Worst-case to average case reductions for the distance to a code

CCC 2018

Eli Ben-Sasson and Swastik Kopparty and Shubhangi Saraf

June 2018

Overview

- motivation
- main results
- applications
- one proof

Motivation

- Arithmetization [LFKN92]
  - reduces computational problems to algebraic problems about low-degree polynomials
  - used in IP, MIP, PCP, ZK, IPCP, IOP, ... protocols
- example: 3SAT formula $\phi \mapsto$ "local" constraints over linear code $V \subset \mathbb{F}^n$, satisfying
  - Completeness: $\phi \in 3SAT \Rightarrow \exists v_1, \ldots, v_k \in V$ that satisfy all "local" constraints
  - Soundness: $\phi \notin 3SAT \Rightarrow \forall \vec{u} = (u_1, \ldots, u_k) \in (\mathbb{F}^n)^k$, if $\vec{u}$ satisfies "local" constraints, then $\exists u^* \in \vec{u}, \Delta(u^*, V) > 0.1$ (Δ is relative Hamming distance).
- This talk discusses
  1. worst-to-average case: $\Delta(u^*, V) > \delta \mapsto$ almost all $u \in \text{span}(\vec{u})$ satisfy $\Delta(u, V) \approx \delta$
  2. local distance amplification: $\Delta(u^*, V) > 0.1 \mapsto \Delta(u^{**}, V) > 0.99$, u^{**} locally computed from u^*.
- Techniques: (i) more interaction, (ii) more randomness; for (2) above, also use automorphisms of V.

Main results on worst-to-average case distance reductions

Let $U, V \subseteq \mathbb{F}^n$. If $u^* \in U$ is δ-far from V ($\Delta(u^*, V) \geq \delta$) ...

Prior state of art — Unique decoding distance [RVW 2013]

Then most $u \in U$ are at least half as far from V as u:

$$\Pr_{u \in U}[\Delta(u, V) < \delta/2] \leq \frac{1}{|\mathbb{F}|-1}.$$

First result — List decoding distance for general spaces V

Then most $u \in U$ are $\approx J(\delta) \triangleq 1 - \sqrt{1-\delta}$ far from V:

$$\Pr_{u \in U}[\Delta(u, V) < J(\delta) - \epsilon] < O_\epsilon\left(\frac{1}{|\mathbb{F}|}\right),$$

For $\delta = 1 - o(1)$, most $u \in U$ have $\Delta(u, V) = 1 - o(1)$.

Second result — Distance preservation for codes V

If moreover V has minimal distance λ and $\delta < J(J(\lambda)) - \epsilon$, then

$$\Pr_{u \in U}[\Delta(u, V) < \delta - \epsilon] < O_\epsilon\left(\frac{1}{|\mathbb{F}|}\right),$$

For $\lambda = 1 - o(1)$, most $u \in U$ have $\Delta(u, V) \approx \delta$.

Main results on local distance amplification

Let $V \subseteq \mathbb{F}^n$ be a subspace

- *q-local map* $M : \mathbb{F}^n \to \mathbb{F}^n$ — ith entry of $M(v)$ depends on $\leq q$ entries of v;
- We are interested in q-local maps that (i) preserve perfect completeness and (ii) amplify soundness
- *Automorphism group* $\mathrm{Aut}(V)$ — group of permutations on $[n]$ that leave V invariant: $\forall v \in V, \pi \in \mathrm{Aut}(V), \pi(v) \in V$
- Example: For $V = \mathrm{RS}[\mathbb{F}, \rho] \triangleq \{f(x) : \mathbb{F} \to \mathbb{F} \mid \deg(f) < \rho|\mathbb{F}|\}$, $\mathrm{Aut}(V)$, $\mathrm{Aut}(\mathrm{RS}[\mathbb{F}, \rho]) = \{x \mapsto ax + b \mid a \in \mathbb{F}^*, b \in \mathbb{F}\}$;

Third result — Distance amplification for RS codes

For $\delta, \epsilon > 0$ there exists $q = q(\rho, \delta, \epsilon)$ such that if $u : \mathbb{F} \to \mathbb{F}$ is δ-far from $\mathrm{RS}[\mathbb{F}, \rho]$ then
$\Pr_{\pi_1, \ldots, \pi_q \in \mathrm{Aut}(\mathrm{RS}[\mathbb{F}, \rho])} \left[\Delta(\sum_{i=1}^{q} \pi_i(u), \mathrm{RS}[\mathbb{F}, \rho]) < J(J(1-\rho)) - \epsilon\right] < \frac{O_{\epsilon,q}(1)}{|\mathbb{F}|}$,
For $\rho = o(1)$, this gives distance amplification up to distance $1 - o(1)$.

Application I: High-error Polishchuk-Spielman theorems

For $A, B \subseteq \mathbb{F}, |A| = |B| = N$ suppose $f_r, f_c : A \times B \to \mathbb{F}$ satisfy

- each row of $f_r : A \times B \to \mathbb{F}$ is a degree d_r polynomial
- each column of f_c is a degree d_c polynomial
- $\Pr_{a,b}[f_r(a,b) = f_c(a,b)] \geq \eta$, η is the *agreement parameter*

Then

- Folklore: $\eta = 1 \Rightarrow f_r = f_c$ is degree-(d_r, d_c) bivariate polynomial
- High degree, high agreement [PS94]: For $\frac{d_r+d_c}{N} + < \frac{1}{2}$ and $\eta > \frac{1}{2}$, we have that f_r, f_c are close to degree-(d_r, d_c) bivar polynomial
- Open: prove for degree $d_r, d_c = \Omega(|A|)$ and $\eta \ll 1/2$
- [CMS17]: for $\eta \ll \frac{1}{2}$ and $d_r, d_c = O(\log N)$, we have that f_r, f_c are close to degree-(d_r, d_c) poly
- New: for $\eta \ll \frac{1}{2}$ and $d_r = O(\log \log n)$ and $d_c = \Omega(N)$ we have that f_r, f_c are close to degree-(d_r, d_c) poly;
- [CMS17] and new result are incomparable
  - [CMS17] holds for larger degree in *both* axes;
  - new result requires lower degree, but only for one axis;
  - different proof techniques.

Application II: Improved IOPPs for Reed-Solomon codes

Plan:

1. Interactive Oracle Proof of Proximity (IOPP) definition
2. Fast RS IOPP (FRI) protocol and prior soundness
3. Improved FRI soundness analysis

Interactive Oracle Proof of Proximity (IOPP) [RRR16,BCS16]

- Proximity testing: given $P \subset \Sigma^S$, oracle $f : S \to \Sigma$, distinguish between $f \in P$ and f is δ-far from P;
- IOPP model generalizes IP [GMR85], IPCP [KR05], and PCPP [BGHSV05, DR06];
- IOPP model (informal definition)
  - Prover sends oracle $f : S \to \Sigma$
  - Verifier sends 1st randomness r_1
  - Prover sends 1st proof oracle $\pi_1 : S_1 \to \Sigma$
  - Verifier sends r_2, prover sends π_2, repeat for R rounds;
  - Verifier queries $f, \pi_1, \ldots, \pi_R$, outputs acc/rej
- soundness+completeness as in the PCPP model
- query complexity q measured over all oracles;
- proof length and prover complexity measured over $\pi_1, \ldots, \pi_R$

Fast RS IOPP (FRI) [BBHR18]

- RS proximity testing: Fix field $\mathbb{F}$, blocklength $N \leq |\mathbb{F}|$, rate ρ, proximity parameter $\delta \leq 1 - \rho$;
- Given oracle $f : S \to \mathbb{F}$
  - accept if $\deg(f) < \rho N$,
  - reject w.p. $\geq 1/2$ if f is δ-far from degree $< \rho N$
- Pay attention to proximity parameter δ_0

Theorem (Informal) [BBHR18] [New]

FRI protocol with blocklength N, and rate $\rho < 1$ has

- $O(N)$ prover arithmetic complexity and proof length
- $O(\log N)$ rounds, verifier arithmetic complexity and queries;
- $\delta - \frac{O(1)}{|\mathbb{F}|}$ rejection pr. for $\delta < \delta_0$, where $\delta_0 \approx \cancel{\frac{1-\rho}{4}}\ 1 - \rho^{\frac{1}{4}}$

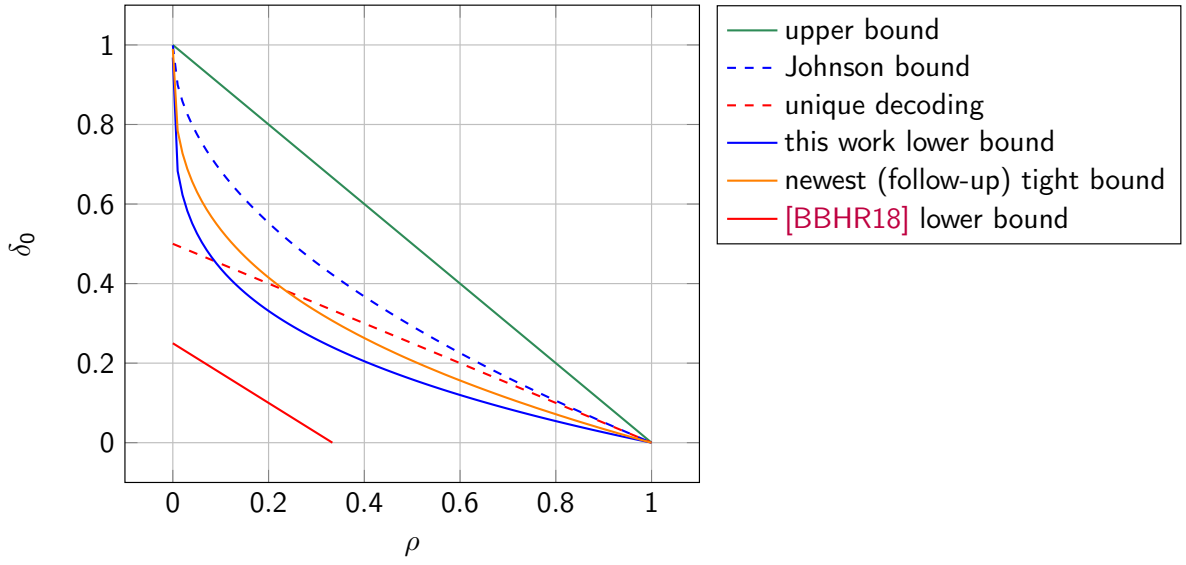
Theorem (followup) [Newest]

FRI protocol has same parameters as in Theorem above, but

- $\delta - \frac{O(1)}{|\mathbb{F}|}$ rejection pr. for $\delta < \delta_0$, where $\delta_0 \approx 1 - \rho^{\frac{1}{3}}$, tight(!)

FRI soundness as function of rate

Higher lines mean higher (better) soundness (rejection prob.):



FRI soundness: example setting

- Example: for $\rho = 2^{-8} = \frac{1}{256}$ and $\delta = 1 - \rho$:
  - old rejection probability $\geq 1/4$
  - new rejection probability $\geq 3/4$
  - follow-up: tight bound (upper+lower): $= 0.842\ldots$

One proof

- Lemma If $\Delta(u^*, V) \geq \delta$, then there are at most $O(1)$ values of $\alpha \in \mathbb{F}$ for which
$$\Delta(u^* + \alpha u, V) \leq J(\delta) - \epsilon.$$

- Key ingredient: **Johnson Bound** If $u, w_1, \ldots, w_t \in \mathbb{F}^n$ are such that $\Delta(w_i, w_j) \geq \delta$ and $\Delta(u, w_i) \leq J(\delta) - \epsilon$, then $t \leq O_\epsilon(1)$.
- Proof: Suppose $\alpha_1, \ldots, \alpha_t \in \mathbb{F}$ and $v_1, \ldots, v_t \in V$ are such that:
$$\Delta(u^* + \alpha_i u, v_i) < J(\delta) - \epsilon.$$

  Then:
$$\Delta(u, \frac{1}{\alpha_i}(v_i - u^*)) < J(\delta) - \epsilon.$$

  But note that:
$$\Delta(\frac{1}{\alpha_i}(v_i - u^*), \frac{1}{\alpha_j}(v_j - u^*)) \geq \Delta(u, V) \geq \delta.$$

  Thus the Johnson bound gives the desired bound on t.

Proof sketch for distance preservation

- Distance Preservation Theorem Suppose V has distance λ, and $\Delta(u^*, V) \geq \delta$, where $\delta \leq J(J(\lambda))$. Then most for most $\alpha \in \mathbb{F}$, we have that $u^* + \alpha u$ is $(\delta - \epsilon)$-far from V.
- Intermediate structure theorem Suppose V has distance λ and $\delta < J(J(\lambda))$.
  For arbitrary $u, u^* \in \mathbb{F}^n$, if there are many $\alpha \in \mathbb{F}$ such that $\Delta(u^* + \alpha u, V) < \delta - \epsilon$, then there is a set $S \subseteq [n]$, and vectors $v, v^* \in V$ with:
  - $|S| < \delta + \epsilon$.
  - $u|_{[n]\setminus S} = v|_{[n]\setminus S}$.
  - $u^*|_{[n]\setminus S} = v^*|_{[n]\setminus S}$.
- In words: the only way to make the line $\{u^* + \alpha u \mid \alpha \in \mathbb{F}\}$ in $\mathbb{F}^n$ have many points close to V is if u^* and u are both close to V *with the set of agreeing coordinates aligned*.
- Immediately implies the distance preservation theorem.
- Intermediate structure theorem proved using (1) two invocations of the Johnson bound[1], and (2) some tools from graph theory.

[1] of course .. see $J(J(\lambda))$

Applications: proof sketch

- RS distance amplification:
  - Want to show that if g = random linear combination of random affine shifts of f, then g far from RS code.
  - Key tool: intermediate structure theorem.
  - If g is often close to low degree, then we get that f and a random affine shift of f must have a large set of coordinates where both agree with RS code.
  - But random affine shifts are quite mixing: This rules out the above possibility.
- High-error Polishcuk-Spielman bivariate testing:
  - Immediately follows from intermediate structure theorem.
- Improved soundness for Fast Reed-Solomon IOPP:
  - Immediately follows from distance preservation theorem.

Final remarks

Summary

- Worst-to-average case reductions for linear spaces
  - New: If some $u^* \in U$ is δ-far from V, then most members of U are $\approx \delta$-far from V
  - Prior [RVW16]: ... most members of U are $\approx \delta/2$-far from V
- q-local distance amplification for RS codes
  - New: If $f : \mathbb{F} \to \mathbb{F}$ is $\frac{1}{100}$-far from degree-$\frac{|\mathbb{F}|}{100}$ polynomials, then w.h.p. over random $a_i \in \mathbb{F}^*, b_i \in \mathbb{F}$,

$$\hat{f}(X) \triangleq \sum_{i=1}^{100} f(a_i X + b_i)$$

    is (i) 100-local and (ii) $\frac{9}{10}$-far from degree-$\frac{|\mathbb{F}|}{100}$ polynomials
- two applications to low-degree testing
  - high-error Polischuk-Spielman bivariate low-degree testing
  - improved RS soundness analysis of FRI protocol