Electronic Colloquium on Computational Complexity, Report No. 100 (2005)

Linear Degree Extractors and the Inapproximability of MAX CLIQUE and CHROMATIC NUMBER

David Zuckerman*
Department of Computer Science
University of Texas at Austin
1 University Station C0500
Austin, TX, 78712
diz@cs.utexas.edu

September 12, 2005

Abstract

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 $\log n + O(1)$ additional random bits for sources with constant entropy rate. We further construct dispersers, which are similar to one-sided extractors, which use an arbitrarily small constant 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 $\alpha > 0$.

We use our dispersers to derandomize the results of Hastad [Hås99] and Feige-Kilian [FK98] and show that for all $\epsilon > 0$, approximating MAX CLIQUE and CHROMATIC NUMBER to within $n^{1-\epsilon}$ are NP-complete. We also derandomize the results of Khot [Kho01] 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 $\mathrm{N\tilde{P}} = \tilde{\mathrm{P}}$.

Our constructions rely on recent results in additive number theory and extractors by Bourgain-Katz-Tao [BKT04], Barak-Impagliazzo-Wigderson [BIW04], Barak-Kindler-Shaltiel-Sudakov-Wigderson [BKS+05], and Raz [Raz05]. We also simplify and slightly strengthen key theorems in the second and third of these papers, and strengthen a related theorem by Bourgain [Bou05].

*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 Grant CCR-0310960.

ISSN 1433-8092

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 [Kar72]. The best approximation algorithms for these problems are of the form $n/\text{polylog}(n)$ [BH92, Hal93], which is not much better than the trivial approximation of n. Yet no strong inapproximability results were known until Feige et al. [FGL+96] discovered a connection between probabilistically checkable proofs (PCPs) and MAX CLIQUE. The celebrated PCP Theorem of Arora et al. [ALM+98] then implied that it is NP-complete to approximate MAX CLIQUE to within n^c for some constant $c > 0$. This ratio was improved in [BS94, BGS98] until Hastad, in a breakthrough, showed a hardness ratio of $n^{1-\epsilon}$, for any $\epsilon > 0$ [Hås99]. The catch is that Hastad's reduction is randomized, so his theorem assumes that NP $\neq$ ZPP. Assuming only NP $\neq$ P, Hastad's hardness ratio becomes $n^{1/2-\epsilon}$. In this paper we derandomize Hastad's randomized reduction:

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

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-complete to approximate CHROMATIC NUMBER to within n^c for some constant $c > 0$ [LY99]. Other reductions ensued, culminating in Feige and Kilian's proof of a hardness ratio of $n^{1-\epsilon}$ [FK98]. This uses Hastad's result, so it assumes that NP $\neq$ ZPP. Assuming only NP $\neq$ P, the best previous hardness ratio explicitly stated appears to be $n^{1/7-\epsilon}$ [BGS98]. Previous work likely implied something better, though certainly no better than $n^{1/2-\epsilon}$. In this paper we derandomize Feige and Kilian's result:

Theorem 2. *For all $\epsilon > 0$, it is* NP*-complete to approximate* CHROMATIC NUMBER *to within $n^{1-\epsilon}$.*

Engebretsen and Holmerin [EH03] improved the hardness ratios for both problems to $n^{1-o(1)}$ under the stronger assumption that NP $\not\subseteq$ ZPTIME$(2^{\text{polylog}(n)})$. Khot [Kho01] 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 $\text{N}\tilde{\text{P}}$-completeness, which is the quasi-polynomial analogue of NP-completeness. As we consider quasi-polynomial time reductions, $\text{N}\tilde{\text{P}}$-completeness is weaker than NP-completeness; see Subsection 2.1 for more details.

Theorem $\tilde{1}$. For some $\gamma > 0$, it is $\text{N}\tilde{\text{P}}$-complete to approximate MAX CLIQUE to within $n/2^{(\log n)^{1-\gamma}}$.

Theorem $\tilde{2}$. For some $\gamma > 0$, it is $\text{N}\tilde{\text{P}}$-complete 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 a variant of a randomness extractor. Good dispersers were known to help derandomize inapproximability results for MAX CLIQUE (e.g., [Zuc96, TZ04], 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 1.1. *The* min-entropy *of a distribution X is $H_\infty(X) = \min_a\{-\log_2 X(a)\}$. 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 an algorithm which extracts randomness from a k-source using a few additional truly random bits.

Definition 1.2. *[NZ96] Let U_ℓ denote the uniform distribution on ℓ bits. A function 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 Ext(X,U_d) is ϵ-close in statistical (variation) distance to U_m. We say Ext is a* strong (k,ϵ)*-extractor if for every k-source X, the distribution Ext$(X,Y) \circ Y$ is ϵ-close to U_{m+d}, where Y is chosen from U_d. Here $\circ$ 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. These include the focus of this paper, inapproximability [Zuc96, Uma99, MU02], as well as pseudorandom generators for space-bounded computation [NZ96], expanders that beat the second eigenvalue bound [WZ99], random sampling using few random bits [Zuc97], cryptography [Lu02, Vad03, CDH+00, DS02], error-correcting codes with strong list decoding properties [TZ04], superconcentrators and non-blocking networks [WZ99], sorting and selecting in rounds [Pip87], time versus space complexities [Sip88], and implicit data structures [FN93, Zuc91].

We wish to construct extractors for any min-entropy k with t, 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" [WZ99]. Starting with the first extractor of Nisan and Zuckerman [NZ96], a lot of effort has been expended constructing good extractors. See the surveys [Sha02, NT99, Nis96] 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 $\text{Ext}(x,y) = z$. 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 [Zuc97] the degree is the number of samples; in the extractor codes of [TZ04] D is the length of the code; in the simulation of BPP using weak sources [Zuc96] the degree is the number of calls to the BPP algorithm. Most relevant for us, in the inapproximability of MAX CLIQUE [Zuc96] the size of the graph is closely related to D.

Before the work of Ta-Shma et al. [TZS01], 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 $D = O(n \log^* n)$, but then 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 $D = n \cdot \text{polylog}(n)$. Our new construction achieves linear degree and linear output length for constant-rate sources.

Theorem 3. *For all $\alpha, \delta, \epsilon > 0$ there is an efficient family of strong $(k = \delta n, \epsilon)$-extractors Ext : $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $m \geq (1-\alpha)\delta n$ and $D = 2^d = O(n)$.*

Dispersers are one-sided analogues of extractors, and are defined in Section 2. When the error is allowed to be very close to 1, say $1 - s$, non-explicit dispersers can have degree even smaller than n, namely $O(n/\log s^{-1})$. In this paper, we succeed in matching this degree for constant-rate sources. These dispersers are the key for our inapproximability results.

Theorem 4. *For all $\alpha, \delta > 0$ 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 $m \geq (1-\alpha)\delta n$.*

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 [AKS87, IZ89, CW89]. This yields a disperser for sources with entropy rate greater than 1/2 [CW89]. By using Chernoff bounds for random walks on expanders [Gil98, Kah97, WX05], 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 [RSW00, TUZ01]. We condense using techniques developed from additive number theory. Bourgain-Katz-Tao [BKT04] showed that, in a prime field $\mathbb{F}_p$, if $|A| \leq p^{.9}$, then $\max(|A+A|, |A \cdot A|) \geq |A|^{1+\alpha}$. Barak-Impagliazzo-Wigderson [BIW04] used these ideas to show that if A, B, and C are independent rate-δ-sources with $\delta \leq .9$, then $AB + C$ is close to a rate-$(1+\alpha')\delta$-source. Using this recursively, as proposed in [Zuc90], they showed how to extract randomness from a constant number of independent sources with constant entropy rate. Barak-Kindler-Shaltiel-Sudakov-Wigderson [BKS+05] and Raz [Raz05] further built on these ideas to extract randomness from three independent sources with constant entropy rate.

A key tool they developed was a condenser, which condenses a source of constant entropy rate to a source with entropy rate at least .9, using only a constant number of additional random bits. The Raz condenser is stronger in that most outputs are good. By iteratively applying the Raz condenser in a manner similar to [WZ99], we can improve the output length to be $1-\alpha$ fraction of the randomness, for any $\alpha > 0$.

We simplify and slightly strengthen key lemmas in Barak-Impagliazzo-Wigderson [BIW04] and Barak et al. [BKS+05]. First, we show that in the $AB+C$ lemma, A and C do not have to be independent. Instead, the lemma follows if (A, C) is independent from B. Second, we strengthen a theorem of Bourgain [Bou05] and show that the function $A(A+B)$ also gives a rate improvement. Third, we show that the basic condenser in [BKS+05] can be made stronger, in that it suffices to have two outputs instead of four. Our proofs of the first and third of these theorems are simple, given the point-line incidence theorem from Bourgain-Katz-Tao [BKT04].

2 Preliminaries

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 $\mathrm{N\tilde{P}}$-Completeness

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 $\mathrm{N\tilde{P}}$-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 $\mathrm{N\tilde{P}} \subseteq \mathrm{TIME}(2^{(\log n)^{c+1}})$, contradicting the hierarchy theorems.

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

Finally, we only prove the hardness directions in our completeness results, as it is easy and well-known that MAX CLIQUE and CHROMATIC NUMBER are in $\mathrm{NP} \subseteq \mathrm{N\tilde{P}}$.

2.2 Distance Between Distributions

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*

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

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

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

Lemma 2.2. *Let $p_x = \Pr[X = x]$. The distance of X to the closest ℓ-source is $\sum_x \max(p_x - 2^{-\ell}, 0)$.*

Of course, only those x with $p_x > 2^{-\ell}$ 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 following lemma shows that it suffices to consider flat k-sources.

Lemma 2.4. *[CG88] A k-source is a convex combination of flat k-sources.*

2.4 Dispersers

Dispersers are usually defined with respect to an error parameter ϵ. For us it is more convenient to use the parameter $s = 1 - \epsilon$.

Definition 2.5. *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 $\Gamma_y(X) = \{\mathrm{DIS}(x, y) | x \in X\}$, and let $\Gamma(X) = \cup_y \Gamma_y(X)$ be the 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, for any $X \subseteq [N]$ with $|X| \geq K$, there is a y such that $|\Gamma_y(X)| \geq sM$.*

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 2.6. *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 will be useful when $D = O(1)$.

Lemma 2.7. *A (K, s)-disperser is also a strong $(K, s/D)$-disperser.*

We also use the following simple lemma.

Lemma 2.8. *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 $\mathrm{DIS} : [N] \times [D_1 D_2] \to [M]$.*

Proof. Take $\mathrm{DIS}(x, (y_1, y_2)) = \mathrm{DIS}_2(\mathrm{DIS}_1(x, y_1), y_2)$. □

2.5 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.

2.6 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 $(k \to \ell, \epsilon)$-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-$(k/n \to \ell/m, \epsilon)$-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 $(k \to \ell, \epsilon)$-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-$(k/n \to \ell/m, \epsilon)$-somewhere-condenser.*

Note that a $(k \to \ell, \epsilon)$-strong-condenser is a $(k \to \ell, \epsilon)$-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 $(k \to \ell, \epsilon)$-somewhere-condenser, then it is a $(2^k, (1-\epsilon)2^{\ell - m})$-disperser.*

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

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, the distribution $(Z_2 | Z_1 = z_1)$ is ϵ_2-close to an ℓ_2-source. Then $Z_1 \circ Z_2$ is $\epsilon_1 + \epsilon_2$-close to an $\ell_1 + \ell_2$-source*

Proof. Denote $p_{z_1} = \Pr[Z_1 = z_1]$ and $p_{z_2|z_1} = \Pr[Z_2 = z_2 | Z_1 = z_1]$. We will use Lemma 2.2 to bound the distance of $Z_1 \circ Z_2$ to the closest $\ell_1 + \ell_2$-source. This distance is the following, where we sum over all z_1 and, for each z_1, the 2^{ℓ_2} strings z_2 with the highest $p_{z_2|z_1}$.

$$\begin{aligned}
\sum_{z_1,z_2} \max(p_{z_1} \cdot p_{z_2|z_1} - 2^{-(\ell_1+\ell_2)}, 0) &\leq \sum_{z_1,z_2} \max(p_{z_1} \cdot p_{z_2|z_1} - p_{z_1} 2^{-\ell_2}, 0) + \sum_{z_1,z_2} \max(p_{z_1} 2^{-\ell_2} - 2^{-(\ell_1+\ell_2)}, 0) \\
&= \sum_{z_1,z_2} p_{z_1} \max(p_{z_2|z_1} - 2^{-\ell_2}, 0) + \sum_{z_1,z_2} 2^{-\ell_2} \max(p_{z_1} - 2^{-\ell_1}, 0) \\
&= \mathrm{E}_{z_1}\left[\sum_{z_2} \max(p_{z_2|z_1} - 2^{-\ell_2}, 0)\right] + \sum_{z_1} \max(p_{z_1} - 2^{-\ell_1}, 0) \\
&\leq \epsilon_2 + \epsilon_1.
\end{aligned}$$

□

We build extractors by first condensing and then applying a weaker extractor. The idea of condensing before extracting was used in [RSW00, TUZ01], and a simple lemma from [TUZ01] shows that this works.

Lemma 2.14. *[TUZ01] Suppose $C : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{n'}$ is an efficient (strong) $(k \to \ell, \epsilon_1)$-condenser, and Ext : $\{0,1\}^{n'} \times \{0,1\}^{d_2} \to \{0,1\}^m$ is an efficient (strong) (ℓ, ϵ_2)-extractor. Then Ext$'(x, y_1 \circ y_2) =$ Ext$(C(x, y_1), y_2)$ is an efficient (strong) $(k, \epsilon_1 + \epsilon_2)$-extractor.*

2.7 Generating Primes

Our $n^{1-o(1)}$-factor inapproximability results, as well as our extractor and disperser constructions for sub-constant entropy rate, require a large prime. Unfortunately, there is no known efficient deterministic algorithm to find a large prime. This won't be a problem for our inapproximability results, because the weaker polynomial-time constructibility suffices. The corresponding extractor and disperser constructions are only efficient under a well-known conjecture about the gaps between consecutive primes.

Definition 2.15. *Let Prime(n) denote the time of the fastest deterministic algorithm which, on input n in unary, outputs with high probability an ℓ-bit prime, where $n \leq \ell \leq n + \sqrt{n}$.*

Of course, Prime$(n) \leq 2^{2n}$. Cramer [Cra37] conjectured that there is always a prime between N and $N + O(\log^2 N)$, but the best known results are of the form $N + N^c$. Under Cramer's conjecture, the deterministic primality test [AKS04] implies that Prime$(n) = \mathrm{poly}(n)$. In particular, under this conjecture there is a polynomial-time algorithm which outputs a prime with exactly n bits.

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 $s = s(n), \alpha > 0$, there is a $\beta > 0$ and an efficient family of strong $(K = N^{1-\beta}, s)$-dispersers* DIS : $[N = 2^n] \times [D] \to [M = 2^m]$ *such that $D \leq n / \lg s^{-1}$ and $m \geq (1-\alpha)n$.*

Proof. We use the disperser of [AKS87]. Set $c = \lg s^{-1}$, and $m = (1-\alpha)n$. Let G be a 2^c-regular expander on $[2^m]$ with $\lambda = \lambda(G) \leq 2^{1-c/2}$. To find the neighbors of a vertex $u \in [2^n]$, use the n bits defining u to choose a random vertex $v_1 \in [2^m]$ and then take a random walk $v_1, \ldots, v_D$ on G. Connect u to $v_1, \ldots, v_D$. Note that $n = m + (D-1)c$, so $D = 1 + (n-m)/c$ and $(n-m)/c \leq D \leq n/c = n/\lg s^{-1}$.

We use the tight analysis given by Kahale [Kah95]. 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 $s = 2^{-c/2}$, this probability is less than $2^{(2-c/2)D} = 2^{-an}$ for some constant a. Therefore, this is a $(K = N^{1-\beta}, s)$-disperser for any $\beta < a$.

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, where each $|S_i| \leq s2^m$. By the result of Kahale mentioned in [Gol97],

$$\Pr[(\forall i) v_i \in S_i] \leq s(s + (1-s)\lambda^2)^{(D-1)/2} < (s+\lambda^2)^{D/2}.$$

This bound is at most square root of Kahale's earlier bound, so this probability is at most $2^{-(a/2)n}$, and we now choose $\beta < a/2$. □

To give a construction for all positive entropy rates, we use the following theorem, which follows from the condenser in [BKS$^+$05] or [Raz05]. While [BKS$^+$05] only gives an ordinary disperser, by Lemma 2.7 it is also a strong disperser for essentially the same parameters, since D is constant.

Theorem 3.2. *[BKS$^+$05, Raz05] For any $\beta, \delta > 0$, there is an efficient family of rate-$(\delta \to 1-\beta, \epsilon = 2^{-\Omega(n)})$-somewhere-condensers $C : [N = 2^n] \times [D] \to [M = 2^m]$ where $D = O(1)$ and $m = \Omega(n)$.*

Remark 3.3. *For subconstant $\delta = \delta(n)$, the dependence on $\delta = \delta(n)$ is $D = (1/\delta)^{O(1)}$ and $m = \delta^{O(1)}n$. However, in this case the construction requires a large prime and runs in time $\mathrm{poly}(n) + \mathrm{Prime}(n)$ (see Subsection 2.7). Hence it is efficient under Cramer's conjecture and polynomial-time constructible without assumptions. Using the reduction in [BKS$^+$05], instead of the prime being generated deterministically, it suffices for the n-bit prime to be generated using $k = \delta n$ random bits.*

Applying Lemmas 2.12 and 2.7, we deduce

Corollary 3.4. *For any $\beta, \delta > 0$, there is an efficient family of strong $(K = N^\delta, M^{-\beta})$-dispersers* $\mathrm{DIS} : [N = 2^n] \times [D] \to [M = 2^m]$ *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* $\mathrm{DIS} : [N = 2^n] \times [D] \to [M = 2^m]$ *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 $m = \delta^{O(1)}n$, but in this case* DIS *is efficient under Cramer's conjecture and polynomial-time constructible without assumptions.*

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 $\mathrm{DIS}_2 : [N'] \times [D_2 \leq n'/\lg s^{-1}] \to [M = 2^m]$, be an efficient strong $(K' = (N')^{.9}, s)$-disperser given by Proposition 3.1, with $m = n'/2$. Applying Lemma 2.8 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 this section, as dispersers suffice to prove those results.

4.1 Basic Construction

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 [Raz05] to reduce to the high-entropy case.

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

Proof. For $m = (1-\alpha)n$ and $c = 3$ (say), let G be a 2^c-regular expander on $[2^m]$ with $\lambda = \lambda(G) \leq 2^{1-c/2}$. On input $x \in [2^n]$ and $y \in [D]$, use the n bits defining x to choose a random vertex $v_1 \in [2^m]$ and then take a random walk $v_1, \ldots, v_D$ on G. Output v_y. Note that $n = m + (D-1)c$, so $D = 1 + (n-m)/c = 1 + \alpha n/c \leq \alpha n$.

Let $S \subseteq [2^m]$ have density $\mu = |S|/2^m$. Further let the random variable $\hat{\mu}$ denote the fraction of v_i which are in S. By a theorem of Gillman [Gil98], which was achieved independently by Kahale [Kah97],

$$\Pr[|\hat{\mu} - \mu| \geq \epsilon] \leq 3\exp((1-\lambda)D/\epsilon^2).$$

This error is 2^{-an} for some constant $a = a(\lambda, \epsilon)$. Thus this is a $(k = (1-\beta)n, \epsilon)$-extractor for $\beta < a$. □

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

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

Proof. We use the same construction. For the proof, we now consider $S \subseteq [D] \times [2^m]$, so $S = \cup_i \{i\} \times S_i$. We now use the Chernoff bound which handles the case when we count the number of times that the ith step of the random walk lands in S_i. This Chernoff bound was recently proved by Wigderson and Xiao [WX05]. □

Theorem 4.3. *[Raz05] For any constants $\beta, \delta, \epsilon > 0$, there is a constant d such that there is an efficient rate-$(\delta \to (1-\beta), \epsilon)$-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 $\delta, \epsilon > 0$ there is an efficient family of strong-$(k = \delta n, \epsilon)$-extractors $Ext : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $m = \Omega(n)$ and $D = 2^d = O(n)$.*

4.2 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 [WZ99]: if the output 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. *For any constants $\alpha, \beta, \delta, \epsilon > 0$, there is a constant d such that there is an efficient rate-$(\delta \to (1-\beta), \epsilon)$-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 βn, we would like to iterate about $1/\beta$ times, but we cannot do this if the initial error is bigger than β. Hence we need an improved version of this condenser, which follows from the improved merger of Dvir and Raz [DR05].

Lemma 4.5. *For any $\delta > 0$, there exists $\beta > 0$, such that for any $\epsilon > 0$, there is a constant d such that there is an efficient rate-$(\delta \to (1-\delta), \epsilon)$-strong condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that $m \geq \beta n$.*

Proof. (Sketch) Start with the somewhere condenser of Theorem 3.2 to boost the "somewhere-rate" to $1 - \delta/2$. The error is exponentially small. Now apply the strong merger of [DR05]. Choose the output length to ensure that the final rate is at least $1 - \delta$. Note that this choice is independent of the error. By the properties of the merger, we can make the error arbitrarily small at the expense of increasing the seed length d. □

The following lemma is the condenser analogue to the corresponding extractor lemma in [WZ99]. The notation $\circ$ denotes concatenation.

Lemma 4.6. *Suppose $C_1 : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{m_1}$ is a strong $(k \to \ell_1, \epsilon_1)$-condenser and $C_2 : \{0,1\}^n \times \{0,1\}^{d_2} \to \{0,1\}^{m_2}$ is a strong $(k - m_1 - s \to \ell_2, \epsilon_2)$-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 $(k \to \ell_1 + \ell_2, \epsilon_1 + \epsilon_2 + 2^{-s})$-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: $S = \{z | \Pr_X[C_1(X, y_1) = z] \leq 2^{-(m_1+s)}\}$. Then $\Pr[C_1(X, y_1) \in S] \leq |S| 2^{-(m_1+s)} \leq 2^{-s}$. 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 $\epsilon_2^{y_2}$ of some ℓ_2-source. Putting this together as in Lemma 2.13, $C(x, y_1 \circ y_2)$ is within $\epsilon_1^{y_1} + 2^{-s} + \epsilon_2^{y_2}$ of some $\ell_1 + \ell_2$-source. Since the average of $\epsilon_i^{y_i}$ is at most ϵ_i, this completes the proof of the lemma. □

Applying this lemma inductively, we can show:

Lemma 4.7. *Suppose $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is an efficient strong $(k \to \ell, \epsilon)$-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 $(k + (t-1)m + s \to t\ell, t\epsilon + (t-1)2^{-s})$-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 4.6 gives the condenser for $t + 1$. □

As Anup Rao pointed out, if we are applying a condenser many times, we can do better in some parameters. The following lemma does this, but it does not dominate the other lemma. We don't use this lemma, but include it for the interested reader.

Lemma 4.8. *Suppose $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a strong $(k \to \ell, \epsilon)$-condenser. Then $C' : \{0,1\}^n \times \{0,1\}^{td} \to \{0,1\}^{tm}$ given by $C'(x, y_1 \circ y_2 \circ \ldots \circ y_t) = C(x, y_1) \circ \ldots \circ C(x, y_t)$ is a strong $(k' = k + t\ell \to t\ell, t\epsilon)$-condenser.*

Proof. (Sketch.) Let X be a k'-source. For fixed $y_1, \ldots, y_i, z_1, \ldots, z_i$, we will be interested in whether $H_\infty(X|C(X, y_1) = z_1, \ldots, C(X, y_i) = z_i) \geq k$. If it is, then we will be able to use the condenser with y_{i+1}. If it is not, then $\Pr[C(X, y_1) = z_1, \ldots, C(X, y_i) = z_i] < 2^{k-k'} = 2^{-t\ell}$, which is also good. □

We can now prove Theorem 5. Note that the min-entropy rate of the output is preserved, and the main loss of entropy is with the initial choice of k versus output m. Therefore, the main point is to choose the initial k small enough.

Proof of Theorem 5. Let $\alpha, \beta, \delta, \epsilon > 0$ be given. By Lemma 4.5, for some $\gamma > 0$ there is an efficient strong rate-$(\alpha\delta \to (1-\beta), \epsilon')$-condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, where ϵ' will be chosen later and $m \geq \gamma n$. Set $i = (1-\alpha)\delta/\gamma$ and apply Lemma 4.7 with an s to be chosen later. This gives an efficient strong $(\delta n - \gamma n + s \to (1-\beta)(im), i\epsilon' + 2^{-s})$-condenser $C' : \{0,1\}^n \times \{0,1\}^{id} \to \{0,1\}^{im}$. Choosing $s = \gamma n$ and ϵ' small enough so $i\epsilon' + 2^{-s} \leq \epsilon$ gives the theorem. □

Combining Theorem 5 with Proposition 4.1, we obtain our main extractor construction:

Theorem 3. For all $\alpha, \delta, \epsilon > 0$ there is an efficient family of strong $(k = \delta n, \epsilon)$-extractors $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with $m \geq (1-\alpha)\delta n$ and $D = 2^d = O(n)$.

Similarly we obtain our disperser:

Theorem 4. For all $\alpha, \delta > 0$ 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 $m = \delta^{O(1)} n$, but in this case DIS is efficient under Cramer's conjecture and polynomial-time constructible without assumptions.

5 The Inapproximability of MAX CLIQUE

In this section, 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. [FGL+96] were the first to show how to obtain inapproximability results using PCPs. Bellare, Goldreich, and Sudan [BGS98] showed that free bit complexity is the measure which gives best inapproximability results.

Definition 5.1. $\mathrm{FPCP}_s(r, f)$ *is the class of promise problems recognized by PCP verifiers using r random bits and f free bits, achieving perfect completeness and soundness s.*

Hastad [Hås99] showed how to reduce the soundness by paying only a tiny amount in the free bit complexity. Specifically, he showed:

Theorem 5.2. *[Hås99] 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 following follows from the reductions in [FGL+96, BGS98] and the amplification of a PCP via a good disperser, as first suggested in [Zuc96].

Lemma 5.3. *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* $\mathrm{NP} \subseteq \mathrm{FPCP}_{K/2^R}(R, Df)$. *Hence it is* NP*-hard to distinguish graphs G with clique number at least 2^R from graphs with clique number at most K, where $|V(G)| = 2^{R+Df}$.*

This suffices to prove our theorem.

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

Proof. Fix $\epsilon > 0$. We first claim that there is a function $R = O(\log n)$ such that $\mathrm{NP} \subseteq \mathrm{FPCP}_{2^{(\epsilon-1)R}}(R, \epsilon R)$. To see this, note that Theorem 4 says that for any $s = s(n)$ there is an efficient family of $(K = N^\epsilon, s)$-dispersers of degree $D \leq c(\log N)/\log s^{-1}$, for some $c = c(\epsilon)$. Let $\bar{f} \leq \epsilon/c$, and apply Theorem 5.2 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 $(K = (2^R)^\epsilon, 2^{-\ell})$-disperser $\mathrm{DIS} : [2^R] \times [D] \to [2^r]$. Apply Lemma 5.3 with this disperser, and note that $Df \leq (cR/\ell) \cdot (\ell\bar{f}) = \bar{f} \cdot cR \leq \epsilon R$. Since the output length is linear in the input length, $R = O(\log n)$. This proves the claim.

Lemma 5.3 then implies that it is NP-complete to distinguish clique size $2^{\epsilon R}$ from 2^R in graphs on $2^{(1+\epsilon)R}$ vertices, which gives the theorem. □

To obtain inapproximability up to an $n^{1-o(1)}$ factor, we can use the following theorem by Hastad and Khot [HK01], which is basically the same as that obtained by Samorodnitsky and Trevisan [ST00] but gives perfect completeness.

Theorem 5.4. *[HK01] 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 $\tilde{1}$. For some $\gamma > 0$, it is $\tilde{\mathrm{NP}}$-complete to approximate MAX CLIQUE to within $n/2^{(\log n)^{1-\gamma}}$.

Proof. Set $\epsilon = \epsilon(n) = 1/\log n$. We prove the theorem by showing that there is a function $R = \mathrm{polylog}(n)$ such that $\mathrm{NP} \subseteq \mathrm{FPCP}_{2^{(\epsilon-1)R}}(R, \epsilon R)$. By Theorem 4, there is a c such that for any $s = s(n)$ there is a polynomial-time constructible family of $(K = N^\epsilon, s)$-dispersers of degree $D \leq (\log n)^c(\log N)/\log s^{-1}$. Let $\ell = 9(\log n)^{2(c+1)}$ and $s = 2^{-\ell}$. We'll use the polynomial-time constructible $(K = (2^R)^\epsilon, 2^{-\ell})$-disperser $\mathrm{DIS} : [2^R] \times [D] \to [2^r]$. Apply Theorem 5.4 to get $r = r(n) = \mathrm{polylog}(n)$ such that $\mathrm{NP} \subseteq \mathrm{FPCP}_s(r, 3\sqrt{\ell})$. Apply Lemma 5.3 with disperser DIS, and note that $Df \leq (R(\log n)^c/\ell) \cdot (3\sqrt{\ell}) = R/\log n = \epsilon R$. Since $R = \mathrm{polylog}(n)$, this proves the claim. □

6 The Inapproximability of CHROMATIC NUMBER

Now we show how our dispersers also imply the NP-completeness of approximating CHROMATIC NUMBER to within $n^{1-\epsilon}$ for any $\epsilon > 0$. We derandomize Feige & Kilian's proof [FK98] 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 [Lov75]. Feige

and Kilian amplify the hardness ratio using randomized graph products. That is, they start with a graph G (from a family of graphs) which has constant hardness ratio, and build a subgraph G' of the product graph G^D which has hardness ratio $|V(G')|^{1-\epsilon}$. G^D is defined with respect to the following "OR" graph product.

Definition 6.1. *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 $\chi_f(G \times H) = \chi_f(G) \cdot \chi_f(H)$ [Fei97].

We derandomize the randomized graph powering. This was done earlier in the clique setting [AFWZ95], 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 $\chi_f(G') \leq \chi_f(G^D) = \chi_f(G)^D$.

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 $\mathrm{DIS}(X) = \{\mathrm{DIS}(x) | x \in X\}$. Now define the graph $\mathrm{DIS}(G^D)$ to be the induced subgraph of G^D on vertex set $\mathrm{DIS}(X)$.

Lemma 6.2. *Given a graph $G = (V, E)$ and a disperser* DIS *with degree D, let $G' = (V', E') = \mathrm{DIS}(G^D)$. Then*

1. *$\chi_f(G') \leq (\chi_f(G))^D$.*
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 $\chi_f(G') \leq \chi_f(G^D) = (\chi_f(G))^D$. For the second part, suppose $\alpha(G') \geq K$, and let X be an independent set in G' of size K. Note that $\Gamma_i(X)$ corresponds to the ith coordinates of X. By the strong disperser property, for some $i \in [D]$, $|\Gamma_i(X)| \geq s|V| > \alpha(G)$. 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. □

Feige and Kilian gave the following reduction:

Theorem 6.3. *[FK98] 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 $\chi_f(G) \leq s^{-\gamma}$;*
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 state their theorem slightly differently: for any $\gamma, \ell > 0$, they can set $s = O(2^{-\ell})$ and if $x \in L$ then $\chi_f(G) \leq 2^{3\gamma\ell+1}$. This is equivalent to our statement above, for a slightly different choice of γ.)

We are now ready to prove our theorem.

Theorem 2. It is NP-complete to approximate CHROMATIC NUMBER to within $n^{1-\epsilon}$ for any $\epsilon > 0$.

Proof. Theorem 4 says that for any $s = s(n)$ there is an efficient family of strong $(K = N^\epsilon, s)$-dispersers of degree $D \leq cn/\log s^{-1}$, for some $c = c(\epsilon)$. Set $\gamma = \epsilon/c$, and use the Feige-Kilian reduction, which comes with an $s = s(\gamma)$. Using this s, apply Lemma 6.2 using an efficient strong $(K = N^\epsilon, s)$-disperser. In polynomial time we construct a graph G' on N vertices such that if $x \in L$,

$$\chi_f(G) \leq s^{-\gamma D} \leq 2^{\gamma c n} = N^\epsilon.$$

If $x \notin L$, then $\alpha(G') \leq N^\epsilon$. Thus it is NP-complete to distinguish graphs with chromatic number N^ϵ from graphs with clique number N^ϵ, and the theorem follows. □

To derandomize Khot's results, we use his reduction:

Theorem 6.4. *[Kho01] 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. $|V| \leq 2^{(\log n)^{1+3\beta}}$;
2. *If $x \in L$ then $\chi_f(G) \leq 2^{(\log n)^\beta}$;*
3. *If $x \notin L$ then $\alpha(G) < 2^{-(\log n)^{2\beta}}|V|$.*

We can now show:

Theorem $\tilde{2}$. For some $\gamma > 0$, it is $\tilde{\text{NP}}$-complete 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 4, with $s = 2^{-(\log n)^{2\beta}}$ and δ to be chosen shortly. This has degree $D \leq (\log N)/(\delta^c (\log n)^{2\beta})$. Set $\delta = (\log n)^{-\beta/2c}$. Applying Lemma 6.2, it is $\tilde{\text{NP}}$-hard to distinguish between graphs on N vertices with chromatic number N^δ from those with chromatic number $2^{(\log n)^\beta D} \leq N^{(\log n)^{-\beta/2}}$. □

7 Simplifying and Strengthening Additive Number Theory Applications

7.1 Point-Line Incidences

The following theorem on point-line incidences will be critical in our improvements of the lemmas from [BIW04, BKS+05]. It was proved by Bourgain, Katz, & Tao [BKT04] for more general fields, but required a lower bound on M. Konyagin [Kon03] eliminated the need for this lower bound over prime fields. The constant 1.9 below may be replaced by any constant less than 2, but the constant α will likely decrease.

Incidence Theorem. [BKT04, Kon03] Let $F = \mathbb{F}_p$ be a prime field and let P, L be sets of points and lines in F^2 of cardinality at most $M \leq p^{1.9}$. Then there exists an $\alpha > 0$ such that the number of incidences

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

7.2 AB+C Theorem from Two Sources

In this section, we will consider a scenario where we have several independent weak sources, but no truly random seed. The basic lemma of [BIW04] shows how to improve the entropy rate with three sources A, B, and C, by computing $AB + C$. Here we show how to do it with just two, by allowing A and C to be correlated. Our proof is also simpler than that in [BIW04]. There is nothing special about the constant .9 below; any constant less than 1 will do.

Theorem 6. *Suppose $\delta \leq .9$. If (A, C) and B are output from independent rate-δ-sources, then $AB + C$ is $2^{-\Omega(\alpha k)}$-close to a rate-$(1+\alpha)\delta$-source. Here α is the constant from the Incidence Theorem, where arithmetic is over a suitable field F.*

We prove this using the Incidence Theorem. 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 lemma. This lemma is key in getting a statistical theorem from the few incidences.

Lemma 7.1. *Suppose X is ϵ-far from a k-source. Then $\exists S \subseteq \mathrm{supp}(X)$, $|S| < 2^k$, such that $\Pr[X \in S] \geq \epsilon$.*

Proof. Take $S = \{s | X(s) > 2^{-k}\}$. The correctness follows from Lemma 2.2. □

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

Proof of Theorem 6. 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 7.1. 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 $\Pr[AB + C \in S] \geq \epsilon$,

$$I(P, L) \geq \epsilon |L| \cdot |B| = \epsilon K^3.$$

On the other hand, since $|L| = K^2$ and $|P| \leq K \cdot K' = K^{2+\alpha}$, by the Incidence Theorem

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

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

7.3 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 can do this from two sources of equal length. We do this by giving a statistical version of a theorem by Bourgain. This theorem shows that the $AB + C$ theorem holds when $C = A^2$ and the entropy rate is measured with respect to the length of A, rather than (A, C), so only half the randomness is required.

Recently Bourgain [Bou05] showed that for a prime p, the function $g : \mathbb{F}_p \times \mathbb{F}_p \to \mathbb{F}_p$ given by $g(x, y) = x(x + y)$ has the following "expanding" property. For $|A| = |B| = p^\delta$, $\delta < 1$, $g(A, B) \geq p^{\delta+\beta}$ for some $\beta = \beta(\delta) > 0$.[1] We use the techniques developed in this section to show a statistical analogue

[1] Bourgain's proof also uses the Incidence Theorem, but it was done independently of our use of the Incidence Theorem in Subsections 7.2 and 7.4.

of this theorem. Note that our new 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 7. *For all $\delta \in (0, 1)$, there is a $\beta = \beta(\delta) > 0$ such that if X, Y are output from independent rate-δ-sources, then $g(X, Y)$ is exponentially close to a $\delta + \beta$-source.*

Proof. Let X and Y be independent random variables uniformly distributed over sets A and B of size p^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 later. Let S be the set of size at most $p^{\delta+\beta}$ guaranteed by Lemma 7.1. We can now follow Bourgain's proof closely (we change the beginning to take a probabilistic perspective).

Let X' and Y' be distributed as X, Y, respectively, but mutually independent. Then $\Pr[g(X, Y) = g(X', Y')] \geq \epsilon^2/|S|$. Since for $a \neq 0$, $|g(a, B)| = |B|$, we get that $\Pr[g(X, Y) \in g(X', B)] \geq \gamma$, where $\gamma = |B|\epsilon^2/|S|$. Let p_a denote $\Pr[g(X, Y) \in g(a, B)]$, so $\mathrm{E}_{a \in A}[p_a] \geq \gamma$. Hence $\Pr_{a \in A}[p_a \geq \gamma/2] \geq \gamma/2$.

Fix any a with $p_a \geq \gamma/2$. Let $q_{x,a}$ denote $\Pr[g(x, Y) \in g(a, B)]$. Since $\mathrm{E}_x[q_{x,a}] \geq \gamma/2$, $\Pr_{x \in A}[q_{x,a} \geq \gamma/4] \geq \gamma/4$. Let $A_1 = \{x \in A | q_{x,a} \geq \gamma/4\}$, and for $x \in A_1$ let $B_x = \{y | g(x, y) \in g(a, B)\}$.

We are now in the same situation as in Bourgain's proof, and the rest of our proof can follow his. □

7.4 Condensing with One Random Bit

We now return to the model of one weak source and a small random seed. Barak et al. [BKS+05] consider a condenser which adds two 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. Again, there is nothing special about the constant .9 below; any constant less than 1 will do.

Theorem 8. *Suppose $\delta \leq .9$. The map $(A, B, C) \mapsto ((A, C), (B, AB + C))$ is a rate-$(\delta \to (1 + \alpha/2)\delta, \epsilon)$-somewhere-condenser, where $\epsilon = 2^{-\Omega(\alpha k)}$. Here α is the constant from the Incidence Theorem.*

Note that the point $(b, ab + c)$ lies on the line $y = ax + c$. The above map can therefore be viewed as follows. Define the point-line incidence graph $G = (V, W, E)$ with vertices $V = F^2$ the set of points, and W the set of lines over F. A point p lies on line ℓ if $p \in \ell$. The above map is then the function from E to $V \times W$ which maps an edge (p, ℓ) to its two endpoints.

As in the previous subsections, we need a lemma to convert the statistical problem to a counting problem. This time the lemma involves a pair of distributions, instead of a single one.

Lemma 7.2. *If (X, Y) is ϵ-far from somewhere-entropy k, then there exists sets $S \subseteq \mathrm{supp}(X)$,$T \subseteq \mathrm{supp}(Y)$, $|S|, |T| < 2^k$, such that*

$$\Pr[X \in S \wedge Y \in T] \geq \epsilon.$$

Proof. Let $S = \{s | X(s) > 2^{-k}\}$ and $T = \{t | Y(t) > 2^{-k}\}$. If $\Pr[X \in S \wedge Y \in T] < \epsilon$, then this probability may be redistributed to ensure that all strings have probability at most 2^{-k}. This contradicts that (X, Y) is ϵ-far from somewhere-entropy k. □

We can now prove the theorem.

Proof of Theorem 8. We may assume that (A, B, C) is uniform on a set of size $K = 2^k$, and set $k' = (1 + \alpha/2)k$. Suppose $((A, C), (B, AB + C))$ is ϵ-far from min-entropy k'. Let $L = S$ and $P = T$ be the sets of size less than $K' = 2^{k'}$ given by Lemma 7.2. Then view L as a set of lines $(A, C)(x) = Ax + C$ and P as a set of points.

As in the previous subsections, we calculate the number of incidences in two different ways. On the one hand, since the line (a, c) is always incident to the point $(b, ab + c)$ and at least ϵ fraction of these pairs lie in $S \times T$, the number of incidences

$$I(P, L) \geq \epsilon K.$$

On the other hand, by the Incidence Theorem,

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

Combining these,

$$K' = \Omega((\epsilon K)^{1/(3/2-\alpha)}) = \Omega((\epsilon K)^{2/3+4\alpha/9}).$$

Thus we may take $\epsilon = K^{-\alpha/9}$ and the theorem is proved. □

Acknowledgements

We are grateful to Avi Wigderson for our joint work on Subsection 4.2, as well as for other useful discussions. I would also like to thank Ronen Shaltiel, Mike Saks, Salil Vadhan, Anup Rao, Madhu Sudan, and Jesse Kamp for helpful discussions and comments.

References

[AFWZ95] N. Alon, U. Feige, A. Wigderson, and D. Zuckerman. Derandomized graph products. *Computational Complexity*, 5:60–75, 1995.

[AKS87] M. Ajtai, J. Komlós, and E. Szemerédi. Deterministic simulation in Logspace. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 132–140, 1987.

[AKS04] M. Agrawal, N. Kayal, and N. Saxena. Primes is in P. *Annals of Mathematics*, 160:781–793, 2004.

[ALM+98] 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.

[BGS98] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs, and nonapproximability – towards tight results. *SIAM Journal on Computing*, 27(3):804–915, 1998.

[BH92] R. Boppana and M. Halldorsson. Approximating maximum independent sets by excluding subgraphs. *Bit*, 32:180–196, 1992.

[BIW04] B. Barak, R. Impagliazzo, and A. Wigderson. Extracting randomness using few independent sources. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 384–393, 2004.

[BKS+05] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson. Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 1–10, 2005.

[BKT04] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geometric and Functional Analysis*, 14:27–57, 2004.

[Bou05] J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1:1–32, 2005.

[BS94] M. Bellare and M. Sudan. Improved non-approximability results. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, pages 184–193, 1994.

[CDH+00] R. Canetti, Y. Dodis, S. Halevi, E. Kushilevitz, and A. Sahai. Exposure-resilient functions and all-or-nothing transforms. In Bart Preneel, editor, *Advances in Cryptology — EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 453–469. Springer-Verlag, May 2000.

[CG88] 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.

[Cra37] H. Cramer. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arithmetica*, pages 23–46, 1937.

[CW89] A. Cohen and A. Wigderson. Dispersers, deterministic amplification, and weak random sources. In *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, pages 14–19, 1989.

[DR05] Z. Dvir and R. Raz. An improved analysis of mergers. Technical Report TR05-025, Electronic Colloquium on Computational Complexity, 2005.

[DS02] Y. Dodis and J. Spencer. On the (non)universality of the one-time pad. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 376–385, 2002.

[EH03] L. Engebretsen and J. Holmerin. Towards optimal lower bounds for clique and chromatic number. *Theoretical Computer Science*, 299:537–584, 2003.

[Fei97] U. Feige. Randomized graph products, chromatic numbers, and the Lovasz θ function. *Combinatorica*, 17:79–90, 1997.

[FGL+96] 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.

[FK98] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57:187–199, 1998.

[FN93] A. Fiat and M. Naor. Implicit $O(1)$ probe search. *SIAM Journal on Computing*, 22:1–10, 1993.

[Gil98] D. Gillman. A chernoff bound for random walks on expander graphs. *SIAM Journal on Computing*, 27:1203–1220, 1998.

[Gol97] O. Goldreich. A sample of samplers – a computational perspective on sampling (survey). Technical Report TR97-020, Electronic Colloquium on Computational Complexity, 1997.

[Hal93] M. Halldorsson. A still better performance guarantee for approximate graph coloring. *Information Processing Letters*, 45:19–23, 1993.

[Hås99] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[HK01] J. Hastad and S. Khot. Query efficient PCPs with perfect completeness. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 610–619, 2001.

[IZ89] R. Impagliazzo and D. Zuckerman. How to recycle random bits. In *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, pages 248–253, 1989.

[Kah95] N. Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM*, 42:1091–1106, 1995.

[Kah97] N. Kahale. Large deviation bounds for Markov chains. *Combinatorics, Probability, and Computing*, 6:465–474, 1997.

[Kar72] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1972.

[Kho01] S. Khot. Improved inapproximability results for MaxClique, Chromatic Number and Approximate Graph Coloring. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 600–609, 2001.

[Kon03] S. Konyagin. A sum-product estimate in fields of prime order. Technical report, Arxiv, 2003. `http://arxiv.org/abs/math.NT/0304217`.

[Lov75] L. Lovasz. On the ratio of the optimal integral and fractional covers. *Discrete Mathematics*, 13:383–390, 1975.

[Lu02] C.-J. Lu. Hyper-encryption against space-bounded adversaries from on-line strong extractors. In *Advances in Cryptology — CRYPTO '02, 22nd Annual International Cryptology Conference, Proceedings*, 2002.

[LY99] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41:960–981, 1999.

[MU02] E. Mossel and C. Umans. On the complexity of approximating the VC dimension. *Journal of Computer and System Sciences*, 65:660–671, 2002.

[Nis96] N. Nisan. Extracting randomness: How and why – a survey. In *Proceedings of the 11th Annual IEEE Conference on Computational Complexity*, pages 44–58, 1996.

[NT99] N. Nisan and A. Ta-Shma. Extracting randomness: A survey and new constructions. *Journal of Computer and System Sciences*, 58:148–173, 1999.

[NZ96] N. Nisan and D. Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996.

[Pip87] N. Pippenger. Sorting and selecting in rounds. *SIAM Journal on Computing*, 16(6):1032–1038, December 1987.

[Raz05] R. Raz. Extractors with weak random seeds. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 11–20, 2005.

[RSW00] O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 22–31, 2000.

[Sha02] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the European Association for Theoretical Computer Science*, 77:67–95, June 2002.

[Sip88] M. Sipser. Expanders, randomness, or time vs. space. *Journal of Computer and System Sciences*, 36:379–383, 1988.

[ST00] A. Samorodnitsky and L. Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 191–199, 2000.

[TUZ01] A. Ta-Shma, C. Umans, and D. Zuckerman. Loss-less condensers, unbalanced expanders, and extractors. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 143–152, 2001.

[TZ04] A. Ta-Shma and D. Zuckerman. Extractor codes. *IEEE Transactions on Information Theory*, 50:3015–3025, 2004.

[TZS01] A. Ta-Shma, D. Zuckerman, and S. Safra. Extractors from Reed-Muller codes. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 638–647, 2001.

[Uma99] C. Umans. Hardness of approximating Σ_2^p minimization problems. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 465–474, 1999.

[Vad03] S. Vadhan. On constructing locally computable extractors and cryptosystems in the bounded storage model. In *Advances in Cryptology — CRYPTO '03, 23rd Annual International Cryptology Conference, Proceedings*, 2003.

[WX05] A. Wigderson and D. Xiao. A randomness-efficient sampler for matrix-valued functions and applications. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.

[WZ99] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999.

[Zuc90] D. Zuckerman. General weak random sources. In *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science*, pages 534–543, 1990.

[Zuc91] D. Zuckerman. *Computing Efficiently Using General Weak Random Sources*. PhD thesis, University of California at Berkeley, 1991.

[Zuc96] D. Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16:367–391, 1996.

[Zuc97] D. Zuckerman. Randomness-optimal oblivious sampling. *Random Structures and Algorithms*, 11:345–367, 1997.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'