**03683170: Expanders, Pseudorandomness and Derandomization** 28/02/16

Lecture 1 – Definitions and basic results

*Amnon Ta-Shma and Dean Doron*

# 1 Resilient functions

In the *collective coin flipping problem* there are $n$ computationally unbounded players that can each broadcast a bit. At the end, some function is applied to the broadcasted bits and the output of this function is the output of the game. The goal of the game is to toss a (nearly) random bit. The catch is that some malicious coalition of players (say of size $q$) may wait to see what the honest players broadcast before broadcasting their own bit, i.e., their behavior is *malicious* and *adaptive* and depends on the values of the good players. The bad players win if after seeing the outcome of the good players they can force the value of the game to be either 0 or 1 as they wish. Since we allow only one round in the above protocol, we can formalize the above:

**Definition 1.** *Let $f$ be any boolean function on $x_1, \ldots, x_n$. The influence of $Q \subseteq [n]$, $I_Q(f)$, is defined to be the probability that $f$ is undetermined after fixing the variables outside $Q$ uniformly at random. For $q > 0$ we define $I_q(f) = \max_{Q \subseteq [n], |Q|=q} I_Q(f)$.*

**Definition 2.** *Let $f$ be any boolean function on $x_1, \ldots, x_n$ and $q$ any integer. We say that $f$ is $(q, \varepsilon)$-resilient if $I_q(f) \leq \varepsilon$.*

The constant function $f$ has $I_1(f) = 0$ and no player has influence. A dictatorship $f(x_1, \ldots, x_n) = x_i$ for some $i$, has $I_1(f) = 1$ and the dictator has full influence. A natural question is how resilient *balanced* functions can be, and whether we can find these resilient functions and force other properties on them (like monotonicity).

## 1.1 The case of one malicious player

**Example 3** (The majority function)**.** $\mathrm{MAJ}(x_1, \ldots, x_n) = 1$ *iff* $\sum_{i=1}^{n} x_i > \frac{n}{2}$.

If we let $X \in [0, n-1]$ be the Hamming weight of the coins tossed by the $n-1$ honest players, then the majority is undetermined if and only if $\frac{n}{2} - 1 \leq X \leq \frac{n}{2} + 1$. However, when $n$ goes to infinity, $\binom{n}{n/2 \pm 1} \approx \binom{n}{n/2} \approx \sqrt{\frac{2}{\pi}} \frac{2^n}{\sqrt{n}}$, so:

**Theorem 4.** *As $n$ goes to infinity, $I_1(\mathrm{MAJ}) = \Theta(\frac{1}{\sqrt{n}})$.*

We note that, similarly, for a fixed $q$ and as $n$ goes to infinity, $I_q(\mathrm{MAJ}) = \Theta(\frac{q}{\sqrt{n}})$.

Another example is:

**Example 5** (Iterated majority)**.** *Suppose $n = 3^k$ and define $\mathrm{IMAJ}_3(x_1, \ldots, x_n)$ as follows. On inputs of length 3, $\mathrm{IMAJ}_3$ is simply the majority function. On inputs of length $3^k$ for $k > 1$,*

$$\mathrm{IMAJ}_3(x_1, \ldots, x_n) = \mathrm{MAJ}\left(\mathrm{IMAJ}_3(x^{(1)}), \mathrm{IMAJ}_3(x^{(2)}), \mathrm{IMAJ}_3(x^{(3)})\right)$$

*where $x^{(1)}, x^{(2)}, x^{(3)}$ is a partition of the bits into three consecutive blocks of equal length.*

The iterated majority is more resilient than the majority function:

**Theorem 6.** *As $n$ goes to infinity, $I_1(\mathrm{IMAJ}_3) \leq \frac{1}{n^\alpha}$, where $\alpha = \log_3 2 \approx 0.63$.*

Similarly, for a fixed $q$, $I_q(\mathrm{IMAJ}_3) \leq \frac{q}{n^\alpha}$. We leave the proof as an exercise.

Our next example is the Tribes function. For a given $b$ let $n$ be the first multiple of $b$ for which $n \geq b \cdot 2^b$. It is easy to check that $\log n - \log\log n \leq b \leq \log n - \log\log n + 1$.

**Example 7** (The Tribes function)**.**

$$\mathrm{Tribes}(x_1, \ldots, x_n) = \bigvee_{j=1}^{n/b} \bigwedge_{k \in \{(j-1)b+1, \ldots, jb\}} x_k.$$

The Tribes function is even more resilient:

**Theorem 8** ([3])**.** $I_q(\mathrm{Tribes}) = O\left(\frac{\log n}{n}\right)$.

*Proof.* It is sufficient to look at a specific coordinate $i$. $x_i$ determines Tribes when everyone else in its conjunct is TRUE and all other conjuncts evaluates to FALSE. As every coordinate is in exactly one conjunct, these two events are independent. The probability of the first event is $2^{-(b-1)}$, The probability of the second event is $(1-2^{-b})^{\frac{n}{b}-1}$. Multiplying, we obtain:

$$2^{-(b-1)}(1-2^{-b})^{\frac{n}{b}-1} \leq 2 \cdot \frac{2^{-b}}{1-2^{-b}} \cdot e^{-\frac{n \cdot 2^b}{b}} \leq 2^{-b} e^{-1} \leq \frac{\log n}{2e}.$$

□

After reviewing the above constructions, one can hope to find a function with influence that is $O(n^{-1})$ for $q = 1$. However, Kahn, Kalai and Linial proved:

**Theorem 9** ([14])**.** *For every* balanced *boolean function $f$ on $n$ variables, $I_1(f) = \Omega(\frac{\log n}{n})$.*

In fact, the proof also holds for unbalanced function (e.g., the constant function) if we account for the *bias* of $f$. Define $bias(f) = \Pr_x[f(x) = 0] - \Pr_x[f(x) = 1]$. Then,

**Theorem 10** ([14])**.** *For every boolean function $f$ on $n$ variables, $I_1(f) = \Omega(\frac{\log n}{n} \cdot (1 - bias(f)^2))$.*

The proof uses Fourier Analysis, which is a powerful and important proof technique. We will not show the proof in full in class. Instead, we will talk about the basics and prove the theorem with the simple (almost trivial) bound $\frac{1}{n}$ instead of the tight bound $\frac{\log n}{n}$. However, the proof will be covered in the seminar ("A seminar on pseudo-randomness", Sunday, 15:00-17:00, Orenstein 111 – you are all welcome).

## 1.2 The case of many malicious player

We will later need a

1. Monotone,

2. almost balanced function,

3. that has a small depth circuit, and,

4. is resilient against coalitions of size $n^\alpha$ for $\alpha > 1/2$.

Let us check the functions we have seen so far:

The majority function is resilient against coalitions of size $n^{1/2-\varepsilon}$, but not against coalitions of size $n^{1/2+\varepsilon}$ (why?). The majority function is monotone and balanced, but it does not have a small depth circuit.

The Tribes function is monotone, almost balanced, and has a low depth circuit. The Tribes function has high resiliency when considering coalitions of size 1. However, verify to yourself that already $I_b(Tribes) = \Omega(1)$. Thus, coalitions of about $\log n$ players have high influence over the result.

We would like to generalize the Tribes function to be resilient against many malicious players. This was first done in a *non-constructive* way by Ajtai and Linial. We will later need and discuss *explicit* variants of the construction.

Ajtai and Linial generalized the Tribes function by considering arbitrary partitions, and arbitrary assignments to the variables. Specifically, let $P$ be a partition of $[n]$ into $\frac{n}{b}$ equal-sized disjoint sets $P_1, \ldots, P_{n/b}$ each of length $b$. Let $g : [n] \to \{0,1\}$ be a boolean function. We define

**Example 11** (The generalized Tribes function).

$$\text{Tribes}(P, g)(x_1, \ldots, x_n) = \bigvee_{j=1}^{n/b} \bigwedge_{k \in P_j} \delta_{x_k, g(k)},$$

*where $\delta$ is the Kronecker delta function.*

Thus, the Tribes function in this notation corresponds to the partition of the $n$ variables to $\frac{n}{b}$ consecutive blocks and using the constant function $g \equiv 1$. Ajtai and Linial proved that if we pick uniformly at random and independently $n$ partitions $P^{(1)}, \ldots, P^{(n)}$ and $n$ functions $g^{(1)}, \ldots, g^{(n)}$ and define:

$$f(x_1, \ldots, x_n) \;=\; \bigwedge_{i=1}^{n} \text{Tribes}(P^{(i)}, g^{(i)})(x_1, \ldots, x_n)$$

then, w.h.p. (over the choice of the partitions and the functions) we get a highly resilient function.

**Theorem 12** ([1]). *There exists a boolean function $f$ on $n$ variables such that for every $\varepsilon > 0$, it holds that $I_q(f) = O(\varepsilon)$ for $q = \frac{\varepsilon n}{\log^2 n}$.*

The proof uses the probabilistic method and is rather complicated. If we take $g \not\equiv 1$, then the resulting function is almost balanced, has a small depth circuit and high resiliency, but is not monotone because the values $g(i) = 0$ correspond to taking $\neg x_i$ in the circuit. However, if we fix the functions $g^{(i)}$ to be constant 1 we do get a monotone small depth circuit.

Chattopadhyay and Zuckerman [4] showed how to make the Ajtai-Linial construction *monotone* and *explicit* for a small loss in the parameters, and Meka later improved this:

**Theorem 13** ([18]). *For some constant $c > 0$ the following holds. There exists an explicit depth-three monotone function $f : \{0,1\}^n \to \{0,1\}$ which can be computed in time $n^c$ such that*

- *$f$ is almost balanced:* $\Pr_{x\sim U_n}[f(x)] = \frac{1}{2} \pm \frac{1}{10}$.
- *$f$ has small influence:* $I_q(f) \leq \frac{c\log^2 n}{\log n} \cdot q$.

As we said we will later need this construction. If you are interested in the construction (which also involves extractors) talk to us, and present it in the seminar.

Finally, we show the resiliency cannot be much improved, by using the KKL lower bound (Theorem 10) on $I_1$, and extending the result to many malicious players:

**Theorem 14.** *Fix a balanced, monotone, boolean function $f$. Then, there exists a coalition $J_0$ (resp. $J_1$) of size $q = O(\frac{n}{\log n})$ that can force the function to be $0$ (resp. $1$) with probability at least $0.9$. In particular, $f$ is not $(2q, 0.8)$-resilient.*

*Proof.* First notice that for a monotone function $f$, if a subset $J \subseteq [n]$ wants to push $f$ to be 0 (resp. 1), then its best strategy is to chose the value 0 (resp. 1) for all its members. Now, assume towards contradiction that there exists such a balanced monotone function $f$ and let us prove the existence of $J_0$.

Pick the most influential variable in $f_n = f$. W.l.o.g it is $x_n$. Then define $f_{n-1}(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, 0)$. Pick the most influential variable in $f_{n-1}$ and w.l.o.g it is $x_{n-1}$. Repeat this process $T$ times defining $f_k(x_1, \ldots, x_k) = f(x_1, \ldots, x_k, 0, \ldots, 0)$ and $p_k = \Pr_x[f_k(x) = 1]$, until $p_T \leq 0.1$. Set $J_0 = \{n - T + 1, \ldots, n\}$.

Denote $p_k = \Pr_x[f_k(x) = 1]$ and $e_k = I_{\{x_k\}}(f_k)$. Then, by KKL, $e_k = \Omega(\frac{\log n}{n} \cdot (1 - bias^2(f_k))) = \Omega(\frac{\log n}{n} \cdot p_k)$ (Check!). As $p_k \geq 0.1$, $e_k = \Omega(\frac{n}{\log n})$. Also, $p_{k+1} = p_k - \frac{e_k}{2}$. Hence, it follows that $T \leq O(\frac{n}{\log n})$. □

We remark that the theorem also holds for non-monotone functions, and that a variant of it can be proved to coalitions that control the function with probability close to 1, see Ryan Odonnell's lecture notes. In particular, linear coalitions (and even those of size substantially more that $\frac{n}{\log n}$) cannot be tolerated in one round protocols.

### 1.3 Another name for the same object: Extractors for non-oblivious sources

**Definition 15.** *For random variables $X$ and $Y$ taking values in $\mathcal{U}$, their statistical distance is $|X - Y| = \max_{T\subseteq \mathcal{U}} |\Pr[X \in T] - \Pr[Y \in T]|$. We say that $X$ and $Y$ are $\varepsilon$-close if $|X - Y| \leq \varepsilon$.*

**Definition 16.** *Let $m \leq n$ be integers and let $\varepsilon \geq 0$. Let $E : \{0,1\}^n \to \{0,1\}^m$ be a function and $\mathcal{C}$ be a class of probability distributions over $\{0,1\}^n$. We say that $E$ is an extractor for $\mathcal{C}$ with $\varepsilon$ error if for every $X \in \mathcal{C}$, the distribution $E(X)$ is $\varepsilon$-close to $U_m$.*

**Definition 17.** *Let $q \leq n$. A random variable $X$ over $\{0,1\}^n$ is an $(n, q)$ non-oblivious bit-fixing source if there exists some $Q \subseteq [n]$ of size $q$ and functions $g_i : \{0,1\}^{n-q} \to \{0,1\}$ for $i \in [q]$, where the distribution $X$ is chosen as follows:*

- *First, the $n - q$ bits outside $Q$ are uniformly chosen, say, $a_1, \ldots, a_{n-q}$,*
- *Then, for every $i \in Q$, the value of the $i$-th bit is set to $g_i(a_1, \ldots, a_{n-q})$.*

Notice that if $f$ is a balanced $(q, \varepsilon)$ resilient function, then $f : \{0,1\}^n \to \{0,1\}$ is an extractor for $(n, q)$ non-oblivious bit fixing sources with error $\varepsilon$.

Note that the new terminology allows outputting many output bits, and this raises the natural question of how many bits can be extracted from a non-oblivious source.

We also have the weaker notion of a disperser.

**Definition 18.** *Let $m \le n$ be integers and let $\varepsilon \ge 0$. Let $E : \{0,1\}^n \to \{0,1\}^m$ be a function and $\mathcal{C}$ be a class of probability distributions over $\{0,1\}^n$. We say that $E$ is a disperser for $\mathcal{C}$ with $\varepsilon$ error if for every $X \in \mathcal{C}$, $|\mathrm{Supp}(E(X))| \ge (1-\varepsilon)2^m$.*

Note that an extractor with $\varepsilon$–error is also a disperser with $\varepsilon$–error, but not necessarily vice versa.

### 1.4 Beyond one round

One can also ask whether having many rounds can improve the resiliency of a function. Intuitively, having many rounds make life much harder for the malicious players, as now they are forced to take decisions based on partial knowledge, whereas in the one round case they have full knowledge. Indeed, having many rounds does help, and with $O(\log n)$ rounds one can handle any $\frac{1}{2} - \varepsilon$ fraction of malicious players. Denote $I_q(P)$ as the maximal probability that $q$ malicious players determine the output of a protocol $P$.

**Theorem 19.** *[11] For every $n$ and $\delta > 0$, fix $q = (\frac{1}{2} - \delta)n$. There exists a protocol $P$ with $\log n$ rounds, such that $I_q(P) = 1 - \Omega(\delta^{1.65})$.*

We will see the above result in the exercise, and already saw that the above result cannot be achieved in a single round. Regarding the tightness of the parameters:

- The fraction $\frac{1}{2} - \varepsilon$ of malicious players is almost optimal, as no protocol can handle malicious majority regardless of the number of rounds [21].
- The number of rounds, $\log n$ in the theorem, may not be optimal, the best lower bound known today is:

  **Theorem 20.** *[20] Fix $\varepsilon > 0$ and $q = \alpha n$ for some constant $\alpha < \frac{1}{2}$. If a protocol $P$ has $I_q(P) \le \varepsilon$ then it has at least $r \ge (\frac{1}{2} - \varepsilon) \log^\star n$ rounds.*

A lower bound on the influence exists for many rounds as well.

**Theorem 21.** *[3] For every protocol $P$ in whatever number of rounds, and every $q$, $I_q(P) = \Omega(\frac{q}{n})$.*

## 2 Extractors for oblivious bit-fixing sources

**Definition 22.** *Let $q \le n$. A random variable $X$ over $\{0,1\}^n$ is an $(n, q)$* oblivious *bit-fixing source, if there exists $Q \subseteq [n]$ of size $q$ such that bits in $Q$ are fixed (to some arbitrary unknown value) and bits outside $Q$ are uniform.*

Oblivious bit-fixing sources naturally arise in Cryptography. For example assume a scenario where Alice and Bob share a uniformly chosen random key $K \in \{0,1\}^n$, and at some stage some of the bits are compromised to an adversary Eve. Suppose Alice and Bob know that at most half of the bits were compromised but do not know which. Can they still make use of their shared partially compromised private key? Indeed, if $E : \{0,1\}^n \to \{0,1\}^m$ is an extractor for $(n, \frac{n}{2})$ oblivious bit-fixing sources, then $K' = E(K)$ is close to uniform even given the knowledge Eve has (see also Exercise 1).

It is trivial that the parity function $\bigoplus_{i=1}^n x_i$ is a zero-error extractor for $(n,q)$ oblivious bit-fixing sources for every $q \leq n-1$. A straight forward generalization is to partition the bits into $\frac{n}{q+1}$ groups each with $q+1$ bits, and outputting the parity of each set. This gives $n/q+1$ output bits. However, this does not work when $q > n/2$, and, in fact, Chor et al. [6] showed that when $q > \frac{2}{3}n$, it is impossible to extract even *two* truly random bits by a linear extractor.

In the solutions so far, the output bits were completely uniform. Now, we are going to abandon this property and we ask whether we can output many bit that are close to uniform:

**Theorem 23.** *Let $n$ be an integer, and $\varepsilon \geq 0$. For every $q \leq n - \log n - 2\log\frac{1}{\varepsilon} - O(1)$ and $m \leq n - q - 2\log\frac{1}{\varepsilon} - O(1)$ there exists an $\varepsilon$-error extractor $E : \{0,1\}^n \to \{0,1\}^m$ for $(n,q)$ oblivious bit-fixing sources.*

*Proof.* As an exercise. □

We now show a construction by Kamp and Zuckerman [15] that extracts many bits from an oblivious bit fixing source. For the construction we shall need the notion of *expanders*. Expanders are graphs with two seemingly contradictory properties: They are sparse, but also well connected. They come in two flavors: *combinatorial* and *algebraic*. For this application we need the algebraic variant. We shall later meet with them again (and discuss more).

## 2.1 A first encounter with expanders

Let $G$ be an undirected, $d$-regular graph on $n$ vertices. Let $A$ be its normalized adjacency matrix, i.e.,

$$A_{i,j} = \begin{cases} 0 & (i,j) \notin E \\ 1/d & (i,j) \in E \end{cases}$$

$A$ is a Hermitian matrix, hence it has an orthonormal basis $\{v_1, \ldots, v_n\}$ with real eigenvalues $\lambda_1, \ldots, \lambda_n$. W.l.o.g., let us arrange the eigenvalues in decreasing order, i.e., $\lambda_n \leq \ldots \leq \lambda_1$.

**Exercise 24.** *Let $G$ be an undirected, regular graph. Prove:*

- *The vector $\mathbf{1}$ is an eigenvector of $A$ with eigenvalue 1.*
- *For all $i \in [n]$, $|\lambda_i| \leq 1$.*
- *$G$ is connected iff $\lambda_2 < 1$.*
- *A connected graph $G$ is bipartite iff $\lambda_n = -1$.*

**Exercise 25.** *Let $G$ be a cycle on $n$ vertices. Show that $\{\chi_k\}_{k=0}^{n-1}$ is an orthonormal eigenvector basis for $G$ with eigenvalues $\cos\left(\frac{2\pi k}{n}\right)$, where $\chi_k(i) = \omega^{ki}$ for $\omega$ a primitive $n$-th root of unity.*

**Definition 26.** *A regular, undirected graph $G = (V, E)$ is a $\lambda$-expander if the second largest eigenvalue (in absolute value) of its normalized adjacency matrix in absolute value is at most $1 - \lambda$.*

The gap $\lambda$ between the largest eigenvalue 1, and the second largest eigenvalue is called the *spectral gap* of the graph. Intuitively (and as we shall later see, also formally) the larger the gap is, the more "connected" the graph is. Usually, we call a family of $d$-regular graphs $\{G_n\}_{n \in \mathbb{N}}$ an expander family, if the number of vertices in the family goes to infinity, yet the spectral gap stays constant.

**Exercise 27.** *consider the family $C = \{C_n\}$ where $C_n$ is an undirected cycle on $n$ vertices. Is $C$ an expander?*

Alon and Boppana proved that:

**Theorem 28** ([2])**.** *Let $d \geq 3$ be a constant and $G = \{G_n\}$ a family of undirected graphs where $G_n$ is a $d$-regular graph on $n$ vertices. Then, $\lim_{n \to \infty} \lambda_2(G_n) \geq \frac{2\sqrt{d-1}}{d}$.*

Random graphs almost match this bound as well, in the sense that in a family $\{G_n\}$ with $G_n$ being a random $d$-regular graph on vertices, w.h.p. (over the choice of the graphs) the second largest eigenvalue tends to the above limit (see [12]).

A graph $G$ is called *Ramanujan* if $\lambda_2(G) = \frac{2\sqrt{d-1}}{d}$ (notice that this is better than just just having the limit tend to the bound). Remarkably, Margulis [17] and later on Lubotzky, Philips and Sarnak [16] showed explicit constructions of Ramanujan graphs.

## 2.2 An extractor for oblivious bit-fixing sources with many output bits

We are now ready to present the Zuckerman-Kamp construction. Fix $d$ and consider the more general notion of a $(n, q, d)$ *symbol fixing source* – an $(n, q)$ oblivious bit-fixing source where instead of bits, the variables are taken from an alphabet $\Sigma$ of size $d$. That is, we have $q$ symbols that are fixed to some constants from $\Sigma$ and the rest are uniformly chosen from $\Sigma$.

We prove:

**Theorem 29.** *[15] For every $d > 2$ there exists a constant $\alpha < 1$ such that for every $n$ and $q$ there exists an explicit extractor $E : \Sigma^n \to \Sigma^m$ for $(n, q, d)$ symbol-fixing sources with error $\varepsilon$ and $m \geq 2\alpha(n - q) - 2\log_d \frac{1}{\varepsilon}$.*

*Proof.* Let $G$ be a $d$-regular, labeled, expander graph on $M = d^m$ vertices and denote $\bar{\lambda} = \max\{\lambda_2(G), -\lambda_M(G)\} \leq d^{-\alpha}$ for some constant $\alpha$. Furthermore, let us assume that the graph is consistently labeled in the following sense: from every vertex there is exactly one edge leaving $v$ labeled $i$, and exactly one edge entering $v$ labeled $i$ (Justify why such a labeling always exist! Also, many explicit constructions already come with such a labeling). Let $v_0$ be some fixed vertex in $G$. Given $x \in \Sigma^n$, $E(x)$ is obtained by:

1. Start at $v_0$.
2. Take a walk of length $n$ on $G$ according to the symbols of $x$.
3. Output the label of the final vertex on the walk.

The intuition, that we will make precise soon, is that the uniform symbols mix us fast in the graph, while the fixed symbols do not hurt us. Therefore, the random walk behaves essentially like a random walk on the random steps only. Because of the rapid mixing properties of expanders, the output will be close to uniform.

Let us first consider a random step:

**Claim 30.** *Let $A$ be a transition matrix of an undirected $d$-regular graph $G$ on $M$ vertices and let $\pi = \frac{1}{M}\mathbf{1}$. Then, for every probability distribution $p$, $\| Ap - \pi \|_2 \leq \bar{\lambda} \| p - \pi \|_2$.*

*Proof.* Let $v_1 = \frac{1}{\sqrt{M}}\mathbf{1} = \sqrt{M}\pi, v_2, \ldots, v_M$ be the orthonormal, eigenvector basis of $A$, with eigenvalues $\lambda_1 = 1 \geq \lambda_2 \geq \ldots \geq \lambda_M$. Express $p = \sum_{i=1}^{M} \alpha_i v_i$. Then $\alpha_1 = \langle p, v_1 \rangle = \frac{1}{\sqrt{M}}$ and $p - \pi = \sum_{i=2} \alpha_i v_i$ (why?). It follows that:

$$\begin{aligned} \| Ap - \pi \|_2 &= \| A(p-\pi) \|_2 = \left\| A \sum_{i=2} \alpha_i v_i \right\|_2 = \left\| \sum_{i=2}^{M} \alpha_i \lambda_i v_i \right\|_2 \\ &\leq \bar{\lambda} \sqrt{\sum_{i=2}^{M} |\alpha_i|^2} = \bar{\lambda} \| p - \pi \|_2 . \end{aligned}$$

□

We now consider what happens in the fixed steps. As the graph is consistently labeled, the transition matrix associated with this step is a permutation matrix, hence, a unitary operator. In particular

$$\| Ap - \pi \|_2 = \| A(p-\pi) \|_2 = \| p - \pi \|_2 .$$

Together this implies that:

**Lemma 31.** *Let $A$ be the transition matrix of $G$ on $M$ vertices and $\pi = \frac{1}{M}\mathbf{1}$. Consider an $n$-steps walk on $G$, where the steps follow an input $x$ from an $(n, q, d)$ symbol-fixing source. Then, for every initial probability distribution $p$ over the vertices,*

$$\left\| \prod_{i=1}^{n} A_i p - \pi \right\|_2 \leq \bar{\lambda}^{n-q} \| p - \pi \|_2 ,$$

*where $A_i$ is $A$ if the $i$-th player is honest and an appropriate permutation matrix otherwise.*

In particular, $\| \prod_{i=1}^{n} A_i p - \pi \|_2 \leq \bar{\lambda}^{n-q}$ (why?). By Cauchy-Schwartz, $| \prod_{i=1}^{n} A_i p - \pi | \leq \sqrt{M} \bar{\lambda}^{n-q} = \varepsilon$. Plugging in $M = d^m$ and $\bar{\lambda} \leq d^{-\alpha}$, we have:

$$\varepsilon \leq d^{\frac{m}{2} - \alpha(n-q)}.$$

Taking the logarithm of both sides, we obtain what is desired. Note that expanders exist with $\alpha \approx \frac{1}{2}$ so we can extract almost all the entropy. □

What about *bit-fixing* sources, where $d = 2$? Regular, undirected graphs of degree 2, are disjoint union of cycles, and, generally, are not good expanders (see Exercise 25). However, let us nevertheless try the cycle. By the same exercise, for a cycle $G$ on an odd number $M$ of vertices, $\bar{\lambda} = \cos(\frac{2\pi}{M}) \leq e^{-\frac{\pi^2}{M^2}}$. Thus, the incurred error this time is at most $\sqrt{M}\bar{\lambda}^{n-q} \leq \sqrt{M}e^{-\frac{\pi^2(n-q)}{M^2}}$. This gives:

**Theorem 32.** *[15] For every $n$, $q$ and odd $M$ there exists an explicit extractor $E : \{0,1\}^n \to [M]$ for $(n,q)$ bit-fixing sources with error $\varepsilon = \sqrt{M}e^{-\frac{\pi^2(n-q)}{M^2}}$.*

Note that because of the inferior expansion, if we want constant error $\varepsilon$ we manage to extract only $\log M = O(\log(n-q))$ bits.

If we look again at the construction we see that if we interpret $x_i \in \{-1,1\}^n$ we simply output $\sum_{i=1}^{n} x_i \bmod M$. It is clear that the fixed sum is a fixed shift, and so it is enough to prove that the uniform part converges to uniform. Roughly speaking, we need $n > M^2$ so that the "drunken walk" converges to uniform (assuming we have an odd number of vertices), i.e., $M \approx \sqrt{n}$ and $m = \log(M) \approx \frac{1}{2}\log n$. When $M$ is even we never converge to uniform as the graph is bipartite and the number of the round determines the parity of the vertex (whether it is an odd or even vertex).

# 3 Two independent sources

## 3.1 Independent source extractors and dispersers

**Definition 33.** *Let $X$ be a random variable over $\mathcal{U}$. The* min-entropy *of $X$ is given by $H_\infty(X) = \min_{x\in\mathcal{U}} \log \frac{1}{\Pr[X=x]}$. If $X$ is distributed over $\{0,1\}^n$ and has min-entropy at least $k$, we say that $X$ is an $(n,k)$ source. Also, we say a random variable is flat if it is uniform over its support.*

**Exercise 34.** *An $(n,k)$ source $X$ can be expressed as a convex combination of flat sources each with at least $k$ min-entropy.*

**Definition 35.** *A function $E : (\{0,1\}^n)^\ell \to \{0,1\}^m$ is a $(k,\varepsilon)$ extractor for $\ell$ independent sources if for every independent $X_1,\ldots,X_\ell$ such that $H_\infty(X_i) \geq k$ it holds that $E(X_1,\ldots,X_\ell)$ is $\varepsilon$-close to $U_m$.*

**Exercise 36.** *Show that there is no 1-source extractor. Namely, for every $E : \{0,1\}^n \to \{0,1\}$ there exists an $(n, k = n-1)$ source $X$ such that $f(X)$ is fixed.*

We can also consider the more general setting where the length of each source or its min-entropy need not be the same among all sources.

A probabilistic argument shows the existence of two-source extractor for $k = O(\log n + \log\frac{1}{\varepsilon})$. We leave it as an exercise.

Next, we show a simple explicit construction by Chor and Goldreich, that shows that if the sum of min-entropies is larger than $n$ then the inner-product function is a two-source extractor.

**Theorem 37.** *[5] For every $n$ and $k > \frac{n}{2}$, $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ given by $E(x,y) = \langle x,y\rangle \bmod 2$ is a $(k, 2^{-(k-n/2)})$ two-source extractor.*

*Proof.* Let $X$ and $Y$ be two $(n,k)$ sources and assume w.l.o.g. (why?) that they are flat. Thus, we think of $X$ and $Y$ as two subsets of $\{0,1\}^n$ of size $2^k$ each.

Let $H$ be the $2^n \times 2^n$ graph defined by $H_{x,y} = (-1)^{\langle x,y \rangle}$ ($H$ is called the Hadamard matrix). Then,

$$\frac{1}{|X| \cdot |Y|} \, \mathbf{1}_X H \mathbf{1}_Y \;=\; \Pr_{x \sim X, y \sim Y}[E(x,y) = 0] - \Pr_{x \sim X, y \sim Y}[E(x,y) = 1],$$

where $\mathbf{1}_X$ (resp. $\mathbf{1}_Y$) is the characteristic vector of $X$ (resp. $Y$). Now, using the fact that $\| H \| = \sqrt{2^n}$ (see Exercise 1),

$$\frac{|\mathbf{1}_X H \mathbf{1}_Y|}{|X| \cdot |Y|} \;\leq\; \frac{\| \mathbf{1}_X \|_2 \cdot \| H \|_2 \cdot \| \mathbf{1}_Y \|_2}{|X| \cdot |Y|} \;=\; \frac{\sqrt{|X|}\sqrt{|Y|}}{|X| \cdot |Y|}\sqrt{2^n} \;=\; \sqrt{\frac{2^n}{|X| \cdot |Y|}}.$$

□

The above construction can be generalized to extract more than one bit.

### 3.2 Ramsey graphs and 2-source dispersers

Ramsey theory is a branch of combinatorics that studies the unavoidable presence of local structure in globally-unstructured objects (or the appearance of some order in global disorder). In a pioneer paper, Ramsey considered an instantiation of this phenomena in graph theory.

**Definition 38.** *A graph on $n$ vertices is called $K$-Ramsey if it contains no clique or independent set of size $K$.*

Ramsey (in the context of propositional logic), followed independently by Erdös and Szekeres, showed:

**Theorem 39** ([19, 10])**.** *There is no graph on $n$ vertices that is $\frac{1}{2}\log n$-Ramsey.*

Erdös later proved:

**Theorem 40** ([9])**.** *There exists a graph on $n$ vertices that is $2\log n$-Ramsey.*

We will see both theorems in the exercise.

Ramsey graphs have an analogous definition for bipartite graphs. A bipartite graph on two sets of $n$ vertices is a bipartite $K$-Ramsey if it has no $K \times K$ complete or empty bipartite subgraph.

**Exercise 41.** *Suppose you are given a bipartite graph on $N$ vertices that is $K$-Ramsey. Show how to construct from it an undirected graph on $N$ vertices that is $2K$-Ramsey.*

Thus, constructing bipartite Ramsey-graphs is at least as hard as constructing Ramsey graphs, and it is believed to be a strictly harder problem. We now see that the bipartite Ramsey problem is a subcase of the two-source extraction problem.

Recall that a 2-source disperser $E : (\{0,1\}^n)^2 \to \{0,1\}^m$ is the weaker analogue of a 2-source disperser, where for any two independent sources $X$ and $Y$ with min-entropy at least $k$, the size of the support of $E(X,Y)$ is at least $(1-\varepsilon)2^m$ for some error parameter $\varepsilon$. It is not hard to see the connection between bipartite Ramsey graphs and zero-error 2-source dispersers.

**Claim 42.** *An explicit $(k,0)$ 2-source disperser $E : (\{0,1\}^n)^2 \to \{0,1\}$ translates into an explicit construction of a bipartite $2^k$-Ramsey graph on $2^n$ vertices.*

*Proof.* Let $E$ be such a disperser. Use $E$ to define the bipartite graph $G_E = (S = \{0,1\}^n, T = \{0,1\}^n, E)$ and $\{u,v\} \in E$ iff $E(u,v) = 1$. $G_E$ then has the property that for every two sets $A, B \subseteq \{0,1\}^n$ of size $2^k$, $E(A,B) = \{0,1\}$, meaning that every induced $2^k \times 2^k$ sub-graph is neither empty nor complete. □

The probabilistic construction in Exercise 1 shows that not only bipartite Ramsey graphs exist (i.e., graphs in which the edges between every two sets of size $K$ are not mono-chromatic) but also almost balanced Ramsey graphs exist (i.e., graphs in which every two sets of size $K$, have about the same number of red and blue edges between them) and for about the same parameters.

Constructing explicit Ramsey graphs (or bipartite Ramsey graphs, two-source dispersers or two-source extractors) is a great challenge. The current record is $k = 2^{(\log\log n)^{O(1)}}$, by Cohen [8], and using different techniques [4]. We will cover (hopefully) in this class the second construction. The challenge of constructing $O(\log n)$-Ramsey graphs is roughly equivalent to constructing 2-source dispersers for min-entropy $\log n + O(1)$, and dispersers for min-entropy $O(\log n)$ would yield $\mathrm{poly}(\log n)$-Ramsey graphs.

### 3.3 Extractors for affine sources

Let $\mathbb{F}$ be a finite field. A set $X \subseteq \mathbb{F}^n$ is an affine subspace if $X = V + b$ for some vector $b \in \mathbb{F}^n$ and subspace $V \subseteq \mathbb{F}^n$. The dimension of the affine space is the dimension of $V$. A distribution $X$ over $\mathbb{F}^n$ is a $k$–affine source if it is uniformly distributed over some affine space of dimension $k$.

Using the probabilistic method:

**Theorem 43.** *There exists an affine extractor $E : \mathbb{F}_2^n \to \{0,1\}^m$ for min-entropy $k \geq 2\log n6O(1)$, where $m \leq k - O(1)$.*

This will be given in the exercise.

## 4 Seeded extractors

Seeded extractors are a relaxation of two-source extractors: we still work with two *independent* sources, but now we are promised that one of the sources is actually *uniform.* The goal is to use this pure randomness to extract even more entropy from the other defective source.

**Definition 44.** *A function $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k,\varepsilon)$ (seeded) extractor if for every distribution $X$ over $\{0,1\}^n$ with $H_\infty(X) \geq k$, $E(X,Y)$ is $\varepsilon$ close to $U_m$, where $Y$ is the uniform distribution over $\{0,1\}^d$. If it also holds that $|E(X,Y) \circ Y - U_m \circ Y| \leq \varepsilon$ we say that the extractor is* strong.

We have a few parameters:

- $n$ – the length of the defective random source,
- $k$ – the amount of entropy in the source distribution

- $d$ – The seed length (Also can be thought as the length and entropy in the second distribution).
- $m$ – the number of output bits. Also, the amount of entropy in the output distribution.
- $\varepsilon$ – the error. The distance of the output distribution from uniform.

**Definition 45.** *A function $E : [N] \times [D] \to [M]$ is a $(k, \varepsilon)$ disperser if for every $A \subseteq [N]$ of size at least $K = 2^k$, $|E(A, [D])| \geq (1 - \varepsilon)M$.*

We first show that non-explicitly such objects exist. For that we need the following Chernoff bound [13, 7]:

**Theorem 46.** *Suppose $Y_1, \dots, Y_n$ are i.i.d. boolean random variables with expectation $\mu$. Then for every $\varepsilon > 0$,*

$$\Pr\left[\sum_{i=1}^{n} Y_i > (\mu + \varepsilon)n\right] \leq e^{-2\varepsilon^2 n}.$$

**Theorem 47.** *For every integers $k \leq n$ and $\varepsilon > 0$ there exists $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ that is a $(k, \varepsilon)$ extractor with $m = k + d - 2\log\frac{1}{\varepsilon} - O(1)$ and $d = \log(n - k) + 2\log\frac{1}{\varepsilon} + O(1)$.*

*Proof.* The proof, again, uses the probabilistic method. Denote $N = 2^n$, $K = 2^k$, $D = 2^d$ and $M = 2^m$. Choose $E$ uniformly at random, by choosing each $E(x, y)$ uniformly at random from $[M]$. We say $E$ is *bad* for $X \subseteq [N]$ and $T \subseteq [M]$ if

$$\frac{\left|\left\{x \in \mathrm{Supp}(X), y \in \{0,1\}^d : E(x,y) \in T\right\}\right|}{K \cdot D} \geq \frac{|T|}{M} + \varepsilon.$$

(Notice that we only care about cases where the set $T$ gets too much weight).

For a fixed $X$ of size $K$, and a fixed $T$, Theorem 46 tells us the probability over $E$ that $(X, T)$ is bad for $E$ is at most $e^{-2\varepsilon^2 KD}$. By a union bound, the probability over $E$ that there exist such bad $X$ and $T$ is at most $\binom{N}{K} 2^M e^{-2\varepsilon^2 KD} < 1$ (the inequality is left as an exercise). Therefore there exists a choice for $E$ with no and bad $(X, T)$, and this $E$ is necessarily a strong $(k, \varepsilon)$ extractor (why?). □

For dispersers, we have both upper and lower bounds, which will be given as an exercise:

**Theorem 48.** *For every integers $n \geq k$ and $\varepsilon > 0$ there exists a $(K, \varepsilon)$ disperser $E : [N] \times [D] \to [M]$ with $m = k + d - \log\log\frac{1}{\varepsilon} - O(1)$ and $d = \log(n - k) + \log\frac{1}{\varepsilon} + O(1)$.*

The quantity $d$ is the seed length. The quantity $k + d - m$ is referred to as the "entropy loss". We see that non-explicitly there are extractors with seed length at most $\log n + 2\log\frac{1}{\varepsilon} + O(1)$ and entropy loss at most $2\log\frac{1}{\varepsilon} + O(1)$ and dispersers with seed length at most $\log n + \log\frac{1}{\varepsilon} + O(1)$ and entropy loss at most $\log\log\frac{1}{\varepsilon} + O(1)$. These bounds are tight (even when not considered simultaneously). We state here a result of this kind for dispersers and we will it in the exercises.

**Theorem 49.** *If $E : [N] \times [D] \to [M]$ is a $(K, \varepsilon)$ disperser, $N \geq K$ and $D \leq \frac{1}{2}(1 - \varepsilon)M$ then $D = \Omega\left(\frac{1}{\varepsilon}\log\frac{N}{K}\right)$.*

## References

[1] Miklós Ajtai and Nathan Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[2] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[3] Michael Ben-Or and Nathan Linial. Collective coin flipping. *Randomness and Computation*, 5:91–115, 1990.

[4] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. *ECCC*, 2015.

[5] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[6] Benny Chor, Oded Goldreich, Johan Hasted, Joel Freidmann, Steven Rudich, and Roman Smolensky. The bit extraction problem or t-resilient functions. In *Foundations of Computer Science, 1985., 26th Annual Symposium on*, pages 396–407. IEEE, 1985.

[7] Vasek Chvátal. The tail of the hypergeometric distribution. *Discrete Mathematics*, 25(3):285–287, 1979.

[8] Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved ramsey graphs. *arXiv preprint arXiv:1506.04428*, 2015.

[9] Paul Erdös. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.

[10] Paul Erdös and George Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2:463–470, 1935.

[11] Uriel Feige. Noncryptographic selection protocols. In *Foundations of Computer Science, 1999. 40th Annual Symposium on*, pages 142–152. IEEE, 1999.

[12] Joel Friedman. On the second eigenvalue and random walks in randomd-regular graphs. *Combinatorica*, 11(4):331–362, 1991.

[13] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

[14] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. In *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pages 68–80. IEEE, 1988.

[15] Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM Journal on Computing*, 36(5):1231–1247, 2006.

[16] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[17] Grigorii Aleksandrovich Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problemy peredachi informatsii*, 24(1):51–60, 1988.

[18] Raghu Meka. Explicit resilient functions matching ajtai-linial. *CoRR*, abs/1509.00092, 2015.

[19] FP Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 2(1):264–286, 1930.

[20] Alexander Russell, Michael Saks, and David Zuckerman. Lower bounds for leader election and collective coin-flipping in the perfect information model. *SIAM Journal on Computing*, 31(6):1645–1662, 2002.

[21] Michael Saks. A robust noncryptographic protocol for collective coin flipping. *SIAM Journal on Discrete Mathematics*, 2(2):240–244, 1989.