Optimal Attacks for Multivariate and Multi-model Side-Channel Leakages

Nicolas Bruneau, Sylvain Guilley, Annelie Heuser, Damien Marion and Olivier Rioul
firstname.lastname@telecom-paristech.fr

Abstract. In practice, a side-channel signal is measured as a trace consisting of *several* samples where *several* sensitive bits are manipulated in parallel, each leaking differently. Therefore, the informed attacker needs to devise side-channel distinguishers that can handle both *multivariate* leakages and *multivariate* models at the same time. In the state of the art, these two issues have two independent solutions: on the one hand, dimensionality reduction can cope with multivariate leakage; on the other hand, online stochastic approaches can cope with multivariate models.
In this work, we combine both solutions to derive closed-form expressions of the resulting *optimal* distinguisher in terms of matrix operations, in all situations where the model can be either profiled offline or regressed online. Optimality here means that the success probability is maximized for a given number of traces. We recover known results for uni- and bi-variate models (including correlation power analysis), and investigate novel distinguishers for multivariate models with more than two parameters. Following ideas from the AsiaCrypt'2013 paper "*Behind the Scene of Side-Channel Attacks*", we also provide fast computation algorithms in which the traces are accumulated prior to computing the distinguisher values.

(full version at PROOFS 2016)

1. Fact

Side-channel leakages are:

- **multi-variate** .... (in time)
- **multi-model** .... (e.g., each bit leaks $\neq$)

2. Matrix Notations

- Q .... number of queries,
- D .... number of samples,
- S .... number of models.

In matrix notation:

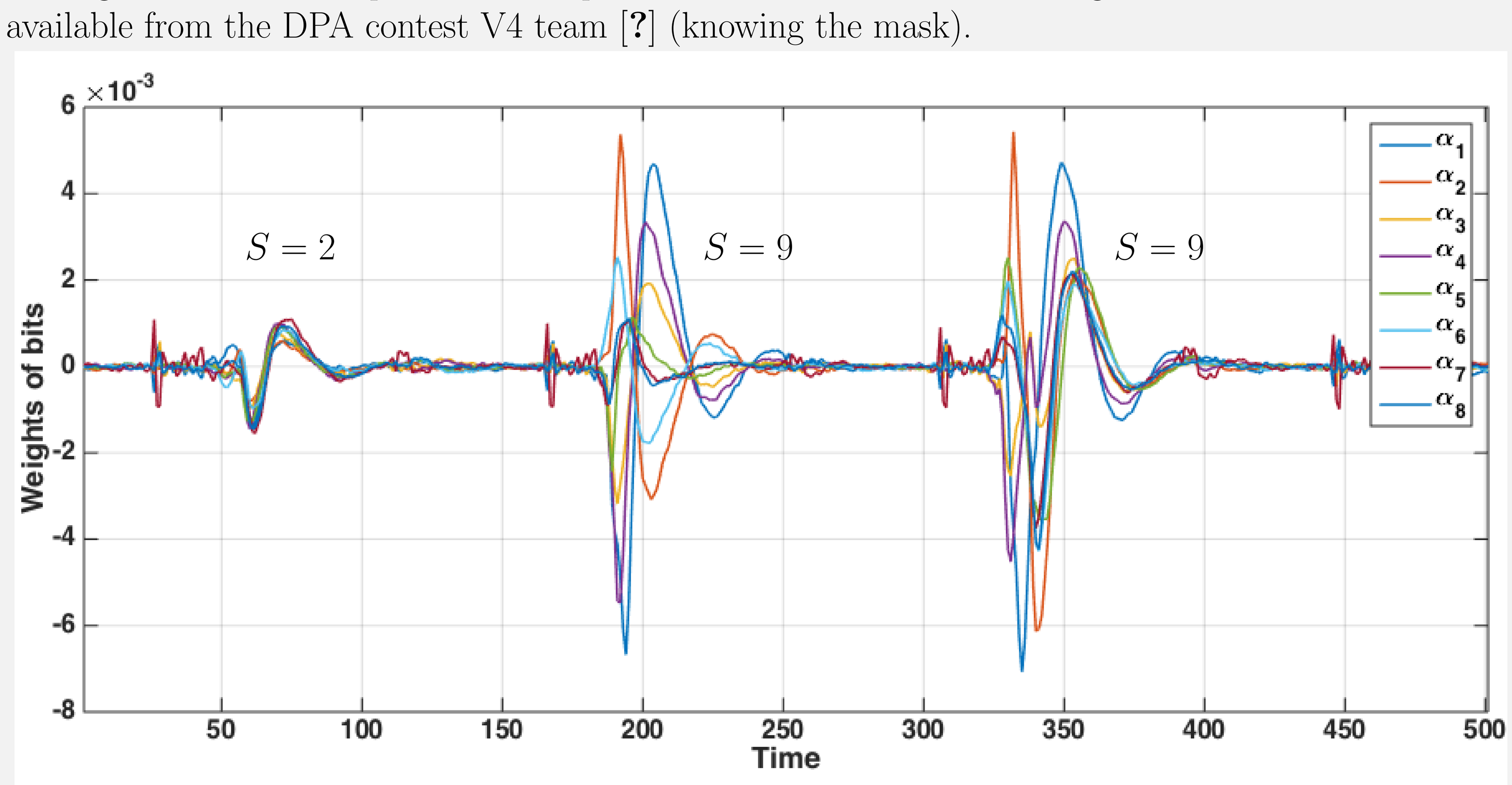
$$\mathbf{X} = \alpha \mathbf{Y}^\star + \mathbf{N} \tag{1}$$

where

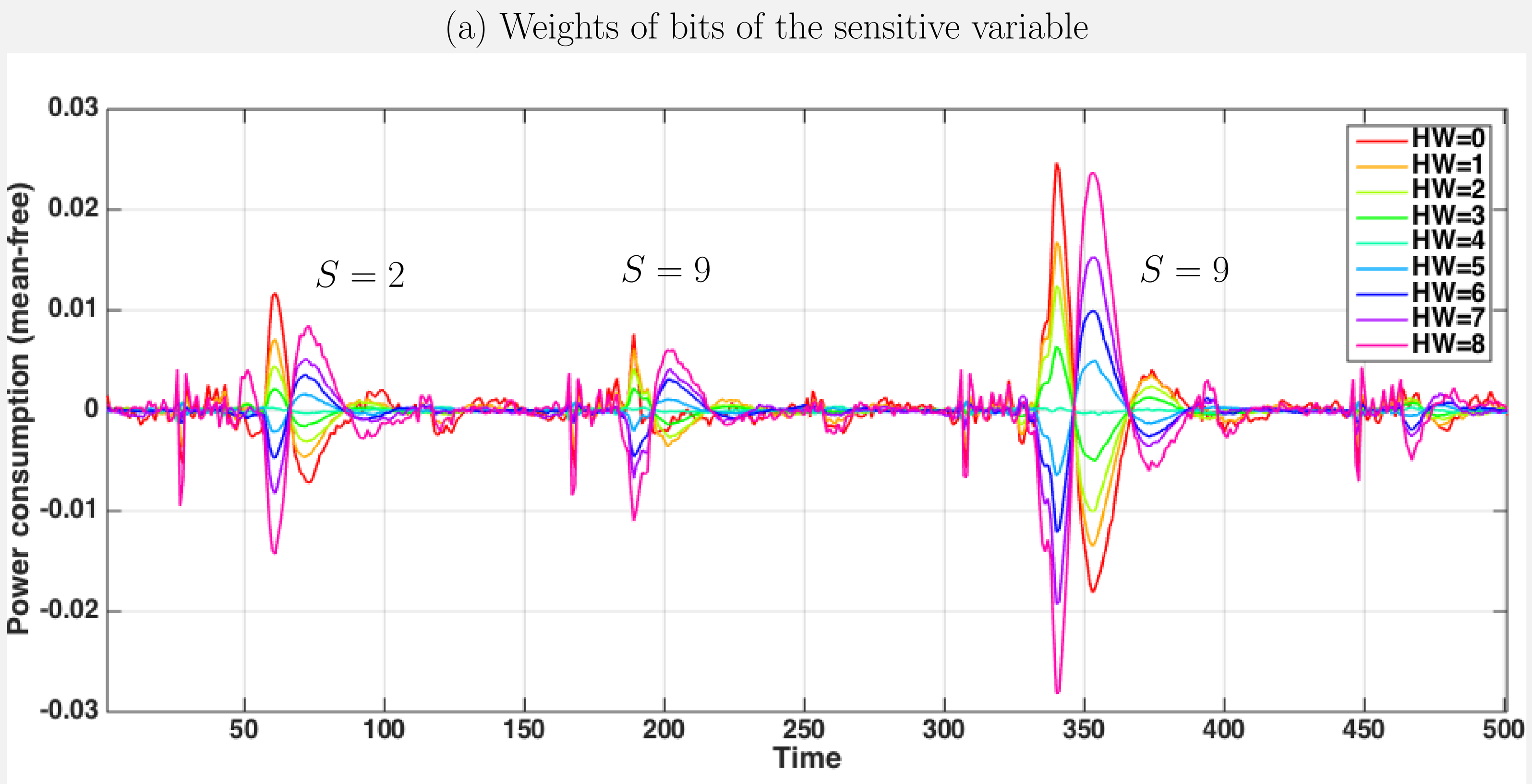
- $\mathbf{X}$ is a matrix of size .... $D \times Q$,
- α is a matrix of size .... $D \times S$,
- $\mathbf{Y}^\star$ (the star means: "for the correct key $k = k^\star$") is a matrix of size .... $S \times Q$,
- $\mathbf{N}$ is a matrix of size .... $D \times Q$.

3. Real World Example

The figures below show power consumption traces taken from an ATMega smartcard—datasets are available from the DPA contest V4 team [?] (knowing the mask).



(a) Weights of bits of the sensitive variable



(b) Mean power consumption for each Hamming weight class

4. Question

What is the optimal distinguisher, when in Equation (1):

- α is known? .... $\mathcal{D}_{\mathrm{ML}}(\mathbf{x}, \mathbf{t})$
- α is unknown? .... $\mathcal{D}_{\mathrm{ML,sto}}(\mathbf{x}, \mathbf{t})$

5. Solution

Theorem 1. *The optimal* maximum likelihood *(ML) distinguisher [?] for Gaussian noise writes*

$$\mathcal{D}_{\mathrm{ML}}(\mathbf{x}, \mathbf{t}) = \operatorname*{argmin}_k \ \mathrm{tr}\left((\mathbf{x} - \alpha\mathbf{y})^{\mathsf{T}} \Sigma^{-1} (\mathbf{x} - \alpha\mathbf{y})\right). \tag{2}$$

Proof. From [?] we have $\mathcal{D}_{\mathrm{ML}}(\mathbf{x}, \mathbf{t}) = \mathrm{argmax}_k \ p(\mathbf{x}|\mathbf{y})$ where from (1) it is easily seen that $p(\mathbf{x}|\mathbf{y}) = p_{\mathbf{N}}(\mathbf{x} - \alpha\mathbf{y})$. From the i.i.d. assumption the noise density $p_{\mathbf{N}}(\mathbf{n})$ is given by

$$p_{\mathbf{N}}(\mathbf{n}) = \prod_{q=1}^{Q} \frac{1}{\sqrt{(2\pi)^D |\det \Sigma|}} \exp -\frac{1}{2} n_q{}^{\mathsf{T}} \Sigma^{-1} n_q \tag{3}$$

$$= \frac{1}{(2\pi)^{DQ/2}} \frac{1}{(\det \Sigma)^{Q/2}} \exp -\frac{1}{2} \left(\sum_{q=1}^{Q} n_q{}^{\mathsf{T}} \Sigma^{-1} n_q \right) \tag{4}$$

$$= \frac{1}{(2\pi)^{DQ/2} (\det \Sigma)^{Q/2}} \exp -\frac{1}{2} \mathrm{tr}\left(\mathbf{n}^{\mathsf{T}} \Sigma^{-1} \mathbf{n}\right). \tag{5}$$

Thus $p_{\mathbf{N}}(\mathbf{x} - \alpha\mathbf{y})$ is maximum when the expression $\mathrm{tr}\left(\mathbf{n}^{\mathsf{T}} \Sigma^{-1} \mathbf{n}\right)$ for $\mathbf{n} = \mathbf{x} - \alpha\mathbf{y}$ is minimum. □

Theorem 2. *The optimal stochastic multivariate attack is given by*

$$\mathcal{D}_{\mathrm{ML,sto}}(\mathbf{x}, \mathbf{t}) = \operatorname*{argmax}_{k \in \mathbb{F}_2^n} \ \mathrm{tr}\left(\mathbf{y}^{\mathsf{T}} (\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1} \mathbf{y} \ \ \mathbf{x}^{\mathsf{T}} \Sigma^{-1} \mathbf{x}\right). \tag{6}$$

for which the optimal value of α is given by

$$\alpha^{opt} = (\mathbf{x}\mathbf{y}^{\mathsf{T}})(\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1}. \tag{7}$$

Proof. Let $\mathbf{x}' = \Sigma^{-1/2} \ \mathbf{x}$ and $\mathbf{y}' = (\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1/2} \ \mathbf{y}$. The optimal distinguisher minimizes the following expression over $\alpha \in \mathbb{R}^{D \times S}$:

$$\mathrm{tr}\left((\mathbf{x} - \alpha\mathbf{y})^{\mathsf{T}} \Sigma^{-1} (\mathbf{x} - \alpha\mathbf{y})\right) = \mathrm{tr}\left((\mathbf{x}' - \alpha'\mathbf{y})(\mathbf{x}' - \alpha'\mathbf{y})^{\mathsf{T}}\right) = \sum_{d=1}^{D} \|\mathbf{x}' - \alpha'_d \mathbf{y}\|^2.$$

The minimization over α'_d yields $\alpha'_d = (\mathbf{x}'_d \mathbf{y}^{\mathsf{T}})(\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1}$ for all $d = 1, \ldots, D$. This gives $\alpha' = (\mathbf{x}'\mathbf{y}^{\mathsf{T}})(\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1}$ hence $\alpha = (\mathbf{x}\mathbf{y}^{\mathsf{T}})(\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1}$, which remarkably does *not* depend on Σ.
The minimized value of the distinguisher is thus

$$\begin{aligned}
\min_\alpha \ \mathrm{tr}\left((\mathbf{x} - \alpha\mathbf{y})^{\mathsf{T}} \Sigma^{-1} (\mathbf{x} - \alpha\mathbf{y})\right) &= \mathrm{tr}\left((\mathbf{x} - \alpha^{\mathrm{opt}}\mathbf{y})^{\mathsf{T}} \Sigma^{-1} (\mathbf{x} - \alpha^{\mathrm{opt}}\mathbf{y})\right) \\
&= \mathrm{tr}\left((\mathsf{Id} - \mathbf{y}^{\mathsf{T}} (\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1})^2 \mathbf{x}^{\mathsf{T}} \Sigma^{-1} \mathbf{x}\right) \\
&= \mathrm{tr}\left(\mathbf{x}^{\mathsf{T}} \Sigma^{-1} \mathbf{x}\right) - \mathrm{tr}\left(\mathbf{y}^{\mathsf{T}} (\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1} \ \mathbf{x}^{\mathsf{T}} \Sigma^{-1} \mathbf{x}\right)
\end{aligned}$$

where Id is the $D \times D$ identity matrix and where $\mathrm{tr}\left(\mathbf{x}^{\mathsf{T}} \Sigma^{-1} \mathbf{x}\right)$ is a constant independent of k. □

6. Summary for $S > 2$ Models

Mathematical expression for multivariate ($D \geq 1$) optimal attacks
with a linear combination of models ($S \geq 1$):

Leakage model:

$\mathbf{x} = \alpha\mathbf{y}^\star + \mathbf{n}$
$\forall q, n_q \sim \mathcal{N}(0, \Sigma)$
$\mathbf{y}^\star = \phi(\mathbf{t}, k^\star)$
$\mathbf{y} = \phi(\mathbf{t}, k)$
$\mathbf{x} \in \mathbb{R}^{D \times Q}, \mathbf{y} \in \mathbb{R}^{S \times Q}$
$\alpha \in \mathbb{R}^{D \times S}, \Sigma \in \mathbb{R}^{D \times D}$

Is α known?

Optimal distinguisher:

yes → $\mathcal{D}_{ML}(\mathbf{x}, \mathbf{t}) = \mathrm{argmin}_k \ \mathrm{tr}\left((\mathbf{x} - \alpha\mathbf{y})^{\mathsf{T}} \Sigma^{-1} (\mathbf{x} - \alpha\mathbf{y})\right)$

no → $\mathcal{D}_{ML,sto}(\mathbf{x}, \mathbf{t}) = \mathrm{argmax}_k \ \mathrm{tr}\left(\mathbf{y}^{\mathsf{T}} (\mathbf{y}\mathbf{y}^{\mathsf{T}})^{-1} \mathbf{y} \ \mathbf{x}^{\mathsf{T}} \Sigma^{-1} \mathbf{x}\right)$

6*bis*. Summary for $S = 2$ Models

Modus operandi for multivariate ($D \geq 1$) optimal attacks
with one model $\mathbf{Y}$ associated to envelope $\alpha \in \mathbb{R}^{D \times 1}$
and a constant offset $\beta \in \mathbb{R}^{D \times 1}$ ($S = 2$):

Leakage model:

$\mathbf{x} = \alpha\mathbf{y}^\star + \beta\mathbf{1} + \mathbf{n}$ $\quad$ $\mathbf{y}^\star = \phi(\mathbf{t}, k^\star)$ $\quad$ $\mathbf{x} \in \mathbb{R}^{D \times Q}, \mathbf{y} \in \mathbb{R}^{1 \times Q}$
$\forall q, n_q \sim \mathcal{N}(0, \Sigma)$ $\quad$ $\mathbf{y} = \phi(\mathbf{t}, k)$ $\quad$ $\alpha, \beta \in \mathbb{R}^{D \times 1}, \Sigma \in \mathbb{R}^{D \times D}$

Are α, β known?

yes → Affine projection:

$$\tilde{\mathbf{x}} = \frac{\alpha^{\mathsf{T}} \Sigma^{-1}}{\alpha^{\mathsf{T}} \Sigma^{-1} \alpha} (\mathbf{x} - \beta\mathbf{1}) \in \mathbb{R}^{1 \times Q}$$

Univariate ML attack:

$$\mathcal{D}_{\mathrm{ML}}^{S=2}(\mathbf{x}, \mathbf{t}) = \mathrm{argmin}_k \ \|\tilde{\mathbf{x}} - \mathbf{y}\|_2^2$$

no → Data transformation:

$$\mathbf{x}' = \Sigma^{-1/2} \mathbf{x}$$

New multivariate CPA attack:

$$\mathcal{D}_{\mathrm{ML,sto}}^{S=2}(\mathbf{x}, \mathbf{t}) = \mathrm{argmax}_k \ \sum_{d=1}^{D} \frac{\widehat{\mathrm{Cov}}(\mathbf{x}'_d, \mathbf{y})^2}{\widehat{\mathrm{Var}}(\mathbf{y})}$$