DIGITAL ACCESS TO
SCHOLARSHIP AT HARVARD

On Extractors and Exposure-Resilient Functions for Sublogarithmic Entropy

(Article begins on next page)

On Extractors and Exposure-Resilient Functions for Sublogarithmic Entropy*

Yakir Reshef[†]   Salil Vadhan[‡]

March 9, 2012

Abstract

We study resilient functions and exposure-resilient functions in the low-entropy regime. A *resilient function* (a.k.a. deterministic extractor for oblivious bit-fixing sources) maps any distribution on n-bit strings in which k bits are uniformly random and the rest are fixed into an output distribution that is close to uniform. With *exposure-resilient functions*, all the input bits are random, but we ask that the output be close to uniform conditioned on any subset of $n - k$ input bits. In this paper, we focus on the case that k is sublogarithmic in n.

We simplify and improve an explicit construction of resilient functions for k sublogarithmic in n due to Kamp and Zuckerman (SICOMP 2006), achieving error exponentially small in k rather than polynomially small in k. Our main result is that when k is sublogarithmic in n, the short output length of this construction ($O(\log k)$ output bits) is optimal for extractors computable by a large class of space-bounded streaming algorithms.

Next, we show that a random function is a resilient function with high probability if and only if k is superlogarithmic in n, suggesting that our main result may apply more generally. In contrast, we show that a random function is a static (resp. adaptive) exposure-resilient function with high probability even if k is as small as a constant (resp. $\log \log n$). No explicit exposure-resilient functions achieving these parameters are known.

Keywords: pseudorandomness, exposure-resilient function, randomness extractor, bit-fixing source

1 Introduction

Randomness extractors are functions that extract almost-uniform bits from weak sources of randomness (which may have biases and/or correlations). Extractors can be used for simulating randomized algorithms and protocols with weak sources of randomness, have close connections to many other "pseudorandom objects" (such as expander graphs and error-correcting codes), and have a variety of other applications in theoretical computer science.

1

The most extensively studied type of extractor is the *seeded extractor*, introduced by Nisan and Zuckerman [NZ]. These extractors are given as additional input a small "seed" of truly random bits to use as a catalyst for the randomness extraction, and this allows for extracting almost-uniform bits from very unstructured sources, where all we know is a lower bound on the min-entropy. In many applications, such as randomized algorithms, the need for truly random bits can be eliminated by trying all possible seeds and combining the results (e.g. by majority vote). However, prior to the Nisan–Zuckerman notion, there was a substantial interest in *deterministic extractors* (which have no random seed) for restricted classes of sources. Over the past decade, there has been a resurgence in the study of deterministic extractors, motivated by settings where enumerating all possible seeds does not work (e.g. distributed protocols) and by other applications in cryptography.

In this paper, we study one of the most basic models: an *oblivious bit-fixing source (OBFS)* is an n-bit source where some $n - k$ bits are fixed arbitrarily and the remaining k bits are uniformly random. Deterministic extractors for OBFSs, also known as *resilient functions (RFs)*, were first studied in the mid-80's, motivated by cryptographic applications [Vaz, BBR, CGH$^+$]. A more relaxed notion is that of an *exposure-resilient function (ERF)*, introduced in 2000 by Canetti et al. [CDH$^+$]. Here all n bits of the source are chosen uniformly at random, but $n - k$ of them are seen by an adversary; an ERF should extract bits that are almost-uniform even conditioned on what the adversary sees. ERFs come in two types: *static* ERFs, where the adversary decides which $n - k$ bits to see in advance, and *adaptive* ERFs, where the adversary reads the $n - k$ bits adaptively. In recent years, there has been substantial progress in giving explicit constructions of both RFs and ERFs [CDH$^+$, DSS, KZ, GRS].

In this paper, we focus on the case when k, the number of random bits unknown to the adversary, is very small, e.g. $k < \log n$. While this case is not directly motivated by applications, it is interesting from a theoretical perspective for a couple of reasons:

- For many other natural classes of sources (such as several independent sources [CG], samplable sources [TV], and affine sources [BKS$^+$]), at least logarithmic min-entropy is necessary for extraction.[1]

- This is a rare case where a random function is *not* an optimal extractor. For example, the parity function extracts one completely unbiased bit from any bit-fixing source with $k = 1$ random bits, but we show that a random function will fail to extract from some such source with high probability.

Our first results concern explicit constructions of extractors for OBFS with k sublogarithmic in n.

- We simplify and improve an explicit construction of extractors for OBFSs with small k by Kamp and Zuckerman [KZ]. In particular, the error parameter of our construction can be exponentially small in k, whereas the

---

[1]For the case of 2 independent sources, the need for logarithmic min-entropy is proven in [CG]. For sources samplable by circuits of size $s = n^2$, it can be shown by noting that the uniform distribution on any 2^k elements of $\{0, 1\}^n$ is samplable by a circuit of size $O(n \cdot 2^k)$ (and we can pick 2^k elements on which the first bit of the extractor is constant). For affine sources, it can be shown by analyzing the k-th Gowers norm of the set of inputs on which the first bit of the extractor is constant (as pointed out to us by Ben Green).

Kamp–Zuckerman construction achieves error that is polynomially small in k. Our extractor (like that of [KZ]) extracts only $\Theta(\log k)$ almost-uniform bits, in contrast to extractors for superlogarithmic k, which can extract nearly k bits [GRS].

- We prove that, when k is sublogarithmic, the $\Theta(\log k)$ output length of our extractor is optimal for extractors for OBFSs computable by space-bounded streaming algorithms with a certain "forgetlessness" property. The class of streaming algorithms we analyze includes our construction as well as many natural random-walk based constructions. This is our main result.

Next, we investigate properties of random functions as extractors for OBFS's and find that $k \approx \log n$ appears to be a critical point for extractors for OBFSs in this setting as well. Specifically, we show that:

- A random function is an extractor for OBFSs (with high probability) *if and only if k is at least roughly $\log n$.*

- In contrast, for the more relaxed concept of exposure-resilient functions, random functions suffice even for sublogarithmic k. For static ERFs, k can be as small as a constant, and for adaptive ERFs, k can be as small as $\log \log n$.

All of the results concerning random functions yield resilient/exposure-resilient functions that output nearly k almost-uniform bits.

2 Preliminaries

Throughout, we will use the convention that a lowercase number (e.g. n) implicitly defines a corresponding capital number (N) as its exponentiation with base 2 (i.e. $N = 2^n$).

Definition 2.1 (Statistical Distance). Let X and Y be two random variables taking values in a set S. The *statistical distance* $\Delta(X, Y)$ between X and Y is

$$\Delta(X, Y) = \max_{T \subset S} |\Pr[X \in T] - \Pr[Y \in T]| = \frac{1}{2} \sum_{w \in S} |\Pr[X = w] - \Pr[Y = w]|$$

We will write $X \approx_\varepsilon Y$ to mean $\Delta(X, Y) \leq \varepsilon$, and we will use U_n to denote the uniform distribution on $\{0,1\}^n$. When U_n appears twice in the same set of parentheses, it will denote the same random variable. For example, a string chosen from the distribution (U_n, U_n) will always be of the form $w \circ w$ for some $w \in \{0,1\}^n$. To avoid confusion, we will explicitly indicate cases in which we refer to multiple independent copies of uniform distributions. For example, (U_n, U'_m) equals U_{n+m}.

Definition 2.2 (Oblivious Symbol-Fixing Source). An (n, k, d) *oblivious symbol-fixing source (OSFS)* X is a distribution on strings of length n with symbols in $[d]$, of which all but k positions are fixed and the rest are chosen independently and uniformly at random.

Definition 2.3 (Oblivious Bit-Fixing Source)**.** An (n, k) *oblivious bit-fixing source (OBFS)* is an $(n, k, 2)$ oblivious symbol-fixing source.

We will use $\binom{[n]}{\ell}$ to denote the set $\{L \subset [n] : |L| = \ell\}$ and, given some $L \in \binom{[n]}{\ell}$ and a string $a \in \{0, 1\}^\ell$, we will write $L^{a,n}$ to denote the oblivious bit-fixing source that has the bits with positions in L fixed to the string a.

Definition 2.4 (Deterministic Randomness Extractor)**.** Let $\mathcal{C}$ be a class of sources on $\{0, 1\}^n$. A *deterministic ε-extractor* for $\mathcal{C}$ is a function $E : \{0, 1\}^n \to \{0, 1\}^m$ such that for every $X \in \mathcal{C}$ we have $E(X) \approx_\varepsilon U_m$.

Here we will focus mainly on deterministic randomness extractors for oblivious bit-fixing sources, also known as *resilient functions* (RFs).

Definition 2.5 (Resilient Function)**.** A (k, ε)-*RF* is a function $f : \{0, 1\}^n \to \{0, 1\}^m$ that is a deterministic ε-extractor for (n, k) oblivious bit-fixing sources.

We can also characterize extractors for OBFSs by their ability to fool a distinguisher: consider a computationally unbounded adversary A that can set some of f's input bits in advance but must allow the rest to be chosen uniformly at random. Then f satisfies Definition 2.5 if and only if A is unable to distinguish between f's output and the uniform distribution regardless of how A changes f's input.

When viewed through this lens, the notion of deterministic extraction from OBFSs has a natural relaxation obtained by restricting A to only *read* (rather than modify) a portion of f's input bits. Functions that are able to fool adversaries of this type are called *exposure-resilient functions* (ERFs). We define below the two simplest variants of exposure-resilient functions, which correspond to whether A reads the bits of f's input all at once or one at a time.

Definition 2.6 (Static Exposure-Resilient Function)**.** A *static (k, ε)-ERF* is a function $f : \{0, 1\}^n \to \{0, 1\}^m$ with the property that for every $L \in \binom{[n]}{n-k}$, f satisfies $(U_n|_L, f(U_n)) \approx_\varepsilon (U_n|_L, U'_m)$.

This definition can be restated in terms of average-case extraction using the following lemma.

Lemma 2.7. *A function $f : \{0, 1\}^n \to \{0, 1\}^m$ is a static (k, ε)-ERF if and only if for every $L \in \binom{[n]}{n-k}$, f satisfies*

$$\underset{a \leftarrow U_{n-k}}{\mathrm{E}} \left[\Delta \left(f \left(L^{a,n} \right), U_m \right) \right] \leq \varepsilon$$

Proof. The lemma follows from the general fact that for two joint distributions (X, Y) and (X, Y'),

$$\Delta \left((X, Y), (X, Y') \right) = \underset{x \leftarrow X}{\mathrm{E}} \left[\Delta \left(Y|_{X=x}, Y'|_{X=x} \right) \right]$$

Setting $X = U_n|_L$, $Y = f(U_n)$, $Y' = U'_m$ gives the result since then $f(U_n)|_{X=a} = f(L^{a,n})$ and $U'_m|_{X=a} = U'_m$. □

Allowing the adversary to adaptively request bits of f's input one at a time gives rise to the strictly stronger notion of an *adaptive ERF*:

Definition 2.8 (Adaptive Exposure-Resilient Function)**.** An *adaptive* (k, ε)-*ERF* is a function $f \colon \{0,1\}^n \to \{0,1\}^m$ with the property that for every algorithm $A \colon \{0,1\}^n \to \{0,1\}^*$ that can (adaptively) read at most $n-k$ bits of its input,[2] f satisfies $(A(U_n), f(U_n)) \approx_\varepsilon (A(U_n), U_m')$.

The following lemma will allow us to restrict our attention to algorithms A that simply output the values of the bits that they request as they receive them (rather than outputting some function of those bits). The lemma is a straightforward application of the fact that application of deterministic functions can only reduce the statistical distance between two random variables.

Lemma 2.9. *Let $A \colon \{0,1\}^n \to \{0,1\}^*$ be an adaptive adversary that reads at most d bits of its input and let $A_r \colon \{0,1\}^n \to \{0,1\}^*$ be the algorithm that adaptively reads the same bits as A and outputs them in the order that they were read. For every function $f \colon \{0,1\}^n \to \{0,1\}^m$, the statistical distance between $(A(U_n), f(U_n))$ and $(A(U_n), U_m')$ is at most the distance between $(A_r(U_n), f(U_n))$ and $(A_r(U_n), U_m')$.*

Proof. First, modify A_r by padding its output with 0's so that its output length is always d. Now define a second algorithm $A_p \colon \{0,1\}^d \to \{0,1\}^*$ as follows: on an input $x \in \{0,1\}^d$, A_p runs A, sequentially feeding it the bits of x in response to A's requests, and then outputs A's output. The fact that $A = A_p \circ A_r$ then implies the desired result. $\qquad\square$

3   A simplification and a lower bound

In this section, we prove that when the entropy parameter k is sublogarithmic in the input length n, an output length of $O(\log k)$ is optimal for a natural class of space-bounded streaming algorithms, including algorithms that use the input bits to conduct a random walk on a graph. Before we state this lower bound, we give a simple improvement on the state of the art in explicit constructions of extractors for oblivious bit-fixing sources (i.e. resilient functions) for sublogarithmic entropy. Our lower bound then shows that the parameters achieved by this construction are optimal.

3.1   The simplification

We start with a simplification of a previous construction due to [KZ]. The construction of [KZ] is based on very good extractors for oblivious *symbol-fixing* sources with $d \geq 3$ symbols obtained by using the symbols of the input string to take a random walk on an expander graph of degree d. Since expander graphs do not exist with degree $d = 2$, this approach could not be used for oblivious bit-fixing sources. However, the construction of [KZ] uses the fact that while a random walk on an expander is not an option, a random walk on a cycle still extracts some randomness even when the entropy k of the input is very small. Our construction is a slight modification of this random walk that simplifies the argument and improves the error parameter.

---

[2]In other words, A is a binary decision tree of depth $n - k - 1$ with leaves labelled by its output strings and each internal node labelled by the position of the bit that A requests at that juncture.

Theorem 3.1. *For every $n \in \mathbb{N}$, $k \in [n]$, $\varepsilon > 0$, and $m = \frac{1}{2}(\log k - \log \log (1/\varepsilon))$, the function $f \colon \{0,1\}^n \to \{0,1\}^m$ defined by*

$$f(w) = \sum_{i=1}^{n} w_i \pmod{2^m}$$

is a (k, ε)-RF. In particular, setting $\varepsilon = 2^{-\sqrt{k}}$ gives output length $m = \frac{1}{4} \log k$.

Proof. We can treat f as computing the endpoint of a walk on $\mathbb{Z}/M\mathbb{Z}$ (where $M = 2^m$) that starts at 0 and either adds 1 or 0 to its state with every bit that it reads. Since the endpoint of this walk does not depend on the order in which the input bits are processed, we may assume without loss of generality that all of the fixed bits in f's input come at the beginning. These bits only change the starting vertex of the random walk and do not affect the distance from uniform of the resulting distribution. Therefore, to bound the distance from uniform of any distribution of the form $f(L^{*,n})$ we need only bound the mixing time of a walk on $\mathbb{Z}/M\mathbb{Z}$ consisting of k random steps. The following claim, whose proof we defer to the appendix, accomplishes this.

Claim 3.2. *Let W_k be the distribution on the vertices of $\mathbb{Z}/M\mathbb{Z}$ (where $M = 2^m$) obtained by beginning at 0 and adding 1 or 0 with equal probability k times. Then*

$$\Delta(W_k, U_m)^2 \leq \frac{e^{-k\pi^2/2M^2}}{2\left(1 - e^{-3k\pi^2/2M^2}\right)}$$

Since $k \geq M^2$, the denominator in Claim 3.2 is bounded from below by $2(1 - e^{-3\pi^2/2}) > 1$ and so we have bounded the distance from uniform by $e^{-k\pi^2/4M^2}$. With our setting of parameters this is at most $\varepsilon^{\log(e)\pi^2/4} \leq \varepsilon$, as desired. $\qquad\square$

The difference between this construction and that of [KZ] is that each step of the random walk carried out by f consists of adding either 1 or 0 rather than 1 or -1 to the current state; that is, the construction of [KZ] involves a simple random walk while ours uses a directed lazy random walk. This has two advantages. First, the random walk in the construction of [KZ] cannot be carried out on a graph of size 2^m since any even-sized cycle is bipartite and the walk traverses an edge at each step. This necessitates an additional lemma about converting the output of the random walk to one that is almost uniformly distributed over $\{0,1\}^m$, which incurs an error polynomially related to k.[3] By eliminating the need for this lemma, the construction of Theorem 3.1 manages to achieve an exponentially small error parameter. Second, setting $m = 1$ in the construction of Theorem 3.1 makes it clear that the idea underlying both it and the [KZ] construction is simply a generalization of bitwise addition modulo 2—the parity function—which extracts 1 uniformly random bit whenever $k \geq 1$.

As discussed previously, this construction achieves output length only logarithmic in k. This is considerably worse than the output length of $k - 2\log(1/\varepsilon) - O(1)$ which we show to be possible both for extractors for OBFSs with $k > \log n$ (Section 4.1) and for ERFs (Section 4.2). The lower bound we prove in the following section shows why this is the case.

---

[3]This additional error was overlooked in [KZ], and their Theorem 1.2 erroneously claims an error exponentially small in k.

3.2 The lower bound

The extractor of Theorem 3.1 is a symmetric function; that is, its output is not sensitive to the order in which the input bits are arranged. We begin building our more general negative result by first showing that extractors for OBFSs with this property cannot have superlogarithmic output length. We then extend this idea to certain types of streaming algorithms with low space by showing that we can fix a subset of the input bits so that on the remaining bits, the computation of such an algorithm can be treated like a walk on a directed cycle that either stays in place or moves at each step. Our general negative result follows from the fact that the ending point of such a walk is a symmetric function.

The following lemma gives a limit on the output length of symmetric extractors for OBFSs. Intuitively, the lemma holds because the output of a symmetric function depends only on the number of 1s in its input string. Since this number can take only $k + 1$ different values for each given OBFS, the function cannot output more than $\log(k + 1)$ random bits.

Lemma 3.3. *Suppose that $X = L^{a,n}$ is an (n, k)-OBFS and that $f \colon \{0, 1\}^n \to \{0, 1\}^m$ is a symmetric function of the input bits in $[n] - L$. (That is, for every permutation $\pi \colon [n] \to [n]$ that fixes L, $f(x_{\pi(1)}, \ldots, x_{\pi(n)}) = f(x_1, \ldots, x_n)$.) Then $f(X) \approx_\varepsilon U_m$ implies that $m \le \log((k + 1)/(1 - \varepsilon))$.*

Proof. By the symmetry of f on the bits in $[n] - L$, the size of the support of $f(X)$ is at most $k + 1$. (The output depends only on the number of input bits in $[n] - L$ that equal 1.) Thus, the distance between $f(X)$ and U_m is at least $(M - k - 1)/M$. Together with $f(X) \approx_\varepsilon U_m$, this implies that $\varepsilon \ge (M - k - 1)/M$, which is equivalent to $m \le \log((k + 1)/(1 - \varepsilon))$. $\qquad\square$

We can use Lemma 3.3 to show that no symmetric function with large output length can be even a static ERF.

Proposition 3.4. *If a symmetric function $f \colon \{0, 1\}^n \to \{0, 1\}^m$ is a static (k, ε)-ERF then $m \le \log((k + 1)/(1 - \varepsilon))$.*

Proof. From Lemma 2.7, we have that for f to be a static ERF, it must satisfy, for all sets $L \in \binom{[n]}{n-k}$,

$$\mathop{\mathrm{E}}_{a \leftarrow U_{n-k}} \left[\Delta \left(f\left(L^{a,n} \right), U_m \right) \right] \le \varepsilon$$

It follows by averaging that there exists a set L and a string a such that $f(L^{a,n}) \approx_\varepsilon U_m$. Application of Lemma 3.3 to the source $L^{a,n}$ then yields the result. $\qquad\square$

Since every deterministic ε-extractor for (n, k)-OBFSs is a static (k, ε)-ERF and every adaptive (k, ε)-ERF is also a static (k, ε)-ERF, Proposition 3.4 applies to extractors for OBFSs and adaptive ERFs as well. Thus, Proposition 3.4 shows that constructions like that of Theorem 3.1 and that of [KZ] are optimal for symmetric functions.

However, there are many natural candidates for extraction from OBFSs that are similar to that of Theorem 3.1 but are not symmetric, such as the analogous random walk on a directed version of a 3-regular or 4-regular expander graph. For instance, we could try the graph with vertex set $\mathbb{F}_p$ (with p prime) where

7

the edge labelled 0 from vertex x goes to $x + 1$ and the edge labelled 1 goes to x^{-1} (or 0 in case $x = 0$). The undirected version of this graph is known to be an expander [Lub], so we might hope that with k random steps we can reach an almost uniform vertex even for $p = 2^{\Omega(k)}$ and thus output $\Omega(k)$ almost-uniform bits.

It turns out that such constructions do no better, as we now show by extending the above lower bound for extractors for OBFSs to a large class of small-source streaming algorithms. We start by defining the model of computation that we will assume.

Definition 3.5 (Streaming Algorithm)**.** A *streaming algorithm* $A \colon \{0,1\}^n \to \{0,1\}^m$ is given by a 5-tuple $(V, v_0, \Sigma^0, \Sigma^1, \varphi)$, where V is the state space, $v_0 \in V$ is the initial state, $\Sigma^0 = (\sigma_1^0, \ldots, \sigma_n^0)$ and $\Sigma^1 = (\sigma_1^1, \ldots, \sigma_n^1)$ are two sequences of functions from V to itself, and φ is a function from V to $\{0,1\}^m$. On an input sequence $(b_1, \ldots, b_n) \in \{0,1\}^n$, A computes by updating its state using the rule $v_{i+1} = \sigma_i^{b_i}(v_i)$. A's output is $A(b_1, \ldots, b_n) = \varphi(v_n)$. The function φ is called the *output function* of A, and the *space* of A is $\log |V|$.

We say that A is *forgetless* if and only if for every i at least one of either $\sigma_i{}^0$ or $\sigma_i{}^1$ is a permutation. (Thus, if the i-th bit is fixed to a certain value, A does not "forget" anything about its state when reading that bit.)

Forgetless streaming algorithms include random walks on 2-regular digraphs that are consistently labelled (meaning that the edges labelled b form a permutation, for each $b \in \{0,1\}$), like the graph on $\mathbb{F}_p$ mentioned above. However, forgetless streaming algorithms are more general in the sense that they can compute random walks in which each step of the walk is conducted on a different graph.

We now show that forgetless streaming algorithms with small space cannot compute extractors for OBFSs with large output length (for small k). This is our main result.

Theorem 3.6. *Suppose that $f \colon \{0,1\}^n \to \{0,1\}^m$ is a deterministic ε-extractor for (n,k)-OBFSs that can be computed by a forgetless streaming algorithm with space $s \leq \log{(n/k)}/k$. Then $m \leq \log{((k+1)/(1-\varepsilon))}$.*

Proof. Fix an ε-extractor for (n,k)-OBFSs $f \colon \{0,1\}^n \to \{0,1\}^m$ and let A be a forgetless streaming algorithm with space $s \leq \log{(n/k)}/k$ that computes f. To show that $m \leq \log{((k+1)/(1-\varepsilon))}$, we will first reduce to a special case in which we can make some simplifying assumptions about A. We will then construct an oblivious bit-fixing source X such that f is symmetric on the set of bit positions not fixed by X. This will allow us to apply Lemma 3.3 to obtain our result since f must map X close to uniform.

Reduction to the special case: Let Σ^0 and Σ^1 be the sequences of functions used by A, and let φ be its output function. We reduce to the special case that every element of Σ^0 is the identity.

Since A is forgetless, we can switch some of the functions σ_i^0 and σ_i^1 to make every function in Σ^0 a permutation while preserving the fact that A computes a (k, ε)-RF. (This corresponds to just negating some input bits.) This allows us to define a new sequence of functions $F = \{f_1, \ldots, f_n\}$ and a new output

8

function ψ by the following relations.

$$\sigma_i^0 \circ \cdots \circ \sigma_1^0 \circ f_i = \sigma_i^1 \circ \sigma_{i-1}^0 \circ \cdots \circ \sigma_1^0$$
$$\psi = \varphi \circ \sigma_n^0 \circ \cdots \circ \sigma_1^0$$

Then $(V, v_0, (\mathrm{id}, \mathrm{id}, \ldots, \mathrm{id}), (f_1, \ldots, f_n), \psi)$ can be verified to be a streaming algorithm that computes the same function as $(V, v_0, \Sigma^0, \Sigma^1, \varphi)$.

Constructing the source X: Letting $S = 2^s$, we can choose a set $F_1 \subset F$ of size at least n/S such that all the functions in F_1 map the initial state v_0 to some common state (call it v_1). We can then choose a set $F_2 \subset F_1$ of size at least n/S^2 such that all functions in F_2 map v_1 to some common state, which we call v_2. Continuing in this way, we obtain a set $F_k \subset F$ of size at least n/S^k and a sequence $(v_0, \ldots, v_k)$ with the property that every $f \in F_k$ satisfies $f(v_i) = v_{i+1}$ for $0 \le i < k$. By our assumption that $s \le \log(n/k)/k$, we have $|F_k| \ge n/S^k \ge k$. Thus, the set of bits at positions corresponding to functions in F_k contains some subset D of size k. We define X to be the oblivious bit-fixing source that has the bits in D un-fixed and the rest of the bits fixed to 0.

Obtaining the desired bound: For any string w in the support of X, f's output will be $\psi(v_{H(w)})$ where $H(w)$ is the Hamming weight of w. Therefore f is a symmetric function of the bits in positions not fixed by X. Since X contains at least k independent, uniformly random bits and f is a (k, ε)-resilient function, Lemma 3.3 yields $m \le \log((k+1)/(1-\varepsilon))$ as desired. $\qquad\square$

What does this theorem tell us about extraction in low-entropy settings? If we set $s = m \le k$ (as in the walk on the cycle of Theorem 3.1) then Theorem 3.6 implies that when $k < \sqrt{\log n - \log\log n}$ we are confined to output length $m \le \log((k+1)/(1-\varepsilon))$. In other words, the output length of $\Omega(\log k)$ offered by Theorem 3.1 is close to optimal for extractors in this model when $k < \sqrt{\log n}$.

We note here a separate, trivial space lower bound that applies even to the forgetful case: since streaming algorithms under our model cannot produce any output bits until they have read all the input bits, we have $s > m - 1$ when $\varepsilon < 1/2$. This bound can in fact be generalized to streaming algorithms that are allowed to output bits at any point in their computation by a simple adaptation of a space lower bound for strong extractors proven in [BRST]. The resulting lower bound says that $s \ge m - 4$ when $\varepsilon \le 1/8$ and $k \le n/2$ for extractors for OBFSs computable by any streaming algorithm.

4 Non-constructive results

We now turn to determining for what values of the entropy parameter k it is possible to achieve output length $m = \Omega(k)$ using the probabilistic method. Here we find that the results are roughly in agreement with our explicit lower bounds from the previous section. That is, a randomly chosen function $f \colon \{0,1\}^n \to \{0,1\}^m$ will almost always be an extractor for OBFSs with output length $m = \Omega(k)$ when k is larger than $\log n$, and this output length cannot be achieved using the probabilistic method when $k < \log n$.

We then show that random functions can do better in the more relaxed realm of exposure-resilient functions: a randomly chosen function is almost always a

static ERF with optimal output length for any k, and an *adaptive* ERF with optimal output length when k is larger than $\log \log n$.

Before we proceed, we state a Chernoff bound and a partial converse to it that we will use in proving these results. A sketch of the proof of Lemma 4.2 is given in the appendix.

Lemma 4.1 (A Chernoff bound [Hoe])**.** *Let $X_1, \ldots, X_t$ be independent random variables taking values in $[0, 1]$, and let $X = (\sum_i X_i)/t$ and $\mu = \mathrm{E}[X]$. Then for every $0 < \varepsilon < 1$, we have*

$$\Pr\left[|X - \mu| > \varepsilon\right] < 2e^{-t\varepsilon^2/4} \leq 2^{-\Omega\left(t\varepsilon^2\right)}$$

Lemma 4.2 (Anti-concentration bound)**.** *Let $X_1, \ldots, X_t$ represent the results of independent, unbiased coin flips, and let $X = (\sum_i X_i)/t$. Then for every $0 \leq \varepsilon \leq 1/2$, we have*

$$\Pr\left[\left|X - \frac{1}{2}\right| \geq \varepsilon\right] \geq 2^{-O\left(t\varepsilon^2\right)}$$

4.1 Deterministic extractors for OBFSs

Theorem 4.3 below, which follows from a straightforward application of the Chernoff bound stated in Lemma 4.1, shows that the probabilistic methods gives extractors for OBFSs with $k > \log n$. Theorem 4.4 then shows that $k > \log n$ is the best we can do using the probabilistic method.

Theorem 4.3. *For every $n \in \mathbb{N}$, $k \in [n]$, and $\varepsilon > 0$, a randomly chosen function $f \colon \{0, 1\}^n \to \{0, 1\}^m$ with $m \leq k - 2\log(1/\varepsilon) - O(1)$ and $k \geq \max\{\log(n - k), \log\log\binom{n}{k}\} + 2\log(1/\varepsilon) + O(1)$ is a deterministic ε-extractor for (n, k)-OBFSs with probability at least $1 - 2^{-\Omega(K\varepsilon^2)}$, where $K = 2^k$.*

Proof. Fix an (n, k)-OBFS X. Choosing the function f consists of independently assigning a string in $\{0, 1\}^m$ to each string in the support of X. In order for f to map X close to uniform, we need to have chosen it such that, for every fixed statistical test $T \subset \{0, 1\}^m$, the fraction of strings in X mapped by f into T is very close to the density of T in $\{0, 1\}^m$. This is expressed formally by the condition below.

$$\left|\frac{|f^{-1}(T)|}{2^k} - \frac{|T|}{2^m}\right| \leq \varepsilon$$

Now fix one specific test $T \subset \{0, 1\}^m$. For each string w in the support of X, define the indicator variable I_w to be 1 if $f(w) \in T$ and 0 otherwise. Then Lemma 4.1 (our Chernoff bound) applied to $\left(\sum_w I_w\right)/2^k = |f^{-1}(T)|/2^k$ shows that f fails the condition above with probability at most $2^{-\Omega(K\varepsilon^2)}$.

There are 2^M possible tests $T \subset \{0, 1\}^m$ (where $M = 2^m$). A union bound over all these tests therefore gives that the probability that f fails to map X to within ε of uniform is at most $2^{M-\Omega(K\varepsilon^2)}$. We can perform a similar union bound over the possible choices of the source X: there are $\binom{n}{k}N/K$ such sources, yielding that the probability that f is not a (k, ε)-RF is at most

$$\binom{n}{k}\frac{N}{K}\, 2^{M-\Omega\left(K\varepsilon^2\right)} = 2^{-\Omega\left(K\varepsilon^2\right)}$$

10

provided $K \geq \max\{\log\left(\frac{N}{K}\right), \log\binom{n}{k}\}c/\varepsilon^2$ for a sufficiently large constant c and $M \leq c'K\varepsilon^2$ for a sufficiently small constant c'. Taking logarithms gives the result. $\qquad\square$

The $\max\{\log(n-k), \log\log\binom{n}{k}\}$ term in the statement of Theorem 4.3 is always at most $\log n$, so the theorem always holds when $k \geq \log n + 2\log(1/\varepsilon) + O(1)$, as discussed earlier. In the following theorem, we prove a limitation on the extraction properties of random functions which shows that this bound on k is in fact nearly tight.

Theorem 4.4. *There is a constant c such that for every $n \in \mathbb{N}$, $k \in [n]$, and $\varepsilon \in [0, 1/2]$ satisfying $k \leq \log(n-k) + 2\log(1/\varepsilon) - c$, a random function $f\colon \{0,1\}^n \to \{0,1\}$ will fail to be a deterministic ε-extractor for (n,k)-OBFSs with probability at least $1 - 2^{-\sqrt{N/K}}$, where $N = 2^n$ and $K = 2^k$.*

Proof. Fix an input size n and a set L of $n-k$ fixed bits (say, $L = [n-k]$). To say that f an ε-extractor for (n,k)-OBFSs is to say that all 2^{n-k} sets S of the form $L^{*,n}$ satisfy the following condition.

$$\left|\Pr_{w \leftarrow S}[f(w) = 1] - \frac{1}{2}\right| \leq \varepsilon$$

Since $f(w)$ is chosen independently for each string $w \in S$, we can use the converse of our Chernoff bound (Lemma 4.2) to say that the probability that f satisfies this condition for a fixed set S is at most $1 - 2^{-O(K\varepsilon^2)}$, where $K = 2^k = |S|$.

Since there are N/K subsets of the form $L^{*,n}$ and they are disjoint, the probability that f will fail the above condition on none of them (which is necessary condition for f to be a resilient function) is at most

$$\left(1 - 2^{-O\left(K\varepsilon^2\right)}\right)^{N/K}$$

This probability is at most $2^{-\sqrt{N/K}}$ provided that $N/K \geq 2^{CK\varepsilon^2}$ for a sufficiently large constant $C = 2^c$. Taking logarithms twice completes the proof. $\qquad\square$

Theorem 4.4 does not establish that extractors for OBFSs with the stated parameters do not exist; indeed, as mentioned earlier, the parity function (i.e. $f(x_1, \ldots, x_n) = \oplus x_i$) is a perfect resilient function for even $k = 1$. What the theorem does show, however, is that $k \approx \log n$ represents a critical point below which these extractors become very rare. This seems consistent with the lower bound on k proven in Theorem 3.6.

4.2 Exposure-resilient functions

We now show that probabilistically constructing exposure-resilient functions is easier than constructing extractors for OBFSs. This is because, while the adversary can choose input sources in the extractor setting, here it can only expose them. The probabilistic constructions of static and adaptive ERFs both proceed by counting the number of adversaries that must be fooled and then applying Lemma 4.5 (below), which is an upper bound on the probability that a randomly chosen function will fail to fool a fixed adversary. This lemma applies

equally both to static and adaptive adversaries; the difference in achievable parameters between static and adaptive ERFs therefore stems solely from the fact that there are many more adversaries in the adaptive setting.

Lemma 4.5. *Let* $A\colon \{0,1\}^n \to \{0,1\}^*$ *be an algorithm that reads at most d bits of its input, let $\varepsilon > 0$, and choose a function $f\colon \{0,1\}^n \to \{0,1\}^m$ uniformly at random with $m = n - d - 2\log(1/\varepsilon) - O(1)$. Then f will fail to satisfy*

$$(A(U_n), f(U_n)) \approx_\varepsilon (A(U_n), U'_m)$$

with probability at most $2^{-\Omega(N\varepsilon^2)}$, where $N = 2^n$.

Proof. Lemma 2.9 allows us to assume without loss of generality that A adaptively reads d bits and outputs them in the order that they were read. Under this assumption, we have $(A(U_n), U'_m) = U_{d+m}$. We therefore need only to bound the probability that $(A(U_n), f(U_n))$ is far from U_{d+m}.

Fix a statistical test $T \subset \{0,1\}^d \times \{0,1\}^m$. In order for $(A(U_n), f(U_n))$ to pass this specific test of uniformity, we need f to satisfy

$$\left| \Pr\left[(A(U_n), f(U_n)) \in T\right] - \frac{|T|}{2^{d+m}} \right| \le \varepsilon \tag{4.1}$$

For every $w \in \{0,1\}^n$, define I_w to be 1 if $(A(w), f(w)) \in T$ and 0 otherwise, and notice that $\Pr[(A(U_n), f(U_n)) \in T] = \frac{1}{2^n}\sum_w I_w$. For $x \in \{0,1\}^d$, let T_x denote $T \cap (\{x\} \times \{0,1\}^m)$. Then, for a fixed w, the expectation of I_w over the choice of f is exactly $|T_{A(w)}|/2^m$, and so by the regularity of A the expectation of $\frac{1}{2^n}\sum_w I_w$ over the choice of f is $|T|/2^{d+m}$. A Chernoff bound (Lemma 4.1) then gives that the probability over the choice of f that Equation (4.1) is not satisfied is at most $2^{-\Omega(N\varepsilon^2)}$.

Since there are 2^{DM} possible choices of T in the above analysis (where $D = 2^d$, $M = 2^m$), a union bound shows that the probability that $(A(U_n), f(U_n))$ will fail one or more of them is at most $2^{DM}2^{-\Omega(N\varepsilon^2)} = 2^{-\Omega(N\varepsilon^2)}$ if $m = n - d - 2\log(1/\varepsilon) - c$ for a sufficiently large constant c. $\square$

Having established that a random function will tend to fool a fixed adversary, we now establish the existence of static and adaptive exposure-resilient functions. In both cases, we do so by taking a union bound over all potential adversaries and applying Lemma 4.5. Thus, the parameters achieved are those that bring the number of adversaries to below $2^{N\varepsilon^2}$.

Theorem 4.6. *For every $n \in \mathbb{N}$, $k \in [n]$, and $\varepsilon \ge c\sqrt{n/2^n}$ where c is a universal constant, a randomly chosen function $f\colon \{0,1\}^n \to \{0,1\}^m$ with $m \le k - 2\log(1/\varepsilon) - O(1)$ is a static (k,ε)-ERF with probability at least $1 - 2^{-\Omega(N\varepsilon^2)}$, where $N = 2^n$.*

Proof. Every static adversary that tries to distinguish the output of f from uniform is an algorithm $A\colon \{0,1\}^n \to \{0,1\}^{n-k}$ that reads exactly $n - k$ bits of its input. We can therefore apply Lemma 4.5 with $d = n - k$ to get that the probability that f will fail to fool any one adversary is at most $2^{-\Omega(N\varepsilon^2)}$. Taking a union bound over the $\binom{n}{k}$ possible adversaries, we get that the probability that f will not fool all adversaries is at most

$$\binom{n}{k}2^{-\Omega(N\varepsilon^2)} \le N2^{-\Omega(N\varepsilon^2)} = 2^{-\Omega(N\varepsilon^2)}$$

12

where the final equality is given by the constraint on ε. $\qquad\square$

Counting the number of adversaries in the adaptive setting is a bit more work, but Lemma 2.9 from our preliminaries simplifies this task.

Theorem 4.7. *For every $n \in \mathbb{N}$, $k \in [n]$, and $\varepsilon > 0$, a randomly chosen function $f \colon \{0,1\}^n \to \{0,1\}^m$ with $m \le k - 2\log(1/\varepsilon) - O(1)$ and $k \ge \log\log n + 2\log(1/\varepsilon) + O(1)$ is an adaptive (k, ε)-ERF with probability at least $1 - 2^{-\Omega(N\varepsilon^2)}$, where $N = 2^n$.*

Proof. The proof is identical to that of Theorem 4.6 except that we have to count the number of adaptive adversaries. We do so below.

First we note that Lemma 2.9 implies that if f fools all adaptive adversaries that output the bits they read as they read them, then f fools all adaptive adversaries. We therefore only need to count this smaller set of adversaries. The process by which such an adversary chooses which bits to request can be modelled by a decision tree of depth $n - k - 1$ whose internal nodes are labelled by elements of $[n]$. Since the number of nodes in such a tree is $2^{n-k-1} - 1 < N/2K$, where $N = 2^n$ and $K = 2^k$, we can bound the total number of trees—and therefore adversaries—by $n^{N/2K}$.

Proceeding with the same kind of union bound as in the proof of Theorem 4.6, we see that the probability that f will not fool all adaptive adversaries is at most $n^{N/2K}2^{-\Omega(N\varepsilon^2)} = 2^{-\Omega(N\varepsilon^2)}$, provided that $K \ge (c\log n)/\varepsilon^2$ for a sufficiently large constant c. Taking logarithms yields the theorem. $\qquad\square$

5 Future work

The general question of whether there exist resilient functions with large output length in the low-entropy range studied here is still unresolved.

Open Question 1. Does there exist, for all $n \in \mathbb{N}$ and some growing function $0 < k(n) < \log n$, a deterministic ε-extractor for $(n, k(n))$-OBFSs with output length $m = \Omega(k(n))$ and ε constant?

Theorem 3.6 shows that to resolve this question in the positive direction requires a function that is either not computable by a forgetless streaming algorithm or uses a considerable amount of space. In the other direction, an interesting step towards a negative result would be to at least remove the forgetlessness condition from the space lower bound proven in that theorem.

We can ask an analogous question for the case of adaptive ERFs with $k < \log\log n$.

Open Question 2. Does there exist, for all $n \in \mathbb{N}$ and some growing function $0 < k(n) < \log\log n$, an adaptive $(k(n), \varepsilon)$-ERF with output length $m = \Omega(k(n))$ and ε constant?

In this case, we cannot even rule out the possibility that a more clever use of the probabilistic method will resolve this question positively. Thus, a first step toward a negative result might be to prove an analogue to Theorem 4.4 that shows that adaptive ERFs with near-optimal output length become very rare when $k < \log\log n$.

A third open problem arising from this work is that of finding an explicit construction of a static ERF with the parameters achieved using the probabilistic method in Theorem 4.6. Currently, an output length of $\Omega(k)$ is achieved in [DSS] using strong extractors, but the construction works only when $k > \log n$. For k smaller than $\log n$, there is no known construction of a static ERF that is not also an RF, making the construction of Theorem 3.6 the current state of the art. This leaves us with the following open question:

Open Question 3. Does there exist, for all $n \in \mathbb{N}$ and some growing function $0 < k(n) < \log n$, an explicit static $(k(n), \varepsilon)$-ERF with output length $m = \Omega(k(n))$ and ε constant?

References

[BBR] Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM Journal on Computing*, 17(2):210–229, April 1988.

[BKS$^+$] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: new constructions of condensers, Ramsey graphs, dispersers, and extractors. In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 1–10. ACM, New York, 2005.

[BRST] Z. Bar-Yossef, O. Reingold, R. Shaltiel, and L. Trevisan. Streaming computation of combinatorial objects. In *Proceedings of 17th Annual IEEE Conference on Computational Complexity (CCC '02)*, pages 165–174, 2002.

[CDH$^+$] R. Canetti, Y. Dodis, S. Halevi, E. Kushilevitz, and A. Sahai. Exposure-resilient functions and all-or-nothing transforms. In *Advances in Cryptography – EUROCRYPT 2000*, volume 1807/2000, pages 453–469, 2000.

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

[CGH$^+$] Benny Chor, Oded Goldreich, Johan Hastad, Joel Friedman, Steven Rudich, and Roman Smolensky. The bit extraction problem or t-resilient functions (preliminary version). In *26th Annual Symposium on Foundations of Computer Science*, pages 396–407, Portland, Oregon, 21–23 October 1985. IEEE.

[Dia] P. Diaconis. Group representations in probability and statistics. In *Lecture Notes–Monograph Series 11, Institute of Mathematical Statistics*, 1988. Hayward, CA.

[DSS] Y. Dodis, A. Sahai, and A. Smith. On perfect and adaptive security in exposure-resilient cryptography. In *Advances in Cryptography – EUROCRYPT 2001*, volume 2045, pages 301–324, 2001.

[GRS] A. Gabizon, R. Raz, and R. Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. *SIAM Journal on Computing*, 36(4):1072–1094, 2006.

[Hoe] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[KZ] J. Kamp and D. Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. *SIAM Journal on Computing*, 36(5):1231–1247, 2003.

[Lub] A. Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*. Birkhauser, Basel, 1994.

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

[Res] Y. Reshef. On resilient and exposure-resilient functions. Undergraduate Thesis, Harvard College, 2009.

[TV] Luca Trevisan and Salil Vadhan. Extracting randomness from samplable distributions (extended abstract). In *41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000)*, pages 32–42. IEEE Comput. Soc. Press, Los Alamitos, CA, 2000.

[Vaz] Umesh V. Vazirani. Strong communication complexity or generating quasirandom sequences from two communicating semirandom sources. *Combinatorica*, 7(4):375–392, 1987.

A   Proof of Claim 3.2

Claim 3.2 (Restated). *Let W_k be the distribution on the vertices of $\mathbb{Z}/M\mathbb{Z}$ (where $M = 2^m$) obtained by beginning at 0 and adding 1 or 0 with equal probability k times. Then*

$$\Delta(W_k, U_m)^2 \leq \frac{e^{-k\pi^2/2M^2}}{2\left(1 - e^{-3k\pi^2/2M^2}\right)}$$

Proof. Consider $\mathbb{Z}/M\mathbb{Z}$ as an additive group, and let P be the probability distribution on $\mathbb{Z}/M\mathbb{Z}$ that equals 0 with probability $1/2$ and 1 otherwise. Then the distribution on $\mathbb{Z}/M\mathbb{Z}$ after k steps of our random walk is P^{*k}, the k-th convolution of P with itself. We will bound the distance $\Delta(P^{*k}, U_m)$ using Lemma 1 from Chapter 3 of [Dia], which we have reproduced below and specialized to the case of finite abelian groups (whose irreducible representations are exactly the homomorphisms to $\mathbb{C}^*$).

Lemma A.1. *Let Q be a probability distribution on a finite abelian group G, and let U be the uniform distribution on the elements of G. For every group homomorphism $\chi : G \to \mathbb{C}^*$ (where $\mathbb{C}^*$ is the non-zero complex numbers under*

multiplication), let $\hat{Q}(\chi) = \sum_{g \in G} Q(g)\chi(g)$ denote the Fourier transform of Q at χ. Then

$$(\Delta(Q, U))^2 \le \frac{1}{4} \sum_{\chi \ne \mathbb{1}} |\hat{Q}(\chi)|^2$$

where the sum is over all the non-trivial homomorphisms $\chi : G \to \mathbb{C}^*$, and $|\cdot|$ denotes the complex norm.

Applying this lemma with $Q = P^{*k}$, and noting that $\widehat{P^{*k}}(\chi) = (\hat{P}(\chi))^k$ (by the Convolution Identity $\widehat{A * B}(\chi) = \hat{A}(\chi)\hat{B}(\chi)$), gives that

$$\left(\Delta(P^{*k}, U)\right)^2 \le \frac{1}{4} \sum_{\chi \ne \mathbb{1}} |\hat{P}(\chi)|^{2k}$$

We then use the fact that in the case of $G = \mathbb{Z}/M\mathbb{Z}$, the non-trivial homomorphisms to $\mathbb{C}^*$ are given by $\chi_j(1) = e^{2\pi i j/M}$ for $j \in 1, 2, \ldots, M-1$. This gives the following upper bound for the squared distance from uniform after k random steps.

$$\frac{1}{4} \sum_{j=1}^{M-1} \left| \frac{1}{2} + \frac{1}{2} e^{\frac{2\pi i j}{M}} \right|^{2k}$$

To bound this sum, we first note that

$$\left| \frac{1}{2} + \frac{1}{2} e^{ix} \right|^2 = \frac{1}{2} + \frac{1}{2}\cos(x) \le e^{-x^2/8}$$

for $x \in [0, \pi]$. This, together with the fact that $M = 2^m$ is even, allows us to write

$$
\begin{aligned}
\frac{1}{4} \sum_{j=1}^{M-1} \left| \frac{1}{2} + \frac{1}{2} e^{\frac{2\pi i j}{M}} \right|^{2k} &= \frac{1}{2} \sum_{j=1}^{(M-2)/2} \left(\frac{1}{2} + \frac{1}{2} \cos\left(\frac{2\pi j}{M} \right) \right)^k \\
&\le \frac{1}{2} \sum_{j=1}^{(M-2)/2} e^{-k\pi^2 j^2/2M^2} \\
&\le \frac{1}{2} e^{-k\pi^2/2M^2} \sum_{j=1}^{\infty} e^{-k\pi^2(j^2-1)/2M^2} \\
&\le \frac{1}{2} e^{-k\pi^2/2M^2} \sum_{j=0}^{\infty} e^{-3k\pi^2 j/2M^2} \\
&= \frac{e^{-k\pi^2/2M^2}}{2\left(1 - e^{-3k\pi^2/2M^2}\right)}
\end{aligned}
$$

which is the desired result. $\qquad\square$

(The above proof is essentially the same as the proof of Theorem 1 in Chapter 3 of [Dia]. The two proofs differ only in the details of the computations.)

B Proof sketch of Lemma 4.2

Lemma 4.2 (Restated). *Let $X_1, \ldots, X_t$ represent the results of independent, unbiased coin flips, and let $X = (\sum_i X_i)/t$. Then for every $0 \le \varepsilon \le 1/2$, we have*

$$\Pr\left[\left|X - \frac{1}{2}\right| \ge \varepsilon\right] \ge 2^{-O(t\varepsilon^2)}$$

Proof Sketch. We address three separate cases: $0 \le \varepsilon < \frac{1}{4\sqrt{t}}$, $\frac{1}{4\sqrt{t}} \le \varepsilon < \frac{1}{5}$, and $\frac{1}{5} \le \varepsilon \le \frac{1}{2}$. In the first case, we upper-bound the probability that $\left|X - \frac{1}{2}\right| < \varepsilon$ using the fact that no term of the binomial distribution exceeds $\sqrt{2/\pi t}$ in probability mass. In the second case, we set $\beta = \frac{1}{2} + 2\varepsilon$ and use Stirling's approximation to lower-bound the probability by

$$
\begin{aligned}
\lfloor \varepsilon t \rfloor \cdot \binom{t}{\lfloor \beta t \rfloor}/2^t &\ge \lfloor \varepsilon t \rfloor \cdot \frac{1}{\sqrt{t}} \frac{1}{2^t \beta^{\beta t}(1-\beta)^{(1-\beta)t}} \\
&\ge 2^{-O(t\varepsilon^2)}
\end{aligned}
$$

where the first inequality is from Stirling's approximation. In the third case, we just lower-bound the probability by 2^{-t}. $\qquad\square$