Statistical Leakage Minimization through Joint Selection of Gate Sizes, Gate Lengths and Threshold Voltage

Sarvesh Bhardwaj, Yu Cao, Sarma Vrudhula

NSF Center for Low Power Electronics

Arizona State University, Tempe, AZ

{sarvesh.bhardwaj, yu.cao, vrudhula}@asu.edu

Motivation

- Leakage has been increasing exponentially with every technology generation
- Process Variations cause large parametric yield loss due to leakage
  - Difference between specified and observed values
- Need for Statistical Circuit Optimization techniques for reducing leakage without paying Performance penalty
- Improve the parametric yield of leakage by performing Gate Sizing, Gate Length biasing and optimal Threshold Voltage selection

Outline

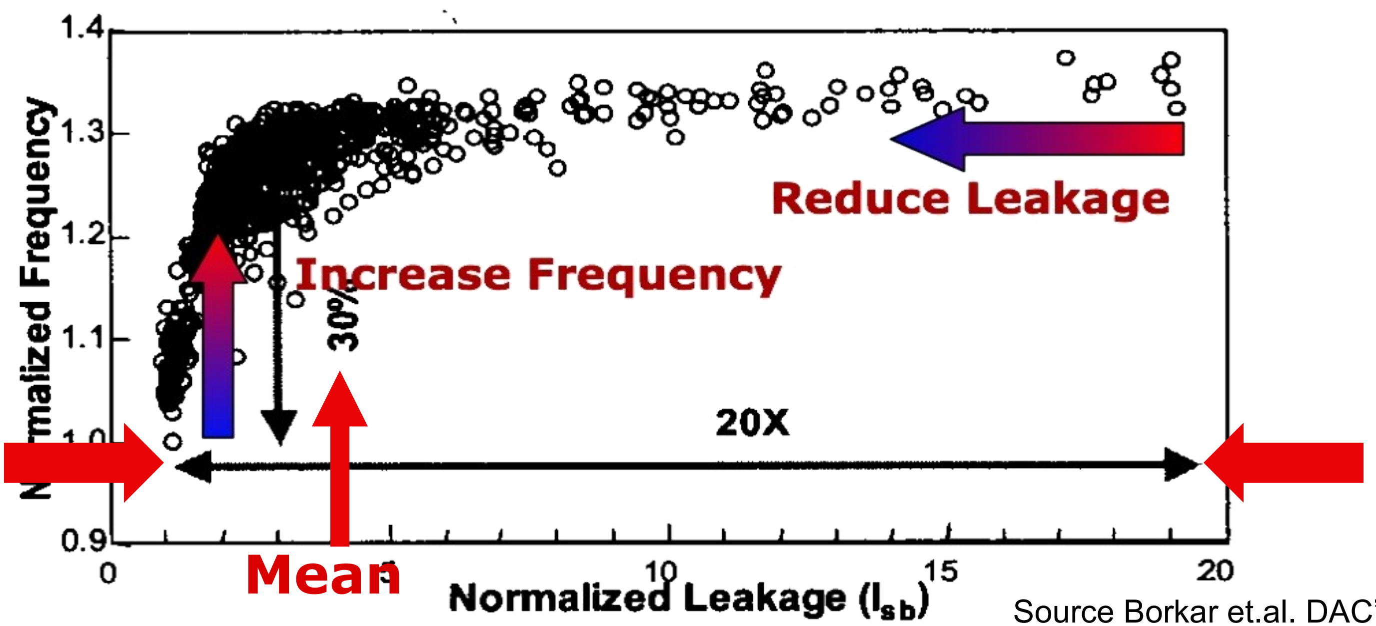
- Statistical Leakage Minimization Problem: What to minimize?
  - Objective: What to minimize?
  - Constraints: Including Performance Constraints
- Computation of the leakage objective
- Representation of the delay constraints
- Formulation as Convex Optimization,
- Experimental Results
- Conclusion and Future Work

Previous Work

- Circuit Design Techniques:
  - Transistor Stacking: *Narendra et. al.* - JSSC'04
  - Sleep Transistor Insertion: *Long et. al.* - DAC'03
  - Body biasing: *Neau and Roy* - ISLPED'03
- Circuit Optimization using Gate Sizing and dual-Vth assignment:
  - **Deterministic**: *Chen et. al.*(TCAS'02), *Ketkar et. al.*(ICCAD'02) and *Sapatnekar et. al.*(TCAD'93)
  - **Statistical**: *Raj et. al.*(DAC'04), *Srivastava et. al.*(DAC'04), *Patil et. al.* (ISQED'05), *Singh et.al.* (DAC'05) and *Mani et.al.*(DAC'05)

Mean Variance Optimization

- Yield limited by both: Delay and Leakage



Source Borkar et.al. DAC'03

- Minimize both Mean and Variance of Leakage

$$\lambda \cdot \mu^2(\ell) + (1-\lambda) \cdot \sigma^2(\ell)$$

Optimization Problem

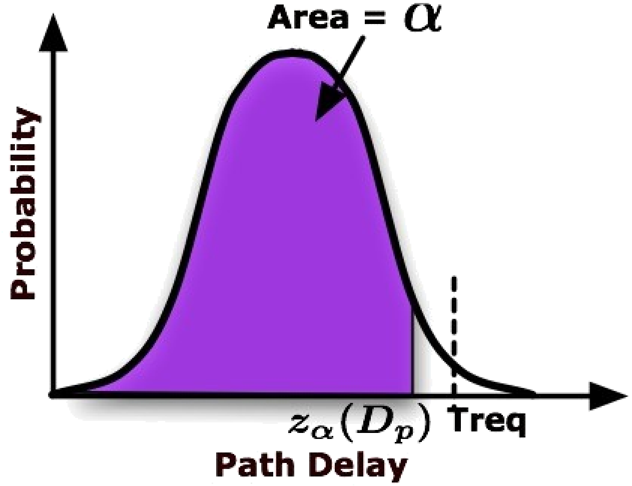
Mean → $\mu^2(\ell)$; Variance → $\sigma^2(\ell)$

$$\min_{W,L,V_t} \quad \lambda \cdot \mu^2(\ell) + (1-\lambda) \cdot \sigma^2(\ell)$$

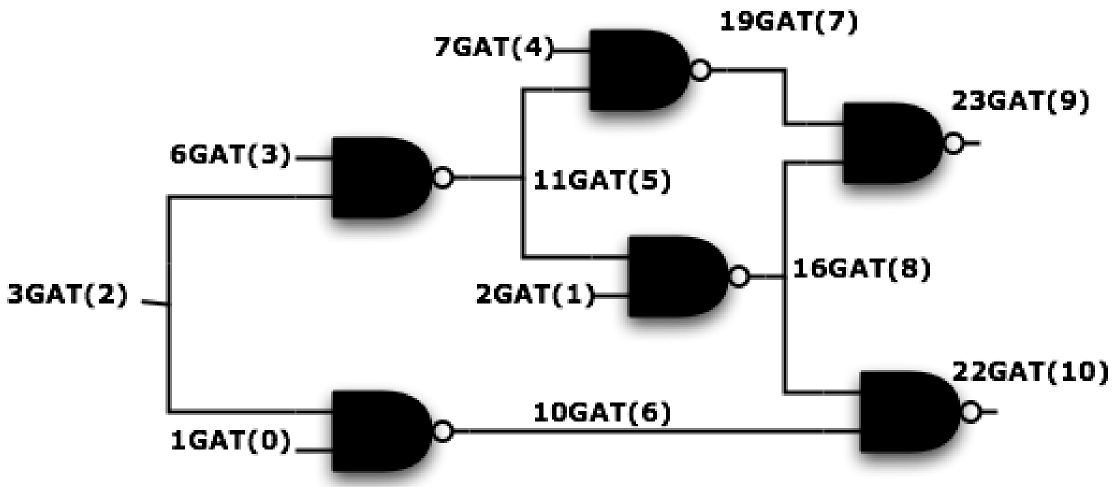
$$\text{subject to} \quad P(D_p \leq T_{req}) \geq \alpha, \ \forall\, p \in \mathcal{P}$$

Constraint on the parametric yield of path delays

$$z_\alpha(D_p) \leq T_{req}$$



Model for Parameter Variations

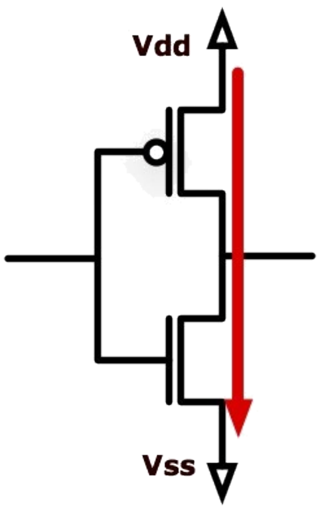


Random Component $\sim N(0, \sigma^2)$

Gate Length:	$L_{e,i} = L_{o,i} + L_{\xi}$	
Threshold Voltage:	$V_{th,i} = V_o + V_{\xi,i}$	
Gate Size:	W_i	

Decision Variables

Statistical Leakage Model



$$\ell = \sum_{i \in N} I(w_i, L_{o,i}, V_o) \cdot \exp\left(\frac{-(\alpha_i L_\xi + a L_\xi^2 + b V_{\xi,i})}{c}\right)$$

Deterministic function (the term $I(w_i, L_{o,i}, V_o)$) — Random function (the exp term)

$$I(w_i, L_{o,i}, V_o) = I_{So} w_i \exp\left(\frac{-(L_{o,i} + a L_{o,i}^2 + b V_o)}{c}\right)$$

also used in Rao et al DAC'04

- **Mean:**

$$E[\ell] = \sum_{i \in N} I(w_i, L_{o,i}, V_o) \cdot E\left[\exp\left(\frac{-(\alpha_i L_\xi + a L_\xi^2 + b V_{\xi,i})}{c}\right)\right]$$

- **Second Moment:**

$$E[\ell^2] = \sum_{i,j \in N} I_i I_j \cdot E\left[\exp\left(\frac{-((\alpha_i + \alpha_j) L_\xi + 2a L_\xi^2 + b(V_{\xi,i} + V_{\xi,j}))}{c}\right)\right]$$

Statistical Gate Delay Model

- Physical Gate Delay Model proposed in *Cao et. al., DAC'05*
- Mean Delay (saturation mode)

$$E[d_i] = \alpha\left(\frac{\beta_1}{w_i} + \beta_2\right)\frac{L_{o,i}V_{dd}}{(V_{dd} - V_o)^2}\left(1 + \frac{(V_{dd} - V_o)}{\gamma L_{o,i}}\right)$$

Size dependent term

- Delay Variance
  - For low Vth, both delay variance and mean delay is high

$$\frac{\sigma(d_i)}{E[d_i]} = k_\sigma\left(E[d_i]\right)^\delta, \quad \delta > 0$$

Computation of α-percentile of Path Delays

- Path delay is sum of gate delays

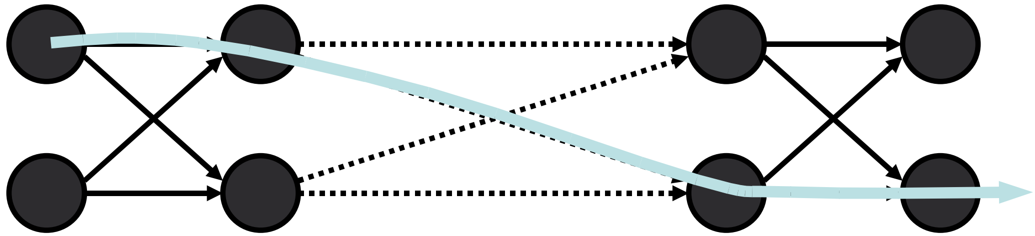
$$D_p = d_1 + d_2 + \cdots + d_n$$

- Using Central Limit Theorem
  - Path delay can be approximated as a Normal Random Variable for circuit depth of ~10
  - α-percentile of the path delay

$$\begin{aligned} z_\alpha[D_p] &= E[D_p] + \Phi^{-1}(\alpha) \cdot \sigma(D_p) \\ &= \sum_{i \in p} E[d_i] + \Phi^{-1}(\alpha) \cdot \left(\sum_{i,j \in p} C_{ij} \right)^{0.5} \end{aligned}$$

Upper bound on α-percentile of path Delay

- Potentially exponential number of paths in the number of nodes



- Avoid enumerating all the paths by obtaining an upper bound on the $z_\alpha[D_p]$ of path delay

$$z_\alpha[D_p] = \sum_{i \in p} E[d_i] + \Phi^{-1}(\alpha) \cdot \left(\sum_{i,j \in p} C_{ij} \right)^{0.5}$$

Upper bound on $z_\alpha[D_p]$

- Simple upper bound on standard deviation of a path delay

$$\left(\sum_{i,j\in p} C_{ij}\right)^{0.5} \le \sum_{i\in p} \sigma(d_i)$$

- Upper bound on the path delay,

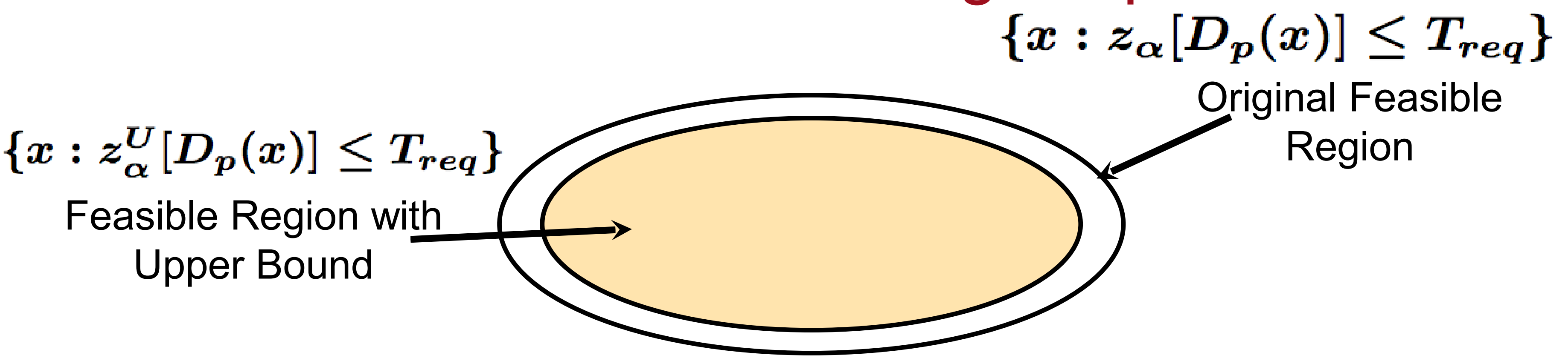
Before $$z_\alpha[D_p] = \sum_{i\in p} E[d_i] + \Phi^{-1}(\alpha)\cdot\left(\sum_{i,j\in p} C_{ij}\right)^{0.5}$$

After $$z_\alpha^U[D_p] = \sum_{i\in p}\left(E[d_i] + \Phi^{-1}(\alpha)\cdot\sigma(d_i)\right)$$

Assign this delay to each gate

Advantages of using Upper Bound

- Transform the path based delay constraints into node based constraints
  - Path based constraint: Exponential in ***N***
  - Node based constraints: Linear in ***N***
- The optimal solution of the new problem is a feasible solution of the original problem



Convexity of Objective Function

- General form of Mean and Second Moment of Leakage (*a*,*b*,*c* and *d* are positive)

$$\sum_{i \in N} k_i e^{\left((aL_{o,i} - b)^2 - c \cdot V_o - d \cdot z_i\right)}, \quad \text{where} \quad z_i = \log w_i$$

Convex Function + Linear Function

Convex Function

- Exponential of a convex function is also convex
- Objective Function $\lambda \cdot \mu^2(\ell) + (1 - \lambda) \cdot \sigma^2(\ell)$
  - Rewrite as $(2\lambda - 1) \cdot \mu^2(\ell) + (1 - \lambda) \cdot E[\ell^2]$
  - Convex for $0.5 \leq \lambda \leq 1.0$

Convexity of Delay Constraints

- Delay constraints

$$z_\alpha^U[D_p] = \sum_{i \in p} \left(E[d_i] + \Phi^{-1}(\alpha) \cdot \sigma(d_i) \right) \leq T_{req}$$

$$E[d_i] = \alpha \left(\frac{\beta_1}{w_i} + \beta_2 \right) \frac{L_{o,i} V_{dd}}{(V_{dd} - V_{th,i})^2} \left(1 + \frac{(V_{dd} - V_{th,i})}{\gamma L_{o,i}} \right)$$

$$\frac{1}{V_{dd} - V_{th,i}} \leq t_i$$

Introduce variable t_i

Mean Delay: Posynomial

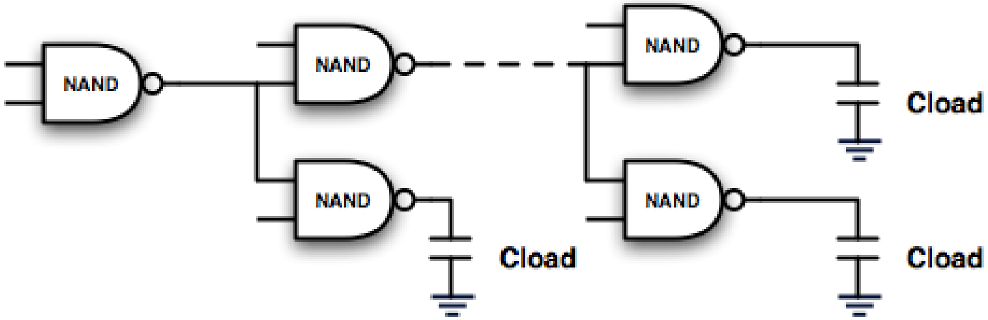
Transform

Convex Function

Valid posynomial Inequality

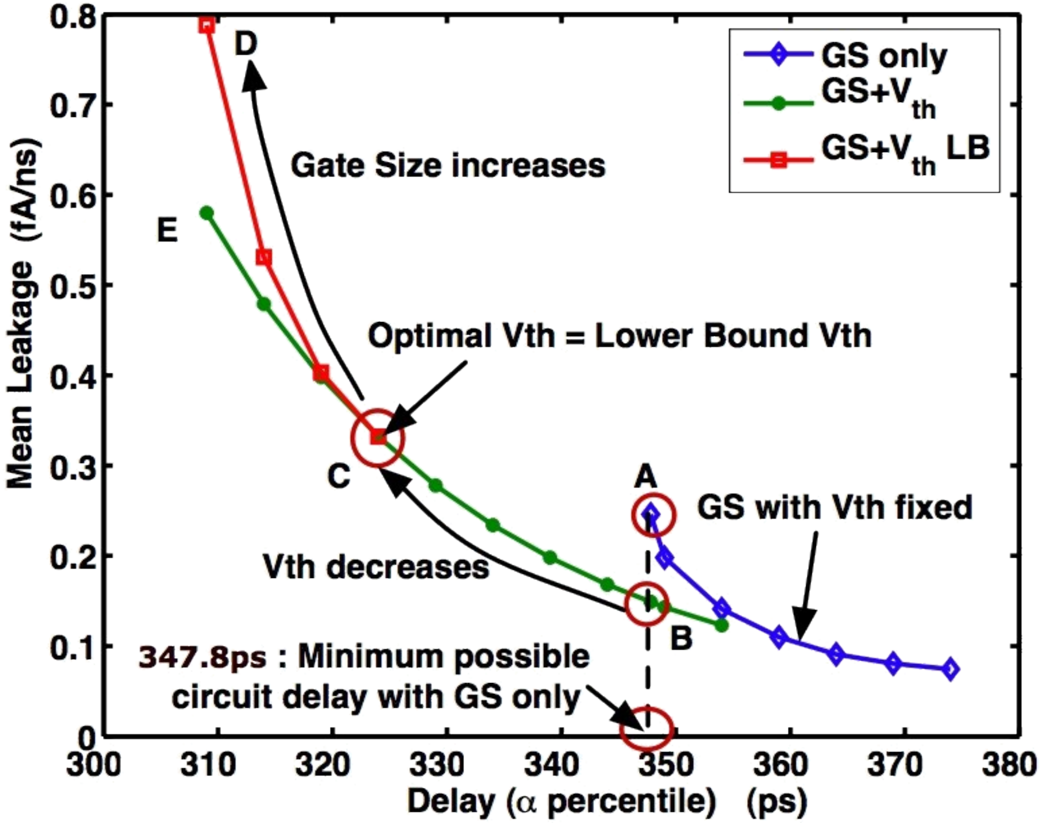
Experimental Results

- Compare the effect of different decision variables
  - GS: Gate Sizing
  - GSV: GS + Threshold Voltage Assignment
  - GSVL: GS + Thresh. Volt. + Gate Length Biasing
- Convex optimization problem solved using LANCELOT
- Single V_{th}, upto 30% gate length bias



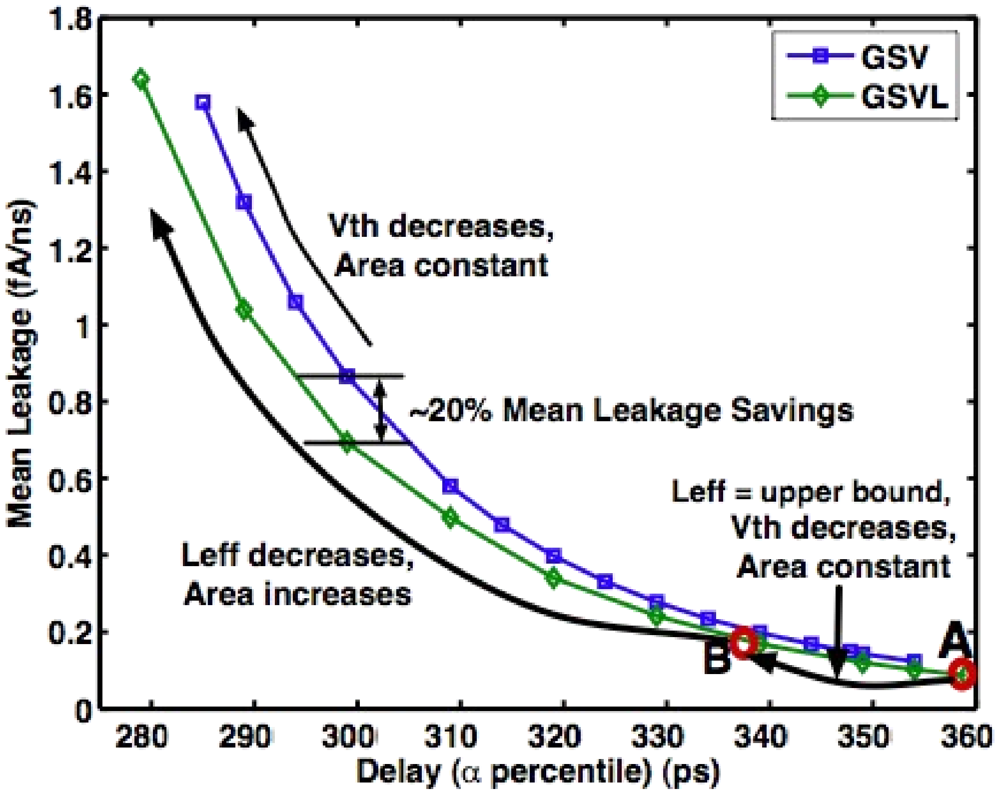
Comparison of GS with GSV

- Leakage savings with V_{th} and gate sizes as decision variables



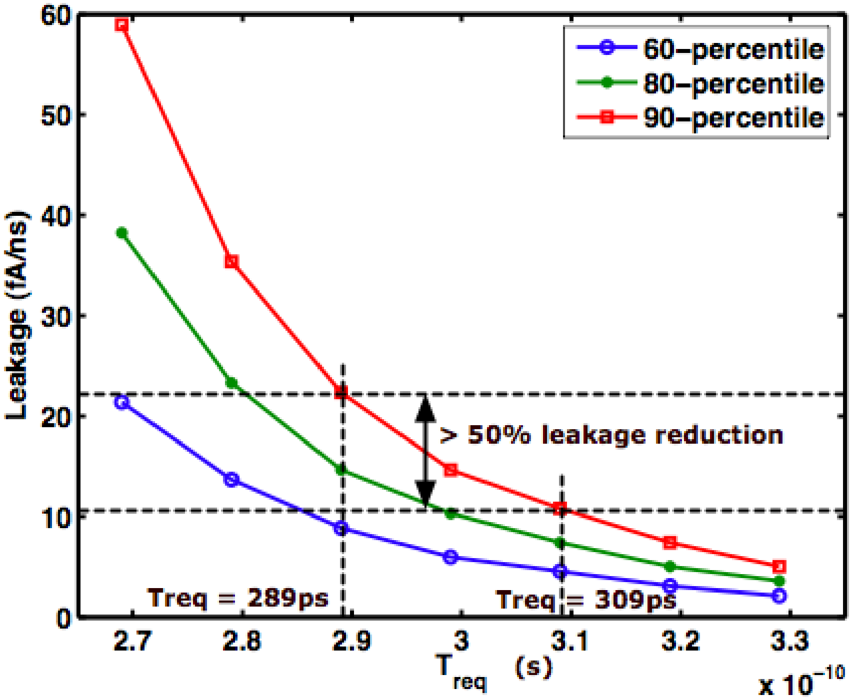
Comparison of GSV and GSVL

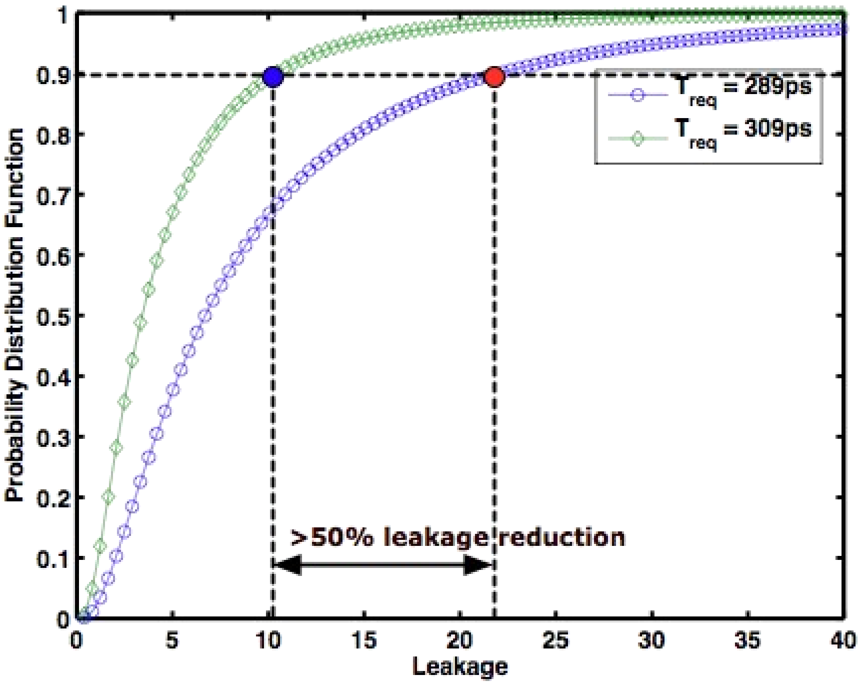
- Leakage savings with L_{eff}, V_{th} and gate sizes as decision variables



Leakage Delay Tradeoffs

- Small increase in delay can provide significant reduction in leakage





Conclusions and Future Work

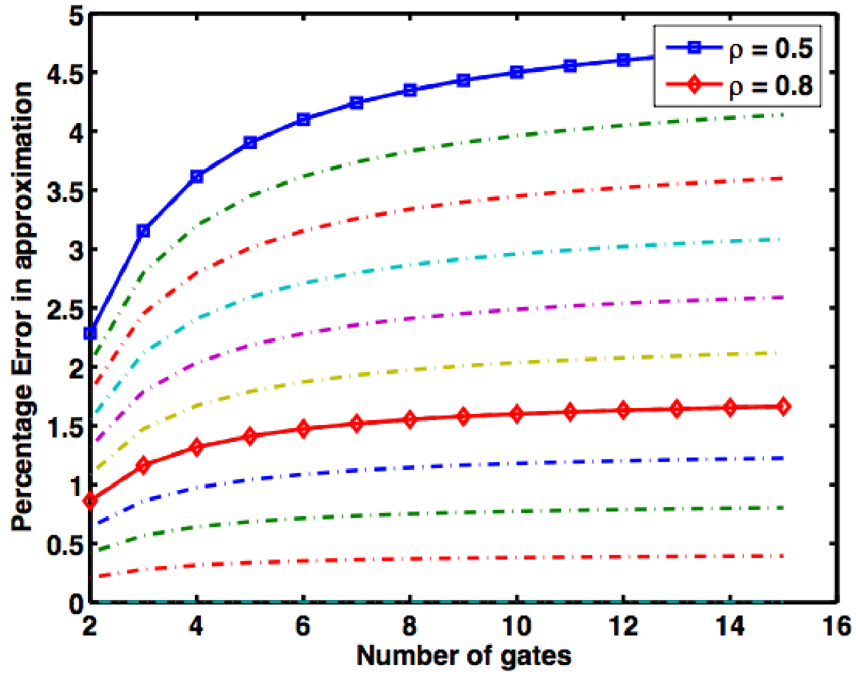
- Minimized leakage using GS, V_{th} and L_{eff} as decision variables
- Minimized both mean and variance of leakage
- Transformed the problem into a convex optimization problem
- Considerable leakage savings are obtained by introduction of V_{th} and L_{eff}

Thank You!

Questions
and
Answers

Tightness of the Upper Bound

- Difference between true ■-percentile and upper bound ~ 5%

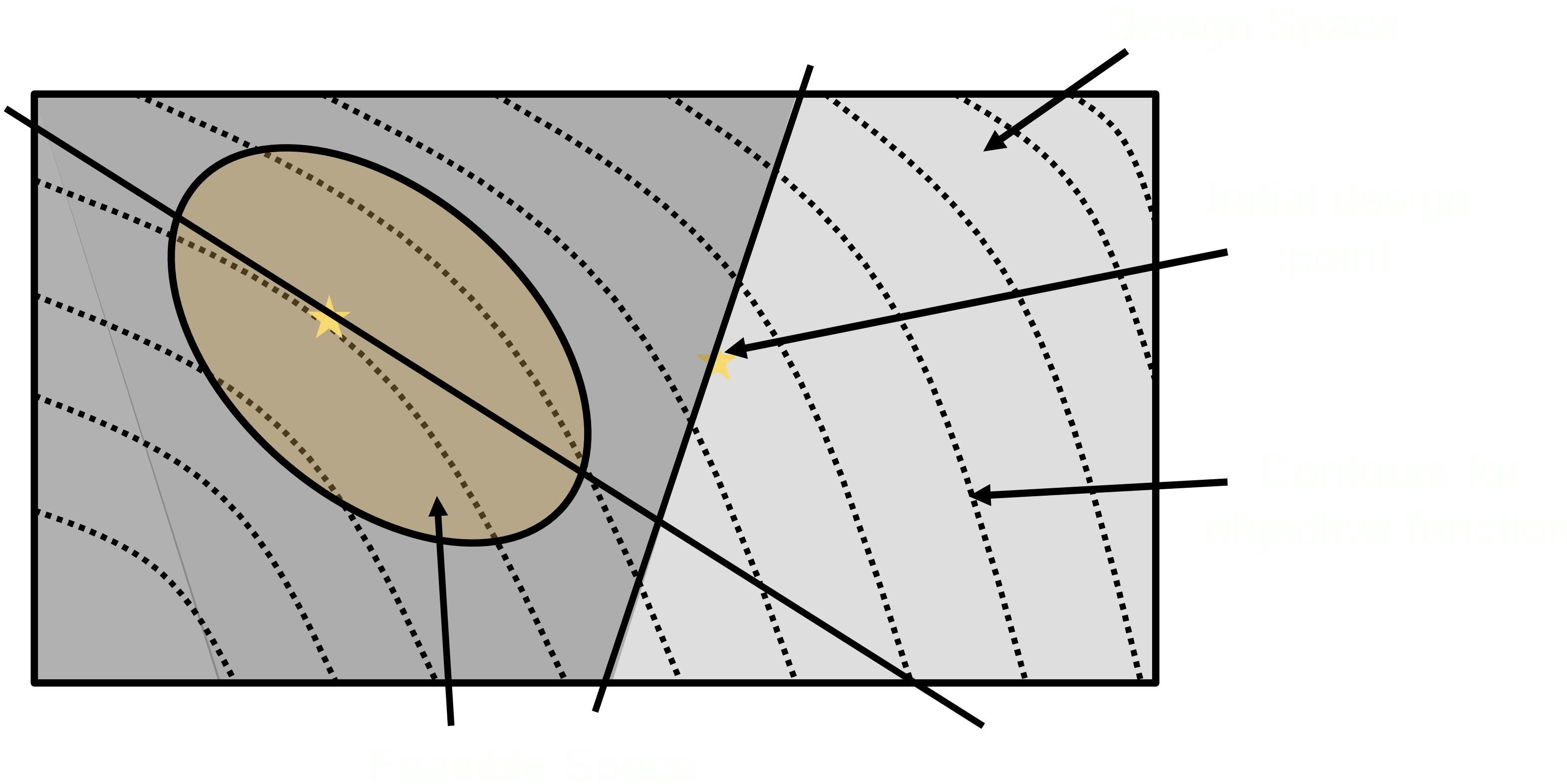


[redacted] of the Longest Path

- Assign each gate its [redacted]-percentile delay,
- Find the longest path using Dijkstra's Algorithm.

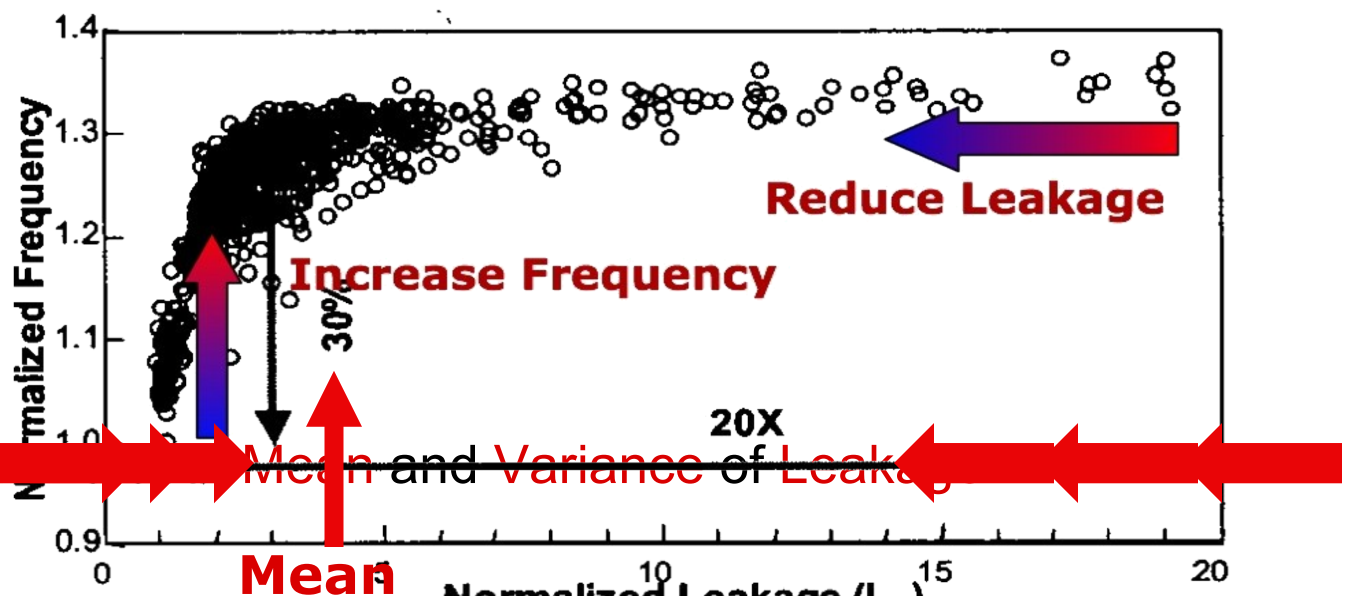
Optimization Algorithm

- Algorithm used in Sapatnekar et al - TCAD'93.



Mean Variance Optimization

- Yield limited by both: Delay and Leakage



Source Borkar et.al. DAC’03

- Min… Mean and Variance of Leak…

$$\lambda \cdot \mu^2(\ell) + (1 - \lambda) \cdot \sigma^2(\ell)$$