Chapter 2
Locally Testable Codes with an Application in Hardness Amplification

Angsheng Li

Institute of Software
Chinese Academy of Sciences

Advanced Algorithms, U CAS
14th, March, 2016

Outline

1. Backgrounds
2. Yao's XOR
3. Impagliazzo's hard core
4. Error correcting codes
5. Decoding ECC
6. Local decoding and hardness amplification
7. Local algorithms in networks

Algebraic fundamental theorem

Given a polynomial $p(x_1, \cdots, x_n)$ of (total) degree d, and a finite field $\mathbb{F}$. If $p \not\equiv 0$, then

$$\Pr_{x \in_{\mathrm{R}} \mathbb{F}^n}[p(x) = 0] \leq \frac{d}{|\mathbb{F}|}.$$

IP=PSPACE

Completeness: If a QSAT instance ϕ is satisfied, then the honest prover convinces the verifier to accept with probability 1.

Soundness: If a prover P is false, then with probability at least

$$(1 - \frac{d}{p})^n$$

an error is kept to the final stage, at which the verifier V works with only constants and certainly detects the error.

Applications of the locally testable codes

1. IP and PCP
2. Hardness amplification
3. Codes and information theory (research project)
4. Networks (research project)

Idea

- Why hardness amplification?
- XOR ensures that:
  mild average-case hard $\Rightarrow$ strong average-case hard

Average-case hardness

Definition

For $f : \{0,1\}^n \to \{0,1\}$ and $\rho \in [0,1]$, we define the *ρ-average case hardness* of f, written $H^{\rho}_{\mathrm{avg}}(f)$, to be the largest size s such that for every circuit C of size $\leq s$,

$$\Pr_{x \in_{\mathrm{R}} \{0,1\}^n}[C(x) = f(x)] < \rho.$$

For $f : \{0,1\}^* \to \{0,1\}$,

$$H^{\rho}_{\mathrm{avg}}(f_n) = H^{\rho}_{\mathrm{avg}}(f)(n),$$

f_n is the restriction of f on $\{0,1\}^n$, denoted by $f_n = f \upharpoonright \{0,1\}^n$.

Worst-case hardness

Definition

We define the *worst-case hardness* of f to be

$$H_{\mathrm{wrs}}(f) = H^{1}_{\mathrm{avg}}(f).$$

We define the *average-case hardness* of f by

$$H_{\mathrm{avg}}(f) = \max\{s : H^{\frac{1}{2}+\frac{1}{s}}_{\mathrm{avg}}(f) \geq s\}.$$

Yao's XOR lemma

Theorem

(Yao, 1982) For every $f : \{0,1\}^n \to \{0,1\}$, $\delta > 0$ and $k \in \mathbb{N}$, if $\epsilon > 2(1-\delta)^k$, then

$$H_{\text{avg}}^{\frac{1}{2}+\epsilon}(f^{\oplus k}) \geq \frac{\epsilon^2}{400n} \cdot H_{\text{avg}}^{1-\delta}(f), \tag{1}$$

where $f^{\oplus k}(x_1, \cdots, x_k) = \sum_{i=1}^{k} f(x_i) \pmod 2$.

Intuition If small circuits fail to compute f with prob better than $1-\delta$, then some smaller circuits fail to compute $f^{\oplus k}$ with pron better than $\frac{1}{2} + 2(1-\delta)^k$.

Intuition

By assumption, for a small circuit C,

$$\begin{aligned} f(x_1) = C(x_1) & \; - < 1 - \delta \\ & \vdots \\ f(x_k) = C(x_k) & \; - < 1 - \delta. \end{aligned} \tag{2}$$

With probability $< (1-\delta)^k$, for each i, $f(x_i) = C(x_i)$, giving the correct $f^{\oplus k}(x_1, \cdots, x_k)$.
Suppose it is not the case, then we can only guess $f^{\oplus k}(x_1, \cdots, x_k)$, for which the prob of correctness is $\frac{1}{2}$.
Therefore, the correctness of computation for $f^{\oplus k}(x_1, \cdots, x_k)$ by small circuits is

$$< (1-\delta)^k + \frac{1}{2}.$$

Hardcore lemma

We say that a distribution H over $\{0,1\}^n$ has density δ, if for every $x \in \{0,1\}^n$,

$$\Pr[H = x] \leq 1/\delta \cdot \frac{1}{2^n}.$$

Lemma

(Impagliazzo, 95) For every $\delta > 0$, $f : \{0,1\}^n \to \{0,1\}$, and $\epsilon > 0$, if $H_{\mathrm{avg}}^{1-\delta}(f) \geq s$, then there is a density δ distribution H such that for every circuit V of size at most $\epsilon^2 s/100n$,

$$\Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)] \leq \frac{1}{2} + \epsilon.$$

Intuition of hardcore

$$\forall C \exists H \Rightarrow \exists H \forall C$$

$\forall C \exists H$
For every small circuit C, C fails to compute f on at least $\delta \cdot 2^n$ inputs.

$\exists H \forall C$
There exists a δ-dense distribution H on which every small circuit fails to compute f slightly better than random guess.

Hardcore implies Yao's XOR lemma

Given $f: \{0,1\}^n \to \{0,1\}$ with $H_{\mathrm{avg}}^{1-\delta}(f) \geq s$, and k, suppose to the contrary that there is a circuit C of size $s' = \frac{\epsilon^2}{400n} s$ such that

$$\Pr_{x_1,\cdots,x_k \in_{\mathrm{R}} U_n}[C(x_1,\cdots,x_k) = \sum_{i=1}^{k} f(x_i) \pmod 2)] \geq \frac{1}{2} + \epsilon,$$

where $\epsilon > 2(1-\delta)^k$.
Fix $k = 2$.

Proof - I

By hardcore lemma, let H be a density δ distribution such that for every circuit C' of size s', C' fails to compute $f^{\oplus k}$ better than $\frac{1}{2} + \frac{\epsilon}{2}$.
Define a distribution G by

$$\Pr[G = x] = \frac{1}{1-\delta} \cdot (2^{-n} - \delta \cdot \Pr[H = x]).$$

Then,

$$U_n = (1-\delta)G + \delta H.$$

$$(U_n)^2 = (1-\delta^2)G^2 + (1-\delta)\delta GH + \delta(1-\delta)HG + \delta^2 H^2.$$

Proof - II

Given a distribution D over $\{0,1\}^{2n}$, let P_D be the probability that

$$C(x_1, x_2) = \sum_{i=1}^{2} f(x_i) \pmod 2$$

where $(x_1, x_2) \in_{\mathrm{R}} D$.
By the assumption of f,

$$\begin{aligned} \frac{1}{2} + \epsilon &\leq P_{(U_n)^2} \\ &= (1-\delta)^2 P_{G^2} + (1-\delta)\delta P_{GH} + \delta(1-\delta) P_{HG} + \delta^2 P_{H^2}. \end{aligned} \tag{3}$$

Proof - III

Since $P_{G^2} \leq 1$ and $\epsilon > 2(1-\delta)^2$,

$$\frac{1}{2} + \frac{\epsilon}{2} < (1-\delta)\delta P_{GH} + \delta(1-\delta)P_{HG} + \delta^2 P_{H^2}.$$

Note $(1-\delta)\delta + \delta(1-\delta) + \delta^2 < 1$. One of the following holds

(i) $P_{GH} > \frac{1}{2} + \frac{\epsilon}{2}$

(ii) $P_{HG} > \frac{1}{2} + \frac{\epsilon}{2}$

(iii) $P_{H^2} > \frac{1}{2} + \frac{\epsilon}{2}$

Proof - IV

Assume WLOG (i). Then,

$$\Pr_{x_1 \in_{\mathrm{R}} G, x_2 \in_{\mathrm{R}} H}[C(x_1, x_2) = f(x_1) + f(x_2) \pmod 2)] > \frac{1}{2} + \frac{\epsilon}{2}.$$

By averaging, there is a fixed x_1 such that

$$\Pr_{x_2 \in_{\mathrm{R}} H}[C(x_1, x_2) + f(x_1) = f(x_2) \pmod 2)] > \frac{1}{2} + \frac{\epsilon}{2}.$$

Therefore is a circuit D of size s' such that

$$\Pr_{x \in_{\mathrm{R}} H}[D(x) = f(x)] > \frac{1}{2} + \frac{\epsilon}{2}.$$

A contradiction.
The same proof works for $k > 2$.

Proof - I

Assume: $H_{\text{avg}}^{1-\delta}(f) \geq s$ and $\epsilon > 0$.
Goal: To show the existence of δ-dense distribution H on which all circuits of size s' fail to compute better than $\frac{1}{2} + \epsilon$.
A game approach:
Alice:
Chooses a δ-dense distribution H,
Bob: Chooses a circuit of size s'.
Alice pays Bob:

$$p(H, C) = \Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)]$$

Alice wants to minimise $p(H, C)$, and Bob wants to maximise $p(H, C)$.
This is a zero-sum game.

Proof - II

Toward a contradiction. Then,
For every δ-dense distribution H chosen by Alice, Bob is able to choose a circuit C of size s' with which his payoff is:

$$\Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)] \geq \frac{1}{2} + \epsilon.$$

That is, Bob always has an advantage ϵ, whatever H Alice chose.
By the min-max theorem,

$$\min_{\mathcal{H}} \max_{\mathcal{C}} \Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)] = \max_{\mathcal{C}} \min_{\mathcal{H}} \Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)].$$

By the hypothesis,

$$\min_{\mathcal{H}} \max_{\mathcal{C}} \Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)] \geq \frac{1}{2} + \epsilon.$$

Proof - III

Therefore

$$\max_{\mathcal{C}} \min_{\mathcal{H}} \Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)] \geq \frac{1}{2} + \epsilon.$$

So, there is a distribution $\mathcal{C}$ of circuits of size s' such that for every H,

$$\Pr_{C \in_{\mathrm{R}} \mathcal{C}, x \in_{\mathrm{R}} H}[C(x) = f(x)] \geq \frac{1}{2} + \epsilon.$$

For $x \in \{0,1\}^n$, we say that x is “bad", if

$$\Pr_{C \in_{\mathrm{R}} \mathcal{C}}[C(x) = f(x)] < \frac{1}{2} + \epsilon,$$

and “good", otherwise.

Proof - IV

Lemma

The number of bad x's is $< \delta \cdot 2^n$.

Towards a contradiction. Let H be the uniform distribution of the bad x's. Then,

$$\Pr_{C \in_{\mathrm{R}} \mathcal{C}, x \in_{\mathrm{R}} H}[C(x) = f(x)] < \frac{1}{2} + \epsilon,$$

A contradiction.
For $t = \frac{50n}{\epsilon^2}$, pick circuits

$$C_1, C_2, \cdots, C_t$$

independently from $\mathcal{C}$,
define

$$C(x) = \text{majority}\{C_i(x) : 1 \leq i \leq t\}.$$

Then $|C| = t \cdot s' < s$.

Proof - V

Lemma

For every good $x \in \{0, 1\}^n$,

$$\Pr[C(x) \neq f(x)] < 2^{-n}.$$

If x is good, then for each i, $1 \leq i \leq t$, the probability that $C_i(x) = f(x)$ is $\geq \frac{1}{2} + \epsilon$.
For i, $1 \leq i \leq t$, define $X_i = 1$ if $C_i(x) = f(x)$, and 0, otherwise.
Let

$$X = \sum_{i=1}^{t} X_i.$$

By Chenoff bound, with probability $< 2^{-n}$, $X < \frac{1}{2}t$.

Proof - VI

Therefore,

$$\Pr_{x \in_{\mathrm{R}} \{0,1\}^n}[C(x) = f(x)] > 1 - \delta,$$

contradicting $H_{\text{avg}}^{1-\delta}(f) \geq s$.

A general method:
For every good x, for a random circuit C' of size s',
$C'(x) = f(x)$ with a nontrivial advantage ϵ.
Then there is a slightly larger circuit C that computes f with prob 1 on all good inputs.

The min-max theorem

Theorem

Let A be the payoff matrix of a zero-sum game. Then

$$\min_p \max_q qAp = \max_q \min_p qAp,$$

p, q are distributions of strategies.

Hardness amplification

Worst-case hardness

$\Downarrow$ (ECC) error correcting codes

Average case hardness.

Error correcting code

Definition

For $x, y \in \{0,1\}^m$, the fractional Hamming distance of x and y, written, $\Delta(x,y)$ is defined ny

$$\frac{1}{m}|\{i : x_i \neq y_i\}|.$$

For $\delta \in [0,1]$, $E : \{0,1\}^n \to \{0,1\}^m$, E is called an *error correcting code with distance* δ, if for every $x \neq y$,

$$\Delta(E(x), E(y)) \geq \delta.$$

$E(x)$: the *codeword* of x.

Intuition of ECC

Why ECC?

- To increase slightly the dimensionality allows us to amplify errors largely
- To amplify errors is to rectify the errors.
- Increasing errors amplifies hardness.

Existence of ECC

Lemma

For every $\delta < \frac{1}{2}$ and large n, there is a function $E : \{0,1\}^n \to \{0,1\}^m$ that is an ECC with distance δ for $m = n/(1 - H(\delta))$, where $H(\delta) = -\delta \log_2 \delta - (1-\delta)\log_2(1-\delta)$, the Shannon entropy of δ.

Proof

Each δ-ball in $\{0,1\}^m$ contains at most $o(1) \cdot 2^{H(\delta)n}$ elements.
$m = n/(1 - H(\delta))$, there are at least 2^n many δ-balls in $\{0,1\}^m$.
Random enumeration of the δ-balls will define an ECC E with distance δ.
Why $\delta < 1/2$?

High-dimensional geometry

The math principle of ECC is a high-dimensional geometry theorem:
The volume of a ball of radius r in m-dimensional space is approximately

$$\frac{\pi^{m/2}}{(m/2)!}r^m.$$

The volume increases exponentially as the dimensionality increases.

Efficient ECC

We will need explicitly defined ECC that are both efficiently encoded and decoded.

Decoding an ECC:
If $\Delta(E(x), y) < \frac{\delta}{2}$, then efficiently compute x.

Walsh-Hadamard code

$$\begin{aligned} WH : \{0,1\}^n &\to \{0,1\}^{2^n} \\ a &\mapsto \langle a \cdot x \rangle, x \in \{0,1\}^n. \end{aligned} \tag{4}$$

Lemma

WH is an ECC of distance $\frac{1}{2}$.

ECC over Σ

Given alphabet Σ, $x, y \in \Sigma^m$,

$$\Delta(x, y) = \frac{1}{m}|\{i : x_i \neq y_i\}|.$$

A function $E : \Sigma^n \to \Sigma^m$ is an ECC with distance δ over Σ if for $x \neq y$, $\Delta(E(x), E(y)) \geq \delta$.

Reed-Solomon code

Let $\mathbb{F}$ be a field and n, m numbers with $n \leq m \leq |\mathbb{F}|$. The Reed-Solomon code is

$$RS: \quad \mathbb{F}^n \to \mathbb{F}^m$$

$$(a_0, a_1, \cdots, a_{n-1}) \mapsto (z_0, z_1, \cdots, z_{m-1}),$$

where $z_j = \sum_{i=0}^{n-1} a_i f_j^i$, f_j is the jth element of $\mathbb{F}$.

Let

$$A(x) = \sum_{i=0}^{n-1} a_i x^i.$$

Then $z_j = A(f_j)$.

RS lemma

Lemma

The Reed-Solomon code $RS : \mathbb{F}^n \to \mathbb{F}^m$ *has distance* $1 - \frac{n}{m}$.

Reed-Muller code

RM maps a polynomial P of l variables and degree d to the values of the polynomial.
That is,

$$P \mapsto \langle P(x_1, \cdots, x_l) \rangle, \;\; x_1, \cdots, x_l \in \mathbb{F}$$

for each l variable degree d polynomial P.

Concatenated codes

Lagrange interpolation

For any set of pairs $(a_1, b_1), \cdots, (a_{d+1}, b_{d+1})$, there exists a unique polynomial $g(x)$ of degree at most d such that $g(a_i) = b_i$, for each $i \in \{1, 2, \cdots, d+1\}$.

Proof.

$$g(x) = \sum_{i=1}^{d+1} b_i \frac{\prod_{j \neq i}(x - a_j)}{\prod_{j \neq i}(a_i - a_j)}. \tag{5}$$

□

Unique decoding for Reed-Solomon

Theorem

There is a polynomial time algorithm that, given a list $(a_1, b_1), \cdots, (a_m, b_m)$ of pairs of elements of a finite field $\mathbb{F}$ such that there is a unique degree d polynomial $G : \mathbb{F} \to \mathbb{F}$ satisfying $G(a_i) = b_i$ for t of the numbers $i \in [m]$, where $t > \frac{m}{2} + \frac{d}{2}$, recovers G.

Let $t \geq \frac{m}{2} + \frac{d}{2} + 1$, let $L = \frac{m}{2} + \frac{d}{2}$, and $l = \frac{m}{2} - \frac{d}{2}$.
Set

$$C(x) = c_0 + c_1 x + \cdots + c_L x^L$$

$$E(x) = e_0 + e_1 x + \cdots + e_{l-1} x^{l-1} + e_l x^l$$

Proofs

For each $i \in [m]$, set

$$C(a_i) = b_i E(a_i)$$

This is a homogenous system of linear equations with m equations, and $m + 2$ unknowns. It must have nonzero solutions.
Solving the system, get polynomials $C(c)$ and $E(x)$.
Consider $C(x) - G(g)E(x)$.
The degree of the polynomial is $\frac{m}{2} + \frac{d}{2}$. However, for every i, if $G(a_i) = b_i$, then $C(a_i) - G(a_i)E(a_i) = 0$, for which the number of such i's is $t \geq \frac{m}{2} + \frac{d}{2} + 1$.
Therefore $C(x) - G(x)E(x) \equiv 0$. Set $P = \frac{C(x)}{E(x)}$. Then

$$P \equiv G.$$

Decoding and hardness amplification

Decoding: Given a string $x \in \{0,1\}^n$,

$$x \Rightarrow E(x) \Rightarrow \text{ corrupted } E(x) \Rightarrow x \tag{6}$$

Worst case hardness to mildly average case hardness: Given a function $f : \{0,1\}^n \to \{0,1\}$, interpreted as a string,

$$f \Rightarrow E(f) \Rightarrow \text{ computes } f \text{ with prob } 1-\rho \Rightarrow \text{perfectly computes } f \tag{7}$$

Local decoder

Let $E : \{0,1\}^n \rightarrow \{0,1\}^m$ be an ECC and let ρ and q be some numbers. A *local decoder* for E handling ρ errors is an algorithm D, that given random access to a string y this is ρ-close to some codeword $E(x)$ for some unknown x, and an index j, runs for poly $\log m$ time and outputs x_j with probability at least $\frac{2}{3}$.

$$x \Rightarrow E(x) \Rightarrow y : \text{corrupted } E(x) \Rightarrow x_j$$

- The decoder D is allowed to randomly read some bits of y only, the corrupted $E(x)$, where x is an unknown.
- The running time of D is poly $\log m$, m is the length of y.

Hardness amplification from local decoder

Theorem

Suppose that there exists an ECC with a poly time encoding algorithm and a local decoding algorithm handling ρ errors. Then for a function $f \in \mathrm{EXP}$ with $H_{\mathrm{wrs}}(f)(n) \geq s(n)$ for some $s(n) \geq n$. There is an $\epsilon > 0$ and $\hat{f}$ such that

$$H_{\mathrm{avg}}^{1-\rho}(\hat{f})(n) \geq (s(\epsilon n))^{\epsilon}.$$

Local decoder for WH

Theorem

For $\rho < \frac{1}{4}$, there exists a local decoder for the Walsh-Hadamard code handling ρ errors.

Input:
(i) $j \in [n]$,
(ii) random access to $f : \{0,1\}^n \to \{0,1\}$, where

$$\Pr_{y \in_R \{0,1\}^n}[f(y) \neq x \cdot y] \leq \rho$$

for $\rho < \frac{1}{4}$ and some unknown x.
Output: A bit b, that is expected to be x_j.

D for WH

The local decoder D proceeds:

1) Let e^j be the vector that is 1 in the jth bit, and 0 on the all other bits
2) Randomly pick $y \in \{0,1\}^n$
3) Query f for $f(y)$ and $f(y+e^j)$
4) Output $b = f(y) + f(y+e^j) \pmod 2$.

Then:

$$f(y) = x \cdot y \text{ with prob } 1-\rho$$

$$f(y+e^j) = x \cdot (y+e^j) \text{ with prob } 1-\rho$$

So with prob $1-2\rho$,

$$b = x_j.$$

Run several times, then with prob almost 1, $b = x_j$.

Computing the correct $f(x)$ from a corrupted f

Compute $f(x)$ as follows:

1. Randomly pick y
2. Let $b = f(y) + f(y + x) \pmod 2$

Then with prob $1 - 2\rho$,
b is the correct value of $f(x)$.
We say that f has the *self-correction property*.

Local decoder for Reed-Muller

Theorem

For every field F and numbers d, l, there is a local decoder for the Reed-Muller code handling $(1-\frac{d}{|F|})/6$ errors.
That is, there is a poly time algorithm D, that given random access to a function $f : F^l \to F$ that agrees with some degree d l variable polynomial P on $1-(1-\frac{d}{|F|})/6$ fraction of the inputs and $x \in F^l$, outputs $P(x)$ with prob $\geq \frac{2}{3}$.

Goal: Given $f : F^l \to F$ that agrees P on $1-(1-\frac{d}{|F|})/6$ fraction of the inputs, then we can compute $P(x)$, for each $x \in F^l$.

Local decoder - I

Let $\rho \leq (1 - \frac{d}{|F|})/6$.
Then

$$\Pr_{y \in_{\mathrm{R}} F^l}[P(y) \neq f(y)] < \rho,$$

where P is an unknown polynomial of degree d and l variables.
Input: $x \in F^l$
Output: $P(x)$.

Local decoder D

D proceeds as follows:

1) Pick a random line L_x, i.e., $z \in_{\mathrm{R}} F^l$, with

$$L_x = \{x + tz : t \in F\}$$

2) Query f to obtain a set of pairs

$$\{(t, f(x + tz)) \mid t \in F\}$$

– agrees with $P(x + tz)$ for many t's

3) Run Reed-Solomon decoding to gain a poly Q such that

$$Q(t) = f(x + tz)$$

for a large number of t's, which solves $P(x + tz)$.

4) Output $Q(0)$, expected to be $f(x)$.

Proofs - I

For $t \neq 0$, $x + tz$ is uniformly distributed, so the expected number of points $x + tz$ at which $f(x + tz) \neq P(x + tz)$ is $\rho|F|$. Define$X_t = 1$ if $f(x + tz) \neq P(x + tz)$, and 0, otherwise. Let $X = \sum_{t \neq 0, t \in F} X_t$. Then

$$E[X] = \rho \cdot |F|.$$

Proofs - II

By Markov inequality

$$\Pr[X > 3E[X]] < \frac{1}{3}.$$

So with prob $\geq \frac{2}{3}$,

$$X \leq 3E[X] = (1 - \frac{d}{|F|})|F|/2$$

By Reed-Solomon, with prob $\geq \frac{2}{3}$,

$$Q(t) \equiv P(x + tz),$$

in which case,

$$Q(0) = P(x).$$

Local decoder for concatenated codes

Reference

Chapters 8, 11, 19, 21, 22
Sanjeev Arora, Boaz Barak, Computational Complexity, A Modern approach, Cambridge University Press, 2010

New directions

1. Local algorithms for networks
   - Network algorithms must be local
2. Structure and algorithms for big data
   - The noisy big data require new theory of structures and algorithms

Grand challenge

We have seen that given a corrupted function f, it is possible to have an algorithm to compute the true f by random access the bits of the corrupted f. Here a function is either a linear function or a low degree polynomial.
What we can do for general functions?
More importantly, in nature, we often have a noisy data structure G, which is evolved by the rules, regulations and laws of specific objects, perturbed by random variations and noises. That is, G is a structured noisy data, consisting of a "true structure" T formed by laws and a noisy part N. A grand challenge is to define the true structure T of a noisy data G, excluding the perturbation by the random variations and noises occurred in G. Once such a true structure T of G is defined, we are able to discover the true knowledge of G.

Exercises 1

1. Let $X_1, \cdots, X_n$ be independent random variables such that X_i is equal to 1 with probability $1-\delta$ and equal to 0 with probability δ. Let $X = \sum_{i=1}^{n} X_i \pmod 2$. Prove that
$$\Pr[X = 1] = \frac{1}{2} + (1-2\delta)^k/2.$$
2. Prove that if there exists a δ-density distribution H such that $\Pr_{x \in_{\mathrm{R}} H}[C(x) = f(x)] \le \frac{1}{2} + \epsilon$ for every circuit c of size at most S with $S \le \sqrt{\epsilon^2 \delta 2^n / 100}$, then there exists a subset $I \subseteq \{0,1\}^n$ of size at least $\frac{\delta}{2} 2^n$ such that
$$\Pr_{x \in_{\mathrm{R}} I}[C(x) = f(x)] \le \frac{1}{2} + 2\epsilon$$
for every circuit C of size at most S.
3. Prove the hardcore lemma for general k.

Exercises 2

1. Let $f : \mathbb{F} \to \mathbb{F}$ be any function. Suppose integer $d \geq 0$ and number $\epsilon > 2\sqrt{\frac{d}{|\mathbb{F}|}}$. Prove that there are at most $2/\epsilon$ degree d polynomials that agree with f on at least an ϵ fraction of its coordinates.
2. Prove that if $Q(x, y)$ is a bivariate polynomial over some field $\mathbb{F}$ and $P(x)$ is a univariate polynomial over $\mathbb{F}$ such that $Q(P(x), x)$ is the zero polynomial, then $Q(x, y) = (y - P(x))A(x, y)$ for some polynomial $A(x, y)$.

Exercises 3

Linear codes We say that an ECC $E : \{0,1\}^n \to \{0,1\}^m$ is *linear*, if for every $x, x' \in \{0,1\}^n$, $E(x + x') = E(x) + E(x')$ (componentwise addition modulo 2). A linear ECC can be seen an $M \times n$ matrix A such that $E(x) = Ax$, thinking of x as a column vector.

1. Prove that the distance of a linear ECC is equal to the minimum over all nonzero $x \in \{0,1\}^n$ of the fraction of 1's in $E(x)$.
2. Prove that for every $\delta > 0$, there exists a linear ECC $E : \{0,1\}^n \to \{0,1\}^m$ for $m = n/(1 - H(\delta))$ with distance δ.
3. Prove that for some $\delta > 0$, there si an ECC $E : \{0,1\}^n \to \{0,1\}^{\mathrm{poly}(n)}$ of distance δ with poly time encoding, and decoding algorithms.