The good, the bad and the statistical

Noel Menezes

Strategic CAD Labs
Design and Technology Solutions
Intel Corp.

Acknowledgements

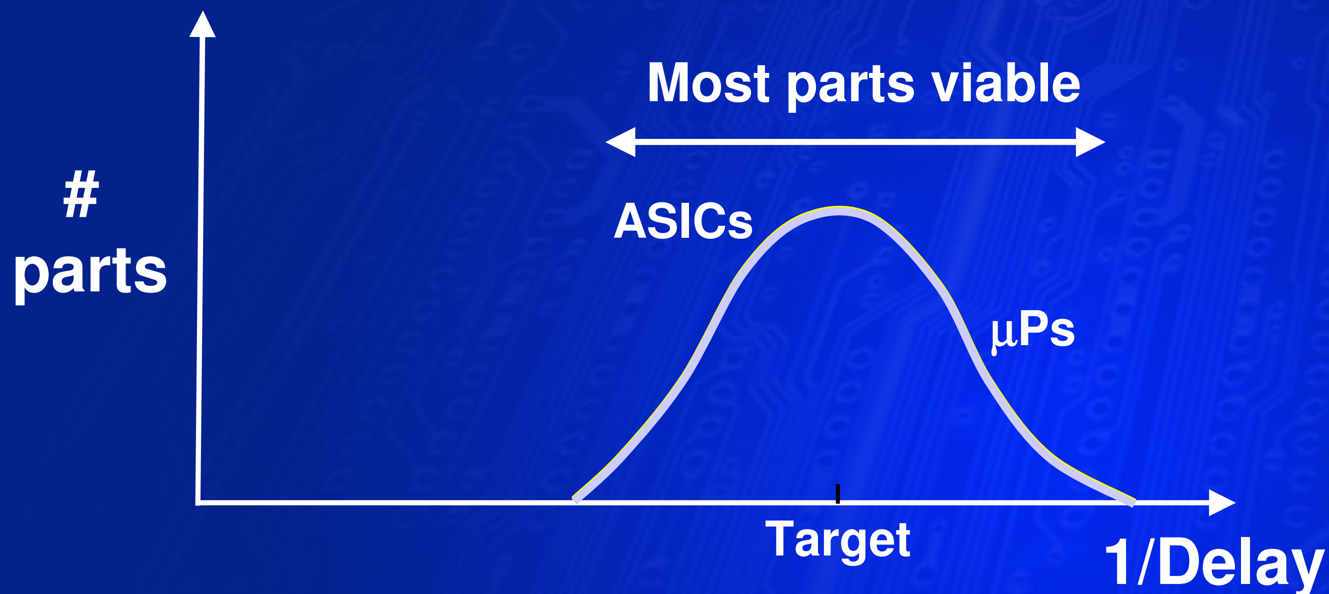
- Keith Bowman
- Yossi Abulafia
- Steve Burns
- Mahesh Ketkar
- Vivek De
- Jim Tschanz
- Chandra Kashyap
- Chirayu Amin
- Nagib Hakim

There are lies, damned lies, and statistics.

-- Benjamin Disraeli

intel Leap ahead™

Variable vs. fixed frequency products



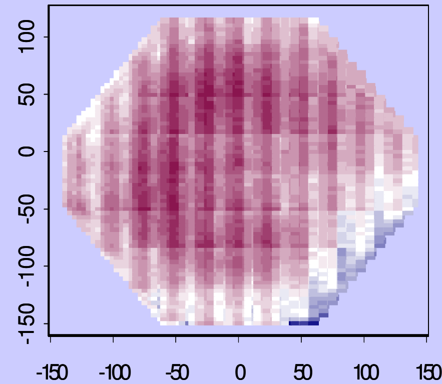
- Microprocessors can be binned
  - At-speed test is not economically feasible for most ASICs

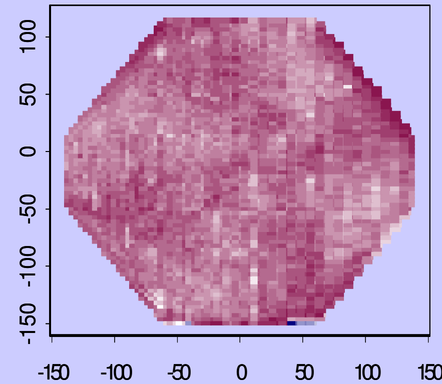
Outline

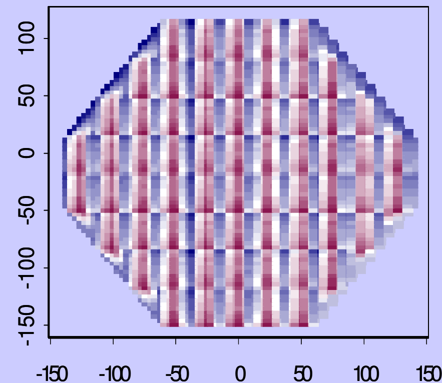
- Statistical CD variation basics
  - Modeling
  - Path distributions
- The hope of SSTA
- Statistical optimization
- Statistical techniques in design
  - Skew computation
  - Min-delay analysis
  - Bin-split prediction
  - FMAX-ISB estimation
- Conclusions

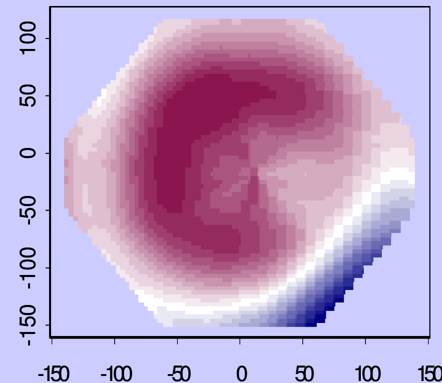
Classification

- Electrical behavior
  - Le
  - VT
  - Width
  - Interconnect
- Spatial behavior
  - Wafer to wafer
  - Die to die
  - Within die





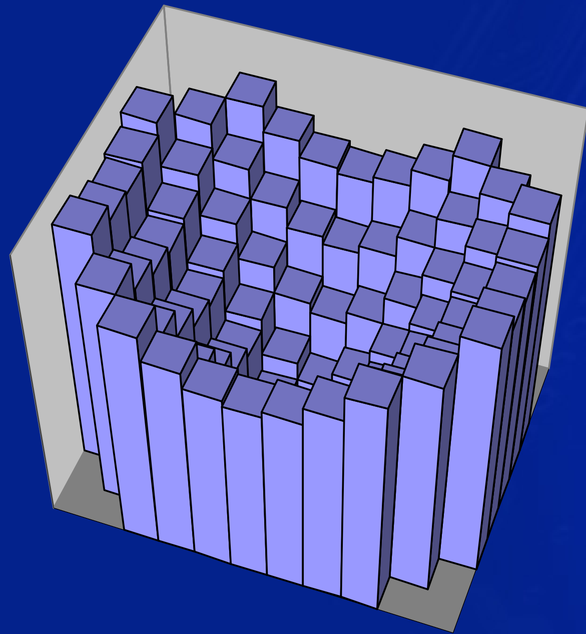




Source: N. Hakim, ICCAD 2004

Scale of Variations

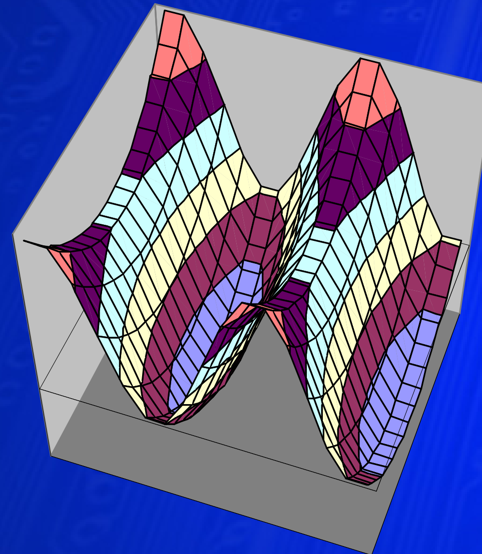
Die-to-Die (D2D) Variations



Wafer Scale

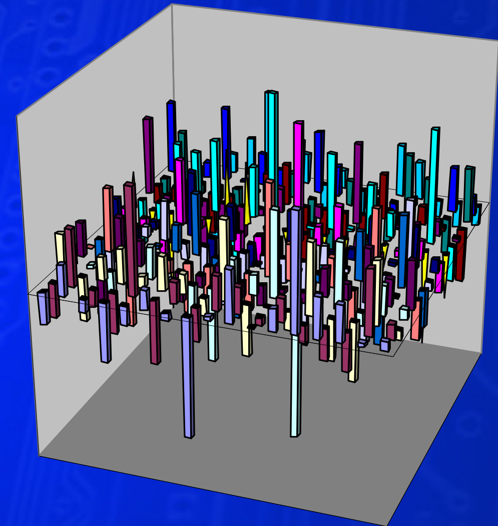
Within-Die (WID) Variations

Systematic



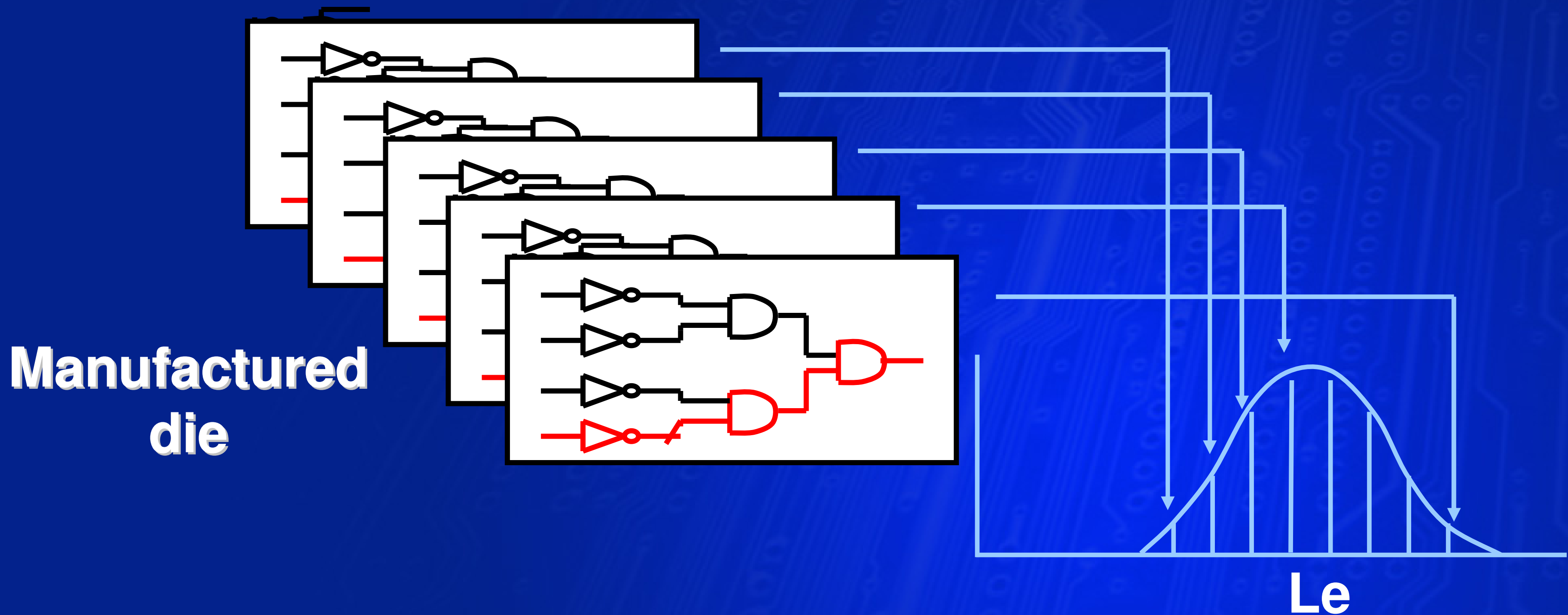
Die Scale

(Uncorrelated) Random



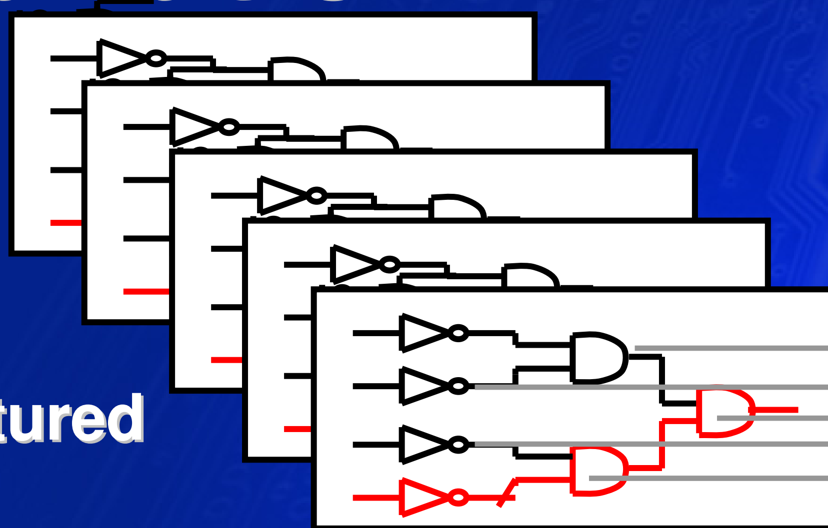
Feature Scale

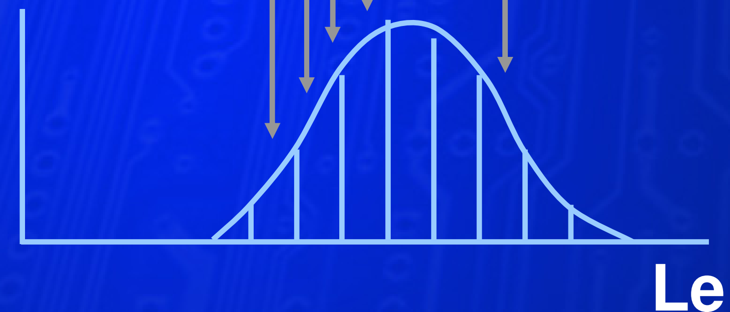
D2D variation



- D2D variation effects are primarily addressed by process engineers

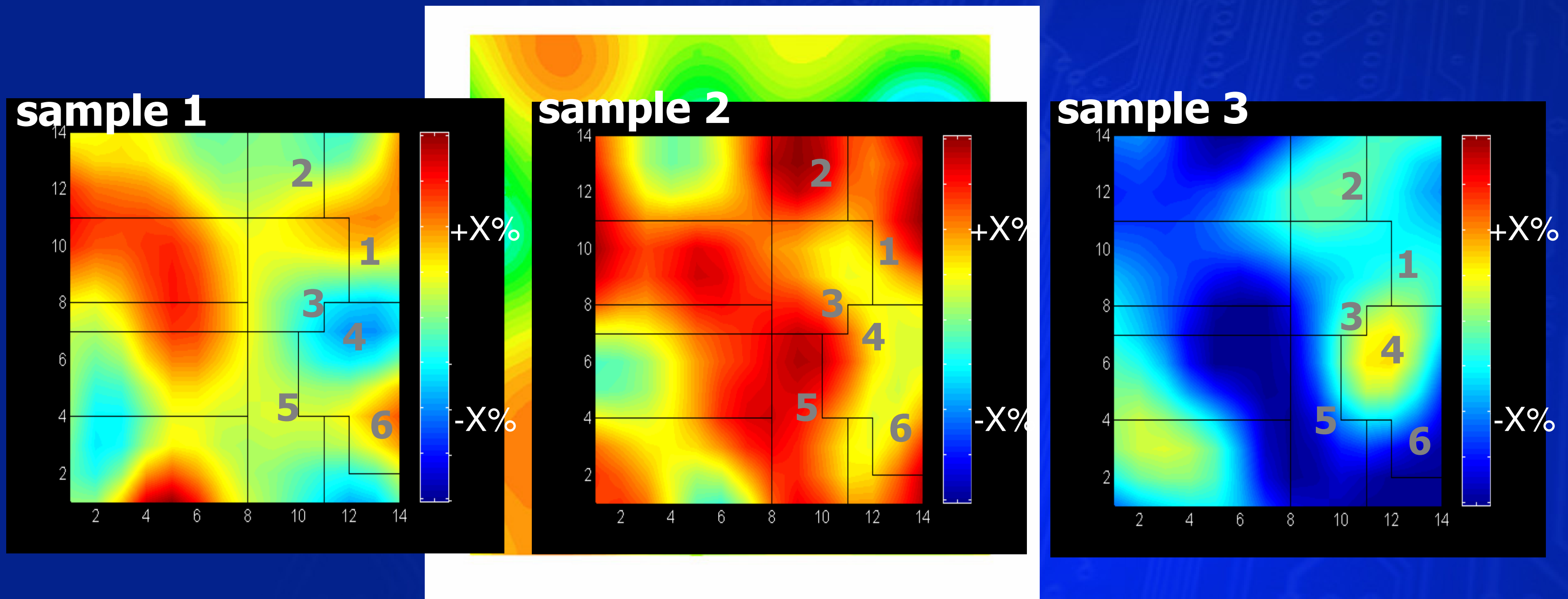
Within-die (WID) variation





- Two WID components
  - Purely random
  - Correlated random -- systematic

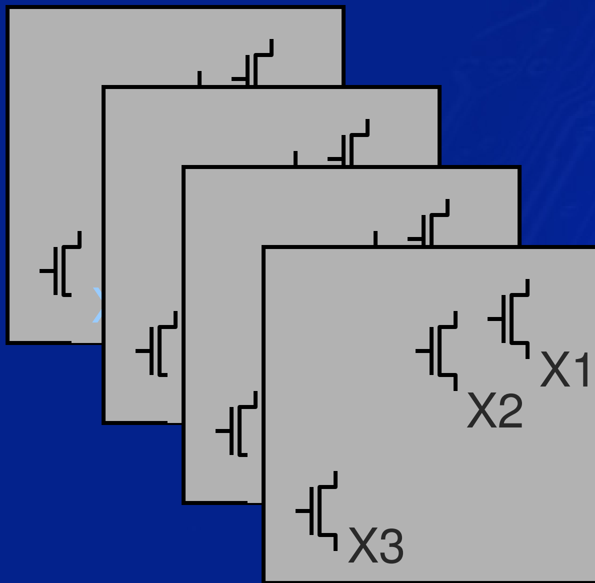
Systematic (correlated random) WID variation

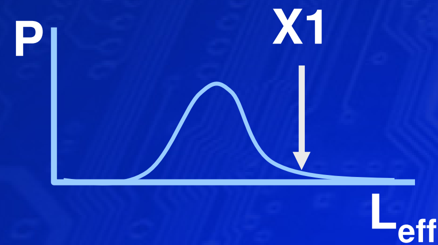


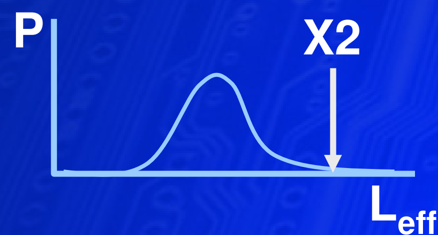
- Models distance-dependent smooth variations
  - Exact shape is unknown

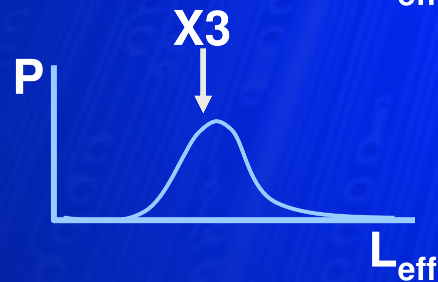
S. Samaan, ICCAD 04

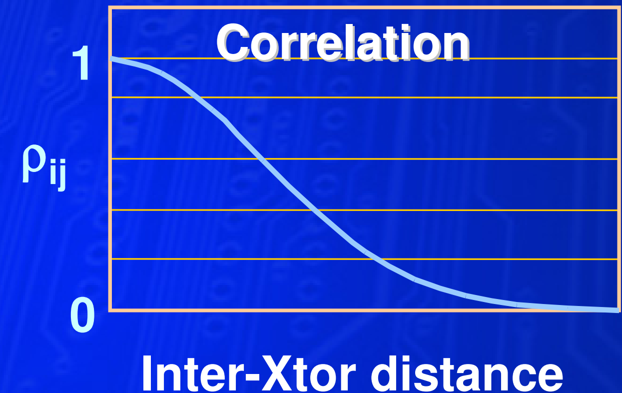
Nature of correlated variation











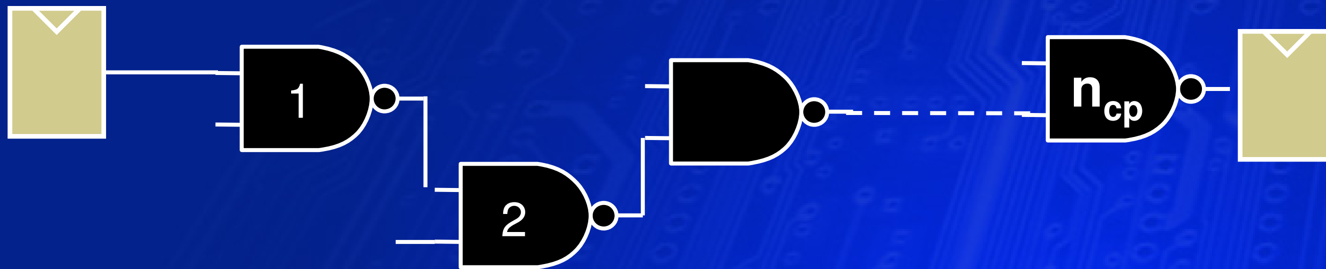
- CDs of transistors that are close track
  - Tracking diminishes with distance

Variations: Trend & Impact on Performance & Power

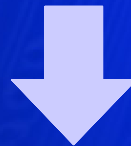
Variation	Delay Impact	Power Impact	Other	Trend
Leff	Large	Large		Flat
Width	Small	Small		Decreasing
Vt	Small	Medium	SRAM	Increasing
Interconnect	Small	Low		Increasing
Other	Variable	Variable		Flat

N. Hakim, ICCAD 2004

Principle I: Path effect of random variations



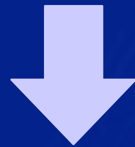
$$\sigma^2_{T_{cp}} = \sigma^2_{T_{nand}} + \sigma^2_{T_{nand}} + \ldots + \sigma^2_{T_{nand}} = n_{cp}\sigma^2_{T_{nand}}$$



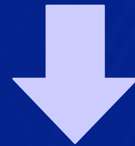
$$\frac{\sigma_{T_{cp}}}{T_{cp}} = \frac{\sqrt{n_{cp}}\sigma_{T_{nand}}}{n_{cp}T_{nand}} = \frac{1}{\sqrt{n_{cp}}}\frac{\sigma_{T_{nand}}}{T_{nand}}$$

Principle II: Delay distributions for independent paths

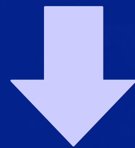
$$t_{chip} = max(t_1, t_2)$$



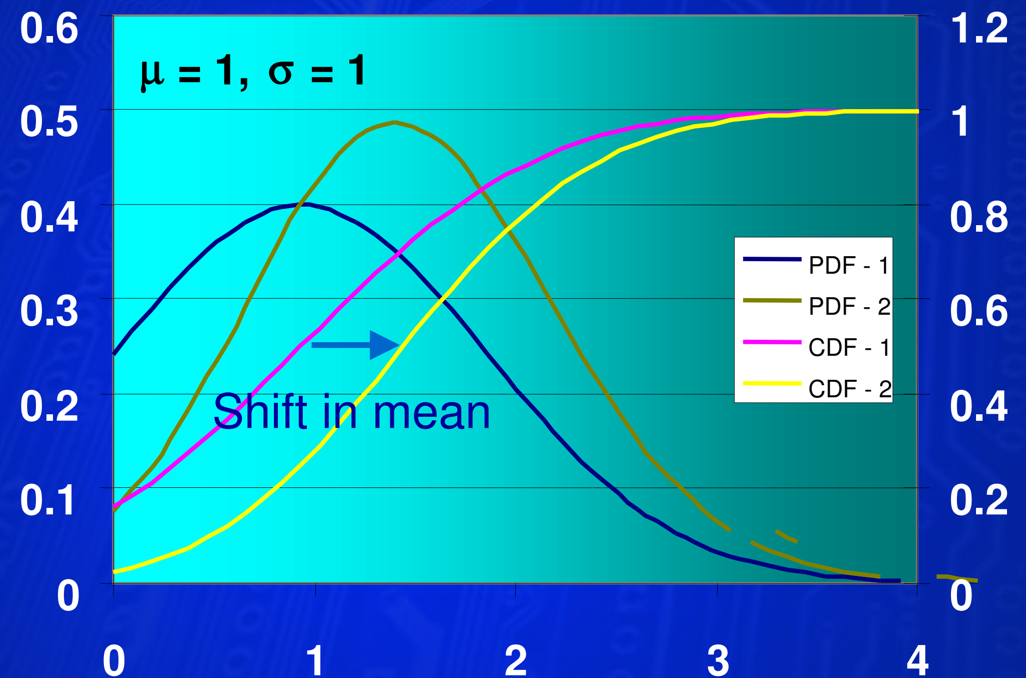
$$t_{chip} \le T = t_1 \le T \wedge t_2 \le T$$



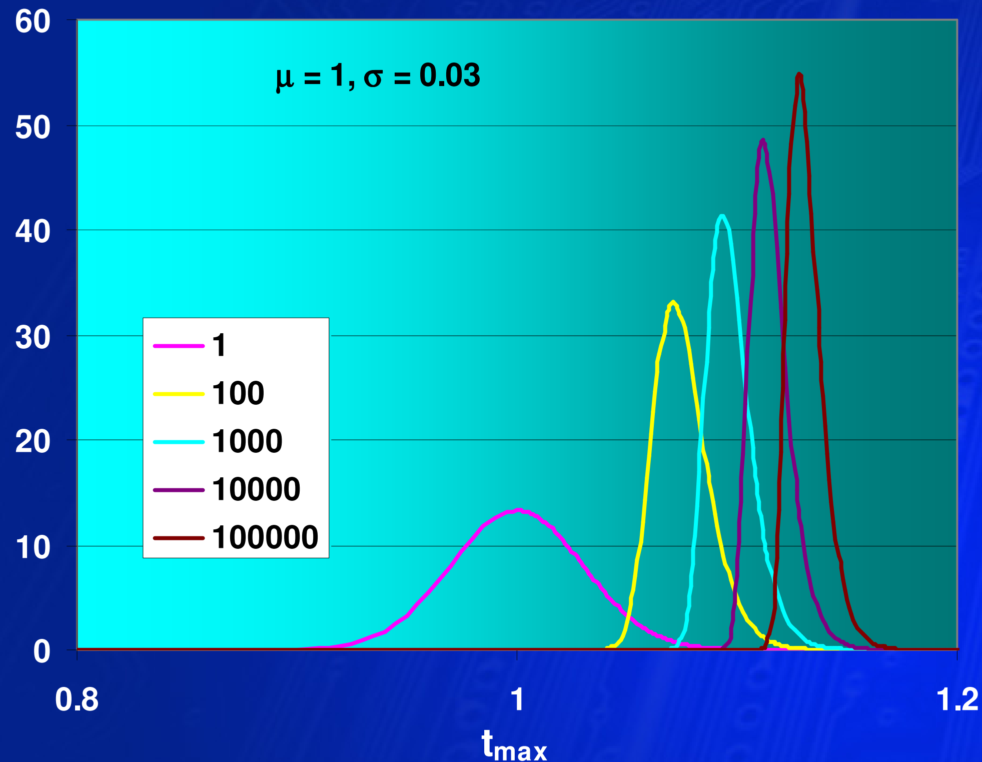
$$P_{chip}(t_{chip} \le T) = P(t_1 \le T)P(t_2 \le T)$$



$$cdf_{chip}(t) = cdf_1(t)cdf_2(t)$$



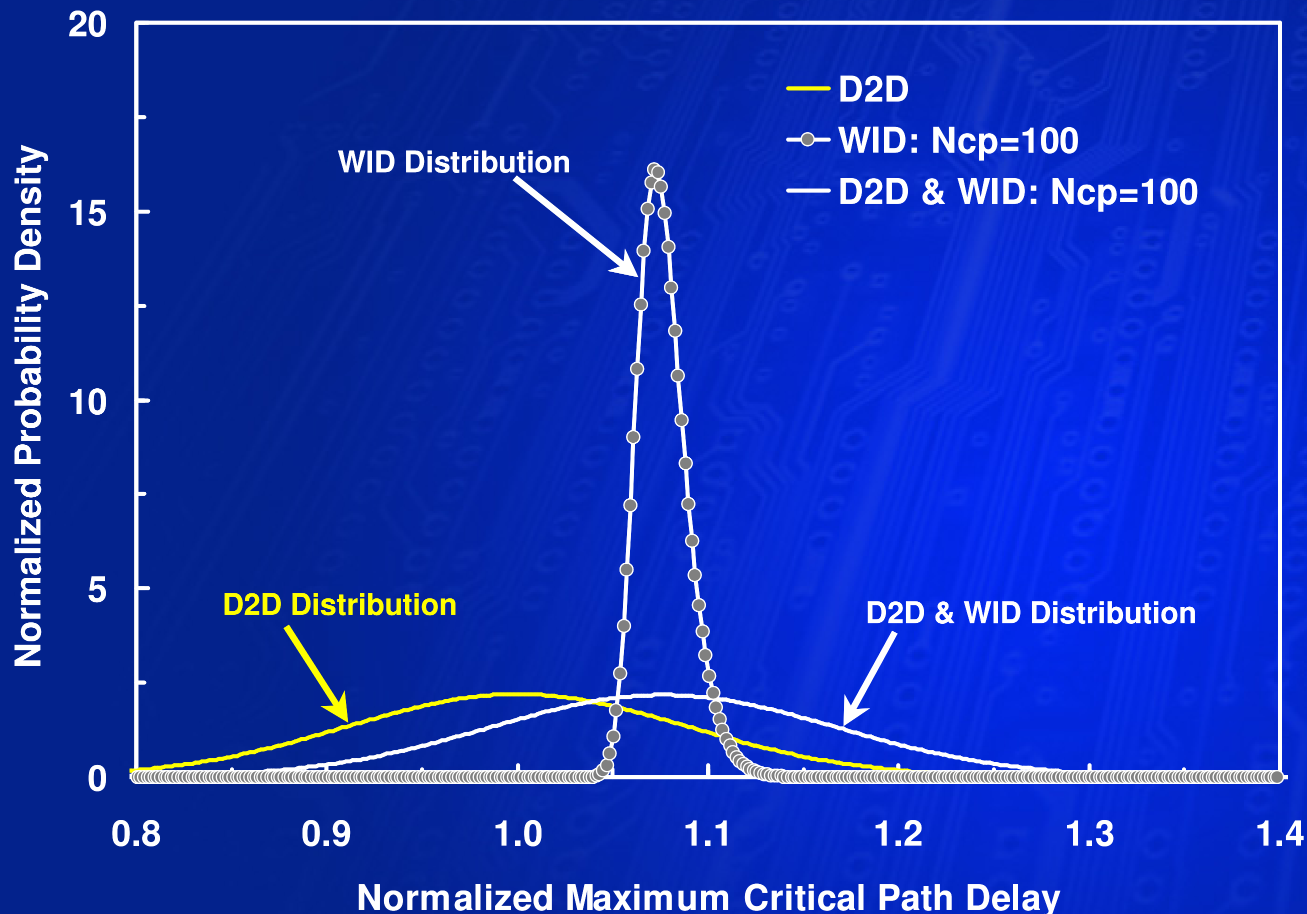
Chip-level delay distributions



Bowman, et al.
ISSCC 2001

- As the number (N_{cp}) of independent critical paths increases, the WID distribution mean increases and the variance decreases
- As N_{cp} increases to larger values, the dependency on N_{cp} decreases

Combining the die-to-die and within-die delay distributions



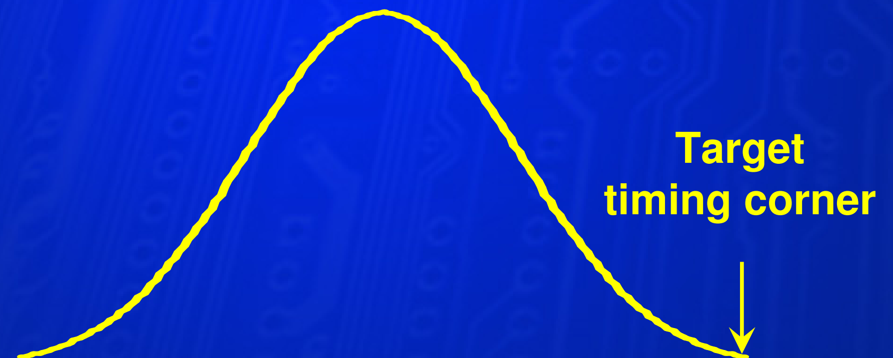
Bowman, et al.
ISSCC 2001

- Within-die variations impact the delay mean
- Die-to-die variations impact the delay variance

SSTA: Reduce pessimism of corner based analysis

- Assumption: Corner-based analysis is too pessimistic
  - The deterministic worst-case corner is derived by simultaneously setting each parameter at its 3σ value
  - Not correctly accounting for die-to-die and within-die random and systematic components during cell characterization leads to unwarranted pessimism during STA

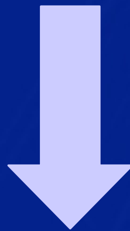
$$\sigma^2_{Xtor} = \sigma^2_{D2D} + \sigma^2_{sys} + \sigma^2_{rand}$$



Recovery of random variation component by SSTA

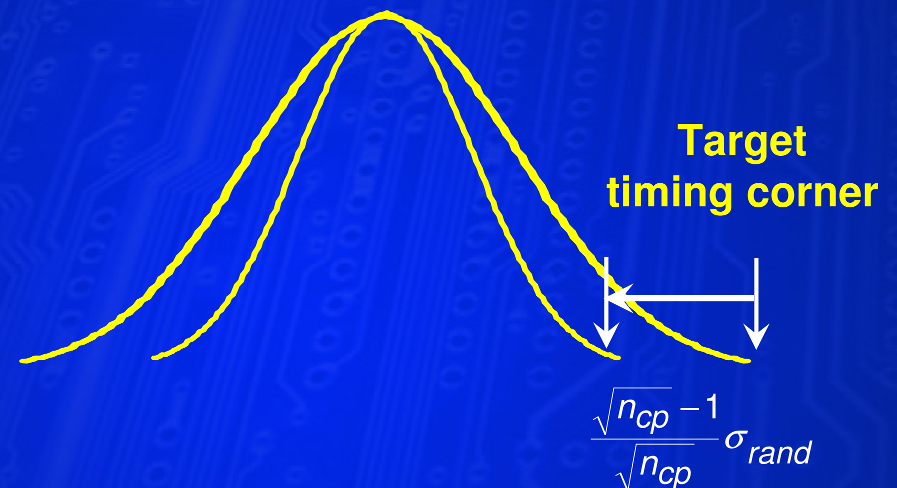
Worst-case process corner

$$\sigma^2_{Xtor} = \sigma^2_{D2D} + \sigma^2_{sys} + \sigma^2_{rand}$$



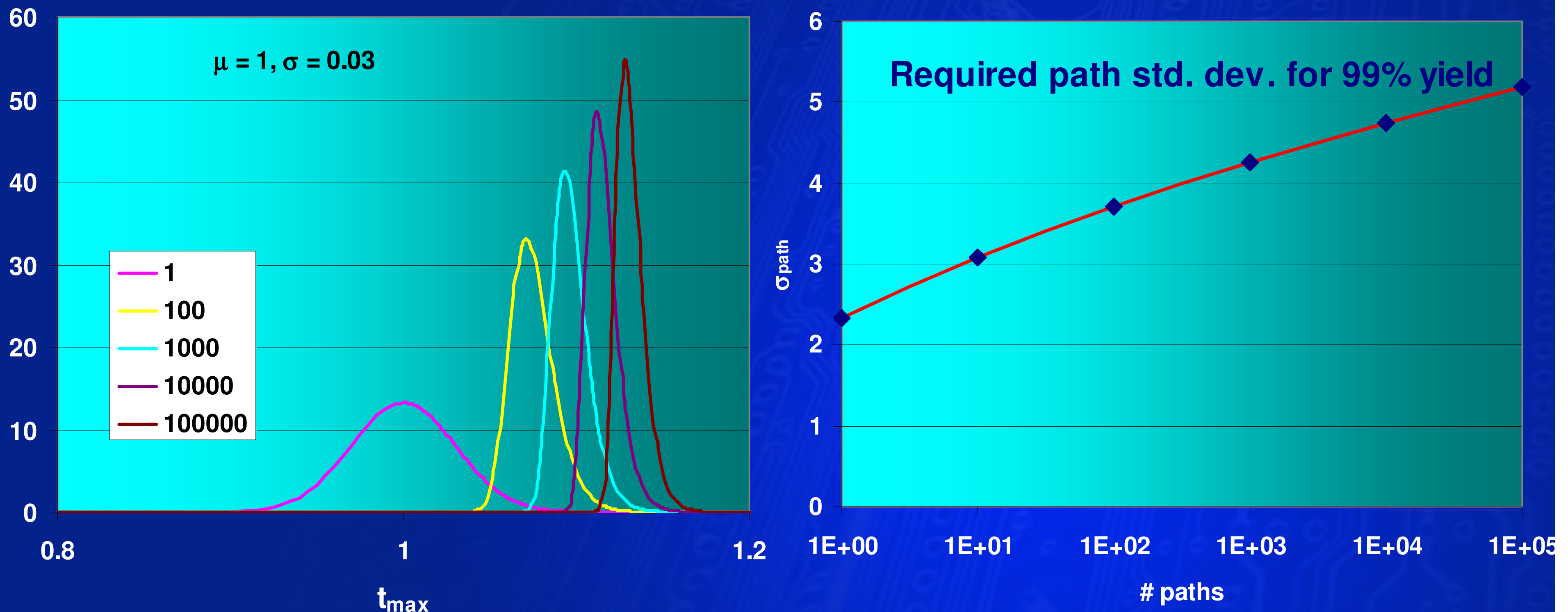
Worst-case product corner (potential)

$$\sigma^2_{path} = \sigma^2_{D2D} + \sigma^2_{sys} + \frac{1}{n_{cp}}\sigma^2_{rand}$$



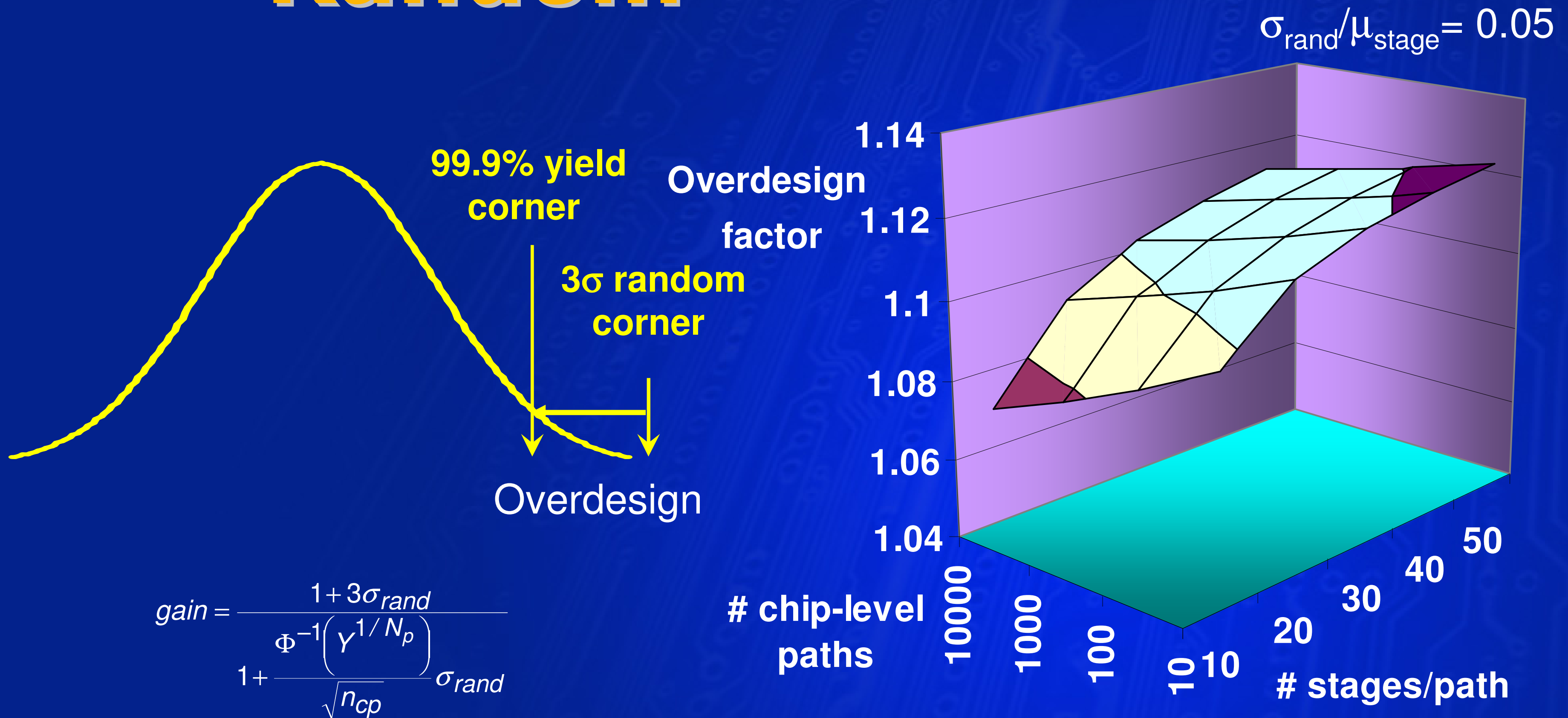
- **The transistor-to-path sigma reduction mitigates the the effect of random variations on product yield**

The critical path effect on random variations



- The critical paths-to-die timing pushout exacerbates the effect of random variations on product yield
- The pushout effect counters the path averaging effect on random variations

SSTA gain relative to 3σ corner analysis: Random

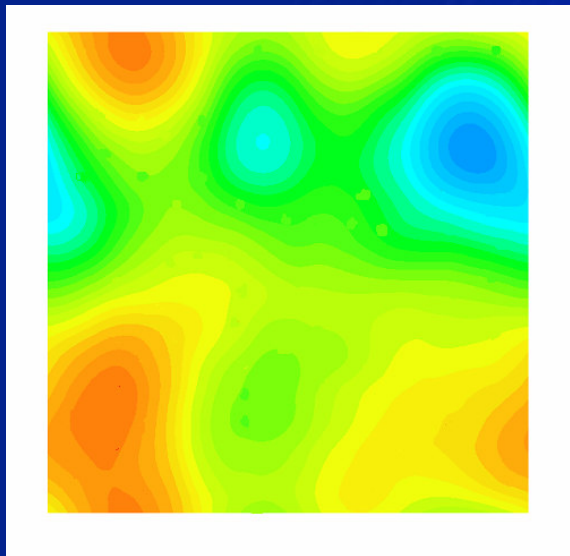


$$gain = \frac{1+3\sigma_{rand}}{1+\frac{\Phi^{-1}\left(Y^{1/N_p}\right)}{\sqrt{n_{cp}}}\sigma_{rand}}$$

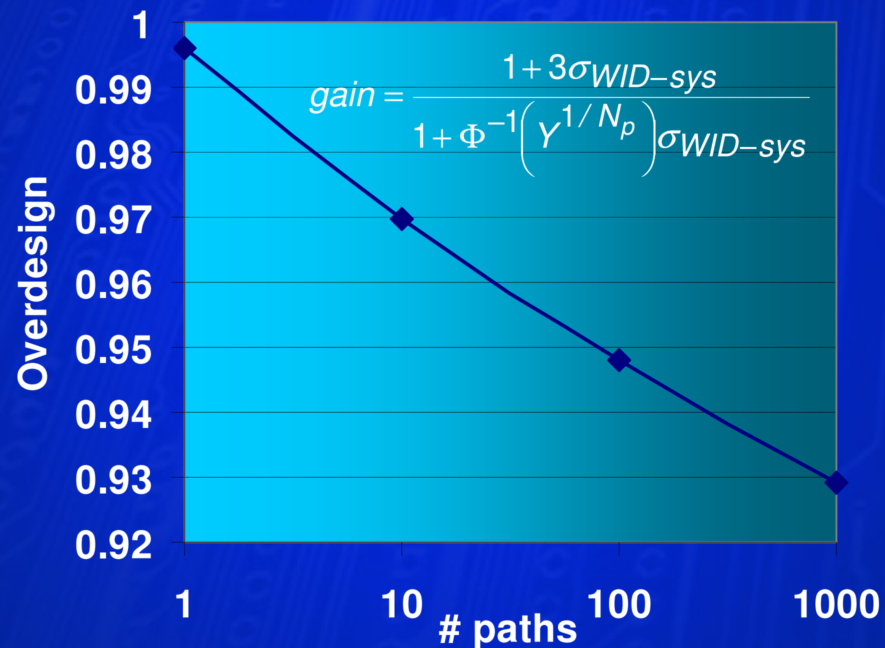
- **SSTA provides a potential gain relative to a 3σ corner-based STA for purely random variation**

SSTA gain relative to 3σ corner analysis: Systematic WID variation

- **Systematic WID affect all stages along a path equally – no averaging effect due to path localization**
- **The number of systematic independent critical paths is determined by the spatial correlation distance**

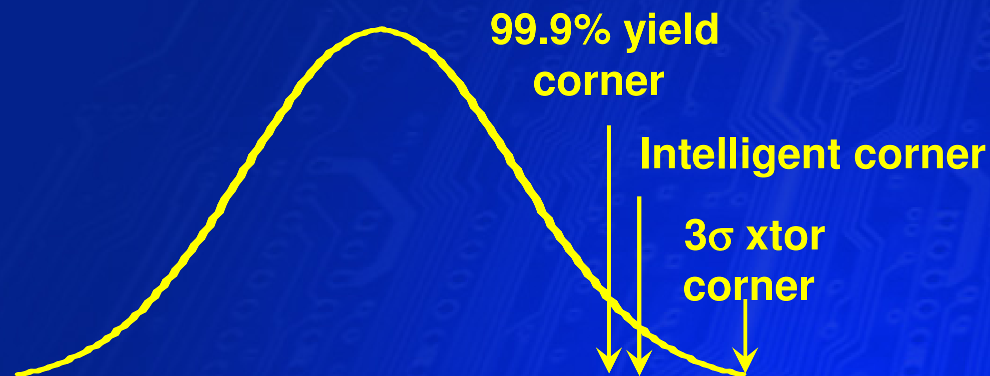


Systematic WID -- sample



➢ **SSTA gain relative to 3σ corner-based STA is reduced for systematic WID CD variation**

The intelligent corner approach

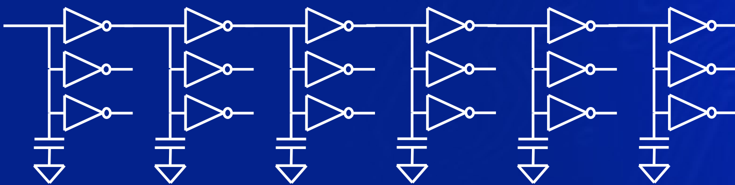


- Instead of SSTA, apply a global technique to pick a yield corner based on product information
  - Number of critical paths
  - Number of stages along critical paths
  - Available process variation data
  - Test methodology
- One such technique is described in Najm, Menezes [DAC 04]

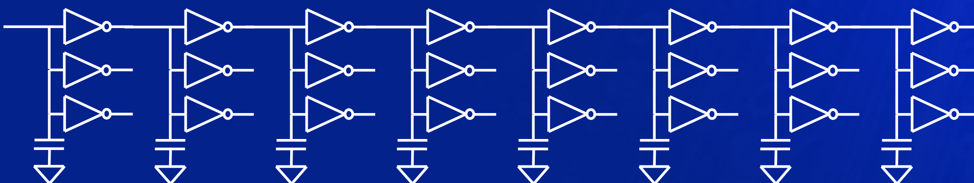
Statistical optimization

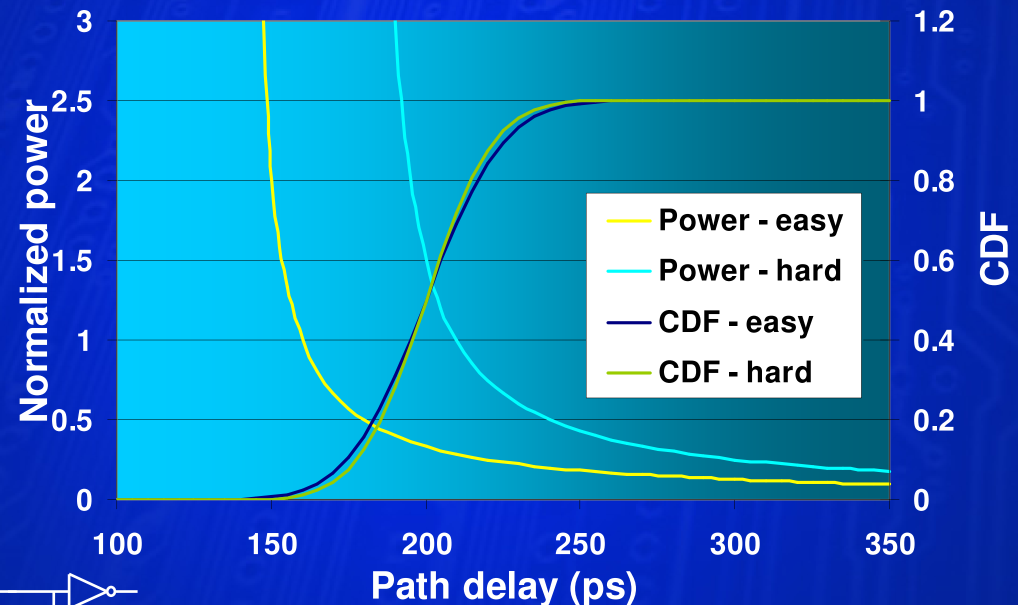
- Differences in delay variations of different paths depend on the number and type of stages
- Due to differences in power sensitivity to delay, power can be improved by budgeting timing yield loss based on power sensitivity

$n_{easy} = 6$



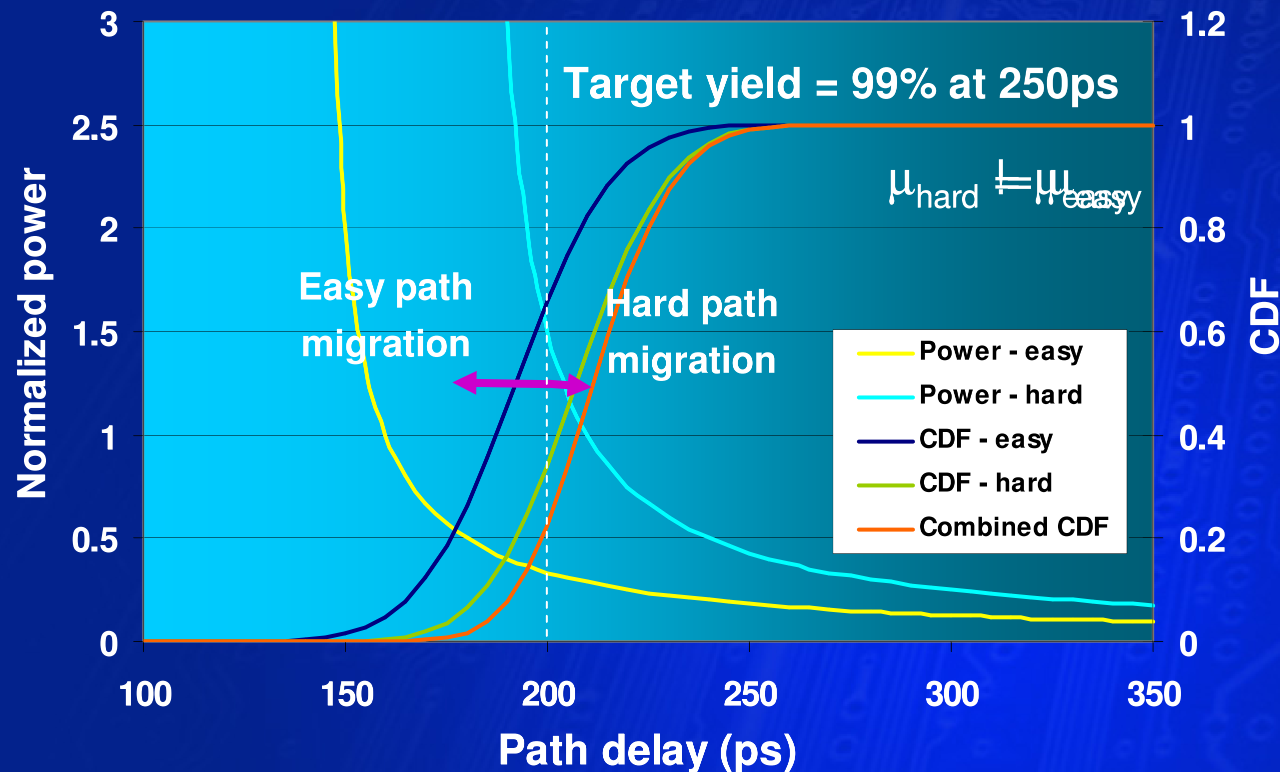
$n_{hard} = 8$





Burns, et al.
DAC 2007

Exploiting power sensitivities



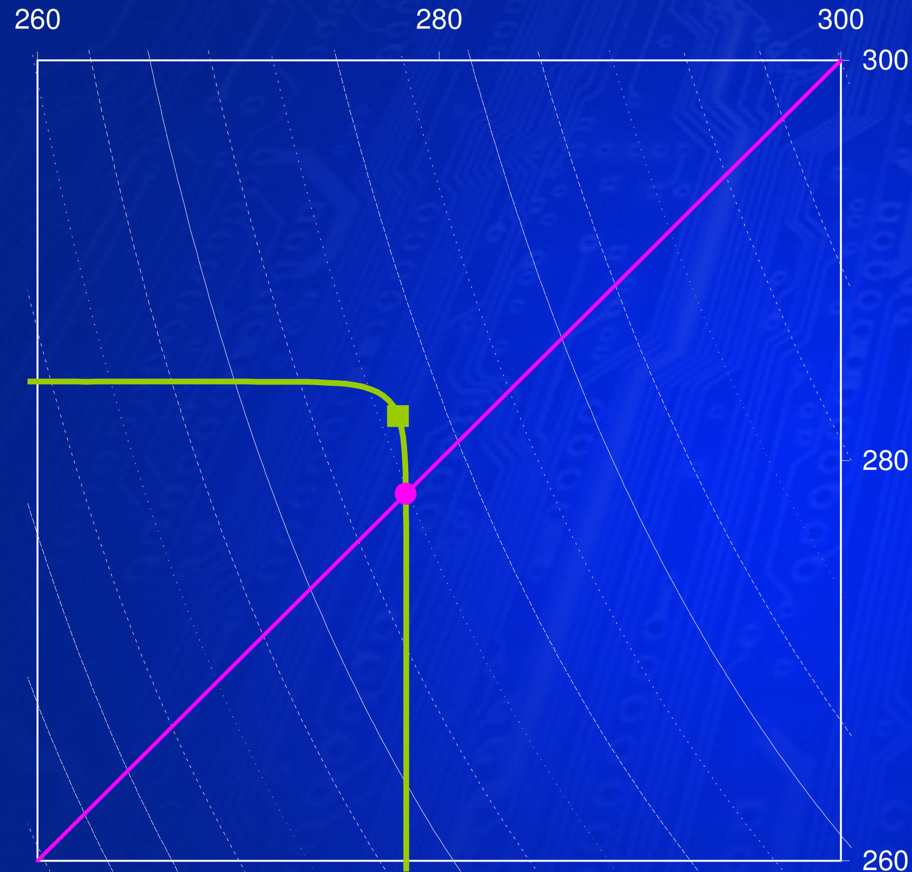
	Conventional optimization	Statistical optimization
μ_{hard}	200 ps	207
μ_{easy}	200ps	192
Power (ave.)	1.83	1.43

Statistical optimization benefits over global guardbanding

Dashed iso-power level curves spaced at 2% intervals.

The green locus represents those design targets that result in a 300ps median max delay

The pink line shows design targets corresponding to some global guardband



N_{hard} = 1000
N_{easy} =10000

Stage sigma = 5%

$p_{total} = n_{hard}\, p_{hard} + n_{easy}\, p_{easy}$

Burns, et al.
DAC 2007

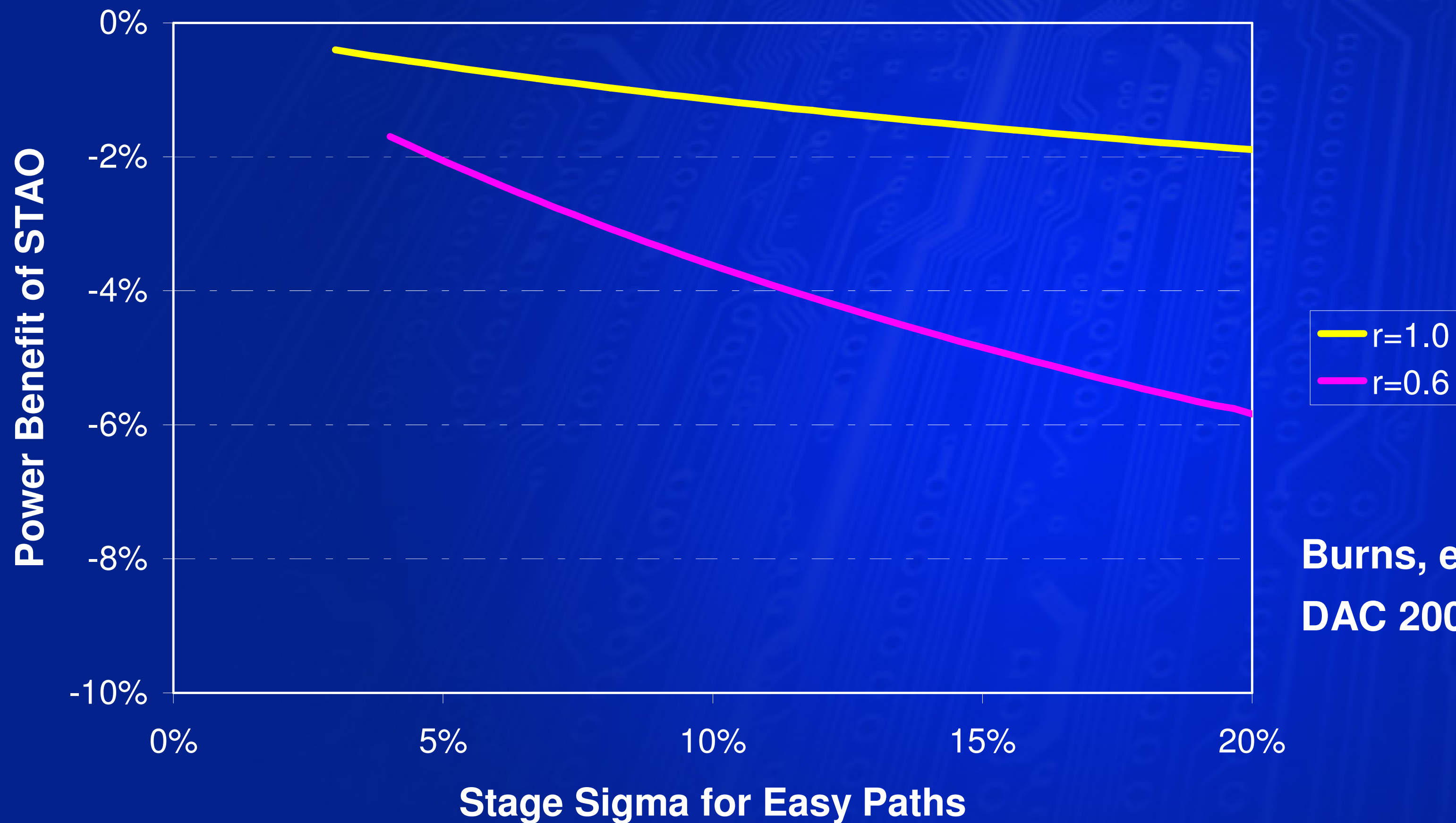
- **Statistical optimization benefits are mitigated because the pushout effect for a large number of paths is not so pronounced**

Power Benefit Sweeping Magnitude of Variation

neasy =10000
nhard = 1000

Easy stage sigma = sweeping
Hard stage sigma = r*(Easy stage sigma)

Iso-hardware intensity.

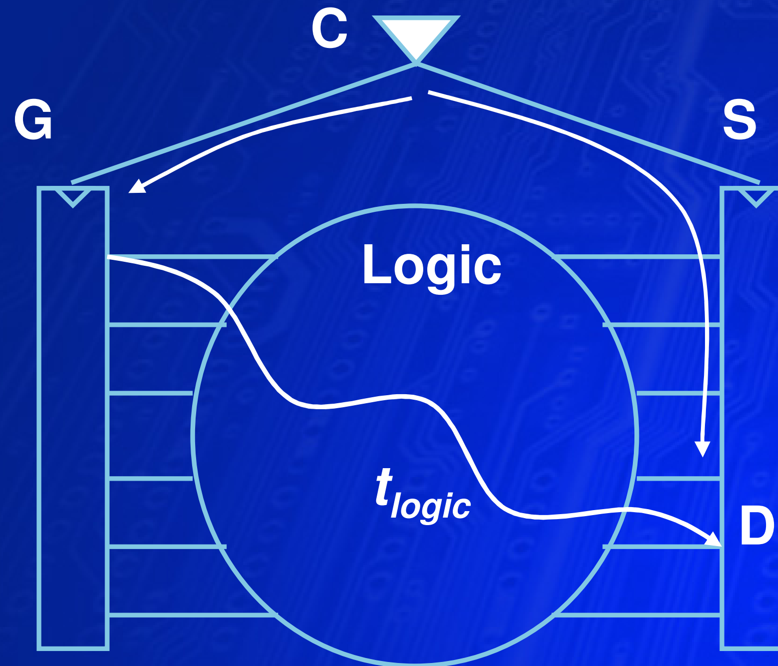


Burns, et al.
DAC 2007

Statistics in design

- Min-delay analysis
- Max skew computation
- Bin split prediction
- Leakage-performance estimation

Basic timing checks: Hold check

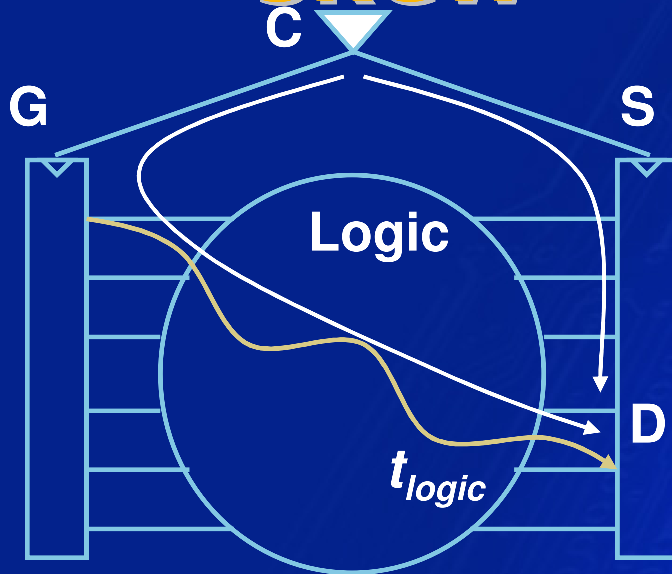


$$t_{CG} + t_{\text{logic}} - t_{CS} - t_{hold} \geq min_skew$$

These are hard numbers in traditional STA

This is statistically calculated

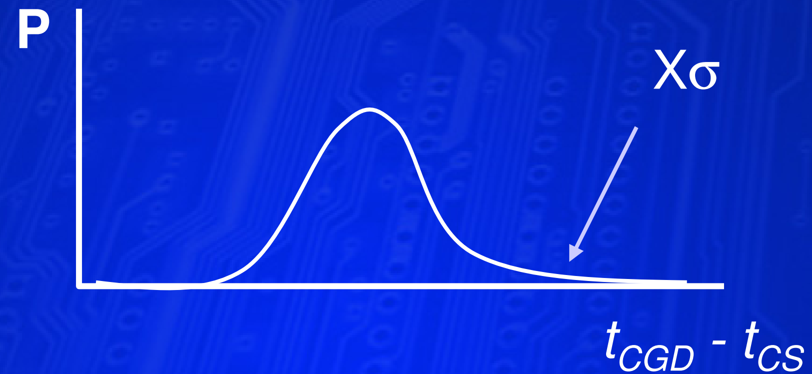
Calculating min-delay skew



$$(t_{CGD} - t_{CS} - t_{hold})_{X\sigma} \geq 0$$

where

$$X\sigma = f(\text{DPM spec})$$



- **The min-delay skew is calculated based on**
  - **Number of min-delay paths on the chip**
  - **Conservatism in hold characterization for sequentials**
  - **Process variation spec at fast corner**
  - **Sophistication of min-delay analysis methodology**

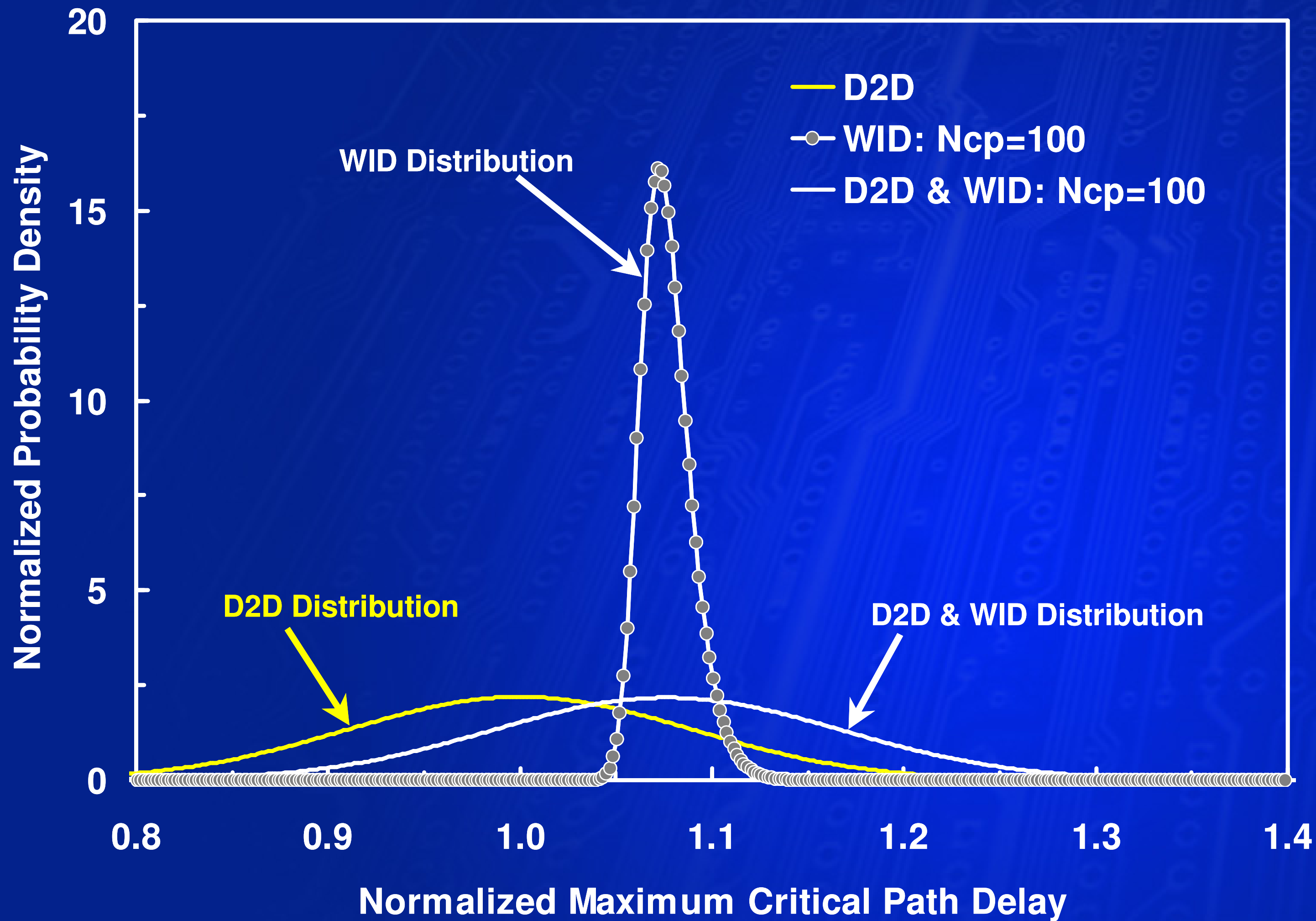
Max skew computation

Clock grid
C
Sampling path, CS
Clock buffers
G
Clock-data path, CGD
S
D

$$\sigma^2_{margin} = \sigma^2(t_{CS}) + \sigma^2(t_{CGD}) - 2\,cov(t_{CS}, t_{CGD})$$

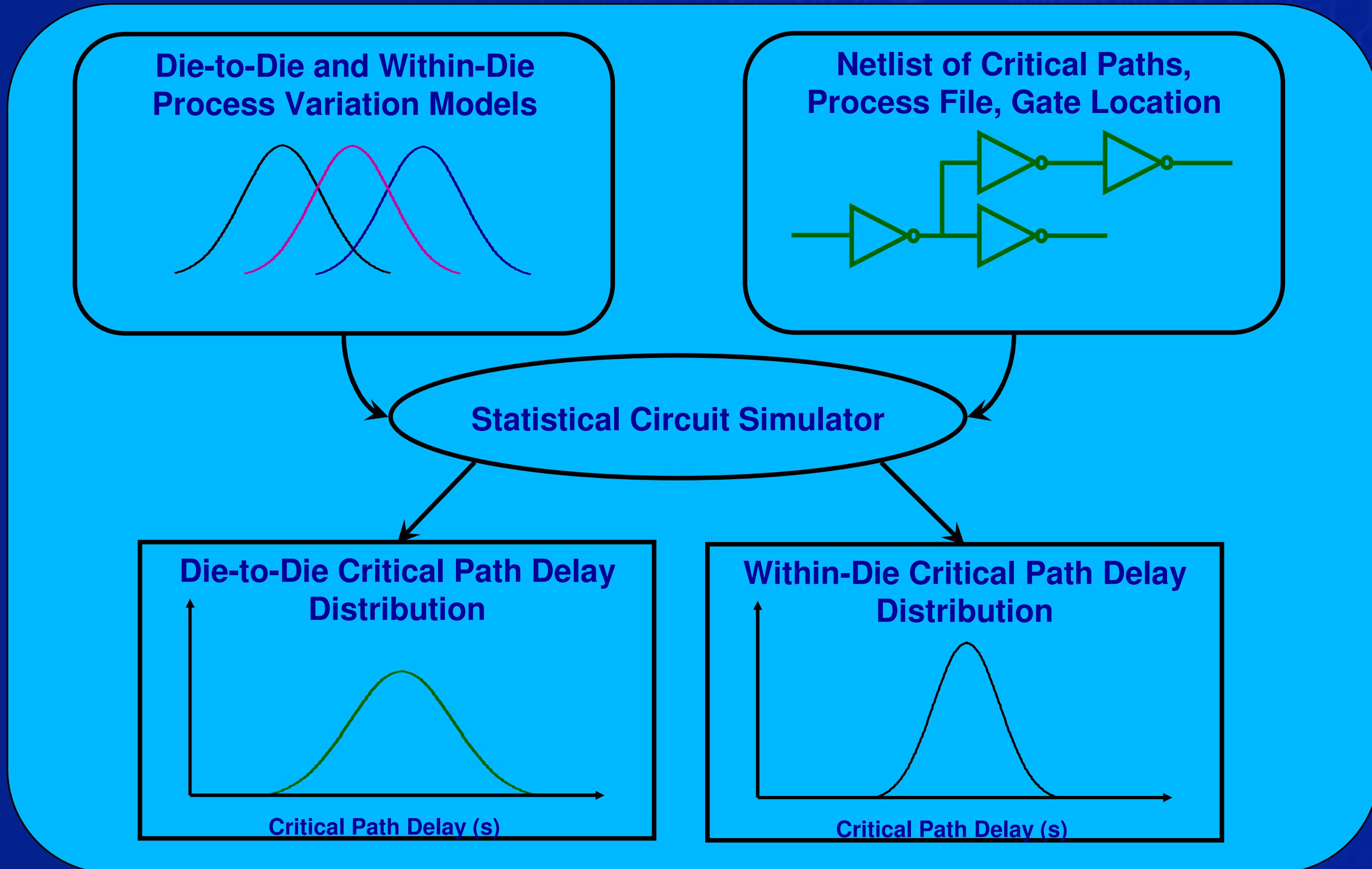
- **Max delay skew ($k\sigma_{margin}$) is based on**
  - **Testing methodology (binning?)**
  - **Process variation specs at nominal and slow corners**
  - **Conservatism in STA methodology**
  - **Clocking and clock distribution methodology**

FMAX prediction model



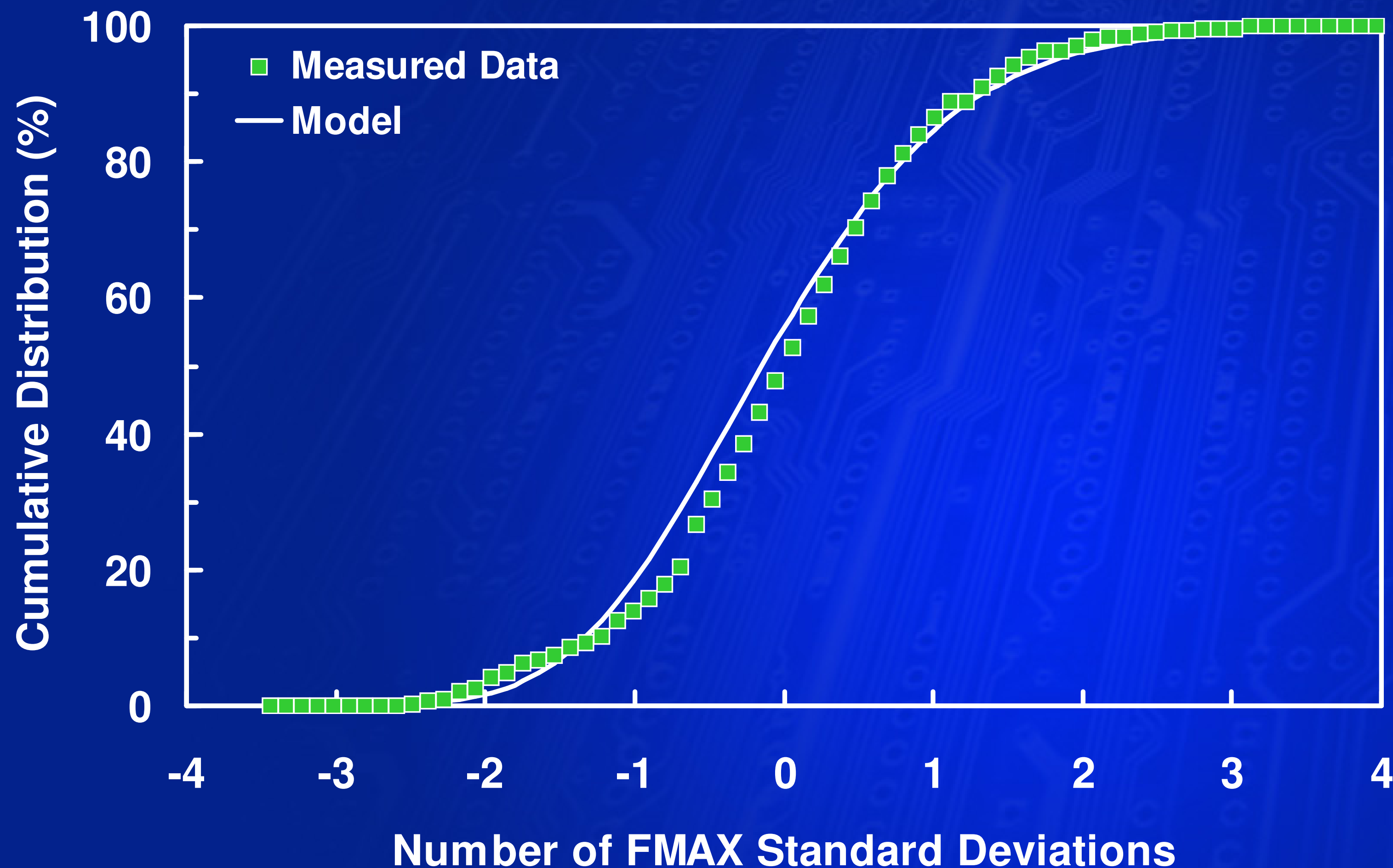
- Within-die variations impact the delay mean
- Die-to-die variations impact the delay variance

FMAX yield prediction



Bowman, et al., ISSCC 2001

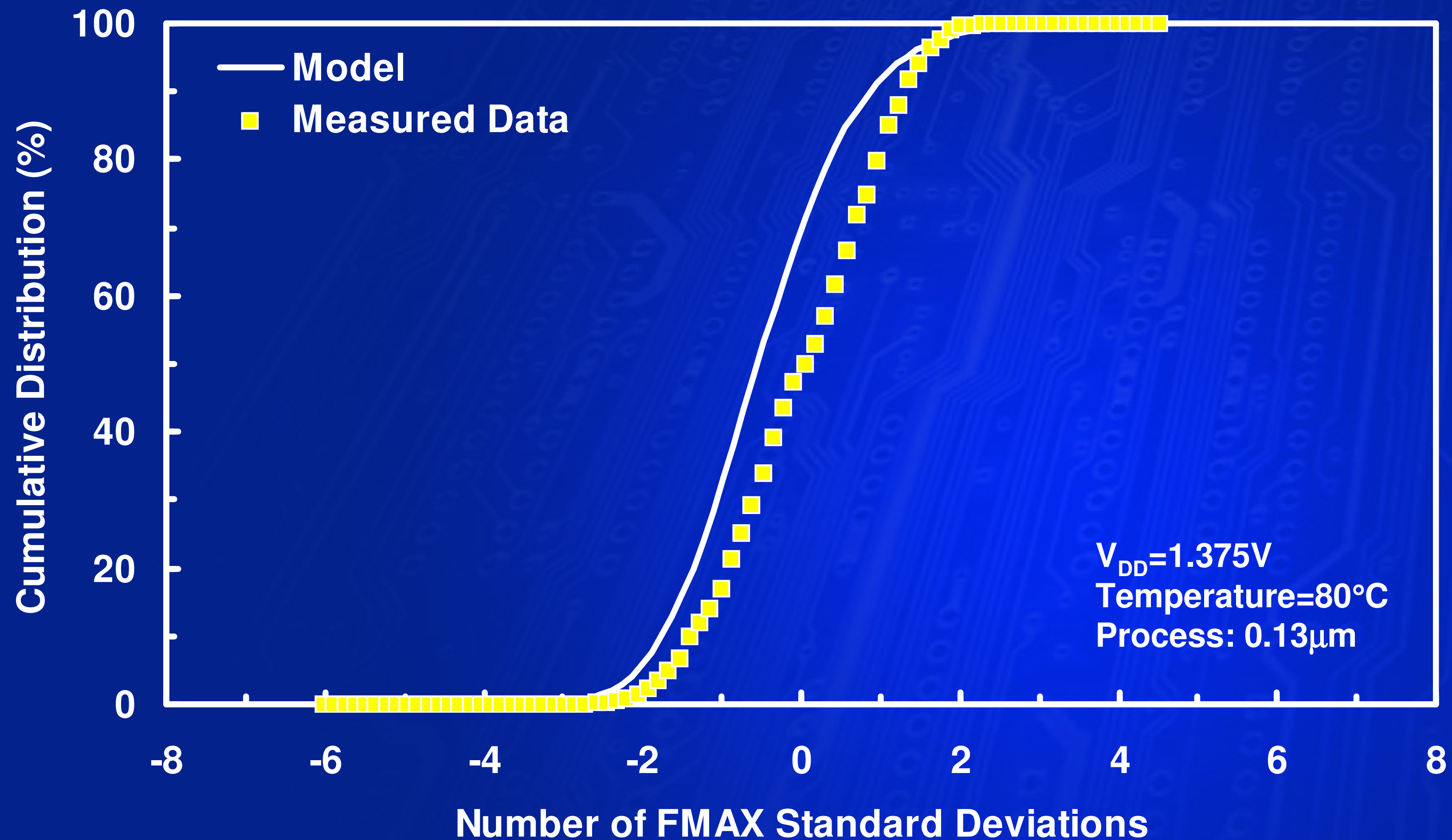
FMAX model comparison with a 0.25 μm processor



- Data represents ~50,000 die measurements taken at sort

Bowman, et al., ISSCC 2001

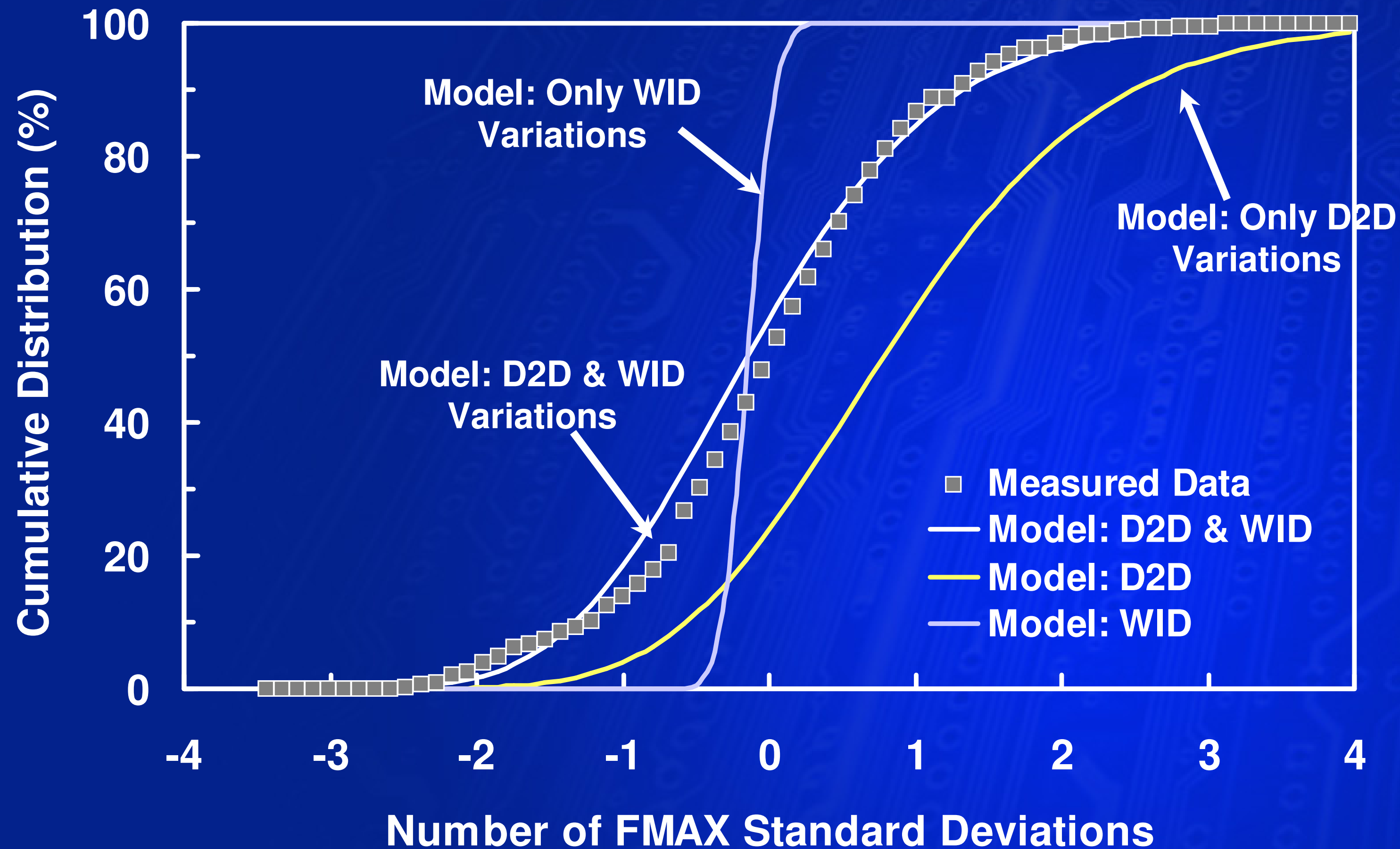
FMAX model comparison with a 0.13 μm processor



- Data represents ~1,000 die measurements taken at class

Bowman, et al., ISSCC 2001

Individual Contributions of D2D and WID Variations



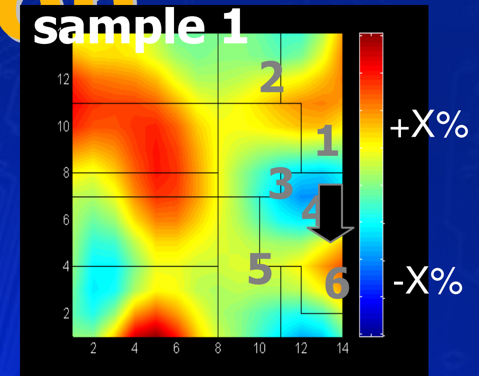
- Within-die variations primarily impact the FMAX mean
- Die-to-die variations primarily impact the FMAX variance

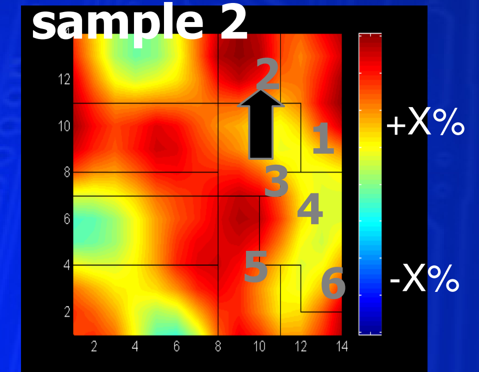
Bowman, et al., ISSCC 2001

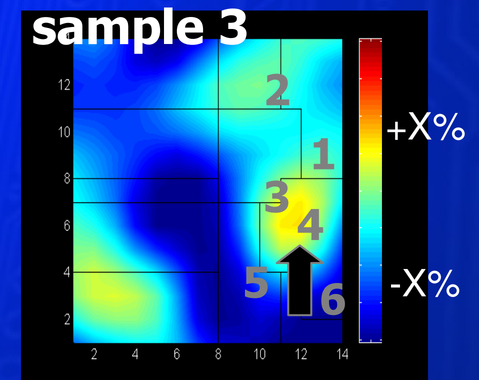
FMAX-ISB prediction

MONTE-CARLO ITERATIONS

1. Draw lots for process variation map
2. Map timing model onto variation map and recalculate path delay to find the slowest path that will define FMAX
3. Calculate the die's total leakage (ISB)





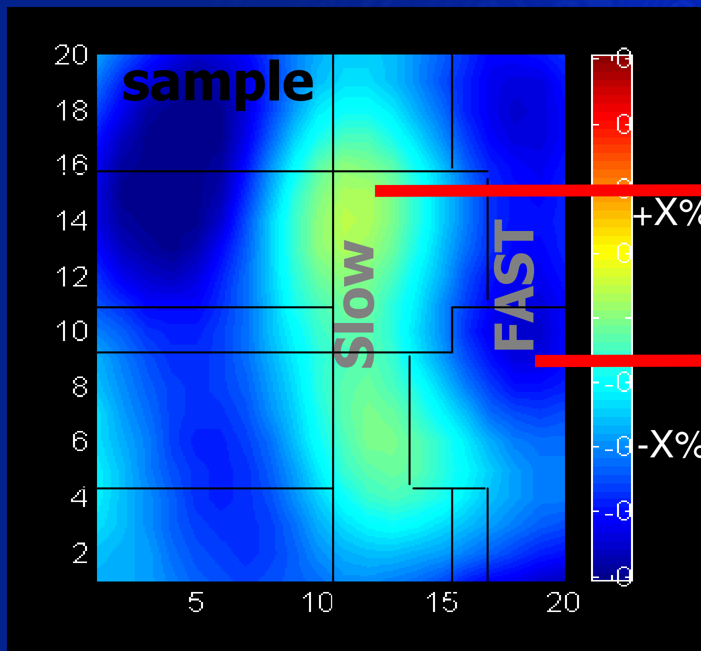


- FMAX is going to be defined by MANY paths (as a result of timing model re-order)
- The spatial distribution of the paths is very important to FMAX distribution

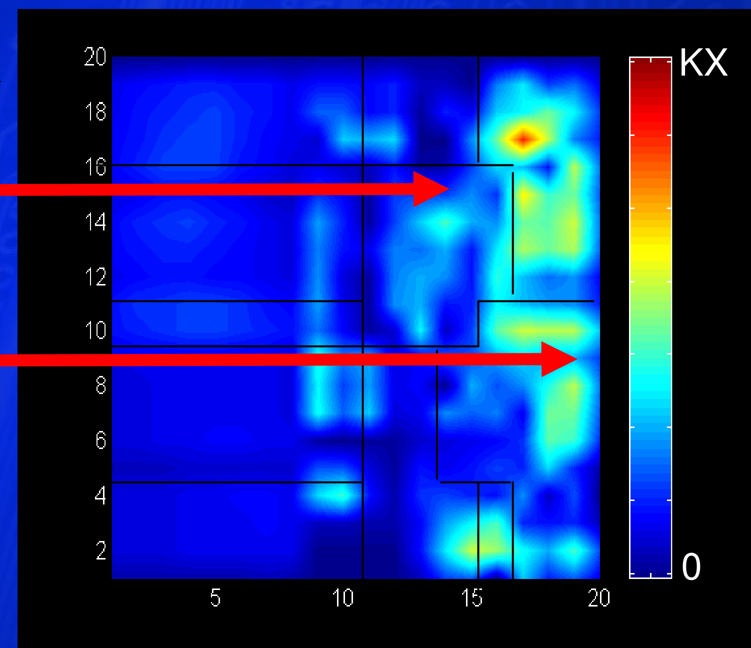
Abulafia, Kornfeld, Trans. VLSI 2005

Leakage modeling

$$I_{speculated_die} = \sum_{i,j} \sum_{type=n,p} \frac{I_{off}(type, L_{i,j}) \cdot W_{i,j}(type)}{SF}$$



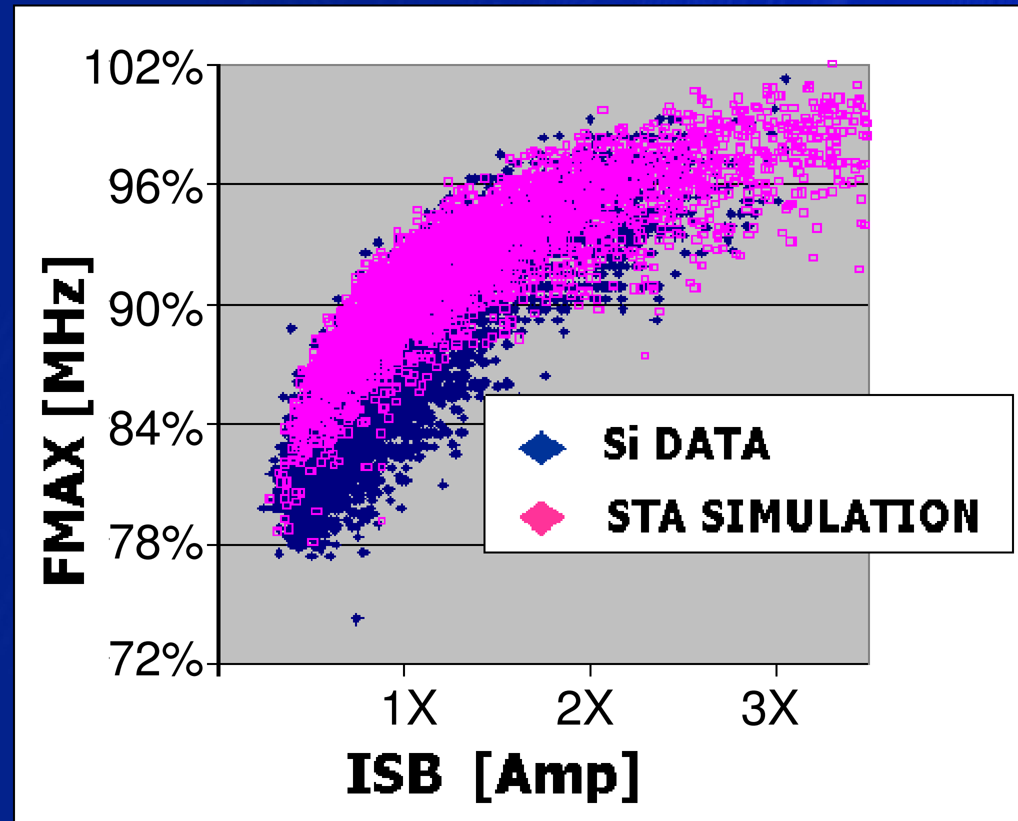
Channel-length variation map



Leakage map

Abulafia, Kornfeld, Trans. VLSI 2005

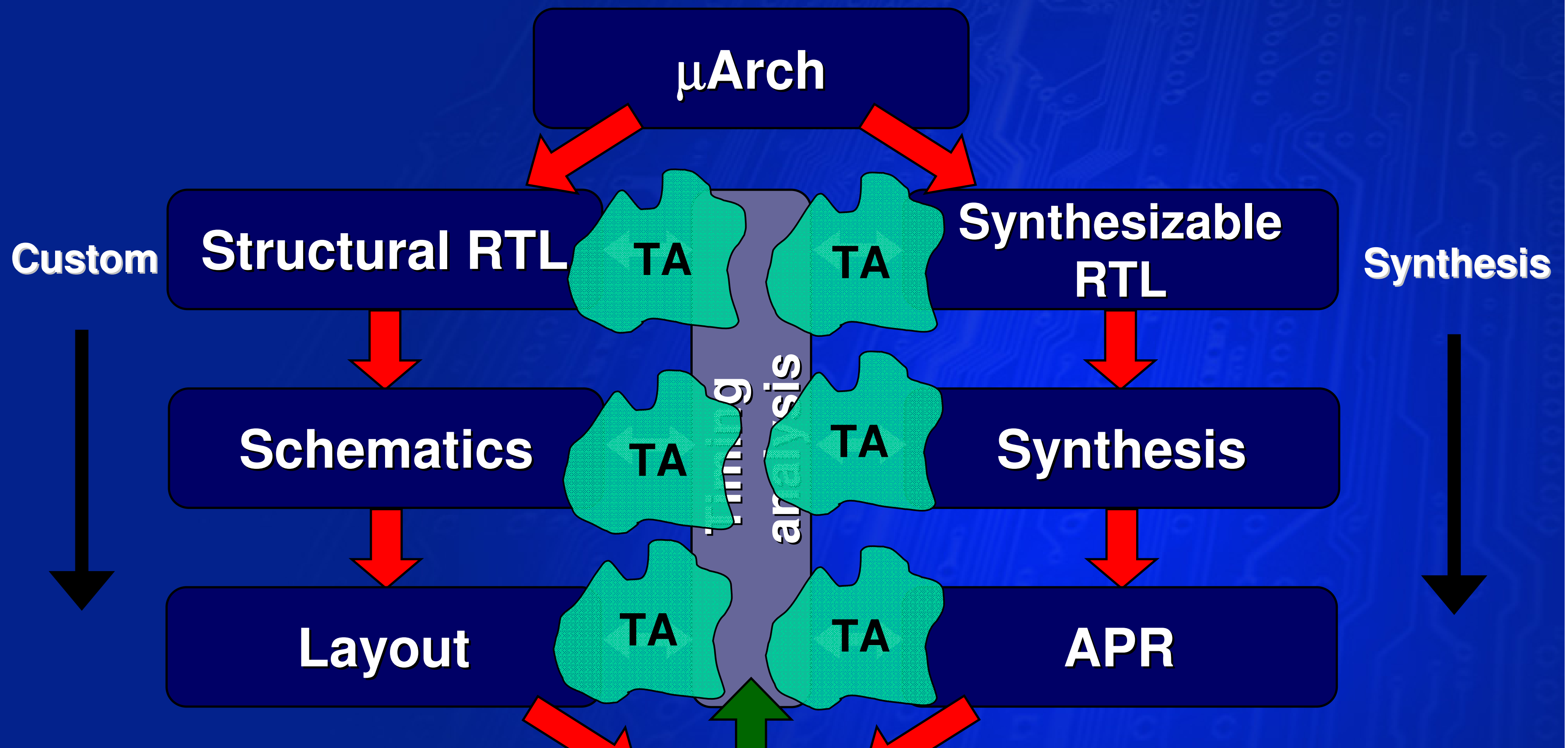
FMAX/ISB prediction comparison with Si



This analysis predicts the log-normal distribution and BANANA shape

Abulafia, Kornfeld, Trans. VLSI 2005

Design methdology



The ROI from any fine-grained analysis flow like SSTA has to be justified in terms of the overall design methodology

Conclusion

- Statistical techniques are applied in design with good silicon correlation
- The application of statistical techniques at fine-grained circuit level i.e. SSTA may provide a benefit relative to a worst-case process corner approach for some products
  - These benefits may not be large compared to an intelligent corner approach
- Statistical circuit optimization benefits are small
- A significant barrier to SSTA adoption is the impact on design methodology which has not been studied

I am convinced that He is not throwing dice.

Albert Einstein in a letter to Max Born, Dec. 12, 1926

Many designers would prefer it that way

References

- K. Bowman, S. Duvall, J. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution," IEEE JSSC, 2001.
- Y. Abulafia, A. Kornfeld, "Estimation of FMAX/ISB in microprocessors," IEEE Trans. VLSI, 2005.
- S. B. Samaan, "The impact of device parameter variations on the frequency and performance of VLSI chips," ICCAD 2004.
- S. Burns, M. Ketkar, N. Menezes, K. Bowman, J. Tschanz, V. De, "Comparative analysis of conventional and statistical design techniques," DAC2 007.
- F. Najm, N. Menezes, "Statistical timing analysis based on a timing yield model," DAC 2004.

Q & A

intel®
Leap ahead™