On the Existence of Seedless Condensers: Exploring the Terrain

Eshan Chattopadhyay* Cornell University eshan@cs.cornell.edu

Mohit Gurumukhani* Cornell University mgurumuk@cs.cornell.edu

Noam Ringach † Cornell University nomir@cs.cornell.edu

Abstract

While the existence of randomness extractors, both seeded and seedless, has been studied for many sources of randomness, currently, not much is known regarding the existence of seedless condensers in many settings. Here, we prove several new results for seedless condensers in the context of three related classes of sources: Non-Oblivious Symbol Fixing (NOSF) sources, one-sided NOSF (oNOSF) sources (originally defined as SHELA sources in [AORSV, EUROCRYPT'20]), and almost Chor-Goldreich (CG) sources as defined in [DMOZ, STOC'23]. We will think of these sources as a sequence of random variables $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ on ℓ symbols where at least g out of these ℓ symbols are "good" (i.e., have some min-entropy requirement), denoted as a (g, ℓ)-source, and the remaining "bad" $\ell - g$ symbols may adversarially depend on these g good blocks. The difference between each of these sources is realized by restrictions on the power of the adversary, with the adversary in NOSF sources having no restrictions.

Prior to our work, the only known seedless condenser upper or lower bound in these settings is due to [DMOZ, STOC'23], where they explicitly construct a seedless condenser for a restricted subset of (g, ℓ)-adversarial CG sources.

The following are our main results concerning seedless condensers for each of these sources.

1. oNOSF sources
   (a) When $g \leq \ell/2$, we prove that condensing with error 0.99 above rate $\frac{1}{\lfloor \ell/g \rfloor}$ is impossible. In fact, we show that this is tight.
   (b) Quite surprisingly, for $g > \ell/2$, we show the existence of excellent condensers for uniform oNOSF sources. In addition, we show the existence of similar condensers for oNOSF sources with only logarithmic min-entropy. Our results are based on a new type of two-source extractors, called *output-light two-source extractors*, that we introduce and prove the existence of.
2. Adversarial CG sources
   (a) We observe that uniform adversarial CG sources are equivalent to uniform oNOSF sources and consequently inherit the same results.
   (b) We show that one cannot condense beyond the min-entropy gap of each block or condense low min-entropy CG sources above rate 1/2.
3. NOSF sources
   (a) We show that condensing with constant error above rate $\frac{g}{\ell}$ is impossible for uniform NOSF sources for any g and ℓ, thus ruling out the possibility of any non-trivial condensing. This shows an interesting distinction between NOSF and oNOSF sources.

These results make progress on several open question from [DMOZ, STOC'23], [AORSV, EUROCRYPT'20], and [KN, RANDOM'23].

*Supported by a Sloan Research Fellowship and NSF CAREER Award 2045576.

†Supported by NSF GRFP grant DGE – 2139899, NSF CAREER Award 2045576 and a Sloan Research Fellowship.

Contents

1 Introduction

One of the most fruitful lines of research in computer science has been that of randomness. From the traditionally more applied areas of algorithm design (e.g., Monte Carlo simulations), error-correcting codes and cryptography to the more theoretical areas of property testing, combinatorics, and circuit lower bounds, randomness has played a key role in seminal discoveries. In many of these works, the use of high-quality random bits, or alternatively, a way to convert low-quality randomness into high-quality randomness, is essential. In cryptography, the authors of [DOPS04] showed that high-quality randomness is essential for tasks such as bit commitment schemes and secure two-party computation. On the other hand, being able to extract uniform bits from low-quality randomness allows us to simulate randomized algorithms [Zuc90].

In most use-cases, randomness takes the form of uniformly random bits. These motivated the construction of randomness extractors,[1] functions that take low-quality randomness (which we often like to think of as natural processes) and convert it into uniformly random bits. It is impossible to extract from the class of all sources and so extractors are constructed with respect to a restricted class of sources.

A number of works [SV86, CG88, Zuc90, RVW04] have shown that deterministic extraction is impossible for many natural classes of randomness sources. The question that arises for such sources then is whether any improvement to their randomness can be made. That is, while it may not be possible to convert a source into uniform bits, maybe it is possible to condense a source into another source with a higher density of randomness. The central focus of our paper is in understanding the possibility of condensing for various natural models of weak sources where it is known that extraction is impossible.

We first introduce the way that we measure randomness and the notions of extractors and condensers. The notion of randomness that is standard in this line of work is that of min-entropy. For a source $\mathbf{X}$ on n bits, we define its *min-entropy* as $H_\infty(\mathbf{X}) = \min_{x\in\{0,1\}^n}\{-\log(\Pr[\mathbf{X}=x])\}$. A source $\mathbf{X}$ over n bits with min-entropy at least k is called an (n,k)-*source*. Given any two distributions $\mathbf{X}$ and $\mathbf{Y}$ on $\{0,1\}^n$, we define their statistical distance or total-variation (TV) distance as $|\mathbf{X}-\mathbf{Y}| = \max_{Z\subseteq\{0,1\}^n} |\Pr_{x\sim\mathbf{X}}[x\in Z] - \Pr_{y\sim\mathbf{Y}}[y\in Z]|$. We also need the notion of *smooth min-entropy*: for a source $\mathbf{X}$ on $\{0,1\}^n$, it=s smooth min-entropy with smoothness parameter ε is $H^\varepsilon_\infty(\mathbf{X}) = \max_{\mathbf{Y}:|\mathbf{X}-\mathbf{Y}|\le\varepsilon} H_\infty(\mathbf{Y})$. Conceptually, smooth min-entropy asks that the source we are looking at be ε-close in TV-distance to some other source with the desired amount of min-entropy. We are now in a position to define randomness extraction and condensing.

Definition 1.1. *Let $\mathcal{X}$ be a family of distributions over $\{0,1\}^n$. A function $\mathsf{Ext}:\{0,1\}^n\to\{0,1\}^m$ is an* extractor *for $\mathcal{X}$ with error $\varepsilon>0$ if for all $\mathbf{X}\in\mathcal{X}$ we have $|\mathsf{Ext}(\mathbf{X})-\mathbf{U}_m|\le\varepsilon$.*

For extractors to exist, we require all sources in $\mathcal{X}$ to have entropy. When each source in $\mathcal{X}$ is an (n,k)-source, we say that Ext is a (k,ε)-extractor for $\mathcal{X}$. For some classes, an extractor may not exist (such as for the class of all $(n,n-1)$-sources). Consequently, we turn to the looser requirements of condensing.

Definition 1.2. *For a family of distributions $\mathcal{X}$ over $\{0,1\}^n$, a function $\mathsf{Cond}:\{0,1\}^n\to\{0,1\}^m$ is a* condenser *with error $\varepsilon\ge 0$ if for all $\mathbf{X}\in\mathcal{X}$ we have that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X}))/m \ge H_\infty(\mathbf{X})/n$. We say that* Cond *has* entropy gap Δ *if $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X}))\ge m-\Delta$. When $\mathcal{X}$ is the class of (n,k)-sources and $k'=m-\Delta$, we say that* Cond *is a (k,k',ε)-condenser.*

[1] In this paper, when we mention extractors/condensers, we usually mean seedless extractors/condensers.

Unfortunately, even this notion is too strong as we cannot condense with error ε from the class of all (n, k)-sources so that the output entropy rate is larger than k/n. [2] We thus study condensing from classes of sources which have some additional structure along with a min-entropy requirement. In this paper, we explore the possibility of condensing from three related models of weak sources. These models, some of which have been studied since the 1980s, are very general and well-motivated by practical considerations.

The rest of the introduction is organized as follows: In Section 1.1, we present the case that condensers have many applications and are hence a natural direction of study, in particular when extraction is not feasible. In Section 1.2, we discuss the models of weak sources that we study, present relevant prior work on these models, and discuss our results for each of them.

1.1 The utility of condensing

We present two viewpoints in motivating our study of condensers. We compare what is possible via condensing in contrast to extracting and consider the utility of condensing for simulating BPP algorithms.

1.1.1 Condensing vs. extracting

Condensers exist in many scenarios when it can be provably shown that deterministic extraction is not possible. Thus, they allow us to obtain randomness that is more useful than what we began with in cases where extracting uniform bits is impossible. One significant example is that of Santha-Vazirani (SV) sources [SV86] and their generalization, Chor-Goldreich (CG) sources [CG88].

Informally, an SV source is a string of random bits such that the conditional distribution of each bit on the bits that come before it is guaranteed to have some minimum amount of min-entropy; a CG source generalizes this to allow each bit to instead be a symbol in $\{0,1\}^n$. It is well known that deterministic extraction is impossible for both SV and CG sources [SV86, CG88, RVW04]. The recent result of [DMOZ23] with regards to condensing from CG sources stands in contrast to these impossibility results for extraction. Other examples of sources for which deterministic extraction is not possible while deterministic condensing are the *somewhat dependent* sources of [BGM22] and block sources [BCDT19].

We briefly mention that seeded condensers are known to achieve parameters unattainable by seeded extractors [RT00]. Further, seeded condensers have been extremely useful in excellent constructions of seeded extractors [RSW06, Zuc07, TUZ07, GUV09].

1.1.2 Condensing for simulating BPP algorithms

Condensers with small entropy gap are useful in simulating randomized algorithms with low overhead [DMOZ23]. There are two ways one can go about this. First, there exists an explicit seeded extractor Ext with seed length $d = O(\log(\Delta))$ that can extract from any (n, k)-source $\mathbf{X}$ with entropy gap $\Delta = n - k$ [RVW02]. Then, to simulate a randomized algorithm $\mathcal{A}$ in BPP, we instead sample $x \sim \mathbf{X}$ and take the majority of the output of $\mathcal{A}$ on $\{\mathsf{Ext}(x, s)\}$ where we cycle over all seeds s [Vad12].

[2] Assuming $m \leq n$, the output entropy can be shown to be most $k + m - n + \log(1/(1-\varepsilon))$. See Lemma 5.20 for a proof of this fact.

For some applications in randomized protocols, cryptography and interactive proofs, one cannot afford to compute Ext all 2^d times by cycling through every seed [BDKPPSY11, DRV12, DY13, DPW14]. Alternatively, we can simulate $\mathcal{A}$ using a "one-shot" method in which we do not iterate over all seeds. A result from [DPW14] allows us to simulate $\mathcal{A}$ on the condensed source $\mathbf{X}$ (with entropy gap Δ) by reducing the error of $\mathcal{A}$ to $2^{-\Delta-1} \cdot \varepsilon$ and then using $\mathbf{X}$ directly to simulate random bits in $\mathcal{A}$. Such a simulation will have error ε.

1.2 Models of weak sources and our results

We consider three adversarial classes of sources motivated by weak sources that appear in practice as well as in various cryptographic settings. These sources are natural generalizations of the well-studied independent sources wherein we allow for an adversarial dependence between sources. Changing the scope and power of the adversary in natural ways gives rise to the three different classes of sources that we will consider.

The three randomness sources that we focus on in this work are all composed of blocks of bits, known as symbols, which vary in how they are permitted to relate to other symbols in the source. In these definitions, we will consider sources $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ of length ℓ where each $\mathbf{X}_i \in \{0,1\}^n$ is called a block. Generally, we will term blocks that have some minimum amount of randomness "good" and blocks that are chosen by an adversary as "bad". Next, we discuss these three models of weak sources, presenting what was known from prior work and our new results for each of these models.

1.2.1 One-sided non-oblivious symbol fixing sources

The first class of adversarial sources that we will define is that of *one-sided non-oblivious symbol fixing (oNOSF) sources*. While these are a restriction of general NOSF sources, which we will define later, we introduce them first since they have the weakest adversary and, consequently, the strongest positive results. Formally, we define oNOSF sources as follows.

Definition 1.3 (oNOSF sources, [AORSV20])**.** *A (g, ℓ, n, k)-*oNOSF source $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ *on $(\{0,1\}^n)^\ell$ is such that g out of the ℓ blocks are independently sampled (n,k)-sources (i.e., good), and the remaining $\ell - g$ bad blocks only depend on blocks with smaller indices (i.e., to their left).*

If $k = n$, we call $\mathbf{X}$ a *uniform (g,ℓ)-oNOSF source*. oNOSF sources form a natural class of sources to study when an adversary is working in real time and cannot predict the future. One such real-world example is that of blockchains. From [GKL15, PSS17], we know that in a sequence of blockchains, there will be some fraction of blocks that are chosen by honest players. Moreover, since these honest players are not working together, their chosen blocks may be considered as independent, fulfilling the requirement for good blocks for oNOSF sources. The adversarial players, on the other hand, can only see blocks added to the blockchain thus far and do not know which values of blocks will be added in the future, fulfilling the requirements for bad blocks for oNOSF sources. For more uses of oNOSF sources, see [AORSV20].

Previous work Prior to our work, the only results for condensing or extracting from oNOSF sources are due to [AORSV20]. In [AORSV20], the authors study Somewhere Honest Entropic Look Ahead (SHELA) sources, which are exactly convex combinations of oNOSF sources (see Proposition 4.15). They (1) transform (not uniform) oNOSF sources into uniform NOSF sources

and (2) show that for any $\gamma \in (0,1)$, there exists an ℓ such that extraction is not possible for $(\lfloor \gamma \ell \rfloor, \ell)$-oNOSF sources.

Our results We prove the existence of condensers with excellent parameters when the majority of the blocks of a uniform oNOSF source are good.

Theorem 1.4 (Informal version of Theorem 6.15)**.** *For all constant g and ℓ such that $g > \ell/2$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g,ℓ)-oNOSF source $\mathbf{X}$, we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m))$ where $m = n$ and $\varepsilon = 1/\Omega(m^{1/4})$.*

For our construction, we introduce a new type of two-source extractor[3] that we call a *R-output-light* two-source extractor. Such a two-source extractor $2\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ satisfies the additional guarantee that each output $z \in \{0,1\}^m$ can only be produced by R inputs $x \in \{0,1\}^{n_1}$ in the first source (see Definition 6.12 for the formal definition). The existence of such extractors is not obvious, and we show that output-light two-source extractors exist with strong parameters in Lemma 6.13. Our proof uses the observation that R-output-lightness is implied by the notion of R-invertibility, which simply bounds $\|\mathsf{Cond}\|_\infty$ by R (see Definition 6.20 for a formal definition). Incidentally, this latter notion has been recently used in a different context, to construct explicit random access linear codes with constant rate and distance [CM24]. While we are unable to explicitly construct such output-light two-source extractors, we do construct an explicit output-light *seeded* extractor, which we use to condense from uniform $(2,3)$-oNOSF sources and more (see Appendix A).

In fact, we can achieve a stronger result and show existence of condensers for oNOSF sources with only logarithmic min-entropy guarantee in the good blocks.

Theorem 1.5 (Informal version of Corollary 6.16)**.** *For any constant g and ℓ such that $g > \ell/2+1$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any (g,ℓ,n,k)-oNOSF source $\mathbf{X}$ with $k \geq 2\log(n)$ we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m))$ where $m = \Omega(k)$ and $\varepsilon = 1/\Omega(m^{1/4})$.*

To accomplish this, we transform logarithmic min-entropy oNOSF sources to uniform oNOSF sources and then apply the condenser for uniform oNOSF sources. We transform logarithmic min-entropy oNOSF sources to uniform oNOSF sources by modifying the construction of a somewhere-extractor for high min-entropy SHELA sources by [AORSV20]. These results imply that oNOSF sources can be useful for low overhead simulation of BPP algorithms. Furthermore, taken in tandem with the result that for all $\gamma > 0$ there exists a large enough ℓ such that one cannot extract from uniform $(\lfloor \gamma \ell \rfloor, \ell)$-oNOSF sources from [AORSV20], we have shown that oNOSF sources are one of the natural classes of sources that admit seedless condensing but not seedless extraction. This adds oNOSF sources to the short list of such natural sources mentioned in Section 1.1.1.

In contrast, condensing in the regime of $g \leq \ell/2$ is more nuanced: some non-trivial condensing beyond rate $\frac{g}{\ell}$ is possible provided g does not divide ℓ, but condensing to a significantly higher rate is not possible.

Theorem 1 (Theorem 5.1, restated)**.** *For any function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$ and $\varepsilon > 0$, there exists a constant δ and uniform (g,ℓ)-oNOSF source $\mathbf{X}$ with $g \leq \ell/2$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{1}{\lfloor \ell/g \rfloor} \cdot m + \delta$.*

[3] See Definition 4.9 for a definition of two-source extractors

This partially resolves[4] a conjecture of [AORSV20]: they conjectured that (g, ℓ)-oNOSF sources cannot be transformed into uniform (g', ℓ')-NOSF sources with $\frac{g'}{\ell'} > \frac{g}{\ell}$. Our condensing impossibility implies $\frac{g'}{\ell'} \leq \frac{1}{\lfloor \ell/g \rfloor}$ for any such transformation. This negative result is tight and we are able to condense uniform (g, ℓ)-oNOSF sources up to rate $\frac{1}{\lfloor \ell/g \rfloor}$.

Theorem 2. *(Informal version of Theorem 6.3) For any constant g and ℓ such that $\lfloor \ell/g \rfloor = r$ and $\ell/g \neq r$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ)-oNOSF source $\mathbf{X}$ we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq \frac{1}{r} \cdot m - O(\log(m))$ where $\varepsilon = 1/\Omega(m^{1/4})$ and $m = \Omega(n)$.*

As before in Theorem 2, we can convert a logarithmic min-entropy oNOSF source to a uniform oNOSF source and then apply Theorem 2. This yields:

Theorem 3. *(Informal version of Theorem 6.1) For all constant g and ℓ such that $\left\lfloor \frac{\ell-1}{g-1} \right\rfloor = r$ and $\frac{\ell-1}{g-1} \neq r$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any (g, ℓ, n, k)-oNOSF source $\mathbf{X}$ with $k \geq 2\log(n)$, we have that $H^\varepsilon_\infty(\mathbf{X}) \geq \frac{1}{r} \cdot m - O(\log(m))$ with $m = \Omega(k)$ and $\varepsilon = 1/\Omega(m^{1/4})$.*

We note that Theorem 1.4 and Theorem 1.5 are special cases of Theorem 2 and Theorem 3 in the case that $\lfloor \ell/g \rfloor = r = 1$, allowing us to state all of our condensing possibility results succinctly.

Put together, our results demonstrate a sharp threshold at $g = \ell/2$ for condensing from oNOSF sources with a small entropy gap. To our knowledge, there is no other set of sources that exhibits such behavior, making oNOSF sources unique among both adversarial sources and general randomness sources.

1.2.2 Adversarial Chor-Goldreich sources

Next, we consider a generalization of oNOSF sources, termed *adversarial Chor-Goldreich (CG) sources*, that we obtain by strengthening the adversary's power. Adversarial CG sources share the motivation from oNOSF sources that the adversary cannot predict the future. Rather than forcing the adversary to have its blocks only depend on blocks in the past (those with smaller indices), aCG sources require that good blocks have some entropy conditioned on all blocks that came before them. In other words, bad blocks cannot expose all of the entropy of future good blocks.

Definition 1.6 (Adversarial CG (aCG) sources, [CG88, DMOZ23])**.** *We define a (g, ℓ, n, k)-aCG source $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ to be a distribution on $(\{0,1\}^n)^\ell$ such that there exists a set of good indices $\mathcal{G} \subseteq [\ell]$ of size at least g for which $H_\infty(\mathbf{X}_i \mid \mathbf{X}_1 = x_1, \ldots, \mathbf{X}_{i-1} = x_{i-1}) \geq k$ for all $i \in \mathcal{G}$ and all prefixes $x_1, \ldots, x_{i-1}$.*

As before, if $k = n$, then we say that $\mathbf{X}$ is a uniform (g, ℓ)-aCG source. Observe that because the good blocks of a oNOSF source are independent of all blocks before it, oNOSF sources are trivially aCG sources as well. As a consequence, our condensing impossibility results from Theorem 1 immediately apply to aCG sources as well. Moreover, a convenient fact that we later show in Proposition 4.17 and will rely on is that uniform (g, ℓ)-aCG sources and uniform (g, ℓ)-oNOSF sources are equivalent.

[4] Our result on the existence of condensers falls short of completely resolving their conjecture as it does not transform uniform oNOSF sources into uniform NOSF sources.

CG sources are a well-studied class of sources introduced by [CG88] as a generalization of Santha-Vazirani sources [SV86]. Hence, the majority of the work done on CG sources has been in the non-adversarial setting in which $g = \ell$. Adversarial CG sources that contain bad blocks were only recently introduced in [DMOZ23] (although they use the terminology "almost" CG sources), in which the authors show several condensing results for CG and adversarial CG sources. Our work can then be seen as a meaningful addition to this long line of research on CG sources and their generalizations.

Previous work The impossibility of extraction from both oNOSF sources and aCG sources due to [AORSV20, CG88] naturally raises the question of whether there is a distinction between these two sources with regards to randomness condensing.

For CG sources, [GP20] showed that errorless condensing is impossible. In contrast, [DMOZ23] proved several possibility results regarding condensing with error for CG sources. Their results are stated as assuming the size of each block is very small (almost constant) compared to the number of blocks.

We also note that the authors of [DMOZ23] considered various other relaxations of the definition of aCG sources that we do not consider here. These include good blocks having only smooth min-entropy conditioned on previous blocks instead of the stronger condition of min-entropy, having smooth min-entropy conditioned on a constant fraction of prefixes of previous blocks instead of all prefixes, and having a Shannon entropy requirement instead of min-entropy requirement.

Our results In [DMOZ23], the authors pose the question of whether it is possible to condense from aCG sources with a constant entropy gap.[5] We give a partially positive answer to this by showing that we can condense from uniform (g, ℓ)-aCG sources with $g > \ell/2$ with logarithmic entropy gap since uniform (g, ℓ)-aCG sources are equivalent to uniform (g, ℓ)-oNOSF sources and we can defer to Theorem 2. Of course, all of Theorem 2 applies to uniform aCG sources, so we can condense any uniform (g, ℓ)-aCG source to rate $\frac{1}{\lfloor \ell/g \rfloor}$. The generalization of these results in Theorem 1.5 do not hold for non-uniform aCG sources since non-uniform aCG sources need not be oNOSF sources. Before our work, no non-trivial condensing was known for uniform (g, ℓ, n)-aCG sources even in the case of $g = \ell - 1$. It is important to note that our results hold for comparatively large block sizes $n = 2^{\omega(\ell)}$, in contrast to the results of [DMOZ23] that hold for constant block sizes and increasing ℓ.

As previously mentioned, since oNOSF sources are a subclass of aCG sources, our condensing impossibility results from Theorem 1 transfer over. Thus, in the $g \leq \ell/2$ regime, we give a negative answer to the question of [DMOZ23] by showing that good condensers do not exist for uniform (g, ℓ)-aCG sources, let alone condensers with a constant entropy gap. Note that unlike our condensing possibility results that only apply to uniform aCG sources, our impossibility result applies to non-uniform aCG sources as well.

In addition, we prove various condensing impossibility results that work even when there are no bad blocks (i.e., for non-adversarial, or just regular, CG sources): the first result of Theorem 4 is based on a reduction from general (n, k)-sources to CG sources and the second result uses a reduction from uniform oNOSF sources to low min-entropy CG sources.

[5] In their paper, they phrase it as removing the requirement of suffix-friendliness.

Theorem 4 (Informal version of Theorem 5.21 and Theorem 5.22)**.** *For all $\Delta > 0$ and for every function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists an (ℓ, ℓ)-aCG source $\mathbf{X}$ satisfying either of the following with $\varepsilon = 0.99$:*

- *The good blocks have min-entropy at least $n - \Delta - \log(\ell) - O(1)$ conditioned on all fixings of previous blocks and $H^\varepsilon_\infty(f(\mathbf{X})) \leq m - \Delta - \max(m - \ell n, 0) + O(1)$.*
- *The good blocks have min-entropy at least $n/\ell - \log(\ell) - O(1)$ conditioned on all fixings of previous blocks and $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{1}{2} \cdot m + O(1)$.*

It is important to note that the first bullet above does not subsume the second. In particular, the second bullet point from above gives a stronger result than the first in the setting when m is much larger than n.

We note that these results do not contradict the condensing result from [DMOZ23] as in the parameter regimes for which Theorem 4 works, the condenser of [DMOZ23] does not result in an entropy increase. This also shows a separation between aCG sources and oNOSF sources since Theorem 3 can condense from oNOSF sources in this parameter regime.

1.2.3 Non-oblivious symbol fixing sources

Finally, we strengthen the adversary one last time by letting the bad blocks depend arbitrarily on all the good blocks. This gives rise to NOSF sources which themselves generalize the setting of non-oblivious bit-fixing (NOBF) sources [CGHFRS85] where each block is a bit (i.e., $n = 1$).

Definition 1.7 (NOSF sources)**.** *A (g, ℓ, n, k)-NOSF source $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ on $(\{0,1\}^n)^\ell$ is such that g out of the ℓ blocks are independently sampled (n, k)-sources (i.e., "good") while the other $\ell - g$ bad blocks may depend arbitrarily on the good blocks.*

When $k = n$ and n is clear from context, we simply call $\mathbf{X}$ a uniform (g, ℓ)-NOSF source. The adversary in NOSF sources clearly has a significant amount of power; every single good block is sampled before the adversary gets to decide what to place in the bad blocks. As NOSF sources are in the setting in which the adversary is the strongest, they are also the sources for which we are most motivated to be able to extract or condense as they are the most general. We note that much of the progress on explicit constructions of two-source extractors and condensers [CZ19, BCDT19], a major problem in the area of randomness extraction, is based on constructing extractors and condensers for NOSF sources (in a parameter regime where it was existentially known that extraction is possible). This further motivates our exploration of condensing from NOSF sources in a more general parameter setting.

Previous work We can trace back study of extracting from NOBF sources to the seminal work of Ben-Or and Linial in [BL89].[6] They made the connection between NOBF extractors and the influence of sets of variables on Boolean functions. Together with the work of Kahn, Kalai, and Linial in [KKL88], in which they demonstrated lower bounds on the influence of variables on Boolean functions, these works show that it is not possible to extract from uniform (g, ℓ)-NOBF sources when the number of bad bits is $b = \ell - g = \Omega(\ell / \log \ell)$. While no analogous result is known for NOSF sources,[7] the extraction impossibility result from [AORSV20] for oNOSF sources also

[6] They used the terminology "collective coin flipping protocol" instead of "NOBF extractor".

[7] Although one is conjectured in [Fri04] that attempts to recover what was initially proposed in [BKKKL92].

applies for NOSF sources: for any $\gamma > 0$ there exists a large constant ℓ such that it is impossible to extract even one bit from uniform $(\gamma\ell, \ell)$-NOSF sources.

To attempt to match these lower bounds on extraction, resilient functions, introduced by [BL85], have yielded the current best results. The resilient function of Ajtai and Linial in [AL93] and its explicit versions constructed by [CZ19, Mek17] achieve extractors for uniform (g, ℓ)-NOBF sources when $b = O(\ell/\log^2 \ell)$, leaving a $1/\log \ell$ gap between the lower and upper bounds.

Noting that a uniform (g, ℓ, n)-NOSF source is a uniform $(ng, n\ell)$-NOBF source, these results imply extractors when $g > \ell(1 - 1/C\log^2(n\ell))$, for some large enough constant C. This still leaves open whether condensing is possible for most settings of parameters.

Related to this, the work of [KN23] explores what they call extracting multimergers, which we may consider as extractors for uniform NOSF sources. For seedless extracting multimergers, their result implies that extracting from uniform $(2, 3)$-NOSF sources is impossible.

Our results As oNOSF sources are also NOSF sources, our condensing impossibility result in Theorem 1 also applies to (g, ℓ)-NOSF sources when $g \leq \ell/2$. However, we are able to show an even stronger result for any setting of g and ℓ and thus extend existing lower bounds of extraction to condensing.

Theorem 5 (Corollary 5.10 restated)**.** *For all constant $g, \ell \in \mathbb{N}$, there exist constant $\varepsilon, \delta > 0$ so the following holds: for all $a, m, n \in \mathbb{N}$ and all functions $f : (\{0,1\}^n)^{a\ell} \to \{0,1\}^m$, there exists a uniform $(ag, a\ell)$-NOSF source $\mathbf{X}$ such that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + \delta$.*

By varying a above, we extend our result for any g and ℓ to any rate g/ℓ uniform NOSF source. These results together put NOSF sources in stark contrast with adversarial CG and oNOSF sources since they can both be condensed in a useful manner for simulating BPP algorithms, while we have shown that NOSF sources cannot be condensed in such a manner.

2 Proof Overview

We present the main ideas and techniques for proving our main condensing impossibility results in Section 2.1 and possibility results in Section 2.2.

2.1 Impossibility results

In this subsection, we will go over the main techniques used in proving the condensing impossibility result for the case that $g \leq \ell/2$ in Section 2.1.1, the condensing impossibility result for uniform NOSF sources when $g > \ell/2$ in Section 2.1.2, and the condensing impossibility result for low min-entropy CG sources in Section 2.1.3.

2.1.1 Impossibility of condensing from uniform (g, ℓ)-oNOSF sources for $g \leq \ell/2$

We prove that when the number of good blocks g is not more than half of the total number of blocks ℓ, then condensing beyond rate $\frac{1}{\lfloor \ell/g \rfloor}$ is impossible. Formally, we will prove the following statement.

Theorem 2.1 (Theorem 5.1, restated)**.** *For all ε, there exists a δ such that for all $g, \ell \in \mathbb{N}$ with $g \leq \ell/2$ and for all $f : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$, there exists a uniform (g, ℓ)-oNOSF source $\mathbf{X}$ such that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{1}{\lfloor \ell/g \rfloor} \cdot m + \delta$.*

The steps we take to achieve the result of Theorem 2.1 are, broadly, as follows:

1. We first reduce proving the theorem to only proving it for the special case of $g = 1$. We show that if it is possible to condense uniform (g, ℓ)-oNOSF sources to entropy-rate more than $\frac{1}{\lfloor \ell/g \rfloor}$, then it is possible to condense uniform $(1, \ell')$-oNOSF sources to rate beyond $\frac{1}{\lfloor \ell'/1 \rfloor} = \frac{1}{\ell'}$ where $\ell' = \lfloor \ell/g \rfloor$. We do this by transforming any uniform $(1, \ell')$-oNOSF source to a uniform (g, ℓ)-oNOSF source.

2. We prove the theorem for the special case of $g = 1$ and arbitrary ℓ. We do this by using an "induct or win" argument. We show either condensing from uniform $(1, \ell)$-oNOSF sources is impossible (win) or we reduce to the case of condensing from uniform $(1, \ell - 1)$-oNOSF sources (induct). Either we will win at some point in our reduction or we will reach the base case of $g = \ell = 1$ where the claim trivially holds. Let f be a candidate condenser and take cases on whether there exists a fixing of the first block in f such that the partial function obtained by fixing f to that values will have small support. If such a fixing exists, then we reduce the problem to condensing from uniform $(1, \ell - 1)$-oNOSF sources. If not, then we directly construct a uniform $(1, \ell)$-oNOSF source where f fails to condense from by reducing to a graph problem.

3. The graph problem we reduce to in the "win" case is the following: Let $G = (U, V)$ be a bipartite graph with $U = [N], V = [M]$ and such that $\deg(u) \geq c_0 M^{\delta}$ for all $u \in U$ where $\delta > 0$ is some constant. Then, show there exists $D \subset V$ such that $|\mathrm{Nbr}(D)| \geq c_1 N$ and $|D| \leq c_2 \cdot M^{1-\delta}$ where c_0, c_1, c_2 are some universal constants.

We expand on these three steps and prove them.

Step 1 In this step, we transform any uniform $(1, \ell')$-oNOSF source $\mathbf{X}$ to a uniform (g, ℓ)-oNOSF source $\mathbf{Y}$ where $\ell' = \lfloor \ell/g \rfloor$. Divide ℓ by ℓ' so that $\ell = a\ell' + r$ where $0 \leq r < \ell'$. We compute that $a \geq g$. We split the blocks of $\mathbf{X}$ as evenly as possible: split up the first r blocks of $\mathbf{X}$ into $a + 1$ blocks and the remaining $\ell' - r$ blocks into a blocks. These $a\ell' + r = \ell$ blocks that we obtained by splitting $\mathbf{X}$ will form $\mathbf{Y}$. If a block in $\mathbf{X}$ is uniform, then all the split up blocks will also be uniform. Similarly, if a block in $\mathbf{X}$ is bad and only depended on blocks appearing before it, so will all the blocks formed after splitting it. Also, as at least one block in $\mathbf{X}$ is good, $\mathbf{Y}$ must have at least $a \geq g$ good blocks in it. Hence, $\mathbf{Y}$ is indeed a uniform (g, ℓ)-oNOSF source.

Step 2 In this step, we execute our induct or win argument. Fix a candidate condenser function $f : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$. We proceed by contradiction and assume f can condense from uniform $(1, \ell)$-oNOSF sources beyond rate $1/\ell$. We either directly construct a uniform $(1, \ell)$-oNOSF source $\mathbf{X}$ where f will fail to condense from or we show how using f, we can obtain a condenser for uniform $(1, \ell - 1)$-oNOSF sources, which is a contradiction.

Case 1. There exists a fixing of the first block x_1 such that $\left|f(x_1, y_1, \ldots, y_{\ell-1}) | (y_1, \ldots, y_{\ell-1}) \in \{0,1\}^{n(\ell-1)}\right| \leq 2^{m(1-1/\ell)}$. Then, by appropriately relabeling outputs, we can define $h : (\{0,1\}^n)^{\ell-1} \to \{0,1\}^{m(1-1/\ell)}$ as $h(y_1, \ldots, y_{\ell-1}) = f(x_1, y_1, \ldots, y_{\ell-1})$. We now show that h will be a condenser for uniform $(1, \ell - 1)$-oNOSF sources. Let $\mathbf{Y}$ be arbitrary uniform $(1, \ell - 1)$-oNOSF source. We transform $\mathbf{Y}$ into a uniform $(1, \ell)$-oNOSF source$\mathbf{Y'}$ by letting the first block of $\mathbf{Y'}$ be fixed to x_1 and the remaining $\ell - 1$ blocks behave as $\mathbf{Y}$. By

assumption, f can condense $\mathbf{Y}'$ so that output entropy is more than $\frac{1}{\ell} \cdot m$. However this implies h can condense $\mathbf{Y}$ to have entropy more than $\frac{1}{\ell} \cdot m = \frac{1}{\ell-1} \cdot m(1-1/\ell)$. As h outputs $m(1-1/\ell)$ bits, this is a contradiction.

Case 2. For every fixing of the first block $x_1 : |f(x_1, y_1, \ldots, y_{\ell-1})| > 2^{m(1-1/\ell)}$. To show f fails to condense from $\mathbf{X}$, it suffices to show that with constant probability, $f(\mathbf{X})$ will lie in a small set $D \subset \{0,1\}^m$ where $|D| = O\left(2^{m(1/\ell)}\right)$ (see Claim 4.3 for a formal version of this). Consider the bipartite graph $H = (U = (\{0,1\}^n), V = \{0,1\}^m)$ where edge (u, v) is included if there exist $y_1, \ldots, y_{\ell-1}$ such that $f(x_1, y_1, \ldots, y_{\ell-1}) = v$. By assumption, for all $u \in U : \deg(u) > 2^{m(1-1/\ell)}$. Our graph theoretic dominating set lemma from Item 3 guarantees that there exists $D \subset \{0,1\}^m$ such that $|D| \leq c_0 2^{m(1/\ell)}$ and $|\mathrm{Nbr}(D)| \geq c_1 2^n$ where c_0, c_1 are universal constants. Now, let $\mathbf{X}$ be uniform $(1, \ell)$-oNOSF source where the first block is uniform and the remaining $\ell - 1$ blocks are adversarial where the value of those $\ell - 1$ blocks (depending on the value of the first block) is set so that f outputs an element from D if possible. By the construction of the bipartite graph and the construction of $\mathbf{X}$, with probability c_1, $f(\mathbf{X})$ will output an element in D. Hence, as f outputs an element from a small set, D, with high probability, it fails to condense from $\mathbf{X}$.

Step 3 We prove the dominating set lemma for bipartite graph in this step to conclude the proof of the "win" argument. We construct D by repeatedly adding the vertex from V that has the highest degree, removing vertices incident to that vertex, and stopping until at least $c_1 N$ many vertices from U are incident to some vertex from D. Whenever we attempt to add a vertex to D, the graph will have at least $(1-c_1)N$ many vertices and so at least $(1-c_1)N \cdot c_0 \cdot M^{1-\delta}$ many edges. This implies there will always be a vertex $v \in V$ such that $\deg(v) \geq c_0(1-c_1) \cdot \frac{N}{M^\delta}$. This is true at each stage and we repeat this until at least $c_1 N$ many vertices are covered. Hence, $|D| \leq c_2 \cdot M^{1-\delta}$ for some universal constant c_2 as desired.

2.1.2 Impossibility of condensing from uniform NOSF sources

We prove much stronger condensing impossibility result for uniform NOSF sources: we prove that no non-trivial condensing is possible. We are able to do so since the bad blocks have no restrictions and can arbitrarily depend on any good block. Formally, we show the following:

Theorem 2.2 (Corollary 5.10 restated)**.** *For all fixed $g, \ell \in \mathbb{N}$, there exist fixed $\varepsilon, \delta > 0$ so that the following holds: for all $a, m, n \in \mathbb{N}$ and all functions $f : (\{0,1\}^n)^{a\ell} \to \{0,1\}^m$, there exists a uniform $(ag, a\ell)$-NOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + \delta$.*

We prove Theorem 2.2 using the following strategy:

1. We reduce the general case to the special case of $a = 1$ and $g > \ell/2$.
2. Our high level strategy for this step is same as in Item 2 from Section 2.1.1. We perform an "induct or win" argument to show it is impossible to condense from uniform (g, ℓ)-NOSF sources where $g > \ell/2$ beyond rate g/ℓ. As earlier, we show either condensing from uniform (g, ℓ)-NOSF sources is impossible (win) or we reduce to the case of condensing from uniform $(g, \ell-1)$-NOSF sources (induct). So, we recursively apply this argument and either win at some point or reach a base case of $g = \ell$ where the claim trivially holds. Let f be a candidate condenser and take cases on whether there exists a block position p such that for constant fraction of fixings of all other blocks, the partial function obtained by fixing f to those values

will have large support. If this holds, then we use the almost-dominating set argument from Item 3 (from Section 2.1.1) to reduce to the case of condensing from uniform $(g, \ell-1)$-NOSF sources. If such a position p with these fixings do not exist, then we directly construct a uniform (g,ℓ)-oNOSF source where f fails to condense from by reducing to a hypergraph problem.

3. The hypergraph problem we reduce to in the "win" case is the following: Let $H = (V_1, \ldots, V_t, E)$ be a t-uniform t-partite hypergraph with $V_1 = \cdots = V_t = [N], |E| = c_0 N^t$. Let the edges of H be colored in M colors in a 'locally light' way: such that for every position $p \in [T]$, and every $(t-1)$ tuples: $(v_1, \ldots, v_{p_1}, v_{p+1}, \ldots, v_t) \in [N]^{t-1}$, the number of distinct colored edges as entries in position p vary is $\leq c_1 M^{\delta}$. Formally, $|\chi(v_1, \ldots, v_{p-1}, y, v_{p+1}, \ldots, v_t) : y \in [N]| \leq c_1 M^{\delta}$. Then, there exists $D \subseteq [M]$ such that $|D| \leq c_2 \cdot M^{t\delta}$ and at least $c_3 N^t$ edges in H are colored in one of the colors from D. Here, c_0, c_1, c_2 are some constants.

We expand on these three steps and prove them.

Step 1 We show how to reduce to the case of $a = 1$. We do this using the same argument as in Item 1: we transform uniform (g,ℓ)-NOSF sources into uniform $(ag, a\ell)$-NOSF sources by splitting up blocks; this way, a condenser for uniform $(ag, a\ell)$-NOSF sources will also condense from uniform (g,ℓ)-NOSF sources.

We next carefully examine the argument made in Item 2 and see that the induct or win argument made there can be generalized to show the following: either condensing from uniform (g,ℓ)-NOSF source is impossible or we reduce to the case of condensing from uniform $(g, \ell-g)$-NOSF source. Applying this recursively to arbitrary g, ℓ, we either win and show impossibility at some step or we end up reducing to showing impossibility for condensing from uniform (g,ℓ)-NOSF sources where $g > \ell/2$.

Combining these two steps, we reduce the general case to the special case of $a = 1$ and $g > \ell/2$.

Step 2 In this step, we execute our induct or win argument. We fix a candidate condenser function $f : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$. Proceed by contradiction and assume f can condense from uniform (g,ℓ)-NOSF sources beyond rate g/ℓ. We either directly construct a uniform (g,ℓ)-NOSF source $\mathbf{X}$ where f will fail to condense from or we show how using f, we can obtain a condenser for uniform $(g, \ell-1)$-NOSF sources, which is a contradiction. For $p \in [\ell]$, let S_p be the set of $\ell-1$ tuples $(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell)$ such that

$$|\{f(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) : y \in \{0,1\}^n\}| \geq c_0 2^{m/\ell}$$

Case 1. There exists $p \in [\ell]$ such that $|S_p| \geq c_1 2^{n(\ell-1)}$ where $c_1 > 0$ is a small constant. Without loss of generality let $p = 1$. Construct a bipartite graph $G = (U, V)$ where $U = S_1, V = \{0,1\}^m$ and edge (u, v) if there exists a fixing y of block p such that $f(u, y) = v$. Then, we see that G satisfies the requirement for Item 3 and hence, there exists $D \subset \{0,1\}^m$ such that $|D| \leq 2^{m(1-1/\ell)}$ which neighbors at least $c_3 2^{n(\ell-1)}$ vertices from U. For the sake of presentation, assume $c_1 = c_3 = 1$. In the full proof, $c_1, c_3 > 0$ are small constants and we need to induct using a stronger inductive hypothesis. Now, define $h : \{0,1\}^{n(\ell-1)} \to \{0,1\}^{m(1-1/\ell)}$ as $h(y_1, \ldots, y_{\ell-1}) = f(x_1, y_1, \ldots, y_{\ell-1})$ where x_1 is such that $f(x_1, y_1, \ldots, y_{\ell-1}) \in D$ (as $c_1 = c_3 = 1$, such x_1 always exists). The output

domain of h can be made $\{0,1\}^{m(1-1/\ell)}$ instead of D by appropriately relabeling the output. We then show, similar to proof of case 1 of Item 2, h will be a condenser for uniform $(g, \ell-1)$-NOSF sources and get a contradiction.

Case 2. For all $p \in [\ell]$, $|S_p| \geq c_1 2^{n(\ell-1)}$. We say $x = (x_1, \ldots, x_\ell) \in \{0,1\}^{n\ell}$ is bad if for some $p \in [\ell]$, removing position p from x makes it an element of S_p. Let B be set of such bad strings. Then, $|B| \leq c_1 \ell \cdot 2^{n\ell}$. Let $H = (V_1, \ldots, V_\ell)$ where $V_i = \{0,1\}^n$ be ℓ-uniform ℓ-partite hypergraph where edge $v = (v_1, \ldots, v_\ell)$ is in H if $v \notin B$. Then, H has at least $(1 - c_1\ell)2^{n\ell}$ edges. By an averaging argument, there exists $x = (x_1, \ldots, x_{\ell-g}) \in \{0,1\}^{n(\ell-g)}$ such that the number of edges in H containing that x is at least $(1 - c_1\ell)2^{ng}$. Consider uniform (g,ℓ)-oNOSF source $\mathbf{Y}$ where the first $\ell - g$ blocks always output x and the remaining g blocks are uniform. To show f fails to condense from $\mathbf{X}$, it suffices to show: constant probability, $f(\mathbf{X})$ will lie in a small set $D \subset \{0,1\}^m$ where $|D| = O\left(2^{m(g/\ell)}\right)$ (see Claim 4.3 for a formal version of this).

Let $H' = (U_1, \ldots, U_g)$ where $U_i = \{0,1\}^n$ be g-uniform g-partite hypergraph where edge $u = (u_1, \ldots, u_g)$ is in H' if $(x_1, \ldots, x_{\ell-g}, u_1, \ldots, u_g)$ is in H. Then, H' has at least $(1 - c_1\ell)2^{ng}$ edges. Now, color H' into 2^m colors by coloring edge $(u_1, \ldots, u_g)$ as $f(x_1, \ldots, x_{\ell-g}, u_1, \ldots, u_g)$. By definition of S_p and construction of H', we see that for every $\ell - 1$ tuples u in $\{0,1\}^{n(\ell-1)}$, the number of distinct colors in H' is at most $c_0 \cdot 2^{m/\ell}$. We apply the hypergraph lemma to H' and infer that there exists $D \subset \{0,1\}^m$ such that $|D| \leq c_2 \cdot 2^{m(g/\ell)}$ at least $c_3 \cdot 2^{ng}$ edges in in H' are colored in one of the colors from D. Hence, we found a small set D such that with constant probability, $f(\mathbf{X})$ lies in D as desired.

Step 3 We finally solve the hypergraph problem to conclude the proof of the "win" argument. We repeatedly pick the color which covers the most edges to D until the number of edges covered is at least $c_3 \cdot N^t$. At the last step of the process, H must have at least $(c_0 - c_3) \cdot N^t$ edges. We show that at that stage, the chosen a color will cover at least $c_4 N^t / M^{t\delta}$ edges. This implies at each step before this, the chosen color must cover at least that many edges and hence, $|D| \leq \frac{1}{c_4} M^{t\delta}$ as desired.

So, our goal is to show that in a t-uniform t-partite hypergraph $H = (V_1, \ldots, V_t)$ having at least $c_5 N^t$ edges and colored in M colors in a 'locally light' manner - on fixing any $t-1$ tuple, the number of colors adjacent to it as last entry varies is at most $c_1 \cdot M^\delta$, there exists a color γ covering at least $\Omega(N^t/M^{t\delta})$ edges. We induct on t and show this. We sketch the idea below for bipartite graphs.

For every $v_2 \in V_2$, let $C_{v_2} \subset [M]$ be the set of colors that have at most $c_6 \cdot (N/M^\delta)$ where c_6 is a very small constant. We remove edge (v_1, v_2) from H if $(v_1, v_2) \in C_{v_2}$. For each v_2, we remove at most $c_1 c_6 \cdot N$ edges incident to it. Overall, we end up removing at most $c_1 c_6 \cdot N^2$ edges from H and it still has $(c_5 - c_1 c_6)N^2$ edges. Doing this ensures that every color incident to every vertex v_2 in V_2 has at least $c_6 \cdot (N/M^\delta)$ edges incident to it. We finally find such a popular color by doing the following: By averaging argument, let $v_1^* \in V_1$ and $\gamma^* \in [M]$ be such that the number of edges incident to v_1^* with color γ is at least $\frac{c_5 - c_1 c_6}{c_1} \cdot (N/M^\delta)$. Let $\mathrm{Nbr}_\gamma(v_1^*) = \{v_2 \in V_2 : (v_1^*, v_2) \text{ is colored with color } \gamma\}$. Moreover, for every $v_2 \in \mathrm{Nbr}_\gamma(v_1^*)$, the number of edges incident to them with color γ is at least $c_6 \cdot N/M^\delta$. We are done as at least $c_6 \cdot \frac{c_5 - c_1 c_6}{c_1} \cdot N^2/M^{2\delta} = \Omega(N^2/M^{2\delta})$ edges in H colored with color γ.

2.1.3 Impossibility of condensing from low min-entropy aCG sources

We provide two impossibility result for (ℓ, ℓ)-aCG source, we only sketch proof for one of them as they both share many ideas. Our impossibility result Theorem 5.21 is based on reduction from general (n, k)-sources and the fact that it is impossible to condense from such sources.

Here, we sketch a proof of the second impossibility result where we show that it is impossible to condense from non-adversarial CG sources when each block's min-entropy, conditioned on previous blocks, is roughly bounded by $n/(\ell + 1)$.

Theorem 2.3 (Theorem 5.22 restated)**.** *For all $0 < \varepsilon < 1$ there exists a $\delta > 0$ such that the following holds: for every function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a (ℓ, ℓ)-aCG source $\mathbf{X}$ where the good blocks have min-entropy at least $\frac{n-\ell\log(2\ell/\varepsilon)}{\ell+1}$ conditioned on all fixings of previous blocks and $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{1}{2} \cdot m + \delta$.*

The bulk of the proof is based on a transformation from a uniform $(1, 2)$-oNOSF source $\mathbf{X} = \mathbf{X}_1, \mathbf{X}_2$ to a source $\mathbf{Y} = \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ that is $\varepsilon/2$-close to an (ℓ, ℓ)-aCG source. With this transformation, applying Theorem 2.1 with $\ell = 2$ and $\varepsilon/2$ to $\mathbf{X}$ then allows us to infer that we also cannot condense from $\mathbf{Y}$ with error ε. Thus, we focus on how to construct $\mathbf{Y}$ next.

Briefly, to construct $\mathbf{Y}$, we will take substrings of $\mathbf{X}_1$ and $\mathbf{X}_2$ to place into each block of $\mathbf{Y}$. From $\mathbf{X}_2$, we will take constant sized chunks of size $t_2 = \frac{n-\ell\log(2\ell/\gamma)}{\ell+1}$ where $\gamma = \frac{\varepsilon}{2\ell}$ to place into each $\mathbf{Y}_i$, and from $\mathbf{X}_1$ we will take blocks of increasing size $i \cdot t_1 - 1$ to place into each $\mathbf{Y}_i$ where $t_1 = t_2 + \log(1/\gamma)$. Our proof then finishes with an inductive argument to claim that $\mathbf{Y}$ is indeed $\varepsilon/2$-close to an (ℓ, ℓ)-aCG source source, as required.

2.2 Possibility results

In this subsection, we will present our existential construction of condensers for oNOSF sources and uniform aCG sources. We begin by describing the construction of our condenser for uniform (g, ℓ)-oNOSF sources and uniform (g, ℓ)-aCG source in the setting of $g > \ell/2$ in Section 2.2.1. Then we generalize this result to any setting of g and ℓ in Section 2.2.2. Finally, we deal with logarithmic min-entropy oNOSF sources in Section 2.2.3.

2.2.1 Condensing from uniform (g, ℓ)-oNOSF sources for $g > \ell/2$

Before we dive into the actual proof, it is instructive to see why a random function fails to be a condenser for uniform (g, ℓ)-oNOSF sources. In particular, let us consider uniform $(2, 3)$-oNOSF sources. For a random function $f : \{0,1\}^{3n} \to \{0,1\}^t$, with high probability over $x_1, x_2 \in \{0,1\}^n$, we have $|f(x_1, x_2, \cdot)| = 2^m$. Hence, if the adversary is in position 3, then it can depend on x_1 and x_2 to ensure the output of f always lies in a small set. To overcome this, one can consider restricting the number of choices adversary has when it is in position 3. This intuition indeed works out and we give further details.

Theorem 2.4 (Theorem 6.15 restated)**.** *For all g, ℓ such that $g > \ell/2$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ)-oNOSF source $\mathbf{X}$, $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - (5^{\ell-g} - 3)\log(gn)$ where $m = n - 2(5^{\ell-g} - 1)\log(gn), \varepsilon = (gn)^{-1/4}$.*

Our construction relies on a (k_1, k_2, ε)-two-source extractor $\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ with a property that we term *output-lightness*, the definition and importance of which we will see

soon, and a clever choice of a partition and prefixes of our input source $\mathbf{X}$. We do not currently know of a construction of a two-source extractor with our desired min-entropy and error parameters that is also output-light, so our construction is currently based on an existential output-light two-source extractor that we show in Lemma 6.13. In particular, if we write $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ and we take $\mathbf{Y}_i$ to be the prefix of $\mathbf{X}_i$ containing the first $5^{\ell-i} \cdot 4\log(gn)$ bits, then we define our two inputs to Ext as $\mathbf{Z}_1 = \mathbf{X}_1, \ldots, \mathbf{X}_g$ and $\mathbf{Z}_2 = \mathbf{Y}_{g+1}, \ldots, \mathbf{Y}_\ell$. Thus, our condenser becomes $\mathsf{Cond}(\mathbf{X}) := \mathsf{Ext}(\mathbf{Z}_1, \mathbf{Z}_2)$.

There are only two cases we must consider: when the adversary places at least one good block in $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ and when all of $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ are adversarial (so $\mathbf{X}_1, \ldots, \mathbf{X}_g$ is uniform). In the latter case, we have that $\mathbf{Z}_1$ is just the uniform distribution on gn bits and $\mathbf{Z}_2$ is fully controlled by the adversary. For $\mathsf{Ext}(\mathbf{Z}_1, \mathbf{Z}_2)$ to condense then, we would require that no element $h \in \{0,1\}^m$ have too much weight placed on it by the adversary. Recalling that $\mathbf{Z}_1$ is uniform in this case, this statement is equivalent to asking that the sum over all settings z_1 of $\mathbf{Z}_1$ of the number of z_2 such that $\mathsf{Ext}(z_1, z_2) = h$ is not larger than $R = 2^{n_1+n_2-m+O(1)}$. This is precisely our definition of R-output-lightness (see Definition 6.12 for a formal definition). With this property, we use Claim 4.5 to get that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq n_1 = \log(R/\varepsilon)$.

In the case that there is at least one good block among $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$, then we notice that there must be one good block among $\mathbf{X}_1, \ldots, \mathbf{X}_g$ because $g > \ell/2$, so $H_\infty(\mathbf{Z}_1) \geq n$. Without loss of generality, we also assume that we only have one good block $\mathbf{X}_j$ for $j \in \{g+1, \ldots, \ell\}$. Consequently, we can define $\mathbf{A} = \mathbf{Y}_{g+1}, \ldots, \mathbf{Y}_{j-1}$ and $\mathbf{B} = \mathbf{Y}_{j+1}, \ldots \mathbf{Y}_\ell$ so that $\mathbf{Z}_2 = \mathbf{A} \circ \mathbf{Y}_j \circ \mathbf{B}$ where the adversary controls both $\mathbf{A}$ and $\mathbf{B}$ but not $\mathbf{Y}_j$. Since $\mathbf{X}$ is a oNOSF source, $\mathbf{Y}_j$ remains uniform regardless of any fixing of $\mathbf{A}$, so $H_\infty(\mathbf{Y}_j \mid A) = H_\infty(\mathbf{Y}_j) = 5^{\ell-j} \cdot 4\log(gn)$. In addition, since we chose $\mathbf{A}$ to be logarithmically small in n, the min-entropy chain rule (Lemma 4.4) gives us that, with high probability over the fixings of $\mathbf{A}$, the min-entropy of $\mathbf{Z}_1$ is not decreased by too much more than the length of $\mathbf{A}$ which is at most n_2. In particular, for any of these good fixings $a \in \mathsf{Supp}(\mathbf{A})$, we chose k_1 to be such that $H_\infty(\mathbf{Z}_1 \mid \mathbf{A} = a) \geq k_1$. Then if we temporarily make the assumption that $\mathbf{B}$ is uniform, we have that $H_\infty(\mathbf{Z}_2 \mid \mathbf{A} = a) = H_\infty(a, \mathbf{Y}_j, \mathbf{B} \mid \mathbf{A} = a) = \sum_{i=j}^{\ell} 5^{\ell-j} \cdot 4\log(gn) = (5\ell - j + 1 - 1)\log(gn)$. Since we can choose k_2 to be smaller than $H_\infty(\mathbf{Z}_2 \mid \mathbf{A} = a)$, we get that $\mathsf{Ext}(\mathbf{Z})$ is ε-close to $\mathbf{U}_m$. Of course, $\mathbf{B}$ may be adversarially chosen. To take this into account, we use Lemma 6.18, which says that if only a few bits of a source are adversarially controlled then we can still condense, to reduce our output entropy by the length of $\mathbf{B}$ and multiplicatively increase our error by $2^{\text{length}(\mathbf{B})}$. Finally, because we constructed $\mathbf{B}$ to have $\sum_{i=j+1}^{\ell} 5^{\ell-j} \cdot 4\log(gn) = (5^{\ell-j} - 1)\log(gn)$ bits, it is still short enough in comparison $\mathbf{Y}_j$ to allow us to condense with our desired error.

2.2.2 Condensing from uniform (g, ℓ)-oNOSF sources for any g and ℓ

While we can condense from uniform (g, ℓ)-oNOSF sources for $g > \ell/2$ as we saw above (Theorem 2.4), we know from Theorem 2.1 that when $g \leq \ell/2$ we cannot condense from uniform (g, ℓ)-oNOSF sources above rate $\frac{1}{\lfloor \ell/g \rfloor}$. Here, we sketch the argument for a matching bound showing that this is indeed tight by generalizing Theorem 2.4.

Theorem 2.5 (Theorem 6.3 restated)**.** *For any g and ℓ such that $\lfloor \ell/g \rfloor = r$ and $r < \ell/g$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ)-oNOSF source $\mathbf{X}$ we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq \frac{1}{r} \cdot m - 2(5^{\ell-g} - 1)\log(gn)$ where $\varepsilon = (gn)^{-1/4}$ and $m = r(n - 2(5^{\ell-g} - 1)\log(gn))$.*

Satisfyingly, we need no new tools to construct this condenser. Instead, we use r instances of the condenser from Theorem 2.4. We will prove this inductively on r, so let us consider the base

case of $r = 1$. Notice that $r = 1$ implies that $g > \ell/2$, so we are exactly in a position to use the condenser $\mathsf{Cond}_1 : (\{0,1\}^n)^\ell \to \{0,1\}^{m_1}$ from Theorem 2.4 without modification. Thus, we define our output block as $\mathbf{O} = \mathbf{O}_1 = \mathsf{Cond}_1(\mathbf{X})$.

To generalize to larger values of r, we perform induction on r and take the inductive hypothesis of $r-1$ to be true. We consider two cases. Beginning with the case that all of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are bad, we notice that $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ is a uniform $(g, \ell-g)$-oNOSF source with $\left\lfloor \frac{\ell-g}{g} \right\rfloor = r-1$ and $\frac{\ell-g}{g} \neq r-1$. Our inductive hypothesis then gives us $r-1$ output blocks $\mathbf{O}_2, \ldots, \mathbf{O}_r$ on $(\{0,1\}^{m_r})^{r-1}$ where at least one is condensed. On the other hand, consider when at least one of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ is good and take Cond_1 to be an instance of the condenser from Theorem 2.4 for $\mathbf{X}$, and define $\mathbf{O}_1$ to be $\mathsf{Cond}_1(\mathbf{X})$ truncated to its first m_r bits. Observe that if $\mathsf{Cond}_1(\mathbf{X})$ succeeds and condenses $\mathbf{X}$ to some min-entropy k source, then $H_\infty(\mathbf{O}_1) \geq k - (m_1 - m_r)$, so we only lose as many bits of entropy in $\mathbf{O}_1$ as we truncate from $\mathsf{Cond}_1(\mathbf{X})$, which we show in Lemma 6.19, and $m_1 - m_r$ is still constant in g and ℓ. Then in this case we again get that $\mathbf{O}_1$ must be properly condensed by $2\mathsf{Ext}_1$ being output-light when all of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are good or by $2\mathsf{Ext}_1$ being a two-source extractor when at least one of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ is good. Thus, if we let our output be $\mathbf{O} = \mathbf{O}_1, \ldots, \mathbf{O}_r$, then at least one block is always condensed in any case.

2.2.3 Condensing from logarithmic min-entropy (g, ℓ)-oNOSF sources

We can extend Theorem 2.4 and Theorem 2.5 to logarithmic min-entropy oNOSF source by converting a logarithmic min-entropy oNOSF source into a uniform oNOSF source via the following theorem.

Theorem 2.6 (Theorem 6.2 restated)**.** *For any g and ℓ, there exists a function $f : (\{0,1\}^n)^\ell \to (\{0,1\}^m)^{\ell-1}$ with $m = \frac{k}{8\ell}$ such that for any (g, ℓ, k)-oNOSF source $\mathbf{X}$ with $k \geq 2\log(n)$ there exists a uniform $(g-1, \ell-1)$-oNOSF source $\mathbf{Y}$ such that $|f(\mathbf{X}) - \mathbf{Y}| \leq (g-1) \cdot 2^{-\Omega(k)}$.*

Thus, if we take a (g, ℓ, n, k)-oNOSF source $\mathbf{X}$ such that $g > \ell/2+1$ so $g-1 > (\ell-1)/2$, we can simply apply f from Theorem 2.6 to $\mathbf{X}$ and then pass the result to our condenser from Theorem 2.5 to condense from logarithmic min-entropy oNOSF source.

Theorem 2.7 (Theorem 6.1 restated)**.** *For all $g, \ell, r \in \mathbb{N}$ such that $\left\lfloor \frac{\ell-1}{g-1} \right\rfloor = r$ and $r < \frac{\ell-1}{g-1}$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any (g, ℓ, n, k)-oNOSF source $\mathbf{X}$ with $k \geq 2\log(n)$, we have that $H^\varepsilon_\infty(\mathbf{X}) \geq \frac{1}{r} \cdot m - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)$ with $m = r\left(\frac{k}{8\ell} - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)\right)$ and $\varepsilon = (g-1) \cdot 2^{-\Omega(k)} + \left(\frac{(g-1)k}{8\ell}\right)^{-1/4}$.*

All that is left then is to show how we convert a low min-entropy oNOSF source to a uniform oNOSF source in Theorem 2.6. Our method here is based on the somewhere extractor for low-entropy oNOSF source from [AORSV20] with two important modifications. First, we use a two-source extractor instead of a seeded extractor which enables us to handle logarithmic min-entropy in the good blocks of a oNOSF source instead of just linear. Second, we require that the output of our function is not just somewhere random, but instead a uniform oNOSF source. To achieve this, we decrease the output length of our two-source extractor (which decreases the block length of our resulting uniform oNOSF source) to show that the good blocks in our resulting source are independent from all adversarial blocks before them.

The construction of f from Theorem 2.6 is quite straightforward. For every $i \in \{2, \dots, \ell\}$, we use the same existential two-source extractor from Lemma 6.13 that we used in the proof of Theorem 2.4 to define $2\mathsf{Ext}_i : (\{0,1\}^n)^{i-1} \times \{0,1\}^n \to \{0,1\}^m$ where $m = \frac{k}{8\ell}$ and $k \geq 2\log(n)$ is the min-entropy requirement of each good block in our (g, ℓ, n, k)-oNOSF source $\mathbf{X} = \mathbf{X}_1, \dots, \mathbf{X}_\ell$. We then define our $\ell - 1$ output blocks as $\mathbf{O}_i = 2\mathsf{Ext}_i((\mathbf{X}_1, \dots, \mathbf{X}_{i-1}), \mathbf{X}_i)$, so $f(\mathbf{X}) = \mathbf{O}_2, \dots, \mathbf{O}_\ell$. Because there are g good blocks in $\mathbf{X}$ at indices $G_1, \dots, G_g$, we are guaranteed that $\mathbf{O}_{G_2}, \dots, \mathbf{O}_{G_g}$ are the outputs of a two-source extractor with a good block in each source. The crux of our argument then is to show that $\mathbf{O}_{G_2}, \dots, \mathbf{O}_{G_g}$ are close to uniform and independent of the adversarial blocks before them. This part of our argument follows that of [AORSV20] closely, so we do not expand on it here except to note that shortening the length of our output blocks from $m = O(k)$, not depending on ℓ, in [AORSV20] to $m = k/8\ell$ is what allows us to show that good output blocks are independent of output blocks before them.

3 Open Questions

There are several natural questions that are raised by our work. A few immediate open questions are:

1. Explicitly construct output-light two-source extractor. This would immediately imply explicit condensers for oNOSF sources and uniform aCG sources by Lemma 6.14.

2. In our condensing possibility results for uniform oNOSF sources and uniform aCG sources in Theorem 2.4 and Theorem 2.5, and our possibility results for logarithmic min-entropy oNOSF sources in Theorem 2.7, we require $\ell = o(\log(n))$, that our block size to be much smaller than the total number of blocks. It would be interesting to extend these results to smaller block sizes, such as the regime achieved for almost CG sources in [DMOZ23].

3. Is it possible to improve our condenser for uniform aCG sources in Theorem 2.4 to have constant entropy gap?

4. Can our condensing impossibility result for CG sources in Theorem 5.22 be strengthened to close the gap with the results in [DMOZ23]?

4 Preliminaries

We will generally denote distributions or sources in a bold font, such as $\mathbf{X}$, and reserve $\mathbf{U}_m$ to be the uniform distribution on m bits. When these sources are actually a sequence of sources, we use subscripts to denote blocks of that source as $\mathbf{X} = \mathbf{X}_1, \dots, \mathbf{X}_\ell$. In addition, since we often consider binary strings of length n and m, we let $N = 2^n$ and $M = 2^m$. Often it is convenient to consider strings as labels, in which case we use the notation $[N] = \{1, 2, \dots, N\}$.

4.1 Basic probability lemmas

Here, we first state a few basic probability facts that will be useful to us throughout. Our first one is a direct reverse Markov style inequality.

Claim 4.1 (Reverse Markov)**.** *Let* $\mathbf{X}$ *be a random variable taking values in* $[0,1]$*. Then, for* $0 \leq d < \mathbb{E}[X]$*, it holds that*

$$\Pr[\mathbf{X} > d] \geq \frac{\mathbb{E}[\mathbf{X}] - d}{1 - d}$$

Proof. Let $\mathbf{Y} = 1 - \mathbf{X}$. Applying Markov's inequality to $\mathbf{Y}$ gives the required bound. □

We will use the following version of the Chernoff bound:

Claim 4.2 (Chernoff Bound)**.** *Let* $\mathbf{X}_1, \ldots, \mathbf{X}_n$ *be independent random variables taking values in* $\{0,1\}$*. Let* $\mathbf{X} = \sum_i \mathbf{X}_i$*. Let* $\mu = \mathbb{E}[\mathbf{X}]$*. Then, for all* $\delta \geq 0$*, the following holds:*

$$\Pr[\mathbf{X} \geq (1+\delta)\mu] \leq \exp(-\delta^2 \mu/(2+\delta))$$

Several of our impossibility results rely on a simple TV distance bound.

Claim 4.3 (TV distance lower bound)**.** *Let* $\mathbf{X} \sim \{0,1\}^n$ *and* $S \subset \{0,1\}^n$ *be such that* $\Pr_{x \sim \mathbf{X}}[x \in S] \geq p$*. Then, for* $0 < \varepsilon < p$*, it holds that* $H^{\varepsilon}_{\infty}(\mathbf{X}) \leq \log\left(\frac{|S|}{p-\varepsilon}\right)$*.*

Proof. Let $k = \log\left(\frac{|S|}{p-\varepsilon}\right)$. Let $\mathbf{Y} \sim \{0,1\}^n$ be an arbitrary distribution with $H_{\infty}(\mathbf{Y}) \geq k$. By the min entropy condition, for all $s \in S$, it holds that $\Pr[\mathbf{Y} = s] \leq 2^{-k}$. Hence,

$$|\mathbf{X} - \mathbf{Y}| \geq \Pr_{x \in \mathbf{X}}[x \in S] - \Pr_{y \in \mathbf{Y}}[y \in S] = p - 2^{-k} \cdot |S| = \varepsilon$$

□

We will utilize the very useful min entropy chain rule in our constructions.

Lemma 4.4 (Min-entropy chain rule)**.** *For any random variables* $\mathbf{X} \sim X$ *and* $\mathbf{Y} \sim Y$ *and* $\varepsilon > 0$*,*

$$\Pr_{y \sim \mathbf{Y}}[H_{\infty}(\mathbf{X} \mid \mathbf{Y} = y) \geq H_{\infty}(\mathbf{X}) - \log|\mathsf{Supp}(\mathbf{Y})| - \log(1/\varepsilon)] \geq 1 - \varepsilon.$$

Lastly, we will later utilize a consequence of upper bounds on smooth min-entropy.

Claim 4.5 (Lemma 8.8 from [Zuc07])**.** *Let* $\mathbf{X} \sim \{0,1\}^n$ *be such that* $H^{\varepsilon}_{\infty}(\mathbf{X}) < k$*. Then, there exists* $D \subset \mathsf{Supp}(\mathbf{X})$ *such that* $|D| < 2^k$ *and* $\Pr[\mathbf{X} \in D] \geq \varepsilon$*.*

4.2 Extractors

Let $\mathbf{A} \approx_{\varepsilon} \mathbf{B}$ mean that $\mathbf{A}$ and $\mathbf{B}$ are ε close in statistical distance. Recall the definition of a seeded extractor.

Definition 4.6. *A* (k,ε)*-seeded extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *satisfies the following: for every* (n,k)*-source* $\mathbf{X}$*, and every* $\mathbf{Y} = \mathbf{U}_d$*,*

$$\mathsf{Ext}(\mathbf{X}, \mathbf{Y}) \approx_{\varepsilon} \mathbf{U}_m.$$

d is called the seed length *of* Ext*.* Ext *is called strong if*

$$\mathsf{Ext}(\mathbf{X}, \mathbf{Y}), \mathbf{Y} \approx_{\varepsilon} \mathbf{U}_m, \mathbf{Y}.$$

A useful fact about strong seeded extractors that they work even when the seed is not fully uniform. (See for example Lemma 6.4 from [CGL20] for a proof.)

Lemma 4.7. *Let* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *be a strong* (k,ε)*-seeded extractor. Let* $\mathbf{X}$ *be a* (n,k)*-source and let* $\mathbf{Y}$ *be a* $(d, d-\lambda)$*-source. Then,*

$$|\mathsf{Ext}(\mathbf{X},\mathbf{Y}), \mathbf{Y} - \mathbf{U}_m, \mathbf{Y}| \leq 2^{\lambda}\varepsilon.$$

We will use the following construction of seeded extractors:

Theorem 4.8 (Theorem 1.5 in [GUV09])**.** *For all constant* $\alpha > 0$ *and all* n, k, ε*, there exists an explicit* (k,ε)*-seeded extractor* $\mathsf{sExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with* $d = O(\log(n/\varepsilon))$ *and* $m \geq (1-\alpha)k$*.*

In addition, we will use a generalization of seeded extractors, two-source extractors, that only require the second source to be independent from the first and not necessarily be uniform.

Definition 4.9. *A function* $2\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a* (k_1, k_2, ε)-two-source extractor *if for every* (n_1, k_1)*-source* $\mathbf{X}_1$ *and* (n_2, k_2)*-source* $\mathbf{X}_2$ *where* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are independent of each other, we have*

$$2\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2) \approx_{\varepsilon} \mathbf{U}_m.$$

It is said to be strong in the first argument *if*

$$2\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2), \mathbf{X}_1 \approx_{\varepsilon} \mathbf{U}_m, \mathbf{X}_1.$$

Similarly, one can define $2\mathsf{Ext}$ that is strong in the second argument. If $2\mathsf{Ext}$ is strong in both arguments, we simply say that it is *strong*. We use the fact that inner product function is a good two source extractor:

Theorem 4.10. *[CG88, Vaz85, ILL89] Let* $\mathbf{X}, \mathbf{Y} \sim \{0,1\}^n$ *with* $H_\infty(\mathbf{X}) = k_1, H_\infty(\mathbf{Y}) = k_2$*. Let* $m = \frac{n}{r}$ *for some* $r \in \mathbb{N}$*. Let* $\mathsf{IP}(x,y) : \{0,1\}^{2n} \to \{0,1\}^m$ *be the function that interprets* x, y *as elements of* $\mathbb{F}_{2^m}^r$ *and outputs the* m *bit string corresponding to* $x \cdot y$*. Then,* $|\mathsf{IP}(\mathbf{X},\mathbf{Y}) - \mathbf{U}_m| \leq 2^{(n+m-k_1-k_2)/2}$.

For a proof of the above theorem, see Theorem 2.5.3 in [Cha16].

4.3 Randomness sources relevant to our work

We now formally introduce the randomness sources that are relevant to our work. We begin with NOSF sources, which have no restrictions on the adversary producing the bad blocks.

Definition 4.11 (NOSF source)**.** *A* (g, ℓ, n, k)-NOSF source (NOSF) $\mathbf{X}$ *with symbols in* $\Sigma = \{0,1\}^n$ *and length* ℓ *is over* Σ^ℓ*, written as* $\mathbf{X} = \mathbf{X_1}, \ldots, \mathbf{X}_\ell$*, and has the following property: There exists a set of good blocks* $\mathcal{G} \subseteq [\ell]$ *such that* $|\mathcal{G}| \geq g$ *and the random variables in* $\{\mathbf{X}_i\}_{i\in\mathcal{G}}$ *are each independently sampled* (n,k)*-sources. We say that a block* $\mathbf{X}_i$ *is* good *if* $i \in \mathcal{G}$ *and* bad *otherwise.*

Note that we have no restrictions on how bad blocks may depend on the good blocks. If $k = n$, we say that $\mathbf{X}$ is a *uniform* (g, ℓ, n)*-NOSF* source. When n is implicit or not relevant, we simply call $\mathbf{X}$ a uniform (g,ℓ)-NOSF source. Next, we introduce oNOSF sources by restricting the NOSF adversary.

Definition 4.12 (One-sided NOSF source)**.** *A (g, ℓ, n, k)-*oNOSF source $\mathbf{X}$ *with symbols in $\Sigma = \{0,1\}^n$ and length ℓ is over Σ^ℓ, written as $\mathbf{X} = \mathbf{X_1}, \ldots, \mathbf{X}_\ell$, and has the following property: There exists a set of good blocks $\mathcal{G} \subseteq [\ell]$ such that $|\mathcal{G}| \geq g$ and the random variables in $\{\mathbf{X}_i\}_{i \in \mathcal{G}}$ are each independently sampled (n,k)-sources such that $\mathbf{X}_i$ is independent of $\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}$. We say that a block $\mathbf{X}_i$ is* good *if $i \in \mathcal{G}$ and* bad *otherwise.*

Remark 4.13. *One-sided NOSF sources are also NOSF sources because the adversary in oNOSF sources is strictly weaker than that of NOSF sources.*

These oNOSF sources are special cases of the SHELA sources from [AORSV20]. We now introduce SHELA sources in their full generality.

Definition 4.14 (SHELA source [AORSV20])**.** *A distribution $\mathbf{X}$ over $(\{0,1\}^n)^\ell$ is a (g, ℓ, n, k)-*Somewhere Honest Entropic Look Ahead (SHELA) source *if there exists a (possibly randomized) adversary $\mathcal{A}$ such that $\mathbf{X}$ is produced by sampling g out of ℓ indices to place independently sampled (n,k)-sources and then placing adversarial blocks in the other $\ell - g$ indices that may depend arbitrarily on any block that comes before it.*

Concretely, there must exist random variables $1 \leq \mathbf{I}_1 < \mathbf{I}_2 < \cdots < \mathbf{I}_g \leq \ell$ with arbitrary joint distribution, denoting the indices of the independent (n,k)-sources, and g independent (n,k)-sources $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_g$ such that $\mathbf{X}$ is generated in the following manner:

1. *Sample $(i_1, i_2, \ldots, i_g) \sim (\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_g)$.*
2. *For all $j \in [g]$ set $\mathbf{B}_{i_j} = \mathbf{Z}_j$.*
3. *For all $i \in [\ell] \setminus \{i_1, i_2, \ldots, i_g]$, the adversary sets $\mathbf{B}_i = \mathcal{A}(\mathbf{B}_1, \ldots, \mathbf{B}_{i-1}, i_1, \ldots, i_g\}$.*
4. *Finally, let $\mathbf{X} = (\mathbf{B}_1, \ldots, \mathbf{B}_\ell)$.*

We will generally call the blocks $\mathbf{Z}_1, \ldots, \mathbf{Z}_g$ the "good" blocks and the remaining blocks "bad" blocks.

Similar to NOSF sources, when $k = n$ we will simply say $\mathbf{X}$ is a (g, ℓ, n)-*uniform SHELA* source, and when n is implicit we will simplify further to a uniform (g, ℓ)-SHELA source.

While working over oNOSF sources is easier than working over general SHELA sources, all of our results still apply to general SHELA sources since SHELA sources are convex combinations of oNOSF sources.

Proposition 4.15. *Every (g, ℓ, n, k)-SHELA source $\mathbf{X}$ is a convex combination of (g, ℓ, n, k)-oNOSF sources.*

Proof. Let $\mathbf{I} = \mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_g$ be the distribution of indices used in the construction of $\mathbf{X}$. For a sample $\mathcal{I} \sim \mathbf{I}$, let $\mathbf{X}_\mathcal{I}$ be the (g, ℓ, n, k)-oNOSF source in the construction of which the adversary chose the good blocks to be at indices $\mathcal{I}$ and the functions describing the bad blocks to be identical to those of $\mathbf{X}$ when the sample of indices from I is $\mathcal{I}$. That is, when $\mathcal{I}$ is sampled in the construction of $\mathbf{X}$ we have for all $j \in [\ell] \setminus \mathcal{I}$ that $\mathbf{X}_j = (\mathbf{X}_\mathcal{I})_j$ as functions.

With this setup, we directly have that $\mathbf{X} = \mathbb{E}_{\mathcal{I} \sim \mathbf{I}}[\mathbf{X}_\mathcal{I}]$, so $\mathbf{X}$ is a convex combination of $\mathbf{X}_\mathcal{I}$'s. □

Lastly, we define adversarial Chor-Goldreich (CG) sources, which have an adversary like that of oNOSF sources that can depend arbitrarily on past blocks, but the adversary of adversarial CG sources can have some effect on future blocks, unlike that of oNOSF sources.

Definition 4.16 (Adversarial CG source)**.** *A* (g, ℓ, n, k)-aCG source $\mathbf{X}$ *with symbols in* $\Sigma = \{0,1\}^n$ *and length* ℓ *is over* Σ^ℓ*, written as* $\mathbf{X} = \mathbf{X_1}, \ldots, \mathbf{X}_\ell$*, and has the following property: There exists a set of good blocks* $\mathcal{G} \subseteq [\ell]$ *such that* $|\mathcal{G}| \geq g$ *and the random variables in* $\{\mathbf{X}_i\}_{i \in \mathcal{G}}$ *have the property that for all prefixes* $(a_1, \ldots, a_{i-1}) \in (\{0,1\}^n)^{i-1}$*,*

$$H_\infty(\mathbf{X}_i \mid \mathbf{X}_1, \ldots, \mathbf{X}_{i-1} = a_1, \ldots, a_{i-1}) \geq k.$$

As before, if $k = n$ then we simply call $\mathbf{X}$ a uniform (g, ℓ, n)-aCG source, and we omit n when it is implicit.

We have introduced all of these definitions since our results resolve open questions for each. The relationship between all these definitions is necessary to clearly see how our lower and upper bounds apply. In line with this, we show an equivalence between uniform oNOSF sources and uniform aCG sources.

Proposition 4.17. *A source* $\mathbf{X}$ *is a uniform oNOSF source if and only if it is a uniform aCG source.*

Proof. Say $\mathbf{X}$ is a uniform (g, ℓ, n)-oNOSF source. Then, because bad blocks may only depend on the good blocks that have a lower index than them and all the good blocks are sampled independently, the good blocks satisfy the prefix condition in Definition 4.16 to give us that $\mathbf{X}$ is a uniform (g, ℓ, n)-aCG source.

On the other hand, say that $\mathbf{X}$ is a uniform (g, ℓ, n)-aCG source. Then the fact that for a good block $\mathbf{X}_i$ we have for all $(a_1, \ldots, a_{i-1}) \in (\{0,1\}^n)^{i-1}$ that $H_\infty(\mathbf{X}_i \mid \mathbf{X}_1, \ldots, \mathbf{X}_{i-1} = a_1, \ldots, a_{i-1}) = n$, so $\mathbf{X}_i$ is uniform given any prefix, means that $\mathbf{X}_i$ is independent of all blocks that come before it. In particular, this means that bad blocks may only depend on the good blocks that come before them. In addition, the good blocks being uniform clearly means that they are independent from each other. Hence, $\mathbf{X}$ is a uniform (g, ℓ, n)-oNOSF source as well. □

Therefore, when we prove a condensing impossibility result by constructing a oNOSF source, that same result applies to NOSF sources and aCG sources sources as well. On the other hand, our condensing possibility results for uniform oNOSF sources also apply to uniform aCG sources, but our results for non-uniform oNOSF sources may not apply to non-uniform aCG sources.

5 Impossibility Results

In this section, we prove condensing impossibility results for uniform NOSF sources and uniform oNOSF sources. First, in Section 5.1 we demonstrate condensing impossibility results for all three classes of sources when $g \leq \ell/2$. Then, in Section 5.2 we show a condensing impossibility result for uniform (g, ℓ)-NOSF sources for arbitrary settings of g and ℓ. Finally, we use a result from Section 5.1 to show the impossibility of condensing from low min-entropy CG sources in Section 5.3.

5.1 Impossibility of condensing when $g \leq \ell/2$

We will prove that for $g \leq \ell/2$, it is impossible to condense from uniform (g, ℓ)-oNOSF sources to rate more than $\frac{1}{\lfloor \ell/g \rfloor}$. As we noted in Remark 4.13 and Proposition 4.17, these results apply to uniform NOSF sources and uniform aCG sources as well.

Theorem 5.1. *For all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $g, \ell \in \mathbb{N}$ with $g \leq \ell/2$ and for all $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g,ℓ)-oNOSF source $\mathbf{X}$ so that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{1}{\lfloor \ell/g \rfloor} \cdot m + \delta$.*

This implies that for the special case when g divides ℓ, any non-trivial condensing is impossible.

Corollary 5.2. *For all $\varepsilon > 0, g, \ell \in \mathbb{N}$ with $g \mid \ell$, there exists a $\delta > 0$ such that: for all functions $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g,ℓ)-oNOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) < \frac{g}{\ell} \cdot m + \delta$.*

Proof. Follows immediately from Theorem 5.1. □

The proof of Theorem 5.1 involves two ingredients. First, we show that for the special case of $g = 1$, condensing above rate $\frac{1}{\ell}$ is impossible for uniform $(1,\ell)$-oNOSF sources. Second, extend these results to uniform (g,ℓ)-oNOSF source with $g \leq \ell/2$ by showing that if it is impossible to condense from uniform $(1,\ell')$-oNOSF sources, then it is impossible to condense above rate $\frac{1}{\ell'}$ from uniform (g,ℓ)-oNOSF sources when $\frac{g}{\ell} \leq \frac{1}{\ell'}$.

Formally, these two lemmas are as follows:

Lemma 5.3. *For all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all functions $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform $(1,\ell)$-oNOSF source $\mathbf{X}$ so that $H^\varepsilon_\infty(f(\mathbf{X})) < \frac{1}{\ell} \cdot m + \delta$.*

Lemma 5.4. *Let $g, \ell, \ell', n', n, m \in \mathbb{N}$ be such that $\ell' \leq \ell, \frac{g}{\ell} \leq \frac{1}{\ell'}, \lceil \ell/\ell' \rceil n < n'$. Let $0 < \varepsilon < 1, \delta > 0$ be such that: for any function $f : (\{0,1\}^{n'})^{\ell'} \to \{0,1\}^m$, there exists a uniform $(1,\ell')$-oNOSF source $\mathbf{Y}$ so that $H^\varepsilon_\infty(f(\mathbf{Y})) < \frac{1}{\ell'} \cdot m + \delta$. Then, for any function $h : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g,ℓ)-oNOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(h(\mathbf{X})) \leq \frac{1}{\ell'} \cdot m + \delta$.*

Our main theorem follows by combining these two lemmas.

Proof of Theorem 5.1. Divide ℓ by g so that $\ell = c \cdot g + r$ where $c \geq 1, r < g \in \mathbb{N}$. We can derive our desired impossibility result by applying Lemma 5.4 to the result of Lemma 5.3 uniform $(1,c)$-oNOSF sources. □

We defer the proof of Lemma 5.4 until Section 5.4. In the next subsubsection, we will focus on proving Lemma 5.3.

5.1.1 Proving main theorem for the case of $g = 1$

We prove this lemma by showing that if one cannot condense from uniform (g,ℓ)-oNOSF sources, then one cannot condense from uniform $(g,\ell+g)$-oNOSF sources.

Lemma 5.5. *Let $c_0, c_1, \varepsilon, \delta \in \mathbb{R}$ and $g, n, \ell \in \mathbb{N}$ be such that $g \leq \ell, 0 < c_0 < 1, \varepsilon < c_1 < 1$. Assume that for all $A \in \mathbb{N}$ and function $f : (\{0,1\}^n)^\ell \to [A]$, there exists a uniform (g,ℓ)-oNOSF source (uniform (g,ℓ)-NOSF source, respectively) $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot \log(A) + \delta$. Then, for all $M \in \mathbb{N}$ and every function $h : (\{0,1\}^n)^{\ell+g} \to [M]$, there exists a uniform $(g,\ell+g)$-oNOSF source (uniform (g,ℓ)-NOSF source, respectively) $\mathbf{Y}$ such that $H^\varepsilon_\infty(h(\mathbf{Y})) \leq \frac{g}{\ell+g} \cdot m + \delta'$ where $\delta' = \max\left(\log\left(\frac{c_1}{(1-c_1)c_0(c_1-\varepsilon)}\right), \delta + \frac{\log(c_0)g}{\ell}\right)$ and $m = \log(M)$.*

We remark that Lemma 5.5 paves the way for an inductive argument and we instantiate it to prove Lemma 5.3 as follows:

Proof of Lemma 5.3. We inductively apply Lemma 5.5 with $g = 1$ and arbitrary ℓ to prove the claim. Notice that for all distributions $\mathbf{X}$ on $\{0,1\}^m$, $H^{\varepsilon}_{\infty}(\mathbf{X}) \leq m$. For the base case of $g = 1, \ell = 1$: for any function f and uniform $(1,1)$-oNOSF source $\mathbf{W}$, it must be that $H^{\varepsilon}_{\infty}(f(\mathbf{W})) \leq m$. Now, inductively apply Lemma 5.5 by setting $c_0 = 1, c_1 = \frac{1+\varepsilon}{2}, \delta = \log\left(\frac{2(1+\varepsilon)}{(1-\varepsilon)^2}\right)$ to infer the claim for uniform $(1,\ell)$-NOSF sources. □

5.1.2 Recursive impossibility lemma

To prove Lemma 5.5, we find a dominating set in dense bipartite graphs with left degree lower bound. We will use it to construct a uniform oNOSF source that will serve as a counterexample for a candidate condenser.

Lemma 5.6 (Small Dominating Set in Bipartite Graph)**.** *Let $c_0 > 0, 0 < c_1 < 1, \delta > 0 \in \mathbb{R}, N, M \in \mathbb{N}$ be arbitrary. Let $G = (U, V, E)$ be a bipartite graph with $|U| = N$, $|V| = M$, such that for all $u \in U : \deg(u) \geq c_0 \cdot M^{\delta}$. Then, there exists $D \subseteq V$ with $|D| \leq \frac{c_1}{(1-c_1)c_0} \cdot M^{1-\delta}$ such that $|Nbr(D)| \geq c_1 N$.*

Using this dominating set lemma, we prove Lemma 5.5.

Proof of Lemma 5.5. Fix a function $h : (\{0,1\}^n)^{\ell+g} \to [M]$. We will construct a uniform $(g, \ell+g)$-oNOSF source (uniform $(g,\ell+g)$-NOSF source respectively) $\mathbf{Y}$ such that $H^{\varepsilon}_{\infty}(f(Y)) < \frac{g}{\ell+g} \cdot m + \delta'$. Let $N = 2^n$. We consider two cases:

Case 1. For all $(x_1, \ldots, x_g) \in (\{0,1\}^n)^g : |\operatorname{Supp}(h(x_1, \ldots, x_g, \mathbf{U}_\ell))| \geq c_0 M^{\ell/(\ell+g)}$.
Consider an undirected bipartite graph $G = (U, V, E)$ where $U = (\{0,1\}^n)^g$ and $V = [M]$ with edge $e = (u, v) \in E$ where $u = (x_1, \ldots, x_g) \in U$ and $v \in V$ iff there exist $x_{g+1}, \ldots, x_{\ell+g}$ such that $h(x_1, \ldots, x_{\ell+g}) = v$. Applying Lemma 5.6 to G, there exists $D \subset [M]$ such that $|D| \leq \frac{c_1}{(1-c_1)c_0} M^{g/(\ell+g)}$ and for $c_1 N^g$ many tuples $(x_1, \ldots, x_g) \in (\{0,1\}^n)^g$, there exist $y_1, \ldots, y_\ell \in (\{0,1\}^n)^\ell$ such that $h(x_1, \ldots, x_g, y_1, \ldots, y_\ell) \in D$. Let $\mathsf{Adv} : (\{0,1\}^n)^g \to (\{0,1\}^n)^\ell$ be defined as:

$$\mathsf{Adv}(x_1, \ldots, x_g) = \begin{cases} (y_1, \ldots, y_\ell) & \text{if there exist } y_1, \ldots, y_\ell \text{ such that } h(x_1, \ldots, x_g, y_1, \ldots, y_\ell) \in D \\ (0^n)^\ell & \text{otherwise} \end{cases}$$

Consider the uniform $(g, \ell+g)$-oNOSF source (uniform $(g,\ell+g)$-NOSF source respectively) $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_{\ell+g})$ such that $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are independent uniform distributions and $(\mathbf{X}_{g+1}, \ldots, \mathbf{X}_{g+\ell}) = \mathsf{Adv}(\mathbf{X}_1, \ldots, \mathbf{X}_g)$. Then, with probability c_1, $h(\mathbf{X}) \in D$. Applying Claim 4.3, we infer that $H^{\varepsilon}_{\infty}(\mathbf{X}) \leq \log\left(\frac{c_1 M^{g/(\ell+g)}}{(1-c_1)c_0(c_1-\varepsilon)}\right) = \frac{g}{\ell+g} \cdot m + \log\left(\frac{c_1}{(1-c_1)c_0(c_1-\varepsilon)}\right) \leq \frac{g}{\ell+g} \cdot m + \delta'$.

Case 2. There exist $x_1, \ldots, x_g \in (\{0,1\}^n)^g$ such that $|\operatorname{Supp}(h(x_1, \ldots, x_g, \mathbf{U}_\ell))| \leq c_0 M^{\ell/(\ell+g)}$.
Let $S = \operatorname{Supp}(h(x_1, \ldots, x_g, \mathbf{U}_\ell))$. Define $f : \{0,1\}^\ell \to S$ by $f(y_1, \ldots, y_\ell) = h(x_1, \ldots, x_g, y_1, \ldots, y_\ell)$. Then, by assumption, there exists uniform (g,ℓ)-oNOSF source (uniform (g,ℓ)-NOSF source, respectively) $\mathbf{Y}$ such that $H^{\varepsilon}_{\infty}(f(\mathbf{Y})) \leq \frac{g}{\ell} \cdot \log(|S|) + \delta$. Consider uniform $(g, \ell+g)$-oNOSF source $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_{\ell+g})$ where distributions $\mathbf{X}_1, \ldots, \mathbf{X}_g$ always output $x_1, \ldots, x_g$ and $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_{\ell+g}$ are distributed as $\mathbf{Y}$. Then,

$$H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot \log(|S|) + \delta \leq \frac{g}{\ell+g} \cdot m + \delta + \frac{\log(c_0) g}{\ell} \leq \frac{g}{\ell+g} \cdot m + \delta'$$

□

5.1.3 Finding small dominating set in bipartite graphs

We now directly prove Lemma 5.6.

Proof of Lemma 5.6. We construct D via a greedy algorithm specified in Algorithm 1. This algorithm greedily chooses right vertices in V with highest degree, adds them to D, and stops once the neighborhood of D, gets large enough. To analyze this algorithm, we can use loose bounds on the

Algorithm 1:

$i \leftarrow 0$
$D \leftarrow \varnothing$
$G_0 = (U_0, V_0, E_0) \leftarrow G = (U, V, E)$
while $|Nbr(D)| < c_1 N$ **do**
  Let $v_i \in V_i$ be the vertex of maximum degree in G_i
  $D \leftarrow D \cup \{v_i\}$
  $V_{i+1} \leftarrow V_i \setminus \{v_i\}$
  $U_{i+1} \leftarrow U_i \setminus \mathrm{Nbr}(v_i)$
  $E_{i+1} \leftarrow E_i \setminus \{(u, v) \in E : v = v_i \text{ or } u \in \mathrm{Nbr}(v_i)\}$
  $G_{i+1} \leftarrow (U_{i+1}, V_{i+1}, E_{i+1})$
end

number of edges and vertices at any one step. As the algorithm stops once at least $c_1 N$ vertices are removed from U, for all iterations i, $|U_i| \geq (1 - c_1)N$. In addition, because left vertices are only removed when one of their neighbors in V is added to D, the remaining vertices in U always have their original degrees intact. So, for all iterations i and for all $u \in U_i$, it holds that $\deg(u) \geq c_0 \cdot M^\delta$. So,

$$|E_i| \geq |U_i|\, c_0 \cdot M^\delta \geq (1 - c_1) N c_0 \cdot M^\delta$$

Observe that for all i, $|V_i| \leq |V| = M$. So,

$$\deg(v_i) \geq \frac{|E_i|}{|V_i|} \geq \frac{(1 - c_1) N c_0 \cdot M^\delta}{M} = \frac{(1 - c_1) c_0 N}{M^{1-\delta}}.$$

The algorithm terminates when at least $c_1 N$ vertices are added to D and at each step $\deg(v_i)$ vertices are added to D. Hence, the number of iterations for which the algorithm runs is at most

$$\frac{c_1 N}{\frac{(1-c_1)c_0 N}{M^{1-\delta}}} = \frac{c_1}{(1 - c_1)c_0} \cdot M^{1-\delta}$$

The claim follows since exactly 1 vertex is added to D in each iteration. □

5.2 Impossibility of condensing from uniform (g, ℓ)-NOSF sources

Our main theorem in this subsection is that it is impossible to condense from uniform (g, ℓ)-NOSF sources where $g \geq \frac{\ell}{2} + 1$. Using it and previous results, we obtain impossibility results for all g, ℓ.

Theorem 5.7. *There exists a universal constant $c > 0$ such that for all $g, \ell, m, n \in \mathbb{N}$ with $\ell/2 < g < \ell$, there exist $\varepsilon = \left(\frac{1}{c\ell}\right)^{\ell - g}, \delta = c \cdot \ell^2 \log(\ell)$ so that the following holds: for any function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g, ℓ)-NOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + \delta$.*

We also infer the following useful corollary that shows that uniform (g,ℓ)-NOSF sources cannot be condensed beyond rate $1-1/\ell'$ with error $O(1/\ell')$ where ℓ' is the smallest integer such that $g/\ell \leq 1-1/\ell'$.

Corollary 5.8. *There exists a universal constant c such that the following holds: For all $g,\ell,\ell',m,n \in \mathbb{N}$ where ℓ' is the smallest integer such that $\frac{g}{\ell} \leq \frac{\ell'-1}{\ell'}$, there exist $\varepsilon = \frac{1}{c\ell'}, \delta = c\cdot(\ell')^2\log(\ell')$ so that the following holds: for all functions $f:(\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g,ℓ)-NOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{1}{\ell'}\cdot m+\delta$.*

We also get a stronger impossibility result for uniform (g,ℓ)-NOSF sources (compared to condensing impossibility for uniform (g,ℓ)-oNOSF sources proved in Theorem 5.1) for the regime $g \leq \ell/2$.

Corollary 5.9. *There exists a universal constant c such that for all $\ell,g,r,m,n \in \mathbb{N}$ with $\ell \mod g = r$, there exist $\varepsilon = \left(\frac{1}{c(g+r)}\right)^r, \delta = c\cdot(r+g)^2\log(g+r)$ so that the following holds: for all functions $f:(\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g,ℓ)-NOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell}\cdot m+\delta$.*

Proof. For $\ell = g+r$, apply Lemma 5.12 (see below) to infer there exists a universal constant c such that the claim holds for $\varepsilon = \left(\frac{1}{c(g+r)}\right)^r$, $\delta = c\cdot(r+g)^2\log(g+r)$. Now, recursively apply Lemma 5.5 with these ε,δ, setting $c_0 = 1, c_1 = \frac{1+\varepsilon}{2}$ to infer the claim. When applying Lemma 5.5, we take advantage of the fact that $\varepsilon < 3/4$ and that c is a large enough constant to get that $\delta \geq \log\left(\frac{2(1+\varepsilon)}{(1-\varepsilon)^2}\right)$. □

We obtain impossibility result for all uniform $(ag,a\ell)$-NOSF sources where g and ℓ are constants and $a \in \mathbb{N}$ is arbitrarily large.

Corollary 5.10. *For all fixed $g,\ell \in \mathbb{N}$, there exist constants $\varepsilon,\delta > 0$ so that the following holds: for all $a,m,n \in \mathbb{N}$ and for all functions $f:(\{0,1\}^n)^{a\ell} \to \{0,1\}^m$, there exists a uniform $(ag,a\ell)$-NOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell}\cdot m+\delta$.*

We also record the special case of when the total number of blocks ℓ is a constant.

Corollary 5.11. *For all fixed $g,\ell \in \mathbb{N}$, there exist constants $\varepsilon,\delta > 0$ so that the following holds: For all $m,n \in \mathbb{N}$ and for all functions $f:(\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g,ℓ)-NOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell}\cdot m+\delta$.*

Proof. Directly follows by setting $a=1$ in Corollary 5.10. □

We prove our main theorem using the following general version of the theorem which we denote as our main lemma:

Lemma 5.12. *There exists universal constants c such that for all $c_0 > 0, g,\ell,M,n \in \mathbb{N}$ with $\ell/2 < g < \ell$, and for all $A \subset (\{0,1\}^n)^\ell$ with $|A| = c_0(2^n)^\ell$, the following holds: for any function $f:(\{0,1\}^n)^\ell \to [M]$, there exists a uniform (g,ℓ)-NOSF source $\mathbf{X}$, $A' \subset A \cap \mathrm{Supp}(\mathbf{X})$ and $D \subset [M]$ such that $f(A') \subset D$ where $|A'| \geq c_0\cdot\left(\frac{1}{c\ell}\right)^{\ell-g}\cdot N^g$, and $|D| \leq (c\ell)^{\ell^2}\cdot\left(\frac{2}{c_0}\right)^g\cdot M^{g/\ell}$.*

Using this main lemma, the theorem follows:

Proof of Theorem 5.7 assuming Lemma 5.12. Applying Lemma 5.12 with $A = M = \{0,1\}^m$, we infer that there exists uniform (g,ℓ)-NOSF source $\mathbf{X}$, universal constant c_0 and $D \subset \{0,1\}^m$ such that $|D| \leq (c_0\ell)^{\ell^2} \cdot 2^g \cdot M^{g/\ell}$ and $\Pr[f(\mathbf{X}) \in D] \geq \left(\frac{1}{c_0\ell}\right)^{\ell-g}$. Applying Claim 4.3 with $\varepsilon = \frac{1}{2} \cdot \left(\frac{1}{c_0\ell}\right)^{\ell-g}$, we infer that

$$\begin{aligned} H^{\varepsilon}_{\infty}(f(\mathbf{X})) &\leq \log\left(\frac{|D|}{\varepsilon/2}\right) \\ &\leq \frac{g}{\ell} \cdot m + \ell^2 \log(c_0\ell) + g + (\ell - g)\log(c_0\ell) + 1 \\ &\leq \frac{g}{\ell} \cdot m + c \cdot \ell^2 \log(\ell) \end{aligned}$$

where c is a large enough universal constant. As $\varepsilon = \frac{1}{2} \cdot \left(\frac{1}{c_0\ell}\right)^{\ell-g} \geq \left(\frac{1}{c\ell}\right)^{\ell-g}$, we infer the claim. □

To prove our corollary regarding condensing uniform (g,ℓ)-NOSF sources where $\frac{g}{\ell}$ is a large constant, we will use the following lemma:

Lemma 5.13. *Let $g, \ell, \ell', n', n, m \in \mathbb{N}$ be such that $\frac{g}{\ell} \leq \frac{\ell'-1}{\ell'}$, $\lceil \ell/\ell' \rceil n < n'$. Let $0 < \varepsilon < 1, \delta > 0$ be such that: for any function $f : (\{0,1\}^{n'})^{\ell'} \to \{0,1\}^m$, there exists a uniform $(\ell'-1,\ell')$-NOSF source $\mathbf{Y}$ so that $H^{\varepsilon}_{\infty}(f(\mathbf{Y})) \leq \frac{\ell'-1}{\ell'} \cdot m + \delta$. Then, for any function $h : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$, there exists a uniform (g,ℓ)-NOSF source $\mathbf{X}$ such that $H^{\varepsilon}_{\infty}(h(\mathbf{X})) \leq \frac{\ell'-1}{\ell'} \cdot m + \delta$.*

We will prove this lemma in a later in Section 5.4. Using it, the corollary immediately follows:

Proof of Corollary 5.8. We apply Theorem 5.7 to uniform $(\ell'-1,\ell')$-NOSF sources and use it in Lemma 5.13 to infer the claim. □

To prove our corollary regarding condensing uniform $(ag, a\ell)$-NOSF sources where g and ℓ are constants and a is arbitrary, we will use the following lemma that allows us to generalize the impossibility result:

Lemma 5.14. *Let $g, \ell \in \mathbb{N}$ and $0 < \varepsilon < 1, \delta > 0$ be such that for all $n, m \in \mathbb{N}$ and all functions $f : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$, there exists an uniform (g,ℓ)-NOSF source $\mathbf{X}$ such that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + \delta$. Then, for all $a, n, m \in \mathbb{N}$ and all functions $f : (\{0,1\}^n)^{a\ell} \to \{0,1\}^m$, there exists an uniform $(ag, a\ell)$-NOSF source $\mathbf{X}$ such that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + \delta$.*

We will also prove this lemma in Section 5.4. Using it, the corollary immediately follows:

Proof of Corollary 5.10. We apply Corollary 5.9 with g, ℓ to infer that there exist $0 < \varepsilon < 1, \delta > 1$ such that for all $n, m \in \mathbb{N}$ and all functions $f : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$, there exists an uniform (g,ℓ)-NOSF source $\mathbf{X}$ such that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + \delta$. Finally, we apply Lemma 5.14 to infer the claim. □

5.2.1 Proving the main lemma

Here, we will prove Lemma 5.12. We first introduce some helpful notation for this part. For an edge $e \in E$, let $\chi(e)$ denote the color of e in H. For a vertex $x \in H$, let

$$\mathrm{Nbr}_H(x) = \{y \in H : (x, y) \in E\}.$$

Similarly, for a vertex $x \in H$, and color $\gamma \in [M]$, let

$$\mathrm{Nbr}_H(x, \gamma) = \{y \in H : (x, y) \in E \text{ and } \chi(x, y) = \gamma\}.$$

To prove the main lemma, we will utilize the following special case of the main lemma, corresponding to the case of $g = \ell - 1$, that we prove later:

Lemma 5.15. *There exists a universal constant $c > 0$ such that for all $M, n, \ell \geq 3 \in \mathbb{N}$, and $A \subset (\{0,1\}^n)^\ell$ with $|A| = c_0(2^n)^\ell$, the following holds: for any function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform $(\ell - 1, \ell)$-NOSF source $\mathbf{X}$, $A' \subset A \cap \mathrm{Supp}(\mathbf{X})$ with $|A'| \geq \frac{1}{c} \cdot \frac{c_0}{\ell} \cdot N^{\ell-1}$, and $D \subset [M]$ with $|D| \leq c \cdot \frac{1}{\ell^2} \cdot \left(\frac{2}{c_0}\right)^{\ell-2} \cdot M^{(\ell-1)/\ell}$ such that $f(A') \subset D$.*

The main lemma follows by an inductive argument where the special case above is the base case.

Remark 5.16. *In proof of Lemma 5.12, one can use $g = \ell$ as the base case as well. However, for clarity's sake we use $g = \ell - 1$ as the base case. For first time readers, it will be helpful to first read the direct non-inductive proof of Lemma 5.15 presented in Section 5.2.2 before reading the proof of Lemma 5.12 as both these proofs share a lot of ideas.*

Proof of Lemma 5.12. Let $N = 2^n$. We will often identify $\{0,1\}^n$ with $[N]$ wherever convenient. We let c be a very large universal constant.

We proceed by induction on $b = \ell - g$. Formally, for $b \geq 1 \in \mathbb{N}$ we will prove the claim for $\ell, g \in \mathbb{N}$ with $\ell/2 < g < \ell$ such that $b = \ell - g$.

For the base case, we let $b = 1$ and apply Lemma 5.15 to infer the claim.

For the inductive step, say we want to prove the hypothesis for $b \geq 2$ assuming the hypothesis holds for $b - 1$. Fix some g, ℓ such that $\ell - g = b$. Let $c_1 = \frac{1}{4\ell}, c_2 = \left(\frac{3}{2}\right)^{1/\ell^2} - 1, c_3 = \frac{2}{3}, c_4 = \frac{c_0}{4}$. By binomial approximation, there exists a constant $\alpha \geq 1$ such that $\frac{1}{\alpha} \cdot \frac{0.4}{\ell^2} \leq c_2 \leq \alpha \cdot \frac{0.6}{\ell^2}$ for all ℓ. For all positions $p \in [\ell]$, let $S_p \subset [N]^{\ell-1}$ be defined as follows: $(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell) \in S_p$ if and only if

$$|\{f(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) : y \in \{0,1\}^n \wedge (x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A\}| \geq c_2 M^{1/\ell}$$

We consider various cases:

Case 1. There exists $p \in [\ell]$ such that $|S_p| \geq c_0 c_1 N^{\ell-1}$.
Consider the bipartite graph $G = (U, V, E)$ where $U = S_p, V = [M]$ and edge $e = (u, v) = ((x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell), z) \in E$ if and only if there exists $y \in \{0,1\}^n$ such that $f(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) = z$ and $(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A$. Then by assumption, for all $u \in U$, it holds that $\deg(u) \geq c_2 M^{1/\ell}$. We apply Lemma 5.6 to G and infer that there exists $D_{end} \subset V$ such that $|D_{end}| \leq \frac{c_3}{c_2(1-c_3)} M^{(\ell-1)/\ell}$ and $\mathrm{Nbr}_G(D_{end}) \geq c_0 c_1 c_3 |U| \geq c_0 c_1 c_3 N^{\ell-1}$.

Let $A'_{end} \subset A$ be defined as follows: for a vertex $u = (x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell) \in \mathrm{Nbr}_G(D_{end})$, we add $(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell)$ to A'_{end} where y is such that $(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A$ and $f(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in D_{end}$ (we only pick one such y per u and if multiple such y exist, we break ties arbitrarily). Let $z \in D_{end}$ be an arbitrary element. Let $f_{end} : (\{0,1\}^n)^{\ell-1} \to D_{end}$ be defined as follows:

$$f_{end}(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell) = \begin{cases} f(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) & \exists y : (x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A'_{end} \\ z & \text{otherwise} \end{cases}$$

We now use inductive hypothesis on the candidate function f_{end} (having range D_{end}), and restriction set $\mathrm{Nbr}_G(D_{end})$. Notice that $|\mathrm{Nbr}_G(D_{end})| \geq c_0^{end} = c_0 c_1 c_3$. We infer there exists uniform $(g, \ell-1)$-NOSF source $\mathbf{X}^{ind}$, $A'_{ind} \subset \mathrm{Nbr}_G(D_{end}) \cap \mathrm{Supp}(\mathbf{X}^{ind})$, and $D_{ind} \subset D_{end}$ such that $f(A'_{ind}) \subset D_{ind}$. Let $\mathsf{Adv} : (\{0,1\}^n)^{\ell-1} \to \{0,1\}^n$ be defined as follows:

$$\mathsf{Adv}_{end}(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell) = \begin{cases} y & \exists y : (x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A'_{end} \\ 0^n & \text{otherwise} \end{cases}$$

Let $\mathbf{X}^{ind} = (\mathbf{X}^{ind}_1, \ldots, \mathbf{X}^{ind}_{p-1}, \mathbf{X}^{ind}_{p+1}, \ldots, \mathbf{X}^{ind}_\ell)$. Now, define

$$\mathbf{X} = (\mathbf{X}^{ind}_1, \ldots, \mathbf{X}^{ind}_{p-1}, \mathsf{Adv}_{end}(\mathbf{X}^{ind}_1, \ldots, \mathbf{X}^{ind}_{p-1}, \mathbf{X}^{ind}_{p+1}, \ldots, \mathbf{X}^{ind}_\ell), \mathbf{X}^{ind}_{p+1}, \ldots, \mathbf{X}^{ind}_\ell)$$

Similarly, define

$$A' = \{(x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \mid (x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A'_{end} \wedge (x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell) \in A'_{ind}\}$$

Let $D = D_{ind}$. By construction, $\mathbf{X}$ is a uniform (g, ℓ)-NOSF source where $A' \subset A \cap \sup(\mathbf{X})$ and $f(A') \in D$. Moreover,

$$\begin{aligned} |A'| &\geq c_0^{prev} \cdot \left(\frac{1}{c(\ell-1)}\right)^{\ell-g-1} \cdot N^g \\ &\geq c_0 \cdot \frac{1}{4\ell} \cdot \frac{2}{3} \cdot \left(\frac{1}{c\ell}\right)^{\ell-g-1} \cdot N^g \\ &\geq c_0 \cdot \left(\frac{1}{c\ell}\right)^{\ell-g} \cdot N^g \end{aligned}$$

Also,

$$\begin{aligned}
|D| &\leq (c(\ell-1))^{(\ell-1)^2} \cdot \left(\frac{1}{c_0^{end}}\right)^g \cdot (|D_{end}|)^{g/(\ell-1)} \\
&\leq (c(\ell-1))^{(\ell-1)^2} \cdot \left(\frac{1}{c_0 c_1 c_3}\right)^g \cdot \left(\frac{c_3}{1-c_3} \cdot \frac{1}{c_2} \cdot M^{(\ell-1)/\ell}\right)^{g/(\ell-1)} \\
&\leq (c\ell)^{(\ell-1)^2} \cdot \left(\frac{1}{c_0} \cdot (4\ell) \cdot \frac{3}{2}\right)^g \cdot \left(2 \cdot \frac{\alpha \cdot \ell^2}{0.4} \cdot M^{(\ell-1)/\ell}\right)^{g/(\ell-1)} \\
&\leq (c\ell)^{(\ell-1)^2} \cdot \left(\frac{1}{c_0}\right)^g \cdot (6\ell)^\ell \cdot (5\alpha \cdot \ell^2) \cdot M^{g/\ell} \\
&\leq (c\ell)^{\ell^2} \cdot \left(\frac{1}{c_0}\right)^g \cdot M^{g/\ell}
\end{aligned}$$

Hence, the inductive step is proven for this case.

Case 2. The above case does not happen, i.e., for all $p \in [\ell]$, $|S_p| < c_0 c_1 N^{\ell-1}$.
Let $S \subset [N]^\ell$ be defined as follows: $(x_1, \ldots, x_\ell) \in S$ if and only if there exists $p \in [\ell]$ such that $x_1, \ldots, x_{p-1}, \ldots, x_{p+1}, \ldots, x_\ell \in S_p$. Then,

$$|S| \leq \sum_{p=1}^{\ell} |S_p| \cdot N \leq c_0 c_1 \ell \cdot N^\ell$$

Consider the ℓ-uniform ℓ-partite hypergraph $H = (V_1, \ldots, V_\ell, E)$ where $e = (v_1, \ldots, v_\ell) \in E$ if and only if $e \in A \setminus S$. As $|S| \leq c_0 c_1 \ell \cdot N^\ell$, it must be that $|E| \geq c_0(1 - c_1\ell) \cdot N^\ell$. This implies there exists $(v_1^*, \ldots, v_{\ell-g}^*) \in (V_1, \ldots, V_{\ell-g})$ such that $\deg(v_1^*, \ldots, v_{\ell-g}^*) \geq c_0(1 - c_1\ell) \cdot N^g$. Consider the g-uniform g-partite hypergraph $H' = (V_{\ell-g+1}, \ldots, V_\ell, E')$ where $e = (v_{\ell-g+1}, \ldots, v_\ell) \in E'$ if and only if $(v_1^*, \ldots v_{\ell-g}^*, v_{\ell-g+1}, \ldots, v_\ell) \in E$. Then, $|E'| \geq c_0(1 - c_1\ell) \cdot N^g$ Now, color the edges of H' with colors from $[M]$ such that $\chi(v_{\ell-g+1}, \ldots, v_\ell) = f(v_1^*, \ldots, v_{\ell-g}^*, v_{\ell-g+1}^*, \ldots, v_\ell)$. By assumption, for every position $p \in [\ell - g + 1, \ell]$, and every $(\ell-1)$ tuple $(v_{\ell-g+1}, \ldots, v_{p_1}, v_{p+1}, \ldots, v_\ell) \in [N]^{\ell-1}$: the number of distinct colored edges as entries in position p vary in H' is $\leq c_2 M^{1/\ell}$. Formally, $|\chi_{H'}(v_{\ell-g+1}, \ldots, v_{p-1}, y, v_{p+1}, \ldots, v_\ell) : y \in [N]| \leq c_2 M^{1/\ell}$. Applying Lemma 5.17 to H, we infer that there exists $D \subset [M]$ such that $|D| \leq \frac{c_4 c_2 (c_2+1)^{g(g+1)/2-1}}{(c_0(1-c_1\ell)-c_4)^g} M^{g/\ell}$ and $c_4 N^g$ edges in H are colored in one of the colors from D.

Let $A' = \{(v_1^*, \ldots, v_{\ell-g}^*, x_{\ell-g+1}, \ldots, x_\ell) : \chi_{H'}(x_{\ell-g+1}, \ldots, x_\ell) \in D\}$. Then, $A' \subset A$ and

$$\left|A'\right| \geq c_4 N^g \geq \frac{1}{c} \cdot c_0 \cdot N^g \geq c_0 \cdot \left(\frac{1}{c\ell}\right)^{\ell-g} \cdot N^g$$

Moreover, $f(A') \subset D$ and

$$|D| \leq \frac{c_4 c_2 (c_2+1)^{g(g+1)/2-1}}{(c_0(1-c_1\ell)-c_4)^g} M^{g/\ell} \leq (c\ell)^{\ell^2} \cdot \left(\frac{1}{c_0}\right)^g \cdot M^{g/\ell}$$

We now create uniform $(\ell-1,\ell)$-NOSF source $\mathbf{X} = (\mathbf{X}_1,\ldots,\mathbf{X}_\ell)$ which will have the desired properties. Let $\mathsf{Adv} : \{0,1\}^{(\ell-1)n} \to \{0,1\}^n$ be defined as follows:

$$\mathsf{Adv}(x_{\ell-g+1},\ldots,x_\ell) = \begin{cases} v_1^*,\ldots,v_{\ell-g}^* & \text{if } (x_{\ell-g+1},\ldots,x_\ell) \in E' \\ 0^n & \text{otherwise} \end{cases}$$

Let $\mathbf{X}_{\ell-g+1} = \cdots = \mathbf{X}_\ell = \mathbf{U}_n$ and let $\mathbf{X}_1,\ldots,\mathbf{X}_{\ell-g} = \mathsf{Adv}(\mathbf{X}_{\ell-g+1},\ldots,\mathbf{X}_\ell)$. Then, $A' \subset \mathrm{Supp}(\mathbf{X})$ as desired, completing the inductive step for this case as well.

□

5.2.2 Proving the main lemma for $g = \ell - 1$

We will prove our main lemma for the case of $g = \ell - 1$ using a color covering lemma for dense t-partite t-uniform hypergraphs colored in some special way:

Lemma 5.17 (Small Color Covering for Hypergraphs)**.** *Let $0 < c_0 \le 1, 0 < c_1, 0 < \varepsilon < c_0$ be arbitrary. Let $H = (V_1,\ldots,V_t,E)$ be a t-uniform t-partite hypergraph with $V_1 = \cdots = V_t = [N], |E| = c_0 N^t$. Let the edges of H be colored in one of M colors so that for every position $p \in [T]$, and every $(t-1)$ tuples: $(v_1,\ldots,v_{p_1},v_{p+1},\ldots,v_t) \in [N]^{t-1}$, the number of distinct colored edges as entries position p vary is $\le c_1 M^\delta$. Formally, $|\chi(v_1,\ldots,v_{p-1},y,v_{p+1},\ldots,v_t) : y \in [N]| \le c_1 M^\delta$. Then, there exists $D \subseteq [M]$ such that $|D| \le \frac{\varepsilon c_1 (c_1+1)^{t(t+1)/2-1}}{(c_0-\varepsilon)^t} \cdot M^{t\delta}$ and at least εN^t edges in H are colored in one of the colors from D.*

We prove this color covering lemma later. Using it, we prove our main lemma for the case of $g = \ell - 1$:

Proof of Lemma 5.15. Let $N = 2^n$. We will often identify $\{0,1\}^n$ with $[N]$ wherever convenient. Let $c_1 = \frac{1}{4\ell}, c_2 = \left(\frac{3}{2}\right)^{1/\ell^2} - 1, c_3 = \frac{2}{3}, c_4 = \frac{c_0}{4}$. We let c be a very large universal constant. By binomial approximation, there exists a constant $\alpha \ge 1$ such that $\frac{1}{\alpha} \cdot \frac{0.4}{\ell^2} \le c_2 \le \alpha \cdot \frac{0.6}{\ell^2}$ for all ℓ. For all positions $p \in [\ell]$, let $S_p \subset [N]^{\ell-1}$ be defined as follows: $(x_1,\ldots,x_{p-1},x_{p+1},\ldots,x_\ell) \in S_p$ if and only if

$$|\{f(x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell) : y \in \{0,1\}^n \wedge (x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell) \in A\}| \ge c_2 M^{1/\ell}$$

We consider various cases:

Case 1. There exists $p \in [\ell]$ such that $|S_p| \ge c_0 c_1 N^{\ell-1}$.
Consider the bipartite graph $G = (U,V,E)$ where $U = S_p, V = [M]$ and edge $e = (u,v) = ((x_1,\ldots,x_{p-1},x_{p+1},\ldots,x_\ell),z) \in E$ if and only if there exists $y \in \{0,1\}^n$ such that $f(x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell) = z$ and $(x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell) \in A$. Then by assumption, for all $u \in U$, it holds that $\deg(u) \ge c_2 M^{1/\ell}$. We apply Lemma 5.6 to G and infer that there exists $D \subset V$ such that $|D| \le \frac{c_3}{c_2(1-c_3)} M^{(\ell-1)/\ell}$ and $\mathrm{Nbr}_G(D) \ge c_0 c_1 c_3 |U| \ge c_0 c_1 c_3 N^{\ell-1}$.

We now construct set $A' \subset A$. For each vertex $u = (x_1,\ldots,x_{p-1},x_{p+1},\ldots,x_\ell) \in \mathrm{Nbr}_G(D)$, we add $(x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell)$ to A' where y is such that $(x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell) \in A$ and $f(x_1,\ldots,x_{p-1},y,x_{p+1},\ldots,x_\ell) \in D$ (we only pick one such y per u and if multiple such y exist, we break ties arbitrarily). By construction, we have that $A' \subset A$ and

$$|A'| \ge c_0 c_1 c_3 N^{\ell-1} \ge \frac{1}{c} \cdot \frac{c_0}{\ell} \cdot N^{\ell-1}$$

Moreover, $f(\mathbf{A}') \in D$ and

$$|D| \leq \frac{c_3}{c_2(1-c_3)} M^{(\ell-1)/\ell} \leq c \cdot \frac{1}{\ell^2} \cdot \left(\frac{2}{c_0}\right)^{\ell-2} \cdot M^{(\ell-1)/\ell}$$

We now construct uniform $(\ell-1,\ell)$-NOSF source $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_\ell)$ with the desired properties. Let $\mathsf{Adv} : (\{0,1\}^n)^{(\ell-1)} \to \{0,1\}^n$ be defined as follows:

$$\mathsf{Adv}(x_1, \ldots, x_{p-1}, x_{p+1}, \ldots, x_\ell) = \begin{cases} y & \exists y : (x_1, \ldots, x_{p-1}, y, x_{p+1}, \ldots, x_\ell) \in A' \\ 0^n & \text{otherwise} \end{cases}$$

Let $\mathbf{X}_1 = \cdots = \mathbf{X}_{p-1} = \mathbf{X}_{p+1} = \cdots = \mathbf{X}_\ell = \mathbf{U}_n$. Let $\mathbf{X}_p = \mathsf{Adv}(\mathbf{X}_1, \ldots, \mathbf{X}_{p-1}, \ldots, \mathbf{X}_{p+1}, \ldots, \mathbf{X}_\ell)$. Then, $\mathbf{A}' \subset \mathrm{Supp}(\mathbf{X})$ as desired.

Case 2. The above case does not happen, i.e., for all $p \in [\ell]$, $|S_p| < c_0 c_1 N^{\ell-1}$.
Let $S \subset [N]^\ell$ be defined as follows: $(x_1, \ldots, x_\ell) \in S$ if and only if there exists $p \in [\ell]$ such that $x_1, \ldots, x_{p-1}, \ldots, x_{p+1}, \ldots, x_\ell \in S_p$. Then,

$$|S| \leq \sum_{p=1}^{\ell} |S_p| \cdot N \leq c_0 c_1 \ell \cdot N^\ell$$

Consider the ℓ-uniform ℓ-partite hypergraph $H = (V_1, \ldots, V_\ell, E)$ where $e = (v_1, \ldots, v_\ell) \in E$ if and only if $e \in A \setminus S$. As $|S| \leq c_0 c_1 \ell \cdot N^\ell$, it must be that $|E| \geq c_0(1 - c_1\ell) \cdot N^\ell$. This implies there exists $v_1^* \in V_1$ such that $\deg(v_1^*) \geq c_0(1 - c_1\ell) \cdot N^{\ell-1}$. Consider the $(\ell-1)$-uniform $(\ell-1)$-partite hypergraph $H' = (V_2, \ldots, V_\ell, E')$ where $e = (v_2, \ldots, v_\ell) \in E'$ if and only if $(v_1^*, \ldots, v_\ell) \in E$. Then, $|E'| = \deg(v_1^*) \geq c_0(1 - c_1\ell) \cdot N^{\ell-1}$ Now, color the edges of H' with colors from $[M]$ such that $\chi(v_2, \ldots, v_\ell) = f(v_1^*, v_2, \ldots, v_\ell)$. By assumption, for every position $p \in [2, \ell]$, and every $(\ell-1)$ tuple $(v_2, \ldots, v_{p_1}, v_{p+1}, \ldots, v_\ell) \in [N]^{\ell-1}$: the number of distinct colored edges as entries in position p vary in H' is $\leq c_2 M^{1/\ell}$. Formally, $|\chi_{H'}(v_2, \ldots, v_{p-1}, y, v_{p+1}, \ldots, v_\ell) : y \in [N]| \leq c_2 M^{1/\ell}$. Applying Lemma 5.17 to H, we infer that there exists $D \subset [M]$ such that $|D| \leq \frac{c_4 c_2 (c_2+1)^{(\ell-1)\ell/2-1}}{(c_0(1-c_1\ell)-c_4)^{\ell-1}} M^{(\ell-1)/\ell}$ and $c_4 N^{\ell-1}$ edges in H are colored in one of the colors from D.
Let $A' = \{(v_1^*, x_2, \ldots, x_\ell) : \chi_{H'}(x_2, \ldots, x_\ell) \in D\}$. Then, $A' \subset A$ and

$$\left|A'\right| \geq c_4 N^{\ell-1} \geq \frac{1}{c} \cdot c_0 \cdot N^{\ell-1} \geq \frac{1}{c} \cdot \frac{c_0}{\ell} \cdot N^{\ell-1}$$

Moreover, $f(\mathbf{A}') \in D$ and

$$|D| \leq \frac{c_4 c_2 (c_2+1)^{(\ell-1)\ell/2-1}}{(c_0(1-c_1\ell)-c_4)^{\ell-1}} M^{(\ell-1)/\ell} \leq c \cdot \frac{1}{\ell^2} \cdot \left(\frac{2}{c_0}\right)^{\ell-2} \cdot M^{(\ell-1)/\ell}$$

We now create uniform $(\ell-1,\ell)$-NOSF source $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_\ell)$ which will have the desired properties. Let $\mathsf{Adv} : \{0,1\}^{(\ell-1)n} \to \{0,1\}^n$ be defined as follows:

$$\mathsf{Adv}(x_2, \ldots, x_\ell) = \begin{cases} v_1^* & \text{if } (x_2, \ldots, x_\ell) \in E' \\ 0^n & \text{otherwise} \end{cases}$$

Let $\mathbf{X}_2 = \cdots = \mathbf{X}_\ell = \mathbf{U}_n$ and let $\mathbf{X}_1 = \mathsf{Adv}(\mathbf{X}_2, \ldots, \mathbf{X}_\ell)$. Then, $A' \subset \mathrm{Supp}(\mathbf{X})$ as desired.

□

5.2.3 Finding a small color covering in locally-light hypergraphs

We consider dense t-uniform t-partite hypergraphs where all edges are colored and the hypergraph satisfies a "locally-light" condition: all $t-1$-tuples are adjacent to a small number of colors. The covering lemma finds small set of colors that covers constant fraction of edges in the hypergraph. We do this by finding a popular color in such a hypergraph.

Lemma 5.18 (Popular Color in Locally-Light Hypergraphs)**.** *Let $0 < c_0 \leq 1, 0 < c_1$ be arbitrary. Let $t \geq 2 \in \mathbb{N}$. Let $H = (V_1, \ldots, V_t, E)$ be a t-uniform t-partite hypergraph with $|V_1| = \cdots = |V_t| = N, |E| = c_0 N^t$. Let the edges of H be colored in one of M colors so that for every position $p \in [T]$, and every $(t-1)$ tuples: $(v_1, \ldots, v_{p_1}, v_{p+1}, \ldots, v_t) \in [N]^{t-1}$, the number of distinct colored edges as entries position p vary is $\leq c_1 M^\delta$. Formally, $|\chi(v_1, \ldots, v_{p-1}, y, v_{p+1}, \ldots, v_t) : y \in [N]| \leq c_1 M^\delta$. Then, there exists a color $\gamma \in [M]$ such that at least $\frac{c_0^t}{c_1(c_1+1)^{t(t+1)/2-1}} \cdot N^t/M^{t\delta}$ edges in H are colored with color γ.*

Using this lemma, our color covering lemma for hypergraph follows by repeatedly finding such popular colors.

Proof of Lemma 5.17. We introduce some additional notation: for a color $\gamma \in [M]$, let

$$\mathsf{count}_H(\gamma) = |\{e \in H : \chi(e) = \gamma\}|.$$

We will construct D by a greedy algorithm where we add the most popular color in H to D, remove all edges of that color, and repeat. Further details are specified in Algorithm 2.

Algorithm 2:

$i \leftarrow 0$, $D \leftarrow \emptyset$
$H^{(0)} = (V_1^{(0)}, \ldots, V_t^{(0)}, E^{(0)}) \leftarrow H = (V_1, \ldots, V_t, E)$
while $\mathsf{count}_H(D) \leq \varepsilon N^t$ **do**
  Let $\gamma_i \in [M]$ be the color that maximizes $\mathsf{count}_{H_i}(\gamma_i)$.
  $D \leftarrow D \cup \{\gamma_i\}$
  $E^{(i+1)} \leftarrow E^{(i)} \setminus \{e \in E \mid \chi(e) = \gamma_i\}$
  $H^{(i+1)} \leftarrow (V_1, \ldots, V_t, E^{(i+1)})$
  $i \leftarrow i + 1$
end

We observe that $\mathsf{count}_H(\gamma_1) \geq \cdots \geq \mathsf{count}_H(\gamma_{|D|})$. At the iteration number $|D|$ of the loop, the number of uncovered edges in H is at least $(c_0 - \varepsilon)N^2$. Applying Lemma 5.18 on $H^{(|D|-1)}$, we infer that

$$\mathsf{count}_H(\gamma_{|D|}) \geq \frac{(c_0 - \varepsilon)^t}{c_1(c_1+1)^{t(t+1)/2-1}} \cdot N^t/M^{t\delta}$$

Hence, the number of edges covered in each iteration of the loop is at least $\frac{(c_0-\varepsilon)^t}{c_1(c_1+1)^{t(t+1)/2-1}} \cdot N^t/M^{t\delta}$ As the loop stops when the number of edges covered is at least εN^t, the number of iterations to terminate is at most

$$\frac{\varepsilon N^t}{\frac{(c_0-\varepsilon)^t}{c_1(c_1+1)^{t(t+1)/2-1}} \cdot N^t/M^{t\delta}} = \frac{\varepsilon c_1(c_1+1)^{t(t+1)/2-1}}{(c_0-\varepsilon)^t} \cdot M^{t\delta}$$

As the number of iterations of the loop equals $|D|$, we indeed infer the claim. □

5.2.4 Finding a popular color in locally-light hypergraphs

For the base case, we find such a popular color in graphs:

Lemma 5.19 (Popular Color in Locally-Light Graphs)**.** *Let $0 < c_0 \le 1, 0 < c_1$ be arbitrary. Let $H = (U, V, E)$ be a bipartite graph with $|U| = |V| = N, |E| = c_0 N^2$. Let the edges of H be colored in one of M colors so that for every vertex $x \in H$, the number of distinct colored edges incident on x is $\le c_1 M^\delta$. Then, there exists a color $\gamma \in [M]$ such that at least $\frac{c_0^2}{(c_1+1)^2 c_1} \cdot N^2/M^{2\delta}$ edges in H are colored with color γ.*

Using this, we inductively find a popular color in locally-light hypergraphs.

Proof of Lemma 5.18. We prove this result by induction on t with the inductive hypothesis for t being that such a popular color exists for graphs with this property.

For the base case, $t = 2$. We apply Lemma 5.19 directly on H and infer the claim.

For the inductive step, assume that we have proven the hypothesis for $t-1$ and using it, we prove the hypothesis for t where $t \ge 3$. Let $c_2 = \frac{c_0}{c_1+1}$. Let

$$E' = \{e = (v_1, \dots, v_t) \in E : |\mathrm{Nbr}_H((v_2, \dots, v_t), \chi(e))| \ge c_2 N/M^\delta\}$$

Let $H' = (V_1, \dots, V_t, E')$. We now lower bound the number of edges in E'. Fix arbitrary $(t-1)$-tuple $v = (v_2, \dots, v_t) \in [N]^{t-1}$. By assumption, $|\{\gamma \in [M] : |\mathrm{Nbr}_H(v, \gamma)| > 0\}| \le c_1 M^\delta$. In H', we excluded all edges in H incident to the tuple v with color γ such that $|\mathrm{Nbr}_H(v, \gamma)| < c_2 N/M^\delta$. Hence, we exclude at most $c_1 c_2 N$ such edges incident to v in H. As there are at most N^{t-1} such tuples, the total number of edges we excluded in E' is at most $c_1 c_2 N^t$. Hence, $|E'| \ge (c_0 - c_1 c_2) N^t$.

As $|V_1| \le N$, there exists $v_1^* \in V_1$ such that $\mathrm{Nbr}_{H'}(v_1^*) \ge |E'|/N \ge (c_0 - c_1 c_2) N^{t-1}$. Let $G = (V_2, \dots, V_t, E_G)$ be a $(t-1)$-uniform $(t-1)$-partite hypergraph where $(v_2, \dots, v_t) \in E_G$ if and only if $(v_1^*, v_2, \dots, v_t) \in E'$. We see that G satisfies the conditions of the inductive hypothesis for $t-1$. Hence, there exists a color $\gamma^* \in [M]$ such that at least $\frac{(c_0 - c_1 c_2)^{t-1}}{c_1 (c_1+1)^{t(t-1)/2-1}} \cdot N^{t-1}/M^{(t-1)\delta}$ edges in G are colored by γ^*. For each edge $e = (v_2, \dots, v_t)$ in G with $\chi(e) = \gamma^*$, by property of E' and the fact that edge $(v_1^*, v_2, \dots, v_t) \in E'$, $|\mathrm{Nbr}_{H'}((v_2, \dots, v_t), \gamma^*)| \ge c_2 N/M^\delta$. Thus, the number of edges in H', and hence H colored with γ^* is at least

$$\frac{(c_0 - c_1 c_2)^{t-1}}{c_1 (c_1+1)^{t(t-1)/2-1}} \cdot N^{t-1}/M^{(t-1)\delta} \cdot c_2 N/M^\delta \ge \frac{c_0^t}{c_1 (c_1+1)^{t(t+1)/2-1}} \cdot N^T/M^{t\delta}$$

Hence, the inductive hypothesis holds for t, completing the inductive step, proving the claim. □

Finally, we directly argue a popular color exists in dense locally-light bipartite graphs.

Proof of Lemma 5.19. Let $c_2 = \frac{c_0}{c_1+1}$. Let

$$E' = \{e = (u, v) \in E : |\mathrm{Nbr}_H(v, \chi(e))| \ge c_2 N/M^\delta\}$$

Let $H' = (U, V, E')$. We now lower bound the number of edges in E'. Fix arbitrary vertex $v \in V$. By assumption, $|\{\gamma \in [M] : |\mathrm{Nbr}_H(v, \gamma)| > 0\}| \le c_1 M^\delta$. In H', we excluded all edges in H incident to v with color γ such that $|\mathrm{Nbr}_H(v, \gamma)| < c_2 N/M^\delta$. Hence, we exclude at most $c_1 c_2 N$ such edges incident to v in H. As $|V| \le N$, the total number of edges we excluded in E' is at most $c_1 c_2 N^2$. Hence, $|E'| \ge (c_0 - c_1 c_2) N^2$.

As $|U| \le N$, there exists $u^* \in U$ such that $\mathrm{Nbr}_{H'}(u^*) \ge |E'|/N \ge (c_0 - c_1c_2)N$. As the number of distinct colored edges incident on u^* is at most c_1M^δ, there exists a color $\gamma^* \in [M]$ such that $|\mathrm{Nbr}_{H'}(u^*, \gamma^*)| \ge \frac{(c_0-c_1c_2)N}{c_1M^\delta}$. For each $v \in \mathrm{Nbr}_{H'}(u^*, \gamma^*)$, by definition of E', $|\mathrm{Nbr}_{H'}(v, \gamma^*)| \ge \frac{c_2N}{M^\delta}$. Hence, the number of edges $e \in H'$ colored with γ^* is at least

$$\sum_{v \in \mathrm{Nbr}_{H'}(u^*,\gamma^*)} |\mathrm{Nbr}_{H'}(v, \gamma^*)| \ge \left|\mathrm{Nbr}_{H'(u^*,\gamma^*)}\right| \cdot \frac{c_2N}{M^\delta} \ge \frac{(c_0 - c_1c_2)c_2N^2}{c_1M^{2\delta}} = \frac{c_0^2N^2}{(c_1+1)^2c_1M^{2\delta}}$$

□

5.3 Impossibility of condensing from CG sources

We prove two impossibility results regarding impossibility of condensing from (ℓ, ℓ)-aCG sources. Our first result Theorem 5.21 states that any candidate condenser cannot decrease the entropy gap present in the blocks of CG sources. Our second result in contrast, states that when blocks have linear entropy, then condenser cannot condense beyond rate 1/2. The latter result is much stronger than the former in regimes where m is comparatively larger than n (say $m = O(n\ell)$ and $\ell = \omega(1)$).

5.3.1 Impossibility of non-trivial condensing beyond min-entropy gap

We will use the fact that it is impossible to condense from general (n, k)-sources.

Lemma 5.20. *For all $n, k, m \in \mathbb{N}$ and $\varepsilon > 0$ the following holds: For all functions $f : \{0,1\}^n \to \{0,1\}^m$, there exists an (n, k) source $\mathbf{X}$ such that $H^\varepsilon_\infty(f(\mathbf{X})) \le m - (n - k) + \log(1/(1-\varepsilon)) - \max(m - n, 0)$.*

We believe a result of this form is well-known but we were unable to find a good reference. Thus, for the sake of completeness, we prove this lemma at the end of this subsection. Using this, we prove our impossibility result for (ℓ, ℓ)-aCG sources.

Theorem 5.21. *For all $0 < \varepsilon < 1, \Delta$ and $\ell, m, n \in \mathbb{N}$, the following holds: for every function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a (ℓ, ℓ)-aCG source $\mathbf{X}$ where the good blocks have min-entropy at least $n - \Delta - \log(\ell/\varepsilon) - O(1)$ conditioned on all fixings of previous blocks and $H^\varepsilon_\infty(f(\mathbf{X})) \le m - \Delta + \log(2/(2-\varepsilon)) - \max(m - \ell n, 0)$.*

Proof. Let $\mathbf{X}$ be an arbitrary (t, k)-source where $t = \ell n$ and $k = n - \Delta$. We transform $\mathbf{X}$ into a source $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell)$ with block lengths n that is $\varepsilon/2$-close to an (ℓ, ℓ)-aCG source where the good blocks have min-entropy at least $n - \Delta - \log(2(\ell-1)/\varepsilon)$ conditioned on all fixings of the previous blocks. We then apply Lemma 5.20 to infer that it is impossible to condense (ℓ, ℓ)-aCG sources so that the output distribution is ε-close to having min-entropy more than $m - \Delta + \log(2/(2-\varepsilon) - \max(m - \ell n, 0)$ as desired.

Let $\gamma = \frac{\varepsilon}{2(\ell-1)}$. For $1 \le i \le \ell$, we define each block $\mathbf{Y}_i$ as $\mathbf{Y}_i = \mathbf{X}[(i-1) \cdot n, i \cdot n]$. We prove that $\mathbf{Y}$ is $\varepsilon/2$-close to a block source with ℓ blocks of length n each and each of them has entropy at least $n - \Delta - \log(1/\gamma)$ conditioned on all previous blocks. We will inductively prove that for all j down from $\ell + 1$ to 1, there exist blocks $\mathbf{Z}_j, \ldots, \mathbf{Z}_\ell$ such that:

1.
$$(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) \approx_{(\ell-j+1)\gamma} (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}_j, \ldots, \mathbf{Z}_\ell)$$

2. For all $w \in \mathbb{N}$ such that $j \leq w \leq \ell$, conditioned on every fixing of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}_j, \ldots, \mathbf{Z}_{w-1}$: the min-entropy of $\mathbf{Z}_w$ is at least $n - \Delta - \log(1/\gamma)$.

3. $$H_\infty((\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1})) \geq n(j-1) - \Delta.$$

The base case of $j = \ell + 1$ is trivially true.

For the inductive step, assume the claim holds for $j+1 \leq \ell+1$. We prove the claim for j. We begin by proving the third requirement is satisfied. By assumption, we know that $H_\infty((\mathbf{Y}_1, \ldots, \mathbf{Y}_j)) \geq jn - \Delta$. Looking ahead, we apply Lemma 6.19 to $(\mathbf{Y}_1, \ldots, \mathbf{Y}_j)$ and to it's projection onto first $n(j-1)$ bits ; we infer that $H_\infty((\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1})) \geq n(j-1) - \Delta$ as desired.

Applying Lemma 4.4, we get that with probability at least $1 - \gamma$ over fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$, the source $(\mathbf{Y}_1, \ldots, \mathbf{Y}_j)$ will have conditional min-entropy at least $n - \Delta - \log(1/\gamma)$. So, with probability at least $1 - \gamma$ over fixings of $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1})$, $\mathbf{Y}_j$ will have conditional min-entropy at least $n - \Delta - \gamma$. Let these good fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$ be S. By the inductive hypothesis, there exist blocks $\mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell$ that satisfy the conditions of the inductive hypothesis. Define the distribution $(\mathbf{Z}'_j, \ldots, \mathbf{Z}'_\ell)$ as being same as the conditional distribution of $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell)$ when $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}) \in S$ and equal to $(\mathbf{U}_n)^{\ell-j+1}$ otherwise. Then,

$$(\mathbf{Y}_1, \ldots, \mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell) \approx_\gamma (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}'_j, \ldots, \mathbf{Z}_\ell)$$

Hence,

$$\begin{aligned}
\left|(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) - (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}'_j, \ldots, \mathbf{Z}'_\ell)\right| &\leq |(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) - (\mathbf{Y}_1, \ldots, \mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell)| \\
&\quad + \left|(\mathbf{Y}_1, \ldots, \mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell) - (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}'_j, \ldots, \mathbf{Z}'_\ell)\right| \\
&\leq (\ell - j)\gamma + \gamma \\
&\leq (\ell - j + 1)\gamma
\end{aligned}$$

We now prove that the second condition of the inductive hypothesis for j is satisfied.

When $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}) \notin S$, for all $w \in \mathbb{N}$ with $j \leq w \leq \ell$, $\mathbf{Z}'_w$ will be independent and uniform and so will have entropy at least $n - \Delta - \log(1/\gamma)$ conditioned on all fixings of blocks before it.

When $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}) \in S$, then the conditional distribution $(\mathbf{Z}'_j, \ldots, \mathbf{Z}'_w)$ equals the conditional distribution $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell)$. By the definition of S, $\mathbf{Z}'_j$ will have min-entropy at least $n - \Delta - \log(1/\gamma)$. Moreover, for all $w \in \mathbb{N}$ with $j+1 \leq w \leq \ell$, by the inductive hypothesis, on every fixing of $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \mathbf{Z}_{w-1})$, we have that $\mathbf{Z}_w$ will have min-entropy at least $n - \Delta - \log(1/\gamma)$. Thus, on every fixing of $(\mathbf{Z}'_j, \ldots, \mathbf{Z}'_{w-1})$, we have that $\mathbf{Z}'_w$ will have min-entropy at least $n - \Delta - \log(1/\gamma)$.

Hence, all 3 conditions are satisfied and the inductive step is proven. □

Lastly, we provide the proof that no non-trivial condensers exist for arbitrary (n, k)-sources.

Proof of Lemma 5.20. Let $N = 2^n, M = 2^m, K = 2^k$. We identify $\{0,1\}^n, \{0,1\}^m$ with $[N], [M]$ respectively. We prove that for $m \leq n$, $H^\varepsilon_\infty(f(\mathbf{X})) \leq k + m - n + \log(1/(1-\varepsilon))$. If $m > n$ then we observe that $|\mathrm{Supp}(f)| \leq N$. So, we relabel f so that it's co-domain is $\{0,1\}^n$, apply the mentioned claim, and infer that $H^\varepsilon_\infty(f(\mathbf{X})) \leq k + \log(1/(1-\varepsilon))$. Hence, it suffices to prove the said claim for $m \leq n$.

For $z \in [M]$, let $\chi(z) = \{x \in [N] : f(x) = z\}$ and $w(z) = |\chi(z)|$. Without loss of generality, let $w(1) \geq w(2) \geq \cdots \geq w(M)$. For $i \in [M]$, let $S_i = \sum_{j=1}^{i} w(j)$. Let $i^* \in [M]$ be the smallest integer such that $S_{i^*} \geq K$. Let $r = K - S_{i^*-1}$. Let $A = \cup_{j=1}^{i^*-1} \chi(j)$. Moreover, add arbitrary r elements from $\chi(i^*)$ into A. Let $\mathbf{X}$ be the (n, k) source that is uniform over the set A. Let $k_{out} = k + m - n + \log(1/(1-\varepsilon))$. Then, we claim that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq k_{out}$. As $w(1) \geq \cdots \geq w(M)$, it must be that for all $1 \leq j \leq M : \frac{S_j}{j} \geq \frac{S_M}{M} = \frac{N}{M}$. Hence, $S_j \geq \frac{jN}{M}$. In particular, if $j \geq \frac{KM}{N}$, then $S_j \geq K$. As i^* is the smallest integer such that $S_{i^*} \geq K$, it must be that $i^* \leq \frac{KM}{N}$. Hence, with probability 1, $f(\mathbf{X}) \in [KM/N]$. Applying Claim 4.3, we infer that $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq k+m-n+\log(1/\varepsilon)$ as desired. □

5.3.2 Impossibility of condensing beyond rate 1/2

Using condensing impossibility result for uniform $(1, 2)$-oNOSF sources, we prove a condensing impossibility result for (ℓ, ℓ)-aCG source (which are just CG sources, with no adversarial blocks) where the good blocks have min-entropy at least $O(n/\ell)$ conditioned on every fixing of previous blocks.

Theorem 5.22. *For all $0 < \varepsilon < 1$ there exists a $\delta > 0$ such that the following holds: for every function $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a (ℓ, ℓ)-aCG source $\mathbf{X}$ where the good blocks have min-entropy at least $\frac{n-\ell\log(2\ell/\varepsilon)}{\ell+1}$ conditioned on all fixings of previous blocks and $H^{\varepsilon}_{\infty}(f(\mathbf{X})) \leq \frac{1}{2} \cdot m + \delta$.*

Proof. Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ be an arbitrary uniform $(1, 2)$-oNOSF source where the length of the blocks is at least $\frac{\ell(\ell+1)}{2} \cdot \left(\frac{n-\ell\log(2\ell/\varepsilon)}{\ell+1} + \log(2\ell/\varepsilon)\right)$. We transform $\mathbf{X}$ into a source $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell)$ with block lengths n that is $\varepsilon/2$-close to an (ℓ, ℓ)-aCG source. We will then apply Lemma 5.3 with $\ell = 2$ and error $\varepsilon/2$ to infer that it is impossible to condense such (ℓ, ℓ)-aCG sources so that the output distribution is ε-close to having min-entropy more than $\frac{1}{2} \cdot m + \delta$.

Let $\gamma = \frac{\varepsilon}{2\ell}, t_2 = \frac{n-\ell\log(1/\gamma)}{\ell+1}, t_1 = t_2 + \log(1/\gamma)$. Define each block $\mathbf{Y}_i$ as follows:

$$\mathbf{Y}_i = \mathbf{X}_2[(i-1) \cdot t_2 + 1, \ldots, (i) \cdot t_2] \circ \mathbf{X}_1[(i)(i-1)/2 \cdot t_1 + 1, \ldots, (i)(i+1)/2 \cdot t_1] \circ 0^{n-(t_2+i\cdot t_1)}$$

We claim that $\mathbf{Y}$ is $\varepsilon/2$-close to a CG source where each good block has min-entropy at least t_2 for every fixing of the previous blocks. We consider cases on the location of the good block in $\mathbf{X}$.

Case 1. Block $\mathbf{X}_2$ is good.
As $\mathbf{X}$ is a uniform oNOSF source, $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent. As $\mathbf{X}_2$ is uniform, this implies each sub-source from $\mathbf{X}_2$ is also uniform conditioned on every fixing of all other bits in $\mathbf{X}$. Hence, for all i, and $(y_1, \ldots, y_{i-1}) \in (\{0,1\}^n)^{i-1} : H_\infty(\mathbf{Y}_i \mid (\mathbf{Y}_1, \ldots, \mathbf{Y}_{i-1}) = (y_1, \ldots, y_{i-1})) \geq t_2$.

Case 2. Block $\mathbf{X}_1$ is good.
We will use the following claim that most fixings of previous blocks preserve min-entropy:

Claim 5.23. *For all $1 \leq i \leq \ell$: with probability at least $1 - \gamma$ over fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{i-1}$: $H_\infty(\mathbf{Y}_i) \geq t_2$*

We will prove this claim later.

Using it, we will inductively to prove that for all j starting from ℓ down to 1, there exist blocks $\mathbf{Z}_j, \ldots, \mathbf{Z}_\ell$ such that:

1. $(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) \approx_{(\ell-j+1)\gamma} (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}_j, \ldots, \mathbf{Z}_\ell)$.
2. For all $t \in \mathbb{N}$ such that $j \leq t \leq \ell$, conditioned on every fixing of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}_j, \ldots, \mathbf{Z}_{t-1}$: the min-entropy of $\mathbf{Z}_t$ is at least t_2.

The overall claim exactly corresponds to the inductive hypothesis for $j = 1$ and hence, it suffices to prove this.

For the base case of $j = \ell$, proceed as follows: Applying Claim 5.23, with probability at least $1 - \gamma$ over fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell-1}$, $\mathbf{Y}_\ell$ will have conditional min-entropy at least t_2. Let these good fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell-1}$ be S. Define the distribution $\mathbf{Z}_\ell$ as being same as the conditional distribution $\mathbf{Y}_\ell$ when $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell-1}) \in S$ and equals $\mathbf{U}_n$ otherwise. Then,

$$(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) \approx_\gamma (\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell-1}, \mathbf{Z}_\ell)$$

Moreover, conditioned on every fixing of $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell-1})$: $\mathbf{Z}_\ell$ will have entropy at least t_2. Hence, both conditions are satisfied and the base case is proven.

For the inductive step, assume the claim holds for $j + 1 \leq \ell$. We prove the claim for j. Applying Claim 5.23, with probability at least γ over fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$, $\mathbf{Y}_j$ will have conditional min-entropy at least t_2. Let these good fixings of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$ be S. By the inductive hypothesis, there exists blocks $\mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell$ that satisfy both conditions laid out in the inductive hypothesis. Define the distribution $(\mathbf{Z}'_j, \ldots, \mathbf{Z}'_\ell)$ as being same as the conditional distribution of $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell)$ when $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}) \in S$ and equal to $(\mathbf{U}_n)^{\ell-j+1}$ otherwise. Then,

$$(\mathbf{Y}_1, \ldots, \mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell) \approx_\gamma (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}'_j, \ldots, \mathbf{Z}_\ell)$$

Hence,

$$\begin{aligned}
\left|(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) - (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}'_j, \ldots, \mathbf{Z}'_\ell)\right| &\leq \left|(\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell) - (\mathbf{Y}_1, \ldots, \mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell)\right| \\
&\quad + \left|(\mathbf{Y}_1, \ldots, \mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell) - (\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}, \mathbf{Z}'_j, \ldots, \mathbf{Z}'_\ell)\right| \\
&\leq (\ell - j)\gamma + \gamma \\
&\leq (\ell - j + 1)\gamma
\end{aligned}$$

We now prove that the second condition of the inductive hypothesis for j is satisfied.

When $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}) \notin S$, for all $t \in \mathbb{N}$ with $j \leq t \leq \ell$, $\mathbf{Z}'_t$ will be independent and uniform and hence will have entropy at least t_2 conditioned on all fixings of blocks before it.

When $(\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}) \in S$, then the conditional distribution $(\mathbf{Z}'_j, \ldots, \mathbf{Z}'_t)$ equals the conditional distribution $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_\ell)$. By the definition of S, $\mathbf{Y}_j$ and hence $\mathbf{Z}'_j$ will have min-entropy at least t_2. Moreover, for all $t \in \mathbb{N}$ with $j + 1 \leq t \leq \ell$, by the inductive hypothesis, on every fixing of $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \mathbf{Z}_{t-1})$, we have that $\mathbf{Z}_t$ will have min-entropy at least t_2. Equivalently, on every fixing of $(\mathbf{Z}'_j, \ldots, \mathbf{Z}'_{t-1})$, we have that $\mathbf{Z}'_t$ will have min-entropy at least t_2.

Hence, both conditions are satisfied and the inductive step is proven.

□

Lastly we prove our claim that most fixings of previous blocks preserve min-entropy in the later block.

Proof of Claim 5.23. Let $s = \frac{i(i-1)}{2} \cdot t_1$. As $\mathbf{X}_1$ is uniform, $\mathbf{X}_1[s+1, \ldots, s+i \cdot t_1]$ remains uniform conditioned on every fixing α of $\mathbf{X}_1[1, \ldots, s]$. By the min-entropy chain rule (Lemma 4.4), with probability at least $1-\gamma$ over fixings β of $\mathbf{X}_2[1, \ldots, i \cdot t_2]$ and every fixing $\alpha \in \{0,1\}^s$:

$$\begin{aligned}
&H_\infty\left(\mathbf{X}_1[s+1, \ldots, s+i \cdot t_1] \mid \mathbf{X}_2[1, \ldots, (i-1) \cdot t_2] = \beta, \mathbf{X}_1[1, \ldots, s] = \alpha\right) \\
&\geq i \cdot t_1 - ((i-1) \cdot t_2 + \log(1/\gamma)) \\
&\geq t_2
\end{aligned}$$

By construction, fixing $\mathbf{X}_2[1, \ldots, (i-1) \cdot t_2]$ and $\mathbf{X}_1[1, \ldots, s]$ fixes $\mathbf{Y}_1, \ldots, \mathbf{Y}_{i-1}$. For every fixing of these blocks, $H_\infty(\mathbf{Y}_i) \geq H_\infty(\mathbf{X}_1[s+1, \ldots, s+i \cdot t_1])$ and the claim follows. □

5.4 Deferred proofs of helpful lemmas

The remaining deferred proofs of lemmas follow from the following results:

Lemma 5.24. *Let $g, \ell, n, g', \ell', n', m \in \mathbb{N}$ be such that $g \leq a \cdot g' + \max(b - (\ell' - g'), 0), (a+1)n \leq n'$ where $a, b \in \mathbb{N}$ are unique integers such that $\ell = a \cdot \ell' + b$ where $0 \leq b < \ell'$. Let $0 < \varepsilon < 1, \delta > 0$ be such that: for any function $f : (\{0,1\}^{n'})^{\ell'} \to \{0,1\}^m$, there exists a uniform (g', ℓ')-oNOSF source (uniform (g', ℓ')-NOSF source, respectively) $\mathbf{Y}$ so that $H^\varepsilon_\infty(f(\mathbf{Y})) \leq \frac{g'}{\ell'} \cdot m + \delta$. Then, for any function $h : (\{0,1\}^n)^\ell \to \{0,1\}^m$, there exists a uniform (g, ℓ)-oNOSF source (uniform (g, ℓ)-NOSF source, respectively) $\mathbf{X}$ such that $H^\varepsilon_\infty(h(\mathbf{X})) \leq \frac{g'}{\ell'} \cdot m + \delta$.*

Proof of Lemma 5.24. For the sake of contradiction, assume there exists a non-trivial condenser $h : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ)-oNOSF source (uniform (g, ℓ)-NOSF source, respectively) $\mathbf{X}$, $H^\varepsilon_\infty(h(\mathbf{X})) \geq \frac{g'}{\ell'} \cdot m + \delta$. We will use h to construct a condenser $f : (\{0,1\}^{n'})^{\ell'} \to \{0,1\}^m$ for uniform (g', ℓ')-oNOSF source (uniform (g', ℓ')-NOSF sources, respectively) to get a contradiction.

Define $f : (\{0,1\}^{n'})^{\ell'} \to \{0,1\}^m$ as: $f(\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell'}) = h(\mathbf{X}_1, \ldots, \mathbf{X}_{\ell = a\ell' + b})$ where the $\mathbf{X}_i$ are constructed by splitting up the $\mathbf{Y}_j$ as evenly as possible. Concretely, from each of the b blocks $\mathbf{Y}_1, \ldots, \mathbf{Y}_b$, construct $a+1$ blocks of length n each to form $b(a+1)$ blocks of the $\mathbf{X}_i$'s. Furthermore, from each of the $\ell' - b$ remaining blocks $\mathbf{Y}_{b+1}, \ldots, \mathbf{Y}'_\ell$ construct a blocks of length n each to form $(\ell' - b)a$ blocks of the $\mathbf{X}_i$'s. In total, we constructed $b(a+1) + (\ell' - b)a = a\ell' + b = \ell$ blocks of $\mathbf{X}$ as desired. As $\mathbf{Y}$ contains at least g' good blocks, $\mathbf{X}$ will contain $a \cdot g'$ good blocks if $g' \leq \ell' - b$ and at least $a \cdot g' + g' - (b - \ell')$ otherwise. By assumed constraints on g, we infer that $\mathbf{X}$ will contain at least g good blocks as desired. We indeed check that this construction preserves one sidedness of the bad blocks —- if $\mathbf{X}$ is a uniform oNOSF source source, then $\mathbf{Y}$ is also a uniform oNOSF source source. Hence, $\mathbf{X}$ is a uniform (g, ℓ)-oNOSF source (uniform (g, ℓ)-NOSF source, respectively). Thus, $H^\varepsilon_\infty(f(\mathbf{Y}) \geq H^\varepsilon_\infty(h(\mathbf{X})) \geq \frac{g'}{\ell'} \cdot m + \delta$, a contradiction. □

The deferred proofs of couple of lemmas follow from this result.

Proof of Lemma 5.4. We apply Lemma 5.24 with $g' = 1$ to infer the claim. □

Proof of Lemma 5.13. We apply Lemma 5.24 with $g' = \ell' - 1$ to infer the claim. □

Proof of Lemma 5.14. We apply Lemma 5.24 with $g = ag', \ell = a\ell'$ to infer the claim. □

6 Condensers for oNOSF Sources

We will prove the following main theorem regarding condensing from oNOSF sources in this section:

Theorem 6.1. *For all $g, \ell, r \in \mathbb{N}$ such that $\left\lfloor \frac{\ell-1}{g-1} \right\rfloor = r$ and $r < \frac{\ell-1}{g-1}$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any (g, ℓ, n, k)-oNOSF source $\mathbf{X}$ with $k \geq 2\log(n)$, we have that $H^\varepsilon_\infty(\mathbf{X}) \geq \frac{1}{r} \cdot m - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)$ with $m = r\left(\frac{k}{8\ell} - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)\right)$ and $\varepsilon = (g-1) \cdot 2^{-\Omega(k)} + 2^{-\log\left(\frac{(g-1)k}{8\ell}\right)/4}$.*

This result is tight up to lower order terms as it asymptotically matches the impossibility results of Theorem 5.1.

We prove this theorem in two steps. First, we show how to transform oNOSF sources to uniform oNOSF sources:

Theorem 6.2. *For any g and ℓ, there exists a function $f : (\{0,1\}^n)^\ell \to (\{0,1\}^m)^{\ell-1}$ with $m = \frac{k}{8\ell}$ such that for any (g, ℓ, k)-oNOSF source $\mathbf{X}$ with $k \geq 2\log(n)$ there exists a uniform $(g-1, \ell-1)$-oNOSF source $\mathbf{Y}$ such that $|f(\mathbf{X}) - \mathbf{Y}| \leq (g-1) \cdot 2^{-\Omega(k)}$.*

Second, we show how to condense from uniform oNOSF sources.

Theorem 6.3. *For any g and ℓ such that $\lfloor \ell/g \rfloor = r$ and $r < \ell/g$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ)-oNOSF source $\mathbf{X}$ we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq \frac{1}{r} \cdot m - 2(5^{\ell-g} - 1)\log(gn)$ where $\varepsilon = 2^{-\log(gn)/4}$ and $m = r(n - 2(5^{\ell-g} - 1)\log(gn))$.*

Using these two ingredients, our main theorem follows:

Proof of Theorem 6.1. Take the transformation function f from Theorem 6.2 and let $\mathbf{X}' = f(\mathbf{X})$ be the resulting source that is $\varepsilon_1 = (g-1) \cdot 2^{-\Omega(k)}$ close to a uniform $\left(g' = g-1, \ell' = \ell-1, n' = \frac{k}{8\ell}\right)$-oNOSF source source.

By assumption, we have $\lfloor \ell'/g' \rfloor = r$ and $r < \ell'/g'$. Consequently, we can apply Theorem 6.3 to get a condenser $\mathsf{Cond}' : (\{0,1\})^{n'})^{\ell'} \to \{0,1\}^m$ where $m = r(n' - 2(5^{\ell'-g'} - 1)\log(g'n'))$ such that $H^{\varepsilon_2}_\infty(\mathsf{Cond}(\mathbf{X}')) \geq \frac{1}{r} \cdot m - 2(5^{\ell'-g'} - 1)\log(g'n')$ with $\varepsilon_2 = 2^{-\log(g'n')/4}$. These expressions simplify to $m = r\left(\frac{k}{8\ell} - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)\right)$, $\varepsilon_2 = 2^{-\log\left(\frac{(g-1)k}{8\ell}\right)/4}$, and $H^{\varepsilon_2}_\infty(\mathsf{Cond}(\mathbf{X}')) \geq m - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)$.

Finally, we put these two steps together to define $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ as $\mathsf{Cond}(\mathbf{X}) := \mathsf{Cond}'(f(\mathbf{X}))$ so that $H^\varepsilon_\infty(\mathbf{X}) \geq m - (5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)$ with $\varepsilon = (g-1) \cdot 2^{-\Omega(k)} + 2^{-\log\left(\frac{(g-1)k}{8\ell}\right)/4}$. □

6.1 Transforming low entropy oNOSF sources to uniform oNOSF sources

We will prove Theorem 6.2 in this subsection. We will use the fact that a random function is a very good two source extractor.

Lemma 6.4. *Let $n_1, n_2, k_1, k_2, m, \varepsilon$ be such that $k_1 \le n_1, k_2 \le n_2, m = k_1 + k_2 - 2\log(1/\varepsilon) - O(1)$, $k_2 \ge \log(n_1 - k_1) + 2\log(1/\varepsilon) + O(1)$, and $k_1 \ge \log(n_2 - k_2) + 2\log(1/\varepsilon) + O(1)$. Then, a random function $\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a (k_1, k_2, ε)-two source extractor with probability $1 - o(1)$.*

We defer proof of this to Section 6.3. We will utilize the well known result that every two-source extractor is also a strong two-source extractor by an argument due to Boaz Barak included in [Rao07]. We rephrase it here in the form that we will apply it.

Lemma 6.5. *If $2\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a (k_1, k_2, ε)-two-source extractor, then it is a $(k_1, k_2', 2^m(\varepsilon + 2^{k_2 - k_2'})$ two-source extractor that is strong in the second argument.*

We will focus on strongness in the second argument since that is all that we will need. Using this, we will prove our main lemma:

Lemma 6.6. *Let $g, \ell, m, n \in \mathbb{N}$ and $k, k_1, k_2, \varepsilon > 0$ be such that $k \ge k_1 + \ell m + \log(1/\varepsilon)$. Suppose there exists a strong (in the second argument) (k_1, k_2, ε)-two-source extractor $2\mathsf{Ext} : \{0,1\}^{(\ell-1)\cdot n} \times \{0,1\}^n \to \{0,1\}^m$. Then we can construct a function $f : (\{0,1\}^n)^\ell \to (\{0,1\}^m)^{\ell-1}$ such that for any (g, ℓ, n, k)-oNOSF source $\mathbf{X}$, there exists a uniform $(g-1, \ell-1)$-oNOSF source $\mathbf{Y}$ such that $|f(\mathbf{X}) - \mathbf{Y}| \le 2(g-1)\varepsilon \cdot 2^{k_2 - k}$.*

Using this main lemma, Theorem 6.2 follows:

Proof of Theorem 6.2. Applying Lemma 6.4 with $k_1 = k - m\ell - \log(1/\varepsilon)$, $k_2 = \frac{2}{3}k$, and $\varepsilon = 2^{-k/15}$, we get that there exists a two-source extractor $2\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ with these parameters.

Lemma 6.5 then gives us that $2\mathsf{Ext}$ is a strong, in the second source, $(k_1, k_2', \varepsilon')$-two-source extractor where $k_2' = 2k_2$ and $\varepsilon' = 2^m(\varepsilon + 2^{k_2 - k_2'}) = 2^{-29k/48} + 2^{-k/240} = 2^{-\Omega(k)}$. Using this strong two-source extractor as input to Lemma 6.6 gives us the desired result.[8] □

One can get an explicit version of this transformation, with polynomial error by using an explicit two-source extractor, such as the one from [CZ19].

We now focus on proving Lemma 6.6. We extend an argument for a somewhere extractor for low entropy oNOSF sources from [AORSV20]. We do this by using a two source extractor instead of a seeded extractor in their construction.

To achieve this result, we use the notion of *average conditional min-entropy* and use some known results about two-source extractors.

Definition 6.7. *For any two distributions $\mathbf{X}$ and $\mathbf{W}$, define the* average conditional min-entropy *of $\mathbf{X}$ given $\mathbf{W}$ as*

$$\widetilde{H}_\infty(\mathbf{X} \mid \mathbf{W}) = -\log\left(\mathop{\mathbb{E}}_{w \sim \mathbf{W}}\left[\max_{x \in \mathsf{Supp}(\mathbf{X})} \Pr[\mathbf{X} = x \mid \mathbf{W} = w]\right]\right).$$

We use this notion of average conditional min-entropy to define notions of average-case strongness in two-source extractors. Let's first recall from Definition 4.9 that if $2\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ is a (k_1, k_2, ε)-two-source extractor, then it is said to be strong in its first argument if $2\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2), \mathbf{X}_1 \approx_\varepsilon \mathbf{U}_m, \mathbf{X}_1$.

We now define average-case strong two-source extractors:

[8] We note that we have not fully optimized our parameters here from Lemma 6.4, and it is possible to get $k_2' = (1+\gamma)k_2$ for some $\gamma > 0$ at the expense of other constants.

Definition 6.8. *We say that* 2Ext *is* average-case strong in the second argument *(with the first argument version being analogous) if*

$$2\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2), \mathbf{X}_2, \mathbf{W} \approx_\varepsilon \mathbf{U}_m, \mathbf{X}_2, \mathbf{W}$$

for every $\mathbf{X}_1$ *and* $\mathbf{W}$ *such that* $\widetilde{H}_\infty(\mathbf{X}_1 \mid \mathbf{W}) \geq k_1$ *with* $\mathbf{X}_2$ *independent of* $\mathbf{X}_1$ *and* $\mathbf{W}$.

One benefit of the average conditional min-entropy in comparison to conditional min-entropy is that the chain rule is simpler:

Lemma 6.9. *[DORS08] Let* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{C}$ *be distributions such that* $\mathsf{Supp}(\mathbf{B}) \leq 2^\lambda$. *Then* $\widetilde{H}_\infty(\mathbf{A} \mid \mathbf{B}, \mathbf{C}) \geq \widetilde{H}_\infty(\mathbf{A}, \mathbf{B} \mid \mathbf{C}) - \lambda \geq \widetilde{H}_\infty(\mathbf{A} \mid \mathbf{C}) - \lambda$.

In addition, Lemma 2.3 of [DORS08] shows that strong two-source extractors are average-case strong two-source extractors with similar parameters.

Lemma 6.10. *[DORS08] For any* $\eta > 0$, *if* 2Ext *is a* (k_1, k_2, ε)*-two-source extractor that is strong in the second argument, then* 2Ext *is a* $(k_1 + \log(1/\eta), k_2, \varepsilon + \eta))$*-two-source extractor that is average-case strong in the second argument.*

Finally, we introduce a useful fact about average-case strong seeded extractors due to [AORSV20] that we slightly generalize to average-case strong two source extractors.

Lemma 6.11. *[AORSV20] Let* $2\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *be a* (k_1, k_2, ε)*-two-source extractor that is average-case strong in the second argument. If* $\mathbf{W}$ *is a random variable such that* $\widetilde{H}_\infty(\mathbf{X}_1 \mid \mathbf{W}) \geq k_1$ *and* $\mathbf{X}_2'$ *is a* (n_2, k_2')*-source independent of* $\mathbf{X}_1$ *and* $\mathbf{X}_2$, *then*

$$2\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2'), \mathbf{X}_2', \mathbf{W} \approx_{\varepsilon \cdot 2^{k_2 - k_2'}} \mathbf{U}_m, \mathbf{X}_2', \mathbf{W}.$$

We will prove this lemma later. We can use it to prove our main theorem in which we provide a general transformation of low min-entropy oNOSF sources to uniform oNOSF sources given a two-source extractor. This transformation is based on a similar transformation in [AORSV20].

Proof of Lemma 6.6. For $i \in \{2, \ldots, \ell\}$, let $2\mathsf{Ext}_i : (\{0,1\}^n)^{i-1} \times \{0,1\}^n \to \{0,1\}^m$ be defined as $2\mathsf{Ext}_i(x, y) = 2\mathsf{Ext}(x \circ 0^{(\ell - i)n}, y)$. Then, $2\mathsf{Ext}_i$ is a (k_1, k_2, ε) strong (in the second argument) two source extractor. Since $2\mathsf{Ext}_i$ is strong in its second argument, using Lemma 6.10 with $\eta = \varepsilon$ gives us that $2\mathsf{Ext}_i$ is an average-case strong (in its second argument) $(\overline{k}_1, k_2, \overline{\varepsilon})$-two-source extractor with $\overline{k}_1 = k_1 + \log(1/\varepsilon)$ and $\overline{\varepsilon} = 2\varepsilon$. By assumption, we have that $k \geq k_1 + \ell m + \log(1/\varepsilon) = \overline{k}_1 + \ell m$, so $k - \ell m \geq \overline{k}_1$.

Next, write $\mathbf{X}$ as $\mathbf{X} = \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_\ell$ with the g good blocks (i.e., independent (n, k)-sources) at indices $G_1, \ldots, G_g$. We define $f(\mathbf{X})$ as $f(\mathbf{X}) = \mathbf{O}_2, \ldots, \mathbf{O}_\ell$ where our $\ell - 1$ output blocks are defined as $\mathbf{O}_i := 2\mathsf{Ext}_i(\mathbf{X}_{1:i-1}, \mathbf{X}_i) \in \{0,1\}^m$. We must show two properties. First, that

$$\mathbf{O}_{G_2}, \ldots, \mathbf{O}_{G_g} \approx_{2(g-1)\varepsilon \cdot 2^{k_2 - k}} \mathbf{U}_m^{\ell - 1}, \tag{1}$$

where $\mathbf{U}_m^{\ell-1}$ is $\ell - 1$ independent copies of the uniform distribution on m bits. Equation (1) says that f outputs a somewhere random source.

Second, to fulfill the uniform oNOSF source requirements, we must show that $\mathbf{O}_{G_r}$ for $r \in \{2, \ldots, g\}$ is still close to uniform even when conditioned on all previous blocks. That is, we must show

$$\mathbf{O}_1, \ldots, \mathbf{O}_{G_r - 1}, \mathbf{O}_{G_r} \approx_{2\varepsilon \cdot 2^{k_2 - k}} \mathbf{O}_1, \ldots, \mathbf{O}_{G_r - 1}, \mathbf{U}_m. \tag{2}$$

Notice that by applying the triangle inequality $g-1$ times to Equation (2), we also achieve Equation (1). Consequently, we now focus on proving Equation (2).

We would like to apply Lemma 6.11 to get Equation (2), so we must show that $\widetilde{H}_\infty(\mathbf{X}_{1:I_r-1} \mid \mathbf{O}_{2:I_r-1}) \geq \overline{k}_1$. We thus compute

$$\begin{aligned}
\widetilde{H}_\infty(\mathbf{X}_{1:G_r-1} \mid \mathbf{O}_{2:G_r-1}) &\geq \widetilde{H}_\infty(\mathbf{X}_{G_{r-1}} \mid \mathbf{O}_{2:G_r-1}) \\
&\geq \widetilde{H}_\infty(\mathbf{X}_{G_{r-1}} \mid \mathbf{O}_{2:G_{r-2}}) - (G_r - 1 - G_{r-2})m && \text{By Lemma 6.9} \\
&\geq \widetilde{H}_\infty(\mathbf{X}_{G_{r-1}} \mid \mathbf{O}_{2:G_{r-2}}) - \ell m \\
&= H_\infty(\mathbf{X}_{G_{r-1}}) - \ell m && \text{By } \mathbf{X}_{G_{r-1}} \text{ independent of } \mathbf{O}_{2:G_{r-2}} \\
&\geq k - \ell m. && \text{By assumption} \\
&\geq (k_1 + \ell m + \log(1/\varepsilon)) - \ell m && \text{By assumption} \\
&= k_1 + \log(1/\varepsilon) \\
&= \overline{k}_1,
\end{aligned}$$

as desired. Now, using the fact that $\widetilde{H}_\infty(\mathbf{X}_{1:G_r-1} \mid \mathbf{O}_{2:G_r-1}) \geq \overline{k}_1$, that $\mathbf{X}_{G_r}$ is independent of $\mathbf{X}_{1:G_r-1}$ and $\mathbf{O}_{2:G_r-1}$, and that Ext_{G_r} is $(\overline{k}_1, k_2, \overline{\varepsilon})$-average-case strong in its second argument, Lemma 6.11 gives us exactly Equation (2).

Finally, we simply let $\mathbf{Y} = \mathbf{Y}_2, \ldots \mathbf{Y}_\ell$ where $\mathbf{Y}_i = U_m$ if $i \in \{G_1, \ldots, G_g\}$ and $\mathbf{Y}_i = \mathbf{O}_i$ otherwise. Notice that the uniform blocks of $\mathbf{Y}$ are independent from each other and from all blocks with a smaller index, meaning that $\mathbf{Y}$ is a uniform $(g-1, \ell-1)$-oNOSF source. Then Equation (1) gives us that $|f(\mathbf{X}) - \mathbf{Y}| \leq 2(g-1)\varepsilon \cdot 2^{k_2-k}$, completing the proof. □

We lastly prove our useful lemma regarding average case strong two-source extractors:

Proof of Lemma 6.11. Without loss of generality, assume that $\mathbf{X}_2'$ is a flat k_2' source. Let $S \subseteq \{0,1\}^{n_2}$ be such that $S \supseteq \mathsf{Supp}(\mathbf{X}_2')$ and $|S| = 2^{k_2}$, and define $\mathbf{X}_2$ to be the flat source on S, so $\mathbf{X}_2$ is a (n_2, k_2)-source. Because $2\mathsf{Ext}$ is average-case strong in its second argument, we have that

$$\begin{aligned}
|2\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2), \mathbf{X}_2, \mathbf{W} - \mathbf{U}_m, \mathbf{X}_2, \mathbf{W}| &= \sum_{x_2 \in \mathsf{Supp}(\mathbf{X}_2)} 2^{-k_2} \left|2\mathsf{Ext}(\mathbf{X}_1, x_2), \mathbf{W} - \mathbf{U}_m, \mathbf{W}\right| \\
&\leq \varepsilon.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\varepsilon &\geq \sum_{x_2 \in \mathsf{Supp}(\mathbf{X}_2)} 2^{-k_2} \left|2\mathsf{Ext}(\mathbf{X}_1, x_2), \mathbf{W} - \mathbf{U}_m, \mathbf{W}\right| \\
&\geq \sum_{x_2 \in \mathsf{Supp}(\mathbf{X}_2')} 2^{-k_2} \left|2\mathsf{Ext}(\mathbf{X}_1, x_2), \mathbf{W} - \mathbf{U}_m, \mathbf{W}\right| \\
&= 2^{k_2'-k_2} \sum_{x_2 \in \mathsf{Supp}(\mathbf{X}_2')} 2^{-k_2'} \left|2\mathsf{Ext}(\mathbf{X}_1, x_2), \mathbf{W} - \mathbf{U}_m, \mathbf{W}\right| \\
&= 2^{k_2'-k_2} \left|\mathsf{Ext}(\mathbf{X}_1, \mathbf{X}_2'), \mathbf{X}_2', \mathbf{W} - \mathbf{U}_m, \mathbf{X}_2', \mathbf{W}\right|.
\end{aligned}$$

Rearranging the last line gives us the desired inequality. □

6.2 Condensing from oNOSF sources using output-light two source extractors

In this subsection, we will prove Theorem 6.3. To obtain the condenser, we will utilize two-source extractors which have an additional property that we call output-light.

Formally, we define output-light two source extractors as follows:

Definition 6.12 (Output-light Two Source Extractor)**.** *Let* $\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *be a* (k_1, k_2, ε)*-two source extractor. Then,* Ext *is* R*-output-light if for every* $z \in \{0,1\}^m$*, it holds that* $|\{x \in \{0,1\}^{n_1} : \exists y \in \{0,1\}^{n_2}(\mathsf{Ext}(x,y) = z)\}| \leq R$*.*

We will show a random function is a output-light two source extractor with strong parameters and we will use it with the following parameters:

Lemma 6.13. *Let* $0 < \delta < 1, C \geq 4$ *be arbitrary constants. Let* n_1*,* k_1*,* n_2*,* k_2*,* m*,* ε *be such that* n_1 *is arbitrary,* $n_2 = C\log(n_1), k_1 = \delta n_1 - 2n_2, k_2 = 4\log(n_1), m = k_1 - 2n_2, \varepsilon = 2^{-k_2/4}$ *(note that if* k_2 *is larger than the minimum requirement, then* ε *gets proportionally smaller). Then, a random function* $\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is an* R*-output-light two-source-extractor where* $R = 2^{n_1+n_2-m+O(1)}$*.*

We defer proof of their existence in Section 6.3. Using such an extractor, we will prove the following general condensing result:

Lemma 6.14. *Let* g, ℓ, r, n *be such that* $r = \lfloor \ell/g \rfloor$ *and* $r < \ell/g$*. Assume that for* $c \in \{1, \ldots, r\}$*, there exists an* R_c*-output-light* $(k_{1,c}, k_{2,c}, \varepsilon_{\mathsf{Ext}_c})$*-two-source extractor* $2\mathsf{Ext}_c : \{0,1\}^{n_{1,c}} \times \{0,1\}^{n_{2,c}} \to \{0,1\}^{m_c}$ *where* $n_{1,c} = gn$*,* $n_{2,c} = \frac{5^{\ell-cg}-1}{4} \cdot 4\log(gn)$*,* $k_{1,c} = n - 2n_{2,c}$*,* $k_{2,c} = 4\log(gn)$*,* $m_c = n - 2n_{2,c}, \varepsilon_{\mathsf{Ext}_c} = 2^{-k_{2,c}/4}$ *and* $\log(R_c/\varepsilon) \leq n_{1,c} + 2n_{2,c} - m_c$*. Then there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform* (g,ℓ)*-oNOSF source* $\mathbf{X}$*, we have* $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq \frac{1}{r} \cdot m - 2n_{2,1}$ *here* $m = r \cdot m_r, \varepsilon = 2^{\log(gn)/4}$*.*

Using this main lemma, the theorem follows:

Proof of Theorem 6.3. We plug in the result of Lemma 6.13 to Lemma 6.14 to infer our claim. □

Before we prove this main lemma, we prove Theorem 6.3 for the special case when $g > \ell/2$.

Theorem 6.15. *For all* g, ℓ *such that* $g > \ell/2$*, there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform* (g,ℓ)*-oNOSF source* $\mathbf{X}$*,* $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - (5^{\ell-g} - 3)\log(gn)$ *where* $m = n - 2(5^{\ell-g} - 1)\log(gn), \varepsilon = 2^{-\log(gn)/4}$*.*

As an application of this theorem, we construct a condenser from a low min-entropy (g,ℓ)-oNOSF source with $g > \ell/2 + 1$. We do this by composing our transformation from Theorem 6.2 with the condenser from Theorem 6.15.

Corollary 6.16. *For all* g *and* ℓ *such that* $g > \ell/2+1$*, there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any* (g,ℓ,n,k)*-oNOSF source* $\mathbf{X}$ *with* $k \geq 2\log(n)$*, we have* $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - (5^{\ell-g} - 3)\log\left(\frac{(g-1)k}{8\ell}\right)$ *where* $m = \frac{k}{8\ell} - 2(5^{\ell-g} - 1)\log\left(\frac{(g-1)k}{8\ell}\right)$ *and* $\varepsilon = (g-1) \cdot 2^{-\Omega(k)} + 2^{-\log\left(\frac{(g-1)k}{8\ell}\right)/4}$*.*

Proof. We begin by transforming $\mathbf{X}$ into a uniform $\left(g' = g-1, \ell' = \ell-1, n' = \frac{k}{8\ell}\right)$-oNOSF source $\mathbf{X}'$ defined as $\mathbf{X}' = f(\mathbf{X})$ by taking f from Theorem 6.2. In this step, we accumulate $\varepsilon_1 = (g-1)\cdot 2^{-\Omega(k)}$ error.

As $g' > \ell'/2$, we apply Theorem 6.15 to get a condenser $\mathsf{Cond}' : (\{0,1\}^{n'})^{\ell'} \to \{0,1\}^m$ such that $H^{\varepsilon_2}_\infty(\mathsf{Cond}(\mathbf{X}')) \geq m - (5^{\ell'-g'}-3)\log(g'n')$ where $m = n' - 2(5^{\ell'-g'}-1)\log(g'n'), \varepsilon_2 = 2^{-\log(g'n')/4}$. Simplifying these expressions then yields $m = \frac{k}{8\ell} - 2(5^{\ell-g}-1)\log\left(\frac{(g-1)k}{8\ell}\right)$, $\varepsilon_2 = 2^{-\log\left(\frac{(g-1)k}{8\ell}\right)/4}$, and $H^{\varepsilon_2}_\infty(\mathsf{Cond}(\mathbf{X}')) \geq m - (5^{\ell-g}-3)\log\left(\frac{(g-1)k}{8\ell}\right)$.

We finish by combining these two steps and defining $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ as $\mathsf{Cond}(\mathbf{X}) := \mathsf{Cond}'(f(\mathbf{X}))$ so that $H^{\varepsilon}_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - (5^{\ell-g}-3)\log\left(\frac{(g-1)k}{8\ell}\right)$ where $\varepsilon = \varepsilon_1 + \varepsilon_2 = (g-1)\cdot 2^{-\Omega(k)} + 2^{-\log\left(\frac{(g-1)k}{8\ell}\right)/4}$. □

6.2.1 Condensing from (g,ℓ)-oNOSF sources with $g > \ell/2$

We will prove Theorem 6.15 that allows us to condense from uniform (g,ℓ)-oNOSF sources when $g > \ell/2$. This theorem allows us to condense to almost full entropy.

We will prove this theorem using the following general lemma:

Lemma 6.17. *Assume that for some g, n there exists an R-output-light $(k_1, k_2, \varepsilon_{\mathsf{Ext}})$-two-source-extractor $\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ where $n_1 = gn, n_2 = \frac{5^{\ell-g}-1}{4}\cdot 4\log(gn), k_1 = n - 2n_2, k_2 = 4\log(gn), m = n - 2n_2, \varepsilon_{\mathsf{Ext}} = 2^{-k_2/4}$ (notice that we require that if k_2 supplied is larger, then $\varepsilon_{\mathsf{Ext}}$ gets proportionally smaller). Then, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g,ℓ)-oNOSF source $\mathbf{X}$ with $g > \ell/2$, $H^{\varepsilon}_\infty(\mathsf{Cond}(\mathbf{X})) = \min(m - n_2, n_1 - \log(R/\varepsilon))$ where $\varepsilon = 2^{-\log(gn)/4}$.*

Using this, our theorem directly follows:

Proof of Theorem 6.15. We use the output-light two source extractor guaranteed from Lemma 6.13 in Lemma 6.17 to get our result. □

Towards proving our general lemma, we show that for any flat distribution $\mathbf{X}$ over n bits, if a function f condenses from $\mathbf{X}$, then f also condenses (with a slight loss in parameters) from a distribution $\mathbf{X}'$ which is the same as the distribution $\mathbf{X}$ on most output bits but some output bits are arbitrarily controlled by an adversary. We note that a lemma similar in spirit to this one was shown as Lemma 28 in [BCDT19].

Lemma 6.18. *Let $\mathbf{X} \sim \{0,1\}^n$ be an arbitrary flat distribution and let $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^m$ be such that $H^{\varepsilon}_\infty(\mathsf{Cond}(\mathbf{X})) = k$. Let $G \subset [n]$ with $|G| = n - b$. Let $\mathbf{X}_G \sim \{0,1\}^{n-b}$ be the projection of $\mathbf{X}$ onto G. Let $\mathbf{X}' \sim \{0,1\}^n$ be the distribution where the output bits defined by G equal $\mathbf{X}_G$ and remaining b bits are deterministic functions of the $n-b$ bits defined by G under the restriction that $\mathrm{Supp}(\mathbf{X}') \subset \mathrm{Supp}(\mathbf{X})$. Then, $H^{\varepsilon'}_\infty(f(\mathbf{X}')) \geq k - b$ where $\varepsilon' = \varepsilon \cdot 2^b$.*

We will prove this result later. Using this result, we use output-light two-source-extractor to prove our general lemma:

Proof of Lemma 6.17. Let the input be $x = (x_1, \ldots, x_\ell)$. For $g+1 \leq i \leq \ell$, let y_i be the first $5^{\ell-i}\cdot(4\log(gn))$ bits of x_i. Let $z_1 = x_1 \circ \cdots \circ x_g, z_2 = y_{g+1} \circ \cdots \circ y_\ell$. Then, let $\mathsf{Cond}(x) = \mathsf{Ext}(z_1, z_2)$.

Let $\mathbf{Y}_i$ be the distribution of y_i and let $\mathbf{Z}_i$ be the distribution of z_i. We consider cases on the position of the adversary and show that for all such oNOSF sources, the output will be condensed:

Case 1. At least one source out of $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ is good.
Let this block be $\mathbf{X}_j$ (so $g+1 \le j \le \ell$). As $g > \ell/2$, at least one source out of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ is good and hence, $H_\infty(\mathbf{Z}_1) \ge n$. Without loss of generality, we assume only these 2 sources are good in $\mathbf{X}$ and remaining sources are bad. Let $\mathbf{A} = \mathbf{Y}_{g+1}, \ldots, \mathbf{Y}_{j-1}, \mathbf{B} = \mathbf{Y}_{j+1}, \ldots, \mathbf{Y}_\ell$. Then, $\mathbf{Z}_2 = \mathbf{A} \circ \mathbf{Y}_j \circ \mathbf{B}$. As $\mathbf{X}_j$ is a good source and $\mathbf{X}$ is a oNOSF source, $\mathbf{X}_j$ remains a uniform source conditioned on any fixing of $\mathbf{A}$, so $\mathbf{Y}_j$ does as well. Also, by the min-entropy chain rule (Lemma 4.4), with probability $1 - \varepsilon/2$ over fixings of $\mathbf{A}$, $H_\infty(\mathbf{Z}_1) \ge n - n_2 - \log(2/\varepsilon) \ge k_1$.

Consider $(\mathbf{Z}|\mathbf{A} = a)$ where a is such a good fixing of $\mathbf{A}$. We will show that for all such good fixings $H_\infty^{\varepsilon/2}\mathsf{Ext}(\mathbf{Z}) \ge m - n_2$. By assumption, $H_\infty(\mathbf{Z}_1|\mathbf{A} = a) \ge k_1$ and $H_\infty(\mathbf{Y}_j|\mathbf{A} = a) = H_\infty(\mathbf{Y}_j) = 5^{\ell-j} \cdot 4\log(gn)$. Moreover, we can without loss of generality assume $H_\infty(\mathbf{Z}_1|\mathbf{A} = a)$ is a flat source (we can express it as convex combination of such flat sources). As $\mathbf{X}$ is a oNOSF source, $(\mathbf{Z}_1|\mathbf{A} = a)$ and $(\mathbf{Y}_j|\mathbf{A} = a)$ are independent distributions. Assume for now that $(\mathbf{B}|\mathbf{A} = a)$ were uniform and independent of $\mathbf{Y}_j$ and $\mathbf{A}$. Then, $(\mathbf{Z}_1|\mathbf{A} = a)$ and $(\mathbf{Z}_2|\mathbf{A} = a) = (a, \mathbf{Y}_j, \mathbf{B}|\mathbf{A} = a)$ will be independent sources with min-entropy at least k_1 and $k_2' = \sum_{i=j}^{\ell} 5^{\ell-j}(4\log(gn)) = \frac{5^{\ell-j+1}-1}{4} \cdot 4\log(gn) \ge k_2$, respectively. Hence, $\mathsf{Ext}(\mathbf{Z})$ will be $\varepsilon_{\mathsf{Ext}}$ close to the uniform distribution over m bits where $\varepsilon_{\mathsf{Ext}} = 2^{-k_2'/4}$. However, in reality, $\mathbf{B}$ might be arbitrarily controlled by an adversary and can depend on $\mathbf{Z}_1, \mathbf{A}, \mathbf{Y}_j$. The number of bits controlled by the adversary is $n_b = \sum_{i=j+1}^{l} 5^{\ell-i}(4\log(gn)) = \frac{5^{\ell-j}-1}{4} \cdot 4\log(gn)$. To overcome this, we apply Lemma 6.18 (using the fact that $(\mathbf{Z}_1, \mathbf{A}, \mathbf{U}_{k_2'})|\mathbf{A} = a$ is a flat distribution) and infer that $H_\infty^{\varepsilon'}(\mathsf{Ext}(\mathbf{Z})) \ge m - n_b$ where $\varepsilon' = \varepsilon_{\mathsf{Ext}} \cdot 2^{n_b} \le 2^{-k_2'/4+n_b}$. As $n_b \le n_2$, the output min-entropy is at least $m - n_b \ge m - n_2$. Moreover, $\varepsilon' \le 2^{-k_2'/4+n_b}$ which is $\le 2^{-\log(gn)} \le \varepsilon/2$ if $j = \ell$ and also which is $\le 2^{-\log(gn)\cdot(5^{\ell-j}+3)/16} \le 2^{-\log(gn)/2} \le \varepsilon/2$ if $j \le \ell - 1$.

Case 2. All of the sources $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ are bad.
This implies sources $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are good. Let $k_{out} = n_1 - \log(R/\varepsilon)$. Let $N_1 = 2^{n_1}, M = 2^m$, and $K_{out} = 2^{k_{out}}$. Assume that there exists a uniform (g, ℓ)-oNOSF source $\mathbf{X}$ such that $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) < k_{out}$. By Claim 4.5, there exists $H \subset \{0,1\}^m$ such that $|H| < K_{out}$ and $\Pr[\mathsf{Cond}(\mathbf{X}) \in H] \ge \varepsilon$. This implies there exist $h \in H$ and $P \subset \{0,1\}^{n_1}$ with $|P| > \frac{\varepsilon N_1}{K_{out}} = R$ such that for all $z_1 \in P$, there exists $z_2 \in \{0,1\}^{n_2}$ so that $\mathsf{Cond}(z_1, z_2) = h$. However, this contradicts the fact that Ext is R-output-light. Hence, for all uniform (g, ℓ)-oNOSF sources(g, ℓ) $\mathbf{X}$, $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \ge n_1 - \log(R/\varepsilon)$.

□

We finally prove our useful lemma that states a condenser for a distribution $\mathbf{X}$ still condenses from a tampered version of $\mathbf{X}$ where some output bits are controlled by an adversary.

Proof of Lemma 6.18. We first claim that as $\mathbf{X}$ is a flat source, for all $x \in \mathrm{Supp}(\mathbf{X}')$, $\Pr[\mathbf{X}' = x] \le 2^b \cdot \Pr[\mathbf{X} = x]$. Indeed, let $S_x = \{z \in \mathrm{Supp}(\mathbf{X}) : x \text{ and } z \text{ equal each other when restricted to bits in } G\}$. Then, $|S_x| \le 2^b$. Hence,

$$\Pr[\mathbf{X}' = x] \le \sum_{z \in S_x} \Pr[\mathbf{X} = z] = |S_x| \cdot \Pr[\mathbf{X} = x] \le 2^b \cdot \Pr[\mathbf{X} = x].$$

We now proceed by contradiction and assume $H_\infty^{\varepsilon'}(f(\mathbf{X}')) < k - b$. Let $O = \{y \in \{0,1\}^m : \Pr[f(\mathbf{X}') = y] > 2^{k-b}\}$. As $H_\infty^{\varepsilon'}(f(\mathbf{X}')) < k - b$, it must be that $\Pr[f(\mathbf{X}') \in O] \geq \varepsilon' + |O| \cdot 2^{b-k}$. Let $I = \{x \in \mathrm{Supp}(\mathbf{X}') : f(x) \in O\}$. We now see that

$$\Pr[f(\mathbf{X}) \in O] \geq \Pr[\mathbf{X} \in I] = \sum_{x \in I} \Pr[\mathbf{X} = x] \geq \sum_{x \in I} \Pr[\mathbf{X}' = x] \cdot 2^{-b} = (\varepsilon' + |O| \cdot 2^{b-k}) \cdot 2^{-b} = \varepsilon + |O| \cdot 2^{-k}$$

where the first inequality follows by our observation. For $y \in O$, let $I_y = \{x \in I : f(x) = y\}$. We see that

$$\Pr[f(\mathbf{X}) = y] \geq \Pr[\mathbf{X} \in I_y] = \sum_{x \in I_y} \Pr[\mathbf{X} = x] \geq \sum_{x \in I_y} \Pr[\mathbf{X}' = x] \cdot 2^{-b} > 2^{-k+b} \cdot 2^{-b} = 2^{-k}$$

Hence, for all $y \in O$, $\Pr[f(\mathbf{X}) = y] > 2^{-k}$ and $\Pr[f(\mathbf{X}) \in O] \geq \varepsilon + |O| \cdot 2^{-k}$. These together imply $H_\infty^{\varepsilon}(f(\mathbf{X})) < k$, a contradiction.

□

6.2.2 Condensing from uniform oNOSF sources in all regimes

We finally prove our main lemma of the section - Lemma 6.14. We will use the following simple claim that guarantees projections of high-entropy distributions have high-entropy.

Lemma 6.19. *Let $\mathbf{X}$ be an arbitrary (n,k)-source and $\pi : \{0,1\}^n \to \{0,1\}^{n-d}$ be a projection onto $n-d$ bits of $\mathbf{X}$ (i.e., removes d bits of $\mathbf{X}$). Then $\pi(\mathbf{X})$ is a $(n-d, k-d)$-source.*

Proof. Because $\mathbf{X}$ is an (n,k)-source, for any $x \in \mathsf{Supp}(\mathbf{X})$, we have that $\Pr[\mathbf{X} = x] \leq 2^{-k}$. Furthermore, for any $y \in \{0,1\}^{n-d}$, there are at most 2^d elements from $\mathrm{Supp}(\mathbf{X})$ that could map to y under π. Thus, for any $y \in \mathsf{Supp}(\pi(\mathbf{X}))$, we can compute that

$$\begin{aligned}
\Pr[\pi(\mathbf{X}) = y] &= \sum_{\substack{x \in \mathsf{Supp}(\mathbf{X}) \\ \pi(x) = y}} \Pr[\mathbf{X} = x] \\
&\leq \sum_{\substack{x \in \mathsf{Supp}(\mathbf{X}) \\ \pi(x) = y}} 2^{-k} \\
&\leq 2^d \cdot 2^{-k} = 2^{d-k}.
\end{aligned}$$

Therefore, $H_\infty(\pi(\mathbf{X})) \geq k - d$, as required. □

We are finally ready to prove the main lemma. The proof of this main lemma uses a similar strategy as in Lemma 6.17.

Proof of Lemma 6.14. We will proceed inductively on $r \in \mathbb{N}$ with the base case of $r = 1$ taken care of by Lemma 6.17. For the inductive step, take Lemma 6.14 to be true for $r - 1$.

We will output r output blocks $\mathbf{O}_1, \ldots, \mathbf{O}_r$ where each $\mathbf{O}_i \sim \{0,1\}^{m_r}$. We begin by defining our first output block $\mathbf{O}_1 \in \{0,1\}^m$ by defining $\mathbf{Y}_i$ to be the distribution after $\mathbf{X}_i$ is projected onto its first $5^{\ell-i} \cdot 4\log(gn)$ bits and setting $\mathbf{Z}_1 = \mathbf{X}_1, \ldots, \mathbf{X}_g$ and $\mathbf{Z}_2 = \mathbf{Y}_{g+1}, \ldots, \mathbf{Y}_\ell$. We let $\mathbf{O}_1$ be the first m_r bits of $2\mathsf{Ext}_1(\mathbf{Z}_1, \mathbf{Z}_2)$.

To define our last $r-1$ output blocks $\mathbf{O}_2, \ldots, \mathbf{O}_r$, we pretend for the moment that $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are bad blocks of $\mathbf{X}$. Since $\lfloor \ell/g \rfloor = r$ and $r < \ell/g$, it must be that $\left\lfloor \frac{\ell-g}{g} \right\rfloor = r-1$ and $r-1 < (\ell-g)/g$. Because $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are bad, we see that $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ is a uniform $(g, \ell-g)$-oNOSF source with $\left\lfloor \frac{\ell-g}{g} \right\rfloor = r-1$ and $r-1 < \frac{\ell-g}{g}$, meaning that we can use existence of $2\mathsf{Ext}_c$ for $c \in \{2, \ldots, r\}$ to apply our inductive hypothesis to get the output blocks $\mathbf{O}_2, \ldots, \mathbf{O}_r$ with the property that $H^\varepsilon_\infty(\mathbf{O}_2, \ldots, \mathbf{O}_r) \geq m_r - n_{2,2}$. These allow us to define $\mathsf{Cond}(\mathbf{X}) = \mathbf{O}_1, \ldots, \mathbf{O}_r$. Of course, we do not necessarily immediately have that this is true, but it will hold in our last case in our case analysis:

Case 1. At least one, but not all, of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ is good. Because not all of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are good, it must be that at least one of $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ is good. Thus, we can use the exact calculations of **Case 1.** from Lemma 6.17 to get that $H^\varepsilon_\infty(2\mathsf{Ext}_1(\mathbf{Z}_1, \mathbf{Z}_2)) \geq m_1 - n_{2,1}$. Then, because $\mathbf{O}_1$ is just $2\mathsf{Ext}_1(\mathbf{Z}_1, \mathbf{Z}_2)$ truncated to its first m_r bits, we use Lemma 6.19 to get $H^\varepsilon_\infty(\mathbf{O}_1) \geq m_1 - n_{2,1} - (m_1 - m_r) \geq m_r - 2n_{2,1}$.

Case 2. All of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are good.

In this case, we get that $\mathbf{O}_1$ is condensed by the R_1-output-lightness of $2\mathsf{Ext}_1$. We achieve this by using the exact calculations of **Case 2.** of Lemma 6.17 to get that $H^\varepsilon_\infty(2\mathsf{Ext}_1(\mathbf{Z}_1, \mathbf{Z}_2)) \geq n_{1,1} - \log(R_1/\varepsilon) \geq m_1 - 2n_{2,1}$. We again conclude by using Lemma 6.19 to get that $H^\varepsilon_\infty(\mathbf{O}_1) \geq m_r - 2n_{2,1}$.

Case 3. All of $\mathbf{X}_1, \ldots, \mathbf{X}_g$ are bad.

In this last case, we do not get that $\mathbf{O}_1$ is condensed because $\mathbf{Z}_1$ can be arbitrarily bad. Instead, we have that $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ is a uniform $(g, \ell-g)$-oNOSF source with $\left\lfloor \frac{\ell-g}{g} \right\rfloor = r-1$ and $\frac{\ell-g}{g} < r-1$, so by our inductive hypothesis it must be that $H^\varepsilon_\infty(\mathbf{O}_2, \ldots, \mathbf{O}_r) \geq m_r - 2n_{2,2} \geq m_r - 2n_{2,1}$.

In all cases, we get that $H^\varepsilon_\infty(\mathbf{O}_1) \geq m_r - 2n_{2,1}$. Thus, we can conclude that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) = H^\varepsilon_\infty(\mathbf{O}_1, \ldots, \mathbf{O}_r) \geq m_r - n_{2,1} = \frac{1}{r} \cdot m - 2n_{2,1}$ as desired. □

6.3 Existence of output-light two-source extractors

In this subsection, we show a random function is an output-light two-source extractor. Towards showing output lightness, we introduce a related notion, of R-invertible functions.

Definition 6.20 (R-invertible function)**.** *A function $f : \{0,1\}^n \to \{0,1\}^m$ is R-invertible if for every $z \in \{0,1\}^m$, it holds that $|\{x \in \{0,1\}^n : f(x) = z\}| \leq R$.*

We record the observation that R-invertible functions are also R-output light.

Observation 6.21. *Let $\mathsf{Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ be a (k_1, k_2, ε)-two source extractor. If Ext is R-invertible, then Ext is R-output-light.*

We now show that a random function is optimally invertible, hence concluding a random function is also output light.

Lemma 6.22. *Let $f : \{0,1\}^n \to \{0,1\}^m$ be a random function where $m \leq n - \log n$. Then, with probability $1 - o(1)$, f will be 2^{n-m+c}-invertible where c is a universal constant.*

Proof of Lemma 6.22. Let $R = (1+\delta)2^{n-m}$ where $\delta > 0$ is a large constant. For $z \in \{0,1\}^m$, let E_z be the event that $|\{x \in \{0,1\}^n : f(x) = z\}| > R$. Fix any such z. Let $B_1, \ldots, B_{2^n}$ be the events corresponding to whether $f(x) = z$. Using Claim 4.2, we infer that

$$\Pr\left[\sum_i B_i > R\right] \leq \exp\left(-\frac{\delta^2}{2+\delta} \cdot 2^{n-m}\right)$$

We union bound over all $z \in \{0,1\}^m$ to obtain that the probability that at least one E_z holds is at most

$$\exp\left(-\frac{\delta^2}{2+\delta} \cdot 2^{n-m}\right) \cdot 2^m \leq \exp\left(-\frac{\delta^2}{2+\delta} \cdot m + m\right) \leq o(1)$$

The claim follows. □

We now prove that a random function is a two source extractor with strong parameters:

Proof of Lemma 6.4. Let $N_1 = 2^{n_1}, N_2 = 2^{n_2}, K_1 = 2^{k_1}, K_2 = 2^{k_2}, M = 2^m$. It suffices to show that a random function is a two source extractor where the two sources have min-entropies exactly k_1, and k_2, and are flat. Using proposition 6.12 in [Vad12], we infer that a random function $\mathsf{Ext} : \{0,1\}^{n_1+n_2} \to \{0,1\}^m$ is a (k_1+k_2, ε) extractor with probability $1 - 2^{-cK_1K_2\varepsilon^2}$ where $c > 0$ is some universal constant. We union bound over all pairs of sources with min-entropies k_1, and k_2 out of n_1, and n_2 bits respectively. The probability that a random function Ext will not be a (k_1, k_2, ε)-two-source-extractor is:

$$\begin{aligned}
\binom{N_1}{K_1}\binom{N_2}{K_2}2^{-cK_1K_2\varepsilon^2} &\leq \left(\frac{eN_1}{K_1}\right)^{K_1}\left(\frac{eN_2}{K_2}\right)^{K_2} 2^{-cK_1K_2\varepsilon^2} \\
&\leq 2^{K_1\log\left(\frac{eN_1}{K_1}\right)+K_2\log\left(\frac{eN_2}{K_2}\right)-cK_1K_2\varepsilon^2} \\
&= 2^{\log(e)K_1(n_1-k_1)+\log(e)K_2(n_2-k_2)-cK_1K_2\varepsilon^2} \\
&\leq o(1)
\end{aligned}$$

where the last inequality follows because $k_2 > \log(n_1 - k_1) + 2\log(1/\varepsilon) + O(1)$, and $k_1 > \log(n_2 - k_2) + 2\log(1/\varepsilon) + O(1)$. Hence, the claim follows. □

Finally, we show output-light two-source extractors exist, as required for our constructions.

Proof of Lemma 6.13. Combining Lemma 6.22 and Lemma 6.4 along with Observation 6.21, we infer that output-light two source extractors with the promised parameters exist. □

References

[AORSV20] Divesh Aggarwal, Maciej Obremski, João Ribeiro, Luisa Siniscalchi, and Ivan Visconti. "How to Extract Useful Randomness from Unreliable Sources". en. In: *Advances in Cryptology – EUROCRYPT 2020*. Ed. by Anne Canteaut and Yuval Ishai. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 343–372. ISBN: 978-3-030-45721-1. DOI: 10.1007/978-3-030-45721-1_13 (cit. on pp. 3–7, 15, 16, 19, 39, 40, 52).

[AL93] Miklós Ajtai and Nathan Linial. "The influence of large coalitions". en. In: *Combinatorica* 13.2 (June 1993), pp. 129–145. ISSN: 1439-6912. DOI: 10.1007/BF01303199 (cit. on p. 8).

[BGM22] Marshall Ball, Oded Goldreich, and Tal Malkin. "Randomness Extraction from Somewhat Dependent Sources". In: *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*. Ed. by Mark Braverman. Vol. 215. Leibniz International Proceedings in Informatics (LIPIcs). ISSN: 1868-8969. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, 12:1–12:14. ISBN: 978-3-95977-217-4. DOI: 10.4230/LIPIcs.ITCS.2022.12 (cit. on p. 2).

[BDKPPSY11] Boaz Barak, Yevgeniy Dodis, Hugo Krawczyk, Olivier Pereira, Krzysztof Pietrzak, François-Xavier Standaert, and Yu Yu. "Leftover Hash Lemma, Revisited". en. In: *Advances in Cryptology – CRYPTO 2011*. Ed. by Phillip Rogaway. Berlin, Heidelberg: Springer, 2011, pp. 1–20. ISBN: 978-3-642-22792-9. DOI: 10.1007/978-3-642-22792-9_1 (cit. on p. 3).

[BCDT19] Avraham Ben-Aroya, Gil Cohen, Dean Doron, and Amnon Ta-Shma. "Two-Source Condensers with Low Error and Small Entropy Gap via Entropy-Resilient Functions". en. In: *DROPS-IDN/v2/document/10.4230/LIPIcs.APPROX-RANDOM.2019.43*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2019.43 (cit. on pp. 2, 7, 43).

[BL85] Michael Ben-Or and Nathan Linial. "Collective coin flipping, robust voting schemes and minima of Banzhaf values". In: *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*. ISSN: 0272-5428. Oct. 1985, pp. 408–416. DOI: 10.1109/SFCS.1985.15 (cit. on p. 8).

[BL89] Michael Ben-Or and Nathan Linial. "Collective Coin Flipping". In: *Advances In Computing Research* 5 (1989), pp. 91–115 (cit. on p. 7).

[BKKKL92] Jean Bourgain, Jeff Kahn, Gil Kalai, Yitzhak Katznelson, and Nathan Linial. "The influence of variables in product spaces". en. In: *Israel Journal of Mathematics* 77.1 (Feb. 1992), pp. 55–64. ISSN: 1565-8511. DOI: 10.1007/BF02808010 (cit. on p. 7).

[Cha16] Eshan Chattopadhyay. "Explicit Two-Source Extractors and More". PhD thesis. Austin, TX: The University of Texas at Austin, May 2016 (cit. on p. 18).

[CGL20] Eshan Chattopadhyay, Vipul Goyal, and Xin Li. "Nonmalleable Extractors and Codes, with Their Many Tampered Extensions". In: *SIAM Journal on Computing* 49.5 (Jan. 2020). Publisher: Society for Industrial and Applied Mathematics, pp. 999–1040. ISSN: 0097-5397. DOI: 10.1137/18M1176622 (cit. on p. 18).

[CZ19] Eshan Chattopadhyay and David Zuckerman. "Explicit two-source extractors and resilient functions". In: *Annals of Mathematics* 189.3 (May 2019). Publisher: Department of Mathematics of Princeton University, pp. 653–705. ISSN: 0003-486X, 1939-8980. DOI: 10.4007/annals.2019.189.3.1 (cit. on pp. 7, 8, 39, 56).

[CG88] Benny Chor and Oded Goldreich. "Unbiased Bits from Sources of Weak Randomness and Probabilistic Communication Complexity". In: *SIAM Journal on Computing* 17.2 (Apr. 1988). Publisher: Society for Industrial and Applied Mathematics, pp. 230–261. ISSN: 0097-5397. DOI: 10.1137/0217015 (cit. on pp. 1, 2, 5, 6, 18).

[CGHFRS85] Benny Chor, Oded Goldreich, Johan Hasted, Joel Freidmann, Steven Rudich, and Roman Smolensky. "The bit extraction problem or t-resilient functions". In: *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*. SFCS '85. USA: IEEE Computer Society, Oct. 1985, pp. 396–407. DOI: 10.1109/SFCS.1985.55 (cit. on p. 7).

[CM24] Joshua Cook and Dana Moshkovitz. *Explicit Time and Space Efficient Encoders Exist Only With Random Access*. en. ISSN: 1433-8092. Feb. 2024 (cit. on pp. 4, 52).

[DOPS04] Y. Dodis, Shien Jin Ong, M. Prabhakaran, and A. Sahai. "On the (im)possibility of cryptography with imperfect randomness". In: *45th Annual IEEE Symposium on Foundations of Computer Science*. ISSN: 0272-5428. Oct. 2004, pp. 196–205. DOI: 10.1109/FOCS.2004.44 (cit. on p. 1).

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. "Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data". In: *SIAM Journal on Computing* 38.1 (Jan. 2008). Publisher: Society for Industrial and Applied Mathematics, pp. 97–139. ISSN: 0097-5397. DOI: 10.1137/060651380 (cit. on p. 40).

[DPW14] Yevgeniy Dodis, Krzysztof Pietrzak, and Daniel Wichs. "Key Derivation without Entropy Waste". en. In: *Advances in Cryptology – EUROCRYPT 2014*. Ed. by Phong Q. Nguyen and Elisabeth Oswald. Berlin, Heidelberg: Springer, 2014, pp. 93–110. ISBN: 978-3-642-55220-5. DOI: 10.1007/978-3-642-55220-5_6 (cit. on p. 3).

[DRV12] Yevgeniy Dodis, Thomas Ristenpart, and Salil Vadhan. "Randomness Condensers for Efficiently Samplable, Seed-Dependent Sources". en. In: *Theory of Cryptography*. Ed. by Ronald Cramer. Berlin, Heidelberg: Springer, 2012, pp. 618–635. ISBN: 978-3-642-28914-9. DOI: 10.1007/978-3-642-28914-9_35 (cit. on p. 3).

[DY13] Yevgeniy Dodis and Yu Yu. "Overcoming Weak Expectations". en. In: *Theory of Cryptography*. Ed. by Amit Sahai. Berlin, Heidelberg: Springer, 2013, pp. 1–22. ISBN: 978-3-642-36594-2. DOI: 10.1007/978-3-642-36594-2_1 (cit. on p. 3).

[DMOZ23] Dean Doron, Dana Moshkovitz, Justin Oh, and David Zuckerman. "Almost Chor-Goldreich Sources and Adversarial Random Walks". In: *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. STOC 2023. New York, NY, USA: Association for Computing Machinery, June 2023, pp. 1–9. ISBN: 978-1-4503-9913-5. DOI: 10.1145/3564246.3585134 (cit. on pp. 2, 5–7, 16).

[Fri04] Ehud Friedgut. "Influences in Product Spaces: KKL and BKKKL Revisited". en. In: *Combinatorics, Probability and Computing* 13.1 (Jan. 2004), pp. 17–29. ISSN: 1469-2163, 0963-5483. DOI: 10.1017/S0963548303005832 (cit. on p. 7).

[GKL15] Juan Garay, Aggelos Kiayias, and Nikos Leonardos. "The Bitcoin Backbone Protocol: Analysis and Applications". en. In: *Advances in Cryptology - EUROCRYPT 2015*. Ed. by Elisabeth Oswald and Marc Fischlin. Vol. 9057. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 281–310. ISBN: 978-3-662-46803-6. DOI: 10.1007/978-3-662-46803-6_10 (cit. on p. 3).

[GP20] Dmitry Gavinsky and Pavel Pudlák. "Santha-Vazirani sources, deterministic condensers and very strong extractors". en. In: *Theory of Computing Systems* 64.6 (Aug. 2020), pp. 1140–1154. ISSN: 1433-0490. DOI: 10.1007/s00224-020-09975-8 (cit. on p. 6).

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. "Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes". In: *Journal of the ACM* 56.4 (July 2009), 20:1–20:34. ISSN: 0004-5411. DOI: 10.1145/1538902.1538904 (cit. on pp. 2, 18).

[ILL89] Russell Impagliazzo, Leonid A. Levin, and Michael Luby. "Pseudo-random Generation from one-way functions (Extended Abstracts)". In: *Proceedings of the 21st Annual ACM Symposium on Theory of Computing, May 14-17, 1989, Seattle, Washington, USA*. Ed. by David S. Johnson. ACM, 1989, pp. 12–24. DOI: 10.1145/73007.73009 (cit. on p. 18).

[KKL88] J. Kahn, G. Kalai, and N. Linial. "The influence of variables on Boolean functions". In: *[Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science*. Oct. 1988, pp. 68–80. DOI: 10.1109/SFCS.1988.21923 (cit. on p. 7).

[KN23] Swastik Kopparty and Vishvajeet N. "Extracting Mergers and Projections of Partitions". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2023, September 11-13, 2023, Atlanta, Georgia, USA*. Ed. by Nicole Megow and Adam D. Smith. Vol. 275. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023, 52:1–52:22. DOI: 10.4230/LIPICS.APPROX/RANDOM.2023.52 (cit. on pp. 8, 57).

[Li16] Xin Li. "Improved Two-Source Extractors, and Affine Extractors for Polylogarithmic Entropy". en. In: *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. New Brunswick, NJ, USA: IEEE, Oct. 2016, pp. 168–177. ISBN: 978-1-5090-3933-3. DOI: 10.1109/FOCS.2016.26 (cit. on p. 56).

[Mek17] Raghu Meka. "Explicit Resilient Functions Matching Ajtai-Linial". In: *Proceedings of the 2017 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. Proceedings. Society for Industrial and Applied Mathematics, Jan. 2017, pp. 1132–1148. DOI: 10.1137/1.9781611974782.73 (cit. on p. 8).

[PSS17] Rafael Pass, Lior Seeman, and Abhi Shelat. "Analysis of the Blockchain Protocol in Asynchronous Networks". en. In: *Advances in Cryptology – EUROCRYPT 2017*. Ed. by Jean-Sébastien Coron and Jesper Buus Nielsen. Cham: Springer International Publishing, 2017, pp. 643–673. ISBN: 978-3-319-56614-6. DOI: 10.1007/978-3-319-56614-6_22 (cit. on p. 3).

[RT00] Jaikumar Radhakrishnan and Amnon Ta-Shma. "Bounds for Dispersers, Extractors, and Depth-Two Superconcentrators". In: *SIAM Journal on Discrete Mathematics* 13.1 (Jan. 2000). Publisher: Society for Industrial and Applied Mathematics, pp. 2–24. ISSN: 0895-4801. DOI: 10.1137/S0895480197329508 (cit. on p. 2).

[Rao07] Anup Rao. *An Exposition of Bourgain's 2-Source Extractor*. en. Tech. rep. TR07-034. ISSN: 1433-8092. Electronic Colloquium on Computational Complexity (ECCC), Apr. 2007 (cit. on p. 39).

[RSW06] Omer Reingold, Ronen Shaltiel, and Avi Wigderson. "Extracting Randomness via Repeated Condensing". In: *SIAM Journal on Computing* 35.5 (Jan. 2006). Publisher: Society for Industrial and Applied Mathematics, pp. 1185–1209. ISSN: 0097-5397. DOI: 10.1137/S0097539703431032 (cit. on p. 2).

[RVW02] Omer Reingold, Salil Vadhan, and Avi Wigderson. "Entropy Waves, the Zig-Zag Graph Product, and New Constant-Degree Expanders". In: *Annals of Mathematics* 155.1 (2002), pp. 157–187. ISSN: 0003486X (cit. on p. 2).

[RVW04] Omer Reingold, Salil Vadhan, and Avi Wigderson. *A Note on Extracting Randomness from Santha-Vazirani Sources*. en. Tech. rep. 2004 (cit. on pp. 1, 2).

[SV86] Miklos Santha and Umesh V Vazirani. "Generating quasi-random sequences from semi-random sources". In: *Journal of Computer and System Sciences* 33.1 (Aug. 1986), pp. 75–87. ISSN: 0022-0000. DOI: 10.1016/0022-0000(86)90044-9 (cit. on pp. 1, 2, 6).

[TUZ07] Amnon Ta-Shma, Christopher Umans, and David Zuckerman. "Lossless Condensers, Unbalanced Expanders, And Extractors". en. In: *Combinatorica* 27.2 (Mar. 2007), pp. 213–240. ISSN: 1439-6912. DOI: 10.1007/s00493-007-0053-2 (cit. on p. 2).

[Vad12] Salil P. Vadhan. "Pseudorandomness". English. In: *Foundations and Trends® in Theoretical Computer Science* 7.1–3 (Dec. 2012). Publisher: Now Publishers, Inc., pp. 1–336. ISSN: 1551-305X, 1551-3068. DOI: 10.1561/0400000010 (cit. on pp. 2, 47).

[Vaz85] Umesh V. Vazirani. "Towards a Strong Communication Complexity Theory or Generating Quasi-Random Sequences from Two Communicating Slightly-random Sources (Extended Abstract)". In: *Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA*. Ed. by Robert Sedgewick. ACM, 1985, pp. 366–378. DOI: 10.1145/22145.22186 (cit. on p. 18).

[Zuc90] David Zuckerman. "General weak random sources". In: *Proceedings of the 31st Annual Symposium on Foundations of Computer Science*. SFCS '90. USA: IEEE Computer Society, Oct. 1990, 534–543 vol.2. ISBN: 978-0-8186-2082-9. DOI: 10.1109/FSCS.1990.89574 (cit. on p. 1).

[Zuc07] David Zuckerman. "Linear Degree Extractors and the Inapproximability of Max Clique and Chromatic Number". In: *Theory of Computing* 3 (Aug. 2007). Number: 6 Publisher: Theory of Computing, pp. 103–128. DOI: 10.4086/toc.2007.v003a006 (cit. on pp. 2, 17).

A Explicit Condensers for oNOSF Sources

In this section, we provide explicit constructions of condensers for uniform (g, ℓ)-oNOSF sources for uniform $(2, 3)$-oNOSF sources directly in Appendix A.1 and for uniform $(6, 9)$-oNOSF sources

along with various other settings of parameters via a recursive nesting method in Appendix A.2.

Recall that [AORSV20] showed that for all $\gamma > 0$ there exists an ℓ large enough such that it is impossible to extract from uniform $(\lfloor \gamma \ell \rfloor, \ell)$-oNOSF sources below error $\frac{1-\gamma}{48}$. One may then wonder whether the explicit condensers for uniform (g, ℓ)-oNOSF sources that we constructed in the previous couple subsections are for some "easy" case of small g and ℓ, such as uniform $(2, 3)$ and $(6, 9)$-oNOSF sources in Theorem A.1 and Theorem A.3. In Appendix B, we dispel such worries by showing that one cannot extract from rate 2/3 uniform NOSF sources.

A.1 An explicit condenser for uniform $(2, 3)$-oNOSF sources

In this section, we construct a condenser for uniform $(2, 3)$-oNOSF sources. The following is our main result.

Theorem A.1. *There exists constant $0 < c_0 < 1$ such that for all $\varepsilon > 2^{-c_0 n}$, we can explicitly construct a condenser $\mathsf{Cond} : (\{0,1\}^n)^3 \to \{0,1\}^m$, where $m = \frac{n}{16}$ such that for any uniform $(2, 3)$-oNOSF source $\mathbf{X}$, $H^{\varepsilon}_{\infty}(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$.*

To prove this, we construct an explicit output-light seeded extractor (see Definition 6.12) that works for somewhere-random sources. We observe that in the proof of Lemma 6.17 for $g = 2, \ell = 3$, it suffices to construct an output-light seeded extractor instead of an output-light two-source extractor. And moreover, this output-lights seeded extractor need only extract from somewhere random sources.

We could in fact use use existing seeded extractors that are known to be invertible, such as Trevisan's extractor [CM24]. However, this requires seed length of $O(\log^2(n))$, which translates into the entropy gap of the condenser. For the specific case of somewhere random sources, we construct a better seeded extractor that has seed length $O(\log(n))$.

Theorem A.2. *There exists constant $0 < c_0 < 1$ such that for all $\varepsilon > 2^{-c_0 n}$, there exists a R-output-light strong linear seeded ε-extractor $\mathsf{Ext} : \{0,1\}^{2n} \times \{0,1\}^d \to \{0,1\}^m$ for the class of distributions $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, each $\mathbf{X}_i$ being a random variable on n bits and at least one of $\mathbf{X}_1$ or $\mathbf{X}_2$ is guaranteed to be uniform, with $d = O(\log n/\varepsilon), m = \frac{n}{16}$ and $R = \frac{2^{2n-m}}{\mathrm{poly}(m, 1/\varepsilon)}$.*

We note that this construction matches the probabilistic bounds (Theorem 6.15) as the m bit output is condensed to entropy $m - O(\log(m))$ with $m = O(n)$. We also remark that we have not tried to optimize the constant appearing in the output length of the extractor.

A.1.1 An explicit output-light seeded extractor for somewhere-random sources

We prove Theorem A.2 in this section and show

Proof of Theorem A.2. We claim that Ext computed by Algorithm 3 computes the desired extractor.

Let $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2)$ be the distribution of the variable Y above. Let $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}'_2$ be the distribution of the variables R_1, R_2, R'_2 above. We will show that $\mathbf{Y}$ is ε_0 close to being a block source. As either $\mathbf{X}_1$ or $\mathbf{X}_2$ is guaranteed to be uniform, $H_{\infty}(\mathbf{Y}_i) \geq \frac{n_i}{2}$. By the min-entropy chain rule Lemma 4.4, with probability $1 - \varepsilon_0$ over fixings of $\mathbf{Y}_1 = \alpha$, it holds that $H_{\infty}(\mathbf{Y}_2 \restriction \mathbf{Y}_1 = \alpha) \geq \frac{n_2}{2} - n_1 - \log(1/\varepsilon_0) = \frac{3n}{4} - \log(1/\varepsilon_0)$. We will add ε_0 to our total error and assume this property about $\mathbf{Y}$ from here on. By property of Ext_{GUV}, it holds that, $\left|\mathbf{R}_2 - \mathbf{U}_{|\mathbf{R}_2|}\right| \leq \varepsilon_0$. We will add ε_0

Algorithm 3: Ext (Output-light Somewhere-extractor)

Input: source $X = (X_1, X_2) \in \{0,1\}^n \times \{0,1\}^n$, seed $S \in \{0,1\}^d$

Let $\mathsf{Ext}_{GUV} : \{0,1\}^{7n/4} \times \{0,1\}^d \to \{0,1\}^{n/2-\log(1/\varepsilon_0)}$ be the GUV extractor from Theorem 4.8 instantiated for entropy $3n/4$ and error $\varepsilon_0 = \varepsilon/4$.

Let $U = X_1, V = X_2$.
Let $n_1 = \frac{n}{4}, n_2 = \frac{7n}{4}$.
Let $Y = (Y_1, Y_2)$ where $Y_1 = \left(U_{[1,n_1/2]}, V_{[1,n_1/2]}\right), Y_2 = \left(U_{[(n_1/2)+1,n]}, V_{[(n_1/2)+1,n]}\right)$.
Let $R_2 = \mathsf{Ext}_{GUV}(Y_2, S)$.
Let R_2' be a length $n/4$ prefix of R_2 with last bit set to 1. Let $R_1 \in \{0,1\}^{n/16} = R_2' \cdot Y_1$ where the operation is done over the finite field $\mathbb{F}_{2^{n/16}}^4$.
Output R_1.

to our total error and assume $\mathbf{R}_2$ is uniform from here on. So, $\mathbf{R}_2'$ is a distribution over $\{0,1\}^{n/4}$ with min entropy $\frac{n}{4} - 1$. As $\mathbf{Y}_1 \sim \{0,1\}^{n/4}$ is such that $H_\infty(\mathbf{Y}_1) \geq \frac{n}{8}$, by Theorem 4.10, it holds that $\left|\mathbf{R}_1 - \mathbf{U}_{|\mathbf{R}_1|}\right| \leq 2^{-n/32+1}$. As $\mathbf{Y}$ is a block source, for each fixing α of $\mathbf{Y}_1$, it holds that:

$$\left|\mathsf{Ext}_{GUV}(\mathbf{Y}_2, S) - \mathbf{U}_{|\mathbf{R}_2|}\right| \leq \varepsilon_0$$

Hence, it must be that

$$\left|(\mathbf{Y}_1, \mathsf{Ext}_{GUV}(\mathbf{Y}_2, S)) - (\mathbf{Y}_1, \mathbf{U}_{|\mathbf{R}_2|})\right| \leq \varepsilon_0$$

and thus,

$$\left|\mathbf{R}_1 - \mathbf{U}_{|\mathbf{R}_1|}\right| \leq 2\varepsilon_0 + 2^{-n/32+1} \leq 3\varepsilon_0,$$

using the fact that $\varepsilon \geq 2^{-c_0 n}$, for some small $c_0 > 0$. The total error of the extractor on input $\mathbf{X}$ is thus bounded by $4\varepsilon_0 = \varepsilon$, as desired.

We now prove that this extractor is indeed output-light. For every fixing of the output R_1 of $\mathbf{R}_1$, β of $\mathbf{Y}_2$ and the seed S, we can uniquely recover R_2'. Given $\frac{3n}{16}$ bits corresponding to first three out of the 4 intermediate outputs of the inner product, we can use R_1 to compute the fourth intermediate outer product and then use R_2' to invert each of the products and recover R_1. Thus for a fixed seed S and output R_1, there can be at most $2^{3n/16+7n/4} = 2^{31n/16}$ such $x \in \{0,1\}^{2n}$ so that $\mathsf{Ext}(x, s) = z$. As there are at most 2^d seeds, for a fixed output $R_1 \in \{0,1\}^{n/16}$, $\left|\left\{x \in \{0,1\}^{2n} : \exists y(\mathsf{Ext}(x, y)) = z\right\}\right| \leq 2^{2n-n/16-\log(n/\varepsilon)} = \frac{2^{2n-m}}{\mathrm{poly}(n,1/\varepsilon)} = \frac{2^{2n-m}}{\mathrm{poly}(m,1/\varepsilon)}$. □

A.2 Recursive condenser compositions

By composing the explicit condenser from Theorem A.1, we can get explicit condensers for other values of g and ℓ as well. We present an explicit computation of parameters for the case of uniform $(6,9)$-oNOSF sources in Appendix A.2.1 and sketch of the general recursive composition in Appendix A.2.2.

A.2.1 An explicit condenser for uniform $(6,9)$-oNOSF sources

We can take our condenser for uniform $(2,3)$-oNOSF sources even further to create a condenser for uniform $(6,9)$-oNOSF sources by nesting it within itself.

Theorem A.3. *There exists a constant $0 < c < 1$ such that for all $\varepsilon > 2^{-cn+2}$, we can explicitly construct a condenser $\mathsf{Cond} : (\{0,1\}^n)^{3^2} \to \{0,1\}^m$ where $m = \frac{n}{16^2}$ such that for any uniform $\left(\frac{2}{3}3^2, 2^2\right)$-oNOSF source $\mathbf{X}$, we have $H^{\varepsilon}_{\infty}(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$.*

Proof. To create a condenser for uniform $\left(\frac{2}{3}3^2, 2^2\right)$-oNOSF sources, we will apply our condenser from Theorem A.1 in a nested fashion. Recall that Theorem A.1 states that there exists a constant $0 < c_1 < 1$ such that for all $\varepsilon_1 > 2^{-c_1 n}$, we can explicitly construct a condenser $\mathsf{Cond}_1 : (\{0,1\}^n)^3 \to \{0,1\}^{m_1}$, where $m_1 = \frac{n}{16}$ such that for any uniform $(6,9)$-oNOSF source $\mathbf{X}$, $H^{\varepsilon}_{\infty}(\mathsf{Cond}_1(\mathbf{X})) \geq m_1 - O(\log(m_1/\varepsilon_1))$.

Let X be a uniform $(6,9)$-oNOSF source. We will apply Cond_1 on the first, second, and last third of $\mathbf{X}$ to get $\mathbf{Z}_1 = \mathsf{Cond}_1(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$, $\mathbf{Z}_2 = \mathsf{Cond}_1(\mathbf{X}_4, \mathbf{X}_5, \mathbf{X}_6)$, and $\mathbf{Z}_3 = \mathsf{Cond}_1(\mathbf{X}_7, \mathbf{X}_8, \mathbf{X}_9)$. To define $\mathsf{Cond}_2 : (\{0,1\}^n)^9 \to \{0,1\}^{m_2}$, we again apply Cond_1 to get $\mathsf{Cond}_2(\mathbf{X}) = \mathsf{Cond}_1(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)$. Now, we analyze the result of this construction.

We begin by noticing that the output length of Cond_2 is $m_2 = \frac{t_1}{16} = \frac{n}{16^2}$ since each of $\mathbf{Z}_1$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$ is on $\{0,1\}^{m_1}$. Next, because $\mathbf{X}$ is a uniform $(6,9)$-oNOSF source, it has at most 3 bad blocks. Consequently, if Cond_1 did not work for one of $\mathbf{Z}_1$, $\mathbf{Z}_2$, or $\mathbf{Z}_3$ (i.e., its conditions were not satisfied because 2 or 3 of its given blocks were bad), call this output $\mathbf{Z}_k$, then there is at most 1 bad block left over as inputs to Cond_1 for $\mathbf{Z}_i$ and $\mathbf{Z}_j$ where $i, j \neq k$ and $i \neq j$. Thus, the conditions for Cond_1 are met in the creation of $\mathbf{Z}_i$ and $\mathbf{Z}_j$ so $H^{\varepsilon_1}_{\infty}(\mathbf{Z}_i), H^{\varepsilon_1}_{\infty}(\mathbf{Z}_j) \geq m_1 - O(\log(m_1/\varepsilon_1))$.

By accumulating $2\varepsilon_1$ error, we can consider $\mathbf{Z}_i$ and $\mathbf{Z}_j$ as having min entropy at least $m_1 - O(\log(m_1/\varepsilon_1))$. Applying Lemma 4.7 allows us to consider $\mathbf{Z}_i$ and $\mathbf{Z}_j$ as having full min entropy in our application of $\mathsf{Cond}_1(\mathbf{Z})$ by accumulating $2 \cdot 2^{O(\log(m_1/\varepsilon_1)} = O(\mathrm{poly}(m_1/\varepsilon_1)) = \left(\frac{m_1}{\varepsilon_1}\right)^p =: \gamma$ error for some exponent $p \geq 1$. Finally, in our application of Theorem A.1 in $\mathsf{Cond}_1(\mathbf{Z})$, we set $\varepsilon_2 = \left(\frac{m_1}{\varepsilon_1}\right)^{-2p} \gamma + 2\varepsilon_1 = \left(\frac{m_1}{\varepsilon_1}\right)^{-p} + 2\varepsilon_1$. Using that $\varepsilon_1 > 2^{-c_1 n}$, we get that $\left(\frac{m_1}{\varepsilon_1}\right)^{-p} > \left(\frac{n}{16}\right)^{-p} 2^{-c_1 np}$. If we take $n > 16$ then we have $2^{-c_1 np} > \left(\frac{n}{16}\right)^{-p} 2^{-c_1 np}$. Thus, we require that $\varepsilon_2 = \left(\frac{m_1}{\varepsilon_1}\right)^{-p} + 2\varepsilon_1 > 2^{-c_1 np} + 2^{-c_1 n+1} > 2^{-c_1 n+2}$ since $2^{-c_1 np} \leq 2^{-c_1 n+1}$. Setting $\varepsilon > 2^{-c_1 n+2}$ and $c = c_1$ in the theorem statement gives us our desired result.

□

A.2.2 General recursive composition

At the expense of shorter output length and larger error, we can generalize our explicit recursive composition from Theorem A.3 to any odd ℓ. We give a proof sketch of this composition here. We first state a simple corollary from Theorem A.3.

Corollary 1. *There exists constant $0 < c_0 < 1$ such that for all $\varepsilon > 2^{-c_0 n}$, we can explicitly construct a condenser $\mathsf{Cond} : (\{0,1\}^n)^{\ell} \to \{0,1\}^m$, where $m = \frac{n}{16}$ such that for any uniform $(\ell - 1, \ell)$-oNOSF source $\mathbf{X}$ with $\ell \geq 3$, we have $H^{\varepsilon}_{\infty}(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$.*

Proof. We simply apply the condenser from Theorem A.3 to $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ and infer the result because at most one of $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ can be bad, so $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ is a uniform $(2,3)$-oNOSF source. □

Next, we give a sketch for what happens when we compose two condensers in a nested manner.

Lemma A.4. *For $i \in \{1,2\}$, say there exists a condenser $\mathsf{Cond}_i : (\{0,1\}^{n_i})^{\ell_i} \to \{0,1\}^{m_i}$ for uniform (g_i, ℓ_i)-oNOSF sources with $m_i = f_i(n_i)$, entropy gap $\Delta_i = O(\log(m_i/\varepsilon_i))$, and error*

$\varepsilon_i = 2^{-\Omega(n_i)}$. *Let* $b_i = \ell_i - g_i$. *Then there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *for any uniform* $(g = \ell - b, \ell)$*-oNOSF source* $\mathbf{X}$ *where* $\ell = \ell_1 \cdot \ell_2$, $b = (b_1+1)(b_2+1) - 1$, $m = \max(f_1(f_2(n)), f_2(f_1(n)))$, *and error* $\varepsilon = 2^{-\Omega(n)}$ *such that* $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$.

Proof. We will first consider defining Cond by nesting Cond_2 within Cond_1, although it will turn out that the number of bad blocks that Cond can handle is independent from the order that we choose to to nest Cond_1 and Cond_2.

Because $\mathbf{X}$ has $\ell = \ell_1 \cdot \ell_2$ blocks, we can split $\mathbf{X}$ up into ℓ_1 chunks $\mathbf{Y}_1, \ldots, \mathbf{Y}_{\ell_1}$ each with ℓ_2 blocks from $\mathbf{X}$ in it. Then, we apply Cond_2 to each of these chunks to get $\mathbf{Z}_j = \mathsf{Cond}_2(\mathbf{Y}_j)$ for $j \in \{1, \ldots, \ell_1\}$. Finally, we define $\mathsf{Cond}(\mathbf{X}) := \mathsf{Cond}_1(\mathbf{Z}_1, \ldots, \mathbf{Z}_{\ell_1})$. We claim that this construction gives us the desired result.

To compute the number of bad blocks b that Cond can handle, we will think adversarially as to the fewest number of blocks that are required to break our construction. To make the output of Cond_1 fail, we require that at least $b_1 + 1$ of $\mathbf{Z}_1, \ldots, \mathbf{Z}_{\ell_1}$ be bad. For a single $\mathbf{Z}_j$ to be bad — that is, for Cond_2 to fail — we require that at least $b_2 + 1$ of the blocks of $\mathbf{X}$ used for $\mathbf{Z}_j$ be bad. Thus, to make the output of Cond fail, we require at least $(b_1+1)(b_2+1)$ bad blocks in $\mathbf{X}$. Conversely, this means that Cond can handle at most $b = (b_1+1)(b_2+1) - 1$ bad blocks. In other words, Cond requires at least $g = \ell - b = \ell - (b_1+1)(b_2+1) + 1$ good blocks to succeed. Notice that this equation is symmetric in b_1 and b_2, demonstrating that the order of composition of Cond_1 and Cond_2 does not matter in computing b.

Next, we focus on the output length m. Each $\mathbf{Z}_j$ will be on $\{0,1\}^{f_2(n)}$, so $\mathsf{Cond}(\mathbf{X})$ is on $\{0,1\}^{f_1(f_2(n))}$. Since the order of composition of Cond_1 and Cond_2 does not matter for b, we can take the optimal order to get $m = \max(f_1(f_2(n)), f_2(f_1(n)))$.

Finally, to compute ε and our final min-entropy gap, we defer to the proof of Theorem A.3 since those computations follow in a similar manner. □

We remark that Theorem A.3 is an explicitly computed version of Lemma A.4 applied to the explicit uniform $(2,3)$-oNOSF source condenser from Theorem A.1.

To finish our generalization, we give a sketch for how one could recursively apply Lemma A.4. For the sake of succinctness, we define a useful prime counting function.

Definition A.5. *For any* $\ell \in \mathbb{N}$, *let* $\varphi(\ell)$ *be the total number of prime factors of* ℓ *counting multiplicity. That is, if* $\ell = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}$, *then* $\varphi(\ell) = \sum_{i=1}^n a_i$.[9] *We use this to define*

$$\Phi(\ell) := \begin{cases} \varphi(\ell) & \ell \text{ odd} \\ \varphi(\ell - 1) & \ell \text{ even} \end{cases}.$$

Essentially, $\Phi(\ell)$ returns $\varphi(\ell)$ if ℓ is odd and $\varphi(\ell - 1)$ if ℓ is even. We now state our main theorem.

Theorem A.6. *Let* $\ell \geq 3$.[10] *Then there exists an explicit condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *for any uniform* $\left(\ell - 2^{\Phi(\ell)} + 1, \ell\right)$*-oNOSF source* $\mathbf{X}$ *such that* $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ *where* $m = \frac{n}{16^{\Phi(\ell)}}$ *and* $\varepsilon = 2^{-\Omega(n)}$.

[9] In number theory, this is usually denoted by the prime Omega function $\Omega(\ell)$, but we don't this notation here to avoid confusion with asymptotics.

[10] Note that we consider ℓ as a constant and $\Phi_1(\ell) \leq \ell$, so we can consider $\Phi_1(\ell)$ as a constant as well.

Proof. Without loss of generality, we take ℓ to be odd. If ℓ is even, we truncate $\mathbf{X}$ to its first $\ell - 1$ blocks. Since $\ell - 1$ is odd, meaning that $\Phi(\ell - 1) = \varphi(\ell - 1)$, we can use our result for the odd case (which we prove below) to get that we can explicitly condense from uniform $\big((\ell-1) - 2^{\varphi(\ell-1)} + 1, \ell + 1\big)$-oNOSF sources. Thus, since we may be removing a good block when we truncate the last block of $\mathbf{X}$, this means that we can condense from uniform $\big(((\ell-1) - 2^{\varphi(\ell-1)} + 1) + 1, \ell\big)$-oNOSF sources which simplifies to uniform $\big(\ell - 2^{\Phi(\ell)} + 1, \ell\big)$-oNOSF sources, matching our claim.

Factor ℓ as $\ell = \ell_1 \cdots \ell_{\varphi(\ell)}$ and note that each factor is at least 3 since ℓ is odd. By Corollary 1, we know that there exists an explicit condenser Cond_i for uniform $(\ell_i - 1, \ell_i)$-oNOSF sources for $i \in [\varphi(\ell)]$. Let $b_i = \ell_i - 1$ and $b'_1 = b_1$. Nesting Cond_1 in Cond_2 by Lemma A.4 gives us a new explicit condenser for uniform $(\ell_1 \cdot \ell_2 - b'_2, \ell_1 \cdot \ell_2)$-oNOSF sources that can handle $b'_2 = (b_2 + 1)(b'_1 + 1) - 1$ bad blocks.

Repeatedly applying Lemma A.4 gives us explicit condensers for uniform $(\ell'_i - b'_i, \ell'_i)$-oNOSF sources where $\ell'_i = \ell_1 \cdot \ell_i$ and $b'_i = (b_i + 1)(b'_{i-1} + 1) - 1$. Taking $i = \varphi(\ell)$ then gives us our desired condenser Cond for $\ell = \ell'_{\varphi(\ell)}$ blocks that can handle at most $b'_{\varphi(\ell)} = 2^{\varphi(\ell)} - 1$ bad blocks, meaning Cond is a condenser for uniform $\big(\ell - 2^{\varphi(\ell)} + 1, \ell\big)$-oNOSF sources, as desired.

The output lengths $m = \frac{n}{16^{\varphi(\ell)}}$ of Cond follows because each application of Lemma A.4 divides the output length by 16 due to our construction in Corollary 1. Furthermore, the final error and entropy gap of Cond again follow similarly from the explicit computations in Theorem A.3. □

As we have done previously in Corollary 6.16, we can prepend our function that transforms low min-entropy oNOSF sources to uniform oNOSF sources from Lemma 6.6 to get a corollary of Theorem A.6 but for low min-entropy oNOSF sources. We do note that to use Lemma 6.6 explicitly in this way, we require using a two-source extractor from [CZ19, Li16] that has polynomial error which ultimately gives us polynomial instead of exponential error as we have in Theorem A.6.

Corollary 2. *Let $\ell > 3$. Then there exists an explicit condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ for any $\big(\ell - 2^{\Phi(\ell-1)} + 1, \ell, n, k\big)$-oNOSF source $\mathbf{X}$ with $k \geq \Omega(\log^C(n))$ for some large enough constant C, such that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \Omega(\mathrm{poly}(k))$ and $\varepsilon = 1/\Omega(\mathrm{poly}(k))$.*

Proof. We instantiate Lemma 6.6 with the explicit two-source extractor from [Li16] which achieves polynomially small error, has output length $\mathrm{poly}(k)$, and can handle min-entropy at least $k \geq \Omega(\log^C(n))$. Prepending this transformation to Theorem A.6 gives us our desired result. □

For a cleaner statement of Theorem A.6, we can truncate our input to a power of 3 instead.

Corollary 3. *Let $\ell \geq 3$ and take the unique $a, r \in \mathbb{N}$ such that $\ell = 3^a + r$ and $r < 2 \cdot 3^a$. Then there exists an explicit condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ for any uniform $(\ell - 2^a + 1, \ell)$-oNOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \frac{n}{16^a}$ and $\varepsilon = 2^{-\Omega(n)}$.*

Proof. Truncate $\mathbf{X}$ to its first 3^a blocks and use that as input to Theorem A.6 where $\Phi(3^a) = a$. This gives us a condensing possibility result for uniform $((\ell - r) + 2^a + 1, \ell - r)$-oNOSF sources. Since we may remove r good blocks when we truncate $\mathbf{X}$, our result holds for uniform $(((\ell - r) + 2^a + 1) + r, \ell)$-oNOSF sources, which simplifies to uniform $(\ell + 2^a + 1, \ell)$-oNOSF sources. □

If we only take powers of 3, then we get:

Corollary 4. *Let $\ell = 3^a$ for some $a \in \mathbb{N}$. Then there exists an explicit condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *for any uniform $(3^a - 2^a + 1, 3^a)$-oNOSF source $\mathbf{X}$ such that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \frac{n}{16^a}$ and $\varepsilon = 2^{-\Omega(n)}$.*

As always, we get similar versions for low min-entropy oNOSF sources. Here, we get them as corollaries from Corollary 2.

Corollary 5. *Let $\ell > 3$ and take the unique $a, r \in \mathbb{N}$ such that $\ell = (3^a+1)+r$ and $r < 2\cdot 3^a+1$. Then there exists an explicit condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *for any $(\ell - 2^a + 1, \ell, n, k)$-oNOSF source $\mathbf{X}$ with $k \geq \Omega(\log^C(n))$ for some large enough constant C, such that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \Omega(\text{poly}(k))$ and $\varepsilon = 2^{-\Omega(\text{poly}(k))}$.*

Proof. Truncate $\mathbf{X}$ to its first $3^a + 1$ blocks and use that as input to Corollary 2 where $\Phi((3^a+1)-1) = a$. This gives us a condensing possibility result for low min-entropy $((\ell - r) + 2^a + 1, \ell - r)$-oNOSF sources. Since we may remove r good blocks when we truncate $\mathbf{X}$, our result holds for low min-entropy $(((\ell - r) + 2^a + 1) + r, \ell)$-oNOSF sources, which simplifies to low min-entropy $(\ell + 2^a + 1, \ell)$-oNOSF sources. □

We can also restrict $r = 0$ to get an analogous result.

Corollary 6. *Let $\ell = 3^a + 1$ for some $a \in \mathbb{N}$. Then there exists an explicit condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *for any $(3^a - 2^a + 2, 3^a, n, k)$-oNOSF source $\mathbf{X}$ with $k \geq \Omega(\log^C(n))$ for some large enough constant C, such that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \Omega(\text{poly}(k))$ and $\varepsilon = 2^{-\Omega(\text{poly}(k))}$.*

B Extraction impossibility for rate 2/3 oNOSF sources

We end our appendix by showing that one cannot extract from rate 2/3 uniform oNOSF sources. Importantly, we note that this result is distinct from a similar result in [KN23] where the authors showed that extracting from rate 2/3 uniform NOSF sources is impossible. Since uniform oNOSF sources are a strict subset of the class of uniform NOSF sources, their impossibility result does not transfer to uniform oNOSF sources and we must prove our own. To do so, we first claim the case of uniform $(2,3)$-oNOSF sources.

Theorem B.1. *For any function $f : (\{0,1\}^n)^3 \to \{0,1\}$ there exists a uniform $(2,3)$-oNOSF source X such that $|f(\mathbf{X}) - \mathbf{U}_1| \geq 0.08$.*

Then our desired result follows as a corollary.

Corollary 7. *For any function $f : (\{0,1\}^n)^\ell \to \{0,1\}$ where ℓ is divisible by 3, there exists a uniform $\left(\frac{2}{3}\cdot\ell, \ell\right)$-oNOSF source X such that $|f(\mathbf{X}) - \mathbf{U}_1| \geq 0.08$.*

Proof. For the sake of contradiction, say there exists such an ℓ and function f such that $|f(\mathbf{X}) - \mathbf{U}_1| < 0.08$ for any uniform $\left(\frac{2}{3}\cdot\ell, \ell, n\right)$-oNOSF source X. Then if we let $\mathbf{X}$ be a uniform $\left(2, 3, \frac{n\ell}{3}\right)$-oNOSF source but consider it as a uniform $\left(\frac{2}{3}\cdot\ell, \ell, n\right)$-oNOSF source by splitting up each block into $\ell/3$ sub-blocks to get $3 \cdot \frac{\ell}{3} = \ell$ total blocks, we get that f is a extractor for uniform $\left(2, 3, \frac{n\ell}{3}\right)$-oNOSF sources, a contradiction to Theorem B.1. □

We now prove the main theorem.

Proof of Theorem B.1. To show that extraction is impossible, we will attempt to fix the output of f with constant probability over its inputs. We begin by classifying the points in the first two coordinates of f as follows.

$$S_0 = \{(x_1, x_2) \in [N]^2 \mid \forall x_3 \in [N],\ f(x_1, x_2, x_3) = 0\}$$
$$S_1 = \{(x_1, x_2) \in [N]^2 \mid \forall x_3 \in [N],\ f(x_1, x_2, x_3) = 1\}$$
$$S_{0,1} = \{(x_1, x_2) \in [N]^2 \mid \exists x_3, x_3' \in [N],\ f(x_1, x_2, x_3) = 0 \text{ and } f(x_1, x_2, x_3') = 1\}.$$

Note that we can write $S_{0,1} = [N]^2 \setminus (S_0 \cup S_1)$. In order, these are the sets of points in $\mathbf{X}_1$ and $\mathbf{X}_2$ that fix the output of f to 0, to 1, and that do not fix the output of f. We now take constants $0.5 \leq c_0, c_1 \leq 1$ and look at two cases that allow us to fix the output of f by putting an adversary in the third coordinate, $\mathbf{X}_3$.

Case 1. We have $|S_0| + |S_{0,1}| \geq c_0 N^2$. Here, we know that for $(x_1, x_2) \in S_0 \cup S_{0,1}$ there exists some x_3 such that $f(x_1, x_2, x_3) = 0$. Define $a(x_1, x_2)$ be this x_3 for $(x_1, x_2) \in S_0 \cup S_{0,1}$ and 0 otherwise. Consequently, if we let $\mathbf{X}_1$ and $\mathbf{X}_2$ be random and define our uniform $(2,3)$-SHELA source as $\mathbf{X} = \mathbf{X}_1, \mathbf{X}_2, a(\mathbf{X}_1, \mathbf{X}_2)$, then we have that $\Pr[f(\mathbf{X}) = 0] \geq c_0$. It follows that $|f(\mathbf{X}) - \mathbf{U}_1| \geq c_0 - \frac{1}{2}$.

Case 2. We have $|S_1| + |S_{0,1}| \geq c_0 N^2$. This case follows similarly since for $(x_1, x_2) \in S_1 \cup S_{0,1}$ there exists some x_3 such that $f(x_1, x_2, x_3) = 1$. Therefore, we can define an adversary $a(x_1, x_2)$ such that when $\mathbf{X} = \mathbf{X}_1, \mathbf{X}_2, a(\mathbf{X}_1, \mathbf{X}_2)$ with $\mathbf{X}_1$ and $\mathbf{X}_2$ uniform we have $|f(\mathbf{X}) - \mathbf{U}_1| \geq c_0 - \frac{1}{2}$.

Case 3. We are in neither of the previous two cases. Thus, because $|S_0| + |S_1| + |S_{0,1}| = N^2$ we have that $(1 - c_0)N^2 < |S_0|, |S_1| < c_0 N^2$ and $(2c_0 - 1)N^2 < |S_{0,1}| < c_0 N^2$. To proceed, we will set up two sub-cases in which we either make $\mathbf{X}_1$ our bad block or $\mathbf{X}_2$ our bad block.

Consider the bipartite graph $H = (U, V)$ with $|U| = N$ left vertices representing the values of $\mathbf{X}_1$ and $|V| = N$ vertices representing the values of $\mathbf{X}_2$. We place an edge (u, v) with label t if $(u, v) \in S_t$ and do not place an edge otherwise. Consequently, the number of edges E in H is at least $E = |S_0| + |S_1| \geq 2(1 - c_0)N^2$. For any $u \in U$, define its normalized degree (counting edges with either label) as $d_u = \deg(u)/N$. We then see that $\mathbb{E}_{u \sim U}[d_u] = E/|U| \geq 2(1 - c_0)$. To split into our two sub-cases, we will consider the set of heavy vertices $U_H = \{u \in U \mid d_u > c_1\}$ in U. By Claim 4.1, we get that $\Pr_{u \in U}[d_u > c_1] \geq \frac{\mathbb{E}_u[d_u] - c_1}{1 - c_1} \geq \frac{2(1-c_0) - c_1}{1 - c_1} =: c_2$, meaning that $|U_H| \geq c_2 N$.

- **Case i.** For all $u \in U_H$ we have $u \in S_0 \cap S_1$ (i.e., u has at least one edge labeled with a 0 and another with a 1). this means that for any $u \in U_H$ there exists an $x_2 \in [N]$ such that for all $x_3 \in [N]$ we have that $f(u, x_2, x_3) = 0$. Let $a(x_1)$ be defined as outputting this x_2 that fixes f to 0 for $x_1 \in U_H$ and to be 0 otherwise. Defining $\mathbf{X} = \mathbf{X}_1, a(\mathbf{X}_1), \mathbf{X}_3$ with $\mathbf{X}_1$ and $\mathbf{X}_3$ uniform gives us a uniform $(2,3)$-SHELA source for which $\Pr[f(\mathbf{X}) = 0] \geq |U_H|/N \geq c_2$, so $|f(\mathbf{X}) - U_1| \geq c_2 - \frac{1}{2}$.
- **Case ii.** There exists a $u \in U_H$ such that $u \notin S_0 \cap S_1$. Without loss of generality, say $u \in S_0$, so all of the edges of u are labeled 0, meaning that for all $x_2 \in \mathcal{N}(u)$ and any $x_3 \in [N]$ we have that $f(u, x_2, x_3) = 0$. Because $u \in U_H$, we have that $d_u > c_1$, so defining $\mathbf{X} = u, \mathbf{X}_2, \mathbf{X}_3$ with $\mathbf{X}_2$ and $\mathbf{X}_3$ uniform gives us that $\Pr[f(\mathbf{X}) = 0] \geq c_1$. Therefore, $|f(\mathbf{X}) - U_1| \geq c_1 - \frac{1}{2}$.

Combining all of our cases and recalling that $c_2 = \frac{2(1-c_0) - c_1}{1 - c_1}$, we have that we can construct a uniform $(2,3)$-oNOSF source $\mathbf{X}$ such that $|f(\mathbf{X}) - U_1| \geq \varepsilon$ where $\varepsilon = \min(c_0, c_1, c_2) - \frac{1}{2}$. Setting $c_0 = 0.58$ and $c_1 = 0.6$ gives us $\varepsilon = 0.58 - 0.5 = 0.08$. □