THEORY OF COMPUTING, Volume 3 (2007), pp. 103–128
http://theoryofcomputing.org

Linear Degree Extractors and the Inapproximability of Max Clique and Chromatic Number

David Zuckerman*

Received: November 7, 2006; revised: July 19, 2007; published: August 6, 2007.

Abstract: We derandomize results of Håstad (1999) and Feige and Kilian (1998) and show that for all $\varepsilon > 0$, approximating MAX CLIQUE and CHROMATIC NUMBER to within $n^{1-\varepsilon}$ are NP-hard. We further derandomize results of Khot (FOCS '01) and show that for some $\gamma > 0$, no quasi-polynomial time algorithm approximates MAX CLIQUE or CHROMATIC NUMBER to within $n/2^{(\log n)^{1-\gamma}}$, unless $\tilde{\mathrm{NP}} = \tilde{\mathrm{P}}$.

The key to these results is a new construction of dispersers, which are related to randomness extractors. A randomness extractor is an algorithm which extracts randomness from a low-quality random source, using some additional truly random bits. We construct new extractors which require only $\log_2 n + O(1)$ additional random bits for sources with constant entropy rate, and have constant error. Our dispersers use an arbitrarily small constant

*Most of this work was done while visiting Harvard University, and was supported in part by a Radcliffe Institute for Advanced Study Fellowship, a John Simon Guggenheim Memorial Foundation Fellowship, a David and Lucile Packard Fellowship for Science and Engineering, and NSF Grants CCR-0310960 and CCF-0634811. A preliminary version of this paper appeared in STOC'06.

ACM Classification: F.2, G.3, G.2

AMS Classification: 68Q17, 68R10, 65C10

Key words and phrases: extractor, disperser, clique, chromatic number, approximation, pseudorandom, explicit construction, NP-hard

 DOI: 10.4086/toc.2007.v003a006

times $\log n$ additional random bits for sources with constant entropy rate. Our extractors and dispersers output $1-\alpha$ fraction of the randomness, for any $\alpha > 0$.

Our constructions rely on recent results in additive number theory and extractors by Bourgain, Katz, and Tao (2004), Barak, Impagliazzo, and Wigderson (FOCS '04), Barak et al. (STOC '05), and Raz (STOC '05). We also simplify and slightly strengthen key theorems in the second and third of these papers, and strengthen a related theorem by Bourgain (2005).

1 Introduction

This work has two sources of motivation: inapproximability and randomness extractors. We begin with inapproximability.

1.1 Inapproximability

MAX CLIQUE and CHROMATIC NUMBER are central optimization problems. Their decision versions were in Karp's original list of NP-complete problems [32]. The best approximation algorithms for these problems give approximation ratios of the form $n/\operatorname{polylog}(n)$ [8, 24], which is not much better than the trivial approximation of n. Yet no strong inapproximability results were known until Feige et al. [18] discovered a connection between probabilistically checkable proofs (PCPs) and MAX CLIQUE. The celebrated PCP Theorem of Arora et al. [3] then implied that it is NP-hard to approximate MAX CLIQUE to within n^c for some constant $c > 0$. This ratio was improved in [7, 6] until Håstad, in a breakthrough, showed a hardness ratio of $n^{1-\varepsilon}$, for any $\varepsilon > 0$ [25]. The catch is that Håstad's reduction is randomized, so his theorem assumes that NP $\neq$ ZPP. Assuming only NP $\neq$ P, Håstad's hardness ratio becomes $n^{1/2-\varepsilon}$. In this paper we derandomize Håstad's randomized reduction:

Theorem 1.1. *For all $\varepsilon > 0$, it is* NP*-hard to approximate* MAX CLIQUE *to within $n^{1-\varepsilon}$.*

The inapproximability of CHROMATIC NUMBER has historically been even harder to prove than MAX CLIQUE, because advances have typically occurred through reductions from MAX CLIQUE. Lund and Yannakakis were the first to show that it is NP-hard to approximate CHROMATIC NUMBER to within n^c for some constant $c > 0$ [37]. Other reductions ensued, culminating in Feige and Kilian's proof of a hardness ratio of $n^{1-\varepsilon}$ [19]. This uses Håstad's result, so it assumes that NP $\neq$ ZPP. Assuming only NP $\neq$ P, the best previous hardness ratio explicitly stated appears to be $n^{1/7-\varepsilon}$ [6]. Previous work likely implied something better, though certainly no better than $n^{1/2-\varepsilon}$. In this paper we derandomize Feige and Kilian's result:

Theorem 1.2. *For all $\varepsilon > 0$, it is* NP*-hard to approximate* CHROMATIC NUMBER *to within $n^{1-\varepsilon}$.*

Engebretsen and Holmerin [16] improved the hardness ratios for both problems to $n^{1-o(1)}$ under the stronger assumption that NP $\not\subseteq$ ZPTIME$(2^{\operatorname{polylog}(n)})$. Khot [34] later improved these $n^{1-o(1)}$ factors to $n/2^{(\log n)^{1-\gamma}}$ for some constant $\gamma > 0$, under the same assumption. We derandomize Khot's results and show $\tilde{\mathrm{NP}}$-hardness with respect to quasi-polynomial time reductions. Because of the quasi-polynomial time reductions, $\tilde{\mathrm{NP}}$-hardness is weaker than NP-hardness; see Subsection 2.1 for more details.

Theorem 1.3. *For some $\gamma > 0$, it is* $\tilde{\mathrm{NP}}$*-hard to approximate* MAX CLIQUE *to within* $n/2^{(\log n)^{1-\gamma}}$.

Theorem 1.4. *For some $\gamma > 0$, it is* $\tilde{\mathrm{NP}}$*-hard to approximate* CHROMATIC NUMBER *to within* $n/2^{(\log n)^{1-\gamma}}$.

The key to our inapproximability results is constructing an appropriate disperser, which is related to a randomness extractor. Good dispersers were known to help derandomize inapproximability results for MAX CLIQUE (e.g., [54, 48]), but it was not known for CHROMATIC NUMBER. Before discussing dispersers, we discuss extractors.

1.2 Randomness extractors

Randomness extractors are motivated by the possibility of using defective sources of randomness. The model for defective random source involves lower bounding the min-entropy:

Definition 1.5. The *min-entropy* of a distribution X is $H_\infty(X) = \min_x\{-\log_2 \Pr[X = x]\}$. A *$k$-source* is a distribution with min-entropy at least k. The *entropy rate* of a k-source on $\{0,1\}^n$ is k/n; we sometimes call a k-source a rate-k/n-source.

A randomness extractor is a function which extracts randomness from a k-source using a few additional uniformly random bits.

Definition 1.6 ([42]). Let U_ℓ denote the uniform distribution on ℓ bits. A function $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a *(k,ε)-extractor* if for every k-source X, the distribution $\mathrm{Ext}(X,U_d)$ is ε-close in statistical (variation) distance to U_m. We say Ext is a *strong* (k,ε)-extractor if the function $\mathrm{Ext}(x,y) \circ y$ is a (k,ε)-extractor, where $\circ$ denotes concatenation.

Besides their straightforward applications to simulating randomized algorithms using weak sources, extractors have had applications to many areas in derandomization that are seemingly unrelated to weak sources, including inapproximability [54, 51, 40]. Nisan and Ta-Shma [41] survey these applications.

Like many objects in the study of pseudorandomness, the existence of excellent extractors is relatively easy to establish via the probabilistic method. However, the explicit construction of efficient extractors has proved to be much more difficult.

We wish to construct extractors for any min-entropy k with d, the number of truly random bits, as small as possible and m, the number of output bits, as large as possible. Different parameter settings are needed for different applications. Constructing good extractors is highly non-trivial, because such constructions beat the "eigenvalue bound" [53]. Starting with the first extractor of Nisan and Zuckerman [42], a lot of effort has been expended constructing good extractors. See Shaltiel's survey [46] for more details.

In many applications, extractors are viewed as highly unbalanced strong expanders. In this view an extractor is a bipartite graph $G = (V,W,E)$ with $V = \{0,1\}^n$, $W = \{0,1\}^m$, and (x,z) is an edge iff there is some $y \in \{0,1\}^d$ such that $\mathrm{Ext}(x,y) = z$. Thus, the degree of each vertex of V is $D = 2^d$, and the extractor hashes the input $x \in V$ to a random neighbor among its D neighbors in W.

Often this degree D is of more interest than $d = \log D$. For example, in the samplers of [55] the degree is the number of samples; in the extractor codes of [48] D is the length of the code; in the simulation of BPP using weak sources [54] the degree is the number of calls to the BPP algorithm. Most relevant for us, in the inapproximability of MAX CLIQUE [54] the size of the graph is closely related to D.

Before the work of Ta-Shma et al. [49], all explicit extractors had degree D at least some unspecified polynomial in $n = \log|V|$. In contrast, a non-explicit construction achieves $D = O(n) = O(\log|V|)$, which matches the lower bound. Ta-Shma et al. were able to achieve degree $D = O(n\log^* n)$ for $k \geq \sqrt{n}\log^2 n$, but could only output about $k/\sqrt{n}$ bits. In the case where $k = \Omega(n)$, they could output $m = \Omega(k)$ bits, but then they achieved degree $D = n \cdot \mathrm{polylog}(n)$. Our new construction achieves linear degree and linear output length for constant-rate sources.

Theorem 1.7. *For all constant $\alpha, \delta, \varepsilon > 0$ there is an efficient family of strong $(k = \delta n, \varepsilon)$-extractors $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $m \geq (1-\alpha)\delta n$ and $D = 2^d = O(n)$.*

We now define the related notion of a disperser. While dispersers are usually defined with respect to an error parameter ε, here it is more convenient to use the parameter $s = 1 - \varepsilon$.

Definition 1.8. We may view a function $\mathrm{DIS} : [N] \times [D] \to [M]$ as a bipartite graph $([N],[M],E)$ where $(x,z) \in E$ iff $\mathrm{DIS}(x,y) = z$ for some $y \in [D]$. For a set $X \subseteq [N]$, let $\Gamma(X) = \{\mathrm{DIS}(x,y) \mid x \in X, y \in [D]\}$ be the set of neighbors of X. We say DIS is a (K,s)-*disperser* if, for any $X \subseteq [N]$ with $|X| \geq K$, $|\Gamma(X)| \geq sM$. We say DIS is a *strong* (K,s)-disperser if the function $\mathrm{DIS}(x,y) \circ y$ is a (K,s)-disperser, where $\circ$ denotes concatenation.

When s is very small (so the error is close to 1), the probabilistic method can be used to show that there exists dispersers with degree even smaller than n, namely $O(n/\log s^{-1})$. In this paper, we succeed in matching this degree explicitly for constant-rate sources. These dispersers are the key for our inapproximability results.

Theorem 1.9. *For all constant $\alpha, \delta > 0$ and $s = s(n) > 0$, there is an efficient family of strong $(K = N^{\delta}, s)$-dispersers $\mathrm{DIS} : [N = 2^n] \times [D] \to [M = 2^m]$ such that $D = O(n/\log s^{-1})$ and $m \geq (1-\alpha)\delta n$. For subconstant $\delta = \delta(n)$, the dependence is $D = (1/\delta)^{O(1)} n/\log s^{-1}$ and $m = \delta^{O(1)} n$.*

1.3 Techniques

Our techniques are based on a combination of random walks on expanders and additive number theory. Random walks on expanders have been used to amplify the success probability of RP and BPP algorithms without using many additional random bits [1, 29, 13]. This yields a disperser for sources with entropy rate greater than 1/2 [13]. By using Chernoff bounds for random walks on expanders [21, 31, 52], we can construct extractors in a similar way. However, random walks provably fail when the entropy rate drops below 1/2, so they were not considered relevant for this case.

We handle entropy rates below 1/2 by first condensing the input until its entropy rate exceeds 1/2, and then applying a random walk on an expander. Condensers have been used before to build extractors [44, 47]. We condense using techniques developed from additive number theory.

Our basic condenser requires only one additional random bit, and is very simple. Choose a prime p and form the line-point incidence graph over $\mathbb{F}_q$, where $q = 2^p$. This bipartite graph has as independent sets the lines and points in the plane $\mathbb{F}_q^2$, with an edge between a point and a line if the point lies on the line. View the input distribution as a distribution over the q^3 edges. On input an edge, use the one random bit to output a random choice of its two endpoints. If the input distribution has min-entropy rate δ, then roughly speaking one of the two outputs will have min-entropy rate $\delta' > \delta$. The proof of correctness is simple, given the line-point incidence theorem from Bourgain-Katz-Tao [11].

This basic condenser improves that of Barak et al. [5], which requires two random bits. The improvement is not necessary for our results, as we apply the basic condenser iteratively to achieve entropy rate .9. In fact, we use Raz's condenser [43], which is strong in the sense that with high probability over the constant-bit seed, the min-entropy rate will increase. We iteratively apply the Raz condenser in a manner similar to [53], to improve the output length to $1-\alpha$ fraction of the input min-entropy, for any $\alpha > 0$.

Although not needed for our other results, we further simplify and slightly strengthen other applications of additive number theory. These applications are based on the important theorem of Bourgain-Katz-Tao [11] and extended by Bourgain-Glibichuk-Konyagin [10]: in a field $\mathbb{F}_q$ where q is either prime or 2^p for p prime, if $|A| \leq p^{.9}$, then $\max(|A+A|, |A \cdot A|) \geq |A|^{1+\alpha}$ for a global constant $\alpha > 0$. (See Section 2.8 for more details, including the recent extension to non-prime fields.) Barak-Impagliazzo-Wigderson [4] used these ideas to show that for $\delta \leq .9$, if A, B, and C are independent rate-δ-sources taking values in $\mathbb{F}_q$, then $AB+C$ is close to a rate-$(1+\alpha')\delta$-source. We show that A and C do not have to be independent. Instead, the lemma follows if (A,C) is independent from B. Our overall proof is simpler than that in [4]. We further strengthen a theorem of Bourgain [9] and show that the function $A(A+B)$ also gives a rate improvement.

This paper is organized as follows. After some preliminaries in Section 2, we give our basic disperser and extractor constructions in Sections 3 and 4, respectively. We next show how to improve the output length of both constructions in Section 5. We then give the inapproximability of Max Clique and Chromatic Number in Sections 6 and 7, respectively. Finally, we improve the additive number theory applications in Section 8.

2 Preliminaries

Some common notation we use is $\circ$ for concatenation and $[n]$ for the set $\{1,2,\ldots,n\}$. All logarithms are to the base 2.

When letters denote integers we often use a capital letter to denote 2 to the corresponding small letter, e.g., $K = 2^k$. When letters denote random variables we often use capital letters for random variables and corresponding small letters for their instantiations.

For readability, we often assume various quantities are integers when they are not necessarily. It is not hard to see that this does not affect our analysis.

We often use the term efficient to denote polynomial-time computable.

2.1 Reductions and quasi-NP-hardness

Our NP-hardness results are with respect to polynomial-time, many-one reductions.

Quasi-polynomial in n means $2^{\mathrm{polylog}(n)}$. $\mathrm{N\tilde{P}}$ and $\tilde{\mathrm{P}}$ are the quasi-polynomial analogues of NP and P, respectively. As usual with inapproximability results, we analyze the appropriate gap problem.

Note that no language is $\tilde{\mathrm{NP}}$-complete with respect to polynomial-time reductions. For if there were such a language, it would be in $\mathrm{TIME}(2^{(\log n)^c})$ for some c; but then $\tilde{\mathrm{NP}} \subseteq \mathrm{TIME}(2^{(\log n)^{c+1}})$, contradicting the time hierarchy theorem.

Therefore, we consider $\tilde{\mathrm{NP}}$-hardness with respect to quasi-polynomial-time, many-one reductions. Then any NP-hard language is also $\tilde{\mathrm{NP}}$-hard. Moreover, if an $\tilde{\mathrm{NP}}$-hard language is in $\tilde{\mathrm{P}}$, then $\tilde{\mathrm{NP}} = \tilde{\mathrm{P}}$. Of course, $\tilde{\mathrm{NP}} = \tilde{\mathrm{P}} \iff \mathrm{NP} \subseteq \tilde{\mathrm{P}}$.

2.2 Distance between distributions

For a probability distribution X, $X(s)$ denotes $\Pr[X = s]$. For a set S, $X(S)$ denotes $\Pr[X \in S]$.

Definition 2.1. Let X_1 and X_2 be two distributions on the same space Ω. The statistical, or variation, distance between them is

$$\|X_1 - X_2\| = \max_{S \subseteq \Omega} |X_1(S) - X_2(S)| = \frac{1}{2} \sum_{s \in \Omega} |X_1(s) - X_2(s)| \,.$$

We say X_1 and X_2 are ε-close if $\|X_1 - X_2\| \le \varepsilon$, and are ε-far otherwise. We say a distribution on $\{0,1\}^n$ is ε-uniform if it is ε-close to U_n, the uniform distribution on n bits.

A useful method of computing the distance to the closest k-source is the following.

Lemma 2.2. *The distance of X to the closest ℓ-source is $\sum_s \max(X(s) - 2^{-\ell}, 0)$.*

Of course, only those s with $X(s) > 2^{-\ell}$ contribute to the above sum.

2.3 Flat sources

Definition 2.3. A source is a probability distribution. A flat source is a source which is uniform on its support. The support of a distribution X is denoted $\mathrm{supp}(X)$.

The following lemma shows that it suffices to consider flat k-sources.

Lemma 2.4 ([12]). *Any k-source is a convex combination of flat k-sources.*

2.4 Dispersers

Dispersers were defined in Definition 1.8. There are two possible notions of efficiency: one relative to the input size $\log N + \log D$ and the other relative to the graph size $N + M$. For the inapproximability results, we only need the second, weaker, notion.

Definition 2.5. We say $\mathrm{DIS} : [N] \times [D] \to [M]$ is *efficient* if it runs in polynomial time in its input size $\log N + \log D$. We say DIS is *polynomial-time constructible* if the disperser graph is constructible in polynomial time in the number of vertices $N + M$.

Of course, efficient implies polynomial-time constructible.

The following simple lemma is useful when $D = O(1)$.

Lemma 2.6. *A (K,s)-disperser is also a strong $(K,s/D)$-disperser.*

We also use the following simple lemma.

Lemma 2.7. *Given an efficient (K,s')-disperser* $\mathrm{DIS}_1 : [N] \times [D_1] \to [N']$ *and an efficient $(K' = s'N', s)$-disperser* $\mathrm{DIS}_2 : [N'] \times [D_2] \to [M]$*, we can build an efficient (K,s)-disperser* $\mathrm{DIS} : [N] \times [D_1 D_2] \to [M]$. *If moreover* DIS_2 *is a strong (K',s)-disperser, then* DIS *is a strong $(K,s/D_1)$-disperser.*

Proof. Take $\mathrm{DIS}(x, y_1 \circ y_2) = \mathrm{DIS}_2(\mathrm{DIS}_1(x,y_1), y_2)$. It is straightforward to verify that DIS is a (K,s)-disperser. To see the final statement of the lemma, suppose DIS_2 is strong. Then $\mathrm{DIS}(x, y_1 \circ y_2) \circ y_2$ is a (K,s)-disperser, so $\mathrm{DIS}(x, y_1 \circ y_2) \circ y_1 \circ y_2$ is a $(K, s/D_1)$-disperser. □

While we need the notion of strong disperser for the inapproximability of Chromatic Number, the notion that suffices for this is captured in the following simple lemma.

Lemma 2.8. *Let* $\mathrm{DIS} : [N] \times [D] \to [M]$ *be a strong (K,s)-disperser. For a set $X \subseteq [N]$, let $\Gamma_y(X) = \{\mathrm{DIS}(x,y) \mid x \in X\}$. Then* DIS *has the property that for any $X \subseteq [N]$ with $|X| \geq K$, there is a y such that $|\Gamma_y(X)| \geq sM$.*

2.5 Expander graphs

Expander graphs are related to dispersers, and we use random walks on expanders to build our dispersers. We define expansion via eigenvalues. Let G be a connected regular undirected graph, and let A be the transition matrix of a random walk on G. (If M is the adjacency matrix and d the degree, then $A = M/d$.) We call G a λ-expander if all eigenvalues of A other than 1 are at most λ in absolute value. Smaller λ mean better expansion. We will need 2^c-regular $2^{-\gamma c}$-expanders on 2^m nodes, for a constant $\gamma > 0$.

Extending earlier constructions which required large primes [36, 38], Morgenstern [39] gave explicit constructions which achieve this with γ approaching $1/2$. However, because the number of vertices is not 2^m and there are restrictions on the degree, it is easier to use expanders by Gabber and Galil [20]. They gave an explicit construction of 8-regular λ-expanders on 2^m nodes, for even m, where $\lambda = 5\sqrt{2}/8 < 1$. (See the survey [28] for the statement in this form, and for many other aspects about expanders.) The neighbors of a vertex may be computed with a constant number of arithmetic operations. By taking the $(c/3)$th power of the graph, we get a 2^c-regular $\lambda^{c/3}$-expander, as we need (though this requires that $3|c$).

2.6 Somewhere-random sources

The concept of somewhere-random sources will be useful in constructing dispersers.

Definition 2.9. An *elementary somewhere-k-source* is a vector of sources $(X_1, \ldots, X_\ell)$, such that some X_i is a k-source. A *somewhere-k-source* is a convex combination of elementary somewhere-k-sources.

Note that there may be arbitrary dependencies among the X_i. Further note that in a somewhere-k-source which is not elementary, all X_i may have low min-entropy.

2.7 Condensers

Condensers and somewhere condensers will be essential in our extractor and disperser constructions, respectively.

Definition 2.10. A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to \ell, \varepsilon)$-*condenser* if for every k-source X, $C(X,U_d)$ is ε-close to some ℓ-source. When convenient, we call C a rate-$(k/n \to \ell/m, \varepsilon)$-condenser. The condenser is strong if the average over $y \in \{0,1\}^d$ of the minimum distance of $C(X,y)$ to some ℓ-source is at most ε.

Definition 2.11. A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to \ell, \varepsilon)$-*somewhere-condenser* if for every k-source X, the vector $\langle C(X,y)\rangle_{y \in \{0,1\}^d}$ is ε-close to a somewhere ℓ-source. When convenient, we call C a rate-$(k/n \to \ell/m, \varepsilon)$-somewhere-condenser.

Note that a $(k \to \ell, \varepsilon)$-strong-condenser is a $(k \to \ell, \varepsilon)$-somewhere-condenser. We will also need the following.

Lemma 2.12. *If $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k \to \ell, \varepsilon)$-somewhere-condenser, then it is a $(2^k, (1-\varepsilon)2^{\ell-m})$-disperser.*

Proof. This follows because a distribution which is ε-close to an ℓ-source must have a support of size at least $(1-\varepsilon)2^\ell$. □

When composing condensers, we will need the following type of lemma.

Lemma 2.13. *Suppose Z_1 is ε_1-close to an ℓ_1-source, and for all $z_1 \in \mathrm{supp}(Z_1)$, the distribution $(Z_2 \mid Z_1 = z_1)$ is ε_2-close to an ℓ_2-source. Then $Z_1 \circ Z_2$ is $\varepsilon_1 + \varepsilon_2$-close to an $\ell_1 + \ell_2$-source.*

Proof. Let W_1 be an arbitrary ℓ_1-source which is ε_1-close to Z_1. For $w_1 \in \mathrm{supp}(Z_1) \cap \mathrm{supp}(W_1)$, define the distribution of $(W_2 \mid W_1 = w_1)$ to be an arbitrary ℓ_2-source which is ε_2-close to $(Z_2 \mid Z_1 = w_1)$. For $w_1 \in \mathrm{supp}(W_1) \setminus \mathrm{supp}(Z_1)$, define the distribution of $(W_2 \mid W_1 = w_1)$ to be the uniform distribution. Then $W_1 \circ W_2$ is an $\ell_1 + \ell_2$-source, which is $\varepsilon_1 + \varepsilon_2$-close to $Z_1 \circ Z_2$, □

We build extractors by first condensing and then applying a weaker extractor. The idea of condensing before extracting was used in [44, 47], and a simple lemma from [47] shows that this works.

Lemma 2.14 ([47]). *Suppose that $C : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{n'}$ is an efficient (strong) $(k \to \ell, \varepsilon_1)$-condenser, and $\mathrm{Ext} : \{0,1\}^{n'} \times \{0,1\}^{d_2} \to \{0,1\}^m$ is an efficient (strong) (ℓ, ε_2)-extractor. Then*

$$\mathrm{Ext}'(x, y_1 \circ y_2) = \mathrm{Ext}(C(x,y_1), y_2)$$

is an efficient (strong) $(k, \varepsilon_1 + \varepsilon_2)$-extractor.

2.8 Sum-product theorem

The following sum-product theorem underlies our condensers.

Theorem 2.15 ([11, 10]). *There is a constant $\alpha > 0$ such that for any field $F = \mathbb{F}_q$ where q is either prime or 2^p for p prime, the following holds. For any non-empty $A \subset F$, $|A| < q^{.9}$,*

$$\max(|A+A|, |A \cdot A|) = \Omega(|A|^{1+\alpha}).$$

Here the .9 can be increased to any constant less than 1, but the constant α will likely decrease. Note that for $q = 2^p$, when $A = \{0, 1\}$ then $A + A = A \cdot A = A$; however, the Ω handles this problematic case. We use results based on earlier versions of this theorem, when the full bounds were not known to hold for fields of size 2^p. Although the result quoted above isn't apparent for such fields in the credited papers, it follows from Corollary 2.56 of [50], which credits those papers. For the best constant α as of this writing see [33]. For a self-contained exposition of the prime case, see [23].

3 Disperser construction

We first use random walks on expanders to construct low-degree dispersers for high min-entropy. This construction could work for any min-entropy rate bigger than $1/2$, but to output almost all the randomness we need rate close to 1.

Proposition 3.1. *For any $\alpha > 0$, there is a $\beta, c_0 > 0$ such that for any $c = c(n) \geq c_0$, there is an efficient family of strong $(K = N^{1-\beta}, 2^{-c})$-dispersers* DIS : $[N = 2^n] \times [D] \to [M = 2^m]$ *such that $D \leq \alpha n/c$ and $m \geq (1-\alpha)n$. (Let $\gamma > 0$ be the constant from Section 2.5. We can take $c_0 = 2/\gamma$ and $\beta = \alpha\gamma/5$.)*

Proof. We use the disperser of [1]. Set $s = 2^{-c}$, and $m = (1-\alpha)n$. Let G be a 2^c-regular $2^{-\gamma c}$-expander on $[2^m]$ (see Section 2.5). To find the neighbors of a vertex $u \in [2^n]$, use the n bits defining u to choose a random vertex $v_0 \in [2^m]$ and then take a random walk $v_1, \ldots, v_D$ on G. Connect u to $v_1, \ldots, v_D$. We ignore v_0 so that we cleanly get $n = m + Dc$, and $D = (n-m)/c = \alpha n/c$.

First consider when the bits describing the random walk are uniformly random. In this case we can use the tight analysis given by Kahale [30]. For $S \subseteq [2^m]$ and $s = |S|/2^m$, Kahale showed that

$$\Pr[(\forall i) v_i \in S] \leq s(s + (1-s)\lambda)^{D-1} < (s+\lambda)^D.$$

Since $s = 2^{-c} < 2^{-\gamma c}$, this probability is less than $2^{(1-\gamma c)D} \leq 2^{-(\gamma/2)cD} \leq 2^{-(\gamma/2)\alpha n}$.

When the bits describing the random walk are chosen from a source with min-entropy $(1-\beta)n$, each string which before had probability 2^{-n} now has probability at most $2^{\beta n} \cdot 2^{-n}$. Therefore, the error probability grows by at most $2^{\beta n}$, and hence is at most $2^{(\beta - \gamma\alpha/2)n}$. Therefore, this is a $(K = N^{1-\beta}, s)$-disperser for any $\beta < \gamma\alpha/2$.

We still need to show that this disperser is strong. To do this, we must consider the situation where instead of one S we now have D such S_i, $|S_i| = s_i 2^m$, where the average of the s_i is at most s. By the result of Kahale given as Theorem A.5 in [22], for a uniformly random walk

$$\begin{aligned}\Pr[(\forall i) v_i \in S_i] &\le \sqrt{s_1 s_D} \prod_{i=2}^{D} \sqrt{s_i + (1-s_i)\lambda^2} \le \sqrt{\prod_{i=1}^{D} (\lambda^2 + (1-\lambda^2)s_i)} \\ &\le \left(\frac{1}{D} \sum_{i=1}^{D} (\lambda^2 + (1-\lambda^2)s_i) \right)^{D/2} \le (\lambda^2 + s)^{D/2}.\end{aligned}$$

The third inequality follows from the arithmetic-geometric mean. This bound $(s+\lambda^2)^{D/2}$ is at most the square root of Kahale's earlier bound, so it's at most $2^{-(\gamma\alpha/4)n}$. By choosing $\beta < \gamma\alpha/4$ the proposition follows. □

To give a construction for all positive entropy rates, we use the following theorem, which follows from the condenser in [5] or [43]. While [5] only gives an ordinary disperser, by Lemma 2.6 it is also a strong disperser for essentially the same parameters, since D is constant.

Theorem 3.2 ([5, 43]). *For any $\beta, \delta > 0$, there is an efficient family of rate-$(\delta \to 1-\beta, \varepsilon = 2^{-\Omega(n)})$-somewhere-condensers $C : [N = 2^n] \times [D] \to [M = 2^m]$ where $D = O(1)$ and $m = \Omega(n)$. For subconstant $\delta = \delta(n)$ the dependence is $D = (1/\delta)^{O(1)}$ and $m = \delta^{O(1)} n$.*

Remark 3.3. In the original paper, the construction for subconstant δ required a large prime. However, there is no longer a need for this, given the new sum-product theorem for fields of size 2^p (see Subsection 2.8).

Applying Lemmas 2.12 and 2.6, we deduce

Corollary 3.4. *For any $\beta, \delta > 0$, there is an efficient family of strong $(K = N^\delta, M^{-\beta})$-dispersers* DIS : $[N = 2^n] \times [D] \to [M = 2^m]$ *where $D = O(1)$ and $m = \Omega(n)$.*

We can now give our disperser construction, although for now we obtain output length a small constant fraction of δn, rather than almost all of it.

Theorem 3.5. *For any $\delta > 0$ and $s = s(n) > 0$, there is an efficient family of strong $(K = N^\delta, s)$-dispersers* DIS : $[N = 2^n] \times [D] \to [M = 2^m]$ *such that $D = O(n/\log s^{-1})$ and $m = \Omega(n)$. For subconstant $\delta = \delta(n)$ the dependence is $D = (1/\delta)^{O(1)} n/\log s^{-1}$ and $m = \delta^{O(1)} n$.*

Proof. Let $\mathrm{DIS}_1 : [N = 2^n] \times [D_1 = O(1)] \to [N' = 2^{n'}]$ be an efficient strong $(K = N^\delta, (N')^{-.1})$-disperser from Corollary 3.4, with $n' = \Omega(n)$. Let $\mathrm{DIS}_2 : [N'] \times [D_2 \le n'/\lg s^{-1}] \to [M = 2^m]$, be an efficient strong $(K' = (N')^{.9}, s)$-disperser given by Proposition 3.1, with $m = n'/2$. Applying Lemma 2.7 yields the desired disperser. □

To improve the output length to $(1-\alpha)\delta n$, we need to use better condensers, and we defer the proof to the next section.

4 Extractor construction

Readers interested solely in the inapproximability results can skip directly to Section 6, as the current dispersers suffice to prove those results.

Our extractor construction is essentially the same as our disperser construction. We first show how to extract when the entropy rate is close to 1, by using random walks on expanders. Then we use Raz's recent condenser [43] to reduce to the high-entropy case.

Proposition 4.1. *For all $\alpha,\varepsilon > 0$, there exists $\beta > 0$ such that there is an efficient family of $(k = (1-\beta)n,\varepsilon)$-extractors* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with $m \geq (1-\alpha)n$ and $D = 2^d \leq \alpha n$.*

Proof. Set $m = (1-\alpha)n$ and $c = 3$ (say). Let G be a 2^c-regular λ-expander on $[2^m]$ with λ bounded away from 1 (see Section 2.5). To find the neighbors of a vertex $u \in [2^n]$, use the n bits defining u to choose a random vertex $v_0 \in [2^m]$ and then take a random walk $v_1,\ldots,v_D$ on G. Connect u to $v_1,\ldots,v_D$. As before, $n = m + Dc$, and $D = (n-m)/c = \alpha n/c$.

Let $S \subseteq [2^m]$ have density $\mu = |S|/2^m$. First consider when the bits describing the random walk are chosen uniformly, and let the random variable $\hat{\mu}$ denote the fraction of v_i which are in S. Gillman [21] (see also Kahale [31]) proved a Chernoff bound for random walks on expanders. We use the improved constants obtained by Healy [27]:

$$\Pr[|\hat{\mu}-\mu| \geq \varepsilon] \leq 2\exp(-(1-\lambda)\varepsilon^2 D/4)\,.$$

(Dinwoodie [14] essentially improved the constant 4 above to 2, but only states it from a worst-case vertex so there is another term.)

For large enough n (to get rid of the multiplicative 2), this error is at most 2^{-an} for $a = (1-\lambda)\varepsilon^2\alpha/(4c)$. When the bits describing the random walk are chosen from a source with min-entropy $(1-\beta)n$, the error probability grows by at most $2^{\beta n}$. Thus this is a $(k = (1-\beta)n,\varepsilon)$-extractor for $\beta < a$. □

We can make these extractors strong by using a better Chernoff bound.

Proposition 4.2. *For all $\alpha,\varepsilon > 0$, there exists $\beta > 0$ such that there is an efficient family of strong-$(k = (1-\beta)n,\varepsilon)$-extractors* $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with $m \geq (1-\alpha)n$ and $D = 2^d \leq \alpha n$.*

Proof. We use the same construction. For the proof, we must now show near uniformity over $[D]\times[2^m]$. We therefore consider $S \subseteq [D]\times[2^m]$, so $S = \cup_i \{i\} \times S_i$. Again consider when the bits describing the random walk are chosen uniformly, and now let the random variable $\hat{\mu}$ denote the fraction of v_i which are in S_i. Wigderson and Xiao [52] improved Gillman's theorem above for this case. We again use Healy's improved constants [27]):

$$\Pr[|\hat{\mu}-\mu| \geq \varepsilon] \leq 2\exp(-(1-\lambda)\varepsilon^2 D/4)\,.$$

We can then conclude with the same argument as above. □

For dispersers, we combined the high-entropy construction with somewhere-condensers from Theorem 3.2. For extractors, we need to use the improved strong condenser due to Raz [43].

Theorem 4.3 ([43]). *For any constants $\beta, \delta, \varepsilon > 0$, there is a constant d such that there is an efficient rate-$(\delta \to (1-\beta), \varepsilon)$-strong condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that $m = \Omega(n)$.*

Applying Lemma 2.14 to Raz's condenser and the extractor above, we obtain the desired theorem, except that the output length is linear instead of the $(1-\alpha)$-fraction we claimed.

Theorem 4.4. *For all $\delta, \varepsilon > 0$ there is an efficient family of strong-$(k = \delta n, \varepsilon)$-extractors* Ext $: \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with $m = \Omega(n)$ and $D = 2^d = O(n)$.*

5 Improving the output length

The results in this section were obtained jointly with Avi Wigderson.

We now would like to obtain output length $(1-\alpha)k$, for an arbitrary $\alpha > 0$, while maintaining the linear degree. The initial idea is to do a construction similar to that by Wigderson and the author [53]: if the output length is significantly less than k, use an independent seed to extract more bits from the same input. We can't do this directly, because even two runs of the extractor gives degree $\Theta(n^2)$, which is too expensive. Yet we can achieve this with the condenser, which uses only a constant number of random bits. Thus, our intermediate goal, which is interesting in its own right, is:

Theorem 5.1. *For any constants $\alpha, \beta, \delta, \varepsilon > 0$, there is a constant d such that there is an efficient rate-$(\delta \to (1-\beta), \varepsilon)$-strong condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that $m \geq (1-\alpha)\delta n$.*

Yet this theorem cannot be achieved by applying the above idea to Theorem 4.3. The reason is that the error cannot be controlled. If the output length is γn, we would like to iterate about $1/\gamma$ times, but we cannot do this if the initial error is bigger than γ. In Theorem 4.3, as well as Theorem 3.2, the output length may depend on the error. Hence we construct an improved condenser, which follows from the improved merger of Dvir and Raz [15]. In this merger, the output length doesn't depend on the error.

Lemma 5.2. *For any $\delta > 0$, there exists $\gamma > 0$, such that for any $\varepsilon > 0$, there is a constant d such that there is an efficient rate-$(\delta \to (1-\delta), \varepsilon)$-strong condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that $m \geq \gamma n$.*

Proof. Fix $\delta > 0$. By Theorem 3.2, there is an efficient rate-$(\delta \to 1-\delta/2, \varepsilon_1 = 2^{-\Omega(n)})$ somewhere-condenser $C_1 : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{m_1}$, where $d_1 = O(1)$ and $m_1 = \Omega(n)$. By the main theorem of [15], there is a "strong merger" $M : (\{0,1\}^{m_1})^{2^{d_1}} \times \{0,1\}^d \to \{0,1\}^m$ with $d = f(d_1) = O(1)$ and $m = \Omega(m_1)$ such that whenever the input X_1 on $(\{0,1\}^{m_1})^{2^{d_1}}$ is a somewhere rate $(1-\delta/2)$-source, then the average over $y \in \{0,1\}^d$ of the distance of $M(X_1, y)$ to the closest rate $(1-\delta)$-source is at most $\varepsilon/2$. The length m may be chosen independently of ε, although d depends on ε. Hence the required strong condenser is $C(x,y) = M(\langle C_1(x,y_1)\rangle_{y_1 \in \{0,1\}^{d_1}}, y)$. □

The following lemma is the condenser analogue to the corresponding extractor lemma in [53].

Lemma 5.3. *Suppose that $C_1 : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{m_1}$ is a strong $(k \to \ell_1, \varepsilon_1)$-condenser and $C_2 : \{0,1\}^n \times \{0,1\}^{d_2} \to \{0,1\}^{m_2}$ is a strong $(k - m_1 - s \to \ell_2, \varepsilon_2)$-condenser. Then $C : \{0,1\}^n \times \{0,1\}^{d_1+d_2} \to \{0,1\}^{m_1+m_2}$, given by*

$$C(x, y_1 \circ y_2) = C_1(x, y_1) \circ C_2(x, y_2),$$

is a strong $(k \to \ell_1 + \ell_2, \varepsilon_1 + \varepsilon_2 + 2^{-s})$*-condenser.*

Proof. Let X be a k-source. For $y \in \{0,1\}^{d_i}$, let ε_i^y denote the minimum distance of $C_i(X,y)$ to some ℓ_i-source. Fix $y_1 \in \{0,1\}^{d_1}$. Let S denote the set of low-probability elements in the output:

$$S = \{z \mid \Pr_X[C_1(X,y_1) = z] \le 2^{-(m_1+s)}\}.$$

Then $\Pr[C_1(X,y_1) \in S] \le |S|2^{-(m_1+s)} \le 2^{-s}$. For $z \notin S$, X conditioned on $C_1(X,y_1) = z$ is a $(k-m_1-s)$-source. Hence, under such conditioning, for each $y_2 \in \{0,1\}^{d_2}$, $C_2(X,y_2)$ is within $\varepsilon_2^{y_2}$ of some ℓ_2-source. Putting this together as in Lemma 2.13, $C(X, y_1 \circ y_2)$ is within $\varepsilon_1^{y_1} + 2^{-s} + \varepsilon_2^{y_2}$ of some $\ell_1 + \ell_2$-source. Since the average of $\varepsilon_i^{y_i}$ is at most ε_i, this completes the proof of the lemma. □

Applying this lemma inductively, we can show:

Lemma 5.4. *Suppose* $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is an efficient strong* $(k \to \ell, \varepsilon)$*-condenser. Then for any positive integers* s,t*, we can construct* $C' : \{0,1\}^n \times \{0,1\}^{td} \to \{0,1\}^{tm}$*, an efficient strong* $(k + (t-1)m + s \to t\ell, t\varepsilon + (t-1)2^{-s})$*-condenser.*

Proof. We prove this by induction on t. For the base case $t = 1$ we can take $C' = C$. Now assume the lemma for a given t. Set C_1 to be the condenser given by the lemma for t, and set $C_2 = C$. Applying Lemma 5.3 gives the condenser for $t+1$. □

We can now prove Theorem 5.1. We would like to condense additional entropy, as long as there is $\alpha\delta n$ entropy left in the source. We also want the output entropy rate to be $1-\beta$, and if in each iteration we have this entropy rate, then overall we do as well. These two goals mean we should use a condenser converting rate $\alpha\delta$ to rate $1-\beta$. This condenser has some output length γn, so we need to iterate $1/\gamma$ times. This determines the error we need, which is why it is crucial we can pick the error after knowing γ.

Proof of Theorem 5.1. Let $\alpha, \beta, \delta, \varepsilon > 0$ be given. By Lemma 5.2, for some $\gamma > 0$ there is an efficient strong rate-$(\alpha\delta \to (1-\beta), \varepsilon')$-condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, where ε' will be chosen later and $m \ge \gamma n$. Set $t = (1-\alpha)\delta/\gamma$ and apply Lemma 5.4 with an s to be chosen later. This gives an efficient strong $(\delta n - \gamma n + s \to (1-\beta)(tm), i\varepsilon' + 2^{-s})$-condenser $C' : \{0,1\}^n \times \{0,1\}^{id} \to \{0,1\}^{tm}$. Choosing $s = \gamma n$ and ε' small enough so $t\varepsilon' + 2^{-s} \le \varepsilon$ gives the theorem. □

Combining our condenser from Theorem 5.1 with our extractor from Proposition 4.1 via Lemma 2.14, we obtain our main extractor construction:

Theorem 1.7. For all constant $\alpha, \delta, \varepsilon > 0$ there is an efficient family of strong $(k = \delta n, \varepsilon)$-extractors $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $m \ge (1-\alpha)\delta n$ and $D = 2^d = O(n)$.

By combining the same condenser with the earlier disperser of Proposition 3.1, we obtain our main disperser construction:

Theorem 1.9. For all constant $\alpha, \delta > 0$ and $s = s(n) > 0$, there is an efficient family of strong $(K = N^\delta, s)$-dispersers $\mathrm{DIS} : [N = 2^n] \times [D] \to [M = 2^m]$ such that $D = O(n/\log s^{-1})$ and $m \ge (1-\alpha)\delta n$. For subconstant $\delta = \delta(n)$, the dependence is $D = (1/\delta)^{O(1)} n/\log s^{-1}$ and $m = \delta^{O(1)} n$.

6 Max Clique

In this section, we show how our dispersers yield inapproximability results for MAX CLIQUE. We assume some familiarity with PCPs. Since the inapproximability of MAX CLIQUE follows from the proof of the inapproximability of CHROMATIC NUMBER, readers not familiar with PCPs may prefer to read the next section, which doesn't use them.

Historically, Feige et al. [18] were the first to show how to obtain inapproximability results for MAX CLIQUE using PCPs. Bellare, Goldreich, and Sudan [6] showed that free bit complexity is the parameter of a PCP which gives the best inapproximability results.

Definition 6.1. $\mathrm{FPCP}_s(r, f)$ is the class of promise problems recognized by PCP verifiers using r random bits and f free bits, achieving perfect completeness and soundness s.

Theorem 6.2 ([6, 18]). *If* $\mathrm{NP} \subseteq \mathrm{FPCP}_s(r, f)$*, then it is NP-hard to distinguish whether a graph on* 2^{r+f} *vertices has clique number at least* 2^r *or at most* $s2^r$.

Håstad [25] showed how to reduce the soundness by paying only a tiny amount in the free bit complexity. Specifically, he showed:

Theorem 6.3 ([25]). *For any* $\bar{f} > 0$*, there is an* ℓ *such that* $\mathrm{NP} \subseteq \mathrm{FPCP}_{2^{-\ell}}(O(\log n), \bar{f}\ell)$.

The quantity $\bar{f}$ is called the *amortized free bit complexity*, and can be less than 1 (Håstad's result shows it can be any positive constant).

The following follows from Theorem 6.2 and the amplification of a PCP via a good disperser, as first suggested in [54].

Lemma 6.4. *Suppose* $\mathrm{NP} \subseteq \mathrm{FPCP}_s(r, f)$ *and there is a polynomial-time constructible* (K, s)*-disperser* $\mathrm{DIS} : [2^R] \times [D] \to [2^r]$*. Then* $\mathrm{NP} \subseteq \mathrm{FPCP}_{K/2^R}(R, Df)$*, and hence it is* NP*-hard to distinguish whether a graph on* 2^{R+Df} *vertices has clique number at least* 2^R *or clique number at most* K.

This suffices to prove our theorem.

Theorem 1.1. It is NP-hard to approximate MAX CLIQUE to within $n^{1-\varepsilon}$ for any $\varepsilon > 0$.

Proof. Equivalently, we will show a factor of $n^{1-2\varepsilon}$. Fix $\varepsilon > 0$. Theorem 1.9 says that for any $s = s(n)$ there is an efficient family of $(K = N^{\varepsilon}, s)$-dispersers of degree $D \le c(\log N)/\log s^{-1}$, for some $c = c(\varepsilon)$. Let $\bar{f} \le \varepsilon/c$, and apply Theorem 6.3 to get an ℓ and $r = r(n) = O(\log n)$ such that $\mathrm{NP} \subseteq \mathrm{FPCP}_{2^{-\ell}}(r, \bar{f}\ell)$. Now let $s = 2^{-\ell}$, so there is an efficient $(K = (2^R)^{\varepsilon}, 2^{-\ell})$-disperser $\mathrm{DIS} : [2^R] \times [D] \to [2^r]$. Apply Lemma 6.4 with this disperser, and note that $Df \le (cR/\ell) \cdot (\ell\bar{f}) = \bar{f} \cdot cR \le \varepsilon R$. Hence it is NP-hard to distinguish clique number at least 2^R from clique number at most $2^{\varepsilon R}$ in graphs on $2^{(1+\varepsilon)R}$ vertices. Moreover, since the output length is linear in the input length, $R = O(\log n)$, so the reduction is polynomial time. □

To obtain inapproximability up to an $n^{1-o(1)}$ factor, we can use the following theorem by Håstad and Khot [26], which is basically the same as that obtained by Samorodnitsky and Trevisan [45] but gives perfect completeness.

Theorem 6.5 ([26]). *For any $\ell = \ell(n)$ which is one less than a perfect square,*

$$\mathrm{NP} \subseteq \mathrm{FPCP}_{2^{-\ell}}(O(\ell \log n), 2\sqrt{\ell+1}).$$

We can now prove:

Theorem 1.3. For some $\gamma > 0$, it is $\tilde{\mathrm{NP}}$-hard to approximate MAX CLIQUE to within $n/2^{(\log n)^{1-\gamma}}$.

Proof. Set $\varepsilon = \varepsilon(n) = 1/\log n$. By Theorem 3.5 (or the stronger Theorem 1.9), there is a c such that for any $s = s(n)$ there is a polynomial-time constructible family of $(K = N^{\varepsilon}, s)$-dispersers of degree $D \le (\log n)^c (\log N)/\log s^{-1}$. Let $\ell = 9(\log n)^{2(c+1)}$ and $s = 2^{-\ell}$. Apply Theorem 6.5 to get $r = r(n) = \mathrm{polylog}(n)$ such that $\mathrm{NP} \subseteq \mathrm{FPCP}_s(r, 3\sqrt{\ell})$. We'll use the polynomial-time constructible $(K = (2^R)^{\varepsilon}, 2^{-\ell})$-disperser $\mathrm{DIS} : [2^R] \times [D] \to [2^r]$. Apply Lemma 6.4 with disperser DIS, and note that $Df \le (R(\log n)^c/\ell) \cdot (3\sqrt{\ell}) = R/\log n = \varepsilon R$. Hence it is NP-hard to distinguishing clique number at least 2^R from clique number at most $2^{R/\log n}$ in graphs on $2^{(1+1/\log n)R}$ vertices. Since $R = \mathrm{polylog}(n)$, the theorem follows. □

7 Chromatic Number

Now we show how our dispersers imply the NP-hardness of approximating CHROMATIC NUMBER to within $n^{1-\varepsilon}$ for any $\varepsilon > 0$. We derandomize Feige's and Kilian's proof [19] of the same inapproximability ratio but under the stronger assumption that NP is not in ZPP. As in their proof, we work with the fractional chromatic number χ_f, which up to logarithmic factors is the same as the chromatic number χ [35]. They also make use of the independence number α. Just as $\alpha(G)\chi(G) \ge |V(G)|$, so too is $\alpha(G)\chi_f(G) \ge |V(G)|$ (here $V(G)$ denotes the vertices of G).

Feige and Kilian start with a graph G (from a family of graphs) which has a constant hardness ratio: either G has chromatic number at least c or at most $c' < c$. They actually need $c' = c^{\gamma}$, where $\gamma > 0$ is arbitrary, as well as a corresponding bound on the independence number α.

Theorem 7.1 ([19]). *For all $\gamma > 0$, there is an $s > 0$, such that there is a polynomial-time reduction from an NP-complete language L to chromatic number with the following properties. On input x, the algorithm outputs a graph $G = (V, E)$ such that*

1. *If $x \in L$ then $\chi_f(G) \le s^{-\gamma}$;*
2. *If $x \notin L$ then $\alpha(G) < s|V|$, and hence $\chi_f(G) > 1/s$.*

(The parameter s is not exactly the soundness of the PCP; rather, it is the soundness times 2^{-f}, where f is the free bit complexity. Also, Feige and Kilian don't state this as a theorem, but it can be deduced from their Lemma 2 and the parameters achieved in their Section 5.6. They state their parameters as: for any $\gamma, \ell > 0$, they can set $s = O(2^{-\ell})$ and if $x \in L$ then $\chi_f(G) \le 2^{3\gamma\ell+1}$. This is equivalent to our statement above, for a slightly different choice of γ.)

Feige and Kilian next amplify the hardness ratio using randomized graph products. That is, they take a suitably-sized random subgraph G' of the product graph G^D which has hardness ratio $|V(G')|^{1-\varepsilon}$. G^D is defined with respect to the following "OR" graph product.

Definition 7.2. For graphs $G = (V,E)$ and $H = (W,F)$, define the graph $G \times H$ as having vertex set $V \times W$, and edges $\{(v,w),(v',w')\}$ where $\{v,v'\} \in E$ or $\{w,w'\} \in F$.

Note that $(v_1,\ldots,v_D)$ is adjacent to $(w_1,\ldots,w_D)$ in G^D if any (v_i,w_i) is an edge in G. It is straightforward to show that $\alpha(G\times H) = \alpha(G)\cdot\alpha(H)$. Using the definition of χ_f as a linear program and linear programming duality, Feige showed that $\chi_f(G\times H) = \chi_f(G)\cdot\chi_f(H)$ [17].

We derandomize the randomized graph powering. This was done earlier in the clique setting [2], but the results there are not tight enough. On the other hand, for cliques, two types of bounds are needed – one if the clique number is large, and one if it's small. For chromatic number, one of the two cases becomes easy. If $\chi_f(G)$ is small, it will suffice to use the trivial bound $\chi_f(G') \le \chi_f(G^D) = \chi_f(G)^D$.

We can define a derandomized graph powering of $G = (V,E)$ with respect to any disperser $\mathrm{DIS} : X \times [D] \to V$ as follows. Define $\mathrm{DIS}(x) = (\mathrm{DIS}(x,1),\mathrm{DIS}(x,2),\ldots,\mathrm{DIS}(x,D))$ and $\mathrm{DIS}(X) = \{\mathrm{DIS}(x) \mid x \in X\}$. Now define the graph $\mathrm{DIS}(G^D)$ to be the induced subgraph of G^D on vertex set $\mathrm{DIS}(X)$.

Lemma 7.3. *Given a graph $G = (V,E)$ and a disperser* DIS *with degree D, let $G' = (V',E') = \mathrm{DIS}(G^D)$. Then*

1. $\chi_f(G') \le (\chi_f(G))^D$.
2. *If $\alpha(G) < s|V|$ and* DIS *is a strong (K,s)-disperser, then $\alpha(G') < K$, and hence $\chi_f(G') > |V'|/K$.*

Proof. The first part follows because $\chi_f(G') \le \chi_f(G^D) = (\chi_f(G))^D$. For the second part, suppose $\alpha(G') \ge K$, and let X be an independent set in G' of size K. Note that $\Gamma_i(X)$, as defined in Lemma 2.8, corresponds to the set of ith coordinates of X. By the strong disperser property, for some $i \in [D]$, $|\Gamma_i(X)| \ge s|V| > \alpha(G)$. Hence $\Gamma_i(X)$ is not an independent set in G, so it contains an edge, say $\{v_i,w_i\}$. If v_i is the ith coordinate of v, and w_i is the ith coordinate of w, then because we are using the OR graph product, $\{v,w\}$ is an edge in G'. Since $v,w \in X$, this contradicts our assumption that X was an independent set. $\square$

We are now ready to prove our theorem.

Theorem 1.2. It is NP-hard to approximate CHROMATIC NUMBER to within $n^{1-\varepsilon}$ for any $\varepsilon > 0$.

Proof. Fix $\varepsilon > 0$. Theorem 1.9 says that for any $s = s(n)$ there is an efficient family of strong $(K = N^\varepsilon, s)$-dispersers of degree $D \le cn/\log s^{-1}$, for some $c = c(\varepsilon)$. Set $\gamma = \varepsilon/c$, and use the Feige-Kilian reduction, which comes with an $s = s(\gamma)$. Using this s, apply Lemma 7.3 using an efficient strong $(K = N^\varepsilon, s)$-disperser. In polynomial time we construct a graph G' on N vertices such that if $x \in L$,

$$\chi_f(G') \le s^{-\gamma D} \le 2^{\gamma c n} = N^\varepsilon .$$

If $x \notin L$, then $\alpha(G') \le N^\varepsilon$, so $\chi_f(G') \ge N^{1-\varepsilon}$. Thus it is NP-hard to distinguish graphs with fractional chromatic number N^ε from graphs with fractional chromatic number $N^{1-\varepsilon}$. Converting to chromatic number loses only a logarithmic factor, so the theorem follows. $\square$

To derandomize Khot's results, we use his reduction in place of Theorem 7.1:

Theorem 7.4 ([34]). *For any $\beta > 0$, there is a quasi-polynomial-time reduction from an NP-complete language L to* CHROMATIC NUMBER *with the following properties. On input x of size n, the algorithm outputs a graph $G = (V,E)$ such that*

1. $|V| \le 2^{(\log n)^{1+3\beta}}$;
2. *If $x \in L$ then $\chi_f(G) \le 2^{(\log n)^\beta}$;*
3. *If $x \notin L$ then $\alpha(G) < 2^{-(\log n)^{2\beta}}|V|$.*

We can now show:

Theorem 1.4. For some $\gamma > 0$, it is $\tilde{\mathrm{NP}}$-hard to approximate CHROMATIC NUMBER to within $n/2^{(\log n)^{1-\gamma}}$.

Proof. We use the polynomial-time constructible (N^δ, s)-strong disperser from Theorem 1.9, with $s = 2^{-(\log n)^{2\beta}}$ and δ to be chosen shortly. This has degree $D \le (\log N)/(\delta^c(\log n)^{2\beta})$. Set $\delta = (\log n)^{-\beta/2c}$. Applying Lemma 7.3, it is $\tilde{\mathrm{NP}}$-hard to distinguish between graphs on N vertices with chromatic number N^δ from those with chromatic number $2^{(\log n)^\beta D} \le N^{(\log n)^{-\beta/2}}$. □

8 Simplifying and strengthening additive number theory applications

We now give our simple one-bit condenser and improve other lemmas from [4, 5, 9]. We first define incidences of lines and points.

Definition 8.1. For P a set of points and L a set of lines, $I(P,L)$ denotes the number of *incidences*, i.e., the number of ordered pairs (p,ℓ) where the point p lies on the line ℓ.

We rely heavily on the following theorem on point-line incidences. Bourgain, Katz, and Tao [11] showed how this theorem follows from the sum-product theorem (see Section 2.8). The constant 1.9 below can be increased to any constant less than 2, but the constant α will likely decrease.

Theorem 8.2 (Incidence Theorem [11, 10]). *Let $F = \mathbb{F}_q$, where q is either prime or 2^p for p prime. Let P, L be sets of points and lines in F^2 of cardinality at most $M \le p^{1.9}$. Then there exists an $\alpha > 0$ such that the number of incidences*

$$I(P,L) = O(M^{3/2-\alpha}).$$

8.1 Condensing with one random bit

Barak et al. [5] consider a condenser which uses two extra bits of randomness; here we show that one bit of randomness suffices. Of course, one bit is necessary, so this is optimal. Our proof is also simpler, proceeding directly from the Incidence Theorem 8.2. There is nothing special about the constant .9 below; any constant less than 1 will do.

Our condenser is simple to describe. We work over a field $F = \mathbb{F}_q$, where $q = 2^p$ for p prime. Define the point-line incidence graph as the bipartite graph $G = (V,W,E)$ with vertices $V = F^2$ the set of points, and W the set of lines over F, and (p,ℓ) is an edge iff p and ℓ are incident. Our condenser is based on the function $h : E \to V \times W$ which maps an edge to its two endpoints. An equivalent view of h is the map from F^3 to $(F^2)^2$ which maps (a,b,c) to $((b,ab+c),(a,c))$. This is because the point $(b,ab+c)$ lies on the line $y = ax+c$.

Our condenser $C : F^3 \times \{0,1\} \to F^2$ is simply $C(e,i) = h(e)_i$. The two-bit condenser of Barak, et al. is very similar: their corresponding h maps (a,b,c) to the length 4 vector $(a,b,c,ab+c)$.

Theorem 8.3. *Suppose $\delta \le .9$ and $q^\delta = \omega(1)$. The function C above is a rate-$(\delta \to (1+\alpha/2)\delta, \varepsilon)$-somewhere-condenser, where $\varepsilon = q^{-\alpha\delta/20}$. Here α is the constant from the Incidence Theorem 8.2.*

Before we proceed, it is convenient to introduce a modified notion of somewhere-random source, which we call somewhere light.

Definition 8.4. A vector of sources $X = (X_1,\dots,X_\ell)$ is *ε-close to somewhere-k-light* if the probability, when $(x_1,\dots,x_\ell)$ is output according to X, that no x_i are light is at most ε. We say x_i is light if $\Pr[X_i = x_i] \le 2^{-k}$.

The following lemma describes the relationship between this notion and that of somewhere-random.

Lemma 8.5. *Assume $\ell 2^{-t} < 1-\varepsilon$. If $X = (X_1,\dots,X_\ell)$ is ε-close to somewhere-k-light, then X is $((\ell-1)2^{-t}+\varepsilon)$-close to a somewhere-$(k-t)$-source.*

Proof. Partition the support of X into $\ell+1$ bins so that bin i contains vectors $(x_1,\dots,x_\ell)$ where x_i is light (break ties arbitrarily), and bin 0 contains vectors with no light coordinates. The probability of a bin is the sum of the probabilities of vectors in the bin. By assumption, bin 0 has probability at most ε. Let $\mathrm{bin}(x)$ denote the bin of x. Consider any bin $i \ne 0$ with probability at least 2^{-t} (since $\ell 2^{-t} < 1-\varepsilon$ there is at least one such bin). For any $(x_1,\dots,x_\ell)$ in bin i,

$$\Pr[X_i = x_i \mid \mathrm{bin}(X) = i] \le \Pr[X_i = x_i]/\Pr[\mathrm{bin}(X) = i] \le 2^t \cdot 2^{-k}.$$

Hence, if we let Y^i denote the distribution of X conditional on $\mathrm{bin}(X) = i$, we get that Y^i_i has min-entropy at least $k-t$, and hence Y^i is a somewhere $(k-t)$-source. For any bin i with probability less than 2^{-t}, and for $i=0$, let Y^i be the uniform distribution. Define the distribution $Y = \sum_i \Pr[\mathrm{bin}(X) = i] Y^i$. Then Y is a somewhere $(k-t)$-source and the distance of X to Y comes only from bin 0 and bins with low probability, and is at most $\varepsilon + (\ell-1)2^{-t}$. □

We now work with the modified notion. The main idea is to convert the statistical problem to a counting problem, which we do via the following lemma.

Lemma 8.6. *If (X,Y) is not ε-close to a somewhere-k-source, then there exists sets $S \subseteq \mathrm{supp}(X), T \subseteq \mathrm{supp}(Y)$, $|S|,|T| < 2^{k+1}/\varepsilon$, such that*

$$\Pr[X \in S \wedge Y \in T] > \varepsilon/2.$$

Proof. Let $r = k + \lg(2/\varepsilon)$. By Lemma 8.5, (X,Y) is not $\varepsilon/2$-close to somewhere-r-light. Setting $S = \{s \mid X(s) > 2^{-r}\}$ and $T = \{t \mid Y(t) > 2^{-r}\}$ yields the lemma. □

We can now prove the theorem.

Proof of Theorem 8.3. Instead of C, we analyze the equivalent function h. We may assume that the input to h is uniform on a set of edges of size $K = 2^k = q^{3\delta}$, and set $k' = (1+\alpha/2)(2k/3)$. Suppose the output (X,Y) of h is not ε-close to a somewhere-k'-source. Let $P = S$ and $L = T$ be the sets of size less than $K_0 = 2^{k'+1}/\varepsilon$ given by Lemma 8.6. Assuming without loss of generality that $\alpha \le .1$, note that $K_0 \le q^{2\delta} 1+\alpha/2 \le q^{1.8 \cdot 1.05} < q^{1.9}$.

We calculate the number of incidences $I(P,L)$ in two ways. On the one hand, since each edge is an incident point-line pair, and at least $\varepsilon/2$ fraction of these pairs lie in $P \times L$, the number of incidences $I(P,L) \ge \varepsilon K/2$. On the other hand, by the Incidence Theorem 8.2,

$$I(P,L) = O(K_0^{3/2-\alpha}) = O(K^{(1+\alpha/2)(3/2-\alpha)2/3}/\varepsilon^2) = O(K^{1-\alpha/6}/\varepsilon^2)\,.$$

Combining these, we get a contradiction for $\varepsilon = K^{-\alpha/20}$, and the theorem is proved. □

8.2 AB+C theorem from two sources

In this section and the next, we consider a scenario where we have several independent weak sources, but no truly random seed. The sum-product theorem implies that if A, B, and C are sets of the same size K, then the set $AB+C$ is noticably bigger than K. Barak et al. [4] showed the significantly stronger statistical statement: if A, B, and C are independent distributions with min-entropy k each, then the entropy rate of $AB+C$ is noticably larger than k.

Here we show how to improve the entropy rate with just two sources, by allowing A and C to be correlated. Our proof is also simpler than that in [4]. Again, there is nothing special about the constant .9 below; any constant less than 1 will do.

Theorem 8.7. *Suppose $\delta \le .9$ and $q^\delta = \omega(1)$. If (A,C) and B are output from independent rate-δ-sources, where A,B,C are elements of a field $F = \mathbb{F}_q$, where q is prime or 2^p where p is prime. Then $AB+C$ is $q^{-\alpha\delta/2}$-close to a rate-$(1+\alpha)\delta$-source, where α is the constant from the Incidence Theorem 8.2.*

We prove this using the Incidence Theorem 8.2. The relevance of lines comes in viewing (a,c) as the line $y = ax+c$. In order to get a suitable set of points, we use the following simple lemma. This lemma is key in deducing a statistical theorem, which is about distributions, from the Incidence Theorem 8.2, which just bounds set sizes.

Lemma 8.8. *Suppose X is ε-far from a k-source. Then $\exists S \subseteq \operatorname{supp}(X)$, $|S| < 2^k$, such that* $\Pr[X \in S] \ge \varepsilon$.

Proof. Take $S = \{s \mid X(s) > 2^{-k}\}$, so $|S| < 2^k$. Lemma 2.2 implies that the distance of X to the closest k-source is $\sum_{s\in S}(X(s) - 2^{-k}) \le \Pr[X \in S]$. □

We can now prove the theorem by taking the set of points to be $B \times S$.

Proof of Theorem 8.7. Let (A,C) be output from a flat $2k$-source, and B from an independent flat k-source. Suppose $AB+C$ is ε-far from a k' source, where $k' = (1+\alpha)k$. Let S be the set of size less than $K' = 2^{k'}$ given by Lemma 8.8. Define the set of lines L to be the support of (A,C), where (a,c) is associated with the line $ax+c$. Let P be the set of points $\mathrm{supp}(B) \times S$.

We calculate the number of incidences in two different ways. On the one hand, note that when the line (a,c) applied to b lands in S, it corresponds to an incidence. Since $\Pr[AB+C \in S] \geq \varepsilon$, and since the distributions are flat,

$$I(P,L) \geq \varepsilon |L| \cdot |\mathrm{supp}(B)| = \varepsilon K^3 ,$$

where $K = 2^k \leq |F|^{.9}$. On the other hand, since $|L| = K^2 \leq |F|^{1.8}$ and $|P| \leq K \cdot K' = K^{2+\alpha} \leq |F|^{1.9}$, by the Incidence Theorem 8.2

$$I(P,L) = O(K^{(2+\alpha)(3/2-\alpha)}) = o(K^{3-\alpha/2}) .$$

Hence we may take $\varepsilon = K^{-\alpha/2}$ and the theorem follows. □

8.3 Rate-improving function for two equal-length sources

Note that the previous theorem improves the rate from two independent sources, where one has twice the length of the other. In this subsection, we do this from two sources of equal length, by giving a statistical version of a theorem by Bourgain. Bourgain [9] showed that for a prime q, the function $g : \mathbb{F}_q \times \mathbb{F}_q \to \mathbb{F}_q$ given by $g(x,y) = x(x+y)$ has the following "expanding" property. For $|A| \geq |B| \geq q^\delta$, $\delta < 1$, $g(A,B) \geq q^{\delta+\beta}$ for some $\beta = \beta(\delta) > 0$.[1] With the new sum-product theorem holding also for $q = 2^p$, p prime, Bourgain's theorem will also hold in this case.

We show the statistical analogue of this theorem. Equivalently, our theorem says that the $AB+C$ theorem holds when $C = A^2$, and furthermore the entropy rate is measured with respect to the length of A, rather than (A,C).

Theorem 8.9. *Suppose $\delta \leq .9$ and $q^\delta = \omega(1)$. If X,Y are output from independent rate-δ-sources on $F = \mathbb{F}_q$, then $g(X,Y)$ is $q^{-\alpha\delta/4}$-close to a rate-$(1+\alpha/2)\delta$-source. Here α is the constant from the Incidence Theorem 8.2.*

Proof. We follow Bourgain's proof, but some care is required to make it statistical. Let X and Y be independent random variables uniformly distributed over sets A and B of size q^δ. Assume without loss of generality that they don't contain 0. Suppose $g(X,Y)$ is not ε-close to a $\delta+\beta$-source, where we will choose β and ε later. Let

$$S = \{z \in F \mid \Pr[g(X,Y) = z] > q^{-(\delta+\beta)}\} ,$$

i. e., S is the set of size less than $q^{\delta+\beta}$ guaranteed by Lemma 8.8, such that $\Pr[g(X,Y) \in S] \geq \varepsilon$.

The difficulty in proving the theorem is that directly, this probability being at least ε does not give many lines (in y), so we cannot apply the Incidence Theorem 8.2. We follow Bourgain and find many more lines by exploiting the linearity in y.

To this end, we begin with a collision probability lower bound. Let

$$T = \{y \mid \Pr[g(X,y) \in S] \geq \varepsilon/2\} .$$

[1] Bourgain's proof also uses the Incidence Theorem 8.2, but it was done independently of our use of the Incidence Theorem 8.2 in Subsections 8.1 and 8.2.

Then $|T| \geq \frac{\varepsilon}{2}|B| \geq \frac{\varepsilon}{2}q^{\delta}$. Fix $y' \in T$. Let Z be distributed as X, but independent of X and Y. Then

$$\Pr_{X,Y,Z}[g(X,Y) = g(Z,y')] > \Pr_{Z}[g(Z,y') \in S]q^{-(\delta+\beta)} \geq \frac{\varepsilon}{2}q^{-(\delta+\beta)}.$$

Let X_1 also be distributed as X, but independent of all previously defined random variables. We now show that a function in both X, X_1, and Y, which is linear in Y, still has significant probability of being in S. This will give us many more lines.

$$\begin{aligned}
\Pr_{X,X_1,Z,Y}[X(X+\frac{X_1}{Z}(X_1+Y)-Z) \in S] &\geq \sum_{y' \in T} \Pr[X(X+y') \in S]\Pr[\frac{X_1}{Z}(X_1+Y)-Z = y'] \\
&= \sum_{y' \in T} \Pr[X(X+y') \in S]\Pr[X_1(X_1+Y) = Z(Z+y')] \\
&> |T|\frac{\varepsilon}{2} \cdot \frac{\varepsilon}{2}q^{-(\delta+\beta)} > \frac{\varepsilon^3}{8}q^{-\beta}.
\end{aligned}$$

Therefore, there is a fixed z such that

$$\Pr_{X,X_1,Y}[X(X+\frac{X_1}{z}(X_1+Y)-z) \in S] > \frac{\varepsilon^3}{8}q^{-\beta}. \tag{8.1}$$

This says there are many lines (linear in y) which, when applied to many values of y, land in S. This will contradict the Incidence Theorem 8.2. In particular, let $\ell_{x,x_1}(y)$ denote the line

$$\frac{xx_1}{z}y + \left(x^2 + \frac{xx_1^2}{z} - zx\right),$$

and let L denote the set of all such lines as x, x_1 range over A.

Of course, $|L| \leq |A|^2 = q^{2\delta}$. We also show $|L| \geq q^{2\delta}/3$ by observing that, for fixed z, w, there are at most 3 nonzero solutions in x, x_1 to

$$\begin{aligned}
z &= \frac{xx_1}{z} \\
w &= x^2 + \frac{xx_1^2}{z} - zx.
\end{aligned}$$

We define the points $P = B \times S$, so $|P| < q^{2\delta+\beta}$. By Equation (8.1) and the fact that the pairs (x, x_1) overcount lines by at most a factor of 3, the number of incidences is at least $\frac{\varepsilon^3}{24}q^{3\delta-\beta}$. By the Incidence Theorem 8.2, the number of incidences is $O(q^{(2\delta+\beta)(3/2-\alpha)})$. Comparing these gives the theorem, with $\beta = \alpha\delta/2$ and $\varepsilon = q^{-\alpha\delta/4}$.

□

Acknowledgements

I am grateful to Avi Wigderson for our joint work on Section 5, as well as for other useful discussions. I also thank Ronen Shaltiel, Anup Rao, Mike Saks, Salil Vadhan, Madhu Sudan, and Jesse Kamp for helpful discussions and comments. Finally, I thank the anonymous referees for careful readings and very helpful comments.

References

[1] * M. AJTAI, J. KOMLÓS, AND E. SZEMERÉDI: Deterministic simulation in LOGSPACE. In *Proc. 19th STOC*, pp. 132–140, 1987. [STOC:28395.28410]. 1.3, 3

[2] * N. ALON, U. FEIGE, A. WIGDERSON, AND D. ZUCKERMAN: Derandomized graph products. *Computational Complexity*, 5:60–75, 1995. [Springer:r591795p150lj86q]. 7

[3] * S. ARORA, C. LUND, R. MOTWANI, M. SUDAN, AND M. SZEGEDY: Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45:501–555, 1998. [JACM:278298.278306]. 1.1

[4] * B. BARAK, R. IMPAGLIAZZO, AND A. WIGDERSON: Extracting randomness using few independent sources. In *Proc. 45th FOCS*, pp. 384–393, 2004. [FOCS:10.1109/FOCS.2004.29]. 1.3, 8, 8.2

[5] * B. BARAK, G. KINDLER, R. SHALTIEL, B. SUDAKOV, AND A. WIGDERSON: Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. In *Proc. 37th STOC*, pp. 1–10, 2005. [STOC:1060590.1060592]. 1.3, 3, 3.2, 8, 8.1

[6] * M. BELLARE, O. GOLDREICH, AND M. SUDAN: Free bits, PCPs, and nonapproximability — towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998. [SICOMP:10.1137/S0097539796302531]. 1.1, 1.1, 6, 6.2

[7] * M. BELLARE AND M. SUDAN: Improved non-approximability results. In *Proc. 26th STOC*, pp. 184–193, 1994. [STOC:195058.195129]. 1.1

[8] * R. BOPPANA AND M. HALLDORSSON: Approximating maximum independent sets by excluding subgraphs. *Bit*, 32:180–196, 1992. 1.1

[9] * J. BOURGAIN: More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1:1–32, 2005. [WorldSci:10.1142/S1793042105000108]. 1.3, 8, 8.3

[10] * J. BOURGAIN, A. GLIBICHUK, AND S. KONYAGIN: Estimates for the number of sums and products and for exponential sums in fields of prime order. *Journal of the London Mathematical Society*, 73:380–398, 2006. [Cambridge:10.1112/S0024610706022721]. 1.3, 2.15, 8.2

[11] * J. BOURGAIN, N. KATZ, AND T. TAO: A sum-product estimate in finite fields, and applications. *Geometric and Functional Analysis*, 14:27–57, 2004. [Springer:s00039-004-0451-1]. 1.3, 2.15, 8, 8.2

[12] * B. CHOR AND O. GOLDREICH: Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988. [SICOMP:10.1137/0217015]. 2.4

[13] * A. COHEN AND A. WIGDERSON: Dispersers, deterministic amplification, and weak random sources. In *Proc.30th FOCS*, pp. 14–19, 1989. 1.3

[14] * I.H. DINWOODIE: A probability inequality for the occupation measure of a reversible markov chain. *Annals of Applied Probability*, 5:37–43, 1995. 4

[15] * Z. DVIR AND R. RAZ: Analyzing linear mergers. Technical Report TR05-025, Electronic Colloquium on Computational Complexity, 2005. [ECCC:TR05-025]. 5, 5

[16] * L. ENGEBRETSEN AND J. HOLMERIN: Towards optimal lower bounds for clique and chromatic number. *Theoretical Computer Science*, 299:537–584, 2003. [TCS:10.1016/S0304-3975(02)00535-2]. 1.1

[17] * U. FEIGE: Randomized graph products, chromatic numbers, and the Lovasz θ function. *Combinatorica*, 17:79–90, 1997. [Springer:x785787h43724566]. 7

[18] * U. FEIGE, S. GOLDWASSER, L. LOVASZ, S. SAFRA, AND M. SZEGEDY: Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43:268–292, 1996. [JACM:226643.226652]. 1.1, 6, 6.2

[19] * U. FEIGE AND J. KILIAN: Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57:187–199, 1998. [JCSS:10.1006/jcss.1998.1587]. 1.1, 7, 7.1

[20] * O. GABBER AND Z. GALIL: Explicit construction of linear sized superconcentrators. *Journal of Computer and System Sciences*, 22:407–420, 1981. [JCSS:10.1016/0022-0000(81)90040-4]. 2.5

[21] * D. GILLMAN: A Chernoff bound for random walks on expander graphs. *SIAM Journal on Computing*, 27:1203–1220, 1998. [SICOMP:10.1137/S0097539794268765]. 1.3, 4

[22] * O. GOLDREICH: A sample of samplers – a computational perspective on sampling (survey). Technical Report TR97-020, Electronic Colloquium on Computational Complexity, 1997. [ECCC:TR97-020]. 3

[23] * B. GREEN: Sum-product estimates. Unpublished lecture notes. Available at author's website, 2005. 2.8

[24] * M. HALLDORSSON: A still better performance guarantee for approximate graph coloring. *Information Processing Letters*, 45:19–23, 1993. [IPL:10.1016/0020-0190(93)90246-6]. 1.1

[25] * J. HÅSTAD: Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Mathematica*, 182:105–142, 1999. [Springer:m68h3576646ll648]. 1.1, 6, 6.3

[26] * J. HASTAD AND S. KHOT: Query efficient PCPs with perfect completeness. In *Proc. 42nd FOCS*, pp. 610–619, 2001. [FOCS:10.1109/SFCS.2001.959937]. 6, 6.5

[27] * A. HEALY: Randomness-efficient sampling within NC^1. In *Proc. 10th Intern. Workshop on Randomization and Computation (RANDOM)*, pp. 398–409, 2006. [Springer:b773545612310728]. 4, 4

[28] * S. HOORY, N. LINIAL, AND A. WIGDERSON: Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43:439–561, 2006. 2.5

[29] * R. IMPAGLIAZZO AND D. ZUCKERMAN: How to recycle random bits. In *Proc.30th FOCS*, pp. 248–253, 1989. 1.3

[30] * N. KAHALE: Eigenvalues and expansion of regular graphs. *Journal of the ACM*, 42:1091–1106, 1995. [JACM:210118.210136]. 3

[31] * N. KAHALE: Large deviation bounds for Markov chains. *Combinatorics, Probability, and Computing*, 6:465–474, 1997. [Cambridge:10.1017/S0963548397003209]. 1.3, 4

[32] * R. M. KARP: Reducibility among combinatorial problems. In R. E. MILLER AND J. W. THATCHER, editors, *Complexity of Computer Computations*, pp. 85–103. Plenum Press, New York, 1972. 1.1

[33] * N. KATZ AND C.-Y. SHEN: Garaev's inequality in finite fields not of prime order. Technical report, Arxiv, 2007. [arXiv:math.NT/0703676]. 2.8

[34] * S. KHOT: Improved inapproximability results for MaxClique, Chromatic Number and Approximate Graph Coloring. In *Proc. 42nd FOCS*, pp. 600–609, 2001. [FOCS:10.1109/SFCS.2001.959936]. 1.1, 7.4

[35] * L. LOVASZ: On the ratio of the optimal integral and fractional covers. *Discrete Mathematics*, 13:383–390, 1975. 7

[36] * A. LUBOTZKY, R. PHILIPS, AND P. SARNAK: Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. [Springer:k285687344657q53]. 2.5

[37] * C. LUND AND M. YANNAKAKIS: On the hardness of approximating minimization problems. *Journal of the ACM*, 41:960–981, 1994. [JACM:185675.306789]. 1.1

[38] * G.A. MARGULIS: Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and superconcentrators. *Problems of Information Transmission*, 24:39–46, 1988. 2.5

[39] * M. MORGENSTERN: Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory, Series B*, 62:44–62, 1994. [Elsevier:10.1006/jctb.1994.1054]. 2.5

[40] * E. MOSSEL AND C. UMANS: On the complexity of approximating the VC dimension. *Journal of Computer and System Sciences*, 65:660–671, 2002. [JCSS:10.1016/S0022-0000(02)00022-3]. 1.2

[41] * N. NISAN AND A. TA-SHMA: Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58:148–173, 1999. [JCSS:10.1006/jcss.1997.1546]. 1.2

[42] * N. NISAN AND D. ZUCKERMAN: Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996. [JCSS:10.1006/jcss.1996.0004]. 1.6, 1.2

[43] * R. RAZ: Extractors with weak random seeds. In *Proc. 37th STOC*, pp. 11–20, 2005. [STOC:1060590.1060593]. 1.3, 3, 3.2, 4, 4, 4.3

[44] * O. REINGOLD, R. SHALTIEL, AND A. WIGDERSON: Extracting randomness via repeated condensing. In *Proc. 41st FOCS*, pp. 22–31, 2000. [FOCS:10.1109/SFCS.2000.892008]. 1.3, 2.7

[45] * A. SAMORODNITSKY AND L. TREVISAN: A PCP characterization of NP with optimal amortized query complexity. In *Proc. 32nd STOC*, pp. 191–199, 2000. [STOC:335305.335329]. 6

[46] * R. SHALTIEL: Recent developments in explicit constructions of extractors. *Bulletin of the European Association for Theoretical Computer Science*, 77:67–95, June 2002. 1.2

[47] * A. TA-SHMA, C. UMANS, AND D. ZUCKERMAN: Lossless condensers, unbalanced expanders, and extractors. *Combinatorica*, 27:213–240, 2007. [Springer:y86m43u236782602]. 1.3, 2.7, 2.14

[48] * A. TA-SHMA AND D. ZUCKERMAN: Extractor codes. *IEEE Transactions on Information Theory*, 50:3015–3025, 2004. [IEEE:10.1109/TIT.2004.838377]. 1.1, 1.2

[49] * A. TA-SHMA, D. ZUCKERMAN, AND S. SAFRA: Extractors from Reed-Muller codes. *Journal of Computer and System Sciences*, 72:786–812, 2006. [JCSS:10.1016/j.jcss.2005.05.010]. 1.2

[50] * T. TAO AND V. VU: *Additive Combinatorics*. Cambridge University Press, 2006. 2.8

[51] * C. UMANS: Hardness of approximating Σ_2^p minimization problems. In *Proc. 40th FOCS*, pp. 465–474, 1999. [FOCS:10.1109/SFFCS.1999.814619]. 1.2

[52] * A. WIGDERSON AND D. XIAO: A randomness-efficient sampler for matrix-valued functions and applications. In *Proc. 46th FOCS*, pp. 397–406, 2005. [FOCS:10.1109/SFCS.2005.8]. 1.3, 4

[53] * A. WIGDERSON AND D. ZUCKERMAN: Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999. [Springer:wcjlnyjmdxf30b9x]. 1.2, 1.3, 5, 5

[54] * D. ZUCKERMAN: Simulating BPP using a general weak random source. *Algorithmica*, 16:367–391, 1996. [Algorithmica:kx95d4u1jvyxh882]. 1.1, 1.2, 6

[55] * D. ZUCKERMAN: Randomness-optimal oblivious sampling. *Random Structures and Algorithms*, 11:345–367, 1997. [Wiley:10.1002/(SICI)1098-2418(199712)11:4¡345::AID-RSA4¿3.0.CO;2-Z]. 1.2

AUTHOR

David Zuckerman [About the author]
Department of Computer Science
University of Texas at Austin
1 University Station C0500
Austin, TX 78712
diz@cs.utexas.edu
http://www.cs.utexas.edu/~diz

ABOUT THE AUTHOR

DAVID ZUCKERMAN received his Ph. D. from U.C. Berkeley in 1991 under the supervision of Umesh Vazirani. Since 1994 he has been on the faculty of the University of Texas at Austin. His research interests lie primarily in the role of randomness in computation, particularly pseudorandomness and its relationship to other topics in complexity theory, coding theory, and cryptography. His hobbies include dancing and Scrabble.