Interactive Oracle Proofs and R1CS-satisfiability

Alexander Lindenbaum

Notes are from Chapters 8 and 10 of [Tha22].

1 IOPs: Definition and Associated Succinct Arguments

PCPs are great for non-interactivity and verifier runtime, but the prover requires massive time/space resources. Interactive Oracle Proofs (IOPs) generalize PCPs and IPs, and are more efficient protocols. An IOP is an IP where in each round the verifier is given query access to the prover's messages' bits. We aim for the IOP verifier to run in time sublinear in the total proof length.

Similarly from public-coin IPs to non-interactive arguments, IOPs can be transformed into non-interactive arguments in the random oracle model, using Merkle-hashing and Fiat-Shamir. Instead of the prover sending the message each round, the prover sends a Merkle-commitment to the IOP prover's message. The verifier simulates the IOP verifier to determine which elements to query, then asks the prover to reveal the relevant symbols and provide authentication. Then apply Fiat-Shamir to that protocol to make it non-interactive. Hence we focus on succinct IOPs, with this transformation in mind.

2 Polynomial IOPs and Associated Succinct Arguments

2.1 R1CS-Satisfiability

First, a detour into R1CS-Satisfiability, a computational problem which, akin to arithmetic circuit satisfiability, is a popular method to encode instances which we want zero knowledge protocols for.

Definition 1. A *rank-1 constraint system (R1CS) instance* is specified by three $m \times n$ matrices A, B, C with entries from a field $\mathbb{F}$ and is satisfiable if and only if there is a vector $z \in \mathbb{F}^n$ with $z_1 = 1$ such that

$$(Az) \circ (Bz) = Cz,$$

where $\circ$ denotes entrywise product.

Some remarks:

- Any z satisfying this equation is analogous to the "correct intermediary evaluations" of an arithmetic circuit, in the context of arithmetic circuit satisfiability.
- We enforce $z_1 = 1$, otherwise $z = 0^n$ trivially satisfies this equation for all matrices A, B, C. This also ensures that there is an efficient transformation from circuit-SAT to R1CS-SAT (not mentioned here).

- The ith rows a_i, b_i, c_i specify a "rank-one constraint" because we must satisfy

$$\langle a_i, z\rangle \cdot \langle b_i, z\rangle = \langle c_i, z\rangle.$$

2.2 Polynomial IOPs

In a *polynomial IOP*, each message that the prover sends is a polynomial h_i over a finite field $\mathbb{F}$, of degree $\leq d_i$. The verifier is given query access to evaluations of h_i, so $\mathcal{V}$ chooses an input r to h_i to learn $h_i(r)$.

We cover a polynomial commitment scheme which will allow the prover's messages to be implemented via a standard IOP. So each polynomial h_i will be specified by a certain commitment string m_i. Then when we take a polynomial IOP for R1CS-satisfiability, and replace each of $\mathcal{P}$'s messages and associated evaluation querys with a polynomial commitment scheme based on a standard IOP, we get a standard IOP for R1CS-satsifiability. We can turn that into a succict argument via Fiat-Shamir.

Even if the polynomial commitment scheme is not a standard IOP, we can still obtain a succinct argument with these steps:

1. Design a public-coin polynomial IOP for circuit-/R1CS-satisfiability.
2. Obtain a public-coin, interactive succinct argument by replacing $\mathcal{P}$'s messages h_i with a polynomial commitment scheme.
3. Remove interaction via Fiat-Shamir.

This is a recipe for SNARKs.

3 A Polynomial IOP for R1CS-satisfiability

3.1 The univariate sum-check protocol

Fact 2. *Let $\mathbb{F}$ be a finite field and $H \subseteq \mathbb{F}$ a multiplicative subgroup of size n. Then for any polynomial q of degree less than n,*

$$\sum_{a \in H} q(a) = q(0) \cdot n.$$

It follows that $\sum_{a\in H} q(a)$ is 0 if and only if $q(0) = 0$.

The proof is in Chapter 10 of [Tha22]. Let H be a multiplicative subgroup of $\mathbb{F}$ of size n. Let p by any univariate polynomial of degree $\leq D$, where D may be greater than n. We denote the vanishing polynomial of H:

$$Z_H(X) = \prod_{a \in H} (X - a) = X^n - 1.$$

Z_H is sparse, and so can be efficiently evaluated at any r with repeated squaring.

Lemma 3. *$sum_{a\in H} p(a) = 0$ if and only if there exists polynomials h^*, f with $deg(h^*) \leq D - n$ and $deg(f) < n - 1$ satisfying*

$$p(X) = h^*(X) \cdot Z_H(x) + X \cdot f(X).$$

This gives us a polynomial IOP for proving that a univariate polynomial p of degree D sums to 0 over a multiplicative subgroup H with $|H| = n$. The prover sends two messages specifying f and h^*. The verifier can confirm that the equation holds by evaluating both sides at a random point $r \in \mathbb{F}$. Since the max degree of all polynomials is $\max(D, n)$, up to soundness error $\max(D, n)/|\mathbb{F}|$, if the equation holds at a random point $r \in \mathbb{F}$ then it is safe for the verifier to believe that the equation holds for all of $\mathbb{F}$.

3.2 Polynomial IOP for R1CS-SAT via Univariate Sum-Check

Recall that we have three $m \times n$ matrices A, B, C, and the prover wants to show that there is a z such that $Az \circ Bz = Cz$. Suppose for simplicity that $m = n$ and that there is a multiplicative subgroup H of size n. Label the n entries of z with elements of H, and let $\hat{z}$ be the unique univariate polynomial of degree $\leq n - 1$ such that $\hat{z}(h) = z_h$ for all $h \in H$. Similarly, let $z_A = Az$, $z_B = Bz$, and $z_C = Cz$ and $\hat{z_A}$, $\hat{z_B}$, $\hat{z_C}$ be the unique polynomials extending z_A, z_B, z_C, resp.

To check that $Az \circ Bz = Cz$, the verifier must confirm that

1. for all $h \in H$, $\hat{z}_A(h) \cdot \hat{z}_B(h) = \hat{z}_C(h)$, and
2. for all $h \in H$, and $M \in \{A, B, C\}$, $\hat{z}_M(h) = \sum_{j \in H} M_{h,j} \cdot \hat{z}(j)$.

2. confirms that $z_M = M$ for all M, and 1. confirms that $Az \circ Bz = Cz$. The prover will send $\hat{z}, \hat{z}_A, \hat{z}_B$, and $\hat{z}_C$.

Checking 1. This check is equivalent to the existence of a polynomial h^* of degree $\leq n$ such that

$$\hat{z}_A(X) \cdot \hat{z}_B(X) - \hat{z}_C(X) = h^*(X) \cdot Z_H(X).$$

The prover sends a message specifying h^*. The verifier checks that the equation holds under a random $r \in \mathbb{F}$.

Checking 2. Fix $M \in \{A, B, C\}$. Let $\hat{M}(X, Y)$ denote the bivariate low-degree extension of M, interpreted in the manner as a function $M(x, y) : H \times H \to \mathbb{F}$ via $M(x, y) = M_{x,y}$. So $\hat{M}(x, y)$ is the unique bivariate polynomial of degree $\leq n$ in that each variable extends M. Equation 2. holds for all $h \in H$ if and only

$$\hat{z}_M(X) = \sum_{j \in H} M(\hat{X}, j)\hat{z}(j).$$

Any two distinct polynomials of degree $\leq n$ can agree on at most n inputs, so the verifier chooses $r' \in \mathbb{F}$ at random, then up to error $n/|\mathbb{F}|$ over choice of r', the equation holds if and only if $\hat{z}_M(r') = \sum_{j \in H} \hat{M}(r', j)\hat{z}(j)$. But this check is not as simple as 1: when we set

$$q(Y) = \hat{M}(r', Y)\hat{z}(Y) - \hat{z}_M(r') \cdot |H|^{-1},$$

then the validity of 2. with r' is equivalent to $\sum_{j \in H} q(Y) = 0$. To finish this check, the verifier requests that the prover establish that $\sum_{j \in H} q(Y) = 0$ via a univariate sum-check protocol. At the end of the sum-check protocol, the verifier evaluates q at a randomly chosen r''. $\hat{z}(r'')$ and $\hat{z}_M(r')$ cqan be evaluated in a constant number of field operations, but $\hat{M}(r', r'')$ is not so obvious.

Holography/computation commitments The goal is for an untrusted prover to efficiently and verifiably reveal $\hat{M}(r', r'')$ to the verifier, where the pre-processing time and the runtime of the prover when revealing $\hat{M}(r', r'')$ should be linear in the number of nonzero entries of M, called K. I do not cover the protocol for acheiving this, but was solved in [CHM+19] and [COS19].

References

[CHM+19] Alessandro Chiesa, Yuncong Hu, Mary Maller, Pratyush Mishra, Psi Vesely, and Nicholas Ward. Marlin: Preprocessing zksnarks with universal and updatable srs. Cryptology ePrint Archive, Paper 2019/1047, 2019. https://eprint.iacr.org/2019/1047.

[COS19] Alessandro Chiesa, Dev Ojha, and Nicholas Spooner. Fractal: Post-quantum and transparent recursive proofs from holography. Cryptology ePrint Archive, Paper 2019/1076, 2019. https://eprint.iacr.org/2019/1076.

[Tha22] Justin Thaler. Proofs, arguments, and zero-knowledge. *Foundations and Trends in Privacy and Security*, 4(2–4):117–660, 2022.