

REscope: High-dimensional Statistical Circuit Simulation towards Full Failure Region Coverage

Wei Wu¹, Wenyao Xu², Rahul Krishnan¹, Yen-Lung Chen³, and Lei He¹

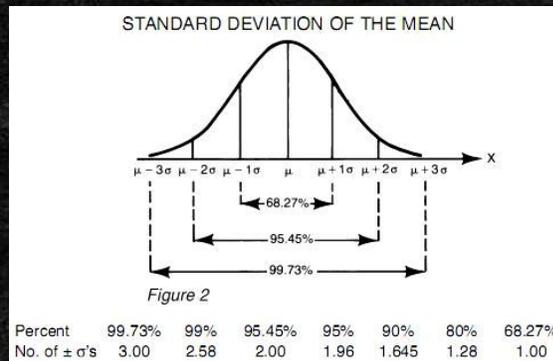
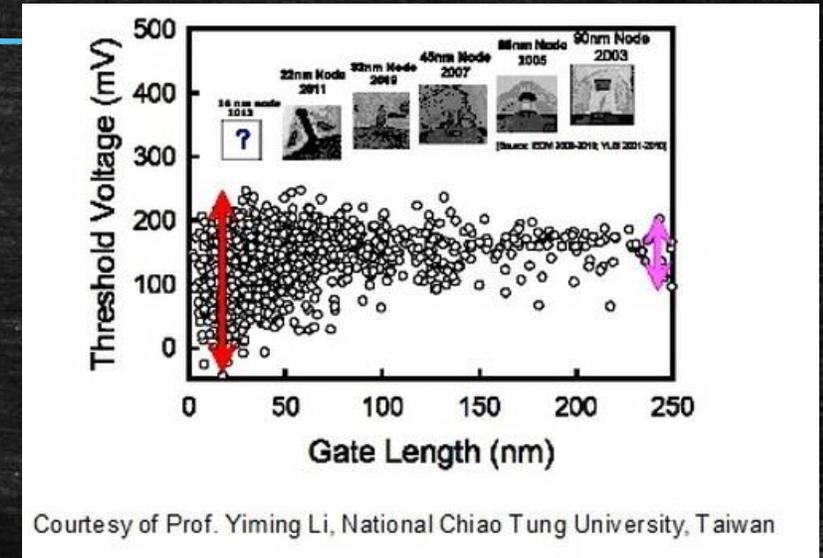
¹ Electrical Engineering Department, UCLA, <http://eda.ee.ucla.edu>

² CSE Dept., University of Buffalo, SUNY, NY, USA

³EE Dept., National Central University, Taiwan, R.O.C

Need & Challenge of High Sigma analysis

- **Need for statistical circuit analysis**
 - Process technology continue to scale 45nm, 32nm, 22nm ...
 - Shrinking devices → more prone to PVT variations
 - Statistical circuit analysis helps to **debug circuits in the pre-silicon phase**, and enhances yield rate
- **Challenges: High Sigma (Rare Failure Event) analysis**
 - Memory cell (e.g. 6+ sigma)
 - Critical circuits (e.g. 4-5 sigma): I/O cell, PLL, etc.



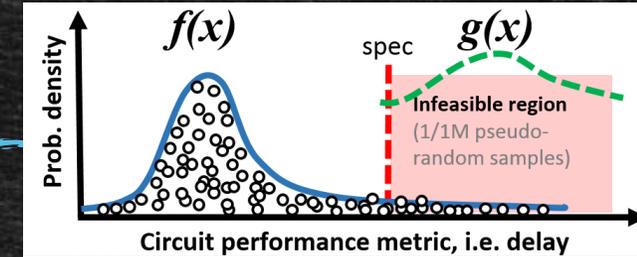
Sigma	Probability	# of Simulations
1	0.15866	700
2	0.02275	4,400
3	0.00135	74,100
4	3.17E-05	3,157,500
5	2.87E-07	348,855,600

Existing Approaches

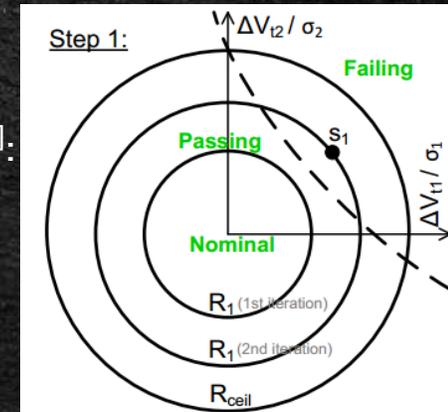
- Monte Carlo Simulation:
 - Monte Carlo is **infeasible** for high sigma analysis in terms of computational complexity
 - One post-layout PLL simulation may take several hours

- Importance Sampling methods:
 - Shift the sample distribution to more “important” region
 - e.g. Spherical sampling^[QTD10], HDIS^[WGC14]

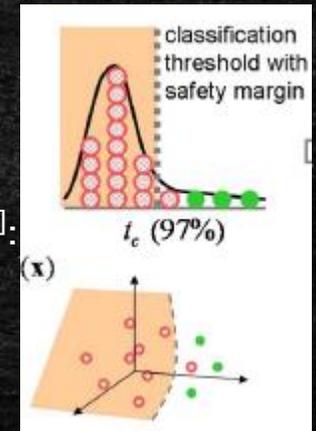
- Classification based methods:
 - Statistical Blockade (Linear classifier)^[SR08]



Spherical sampling^[QTD10]:



Statistical Blockade^[QTD10]:



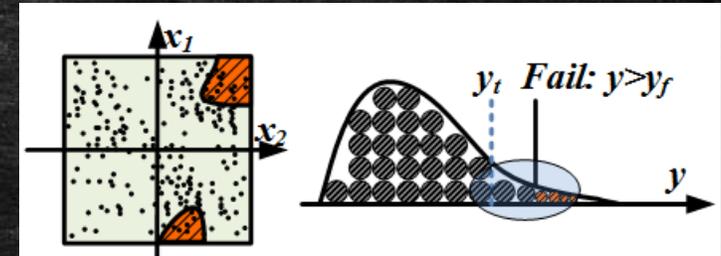
[QTD10] M. Qazi, M. Tikekar, L. Dolecek, D. Shah, and A. Chandrakasan, “Loop flattening and spherical sampling: Highly efficient model reduction techniques for SRAM yield analysis,” in DATE’2010, pp. 801–806.

[12] W. Wu, F. Gong, G. Chen, and Lei He, “A Fast and Provably Bounded Failure Analysis of Memory Circuits in High Dimensions”, ASPDAC’2014

[SR08] A Singhee, and RA Rutenbar. “Statistical blockade: Very fast statistical simulation and modeling of rare circuit events and its application to memory design,” in DATE’2008, pp. 235-251.

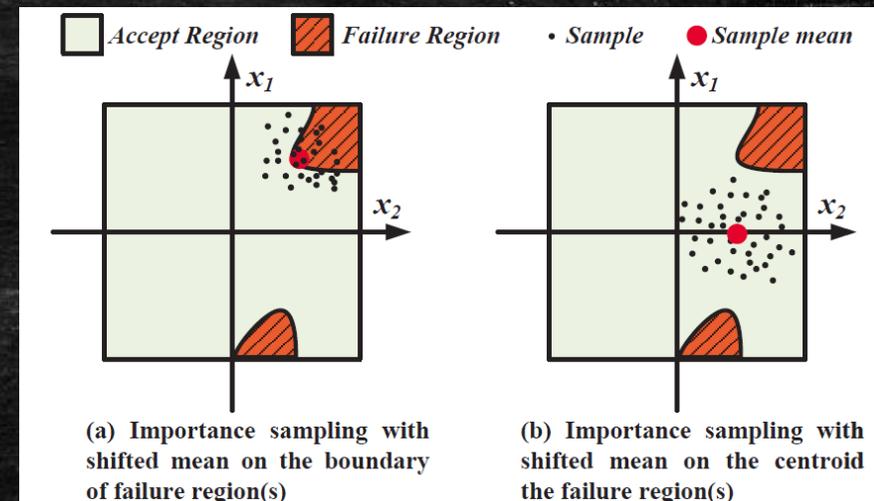
What if there are multiple failure regions ?

- What if failure samples fall in multiple disjoint regions

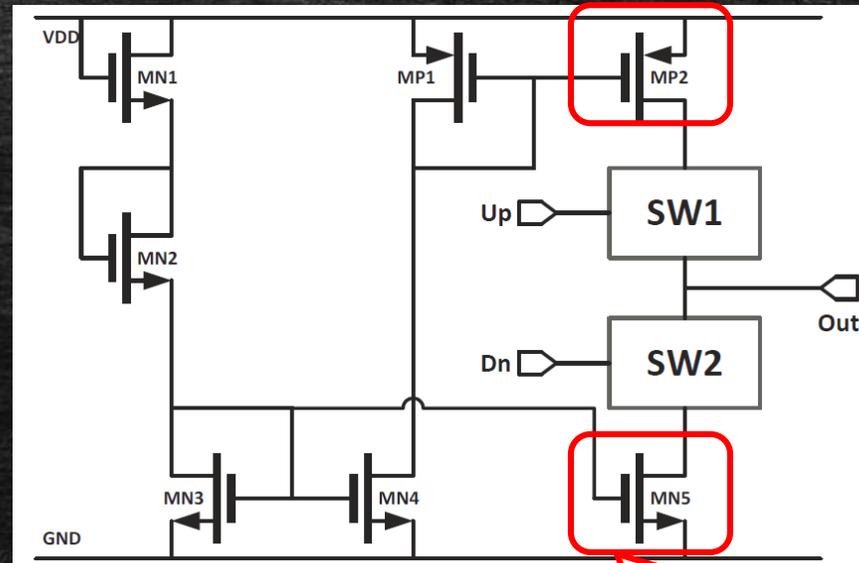
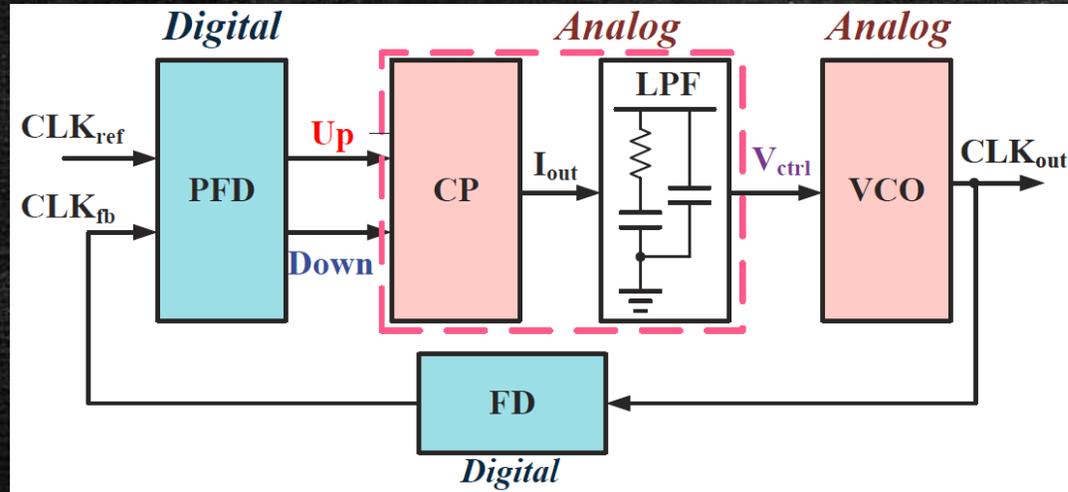


- Importance sampling works like this:

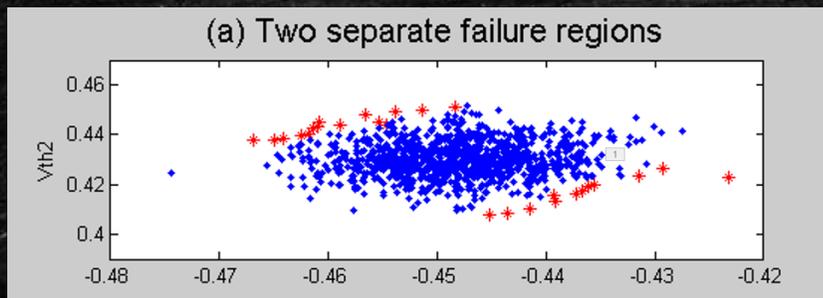
- Statistical blockade can't identify multiple failure regions. Even worse, **the classifier scales poorly** with the number of variation variables



Unfortunately, multiple failure regions do exist



PFD: phase frequency detector; CP: Charge pump
FD: frequency divider; VCO: voltage controlled oscillator



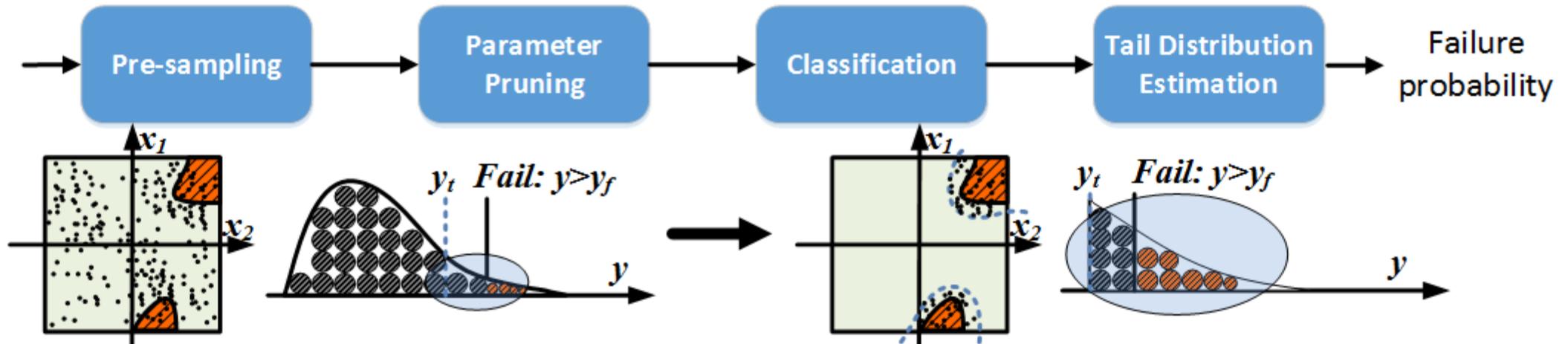
V_{th} s of MP2 and MN5 are variation parameters

Mismatch between MP2 and MN5 may result in fluctuation of control voltage, which will lead to "jitter" in the clock.

Rare-Event Microscope (REscope)

- Make the classification based approach more practical, in terms of **accuracy** and **efficiency**.
 - Identifies **multiple failure regions**
 - Handles **high dimensional** problems
 - Approximates the tail as a **generalized pareto distribution (GPD)**
 - GPD: a good model for the distribution of the exceedence to a certain threshold in another distribution, i.e. the tail of $\text{PDF}(y)$

Distribution of process variation parameters



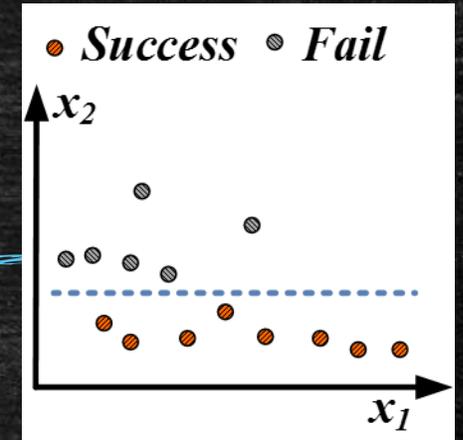
Pruning and Classification

- **Parameter pruning (Feature selection)**

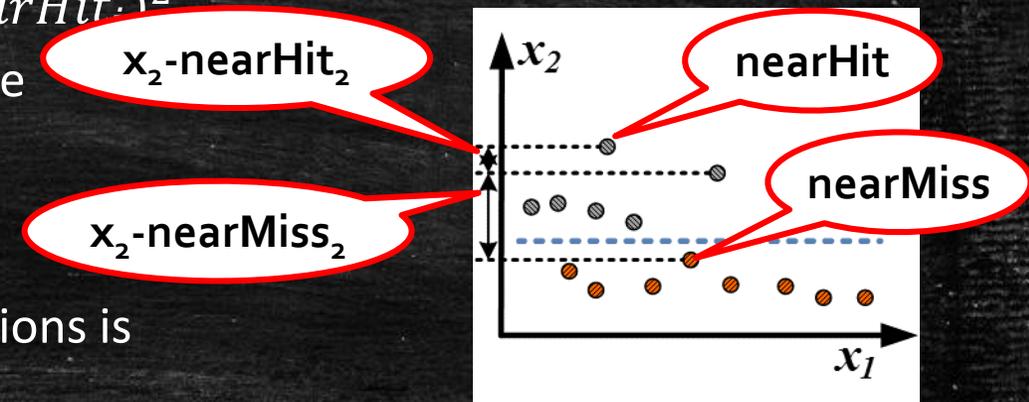
- Evaluate the “Importance” of each parameter (feature)
- RELIEF-F algorithm:
 - $$W_i = W_i + (x_i - nearMiss_i)^2 - (x_i - nearHit_i)^2$$
- Sort parameters by “Importance”, and prune the unimportant parameters

- **Classification**

- The boundary separating accept and failure regions is usually nonlinear
- Apply support vector machine (SVM) with radial basis function (RBF) kernel yields a good nonlinear boundary



Calculate W_2



Distribution of process variation parameters

Pre-sampling

Parameter Pruning

Classification

Tail Distribution Estimation

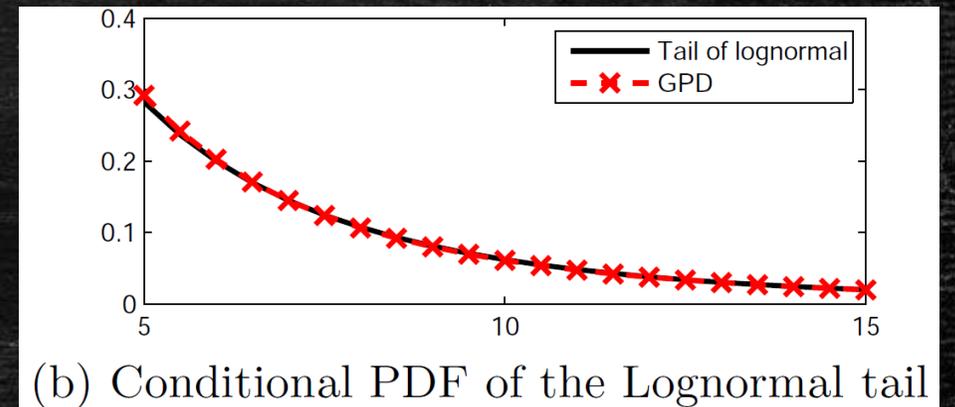
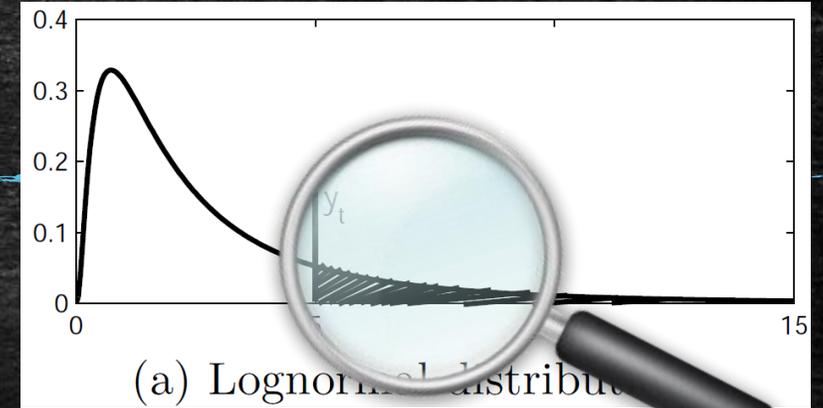
Failure probability

GPD Approximation

- **Approximate the tail of PDF by GPD with only 2 parameters**
 - Shape parameter(ξ), scale parameter(σ), starting point of the tail(μ)

$$F_{(\xi, \mu, \sigma)}(y) = \begin{cases} 1 - (1 - \frac{\xi(y-\mu)}{\sigma})^{\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ 1 - \exp(-\frac{(y-\mu)}{\sigma}) & \text{for } \xi = 0 \end{cases} \quad (5)$$

- Calculate the initial solution by probability weighted moment matching
- Refine the solution by Newton's method



Distribution of process variation parameters

Pre-sampling

Parameter Pruning

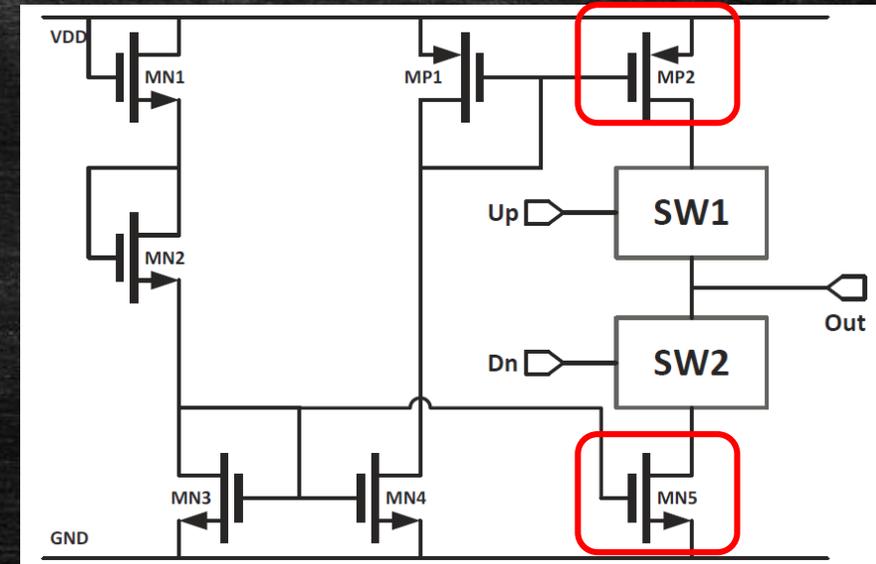
Classification

Tail Distribution Estimation

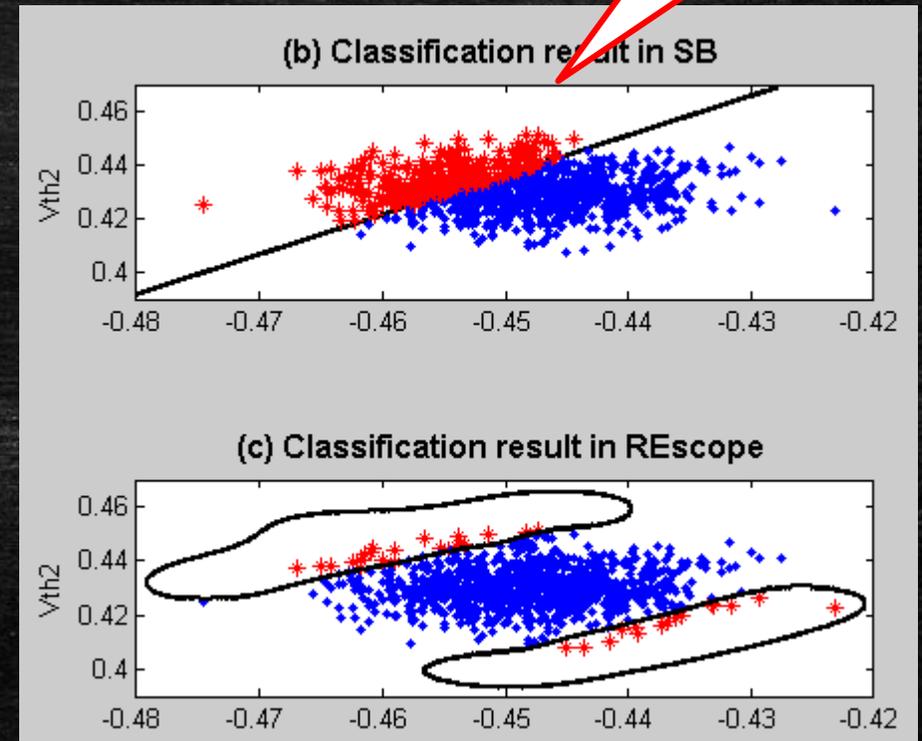
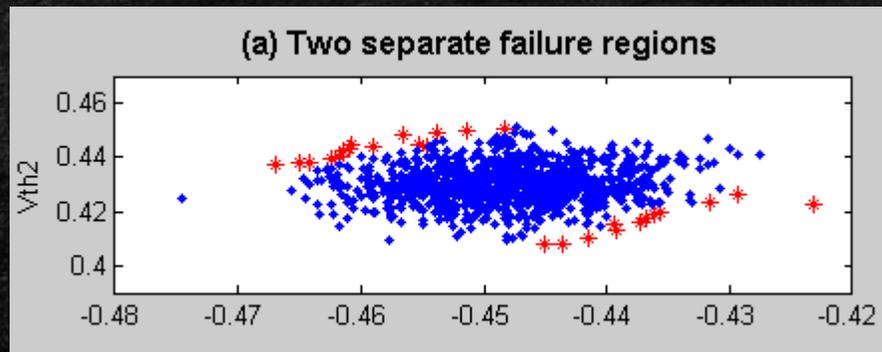
Failure probability

Experiments: REscope for Charge Pump

- One circuit, two configurations
 - An **illustrative** example: only **2** variation parameters
 - V_{th} of MP2 and MN5
 - A high dimensional example: **108** process variation parameters (27 transistors, 4 process variation parameters for each)
 - Channel length
 - Channel weight
 - Gate oxide thickness
 - Flat-band voltage
- Failure: current mismatch.

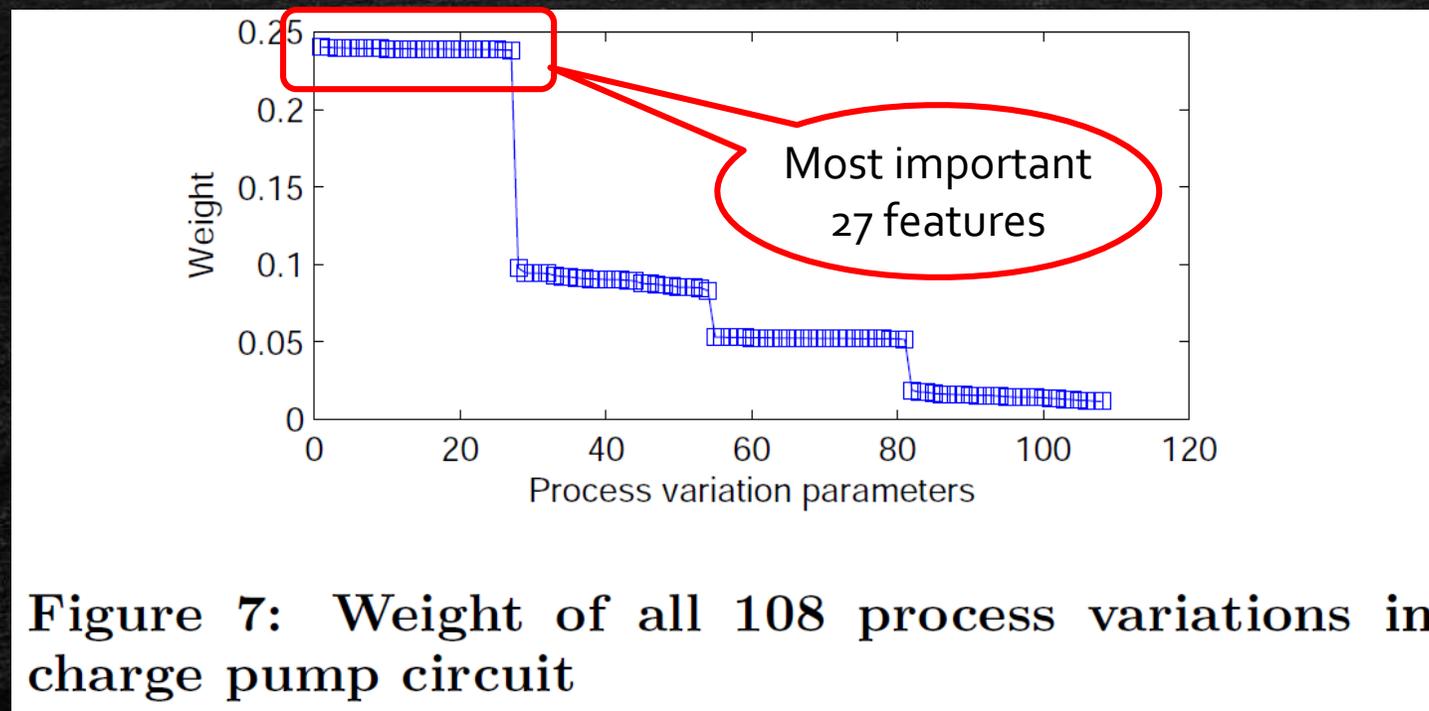


Experiment 1: Illustration of Classification results



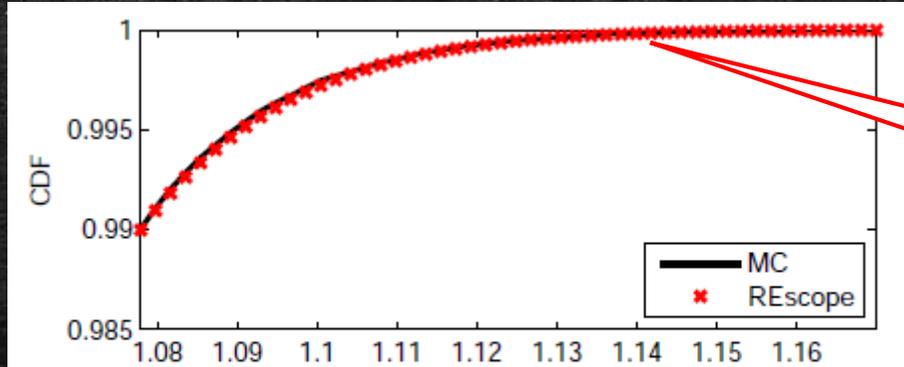
Experiment 2: Handling high dimensional cases

- Parameter pruning results:
 - 27 important features are selected out of 108 variables



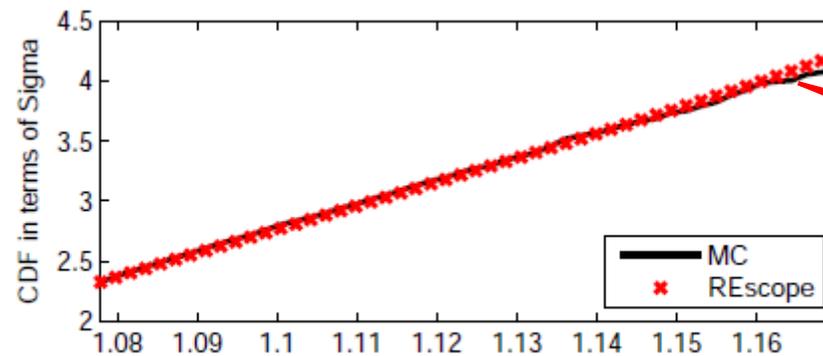
Accuracy of REscope

- Tail modeling after parameter pruning:



(a) CDF tail in linear scale

The estimated P_{fail}
overlap with MC



(b) CDF tail in log scale

Mismatch after 4.2 sigma
because we don't have
enough MC samples so far.

Comparison with existing methods

Table 1: Comparison of the accuracy and efficiency on charge pump circuit

	Monte Carlo (MC)	Importance sampling (HDIS)[12]	Proposed approach (REscope)
failure probability	2.279e-5 (0%)	1.136e-3	2.256e-5 (+1.05%)
#sim. runs	1.4e+6 (389x)	2e+4 (5.6x)	3.6e+3 (1x)

- **389x** speedup on MC with **1.05% error at 4 sigma**, while importance sampling and statistical blockade fail.

Conclusions

- Strength:
 - Dimension reduction algorithm keeps the “important” parameters only.
 - Nonlinear classifier works perfectly on detecting multiple failure regions.
 - The tail is explicitly matched to a known distribution, such as GPD, to further reduce the number of sample.
 - Experiments show 389x speedup on MC with only 1.05% error, which importance sampling and statistical blockade fail.

- Limitation and Future direction:
 - Considering the correlation between variations sources
 - Algorithm to automatically adjust the classifier configuration



Thanks!

Address Comments to weiw@seas.ucla.edu