3. Proposed iterative failure region diagnosis flow

### A. Smart sample selection scheme

The simulation cost is a major bottleneck in the statistical analysis of the circuit. The proposed iterative failure diagnosis method can lead to an extra simulation cost in each iteration. To mitigate this problem, we propose the smart sampling scheme, which significantly reduce the number of samples required. Two factors can be considered for the sample selection process. The effectiveness of the sample group is the first factor. Each sample consists of the combination of process parameters, which affect differently on simulation results Therefore, the sensitivity of each parameter should be considered for the sample selection. The proposed selection method calculates correlation coefficients between parameters and simulation results for the sensitivity analysis as follows:

$$\begin{aligned} &x_n \in R^m, y_n \in R \\ &\rho_{x_n,y_n} = \frac{cov(x_n, y_n)}{\sigma_{x_n}\sigma_{y_n}} \in R^m \end{aligned} \tag{5}$$

where $x_n$,$y_n$ are $n$ process parameter vectors with $m$ dimension and $n$ simulation results, respectively.

The second factor is the coverage ratio of parameters search space by selected samples. The diversity of samples can be calculated by euclidean distances with the reference sample, which is the median from simulation results. Samples around the median can be chosen as the median is located on the highest probability region in the distribution of simulation results. Simultaneously, samples found in the boundary region of the search space can be selected as these samples represent the maximum and minimum conditions of parameters. Thus, the proposed sampling method can calculates two distance factors to cover both central and boundary regions of the search space as follows:

$$\begin{aligned} &\tilde{y} = median(y_n) \\ &x_{ref} = x : \tilde{y} = f(x), x \in R^m \\ &D_{central}(x) = \frac{1}{\left|\frac{x - x_{ref}}{max(x_n) - min(x_n)}\right|} \in R^m \\ &D_{boundary}(x) = \left|\frac{x - x_{ref}}{max(x_n) - min(x_n)}\right| \in R^m \end{aligned} \tag{6}$$

We define the sample's weight by multiplying the correlation coefficient to distance factors of each sample as

$$\begin{aligned} &W_{central}(x) = {\rho_{x_n,y_n}}^T \cdot D_{central}(x) \\ &W_{boundary}(x) = {\rho_{x_n,y_n}}^T \cdot D_{boundary}(x) \end{aligned} \tag{7}$$

The final set of selected samples can be obtained as

$$E(n,r) = S(\frac{nr}{2}, W_{central}(x_n)) \cup S(\frac{nr}{2}, W_{boundary}(x_n)) \tag{8}$$

where $S(n, W(x_n))$ is the set of $n$ samples $x_n$ sorted by $W(x_n)$ and $r$ is the selection ratio, which determines the number of selected samples.

### B. In-spec guidance of process parameters

In order to improve yield of a circuit, designers need to know good ranges of process parameters with regards to the circuit pearformance specification. However, applying all possible combinations of parameters is impossible due to exponential possibilities with sizes of parameters.

The proposed method ranks priorities of process parameters based on its variances because the parameters with huge variance easily lead to spread samples in a failure region. We can easily calculate the variances and MIN,MAX ranges of parameters as by-products from simulation data of updated failure region in our iterative framework. Our method redraws samples with the distributions of high ranked parameters. The rest of parameters can be assigned as predefined nominal values. We use SOBOL algorithm [15] to redraw these sample. It uses a quasi-random low-discrepancy sequence, so these sample can cover the search spaces of parameters more uniformly than the previous samples for simulation. As a result, we can generate samples with not only reduced dimensionality but also well-coverage of the failure region. The classifier with the updated threshold can filter out these samples to determine pass or fail condition of process parameters. With the classification result of samples, the proposed method induces *if-then* rules from the highest-ranked parameters and finally, the combination of in-spec conditions for all parameters can be calculated. The overall steps of the new in-spec guidance method are explained in Fig. 4.

### C. *Iterative calculation of failure probabilities*

In the proposed iterative diagnosis method, the probability of updated failure region should be repeatedly calculated. In this subsection, we derive the complete formula, which was not given explicitly in existing works [3], [5].

The failure probability with given threshold $t$ in (3) (which denotes $[1 - P(Y \leq t])$ can be replaced by the PDF from simulation result with the threshold $t_i$ in each iteration. So, the equation can be rewritten with $P_{MC}(Y \geq t_1)$, which denotes the PDF of initial MC sampling simulation with the threshold $t_1$.

$$\begin{aligned} &P(Y \geq t_c) = P_{MC}(Y \geq t_1) \cdot P(Y \geq t_c | Y \geq t_1) \\ &P(Y \geq t_c | Y \geq t_1) = \frac{P(Y \geq t_c, Y \geq t_1)}{P(Y \geq t_1)} \end{aligned} \tag{9}$$

The conditional probability part in (9) can be estimated by GPD fitting using simulated failure samples. Therefore, we can reformulate (9) as

$$P(Y > t_c) = P_{MC}(Y \geq t_1) \cdot P_{IS}(Y \geq t_c | Y \geq t_1) \tag{10}$$

where $P_{IS}$ represents the conditional probability in the updated distribution by GPD fitting. If the proposed method iterates twice with $t_1$ and $t_2$ as threshold bounds, failure samples for the second simulation were filtered out again by the updated classifier based on $t_2$. The failure probability with $t_2$ can be computed by multiplying the first failure probability with $t_1$. Thus, the failure probability in each step can be calculated as

$$\begin{aligned} &P(Y \geq t_c) = P_{(2)}(Y \geq t_c) \\ &P_{(1)}(Y \geq t_2) = P_{MC}(Y \geq t_1) \cdot P_{IS(1)}(Y \geq t_2 | Y \geq t_1) \\ &P_{(2)}(Y \geq t_c) = P_{MC}(Y \geq t_1) \cdot P_{IS(1)}(Y \geq t_2) \\ &\qquad\qquad \cdot P_{IS(2)}(Y \geq t_c | Y \geq t_2) \\ &P_{IS(i)}(Y \geq t_c | Y \geq t_i) = \frac{P_{IS(i)}(Y \geq t_c)}{P_{IS(i)}(Y \geq t_i)} \end{aligned} \tag{11}$$

Without the loss of generality, we can formulate the iterative failure probability calculation as follows. Suppose that the iteration loops $k$ times,

$$\begin{aligned} &P_{(i)}(Y \geq t_c) \\ &= \begin{cases} P_{MC}(Y \geq t_i) \cdot P_{IS(i)}(Y \geq t_c | Y \geq t_i) & \text{for } i = 1 \\ P_{MC}(Y \geq t_i) \cdot \prod_{i=1}^{k-1} (P_{IS(i)}(Y \geq t_{i+1})) & \\ \qquad \cdot P_{IS(k)}(Y \geq t_c | Y \geq t_k) & \text{for } i > 1 \end{cases} \end{aligned} \tag{12}$$

The final failure probability can be obtained by recursively applying (12) in each iteration.

## IV. Experimental results and discussions

The proposed method (IFRD) was implemented in Python 2 and tested on a Linux workstation with 32 CPUs (2.6GHz Xeon processors) and 64GB RAM. The performance and accuracy of proposed method has been evaluated by verifying the failure rate of 6T-SRAM single-bit cell and charge pump circuit in PLL, which are highly replicated instances for system-on-chip (SOC) designs. Both of circuits were designed with BSIM4 transistor model and simulated in NGSPICE [16]. Table. I shows 9 major process parameters of MOSFET. We performed four different methods (MC, REscope [7], RSB [5] and IFRD) to compare their accuracy and performance.

TABLE I
Process parameters of MOSFET

| **Variable name** | **Std($\sigma$)** | **Unit** |
|---|---|---|
| Flat-band voltage($V_{fb}$) | 0.1 | $V$ |
| Gate oxide thickness($t_{ox}$) | 0.05 | $m$ |
| Mobility ($\mu_0$) | 0.1 | $m^2/Vs$ |
| Doping concentration at depletion $N_{dep}$) | 0.1 | $cm^{-3}$ |
| Channel-length offset ($\Delta L$) | 0.05 | $m$ |
| Channel-width offset ($\Delta Q$) | 0.05 | $m$ |
| Source/drain sheet resistance($R_{sh}$) | 0.1 | $Ohm/mm^2$ |
| Source-gate overlap unit capacitance($C_{gso}$) | 0.1 | $F/m$ |
| Drain-gate overlap unit capacitance($C_{gdo}$) | 0.1 | $F/m$ |

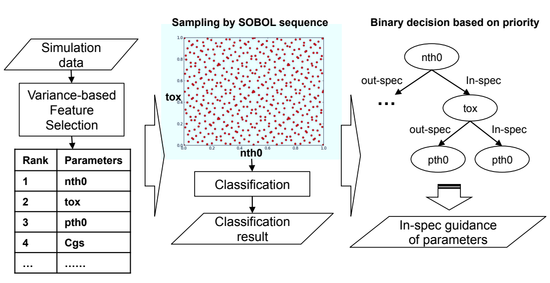



Fig. 4. In-spec guidance of parameters calculation flow 1) Feature selection 2) Sampling and Classification 3) Calculate the boundaries for in-spec conditions

### A. *6T-SRAM failure rate diagnosis*

The schematic design of the single-bit 6T-SRAM cell is shown in Fig. 5. The read operation of 6T-SRAM fails when the voltage gap between $BL$ and $\bar{BL}$ is not enough to be determined by sense amplifiers in certain period. We measure the delay of discharging $\bar{BL}$ as the failure criterion. Our experimental setup for initial conditions are: $\bar{Q}$='L', $\bar{Q}$='H', $BL$ and

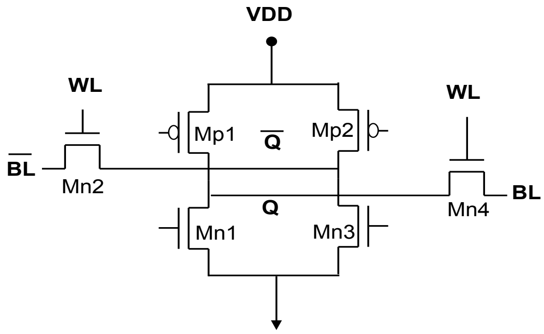



Fig. 5. The schematic of 6T-SRAM single-bit cell

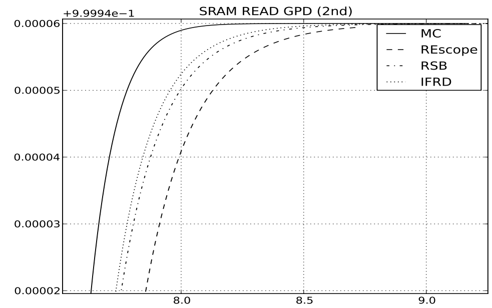



Fig. 6. Estimating the CDF of 6T-SRAM readtime around 5-sigma region

$\bar{B}L$ = 'H.' When $WL$ turns on, $\bar{B}L$ is discharged by MN2 and MN1 and $BL$ charged by MP2. So, the number of the process parameters is $27(3*9)$ in this reading operation. The initial number of sampling for capturing the circuit behavior is 2,000. We set the failure criterion $t_c$ as P(Y$\geq t_c$)=0.00023, which means 5-sigma in terms of the yield level. The proposed method is iterated twice with 97% percentile bound as a slope guard to separate the failure region from initial distribution. Hence, threshold bounds $t_1$ and $t_2$ are calculated as P(Y$\geq t_1$)=0.03 and P(Y$\geq t_2$)=0.0009, respectively. Table II shows the accuracy and performance of failure analysis performed by different approaches. Comparing the traditional MC method, IFRD shows 63x speedup gain with 0.005% error. With using proposed sampling reduction method, IFRD is also 2x faster than recursive statistical blockade approach with only additional 0.0007% error. For the estimated specification guidance for parameters, we find that only 1.2% of samples, which meet the in-spec guidance, are determined as the failure by the classifier in Table III. Fig. 6 shows that our proposed method is more accurate than previous methods since the tail of CDF depicting the 5-sigma failure region is more correlated to the golden reference (MC).

TABLE II
COMPARISON OF THE ACCURACY AND EFFICIENCY ON 6T-SRAM CIRCUIT

| | **Failure probability** | **# Sim. runs** | **Speed -up(x)** | **Error** |
|---|---|---|---|---|
| Monte Carlo(MC) | 2.300E-04 | 1M | - | - |
| Rare Event Microscope (REscope) | 3.786E-04 | 5009 | 199.6 | 1.486E-04 |
| Recursive Statistical Blockade (RSB) | 2.775E-04 | 29260 | 34.2 | 4.754E-05 |
| Proposed method (IFRD) | 2.852E-04 | 15730 | 63.6 | 5.515E-05 |

TABLE III
ESTIMATED IN-SPEC GUIDANCE OF PARAMETERS ON 6T-SRAM CIRCUIT

| **Rank** | **Parameter @MOSFET** | **Initial condition ($\mu, \sigma$)** | **In-spec Guidance [MIN,MAX]** |
|---|---|---|---|
| 1 | $vfb$@MN1 | (-5.5E-01,0.1) | [-7.31E-01,-3.78E-01] |
| 2 | $vfb$@MP2 | (5.5E-01,0.1) | [3.54E-01,7.15E-01] |
| 3 | $ndep$@MN1 | (2.8E+18,0.1) | [1.84E+18,3.76E+18] |
| 4 | $ndep$@MN2 | (2.8E+18,0.1) | [1.70E+18,3.84E+18] |
| 4 | $ndep$@MP2 | (2.8E+18,0.1) | [1.89E+18,3.78e+18] |
| **Failure probability** | 0.0009($= t_2$) | **Estimation Error (%)** | **1.21** |

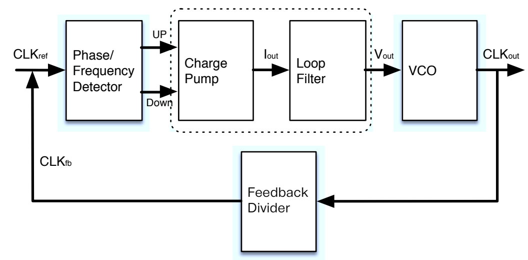



Fig. 7. A functional diagram of PLL

### B. Charge pump failure rate diagnosis

In a large logic circuit, a clock is frequently distributed to several sub-clocks, so frequencies of sub-clocks are easy to be inaccurate due to propagation delays. A PLL is frequently used to adjust the phase of clock. The functional block diagram of PLL is shown in fig. 7. After comparing the output clock ($CLK_{out}$) with the reference clock ($CLK_{ref}$) by phase detector, a charge pump circuit actually adjusts the frequency of clock signal by charging and discharging capacitors rely on input signals ($UP$ and $DN$). The mismatch of MOSFETS in a charge pump can cause the unbalanced timing and phase jitter between two different operation modes. Hence, we measure the timing ratio of charging and discharging operations, which can be formulated mathematically as

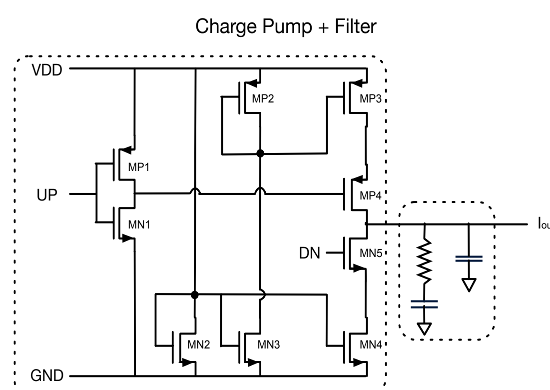



Fig. 8. Schematic representations of charge pump and filter

$r_{min} \leq \frac{t_{discharge}}{t_{charge}} \leq r_{max}$ ($r_{min,max}$ represents the minimum and maximum ratio to determine failures). A charge pump circuit consists of 9 MOSFETS as shown in fig. 8. The total number of process parameter is $81(9*9)$, so the dimensionality of parameters is much higher than 6T-SRAM case. We perform 3,000 initial sampling and simulation to model the initial performance distribution accurately. Similar to the 6T-SRAM testcase, we perform our algorithm twice with 97% percentile bound (P(Y$\geq$ $t_1$)=0.03, P(Y$\geq$ $t_2$)=0.0009). The result is summarized in Table IV. Our approach only required 6,263 sampling for estimating the failure probability of 5-sigma region with 0.0025% error, while MC needs 1 million samples. Even RSB and IFRD magnify the failure region with same confidence level, our approach can achieve 1.7 times speed gain with nearly same accuracy. Table V shows that the proposed method makes the decision for in-spec conditions of process parameters with 99.98% confidence level by managing only the first 5 ranked parameters of 81. The tail distribution of IFRD in 5-sigma failure region is much close to MC than REscope as we can see in Fig. 9.

TABLE IV
COMPARISON OF THE ACCURACY AND EFFICIENCY ON CHARGE PUMP CIRCUIT

| | Failure probability | # Sim. runs | Speed -up(x) | Error |
|---|---|---|---|---|
| Monte Carlo(MC) | 2.300E-04 | 1M | - | - |
| Rare Event Microscope (REscope) | 3.337E-04 | 4875 | 205.1 | 1.037-04 |
| Recursive Statistical Blockade (RSB) | 2.245-04 | 10432 | 95.9 | 5.471E-06 |
| Proposed method (IFRD) | 2.052-04 | 6263 | 159.7 | 2.477E-05 |

TABLE V
ESTIMATED IN-SPEC GUIDANCE OF PARAMETERS ON CHARGE PUMP CIRCUIT

| Rank | Parameter @MOSFET | Initial condition $(\mu, \sigma)$ | In-spec Guidance [MIN,MAX] |
|---|---|---|---|
| 1 | $ndep$@MN1 | (2.8E+18,0.1) | [1.74E+18,3.78E+18] |
| 2 | $ndep$@MP2 | (2.8E+18,0.1) | [1.68E+18,3.73E+18] |
| 3 | $ndep$@MP4 | (2.8E+18,0.1) | [1.78E+18,3.81E+18] |
| 4 | $ndep$@MN5 | (2.8E+18,0.1) | [1.78E+18,3.77E+18] |
| 4 | $ndep$@MP3 | (2.8E+18,0.1) | [1.88E+18,3.80E+18] |
| **Failure probability** | 0.0009($= t_2$) | **Estimation Error (%)** | **0.02** |

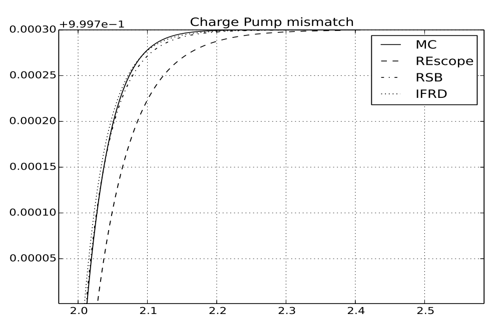



Fig. 9. Estimating the CDF of charge pump mismatch around 5-sigma region

## V. CONCLUSION

In this paper, we presented the novel statistical diagnosis method for rare failure events. The proposed method introduced two new techniques to speed up the failure analysis while providing the in-spec guidance of process parameters. First, the proposed method applies the smart sample selection method to reduce the additional simulation cost during iterative failure region locating process. Second, the new approach can provide safe design space or in-spec boundaries of parameters, which can help design to improve the yield and meet the target performance of design. Experimental results show that this method is 159x faster than the traditional MC method with only 0.002% error. Suggested in-spec guidance method also provides the safe conditions of process parameters with 0.02% estimation error. Furthermore, the new method reduced the simulation cost by 2x than recently published approaches without sacrificing accuracy.

## REFERENCES

[1] B. Calhoun, Y. Cao, X. Li, K. Mai, L. Pileggi, R. Rutenbar, and K. L. Shepard, "Digital circuit design challenges and opportunities in the era of nanoscale cmos," *Proceedings of the IEEE*, vol. 96, pp. 343–365, Feb 2008.

[2] H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: variability characterization and modeling for 65- to 90-nm processes," in *Custom Integrated Circuits Conference, 2005. Proceedings of the IEEE 2005*, pp. 593–599, Sept 2005.

[3] A. Singhee and R. Rutenbar, "Statistical blockade: Very fast statistical simulation and modeling of rare circuit events and its application to memory design," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 28, pp. 1176–1189, Aug 2009.

[4] D. Hocevar, M. Lightner, and T. N. Trick, "A study of variance reduction techniques for estimating circuit yields," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 2, pp. 180–192, July 1983.

[5] A. Singhee, J. Wang, B. Calhoun, and R. Rutenbar, "Recursive statistical blockade: An enhanced technique for rare event simulation with application to sram circuit design," in *VLSI Design, 2008. VLSID 2008. 21st International Conference on*, pp. 131–136, Jan 2008.

[6] W. Wu, F. Gong, G. Chen, and L. He, "A fast and provably bounded failure analysis of memory circuits in high dimensions," in *Design Automation Conference (ASP-DAC), 2014 19th Asia and South Pacific*, pp. 424–429, Jan 2014.

[7] W. Wu, W. Xu, R. Krishnan, Y.-L. Chen, and L. He, "Rescope: High-dimensional statistical circuit simulation towards full failure region coverage," in *Design Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE*, pp. 1–6, June 2014.

[8] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Statistical design and optimization of sram cell for yield enhancement," in *Computer Aided Design, 2004. ICCAD-2004. IEEE/ACM International Conference on*, pp. 10–13, Nov 2004.

[9] K. Agarwal and S. Nassif, "Statistical analysis of sram cell stability," in *Design Automation Conference, 2006 43rd ACM/IEEE*, pp. 57–62, 2006.

[10] 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 *Design, Automation Test in Europe Conference Exhibition (DATE), 2010*, pp. 801–806, March 2010.

[11] S. N. R. Kanj, R. Joshi, "Mixture important sampling and its application to the analysis of sram designs in the presence of rare failure events," in *Proc. IEEE/ACM Design Automation Conference (DAC)*, pp. 69–72, 2006.

[12] T.Hastie, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[13] J. J.R.Hosking, "Parameter and quantile estimation for the generalized pareto distribution," *Technometrics*, vol. 29, no. 3, pp. 339–349, 1987.

[14] J.R.Hosking, "Algorithm as 215: Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 34, no. 3, pp. 301–310, 1985.

[15] S. Joe and F. Y. Kuo, "Remark on algorithm 659: Implementing sobol's quasirandom sequence generator," *ACM Trans. Math. Softw.*, vol. 29, pp. 49–57, Mar. 2003.

[16] P. Nenzi and V. Holger, "Ngspice users manual," *Version 25plus*, 2010.