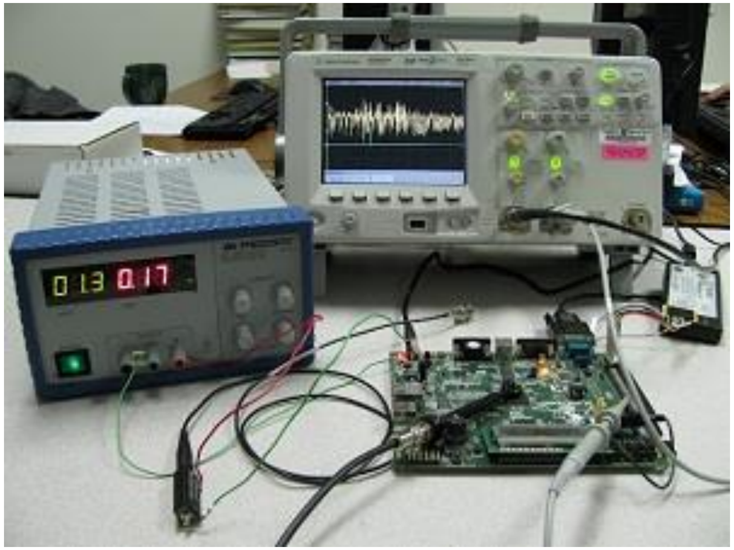
# How (not) to Use Welch's T-test in Side-Channel Security Evaluations



F.-X. Standaert

UCL Crypto Group, Belgium

**CARDIS 2018, Montpellier, France**

# Motivation for leakage detection 

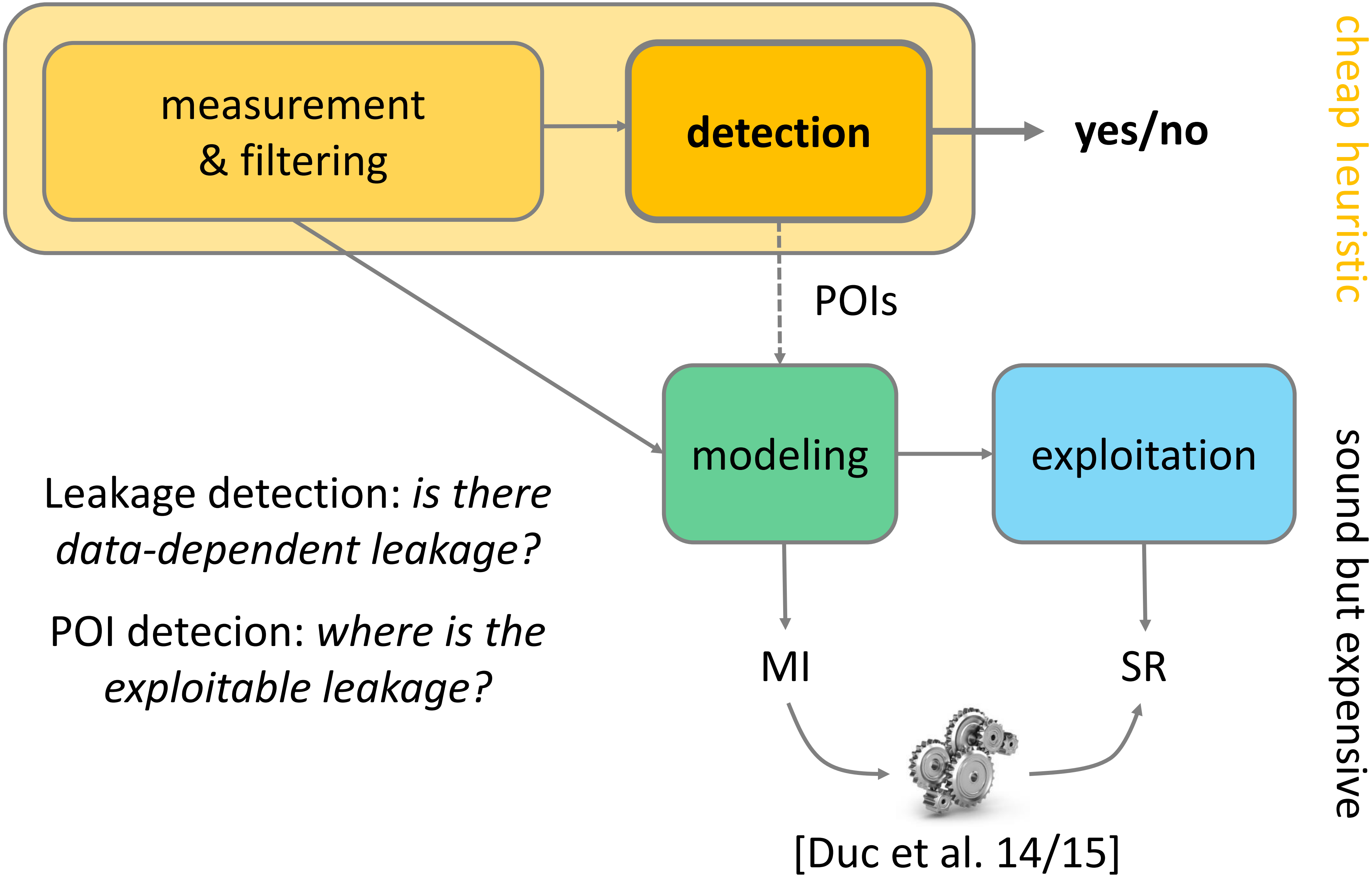
- Real World Crypto 2017 (Helena Handschuh)
  - *DPA resistance for real people*
  - https://www.youtube.com/watch?v=qvwwz8V9XRo

- Provide test methods that are
  - Repeatable
  - Precise
  - Automated
  - Less subjective
  - Low cost

} *conformance-style* testing

# Big picture (side-channel evaluations)



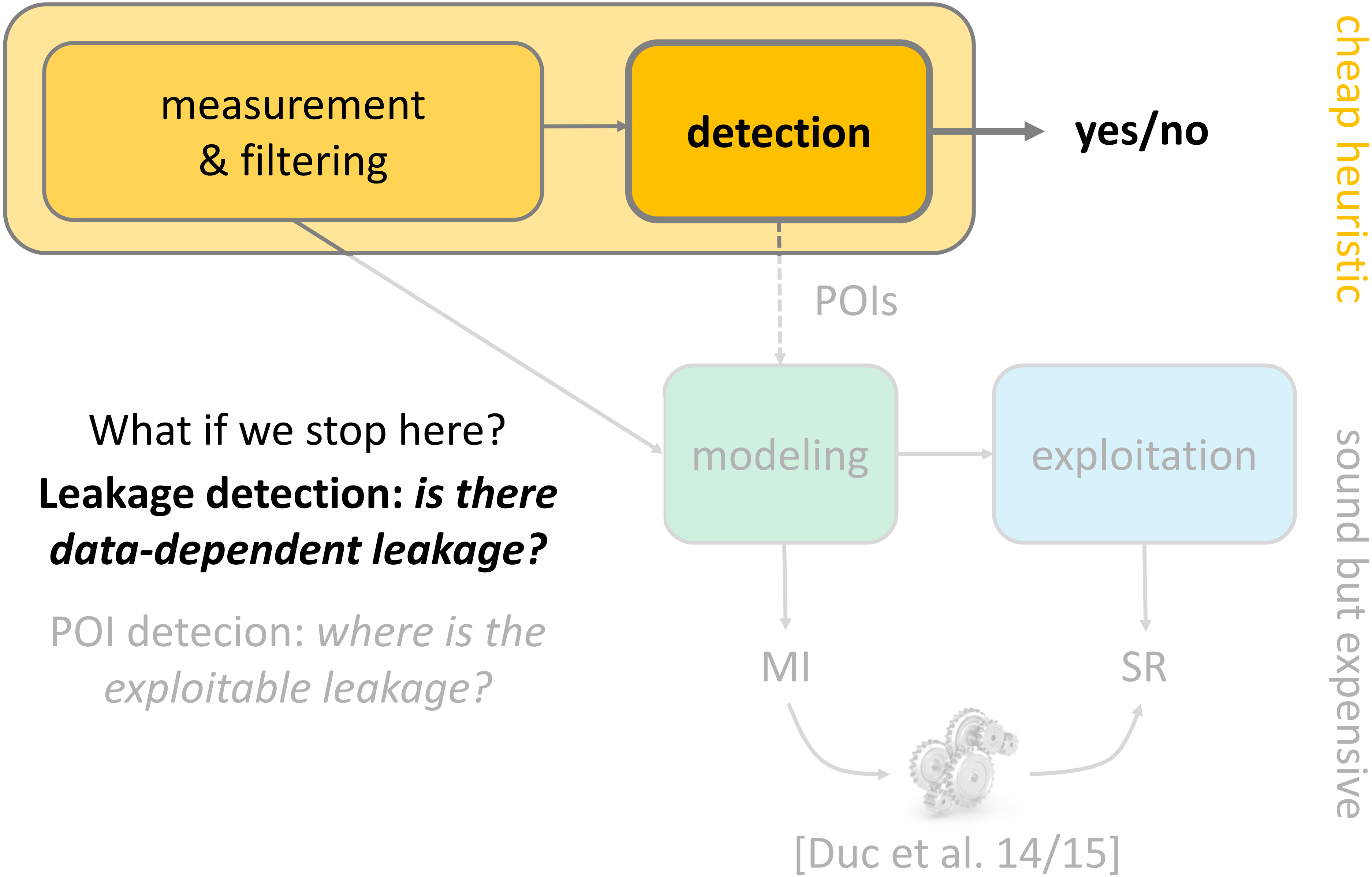



Leakage detection: *is there data-dependent leakage?*

POI detecion: *where is the exploitable leakage?*

# Bigger picture (this talk)






What if we stop here?

**Leakage detection: *is there data-dependent leakage?***

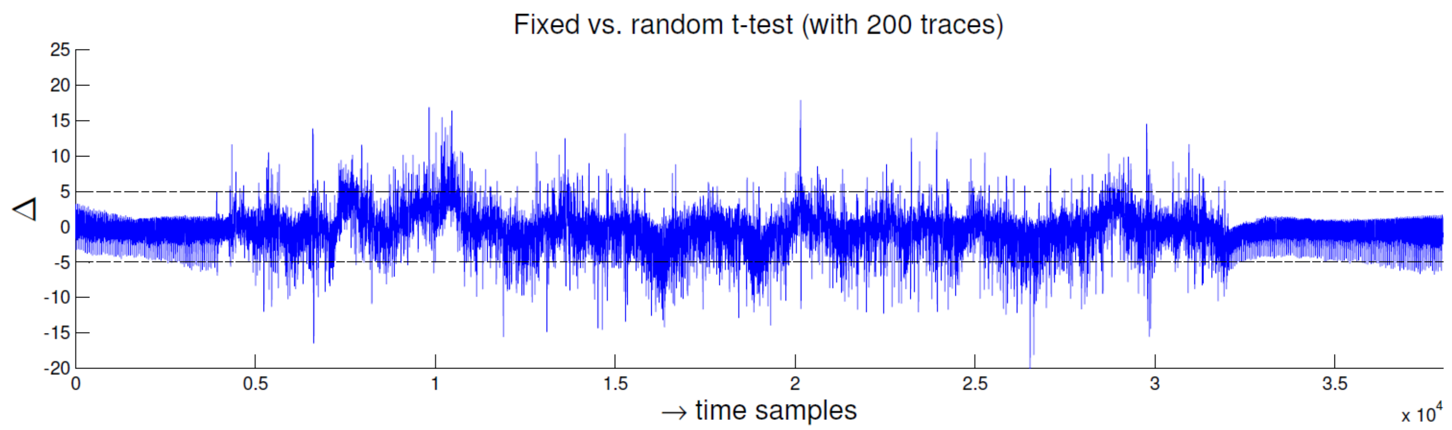
POI detecion: *where is the exploitable leakage?*

# Example: detection with Welch's t-test



- AES Rijndael example
  - 128-bit key fixed
  - $N_f$ traces with a fixed plaintext
  - $N_r$ traces with random plaintexts
  - Apply Welch's t-test to the f&r classes:

$$\Delta(t) = \left[\hat{\mu}_f(t) - \hat{\mu}_r(t)\right] / \left[\left(\hat{\sigma}_f^2(t)/N_f\right) + \left(\hat{\sigma}_r^2(t)/N_r\right)\right]$$

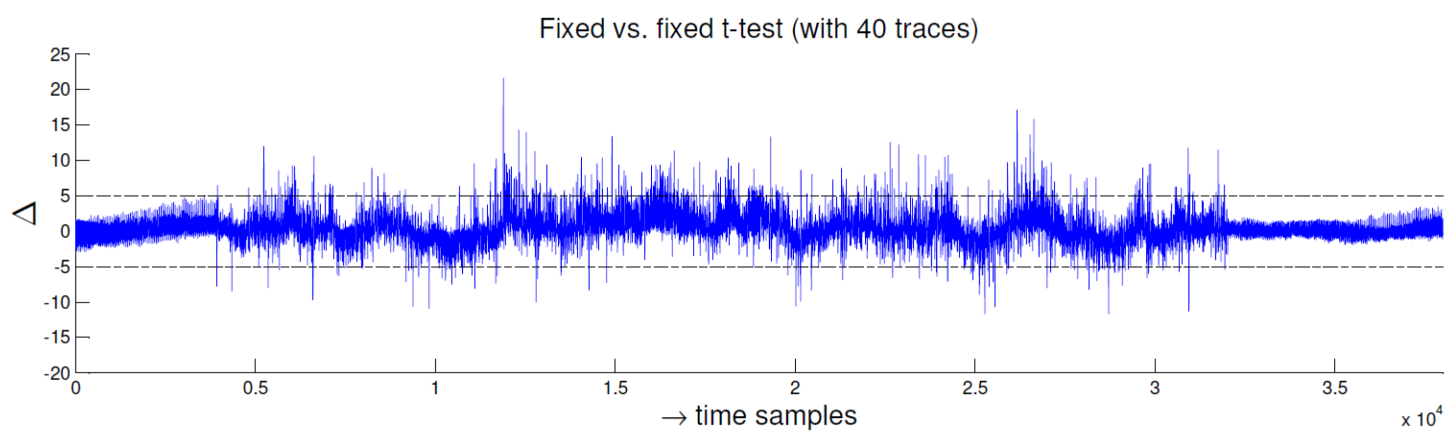


# Example: detection with Welch's t-test



- AES Rijndael example
  - 128-bit key fixed
  - $N_{f_1}$ traces with a fixed plaintext
  - $N_{f_2}$ traces with another fixed plaintext
  - Apply Welch's t-test to the f&f classes:

$$\Delta(t) = \left[\hat{\mu}_{f_1}(t) - \hat{\mu}_{f_2}(t)\right] / \left[\left(\hat{\sigma}^2_{f_1}(t)/N_{f_1}\right) + \left(\hat{\sigma}^2_{f_2}(t)/N_{f_2}\right)\right]$$



# Tempting shortcoming



- No detection ⇒ there is anyway no attack
  - Are there false negatives that contradict this?

# Tempting shortcoming



- No detection ⇒ there is anyway no attack
  - Are there false negatives that contradict this?

- *Exemple of false negative #1*

- $y = x \oplus k, z = \mathrm{S}(x \oplus k), l = \mathrm{HW}(z) + n$
  - $\hat{\mu}_r = 4$ anyway
  - Say $\hat{\mu}_f = 4$ $(z = 15)$

  } no detection possible
- Not all leaking samples can be detected

# Tempting shortcoming



- No detection ⇒ there is anyway no attack
  - Are there false negatives that contradict this?

- *Exemple of false negative #1*

- $y = x \oplus k, z = \mathrm{S}(x \oplus k), l = \mathrm{HW}(z) + n$
  - $\hat{\mu}_r = 4$ anyway
  - Say $\hat{\mu}_f = 4$ ($z = 15$)

  } no detection possible
- Not all leaking samples can be detected

- But not a problem if applied to long traces

# Tempting shortcoming



- No detection ⇒ there is anyway no attack
  - Are there false negatives that contradict this?

- *Exemple of false negative #2*

- Highly multivariate attacks
  - Static leakages (slow clock) [M14,M+15]
  - Horizontal attacks, SASCA [B+16,GS18]

[M14] Amir Moradi: *Side-Channel Leakage through Static Power - Should We Care about in Practice?* CHES 2014: 562-579. [M+15] Santos Merino Del Pozo, François-Xavier Standaert, Dina Kamel, Amir Moradi: *Side-channel attacks from static power: when should we care?* DATE 2015: 145-150 [B+16] Alberto Battistello, Jean-Sébastien Coron, Emmanuel Prouff, Rina Zeitoun: *Horizontal Side-Channel Attacks and Countermeasures on the ISW Masking Scheme.* CHES 2016: 23-39 [GS18] Vincent Grosso, François-Xavier Standaert: *Masking Proofs Are Tight and How to Exploit it in Security Evaluations.* EUROCRYPT (2) 2018: 385-412

# Tempting shortcoming



- No detection ⇒ there is anyway no attack
  - Are there false negatives that contradict this?

- *Exemple of false negative #2*

- Highly multivariate attacks
  - Static leakages (slow clock) [M14,M+15]
  - Horizontal attacks, SASCA [B+16,GS18]

- But these are highly sophisticated attacks 🤔

[M14] Amir Moradi: *Side-Channel Leakage through Static Power - Should We Care about in Practice?* CHES 2014: 562-579. [M+15] Santos Merino Del Pozo, François-Xavier Standaert, Dina Kamel, Amir Moradi: *Side-channel attacks from static power: when should we care?* DATE 2015: 145-150 [B+16] Alberto Battistello, Jean-Sébastien Coron, Emmanuel Prouff, Rina Zeitoun: *Horizontal Side-Channel Attacks and Countermeasures on the ISW Masking Scheme.* CHES 2016: 23-39 [GS18] Vincent Grosso, François-Xavier Standaert: *Masking Proofs Are Tight and How to Exploit it in Security Evaluations.* EUROCRYPT (2) 2018: 385-412

# Research problem

- Can we design an implementation
  - For which detection is hard / impossible
  - That is trivial to break (e.g., with 1 trace)
    - Exploiting a simple (univariate) attack

?

# Experimental setting

- Masked encoding (parallel implementation)
  - $x = x_1 \oplus x_2 \oplus \cdots \oplus x_m$
  - Vector of shares $\bar{x} = (x_1, x_2, \ldots, x_m)$

- Linear (or quadratic) leakage function

$$\mathrm{L}(\bar{x}) = \left(\sum_{i=1}^{m} \alpha_i . x_i\right) + n \,, \quad \mathrm{L}(\bar{x}) = \left(\sum_{i=1}^{m} \alpha_i . x_i\right) + \left(\sum_{i,j=1}^{m} \beta_{i,j} . (x_i \wedge x_j)\right) + n$$

- Compute t-test statistic and MI (worst-case) metric

$$\mathrm{MI}(X; \boldsymbol{L}) = \mathrm{H}(X) + \sum_{x} \Pr[x] + \sum_{\boldsymbol{l}} \mathrm{f}(\boldsymbol{l}|x) . \log_2(\Pr[x|\boldsymbol{l}])$$

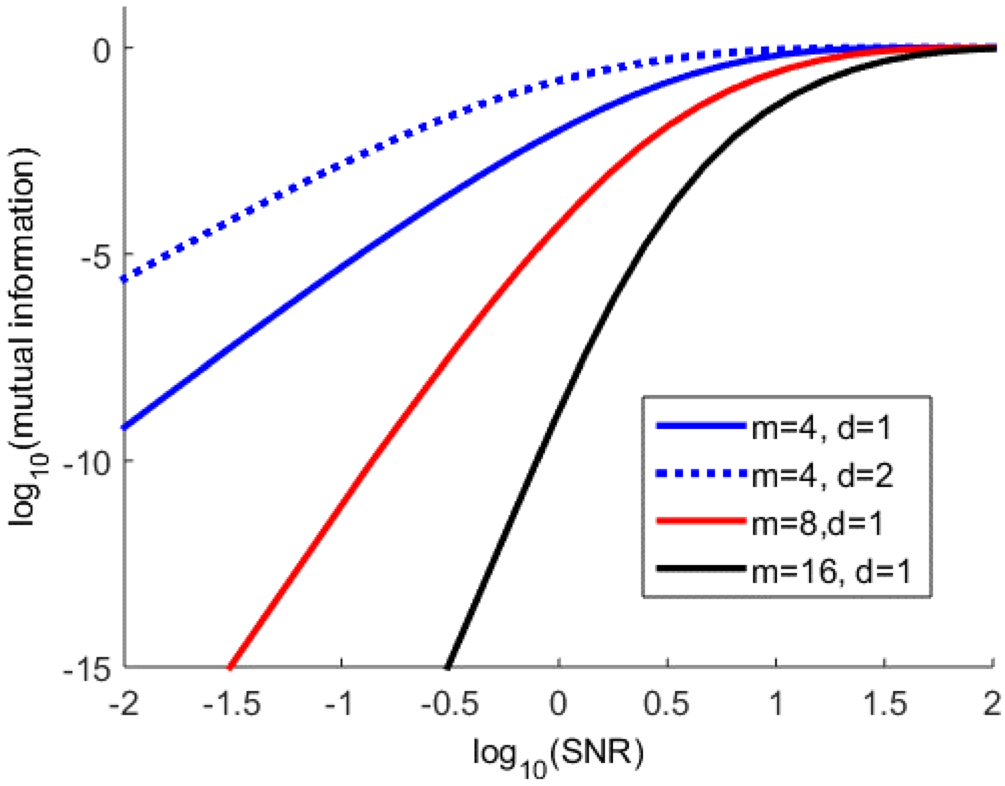
# Parameters



- The number of shares = bus size *m*
- The degree of the leakage function (*d*=1,2)
- The order of the leakage detection ($o \leq m$)
  - Pre-processed samples $\overline{L'}(i) = \left(\bar{L}(i) - \hat{\mu}\left(\bar{L}(i)\right)\right)^{o}$
- The amount of noise in the leakages

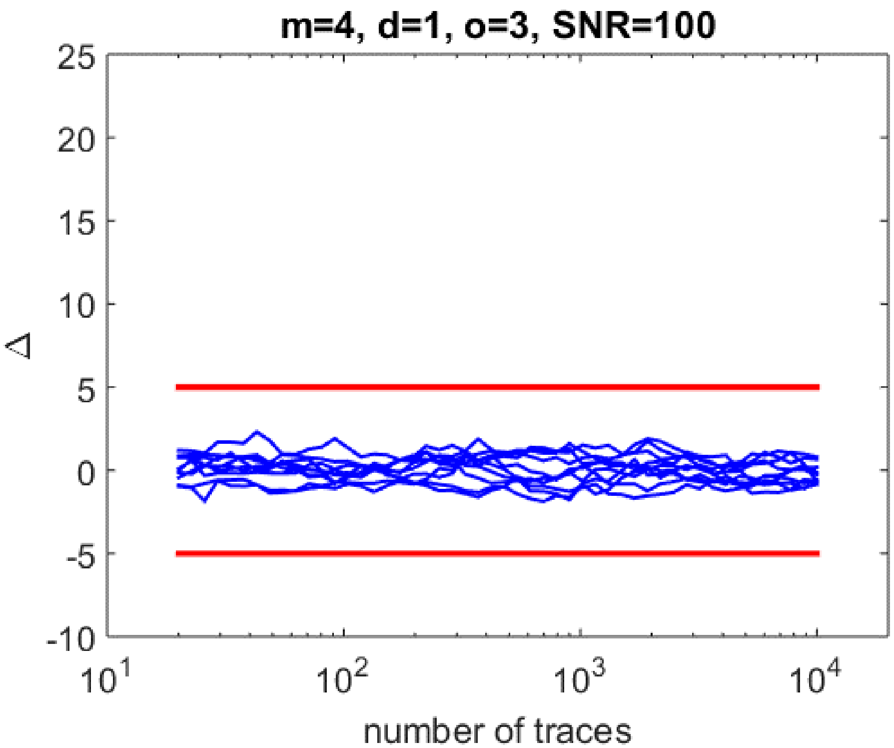
$$\mathrm{SNR} = \frac{m/4}{\sigma_n^2}$$

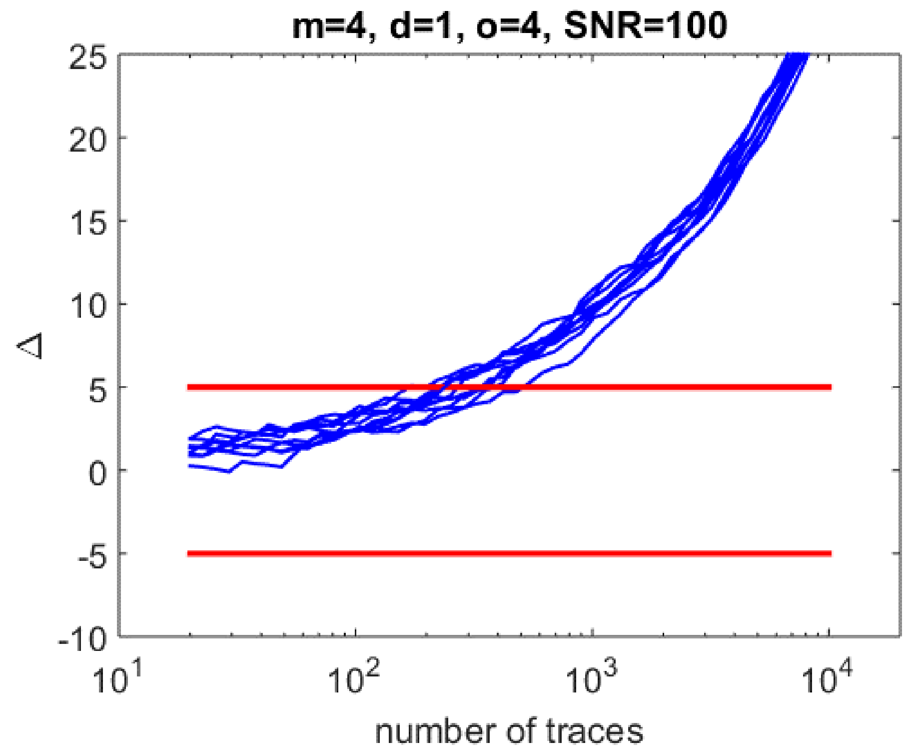
# Information theoretic analysis




- Very weak security for high SNRs
  - Trivial attack: check whether HW is even or odd

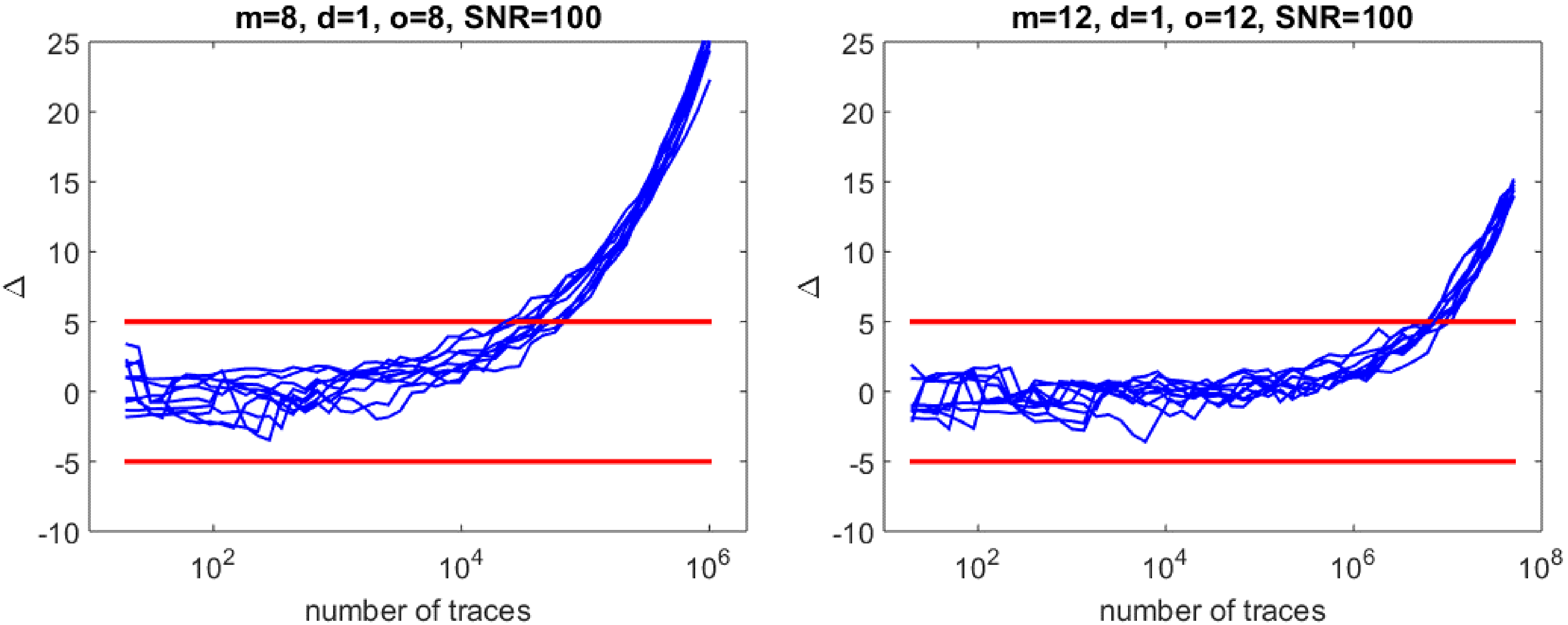
# Leakage detection #1




- Detection starts at order 4 (as expected)
- But it is already not trivial with 4 shares!

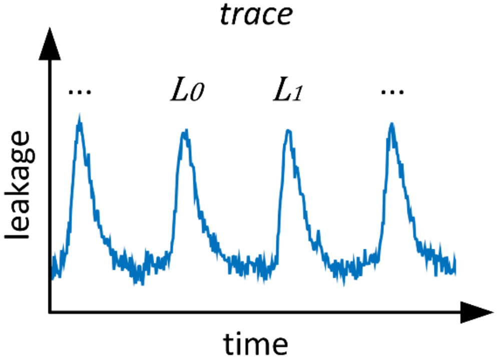
# Leakage detection #2




- Things get worse as the # of shares increase
- Why: detection assumes an Adv. strategy
  - Estimating moments is suboptimal with high SNR

# Interpretation (security models)

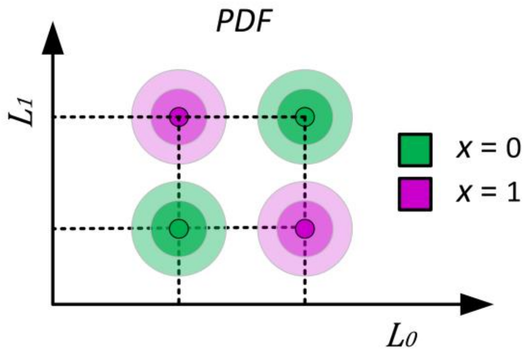





**Noisy leakages security:** $N \propto \frac{c}{\mathrm{MI}(X;\boldsymbol{L})}$

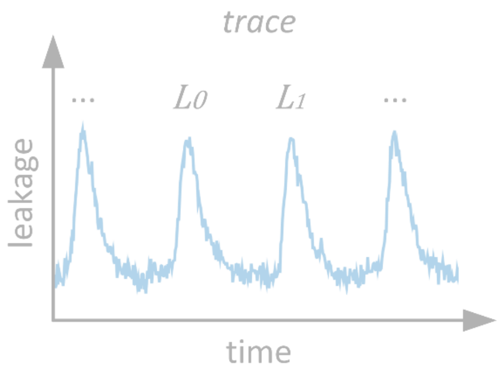
Goal (ideally): $\mathrm{MI}(X;\boldsymbol{L}) < \mathrm{MI}(X_i;L_i)^m$

# Interpretation (security models)




**Bounded moment security:**

$$\underbrace{\prod_{i_1,i_2,\dots,i_{m-1}} L_i}_{(m\text{-}1)\text{th order statistical moment (ideally)}} \perp\!\!\!\perp X$$

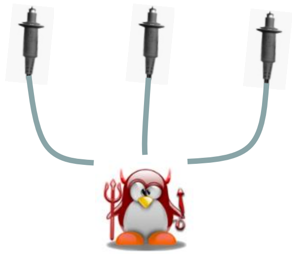



**Noisy leakages security:** $N \propto \frac{c}{\mathrm{MI}(X;\boldsymbol{L})}$

Goal (ideally): $\mathrm{MI}(X;\boldsymbol{L}) < \mathrm{MI}(X_i;L_i)^m$

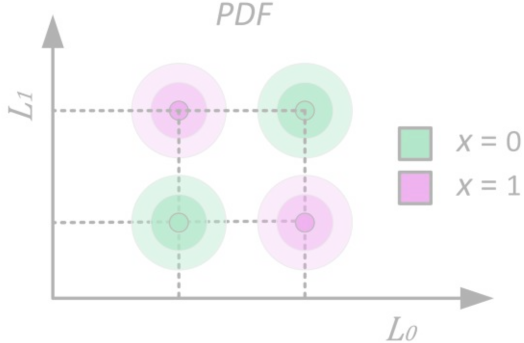
# Interpretation (security models)

$x = x_1 + x_2 + \cdots + x_m$



**Probing security:**

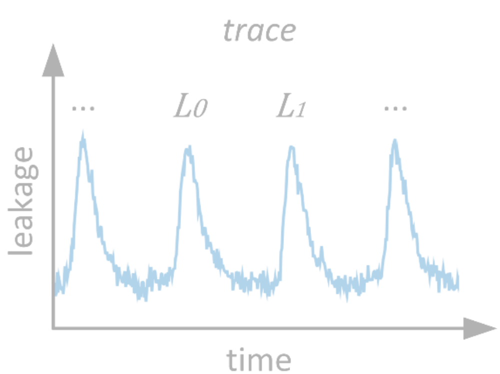
Sets of (m-1) probes are $\perp\!\!\!\perp$ of $X$ (ideally)




**Bounded moment security:**

$$\prod_{i_1, i_2, \ldots, i_{m-1}} L_i \perp\!\!\!\perp X$$

($m$-1)th order statistical moment (ideally)

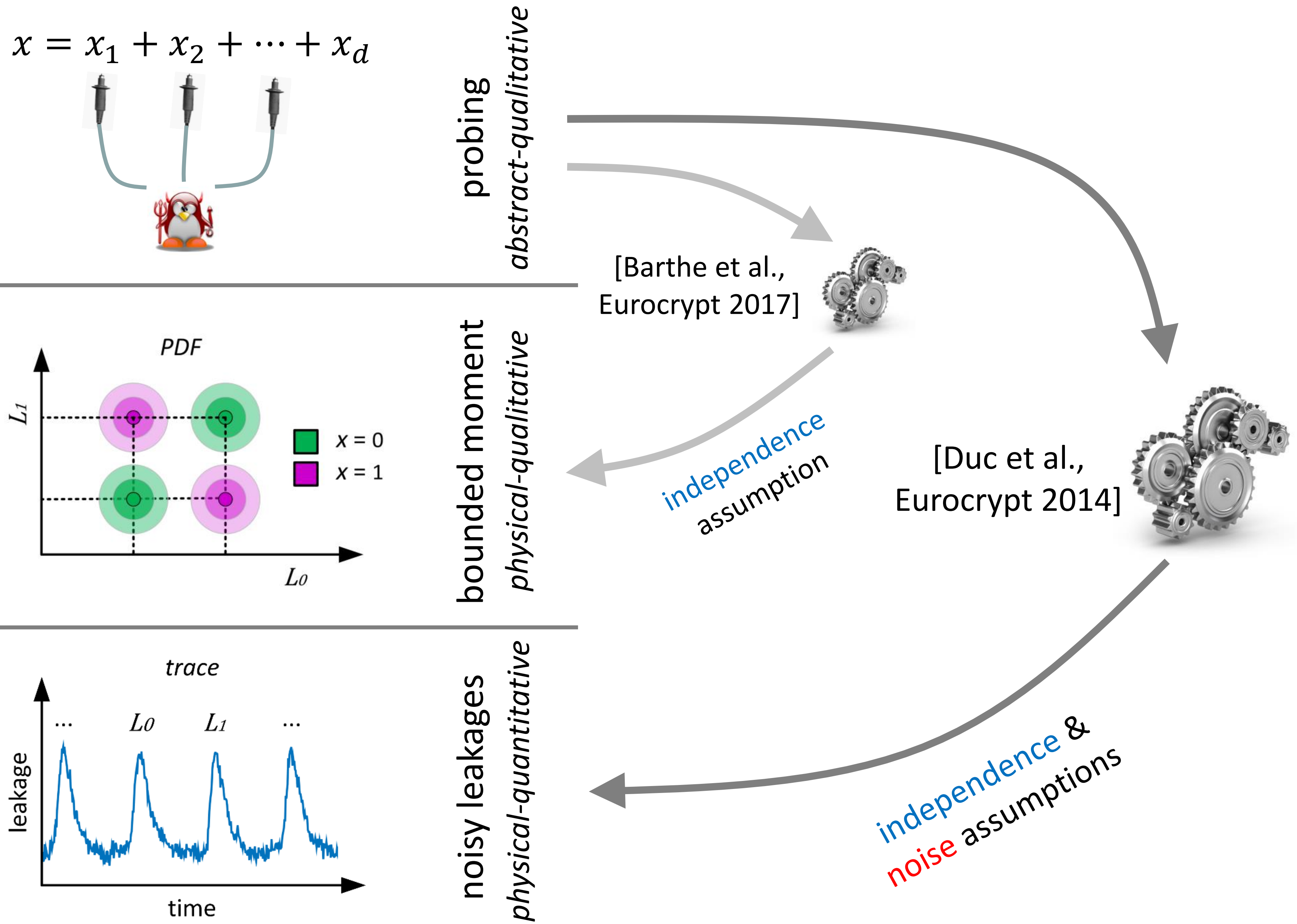



**Noisy leakages security:** $N \propto \frac{c}{\mathrm{MI}(X;\boldsymbol{L})}$

Goal (ideally): $\mathrm{MI}(X;\boldsymbol{L}) < \mathrm{MI}(X_i;L_i)^m$

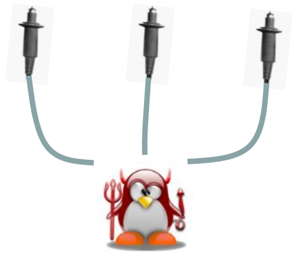
# Interpretation (security reductions)



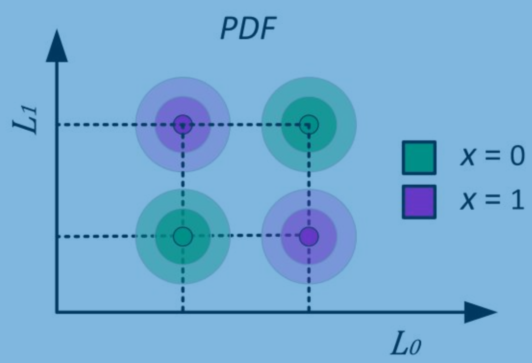


# Take home message: separation of duties

$$x = x_1 + x_2 + \cdots + x_d$$

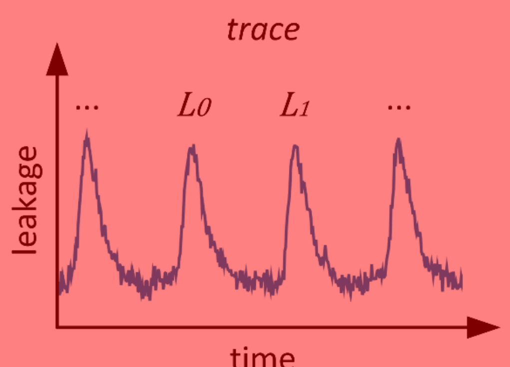


probing




bounded moment

Can be evaluated with Welch's t-test
(or any moment-based tool)




noisy leakages

Cannot be evaluated with Welch's t-test
(needs SNR or distribution-based tool)

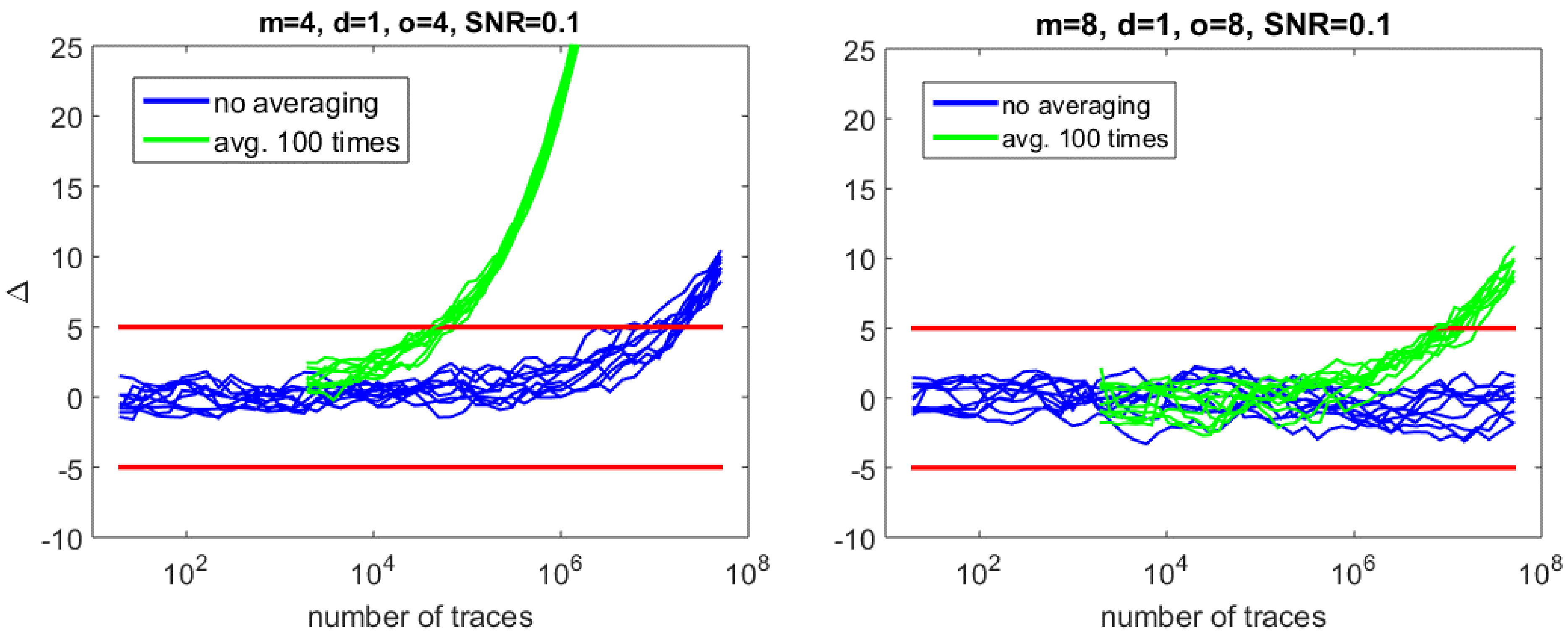
# Last remark



- Say you want to evaluate the security order
  - Smallest leaking moment of $\mathrm{f}(x|l)$
- But noise is large (SNR is low)
- Hence detection complexity grows exp. in *m*

- If masks are under control, an improved detection is obtained by averaging $l|\bar{x}$
  - Intuition: prevents noise amplification

# Illustration




- The improved detection level is even less correlated with the security level (but it wasn’t anyway…)

# Conclusions



- “Detection-only” evaluations are risky
  - Have a limited quantitative meaning
    - Especially in the case of masking
- This paper discusses the noise issue
- But the multivariate issue is as important

# Conclusions



- “Detection-only” evaluations are risky
  - Have a limited quantitative meaning
    - Especially in the case of masking
- This paper discusses the noise issue
- But the multivariate issue is as important

- Limitations are less critical if detection occurs
- But interpreting “no detection” is very hard
  - It certainly does not mean the device is secure

# Conclusions



- “Detection-only” evaluations are risky
  - Have a limited quantitative meaning
    - Especially in the case of masking
- This paper discusses the noise issue
- But the multivariate issue is as important

- Limitations are less critical if detection occurs
- But interpreting “no detection” is very hard
  - It certainly does not mean the device is secure

- (Improved) detection is a useful ingredient though
  - To assess an implementation’s “security order”
  - As a first step before other analyzes

# THANKS

http://perso.uclouvain.be/fstandae/