Electronic Colloquium on Computational Complexity, Report No. 171 (2024)

Condensing Against Online Adversaries

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

We investigate the task of deterministically condensing randomness from Online Non-Oblivious Symbol Fixing (oNOSF) sources, a natural model of defective random sources for which it is known that extraction is impossible [AORSV, EUROCRYPT'20]. A (g, ℓ)-oNOSF source is a sequence of ℓ blocks $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_\ell) \sim (\{0,1\}^n)^\ell$, where at least g of the blocks are *good* (are independent and have some min-entropy), and the remaining *bad* blocks are controlled by an *online adversary* where each bad block can be arbitrarily correlated with any block that appears before it.

The existence of condensers was recently studied in [CGR, FOCS'24]. They proved condensing impossibility results for various values of g and ℓ, and they showed the existence of condensers matching the impossibility results in the special case when n is extremely large compared to ℓ (i.e., the setting of few blocks of large length).

In this work, we make significant progress on proving the existence of condensers with strong parameters in almost all parameter regimes, even when n is a large enough constant and ℓ is growing. This almost resolves the question of the existence of condensers for oNOSF sources, except when n is a small constant.

As our next result, we construct the first explicit condensers for oNOSF sources and achieve parameters that match the existential results of [CGR, FOCS'24]. We also obtain a much improved construction for transforming low-entropy oNOSF sources (where the good blocks only have min-entropy, as opposed to being uniform) into uniform oNOSF sources.

We find interesting connections and applications of our results on condensers to collective coin flipping and collective sampling, problems that are well-studied in fault-tolerant distributed computing. We use our condensers to provide very simple protocols for these problems.

Finally, to understand the case of small n, we focus on $n = 1$ which corresponds to online non-oblivious bit-fixing (oNOBF) sources. We introduce and initiate a systematic study of a new, natural notion of the influence of Boolean functions, which we call *online influence*, and believe is of independent interest. Using tools from Boolean Fourier analysis, we establish tight bounds on the total online influence of Boolean functions, which imply extraction lower bounds. Several problems remain open regarding this new measure of influence; progress on these will lead to improved extractors and condensers for oNOBF sources or further strengthen our lower bounds.

*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.

ISSN 1433-8092

Contents

1 Introduction

Randomness is extremely useful in computation with wide-ranging applications in algorithm design, cryptography, distributed computing protocols, machine learning, error-correcting codes, and much more [MR95, Vad12]. Most of these applications require access to high quality randomness. However in a lot of settings, especially arising in practice, algorithms only have access to low quality source of randomness. This motivates the notion of *condensers*: functions that transform weak random sources into strong random sources that are of *better quality*.

In this line of work, the standard way of measuring the amount of randomness is using min-entropy. Formally, for a source (distribution) $\mathbf{X}$ with support Ω, define its min-entropy as $H_\infty(\mathbf{X}) = \min_{x\in\Omega} \log_2(1/\Pr[\mathbf{X} = x])$. We will also need the notion of smooth min-entropy, which measures how close a distribution is to having high entropy. Formally, for a source $\mathbf{X}$, its smooth min-entropy with parameter ε is defined as $H^\varepsilon_\infty(\mathbf{X}) = \max_{\mathbf{Y}:|\mathbf{X}-\mathbf{Y}|\le\varepsilon}\{H_\infty(\mathbf{Y})\}$, where $|\cdot|$ denotes the statistical distance (Definition 3.1).

With this, we are ready to formally define condensers:

Definition 1.1. *A function* $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^m$ *is a* $(k_{in}, k_{out}, \varepsilon)$*-condenser for a family of distributions* $\mathcal{X}$ *if for all* $\mathbf{X} \in \mathcal{X}$ *with* $\mathbf{X} \sim \{0,1\}^n$ *and* $H_\infty(\mathbf{X}) \ge k_{in}$*, we have that* $H^\varepsilon_\infty(\mathbf{X}) \ge k_{out}$*.*

We say $\frac{k_{in}}{n}$ *is the input entropy rate,* $\frac{k_{out}}{m}$ *is the output entropy rate, and* $m - k_{out}$ *is the entropy gap of* Cond*.*

The task of the condenser is to make the output entropy rate as high as possible compared to the input entropy rate, or, in other words, to make the output distribution more "condensed". Related to this, it is also desirable to have as small entropy gap as possible. Notice that if the entropy gap is 0, the output distribution is ε-close to the uniform distribution. Such condensers with entropy gap 0 are known as *randomness extractors*—a topic that has been extensively studied in theoretical computer science.

When $\mathcal{X}$ is the family of all distributions, it is folklore that no non-trivial condensing is possible.[1] So, we additionally assume that $\mathcal{X}$ is a structured family of sources.[2] Since extractors are the highest quality condensers, a significant amount of work has focused on constructing extractors for interesting family of sources, such as: sources generated by small circuits, two independent sources, algebraically generated sources, sources generated by small space sources, and many more [TV00, CZ19, DGW09, KZ07].

However, for many natural family of sources, one can provably show that no extractor can exist. In such situations, one can still hope to show that high quality condensers exist. We note that condensers (and sources with high min-entropy rate) are very useful: the condensed distribution can be used to efficiently simulate randomized algorithms with small overhead, perform one-shot simulations for randomized protocols, cryptography and interactive proofs, and much more. [DPW14] showed these condensers are equivalent to 'unpredictability extractors' that can simulate cryptographic protocols against biased distinguishers. For details on these applications and more, see [AORSV20, DMOZ23, CGR24].

In this work, we focus on one natural family of sources where it is known that extraction is impossible. The family we consider are known as online non-oblivious symbol fixing sources

[1] Assuming $m \le n$ (wlog this holds since $|\mathsf{Cond}(\{0,1\}^n)| \le 2^n$), $m - k_{out} \ge (n - k_{in}) - \log(1/(1-\varepsilon))$ and hence the output entropy rate cannot be more than the input entropy rate without incurring extremely large error (> 0.999).

[2] A different route, that has been widely studied, is to assume access to a short independent seed. In this work, we will limit ourselves to the *seedless setting*.

(oNOSF sources).[3] Formally:

Definition 1.2. *A (g, ℓ, n, k)-oNOSF source $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_\ell)$ is such that each block $\mathbf{X}_i$ is over $\{0,1\}^n$, g of the blocks are are independent sources with min-entropy k ("good blocks"), and each "bad block" is an arbitrary function of the blocks with an index smaller than it. When $k = n$, we will call such sources uniform (g, ℓ, n)-oNOSF sources.*

These sources are inspired by real-time randomness generation settings such as in blockchains. There, each subsequent block is random or controlled by an adversary. Since these sources are generated in real time, a bad block can only be a function of the blocks that have appeared so far, and it is reasonable to assume that the good blocks contain entropy and are independent. Further, there are natural cryptographic settings, such as creating a Common Reference String, that are widely used in various cryptographic protocols where oNOSF source sources naturally arise (see [AORSV20] for a discussion).

The remainder of our introduction is structured as follows. We give an overview of previous work in Section 1.1 before presenting our main existential and explicit condenser results in Section 1.2. In Section 1.3, we show how our results have implications for collective coin flipping and sampling protocols. Next, we introduce our notion of online influence and related results in Section 1.4. We end by defining a local version of oNOSF sources and give explicit extractors for them in Section 1.5.

1.1 Previous work

The study of condensers for oNOSF sources was initiated by [AORSV20].[4] Their results include the following:

- It is impossible to extract from uniform oNOSF sources (even when the fraction of good blocks is an arbitrary constant).
- An explicit transformation from $(g, \ell, n, 0.9n)$-oNOSF source into a source over $(\{0,1\}^{O(n)})^{\ell-1}$ where $g-1$ of the blocks are uniform and independent.
- An explicit transformation from $(g, \ell, n, 0.1n)$-oNOSF source into a source over $(\{0,1\}^{O(n)})^{100\ell}$ where $g-1$ of the blocks are uniform and independent.

Even though the output entropy rate is only slightly more than the input-entropy rate in the second result and smaller in the third result, the fact that a lot of the blocks are truly uniform is very useful, and they find interesting cryptographic applications of these somewhere-extractors.

oNOSF sources were further studied by [CGR24], where they obtained the following results:

- When $n \ge k \ge \ell$, there exist functions that can transform a (g, ℓ, n, k)-oNOSF source into a uniform $(g-1, \ell-1, O(k/\ell))$-oNOSF source (this function can be made explicit with slightly worse dependence on output length).

[3] These sources are in contrast to non-oblivious symbol fixing (NOSF) sources where bad blocks can be arbitrary functions of all the good blocks. These sources were introduced in [CGHFRS85] with applications in leakage-resilient cryptography, and have been well-studied.

[4] In [AORSV20], these sources were called SHELA (Somewhere Honest Entropic Look Ahead) sources.

- When $n \geq 2^{\omega(\ell)}$, there exists condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^{m=O(n\cdot\ell/g)}$ such that for any uniform (g,ℓ,n)-oNOSF source $\mathbf{X}$, $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq \frac{1}{\lceil \ell/g \rceil} \cdot m - O(\log(n/\varepsilon))$. Their result is not explicit.

- It is impossible to condense from uniform (g,ℓ,n)-oNOSF sources with output entropy rate more than $\frac{1}{\lceil \ell/g \rceil}$.

We also mention a related family of sources, namely adversarial Chor-Goldreich sources. uniform oNOSF sources can be seen as a special case of adversarial Chor-Goldreich sources where the good blocks are uniform. Constructing condensers where the output entropy rate is g/ℓ for adversarial Chor-Goldreich sources is already a challenging task, although such condensers in various parameter regimes have been recently constructed [DMOZ23, GLZ24]. [DMOZ24] recently constructed condensers for a related more general model.

1.2 New condenser constructions

Previous works only showed the existence of condensers for oNOSF sources when $n \geq 2^{\omega(\ell)}$. We vastly improve on this result in two ways. First, we show that for almost all values of n, ℓ, even when n is a small constant, excellent condensers exist. Second, we provide explicit condensers for oNOSF sources when $n \geq 2^{\omega(\ell)}$. We also obtain much better transformation from low-entropy oNOSF sources to uniform oNOSF sources that work even when $k \ll \ell$. These results show condensers always exist, except when n is a very small constant (such as $n = 1$). To further our understanding of this case, we initiate the study of *online influence* of Boolean functions, a natural generalization of influence that captures the one-sided nature of our online adversary. We also discover surprising connections between condensers for oNOSF sources and protocols for natural problems in distributed computing, such as collective coin flipping and collective sampling. We now discuss our result in details below.

1.2.1 Existential condensers

We show how to condense from uniform (g,ℓ,n)-oNOSF sources for almost all settings of ℓ and n when $g \geq 0.51\ell$. In particular, we show:

Theorem 1 (Informal version of Theorem 4.1)**.** *For all ℓ, ε where $\ell \geq O(\log(1/\varepsilon))$, and $n = 10^4$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform $(0.51\ell,\ell,n)$-oNOSF source $\mathbf{X}$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq 0.99m$ where $m = O(\ell + \log(1/\varepsilon))$. Furthermore, when $n = \omega(1)$, the output entropy rate becomes $1 - o(1)$.*

This is tight since [CGR24] showed it is impossible to condense uniform $(0.5\ell,\ell,n)$-oNOSF sources beyond output entropy rate 0.5.

Using our new results regarding transforming oNOSF sources to uniform oNOSF sources, we also obtain condensers for $(0.51\ell,\ell,n,k)$-oNOSF sources when $n \geq \mathrm{poly}(\log(\ell))$,

Theorem 2. *For all ℓ, n, ε where $n = \mathrm{poly}(\log(\ell/\varepsilon)), k = O(\log(\ell/\varepsilon))$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any $(0.51\ell,\ell,n,k)$-oNOSF source $\mathbf{X}$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq m - O(m/\log(m)) - O(\log(1/\varepsilon))$ where $m = \Omega(k)$.*

We can also extend our result to condense from uniform (g, ℓ, n)-oNOSF sources for all g, ℓ and constant n where the output entropy rate is $1/\lfloor \ell/g \rfloor - 0.001$. This is tight since [CGR24] showed it is impossible to condense such sources beyond output entropy rate $1/\lfloor \ell/g \rfloor$.

Previously, [CGR24] showed how to existentially condense from uniform (g, ℓ, n)-oNOSF sources when $g \geq 0.51\ell$, provided $n \geq 2^{\omega(\ell)}$. As n gets smaller, condensing becomes harder since a uniform (g, ℓ, n)-oNOSF source is also a uniform $(g \cdot n/1000, \ell \cdot n/1000, 1000)$-oNOSF source. Hence, we greatly improve the parameters while using different and much simpler techniques.

1.2.2 Explicit condensers

We construct the first explicit condensers for oNOSF sources. Our explicit condenser construction achieves the same parameters as the existential condenser construction of [CGR24]. We show how to explicitly condense from uniform (g, ℓ, n)-oNOSF sources when $n \geq 2^{\omega(\ell)}$ and $g \geq 0.5\ell + 1$. We state the results for constant ℓ since that is cleaner:

Theorem 3 (Informal version of Theorem 5.1)**.** *For all n, ε and constant ℓ, there exists an* explicit *condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform $(0.5\ell + 1, \ell)$-oNOSF source $\mathbf{X}$, we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \Omega(n)$.*

Just like earlier, since condensing when $g = 0.5\ell$ is impossible, this result is also tight. Using our new results regarding transforming oNOSF sources to uniform oNOSF sources, we also obtain explicit condensers for $(0.51\ell, \ell, n, k)$-oNOSF sources for the same parameter regime:

Corollary 1.3 (Corollary 5.2, simplified)**.** *For all ℓ, n, ε with constant ℓ and $n \geq O(\log(1/\varepsilon))$, there exists an* explicit *condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any $(0.5\ell + 2, \ell, n, \mathrm{poly}(\log n))$-oNOSF source $\mathbf{X}$, we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ where $m = \mathrm{poly}(\log n)$.*

Similar to earlier, we can also extend our result to explicitly condense from uniform (g, ℓ, n)-oNOSF sources in the same parameter regime so that the output entropy rate is $1/\lfloor \ell/g \rfloor - o(1)$. Just like earlier, this is tight as well.

Previously, [CGR24] showed how to existentially condense from uniform (g, ℓ, n)-oNOSF sources in this parameter regime. However, they relied on the existence of a very strong pseudorandom object: "output-light" low-error two-source extractors. Such extractors, even without the output-lightness requirement, are extremely hard to construct and it is a major open problem to obtain such extractors. We are able to make this condenser explicit by building up on their ideas, making interesting observations regarding oNOSF sources, and stitching them together so that the base pseudorandom object we rely on are seeded extractors that we know how to explicitly construct with near optimal parameters.

1.2.3 Transforming low-entropy oNOSF sources to uniform oNOSF sources

We show how to existentially, as well as explicitly, with a slight loss in parameters, transform (g, ℓ, n, k)-oNOSF sources into uniform $(0.99g, \ell - 1, n)$-oNOSF sources. Formally, we show:

Theorem 1.4 (Informal version of Theorem 6.1)**.** *For all ℓ, n, k, ε where $n = \mathrm{poly}(\log(\ell)), k = O(\log(\ell/\varepsilon))$, there exists a function f such that f transforms $(0.51\ell, \ell, n, k)$-oNOSF sources into uniform $(0.509\ell, \ell, m)$-oNOSF sources with error ε where $m = O(k)$.*

Our construction can also be made explicit with slightly worse dependence on m and ε. See Corollary 6.4 for the full tradeoff.

Previously, [CGR24] provided such a transformation only for $n \geq k \geq \Omega(\ell)$. Hence, our transformation makes a major improvement on their parameters. Such an improvement allows us to obtain better condensers for low-entropy oNOSF sources in the regime $n = \mathrm{poly}(\log(\ell/\varepsilon))$ (see Theorem 2).

1.3 Application to collective coin flipping and collective sampling

Condensing from oNOSF sources can be viewed as a special case of coin flipping and collective sampling protocols in the full information model that arise in fault-tolerant distributed computing.

1.3.1 Background

Say there are ℓ players who have a common broadcast channel and want to jointly perform a task such as collectively flipping a coin. Some b players out of them are "bad" and want to deter the task. We assume the bad players are computationally unbounded so cryptographic primitives are of no use. We further assume that each player has private access to uniform randomness. [BL89] initiated the study of this model and aptly termed this task as "collective coin flipping."

The simplest way to collectively flip a coin would be for all the players to initially agree on a function $f : \{0,1\}^\ell \to \{0,1\}$, then synchronously broadcast one random bit r_i, and to finally agree on the output being $f(r_1, \ldots, r_\ell)$. However, synchronizing broadcasts is hard, and it could be that the bad players set their output as function of the bits of the good players. [KKL88] showed that no function f can handle more than $O\left(\frac{\ell}{\log \ell}\right)$ corruptions.

One way to allow for more corruptions (almost linear) amongst players is to consider "protocols" that allow more rounds of communication. In particular, a protocol can be thought of as a tree where each vertex represents a "round" where in every round the following happens: all good players sends their bits, then all bad players send their bits as a function of the bits of the good players, and they jointly compute a function of these bits. Depending on the outcome of the function, everyone branches on one branch in this tree. Furthermore, every leaf is labelled with final outcomes (say 0 or 1) and, once you reach a leaf, that is the outcome that everybody agrees on. [GGL98] initiated the study of protocols where the outcomes are from a larger range and where the bad players are trying to minimize the largest probability of any outcome. They called this problem "collective sampling."

1.3.2 Known results

[BL89] showed that for protocols with outcomes $\{0,1\}$, b bad players can always ensure some outcome occurs with probability at least $\frac{1}{2} + \frac{b}{2\ell}$. [AN93] first constructed a protocol that can handle $O(n)$ corruptions. Follow-up works tried to reduce the number of rounds in this protocol where, in some settings, players were allowed to send more bits per round [RZ01, Fei99].

[GGL98] showed that for all collective sampling protocols and all outcomes, there exists a way for b bad players to coordinate and ensure that an outcome that happens without corruption with probability p, now happens with probability $p^{1-(b/n)} \geq p\left(1 + \frac{b}{n}\log(1/p)\right)$. For further results and bounds, see [Dod06].

1.3.3 Connection to oNOSF sources

The problem of extracting or condensing from oNOSF sources can be seen as special cases or variants of collective coin flipping and collective sampling that provide very simple protocols. For instance, suppose one has an extractor or condenser f for uniform (g, ℓ, n)-oNOSF sources. Then, consider a protocol where all ℓ players take turns and output n random bits. The agreed final outcome is f applied on these ℓn bits. This leads to protocols that are structurally much simpler since players don't have to carefully compute whose turn is it to go in various rounds and can obliviously prepare for their turn.

The above protocol can also be viewed as a relaxed version of a 1-round protocol where instead of everyone providing their output asynchronously, they take turns and provide outputs one after another in a simple sequential manner.

1.3.4 Previous results interpreted in oNOSF source context

Previous impossibility results can be interpreted in the context of extracting / condensing from uniform oNOSF sources. For instance, collective coin flipping impossibility results of [BL89] imply extraction impossibility results for uniform (g, ℓ, n)-oNOSF sources when $n = 1$. In particular, they imply:

Corollary 1.5. *There does not exist an ε-extractor for uniform $(g, \ell, 1)$-oNOSF sources where $\varepsilon = \frac{b}{2\ell}$.*

Similarly, we observe that the notion of collective sampling is equivalent to 0-error condensing. Hence, lower bounds of [GGL98] imply zero-error condensing lower bounds for uniform (g, ℓ, n)-oNOSF sources when $n = 1$. Formally:

Corollary 1.6. *There does not exist a condenser $\mathsf{Cond} : \{0,1\}^\ell \to \{0,1\}^m$ for uniform $(g, \ell, 1)$-oNOSF sources that can guarantee output smooth min-entropy (with parameter $\varepsilon = 0$) more than $k = \frac{g}{\ell} \cdot m$.*

1.3.5 ε-collective sampling

Since collective sampling lower bounds show that for any protocol, 0-error condensing beyond rate g/ℓ is impossible, one can naturally ask whether condensing with small error ε is possible: We call this problem ε-collective sampling where the goal is to output a distribution which is ε-close to a distribution where every output has small probability.

Interpreted this way, this is exactly what protocols arising out of our condensers for uniform oNOSF sources provide: Using Theorem 1, when each player has access to 10^4 random bits, there exists a simple protocol that can handle 0.49ℓ corrupt players such that the players can collectively sample a distribution over $m = O(\ell)$ bits which is $2^{-\Omega(\ell)}$-close to having entropy $0.99m$. As far as we are aware, such a protocol cannot be implied using any other previous protocol (a lot of previous protocols are obtained through *leader election* protocols which do not seem useful here since the leader has access to only constant number of bits).

We similarly obtain explicit protocols using Theorem 3 for the case when each player has access to $n \geq 2^{\omega(\ell)}$ bits.

1.3.6 Collective coin flipping and sampling with weak random sources

A natural extension to collective coin flipping and sampling in the full information model is when all players only have access to weak source of randomness (that are independent from each other) instead of true uniform randomness. This question was first studied by [GSV05]. [KLRZ08] used network extractor protocol to transform weak random sources of each player into independent private random sources. This way, after using the network extraction protocol, players can follow the usual collective coin flipping / sampling protocol. [GSZ21] improved the network extraction protocol using two-source non-malleable extractors.

Using our (g, ℓ, n, k)-oNOSF source condensers, we obtain alternative, simple ε-collective sampling protocols in the setting where players have access to weak sources of randomness. We obtain such an existential protocol using Theorem 2, and explicit protocol using Corollary 1.3.

1.4 Online influence

When n is a large enough constant, we showed that great condensers do exist for oNOSF sources. We now study the case of small n. In particular, we focus on the case of $n = 1$. We call such uniform $(g, \ell, 1)$-oNOSF sources as (g, ℓ)-oNOBF sources; oNOBF stands for online non-oblivious bit-fixing sources. We ask what is the exact tradeoff between g, ℓ, and ε for extracting / condensing from oNOBF sources.

Towards this, we introduce a new notion of influence for Boolean functions which we call online influence.

Definition 1.7 (Online influence)**.** *For a function $f : \{0,1\}^\ell \to \{0,1\}$, the* online influence *of the i-th bit is*

$$\mathbf{oI}_i[f] = \mathop{\mathbb{E}}_{x \sim \mathbf{U}_{i-1}} \left[\left| \mathop{\mathbb{E}}_{y \sim \mathbf{U}_{\ell-i}} [f(x,1,y)] - \mathop{\mathbb{E}}_{y \sim \mathbf{U}_{\ell-i}} [f(x,0,y)] \right| \right]$$

and the total online influence *is*

$$\mathbf{oI}[f] = \sum_{i=1}^{\ell} \mathbf{oI}_i[f].$$

We first observe that for monotone functions, online influence equals the usual notion of influence (see Lemma 7.4 for a proof).

Next, we show a Poincaré style inequality for total online influence:

Theorem 4 (Theorem 7.5, restated)**.** *For any $f : \{0,1\}^\ell \to \{0,1\}$, we have* $\mathrm{Var}(f) \le \mathbf{oI}[f] \le \sqrt{\ell \, \mathrm{Var}(f)}$.

These inequalities are tight, with tightness exhibited by PARITY and MAJORITY respectively (see Example 7.6 for more details).

We then ask the natural question that for a function f, what is the maximum online influence out of all n bits? For the usual notion of influence, this question was resolved by the famous theorem of [KKL88] who showed there always exists a bit with influence at least $\mathrm{Var}(f) \cdot \Omega\left(\frac{\log \ell}{\ell}\right)$. We show that surprisingly, there exists a balanced function, namely the address function, where every bit has online influence at most $O\left(\frac{1}{\ell}\right)$ (see Lemma 7.12 for a proof). This provides a separation between the usual notion of influence and online influence. Additionally, in Lemma 7.13, we show that the address function is an extractor for uniform $(\ell-1,\ell)$-oNOSF sources with error $1/\ell$.

Lastly, we show that using this notion of online influence, we can obtain the following extraction lower bound:

Theorem 1.8 (Informal version of Theorem 7.17)**.** *For $\varepsilon < 0.01$, there do not exist extractors for $(0.98\ell,\ell)$-oNOBF sources with error at most ε.*

A similar extraction lower bound was shown in [AORSV20] using different techniques.

1.5 Extracting from local oNOSF sources

A natural variation on our definition of oNOSF sources is to consider the case where the adversary cannot remember the value of every good block in the past; rather, it can only remember the value of the most recent s blocks. Arguably, this is a realistic assumption in the setting of many short blocks, where it could be difficult to introduce long range correlation.

Definition 1.9 (Local oNOSF sources)**.** *We call a (g,ℓ,n,k)-oNOSF source $\mathbf{X}=(\mathbf{X}_1,\dots,\mathbf{X}_\ell)$ an* s-local (g,ℓ,n,k)-oNOSF source *if each bad block $\mathbf{X}_i$ can only depend on at most s blocks $\mathbf{X}_{i-s},\dots,\mathbf{X}_{i-1}$ that come before it.*

Interestingly, weakening the adversary in this way converts our oNOSF source into a small-space source. These sources were first studied by [KRVZ11] and we refer the reader to them for a definition and background. Since the adversarial blocks of an s-local (g,ℓ,n,k)-oNOSF source can only depend on the binary string of length at most sn to its left, we easily see that an s-local (g,ℓ,n,k)-oNOSF source is samplable by a space-sn source.

Using recent explicit extractors for low-space sources provided by [CL22, Li23] and the fact that a (g,ℓ,n,k)-oNOSF source has entropy at least gk, we get the following extraction result for these local online sources.

Theorem 1.10 (Using the explicit extractor of [CL22])**.** *There exists a universal constant C such that for every s and $k \geq \frac{2sn+\log^C(n\ell)}{g}$ there is an explicit extractor $\mathsf{Ext} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ with error $\varepsilon=(n\ell)^{-\Omega(1)}$ and output length $m=(gk-2sn)^{\Omega(1)}$ for every s-local (g,ℓ,n,k)-oNOSF source.*

A similar result with slightly better entropy requirement, but constant error, can be obtained using the small-space extractor from [Li23].

2 Proof overview

Our proof overview begins by outlining our new existential results for condensers in Section 2.1 that is able to handle even constant block length. Next, we present our explicit condenser results in Section 2.2 before discussing of our low-entropy to uniform oNOSF source conversion in Section 2.3. Finally, we end with a summary of the key ideas behind our results regarding online influence in Section 2.4.

2.1 Existence of oNOSF condensers for all ℓ and n

Here we sketch the proof of Theorem 1. This result states that when $g = 0.51\ell$ and $n = 1000$, there exists a condenser Cond for uniform (g, ℓ, n)-oNOSF sources so that the output entropy rate is 0.99, the number of output bits is $m = O(\ell + \log(1/\varepsilon))$, and the error of the condenser is ε where $\varepsilon \leq 2^{-\Omega(\ell)}$ is arbitrary.

Our construction uses amazing seeded condensers (see Definition 3.3) with $1 \cdot \log(1/\varepsilon)$ dependence on seed length. We slightly modify our source and then apply such seeded condenser. Here is a proof sketch:

Proof sketch for Theorem 1. Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_\ell)$ be such a source. Let $\mathbf{Y}_1 \sim (\{0,1\}^n)^{0.5\ell}$ be the source obtained by concatenating the first 0.5ℓ blocks of $\mathbf{X}$. Since 0.51ℓ blocks are good, there exist at least 0.01ℓ uniform blocks in $\mathbf{Y}_1$. We treat $\mathbf{Y}_1$ as a single distribution over $n\ell$ bits with min-entropy $\geq 0.01\ell n$.

Let $\mathbf{Y}_2 \sim \{0,1\}^{0.5\ell}$ be the source obtained by concatenating 1 bit from each of the last 0.5ℓ blocks of $\mathbf{X}$. Once again, since 0.51ℓ blocks are good, there exist at least 0.01ℓ uniform bits in $\mathbf{Y}_1$.

We will use the following seeded condenser:

Theorem 2.1 (Theorem 4.8, simplified)**.** *For all d, ε such that $d \geq \log(\ell n/\varepsilon) + O(1)$, there exists a seeded condenser $\mathsf{sCond} : \{0,1\}^{0.5\ell n} \times \{0,1\}^d \to \{0,1\}^m$ s.t. for all $\mathbf{X} \sim \{0,1\}^{0.5\ell n}$ with $H_\infty(\mathbf{X}) \geq 0.01\ell n$, we have $H_\infty^\varepsilon(\mathsf{sCond}(\mathbf{X}, \mathbf{U}_d)) \geq 0.01\ell n + d$ where $m = 0.01\ell n + d + \log(1/\varepsilon) + O(1)$.*

Our condenser Cond will output $\mathsf{sCond}(\mathbf{Y}_1, \mathbf{Y}_2)$. Observe that not only is $\mathbf{Y}_2$ not uniform, there could be as many as 0.49ℓ "bad bits" in $\mathbf{Y}_2$ that can depend on $\mathbf{Y}_1$. To remedy this, we use the well known fact that the behavior of such adversarial $\mathbf{Y}_2$ cannot be far worse than the behavior if $\mathbf{Y}_2$ were uniform. In particular, say if $\mathbf{Y}_2$ were uniform then the output entropy and error are k and ε respectively. Then for the actual $\mathbf{Y}_2$, the output entropy will be $k - 0.49\ell$ and error will be $\varepsilon \cdot 2^{0.49\ell}$. See Lemma 4.9 for the formal statement.

For us, it means the following: let $\varepsilon_{\mathsf{sCond}}, k_{\mathsf{sCond}}$ be such that $H_\infty^{\varepsilon_{\mathsf{sCond}}}(\mathsf{sCond}(\mathbf{Y}_1, \mathbf{U}_{0.5\ell})) \geq k_{\mathsf{sCond}}$. Then, it must be that $H_\infty^{2^{0.49\ell} \cdot \varepsilon_{\mathsf{sCond}}}(\mathsf{sCond}(\mathbf{Y}_1, \mathbf{Y}_2)) \geq k_{\mathsf{sCond}} - 0.49\ell$. So, for our final error to be some ε, we need to have $\varepsilon_{\mathsf{sCond}} = \varepsilon \cdot 2^{-0.49\ell}$. For seeded condensers to exist, we need $0.5\ell \geq \log(\ell n/\varepsilon_{\mathsf{sCond}}) + O(1)$ and we check that such an inequality can indeed be satisfied if $\varepsilon \geq 2^{-0.01\ell}$.

Hence, we finally obtain that our seeded condenser will output $0.01\ell n + O(\ell)$ bits and will have output entropy $m - \Delta$ where $\Delta = O(\ell)$. Hence, if n is a large enough constant, our output entropy rate, $\frac{m-\Delta}{m}$, will be ≥ 0.99 as desired.

Remark 2.2. *Here we crucially used the fact that there exist seeded condensers with seed length dependence $1 \cdot \log(1/\varepsilon)$. Currently, we do not have explicit constructions with this dependence. We also couldn't have used a seeded extractor since for them, the seed length dependence is $2 \cdot \log(1/\varepsilon)$. For that to work, we would need to assume $g \geq 0.76\ell$.*

□

2.2 Explicit condensers for uniform oNOSF sources

Here we sketch the proof of Theorem 3: we construct explicit condensers for uniform $(0.5\ell + 1, \ell, n)$-oNOSF sources where ℓ is a constant and n is arbitrarily growing.

Our construction will be similar to that of the existential construction of condensers from [CGR24]. We will use the online nature of these sources to make more observations that will allow us to obtain an explicit construction of such condensers using explicit seeded extractors (see Definition 3.4 for a definition) as our primitive:

Proof sketch for Theorem 3. Let $\mathbf{X}$ be such a source. Let's review the [CGR24] construction at a high level: they take the first $\lceil \ell/2 \rceil$ blocks and treat them as a single entity. From the remaining blocks, they take first few bits of each of the blocks with the number of such bits geometrically decreasing per block and concatenate them to obtain a second entity. They then pass these two sources to an "output-light" two-source extractor to obtain their final output.

We first split each block in $\mathbf{X}$ into two parts of equal sizes. The resultant source is uniform $(\ell+1, 2\ell, n/2)$-oNOSF source. We call this source $\mathbf{X}$ as well since we are just re-interpreting $\mathbf{X}$ as this source. This simple trick turns out to be very useful since it allows us to only focus on the situation where the number of blocks is an even number.

Remark 2.3. *The construction of [CGR24] had to introduce the notion of "output-lightness" to deal with the case of odd number of blocks. For instance, say $\ell = 5$. Then, their first entity is obtained by concatenating the first 3 blocks and second entity by taking careful number of bits from the remaining 2 blocks. To handle scenarios where the first 3 blocks were uniform and last 2 were bad, the output-lightness property was imposed on two-source extractors, something which we do not know how to explicitly construct.*

Just reducing to even cases is not enough since the construction of [CGR24] required low-error two-source extractors with excellent parameters, and we do not know how to explicitly construct them. We bypass this requirement by further exploiting the fact that our adversary is online.

Let $\mathbf{W} \sim (\{0,1\}^n)^\ell$ be the concatenation of the first ℓ blocks of $\mathbf{X}$. Let $\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell$ be the sources obtained by carefully choosing the first few bits from each of the blocks $\mathbf{X}_{\ell+1}, \ldots, \mathbf{X}_{2\ell}$. Our final construction will be the parity of the outputs of seeded extractors applied with source $\mathbf{W}$ and seeds $\mathbf{Y}_i$. More formally, we output

$$\bigoplus_{i=1}^{\ell} \mathsf{sExt}_i(\mathbf{W}, \mathbf{Y}_i)$$

where sExt_i is any explicit near optimal seeded extractor (such as the extractor from Theorem 3.5).

Since the number of blocks, 2ℓ, is even and number of good blocks is $\ell+1$, both $\mathbf{W}$ and $\mathbf{Y}$ will obtain some bits from a good block. This means, $H_\infty(\mathbf{W}) \geq n$ and that there exists $j \in [\ell]$ such that $\mathbf{Y}_j$ is uniform. We now condition on fixing blocks $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$. Since these blocks can depend on $\mathbf{W}$, $\mathbf{W}$ will lose some small amounts of entropy (the amount will be very small since these blocks are tiny compared to the amount entropy in $\mathbf{W}$). Moreover, since the adversary is online, $\mathbf{Y}_j$ remains uniform even after doing this conditioning. We now view our construction as

$$g(\mathbf{W}) \oplus \bigoplus_{i=j}^{\ell} \mathsf{sExt}_i(\mathbf{W}, \mathbf{Y}_i)$$

where g is the fixed function obtained by fixing $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$.

We now compare two scenarios: (1) Where all $\mathbf{Y}_j, \ldots, \mathbf{Y}_\ell$ are uniform (2) Only $\mathbf{Y}_j$ is uniform and $\mathbf{Y}_{j+1}, \ldots, \mathbf{Y}_\ell$ are arbitrarily controlled by an adversary and can even depend on $\mathbf{W}$:

In the first scenario, we further condition on fixing $\mathbf{Y}_{j+1}, \ldots, \mathbf{Y}_\ell$. Since in this scenario these are independent and random, $\mathbf{W}$ retains the same entropy and $\mathbf{Y}_j$ remains uniform. So our overall output is of the form $h(\mathbf{W}) \oplus \mathsf{sExt}_j(\mathbf{W}, \mathbf{Y}_j)$ for some fixed function h. We condition on fixing output $h(\mathbf{W})$. Since $m \ll H_\infty(\mathbf{W})$, we infer that $\mathbf{W}$ still has lots of entropy when we do this fixing. So, the output is just $z \oplus \mathsf{sExt}_j(\mathbf{W}, \mathbf{Y}_j)$ where z is a fixed string, and hence the output distribution is uniform.

The second scenario is more realistic and, in the worst case, this is what can actually happen. We then use the result that if an adversary controls few bits in the input distribution, then they cannot make the output of the condenser too bad (see Lemma 4.9 for full statement). With this, and by carefully choosing geometrically decreasing lengths of $\mathbf{Y}_i$ to help control the error, we indeed obtain that the output will be condensed. □

2.3 Converting low-entropy oNOSF sources to uniform oNOSF sources

They key part of our proof for condensing from low-entropy oNOSF sources is a transformation from low-entropy oNOSF sources to uniform oNOSF sources. Here, we sketch the proof for our transformation in Theorem 1.4 and compare it to that of [CGR24]. Both these transformations rely on two-source extractors (see Definition 3.6 for definition) as a basic primitive.

Given a (g, ℓ, n, k)-oNOSF source $\mathbf{X} = \mathbf{X}_1, \dots, \mathbf{X}_\ell$, [CGR24] uses excellent existential two-source extractors (such as from Lemma 6.6) to define output blocks $\mathbf{O}_i = \mathsf{2Ext}(\mathbf{X}_1 \circ \cdots \circ \mathbf{X}_{i-1}, \mathbf{X}_i)$ for $i \in \{2, \dots, \ell\}$ and define their transformation as $f(\mathbf{X}) = \mathbf{O}_2, \dots, \mathbf{O}_\ell$. They show that $\mathbf{O}_i$ is a good block if: (1) $\mathbf{X}_i$ is a good block and (2) at least one block amongst $\mathbf{X}_1, \dots, \mathbf{X}_{i-1}$ is a good block. They showed that such a good block will be uniform and independent of the blocks $\mathbf{O}_2, \dots, \mathbf{O}_{i-1}$ and argued there will be $g-1$ such good output blocks. This indeed shows their output is a uniform $(g-1, \ell-1, m)$-oNOSF source. However, each of their output blocks has length $m = O\left(\frac{k}{\ell}\right) \leq O\left(\frac{n}{\ell}\right)$, and so they were not able to handle the case of $n = o(\ell)$. We improve on their construction by using a "sliding window" based technique to obtain a much better transformation that can even handle $n = \mathrm{poly}(\log(\ell))$.

Theorem 2.4 (Theorem 6.1 restated)**.** *Let $d, g, g_{out}, \ell, n, m, k, \varepsilon$ be such that $g_{out} \leq g - \frac{\ell-g+2}{d}, n \geq k \geq \log(nd-k) + md + 2\log(2g_{out}/\varepsilon)$. Then, there exists 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 uniform $(g_{out}, \ell-1, m)$-oNOSF source $\mathbf{Y}$ for which $|f(\mathbf{X}) - \mathbf{Y}| \leq \varepsilon$.*

The parameter d in our theorem statement above is the width of our sliding window. When we set $d = \ell$ we recover the analysis of [CGR24]. The true advantage of our transformation emerges when d is very small compared to ℓ. For instance, when $g = 0.51\ell, n = \mathrm{poly}(\log(\ell))$ and $k = \mathrm{poly}(\log(\ell))$, we set d to be a large constant and conclude that the output distribution is a uniform $(0.509\ell, \ell, \mathrm{poly}(\log(\ell))$-NOSF source.

Proof sketch of Theorem 2.4. Define $\mathbf{O}_i = \mathsf{2Ext}(\mathbf{X}_{i-d} \circ \cdots \circ \mathbf{X}_{i-1}, \mathbf{X}_i)$. We call $\mathbf{O}_i$ to be a good output block when $\mathbf{X}_i$ is good and there's at least one good block amongst $\{\mathbf{X}_{i-d}, \dots, \mathbf{X}_{i-1}\}$.

We first compute the number of good output blocks g_{out}. Let $j_1, \dots, j_g$ be the indices of the good input blocks in $\mathbf{X}$ and $d_i = j_{i+1} - j_i$ be the gap between the i-th good block and the next $(i+1)$-th good block. If the gap d_i is at most d, then $\mathbf{O}_{i+1}$ must be a good output block. So, g_{out} is the number of i such that $d_i \leq d$. Since $g \geq 0.51\ell$, such large gaps can't appear too often and we can calculate that $g_{out} = g - \frac{\ell-g+2}{d}$ as desired.

Next, we show that the good output blocks are indeed uniform conditioned on all previous output blocks. With this, we will obtain that the output distribution will be uniform $(g_{out}, \ell-1, m)$-oNOSF source as desired. Let i be the index of a good output block. We want to show that $\mathbf{O}_i$ is uniform conditioned on $\mathbf{O}_1, \ldots, \mathbf{O}_{i-1}$. To do this, we first observe that any input block contributes to at most $d+1$ good output blocks. This means that$(\mathbf{X}_{i-d} \circ \cdots \circ \mathbf{X}_{i-1})$, which has min-entropy at least k, loses at most $d \cdot m$ min-entropy conditioned on fixing $\mathbf{O}_1, \ldots, \mathbf{O}_{i-1}$. Moreover, $\mathbf{X}_i$ still remains uniform and independent of $(\mathbf{X}_{i-d} \circ \cdots \circ \mathbf{X}_{i-1})$ when fixing these previous output blocks. Hence, the output of the two-source extractor will indeed be uniform as desired. □

We can make Theorem 2.4 explicit by using the explicit two-source extractors of Theorem 6.7 at a slight cost of dependence on m and ε as seen in Corollary 6.4.

2.4 Online influence

In this subsection, we provide a brief overview of our results regarding online influence and contrast it with the established notion of influence for Boolean functions.

A Poincaré inequality and extractor lower bounds One fundamental inequality of regular influence is that of the Poincaré inequality stating that $\mathrm{Var}(f) \leq \mathbf{I}[f]$. We prove a similar result for online influence.

Theorem 2.5 (Theorem 7.5 restated)**.** *For any* $f : \{0,1\}^\ell \to \{0,1\}$*, we have* $\mathrm{Var}(f) \leq \mathbf{oI}[f] \leq \sqrt{\ell \, \mathrm{Var}(f)}$.

It is not hard to derive extractor lower bounds for oNOBF sources from the above result. The high level idea is to collect bits with high online influence, which is guaranteed by the above theorem (using an averaging argument) to form a *coalition of coordinates* that has enough online influence to bias the claimed extractor. We refer the reader to Theorem 7.17 for more details.

The proof of Theorem 2.5 is based on techniques from the Fourier analysis of Boolean functions.[5] The following key result that implies Theorem 2.5 in a straightforward way.

Lemma 2.6 (Lemma 7.7 restated)**.** *For any* $f : \{0,1\}^\ell \to \{0,1\}$ *and* $i \in [\ell]$, $\mathbf{oI}_i(f)^2 \leq \sum_{\substack{S \subseteq [i] \\ S \ni i}} \widehat{f}(S)^2 \leq \mathbf{oI}_i(f)$.

The above bound is established using the following Fourier analytic characterization of online influence.

Claim 2.7 (Claim 7.8 restated)**.** *For any* $f : \{0,1\}^\ell \to \{0,1\}$*, we can write the online influence of its* i*-th bit as*

$$\mathbf{oI}_i[f] = \mathop{\mathbb{E}}_{x \sim \mathbf{U}_{i-1}} \left[\left| \sum_{\substack{T \subseteq [i] \\ T \ni i}} \widehat{f}(T) \chi_{T \setminus \{i\}}(x) \right| \right].$$

The proof of the above result mostly standard Fourier analytic computation and we refer the reader to Section 7 for more details.

[5] We give a very brief recap of necessary notions from Fourier analysis of Boolean functions in Section 7.2.

Influence vs Online Influence It is not hard to see that for all $i \in [\ell]$, $\mathbf{oI}_i[f] \leq \mathbf{I}_i[f]$, with equality always holding for $i = \ell$ as an adversarial online bit in the last index can see every good bit. Moreover, we observe that for monotone functions, the notion of online influence is equivalent to that of regular influence, demonstrating that any separation between the two notions must come from non-monotone functions.

We exactly exhibit such a separation via the non-monotone address function $\mathrm{Addr}_\ell : \{0,1\}^{\log \ell + \ell} \to \{0,1\}$ which considers its first $\log \ell$ bits as an index in $\{1, \ldots, \ell\}$ and then outputs the value of the chosen index. It is easy to show (as we do in Lemma 7.12) that the first $\log \ell$ bits of Addr_ℓ have no online influence, while the remaining bits have online influence of $O\left(\frac{1}{\ell}\right)$. This is in contrast to the well known result of [KKL88] showing that, for a balanced function such as Addr_ℓ, there must exist a bit with influence at least $\Omega\left(\frac{\log \ell}{\ell}\right)$.

2.5 Organization

In the remainder of our paper, we give some preliminaries in Section 3 before moving on to our core results. Section 4 details our proofs for the existence of seedless condensers for oNOSF sources for all regimes of ℓ and n, while Section 5 provides proofs for our explicit constructions of condensers. Next, Section 6 shows how to handle converting low-entropy oNOSF sources to uniform oNOSF source for a broader range of parameters. In Section 7 we introduce the notion of online influence and use it to provide an extraction lower bound for oNOBF sources. We conclude with a list of some open questions in Section 8.

3 Preliminaries

In this section we give some basic background and facts used throughout our paper. We use boldfaced font to indicate a random variable such as $\mathbf{X}$. Often we will use $\circ$ or , to indicate concatenation of blocks. So if $\mathbf{X}_1 \sim \{0,1\}^n$ and $\mathbf{X}_2 \sim \{0,1\}^n$, then $\mathbf{X}_1, \mathbf{X}_2$ will be the concatenated random variable over $\{0,1\}^{2n}$. We will use the notation $[n]$ as shorthand for $\{1, \ldots, n\}$. Also all logs in this paper will have base 2 unless stated otherwise.

3.1 Basic probability notions

We measure the distance between two distributions via statistical distance:

Definition 3.1 (Statistical Distance)**.** *For any two distributions $\mathbf{X}, \mathbf{Y}$ over Ω, we define the statistical distance or total-variation distance (TV) distance as:*

$$|\mathbf{X} - \mathbf{Y}| = \max_{S \subset \Omega} |\Pr[\mathbf{X} \in S] - \Pr[\mathbf{Y} \in S]| = \frac{1}{2} \sum_{s \in \Omega} |\Pr[\mathbf{X} = s] - \Pr[\mathbf{Y} = s]|$$

We use the notation $\mathbf{X} \approx_\varepsilon \mathbf{Y}$. to denote the fact that $|\mathbf{X} - \mathbf{Y}| \leq \varepsilon$.

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

Lemma 3.2 (Min-entropy chain rule)**.** *For any random variables $\mathbf{X} \sim X$ and $\mathbf{Y} \sim Y$ and $\varepsilon > 0$,*

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

3.2 Condensers and extractors

Recall the definition of a seeded condenser.

Definition 3.3. *A* $(k_{in}, k_{out}, \varepsilon)$*-seeded condenser* $\mathsf{sCond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *satisfies the following: for every source* $\mathbf{X} \sim \{0,1\}^n$ *with* $H_\infty(\mathbf{X}) \geq k_{in}$*, and* $\mathbf{Y} = \mathbf{U}_d$*,*

$$H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X}, \mathbf{Y})) \geq k_{out}.$$

Here, d *is called the* seed length *of* sCond*.*

Seeded extractor is the special case of seeded condenser where $k_{out} = m$. We here record the full definition for completeness sake:

Definition 3.4. *A* (k, ε)*-seeded extractor* $\mathsf{sExt} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *satisfies the following: for every source* $\mathbf{X} \sim \{0,1\}^n$ *with* $H_\infty(\mathbf{X}) \geq k$*, and* $\mathbf{Y} = \mathbf{U}_d$*,*

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

Here, d *is called the* seed length *of* sExt*.* sExt *is called strong if*

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

We will use the following near optimal explicit construction of seeded extractors:

Theorem 3.5 (Theorem 1.5 in [GUV09])**.** *For all constant* $0 < \alpha < 1$*, there exists a constant* C *such that for 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* $d = C \log(n/\varepsilon)$ *and* $m \geq (1-\alpha)k$*.*

Next, we recall the definition of two-source extractors.

Definition 3.6. *A function* $\mathsf{2Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a* (k_1, k_2, ε)-two-source extractor *if for every source* $\mathbf{X}_1 \sim \{0,1\}^{n_1}$ *with* $H_\infty(\mathbf{X}_1) \geq k_1$ *and* $\mathbf{X}_2 \sim \{0,1\}^{n_2}$ *with* $H_\infty(\mathbf{X}_2) \geq k_2$ *where* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are independent of each other, we have*

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

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

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

4 Existence of condensers for all values of ℓ, n

We will show that there exist condensers for uniform (g, ℓ, n)-oNOSF sources for almost all settings of ℓ, n, provided $g > 0.5\ell$. Observe that a uniform (g, ℓ, n)-oNOSF source is also a uniform $(g \cdot s, \ell \cdot s, n/s)$-oNOSF source by simply dividing up all blocks into s parts. This implies that as n becomes smaller (relative to ℓ), it gets harder to condense with the hardest case being $n = 1$. Our condenser will also be handle the case of $n = O(1)$ and ℓ arbitrarily growing:

Theorem 4.1 (Simplified version of Corollary 4.7). *For all $g, \ell, n, \varepsilon, \delta$ where $g = 0.51\ell$, and $0.01\ell n \geq 2\log(\ell n/2\varepsilon) + O(1)$, there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform (g, ℓ, n)-oNOSF source $\mathbf{X}$, we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - \Delta$ where $m = 0.005\ell n + 200(\ell + \log(\ell n/2\varepsilon)) + O(1)$ and $\Delta = 200(\ell + \log(\ell n/2\varepsilon)) + O(1)$.*

Note that when n is a large enough constant, $m \geq 100\Delta$ and hence, the output entropy rate is at least 0.99.

In fact, we obtain a general result for all values of n, ℓ and when $g = 0.5\ell + e$ where $e \in \mathbb{N}$ is arbitrary. See Lemma 4.4 for the full tradeoff; to get slightly better parameters for small n, see Corollary 4.6.

We combine the above condenser for uniform oNOSF sources with the transformation for low-entropy oNOSF sources to uniform oNOSF sources from Corollary 6.2 to obtain the following condenser for low-entropy oNOSF sources:

Corollary 4.2. *Let $g, \ell, n, m, k, \varepsilon$ be such that $g = 0.51\ell, n = \mathrm{poly}(\log(\ell/\varepsilon)), k = \Omega(\log(\ell/\varepsilon)), m = \Omega(\ell \log(\ell/\varepsilon))$. Then, we can construct condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any (g, ℓ, n, k)-oNOSF source $\mathbf{X}$, we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - \Delta$ where $\Delta = O(\ell + \log(1/\varepsilon))$.*

Remark 4.3. *Previous condensers from [CGR24] could only show that condensers exist for uniform oNOSF sources when $\ell = o(\log n)$. They relied on existence of low-error two source extractors equipped with an additional "regularity" property. Our constructions are much simpler, recover all their results with even better parameters, and work for all values of n and ℓ, including the hardest case of $n = O(1)$.*

We provide our general construction of condensers in Section 4.1. To do that, we will require another type of condenser for two uniform oNOSF sources where the bad bits of the second block are allowed to depend on the bits of the first block. We provide this construction in Section 4.2.

4.1 Constructing condensers for uniform oNOSF sources

In this subsection, we will construct the following general condenser for uniform oNOSF sources:

Lemma 4.4 (General uniform oNOSF source condensing). *For all $g, \ell, n, \varepsilon, e$ where $g \geq (\ell/2) + e$, and $en \geq 2\log(\ell n/2\varepsilon) + O(1)$, there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform (g, ℓ, n)-oNOSF source $\mathbf{X}$, we have $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq m - \Delta$ where $m = \frac{en}{2} + (2\ell - e)\left\lceil \frac{\log(\ell n/2\varepsilon) + O(1)}{e} \right\rceil + \log(1/\varepsilon) + O(1)$ and $\Delta = (2\ell - 2e)\left\lceil \frac{\log(\ell n/2\varepsilon) + O(1)}{e} \right\rceil + \log(1/\varepsilon) + O(1)$.*

To do this, we will use a condenser for two distinct uniform oNOSF sources where one source can depend on the other:

Lemma 4.5. *For all $g, \ell, n_x, n_y, \varepsilon$ where $n_x \geq n_y$ and $gn_y \geq \log(\ell n_x/\varepsilon) + O(1)$, there exists a condenser* $\mathsf{Cond} : (\{0,1\}^{n_x})^\ell \times (\{0,1\}^{n_y})^\ell \to \{0,1\}^m$ *such that: For any uniform (g, ℓ, n_x)-oNOSF source $\mathbf{X}$ and uniform (g, ℓ, n_y)-oNOSF source $\mathbf{Y}$ with the additional property that bad blocks in $\mathbf{Y}$ can depend on $\mathbf{X}$ as well, we have that $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X}, \mathbf{Y})) \geq m - \Delta$ where $m = gn_x + (2\ell - g)n_y + \log(1/\varepsilon) + O(1)$ and $\Delta = (2\ell - 2g)n_y + \log(1/\varepsilon) + O(1)$.*

We construct this condenser in Section 4.2. Using this, our main general condenser can be constructed as follows:

Proof of Lemma 4.4. We split each block in $\mathbf{X}$ into 2 parts to obtain a uniform $(2g, 2\ell, n/2)$-oNOSF source. We call this resultant source $\mathbf{X}$ as well since it is the same distribution, just viewed differently. Let $\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_\ell)$ and where for $1 \le i \le \ell$, $\mathbf{U}_i = \mathbf{X}_i$. Let $\mathbf{V} = (\mathbf{V}_1, \ldots, \mathbf{V}_\ell)$ where for $1 \le i \le \ell$, we define $\mathbf{V}_i$ to be prefix of length n_v of $\mathbf{X}_{\ell+i}$ where $n_v = \left\lceil \frac{\log(\ell n/2\varepsilon)+O(1)}{e} \right\rceil$.

We observe that $\mathbf{U}$ is a uniform $(e, \ell, n/2)$-oNOSF source and $\mathbf{V}$ is a uniform (e, ℓ, n_v)-oNOSF source where bad bits in $\mathbf{V}$ can depend on $\mathbf{U}$ and the good bits in both sources are independent. We now define our condenser Cond to be the condenser from Lemma 4.5 applied to sources $\mathbf{U}, \mathbf{V}$. Hence, we will have that $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{U}, \mathbf{V})) \ge m - \Delta$ where $m = en/2 + (2\ell - e)n_y + \log(1/\varepsilon) + O(1)$ and $\Delta = (2\ell - 2e)n_y + \log(1/\varepsilon) + O(1)$ as desired. □

Our first corollary will apply to the regime that his the hardest to condense from, namely when n is very small compared to ℓ, even when $n = O(1)$ and ℓ is arbitrarily growing:

Corollary 4.6 (Small n)**.** *For all $g, \ell, n, \varepsilon, \delta$ where $g \ge (0.5 + \delta)\ell, \varepsilon \ge 2^{-\delta\ell + O(1)}$, and $n \le 2^{\delta\ell/2}$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ, n)-oNOSF source $\mathbf{X}$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \ge m - \Delta$ where $m = \delta\ell n/2 + (2-\delta)\ell + \log(1/\varepsilon) + O(1)$ and $\Delta = (2-\delta)\ell + \log(1/\varepsilon) + O(1)$.*

Proof. We observe that $\left\lceil \frac{\log(\ell n/2\varepsilon)+O(1)}{e} \right\rceil = 1$ and directly apply Lemma 4.4. □

We also obtain the following general tradeoff for larger n that may be growing with ℓ or even when $\ell = O(1)$ and n growing alone (this applies to all n but is most interesting when n is large since Corollary 4.6 provides better tradeoff for small n).

Corollary 4.7 (Larger n)**.** *For all $g, \ell, n, \varepsilon, \delta$ where $g \ge (0.5+\delta)\ell$, and $\delta\ell n \ge 2\log(\ell n/2\varepsilon) + O(1)$, there exists a condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ such that for any uniform (g, ℓ, n)-oNOSF source $\mathbf{X}$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \ge m - \Delta$ where $m = \frac{\delta\ell n}{2} + (2/\delta - 1)(\log(\ell n/2\varepsilon) + O(1)) + (2 - \delta)\ell + \log(1/\varepsilon) + O(1)$ and $\Delta = (2/\delta - 1)(\log(\ell n/2\varepsilon) + O(1)) + 2(2-\delta)\ell + \log(1/\varepsilon) + O(1)$.*

Proof. We observe that $\left\lceil \frac{\log(\ell n/2\varepsilon)+O(1)}{e} \right\rceil \le 1 + \frac{\log(\ell n/2\varepsilon)+O(1)}{e}$ and apply that to the condenser from Lemma 4.4. □

4.2 Condenser for two uniform oNOSF sources

In this subsection, we will prove Lemma 4.5. To construct the claimed condenser, we will use the following folklore result regarding existence of excellent seeded condensers (e.g., see Corollary 3 of [GLZ24]).

Theorem 4.8. *For all n, k, d, ε such that $d \ge \log(n/\varepsilon) + O(1)$, there exists a seeded condenser $\mathsf{sCond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that for all $\mathbf{X} \sim \{0,1\}^n$ with $H_\infty(\mathbf{X}) = k$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \ge k + d$ where $m = k + d + \log(1/\varepsilon) + O(1)$.*

We will also use the following result from [CGR24] that states an adversary can't make things too bad if it controls very few bits. We note that similar lemmas have been useful in previous construction of condensers [BCDT19, BGM22, GLZ24]:

Lemma 4.9 (Lemma 6.18 in [CGR24])**.** *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* $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq k$*. Let* $G \subset [n]$ *with* $|G| = n - b$ *be arbitrary. 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,* $H_\infty^{\varepsilon'}(f(\mathbf{X}')) \geq k - b$ *where* $\varepsilon' = \varepsilon \cdot 2^b$*.*

With this, we are ready to provide the construction of condensers for two uniform oNOSF sources:

Proof of Lemma 4.5. Let $\mathsf{sCond} : (\{0,1\}^{n_x})^\ell \times (\{0,1\}^{n_y})^\ell \to \{0,1\}^m$ be lossless condenser guaranteed from Theorem 4.8 with $\varepsilon_{\mathsf{sCond}} = \varepsilon \cdot 2^{-(\ell-g)n_y}$. We define $\mathsf{Cond}(x,y) = \mathsf{sCond}(x,y)$.

Let $\mathbf{O}_{unif} = \mathsf{Cond}(\mathbf{X}, \mathbf{U}_{\ell n_y})$ and $\mathbf{O}_{adv} = \mathsf{Cond}(\mathbf{X}, \mathbf{Y})$. We argue that $\mathbf{O}_{unif}$ will be highly condensed and since the adversary controls so few bits in $\mathbf{Y}$, $\mathbf{O}_{adv}$ will be condensed as well.

We first see that by the property of the seeded condenser, $H_\infty^{\varepsilon_{\mathsf{sCond}}}(\mathbf{O}_{unif}) \geq gn_x + \ell n_y$. Next we observe that $\mathbf{O}_{adv}$ can be obtained from $\mathbf{O}_{unif}$ by an adversary controlling $b = (\ell - g)n_y$ bits from $(\mathbf{X}, \mathbf{U}_{\ell n_y})$ to obtain $(\mathbf{X}, \mathbf{Y})$ and considering the output of sCond. We apply Lemma 4.9 which allows us to compare output entropy in such scenarios and obtain that

$$H_\infty^{\varepsilon_{\mathsf{sCond}} \cdot 2^b}(\mathbf{O}_{adv}) \geq H_\infty^{\varepsilon}(\mathbf{O}_{unif}) - b \geq (gn_x + \ell n_y) - ((\ell - g)n_y) = m - \Delta.$$

As $\varepsilon_{\mathsf{sCond}} \cdot 2^b = \varepsilon$, we indeed have that $H_\infty^\varepsilon(\mathbf{O}_{adv}) \geq m - \Delta$ as desired. □

5 Explicit Condensers for uniform oNOSF sources

In this section, we will prove the following main result regarding condensing from uniform oNOSF sources, matching the existential condenser parameters of [CGR24]. We state this for constant ℓ but our condenser can handle any $\ell = o(\log n)$.

Theorem 5.1 (Clean version of Theorem 5.5)**.** *For constant* g, ℓ *where* $g > \ell/2$*, and all* n, ε*, there exists an* explicit *condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform* (g,ℓ)*-oNOSF source* $\mathbf{X}$*, we have* $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq m - O(\log(m/\varepsilon))$ *where* $m = \Omega(n)$*.*

Using the transformation of low-entropy oNOSF sources to uniform oNOSF sources from Corollary 6.4 by setting $d = \ell$ (for this parameter regime, such a transformation can also be obtained using results from [CGR24]), we get an explicit condenser for low-entropy oNOSF sources:

Corollary 5.2. *For constant* g, ℓ *where* $g > \ell/2 + 1$*, and all* n, k, ε *with* $k \geq \mathrm{poly}(\log(n)) + \log(n/\varepsilon) + O(1)$*,* $\varepsilon \geq n^{-\Omega(1)}$*, there exists an* explicit *condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any* (g, ℓ, n, k)*-oNOSF source* $\mathbf{X}$*, we have* $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) = m - O(\log(m/\varepsilon))$ *where* $m = k^{\Omega(1)}$*.*

We now use a reduction from Lemma 5.12 in [CGR24]. The reduction shows how to use condensers for $g > \ell/2$ to construct condensers for all g, ℓ that condense up to rate $1/\lfloor \ell/g \rfloor$.[6] Using this, we construct explicit condensers for all (g, ℓ). Particularly, we get the following:

[6] The statement of the lemma in [CGR24] does not explicitly state it as a reduction but such a reduction easily follows from the proof of the lemma.

Corollary 5.3. *For any constant g and ℓ and all n, ε, there exists an* explicit *condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any uniform (g,ℓ)-oNOSF source $\mathbf{X}$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq 1/\lfloor \ell/g \rfloor \cdot m - O(\log(m/\varepsilon))$ where $m = \Omega(n)$.*

We again use the transformation of low-entropy oNOSF sources to uniform oNOSF sources (Corollary 6.4) to construct a condenser for low-entropy oNOSF sources for all g and ℓ:

Corollary 5.4. *For all constant g, ℓ, and all n, k, ε with $k \geq \mathrm{poly}(\log n)$, $\varepsilon \geq n^{-\Omega(1)}$, there exists an* explicit *condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *such that for any (g,ℓ,n,k)-oNOSF source $\mathbf{X}$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) = 1/\lfloor (\ell-1)/(g-1) \rfloor \cdot m - O(\log(m/\varepsilon))$ where $m = \Omega(k^{-\Omega(1)})$.*

5.1 Proving the main theorem

Here we prove the following full version of Theorem 5.1:

Theorem 5.5. *There exists a universal constant C such that for all n, g, ℓ, ε where $g > \ell/2$, there exists an explicit condenser $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ satisfying: for any uniform (g,ℓ)-oNOSF source $\mathbf{X}$ with $g > \ell/2$, we have $H_\infty^\varepsilon(\mathsf{Cond}(\mathbf{X})) \geq m - (3C)^\ell \log(2\ell n/\varepsilon)$ where $m = \frac{(n/2)-(3C)^\ell \log(2\ell n/\varepsilon)}{3}$.*

The proof of this theorem follows by explicitly constructing an extractor for multiple independent uniform sources, where few of them are set to constant, with the error depending on the longest source that is non-constant. Formally, we use the following:

Lemma 5.6. *There exists universal constant C such that for all $n_k, k_x, n_{y,1}, \ldots, n_{y,t}, m, 0 < \varepsilon_1 \leq \cdots \leq \varepsilon_t < 1$ satisfying $n_{y,i} \geq C\log(2n_x/\varepsilon_i)$ and $m = \frac{k_x - \log(2/\varepsilon_1)}{3}$, the following holds: There exists an explicit extractor $\mathsf{Ext} : \{0,1\}^{n_x} \times \{0,1\}^{n_{y,i}} \times \cdots \times \{0,1\}^{n_{y,t}} \to \{0,1\}^m$ satisfying: For all $1 \leq j \leq t$ and all independent sources $\mathbf{X} \sim \{0,1\}^{n_x}, \mathbf{Y}_1 \sim \{0,1\}^{n_{y,1}}, \ldots, \mathbf{Y}_t \sim \{0,1\}^{n_{y,t}}$ where $H_\infty(\mathbf{X}) = k_x$, each of $\mathbf{Y}_1, \ldots, \mathbf{Y}_{j-1}$ are fixed constants and all $\mathbf{Y}_j, \ldots, \mathbf{Y}_t$ are uniform, we have that $\mathsf{Ext}(\mathbf{X}, \mathbf{Y}_1, \ldots, \mathbf{Y}_t)$ is ε_j-close to $\mathbf{U}_m$.*

We will prove this lemma at the end of this section. Along with this extractor, we use the fact that if a function f condenses any flat distribution $\mathbf{X}$, then f also condenses, with a small loss in parameters, any distribution $\mathbf{X}'$ which is the same as $\mathbf{X}$ but with a few bits controlled by an adversary. Lemma 4.9 elaborates and provides the result we need. Using this and our previously stated extractor, our condenser result follows:

Proof of Theorem 5.5. First, we split each block in $\mathbf{X}$ into two to obtain uniform $(2g, 2\ell, n/2)$-oNOSF source. This source is still $\mathbf{X}$, just with this new parameters. For $1 \leq i \leq \ell$, let $n_{y,i} = 2C(3C)^{\ell-i}\log(2\ell n/\varepsilon)$ and let $n_y = \sum_{i=1}^{\ell} n_{y,i}$. Let $\mathbf{Y}_i$ be the length $n_{y,i}$ prefix of the block $\mathbf{X}_{\ell+i}$. Let $\mathbf{W} = (\mathbf{X}_1, \ldots, \mathbf{X}_\ell)$ and let $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell)$. We use the extractor Ext from Lemma 5.6 with $k_x = n/2 - n_y - \log(2/\varepsilon)$, m from the lemma statement, and for $1 \leq i \leq \ell$, we set $\varepsilon_i = \left(\frac{\varepsilon}{2\ell n}\right)^{(3C)^{\ell-i}}$. With this, we define our condenser as:

$$\mathsf{Cond}(\mathbf{X}) = \mathsf{Ext}(\mathbf{W}, \mathbf{Y}_1, \ldots, \mathbf{Y}_\ell).$$

We easily compute and check that our parameter settings satisfy the requirements of Lemma 5.6. We will show that the output entropy (with error ε) is at least $m - n_y$. We compute that $n_y \leq$

$(3C)^\ell \log(2\ell n/\varepsilon)$, the output entropy gap. Hence if we show this, then our condenser will indeed have the claimed property.

We now show that our condenser construction is correct. First, since $2g > \ell$, there exists at least one good block amongst $\mathbf{X}_{\ell+1}, \dots, \mathbf{X}_{2\ell}$ and hence, at least one good block amongst $\mathbf{Y}_1, \dots, \mathbf{Y}_\ell$. Let this good block appear at index $j \in [\ell]$. Similarly, there exists at least one good block amongst $\mathbf{X}_1, \dots, \mathbf{X}_\ell$ and so, $H_\infty(\mathbf{W}) \geq n/2$. Let $\mathbf{A} = \mathbf{Y}_1, \dots, \mathbf{Y}_{j-1}$ and let $\mathbf{B} = \mathbf{Y}_{j+1}, \dots, \mathbf{Y}_\ell$. Then, $\mathbf{Y} = (\mathbf{A}, \mathbf{Y}_j, \mathbf{B})$. We will show that $H_\infty^\varepsilon(\mathsf{Ext}(\mathbf{W}, (\mathbf{A}, \mathbf{Y}_j, \mathbf{B}))) \geq m - n_y$.

We will now consider fixings of $\mathbf{A}$. We say a fixing of $\mathbf{A} = a$ is good if $H_\infty(\mathbf{W} | \mathbf{A} = a) \geq n/2 - n_y - \log(2/\varepsilon) = k_x$. By the min-entropy chain rule (Lemma 3.2), at least $1 - \varepsilon/2$ fraction of fixings of $\mathbf{A}$ are good. As $\mathbf{X}$ is an oNOSF source, $\mathbf{X}_{\ell+j}$ is independent of blocks $\mathbf{X}_1, \dots, \mathbf{X}_{\ell+j-1}$. Hence, $\mathbf{Y}_j$ remains uniform and independent of $\mathbf{W}$, for every fixing of $\mathbf{A}$. We will show that for every good fixing of $\mathbf{A} = a$, $H_\infty^{\varepsilon/2} \mathsf{Ext}(\mathbf{W}, \mathbf{Y}) \geq m - \sum_{i=j+1}^n n_{y,i} \geq m - n_y$. This will prove our result as our total error will be ε and the min-entropy guarantee will be $m - n_y$, as desired.

Consider the best case scenario when $(\mathbf{B} | \mathbf{A} = a) = \mathbf{U}_{|\mathbf{B}|}$. This is unrealistic since it is possible that all bits in $\mathbf{B}$ are bad and arbitrarily depend on the remaining bits. Nevertheless, it is instructive to see what happens in this scenario. In this case, $\mathbf{W}, \mathbf{Y}$ are independent distributions and we can infer that $\mathsf{Ext}(\mathbf{W}, \mathbf{Y}) \approx_{\varepsilon_j} \mathbf{U}_m$. However, as alluded before, all bits in $\mathbf{B}$ can be adversarially set. To overcome this, we invoke Lemma 4.9 that allows us to compare how worse off our output distribution can be compared to the best case scenario. We conclude that even when $\mathbf{B}$ is completely adversarially controlled, $H_\infty^{\varepsilon'}(\mathsf{Ext}(\mathbf{W}, \mathbf{Y})) \geq m - |\mathbf{B}| = m - \sum_{i=j+1} n_{y,i}$ where

$$
\begin{aligned}
\varepsilon' &= \varepsilon_j \cdot 2^{|\mathbf{B}|} \\
&= \left(\frac{\varepsilon}{2\ell n}\right)^{(3C)^{\ell-j}} \cdot 2^{\sum_{i=j+1}^{\ell} n_{y,i}} \\
&= \left(\frac{\varepsilon}{2\ell n}\right)^{(3C)^{\ell-j}} \cdot 2^{2C \log(2\ell n/\varepsilon) \sum_{i=j+1}^{\ell} (3C)^{\ell-i}} \\
&= \left(\frac{\varepsilon}{2\ell n}\right)^{(3C)^{\ell-j}} \cdot \left(\frac{2\ell n}{\varepsilon}\right)^{2C \frac{(3C)^{\ell-j}-1}{3C-1}} \\
&\leq \left(\frac{\varepsilon}{2\ell n}\right)^{(3C)^{\ell-j}} \cdot \left(\frac{2\ell n}{\varepsilon}\right)^{(3C)^{\ell-j}-1} \\
&\leq \frac{\varepsilon}{2\ell n} \\
&\leq \varepsilon/2
\end{aligned}
$$

This proves our claim, showing that for all good fixings, our output is highly condensed.

We need to be careful when invoking Lemma 4.9 since it requires that $(\mathbf{W}, \mathbf{A}, \mathbf{Y}_j, \mathbf{U}_{|\mathbf{B}|})$ should be a flat distribution. While that may not be true, we can express $\mathbf{W}$ as a convex combination of flat sources with same min-entropy and since $\mathbf{A}$ is fixed and $\mathbf{Y}_j$ and $\mathbf{U}_{|\mathbf{B}|}$ are independent and uniform, we can express the joint distribution as convex combination of flat sources, for each of them invoke the lemma, and conclude that the original distribution will be condensed as well. □

Lastly, we show how to construct our multi-source extractor with the desired properties. Using various seeded extractors, we construct our final extractor as follows:

Proof of Lemma 5.6. For $1 \leq i \leq t$, let $\mathsf{sExt}_i : \{0,1\}^{n_x} \times \{0,1\}^{n_{y,i}} \to \{0,1\}^m$ be explicit $(\varepsilon_i/2)$-seeded-extractor guaranteed by Theorem 3.5. Our extractor construction is:

$$\mathsf{Ext}(x, y_1, \ldots, y_t) = \bigoplus_{i=1}^{t} \mathsf{sExt}_i(x, y_i).$$

Let $\mathbf{Z}_{good} = \mathsf{sExt}_j(\mathbf{X}, \mathbf{Y}_j)$ and let $\mathbf{Z}_{rest} = \bigoplus_{1 \leq i \leq t, i \neq j} \mathsf{sExt}_i(\mathbf{X}, \mathbf{Y}_i)$. Notice that our final output distribution is $\mathbf{Z}_{good} \oplus \mathbf{Z}_{rest}$. We will argue that on most fixings of $\mathbf{Z}_{rest}$, the output will be close to uniform.

By Lemma 3.2, we have that

$$\Pr[H_\infty(\mathbf{X}|\mathbf{Z}_{rest} = z_{rest}) \geq k_x - m - \log(2/\varepsilon_j)] \geq 1 - \varepsilon_j/2.$$

Call the fixings z_{rest} of $\mathbf{Z}_{rest}$ that satisfy the above property of leaving $\mathbf{X}$ with a lot of entropy when conditioning on them, as the "good fixings." As $\mathbf{Z}_{rest}$ is independent of $\mathbf{Y}_j$ and $\mathbf{X}$ is left with a lot of entropy conditioning on a good fixing z_{rest}, we have that

$$\mathsf{sExt}_j((\mathbf{X}|\mathbf{Z}_{rest} = z_{rest}), (\mathbf{Y}_j|\mathbf{Z}_{rest} = z_{rest})) \approx_{\varepsilon_j/2} \mathbf{U}_m.$$

As $1 - \varepsilon_j/2$ fraction of fixings of $\mathbf{Z}_{rest}$ are good, we conclude that $\mathsf{Ext}(\mathbf{X}, \mathbf{Y}_1, \ldots, \mathbf{Y}_t) \approx_{\varepsilon_j} \mathbf{U}_m$ as desired. ☐

6 Transforming low-entropy oNOSF sources to uniform oNOSF sources

In this section, we show how to transform low-entropy oNOSF sources into uniform oNOSF sources. Such a transformation was also provided in [CGR24]. Here, we obtain improved bounds using a generalized construction that allows us to obtain better tradeoffs and parameters in many more regimes of n, ℓ. Our main theorem is:

Theorem 6.1. *Let $d, g, g_{out}, \ell, n, m, k, \varepsilon$ be such that $g_{out} \leq g - \frac{\ell - g + 2}{d}, n \geq k \geq \log(nd - k) + md + 2\log(2g_{out}/\varepsilon)$. Then, there exists 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 uniform $(g_{out}, \ell - 1, m)$-oNOSF source $\mathbf{Y}$ for which $|f(\mathbf{X}) - \mathbf{Y}| \leq \varepsilon$.*

Our construction's power comes from the flexibility of setting d to any desired value. For instance by setting d to be a large constant, we can get the following transformation that works even when n is very small compared to ℓ:

Corollary 6.2 (Transformation for small n)**.** *Let $g, \ell, n, m, k, \varepsilon, \delta$ be such that $\delta \leq 0.99, g = \delta\ell, n = \mathrm{poly}(\log(\delta\ell/\varepsilon)), k = \Omega(\log(\delta\ell/\varepsilon)), m = \Omega(k)$. 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 uniform $(0.99\delta\ell, \ell - 1, m)$-oNOSF source $\mathbf{Y}$ such that $|f(\mathbf{X}) - \mathbf{Y}| \leq \varepsilon$.*

We additionally note that when we set $d = \ell$, we recover the same construction as in [CGR24], matching its parameters. This is most interesting in the regime when say $\ell = O(1)$ and n is arbitrarily growing.

Corollary 6.3 (similar parameters as Theorem 5.2 from [CGR24])**.** *Let* $g, \ell, n, m, k, \varepsilon$ *be such that* $k \geq 1.01(\log(n\ell) + 2\log(2(g-1)/\varepsilon)), m = k/200\ell$. *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 uniform* $(g-1, \ell-1, m)$*-oNOSF source* $\mathbf{Y}$ *such that* $|f(\mathbf{X}) - \mathbf{Y}| \leq \varepsilon$.

To obtain these transformations, we will use two-source extractors. In fact, using explicit construction of two-source-extractors, we also obtain an explicit transformation:

Corollary 6.4 (Explicit Transformation)**.** *There exists a universal constant* C *such that for all* $d, g, g_{out}, \ell, n, m, k, \varepsilon$ *satisfying* $g_{out} \leq g - \frac{\ell-g+2}{d}, k \geq \text{poly}(\log(n)) + md + 2\log(2g_{out}/\varepsilon) + O(1), m \leq \text{poly}(\log n), \varepsilon \geq n^{-\Omega(1)}/2g_{out}$. *the following holds: There exists an explicit 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 uniform* $(g_{out}, \ell-1, m)$*-oNOSF source* $\mathbf{Y}$ *for which* $|f(\mathbf{X}) - \mathbf{Y}| \leq \varepsilon$.

We can instantiate this lemma even in the case of constant d and get an explicit transformation similar to Corollary 6.2 with fewer output bits per block.

We will use the following main technical lemma that shows how to use two-source extractors to obtain these transformations:

Lemma 6.5 (Main Lemma)**.** *Let* $d, g, g_{out}, \ell, n, m, k_{\mathsf{2Ext}}, k, \varepsilon_{\mathsf{2Ext}}$ *be such that* $k \geq k_{\mathsf{2Ext}} + m \cdot d + \log(1/\varepsilon_{\mathsf{2Ext}}), g_{out} \leq \frac{g(d+1)-\ell-2}{d}$. *Let* $\mathsf{2Ext} : \{0,1\}^{d \cdot n} \times \{0,1\}^n \to \{0,1\}^m$ *be* $(k_{\mathsf{2Ext}}, \varepsilon_{\mathsf{2Ext}})$*-average-case-strong two-source extractor. 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* $(g_{out}, \ell-1, m)$*-oNOSF source* $\mathbf{Y}$ *such that* $|f(\mathbf{X}) - \mathbf{Y}| \leq \varepsilon$ *where* $\varepsilon = 2g_{out} \cdot \varepsilon_{\mathsf{2Ext}}$.

Existentially, two-source-extractors with following parameters exist:

Lemma 6.6 (Lemma 5.4 from [CGR24])**.** *Let* $n_1, n_2, k_1, k_2, m, \varepsilon$ *be such that* $k_1 \leq n_1, k_2 \leq n_2, m = k_1 + k_2 - 2\log(1/\varepsilon) - O(1)$*,* $k_2 \geq \log(n_1 - k_1) + 2\log(1/\varepsilon) + O(1)$*, and* $k_1 \geq \log(n_2 - k_2) + 2\log(1/\varepsilon) + O(1)$. *Then, a random function* $\mathsf{2Ext} : \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^m$ *is a* (k_1, k_2, ε)*-two source extractor with probability* $1 - o(1)$.

Using this, our main result follows:

Proof of Theorem 6.1. We use the two-source-extractors from Theorem 6.1 and apply it in Lemma 6.5. □

To make this transformation explicit, we can use the following construction of a two-source-extractor:

Theorem 6.7 ([CZ19, Mek17, Li16])**.** *There exists a universal constant* $C \geq 1$ *such that for all* n, k, m, ε *with* $k \geq \log^C(n), m \leq n^{1/C}, \varepsilon \geq n^{-1/C}$*, the following holds: There exists an explicit* (n, k) *two-source-extractor* $\mathsf{2Ext} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$.

With this our explicit transformation follows:

Proof of Corollary 6.4. We use the explicit two-source-extractors from Theorem 6.7 and apply it in Lemma 6.5. □

6.1 Low-entropy oNOSF source to uniform using two-source-extractors

In this subsection, we will prove Lemma 6.5. To do this, we will use two-source-extractors and average-case two-source-extractors. Let's first define them:

Definition 6.8. *We say that* $\mathsf{2Ext}$ *is* (k_1, k_2, ε) average-case *strong if*

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

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* $H_\infty(\mathbf{X}_2) \geq k_2$ *and* $\mathbf{W}$*.*

This notion of average-case two-source-extractors allows us obtain a simpler chain rule:

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

Lemma 2.3 of [DORS08] shows that all two-source extractors are average-case-two-source extractors with similar parameters.

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

With this, we will finally prove our main lemma that shows how to use two-source-extractors to obtain our transformation:

Proof of Lemma 6.5. For $-d \leq i \leq 0$, define $\mathbf{X}_i$ to be the random variable that always outputs 0^n. For $2 \leq i \leq \ell$, we output $\mathbf{O}_i = \mathsf{2Ext}(\mathbf{X}_{i-d} \circ \cdots \circ \mathbf{X}_{i-1}, \mathbf{X}_i)$.

For $2 \leq i \leq \ell$, we say that $\mathbf{O}_i$ is good if (1) $\mathbf{X}_i$ is good and (2) there exists a block amongst $\mathbf{X}_{i-d}, \ldots, \mathbf{X}_{i-1}$ that is good. We observe that if $\mathbf{O}_i$ is good, then $|\mathbf{O}_i - \mathbf{U}_m| \leq \varepsilon_{\mathsf{2Ext}}$. Let g' be the number of such good $\mathbf{O}_i$. Let $j_1, \ldots, j_g$ be the indices of the good blocks in $\mathbf{X}$. For $1 \leq i \leq g-1$, let $d_i = j_{i+1} - j_i$. We observe that g' equals number of i such that $d_i \leq d$. As $\sum_{i=1}^{g-1} d_i \leq \ell$ and $d_i \geq 1$, we infer that $g' \geq \frac{(g-1)(d+1)-\ell}{d}$. Hence, as long as $g_{out} \leq \lceil g' \rceil$, we can guarantee the desired number of good blocks in the output. This holds as long as $g_{out} \leq \frac{g(d+1)-\ell-2}{d}$.

Using Lemma 6.10, we infer that $\mathsf{2Ext}$ is $(k_{\mathsf{2Ext}} + \log(1/\varepsilon_{\mathsf{2Ext}}), 2\varepsilon_{\mathsf{2Ext}})$-average-case-two-source extractor. We will use this property below.

Now, using a hybrid argument we will show that

$$(\mathbf{O}_2, \ldots, \mathbf{O}_\ell) \approx_{2g_{out} \cdot \varepsilon_{\mathsf{2Ext}}} (\mathbf{Y}_2, \ldots, \mathbf{Y}_\ell)$$

where $\mathbf{Y} = (\mathbf{Y}_2, \ldots, \mathbf{Y}_\ell)$ is a uniform (g_{out}, ℓ, m)-oNOSF source that we will define as the proof goes. Let $\mathbf{Y}^{(1)} = (\mathbf{O}_2, \ldots, \mathbf{O}_\ell)$ and for $2 \leq i \leq \ell$, let $\mathbf{Y}^{(i)} = (\mathbf{O}_2, \ldots, \mathbf{O}_i, \mathbf{Y}_{i+1}, \ldots, \mathbf{Y}_\ell)$. Hence, $\mathbf{Y}^{(\ell)} = \mathbf{Y}$. We proceed by induction. We will show that for $2 \leq i \leq \ell$,

$$\left|\mathbf{Y}^{(i)} - \mathbf{Y}^{(i-1)}\right| \leq 2\varepsilon_{\mathsf{2Ext}}$$

whenever $\mathbf{O}_i$ is good and

$$\mathbf{Y}^{(i)} = \mathbf{Y}^{(i-1)}$$

whenever $\mathbf{O}_i$ is bad. By repeated applications of the triangle inequality, we will have shown that our output is indeed close to some uniform oNOSF source with desired parameters.

We proceed by induction and let $i \geq 2$ be arbitrary. If $\mathbf{O}_i$ is bad, then we let $\mathbf{Y}_i = \mathbf{O}_i$. Then, we indeed have that $\mathbf{Y}^{(i)} = \mathbf{Y}^{(i-1)}$ as desired. Otherwise, we assume $\mathbf{O}_i$ is good. Then, it must be that $\mathbf{X}_i$ is good. Let i_{prev} be the index of the good block before $\mathbf{X}_i$ in $\mathbf{X}$. Then, we know that $i - i_{prev} \leq d$. We first claim that

$$\widetilde{H}_\infty(\mathbf{X}_{i_{prev}} | \mathbf{O}_1, \dots, \mathbf{O}_{i-1}) \geq k_{\mathsf{2Ext}} = k - m \cdot d$$

Firstly, by construction, blocks $\mathbf{O}_2, \mathbf{O}_{i_{prev}-1}$ are functions of blocks $\mathbf{X}_1, \dots, \mathbf{X}_{i_{prev}-1}$. As $\mathbf{X}_{i_{prev}}$ is independent of $\mathbf{X}_1, \dots, \mathbf{X}_{i_{prev}-1}$, we infer that $\mathbf{X}_{i_{prev}}$ is independent of $\mathbf{O}_2, \mathbf{O}_{i_{prev}-1}$. As $\mathsf{2Ext}$ is average-case-strong, we apply Lemma 6.9 to get that

$$\widetilde{H}_\infty(\mathbf{X}_{i_{prev}} | \mathbf{O}_2, \dots, \mathbf{O}_{i-1}) \geq k - m \cdot (i - i_{prev}) \geq k - m \cdot d = k_{\mathsf{2Ext}} + \log(1/\varepsilon)$$

where for the second last inequality, we used the fact that $i - i_{prev} \leq d$. Moreover, as $\mathbf{X}_i$ is independent of $\mathbf{X}_1, \dots, \mathbf{X}_{i-1}$ and $\mathbf{O}_2, \dots, \mathbf{O}_{i-1}$ are solely functions of $\mathbf{X}_1, \dots, \mathbf{X}_{i-1}$, we infer that $\mathbf{X}_i$ is independent of $\mathbf{O}_2, \dots, \mathbf{O}_{i-1}$. Hence, conditioned on fixing $\mathbf{O}_2, \dots, \mathbf{O}_{i-1}$, $\mathbf{O}_i$ will be $2\varepsilon_{\mathsf{2Ext}}$ close to $\mathbf{U}_m$. This implies $\mathbf{Y}^{(i-1)} \approx_{2\varepsilon_{\mathsf{2Ext}}} \mathbf{Y}^{(i)}$ as desired. This shows that a good block in $\mathbf{Y}$ is uniform conditioned on all previous blocks, .i.e., it is independent of all the blocks before it. This shows all bad blocks can only depend on good blocks appearing before them and that good blocks are independent of each other. This implies $\mathbf{Y}$ is indeed a uniform oNOSF source as desired. □

7 Online influence

Our results from Section 4 provide an almost complete picture on the possibility of condensing from uniform (g, ℓ, n)-oNOSF source, when the length of each symbol (i.e., n) is a large enough constant. In this section, we focus on the class of oNOBF sources, which corresponds to uniform $(g, \ell, n = 1)$-oNOSF source and, as noted before, provides the adversary with the most amount of power. We prove lower bounds on the possibility of extraction for oNOBF sources. To this end, we introduce a new, natural notion of influence of Boolean functions, which we call *online influence*.

We initiate a study of online influence, that we believe is of independent interest. We formally define the notion and discuss some basic properties in Section 7.1. We establish tight bounds on the online influence for general functions, including a Poincaré style inequality, in Section 7.2. We provide an example exhibiting a separation between maximum (standard) influence and online influence in Section 7.3. Finally, we present lower bounds for extraction and condensing from oNOBF sources in Section 7.4.

7.1 Basic Properties

In this section, for a function $f : \{0,1\}^\ell \to \{0,1\}$, we will freely use commas to indicate concatenation in its input. For example, for $x \in \{0,1\}^{i-1}$ and $y \in \{0,1\}^{\ell-i+1}$, we write $f(x, 1, y)$ to indicate f applied to the tuple $(x_1, \dots, x_{i-1}, 1, y_1, \dots, y_{\ell-i+1})$.

When asking about the influence of a single bit, such as the i-th bit, previous work has specifically looked at whether the i-th bit still has the ability to change the output of some function $f : \{0,1\}^\ell \to \{0,1\}$ after all other $\ell - 1$ bits have been set. In other words, if the i-th bit is a non-oblivious adversary (that is, it can look at the values of all the other bits before setting its own value), how much power does it have? This has led to a standard notion of influence defined below.

Definition 7.1 (Influence)**.** *For a function* $f : \{0,1\}^\ell \to \{0,1\}$*, the* influence *of the* i*-th bit is*

$$\mathbf{I}_i[f] = \mathop{\mathbb{E}}_{\substack{x \sim \mathbf{U}_{i-1} \\ y \sim \mathbf{U}_{n-i}}} [|f(x,1,y) - f(x,0,y)|]$$

and the total influence *is*

$$\mathbf{I}[f] = \sum_{i=1}^{\ell} \mathbf{I}_i[f].$$

However, in our setting of oNOSF sources and oNOBF sources, an adversarial bit can only depend on the bits that come before it. This motivates our new definition of online influence, where we prevent the i-th bit from depending on bits that come after it by independently sampling subsequent bits.

Definition 7.2 (Online influence)**.** *For a function* $f : \{0,1\}^\ell \to \{0,1\}$*, the* online influence *of the* i*-th bit is*

$$\mathbf{oI}_i[f] = \mathop{\mathbb{E}}_{x \sim \mathbf{U}_{i-1}} \left[\left| \mathop{\mathbb{E}}_{y \sim \mathbf{U}_{\ell-i}} [f(x,1,y)] - \mathop{\mathbb{E}}_{y \sim \mathbf{U}_{\ell-i}} [f(x,0,y)] \right| \right]$$

and the total online influence *is*

$$\mathbf{oI}[f] = \sum_{i=1}^{\ell} \mathbf{oI}_i[f].$$

Remark 7.3. *It is easy to see that for any* $f : \{0,1\}^\ell \to \{0,1\}$*, and any* $i \in \ell$*,* $\mathbf{oI}_i(f) \le \mathbf{I}_i(f)$*. Further, they are the same for the last bit:* $\mathbf{I}_\ell[f] = \mathbf{oI}_\ell[f]$*.*

Many results for the influence of a function are based on working with monotone functions. In contrast, it turns out that monotone functions are not very interesting for online influence as the definition collapses to that of regular influence.

Lemma 7.4. *If* $f : \{0,1\}^\ell \to \{0,1\}$ *is monotone, then* $\mathbf{oI}_i[f] = \mathbf{I}_i[f]$ *for all* $i \in [\ell]$*.*

Proof. Using the monotonicity of f, note that for any $x \in \{0,1\}^{i-1}$ and any $y \in \{0,1\}^{\ell-i}$, $f(x,1,y) \ge f(x,0,y)$. Thus, $\mathbf{oI}_i[f] = \mathbb{E}_{x \sim \mathbf{U}_{i-1}, y \sim \mathbf{U}_{\ell-i}}[f(x,1,y)] - f(x,0,y)] = \mathbf{I}_i(f)$. □

Thus, any difference between influence and online influence can only be demonstrated by non-monotone functions.

7.2 A Poincaré inequality for online influence

Similar to regular influence, we prove a Poincaré-style inequality holds for online influence, and also provide an upper bound on online influence. The following is the main result of this subsection.

Theorem 7.5. *For any* $f : \{0,1\}^\ell \to \{0,1\}$*, we have* $\mathrm{Var}(f) \le \mathbf{oI}[f] \le \sqrt{\ell \, \mathrm{Var}(f)}$*.*

Before proving the above result, we observe that MAJORITY and PARITY functions provide tight examples (up to constants) for the upper and lower bound respectively for Theorem 7.5.

Example 7.6. *The majority function on ℓ bits $\mathrm{Maj}_\ell : \{0,1\}^\ell \to \{0,1\}$, is monotone, and hence by by Lemma 7.4, has total online influence $\mathbf{oI}[\mathrm{Maj}_\ell] = \mathbf{I}[\mathrm{Maj}_\ell] = \sqrt{2\ell/\pi} + O(1/\sqrt{\ell})$, achieving the upper bound (up to constants).*

The PARITY function on ℓ bits $\bigoplus_\ell : \{0,1\}^\ell \to \{0,1\}$ for $i \in [\ell-1]$ has online influence $\mathbf{oI}_i[\bigoplus_\ell] = 0$, while $\mathbf{oI}_\ell[\bigoplus_\ell] = 1$. Thus, PARITY meets the lower bound of Theorem 7.5. We note that this is starkly different from regular influence where $\mathbf{I}_i[\bigoplus_n] = 1$ for all i.

To prove Theorem 7.5, we will use Boolean Fourier analysis. Recall that any Boolean function $f : \{0,1\}^\ell \to \{0,1\}$ has a unique Fourier expansion: $f(x) = \sum_{S\subseteq[\ell]} \widehat{f}(S)\chi_S(x)$, where $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$ and $\widehat{f}(S) = \mathbb{E}_{y\sim\mathbf{U}_\ell}[(-1)^{f(y)}\chi_S(y)]$. Also recall that $\widehat{f}(\emptyset) = \mathbb{E}[f(\mathbf{U}_\ell)]$, $\mathrm{Var}(f) = \sum_{S\subseteq[\ell], S\neq\emptyset} \widehat{f}(S)^2$, and for any $S \neq T$, $\mathbb{E}_{x\sim\mathbf{U}_\ell}[\chi_S(x)\chi_T(x)] = 0$. For more background, we refer the reader to the excellent book by O'Donnell [ODo14].

The following is our key lemma, from which Theorem 7.5 is easy to derive.

Lemma 7.7. *For any $f : \{0,1\}^\ell \to \{0,1\}$ and $i \in [\ell]$, $\mathbf{oI}_i(f)^2 \le \sum_{\substack{S\subseteq[i]\\ S\ni i}} \widehat{f}(S)^2 \le \mathbf{oI}_i(f)$.*

We first derive Theorem 7.5 using Lemma 7.7.

Proof of Theorem 7.5. We first show the lower bound $\mathbf{oI}[f] \ge \mathrm{Var}(f)$. We have,

$$\begin{aligned}
\mathbf{oI}[f] &= \sum_{i=1}^{\ell} \mathbf{oI}_i[f] \\
&\ge \sum_{i=1}^{\ell} \sum_{\substack{S\subseteq[i]\\ S\ni i}} \widehat{f}(S)^2 && \text{(Lemma 7.7)} \\
&= \sum_{\substack{S\subseteq[\ell]\\ S\neq\varnothing}} \widehat{f}(S)^2 \\
&= \mathrm{Var}(f).
\end{aligned}$$

The upper bound is easy to derive as well.

$$\begin{aligned}
\mathbf{oI}[f] &= \sum_{i=1}^{\ell} \mathbf{oI}_i[f] \\
&\le \sqrt{\ell \sum_{i=1}^{\ell} (\mathbf{oI}_i[f])^2} && \text{(Cauchy-Schwarz inequality)} \\
&\le \sqrt{\ell \sum_{i=1}^{\ell} \sum_{\substack{S\subseteq[i]\\ S\ni i}} \widehat{f}(S)^2} && \text{(Lemma 7.7)} \\
&= \sqrt{\ell\,\mathrm{Var}(f)}.
\end{aligned}$$

This completes the proof. ∎

We now focus on proving Lemma 7.7. We need the following useful characterization of $\mathbf{oI}_i(f)$.

Claim 7.8. *For any $f : \{0,1\}^\ell \to \{0,1\}$, we can write the online influence of its i-th bit as*

$$\mathbf{oI}_i[f] = \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)\right|\right].$$

Assuming the above claim, let us prove Lemma 7.7. We supply the proof of Claim 7.8 below.

Proof of Lemma 7.7. We first prove the inequality $\mathbf{oI}_i(f) \geq \sum_{\substack{S\subseteq[i]\\ S\ni i}} \widehat{f}(S)^2$. Note that $\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)$ is in $[0,1]$ since $\mathbb{E}_{y\sim\mathbf{U}_{\ell-i}}[f|_{x,b}(y)]$ is in $[0,1]$ for all $x \in \{0,1\}^{i-1}, b \in \{0,1\}$. Moreover, for any $x \in \{0,1\}^{i-1}$ we have $\left|\mathbb{E}_{y\sim\mathbf{U}_{\ell-i}}[f|_{x,1}(y)] - \mathbb{E}_{y\sim\mathbf{U}_{\ell-i}}[f|_{x,0}(y)]\right| = \left|\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)\right|$ by Claim 7.8.

Thus,

$$\begin{aligned}
\mathbf{oI}_i[f] &= \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)\right|\right] \\
&\geq \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left(\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)\right)^2\right] \\
&= \sum_{\substack{T\subseteq[i]\\ T\ni i}}\sum_{\substack{S\subseteq[i]\\ S\ni i}} \widehat{f}(T)\widehat{f}(S)\mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}[\chi_{T\setminus\{i\}}(x)\chi_{S\setminus\{i\}}(x)] \\
&= \sum_{\substack{S\subseteq[i]\\ S\ni i}} \widehat{f}(S)^2.
\end{aligned}$$

Next, we prove $\mathbf{oI}_i(f)^2 \leq \sum_{\substack{S \subseteq [i] \\ S \ni i}} \widehat{f}(S)^2$. We have,

$$\begin{aligned}
\mathbf{oI}_i[f]^2 &= \left(\mathop{\mathbb{E}}_{x \sim \mathbf{U}_{i-1}} \left[\left| \sum_{\substack{T \subseteq [i] \\ T \ni i}} \widehat{f}(T) \chi_{T \setminus \{i\}}(x) \right| \right] \right)^2 && \text{(Claim 7.8)} \\
&\leq \mathop{\mathbb{E}}_{x \sim \mathbf{U}_{i-1}} \left[\left(\sum_{\substack{T \subseteq [i] \\ T \ni i}} \widehat{f}(T) \chi_{T \setminus \{i\}}(x) \right)^2 \right] \\
&= \sum_{\substack{S \subseteq [i] \\ S \ni i}} \widehat{f}(S)^2 && \text{(derived above)}
\end{aligned}$$

□

Next, we show how to rewrite $\mathbf{oI}_i[f]$ in terms of the Fourier coefficients of f.

Proof of Claim 7.8. We begin by defining the restriction $f|_{x,b}(y) = f(x, b, y)$ for $x \in \{0,1\}^{i-1}$, $b \in \{0,1\}$, and $y \in \{0,1\}^{\ell - i}$. Thus, can we rewrite $\mathbf{oI}_i[f]$ as

$$\mathbf{oI}_i[f] = \mathop{\mathbb{E}}_{x \sim \mathbf{U}_{i-1}} \left[\left| \mathop{\mathbb{E}}_{y \sim \mathbf{U}_{\ell-i}} [f|_{x,1}(y)] - \mathop{\mathbb{E}}_{y \sim \mathbf{U}_{\ell-i}} [f|_{x,0}(y)] \right| \right]. \tag{1}$$

We would like to put the above expression in terms of Fourier coefficients of f. This motivates us to find the Fourier coefficients of $f|_{x,b}(y)$ in terms of those of f, which we do via computation. We manipulate the Fourier expansion of $f(z)$ for $z = (x, b, y) \in \{0,1\}^\ell$ to get

$$\begin{aligned}
f(z) &= \sum_{S \subseteq [\ell]} \widehat{f}(S) \chi_S(z) \\
&= \sum_{S \subseteq [\ell]} \widehat{f}(S) \chi_S(x, b, y) \\
&= \sum_{S \subseteq [\ell]} \widehat{f}(S) \chi_{S \cap [i]}(x, b) \chi_{S \setminus [i]}(y) \\
&= \sum_{S \subseteq \{i+1, \ldots, \ell\}} \left(\sum_{T \subseteq [i]} \widehat{f}(S \cup T) \chi_T(x, b) \right) \chi_S(y).
\end{aligned} \tag{2}$$

We also have that

$$\begin{aligned}
f(z) &= f(x, b, y) \\
&= f|_{x,b}(y) \\
&= \sum_{S \subseteq \{i+1, \ldots, \ell\}} \widehat{f|_{x,b}}(S) \chi_S(y).
\end{aligned} \tag{3}$$

Therefore, Equation (2) and Equation (3) allow us to conclude that

$$\widehat{f|_{x,b}}(S) = \sum_{T\subseteq[i]} \widehat{f}(S \cup T)\chi_T(x,b).$$

This allows us to easily take the expectation of $f|_{x,b}$ and get a result in terms of the Fourier coefficients of f.

$$\begin{aligned}\mathop{\mathbb{E}}_{y\sim\mathbf{U}_{\ell-i}}[f|_{x,b}(y)] &= \widehat{f|_{x,b}}(\varnothing)\\ &= \sum_{T\subseteq[i]} \widehat{f}(T)\chi_T(x,b)\\ &= \sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)b + \sum_{T\subseteq[i-1]} \widehat{f}(T)\chi_T(x).\end{aligned}$$

We now plug this in to our definition of $\mathbf{oI}_i[f]$ in Equation (1) to get a simplified expression.

$$\begin{aligned}\mathbf{oI}_i[f] &= \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\mathop{\mathbb{E}}_{y\sim\mathbf{U}_{\ell-i}}[f|_{x,1}(y)] - \mathop{\mathbb{E}}_{y\sim\mathbf{U}_{\ell-i}}[f|_{x,0}(y)]\right|\right]\\ &= \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\left(-\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x) + \sum_{T\subseteq[i-1]} \widehat{f}(T)\chi_T(x)\right) - \left(\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{b\}}(x) + \sum_{T\subseteq[i-1]} \widehat{f}(T)\chi_T(x)\right)\right|\right]\\ &= \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\sum_{\substack{T\subseteq[i]\\ T\ni i}} \widehat{f}(T)\chi_{T\setminus\{i\}}(x)\right|\right].\end{aligned}$$

□

7.3 A tight example for maximum online influence

The lower bound on total online influence from Theorem 7.5 allows us to conclude that for balanced functions, there must be at least one bit with online influence $\Omega(1/\ell)$. We can phrase this in terms of maximum influence.

Definition 7.9 (Maximum influence)**.** *For a function* $f : \{0,1\}^\ell \to \{0,1\}$*, we define its* maximum influence *as* $\mathbf{I}_{\max}[f] = \max_{i\in[\ell]} \mathbf{I}_i[f]$ *and its* maximum online influence *as* $\mathbf{oI}_{\max}[f] = \max_{i\in[\ell]} \mathbf{oI}_i[f]$.

In terms of maximum online influence, we get the following corollary from Theorem 7.5.

Corollary 7.10. *For a function* $f : \{0,1\}^\ell \to \{0,1\}$*, we have* $\mathbf{oI}_{\max}[f] \geq \mathrm{Var}(f)/\ell$.

Proof. By Theorem 7.5 we have that $\mathbf{oI}[f] = \sum_{i=1}^{\ell} \mathbf{oI}_i[f] \geq \mathrm{Var}(f)$, and the conclusion follows via an averaging argument. □

We show that the bound in Corollary 7.10 is in fact tight (up to constants), as witnessed by the address function.

Definition 7.11. *We define the* address function $\mathrm{Addr}_\ell : \{0,1\}^{\log(\ell)+\ell} \to \{0,1\}$ *as follows: For* $z \in \{0,1\}^{\log(\ell)+\ell}$*, split* z *up as* $z = (x, y)$ *with* x *of length* $\log(\ell)$ *and* y *of length* ℓ*. Then interpret* x *as a binary number which gives us an index* $i(x) \in [\ell]$*. The output of* Addr_ℓ *is the* $i(x)$*-th bit of* y*, so* $\mathrm{Addr}_\ell(x, y) = y_{i(x)}$.

Lemma 7.12. *Let* $m = \ell + \log \ell$ *and* Addr_ℓ *be the function defined above. Then,*

- *for* $1 \le i \le \log \ell$*,* $\mathbf{oI}_i[\mathrm{Addr}_\ell] = 0$.
- *for* $\log \ell < i \le m$*,* $\mathbf{oI}_i[\mathrm{Addr}_\ell] = 1/\ell$.

Thus, $\mathbf{oI}_{\max}(\mathrm{Addr}_\ell) = \Theta(1/m)$.

Proof. For $i \in [\log \ell]$, no matter what the value of the i-th bit of Addr_ℓ is set to, the output bit will be a uniform bit, so we immediately get that $\mathbf{oI}_i[f] = 0$. Formally, we see this by computing

$$\begin{aligned}\mathbf{oI}_i[f] &= \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\mathop{\mathbb{E}}_{y\sim\mathbf{U}_{\ell-i}}[f(x,1,y)] - \mathop{\mathbb{E}}_{y\sim\mathbf{U}_{\ell-i}}[f(x,0,y)]\right|\right] \\ &= \mathop{\mathbb{E}}_{x\sim\mathbf{U}_{i-1}}\left[\left|\frac{1}{2} - \frac{1}{2}\right|\right] \\ &= 0.\end{aligned}$$

For $i \in \{\log \ell + 1, \ldots, m\}$, the i-th bit only has control if it's selected by the first $\log \ell$ address bits, meaning it has a $1/\ell$ chance of controlling the output (and otherwise the output is uniform). Hence, $\mathbf{oI}_i[f] = \frac{1}{\ell}$. ☐

Compared with the result of [KKL88] that $\mathbf{I}_{\max}[f] \ge \mathrm{Var}(f) \cdot \Omega\left(\frac{\log \ell}{\ell}\right)$, this exhibits a separation between maximum (standard) influence and the online influence (of balanced functions).

Moreover, this analysis of the address function also shows us that it is an extractor for uniform $(\ell-1, \ell)$-oNOSF sources.

Lemma 7.13. *For all* ℓ, n *where* $\ell \ge 2$ *and* $n \ge \log(\ell - 1)$*, there exists an explicit extractor* $\mathsf{Ext} : (\{0,1\}^n)^\ell \to \{0,1\}^n$ *such that for any uniform* $(\ell-1, \ell, n)$*-oNOSF source* $\mathbf{X}$*, we have* $\mathsf{Ext}(\mathbf{X}) \approx_\varepsilon \mathbf{U}_n$ *where* $\varepsilon = \frac{1}{\ell-1}$.

Proof. Let Ext be defined as follows: From the first block, use the first $\log(\ell-1)$ bits and interpret them as an index $j \in [\ell-1]$. Then, output block with index $j+1$. For a source $\mathbf{X}$ with first block controlled by an adversary, the output will be truly uniform and for a source $\mathbf{X}$ with adversary controlling one of the last $\ell - 1$ blocks, that block will be outputted with probability $\frac{1}{\ell-1}$ while a uniform block will be outputted otherwise. This makes our total error at most $\frac{1}{\ell-1}$ as desired. ☐

7.4 Online influence of sets and extraction lower bounds

To prove lower bounds on extraction for oNOBF sources, we extend the definition of online influence to set in the natural way.

Definition 7.14 (Online influence of sets and online-resilient functions)**.** *For any function* $f : \{0,1\}^\ell \to \{0,1\}$*, and any* $B \subset [\ell]$*, where* $B = \{i_1 < i_2 < \ldots < i_k\}$*, define* $\mathbf{oI}_B(f)$ *as follows: an online adversary* $\mathcal{A}$ *samples a distribution* $\mathbf{X}$ *in online manner, starting with sampling the* *variables $x_1, x_2, \ldots, x_{i_1-1}$ independently and uniformly, and* $\mathbf{A}$ *picks the value of x_{i_1} depending on $x_{<i_1}$; next the variables $x_{i_1+1}, \ldots, x_{i_2-1}$ are sampled independently and uniformly, and* $\mathbf{A}$ *sets the value of x_{i_2} based on all set variables so for, and so on. Define the advantage of $\mathcal{A}$, $adv_{f,B}(A) = |\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{U}_\ell)]|$. Then, $I_B(f) = \max_{\mathcal{A}}\{adv_{f,B}(A)\}$, where the maximum is taken over all online adversaries $\mathcal{A}$ (that control bits in B).*

We say a function f is (b, ε)-online-resilient if $\mathbf{oI}_B(f) \leq \varepsilon$ for every B of size at most b.

Online-resilient functions can be used to construct extractors for oNOBF sources.

Lemma 7.15. *Let $f : \{0,1\}^\ell \to \{0,1\}$ be a (b, ε_1)-online-resilient function with the property that $|f(\mathbf{U}_\ell) - \mathbf{U}_1| \leq \varepsilon_2$. Then f can extract from $(g = \ell - b, \ell)$-oNOBF sources with error at most $\varepsilon_1 + \varepsilon_2$.*

Proof. Consider a $(g = \ell - b, \ell)$-oNOBF source $\mathbf{X}$. Recall that $\mathbf{X}$ is created by choosing some set of bad indices B of size b, letting the bits in $\overline{B}$ be uniform, and finally setting the bits in B adversarially while only depending on uniform bits to the left of them. Using the triangle inequality for total variation distance, we get that

$$\begin{aligned} |f(\mathbf{X}) - \mathbf{U}_1| &\leq |f(\mathbf{X}) - f(\mathbf{U}_\ell)| + |f(\mathbf{U}_\ell) - \mathbf{U}_1| \\ &\leq \varepsilon_1 + \varepsilon_2, \end{aligned}$$

as claimed. □

Remark 7.16. *Our results below on oNOBF extraction impossibility can be interpreted as a limit on online-resilience of balanced Boolean functions.*

For $B \subset [\ell]$, we use the notation $f|_{\overline{B}}$ to indicate the function obtained from f by letting an online adversary control the indices in B.

Theorem 7.17. *Let $f : \{0,1\}^\ell \to \{0,1\}$ be such that $\mathbb{E}_{x \sim \mathbf{U}_\ell}[f(x) = 1] = \alpha$. Then for any $1 \geq \beta > \alpha$, there exists a coalition $B \subseteq [\ell]$ such that $\mathbf{oI}_B(f) \geq \beta - \alpha$, where $|B| \leq \gamma n$ and $\gamma = \frac{\beta - \alpha}{\alpha(1-\beta)}$.*

Proof. We greedily collect the bits with the most online influence and add them to B until our goal of $\mathbb{E}_{x \sim \mathbf{U}_\ell|_{\overline{B}}}[f|_{\overline{B}}(x) = 1] \geq \beta$ is achieved. Our first step is as follows: let $B_0 = \varnothing$, $f_0 = f$, and $i_1 = \operatorname{argmax}_{i \in [\ell]}\{\mathbf{oI}_i[f]\}$. Corollary 7.10 tells us that $\mathbf{oI}_{i_1} \geq \operatorname{Var}(f_0)/n$. Recall that if $\mathbb{E}_{x \sim \mathbf{U}_\ell}[f(x) = 1] = p$ then $\operatorname{Var}(f) = p(1-p)$. Because we have not yet achieved our goal of $\mathbb{E}_{x \sim \mathbf{U}_\ell|_{\overline{B}}}[f|_{\overline{B}}(x) = 1] \geq \beta$, we have that $\operatorname{Var}(f_0) \geq \alpha(1-\beta)$. Thus, we collect i_1 as $B_1 = \{i_1\}$, let $f_1 = f_0|_{\overline{B_1}}$ and see that $\mathbb{E}_x[f_1(x)] \geq \mathbb{E}_x[f_0] + \mathbf{oI}_{i_1}[f_0] \geq \alpha + \frac{\alpha(1-\beta)}{\ell}$.

We now repeat this process t times to get $B_t = \{i_1, \ldots, i_t\}$ until our goal is achieved. For general t, let $f_t = f|_{B_t}$ where $B_t = B_{t-1} \cup \{i_t\}$ and $i_t = \operatorname{argmax}_{i \in [n] \setminus B_{t-1}}\{\mathbf{oI}_i[f_{t-1}]\}$. At the $(t-1)$-th step, since we have not stopped, it means that $\mathbb{E}_x[f_{t-1}(x) = 1] < \beta$, but we of course have $\mathbb{E}_x[f_{t-1}(x) = 1] \geq \alpha$ as well. Thus, by Corollary 7.10, collecting i_t as a bad bit gives us that

$$\begin{aligned} \mathop{\mathbb{E}}_x[f_t(x)] &\geq \mathop{\mathbb{E}}_x[f_{t-1}(x)] + \mathbf{oI}_{i_t}[f_{t-1}] \\ &\geq \alpha + \frac{\alpha(1-\beta)}{\ell}(t-1) + \frac{\alpha(1-\beta)}{\ell} \\ &= \alpha + \frac{\alpha(1-\beta)}{\ell} \cdot t. \end{aligned}$$

We repeat this process until $\Pr_x[f_t(x) = 1] \geq \beta$. Therefore, the number of steps is the smallest b such that $\alpha + \frac{\alpha(1-\beta)}{\ell} \cdot b \geq \beta$, meaning that the number of steps is at most $b \leq \ell \cdot \frac{\beta-\alpha}{\alpha(1-\beta)}$. We let $B = B_b$ and get the desired coalition. □

We can also ask the dual question of how large we are able to make β given some budget b of bad bits.

Corollary 7.18. *Let $f : \{0,1\}^\ell \to \{0,1\}$ be such that $\Pr_{x\sim\mathbf{U}_\ell}[f(x) = 1] \geq \alpha$. If we are able to control b bits in an online adversarial manner, then there exists a set $B \subseteq [\ell]$ of indices of size $|B| = b$ such that $\Pr_{x\sim\mathbf{U}_\ell|_{\overline{B}}}[f|_{\overline{B}}(x) = 1] \geq \beta$ where $\beta \geq \frac{\alpha(\ell+b)}{\ell+\alpha b}$.*

Proof. For a fixed β, Theorem 7.17 tells us that $b \leq \ell \cdot \frac{\beta-\alpha}{\alpha(1-\beta)}$. Solving for β, gives the desired bound. □

We now immediately obtain our oNOBF extraction impossibility result.

Corollary 7.19. *For any balanced function $f : \{0,1\}^\ell \to \{0,1\}$ and $0 < \varepsilon < 1/6$, there exists a $(g = \ell - b, \ell)$-oNOBF source $\mathbf{X}$ with $b \leq 6\varepsilon n$ such that $|f(\mathbf{X}) - \mathbf{U}_1| \geq \varepsilon$.*

Proof. It is enough to find a set B of indices such that $\mathbf{oI}_B(f) \geq \beta$. By Theorem 7.17, there exists such a set B of size $b = |B| \leq \ell \cdot \frac{2\varepsilon}{\frac{1}{2}-\varepsilon}$. The bound on $|B|$ follows since $\varepsilon \leq \frac{1}{6}$. □

Remark 7.20. *We can similarly obtain a Poincaré inequality for uniform oNOSF sources and show there always exists an influential block. To obtain extraction impossibility for such uniform oNOSF sources with constant δ fraction corrupt blocks, we do the following: Let f be candidate extractor for uniform $((1-\delta)\ell, \ell, n)$-oNOSF source. Then, f also extracts from uniform $(\lceil 1/\delta \rceil - 1, \lceil 1/\delta \rceil, \ell n/\lceil 1/\delta \rceil)$-oNOSF source. Since there exists an influential coordinate with influence $O(\delta)$, we let the adversary control that coordinate and infer there exists constant $\varepsilon = O(\delta)$ for which it is impossible to extract with error less than ε.*

8 Open Problems

We list here some interesting open problems left by our work:

- Explicitly construct a seeded condenser with dependence on seed length being $1 \cdot \log(1/\varepsilon)$. This will immediately make all our existential condenser constructions explicit.
- All our condensers have entropy gap much larger than a constant. It will be interesting to show there exist condensers with constant entropy gap (for any values of n, ℓ) for uniform oNOSF sources. A slightly weaker but equally interesting question is to construct seeded extractors for uniform oNOSF sources with constant seed length.
- Show that there exist non-trivial condensers for oNOBF sources or show no such condenser exists. We conjecture that no condenser exists with output entropy rate larger than the input entropy rate for such sources.
- Construct a function that is t-online resilient for $t = \omega\left(\frac{n}{\log^2(n)}\right)$ or show that $O\left(\frac{n}{\log^2(n)}\right)$ is tight. Note that one can obtain $t = O\left(\frac{n}{\log^2(n)}\right)$-online resilient functions using the Ajtai-Linial function [AL93].

- Construct ε-collective sampling protocols with lesser number of rounds than the ones obtained using uniform oNOSF source condensers. It will also be interesting to explicitly construct such protocols when number of players is very large compared to the number of bits each player has access to. Also proving lower bounds for ε-collective sampling protocols will be very interesting.

Acknowledgments

We thank Rocco Servedio for extremely useful conversations about the notion of online influence. In particular, the result in Section 7.3 was obtained jointly with Rocco. We also thank Madhur Tulsiani for asking a question during our talk at FOCS'24 that motivated us to consider the model of local oNOSF sources in Section 1.5.

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. 1, 2, 8).

[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. 31).

[AN93] Noga Alon and Moni Naor. "Coin-Flipping Games Immune Against Linear-Sized Coalitions". In: *SIAM J. Comput.* 22.2 (1993), pp. 403–417. DOI: 10.1137/0222030 (cit. on p. 5).

[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. 16).

[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 p. 16).

[BL89] Michael Ben-Or and Nathan Linial. "Collective Coin Flipping". In: *Advances In Computing Research* 5 (1989), pp. 91–115 (cit. on pp. 5, 6).

[CGR24] Eshan Chattopadhyay, Mohit Gurumukhani, and Noam Ringach. "On the Existence of Seedless Condensers: Exploring the Terrain". In: *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. To appear. 2024 (cit. on pp. 1–5, 9–11, 15–17, 20, 21).

[CL22] Eshan Chattopadhyay and Jyun-Jie Liao. "Extractors for sum of two sources". In: *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2022. New York, NY, USA: Association for Computing Machinery, June 2022, pp. 1584–1597. ISBN: 978-1-4503-9264-8. DOI: 10.1145/3519935.3519963 (cit. on p. 8).

[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. 1, 21).

[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. 2).

[Dod06] Yevgeniy Dodis. *Fault-tolerant leader election and collective coin-flipping in the full information model*. 2006 (cit. on p. 5).

[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. 22).

[DPW14] Yevgeniy Dodis, Krzysztof Pietrzak, and Daniel Wichs. "Key Derivation without Entropy Waste". In: *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*. Ed. by Phong Q. Nguyen and Elisabeth Oswald. Vol. 8441. Lecture Notes in Computer Science. Springer, 2014, pp. 93–110. DOI: 10.1007/978-3-642-55220-5_6 (cit. on p. 1).

[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. 1, 3).

[DMOZ24] Dean Doron, Dana Moshkovitz, Justin Oh, and David Zuckerman. "Online Condensing of Unpredictable Sources via Random Walks". In: *Electron. Colloquium Comput. Complex.* TR24-165 (2024). ECCC: TR24-165 (cit. on p. 3).

[DGW09] Zeev Dvir, Ariel Gabizon, and Avi Wigderson. "Extractors And Rank Extractors For Polynomial Sources". In: *Comput. Complex.* 18.1 (2009), pp. 1–58. DOI: 10.1007/S00037-009-0258-4 (cit. on p. 1).

[Fei99] U. Feige. "Noncryptographic selection protocols". In: *40th Annual Symposium on Foundations of Computer Science (Cat. No.99CB37039)*. ISSN: 0272-5428. Oct. 1999, pp. 142–152. DOI: 10.1109/SFFCS.1999.814586 (cit. on p. 5).

[GGL98] Oded Goldreich, Shafi Goldwasser, and Nathan Linial. "Fault-Tolerant Computation in the Full Information Model". In: *SIAM J. Comput.* 27.2 (1998), pp. 506–544. DOI: 10.1137/S0097539793246689 (cit. on pp. 5, 6).

[GSV05] Shafi Goldwasser, Madhu Sudan, and Vinod Vaikuntanathan. "Distributed Computing with Imperfect Randomness". In: *Distributed Computing, 19th International Conference, DISC 2005, Cracow, Poland, September 26-29, 2005, Proceedings*. Ed. by Pierre Fraigniaud. Vol. 3724. Lecture Notes in Computer Science. Springer, 2005, pp. 288–302. DOI: 10.1007/11561927_22 (cit. on p. 7).

[GLZ24] Jesse Goodman, Xin Li, and David Zuckerman. "Improved Condensers for Chor-Goldreich Sources". In: *Proceedings of the 65th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. To appear. 2024 (cit. on pp. 3, 16).

[GSZ21] Vipul Goyal, Akshayaram Srinivasan, and Chenzhi Zhu. "Multi-source Non-malleable Extractors and Applications". In: *Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part II*. Ed. by Anne Canteaut and François-Xavier Standaert. Vol. 12697. Lecture Notes in Computer Science. Springer, 2021, pp. 468–497. DOI: 10.1007/978-3-030-77886-6_16 (cit. on p. 7).

[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 p. 14).

[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 pp. 5, 7, 13, 29).

[KLRZ08] Yael Tauman Kalai, Xin Li, Anup Rao, and David Zuckerman. "Network Extractor Protocols". In: *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*. IEEE Computer Society, 2008, pp. 654–663. DOI: 10.1109/FOCS.2008.73 (cit. on p. 7).

[KRVZ11] Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. "Deterministic extractors for small-space sources". In: *Journal of Computer and System Sciences*. Celebrating Karp's Kyoto Prize 77.1 (Jan. 2011), pp. 191–220. ISSN: 0022-0000. DOI: 10.1016/j.jcss.2010.06.014 (cit. on p. 8).

[KZ07] Jesse Kamp and David Zuckerman. "Deterministic Extractors for Bit-Fixing Sources and Exposure-Resilient Cryptography". en. In: *SIAM Journal on Computing* 36.5 (Jan. 2007), pp. 1231–1247. ISSN: 0097-5397, 1095-7111. DOI: 10.1137/S0097539705446846 (cit. on p. 1).

[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. 21).

[Li23] Xin Li. "Two Source Extractors for Asymptotically Optimal Entropy, and (Many) More". In: *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2023, Santa Cruz, CA, USA, November 6-9, 2023*. IEEE, 2023, pp. 1271–1281. DOI: 10.1109/FOCS57990.2023.00075 (cit. on p. 8).

[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. 21).

[MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized algorithms*. Cambridge University Press, 1995 (cit. on p. 1).

[ODo14] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014 (cit. on p. 25).

[RZ01] Alexander Russell and David Zuckerman. "Perfect Information Leader Election in log* n+O (1) Rounds". In: *J. Comput. Syst. Sci.* 63.4 (2001), pp. 612–626. DOI: 10.1006/JCSS.2001.1776 (cit. on p. 5).

[TV00] Luca Trevisan and Salil P. Vadhan. "Extracting Randomness from Samplable Distributions". In: *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*. IEEE Computer Society, 2000, pp. 32–42 (cit. on p. 1).

[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 p. 1).

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il