# Explicit resilient functions matching Ajtai-Linial

Raghu Meka

Department of Computer Science

University of California, Los Angeles

### Abstract

A Boolean function on $n$ variables is $q$-resilient if for any subset of at most $q$ variables, the function is very likely to be determined by a uniformly random assignment to the remaining $n - q$ variables; in other words, no coalition of at most $q$ variables has significant influence on the function. Resilient functions have been extensively studied with a variety of applications in cryptography, distributed computing, and pseudorandomness. The best known resilient function on $n$ variables due to Ajtai and Linial [AL93] has the property that only sets of size $\Omega(n/(\log^2 n))$ can have influence bounded away from zero. However, the construction of Ajtai and Linial is by the probabilistic method and does not give an efficiently computable function.

We construct an explicit monotone depth three almost-balanced Boolean function on $n$ bits that is $\Omega(n/(\log^2 n))$-resilient matching the bounds of Ajtai and Linial. The best previous explicit constructions of Meka [Mek09] (which only gives a logarithmic depth function), and Chattopadhyay and Zuckerman [CZ15] were only $(n^{1-\beta})$-resilient for any constant $0 < \beta < 1$. Our construction and analysis are motivated by (and simplifies parts of) the recent breakthrough of [CZ15] giving explicit two-sources extractors for polylogarithmic min-entropy; a key ingredient in their result was the construction of explicit constant-depth resilient functions.

An important ingredient in our construction is a new randomness-optimal oblivious sampler that preserves moment generating functions of sums of variables and could be useful elsewhere.

## 1 Introduction

In this work we study *resilient functions* introduced by Ben-Or and Linial [BL85] in the context of collective-coin flipping. Consider the following game: There are $n$ players who communicate by broadcast and want to agree on a random coin-toss. If all the players are honest, this is trivial: pick a player, have the player toss a coin and use the resulting value as the collective coin-toss. Now suppose that there are a few *bad* players who are computationally unbounded, can collude amongst themselves, and broadcast last in each round, i.e., they broadcast after observing the bits broadcast by the good players in each round. The problem of *collective coin-flipping* is to design protocols so that the bad players cannot bias the collective-coin too much. An important and well-studied case of protocols are *one-round* collective coin-flipping protocols. We will adopt the notation of boolean functions instead of protocols, as both are equivalent for a single round.

**Definition 1.1.** *For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, and $Q \subseteq [n]$, let $I_Q(f)$ be the probability that $f$ is not-determined by a uniformly random partial assignment to the bits not in $Q$. Let $I_q(f) = \min_{Q \subseteq [n], |Q| \leq q} I_Q(f)$. We say $f$ is $(q, \delta)$-resilient if $I_q(f) \leq \delta$. In addition, for $0 < \tau < 1$, we say $f$ is $\tau$-strongly resilient if for all $1 \leq q \leq n$, $I_q(f) \leq \tau \cdot q$.*

Intuitively, $I_q(f)$ quantifies the amount of influence any set of $q$ variables can exert on the evaluation of the function $f$. If $f$ is almost-balanced and $I_q(f)$ is small, say $o(1)$, then evaluating

$f$ gives a one-round coin-flipping protocol that outputs a nearly unbiased bit even in the presence of up to $q$ bad players. More information and discussion of other models can be found in the survey of Dodis [Dod06].

In their work introducing the problem, Ben-Or and Linial constructed an explicit balanced[1] $(1/n^\alpha)$-strongly resilient function for $\alpha = \log_3 2$. Subsequently, the seminal work of Kahn, Kalai, and Linial [KKL88] showed that for any balanced function $f$, $I_q(f) = 1 - o(1)$ for $q = \omega(n/\log n)$. Following this, Ajtai and Linial [AL93] showed the *existence* of a balanced function that is $\tau$-strongly resilient for $\tau = O(\log^2 n/n)$; this in particular implies the existence of a $(\Omega(n/\log^2 n), 1/3)$ resilient function. However, the construction in Ajtai and Linial is probabilistic and does not lead to an efficiently computable resilient function—it only gives a non-uniform polynomial-size circuit for computing such a resilient function. We construct an efficiently computable function matching the existential result of Ajtai and Linial:

**Theorem 1.2** (Main). *For some universal constants $c_1, c_2 \geq 1$ the following holds. There exists an efficiently computable function $f : \{0,1\}^n \to \{0,1\}$ such that[2]*

- *$f$ is almost-balanced:* $\Pr_{x \in_u \{0,1\}^n}[f(x) = 1] = 1/2 \pm 1/10$.

- *$f$ is $(c_1(\log^2 n)/n)$-strongly resilient.*

- *$f$ has a* uniform *depth $3$ monotone circuit of size at most $n^{c_2}$.*

The best previously known explicit resilient functions as above could only tolerate roughly at most $q \ll n^{1-\beta}$ bad players for all constants $0 < \beta < 1$: [Mek09] gave such a function of logarithmic-depth while the recent breakthrough of [CZ15] gave such a function of depth 4.

The existential guarantee of [AL93] is slightly stronger than the above; they show the existence of a balanced constant-depth function with similar resilience; however, their function is not monotone[3]. Our construction essentially matches theirs while being efficiently computable. We can have the bias of the function be $1/2 \pm o(1)$ at the expense of reducing the resilience; see Corollary 7.4 for one such trade-off.

## 1.1 Two-source extractors

The present work builds on a recent breakthrough of Chattopadhyay and Zuckerman [CZ15] who gave an explicit *two-source extractor* for poly-logarithmic *min-entropy* sources—resolving a longstanding problem in pseudorandomness. One of the main building blocks of their work is an efficiently computable resilient function with a stronger guarantee described below. We noticeably simplify the construction and analysis of [CZ15] and obtain better quantitative bounds. We explain these next starting with the definitions of extractors.

For a random variable $X$, the min-entropy of $X$ is defined by

$$H_\infty(X) = \min_{x \in Support(X)} (\log_2(1/\Pr[X = x])).$$

A two-source extractor is a function that takes two high min-entropy sources and outputs a nearly uniform random bit:

**Definition 1.3.** *A function $Ext : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is a $(n, k)$ two-source extractor with error $\varepsilon$ if for any two independent $X, Y$ with $H_\infty(X), H_\infty(Y) \geq k$, $Ext(X, Y)$ is $\varepsilon$-close to a uniformly random bit. If $Ext(X, Y)$ has full support for all such sources $X, Y$, then we say $Ext$ is a* two-source disperser.

---

[1] We say $f : \{0,1\}^n \to \{0,1\}$ is balanced if $\Pr_{x \in_u \{0,1\}^n}[f(x) = 1] = 1/2$.

[2] Henceforth, for a multi-set $S$, $x \in_u S$ denotes a uniformly random element of $S$.

[3] While it is possible to make their construction monotone, this blows up the depth.

Extractors have many applications across several areas including cryptography, error-correcting codes, randomness amplification; we refer the reader to the history of the problem in [CZ15] (the references are too many). By the probabilistic method, there exist two-source extractors with error $2^{-\Omega(k)}$ for all $k \geq 2 \log n$ (even outputting $\Omega(k)$ bits; as in [CZ15] we only focus on extracting one bit in this work). Constructing such functions explicitly, as is required in most applications, is much harder and has been studied extensively.

Until very recently, the best explicit two-source extractor due to Bourgain [Bou05] required min-entropy at least $k \geq cn$ for some constant $c = 0.49... < 0.5$; the best explicit two-source disperser due to [BRSW12] required min-entropy at least $\exp(\text{poly}(\log \log n))$. Chattopadhyay and Zuckerman broke the barrier for two-source extractors and gave an explicit construction for min-entropies at least $C(\log n)^{74}$; independently, Cohen [Coh15] gave an explicit two-source disperser for min-entropy $\log^C n$ for some (unspecified) constant $C$. We show the following:

**Theorem 1.4.** *For a sufficiently big constant $C$, there exists an explicit $(n, k)$ two-source extractor $D : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with constant-error for $k \geq C \log^{10} n$. In particular, we get an explicit $(n, k)$ two-source disperser for $k \geq C \log^{10} n$.*

**Theorem 1.5.** *For any constant $c \geq 1$, there exists a constant $C$ such that there exists an explicit $(n, k)$ two-source extractor $Ext : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ with error $1/n^c$ for $k \geq C \log^{18} n$.*

**Two-source extractors and resilient functions** Along with the above quantitative improvements, our construction and analysis simplify [CZ15]:

- [CZ15] uses several known extractors in their construction of $n^{1-\beta}$-resilient functions for all constant $0 < \beta < 1$; we present a simpler construction based on Reed-Solomon codes.

- More importantly, the analysis of [CZ15] uses Braverman's celebrated result—[Bra10]—that polylog-wise independence fools constant-depth circuits. We present a direct and self-contained analysis without recourse to Braverman's result that in turn uses several non-trivial tools from the study of constant-depth circuits, e.g., [Raz87], [Smo87], [LMN93].

Chattopadhyay and Zuckerman reduce the problem of computing two-source extractors to that of constructing an explicit $n^{1-\delta}$-resilient function for some constant $\delta > 0$ with the following stronger property; this reduction was also implicit in [Li15]. A distribution $\mathcal{D}$ on $\{0,1\}^n$ is $t$-wise independent if for $X \leftarrow \mathcal{D}$, and all $I \subseteq [n]$ with $|I| \leq t$, the projection of $X$ onto the coordinates in $I$, $X_I$, is uniformly distributed over $\{0,1\}^I$.

**Definition 1.6.** *For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, $Q \subseteq [n]$, and a distribution $\mathcal{D}$ on $\{0,1\}^{[n]\setminus Q}$, let $I_{Q,\mathcal{D}}(f)$ be the probability that $f$ is not-determined by setting the bits not in $Q$ according to $\mathcal{D}$. Let $I_{q,t}(f) = \min\{I_{Q,\mathcal{D}}(f) : Q \subseteq [n], |Q| \leq q, \mathcal{D}$ is $t$-wise independent$\}$. We say $f$ is $t$-wise $(q, \varepsilon)$-resilient if $I_{q,t}(f) \leq \varepsilon$. In addition, for $0 < \tau < 1$, we say $f$ is $t$-wise $\tau$-strongly resilient if for all $1 \leq q \leq n$, $I_{q,t}(f) \leq \tau \cdot q$.*

The core of [CZ15] is the construction of a $\text{poly}(\log n)$-wise $(n^{1-\beta}, 1/n^{\Omega(1)})$-resilient function for all $\beta > 0$. They achieve this as follows: a) Construct an explicit constant-depth monotone $(n^{1-\beta}, 1/n^{\Omega(1)})$-resilient function. b) Apply Braverman's result to conclude that for all $q, \varepsilon > 0$, a constant-depth monotone $(q, \varepsilon)$-resilient function is also $t$-wise $(q, 2\varepsilon)$-resilient for $t = \text{poly}(\log(n/\varepsilon))$. In contrast, our analysis of resilience is robust enough to imply resilience even under limited independence with little extra work. As a corollary, we get the following:

**Theorem 1.7.** *For some universal constant $c \geq 1$ the following holds. There exists an efficiently computable function $f : \{0,1\}^n \to \{0,1\}$ such that,*

- *$f$ is almost-balanced: For any $(c \log^2 n)$-wise independent distribution $\mathcal{D}$ on $\{0,1\}^n$, $\Pr_{x \sim \mathcal{D}}[f(x) = 1] = 1/2 \pm 1/9$.*

3

- $f$ *is* $(c \log^2 n)$*-wise* $(c(\log^2 n)/n)$*-strongly resilient.*

- $f$ *has a* uniform *depth* 3 *monotone circuit of size at most* $n^c$.

As in the case of Theorem 1.2, we can have the bias of the function be $1/2 \pm o(1)$ at the expense of reducing the resilience.

## 1.2  Oblivious samplers preserving moment generating functions

A critical ingredient in our proof of Theorem 1.2 is an explicit *oblivious sampler* with optimal—up to constant factors—seed-length that approximates *moment generating functions* (MGF). We state this result next which may be of independent interest.

**Theorem 1.8.** *For all* $0 < \mu \le 1$, $1 \le v, w$, *there exists an explicit generator* $G : \{0,1\}^r \to [v]^w$ *such that for all functions* $f_1, \ldots, f_w : [v] \to [0,1]$ *with* $\sum_{i=1}^{w} \mathbb{E}_{x \in_u [v]}[f_i(x)] = \mu$,

$$\mathbb{E}_{y \in_u \{0,1\}^r} \left[ 2^{\sum_{i=1}^{w} f_i(G(y)_i)} \right] = 1 + O(\mu).$$

*The seed-length of the generator is* $r = O(w + (\log v) + w((\log \log v) + \log(1/\mu))/(\log w))$.

We state a more precise version which works for estimating $\mathbb{E}\left[\exp\left(\theta \cdot \sum_i f_i(\ )\right)\right]$ for all $\theta > 0$, i.e., the MGF, in Section 6; here we focus on the above for simplicity and as it captures the main ideas.

Let us first compare the above with known randomness efficient samplers such as those of [Zuc97, Gil98, Kah97]. For concreteness, let us consider the special case when $v = 2^{O(w)}$ and $\mu = 1$; these are the parameters we face in our application. The seed-length of Theorem 1.8 in this case is $O(w)$. On the other hand, the samplers of [Zuc97, Gil98, Kah97] when instantiated to obtain a guarantee as above require a seed-length of $\Omega(w \log w)$. The improvement from $O(w \log w)$ to $O(w)$ is critical in our application; as we enumerate our all possible seeds eventually, the improvement from $O(w \log w)$ to $O(w)$ in seed-length translates to an improvement from super-polynomial running-time $(n^{O(\log \log n)})$ to polynomial running-time of the resilient function.

To illustrate the gap further, with the above parameters, let $X_i = f_i(x_i)$ for an independent uniformly random $x_i \in_u [v]$ and let $Y_i = f_i(G(y)_i)$, where $G$ is as in the theorem. In this case, standard Chernoff bounds imply that for a sufficiently big constant $C$,

$$\Pr[X_1 + \cdots + X_w > Cw/(\log w)] \le \exp(-w). \tag{1.1}$$

Our argument in Section 6 shows that our generator satisfies the same property: $\Pr[Y_1 + \cdots + Y_w > Cw/(\log w)] \le \exp(-w)$. Note that the seed-length used by our generator for this setting is $O(w)$ which is optimal[4].

In contrast, if one uses the *expander sampler* as in [Gil98, Kah97, Hea08] on an expander graph with degree $D$ the best one could get using the current analyses is (cf. Corollary 23 of [Hea08]),

$$\Pr[Y_1 + \cdots + Y_w > Cw/(\log w)] \le \exp(-w \cdot (1 - (w^2/D^4))).$$

In other words, to get an $\exp(-w)$ bound on the tail-bound, one needs the degree $D$ of the expander to be $w^{\Omega(1)}$; as the number of random bits needed by such a generator is $(\log v) + w(\log D)$, the total seed-length will be $O(w \log w)$. Similarly, applying the analysis of expander Chernoff bounds from [Gil98, Kah97, Hea08] to get a bound on the MGF as in the theorem requires the degree of the expander to be at least $w^{\Omega(1)}$. This in turn requires the total seed-length to be $\Omega(w \log w)$.

---

[4]To get a tail bound of $\exp(-w)$, the sample space has to have $\exp(w)$ points.

Another potential approach for constructing samplers as in theorem is to apply the pseudorandom generators (PRGs) for small-space machines of Nisan or Impagliazzo, Nisan, and Wigderson [Nis92, INW94]. However, to obtain a guarantee on the MGF as in the theorem or to satisfy Equation 1.1 in the special-case above, we need to instantiate the generators with error $\ll \exp(-w)$. This in turn forces the seed-length of the generators to be $\Omega(w \log w)$.

## 2 Overview of construction

We next give a high level overview of our main construction and analysis. First, some notations:

- Throughout, by a partition $P$ of $[n]$ we mean a division of $[n]$ into $w$-sized blocks $P_1, \ldots, P_v$, where $v = n/w$ (we assume $w$ divides $n$).

- Let $\mathcal{D}_p$ denote the product distribution on $\{0,1\}^n$ where each bit is $p$-biased.

As in [AL93] and [CZ15] our construction will be an *AND* of several *Tribe* functions:

**Definition 2.1.** *For a partition* $P = \{P_1, \ldots, P_v\}$ *of* $[n]$ *into* $w$-*sized blocks, the associated* Tribes *function is the DNF defined by* $T_P = \vee_{j=1}^{v}\left(\wedge_{\ell \in P_j} x_\ell\right)$. *A collection of partitions* $\mathcal{P} = \{P^1, \ldots, P^u\}$ *defines a function* $f \equiv f_{\mathcal{P}} : \{0,1\}^n \to \{0,1\}$ *as follows:*

$$f_{\mathcal{P}}(x) := \bigwedge_{i=1}^{u} \bigvee_{j=1}^{v} \left(\bigwedge_{k \in P_j^i} x_k\right) = \bigwedge_{i=1}^{u} T_{P^i}(x).$$

The final function satisfying Theorem 1.2 will be $f_{\mathcal{P}}$ for a suitably chosen set of partitions. To analyze such functions, we first state two abstract properties that allow us to analyze the bias as well as influences of such functions; we then design partitions that satisfy the properties. The properties we define are motivated by [CZ15] and abstracting them in this way allows us to give a modular analysis of the construction. The partitions themselves will be designed using the sampler we construct in Theorem 1.8 which forms the core of our analysis and construction. The analysis of resilience under limited independence follows a similar approach in addition to some careful, but elementary, calculations involving elementary symmetric polynomials.

**Analyzing bias** The first condition allows us to approximate the bias of functions of the form $f_{\mathcal{P}}$.

**Definition 2.2.** *Let* $\mathcal{P} = \{P^1, \ldots, P^u\}$ *be a collection of partitions of* $[n]$ *into* $w$-*sized blocks. For* $d \leq w$, *we say* $\mathcal{P}$ *is a* $d$-*design if no two blocks across any of the partitions overlap in more than* $w - d$ *elements: formally, for all* $\alpha \neq \beta \in [u]$, *and* $i, j \in [v]$, $|P^\alpha(i) \cap P^\beta(j)| \leq w - d$. *In addition, for* $d \leq k \leq w$ *and* $\delta \in (0,1)$, *we say* $\mathcal{P}$ *is a* $(d, k, \delta)$-*design if it is a* $d$-*design and for all* $\alpha \in [u]$, *and* $i, j \in [v]$,

$$\Pr_{\beta \in_u [u]}\left[|P_i^\alpha \cap P_j^\beta| \leq w - k\right] \geq 1 - \delta.$$

We should think of $k \gg d$. Intuitively, the first condition says that any two blocks arising in our partitions differ in at least $d$ elements (i.e., do not overlap completely); in contrast, the second condition says that with probability $1 - \delta$, two random blocks differ in at least $k$ elements (i.e, have very little overlap if $k \gg d$).

When a collection of partitions $\mathcal{P} = \{P^1, \ldots, P^u\}$ satisfies the above condition, the following claim gives a formula for the bias of $f_{\mathcal{P}}$. To parse the formula, even if clearly false, suppose that the tribes involved in $f_{\mathcal{P}}$ were on disjoint sets of variables. Then, we would have

$$\Pr_{x \in_u \{0,1\}^n}[f_{\mathcal{P}}(x) = 1] = \prod_{\alpha=1}^{u} \Pr_{x \in_u \{0,1\}^n}[T_{P^\alpha}(x) = 1] = (1 - (1 - 2^{-w})^v)^u := \mathsf{bias}(u, v, w).$$

The next claim shows that when $\mathcal{P}$ forms a $(d, k, \delta)$-design, the different tribes behave as if they are on disjoint sets of variables and the above formula for the bias is approximately correct. As is sufficient in our applications, we specialize to the case where $\mathsf{bias}(u, v, w)$ is close to $1/2$; this corresponds to choosing[5] $v = \Theta(1)w2^w$ and $u$ such that $1/3 \leq \mathsf{bias}(u, v, w) \leq 2/3$.

**Theorem 2.3.** *Let* $\mathcal{P} = \{P^1, \ldots, P^u\}$ *be a collection of partitions of* $[n]$ *into* $w$-*sized blocks that is a* $(d, k, \delta)$-*design. Let* $u, v, w$ *be such that* $v = \Theta(1)w2^w$ *and* $1/3 \leq \mathsf{bias}(u, v, w) = (1 - (1 - 2^{-w})^v)^u \leq 2/3$. *Then,*

$$\left| \Pr_{x \in_u \{0,1\}^n} [f_{\mathcal{P}}(x) = 1] - \mathsf{bias}(u, v, w) \right| \leq \min \begin{cases} O(1)w \exp(-\Omega(d)) \\ O(1) \cdot (w \exp(-\Omega(k)) + \exp(-\Omega(d)) + 2^w \delta) \end{cases} .$$

The proof of Theorem 2.3 is similar to the arguments in [CZ15] and relies on Janson's inequality. However, our argument is more subtle as we need to handle $(d, k, \delta)$-designs and not just $d$-designs as is done there. In particular the theorem implies that when $\mathcal{P}$ is a $d$-design with $d \gg \log w$, the error is at most $\exp(-\Omega(d))$. On the other hand, when $d \ll \log w$ we can use the second formula if $k \gg \log w$. We need the more refined statement above where only most blocks are far from each other as this is what our construction achieves.

**Analyzing influences** We next specify a sufficient condition on a collection of partitions $\mathcal{P} = \{P^1, \ldots, P^u\}$ to guarantee that small coalitions have small influence on $f_{\mathcal{P}}$.

**Definition 2.4.** *Let* $\mathcal{P} = \{P^1, \ldots, P^u\}$ *be a collection of partitions of* $[n]$ *into* $w$-*sized blocks. We say* $\mathcal{P}$ *is* $(q, \tau)$-*load balancing if for all* $Q \subseteq [n]$ *with* $|Q| \leq q$, *and* $j \in [v]$,

$$\mathbb{E}_{\alpha \in_u [u]} \left[ 1(Q \cap P_j^\alpha \neq \emptyset) 2^{|Q \cap P_j^\alpha|} \right] \leq \tau \cdot (q/v).$$

To gain some intuition for the definition and the use of the name *load balancing* we first view partitions as hash functions. A partition $P = \{P_1, \ldots, P_v\}$ can be seen as a hash function $h_P : [n] \to [v]$: $h_P(i) = j$ if $i \in P_j$. We can then also view a collection of partitions $\mathcal{P}$ as defining a family of hash functions $\mathcal{H} = \{\langle_{\mathcal{P}} : \mathcal{P} \in \mathcal{P}\}$. The definition above then says that a random hash function from the family $\mathcal{H}$ is *load-balancing* in a certain concrete way. In particular, for any subset $Q \subseteq [n]$, and any particular bin $j \in [v]$, the number of items hashed into bin $j$ satisfy the following inequality:

$$\mathbb{E}_{h \in_u \mathcal{H}} \left[ 1(Q \cap h^{-1}(j) \neq \emptyset) 2^{|Q \cap h^{-1}(j)|} \right] \leq \tau \cdot (|Q|/v).$$

Indeed, standard Chernoff bounds imply that the above property is clearly satisfied by truly random hash functions. Thus, the partitions we construct can be seen as imitating this property of truly random hash functions but with a much smaller size family.

We show that load-balancing partitions give us fine control on the influences:

**Theorem 2.5.** *Let* $\mathcal{P} = \{P^1, \ldots, P^u\}$ *be a collection of partitions of* $[n]$ *into* $w$-*sized blocks that is* $(q, \tau)$-*load balancing. Then,*

$$I_q(f_{\mathcal{P}}) \leq (u(1 - 2^{-w})^{v-q}) \cdot (\tau 2^{-w}) \cdot q.$$

A similar claim is used in the analysis of [AL93, CZ15]. However, [CZ15] work with the stronger condition that for any $Q \subseteq [n]$ with $|Q| \ll q$, $|Q \cap P_j^\alpha| \ll w$ for most $\alpha$ (as opposed to just having a bound on the expectation of $2^{|Q \cap P_j^\alpha|}$). The above generalization, while straightforward, is important as one cannot hope to satisfy their stronger requirement for $Q$ very large ($n^{1-o(1)}$) as needed for the proof of Theorem 1.2.

---

[5]Our arguments also give analogous, albeit more cumbersome, bounds for all $u, v, w$ and even other product distributions.

## 2.1 Constructing nice partitions

We next outline how to construct a collection of partitions which is a *good* design as well as *sufficiently* load-balancing. Fix $v, w$. For a string $\alpha \in [v]^w$, define an associated partition $P^\alpha$ of $[n] \equiv [vw]$ into $w$-sized blocks as follows:

- Write $\{1, \ldots vw\}$ from left to right in $w$ blocks of length $v$ each. Now, permute the $k$'th block by shifting the integers in that block by adding $\alpha_k$ modulo $v$.

- The $i$'th part now comprises of the elements in the $i$'th position in each of the $w$ blocks.

Formally, for $i \in [v]$ $P_i^\alpha = \{(k-1)v + ((i - \alpha_k) \mod v) : k \in [w]\}$. As in [CZ15], our final function will be $f_{\mathcal{P}}$ for $\mathcal{P} := \mathcal{P}_{\mathcal{U}} = \{P^\alpha : \alpha \subseteq \mathcal{U}\}$ for a suitably chosen set of strings $\mathcal{U} \subseteq [v]^w$; in their work, $\mathcal{U}$ is chosen by using known *seeded* extractors.

For intuition, fix a constant $0 < \beta < 1$ and first consider the case where $\mathcal{U}$ is the set of Reed-Solomon codewords corresponding to degree $\ell \geq 1/\beta$ polynomials over $[v]$ [6]. Then, a simple calculation shows that $\mathcal{P}_{\mathcal{U}}$ is a $(w - \ell)$-design; that is, no two blocks in the partitions of $\mathcal{P}$ overlap in more than $\ell$ positions. Further, note that a random element of $\mathcal{U}$ is $\ell$-wise independent; combining this with standard Chernoff-type bounds for $\ell$-wise independent hash functions implies that $\mathcal{P}$ is $(q, O_\beta(1))$-load balancing for $q \ll n^{1-\beta}$. Setting the parameters appropriately and applying Theorems 2.3 and 2.5 shows that $f_{\mathcal{P}}$ is almost-balanced and $(n^{1-\beta})$-resilient—giving a simpler construction matching [CZ15].

For the main theorem, Theorem 1.2, we follow the same outline as above using Theorem 1.8 to get a suitable set of partitions. Indeed, a look at the definition of load-balancing suggests that Theorem 1.8 should be relevant as it also involves a similar expression. Concretely, as a first attempt, we take $\mathcal{U}$ to be the range of the generator from Theorem 1.8 and study the bias and influence of $f \equiv f_{\mathcal{P}_{\mathcal{U}}}$. Theorem 1.8 immediately implies that $\mathcal{P}_{\mathcal{U}}$ is $(\Omega(n/\log n), O(1))$-load balancing. This combined with Theorem 2.5 implies that $f$ is $O((\log^2 n)/n)$-strongly resilient.

To complete the proof, we need to show that $f$ is almost-balanced. Unfortunately, $\mathcal{P}_{\mathcal{U}}$ need not be a good-enough design to apply Theorem 2.3 directly. We get around this at a high-level by encoding parts of the output of the oblivious sampler in Theorem 1.8 using a Reed-Solomon code; doing so, we get $\mathcal{P}_{\mathcal{U}}$ to be a $(C, w/2, 2^{-Cw})$-design for any large constant $C$. We then apply Theorem 2.3 to show that $f$ is almost-balanced. We leave the details of the encoding to the actual proof.

## 2.2 Construction of the oblivious sampler

The generator is obtained by instantiating the Nisan-Zuckerman [NZ96] PRG for small-space machines with $k$-wise independent seeds being fed into the *extractor* rather than truly independent ones. Concretely, let $E : [v^c] \times [D] \to [v]$ be a $(c(\log v)/2, \varepsilon)$-extractor (see Section 3 for formal definitions) with error $\varepsilon \approx 1/w$. For $\ell = \Theta(w/(\log w))$, let $G_\ell : [D]^\ell \to [D]^w$ generate a $\ell$-wise independent distribution. Then, our generator satisfying Theorem 1.8, $G : [v^c] \times [D]^\ell \to [v]^w$ is defined as follows:

$$G(x, y) = (E(x, G_\ell(y)_1), E(x, G_\ell(y)_2), \cdots, E(x, G_\ell(y)_w)). \tag{2.1}$$

While the above construction serves as the base, in our proof of Theorem 1.2 we need a generator that satisfies certain additional constraints: output strings on different seeds should be far from each other. We satisfy these constraints by careful modifications of the above construction.

---

[6] Assuming for simplicity that $v$ is a prime.

## 2.3 Analyzing bias and resilience under limited independence

The extension to obtain resilient functions under limited independence, as in Theorem 1.7, follows from our analysis of Theorem 2.3 along with the following claim. The core of the proof of Theorem 2.3 involves calculating the biases of disjunctions of a small number of Tribes, i.e., functions of the form $h = \vee_{\alpha \in I} T_{P^\alpha}$, where $P^\alpha$ are partitions of $[n]$ and $|I|$ is not too large. We directly show that such functions are *fooled* to within error $\varepsilon$ by $O(|I|^2 w^2 + |I| w \log(1/\varepsilon))$-wise independence; that is, for any $O(|I|^2 w^2 + |I| w \log(1/\varepsilon))$-wise independent distribution $\mathcal{D}$ on $\{0,1\}^n$,

$$\left| \Pr_{x \in_u \{0,1\}^n} [\vee_{\alpha \in I} T_{P^\alpha}(x) = 1] - \Pr_{x \sim \mathcal{D}} [\vee_{\alpha \in I} T_{P^\alpha}(x) = 1] \right| < \varepsilon.$$

The proof of the above claim involves standard approximations based on the inclusion-exclusion principle and some inequalities involving elementary symmetric polynomials.

# 3 Preliminaries

## 3.1 Pseudorandomness

We first recall some standard notions from pseudorandomness.

**Definition 3.1.** *A collection of random variables $\{X_1, \ldots, X_m\}$ is $k$-independent if for any $I \subseteq [m], |I| \leq k$, the random variables $\{X_i : i \in I\}$ are independent of each other.*

**Definition 3.2.** *For a distribution $X$, $H_\infty(X) = \min_{x \in Support(X)} \log(1/\Pr[X = x])$.*

**Definition 3.3** ([NZ96]). *A function $E : [N] \times [D] \to [M]$ is a $(k, \varepsilon)$-strong extractor if for every distribution $X$ over $[N]$ with $H_\infty(X) \geq k$, and $Y \in_u [D]$, $(Y, E(X, Y))$ is $\varepsilon$-close in statistical distance to the uniform distribution over $[D] \times [M]$.*

We need the following sampling properties of strong extractors, c.f., [Zuc97].

**Lemma 3.4.** *Let $E : [N] \times [D] \to [M]$ be a $(k, \varepsilon)$-strong extractor. Then, for all functions $g_1, \ldots, g_D : [M] \to \{0, 1\}$, with $\mu = (1/D) \sum_i \mathbb{E}_{x \in_u [M]}[g_i(x)]$, there are at most $2^k$ elements $x \in [N]$ such that*

$$\left| \frac{1}{D} \sum_{z \in [D]} g_z(E(x, z)) - \mu \right| \geq \varepsilon.$$

We also need the following explicit extractor construction due to Zuckerman [Zuc97]:

**Theorem 3.5.** *There exists a constant $C \geq 1$ such that for all $\varepsilon > 0$, and $1 \leq M \leq N^{1/3}$, there is an explicit $((\log N)/2, \varepsilon)$-strong extractor $E : [N] \times [D] \to [M]$ with $D = ((\log N)/\varepsilon)^C$. We also assume without loss of generality that $E(X, Y)$ is uniformly random over $[M]$ when $X, Y$ are uniformly random over $[N]$ and $[D]$ respectively.*

Finally, we also need the following explicit generator of $\ell$-wise independent distributions which follows for instance from the *Reed-Solomon* code.

**Definition 3.6** (Shift-Hamming distance). *For two sequences $x, x'$ over an alphabet $[B]^d$, let $d_H(x, x') = \min_{a \in [B]} |\{i \in [d] : x_i - x'_i \neq a \mod B\}|$.*

**Lemma 3.7.** *For all prime $v$ and $1 \leq \ell \leq m \leq v$, there exists an explicit function $G_\ell : [v]^\ell \to [v]^m$ such that*

- *For any $y \neq y' \in [v]^\ell$, $d_H(y, y') \geq m - \ell$.*
- *For $y \in_u [v]^\ell$, $G_\ell(y)$ is $\ell$-wise independent.*

*Proof.* Follows by using the Reed-Solomon code over $[v]$ of degree $\ell$ and length $m$. $\qquad\square$

We need the following theorem of [DETT10] on fooling read-once Conjunctive Normal Formulas (CNFs) using limited independence. A CNF $f : \{0,1\}^n \to \{0,1\}$ is said to be a readonce formula if each variable appears in at most one clause. The width of a CNF is the maximum length of any clause in the CNF.

**Theorem 3.8.** *There exists a constant $C$ such that the following holds for all $0 < \varepsilon < 1$ and $t \geq Cw \log(1/\varepsilon)$: For any width $w$ read-once CNF $f : \{0,1\}^n \to \{0,1\}$ and $t$-wise independent distribution $\mathcal{D}$ on $\{0,1\}^n$,*

$$\left| \Pr_{x \leftarrow \mathcal{D}}[f(x) = 1] - \Pr_{x \in_u \{0,1\}^n}[f(x) = 1] \right| \leq \varepsilon.$$

## 3.2 Probability

We next review some results from probability theory.

We need Janson's inequality from probability theory, c.f., [AS11] that allows us to bound the bias of *monotone Disjunctive Normal Formulas (DNFs)*.

**Theorem 3.9.** *Let $S_1, \ldots, S_m \subseteq [n]$ be a collection of sets and define $f_i : \{0,1\}^n \to \{0,1\}$ by $f_i(x) = \wedge_{j \in S_i} x_j$. Let $x \in_u \{0,1\}^n$. Then,*

$$\prod_{i=1}^m \Pr\left[\neg f_i(x) = 1\right] \leq \Pr\left[\bigwedge_{i=1}^m (\neg f_i(x)) = 1\right] \leq \exp\left(\frac{\Delta}{1-\gamma}\right) \prod_{i=1}^m \Pr\left[\neg f_i(x) = 1\right],$$

*where $\gamma = \max_{i=1}^m \Pr[f_i(x) = 1]$ and*

$$\Delta = \sum_{i \neq j : S_i \cap S_j \neq \emptyset} \Pr[f_i(x) \wedge f_j(x) = 1].$$

The following immediate corollary is more convenient for us.

**Theorem 3.10.** *Let $S_1, \ldots, S_m \subseteq [n]$ be a collection of sets and define $f_i : \{0,1\}^n \to \{0,1\}$ by $f_i(x) = \wedge_{j \in S_i} x_j$. Let $x \in_u \{0,1\}^n$, and $Z_i = f_i(x)$ for $i \in [m]$. Then,*

$$\prod_{i=1}^m \mathbb{E}[(1 - Z_i)] \leq \mathbb{E}\left[\prod_{i=1}^m (1 - Z_i)\right] \leq \exp\left(\frac{\Delta}{1-\gamma}\right) \cdot \prod_{i=1}^m \mathbb{E}[(1 - Z_i)],$$

*where $\gamma = \max_{i=1}^m \mathbb{E}[Z_i]$ and*

$$\Delta = \sum_{i \neq j : S_i \cap S_j \neq \emptyset} \mathbb{E}[Z_i Z_j].$$

We need the following instantiation of Holder's inequality.

**Fact 3.11.** *Let $Y_1, \ldots, Y_r$ be real-valued random variables. Then,*

$$\mathbb{E}\left[|Y_1 Y_2 \cdots Y_r|\right] \leq \prod_{i=1}^r \mathbb{E}\left[|Y_i|^r\right]^{1/r}.$$

**Definition 3.12.** *For any collection of variables $x_1, \ldots, x_m$ and $1 \leq a \leq m$, $S_a(x_1, \ldots, x_m) = \sum_{I \subseteq [m], |I| = a} \prod_{i \in I} x_i$ denotes the $a$'th symmetric polynomial.*

We will use the following bound on moments of symmetric polynomials of independent Bernoulli random variables; see the appendix for proof.

9

**Lemma 3.13.** *Let $0 < p < 1/2$ and $X_1, \ldots, X_v$ be independent indicator random variables with $\Pr[X_i = 1] = p$. Then, for $k \geq 2e^2 vp$, and all $r \geq 1$,*

$$\mathbb{E}[S_k(X_1, \ldots, X_v)^r] \leq (1/2)^k.$$

We will use the following standard fact about symmetric polynomials.

**Fact 3.14.** *For all $1 \leq a \leq m$ and $0 \leq q_1, \ldots, q_m$ with $\sum_i q_i \leq \mu$, $S_a(q_1, \ldots, q_m) \leq \binom{m}{a} \cdot (\mu/m)^a$.*

We need the following elementary approximations; see the appendix for proofs.

**Fact 3.15.** *For all $x \geq 2$, $e^{-1}(1 - 1/x) \leq (1 - 1/x)^x \leq e^{-1}$.*

**Fact 3.16.** *Let $1 \leq w \leq v \leq u$, $B \geq 1$ with $0 \leq v - 2^w(\ln(u/\ln 2)) \leq B$, and $\theta = (1 - 2^{-w})^v$. Then, $(1 + \theta)^u = O(1)$ and $(1 - \theta)^u = 1/2 \pm O(B \ln u)2^{-w}$.*

# 4 Analyzing bias

Here we prove [Theorem 2.3](#). Let $\mathcal{P} = \{P^1, \ldots, P^u\}$, $f_{\mathcal{P}}$, be as in the theorem statement. Let $x \in_u \{0, 1\}^n$ and for $\alpha \in [u]$ let $Z_\alpha = T_{P^\alpha}(x)$. Then,

$$f_{\mathcal{P}}(x) = \prod_{\alpha \in [u]} T_{P^\alpha}(x) = \prod_{\alpha \in [u]} Z_\alpha.$$

The theorem essentially states that $\mathbb{E}[\prod_\alpha Z_\alpha] \approx \prod_\alpha \mathbb{E}[Z_\alpha]$. We will prove this by showing that $Z_\alpha$'s satisfy a weak notion of limited independence. Roughly speaking, we will show that for $I \subseteq [u], |I| \leq d$,

$$\mathbb{E}\left[\prod_{\alpha \in I}(1 - Z_\alpha)\right] \approx \prod_{\alpha \in I} \mathbb{E}[(1 - Z_\alpha)]. \tag{4.1}$$

The proof of [Equation 4.1](#) uses Janson's inequality and the design properties of $\mathcal{P}$. The theorem then follows by writing

$$\prod_\alpha Z_\alpha = \prod_\alpha (1 - (1 - Z_\alpha)) = \sum_{a=1}^u S_a(\{1 - Z_\alpha : \alpha \in [u]\})$$

and applying Bonferroni inequalities to truncate the latter expansion to the first $d$ terms; the error from the truncation is bounded by applying [Equation 4.1](#). We next gives a concrete form of [Equation 4.1](#).

**Lemma 4.1.** *Let $\mathcal{P}$ be as in [Theorem 2.3](#) and $\{Z_\alpha\}$'s be as defined above. Then, for all $a \leq d$,*

$$\binom{u}{a} \cdot (1 - 2^{-w})^{va} \leq \mathbb{E}\left[S_a(1 - Z_1, 1 - Z_2, \ldots, 1 - Z_u)\right] \leq \gamma \cdot \binom{u}{a} \cdot (1 - 2^{-w})^{va},$$

*where $\gamma = \exp\left(O(1)wd2^{-\Omega(k)}\right) + 2^{O(w)}\delta$.*

Note that $\mathbb{E}[Z_\alpha] = (1 - 2^{-w})^v$ for all $\alpha \in [u]$. Therefore, if the $Z_\alpha$'s were actually independent of each other, then $\mathbb{E}[S_a(1 - Z_1, 1 - Z_2, \ldots, 1 - Z_u)] = \binom{u}{a}(1 - 2^{-w})^{va}$. Thus, the lemma says that for a design, the expectation of the symmetric polynomial is close to what it would be if $\{Z_\alpha\}$'s were independent of each other.

Before proving the lemma, we first show it implies [Theorem 2.3](#).

*Proof of Theorem 2.3.* Let $Z_\alpha$'s be as above; for brevity, let $Y_\alpha = 1 - Z_\alpha$ and $\theta = (1 - 2^{-w})^v$. Note that

$$\Pr[f_{\mathcal{P}}(x) = 1] = \mathbb{E}\left[\prod_{\alpha \in [u]} Z_\alpha\right], \qquad \prod_{\alpha \in [u]} \mathbb{E}[Z_\alpha] = (1 - \theta)^u = \mathsf{bias}(u, v, w). \tag{4.2}$$

Now, by Bonferroni inequalities, we have

$$\prod_\alpha Z_\alpha = \prod_\alpha (1 - Y_\alpha) = \sum_{a=0}^{d-1} (-1)^t S_a(Y_1, \ldots, Y_u) \pm S_d(Y_1, \ldots, Y_u).$$

Thus,

$$\left| \mathbb{E}\left[\prod_\alpha Z_\alpha\right] - \sum_{a=0}^{d-1} (-1)^t \mathbb{E}[S_a(Y_1, \ldots, Y_u)] \right| \leq \mathbb{E}[S_d(Y_1, \ldots, Y_u)].$$

Further, by Lemma 4.1, for all $1 \leq a \leq d$,

$$\binom{u}{a} \theta^a \leq \mathbb{E}[S_a(Y_1, \ldots, Y_u)] \leq \gamma \binom{u}{a} \theta^a.$$

Let $Y_1', \ldots, Y_u'$ be independent random variables with $Y_i'$ having the same marginal as $Y_i$. Clearly, $\mathbb{E}[S_a(Y_1', \ldots, Y_u')] = \binom{u}{a} \theta^a$. Therefore, for any $a \leq d$,

$$|\mathbb{E}[S_a(Y_1, \ldots, Y_u)] - \mathbb{E}[S_a(Y_1', \ldots, Y_u')]| \leq (\gamma - 1) \binom{u}{a} \theta^a.$$

Combining the above inequalities, we get

$$\left| \mathbb{E}\left[\prod_i Z_i\right] - \sum_{a=0}^{d-1} (-1)^t \mathbb{E}[S_a(Y_1', \ldots, Y_u')] \right| \leq \mathbb{E}[S_d(Y_1, \ldots, Y_u)] + (\gamma - 1) \sum_{a=1}^{d-1} \binom{u}{a} \theta^a$$

$$\leq \gamma \binom{u}{d} \theta^d + (\gamma - 1)(1 + \theta)^u.$$

Note that the above arguments also apply to the case when $Z_i$'s were truly independent of each other with $\gamma = 1$. Therefore,

$$\left| \mathbb{E}[Z_1 \cdots Z_u] - \prod_{i \in [u]} \mathbb{E}[Z_i] \right| \leq 2\gamma \binom{u}{d} \theta^d + 2(\gamma - 1)(1 + \theta)^u.$$

We next simplify the error bound by plugging in the values of $\gamma$ and $\theta$. Recall that $\mathsf{bias}(u, v, w) = (1 - \theta)^u \in [1/3, 2/3]$ so that $(1 + \theta)^u = O(1)$ and $\theta = O(1/u)$. Therefore,

$$\left| \mathbb{E}[\prod_{\alpha \in u} Z_\alpha] - \prod_{\alpha \in u} \mathbb{E}[Z_\alpha] \right| = \exp(-\Omega(d)) + O(\gamma - 1). \tag{4.3}$$

Note that 1 is a trivial upper bound on the left-hand side; further, a simple calculation shows that

$$\min(1, \gamma - 1) = \min\left(1, \exp\left(O(wd)2^{-\Omega(k)}\right) + 2^{O(w)}\delta - 1\right)$$

$$\leq \min\left(1, \exp\left(O(wd)2^{-\Omega(k)}\right) - 1\right) + 2^{O(w)}\delta$$

$$= O(wd)2^{-\Omega(k)} + 2^{O(w)}\delta = O(w)2^{-\Omega(k)} + 2^{O(w)}\delta,$$

11

where we used the fact that for all $\lambda > 0$ $\min(1, e^\lambda - 1) = O(\lambda)$ and $d2^{-\Omega(k)} \leq k2^{-\Omega(k)} = O\left(2^{-\Omega(k)}\right)$.

Combining the above with Equation 4.2 and Equation 4.3 we get

$$|\Pr[f_{\mathcal{P}}(x) = 1] - \mathsf{bias}(u, v, w)| \leq \exp(-\Omega(d)) + O(w)\exp(-\Omega(k)) + 2^{O(w)}\delta.$$

This proves the second bound of the theorem. The first bound is a special case as any $(d, k, \delta)$-design is also a $(d, d, 0)$-design by definition. $\square$

*Proof of Lemma 4.1.* We first further break up each $Z_\alpha$ according to the blocks of the partition $P^\alpha$; for each $i \in [v]$, let $Z_{\alpha i} = \wedge_{j \in P_i^\alpha} x_j$ (recall $x \in_u \{0, 1\}^n$). Then, $Z_\alpha = \vee_{i \in [v]} Z_{\alpha i}$ so that $1 - Z_\alpha = \prod_{i \in [v]}(1 - Z_{\alpha i})$. Therefore, for any $I \subseteq [u]$,

$$\prod_{\alpha \in I}(1 - Z_\alpha) = \prod_{\alpha \in I}\prod_{i \in v}(1 - Z_{\alpha i}).$$

Note that we can apply Theorem 3.10 to the latter product. Fix a set $I \subseteq [u]$. Define $(\alpha, i) \neq (\beta, j)$ to be adjacent, $(\alpha, i) \sim (\beta, j)$, if $P_i^\alpha \cap P_j^\beta \neq \emptyset$. Let,

$$\Delta_I = \sum_{(\alpha, i) \sim (\beta, j) \in I \times [v]} \mathbb{E}[Z_{\alpha i} \cdot Z_{\beta j}].$$

We next bound $\Delta_I$. Let $d_I = w - \max\{|P_i^\alpha \cap P_j^\beta| : \alpha \neq \beta \in I, i, j \in [v]\}$ quantify the maximum overlap among any two blocks of the partitions $P^\alpha, \alpha \in I$. Then, as $\mathcal{P}$ is a $d$-design, $d_I \geq d$ for all $I$. Fix $\alpha \neq \beta \in I$ and an in index $i \in [v]$, and let $j_1, \ldots, j_b \in [v]$ be the indices such that $(\alpha, i) \sim (\beta, j)$, and let $w_\ell = |P_i^\alpha \cap P_{j_\ell}^\beta|$. Then, $1 \leq w_1, \ldots, w_b \leq w - d_I$. Now,

$$\sum_{j \in [v]: (\beta, j) \sim (\alpha, i)} \mathbb{E}[Z_{\alpha i} Z_{\beta j}] = \sum_{\ell=1}^{b} 2^{-|P_i^\alpha \cup P_{j_\ell}^\beta|} = \sum_{\ell=1}^{b} 2^{-(2w - w_\ell)} = 2^{-2w} \cdot \sum_{\ell=1}^{b} 2^{w_\ell}.$$

As $w_i \in [w - d_I]$, and $\sum_i w_i = w$, the above expression is maximized by setting as many of the $w_i$'s to $w - d_I$ as possible. Thus,

$$\sum_{\ell=1}^{b} 2^{w_\ell} \leq \lceil w/(w - d_I) \rceil 2^{w - d_I} = 2^w \lceil w/(w - d_I) \rceil 2^{-d_I} \leq 2^{w+1} 2^{-d_I/2}.$$

Therefore,

$$\sum_{j \in [v]: (\beta, j) \sim (\alpha, i)} \mathbb{E}[Z_{\alpha i} Z_{\beta j}] = 2^{1 - w - d_I/2}.$$

Summing over all indices $(\alpha, i)$, we get

$$\Delta_I \leq (vd) \cdot 2^{1 - w - d_I/2} = O(wd2^{-d_I/2}).$$

Finally, observe that $\max_{\alpha, i} \mathbb{E}[Z_{\alpha i}] = 2^{-w} \leq 1/2$. Thus, by Theorem 3.10,

$$\prod_{\alpha \in I} \mathbb{E}[(1 - Z_\alpha)] \leq \mathbb{E}\left[\prod_{\alpha \in I}(1 - Z_\alpha)\right] \leq \exp(2\Delta_I) \cdot \prod_{\alpha \in I} \mathbb{E}[(1 - Z_\alpha)]. \tag{4.4}$$

Note that $\mathbb{E}[(1 - Z_\alpha)] = (1 - 2^{-w})^v$. Therefore, summing over all sets $I \subseteq [u], |I| = a$ gives us the lower bound of the claim. For the upper bound, we have

$$
\mathbb{E}\left[S_a(1 - Z_1, \ldots, 1 - Z_u)\right] \leq (1 - 2^{-w})^{va} \cdot \sum_{I \in \binom{[u]}{a}} \exp(2\Delta_I)
$$

$$
\leq (1 - 2^{-w})^{va} \cdot \sum_{I \in \binom{[u]}{a}} \exp(O(wd2^{-d_I/2})))
$$

$$
= (1 - 2^{-w})^{va} \cdot \binom{u}{a} \cdot \mathbb{E}_{I \in_u \binom{[u]}{a}}\left[\exp(O(wd2^{-d_I/2})))\right].
$$

We next bound the last expectation. From the design properties of $\mathcal{D}$ and a union bound applied to all possible pairs $(\alpha, i), (\beta, j) \in I \times [v]$, it follows that for $I \in_u \binom{[u]}{a}$,

$$
\Pr[d_I < k] = \Pr\left[\exists \alpha \neq \beta \in I, \ \ i, j \in [v] \ \ |P_i^\alpha \cap P_j^\beta| \geq w - k\right]
$$

$$
\leq \sum_{i,j \in [v]} \Pr\left[\exists \alpha \neq \beta \in I \ \ |P_i^\alpha \cap P_j^\beta| \geq w - k\right]
$$

$$
\leq v^2 d^2 \delta.
$$

Hence,

$$
\mathbb{E}\left[\exp\left(O(1)wd2^{-d_I/2}\right)\right] = \Pr[d_I \leq k] \cdot \mathbb{E}\left[\exp\left(O(1)wd2^{-d_I/2}\right) \mid d_I \leq k\right] +
$$

$$
\Pr[d_I > k] \cdot \mathbb{E}\left[\exp\left(O(1)wd2^{-d_I/2}\right) \mid d_I > k\right]
$$

$$
\leq \exp\left(O(1)wd2^{-k/2}\right) + (v^2 d^2 \delta) \cdot \exp\left(O(1)wd2^{-d/2}\right)
$$

$$
= \exp\left(O(1)wd2^{-k/2}\right) + 2^{O(w)}\delta.
$$

This proves the lemma. $\qquad\qquad\square$

## 4.1 Analyzing bias under limited independence

We next a state a version of Theorem 2.3 for distributions with limited independence.

**Corollary 4.2.** *Let $\mathcal{P} = \{P^1, \ldots, P^u\}$ be a collection of partitions of $[n]$ into $w$-sized blocks that is a $(d, k, \delta)$-design and $\varepsilon \in (0, 1)$. Let $u, v, w$ be such that $v = \Theta(1)w2^w$ and $1/3 \leq \mathsf{bias}(u, v, w) = (1 - (1 - 2^{-w})^v)^u \leq 2/3$. Let $\mathcal{D}$ be a $(C(w^2 d^2 + wd \log(1/\varepsilon)))$-wise independent distribution for $C$ a sufficiently large constant. Then,*

$$
\left|\Pr_{x \sim \mathcal{D}}[f_{\mathcal{P}}(x) = 1] - \mathsf{bias}(u, v, w)\right| \leq \varepsilon + \min \begin{cases} O(1)w \exp(-\Omega(d)) \\ O(1) \cdot (w \exp(-\Omega(k)) + \exp(-\Omega(d)) + 2^w \delta) \end{cases}.
$$

We prove the corollary by repeating the proof of Theorem 2.3 while using an analogue of Lemma 4.1 for $x \sim \mathcal{D}$. To do so, we will use the following claim saying that limited independence is sufficient to fool disjunctions of a few Tribes.

**Claim 4.3.** *The following holds for some constant $C > 1$. Let $\{P^1, \ldots, P^r\}$ be a collection of partitions of $[n]$ into $v$ blocks of length $w$ each with $v = \Theta(1)w2^w$. Then, for any $Cwr(w + r \log(1/\varepsilon) + r^2)$-wise independent distribution $\mathcal{D}$ over $\{0, 1\}^n$,*

$$
\left|\Pr_{x \in_u \{0,1\}^n}\left[\vee_{\alpha \in [r]} T_{P^\alpha}(x) = 0\right] - \Pr_{x \sim \mathcal{D}}\left[\vee_{\alpha \in [r]} T_{P^\alpha}(x) = 0\right]\right| \leq \varepsilon.
$$

13

*Proof.* The lemma follows from some careful approximations of elementary symmetric polynomials. We first setup some notation. Let $k > \max(2r, 2e^2 v 2^{-w})$ be an even integer to be chosen later and let $x \sim \mathcal{D}$ where $\mathcal{D}$ is a $t$-wise independent distribution for $t \geq kwr$. For $\alpha \in [r]$, let $X^\alpha = (X_1^\alpha, \ldots, X_v^\alpha)$ where $X_i^\alpha = \wedge_{\ell \in P_i^\alpha} x_\ell$. Note that $T_{P^\alpha}(x) = \vee_{i=1}^u X_i^\alpha$. We will use inclusion-exclusion to approximate each $T_{P^\alpha}(x)$. Let $p_k, e_k : \{0,1\}^v \to \{0,1\}$ be defined by

$$p_k(z_1, \ldots, z_v) = \sum_{\ell=0}^{k-1} (-1)^\ell S_\ell(z_1, \ldots, z_v), \quad e_k(z_1, \ldots, z_v) = \sum_{\ell=k}^{n} (-1)^\ell S_\ell(z_1, \ldots, z_v).$$

By the inclusion-exclusion formula, for all $z \in \{0,1\}^v$, $\neg OR(z) = p_k(z) + e_k(z)$. Further, by Bonferroni inequalities, for all $z \in \{0,1\}^v$,

$$0 \leq e_k(z) \leq S_k(z). \tag{4.5}$$

Now,

$$1 - \vee_{\alpha \in [r]} T_{P^\alpha}(x) = \prod_{\alpha \in [r]} \neg OR(X^\alpha) = \prod_{\alpha \in [r]} (p_k(X^\alpha) + e_k(X^\alpha)) \tag{4.6}$$

$$= \underbrace{\prod_{\alpha \in [r]} p_k(X^\alpha)}_{:=P_k(x)} + \underbrace{\sum_{I \neq \emptyset \subseteq [r]} \prod_{\alpha \notin I} p_k(X^\alpha) \prod_{\alpha \in I} e_k(X^\alpha)}_{:=E_k(x)}.$$

We view $P_k$ as a low-degree polynomial approximation for the left-hand-side and $E_k$ as the error term. Indeed, $P_k$ is of degree at most $kwr$ in $x$ as each $p_k(X^\alpha)$ is of degree at msot $k$ in $\{X_1^\alpha, \ldots, X_v^\alpha\}$'s which in-turn are of degree at most $w$ in $x$. Therefore, for $t \geq kwr$, $\mathbb{E}[P_k(x)]$ is the same under all $t$-wise independent distributions. We next bound the expectation of $E_k$ under $t$-wise independent distributions. By Equation 4.5, $0 \leq e_k(z) \leq S_k(z)$; further, $|p_k(z)| = |\neg OR(z) - e_k(z)| \leq 1 + e_k(z) \leq 1 + S_k(z)$. Thus, for any fixed $\emptyset \neq I \subseteq [r]$,

$$\mathbb{E}\left[\left|\prod_{\alpha \in I} e_k(X^\alpha) \prod_{\alpha \notin I} p_k(X^\alpha)\right|\right] \leq \mathbb{E}\left[\prod_{\alpha \in I} S_k(X^\alpha) \prod_{\alpha \notin I} (1 + S_k(X^\alpha))\right]$$

$$\leq \prod_{\alpha \in I} \mathbb{E}[S_k(X^\alpha)^r]^{1/r} \cdot \prod_{\alpha \notin I} \mathbb{E}[(1 + S_k(X^\alpha))^r]^{1/r}$$

(by Fact 3.11)

$$\leq \prod_{\alpha \in I} \mathbb{E}[S_k(X^\alpha)^r]^{1/r} \cdot \prod_{\alpha \notin I} \left(1 + \mathbb{E}[S_k(X^\alpha)^r]^{1/r}\right)$$

(by Minkowski's inequality).

Note that $S_k(X^\alpha)^r$ is of degree at most $kwr$ as a polynomial in $x$. Therefore, $\mathbb{E}[S_k(X^\alpha)^r]$ is the same under all $t$-wise independent distributions and in particular the same as for $x \in_u \{0,1\}^n$. However, in this case $(X_1^\alpha, \ldots, X_v^\alpha)$ are independent indicator random variables with $\mathbb{E}[X_i^\alpha] = 2^{-w}$. Therefore, as $k \geq 2e^2 v 2^{-w}$, by Lemma 3.13,

$$\mathbb{E}[S_k(X^\alpha)^r] \leq (1/2)^k \leq 1.$$

Combining the above estimates we get that for all $I \neq \emptyset \subseteq [r]$,

$$\mathbb{E}\left[\left|\prod_{\alpha \in I} e_k(X^\alpha) \prod_{\alpha \notin I} p_k(X^\alpha)\right|\right] \leq (1/2)^{(|I|)(k/r)} \cdot 2^{r-|I|} \leq 2^r \cdot (1/2)^{k/r}.$$

14

Plugging the above into equation [Equation 4.6](#) we get $\mathbb{E}[|E_k(x)|] \leq 2^{2r}(1/2)^{k/r}$. Finally, as the above estimates hold under any $t$-wise independent distribution and $P_k$ is of degree at most $kwr$, we get that

$$\left| \Pr_{x \in_u \{0,1\}^n} \left[ \vee_{\alpha \in [r]} T_{P^\alpha}(x) = 0 \right] - \Pr_{x \sim \mathcal{D}} \left[ \vee_{\alpha \in [r]} T_{P^\alpha}(x) = 0 \right] \right| \leq \mathbb{E}_{x \in_u}[|E_k(x)|] + \mathbb{E}_{x \sim \mathcal{D}}[|E_k(x)|] \leq 2^{2r+1}(1/2)^{k/r}.$$

Plugging in $k = \max(3r(r + \log(1/\varepsilon)), 2ev2^{-w})$ we get the above error to be at most $\varepsilon$. It follows that it suffices for $t = kwr$ to be $C(wr(r^2 + r\log(1/\varepsilon) + w))$ for a sufficiently big constant $C$. $\quad\square$

*Proof of [Corollary 4.2](#).* The proof of the corollary is similar to that of [Lemma 4.1](#) and [Theorem 2.3](#) with one change. Using the notation from the above proof, for any $I \subseteq [u]$, the event $\prod_{\alpha \in I}(1 - Z_\alpha)$ corresponds to the unsatisfiability of $\vee_{\alpha \in I} T_{P^\alpha}$. By using the above theorem, we get an analogue of [Equation 4.4](#) for the present case as well: for any $I \subseteq [u]$ with $|I| \leq d$,

$$\prod_{\alpha \in I} \mathbb{E}[(1 - Z_\alpha)] - \varepsilon \leq \mathbb{E}\left[ \prod_{\alpha \in I}(1 - Z_\alpha) \right] \leq \exp(2\Delta_I) \cdot \prod_{\alpha \in I} \mathbb{E}[(1 - Z_\alpha)] + \varepsilon. \tag{4.7}$$

The corollary now follows by using the above inequality in place of [Equation 4.4](#) in the rest of the proof of [Theorem 2.3](#). $\quad\square$

# 5 Analyzing influence

*Proof of [Theorem 2.5](#).* Let $Q \subseteq [n]$ with $|Q| = q$. Note that for every $\alpha \in [u]$, a partial assignment $x$ to the variables not in $Q$ leaves $T_{P^\alpha}$ undetermined if and only if

1. For every part $P_j^\alpha$ that does not intersect $Q$, $x_i = 0$ for some $i \in P_j^\alpha$.
2. For some $j \in [v]$ with $P_j^\alpha \cap Q \neq \emptyset$, $x_i = 1$ for every $i \in (P_j^\alpha \setminus Q)$.

The above two events are independent of each other. The probability of (1) is at most $(1 - 2^{-w})^{v-q}$ as there are at least $v - q$ parts of $P^\alpha$ that do not intersect $Q$. The probability of (2) is at most

$$\sum_{j \in [v]: P_j^\alpha \cap Q \neq \emptyset} 2^{-\left(w - |P_j^\alpha \cap Q|\right)}.$$

Therefore, for $\alpha \in_u [u]$,

$$I_Q(f_\mathcal{P}) \leq \sum_{\alpha \in [u]} I_Q(T_{P^\alpha}) \leq u \cdot \mathbb{E}_{u \in_u [u]}[I_Q(T_{P^\alpha})]$$

$$\leq u \cdot \mathbb{E}\left[ (1 - 2^{-w})^{v-q} \cdot \left( \sum_{j \in [v]: P_j^\alpha \cap Q \neq \emptyset} 2^{-w + |P_j^\alpha \cap Q|} \right) \right]$$

$$\leq u(1 - 2^{-w})^{v-q} 2^{-w} \cdot \left( \sum_{j \in [v]} \mathbb{E}_{\alpha \in_u [u]}\left[ 1(P_j^\alpha \cap Q \neq \emptyset) 2^{|P_j^\alpha \cap Q|} \right] \right)$$

$$\leq u(1 - 2^{-w})^{v-q} \cdot 2^{-w} \cdot \tau q.$$

$\square$

We next a state a version of [Theorem 2.5](#) for distributions with limited independence.

**Corollary 5.1.** *Let $\mathcal{P} = \{P^1, \ldots, P^u\}$ be a collection of partitions of $[n]$ into $w$-sized blocks that is $(q, \tau)$-load balancing. Then, for all $t \geq Cw\log(1/\varepsilon)$ for some sufficiently big constant $C$,*

$$I_{q,t}(f_{\mathcal{P}}) \leq (u(1 - 2^{-w})^{v-q}) \cdot (\tau 2^{-w})q + (uv)\varepsilon.$$

*Proof.* The argument is similar to that of Corollary 4.2. Let $\mathcal{D}$ be a $t$-wise independent distribution on $\{0,1\}^n$. As in the above proof, observe that $Q$ leaves $T_{P^\alpha}$ undetermined if and only if

1. For every part $P_j^\alpha$ that does not intersect $Q$, $x_i = 0$ for some $i \in P_j^\alpha$.

2. For some $j \in [v]$ with $P_j^\alpha \cap Q \neq \emptyset$, $x_i = 1$ for every $i \in (P_j^\alpha \setminus Q)$.

However, the above two events are no longer independent. For a fixed $\alpha$, let $J \subseteq [v]$ be all parts $P_j^\alpha$ that do not intersect $Q$. Then, the first condition above is equivalent to

$$f_1(x) := \bigwedge_{j \in J} \left( \bigvee_{i \in P_j^\alpha} (\neg x_i) \right).$$

Similarly, the second condition is equivalent to

$$f_2(x) := \bigvee_{j \notin J} \left( \bigwedge_{i \in P_j^\alpha \setminus Q} x_j \right).$$

We are interested in

$$\Pr_{x \leftarrow \mathcal{D}}[f_1(x) \wedge f_2(x) = 1] = \Pr_{x \leftarrow \mathcal{D}}[f_1(x) = 1] - \Pr_{x \leftarrow \mathcal{D}}[f_1(x) \wedge (\neg f_2(x)) = 1].$$

Now, observe that

$$f_3(x) := f_1(x) \wedge (\neg f_2(x)) = \bigwedge_{j \in J} \left( \bigvee_{i \in P_j^\alpha} (\neg x_i) \right) \bigwedge_{j \notin J} \left( \bigvee_{i \in P_j^\alpha \setminus Q} (\neg x_j) \right).$$

Now, $f_1, f_3$ are both read-once CNFs, that is CNF formulas where each variable appears at most once. Therefore, by Theorem 3.8, for $t \gg w\log(1/\varepsilon)$,

$$\left| \Pr_{x \leftarrow \mathcal{D}}[f_1(x) = 1] - \Pr_{x \in_u \{0,1\}^n}[f_1(x) = 1] \right| \leq \varepsilon, \quad \left| \Pr_{x \leftarrow \mathcal{D}}[f_3(x) = 1] - \Pr_{x \in_u \{0,1\}^n}[f_3(x) = 1] \right| \leq \varepsilon.$$

Therefore,

$$\Pr_{x \leftarrow \mathcal{D}}[f_1(x) \wedge f_2(x) = 1] \leq \Pr_{x \in_u \{0,1\}^n}[f_1(x) \wedge f_2(x) = 1] + 2\varepsilon \leq$$

$$(1 - 2^{-w})^{v-q} \cdot \left( \sum_{j \in [v]: P_j^\alpha \cap Q \neq \emptyset} 2^{-w + |P_j^\alpha \cap Q|} \right) + 2\varepsilon,$$

where the last inequality follows from the arguments of Theorem 2.5. The main statement now follows by repeating the calculations of Theorem 2.5 with the above equation leading to an additional error of $2(uv)\varepsilon$. $\qquad\square$

# 6 Oblivious sampler preserving the moment generating function

Here we prove Theorem 1.8; it will be the main building block in our final construction of resilient functions. With a view towards future use, we modify the construction presented in the introduction (Equation 2.1) even if the simpler construction described there suffices for Theorem 1.8.

Let $\varepsilon = \mu/w$, and $c$ a sufficiently big constant to be chosen later. Let $E : [v^c] \times [D] \to [v/D]$ be a $((c \log v)/2, \varepsilon)$-strong extractor as in Theorem 3.5 with $D = ((\log v)/\varepsilon)^C$ for some universal constant $C$. Without loss of generality, suppose that $D$ is prime. For a parameter $\ell \geq 1$ to be chosen later, let $G_\ell : [D]^\ell \to [D]^w$ generate a $\ell$-wise independent distribution as in Lemma 3.7.

Define $G : [v^c] \times [D]^\ell \to [v]^w$ as follows:

$$G(x, y)_i = G_\ell(y)_i \circ E(x, G_\ell(y)_i), \tag{6.1}$$

where we associate $[D] \times [v/D]$ with $[v]$ in a straightforward manner.

To analyze the generator we shall use the following lemma about random variables with limited independence. A similar statement appears in [GKM15]; however, our setting is considerably simpler and we give a direct proof in the appendix.

**Lemma 6.1.** *Let $Y_1, \ldots, Y_w$ be $\ell$-wise independent random variables supported on $[0, 1]$. Then, for all $\theta > 0$,*

$$\mathbb{E}\left[\exp(\theta(Y_1 + \cdots + Y_w))\right] \leq \prod_i \mathbb{E}[\exp(\theta Y_i)] + \exp(2\theta w) \left(\frac{e \sum_i \mathbb{E}[Y_i]}{\ell}\right)^\ell.$$

Towards proving Theorem 1.8, we first prove a lemma with some precise but cumbersome bounds on the moment generating function.

**Lemma 6.2.** *Let $G : [v^c] \times [D]^\ell \to [v]^w$ be as in Equation 6.1 and let $f_1, \ldots, f_w : [v] \to [0, 1]$ be functions with $\sum_i \mathbb{E}[f_i] = \mu$. Let $\alpha = G(x, y)$ for $(x, y) \in_u [v^c] \times [D]^\ell$. Then, for all $\theta \geq 0$,*

$$\mathbb{E}_\alpha\left[\exp\left(\theta \sum_i f_i(\alpha_i)\right)\right] \leq \exp(e^\theta w\varepsilon) \cdot \exp(e^\theta \mu) + \exp(2\theta w) \left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell + \exp(\theta w) \cdot (wv^{-c/2}).$$

*Proof.* Let $Y_i = f_i(\alpha_i)$. Note that for any fixed $x$, the random variables $Y_1, \ldots, Y_w$ are $\ell$-wise independent with respect to the randomness of $y$. We bound the expectation of $\exp\left(\theta \sum_i f_i(\alpha_i)\right) = \exp\left(\theta \sum_i Y_i\right)$ as follows:

- Using sampling properties of extractors, Lemma 3.4, w show that with probability at least $1 - wv^{-c/2}$ over $x \in_u [v^c]$, $\sum_i \mathbb{E}_y[Y_i] \leq \mu + w\varepsilon$.

- We then apply the previous lemma to the $Y_i$'s conditioned on $x$ satisfying the above event.

For $i \in [w]$, call $x \in [v^c]$ $i$-bad if

$$|\mathbb{E}[Y_i|x] - \mu_i| = \left|\mathbb{E}_{z \in_u [D]}[f_i(z \circ E(x, z))] - \mu_i\right| \geq \varepsilon.$$

Call $x \in [v^c]$ *bad* if it is $j$-bad for some $j \in [w]$ and *good* otherwise. Fix $j \in [w]$. For $z \in [D]$, define $g_z : [v/D] \to [0, 1]$ by $g_z(x') = f_j(z \circ x')$. Then, $\sum_{z \in [D]} \mathbb{E}_{x' \in_u [v/D]}[g_z(x')] = D \mathbb{E}_j[f_j] = D\mu_j$; thus, $x$ is $j$-bad if and only if

$$\left|(1/D) \sum_{z \in [D]} g_z(E(x, z)) - \mu_j\right| \geq \varepsilon.$$

17

Therefore, by Lemma 3.4, for every $j \in [w]$, there are at most $v^{c/2}$ bad strings. Thus, $\Pr_{x \in_u [v^c]}[x \text{ is } j\text{-bad}] \leq v^{-c/2}$. Then, by a union bound, $\Pr_{x \in_u [v^c]}[\text{x is bad}] \leq wv^{-c/2}$. Finally, conditioned on $x$ being good,

$$\sum_j \mathbb{E}_y[Y_j] = \sum_j \mathbb{E}_{z \in_u [D]}[f_j(z \circ E(x,z))] \leq \sum_j (\mu_j + \varepsilon) = \mu + w\varepsilon. \tag{6.2}$$

Now, conditioned on $x$, the random variables $Y_1, \ldots, Y_w$ are $\ell$-wise independent by the definition of $G(x,y)$. Therefore, by Lemma 6.1, for good $x$,

$$\mathbb{E}_y\left[\exp\left(\theta \sum_i f_i(\alpha_i)\right)\right] = \mathbb{E}_y\left[\exp\left(\sum_i Y_i\right)\right]$$

$$\leq \prod_i \mathbb{E}_y[\exp(\theta Y_i)] + \exp(2\theta w)\left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell$$

$$\leq \prod_i \left(1 + (e^\theta)\mathbb{E}_y[Y_i]\right) + \exp(2\theta w)\left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell$$

$$\leq \exp\left((e^\theta)\sum_i \mathbb{E}_y[Y_i]\right) + \exp(2\theta w)\left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell$$

$$\leq \exp((e^\theta)(\mu + w\varepsilon)) + \exp(2\theta w)\left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell.$$

Finally, as $\exp(\theta \sum_i f_i(\alpha_i)) \leq \exp(\theta w)$ always, we get that

$$\mathbb{E}_{x,y}\left[\exp\left(\theta \sum_i f_i(\alpha_i)\right)\right] \leq \exp((e^\theta)(\mu + w\varepsilon)) + \exp(2\theta w)\left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell + \Pr[\text{x is bad}]\exp(\theta w)$$

$$\leq \exp(e^\theta w\varepsilon) \cdot \exp(e^\theta \mu) + \exp(2\theta w)\left(\frac{e(\mu + w\varepsilon)}{\ell}\right)^\ell + \exp(\theta w) \cdot (wv^{-c/2}).$$

$\square$

*Proof of Theorem 1.8.* The theorem follows by applying the above lemma with $\varepsilon = \mu/w$, $\theta = \ln 2$, $\ell = 12w/(\log w)$. Note that for this setting, there exist extractors as in Equation 6.1 with $D = ((\log v)/\varepsilon)^{O(1)}$ so that

$$\log D = O(\log \log v + \log w + \log(1/\mu)).$$

With this setup, as $\mu \leq 1$, we get

$$\mathbb{E}_{z \in_u \{0,1\}^r}\left[\exp\left(\theta \sum_i f_i(G(z)_i)\right)\right] \leq \exp(2w\varepsilon) \cdot \exp(2\mu) + \exp(2w)\left(\frac{2e\mu}{\ell}\right)^\ell + \exp(w) \cdot wv^{-c/2}$$

$$\leq 1 + O(\mu) + \exp(w) \cdot wv^{-c/2},$$

as $\varepsilon \leq \mu/w$. We now set $c = C \max(1, (w + \log(1/\mu))/(\log v))$ for a sufficiently large constant $C$ so that the last term is also $O(\mu)$.

The seed-length of the generator is

$$r = c \log v + \ell(\log D) = O(\log v + w + w((\log \log v) + (\log(1/\mu)))/(\log w)).$$

The theorem now follows. $\square$

# 7 Explicit resilient functions

Here we present our main construction proving Theorem 1.2. Fix $v, w$. For a string $\alpha \in [v]^w$, define an associated partition $P^\alpha$ of $[n] \equiv [vw]$ into $w$-sized blocks as follows:

- Write $\{1, \ldots vw\}$ from left to right in $w$ blocks of length $v$ each. Now, permute the $k$'th block by shifting the integers in that block by adding $\alpha_k$ modulo $v$.

- The $i$'th part now comprises of the elements in the $i$'th position in each of the $w$ blocks.

Formally, for $i \in [v]$ $P_i^\alpha = \{(k-1)v + ((i - \alpha_k) \mod v) : k \in [w]\}$. As in [CZ15], our final function will be $f_\mathcal{P}$ for $\mathcal{P} = \{P^\alpha : \alpha \subseteq \mathcal{U}\}$ for a suitably chosen set of strings $\mathcal{U} \subseteq [v]^w$.

## 7.1 Polynomially resilient functions from Reed-Solomon code

For intuition, we first use our arguments to present a simpler variant of the construction of [CZ15] (e.g., the function below is depth 3 as opposed to the depth 4 construction of [CZ15]) to get a $(n^{1-\delta})$-resilient function from Reed-Solomon codes as alluded to in the introduction.

Let $1 \leq w \leq v$, where $v$ is prime. For some parameter $\ell \geq 1$ to be chosen later, let $G_\ell : [v]^\ell \to [v]^w$ be an $\ell$-wise independent generator as in Lemma 3.7 and let $\mathcal{RS} = \{G_\ell(x) : x \in [v]^\ell\}$. Let $f \equiv f_\mathcal{RS} = f_\mathcal{P}$, where $\mathcal{P} = \{P^\alpha : \alpha \in \mathcal{RS}\}$. We show that for any constant $0 < \beta < 1$, and $\ell \geq 1/2\beta$, $f_\mathcal{RS}$ is $\Omega(n^{1-\beta})$-resilient and has bias $1/2 \pm n^{-\Omega(1)}$.

**Lemma 7.1.** *For $1 \leq \ell \leq w$, $f_\mathcal{RS}$ as defined above is a $(w - \ell)$-design.*

*Proof.* First note that for any $\alpha, \beta \in [v]^w$ and $i, j \in [v]$,, $|P_i^\alpha \cap P_j^\beta| = |\{k \in [w] : \beta_k - \alpha_k = (j - i) \mod v\}| \leq w - d_H(\alpha, \beta)$. From the properties of $\mathcal{RS}$ as in Lemma 3.7, for $\alpha \neq \beta \in \mathcal{RS}$, $d_H(\alpha, \beta) \geq w - \ell$. Therefore, for $\alpha \neq \beta \in \mathcal{RS}$, $|P_i^\alpha \cap P_j^\beta| \leq w - d_H(\alpha, \beta) \leq \ell$. The claim now follows from the definition of design. $\square$

**Lemma 7.2.** *For $1 \leq \ell \leq w/2$, $\mathcal{RS}$ is $(q, \tau)$-load balancing for $\tau = 2^\ell + 2^w(q/v)^{\ell-1}$.*

*Proof.* The argument here is similar to the proofs of Chernoff bounds for random variables with limited independence (especially those typically used in analyzing limited independence hash functions).

Let $Q \subseteq [n]$ with $|Q| = q \leq v$ and fix an index $j \in [v]$. Let $\alpha \in_u \mathcal{RS}$ and let $X = |Q \cap P_j^\alpha|$; note that $\mathbb{E}[X] = q/v$. We are interested in estimating $\mathbb{E}[1(X > 0)2^X]$. We do so by first proving a tail bound on $X$. To this end, for $1 \leq i \leq w$, let

$$X_i = \begin{cases} 1 & \text{if } j - \alpha_i \in Q \cap \{(i-1)v + 1, (i-1)v + 2, \ldots, (iv)\} \\ 0 & \text{otherwise} \end{cases}.$$

Then, from the definition of the parition $P_j^\alpha$, $X = X_1 + X_2 + \cdots + X_w$. Further, as each $X_i$ only depends on $\alpha_i$, $X_1, \ldots, X_w$ are $\ell$-wise independent. Therefore, by a standard calculation,

$$\Pr[X \geq \ell] \leq \mathbb{E}[S_\ell(X_1, \ldots, X_w)] = \sum_{I \subseteq [w], |I| = \ell} \mathbb{E}\left[\prod_{i \in I} X_i\right]$$

$$= \sum_{I \subseteq [w], |I| = \ell} \prod_{i \in I} \mathbb{E}[X_i] \leq \binom{w}{\ell}\left(\frac{\sum_{i=1}^w \mathbb{E}[X_i]}{w}\right)^\ell.$$

where the last inequality follows from Fact 3.14. Now, as $\mathbb{E}[X] = q/v$, the above expression simplifies to

$$\Pr[X > \ell] \leq \binom{w}{\ell} \cdot (q/wv)^\ell \leq \left(\frac{eq}{v\ell}\right)^\ell.$$

Thus, for $\ell \geq 3$,

$$
\begin{aligned}
\mathbb{E}\left[1(Q \cap P_j^\alpha \neq \emptyset)2^{|Q \cap P_j^\alpha|}\right] &= \mathbb{E}\left[1(X > 0)2^X\right] \\
&= \Pr[X < \ell]\,\mathbb{E}\left[1(X > 0)2^X | X < \ell\right] + \Pr[X \geq \ell]\,\mathbb{E}\left[1(X > 0)2^X | X \geq \ell\right] \\
&\leq 2^\ell \Pr[X > 0] + 2^w \Pr[X > \ell] \\
&\leq 2^\ell(q/v) + 2^w(q/v)^\ell.
\end{aligned}
$$

The claim now follows from the definition of load balancing. $\qquad\square$

We next use the above claims along with Theorems 2.3 and 2.5 for a suitable setting of parameters.

**Lemma 7.3.** *For all $0 < \delta < 1$, there exists a constant $c_\delta \geq 1$ and a suitable choice of $v = \Theta_\delta(2^w w)$ such that the following holds. For $\ell \geq 1/\delta$, the function $f_{\mathcal{RS}}$ as defined above is $c_\delta 2^{-w}$-strongly resilient and $\Pr_{x \in_u \{0,1\}^n}[f_{\mathcal{RS}}(x) = 1] = 1/2 \pm 2^{-\Omega(w)}$.*

*Proof.* Let $\ell = \max(3, \lceil 1/\delta \rceil)$. For $v$ to be chosen in a little bit, let $u = v^\ell$ and $f \equiv f_{\mathcal{RS}}$. We would like our choice of $v$ to minimize $|v - 2^w \ln((\ln 2)/v^\ell)|$ so that we can get an almost-balanced function using Fact 3.16. To this end, let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be defined by

$$
\phi(x) = x - 2^w \ln((\ln 2)x^\ell)
$$

and let $x^* \geq 1$ be such that $\phi(x^*) = 0$. There exists such an $x^*$ by the continuity of $\phi$. It is also easy to check that for $w$ sufficiently large, $\phi'(y) \geq 0$ for all $y \geq x^*$. We set $v$ to be the smallest prime larger than $x^*$.[7] Note that $v \leq x^* + B$ where $B = 2^{c_1 w}$ for some universal constant $c_1 < 1$ (see [Wik] for instance), so that $0 = \phi(x^*) \leq \phi(v) \leq \phi(x^*) + B$. Let $\theta = (1 - 2^{-w})^v$. Then, by Fact 3.16, $(1 + \theta)^u = O(1)$ and

$$
\mathsf{bias}(u, v, w) = (1 - \theta)^u = 1/2 \pm O(1)2^{-\Omega(w)}.
$$

Now, by Lemma 7.1 and Theorem 2.3,

$$
\Pr_{x \leftarrow \mathcal{D}_{1/2}}[f_{\mathcal{RS}}(x) = 1] = \mathsf{bias}(u, v, w) \pm \exp(-\Omega(w - \ell)) = \mathsf{bias}(u, v, w) \pm \exp(-\Omega(w)).
$$

Next, by Lemma 7.2 and Theorem 2.5, for any $q \leq 2^{w(1-\delta)}$, as $\ell \geq \lceil 1/\delta \rceil$,

$$
\begin{aligned}
I_q(f_{\mathcal{RS}}) &\leq u(1 - 2^{-w})^{v-q} \cdot (2^{-w}q) \cdot \left(2^\ell + 2^w(q/v)^{\ell-1}\right) \\
&\leq O(1) \cdot (2^{-w}q) \cdot (2^\ell + 2^w 2^{-\ell\delta w}) \\
&= O_\delta(2^{-w}q).
\end{aligned}
$$

Therefore, $f_{\mathcal{RS}}$ is $O_\delta(2^{-w})$-strongly resilient. $\qquad\square$

**Corollary 7.4.** *For all $0 < \delta < 1$, there exists a constant $c_\delta \geq 1$ such that the following holds. There exists an explicit depth-three monotone function $f : \{0,1\}^n \to \{0,1\}$ which can be computed in time $n^{c_\delta}$ such that for $t \geq c_\delta(\log n)^4$*

- *$f$ is almost balanced: for any $t$-wise independent distribution $\mathcal{D}$ on $\{0,1\}^n$, $\Pr_{x \leftarrow \mathcal{D}}[f(x) = 1] = 1/2 \pm n^{-\Omega(1)}$.*
- *$f$ is $t$-wise $c_\delta n^{-(1-\delta)}$-strongly resilient.*

---

[7]We can find such a prime in time $2^{O(w)}$ which is fine for us.

*Proof.* We instantiate the previous lemma to get $f \equiv f_{\mathcal{RS}}$ for $\delta' = \delta/2$. Then, $n = O(2^w w^2)$ so that $2^{w(1-\delta')} = \Omega(n^{1-\delta})$. To analyze the bias under $t$-wise independent distributions, we apply Corollary 4.2 with $d = w/2$, $\varepsilon = 1/n$ instead of Theorem 2.3 in the above argument. Similarly, to analyze the influence under $t$-wise independent distributions we use Corollary 5.1 with $\varepsilon' = 1/(n^3)$ instead of Theorem 2.5. Then, the amount of independence needed is $O(w^2 d^2 + wd \log(1/\varepsilon)) = O(\log^4 n)$. $\qquad\square$

## 7.2  Proof of Theorem 1.2

We now prove Theorem 1.2. The approach is similar to the above with one crucial difference: we use the output of the generator from Theorem 1.8 instead of the Reed-Solomon code. We ensure the requisite design properties to apply Theorem 2.3, at a high-level, by *padding* the output of the generator with a Reed-Solomon code.

Let $c$ be a sufficiently big constant to be chosen later and suppose that $v, D$ are prime numbers below. Let $E : [v^c] \times [D] \to [v/D]$ be a strong extractor with error $\varepsilon = \text{poly}(1/w)$ and $D = \text{poly}(w)$ to be chosen later. Let $\ell = \Theta(w/(\log w))$ be a parameter to be chosen later. Let $G_c : [v]^c \to [v]^{2c}$, $G_\ell : [D]^\ell \to [D]^{w-2c}$ generate a $c$-wise independent distribution over $[v]$ and a $\ell$-wise independent distribution over $[D]$ respectively as guaranteed by Lemma 3.7.

Now, define $\mathcal{U} : [v^c] \times [D]^\ell \to [v]^w$ as follows:

$$\mathcal{U}(x,y)_i = \begin{cases} G_c(x)_i & \text{if } 1 \le i \le 2c \\ G_\ell(y)_{i-2c} \circ E(x, G_\ell(y)_{i-2c}) & \text{if } 2c < i \le w \end{cases}. \tag{7.1}$$

(Here, we associate an element of $[D] \times [v/D]$ with an element of $v$ in a straightforward bijective manner.)

Abusing notation, we let $\mathcal{U} = \{\mathcal{U}(x,y) : x \in [v^c], y \in [D]^\ell\} \subseteq [v]^w$ as well. Our final function will be $f \equiv f_{\mathcal{P}}$ for $\mathcal{P} = \{P^\alpha : \alpha \in \mathcal{U}\}$. The following claims help us apply Theorem 2.3 and Theorem 2.5 to analyze $f_{\mathcal{P}}$.

**Lemma 7.5.** *For all $c < (w - \ell)/3$, $\mathcal{P}$ is a $(c, w - 2c - \ell, 1/D^\ell)$-design.*

*Proof.* Note that for any $\alpha, \beta \in [v]^w$, and any $i, j \in [v]$, $|P_i^\alpha \cap P_j^\beta| = |\{k \in [w] : \beta_k - \alpha_k = (j-i) \bmod v\}| \le w - d_H(\alpha, \beta)$.

Let $\alpha = G(x, y) \neq \beta = G(x', y') \in \mathcal{U}$. We consider two cases depending on whether $y \neq y'$. If $y \neq y'$, then $d_H(\alpha, \beta) \ge d_H(G_\ell(y), G_\ell(y')) \ge w - 2c - \ell$; similarly, if $x \neq x'$, then $d_H(\alpha, \beta) \ge d_H(G_c(x), G_c(x')) \ge c$. Therefore, $d_H(\alpha, \beta) \ge \min(c, w - 2c - \ell) = c$. Hence, $\mathcal{P}$ is a $c$-design.

Now, consider any fixed $\alpha = G(x, y) \in \mathcal{U}$, $i, j \in [v]$, and $\beta = G(x', y') \in_u \mathcal{U}$. Then, by the above argument, if $y \neq y'$, $d_H(\alpha, \beta) \ge w - 2c - \ell$ so that $|P_i^\alpha \cap P_j^\beta| \le 2c + \ell$. On the other hand, $\Pr[y' \neq y] \ge 1 - 1/D^\ell$. Thus,

$$\Pr_{\beta \in_u \mathcal{U}}[|P_i^\alpha \cap P_j^\beta| \ge w - (w - 2c - \ell)] \le \Pr[y' = y] \le 1/D^\ell.$$

Therefore, $\mathcal{P}$ is a $(c, w - 2c - \ell, 1/D^\ell)$-design as needed. $\qquad\square$

We next use Lemma 6.2 to analyze the load-balancing properties of $\mathcal{P}$.

**Lemma 7.6.** *Let $\mathcal{U}$ be as in Equation 7.1 for $E$ being a $((c \log v)/2, \varepsilon)$-extractor for $\varepsilon < 1/w$ and $\ell \ge 6w/(\log w)$. Then, $\mathcal{P} = \{P^\alpha : \alpha \in \mathcal{U}\}$ is $(q, \tau)$-load balancing for $q \le v$ and $\tau = O(2^{2c} + (vw2^w)v^{-c/2})$.*

*Proof.* Let $Q \subseteq [n]$ with $|Q| = q \leq v$ and fix an index $j \in [v]$. Let $\alpha \in_u \mathcal{U}$ and let $X = |Q \cap P_j^\alpha|$; note that $\mathbb{E}[X] = q/v$. For $1 \leq i \leq w$, let

$$X_i = \begin{cases} 1 & \text{if } j - \alpha_i \in Q \cap \{(i-1)v + 1, (i-1)v + 2, \dots, (iv)\} \\ 0 & \text{otherwise} \end{cases}.$$

Then, from the definition of the parition $P_j^\alpha$, $X = X_1 + X_2 + \cdots + X_w$.

For brevity, let $\mathbb{E}[X] = \mu = q/v \leq 1$, $Y = \sum_{i=1}^{2c} X_i$, and $Z = \sum_{i=2c+1}^{w} X_i$. Note that $X = Y + Z$. We are interested in estimating $\mathbb{E}[1(X > 0)2^X]$. We do so mainly by applying Lemma 6.2 to the random variable $Z$ combined with the trivial observation that $Y$ is at most $2c$:

$$\begin{aligned}
\mathop{\mathbb{E}}_{\alpha \in_u U} \left[ 1(Q \cap P_k^\alpha \neq \emptyset)2^{|Q \cap P_k^\alpha|} \right] &= \mathbb{E}\left[1(X > 0)2^X\right] = \mathbb{E}\left[2^X\right] - \Pr[X = 0] \\
&= \mathbb{E}\left[2^X\right] - 1 + \Pr[X \geq 1] \\
&\leq \mathbb{E}\left[2^X\right] - 1 + \mathbb{E}[X] \\
&= \mathbb{E}[2^Y - 1] + \mathbb{E}[2^Y(2^Z - 1)] + \mu \\
&\leq \mathbb{E}[2^Y - 1] + 2^{2c} \cdot \mathbb{E}[2^Z - 1] + \mu \\
&\leq \mathbb{E}[2^{2c}Y] + 2^{2c} \cdot \mathbb{E}[2^Z - 1] + \mu.
\end{aligned}$$

By Lemma 6.2 applied to $Z$ with $\theta = \ln 2$, $\ell = 12w/(\log w)$, and $\varepsilon < 1/2w$, we get that

$$\begin{aligned}
\mathbb{E}[2^Z] &= O(1)\exp(O(\mu)) + 2^w \left(\frac{2e}{\ell}\right)^\ell + O(w2^w)v^{-c/2} \\
&\leq 1 + O(\mu) + 2^{-w} + O(w2^v)v^{-c/2}.
\end{aligned}$$

Further, $\mathbb{E}[Y] \leq \mathbb{E}[X] = \mu$. Therefore, as $\mu \geq 1/v$,

$$\mathbb{E}\left[1(Q \cap P_k^\alpha \neq \emptyset)2^{|Q \cap P_k^\alpha|}\right] \leq O(2^{2c})\mu + O(1)w2^w v^{-c/2} \leq O(1) \cdot \mu \cdot (2^{2c} + w2^w v^{-c/2+1}).$$

$\square$

*Proof of Theorem 1.2.* The theorem follows essentially by combining the above two lemmas and Theorem 2.3, Theorem 2.5. We first set up some parameters. Let $c \geq 2$ be a sufficiently large constant to be chosen later. Let $w \geq 1$ be arbitrary and $\varepsilon = 1/w^3$. Let $D = ((c\log v)w)^C$ for some universal constant $C$ so that there exists an explicit $(c(\log v)/2, \varepsilon)$-strong extractor $E : [v^c] \times [D] \to [v/D]$ for all $c \geq C$ as in Theorem 3.5. Set $v = \Theta(2^w w)$ to be chosen precisely in a little bit. For this setting of $v$, let $\mathcal{U} \subseteq [v]^w$ be as defined in Equation 7.1 with $\ell = 12w/(\log w)$ and $E$ as the extractor. Let $\mathcal{P} = \mathcal{P}_{\mathcal{U}}$. Then, $|\mathcal{U}| := u = v^c \times D^\ell$. We will show that $f \equiv f_{\mathcal{P}}$ satisfies the conditions of Theorem 1.2 for $c$ sufficiently large.

As in the proof of Lemma 7.3, we would like $v$ to be as close as possible to $2^w \ln((\ln 2)u)$. To this end, let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be defined by

$$\phi(x) = x - 2^w \left(c \ln x + C\ell \ln(\log x) + C\ell \ln(cw) + \ln \ln 2\right)$$

and let $x^* \geq 1$ be such that $\phi(x^*) = 0$. There exists such an $x^*$ by the continuity of $\phi$. Let $v$ be the smallest prime larger than $x^*$. Note that $v \leq x^* + B$ where $B = 2^{c_1 w}$ for some universal constant $c_1 < 1$ (see [Wik] for instance), so that $0 = \phi(x^*) \leq \phi(v) \leq \phi(x^*) + B$. Let $\theta = (1 - 2^{-w})^v$. Then, by Fact 3.16, $(1 + \theta)^u = O(1)$ and

$$\mathsf{bias}(u, v, w) = (1 - \theta)^u = 1/2 \pm O(1)2^{-\Omega(w)}. \tag{7.2}$$

**Analyzing bias:** By Lemma 7.5, $\mathcal{P}$ is a $(c, w - 2c - \ell, 1/D^\ell)$-deisgn. Therefore, by Theorem 2.3,

$$\Pr_{x \in_u \{0,1\}^n}[f_\mathcal{P}(x) = 1] = \mathsf{bias}(u, v, w) + w \exp(-\Omega(w)) + \exp(-\Omega(c)) + 2^w/D^\ell$$

$$= \mathsf{bias}(u, v, w) + \exp(-\Omega(c)),$$

as $D^\ell \gg 2^w$ for $C$ a sufficiently large constant.

**Analyzing influence:** We claim that $f_\mathcal{P}$ has small influence for coalitions of size $o(2^w)$. Let $1 \le q \le v$ so that $q/v \le 1$. Then, by Lemma 7.6, $\mathcal{P}$ is $(q, \tau)$-load balancing for

$$\tau = O(2^{2c})(1 + v^{-c/2+1}) = O(2^{2c}).$$

Therefore, by Theorem 2.5, for all $q \le v$,

$$I_q(f_\mathcal{P}) \le u(1 - 2^{-w})^{v-q} \cdot 2^{2c}(2^{-w}q) = O(2^{2c})2^{-w}q = O(2^{2c}) \cdot \left((\log^2 n)/n\right) \cdot q.$$

Here, the last inequality follows as $2^w = \Theta(n/(\log^2 n))$.

The theorem now follows by choosing $c$ to be a sufficiently large constant. □

We next prove Theorem 1.7.

*Proof of Theorem 1.7.* The proof is exactly the same as the above argument for Theorem 1.2 but instead of using Theorem 2.3 we use Corollary 4.2 with $d = c$, $\varepsilon = 1/u$ and instead of Theorem 2.5 we use Corollary 5.1 with $\varepsilon = 1/(uv^2)$. The amount of independence we need is $t \gg O(w^2 d^2 + wd \log(1/\varepsilon)) = O(\log^2 n)$ as required for the theorem. Note that the *error* $1/9$ can be made to be an arbitrary small constant. □

# 8 Better two-source extractors

Our improved quantitative bounds for two-source extractors follow immediately by using our resilient functions, Theorem 1.7 and Corollary 7.4 in the reduction of [CZ15]. Concretely, [CZ15] show the following for some universal constants $0 < c < 1$ and $C, C' \ge 1$. Suppose for some functions $\varepsilon : \mathbb{Z}_+ \to [0,1]$, and $t : \mathbb{Z}_+ \to \mathbb{Z}_+$, the following holds: for all $m \ge 1$, there exists an explicit $t(m)$-wise $(m^{1-c}, \varepsilon)$-resilient function $f : \{0,1\}^m \to \{0,1\}$ with $\Pr_{x \in_u \{0,1\}^m}[f(x) = 1] = 1/2 \pm \varepsilon(m)$. Then, there exists an explicit $(n, k)$ two-source extractor with error at most $\varepsilon(n^C)$ and $k \ge C' \cdot (t(n^C))^4 \log^2 n$. Instantiating this reduction with the resilient function from Theorem 1.7 (with $1/9$ replaced with a sufficiently small constant) gives Theorem 1.4; similarly, using Corollary 7.4 gives Theorem 1.5.

# References

[AL93]   Miklós Ajtai and Nathan Linial. The influence of large coalitions. *Combinatorica*, 13(2):129–145, 1993.

[AS11]   N. Alon and J.H. Spencer. *The Probabilistic Method.* Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, 2011.

[BL85]   Michael Ben-Or and Nathan Linial. Collective coin flipping, robust voting schemes and minima of banzhaf values. In *26th Annual Symposium on Foundations of Computer Science, Portland, Oregon, USA, 21-23 October 1985*, pages 408–416, 1985.

[Bou05]    J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 01(01):1–32, 2005.

[Bra10]    Mark Braverman. Polylogarithmic independence fools $AC^0$ circuits. *J. ACM*, 57(5), 2010.

[BRSW12]   Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2-source dispersers for $n^{o(1)}$ entropy, and Ramsey graphs beating the Frankl-Wilson construction. *Ann. Math. (2)*, 176(3):1483–1544, 2012.

[Coh15]    Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved ramsey graphs. *CoRR*, abs/1506.04428, 2015. To appear in STOC 2015.

[CZ15]     Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:119, 2015. To appear in STOC 2015.

[DETT10]   Anindya De, Omid Etesami, Luca Trevisan, and Madhur Tulsiani. Improved pseudorandom generators for depth 2 circuits. In *APPROX-RANDOM*, pages 504–517, 2010.

[Dod06]    Yevgeniy Dodis. Fault-tolerant leader election and collective coin-flipping in the full information model, 2006.

[Gil98]    David Gillman. A chernoff bound for random walks on expander graphs. *SIAM J. Comput.*, 27(4):1203–1220, 1998.

[GKM15]    Parikshit Gopalan, Daniek Kane, and Raghu Meka. Pseudorandomness via the discrete fourier transform. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 903–922, 2015.

[Hea08]    Alexander Healy. Randomness-efficient sampling within nc$^1$. *Computational Complexity*, 17(1):3–37, 2008.

[INW94]    Russell Impagliazzo, Noam Nisan, and Avi Wigderson. Pseudorandomness for network algorithms. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montréal, Québec, Canada*, pages 356–364, 1994.

[Kah97]    Nabil Kahale. Large deviation bounds for markov chains. *Combinatorics, Probability & Computing*, 6(4):465–474, 1997.

[KKL88]    Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions (extended abstract). In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 68–80, 1988.

[Li15]     Xin Li. Three-source extractors for polylogarithmic min-entropy. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, 2015.

[LMN93]    Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, 1993.

[Mek09]    Raghu Meka. Explicit coin flipping protocols, 2009. Unpublished manuscript.

[Nis92]    Noam Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992.

[NZ96]     Noam Nisan and David Zuckerman. Randomness is linear in space. *J. Comput. Syst. Sci.*, 52(1):43–52, 1996.

[Raz87]    A. A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition, 1987.

[Smo87]   Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 77–82, 1987.

[Wik]     Wikipedia. Prime gap — Wikipedia, the free encyclopedia.

[Zuc97]   David Zuckerman. Randomness-optimal oblivious sampling. *Random Struct. Algorithms*, 11(4):345–367, 1997.

# A  Missing proofs

*Proof of Lemma 3.13.* For a set $T \subseteq [v]$, let $n(T) = |\{(I_1, \ldots, I_r) : |I_j| = k, \forall j \in [r]$ and $\cup_{j=1}^r I_j = T\}|$. Then,

$$S_k(X_1, \ldots, X_v)^r = \sum_{I_1, \ldots, I_r \in \binom{[v]}{k}} \prod_{i \in \cup_{j=1}^r I_j} X_i = \sum_{T \subseteq [v]} n(T) \prod_{i \in T} X_i.$$

Thus,

$$\mathbb{E}[S_k(X_1, \ldots, X_v)^r] = \sum_{T \subseteq [v]} n(T) p^{|T|}.$$

Finally, observe that for any $T$ with $|T| = t$, $n(T) \leq \binom{t}{k}^r \leq (et/k)^r$. Thus,

$$\mathbb{E}[S_k(X_1, \ldots, X_v)^r] \leq \sum_{t=k}^{kr} \binom{v}{t} \cdot p^t \left(\frac{et}{k}\right)^r$$

$$\leq \sum_{t=k}^{kr} \left(\frac{evp}{t}\right)^t \cdot \left(\frac{et}{k}\right)^r \leq \sum_{t=k}^{kr} \left(\frac{1}{2}\right)^r \cdot \left(\frac{evp}{t}\right)^{t-r} \leq (1/2)^k,$$

as $evp/t \leq 1/4$, and $e^2 vp/k \leq 1/2$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

*Proof of Fact 3.16.* Note that under the assumptions we must have $v = \Omega(2^w w)$. Then, by Fact 3.15,

$$(1 - 2^{-w})^{B + 2^w \ln(u/(\ln 2))} \leq \theta \leq (1 - 2^{-w})^{2^w \ln(u/(\ln 2))} \leq (\ln 2)/u.$$

Therefore, $(1 + \theta)^u = O(1)$ and

$$(1 - \theta)^u \geq (1 - (\ln 2)/u)^u \geq 1/2(1 - (\ln 2)/u)^{\ln 2} \geq 1/2(1 - O(1)/u).$$

Further,

$$\theta \geq ((\ln 2)/u) \cdot (1 - 2^{-w})^{B + \ln((\ln 2)u)} \geq ((\ln 2)/u) \cdot (1 - O(B + \ln u)2^{-w}).$$

Thus,

$$(1 - \theta)^u \leq \exp(-u\theta) \leq (1/2)^{\left(1 - O(B \ln u)2^{-w}\right)} \leq 1/2 \pm O(B \ln u)2^{-w}.$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$$

*Proof of Lemma 6.1.* The proof relies on the following elementary inequality about symmetric polynomials: for all $0 \leq a_1, \ldots, a_w \leq A$,

$$\prod_i (1 + a_i) \leq \sum_{i=1}^{\ell-1} S_i(a_1, \ldots, a_w) + (1 + A)^w \cdot S_\ell(a_1, \ldots, a_w).$$

Let $Z_i = \exp(\theta Y_i) - 1$. Let $p(Z) = \sum_{i=1}^{\ell-1} S_i(Z_1, \ldots, Z_w)$. Applying the above inequality to $Z = (Z_1, \ldots, Z_w)$ we get

$$\mathbb{E}\left[\exp(\theta(Y_1 + \cdots + Y_w))\right] = \mathbb{E}\left[\prod_i (1 + Z_i)\right] \leq \mathbb{E}[p(Z)] + \exp(\theta w) \cdot \mathbb{E}[S_\ell(Z)] \leq$$

$$\mathbb{E}[p(Z)] + \exp(\theta w) \cdot \binom{w}{\ell} \cdot \left(\frac{\sum_i \mathbb{E}[Z_i]}{w}\right)^\ell,$$

where the last inequality follows because $Z$'s are $\ell$-wise independent so that by Fact 3.14

$$\mathbb{E}[S_\ell(Z)] = \sum_{I \subseteq [w], |I| = \ell} \prod_{i \in I} \mathbb{E}[Z_i] \leq \binom{w}{\ell} \cdot \left(\frac{\sum_i \mathbb{E}[Z_i]}{w}\right)^\ell.$$

Now, for any $i \in [w]$,

$$\mathbb{E}[Z_i] = \sum_{j=1}^\infty \frac{\theta^j \, \mathbb{E}[Y_i^j]}{j!} \leq \sum_{j=1}^\infty \frac{\theta^j \, \mathbb{E}[Y_i]}{j!} = \mathbb{E}[Y_i](e^\theta - 1).$$

Therefore, $\sum_i \mathbb{E}[Z_i] \leq e^\theta(\sum_i \mathbb{E}[Y_i]) = e^\theta \cdot \mu$. Finally, note that as $\{Z_1, \ldots, Z_w\}$ are $\ell$-wise independent, the expectation of $p(Z)$ would be the same as when the $Z$'s were independent. However, in this case,

$$\mathbb{E}[p(Z)] \leq \mathbb{E}\left[\prod_i (1 + Z_i)\right] = \prod_i \mathbb{E}[(1 + Z_i)].$$

Combining the above equations, we get

$$\mathbb{E}\left[\exp(\theta(Y_1 + \cdots + Y_w))\right] \leq \prod_i \mathbb{E}[\exp(\theta Y_i)] + \exp(\theta w) \cdot \binom{w}{\ell} \cdot (e^\theta \mu/w)^\ell$$

$$\leq \prod_i \mathbb{E}[\exp(\theta Y_i)] + \frac{\exp(2\theta w)(e\mu)^\ell}{\ell^\ell}.$$

This finishes the claim.

$\square$