Wave: A New Family of Trapdoor One-Way Preimage Sampleable Functions Based on Codes

Thomas Debris-Alazard, Nicolas Sendrier and Jean-Pierre Tillich

Inria Saclay,
EPI GRACE

Code-based Signatures

- Stern Zero Knowledge Protocol 93' + Fiat-Shamir transform 87'

  Long signatures $\approx \Theta(\lambda^2)$ bits

- KKS [Kabatianskii, Krouk, Smeets] 97', $\approx$ Schnorr signature

  At best one-time

- CFS [Courtois, Finiasz, Sendrier] 01', hash and sign,

  Poor scaling, key several gigabytes for 128 bits of security

- RankSign [GRSZ] 14', hash and sign (rank metric),

  Broken by a polynomial time attack

- No code-based signature in the NIST-PQC round 2

- : Durandal [ABGHZ] 19' Eurocrypt (rank metric), Schnorr-Lyubashevsky signature

  Leakage-freeness not proven

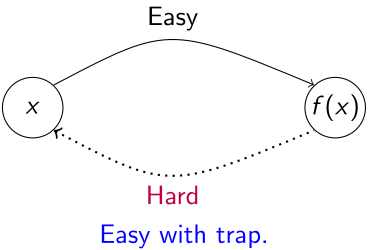
Results

- A code-based "hash-and-sign" ;
- Security reduction to two NP-complete problems in coding theory:
  - Generic decoding of a linear code;
  - Distinguish between random codes and generalized permuted $(U, U+V)$-codes.
- We follow the lattice-based strategy of Gentry-Peikert-Vaikuntanathan (GPV)

  $\rightarrow$ We avoid information leakage
- Nice feature: uniform signatures through an efficient rejection sampling, one rejection every ≈ 100 signatures.
- Key Size $\approx$3MB, signature size $\approx$900B, signing time ≈ 0.1s, implementation available at `http://wave.inria.fr`;

Full Domain Hash Signature

- $\mathcal{H}(\cdot)$ hash function,
- f trapdoor one-way function



- To sign m:

$$\text{Compute } \sigma \in f^{-1}(\mathcal{H}(\text{m})).$$

Code-Based One-Way Function

- $|\cdot|$ denotes the Hamming weight
- H matrix over $\mathbb{F}_q$ with $n-k$ rows and n columns
- w an integer (weight)

One-way in code-based crypto. is:

$$
\begin{array}{rccc}
f_{w,\mathrm{H}} : & \{\mathrm{e} \in \mathbb{F}_q^n : |\mathrm{e}| = w\} & \longrightarrow & \mathbb{F}_q^{n-k} \\
 & \mathrm{e} & \longmapsto & \mathrm{He}^{\mathsf{T}}
\end{array}
$$

To hope $f_{w,\mathrm{H}}$ surjective, choose w big enough

$w \geq (1+\varepsilon)w_{\mathrm{GV}}$ where $q^{n-k} \approx \binom{n}{w_{\mathrm{GV}}}(q-1)^{w_{\mathrm{GV}}}$

Typically we expect an exponential number of pre-images...

Gentry-Peikert-Vaikuntanathan (GPV) Approach

Add properties to $f_{w,\mathrm{H}}$: preimage sampleable function!

- $\stackrel{\$}{\leftarrow}$ means uniformly picked,
- S_w words of Hamming weight w.

1. Trap. algo: $\forall \mathrm{s}$, $\mathrm{e} \leftarrow f_{w,\mathrm{H}}^{-1}(\mathrm{s})$ distributed as $\mathrm{e} \stackrel{\$}{\leftarrow} S_w \cap f_{w,\mathrm{H}}^{-1}(\mathrm{s})$.

Gentry-Peikert-Vaikuntanathan (GPV) Approach

Add properties to $f_{w,\mathrm{H}}$: preimage sampleable function!

- $\stackrel{\$}{\leftarrow}$ means uniformly picked,
- S_w words of Hamming weight w.

1. ~~Trap. algo: $\forall$s, e $\leftarrow f^{-1}_{w,\mathrm{H}}(\mathrm{s})$ distributed as e $\stackrel{\$}{\leftarrow} S_w \cap f^{-1}_{w,\mathrm{H}}(\mathrm{s})$.~~

   We relax to: $f^{-1}_{w,\mathrm{H}}(\mathrm{s}^{\mathrm{unif}}) \stackrel{\$}{\leftarrow} S_w$ for $\mathrm{s}^{\mathrm{unif}}$ uniformly distributed.

   $\rightarrow$ Enough for a security reduction in the ROM

2. $f_{w,\mathrm{H}}(\mathrm{e})$ uniformly distributed when e $\stackrel{\$}{\leftarrow} S_w$,

1 Context

2 Decoding with Our Trapdoor

3 Leakage-Free Signatures

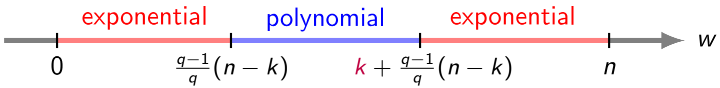
Prange Algorithm

Given: $\mathbf{H} \in \mathbb{F}_q^{(n-k)\times n}$ and $\mathbf{s}$ uniformly distributed over $\mathbb{F}_q^{n-k}$;

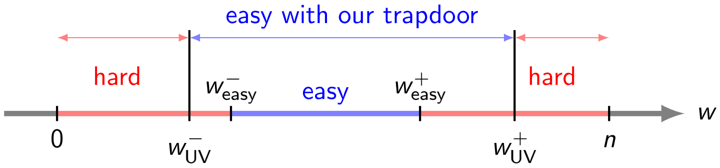
Find: $\mathbf{e} \in \mathbb{F}_q^{n}$ such that (i) $|\mathbf{e}| = w$ and (ii) $\mathbf{H}\mathbf{e}^{\intercal} = \mathbf{s}^{\intercal}$.

$\rightarrow$ Linear system with n unknowns $>$ $n-k$ equations

$\mathbf{e}'$	$\mathbf{e}''$
k symbols (to choose, guide the weight)	$n-k$ symbols (Uniformly Distributed)



Decoding with our Trapdoor



Our Trapdoor (I)

We use special matrices:

$$\mathbf{H}_{\text{sec}} \stackrel{\triangle}{=} \underbrace{\overbrace{\begin{pmatrix} \mathbf{H}_U & 0 \\ -\mathbf{H}_V & \mathbf{H}_V \end{pmatrix}}^{n}}_{n/2} \begin{matrix} \updownarrow n/2 - k_U \\ \updownarrow n/2 - k_V \end{matrix}$$

where $\mathbf{H}_U$ and $\mathbf{H}_V$ are random!

To hide our trapdoor: $\mathbf{P}$ permutation, $\mathbf{S}$ invertible and

$$\mathbf{H}_{\text{pub}} \stackrel{\triangle}{=} \mathbf{S}\mathbf{H}_{\text{sec}}\mathbf{P} \text{ : public}$$

Security Assumption: Distinguishing $\mathbf{H}_{\text{pub}}$/random matrix (same size) is computationally hard.

Proposition

The underlying decision problem is NP-complete.

Our Trapdoor (II)

Let,

$$\mathrm{e} = (\mathrm{e}_U, \mathrm{e}_U + \mathrm{e}_V) \quad ; \quad \mathrm{s} = (\mathrm{s}_U, \mathrm{s}_V)$$

$$\boxed{\mathrm{H}_{\mathrm{sec}}\mathrm{e}^\mathsf{T} = \mathrm{s}^\mathsf{T} \iff \begin{cases} \mathrm{H}_U \mathrm{e}_U^\mathsf{T} = \mathrm{s}_U^\mathsf{T} \\ \mathrm{H}_V \mathrm{e}_V^\mathsf{T} = \mathrm{s}_V^\mathsf{T} \end{cases}}$$

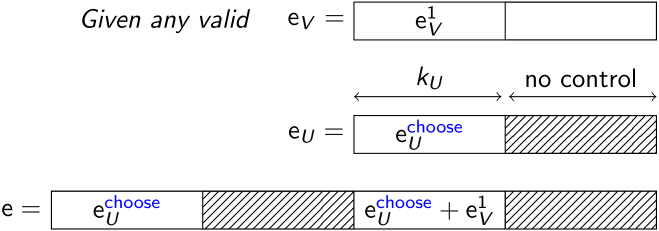
$$k_U + k_V = \mathrm{Ncols}(\mathrm{H}_{\mathrm{sec}}) - \mathrm{Nrows}(\mathrm{H}_{\mathrm{sec}})$$

$$k_U = \mathrm{Ncols}(\mathrm{H}_U) - \mathrm{Nrows}(\mathrm{H}_U) \quad \text{and} \quad k_V = \mathrm{Ncols}(\mathrm{H}_V) - \mathrm{Nrows}(\mathrm{H}_V)$$

$\rightarrow$ Prange directly on $\mathrm{H}_{\mathrm{sec}}$ chooses $k_U + k_V$ symbols of e but here e_U appears twice ($k_U > k_V$)...

Our Decoder

Final error $\mathrm{e} = (\mathrm{e}_U, \mathrm{e}_U + \mathrm{e}_V) \in \mathbb{F}_q^n$ of shape:



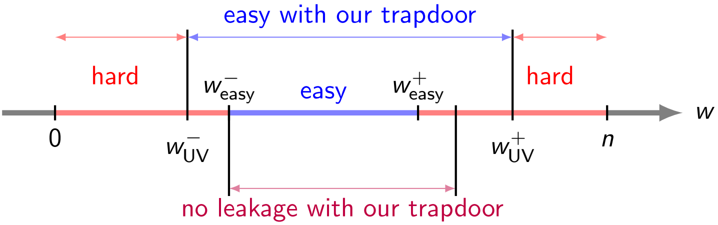
To reach an error of maximum weight

- Choose k_U symbols $\mathrm{e}_U^{\text{choose}}(i)$ s.t: $\begin{cases} \mathrm{e}_U^{\text{choose}}(i) \neq 0 \\ \mathrm{e}_U^{\text{choose}}(i) + \mathrm{e}_V^1(i) \neq 0 \end{cases}$

$\rightarrow$ Possible as we work in $\mathbb{F}_q$ with $q \geq 3$

$\rightarrow$ We gain by choosing $2k_U > k_U + k_V$

Leakage-Free Signatures



We will now work with $q = 3$.

Leakage-Free Signatures

$\mathbf{e}^{\text{sgn}} \stackrel{\triangle}{=} (\mathbf{e}_1^{\text{sgn}}, \mathbf{e}_2^{\text{sgn}})$ signature, $\mathbf{e}^{\text{unif}} \stackrel{\triangle}{=} (\mathbf{e}_1, \mathbf{e}_2)$ unif word of weight w.

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

We would like,

$$\mathbf{e}^{\text{sgn}} \sim \mathbf{e}^{\text{unif}}$$

In a first step we want,

$$\mathbf{e}_V \sim \mathbf{e}_2 - \mathbf{e}_1 \quad \text{where} \quad \mathbf{e}_V = \text{Prange}\left(\mathbf{H}_V, \mathbf{s}_V\right)$$

First approximation, distribution of Prange algorithm, only function of the weight:

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

$\rightarrow$ Uniformity property is enough $|\mathbf{e}_V| \sim |\mathbf{e}_2 - \mathbf{e}_1|$

Guide the Weight of $\mathbf{e}_V$

- We first look for $\mathbb{E}(|\mathbf{e}_V|) = \mathbb{E}(|\mathbf{e}_2 - \mathbf{e}_1|)$ where $\mathbf{e}^{\text{unif}} \stackrel{\triangle}{=} (\mathbf{e}_1, \mathbf{e}_2)$

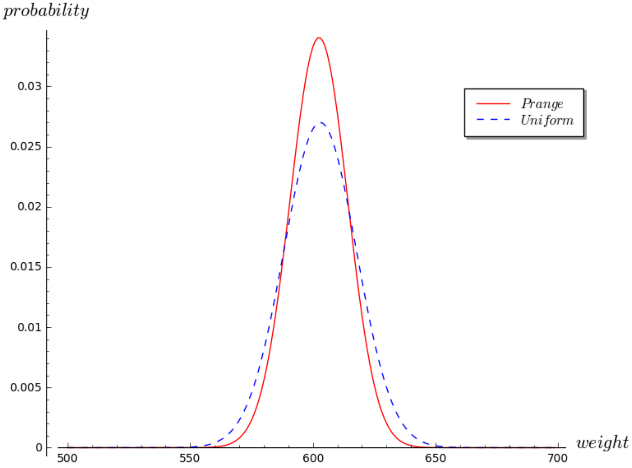
$\mathbf{e}_V =$

$\mathbf{e}'_V$	$\mathbf{e}''_V$
⟵ k_V symbols ⟶	⟵ $n/2 - k_V$ symbols ⟶

- $\mathbf{e}''_V$ follows a uniform law over $\mathbb{F}_3^{n/2-k}$: $\mathbb{E}(|\mathbf{e}''_V|) = \frac{2}{3}(n/2 - k_V)$
- $\mathbf{e}'_V$ can be chosen.

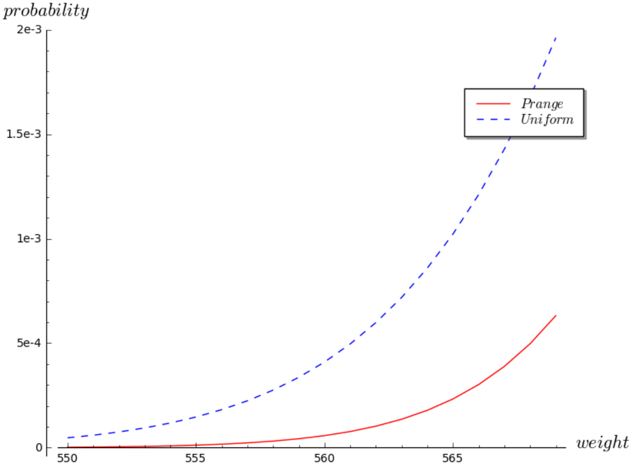
$\rightarrow k_V$ is fixed as: $\mathbb{E}(|\mathbf{e}'_V|) + \frac{2}{3}(n/2 - k_V) = \mathbb{E}\left(|\mathbf{e}_2 - \mathbf{e}_1|\right)$

Rejection Sampling



$$\mathbb{P}(\text{accept}) = \min_j \frac{\mathbb{P}(|e_V| = j)}{\mathbb{P}(|e_2 - e_1| = j)}$$

Rejection Sampling: Tail



$$\mathbb{P}(\text{accept}) = \min_j \frac{\mathbb{P}(|e_V| = j)}{\mathbb{P}(|e_2 - e_1| = j)}$$

Probabilistic Choice of e'_V

$$e_V = \left[\; e'_V \;\middle|\; e''_V \;\right]$$

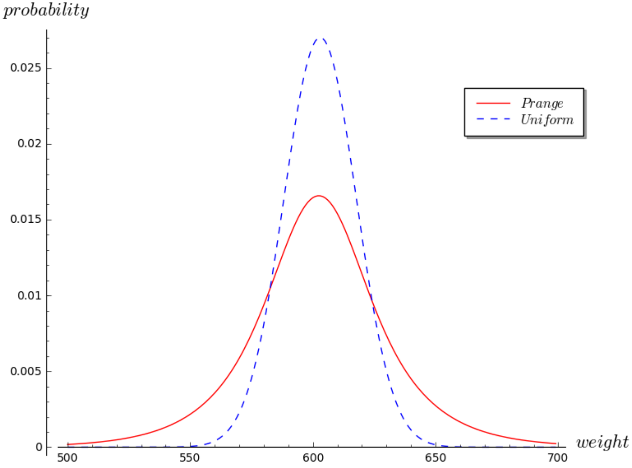
k_V symbols — $n/2 - k_V$ symbols

- e''_V follows a uniform law: its variance is fixed,

Choose the weight of e'_V as a random variable!

- $|e'_V|$ s.t: $\begin{cases} \mathbb{E}(|e'_V|) + \frac{2}{3}(n/2 - k_V) = \mathbb{E}\left(|e_2 - e_1|\right) \\ |e'_V| \text{ high variance!} \end{cases}$

Rejection Sampling



$$\mathbb{P}(\text{accept}) = \min_j \frac{\mathbb{P}(|e_V| = j)}{\mathbb{P}(|e_2 - e_1| = j)}$$

Non-Uniformity of Prange

$$\mathbb{P}(\text{Prange}(\cdot) = \mathbf{x} \mid |\text{Prange}(\cdot)| = |\mathbf{x}|) = \frac{1}{\#\{\mathbf{y} : |\mathbf{y}| = |\mathbf{x}|\}} \quad : \text{only} \approx.$$

Given $\mathbf{H} \in \mathbb{F}_3^{(n-k)\times n}$ and $\mathbf{s} \in \mathbb{F}_3^{n-k}$ find $\mathbf{e} \in \mathbb{F}_3^{n}$ s.t $\mathbf{H}\mathbf{e}^{\intercal} = \mathbf{s}^{\intercal}$.

$\rightarrow$ Linear system with n unknowns $> n-k$ equations

$\mathbf{e}_1^{\text{choose}}$	Unif. Distrib.
⟷ k symbols	⟷ $n-k$ columns where $\mathbf{H}$ invertible

Non-Uniformity of Prange

$$\mathbb{P}(\text{Prange}(\cdot) = \mathbf{x} \mid |\text{Prange}(\cdot)| = |\mathbf{x}|) = \frac{1}{\#\{\mathbf{y} : |\mathbf{y}| = |\mathbf{x}|\}} \quad : \text{only } \approx.$$

Given $\mathbf{H} \in \mathbb{F}_3^{(n-k)\times n}$ and $\mathbf{s} \in \mathbb{F}_3^{n-k}$ find $\mathbf{e} \in \mathbb{F}_3^n$ s.t $\mathbf{H}\mathbf{e}^{\intercal} = \mathbf{s}^{\intercal}$.

$\rightarrow$ Linear system with n unknowns $> n-k$ equations

$\mathbf{e}_1^{\text{choose}}$	Unif. Distrib.

k symbols — $n-k$ columns where $\mathbf{H}$ invertible

$\mathbf{e}_1^{\text{choose}}$	

$k-d$ symbols — $n-k+d$ columns where $\mathbf{H}$ of rank $n-k$

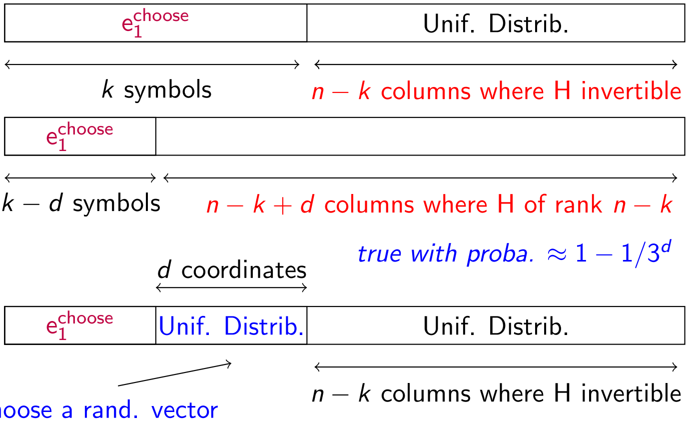
true with proba. $\approx 1 - 1/3^d$

Non-Uniformity of Prange

$$\mathbb{P}(\text{Prange}(\cdot) = \mathbf{x} \mid |\text{Prange}(\cdot)| = |\mathbf{x}|) = \frac{1}{\#\{\mathbf{y} : |\mathbf{y}| = |\mathbf{x}|\}} \quad : \text{only} \approx.$$

Given $\mathbf{H} \in \mathbb{F}_3^{(n-k)\times n}$ and $\mathbf{s} \in \mathbb{F}_3^{n-k}$ find $\mathbf{e} \in \mathbb{F}_3^n$ s.t $\mathbf{H}\mathbf{e}^\intercal = \mathbf{s}^\intercal$.

$\rightarrow$ Linear system with n unknowns $> n-k$ equations



Reaching Uniform Signatures

Theorem

Let $\mathbf{e}^{\text{sgn}}$ be a signature, $\mathbf{e}^{\text{unif}}$ be a uniformly distributed error of weight w. We have for P, Q polynomials and Δ statistical. dist.

$$\mathbb{P}_{\mathbf{H}_{\text{pub}}}\left(\Delta(\mathbf{e}^{\text{sgn}}, \mathbf{e}^{\text{unif}}) > Q(d)3^{-d/2}\right) \leq P(d)3^{-d/2}.$$

We can improve $d/2 \longrightarrow d$

We also prove:

$$\Delta(\mathbf{H}_{\text{pub}}\mathbf{e}^{\intercal}, \mathbf{s}^{\text{unif}}) \text{ negligible} \quad \text{where} \quad \mathbf{e} \overset{\$}{\leftarrow} S_w \quad \text{and} \quad \mathbf{s}^{\text{unif}} \overset{\$}{\leftarrow} \mathbb{F}_3^{n-k}$$

Conclusion

- The first code-based "hash-and-sign" based on NP-complete problems that follows the GPV strategy;
- Scalability of the scheme (in bits):

$$\text{signature length} = 105\lambda \quad \text{and} \quad \text{keySize} = 1565\lambda^2$$

Ongoing Work:

- Algorithms to distinguish permuted generalized $(U, U + V)$-codes and random codes: currently decoding algorithms;
- Hope to remove the rejection sampling

  $\rightarrow$ Many degrees of freedom in the Prange algorithm!

Conclusion

- The first code-based "hash-and-sign" based on NP-complete problems that follows the GPV strategy;

- Scalability of the scheme (in bits):

$$\text{signature length} = 105\lambda \quad \text{and} \quad \text{keySize} = 1565\lambda^2$$

Ongoing Work:

- Algorithms to distinguish permuted generalized $(U, U+V)$-codes and random codes: currently decoding algorithms;

- Hope to remove the rejection sampling

  $\rightarrow$ Many degrees of freedom in the Prange algorithm!

Thank You!