Statistical Analysis and Optimization of Asynchronous Digital Circuits

Tsung-Te Liu and Jan M. Rabaey

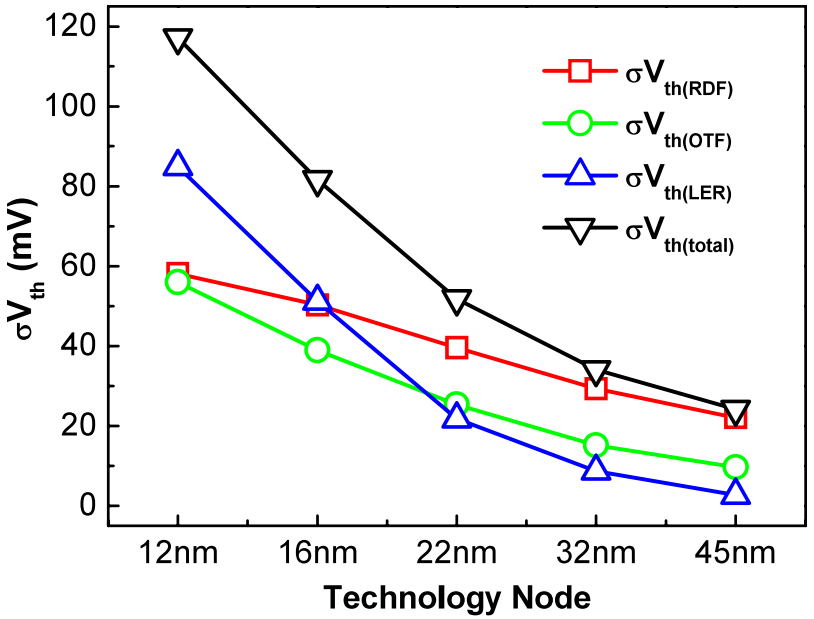
University of California, Berkeley

Outline

- Motivation
- Variability model of CMOS digital circuit
- Performance model for different timing schemes
- Performance comparison
- Conclusion

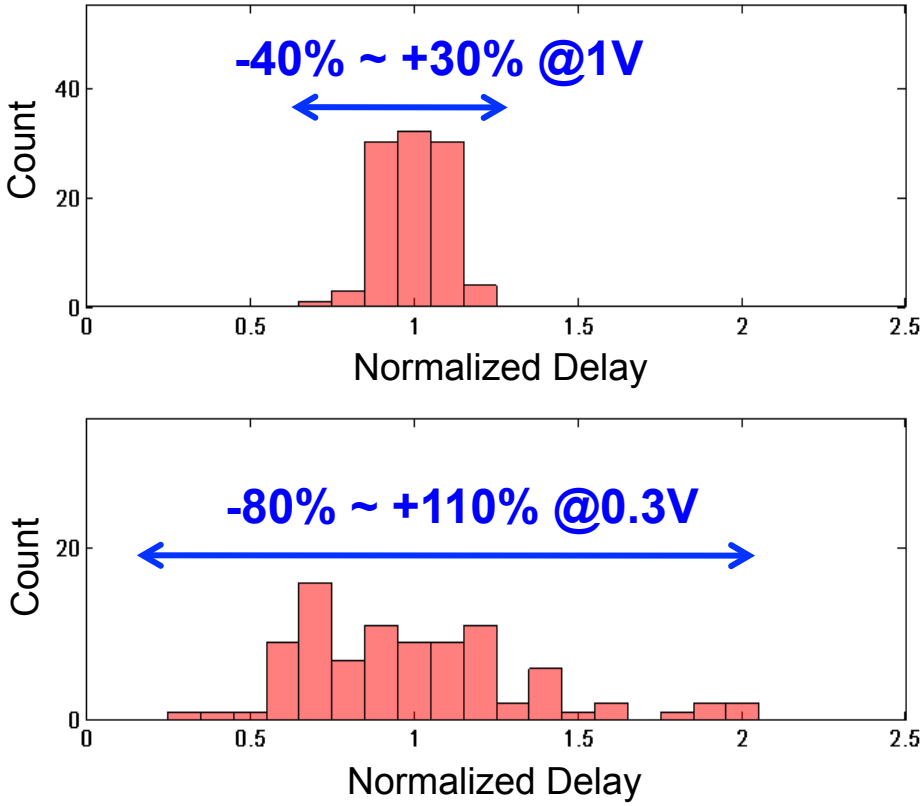
Variability Continues to Increase as Technology and Voltage Scales Down

Device variability vs. Technology node



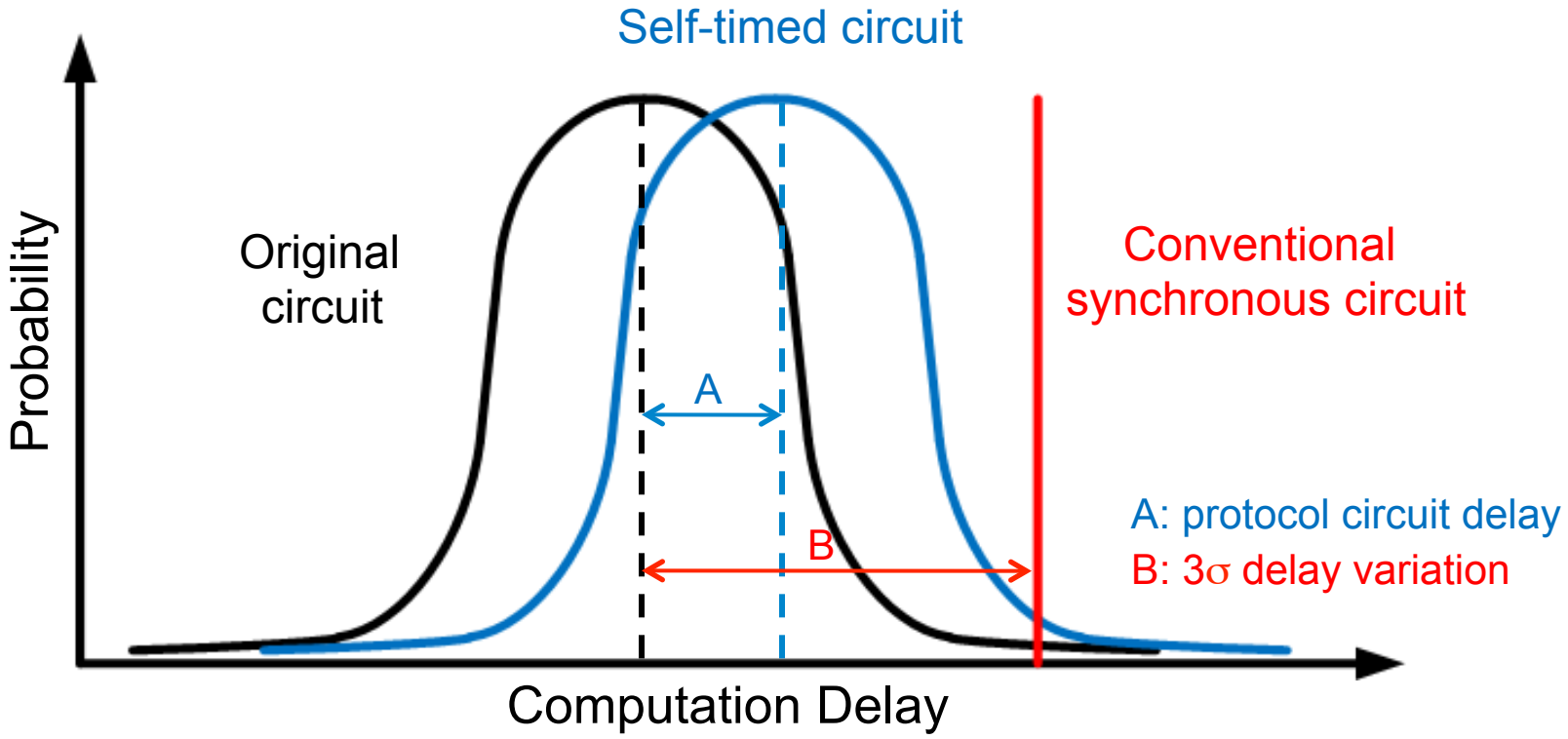
[Cao, ASU]

Delay spread due to process variations



- Higher variability with finer design rules and larger wafers
- Higher variability with lower supply voltages

Circuit Performance Characteristics with Different Timing Schemes



- Self-timed circuit is a variation-monitoring circuit by itself
- Becomes advantageous when the variation is large (B>A)
- Statistical analysis framework is necessary

Statistical Analysis Framework

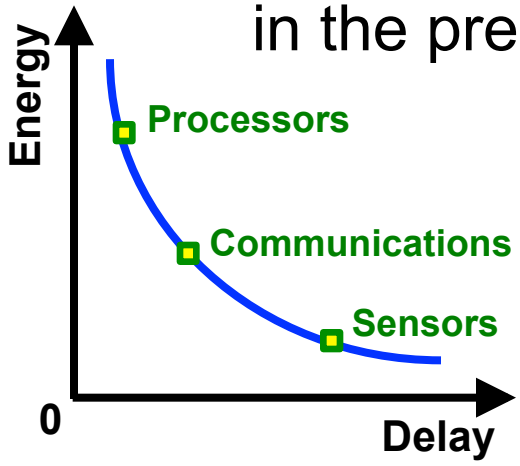
Circuit Variability Model

- Supply voltage
- Logic depth
- Width and length
- Body bias

Performance Model

- Computation overhead
- Communication overhead
- Delay and energy performance

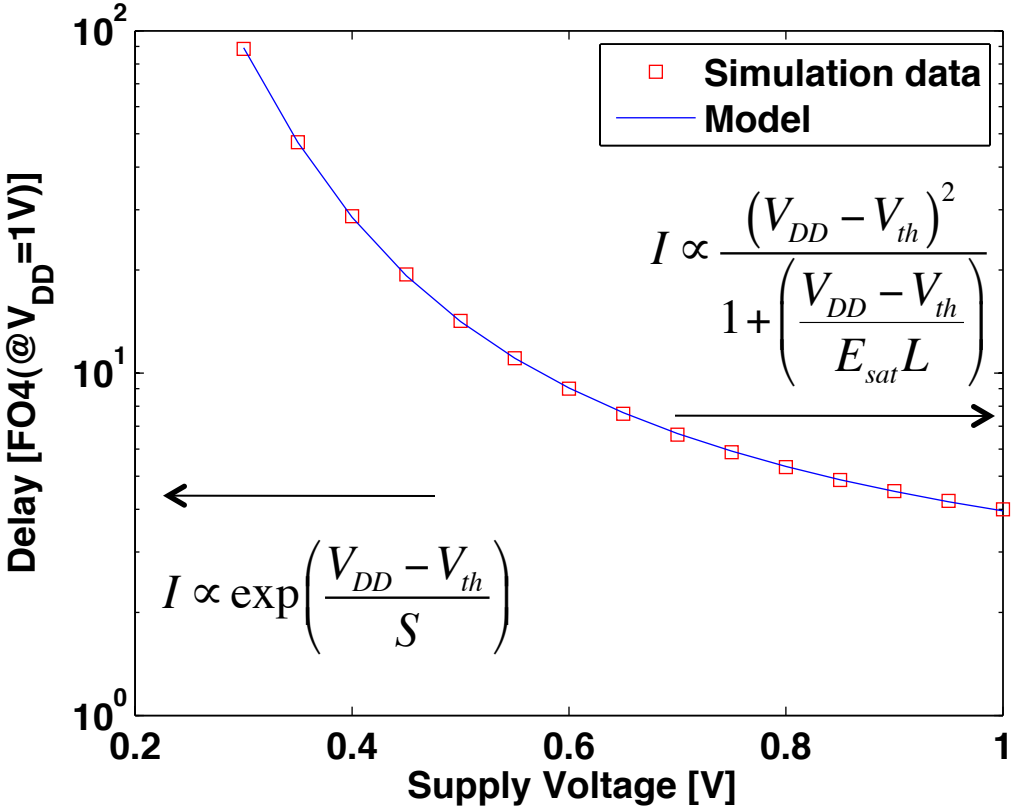
Determine the optimal timing strategy in the presence of variability



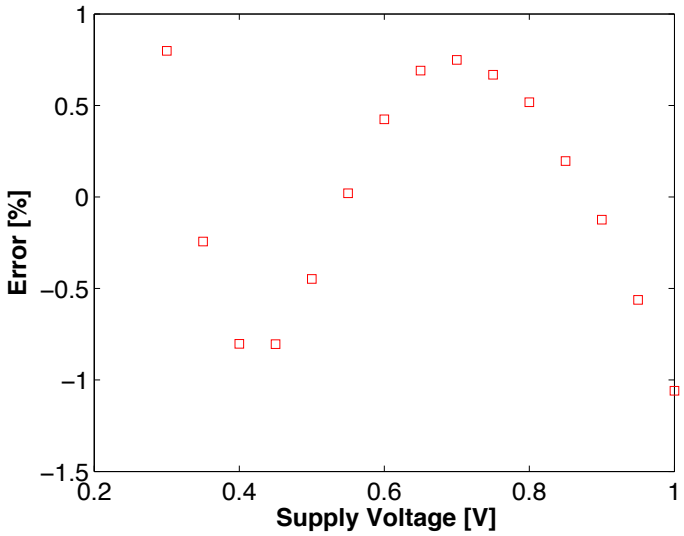
Outline

- Motivation
- Variability model of CMOS digital circuit
- Performance model for different timing schemes
- Performance comparison
- Conclusion

Delay Model of CMOS Digital Circuit



4-stage FO4 INV chain



$$I \propto \frac{\left\{\ln\left[1+\exp\left(\frac{V_{DD}-V_{th}}{2S}\right)\right]\right\}^2}{\left\{1+\ln\left[1+\exp\left(\frac{V_{DD}-V_{th}}{E_{sat}L}\right)\right]\right\}}$$

- One unified current model across different operating regions
- Model error <2% from 0.3V to 1V

Delay Variability Model

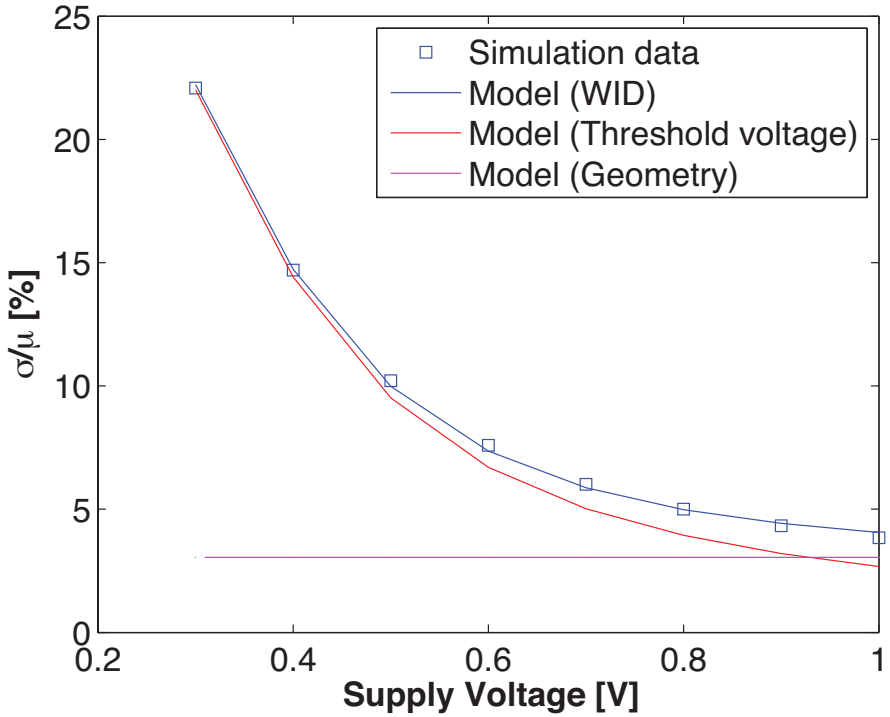
$$\frac{\sigma_{T_d}}{\mu_{T_d}} = \sqrt{\left(S_{T_d}^{V_{th}}\right)^2 \cdot \left(\frac{\sigma_{V_{th}}}{\mu_{V_{th}}}\right)^2 + \left(S_{T_d}^{K}\right)^2 \cdot \left(\frac{\sigma_{K}}{\mu_{K}}\right)^2}$$

$$S_{T_d}^{V_{th}} = \frac{\partial V_{th} / V_{th}}{\partial T_d / T_d}$$

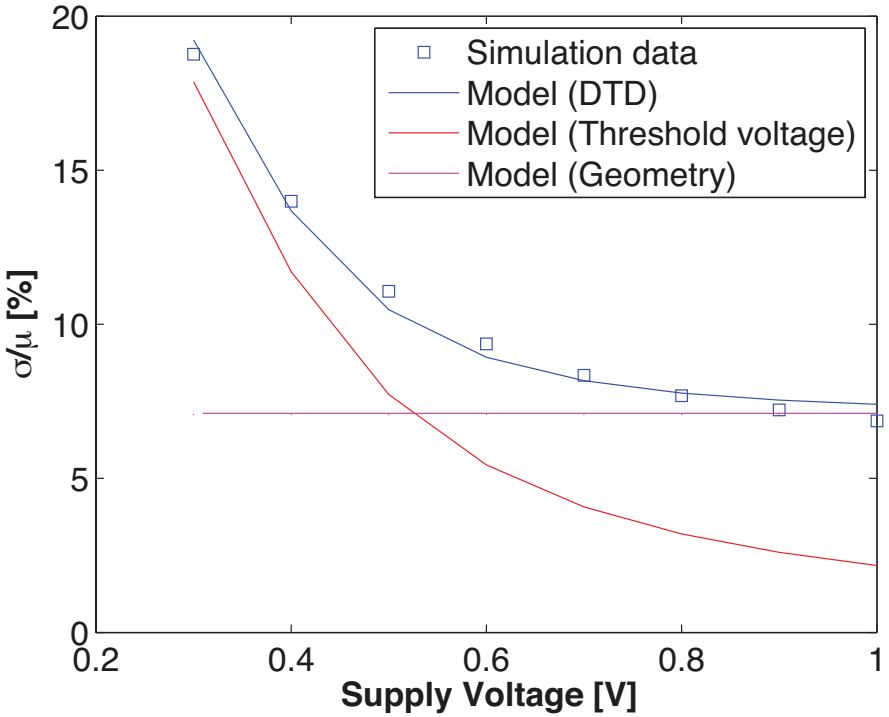
Threshold voltage

$$S_{T_d}^{K} = \frac{\partial K / K}{\partial T_d / T_d}.$$

Geometry

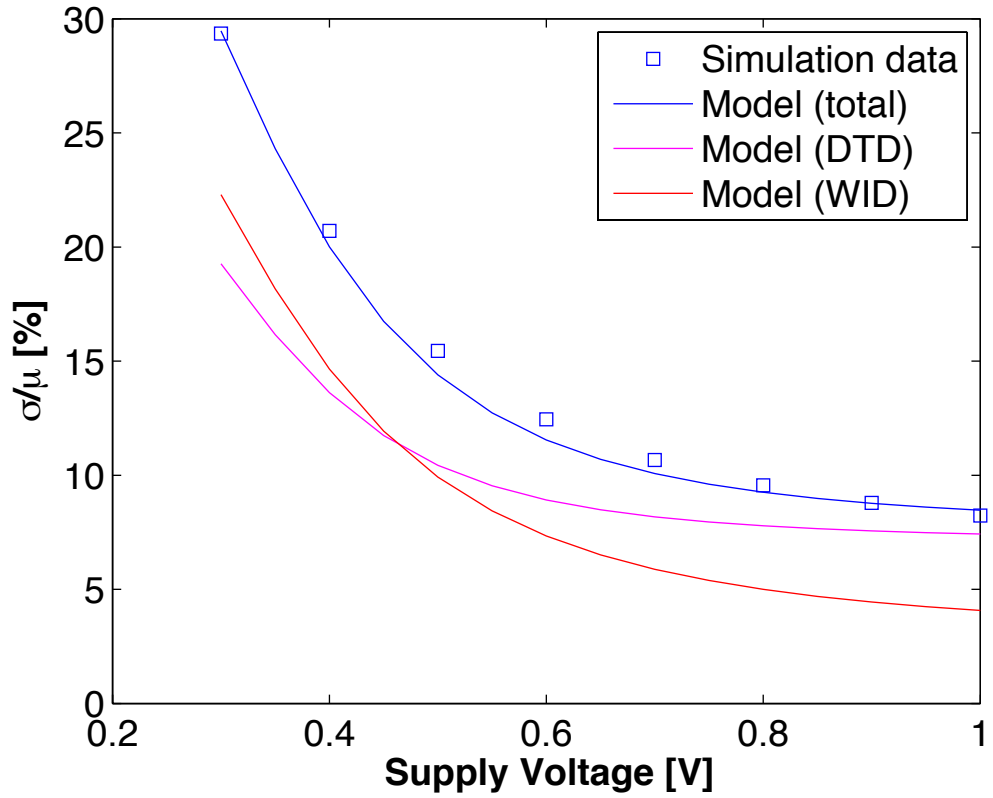


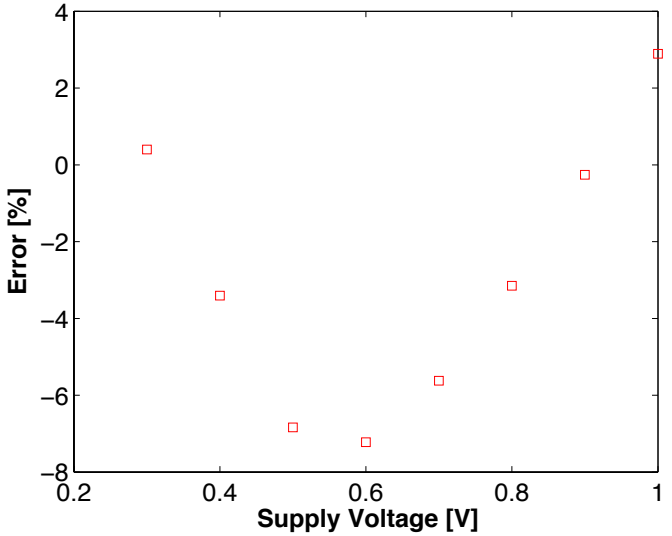
Within die variation (WID)
"Local mismatch"



Die-to-die variation (DTD)
"Global variation"

Delay Variability Model

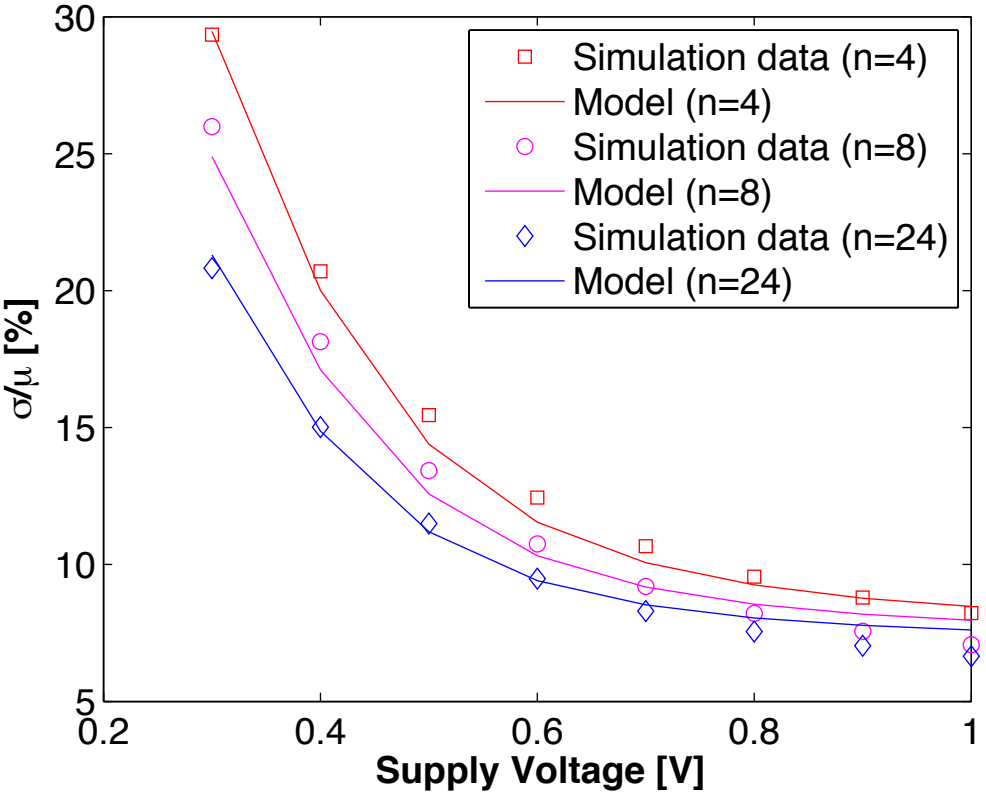


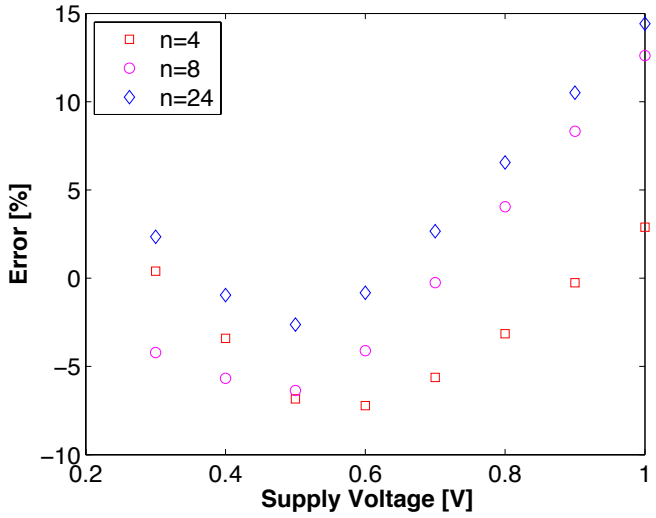


$$\frac{\sigma_{T_{d,total}}}{\mu_{T_{d,total}}} = \sqrt{\left(\frac{\sigma_{T_{d,DTD}}}{\mu_{T_{d,DTD}}}\right)^2 + \left(\frac{\sigma_{T_{d,WID}}}{\mu_{T_{d,WID}}}\right)^2}$$

- Model error <8% from 0.3V to 1V
- Local mismatch dominates at low supply voltages

Delay Variability Model with Different Logic Depths





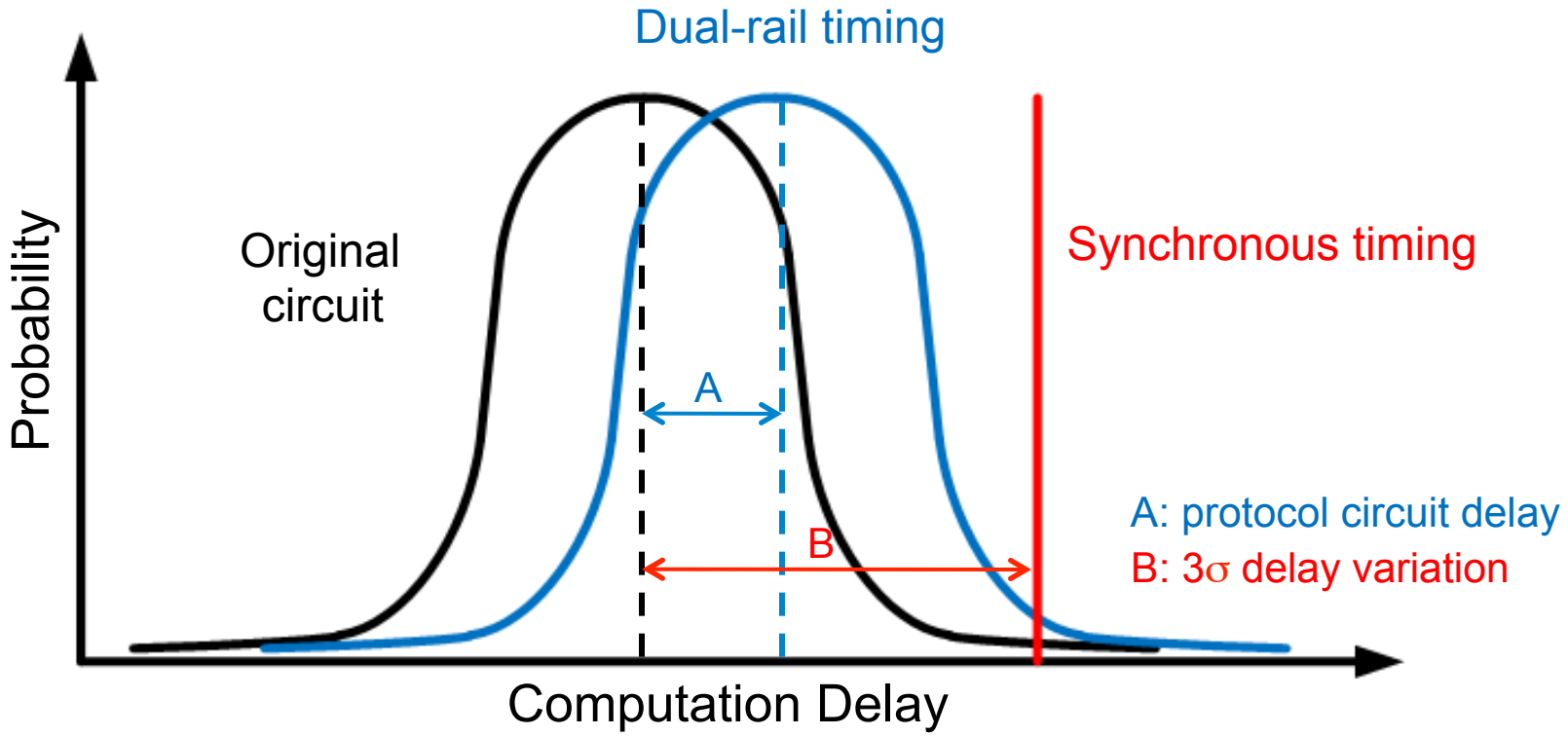
$$\frac{\sigma_{T_{d,total\_n}}}{\mu_{T_{d,total\_n}}} = \sqrt{\left(\frac{\sigma_{T_{d,DTD\_4}}}{\mu_{T_{d,DTD\_4}}}\right)^2 + \left(\frac{4}{n}\right)\cdot\left(\frac{\sigma_{T_{d,WID\_4}}}{\mu_{T_{d,WID\_4}}}\right)^2}$$

- Use 4-stage inverter chain model as baseline model
- Model error <13% for n=8 and <15% for n=24

Outline

- Motivation
- Variability model of CMOS digital circuit
- Performance model for different timing schemes
- Performance comparison
- Conclusion

Delay Overhead Evaluation

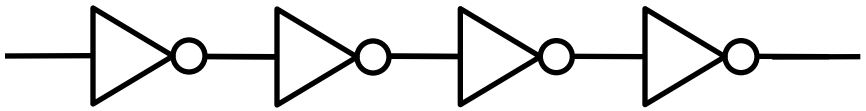


- Assumption: Process variation follows Gaussian distribution
- Dual-rail approach: have only protocol overhead but no delay overhead
- Synchronous approach: have only delay overhead

For 99.7% yield: $D_{sync} = \frac{3\sigma_{logic,total}}{\mu_{logic,total}}$

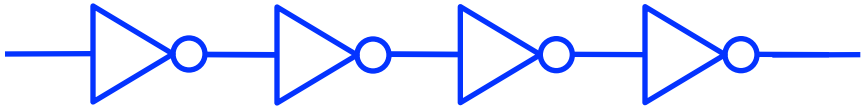
Bundled-Data Self-Timed Approach

Main data path

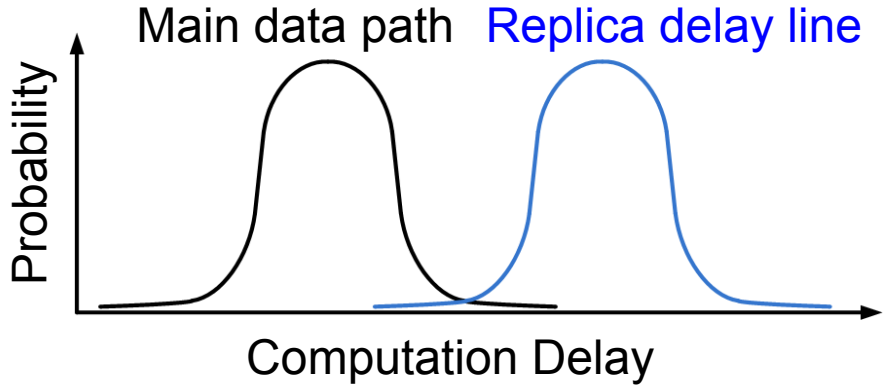


$$f_{logic}(t) = N\left(\mu_{logic}, \sigma^2_{logic}\right)$$

Replica delay line



$$f_{delay-line} = N(\mu_{delay-line}, \sigma^2_{delay-line})$$



Goal: $P\left(t_{logic} \le t_{delay-line}\right) \approx 1$

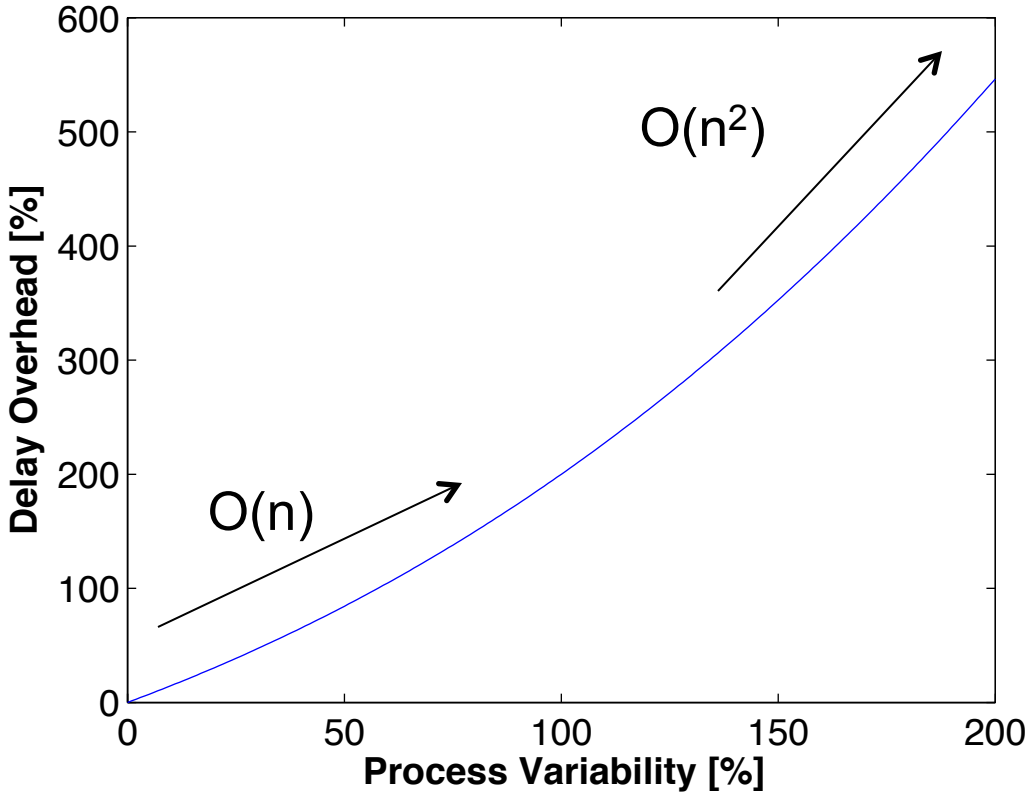
Assume main data path and replica delay line exhibit similar statistics:

For 99.7% yield: $D_{bundled-data} = D^2_{variation} \cdot \left(0.5 + \sqrt{0.25 + \frac{2}{D^2_{variation}}}\right)$

where $D_{bundled-data} = \frac{\mu_{delay-line} - \mu_{logic}}{\mu_{logic}}$ $\quad D_{variation} = \frac{3\sigma_{logic,WID}}{\mu_{logic,WID}}$

Bundled-Data Delay Overhead

$$D_{bundled-data} \propto \begin{cases} \sqrt{2} \cdot D_{variation}, & when \quad D_{variation} \to 0 \\ D_{variation}^2, & when \quad D_{variation} \to \infty \end{cases}.$$



- Delay overhead becomes much larger as process variability increases!

Performance Model under Variations

Original delay and energy model

$$T_{comp} = T_{delay}$$

$$E_{dynamic} = \alpha C_{switch} V^2$$

$$E_{leakage} = V I_{leakage} T_{delay}$$

$$E_{total} = \alpha C_{switch} V^2 + V I_{leakage} T_{delay}$$

Statistical delay and energy model

$$T_{comp} = T_{delay} (1+P+D)$$

$$E_{dynamic} = \alpha C_{switch} (1+P) V^2$$

$$E_{leakage} = V I_{leakage} (1+P) T_{delay} (1+P+D)$$

$$E_{total} = \alpha C_{switch} (1+P) V^2 + V I_{leakage} (1+P) T_{delay} (1+P+D)$$

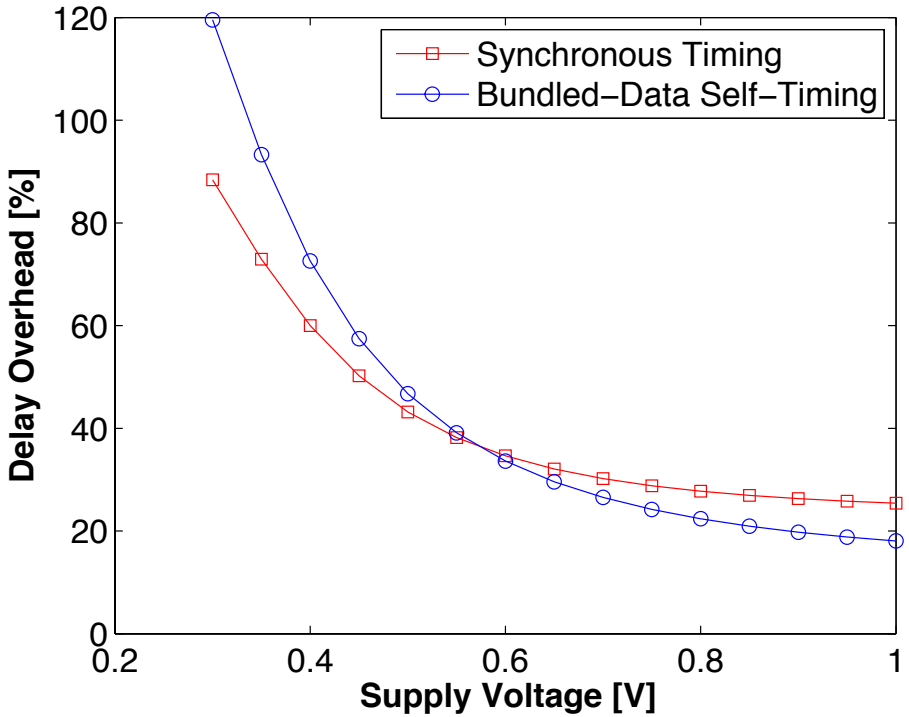
Timing scheme	Synchronous	Bundled-Data	Dual-Rail
Delay Overhead (D)	D_{sync}	$D_{bundled\text{-}data}$	0
Protocol Overhead (P)	0	$P_{bundled\text{-}data}$	$P_{dual\text{-}rail}$

- Evaluate computation delay and energy under variations
- Overhead changes with supply voltage and logic depth

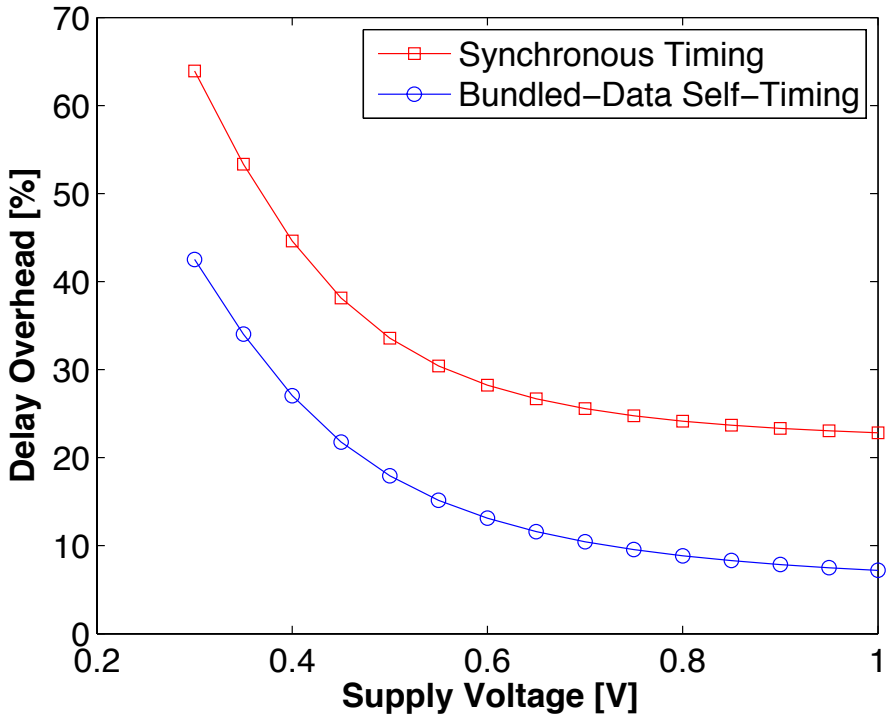
Outline

- Motivation
- Variability model of CMOS digital circuit
- Performance model for different timing schemes
- Performance comparison
- Conclusion

Delay Overhead Comparison



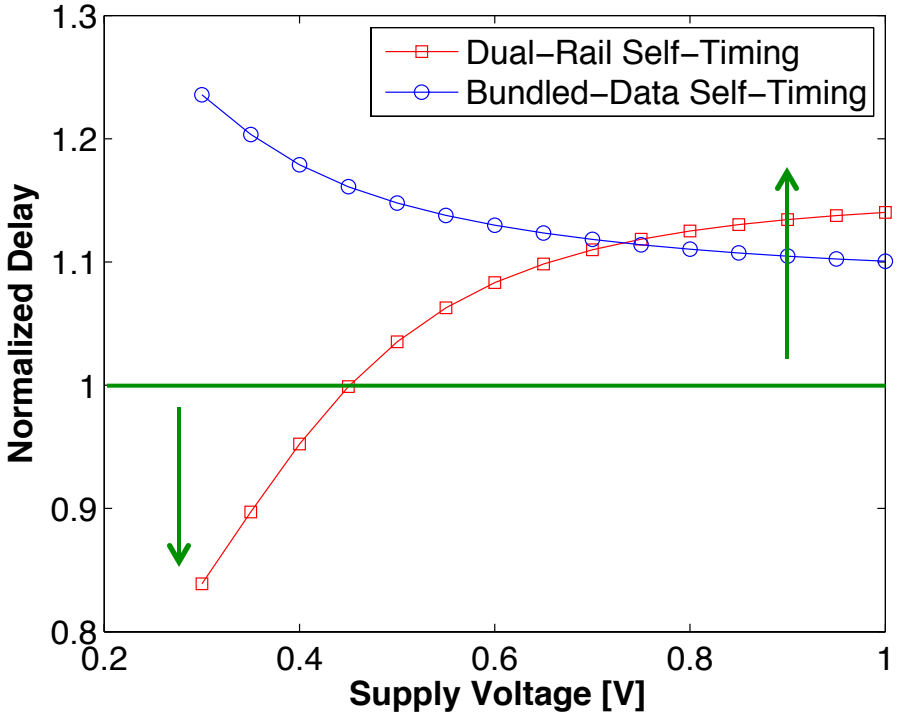
4-stage FO4 INV chain



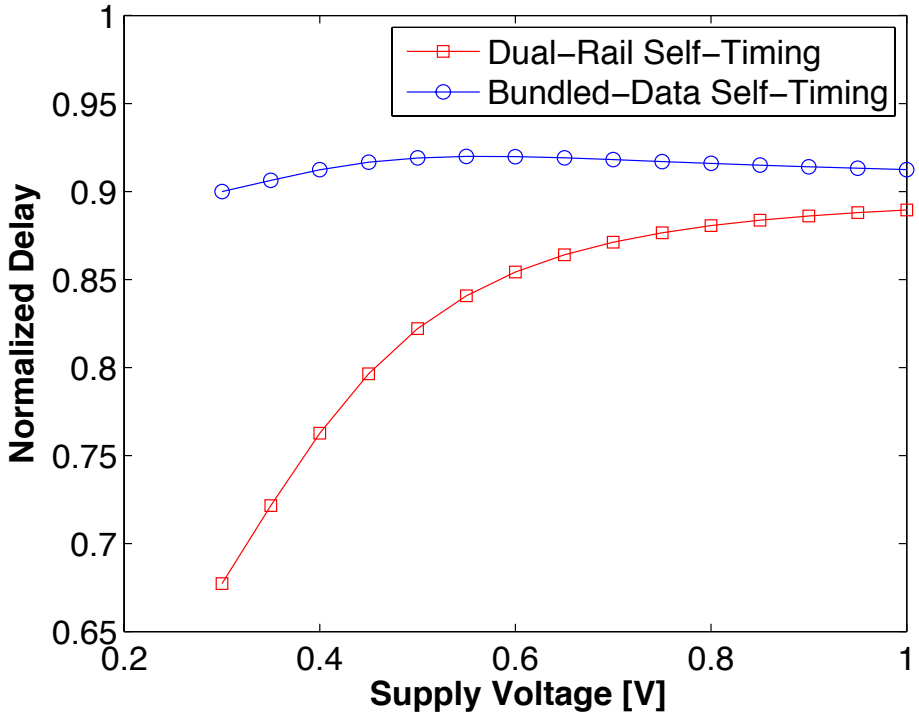
24-stage FO4 INV chain

- Global variation affects only synchronous approach
- Local mismatch dominates at low supply voltages
- Local mismatch has less impact on longer critical path

Speed Performance Comparison



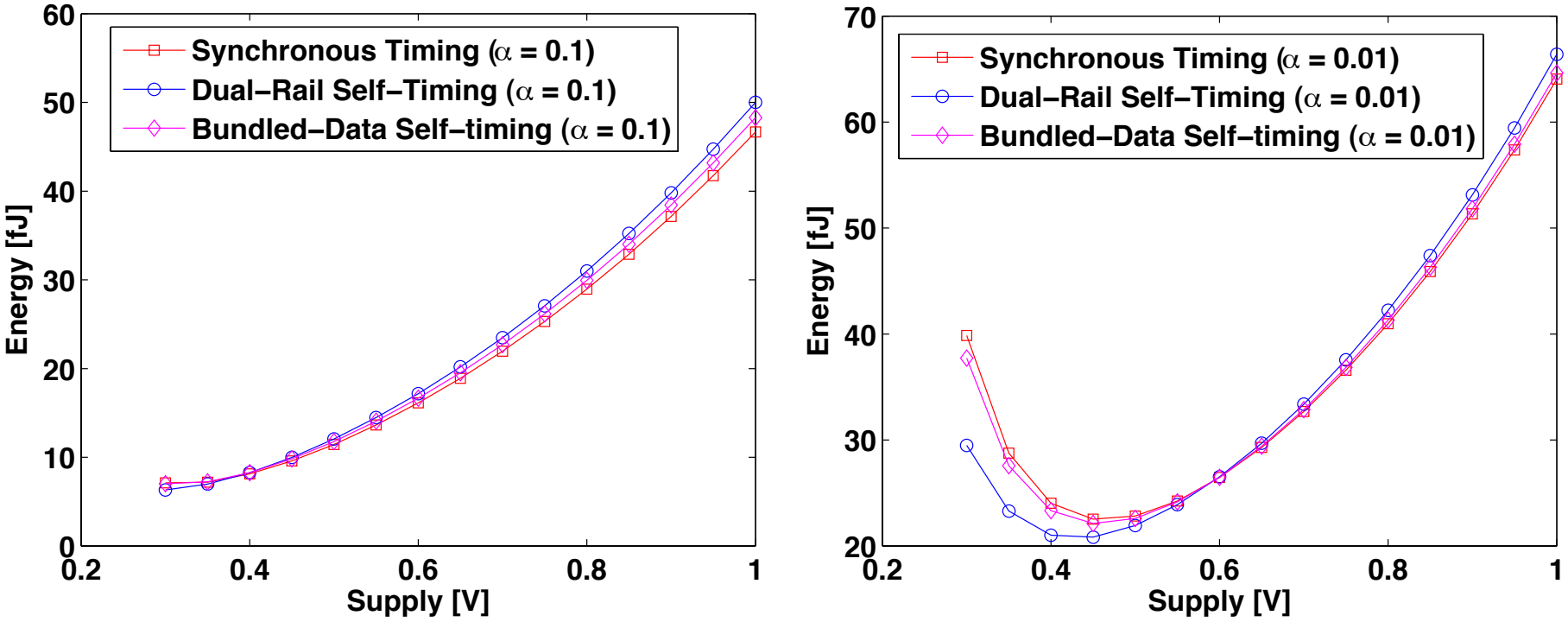
4-stage FO4 INV chain



24-stage FO4 INV chain

- Assumption: $P_{bundled\text{-}data} = 1T_{FO4}$; $P_{dual\text{-}rail} = 2T_{FO4}$
- Synchronous scheme is better for small critical path at high supply voltages
- Dual-rail scheme is better for large critical path at low supply voltages

Energy Performance Comparison



24-stage FO4 INV chain

- Synchronous scheme is better for high activity at high supply voltages
- Dual-rail scheme is better for low activity at low supply voltages
- Leakage dominates for low activity at low supply voltages

Conclusion

- A statistical analysis framework is proposed to evaluate performance of CMOS digital circuit in the presence of process variations.
- Designer can efficiently determine the optimal timing strategy, pipeline depth and supply voltage based on the proposed variability and statistical performance models.
- Asynchronous design exhibits better energy and delay characteristics for circuits with low activity and larger critical path delay under process variations

Acknowledgement

- Berkeley Wireless Research Center
- NSF Infrastructure Grant
- STMicroelectronics
- Multiscale System Center

Thank you!