Tight bounds for depth-two superconcentrators

Jaikumar Radhakrishnan

Tata Institute, Mumbai, India
and
Hebrew University, Jerusalem, Israel

Amnon Ta-Shma

Deptartment of Computer Science
Weizmann Institute of Science
Rehovot, Israel

Abstract

We show that the minimum size of a 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. We present two different methods for showing lower bounds. First, we show that superconcentrators contain several disjoint disperser graphs. When combined with the lower bound for disperser graphs due to 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 the Kővari, Sós and Turán can be extended to give tight lower bounds for extractors, both in terms of the number of truly random bits needed to extract one additional bit and in terms of 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 a constant number of additional bits, 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 $\log\log(1/\epsilon)$.

1 Introduction

An *N-superconcentrator* is a directed graph with N distinguished vertices called *inputs*, and N other distinguished vertices called *outputs*, such that for any $1 \le k \le N$, 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; recently, Spielman [Spi96] discovered a connection between superconcentrators and the design of error-correcting codes. Valiant [Val76] showed that there exist N-superconcentrators of size $O(N)$; Pippenger [Pip77] showed that there exist N-superconcentrators of size $O(N)$ and depth $O(\log N)$. On the other hand, Pippenger [Pip82] showed that there are no linear size depth-two superconcentrators. This raised the question of the exact tradeoff between depth and size, which attracted much research during the last two decades [Pip82, DDPW83, Pud94, AP94]. The following table gives a summary of the results.

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

Here $\lambda(d, N)$ is the inverse of 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 [DDPW83] 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 (Main result)

$$\text{size}(N) \;=\; \Theta(N \cdot \frac{\log^2 N}{\log\log N}).$$

For the upper bound, we use the method of Wigderson and Zuckerman [WZ93], who showed how superconcentrators can be constructed using a type of expander graphs called *dispersers*.

Definition 1.1 (Sipser [Sip88])
A bipartite graph $G = (V_1 = [N], V_2 = [M], E)$ is a (K, ϵ)-disperser, 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 the disperser G is $|E(G)|$.*

Following a hint given by Nisan and Wigderson in their course, *Research Pearls in Theory of Computation* [NW], we refine the construction in [WZ93] and get the upper bound by putting together a small number of dispersers. These dispersers are obtained by probabilistic arguments; the best explicit construction known uses dispersers constructed by Ta-Shma [T96a], and gives N-superconcentrators of size $O(N(\log N)^{\text{poly}\log\log n})$ (see [T96b, Nis96]).

We also observe a connection in the opposite direction: every depth-two superconcentrator contains many disjoint dispersers. Thus, lower bounds for dispersers 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 [Han64] (see also Katona and Szemerédi [KS67]).

Dispersers and extractors. One special case of disperser graphs is the class of *highly-expanding graphs* [Pip87] , sometimes called *a-expanding graphs* [WZ93]. 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. 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 [Bol78, pages 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 dispersers, this method gives the following lower bounds.

Theorem 2 (Lower bounds for dispersers) *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 $\lceil \overline{D} \rceil \leq \frac{(1-\epsilon)M}{2}$ (i.e. G is not trivial). If $\frac{1}{M} \leq \epsilon \leq \frac{1}{2}$ 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 $K \leq \frac{N}{2}$ and $\overline{D} \leq M/4$. Then, $\frac{\bar{D}K}{M} = \Omega(\log \frac{1}{\epsilon})$.

Although the case $\epsilon \leq 1/2$ is the one usually studied, (and also the one more similar to a-expanding graphs, where ϵ tends to 0), for the superconcentrator lower bound we need the case where $\epsilon > 1/2$.

We further show that this method gives good lower bounds for extractors.

Definition 1.2 (Extractors) *$E : \{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^m$ is a (k, ϵ)-extractor, if for any distribution X on $\{0,1\}^n$ with $H_\infty(X) \geq k$, we have that $d(E(X, U_d), U_m) < \epsilon$, where $d(\cdot,\cdot)$ is the variational distance, U_l is the uniform distribution over $\{0,1\}^l$, and $H_\infty(X) = \min_{y \in \{0,1\}^n} \log(\frac{1}{X(y)})$.*

In this view an extractor uses d random bits to extract m quasi-random bits from a source with min-entropy k. Extractors correspond to bipartite graphs with properties similar to random graphs.

Definition 1.3 (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 $Y \subseteq V_2$,*

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

Clearly, every (K, ϵ)-extractor is a (K, ϵ)-disperser. Dispersers and extractors play important roles in the deterministic simulation of randomized algorithms; for a survey see the paper of Nisan [Nis96].

Theorem 3 (Lower bounds for extractors) *Let $G = (V_1 = [N], V_2 = [M], E)$ be a (K, ϵ)-extractor with degree $D \leq \frac{M}{4}$ and $K \leq \frac{N}{C}$, where C is a constant.*

(a) If $\epsilon \leq \frac{1}{2}$ then $D = \Omega((\frac{1}{\epsilon})^2 \cdot \log(\frac{N}{K}))$.

(b) $\frac{DK}{M} = \Omega((\frac{1}{\epsilon})^2)$.

In the function terminology this means that if $E : \{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^m$ 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 two extra random bits (i.e. having $m \geq d + 2$) we need to invest at least $d \geq \log(n-k) + \log(\frac{1}{\epsilon}) - O(1)$ 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 $2\log\frac{1}{\epsilon}$). Goldreich and Wigderson [GW94, section 6] show the existence of explicit (k,ϵ)-extractors $E : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^n$, with $d = n - k + 2\log\frac{1}{\epsilon} + 2$. Theorem 3 shows that the entropy loss of $2\log\frac{1}{\epsilon}$ in their extractor unavoidable, and the result of Wigderson and Zuckerman [WZ93] is tight up to a constant additive term.

Theorems 2 and 3 improve the lower bounds shown by Nisan and Zuckerman [NZ93]; they showed that $D \geq \max\{\log(\frac{N}{K}), \frac{1}{\epsilon}\}$, and $\frac{DK}{M} \geq 1 - \epsilon$. Furthermore, our lower bounds match the upper bounds (obtained using well-known probabilistic arguments) up to constant factors.

Theorem 4 (Probabilistic constructions) *For every $1 < K \leq N$, $M > 0$ and $\epsilon > 0$ there exists a*

(a) (K,ϵ)-disperser $G = (V_1 = [N], V_2 = [M], E)$ with degree $D = 2\cdot\max\{\frac{1}{\epsilon}(\ln(\frac{N}{K})+1)\ ,\ \frac{M}{K}(\ln(\frac{1}{\epsilon})+1)\}$.

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

1.1 Organization of the paper

In Section 2, we prove the lower bounds for dispersers and extractors. In Section 3, we present the new upper and lower bounds for depth-two superconcentrators. In the appendix, we describe probabilistic arguments that give Theorem 4.

2 Bounds for dispersers and extractors

In this section we present the lower bounds for dispersers and extractors. As stated earlier, the lower bounds for dispersers claimed in Theorem 2 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 the Theorem 3, where we use the same method to show lower bounds for extractors.

Dispersers. We now describe the proof of Theorem 2. Suppose $G = (V_1 = [N], V_2 = [M], E)$ is a (K,ϵ)-disperser.

We first observe that 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

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

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 X misses v) 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

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

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

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

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

$$N\binom{M - \overline{D}}{L}\binom{M}{L}^{-1} \leq K - 1. \qquad (2)$$

Our lower bounds will follow from this inequality.

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

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

that is,

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

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

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

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

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

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

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

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

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

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

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

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

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

Thus, $\lceil\overline{D}\rceil \geq \dfrac{\log(N/(K-1))}{\log(2/(1-\epsilon))}$. □

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. In fact, for the lower bound to apply, we do not need the full power of an extractor. We identify what we need in the definition of *week-extractor* give below.

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

A weak-extractor is 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.

Claim 2.1 *If $G = (V_1 = [N], V_2 = [M], E)$ is (K,ϵ)-extractor, then G is also a $(2K,\epsilon,p)$-weak-extractor for all p.*

Proof: Suppose G is not a $(2K,\epsilon,p)$-weak-extractor. Then, some $B \subseteq V_2$ of size $\lceil pM\rceil$ ϵ-misses at least $2K$ vertices of V_1. Let $A^- = \{v \in V_1 : \mathrm{disc}(v,B) \leq -\epsilon\}$ and $A^+ = \{v \in V_1 : \mathrm{disc}(v,B) \geq \epsilon\}$. Thus, $|A^- \cup A^+| \geq 2K$; so one of these two sets has size at least K. Suppose it is A^- (the other case is similar). Recall that all vertices in V_1 have the same degree, say D. Now, for $v \in A^-$, we have $|E(v,B)| \leq D(\frac{|B|}{M}-\epsilon)$. We then have $|E(A^-,V_2)| = |A^-|\cdot D$ and $|E(A^-,B)| \leq |A^-|D(\frac{|B|}{M}-\epsilon)$. Thus,

$$\frac{|E(A^-,B)|}{|E(A^-,B)|} - \frac{|B|}{M} \leq -\epsilon.$$

But this contradicts our assumption that G is a (K,ϵ)-extractor. □

Now we prove a lower bound for weak-extractors.

Lemma 2.1 *There exists a constant $C > 0$, such that if $G = (V_1 = [N], V_2 = [M], E)$ is a (K,ϵ,p)-weak-extractor with average degree $\bar{D}$, $K \leq N/C$ and $p \leq 1/2$, then $\bar{D} = \Omega(\frac{p}{\epsilon^2}\log(\frac{N}{K}))$.*

Proof: We proceed as in the case of dispersers, by picking a random $\lceil pM \rceil$ sized subset $R \subseteq V_2$. We will use the following estimate on the probability that R ϵ-misses a vertex $v \in V_1$. We do not justify this here; a proof will be supplied in the final version of the paper.

Claim 2.2 *$\Pr[R$ ϵ-misses $v] \geq C^{-1}\exp(-C\epsilon^2 d_v/p)$, where d_v is the degree of v and $C > 0$ is a constant (independent of d_v).*

It follows that the expected number of vertices missed by R is at least $\sum_{v \in V_1} C^{-1}\exp(-C\frac{\epsilon^2}{p}d_v)$. Since R never misses K vertices, we have

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

Since $\exp(-Cx)$ is a convex function of x, the left hand side is minimized when all the degrees d_v are the same, i.e. they are equal to $\bar{D}$. Thus, we have $N\,C^{-1}\exp(-C\frac{\epsilon^2}{p}\bar{D}) \leq K-1$. By taking lns we obtain

$$\bar{D} \geq \frac{p}{C\epsilon^2}\ln\frac{N}{CK} = \Omega(\frac{p}{\epsilon^2}\ln\frac{N}{K}).$$

□

Now, we return to the proof of Theorem 3. Part (a) follows immediately from Lemma 2.1 by taking $p = 1/2$. We next show that, as in the proof of Theorem 2, part (b) follows from part (a) by reversing the roles of V_1 and V_2.

Claim 2.3 *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. Then for all $T > 2$ and $\frac{K}{N} \leq p \leq \frac{1}{2}$, G' is a $(\frac{4M}{T}, \epsilon' = T\epsilon p, p)$-weak-extractor.*

Proof: Suppose G' is not a $(\frac{4M}{T}, \epsilon' = T\epsilon p, p)$-weak-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

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

Therefore,

$$\begin{aligned} \frac{|E'(A^-,B)|}{|B|D} &\leq \frac{|E'(A^-,V_2')|}{|B|D}(\frac{|B|}{N} - \epsilon') \\ &= \frac{|E(V_1,A^-)|}{ND}(1 - \frac{N\epsilon'}{|B|}). \end{aligned} \quad (3)$$

Since G is an extractor, we have

$$\frac{|E(V_1,A^-)|}{ND} \leq \frac{|A^-|}{M} + \epsilon. \quad (4)$$

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 (3) and (4), we obtain

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

[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 3, we choose $p = K/N$ and $T = 8C$ in the above claim, and conclude that G' is a $(M/(2C), 8\epsilon CK/N, K/N)$-weak-extractor. Then, by Lemma 2.1 we have $\bar{D}(G') \geq \frac{p}{C\epsilon'^2}\ln(\frac{M}{4CM/T})$; that is,

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

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

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 4 (a). Next, we show that the lower bounds on dispersers, shown in Theorem 2 (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.

3.1 The upper bound

Theorem 5 $\text{size}(N) = O(N \log^2 N / \log\log N)$.

Proof: Our construction is based on a similar construction due to Wigderson and Zuckerman [WZ93]. We build a depth-two graph $(A = [N], C, B = [N], E)$, where C is the disjoint union of $C_i = [2\log^{i+1} N]$, $i = 0, \ldots, \log_{\log N} N - 1$. 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 $\log^i N \le K \le \log^{i+1} N$, 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. Thus:

Claim 3.1 *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 [Mes84], Menger's theorem implies that this is sufficient for G to be a superconcentrator (clearly, it is necessary). Thus, all that is left is to check the sizes of the edge sets E_i. By Theorem 4 (a), we may take $|E_i| = O(N \log N)$. Thus, we have $|E(G)| = \sum_i 2|E_i| = O(N \log^2 N / \log\log N)$. $\square$

Remark. This construction differs from the one in [WZ93] in only one respect: in their construction each C_i takes care of all K-sets for $2^i \le K < 2^{i+1}$, whereas in ours each C_i takes care of all K-sets for $\log^i N \le K < \log^{i+1} N$. 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 2). Furthermore, we can build dispersers that recover almost all the random bits we invest (see Theorem 4 (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 $1 \le K \le N$, 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 *any* subset $X \subseteq A$ of cardinality K and look at $\Gamma(X)$, we see that any 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. By applying the disperser lower bound to each of them, we get our lower bound.

Theorem 6 $\text{size}(N) = \Omega(N \cdot (\log N / \log\log N)^2)$.

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

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

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 arbitrarily). Let $Z_i = \Gamma(X_i)$. Clearly $|Z_i| \le K_i \bar{D}$. Let $C_i = Z_i \setminus (Z_1 \bigcup Z_2 \bigcup \ldots \bigcup Z_{i-1})$. 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. 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 $|C_i| \le |Z_i| \le K_i\bar{D}$. (*End of claim.*)

By Theorem 2(a), the number of edges between A and C_i is at least $\Omega(N \cdot \log(N/K_i)/\log(1/(1-\epsilon)))$. 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. $\square$

3.3 The improved lower bound

If we look at the construction of Theorem 5, 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. We show that this is the only way efficient depth-two superconcentrators can be built.

Theorem 7 $\mathrm{size}(N) = \Omega(N \log^2 N / \log\log N)$.

Proof: Let $G = (A = [N], C, B = [N], E)$ be a depth-two N-superconcentrator. As in the proof of Theorem 6, we proceed in stages. In stage i ($i = 1, 2, \ldots$), we consider sets of size $K = K_i = \log^{4i} N$. Let

$$\begin{aligned} C_i &= \{w \in C : \frac{N}{K}\frac{1}{\log^2 N} \leq \deg(w) < \frac{N}{K}\log^2 N\}; \\ D_i &= \{w \in C : \deg(w) < \frac{N}{K}\frac{1}{\log^2 N}\}. \end{aligned}$$

We will show that if $K \leq N^{3/4}$ (i.e. $i \leq \frac{3}{16}(\log N/\log\log N)$), 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

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

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

$$|E(G)| \leq \frac{1}{10}N\log^2 N, \tag{5}$$

for otherwise the theorem follows immediately.

Claim 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 claim. All vertices outside $C_i \cup D_i$ have degree at least $(N/K)\log^2 N$. Then, assumption (5) implies that the number of vertices outside $C_i \cup D_i$ is at most

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

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 $C_i \cup D_i$. (*End of Claim A.*)

We wish to show that G restricted to $A \cup B$ and C_i cannot be sparse. We know from Claim 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 (5)), 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 [Han64]. 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.
- Our second task is 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 $A \cup B$.

Claim 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 claim. For $u \in A \cup B$, let d_u be the number neighbors of u in C_i. Let A' be the set of $N/2$ vertices $u \in A$ with smallest d_u, and B' be the set of $N/2$ vertices $u \in B$ with smallest d_u. We will prove that there is some $u \in A' \cup B'$ with $d_u \geq \frac{1}{5}\log N$. Suppose $u \in A'$. Then, for all $u \in A \setminus A'$ $d_u \geq \frac{1}{5}\log N$, and there are more than $\frac{1}{10}N\log N$ edges incident on C_i, as we want.

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

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

- with probability 1/2, delete all neighbors of w from A';
- with probability 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. Claim 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 3.1 below) that $|E(H)| \geq \dfrac{|S| \cdot |T|}{K} - |S| - |T|$. 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 ($\frac{N2^{-d}}{2} \sim N^{4/5} \gg K$). 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 3.2

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

This gives a lower bound on the average number of connections required between S and T. On the other hand, 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 3.3 $\mathbf{E}[|E(H)|] \;\leq\; |E(G)| \cdot \frac{N}{K\log^2 N} \cdot 2^{-2d}$.

Before proceeding to the proof of these claims, let us complete the proof of Claim B. Putting the two claims together we obtain $|E(G)| \cdot \frac{N}{K\log^2 N} \cdot 2^{-2d} > \frac{N^2 2^{-2d}}{10K}$, i.e., $|E(G)| > \frac{N}{10}\log^2 N$. But then G has too many edges, contradicting (5). (*End of Claim B.*)

Proof of Claim 3.2. 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; in that case, $\mathbf{E}[X_u Y_v] = 2^{-2d}$. 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

$$\mathbf{E}[\frac{|S|\cdot|T|}{K} - |S| - |T|] \;\geq\; \frac{N^2 2^{-2d}}{8K} - N2^{-d} \;>\; \frac{N^2 2^{-2d}}{10K},$$

where the last inequality holds because $2^d K \leq N^{1/5} N^{3/4} = o(N)$. (*End of Claim 3.2.*)

Proof of Claim 3.3. 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

$$\begin{aligned}\sum_{w\in D_i} \deg(w)^2 &\leq \frac{N}{K\log^2 N} \sum_{w\in D_i} \deg(w) \\ &\leq \frac{N}{K\log^2 N} |E(G)|. \qquad (6)\end{aligned}$$

Fix a pair $(u, v) \in A' \times B'$. If u and v have a common neighbor in C_i, then one of them will be deleted and $\Pr[(u, v) \in S \times T] = 0$; otherwise, u and v don't have a common neighbor in C_i and $\Pr[(u, v) \in S \times T] = \Pr[u \in S] \cdot \Pr[v \in T]$. Thus, in both cases, $\Pr[(u, v) \in S \times T] \leq 2^{-2d}$. The claim follows from this and (6) by the linearity of expectation. (*End of Claim 3.3.*) $\square$

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

Lemma 3.1 *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. Each such set cannot miss more than K vertices in V_2; in particular, there are at least $N_2 - K$ edges leaving it. Collecting the contributions from the different sets, we get at least $(\frac{N_1}{K} - 1)(N_2 - K) > \frac{N_1 N_2}{K} - N_1 - N_2$ edges. $\square$

Acknowledgments

We thank Roy Armoni, Oded Goldreich, Nati Linial, Avner Magen, Noam Nisan, Stephen Ponzio, Avi Wigderson and Shiyu Zhou for many helpful discussions. We thank Oded for asking us about a lower bound for the entropy loss of extractors, and Aravind Srinivasan for helping us with the tail estimates.

References

[AP94] N. Alon and P. Pudlak. Superconcentrators of depths 2 and 3; odd levels help (rarely). *Journal of Computer and System Sciences*, 48:194–202, 1994.

[Bol78] B. Bollobás. *Extremal Graph Theory*. Academic Press, London, 1 edition, 1978.

[DDPW83] D. Dolev, C. Dwork, N. Pippenger, and A. Wigderson. Superconcentrators, generalizers and generalized connectors with limited depth. In *ACM Symposium on Theory of Computing (STOC)*, pages 42–51, 1983.

[GW94] O. Goldreich and A. Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. In *Proceedings of the 26th Annual ACM Symposium on the Theory of Computing*, ACM, pages 574–583, 1994.

[Han64] G. Hansel. Nombre minimal de contacts de fermature nécessaires pour réaliser une fonction booléenne symétrique de n variables. *C. R. Acad. Sci. Paris*, 258:6037–6040, 1964.

[KS67] G. Katona and E. Szemerédi. On a problem of graph theory. *Studia Sci. Math. Hungarica*, 2:23–28, 1967.

[Mes84] R. Meshulam. A geometric construction of a superconcentrator of depth 2. *Theoretical Computer Science*, 32:215–219, 1984.

[Nis96] N. Nisan. Refining randomness: How and why. In *IEEE Conference on Computational Complexity*, pages 44–58, 1996.

[NW] N. Nisan and A. Wigderson. Research pearls in theory of computation: Homework 2, problem 2. http://www.cs.huji.ac.il/course/pearls/.

[NZ93] N. Nisan and D. Zuckerman. More deterministic simulation in logspace. In *Proceedings of the 25th Annual ACM Symposium on the Theory of Computing*, ACM, pages 235–244, 1993.

[Pip77] N. Pippenger. Superconcentrators. *SIAM Journal on Computing*, pages 298–304, 1977.

[Pip82] N. Pippenger. Superconcentrators of depth 2. *Journal of Computer and System Sciences*, 24:82–90, 1982.

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

[Pud94] P. Pudlak. Communication in bounded depth circuits. *Combinatorica*, 14(2):203–216, 1994.

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

[Spi96] D. Spielman. Linear time encodable and decodable error-correcting codes. *IEEE Trans. on Information Theory*, 42(6):1723–1731, November 1996.

[T96a] A. Ta-Shma. On extracting randomness from weak random sources. In *STOC 1996*, pages 276–285, May 1996.

[T96b] A. Ta-Shma. *Refining Randomness*. PhD thesis, The Hebrew University, Jerusalem, June 1996.

[Val76] L. Valiant. Graph-theoretic properties in computational complexity. *Journal of Computer and System Sciences*, 13:278–285, 1976.

[WZ93] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. In *Proceedings of the 25th Annual ACM Symposium on the Theory of Computing*, ACM, pages 245–251, 1993.

Appendix

Existence of dispersers and extractors

In this section, we show that the probabilistic arguments used in Sipser [Sip88] yield dispersers and extractors with parameters close to the lower bound shown in Section 2.

Dispersers. Now we turn to the upper bound stated in Theorem 4 (a). We construct a 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 if there is some subset of V_2 of size ϵM that misses some K vertices of V_1. vertices of V_2. Thus, $\Pr[G \text{ is not a } (K, \epsilon) \text{ disperser}]$ is at most

$$\begin{aligned} & \binom{N}{K}\binom{M}{L}(1 - L/M)^{KD} \\ < \quad & (eN/K)^K (eM/L)^L (1 - L/M)^{KD} \\ \leq \quad & (eN/K)^K (eM/L)^L \exp(-LKD/M)). \quad (7) \end{aligned}$$

[To justify the last inequality we use $1 - x \leq e^{-x}$.]

Since $D \geq \frac{2M}{L}(1+\ln(N/K))$ we have $\frac{LD}{2M} \geq 1 + \ln(N/K)$; hence, $(eN/K)^K = \exp(1+\ln(N/K)) \leq exp(LD/(2M))$. Thus

$$(eN/K)^K \leq \exp(LKD/(2M)).$$

Similarly, since $D \geq \frac{2M}{K}(1+\ln(M/L))$, we have

$$(eM/L)^K \leq \exp(LKD/(2M)).$$

Thus, $(eN/K)^K \cdot (eM/L)^K \leq \exp(LKD/M)$. Therefore, by (7), 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 will use Definition 1.3. We will obtain a bipartite graph $G = (V_1 = [N], V_2 = [M], E)$ where all vertices in V_1 will have the same degree D, such that for every $S \subseteq V_1$ of cardinality K, and any $R \subseteq V_2$,

$$\left| \frac{|E(S,R)|}{KD} - \frac{|R|}{N} \right| < \epsilon. \tag{8}$$

We first observe that (8) can be replaced by a seemingly weaker condition.

Claim .4 *If for every $S \subseteq V_1$ of size K, and any $R \subseteq V_2$, $\frac{|E(S,R)|}{<} KD(\frac{|R|}{N}+\epsilon)$, then G is a (K,ϵ)-extractor.*

For, if there exist some S and R wth $|E(S,R)| \leq KD(|R|/N - \epsilon)$, then consider the set $\bar{R} = V_2 \setminus R$. We have $|E(S,\bar{R})| > KD(\frac{|\bar{R}|}{N}+\epsilon)$, 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 for each $v \in V_1$.

Fix $S \subseteq V_1$ of size K and $R \subseteq V_2$. Let $p = \frac{|R|}{N}$. We wish to estimate the probability (over the choices of the edges) that $|E(S,R)| > KD(p+\epsilon)$. 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 $p = |R|/M$ and the value 0 with probability $1-p$. By Chernoff's inequality, we have

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

Thus, $\Pr[G \text{ is not a } (K,\epsilon)\text{-extractor}]$ 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. □