MAYO: Overview + Updates

Ward Beullens

IBM Research Europe - Zurich

NIST PQC seminar
24/09/2024

Part 1: Short overview of Multivariate Crypto

Multivariate Quadratic Cryptography

Cryptography based on the hardness of finding solutions to systems of multivariate quadratic equations.

Example: Solve for integers x and y:

$$\begin{aligned} x + 5x^2 + 3xy &= 4 \mod 7 \\ x^2 + 5xy + 5y^2 &= 1 \mod 7 \end{aligned}$$

Solution: $x = 6$ and $y = 0$.

For only 2 variables this is still doable, but for more variables this problem quickly becomes very difficult.

E.g. The current record mod 31 is solving a system of only 22 equations in 22 variables! ($\sim$1 core-year of computation effort)

A taxonomy of multivariate signatures

Multivariate Signatures

Pure MQ

- MQDSS/SOFIA
- MUDFiSh
- Mesquite
- **MQOM**
- **Biscuit***
- KuMQuat

Trapdoors

Oil and Vinegar-like

- **Oil & Vinegar**
- **MAYO**
- **PROV**
- **QR-UOV**
- **SNOVA**
- **TUOV**
- **VOX**

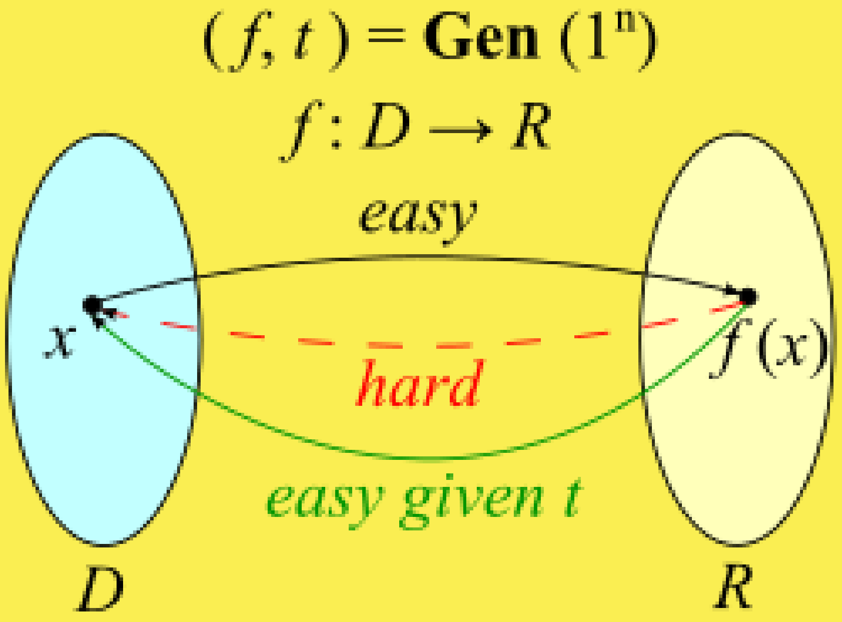
HFE-like

- C^* (1988) ✝
- HFE (1996) ✝
- FHEv- (2001) ✝
- …

Multivariate trapdoor signatures

Based on *trapdoored* multivariate maps.
I.e. quadratic functions $P(x)\colon \mathbb{F}_q^n \to \mathbb{F}_q^m$.

Maps look random (difficult to find preimages), but that have some hidden structure, that allows to compute preimages efficiently.



Full-Domain-Hash signatures (think RSA signatures)

<u>PK:</u> P

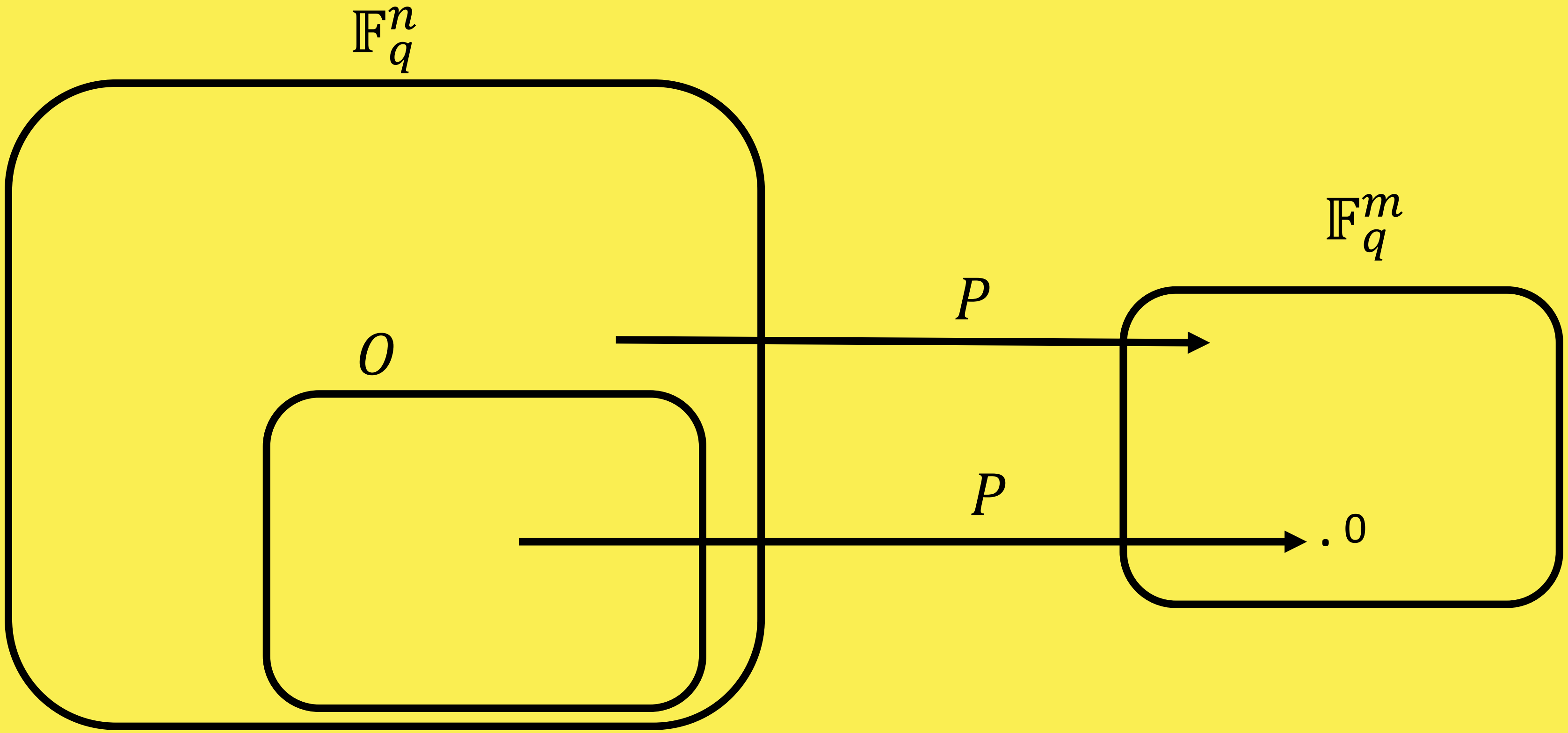
<u>SK:</u> trapdoor information

<u>Signature:</u> s such that $P(s) = H(m)$, where $H(m) \in \mathbb{F}_q^m$ is a hash digest of message m.

Oil & Vinegar Trapdoor

Public key is a quadratic map: $P = (p_1(x), \ldots, p_m(x)) \colon \mathbb{F}_q^n \to \mathbb{F}_q^m$

Trapdoor is a subspace $O \subset \mathbb{F}_q^n$ of dimension m on which P vanishes.



Definition of polar form:

Let $P\colon \mathbb{F}_q^n \to \mathbb{F}_q^m$, then we define its polar form as:

$$P'(x,y)\colon \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q^m\colon P'(x,y) = P(x+y) - P(x) - P(y)$$

This is symmetric:

$$P'(x,y) = P'(y,x)$$

And bilinear. I.e., for all $\alpha, \beta \in \mathbb{F}_q^n$

$$P'(\alpha x + \beta x', y) = \alpha P'(x,y) + \beta P'(x',y)$$

Using the trapdoor O

Given $P: \mathbb{F}_q^n \to \mathbb{F}_q^m, O \subset \mathbb{F}_q^n, y \in \mathbb{F}_q^m$.
We want to find x s.t. $P(x) = y$.

1. Pick $v \in \mathbb{F}_q^n$ uniformly at random.
2. Solve for $o \in O$ s.t. $P(v + o) = y$.

Using the trapdoor O

Given $P: \mathbb{F}_q^n \to \mathbb{F}_q^m, O \subset \mathbb{F}_q^n, y \in \mathbb{F}_q^m$.
We want to find x s.t. $P(x) = y$.

1. Pick $v \in \mathbb{F}_q^n$ uniformly at random.
2. Solve for $o \in O$ s.t. $P(v + o) = y$.

$$P(v + o) = P(v) + \cancel{P(o)} + P'(o, v) = y$$

Is a linear system of m equations in m variables.

Using the trapdoor O

Given $P: \mathbb{F}_q^n \to \mathbb{F}_q^m, O \subset \mathbb{F}_q^n, y \in \mathbb{F}_q^m$.
We want to find x s.t. $P(x) = y$.

1. Pick $v \in \mathbb{F}_q^n$ uniformly at random.
2. Solve for $o \in O$ s.t. $P(v + o) = y$.

$$P(v + o) = P(v) + \cancel{P(o)} + P'(o, v) = y$$

Is a linear system of m equations in m variables.

If the system does not have solutions, try again with new v

Parameters (NIST SL 1)

2 constraints:
- Finding oil space O should be hard
- It should be hard to solve $P(x) = y$ without O

Attacks:
Exponential in $n - 2m$
Exponential in m

	UOV-Ip	UOV-Is
# Variables n	112	160
# Equations m	44	64
Finite Field	GF(256)	GF(16)
Pk size	44 KB	67 KB
Signature size	128 B	96 B

Pros and Cons of *Oil and Vinegar*

Advantages:

- Old and well studied (1997)
- Small signatures (96B)
- Fast (100 Kcyc sign,
  150 Kcyc verify)

Limitations:

- Somewhat large public keys
  (44KB)

Classification of variants of Oil and Vinegar

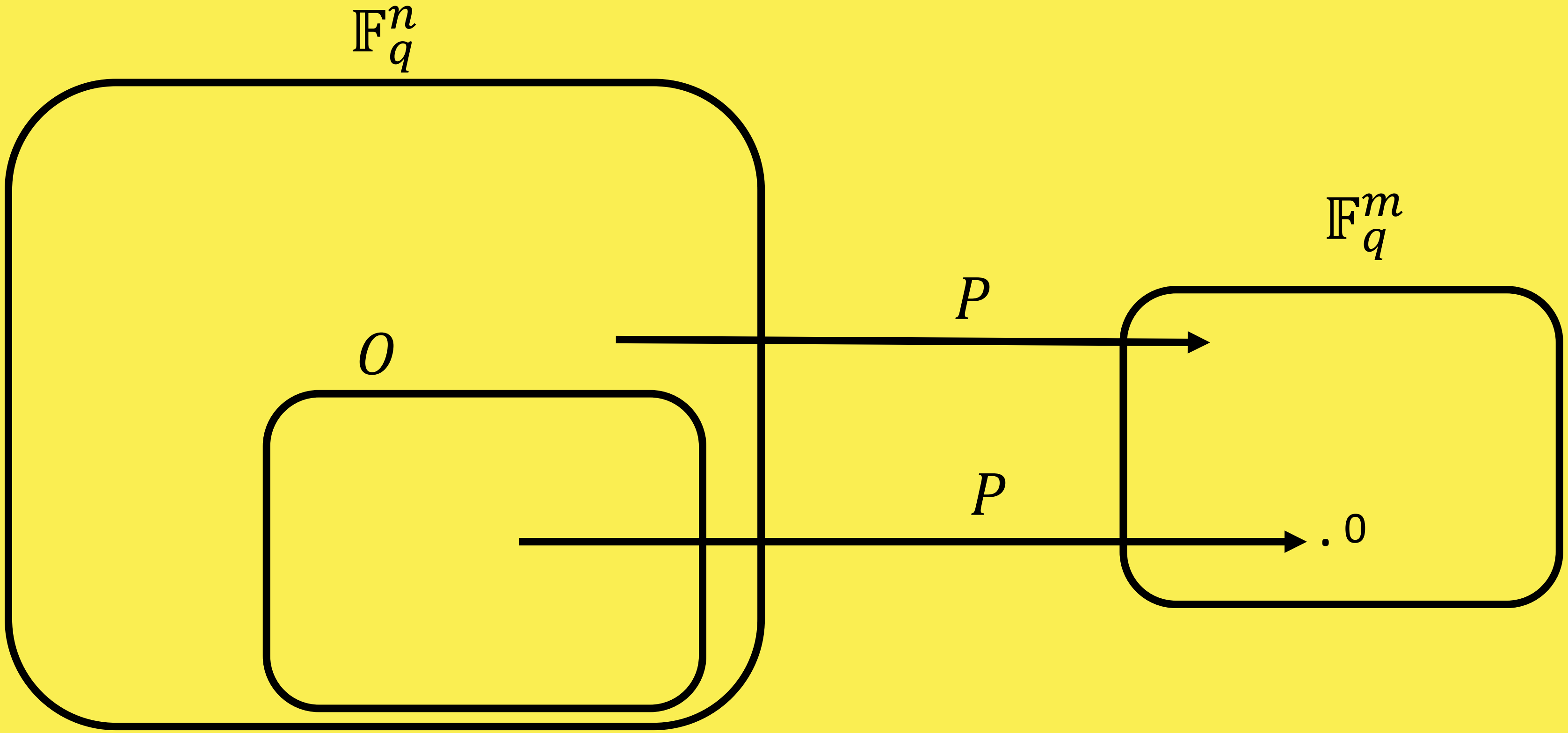
Add “whipping” structure

Add extension field structure

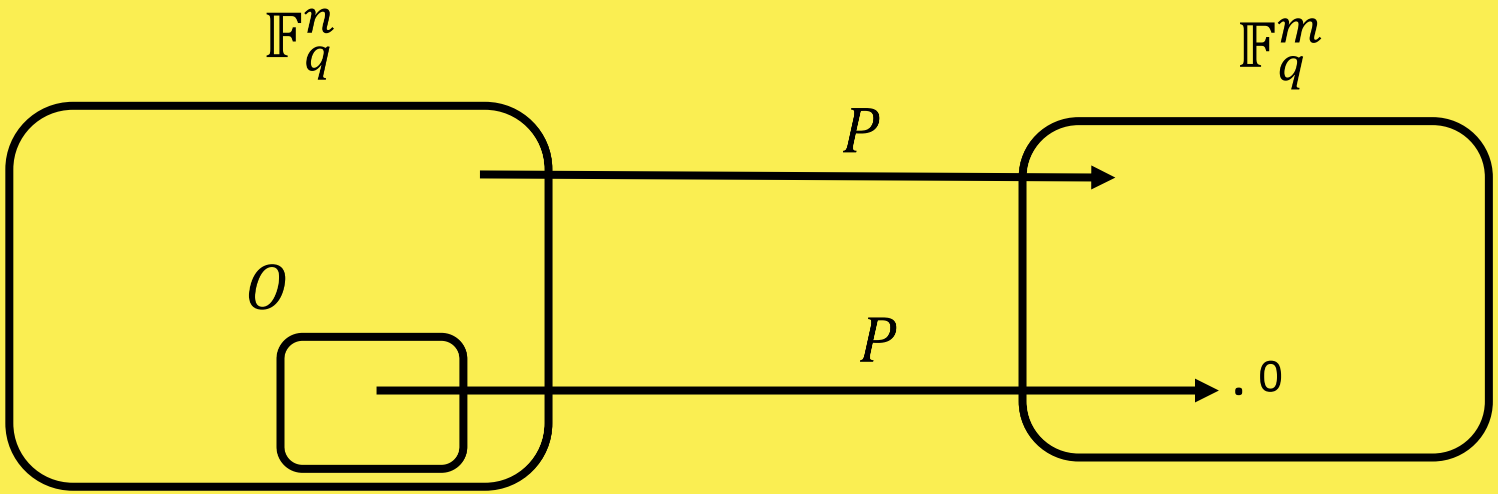
UOV PROV TUOV FOX	MAYO
QR-UOV	SNOVA

Part 2: MAYO in a nutshell

Oil and Vinegar Public Key



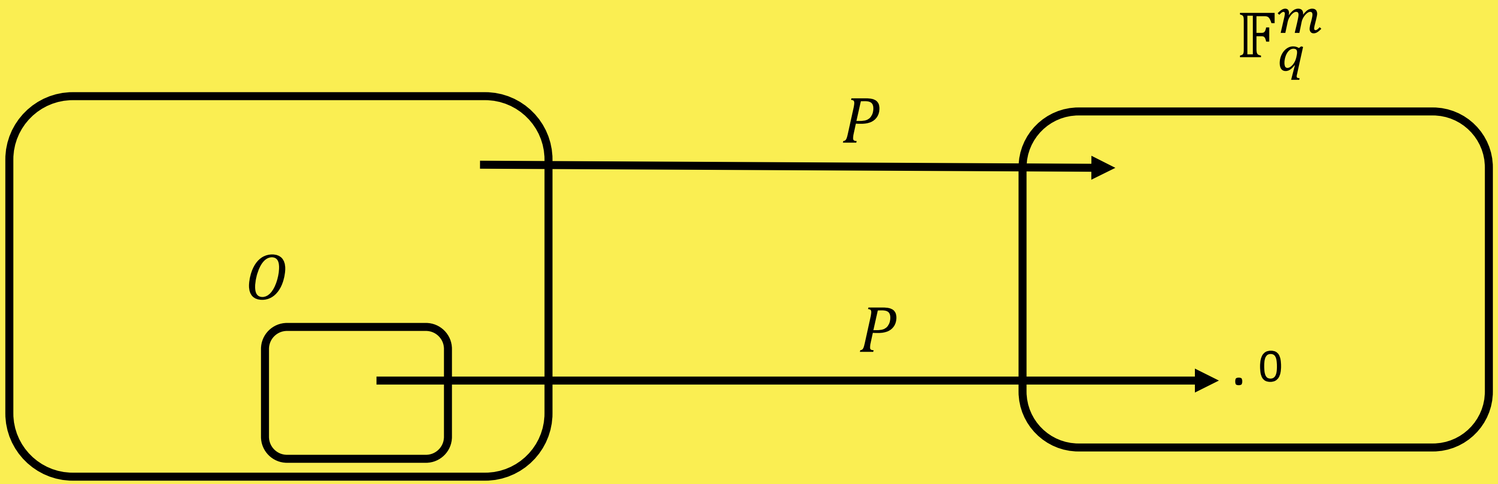
MAYO Public Key



Making O smaller has 2 benefits:

- We can use smaller n (key recovery attack exponential in $n - 2o$)
- Public key becomes smaller: $O(o^2 m)$ instead of $O(m^3)$

MAYO Public Key



But, if $\dim(O) < m$ the signing algorithm fails:

$P(v+o) = P(v) + P'(o,v) = t \in \mathbb{F}_q^m$: m equations, $\dim(O)$ variables.

A little oil can go a long way

Whip map $P: \mathbb{F}_q^n \to \mathbb{F}_q^m$ with small space O up to a larger map $P^\star: \mathbb{F}_q^{kn} \to \mathbb{F}_q^m$, that vanishes on a larger oil space O^k.

$P: \mathbb{F}_q^n \to \mathbb{F}_q^m$

"Whip up" $\times\, k$

$P^\star: \mathbb{F}_q^{kn} \to \mathbb{F}_q^m$

Whipping Oil-and-Vinegar: Attempt 1

Let $P^\star(x_1, \dots, x_k) = P(x_1) + P(x_2) + \cdots + P(x_k)$.

Then $P^\star\colon\ \mathbb{F}_q^{kn} \to \mathbb{F}_q^m$ vanishes on a large oil space

$$O^k = \{\, (\mathrm{o}_1, \dots, o_k) \mid \mathrm{o}_1, \dots, o_k \in O\}$$

So, if $\dim\left(O^k\right) = ko \geq m$, then we can sample preimages for $P^\star$.

Bonus slide: Why $P^\star(x_1, x_2) = P(x_1) + P(x_2)$ is not preimage resistant.

We want to solve $P(x_1) + P(x_2) = t$ for $x_1, x_2 \in \mathbb{F}_q^n$,
given arbitrary $t \in \mathbb{F}_q^m$ (e.g., $t = H(M)$)

For simplicity, assume $-1 = \alpha^2$ is a square, and P is homogeneous.

Set $x_2 := \alpha x_1 + r$ for random $r \in \mathbb{F}_q^m$

$$P(x_1) + P(\alpha x_1 + r) = t$$
$$P(x_1) + P(\alpha x_1) + P(r) + P'(\alpha x_1, r) = t$$
$$P(r) + P'(\alpha x_1, r) = t$$

Is just a system of linear equations.

Whipping Oil-and-Vinegar: Attempt 2

Choose matrices $E_{i,j}$ for all $0 \leq i \leq j \leq k$ and set

$$P^{\star}(x_1, \ldots, x_k) = \sum_{i} E_{ii} P(x_i) + \sum_{i<j} E_{ij} P'(x_i, x_j)$$

New hardness assumption:

Systems P^* of this form are preimage resistant when P is uniformly random.

Security Analysis

Assume that:

1) Oil-and-Vinegar maps P are indistinguishable from random MQ maps. ☺
2) Whipping up a random map P, results in a (multi-target) preimage resistant MQ map $P^\star$.

Then the MAYO signature scheme is EUF-CMA secure.
(for appropriately chosen parameters)

In particular, we proved that signatures do not leak information about the secret key.

MAYO parameters

	Oil & Vinegar GF(16)	Oil & Vinegar GF(256)	MAYO 1 $o = 8$	MAYO2 $o = 18$
# Variables	160	112	66 x 9	66 x 16
# Equations	64	44	64	69
Finite Field	GF(16)	GF(256)	GF(16)	GF(16)
Pk size	67 KB	44 KB	1.1 KB	5.4 KB
Signature size	96 B	128 B	321 B	180 B

Size of O gives a trade-off between signature size and pk size.

Advantages:

- Short signatures (180B)
- Short keys (1.1KB)
- Fast ($111\mu s$ signing, $30\mu s$ verify)

Limitations:

- New hardness assumption (2021)

Update 1:

New representation of public key for faster implementations

Bitsliced vs. Nibblesliced representations

How to represent matrices over $GF(16)$? Representation is irrelevant for security, but important for interoperability and efficient implementation.

$a_0 + a_1x + a_2x^2 + a_3x^3$	$d_0 + d_1x + d_2x^2 + d_3x^3$	$g_0 + g_1x + g_2x^2 + g_3x^3$
$b_0 + b_1x + b_2x^2 + b_3x^3$	$e_0 + e_1x + e_2x^2 + e_3x^3$	$h_0 + h_1x + h_2x^2 + h_3x^3$
$c_0 + c_1x + c_2x^2 + c_3x^3$	$f_0 + f_1x + f_2x^2 + f_3x^3$	$i_0 + i_1x + i_2x^2 + i_3x^3$

(Column major) bitsliced representation:

$$a_0b_0c_0 \;\; a_1b_1c_1 \;\; a_2b_2c_2 \;\; a_3b_3c_3 \;\; \ldots$$

Good for bitsliced arithmetic on embedded platforms.

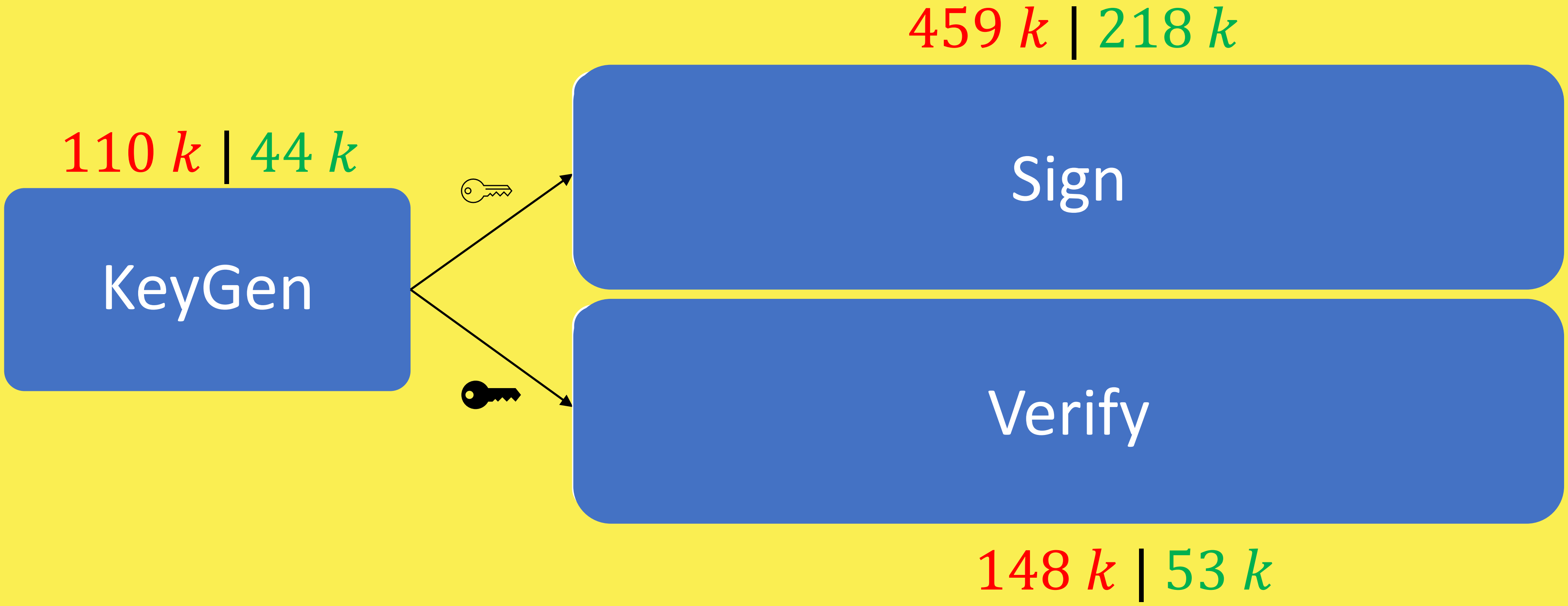
(Column major) nibblesliced representation:

$$a_0a_1a_2a_3 \;\; b_0b_1b_2b_3 \;\; c_0c_1c_2c_3 \;\; \ldots$$

Good for AVX2 shuffle-based arithmetic on "big" CPUs **and table-lookup-based multiplication on embedded platforms.**

Ice Lake performance MAYO 1

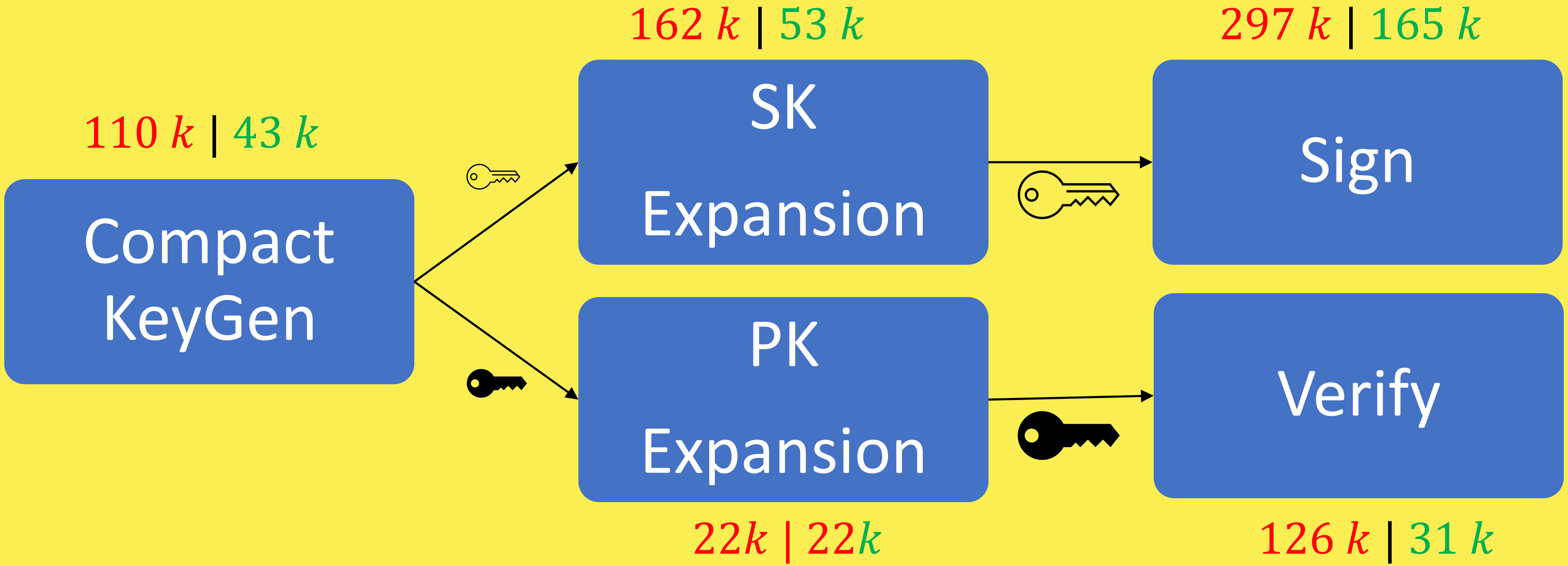
AVX2 + AESNI Bitsliced | Nibble-sliced implementation



Dilithium2: KeyGen 81 k, Sign 219 k, Verify 79 k

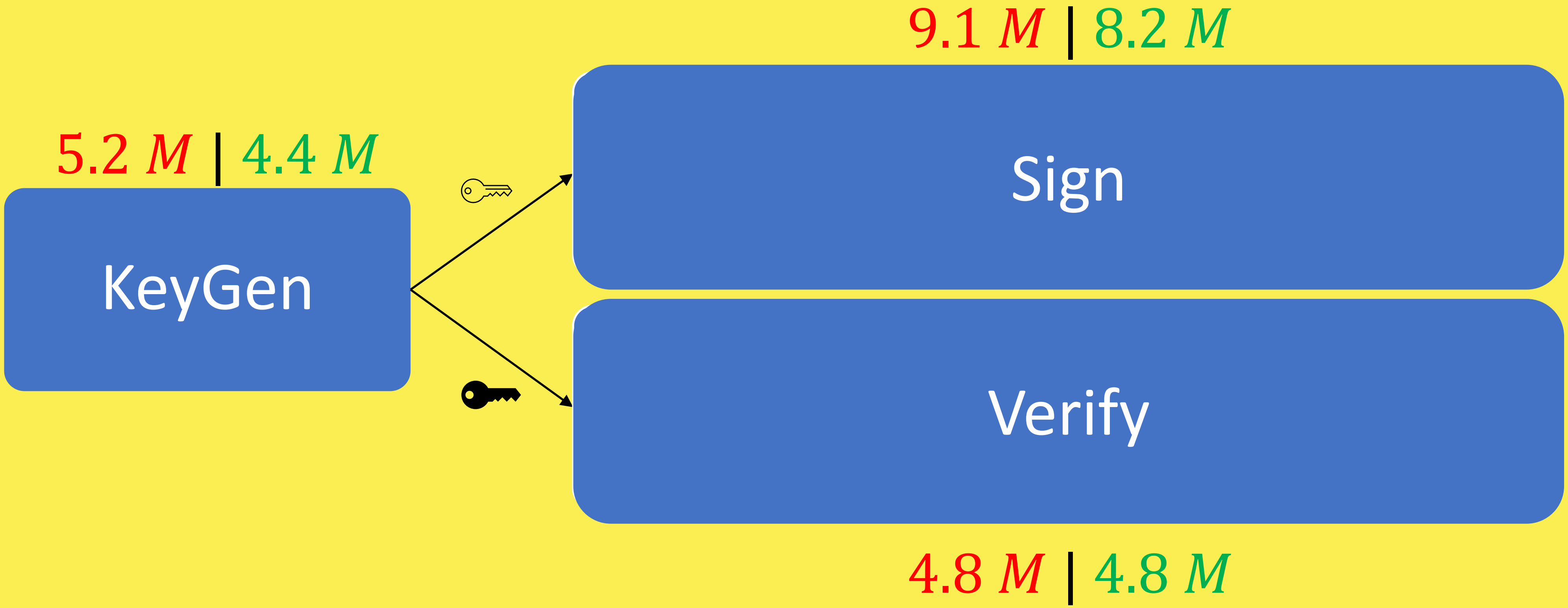
Ice Lake performance MAYO 1

AVX2 + AESNI Bitsliced | Nibble-sliced implementation



Cortex-M4 performance MAYO 1

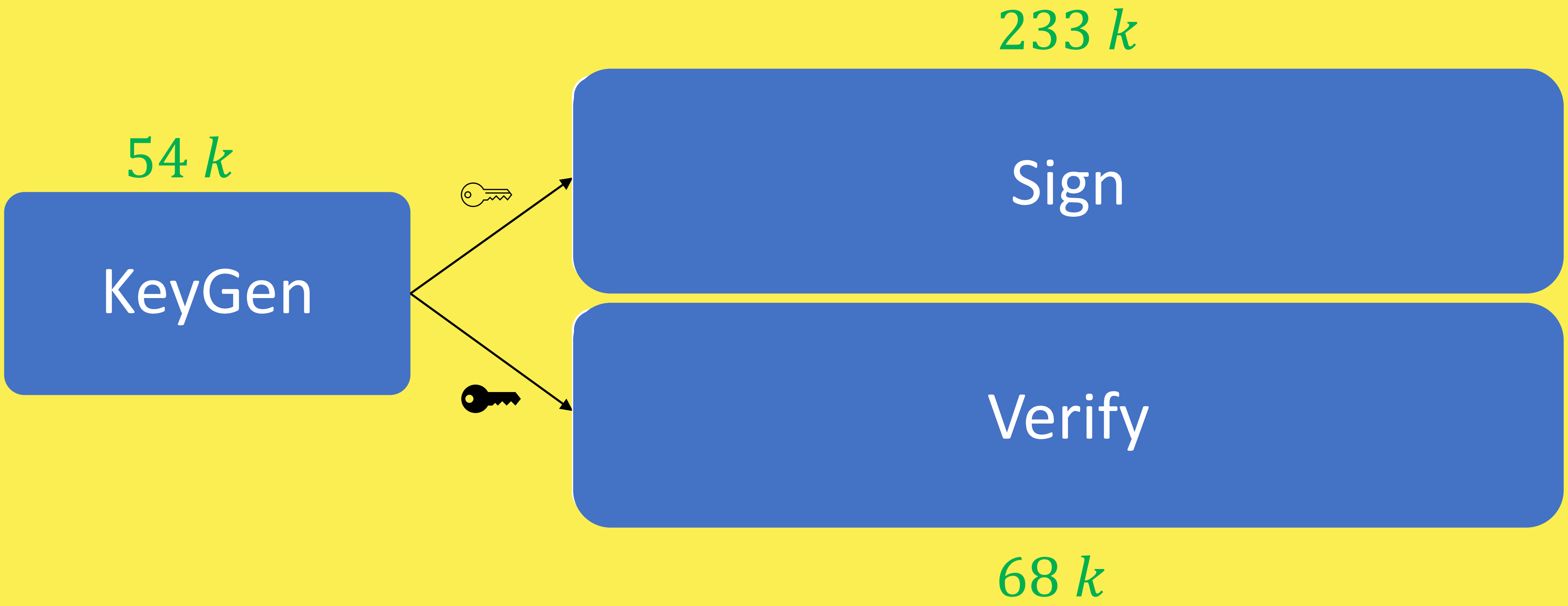
ST NUCLEO-L4R5ZI @ 20 MHz Bitsliced | Nibble-sliced



Dilithium2: KeyGen 1.6 M, Sign 4.0 M, Verify 1.6 M

Work in progress: ARM Neon

Nibble-sliced



Dilithium2: KeyGen 71 k, Sign 224 k, Verify 69 k

Update 2:

New parameters

Why new parameters?

Improved method for solving underdetermined systems by Hashimoto (2023) reduced security margin by between 14 bits (MAYO1) and 2 bits (MAYO2)	Add more security margin against generic system solving attacks.
Parameters with $o > n - m$ mean that the $P(x) = 0$ variety has a larger dimension than generic varieties. Related to new Minrank attack of Furue and Ikematsu (2023)	Pick parameters with $o \leq n - m$. **Side effect:** we get much more security against known key-recovery attacks.
Restart probability 2^{-36} makes it hard to cover all corner cases of implementation with KATs.	Increase restart probability **Side effect:** Small reduction in key sizes

New (tentative) MAYO parameters

	MAYO1	MAYO2	MAYO3	MAYO5
Security Level	1	1	3	5
(n,m,o,k)	(86,78,8,10)	(81,64,17,4)	(117, 107, 10, 11)	(153,141,12,12)
Signature size	454 B	186 B	676 B	958 B
Pk size	1420 B	4912 B	2959 B	5515 B

New (tentative) MAYO parameters

	MAYO1	MAYO2	MAYO3	MAYO5
Security Level	1	1	3	5
(n,m,o,k)	(86,78,8,10)	(81,64,17,4)	(117, 107, 10, 11)	(153,141,12,12)
Signature size	454 B	186 B	676 B	958 B
Pk size	1420 B	4912 B	2959 B	5515 B
Restart Prob	2^{-12}	2^{-20}	2^{-16}	2^{-16}
Forgery Attacks	156	155*	222	295
Key Recovery	197	167	260	332

*ignoring an attack on hash function collision with bit cost $\sim 2^{143}$

Work in progress

- Implementing new parameter sets
- Low-memory implementation
- Formally verified security proof
- Formally verified implementation
- Thresholdized signing for MAYO
- ...