A Heuristic Method for Statistical Digital Circuit Sizing

Stephen Boyd   Seung-Jean Kim   Dinesh Patil
Mark Horowitz

Statistical variation in digital circuits

• growing in importance as devices shrink

• modeling still open

 – many sources: environmental, process parameter variation,
   lithography
 – intrachip, interchip variation
 – distributions, correlations not well known, change as process matures

Statistical digital circuit sizing

- standard design approaches: margining, guardbanding, design over corners

- **statistical design** explicitly takes statistical variation into account (combines circuit design with design for manufacturing, yield optimization, design centering, . . .)

- statistical design is **very hard problem** (even for small circuits)

- this talk: a (relatively) simple heuristic method for statistical design that appears to work well

Outline

- A quick example

- Digital circuit sizing: models and optimization

- New method for statistical digital circuit sizing

- Digital circuit sizing example

- Conclusions and future work

A Quick Example

Example: Ladner-Fisher 32-bit adder

- 64 inputs, 33 outputs, 451 gates, 3214 paths, max depth 8

- simplified RC delay model

- design variables: 451 scale factors for gates

- cycle time T_{cycle} is max path delay

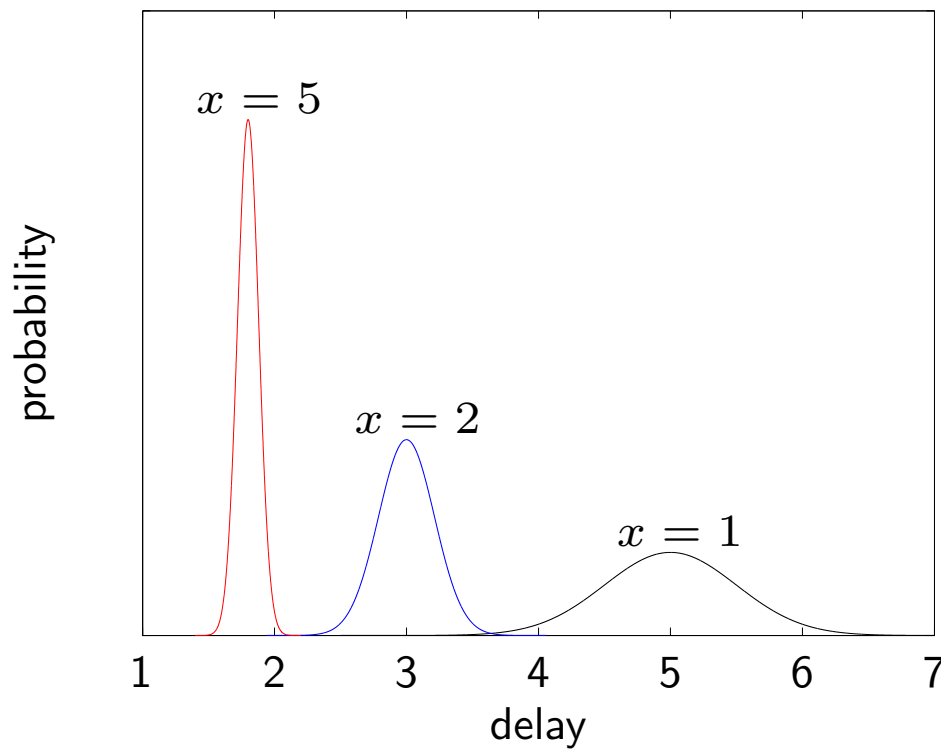- minimize cycle time subject to limits on area, min/max scale factor

Optimization results (no statistical variation)

path delays with optimized & uniform scale factors (same total area)
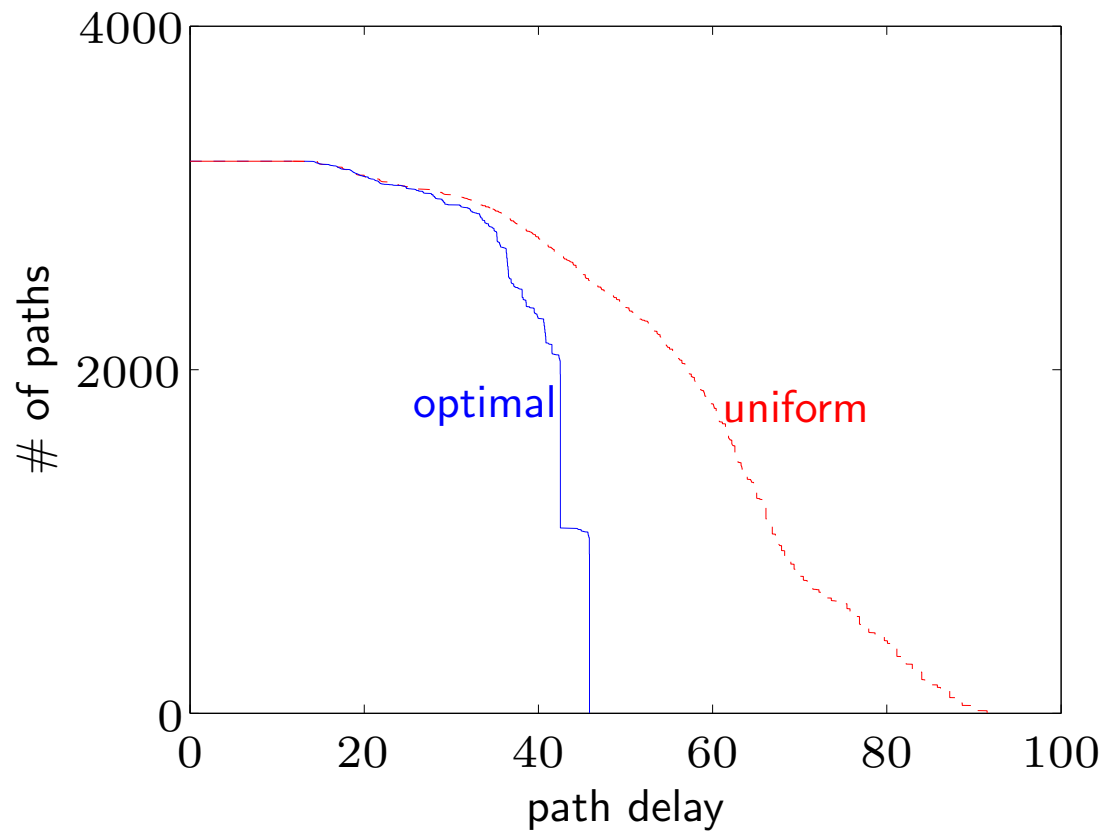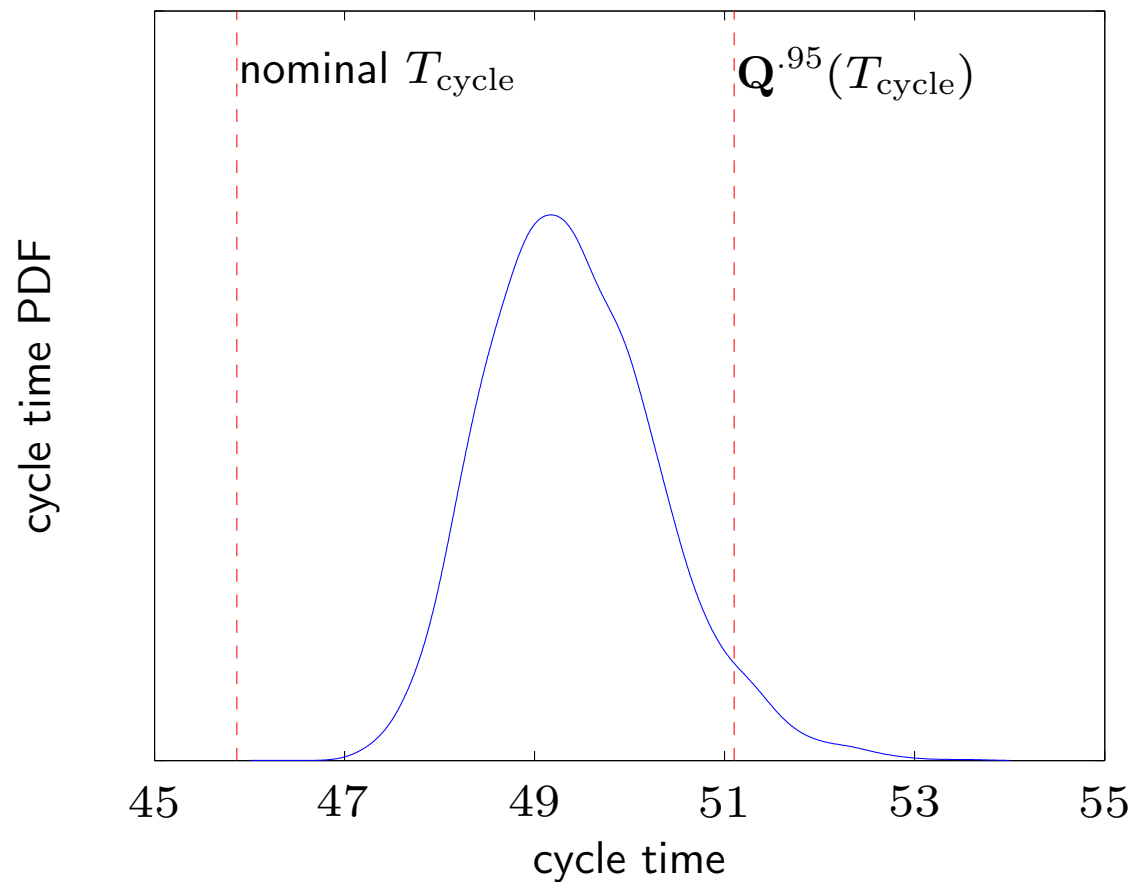
Statistical variation in gate delay

- simple Pelgrom model; larger gates have less (relative) variation in delay

- min sized gate has 10% variation

Effects of statistical variation on nominal optimal design

T_{cycle} PDF estimated via Monte Carlo

Why isn't T_{cycle} PDF centered around nominal value?

- T_{cycle} is **max** of 3214 random path delays

- **max of RVs** behaves differently from **sum of RVs**

  - in sum, negative and positive deviations tend to cancel out;
    PDF is centered, has *smaller relative variation*
  - in max, large deviation of *any* leads to large value;
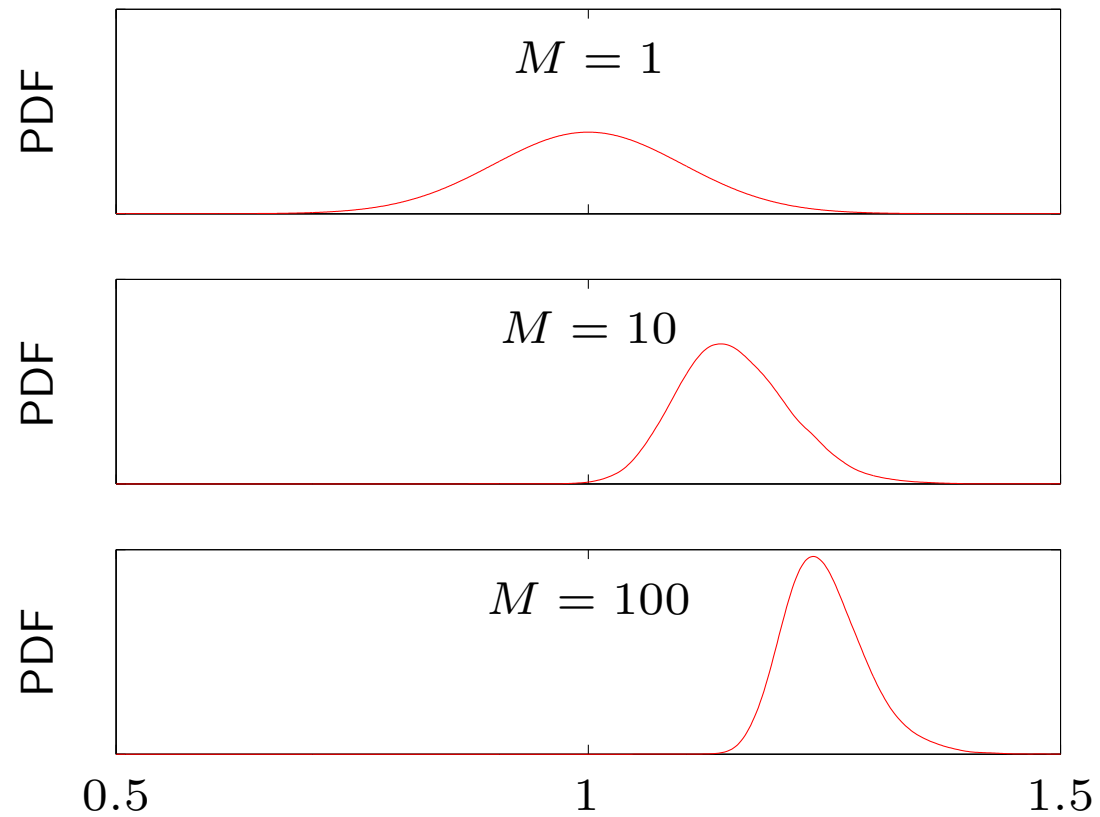   PDF is *shifted, skewed* to right, has *large relative deviation*

PDF of sum of random variables

$Z = \sum_{i=1}^{M} X_i$, $X_i \sim \mathcal{N}(1, 0.1)$ independent
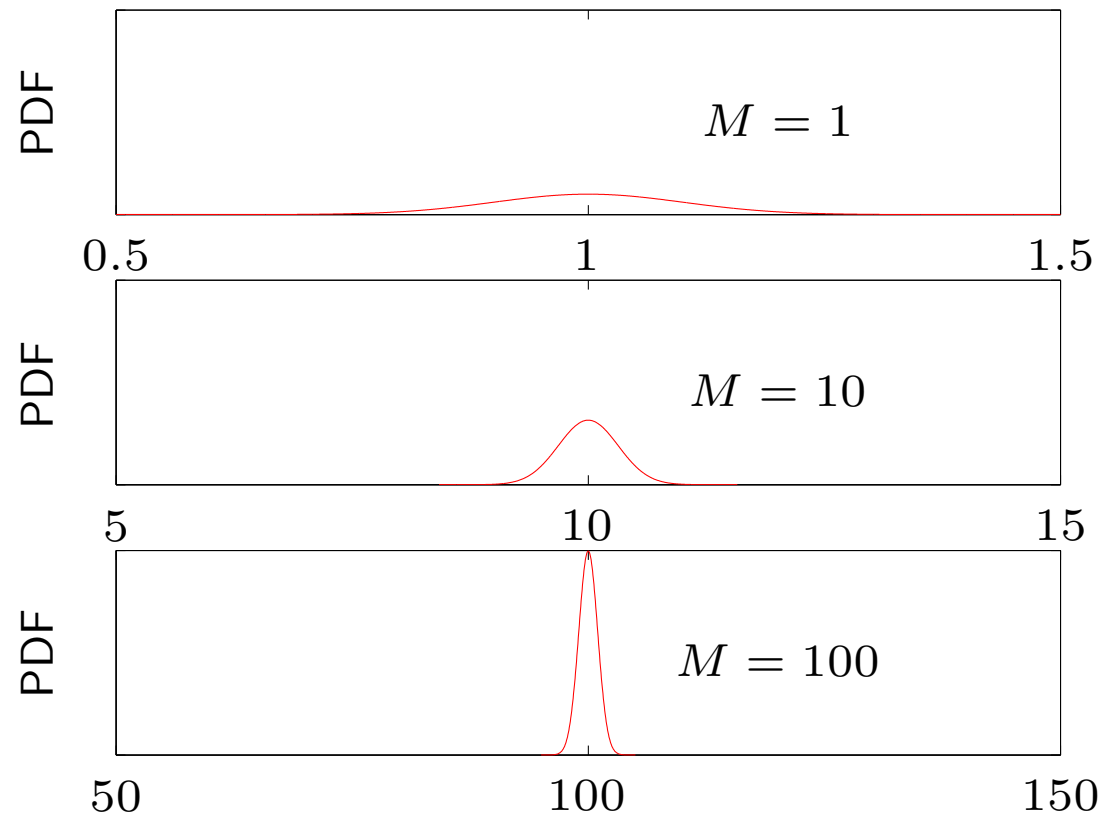
PDF of max of random variables

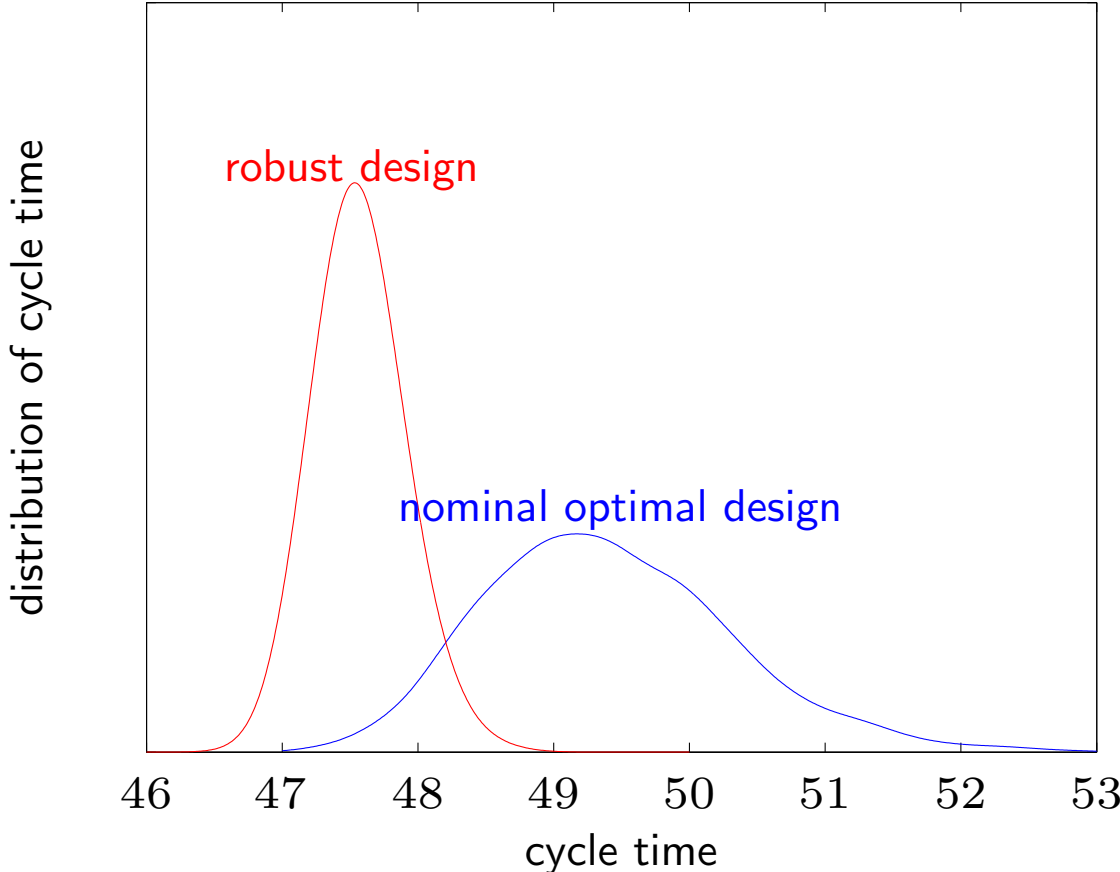$Z = \max\{X_1, \ldots, X_M\}$, $X_i \sim \mathcal{N}(1, 0.1)$ independent

Simple worst-case design

- use slow model for all gates, e.g., $1.2D_i$

- gives **same** design

- can we do better?

Statistically robust design via new method

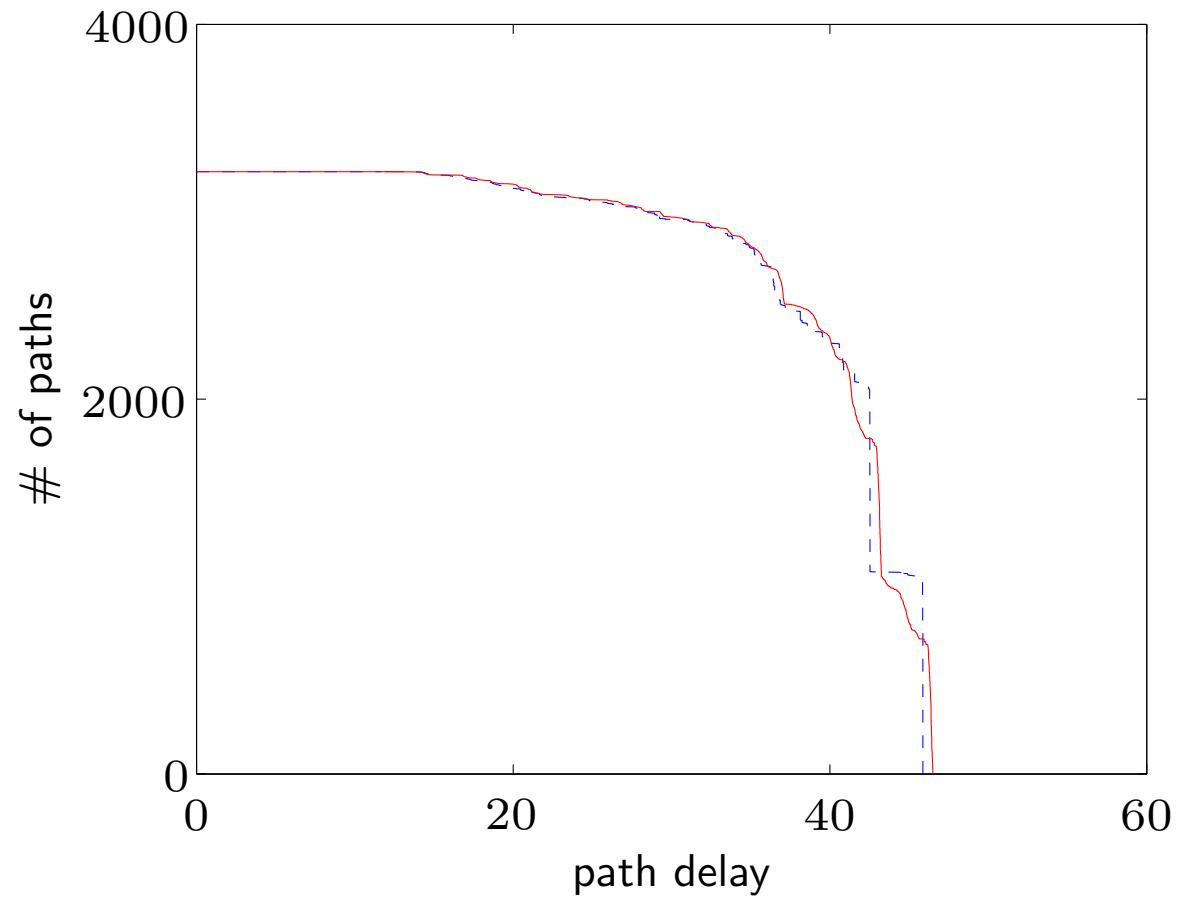same circuit, uncertainty model, and constraints

Statistically robust design via new method

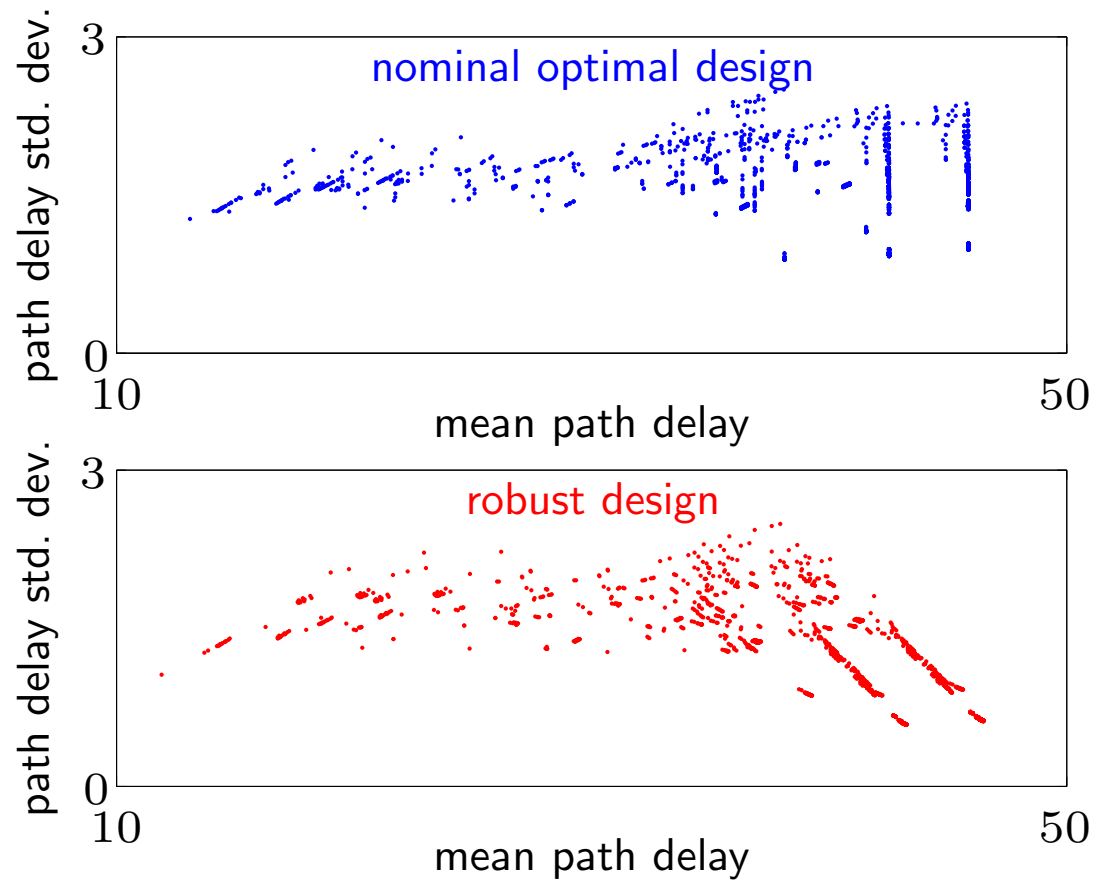|	nominal delay	$\mathbf{E}\,\mathbf{D}$	$\sigma_{\mathbf{D}}$	$\mathbf{Q}^{.95}(\mathbf{D})$
nominal optimal	45.9	49.4	0.91	51.1
robust	46.5	47.6	0.29	48.1

- same circuit, uncertainty model, and constraints

- compared to nominal optimal design, some gates are upsized, others are downsized
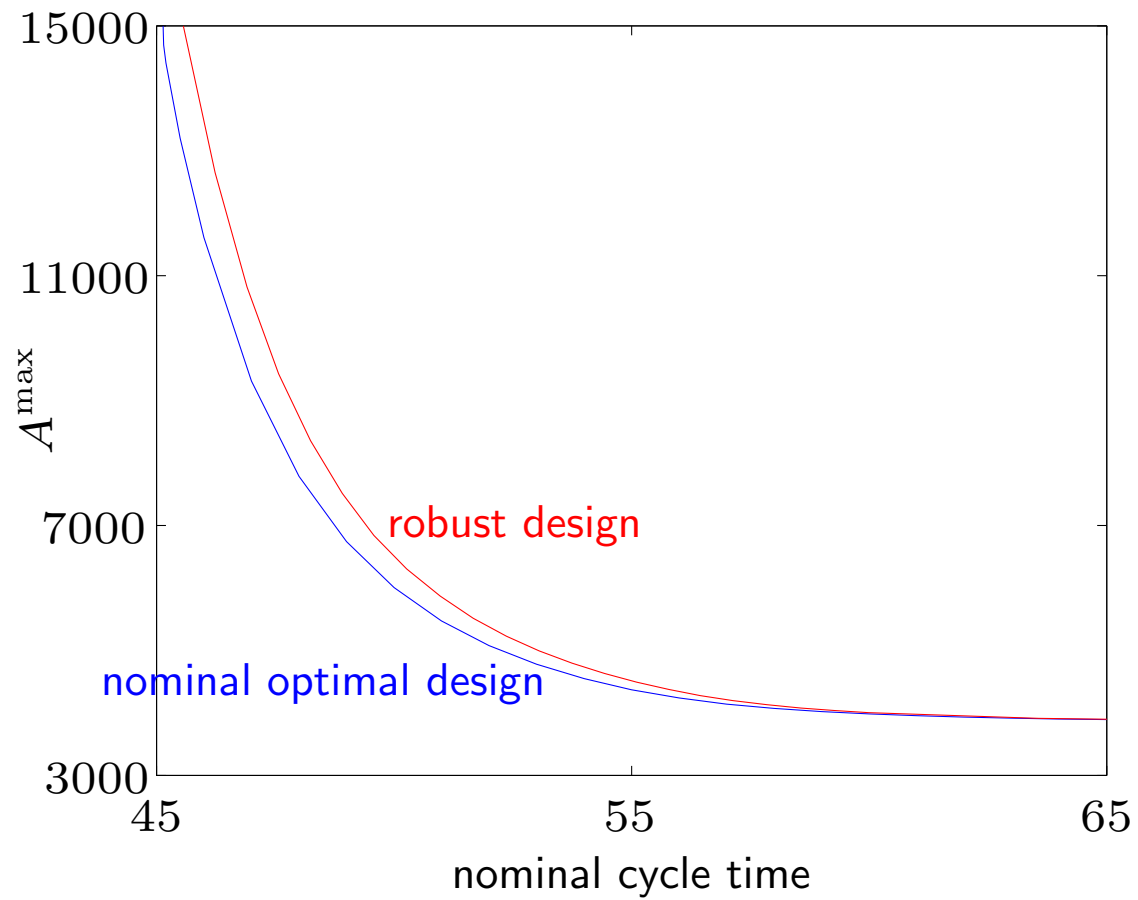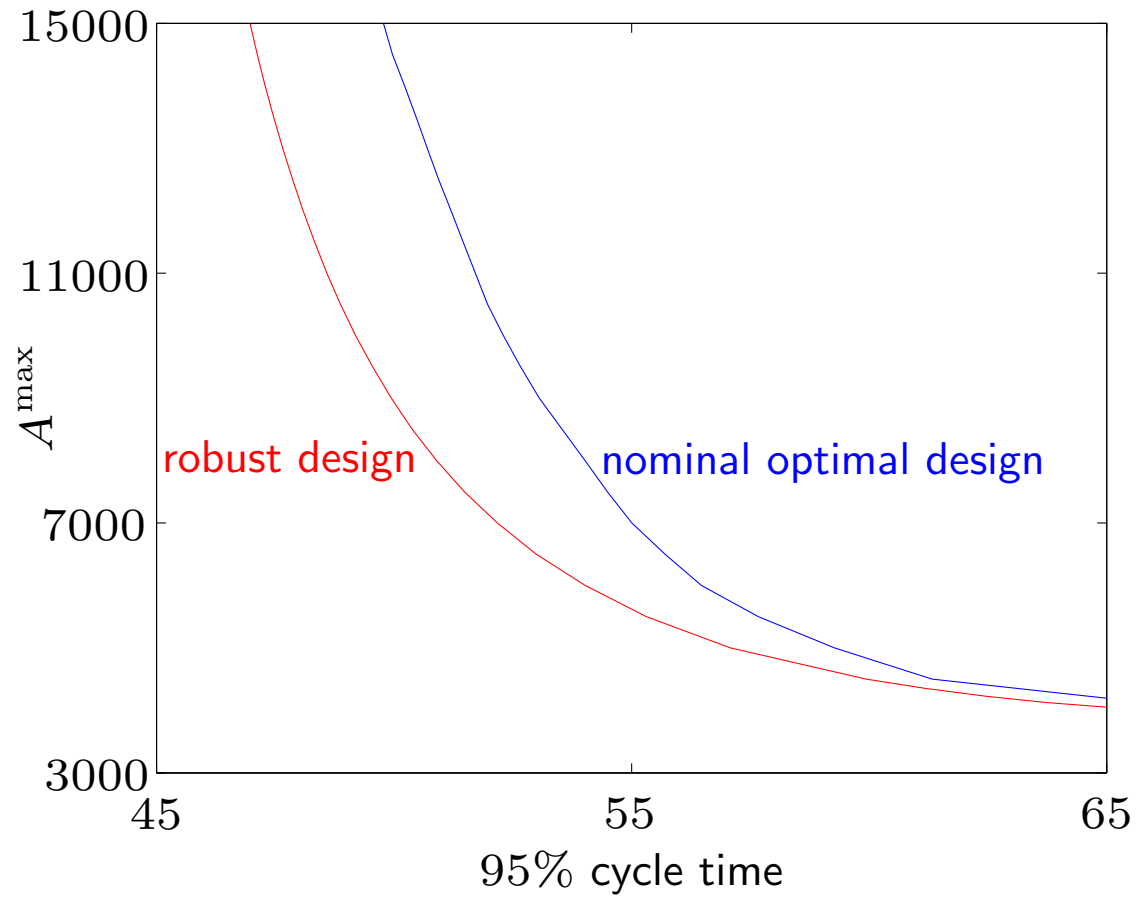
Nominal vs. statistical robust designs

Path delay mean/std. dev. scatter plots

Area/delay trade-off analysis

Area/delay trade-off analysis

Digital Circuit Sizing: Models

Gate scaling

input flip flops   combinational logic block   output flip flops
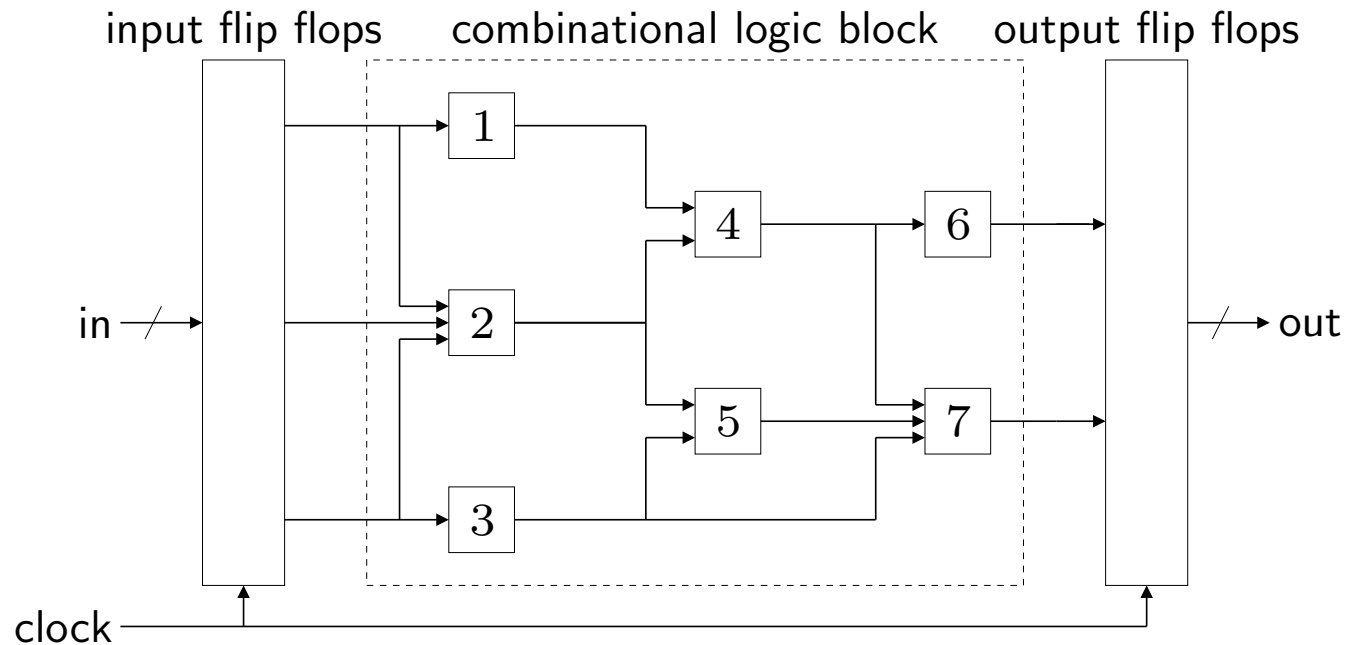
- combinational logic; circuit topology & gate types given

- gate sizes (scale factors $x_i \geq 1$) to be determined

- scale factors affect total circuit area, power and delay

RC gate delay model
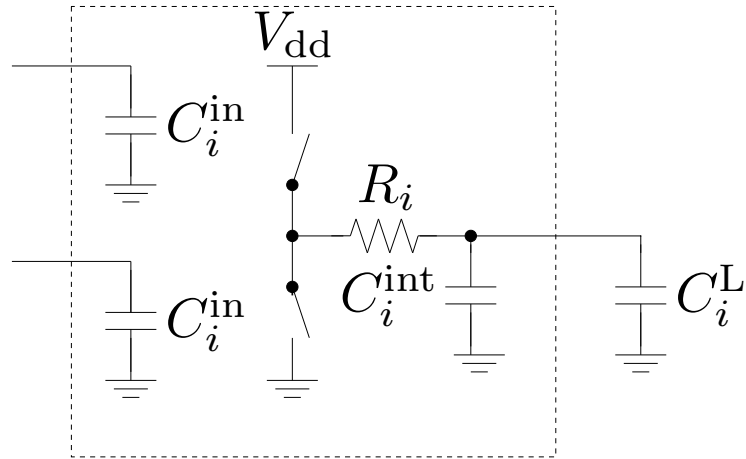
- input & intrinsic capacitances, driving resistance, load capacitance

$$C_i^{\mathrm{in}} = \bar{C}_i^{\mathrm{in}} x_i, \qquad C_i^{\mathrm{int}} = \bar{C}_i^{\mathrm{int}} x_i, \qquad R_i = \bar{R}_i / x_i, \qquad C_i^{\mathrm{L}} = \sum_{j \in \mathrm{FO}(i)} C_j^{\mathrm{in}}$$

- RC gate delay:

$$D_i = 0.69 R_i (C_i^{\mathrm{L}} + C_i^{\mathrm{int}})$$

Path and circuit delay
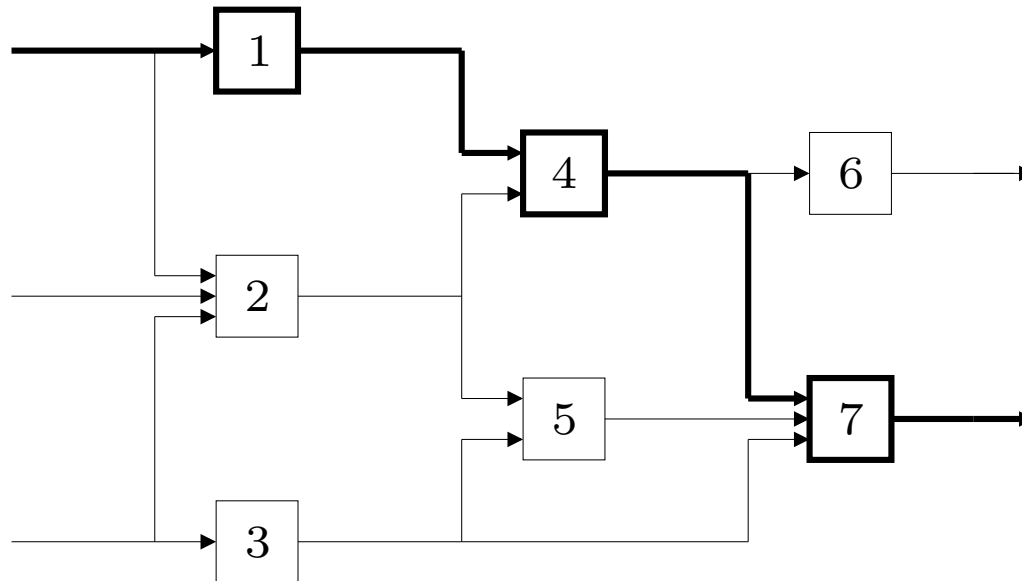
- delay of a path: sum of delays of gates on path

- circuit delay (cycle time): maximum delay over all paths

Area & power

- total circuit area: $A = x_1 \bar{A}_1 + \cdots + x_n \bar{A}_n$

- total power is $P = P_{\mathrm{dyn}} + P_{\mathrm{stat}}$

  - dynamic power $P_{\mathrm{dyn}} = \displaystyle\sum_{i=1}^{n} f_i (C_i^{\mathrm{L}} + C_i^{\mathrm{int}}) V_{\mathrm{dd}}^2$

    f_i is gate switching frequency

  - static (leakage) power $P_{\mathrm{stat}} = \displaystyle\sum_{i=1}^{n} I_i^{\mathrm{leak}} V_{\mathrm{dd}}$

    I_i^{leak} is leakage current (average over input states)

Parameters used in example

- model parameters:

gate type	$\bar{C}^{\text{in}}$	$\bar{C}^{\text{int}}$	$\bar{R}$	$\bar{A}$
INV	3	3	0.48	3
NAND2	4	6	0.48	8
NOR2	5	6	0.48	10
AOI21	6	7	0.48	17
OAI21	6	7	0.48	16

- time unit is τ, delay of min-size inverter $(0.69 \cdot 0.48 \cdot 3 = 1)$

- area (total width) unit is width of NMOS in min-size inverter

Statistical variation in threshold voltage

- we focus on statistical variation in threshold voltage V_{th}
  (can also model variations in other parameters, $e.g.$, t_{ox}, L_{eff}, ...)

- *Pelgrom model*:
$$\sigma_{V_{\mathrm{th}}} = \bar{\sigma}_{V_{\mathrm{th}}} x^{-1/2}$$
  where $\overline{\sigma}^2_{V_{\mathrm{th}}}$ is V_{th} variance for unit scaled gate

- larger gates have less V_{th} variation

Statistical gate delay model

- alpha-power law model:

$$D \propto \frac{V_{\mathrm{dd}}}{(V_{\mathrm{dd}} - V_{\mathrm{th}})^{\alpha}}$$

$(\alpha \approx 1.3)$

- for small variation in V_{th},

$$\sigma_D \approx \left| \frac{\partial D}{\partial V_{\mathrm{th}}} \right| \sigma_{V_{\mathrm{th}}} = \alpha (V_{\mathrm{dd}} - V_{\mathrm{th}})^{-1} \bar{\sigma}_{V_{\mathrm{th}}} x^{-0.5} D$$

- gate scaling affects mean delay and relative variation differently

- relative variation decreases as gate scale factor increases:
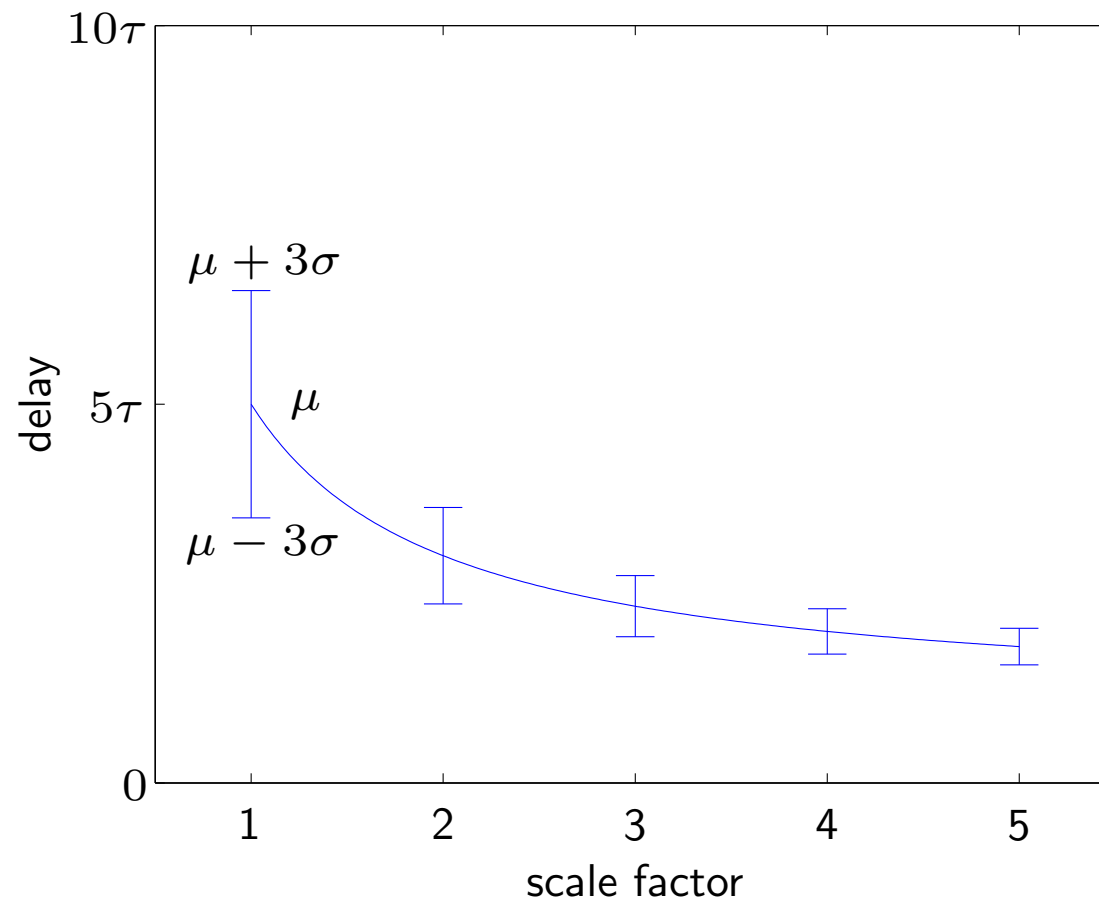
$$\sigma_D / D \propto x^{-0.5}$$

Statistical variation in gate delay

10% relative variation for min sized gate ($\sigma_D/D = 0.1$)
inverter driving $C_{\mathrm{L}} = 4$
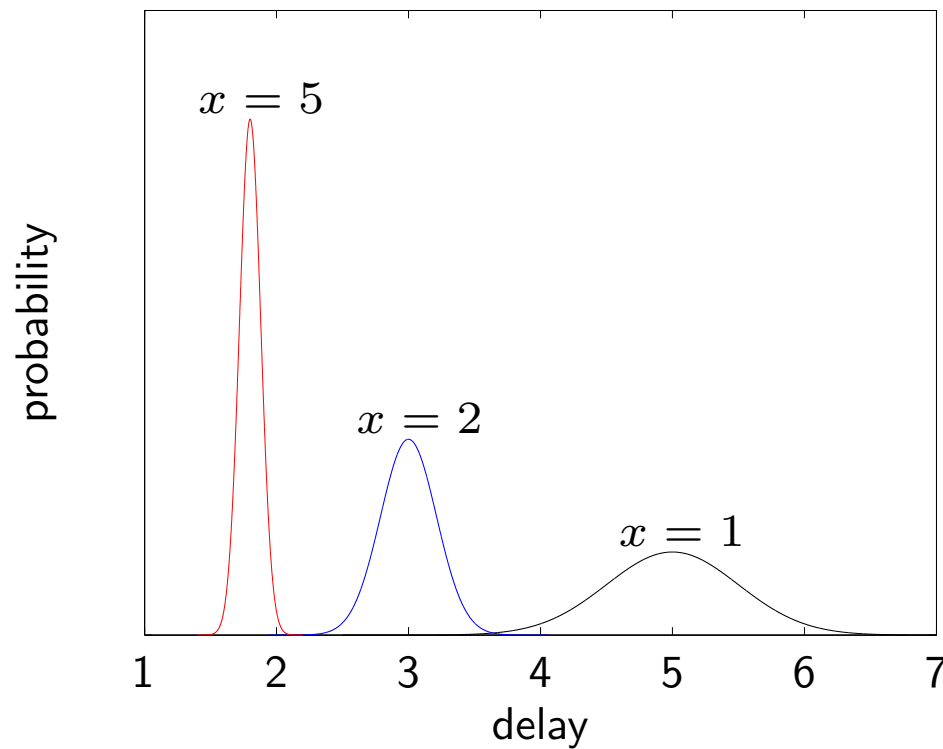
Statistical variation in gate delay

inverter driving $C_{\mathrm{L}} = 4$

Statistical leakage power model

- leakage current

$$I^{\text{leak}} \propto x e^{-V_{\text{th}}/V_0}$$
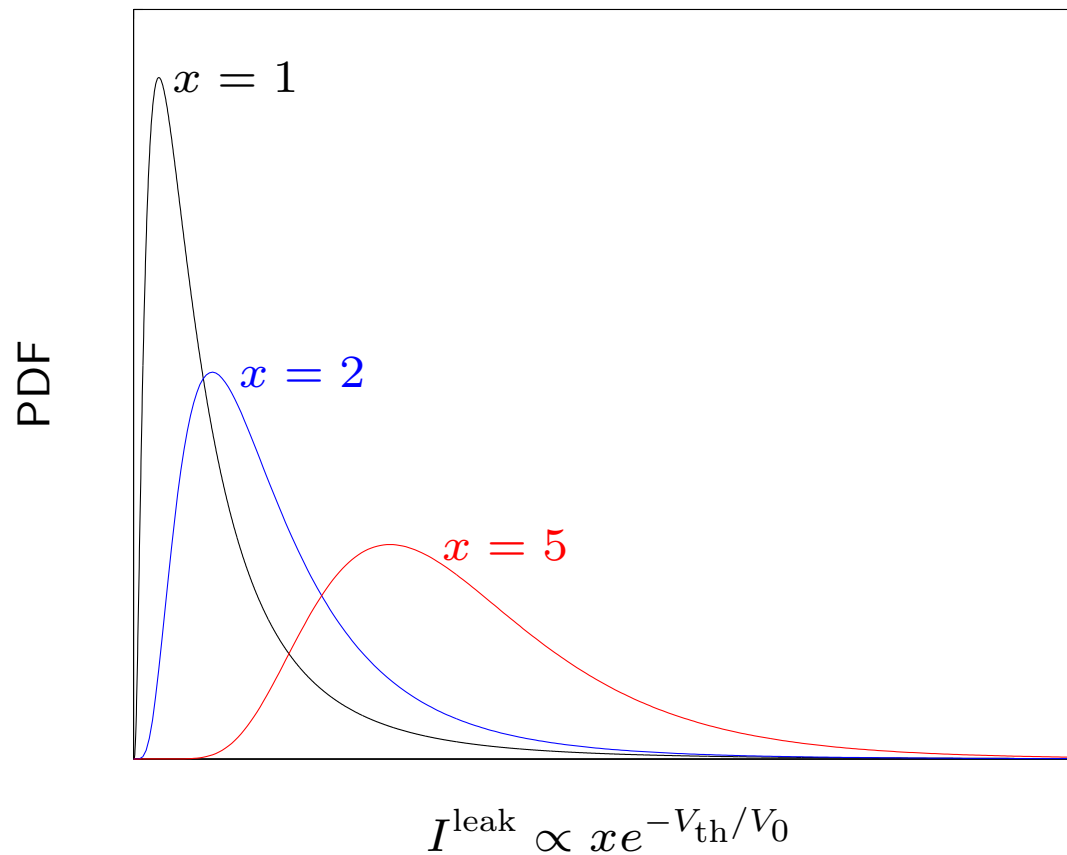
$(V_0 \approx 0.04)$

- linearization does not give accurate prediction of $\mathbf{E}\, I^{\text{leak}}$, $\sigma_{I^{\text{leak}}}$

- exact values for V_{th} Gaussian:

$$\mathbf{E}\, I^{\text{leak}} = I^{\text{leak,nom}} e^{\overline{\sigma}^2_{V_{\text{th}}}/(2V_0^2 x)}, \qquad \sigma_{I^{\text{leak}}} = \left(e^{\overline{\sigma}^2_{V_{\text{th}}}/(V_0^2 x)} - 1 \right)^{1/2} \mathbf{E}\, I^{\text{leak}}$$
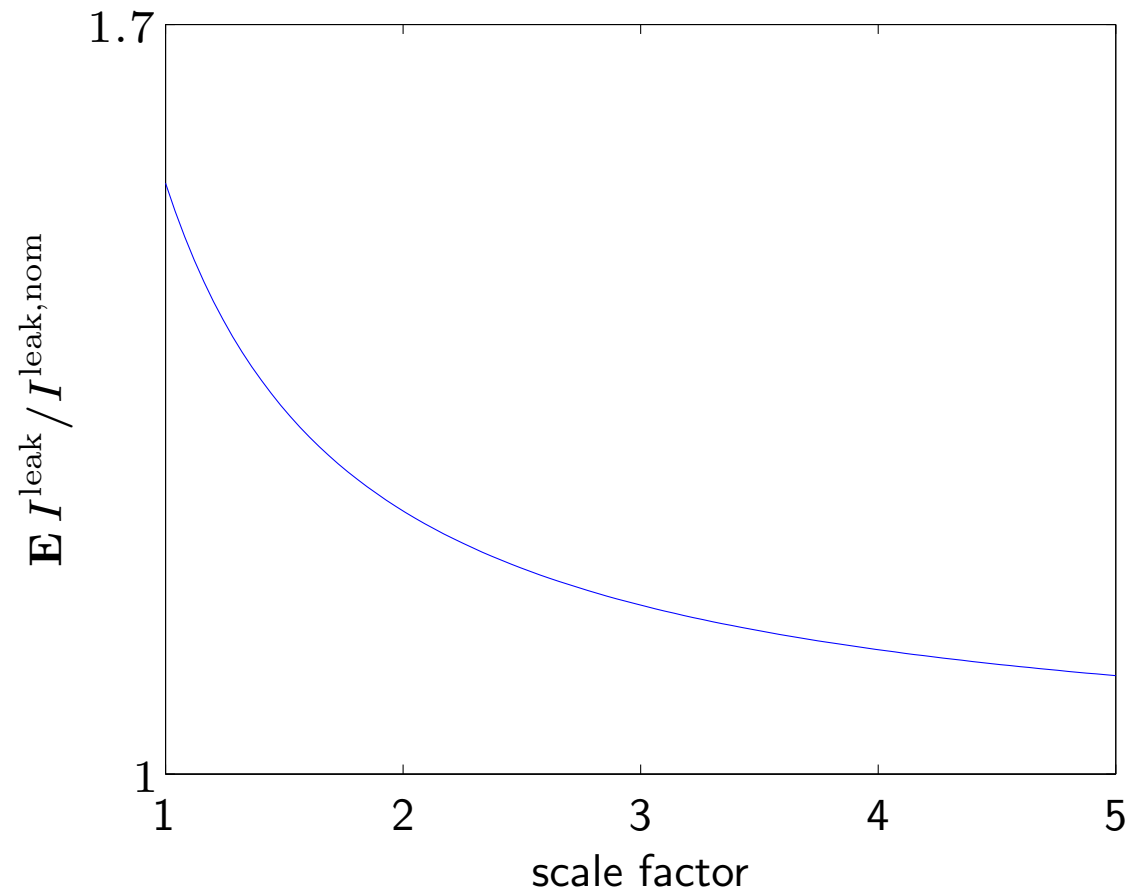
$I^{\text{leak,nom}}$ is leakage current when statistical variation is ignored

Effects of statistical variation on leakage power

$$V_{\mathrm{th}} \sim \mathcal{N}(\bar{V}_{\mathrm{th}}, 0.15\bar{V}_{\mathrm{th}}),\ \bar{V}_{\mathrm{th}} = 0.25,\ V_0 = 0.04$$

$$I^{\mathrm{leak}} \propto x e^{-V_{\mathrm{th}}/V_0}$$

Statistical variation in leakage power

Digital Circuit Sizing: Optimization

Basic gate scaling problem (no statistical variation)

$$\begin{aligned} \text{minimize} \quad & D \\ \text{subject to} \quad & P \leq P^{\max}, \quad A \leq A^{\max} \\ & 1 \leq x_i, \quad i = 1, \ldots, n \end{aligned}$$

a **geometric program** (GP); can be solved efficiently

extensions/variations:

- minimize area, power, or some combination

- maximize clock frequency subject to area, power limits

- add other constraints

- optimal trade-off of area, power, delay

Statistical parameter variation

- now model gate delay & power as **random variables**

- circuit performance measures P, D become random variables $\mathbf{P}$, $\mathbf{D}$

- **distributions** of $\mathbf{P}$, $\mathbf{D}$ depend on gate scalings x_i

- for fixed design, can estimate PDFs of $\mathbf{P}$, $\mathbf{D}$ via Monte Carlo
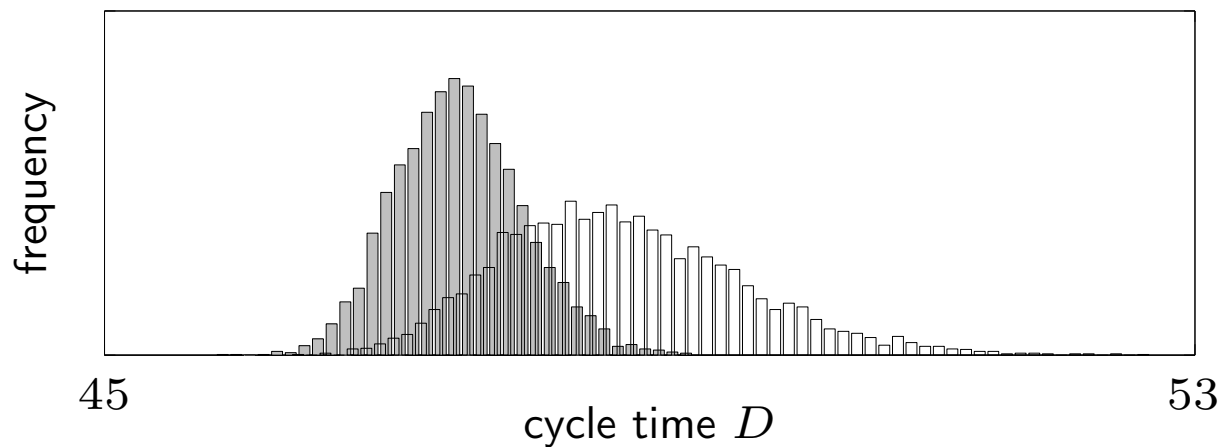
Statistical design

- measure random performance measures by 95% quantile (say)

$$\begin{aligned}
\text{minimize} \quad & \mathbf{Q}^{.95}(\mathbf{D}) \\
\text{subject to} \quad & \mathbf{Q}^{.95}(\mathbf{P}) \leq P^{\max}, \qquad A \leq A^{\max} \\
& 1 \leq x_i, \quad i = 1, \ldots, n
\end{aligned}$$

- **extremely difficult** stochastic optimization problem; almost no analytic/exact results

- but, simple heuristic method works well

The New Method

Statistical power constraint

- total power is sum of gate powers

$$\mathbf{E}\,\mathbf{P} = \sum_{i=1}^{n} \mathbf{E}\,\mathbf{P}_i$$

- if n is large and $\mathbf{P}_1, \ldots, \mathbf{P}_m$ are independent (enough),

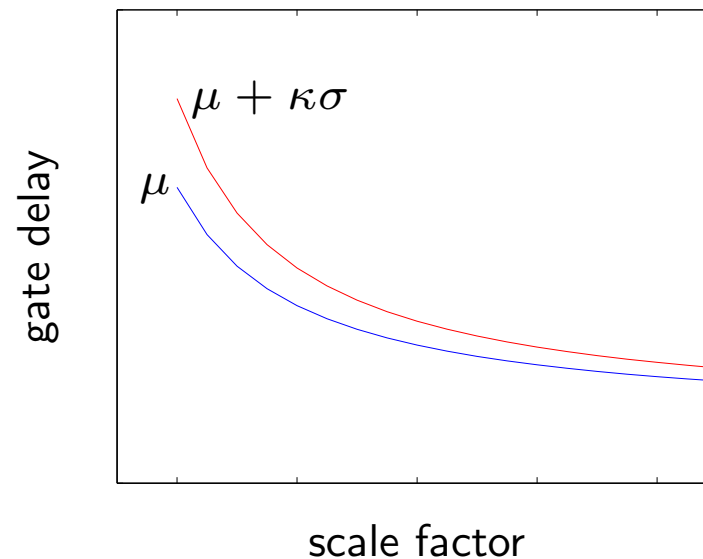$$\mathbf{P} \approx \sum_{i=1}^{n} \mathbf{E}\,\mathbf{P}_i$$

- can use $\mathbf{E}\,\mathbf{P} \leq P^{\mathrm{max}}$ as reasonable approximation of $\mathbf{Q}^{.95}(\mathbf{P}) \leq P^{\mathrm{max}}$

Surrogate gate delay

- define **surrogate gate delays**

$$\tilde{D}_i(x) = D_i(x) + \kappa_i \sigma_i(x)$$

$\kappa_i \sigma_i(x)$ is **margin** on gate delay (κ_i is typically 2)

- gives more margin to smaller gates

Interpretation of gate delay margins

- margins $\kappa_i \sigma_i(x)$ take statistical gate delay variation into account

- κ_i related to $\mathbf{Prob}\left(\mathbf{D}_i \leq \mu_i + \kappa_i \sigma_i\right)$

  – Chebyshev inequality:

$$\mathbf{Prob}\left(\mathbf{D}_i \leq \mu_i + \kappa_i \sigma_i\right) \geq \frac{\kappa_i^2}{1 + \kappa_i^2}$$

  – if $\mathbf{D}_i$ is Gaussian

$$\mathbf{Prob}\left(\mathbf{D}_i \leq \mu_i + \kappa_i \sigma_i\right) = \frac{1}{\sqrt{2\pi}} \int_{\kappa_i}^{\infty} e^{-t^2/2} \, dt$$

Heuristic for statistical design

- use modified (leakage) power model taking into account statistical variation

- use **surrogate gate delays** $\tilde{D}_i(x) = D_i(x) + \kappa_i \sigma_i(x)$

- now solve resulting (deterministic) gate scaling problem

- verify statistical performance via Monte Carlo analysis
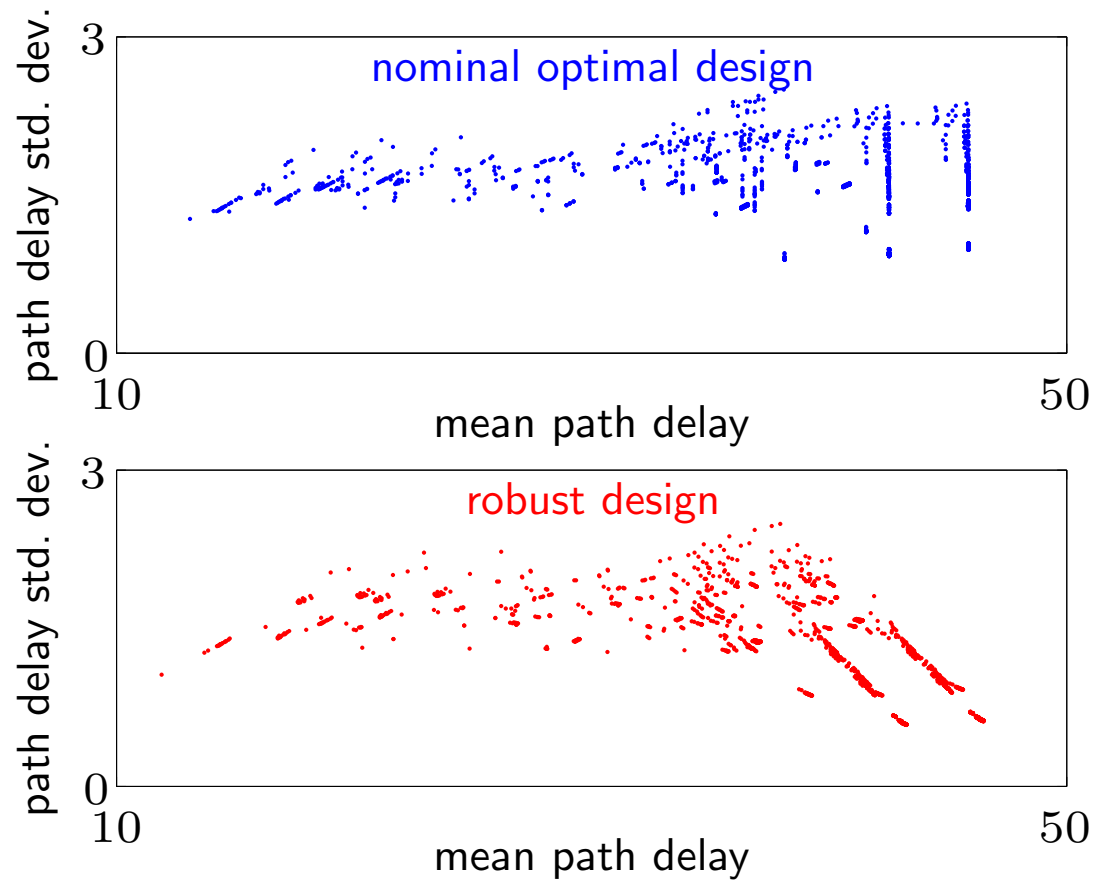
(can update κ_i's and repeat)

Digital Circuit Sizing Example
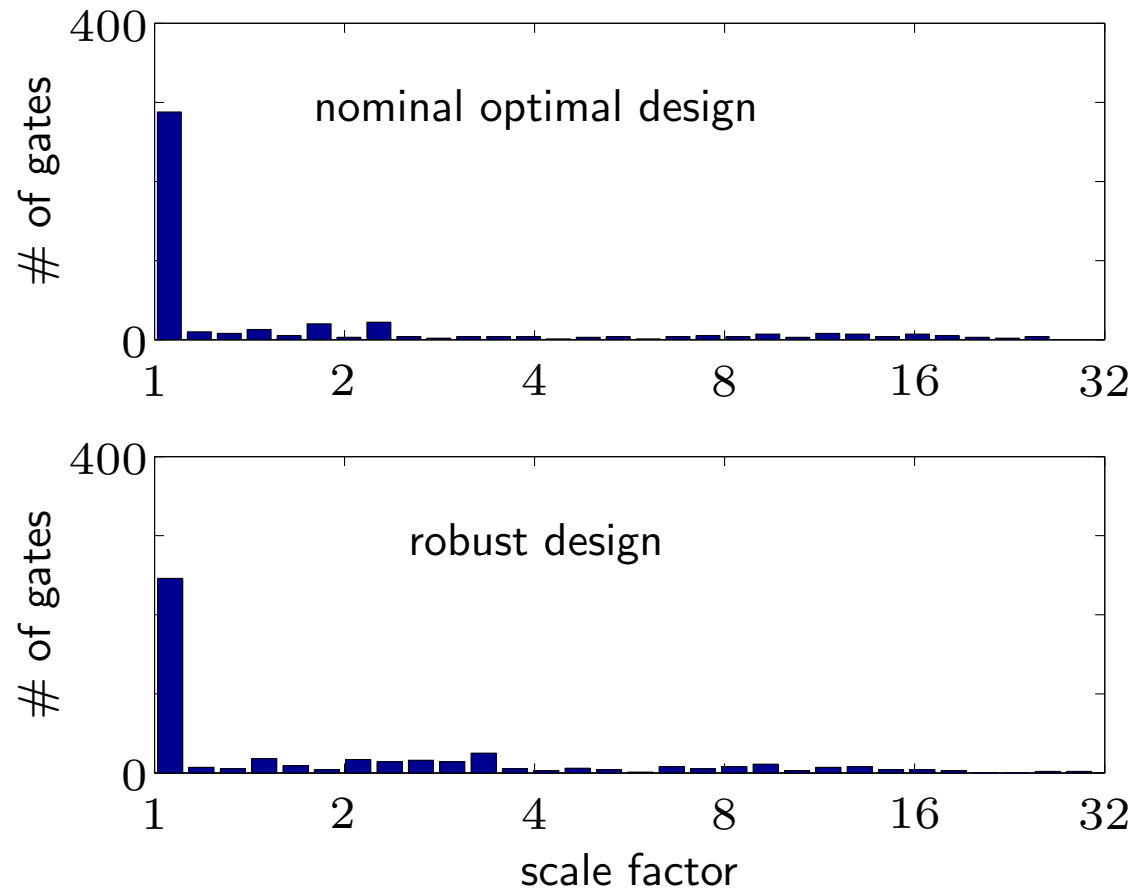
Statistically robust design via new method
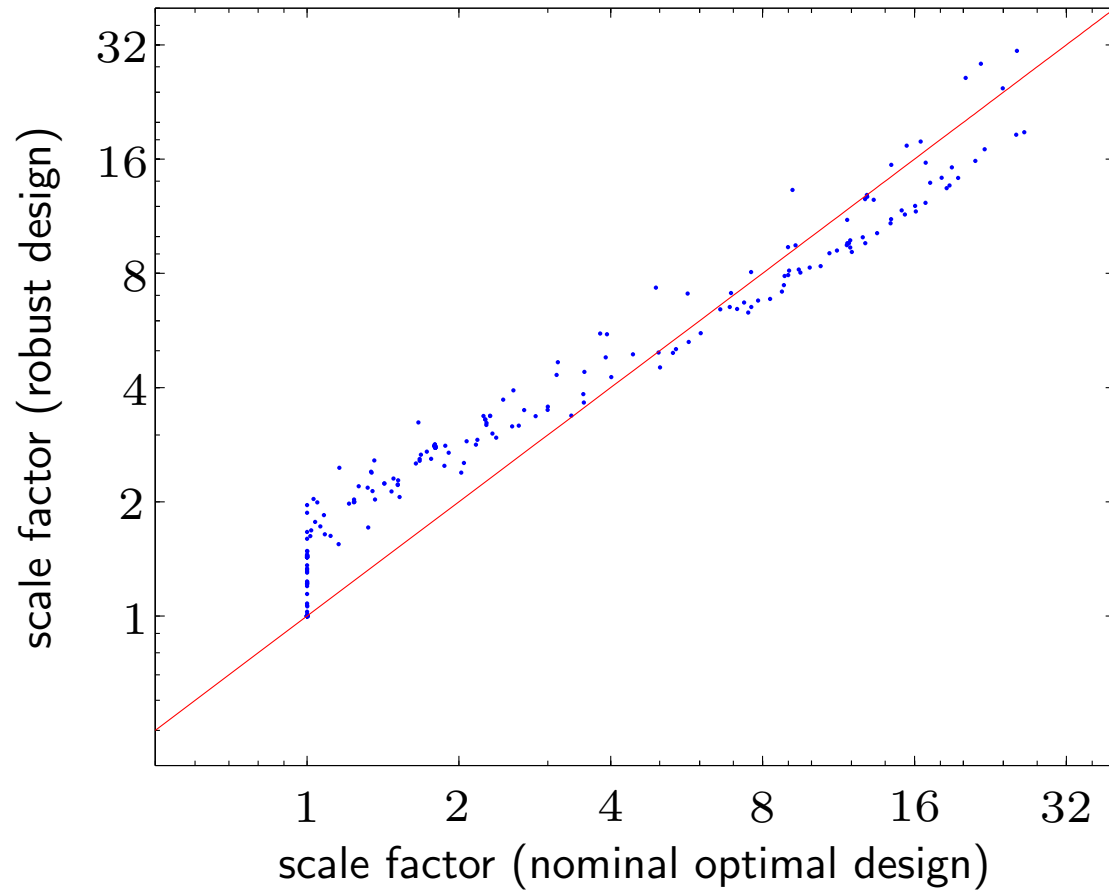
same circuit, uncertainty model, and constraints
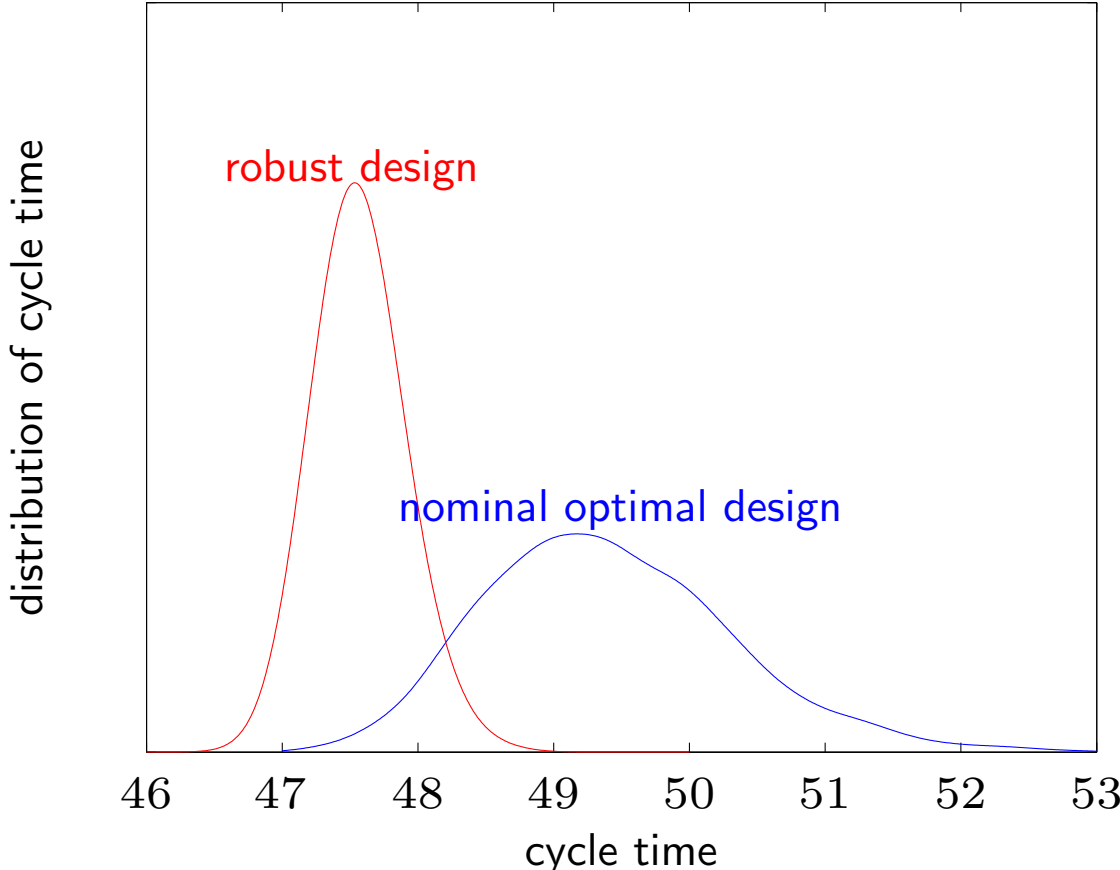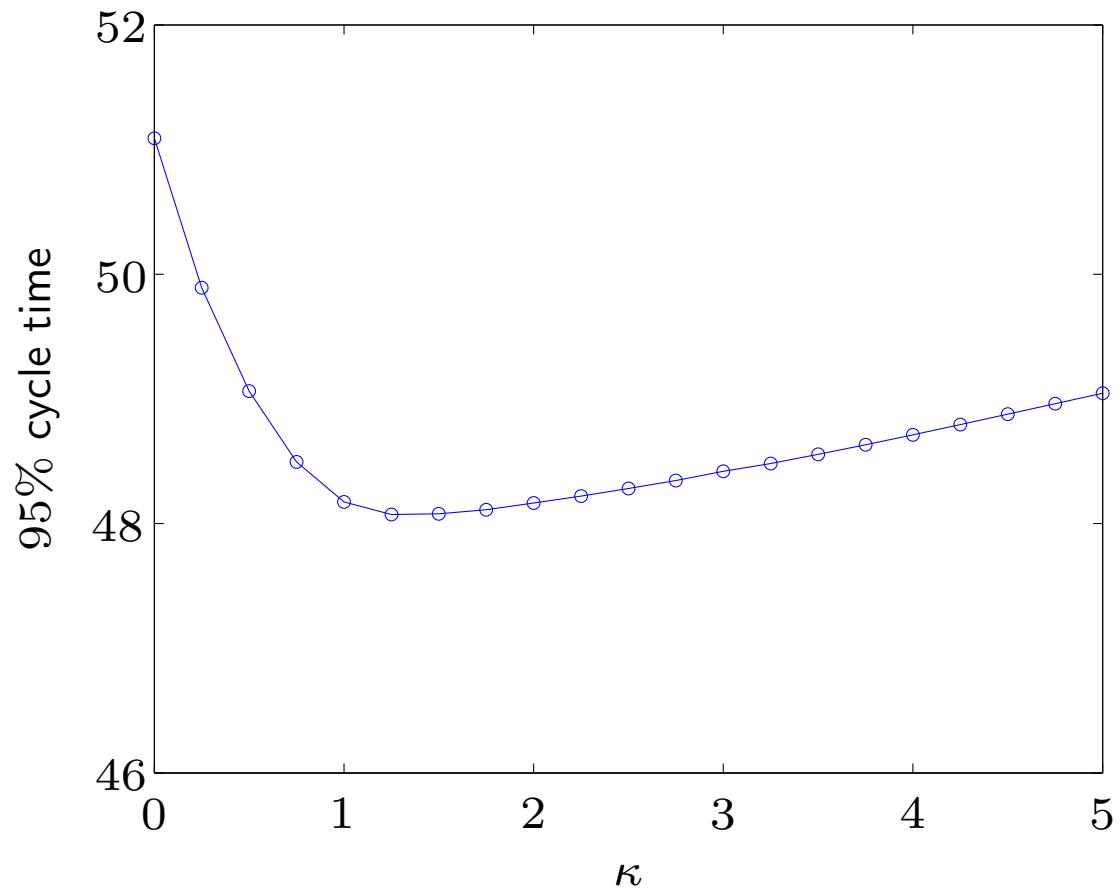
Path delay mean/std. dev. scatter plots

Comparison of nominal optimal and robust designs

Comparison of nominal optimal and robust designs

Effect of margin coefficients
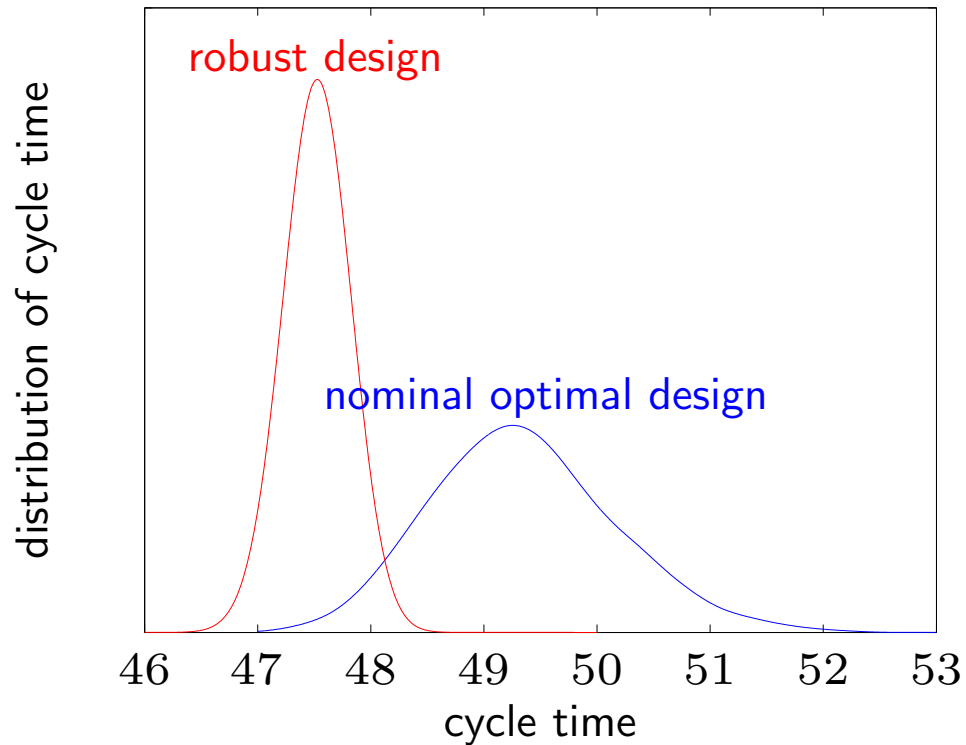
Sensitivity to model assumptions

question: how sensitive is robust design to our model of process variation?

- distribution shape

- correlation between gates

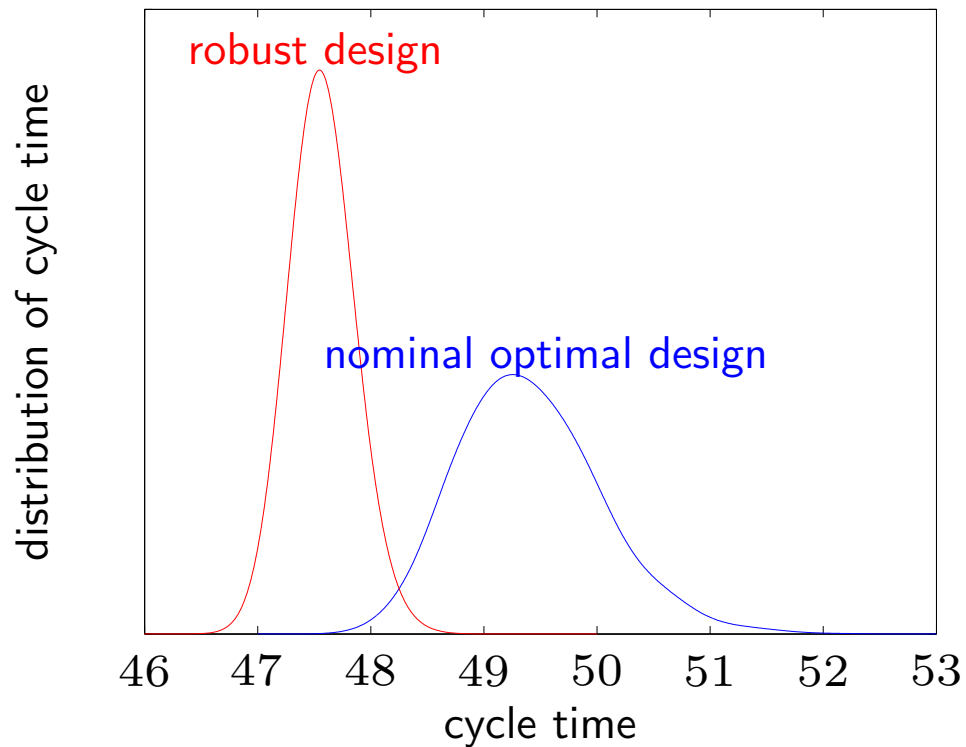- Pelgrom model of variance vs. scale factor

answer: not very

Simulation with uniform gate delay distributions

compared with Gaussian gate delays:
nominal optimal design not quite as bad; robust design still quite good

Simulation with correlated gate delays

connected gates have delays that are 30% correlated

nominal optimal not as bad; but robust design still quite good

Conclusions and Future Work

Conclusions

- statistically robust design is subtle; cannot be done by hand

- exact or direct methods will not work well

  - computationally intractable
  - depend on details of statistical models

- heuristic method is relatively simple, scales well, gives good designs

  - reduces problem to a deterministic one

References

- Boyd, Kim, and Mohan, *DATE Tutorial 2005*
  Geometric programming and its applications to EDA problems

- Boyd, Kim, Patil, and Horowitz, *SPIE ML 2006*
  A heuristic method for statistical digital circuit sizing

- Kim, Boyd, Patil, and Horowitz, *Optimization and Engineering*, 2006
  A heuristic for optimizing stochastic activity networks with applications
  to statistical digital circuit sizing

- Boyd, Kim, Patil, and Horowitz, *Operations Research*, 2005
  Digital circuit optimization via geometric programming

- Patil, Yun, Kim, Cheung, Horowitz, and Boyd, *ISQED 2005*
  A new method for design of robust digital circuits

all available from `www.stanford.edu/~boyd/research.html`

References (continued)

- Mani, Devgan, Orshansky, *DAC 2005*
  An efficient algorithm for statistical minimization of total power under timing yield constraints

- Satish, Ravindran, Moskewicz, Chinnery, and Keutzer, *UCB tech. report, 2005*
  Evaluating the effectiveness of statistical gate sizing for power optimization

- Bhardwaj and Vrudhula, *DAC 2005*
  Leakage minimization of nano-scale circuits in the presence of systematic and random variations