A Post Quantum Digital Signature from QC-LDPC Codes

Christian Picozzi and Giovanni Tognolini

University of Trento

14 October, 2022

In a nutshell

What will we see?

The Scheme
Security
Future Directions

The Main Idea

Some PKE/KEM from NIST PQC
(LEDACrypt, BIKE, HQC)

$\downarrow$

The result

- Post Quantum code-based digital signature;
- QC-codes (for compact key-size);
- LDPC-codes (for good performance).

Setup Phase

We don't really care about it for now...

Describe each parameter as soon as it is involved in the scheme.

↓

We will denote in blue the parameters coming from the setup phase;

KeyGen Phase

Randomly generate the following elements:

- $x, y \in R := \mathbb{F}_2[X]/(X^n - 1)$, with $\mathrm{w}(x) = \mathrm{w}(y) = w$;
- $p, q \in R$, with $\mathrm{w}(p) = \mathrm{w}(q) = w_{pq}$.

Define the polynomials:

- $h := pq^{-1}$;
- $s := x + hy$.

With this notation the private and public keys are given by

$$\begin{cases} sk = (x, y, p, q) \\ pk = (h, s) \end{cases} .$$

Sufficient Conditions for q
(to be invertible)

Known in literature

If we take

- n prime;
- 2 is a primitive root modulo n;
- w_{pq} odd.

then it works.

Observation (Why?)

No deails:

- Same idea behind BIKE and LEDACrypt;
- interesting;
- not our focus now.

Signature Phase

Take as input a message m to be signed and the secret key sk. Generate:

$$r := \mathcal{H}_{w_r}(m \,||\, pk \,||\, \texttt{nonce})$$

$$t \in R \text{ such that } \mathrm{w}(t) \in I_t$$

$$\begin{cases} \alpha := qt + ry \text{ and } \mathrm{w}(\alpha) \in I \\ \beta := \alpha h + sr \text{ and } \mathrm{w}(\beta) \in I \end{cases}$$

With this notation the signature is given by

$$(\alpha, \texttt{nonce}).$$

How to contruct I ?

A genuine signer must be able to sign efficiently

$$\alpha = qt + ry$$

Observation (from HQC...)

- q, t of given weights;
- $z := qt$.

Then z is distributed as a binomial r.v. of *known* parameter $\tilde{p}$.

$$\downarrow$$

The public parameters determine the probability distribution of α.

$$\downarrow$$

We can find an interval I such that, if the scheme is executed honestly, the failure probability is negligible.

Why $\mathrm{w}(\beta) \in I$?

$$\begin{aligned}\beta &= h\alpha + sr \\ &= h(qt + ry) + (x + hy)r \\ &= hqt + hry + xr + hry \\ &= pt + xr\end{aligned}$$

$\downarrow$

Same distribution as α:

$$\begin{cases}\alpha = qt + ry \\ \beta = pt + rx\end{cases}$$

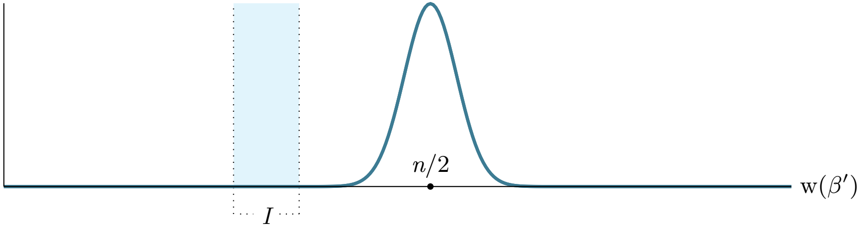
Conclusion

A genuine signer is able to generate a pair $(\alpha, \texttt{nonce})$ such that $\mathrm{w}(\alpha), \mathrm{w}(\beta) \in I$.

What's the idea behind?

$n = 17669, w_r = 74, w_{pq} = 31, w = 64, I_t = [200, 266], I = [6000, 7200]$

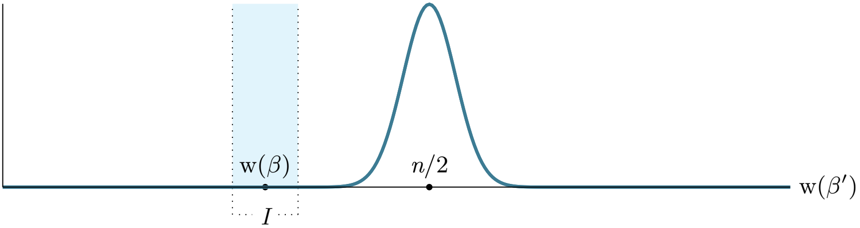
Let's just guess $(\alpha', \texttt{nonce'})$ and compute $\beta' := \alpha' h + s r'$.



What's the idea behind?

$n = 17669, w_r = 74, w_{pq} = 31, w = 64, I_t = [200, 266], I = [6000, 7200]$

Choose α honestly.



A genuine signer is the only one who can *efficiently* produce a signature.

Verification Phase

Take as input the signed message $(m, (\alpha, \texttt{nonce}))$.
Compute $r := \mathcal{H}_{w_r}(m \,||\, pk \,||\, \texttt{nonce})$ and $\beta := h \cdot \alpha + s \cdot r$ and check that

$$\begin{cases} \mathrm{w}(\alpha) \in I \\ \mathrm{w}(\beta) \in I \end{cases} .$$

If these conditions are satisfied the verifier accepts the signature, otherwise it rejects.

Security

some considerations about the hardness of:

Recovering (p, q);
Recovering (x, y);
Forging a signature.

Before doing that...

Well known:

$$\begin{array}{ccc} \mathbb{F}_2[X]/(X^n - 1) & \longleftrightarrow & (\mathbb{F}_2)^n \\ a := a_0 + a_1 X + \ldots + a_{n-1} X^{n-1} & \longleftrightarrow & (a_0, a_1, \ldots, a_{n-1})^\top =: \bar{p} \end{array}$$

We can express the product $a \cdot b$ as

$$a \cdot b = \underbrace{\begin{pmatrix} a_0 & a_{n-1} & \cdots & a_1 \\ a_1 & a_0 & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & a_0 \end{pmatrix}}_{\mathrm{circ}(a)} \cdot \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{pmatrix}$$

What does this representation allows?

We can relate our scheme to some lattice and coding problems.

Hardness of recovering (p, q)

Public key: (h, s)
where $h = pq^{-1}$

$$\downarrow$$

Observation (from NTRU)

$$(q, p) = (q_0, q_1, \ldots, q_{n-1}, p_0, p_1, \ldots, p_{n-1})$$

is very likely the shortest vector of the lattice

$$\mathcal{L}_h := \left\{ X \cdot M_h \mid X \in \mathbb{F}_2^{2n} \right\}, \text{ where}$$

$$M_h := \begin{pmatrix} I_n & \mathrm{circ}(h) \\ 0 & 2I_n \end{pmatrix}$$

Seems difficult to retrieve (p, q).

Hardness of recovering (x, y)

Public key: (h, s)
where: $s = x + hy$ and $\mathrm{w}(x), \mathrm{w}(y) = w$

Said otherwise...

$$\begin{cases} \bar{s} = \left[\mathbb{1} \mid \mathrm{circ}(h)\right] \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \\ \mathrm{w}(x), \mathrm{w}(y) = w \end{cases}$$

Observation

This problem is strictly related to the Maximum Likelihood Decoding problem (MLD), which is known to be difficult.

To be More Precise

this is a particular instance of MLD

Observation

MLD is NP-complete in the general case (random matrices), but in our case the matrix has a particular structure, given by

$$\left[\mathbb{1} \mid \mathrm{circ}(h)\right].$$

As far as we know, there are no weaknesses linked to this particular structure.

$\downarrow$

Seems difficult to retrieve (x, y).

Hardness of forging a signature

An adversary has to create a pair (α, **nonce**) such that:

$$\begin{cases} \mathrm{w}(\alpha h + sr) \in I \\ \mathrm{w}(\alpha) \in I \end{cases}$$

Observation

If we were able to forge a single message, we would be able to solve a particular instance of MLD.

Indeed, if we fix a nonce

$$\begin{cases} \mathrm{w}(\alpha h + sr) \le t_1 \\ \mathrm{w}(\alpha) \le t_2 \end{cases} \iff \begin{cases} \beta := \alpha h + sr \\ \mathrm{w}(\beta) \le t_1 \\ \mathrm{w}(\alpha) \le t_2 \end{cases} \implies \begin{cases} sr = \left(\mathbf{1} \mid\mid \mathrm{circ}(h)\right) \begin{pmatrix} \beta \\ \alpha \end{pmatrix} \\ \mathrm{w}(\beta \mid\mid \alpha) \le t_1 + t_2 \end{cases} .$$

Future Direction

Complexity reduction
Derivation from IP

Thanks

Setup Phase (finally)

Generate the parameters $(n, w, w_{pq}, w_r, I, I_t, \mathcal{H}_{w_r})$, where:

- n is a prime such that 2 is a primitive root modulo n;
- w_{pq} is an odd integer;
- w, w_{pq}, w_r are integers smaller than n;
- I and $I_t \subseteq \mathbb{N}$ are two interval;
- $\mathcal{H}_{w_r}$ is a hash function which produces digests of weight w_r.

Invertibility of q (part 1)

(Same idea as BIKE and LEDACrypt)

n positive integer not divisible by 2. Then, over $\mathbb{F}_2[X]$:

$$X^n - 1 = \prod_{d|n} \phi_d(X).$$

$$n \text{ is prime} \implies X^n - 1 = \phi_1(x)\phi_n(X)$$
$$\text{where } \phi_1(X) = X + 1 \text{ and } \phi_n(x) = 1 + X + X^2 + \ldots + X^{n-1}.$$

If $(n, 2) = 1$, then $\phi_n(X)$ factors into $\varphi(n)/d$ distinct monic irreducible polynomials in $\mathbb{F}_2[X]$, where d is the least positive integer such that $2^d \equiv 1 \pmod{n}$.

$$2 \text{ primitive root } \pmod{n} \implies d = \varphi(n)$$
$$\implies \phi_1(X), \phi_n(X) \text{ irreducible.}$$

Invertibility of q (part 2)

Observation

In our setting, an element is *not* invertible in $R = \mathbb{F}_2[X]/(X^n - 1)$ if and only if it is divisible by $X + 1$ or $1 + X + X^2 + \ldots + X^{n-1}$.

- $1 + X + X^2 + \ldots + X^n$ divides only itself;
- $1 + X$ divides only polynomials of even weight.

Any element of odd weight, different from $1 + X + X^2 + \ldots + X^{n-1}$, is invertible in R.

Conclusion:
If we take n prime, 2 primitive root (mod n), and w_{pq} odd,
we can be sure q is invertible in R.

Recovering (p, q): main idea

- $(q, p) \in \mathcal{L}_h$;
- $||(q, p)|| = \sqrt{2 w_{pq}}$.
- According to the Gaussian heuristic:

$$\sigma(\mathcal{L}_h) = \sqrt{\frac{n}{2\pi e}} \det(\mathcal{L}_h)^{1/n} = \sqrt{\frac{2n}{\pi e}} \approx 0,484 \cdot \sqrt{n}.$$

Observation

If we take $w_{pq} \approx 3 \ln(n)$, then

$$\frac{||(q, p)||}{\sigma(\mathcal{L}_h)} \approx 6,2 \cdot \frac{\ln(n)}{\sqrt{n}} \in O(\frac{1}{\sqrt{n}}).$$

$\longrightarrow$ $||(q, p)||$ is shorter than predicted by the Gaussian Heuristic.
$\longrightarrow$ (q, p) is very likely a shortest vector of the lattice.

Seems difficult to retrieve (p, q).