On the Complexity of Decomposable Randomized Encodings, or: How Friendly Can a Garbling-Friendly PRF be?

Marshall Ball
Columbia University
marshall@cs.columbia.edu

Justin Holmgren
Simons Institute
holmgren@alum.mit.edu

Yuval Ishai
Technion
yuvali@cs.technion.ac.il

Tianren Liu
University of Washington
liutr@mit.edu

Tal Malkin
Columbia University
tal@cs.columbia.edu

Abstract

Garbling schemes, also known as decomposable randomized encodings (DRE), have found many applications in cryptography. However, despite a large body of work on constructing such schemes, very little is known about their limitations.

We initiate a systematic study of the DRE complexity of Boolean functions, obtaining the following main results:

- **Near-quadratic lower bounds**. We use a classical lower bound technique of Nečiporuk [Dokl. Akad. Nauk SSSR '66] to show an $\Omega(n^2/\log n)$ lower bound on the size of any DRE for many explicit Boolean functions. For some natural functions, we obtain a corresponding upper bound, thus settling their DRE complexity up to polylogarithmic factors. Prior to our work, no superlinear lower bounds were known, even for non-explicit functions.
- **Garbling-friendly PRFs.** We show that any exponentially secure PRF has $\Omega(n^2/\log n)$ DRE size, and present a plausible candidate for a "garbling-optimal" PRF that nearly meets this bound. This candidate establishes a barrier for super-quadratic DRE lower bounds via natural proof techniques. In contrast, we show a candidate for a *weak* PRF with near-exponential security and linear DRE size.

Our results establish several qualitative separations, including near-quadratic separations between computational and information-theoretic DRE size of Boolean functions, and between DRE size of weak vs. strong PRFs.

2012 ACM Subject Classification Theory of computation → Computational complexity and cryptography

Keywords and phrases Randomized Encoding, Private Simultaneous Messages

Digital Object Identifier 10.4230/LIPIcs...

Funding The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either express or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

Marshall Ball: Supported by an IBM Research PhD Fellowship. This work is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA) via Contract No. 2019-1902070006.
Yuval Ishai: Supported by ERC Project NTSC (742754), NSF-BSF grant 2015782, BSF grant 2018393, and a grant from the Ministry of Science and Technology, Israel and Department of Science and Technology, Government of India.
Tianren Liu: This work was mostly done in Massachusetts Institute of Technology. Research supported by NSF Grants CNS-1350619, CNS-1414119 and CNS-1718161, an MIT-IBM grant and Vinod Vaikuntanathan's DARPA Young Faculty Award.
Tal Malkin: This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA) via Contract No. 2019-1902070006.

1 Introduction

Originating from Yao's garbled circuit construction [65], garbling schemes have played an important role in different sub-areas of cryptography. A garbled representation of $f(x)$ is a randomized function $\hat{f}(x;r)$ such that: (1) a sample from the output of $\hat{f}(x;r)$ reveals $f(x)$ and no additional information about x; and (2) each output bit of $\hat{f}$ depends on at most *one* bit of x (but can depend arbitrarily on r); equivalently, each bit of x acts as a selector between two strings that are determined by r. We refer to such a garbled representation $\hat{f}$ for f as a *decomposable randomized encoding* (DRE)[1] of f, and refer to the output length of $\hat{f}$ as its *size*.

Garbling schemes were initially motivated by the goal of efficient secure computation [65, 44, 30, 40]. This still serves as a primary motivation for their study, which has led to many optimized constructions (see, e.g., [12] and references therein).

Over the years, different flavors of garbling schemes have found applications in many other areas of cryptography, including parallel cryptography [8], one-time programs and leakage-resilient cryptography [36], verifiable computation [33, 10], key-dependent message security [13, 5], identity-based encryption [29], and more. See [18, 39, 6] for surveys.

Despite the large body of work on constructing and applying such garbling schemes, very little is known about their *limitations*. Previous relevant works show very limited lower bounds for more liberal notions of garbling. These include either conditional lower bounds that apply to computationally efficient garbling of intractable functions [5, 1] or linear size lower bounds for so-called "2-party PSM protocols" [30, 25, 7].

In this work, we initiate a complexity theoretic study of standard ("DRE-style") garbling schemes, providing *lower bounds* in both *information-theoretic* and *computational settings*.

1.1 Our Contribution

We make two types of contributions: (1) obtaining the first super-linear lower bounds on the DRE size of Boolean functions (with some matching upper bounds), and (2) studying the minimal DRE size of (strong and weak) pseudorandom functions. We now detail both types of results.

[1] This notion of garbling roughly corresponds to a *projective garbling scheme* in the terminology of Bellare et al. [18]. We use the DRE terminology when we want to emphasize that we are not interested in the process of "garbling" a given representation of f, but only in the *existence* of a garbled representation $\hat{f}$ with a given complexity.

1.1.1 Near-quadratic lower bounds and matching upper bounds

We adapt a classical lower bound technique of Nečiporuk [49] to show an $\Omega(n^2/\log n)$ lower bound on the size of any DRE for many explicit Boolean functions $f : \{0,1\}^n \to \{0,1\}$. Nečiporuk showed that functions with many subfunctions cannot have small formulas or branching programs. We provide matching lower bounds on DRE for the same class. In particular, this yields $\Omega(n^2/logn)$ lower bounds on DRE size for almost all functions, including the explicit examples of Element Distinctness, Indirect Storage Access, Clique, Determinant, Matching, and others. These bounds hold in both the information theoretic setting and the exponentially-secure computational setting, provided the DRE admits a sub-exponential decoding algorithm in the latter case.

For the explicit example of Element Distinctness, we obtain a corresponding upper bound, thus settling its DRE complexity up to polylogarithmic factors. Furthermore, since some of the functions that admit nearly quadratic lower bounds on DRE size can be computed by linear-size circuits, our lower bounds establish a near-quadratic gap between the size of computationally secure and information-theoretic DRE in a setting where the input size is polynomially bigger than the computational security parameter. In fact, given that our nearly quadratic lower bounds also apply to computational DREs with security parameter nearly that of the input size, this means, in a concrete sense, that a tradeoff between DRE size and security parameter is inherent!

The only previous lower bounds we are aware of are *linear* lower bounds that also apply to the more liberal 2-party Private Simultaneous Messages (PSM)[2] setting [30, 25, 7] and quadratic lower bounds for *non-Boolean* functions $f : \{0,1\}^n \to \{0,1\}^n$ that follow from the input locality lower bounds of [9]. In contrast to the other classes lower bounded by Nečiporuk's method, such as formulas and branching programs, no superlinear lower bounds on DRE size were known prior to our work, even for a *non-explicit* (e.g., random or worst-case) Boolean function.

1.1.2 Garbling-friendly PRFs

There is a recent line of work on "MPC-friendly" block ciphers [3, 37, 54, 28, 27, 2] and pseudorandom functions (PRFs) [48, 32, 37, 19]. In this context, DRE size is a highly relevant complexity measure that is often used as a benchmark. The question of minimizing the DRE size of PRFs is motivated by the goal of securely evaluating a PRF in a setting where the input (and possibly also the key) is secret-shared between two or more parties. This is useful for natural applications that involve secure keyword search and distributed forms of searchable symmetric encryption; see [19] for discussion.

For the case of exponentially secure (strong) PRFs, we show that the DRE size must be $\Omega(n^2/\log n)$.[3] Finally, we conjecture that a candidate PRF based on the "hidden shift problem" is exponentially secure PRF with almost matching DRE size $O(n^2)$. That is, the function outputs the quadratic character of a hidden shift of the input, determined by the secret key. To defeat known attacks (both quantum and classical), we restrict inputs bounded interval rather than the entire domain. A similar PRF (without the input restriction) has been proposed in [37] as an attractive candidate for MPC-friendly PRF, but in an interactive

[2] A DRE can be viewed as an n-party PSM protocol in which each party holds just one bit. Any 2-party PSM lower bound implies a similar DRE lower bound, but the converse is not true.

[3] This is almost immediate in the non-uniform setting, given our lower bounds. In the appendix we give a constructive proof for this fact in the uniform setting by exhibiting a sublinear test for an average-case variant of the natural property used in Nečiporuk's method.

setting of arithmetic MPC, rather than in the context of garbling. We also present a similar PRF construction with $\Omega(n)$ bits of output, for which we can still obtain a near-quadratic DRE size upper bound.[4] Consequently, modulo the validity of the conjectured security, these PRFs are nearly garbling-optimal.

Interpreted differently, our garbling-friendly PRF candidate establishes a barrier for super-quadratic DRE lower bounds on *explicit* Boolean functions via *natural proof* techniques [53]. In contrast, we show that a recent candidate for a *weak* PRF with near-exponential security due to Boneh et al. [19] has a *linear* DRE size.

Our results imply several qualitative separations, including near-quadratic separations between computational and information-theoretic DRE size of Boolean functions, and between the DRE size of weak vs. strong PRFs.

2 Preliminaries

2.1 Cryptography

▶ **Definition 1** (Pseudorandom Functions [35]). *An* $(s(\cdot), \delta(\cdot))$-secure pseudorandom function (PRF) family *is an ensemble* $\mathcal{F} = \{\mathcal{F}_\lambda\}_{\lambda \in \mathbb{Z}^+}$, *where each* $\mathcal{F}_\lambda$ *is a keyed family of functions* $\mathcal{F}_\lambda = \{f_k : \{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)}\}_{k \in \{0,1\}^{\kappa(\lambda)}}$, *satisfying the following security property:*

Pseudorandomness *For every* $\lambda \in \mathbb{Z}^+$ *and every size-*s *(ensemble) of oracle circuits* $\mathcal{A}$ *(with output in* $\{0,1\}$*),*

$$\left| \mathop{\mathbb{E}}_{\substack{k \leftarrow \{0,1\}^{\kappa(\lambda)} \\ U:\{0,1\}^{n(\lambda)} \to \{0,1\}^{m(\lambda)}}} [\mathcal{A}^{f_k}(1^\lambda) - \mathcal{A}^U(1^\lambda)] \right| \le \delta(\lambda).$$

$n(\cdot)$, $m(\cdot)$, *and* $\kappa(\cdot)$ *are respectively called the* input length, output length, *and* key length *of* $\mathcal{F}$.

▶ **Definition 2** (Weak PRFs [34]). *An* $(s(\cdot), \delta(\cdot))$-secure weak PRF family *is a relaxation of a PRF family as in Definition 1, with the "pseudorandomness" security property replaced by the following notion of "weak pseudorandomness":*

Weak Pseudorandomness *For every* λ, *the tuples*

$$\left(X_1, \ldots, X_{s(\lambda)}, f_K(X_1), \ldots, f_K(X_{s(\lambda)})\right)$$

and

$$\left(X_1, \ldots, X_{s(\lambda)}, Y_1, \ldots, Y_{s(\lambda)}\right)$$

are $(s(\lambda), \delta(\lambda))$*-indistinguishable in the probability space defined by sampling*

$$\begin{aligned} &K \leftarrow \{0,1\}^{\ell(\lambda)} \\ &X_1, \ldots, X_{s(\lambda)} \leftarrow \{0,1\}^{n(\lambda)} \\ &Y_1, \ldots, Y_{s(\lambda)} \leftarrow \{0,1\}^{m(\lambda)}. \end{aligned}$$

▶ **Definition 3.** *Random variables* X *and* Y *are* (s, ϵ)*-indistinguishable if the advantage of every size-*s *circuit in distinguishing* X *from* Y *is at most* ϵ. *We denote this by* $X \approx^{(s,\epsilon)} Y$.

[4] For this case, multi-bit output, we use input locality bounds of [9] to prove a slightly stronger (and nearly tight) quadratic lower bound (contrast with our $\Omega(n^2/\log n)$ bounds for single bit output).

2.2 Information Theory

▶ **Definition 4.** *The* min-entropy *of a random variable X is* $H_\infty(X) \stackrel{\text{def}}{=} \min_{x \in \text{Supp}(X)} \log_2 \left(\frac{1}{\Pr[X=x]} \right)$.

2.3 Decomposable Randomized Encodings

▶ **Definition 5** (Randomized Encodings)**.** *A* randomized encoding *for a function $f : \{0,1\}^n \to \mathcal{Y}$ consists of a "randomness" distribution $\mathcal{R}$, an encoding function* $\mathsf{Enc} : \{0,1\}^n \times \mathcal{R} \to \{0,1\}^\ell$, *and a decoding function* $\mathsf{Dec} : \{0,1\}^\ell \to \mathcal{Y}$. *$\ell$ is called* the size of the randomized encoding.

A randomized encoding $(\mathcal{R}, \mathsf{Enc}, \mathsf{Dec})$ *for function $f : \{0,1\}^n \to \mathcal{Y}$ should satisfy:*

Correctness *For any input $x \in \{0,1\}^n$,*

$$\Pr_{R \leftarrow \mathcal{R}}[\mathsf{Dec}(\mathsf{Enc}(x,R)) = f(x)] = 1.$$

Security *For all $x, y \in \{0,1\}^n$ with $f(x) = f(y)$, the distribution of* $\mathsf{Enc}(x,R)$ *is identical to the distribution of* $\mathsf{Enc}(y,R)$ *when sampling $R \leftarrow \mathcal{R}$.*

The security can be relaxed to require only that $\mathsf{Enc}(x,R)$ *and* $\mathsf{Enc}(y,R)$ *cannot be effectively distinguished by small circuits.*

(s,δ)-***Security*** *For all $x, y \in \{0,1\}^n$ such that $f(x) = f(y)$, for any circuit $\mathcal{A} : \{0,1\}^\ell \to \{0,1\}$ of size at most s,*

$$\Big| \Pr_{R \leftarrow \mathcal{R}}[\mathcal{A}(\mathsf{Enc}(x,R)) = 1] - \Pr_{R \leftarrow \mathcal{R}}[\mathcal{A}(\mathsf{Enc}(y,R)) = 1] \Big| \le \delta.$$

In this paper, we focus on decomposable randomized encoding (DRE)*, which is a randomized encoding that also satisfies an additional property:*

Decomposability *Each output bit of* $\mathsf{Enc}(x,r)$ *is determined by r and 1 bit of input x.*

To ease presentation, we also introduce an equivalent definition of DRE. The equivalent definition is used when we prove lower bounds on the size of DRE.

▶ **Definition 6.** *An (s,δ)-secure* decomposable randomized encoding (DRE) for a function $f : \{0,1\}^n \to \{0,1\}$ *is a family of random variables*

$$\mathcal{X} = \begin{pmatrix} \mathcal{X}_0^1, \ldots, \mathcal{X}_0^n \\ \mathcal{X}_1^1, \ldots, \mathcal{X}_1^n \end{pmatrix}$$

such that

Correctness *There is an algorithm* Dec *such that for every $x \in \{0,1\}^n$,*

$$\Pr[\mathsf{Dec}(\mathcal{X}_{x_1}^1, \ldots, \mathcal{X}_{x_n}^n) = f(x)] = 1.$$

Dec *is called a* decoding algorithm *for $\mathcal{X}$.*

(s,δ)-***Security*** *For all $x, y \in \{0,1\}^n$ such that $f(x) = f(y)$,*

$$(\mathcal{X}_{x_1}^1, \ldots, \mathcal{X}_{x_n}^n) \approx^{(s,\delta)} (\mathcal{X}_{y_1}^1, \ldots, \mathcal{X}_{y_n}^n).$$

The size of $\mathcal{X}$ *is*

$$|\mathcal{X}| \stackrel{\text{def}}{=} \sum_{i \in [n], b \in \{0,1\}} \log_2 \big| \text{Supp}(\mathcal{X}_b^i) \big|.$$

2.4 Function Restrictions

▶ **Definition 7** ([50]). *For any function $f : X^n \to Y$, any set $S \subseteq [n]$ with complement $\bar{S}$, and any $z \in X^{\bar{S}}$, the* restriction of f to S using z *is the function*

$$f_{S|z} : X^S \to Y$$

defined by fixing the coordinates in $\bar{S}$ to the value z. More formally, for any $x \in X^S$, we define

$$f_{S|z}(x) \stackrel{\mathsf{def}}{=} f(x'),$$

where for each $i \in [n]$,

$$x'_i = \begin{cases} x_i & \text{if } i \in S \\ z_i & \text{otherwise.} \end{cases}$$

3 Lower Bounds on DRE Size

Over 50 years ago, Nečiporuk published a two-page note titled "On a boolean function." [49] Within these two pages, Nečiporuk introduced an elegant combinatorial measure of a function related to the number of ways a function can be restricted distinctly. To this day, Nečiporuk's method still provides the strongest lower bounds known for formulas over arbitrary finite bases, deterministic branching programs, non-deterministic branching programs, parity branching programs, switching networks, span programs, and more [16].

In this section we recall Nečiporuk's measure and add decomposable randomized encoding (DRE) size to the list of complexity measures that are lower bounded by Nečiporuk's measure. Specifically, we show that for any function f, the DRE complexity of f is at least Nečiporuk's measure (which for explicit functions is as large as $n^2/\log n$). Prior to this work no super linear lower bounds on DRE size were known.

3.1 Technical Overview

To lower bound the DRE size of a function $f : \{0,1\}^n \to \{0,1\}$, we first consider all possible restrictions of f, using notation as in Definition 7. For simplicity, suppose that

$$\mathcal{X} = \begin{pmatrix} \mathcal{X}_0^1, \ldots, \mathcal{X}_0^n \\ \mathcal{X}_1^1, \ldots, \mathcal{X}_1^n \end{pmatrix}$$

is a *perfect* DRE for f. Then for all $S \subseteq [n]$ (with complement denoted by $\bar{S}$), we observe that:

1. The distribution of $(\mathcal{X}_{z_i}^i)_{i \in \bar{S}}$ does not depend on $z \in \{0,1\}^{\bar{S}}$ (as long as $f_{S|z}$ is non-constant). This follows from DRE security.
2. Given $(\mathcal{X}_{z_i}^i)_{i \in \bar{S}}$, the values $(X_b^i)_{i \in S, b \in \{0,1\}}$ are sufficient to reconstruct the truth table of $f_{S|z}$. This follows from DRE correctness.

Together, these properties imply that the size of the support of $(X_b^i)_{i \in S, b \in \{0,1\}}$ is at least the number of non-constant truth tables of the form $f_{S|z}$ for some $z \in \{0,1\}^{\bar{S}}$. We obtain a bound on the size of $\mathcal{X}$ by partitioning $[n]$ into sets $S_1, \ldots, S_m$, and lower bounding the size of each $(\mathcal{X}_b^i)_{i \in S_j, b \in \{0,1\}}$. The maximum bound on the *bit length* of $\mathcal{X}$ that can be achieved in this way is essentially Nečiporuk's measure of f.

We elaborate further below, defining a somewhat more general computational analogue of Nečiporuk's measure (that will suffice for lower bounds on computationally secure DREs).

3.2 Nečiporuk's Measure

Let $f : \{0,1\}^n \to \{0,1\}$ be any boolean function. For any subset $S \subseteq [n]$, let $\bar{S}$ denote $[n] \setminus S$, and define

$$g_S(f) \stackrel{\text{def}}{=} \log(\#\{f_{S|z} : z \in \{0,1\}^{\bar{S}}\}).$$

Let $V = (V_1, \dots, V_m)$ denote a partition of $[n]$. That is, $V_1, \dots, V_m$ are pairwise disjoint subsets of $[n]$ whose union is $[n]$. Then, the Nečiporuk measure of f is

$$G(f) \stackrel{\text{def}}{=} \max_V \sum_{V_i \in V} g_{V_i}(f).$$

▶ Remark 8. It is well known that for any function f, $G(f) \le n^2/\log n$ [57].

3.3 Functions with Maximal Measure

We recall several functions whose Nečiporuk measures are known to be as high as possible ($\Omega(n^2/\log n)$, where n is the bit-length of the input).

Element Distinctness.

Element Distinctness is a function $\mathrm{ED}_m : [m^2]^m \to \{0,1\}$ which given a vector $(x_1, \dots, x_m) \in [m^2]^m$ and outputs 1 if all x_i are distinct and 0 otherwise ($\exists i \ne j$ such that $x_i = x_j$).

Others.

Clique, matching, and determinant all have measure $\Omega(n^2/\log n)$ [57].

Random.

Finally, and perhaps unsurprisingly, we note that a random function has measure at least $\frac{n(n-2)}{\log n}$) with overwhelming probability (for n large enough). See Appendix B for proof.

3.4 DRE Size Lower Bounds via Nečiporuk

We define a pseudo-min-entropic analogue of Nečiporuk's measure, with an additional non-constantness restriction that is tailored for use in DRE lower bounds.

▶ **Definition 9.** *The (s,ϵ)-pseudo min-entropy of a random variable X, which we will denote by $\tilde{H}_\infty^{(s,\epsilon)}(X)$, is the supremum of $H_\infty(\tilde{X})$ over all random variables $\tilde{X}$ that are (s,ϵ)-indistinguishable from X.*

▶ **Definition 10.** *For any function $f : \{0,1\}^n \to \{0,1\}$ and any subset $\emptyset \ne V \subseteq [n]$, define*

$$\tilde{G}_V^{(s,\epsilon)}(f) \stackrel{\text{def}}{=} \sup\left(\tilde{H}_\infty^{(s,\epsilon)}(f_{V|Z})\right),$$

where the supremum is taken over all $\{0,1\}^{\bar{V}}$-valued random variables Z whose support only consists of values z that make $f_{V|z}$ non-constant.

We define $\tilde{G}^{(s,\epsilon)}(f)$ to be the maximum over all partitions $[n] = V_1 \cup \cdots \cup V_m$ of $\sum_{i\in[m]} \tilde{G}_{V_i}^{(s,\epsilon)}(f)$.

▶ Remark 11. If not for the non-constantness constraint on $f_{V|Z}$, the measure $\tilde{G}^{(\infty,0)}$ is the same as Nečiporuk's original measure. Reducing s or increasing ϵ only increases this measure. Taking the non-constantness restriction into account, our measure cannot be smaller than Nečiporuk's measure by more than $O(n)$ (so superlinear lower bounds on Nečiporuk's measure imply an asymptotically identical lower bound on our measure).

Beyond a certain threshold, increasing s no longer changes the value of $\tilde{G}^{(s,\epsilon)}$:

▷ Claim 12. For any function $f : \{0,1\}^n \to \{0,1\}$ and any subset $V \subseteq [n]$, we have $\tilde{G}_V^{(\infty,\epsilon)}(f) = \tilde{G}_V^{\left(2^{2^{|V|}},\epsilon\right)}(f)$.

Proof. Any function of n bits can be computed by a circuit of size 2^n. In fact this can be strengthened to $O(\frac{2^n}{n})$ [59, 45], but we prefer the simpler bound 2^n. Apply this to the (s,ϵ)-indistinguishability in the definition of pseudo-min-entropy of $f_{V|Z}$ (which is a truth table of bit length $n = 2^{|V|}$). ◀

Our main lower bound is given by the following theorem.

▶ **Theorem 13.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a function, and let $\mathcal{X}$ be a $(s^*_{\mathsf{DRE}}, \frac{1}{3})$-secure DRE for f with a decoding algorithm of size s_{Dec}.*

Then for all $V \subseteq [n]$, we have

$$|\mathcal{X}^V| \geq \min\left(\log_2\left(\frac{s^*_{\mathsf{DRE}}}{s_{\mathsf{Dec}} \cdot 2^{|V|}}\right), \tilde{G}_V^{(s^*_{\mathsf{DRE}},\frac{1}{3})}(f) - 2\right).$$

Proof. Suppose otherwise — that $|\mathcal{X}^V| < \log_2\left(\frac{s^*_{\mathsf{DRE}}}{s_{\mathsf{Dec}} \cdot 2^{|V|}}\right)$ and $|\mathcal{X}^V| < \tilde{G}_V^{(s^*_{\mathsf{DRE}},\frac{1}{3})}(f) - 2$.

Let Z be a $\{0,1\}^{\bar{V}}$-valued random variable that maximizes $\tilde{H}_\infty^{(s^*_{\mathsf{DRE}},\frac{1}{3})}(f_{V|Z})$, supported by values z for which $f_{V|z}$ is non-constant, and let $\tilde{F}_V$ denote a random variable that is $(s^*_{\mathsf{DRE}}, \frac{1}{3})$-indistinguishable from $f_{V|Z}$ and satisfies $H_\infty(\tilde{F}_V) = \tilde{H}_\infty^{(s^*_{\mathsf{DRE}},\frac{1}{3})}(f_{V|Z})$. Let Z' be an independent copy of Z.

We first claim that $(\mathcal{X}_Z^{\bar{V}}, f_{V|Z}) \approx^{(s^*_{\mathsf{DRE}},\frac{1}{3})} (\mathcal{X}_{Z'}^{\bar{V}}, f_{V|Z})$. To see why, suppose for contradiction that there is size-s^*_{DRE} circuit $\mathcal{A}$ that distinguishes $(\mathcal{X}_Z^{\bar{V}}, f_{V|Z})$ from $(\mathcal{X}_{Z'}^{\bar{V}}, f_{V|Z})$ with advantage better than $\frac{1}{3}$. Then in particular there exist $z, z' \in \{0,1\}^{\bar{V}}$ such that $\mathcal{A}$ distinguishes $(\mathcal{X}_z^{\bar{V}}, f_{V|z})$ from $(\mathcal{X}_{z'}^{\bar{V}}, f_{V|z})$ with the same advantage. Hardwiring $f_{V|z}$ into $\mathcal{A}$, this gives a circuit $\mathcal{B}$ of size[5] $|\mathcal{B}| \leq |\mathcal{A}|$ for distinguishing $\mathcal{X}_z^{\bar{V}}$ from $\mathcal{X}_{z'}^{\bar{V}}$ with the same advantage. But this contradicts the $(s^*_{\mathsf{DRE}}, \frac{1}{3})$-indistinguishability that is guaranteed by DRE security.

We also know that $(\mathcal{X}_{Z'}^{\bar{V}}, f_{V|Z}) \approx^{(s^*_{\mathsf{DRE}},\frac{1}{3})} (\mathcal{X}_{Z'}^{\bar{V}}, \tilde{F}_V)$, so together with the previous claim, we have $(\mathcal{X}_Z^{\bar{V}}, f_{V|Z}) \approx^{(s^*_{\mathsf{DRE}},\frac{2}{3})} (\mathcal{X}_{Z'}^{\bar{V}}, \tilde{F}_V)$. However, there is a distinguisher that contradicts this. Specifically, try all possible values of $\left(\mathcal{X}_b^i\right)_{i \in V, b \in \{0,1\}}$ (there are at most $\frac{s^*_{\mathsf{DRE}}}{s_{\mathsf{Dec}} \cdot 2^{|V|}}$ possibilities), and apply the DRE decoding algorithm ($2^{|V|}$ times per possibility) to see whether any possibility "explains" the given truth table.

By correctness of the DRE, there will always exist a value that explains $f_{V|Z}$ given $\mathcal{X}_Z^{\bar{V}}$, but because $H_\infty(\tilde{F}_V) > \log_2 |\mathcal{X}^V| + 2$, the probability that any value explains $\tilde{F}_V$ is at most $\frac{1}{4}$. Hence the distinguisher succeeds with probability $\frac{3}{4} > \frac{2}{3}$, which is a contradiction. ◀

[5] Recall that the size of a circuit is measured in number of gates, and all gates of $\mathcal{A}$ whose inputs are the hard-wired value $f_{V|z}$ can be simplified or eliminated.

3.5 The Nečiporuk Measure of PRFs

In this section, we prove lower bounds on the Nečiporuk measure of PRFs (of varying security levels), which imply corresponding lower bounds on the size of DREs.

▶ **Proposition 14.** *If $E : \{0,1\}^{\kappa+n} \to \{0,1\}$ is an (s,ϵ)-secure PRF with key length κ and input length n satisfying $s \geq 4$ and $\epsilon \leq \frac{1}{6}$, then for any subset $V \subseteq [\kappa+1, \kappa+n]$ with $|V| \geq 2$, we have $\tilde{G}_V^{(s,\epsilon')}(E) = 2^{|V|}$ for $\epsilon' = 3\epsilon + 2^{-s+1} + 2^{-2^{|V|}+1}$.*

Proof. Let Z' be a $\{0,1\}^{\bar{V}}$-valued random variable whose first κ coordinates are independent and uniformly random, and the rest of whose coordinates are 0. By PRF security, the probability that $E_{V|Z'}$ is constant is at most $\delta \stackrel{\text{def}}{=} \epsilon + 2^{-\min(s,2^{|V|})+1} \leq \epsilon + 2^{-s+1} + 2^{-2^{|V|}+1} \leq \frac{1}{2}$.

Let $\mathcal{A}$ be an arbitrary size-s circuit. Suppose for contradiction that $\mathcal{A}$ distinguishes $E_{V|Z'}$ from a uniformly random truth table with advantage greater than ϵ. Then each input wire of $\mathcal{A}$ can be replaced by an oracle gate to yield a circuit that distinguishes oracle access to $E(K,\cdot)$ (for uniform K) from oracle access to a uniformly random function with the same advantage ϵ. This contradicts (s,ϵ)-security of the PRF. So $E_{V|Z'}$ is (s,ϵ)-indistinguishable from a uniformly random truth table.

Conditioned on $E_{V|Z'}$ being non-constant, the advantage of any $\mathcal{A}$ in distinguishing $E_{V|Z'}$ from a uniformly random truth table can increase to at most

$$\frac{\frac{1}{2}+\epsilon}{1-\delta} - \frac{1}{2} \leq (\frac{1}{2}+\epsilon)\cdot(1+2\delta) - \frac{1}{2} = \epsilon + \delta + 2\epsilon\cdot\delta \leq 3\epsilon + 2^{-s+1} + 2^{-2^{|V|}+1}.$$

Thus if Z denotes the random variable Z' conditioned on $E_{V|Z'}$ being non-constant, we have $\tilde{H}^{(s,\epsilon')}(E_{V|Z}) = 2^{|V|}$ for $\epsilon' = 3\epsilon + 2^{-s+1} + 2^{-2^{|V|}+1}$. ◀

▶ **Corollary 15.** *If $E : \{0,1\}^{\kappa+n} \to \{0,1\}$ is the evaluation algorithm for an (s,ϵ)-secure PRF family with key length κ and input length n satisfying $s \geq 4$ and $\epsilon \leq \frac{1}{6}$, then $\tilde{G}^{(\infty,\epsilon')}(E) \geq \Omega\left(\frac{n\log s}{\log\log s}\right)$ for $\epsilon' = 3\epsilon + 2^{-s+2}$. In particular, if the PRF family is exponentially secure, then $\tilde{G}^{(\infty,\epsilon')}(E) \geq \Omega\left(\frac{n^2}{\log n}\right)$.*

Proof. For every $V \subseteq [\kappa+1, n]$ of size $|V| = \log\log s$, Proposition 14 implies that there exists a random variable Z such that $\tilde{H}^{(s,\epsilon')}(E_{V|Z}) = 2^{|V|} = \log s$ for $\epsilon' = 3\epsilon + 2^{-s+2}$. But by Claim 12, $\tilde{H}^{(s,\epsilon')}(E_{V|Z}) = \tilde{H}^{(\infty,\epsilon')}(E_{V|Z})$.

The lower bound on $\tilde{G}^{(\infty,\epsilon')}(E)$ follows by partitioning $[\kappa+n]$ into $V_0 \cup V_1 \cup \cdots \cup V_{n/\log\log s}$, where $V_0 = [\kappa]$ and each V_i has size $|V_i| = \log\log s$ for $1 \leq i \leq n/\log\log s$. ◀

▶ Remark 16. We obtain a similar result to Corollary 15 in Appendix A that applies to uniformly secure PRFs.

▶ **Corollary 17.** *If E is the evaluation algorithm for an exponentially secure PRF family with input length n, then any statistically secure DRE for E has size at least $\Omega\left(\frac{n^2}{\log n}\right)$.*

3.6 A Truly Quadratic Lower Bound

We observe that for exponentially secure PRFs with n-bit output, even computationally secure DREs require size $\Omega(n^2)$.

▶ **Theorem 18.** *Any computational DRE of an exponentially-secure PRF with n-bits of output must have size $\Omega(n^2)$.*

To prove this theorem we will rely on the following result of Applebaum et al. [9].

▶ **Theorem 19.** *Let $S(k,x,r)$ be a one-time MAC with key k, message x, and randomness r. Let $\ell(n)$ denote the input locality of $S_k(x,r)$ and let $s(n)$ denote the length of a tag, where n is the security parameter. (A function has input locality ℓ if no input bit affects more than ℓ output bits.) Then, there is an efficient attack on $S(k,x,r)$ that succeeds with probability $1/\binom{s(n)}{\ell(n)} \cdot 2^{-\ell(n)}$.*

Proof. Recall that an exponentially secure PRF $f_k : \{0,1\}^n \to \{0,1\}^n$ is also an exponentially secure one-time MAC [42]. Moreover, a DRE of a MAC preserves unforgeability [9]. Because $1/\binom{n}{\ell(n)} \cdot 2^{-\ell(n)} \leq 2^{\ell(n)}$, it follows Theorem 19 that any DRE of an exponentially-secure f_k must have input locality $\Omega(n)$. By decomposability, any such DRE must have size $\Omega(n^2)$. ◀

4 Upper Bounds on DRE Size

In this section we present nearly matching upper bounds for some of the explicit functions to which our lower bounds apply. We explicitly conjecture two variants of the "hidden shift problem" are exponentially secure PRFs and show that they admit nearly quadratic size (efficient, perfect) DREs. Finally, we show a recent *weak* PRF candidate due to Boneh et al. [19], conjectured to be nearly exponentially-secure, admits a linear-size (efficient and perfect) DRE

4.1 Almost Tight Quadratic Upper Bounds

Partial Decomposability.

We introduce the notation of a *partially decomposable randomized encoding*, so that later we can construction DRE by composing a randomized encoding and a partially decomposable randomized encoding. A randomized encoding $(\mathsf{Enc}, \mathsf{Dec})$ for a function $f : \{0,1\}^n \times \mathcal{W} \to \mathcal{Y}$ is a *partially decomposable randomized encoding (PDRE)* if every bit of $\mathsf{Enc}(x,w,r)$ is determined by $w \in \mathcal{W}, r \in \mathcal{R}$ and only 1 bit of $x \in \{0,1\}^n$.

▶ **Lemma 20** (Composition of randomized encodings). *Let $\mathsf{Enc} : \{0,1\}^n \times \mathcal{W} \to \{0,1\}^\ell$ be (the encoding function of) a randomized encoding $(\mathsf{Enc}, \mathsf{Dec})$ for function $f : \{0,1\}^n \to \mathcal{Y}$. Let $\mathsf{Enc}' : (\{0,1\}^n \times \mathcal{W}) \times \mathcal{R} \to \{0,1\}^{\ell'}$ be the encoding function of a PDRE $(\mathsf{Enc}', \mathsf{Dec}')$ for function Enc. Then $\mathsf{Enc}' : \{0,1\}^n \times (\mathcal{W} \times \mathcal{R}) \to \{0,1\}^{\ell'}$ is the encoding function of a DRE for function f.*

Proof. The corresponding decoding function is $\mathsf{Dec}''(c) := \mathsf{Dec}(\mathsf{Dec}'(c))$. It's easy to verify $(\mathsf{Enc}', \mathsf{Dec}'')$ is a DRE, as each bit of $\mathsf{Enc}'(x,r,w)$ is determined by (r,w) and only 1 bit of x.

A DRE for Element Distinctness.

Choose an $O(\log n)$-bit prime p with $p > \binom{n}{2}$. For all $1 \leq i < i' \leq n$, define indicator $\delta_{i,i'} \in \{0,1\}$ that captures whether $x_i = x_{i'}$,

$$\delta_{i,i'} := \begin{cases} 1, & \text{if } x_i = x_{i'}, \\ 0, & \text{if } x_i \neq x_{i'}. \end{cases}$$

Sample $a \leftarrow \mathbb{Z}_p \setminus \{0\}$ for the CRS. For all $1 \leq i < i' \leq n$, sample random $r_{i,i'} \in \mathbb{Z}_p$ from CRS such that $\sum_{1 \leq i < i' \leq n} r_{i,i'} = 0$. Define $\hat{r}_{i,i'} \in \mathbb{Z}_p$ as $\hat{r}_{i,i'} := r_{i,i'} + a \cdot \delta_{i,i'}$.

Then a DRE for element distinctness is induced by composing the following two claims:

Claim 1. $(\hat{r}_{i,i'})_{1\le i<i'\le n}$ is a randomized encoding of the functionality output.

Proof. It's obvious that $(\hat{r}_{i,i'})_{1\le i<i'\le n}$ is a randomized encoding of $a \cdot \sum_{1\le i<i'\le n} \delta_{i,i'}$. The later is a randomized encoding of the functionality output because: when $(x_i)_{1\le i\le n}$ are all distinct, $a \cdot \sum_{1\le i<i'\le n} \delta_{i,i'}$ is zero; when there is a collision, $a \cdot \sum_{1\le i<i'\le n} \delta_{i,i'}$ is uniformly random in $\mathbb{Z}_p \setminus \{0\}$.

Claim 2. For all $1 \le i < i' \le n$, there exists a PDRE for $\hat{r}_{i,i'}$ of size $O(\log^4 n)$.

Proof. For any $v \in \mathbb{Z}_p$, let $v[k]$ denote the k-th bit of its binary representation. Then the k-th bit of $r_{i,i'}$ can be computed from

$$\begin{aligned}\hat{r}_{i,i'}[k] &= \begin{cases} r_{i,i'}[k], & \text{if } \delta_{i,i'} = 0 \\ (r_{i,i'} + a)[k], & \text{if } \delta_{i,i'} = 1 \end{cases} \\ &= r_{i,i'}[k] \oplus (r_{i,i'}[k] \oplus (r_{i,i'} + a)[k]) \cdot \bigvee_{j=1}^{\log p} (x_i[j] \oplus x_{i'}[j]),\end{aligned}$$

which, as a function of $(x_i, x_{i'})$, is a binary branching program of size $O(\log n)$. Thus there is a PDRE for $\hat{r}_{i,i'}$ of size $O(\log^3 n)$ [8][6]. As $\hat{r}_{i,i'}$ has $\log n$ bits, there exists a PDRE for $\hat{r}_{i,i'}$ of size $O(\log^4 n)$.

4.2 A PRF Candidate With A Nearly Optimal DRE

Now we can present the almost-optimally-garble-able candidate PRF. Modulo a conjecture on its hardness, this simple algebraic PRF candidate admits a (perfect) DRE of size at most a $\log n$ factor from the minimum. Moreover, a simple generalization of this candidate yields linear output length with the same DRE complexity. Thus, if this candidate is exponentially secure, it is indeed optimally-garble-able.

In addition to applications in efficient MPC, this candidate can be conversely interpreted through Razborov and Rudich's natural proof framework as barrier to proving super quadratic bounds on DRE size [53].

An Exponentially-Secure PRF Candidate.

Our starting point is an algebraic object that has received considerable attention in both cryptography and mathematics: Legendre sequences. A Legendre sequence is a sequence of the form:

$$(x+1)^{(p-1)/2}, (x+2)^{(p-1)/2}, (x+3)^{(p-1)/2}, \ldots$$

where all operations are over $\mathbb{Z}_p$ for some prime p.

The pseudorandomness of sequences of quadratic characters have a long history in both cryptography and mathematics [4, 20, 24, 26, 38, 46, 47, 52, 55]. These sequences have been shown to behave as if random with respect to a variety of statistical tests designed for randomness.

Recent work has considered the so-called "hidden shift problems" and their generalizations. In the quadratic character variant of the hidden shift problem, algorithms are given oracle

[6] For a branching program of size s and has t input bits, there is a DRE for the branching program of size s^2t.

access to a function $\phi_k : \mathbb{Z}_p \to \{-1, 0, +1\}$ where $\phi_k(x) = (k + x)^{(p-1)/2}$ from some $k \in \mathbb{Z}_p$. The task is then to recover k. Efficient quantum algorithms for this problem are known [61, 62, 63, 56, 41]. However, the best classical algorithms to date are still just subexponentional (under an assumption on the density of smooth integers) [20, 56, 43]. Indeed, Dam, Hallgren, and Ip [63] have explicitly conjectured that ϕ_k is a PRF with respect to polytime classical algorithms. Grassi et al. [37] additionally proposed this function specifically as an "MPC-friendly" PRF. Recently, cryptanalytic bounties have been announced on this PRF [31].

With the known attacks in mind, we give a twist on the hidden shift problem restricting evaluation to a short interval. So far as we know this confounds all existing techniques (including quantum algorithms) and the best algorithm[7] runs in $2^{(1+o(1))n}$-time [60].

We actually make two conjectures: (1) restricted hidden shift yields an exponentially-secure PRF with one bit of output, (2) a natural generalization is an exponentially-secure PRF with many bits of output. But first, we define the restricted hidden shift function.

For any $m \in \mathbb{Z}^+$, let $p \equiv 1 \pmod{m}$ be a prime with $p \geq 2^{2n}$, and let $\langle \zeta_m \rangle$ denote the group of m^{th} roots of unity in $\mathbb{Z}_p^\times$. For $k \in \mathbb{Z}_p$ define

$$\mathsf{Char}_k^{p,m,n} : [0, 2^n - 1] \to \langle \zeta_m \rangle$$
$$\mathsf{Char}_k^{p,m,n} : x \mapsto (k + x)^{\frac{p-1}{m}} \pmod{p}.$$

Note that $\mathsf{Char}_k^{p,2,n}(x) = 0$ for $k + x = p$. In order to achieve single bit output (just two possible output values) we restrict the key space in addition to the input space, so that this equation cannot be satisfied.

▷ Conjecture 21. Let $p = p(n)$ be any prime sequence satisfying $p \equiv 1 \pmod{m}$, $p > 2^{n+1}$. Then, $\left\{\{\mathsf{Char}_k^{p,2,n}\}_{k \in \{1,\ldots,2^n\}}\right\}_{n \in \mathbb{Z}^+}$ is, for some $s(n) = 2^{\Omega(n)}$, an $(s(\cdot), s(\cdot)^{-1})$-secure PRF family.

Next, we present a variant with long output by applying an input restriction to the "hidden power problem" [21] or "hidden root problem" [64]. In this problem, the goal is to recover k using query access to $x \mapsto (k + x)^e$ for more general $e | p - 1$ (the shift problem discussed above is simply the specific case of $e = \frac{p-1}{2}$). Notably, [21] demonstrated (classical) algorithms for this problem that make $O(1)$ queries and recover k in time $e^{1+\epsilon} \log^{O(1)} p$. With this in mind, we make the following conjecture.

▷ Conjecture 22. Let $p = p(n)$ be any prime sequence and $m = m(n)$ be any positive integer sequence satisfying $p \equiv 1 \pmod{m}$, $p \geq 2^{2n}$, and $\frac{p}{m} \geq 2^n$. Then $\left\{\{\mathsf{Char}_k^{p,m,n}\}_{k \in \mathbb{Z}_p}\right\}_{n \in \mathbb{Z}^+}$ is, for some $s(n) = 2^{\Omega(n)}$, an $(s(\cdot), s(\cdot)^{-1})$-secure PRF family.

An $O(n^2)$ DRE for the Candidate PRF

We now show that there is a DRE for $\mathsf{Char}_{(\cdot)}^{n,m,p}(\cdot)$ of size $O(n^2)$. Assuming the above conjectures, it follows from Corollary 30 that this DRE has essentially optimal size, not just for $\mathsf{Char}_{(\cdot)}^{n,m,p}(\cdot)$, but among DREs for *any* exponentially-secure PRF.

[7] The algorithm is to simply guess k and test on enough x. However it is worth noting that even this is not known to work, and requires making a conjecture on the distribution of Legendre sequences generated by random k [60]. The best *provable* distinguisher that we know of runs in time $2^{(3/2+o(1))n}$-time by simply exhaustively enumerate all sequences of length $2^{n/2}$ and comparing [60]

For clarity, we present a DRE for $\mathsf{Char}_k^{p,2,n}$ and note that the construction can easily be extended to the multi-bit output case.

Our starting point is a simple perfect randomized encoding for quadratic residue[8]:

$$\mathsf{Enc} : x \mapsto x \cdot r^2, \text{ for uniformly sampled } r \leftarrow \mathbb{Z}_p$$
$$\mathsf{Dec} : y \mapsto y^{(p-1)/2}$$

Security follows from the fact that any quadratic residue is mapped to a uniformly random quadratic residue, and any non-residue is mapped to a uniformly random non-residue. Note that this randomized encoding has size $O(n)$.

However, we would like a randomized encoding of the quadratic residuosity of $x + k$ and moreover we would like it to be decomposable. This is easily remedied via bit decomposition and the fact that the above encoding is linear with respect to the input.

$$\begin{aligned} \mathsf{Enc} : x_i &\mapsto x_i \cdot 2^{i-1} \cdot r^2 + s_i \\ k_i &\mapsto k_i \cdot 2^{i-1} \cdot r^2 + t_i \\ &\quad \text{where } r, s_1, \ldots, s_n, t_1, \ldots, t_{2n+1} \text{ are drawn uniformly from } \mathbb{Z}_p \\ &\quad \text{such that } s_1 + \cdots + s_n + t_1 + \cdots + t_{2n+1} = 0. \end{aligned}$$
$$\mathsf{Dec} : y_1, \ldots, y_{3n+1} \mapsto \left(\sum y_i\right)^{(p-1)/2}$$

Similarly, correctness and security follow from the fact that an encoding is simply $3n + 1$ random elements, conditioned on the fact that their sum is a random element with the quadratic residuosity of x. Note that because the encoding consists of $3n + 1$ elements, each of bit length $2n + 1$, the size of this DRE is $O(n^2)$.

4.3 A WPRF Candidate With A Nearly Optimal DRE

In this section, we observe that a recent weak pseudorandom function candidate put forward by Boneh et al. admits a DRE of quasi-linear size [19].

Boneh et al. [19] have put forward the following WPRF candidate related to both the learning parity with noise problem (with "deterministic" noise) and learning with rounding problem (over constant-size modulus). Given a key $k \in \{0, 1\}^n$, they define

$$\mathsf{LWR}_k^6 : \{0,1\}^n \to \{0,1\}$$
$$\mathsf{LWR}_k^6(x) = \begin{cases} 0 & \text{if } \langle x, k \rangle \equiv 0,\ 1, \text{ or } 2 \pmod 6 \\ 1 & \text{if } \langle x, k \rangle \equiv 3,\ 4, \text{ or } 5 \pmod 6. \end{cases}$$

This candidate was proposed with efficient secure function evaluation protocols in mind; however, the protocol presented in [19] requires two phases of interaction: first it applies a DRE-based subprotocol for computing shares of the mod-6 inner product, and then another subprotocol for rounding. Here we show that LWR^6 has a DRE of size $O(n)$.[9]

---

[8] A similar randomization technique for quadratic characters was previously used in related contexts in [30, 5, 1, 37].

[9] In contrast to the PRF candidate proposed above, this WPRF candidate is at most $2^{n/\log n}$-secure. Assuming it is indeed $2^{n/\log n}$ secure, an $O(\lambda \log \lambda)$ size DRE is needed to get 2^λ security.

Protocol

Let S_6 and f act on $\mathbb{Z}_6$ on the right in the natural way. Let $\sigma \in S_6$ denote the permutation that maps x to $x+1$. Let $L = 0 \in \mathbb{Z}_6$.

Randomness

- Sample $r_1, \dots, r_{n-1} \leftarrow S_6$.
- Define $(R_{1,0}, R_{1,1}) = (L \cdot r_1, L \cdot \sigma \cdot r_1)$
- For $2 \leq i \leq n-1$, define $(R_{i,0}, R_{i,1}) = (r_{i-1}^{-1} \cdot r_i, r_{i-1}^{-1} \cdot \sigma \cdot r_i)$
- Define $(R_{n,0}, R_{n,1}) = (r_{n-1}^{-1} \cdot f, r_{n-1}^{-1} \cdot \sigma \cdot f)$

Encoding For $1 \leq i \leq n$, $\mathsf{Enc}_i(z_i, R_i) = M_i = R_{i,z_i}$

Decoding $M_1 \cdots M_n$

Figure 1 A DRE for a function of a sum mod 6 [17]

Let $\lfloor \cdot \rceil : \mathbb{Z}_6 \to \{0,1\}$ denote the function

$$\lfloor x \rceil = \begin{cases} 0 & \text{if } x \in \{0,1,2\} \\ 1 & \text{otherwise.} \end{cases}$$

We obtain our DRE for LWR_k^6 by composing two DREs ([8, 11]); the first is for a function that maps $(z_1, \dots, z_n) \mapsto \lfloor \sum_i z_i \pmod 6 \rceil$ for $z_1, \dots, z_n \in \{0,1\}$, and the second is for the AND function mapping $(k_i, x_i) \in \{0,1\}^2$ to $k_i \cdot x_i$.

The DRE for the first function is obtained as a special case of a result on symmetric functions due to Beimel et al. [17, Theorem 7.2, Figure 9] that refines a group-based DRE due to Kilian [44]:

▷ Imported Theorem 23 ([17]). For any function $f : \mathbb{Z}_6 \to \{0,1\}$, the scheme of Figure 1 is a size-$O(n)$ DRE of the function h that maps $(z_1, \dots, z_n) \mapsto f\big(\sum z_i \pmod 6\big)$.

The second function is constant-sized, and thus has a constant-sized DRE by Barrington's theorem [14] and Kilian's rerandomization.

Acknowledgements.

We thank Igor Shparlinski for helpful discussions and pointers to the literature on the hidden shift problem. We also thank Siyao Guo, Lucas Kowlaczyk, Ron Rothblum, Jonathan Ullman, and Vinod Vaikuntanathan for related discussions and collaborations.

References

1 Shweta Agrawal, Yuval Ishai, Dakshita Khurana, and Anat Paskin-Cherniavsky. Statistical randomized encodings: A complexity theoretic view. In Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors, *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, volume 9134 of *Lecture Notes in Computer Science*, pages 1–13. Springer, 2015. doi:10.1007/978-3-662-47672-7_1.

2 Martin R. Albrecht, Lorenzo Grassi, Léo Perrin, Sebastian Ramacher, Christian Rechberger, Dragos Rotaru, Arnab Roy, and Markus Schofnegger. Feistel structures for mpc, and more. *IACR Cryptology ePrint Archive*, 2019:397, 2019. URL: https://eprint.iacr.org/2019/397.

3 Martin R. Albrecht, Christian Rechberger, Thomas Schneider, Tyge Tiessen, and Michael Zohner. Ciphers for MPC and FHE. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I*, volume 9056 of *Lecture Notes in Computer Science*, pages 430–454. Springer, 2015. doi:10.1007/978-3-662-46800-5_17.

4 Michael Anshel and Dorian Goldfeld. Zeta functions, one-way functions, and pseudorandom number generators. *Duke Math. J.*, 88(2):371–390, 06 1997. doi:10.1215/S0012-7094-97-08815-3.

5 Benny Applebaum. Key-dependent message security: Generic amplification and completeness. *J. Cryptology*, 27(3):429–451, 2014. doi:10.1007/s00145-013-9149-6.

6 Benny Applebaum. Garbled circuits as randomized encodings of functions: a primer. In Yehuda Lindell, editor, *Tutorials on the Foundations of Cryptography*, pages 1–44. Springer International Publishing, 2017. doi:10.1007/978-3-319-57048-8_1.

7 Benny Applebaum, Thomas Holenstein, Manoj Mishra, and Ofer Shayevitz. The communication complexity of private simultaneous messages, revisited. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *Lecture Notes in Computer Science*, pages 261–286. Springer, 2018. doi:10.1007/978-3-319-78375-8_9.

8 Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. Cryptography in nc^0. *SIAM J. Comput.*, 36(4):845–888, 2006. doi:10.1137/S0097539705446950.

9 Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. Cryptography with constant input locality. *J. Cryptology*, 22(4):429–469, 2009. doi:10.1007/s00145-009-9039-0.

10 Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. From secrecy to soundness: Efficient verification via secure computation. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, editors, *Automata, Languages and Programming, 37th International Colloquium, ICALP 2010, Bordeaux, France, July 6-10, 2010, Proceedings, Part I*, volume 6198 of *Lecture Notes in Computer Science*, pages 152–163. Springer, 2010. doi:10.1007/978-3-642-14165-2_14.

11 Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. How to garble arithmetic circuits. *SIAM J. Comput.*, 43(2):905–929, 2014.

12 Marshall Ball, Brent Carmer, Tal Malkin, Mike Rosulek, and Nichole Schimanski. Garbled neural networks are practical. *IACR Cryptology ePrint Archive*, 2019:338, 2019. URL: https://eprint.iacr.org/2019/338.

13 Boaz Barak, Iftach Haitner, Dennis Hofheinz, and Yuval Ishai. Bounded key-dependent message security. In Henri Gilbert, editor, *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Monaco / French Riviera, May 30 - June 3, 2010. Proceedings*, volume 6110 of *Lecture Notes in Computer Science*, pages 423–444. Springer, 2010. doi:10.1007/978-3-642-13190-5_22.

14 David Arno Barrington. *Width-3 permutation branching programs.* Laboratory for Computer Science, Massachusetts Institute of Technology, 1985.

15 Tugkan Batu, Sanjoy Dasgupta, Ravi Kumar, and Ronitt Rubinfeld. The complexity of approximating the entropy. *SIAM J. Comput.*, 35(1):132–150, 2005.

16 Paul Beame, Nathan Grosshans, Pierre McKenzie, and Luc Segoufin. Nondeterminism and an abstract formulation of nečiporuk's lower bound method. *TOCT*, 9(1):5:1–5:34, 2016.

17 Amos Beimel, Ariel Gabizon, Yuval Ishai, Eyal Kushilevitz, Sigurd Meldgaard, and Anat Paskin-Cherniavsky. Non-interactive secure multiparty computation. In *CRYPTO (2)*, volume 8617 of *Lecture Notes in Computer Science*, pages 387–404. Springer, 2014.

18 Mihir Bellare, Viet Tung Hoang, and Phillip Rogaway. Foundations of garbled circuits. In Ting Yu, George Danezis, and Virgil D. Gligor, editors, *the ACM Conference on Computer and*

Communications Security, CCS'12, Raleigh, NC, USA, October 16-18, 2012, pages 784–796. ACM, 2012. doi:10.1145/2382196.2382279.

19 Dan Boneh, Yuval Ishai, Alain Passelègue, Amit Sahai, and David J. Wu. Exploring crypto dark matter: - new simple PRF candidates and their applications. In *TCC (2)*, volume 11240 of *Lecture Notes in Computer Science*, pages 699–729. Springer, 2018.

20 Dan Boneh and Richard J. Lipton. Algorithms for black-box fields and their application to cryptography. In Neal Koblitz, editor, *Advances in Cryptology — CRYPTO '96*, pages 283–297, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg.

21 Jean Bourgain, Moubariz Z. Garaev, Sergei Konyagin, and Igor E. Shparlinski. On the hidden shifted power problem. *SIAM J. Comput.*, 41(6):1524–1557, 2012.

22 Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonina Kolokolova. Learning algorithms from natural proofs. In *Conference on Computational Complexity*, volume 50 of *LIPIcs*, pages 10:1–10:24. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016.

23 Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonina Kolokolova. Agnostic learning from tolerant natural proofs. In *APPROX-RANDOM*, volume 81 of *LIPIcs*, pages 35:1–35:19. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

24 Ivan Damgård. On the randomness of legendre and jacobi sequences. In *CRYPTO*, volume 403 of *Lecture Notes in Computer Science*, pages 163–172. Springer, 1988.

25 Deepesh Data, Manoj Prabhakaran, and Vinod M. Prabhakaran. On the communication complexity of secure computation. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, volume 8617 of *Lecture Notes in Computer Science*, pages 199–216. Springer, 2014. doi:10.1007/978-3-662-44381-1_12.

26 Cunsheng Ding. Pattern distributions of legendre sequences. *IEEE Trans. Information Theory*, 44(4):1693–1698, 1998.

27 Itai Dinur, Daniel Kales, Angela Promitzer, Sebastian Ramacher, and Christian Rechberger. Linear equivalence of block ciphers with partial non-linear layers: Application to lowmc. In Yuval Ishai and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part I*, volume 11476 of *Lecture Notes in Computer Science*, pages 343–372. Springer, 2019. doi:10.1007/978-3-030-17653-2_12.

28 Christoph Dobraunig, Maria Eichlseder, Lorenzo Grassi, Virginie Lallemand, Gregor Leander, Eik List, Florian Mendel, and Christian Rechberger. Rasta: A cipher with low anddepth and few ands per bit. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part I*, volume 10991 of *Lecture Notes in Computer Science*, pages 662–692. Springer, 2018. doi:10.1007/978-3-319-96884-1_22.

29 Nico Döttling and Sanjam Garg. Identity-based encryption from the diffie-hellman assumption. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part I*, volume 10401 of *Lecture Notes in Computer Science*, pages 537–569. Springer, 2017. doi:10.1007/978-3-319-63688-7_18.

30 Uriel Feige, Joe Kilian, and Moni Naor. A minimal model for secure computation (extended abstract). In *STOC*, pages 554–563. ACM, 1994.

31 Dankrad Feist. Legendre prf bounties. URL: https://legendreprf.org/bounties.

32 Michael J. Freedman, Yuval Ishai, Benny Pinkas, and Omer Reingold. Keyword search and oblivious pseudorandom functions. In *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, pages 303–324, 2005.

33 Rosario Gennaro, Craig Gentry, and Bryan Parno. Non-interactive verifiable computing: Outsourcing computation to untrusted workers. In Tal Rabin, editor, *Advances in Cryptology -*

CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings, volume 6223 of *Lecture Notes in Computer Science*, pages 465–482. Springer, 2010. doi:10.1007/978-3-642-14623-7_25.

34 Oded Goldreich. *The Foundations of Cryptography - Volume 1: Basic Techniques*. Cambridge University Press, 2001. URL: http://www.wisdom.weizmann.ac.il/%7Eoded/foc-vol1.html, doi:10.1017/CBO9780511546891.

35 Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1986. doi:10.1145/6490.6503.

36 Shafi Goldwasser, Yael Tauman Kalai, and Guy N. Rothblum. One-time programs. In David A. Wagner, editor, *Advances in Cryptology - CRYPTO 2008, 28th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2008. Proceedings*, volume 5157 of *Lecture Notes in Computer Science*, pages 39–56. Springer, 2008. doi:10.1007/978-3-540-85174-5_3.

37 Lorenzo Grassi, Christian Rechberger, Dragos Rotaru, Peter Scholl, and Nigel P. Smart. MPC-friendly symmetric key primitives. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, pages 430–443, 2016.

38 Jeffrey Hoffstein and Daniel Lieman. The distribution of the quadratic symbol in function fields and a faster mathematical stream cipher. In Kwok-Yan Lam, Igor Shparlinski, Huaxiong Wang, and Chaoping Xing, editors, *Cryptography and Computational Number Theory*, pages 59–68, Basel, 2001. Birkhäuser Basel.

39 Yuval Ishai. Randomization techniques for secure computation. In *Secure Multi-Party Computation*, volume 10 of *Cryptology and Information Security Series*, pages 222–248. IOS Press, 2013.

40 Yuval Ishai and Eyal Kushilevitz. Randomizing polynomials: A new representation with applications to round-efficient secure computation. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 294–304. IEEE Computer Society, 2000. doi:10.1109/SFCS.2000.892118.

41 Gábor Ivanyos, Marek Karpinski, Miklos Santha, Nitin Saxena, and Igor E. Shparlinski. Polynomial interpolation and identity testing from high powers over finite fields. *Algorithmica*, 80(2):560–575, 2018.

42 Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography, Second Edition*. CRC Press, 2014.

43 Dmitry Khovratovich. Key recovery attacks on the legendre prfs within the birthday bound. *IACR Cryptology ePrint Archive*, 2019:862, 2019.

44 Joe Kilian. Founding cryptography on oblivious transfer. In Janos Simon, editor, *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 20–31. ACM, 1988. doi:10.1145/62212.62215.

45 O B Lupanov. A method for synthesizing circuits. *Radiofizika*, 1958.

46 Christian Mauduit. Finite and infinite pseudorandom binary words. *Theor. Comput. Sci.*, 273(1-2):249–261, 2002.

47 Christian Mauduit and András Sárközy. On finite pseudorandom binary sequences, vi,(on sequences). *Monatshefte für Mathematik*, 130(4):281–298, Sep 2000. doi:10.1007/s006050070028.

48 Moni Naor and Omer Reingold. Number-theoretic constructions of efficient pseudo-random functions. In *38th Annual Symposium on Foundations of Computer Science, FOCS '97, Miami Beach, Florida, USA, October 19-22, 1997*, pages 458–467, 1997.

49 Eduard Ivanovich Nečiporuk. On a boolean function. In *Dokl. Akad. Nauk SSSR*, volume 169, pages 765–766, 1966.

50 Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, New York, NY, USA, 2014.

51 Igor Carboni Oliveira and Rahul Santhanam. Conspiracies between learning algorithms, circuit lower bounds, and pseudorandomness. In *Computational Complexity Conference*, volume 79 of *LIPIcs*, pages 18:1–18:49. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

52 Rene Peralta. On the distribution of quadratic residues and nonresidues modulo a prime number. *Mathematics of Computation*, 58(197):433–440, 1992.

53 Alexander A. Razborov and Steven Rudich. Natural proofs. *J. Comput. Syst. Sci.*, 55(1):24–35, 1997.

54 Christian Rechberger, Hadi Soleimany, and Tyge Tiessen. Cryptanalysis of low-data instances of full lowmcv2. *IACR Trans. Symmetric Cryptol.*, 2018(3):163–181, 2018. doi:10.13154/tosc.v2018.i3.163-181.

55 Joël Rivat and András Sárközy. On pseudorandom sequences and their application. In Rudolf Ahlswede, Lars Bäumer, Ning Cai, Harout K. Aydinian, Vladimir M. Blinovsky, Christian Deppe, and Haik Mashurian, editors, *General Theory of Information Transfer and Combinatorics*, volume 4123 of *Lecture Notes in Computer Science*, pages 343–361. Springer, 2006. doi:10.1007/11889342_19.

56 Alexander Russell and Igor E. Shparlinski. Classical and quantum function reconstruction via character evaluation. *J. Complexity*, 20(2-3):404–422, 2004. doi:10.1016/j.jco.2003.08.019.

57 John E. Savage. *Models of computation - exploring the power of computing.* Addison-Wesley, 1998.

58 Rocco A. Servedio and Li-Yang Tan. What circuit classes can be learned with non-trivial savings? In *ITCS*, volume 67 of *LIPIcs*, pages 30:1–30:21. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

59 C. E. Shannon. The synthesis of two-terminal switching circuits. *The Bell System Technical Journal*, 28(1):59–98, Jan 1949. doi:10.1002/j.1538-7305.1949.tb03624.x.

60 Igor E. Shparlinski. Private Communication, 2019.

61 Wim van Dam. Quantum algorithms for weighing matrices and quadratic residues. *Algorithmica*, 34(4):413–428, 2002.

62 Wim van Dam and Sean Hallgren. Efficient quantum algorithms for shifted quadratic character problems. *CoRR*, quant-ph/0011067, 2000.

63 Wim van Dam, Sean Hallgren, and Lawrence Ip. Quantum algorithms for some hidden shift problems. *SIAM J. Comput.*, 36(3):763–778, 2006.

64 Frederik Vercauteren. The hidden root problem. In *Pairing*, volume 5209 of *Lecture Notes in Computer Science*, pages 89–99. Springer, 2008.

65 Andrew Chi-Chih Yao. How to generate and exchange secrets (extended abstract). In *27th Annual Symposium on Foundations of Computer Science, Toronto, Canada, 27-29 October 1986*, pages 162–167. IEEE Computer Society, 1986. doi:10.1109/SFCS.1986.25.

A PRF Bounds in the Uniform Setting

In this appendix, we give improved lower bounds on the complexity of garbling pseudorandom functions (PRFs). In particular, the attack presented here is uniform, as opposed to non-uniform bounds in Corollary 15. Our results follow from applying the natural proof framework of [53]. However, we achieve improved bounds by demonstrating the existence of a property tester for a relaxation of Nečiporuk's measure. By combining our results with those of Section 3.4 we show any exponentially-secure PRF has DRE size $\Omega(n^2/\log n)$.

We then discuss a candidate PRF with a DRE construction of size almost matching the lower bound.

A.1 PRFs are complex under (average-case) Nečiporuk

Intuitively, because a random function has high measure under Nečiporuk, so should a pseudorandom function.[10] In fact, Servedio and Tan have recently shown how to exactly learn functions with low ($O(n^{1.99})$) measure under Nečiporuk in time $2^{n-n^{\delta}}$ (via membership and equivalence queries) [58]. We show that the much simpler task of simply distinguishing a function with low measure can be done much more quickly (and without equivalence queries, which do not fit into the usual PRF game).

We accomplish this via an average case variant of Nečiporuk. Recall that Nečiporuk is ultimately statement about the number of functions that can be generated under some restriction. Viewed differently, this can be framed as a statement about the *maximum entropy* of the random variable defined by sampling a restricted function uniformly at random. Our observation is that for the special case of distinguishing from a random function it suffices to look at the *Shannon entropy* of the same variable. Consequently, instead of bounding the support size we can focus on much easier task of bounding the entropy.

An "average-case" notion of Nečiporuk.

We begin by introducing our average-case variant of Nečiporuk's measure that relies on Shannon entropy as opposed to maximum entropy.

For a function $f : \{0,1\}^n \to \{0,1\}$, and a set $S \subset [n]$, let $Z^{f,S}$ denote the variable distributed according to $f_{S|z}$ for uniformly drawn $z \leftarrow \{0,1\}^{\bar{S}}$. Define,

$$h_S(f) \stackrel{\mathsf{def}}{=} H(Z^{f,S}).$$

Notice that $H_{\max}(Z^{f,S}) = g_S(f)$, thus it follows that $h_S(f) \leq g_S(f)$.

Random functions are complex (under h_S)

Next we observe that random functions have high complexity with respect to the average-case variant of Nečiporuk we defined above.

▶ **Proposition 24.** *For any set $S \subseteq [n]$ and a uniformly random function $F : \{0,1\}^n \to \{0,1\}$,*

$$\Pr[h_S(F) \leq 2^{|S|} - t] < \exp(-\frac{2t^2}{|\bar{S}| + \ln(2)})$$

We can apply the same style of balls/bins argument used for Nečiporuk's original measure again here.

Proof. First, we bound $\mathbb{E}[H(Z^{F,S})]$ from below. We will omit S from the superscript in this proof ($Z^F = Z^{F,S}$). Additionally, we will take Z^F_ϕ to denote $\Pr_F[Z^F = \phi]$. Note that for

[10] Statements of this form indeed were at the heart of Razoborov and Rudich's natural proof framework and its recent extensions [53, 22, 23, 51]. Unfortunately, because Nečiporuk's measure seems to behave poorly under known pseudorandom function generators, its not clear how to apply their framework here to get strong bounds on pseudorandom generators with simple DREs.

any ϕ, $\mathbb{E}[Z^F_\phi] = 1/\#\{\phi : \{0,1\}^{|S|} \to \{0,1\}\} = 2^{-2^{|S|}}$.[11]

$$\begin{aligned}
\mathbb{E}_F[H(Z)] &= \mathbb{E}_F\left[\sum_\phi Z^F_\phi \log(1/Z^F_\phi)\right] \\
&= \sum_\phi \mathbb{E}_F[Z^F_\phi \log(1/Z^F_\phi)] \\
&\geq \sum_\phi \mathbb{E}_F[Z^F_\phi] \log(1/\mathbb{E}_F[Z^F_\phi]) \\
&= 2^{2^{|S|}} \cdot \frac{1}{2^{2^{|S|}}} \log(2^{2^{|S|}}) \\
&= 2^{|S|}
\end{aligned}$$

Note that the third line follows from Jensen's inequality.

Next, we show concentration around the mean in the standard way. Consider $H(Z^F)$ as a Doob martingale on the independent random variables $F_{S|z}$ for $z \in \{0,1\}^{\bar{S}}$. Clearly, if F and F' only differ on single restriction of f to z, then $|H(Z^F) - H(Z^{F'})| \leq \frac{\log(2^{|\bar{S}|}) + \ln(2)}{2^{|\bar{S}|}}$. Moreover, because F is random, these variables are independent. So, we can apply McDiarmid/Azuma's inequality to get, for any $t > 0$:

$$\Pr_F[\mathbb{E}[H(Z^F)] - H(Z^F) \geq t] \leq \exp(\tfrac{2t^2}{|\bar{S}| + \ln(2)}).$$

◀

Plugging $|S| = \log n$ and $t = n/2$ into the above proposition we immediately get the following corollary.

▶ **Corollary 25.** *For any set $S \subseteq [n]$ such that $|S| = \log n$, if $F : \{0,1\}^n \to \{0,1\}$ is a uniformly random function, then*

$$\Pr[h_S(F) \leq n/2] < \exp(-\frac{n^2}{2(n - \log n + \ln(2))}) < \exp(-n/2).$$

A.2 Low Nečiporuk measure can be distinguished from random

Next, we use the above to show that any function with Nečiporuk measure that is slightly less than maximal can be distinguished from a random function in time $O(2^{n/10})$. It immediately follows that none of the classes whose functions have bounded Nečiporuk measure can contain exponentially-secure PRFs.

The following theorem is implicit in Batu et al. [15].

▷ Imported Theorem 26. There is an algorithm that given sample access to a distribution X supported on $[N]$, promised to either have "high" entropy (at least $N/2$) or "low" entropy (at most $N/21$), runs in time $\tilde{O}(N^{1/100})$ and distinguishes which is the case with overwhelming probability.

---

[11] In more detail: Let $M = \#\{0,1\}^{\bar{S}}$ (number of balls) and $N = \#\{0,1\}^{\{0,1\}^S}$ (number of bins). Then, for $k \in \mathbb{N}$ we can see that $\Pr[Z^F_\phi = k/M]$ is the probability that exactly k out M balls (or restrictions $z \in \{0,1\}^{\bar{S}}$) hit the bin ϕ (which happens with probability $1/N$). Thus, $\Pr[Z^F_\phi = k/M] = \binom{M}{k} N^{-k}(1 - \frac{1}{N})^{M-k}$. Because this is simply a rescaled binomial distribution it follows that $\mathbb{E}[Z^F_\phi] = \frac{1}{M} \cdot \frac{M}{N} = \frac{1}{N}$.

▶ Remark 27. Batu et al. actually show how to multiplicatively approximate entropy within a factor of $(1+2\epsilon)\gamma$ ($\gamma > 1, \epsilon \in (0, 1/2]$) given sample access in time $O(N^{1/\gamma^2}/\epsilon^2 \log n)$ with constant failure probability when the distribution has entropy at least $\Omega(\gamma/\eta)$ for some small constant η ([15, Theorem 2]).

To apply this to the low entropy case, it suffices to suffices to show min-entropy is greater than the constant assumed above. For these parameters, empirical estimates are more than efficient enough. In fact, [15, Lemma 2] says just that. Finally, correctness of these estimates can be amplified by taking the median/majority after poly $\log n$ repetitions.

▶ **Theorem 28.** *There is an algorithm running in time $\tilde{O}(2^{n/100})$ that given oracle access to either a random function $F : \{0,1\}^n \to \{0,1\}$ or any $f : \{0,1\}^n \to \{0,1\}$ such that $G(f) < \frac{n^2}{21 \log n}$ can distinguish between the two cases with overwhelming advantage.*

Proof. Note that if $G(f) < \frac{n^2}{21 \log n}$, then in particular $\sum_{V_i} g_{V_i}(f) < \frac{n^2}{21 \log n}$ for the partition $(V_1, \ldots, V_{n/\log n})$ of $[n]$ into consecutive $\log n$-bit blocks. Moreover, there must be some V_i such that $h_{V_i}(f) \leq g_{V_i}(f) < n/21$.

In contrast, Corollary 25 implies that for a uniformly random F, it holds with overwhelming probability that for all i, $h_{V_i}(F) \geq n/2$.

Additionally, for any i, it is possible to efficiently sample Z^{f,V_i} by simply drawing $z \leftarrow \{0,1\}^{|\bar{V}_i|}$ uniformly at random and evaluating $f_{V_i|z}$ on all $x \in \{0,1\}^{V_i}$. Because $|V_i| = \log n$, this procedure takes time poly(n).

It follows that we can run the procedure from Imported Theorem 26 on all V_i in time $\tilde{O}(2^{n/100})$. If the procedure outputs "High" on all V_i, then output "F." Otherwise, output "f." By Theorem 26 and the above observations, the procedure described will err with at most negligible probability. ◀

▶ Remark 29. We note that for $\epsilon > 0$ the above distinguisher can be modified to test on the partition $V = (V_1, \ldots, V_m)$ where each V_i is a block of size $\epsilon \log n$ ($m = \frac{n}{\epsilon n}$) and again distinguish entropy that differs by constant factor in any block from $n^\epsilon/2$, taking time $O(2^{n^\epsilon})$ overall. By Proposition 24 a random function will have Nečiporuk measure $h_{V_i}(f) \geq n^\epsilon/2$ for all V_i with high probability. It follows that an $O(2^{n^\epsilon})$-secure PRF must have DRE complexity $\Omega(n^{1+\epsilon}/\log n)$.

PRFs have high complexity.

From Theorem 28, it almost immediately follows that there can be no exponentially-secure PRFs in any class to which Nečiporuk applies. This yields a host of lower bounds on PRF complexity that, to our knowledge, were not known before now.

▶ **Corollary 30.** *No exponentially-secure PRF has*

- *Decomposable Randomized Encodings of size $o(n^2/\log n)$,*
- *Binary formulas of size $o(n^2/\log n)$ over arbitrary basis,*
- *Deterministic branching programs of size $o(n^2/\log^2 n)$,*
- *Switching networks of size $o(n^2/\log^2 n)$,*
- *Non-deterministic branching programs of size $o(n^{3/2}/\log n)$,*
- *Parity branching programs of size $o(n^{3/2}/\log n)$,*
- *Span programs of size $o(n^{3/2}/\log n)$,*
- *Switching-and-rectifier networks of size $o(n^{3/2}/\log n)$.*

B Deferred Proofs

▶ **Proposition 31.** *For any set* $S \subseteq [n]$ *and a random function* $f : \{0,1\}^n \to \{0,1\}$, $\Pr_f[2^{g_S(f)} \leq 2^{n-|S|} - t] < \exp(-\frac{2t^2}{2^{n-|S|}})$

This follows from a standard balls & bins argument, reproduced here for completeness.

Proof. Recall that $2^{g_V(f)} = \#\{f_{S|z} : z \in \{0,1\}^{\bar{S}}\}$. If we let Y_ϕ for $\phi : \{0,1\}^S \to \{0,1\}$ be the indicator random variable such that

$$Y_\phi := \begin{cases} 1 & \text{if } \exists z \in \{0,1\}^{\bar{S}} : f_{S|z} = \phi \\ 0 & \text{otherwise} \end{cases}$$

Then we can rewrite the above as,

$$2^{g_S(f)} = \sum_{\phi:\{0,1\}^S \to \{0,1\}} Y_\phi.$$

By linearity of expectation,

$$\mathbb{E}[2^{g_S(f)}] = \mathbb{E}[\sum_\phi Y_\phi] = \sum_\phi \mathbb{E}[Y_\phi] = 2^{2^{|S|}} \cdot \frac{2^{|\bar{S}|}}{2^{2^{|S|}}} = 2^{n-|S|}.$$

Finally, we consider $2^{g_S(f)}$ as a doob martingale on the independent random variables $f_{S|z}$ for $z \in \{0,1\}^{\bar{S}}$. Clearly, if f and f' only differ on single restriction of f to z, then $|g_S(f) - g_S(f')| \leq 1$. Moreover, because f is random, these variables are independent. So, we can apply McDiarmid/Azuma's inequality to get, for any $t > 0$:

$$\Pr_f[\mathbb{E}[2^{g_S(f)}] - 2^{g_S(f)} \geq t] \leq \exp(\frac{2t^2}{2^{|\bar{S}|}}).$$

◀

In particular, if we take $|S| = \log n$ and $t = 2^{n-\log n-1}$, then $\Pr_f[g_S(f) \leq n - \log n - 1] \leq \exp(-2^{n-\log n-1})$. This yields the following corollary via a union bound.

▶ **Corollary 32.** *For a random function* f, $\Pr_f[G(f) \leq n^2/\log n - 2n] \leq \frac{n}{\log n} \cdot \exp(-2^{n-1})$