Code-Based Cryptography with Restricted Errors

Anna-Lena Horlemann

University of St. Gallen, Switzerland

Monash Cybersecurity Seminar
April 26th, 2021

Joint work with Marco Baldi, Massimo Battaglioni, Franco Chiaraluce, Edoardo Persichetti, Paolo Santini and Violetta Weger.

Post-quantum cryptography

- Most asymmetric cryptosystems in use are based on one of the following hard problems:
  - integer factorization
  - discrete logarithm
  - elliptic curve discrete logarithm
- These problems are not "hard enough" anymore on quantum computers, due to Shor's algorithm (1994).
- This calls for different cryptosystems, based on other hard mathematical problems. This is the field of *post-quantum cryptography*.
- *Time to change:* NSA announced in 2015 to change to post-quantum cryptography. NIST currently runs a standardization project.

Post-quantum cryptography

The most studied proposed hard problems are

- the general syndrome decoding problem (code-based cryptography)
- the lattice shortest vector problem (lattice-based cryptography)
- inverting hash functions (hash-based cryptography)
- solving systems of multivariate polynomial equations (multivariate cryptography)
- walks in a supersingular isogeny graph (supersingular elliptic curve isogeny cryptography)

Code-based cryptography

- Two general schemes for public key cryptosystems: *McEliece* and *Niederreiter*.
- The original McEliece proposal uses binary Goppa codes and has been unbroken for 30 years.
- Advantage: Computations are fast.
- Disadvantage: Public key size is a lot larger than in other systems.

Code-based cryptography

- Two general schemes for public key cryptosystems: *McEliece* and *Niederreiter*.
- The original McEliece proposal uses binary Goppa codes and has been unbroken for 30 years.
- Advantage: Computations are fast.
- Disadvantage: Public key size is a lot larger than in other systems.

- Related *signature scheme* by Courtois, Finiasz, Sendrier (2001), and other variants.
- Wave signature scheme (Debris-Alazard, Sendrier, Tillich), using high weight errors (2019).

Code-based cryptography

- Two general schemes for public key cryptosystems: *McEliece* and *Niederreiter*.
- The original McEliece proposal uses binary Goppa codes and has been unbroken for 30 years.
- Advantage: Computations are fast.
- Disadvantage: Public key size is a lot larger than in other systems.

- Related *signature scheme* by Courtois, Finiasz, Sendrier (2001), and other variants.
- Wave signature scheme (Debris-Alazard, Sendrier, Tillich), using high weight errors (2019).
- Code-based *zero-knowledge identification schemes*: Stern (1994), . . . , Cayrel-Véron-ElYousfiAlaoui (2011), . . .
- Can create signatures via the Fiat-Shamir transform (1986).

Main idea of code-based identification schemes

- Decoding a random linear code is a hard problem.
- *Public key/information*: the parity check matrix of a random linear code, and a syndrome
- *Secret*: the solution to the corresponding syndrome decoding problem: usually a low-weight (error) vector
- The *prover* wants to prove to the *verifier* that he/she knows the secret.

$$\underbrace{\mathbf{s}}_{\text{syndrome}} = \underbrace{\mathbf{e}}_{\text{error vector}} \cdot \underbrace{\mathbf{H}^\top}_{\text{PC matrix}}$$

The CVE scheme

Public Data	Parameters $q, n, k, \omega \in \mathbb{N}$, parity-check matrix $\mathbf{H} \in \mathbb{F}_q^{(n-k)\times n}$
Private Key	$\mathbf{e} \in \mathrm{S}^{\mathrm{H}}_{n,q,\omega}$
Public Key	$\mathbf{s} = \mathbf{e}\mathbf{H}^\top \in \mathbb{F}_q^{n-k}$

PROVER		VERIFIER
Choose $\mathbf{u} \xleftarrow{\$} \mathbb{F}_q^n$, $\tau \xleftarrow{\$} \mathfrak{M}_n$		
Set $c_0 = \mathsf{Hash}\big(\tau, \mathbf{u}\mathbf{H}^\top\big)$		
Set $c_1 = \mathsf{Hash}\big(\tau(\mathbf{u}), \tau(\mathbf{e})\big)$		
	$\xrightarrow{c_0, c_1}$	
		Choose $z \xleftarrow{\$} \mathbb{F}_q^*$
	$\xleftarrow{z}$	
Set $\mathbf{y} = \tau(\mathbf{u} + z\mathbf{e})$		
	$\xrightarrow{\mathbf{y}}$	
		Choose $b \xleftarrow{\$} \{0, 1\}$
	$\xleftarrow{b}$	
If $b = 0$, set $f := \tau$		
If $b = 1$, set $f := \mathbf{e}' = \tau(\mathbf{e})$		
	$\xrightarrow{f}$	
		If $b = 0$, accept if $c_0 = \mathsf{Hash}\big(\tau, \tau^{-1}(\mathbf{y})\mathbf{H}^\top - z\mathbf{s}\big)$
		If $b = 1$, accept if $\mathrm{wt}_{\mathrm{H}}(\mathbf{e}') = \omega$ and $c_1 = \mathsf{Hash}\big(\mathbf{y} - z\mathbf{e}', \mathbf{e}'\big)$

The CVE scheme

- The scheme is *complete*, i.e., an honest prover always gets accepted.
- It is *zero-knowledge*, because neither the verifier nor an observer can learn any information about the secret.
- It is *sound*, since an impersonator has only a small probability of getting accepted. This *cheating probability* is $\frac{q}{2(q-1)}$. Repeating the protocol several times decreases this to the wanted security level.

The CVE scheme

- The scheme is *complete*, i.e., an honest prover always gets accepted.
- It is *zero-knowledge*, because neither the verifier nor an observer can learn any information about the secret.
- It is *sound*, since an impersonator has only a small probability of getting accepted. This *cheating probability* is $\frac{q}{2(q-1)}$. Repeating the protocol several times decreases this to the wanted security level.
- For parameter choices, we assume that a randomly chosen code achieves the *Gilbert-Varshamov bound*, and choose the weight w to be $(d-1)/2$.
- The security level depends on the fastest generic decoder to find the secret.
- The *communication cost* is the sizes of the information sent from prover to verifier or vice versa. It corresponds to the signature size in the Fiat-Shamir transform and should hence be minimized.

Our idea: secret lives somewhere else

Our idea: secret lives somewhere else

- *Original idea:* replace Hamming sphere $\mathrm{S}^{\mathrm{H}}_{n,q,\omega}$ with Lee sphere $\mathrm{S}^{\mathrm{L}}_{n,q,\omega} := \{\mathbf{v} \in \mathbb{Z}_q^n \mid \mathrm{wt}_L(\mathbf{v}) = w\}$.

Our idea: secret lives somewhere else

- *Original idea:* replace Hamming sphere $\mathrm{S}^{\mathrm{H}}_{n,q,\omega}$ with Lee sphere $\mathrm{S}^{\mathrm{L}}_{n,q,\omega} := \{\mathbf{v} \in \mathbb{Z}_q^n \mid \mathrm{wt}_L(\mathbf{v}) = w\}$.
- *Problem:* the only linear isometries for the Lee metric are monomial matrices with entries in $\{0, \pm 1\}$. But these are not transitive on $\mathrm{S}^{\mathrm{L}}_{n,q,\omega}$.
  $\Longrightarrow$ Attacker knows the exact number of $a \in \mathbb{Z}_q$ that appear in secret $\mathbf{e}$.

Our idea: secret lives somewhere else

- *Original idea:* replace Hamming sphere $\mathrm{S}^{\mathrm{H}}_{n,q,\omega}$ with Lee sphere $\mathrm{S}^{\mathrm{L}}_{n,q,\omega} := \{\mathbf{v} \in \mathbb{Z}_q^n \mid \mathrm{wt}_L(\mathbf{v}) = w\}$.
- *Problem:* the only linear isometries for the Lee metric are monomial matrices with entries in $\{0, \pm 1\}$. But these are not transitive on $\mathrm{S}^{\mathrm{L}}_{n,q,\omega}$.
  $\Longrightarrow$ Attacker knows the exact number of $a \in \mathbb{Z}_q$ that appear in secret $\mathbf{e}$.
- Then the ambient set for the error becomes the orbit of these linear isometries.
  $\Longrightarrow$ Let's take the simplest (and "cheapest") orbit, where $\mathbf{e} \in \{0, \pm 1\}^n$.

Our idea: secret lives somewhere else

- *Original idea:* replace Hamming sphere $\mathrm{S}^{\mathrm{H}}_{n,q,\omega}$ with Lee sphere $\mathrm{S}^{\mathrm{L}}_{n,q,\omega} := \{\mathbf{v} \in \mathbb{Z}_q^n \mid \mathrm{wt}_L(\mathbf{v}) = w\}$.
- *Problem:* the only linear isometries for the Lee metric are monomial matrices with entries in $\{0, \pm 1\}$. But these are not transitive on $\mathrm{S}^{\mathrm{L}}_{n,q,\omega}$.
  $\Longrightarrow$ Attacker knows the exact number of $a \in \mathbb{Z}_q$ that appear in secret $\mathbf{e}$.
- Then the ambient set for the error becomes the orbit of these linear isometries.
  $\Longrightarrow$ Let's take the simplest (and "cheapest") orbit, where $\mathbf{e} \in \{0, \pm 1\}^n$.

Restricted errors!
(and Lee metric equals Hamming metric)

Coding theory for restricted errors

The SDP for restricted errors

Restricted Hamming ball of radius w and parameter a:

$$E_{n,q,w}^{(a)} := \left\{ \mathbf{e} \in \mathbb{F}_q^n \mid \mathrm{wt}_{\mathrm{H}}(\mathbf{e}) \leq w, \;\; \mathbf{e} \in \{0, \pm x_1, \cdots, \pm x_a\}^n \right\}.$$

Restricted Syndrome Decoding Problem (R-SDP)

Let $q = p^m$, with $p \neq 2$ being a prime. On input $\mathbf{H} \in \mathbb{F}_q^{(n-k)\times n}$, $\mathbf{s} \in \mathbb{F}_q^{n-k}$ and $w \in \mathbb{N}$, decide whether there exists an $\mathbf{e} \in E_{n,q,w}^{(a)}$, such that $\mathbf{e}\mathbf{H}^\top = \mathbf{s}$.

The SDP for restricted errors

Restricted Hamming ball of radius w and parameter a:

$$E_{n,q,w}^{(a)} := \left\{ \mathbf{e} \in \mathbb{F}_q^n \mid \mathrm{wt}_{\mathrm{H}}(\mathbf{e}) \leq w, \;\; \mathbf{e} \in \{0, \pm x_1, \cdots, \pm x_a\}^n \right\}.$$

Restricted Syndrome Decoding Problem (R-SDP)

Let $q = p^m$, with $p \neq 2$ being a prime. On input $\mathbf{H} \in \mathbb{F}_q^{(n-k)\times n}$, $\mathbf{s} \in \mathbb{F}_q^{n-k}$ and $w \in \mathbb{N}$, decide whether there exists an $\mathbf{e} \in E_{n,q,w}^{(a)}$, such that $\mathbf{e}\mathbf{H}^\top = \mathbf{s}$.

Theorem

The R-SDP is NP-complete.

Proof: Reduction from the classical SDP.

Now we restrict to $\mathbf{e} \in \{0, \pm 1\}^n$...

Restricted weight and minimum distance

Definition

Let $\mathbb{F}_q$ have characteristic ≥ 5. For $\mathbf{a} \in \mathbb{F}_q^n$, we define the *restricted weight* of $\mathbf{a}$ as

$$\widetilde{\mathrm{wt}}(\mathbf{a}) := \begin{cases} \#_1(\mathbf{a}) + 2 \cdot \#_2(\mathbf{a}) & \text{if } \mathbf{a} \in \{0, \pm 1, \pm 2\}^n, \\ \infty & \text{otherwise,} \end{cases}$$

where $\#_1(\mathbf{a})$ is the number of entries of $\mathbf{a}$ equal to ± 1 and $\#_2(\mathbf{a})$ is the number of entries equal to ± 2.
Let $\mathcal{C}$ be a linear code with length n over $\mathbb{F}_q$. We define its *restricted minimum distance* as

$$\widetilde{d} := \min\left\{\widetilde{\mathrm{wt}}(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C} \cap (\{0, \pm 1, \pm 2\}^n \setminus \{\mathbf{0}_n\})\right\}.$$

If $\mathcal{C} \cap \{0, \pm 1, \pm 2\}^n = \{\mathbf{0}_n\}$, then we set $\widetilde{d} = \infty$.

(restricted minimum distance = min. Lee distance of restricted code)

Unique decodability

$E_{n,q,w} := \{\mathbf{e} \in \{0, \pm 1\} \mid \mathrm{wt}_H(\mathbf{e}) \leq w\}$

Theorem

Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ have restricted minimum distance $\widetilde{d}$. For any parity-check matrix $\mathbf{H}$ for $\mathcal{C}$, and for all $w < \widetilde{d}/2$, there cannot exist two distinct vectors $\mathbf{e}, \mathbf{e}' \in E_{n,q,w}$ such that $\mathbf{e}\mathbf{H}^\top = \mathbf{e}'\mathbf{H}^\top$.

Unique decodability

$E_{n,q,w} := \{\mathbf{e} \in \{0, \pm 1\} \mid \mathrm{wt}_H(\mathbf{e}) \leq w\}$

Theorem

Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ have restricted minimum distance $\widetilde{d}$. For any parity-check matrix $\mathbf{H}$ for $\mathcal{C}$, and for all $w < \widetilde{d}/2$, there cannot exist two distinct vectors $\mathbf{e}, \mathbf{e}' \in E_{n,q,w}$ such that $\mathbf{e}\mathbf{H}^\top = \mathbf{e}'\mathbf{H}^\top$.

Proof:

- $\mathbf{e}\mathbf{H}^\top = \mathbf{e}'\mathbf{H}^\top \iff (\mathbf{e} - \mathbf{e}')$ is codeword;
- $\mathbf{e} - \mathbf{e}' \in \{0, \pm 1, \pm 2\}$ and has restricted weight less than $\widetilde{d}$;
- $\mathbf{e} - \mathbf{e}' = \mathbf{0} \iff \mathbf{e} = \mathbf{e}'$.

Unique decodability

$E_{n,q,w} := \{\mathbf{e} \in \{0, \pm 1\} \mid \mathrm{wt}_H(\mathbf{e}) \leq w\}$

Theorem

Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ have restricted minimum distance $\widetilde{d}$. For any parity-check matrix $\mathbf{H}$ for $\mathcal{C}$, and for all $w < \widetilde{d}/2$, there cannot exist two distinct vectors $\mathbf{e}, \mathbf{e}' \in E_{n,q,w}$ such that $\mathbf{e}\mathbf{H}^\top = \mathbf{e}'\mathbf{H}^\top$.

Proof:

- $\mathbf{e}\mathbf{H}^\top = \mathbf{e}'\mathbf{H}^\top \iff (\mathbf{e} - \mathbf{e}')$ is codeword;
- $\mathbf{e} - \mathbf{e}' \in \{0, \pm 1, \pm 2\}$ and has restricted weight less than $\widetilde{d}$;
- $\mathbf{e} - \mathbf{e}' = \mathbf{0} \iff \mathbf{e} = \mathbf{e}'$.

Note: If $\mathrm{char}(\mathbb{F}_q) = 3$, then $\mathbf{e} - \mathbf{e}' \in \mathbb{F}_3$ and similar result holds.

Gilbert-Varshamov bound

Theorem

For a given finite restricted minimum distance $\widetilde{d}$ and length n, there exists a code in $\mathbb{F}_q^n$ of dimension $\tilde{k}$, where

$$\begin{aligned}\tilde{k} &\geq n - 1 - \log_q \left(\widetilde{V}(n, \widetilde{d} - 1) \right) \\ &= n - 1 - \log_q \left(\sum_{i=0}^{\widetilde{d}-1} \sum_{j=\max\{0, i-n\}}^{\lfloor i/2 \rfloor} \binom{n}{j} \binom{n-j}{i-2j} 2^{i-j} \right).\end{aligned}$$

Gilbert-Varshamov bound

Theorem

For a given finite restricted minimum distance $\widetilde{d}$ and length n, there exists a code in $\mathbb{F}_q^n$ of dimension $\tilde{k}$, where

$$\begin{aligned}\tilde{k} &\geq n - 1 - \log_q\left(\widetilde{V}(n, \widetilde{d} - 1)\right) \\ &= n - 1 - \log_q\left(\sum_{i=0}^{\widetilde{d}-1} \sum_{j=\max\{0,i-n\}}^{\lfloor i/2 \rfloor} \binom{n}{j}\binom{n-j}{i-2j} 2^{i-j}\right).\end{aligned}$$

Definition

For a code with length n and dimension k over $\mathbb{F}_q$, we define the *restricted Gilbert-Varshamov distance* as

$$\widetilde{d}_{\mathsf{GV}} := \begin{cases} \infty & \text{if } \widetilde{V}(n, 2n) < q^{n-k-1}, \\ \max\left\{\widetilde{d} > 0 \;\middle|\; \widetilde{V}(n, \widetilde{d}) \leq q^{n-k-1}\right\} & \text{otherwise.} \end{cases}$$

Random codes and the GV-bound

Note that $\widetilde{V}(n, 2n) = 5^n$, and hence for large values of q and/or n (if $q > 5$), the restricted Gilbert-Varshamov distance is equal to ∞, as long as $k < n(1 - \log_q(5) - \frac{1}{n})$.

Theorem

Let $q > 5$ be an odd prime power and $k \leq n\left(1 - \log_q(5) - \epsilon\right)$, for $0 < \epsilon < 1 - \log_q(5)$. Let $\mathbf{G} \xleftarrow{\$} \mathbb{F}_q^{k \times n}$ with rank k. Then the code generated by $\mathbf{G}$ has restricted minimum distance $\widetilde{d} = \infty$ with probability at least $1 - q^{-\epsilon n}$.

Random codes and the GV-bound

Note that $\widetilde{V}(n, 2n) = 5^n$, and hence for large values of q and/or n (if $q > 5$), the restricted Gilbert-Varshamov distance is equal to ∞, as long as $k < n(1 - \log_q(5) - \frac{1}{n})$.

Theorem

Let $q > 5$ be an odd prime power and $k \leq n\left(1 - \log_q(5) - \epsilon\right)$, for $0 < \epsilon < 1 - \log_q(5)$. Let $\mathbf{G} \xleftarrow{\$} \mathbb{F}_q^{k \times n}$ with rank k. Then the code generated by $\mathbf{G}$ has restricted minimum distance $\widetilde{d} = \infty$ with probability at least $1 - q^{-\epsilon n}$.

$\implies$ In a random code with $k \leq n\left(1 - \log_q(5) - \epsilon\right)$ we can uniquely decode ANY $\mathbf{e} \in \{0, \pm 1\}^n$.

Now we restrict to prime fields $\mathbb{F}_p$ and $\mathbf{e} \in \{\pm 1\}^n$ (full support)...

The CVE scheme with restricted errors

Public Data	Parameters $p, n, k \in \mathbb{N}$, parity-check matrix $\mathbf{H} \in \mathbb{F}_p^{(n-k)\times n}$
Private Key	$\mathbf{e} \in \mathrm{E}_{n,p,n}$
Public Key	$\mathbf{s} = \mathbf{e}\mathbf{H}^\top \in \mathbb{F}_p^{n-k}$

PROVER		VERIFIER
Choose $\mathbf{u} \xleftarrow{\$} \mathbb{F}_p^n$, $\tau \xleftarrow{\$} \widetilde{\mathfrak{M}}_n$		
Set $c_0 = \mathsf{Hash}\big(\tau, \mathbf{u}\mathbf{H}^\top\big)$		
Set $c_1 = \mathsf{Hash}\big(\tau(\mathbf{u}), \tau(\mathbf{e})\big)$		
	$\xrightarrow{c_0, c_1}$	
		Choose $z \xleftarrow{\$} \mathbb{F}_p^*$
	$\xleftarrow{z}$	
Set $\mathbf{y} = \tau(\mathbf{u} + z\mathbf{e})$		
	$\xrightarrow{\mathbf{y}}$	
		Choose $b \xleftarrow{\$} \{0, 1\}$
	$\xleftarrow{b}$	
If $b = 0$, set $f := \tau$		
If $b = 1$, set $f := \mathbf{e}' = \tau(\mathbf{e})$		
	$\xrightarrow{f}$	
		If $b = 0$, accept if $c_0 = \mathsf{Hash}\big(\tau, \tau^{-1}(\mathbf{y})\mathbf{H}^\top - z\mathbf{s}\big)$
		If $b = 1$, accept if $\mathbf{e}' \in \mathrm{E}_{n,p,n}$ and $c_1 = \mathsf{Hash}\big(\mathbf{y} - z\mathbf{e}', \mathbf{e}'\big)$

Security analysis

- *Completeness* (as before)
- *Zero-knowledge* (as before)
- *Soundness:* Adversary $\mathcal{A}$ can have two strategies.
  1. Strategy 1: Find any $\hat{\mathbf{e}}$ with $\hat{\mathbf{e}}\mathbf{H}^\top = \mathbf{s}$. Can answer challenge $b = 0$ successfully.
  2. Strategy 2: Choose any $\hat{\mathbf{e}}$ with $\widetilde{\mathrm{wt}}(\hat{\mathbf{e}}) = n$. Can answer challenge $b = 1$ successfully.

  Moreover, if $\mathcal{A}$ can guess z correctly, it can solve the other challenge, respectively.

Security analysis

- *Completeness* (as before)
- *Zero-knowledge* (as before)
- *Soundness:* Adversary $\mathcal{A}$ can have two strategies.
  1. Strategy 1: Find any $\hat{\mathbf{e}}$ with $\hat{\mathbf{e}}\mathbf{H}^\top = \mathbf{s}$. Can answer challenge $b = 0$ successfully.
  2. Strategy 2: Choose any $\hat{\mathbf{e}}$ with $\widetilde{\mathrm{wt}}(\hat{\mathbf{e}}) = n$. Can answer challenge $b = 1$ successfully.

  Moreover, if $\mathcal{A}$ can guess z correctly, it can solve the other challenge, respectively.
  Cheating probability:

$$\begin{aligned}\mathsf{Pr}[\mathcal{A} \text{ is accepted}] &= \sum_{i=0}^{1} \frac{1}{2}(\mathsf{Pr}[b = i] + \mathsf{Pr}[b = 1 - i]\mathsf{Pr}[z = \hat{z}]) \\ &= \frac{p}{2(p-1)} \xrightarrow{p\to\infty} \frac{1}{2}\end{aligned}$$

Communication cost

- To have cheating probability at most 2^{-t}, we need N rounds with

$$N = \left\lceil \frac{-t}{\log_2\left(\frac{p}{2(p-1)}\right)} \right\rceil \approx t.$$

Communication cost

- To have cheating probability at most 2^{-t}, we need N rounds with

$$N = \left\lceil \frac{-t}{\log_2\left(\frac{p}{2(p-1)}\right)} \right\rceil \approx t.$$

- *Average communication cost* (with seeds for random elements and compression techniques):

$$l_{\mathsf{Hash}} + N\left(\lceil \log_2(p-1) \rceil + n \lceil \log_2(p) \rceil + 1 + l_{\mathsf{Hash}} + \frac{n + l_{\mathsf{Seed}}}{2} \right).$$

- *Maximal communication cost*:

$$l_{\mathsf{Hash}} + N\left(\lceil \log_2(p-1) \rceil + n \lceil \log_2(p) \rceil + 1 + l_{\mathsf{Hash}} + \max\{n\ ,\ l_{\mathsf{Seed}}\} \right).$$

Generic decoder

To estimate the difficulty of finding the secret $\mathbf{e}$ from $\mathbf{H}$ and $\mathbf{s}$, we adapt known generic decoders to the restricted error setting.

- Information set decoders (ISD) are not useful, since we have full support in the error vector.
- Generalized birthday decoders are the most efficient.
- We use partial Gaussian elimination to reduce the R-SDP into a smaller R-SDP, which we solve with Wagner's algorithm ($\approx$ subset sum solver).

Partial Gaussian elimination and small R-SDP approach

1) *Permutation*: pick a random $\pi \in \mathfrak{S}_n$ and apply $\pi(\mathbf{H})$.
2) *PGE*: find $\mathbf{S} \in \mathbb{F}_p^{(n-k)\times(n-k)}$, such that

$$\mathbf{S}\pi(\mathbf{H}) = \begin{bmatrix} \mathbf{I}_{n-k-\ell} & \mathbf{H}' \in \mathbb{F}_p^{(n-k-\ell)\times(k+\ell)} \\ \mathbf{0}_{\ell\times(n-k-\ell)} & \mathbf{H}'' \in \mathbb{F}_p^{\ell\times(k+\ell)} \end{bmatrix}.$$

   Write $\mathbf{sS} = [\mathbf{s}', \mathbf{s}'']$, with $\mathbf{s}' \in \mathbb{F}_p^{n-k-\ell}$ and $\mathbf{s}'' \in \mathbb{F}_p^{\ell}$.
3) *Small R-SDP*: produce a set $\mathcal{E} \subseteq \{\pm 1\}^{k+\ell}$ containing (some) solutions to the R-SDP instance represented by $\mathbf{H}''$ and $\mathbf{s}''$, i.e.,

$$\mathbf{s}'' = \mathbf{H}''\mathbf{e}''^{\top}, \ \ \forall \mathbf{e}'' \in \mathcal{E}.$$

4) *Test*: for each $\mathbf{e}'' \in \mathcal{E}$, test whether $\mathbf{e}' = \mathbf{s}' - \mathbf{e}''\mathbf{H}'^{\top}$ has entries over $\{\pm 1\}$. If such a vector $\mathbf{e}'$ is found, return $\pi^{-1}\big([\mathbf{e}', \mathbf{e}'']\big)$; otherwise, restart from step 3.

Wagner's algorithm for small R-SDP (one level)

Solve $\mathbf{s}'' = \mathbf{H}''\mathbf{e}''^{\top}$ for $\mathbf{e}'' \in \{\pm 1\}^{k+\ell}$.

1) Choose random subsets $\mathcal{R}_0, \mathcal{R}_1 \subseteq \{\pm 1\}^{(k+\ell)/2}$, each of size 2^v.
2) Build lists

$$\mathcal{L}_0 = \left\{ (\mathbf{z} = \mathbf{p}\mathbf{H}_0''^{\top}, \mathbf{p}) \mid \mathbf{p} \in \mathcal{R}_0 \right\},$$

$$\mathcal{L}_1 = \left\{ (\mathbf{z} = \mathbf{p}\mathbf{H}_1''^{\top} - \mathbf{s}'', \mathbf{p}) \mid \mathbf{p} \in \mathcal{R}_1 \right\}.$$

3) For $(\mathbf{z}_1, \mathbf{p}_1) \in \mathcal{L}_1$:
   Search for $(\mathbf{z}_0, \mathbf{p}_0) \in \mathcal{L}_0$ such that $\mathbf{z}_0 + \mathbf{z}_1 = \mathbf{0}$.
   Store $(\mathbf{p}_0, \mathbf{p}_1)$ in $\mathcal{L}$.
4) For $\mathbf{e}'' \in \mathcal{L}$: Compute $\mathbf{e}' = \mathbf{s}' - \mathbf{e}''\mathbf{H}''^{\top}$
   If $\mathbf{e}' \in \{\pm 1\}^{n-k-\ell}$ then return $\pi^{-1}\big([\mathbf{e}', \mathbf{e}'']\big)$.
   Otherwise, restart from step 1.

Overall complexity analysis

Assuming that there are $M = 1 + 2^{n\left(1-(1-R)\log_2(p)\right)}$ solutions, the previous algorithm finds one with an approximate cost of

$$C_{\text{PGE}} + \frac{C_{\text{List}} + N_{\text{Test}} C_{\text{Test}}}{1 - \left(1 - 2^{2v-k-\ell}\right)^M},$$

where

$$C_{\text{PGE}} = \frac{(n-k-\ell)^2 (n-k+1) \left\lceil \log_2(p) \right\rceil^2}{\prod_{j=1}^{n-k} (1 - p^{-j})},$$

$$C_{\text{List}} = 2^{v+1} \left((v+1) + \frac{k+\ell}{2} \ell \left\lceil \log_2(p) \right\rceil \right),$$

$$C_{\text{Test}} = \frac{p}{p-2} (k+\ell) \left\lceil \log_2(p) \right\rceil,$$

$$N_{\text{Test}} = \left(1 - 2^{2v-k-\ell}\right)^M 2^{2v-\ell\log_2(p)} + \left(1 - \left(1 - 2^{2v-k-\ell}\right)^M\right) \frac{m' + (2^{2v} - m') p^{-\ell}}{(1+m')},$$

being

$$m' = M \frac{2^{2v-k-\ell}}{1 - \left(1 - 2^{2v-k-\ell}\right)^M}.$$

We now have all the ingredients to propose parameters for the restricted CVE system.

Comparison of code-based ZK-ID schemes[1]

- security parameter $\lambda = 87$ and a cheating probability 2^{-16}
- seeds and hashes of, respectively, 128 and 160 bits
- Restricted CVE parameters: $p = 29$, $n = 167$, $k = 132$

	CVE	AGS	Rest. CVE
Number of rounds	17	16	17
Public key size (bits)	512	1094	175
Total average comm. cost (kB)	3.472	3.463	2.389
Total max comm. cost (kB)	4.117	4.894	2.430

CVE (Cayrel-Véron-ElYousfiAlaoui, 2011)
AGS (Aguilar-Gaborit-Schrek, 2011)

[1]Disclaimer: We are currently recomputing the parameters due to new subset sum solvers. The final comm. cost will be slightly larger, but smaller than CVE/AGS.

Comparison of signatures from ZK-ID schemes[2]

- security level $\lambda = 128$ bits
- seeds and hashes of 256 bits
- Restricted CVE parameters: $p = 31$, $n = 256$, $k = 204$

	CVE	AGS	Rest. CVE
Number of rounds	129	128	135
Public key size (bits)	832	1574	260
Average sig. size (kB)	43.263	41.040	30.373
Max sig. size (kB)	51.261	56.992	30.373

[2]Disclaimer: We are currently recomputing the parameters due to new subset sum solvers. The final sig. size will be slightly larger, but smaller than CVE/AGS.

More comparisons

- Restricted CVE: security 128 bits, public key size 260 bits, signature size 30.373 kB
- cRVDC-125 (Bellini et al., 2019, rank-metric code-based): security 125 bits, public key 1212 bits, average signature size 22.482 kB
- Durandal (Aragon et al., 2019, rank-metric code-based): security 128 bits (with some security concerns), public key size 121961 bits, signature size ca. 4 kB
- LESS (Biass et al., 2020, code-equivalence-based): security 128 bits, public key size and signature size ca. 15 kB
- Wave (Debris-Alazard et al., 2019, code-based hash-and-sign): security 128 bits, public key size 3.2 MB, signature size ca. 1.6 kB

More comparisons

- Restricted CVE: security 128 bits, public key size 260 bits, signature size 30.373 kB
- cRVDC-125 (Bellini et al., 2019, rank-metric code-based): security 125 bits, public key 1212 bits, average signature size 22.482 kB
- Durandal (Aragon et al., 2019, rank-metric code-based): security 128 bits (with some security concerns), public key size 121961 bits, signature size ca. 4 kB
- LESS (Biass et al., 2020, code-equivalence-based): security 128 bits, public key size and signature size ca. 15 kB
- Wave (Debris-Alazard et al., 2019, code-based hash-and-sign): security 128 bits, public key size 3.2 MB, signature size ca. 1.6 kB

$\Longrightarrow$ Restricted CVE minimizes public key size!

Summary and conclusions

- We studied the syndrome decoding problem with restricted errors and showed that it is NP-complete.
- We derived GV bound to estimate minimum restricted distance of random codes.
- We derived generic decoder for special instance of restricted SDP.
- We used this in the CVE zero-knowledge identification scheme, and could reduce the used field size, for otherwise comparable parameters (length, dimension, security level).
- With Fiat-Shamir we can create digital signatures that have very small public key sizes.
- Due to smaller field size, and only ± 1-multiplication, implementation is very fast.

Summary and conclusions

- We studied the syndrome decoding problem with restricted errors and showed that it is NP-complete.
- We derived GV bound to estimate minimum restricted distance of random codes.
- We derived generic decoder for special instance of restricted SDP.
- We used this in the CVE zero-knowledge identification scheme, and could reduce the used field size, for otherwise comparable parameters (length, dimension, security level).
- With Fiat-Shamir we can create digital signatures that have very small public key sizes.
- Due to smaller field size, and only ± 1-multiplication, implementation is very fast.

Thank you for your attention!

Questions? – Comments?