Randomness Conductors and Constant-Degree Expansion Beyond the Degree / 2 Barrier

PRELIMINARY VERSION

Michael Capalbo[*] Omer Reingold[†] Salil Vadhan[‡] Avi Wigderson[§]

November 15, 2001

Abstract

The main concrete result of this paper is the first explicit construction of constant degree *lossless* expanders. In these graphs, the expansion factor is almost as large as possible: $(1-\epsilon)D$, where D is the degree and ϵ is an arbitrarily small constant. Such graphs are essential components in networks that can implement fast distributed, routing algorithms e.g. [PU89, ALM96, BFU99]. They are useful in expander-based linear codes e.g. [SS96, Spi96, LMSS01]. Their highly unbalanced relatives are used in various storage schemes [UW87, BMRS00] and are the basis of hard tautologies for various proof systems [BW99, ABRW00, AR01]. The best previous explicit constructions gave expansion factor $D/2$, which is too weak for the above applications. The $D/2$ bound was obtained via the eigenvalue method, and as shown in [Kah95], this method cannot give better bounds.

The main abstract contribution of this paper is the introduction and initial study of *randomness conductors*, a notion which generalizes extractors, expanders, condensers and other similar objects. In all these functions, certain guarantee on the input "entropy" is converted to a guarantee on the output "entropy". For historical reasons, specific objects used specific guarantees of different flavors (eg in expanders entropy means "support size", and their property is satisfied whenever input entropy is small. In contrast, in extractors, entropy means "min-entropy" and their property is satisfied whenever input entropy is large). We show that the flexibility afforded by the conductor definition leads to interesting combinations of these objects, and to better constructions such as those above.

The main technical tool in these constructions is a natural generalization to conductors, of the zig-zag products defined by [RVW00] for expanders and extractors. This product maintains certain conductivity properties of the conductors it combines, leading to new iterative constructions.

Keywords: expander graphs, extractors, condensers, graph products.

[*] Institute for Advanced Study, School of Mathematics, Einstein Drive, Princeton, NJ 08540. E-mail: `mrc@ias.edu`.

[†] AT&T Labs - Research. Room A243, 180 Park Avenue, Bldg. 103, Florham Park, NJ, 07932, USA. E-mail: `omer@research.att.com`. Part of this research was performed while visiting the Institute for Advanced Study, Princeton, NJ.

[‡] Harvard University, Division of Engineering and Applied Sciences, Maxwell Dworkin 337, 33 Oxford Street Cambridge, MA 02138, USA. E-mail: `salil@eecs.harvard.edu.` URL: `http://www.eecs.harvard.edu/~salil`. Work begun while at MIT and the Institute for Advanced Study, supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

[§] Institute for Advanced Study, Princeton and the Hebrew University, Jerusalem. Address: Institute for Advanced Study, School of Mathematics, Einstein Drive, Princeton, NJ 08540. E-mail: `avi@ias.edu`.

1 Introduction

The quest for explicit construction of extractors, expanders, and their relative functions which "enhance" randomness, has been one of the richest areas in the interaction between computer science and pure mathematics. Moreover, the huge by now, and still growing, diverse set of applications of such functions in both computer science and pure mathematics makes them central objects of further understanding. We will not elaborate here on either the constructions nor the applications which are not directly relevant to this paper.

Our paper can certainly be viewed as another step in this important process. The progress made here is of two types. The first is in resolving a long-standing open problem in this area — the explicit construction of "lossless" expanders and their unbalanced relatives. The second is in suggesting a framework (which we call *conductors* that encompasses all previously studied "randomness enhancing" functions. Needless to say, the two are related – our new construction was discovered in, and is best described by, the framework of conductors. The rest of the introduction describe both.

1.1 Lossless Expanders

In this subsection we define lossless expanders[1], and explain why they were hard to construct by existing techniques. We then briefly discuss their applicability in several areas.

Consider a bipartite graph G with N inputs I, M outputs O, and every input connected to D outputs. G is called an (K, A)-*expander* if every set X of at most K inputs is connected to at least $A \cdot |X|$ outputs.

Clearly, the best one can hope for with these parameters is A as close as possible to D; when $A = (1-\epsilon)\cdot D$ for a small ϵ we call the expander *lossless*. We can hope for such expansion factor only up to $K = M/D$. A nonconstructive probabilistic argument shows that such graphs do exist with $D = O(\log(N/M))$, and this value of D is best possible.

Our main result is an explicit construction of such "lossless" expanders for any setting of the parameters N, M. When they are within a constant factor of each other, the degree of our graphs is constant, and linear-sized subsets of N expand losslessly. More specifically, the degree of our graphs is $D = \mathrm{polylog}(N/M)$ when N/M is relatively small (so that an optimal graph of size $\mathrm{poly}(N/M)$ can be found by, say, exhaustive search) and $D = \exp(\mathrm{polyloglog}(N/M))$ in general. (Here, for simplicity, we fix ϵ to be an arbitrarily small constant). The size of sets that expand losslessly in all cases is $\Omega(M/D)$, which is the best possible up to a constant factor.

1.2 Previous Work

The best previous construction of constant degree, even for the special case $N = M$, achieved only expansion $A = D/2$. It is obtained from expanders which have optimal second largest eigenvalue — namely, the Ramanujan graphs of [LPS88, Mar88]. Moreover, Kahale [Kah95] showed that some Ramanujan graphs do not expand by more than $D/2$, showing that to get lossless expanders one has to bypass the eigenvalue method.

Lossless expanders with weaker parameters were obtained before. Ta-Shma, Umans and Zuckerman [TUZ01] coined the term "lossless condenser"(which we call here "lossless conductors"), and gave a very elegant construction for the very unbalanced case ($N \gg M$), with almost optimal degree $D = \mathrm{polylog}(N)$ (though with a suboptimal bound $K < M^\epsilon$ on the size of sets which losslessly expand). The only constant-degree lossless expanders (with $N = M$) were obtained by Alon [Alo95] based on graphs of high girth, but again only very small sets ($A = N^\alpha$ for some constant α) expand losslessly.

[1] We use this term here for both the balanced and unbalanced variety. We later use the term lossless conductors for them.

Also, weaker objects of a similar nature were constructed before. Raz and Reingold [RR99] introduced a method which appends a buffer to a "lossy" extractor, which retains the "lost" entropy. This translates to highly unbalanced, nonconstant degree graphs which are lossless for sets of size of a given size K (rather than for all sets of size up to K). Their technique is essential in our construction. Very recently, Capalbo [Cap01] constructed explicit *unique neighbor expanders* (for the case $N = M$) of constant degree. In these graphs, for any set X, $|X| \leq K$ of inputs, a constant fraction of the vertices in $N(X)$ has a unique neighbor in X. This property trivially holds in lossless expanders, but turns out to be sufficient in some of their applications. Capalbo's construction uses the high min-entropy extractors of [RVW00] and graph products. Our paper may be viewed also as a significant extension of his construction, as well as that of the "min-entropy" expanders in [RVW00].

1.3 Applications

We now turn to list a wide variety of known applications of (balanced and unbalanced) lossless expanders. In almost all of them the application depended on a probabilistic construction of such an object, and our construction is the first to make them explicit.

Recall from above that a (K, A) expander is a bipartite graph with N input nodes of degree D and M output nodes, such that every input set X of size at most K has at least $A \cdot |X|$ output nodes as neighbors, and a lossless expander is one in which $A = (1 - \epsilon)D$. We will maintain this notation throughout the applications below.

- Distributed Routing in Networks

  There has been substantial interest, and literature, on constructing networks in which many pairs of nodes can be connected via vertex or edge disjoint paths, and furthermore so that these paths may be found efficiently, hopefully in a distributed manner and even if requests for connections arrive on-line. Examples are the papers [PU89, ALM96, BFU99] In essentially all of them the networks are lossless expanders, or at least contain them as components. To see why consider the following easier problem, which is actually at the heart of most of these algorithms.

  Assume that the network is itself a lossless expander G, that X is some set of inputs with $|X| < K$, and that these inputs wish to find a matching to the output quickly and distributively (this may be viewed as the first step in constructing vertex disjoint paths). Note that in constant time, by sending messages to their direct neighbors, a constant fraction of the elements in X can match themselves to their unique neighbors, and remove themselves from X. Iterating this will end in $O(\log |X|)$ steps (which in a more general algorithm will be pipelined) and linear work.

- Linear Time Decodable Error-Correcting Codes

  Again, a large body of work, best known under LDPC (Low Density Parity Check) codes, constructs good codes from graphs with good expansion properties. (See [LMSS01, SS96, Spi96] and the references therein.) The following, which is from [SS96], illustrates the power of lossless expanders in this context. Specifically, they yield asymptotically good linear codes of *every* constant rate, with a trivial linear-time (and $O(\log n)$ parallel steps) decoding algorithm.

  Let G be a lossless expander. A codeword (of length N) is an assignment of bits to the inputs, so that every output has zero parity of the inputs it is connected to (in other words the outputs describe the parity check matrix of the code). The rate is $1 - M/N$, which can be made an arbitrarily close to 1 with constant degree D. We now show how to correct any set of at most K errors. The (possibly corrupted) message is an assignment to the input nodes which induces some values on the output nodes via parity. While not all outputs have a zero value, every input node (independently) acts as follows: if flipping its value will decrease the number of 1's in the outputs by (say) $D/3$, it flips. It is easy to see, that the total number of 1's in the output will shrink by a constant factor each round, and that every input needs only a constant time to decide what to do.

By using our lossless expanders in this construction, the resulting codes have relative rate $1 - \delta$ and minimum distance $\delta/\text{polylog}(1/\delta)$, which, for small δ, beats the Zyablov bound and is quite close to the Gilbert-Varshamov bound.

- Bitprobe Complexity of Storing Subsets

  An ingenious scheme for storing K-subsets of $[N]$ in binary vectors of length M was recently proposed by [BMRS00]. The scheme allows to determine (with high probability) membership of any element $v \in [N]$ in the stored set by querying only *one* random bit in the vector. The optimal construction of smallest value of M (for constant error) relies on lossless expanders. Let us see how.

  Let G be an $(2K, (1 - \epsilon)D)$ expander with N inputs, M outputs, and degree D (this value is not important for this application). Given a set X of inputs of size K, we will label the outputs with binary values so that the vast majority of neighbors of each neighbors would correctly give membership. Namely, every element of X has most (say) $3\epsilon D$ of its neighbors labelled 0, and every element not in X will have at most $3\epsilon D$ neighbors labelled 1. If we achieve that, querying a random neighbor of $v \in [M]$ will only err with probability 3ϵ about membership of v in X.

  Why should such a labelling exist? As in the previous examples, let's attempt to get it greedily. First, label all output vertices $N(X)$ that are connected to X by 1, and the rest by 0. This classifies correctly vertices in X, but might misclassify vertices in a set Y disjoint from X, each of which has at least $3\epsilon D$ neighboring outputs in $N(X)$. Fix the problem by (re)labelling all outputs in $N(Y)$ by 0. This certainly fix the problem in Y, but may create one in a subset Z of A. Fix $N(Z)$ to 1, etc. The key point is that losslessness implies that the size of the new "problematic set" is at most half the size of the previous one. This is indeed an efficient algorithm for finding the labelling.

- Fault-tolerance and a Distributed Storage Method

  Lossless expanders have an incredible fault-tolerance: an adversary can remove *most* of the neighbors of *every* input, and the graph will remain a lossless expander! More precisely, let G be an $(K, (1 - \epsilon)D)$-expander of degree D. For *every* subgraph of input degree $D' = \gamma D$, is a $(K, (1 - \epsilon/\gamma)D')$-expander.

  This property was used (with $\gamma \approx 1/2$) in a distributed storage scheme due to Upfal and Wigderson [UW87], who gave a near-optimal *deterministic* simulation of PRAMs by a network of communicating processors. Both models have M processors, but they differ in that in the first, every data item can be accessed in unit time, whereas in the second, items which reside in the same processor cannot be accessed simultaneously. If the number of data items N used (read and updated) by the PRAM program is much larger than M, any naive method for distributing the data items will fail.

  The idea in [UW87] is to use a lossless expander where the inputs represent the N data items, and the outputs represent (the memories) of the M processors. Each item has $D = 2c - 1$ "copies", which are distributed to its neighbors in the. When attempting to read or update a data item, each processor is required to access only c copies. When updating, it updates all c (and time-stamps them). When reading, it takes the value of the most recently updated among the c "copies" it has. Intersection of any two c subsets implies consistency.

  How about efficiency? In each phase (simulating one PRAM step) all processors try to access some data item each. They follow a (nontrivial) distributed algorithm to collect c copies each. The analysis heavily uses the fact that at any point in this algorithm, the residual graph (between data items and their yet unobtained copies) is still a lossless expander (by the above fault-tolerance and the fact that at least c neighbors still remain for each unrecovered data item). This guarantees convergence in $O(c)$ rounds.

- Hard Tautologies in Proof Complexity

  The field of Proof Complexity studies propositional proof systems, and tries to prove lower bounds on the size of such proofs for concrete tautologies. There has been significant progress in this field, especially in accomplishing this task for relatively simple proof systems. (See, e.g., the excellent survey [BP98]. A sequence of works [Tse68, Urq87, CS88, BW99, ABRW00, AR01] has gradually elucidated expansion as a key to hard tautologies for several complexity measures (width/degree, size, space) in the important (simple) proof systems Resolution and Polynomial Calculus.

  Here is a general recipe for constructing such tautologies from any lossless expander G, as suggested in [ABRW00]. View G as a circuit, where each input node computes some function of the output nodes connected to it. (Note that this is in reverse to the parity circuit in the coding application.) The choice of functions depends on the proof system, and these choices are made in [ABRW00, AR01] for resolution and the polynomial calculus, respectively. Now the tautology, whose propositional variables are the potential values on output nodes, encodes a statement such as "the zero vector is *not* generated by this circuit for any setting of values on the output nodes". The unique neighbor property is used (nontrivially) to show that, in some sense, the statements for every input are "nearly independent" (in much the same way near-disjointness is used in the Nisan-Wigderson pseudorandom generator [NW94]), and thus every proof will have to examine exponentially many "possibilities".

  It should be noted that the probabilistic proof of existence for lossless expanders was sufficient for this application; in proof complexity, unlike computational complexity, existential lower bounds are as interesting as constructive ones. Still, explicit examples are always more informative.

1.4 Randomness Conductors

In this subsection we try to motivate a general framework for studying "randomness enhancing" functions. Each of the many variants on this theme: expanders, concentrators, dispersers, extractors, condensers, ... may be viewed a function $f : [N] \times [D] \to [M]$. Each function guarantees some randomness properties of the distribution $f(X, U)$ given some guarantees on the randomness in the distribution X, where U is the uniform distribution on $[D]$.

As these objects were originally defined for different sets of applications and motivations, what is exactly meant by "randomness" and "guarantees" in the above description can vary quite a bit. These choices have different advantages and disadvantages. For example:

Vertex Expansion This is the most classical measure of expansion — the support of $f(X, U)$ should be larger than the support of X, provided the latter is not too large. It is also used to define *dispersers* [Sip88] and *a-expanding graphs* [Pip87], though these refer to X of given size support. While this measure is often the one that we want in applications, it tends to be too weak for compositions.

Eigenvalue Expansion It is well-known that the second largest eigenvalue of a graph is a good measure of its expansion [Tan84, AM84, Alo86]. This measure turns out to be equivalent to measuring the *Renyi entropy* of $f(X, U)$ as a function of the Renyi entropy of X. The eigenvalue was very convenient for analyzing algebraic constructions of expanders, and indeed was the measure of choice for almost all previous constructions of constant-degree expanders. However, in some ways it is too strong. As mentioned earlier, it cannot give expansion greater than $D/2$ and also cannot achieve small degree for very unbalanced graphs.

Extractors Extractors, introduced by Nisan and Zuckerman [NZ96], ask that $f(X, U)$ is close (in statistical difference) to the uniform distribution on $[M]$ provided that X has sufficient *min-entropy*. This turns out to overcome most of the deficiencies of the notions mentioned above — extractors can have very small degree for unbalanced graphs, and are eminently composable. (For example, since the output of an extractor is close to uniform, it is very natural to use the output of one extractor as the second input to another.) On the other hand, extractors cannot be lossless [NZ96, RT97], and also their definition only discusses X whose min-entropy is at least some value and thereby does not guarantee expansion of small sets.

Condensers Condensers differ from extractors in that, instead of asking that $f(X, U)$ is close to uniform, they only require $f(X, U)$ is close to some distribution having at least a certain amount of min-entropy. Various formalizations of this basic idea have appeared in the recent extractor literature [RR99, RSW00, TUZ01], where it has been seen that condensers can be lossless, and can be used to discuss expansion of small sets. However, since condensers do not provide an almost-uniform output, they are not quite as composable as extractors (though they compose quite nicely with extractors).

Not surprisingly, there are numerous connections and reductions between the above objects, and our paper could be viewed as making more specific connections of this type. However, we feel that a global view of all these objects (as conductors) is a better description of how we came about our construction, and that these objects merit more study in the future. In particular, it seems useful to have a single notion which captures both extraction (which must be lossy) and lossless condensers simultaneously.

In the most general form, *randomness conductors*[2] capture all of the above objects: every function f is a conductor, and its quality is measured for all values of parameters: for every two values of entropy, k_{in} and k_{out}, we measure the statistical difference of $f(X, U)$ to the nearest distribution of entropy k_{out}, taking worst case over all sources X of entropy k_{in}. In this paper we choose "entropy" above to mean min-entropy, but we suspect that similar results can be obtained when it means Renyi's 2-entropy.

In this work we'll restrict ourselves to *simple conductors*[3]. In these we fix both error (statistical difference) as well as the difference between the input and output entropies ($k_{out} - k_{in}$) to fixed values. When this difference is $d = \log_2 D$, the conductor is *lossless*, as all the incoming entropy (from the source X and the uniform distribution U is close to being preserved at the output. When the output entropy can reach $m = \log_2 M$, we have an *extracting conductor*, which combines extractors and condensers in one. Needless to say, there are other types of conductors we use, but we'll delay their description to the technical section.

It turns out that quite a few known objects, such as expanders, hash functions, and some special constructions of extractors (e.g. those of Trevisan [Tre99, RRV99]) are conductors of good parameters. We just need to combine them in the right way! Our main technical result is a new zig-zag product for conductors which, like in the zig-zag products for expanders and extractors in [RVW00], combines (three) conductors into a larger one, maintaining their conductivity properties. This leads in particular to our construction of constant-degree lossless expanders. While the intuition behind the new zigzag product is similar to the two old ones in [RVW00], the technical details involved in proving its properties are more delicate, due to the higher requirements from conductors.

[2] The analogy with water, heat or electricity conductors is meant to be suggestive.

[3] which are still more general than what we really need.

2 Preliminaries

Expanders are graphs which are sparse but nevertheless highly connected. The standard definition of expanders is based on the notion of set expansion – every (not too large) subset of vertices in an expander should be connected to a large number of neighbors. A quantitative version of such a definition follows:

Definition 2.1 *A bipartite graph $G = ([N], [M], E)$ is a (K, A)-***expander** *if every subset $X \subseteq [N]$ of at most K vertices is connected to at least $A \cdot |X|$ neighbors. (Note that the parameters K and A must satisfy $M/K \leq A \leq N$.)*

Typically we are interested maximizing the expansion factor A while minimizing the left-degree D. Every bipartite graph as above can be viewed as a function $\mathrm{E} : [N] \times [D] \to [M]$, where $\mathrm{E}(x, r)$ is the r'th neighbor of x, and conversely. (We allow multiple edges between two vertices.) In this representation, a (somewhat convoluted) way of viewing set expansion is to say that for every probability distribution X on $[N]$ whose support $\mathrm{Supp}(X)$ is of size at most K, $\mathrm{E}(X, U)$ has support of size at least $A \cdot \mathrm{Supp}(X)$. Thus, if we think of "support size" as a measure of randomness, then expanders can be viewed as "randomness enhancing" functions. However, it turns out to be extremely useful to adopt stronger measures of "randomness" than support size, and to do so we need some definitions.

Let X and Y be random variables over a set S (throughout the paper we identify random variables and their distributions). The **min-entropy** of X is defined to be

$$\mathrm{H}_\infty(X) \overset{\mathrm{def}}{=} \log(1/\max_{a \in S} \Pr[X = a]),$$

where here and throughout this paper, all logarithms are base 2. X is a k-**source** if $\mathrm{H}_\infty(X) \geq k$. (In particular, the uniform distribution on a set of size 2^k is a k-source.) We say that X and Y are ε-**close** if the statistical difference between X and Y is at most ε. That is, if

$$\max_{P \subseteq S} |\Pr[X \in P] - \Pr[Y \in P]| = \frac{1}{2} \sum_{a \in S} |\Pr[X = a] - \Pr[Y = a]| \leq \varepsilon$$

X is a (k, ε)-**source** if it is ε-close to some k-**source**. In Appendix A, we discuss another useful entropy measure, namely *Renyi entropy*, and its relationship to min-entropy.

Now we are prepared to discuss other kinds of "randomness enhancing" functions, and in doing so, it will be convenient to represent everything in bits. For any integer n, we denote by (n) the set of all n-bit strings, $\{0, 1\}^n$. Denote by U_n the uniform distribution over (n).

Definition 2.2 ([NZ96]) *A function $\mathrm{E} : (n) \times (d) \mapsto (m)$ is a (k, ε)-***extractor** *if for any k-source X over (n), the distribution $\mathrm{E}(X, U_d)$ is ε-close to U_m.*

Viewed as a bipartite graph, an extractor guarantees that all subsets of the left-hand side $[N]$ of size *at least* K vertices have a neighborhood of size at least $(1 - \varepsilon) \cdot M$ (and even more, that the edges are distributed almost uniformly among the neighbors.) Here, and throughout the paper, we adopt that capital letters are 2 taken to the corresponding lowercase letter, e.g. $N = 2^n$, $K = 2^k$, $M = 2^m$.

3 Conductors

3.1 Conductors, a broad-spectrum approach

As discussed in the introduction, we consider one of the main contributions of this paper to be the introduction of randomness conductors. In this definition we would like to to encompass a wide spectrum of "randomness enhancing" combinatorial objects. Loosely, all of these objects can be viewed as functions $\mathrm{E} : (n) \times (d) \to (m)$ with some relation between the randomness guarantee on the distribution X of their first input and the randomness guarantee of their output distribution $\mathrm{E}(X, U_d)$.

Definition 3.1 (randomness conductors) *Let ε be a real valued function $\varepsilon : [0, n] \times [0, m] \mapsto [0, 1]$, (where $[a, b]$ denotes the real interval between a and b). A function $\mathrm{E} : (n) \times (d) \mapsto (m)$ is an $\varepsilon(\cdot, \cdot)$* **randomness conductor** *if for any $k_{in} \in \{0, 1, 2, \ldots, n\}$, any $k_{out} \in \{0, 1, 2, \ldots, m\}$ and any k_{in}-source X over (n), the distribution $\mathrm{E}(X, U_d)$ is a $(k_{out}, \varepsilon(k_{in}, k_{out}))$-source.*

Note that any function $\mathrm{E} : (n) \times (d) \mapsto (m)$, is an $\varepsilon(\cdot, \cdot)$ randomness conductor for ε which is identically one. More generally, the requirement of Definition 3.1 from E, is moot for any specific pair (k_{in}, k_{out}) such that $\varepsilon(k_{in}, k_{out}) = 1$. This property makes objects as extractors or condensers a restricted special case of conductors.

Entropy Measures. Definition 3.1 is flexible enough to handle a wide variety of settings, previously dealt with by expanders, extractors, condensers, hash functions and other objects. In particular, the definition can handle (a) The balanced case ($m = n$) and the imbalanced case ($m < n$). (b) The lossless case ($k_{out} = k_{in} + d$) and the lossy case ($k_{out} < k_{in} + d$). (c) The extracting scenario (where $k_{out} = m$) and the condenser scenario (where k_{out} may be much smaller than m). Nevertheless, for our definition we fix a particular measure of randomness. Namely, we use a combination of statistical difference (L_1 norm) and min-entropy. In this we follow the definition of extractors. As discussed in the introduction, our motivation for such a definition includes the following considerations:

- This definition is strong enough to imply vertex expansion: For every k_{in}, k_{out}, if the support of a distribution X is at least $2^{k_{in}}$, then the support of $\mathrm{E}(X, U_d)$ is at least $(1 - \varepsilon(k_{in}, k_{out})) \cdot 2^{k_{out}}$.
- This measure if randomness is very amenable to composition as demonstrated by the extractor literature and by the results of this paper.
- This definition is not "too strong" in the sense that it allows relatively small seed length (the parameter d) even in the unbalanced case where $m < n$. The corresponding definitions that are based solely on min-entropy or Renyi entropy require very high degree in this case (cf., Remark 3.2).

We note that an alternative definition may involve a combination of statistical difference and Renyi entropy (rather than min-entropy). We suspect that similar results can be obtained with this definition. (In fact, the two definitions are closely related — see Appendix A.)

Remark 3.2 It is tempting to define conductors in terms of min-entropy alone. That is, for $\mathrm{E} : (n) \times (d) \mapsto (m)$, and some collection of pairs $\{(k_{in}, k_{out})\}$ to require that every k_{in}-source X will be transformed into a k_{out}-source $\mathrm{E}(X, U_d)$ (this is equivalent to restricting $\varepsilon(\cdot, \cdot)$ in Definition 3.1 to only assume $\{0, 1\}$ values). However, such a definition is too strong unless the setting of parameters is trivial (i.e., $k_{out} \leq d$ or $k_{out} \leq k_{in} + m - n$): Let z be any value in (m) and define the source X_z to be the uniform distribution over $S_z = \{x \in (n) | E(x, y) = z \text{ for some } y \in (d)\}$. If for some $z \in (m)$, the set S_z is of size at least $2^{k_{in}}$, then we have that X_z is a k_{in}-source while $\mathrm{H}_\infty(\mathrm{E}(X_z, U_d)) \leq d$ (since $\Pr[\mathrm{E}(X_z, U_d) = z] \geq 2^{-d}$). Otherwise, let $z \in (m)$ be such that the set of pairs $\{(x \in (n), y \in (d)) | E(x, y) = z\}$ is of size at least 2^{n+d-m}. (Such a z exists by a simple counting argument.) Let X be the uniform distribution over some set of size $2^{k_{in}}$ that contains S_z. Then X is a k_{in}-source while $\mathrm{H}_\infty(\mathrm{E}(X, U_d)) \leq k_{in} + m - n$ since $\Pr[\mathrm{E}(X, U_d) = z] \geq 2^{n+d-m} / 2^{k_{in}+d}$.

We conclude that employing statistical difference in the definition of conductors is indeed essential.

3.2 Special cases of interest

For the constructions of this paper, it will simplify notation to work with several special cases of conductors. In these special cases, both the error parameter and the difference between the input and output min-entropies will be fixed rather than varying as in the general definition.

Definition 3.3 (simple conductors) *A function* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is an* $(k_{max}, \varepsilon, a)$ simple conductor *if for any* $0 \leq k \leq k_{max}$*, and any* k*-source* X *over* (n)*, the distribution* $\mathrm{E}(X, U_d)$ *is a* $(k + a, \varepsilon)$*-source.*

In the above definition, *we allow* a *to be negative*, so that we can discuss conductors which lose more than d bits of entropy. Now we look at two further restrictions. The first of these requires that the conductor is an extractor when the input min-entropy is large. This forces $k_{max} = m - a$, so we drop k_{max} from the notation.

Definition 3.4 (extracting conductors) *A function* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is an* (ε, a) extracting conductor *if for any* $0 \leq k \leq m - a$*, and any* k*-source* X *over* (n)*, the distribution* $\mathrm{E}(X, U_d)$ *is a* $(k + a, \varepsilon)$*-source.*

Note that if $\mathrm{E} : (n) \times (d) \mapsto (m)$ is an (ε, a) extracting conductor then it is also an $(m - a, \varepsilon)$ extractor.

The second restriction is that the conductor is *lossless*. That is, the output min-entropy equals the total amount of randomness invested, namely the input min-entropy plus the number of truly random bits. In other words, $a = d$, so we drop a from the notation.

Definition 3.5 (lossless conductors) *A function* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is a* (k_{max}, ε) lossless conductor *if for any* $0 \leq k \leq k_{max}$*, and any* k*-source* X *over* (n)*, the distribution* $\mathrm{E}(X, U_d)$ *is a* $(k + d, \varepsilon)$*-source.*

As observed by Ta-Shma, Umans, and Zuckerman [TUZ01], lossless conductors[4] are equivalent to bipartite graphs of left-degree $D = 2^d$ such that every set of left vertices of size at most 2^k expands by a factor $(1 - \varepsilon) \cdot D$.

The last two special cases combine the above two cases, by requiring that we have a lossless conductor such that a prefix of the output is an extracting conductor.

Definition 3.6 (buffer conductors) *A pair of functions* $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \mapsto (m) \times (b)$ *is an* $(k_{max}, \varepsilon, a)$ buffer conductor *if* E *is a* (ε, a) *extracting conductor and* $\mathrm{E}' = \langle \mathrm{E}, \mathrm{C} \rangle$ *is an* (k_{max}, ε) *lossless conductor.*

It will also be useful for our construction to consider a restricted type of buffer conductors, where $\mathrm{E}' = \langle \mathrm{E}, \mathrm{C} \rangle$ is a permutation (note that in this case, E' is trivially also an $(n, 0)$ lossless conductor).

Definition 3.7 (permutation conductors) *A pair of functions* $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \mapsto (m) \times (b)$*, where* $n + d = m + b$ *is a* (ε, a) permutation conductor *if* E *is a* (ε, a) *extracting conductor and* $\mathrm{E}' = \langle \mathrm{E}, \mathrm{C} \rangle$ *is a permutation over (n+d).*

4 Some Constructions and lower bounds

In this section we describe some basic conductors. The first set of conductors are constructed probabilistically. The parameters achieved by these probabilistic constructions are essentially optimal, as demonstrated by the lower bounds given in Appendix B (which also contains the proofs of most results from this section). These will serve us in two ways. They will be components in our zig-zag product when their size is a fixed

[4] They referred to such objects as *lossless condensers*.

constant (as they can be found by brute force). Furthermore, we'll see that for every size, our final explicit constructions come very close to the performance of these random constructions.

The second set are "known" explicit conductors. By this we mean past constructions of random-like objects, such as expanders, extractors and hash functions, which happen to have useful parameters as conductors for our zig-zag.

We later extend known composition techniques from extractors and condensers to conductors. This will help improve the parameters of the above.

The proofs for Theorems 4.1–4.3 employ standard probabilistic arguments and are deferred to the final version. Closely matching lower bounds can be obtained by reductions to the known lower bounds for extractors [NZ96, RT97]; details are given in Appendix B.

Lemma 4.1 (nonconstructive extracting conductor) *For every n, $m \leq n$, and $\varepsilon > 0$, there is an (ε, a) extracting conductor* $\mathrm{E} : (n) \times (d) \rightarrow (m)$ *with*

- $d = \log(n - m + 1) + 2\log(1/\varepsilon) + O(1)$,
- $a = d - 2\log(1/\varepsilon) - O(1)$

In terms of graphs, these parameters say the degree is $D = \Theta(\log(2N/M)/\varepsilon^2)$, and the expansion factor is $A = \Theta(\varepsilon^2 D)$. The expression for a says that even these optimal extracting conductors lose $2\log(1/\varepsilon)$ bits of entropy.

Lemma 4.2 (nonconstructive lossless conductor) *For every n, $m \leq n$, and $\varepsilon > 0$, there is a (k_{max}, ε) lossless conductor* $\mathrm{E} : (n) \times (d) \rightarrow (m)$ *with*

- $d = \log(n - m + 1) + \log(1/\varepsilon) + O(1)$.
- $k_{max} = m - d - \log(1/eps) - O(1)$.

In terms of graphs, these parameters say that $D = \Theta(\log(2N/M)/\varepsilon)$ and the size of sets that expand losslessly is $K_{max} = \Theta(\varepsilon M/D)$.

The above two can be combined into one, as a buffer conductor.

Lemma 4.3 (nonconstructive buffer conductors) *For every $n, m, b \leq n - m, \epsilon > 0$, there is a $(k_{max}, \varepsilon, a)$ buffer conductor* $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \rightarrow (m) \times (b)$ *with*

- $d = \log(n - (m + b) + 1) + 2\log(1/\varepsilon) + O(1)$, *and*
- $a = d - 2\log(1/\varepsilon) + O(1)$
- $k_{max} = m + b - d - \log(1/\varepsilon) - O(1)$

We now describe several *explicit* conductors implied by a number of known constructions. The first two are well-known, based on expanders with bounded eigenvalue and almost-pairwise independent hashing, respectively. The analysis merely involves converting the guarantees on Renyi entropy directly provided by these objects into ε-closeness to min-entropy (via Lemma A.3). For the case of extraction (output min-entropy equals output length), this kind of analysis was done for expanders in [GW97] and for hashing in [GW97, SZ98] (generalizing Rackoff's proof of the Leftover Hash Lemma [HILL99, IZ89]). What follows is a generalization to lower min-entropies.

Any constant-degree expander on (n) with bounded second eigenvalue yields a conductor which uses the d random bits to do a random walk on the graph. Roughly speaking, each step adds $\Omega(1)$ bits of entropy, so this gives $a = \Omega(d)$. We get a permutation conductor, by letting the buffer "remember" the sequence of edges taken (equivalently, take the *rotation map* of graph in the sense of [RVW00]). This gives:

Lemma 4.4 (eigenvalue-based conductors) *For every n, $a \leq n$, and $\varepsilon > 0$, there is an explicit (ε, a)* permutation conductor $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \rightarrow (n) \times (d)$ *with* $d = O(a + \log(1/\varepsilon))$.

The key feature of the above conductors is that d does not depend on n.

Using existing constructions of almost pairwise independent families of hash functions [NN93, AGHP92], we also obtain interesting conductors (both extracting and lossless). Here the d truly random bits are used to select a hash function h from the family, and $\mathrm{E}(x, h) = (h, h(x))$. Actually, we divide $h(x)$ into two pieces, using one as part of the extracting conductor, and one to create a buffer.

Lemma 4.5 (conductors from hashing) *For every n, m, b, and $\varepsilon > 0$, there is an explicit $(k_{max}, \varepsilon, a)$ buffer conductor* $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \rightarrow (d + m) \times (b)$ *with* $d = O(\log n + m + b + \log(1/\varepsilon))$, $a = d - 2\log(1/\varepsilon) - 1$, *and* $k_{max} = m + b - \log(1/\varepsilon) - 2$.

The above conductors are optimal (up to additive constants) in terms of a and k_{max}, but d is quite large, growing linearly with m and b. To improve on d, we obtain explicit conductors come from the extractor literature, in particular the recent series of works building on the construction of Trevisan [Tre99]. For the first ones, we exploit the fact that the extractors of [RRV99] have the property that for any $m_1 \leq m_2$ the extractor for min-entropy m_1 can be obtained by taking a subset of the output bits of the extractor for min-entropy m_2. Hence, we have guarantees on the output even when the min-entropy is smaller than needed for a uniform output.

Lemma 4.6 ([Tre99, RRV99]) *For any n, m, $\varepsilon > 0$, there exist explicit (ε, a) extracting conductor* $\mathrm{E} : (n) \times (d) \rightarrow (m)$, *with* $d = O(\log^2(n/\varepsilon) \cdot \log m)$ *and* $a = -O(d)$.

Proof Sketch: In [RRV99, Thm. 22], (k, ε) *extractors* $\mathrm{E} : (n) \times (d) \rightarrow (m)$ with $d = O(\log^2(n/\varepsilon) \cdot \log m)$ and $k = m + O(d)$ are given. For $m_1 \leq m_2 \leq n$, this construction has the property that the extractor for output length m_1 can be obtained by taking a subset of the output bits of the extractor for output length m_2.[5] Thus, if we take the extractor $\mathrm{E} : (n) \times (d) \rightarrow (m)$ for output length m then for every $k \leq m - a$, every k-source X gets transformed into a $(k + a, \varepsilon)$-source for $a = -O(d)$. □

The next ones are the “lossless condensers” of Ta-Shma, Umans, and Zuckerman.

Lemma 4.7 ([TUZ01]) *For any n, $k_{max} \leq n$, and $\varepsilon > 0$, there is an explicit (k_{max}, ε) lossless conductor* $\mathrm{E} : (n) \times (d) \rightarrow (m)$ *with* $d = O(\log n)$, $m = (k_{max}/\varepsilon)^2$.

In [RRV99, TUZ01], there are also other constructions of conductors for different settings of the parameters, but we won't need them here. Also, other recent extractor constructions, such as those in [TZS01, SU01], have the property that prefixes of the output correspond to the same extractor construction for lower min-entropies. This implies that they give rise to interesting conductors, which are better than the above ones for some settings of parameters (e.g. when one is willing to lose more than a constant fraction of the input min-entropy).

[5] This “restriction property” is stated explicitly in [RRV99, Remk.] for their extractors which extract up to a constant fraction of the min-entropy; for those, even an appropriate *prefix* of the output bits are enough. The extractors used here, which extract nearly all the min-entropy, are obtained by repeatedly applying the extractors which extract half of the min-entropy, and thereby inherit the desired property (though not necessarily for a prefix).

4.1 Simple Compositions

In this section we describe two simple but useful methods for composing conductors. The first is for composing any two simple conductors with each other, which generalizes the popular "repeated condensing" and "condense-then-extract" paradigms used in the recent extractor literature [RSW00, TUZ01].

Lemma 4.8 *Let* $\mathrm{E}_1 : (n_1) \times (d_1) \to (m_1)$ *be a* $(k_1, \varepsilon_1, a_1)$ *simple conductor, and let* $\mathrm{E}_2 : (m_1) \times (d_2) \to (m_2)$ *be a* $(k_1 + a_1, \varepsilon_2, a_2)$ *simple conductor, and define* $\mathrm{E} : (n_1) \times (d_1 + d_2) \to (m_2)$ *by* $\mathrm{E}(x, (y_1, y_2)) = \mathrm{E}_2(\mathrm{E}_1(x, y_1), y_2)$*. Then* E *is an* $(k_1, \varepsilon_1 + \varepsilon_2, a_1 + a_2)$ *simple conductor.*

Taking E_1 to be the lossless conductor of Lemma 4.7 and E_2 to be the extracting conductor of Lemma 4.6, we obtain:

Lemma 4.9 *For every* n*,* m*, and* $\varepsilon > 0$*, there is an* (ε, a) *extracting conductor* $\mathrm{E} : (n) \times (d) \to (m)$ *with* $d = O(\log n + \log^3(m/\varepsilon))$ *and* $a = -O(\log^3(m/\varepsilon))$*.*

The next composition shows that we can "improve" an buffer conductor by applying another conductor to just the buffer.

Lemma 4.10 *Let* $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle : (n) \times (d_1) \to (m_1) \times (b_1)$ *be a* $(k_1, \varepsilon_1, a_1)$ *buffer conductor and let* $\mathrm{E}_2 : (b_1) \times (d_2) \to (m_2)$ *be a* $(k_2, \varepsilon_2, a_2)$ *simple conductor, with* $k_2 = \min\{k_1 + d_1 - m_1, d_1 - a_1\}$*. Define* $\mathrm{E} : (n) \times (d_1 + d_2) \to (m_1 + m_2)$ *by*

$$\mathrm{E}(x, (y_1, y_2)) = \mathrm{E}_1(x, y_1) \circ \mathrm{E}_2(\mathrm{C}_1(x, y_1), y_2)$$

Then E *is a* $(k_1, 2\varepsilon_1 + \varepsilon_2, d_1 + a_2)$ *simple conductor.*

Proof: Let X be any k-source on (n_1) with $k \le k_1$, and let $(Y, Z) = \langle \mathrm{E}_1, \mathrm{C}_1 \rangle (X, U_d)$. Since $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$ is a $(k_1, \varepsilon_1, a_1)$ buffer conductor, it follows that (Y', Z') is $2\varepsilon_1$-close to a distribution (Y', Z') where Y' is a $\min\{m_1, k + a_1\}$-source and (Y', Z') is a $k + d_1$-source.

We now analyze the effect of E_2 on Z'. Since (Y', Z') is a $k + d_1$-source, for any y in the support of Y', $(Z'|Y' = y)$ is a k_y-source, for

$$\begin{aligned} k_y &= k + d_1 - \log_2 \frac{1}{\Pr[Y' = y]} \\ &\le k + d_1 - \min\{m_1, k + a_1\} \\ &\le k_2. \end{aligned}$$

This implies that $\mathrm{E}_2((Z'|Y' = y), U_{d_2})$ is a $(k_y + a_2, \varepsilon_2)$-source, which implies that $Y' \circ \mathrm{E}_2(Z', U_{d_2})$ is a $(k + d_1 + a_2, \varepsilon_2)$-source. Recalling that (Y', Z') is $2\varepsilon_1$-close to (Y, Z), we conclude that $\mathrm{E}(X, U_{d_1 + d_2}) = Y \circ \mathrm{E}_2(Z, U_{d_2})$ is a $(k + d_1 + a_2, 2\varepsilon_1 + \varepsilon_2)$-source, as desired. ■

Now we show how to use the above composition to construct buffer conductors from extracting conductors. First, note that a trivial way to extend an extracting conductor $\mathrm{E} : (n) \times (d) \to (m)$ to an buffer conductor $\langle \mathrm{E}, \mathrm{C} \rangle$ is to define $\mathrm{C}(x, y) = (x, y)$. Of course this creates a very large buffer. The above lemma shows that we can shrink the size of this buffer by applying a lossless conductor to it. If we take the lossless conductor to be one based on hashing from Corollary B.5, we obtain exactly the Raz–Reingold method for converting extractors into "extractor-condenser pairs" [RR99]. The above lemma is more general, however, and also shows that if we apply an extracting conductor to the buffer, we get an extracting conductor. It thus also generalizes the method of Wigderson and Zuckerman [WZ99] for decreasing the entropy loss of an extractor.

The two cases (lossless and extracting) can be combined to yield the following composition for buffer conductors (which is proven by appealing to the previous lemma twice):

Lemma 4.11 *Let* $\langle \mathrm{E}_1, \mathrm{C}_1\rangle : (n) \times (d_1) \to (m_1) \times (b_1)$ *be a* $(k_1, \varepsilon_1, a_1)$ *buffer conductor and let* $\langle \mathrm{E}_2, \mathrm{C}_2\rangle : (b_1) \times (d_2) \to (m_2) \times (b_2)$ *be a* $(k_2, \varepsilon_2, a_2)$ *buffer conductor, with* $k_2 = k_1 + d_1 - m_1$, $m_2 = (d_1 - a_1) + a_2$. *Define* $\langle \mathrm{E}, \mathrm{C}\rangle : (n) \times (d_1 + d_2) \to (m_1 + m_2) \times (b_2)$ *by*

$$\begin{aligned} \mathrm{E}(x, (y_1, y_2)) &= \mathrm{E}_1(x, y_1) \circ \mathrm{E}_2(\mathrm{C}_1(x, y_1), y_2), \text{ and} \\ \mathrm{C}(x, (y_1, y_2)) &= \mathrm{C}_2(\mathrm{C}_1(x, y_1), y_2) \end{aligned}$$

Then $\langle \mathrm{E}, \mathrm{C}\rangle$ *is a* $(k_1, 2\varepsilon_1 + \varepsilon_2, d_1 + a_2)$ *buffer conductor.*

Using this composition to combine the conductors of Lemmas 4.9 and 4.5, we get:

Lemma 4.12 *For any* n, m *and* $\varepsilon > 0$, *there exists an explicit* $(m - a, \varepsilon, a)$ *buffer conductor* $\langle \mathrm{E}, \mathrm{C}\rangle : (n) \times (d) \to (m) \times (b)$ *with* $d = O(\log n + \log^3(m/\varepsilon))$, $a = d - 2\log(1/\varepsilon) - O(1)$, *and* $b = 3\log(1/\varepsilon) + O(1)$.

Proof Sketch: Consider the conductors of Lemma 4.9. They have the desired seed length d, but have a small, but suboptimal entropy loss $d - a = O(d)$, and no buffer. We will take the trivial buffer, i.e. the entire input and seed, and then apply the hashing-based conductors of Lemma 4.5 to the reduce the entropy loss to optimal and also obtain a very short buffer. Since we are only using the hashing-based conductor to recover up to $O(d)$ bits of unextracted randomness, we only need to take its output length to be $O(d) + \log(1/\varepsilon)$, and hence it only costs us $O(\log n + d + \log(1/\varepsilon))$ bits of additional seed length. A more detailed proof is given in Appendix B □

Finally, combining these conductors with themselves via Lemma 4.10, we get:

Lemma 4.13 *For any* $k_{max} \leq n$, m, *and* $\varepsilon > 0$, *there exists an explicit* $(k_{max}, \varepsilon, a)$ *buffer conductor* $\langle \mathrm{E}, \mathrm{C}\rangle : (n) \times (d) \to (m) \times (b)$ *with* $d = O(\log n + \log^3(m/\varepsilon) + \log^3(k_{max}/\varepsilon))$, $a = d - 2\log(1/\varepsilon) - O(1)$, *and* $m + b = k_{max} + d + \log(1/\varepsilon) + O(1)$.

5 The Zig-Zag Product for Conductors

In this section we show how to compose conductors via the zig-zag product of [RVW00]. When applied to the conductors described in Section 4, this composition will imply constant-degree, lossless expanders (details appear in Section 7).

The analysis of [RVW00]. Let us first briefly recall the intuition of the zig-zag product for expanders from [RVW00]. This intuition relies on the view of expanders as graphs that "increase entropy". This roughly means that a random step on an expander, starting from a distribution X on the vertices, arrives at a distribution X' with "higher entropy" (as long as X did not contain "too much" entropy to begin with). The analysis in [RVW00] as in most previous constructions of expanders interprets "entropy" as Renyi's H_2-entropy. In this section we analyze zig-zag with respect to a combination of L_1 distance and min-entropy (which indeed gives us much more flexibility). Nevertheless, the intuition in both cases can be described using a very abstract notion of entropy.

Let $N = 2^{n_1}$, $D_1 = 2^{d_1}$, and $D_2 = 2^{d_2}$. Let $\mathrm{E}_1 : (n_1) \times (d_1) \to (n_1)$ be the neighbor function of a D_1-regular expander graph G_1 on N_1 vertices and let $\mathrm{E}_2 : (d_1) \times (d_2) \to (d_1)$ be the neighbor function of a D_2-regular expander graph G_2 on D_1 vertices. For simplicity, let us assume that for every $x_1 \in (n_1)$, $r_1 \in (d_1)$ the function $\mathrm{E}_1(\mathrm{E}_1(x_1, r_1), r_1) = x_1$ and similarly for E_2 (i.e., every edge has the same label when viewed from either of its endpoints). The zig-zag product of G_1 and G_2 is a $(D_2)^2$-regular expander graph $G = G_1 ⓩ G_2$ on $N_1 \cdot D_1$ vertices. The neighbor function E of G is defined as follows: For any $x_1 \in (n_1)$, $x_2 \in (d_1)$,$r_2 \in (d_2)$ and $r_3 \in (d_2)$ define

$$\mathrm{E}(x_1 \circ x_2, r_2 \circ r_3) \stackrel{\mathrm{def}}{=} y_1 \circ y_2, \text{ where}$$

$$\begin{aligned} r_1 &\stackrel{\mathrm{def}}{=} \mathrm{E}_2(x_2, r_2), \\ y_1 &\stackrel{\mathrm{def}}{=} \mathrm{E}_1(x_1, r_1), \text{ and} \\ y_2 &\stackrel{\mathrm{def}}{=} \mathrm{E}_2(r_1, r_3). \end{aligned}$$

Note that a random step on G consists of a (random) step on G_2 followed by a (deterministic) step on G_1 and finally another (random) step on G_2.

Assume now that G_1 and G_2 are both expander graphs and that for $i = 1, 2$, a random step on G_i "adds a_i bits of entropy". Further assume that $D_1 \ll N_1$ (hence we will refer to G_1 as the "large graph" and to G_2 as the "small graph"). The analysis of [RVW00] shows that G is also an expander and more specifically that a random step on G "adds $a = \min\{a_1, a_2\}$ bits of entropy":

Consider a random step starting at a distribution $X = (X_1, X_2)$ on the vertices of X (that is missing at least a bits of entropy). It can be shown that it is sufficient to consider two extreme cases, based on the conditional distributions of X_2 induced by particular assignments $X_1 = x_1$ (where x_1 is in the support of X_1). In the first case, all of these conditional distributions of X_2 are far from uniform (i.e., missing at least a bits of entropy). In this case, the first step on G_2 (in the evaluation $r_1 = \mathrm{E}_2(x_2, r_2)$) will add a bits of entropy (taken from the randomness in r_2). It is easy to verify that the next two steps preserve this entropy. In the second case, the conditional distributions of X_2 are all uniform. In this case, the first step on G_2 is useless. However, now the second step ($y_1 = \mathrm{E}_1(x_1, r_1)$) is in fact a random step on G_1. This step shifts a bits of entropy from r_1 to x_1. Finally, the last step ($y_2 = \mathrm{E}_2(r_1, r_3)$) on G_1 adds (the now missing) a bits of entropy to r_1 (taken from the fresh randomness in r_3).

The new zig-zag product. The zig-zag product discussed above combines a large graph with a small graph, the resulting graph inherits (roughly) its size from the large one, its degree from the small one, and its expansion properties from both. Iteration of this powerful product was used in [RVW00] to construct constant-degree expanders with an elementary analysis. However, these expanders have the disadvantage that their expansion is suboptimal (as a function of their degree). This deficiency can easily be traced back to the expander composition itself: Although the degree of G is quadratic in the degree of G_2, the expansion of G is at most that of G_2. Indeed, the analysis sketched above only guarantees that one of the random steps on G_2 adds entropy. The zig-zag theorem for conductors presented here manages to avoid exactly this problem. For that we use a different variant of the zig-zag product that applies to (possibly unbalanced) bipartite graphs (essentially the same variant was used in [RVW00] to construct extractors for high min-entropy). The main differences are: we will augment the first application of E_2 so that we also obtain a *buffer* which will retain any entropy that would have otherwise been lost in the first step, and we will replace the application of E_2 in step 3 with an application of a conductor E_3 to both buffers from the earlier steps to carry all of the remaining entropy to the output. (In the description above, r_1 plays the role of the buffer in the application of E_1, but the product below will be more general.) This generalization forces us to work with unbalanced graphs, and in a sense, the advantage of the new zig-zag theorem is in the ability to perform the "expander analysis" in the setting of unbalanced graphs.

Definition 5.1 (zig-zag product [RVW00]) *Let* $\langle \mathrm{E}_1, \mathrm{C}_1\rangle : (n_1) \times (d_1) \mapsto (m_1) \times (b_1)$, $\langle \mathrm{E}_2, \mathrm{C}_2\rangle : (n_2) \times (d_2) \mapsto (d_1) \times (b_2)$, *and* $\mathrm{E}_3 : (b_1 + b_2) \times (d_3) \mapsto (m_3)$ *be three functions. Set the parameters*

$$\begin{aligned} n &= n_1 + n_2, \\ d &= d_2 + d_3, \\ m &= m_1 + m_3 \end{aligned}$$

and define the **zig-zag product**

$$\mathrm{E} : (n) \times (d) \mapsto (m)$$

of these functions as follows: For any $x_1 \in (n_1)$, $x_2 \in (n_2)$,$r_2 \in (d_2)$ *and* $r_3 \in (d_3)$ *define*

$$\mathrm{E}(x_1 \circ x_2, r_2 \circ r_3) \stackrel{\mathrm{def}}{=} y_1 \circ y_2, \text{ where}$$

$$\begin{aligned} \langle r_1, z_1\rangle &\stackrel{\mathrm{def}}{=} \langle \mathrm{E}_2, \mathrm{C}_2\rangle(x_2, r_2) \\ \langle y_1, z_2\rangle &\stackrel{\mathrm{def}}{=} \langle \mathrm{E}_1, \mathrm{C}_1\rangle(x_1, r_1), \text{ and} \\ y_2 &\stackrel{\mathrm{def}}{=} \mathrm{E}_3(z_1 \circ z_2, r_3). \end{aligned}$$

In Section 6 we analyze the way the zig-zag product operates on conductors. We consider a very wide setting of the parameters, and show how the zig-zag product can produce both extracting conductors and lossless conductors (in fact, it can produce conductors that are simultaneously extracting and, for different parameters, lossless). The remaining of this section however, discusses a particular, simplified example that demonstrates the operation of the zig-zag product (and in particular, how this product can imply constant-degree lossless expanders).

Since this example is solely for demonstrational purposes, its parameters are quite inferior to those we actually obtain. Let e be some fixed large constant multiple of $\log^3(1/\epsilon)$. The graphs we use in the composition are the following:

- The "big graph" $\langle \mathrm{E}_1, \mathrm{C}_1\rangle : (n_1) \times (100e) \mapsto (n_1) \times (100e)$, is an $(\varepsilon, 5e)$ permutation conductor (that can be taken from Lemma 4.4).

- The first small graph $\langle \mathrm{E}_2, \mathrm{C}_2\rangle : (106e) \times (e) \mapsto (100e) \times (8e)$, is an $(\varepsilon, 0)$ extracting conductor and $\mathrm{E}_3 : (108e) \times (e) \mapsto (104e)$ is a $(102e, \varepsilon)$-lossless conductor. (Both $\langle \mathrm{E}_2, \mathrm{C}_2\rangle$ and E_3 can be taken from Lemma 4.13.)

Let $\mathrm{E} : (n_1 + 106e) \times (2e) \mapsto (n_1 + 104e)$ be the zig-zag product of these functions. The zig-zag theorem for conductors implies that E is an $(n_1 + 100e, O(\varepsilon))$-lossless conductor (as a "bonus", E is also unbalanced).

A rather good intuition for why this is true can be obtained by a simple (though informal) "bookkeeping". As with the description of the expander composition earlier, we try to follow the "entropy flow" from the input $(X_1 \circ X_2, R_2 \circ R_3)$ to the output $Y_1 \circ Y_2$ through the computation of E. Where $X_1 \circ X_2$ is a k-source for some $k \leq n_1 + 100k$, and R_2, R_3 are both uniform over (e). The intermediate steps in this computation are $(R_1, Z_1) = \mathrm{E}_2(X_2; R_2)$, $(Y_1, Z_2) = (X_1, R_1)$, and $Y_2 = \mathrm{E}_3(Z_1 \circ Z_2, R_3)$. The output is $Y_1 \circ Y_2$.

As in the expander analysis, it is sufficient to consider two extreme cases, based on the conditional distributions of X_2 induced by particular assignments $X_1 = x_1$ (where x_1 is in the support of X_1):

- Case I: For every x_1 in the support of X_1, there are less than $100e$ bits of entropy in $X_2 |\ X_1 = x_1$.

- Case II: For every x_1 in the support of X_1, there are at least than $100e$ bits of entropy in $X_2 |\ X_1 = x_1$.

In the first case, applying E_2 squeezes the entropy of X_2 into R_1 (the progress we have made is condensing the source by $6e$ bits). Therefore, $X_1 \circ R_1$ contains some $k' \geq k$ bits of entropy and Z_1 contains the remaining $k + e - k' \leq e$ bits. Since $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$ is a permutation, (Y_1, Z_2) still contains k' bits of entropy, from which at most $100e$ are in Z_2 (since it is $100e$ long) and the rest are in Y_1. We can conclude that Y_1 contains some $k'' \geq k - 100e$ bits of entropy and $Z_1 \circ Z_2$ contains the remaining $k + e - k'' \leq 101e$ bits. Finally, applying E_3 squeezes all of the entropy from $Z_1 \circ Z_2$ and the additional e bits from R_3 (coming to a total of $k + 2e - k''$) into Y_2. We can therefore conclude that, $Y_1 \circ Y_2$ contains the desired $k + 2e$ bits.

In the second case, R_1 is close to uniform (even conditioned on X_1). Furthermore, X_1 contains at least $k - 106e$ bits of entropy (since X_2 may contain at most $106e$ bits of entropy). Since $k \leq n_1 + 100e$, we have that $k - 106e \leq n_1 - 6e$. Therefore, the application of E_1 pushes $5e$ bits of entropy from R_1 to Y_1 (in addition to the $k - 106e$ bits from X_1). We can deduce that Y_1 contains some $k' \geq k - 101e$ bits of entropy. It is easy to see that the remaining $k + e - k' \leq 102e$ bits are in $Z_1 \circ Z_2$. Therefore, as before, applying E_3 squeezes all of the entropy from $Z_1 \circ Z_2$ and the additional e bits from R_3 (coming to a total of $k + 2e - k'$) into Y_2. In this case too, $Y_1 \circ Y_2$ contains the desired $k + 2e$ bits of entropy.

6 The Zig-Zag Theorem

In this section, we state the general zig-zag theorem for conductors. It treats a much more general setting of parameters than the one needed for our constructions. We try to clarify the interaction between parameters below, but still it may help to think of the components of the composition as follows (which is what we will use to obtain our main results):

- The "big graph" $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$ will be the permutation conductor of Lemma 4.4, $m_1 = n_1$, $b_1 = d_1$, and d_1 will be taken to be a sufficiently large constant,
- The "small graphs" $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle$ and E_3 will be optimal "constant-size" nonconstructive conductors from Lemmas 4.1, 4.2, 4.3 or explicit ones from Lemma 4.13, so $d_2 = \mathrm{polylog} n_2$, $d_3 = \mathrm{polylog}(b_1 + b_2)$.

Theorem 6.1 *Let* $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle : (n_1) \times (d_1) \mapsto (m_1) \times (b_1)$ *be an* (ε, a_1) *permutation conductor. Let* $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle : (n_2) \times (d_2) \mapsto (d_1) \times (b_2)$ *be an* (n_2, ε, a_2) *buffer conductor. Let* $\mathrm{E}_3 : (b_1 + b_2) \times (d_3) \mapsto (m_3)$ *be an* (ε, a_3) *extracting conductor.*

Let $\mathrm{E} : (n) \times (d) \mapsto (m)$ *be the zig-zag product of* $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$, $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle$ *and* E_3 *and set*

$$a = \min\{d_2 + a_3, a_1 - (n_2 - m_3) - \log 1/\varepsilon, m_3 + a_2 - d_1 - (n_1 - m_1) - \log 1/\varepsilon\}.$$

Then E *is an* $(5\varepsilon, a)$ *extracting conductor.*

In Theorem 6.1, all of the conductors E_1, E_2 and E_3 are *extracting* conductors. This is necessary if we want E to also be an extracting conductor. However, the composition still gives meaningful results even if the only extracting conductor is E_2 (this conductor must be extracting so that when X_2 has large entropy R_1 will indeed be close to uniform and the application of E_1 useful). A particularly useful case is when E_3 is a lossless conductor, as then we can obtain a lossless conductor (which is how we will beat the degree $/2$ barrier).

Theorem 6.2 *Let* $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle : (n_1) \times (d_1) \mapsto (m_1) \times (b_1)$ *be an* (ε, a_1) *permutation conductor. Let* $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle : (n_2) \times (d_2) \mapsto (d_1) \times (b_2)$ *be an* (n_2, ε, a_2) *buffer conductor. Let* $\mathrm{E}_3 : (b_1 + b_2) \times (d_3) \mapsto (m_3)$ *be an* $(m_3 - a_3', \varepsilon)$ *lossless conductor.*

Let $\mathrm{E} : (n) \times (d) \mapsto (m)$ *be the zig-zag product of* $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$, $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle$ *and* E_3 *. If the following conditions hold:*

- $a_1 \geq d_2 + a_3' + (n_2 - m_3) + \log 1/\varepsilon$.
- $m_3 \geq d_1 + (n_1 - m_1) + (d_2 - a_2) + a_3' + \log 1/\varepsilon$.

Then E *is also a* $(k_{max}', 5\varepsilon)$ *lossless conductor, for* $k_{max}' = m - a_3' - d_2$.

Interpretation. As discussed above, the goal of the conductor composition is to avoid the inherent (factor of 2) entropy loss of the expander composition. For *any* distribution on the vertices $X = X_1 \circ X_2$ (of min-entropy $k \leq m - a$), we want the output $Y_1 \circ Y_2 = \mathrm{E}(X_1 \circ X_2, R_2 \circ R_3)$ to gain entropy from *both parts of the random input* $R_2 \circ R_3$ (the "edge label"). In fact, in some settings of the parameters (which will be ones we use), Theorem 6.1 implies that the entropy in $Y_1 \circ Y_2$ is $k + d_2 + a_3$. That is, we gain all of the entropy in R_2 and all of the entropy that E$_3$ is capable of adding. Furthermore, if E$_3$ is a lossless conductor then Theorem 6.2 will imply that E is also lossless. An additional useful feature of the product is that the output length m of E may be shorter than its input length n (which is naturally more difficult for achieving losslessness).

Under which conditions will the resulting conductor E in Theorems 6.1 and 6.2 be able to add the desired $d_2 + a_3$ (or $d_2 + a_3'$) bits of entropy? First, we will need that the first part of the output (Y_1) together with the two buffers (Z_1, Z_2), will contain all of the entropy in the system so far. That is, they contain k bits of entropy from the source plus the d_2 bits of R_2. This will follow easily from the losslessness of $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$ and $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle$. The next condition is nontrivial. We need Y_1 to contain enough entropy so that the conditional entropy left in the buffers (Z_1, Z_2) will be less that $m_3 - a_3$. In such a case, E$_3$ will manage to condense into Y_2 all of the entropy from the buffers plus a_3 additional bits. As our analysis will show, this condition can be translated into two (more concrete) conditions:

- When R_1 is close to uniform, the conductor E$_1$ must "push" enough entropy into Y_1. More specifically we need that:
$$a_1 \geq d_2 + a_3 + (n_2 - m_3) + \log 1/\varepsilon.$$
- The length of m_3 must be large enough to contain the entropy that may remain in Z_1, Z_2 and the additional a_3 bits from R_3. More specifically we need that:
$$m_3 \geq d_1 + (n_1 - m_1) + (d_2 - a_2) + a_3 + \log 1/\varepsilon$$

Under these conditions, we indeed have $a = d_2 + a_3$. Otherwise, the application of E$_3$ loses some entropy as can be seen in the (somewhat complex) definition of a. Note, that under almost the same conditions, the application of E$_3$ in the setting of Theorem 6.2 is indeed lossless (which allows E to be lossless as well).

Proof Intuition. A rather good intuition for the operation of the zig-zag product on conductors can be obtained by a simple (though informal) "bookkeeping." In this we generalize the intuition for the concrete example which appears in Section 5 (which follows the intuition of the zig-zag product for expanders). As before, we try to follow the "entropy flow" from the input $(X_1 \circ X_2, R_2 \circ R_3)$ to the output $Y_1 \circ Y_2$ through the computation of E. For most of the intuition we concentrate on the proof of Theorem 6.1 (the proof of Theorem 6.2 only differs in the last step of the argument).

As in the expander analysis, we can first show that it is sufficient to consider two extreme cases, based on the conditional distributions of X_2 induced by particular assignments $X_1 = x_1$ (where x_1 is in the support of X_1):

- Case I: For every x_1 in the support of X_1, the distribution $X_2 | X_1 = x_1$ contains less than $d_1 - a_2$ bits of entropy.

- Case II: For every x_1 in the support of X_1, the distribution $X_2|\, X_1 = x_1$ contains at least $d_1 - a_2$ bits of entropy.

In the first case, applying E_2 squeezes the entropy of X_2 plus a_2 bits of entropy from R_2 into R_1. It is important to note that we have made progress in two ways: adding entropy from R_2 and condensing the source by $n_2 - d_1$ bits. Since $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$ is a permutation, the entropy in $Y_1 \circ Z_2$ equals the entropy in $X_1 \circ R_1$, which is $k + a_2$ as argued above. Since Z_2 is of length $b_1 = d_1 + n_1 - m_1$ it may contain at most this amount of entropy. We can therefore conclude that Y_1 contains at least $k + a_2 - d_1 - (n_1 - m_1)$ bits of entropy.

In the second case, R_1 is close to uniform (even conditioned on X_1). Furthermore, X_1 contains at least $k - n_2$ bits of entropy (since X_2 may contain at most n_2 bits of entropy). From the properties of E_1 we can deduce that Y_1 contains at least $(\min\{m_1, k - n_2 + a_1\})$ bits of entropy.

We can conclude from this case analysis that Y_1 contains at least $(\min\{m_1, k - n_2 + a_1, k + a_2 - d_1 - (n_1 - m_1)\})$ bits of entropy. In addition, Y_1, Z_1, Z_2 contain together $k + d_2$ bits of entropy, since $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle$ is a buffer conductor and $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle$ is a permutation conductor. Furthermore, as long as the entropy in Z_1, Z_2 (conditioned on Y_1) is less than $m_3 - a_3$ then Y_2 will contain all the entropy in Z_1, Z_2 plus a_3 bits of entropy from R_3. Otherwise, Y_2 is close to uniform (i.e., has m_3 bits of entropy). We can therefore conclude that Y_1, Y_2 has at least $(\min\{k + d_2 + a_3, m = m_3 + m_1, m_3 + k - n_2 + a_1, m_3 + k + a_2 - d_1 - (n_1 - m_1)\})$ bits of entropy.

For the proof of Theorem 6.2, we must be guaranteed that the conditional entropy in Z_1, Z_2 (given Y_1) has at most $m_3 - a'_3$ bits of entropy. It is not hard to see that if $k \leq k'_{max}$ then under the conditions stated in the theorem this is indeed the case.

To carry on the manipulations of (min) entropy that are required by the informal arguments above, we use known techniques from the extractor literature. In particular, the arguments we use are strongly influenced by the notion of block sources [CG88] and its usefulness for randomness extraction [NZ96] (see more details in [RVW00]).

Proof Sketch: (of Theorem 6.1 and Theorem 6.2) We combine the proofs of the two theorems since they only diverge at the final step of the argument. For all possible values $x \in (n)$, $r_2 \in (d_2)$ and $r_3 \in (d_3)$, the computation of $\mathrm{E}(x, r_2 \circ r_3)$ produces the following intermediate values: $x_1, x_2, r_1, z_1, y_1, z_2$, and y_2. Let $X_1, X_2, R_1, Z_1, Y_1, Z_2$, and Y_2 be the corresponding random variables in the computation $\mathrm{E}(X, R_2 \circ R_3)$, where X is some k-source over (n) with $k \leq m - a$, R_2 is uniformly distributed in (d_2) and R_3 is uniformly distributed in (d_3). Our goal is to prove that (Y_1, Y_2) is a $(k + a, 5\varepsilon)$-source. The heart of this proof is showing that Y_1 contains enough entropy.

Lemma 6.3 *Y_1 is a $(k^1_{out}, 3\varepsilon)$-source, where*

$$k^1_{out} \overset{\mathrm{def}}{=} \min\{m_1, k - n_2 + a_1 - \log(1/\varepsilon), k + a_2 - d_1 - (n_1 - m_1) - \log(1/\varepsilon)\}.$$

Proof: (of Lemma 6.3) We first consider two special ("good") cases:

Case I: For every x_1 in the support of X_1, $\mathrm{H}_\infty(X_2|\, X_1 = x_1) < d_1 - a_2$.

Case II: For every x_1 in the support of X_1, $\mathrm{H}_\infty(X_2|\, X_1 = x_1) \geq d_1 - a_2$.

We then show that X can be essentially "decomposed" into a convex combination of these cases (at the price of losing $\log 1/\varepsilon$ bits of entropy). This will allow us to deduce the lemma.

In Case I:

Lemma 6.4 *Assume that X is a k' source (for some k') and that for every x_1 in the support of X_1, $\mathrm{H}_\infty(X_2|\ X_1 = x_1) < d_1 - a_2$. Then Y_1 is a $(k' + a_2 - d_1 - (n_1 - m_1), \varepsilon)$-source.*

Proof: For any x_1 in the support of X_1, define $k_{x_1} = \mathrm{H}_\infty(X_2|\ X_1 = x_1)$. Since for every such x_1 we have that $k_{x_1} < d_1 - a_2$, we can conclude from the properties of E_2 that $(R_1|\ X_1 = x_1) = \mathrm{E}_2((X_2|\ X_1 = x_1), R_2)$ is a $(k_{x_1}+a_2, \varepsilon)$-source. Since for every x_1 as above, we have that $k_{x_1} \geq k' - \log(1/\Pr[X_1 = x_1])$, it is not hard to argue that (X_1, R_1) is a $(k' + a_2, \varepsilon)$-source. Since $\langle \mathrm{E}_1, \mathrm{C}_1\rangle$ is a permutation we have that (Y_1, Z_2) is also a $(k' + a_2, \varepsilon)$-source.

Let (Y_1', Z_2') be a $(k' + a_2)$-source that is ε-close to (Y_1, Z_2). Since Z_2' is a distribution on $b_1 = d_1 + n_1 - m_1$ bit strings it follows that Y_1' is a $(k' + a_2 - d_1 - (n_1 - m_1))$-source and therefore Y_1 is a $(k' + a_2 - d_1 - (n_1 - m_1), \varepsilon)$-source. ■

In Case II:

Lemma 6.5 *Assume that X is a k' source (for some k') and that for every x_1 in the support of X_1, $\mathrm{H}_\infty(X_2|\ X_1 = x_1) \geq d_1 - a_2$. Then Y_1 is a $(\min\{m_1, k' - n_2 + a_1\}, 2\varepsilon)$-source.*

Proof: Since for every x_1 in the support of X_1, $\mathrm{H}_\infty(X_2|\ X_1 = x_1) \geq d_1 - a_2$, we can conclude from the properties of E_2 that for every such x_1, $(R_1|\ X_1 = x_1) = \mathrm{E}_2((X_2|\ X_1 = x_1), R_2)$ is ε-close to uniform. This implies that (X_1, R_1) is ε-close to (X_1, U_{d_1}) and thus that $Y_1 = \mathrm{E}_1(X_1, R_1)$ is ε-close to $\mathrm{E}_1(X_1, U_{d_1})$.

Since (X_1, X_2) is a k'-source and X_2 is distributed over n_2-bit strings, we can deduce that X_1 is a $(k' - n_2)$-source. Therefore, since E_1 is an (ε, a_1)-extractor we have that $\mathrm{E}_1(X_1, U_{d_1})$ is a $(\min\{m_1, k' - n_2 + a_1\}, \varepsilon)$-source. Therefore, Y_1 is a $(\min\{m_1, k' - n_2 + a_1\}, 2\varepsilon)$-source. ■

Decomposing X into the good cases: To complete the proof of Lemma 6.3, it remains to show how to decompose X into the two cases (without losing too much in entropy):

Lemma 6.6 *Let X be any k-source then X is ε-close to a convex combination $\alpha(X_1', X_2') + \beta(X_1'', X_2'')$ (with $\alpha + \beta = 1$) of two k'-sources, where*

1. *$k' = k - \log 1/\varepsilon$.*
2. *For every x_1 in the support of X_1', $\mathrm{H}_\infty(X_2'|\ X_1' = x_1) < d_1 - a_2$.*
3. *For every x_1 in the support of X_1'', $\mathrm{H}_\infty(X_2''|\ X_1'' = x_1) \geq d_1 - a_2$.*

Proof: Let B be the set of strings x_1 in the support of X_1', such that $\mathrm{H}_\infty(X_2'|\ X_1' = x_1) < d_1 - a_2$. Define the sources:

$$\begin{aligned}(X_1', X_2') &\stackrel{\text{def}}{=} (X_1, X_2)|\ X_1 \in \mathrm{B}\\ (X_1'', X_2'') &\stackrel{\text{def}}{=} (X_1, X_2)|\ X_1 \notin \mathrm{B}\end{aligned}$$

Finally, define $\alpha = 0$ if $\Pr[X_1 \in \mathrm{B}] < \varepsilon$, $\alpha = 1$ if $\Pr[X_1 \in \mathrm{B}] > 1 - \varepsilon$ and $\alpha = \Pr[X_1 \in \mathrm{B}]$ otherwise. Define $\beta = (1 - \alpha)$.

By these definitions, it is clear that conditions (2) and (3) above are satisfied. Furthermore, it follows easily that X is ε-close to $\alpha(X_1', X_2') + \beta(X_1'', X_2'')$. Finally, it is not hard to verify that in case $\alpha > 0$ then (X_1', X_2') is a k'-source and in case $\beta > 0$ then (X_1'', X_2'') is a k'-source. ■

From Lemma 6.4, Lemma 6.5 and Lemma 6.6 we can conclude that Y_1 is 3ε-close to a convex combination $\alpha\tilde{Y}_1' + \beta\tilde{Y}_1''$ (with $\alpha + \beta = 1$), where $\tilde{Y}_1'$ is a $(k + a_2 - d_1 - (n_1 - m_1) - \log(1/\varepsilon))$-source and $\tilde{Y}_1''$ is a $(\min\{m_1, k - n_2 + a_1\} - \log(1/\varepsilon))$-source. Therefore both $\tilde{Y}_1'$ and $\tilde{Y}_1''$ are k^1_{out} sources and so is their convex combination $\alpha Y_1' + \beta Y_1''$. This implies Lemma 6.3. ■

We now show that altogether (Y_1, Z_1, Z_2) contain all the entropy of (X, R_2):

Lemma 6.7 *(Y_1, Z_1, Z_2) is an $(k + d_2, \varepsilon)$-source.*

Proof: For any x_1 in the support of X_1, define $k_{x_1} = \mathrm{H}_\infty(X_2|\ X_1 = x_1)$. Since $\langle \mathrm{E}_2, \mathrm{C}_2\rangle$ is an (n_2, ε) lossless conductor, we have that for every such x_1, the conditional distribution $((R_1, Z_1)|\ X_1 = x_1) = \langle \mathrm{E}_2, \mathrm{C}_2\rangle((X_2|\ X_1 = x_1), R_2)$ is a $(k_{x_1} + d_2, \varepsilon)$-source. For every x_1 as above, we have $k_{x_1} \geq k - \log(1/\Pr[X_1 = x_1])$, and it follows that (X_1, R_1, Z_1) is a $(k + d_2, \varepsilon)$-source. Since $\langle \mathrm{E}_1, \mathrm{C}_1\rangle$ is a permutation we have that (Y_1, Z_2, Z_1) is also a $(k + d_2, \varepsilon)$-source. (Note: it is important that $\langle \mathrm{E}_1, \mathrm{C}_1\rangle$ is one-to-one and not just (n, ε)-lossless, because its seed (R_1) is not necessarily random.) ■

From Lemma 6.3 and Lemma 6.7 it is possible to deduce the following:

Lemma 6.8 *(Y_1, Z_1, Z_2) is 4ε-close to a source $(\tilde{Y}_1, \tilde{Z}_1, \tilde{Z}_2)$, where*

- *$\tilde{Y}_1$ is a k^1_{out} source, and*
- *$(\tilde{Y}_1, \tilde{Z}_1, \tilde{Z}_2)$ is a $k + d_2$ source.*

This is the point where the proofs of Theorem 6.1 and Theorem 6.2 diverge.

Completing the proof of Theorem 6.1. Here we assume that E_3 is an (ε, a_3) extracting conductor and $k \leq m - a$. Let us define $\tilde{Y}_2 = \mathrm{E}_3(\tilde{Z}_1 \circ \tilde{Z}_2, R_2)$. To conclude the theorem it is enough to prove that $(\tilde{Y}_1, \tilde{Y}_2)$ is a $(k + a, \varepsilon)$-source, which implies that (Y_1, Y_2) is a $(k + a, 5\varepsilon)$-source (since (Y_1, Y_2) and $(\tilde{Y}_1, \tilde{Y}_2)$ are 4ε-close).

For any y_1 in the support of $\tilde{Y}_1$, define $k_{y_1} = \mathrm{H}_\infty(\tilde{Z}_1, \tilde{Z}_2|\ \tilde{Y}_1 = y_1)$. Note that $k_{y_1} + \log(1/\Pr[\tilde{Y}_1 = y_1]) \geq k + d_2$. Furthermore, if $k_{y_1} < m_3 - a_3$ then $(\tilde{Y}_2|\ \tilde{Y}_1 = y_1)$ is a $(k_{y_1} + a_3, \varepsilon)$ source. Otherwise, $(\tilde{Y}_2|\ \tilde{Y}_1 = y_1)$ is ε-close to uniform (i.e., it is an (m_3, ε) source). We can conclude that $(\tilde{Y}_1, \tilde{Y}_2)$ is a $(\min\{k + d_2 + a_3, k^1_{out} + m_3\}, \varepsilon)$-source. Taking into account the definition of k^1_{out} (which is $\min\{m_1, k - n_2 + a_1 - \log(1/\varepsilon), k + a_2 - d_1 - (n_1 - m_1) - \log(1/\varepsilon)\}$) and the fact that $k \leq m - a$ we can conclude that $(\tilde{Y}_1, \tilde{Y}_2)$ is a $(k + a, \varepsilon)$-source as desired.

Completing the proof of Theorem 6.2. Here we assume that E_3 is an $(m_3 - a_3', \varepsilon)$ lossless conductor, $k \leq m - d_2 - a_3'$ and that the following conditions hold:

- $a_1 \geq d_2 + a_3' + (n_2 - m_3) + \log 1/\varepsilon$.
- $m_3 \geq d_1 + (n_1 - m_1) + (d_2 - a_2) + a_3' + \log 1/\varepsilon$.

Under these conditions we have that

$$\begin{aligned} k^1_{out} &= \min\{m_1, k - n_2 + a_1 - \log(1/\varepsilon), k + a_2 - d_1 - (n_1 - m_1) - \log(1/\varepsilon)\} \\ &\geq \min\{m_1, k - m_3 + d_2 + a_3'\} \end{aligned}$$

Since $k \leq m - d_2 - a_3' = m_1 + m_3 - d_2 - a_3'$ we can conclude that $\tilde{Y}_1$ is a $(k - m_3 + d_2 + a_3')$-source. The rest of the proof follows very similarly to the completion of the proof for Theorem 6.2.

□

7 Putting it Together

In this section, we apply the zig-zag product for conductors to the conductors in Section 4 to obtain our main results. In all cases, we will take $\langle \mathrm{E}_1, \mathrm{C}_1\rangle$ to be the permutation conductor obtained from Lemma 4.4 (i.e., the rotation map of a power of a constant-degree expander). By taking $\langle \mathrm{E}_2, \mathrm{C}_2\rangle$ to be an optimal buffer conductor (as in Lemma 4.3) and E_3 to be an optimal lossless conductor (as in Lemma 4.2), we get the following:

Theorem 7.1 *For every $n, t \le n, \varepsilon > 0$, there exists a (k_{max}, ε) lossless conductor $\mathrm{E} : (n) \times (d) \to (n-t)$ with*

- $d = O(\log(t+1) + \log(1/\varepsilon))$, *and*
- $k_{max} = (n-t) - d - \log(1/\varepsilon) - O(1)$,

Moreover, E *can be computed in time* $\mathrm{poly}(n, \log(1/\varepsilon))$ *given two appropriate conductors of size* $S = \mathrm{poly}(2^t, 1/eps)$, *which can be found probabilistically in time* $\mathrm{poly}(S)$ *or deterministically in time* $2^{\mathrm{poly}(S)}$.

Note that these parameters are optimal (matching Lemma 4.2) up to the constants hidden in the O-notation. In graph-theoretic terms, this lemma gives bipartite graphs with N vertices on the left, $M = N/T$ vertices on the right, of degree $D = \mathrm{poly}(\log T, 1/\varepsilon)$ with sets of size $K = \Omega(\varepsilon M/D)$ expanding by a factor $(1-\varepsilon)D$. The graphs can be computed efficiently provided t and $1/\varepsilon$ are small (e.g. in the constant-degree case).

Proof Theorem 7.1: Let $a_1 = t + c \cdot (\log(t+1) + \log(1/\varepsilon)$ and $n_2 = c \cdot a_1$, where c will be chosen later to be a sufficiently large constant. Let $n_1 = n - n_2$, and let $\langle \mathrm{E}_1, \mathrm{C}_1\rangle : (n_1) \times (d_1) \to (n_1) \times (d_1)$ be the $(\varepsilon/5, a_1)$ permutation conductor of Corollary 4.4, so

$$d_1 = O(a_1 + \log(1/\varepsilon)).$$

Let $\langle \mathrm{E}_2, \mathrm{C}_2\rangle : (n_2) \times (d_2) \to (d_1) \times (b_2)$ be the $(n_2, a_2, \varepsilon/5)$ buffer conductor of Lemma 4.3, so

$$\begin{aligned} d_2 &= \log n_2 + 2\log(1/\varepsilon) + O(1) = O(\log(t+1) + \log(1/\varepsilon) + \log c) \\ a_2 &= d_2 - 2\log(1/\varepsilon) - O(1) \\ b_2 &= n_2 + d_2 - d_1 + \log(1/\varepsilon) + 2 \end{aligned}$$

Let $m_3 = n_2 - t \ge 0$, and let $\mathrm{E}_3 : (d_1 + b_2) \times (d_3) \to (m_3)$ be the $(m_3 - a_3', \varepsilon/5)$ lossless conductor of Lemma 4.2, so

$$\begin{aligned} d_3 &= \log(d_1 + b_2) + \log(1/\varepsilon) + O(1) = O(\log(t+1) + \log(1/\varepsilon) + \log c) \\ a_3' &= d_3 + \log(1/\varepsilon) + O(1). \end{aligned}$$

Let $\mathrm{E} : (n) \times (d) \to (m)$ be the zig-zag product of these three conductors, so

$$\begin{aligned} d &= d_2 + d_3 = O(\log(t+1) + \log(1/\varepsilon) + \log c), \\ m &= n_1 + m_3 = n - t \end{aligned}$$

By Theorem 6.2, E is a (k_{max}', ε) lossless conductor for

$$k_{max}' = m - a_3' - d_2 = (n-t) - d - \log(1/\varepsilon) - O(1),$$

provided the following two conditions hold:

- $a_1 \geq d_2 + a'_3 + (n_2 - m_3) + \log 1/\varepsilon$.
- $m_3 \geq d_1 + (n_1 - n_1) + (d_2 - a_2) + a'_3 + \log 1/\varepsilon$.

In the first condition, all the terms on the right-hand side are $O(\log(t+1) + \log 1/\varepsilon + \log c)$, except $(n_2 - m_3) = t$. The left-hand side is $a_1 = t + c \cdot (\log(t+1) + \log(1/\varepsilon))$, so it is satisfied for a sufficiently large constant c. In the second condition, all the terms on the right-hand side are $O(\log(t+1) + \log 1/\varepsilon + \log c)$, except $d_1 = O(a_1 + \log(1/\varepsilon))$. The left-hand side is $m_3 = n_2 - t = c \cdot a_1 - t$. Recalling that $a_1 > t$, the second condition also holds for a sufficiently large constant c. ■

In a similar fashion, we get an extracting conductor using Theorem 6.1, taking E_3 to be the optimal extracting conductor of Lemma 4.1.

Theorem 7.2 *For every n, $t \leq n$, $\varepsilon > 0$, there exists an (ε, a) extracting conductor* $\mathrm{E} : (n) \times (d) \rightarrow (n-t)$ *with*

- $d = O(\log(t+1) + \log(1/\varepsilon))$*, and*
- $a = d - 2\log(1/\varepsilon) - O(1)$*,*

Moreover, E *can be computed in time* $\mathrm{poly}(n, \log(1/\varepsilon))$ *given two appropriate conductors of size* $S = \mathrm{poly}(2^t, 1/eps)$*, which can be found probabilistically in time* $\mathrm{poly}(S)$ *or deterministically in time* $2^{\mathrm{poly}(S)}$.

The above conductors are near-optimal in terms of parameters, and are efficiently constructible in the case of low or constant degree case. To improve the computation time for general parameters, we can use instead the explicit conductors of Lemma 4.13. This gives:

Theorem 7.3 *For every n, $t \leq n$, $\varepsilon > 0$, there is an explicit (k_{max}, ε) lossless conductor* $\mathrm{E} : (n) \times (d) \rightarrow (n-t)$ *with*

- $d = O(\log^3(t/\varepsilon))$*, and*
- $k_{max} = (n-t) - d - \log(1/\varepsilon) - O(1)$

Moreover, E *can be computed in time* $\mathrm{poly}(n, \log(1/\varepsilon))$.

Theorem 7.4 *For every n, $t \leq n$, $\varepsilon > 0$, there is an explicit (ε, a) extracting conductor* $\mathrm{E} : (n) \times (d) \rightarrow (n-t)$ *with*

- $d = O(\log^3(t/\varepsilon))$*, and*
- $a = d - 2\log(1/\varepsilon) - O(1)$*,*

Moreover, E *can be computed in time* $\mathrm{poly}(n, \log(1/\varepsilon))$.

Without much additional work, we can also combine Theorems 7.2 and 7.1 in a couple of ways: First, we can contruct buffer conductors $\langle \mathrm{E}, \mathrm{C} \rangle$ where E has the parameters of Theorem 7.2 and $\langle \mathrm{E}, \mathrm{C} \rangle$ has the parameters of Theorem 7.1. Second, we can construct a *single* function E that is an extracting conductor with the parameters of Theorem 7.2 and, for slightly lower min-entropies, is also a lossless conductor with the parameters of Theorem 7.1. And, we can do both — have a pair $\langle \mathrm{E}, \mathrm{C} \rangle$, where E is a lossless conductor and $\langle \mathrm{E}, \mathrm{C} \rangle$ is an extracting conductor. Similar combinations can be done for Theorems 7.3 and 7.4. We omit formal statements of all these combinations here.

Acknowledgments

We are deeply grateful to David Zuckerman for stimulating discussions which were the start of our investigation into lossless expanders. We thank Alexander Razborov for pointing out the applications of our results to Proof Theory, and thank Venkatesan Guruswami and Amin Shokrollahi for discussions about the applications to codes. We thank Ronen Shaltiel and Oded Goldreich for helpful conversations.

References

[ABRW00] M. Alekhnovich, E. Ben-Sasson, A. Razborov, and A. Wigderson. Pseudorandom generators in propositional proof complexity. In *41st Annual Symposium on Foundations of Computer Science*, pages 43–53. IEEE, 2000.

[AR01] M. Alekhnovich and A. Razborov. Lower bounds for polynomial calculus: Non-binomial case. In *42nd Annual Symposium on Foundations of Computer Science*, pages 190–199. IEEE, 2001.

[Alo95] N. Alon. Private communication, 1995.

[Alo86] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. Theory of computing (Singer Island, Fla., 1984).

[AGHP92] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.

[AM84] Noga Alon and V. D. Milman. Eigenvalues, expanders and superconcentrators (extended abstract). In *25th Annual Symposium on Foundations of Computer Science*, pages 320–322, Singer Island, Florida, 24–26 October 1984. IEEE.

[ALM96] Sanjeev Arora, F. T. Leighton, and Bruce M. Maggs. On-line algorithms for path selection in a nonblocking network. *SIAM J. Comput.*, 25(3):600–625, 1996.

[BP98] P. Beame and T. Pitassi. Propositional proof complexity: Past, present and future. Technical Report TR98-067, Electronic Colloquium on Computational Complexity, 1998.

[BW99] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow—resolution made simple. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)*, pages 517–526. ACM, New York, 1999.

[BFU99] Andrei Z. Broder, Alan M. Frieze, and Eli Upfal. Static and dynamic path selection on expander graphs: a random walk approach. *Random Structures Algorithms*, 14(1):87–109, 1999.

[BMRS00] Harry Buhrman, Peter Bro Miltersen, Jaikumar Radhakrishnan, and Venkatesh Srinivasan. Are bitvectors optimal? In *Proceedings of the Thirty-Second Annual ACM Symposium on the Theory of Computing*, 2000.

[Cap01] M. Capalbo. Explicit constant-degree unique-neighbor expanders. 2001.

[CG88] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, April 1988.

[CS88] V. Chvátal and E. Szemerédi. Many hard examples for resolution. *Journal of the ACM*, 35(4):759–768, 1988.

[GW97] Oded Goldreich and Avi Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. *Random Structures & Algorithms*, 11(4):315–343, 1997.

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396 (electronic), 1999.

[IZ89] Russell Impagliazzo and David Zuckerman. How to recycle random bits. In *30th Annual Symposium on Foundations of Computer Science*, pages 248–253, Research Triangle Park, North Carolina, 30 October–1 November 1989. IEEE.

[Kah95] Nabil Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM*, 42(5):1091–1106, September 1995.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[LMSS01] Michael G. Luby, Michael Mitzenmacher, M. Amin Shokrollahi, and Daniel A. Spielman. Improved low-density parity-check codes using irregular graphs. *IEEE Trans. Inform. Theory*, 47(2):585–598, 2001.

[Mar88] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, August 1993.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49(2):149–167, October 1994.

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

[PU89] D. Peleg and E. Upfal. Constructing disjoint paths on expander graphs. *Combinatorica*, 9(3):289–313, 1989.

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

[RT97] Jaikumar Radhakrishnan and Amnon Ta-Shma. Tight bounds for depth-two superconcentrators. In *38th Annual Symposium on Foundations of Computer Science*, pages 585–594, Miami Beach, Florida, 20–22 October 1997. IEEE.

[RR99] Ran Raz and Omer Reingold. On recycling the randomness of the states in space bounded computation. In *Proceedings of the Thirty-First Annual ACM Symposium on the Theory of Computing*, Atlanta, GA, May 1999.

[RRV99] Ran Raz, Omer Reingold, and Salil Vadhan. Extracting all the randomness and reducing the error in Trevisan's extractors. In *Proceedings of the Thirty-First Annual ACM Symposium on the Theory of Computing*, pages 149–158, Atlanta, GA, 1999. See preprint of journal version, revised July 2001 for *J. Computer and System Sci.*

[RVW00] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proceedings of 41st Annual Symposium on Foundations of Computer Science*, pages 3–13, 2000.

[RSW00] Omer Reingold, Ronen Shaltiel, and Avi Wigderson. Extracting randomness via repeated condensing. In *41st Annual Symposium on Foundations of Computer Science*, Redondo Beach, California, 12–14 November 2000. IEEE. To appear.

[Ren70] A. Renyi. Probability theory, 1970.

[SU01] Ronen Shaltiel and Christopher Umans. Simple extractors for all min-entropies and a new pseudo-random generator. In *42nd Annual Symposium on Foundations of Computer Science*. IEEE, 14–17 October 2001. To appear.

[Sip88] Michael Sipser. Expanders, randomness, or time versus space. *Journal of Computer and System Sciences*, 36(3):379–383, June 1988.

[SS96] Michael Sipser and Daniel A. Spielman. Expander codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1710–1722, 1996. Codes and complexity.

[Spi96] Daniel A. Spielman. Linear-time encodable and decodable error-correcting codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1723–1731, 1996. Codes and complexity.

[SZ98] Aravind Srinivasan and David Zuckerman. Computing with very weak random sources. To appear in *SIAM Journal on Computing*, 1998. Preliminary version in *FOCS '94*.

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

[TZS01] Amnon Ta-Shma, David Zuckerman, and Shmuel Safra. Extractors from Reed–Muller codes. In *42nd Annual Symposium on Foundations of Computer Science*. IEEE, 14–17 October 2001. To appear.

[Tan84] Michael R. Tanner. Explicit concentrators from generalized n-gons. *SIAM Journal on Algebraic Discrete Methods*, 5(3):287–293, 1984.

[Tre99] Luca Trevisan. Construction of extractors using pseudo-random generators. In *Proceedings of the Thirty-First Annual ACM Symposium on the Theory of Computing*, pages 141–148, Atlanta, GA, May 1999. See also ECCC TR98-55.

[Tse68] G. C. Tseitin. On the complexity of derivations in propositional calculus. In *Studies in constructive mathematics and mathematical logic, Part II*. Consultants Bureau, New-York-London, 1968.

[UW87] Eli Upfal and Avi Wigderson. How to share memory in a distributed system. *J. Assoc. Comput. Mach.*, 34(1):116–127, 1987.

[Urq87] A. Urquhart. Hard examples for resolution. *Journal of the ACM*, 34(1):209–219, 1987.

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

A Renyi Entropy

Another useful measure of entropy is Renyi's (H_2) entropy [Ren70]. Let D be a distribution over a set S. Define the collision probability of D to be $\mathrm{Col}(D) \overset{\text{def}}{=} \Pr[X = Y]$, where X and Y are independently distributed according to D. The **Renyi entropy** of D is defined to be

$$\mathrm{H}_{Ren}(D) \overset{\text{def}}{=} \log(1/\mathrm{Col}(D)) = \log\left(\frac{1}{\mathrm{E}_{x \leftarrow D}\left[\Pr\left[D = x\right]\right]}\right).$$

Thus, Renyi entropy is obtained by switching the order of the expectation and $\log(1/z)$ in the definition of standard Shannon entropy.

The relation between min-entropy and Renyi entropy is given by the following simple proposition:

Proposition A.1 *Let X be any random variable over a set S then* $\mathrm{H}_\infty(X) \leq \mathrm{H}_{Ren}(X) \leq 2\mathrm{H}_\infty(X)$.

Proof: The proof follows from the following inequalities:

$$\left(\max_{a \in S} \Pr[X = a]\right)^2 \leq \mathrm{Col}(X) = \sum_{a \in S} \Pr[X = a]^2 \leq \max_{a \in S} \Pr[X = a].$$

$\blacksquare$

The Renyi entropy of a distribution gives a relatively weak lower bound on its min-entropy. The bound is much stronger if we consider an approximated version of min-entropy (based on statistical difference):

Lemma A.2 *Let X be a random variable over a set S with $H_{Ren}(X) = k$, then for any $\varepsilon > 0$, X is a $(k - \log 1/\varepsilon, \varepsilon)$ source.*

Proof: Define $S_{\mathrm{B}} \overset{\text{def}}{=} \{a \in S | \Pr[X = a] \geq 2^{-k}/\varepsilon\}$. In order to prove that X is a $(k - \log 1/\varepsilon, \varepsilon)$ source it is enough to show that $\Pr[X \in S_{\mathrm{B}}] \leq \varepsilon$. To do so we observe that $2^{-k} = \mathrm{Col}(X) \geq \Pr[X \in S_{\mathrm{B}}] \cdot (2^{-k}/\varepsilon)$. $\blacksquare$

In case the Renyi entropy of a distribution is very close to its maximal value, we can obtain a better bound than that implied by Lemma A.2:

Lemma A.3 *Let X be a random variable over a set S of size N. For any $\varepsilon \geq 0$ and $K = 2^k \leq N$, if*

$$Col(X) \leq 1/N + \varepsilon^2/K + 2\varepsilon(1/K - 1/N),$$

then X is a (k, ε) source.

Proof: For every $a \in S$, define $p_a \overset{\text{def}}{=} \Pr[X = a]$. Define $S_{\mathrm{B}} \overset{\text{def}}{=} \{a \in S | p_a > 1/K\}$. Note that $|S_{\mathrm{B}}| < K$. Assume towards contradiction that X is not a (k, ε) source. This implies that $\sum_{a \in S_{\mathrm{B}}}(p_a - 1/K) > \varepsilon$. It follows that:

$$\begin{aligned} \mathrm{Col}(x) &= \sum_{a \in S} p_a^2 \\ &= 1/N + \sum_{a \in S} (p_a - 1/N)^2 \end{aligned}$$

$$
\begin{aligned}
&\geq 1/N + \sum_{a \in S_{\mathrm{B}}} (p_a - 1/N)^2 \\
&= 1/N + \sum_{a \in S_{\mathrm{B}}} (p_a - 1/K)^2 + \sum_{a \in S_{\mathrm{B}}} (1/K - 1/N)^2 + 2 \sum_{a \in S_{\mathrm{B}}} (p_a - 1/K)(1/K - 1/N) \\
&> 1/N + \sum_{a \in S_{\mathrm{B}}} (p_a - 1/K)^2 + \sum_{a \in S_{\mathrm{B}}} (1/K - 1/N)^2 + 2\varepsilon(1/K - 1/N) \\
&\geq 1/N + \sum_{a \in S_{\mathrm{B}}} (p_a - 1/K)^2 + 2\varepsilon(1/K - 1/N)
\end{aligned}
$$

By Jensen's inequality and the hypothesis we have that

$$\sum_{a \in S_{\mathrm{B}}} (p_a - 1/K)^2 \geq \frac{1}{|S_{\mathrm{B}}|} \left(\sum_{a \in S_{\mathrm{B}}} (p_a - 1/K) \right)^2 > \varepsilon^2/K,$$

which concludes the proof. ■

B Lower Bounds and Constructions of Conductors

We begin by deriving lower bounds on conductors from those known for extractors.

Lemma B.1 *If* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is an* (ε, a) *extracting conductor with with* $m \leq n + a - O(1)$*,* $m \geq d + 2$ *and* $\varepsilon \leq 1/2$*. Then* $d \geq \max\{a, \log(n - m + a)\} + 2\log 1/\varepsilon - O(1)$*.*

Proof: By definition, E is also an $(m - a, \varepsilon)$-extractor. The proof follows by the lower bounds on extractors of Nisan and Zuckerman [NZ96] and their improvement by Radhakrishnan and Ta-Shma [RT97]. ■

Lemma B.2 *If* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is a* (k_{max}, ε) *loss-less conductor with* $\varepsilon \leq 1/4$*,* $k_{max} + d + O(1) < m$ *and* $n + d \leq m - 1$*. Then* $d \geq \max\{(\log((n - k_{max})/(m - k_{max} - d)), \log 1/\varepsilon)\} - O(1)$*.*

Proof: Let $\mathrm{E}' : (m) \times (d') \mapsto (d + d' + 2)$ be a $(k_{max} + d, \varepsilon)$-extractor with $d' = \log(m - k_{max} - d) + 2\log 1/\varepsilon + O(1)$ (such an extractor exists by the probabilistic method [NZ96, RT97]). Consider the extractor $\mathrm{E}'' : (n) \times (d + d') \mapsto (d + d' + 2)$ that is obtained by composing E and E' as follows: $\forall x \in (n), y_1 \in (d), y_2 \in (d'), \mathrm{E}''(x, (y_1, y_2)) = \mathrm{E}'(\mathrm{E}(x, y_1), y_2)$. It is not hard to prove (as in Lemma 4.8) that E'' is a $(k_{max}, 2\varepsilon)$ extractor. Therefore, by the lower bounds of [NZ96, RT97]) we have that $d + d' \geq \log(n - k_{max}) + 2\log(1/\varepsilon) - O(1)$. This implies that $d \geq \log((n - k_{max})/(m - k_{max} - d)) - O(1)$.

Let x_1 and x_2 be any two n-bit strings such that for some d-bit strings r_1 and r_2, we have $E(x_1, r_1) = E(x_2, r_2)$ (such strings exist since $n + d \leq m - 1$. Let X be the uniform distribution over $\{x_1, x_2\}$. Then it is easy to see that $E(X, U_d)$ is at least $1/2^{d+1}$ far from any $(d + 1)$-source. We can therefore conclude that $d \geq \log 1/\varepsilon - 1$. ■

The next lemma describes the conductors obtained from expanders with bounded second eigenvalue.

Lemma B.3 *Let* $N = 2^n$*,* $D = 2^d$*, and let* $\mathrm{E} : (n) \times (d) \rightarrow (n)$ *be the neighbor function of a* D*-regular expander graph on* N *vertices whose adjacency matrix has 2nd largest eigenvalue* $\leq \lambda D$*. Then, for any* $k \leq k' \leq n$*,* $\varepsilon > 0$*, and* k*-source* X *on* (n)*,* $\mathrm{E}(X, U_d)$ *is a* (k', ε)*-source if:*

1. $2\log(1/\lambda) \geq k' - k + 2\log(1/\varepsilon)$ *and* $k' = n$*; or*

2. $2\log(1/\lambda) \geq k' - k + \log(1/\varepsilon)$ *and* $k' \leq n-1$.

Proof: Let $K = 2^k$, $K' = 2^{k'}$, $Y = \mathrm{E}(X, U_d)$. Since the collision probability of a probability distribution is the square of its L_2-norm, and a random step on an expander shrinks the L_2-norm of the non-uniform component (the component perpendicular to the all one vector) of any probability distribution by a factor of λ^2, we have:

$$\mathrm{Col}(Y) - \frac{1}{N} \leq \lambda^2 \left(\mathrm{Col}(X) - \frac{1}{N}\right) \leq \lambda^2 \cdot \left(\frac{1}{K} - \frac{1}{N}\right) \leq \frac{\lambda^2}{K}.$$

For Part 1, we have:

$$\mathrm{Col}(Y) - \frac{1}{N} \leq \frac{\lambda^2}{K} \leq \frac{\varepsilon^2}{K'} = \frac{\varepsilon^2}{K'} + 2\varepsilon\left(\frac{1}{N} - \frac{1}{N}\right) = \frac{\varepsilon^2}{K'} + 2\varepsilon\left(\frac{1}{K'} - \frac{1}{N}\right).$$

By Lemma A.3, Y is a (k', ε)-source.

Similarly, for Part 2 we have:

$$\mathrm{Col}(Y) - \frac{1}{N} \leq \frac{\lambda^2}{K} \leq \frac{\varepsilon}{K'} \leq \frac{\varepsilon^2 + \varepsilon}{K'} \leq \frac{\varepsilon^2}{K'} + 2\varepsilon\left(\frac{1}{K'} - \frac{1}{N}\right),$$

where the last inequality uses $k' \leq n-1$. ■

Proof Lemma 4.4: (sketch) We apply Lemma B.3 to an appropriate power of any explicit constant-degree expander graph G on $[N]$. Specifically, suppose G comes from any family with degree $D = 2^d$ and second largest eigenvalue $\lambda \cdot D < D$. Then when we take t'th power of G, both d and $\log(1/\lambda)$ grow linearly with t. Taking $t = O(a + \log(1/\varepsilon))$, Lemma B.3 implies that every k-source for $k \leq n - a$ gets transformed to a $(k + a, \varepsilon)$-source. ■

Remark B.4 Using *Ramanujan* graphs [LPS88, Mar88], where $\lambda D = 2\sqrt{D-1}$, the bound on t can be improved to $d = n - k + 2\log(1/\varepsilon) + O(1)$.

The following lemma specifies the conductors implied by families of hash functions.

Lemma B.5 *Let $\mathcal{H}$ be a family of hash functions mapping $(n) \to (m)$ with collision probability*[6] *at most $(1+\delta)/2^m$ such that $\mathcal{H}$ can be sampled using d random bits. Define $\mathrm{E} : (n) \times (d) \to (m+d)$ by $\mathrm{E}(x,h) = (h, h(x))$. Then, for any $k \leq n$, $k' \leq m + d$, $\varepsilon > 0$, and k-source X on (n), $\mathrm{E}(X, U_d)$ is a (k', ε)-source if:*

1. $k' \leq k + d - 2\log(1/\varepsilon) - 1$, $k' = m + d$, *and* $\delta \leq \varepsilon^2/2$*; or*

2. $k' \leq k + d - \log(1/\varepsilon) - 1$, $k' \leq m + d - 1$, *and* $\delta \leq \varepsilon^2/2$*; or*

3. $k' \leq k + d$, $k \leq m - \log(1/\varepsilon) - 2$, $\delta \leq 1$, *and* $\varepsilon \leq 1/2$.

Proof: Let $K = 2^k$, $K' = 2^{k'}$, $M = 2^m$, $Y = \mathrm{E}(X, U_d)$. By the collision probability of $\mathcal{H}$,

$$\mathrm{Col}(Y) \leq \frac{1}{D} \cdot \left(\mathrm{Col}(X) + \frac{1+\delta}{M}\right) \leq \frac{1}{DK} + \frac{1+\delta}{DM}.$$

[6] This means that for every $x_1 \neq x_2 \in (n)$, $\Pr_h[h(x_1) = h(x_2)] \leq (1+\delta)/2^m$.

For Part 1, we have:

$$\mathrm{Col}(Y) \leq \frac{\varepsilon^2}{2K'} + \frac{1}{DM} + \frac{\varepsilon^2}{2K'} = \frac{1}{DM} + \frac{\varepsilon^2}{K'} + 2\varepsilon \cdot \left(\frac{1}{K'} - \frac{1}{DM}\right).$$

By Lemma A.3, Y is a (k', ε)-source.

Similarly, for Part 2, we have

$$\mathrm{Col}(Y) \leq \frac{\varepsilon}{2K'} + \frac{1}{DM} + \frac{\varepsilon}{2K'} \leq \frac{1}{DM} + \frac{\varepsilon^2}{K'} + 2\varepsilon \cdot \left(\frac{1}{K'} - \frac{1}{DM}\right),$$

where the last inequality uses $k' \leq n - 1$. Again, Lemma A.3 says that Y is a (k', ε)-source.

For Part 3, we restrict wlog to X which are uniform over a set S_X of size K (as all k-sources are a convex combination of such X). As observed in [TUZ01], to show that $\mathrm{E}(X, U_d)$ is a $(k + d, \varepsilon)$-source for such X is equivalent to showing that $|\,\mathrm{E}(S_X \times (d))| \geq (1-\varepsilon) \cdot |S_x \times (d)|$. Since $|\,\mathrm{E}(S_X \times (d)| \geq 1/\mathrm{Col}(Y)$, it suffices to show $\mathrm{Col}(Y) \leq \frac{1}{(1-\varepsilon)\cdot KD}$. This we argue as follows:

$$\mathrm{Col}(Y) \leq \frac{1}{DK} + \frac{1+\delta}{DM} \leq \frac{1}{K'} + \frac{\varepsilon}{2K'} \leq \frac{1}{(1-\varepsilon)K'}.$$

■

Lemma 4.12 (restated) *For any n, m and $\varepsilon > 0$, there exists an explicit $(m - a, \varepsilon, a)$ buffer conductor $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \to (m) \times (b)$ with $d = O(\log n + \log^3(m/\varepsilon))$, $a = d - 2\log(1/\varepsilon) - O(1)$, and $b = 3\log(1/\varepsilon) + O(1)$.*

Proof Lemma 4.12: Set $\varepsilon' = \varepsilon/3$. Let $m_1 \leq m$; we will set the value of m_1 at the end of the proof. From Lemma 4.9, we have an explicit (ε', a_1) extracting conductor $\mathrm{E}_1 : (n) \times (d_1) \to (m_1)$ with $d_1 = O(\log n + \log^3(m/\varepsilon'))$ and $a_1 = -O(\log^3(m/\varepsilon'))$. Let $k_1 = m_1 - a_1$. We can extend E_1 to a (k_1, ε', a_1) buffer conductor $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle : (n) \times (d_1) \to (m_1) \times (b_1)$ by using the trivial buffer with $b_1 = n + d_1$. Set $b_2 = 3\log(1/\varepsilon') + 3$, $m_2' = d_1 - a_1 - 2\log(1/\varepsilon') - 1 = O(\log^3(m/\varepsilon'))$. Lemma 4.5 gives us a $(k_2, \varepsilon_2, a_2)$ buffer conductor $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle : (b_1) \times (d_2) \to (m_2) \times (b_2)$ with $d_2 = O((\log b_1) + b_2 + m_2' + \log(1/\varepsilon')) = O(\log n + \log^3(m/\varepsilon'))$, $m_2 = d_2 + m_2'$, $a_2 = d_2 - 2\log(1/\varepsilon) - 1$, and $k_2 = m_2' + b_2 - \log(1/\varepsilon) - 2$.

It can be verified that $k_2 = d_1 - a_1 = k_1 + d_1 - m_1$, and $m_2 = d_1 - a_1 + a_2$. So Lemma 4.10 applies, and we get a $(k_1, 3\varepsilon', d_1 + a_2)$ buffer conductor $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d_1 + d_2) \to (m_1 + m_2) \times (b_2)$. This does indeed have error $3\varepsilon' = \varepsilon$, seed length $d = d_1 + d_2 = O(\log n + \log^3(m/\varepsilon'))$, entropy loss $(d_1 + d_2) - (d_1 + a_2) = 2\log(1/\varepsilon') + 1$, and buffer size $b_2 = 3\log(1/\varepsilon') + 3$. By inspecting the above proof, it can be verified that the value of $m_2 = O(\log n + \log^3(m/\varepsilon'))$ can be set independent of m_1 (provided $m_1 \leq m$), so we may indeed guarantee $m_1 + m_2 = m$ as needed. ■

Lemma 4.13 (restated) *For any $k_{max} \leq n$, m, and $\varepsilon > 0$, there exists an explicit $(k_{max}, \varepsilon, a)$ buffer conductor $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \to (m) \times (b)$ with $d = O(\log n + \log^3(m/\varepsilon) + \log^3(k_{max}/\varepsilon))$, $a = d - 2\log(1/\varepsilon) - O(1)$, and $m + b = k_{max} + d + \log(1/\varepsilon) + O(1)$.*

Proof Lemma 4.13: We compose the conductor of Lemma 4.12 with itself via Lemma 4.5.

Let $\varepsilon' = \varepsilon/3$. Taking Lemma 4.12 without the buffer, for any $m_1 \leq m$, we have an explicit (ε', a_1) extracting conductor $\mathrm{E}_1 : (n) \times (d_1) \to (m_1)$ with $d_1 = O(\log n + \log^3(m/\varepsilon'))$, $a = d_1 - 2\log(1/\varepsilon') - O(1)$. We can extend this to extend to a $(k_{max}, \varepsilon', a_1)$ buffer conductor $\langle \mathrm{E}_1, \mathrm{C}_1 \rangle : (n) \times (d_1) \to (m_1) \times (b_1)$ by using the trivial buffer with $b_1 = n + d_1$.

Let $k_2 = k_{max} + d_1 - m_1$. Now, by shifting all of the output except the seed into the buffer in Lemma 4.12, we have a (k_2, ε', a_2) buffer conductor $\langle \mathrm{E}_2, \mathrm{C}_2 \rangle : (b_1) \times (d_2) \to (d_2) \times (b_2)$ with $d_2 = O(\log b_1 + \log^3(k_2/\varepsilon')) = O(\log n + \log^3(k_{max}/\varepsilon') + O(\mathrm{loglog} m))$, $a_2 = d_2 - 2\log(1/\varepsilon') - O(1)$, and $b_2 = k_2 + \log(1/\varepsilon') + O(1)$, and it can be verified that the conditions of Lemma 4.10 are satisfied.

Composing these two buffer conductors via Lemma 4.10 gives a $(k_{max}, 3\varepsilon', d_1 + a_2)$ buffer conductor $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d_1 + d_2) \to (m_1 + d_2) \times (b_2)$ with seed length $d_1 + d_2 = O(\log n + \log^3(m/\varepsilon') + \log^3(k_{max}/\varepsilon'))$, entropy loss $(d_1 + d_2) - (d_1 - a_2) = 2\log(1/\varepsilon') + O(1)$, and buffer size $b_2 = (k_{max} + d_1 - m_1) + \log(1/\varepsilon') + O(1) = k_{max} + d - (m_1 + d_2) + \log(1/\varepsilon') + O(1)$, as desired. Again the value of d_2 can be set independently of m_1, so we can ensure $m_1 + d_2 = m$. ■