Probabilistic Evaluation of Solutions in Variability-Driven Optimization

Azadeh Davoodi and Ankur Srivastava
University of Maryland,
College Park.

Presenter: Vishal Khandelwal

ISPD2006

Outline

- Motivation
  - Challenge in probabilistic optimization considering process variations
- Pruning Probability
  - Metric for comparison of potential solutions
- Computing the Pruning Probability
- Application
  - Dual-Vth assignment considering process variations
- Results

Motivation

- Many VLSI CAD optimization problems rely on comparison of potential solutions
  - *To identify the solution with best quality, or to identify a subset of potentially good solutions*

- Any potential solution S_i has a corresponding timing r_i & cost c_i:
  - e.g., A solution to the gate-sizing problem has:
    - Timing: Delay of the circuit
    - Cost: Overall sizes of the gates

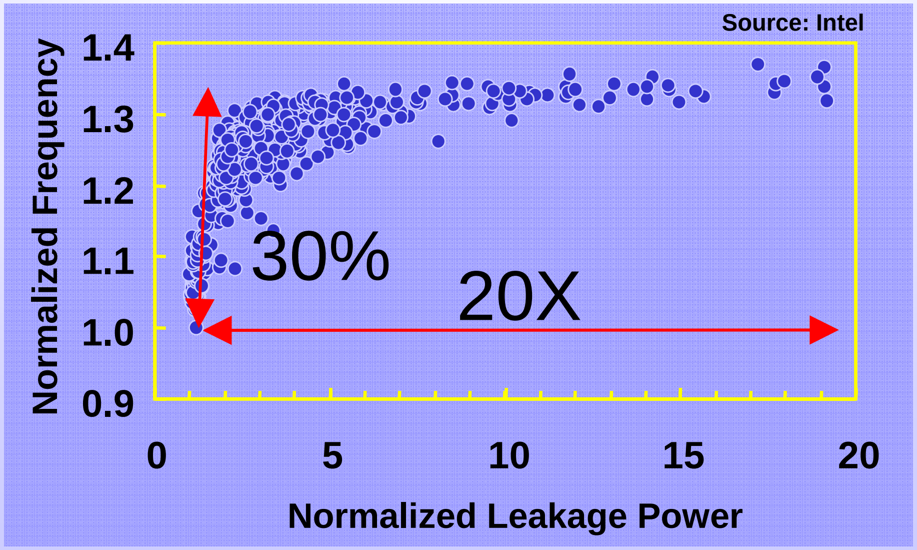
Motivation

- A good solution is the one with better timing and cost

$$S_i \text{ superior } S_j \Leftrightarrow r_i \le r_j \;\&\; c_i \le c_j$$

- Process variations randomize the timing and cost associated with a potential solution

$$S_i \text{ superior } S_j \Leftrightarrow P(R_i \le R_j \;\&\; C_i \le C_j) \approx 1$$



Pruning Probability

$$S_i \text{ superior } S_j \Leftrightarrow P(R_i \le R_j \,\&\, C_i \le C_j) \approx 1$$

- Let $C = C_j - C_i$ and $R = R_j - R_i$

$$P(R \ge 0 \,\&\, C \ge 0) = \int_0^{\infty}\int_0^{\infty} f_{R,C}(r,c)\,dr\,dc$$

$f_{R,C}$: joint probability density function (*jpdf*) of random variables *R* and *C*

Computing the Pruning Probability: Challenges

$$P(R \geq 0 \,\&\, C \geq 0) = \int_0^\infty \int_0^\infty f_{R,C}(r,c)\,drdc$$

- Accuracy
  - Might not have an analytical expression for $f_{R,C}$
  - Might require numerical methods to compute the probability

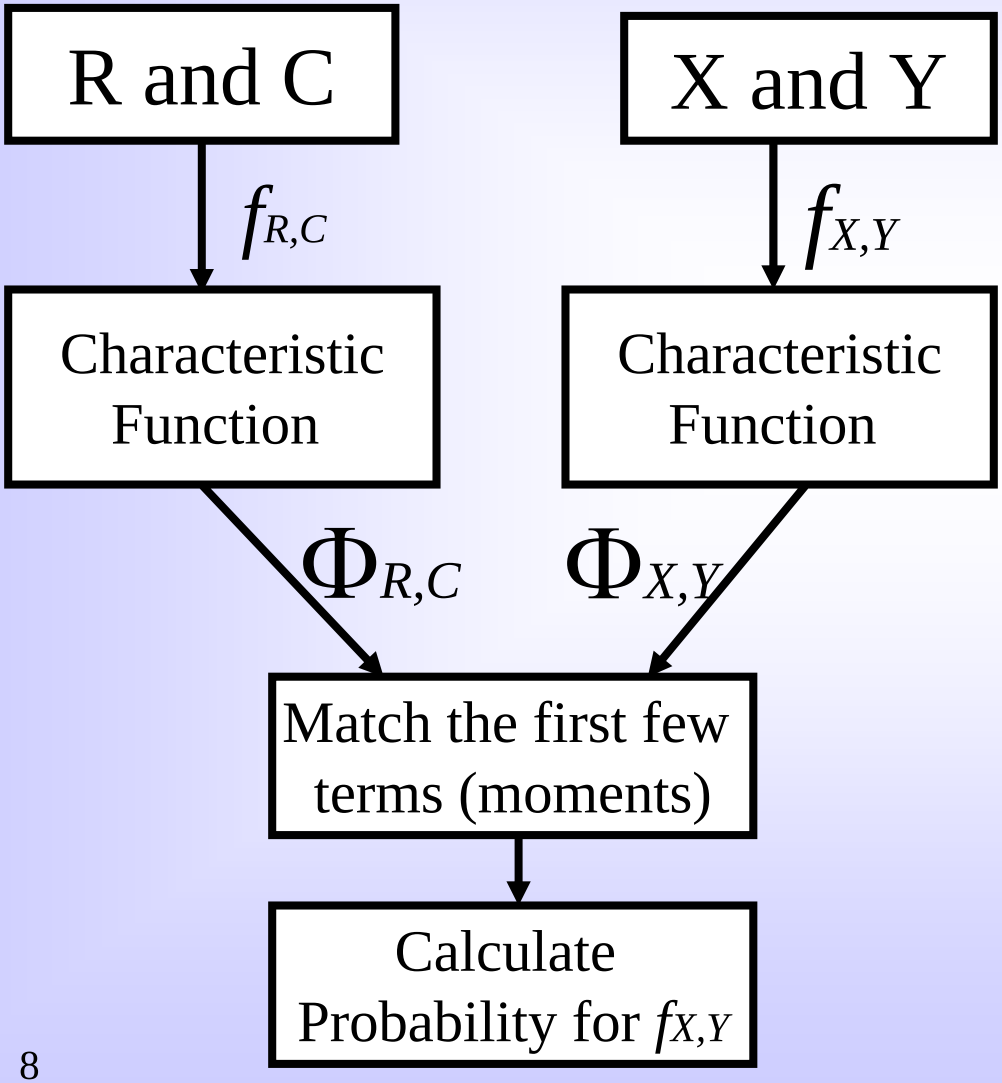
- Fast computation
  - Necessary in an optimization framework
  - Makes the use of numerical techniques such as Monte Carlo simulation impractical

Computing the Pruning Probability: Methods

- Based on analytical approximation of the *jpdf* ($f_{R,C}$)
  - With a well studied *jpdf*
  - For which computing the probability integral is analytically possible

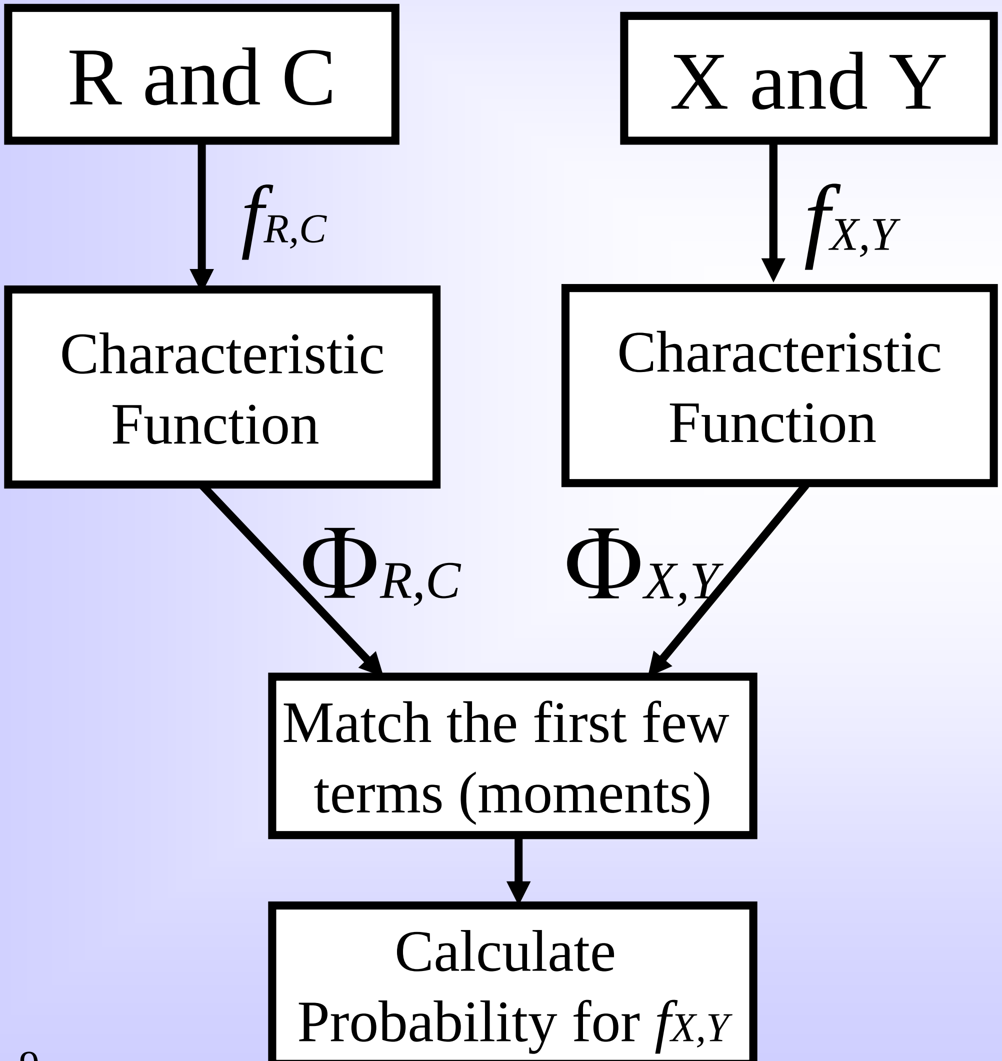
- Using Conditional Monte Carlo simulation
  - Bound-based numerical evaluation of the probability
  - Potentially much faster than Monte Carlo

Computing the Pruning Probability: Approximating *jpdf* by Moment Matching



- Approximate R,C with new random variables X,Y where the type of *jpdf* of X,Y is known
- Compute the first few terms of the characteristic functions (Fourier transform) of the two *jpdf*s (i.e., moments)
- Match the first few moments and determine the parameters of $f_{X,Y}$
- Compute the pruning probability for X and Y

Computing the Pruning Probability:
Approximating *jpdf* by Moment Matching



$$\Phi_{X,Y}(t_1,t_2) = \iint e^{i(t_1x+t_2y)} f_{X,Y}(x,y)dxdy$$

$$= 1 + it_1m_{10} + it_2m_{01} - \frac{t_1^2}{2}m_{20} - \ldots$$

$$m_{ij} = \iint x^i y^j f_{X,Y}(x,y)dxdy = E[X^iY^j]$$

Computing the Pruning Probability: Approximating *jpdf* by Moment Matching

- Challenges:
  - Very few bivariate *jpdf*s have closed form expressions for their moments
  - Integration of very few known *jpdf*s over the quadrant are analytically possible

- Will study the example of bivariate Gaussian approximation given polynomial representation of R and C

Example: Bivariate Gaussian *jpdf* for Polynomials

Polynomial representation of R and C under process variations

- Can represent R and C as polynomials
  - By doing Taylor Series expansion of the R and C expressions in terms of random variables representing the varying parameters due to process variations (e.g., L_{eff}, T_{ox}, etc.)
  - Higher accuracy needs higher order of expansion
  - These r.v.s can be assumed to be independent
    - Using Principal Component Analysis (PCA)

$$R = f_1(L_{eff}, T_{ox}, \ldots)$$
$$C = f_2(L_{eff}, T_{ox}, \ldots)$$

→ PCA and Taylor Series Expansion

$$R = Poly_1(X_1, X_2, \ldots)$$
$$C = Poly_2(X_1, X_2, \ldots)$$

Example: Bivariate Gaussian *jpdf* for Polynomials

$$R = Poly_1(X_1, X_2, \ldots) \approx r_0 + \sum r_i X_i \qquad C = Poly_2(X_1, X_2, \ldots) \approx c_0 + \sum c_i X_i$$

- Assuming {X1,X2,…} are independent r.v.s with Gaussian density functions
  - The *jpdf* ($f_{R,C}$) is approximated to be bivariate Gaussian
  - Using linear approximation of R and C

$$f_{X,Y} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp[-\frac{z}{2(1-\rho^2)}]$$

$$z = \frac{(x_1-\mu_1)^2}{\sigma_1{}^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2{}^2}$$

- Moments of bivariate Gaussian *jpdf* are related to $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$
  - Need to specify the values of these parameters using moment matching

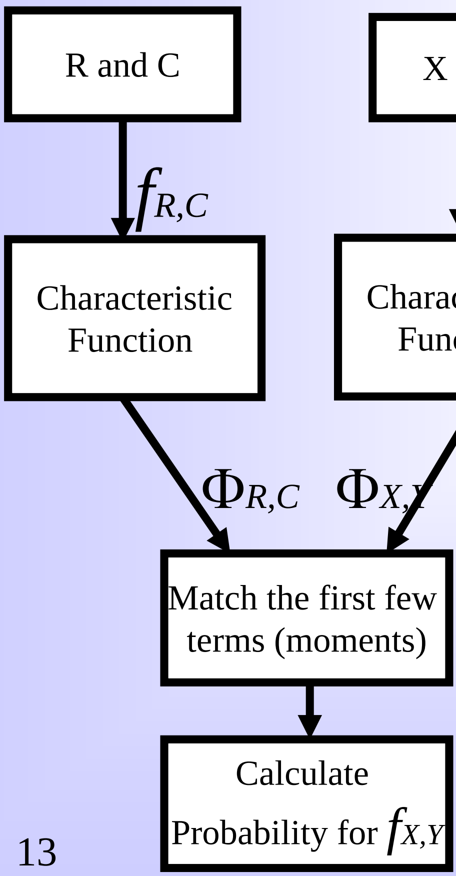
Example: Bivariate Gaussian *jpdf* for Polynomials

$$f_{X,Y} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp[-\frac{z}{2(1-\rho^2)}]$$

$$z = \frac{(x_1-\mu_1)^2}{\sigma_1^{\ 2}} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^{\ 2}}$$

$$R = Poly_1(X_1, X_2,...) \approx r_0 + \sum r_i X_i$$

$$C = Poly_2(X_1, X_2,...) \approx c_0 + \sum c_i X_i$$



$$\mu_2 = E[C] \qquad \sigma_1^{\ 2} + \mu_1^{\ 2} = E[R^2]$$

$$\mu_1 = E[R] \qquad \sigma_2^{\ 2} + \mu_2^{\ 2} = E[C^2]$$

$$\rho\sigma_x\sigma_y + \mu_x\mu_y = E[RC]$$

Analytical expression for probability integral of bivariate Gaussian *jpdf* is available (Hermite Polynomials)*

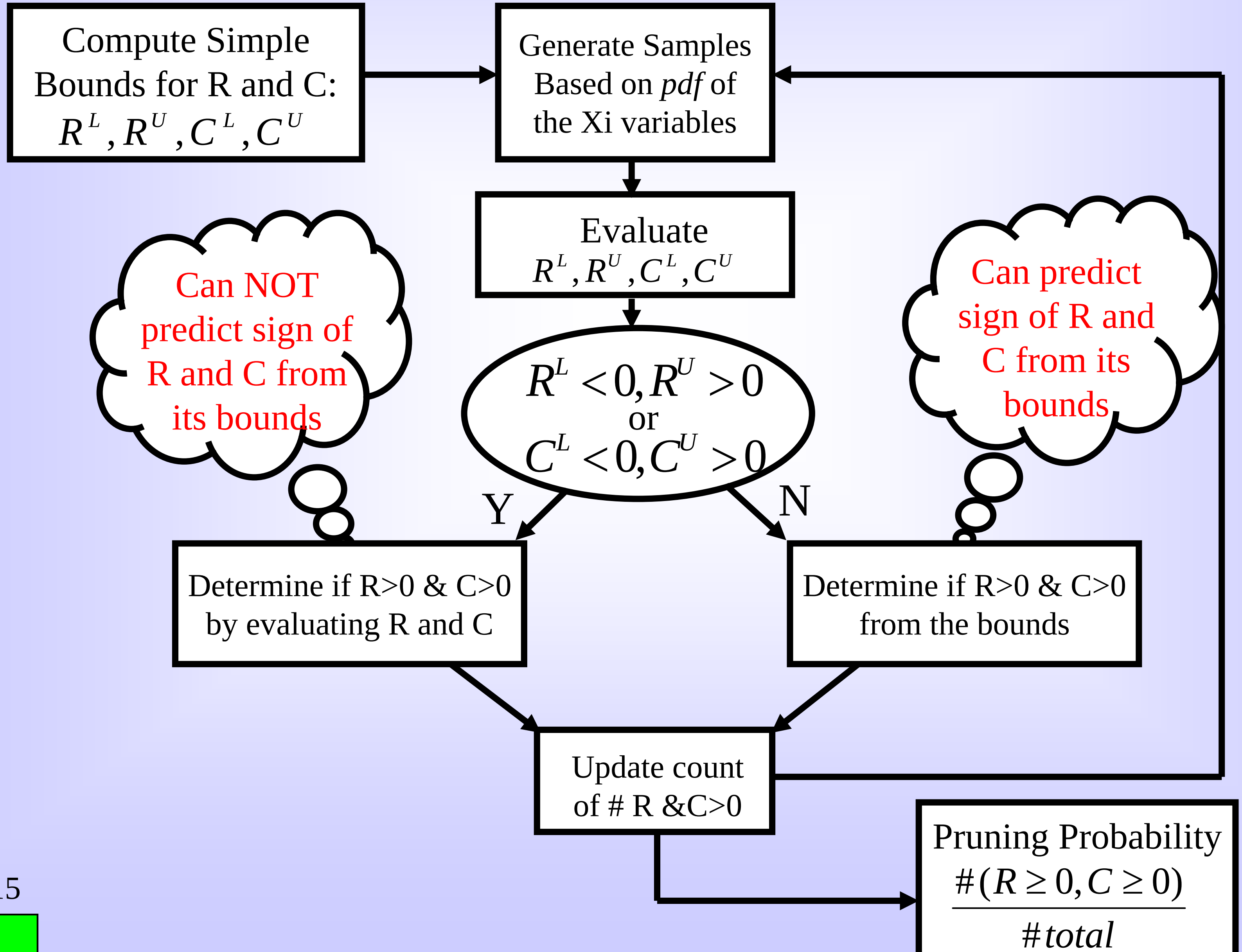
*[Vasicek 1998]

Computing the Pruning Probability: Conditional Monte Carlo (CMC)

$$P(R \geq 0 \,\&\, C \geq 0) = \int_0^\infty \int_0^\infty f_{R,C}(r,c)\,dr\,dc$$

- CMC is similar to MC but:
  - Uses simple bounds that can evaluate the sign of R and C for most of the MC samples
    - *Evaluation of simple bounds are much more efficient than polynomial expressions that are potentially of high order*
  - Only in the cases that the simple bounds can not predict the sign of R and C, the complicated polynomial expressions are evaluated

Computing the Pruning Probability: Conditional Monte Carlo (CMC)



Computing the Pruning Probability: Conditional Monte Carlo (CMC)

- Accurately predicts the probability value

- Speedup is due to the following intuition:
  - Evaluation of simple bounds are much faster than high-order polynomials
  - If the bounds are accurate, they predict the sign of the polynomials very frequently resulting in significant speedup

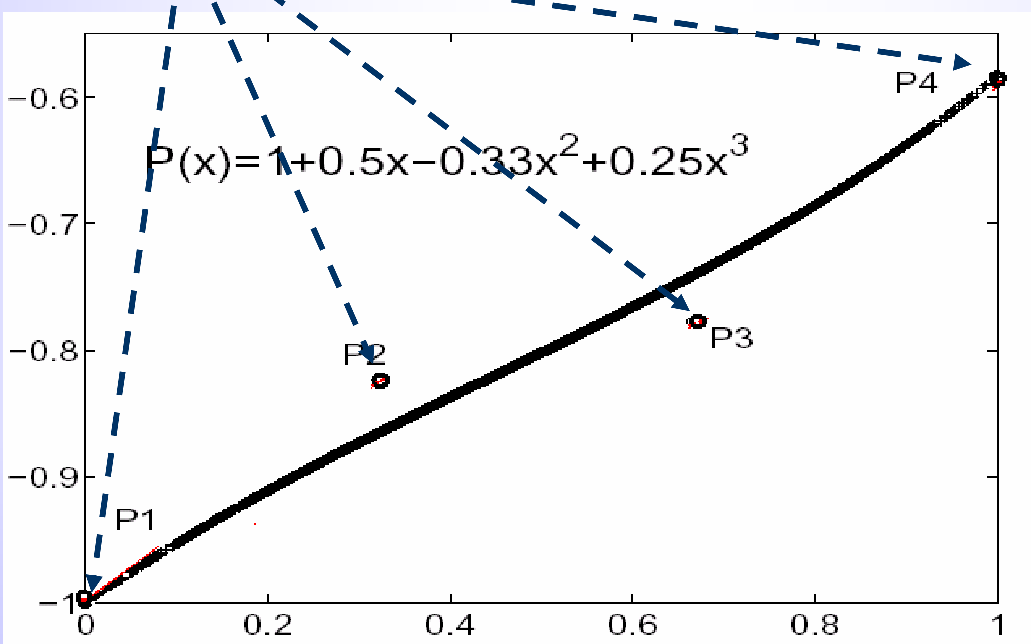
Example: Computing Bounds on Polynomials

$$Poly(x_1,\ldots,x_n) = \sum_{i_1=0}^{l_1}\cdots\sum_{i_n=0}^{l_n} a_{i_1,\ldots,i_n} x_1^{\,i_1} x_2^{\,i_2}\ldots x_n^{\,i_n}$$

- Bernstein coefficients define convex hull for any polynomial*

$$\left(\frac{i_1}{l_1};\ldots;\frac{i_n}{l_n};b_{i_1,\ldots,i_n}\right)$$

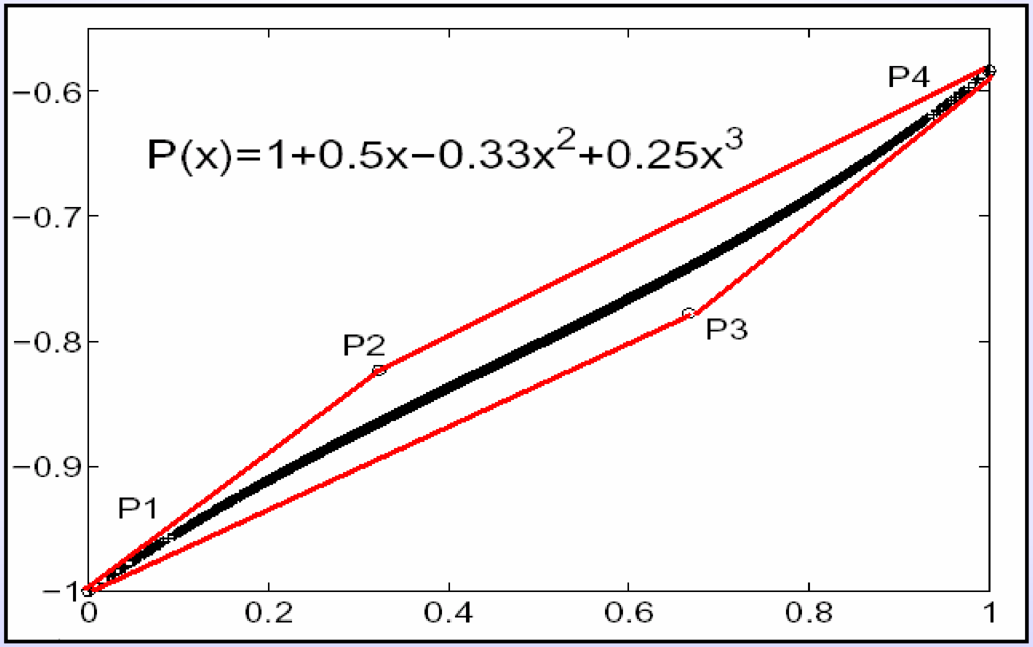
$$b_{i_1,\ldots,i_n} = \sum_{j_1=0}^{i_1}\cdots\sum_{j_n=0}^{j_n} \frac{\binom{i_1}{j_1}\cdots\binom{i_n}{j_n}}{\binom{l_1}{j_1}\cdots\binom{l_n}{j_n}} a_{i_1,\ldots,i_n}$$



*[Cargo, Shisha 1966]

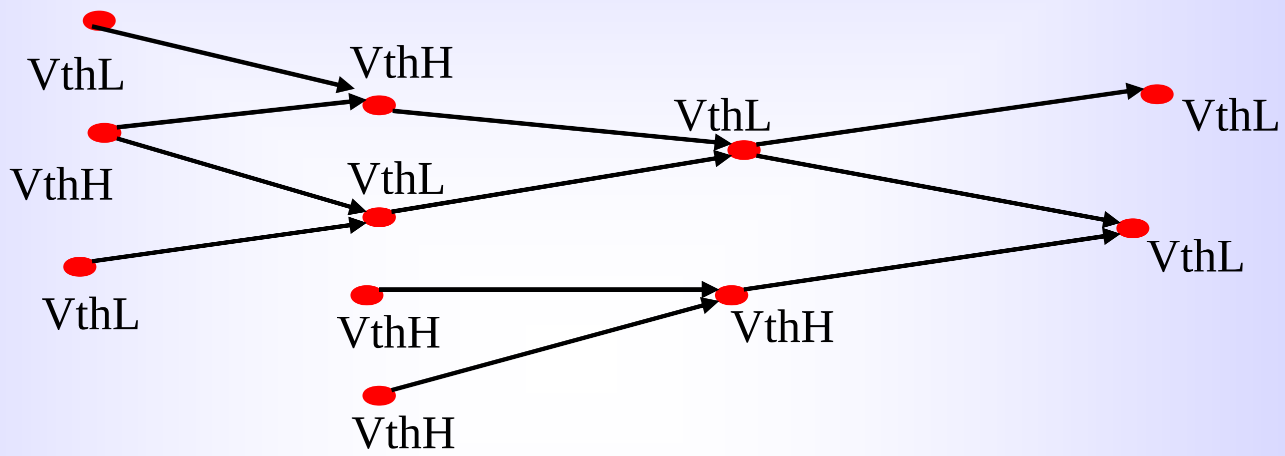
Example: Computing Bounds on Polynomials

- Simple hyper-plane lower-bounds are defined for each polynomial from its Bernstein coefficients*



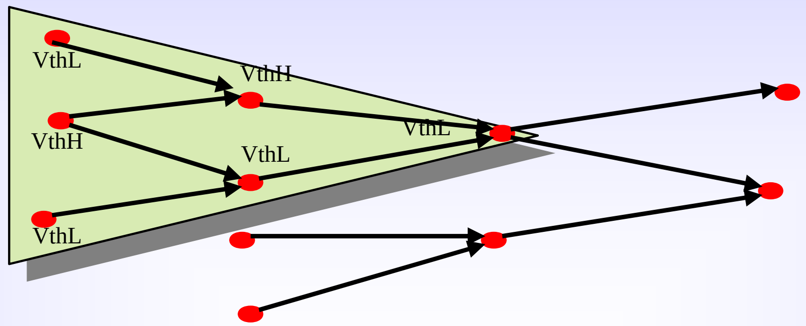
*[Garloff, Jansson, Smith 2003]

Application:
Dual-V_{th} Assignment for Leakage Optimization Under Process Variations



- Assignment of either high or low threshold voltage to gates in a circuit (represented as nodes in a graph)
  - Higher threshold (slow), lower threshold (leaky)
- Under process variations the goal is:
  - To minimize expected value of overall leakage (E[L])
  - Subject to bounding the maximum probability of violating a Timing Constraint (T_{cons}) at the Primary Output

Dual-Vth Assignment for Leakage Optimization Under Process Variations



- Dynamic programming based formulation

  - Topological traversal from PIs to POs

  - Solution at a node:

    -Vth assignment to sub-tree rooted at the node

  - S

    Using the Pruning Probability

    de's children + node's Vth possibilities

    -Pareto-optimal set identified & stored*

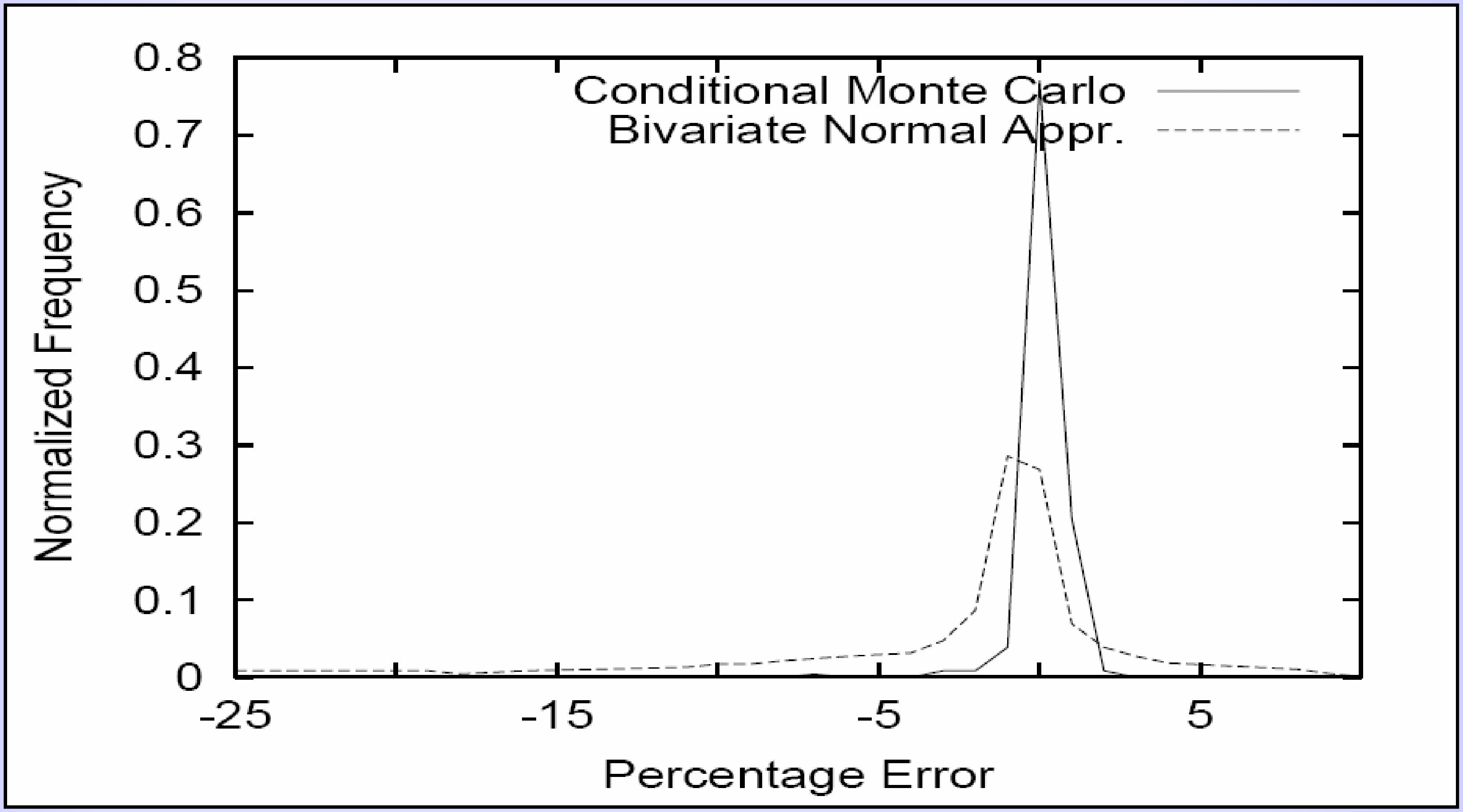
- Each solution:

  - Overall leakage at the node's subtree: $L_k = l_0^{(k)} + \sum l_i^{(k)} X_i + \sum \sum l_{ij}^{(k)} X_i X_j + \ldots$

  - Arrival time of the node's subtree: Approximated as a linear combination of parameters

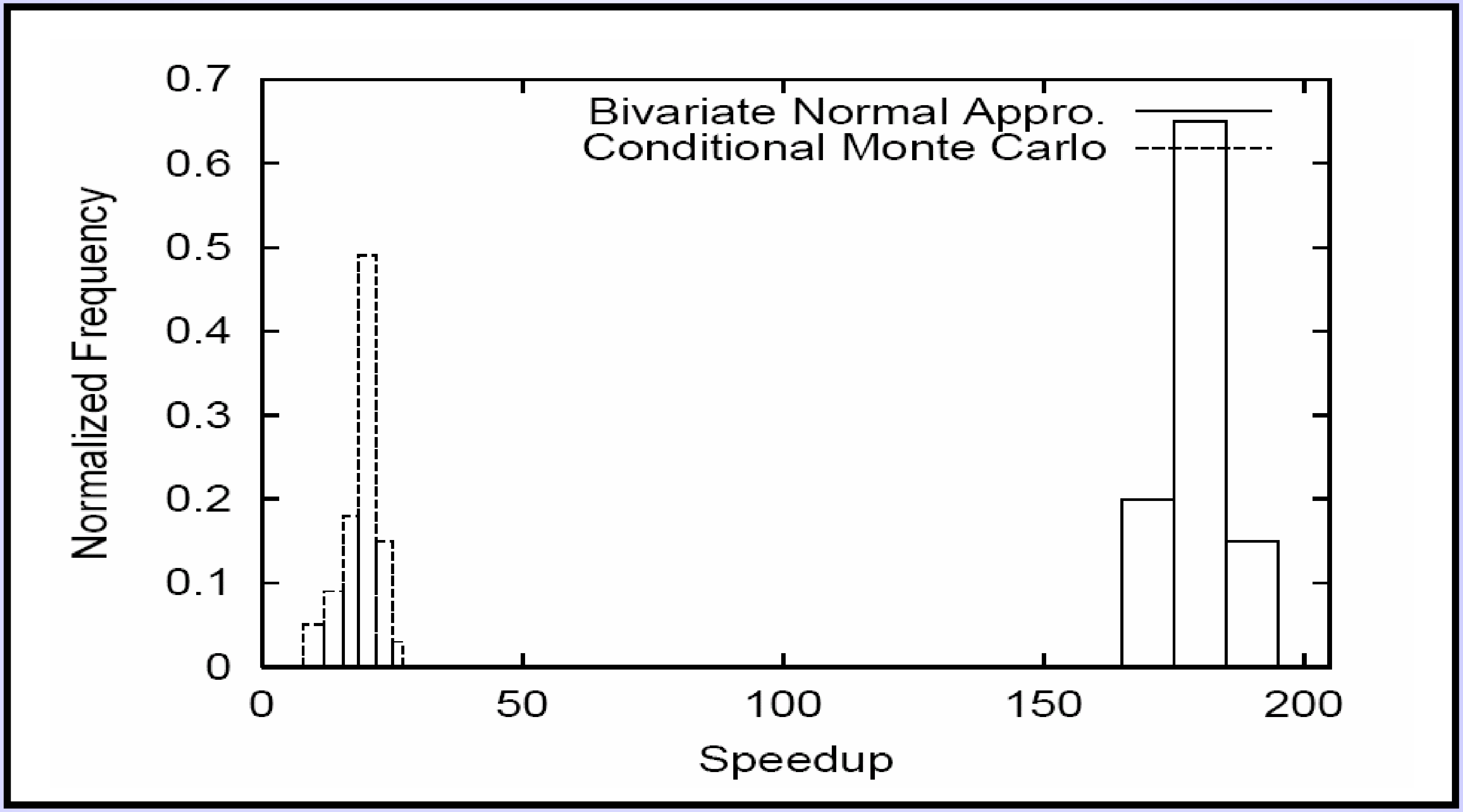
*[Davoodi, Srivastava ISLPED05]

%Error in Estimation of Pruning Probability



For 2600 solution pairs from the dual-V_{th} framework

Speedup in Computing the Pruning Probability



Comparing Quality of Solution in Dual-V_{th} Assignment

	T_{cons} (nsec)	Worst-Case Deterministic			*jpdf* Appr.			Conditional MC		
		$E[I]$	$P_v(T)$	t	%imp	$P_v(T)$	t	%imp	$P_v(T)$	t
C432	33.0	10634	0.11	10	35	0.27	13	46	0.25	552
C499	17.5	14285	0.14	29	15	0.17	51	22	0.14	1582
C880	32.0	16650	0.11	12	39	0.29	14	49	0.30	610
C1355	18.0	17182	0.08	40	29	0.11	50	36	0.09	1572
C1908	29.0	13768	0.13	37	33	0.18	40	39	0.16	1025
C3540	42.0	38561	0.18	123	23	0.23	181	42	0.22	23582
C5315	31.0	42032	0.12	160	5	0.13	173	36	0.16	21449
C6288	110.0	45343	0.19	1131	2	0.19	1699	2	0.19	10539
alu2	8.0	13340	0.03	13	23	0.04	20	35	0.03	753
alu4	12.0	23317	0.06	65	16	0.07	70	40	0.07	1525
dalu	27.0	35812	0.12	68	21	0.15	104	39	0.17	1419
Ave.			0.12		21.9	0.17		35.1	0.16	

Run Time (sec)
E[I] in pA

Maximum allowed risk (probability) for violating the timing constraint: 0.3

Conclusions

- Introduced pruning probability as metric to compare potential solutions in a variability-driven optimization framework
- Illustrated computing of pruning probability:
  - Using efficient *jpdf* approximation
  - Using accurate Conditional Monte Carlo simulation
  - Both methods significantly faster the MC

Thank You For Your Attention!

For details please contact the authors:
{azade,ankurs}@eng.umd.edu