北京大学
PEKING UNIVERSITY

北京大学高能效计算与应用中心
Center for Energy-efficient Computing and Applications

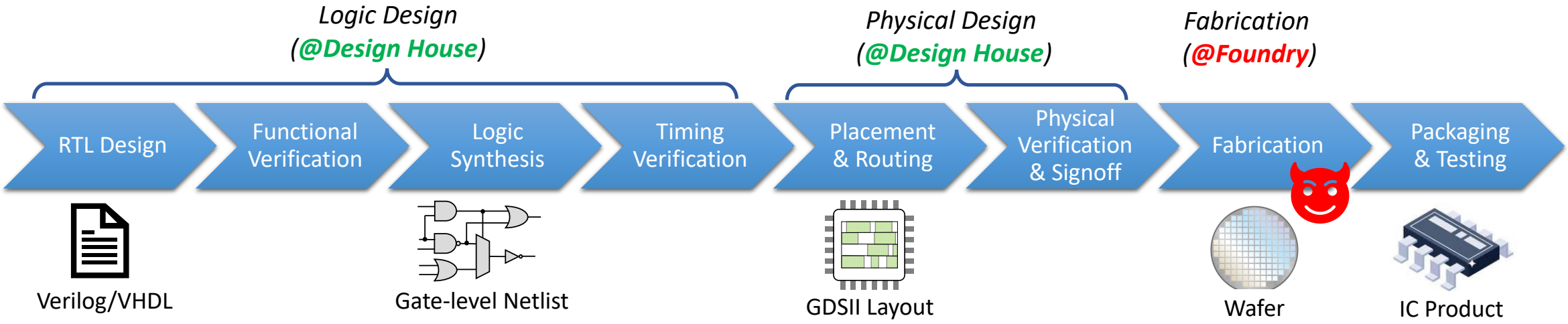
Rethinking IC Layout Vulnerability: Simulation-Based Hardware Trojan Threat Assessment with High Fidelity

Xinming Wei, Jiaxi Zhang, Guojie Luo
Peking University

{weixinming, zhangjiaxi, gluo}@pku.edu.cn

IC Design Process

- IC design flows are mostly **Fabless**

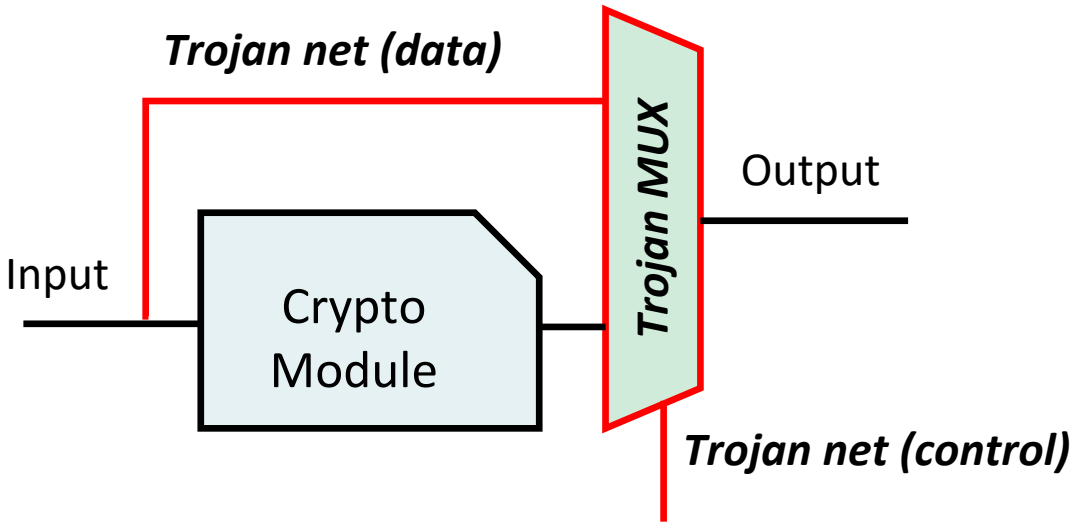


Should the foundry be trusted not to modify the design layout?

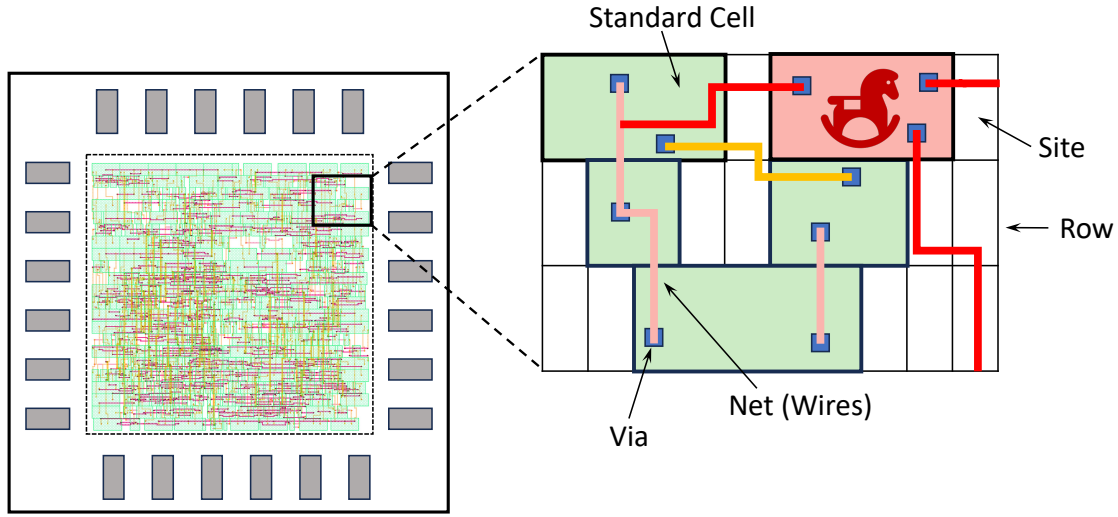
Fabrication-Time Trojan Attacks

What is fabrication-time Trojan?

Malicious IC layout modifications in foundry



Schematic View

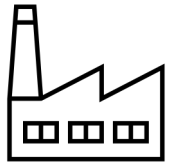


Layout View

Untrusted Foundry Defenses

Prevention

1. Layout Adjustment
2. Split Manufacturing
3. Logic Locking



Fabrication

Detection

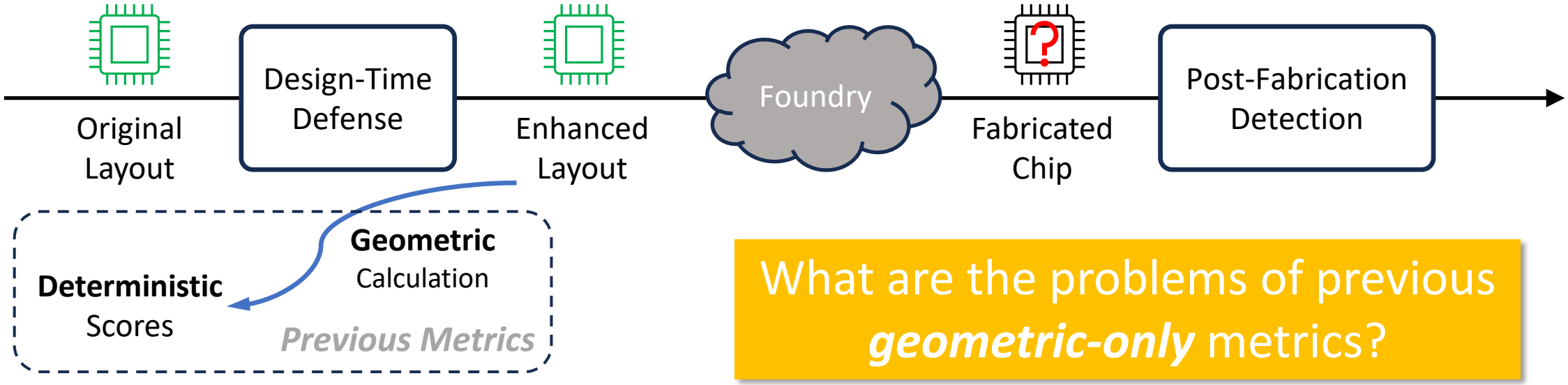
1. Side-Channel
2. Functional Testing
3. Visual Inspection

How to evaluate the vulnerability of a (strengthened) IC layout?

Layout Security Metrics

- Previous metrics
  - Regional vulnerability [Hossein-Talaee *et al.*, ISVLSI'17]
  - Trigger space; net blockage; route distance [Trippel *et al.*, S&P'20]
  - Exploitable place & route resources [Knechtel *et al.*, ISPD'22]

Compute unused layout resources



What are the problems of previous ***geometric-only*** metrics?

Limitations of Previous Metrics: A Case Study

Adder-V1

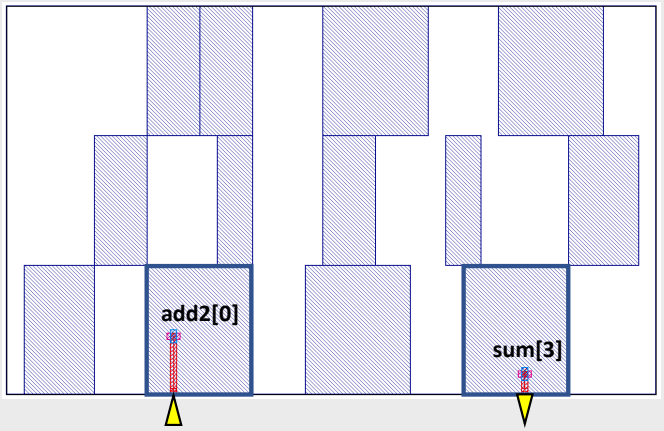
Loosely placed

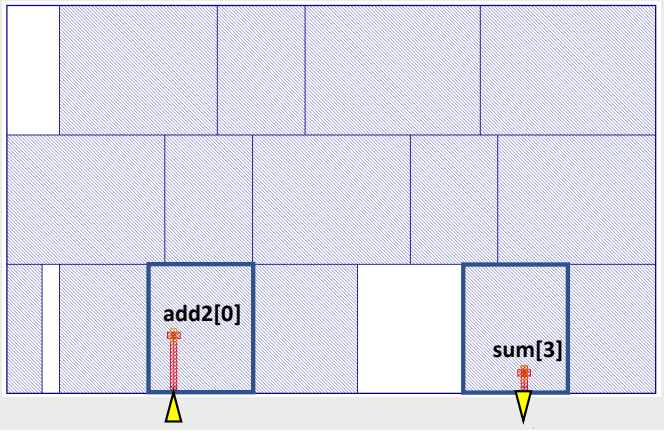
	Prev. Metrics	Reality
Adder-V1	☹	☺
Adder-V2	☺	☹

Adder-V2

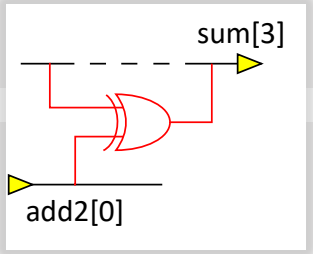
Densely placed

Trojan-free Layouts

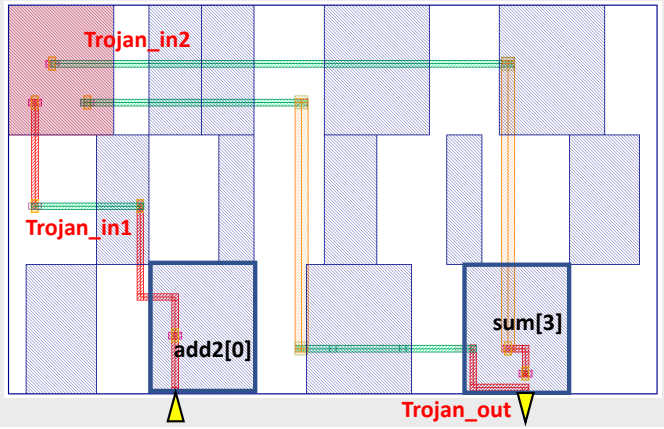


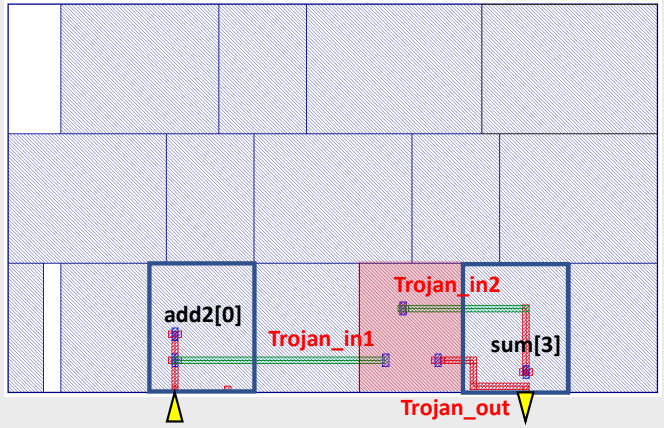


Insert Trojan Gate

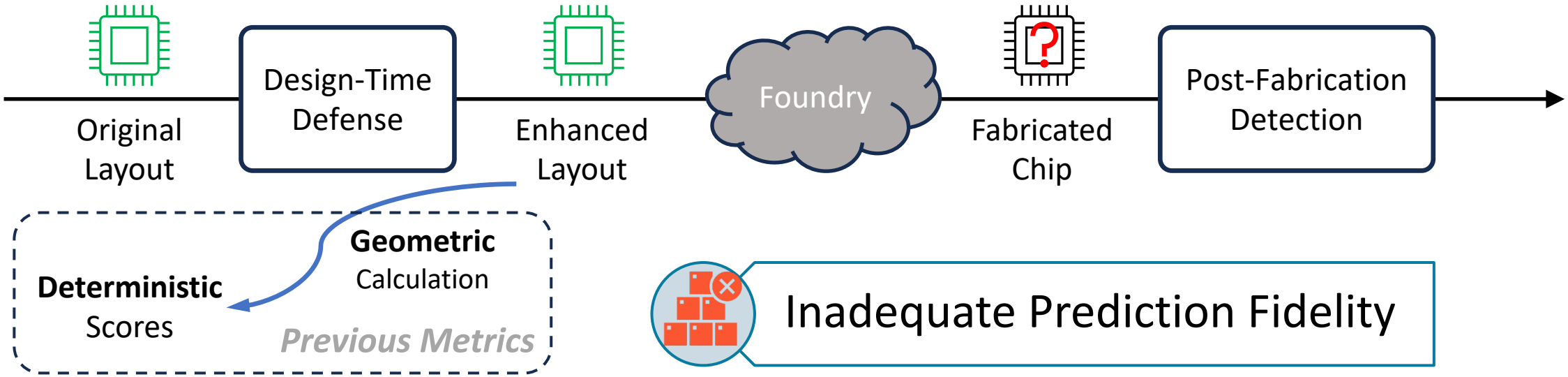


Trojan-in Layouts



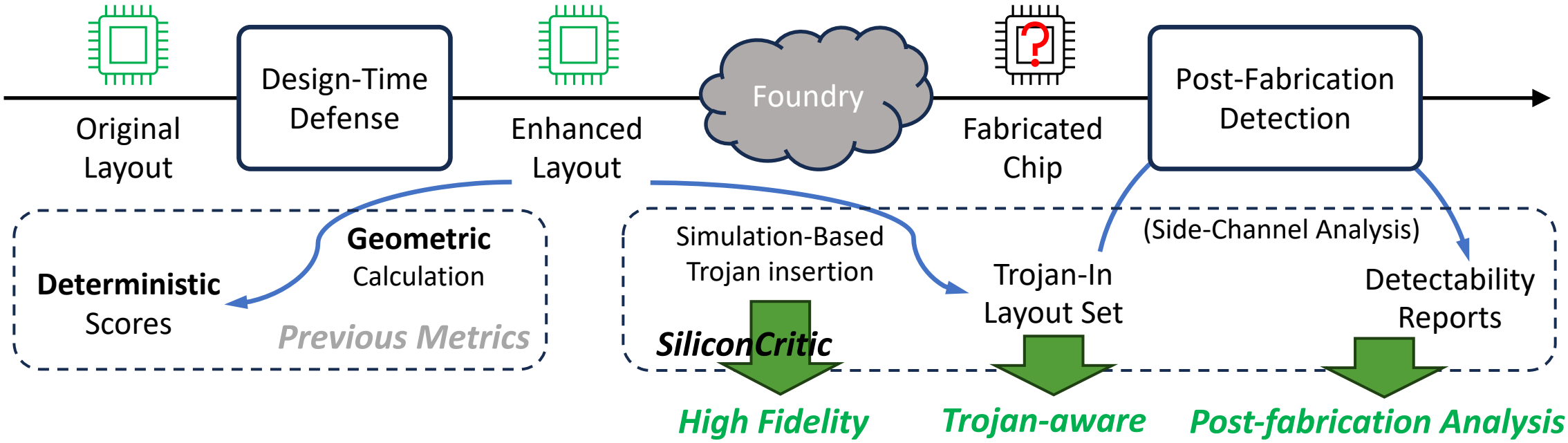


Limitations of Previous Metrics: A Case Study



- Inadequate Prediction Fidelity
- Trojan-Obliviousness
- Lack of Post-Trojan Analysis

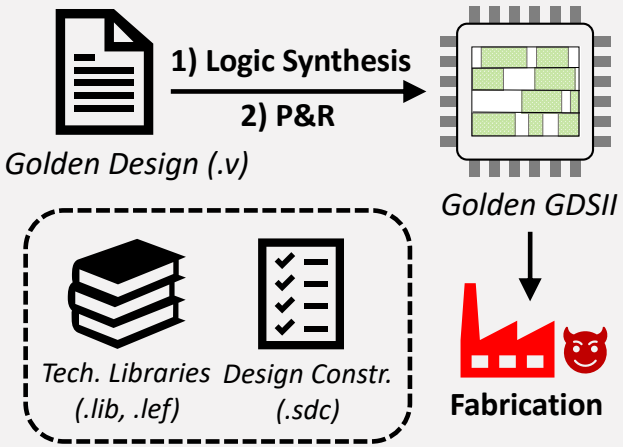
Our Solution



SiliconCritic: From *prior* geometric calculation to *posterior* simulation

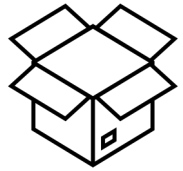
SiliconCritic

Typical IC Process



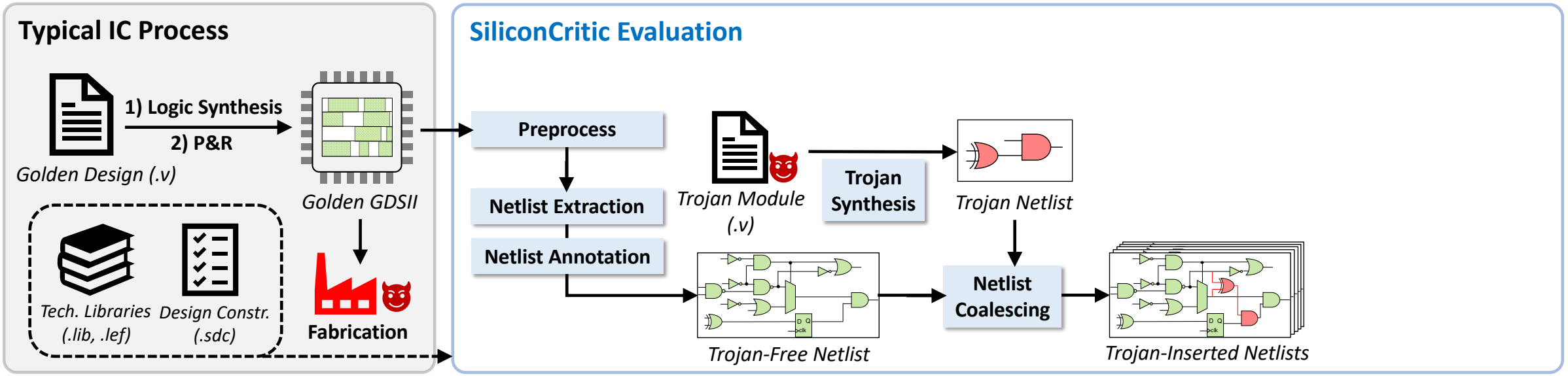
SiliconCritic Evaluation

Blackbox Fabrication-time Trojan Attacks



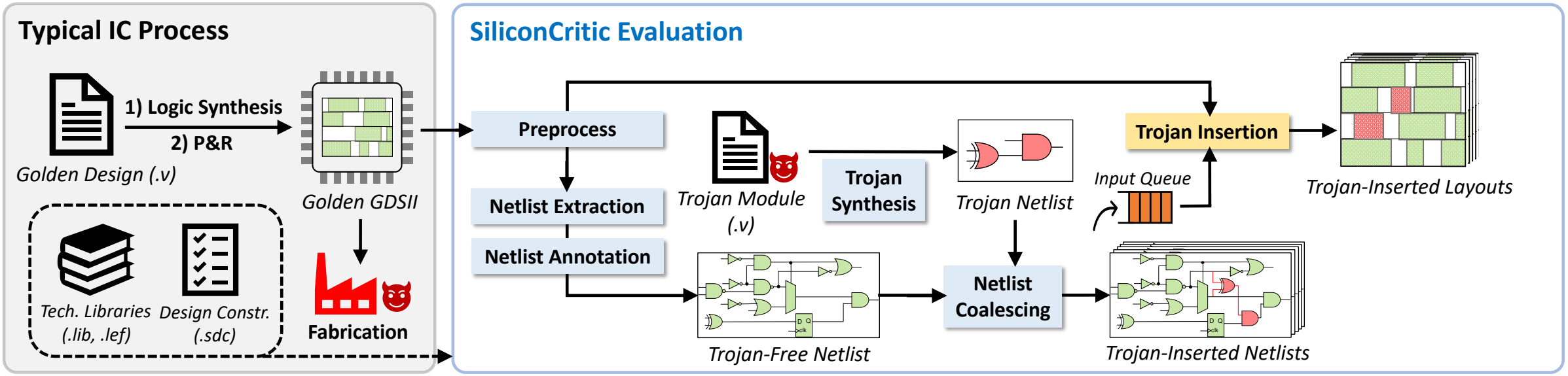
Whitebox Design-time Simulations

SiliconCritic



- **Netlist Preparation**
  - Extract netlist from tapeout-ready layout
  - Locate security-critical signals
  - Append synthesized Trojan for a collection of attack schemes

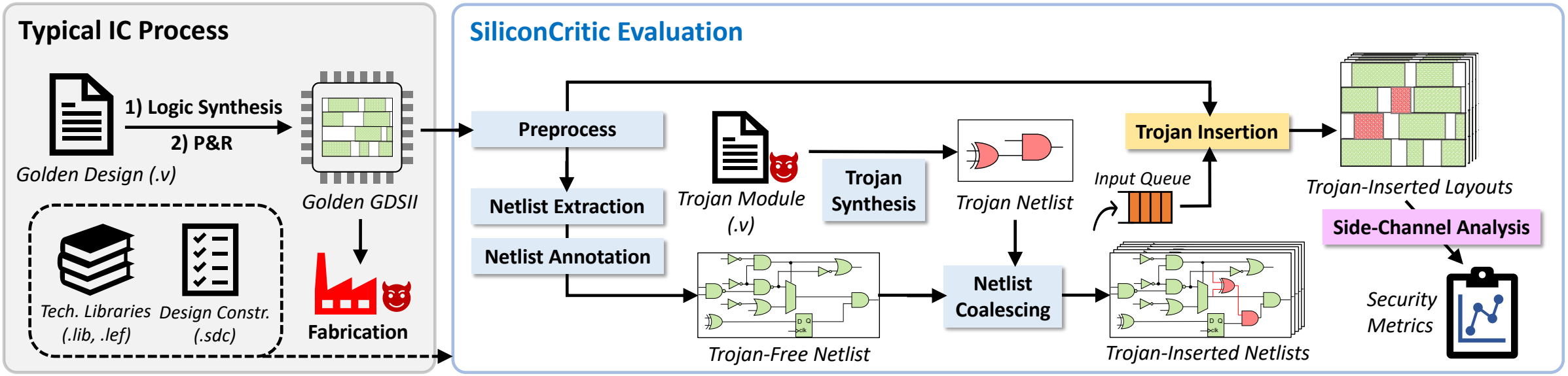
SiliconCritic



Trojan Insertion

- Implement netlist-level Trojan insertion at layout-level in batch
- Attach Trojan cells and nets while maintaining existing layout
- Via Engineering Change Order (ECO) operators provided by commercial CAD tools

SiliconCritic



- Side-channel Analysis
  - Determine layout vulnerability based on Trojan detectability
  - Well-crafted timing/power related metrics

SiliconCritic: Side-Channel Analysis

- Timing Analysis
  - Total Negative Slacks (TNS)
  - Max. Path Delay Rise (**MPDR**)
- Power Analysis
  - Total Power
  - Max. Regional Power Rise (**MRPR**)

$L_{Tj-free}$ / L_{Tj-in}: Layout before/after Trojan insertion
$\boldsymbol{p}$: A critical timing path
$\boldsymbol{r}$: A functional region within the chip core

$$MPDR(L_{Tj-in}) = \max_{\boldsymbol{p} \in \{critical\ paths\}} \frac{delay(\boldsymbol{p}, L_{Tj-in})}{delay(\boldsymbol{p}, L_{Tj-free})}$$

$$MRPR(L_{Tj-in}) = \max_{\boldsymbol{r} \in \{all\ regions\}} \frac{power(\boldsymbol{r}, L_{Tj-in})}{power(\boldsymbol{r}, L_{Tj-free})}$$

Experimental Setup

- Attack ⇒ Design Pairs

Trojan	# Std Cells	Trojan Properties	Design	# Std Cells	Trojan Footprint
Key Leak [1]	80	Sequential, Digital	risc16f84	1290	**6.2016%**
Bus Hijacker [2, 3]	23	Combinational, Digital	Conmax	16537	**0.1391%**
Timebomb [2, 3]	33	Sequential, Digital	AES	189112	**0.0174%**
A2 [4]	2	Combinational, Analog&Digital	OR1200	317296	**0.0006%**

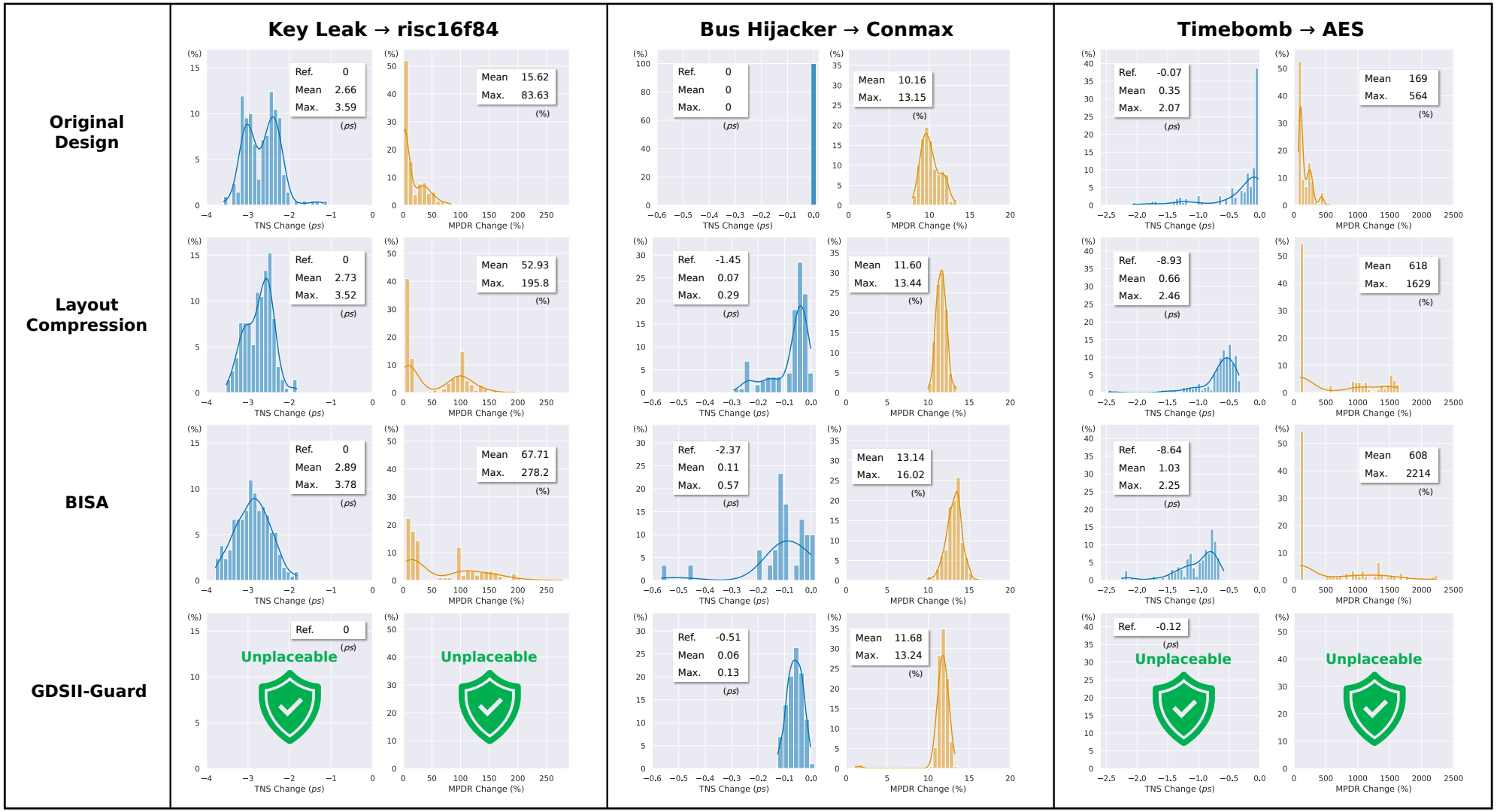
***Trojan footprint**: # Trojan cells / # design cells*

[1] King et al., "Designing and implementing malicious hardware", LEET, 2008.
[2] Salmani et al., "On design vulnerability analysis and trust benchmarks development", ICCD, 2013
[3] Shakya et al., "Benchmarking of hardware Trojans and maliciously affected circuits ," Journal of Hardware and Systems Security, 2017
[4] Yang et al., "A2: analog malicious hardware", IEEE S&P, 2016

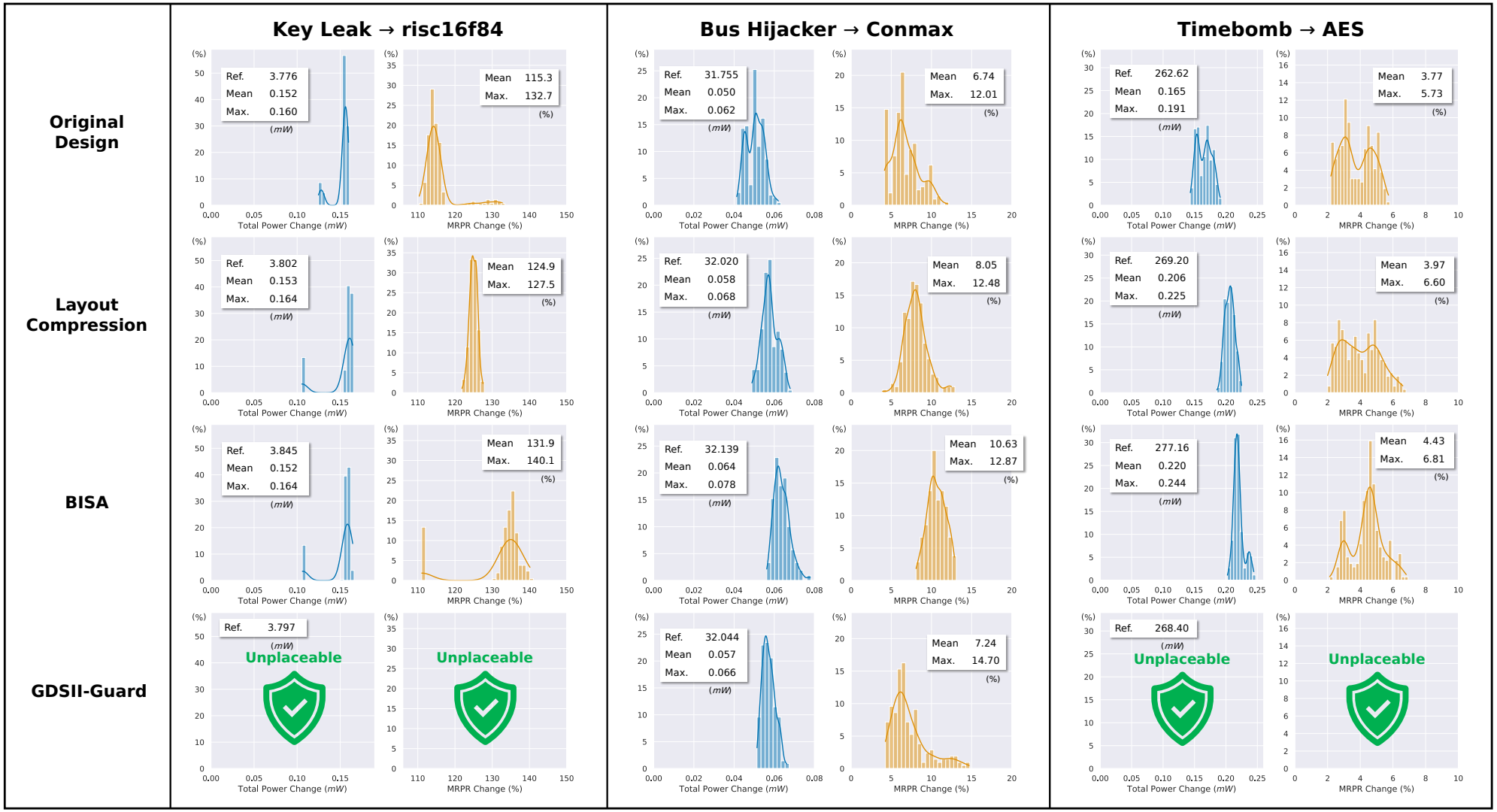
Experimental Setup

- Assessed Defenses
  - **None**
    - Original Design
  - **Layout Compression**
    - Increase core utilization
  - **Built-in Self Authentication (BISA)** [Xiao *et al.*, HOST'13; Ba *et al.*, ISVLSI'16]
    - Occupy unused spaces with tamper-evident logic
  - **GDSII-Guard** [Wei *et al.*, DAC'23]
    - Layout rearrangement to eliminate spatially continuous empty regions

Timing Analysis



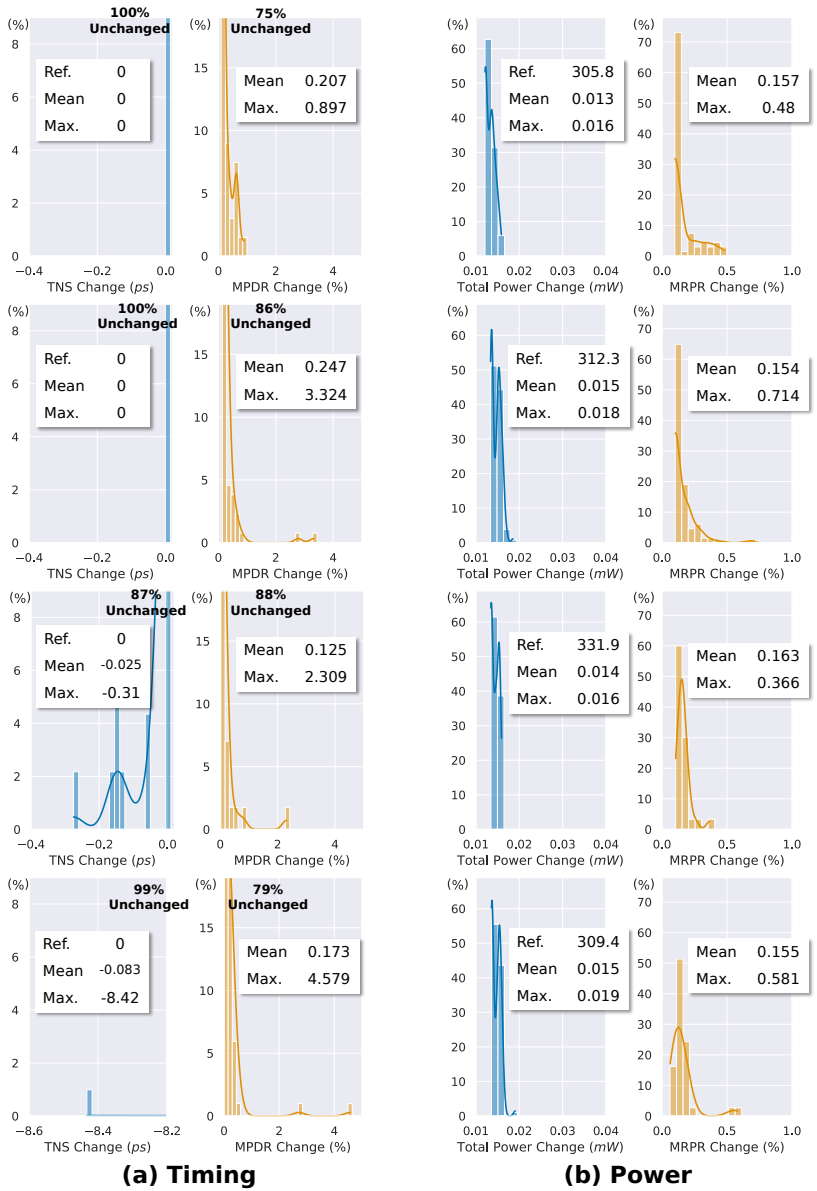
Power Analysis



A2

- The stealthiest hardware Trojan to date [Yang *et al.*, S&P'16]
  - Evades every known detection
  - Its analog trigger is free from timing constraints
  - Its digital payload consists of simply 2 gates

Result: Timing/Power variations are *negligible*



Conclusion

- SiliconCritic can shed light on future design-time defenses
  - IC designers can tailor their defenses for the most adverse Trojans to the design
  - For A2, *routing-centric* defenses will be more promising than *placement-centric*
- *Absolute* defense does not exist
  - Protections should increase attack difficulty, instead of completely blocking Trojan insertion at the expenses of performance, power, or area (PPA)
- SiliconCritic can be extended to
  - 1) Customized side-channel metrics, 2) advanced process, 3) various CAD tools 4) attacker's perspective
- Limitation of SiliconCritic
  - Process variation introduces gap between simulation and fabrication

北京大学
PEKING UNIVERSITY

北京大学高能效计算与应用中心
Center for Energy-efficient Computing and Applications

Thank you!

Welcome to my poster for more details

weixinming@pku.edu.cn

https://github.com/xinming-wei/SiliconCritic