Practical Sublinear Proofs for R1CS from Lattices

Ngoc Khanh Nguyen, **Gregor Seiler**

IBM Research Europe

August 16, 2022

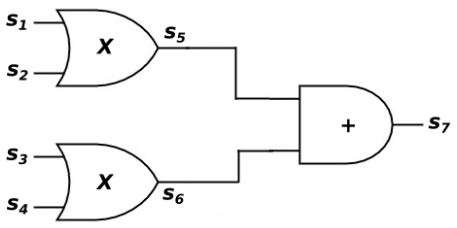
Zero-Knowledge Proof Systems

General Structure:

- Commit to secret vector $\vec{s} \in \mathbb{Z}_q^n$
- Prove linear and product relations between coefficients of $\vec{s}$, e.g.

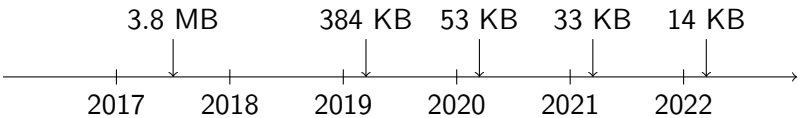
$$a_1 s_1 + a_2 s_2 = 0 \qquad s_1(s_1 - 1) = 0$$

This allows to prove that ***s*** satisfies an arbitrary arithmetic circuit



History and Status of Lattice-Based Proof Systems

Big improvements in *linear-sized* lattice-based proofs over the last three years



Advantages of lattice-based proofs outside of size:

- Lattice-based schemes can be very fast (highly parallelizable polynomial arithmetic)
- No large overhead in memory requirement

Beyond Linear Size

Linear-sized proof systems sufficient e.g. for constructing schemes for privacy-preserving crypto

But: Need standard-model underlying schemes that would never be used elsewhere

This paper: First practical *sublinear* lattice-based proof system

Fortunately, techniques developed for linear-sized proof systems can be transported to sublinear systems

Approximate Proofs and Weak Openings

Given commitment $\boldsymbol{t} = \boldsymbol{A}\boldsymbol{s}$ with short $\boldsymbol{s}$, want to prove knowledge of opening

Approximate Schnorr-Lyubashevsky Proof:

1. Prover commits to mask $\boldsymbol{w} = \boldsymbol{A}\boldsymbol{y}$
2. Verifier sends challenge polynomial $\boldsymbol{c} \in \mathcal{C}$
3. Prover sends masked opening $\boldsymbol{z} = \boldsymbol{y} + \boldsymbol{c}\boldsymbol{s}$ (rejects if $\boldsymbol{z}$ reveals secret information)
4. Verifier checks $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{w} + \boldsymbol{c}\boldsymbol{t}$ and $\|\boldsymbol{z}\| \leq \beta$

Extraction:

Special sound: Two accepting transcripts lead to $\bar{\boldsymbol{c}}\boldsymbol{t} = \boldsymbol{A}\bar{\boldsymbol{z}}$.
Removing $\bar{\boldsymbol{c}}$ gives a *weak opening* such that

$$\boldsymbol{t} = \boldsymbol{A}\boldsymbol{s}^* \quad \text{and} \quad \|\bar{\boldsymbol{c}}\boldsymbol{s}^*\| \leq 2\beta$$

Amortized Opening Proofs

Given many commitments $\boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_i$ for $i = 1, \ldots, n$, want to prove knowledge of all openings at once with small total cost

Similar approximate opening proof where prover sends *amortized masked opening*

$$\boldsymbol{z} = \boldsymbol{y} + \boldsymbol{c}_1\boldsymbol{s}_1 + \cdots + \boldsymbol{c}_n\boldsymbol{s}_n$$

Extraction:

Not special-sound but two transcripts that differ only in one $\boldsymbol{c}_i$ lead to weak opening for $\boldsymbol{t}_i$

Proving Additional Relations

Want to use masked opening $\boldsymbol{z} = \boldsymbol{y} + \boldsymbol{c}_1\boldsymbol{s}_1 + \cdots + \boldsymbol{c}_n\boldsymbol{s}_n$ to prove additional relations on secret $\boldsymbol{s}$

Standard approach:

- Evaluating relations on $\boldsymbol{z}$ leads to multivariate polynomial in $\boldsymbol{c}_i$
- Prover commits to coefficients before $\boldsymbol{c}_i$ are known ("garbage terms")

For example ($n = 2$),

$$\langle \boldsymbol{z}, \boldsymbol{z} \rangle = \boldsymbol{g}_0 + \boldsymbol{g}_1\boldsymbol{c}_1 + \boldsymbol{g}_2\boldsymbol{c}_2 + \boldsymbol{g}_3\boldsymbol{c}_1\boldsymbol{c}_2 + \boldsymbol{g}_4\boldsymbol{c}_1^2 + \boldsymbol{g}_5\boldsymbol{c}_2^2$$

Interested only in highest-degree coefficients, e.g. $\boldsymbol{g}_4 = \langle \boldsymbol{s}_1, \boldsymbol{s}_1 \rangle$.

Two Problems:

1. Number of garbage terms
2. Form of $\boldsymbol{z}$ in extraction

Interactive Schwartz-Zippel

Quadratic expressions in $\boldsymbol{z}$ are quadratic multivariate polynomials $f(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)$ in the challenges: Need $O(n^2)$ garbage terms — prohibitively expensive

Solution via Interaction:

- View f as sum of univariate polynomials in $\boldsymbol{c}_i$ with implicit dependence on earlier challenges: $f(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n) = f_0 + f_1(\boldsymbol{c}_1) + f_2^{(\boldsymbol{c}_1)}(\boldsymbol{c}_2) + \cdots + f_n^{(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{n-1})}(\boldsymbol{c}_n)$
- In $(2n+1)$-round protocol prover sends coefficients of f_i before getting $\boldsymbol{c}_i$

Example

$$\langle \boldsymbol{z}, \boldsymbol{z} \rangle = \underbrace{\boldsymbol{g}_0}_{f_0} + \underbrace{(\boldsymbol{g}_1 \boldsymbol{c}_1 + \boldsymbol{g}_4 \boldsymbol{c}_1^2)}_{f_1(\boldsymbol{c}_1)} + \underbrace{(\boldsymbol{g}_{23} \boldsymbol{c}_2 + \boldsymbol{g}_5 \boldsymbol{c}_2^2)}_{f_2^{(\boldsymbol{c}_1)}(\boldsymbol{c}_2)} \quad \text{where} \quad \boldsymbol{g}_{23} = \boldsymbol{g}_2 + \boldsymbol{g}_3 \boldsymbol{c}_1$$

$$\xrightarrow{\boldsymbol{g}_0, \boldsymbol{g}_1, \boldsymbol{g}_4}$$

$$\xleftarrow{\boldsymbol{c}_1}$$

$$\xrightarrow{\boldsymbol{g}_{23}, \boldsymbol{g}_5}$$

$$\xleftarrow{\boldsymbol{c}_2}$$

Intuition for Security:

The f_i are multiples of $\boldsymbol{c}_i$. Prover can not correct errors in earlier f_i as corrections would get randomly distorted by next challenge

Well-Formedness

Analysis of soundness requires well-formedness of z, i.e. $z = y^* + c_1 s_1^* + \cdots + c_n s_n^*$ in (almost) all accepting transcripts.

Non-amortized case ($n = 1$):

We have $\bar{c}$ such that $\bar{s}_1^*$ is short. Then

$$\bar{c} y^* = \bar{c} z - c_1 \bar{c} s_1^*$$

is also short and therefore a weak opening for w

Amortized case:

The same argument doesn't work in the amortized case as we would need to multiply by many different $\bar{c}$

Well-Formedness in Amortization

Combination with shortness (e.g. binary) proof allows to prove well-formedness

Sketch of argument:

- If one of the s_i^* is non-binary, let $s_{i_0}^*$ be the last non-binary vector. Write

$$z = y^* + c_{i_0} s_{i_0}^* + c_{i_0+1} s_{i_0+1}^* + \cdots + c_n s_n^*$$

  in every accepting transcript with fixed challenges $c_1, \ldots, c_{i_0-1}$. Then

$$\bar{c} y^* = \bar{c} z - c_{i_0} \bar{c} s_{i_0}^* - \bar{c} c_{i_0+1} s_{i_0+1}^* - \cdots - \bar{c} c_n s_n^*$$

  is short and hence y^* is a weak opening for $w + c_1 t_1 + \cdots + c_{i_0-1} t_{i_0-1}$
- Therefore, y^* must be the same in all accepting transcripts with fixed $c_1, \ldots, c_{i_0-1}$
- Binary proof collapses to binary proof for trailing vectors
- Prover has small success probability in it

High-level Overview of our Proof System

Basic square root proof system:

- Split vector $\boldsymbol{s}$ of dimension mn into m vectors $\boldsymbol{s}_i$ of dimension n where $m \approx n$
- Commit to the parts, $\boldsymbol{t}_i = \boldsymbol{A}\boldsymbol{s}_i$
- Run amortized proof

Concretely efficient proof system:

- Recommit to $\boldsymbol{t}_i$ in Merkle tree
- Proof tree with one amortized proof per level (see paper)

Comparison with Ligero

		Proof Size	
Number of constraints	Soundness error	Ligero	**Our System**
2^{19}	2^{-115}	4.58 MB	**4.53 MB**
2^{20}	2^{-114}	8.35 MB	**5.22 MB**
2^{21}	2^{-113}	8.90 MB	**6.08 MB**
2^{22}	2^{-112}	16.23 MB	**7.19 MB**
2^{23}	2^{-111}	17.39 MB	**10.79 MB**
2^{24}	2^{-110}	31.83 MB	**13.21 MB**
2^{25}	2^{-109}	34.15 MB	**16.59 MB**
2^{26}	2^{-108}	62.14 MB	**21.68 MB**
2^{27}	2^{-107}	66.03 MB	**29.04 MB**
2^{28}	2^{-106}	121.90 MB	**42.42 MB**

Thank you!