Randomness Conductors and Constant-Degree Lossless Expanders

[Extended Abstract] *

Michael Capalbo
Institute for Advanced Study
Princeton, NJ
mrc@ias.edu

Omer Reingold†
AT&T Labs — Research
Florham Park, NJ
omer@research.att.com

Salil Vadhan‡
Harvard University
Cambridge, MA
salil@eecs.harvard.edu

Avi Wigderson
Hebrew University
Jerusalem, Israel, and
Institute for Advanced Study
Princeton, NJ
avi@ias.edu

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. The best previous explicit constructions gave expansion factor D/2, which is too weak for many applications. The D/2 bound was obtained via the eigenvalue method, and is known that that 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. 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 graph product, previously defined for expanders and extractors.

*A full version of this paper will be posted on the *Electronic Colloquium on Computational Complexity*, `http://www.eccc.uni-trier.de/eccc/`.

†Part of this research was performed while visiting the Institute for Advanced Study, Princeton, NJ.

‡Work begun while at MIT and the Institute for Advanced Study, supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

STOC'02, May 19-21, 2002, Montreal, Quebec, Canada.

Categories and Subject Descriptors

G.2.1 [**Discrete Mathematics**]: Graph Theory; G.3 [**Probability and Statistics**]: Random Number Generation

General Terms

Theory, Algorithms

Keywords

expander graphs, extractors, condensers, graph products

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 and diverse set of applications of such functions in both computer science and pure mathematics makes them central objects for 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 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 general notion of a *randomness conductor* that encompasses all the 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 $K \approx M/D$. A nonconstructive probabilistic argument shows that such graphs do exist with $D = O(\log(N/M))$, and this value of D is best possible. (Here and below, we fix ϵ to be an arbitrarily small constant for simplicity.)

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 $D = \text{polylog}(N/M)$ when N/M is relatively small (so that an optimal graph of size $\text{poly}(N/M)$ can be found by, say, exhaustive search) and $D = \exp(\text{polyloglog}(N/M))$ 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 [16, 18]. Moreover, Kahale [15] showed that some (asymptotically) 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 [31] 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 $D = \text{polylog}(N)$ (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 [4] based on graphs of high girth, but again only very small sets ($A = N^\alpha$ for some constant α) expand losslessly.

Also, weaker objects of a similar nature were constructed before. Raz and Reingold [23] introduced a method which appends a buffer to a "lossy" extractor, which retains the "lost" entropy. This translates to highly unbalanced, non-constant degree graphs which are lossless for sets 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 [11] 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 [26] 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 suggested in [26].

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. First, it will be useful to deduce a few properties of lossless expanders.

LEMMA 1.1. *Let G be a bipartite graph with N inputs, M outputs, every input vertex of degree D, and every subset of input vertices S of size at most K have at least $(1-\epsilon)D|S|$ output neighbors $\Gamma(S)$ for $\epsilon \leq 1/2$. Then for every such subset S we have*

1. *At most a 2ϵ fraction of the vertices in $\Gamma(S)$ have degree ≥ 2 into S.*
2. *A least a $(1 - 2\epsilon)D|S|$ vertices in $\Gamma(S)$ are* unique neighbors*, namely have degree 1 into S.*
3. *At least a $(1 - 2\epsilon)$ fraction of the vertices in S have a unique neighbor.*
4. *For every $\delta \geq 2\epsilon$, at least a $1-\delta$ fraction of the vertices in S each have more than $(1-2\epsilon/\delta)D$ unique neighbors.*
5. *For every $\delta \geq 2\epsilon$, at most $2\epsilon/(\delta - 2\epsilon)|S|$ inputs* not *in S each have at least δD neighbors in $\Gamma(S)$, provided $|S| \leq (1 - 2\epsilon/\delta)K$.*
6. *Any δ fraction of edges from S touches at least $(\delta - \epsilon)D|S|$ outputs.*

Now we can see how different properties are used in different applications. We use the same parameters as in the lemma.

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 [20, 6, 9]. 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 G describes a distributed network. Assume that a set S of inputs in G needs to find a complete matching into the outputs, i.e. a matching which matches all the vertices in S. By utilizing Property (3), an iterative distributed algorithm in which these vertices look for unique neighbors converges in $O(|S|)$ work and communication, and $O(\log |S|)$ parallel phases. This may be viewed as a first step in constructing disjoint paths.

Linear-Time Decodable Error-Correcting Codes. A large body of work, best known under Low Density Parity Check (LDPC) codes, constructs good codes from graphs with good

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

expansion properties (e.g. [17, 28, 29] and the references therein). The following, which is from [28], illustrates the power of lossless expanders in this context. Specifically, they yield asymptotically good linear codes of *every* constant rate, with trivial linear time (and $O(\log n)$ parallel steps) decoding algorithm.

Here G describes the parity check matrix of the code. 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. The rate is $\geq 1 - M/N$, 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 more than $2D/3$ of its neighbors have value 1, it flips its value. It is easy to see, via Properties (3) and (5), that the total number of corrupted inputs will shrink by a constant factor each round.

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 [10]. 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 the smallest value of M (for constant error) relies on lossless expanders. Let us see how.

G will determine the storage scheme as follows. Given a set S of inputs of size K, we will represent it by labelling the outputs with binary values. This will be done in such a way that the vast majority of neighbors $(1-\delta)|D|$ of each vertex in N will correctly indicate whether $v \in S$. Thus querying a random neighbor of $v \in [N]$ will only err with probability δ.

Why should such a labelling exist? As in the previous examples, lets attempt to get it greedily. First, label all output vertices in $\Gamma(S)$ by 1, and the rest by 0. This classifies correctly vertices in S, but might misclassify vertices in a set T disjoint from S, each of whose vertices has at least δD neighboring outputs in $\Gamma(S)$. Fix the problem by (re)labelling all outputs in $\Gamma(T)$ by 0. This certainly fixes the problem in T, but may create one in a subset U of S. Fix $\Gamma(U)$ to 1, etc. By picking $\delta = 6\epsilon$, Property (5) guarantees that the sizes of the problematic sets shrink by a factor of 2 in each iteration; hence the procedure terminates.

Fault-tolerance and a Distributed Storage Method. By Property (6), lossless expanders have incredible fault-tolerance: removing all but δD neighbors of every input in G in an arbitrary way is still a lossless expander (with the same K and new $\epsilon\prime = \epsilon/\delta$. This property was used (with $\delta \approx 1/2$) in a distributed storage scheme due to [35], who gave a near-optimal *deterministic* simulation of PRAMs by a network of communicating processors. Both models have M processors, and 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, naive methods for distributing the items fail.

The idea in [35] is to use G so that inputs represent the N data items, and the outputs represent (the memories) of the N processors. Each item has $D = 2c-1$ "copies", which are distributed in its neighbors. 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 timestamps 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. Efficiency follows from a similar (but more complex and delicate) argument to the distributed routing construction, since here *many* (c) disjoint complete matchings are needed.

Hard Tautologies in Proof Complexity. The field of Proof Complexity studies propositional proof systems and tries to prove lower bounds on the sizes 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, for example, the excellent survey [7].) A sequence works [34, 36, 13, 8, 1, 2] have 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.

Some tautologies are constructed in [1] from a graph G by letting every *input* vertex be viewed as a function of the bits labelling the *outputs* of G. (Note that this is opposite to LDPC codes and moreover these functions may not be parities). The tautology expresses the statement that no M-bit sequence in the outputs can make all functions zero simultaneously. Losslessness is essential to the lower bound proofs.

It should be noted that in proof complexity, unlike computational complexity, existential lower bounds are interesting. Still, explicit examples are always more informative.

1.4 Randomness Conductors

In this subsection, we 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:

Set 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* [27] and *a-expanding graphs* [21], 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 [32, 5, 3]. 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.

Extraction. Extractors, introduced by Nisan and Zuckerman [19], 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 [19, 22], and also their definition only discusses X whose min-entropy is at least some value and thereby does not guarantee expansion of small sets.

Condensing. 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 sufficient min-entropy. Various formalizations of this basic idea have appeared in the recent extractor literature [23, 25, 31], 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 is a better description of how we came about our construction, and that these objects merit a more unified 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 primarily restrict ourselves to *simple conductors* (which are still more general than what we really need). In these we fix both error (statistical difference) as well as the difference between the input and output entropies ($k_{out} - k_{in}$) to given 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. We use other types of conductors, too, 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 [33, 24]) 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 [26], 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 [26], 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.

2. PRELIMINARIES

Expanders are graphs which are sparse but nevertheless highly connected. The standard definition of expanders is in terms 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.*

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

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

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

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

X is a (k, ε)**-source** if it is ε-close to some k-source.

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 ([19]). *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

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* $k_{in} \in [0,n]$*, any* $k_{out} \in [0,m]$ *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 like extractors and condensers restricted special cases of conductors.

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 unbalanced case ($m < n$), (b) the lossless case ($k_{out} = k_{in} + d$) and the lossy case ($k_{out} < k_{in} + d$), (c) the extractor scenario (where $k_{out} = m$) and the condenser scenario (where k_{out} may be much smaller than m).

One can also consider defining randomness conductors using a variety of different measures of randomness. Nevertheless, for our definition we fix a particular measure. 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.

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.)

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 3.2 (SIMPLE CONDUCTORS). *A function* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is a* $(k_{max}, \varepsilon, a)$ simple conductor *if for any* $0 \le k \le k_{max}$*, 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 3.3 (EXTRACTING CONDUCTORS). *A function* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is an* (ε, a) extracting conductor *if for any* $0 \le k \le m - a$*, 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 3.4 (LOSSLESS CONDUCTORS). *A function* $\mathrm{E} : (n) \times (d) \mapsto (m)$ *is a* (k_{max}, ε) lossless conductor *if for any* $0 \le k \le k_{max}$*, 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 [31], lossless conductors[3] 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. We will use the notation $\langle \mathrm{E}, \mathrm{C} \rangle : (n) \times (d) \mapsto (m) \times (b)$ to indicate that $\mathrm{E} : (n) \times (d) \mapsto (m)$, $\mathrm{C} : (n) \times (d) \mapsto (b)$, and $\langle \mathrm{E}, \mathrm{C} \rangle$ is the concatenation of these two functions (i.e., $\langle \mathrm{E}, \mathrm{C} \rangle(x, r) = \mathrm{E}(x, r) \circ \mathrm{C}(x, r)$).

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

Definition 3.5 (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 3.6 (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 an* (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

In this section, we describe some basic conductors. The first set of them will be shown to exist using the Probabilistic Method. These will serve us in two ways. They will be components in our zig-zag product when their size is a fixed 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 then extend known composition techniques from extractors and condensers to conductors. This will help improve the parameters of the above constructions.

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 [19, 22]; details are given in the full version of the paper.

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

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

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*,* $m \leq n$*,* b*,* $\epsilon > 0$*, there is a* $(k_{max}, \varepsilon, a)$ *buffer conductor* $\langle \mathrm{E}, \mathrm{C}\rangle : (n) \times (d) \to (m) \times (b)$ *with*

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

We now describe some *explicit* conductors implied by existing constructions. The first is based on expanders with bounded second eigenvalue. The analysis merely involves converting the guarantees on Renyi entropy directly provided by the eigenvalue bound into ε-closeness to min-entropy. For the case of extraction (output min-entropy equals output length), this kind of analysis was done in [14]. 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 [26]). 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 $\langle \mathrm{E}, \mathrm{C}\rangle : (n) \times (d) \to (n) \times (d)$ *with* $d = O(a + \log(1/\varepsilon))$*.*

The key feature of the above conductors is that d does not depend on n. They are suboptimal in that a constant fraction of the entropy in d is lost. In order to achieve losslessness, we obtain explicit conductors from the extractor literature, using both constructions and composition techniques from this literature. In particular, drawing upon [14, 30, 33, 23, 24, 31], we obtain the following.

Lemma 4.5. *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* $m + b = k_{max} + d + \log(1/\varepsilon) + O(1)$*.*

The above conductors are optimal in all parameters except for d. The expression for d is suboptimal in that in two respects: Most importantly for us, it depends on n, m, and k_{max} rather than just on $n - m$, which means it cannot give constant-degree conductors. This problem will be solved by using our new zig-zag product to combine these extractors and the ones of Lemma 4.4. A second deficiency, which we don't solve, is that this dependence is polylogarithmic rather than logarithmic. For lack of space, the constructions and proofs underlying Lemma 4.5 is deferred to the final version.

5. ZIG-ZAG FOR CONDUCTORS

In this section we show how to compose conductors via the zig-zag product of [26]. When applied to the conductors described in Section 4, this composition will imply constant-degree, lossless expanders. (Details appear in Section 7.)

The original zig-zag product. Let us first briefly recall the intuition of the zig-zag product for expanders from [26]. 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 [26] 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

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

$r_1 \stackrel{\text{def}}{=} \mathrm{E}_2(x_2, r_2)$, $y_1 \stackrel{\text{def}}{=} \mathrm{E}_1(x_1, r_1)$, and $y_2 \stackrel{\text{def}}{=} \mathrm{E}_2(r_1, r_3)$. 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 [26] 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 G 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 given particular assignments $X_1 = x_1$ (for x_1 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 show 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 at least a bits of entropy from r_1 into y_1 (in addition to the entropy present in 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, and 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 product was used in [26] 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 [26] 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 second application of E_2 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 original product, 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" (as opposed to the easier "extractor analysis") of [26] in the setting of unbalanced graphs.

Definition 5.1 (zig-zag product [26]). *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* $n = n_1 + n_2$, $d = d_2 + d_3$, $m = m_1 + m_3$, *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*

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

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

Note that r_1 and y_1 are computed in exactly the same way as in the original zig-zag product described above, except that at the same time we produce the buffers $z_1 = \mathrm{C}_2(x_2, r_2)$, $z_2 = \mathrm{C}_1(x_1, r_1)$ to hold any leftover entropy. The second application of E_2 has been replaced with an application of E_3 to collect the entropy from the buffers. In the full version, 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 remainder 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 $\log^3(1/\varepsilon)$. 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 a $(106e, \varepsilon, 0)$ buffer 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 4.5.)

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 $k \leq n_1 + 100e$, and R_2, R_3 are both uniform over (e). The intermediate steps in this computation are $(R_1, Z_1) = \langle \mathrm{E}_2, \mathrm{C}_2 \rangle (X_2, R_2)$, $(Y_1, Z_2) = \langle \mathrm{E}_1, \mathrm{C}_1 \rangle (X_1, R_1)$, and $Y_2 = \mathrm{E}_3(Z_1 \circ Z_2, R_3)$. The output is $Y_1 \circ Y_2$.

As in the original zig-zag 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 $100e$ 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 $100e$ bits 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, the extracting property of E_2 implies that R_1 is close to uniform (even conditioned on X_1). Thus E_1, being an extracting conductor, will either push $5e$ bits of entropy from R_1 into Y_1, or will fill Y_1 up (if there is no room). Since before this step, we have at least $k - 106e$ entropy bits in X_1 (as at most $106e$ can be in X_2), the former will ensure that there are at least $k - 101e$ bits of entropy in Y_1, while the latter will ensure at least n_1. Given that $k \leq n_1 + 100e$, we see that Y_1 will have entropy at least $k - 101e$. Thus $Z_1 \circ Z_2$, which contains all the remaining entropy, has at most $101e + e = 102e$ bits of entropy. Therefore, as before, E_3 will squeeze all this entropy plus its seed length e into Y_2. So, 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 4.5, 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 $d_2 + a_3$ 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" a_1 bits of entropy into Y_1 (if there is room for them) or fill Y_1 up (if there is no room). 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).

The intuition for the proof of Theorems 6.1 and 6.2 follows the same kind of entropy "bookkeeping" that was given for the example in Section 5. To carry out the manipulations of (min-)entropy that are required to formalize these arguments, we use known techniques from the extractor literature. In particular, the arguments we use are strongly influenced by the notion of block sources [12] and its usefulness for randomness extraction [19] (cf., [26]). The actual proof is deferred to to the full version of the paper.

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. (We just state the results here; the straightforward but tedious proofs are deferred to the full version.) 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, $t \leq n$, $\varepsilon > 0$, 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* $S = \mathrm{poly}(2^t, 1/\varepsilon)$, *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).

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* $\mathrm{E} : (n) \times (d) \to (n-t)$ *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* $S = \mathrm{poly}(2^t, 1/\varepsilon)$, *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 4.5. This gives:

THEOREM 7.3. *For every n, $t \leq n$, $\varepsilon > 0$, there is an explicit (k_{max}, ε) lossless conductor* $\mathrm{E} : (n) \times (d) \to (n-t)$ *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* $\mathrm{E} : (n) \times (d) \to (n-t)$ *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 construct 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. 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 Ronen Shaltiel and Oded Goldreich for helpful conversations, Alexander Razborov for pointing out the applications of our results to Proof Theory, Venkatesan Guruswami and Amin Shokrollahi for discussions about the applications to codes, and the referees for helpful comments on the presentation.

8. REFERENCES

[1] 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.

[2] 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.

[3] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.

[4] N. Alon. Private communication. 1995.

[5] N. 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 Oct. 1984. IEEE.

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

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

[8] E. Ben-Sasson and A. Wigderson. Short proofs are narrow—resolution made simple. *Journal of the ACM*, 48(2):149–169, March 2001.

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

[10] H. Buhrman, P. B. Miltersen, J. Radhakrishnan, and V. Srinivasan. Are bitvectors optimal? In *Proceedings of the Thirty-Second Annual ACM Symposium on the Theory of Computing*, 2000.

[11] M. Capalbo. Explicit constant-degree unique-neighbor expanders. Submitted, 2001.

[12] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, Apr. 1988.

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

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

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

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

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

[18] 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.

[19] N. Nisan and D. Zuckerman. Randomness is linear in space. *J. Comput. Syst. Sci.*, 52(1):43–52, Feb. 1996.

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

[21] N. Pippenger. Sorting and selecting in rounds. *SIAM J. Comput.*, 16(6):1032–1038, Dec. 1987.

[22] J. Radhakrishnan and A. 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 Oct. 1997. IEEE.

[23] R. Raz and O. 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.

[24] R. Raz, O. Reingold, and S. 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.*

[25] O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. In *41st Annual Symposium on Foundations of Computer Science*, Redondo Beach, California, 12–14 Nov. 2000.

[26] 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.

[27] M. Sipser. Expanders, randomness, or time versus space. *J. Comput. Syst. Sci.*, 36(3):379–383, June 1988.

[28] M. Sipser and D. A. Spielman. Expander codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1710–1722, 1996.

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

[30] A. Srinivasan and D. Zuckerman. Computing with very weak random sources. *SIAM J. Comput.*, 28(4):1433–1459 (electronic), 1999.

[31] 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.

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

[33] L. Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48(4):860–879, July 2001.

[34] 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.

[35] E. Upfal and A. Wigderson. How to share memory in a distributed system. *Journal of the ACM*, 34(1):116–127, 1987.

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