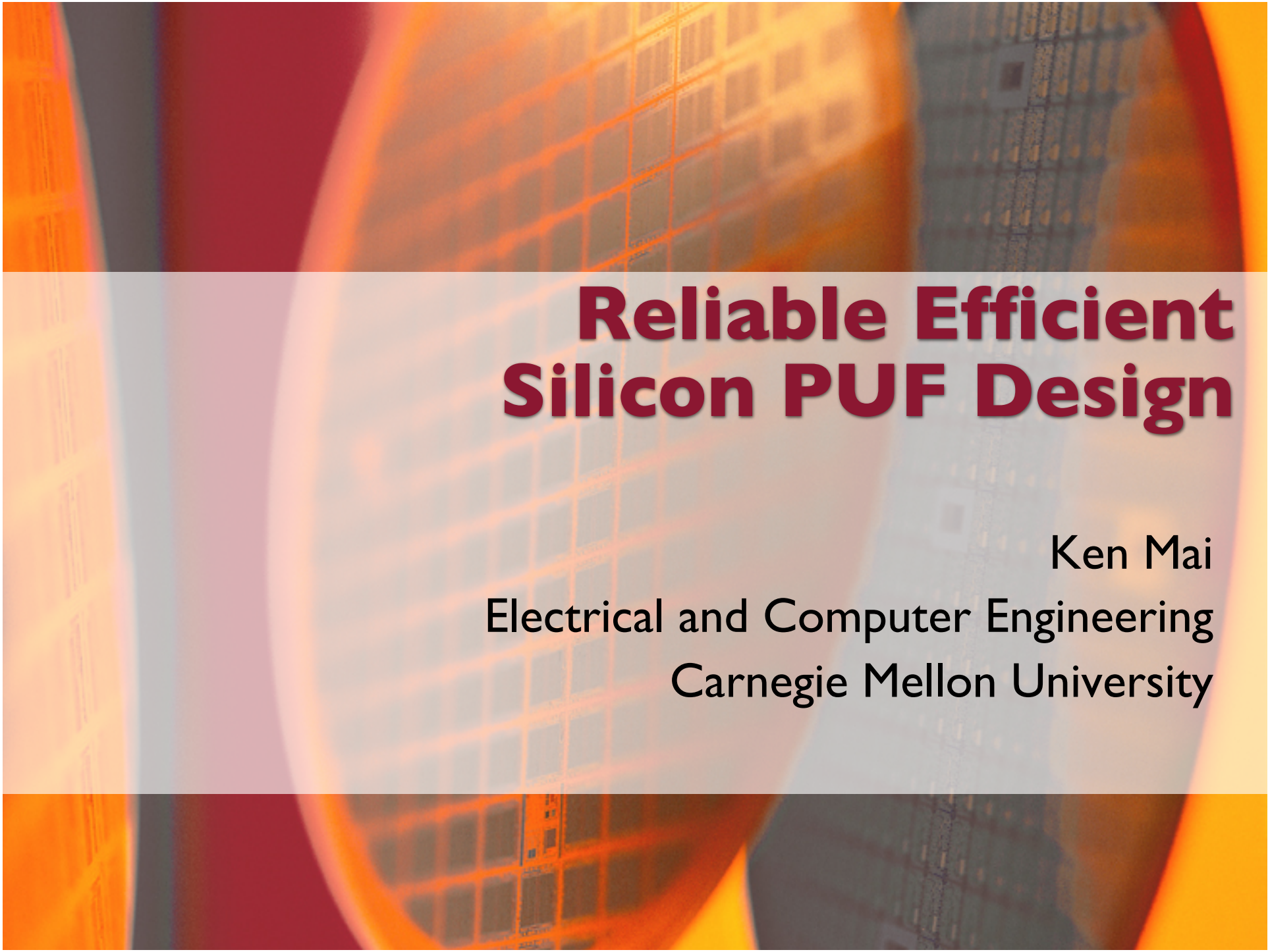
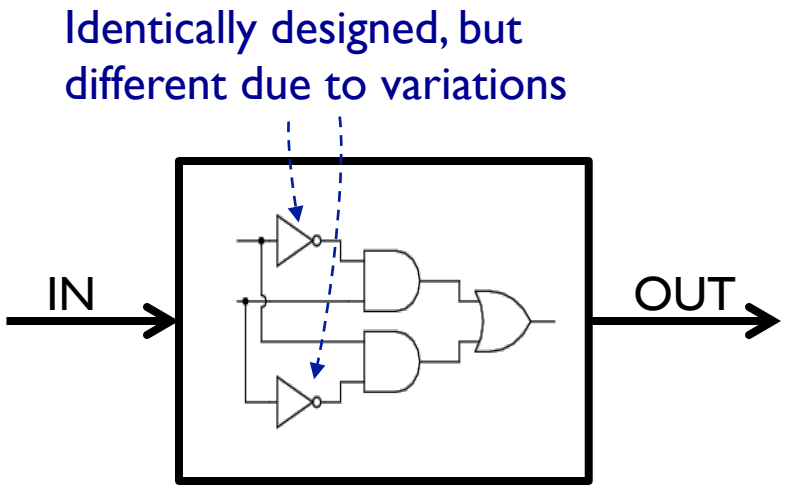


Reliable Efficient Silicon PUF Design

Ken Mai
Electrical and Computer Engineering
Carnegie Mellon University

PUF Design vs. Typical Digital Design

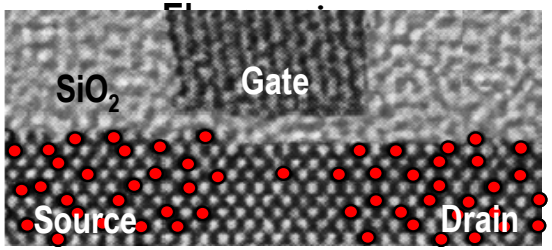


	Digital Design	PUF Design
OUT depends on device variations?		
OUT is a predictable function of IN?		

- PUFs amplify electrical differences in two nominally identical circuits
- The difference is NOT inherent in the design
  - Due to random process variations

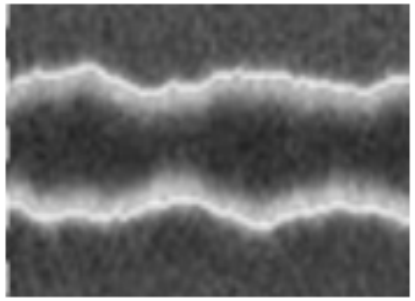
Sources of Variability

Random Dopant

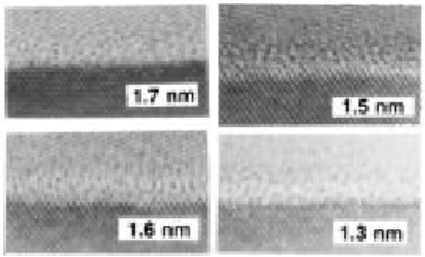


A. Brown *et al.*, *IEEE Trans. Nanotechnology*, p. 195, 2002

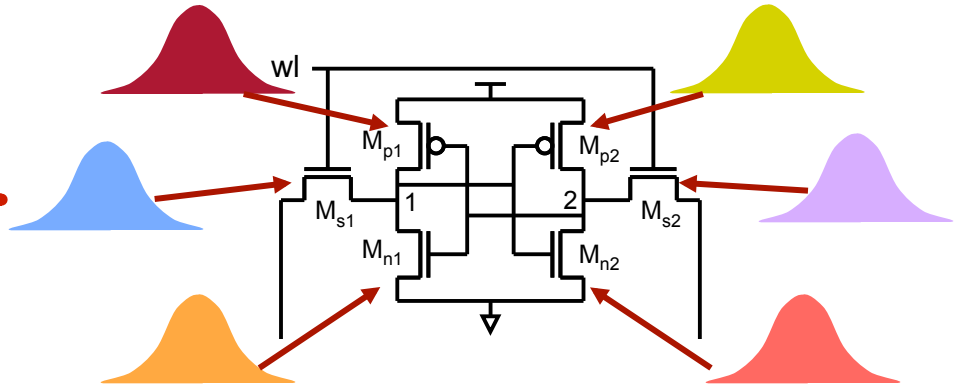
Line Edge Roughness



Gate Oxide Variation

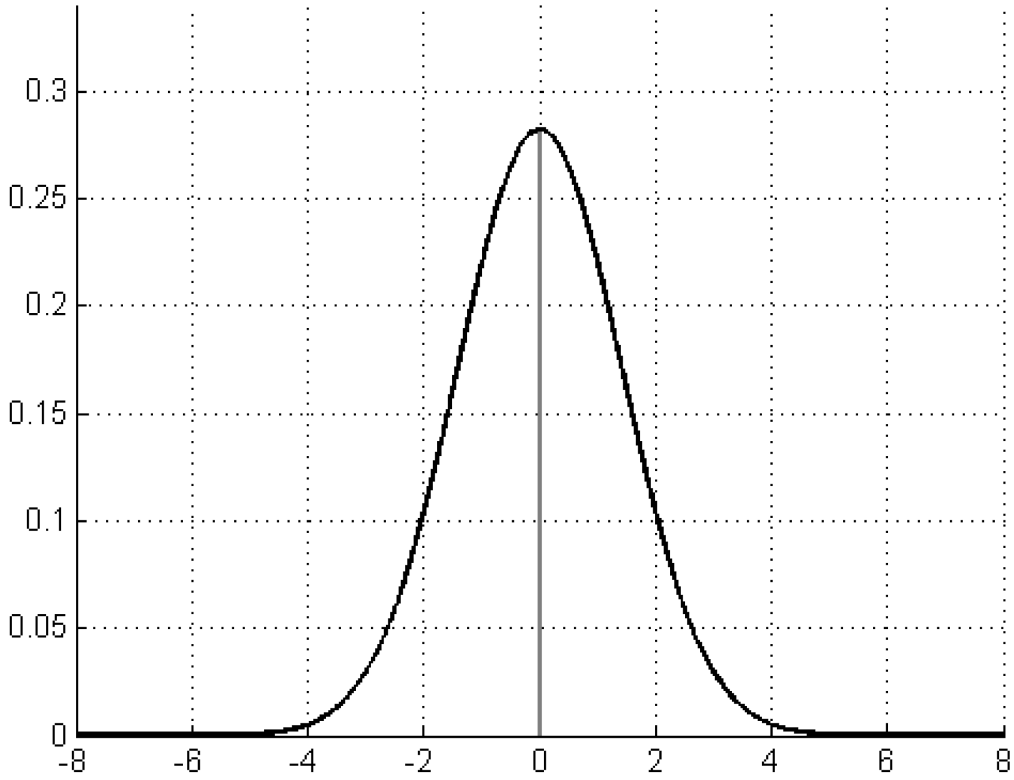


Momose et al, *IEEE Trans. Electron Devices*, 45(3), 1998

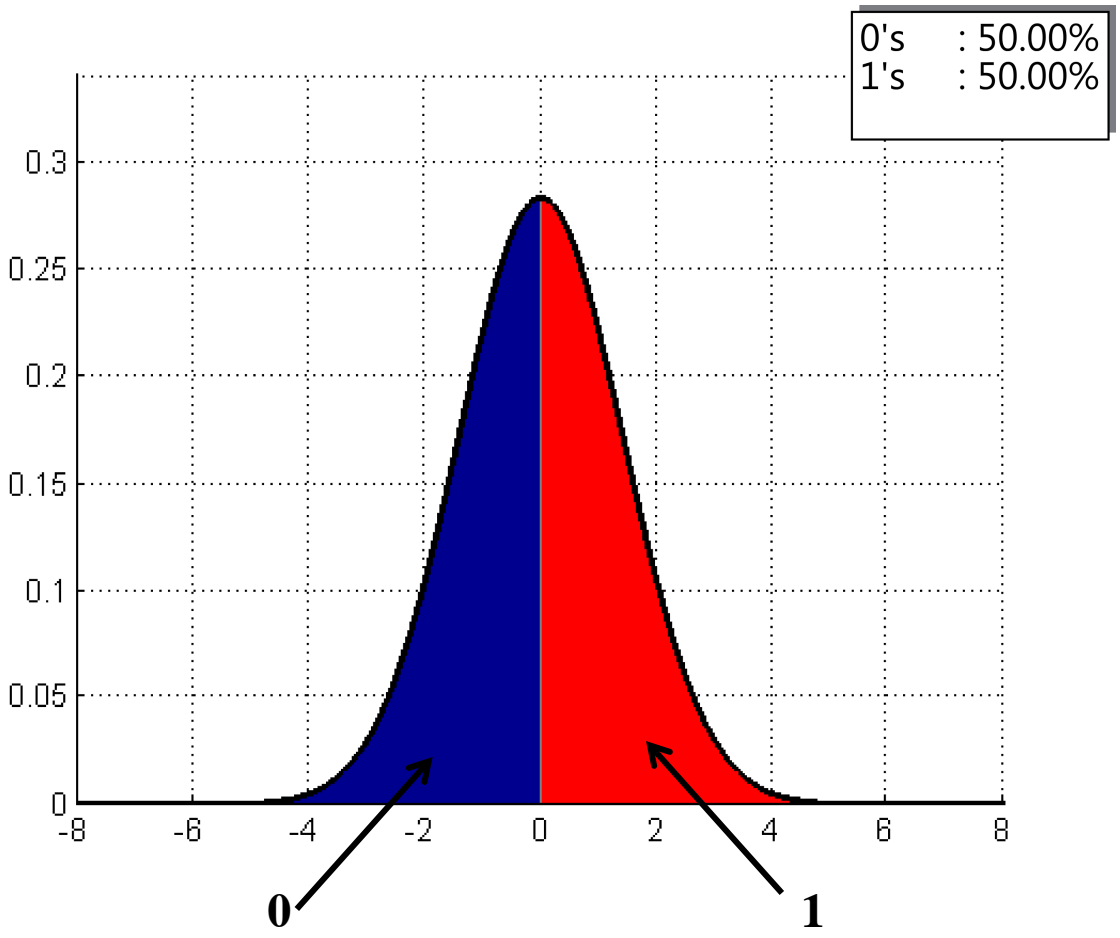


- Nothing is deterministic anymore
- Everything is statistical

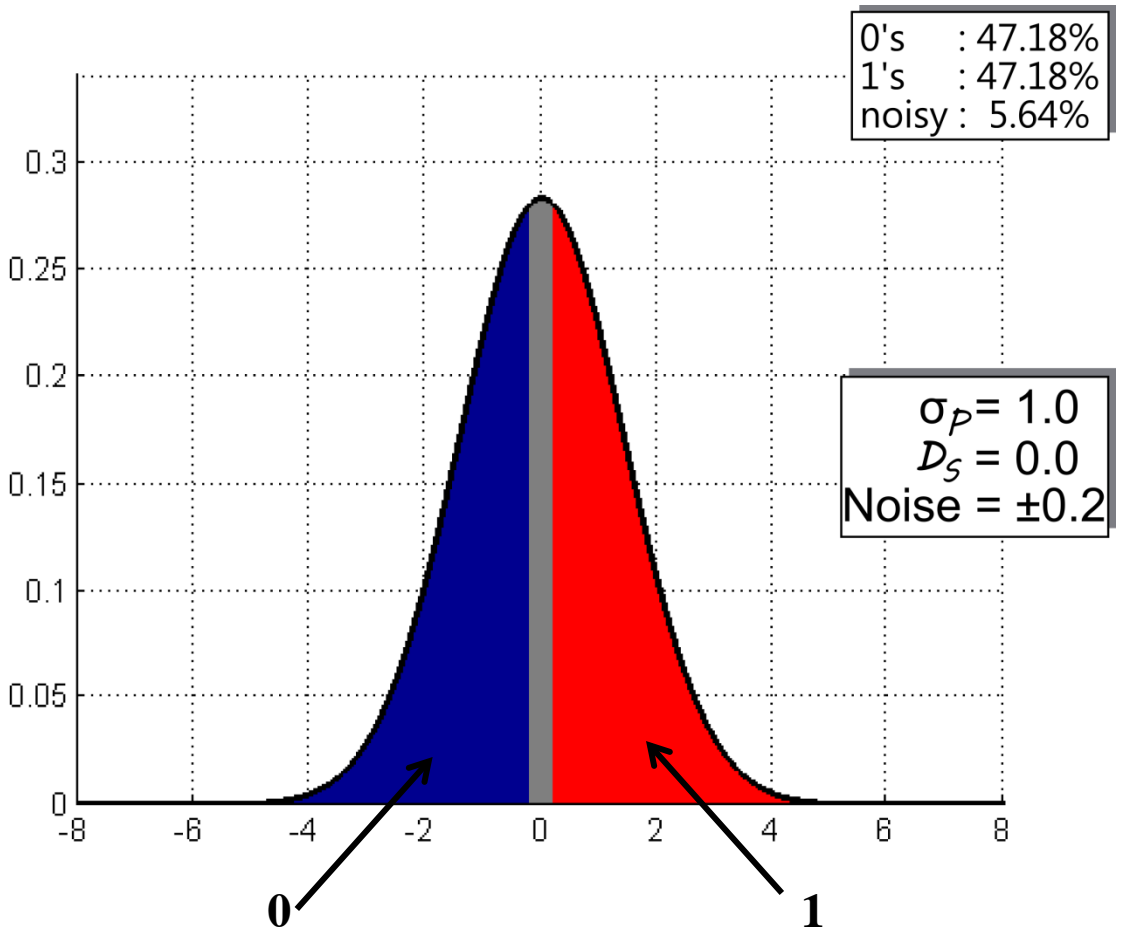
PUF Design Concept



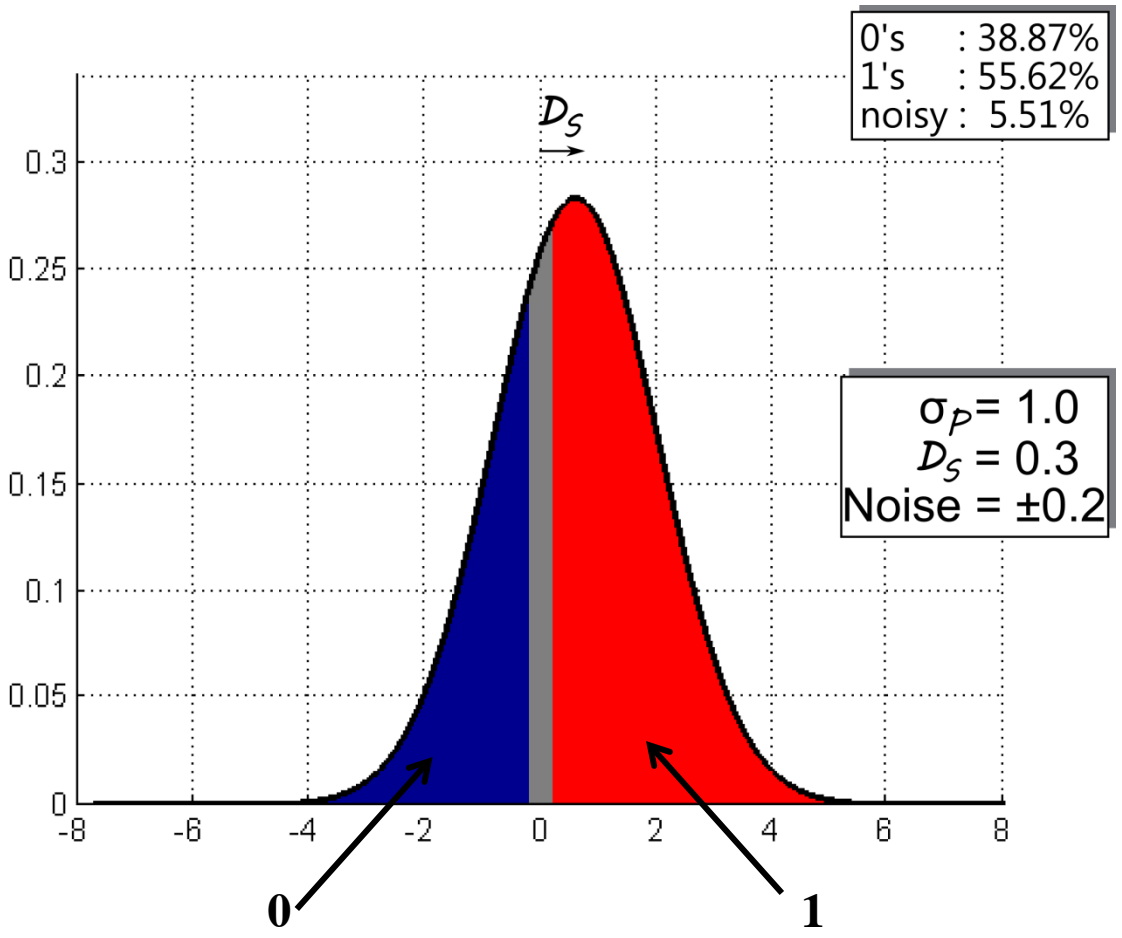
PUF Design Concept



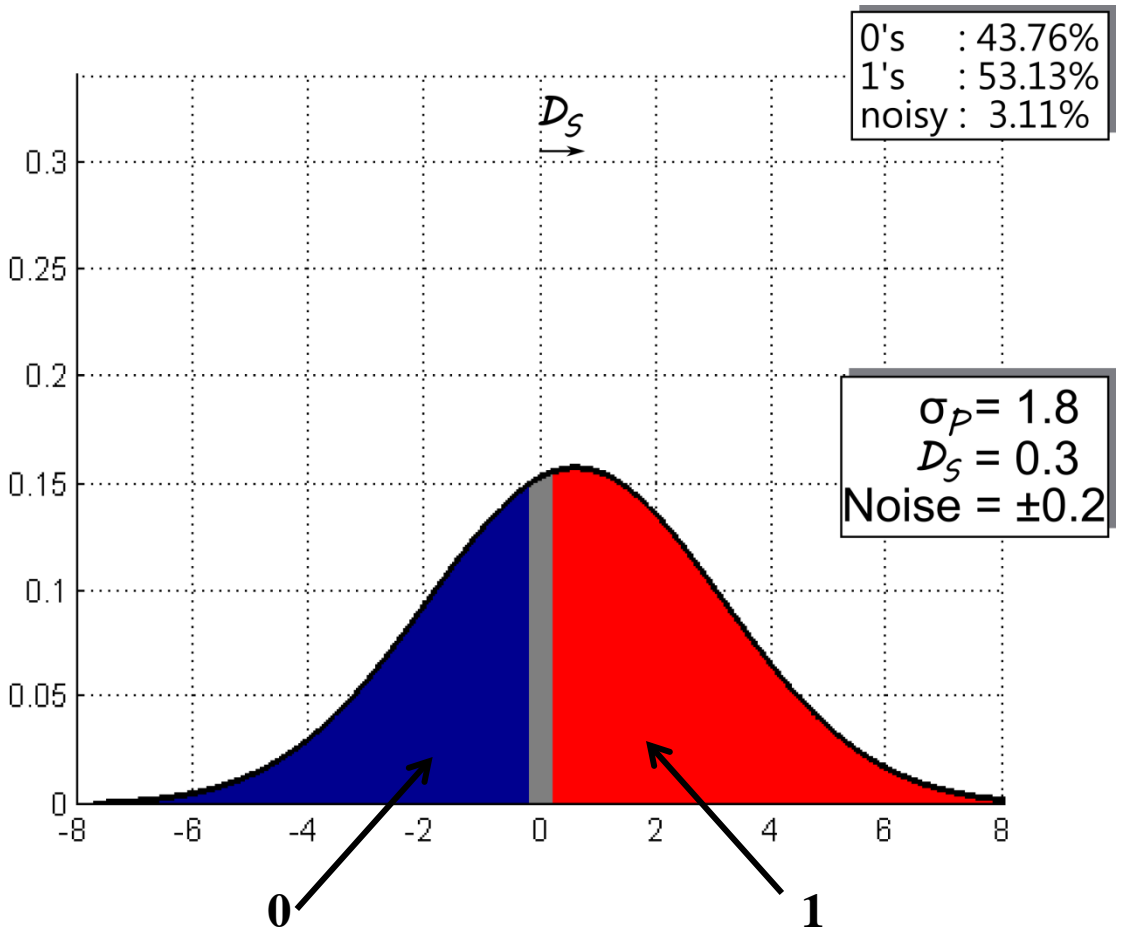
PUF Design Concept



PUF Design Concept

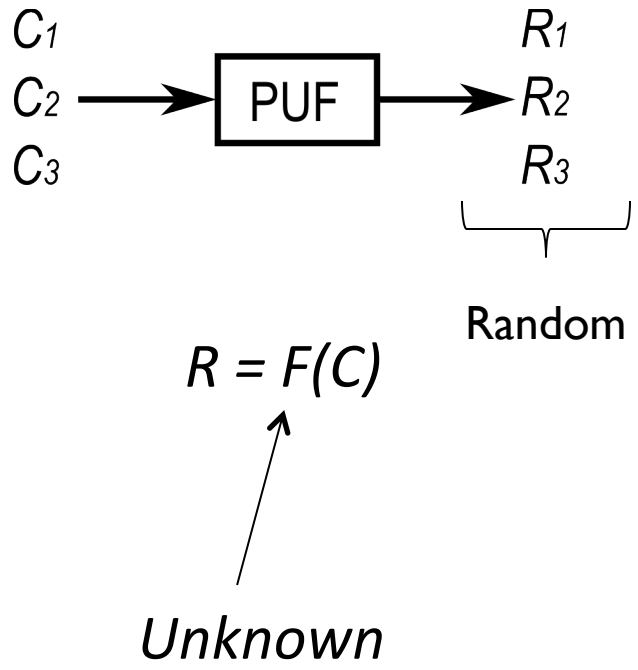


PUF Design Concept



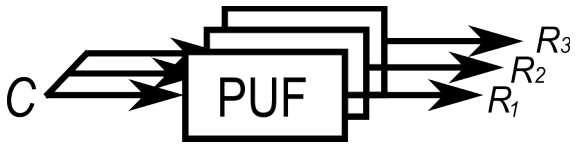
PUF Design Goals

- Randomness
  - PUF cannot be modeled
- Uniqueness
  - Across different dies of same design
- Reliability
  - Across environmental variations, ambient noise, aging



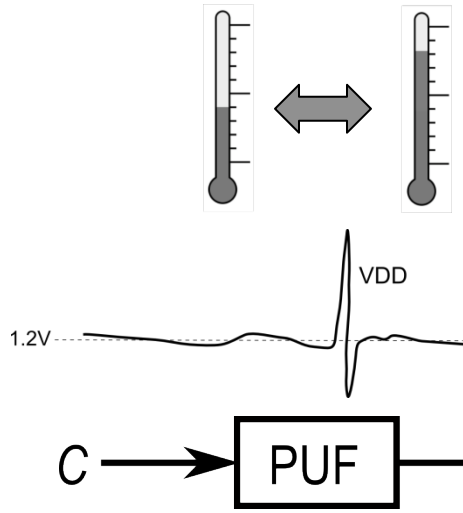
PUF Design Goals

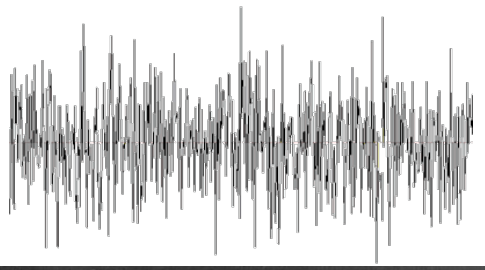
- Randomness
  - PUF cannot be modeled
- Uniqueness
  - Across different dies of same design
- Reliability
  - Across environmental variations, ambient noise, aging

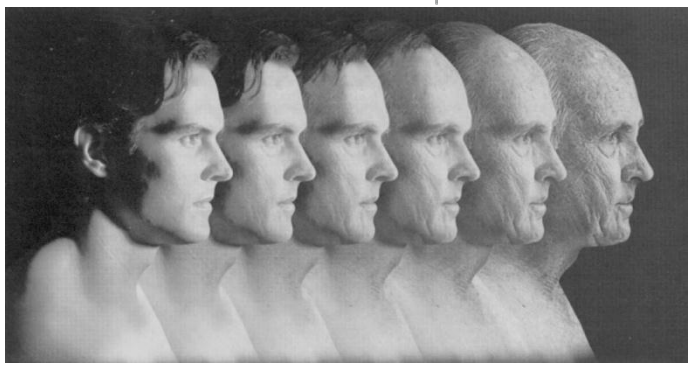


PUF Design Goals

- Randomness
  - PUF cannot be modeled
- Uniqueness
  - Across different dies of same design
- Reliability
  - Across environmental variations, ambient noise, aging







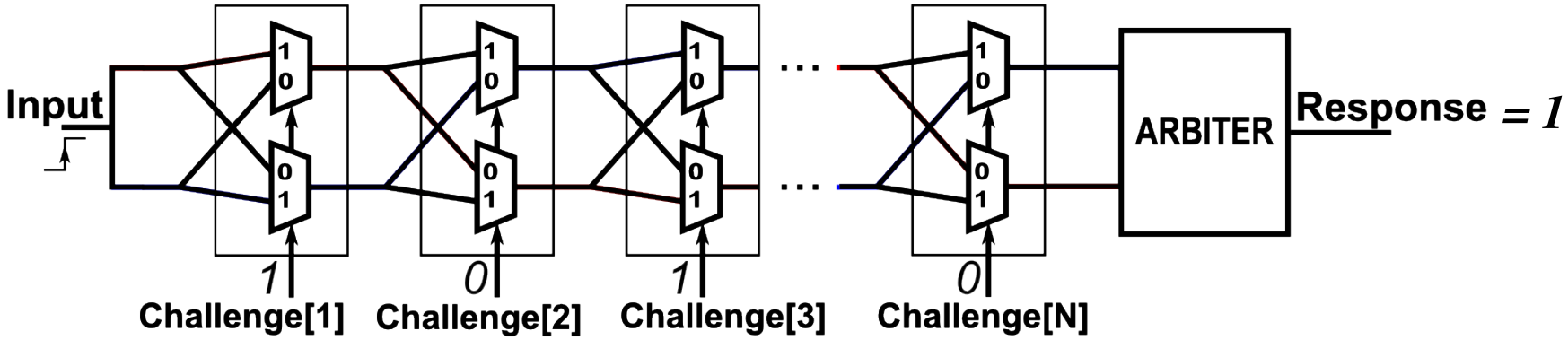
PUF Implementations

- Delay based PUF
  - Arbiter PUF[1] (2002)
  - Ring oscillator PUF[2] (2007)

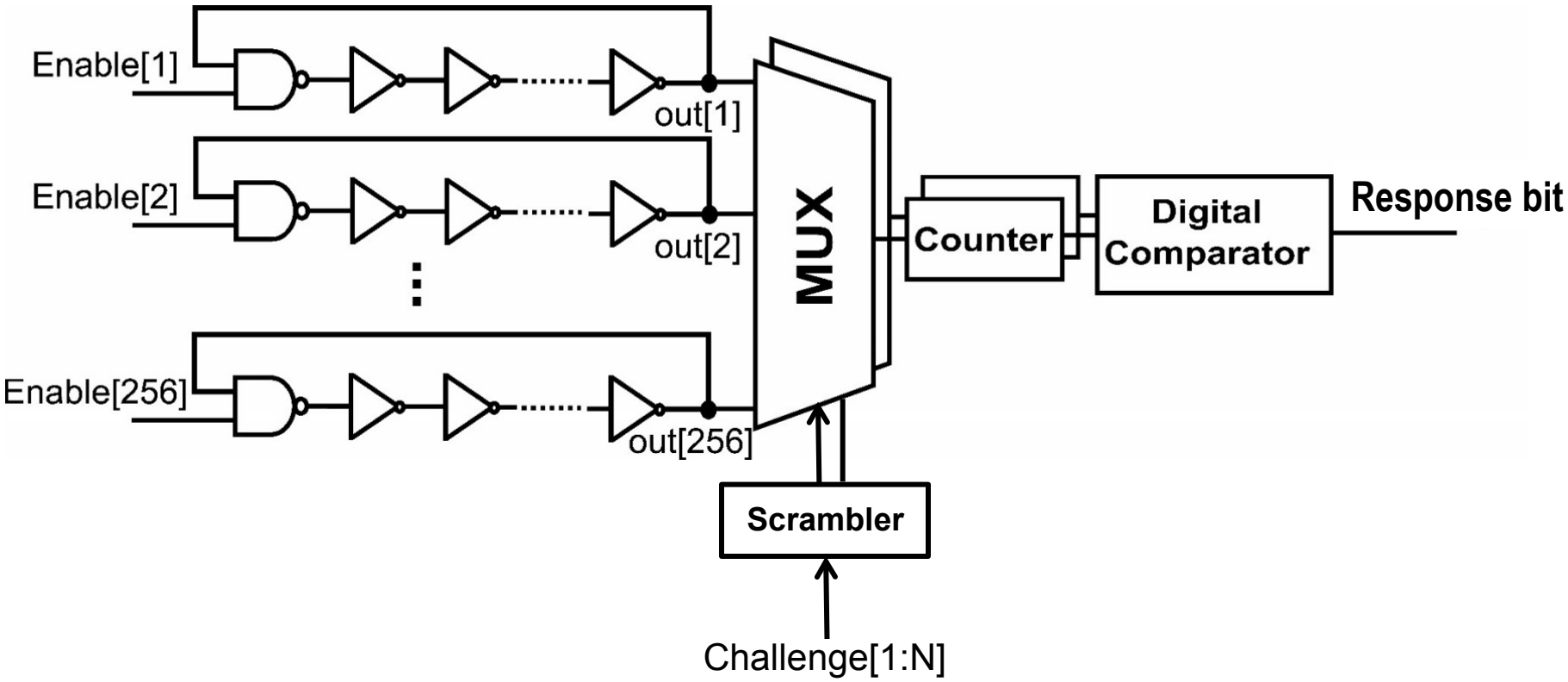
- Bi-stable element based PUF
  - SRAM power-up state PUF[3] (2008)
  - Sense amplifier PUF[4] (2010)

1. B. Gassend et al., *Concurrency and Computation: Practice and Experience*, 2004.
2. G. E. Suh et al., ACM/IEEE Design Automation Conference, 2007.
3. Guajardo et al., CHES 2007 | D. E. Holcomb et al., RFIDSec 2007
4. M. Bhargava et al., IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), 2010.

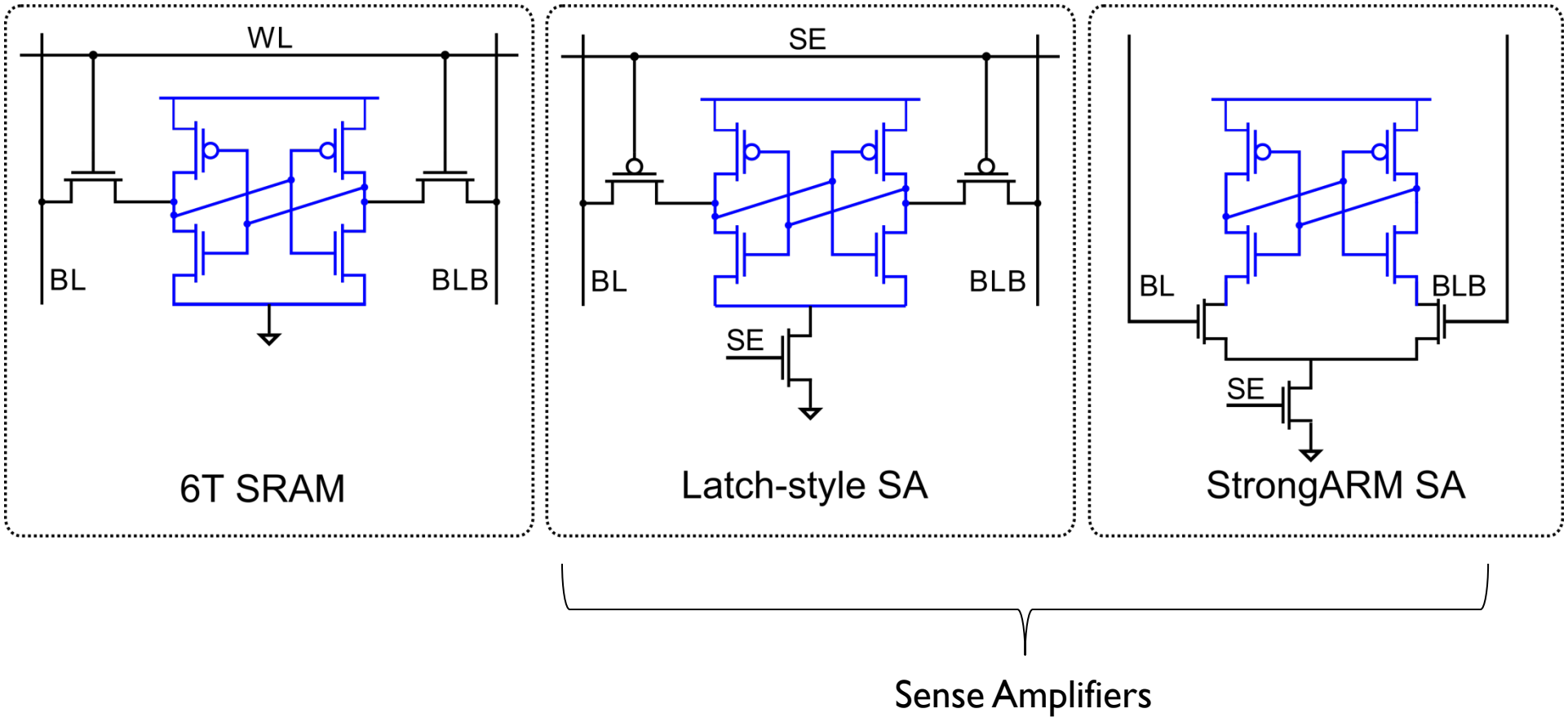
Arbiter PUF


Input
1
0
1
0
1
0
1
0
1
0
1
0
1
0
1
0
ARBITER
Response = 1
1
0
1
0
Challenge[1]
Challenge[2]
Challenge[3]
Challenge[N]

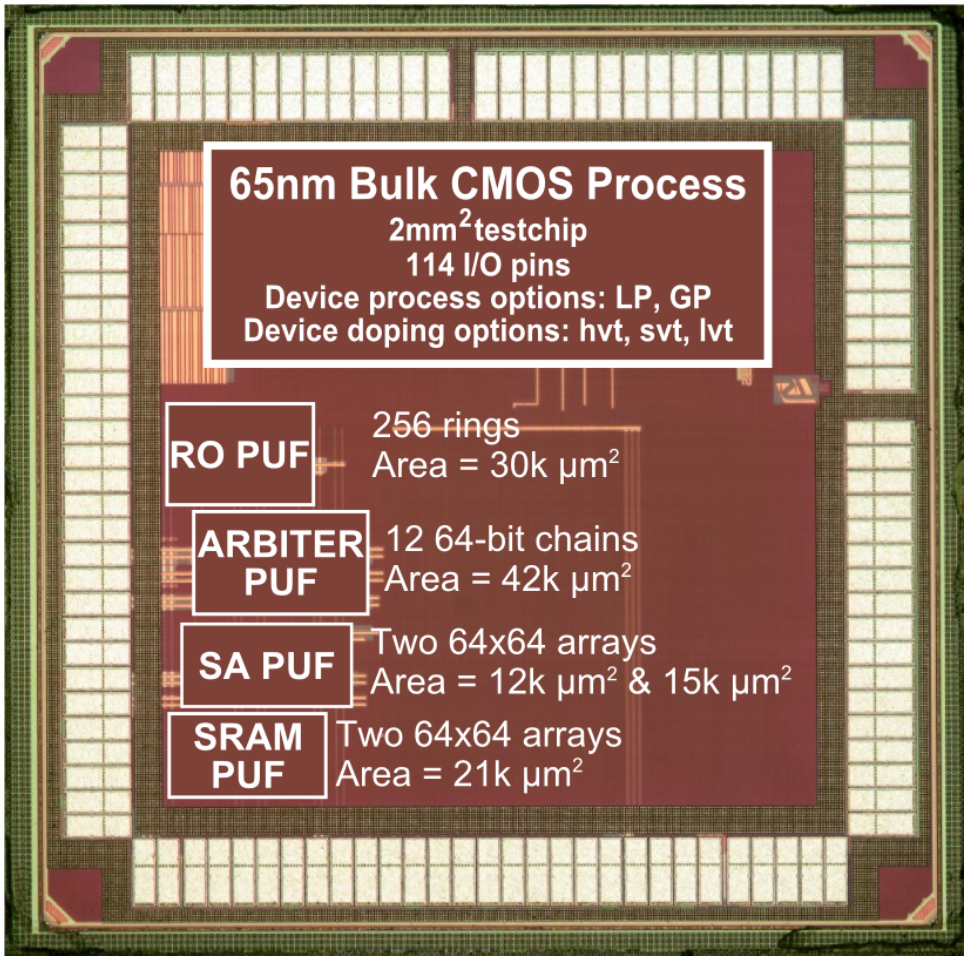
Ring Oscillator PUF



Bi-Stable PUFs

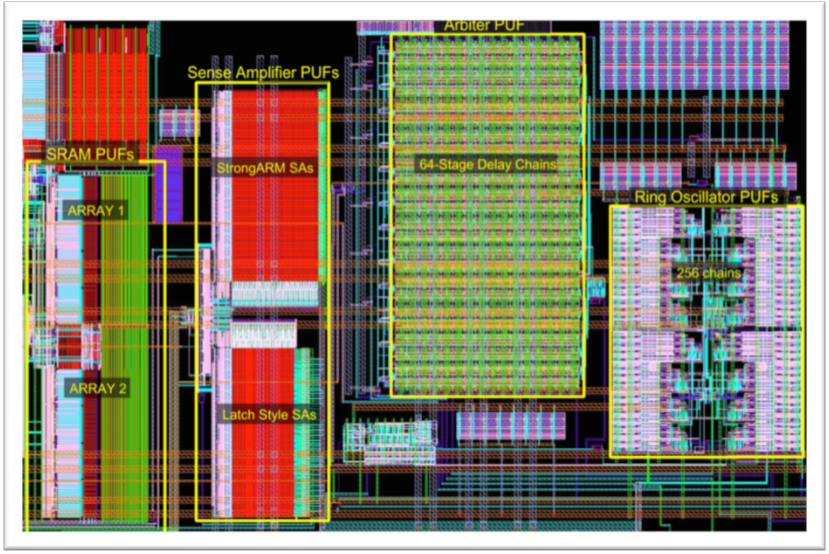


PUF Comparison Testchip



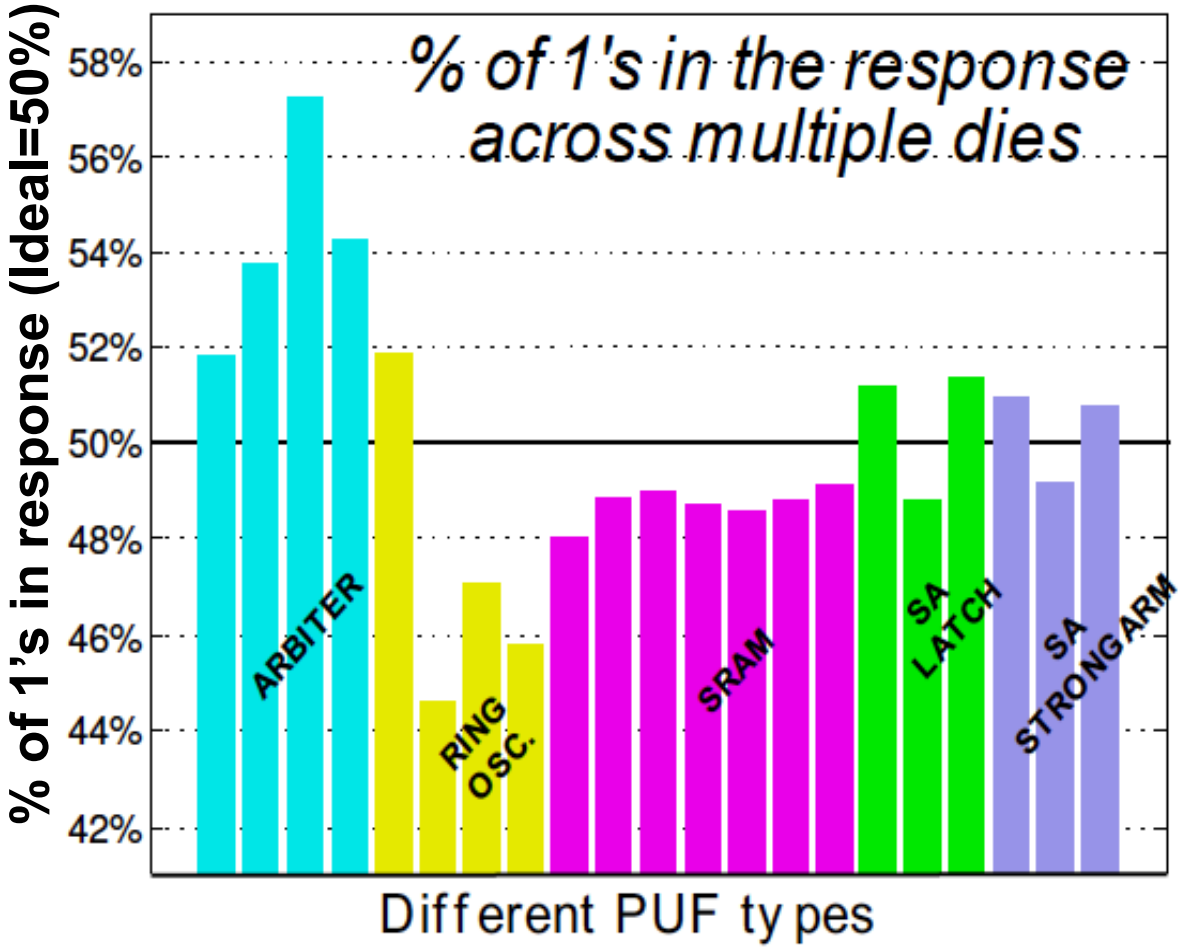
4 PUF implementations

- Arbiter
- Ring oscillators
- SRAM
- Sense amplifier

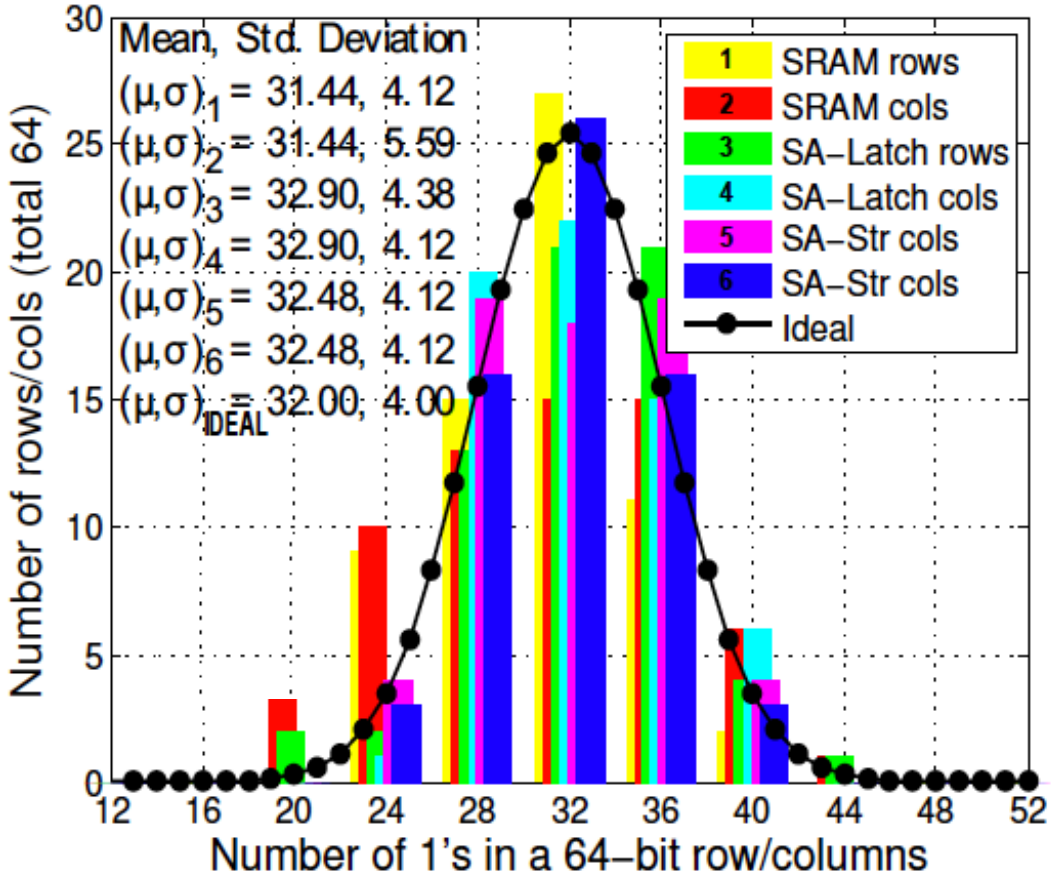


[Bhargava CICC 2012]

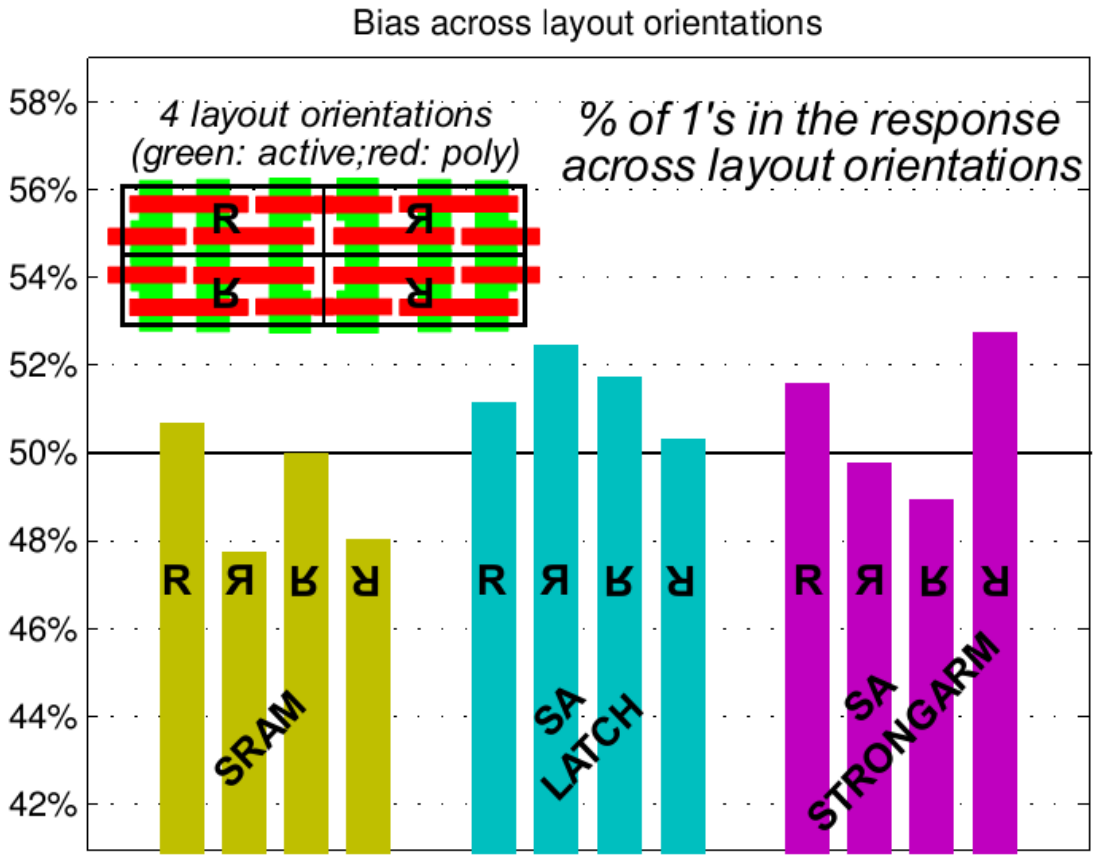
Comparison: Randomness



Comparison: Randomness



Comparison: Randomness



Comparison: Randomness

PUF Types ⇒		SRAM				SA-S				SA-L			
Total Samples (N) ⇒		28672				8192				8192			
		← # Sequences (m) →				← # Sequences (m) →				← # Sequences (m) →			
Test	n_{min}	m_{total}	m_{min}	m_{pass}	*pass?*	m_{total}	m_{min}	m_{pass}	*pass?*	m_{total}	m_{min}	m_{pass}	*pass?*
Freq	100	286	278	284	✔	81	77	77	✔	81	77	77	✔
Block Freq	100	286	278	282	✔	81	77	81	✔	81	77	80	✔
Cusums (f)	100	286	278	286	✔	81	77	79	✔	81	77	77	✔
Cusums (b)	100	286	278	286	✔	81	77	78	✔	81	77	79	✔
Runs	100	286	278	284	✔	81	77	80	✔	81	77	81	✔
Entropy	100	286	278	286	✔	81	77	80	✔	81	77	80	✔
Longest Run	128	224	217	222	✔	64	60	62	✔	64	60	64	✔
Spectral	1000	28	26	27	✔	8	7	8	✔	8	7	8	✔

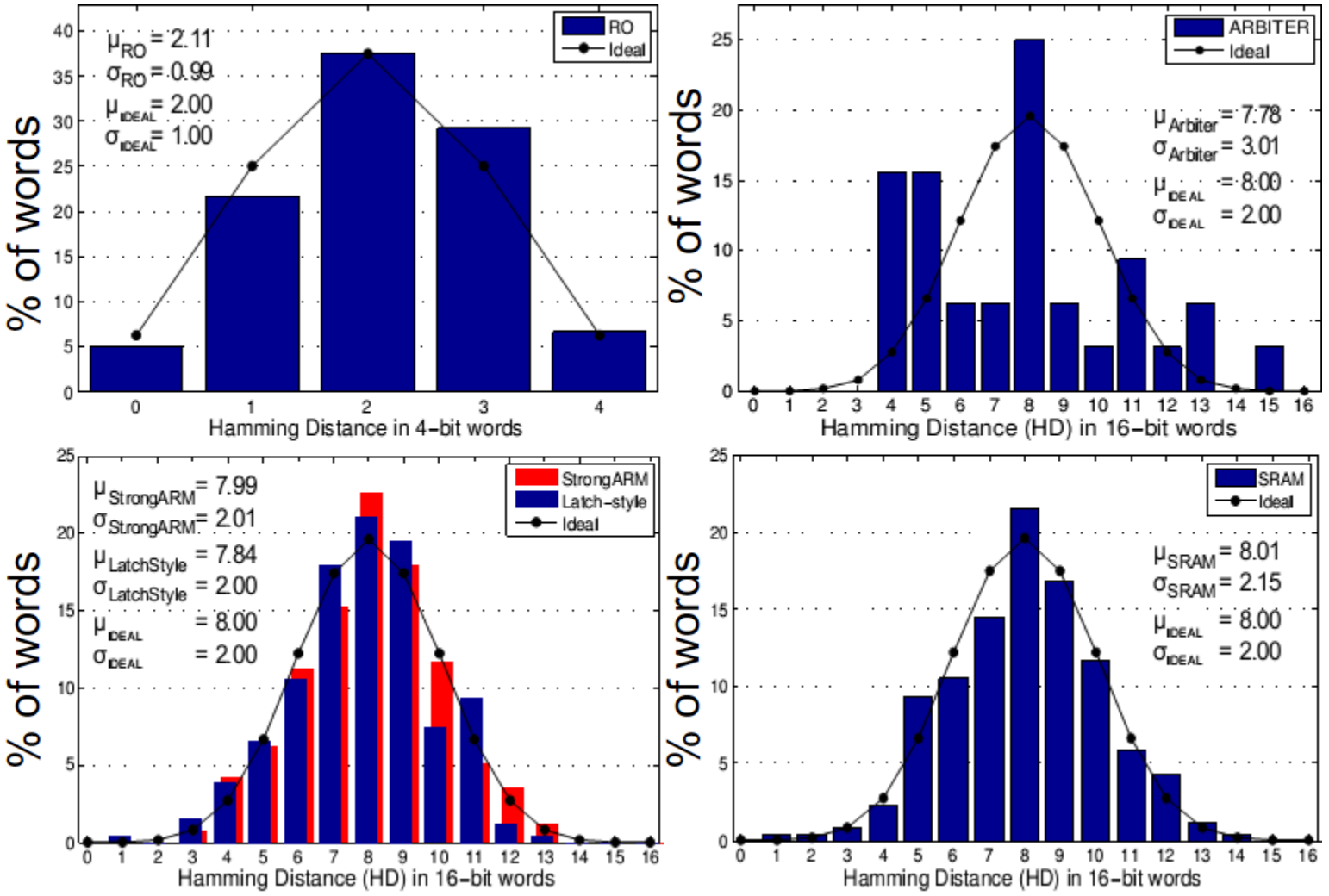
NIST tests on bi-stable PUFs

Comparison: Randomness

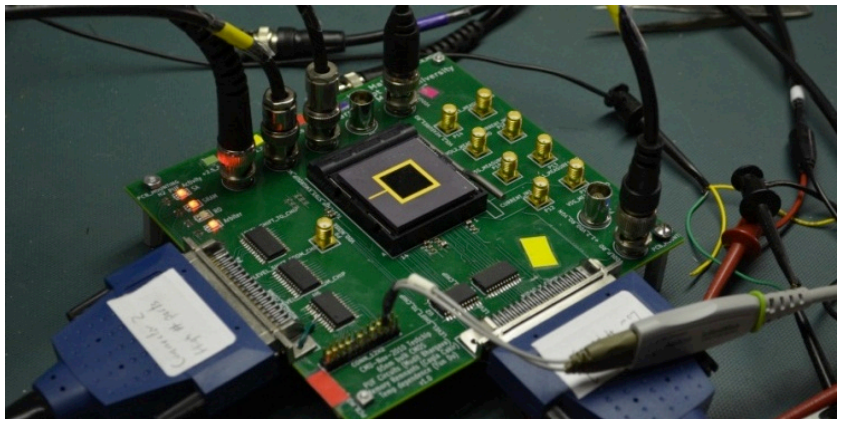
PUF Types ⇒		Arbiter				Ring Oscillator			
Total Samples (N) ⇒		4704				11873			
		← # Sequences (m) →				← # Sequences (m) →			
Test	n_{min}	m_{total}	m_{min}	m_{pass}	$pass?$	m_{total}	m_{min}	m_{pass}	$pass?$
Freq	100	47	44	46	✔	118	113	118	✔
Block Freq	100	47	44	47	✔	118	113	118	✔
Cusums (f)	100	47	44	46	✔	118	113	118	✔
Cusums (b)	100	47	44	47	✔	118	113	118	✔
Runs	100	47	44	47	✔	118	113	116	✔
Entropy	100	47	44	47	✔	118	113	118	✔
Longest Run	128	36	33	36	✔	92	88	91	✔
Spectral	1000	4	3	4	✔	11	9	11	✔

NIST tests on delay based PUFs

Comparison: Uniqueness

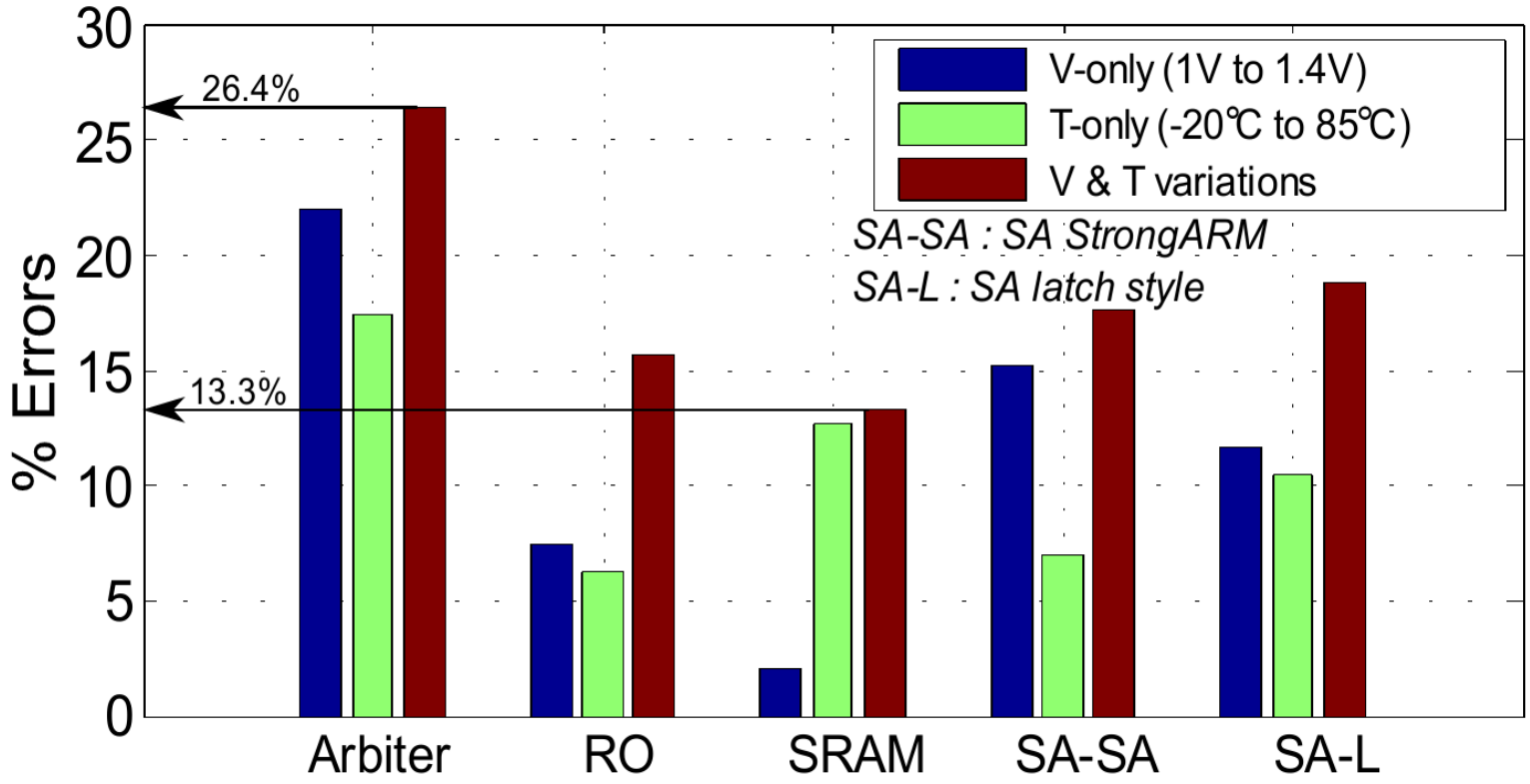


Reliability Measurement



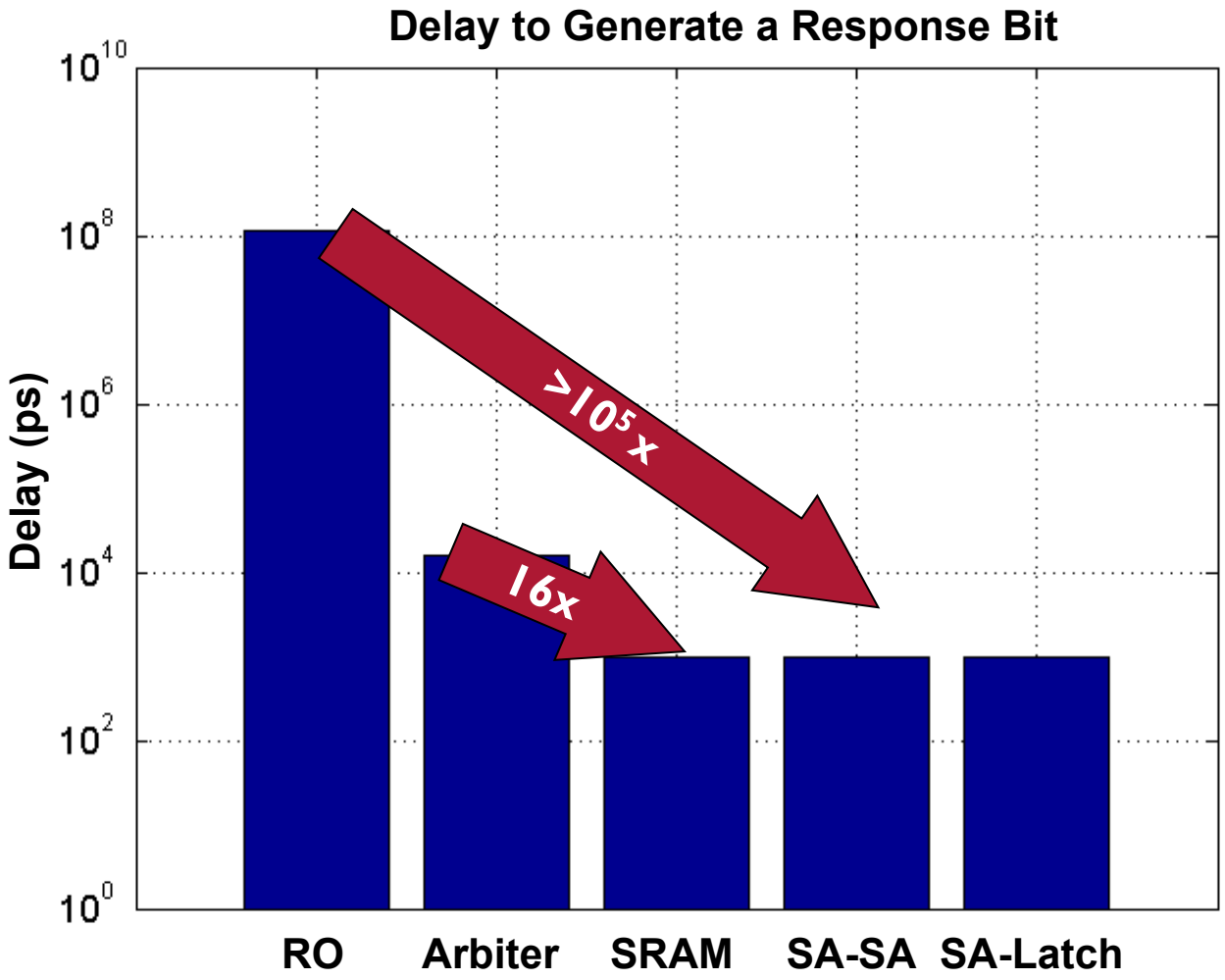
- Chips and board placed in temperature controlled chamber
- -20°C to 85°C
- 1.0V to 1.4V (1.2V nominal)
- Any response bit that flips is marked as erroneous

Comparison: Reliability



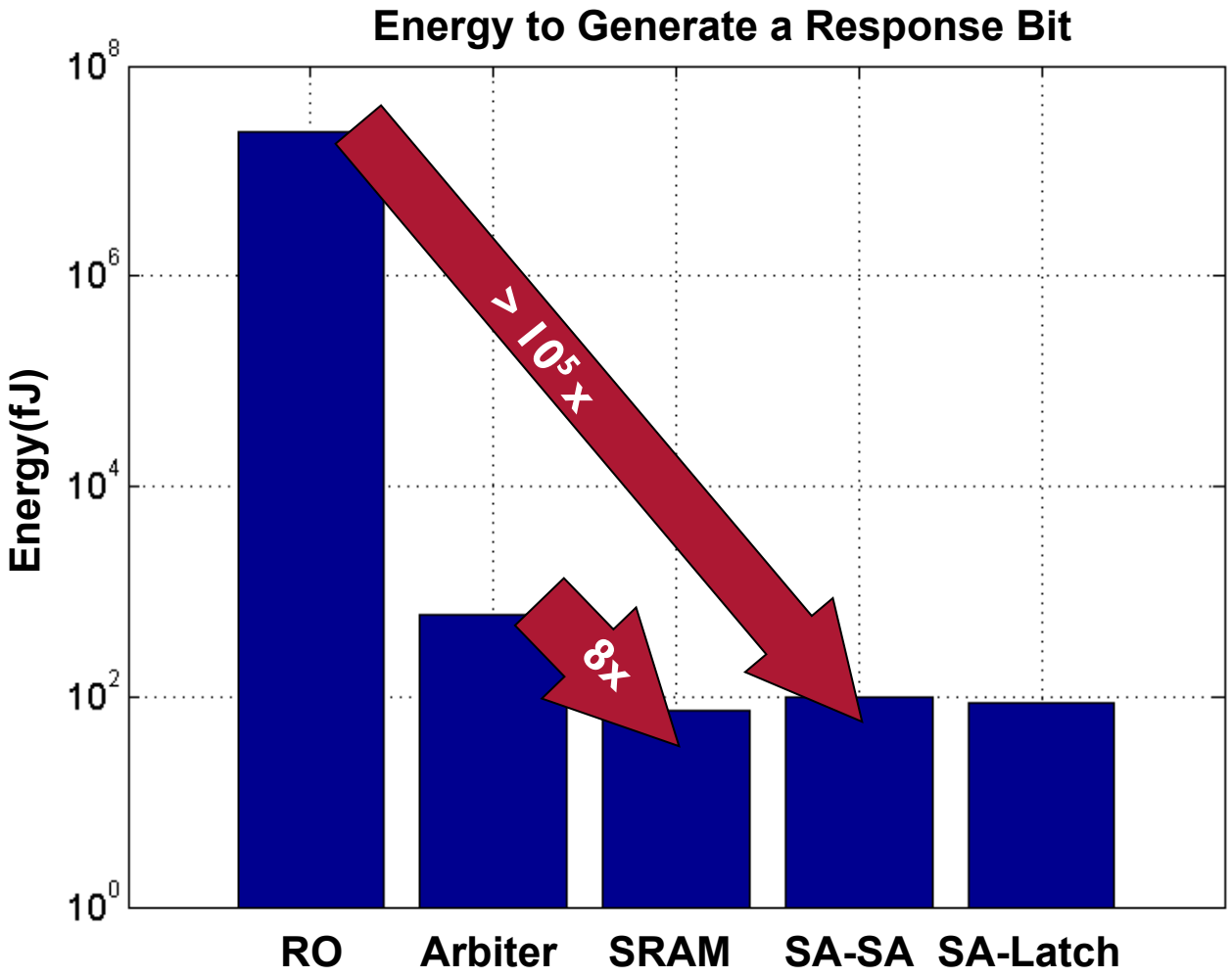
Voltages : 1.0V, 1.2V, 1.4V
Temperatures : -20^{0}C, 27^{0}C, 85^{0}C

Comparison: Delay



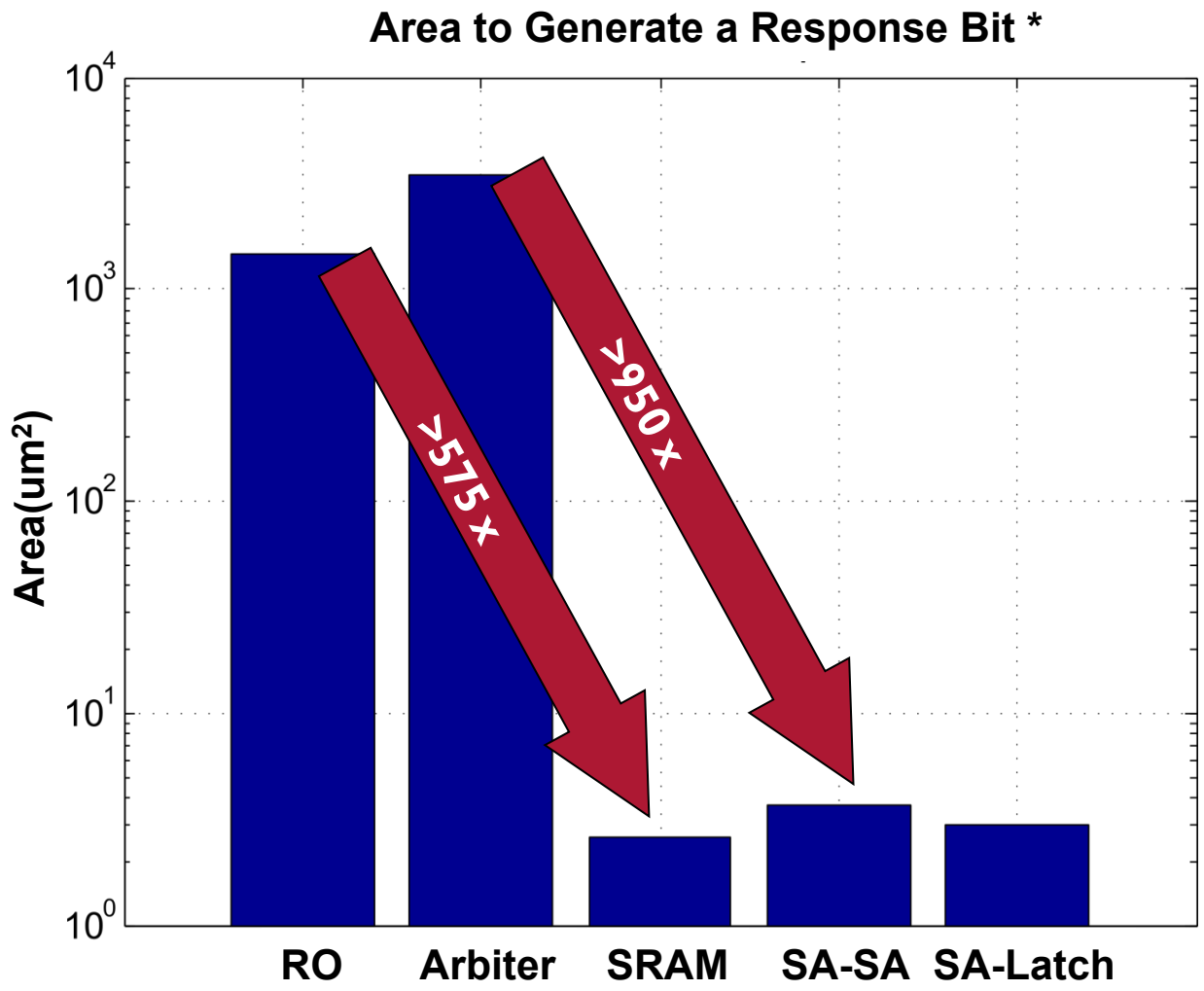
RO	112µs
Arbiter	16ns
SRAM	< 1ns
SA-SA	< 1ns
SA-Latch	< 1ns

Comparison: Energy



RO	23nJ
Arbiter	0.6pJ
SRAM	0.07pJ
SA-SA	0.10pJ
SA-Latch	0.09pJ

Comparison: Area



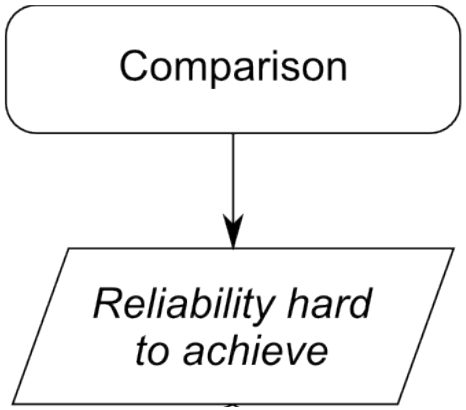
RO	1475μm²
Arbiter	3500μm²
SRAM	2.6μm²
SA-SA	3.6μm²
SA-Latch	2.9μm²

*Amortizes when same chain is re-used (using a different challenge), but at the cost of delay (and area). However, multiple studies have concluded that arbiter PUF is vulnerable to modeling attacks by eavesdropping on several challenge-response pairs (CRPs).

PUF Comparison Summary

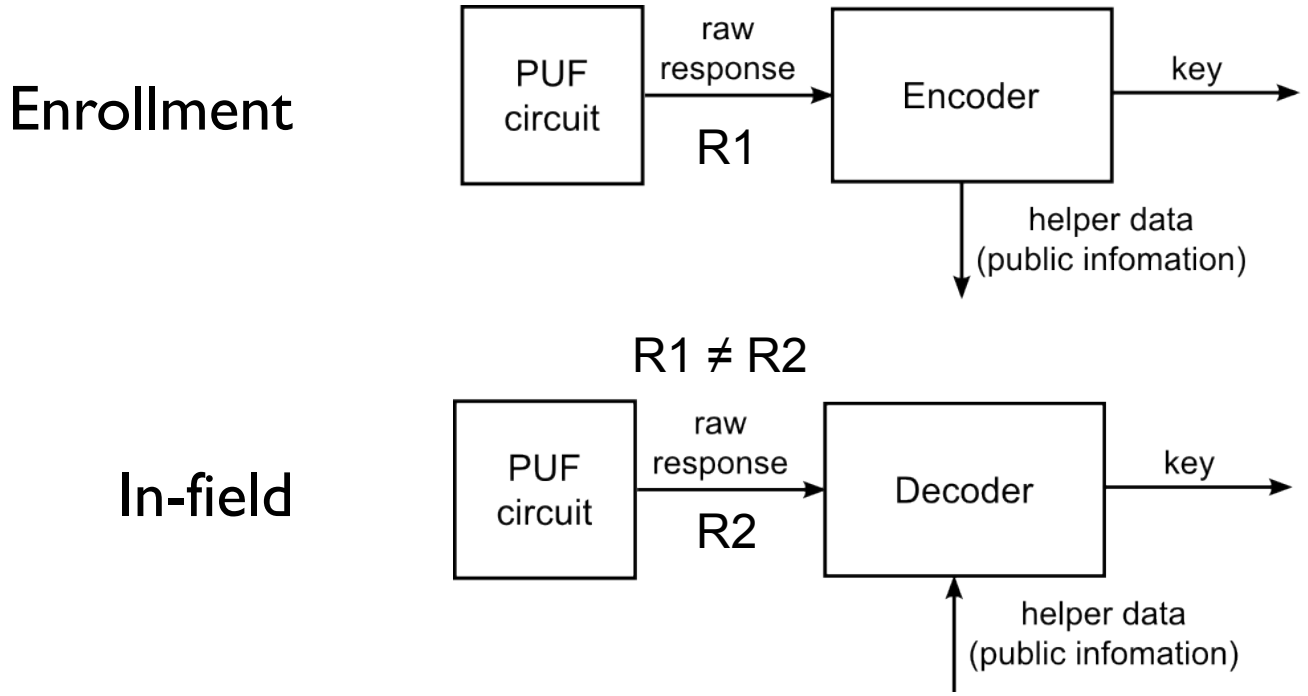
	Security Metrics			VLSI Metrics		
	Rand	Uniq	Reliab	Area	Delay	Energy
Arbiter	✓	✓	✗	✗	✗	✗
RO	✓	✓	✗	✗	✗	✗
SRAM	✓	✓	✗	✓	✓	✓
SA	✓	✓	✗	✓	✓	✓

Reliability Enhancement Options



Dodis et al, 2004
Guajardo et al., 2007
Bosch et al., 2008
Maes et al., 2009
Yu et al., 2010
Yu et al., 2011
Paral et al., 2011
Leest et al., 2012
Maes et al., 2012

Conventional Solution: Error Correction Codes



- Significant overheads
  - Delay, power, and area
  - Complexity scale quickly with number of correctable errors
  - For BER=15%, need 20-80 response bits/key bit
- Requires helper data
  - Can leak information
- Decode incurs delay
  - Often thousands of cycles
  - Micro- or milli-second timescales

Error Reduction Techniques

Extrinsic Techniques

- Multiple Evaluation (ME)
- Activation Control (AC)

Up to 70-80% reduction in errors

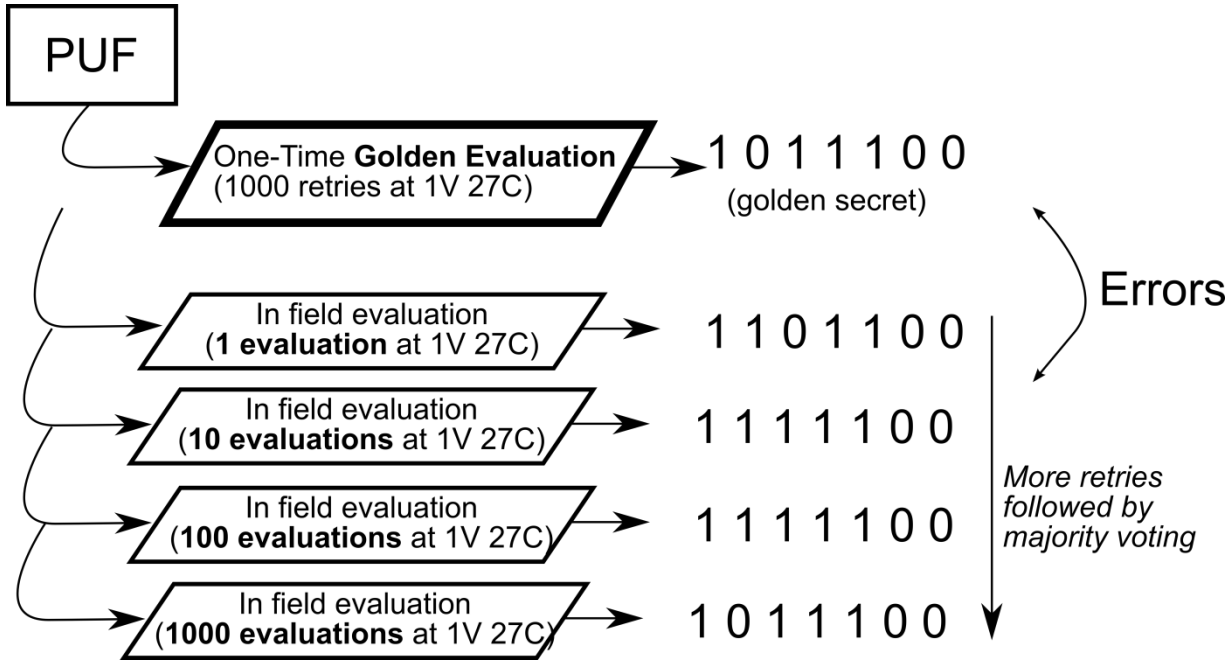
Intrinsic Techniques

- Post-Silicon Selection (PSS)
- Response Reinforcement (RR)

~ 100% reduction in errors

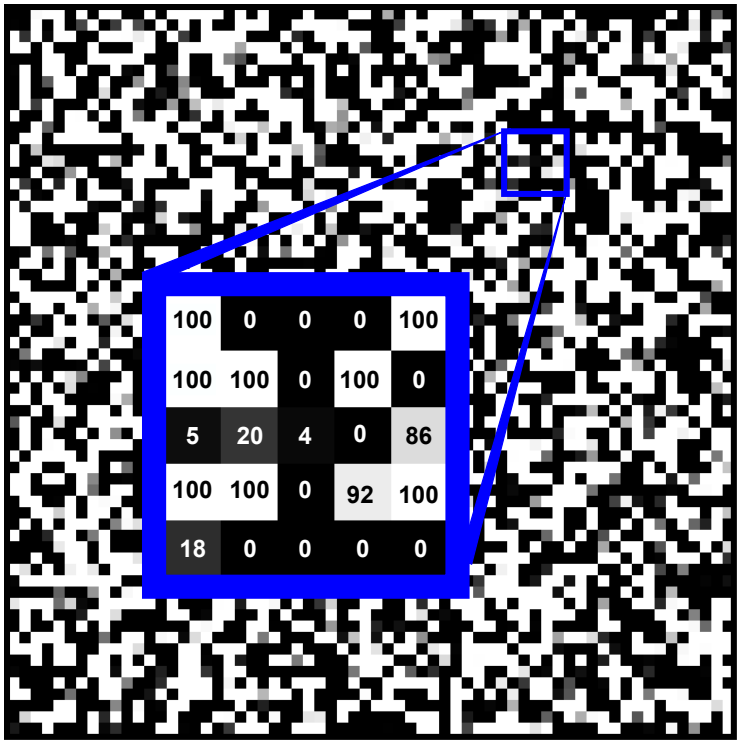
[Bhargava HOST 2012]

Multiple Evaluations (ME)

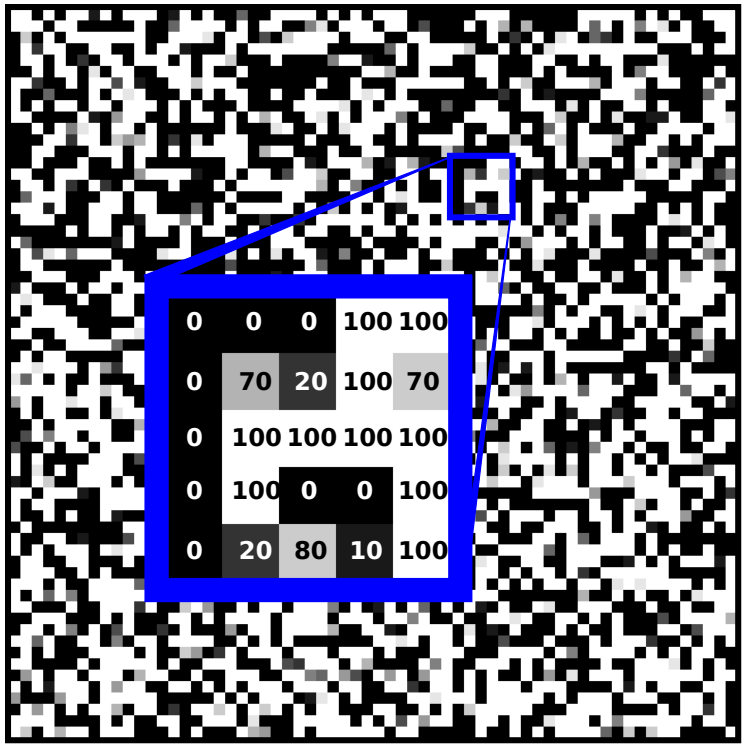


- **Multiple Evaluations done 1-1000 times**
  - Final bit value based on majority voting
- **Errors reduce, but at the cost of increased evaluation time**

ME in SRAM and SA PUF



SRAM Array Bitmap

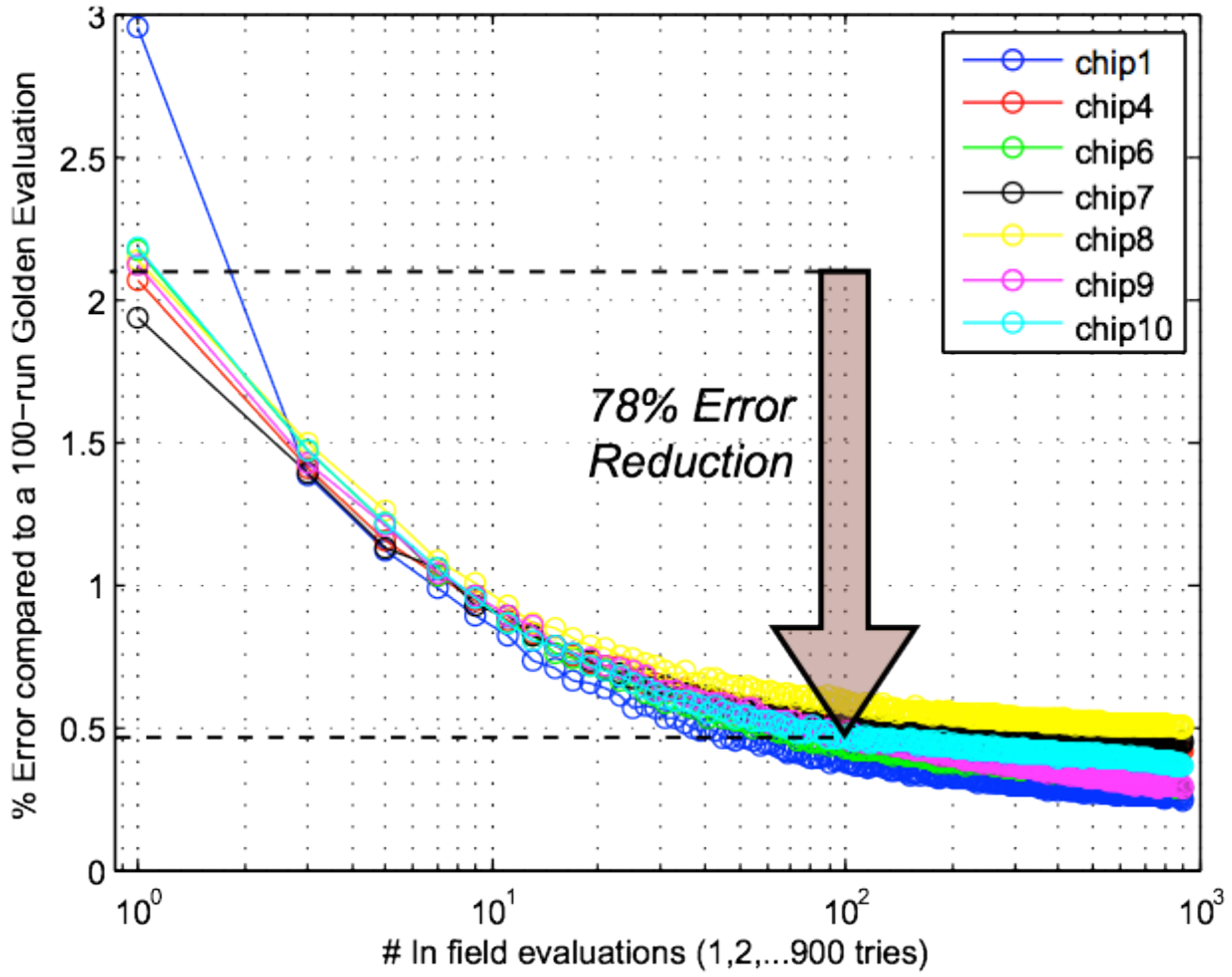


Sense Amplifier Array Bitmap

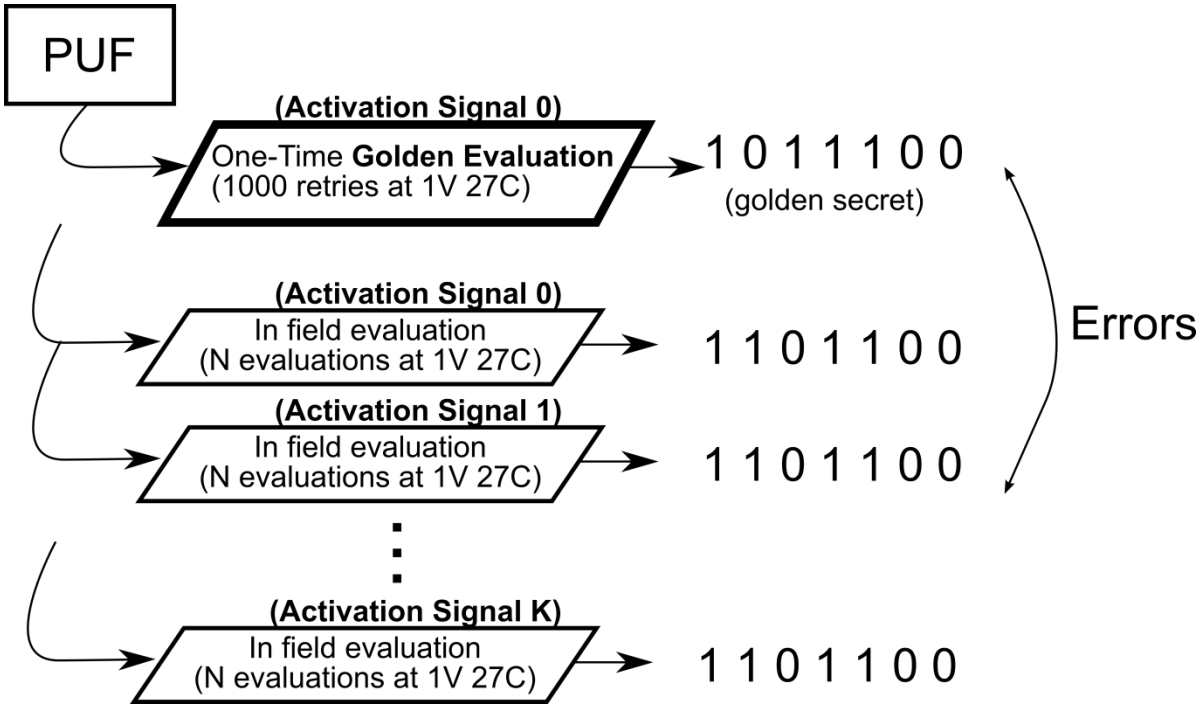
Squares represents % times a bit evaluated to '1' across multiple evaluations.
White ⇒ 'always 1'. Black ⇒ 'always 0'. For the rest, the greyscale value indicates the relative ratio of '1' and '0'.

- **Multiple evaluations at a single operating point**
- **Majority vote to decide final result**
- **Trades-off energy and delay for reliability**

ME Results

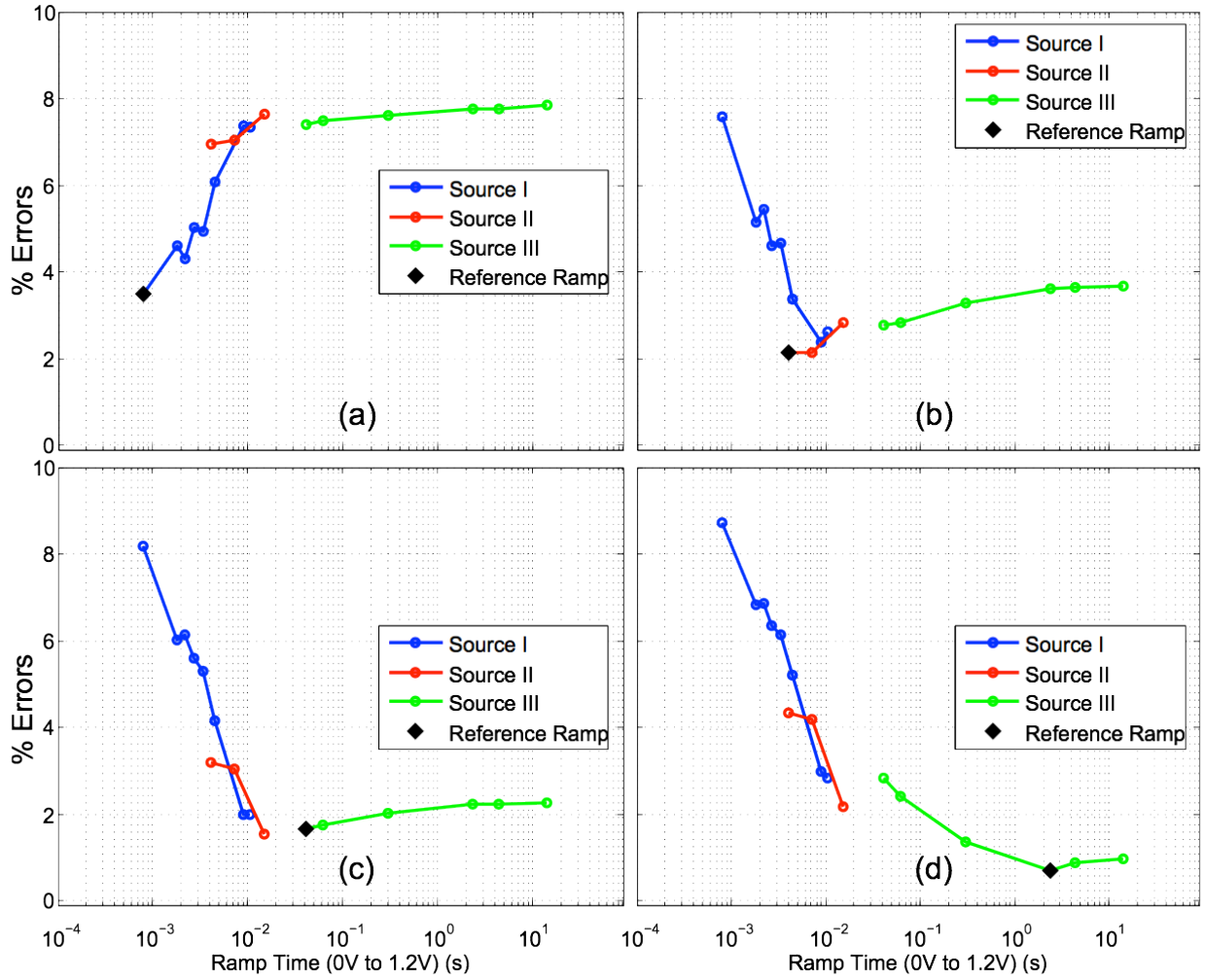

chip1
chip4
chip6
chip7
chip8
chip9
chip10
78% Error Reduction
% Error compared to a 100−run Golden Evaluation
# In field evaluations (1,2,...900 tries)
0
0.5
1
1.5
2
2.5
3
10^0
10^1
10^2
10^3

Activation Control (AC)



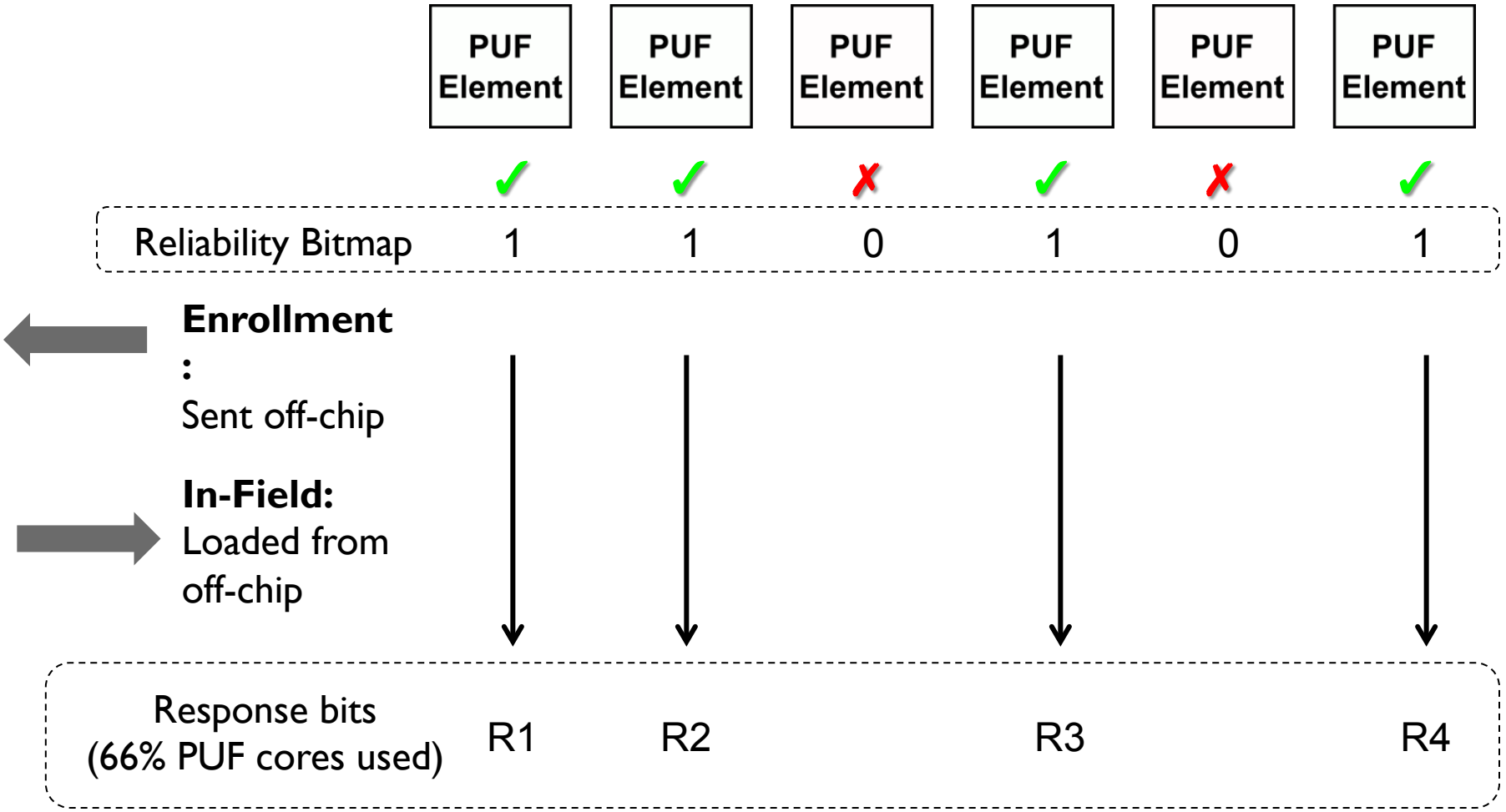
- **Activation Control**
  - SRAM: VDD ramp
  - Sense Amplifier: Sense enable ramp
- **Slow ramps result in more reliable SRAM PUFs**
  - Ramp rate variation from golden results in more errors
  - Tight control of activation control required for high reliability

AC (SRAM) Results



Reliability of an SRAM as a function of VDD ramp rate, but when the golden values were generated at a given ramp rate. The reference ramps are chosen at (a) 0.8ms, (b) 3ms, (c) 40ms, and (d) 25s

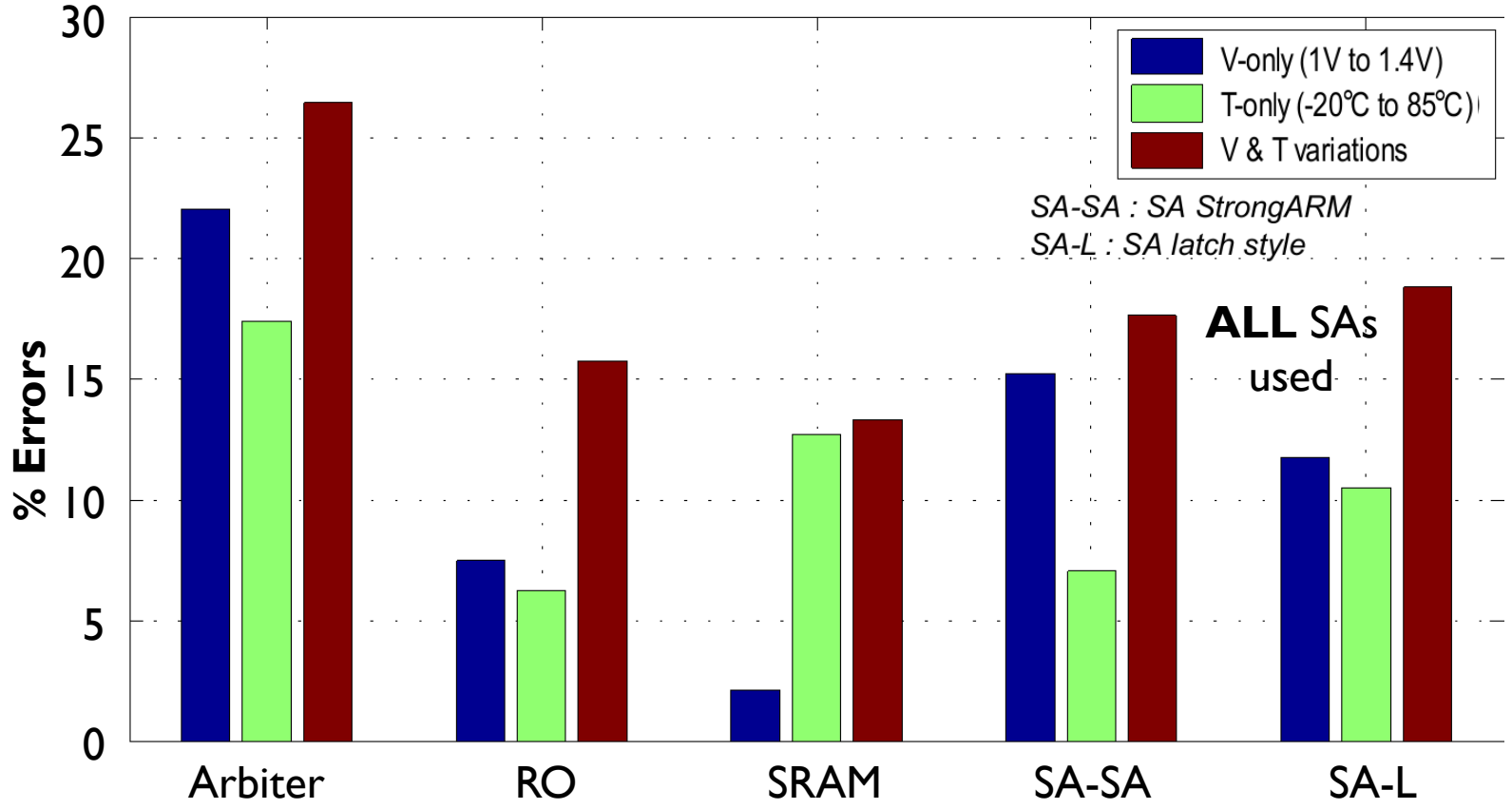
Post-Silicon Selection (PSS)



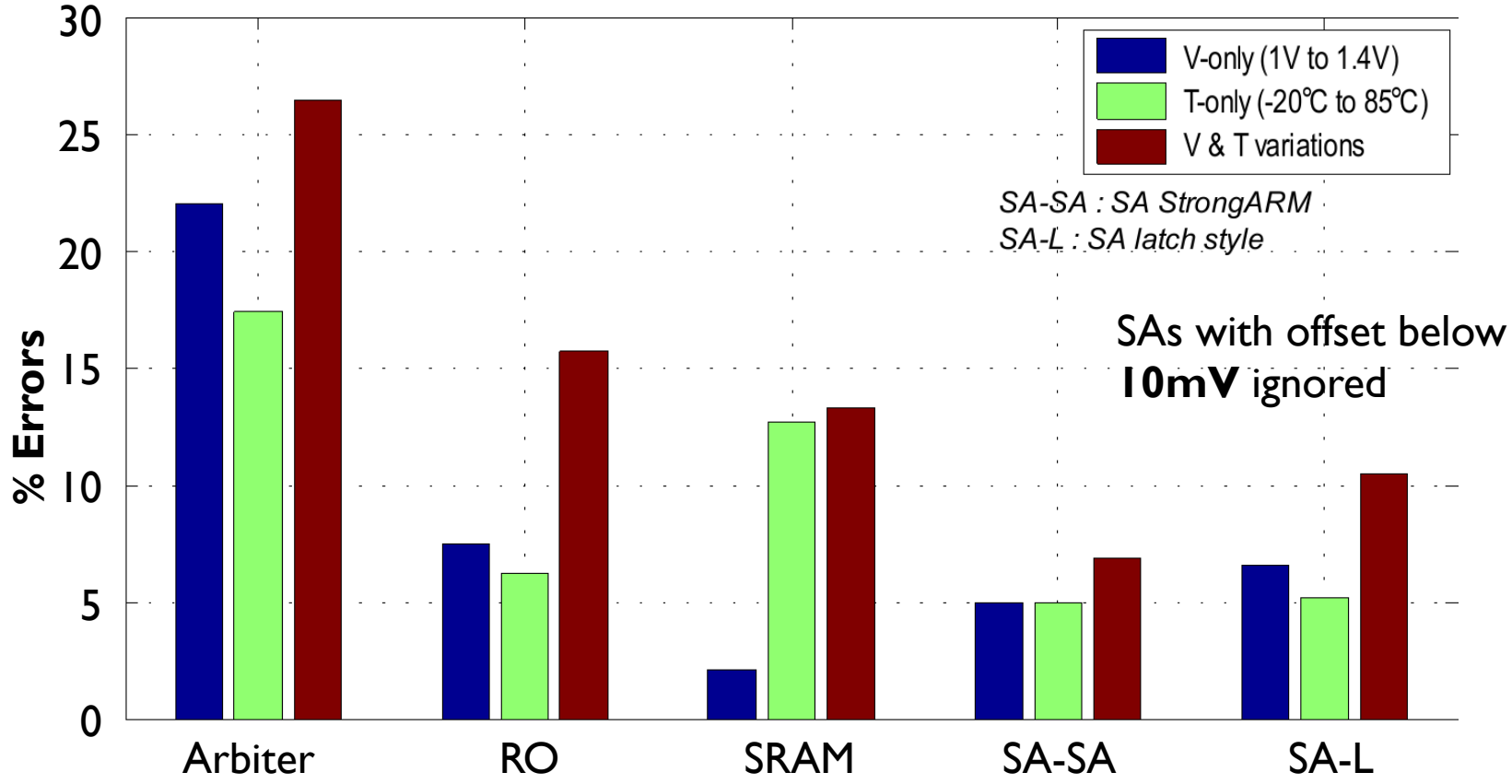
PSS: Reliability Estimation

- Direct estimation of reliability
  - Extensive evaluations at multiple voltage-temperature corners and noise scenarios
  - High tester time, no insights of the safety margin in selected bits
- Indirect estimation of reliability
  - "Soft"-information from PUF elements correlated with reliability
  - Use SA offset voltage to estimate reliability

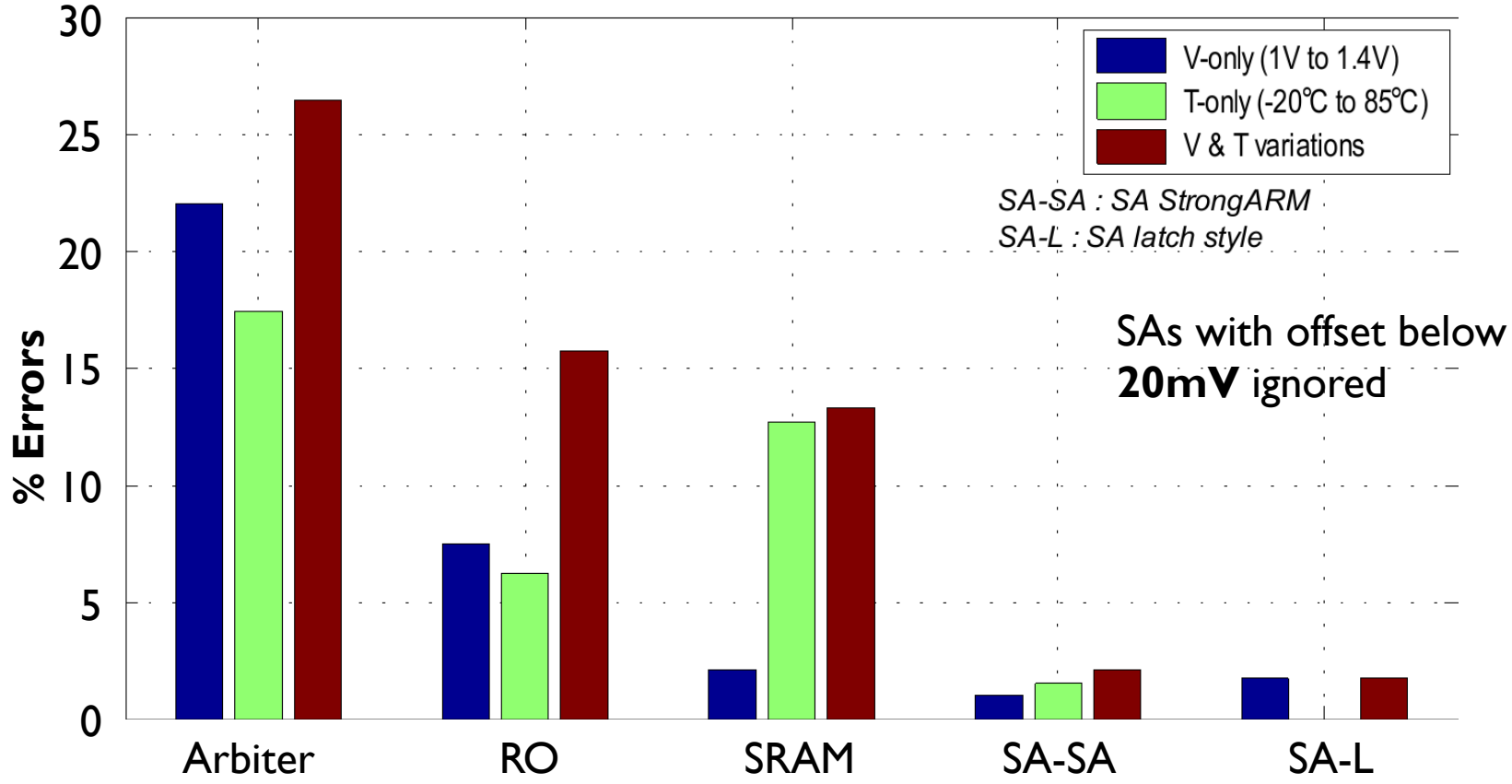
Comparison: Reliability



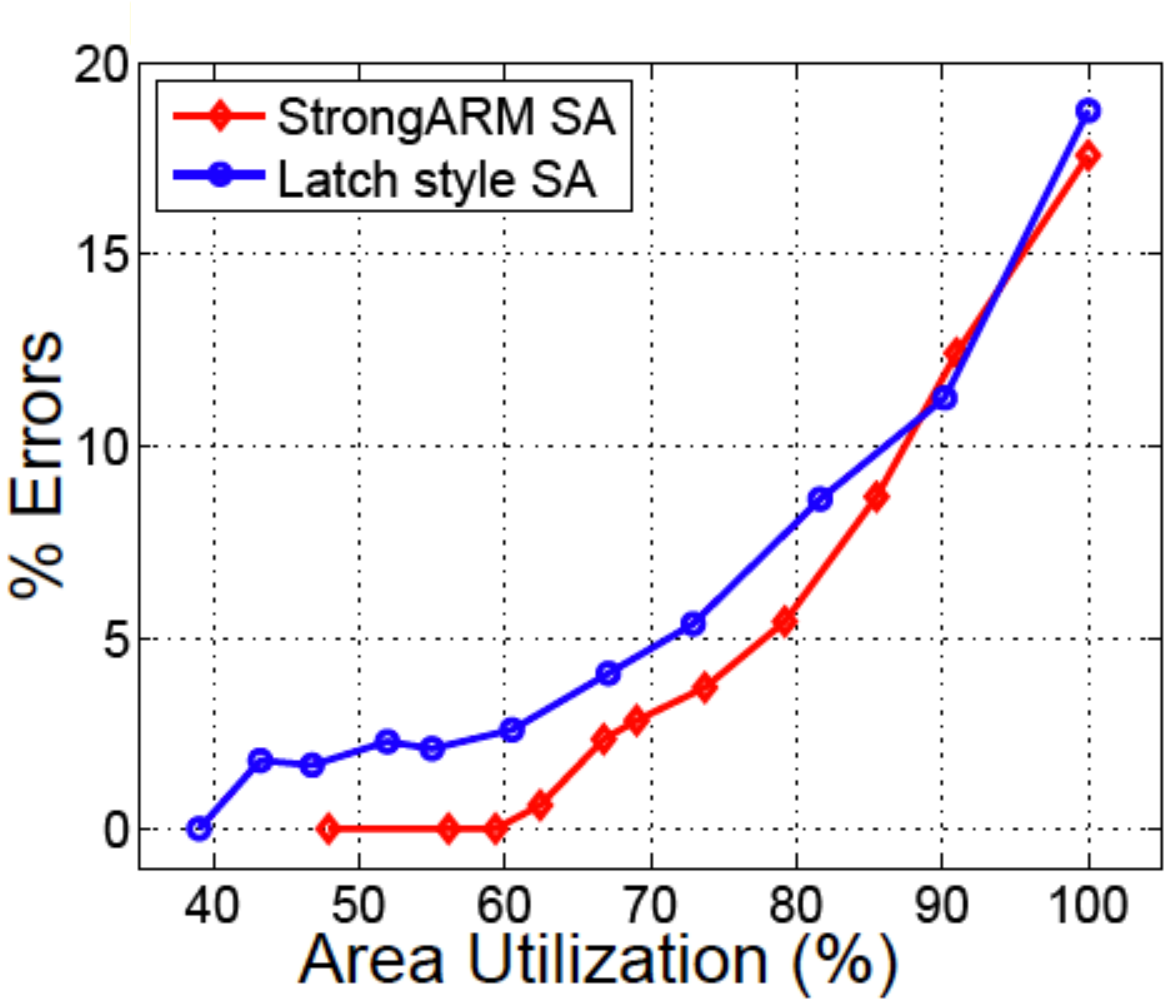
Comparison: Reliability



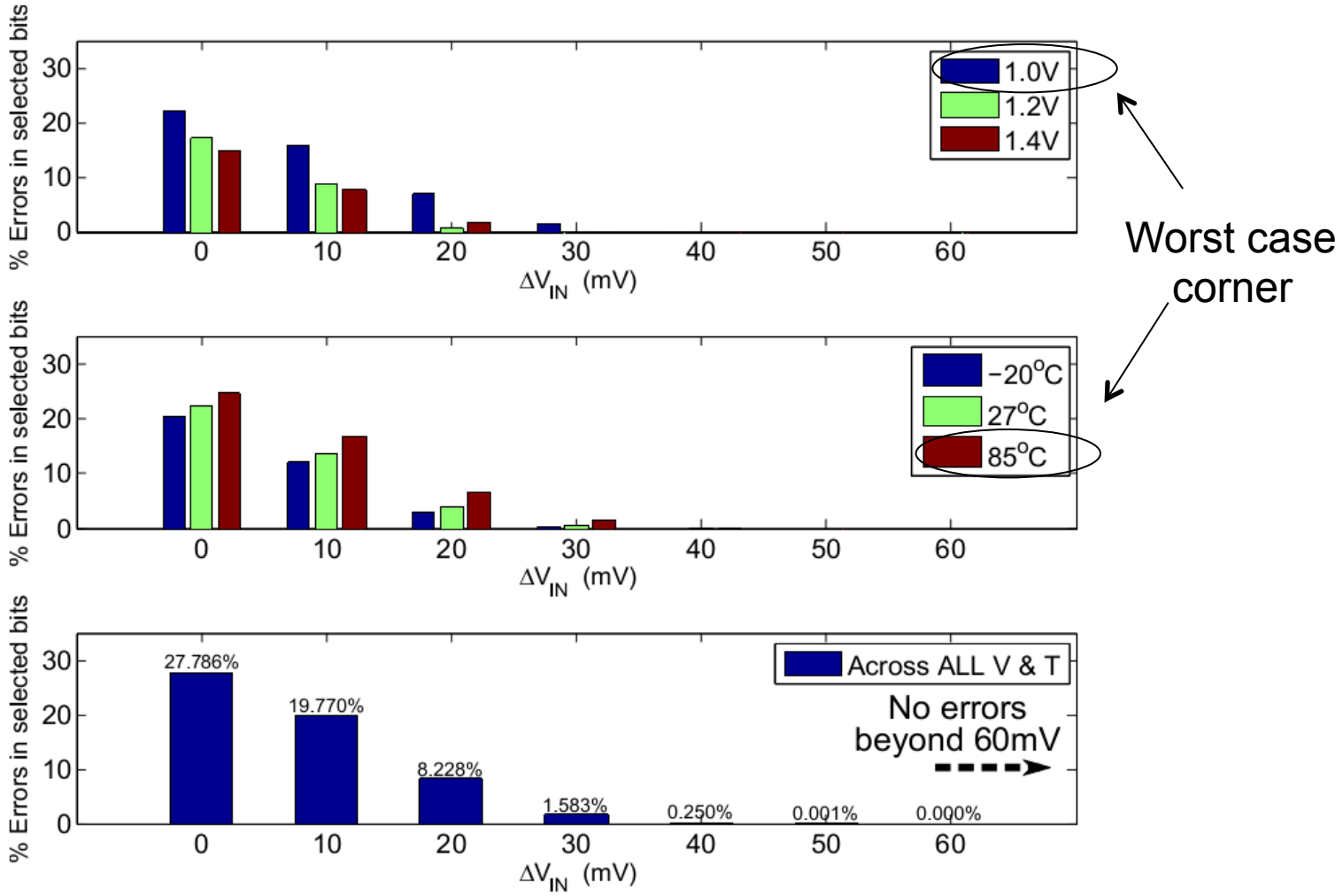
Comparison: Reliability



Reliability vs. Area Tradeoff for SAs



PSS: Reliability Measurements Across V/T

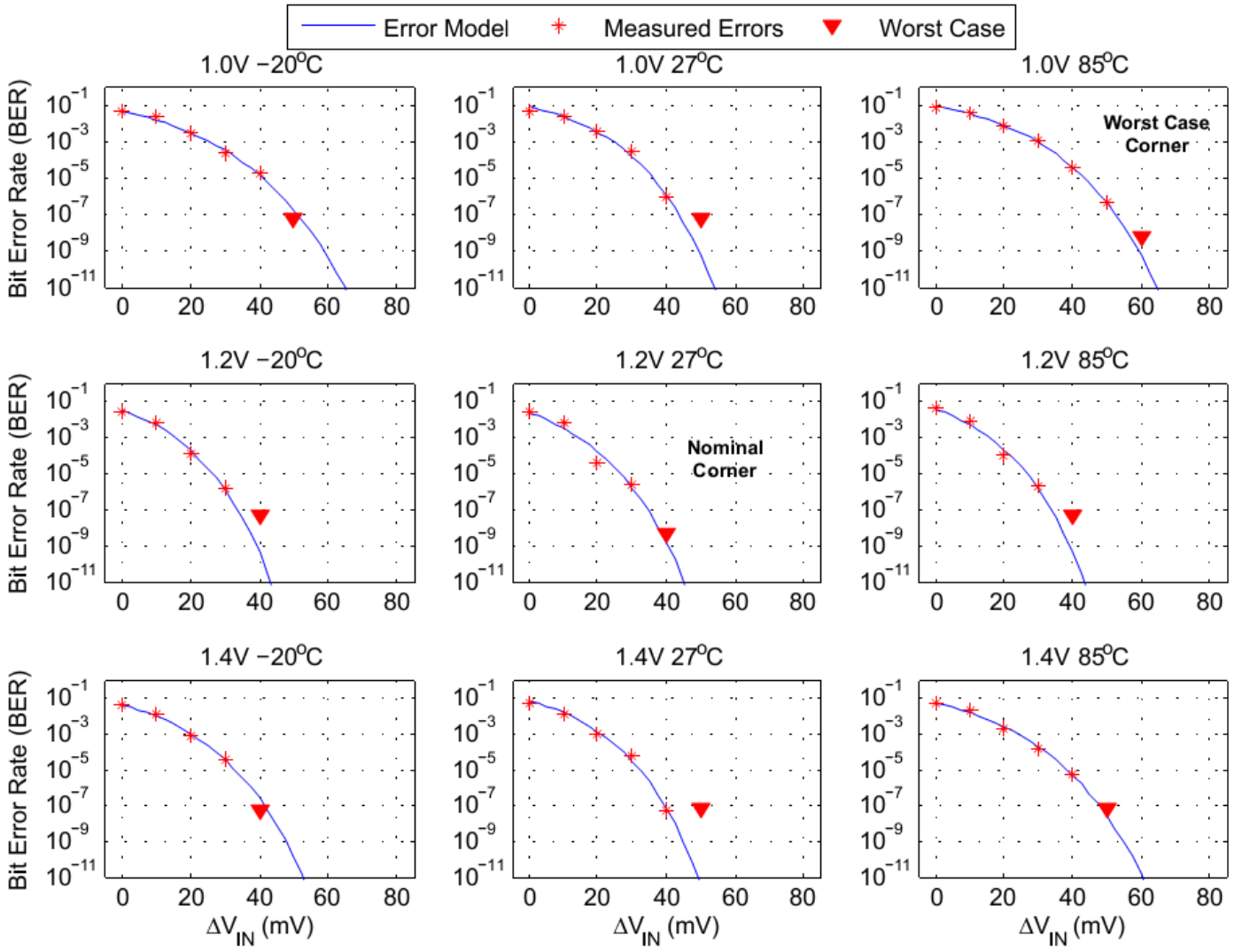


10,000 measurements each at all 9 voltage & temperature combinations
Voltages : 1.0V, 1.2V, 1.4V Temperatures = -20^{0}C, 27^{0}C, 85^{0}C

PSS: Large-Scale Reliability Measurements

- Large-scale measurements
  - 180k at worst case: 1.0V 85^{0}C
  - 100k at nominal: 1.2V 27^{0}C
- No errors in 1213 (29.6%) selected SAs with $\triangle V_{IN}$ = 60mV
- Bit error rate < 4.6 * 10^{-9}
  - 128-bit key error rate < 0.6 * 10^{-6}
- Key error rate < 10^{-6} : typical target in ECC papers

PSS: Large Scale Test Results



Response Reinforcement

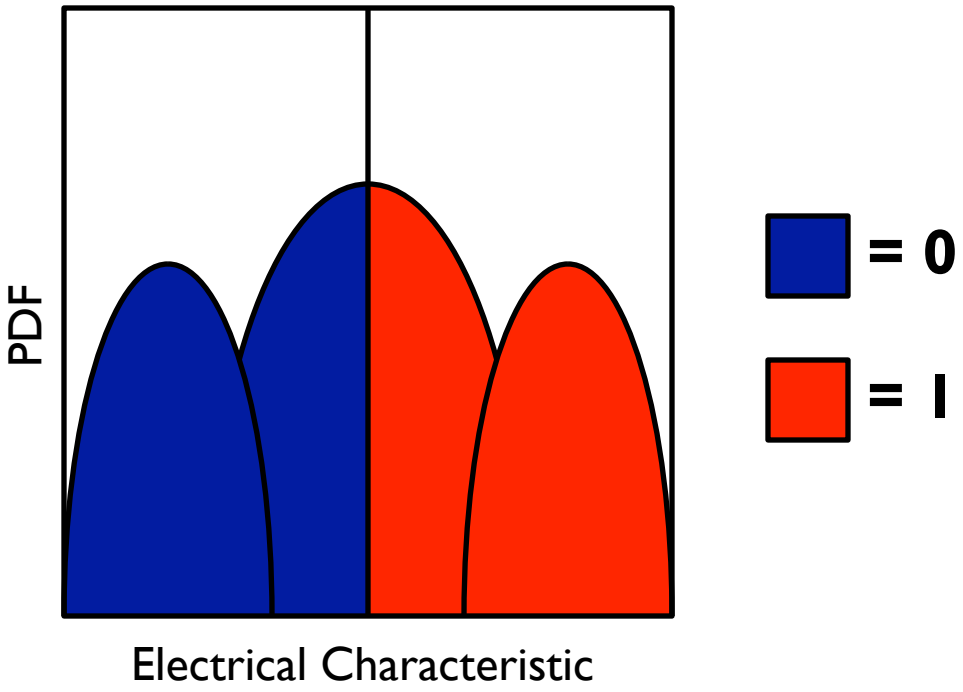
Response reinforcement

- Increase the baseline reliability of the PUF core circuit
- Post-manufacturing amplification of random variations
- Minimize or eliminate the need for ECC
- No helper data

Implementation

- Measure PUF "golden" response
- Reinforce golden response using directed accelerated aging
- Artificially induce IC aging phenomena to amplify PUF circuit random variation for increased reliability

Response Reinforcement Concept



- Can improve PUF reliability by increasing variability
- Optimal distribution is bi-modal

Integrated Circuit Aging Phenomena

Many IC aging effects

- Negative Bias Temperature Instability (NBTI)
- Time Dependent Dielectric Breakdown (TDDB)
- Metal electro-migration (EM)
- Hot Carrier Injection (HCI)

Desired characteristics

- Easy to artificially induce
- Short reinforcement time
- Strong reinforcement effect
- High permanence

Integrated Circuit Aging Phenomena

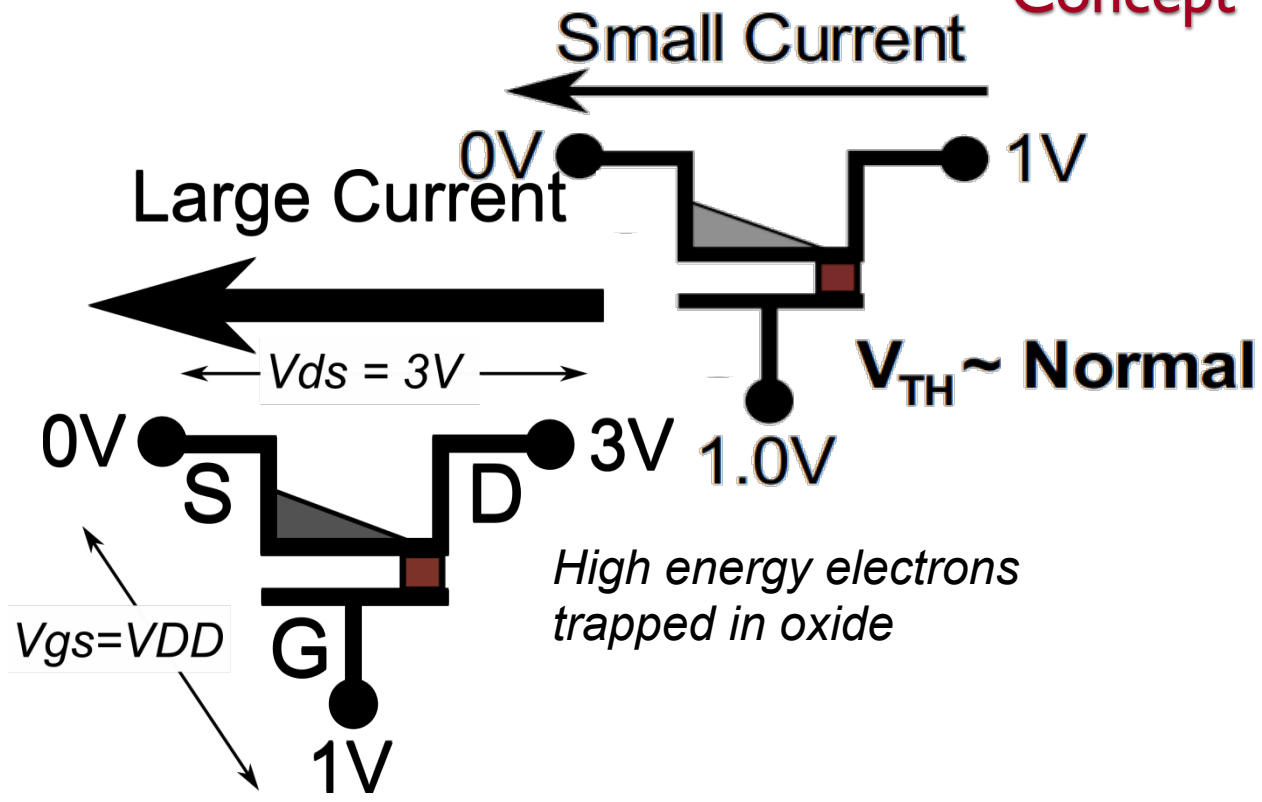
Many IC aging effects

- Negative Bias Temperature Instability (NBTI) [Bhargava HOST 2012]
- Time Dependent Dielectric Breakdown (TDDB)
- Metal electro-migration (EM)
- Hot Carrier Injection (HCI) [Bhargava CHES 2013]

Desired characteristics

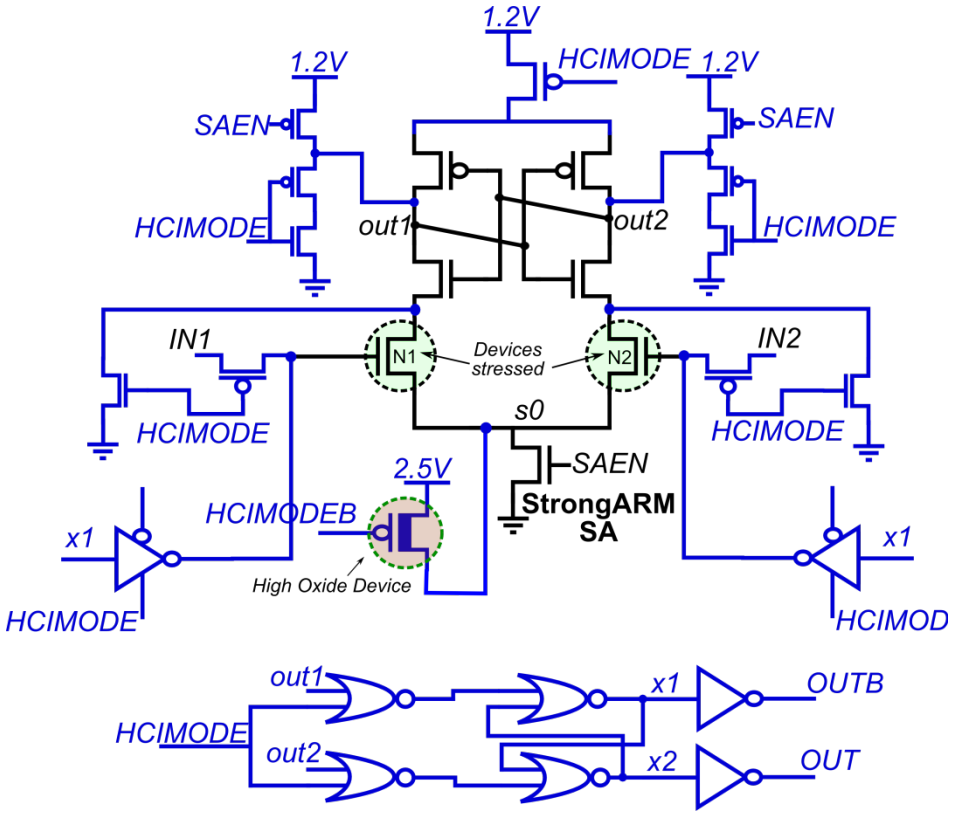
- Easy to artificially induce — Only need a raised voltage ~3V
- Short reinforcement time — ~10s reinforcement (one time)
- Strong reinforcement effect — Shifts transistor V_{TH} by >50mV
- High permanence — Effect lasts for years

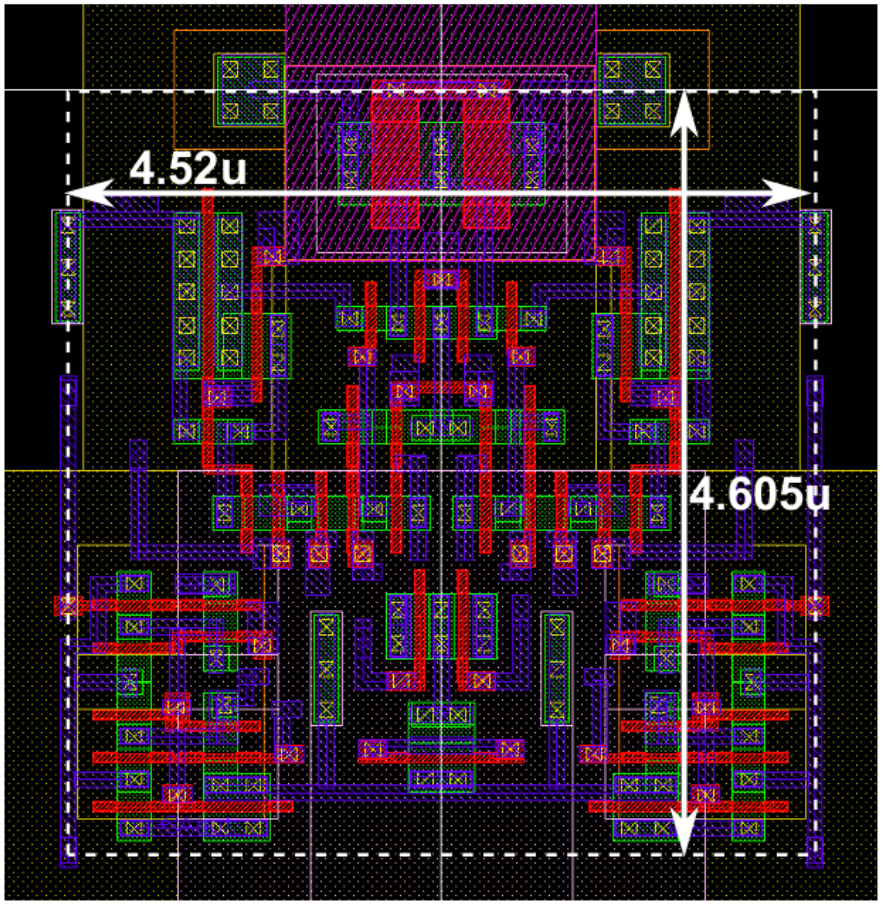
Hot Carrier Injection
Concept



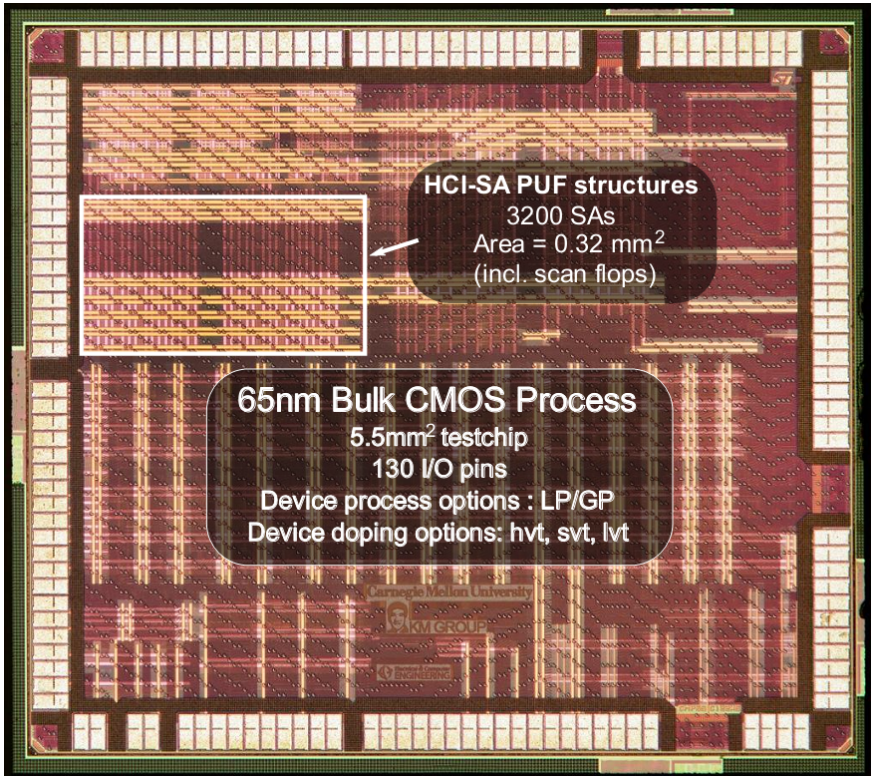
- Small increase in V_{TH} if current in same direction
- High increase in V_{TH} (~ 100 mV) if current in opposite direction

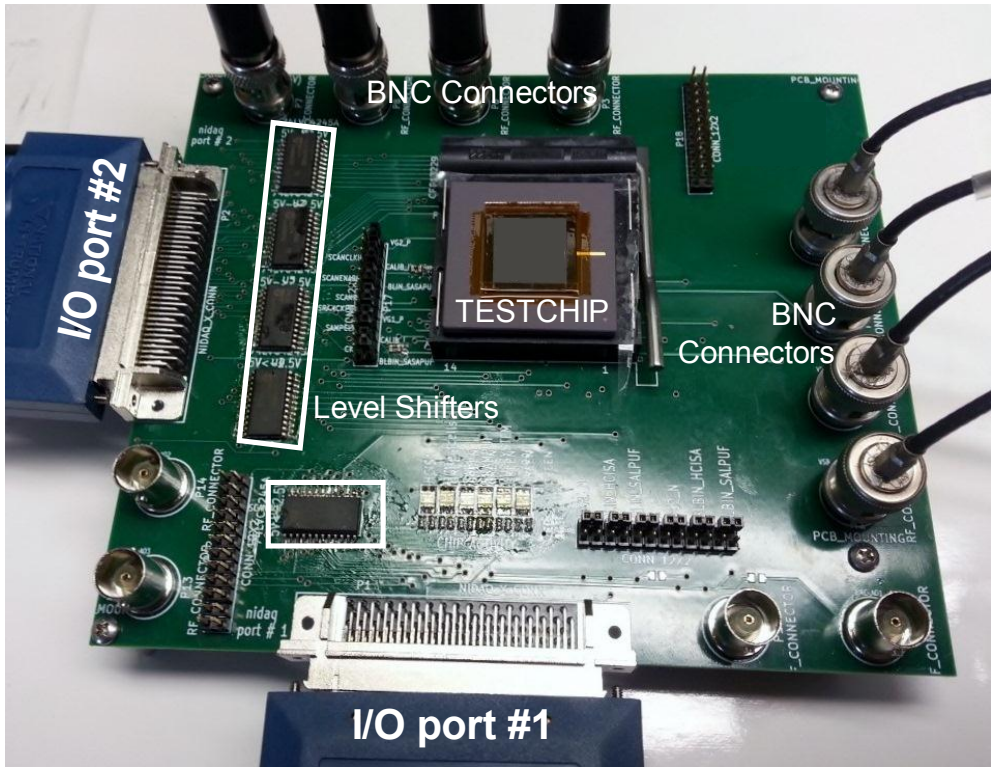
Hot Carrier Injection Sense Amplifier (HCI-SA)





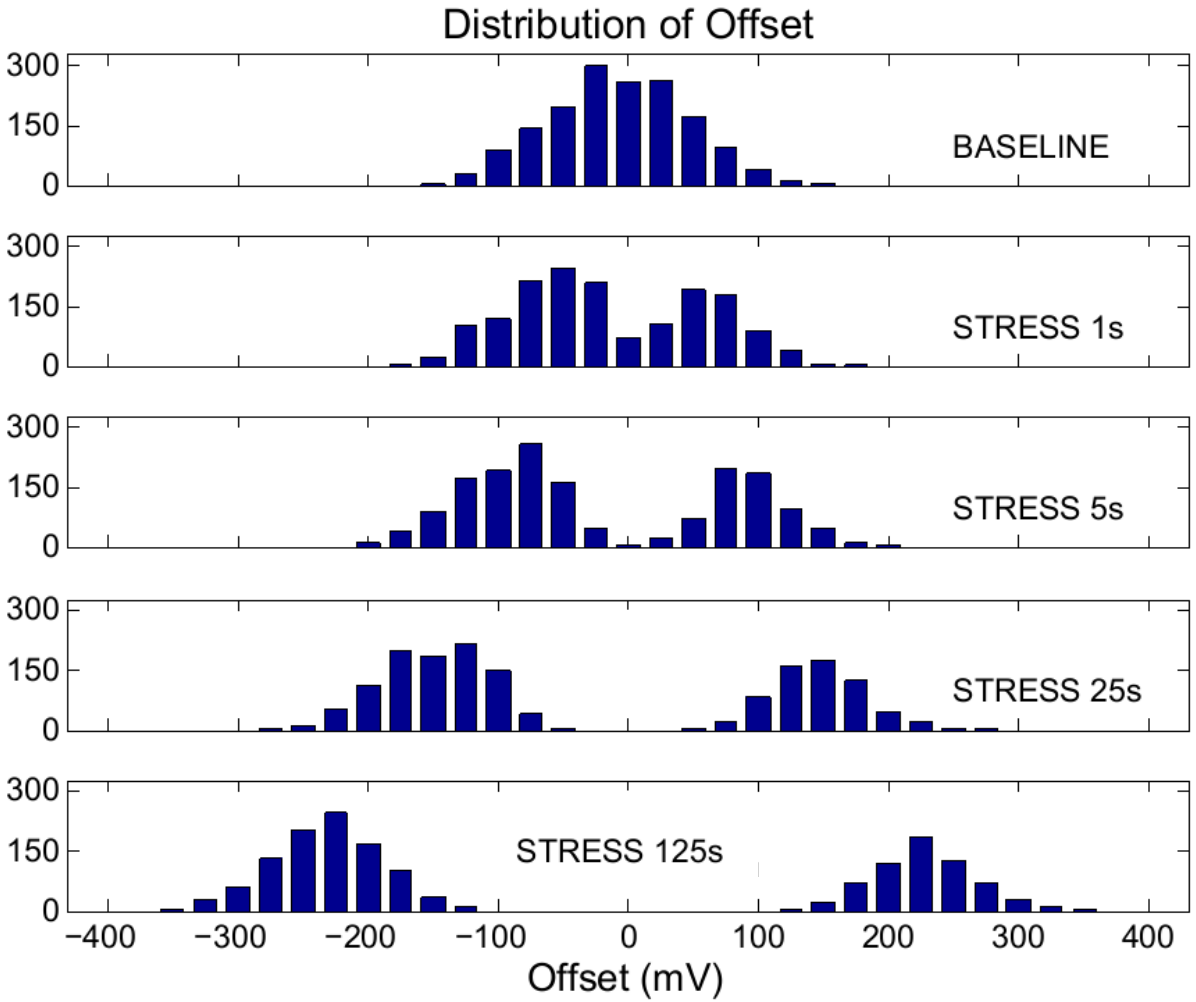
HCI-SA Testchip



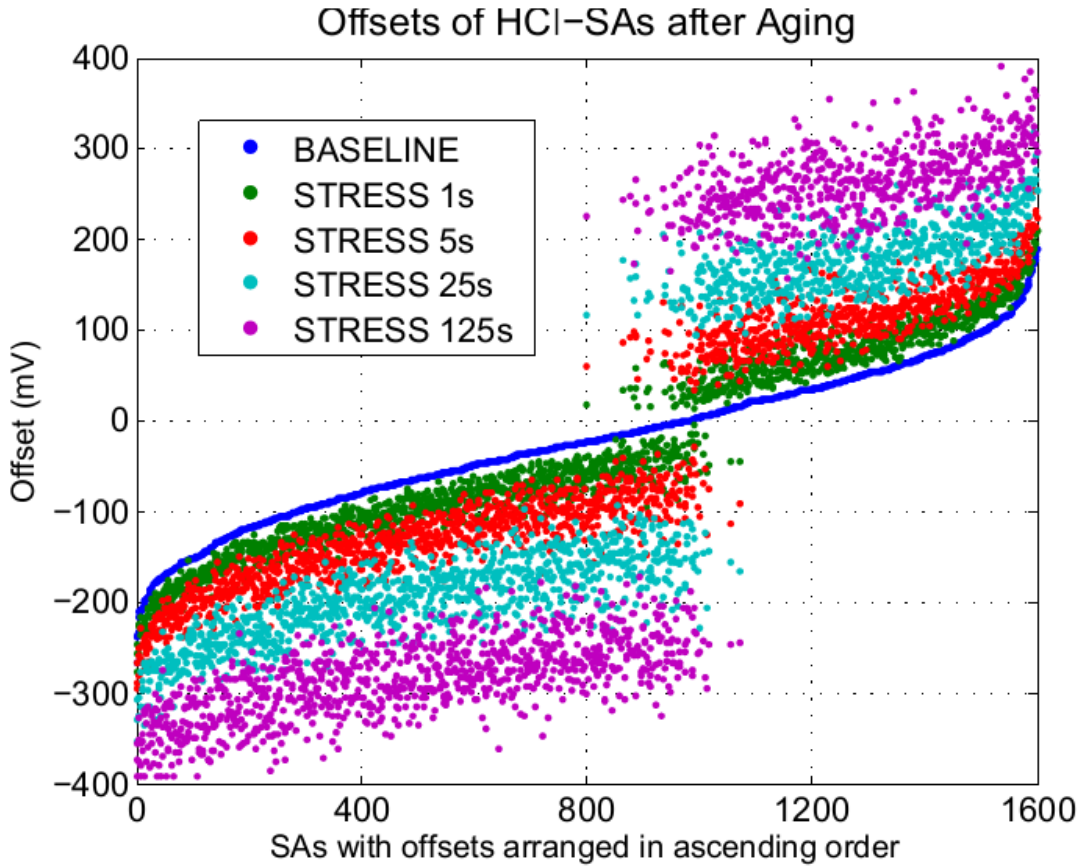


- 1600 self-reinforcing HCI-SA
- 1600 manually controlled HCI-SA
- Tested across 9 voltage/temperature corners
- HCI stress times of 1s, 5s, 25s, 125s

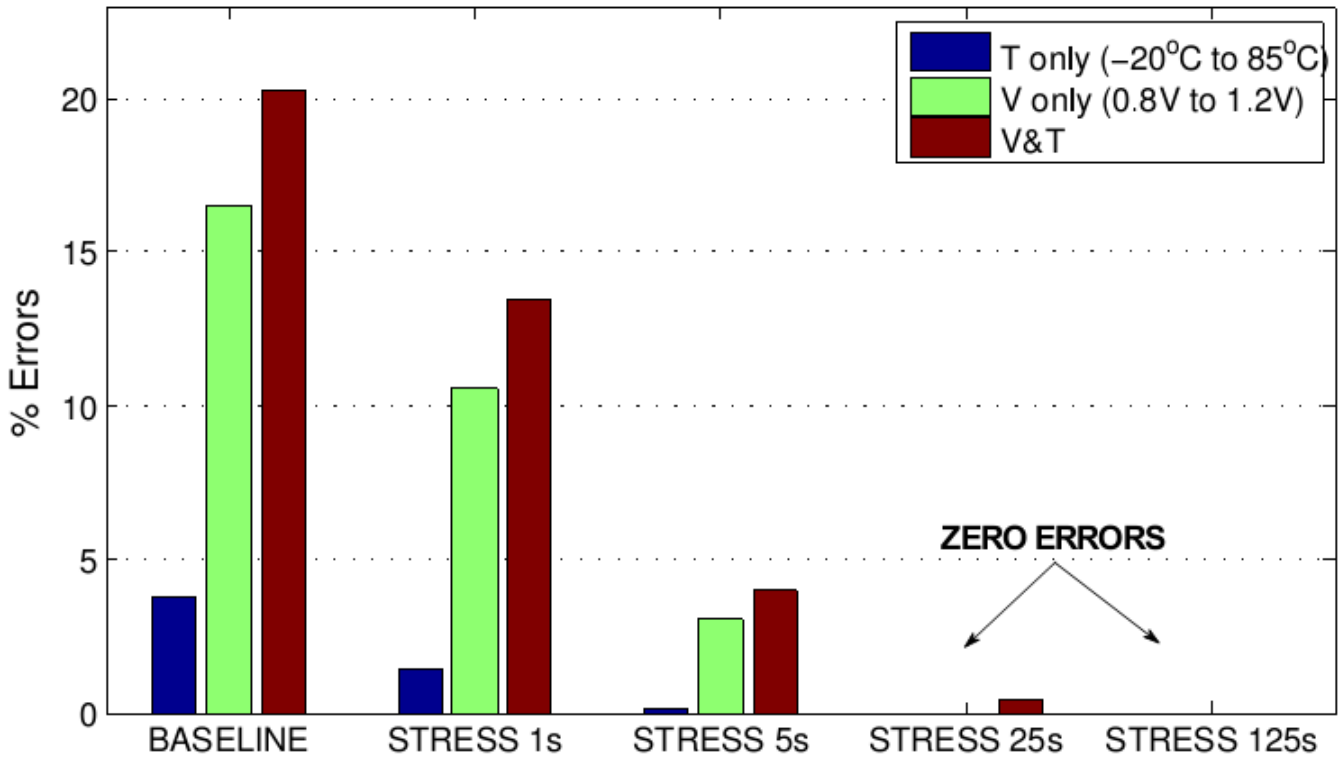
HCI-SA Offset Shift



HCI-SA Offset Shift

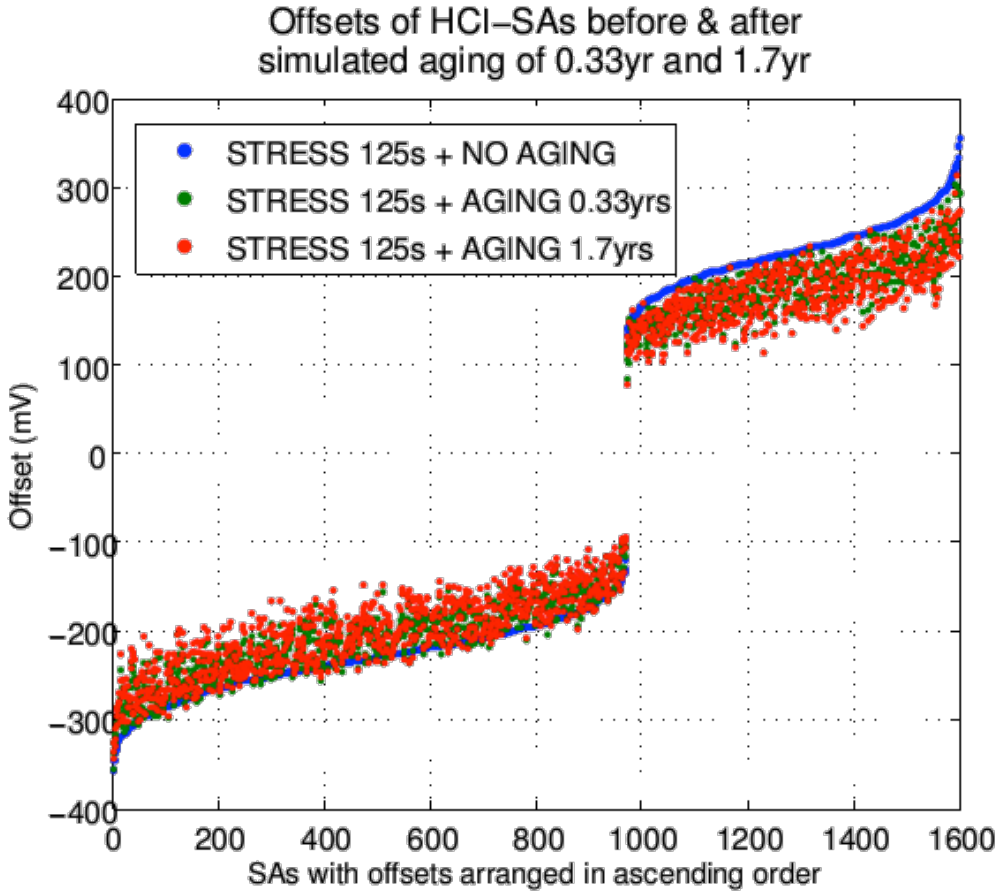


HCI-SA Reliability Measurements



100 runs at all 9 voltage/temperature corners
→ No errors found after stress of 125 seconds

HCI-SA: Permanence of Offset Shift



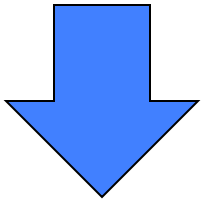
Baked chips at 1.5V and 100^{0}C

- 18 hours ➔ 0.33 years
- 93 hours ➔ 1.7 years

Large-Scale Reliability Measurements

Measured 125k evaluations (125s HCI stress)

- At nominal corner (1.2V 27^{0}C)
- At worst case corner (1.0V -20^{0}C)
- No errors observed in any of the 1600 HCI-Sas



- Bit error rate BER < $5 * 10^{-9}$
- Key error rate KER < $0.6 * 10^{-6}$ (128-bit)
- KER target < 10^{-6} for reliable key generation

HCI Response Reinforcement Summary

HCI-SA PUF

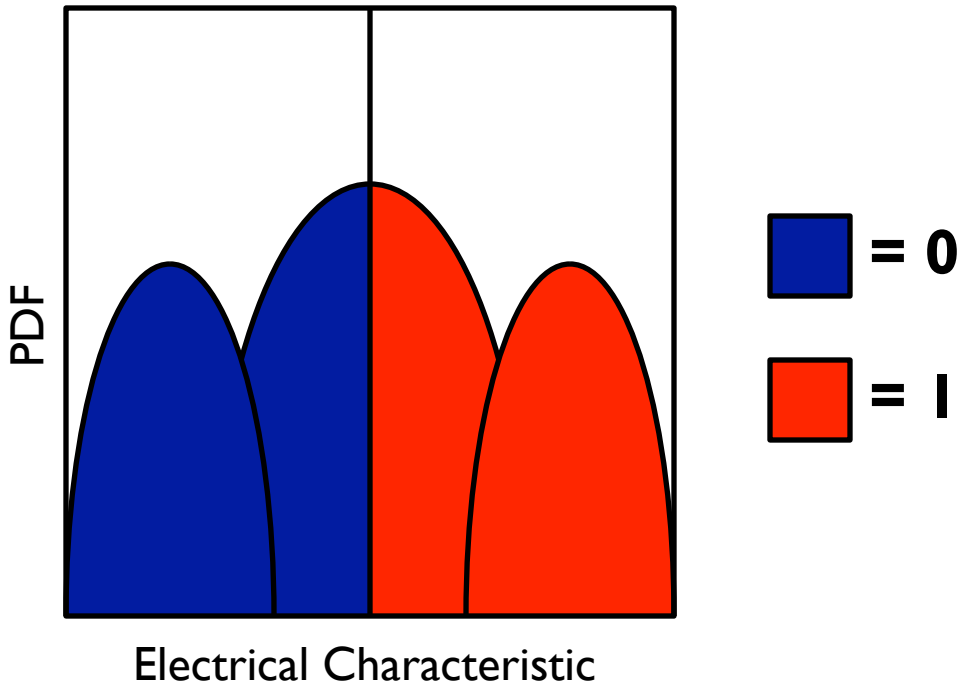
- Reliable – BER < 5 * 10^{-9} without ECC
- Secure – No helper data
- Fast – Response generation in 1 cycle (~1ns)
- Simple – One-time short reinforcement step (125s)
- High Permanence – Small change after ~2yr simulated aging

Can use RR on other types of PUFs as well

Conclusions

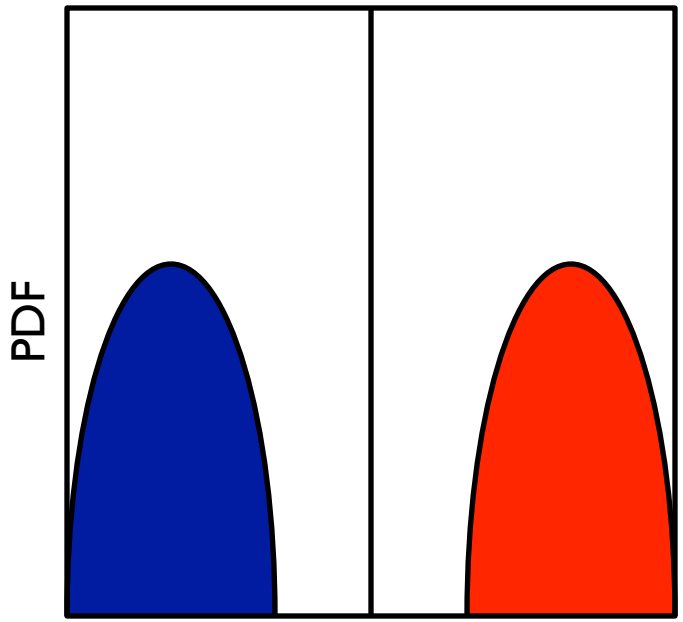
- Bi-stable PUFs have significant VLSI advantages
  - Despite being "weak"
- Multiple ways of decreasing error rate using COTS SRAM
  - Before using ECC / fuzzy extractors
- With custom PUF core can significantly reduce error rate
  - Minimal (or no?) ECC / fuzzy extractors / helper data needed

Parting Thought: Response Reinforcement Redux

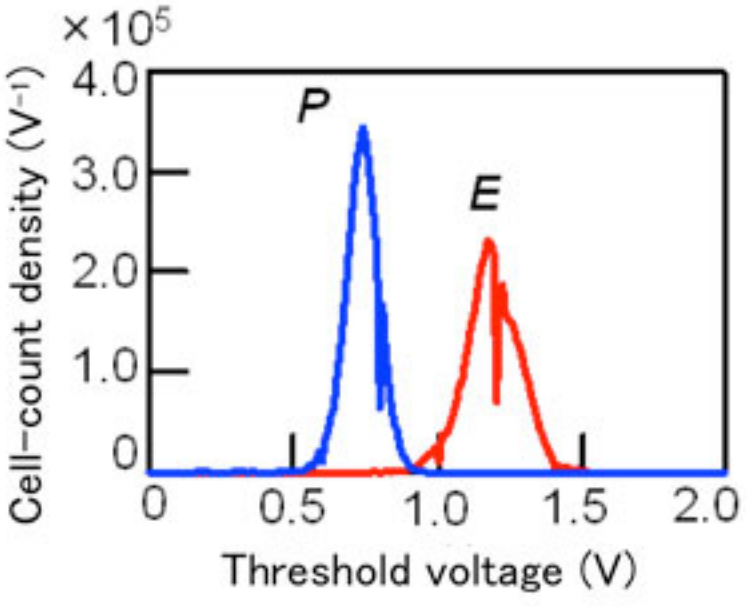


- Can improve PUF reliability by increasing variability
- Optimal distribution is bi-modal

Parting Thought: Through the Looking Glass



Electrical Characteristic



NAND Flash cell voltage distribution

- How different is a PUF from a non-volatile memory?
- Trade-off between PUF robustness and security

Acknowledgments

Students

- Mudit Bhargava
- Cagla Cakir
- Mark McCartney
- Craig Teegarden

Funding

- National Science Foundation
- Semiconductor Research Corporation (SRC)
- Carnegie Mellon CyLab

References

- Bhargava, M.; Ken Mai, "A High Reliability PUF Using Hot Carrier Injection Based Response Reinforcement," *Workshop on Cryptographic Hardware and Embedded Systems (CHES)*, 2013.
- Bhargava, M.; Cakir, C.; Ken Mai, "Comparison of bi-stable and delay-based Physical Unclonable Functions from measurements in 65nm bulk CMOS," *Custom Integrated Circuits Conference (CICC), 2012 IEEE* , vol., no., pp.1,4, 9-12 Sept. 2012.
- Cakir, C.; Bhargava, M.; Ken Mai, "6T SRAM and 3T DRAM data retention and remanence characterization in 65nm bulk CMOS," *Custom Integrated Circuits Conference (CICC), 2012 IEEE* , vol., no., pp.1,4, 9-12 Sept. 2012.
- Bhargava, M.; Cakir, C.; Ken Mai, "Reliability enhancement of bi-stable PUFs in 65nm bulk CMOS," *Hardware-Oriented Security and Trust (HOST), 2012 IEEE International Symposium on* , vol., no., pp.25,30, 3-4 June 2012.
- Bhargava, M.; Cakir, C.; Mai, K., "Attack resistant sense amplifier based PUFs (SA-PUF) with deterministic and controllable reliability of PUF responses," *Hardware-Oriented Security and Trust (HOST), 2010 IEEE International Symposium on* , vol., no., pp.106,111, 13-14 June 2010

Thank You