Wave: A new code-based signature scheme

Thomas Debris-Alazard[1,2] Nicolas Sendrier [2] Jean-Pierre Tillich [2]

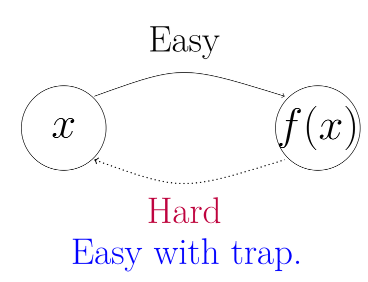
[1] Sorbonne Universités, UPMC Univ Paris 06 [2] Inria, Paris

Results

- The first code-based "hash-and-sign" that strictly follows the GPV strategy (Trapdoor Preimage Sampleable functions) ;
- Security reduction to two problems (NP-complete) of coding theory:
  - Generic decoding of a linear code;
  - Distinguish between random codes and generalized $(U, U+V)$-codes.
- Key Size ≈3MB and signature size ≈13Kbits;
- Feature: uniform signatures through an efficient rejection sampling, one rejection every ≈ 80 signatures.

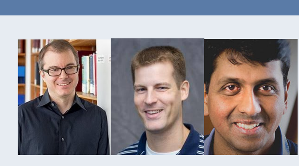
Full Domain Hash (FDH) Signature Schemes

- f be a trapdoor one-way function



- To sign $\mathbf{m}$ one computes $\mathbf{y} = \mathcal{H}(\mathbf{m})$ (hash) and $\sigma \in f^{-1}(\mathbf{y})$.
  → It is required to invert f on all vectors (full domain).
- Verification $f(\sigma) = \mathcal{H}(\mathbf{m})$?

GPV Strategy



It is based on trapdoor one-way preimage sampleable functions!

A family of trapdoor one way-functions $(f_a)_a$ such that distributions:

- $f_a(x)$ is uniformly distributed when

$$x \approx \begin{cases} \text{uniform over words of fixed weight in our case} \\ gaussian \quad \text{for lattices} \end{cases}$$

- algorithm computing f_a^{-1} with the trapdoor

$$\approx \begin{cases} \text{uniform over words of fixed weight in our case} \\ gaussian \quad \text{for lattices} \end{cases}$$

Our Candidate in Code-Based Cryptography

$$f_{\mathbf{H}} : \begin{array}{ccc} \{\mathbf{e} \in \mathbb{F}_q^n : |\mathbf{e}| = w\} & \longrightarrow & \mathbb{F}_q^{n-k} \\ \mathbf{e} & \longmapsto & \mathbf{He}^\intercal \end{array}$$

Inverting $f_{\mathbf{H}}$ amounts to solve the following problem:

Syndrome Decoding Problem :

- Given: $\mathbf{H} \in \mathbb{F}_q^{(n-k)\times n}$, $\mathbf{s} \in \mathbb{F}_q^{n-k}$, and an integer w,
- Find: $\mathbf{e} \in \mathbb{F}_q^n$ such that $\mathbf{He}^\intercal = \mathbf{s}^\intercal$ and $|\mathbf{e}| = w$.

→ Generic problem upon which all code-based cryptography relies.

→ A trapdoor on $f_{\mathbf{H}}$ consists in putting a structure on $\mathbf{H}$!

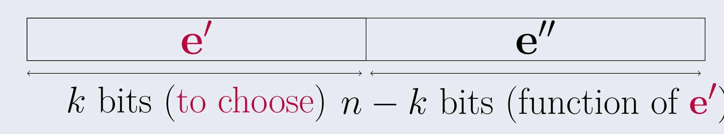
Public-Key: $\mathbf{H}_{\text{pk}}$

Signature of $\mathcal{H}(\mathbf{m})$: $\mathbf{e}$ of weight w with $\mathbf{H}_{\text{pk}}\mathbf{e}^\intercal = \mathcal{H}(\mathbf{m})$.

Hardness of Decoding: Prange Algorithm

Given: $\mathbf{H}$ random of size $(n-k)\times n$, rank $n-k$ and $\mathbf{s} \in \mathbb{F}_q^{n-k}$ random;

Find: $\mathbf{e} \in \mathbb{F}_q^n$ such that $\mathbf{He}^\intercal = \mathbf{s}^\intercal$.



- $\mathbf{e}''$ follows a uniform law over $\mathbb{F}_q^{n-k}$, therefore $\forall \varepsilon > 0, \exists \alpha > 0$:

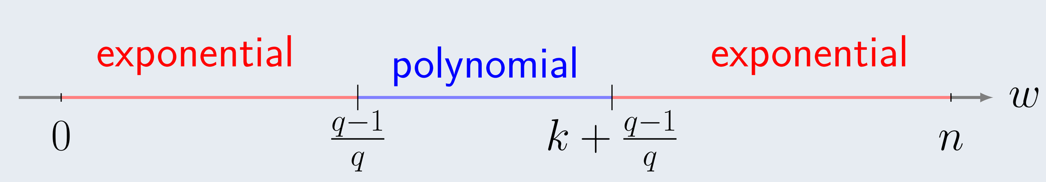
$$\mathbb{E}(|\mathbf{e}''|) = \frac{q-1}{q}(n-k)$$

$$\mathbb{P}\left(\left| |\mathbf{e}''| - \frac{q-1}{q}(n-k) \right| \geq \varepsilon n \right) = e^{-\alpha n}$$

- We get an error $\mathbf{e} = (\mathbf{e}', \mathbf{e}'')$ such that for some $\beta > 0$:

$$\mathbb{E}(|\mathbf{e}|) = \mathbb{E}(|\mathbf{e}'|) + \frac{q-1}{q}(n-k)$$

$$\mathbb{P}\left(|\mathbf{e}| \geq (1+\varepsilon)\left(\mathbb{E}(|\mathbf{e}'|) + \frac{q-1}{q}(n-k)\right)\right) = e^{-\beta n}$$



Our trapdoor: generalized $(U, U+V)$-codes

Let U (resp. V) and be a code over $\mathbb{F}_q$ of length $n/2$, of dimension k_U (resp. k_V) and of parity-check matrix $\mathbf{H}_U$ (resp. $\mathbf{H}_V$).

$$(U, U+V) \triangleq \{(\mathbf{u}, \mathbf{u}+\mathbf{v}) : \mathbf{u} \in U \text{ and } \mathbf{v} \in V\}$$

is code of dimension $k_U + k_V$ and of parity-check matrix:

$$\mathbf{H}_{\text{UV}} \triangleq \begin{pmatrix} \mathbf{H}_U & \mathbf{0} \\ -\mathbf{H}_V & \mathbf{H}_V \end{pmatrix}$$

We restricted our work to the case of: $q = 3$

$$\mathbf{H}_{\text{UV}}\mathbf{e}^\intercal = \mathbf{s}^\intercal \iff \begin{cases} \mathbf{H}_U\mathbf{e}_U^\intercal = \mathbf{s}_U^\intercal \\ \mathbf{H}_V\mathbf{e}_V^\intercal = \mathbf{s}_V^\intercal \end{cases}$$

where: $\mathbf{e} = (\mathbf{e}_U, \mathbf{e}_U + \mathbf{e}_V)$; $\mathbf{s} = (\mathbf{s}_U, \mathbf{s}_V)$

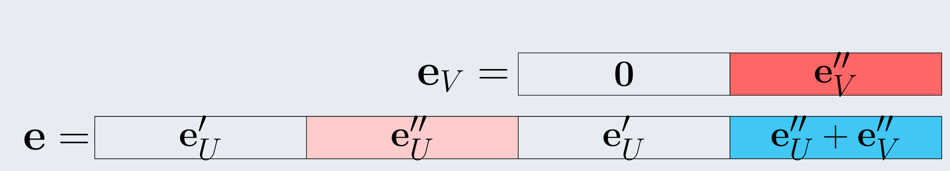
→ Codes U and V are random : we use the Prange algorithm!

(i) firstly to decode in V to get $\mathbf{e}_V$;
(ii) then to decode in U to get $\mathbf{e}_U$ using the knowledge of $\mathbf{e}_V$

We have the freedom to choose:

- k_V (dimension of V) symbols of $\mathbf{e}_V$;
- k_U (dimension of U) symbols of $\mathbf{e}_U$.

We get a final error $\mathbf{e} = (\mathbf{e}_U, \mathbf{e}_U + \mathbf{e}_V) \in \mathbb{F}_3^n$ of shape up to a permutation ($\mathbf{e}_V''$ has only non-zero symbols):



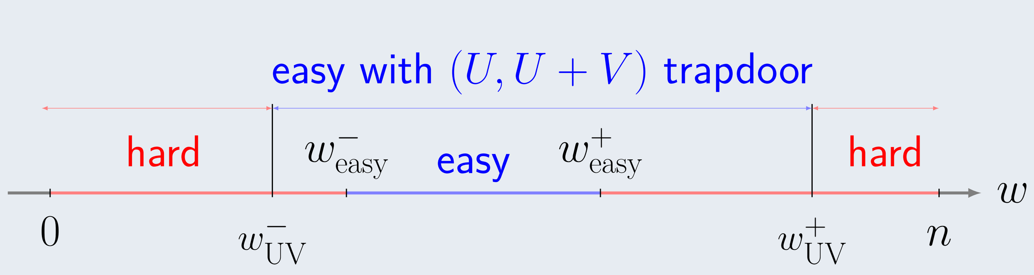
- To reach an error of minimum weight:

Put as many 0's as possible in $\mathbf{e}_U'(i)$ (they are doubled in $\mathbf{e}$).

- To reach an error of maximum weight

Choose k_U symbols $\mathbf{e}_U(i)$ such that: $\begin{cases} \mathbf{e}_U(i) \neq 0 \\ \mathbf{e}_U(i) + \mathbf{e}_V(i) \neq 0 \end{cases}$

→ Possible as $q = 3$ and do not depend of $\mathbf{e}_V(i)$!



The gain is better for large weights!

Achieving the Uniform Distribution

$$\mathbf{e}^{\text{sgn}} \triangleq (\mathbf{e}_U, \mathbf{e}_U + \mathbf{e}_V) \quad (\text{resp. } \mathbf{e}^{\text{unif}} \triangleq (\mathbf{e}_1, \mathbf{e}_2))$$

be a signature (resp. be a uniform word of weight w).

Our goal:

$$\mathbf{e}^{\text{sgn}} \sim \mathbf{e}^{\text{unif}} \iff \begin{cases} \mathbf{e}_U \sim \mathbf{e}_1 \\ \mathbf{e}_V \sim \mathbf{e}_2 - \mathbf{e}_1 \end{cases}$$

→ Having signatures is useless to mount an attack!

Idea for $\mathbf{e}_V \sim \mathbf{e}_2 - \mathbf{e}_1$: rejection sampling.

$$\mathbf{e}_V = \text{Prange}(\mathbf{H}_V, \mathbf{s}_V)$$

Distribution of the Prange algorithm is only depends of the weight:

$$\mathbb{P}(\text{Prange}(\cdot) = \mathbf{e} \mid |\text{Prange}(\cdot)| = |\mathbf{e}|) = \frac{1}{\#\{\mathbf{x} : |\mathbf{x}| = |\mathbf{e}|\}}$$

It is enough to ensure:

$$|\mathbf{e}_V| \sim |\mathbf{e}_2 - \mathbf{e}_1|.$$

By making a rejection sampling on $|\mathbf{e}_V|$:

$$\text{“accept } |\mathbf{e}_V| = i\text{” with probability: } \frac{1}{M}\frac{\mathbb{P}(|\mathbf{e}_2 - \mathbf{e}_1| = i)}{\mathbb{P}(|\mathbf{e}_V| = i)}$$

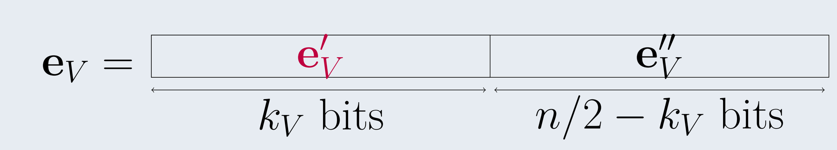
$$M \triangleq \max_j \frac{\mathbb{P}(|\mathbf{e}_2 - \mathbf{e}_1| = j)}{\mathbb{P}(|\mathbf{e}_V| = j)}$$

→ $\frac{1}{M}$ is the average number of reject.

We proceed in essentially the same way for $\mathbf{e}_U$ to get $\mathbf{e}_U \sim \mathbf{e}_1$.

A feasible rejection sampling on $|\mathbf{e}_V|$

- A first Step : $\mathbb{E}(|\mathbf{e}_V|) = \mathbb{E}(|\mathbf{e}_2 - \mathbf{e}_1|)$.

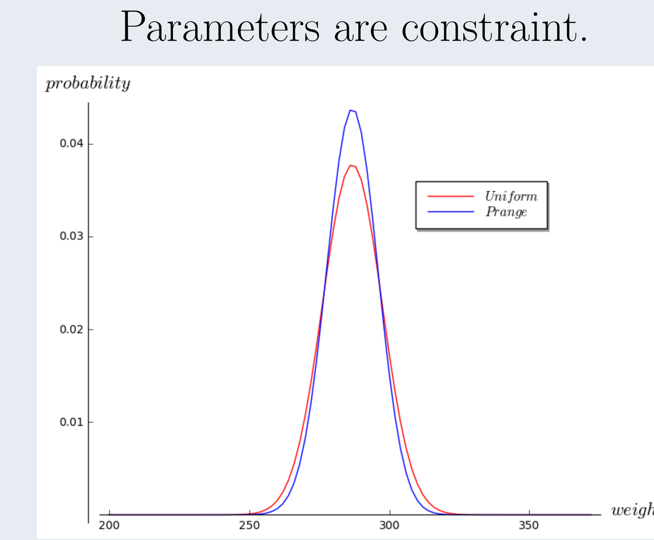


- $\mathbf{e}_V''$ follows a uniform law over $\mathbb{F}_3^{n/2-k}$: $\mathbb{E}(|\mathbf{e}_V''|) = \frac{2}{3}(n/2 - k_V)$
- $\mathbf{e}_V'$ such that: $\mathbb{E}(|\mathbf{e}_V'|) = (1-\alpha)k_V$ with a fixed α.

→ Choose k_V such that:

$$(1-\alpha)k_V + \frac{2}{3}(n/2 - k_V) = \mathbb{E}(|\mathbf{e}_2 - \mathbf{e}_1|).$$

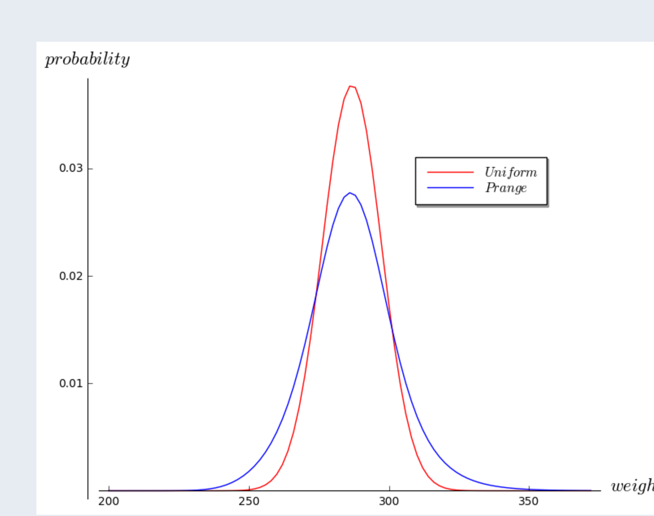
Parameters are constraint.



→ Exponential number of rejects!

In the queue of distribution :
$\mathbb{P}(|\mathbf{e}_V| = i) \ll \mathbb{P}(|\mathbf{e}_2 - \mathbf{e}_1| = i)$

- $\mathbf{e}_V''$ follows a uniform law: its variance is fixed
- Choose $\mathbf{e}_V'$ such that: $\mathbb{E}(|\mathbf{e}_V'|) = (1-\alpha)k_V$ and high variance!



Choice for distribution $|\mathbf{e}_V'|$: large degree of freedom!