A summary on the FRI low degree test

Ulrich Haböck*

Orbis Labs
team@orbislabs.com

September 13, 2022

Abstract

This document is an informal summary on the FRI low degree test [BSBHR18a], [BSCI+20], and DEEP algebraic linking from [BSGKS20]. Based on its most recent soundness analysis [BSCI+20], we discuss parameter settings for practical security levels, how FRI is turned into a polynomial commitment scheme, and the soundness of DEEP sampling in the list decoding regime. In particular, we illustrate the DEEP method applied to proving satisfiability of algebraic intermediate representations and prove a soundness error bound which slightly improves the one in [Sta21].

Contents

*This work is supported by Horizenlabs Italia and Orbis Labs.

Introduction

FRI, in full length *Fast Reed-Solomon Code Interactive Oracle Proof of Proximity*, is a low-degree test for functions on an FFT domain, i.e. a smooth multiplicative subgroup D of a finite field F. Given a function

$$f : D \longrightarrow F$$

FRI proves that f corresponds to a polynomial of low degree with respect to the size of D.

The oracles provided by the FRI prover are again functions on D, or a subdomain of it, and the verifier queries the values at points from their domain only. Due to the small size of D (compared to the cryptographically large sampling spaces of polynomial IOPs) the key tool for distinguishing one polynomial from another is statistical sampling. However, a statistical test can only assure *proximity*, which we measure by the fractional Hamming distance

$$\delta(f, g) = \frac{1}{|D|} \cdot |\{x \in D : f(x) \neq g(x)\}|.$$

In FRI the prover convinces the verifier that a given function $f : D \longrightarrow F$ is θ-*close* (and not necessarily equal) to a low-degree polynomial, i.e.

$$\delta(f, p) \leq \theta,$$

for some polynomial $p(X)$ of specified maximum degree. In words, f agrees with $p(X)$ on a set $A \subseteq D$ of density $\frac{|A|}{|D|} \geq 1 - \theta$. In applications the agreement set is chosen large enough to infer global properties on the low degree polynomial. It is exactly this inference principle which makes FRI applicable to proving algebraic relations between a set of low-degree polynomials, might it be circuit satisfiability or the evaluation identities for building a polynomial commitment scheme.

We stress that fact that this summary does not present any novelties. Instead it is an outcome of my learnings when reading the papers [BSCI+20], [BSGKS20], [BSBHR18a], [KPV19] and [Sta21]. The document provides an overview of FRI and its soundness analysis, including some background on decoding Reed-Solomon codes. It discusses the DEEP method and how it is related to polynomial commitment schemes and we sketch the more general notion of list polynomial commitment schemes [KPV19]. Finally we illustrate how soundness error bounds are proven for the DEEP method in the list decoding regime. In the course the latter we clarify two points of [Sta21], which are the usage of degree correction factors

(these are not needed for the DEEP method), and the quadratic occurence of the decoder list size bound in their soundness error formula, which can be replaced by a linear term.

We assume that the reader knows (public-coin) interactive oracle proofs and their security notions [BSCS16], such as soundness, proof of knowledge, and statistical (i.e. perfect) honest verifier zero-knowledge. Any IOP with these security properties can be compiled into a succinct non-interact-ive argument of knowledge in the random oracle model [BSCS16]: The prover oracle messages are committed by Merkle roots using the random oracle, and the verifier coins are the answers of the random oracle given the prover messages as its input.

Notation

Throughout the document we assume that the size of the sampling domain D and the number of coefficients k are both powers of two, and that the multiplicative subgroup F^* of the finite field F is smooth enough to contain a subgroup of order k and $|D|$. The absolute Hamming distance between two function $f, g \in F^D$ is

$$\Delta(f,g) = |\{x \in D : f(x) \neq g(x)\}|,$$

and we shall write

$$\delta(f,g) = \frac{1}{|D|} \cdot \Delta(f,g)$$

for its fractional variant. Given any subset $V \subseteq F^D$, we denote by

$$\Delta(f,V) = \min_{v \in V} \Delta(f,v)$$

the minimal distance of $f \in F^D$ to V, and likewise we define the minimal fractional Hamming distance. We denote by

$$\mathsf{RS}_k[F,D] = \{\, p(x)|_{x \in D} : p(X) \in F[X], \deg p(X) \leq k-1\}$$

the Reed-Solomon code of rate $\rho = \frac{k}{n}$ over the domain of definition $D \subseteq F^*$. (Here, $p(x)|_{x \in D}$ denotes the domain evaluation, i.e. the functional restriction of $p(x)$ to D.) Whenever we say that a polynomial $p(X)$ belongs to $\mathsf{RS}_k[F,D]$, we mean that its domain evaluation $p(x)|_{x \in D}$ is a code word.

In the context of oracle proofs, we denote oracles for functions $f \in F^D$ by $[f]$, and occasionally call them *domain evaluation oracles* to distinguish from the oracle notion of univariate polynomial IOPs [BFS20], which models an ideal polynomial commitment scheme. In order to a closer alignment with the compiled protocol in the random oracle model, we prefer to say that a party P (the prover) "sends" $[f]$ to another party V (the verifier), meaning that P sets up the oracle for f and V obtains oracle access for it.

Correlated agreement

As in polynomial IOPs, building random linear combinations is the core reduction argument in FRI . While the soundness of it is easily proven in the polynomial model, this is not the case for domain evaluations. Even in the most elementary case, proving that if with noticeable probability a random linear combination of two given functions f_0, f_1 is θ-close to a Reed-Solomon codeword, i.e.

$$\delta\big(f_0 + \lambda \cdot f_1, \mathsf{RS}_k[F,D]\big) \leq \theta,$$

then a similar proximity would hold for f_0 and f_1 , is non-trivial, in particular when targeting only a small increase in the distance bound θ. The most advanced result is the correlated agreement theorem (or *proximity gap theorem*) of Ben-Sasson, et al. [BSCI$^+$20]. We state it for the case of algebraic curves, which is typically favored in the context of proof composition.

Theorem 1. *(Correlated agreement theorem, [BSCI+20], Theorem I.5) Let* $\mathsf{RS}_k = \mathsf{RS}_k[F, D]$ *be the Reed-Solomon code over a a finite field* F *with defining set* $D \subseteq F$ *and rate* $\rho = \frac{k}{|D|}$*. Given a proximity parameter* $\theta \in (0, 1 - \sqrt{\rho})$ *and words* f_0,f_1,...,$f_{N-1} \in F^D$ *for which*

$$\frac{\left|\left\{\lambda \in F : \delta(f_0 + \lambda \cdot f_1 + \ldots + \lambda^{N-1} \cdot f_{N-1}, \mathsf{RS}_k) \leq \theta\right\}\right|}{|F|} > \varepsilon,$$

where ε *is as in* (1) *and* (2) *below. Then there exist polynomials* $p_0(X)$, $p_1(X)$,...,$p_{N-1}(X)$ *belonging to* RS_k*, and a set* $A \subseteq D$ *of density* $\frac{|A|}{|D|} \geq 1 - \theta$ *on which* $f_0, \ldots, f_{N-1}$ *jointly coincide with* $p_0, \ldots, p_{N-1}$*, respectively. In particular,*

$$\delta(f_0 + \lambda \cdot f_1 + \ldots + \lambda^{N-1} \cdot f_{N-1}, \mathsf{RS}_k) \leq \theta$$

for every $\lambda \in F$*.*

The proof of the correlated agreement theorem, including concrete values for the soundness error bound ε, is an algebraic analysis of the Berlekamp-Welch or the Guruswami-Sudan list decoder over the rational function field $K = F(Z)$. It uses the Polichuk-Spielmann lemma to "glue together" the outputs of the decoder for $f_0 + \lambda \cdot f_1 + \ldots + \lambda^{N-1} \cdot f_{N-1}$ over the "small" field F by means of the decoder result for the word

$$f_0 + Z \cdot f_1 + \ldots + Z^{N-1} \cdot f_{N-1} \in K^D$$

over the infinite field K: If for a noticeable fraction of λâĂŹs the distance to the Reed-Solomon code is $\leq \theta$, then the same holds over $F(Z)$.

Depending on the decoding regime the following values for ε are obtained by [BSCI+20]:

1. *Unique decoding regime.* For $\theta \in \left(0, \frac{1-\rho}{2}\right]$, Theorem 1 holds with

$$\varepsilon = (N-1) \cdot \frac{|D|}{|F|}. \tag{1}$$

2. *List decoding regime.* For $\theta \in \left(\frac{1-\rho}{2}, 1 - \sqrt{\rho}\right)$ and setting $\theta = 1 - \sqrt{\rho} \cdot \left(1 + \frac{1}{2m}\right)$, with $m \geq 3$, Theorem 1 holds with

$$\varepsilon = (N-1) \cdot \frac{k^2}{|F| \cdot \min\left(\frac{1}{m}, \frac{1}{10}\right)^7} \approx (N-1) \cdot m^7 \cdot \rho^{-\frac{3}{2}} \cdot \frac{|D|^2}{|F|}. \tag{2}$$

For linear varieties of the form $f_0 + \lambda_1 \cdot f_1 + \ldots + \lambda_{N-1} \cdot f_{N-1}$ a similar result holds, with the $(N-1)$-term in (1) and (2) replaced by 1. See [BSCI+20], Theorem I.2.

Note that in contrast to the unique decoding regime, the sampling domain size $|D|$ occurs quadratically in the error bound, and therefore the field needs to be significantly larger to obtain the same magnitude of soundness as in the unique decoding regime. This quadratic occurrence is inherently connected with the Guruswami-Sudan-Johnson list size bound. It is conjectured by [BSGKS20] that Reed-Solomon codes over prime fields F are more "nicely" list decodable, even up to capacity bound $1 - \rho$, and that the sampling domain size occurs only linearly in the error bound. We will discuss this conjecture in Section 5.4.

FRI proof of proximity

Given a function $f \in F^D$ and its domain evaluation oracle $[f(x)|_{x \in D}]$, FRI is an interactive oracle proof for f being close to a word from $\mathsf{RS}_k[F, D]$,

$$\delta(f, \mathsf{RS}_k[F, D]) \leq \theta,$$

given a *proximity parameter* θ of at most the Johnson list decoding bound. As most interactive oracle proofs, the FRI protocol is comprised of a *commit phase* and a *query phase*. The commit phase consists of one or several rounds, in which the prover sends domain evaluation oracles to the verifier, who then responds with a random challenge. That phase of FRI performs a random reduction similar to the one of an inner product argument [BCC+16], at least halving the instance size with each step by a linear folding procedure. In the concluding query phase, the verifier asks for openings of the oracles at random points from their domain of definition. These openings are then used to check consistency of each reduction step of the commit phase.

Reduction

The commit phase of FRI starts with the instance to proven, i.e. the polynomial $p_0(X) = p(X)$ and its domain evaluation over $D_0 = D$. This instance is stepwised reduced by means of a random folding procedure, yielding a sequence of polynomials

$$p_0(X), p_1(X), \ldots, p_r(X) \in F[X]$$

as words over the domains

$$D_0 \supseteq D_1 \supseteq \ldots \supseteq D_r,$$

respectively, wheras their degree bounds k_i, $\deg p_i(X) < k_i$, decrease with the same ratio as the domains. The quotients

$$a_i = \frac{k_{i-1}}{k_i} = \frac{|D_{i-1}|}{|D_i|}$$

are the *reduction factors*, and we throughout assume that $a_i \geq 2$. (By our assumptions on $|D|$ and k the a_i are again powers of two.) The number of rounds $r \geq 1$, their reduction factors $a_1, \ldots, a_r$ and therefore the decreasing sequence of domains $D_0, \ldots, D_r$, are parameters of FRI.

Protocol 1 (FRI commit phase)**.** *Given the domain evaluation* $[p_0(x)|_{x \in D_0}]$ *for the polynomial* $p_0(X) \in F[X]$, $\deg p_0(X) < k_0$*, the commit phase consists of the following* r *rounds.*

- *In each round* i, $1 \leq i \leq r$*, the prover decomposes the previous polynomial* $p_{i-1}(X)$ *of* $\deg p_{i-1}(X) < k_{i-1}$*, according to*

$$p_{i-1}(X) = F_0(X^{a_i}) + X \cdot F_1(X^{a_i}) + \ldots + X^{a_{i-1}} \cdot F_{a_i - 1}(X^{a_i}), \tag{3}$$

  where each

$$\deg F_i(Y) < \frac{k_{i-1}}{a_i} = k_i.$$

  (For $a_i = 2$ *this is the decomposition into odd and even parts.) The verifier samples a random challenge* $\lambda_i \leftarrow\!\!\$\, F$*, sends it to the prover, which in turn responds with the linear combination*

$$p_i(Y) = F_0(Y) + \lambda_i \cdot F_1(Y) + \ldots + \lambda_i^{a_{i-1}} \cdot F_{a_i - 1}(Y)$$

  as a word on the reduced domain $D_i = D_{i-1}^{a_i} = \{x^{a_i} : x \in D_{i-1}\}$*. That is, it sends* $[p_i(y)|_{y \in D_i}]$ *to the verifier. In the last step however,* $i = r$*, the polynomial* $p_r(X) \in F[X]$ *is revealed in full length instead.*

Let us elaborate on the decomposition (3) in terms of the reduction map

$$\pi_i : D_{i-1} \longrightarrow D_i, \quad x \mapsto x^{a_i}.$$

Notice that for each y in D_i, $y = x^{a_i}$, the values of $F_0(y),\ldots,\ F_{a_i-1}(y)$ are uniquely determined by the values of

$$F_0(y) + F_1(y) \cdot X + ... + F_{a_i-1}(y) \cdot X^{a_i-1}$$

on the coset $\pi_i^{-1}(y) = x \cdot \ker(\pi_i)$, and these values are exactly the ones given by $p_{i-1}(X)$. Hence if τ is a generator of $\ker(\pi_i) = \{1, \tau, \ldots, \tau^{a_i-1}\}$, then

$$\begin{aligned} p_i(\pi_i(x)) = & L_0\left(p_{i-1}\left(\tau^0 \cdot x\right), \ldots, p_{i-1}\left(\tau^{a_i-1} \cdot x\right)\right) \\ & + \lambda_i \cdot L_1\left(p_{i-1}\left(\tau^0 \cdot x\right), \ldots, p_{i-1}\left(\tau^{a_i-1} \cdot x\right)\right) \\ & + \ldots \\ & + \lambda_i^{a_i-1} \cdot L_{a_i-1}\left(p_{i-1}\left(\tau^0 \cdot x\right), \ldots, p_{i-1}\left(\tau^{a_i-1} \cdot x\right)\right) \end{aligned}$$

where $(L_0, ..., L_{a_i-1})$ is the Lagrange interpolation map for the coset $x \cdot ker(\pi_i)$. In other words,

$$p_i(\pi_i(x)) = \mathsf{FFT}_{\lambda_i/x}\left(p_{i-1}\left(\tau^0 \cdot x\right), \ldots, p_{i-1}\left(\tau^{a_i-1} \cdot x\right)\right), \tag{4}$$

that is the Fourier transform of the vector $\left(p_{i-1}\left(\tau^0 \cdot x\right), \ldots, p_{i-1}\left(\tau^{a_i-1} \cdot x\right)\right)$, evaluated at $\frac{\lambda_i}{x}$. This equation will be used to check consistency between the provided oracles.

In some situations it is more efficient to compute the values of $p_i(y)$ over D_i directly from the ones of $p_{i-1}(x)$, $x \in D_{i-1}$, using (4). In terms of field additions A, multiplications M, and FFT operations $\mathsf{FFT}(a_i)$ of size a_i, the cost of this would be

$$|D_i| \cdot \left((a_i - 1)\ \mathsf{A} + (2a_i - 3)\ \mathsf{M} + \mathsf{FFT}(a_i)\right) \approx |D_i| \cdot a_i \cdot (2 + \log_2(a_i))\ \mathsf{M},$$

compared to

$$a_i \cdot k_i\ (\mathsf{M} + \mathsf{A}) + \mathsf{FFT}(|D_i|) \approx |D_i| \cdot (1 + \log_2 |D_i|)\ \mathsf{M}$$

when computing the domain evaluation of the random linear combination $p_i(X)$. Hence using equation (4) is more efficient whenever

$$a_i \cdot (2 + \log_2(a_i)) < 1 + \log_2 |D_i|, \tag{5}$$

which holds for most domain sizes when $a_i = 2$. Already the closest larger choice $a_i = 2^2$ leads to $|D_i| > 2^{15}$, and hence no improvement in computations.

Sampling phase

In the query phase the verifier samples at random points from the defining domains of the oracles, and use the returned values to check the consistency of all reduction steps.

Protocol 2 (FRI query phase)**.** *The query phase consists of $s \geq 1$ many rounds.*

- *In each round the verifier samples an $x_0 \in D_0$ uniformly at random, computes $x_1, \ldots, x_r$ recursively via $x_i = \pi_i(x_{i-1})$, and checks if*

$$p_i(x_i) = \mathsf{FFT}_{\lambda_i/x_i}\left(p_{i-1}(x_{i-1}), p_{i-1}(\tau \cdot x_{i-1}), \ldots, p_{i-1}(\tau^{a_{i-1}} \cdot x_{i-1})\right),$$

  for every $i = 1, \ldots, r$, by querying the values of each p_{i-1} over the coset $x_{i-1} \cdot \ker \pi_i$.

Notice that unlike in [BSCI+20] we choose x_0 uniformly from D_0, and form the x_i by projecting x_{i-1} onto D_i. In distribution, this way of sampling is equivalent to the one in the paper, which starts with $x_r \leftarrow\!\!\$\ D_r$, and then samples x_{i-1} uniformly from the coset $\pi_i^{-1}(x_i)$.

Batching

As for linear polynomial commitment schemes, batching is done via random linear combinations. We will only discuss the algebraic variant, which uses powers of a single random challenge. (Again, this is the one favored in the context of proof composition.)

Given a batch of L low-degree polynomials $q_0(X), \ldots, q_{L-1}(X)$, the verifier samples a random challenge $\lambda \leftarrow_\$ F$. The prover computes the linear combination

$$h(X) = \sum_{i=0}^{L-1} \lambda^i \cdot q_i(X), \tag{6}$$

sends the oracle of it,

$$[h(x)|_{x \in D}],$$

to the verifier. Then both prover and verifier continue with FRI for h. Each $x_0 \leftarrow_\$ D_0 = D$ from the query phase of FRI is used to additionally check consistency between the oracle for $h(X)$ and the ones in the batch, $q_0(X), \ldots, q_{L-1}(X)$, using (6).

Soundness

The soundness analysis of FRI is based on a strengthening of the correlated agreement theorem, which allows to additionally keep track of the success probability for the FRI query phase by a sub-probability measure μ. We state that *weighted correlated agreement theorem* in Appendix A.3. For proximity parameters close to the Johnson bound, the soundness error of the batched FRI oracle proof is as follows:

Theorem 2 (Batched FRI soundness error, [BSCI$^+$20], Theorem 3.8)**.** *Suppose that $q_i \in F^D$, $i = 0, \ldots, L-1$, is a batch of functions given by their domain evaluation oracles. If an adversary passes batched FRI for $\mathsf{RS}_k[F, D]$ and proximity parameter $\theta = 1 - \sqrt{\rho} \cdot \left(1 + \frac{1}{2m}\right)$, $m \geq 3$, with a probability larger than*

$$\varepsilon = \left(L - \frac{1}{2}\right) \cdot \frac{\left(m + \frac{1}{2}\right)^7}{2 \cdot \sqrt{\rho}^3} \cdot \frac{|D_0|^2}{|F|} + \frac{(2m+1) \cdot (|D_0| + 1) \cdot \sum_{i=1}^{r} a_i}{|F|} + (1-\theta)^s, \tag{7}$$

then the functions $q_i \in F^D$, $i = 0, \ldots, L-1$, have correlated agreement with $RS_k[D, F]$ on a set of density of at least $\alpha > \left(1 + \frac{1}{2m}\right) \cdot \sqrt{\rho}$. (We notice that, using affine batching then $L - \frac{1}{2}$ is replaced by $\frac{3}{2}$, see [BSCI$^+$ 20].)

The first two terms in (7),

$$\varepsilon_C = \left(L - \frac{1}{2}\right) \cdot \frac{\left(m + \frac{1}{2}\right)^7}{2 \cdot \sqrt{\rho}^3} \cdot \frac{|D_0|^2}{|F|} + \frac{(2m+1) \cdot (|D_0| + 1) \cdot \sum_{i=1}^{r} a_i}{|F|},$$

correspond to soundness error of the commit phase, reflecting the systematic error estimated by the correlated agreement theorem and collected over the batching step and the reduction rounds. In words, if the oracles in the batch do not share the claimed correlated agreement for $\alpha = 1 - \theta$, then except with probability ε_C, the oracles produced during the commit phase cannot be "nice". That is, the set where all consistency checks would hold is *at most* of density α. The remaining term,

$$\varepsilon_Q = (1-\theta)^s,$$

is the soundness error of the query phase with s rounds. This is the probability not to detect such a set of non-"nice" oracles using s independent samples.

Example parameters

One way to settle the parameters is as follows. For target security level $2^{-\lambda}$, we assure that

1. the soundness error for the commit phase is bounded by $\frac{1}{2} \cdot 2^{-\lambda}$. For that we choose the maximum Johnson proximity $m \geq 3$ so that
$$\varepsilon_C \leq \frac{1}{2} \cdot 2^{-\lambda},$$
2. the soundness error of the query phase is bounded by $\frac{1}{2} \cdot 2^{-\lambda}$. Using m from the first step, we determine the number s of query rounds via
$$\varepsilon_Q = \sqrt{\rho}^s \cdot \left(1 + \frac{1}{2m}\right)^s \leq \frac{1}{2} \cdot 2^{-\lambda}.$$

74 bits of security

Such a configuration is interesting in practice, as its security can be increased by *grinding* (see [Sta21]): Another 16 bits proof of work bound to the proof generation, and one obtains overall 90 bits of security.

For example, let $|F| = 2^{64}$, $|D_0| = 2^{12} \cdot \rho^{-1}$, $L \approx 300$, and we assume that the polynomials are grouped into
$$\{100, 100, 100\}$$
polynomials, each committed by a single tree using Merkle caps. This situation is similar to the one in plonky2. The field is definitely too small to achieve reasonable security, hence one works with field extensions.

- With a degree 2 extension of F, hence a field size of 128 bits, the best security level one can obtain for $\rho = 2^{-5}$ is about 74 bits. The commit phase error is
$$\varepsilon_C \approx 2^{-74.84},$$
with Johnson proximity $m = 3$. To have about the same soundness error in the query phase, we demand $s = 33$ samples, yielding
$$\varepsilon_Q \approx 2^{-75.16}.$$
With a reduction strategy $\{a_1, a_2\} = \{2^4, 2^3\}$ we obtain proof sizes of about 118 kB.

$-\log_2(\rho)$	m	s	T in hashes	$\lvert\pi\rvert$ / kB
3	7	51	17,408 H	174 kB
4	6	39	34,800 H	136 kB
5	3	33	69,632 H	118 kB

- With a degree 3 extension of F, hence a field size of 192 bits, one can choose higher blow-up factors. For $\rho = 2^{-6}$ we obtain 74 bits security by
$$\varepsilon_C \approx 2^{-74.00},$$
where the Johnson proximity is $m = 1,521$. To have about the same soundness error in the query phase, we need only $s = 25$ samples, yielding
$$\varepsilon_Q \approx 2^{-74.98}.$$
With the same reduction strategy as before, we reduce the proof size down to 97 kB. However, this comes at the cost of about tripling the prover cost.

$-\log_2(\rho)$	m	s	T in hashes	$\lvert\pi\rvert$ / kB
6	1,521	25	208,896 H	97 kB
8	760	19	835,584 H	76.4 kB
10	379	15	3,342,336 H	62.5 kB

112 bits security

As in the previous setting, we discuss this level of security as it can be improved by grinding, typically up to 128 bits. All configurations use degree 3 extensions of F.

$-\log_2(\rho)$	m	s	T in hashes	$\lvert\pi\rvert$ / kB
6	34	25	208,896 H	145 kB
8	17	19	835,584 H	115.4 kB
10	379	15	3,342,336 H	95 kB

128 bits security

These configurations do not use grinding. Again, they use degree 3 extensions of F.

$-\log_2(\rho)$	m	s	T in hashes	$\lvert\pi\rvert$ / kB
6	6	45	208,896 H	171 kB
8	17	19	835,584 H	135.4 kB
10	–	–	–	–

Conjectured security

In their line of work on FRI [BSBHR18a, BSGKS20, BSCI$^+$20] the authors make several conjectures on the soundness of FRI for proximity parameters above the Johnson bound. In the most recent one, they state the following.

Conjecture 1 (Full version of [BSCI$^+$20], Conjecture 8.4)**.** *There exist constants c_1, c_2 such that for all $\theta = 1 - \rho - \eta$, $\eta > 0$, the soundness error in the correlated agreement theorem on $f_0, \ldots, f_{N-1}$ is bounded by*

$$\varepsilon \leq \frac{1}{(\eta \cdot \rho)^{c_1}} \cdot \frac{(N \cdot n)^{c_2}}{|F|}.$$

Remark 3. For purely linear batching, a similar conjecture is stated.

We point out that the above conjecture (as well as its corresponding one in [BSBHR18a]) is stated isolated from any general conjectured properties on Reed-Solomon codes, such as list decodability up to capacity bound (as done for DEEP method, see Section 5.4). Instead it is rather justified by "*[to the best of our knowledge...] nothing seems to contradict*" . The authors consider the choice of $c_1 = c_2 = 2$ reasonable, and for fields of characteristic $q > n$ they estimate that $c_1 = c_2 = 1$.

The $c_1 = c_2 = 1$ assumption is of particular interest for practitioners, as it yields proofs of halve the size as in the $c_1 = c_2 = 2$ case. For example, it is used by the ethSTARK [Sta21] (besides its provably secure parameter setting), as well as by plonky2 [Pol].

Adding zero-knowledge

Zero-knowledge for FRI has to be provided on application level. In our use cases, the witnesses of an argument correspond to the values of some polynomial $q(X)$ on a given domain H (the proving domain for Plonk, say). To protect it from being leaked by the queries of the s query rounds (as well as by the final reduction polynomial), one uses a an H-disjoint *coset* $a \cdot D$ of the FRI domain, and randomizes $q(X)$ outside the domain H. That is, the batching and the entire FRI reduction takes place on

$$a \cdot D_0 \supseteq a \cdot D_1 \supseteq \ldots \supseteq a \cdot D_r,$$

instead of $D_0 \supseteq D_1 \supseteq \ldots \subseteq D_r$, where $(a \cdot D_0) \cap H = \emptyset$. This leads to running batched FRI for $q_i(X)$, $i = 0, \ldots, L-1$, over the non-shifted domain D_0 on the shifted polynomials

$$q_i(a \cdot X),$$

$i = 0, \ldots, L-1$, instead.

The number of linear functionals of a polynomial $q_i(X)$ revealed in the course of a single FRI query are: One in the batching step, a_i many for the coset evaluations in each of the reduction steps $i = 1, .., r-1$, and

$$1 + \deg p_r(X) = \rho \cdot |D_r| = \rho \cdot \frac{|D_0|}{a_1 \cdot a_2 \cdot \ldots \cdot a_{r-1}}$$

linear functionals corresponding to the coefficients of the final reduction polynomial. With s queries this leads to overall

$$b = s \cdot (1 + \sum_{i=1}^{r-1} a_i + \rho \cdot \frac{|D_0|}{a_1 \cdot a_2 \cdot \ldots \cdot a_{r-1}}) \tag{8}$$

linear functionals. To reduce this number, one can add a blinding polynomial

$$h(x) \in \mathsf{RS}_k[F, D]$$

to the batch (coming with the cost of an extra commitment). Then the number of linear functionals revealed on a witness polynomial is reduced to $b = s$.

In both cases, the randomization can be done without moving beyond $|H| - 1$ in degree whenever a subset $B \subseteq H$ with $|B| = b$ remains "unused", i.e. unconstrained: Instead of taking

$$p(X) = p(X) + r(X) \cdot v_H(X),$$

where $v_H(X)$ is the vanishing polynomial of H and $r(X)$ a random polynomial of $\deg r(X) = b - 1$, one takes $p(X)$ as the polynomial interpolated from the witness values on $H \setminus B$ and randomly chosen values on B.

FRI as a polynomial commitment scheme

FRI can be turned into a polynomial commitment scheme by means of the evaluation quotients

$$h(x) = \frac{f(x) - v}{x - z}$$

of a committed word $f \in F^D$. This approach, called the DEEP method in [BSGKS20] corresponds to the algebraic linking of the evaluation identity

$$f(X) = v + h(x) \cdot (X - z)$$

with a low-degree problem on the sampling domain D, assuming that $z \notin D$. (For $z \in D$ the oracle can directly answer with the queried value. We will omit this case throughout our discussion.)

For proximity parameters θ up to the unique decoding radius one obtains a polynomial commitment scheme in the classical sense (when compiling the oracle proof into an argument using a secure partially disclosable vector commitment). In the list decoding regime the situation is a bit more subtle due to the non-uniqueness of θ-close code words. In this case the DEEP method can be viewed as an oracle proof for a more general type of polynomial commitment scheme, called *list polynomial commitment scheme* in [KPV19]. However, as their notion does not cover the power of correlated agreement, we shall only sketch list polynomial commitment schemes.

In the unique decoding regime

For a proximity bound up to the unique decoding radius, i.e. $\theta < \frac{1-\rho}{2}$, the situation is quite simple. However, there are several ways to algebraically link the evaluation identity with a low-degree test.

A first construction

We first discuss a naive scheme, in which the maximum degree corresponds to the degree proven by FRI.

- *Setup:* The maximum degree $d = k - 1$ is chosen as the maximum degree of polynomials belonging to $\mathsf{RS}_k[F, D]$.
- *Commit:* Given a polynomial $p(X)$ of degree $\deg p(X) \leq d$, the prover commits its domain evaluation over D, i.e.
$$\mathsf{Com}(p(X)) = [p(x)|_{x \in D}].$$
- *Evaluation proof:* Given an opening claim (z, v) with $z \notin D$, the prover engages with the verifier in a batched FRI argument on
$$f_1(x) = \frac{p(x) - v}{x - z},$$
$$f_2(x) = x \cdot f_1(x) = x \cdot \frac{p(x) - v}{x - z}.$$
with proximity bound $\theta = \frac{1-\rho}{2}$. This proof batches the functions into a random linear combination $f_1(x) + \lambda \cdot f_2(x) = (1 + \lambda \cdot x) \cdot \frac{p(x)-v}{x-z}$, and then runs FRI on it. The linear term $\lambda \cdot x$ is called *degree correction factor*.

We point out that the two functions f_1, f_2 are not needed to be provided by another oracle, as their evaluations on D can be computed from the values of $p(x)$.

Let us discuss that the evaluation proofs in fact provide a view on a unique polynomial of degree $\leq d$, determined by the values committed in $[p(x)|_{x \in D}]$. First of all, if the prover passes with a probability p greater than the soundness error of batched FRI on f_1, f_2 as above, then there exist two polynomials $p_1(X)$, $p_2(X)$ of degree $\leq d$, and a correlated agreement set A of density $1 - \theta \geq \frac{1+\rho}{2}$ such that
$$f_1(x) = p_1(x)\big|_{x \in A},$$
$$x \cdot f_1(x) = p_2(x)\big|_{x \in A},$$

and hence also $x \cdot p_1(x) = p_2(x)\big|_{x \in A}$. As the density of A is strictly greater than ρ, the polynomial $X \cdot p_1(X) - p_2(X)$ has at least $k+1 = d+2$ zeroes and hence must be trivial, i.e. $X \cdot p_1(X) = p_2(X)$. This implies that $\deg p_1(X) \leq d-1$, and hence $p(x)$ coincides on A with the degree d polynomial

$$P(X) = v + (X - z) \cdot p_1(X),$$

which evaluates to v at z. Notice that $\delta(p(x), P(X)) < \frac{1-\rho}{2}$, hence a single evaluation proof implies distance to a degree $\leq d$ polynomial of at most the unique decoding radius. As a consequence, any other evaluation proof (on the same or any other query) is consistent with that unique degree $\leq d$ polynomial, showing that we indeed have a polynomial commitment scheme.

The refined scheme

By similar reasoning (based on a degree $k = d+1$ polynomial vanishing on a set of density $> \rho$) we can remove the degree correction factor in the above naive scheme, running FRI for a proximity parameter $\theta < \frac{1-\rho}{2}$, only on the evaluation quotient of the claim: For any two evaluation claims (z_1, v_1) and (z_2, v_2) we conclude the existence of polynomials $p_1(X)$, $p_2(X)$ of degree $\leq k-1$ and sets A_1, A_2 of density $1-\theta > \frac{1+\rho}{2}$ such that

$$v_1 + (X - z_1) \cdot p_1(X),$$
$$v_2 + (X - z_2) \cdot p_2(X),$$

agree with $p(x)$ on A_1 and A_2, respectively. Since the density of $A_1 \cap A_2$ is at least $1 - 2 \cdot \theta > \rho$, it contains at least $k+1$ points, and by degree we may conclude the formal identity

$$v_1 + (X - z_1) \cdot p_1(X) = v_2 + (X - z_2) \cdot p_2(X).$$

This leads to the following optimized scheme:

- *Setup:* The maximum degree is $d^+ = k$, where k is the absolute rate of $\mathsf{RS}_k[F, D]$.
- *Commit:* Given a polynomial $p(X)$ of degree $\deg p(X) \leq d^+$, the prover commits its domain evaluation over D, i.e.
$$\mathsf{Com}(p(X)) = [p(x)|_{x \in D}].$$
- *Evaluation proof:* Given an opening claim (z, v) with $z \notin D$, the prover engages with the verifier in a batched FRI argument on
$$\frac{p(x) - v}{x - z}$$
with proximity bound $\theta < \frac{1-\rho}{2}$.

Multi-point queries

Instead of batching several point evaluation quotients, queries for the values of a polynomial $p(X)$ over a small set $\Omega = \{z_1, ..., z_m\} \subset F \setminus D$ can be also proven via the multi-evaluation identity

$$\sum_{i=1}^{m} (p(X) - v_i) \cdot L(z_i, X) = 0 \mod v_\Omega(X), \tag{9}$$

where $v_\Omega(X) = \prod_{j=1}^{m} (X - z_j)$ is the vanishing polynomial of Ω and $L(z_i, X) = \prod_{j \neq i} \frac{X - z_j}{z_i - z_j}$ is the Lagrange polynomial at z_i. Similar to the single query case, one argues using the quotient

$$h(x) = \mathsf{Quotient}(p, \{(z_i, v_i) : i = 1, \ldots, m\}) = \frac{p(x) - V(x)}{v_\Omega(x)}, \tag{10}$$

where

$$V(X) = \sum_{i=1}^{m} v_i \cdot L(z_i, X)$$

is the unique degree $\leq m-1$ polynomial that interpolates the claim.

Alternatively, as in the batch evaluation protocol of Boneh, et al. [BDFG21], one can replace the Lagrange kernel with the non-normalized variant $D(z_i, X) = \prod_{j \neq i}(X - z_j)$

$$\sum_{i=1}^{m}(p(X) - v_i) \cdot D(z_i, X) = 0 \mod v_\Omega(X), \tag{11}$$

and work with the quotient

$$h'(x) = \sum_{i=1}^{m} \frac{p(x) - v_i}{x - z_i}$$

instead.

In both cases one has to limit the number m of simultaneous queries to some maximum value m_{max}, satisfying

$$k + m_{max} < (1 - \theta) \cdot n.$$

For this it is sufficient to choose $k + m_{max} \leq (1 - \theta_0) \cdot n = \frac{k+n}{2}$, and hence $m_{max} \leq \frac{n-k}{2}$. Even with the lowest blow-up factor we have $n \geq 2 \cdot k$, it is thus enough to demand

$$m_{max} \leq \frac{k}{2}. \tag{12}$$

In our applications the bound on m_{max} is trivially met, as only few values are queried in the run of the proof. Furthermore, given a polynomial we use multi-point queries of fixed given size $m \leq m_{max}$. As a consequence the maximum degree in the setup can be enlarged to $d_{max} = k + m - 1$.

List commitments

In the list decoding regime the situation is a bit more subtle. Running FRI for $\mathsf{RS}_k[F, D]$ with a proximity parameter $\frac{1-\rho}{2} < \theta < 1 - \sqrt{\rho}$ on an evaluation quotient

$$h(x) = \frac{p(x) - v}{x - z}$$

only proves agreement of p with an evaluation-claim-consistent polynomial of degree $d^+ = k$ on a set of density greater than $\alpha = 1 - \theta$. This might be not large enough for proving the polynomials of different runs of FRI being equal. In fact, they might differ from claim to claim, unless one runs a joint FRI argument on them. Assuming $\alpha > \sqrt{\rho^+}$, where $\rho^+ = \frac{k+1}{|D|}$, the Guruswami-Sudan list decoding bound shows that there might be

$$L \leq \frac{1}{2 \cdot \eta \cdot \rho^+}$$

such code words. This leads to the idea of list polynomial commitment schemes as in [KPV19] with the following information-theoretic model: The prover sets up an oracle which contains a list of l, $1 \leq l \leq L$, low-degree polynomials, and the oracle is allowed to choose which one to evaluate on a given query. Such extended notion is practical as security proofs in the oracle model are similar to polynomial oracle proofs. However, the notion of list polynomial oracles as given in [KPV19] is not strong enough to capture correlated agreement, and as a consequence soundness error bounds are too coarse. For this reason we do not dive into formal details of that model, and instead directly work with DEEP algebraic linking.

DEEP-ALI

In this section we discuss the *DEEP algebraic linking (DEEP-ALI)* [BSGKS20] and demonstrate its application to proving satisfiability of algebraic intermediate representations (AIR). Other representations such as randomized AIR or Plonk [GWC19] can be treated similarly.

Algebraic linking and the DEEP method

Algebraic linking transforms satisfiability of algebraic identities over algebraic subsets of F into proximity problems of low-degree extensions to Reed-Solomon codes over "outside" domains (i.e. disjoint to the algebraic subset) . A family of functions $g_1, \ldots, g_N$ on $\Omega = \{x_1, ..., x_n\}$ satisfies an algebraic identity

$$P(x, g_1(x), \ldots, g_N(x)) = 0$$

on Ω (P is a polynomial), if and only if their low-degree extensions $p_1(X)$, ..., $p_N(X)$ satisfy that $P(X, p_1(X), \ldots, p_N(X))$ is divisible by the vanishing polynomial $v_\Omega(X) = \prod_{i=1}^{n}(X - x_i)$ of Ω , i.e. the quotient

$$h(X) = \frac{P(X, p_1(X), ..., p_N(X))}{v_\Omega(X)}$$

is a low-degree polynomial. This divisibility criterion is translated to the proximity of given code words

$$f_1, ..., f_N, h \in F^D,$$

(the honest prover chooses the domain evaluations of $p_1(X), \ldots, p_N(X)$ and $h(X)$ over D) to low-degree polynomials, i.e. a Reed-Solomon code words[1]. For this the proximity parameter needs to be chosen so that the agreement sets are large enough to infer from local satisfiability of algebraic identities to their satisfiability over the entire field F. This means that the sampling domain D is such that the notion of low-degree is determined by the degree of $P(X, p_1(X), ..., p_N(X))$. DEEP-ALI instead allows for decoupling the sampling domain size from the degree of P.

DEEP-ALI is very much in alignment with a polynomial IOP for proving that

$$P(X, p_1(X), ..., p_N(X)) = h(X) \cdot v_\Omega(X). \tag{13}$$

Instead of showing proximity of the quotient

$$h(x) = \frac{P(x, f_1(x), ..., f_N(x))}{v_\Omega(x)}$$

to a low-degree polynomial, one samples a random point $z \leftarrow\!\!\$\, F$ outside the domain D, and let the prover provide evaluations claims v_i, $i = 1, \ldots, w$ for p_i, and v for h, which are used to check the identity (13) at $X = z$. The validity of the values are supported by proving proximity of the point evaluation quotients

$$\frac{f_i(x) - v_i}{x - z}, \quad i = 1, \ldots, w,$$

as well as

$$\frac{h(x) - v}{x - z}$$

[1] In the case of a single batched FRI proof for the f_i together with h, one needs to use degree correction factors as in Section 4.1.1.

to corresponding low-degree polynomials. Furthermore, by decomposing $h(X)$ into into polynomials of degree $|\Omega| - 1$, e.g.

$$h(X) = h_0(X) + X^{|\Omega|} \cdot h_1(X) + \ldots + X^{(d-1)\cdot|\Omega|} \cdot h_{d-1}(X), \tag{14}$$

one can even use a sampling domain the size of which is not determined by the degree of $h(X)$. (We use a different decomposition as in [BSGKS20, Sta21], which does not imply any further constraints on the sampling space for z.)

In the unique decoding regime, the DEEP-ALI approach is equivalent to a (univariate) polynomial IOP using FRI as a polynomial commitment scheme as described in Chapter 4. For larger proximity parameters, one can generalize the polynomial oracle model to list polynomial commitment schemes as done in [KPV19], but their approach does not yield soundness bounds which are as tight as given by the correlated agreement theorem. In order not to introduce yet another oracle model which reflects this specific correlated agreement property of batched FRI, we directly show how to apply the DEEP method to proving satisfiability of an algebraic intermediate representation.

DEEP-ALI of an AIR

An *algebraic intermediate representation (AIR)*, see [BSBHR18b, BSGKS20, Sta21], is defined over an FFT domain $H \subset F$ with generator g. Each x in H carries a "row" of w witnesses (or, "columns")

$$(g_1(x), .., g_w(x)),$$

on which a certain number of algebraic constraints are imposed. For simplicity we restrict ourselves to constraints between neigboring rows only, i.e. polynomials

$$P_1, \ldots, P_C \in F[X_1, ..., X_w, Y_1, ..., Y_w],$$

each P_i being imposed on a specified coset $a_i \cdot H_i \subseteq H$, where H_i is a subgroup of H. Hence satisfiability of the AIR is defined by

$$P_i(x, g_1(x), \ldots, g_w(x), g_1(g \cdot x), \ldots, g_w(g \cdot x)) = 0 \quad \forall x \in a_i \cdot H_i, \tag{15}$$

for every $i = 1, \ldots, C$. In terms of polynomials $p_1(X), \ldots, p_w(X) \in F[X]$ extending the witness functions $g_1, \ldots, g_w$, satisfiability of an AIR constraint P_i on $a_i \cdot H_i$ can be expressed by demanding the quotient

$$\frac{P_i(p_1(X), \ldots, p_w(X), p_1(g \cdot X), \ldots, p_1(g \cdot X))}{v_{a_i \cdot H_i}(X)}$$

where $v_{a_i \cdot H_i}(X) = z^{|H_i|} - a_i^{|H_i|}$ is the vanishing polynomial of the coset $a_i \cdot H_i$, being again a polynomial. This is the approach [BSBHR18b, BSGKS20, Sta21]. However, instead of working with these quotients we prefer using polynomial identities similar to Plonk [GWC19]: Satisfiability of an AIR constraint P_i imposed on $a_i \cdot H_i$ is equvialent to

$$s_i(X) \cdot P_i(p_1(X), \ldots, p_w(X), p_1(g \cdot X), \ldots, p_1(g \cdot X)) = 0 \bmod v_H(X), \tag{16}$$

where

$$s_i(X) = \frac{v_H(X)}{v_{a_i \cdot H_i}(X)} \in F[X] \tag{17}$$

is the *selector polynomial*[2]for the constraint P_i. The *overall degree* of the AIR is defined as

$$d = \max_i \deg(P_i), \tag{18}$$

[2]Notice that, although $\deg s_i(X) \leq |H| - 1$, the polynomial $s_i(X)$ can be succinctly evaluated outside H using the rational representation from (17). Therefore no evaluation has to be provided by the prover.

where $\deg(P_i)$ is the total degree of P_i.

The sampling domain D for FRI is chosen so that $|D| = \beta \cdot |H|$, with a blow-up factor $\beta = 1/\rho$ being a power of two, and $\mathsf{RS}_k[D, F]$ is the Reed-Solomon code of length $n = |D|$ and rate $\rho = \frac{k}{n}$, with

$$k = |H|. \tag{19}$$

However, the agreement parameter used for FRI is taken slightly larger than $\alpha = \left(1 + \frac{1}{2m}\right) \cdot \sqrt{\rho}$, $m \geq 3$, namely

$$\alpha^+ = \left(1 + \frac{1}{2m}\right) \cdot \sqrt{\rho^+}, \quad m \geq 3, \tag{20}$$

where

$$\rho^+ = \frac{|H| + 2}{|D|}. \tag{21}$$

The reason for this slightly larger choice is due to the evaluation quotients of the protocol, which are subject to the FRI proof. Their denominators are at most quadratic and hence the degree of the non-quotients is bounded by $|H| - 1 + 2$. The low-degree extensions $p_i(X) \in F[X]$ of the witness functions g_i on H are provided as code words over D, and to use again the same code for a polynomial $h(X)$ of larger degree, we split it into segment polynomials as in (14).

The DEEP-ALI protocol (for simplicity without zero-knowledge) for our AIR is as follows:

Protocol 3 (IOP for AIR using DEEP-ALI)**.** *Let $p_1(X), \ldots, p_w(X) \in F[X]$ be polynomials of degree $\deg p_i(X) \leq |H| - 1$ satisfying the AIR constraints (15), $i = 1, \ldots, C$.*

1. *The prover sends the domain evaluation oracles $[p_1], \ldots, [p_w]$ for $p_1(X)$, …, $p_w(X)$ to the verifier, who responds with a randomness $\lambda \leftarrow\!\!\$\, F$.*

2. *The prover computes $h_\lambda(X) \in F[X]$ of degree $\leq d \cdot (|H| - 1)$ satisfying the identity*

$$\sum_{i=1}^{C} \lambda^{i-1} \cdot s_i(X) \cdot P_i(p_1(X), \ldots, p_w(X), p_1(gX), \ldots, p_w(gX)) = h_\lambda(X) \cdot v_H(X),$$

splits it into its segment polynomials $h_{\lambda,j}(X)$, $j = 0, \ldots, d-1$, each of degree $\leq |H| - 1$, as in (14), and sends their domain evaluation oracles $[h_{\lambda,0}]$, …, $[h_{\lambda,d-1}]$ to the verifier. The overall identity to be proven is therefore

$$\begin{aligned} \sum_{i=1}^{C} \lambda^{i-1} \cdot s_i(X) \cdot P_i(p_1(X), \ldots, p_w(X), p_1(gX), \ldots, p_w(gX)) \\ = v_H(X) \cdot \sum_{j=0}^{d-1} X^{j \cdot |H|} \cdot h_{\lambda,j}(X). \end{aligned} \tag{22}$$

The verifier answers with a DEEP query, i.e. a random $z \leftarrow\!\!\$\, F \setminus (D \cup H)$.

3. *Upon receiving the DEEP query z, the prover sends the evaluation claims $(z, v_{i,1})$, $(g \cdot z, v_{i,2})$, $i = 1, ..., w$, for the witness polynomials $p_i(X)$, and (z, v_j), $j = 0, ..., d-1$, for the segment polynomials $h_{\lambda,j}(X)$, to the verifier.*

4. *Eventually, prover and verifier run batched FRI for proximity of the evaluation quotients*

$$\frac{p_i(x) - V_i(x)}{(x - z) \cdot (x - gz)},$$

where $V_i(x)$ is determined from the evaluation claims as described in Section 4.1.3, $i = 1, \ldots, w$, and

$$\frac{h_{\lambda,j}(x) - v_j}{x - z},$$

$j = 0, .., d-1$, to $RS_k[F, D]$, where the chosen agreement parameter is α^+ as defined above. If FRI passes, and if the evaluation claims satisfy the overall identity (22) at $X = z$, the verifier accepts. (Otherwise, it rejects.)

Remark 4. Notice that the polynomial $s_i(X)$ can only be succinctly evaluated outside H. For this reason that H is excluded from the samping space of z.

Remark 5. As discussed above, our definition of AIR is equivalent to the one from [BSBHR18b, BSGKS20, Sta21] (besides that we restricted to constraints between neighboring rows in order to keep the presentation simple). In particular the quotient polynomial $h_\lambda(X)$ in our protocol is the same as

$$\sum_{i=1}^{C} \lambda^{i-1} \cdot \frac{s_i(X)}{v_H(X)} \cdot P_i(p_1(X), \ldots, p_w(X), p_1(gX), \ldots, p_w(gX)) = \sum_{i=1}^{C} \lambda^{i-1} \cdot \frac{P_i(p_1(X), \ldots, p_w(X), p_1(gX), \ldots, p_w(gX))}{v_{a_i \cdot H_i}(X)},$$

which is the batched rational function used in their line of work.

Remark 6. Let us point us the difference of Protocol 3 to the IOP given in [Sta21]. Instead of using a purely linear batching strategy, we use the powers of a single randomness λ, which is the favoured choice in the context of proof composition. Secondly, as in [BSGKS20] we use multi-point quotients for the witness polynomials which are queried at z and gz. This reduces the number of polynomials on which FRI is applied, at the cost of only a slight increase in the choice of k^+. Thirdly, the way we decompose $h_\lambda(X)$ into segment polynomials (14) does not further reduce the sampling space for z, as is needed when using a FRI-like decomposition.

We finally state the soundness error of Protocol 3 in the oracle model.

Theorem 7 (DEEP-ALI soundness)**.** *The above oracle proof for AIR satisfiability has soundness error*

$$\varepsilon \leq L^+ \cdot \left(\frac{C}{|F|} + \frac{d \cdot (k^+ - 1) + (k - 1)}{|F| - |D \cup H|} \right) + \varepsilon_{FRI}, \tag{23}$$

with $k^+ = k + 2$, $L^+ = \frac{m + \frac{1}{2}}{\sqrt{\rho^+}}$, $\rho^+ = \frac{k^+}{n}$, and ε_{FRI} being the soundness error for batched FRI for α^+-agreement with $RS_k[F, D]$, Theorem 2.

Remark 8. We point out some differences to the error bound in [Sta21], Theorem 4. In our bound the list size bound L^+ only occurs linearly instead of quadratically. This due to our more careful analysis of the consequences of the correlated agreement enforced on polynomials produced in different rounds of the protocol. Secondly, as mentioned above, the alternative decomposition of $h_\lambda(X)$ into segment polynomials does not reduce the sampling space for z by a factor d larger domain. Less importantly, since we use do algebraic batching using the powers of λ, the first term incorporates the number of constraints C. A purley linear batching strategy, as used in [Sta21] leads to $\frac{1}{|F|}$ instead.

Remark 9. In the soundness error formula in [BSGKS20], Theorem 15 , the list bound L^+ occurs quadratically. This is due to the application of two separate FRI arguments, one for the batched quotients of

the witness polynomials, and another one for the overall quotient polynomial. (However, the splitting technique for h is outlined in Section 5.5. therein.) For the same reason, the notion of list polynomial commitment schemes from [KPV19] would lead to the w-th power of L^+, w being the number of witness columns. This might be acceptable for proving soundness of standard Plonk in the list polynomial oracle model, but not for a larger number of witness columns.

Proof of Theorem 7. Let us denote $\varepsilon_1 = L^+ \cdot \frac{C}{|F|}$, $\varepsilon_2 = L^+ \cdot \frac{d\cdot(k^+-1)+(k-1)}{|F|-|D\cup H|}$, and $\varepsilon_3 = \varepsilon_{FRI}$. Suppose that P^* is an adversary which succeeds the verifier with a probability exceeding $\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$. Then there exists a first message of P^*, i.e. words $f_1, \ldots, f_w$ on D, on which P^* succeeds with probability $> \varepsilon$, and hence

$$\Pr\left[\lambda : \Pr\left(P^* \text{ succeeds } |\lambda\right) > \varepsilon_2 + \varepsilon_3\right] > \varepsilon_1.$$

(Otherwise $\Pr[P^* \text{ succeeds }] \leq 1 \cdot \varepsilon_1 + (\varepsilon_2 + \varepsilon_3) \cdot (1 - \varepsilon_1) < \varepsilon_1 + \varepsilon_2 + \varepsilon_3$.) Likewise, for every such "good" λ (by the definition of ε_1, there are at least $L^+ \cdot C$ many) there exists a second message of P^*, i.e. words $h_{\lambda,0}, \ldots, h_{\lambda,d-1}$ on D such that

$$\Pr\left[z \in F \setminus D : \Pr(P^* \text{ succeeds } |z) > \varepsilon_3\right] > \varepsilon_2.$$

For each such "good" $z \in F\setminus(D\cup H)$ (by the definition of ε_2, there are more than $L^+\cdot(d \cdot (k^+ - 1) + (k - 1))$ many) the evaluation claims pass the verifier checks, and moreover the soundness of FRI enforces the evaluation quotients

$$\left(\frac{f_1(x) - V_1(x)}{(x - z) \cdot (x - g \cdot z)}, \ldots, \frac{f_w(x) - V_w(x)}{(x - z) \cdot (x - g \cdot z)}, \frac{h_{\lambda,0}(x) - v_0}{x - z}, \ldots, \frac{h_{d-1}(x) - v_{d-1}}{x - z}\right)$$

to have correlated agreement with some $q_i(X) \in F[X]$, $i = 1, ..., w + d$, of degree $\deg q_i(X) \leq |H| - 1$ on a set A of density at least $\alpha^+ > \sqrt{\rho^+}$. Cancelling out the denominators, we see that

$$(f_1, \ldots, f_w, h_{\lambda,0}, \ldots, h_{\lambda,d-1})$$

have correlated agreement on a set of density $\geq \alpha^+$ with some element from $F[X]^{w+d}$ where each component polynomial is of degree $\leq |H| - 1 + 2 = k^+ - 1$, and satisfies the evaluation claim.

In what follows we shall call each element from $F[X]^l$, with component polynomials of degree $\leq k^+ - 1$, having correlated agreement with a vector of functions on a set of density $\geq \alpha^+$, an α^+-*configuration* for that vector of functions.

Let us keep a combination of "good" first and second messages $(f_1, \ldots, f_w)$, $(h_{\lambda,0}, \ldots, h_{\lambda,d-1})$ fixed. We have seen above that the existence of a single "good" z implies the existence of an α^+-configuration for $(f_1, \ldots, f_w, h_{\lambda,0}, \ldots, h_{\lambda,d-1})$. By the Guruswami-Sudan list size bound[3](see Appendix A.2) there are at most L^+ such α^+-configurations. However, since there are more than $L^+ \cdot \left(d \cdot (k^+ - 1) + (k - 1)\right)$ many "good" z, and each establishes an α^+-configuration which smoreover evaluates to the claimed values, we conclude from the pigeon-hole principle that there is at least one α^+-configuration,

$$(p_1, \ldots, p_w, q_{\lambda,0}, \ldots, q_{\lambda,d-1}) \in F[X]^{w+d},$$

for which the overall identity (22) (taking the $q_{\lambda,j}$ as $h_{\lambda,j}$ therein) holds at more than $d \cdot (k^+ - 1) + (k - 1)$ many z. By the degree of the identity, this configuration is a solution of it, hence $(p_1, \ldots, p_w) \in F[X]^w$ is an α^+-configuration for $(f_1, \ldots, f_w)$ which satisfies

$$\sum_{i=1}^{C} \lambda^{i-1} \cdot s_i(X) \cdot P_i(p_1(X), \ldots, p_w(X), p_1(gX), \ldots, p_w(gX)) = 0 \bmod v_H(X). \tag{24}$$

[3]By correlated agreement each of the component functions is $(1 - \alpha^+)$-close to $\mathsf{RS}_{k^+}[F, D]$.

Now let us keep a "good" first message $(f_1, \ldots, f_w)$ fixed. We have seen that for each "good" λ there exists an α^+-*configuration* for $(f_1, \ldots, f_w)$ which is a solution of (24). Again, by the Guruswami-Sudan list size bound, there can be at most L^+ many w-configurations. Since there are at least $L^+ \cdot C$ many "good" λ, we conclude again from the pigeon-hole principle that there is at least one α^+-configuration, which we again denote by $(p_1, \ldots, p_w)$, for which there are at least C many "good" λ for which (24) holds. By linear algebra (the Vandermonde matrix is invertible) we conclude that this configuration satisfies

$$s_i(X) \cdot P_i(X, p_1(X), \ldots, p_w(X), p_1(gX), \ldots, p_w(gX)) = 0 \bmod v_H(X)$$

for every $i = 1, \ldots, C$. The values of $(p_1, \ldots, p_w)$ over H satisfy the constraints the AIR. This completes the proof. □

We note that in Equation (23), the term in the brackets is exactly the soundness error bound of the protocol in the (univariate) polynomial IOP model [BFS20]. As soundness in this model is essentially based on the Schwartz-Zippel lemma, we believe that the blow-up by the factor L^+ holds in general for every (public coin) polynomial IOP when replacing polynomial oracles by domain-evaluation oracles. (At least for the polynomial IOPs we know, such as [GWC19, MBKM19, CHM+20] or [HGdB21], this is the case.) Such a general transformation of (univariate) polynomial IOPs into ordinary (i.e. domain-evaluation) IOPs would be of interest, as the polynomial IOP model is widely used by practicioners. The protocol design as well as its security analysis is much easier to understand in the polynomial oracle model, and their soundness error bounds could be easily taken over. We plan to elaborate on this in a separate document.

Extractability

We only provide a brief sketch how to build the extractor in the oracle model, given a prover P^* which succeeds with a probability of that exceeds the soundness error bound from Theorem 7:

1. Sample a "good" first message $[f_1], \ldots, [f_w]$ on which the prover succeeds with a probability greater than the soundness error bound from Theorem 7.

2. In this step we build a straight-line extractor from the "good" first message $[f_1], \ldots, [f_w]$ pbtained in Step (1): Read $f_1, \ldots, f_w$ from the oracles. By the proof of Theorem 7, $(f_1, \ldots, f_w)$ agrees with an AIR solution $(p_1(X), \ldots, p_w(X)) \in F[X]^w$ on a set A of density $\geq \alpha^+$. To obtain this solution, one repeatedly applies the Guruswami-Sudan list decoder[4]and "intersects" their outputs as described in [Sta21], Section 5.5. One of the resulting configurations must be the one that satisfies the AIR.

The first step takes expected time $O(1/\varepsilon)$, and the Gurswami-Sudan decoder consumes at most $O(|D|^{15})$ field operations, see Remark 13. To obtain strict polynomial running time, at the cost of having a success probability < 1, one may stop the sampling after an appropriate multiple of $1/\varepsilon$.

Boosting soundness

In this section we outline standard techniques to lower the DEEP-ALI soundness error for AIRs over small fields F. (See [Sta21], or [Pol].)

[4]Alternatively one could run the Guruswami-Sudan list decoder over $K = F(Z)$. However, its run-time analysis in the number of operations over F is probably more difficult.

Using extension fields

One simply draws queries (for example the DEEP queries and the FRI challenges) from a suitable large extension field F_e of F. The soundness error bound lowers accordingly, replacing $|F|$ with $|F_e|$. (Notice that the disadvantage of applying this approach to the entire protocol is that all FRI quotients have to be computed over F_e.)

Increasing the number of protocol challenges

Instead of drawing protocol challenges from an extension field, one may repeatedly sample a challenge and run the remaining protocol for them in parallel. For instance, the first verifier challenge λ can be sampled N_1 times, $\lambda_1, \ldots, \lambda_{N_1} \leftarrow\!\!\$\, F$, and prove the overall polynomial identity (22) for all of these cases. This yields a lowered soundness error bound of the first round,

$$\varepsilon_1 = \left(L^+ \cdot \frac{C}{|F|} \right)^{N_\lambda},$$

and increases only the number of $h_{\lambda,j}$ polynomials (by the factor N_λ) that are subject to the DEEP queries in the second round. Likewise, one may also take several DEEP queries $z_1, \ldots, z_{N_z}$ from $F \setminus D$, and apply FRI to the batch of all resulting quotients, lowering the soundness error bound of the second round to

$$\varepsilon_2 = \left(L^+ \cdot \frac{d \cdot (k^+ - 1) + (k-1)}{|F \setminus (D \cup H)|} \right)^{N_z}.$$

However, this comes at the cost of increasing the entire batch for FRI by the factor N_z (which might be acceptable in some applications, though). On the contrary, resampling of FRI challenges would increase the proof size too much. Hence for FRI extension field sampling is preferable.

Beyond the Johnson bound?

The conjectured soundness error for FRI alone (Conjecture 1) is not good enough to argue the security of DEEP-ALI beyond the Guruswami-Sudan list decoding bound. For that reason we also cite a general conjecture on the list decodability of Reed-Solomon codes, which is used by Ben-Sasson et al. to conjecture the soundness error of DEEP-FRI up to capacity bound.

Conjecture 2. *([BSGKS20], Conjecture 21) Let $RS_k[F, D]$ be the Reed-Solomon code over a prime field $F = F_q$ with defining domain D and rate $\rho = \frac{k}{|D|}$. Then there exists a constant C_ρ such that for every $\theta = 1 - \rho - \eta$, with $\eta > 0$, $RS_k[F, D]$ is list-decodable from a fraction of θ errors with list size*

$$L \leq \left(\frac{|D|}{\eta} \right)^{C_\rho}.$$

Remark 10. No concrete assumptions on the constant C_ρ are made in [BSGKS20].

For quite large fields F (compared to the block length $|D| = n$) there are linear codes which are list decodable up to capacity bound $1 - \rho$, such as the *folded Reed-Solomon codes* (see [Gur07], e.g.). In the case of a bounded alphabet, Guruswami [Gur07] demonstrates binary linear codes which are list decodable to the Zyablow bound $\frac{1-\rho}{H}$ (here, H is the entropy of the code) and uses such codes to construct examples that approach capacity bound, having list size $L = O\left(\frac{1}{\eta}\right)$.

However, practitioners seem to avoid this conjecture. The ethSTARK documentation [Sta21] takes a toy protocol as a representative for the entire DEEP-ALI of AIR, whereas the plonky2 writeup [Pol] only sketches soundness in the polynomial oracle model, with no reference to list bounds.

References

[BCC+16] Jonathan Bootle, Andrea Cerulli, Pyrros Chaidos, Jens Groth, and Christophe Petit. Efficient zero-knowledge arguments for arithmetic circuits in the discrete log setting. In M. Fischlin and J.S. Coron, editors, *EUROCRYPT 2016*, volume 9666 of *LNCS*. Springer, 2016. Full paper: `https://eprint.iacr.org/2016/263`.

[BDFG21] Dan Boneh, Justin Drake, Ben Fisch, and Ariel Gabizon. Halo Infinite: Recursive zk-snarks from additive polynomial commitments. In *CRYPTO 2021*, volume 12825 of *LNCS*. Springer, 2021. Full paper: `https://eprint.iacr.org/2020/1536`.

[BFS20] Benedikt Bünz, Ben Fisch, and Alan Szepieniec. Transparent SNARKs from DARK compilers. In *EUROCRYPT 2020*, 2020. Full paper: `https://eprint.iacr.org/2019/1229`.

[BSBHR18a] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Fast Reed-Solomon interactive oracle proofs of proximity. In *ICALP 2018*, 2018. Full paper: `https://eccc.weizmann.ac.il/report/2017/134/`.

[BSBHR18b] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable, transparent, and post-quantum secure computational integrity. In *IACR ePrint Archive 2018/046*, 2018. `https://eprint.iacr.org/2018/046`.

[BSCI+20] Eli Ben-Sasson, Dan Carmon, Yuval Ishai, Swastik Kopparty, and Shubhangi Saraf. Proximity gaps for Reed-Solomon codes. In *FOCS 2020*, 2020. Full paper: `https://eprint.iacr.org/2020/654`.

[BSCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive oracle proofs. In *TCC 2016*, pages 31–60, 2016.

[BSGKS20] Eli Ben-Sasson, Lior Goldberg, Swastik Kopparty, and Shubhangi Saraf. DEEP-FRI: Sampling outside the box improves soundness. In *ITCS 2020*, 2020. Full paper: `https://eprint.iacr.org/2019/336`.

[CHM+20] Alessandro Chiesa, Yuncong Hu, Mary Maller, Pratyush Mishra, Noah Vesely, and Nicholas Ward. Marlin: Preprocessing zkSNARKs with universal and updatable SRS. In *EUROCRYPT 2020*, volume 12105 of *LNCS*, 2020. Full paper: `https://eprint.iacr.org/2019/1047.pdf`.

[GS99] Venkatesan Guruswami and Madhu Sudan. Improved decoding of Reed-Solomon and algebraic-geometry codes. In *IEEE Trans. on Information Theory*, volume 45(6), 1999.

[Gur07] Venkatesan Guruswami. Algorithmic results in list decoding. In *Foundation and Trands in Theoretical Computer Science*, volume 2(2), 2007.

[GWC19] Ariel Gabizon, Zachary J. Williamson, and Oana Ciobotaru. PLONK: Permutations over Lagrange-bases for oecumenical non-interactive arguments of knowledge. In *IACR ePrint Archive 2019/953*, 2019. `https://eprint.iacr.org/2019/953`.

[HGdB21] Ulrich Haböck, Alberto Garoffolo, and Daniele di Benedetto. Darlin: Recursive proofs based on Marlin. In *IACR preprint archive 2021/930*, 2021. `https://eprint.iacr.org/2021/930`.

[KPV19] Assimakis Kattis, Konstantin Panarin, and Alaxander Vlasov. REDSHIFT: Transparent snarks from list polynomial commitment IOPs. In *IACR preprint archive*, 2019. `https://eprint.iacr.org/2019/1400`.

[MBKM19] Mary Maller, Sean Bowe, Markulf Kohlweiss, and Sarah Meiklejohn. Sonic: Zero-knowledge SNARKs from linear-size universal and updatable structured reference strings. In *ACM SIGSAC Conference on Computer and Communication Security*, pages 2111–2128, 2019. Full paper: `https://eprint.iacr.org/2019/099`.

[Pol] Polygon Zero Team: plonky2. `https://github.com/mir-protocol/plonky2`.

[Sta21] StarkWare Team. ethSTARK documentation – version 1.1. In *IACR preprint archive 2021/582*, 2021. `https://eprint.iacr.org/2021/582`.

[Sud97] Madhu Sudan. Decoding of Reed Solomon codes beyond the error-correction bound. In *Journal of Complexity*, volume 13(1), 1997.

[WB86] Lloyd R. Welch and Elwyn R. Berlekamp. Error correction for algebraic block codes. US Patent 4633470, 1986. `https://patents.google.com/patent/US4633470A`.

Appendix

In this section we recap well-known facts on decodability of Reed-Solomon codes[5], and describe the weighted variant of Theorem 1, which is used by the soundness analysis of FRI.

Unless contrary stated, we assume that K is a *general* field (finite, or infinite), and as for finite fields we shall call

$$\mathsf{RS}_k[K, D] = \{p(x)|_{x\in D} \ : \ p(X) \in K[X], \deg p(X) \leq k-1\}$$

the Reed-Solomon code with rate $\rho = \frac{k}{|D|}$ and blocklength $n = |D|$. We say that a family of codes $\{V(n)\}$ of increasing blocklength n is list decodable up to distance $\theta \in (0, 1)$, if the maximum possible number of θ-close codewords,

$$L = \sup_{f\in K^D} \big|B(f, \theta) \cap V(n)\big|,$$

is polynomial in the blocklength n. (Here, $B(f, \theta) = \{w \in \mathsf{RS}_k[K, D] : \delta(f, w) < \theta\}$ is the open -ball around f, and δ is the fractional Hamming distance.) As in the main part of the document, we throughout assume that both n and k are even.

Berlekamp-Welch decoder

Assume that $f \in K^D$ is at most θ_0–close to V, with $\theta_0 = \frac{1-\rho}{2}$ being the unique decoding radius, and let $p(X)$ be the unique polynomial of degree $\leq k-1$ such that $\delta(f, p) \leq \theta_0$. Then the number of points of disagreement is at most $e = \frac{n-k}{2}$. The Berlekamp-Welch decoder [WB86] is based on the observation that if $\Omega = \{x_1, \ldots, x_e\}$ is the set of errors, and $E(x) = \prod_{x\in\Omega}(X - x)$ is its vanishing polynomial, then we have

$$E(x) \cdot f(x) = E(x) \cdot p(x)$$

for all $x \in D$.

Protocol 4 (Welch-Berlekamp decoding)**.** *Let K be a general field, and $V = \mathsf{RS}_k[K, D]$ be the Reed-Solomon code of length $n = |D|$ and rate $\rho = \frac{k}{n}$. Assume any word $f \in K^D$.*

1. *Find the coefficients of polynomials $E(X)$, $G(X)$ over K with $\deg E(X) \leq e$, $\deg G(X) \leq k-1+e$, where $e = \frac{n-k}{2}$, such that*

   $$E(x) \cdot w(x) = G(x) \text{ for all } x \in D.$$

   This linear system has at least one non-trivial solution which can be found in at most $O\left(n^3\right)$ field operations.

   This is a homogeneous linear system of $|D| = n$ equations in $k + 2 \cdot e + 1 = n + 1$ unknown: The $e + 1$ coefficients of $E(X)$ and the $k + e$ coefficients of $G(X)$.

   Notice that for any such non-trivial solution $(E(X), G(X))$ both $E(X)$ and $G(X)$ must be non-trivial.

   If one of the two would be identically zero, the size of D the same is true for the other.

2. *For any such non-trivial solution $(E(X), G(X))$ obtained in step 1, check if $G(X)$ is divisible by $E(X)$. If yes, then output $p(X) = \frac{G(X)}{E(X)}$. (If not, then abort.)*

For a word $f \in K^D$ with fractional Hamming distance of at most θ_0, Step (2) of Protocol 4 always succeeds: Let $p(X)$ be the (unique) θ_0-close code word. This polynomial agrees with f on a set of size $a \geq \frac{n+k}{2}$. Consider the bivariate polynomial

$$Q(X, Y) := Y \cdot E(X) - G(X).$$

[5]The survey by Guruswami [Gur07] is a recommended source.

Then $Q(X, p(X))$ is a univariate polynomial of degree

$$\deg Q(X, p(X)) \leq k - 1 + e = \frac{n+k}{2} - 1,$$

which by the assumption on $p(X)$ has at least a zeroes. Consequently $Q(X, p(X))$ is trivial and $p(X) \cdot E(X) = G(X)$ holds as a formal identity. Since $E(X)$ is non-trivial, we conclude divisibility.

List decoding

The Sudan decoder

The Sudan list decoder [Sud97] generalizes the Berlekamp-Welch procedure by searching for general bivariate polynomials $Q(X, Y) \in K[X, Y]$ which satisfy

$$Q(x, f(x)) = 0 \text{ for all } x \in D.$$

In order that $Y - p(X)$ is a factor of $Q(X, Y)$ for every polynomial $p(X)$ of degree $\leq d = k - 1$ which has the claimed agreement set size with f, one looks for such bivariate Q so that the degree of $Q(X, p(X))$ for any such polynomial is smaller than the targeted agreement set size.

Definition 11. The $(1, d)$*–weighted degree* (in short, $(1, d)$*–degree*) of a monomial $X^i \cdot Y^j$ is $i + d \cdot j$. More generally, the $(1, d)$–weighted degree of a bivariate polynomial $Q(X, Y)$ is the maximum of the weighted degrees of its monomials.

A polynomial $Q(X, Y)$ of $(1, d)$–weighted degree W is of the form

$$Q(X, Y) = \sum_{i + d \cdot j \leq W, i, j \geq 0} c_{i,j} \cdot X^i \cdot Y^j,$$

and its number of coefficients is

$$\begin{aligned}
\sum_{j=0}^{\lfloor W/d \rfloor} W - d \cdot j + 1 &= (W+1) \cdot \left(\left\lfloor \frac{W}{d} \right\rfloor + 1 \right) - d \cdot \frac{\left\lfloor \frac{W}{d} \right\rfloor \cdot \left(\left\lfloor \frac{W}{d} \right\rfloor + 1 \right)}{2} \\
&\geq \left(\left\lfloor \frac{W}{d} \right\rfloor + 1 \right) \cdot \left(W + 1 - \frac{W}{2} \right) \\
&\geq \frac{(W+1) \cdot (W+2)}{2 \cdot d}
\end{aligned}$$

As a consequence, if this lower bound exceeds the number of linear equations $n = |D|$, the linear system has a non-trivial solution. In particular this holds for any

$$W \geq \left\lfloor \sqrt{2 \cdot d \cdot n} \right\rfloor.$$

Protocol 5 (Sudan list decoder)**.** *Assume that K is a general field, and $\mathsf{RS}_k[K, D]$ is the Reed-Solomon code of lenght $n = |D|$ and rate $\rho = \frac{k}{n}$. Let $f \in K^D$, and choose an agreement parameter $a \in [0, n]$, $a > \sqrt{2 \cdot d \cdot n}$, where $d = k - 1$.*

1. *Solve the linear system on the coefficients of $Q(X, Y)$ with $(1, d)$–degree $W = \left\lfloor \sqrt{2 \cdot d \cdot n} \right\rfloor$, given by the interpolation constraints*

$$Q(x, f(x)) = 0, \quad x \in D.$$

This system has a non-trivial solution which is found in at most $O(n^3)$ field operations.

Note that by construction, for any $\left(1 - \frac{a}{n}\right)$–close code word $p(X)$ the irreducible polynomial $Y - p(X)$ divides $Q(X, Y)$. This already proves that the list size $L \leq \left\lfloor \frac{W}{d} \right\rfloor \leq \frac{\sqrt{2 \cdot d \cdot n}}{d} = \sqrt{\frac{2 \cdot n}{d}}$.

2. *Find all factors of $Q(X,Y)$ which are of the form*

$$Y - p(X),$$

with $p(X)$ being a polynomial over K of degree at most $k-1$. There are at most $\sqrt{\frac{2\cdot n}{d}}$ such factors. Filter out those which agree with f on at least a points.

The efficiency of Step (2) depends on the field K. If K is a finite field, then there are polynomial time algorithms (both probabilistic or deterministic) for finding such factors of the form $Y - p(X)$. (They both rely on univariate factorization, see [Gur07], e.g.) If K is infinite, then this might not be true in general.

The Guruswami-Sudan decoder

To extend the interpolation technique to the Johnson limit $1-\sqrt{\rho}$, one takes into account that several close codewords might coincide at some points. One therefore looks for polynomials $Q(X,Y)$ the $(1,d)$–degree of which is m times as large as the targeted agreement set would suggest, and which have a zero of order m at every interpolating point $(x, f(x))$, $x \in D$. The parameter $m \geq 1$ is called *multiplicity parameter*.

Definition 12. A polynomial $Q(X,Y) \in K[X,Y]$ is said to have a *zero of order m* at the point (x,y), if the polynomial $Q(X-x, Y-y)$ has no monomial of absolute degree m.

Such polynomials $Q(X,Y)$ of $(1,d)$-weighted degree W have still the property, that if $p(X)$ is a polynomial of $\deg p(X) \leq d$, then

$$\deg Q(X, p(X)) \leq \frac{W}{m}.$$

Again, counting the number of coefficients and comparing with the number of interpolation constraints yields that whenever

$$\frac{(W+1)\cdot(W+2)}{d\cdot m\cdot(m+1)} > n,$$

and hence in particular for $W \geq \left\lfloor \sqrt{m\cdot(m+1)\cdot d\cdot n} \right\rfloor$ there always exists such a (non-trivial) polynomial $Q(X,Y)$. (For details, see [Gur07], e.g.)

Protocol 6 (Guruswami-Sudan list decoder [GS99])**.** *Assume that K is a general field, and $\mathsf{RS}_k[K,D]$ the Reed-Solomon code of length $n = |D|$ and rate $\rho = \frac{k}{n}$. Let $f \in K^D$, and choose an agreement parameter $a \in [0,n]$, $a > \sqrt{\left(1+\frac{1}{m}\right)\cdot d\cdot n}$, where m is a positive integer (the multiplicity parameter).*

1. *Solve the linear system on the coefficients of $Q(X,Y)$ with $(1,d)$–degree $W = \left\lceil \sqrt{m\cdot(m+1)\cdot d\cdot n} \right\rceil$: For each $x \in D$,*

$$Q(X,Y) \text{ has a zero of order } m \text{ at } (x, f(x)).$$

Such a solution always exists and can be found in polynomially many field operations.

By construction, again for any $\left(1-\frac{a}{n}\right)$–close code word $p(X)$ the irreducible polynomial $Y - p(X)$ divides $Q(X,Y)$. This already proves that the list size $L \leq \frac{\sqrt{m\cdot(m+1)\cdot d\cdot n}}{d} < \sqrt{\frac{m\cdot(m+1)}{\rho}}$.

2. *Find all factors of $Q(X,Y)$ which are of the form*

$$Y - p(X),$$

with $p(X)$ being a polynomial over K of degree at most $d = k-1$. There are at most $\sqrt{\frac{m\cdot(m+1)}{\rho}}$ many. Filter out those which agree with f on at least a points.

As before, this step might be efficient or not, depending on the field K.

Remark 13. Choosing the discriminant method to find factors of the form $Y - p(X)$, the Guruswami-Sudan list decoder taking at most

$$O\left(\max\left\{\frac{d^3 \cdot n^6 \cdot a^6}{(a^2 - d \cdot n)^6}, \frac{a^6}{k^3}\right\}\right)$$

field operations over K, see [GS99]. This is at most of order $O\left(|D|^{15}\right)$.

Note that choosing

$$\alpha = \frac{a}{n} \geq \sqrt{\left(1 + \frac{1}{m}\right) \cdot \rho}$$

implies a large enough agreement parameter for the Protocol 6. In particular the choice $\alpha = \left(1 + \frac{1}{2m}\right) \cdot \sqrt{\rho}$ used throughout the main part of the document is strong enough, since

$$\left(1 + \frac{1}{2 \cdot m}\right)^2 = 1 + \frac{1}{m} + \frac{1}{4 \cdot m^2} \geq 1 + \frac{1}{m}.$$

Let us summarize the consequences of Protocol 6.

Theorem 14 (Guruswami-Sudan)**.** *Let K be a general (possibly infinite) field, and*

$$\mathsf{RS}_k[K, D] = \{p(x)|_{x \in D} \,:\, p(X) \in K[X], \deg(p) < |D|\}$$

the Reed-Solomon code of block length $n = |D|$ and rate $\rho = \frac{k}{n}$. Choos a proximity parameter $\theta = 1 - \left(1 + \frac{1}{2 \cdot m}\right) \cdot \sqrt{\rho}$ for some integer $m \geq 1$. Then $\mathsf{RS}_k[K, D]$ is list decodable for θ with list bound

$$L \leq \sqrt{\frac{m \cdot (m+1)}{\rho}} \leq \frac{m + \frac{1}{2}}{\sqrt{\rho}}. \tag{25}$$

If K is finite, then the Guruswami-Sudan decoder runs in polynomial time.

Weighted correlated agreement

We say that a function $f \in F^D$ has *μ-agreement* of at least α with another function $g \in F^D$,

$$\mathsf{agree}_\mu(f, g) > \alpha,$$

if there is a set $A \subseteq D$ of measure $\mu(A) > \alpha$ on which both functions agree. Likewise we say that

$$\mathsf{agree}_\mu(f, \mathsf{RS}_k) > \alpha,$$

if there exists a $p \in \mathsf{RS}_k[F, D]$ for which $agree_\mu(f, p) > \alpha$.

Theorem 15. *(Full version of [BSCI+20], Theorem 7.1) Let $\theta \in \left(\frac{1-\rho}{2}, 1 - \sqrt{\rho}\right)$, where $\theta = 1 - \sqrt{\rho} \cdot \left(1 + \frac{1}{2m}\right)$, for some integer $m \geq 3$, and assume that μ is a sub-probability measure on D with common denominator M, i.e. for all x in D*

$$\mu(\{x\}) = \frac{a_x}{M},$$

for an integer value a_x . Suppose that for f_0, f_1, ..., $f_{N-1} \in F^D$,

$$\frac{\left|\{\lambda : \mathsf{agree}_\mu(f_0 + \lambda \cdot f_1 + \ldots + \lambda^{N-1} \cdot f_{N-1}, \mathsf{RS}_k) > \alpha\}\right|}{|F|} > \max\left(\varepsilon_J, (N-1) \cdot \frac{M \cdot |D| + 1}{|F|} \cdot \frac{2m+1}{\sqrt{\rho}}\right),$$

with ε_J as in (2). Then there exist polynomials $p_0(X)$, $p_1(X)$, ..., $p_{N-1}(X)$ from $\mathsf{RS}_k[F, D]$, and a set A of density $\mu(A) > \alpha$ on which f_i coincides with p_i for all $i = 0, \ldots, N-1$.