



#### A High-Performance Hardware Implementation of the LESS Digital Signature Scheme

<u>Luke Beckwith</u><sup>1,2</sup>, <u>Robert Wallace</u><sup>1</sup>, Kamyar Mohajerani<sup>1</sup>, and Kris Gaj<sup>1</sup> (1) Cryptographic Engineering Research Group (CERG) at George Mason University (2) PQSecure Technologies

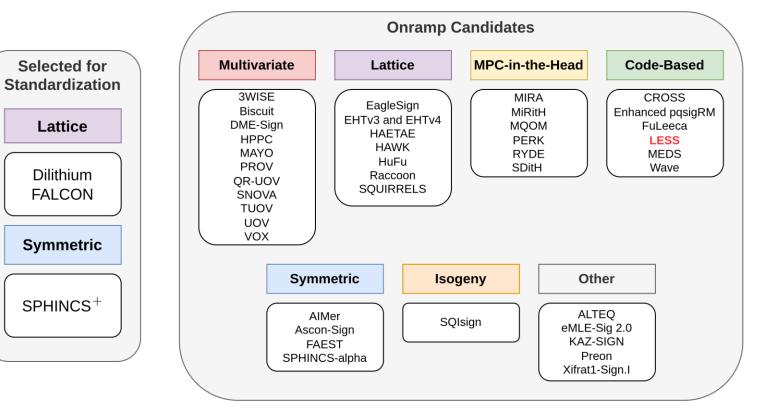
## Outline

- Brief overview of PQC status
- Introduction to LESS
  - Mathematical background
  - Algorithm details
  - Parameters
- Hardware architecture
  - Top-level structure
  - Details of RREF implementation
- Results and comparison

## **PQC Signatures**

#### Winners:

- 3 algorithms
- 2 types of cryptography New Candidates:
- 40 algorithms
- 7+ types



#### LESS

- LESS (Linear Equivalence Signature Scheme):
  - Code-based algorithm based on the difficulty of the linear equivalence problem
  - Constructed using Fiat-Shamir
  - Main elements are large matrices with elements in  $F_q$
- Core Operation: RREF( RREF(Generator) x Monomial Matrix )

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 5 \\ 0 & 0 & 1 & 0 & 6 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 \\ 0 & 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 & 1 & 0 \\ 0 & 5 & 2 & 0 & 0 \\ 0 & 4 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$RREF(Generator) \qquad Monomial_{Matrix}$$

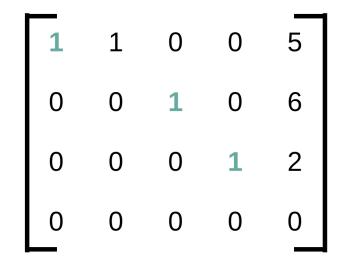
## Background - RREF

#### **Reduced Row Echelon Form (RREF):**

A matrix is said to be in RREF if:

- 1. Rows with only zeros are at the bottom of the matrix
- 2. The leftmost non-zero (leading) entry of each row is to the right of the leading entry of all rows above it
- 3. All leading entries are 1
- 4. Each column containing a leading 1 has zeros in all other entries

The leading entries are also referred to as "pivots" and the corresponding columns as "pivot columns"



Example of a matrix in RREF. Pivots are in green.

#### Background – Monomial Matrix

#### **Monomial Matrix:**

A monomial matrix is a combination of a scalar matrix and a permutation matrix.

Each column and row have only one nonzero entry which is in  $F_q^*$ . The set of monomial matrices is referred to as  $M_n$ .

0	0	0	1	0
3	0	0	0	0
0	0	2	0	0
0	0	0	0	6
0	2	0	0	0

#### Background – Generator Matrix and LEP

#### **Generator Matrix:**

A generator matrix is a matrix whose rows form the basis for a linear code. So, for generator G of code C, the codeword c of message m is calculated by:

c = mG

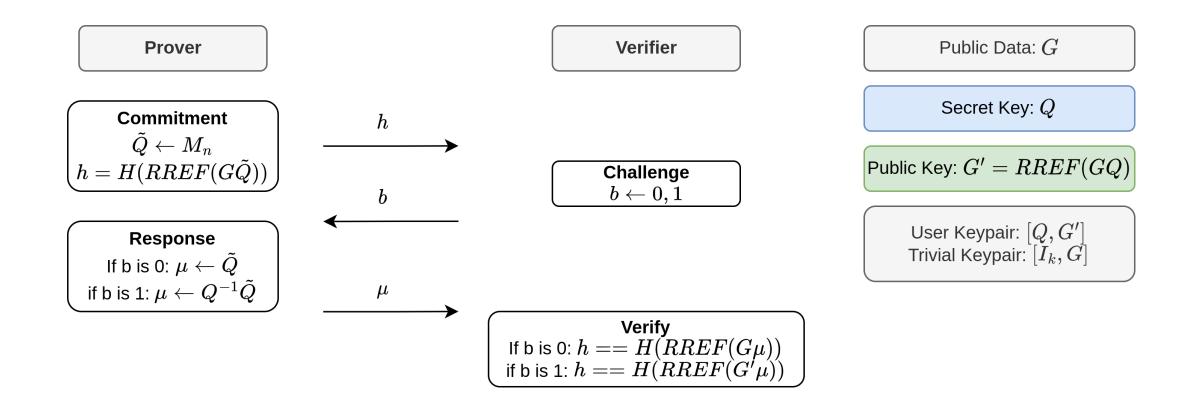
Two generator matrices are said to be *linearly equivalent* if there exist a monomial matrix Q and an invertible matrix S, such that

$$G' = SGQ$$

1	0	0	5	0
0	1	0	0	0
0	0	1	0	5
0	0	0	0	0

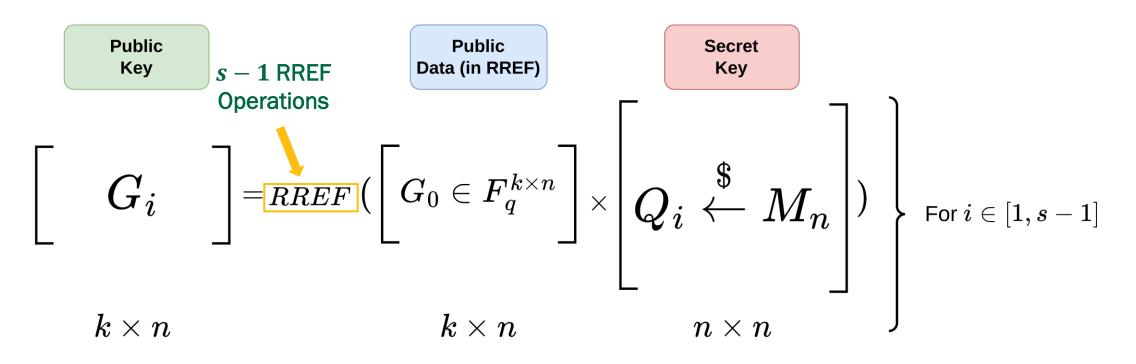
#### Linear Equivalence Problem: Given G' and G, it is difficult to find Q

## **LEP Sigma Identification Protocol**



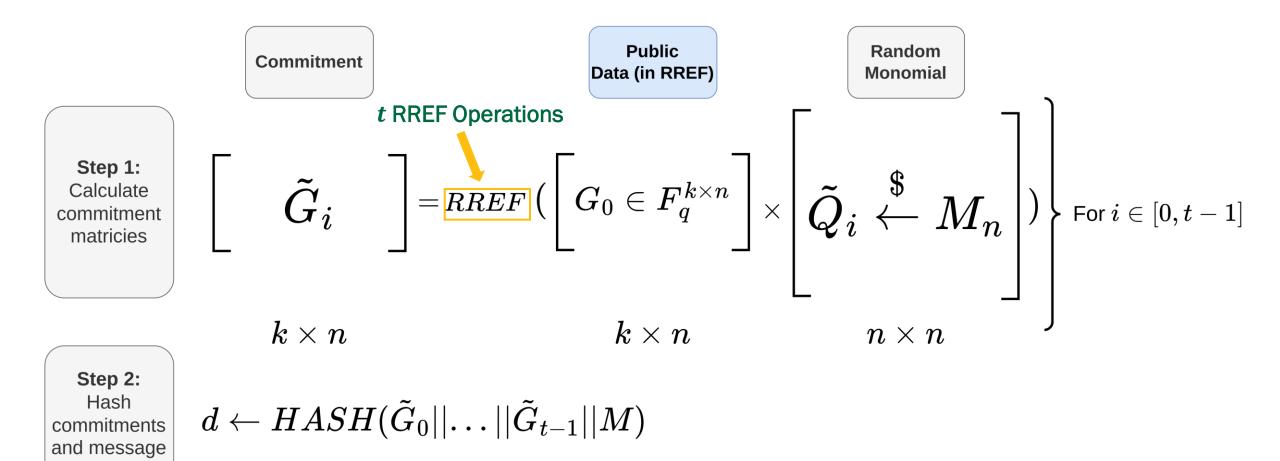
Difficulty can be increased by performing multiple rounds or by using multiple keypairs

## LESS Key Generation (Simplified)

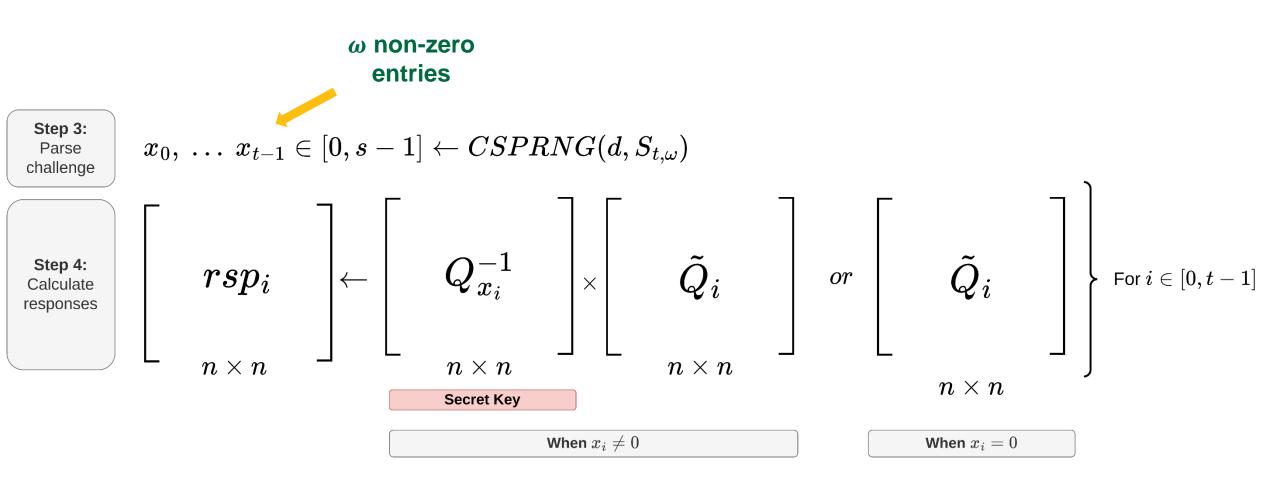


- Each keypair is an instance of LEP
- Multiple keypairs can be used to lower number of rounds needed
  - First keypair is trivial keypair  $(I_k, G_0)$
  - s-1 additional keypairs generated

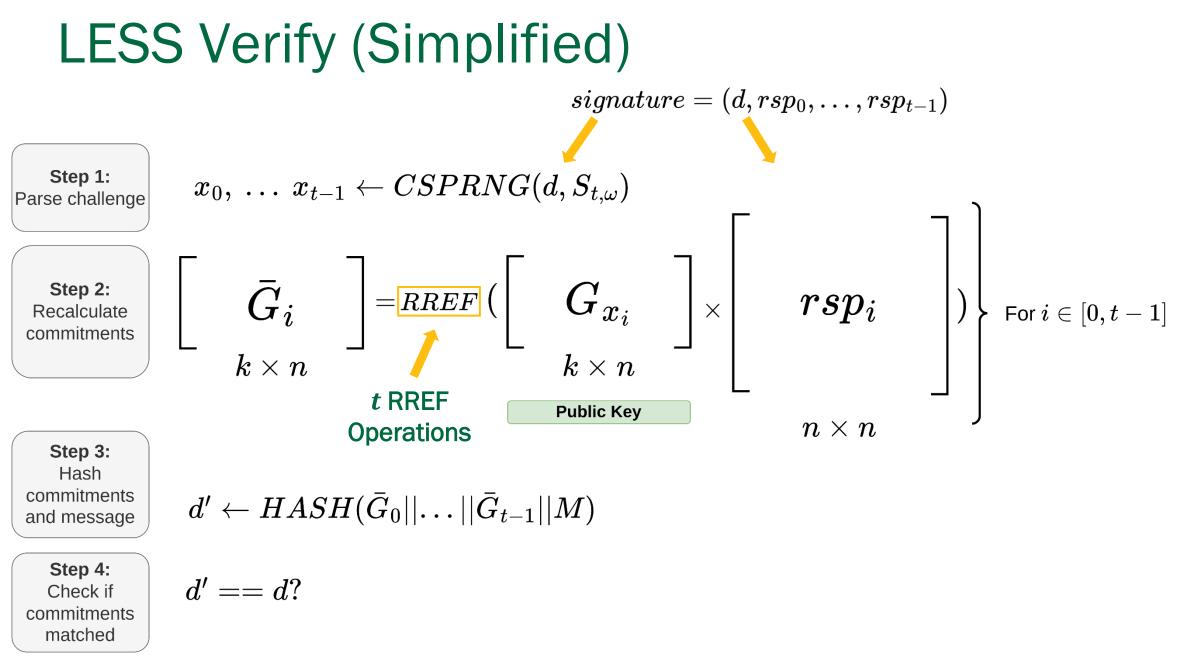
#### LESS Sign Part 1 (Simplified)



#### LESS Sign Part 2 (Simplified)

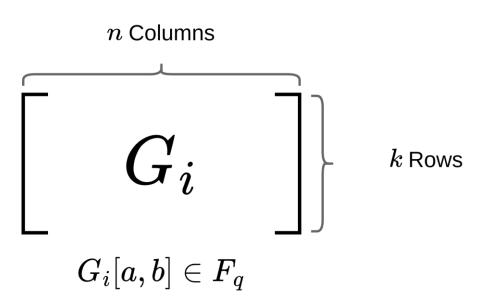


$$signature = (d, rsp_0, \dots, rsp_{t-1})$$



#### **LESS** Parameters

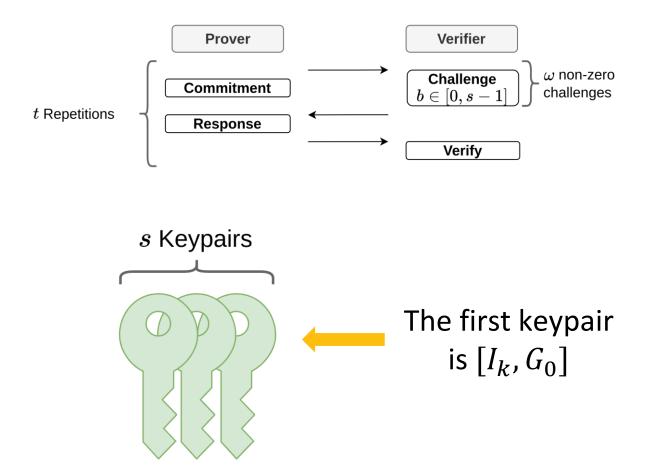
NIST Security	Parameter	Code Parameters				
Level	Set	n	k	q		
1	LESS-1b					
	LESS-1i	252	126	127		
	LESS-1s					
3	LESS-3b	400	200	127		
	LESS-3s	400	200	127		
5	LESS-5b	548	274	127		
	LESS-5s	540	274	127		



 $s \rightarrow$  "Short" - minimize Sig size  $b \rightarrow$  "Balanced" - minimize PK + Sig size  $i \rightarrow$  "Intermediate" - in between b and s

#### **LESS** Parameters

NIST Security	Paramete r	Protocol Parameters			
Level	Set	t	ω	S	
1	LESS-1b	247	30	2	
-	LESS-1i	244	20	4	
	LESS-1s	198	17	8	
3	LESS-3b	759	33	2	
	LESS-3s	895	26	3	
5	LESS-5b	1352	40	2	
	LESS-5s	907	37	3	



#### **LESS** Parameters

#### In LESS:

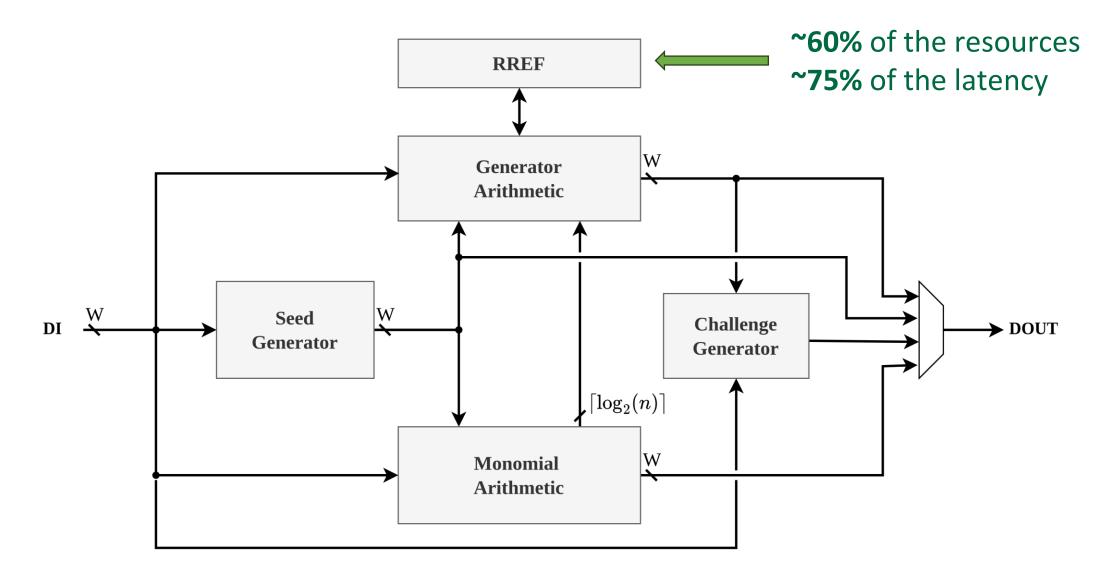
- Signature size  $\nearrow$  when  $k \nearrow$ ,  $t \nearrow$  and  $\omega \nearrow$
- Public key size  $\nearrow$  when  $n \nearrow$ ,  $k \nearrow$ , and  $s \nearrow$
- Secret key size is independent of parameters

#### In this architecture:

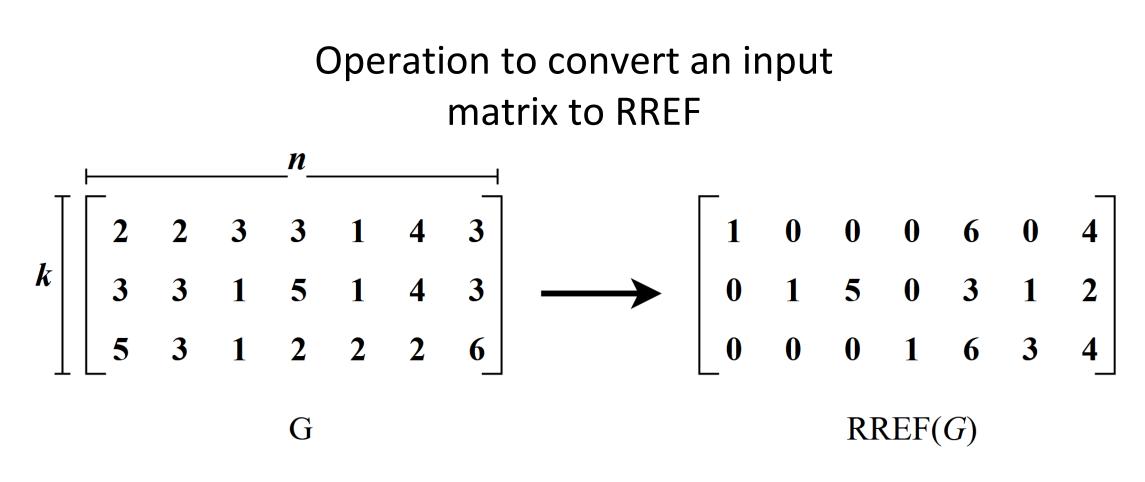
- Area ≯ to n ≯
- Keygen Cycle Latency  $\nearrow$  when  $k \nearrow$  and  $s \nearrow$
- Sign/Verify Cycle Latency  $\nearrow$  when  $k \nearrow$  and  $t \nearrow$

Parameter Meaning
Code Parameters
n: Generator columns
k: Generator rows
<i>q</i> : Generator element modulus
Protocol Parameters
t: Protocol repetitions
$\omega$ : Non-zero challenges
s: Number of keypairs

#### **Top-Level Architecture**



RREF

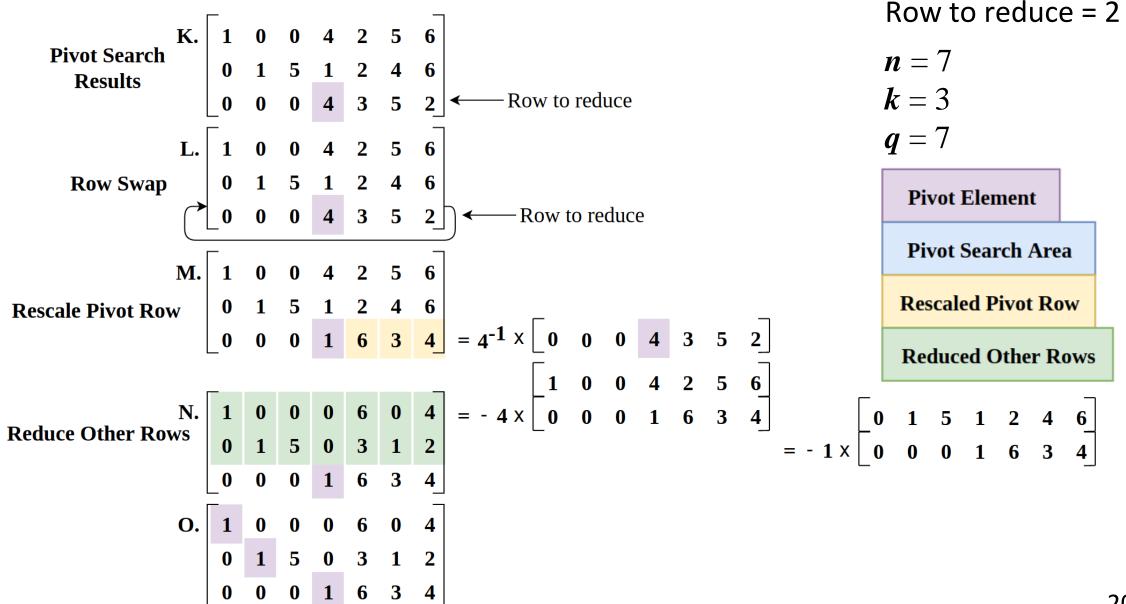


All arithmetic operations performed modulo q = 7

RREF – E	X	al	m	p	le	ļ		
				·			Row to reduce = 0	
А.	2	2	3 1	3 5	1	4	$ \begin{array}{c} 3 \\ 3 \\ 3 \\ 6 \end{array} $ Row to reduce $n = 7$ k = 3	
Direct Cooreh	5	3	1	2	2	2		
Pivot Search B.	2	2	3	3	1	4	$\mathbf{a}$ Row to reduce $\mathbf{q} = 7$	
	3				1		Fivot Element	
C	_	3						
Row Swap	2	2	3 1	3 5	1 1	4 4	3 Rescaled Pivot Row	
Row Swap	_5	3	1	2	2	2	6	
	_						$5 = 2^{-1} \times \begin{bmatrix} 2 & 2 & 3 & 3 & 1 & 4 & 3 \end{bmatrix}$ Reduced Other Rows	
D. Rescale Pivot Row	л З				4			
		3						
E. Reduce Other Rows	1	1	5	5	4	2	$\begin{bmatrix} 5 \\ 5 \\ 2 \\ 2 \end{bmatrix} = -3 \times \begin{bmatrix} 3 & 3 & 1 & 5 & 1 & 4 & 3 \\ 1 & 1 & 5 & 5 & 4 & 2 & 5 \end{bmatrix} \begin{bmatrix} 5 & 3 & 1 & 2 & 2 & 2 & 6 \\ 1 & 1 & 5 & 5 & 4 & 2 & 5 \end{bmatrix} = -5 \times \begin{bmatrix} 1 & 1 & 5 & 5 & 4 & 2 & 5 \\ 1 & 1 & 5 & 5 & 4 & 2 & 5 \end{bmatrix}$	
+	0	0	0	4	3	5	$2 = -3 \times \begin{bmatrix} 1 & 1 & 5 & 5 & 4 & 2 & 5 \end{bmatrix} \begin{bmatrix} 5 & 3 & 1 & 2 & 2 & 6 \end{bmatrix}$	
<b>Pivot Search</b>	_0	5	4	5	3	6	$2 = -5 \times [1 \ 1 \ 5 \ 5 \ 4 \ 2 \ 5]$	0

RREF – E	X	a	m	p	le	Ś	
	<b>1</b>	1	5	5	4	2	<sup>5</sup> Row to reduce = 1
Pivot Search Results	0	0	0	4	3	5	$ \begin{array}{c c}     2 \\     2 \\     2 \\   \end{array} \\          $
Kesuits	_0	5	4	5	3	6	k = 3
G.	1	1	5	5	4	2	$\overline{\mathbf{q}} = 7$
	0	0	0	4	3	5	2 2 Row to reduce Pivot Element
Row Swap	_0	5	4	5	3	6	2 Pivot Element
H.	1	1	5	5	4	2	5 Pivot Search Area
	0	5	4	5	3	6	2 ← Row to reduce Rescaled Pivot Row
	_0	0	0	4	3	5	2 Reduced Other Rows
I.	1	1	5	5	4	2	5
<b>Rescale Pivot Row</b>	0	1	5	1	2	4	$6 = 5^{-1} \times \begin{bmatrix} 0 & 5 & 4 & 5 & 3 & 6 & 2 \end{bmatrix}$
	0_	0	0	4	3	5	$2 \qquad \begin{bmatrix} 1 & 1 & 5 & 5 & 4 & 2 & 5 \end{bmatrix}$
J. Deduce Other Devis	1	0	0	4	2	5	5 6 = $5^{-1} \times \begin{bmatrix} 0 & 5 & 4 & 5 & 3 & 6 & 2 \end{bmatrix}$ 2 $\begin{bmatrix} 1 & 1 & 5 & 5 & 4 & 2 & 5 \\ 6 & = -1 \times \begin{bmatrix} 0 & 1 & 5 & 1 & 2 & 4 & 6 \end{bmatrix}$
Reduce Other Rows +	0	1	5	1	2	4	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
<b>Pivot Search</b>	_0	0	_0_	4	3	5	$= -0 \times \begin{bmatrix} 0 & 1 & 5 & 1 & 2 & 4 & 6 \end{bmatrix}$

#### RREF – Example



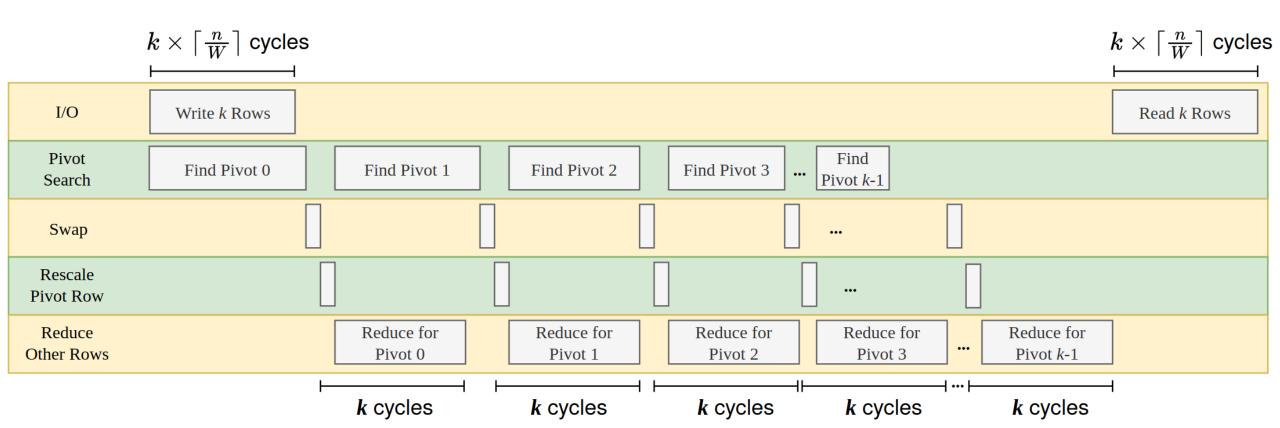
### RREF – Algorithm

#### • Four major operations:

- 1. Pivot Search
- 2. Row Swap
- 3. Rescale Pivot Row
- 4. Reduce Other Rows
- Opportunities for parallelization:
  - 1. Arithmetic performed on entire row
  - 2. Pivot Search while Reduce Other Rows
  - 3. Row operations in *Reduce Other Rows* are independent of each other
- Constant-time implementation

	Input: Matrix $G \in \mathbb{Z}_q^{k \times n}$ Output: Matrix $G \in \mathbb{Z}_q^{k \times n}$								
	Pivot Search	$ \begin{array}{c} 1 & 10 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} $	for $\operatorname{rtr} \in [0, k-1]$ do $\operatorname{vld\_piv} \leftarrow 0$ for $\operatorname{col} \in [\operatorname{rtr}, n-1]$ do $ $ for $\operatorname{row} \in [\operatorname{rtr}, k-1]$ do $ $ if $(G[\operatorname{row}][\operatorname{col}] > 0)$ and $(\operatorname{vld\_piv} = 0)$ then						
		6 7 8 9	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$						
W	Rescale Pivot Row	10 11 12	$m \leftarrow G[\operatorname{rtr}][\operatorname{piv}\_\operatorname{col}]^{-1} \mod q$ <b>for</b> col $\in [0, n - 1]$ <b>do</b> $\mid G[\operatorname{rtr}][\operatorname{col}] \leftarrow m \cdot G[\operatorname{rtr}][\operatorname{col}] \mod q$						
vv	Reduce Other Rows	13 14 15 16 17 18	$ \begin{array}{c c} \mathbf{for} \ \mathrm{row} \in [0, k-1] \ \mathbf{do} \\ \mathbf{if} \ \mathrm{row} \neq \mathrm{rtr} \ \mathbf{then} \\ & m \leftarrow G[\mathrm{row}][\mathrm{piv\_col}] \\ \mathbf{for} \ \mathrm{col} \in [\mathrm{piv\_col}, n-1] \ \mathbf{do} \\ & for \ \mathrm{col} \in [\mathrm{piv\_col}, n-1] \ \mathbf{do} \\ & for \ \mathrm{col} \in [\mathrm{row}][\mathrm{piv\_col}] \cdot G[\mathrm{rtr}][\mathrm{col}] \ \mathrm{mod} \ q \\ & G[\mathrm{row}][\mathrm{col}] \leftarrow G[\mathrm{row}][\mathrm{col}] - \mathrm{tmp} \ \mathrm{mod} \ q \end{array} $						

#### **RREF – Operational Flow**



Latency (clock cycles) =  $k^2 + 3k + 58$ (including implemented pipeline stages, not including I/O)

## **RREF – Column Memory**

•

•

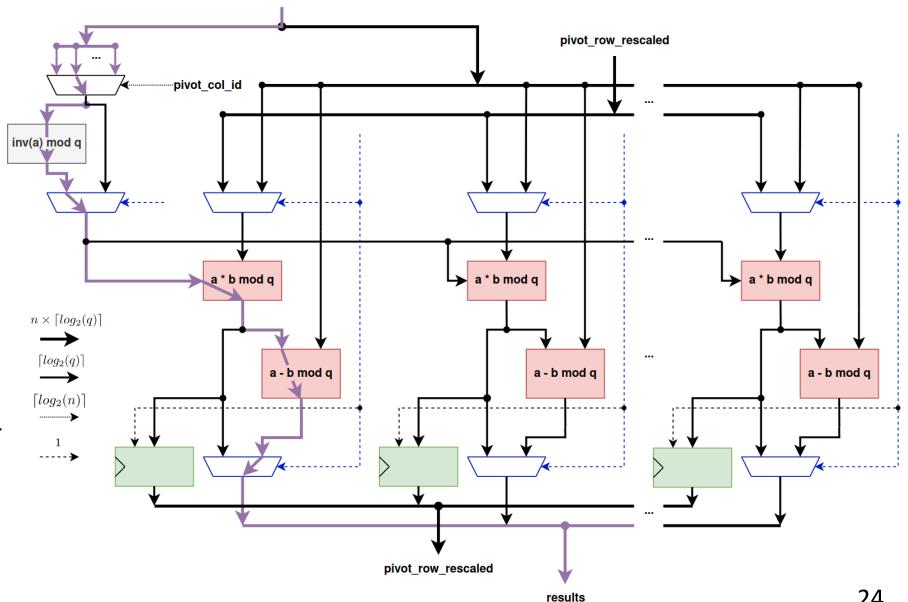
•

ports

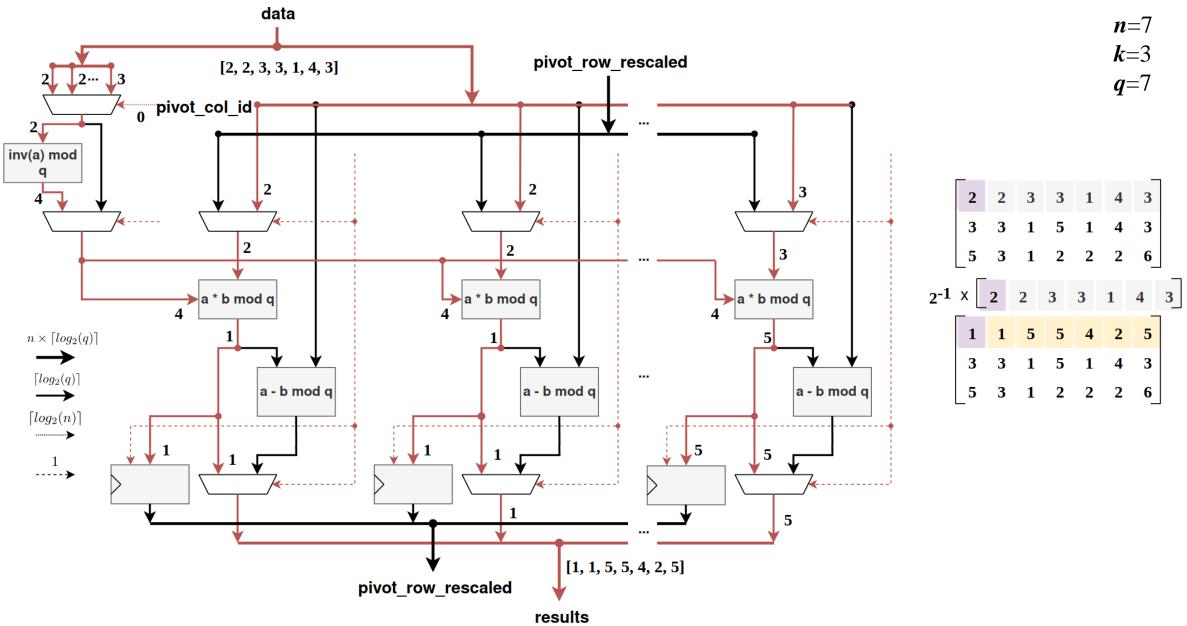
data i  $n \times \lceil log_2(q) \rceil$ w row id *n* **RAMs** to hold one  $\lceil log_2(q) \rceil$  $\rightarrow$ column of the matrix, each  $\left[ log_2(k) \right]$ →wdata a rdata a Parallel memory units to w row id-➤ addr\_a 1 access entire row of matrix ----> Address ----> Translation RAM clog2(k) x clog2(k) in one cycle →wdata b **Address translation tables** pivot\_row\_id -→addr b rdata b data i data o 🔶 data i data o 🔶 data i data o for constant time waddr waddr waddr conditional row swap RAM 0 ---->we RAM 1 .... RAM n-1 ----> we ---->we k x clog2(q) k x clog2(q) k x clog2(q) Separate input and output r\_row\_id raddr raddr raddr rdata a ►wdata a ►addr a r row id-.... Address ----> Translation RAM clog2(k) x clog2(k) →wdata b pivot\_row\_id · ►addr b rdata b data o 23

## **RREF** – Row Arithmetic

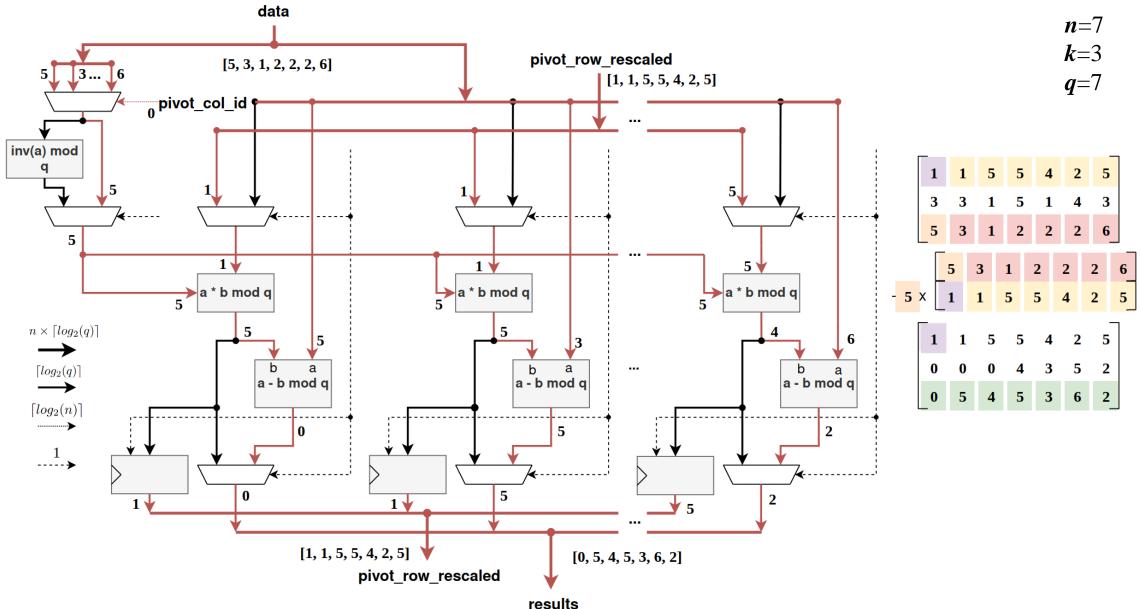
- Hardware re-use for • *Rescale Pivot Row* and *Reduce Other Rows*
- Parallel arithmetic units to ٠ operate on entire row at a time.
- Long feed forward critical ulletpath – good for pipelining
- **Registers** to hold result • from Rescale Pivot Row to be used during *Reduce Other* Rows



#### RREF – Row Arithmetic – Rescale Pivot Row



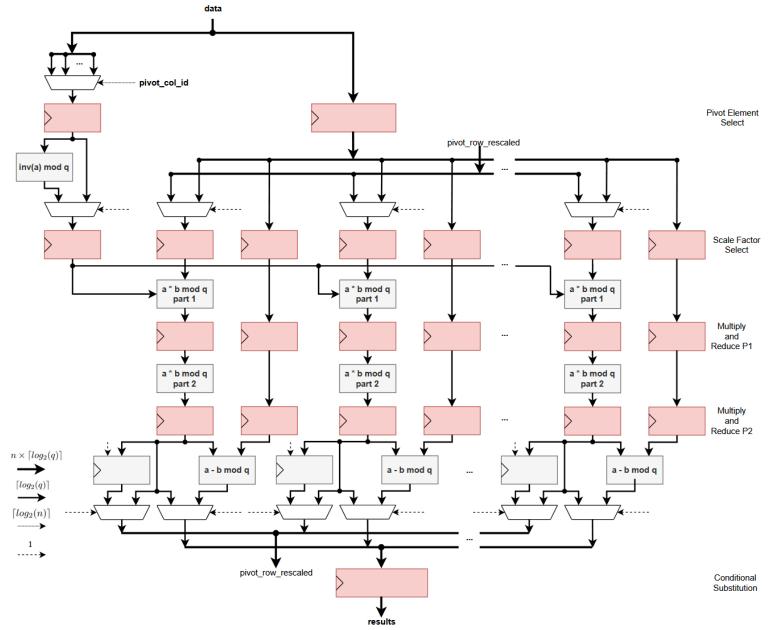
#### RREF – Row Arithmetic – Reduce Other Rows



## **RREF – Row Arithmetic Pipeline**

- Enables higher clock frequency
- Generates new result every clock cycle
- Increases flip-flop utilization

NIST Security Level	n	k	Frequency (MHz)
1	252	126	200
3	400	200	167
5	548	274	142

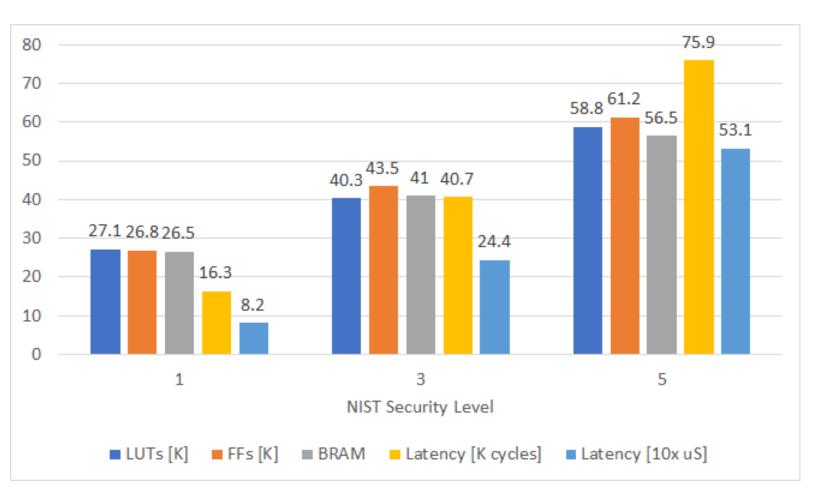


27

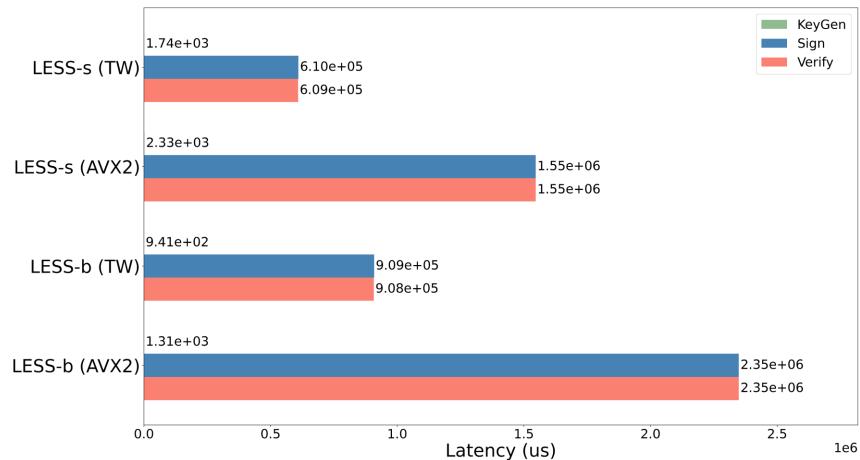
#### RREF – Results

- Area ~ *n*
- *n* Frequency
- Latency (Cycles) ~  $k^2$

NIST Security Level	n	k	Frequency (MHz)
1	252	126	200
3	400	200	167
5	548	274	142



## Improvement Over AVX2 [Level 5]



KeyGen

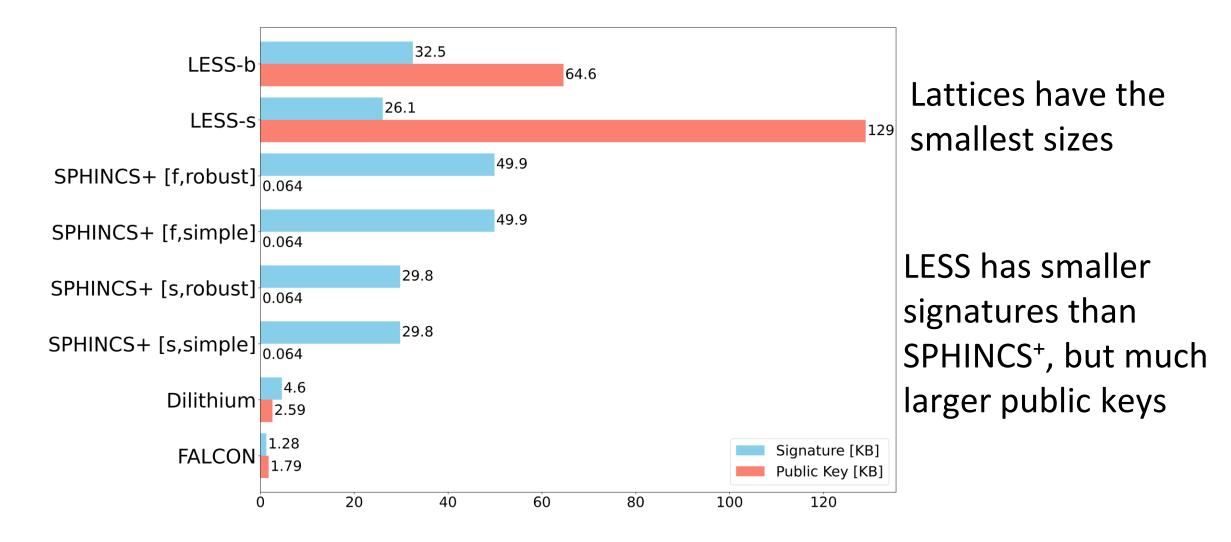
TW (Hardware): **Evaluated on Artix-7** 

AVX2 (Software): Evaluated on Ryzen 5 5600G running @3.9 GHz  $27.3 \times higher frequency$ than HW

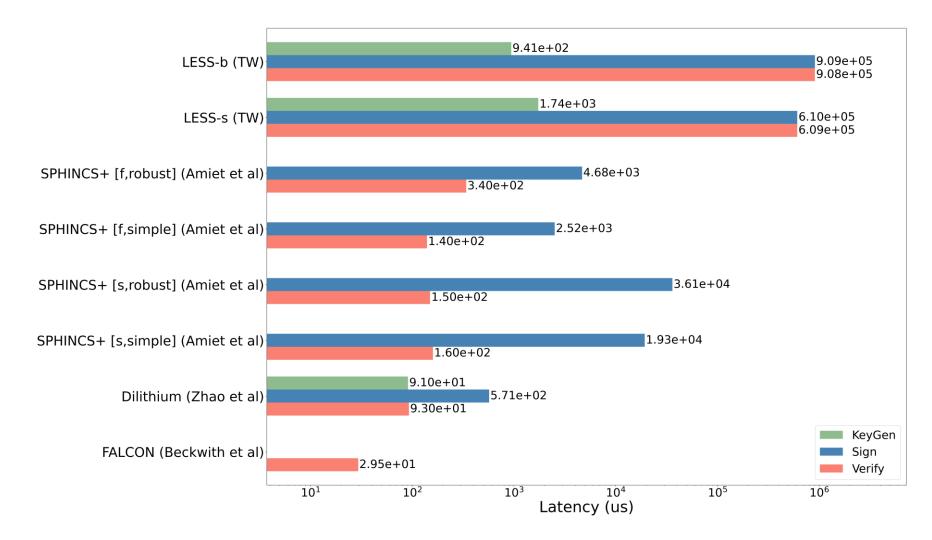
#### HW is faster by a factor of

- **1.4x** for Keygen ٠
- **2.5x** for Sign and Verify  $\bullet$

#### Transmission Cost [Level 5]



## Latency [Level 5]

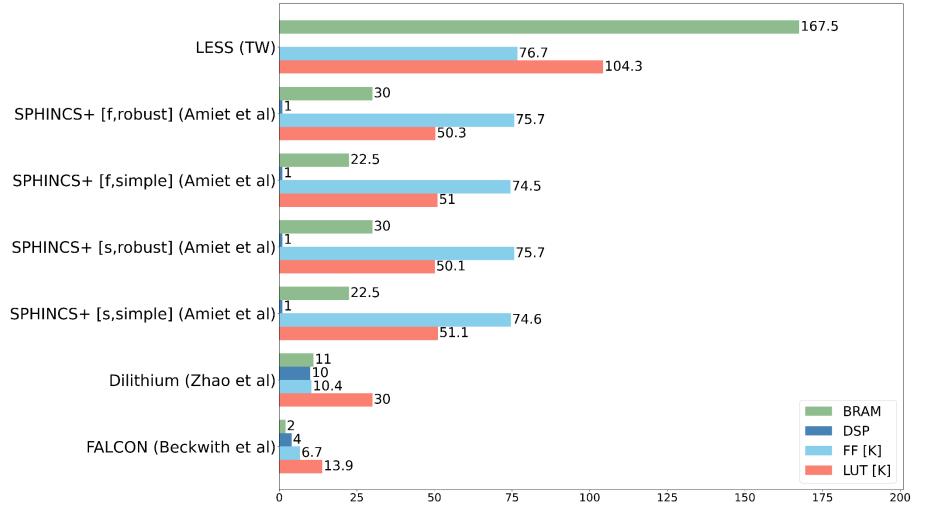


## Compared to SPHINCS+

Order of magnitude slower signing

**Several** orders of magnitude slower for verification

## Area [Level 5]



Compared to SPHINCS+ 2 × more LUTs Similar number of DSP/FF 6 × more BRAM

#### Conclusion

- This work represents the first hardware work on the new candidate LESS
- Our implementation running on an Artix-7 FPGA outperforms optimized AVX2 by  ${\sim}2~{\times}$
- LESS provides smaller signature sizes than SPHINCS<sup>+</sup>, but at the cost of larger public keys and slower signing/verification

# Questions?





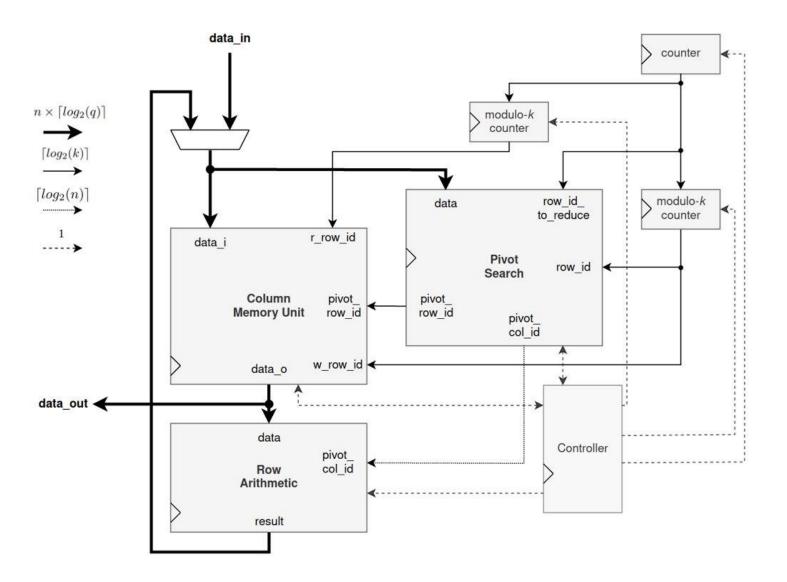


https://cryptography.gmu.edu/athena

Page: PQC

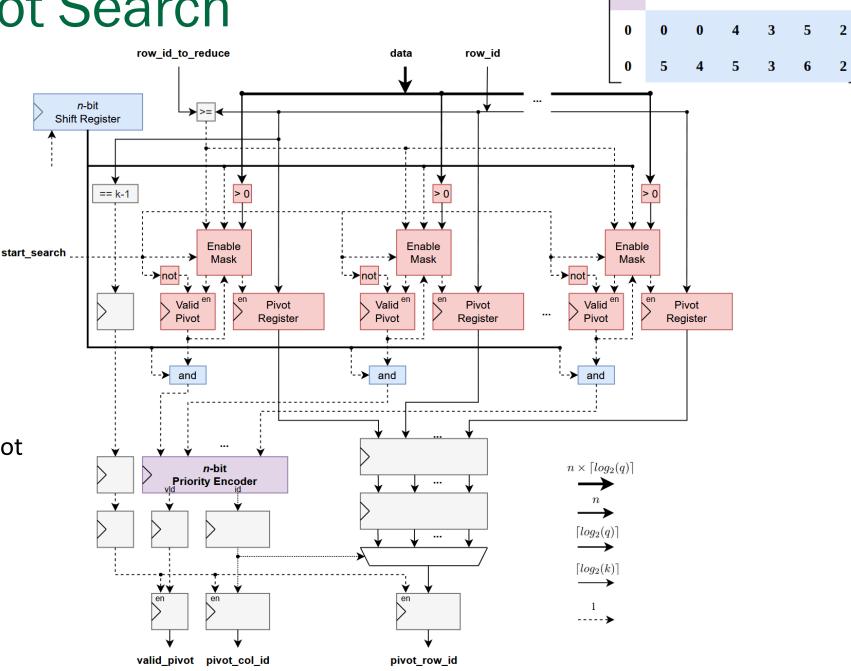
## RREF – Top-Level Unit

- Operates on entire row of the input matrix at a time
- Constant time *Pivot Search* implementation
- Parallel *Pivot Search* with initialization and *Reduce Other Rows*
- Pipelined arithmetic for increased clock frequency and high throughput



## RREF – Pivot Search

- Constant-time search hardware
- Search area decreases as algorithm progresses
- **Parallel units** to identify the non-zero element in every column of a row
- Large priority encoder to identify "left-most" non-zero element in search area as pivot element

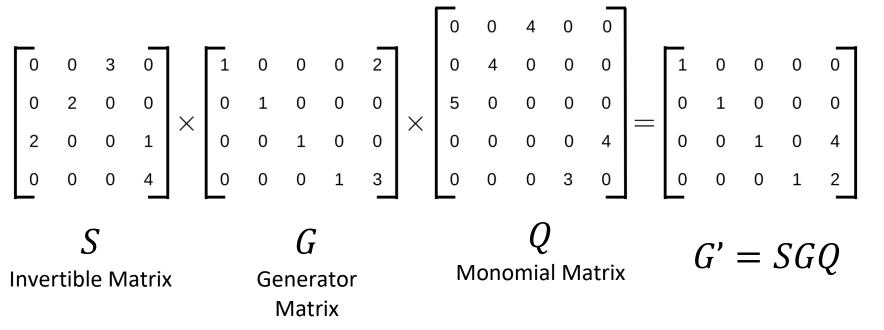


37

### Linear Equivalence Problem

### Linear Equivalence Problem (LEP):

Given two matrices  $G, G' \in F_q^{k \times n}$  which generate codes C, C', determine if the two corresponding codes are linearly equivalent. That is, does there exist matrices  $Q \in M_n$  and  $S \in GL(k)$  such that G' = SGQ where GL(k) is the set of invertible matrices.



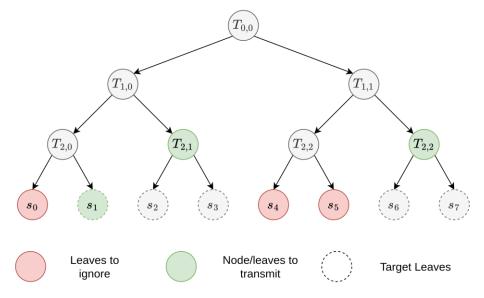
## Introduction: LESS Optimizations

#### **Commitment Seed Tree:**

- Commitment matrices are sampled using a leaf of a tree as the seed
- Benefit: Rather than sending the seeds of all zero-challenges ( $\tilde{Q}$ ), we can send the path nodes needed to generate them

#### **Information Sets:**

- For nonzero-challenges  $(Q^{-1} \times \tilde{Q})$ , send only the k columns of the monomial which are needed to calculate the pivot columns of the commitment
- Non-pivot columns are minimized and sorted to account for lack of scaling/permuting
- Benefit: Cost of non-zero transmissions is cut in half



Example of path nodes saving transmission cost in seed tree

## **Computational Bottlenecks:**

#### **Conversion to RREF:**

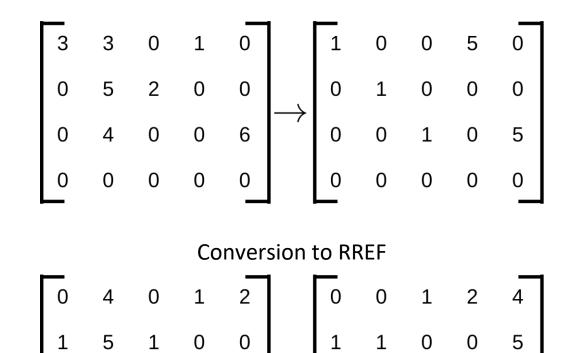
- Requires  $k^2 * n$  operations
- ~80% of the latency in software

#### Column Sorting:

- Non-pivot columns are sorted before hashing commitment
- Column-wise sorting requires transposition before and afterwards for optimal performance

#### **Generator Sampling:**

- On-the-fly sampling used to reduce BRAM requirement
- $K^2$  coefficients (up to 75K) coefficients needed



Column sorted using element-wise comparison

6

8

3

0

0

0

6

0

4

3

0

6

8

0

0

6

### Results

### Hardware Comparison Platform:

- Device: Artix-7 FPGAs
- Area: LUTs, FFs, DSP, BRAM
- Performance: Latency in  $\mu s$

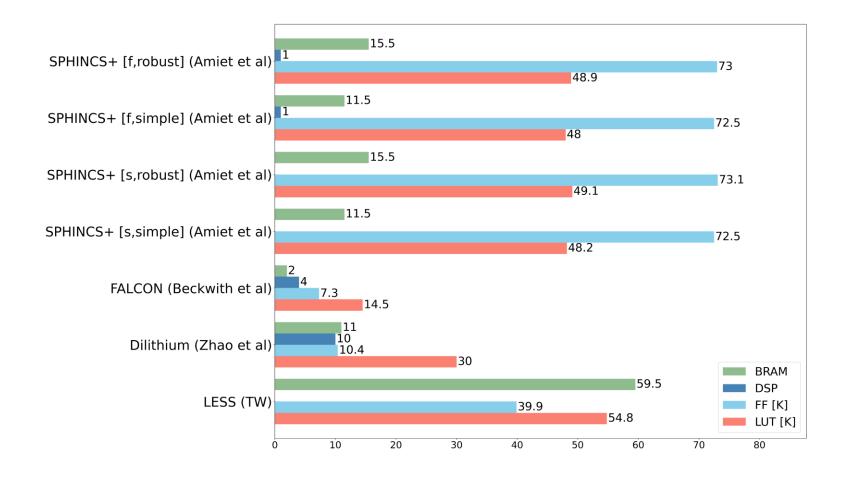
#### Software Comparison Platform:

- Device: Ryzen 5 5600G
- Implementation: AVX2
- Performance: Latency in  $\mu s$

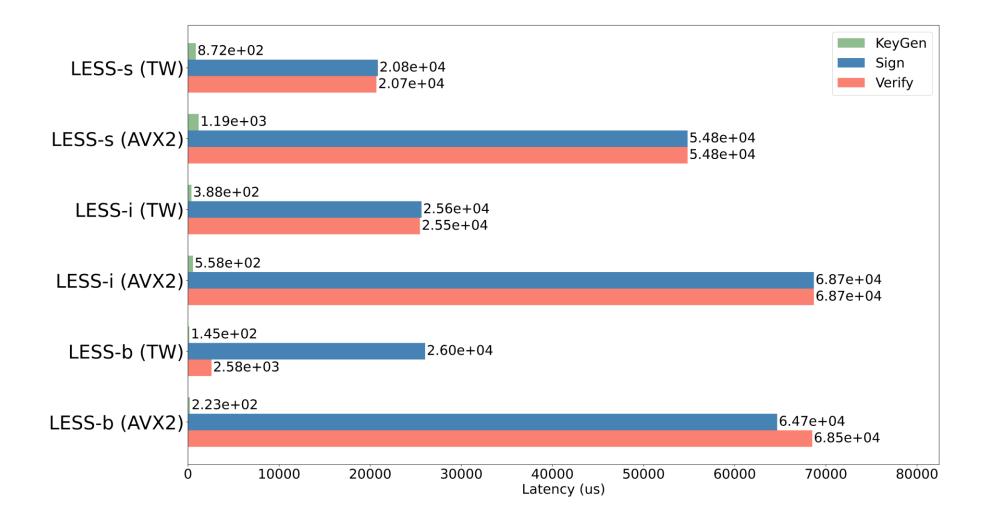
Algorithm	Designer	Platform	Parameter Set Selection	Keygen	Sign	Verify
LESS	TW	Artix-7 FPGA	Synthesis	Yes	Yes	Yes
SPHINCS+	Amiet		Synthesis	Νο	Yes	Yes
Dilithium	Zhao		Runtime	Yes	Yes	Yes
FALCON	Beckwith		Synthesis	Νο	Νο	Yes

 $TW \rightarrow This Work$ 

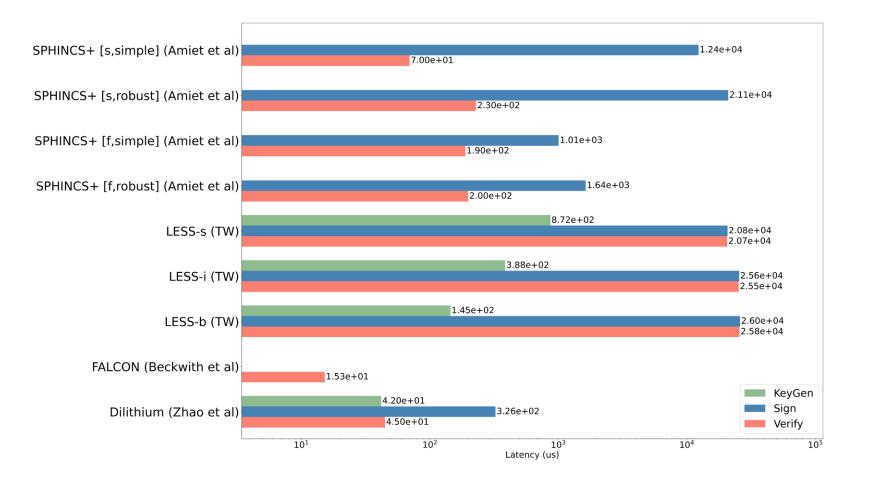
# Area [Level 1]



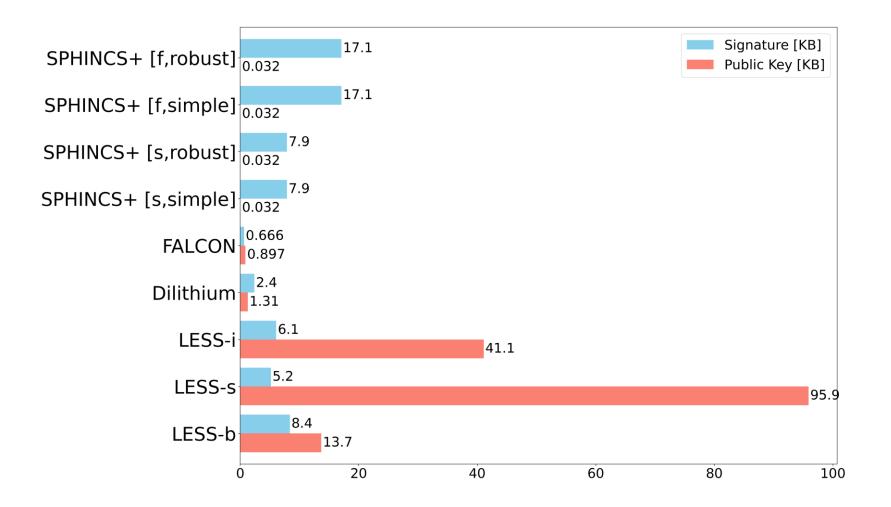
### AVX2 Comparison [Level 1]



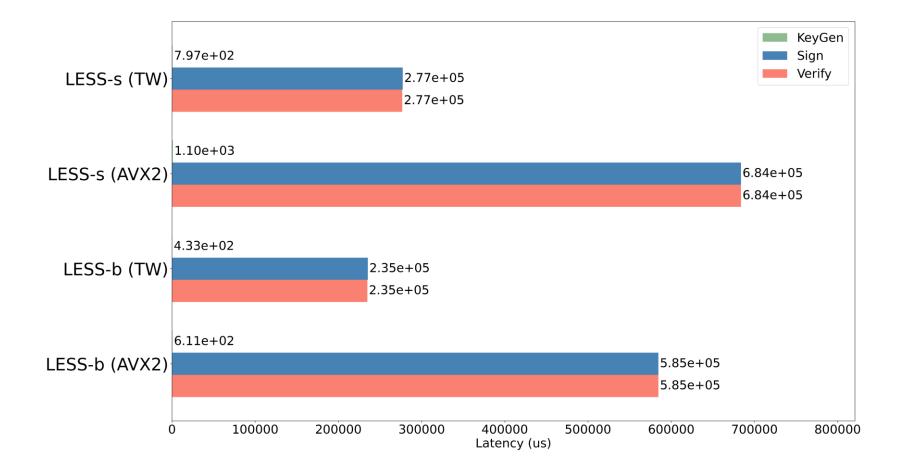
# Latency [Level 1]



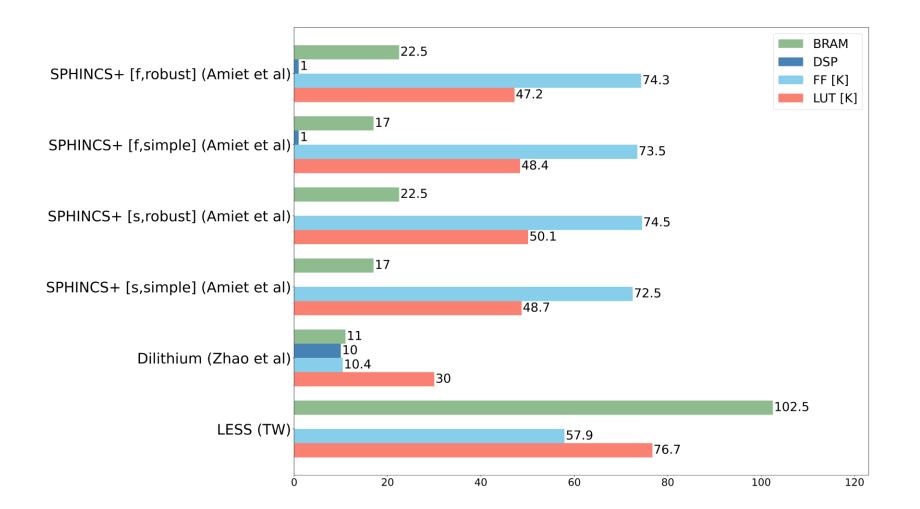
## Transmission Cost [Level 1]



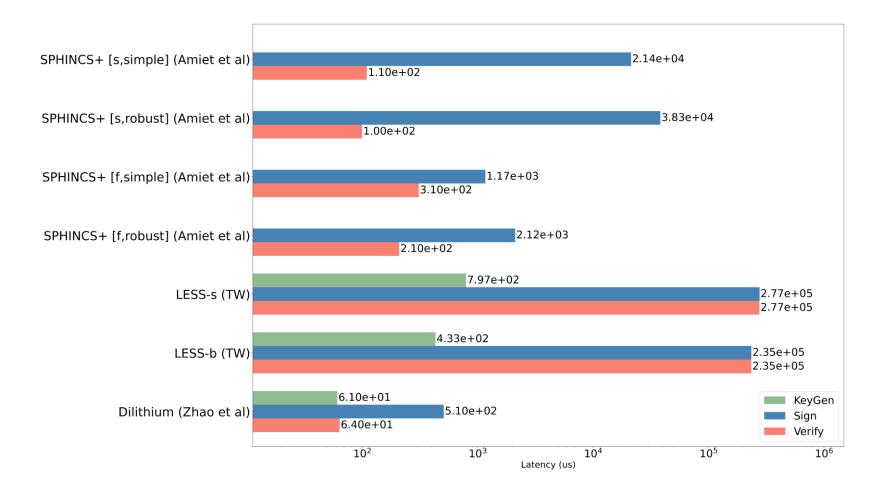
### AVX2 Comparison [Level 3]



# Area [Level 3]



# Latency [Level 3]



### Transmission Cost [Level 3]

