SIAM J. DISCRETE MATH.
Vol. 13, No. 1, pp. 2–24

BOUNDS FOR DISPERSERS, EXTRACTORS, AND DEPTH-TWO SUPERCONCENTRATORS*

JAIKUMAR RADHAKRISHNAN† AND AMNON TA-SHMA‡

Abstract. We show that the size of the smallest depth-two N-superconcentrator is

$$\Theta(N \log^2 N / \log \log N).$$

Before this work, optimal bounds were known for all depths except two. For the upper bound, we build superconcentrators by putting together a small number of disperser graphs; these disperser graphs are obtained using a probabilistic argument. For obtaining lower bounds, we present two different methods. First, we show that superconcentrators contain several disjoint disperser graphs. When combined with the lower bound for disperser graphs of Kővari, Sós, and Turán, this gives an almost optimal lower bound of $\Omega(N(\log N/\log\log N)^2)$ on the size of N-superconcentrators. The second method, based on the work of Hansel, gives the optimal lower bound.

The method of Kővari, Sós, and Turán can be extended to give tight lower bounds for extractors, in terms of both the number of truly random bits needed to extract one additional bit and the unavoidable entropy loss in the system. If the input is an n-bit source with min-entropy k and the output is required to be within a distance of ϵ from uniform distribution, then to extract even one additional bit, one must invest at least $\log(n-k) + 2\log(1/\epsilon) - O(1)$ truly random bits; to obtain m output bits one must invest at least $m - k + 2\log(1/\epsilon) - O(1)$. Thus, there is a loss of $2\log(1/\epsilon)$ bits during the extraction. Interestingly, in the case of dispersers this loss in entropy is only about $\log\log(1/\epsilon)$.

Key words. dispersers, extractors, superconcentrators, entropy loss

AMS subject classifications. 94C15, 05C35

PII. S0895480197329508

1. Introduction.

Superconcentrators. An N*-superconcentrator* is a directed graph with N distinguished vertices called *inputs*, and N other distinguished vertices called *outputs*, such that for any $1 \leq k \leq N$, any set X of k inputs and any set Y of k outputs, there exist k vertex-disjoint paths from X to Y. The *size* of a superconcentrator G is the number of edges in it, and the *depth* of G is the number of edges in the longest path from an input to an output.

Superconcentrators were studied originally to show lower bounds in circuit complexity. Valiant [23] showed that there exist N-superconcentrators of size $O(N)$; Pippenger [16] showed that there exist N-superconcentrators of size $O(N)$ and depth $O(\log N)$. On the other hand, Pippenger [17] showed that every depth-two N-superconcentrators has size $\Omega(N \log^2 N)$. This raised the question of the exact tradeoff between depth and size, which attracted much research during the last two decades [17, 7, 19, 1]. Table 1.1 gives a summary of the results. Here $\lambda(d, N)$ is the inverse of

*Received by the editors November 3, 1997; accepted for publication (in revised form) July 6, 1999; published electronically January 13, 2000. A preliminary version of this paper appeared as *Tight bounds for depth-two superconcentrators*, in 38th Annual IEEE Symposium on Foundations of Computer Science, 1997, pp. 585–594.

http://www.siam.org/journals/sidma/13-1/32950.html

†Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India (jaikumar@tcs.tifr.res.in). This work of this author was done while visiting the Institute of Computer Science, Hebrew University, Jerusalem, Israel.

‡International Computer Science Institute, 1947 Center Street, Berkeley, CA 94704 (amnon@icsi.berkeley.edu). The work of this author was done while at the Department of Computer Science, Weizmann Institute of Science, Rehovot, Israel. The research of this author was supported by a Phil Zacharia postdoctoral fellowship.

Table 1.1

Depth	Size
2	$O(N \log^2 N)$ [17], $\Omega(N \log^{3/2} N)$ [1]
3	$\Theta(N \log\log N)$ [1]
$4, 5$	$\Theta(N \log^* N)$ [7, 19]
$2d$, $2d+1$	$\Theta(N\lambda(d, N))$ [7, 19]
$\Theta(\beta(N))$	$\Theta(N)$ [7]

functions in the Ackerman hierarchy: $\lambda(1, N)$ behaves like $\log N$, $\lambda(2, N)$ behaves like $\log^* N$. In general, $\lambda(d, N)$ decays very rapidly as d grows; β grows more slowly than the inverse of any primitive recursive function. We refer the reader to [7] for the definition of λ and β. Thus, the dependence of the size on the depth was well understood for all depths except two. In this paper, we close this gap.

Let size(N) denote the size of the smallest depth-two N-superconcentrator.

Theorem 1.1 (main result). Size$(N) \;=\; \Theta\left(N \cdot \frac{\log^2 N}{\log\log N}\right)$.

For the upper bound, we use the method of Wigderson and Zuckerman [24], who showed how superconcentrators can be constructed using a type of expander graphs called *disperser graphs*.

Definition 1.2 (disperser graphs [20, 6]). *A bipartite graph $G = (V_1 = [N], V_2 = [M], E)$ is a (K, ϵ)-disperser graph, if for every $X \subseteq V_1$ of cardinality K, $|\Gamma(X)| > (1-\epsilon)M$ (i.e., every large enough set in V_1* misses *less than an ϵ fraction of the vertices of V_2). The size of G is $|E(G)|$.*

Nisan and Wigderson suggested (see [14]) that it might be possible to choose better parameters in the construction given in [24]. We implement their suggestion to obtain superconcentrators by putting together a smaller number of disperser graphs. These disperser graphs are obtained by probabilistic arguments.

Remark. The best explicit construction known gives N-superconcentrators of size $O(N(\log N)^{\mathrm{poly}(\log\log n)})$ (see [22, 13]).

We also observe a connection in the opposite direction: every depth-two superconcentrator contains many disjoint disperser graphs. Thus, lower bounds for disperser graphs imply lower bounds for depth-two superconcentrators. Using this method, we derive a simple $\Omega(N \cdot (\log N / \log\log N)^2)$ lower bound for depth-two N-superconcentrators; this is only a factor of $\log\log N$ away from the upper bound. To obtain the optimal lower bound, we use a method based on the work of Hansel [9] (see also Katona and Szemerédi [11]).

Dispersers and extractors. Disperser graphs arise from disperser functions. For a random variable X taking values in $\{0,1\}^n$, the *min-entropy* of X is given by

$$H_\infty(X) = \min_{x \in \{0,1\}^n} \log(1/\Pr[X = x]).$$

Definition 1.3 (dispersers). *$F : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, ϵ)-disperser if for all random variables X taking values in $\{0,1\}^n$ with $H_\infty(X) \geq k$ and all $W \subseteq \{0,1\}^m$ of size at least $\epsilon 2^m$, we have*

$$\Pr[F(X, Z) \in W] \;>\; 0,$$

where Z is uniformly distributed over $\{0,1\}^d$.

With $F : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, we associate the bipartite graph $G_F = (V_1, V_2, E)$, where $V_1 = \{0,1\}^n$, $V_2 = \{0,1\}^m$, and there is one edge of the form (x, w)

for each z such that $f(x,z)=w$. In this graph, the degree of every vertex in V_1 is exactly 2^d.

It is then easy to verify the following.

PROPOSITION 1.4. *F is a (k,ϵ)-disperser iff G_F is a $(2^k,\epsilon)$-disperser graph.*

One special case of disperser graphs is the class of *highly-expanding graphs* [18], sometimes called *a-expanding graphs* [24]. These are bipartite graphs $G=(A=[N], B=[N], E)$, where for any two subsets $X \subseteq A$, $Y \subseteq B$ of size a, there is an edge between X and Y. This is clearly equivalent to saying that G is a $(K=a, \epsilon=\frac{a}{N})$-disperser graph. If G is an a-expanding graph, then $\bar{G}$ (the bipartite complement of G) has no subgraph isomorphic to $K_{a,a}$. Such graphs have been studied extensively, and the problem of determining the maximum possible number of edges in these graphs is known as *the Zarankiewicz problem* (see [2, pp. 309–326]). An elegant averaging argument, due to Kővari, Sós, and Turán, gives good upper bounds on the number of edges in these graphs. When applied to disperser graphs, this method gives the following lower bounds.

THEOREM 1.5 (lower bounds for disperser graphs). *Let $G=(V_1=[N], V_2=[M], E)$ be a (K,ϵ)-disperser. Denote by $\bar{D}$ the average degree of a vertex in V_1.*

(a) *Assume that $K<N$ and $\lceil \bar{D} \rceil \le \frac{(1-\epsilon)M}{2}$ (i.e., G is not trivial). If $\frac{1}{M} \le \epsilon \le \frac{1}{2}$, then $\bar{D}=\Omega(\frac{1}{\epsilon}\cdot \log\frac{N}{K})$, and if $\epsilon > \frac{1}{2}$, then $\bar{D}=\Omega(\frac{1}{\log(1/(1-\epsilon))}\cdot\log\frac{N}{K})$.*

(b) *Assume that $K \le \frac{N}{2}$ and $\bar{D} \le M/4$. Then, $\frac{\bar{D}K}{M}=\Omega(\log\frac{1}{\epsilon})$.*

Dispersers play an important role in reducing the error probability of algorithms that make *one-sided error.* In such applications, we typically have $\epsilon \le 1/2$. Also, a-expanding graphs fall in this category, because there ϵ tends to 0. Hence, the case $\epsilon \le 1/2$ is the one usually studied. However, for showing lower bounds for superconcentrators, we need to consider the case $\epsilon > 1/2$.

For reducing the error in algorithms that make *two-sided* error, one requires the function to satisfy stronger properties. Such functions are called extractors. For a survey of constructions and applications of dispersers and extractors, see the paper of Nisan [13].

For distributions D_1 and D_2 on $\{0,1\}^n$, the *variational distance* between D_1 and D_2 is given by

$$d(D_1,D_2) \;=\; \max_{S\subseteq\{0,1\}^n} |D_1(S)-D_2(S)|.$$

DEFINITION 1.6 (extractors). *$F:\{0,1\}^n\times\{0,1\}^d \to \{0,1\}^m$ is a (k,ϵ)-extractor, if for any distribution X on $\{0,1\}^n$ with $H_\infty(X)\ge k$, we have that $d(F(X,U_d),U_m)<\epsilon$, where U_d, U_m are random variables uniformly distributed over $\{0,1\}^d$ and $\{0,1\}^m$, respectively.*

In this view, an extractor uses d random bits to extract m quasi-random bits from a source with min-entropy k. Graphs arising from extractors have uniformity properties similar to random graphs.

DEFINITION 1.7 (extractor graphs). *A bipartite multigraph $G=(V_1=[N], V_2=[M], E)$ is a (K,ϵ)-extractor with (left) degree D, if every $x\in V_1$ has degree D and for every $X\subseteq V_1$ of size K, and any $W\subseteq V_2$,*

$$\left|\frac{|E(X,W)|}{|E(X,V_2)|}-\frac{|W|}{|V_2|}\right| < \epsilon.$$

Here, $E(V,W)$ is the set of edges between V and W in G.

We then have the following analogue of Proposition 1.4 (see Chor and Goldreich [4] and Zuckerman [25]).

PROPOSITION 1.8. *F is a (k, ϵ)-extractor iff G_F is a $(2^k, \epsilon)$-extractor graph.*

THEOREM 1.9 (lower bounds for extractors). *There is a constant $C > 0$ such that the following holds. Let $G = (V_1 = [N], V_2 = [M], E)$ be a (K, ϵ)-extractor with $K \leq \frac{N}{C}$. Then,*

(a) *if $\epsilon \leq \frac{1}{2}$ and $D \leq \frac{M}{2}$, then $D = \Omega(\frac{1}{\epsilon^2} \cdot \log(\frac{N}{K}))$;*

(b) *if $D \leq \frac{M}{4}$, then $\frac{DK}{M} = \Omega((\frac{1}{\epsilon})^2)$.*

In the terminology of functions this means that if $F : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k, ϵ)-extractor, then $d \geq \log(n-k) + 2\log(\frac{1}{\epsilon}) - O(1)$ and $d + k - m \geq 2\log(\frac{1}{\epsilon}) - O(1)$. These two bounds have the following interpretation.

(a) In order to extract one extra random bit (i.e., having $m \geq d+1$) we need to invest at least $d \geq \log(n-k) + 2\log(\frac{1}{\epsilon}) - O(1)$ truly random bits (and $d \geq \log(n-k) + \log(\frac{1}{\epsilon}) - O(1)$ for dispersers).

(b) There is an unavoidable entropy loss in the system. The input to the extractor has entropy at least $k + d$ (k in X and d in the truly random bits that we invest), while we get back only m quasi-random bits. Thus, there is a loss of $k + d - m \geq 2\log(\frac{1}{\epsilon}) - O(1)$ bits. In the case of dispersers we have $d + k - m \geq \log\log(\frac{1}{\epsilon}) - O(1)$.

Surprisingly, the entropy loss (which can be compared to the heat wasted in a physical process) has different magnitudes in dispersers (about $\log\log\frac{1}{\epsilon}$) and extractors (about $2\log\frac{1}{\epsilon}$). In [8, 21], explicit (k, ϵ)-extractors $F : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^n$, with $d = n - k + 2\log\frac{1}{\epsilon} + 2$ are constructed. Theorem 1.9 shows that the entropy loss of $2\log\frac{1}{\epsilon}$ in these extractors is unavoidable.

Theorems 1.5 and 1.9 improve the lower bounds shown by Nisan and Zuckerman [15]; they showed that $D \geq \max\{\log(\frac{N}{K}), \frac{1}{2\epsilon}\}$, and $\frac{DK}{M} \geq 1 - \epsilon$. Furthermore, our lower bounds match the upper bounds up to constant factors. Using standard probabilistic arguments [20, 26] one can show that our lower bounds are tight up to constant factors (for completeness we include the proofs in Appendix C).

THEOREM 1.10 (probabilistic constructions). *For every $1 < K \leq N$, $M > 0$ and $\epsilon > 0$ there exists a*

(a) *(K, ϵ)-disperser graph $G = (V_1 = [N], V_2 = [M], E)$ with degree $D = \lceil \frac{1}{\epsilon}(\ln(\frac{N}{K}) + 1) + \frac{M}{K}(\ln(\frac{1}{\epsilon}) + 1) \rceil$,*

(b) *(K, ϵ)-extractor graph $G = (V_1 = [N], V_2 = [M], E)$ with $D = \lceil \max\{\frac{1}{\epsilon^2}(\ln(\frac{N}{K}) + 1),\ \ln 2 \cdot \frac{M}{K} \cdot \frac{1}{\epsilon^2}\} \rceil$.*

1.1. Organization of the paper. In section 2, we first describe the lower bounds for dispersers. We describe the argument informally, leaving the formal proof for the appendix. Then we derive the lower bounds for extractors assuming a technical lemma on hypergeometric distributions. In section 3, we present the new upper and lower bounds for depth-two superconcentrators. The appendix has three parts. In the first we give the formal proof of the lower bounds for dispersers; in the second, we give the proof of the technical lemma used in section 2; in the third, we prove Theorem 1.10.

2. Bounds for dispersers and extractors. In this section we present the lower bounds for disperser and extractor graphs. In the rest of this section, we will drop the word "graphs," and refer to them as dispersers and extractors. As stated earlier, the lower bounds for dispersers claimed in Theorem 1.5 follows from the bounds obtained by Kővari, Sós, and Turán for the Zarankiewicz problem. Instead of quoting their result directly, we will present the complete proof based on their method. This will help clarify the proof of Theorem 1.9, where we use the same method to show lower bounds for extractors.

In the rest of this section, we will use the following notation. For a bipartite graph $G = (V_1, V_2, E)$, $D(G)$ will denote the maximum degree of a vertex in V_1 and $\bar{D}(G)$ will denote the average degree of a vertex in V_1.

2.1. Dispersers. We now describe the proof of Theorem 1.5. Suppose $G = (V_1 = [N], V_2 = [M], E)$ is a (K, ϵ)-disperser.

We first observe that part (b) follows from part (a). Let $G' = ([M], [N], E')$ be the graph obtained from G by interchanging the roles of V_1 and V_2. Then, G' is a $(\lceil \epsilon M \rceil, \frac{K}{N})$-disperser. Hence, by the first half of part (a) we have

$$\bar{D}(G') = \Omega\left(\frac{N}{K} \cdot \log\left(\frac{M}{\epsilon M}\right)\right).$$

Now $\bar{D}(G') = \bar{D}(G)N/M$; thus $\frac{\bar{D}(G)K}{M} = \Omega(\log(\frac{1}{\epsilon}))$, as claimed in part (b).

Now consider part (a). For a vertex v of G and a subset X of vertices of G we say that X *misses* v (and also v misses X) if $\Gamma(v) \cap X = \emptyset$. Now, we let B be a random subset of V_2 of size $L = \lceil \epsilon M \rceil$. For $v \in V_1$ with degree d_v, we have

$$\Pr[B \text{ misses } v] = \binom{M - d_v}{L}\binom{M}{L}^{-1}. \tag{2.1}$$

The expected number of vertices missed by B (that is, $\mathbf{E}[|V_1 \setminus \Gamma(B)|]$) is the sum of these probabilities. Since B can miss at most $K - 1$ vertices, we have

$$\sum_{v \in V_1} \binom{M - d_v}{L}\binom{M}{L}^{-1} \le K - 1.$$

Note that $f(u) = \binom{u}{t}$ is a convex function of u. By applying Jensen's inequality, $f(\mathbf{E}[X]) \le \mathbf{E}[f(X)]$, to the left-hand side above, we obtain

$$N\binom{M - \bar{D}}{L}\binom{M}{L}^{-1} \le K - 1. \tag{2.2}$$

Our lower bounds follow from this inequality. We will now informally sketch the main points of the derivation; the formal proof is in the appendix.

For $\epsilon \le 1/2$, the left-hand side of (2.2) is approximately $N\exp(-\epsilon\bar{D})$. Thus, we obtain the lower bound

$$\bar{D} \;\ge\; \frac{1}{\epsilon}\ln\frac{N}{K-1}.$$

This strengthens the previous lower bound $D \ge \max\{1/(2\epsilon), \log(N/K)\}$ (due to Nisan and Zuckerman [15]).

For $\epsilon > 1/2$, the left-hand side of (2.2) is approximated better by $N\left(\frac{1-\epsilon}{2}\right)^{\bar{D}}$, i.e.,

$$\bar{D} \;\ge\; \frac{\log(N/(K-1))}{\log(1/(1-\epsilon)) + 1}. \qquad \square$$

2.2. Extractors. Since a (K,ϵ)-extractor is also a (K,ϵ)-disperser, the lower bounds for dispersers apply to extractors as well. We will now improve these bounds by exploiting the stronger properties of extractors. As in the proof of Theorem 1.5 we will show that Theorem 1.9(a) implies Theorem 1.9(b). To that end we define "slice" extractors.

2.2.1. Slice-extractors.

DEFINITION 2.1. *Let $G = (V_1 = [N], V_2 = [M], E)$ be a bipartite graph. For $v \in V_1$ and $B \subseteq V_2$, let*

$$\mathrm{disc}(v,B) = \Pr_{w\in\Gamma(v)}[w \in B] - \frac{|B|}{M},$$

*where w is generated by picking a random edge leaving v. We say that v ϵ-*misses *B (and also B ϵ-*misses *v) when $|\mathrm{disc}(v,B)| \geq \epsilon$.*

DEFINITION 2.2 (slice-extractor). *G is a (K,ϵ,p)-slice-extractor, if every $B \subseteq V_2$ of size $\lceil pM\rceil$, ϵ-misses fewer than K vertices of V_1.*

A slice-extractor seems to be weaker than an extractor because it is required to handle only subsets of V_2 of one fixed size, whereas an extractor must handle sets of all sizes. It is simple to show (see, e.g., [26])

Claim 2.3. *If $G = (V_1 = [N], V_2 = [M], E)$ is (K,ϵ)-extractor, then G is also a $(2K,\epsilon,p)$-slice-extractor for all p.*

In fact, we have the following lemma.

LEMMA 2.4. *Suppose $\lceil qM\rceil \leq \lceil pM\rceil < M/2$ and $G = (V_1 = [N], V_2 = [M], E)$ is a (K,ϵ,p)-slice-extractor. Then G is also a $(2K,2\epsilon,q)$-slice-extractor.*

Proof. We will show that for any $A \subseteq V_1$ of size K and any $S \subseteq V_2$ of size $\lceil qM\rceil$,

$$\left|\Pr_{w\in\Gamma(A)}[w \in S] - \frac{\lceil qM\rceil}{M}\right| \leq 2\epsilon. \tag{2.3}$$

This implies that S can not 2ϵ-miss more than $2K$ vertices, and G is a $(2K,2\epsilon,q)$-slice-extractor.

We now show (2.3). Let $S_1, S_2 \subseteq V_2$ be subsets of size $\lceil qM\rceil$ and $T \subseteq V_2$ a subset of size $\lceil pM\rceil - qM$ disjoint from $S_1 \cup S_2$. Denote $T_1 = T \cup S_1$ and $T_2 = T \cup S_2$. Since G is a (K,ϵ,p) slice-extractor, $|\Pr_{w\in\Gamma(A)}[w \in T_i] - p| \leq \epsilon$, for $i = 1,2$. This implies that

$$\left|\Pr_{w\in\Gamma(A)}[w \in S_1] - \Pr_{w\in\Gamma(A)}[w \in S_2]\right| \leq 2\epsilon, \tag{2.4}$$

for every two subsets $S_1, S_2 \subseteq V_2$ of size $\lceil qM\rceil$. Now, pick $S \subseteq V_2$ of size $\lceil qM\rceil$ randomly and uniformly. Then,

$$\mathop{\mathbf{E}}_S[\Pr_{w\in\Gamma(A)}[w \in S]] = \frac{|S|}{M}.$$

Hence, exist sets S^+ and S^- such that

$$\Pr_{w\in\Gamma(A)}[w \in S^-] \leq \frac{\lceil qM\rceil}{M} \leq \Pr_{w\in\Gamma(A)}[w \in S^+].$$

This, when combined with (2.4), implies (2.3). ∎

2.2.2. Proof of Theorem 1.9(a).

LEMMA 2.5. *There exists a constant $C > 0$, such that if $G = (V_1 = [N], V_2 = [M], E)$ is a (K, ϵ, p)-slice-extractor with $D \leq M/2$, $K \leq N/C$, $p \leq 1/10$, $pM \geq 1$ and $\epsilon \leq p/25$, then*

$$\bar{D} \geq \frac{p}{C\epsilon^2} \ln \frac{N}{CK}.$$

Proof. We proceed as in the case of dispersers, by picking a random $\lceil pM \rceil$-sized subset $R \subseteq V_2$. To bound from below the probability that R ϵ-misses a vertex $v \in V_1$, we will use the following lemma, whose proof appears in the appendix.

LEMMA 2.6. *Let R be a random subset of $[M]$ of size qM, Γ be a nonempty subset of $[M]$ of size D, and $w : \Gamma \to [0,1]$ be a weight function such that $w(\Gamma) \stackrel{\text{def}}{=} \sum_{i \in \Gamma} w(i) = 1$. Suppose $\delta \leq 1/25$, $q \leq 1/4$ and $D \leq M/2$. Then,*

$$\Pr[|w(\Gamma \cap R) - q| \geq \delta q] \geq C^{-1} \exp(-C\delta^2 qD).$$

Here C is a constant independent of δ, q, D, and w.

Fix $v \in V_1$. Denote the set of v's neighbors by Γ. Since in our definition of extractors we allow multiple edges, $|\Gamma|$ can be smaller than d_v (the degree of v); so when we pick a random edge leaving v, some vertices in Γ might be more likely to be visited than others. Let us define a weight function $w : \Gamma \to [0,1]$ by letting $w(b)$ be the number of multiple edges between v and b divided by d_v. Then, by definition

$$v \ \epsilon\text{-misses } R \ \text{ iff } \ |w(R \cap \Gamma) - p| \geq \epsilon = (\epsilon/p)p.$$

We take $D = d_v$, $\lceil pM \rceil = qM$ (so $q \leq \frac{1}{4}$) and $\delta = \frac{\epsilon}{p} \leq \frac{1}{25}$. For C a large enough constant we have $\Pr[R \ \epsilon\text{-misses } v] \geq C^{-1} \exp(-C\delta^2 q d_v) \geq C^{-1} \exp(-C\epsilon^2 d_v / p)$. It follows that the expected number of vertices missed by R is at least

$$\sum_{v \in V_1} C^{-1} \exp\left(-C\frac{\epsilon^2}{p} d_v\right).$$

Since R never misses K vertices, we have

$$\sum_{v \in V_1} C^{-1} \exp\left(-C\frac{\epsilon^2}{p} d_v\right) \leq K - 1.$$

Since $\exp(-Cx)$ is a convex function of x, Jensen's inequality implies that

$$N\ C^{-1} \exp\left(-C\frac{\epsilon^2}{p} \bar{D}\right) \leq K - 1.$$

By taking logarithms we obtain

$$\bar{D} \geq \frac{p}{C\epsilon^2} \ln \frac{N}{CK}. \qquad ∎$$

We now show Theorem 1.9(a). Note that we may assume that $\epsilon \leq 10^{-3}$ (say); otherwise the claim follows from the lower bound for dispersers proved in Theorem 1.5(a). But, if $\epsilon \leq 10^{-3}$, our claim follows immediately from Lemma 2.5 by taking $p = 1/10$.

2.2.3. Proof of Theorem 1.9(b). We next show that, as in the proof of Theorem 1.5, part (b) follows from part (a) by reversing the roles of V_1 and V_2.

Claim 2.7. *Suppose $G = (V_1 = [N], V_2 = [M], E)$ is a (K, ϵ)-extractor. Let $G' = (V_1' = [M], V_2' = [N], E')$ be the graph obtained from G by reversing the roles of V_1 and V_2. Suppose $p \geq K/N$ and pN is an integer. Then, for all $T > 2$, G' is a $(\frac{4M}{T}, \epsilon' = T\epsilon p, p)$-slice-extractor.*

Proof. Suppose G' is not a $(\frac{4M}{T}, \epsilon' = T\epsilon p, p)$-slice-extractor. Then, there is some $B \subseteq V_2'$ of size pN that ϵ'-misses at least $\frac{4M}{T}$ vertices of V_1'. Let $A^- = \{v \in V_1' : \mathrm{disc}(v, B) \leq -\epsilon'\}$ and $A^+ = \{v \in V_1' : \mathrm{disc}(v, B) \geq \epsilon'\}$. One of these sets must have size at least $\frac{2M}{T}$, say A^-. Then

$$\begin{aligned} |E'(A^-, B)| &= \sum_{v \in A^-} |E'(v, B)| \\ &\leq \sum_{v \in A^-} d_v \left(\frac{|B|}{N} - \epsilon'\right) \\ &\leq |E'(A^-, V_2')| \cdot \left(\frac{|B|}{N} - \epsilon'\right). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{|E'(A^-, B)|}{|B|D} &\leq \frac{|E'(A^-, V_2')|}{|B|D} \left(\frac{|B|}{N} - \epsilon'\right) \\ &= \frac{|E(V_1, A^-)|}{ND} \left(1 - \frac{N\epsilon'}{|B|}\right). \end{aligned} \tag{2.5}$$

Since G is an extractor, we have

$$\frac{|E(V_1, A^-)|}{ND} \leq \frac{|A^-|}{M} + \epsilon. \tag{2.6}$$

We will now consider the sets $B \subseteq V_1$ and $A^- \subseteq V_2$ in the extractor G, and obtain a contradiction by showing that there are fewer edges between them than required. By combining (2.5) and (2.6), we obtain

$$\begin{aligned} \frac{|E(B, A^-)|}{|B|D} &\leq \left(\frac{|A^-|}{M} + \epsilon\right)\left(1 - \frac{N\epsilon'}{|B|}\right) \\ &\leq \frac{|A^-|}{M} + \epsilon - \frac{|A^-|}{M}\frac{N\epsilon'}{|B|} \\ &\leq \frac{|A^-|}{M} + \epsilon - 2\epsilon \\ &= \frac{|A^-|}{M} - \epsilon. \end{aligned}$$

(For the third inequality, we used $|A^-| \geq 2M/T$, $|B| = pN$ and $\epsilon' = T\epsilon p$.) But this contradicts our assumption that G is a (K, ϵ)-extractor. ∎

Finally, to obtain part (b) of Theorem 1.9, we choose $p = K/N$ and $T = 16C$ in the above claim and conclude that G' is a $(M/(4C), 16\epsilon CK/N, K/N)$-slice-extractor. Since $D(G) \leq M/4$, we have $\bar{D}(G') \leq N/4$. By Markov's inequality, at least half the vertices of $V_1(G')$ have degree at most $N/2$. By restricting ourselves to the vertices of lowest degree, we obtain an $(M/(4C), \epsilon' = 16\epsilon CK/N, p = K/N)$-slice-extractor G'' with $|V_1(G'')| \geq M/2$, $V_2(G'') = N$, $D(G'') \leq N/2$ and $\bar{D}(G'') \leq \bar{D}(G')$. If $\epsilon < 1/(400C)$, we have $\epsilon' \leq p/25$. Then, by Lemma 2.5, we have $\bar{D}(G'') \geq \frac{p}{C\epsilon'^2} \ln(\frac{M/2}{CM/(4C)})$; i.e.,

$$\bar{D}(G'') = \Omega\left(\frac{p}{\epsilon^2 p^2}\right) = \Omega\left(\frac{N}{K}\frac{1}{\epsilon^2}\right).$$

Since $\bar{D}(G'') \leq N\bar{D}(G)/M$, we get $\frac{K\bar{D}(G)}{M} = \Omega(\frac{1}{\epsilon^2})$.

3. Superconcentrators of depth two. In this section, we show bounds on the size of depth-two superconcentrators. First, we show an $O(N \log^2 N / \log\log N)$ upper bound using the upper bounds for dispersers from Theorem 1.10 (a). Next, we show that the lower bounds on dispersers, shown in Theorem 1.5 (a), imply an (almost tight) $\Omega(N(\log N/\log\log N)^2)$ lower bound for superconcentrators. Finally, by using a different method, we improve this lower bound to $\Omega(N \log^2 N/\log\log N)$, matching the upper bound (up to constant factors).

Recall that size(N) is the size of the smallest depth-two N-superconcentrator. It is enough to establish the claimed bounds assuming that N is a power of two. For, let $2^n \leq N < 2^{n+1}$, where $n \geq 1$; then size$(2^n) \leq$ size$(N) \leq$ size(2^{n+1}).

3.1. The upper bound.

THEOREM 3.1. *Size$(N) = O(N \log^2 N/\log\log N)$.*

Proof. Our construction is based on a similar construction due to Wigderson and Zuckerman [24]. We build a depth-two graph $(A = [N], C, B = [N], E)$, where C is the disjoint union of C_i, $i = 0, \ldots, \lceil \log_{\log N} N \rceil - 1$, where $|C_i| = 2\log^{i+1} N$. For every i put a $(K = \log^i N, \epsilon = 1/4)$-disperser $D_i = (A, C_i, E_i)$, and another (K, ϵ)-disperser between B and C_i.

For $\log^i N \leq K \leq \log^{i+1} N$, for every K-set $X \subseteq A$, $\Gamma(X)$ covers at least 3/4 of C_i. Similarly, for every K-set $Y \subseteq B$, $\Gamma(Y)$ covers at least 3/4 of C_i. So, at least half of the vertices of C_i are common neighbors of X and Y. We thus have the following claim.

Claim 3.2. *Any two sets $X \subseteq A$, $Y \subseteq B$ of size K have at least K common neighbors in C.*

However, as shown by Meshulam [12], Menger's theorem implies that this is sufficient for G to be a superconcentrator (clearly, it is necessary). All that remains is to count the number of edges in G. By Theorem 1.10 (a), we may take $|E_i| = O(N \log N)$. Thus, we have $|E(G)| = \sum_i 2|E_i| = O(N \log^2 N/\log\log N)$. □

Remark. This construction differs from the one in [24] in only one respect: in their construction each C_i takes care of all K-sets for $2^i \leq K < 2^{i+1}$, whereas in ours each C_i takes care of all K-sets for $\log^i N \leq K < \log^{i+1} N$. The point is that constructing a disperser that works for just one K, in itself, requires average degree $\log N$ (as can be seen from Theorem 1.5). Furthermore, we can build dispersers that recover almost all the random bits we invest (see Theorem 1.10(a)). This enables us to hash sets of size K into sets of size $K \log N$ and use fewer C_i's.

3.2. Lower bound via dispersers. Now we show that any depth-two superconcentrator must contain $\Omega(\log N/\log\log N)$ disjoint disperser graphs and derive from this an $\Omega(N(\log N/\log\log N)^2)$ lower bound. The idea is as follows. Consider any depth-two superconcentrator $G = (A = [N], C, B = [N], E)$. By definition, for any $1 \leq K \leq N$, and any two subsets $X \subseteq A$, $Y \subseteq B$ of cardinality K, X, and Y

have at least K common neighbors in C. In particular, if we fix a subset $X \subseteq A$ of cardinality K and look at $\Gamma(X)$, we see that every K-subset of B must have at least K neighbors in $\Gamma(X)$. In other words, the induced graph on $\Gamma(X)$ and B is a disperser. By doing this for different K's we get several disjoint dispersers. Our lower bound then follows by applying the disperser lower bound to each of them.

THEOREM 3.3. $\mathrm{Size}(N) = \Omega(N \cdot (\log N/\log\log N)^2)$.

Proof. Let $G = (A = [N], C, B = [N], E)$ be a depth-two N-superconcentrator. We will proceed in stages. In stage i, we will consider subsets of A and B of size $K_i = \log^{3i} N$. If $K_i \le \sqrt{N}$ (i.e., $i \le (1/6)\log N/\log\log N$), then we will show that there is a subset C_i of the middle layer C such that the number of edges between B (the output vertices) and C_i is at least $N\log N/\log\log N$. The sets C_i will be *disjoint* for different values of i. Collecting the edges from the different C_i's, we have

$$\begin{aligned}|E(G)| &\ge \left\lfloor \frac{\log N}{6\log\log N}\right\rfloor \cdot \Omega\left(N\frac{\log N}{\log\log N}\right)\\ &= \Omega\left(N\left(\frac{\log N}{\log\log N}\right)^2\right).\end{aligned}$$

Suppose the average degree in A is $\bar{D}$. If $\bar{D} \ge \log^2 N$, then the number of edges between A and C is $N\bar{D} \ge N\log^2 N$ and we are done. So assume $\bar{D} \le \log^2 N$.

Let $X_i \subseteq A$ be the set of $K_i = \log^{3i} N$ vertices with smallest degrees (breaking ties using some order on the vertices). Let $Z_i = \Gamma(X_i)$. Clearly $|Z_i| \le K_i\bar{D}$. Let $C_i = Z_i \setminus (Z_1 \bigcup Z_2 \bigcup \cdots \bigcup Z_{i-1})$. Since $X_i \subseteq X_{i+1}$ for all i, we also have $Z_i = \Gamma(X_i) \subseteq \Gamma(X_{i+1}) = Z_{i+1}$; thus, $C_i = Z_i \setminus Z_{i-1}$.

Claim 3.4. *G restricted to B and C_i is a $(K_i,\ \epsilon = 1 - \frac{1}{2\bar{D}})$-disperser.*

Proof of claim. Any two sets $X \subseteq A$, $Y \subseteq B$ of cardinality K_i must have K_i common neighbors in C. In particular, any set $Y \subseteq B$ of size K_i has K_i distinct neighbors in $Z_i = \Gamma(X_i)$, and therefore at least $K_i - |Z_{i-1}|$ distinct neighbors in C_i.

Notice that $K_i - |Z_{i-1}| \ge K_i - K_{i-1}\bar{D} > K_i/2$. Thus, in G restricted to B and C_i, any subset in B of size K_i has more than $K_i/2$ distinct neighbors. Thus, the claim follows if $K_i/2 \ge (1-\epsilon)|C_i|$. Indeed

$$(1-\epsilon)|C_i| = \frac{|C_i|}{2\bar{D}} \le \frac{K_i\bar{D}}{2\bar{D}} = \frac{K_i}{2},$$

where the inequality follows from $|C_i| \le |Z_i| \le K_i\bar{D}$.

By Theorem 1.5(a), the number of edges between A and C_i is

$$\Omega\left(\frac{N\cdot\log\left(\frac{N}{K_i}\right)}{\log\left(\frac{1}{(1-\epsilon)}\right)}\right).$$

As long as $K_i \le \sqrt{N}$, this is at least $\Omega(N\log N/\log\log N)$. Since the C_i's are disjoint, we have obtained $\Omega(\log N/\log\log N)$ disjoint dispersers, each having $\Omega(N\log N/\log\log N)$ edges. ∎

3.3. The improved lower bound. If we look at the construction of Theorem 3.1, we see that sets of cardinality K_i *communicate* mainly through a specific subset of C denoted C_i. Furthermore, the vertices of C_i can be identified using their degree: vertices in C_i have degree about N/K_i. In our proof we will find this structure in the superconcentrator.

THEOREM 3.5. $\mathrm{Size}(N) = \Omega(N \log^2 N / \log\log N)$.

Proof. Let $G = (A = [N], C, B = [N], E)$ be a depth-two N-superconcentrator. We assume that N is large. As in the proof of Theorem 3.3, we proceed in stages. In stage i ($i = 1, 2, \ldots$), we consider sets of size $K = K_i = \log^{4i} N$. Let

$$C_i = \left\{ w \in C : \frac{N}{K} \frac{1}{\log^2 N} \ \le\ \deg(w) \ <\ \frac{N}{K} \log^2 N \right\};$$
$$D_i = \left\{ w \in C : \deg(w) \ <\ \frac{N}{K} \frac{1}{\log^2 N} \right\}.$$

We will show that if $K \le N^{3/4}$ (i.e., $i \le \frac{3}{16}(\log N / \log\log N)$), then there are at least $\frac{1}{10} N \log N$ edges incident on C_i. Since the sets C_i are disjoint for different values of i, we have

$$\begin{aligned} |E(G)| &\ge \left\lfloor \frac{3}{16}(\log N / \log\log N) \right\rfloor \cdot \frac{1}{10} N \log N \\ &= \Omega(N \log^2 N / \log\log N). \end{aligned}$$

It remains to show that the number of edges incident on C_i is at least $\frac{1}{10} N \log N$. We assume that

$$|E(G)| \le \frac{1}{10} N \log^2 N; \tag{3.1}$$

otherwise the theorem follows immediately.

LEMMA A. *For every pair of K-sets $X \subseteq A$ and $Y \subseteq B$, there is a common neighbor in $C_i \cup D_i$.*

Proof of lemma. All vertices outside $C_i \cup D_i$ have degree at least $(N/K)\log^2 N$. Then, assumption (3.1) implies that the number of vertices outside $C_i \cup D_i$ is at most

$$\frac{|E(G)|}{(N/K)\log^2 N} \ \le\ \frac{K}{10}.$$

Since X and Y have at least K common neighbors in the original graph, they must in fact have at least $\frac{9}{10}K$ common neighbors in $C_i \cup D_i$.

We wish to show that G restricted to $A \cup B$ and C_i cannot be sparse. We know from Lemma A that every pair of K-sets in $A \cup B$ has a common neighbor in $C_i \cup D_i$. Suppose that G restricted to $A \cup B$ and C_i is sparse. We will first obtain sets $S \subseteq A$ and $T \subseteq B$ such that S and T have no common neighbors in C_i. Then, all pairs of K-sets in S and T have to communicate via D_i; in other words, the bipartite graph induced on S and T by the connections via D_i is a K-expanding graph. Since the number of edges incident on D_i is small (because of (3.1)), D_i cannot provide enough connections for such a K-expanding graph, leading to a contradiction. We thus have three tasks ahead of us.

- First, we need to show how to obtain sets S and T. For this we use a method based on the work of Hansel [9]. We go through all vertices in C_i, and for each, either delete all its neighbors in A or all its neighbors in B. Clearly, after this the surviving vertices in A and B do not have any common neighbor in C_i. It is remarkable that even after this severe destruction, we expect large subsets of vertices $S \subseteq A$ and $T \subseteq B$ to survive.

- Second, we need to show that the number of connections required between S and T is large. This follows from the fact that the bipartite graph induced on S and T by the connections via D_i is a K-expanding graph.
- Finally, we need to show that the low degree vertices in D_i cannot provide the required number of connections between S and T. This will follow from the definition of D_i and the fact that S and T are sufficiently random subsets of $A \cup B$.

LEMMA B. *If $K = K_i \leq N^{3/4}$, then there are more than $\frac{1}{10} N \log N$ edges incident on C_i.*

Proof of lemma. For $u \in A \cup B$, let d_u be the number of neighbors of u in C_i. Let A' be the set of $\frac{N}{2}$ vertices $u \in A$ with smallest d_u, and B' be the set of $\frac{N}{2}$ vertices $v \in B$ with smallest d_v. We will prove that there is some $u \in A' \cup B'$ with $d_u > \frac{1}{5} \log N$. This implies the lemma for, say, $u \in A'$. Then, for all $u \in A \setminus A'$, $d_u > \frac{1}{5} \log N$, and we have more than $\frac{1}{10} N \log N$ edges incident on C_i, as required.

Now we have to prove that there is some $u \in A' \cup B'$ with large d_u. Otherwise, for all $u \in A' \cup B'$, $d_u \leq \frac{1}{5} \log N$. We will show that this contradicts (3.1).

For each $w \in C_i$, perform the following action (independently for each w):

– with probability $\frac{1}{2}$, delete all neighbors of w from A';

– with probability $\frac{1}{2}$, delete all neighbors of w from B'.

Set $d = \frac{1}{5} \log N$. For each vertex u in $A' \cup B'$ that survives, delete it independently with probability $1 - 2^{-(d - d_u)}$.

It is clear that after the above process, the probability that a vertex in $A' \cup B'$ survives is exactly 2^{-d}. Let S be the subset of vertices of A' that survive and T the subset of vertices of B' that survive. Our construction ensures that S and T do not have a common neighbor in C_i. Lemma A then implies that every pair of K-sets in S and T has a common neighbor in D_i.

Consider the bipartite graph $H = (S, T, E)$, where E consists of pairs $(u, v) \in S \times T$ such that u and v have a common neighbor in D_i. Then H is a K-expanding graph (i.e., there is an edge joining every pair of K-sets in S and T). It follows (see Lemma 3.8 below) that

$$|E(H)| \geq \frac{|S| \cdot |T|}{K} - |S| - |T|.$$

Thus, if S and T are large, the required number of edges in H is also large. It is not hard to see that the expected size of S and T is large ($\frac{N 2^{-d}}{2} \sim N^{4/5} \gg K$). But we need S and T to be large *simultaneously.* Instead of ensuring this, it will be easier to directly estimate the average number of edges needed by H.

Claim 3.6.

$$\mathbf{E}[|E(H)|] \;\geq\; \mathbf{E}\left[\frac{|S| \cdot |T|}{K} - |S| - |T|\right] \;>\; \frac{N^2 2^{-2d}}{10K}.$$

This gives a lower bound on the average number of connections required between S and T. Conversely, our next claim shows that if the number of edges in G is small, then the average number of edges in H (which is the number of connections between S and T passing via D_i) is small.

Claim 3.7.

$$\mathbf{E}[|E(H)|] \;\leq\; |E(G)| \cdot \frac{N}{K \log^2 N} \cdot 2^{-2d}.$$

Before proceeding to the proofs of these claims, let us complete the proof of Lemma B. Putting the two claims together we obtain

$$|E(G)| \cdot \frac{N}{K \log^2 N} \cdot 2^{-2d} > \frac{N^2 2^{-2d}}{10K},$$
$$\text{i.e.,} \quad |E(G)| > \frac{N}{10} \log^2 N.$$

But then G has too many edges, contradicting (3.1).

Proof of Claim 3.6. We have $|S| = \sum_{u \in A'} X_u$ and $|T| = \sum_{v \in B'} Y_v$, where X_u and Y_v are the 0-1 indicator variables for the events "$u \in S$" and "$v \in T$," respectively.

For $u \in A'$ and $v \in B'$, X_u and Y_v are independent whenever u and v don't have a common neighbor in C_i, and $\mathbf{E}[X_u Y_v] = 2^{-2d}$. Since $d_u \leq \frac{1}{5} \log N$ and vertices in C_i have degree at most $\frac{N}{K} \log^2 N$, each X_u is independent of all but $\frac{N}{K} \log^2 N \cdot \frac{\log N}{5} < \frac{N (\log N)^3}{5K} < \frac{N}{4}$ of the Y_v's. Therefore,

$$\begin{aligned}
\mathbf{E}[|S| \cdot |T|] &= \sum_{u \in A', v \in B'} \mathbf{E}[X_u Y_v] \\
&\geq |A'| \cdot \frac{N}{4} \cdot 2^{-2d} \\
&= \frac{1}{8} N^2 2^{-2d}.
\end{aligned}$$

Clearly, $\mathbf{E}[|S|] = \sum_{u \in A'} E[X_u] = \frac{N}{2} 2^{-d}$ and similarly $\mathbf{E}[|T|] = \frac{N}{2} 2^{-d}$. Thus we have

$$\mathbf{E}\left[\frac{|S| \cdot |T|}{K} - |S| - |T|\right] \geq \frac{N^2 2^{-2d}}{8K} - N 2^{-d} = \frac{N^2 2^{-2d}}{K}\left(\frac{1}{8} - \frac{2^d K}{N}\right) > \frac{N^2 2^{-2d}}{10K},$$

where the last inequality holds because $2^d K \leq N^{1/5} N^{3/4} = o(N)$ and N is large.

Proof of Claim 3.7. Consider all pairs $(u, v) \in A' \times B'$, such that u and v have a common neighbor in D_i. Since the degree of a vertex in D_i is at most $(N/K) \log^{-2} N$, the number of such pairs is at most

$$\begin{aligned}
\sum_{w \in D_i} \deg(w)^2 &\leq \frac{N}{K \log^2 N} \sum_{w \in D_i} \deg(w) \\
&\leq \frac{N}{K \log^2 N} |E(G)|. \qquad (3.2)
\end{aligned}$$

As argued in the proof of Claim 3.6, for every pair $(u, v) \in A' \times B'$, if u and v don't have a common neighbor in C_i, then $\Pr[(u, v) \in S \times T] = 2^{-2d}$; conversely, if u and v have a common neighbor in C_i, then one of them will be deleted, and $\Pr[(u, v) \in S \times T] = 0$. Thus, in both cases $\Pr[(u, v) \in S \times T] \leq 2^{-2d}$. Our claim follows from this and (3.2) by linearity of expectation. □

Finally, we prove the density bound for K-expanding graphs.

Claim 3.8. *If* $(V_1 = [N_1], V_2 = [N_2], E)$ *is a* K*-expanding graph, then* $|E| \geq \frac{N_1 N_2}{K} - N_1 - N_2$

Proof. If either N_1 or N_2 is less than K, then the claim is trivial. Otherwise, obtain $\lfloor \frac{N_1}{K} \rfloor$ disjoint sets of size K from V_1. No set of size K can miss more than K vertices in V_2; in particular, each such set has $N_2 - K$ edges incident on it. Collecting the contributions from the $\lfloor \frac{N_1}{K} \rfloor$ sets, we get at least $(\frac{N_1}{K}-1)(N_2-K) > \frac{N_1 N_2}{K} - N_1 - N_2$ edges. ∎

Appendix A. Lower bounds for dispersers.

We now complete the proof of Theorem 1.5. We will use inequality (2.2) derived in section 2.1.

First, consider the case $\epsilon \leq \frac{1}{2}$. To simplify the left-hand side of (2.2), we will use the inequality $\binom{a-c}{b}\binom{a}{b}^{-1} \geq (\frac{a-b-c+1}{a-b+1})^b$, valid whenever $a-b-c+1 \geq 0$. In our application, we have $\bar{D} \leq (1-\epsilon)M/2 \leq (1-\epsilon)M$ (an assumption in Theorem 1.5) and $L = \lceil \epsilon M \rceil \leq \epsilon M + 1$; thus, $M - \bar{D} - L + 1 \geq M - (1-\epsilon)M - \epsilon M - 1 + 1 = 0$. Then, (2.2) gives

$$N\left(\frac{M - L - \bar{D} + 1}{M - L + 1}\right)^L \leq K - 1;$$

i.e.,

$$\begin{aligned}\frac{N}{K-1} &\leq \left(\frac{M-L+1}{M-L-\bar{D}+1}\right)^L \\ &= \left(1 + \frac{\bar{D}}{M-L-\bar{D}+1}\right)^L \\ &\leq \exp(\bar{D}L/(M-\bar{D}-L+1)).\end{aligned}$$

On taking lns and solving for $\bar{D}$, we obtain

$$\bar{D} \geq \frac{(M-L+1)\ln(N/(K-1))}{L + \ln(N/(K-1))}.$$

If $\ln(N/(K-1)) > L$, then

$$\bar{D} > \frac{M-L+1}{2} \geq \frac{(1-\epsilon)M}{2},$$

contradicting our assumption. Thus, we may assume that $\ln(N/(K-1)) \leq L$. Then,

$$\begin{aligned}\bar{D} &\geq \frac{M-L+1}{2L}\ln\frac{N}{K-1} \\ &\geq \frac{(1-\epsilon)M}{2\epsilon M+2}\ln\frac{N}{K-1} \\ &\geq \frac{1}{8\epsilon}\ln\frac{N}{K-1}.\end{aligned}$$

For the case $\epsilon > \frac{1}{2}$, we must approximate the left-hand side of (2.2) differently. Since $\binom{a}{b}$ is a nondecreasing function of a, we have from (2.2) that

$$N\binom{M - \lceil \bar{D} \rceil}{L}\binom{M}{L}^{-1} \leq K - 1.$$

Since $\binom{a-b}{c}\binom{a}{c}^{-1} = \binom{a-c}{b}\binom{a}{b}^{-1}$, we have

$$N\binom{M-L}{\lceil \bar{D} \rceil}\binom{M}{\lceil \bar{D} \rceil}^{-1} \leq K - 1.$$

Now we use the inequality $\binom{a-b}{c}\binom{a}{c}^{-1} \geq (\frac{a-b-c+1}{a})^c$ and obtain

$$\begin{aligned} K-1 &\geq N\left(\frac{M-\lceil\bar{D}\rceil-L+1}{M}\right)^{\lceil\bar{D}\rceil} \\ &\geq N\left(\frac{M-L+1}{2M}\right)^{\lceil\bar{D}\rceil} \\ &\geq N\left(\frac{1-\epsilon}{2}\right)^{\lceil\bar{D}\rceil}. \end{aligned}$$

Thus,

$$\lceil\bar{D}\rceil \geq \frac{\log(N/(K-1))}{\log(2/(1-\epsilon))}. \qquad \square$$

Appendix B. Lower bounds on deviation.

This section is devoted to the proof of Lemma 2.6. We reproduce the lemma below for easy reference. Note that in the version below we use ϵ instead of δ and p instead of q.

Lemma 2.6. *Let R be a random subset of $[M]$ of size pM, Γ be a nonempty subset of $[M]$ of size D, and $w : \Gamma \to [0,1]$ be a weight function such that $w(\Gamma) \stackrel{\text{def}}{=} \sum_{i\in\Gamma} w(i) = 1$. Suppose $\epsilon \leq 1/25$, $p \leq 1/4$, and $D \leq M/2$. Then,*

$$\Pr[|w(\Gamma \cap R) - p| \geq \epsilon p] \geq C^{-1}\exp(-C\epsilon^2 pD).$$

Here C is a constant independent of ϵ, p, D, and w.

B.1. Overview of the proof. We have two cases based on the value of p.

Case 1 (*small* p). We first assume that $pD \leq 12$. In this case, we show that with constant probability $\Gamma \cap R = \emptyset$.

Lemma B.1. $\Pr[\Gamma \cap R = \emptyset] \geq \exp(-50)$.

Case 2 (*large* p). We now assume that $pD > 12$. In this case, the proof has two main parts.

Part 1. The expected value of $|\Gamma \cap R|$ is easily seen to be pD. We first show lower bounds on the probability that $|\Gamma \cap R|$ deviates from this expected value by at least ϵpD.

Lemma B.2. *If $pD \geq 12$, then for some constant C_0 (independent of p, D, M, and ϵ)*

(a) $\Pr[|\Gamma \cap R| \leq (1-\epsilon)pD] \;\geq\; C_0^{-1}\exp(-C_0\epsilon^2 pD)$,
(b) $\Pr[|\Gamma \cap R| \geq (1+\epsilon)pD] \;\geq\; C_0^{-1}\exp(-C_0\epsilon^2 pD)$.

Part 2. Note that Part 1 suffices when the weights are all equal. Next, we consider a general distribution of weights. We show that when $|\Gamma \cap R|$ differs significantly from its expected value, then $w(\Gamma \cap R)$ is also likely to differ from its expected value. Note that the expected value of $w(\Gamma \cap R)$ is p.

Lemma B.3. *Let R^+ be a random subset of Γ of size $\lceil pD\rceil$ and R^- be a random subset of size $\lfloor(1-4\epsilon)pD\rfloor$. Then, at least one of the following two statement holds:*

(a) $\Pr[w(R^-) \leq p(1-\epsilon)] \geq 1/25$,
(b) $\Pr[w(R^+) \geq p(1+\epsilon)] \geq 1/4$.

We first complete the proof of Lemma 2.6 assuming that Lemmas B.1, B.2, and B.3 hold. We shall justify Lemmas B.1, B.2, and B.3 after that.

Proof of Lemma 2.6. If $pD \leq 12$, then with constant probability, $\Gamma \cap R$ is empty. Clearly, whenever $\Gamma \cap R$ is empty, $|w(\Gamma \cap R) - p| \geq \epsilon p$. The claim follows from this.

Next, assume that $pD \geq 12$. We now use Lemma B.3 and conclude that at least one of the two statements, (a) and (b), of the lemma holds. Suppose (a) holds. Then, for all sizes $k \leq (1-4\epsilon)pD$, we have

$$\Pr[w(R_k) \leq p(1-\epsilon)] \;\geq\; \frac{1}{25}, \tag{B.1}$$

where R_k is a random k-sized subset of Γ. Let $\mathcal{E}$ denote the event $|\Gamma \cap R| \leq (1-4\epsilon)pD$.

$$\begin{aligned}
\Pr[|w(\Gamma \cap R) - p| \geq \epsilon p] &\geq \Pr[\mathcal{E}] \cdot \Pr[|w(\Gamma \cap R) - p| \geq \epsilon p \mid \mathcal{E}] \\
&\geq \Pr[\mathcal{E}] \cdot \Pr[w(\Gamma \cap R) \leq p(1-\epsilon) \mid \mathcal{E}] \\
&= \Pr[\mathcal{E}] \cdot \sum_{k=0}^{(1-4\epsilon)pD} \Pr[|\Gamma \cap R| = k \mid \mathcal{E}] \\
&\quad \times \Pr[w(\Gamma \cap R) \leq p(1-\epsilon) \mid |\Gamma \cap R| = k].
\end{aligned}$$

By Lemma B.2 (a), we have $\Pr[\mathcal{E}] \geq C_0^{-1} \exp(-C_0 \epsilon^2 pD)$, for some constant C_0. Also, under the condition $|\Gamma \cap R| = k$, the random set $\Gamma \cap R$ has the same distribution as R_k. Then, using (B.1), we have

$$\begin{aligned}
\Pr[|w(\Gamma \cap R) - p| \geq \epsilon p] \geq \Pr[\mathcal{E}] \cdot \sum_{k=0}^{(1-4\epsilon)pD} &\Pr[|\Gamma \cap R| = k \mid \mathcal{E}] \cdot \Pr[w(R_k) \leq p(1-\epsilon)] \\
\geq C_0^{-1} \exp(-C_0 \epsilon^2 pD) &\cdot \frac{1}{25} \\
\geq C^{-1} \exp(-C \epsilon^2 pD)&
\end{aligned}$$

for $C = 25\, C_0$.

In the remaining case, statement (b) of Lemma B.3 holds, and the claim follows by a similar argument, this time using Lemma B.2 (b). ∎

B.2. Proofs. *Proof of Lemma* B.1. We have

$$\Pr[\Gamma \cap R| = \emptyset] \;=\; \binom{M-D}{pM}\binom{M}{pM}^{-1} \;\geq\; \left(\frac{M-D-pM}{M-pM}\right)^{pM} \;=\; \left(1 - \frac{D}{M-pM}\right)^{pM}.$$

Now, we use the inequality $1 - x \geq \exp(-x/(1-x))$, valid whenever $x < 1$. Thus,

$$\Pr[\Gamma \cap R| = \emptyset] \;\geq\; \exp\left(-\frac{DpM}{M-D-pM}\right) \;\geq\; \exp(-4pD) \;\geq\; \exp(-48).$$

For the second inequality, we use $D \leq M/2$ and $p \leq 1/4$, and for the last inequality, we use $pD \leq 12$. ∎

Part 1.

Proof of Lemma B.2. We will present the detailed argument only for part (a); the argument for part (b) is similar.

$$\Pr[|\Gamma \cap R| \geq (1-\epsilon)pD] = \sum_{k \leq (1-\epsilon)pD} \Pr[|\Gamma \cap R| = k].$$

We will be interested only in the last approximately $\sqrt{pqD}$ terms. Let

$$A = \lfloor (1-\epsilon)pD \rfloor \quad \text{and} \quad B = \left\lceil A - \sqrt{pqD} \right\rceil + 1.$$

Then,

$$\Pr[|\Gamma \cap R| \le (1-\epsilon)pD] \ge \sum_{k=A}^{B} \Pr[|\Gamma \cap R| = k].$$

It can be verified that $\Pr[|\Gamma \cap R| = k]$ is an increasing function of k, for $A \le k \le B$. Thus,

$$\Pr[|\Gamma \cap R| \le (1-\epsilon)pD] \ge (A - B + 1) \Pr[|\Gamma \cap R| = B]. \tag{B.2}$$

We will now estimate the two factors on the right-hand side.

- First, $A - B + 1 \ge \sqrt{pqD} - 2$, and since $pD \ge 12$ and $q \ge \frac{3}{4}$, we have $A - B + 1 \ge \frac{1}{3}\sqrt{pqD}$.
- To estimate the second term, we write $B = pd - h$; then we have
$$\epsilon pD + \sqrt{pqD} - 2 \ \le \ h \ \le \ \epsilon pD + \sqrt{pqD}.$$
Since $pD \ge 12$ and $q \ge \frac{3}{4}$, we have $1 \le h < pD < qD$. Claim B.4 below shows that
$$\Pr[|\Gamma \cap R| = B] \ge \frac{1}{\sqrt{4\pi pqD}} \exp\left(-\frac{4h^2}{pqD}\right).$$

Substituting these two estimates in (B.2), we obtain

$$\Pr[|\Gamma \cap R| \ \le \ (1-\epsilon)pD] \ge \frac{1}{6\sqrt{\pi}} \exp\left(-\frac{4h^2}{pqD}\right).$$

To finish the proof of the lemma we consider two cases.

- If $\epsilon pD \le \sqrt{pqD}$, then $h \le 2\sqrt{pqD}$, and the bound above gives
$$\Pr[|\Gamma \cap R| \le (1-\epsilon)pD] \ \ge \ \frac{1}{6\sqrt{\pi}} \exp\left(-\frac{4 \cdot 4pqD}{pqD}\right) \ = \ \frac{1}{6\sqrt{\pi}} \exp(-16).$$
- Conversely, if $\epsilon pD \ge \sqrt{pqD}$, then $h \le 2\epsilon pD$, and
$$\Pr[|\Gamma \cap R| \le (1-\epsilon)pD] \ \ge \ \frac{1}{6\sqrt{\pi}} \exp\left(-\frac{4 \cdot 4(\epsilon pD)^2}{pqD}\right) \ = \ \frac{1}{6\sqrt{\pi}} \exp(-32\epsilon^2 pD).$$

Now we return to the claim.

Claim B.4. *If* $1 \le h < pD$, *then* $\Pr[|\Gamma \cap R| = pD - h] \ge \frac{1}{\sqrt{4\pi pqD}} \exp(-\frac{4h^2}{pqD})$.

Proof of claim. We have

$$\begin{aligned}
\Pr[|\Gamma \cap R| = pD - h] &= \frac{\binom{D}{pD-h}\binom{n-D}{p(n-D)+h}}{\binom{n}{pn}} \\
&\ge \sqrt{2\pi pqn}\ 2^{-nH(p)} \\
&\quad \times \frac{1}{\sqrt{2\pi pqD}}\ 2^{DH(p)}\ \exp\left(-\frac{2h^2}{pqD}\right) \exp\left(-\frac{1}{6}\right) \\
&\quad \times \frac{1}{\sqrt{2\pi pq(n-D)}}\ 2^{(n-D)H(p)}\ \exp\left(-\frac{2h^2}{pq(n-D)}\right) \exp\left(-\frac{1}{6}\right) \\
&\ge \frac{1}{\sqrt{4\pi pqD}} \exp\left(-\frac{4h^2}{pqD}\right).
\end{aligned}$$

For the last inequality we used the assumption $n - D \geq D$. For the first inequality we used the following bounds on binomial coefficients:

- For the denominator, we used the bound (see [3, p. 4])

$$\binom{n}{pn} \leq \frac{2^{nH(p)}}{\sqrt{2\pi pqn}}.$$

- For the numerator, we used

$$\binom{n}{pn+h}, \binom{n}{pn-h} \;>\; \frac{2^{nH(p)}}{\sqrt{2\pi pqn}} \exp\left(-\frac{2h^2}{pqn}\right) \exp\left(-\frac{1}{6}\right),$$

valid for $1 \leq h < \min\{pn, qn\}$. To justify this, we start from (see [3, p. 12])

$$\binom{n}{pn+h} > \frac{2^{nH(p)}}{\sqrt{2\pi pqn}} \exp\left(-\frac{h^2}{pqn}(\frac{1}{2} + \frac{hp}{2qn} + \frac{h^2q}{3p^2n^2} + \frac{q}{n}) - \frac{1}{6}\right).$$

Since $h < qn$, we have $\frac{hp}{qn} < p \leq 1$; since $h \leq pn$, we have $h^2q \leq h^2 \leq p^2n^2$; since $\min\{pn, qn\} \geq 1$, we have $n \geq 2$ and $\frac{q}{n} \leq \frac{1}{n} \leq \frac{1}{2}$. Thus, we get the required bound

$$\binom{n}{pn+h} > \frac{2^{nH(p)}}{\sqrt{2\pi pqn}} \exp\left(-\frac{2h^2}{pqn}\right) \exp\left(-\frac{1}{6}\right).$$

The bound on $\binom{n}{pn-h}$ follows from the above inequality because

$$\binom{n}{pn-h} = \binom{n}{n-pn+h} = \binom{n}{qn+h} > \frac{2^{nH(p)}}{\sqrt{2\pi pqn}} \exp\left(-\frac{2h^2}{pqn}\right) \exp\left(-\frac{1}{6}\right). \qquad \square$$

Part 2. We now present the proof of Lemma B.3. We shall first consider the case $p = 1/2$. The general case will follow from this.

LEMMA B.5. *Let S be a random subset of Γ of size $\ell = (\frac{1}{2} - 2\delta)D$. Assume $\delta \leq 1/12$. Then,*

$$\Pr\left[w(S) \leq \frac{1}{2} - \delta\right] \;\geq\; \frac{1}{12}.$$

Before we proceed to the proof of this lemma, let us deduce Lemma B.3 from it.

Proof of Lemma B.3. Let X be a random subset of Γ of size $2\lceil pD\rceil$. Let

$$\rho \;=\; \Pr[w(X) < 2p(1+\epsilon)].$$

We consider two cases based on the value of ρ:

1. $\rho \leq 1/2$. Let X be as above. Let Y be a random subset of X of size $\lfloor(\frac{1}{2} - 4\epsilon)|X|\rfloor$. Thus, by Lemma B.5, for each value of X

$$\Pr\left[w(Y) \leq \left(\frac{1}{2} - 2\epsilon\right) w(X)\right] \geq \frac{1}{12}.$$

Since $\rho \geq 1/2$, with probability $1/24$, we have $w(Y) < (\frac{1}{2} - 2\epsilon) \cdot 2p(1+\epsilon) < p(1-\epsilon)$. This implies that

$$\Pr[w(R^-) < p(1-\epsilon)] \;\geq\; \frac{1}{24}.$$

2. $\rho < 1/2$. In this case, let Y be a random subset of X of size $\lceil pD \rceil$. Whenever $w(X) \geq 2p(1+\epsilon)$, at least one of Y and $X \setminus Y$ has weight of at least $p(1+\epsilon)$. Thus

$$\Pr[w(Y) > p(1+\epsilon) \mid w(X) \geq 2p(1+\epsilon)] \;\geq\; \frac{1}{2}.$$

Since $\rho < 1/2$, we have that $\Pr[w(X) \geq 2p(1+\epsilon)] > 1/2$. Thus

$$\Pr[w(Y) \geq p(1+\epsilon)] \;\geq\; \frac{1}{4}.$$

Now Y is a random subset of Γ of size $\lceil pD \rceil$; therefore,

$$\Pr[w(R^+) \geq p(1+\epsilon)] \;\geq\; \Pr[w(Y) \geq p(1+\epsilon)]. \qquad \square$$

Proof of Lemma B.5. Let $k = D - 2\ell$, Since $\delta \leq 1/12$, we have that

$$\ell = \left(\frac{1}{2} - 2\delta\right) D \;\geq\; \frac{1}{3}D \quad \text{and} \quad k = 4\delta D \;\leq\; \frac{1}{3}D.$$

Thus, $\ell \geq k$, and we may write $\ell = mk + k'$, where m and k are integers such that $1 \leq m \leq \ell/k$ and $0 \leq k' < k$. We now describe a procedure for generating sets of size ℓ.

Step 1. Let $\pi = \{E_1, E_2, B_1, B_2, \ldots, B_{2m+1}\}$ be a partition of Γ such that $|E_1|, |E_2| = k'$, and for $i = 1, 2, \ldots, 2m+1$, $|B_i| = k$. E_1 and E_2 will be referred to as the *exceptional* blocks.

Step 2. Pick a random permutation σ of $[2m+1]$ and arrange the blocks in the order

$$\langle E_1, B_{\sigma(1)}, B_{\sigma(2)}, \ldots, B_{\sigma(2m+1)}, E_2 \rangle.$$

Step 3. Let

$$\begin{aligned}
\mathsf{Prefix}(\pi, \sigma) &= E_1 \cup B_{\sigma(1)} \cup B_{\sigma(2)} \cup \cdots \cup B_{\sigma(m)};\\
\mathsf{Middle}(\pi, \sigma) &= B_{\sigma(m+1)}; \text{ and}\\
\mathsf{Suffix}(\pi, \sigma) &= B_{\sigma(m+2)} \cup B_{\sigma(m+3)} \cup \cdots \cup B_{\sigma(2m+1)} \cup E_2.
\end{aligned}$$

If π and σ are chosen randomly, then $\mathsf{Prefix}(\pi, \sigma)$ and $\mathsf{Suffix}(\pi, \sigma)$ are random sets of size ℓ (i.e., they have the same distribution as the random set S in the statement of the lemma).

We will prove the lemma by contradiction. Suppose the claim of the lemma is false.

$$\Pr_{\pi,\sigma}\left[w(\mathsf{Prefix}(\pi,\sigma)) > \frac{1}{2} - \delta\right] > \frac{11}{12} \quad \textit{and} \quad \Pr_{\pi,\sigma}\left[w(\mathsf{Suffix}(\pi,\sigma)) > \frac{1}{2} - \delta\right] > \frac{11}{12}.$$

It follows that with probability more than $5/6$, both Prefix and Suffix are *heavy* and consequently $w(\mathsf{Middle}(\pi, \sigma)) < 2\delta$. Let $\mathcal{E}(\pi, \sigma)$ denote this event; then,

$$\Pr_{\pi,\sigma}[\mathcal{E}(\pi, \sigma)] \;>\; \frac{5}{6}. \tag{B.3}$$

$\mathsf{Middle}(\pi, \sigma)$ is a random subset of Γ of size k, and $E_1(\pi)$ and $E_2(\pi)$ are random subsets of Γ of size k'. Since $k' < k$, we may conclude that

$$\Pr_{\pi}[w(E_1(\pi)) < 2\delta] > \frac{5}{6} \qquad \text{and} \qquad \Pr_{\pi}[w(E_2(\pi)) < 2\delta] > \frac{5}{6}.$$

It follows that with probability 2/3, both E_1 and E_2 are light. Let $\mathcal{F}(\pi)$ denote this event; then,

$$\Pr_{\pi}[\mathcal{F}(\pi)] \ > \ \frac{2}{3}. \tag{B.4}$$

Let B^- be the block in π with the smallest weight; let its weight be w^-. Let B^+ be the block in $\hat{\pi}$ with the largest weight; let its weight be w^+. For a partition π and an ordering σ, let σ' be the ordering derived from σ by interchanging the positions of B^+ an B^-. Clearly, if σ is chosen uniformly from the set of all permutations of $[2m+1]$, then σ' is a random ordering with the same distribution as σ. It then follows from (B.3) that

$$\Pr_{\pi,\sigma}[\mathcal{E}(\pi, \sigma')] \ > \ \frac{5}{6}. \tag{B.5}$$

Now, from (B.3), (B.4), and (B.5) we have

$$\Pr_{\pi,\sigma}[\mathcal{E}(\pi, \sigma) \wedge \mathcal{F}(\pi) \wedge \mathcal{E}(\pi, \sigma')] > \frac{1}{3}. \tag{B.6}$$

The probability that $B^- \neq \mathsf{Middle}(\pi, \sigma)$ and B^+ does not appear on the same side of Middle as B^- is

$$\frac{2m}{2m+1} \cdot \frac{m+1}{2m} = \frac{m}{2m+1} \geq \frac{2}{3},$$

where the inequality holds since $m \geq 1$. By combining this with (B.6), we conclude that with nonzero probability the following events take place simultaneously.

(a) $\mathsf{Prefix}(\pi, \sigma)$, $\mathsf{Suffix}(\pi, \sigma)$, $\mathsf{Prefix}(\pi, \sigma')$, and $\mathsf{Suffix}(\pi, \sigma')$ are all heavy;

(b) $E_1(\pi)$ and $E_2(\pi)$ are both light (i.e., have weight less than 2δ); and

(c) $B^- \neq \mathsf{Middle}(\pi, \sigma)$, and B^- and B^+ do not appear on the same side of Middle.

We will show that this is impossible. Suppose (say) $B^- \in \mathsf{Prefix}(\pi, \sigma)$. Then, from the definition of σ' and (a), we have

$$w(\mathsf{Prefix}(\pi, \sigma')) \ = \ w(\mathsf{Prefix}(\pi, \sigma)) + w^+ - w^- \ > \ \frac{1}{2} - \delta + w^+ - w^-.$$

Since $\mathsf{Prefix}(\pi, \sigma')$ and $\mathsf{Suffix}(\pi, \sigma')$ are heavy, we have

$$\begin{aligned} 1 &= w(\mathsf{Prefix}(\pi, \sigma')) + w(\mathsf{Middle}(\pi, \sigma')) + w(\mathsf{Suffix}(\pi, \sigma')) \\ &> \left(\frac{1}{2} - \delta + w^+ - w^-\right) + w^- + \left(\frac{1}{2} - \delta\right) \\ &= 1 + w^+ - 2\delta. \end{aligned}$$

This is impossible, since, as we now show, $w^+ \geq 2\delta$ whenever (a) and (b) hold. For, by (a) $w(\mathsf{Prefix}(\pi, \sigma)) > 1/2 - \delta$, and by (b) $w(E_1(\pi)) < 2\delta$. Hence, one of the blocks $B_1, B_2 \ldots, B_m$, has weight more than

$$\frac{1}{m}\left(\frac{1}{2} - 3\delta\right) \ \geq \ \frac{k}{\ell}\left(\frac{1}{2} - 3\delta\right) \ \geq \ \frac{4\delta}{1/2 - \delta}\left(\frac{1}{2} - 3\delta\right) \ = \ 4\delta\,\frac{1 - 6\delta}{1 - 2\delta}.$$

Since $\delta \le 1/12$, this is at least 2δ. ∎

Appendix C. Existence of dispersers and extractors. In this section we prove Theorem 1.10.

Dispersers. First, consider the part (a) of Theorem 1.10. Sipser [20] showed the existence of disperser graphs with parameters ($N = m^{\log m}, M = m, K = m, D = 2\log^2 m, \epsilon = 1/2$); we use his argument and obtain disperser graphs with parameters close to the lower bounds shown in section 2.

We construct a random graph $G = (V_1 = [N], V_2 = [M], E)$ by choosing D random neighbors for each $v \in V_1$. Fix $\epsilon > 0$ and let $L = \lceil \epsilon M \rceil$. If G is not a (K,ϵ)-disperser, there is some subset of V_2 of size L that misses some K vertices of V_1. Thus, $\Pr[G$ is not a (K,ϵ)-disperser$]$ is at most

$$\begin{aligned} &\binom{N}{K}\binom{M}{L}(1-L/M)^{KD} \\ &< (eN/K)^K (eM/L)^L (1-L/M)^{KD} \\ &\le (eN/K)^K (eM/L)^L \exp(-LKD/M)). \qquad \text{(C.1)} \end{aligned}$$

(To justify the last inequality we use $1 - x \le e^{-x}$.)

Plugging D we have $(eN/K)^K \cdot (eM/L)^K \le \exp(LKD/M)$. Therefore, by (C.1), we have

$$\Pr[G \text{ is not a disperser}] < 1.$$

So there is at least one instance of G that meets our requirements. ∎

Extractors. We now consider part (b) of Theorem 1.10. Essentially the same bounds were derived by Zuckerman [26]. We derive it again for completeness. We will use Definition 1.7. We will obtain a bipartite graph $G = (V_1 = [N], V_2 = [M], E)$ where all vertices in V_1 have the same degree D such that for every $S \subseteq V_1$ of cardinality K, and any $R \subseteq V_2$,

$$\left| \frac{|E(S,R)|}{KD} - \frac{|R|}{M} \right| < \epsilon. \qquad \text{(C.2)}$$

We first observe that (C.2) can be replaced by a seemingly weaker condition.

Claim C.1. *If for every $S \subseteq V_1$ of size K, and any $R \subseteq V_2$, $|E(S,R)| < KD(\frac{|R|}{M} + \epsilon)$, then G is a (K,ϵ)-extractor.*

For, if there exist some S and R with $|E(S,R)| \le KD(\frac{|R|}{M-\epsilon})$, then consider the set $\bar{R} = V_2 \setminus R$. We have $|E(S,\bar{R})| > KD(\frac{|\bar{R}|}{M} + \epsilon)$, contradicting the hypothesis of the claim.

Now we prove the existence of extractors. Consider the random graph $G = (V_1 = [N], V_2 = [M], E)$ obtained by choosing D random neighbors with replacement for each $v \in V_1$.

Fix $S \subseteq V_1$ of size K and $R \subseteq V_2$. Let $p = \frac{|R|}{M}$. We wish to estimate the probability (over the choices of the edges) that $|E(S,R)| \ge KD(p+\epsilon)$. The number of edges between S and R is the sum of KD identically distributed independents random variables $X_1, X_2, \ldots, X_{KD}$, each taking the value 1 with probability $p = \frac{|R|}{M}$ and the value 0 with probability $1-p$. Thus, we can bound the probability of deviation using standard estimates for the binomial distribution:

$$\Pr[|E(S,R)| \ge (p+\epsilon)KD] \le \exp(-2\epsilon^2 KD).$$

(This version of Chernoff's bounds appears in Chvátal [5] and Hoeffding [10].) Thus, $\Pr[G$ is not a (K,ϵ)-extractor] is at most

$$\begin{aligned}
&\binom{N}{K} 2^M \exp(-2\epsilon^2 KD) \\
&< \left(\frac{eN}{K}\right)^K 2^M \exp(-2\epsilon^2 KD) \\
&= (e^{K(1+\ln(N/K))} \cdot e^{-\epsilon^2 KD}) \cdot (e^{M \ln 2} \cdot e^{-\epsilon^2 KD}).
\end{aligned}$$

Since $D \geq \frac{1}{\epsilon^2}(1 + \ln(N/K))$ the first factor is at most 1; similarly, since $D \geq \frac{M \ln 2}{\epsilon^2 K}$ the second factor is at most 1. It follows that

$$\Pr[G \text{ is not an extractor}] < 1.$$

Hence, there is an instance of G that satisfies our requirements. $\square$

Acknowledgments. We thank Roy Armoni, Oded Goldreich (who first asked us about entropy loss), Nati Linial, Avner Magen, Noam Nisan, Avi Wigderson, and Shiyu Zhou for many helpful discussions. We thank Aravind Srinivasan for helping us with the tail estimates. We are also grateful to the anonymous referee for many helpful comments. We thank David Zuckerman for Lemma 2.4 and its proof.

REFERENCES

[1] N. Alon and P. Pudlák, *Superconcentrators of depth* 2 *and* 3*; odd levels help (rarely)*, J. Comput. System Sci., 48 (1994), pp. 194–202.
[2] B. Bollobas, *Extremal Graph Theory*, Academic Press, London, 1978.
[3] B. Bollobas, *Random Graphs*, Academic Press, New York, 1985.
[4] B. Chor and O. Goldreich, *Unbiased bits from sources of weak randomness and probabilistic communication complexity*, SIAM J. Comput., 17 (1988), pp. 230–261.
[5] V. Chvátal, *The tail of the hypergeometric distribution*, Discrete Math., 25 (1979), pp. 285–287.
[6] A. Cohen and A. Wigderson, *Dispersers, deterministic amplification and weak random sources*, in Proceedings IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamos, CA, 1989, pp. 14–19.
[7] D. Dolev, C. Dwork, N. Pippenger, and A. Wigderson, *Superconcentrators, generalizers and generalized connectors with limited depth*, in Proceedings ACM Symposium on Theory of Computing (STOC), ACM, New York, 1983, pp. 42–51.
[8] O. Goldreich and A. Wigderson, *Tiny families of functions with random properties: A quality-size trade-off for hashing*, Random Structures Algorithms, 11 (1997), pp. 315–343.
[9] G. Hansel, *Nombre minimal de contacts de fermature nécessaires pour réaliser une fonction booléenne symétrique de n variables*, C. R. Acad. Sci. Paris, 258 (1964), pp. 6037–6040.
[10] W. Hoeffding, *Probability inequalities for sums of bounded random variables,* J. Amer. Statist. Assoc., 58 (1963), pp. 13–30.
[11] G. Katona and E. Szemerédi, *On a problem of graph theory*, Studia Sci. Math. Hungar., 2 (1967), pp. 23–28.
[12] R. Meshulam, *A geometric construction of a superconcentrator of depth* 2, Theoret. Comput. Sci., 32 (1984), pp. 215–219.
[13] N. Nisan, *Refining randomness: How and why*, in Proceedings IEEE Symposium on Computational Complexity, IEEE Computer Society Press, Los Alamitos, CA, 1996, pp 44–58.
[14] N. Nisan and A. Wigderson, *Research pearls in theory of computation: Homework* 2*, problem* 2*;* also available online from http://www.cs.huji.ac.il/course/pearls/.
[15] N. Nisan and D. Zuckerman, *Randomness is linear in space*, J. Comput. System Sci., 52 (1996), pp 43–52.
[16] N. Pippenger, *Superconcentrators*, SIAM J. Comput., 6 (1977), pp. 298–304.
[17] N. Pippenger, *Superconcentrators of depth* 2, J. Comput. System Sci., 24 (1982), pp 82–90.
[18] N. Pippenger, *Sorting and selecting in rounds*, SIAM J. Comput., 16 (1987), pp. 1032–1038.

[19] P. Pudlák, *Communication in bounded depth circuits*, Combinatorica, 14 (1994), pp. 203–216.
[20] M. Sipser, *Expanders, randomness, or time versus space*, J. Comput. System Sci., 36 (1988), pp. 379–383.
[21] A. Srinivasan and D. Zuckerman, *Computing with very weak random sources*, in Proceedings IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 1994, pp. 264–275.
[22] A. Ta-Shma, *Refining Randomness*, Ph.D. thesis, The Hebrew University, Jerusalem, 1996.
[23] L. Valiant, *Graph-theoretic properties in computational complexity*, J. Comput. System Sci., 13 (1976), pp. 278–285.
[24] A. Wigderson and D. Zuckerman, *Expanders that beat the eigenvalue bound: Explicit construction and applications*, in Proceedings ACM Symposium on the Theory of Computing, ACM, New York, 1993, pp. 245–251.
[25] D. Zuckerman, *Simulating BPP using a general weak random source*, Algorithmica, 16 (1996), pp. 367–391.
[26] D. Zuckerman, *Randomness-optimal oblivious sampling*, Random Structures Algorithms, 11 (1997), pp. 345–367.