Bridging the Gap: Advanced Tools for Side-Channel Leakage Estimation beyond Gaussian Templates and Histograms

Tobias Schneider[1], Amir Moradi[1],
François-Xavier Standaert[2], Tim Güneysu[3]

[1] Ruhr-Universität Bochum
[2] Université catholique de Louvain
[3] University of Bremen & DFKI

UbiCrypt
Cryptography in Ubiquitous Computing

hgi Horst Görtz Institute for IT-Security
RUHR-UNIVERSITÄT BOCHUM

Wednesday, July 13th, 2016

Outline

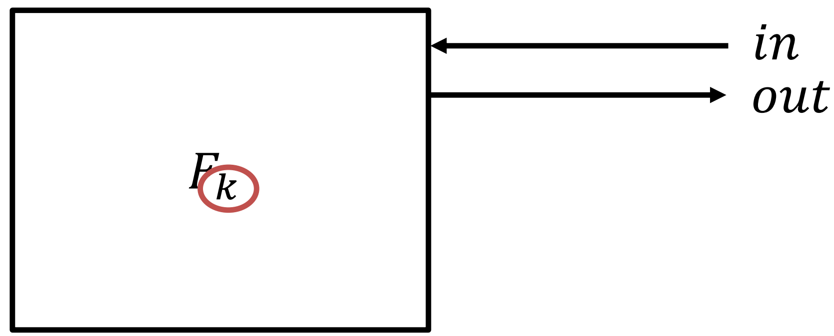
- Introduction
- Background
- New Tools
- Results and Comparison
- Conclusion

Outline

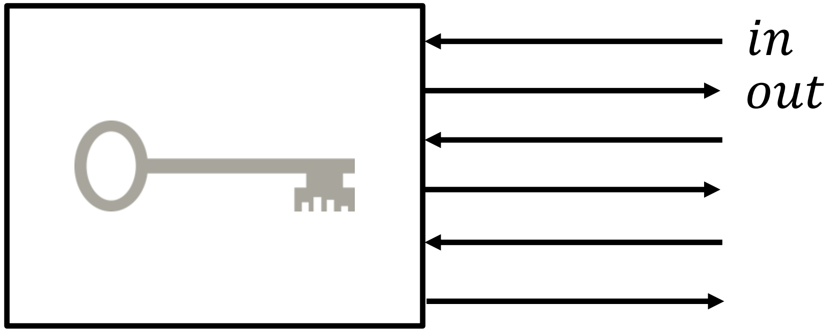
- **Introduction**
- Background
- New Tools
- Results and Comparison
- Conclusion

Side-Channel Analysis

Side-Channel Analysis



Side-Channel Analysis



Side-Channel Analysis

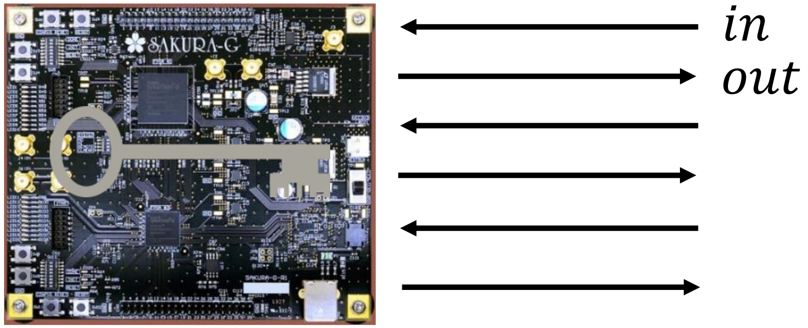


Image from http://satoh.cs.uec.ac.jp/SAKURA/hardware/SAKURA-G.html

Side-Channel Analysis

Timing

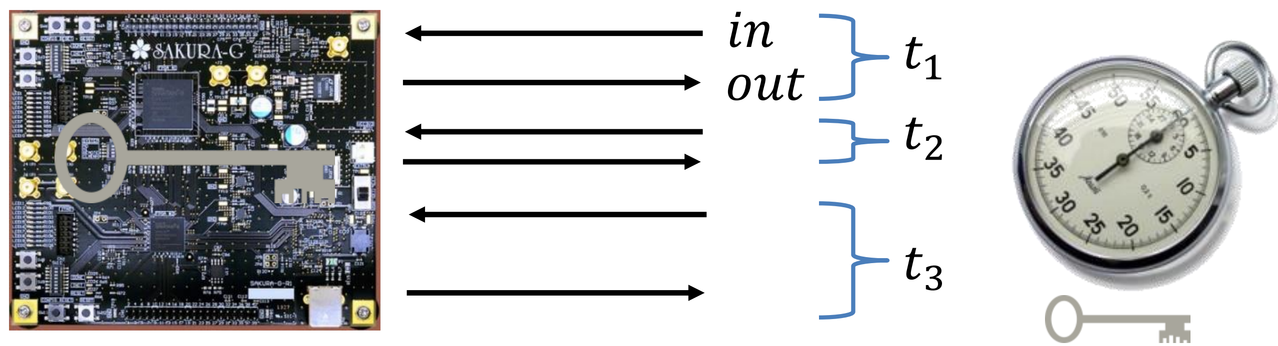


Image from http://satoh.cs.uec.ac.jp/SAKURA/hardware/SAKURA-G.html

Side-Channel Analysis

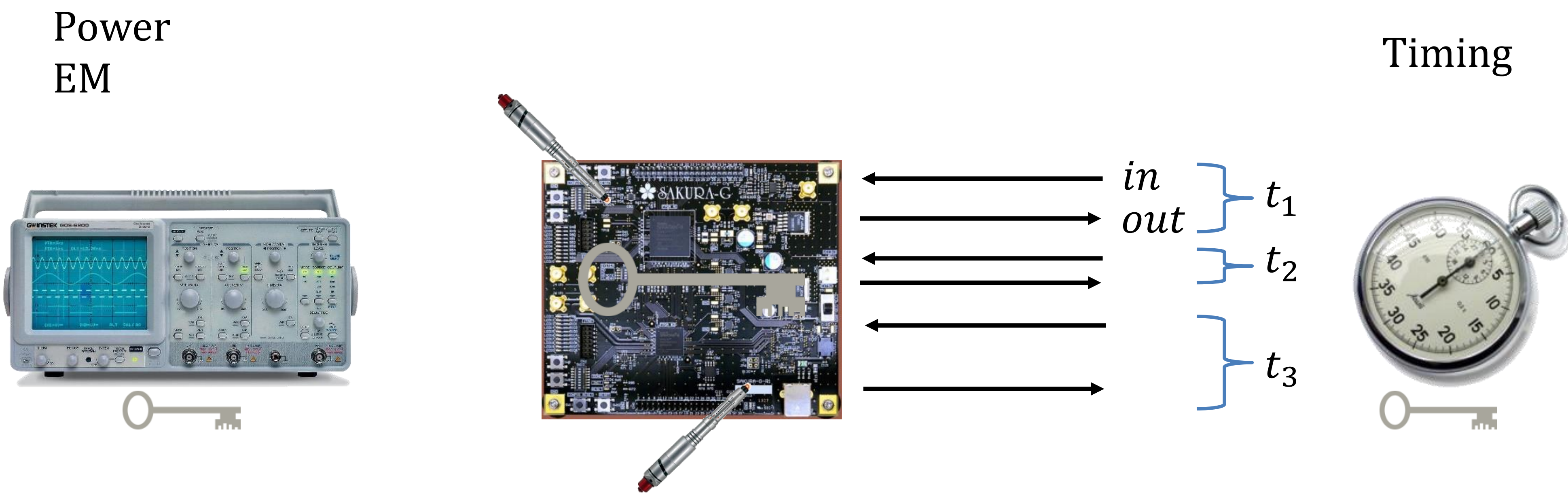
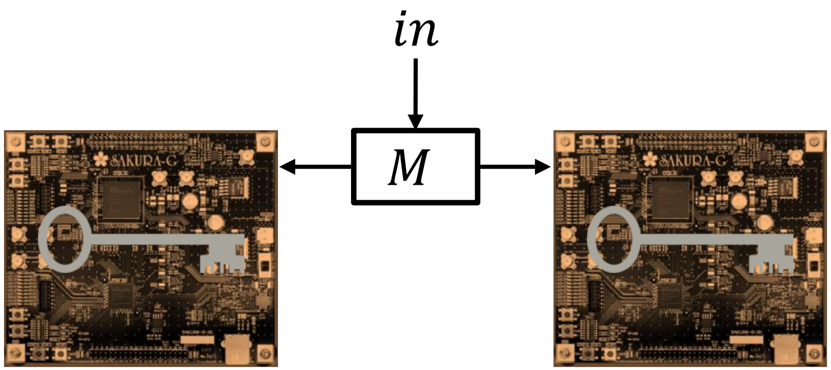


Image from http://satoh.cs.uec.ac.jp/SAKURA/hardware/SAKURA-G.html

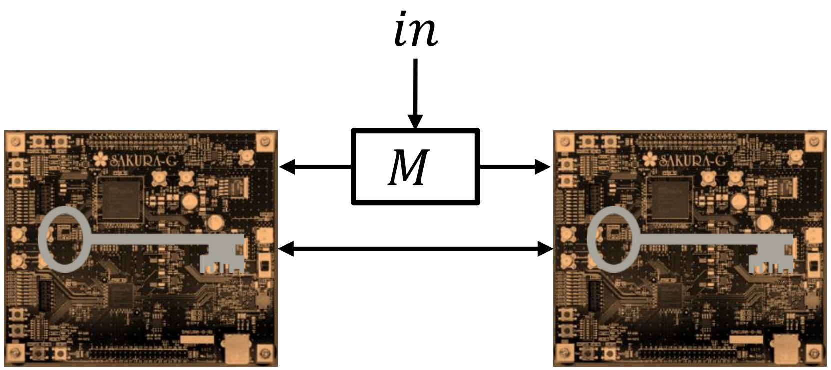
Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



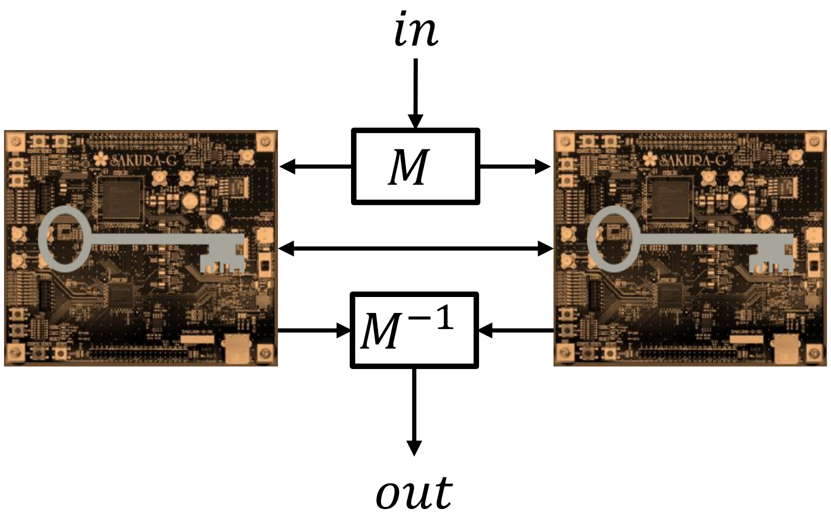
Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



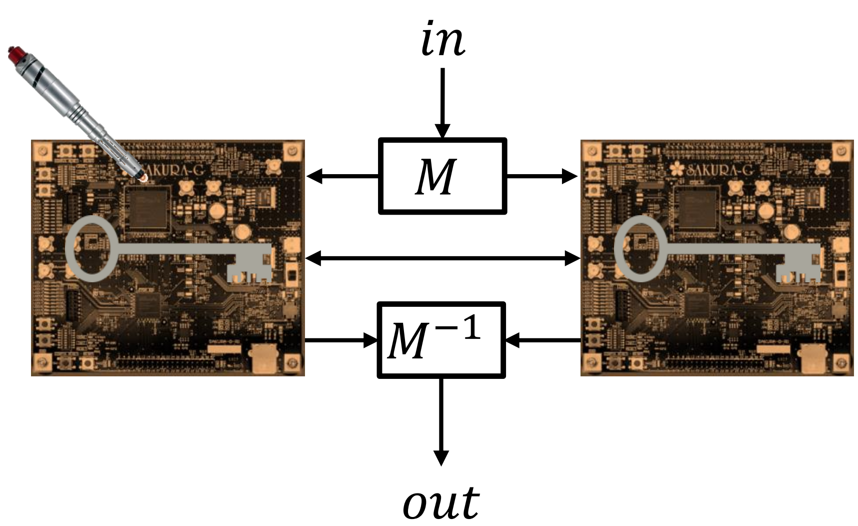
Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



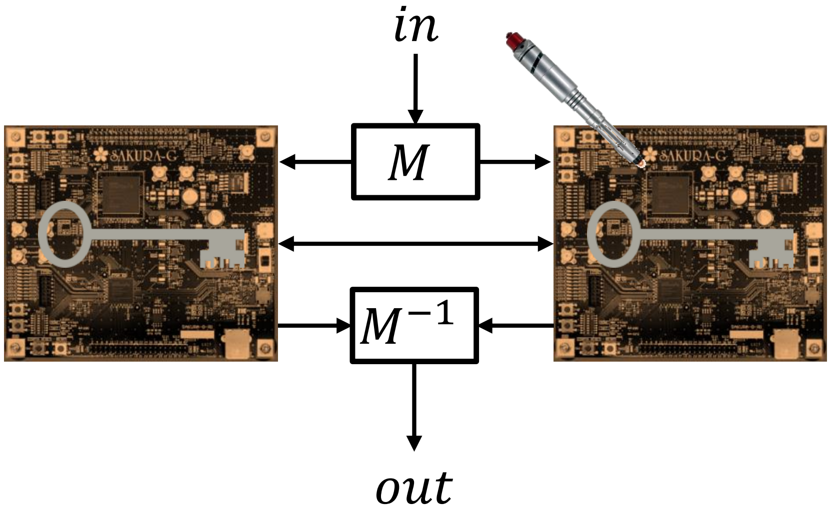
Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



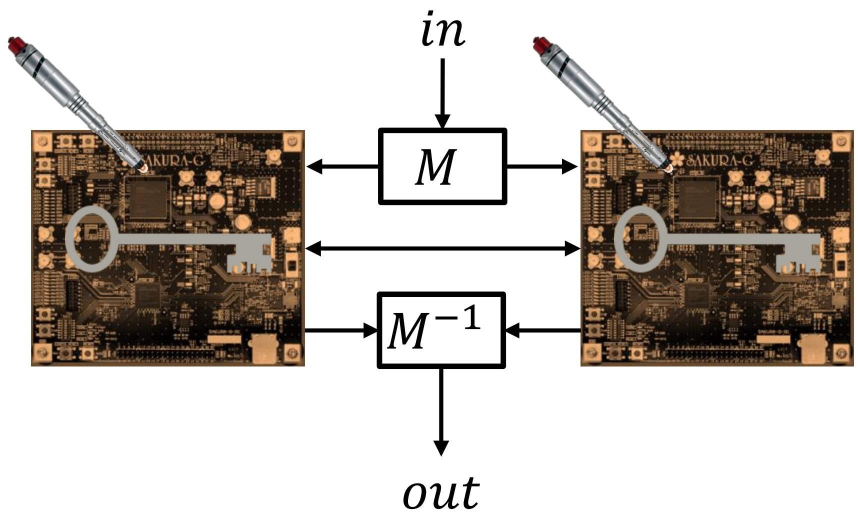
Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



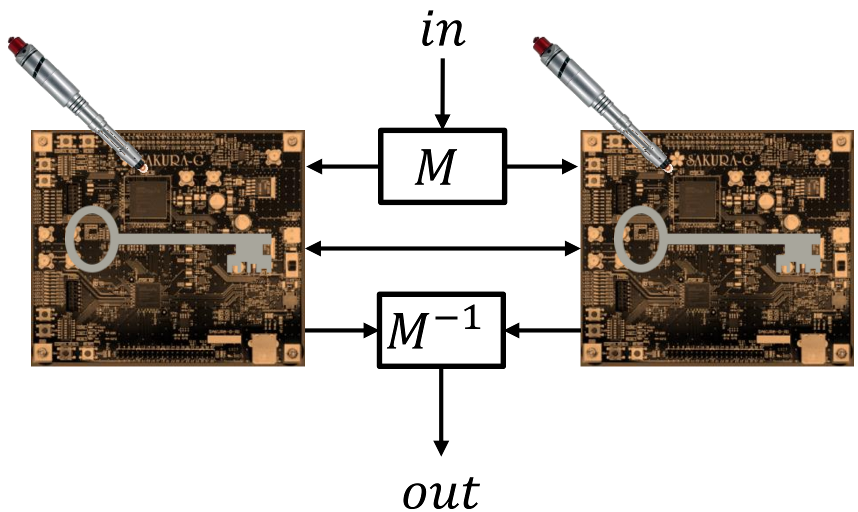
Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



Masking

- Most investigated and best understood protection against SCA
- Every sensible variable is encoded into d shares
- Computation is performed on these shares
- Attacker (ideally) needs to combine leakage of all shares to extract information



- **Problem:** Schemes require significant overhead

Leakage Assessment

- Compare security and performance on a sound basis
- Various different evaluation methodologies
- Some require estimation of leakage Probability Density Function (PDF)

	Profiled	**Non-Profiled**
PDF-Based	Template Attack	MIA
Per Moment	MCP-DPA	MCC-CPA t-Test

- Comprehensive understanding of the leakage behavior is essential
  - E.g., Threshold Implementations (TI) can require more shares to achieve d-th order security due to glitches
- t-test-based leakage detection gives only limited information

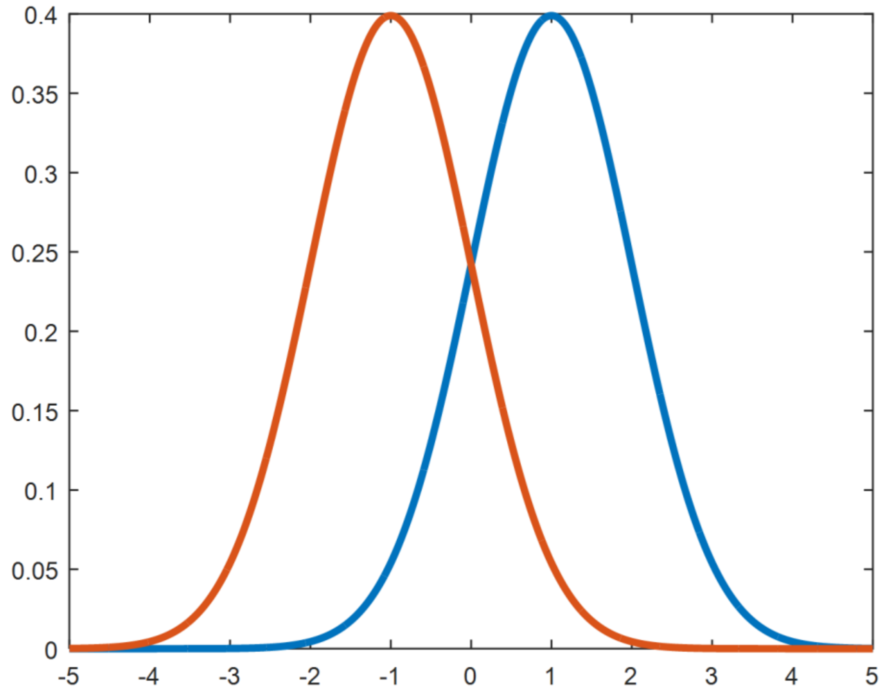
Our Contribution

1) Extend SCA evaluation toolbox with three PDF estimation tools
    - Current state-of-the-art tools used for SCA have limited applicability or slow convergence
2) Introduce per-moment computation for our PDF-based methods and attacks that use a combination of multiple moments
    - Enable thorough leakage profiling
    - More efficient attacks
3) Analyze masked HW design of PRESENT as a case study
    - Profiled setting
    - Non-profiled setting

Outline

- Introduction
- **Background**
- New Tools
- Results and Comparison
- Conclusion

Density Estimation



- Leakage PDF gives information about $\Pr(l|s)$ where l is the leakage for a specific sensible variable s
- Exact PDF is unknown but can be estimated using measurements

Density Estimation

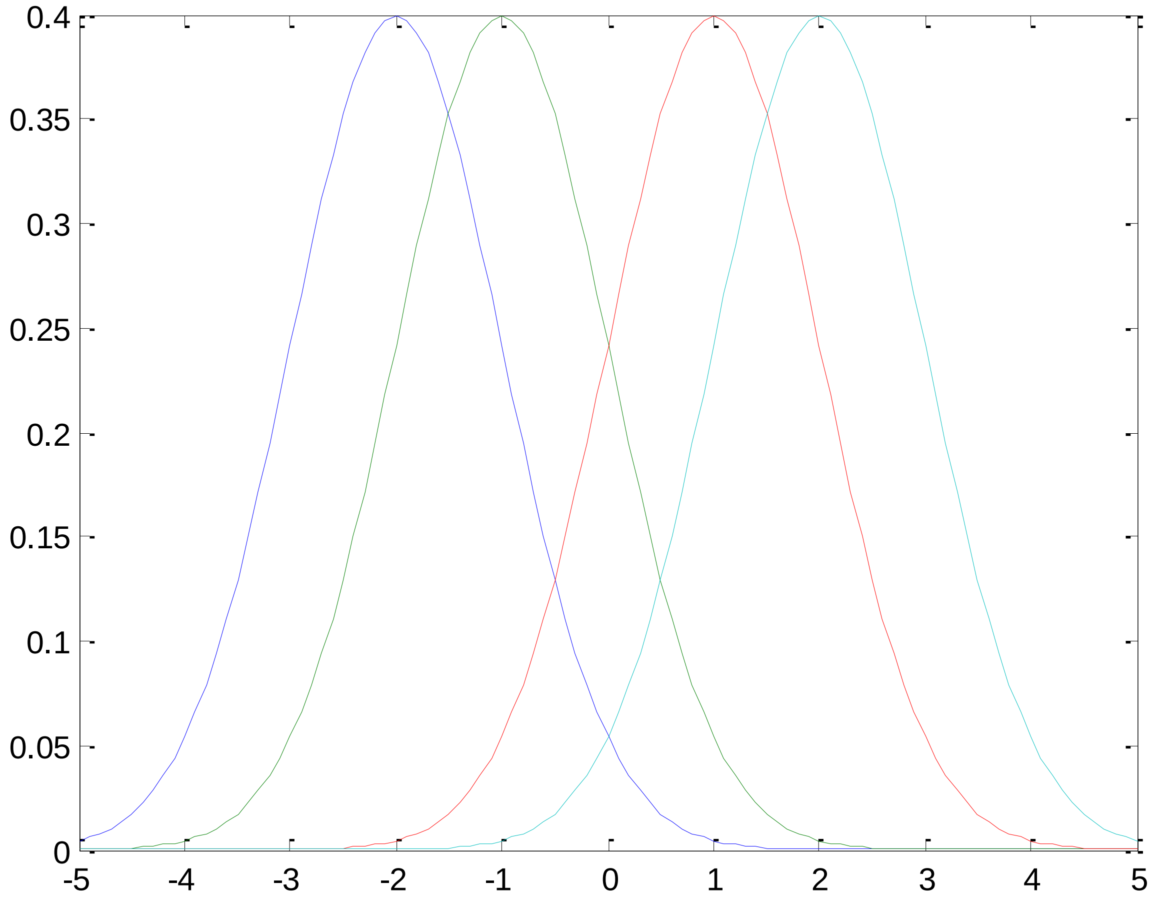
- Two major categories: *Non-Parametric* and *Parametric*
- **Non-parametric**
  - No assumptions about the form of the distribution
  - Examples: histogram, kernel
- **Parametric**
  - Assumes certain distribution form (e.g., symmetric)
  - Example: Gaussian distribution
  - Can be parametrized with statistical moments

$$M_d = E\left(X^d\right) \text{ (Raw Moments, } d \geq 1\text{)}$$

$$CM_d = E\left((X-\mu)^d\right) \text{ (Central Moments, } d \geq 2\text{)}$$

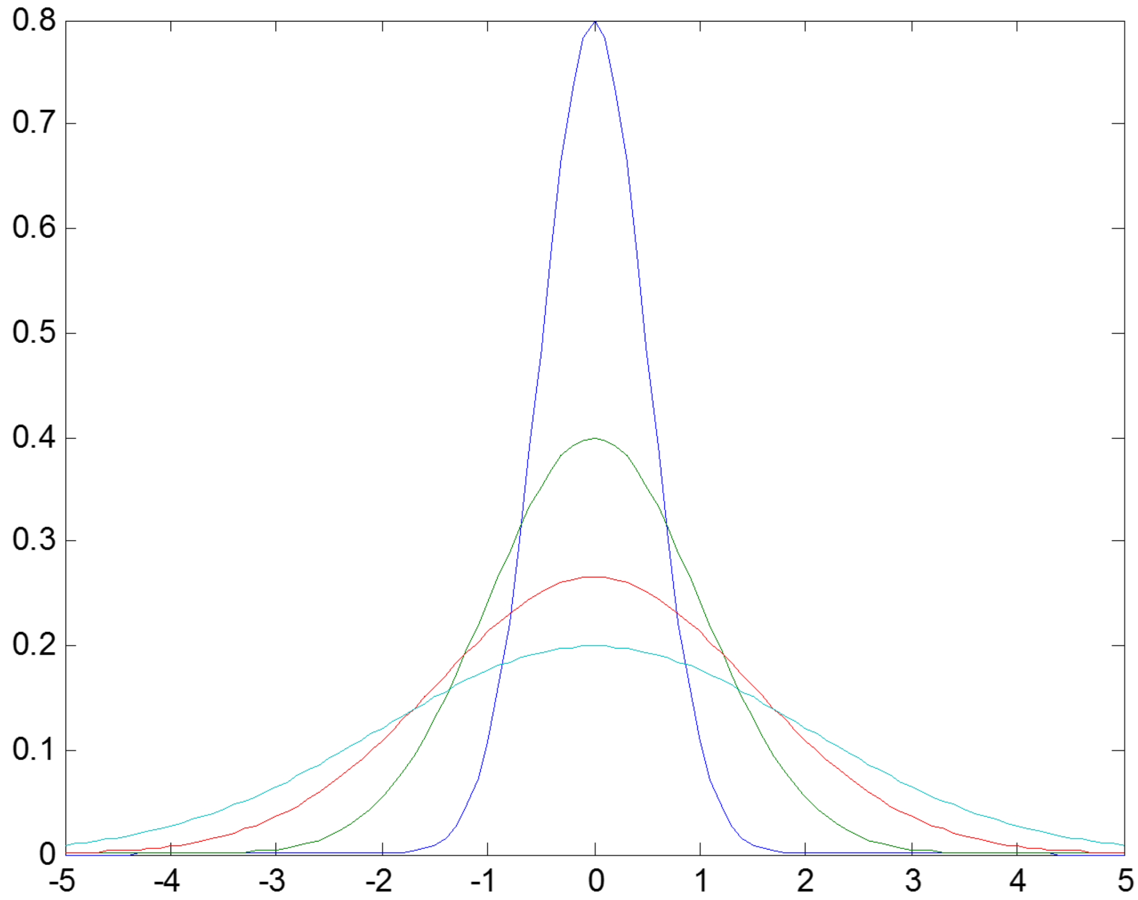
$$SM_d = E\left(\left(\frac{X-\mu}{\sigma}\right)^d\right) \text{ (Standarized Moments, } d \geq 3\text{)}$$

Statistical Moments



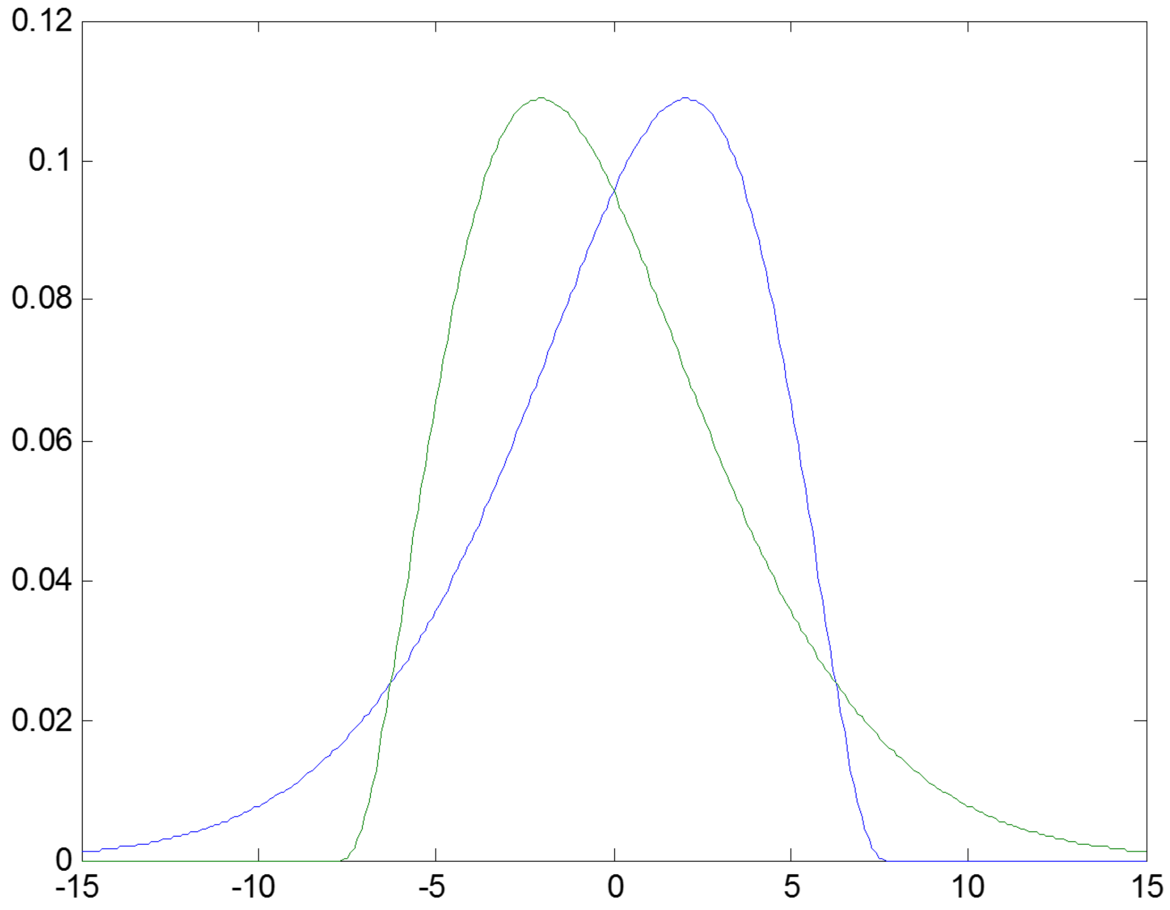
1st moment: Mean (raw)

Statistical Moments



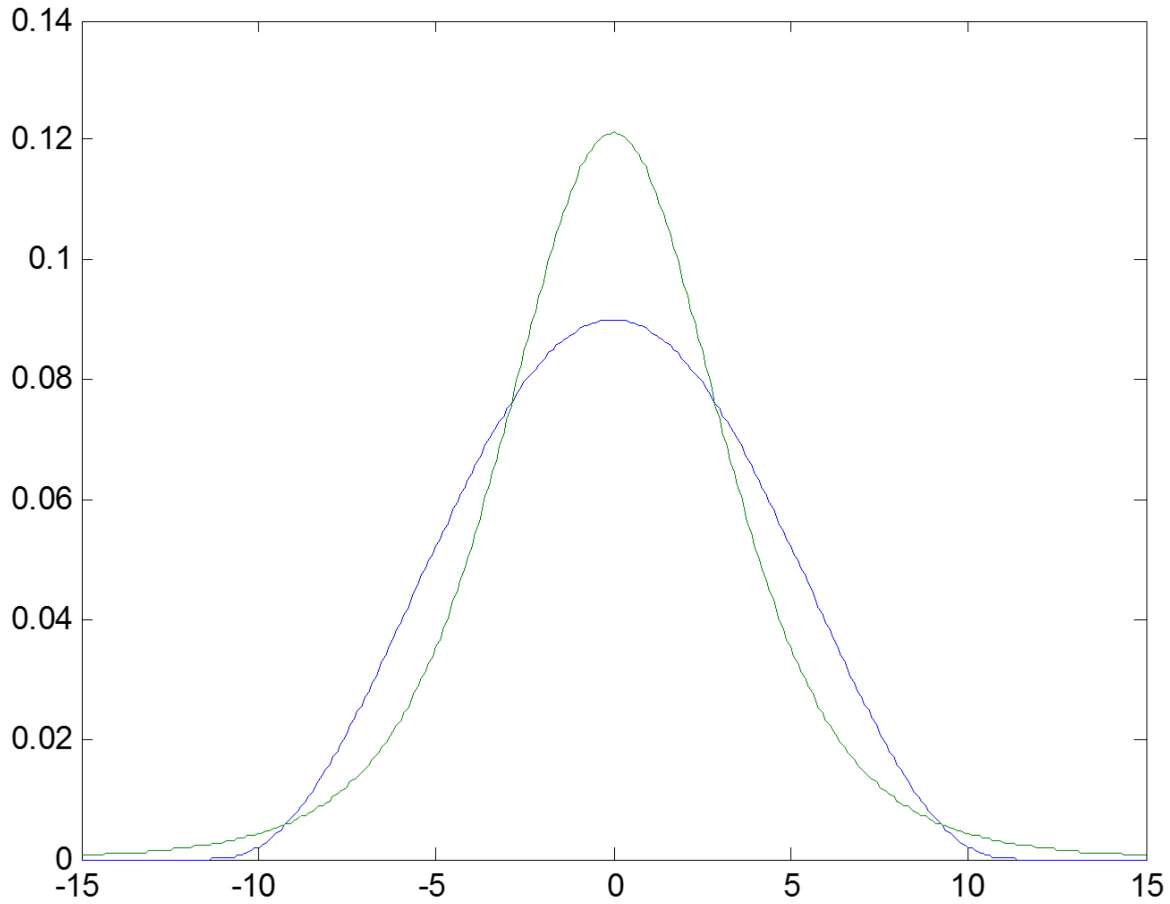
2nd moment: Variance (central)

Statistical Moments



3rd moment: Skewness (standardized)

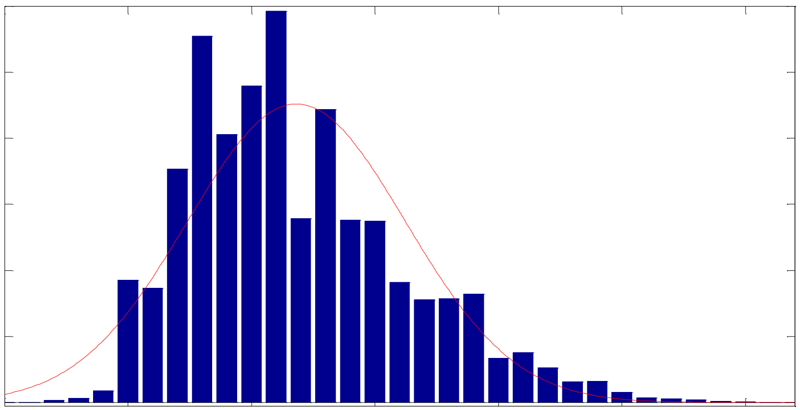
Statistical Moments



4th moment: Kurtosis (standardized)

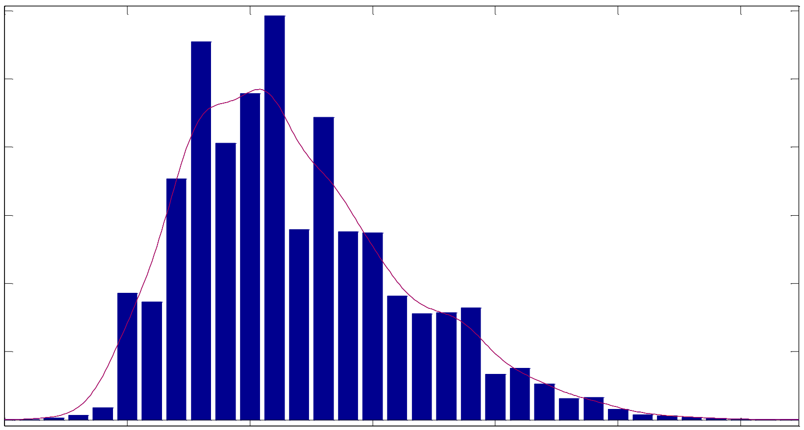
Density Estimation

Gaussian



- Assumes leakage follows a Gaussian distribution
- PDF: $F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Distributions considers only first two moments (μ,σ)

Kernel



- Approximate PDF as sum of kernel functions
- PDF: $F(x) = \frac{1}{n\,h} \sum_{i=0}^{n-1} K\left(\frac{x - l_i}{h}\right)$
- Considers all available leakage
- Parameters: bandwidth h, kernel function $K(.)$

Problems

- **Gaussian**
  - Fast and efficient
  - Not suited for implementations with more than two shares
- **Histogram/Kernel**
  - Can estimate all types of leakage PDF
  - Slow convergence
  - No intuitions about separate moments
- **Our new tools**
  - Faster convergence than kernels
  - Higher flexibility than Gaussian
  - Consider more than the first two moments (up to four)

Outline

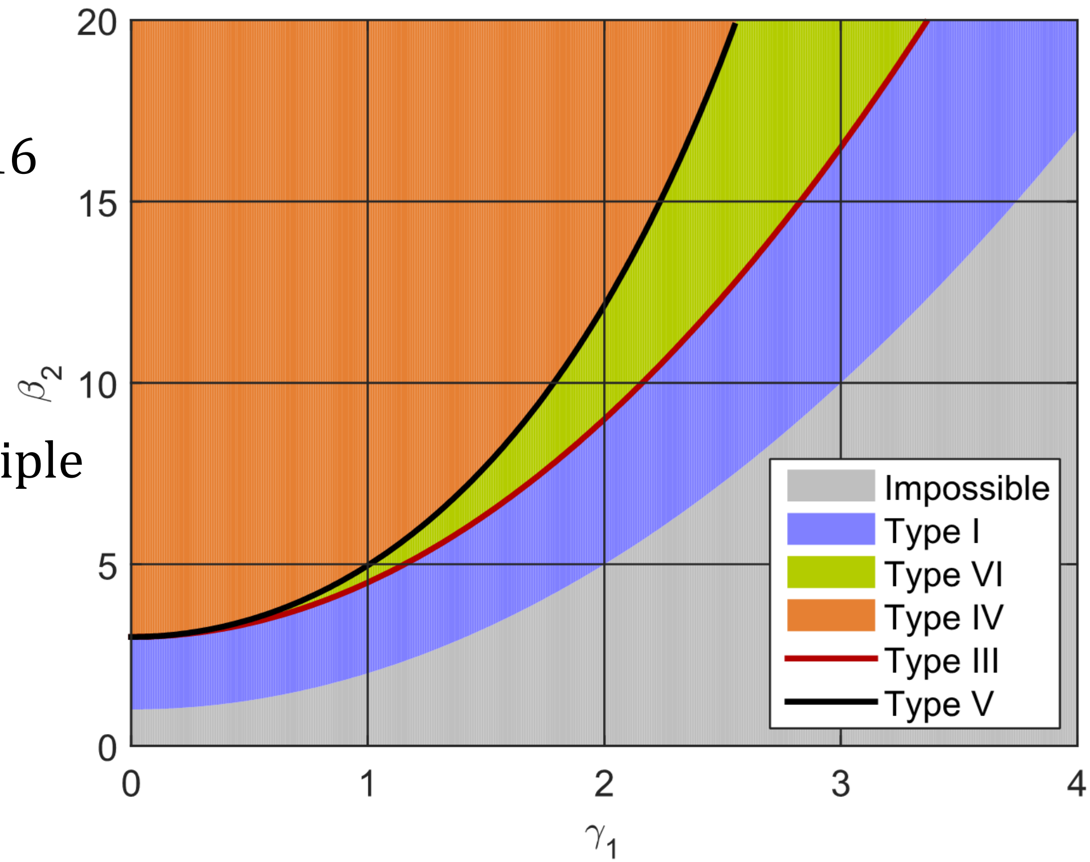
- Introduction
- Background
- **New Tools**
- Results and Comparison
- Conclusion

Exponentially Modified Gaussian

- Exponentially Modified Gaussian (EMG) distribution has been used in other fields (e.g., psychology, physics)
- Similar to Gaussian, but with non-zero skewness (three moments)
- PDF: $F(x) = \frac{\lambda_3}{2} e^{\frac{\lambda_3}{2}(2\lambda_1 + \lambda_3 \lambda_2^2 - 2x)} erfc\left(\frac{\lambda_1 + \lambda_3 \lambda_2^2 - x}{\sqrt{2}\lambda_2}\right)$
- Complementary error function: $erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt.$
- $\lambda_1, \lambda_2, \lambda_3$ can be efficiently computed from the first three moments

Pearson Distribution System

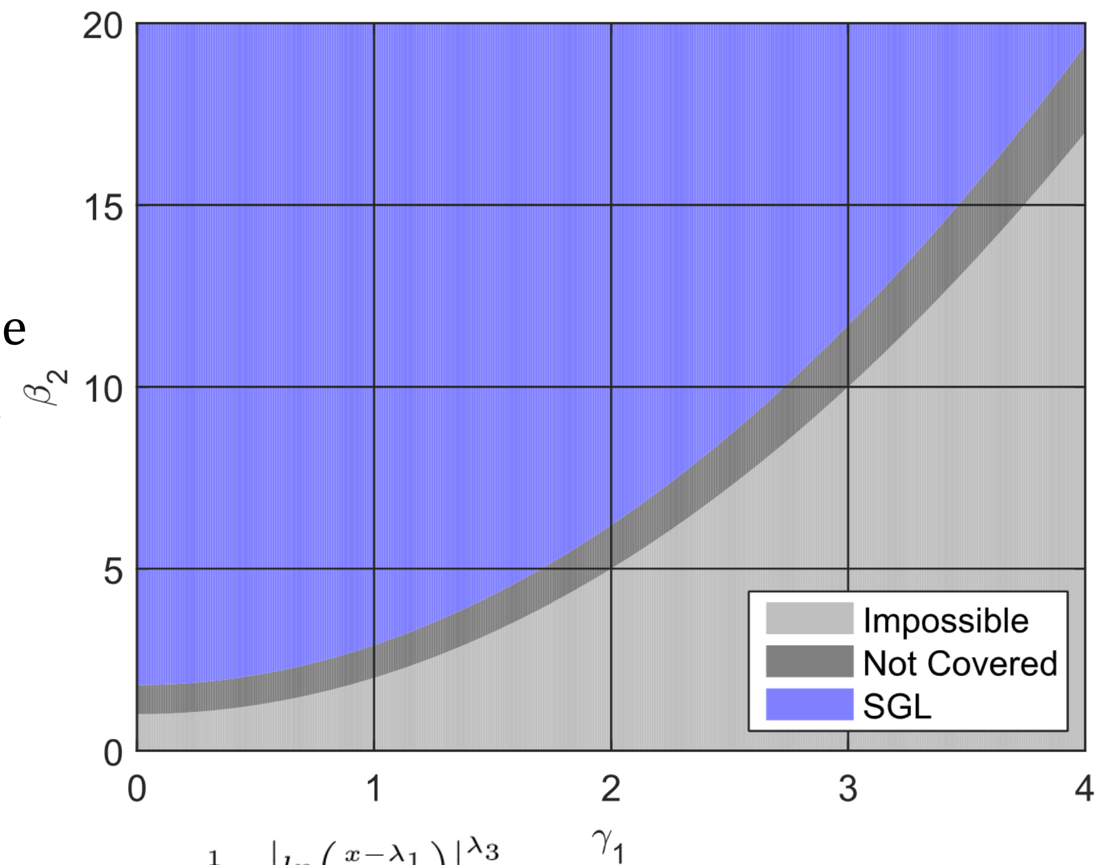
- System of twelve distributions introduced by Pearson in 1895-1916
- Type determined by four moments
- We only used types I, IV, VI



Problem: Requires estimation of multiple PDFs and may face stability issues at transitions between types

Shifted Generalized Lognormal

- SGL introduced by Low in 2013
- Alternative to Pearson
- $\lambda_1, \lambda_2, \lambda_3$ can be computed from the first four statistical moments using Newton's method



- PDF: $F(x) = \frac{1}{2\lambda_3^{1/\lambda_3}\lambda_4 \Gamma(1+1/\lambda_3)(x-\lambda_1)} e^{-\frac{1}{\lambda_3 \lambda_4^{\lambda_3}}\left|ln\left(\frac{x-\lambda_1}{\lambda_2}\right)\right|^{\lambda_3}}$

Comparison

Performance:

- 100 randomly generated sets of moments
- Average computation time over 1000 executions on Intel i5-4200M CPU

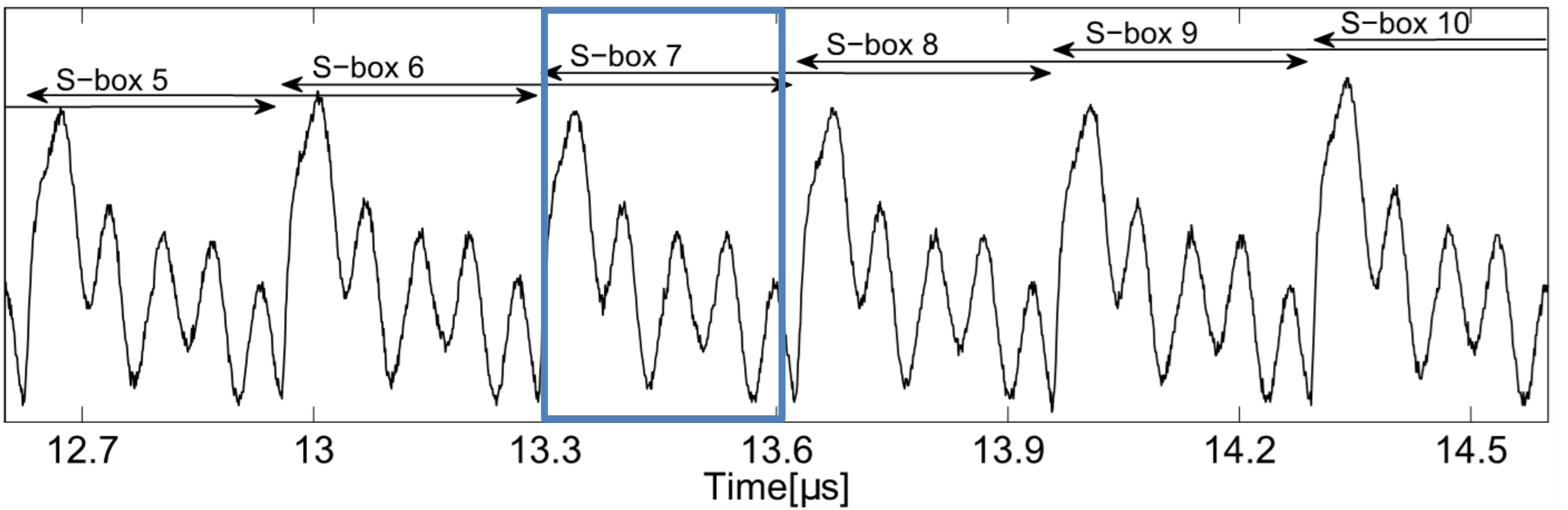
Gaussian	EMG	Pearson	SGL
0.0034 s	0.0082 s	0.029 s	1.70 s

Outline

- Introduction
- Background
- New Tools
- **Results and Comparison**
- Conclusion

Case Study: PRESENT TI

- Threshold implementation of PRESENT
- 1st-order secure with three shares
- 100,000,000 measurements on SASEBO (Xilinx Virtex-II Pro)

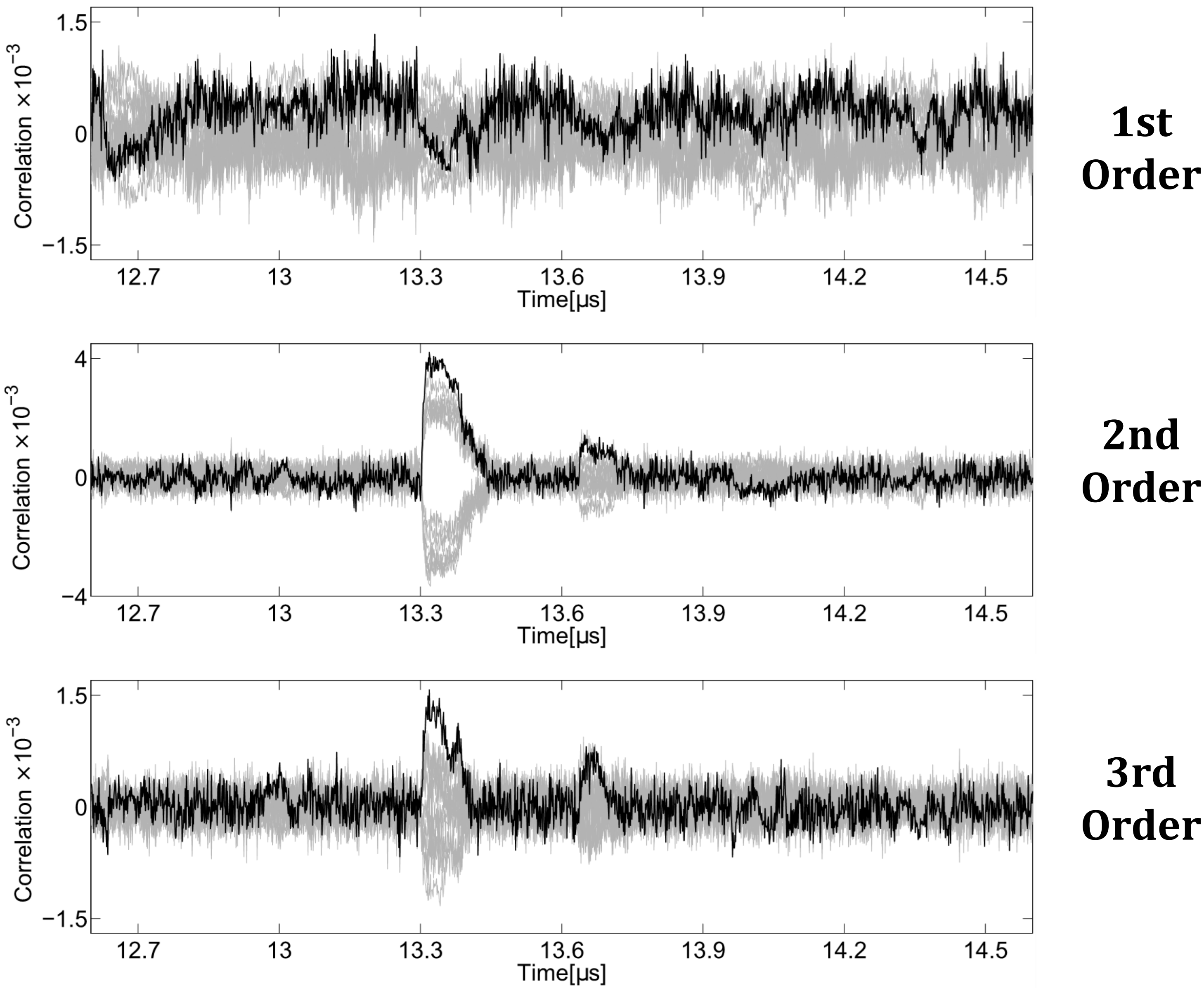


Case Study: PRESENT TI

MCP-DPA by Moradi and Standaert in 2014

Open Questions:

1) Information on a more formal basis
2) Attacking multiple moment jointly



Profiled Evaluation & Attacks

- Information-theoretic metric introduced by Standaert *et al.* in 2009
- Based on mutual information (MI) between sensible variable *S* and leakage *L*
- Later refined to perceived information (PI) to incorporate estimated leakage distributions
- Linked with the success rate of profiled attacks by Duc *et al.* in 2015

$$\hat{P}I(S;L) = H[S] - \sum_{s \in \mathcal{S}} Pr[s] \sum_{l \in \mathcal{L}} \boxed{Pr_{\text{chip}}[l|s]} \cdot log_2 \boxed{\hat{P}r_{\text{model}}[s|l]}$$

Profiled Evaluation & Attacks

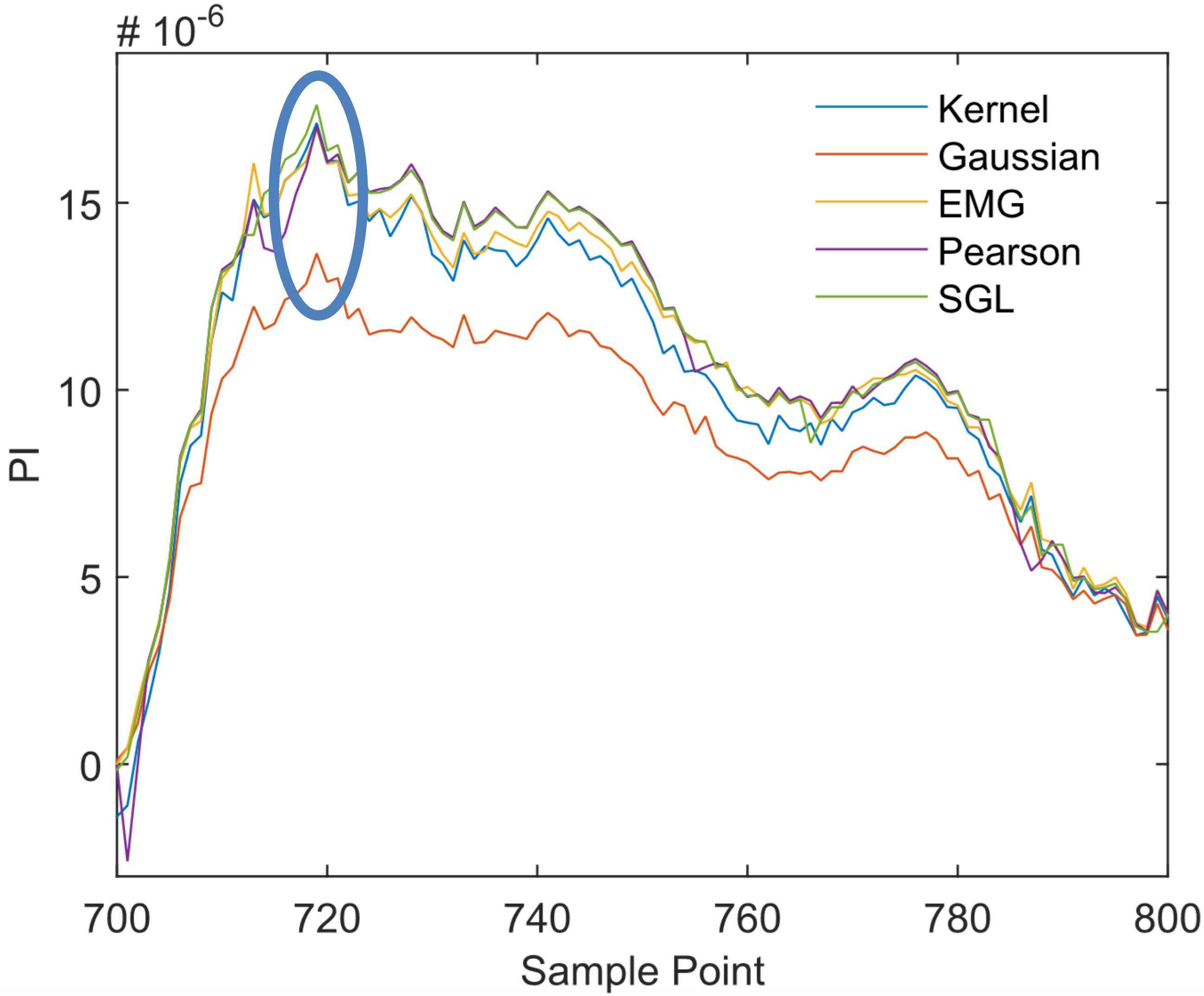
- Information-theoretic metric introduced by Standaert *et al.* in 2009
- Based on mutual information (MI) between sensible variable S and leakage L
- Later refined to perceived information (PI) to incorporate estimated leakage distributions
- Linked with the success rate of profiled attacks by Duc *et al.* in 2015

$$\hat{PI}(S;L) = H[S] - \sum_{s \in \mathcal{S}} Pr[s] \sum_{l \in \mathcal{L}} \boxed{Pr_{\text{chip}}[l|s]} \cdot log_2 \boxed{\hat{Pr}_{\text{model}}[s|l]}$$

- 10-fold cross-validation (90M for model estimation, 10M for chip distr.)

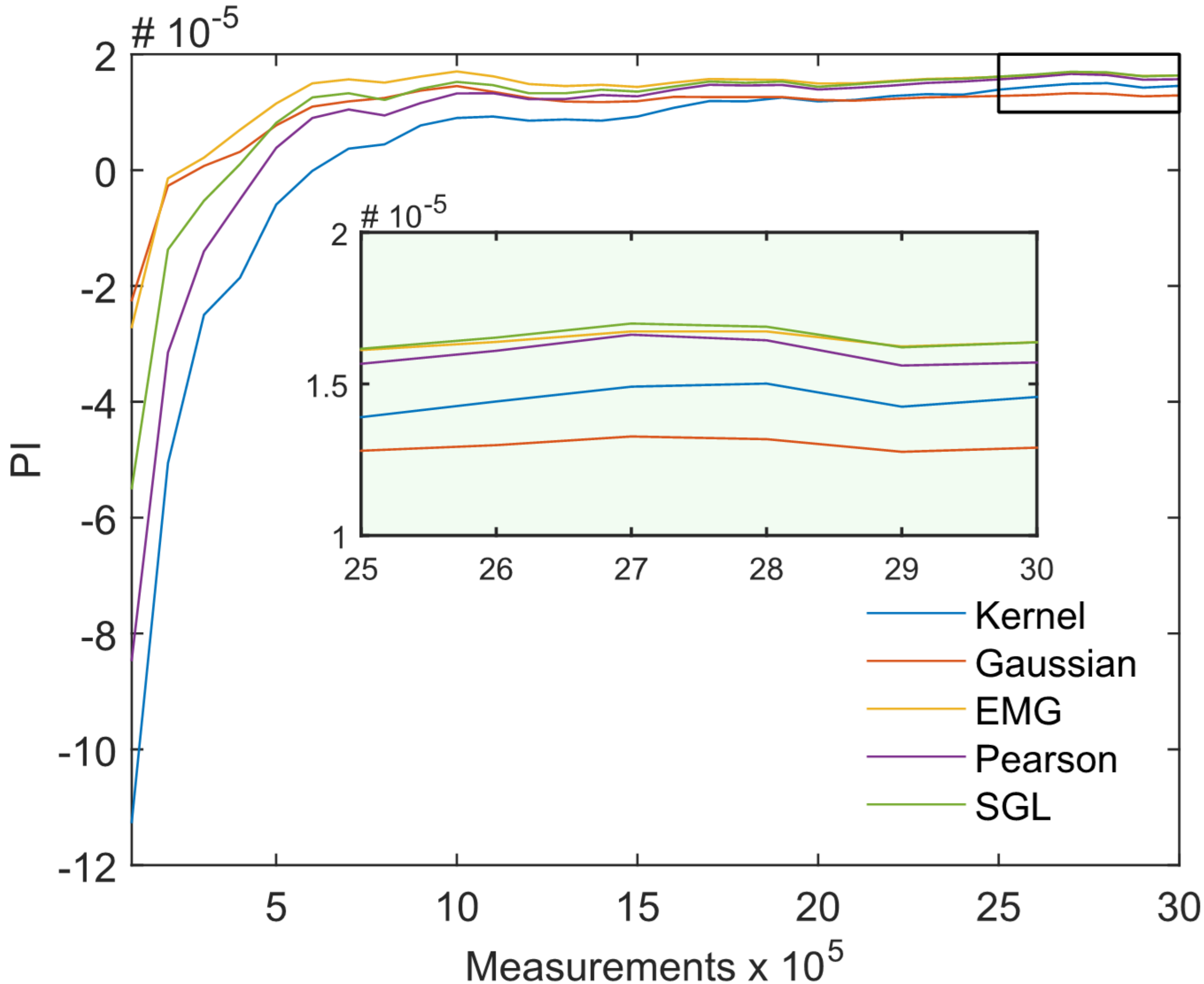
Profiled Evaluation & Attacks

Combined Moments



Profiled Evaluation & Attacks

Combined Moments (Sample Point 719)



Profiled Evaluation & Attacks

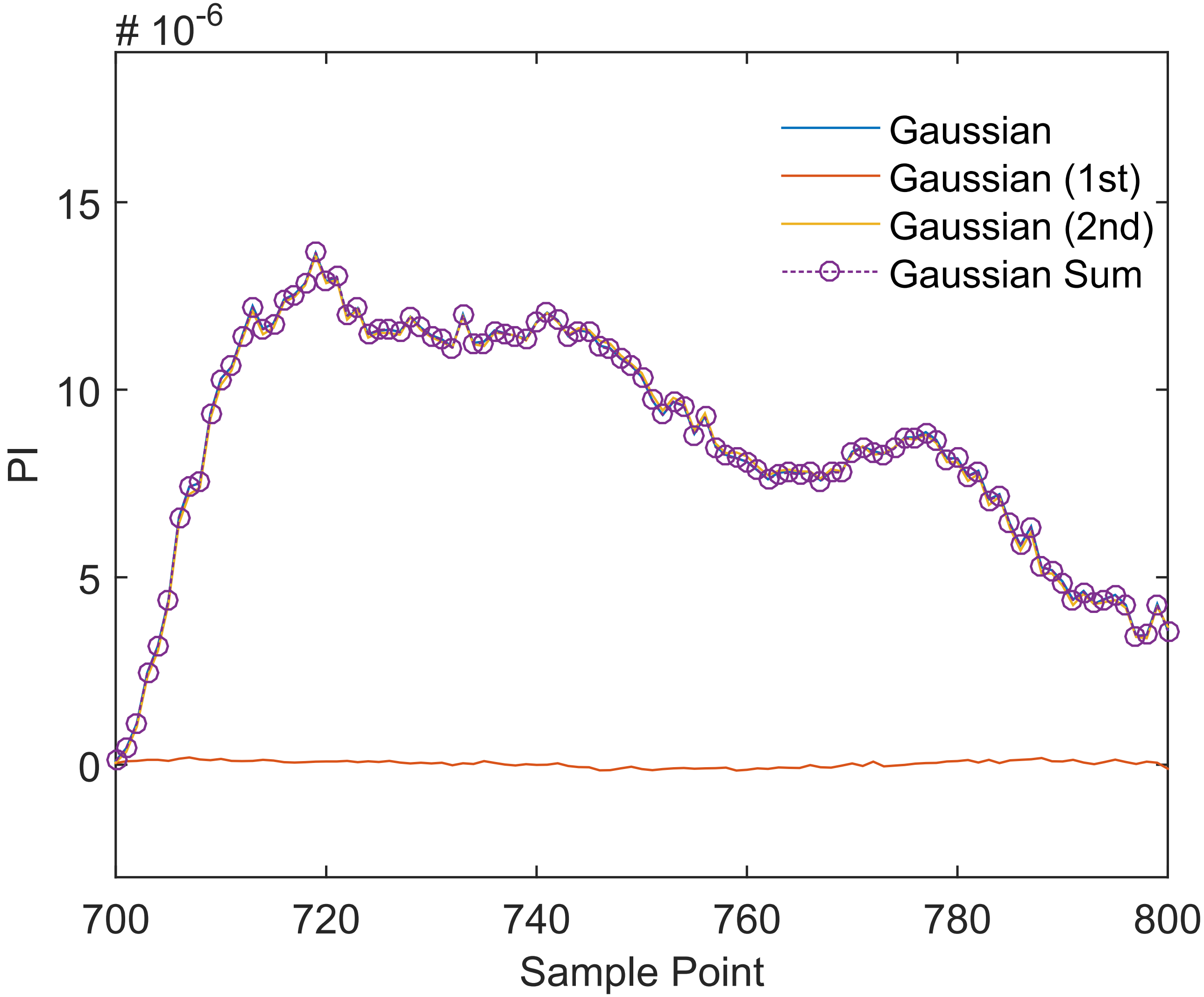
Separate Moments

- Fix all but one of the moments to a fixed value
- Removes all information in these moments
- Should not change the overall form of the distribution
- Average over all classes works well for our case-study

	Dist. 1	Dist. 2	Dist. 3	Dist. 4	Average
Mean	-27.97343	-27.98114	-27.98279	-27.97826	-27.97890
Variance	22.36243	21.99796	22.21650	22.26601	22.21073
Skewness	0.00750	0.00531	0.01310	-0.00007	0.00646
Kurtosis	3.01775	3.02025	3.02192	3.01835	3.01957

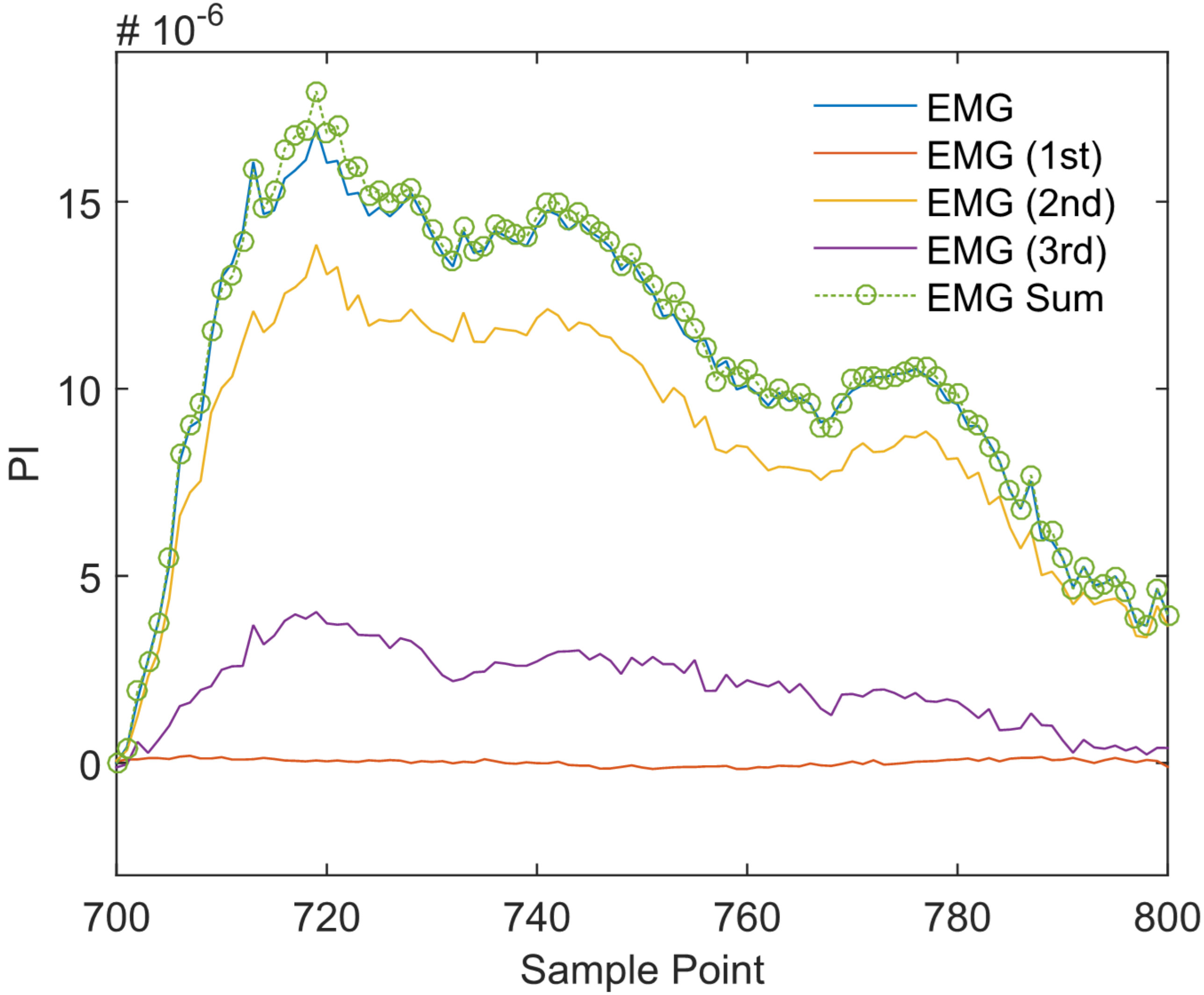
Profiled Evaluation & Attacks

Separate Moments (Gaussian)



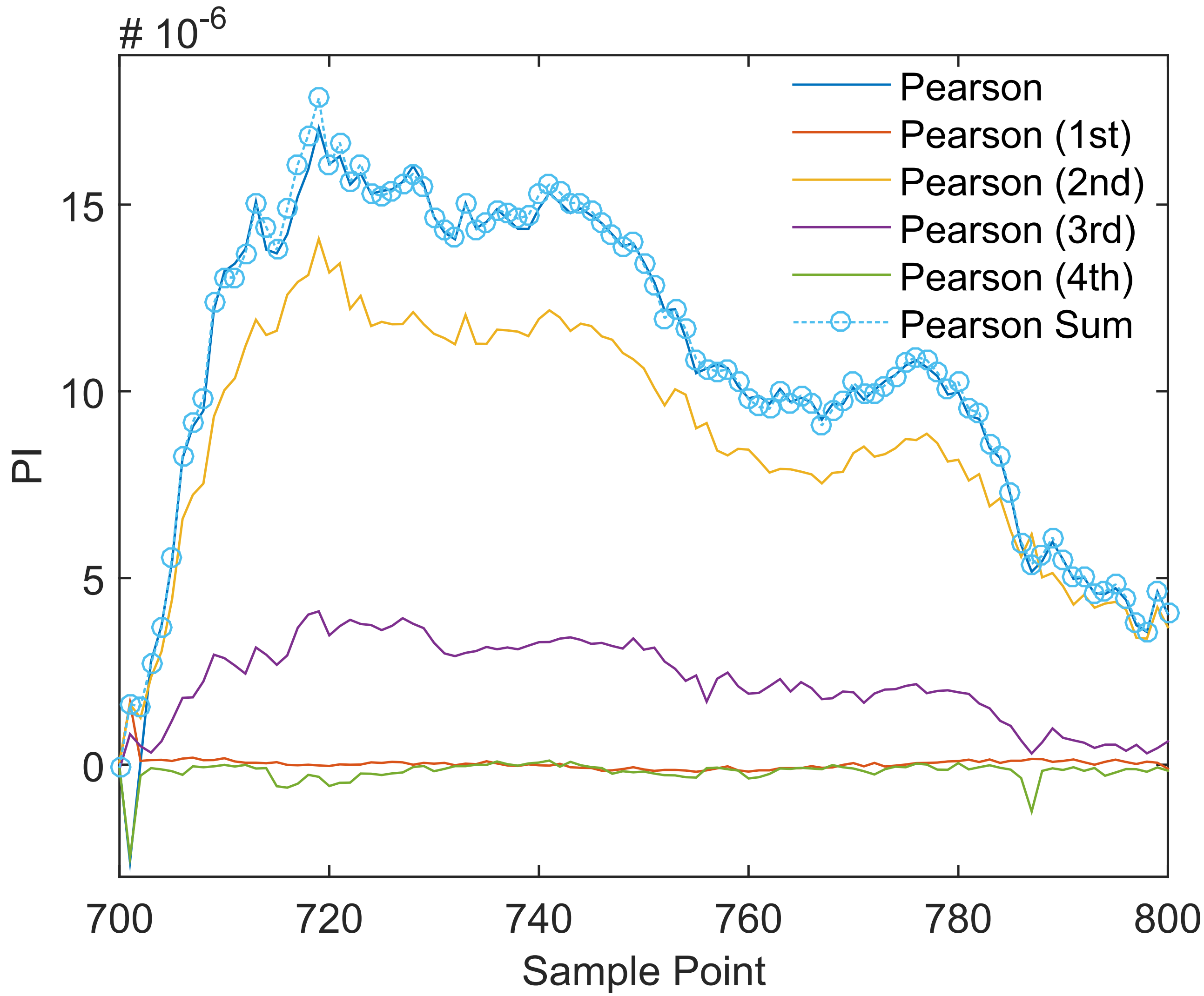
Profiled Evaluation & Attacks

Separate Moments (EMG)



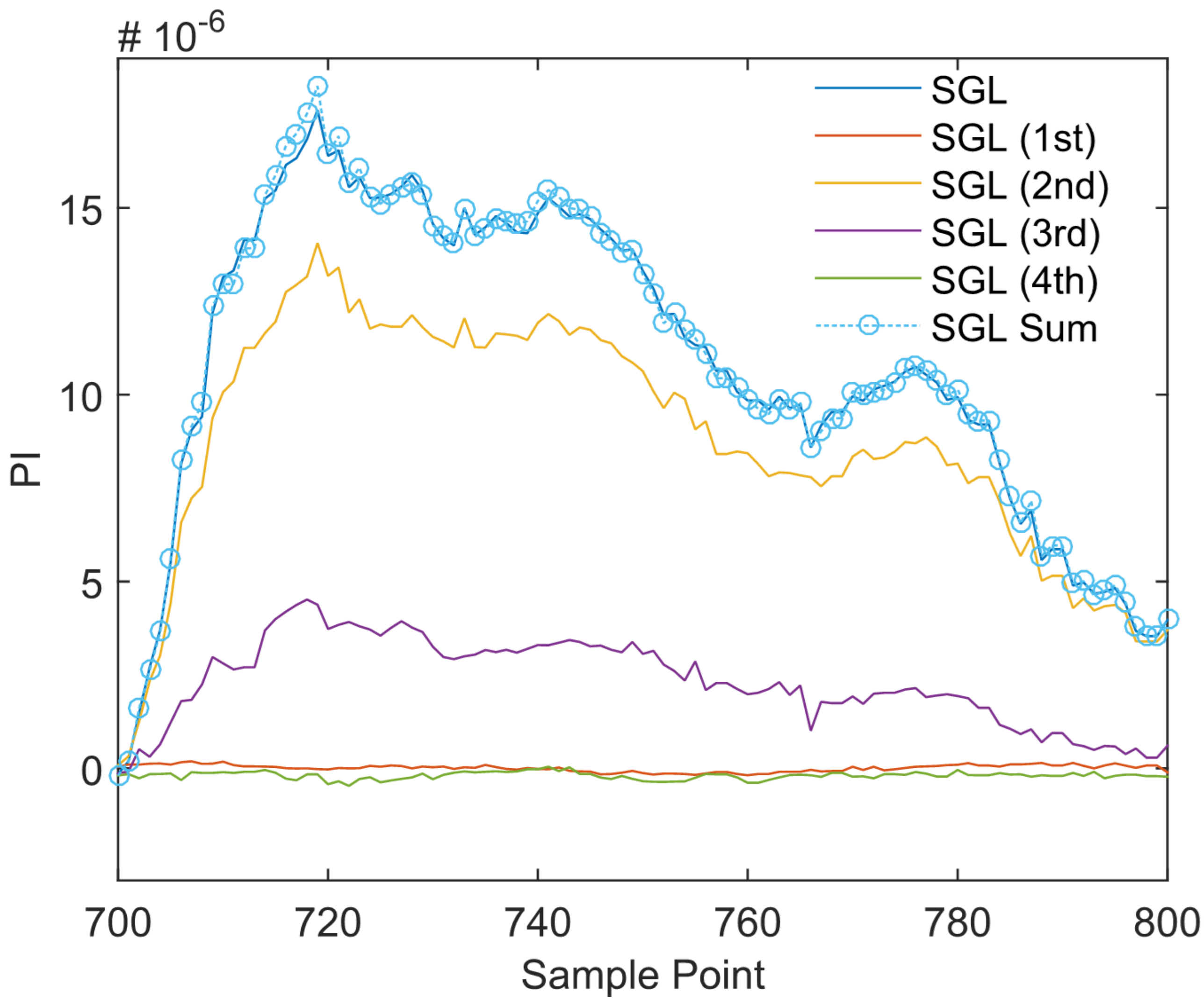
Profiled Evaluation & Attacks

Separate Moments (Pearson)



Profiled Evaluation & Attacks

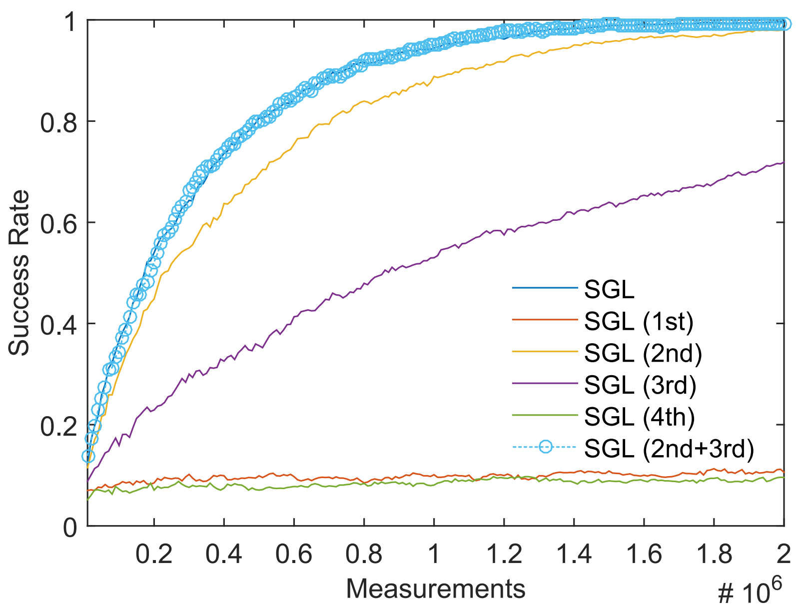
Separate Moments (SGL)

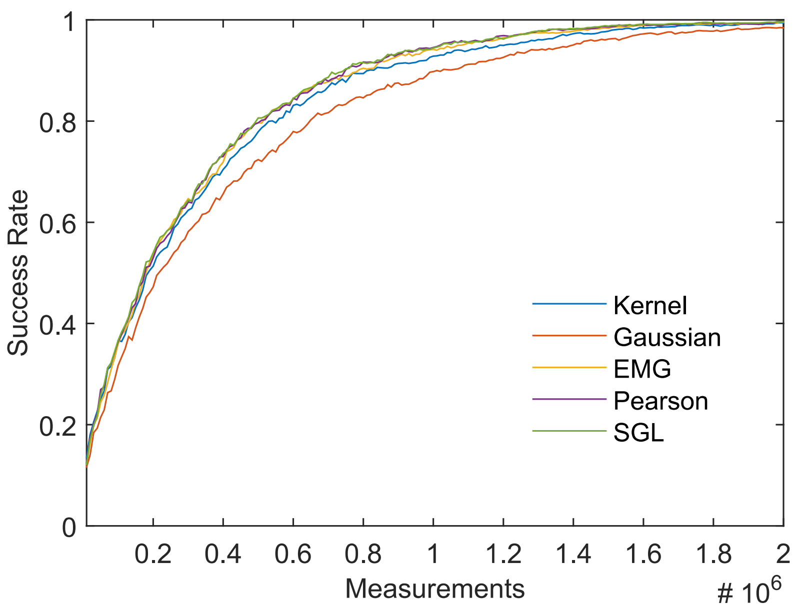


Profiled Evaluation & Attacks

Template Attack (Sample Point 719)

- 90,000,000 used in profiling phase
- Uses the leakage PDF as key distinguisher
- 1000 experiments to compute success rate for different number of traces

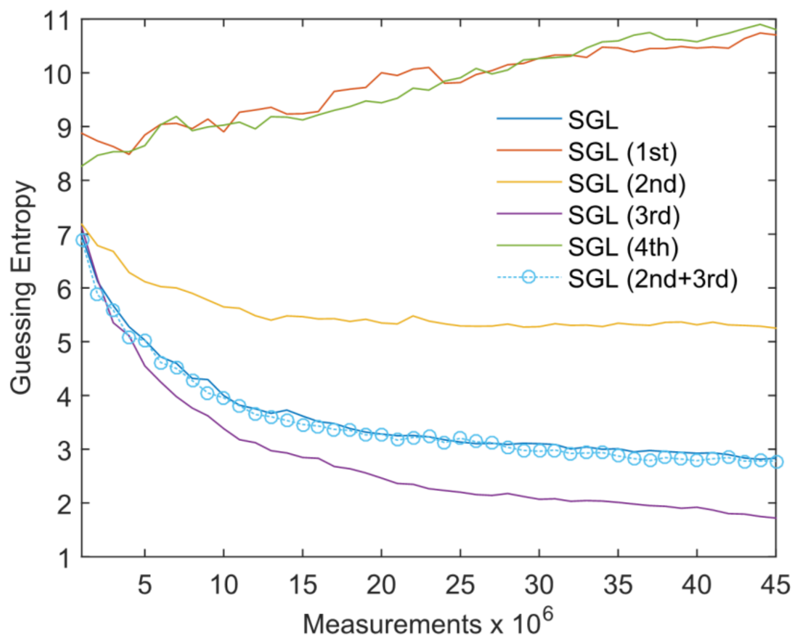


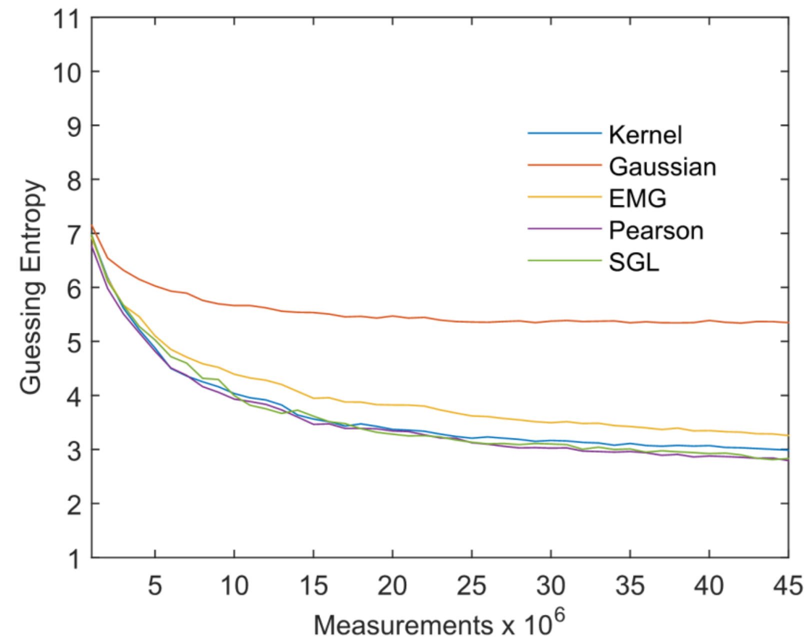


Non-Profiled Evaluation & Attacks

Mutual Information Analysis (Sample Point 719)

- Requires a leakage model for attacks on the first round
- Used 3 MSB of S-box output
- 1000 experiments to compute guessing entropy (average rank of correct key)





Tool Selection

- Gaussian is still very efficient for unprotected devices and simple first-order masking schemes
- New tools can be used for thorough leakage profiling of more complex designs
- The least complex but still applicable distribution should be used
  1) Moments 1-2: Gaussian
  2) Moments 1-3: EMG
  3) Moments 1-4: Pearson or SGL depending on type of leakage and computational limitations

Outline

- Introduction
- Background
- New Tools
- Results and Comparison
- **Conclusion**

Conclusion

- Extended SCA evaluation toolbox
- Introduced new tools which offer high flexibility and fast convergence
- Enable thorough leakage profiling of a majority of current relevant masked HW designs
- Powerful profiled and non-profiled attacks using multiple moments

Future Work:

- Combination of new methods with simplifying approaches
- Extension to multivariate scenario
- Formal investigation of “summing rule”

Thanks for Listening!

Any Questions?