Robust Gate Sizing via Mean-Excess Delay Minimization

Jason Cong
John Lee
Lieven Vandenberghe
UCLA

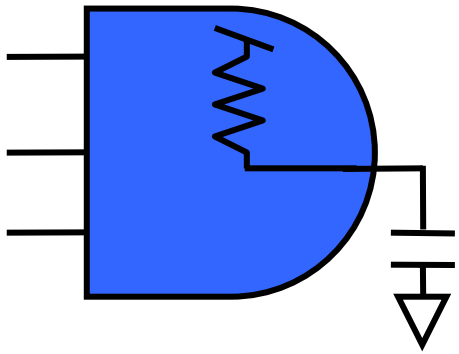
Outline

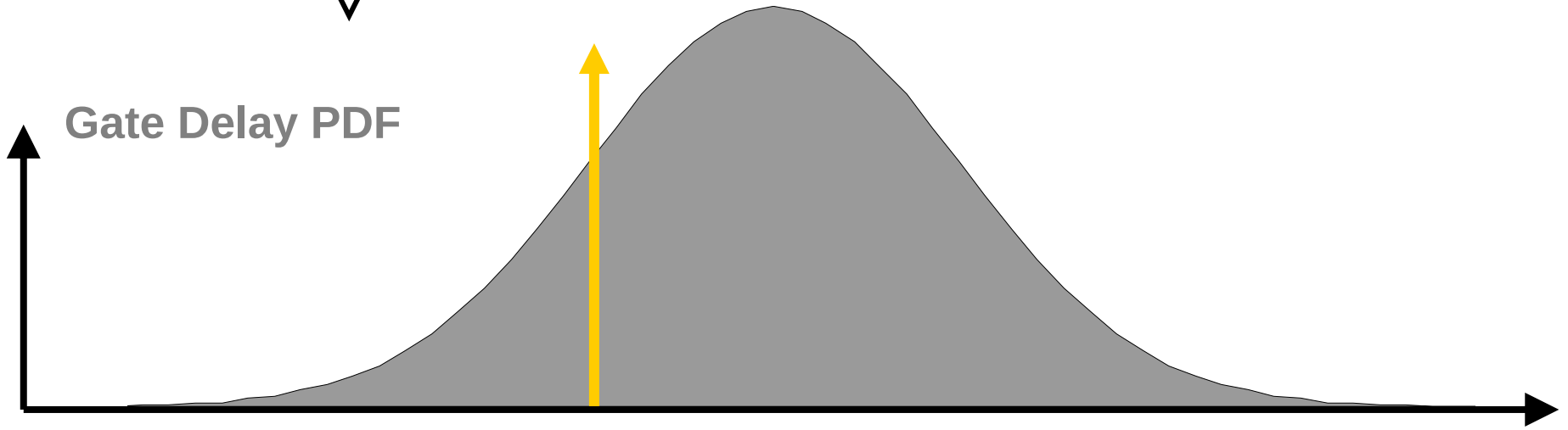
1. Gate Sizes and Delay Variations
2. Robust Circuit Sizing
3. Approaches to Robust Sizing
4. The Mean Excess Delay
5. Mean-Excess Delay Optimization
6. Results
7. Conclusion

UCLA

Gate Sizes and Delay Variations

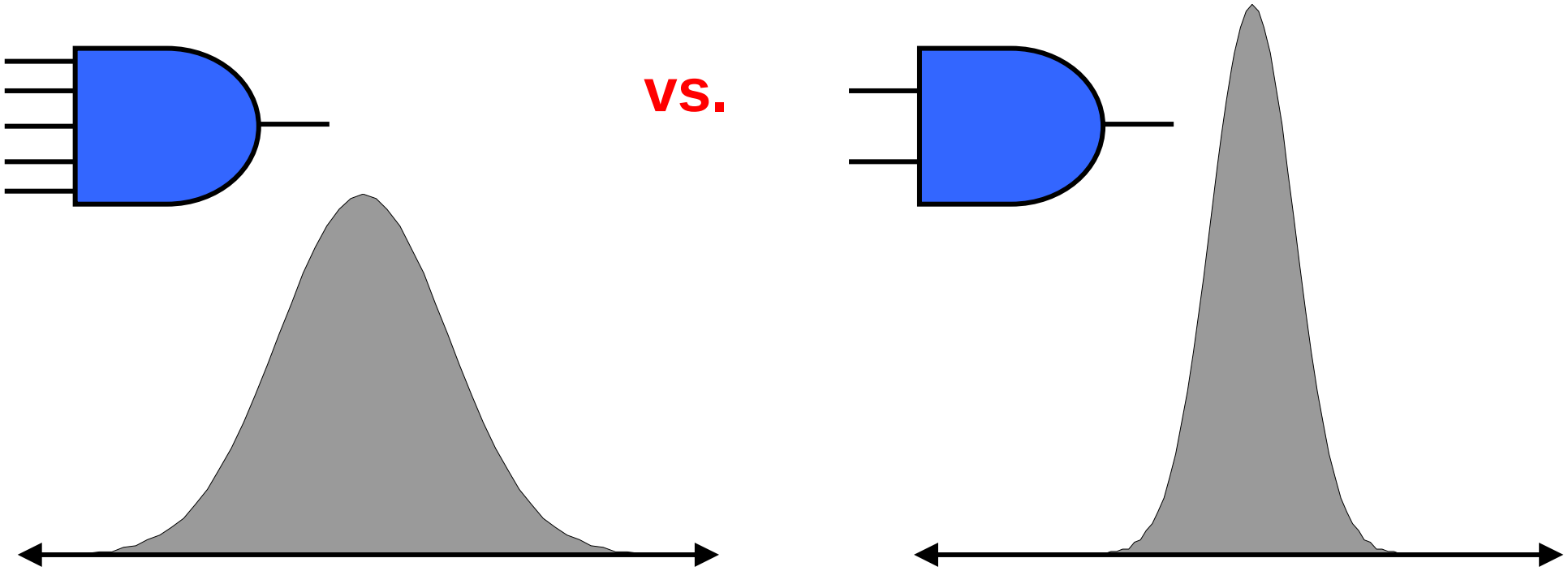
- Process Variations cause the gate delay to vary





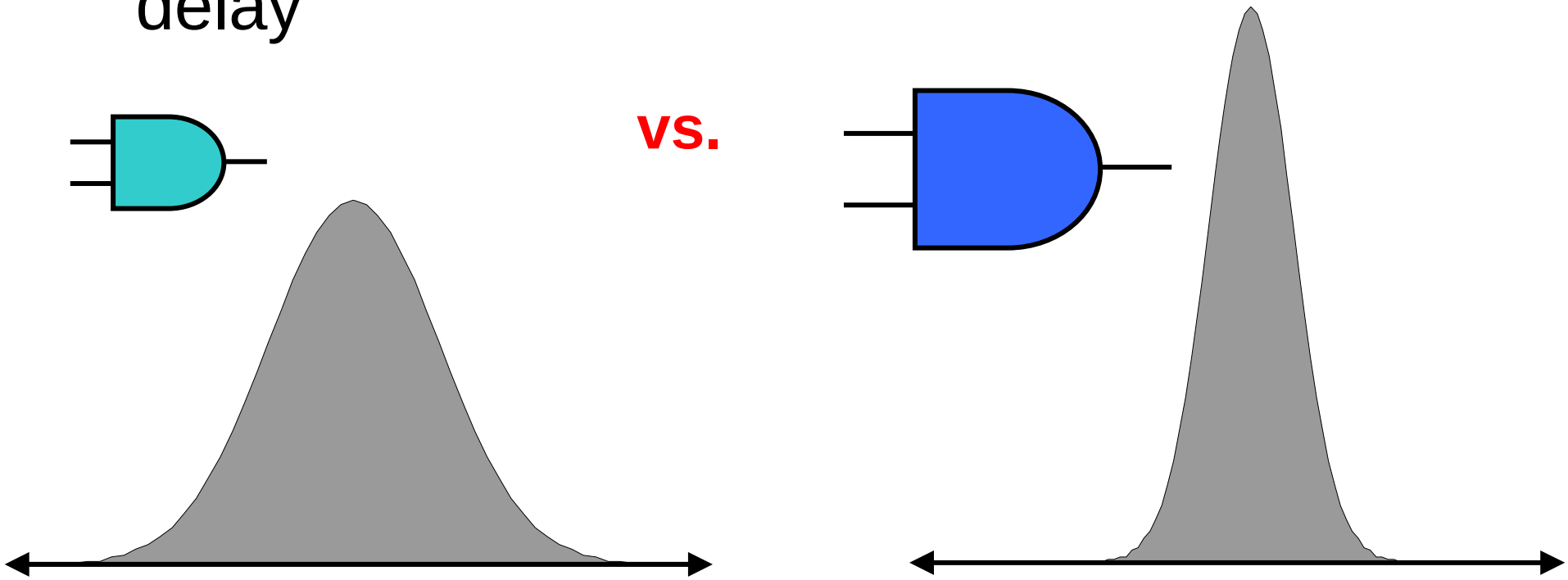
Gate Sizes and Delay Variations

- Different types of gates may have different sized variations.



Gate Sizes and Delay Variations

- Larger gates may have less variation in the delay

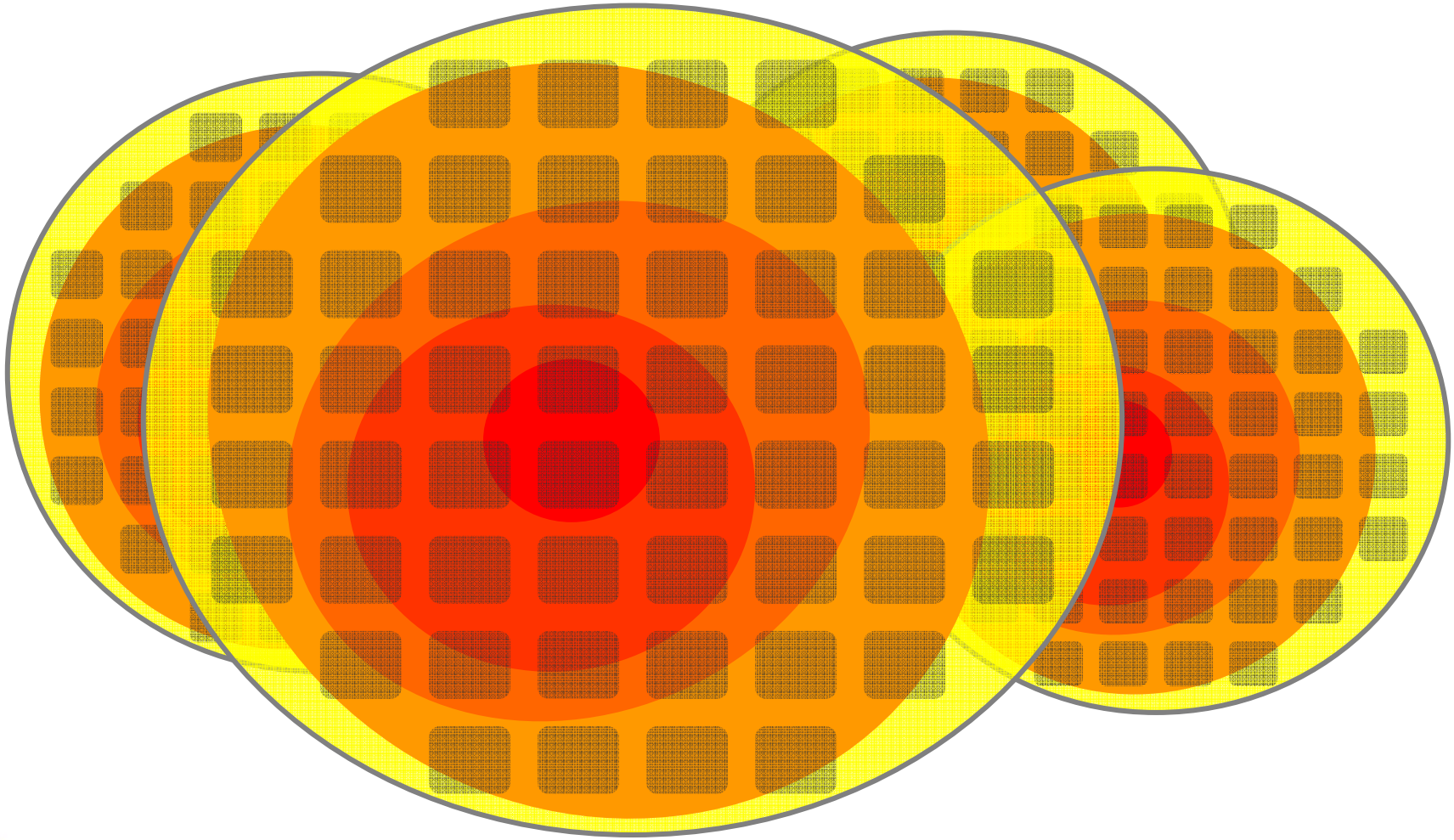


$$\sigma = \frac{1}{\text{size}^{\alpha}}, \alpha \approx 0.3 - 0.5$$

[Pelgrom 85], [Wang et. al 08]

Gate Sizes and Delay Variations

- These variations may be arbitrarily correlated across dies, and within a die



UCLA

Gate Sizes and Delay Variations

Robust Circuit Sizing:

- Correct for gates and critical paths with large variations
- Apply Pelgrom's Law to reduce delay variations in critical gates
- Account for correlations

Robust Circuit Sizing

Select Gate Sizes to balance:

Delay

Power

Yield

UCLA

Robust Circuit Sizing

Topic of this research:

Delay **Minimize**

Power **< p_0 mW**

Yield **> 95%**

UCLA

Approaches to Robust Sizing

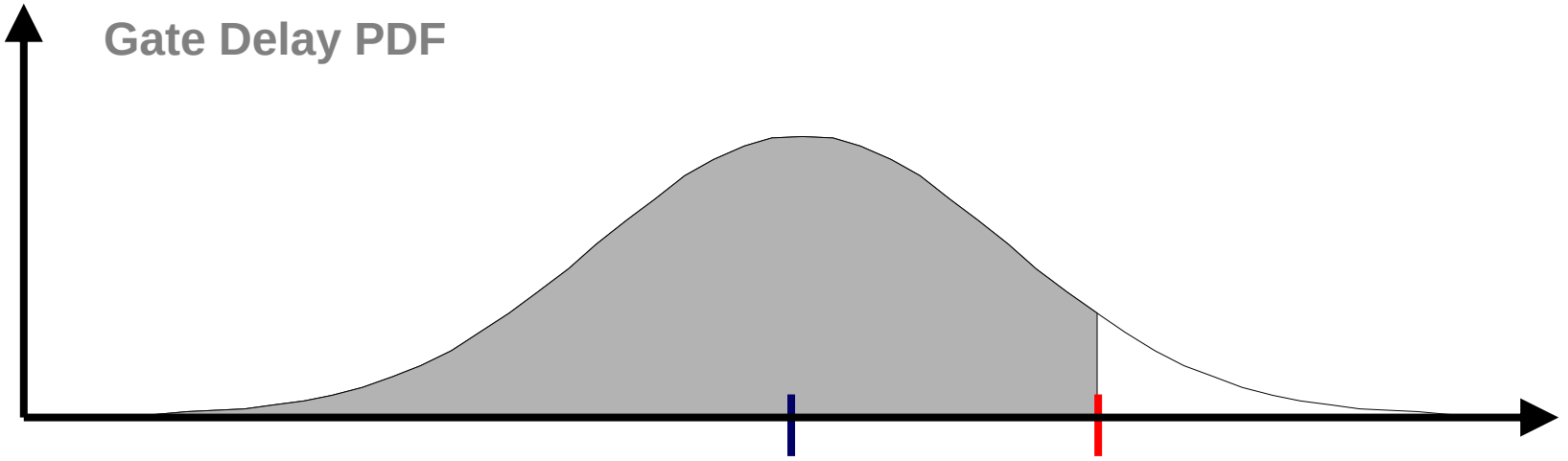
1. Scenario Based
   - Corners
     - Identify process and environment parameters “corners”
     - Design must meet constraints at each corner
   - Multimode
     - Similar to Corners
     - Design must meet a probabilistic blend of the corners
     - More corners improve the design

UCLA

cf. [Boyd et. al 05]

Approaches to Robust Sizing

2. Deterministic Estimates of gate delay
   - Add “padding” – a multiple of the standard deviation – to each ***gate delay*** to account for variation



“Padded” delay = mean + k · standard deviation

Approaches to Robust Sizing

2. Deterministic Estimates of gate delay

Geometric Program

- Patil et. al, “A New Method for Design of Robust Digital Circuits”, ISQED ’05

Geometric Program with linearized constraints

- Singh et. al., “Robust Gate Sizing by Geometric Programming”, DAC ‘05

Linear Program

- Mani and Orshansky, “A New Statistical Optimization Algorithm for Gate Sizing”, DAC ‘04

Approaches to Robust Sizing

3. Stochastic Programming Methods
   - Gate delays left as **distributions**
   - Statistical Static Timing Analysis (SSTA) used to work with circuit delay distribution

Approaches to Robust Sizing

3. Stochastic Programming Methods

Yield Maximization

- Sinha, et al. “Statistical Gate Sizing for Timing Yield Optimization”, ICCAD ’05
- Davoodi and Srivastava, “Variability Driven Gate Sizing for Binning Yield Optimization”, DAC ’06
- Chopra et al. “Parametric Yield Maximization using Gate Sizing based on Efficient Statistical Power and Delay Gradient Computation”, ICCAD ‘05

Power Minimization with Statistical Delay

- Guthaus et al, “Gate Sizing Using Incremental Parameterized Statistical Timing Analysis”, ICCAD ‘05

The Mean-Excess Delay: Quantiles

Delay	**Minimize**
Power	**< p_0 mW**
Yield	**> 95%**

- **> 95%** is a ***quantile*** *constraint*
  - ***Quantiles*** *are not convex*!*
  - *No easy expression!*

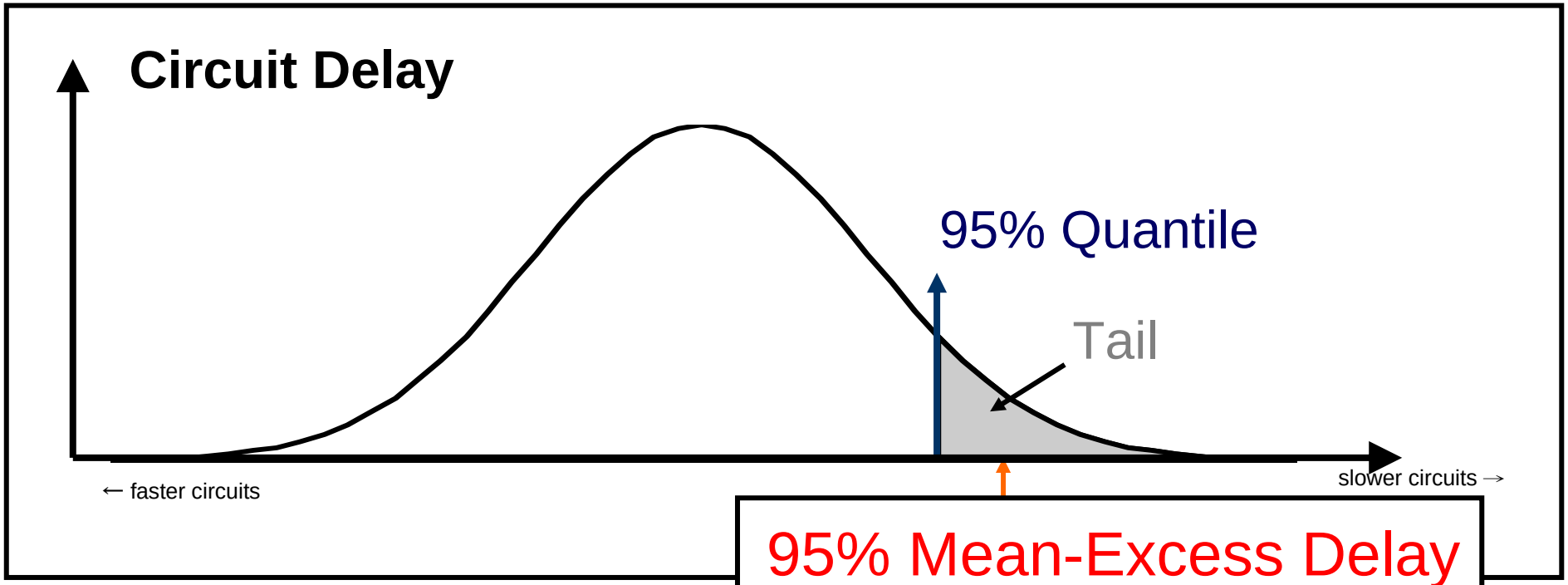
UCLA

**in general*

The Mean-Excess Delay

The Mean-Excess Delay

- Associated with a given quantile
- “Optimal upper bound” on the quantile



UCLA

The Mean-Excess Delay: Theorem

$$m_\beta(x) = \min_t \left\{ t + \frac{1}{1-\beta} \mathrm{E}_v \left[\left[T(x,v) - t \right]^+ \right] \right\}$$

- $m_\beta(x) =$ the Mean-Excess Delay with sizes x
- t : a helper variable β : The associated quantile
- $T(x, v)$: the circuit delay with gate sizes x and variation v
- The function $[u]^+ = \max\{0, u\}$

UCLA

The Mean-Excess Delay: Theorem

$$m_\beta(x) = \min_t \left\{ t + \frac{1}{1-\beta} \mathrm{E}_v \left[\left[T(x,v) - t \right]^+ \right] \right\}$$

Properties:

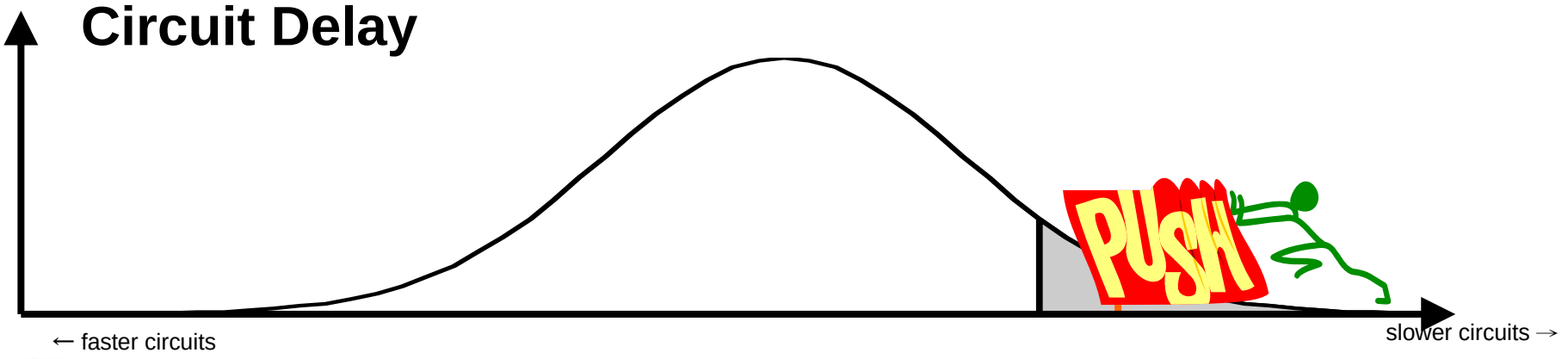
- If $T(x, v)$ is convex in x for fixed v then the problem $\min_x \left\{ m_\beta(x) \right\}$ is jointly convex in t, and x
- The minimizer t is the β-quantile

UCLA

The Mean-Excess Delay

This allows us to write the **convex** problem:

$$\begin{aligned} \underset{x,\,t}{\text{Minimize}} \quad & t + \frac{1}{1-0.95}\,\mathrm{E}_v\left[\left[T(x,v)-t\right]^+\right] \\ \text{Subject to} \quad & \text{Power}(x) \;\le\; p_0 \\ & 1 \le\; x \;\le\; x_{\max} \end{aligned}$$



UCLA

The Mean-Excess Delay & BYL

$$m_\beta(x) = \min_t \left\{ t + \frac{1}{1-\beta} \mathrm{E}_v \left[\left[T(x,v) - t \right]^+ \right] \right\}$$

- For fixed t this is equivalent to the Bin-Yield Loss function*:

$$\mathrm{BYL}(x) = \mathrm{E}_v \left[\left[T(x,v) - t \right]^+ \right]$$

(Used to Maximize Yield)

UCLA

** [Davoodi and Srivastava 06]*

The Mean-Excess Delay: Background

- Adapted from the finance and insurance industries
  - Applied to the risk in a portfolio
  - Called the “Mean-Excess Loss” or “Conditional Value-at-Risk”
    - Convex version of the quantile (“Value-at-Risk” (VaR))

Example

A “95% Value-at-Risk of $1,000,000” means there is a 5% chance of losing $1 million

UCLA

Mean-Excess Delay Optimization

$$\begin{aligned} \underset{x,\,t}{\text{Minimize}} \quad & t + \frac{1}{1-\beta}\,\mathrm{E}_v\left[\left[T(x,v)-t\right]^+\right] \\ \text{Subject to} \quad & \text{Power}(x) \;\le\; p_0 \\ & 1 \le\; x \;\le\; x_{\max} \end{aligned}$$

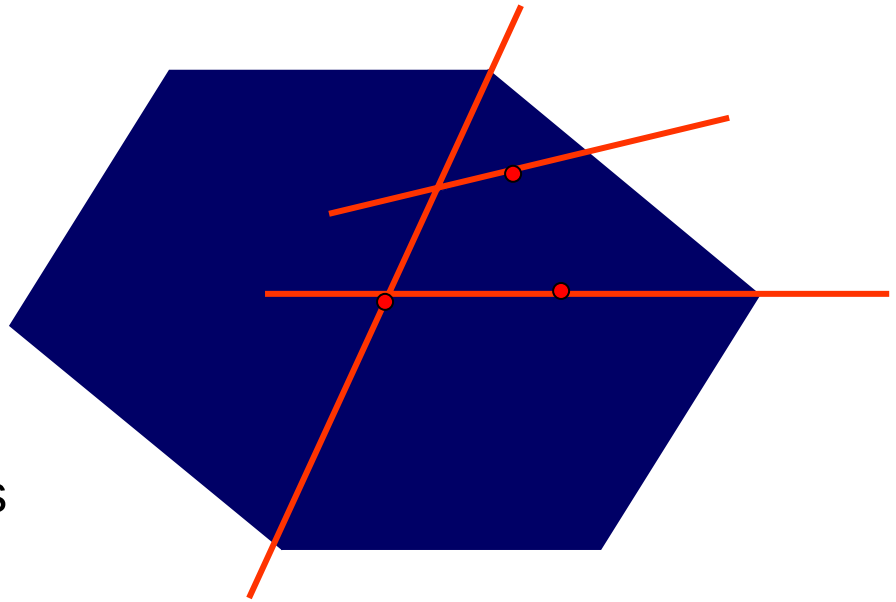
- Solved using the Analytic Center Cutting Plane Method (ACCPM)
  - Same class of algorithms used to solve the Bin-Yield Loss

UCLA

Mean-Excess Delay Optimization

Analytic Center Cutting Plane Method

1. The analytic “center” of the possible solutions is computed
2. Region is “cut” away using the gradient
3. Repeat until solution is found



Possible solutions

UCLA

Mean-Excess Delay Optimization

- Stochastic Programming Methods have used SSTA to compute the gradients
  - Rely heavily on approximations
  - Limits the type of distributions
- We use a Monte Carlo method to evaluate the gradient

Mean-Excess Delay Optimization

Monte Carlo based gradient computation

+

- No limit on distribution types or correlation types
- Sampling the distribution keeps the problem as a Geometric Program
- Faster performance if # samples ≈ # gates

–

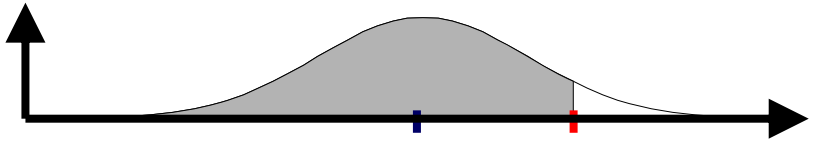
- It is a **randomized algorithm**, so performance may vary

UCLA

Results: Experiment

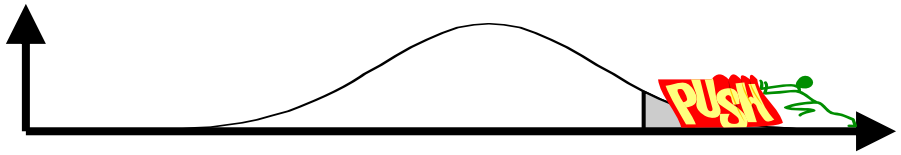
ISCAS ’85 Circuits

1. **Nominal Sizing (ignore variations)**
2. **Padded Sizing**



“Padded” delay = mean+k·standard deviation

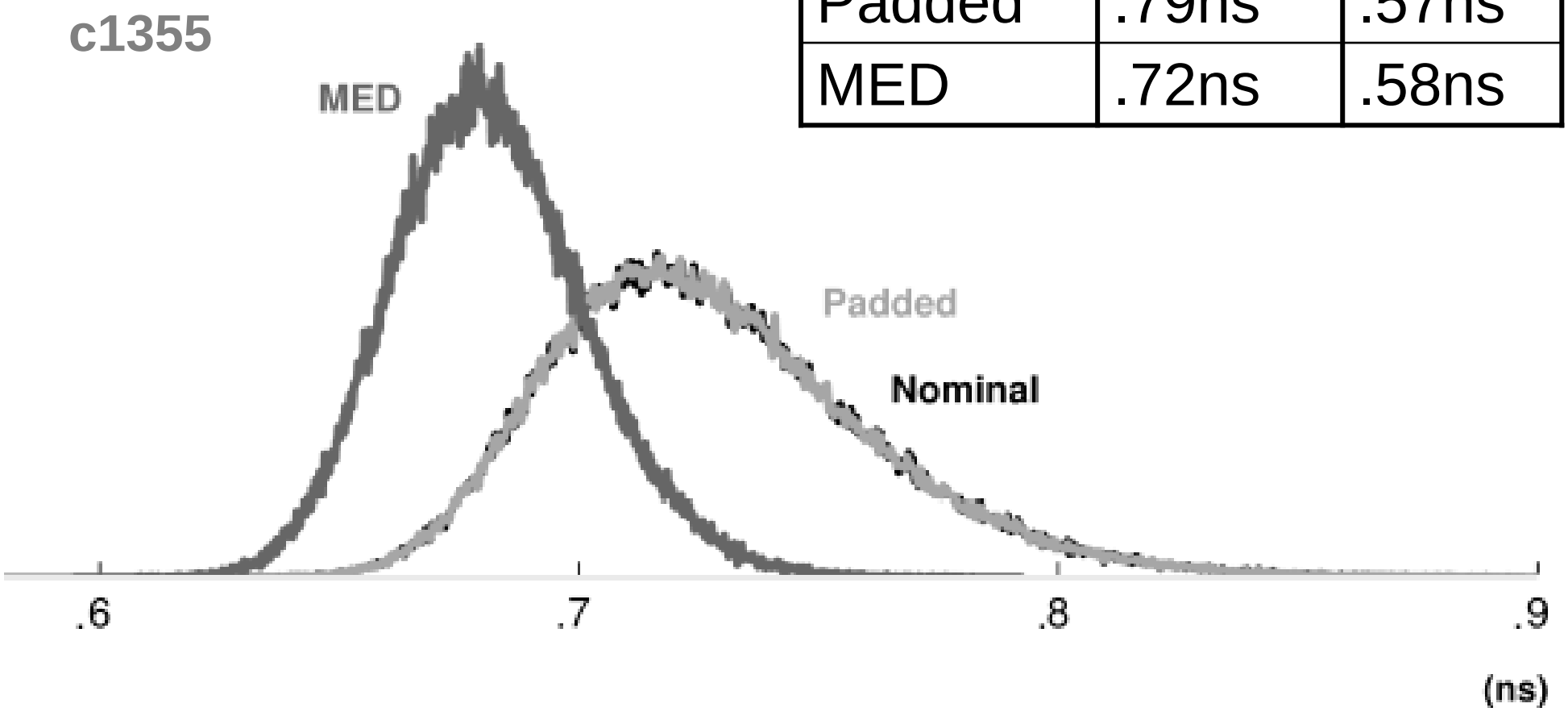
3. **Mean-Excess Delay Sizing**



UCLA

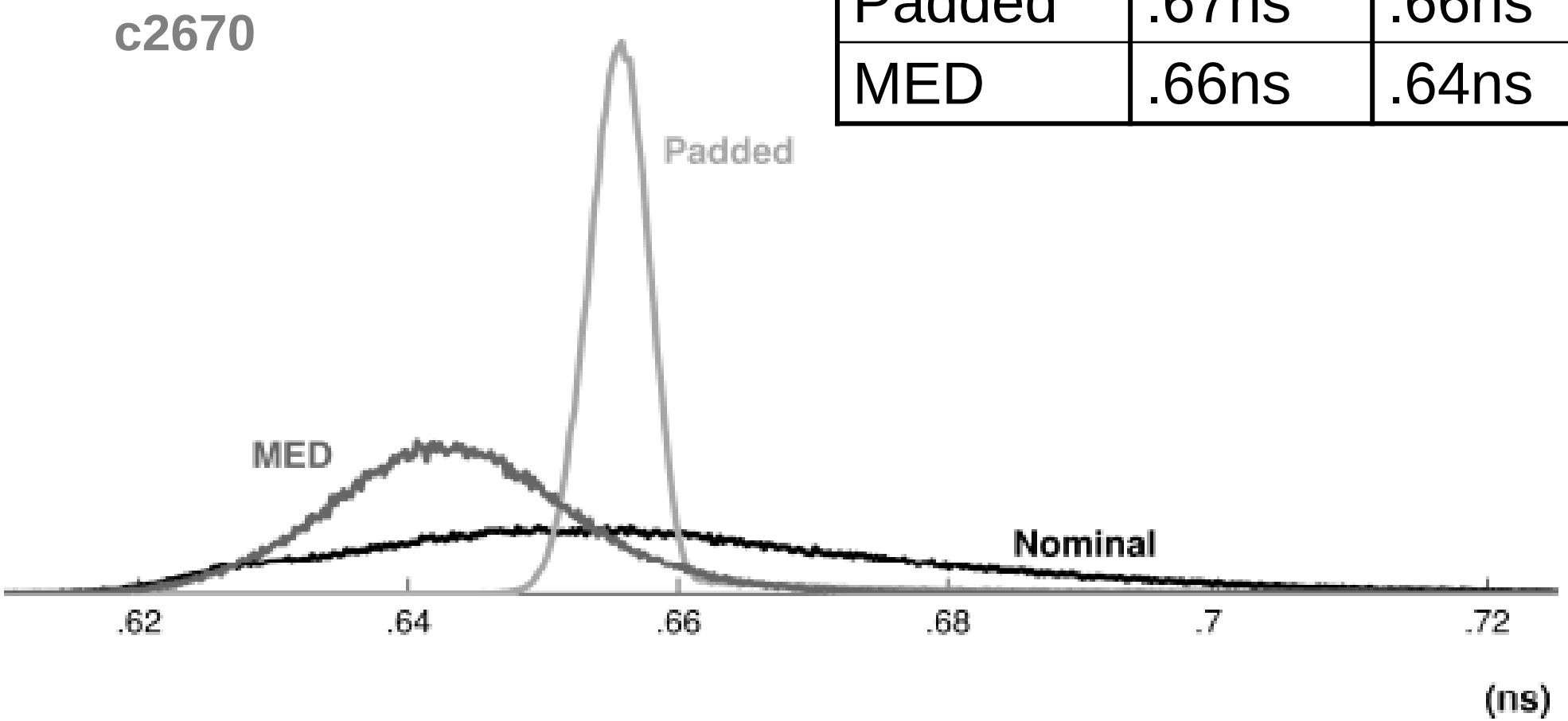
Results: Size independent variations

	c1355	
Method	$q_{0.95}$	d_{nom}
Nominal	.79ns	.57ns
Padded	.79ns	.57ns
MED	.72ns	.58ns



Results: Size dependent variations

	c2670	
Method	$q_{0.95}$	d_{nom}
Nominal	.70ns	.64ns
Padded	.67ns	.66ns
MED	.66ns	.64ns



Summary

- Applied the Mean-Excess Delay to the circuit sizing problem
- For size independent variations, “Padded” sizing is usually similar to nomimal sizing
- For size dependent variations, “Padded” methods are excellent at reducing the variance
- MED sizing can give improvements over the “padded” methods under both variation types

UCLA

Acknowledgements

Financial Support

- SRC contract 2006-TJ-1460
- NSF Award CCF-0528583

UCLA