On the Existence of Seedless Condensers: Exploring the Terrain

Eshan Chattopadhyay [1] Mohit Gurumukhani[1] Noam Ringach [1]

[1]Department of Computer Science, Cornell University

Extractors vs. Condensers

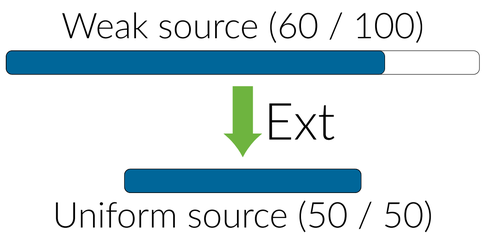
Definition (Min-Entropy). *The* min-entropy *of a distribution* $\mathbf{X}$ *is* $H_\infty(\mathbf{X}) = -\log(\max_{x \in support(\mathbf{X})} \Pr[\mathbf{X} = x])$.

Definition (Smooth Min-Entropy). *The* ε-smooth min-entropy *of* $\mathbf{X}$ *is* $H^\varepsilon_\infty(\mathbf{X}) = \max_{\mathbf{Y}:|\mathbf{X}-\mathbf{Y}| \leq \varepsilon} H_\infty(\mathbf{Y})$, *where* $|\cdot|$ *is the statistical distance.*

Definition (Extractor). *A function* $\mathsf{Ext} : \{0,1\}^n \to \{0,1\}^m$ *is a* ε-extractor *for a class* $\mathcal{X}$ *of distributions over* $\{0,1\}^n$ *if for all* $\mathbf{X} \in \mathcal{X}$ *we have*

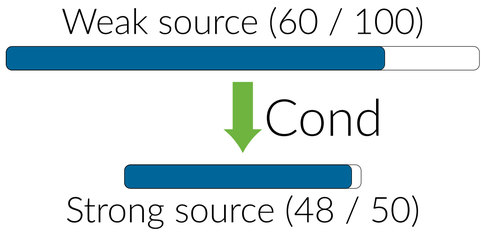
$$|\mathsf{Ext}(\mathbf{X}) - \mathbf{U}_m| \leq \varepsilon,$$

where $\mathbf{U}_m$ *the uniform distribution on* $\{0,1\}^m$.



Definition (Condenser). *A function* $\mathsf{Cond} : \{0,1\}^n \to \{0,1\}^m$ *is a* $(k_{in}, k_{out}, \varepsilon)$-condenser *for a class of sources* $\mathcal{X}$ *if for all* $\mathbf{X} \in \mathcal{X}$ *we have* $H_\infty(\mathbf{X}) \geq k_{in}$ *and*

$$H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq k_{out}.$$



The goal of condensers is to increase the **entropy rate**

$$\frac{k_{in}}{n} \lll \frac{k_{out}}{m}$$

and decrease the **entropy gap**

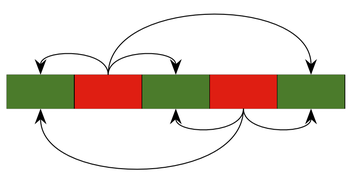
$$\Delta_{out} = m - k_{out}.$$

Remark. *Condensers can exist for certain classes of sources for which extractors* ***provably*** *can't.*

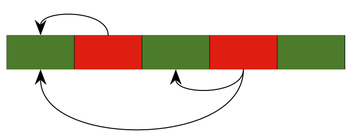
Adversarial Sources

Our main results are on two classes of sources where the difference between them is the power of the adversary.

Definition (Non-Oblivious Symbol Fixing Sources). *A* (g,ℓ)-NOSF source $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ *is a distribution on* $(\{0,1\}^n)^\ell$ *divided as* ℓ *blocks each of length* n *for which* g *of the blocks are "good" (independent and uniform on* $\{0,1\}^n$*) and* $\ell - g$ *blocks are "bad" (can depend arbitrarily on the* g *good blocks).*



Definition (online NOSFs). *A* (g,ℓ)-oNOSF source $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_\ell$ *is a* (g,ℓ)*-NOSF source where the bad blocks can only depend on the good blocks that come before them.*



Background

- Aggarwal, Obremski, Ribeiro, Siniscalchi, and Visconti (EUROCRYPT'20): extractors don't exist for (o)NOSFs with 99% good blocks.
- Previously, the existence of condensers for (o)NOSFs were completely unknown, even with 99% of blocks being good!

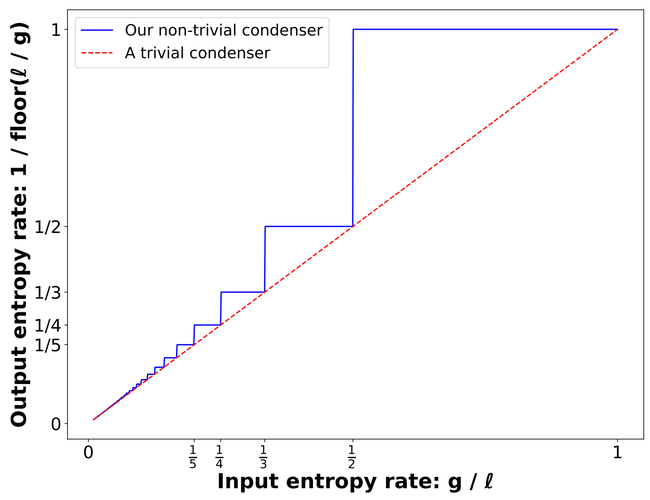
Result 1: Can't condense NOSFs

Theorem. *For all constant* g, $\ell \in \mathbb{N}$, $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *there exists a* (g,ℓ)*-NOSF* $\mathbf{X}$ *such that* $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{g}{\ell} \cdot m + O(1)$ *with* $\varepsilon = 0.99$.

Combined with the result of [AORSV'20], this shows the adversary of NOSFs is too strong for deterministic extraction or condensing.

Result 2: Can't condense oNOSFs above rate $1/\lfloor \ell/g \rfloor$

Theorem. *For all constant* g, $\ell \in \mathbb{N}$, $f : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *there exists a* (g,ℓ)*-oNOSF* $\mathbf{X}$ *such that* $H^\varepsilon_\infty(f(\mathbf{X})) \leq \frac{1}{\lfloor \ell/g \rfloor} \cdot m + O(1)$ *with* $\varepsilon = 0.99$.



Idea 1: Use dominating sets to prove impossibility for $(1,2)$-oNOSF

Goal: Construct a $(1,2)$-oNOSF $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ s.t. w.p. **0.99**, $f(\mathbf{X}_1, \mathbf{X}_2) \in D$ where $|D| \leq \boldsymbol{O}\left(\mathbf{2^{m/2}}\right)$.

Lemma (Dominating set lemma). *Let* $G = (U, V)$ *be bipartite graph with* $U = [N], V = [M]$ *s.t. for all* $u \in U : \deg(u) \geq \sqrt{M}$. *Then, there exists* $D \subset V$ *s.t.* $|D| = O(\sqrt{M})$ *and* $|Nbr(D)| \geq 0.99N$.

Let $G = (U, V) : U = \{0,1\}^n, V = \{0,1\}^m$, with the edge $(u, v) \in G$ if there exists a s.t. $f(u, a) = v$.

Case 1: Exists $u^* \in U$ s.t. $\deg(u^*) \leq 2^{m/2}$

Let $\mathbf{X} = (u^*, \mathbf{U}_n)$. With probability 1, $f(\mathbf{X}) \in Nbr(u^*)$ and $|Nbr(u^*)| \leq 2^{m/2}$.

Case 2: For all $u \in U$, $\deg(u) \geq 2^{m/2}$, so use lemma

Exists $D \subset V$ s.t. $|D| \leq O\left(2^{m/2}\right)$ and $|Nbr(D)| \geq 0.99 \cdot 2^n$. Let $\mathbf{X} = (\mathbf{U}_n, \mathsf{Adv}(\mathbf{U}_n))$ where $\mathsf{Adv}(u) = a$ s.t. $f(u, a) \in D$ (if it exists). W.p. 0.99, $f(\mathbf{X}) \in D$.

Open questions

- Can these condensers be made explicit?
- What's the situation for other regimes, like $n = O(1)$ and increasing ℓ?

Result 3: Can condense oNOSFs to rate $1/\lfloor \ell/g \rfloor$

Theorem. *For any constant* g, $\ell \in \mathbb{N}$ *and* ε, *there exists a condenser* $\mathsf{Cond} : (\{0,1\}^n)^\ell \to \{0,1\}^m$ *s.t. for any* (g,ℓ)*-oNOSF* $\mathbf{X}$ *we have* $H^\varepsilon_\infty(\mathsf{Cond}(\mathbf{X})) \geq \frac{1}{\lfloor \ell/m \rfloor} \cdot m - O(\log(m))$ *where* $m = \Omega(n)$ *and* $\varepsilon = \Omega(m^{-1/4})$.

Corollary. *For constant* g, ℓ *such that* $g > \ell/2$, *we can condense* (g,ℓ)*-oNOSFs to rate* $1 - o(1)$.

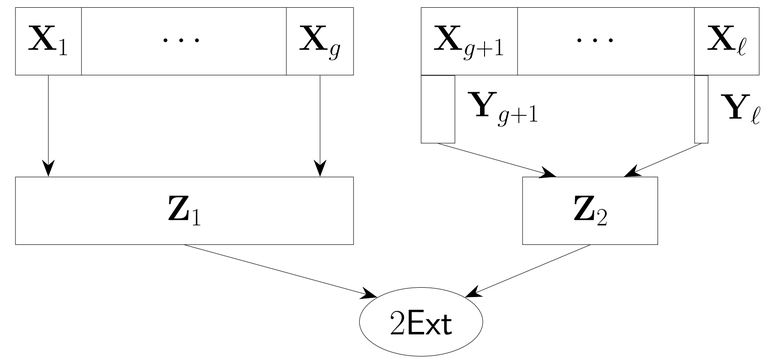
Idea 2: Weaken adversary in later blocks

Goal: Create a function $\mathsf{Cond}(x_1, \ldots, x_\ell)$ that condenses to entropy rate $1 - o(1)$ when $g > \ell/2$.

Problem: The adversarial blocks with higher indices are more "powerful" since they are able to depend on more good blocks than earlier ones.

Solution: To weaken the bad blocks, we take prefixes of geometrically decreasing length.

- For $i > g$, let $y_i := x_i[1] \circ \cdots \circ x_i[20^{\ell - i + 1}(\log n)]$.
- Let $z_1 = x_1 \circ \cdots \circ x_g$, $z_2 = y_{g+1} \circ \cdots \circ y_\ell$.
- Let $\mathsf{Cond}(x_1, \ldots, x_\ell) = 2\mathsf{Ext}(z_1, z_2)$.



Idea 3: Use an output-light two-source extractor

Definition (Output-light two-source extractor). $\mathsf{2Ext} : \{0,1\}^{n_1+n_2} \to \{0,1\}^m$ *is a* R-output-light (k_1, k_2, ε)-two-source extractor *if for independent* (n_1, k_1)*-source* $\mathbf{X}_1$ *and* (n_2, k_2)*-source* $\mathbf{X}_2$*:*

$$|\mathsf{2Ext}(\mathbf{X}_1, \mathbf{X}_2) - \mathbf{U}_m| \leq \varepsilon$$

and for every $w \in \{0,1\}^m$ *we have* $\left|\mathsf{2Ext}^{-1}(w)\right| \leq R$.

Goal: Condense any (g,ℓ)-oNOSF source $\mathbf{X}$ with $g > \ell/2$ to rate $1 - o(1)$.

Case 1: $\mathbf{X}_{g+1}, \ldots, \mathbf{X}_\ell$ are bad

1. For the sake of contradiction, can't condense, so output is in some small set D with ε weight.
2. Implies many inputs to $2\mathsf{Ext}$ with output in D.
3. By an averaging argument, exists an elt in D with large preimage, contradicting output-lightness.

Case 2: Exists $j > g$ s.t. $\mathbf{X}_j$ is uniform

1. $g > \ell/2$, so exists $i \leq g$ s.t. $\mathbf{X}_i$ is uniform.
2. Fix output of $\mathbf{Y}_{g+1}, \ldots, \mathbf{Y}_{j-1}$.
3. By chain rule $H_\infty(\mathbf{Z}_1) \geq k_1$.
4. $\mathbf{X}$ is oNOSF implies $\mathbf{Y}_j$ uniform.
5. If $\mathbf{Y}_{j+1}, \ldots, \mathbf{Y}_j$ uniform, then output is uniform!
6. However, they may all be bad, so output is high entropy instead of uniform.