Electronic Colloquium on Computational Complexity, Report No. 125 (2014)

Beyond the Central Limit Theorem: asymptotic expansions and pseudorandomness for combinatorial sums

Anindya De[*]

DIMACS, Rutgers and Institute for Advanced Study

October 7, 2014

Abstract

In this paper, we construct pseudorandom generators for the class of *combinatorial sums*, a class of functions first studied by [GMRZ13] and defined as follows: A function $f : [m]^n \to \{0,1\}$ is said to be a combinatorial sum if there exists functions $f_1, \ldots, f_n : [m] \to \{0,1\}$ such that $f(x_1, \ldots, x_n) = f_1(x_1) + \ldots + f_n(x_n)$. Derandomization of combinatorial sums generalize previously studied classes such as combinatorial rectangles [EGL$^+$92], small-biased spaces [NN93] and modular sums [LRTV09] among others.

In this work, we present a pseudorandom generator for combinatorial sums with seed length $O(\log m + \log^{3/2}(n/\epsilon))$, thus improving upon [GMRZ13] whenever $\epsilon \leq 2^{-(\log n)^{3/4}}$. As a corollary, this gives the first improvement over the INW generator [INW94] for *fooling* the simple majority function in the case of inverse polynomial error.

The main technical ingredient in our result is the use of *asymptotic expansions* which roughly speaking, are refinements of the classical central limit theorem which achieve a faster convergence rate than the central limit theorem by using more moments for the approximation. While asymptotic version of such theorems have been known in the probability theory literature for some time, explicit bounds were not known for sum of lattice-valued i.i.d. random variables. In the main technical ingredient of this paper, we prove a new asymptotic expansion theorem for sums of lattice-valued independent random variables (even for the non i.i.d. case). Given the far-reaching consequences of the central limit theorem in theoretical computer science, we hope that the new *asymptotic expansions* will be of independent interest.

[*] anindya@math.ias.edu. Most of the work was done while the author was a member at the School of Mathematics, Institute for Advanced Study and during a visit to Columbia University.

ISSN 1433-8092

1 Introduction

Derandomization of randomized logspace (or RL) is one of the holy grails of computational complexity theory. Unlike BPP, where its known that even mild derandomization yields non-trivial circuit lower bounds [IKW02, KI04] and thus likely to be out of reach of current complexity theory, no such *barrier* is known for derandomization of RL. In fact, from the classical result of Savitch [Sav70], it follows that $\mathsf{RL} \subseteq \mathsf{L}^2$ which was improved upon by Saks and Zhou [SZ95] to show that $\mathsf{RL} \subseteq \mathsf{L}^{3/2}$. Further, following the pathbreaking result of Reingold [Rei08] (which showed $\mathsf{SL} = \mathsf{L}$), there were follow-up works [RTV06, CRV11] which suggested that we may not be far from proving $\mathsf{RL} = \mathsf{L}$.

However, in spite of the optimism, the best known derandomization for RL remains the result of [SZ95]. In fact, the state of oblivious derandomization remains significantly worse. To discuss this, we recall the definition of read-once branching programs (ROBPs) and pseudorandom generators (PRGs).

Definition 1. *A (S, D, T) read-once branching program (ROBP) M is a layered directed multi-graph with $T + 1$ layers and at most 2^S vertices in each layer. For $0 < i < T$, a vertex v in layer i of M has at most 2^D outgoing edges labeled with distinct elements of $\{0, 1\}^D$ all leading to a vertex in layer $i+1$.*

We can associate a function $g_M : (\{0, 1\}^D)^T \to [2^S]$ with M. Namely, we label each vertex in the T^{th} layer with a number in $[2^S]$ (not necessarily distinct). On input $x = (x_1, \ldots, x_T)$ where $x_i \in \{0, 1\}^D$ for all $i \in [T]$, $g_M(x)$ is the label of the vertex obtained by starting at v_0 and walking along the edges $x_1, \ldots, x_T$ (in order).

For the reader familiar with branching programs, we remark that typically the vertices of the last layer are labeled with an element in $\{0, 1\}$, thus the resulting $g_M(\cdot)$ becomes a Boolean function. In this paper, we go for a more general definition as it is more convenient for us.

It is not difficult to see that derandomization of (S, D, T) ROBP for $S = O(\log n)$, $T = n$ and $D = 1$ suffices for derandomization of RL. Next, we define the notion of pseudorandomness for a complexity class $\mathcal{C}$. Towards this, let us adopt the notation that U_m denotes the uniform distribution on m-bit strings. In general, for a domain X, we use U_X to denote the uniform distribution on the elements of X. Also, for distributions D_1 and D_2 on the same domain, $\|D_1 - D_2\|_1$ denotes the ℓ_1 distance between D_1 and D_2.

Definition 2. *Let $\mathcal{C} \subseteq \{h : X \to X'\}$. Then, $g : \{0, 1\}^t \to X$ is said to be a ϵ-PRG for $\mathcal{C}$ if for every $h \in \mathcal{C}$,*

$$\|h(g(U_t)) - h(U_X)\|_1 \leq \epsilon.$$

Note that $h(g(U_t))$ (and likewise, $h(U_X)$) defines a random variable on the domain X'. The seed length of g is t and in this paper, whenever we mention PRGs, we mean they are explicit PRGs i.e. g is computable efficiently.

The currently best known PRG for ROBP is due to Impagliazzo, Nisan and Wigderson [INW94] which builds upon the seminal work of Nisan [Nis92].

Theorem 1. *There is an explicit PRG $G_{\mathsf{INW}} : \{0, 1\}^t \to (\{0, 1\}^D)^n$ which is ϵ-pseudorandom for (S, D, T)-ROBP where $O(D + (S + \log(T/\epsilon)) \cdot \log(T))$.*

For the sake of completeness, we mention that Nisan and Zuckerman [NZ96] obtained a ϵ-PRG with seed length $O(S + D)$ if $T = \mathsf{poly}(S, D)$ and $\epsilon \geq 2^{\log^{1-\gamma}(S+D)}$ for $\gamma > 0$. As can be easily seen, this beats G_{INW} only when T is small compared to S and D.

The problem of beating [INW94] for general ROBPs has so far resisted all attacks and in fact, even for $S = 3$, $D = 1$ and $\epsilon < 1/2$, the best known PRG has seed length $O(\log^2 T)$. To get around this, researchers have looked at restricted class of ROBPs: One line of research has dealt with structural restrictions on ROBPs (cf.[BV10, BRRY10, KNP11, De11, Ste12, RSV13, SVW14]). The second line of research, which is also the focus of this paper (and predates the first line of research) has dealt with ROBPs computing *semantically restricted* classes of functions. These include small biased spaces [NN93], combinatorial rectangles [EGL$^+$92, ASWZ96, Lu02, GMR$^+$12], combinatorial checkerboards [Wat13] and modular sums [LRTV09] among others.

A common generalization which includes all these classes is the class of combinatorial sums which was introduced in the work of Gopalan *et al.* [GMRZ13]. They are defined as follows.

Definition 3. *The class of combinatorial sums (denoted by* $\mathsf{Csum}(m,n)$*) consists of* $f : [m]^n \to \mathbb{Z}$ *of the form*

$$f(x_1,\ldots,x_n) = f_1(x_1) + \ldots + f_n(x_n) \textsf{ where for all } i \in [n],\ f_i : [m] \to \{0,1\}.$$

Sometimes we will use the tuple $(f_1,\ldots,f_n)$ *to refer to the combinatorial sum* f*.*

The main result of [GMRZ13] is an ϵ-PRG with seed length $O(\log m + \log n + \log^2(1/\epsilon))$. The main result of this paper is the following theorem.

Theorem 2. *There is a polynomial time computable PRG* $G_{csum} : \{0,1\}^{t_{\mathsf{csum}}} \to [m]^n$ *which* ϵ*-fools* $\mathsf{Csum}(m,n)$*. Here* $t_{\mathsf{csum}} = O(\log m + \log^{3/2}(n/\epsilon))$*.*

When $\epsilon = n^{-\Theta(1)}$, this provides the first improvement over [INW94] for the class $\mathsf{Csum}(m,n)$. Previously, even for $m = 2$, the best known PRG had a seed length of $O(\log^2 n)$ for $\epsilon = n^{-\Theta(1)}$. To understand the result, we begin with a high level description of the result of [GMRZ13].
Let $U_{[m]^n}$ be the uniform distribution on $[m]^n$. Then, note that for $(y_1,\ldots,y_n) \sim U_{[m]^n}$, each $f_i(y_i)$ is an independent $\{0,1\}$ random variable. Let us define $Z = \sum f_i(y_i)$ and $\mathrm{Var}(Z) = \sigma^2$. The key technical component in their result is the use of the so-called *discrete central limit theorem.* Much like the basic central limit theorem, there are several versions of the discrete central limit theorem (discrete CLT) as well (see [CGS10] for a sampling of such results). However, as a corollary, we have the following (which is the version relevant to us).

Theorem 3. ***Discrete central limit theorem:*** *For* Z *as defined above, let* Z' *be the discretized normal with mean* $\mathbf{E}[Z]$ *and variance* $\mathrm{Var}(Z)$*. Then,* $\|Z - Z'\|_1 = O(1/\sigma)$*.*

We now explain the key idea behind using Theorem 3 for derandomization of combinatorial sums is as follows. The derandomization problem is split into two cases:

- low variance case: $\mathrm{Var}(Z) \le \Theta(1/\epsilon^2)$ and
- high variance case: $\mathrm{Var}(Z) \ge \Theta(1/\epsilon^2)$.

We now describe the PRG for these cases. In case (i), we set $t = \mathsf{poly}(1/\epsilon)$, so that $\mathrm{Var}(Z)/t \le \epsilon$. Let $\mathcal{H}_{2,n,t}$ be a family of pairwise of pairwise independent hash functions. The PRG first samples $h \sim \mathcal{H}_{2,n,t}$ to partition $[n]$ into t *buckets.* Within each bucket, the PRG uses $O(1)$-wise independence (the distribution across the buckets is independent). The authors use the notion of sandwiching polynomials (cf. [Baz07]) to show that this PRG construction indeed fools f (in case (i)) up to an error ϵ. However, as the seed across the buckets are independent, hence the seed required grows as $t\cdot(\log m+\log n)$. This has a poor dependence on ϵ. To get around this, the authors observe that the computation of f composed with the PRG can be seen as ROBP (with significantly smaller size than n). Thus, they derandomize this using Theorem 1 to bring down the seed length requirement to $O(\log m + \log n + \log^2(1/\epsilon))$.
For case (ii), note that by Theorem 3, Z is ϵ-close to an appropriate discretized normal in ℓ_1 distance. Thus, it suffices to produce a pseudorandom distribution $(y'_1,\ldots,y'_n)$ so that the induced distribution $Z' = \sum f_i(y'_i)$ is $O(\epsilon)$-close to same discretized normal. In other words, it suffices to *fool* the proof of

Theorem 3. The problem is that proofs of Theorem 3 or its variants are often significantly involved (and based on Fourier analysis or Stein's method) and thus are not readily amenable to derandomization. Rather, the authors come up with a new and simple (albeit quantitively slightly weaker) proof for the discrete CLT. This proof is much simpler and amenable to derandomization. This gives a very high level overview of the utility of Theorem 3 for derandomization of combinatorial shapes (ignoring several details and complications).

We now list a shortcoming of the approach employed in [GMRZ13]. Note that [GMRZ13] gets $O(\log^2 n)$ seed for any inverse polynomial error which is what is guaranteed by [INW94]. The main conceptual barrier in extending their techniques to get an error $o(n^{-1/2})$ with $o(\log^2 n)$ seed is as follows: One of the main ideas that the authors bring in is to fool the proof of the discrete CLT up to error ϵ. As long as the discrete CLT has error $O(\epsilon)$, this automatically guarantees fooling the combinatorial shape to error $O(\epsilon)$. However, note that the optimal error rate for discrete CLTs with variance σ^2 is $O(\sigma^{-1})$. Since, $\sigma \leq n^{1/2}$, this approach is not useful for getting error rate $o(n^{-1/2})$. In fact, the error rate of $o(n^{-1/2})$ is optimal for CLTs with weaker metrics such as the Kolmogorov distance.

While [GMRZ13] requires $O(\log^2 n)$ seed for any inverse polynomial error (including $\Omega(n^{-1/2})$), the above discussion naturally motivates the question of looking at possible extensions of the usual central limit theorem where the rate of error is less than $n^{-1/2}$. With this motivation, we are led to the theory of asymptotic expansions.

1.1 Asymptotic expansions

Our aim is to get better convergence rates for sums of independent integer random variables with respect to the ℓ_1 metric. However, for the moment, we start with the weaker criterion of Kolmogorov metric. Also, to motivate the solution for discrete random variables, we first start with the problem of getting better convergence rates for sums of independent (not necessarily discrete) random variables. Towards this, let us assume that $X_1, \ldots, X_n$ are i.i.d. random variables (with common distribution X) with mean 0 and variance 1. Let $Z = (\sum X_i)/\sqrt{n}$. The classical central limit theorem [Fel68] states that

$$d_K\left(Z, \mathcal{N}(0,1)\right) = O\left(\frac{1}{\sqrt{n}}\right).$$

Here the constant in $O(\cdot)$ depends on X_1. It is natural to ask if instead of approximating Z by a normal, one can get an even better rate of convergence. More precisely, a normal distribution is parameterized by its first two moments. It is natural to investigate the possibility that approximating Z using its first $k > 2$ moments may get us a faster convergence rate.

In answering this question, a basic distinction needs to be made between discrete random variables and continuous random variables. To see why this distinction is necessary, additionally let $X_1, \ldots, X_n$ be supported on $\{-1, 1\}$. Then the cdf of Z has a discontinuity of size $O(n^{-1/2})$ at 0. As a result, no continuous measure can approximate the cdf of Z with error $o(n^{-1/2})$. This shows that there is a fundamental gap between approximations with error $o(n^{-1/2})$ and those with error $\Omega(n^{-1/2})$. We can also ask whether the discrete/continuous dichotomy is the only real barrier?

The answer is affirmative. The theory of asymptotic expansions (cf.[Pet75]) is fairly well-understood for sums of continuous i.i.d. random variables. In particular, as long as $\limsup |\widehat{X}(t)| < 1$,

$$\sup_x |\Pr[Z \leq x] - \Pi_{n,k}(x)| = O\left(\frac{1}{n^{(k-1)/2}}\right), \tag{1}$$

where

$$\Pi_{n,k}(x) = \Phi(x)\left(1 + \sum_{\nu=1}^{k-2} Q_\nu(x)\right).$$

In the above, for all ν, Q_ν is a polynomial whose coefficients are given in terms of the first $\nu+2$ moments of Z. The exact polynomial used here is difficult to state succinctly (see Section 2 for the explicit description of Q_ν). However, the only thing relevant for the discussion is that Q_ν is fully specified knowing just the first $\nu+2$ moments of Z. The condition that $\limsup|\widehat{X}(t)| < 1$ is known in literature as Cramér's condition and the resulting expansion (given in $\Pi_{n,k}(x)$) is known as an Edgeworth expansion. This expansion was first obtained by Cramér [Cra28]. For a more modern reference, see [Fel68]. It is important to mention that $O(\cdot)$ in the above theorem depends upon the random variable X and in particular, incorporates the quantitative aspect of Cramér's condition. The quantitative dependence is as follows (see Petrov [Pet75]):

$$\sup_x |\Pr[Z \leq x] - \Pi_{n,k}(x)| = \frac{k^{O(k)} \cdot \beta_k}{\sigma^k \cdot n^{-\frac{k-1}{2}}} + n^{\frac{k(k+1)}{2}} \cdot \left(\frac{1}{2n} + \sup_{|\xi| > \frac{\sigma^2}{12\beta_3}} |\widehat{X}(\xi)|\right)^n \tag{2}$$

where

$$\Pi_{n,k}(x) = \Phi(x)\left(1 + \sum_{\nu=1}^{k-2} Q_\nu(x)\right).$$

Here σ is the standard deviation of X and $\beta_i = \mathbf{E}[|X|^i]$. To get an idea of the error term, assume that $\beta_k/\sigma^k = O_k(1)$. This is indeed true for a large class of random variables. With this assumption, the first error term is $O_k(1) \cdot n^{-\frac{k-1}{2}}$. For the second term, provided that $\sup_{|\xi|>\sigma^2/(12\beta_3)} |\widehat{X}(\xi)|$ is bounded away from 1 by a constant, for any $k = \tilde{o}(\sqrt{n})$, the term is exponentially small.

It is also worthwhile to mention that our interest in the exact dependence of the error term on the random variables $\{X_i\}$ is because in our application, our random variables $\{X_i\}$ will not be fixed but rather have a dependence on n. Thus, knowing the precise dependence is necessary to get any non-trivial bounds.

At this point, we face two problems. The first is that, in general, we will be dealing with sums of independent but not necessarily identically distributed random variables. In fact, while [Pet75] has (nearly) explicit estimates for the convergence rate for i.i.d. variables, to the best of our knowledge, none of the standard references (e.g. [Pet75, BR86]) on asymptotic expansions have explicit estimates for non i.i.d. variables. The second problem is even more serious. Even if we restrict ourselves to i.i.d. variables which are discrete, Cramér's condition is not satisfied. For the rest of discussion to make sense, it is useful to define the notion of a lattice-valued random variable.

Definition 4. *A random variable X is said to lie on the lattice $\mathcal{L} = \{a + b \cdot h\}_{b \in \mathbb{Z}}$ if $\Pr[X \notin L] = 0$. The maximum h for which there exists such a lattice $\mathcal{L}$ is said to be the maximal span of X.*

For the rest of the discussion, if $X_1, \ldots, X_n$ are discrete, we will assume they are all lattice-valued with a common maximal span. The theory for non-lattice valued random variables is not nearly as well-understood (cf. [BHW94] for a discussion). Assuming that $X_1, \ldots, X_n$ are i.i.d. and lattice valued, correction terms (which are discontinuous) can be added in (1) to get a similar approximation for sums of i.i.d. lattice valued random variables. For example, the following theorem can be found in [IL71].

Theorem 4. *Let $X_1, \ldots, X_n$ be i.i.d. variables such that for all $i \in [n]$, $\mathbf{E}[X_i] = 0$, $\mathrm{Var}(X_i) = 1$ and X_i are lattice valued supported on $\mathbb{Z}$ with maximal span 1. Then,*

$$\sup_x \left|\Pr[Z \leq x] - \Pi_{n,k}(x) - \sum_{\nu=1}^{k-2} (-1)^{\lfloor(\nu-1)/2\rfloor} \cdot n^{-\nu/2} \cdot S_\nu(x \cdot \sqrt{n}) \cdot \frac{d^\nu \Pi_{n,k}(x)}{dx^\nu}\right| = O\left(\frac{1}{n^{(k-2)/2}}\right),$$

where

$$S_{2j}(x) = 2 \cdot \sum_{k=1}^{\infty} (2\pi k)^{-2j} \cdot \cos(2\pi kx) \quad \textit{and} \quad S_{2j+1}(x) = 2 \cdot \sum_{k=1}^{\infty} (2\pi k)^{-2j-1} \cdot \sin(2\pi kx)$$

While this provides an analogue of (1) for the case of lattice valued random variables, the dependence of the error on the random variable X_1 is not explicit in any of the references. The situation is even worse for non identically distributed lattice valued random variables. While in principle, it is entirely possible that the machinery used in [IL71, Pet75] can be used to give fully-explicit bounds even for the case of non-identical lattice random variables, previous literature seems to have avoided this approach noting that this approach seems unwieldy for getting explicit bounds. In fact, Barbour and Čekanvičius [BČ02] use this as their motivation for getting asymptotic expansions for sums of independent integer valued random variables. Their approach is based on using Stein's method for approximations using signed compound Poissons (SCP). While the approach in [BČ02] gives fairly explicit bounds and works for sums of independent (not necessarily identical) integer random variables, the bounds do not scale well when the variance of the random variables increase. This will be true in our case which forces us to seek an alternate approach.

1.2 Our approach

Let $X_1, \ldots, X_n$ be independent lattice valued random variables and $Z = \sum X_i$. We follow the approach of Esséen [Ess45], which treats the case of i.i.d. random variables (both lattice and non-lattice) and combine it with ideas described in Petrov [Pet75] which obtains explicit estimates for i.i.d. non-lattice random variables. In particular, the common approach in both these works is to first study the characteristic function of the random variable Z (i.e. the Fourier transform) and show that the Fourier transform can be well-approximated pointwise in terms of the first ν moments with an error of $n^{-(\nu-1)/2}$. In other words, we obtain an expression $\widetilde{Z}(\xi)$ which is described only in terms of the first ν moments of Z and that $\|\widehat{Z}(\xi) - \widetilde{Z}(\xi)\|_\infty = O(n^{-(\nu-1)/2})$.

The next step is to show that closeness in $\|\cdot\|_\infty$ norm in the Fourier transform can translate to closeness in total variation norm for the actual distance. This is somewhat different in the two works. In particular, for lattice valued random variables, weaker but significantly easier-to-obtain estimates can be obtained via the technique in Esseen [Ess45] which still suffices for our purposes. The approach in Petrov [Pet75] involves more Fourier analysis and is harder to do but sometimes obtain better estimates.

Obtaining the expression $\widetilde{Z}(\xi)$ is most of the technical work for getting the asymptotic expansions theorem and in this, we largely follow the ideas laid out in Petrov [Pet75]. Of course, because we deal with non i.i.d. variables, our calculations become more complicated (Further, at some points, we do need to make modifications as the quantitative estimates given by [Pet75] are useful for us). Another significant point of departure is that we deal with lattice valued variables as opposed to [Pet75] whose explicit bounds only deal with non-lattice valued variables.

Having obtained the expression for $\widetilde{Z}(\xi)$, we follow the routine to go from ℓ_∞ distance in the Fourier transform to ℓ_1 distance among the distributions in a fairly verbatim manner, thus obtaining our final bound. The actual theorem obtained is somewhat involved to state here and requires a fair bit of notation. Thus, we defer it to the next section.

1.3 Application to fooling combinatorial sums

Armed with this asymptotic expansion theorem, we recall our main result concerning derandomization of combinatorial sums is as follows.

Theorem. *There is a efficiently computable PRG $G_{csum} : \{0,1\}^t \to [m]^n$ which ϵ-fools every $f \in \mathsf{Csum}(m, n)$ and $t_{\mathsf{csum}} = O(\log m + \log^{3/2}(n/\epsilon))$.*

The approach is best demonstrated by considering the setting when $\epsilon = n^{-\Theta(1)}$. Similar to [GMRZ13], let us assume that the underlying combinatorial sum is specified by the tuple $(f_1, \ldots, f_n)$. Let $(y_1, \ldots, y_n) \sim U_{[m]^n}$ and $Z = \sum f_i(y_i)$.

- Low variance case: $\mathrm{Var}(Z) \leq 2^{c \cdot \sqrt{\log n}}$.
- High variance case: $\mathrm{Var}(Z) > 2^{c \cdot \sqrt{\log n}}$.

Here c is some fixed positive constant. We remark that if $\epsilon = n^{-\omega(1)}$, then we do an *alphabet reduction* step which shows that it suffices to treat the case when $m = \mathsf{poly}(n/\epsilon)$. The seed length required for this step is $O(\log(m \cdot n/\epsilon))$. This is a fairly easy step and hence, we do not describe it here.

Coming back to the case at hand, let $\epsilon = n^{-\Theta(1)}$ and assume that we are in case (i). In this case, we set $t = 2^{C\sqrt{\log n}}$ where $C > c$ is a large constant. We also set $k = \Theta(\sqrt{\log n})$. Let $H_{k,n,t}$ be a family of k-wise independent hash functions mapping $[n]$ to $[t]$. Note that the seed for $H_{k,n,t}$ is $t_{\mathsf{hash}} = O(\log^{3/2} n)$. The PRG first samples $h \sim \mathcal{H}_{k,n,t}$ to partition $[n]$ into t *buckets.* Within each bucket, the PRG uses k-wise independence (the distribution across the buckets is independent). Mimicking the machinery for low-variance case from [GMRZ13] (described earlier) and combining it with fairly standard techniques for obtaining concentration bounds for sums of k-wise independent random variables, we obtain the the PRG described here ϵ-fools $f \in \mathsf{Csum}(m,n)$ with error ϵ. Unfortunately, as before, since the seed across the buckets are independent, and hence the seed length is prohibitively large. However, as in [GMRZ13], f composed with the PRG itself can be seen as a ROBP with much smaller size. Derandomizing this using Theorem 1 gets us the final seed length of $O(\log m + \log^{3/2}(n/\epsilon))$.

We next describe the approach for the high variance case. As the reader might have guessed, the high variance case is where the power of asymptotic expansions is used. Here we set $t = 2^{C\sqrt{\log n}}$ where $C < c$ is a constant. We also set $k = \Theta(\sqrt{\log n})$. The PRG first samples $h \sim H_{k,n,t}$ to partition $[n]$ into t buckets. Note that the seed required for this step is $t_{\mathsf{hash}} = O(\log^{3/2}(n))$. Let $G_{k,n/t,m} : \{0,1\}^{t_1} \to [m]^{n/t}$ be a k-wise independent generator for the domain $[m]^{n/t}$. Further, let $G^{(cs)}_{m,n/t} : \{0,1\}^{t_2} \to [m]^{n/t}$ be the PRG from [GMRZ13] which δ-fools $\mathsf{Csum}(m, n/t)$ with δ set to a very small constant. Let $G^{(cs)}_{m,n/t} \otimes G_{k,n/t,m} : \{0,1\}^{t_1+t_2} \to [m]^{n/t}$ be defined as

$$G^{(cs)}_{m,n/t} \otimes G_{k,n/t,m} : (z_1, z_2) \mapsto G_{k,n/t,m}(z_1) \oplus_m G^{(cs)}_{m,n/t}(z_2).$$

The final PRG applies t independent copies of $G^{(cs)}_{m,n/t} \otimes G_{k,n/t,m}$ across the different buckets. First of all, we note that the seed for each bucket is $O(\log^{3/2}(n))$. While the seed length for this PRG is large, by applying the same trick (of derandomizing f composed with the PRG using Theorem 1), we note that the total seed length for the PRG can be bounded by $O(\log^{3/2}(n))$. What remains to be proven is that the basic PRG described here ϵ-fools $f \in \mathsf{Csum}(m,n)$ when $\mathrm{Var}(Z) > 2^{c \cdot \sqrt{\log n}}$.

First of all, let $Z_{f,h^{-1}(i)} = \sum_{j \in h^{-1}(i)} f_j(y_j)$. Further, let $Z'_{f,h^{-1}(i)} = \sum_{j \in h^{-1}(i)} f_j(y_j)$ where $y'_1, \ldots, y'_n$ are sampled from the output of the PRG. We first observe that the random variables $\{Z'_{f,h^{-1}(i)}\}$ are independent for $i \in [t]$ (as are $\{Z_{f,h^{-1}(i)}\}$). Further, by our construction, we have that for any $j \in [1, \ldots, k]$, the j^{th} moment of $Z'_{f,h^{-1}(i)}$ is identical to that of $Z_{f,h^{-1}(i)}$ for $i \in [t]$. In fact, because our PRG *contains an independent copy of* $G^{(cs)}_{m,n/t}$, it ensures the Fourier spectrum of $Z'_{f,h^{-1}(i)}$ is also δ-close to the Fourier spectrum $Z_{f,h^{-1}(i)}$ (for $i \in [t]$) in ℓ_∞ distance. Further, an application of concentration bounds for sums of k-wise independent random variables implies that with probability $1 - n^{-\theta(1)}$ for $h \sim H_{k,n,t}$, for all $i \in [t]$,

$$\frac{\mathrm{Var}(Z)}{2 \cdot t} \leq \mathrm{Var}\big(Z_{f,h^{-1}(i)}\big) \leq \frac{3 \cdot \mathrm{Var}(Z)}{2 \cdot t}.$$

It turns out that the above conditions are sufficient to apply the asymptotic expansions theorem and imply that

$$\left\| \sum_{i=1}^{t} Z'_{f,h^{-1}(i)} - \sum_{i=1}^{t} Z_{f,h^{-1}(i)} \right\|_1 \approx t^{-\Omega(k)} \leq \epsilon.$$

This concludes our informal description of the PRG.

2 Asymptotic expansion for sums of independent random variables

2.1 Preliminaries for the asymptotic expansion

We start by setting the notation which will be used throughout this section. For any random variable X supported on $\mathbb{R}$,

- $\widehat{X} : \mathbb{R} \to \mathbb{C}$ denote its characteristic function. In other words, $\widehat{X}(\xi) = \mathbf{E}_{x \sim X}[e^{i \cdot \xi \cdot x}]$.
- $\alpha_{X,k}$ denotes the k^{th} moment of X.
- $\beta_{X,k}$ dentoes the k^{th} absolute moment of X.
- $\gamma_{X,k}$ denotes the k^{th} cumulant of X.

We now let $X_1, \ldots, X_n$ be n independent random variables. We will use the following shorthands.

- The variables $X_1, \ldots, X_n$ are centered.
- $\alpha_{i,k}$ will denote $\alpha_{X_i,k}$. Likewise for $\beta_{i,k}$ and $\gamma_{i,k}$.
- Also, $\beta_k = \sum_{i=1}^{n} \beta_{i,k}$.
- $\sigma^2 = \sum_{i=1}^{n} \alpha_{i,2}$ and $Z = (X_1 + \ldots + X_n)/\sigma$.

We will need to define a family of polynomials $\{P_\nu(i\xi)\}_{\nu \in \mathbb{N}}$. The coefficients of these polynomials will be determined by the moments of Z. For defining these, it will be convenient to assume that $\alpha_{i,k}$, $\beta_{i,k}$ and $\gamma_{i,k}$ exists for all $i \in [n]$ and $k \in \mathbb{N}$. After we finish defining the polynomials, we will see that this assumption is not necessary. Rather, for defining P_ν, we will only require existence of the first $\nu + 2$ moments. But it will help to gain the intuition behind defining these polynomials by assuming all moments, cumulants and absolute moments exist. We now state the main theorem (the exact expression for P_ν is not very important at this stage).

Theorem 5. *Let $X_1, \ldots, X_n$ independent centered random variables such that for $i = 1, \ldots, n$ supported on lattices of span 1. Further, let $Z = (X_1 + \ldots + X_n)/\sigma$. Note that Z is supported on a lattice with span $1/\sigma$. Call the lattice L_Z. Let us assume that $I \geq 2\sqrt{s \cdot \log s}$. Then,*

$$\left| \Pr[Z = z] - \frac{1}{2\pi\sigma} \cdot \int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right| \leq \eta_{\mathsf{low}} + \eta_{\mathsf{med}} + \eta_{\mathsf{high}}$$

where

$$\eta_{\mathsf{low}} = s^{O(s)} \cdot L_{3s}^{s-2}, \quad \eta_{\mathsf{med}} = e^{-\frac{I^2}{6}} + s^{O(s)} \cdot e^{-\frac{I^2}{4}} \text{ and } \eta_{\mathsf{high}} = \sup_{|\zeta| \in [\sigma^2/\beta_3, \pi]} \left| \prod_{i=1}^{n} \widehat{X_i}(\zeta) \right| + s^{O(s)} \cdot e^{-\frac{I^2}{4}}.$$

2.2 Description of $\{P_\nu\}$

Having stated the main theorem, we now start the process of describing the polynomials $\{P_\nu\}$. Toward this, we also recall the following connection between moments and cumulants.

Fact 1. *There are polynomials $p_{k,CM}$, $p_{k,MC} : \mathbb{R}^k \to \mathbb{R}$ such that for any random variable X,*

$$\alpha_{X,k} = p_{k,CM}(\gamma_{X,1}, \ldots, \gamma_{X,k}) \qquad \gamma_{X,k} = p_{k,MC}(\alpha_{X,1}, \ldots, \alpha_{X,k}).$$

Further, if a monomial $\prod_{i=1}^{k} \gamma_{X,i}^{s_i}$ appears in $p_{k,CM}$, then $\sum i \cdot s_i = k$. Likewise, for $p_{k,MC}$.

We next recall the cumulant generating function $\widetilde{X}$ defined as $\widetilde{X}(\xi) = \log \widehat{X}(\xi)$. Note that this definition means that $\widetilde{X}$ may not be defined on the entire $\mathbb{R}$. However, it is easy to see that for every X, there exists $c > 0$, such that $\widetilde{X}$ is defined on $[-c, c]$. We recall the following fact connecting the k^{th} derivative of $\widetilde{X}$ and the $\gamma_{X,k}$.

Fact 2. *The k^{th} derivative of $\widetilde{X}(\xi)$ exists at $\xi = 0$ if and only if $\gamma_{X,k}$ is finite. Further,*

$$\gamma_{X,k} = (-i)^k \cdot \frac{d^k \widetilde{X}(\xi)}{d\xi^k}\bigg|_{\xi=0}$$

Now by definition, we have that $\widetilde{Z}(\xi) = \sum_{i=1}^{n} \widetilde{X_i}(\xi/\sigma)$. For the moment, assume that the Maclaurin expansion of $\widetilde{X_i}$ exists in interval $[-c_i, c_i]$ for some $c_i > 0$. If $c = \min c_i$, then we see that the Maclaurin series expansion of Z exists in the interval $[-\sigma c, \sigma c]$. Further, for $\xi \in [-\sigma c, \sigma c]$, we have

$$\widetilde{Z}(\xi) = \sum_{j=1}^{n} \widetilde{X_j}(\xi/\sigma) = \sum_{\nu=1}^{\infty} \sum_{j=1}^{n} \frac{\gamma_{j,\nu} \cdot i^{\nu} \cdot \xi^{\nu}}{\sigma^{\nu} \cdot \nu!}$$

Next, we notice that $\gamma_{j,1} = 0$ for all $j \in [n]$. Also, $\sum_{j=2}^{n} \gamma_{j,2} = \sigma^2$. This result in the simplified expression

$$\widetilde{Z}(\xi) = -\frac{\xi^2}{2} + \sum_{\nu=1}^{\infty} \frac{\lambda_{\nu+2} \cdot (i\xi)^{\nu+2}}{(\nu+2)! \cdot \sigma^{\nu+2}},$$

where $\lambda_{\nu+2} = \sum_{j=1}^{n} \gamma_{j,\nu+2}$. This implies

$$\widehat{Z}(\xi) = e^{-\frac{\xi^2}{2}} \cdot \exp\left(\sum_{\nu=1}^{\infty} \frac{\lambda_{\nu+2} \cdot (i\xi)^{\nu+2}}{(\nu+2)! \cdot \sigma^{\nu+2}}\right).$$

We define $P_\nu(i\xi)$ to denote the coefficient of w^ν in the formal expansion of

$$\exp\left(\sum_{\nu=1}^{\infty} \frac{\lambda_{\nu+2} \cdot (i\xi)^{\nu+2}}{(\nu+2)! \cdot \sigma^{\nu+2}} \cdot w^{\nu}\right).$$

Now, observe that $\lambda_{\nu+2}/\sigma^{\nu+2} = \gamma_{Z,\nu+2}$. Then, we immediately see that $P_\nu(i\xi)$ is a polynomial whose coefficients are given by polynomials in $\gamma_{Z,3}, \ldots, \gamma_{Z,\nu+2}$. Observe that as we said before, for defining $P_\nu(i\xi)$, we do not need the Maclaurin series expansion of $\widetilde{Z}(\xi)$ to exist in any open set around 0. Neither do we require $\gamma_{j,k}$ to exist for all j and k. Rather, we require $\gamma_{j,k}$ to exist for all j and $1 \le k \le \nu + 2$. Note however that if $\gamma_{j,k}$ exists for all k and the Maclaurin series expansion for $\widehat{Z}(\xi)$ is valid, then

$$\widehat{Z}(\xi) = e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{\infty} P_\nu(i\xi)\right).$$

2.3 Alternate formulation for $\{P_\nu\}$

We now seek an alternate formulation of the polynomials P_ν which will be useful to us later on. Towards this, we make the following definition:

$$V_w(\xi) = \frac{\xi^2}{2} + \frac{1}{w^2} \cdot \log \widehat{Z}(\xi \cdot w) \tag{3}$$

With this, we make the following claim.

Claim 3. *For $V_w(\xi)$ defined as above, the following holds:*

$$\sum_{\nu=1}^{s-3} P_\nu(i\xi) \cdot w^\nu = \sum_{j=1}^{s-3} \sum_{k=0}^{2j+s-3} \frac{d^k V_w^j(\xi)}{d\xi^k}\bigg|_{\xi=0} \cdot \frac{\xi^k}{k!} \tag{4}$$

We begin with a few elementary claims about V_w. First, note that by definition, we have

$$\frac{d^k \log \widehat{Z}(\xi \cdot w)}{d\xi^k}\bigg|_{\xi=0} = \sum_{j=1}^{n} \frac{d^k \log \widehat{X_j}(\xi \cdot w/\sigma)}{d\xi^k}\bigg|_{\xi=0} = \sum_{j=1}^{n} \frac{w^k \cdot \gamma_{j,k} \cdot i^k}{\sigma^k} = \frac{w^k \cdot \lambda_k \cdot i^k}{\sigma^k}$$

The last but one equality follows by definition of cumulants. Using this, we easily have the following equations.

$$\text{For } k \le 2, \quad V^{(k)}(0) = 0. \tag{5}$$

$$\text{For } k > 2, \quad V^{(k)}(0) = \frac{w^{k-2} \cdot \lambda_k \cdot i^k}{\sigma^k}. \tag{6}$$

Towards achieving the characterization of $P_\nu(\xi)$, we note that,

$$P_\nu(i\xi) \cdot w^\nu = S_{1,\nu} + \ldots + S_{\nu,\nu}, \tag{7}$$

where

$$S_{\ell,\nu} = \sum_{\nu_1+\ldots+\nu_\ell=\nu:\, \prod \nu_j > 0} \prod_{j=1}^{\ell} \frac{\lambda_{\nu_j+2} \cdot (i\xi)^{\nu_j+2}}{(\nu_j+2)! \cdot \sigma^{\nu_j+2}} \cdot w^{\nu_j}$$

Note that in the above, different permutations of the tuple $(\nu_1, \ldots, \nu_i)$ are counted as distinct. We then have the following claim.

Claim 4.

$$S_{i,\nu} = \frac{d^{\nu+2i} V_w^i(\xi)}{d\xi^{\nu+2i}}\bigg|_{\xi=0} \cdot \frac{\xi^{\nu+2i}}{(\nu+2i)!}$$

Proof. By the Leibniz rule for differentiation, we have,

$$\frac{d^{\nu+2i} V_w^i(\xi)}{d\xi^{\nu+2i}}\bigg|_{\xi=0} \cdot \frac{\xi^{\nu+2i}}{(\nu+2i)!} = \sum_{\nu_1+\ldots+\nu_i=\nu+2i} \binom{\nu+2i}{\nu_1 \ \nu_2 \ \ldots \ \nu_i} \prod_{j=1}^{i} \frac{d^{\nu_j} V_w(\xi)}{d\xi^{\nu_j}}\bigg|_{\xi=0} \cdot \frac{\xi^{\nu+2i}}{(\nu+2i)!}$$

However, note that using (5), we can say that for any term indexed $(\nu_1, \ldots, \nu_j)$, unless all the $\nu_j > 2$, the term will vanish. Hence, we have

$$\begin{aligned} \frac{d^{\nu+2i} V_w^i(\xi)}{d\xi^{\nu+2i}}\bigg|_{\xi=0} \cdot \frac{\xi^{\nu+2i}}{(\nu+2i)!} &= \sum_{\nu_1+\ldots+\nu_i=\nu:\, \prod \nu_j > 0} \binom{\nu+2i}{\nu_1+2 \ \nu_2+2 \ \ldots \ \nu_i+2} \prod_{j=1}^{i} \frac{d^{\nu_j+2} V_w(\xi)}{d\xi^{\nu_j+2}}\bigg|_{\xi=0} \cdot \frac{\xi^{\nu+2i}}{(\nu+2i)!}, \\ &= \sum_{\nu_1+\ldots+\nu_i=\nu+2i} \binom{\nu+2i}{\nu_1 \ \nu_2 \ \ldots \ \nu_i} \prod_{j=1}^{i} \frac{d^{\nu_j} V_w(\xi)}{d\xi^{\nu_j}}\bigg|_{\xi=0} \cdot \frac{\xi^{\nu+2i}}{(\nu+2i)!}, \end{aligned}$$

which proves the claim. □

Proof of Claim 3: We combine Claim 4 and (7), we have

$$\begin{aligned}\sum_{\nu=1}^{s-3} P_\nu(i\xi)\cdot w^\nu &= \sum_{\nu=1}^{s-3}\sum_{j=1}^{\nu} \left.\frac{d^{\nu+2j}V_w^j(\xi)}{d\xi^{\nu+2j}}\right|_{\xi=0}\cdot\frac{\xi^{\nu+2j}}{(\nu+2j)!}\\ &= \sum_{j=1}^{s-3}\sum_{\nu=j}^{s-3} \left.\frac{d^{\nu+2j}V_w^j(\xi)}{d\xi^{\nu+2j}}\right|_{\xi=0}\cdot\frac{\xi^{\nu+2j}}{(\nu+2j)!} \qquad (8)\end{aligned}$$

To simplify this into our final form, we make the following claim.

Claim 5. *For any $j \in \mathbb{N}$ and $k < 3j$,*

$$\left.\frac{d^k V_w^j(\xi)}{d\xi^k}\right|_{\xi=0} = 0.$$

Proof.

$$\left.\frac{d^k V_w^j(\xi)}{d\xi^k}\right|_{\xi=0} = \sum_{i_1+\ldots+i_j=k}\binom{k}{i_1\ \ldots\ i_j}\prod_{\ell=1}^{j}\left.\frac{d^{i_\ell}V_w(\xi)}{d\xi^{i_\ell}}\right|_{\xi=0} = 0$$

The first equality follows by Leibniz rule for differentiation while the second equality uses the fact that since $k < 3j$, there is at least one ℓ such that $i_\ell < 3$. Using (5), we get that each of the summands and hence the entire sum is zero. □

Combining Claim 5 with (8), we get Claim 3. □

2.4 Approximation of the Fourier spectrum using moments

Let us begin by defining $L_k = (\beta_k/\sigma^k)^{1/(k-2)}$. Let us assume that $L_k \le 1$ for $k \ge 3$. This can be done without loss of generality as otherwise, Theorem 5 holds trivially. Let us now define

$$I = \frac{1}{C}\cdot\min\left\{\min_i \frac{\sigma}{\sigma_i}, \frac{1}{L_{3s}}\right\}, \qquad (9)$$

for a sufficiently large constant $C > 0$. The exact value of C is not important as long as it is sufficiently large. We now state the main lemma of this section which is also the main workhorse for proving Theorem 5.

Lemma 6. *For $|\xi| \le I$,*

$$\left|\widehat{Z}(\xi) - e^{-\frac{\xi^2}{2}}\cdot\left(1+\sum_{\nu=1}^{s-3}P_\nu(i\xi)\right)\right| \le c(s)\cdot L_{3s}^{s-2}\cdot\left(|\xi|^s + |\xi|^{3s-8)}\right)\cdot e^{-2\xi^2/5}$$

where $c(s) = s^{O(s)}$.

We will need a few preliminaries before we start with the proof of Lemma 6. The following inequality follows easily from monotonicity of norms.

Claim 7. *Let X be a real-valued random variable. Assuming that β_{X,k_1+k_2} exists, $\beta_{X,k_1}\cdot\beta_{X,k_2} \le \beta_{X,k_1+k_2}$.*

The following lemma states that $\{L_k\}$ is a monotonically non-decreasing sequence.

Claim 8. *$\{L_k\}$ is a monotonically non-decreasing sequence.*

Proof. Let $k_1 > k_2 \geq 2$ and Y be any real valued random variable. Note that for $\lambda = \frac{k_2-2}{k_1-2}$, $k_2 = \lambda k_1 + (1-\lambda)2$. Hence, applying Claim 35, we have

$$\|Y\|_{k_1}^{\lambda k_1} \cdot \|Y\|_2^{(1-\lambda)2} \geq \|Y\|_{k_2}^{k_2}.$$

Exponentiating all sides by $(k_2-2)^{-1}$ and rearranging powers of $\|Y\|_2$, we have

$$\frac{\|Y\|_{k_1}^{\frac{k_1}{k_1-2}}}{\|Y\|_2^{\frac{k_1}{k_1-2}}} \geq \frac{\|Y\|_{k_2}^{\frac{k_2}{k_2-2}}}{\|Y\|_2^{\frac{k_2}{k_2-2}}}.$$

We now instantiate the random variable Y as follows: Sample $i \in [n]$ u.a.r. and then sample X_i. Applying the previous inequality on the random variable Y, we have

$$\frac{\left(\mathbf{E}_{i\in[n]}\beta_{i,k_1}\right)^{\frac{1}{k_1-2}}}{\left(\mathbf{E}_{i\in[n]}\beta_{i,2}\right)^{\frac{k_1}{2(k_1-2)}}} \geq \frac{\left(\mathbf{E}_{i\in[n]}\beta_{i,k_2}\right)^{\frac{1}{k_2-2}}}{\left(\mathbf{E}_{i\in[n]}\beta_{i,2}\right)^{\frac{k_2}{2(k_2-2)}}}$$

and this immediately implies

$$\frac{\left(\sum_{i\in[n]}\beta_{i,k_1}\right)^{\frac{1}{k_1-2}}}{\left(\sum_{i\in[n]}\beta_{i,2}\right)^{\frac{k_1}{2(k_1-2)}}} \geq \frac{\left(\sum_{i\in[n]}\beta_{i,k_2}\right)^{\frac{1}{k_2-2}}}{\left(\sum_{i\in[n]}\beta_{i,2}\right)^{\frac{k_2}{2(k_2-2)}}}$$

which proves that $\{L_k\}$ is monotonically non-decreasing sequence. □

We will also need a bound on $P_\nu(i\xi)$. Towards that, we establish the following simple bounds.

Claim 9. *Let X be any centered random variable. Then, for any k, (with the notations as before), $|\gamma_{X,k}| \leq 2^k \cdot \beta_{X,k}$.*

Claim 10. *For Z defined as above, $\frac{|\lambda_j|}{\sigma^j} \leq 2^j \cdot j^j$.*

Proof. We consider the random variable $Z' = \sigma \cdot Z = X_1 + \ldots + X_n$. Note that $\lambda_j = \gamma_{Z',j}$. Thus, using Claim 9, it suffices to bound $\beta_{Z',j}$. Assume for the moment that j is even. For an integer j, let $P(j)$ denote the set of all partitions of j none of which are 1. Then, we have

$$\begin{aligned}
\mathbf{E}[|X_1+\ldots+X_n|^j] &\leq \sum_{(a_1,\ldots,a_t)\in P(j)} \binom{j}{a_1\ a_2\ \ldots\ a_t} \cdot \prod_{i=1}^{t} \mathbf{E}\left[\sum_j |X_j|^{a_i}\right] \\
&= \sum_{(a_1,\ldots,a_t)\in P(j)} \binom{j}{a_1\ a_2\ \ldots\ a_t} \cdot \prod_{i=1}^{t} \sigma^{a_i} \cdot L_{a_i}^{a_i-2} \\
&\leq \sigma^j \cdot j^j.
\end{aligned}$$

Note that the last inequality uses that $L_t \leq 1$ for all $t > 2$. This finishes the proof for even j. For odd j, it follows from monotonicity of norms. □

We next make the following claim which bounds the value of $P_\nu(i\xi)$.

Claim 11.

$$|P_\nu(i\xi)| \leq \nu^{O(\nu)} \cdot (|\xi|^{\nu+2} + |\xi|^{3\nu})$$

Proof. Begin by noting that

$$P_\nu(i\xi) = \sum_{\ell=1}^{\nu} \sum_{\nu_1+\ldots+\nu_\ell=\nu:\prod \nu_j>0} \prod_{j=1}^{\ell} \frac{\lambda_{\nu_j+2}\cdot(i\xi)^{\nu_j+2}}{(\nu_j+2)!\cdot\sigma^{\nu_j+2}}$$

Thus,

$$|P_\nu(i\xi)| \le \sum_{\ell=1}^{\nu} \sum_{\nu_1+\ldots+\nu_\ell=\nu:\prod \nu_j>0} \prod_{j=1}^{\ell} \frac{|\lambda_{\nu_j+2}|\cdot|\xi|^{\nu_j+2}}{(\nu_j+2)!\cdot\sigma^{\nu_j+2}} \le \sum_{\ell=1}^{\nu} \sum_{\nu_1+\ldots+\nu_\ell=\nu:\prod \nu_j>0} \prod_{j=1}^{\ell} (2e)^{\nu_j+2}\cdot|\xi|^{\nu_j+2},$$

where the last inequality uses Claim 10. The final term can be easily bound by $\nu^{O(\nu)}\cdot(|\xi|^{\nu+2}+|\xi|^{3\nu})$. □

As a result, we also get the following corollary.

Corollary 12.

$$\left|e^{-\frac{\xi^2}{2}}\cdot\left(1+\sum_{\nu=1}^{s-3} P_\nu(i\xi)\right)\right| \le s^{O(s)}\cdot e^{-\xi^2/2}\cdot(|\xi|^3+|\xi|^{3s}).$$

Proof of Lemma 6: We first start by showing that for $|\xi|\le I$, the function $V_w(\xi)$ is well-defined. Note that $V_w(\xi)$ is well defined as long as $\widehat{Z}(\xi\cdot w)\neq 0$. The following claim states the interval in which this indeed holds and thus $V_w(\xi)$ is well defined.

Claim 13. *For $|\xi|\le\frac{\sigma}{\sigma_i}$, $|\widehat{X_i}(\xi\cdot w)|\ge 1/2$.*

Proof. Using Taylor's theorem (for complex valued functions), we have

$$\left|\widehat{X}_i\left(\frac{\xi\cdot w}{\sigma}\right)-1\right| = \left|\mathbf{E}\left[\exp\left(\frac{i\cdot w\cdot\xi\cdot X_i}{\sigma}\right)-1-\frac{i\cdot\xi\cdot w\cdot X_i}{\sigma}\right]\right| \le \frac{\xi^2 w^2\cdot\mathbf{E}[X_i^2]}{2\cdot\sigma^2}.$$

For $|\xi|\le\frac{\sigma}{\sigma_i}$, the right hand side is at most 1/2, finishing the proof. □

Claim 14. *For $|\xi|\le\min_i\frac{\sigma}{\sigma_i}$, $\widehat{Z}(\xi\cdot w)\neq 0$.*

Proof. Note that $\widehat{Z}(\xi\cdot w)=\prod_{i=1}^n\widehat{X_i}((\xi\cdot w)/\sigma)$. Using Claim 13, we get the claim. □

We now establish an upper bound on $V_w(\xi)$. For this, our strategy is to use Taylor's theorem around $\xi=0$. In particular, we will show the V_w and its first two derivatives are zero and then establish a bound on the third order derivative. Using Taylor's theorem, this will lead us to an upper bound on $V_w(\xi)$. We start with the following simple claims. Having established that the function $V_w(\xi)$ as well as its first and second order derivatives vanish at 0, we now establish a bound on the supremum of the third order derivative. To do this, we first prove the following simple claim.

Claim 15. *For any $\xi\in\mathbb{R}$, $\left|\frac{\partial^k\widehat{X_i}(\xi)}{\partial\xi^k}\right|\le\beta_{i,k}$.*

Proof. Note that $\widehat{X_i}(\xi)=\mathbf{E}_{x\in X_i}[\exp(ix\xi)]$. As a consequence, we get that

$$\frac{\partial^k\widehat{X_i}(\xi)}{\partial\xi^k}=\mathbf{E}_{x\in X_i}[(i\cdot x)^k\exp(ix\xi)].$$

This immediately implies the claim. □

To establish the upper bound on the third order derivative of $V_w(\xi)$, we establish the more general bound on the s^{th} order derivative for $s > 2$. This will be useful later on. For this, we observe the following simple fact.

$$\frac{d^k \log u}{dx^k} = \sum_{i_1,\ldots,i_k : \sum j \cdot i_j = k} c_{i_1,\ldots,i_k} \cdot \frac{1}{u^{\|(i_1,\ldots,i_k)\|_1}} \prod_{\ell=1}^{k} \left(\frac{d^\ell u}{dx^\ell} \right)^{i_\ell} \tag{10}$$

where a simple induction can be used to show that

$$\sum_{i_1,\ldots,i_k : \sum j \cdot i_j = k} |c_{i_1,\ldots,i_k}| \le (2k)!. \tag{11}$$

Claim 16. *For $s > 2$ and $|\xi| \le I$,*

$$\left| \frac{d^s V_w(\xi)}{d\xi^s} \right| \le (4s)! \cdot w^{s-2} \cdot L_s^{s-2}.$$

Proof.

$$\begin{aligned}
\frac{d^s V_w(\xi)}{d\xi^s} &= \frac{1}{w^2} \cdot \left(\sum_{\ell=1}^{n} \frac{d^s \log \widehat{X_\ell}(\xi \cdot w/\sigma)}{d\xi^s} \right) \\
&= \frac{1}{w^2} \cdot \left(\sum_{\ell=1}^{n} \sum_{i_1,\ldots,i_s : \sum j \cdot i_j = s} c_{i_1,\ldots,i_s} \cdot \frac{1}{\widehat{X_\ell}(\xi \cdot w/\sigma)^{\|(i_1,\ldots,i_s)\|_1}} \cdot \prod_{j=1}^{s} \left(\frac{d^j \widehat{X_\ell}(\xi \cdot w/\sigma)}{d\xi^j} \right)^{i_j} \right)
\end{aligned}$$

The last equality uses (10) and (11). Using Claim 15 and Claim 13, we get

$$\left| \frac{d^s V_w(\xi)}{d\xi^s} \right| \le \frac{1}{w^2} \cdot \left(\sum_{\ell=1}^{n} \sum_{i_1,\ldots,i_s : \sum j \cdot i_j = s} |c_{i_1,\ldots,i_s}| \cdot \frac{1}{|\widehat{X_\ell}(\xi \cdot w/\sigma)|^{\|(i_1,\ldots,i_s)\|_1}} \cdot \prod_{j=1}^{s} \left| \left(\frac{d^j \widehat{X_\ell}(\xi \cdot w/\sigma)}{d\xi^j} \right) \right|^{i_j} \right)$$

Note that $|\widehat{X_\ell}(\xi \cdot w/\sigma)| \ge 1/2$ for $|\xi| \le I$ (Claim 13), we get

$$\begin{aligned}
\left| \frac{d^s V_w(\xi)}{d\xi^s} \right| &\le \frac{1}{w^2} \cdot \left(\sum_{\ell=1}^{n} \sum_{i_1,\ldots,i_s : \sum j \cdot i_j = s} |c_{i_1,\ldots,i_s}| \cdot 2^{\|(i_1,\ldots,i_s)\|_1} \cdot \prod_{j=1}^{s} \left| \left(\frac{d^j \widehat{X_\ell}(\xi \cdot w/\sigma)}{d\xi^j} \right) \right|^{i_j} \right) \\
&\le \frac{1}{w^2} \cdot \left(\sum_{\ell=1}^{n} \sum_{i_1,\ldots,i_s : \sum j \cdot i_j = s} |c_{i_1,\ldots,i_s}| \cdot 2^{\|(i_1,\ldots,i_s)\|_1} \cdot \prod_{j=1}^{s} \frac{\beta_{\ell,j}^{i_j} \cdot w^{j \cdot i_j}}{\sigma^{j \cdot i_j}} \right) \quad \text{(Claim 15)} \\
&\le \frac{1}{w^2} \cdot \frac{w^s}{\sigma^s} \left(\sum_{\ell=1}^{n} \sum_{i_1,\ldots,i_s : \sum j \cdot i_j = s} |c_{i_1,\ldots,i_s}| \cdot 2^{\|(i_1,\ldots,i_s)\|_1} \cdot \prod_{j=1}^{s} \beta_{\ell,j}^{i_j} \right) \\
&\le \frac{w^{s-2}}{\sigma^s} \left(\sum_{\ell=1}^{n} \sum_{i_1,\ldots,i_s : \sum j \cdot i_j = s} |c_{i_1,\ldots,i_s}| \cdot 2^{\|(i_1,\ldots,i_s)\|_1} \cdot \beta_{\ell,s} \right) \quad \text{(Claim 7)} \\
&\le \frac{w^{s-2} \cdot \beta_s}{\sigma^s} \cdot 2^s \cdot (2s)! \le (4s)! \cdot w^{s-2} \cdot L_s^{s-2}.
\end{aligned}$$

□

Using this, we have the following corollary.

$$\text{For } |\xi| \le I, \ \left| \frac{d^3 V_w(\xi)}{d\xi^3} \right| \le 12! \cdot w \cdot L_3. \tag{12}$$

Also, repating nearly the same calculation as Claim 16, we also have

$$\text{For } |\xi| \le I, \ \left| \frac{d^2 V_w(\xi)}{d\xi^2} \right| \le 8! \tag{13}$$

Equipped with this inequality, we now establish an upper bound $V_w(\xi)$. Using Taylor's theorem and (5), we have that for $|\xi| \le I$,

$$|V_w(\xi)| \le \frac{1}{6}|\xi^3| \sup_{|\xi'| \le |\xi|} \left| \frac{d^3 V_w(\xi')}{d\xi'^3} \right| \le \Theta(1) \cdot w \cdot |\xi|^3 \cdot L_3 \tag{14}$$

Likewise, we have,

$$\left| \frac{dV_w(\xi)}{d\xi} \right| \le \Theta(1) \cdot w \cdot |\xi|^2 \cdot L_3 \quad \text{and} \quad \left| \frac{d^2 V_w(\xi)}{d\xi^2} \right| \le \Theta(1) \cdot w \cdot |\xi| \cdot L_3. \tag{15}$$

By choosing $C > 0$ sufficiently large in (9) and using $L_{3s} > L_3$, we also get that

$$|V_w(\xi)| \le \frac{w \cdot \xi^2}{10}, \quad \left| \frac{dV_w(\xi)}{d\xi} \right| \le \frac{3w \cdot \xi}{10} \quad \text{and} \quad \left| \frac{d^2 V_w(\xi)}{d\xi^2} \right| \le \frac{3w}{5}. \tag{16}$$

Note that by definition, we have

$$\left(\widehat{X}(\xi \cdot w) \right)^{1/w^2} = e^{-\xi^2/2} \cdot e^{V_w(\xi)}.$$

We seek to control the quantity

$$\left| e^{-\frac{\xi^2}{2}} \cdot e^{V_w(\xi)} - e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} w^\nu \cdot P_\nu(i\xi) \right) \right|$$

and then finally put $w = 1$. To control this quantity, we break this the difference into two parts.

$$R_1(w,\xi) = \sum_{\nu=s-2}^{\infty} \frac{e^{-\frac{\xi^2}{2}} \cdot V_w^\nu(\xi)}{\nu!} \qquad R_2(w,\xi) = e^{-\frac{\xi^2}{2}} \cdot \left(\sum_{\nu=1}^{s-3} \left(\frac{V_w^\nu(\xi)}{\nu!} - \frac{w^\nu \cdot P_\nu(i\xi)}{\nu!} \right) \right)$$

It is easier to control $R_1(w,\xi)$, so we begin with that.

$$\begin{aligned} |R_1(w,\xi)| &\le \sum_{\nu=s-2}^{\infty} \frac{e^{-\frac{\xi^2}{2}} \cdot |V_w(\xi)|^\nu}{\nu!} \le \frac{e^{-\frac{\xi^2}{2}} \cdot |V_w(\xi)|^{s-2}}{(s-2)!} \cdot e^{|V_w(\xi)|} \\ &\le \frac{|V_w(\xi)|^{s-2}}{(s-2)!} \cdot e^{-\frac{\xi^2}{2} + \frac{w\xi^2}{10}} \le \frac{2^{O(s)} \cdot w^{s-2} \cdot \xi^{3(s-2)} \cdot L_3^{s-2}}{(s-2)!} \cdot e^{-\frac{\xi^2}{2} + \frac{w\xi^2}{10}}. \end{aligned} \tag{17}$$

The penultimate inequality uses (16) and the last inequality uses (14). We will now control $R_2(w,\xi)$ which is slightly more tricky to control. Our calculation for this part is somewhat different from those in Petrov [Pet75] or Bhattacharya and Rao [BR86]. In particular, these two works truncate after $\nu = s-3$ (i.e. do an approximation in terms of the first $s-1$ moments). However, naively going through their calculations, it seems one needs to pay a factor of $s^{O(s^2)}$. On the other hand, we pay a factor of $s^{O(s)}$

but instead need to assume that the first $3s+2$ moments exists and the error bound is in terms of L_{3s}. Towards bounding $R_2(w,\xi)$, we have

$$\begin{aligned} R_2(w,\xi) &= e^{-\frac{\xi^2}{2}} \cdot \sum_{\nu=1}^{s-3} \left(\frac{V_w^\nu(\xi)}{\nu!} - \frac{w^\nu \cdot P_\nu(i\xi)}{\nu!} \right) \\ &= e^{-\frac{\xi^2}{2}} \cdot \sum_{\nu=1}^{s-3} \frac{1}{\nu!} \cdot \left(V_w^\nu(\xi) - \sum_{k=0}^{2\nu+s-3} \left.\frac{d^k V_w^\nu(\xi)}{d\xi^k}\right|_{\xi=0} \cdot \frac{\xi^k}{k!} \right) \end{aligned}$$

To get the second equality, we use (4). We next use Taylor's theorem to get

$$|R_2(w,\xi)| \le \sum_{\nu=1}^{s-3} \frac{e^{-\frac{\xi^2}{2}}}{\nu!} \cdot \frac{|\xi|^{2\nu+s-2}}{(2\nu+s-2)!} \kappa_{\nu,2\nu+s-2}(\xi), \tag{18}$$

where

$$\kappa_{\nu,2\nu+s-2}(\xi) = \sup_{|\xi|\le I} \left| \frac{d^{2\nu+s-2} V_w^\nu(\xi)}{d\xi^{2\nu+s-2}} \right|.$$

Thus our task reduces to bounding $\kappa_{\nu,2\nu+s-2}(\xi)$ for $\nu \in [1,\ldots,s-3]$. To bound this number, we recall the following basic fact about higher order derivatives of products of functions. Let $\mathbb{Z}^+ = \mathbb{N} \cup \{0\}$. Using the Leibniz rule, we have the following:

Lemma 17.

$$\frac{d^k \prod_{i=1}^{\ell} u_i}{dx^k} = \sum_{a \in {\mathbb{Z}^+}^{\ell} : \|a\|_1 = k} \binom{k}{a_1,\ldots,a_\ell} \prod_{i=1}^{\ell} \frac{d^{a_i} u_i}{dx^{a_i}}$$

Using Lemma 17, we have

$$\begin{aligned} \left| \frac{d^{2\nu+s-2} V_w^\nu(\xi)}{d\xi^{2\nu+s-2}} \right| &= \left| \sum_{a \in {\mathbb{Z}^+}^{\nu} : \|a\|_1 = 2\nu+s-2} \binom{2\nu+s-2}{a_1,\ldots,a_\nu} \prod_{i=1}^{\nu} \frac{d^{a_i} V_w(\xi)}{d\xi^{a_i}} \right|, \\ &\le \sum_{a \in {\mathbb{Z}^+}^{\nu} : \|a\|_1 = 2\nu+s-2} \binom{2\nu+s-2}{a_1,\ldots,a_\nu} \prod_{i=1}^{\nu} \left| \frac{d^{a_i} V_w(\xi)}{d\xi^{a_i}} \right|, \\ &= \sum_{a \in {\mathbb{Z}^+}^{\nu} : \|a\|_1 = 2\nu+s-2} \binom{2\nu+s-2}{a_1,\ldots,a_\nu} \prod_{i : a_i \in \{0,1\}} \left| \frac{d^{a_i} V_w(\xi)}{d\xi^{a_i}} \right| \prod_{i : a_i \ge 2} \left| \frac{d^{a_i} V_w(\xi)}{d\xi^{a_i}} \right|. \end{aligned} \tag{19}$$

Using Claim 16 and (13), we recall that for $a_i > 1$ and $|\xi| \le I$,

$$\left| \frac{d^{a_i} V_w(\xi)}{d\xi^{a_i}} \right| \le (4 \cdot a_i)! \cdot L_{a_i}^{a_i-2} \cdot w^{a_i-2}.$$

On the other hand, for $a_i \in \{0,1\}$ and $|\xi| \le I$, using (16) we have,

$$\left| \frac{d^{a_i} V_w(\xi)}{d\xi^{a_i}} \right| \le \frac{w \cdot \xi^{2-a_i}}{10}.$$

Applying the last two inequalities to (19),

$$\begin{aligned} \left| \frac{d^{2\nu+s-2} V_w^\nu(\xi)}{d\xi^{2\nu+s-2}} \right| &\le \sum_{a \in {\mathbb{Z}^+}^{\nu} : \|a\|_1 = 2\nu+s-2} \binom{2\nu+s-2}{a_1,\ldots,a_\nu} \prod_{i : a_i \in \{0,1\}} \frac{w \cdot \xi^{2-a_i}}{10} \cdot \prod_{i : a_i \ge 2} (4 \cdot a_i)! \cdot L_{a_i}^{a_i-2} \cdot w^{a_i-2}, \\ &\le \sum_{a \in {\mathbb{Z}^+}^{\nu} : \|a\|_1 = 2\nu+s-2} \binom{2\nu+s-2}{a_1,\ldots,a_\nu} \prod_{i : a_i \in \{0,1\}} \frac{w \cdot \xi^{2-a_i}}{10} \cdot \prod_{i : a_i \ge 2} (4 \cdot a_i)! \cdot L_{2\nu+s-2}^{a_i-2} \cdot w^{a_i-2} \end{aligned}$$

The last inequality uses that $L_{a_i} \leq L_{\|a\|_\infty} \leq L_{\|a\|_1}$ (using Claim 8). For any given $a \in \mathbb{Z}^{+\nu}$, use $\#a(0)$ to denote the number of zero entries in a and $\#a(1)$ to denote the one entries. Note that for any term in the above summation, we have

- The exponent of w is $\|a\|_1 - 2\|a\|_0 + 2\#a(1) + \#a(0)$.
- The exponent of $L_{2\nu+s-2}$ is $\|a\|_1 - 2\|a\|_0 + \#a(1)$.
- The exponent of ξ is $2\#a(0) + \#a(1)$.
- $\frac{1}{10} \cdot \prod_{i:a_i \geq 2}(4 \cdot a_i)! \leq (4(2\nu + s - 2))! \leq s^{O(s)}$.

Using elementary combinatorics, it is easy to show that

$$\sum_{a \in \mathbb{Z}^{+\nu}:\|a\|_1 = 2\nu+s-2} \binom{2\nu + s - 2}{a_1, \ldots, a_\nu} = \nu^{2\nu+s-2}.$$

Note that $|\xi| \leq L_{3s}^{-1}$, using Claim 7, we get that $|\xi| \leq L_{2\nu+s-2}^{-1}$. Thus,

$$L_{2\nu+s-2}^{\|a\|_1 - 2\|a\|_0 + \#a(1)} \cdot |\xi|^{2\#a(0)+\#a(1)} \leq L_{2\nu+s-2}^{\|a\|_1 - 2\|a\|_0 + \#a(1) - 2\#a(0) - \#a(1)} \leq L_{2\nu+s-2}^{s-2} \leq L_{3s}^{s-2}.$$

Using the above and that $|w| \leq 1$, we get

$$\left| \frac{d^{2\nu+s-2} V_w^\nu(\xi)}{d\xi^{2\nu+s-2}} \right| \leq \nu^{2\nu+s-2} \cdot s^{O(s)} \cdot L_{3s}^{s-2} \cdot w^{2\nu+s-2}.$$

Applying this bound in (18), we get

$$\begin{aligned} |R_2(w,\xi)| &\leq e^{-\frac{\xi^2}{2}} \cdot \sum_{\nu=1}^{s-3} \frac{1}{\nu!} \cdot \frac{|\xi|^{2\nu+s-2}}{(2\nu+s-2)!} \cdot \nu^{2\nu+s-2} \cdot s^{O(s)} \cdot L_{3s}^{s-2} \cdot w^{2\nu+s-2}, \\ &\leq e^{-\frac{\xi^2}{2}} \cdot w^s \cdot s^{O(s)} \cdot L_{3s}^{s-2} \cdot (|\xi|^s + |\xi|^{3s-8}). \end{aligned}$$

Using the bound on $|R_1(w,\xi)|$ (and using that $|w| \leq 1$), we get that

$$|R_1(w,\xi)| + |R_2(w,\xi)| \leq s^{O(s)} \cdot L_{3s}^{s-2} \cdot \left(e^{-\frac{\xi^2}{2}} \cdot (|\xi|^s + |\xi|^{3s-8}) + e^{-\frac{2\xi^2}{5}} \cdot |\xi|^{3s-6} \cdot \right)$$

Finally, plugging in $w = 1$, in the above, we complete the proof. □

3 From Fourier closeness to ℓ_1 closeness

We start with some basics of Fourier analysis.

Definition 5. *A distribution* $\mathbf{p}$ *is said to be supported on the lattice* $\mathcal{L} = \{a + b \cdot h\}_{b \in \mathbb{Z}}$ *if* $\mathsf{supp}(\mathbf{p}) \subseteq \mathcal{L}$. *If* h *is the maximum possible number such that there exists a lattice* $\mathcal{L}$ *and* $\mathsf{supp}(\mathbf{p}) \subseteq \mathcal{L}$, *then* h *is said to be the maximal span of the lattice and* a *its offset.*

We need a couple of more facts about Fourier transform of distributions.

Fact 18. *Shifting the distribution by a quantity* λ *multiplies the Fourier transform at point* ξ *by* $e^{i\xi\lambda}$.

Since our distributions will be supported on lattices, we first recall the Fourier inversion formula for probability distributions on lattices. In particular, let $\mathbf{p}$ be any distribution over $\mathcal{L}$. Then,

$$\widehat{\mathbf{p}}(\xi) = \int e^{i\xi t} d\mathbf{p}(t) = \sum_{\nu=-\infty}^{\infty} \mathbf{p}(\nu \cdot h + a) \cdot e^{i\xi(\nu \cdot h + a)}$$

As a consequence, we get

$$\frac{h}{2\pi} \cdot \int_{-\pi/h}^{\pi/h} \widehat{\mathbf{p}}(\xi) \cdot e^{-i\xi(a+\nu\cdot h)} \quad = \quad \frac{h}{2\pi} \cdot \sum_{\nu'=-\infty}^{\infty} \mathbf{p}(\nu' \cdot h + a) \cdot \int_{-\pi/h}^{\pi/h} e^{i\xi(\nu'-\nu)h} = \mathbf{p}(\nu \cdot h + a). \qquad (20)$$

We now move to stating the main theorem of this section. For this, we assume that $X_1, \ldots, X_n$ are independent centered random variables such that for $i = 1, \ldots, n$ supported on lattices of span 1. Further, let $Z = (X_1 + \ldots + X_n)/\sigma$. Note that Z is supported on a lattice with span $1/\sigma$. Call the lattice L_Z. We now state our main theorem.

Theorem 6. *Let $X_1, \ldots, X_n$ and Z be as defined above and let $z \in L_Z$. Let us assume that $I \geq 2\sqrt{s \cdot \log s}$. Then,*

$$\left| \Pr[Z = z] - \frac{1}{2\pi\sigma} \cdot \int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right| \leq \eta_{\mathsf{low}} + \eta_{\mathsf{med}} + \eta_{\mathsf{high}}$$

where

$$\eta_{\mathsf{low}} = s^{O(s)} \cdot L_{3s}^{s-2}, \quad \eta_{\mathsf{med}} = e^{-\frac{I^2}{6}} + s^{O(s)} \cdot e^{-\frac{I^2}{4}} \text{ and } \eta_{\mathsf{high}} = \sup_{|\zeta| \in [\sigma^2/\beta_3, \pi]} \left| \prod_{i=1}^{n} \widehat{X_i}(\zeta) \right| + s^{O(s)} \cdot e^{-\frac{I^2}{4}}.$$

Now, consider the function

$$\widehat{P}(\xi) = e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right).$$

Given (20), it suffices to show that $\widehat{P}(\xi)$ is appropriately close to $\widehat{Z}(\xi)$ in the interval $[-\pi/h, \pi/h]$. We do this in the following proof.

Proof of Theorem 6: We begin by noting that

$$\left| \Pr[Z = z] - \frac{1}{2\pi\sigma} \cdot \int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right|$$
$$= \left| \frac{1}{2\pi\sigma} \cdot \int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z} \cdot \widehat{Z}(\xi) \cdot d\xi - \frac{1}{2\pi\sigma} \cdot \int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right|$$

Towards bounding the quantity on the right hand side, let us divide the interval $[-\pi\sigma, \pi\sigma]$ into three parts. Define $I_{\mathsf{low}} = [-I, I]$, $I_{\mathsf{med}} = [-\frac{1}{10 \cdot L_3}, \frac{1}{10 \cdot L_3}] \setminus I_{\mathsf{low}}$, $I_{\mathsf{high}} = [-\pi\sigma, \pi\sigma] \setminus (I_{\mathsf{low}} \cup I_{\mathsf{med}})$. We control the errors in these regions separately. We define

$$\eta_{\mathsf{low}} = \left| \frac{1}{2\pi\sigma} \cdot \int_{\xi \in I_{\mathsf{low}}} e^{-i\xi z} \cdot \widehat{Z}(\xi) \cdot d\xi - \frac{1}{2\pi\sigma} \cdot \int_{\xi \in I_{\mathsf{low}}} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right|.$$

η_{med} and η_{high} are defined in an analogous manner.

Bounding in I_{low}: We begin by observing that for any $s \geq 2$,

$$\max_{\xi \in \mathbb{R}} \left(|\xi|^s + |\xi|^{(3s-8)}\right) \cdot e^{-\frac{2\xi^2}{5}} \leq s^{2s}. \qquad (21)$$

This can be easily deduced from considering two cases: $|\xi| \leq 3\sqrt{s}$ and $|\xi| > 3\sqrt{s}$. In the former case,

$$\left(|\xi|^s + |\xi|^{(3s-8)}\right) \cdot e^{-\frac{2\xi^2}{5}} \leq \left(|\xi|^s + |\xi|^{(3s-8)}\right) \leq s^{2s}.$$

In the latter case, both $|\xi|^s$ and $|\xi|^{(3s-8)}$ are bounded by $e^{\frac{2\xi^2}{5}}$ and hence

$$\left(|\xi|^s + |\xi|^{(3s-8)}\right) \cdot e^{-\frac{2\xi^2}{5}} \leq 2.$$

Armed with (21) and applying Lemma 6,

$$\begin{aligned} \left| \frac{1}{2\pi\sigma} \cdot \int_{-I}^{I} e^{-i\xi z} \cdot \widehat{Z}(\xi) \cdot d\xi - \frac{1}{2\pi\sigma} \cdot \int_{-I}^{I} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right| &\leq s^{O(s)} \cdot L_{3s}^{s-2} \cdot \frac{|I|}{\sigma} \\ &\leq s^{O(s)} \cdot L_{3s}^{s-2}. \end{aligned}$$

Bounding in I_{med}: For this, we begin by proving the following useful estimate.

Lemma 19. *Let X be a centered random variable with variance $\sigma^2 > 0$ and $\mathbf{E}[|X|^3] = \beta$. Then, for all t,*

$$\left|\mathbf{E}[e^{iXt}]\right| \leq \exp\left(\frac{-t^2 \cdot \sigma^2}{2} + \frac{|t|^3 \cdot \beta}{3}\right).$$

Proof. Consider the random variable $Z = X - X'$ where X' is an independent copy of X. Note that

$$\mathbf{E}[e^{iZt}] = \mathbf{E}[e^{iXt}] \cdot \mathbf{E}[e^{-iXt}] = |\mathbf{E}[e^{iXt}]|^2$$

Thus, for our purposes, it suffices to bound $\mathbf{E}[e^{iZt}]$. Also, the above shows that $0 \leq \mathbf{E}[e^{iZt}] \leq 1$. Using the elementary inequality,

$$\text{for all } 0 \leq x \leq 1, \quad \log x \leq x - 1,$$

we get that

$$\log \mathbf{E}[e^{iZt}] \leq \mathbf{E}[e^{iZt}] - 1. \tag{22}$$

Further, using that $\mathbf{E}[e^{iZt}]$ is real and the Taylor's theorem, we have

$$\mathbf{E}[e^{iZt}] - 1 = \mathbf{E}[\cos(Zt) - 1] \leq -\frac{t^2\mathbf{E}[Z^2]}{2} + \frac{1}{6}|t|^3\mathbf{E}[|Z|^3].$$

Combining this with (22), we have

$$\log \mathbf{E}[e^{iZt}] \leq -\frac{t^2\mathbf{E}[Z^2]}{2} + \frac{1}{6}|t|^3\mathbf{E}[|Z|^3] = -t^2\sigma^2 + \frac{1}{6}|t|^3\mathbf{E}[|Z|^3].$$

Thus, it remains to bound $\mathbf{E}[|Z|^3]$. To do this, note that

$$\mathbf{E}[|Z|^3] = \mathbf{E}[|X - X'|^3] \leq 2\mathbf{E}[|X| \cdot (X - X')^2] \leq 4\mathbf{E}[|X|^3].$$

Thus, we have

$$\log \mathbf{E}[e^{iZt}] \leq -t^2\sigma^2 + \frac{2}{3} \cdot |t|^3 \cdot \beta,$$

which finishes the proof. □

Thus, we get that that for the random variable $Z' = X_1 + \ldots + X_n$, we have

$$|\mathbf{E}[e^{iZ'\xi}]| \le e^{-\frac{\xi^2\sigma^2}{2}+\frac{|\xi|^3\beta_3}{3}},$$

and hence

$$|\mathbf{E}[e^{iZ\xi}]| \le e^{-\frac{\xi^2}{2}+\frac{|\xi|^3\cdot L_3}{3}}.$$

Thus, for $|\xi| \le L_3^{-1}$, $|\mathbf{E}[e^{iZ\xi}]| \le e^{-\frac{\xi^2}{6}}$. Thus, combining this and Corollary 12,

$$\begin{aligned} & \left| \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{med}}} e^{-i\xi z} \cdot \widehat{Z}(\xi) \cdot d\xi - \frac{1}{2\pi\sigma} \cdot \int_{-I}^{I} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right| \\ \le \;& \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{med}}} e^{\frac{-\xi^2}{6}} \cdot d\xi + \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{med}}} s^{O(s)} \cdot e^{-\frac{\xi^2}{2}} \cdot (|\xi|^3 + |\xi|^{3s}) \cdot d\xi \\ \le \;& \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{med}}} e^{\frac{-\xi^2}{6}} \cdot d\xi + \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{med}}} s^{O(s)} \cdot e^{\frac{-\xi^2}{4}} \cdot d\xi. \end{aligned}$$

The last inequality uses $I \ge 2\sqrt{s \cdot \log s}$ and hence for all $\xi \in I_{\mathsf{med}}$, $|\xi| \ge 2\sqrt{s \cdot \log s}$. This easily implies that $\eta_{\mathsf{med}} \le e^{-\frac{I^2}{6}} + s^{O(s)} \cdot e^{-\frac{I^2}{4}}$. This leaves us with bounding η_{high}.

Bounding in I_{high}:

$$\begin{aligned} & \left| \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{high}}} e^{-i\xi z} \cdot \widehat{Z}(\xi) \cdot d\xi - \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{high}}} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right| \\ \le \;& \left| \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{high}}} e^{-i\xi z} \cdot \widehat{Z}(\xi) \cdot d\xi \right| + \left| \frac{1}{2\pi\sigma} \cdot \int_{I_{\mathsf{high}}} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right|. \end{aligned}$$

The latter summand (similar to the case for η_{high}) can be bounded by $s^{O(s)} \cdot e^{-\frac{I^2}{4}}$. This leaves us with bounding the first summand. Let us call the first summand as $\mathsf{err}_{\mathsf{high}}$. Towards this, we observe that $\widehat{Z}(\xi) = \prod_{i=1}^n \widehat{X_i}(\xi/\sigma)$. As a result,

$$\mathsf{err}_{\mathsf{high}} \le \sup_{\xi \in I_{\mathsf{high}}} |\widehat{Z}(\xi)| = \sup_{|\zeta| \in [\sigma^2/\beta_3, \pi]} \left| \prod_{i=1}^n \widehat{X_i}(\zeta) \right|.$$

This finishes our proof. □

It is instructive to see the corollary of this theorem in the setting of i.i.d. centered lattice valued random variables. Towards this, assume that $X_1, \ldots, X_n$ are i.i.d. lattice valued random variables of maximal span 1 (call the common distribution X). Further, let us use $\beta_{(k)}$ to denote $\mathbf{E}[|X|^k]$ and $\alpha_{(k)}$ to denote $\mathbf{E}[X^k]$. Further, let us assume that X is $c(k)$-hypercontractive i.e. $\beta_{(k)}^{1/k} \le c(k) \cdot \alpha_{(2)}^{1/2}$.

Using this notation,

$$L_k = \left(\frac{\beta_k}{\sigma^k}\right)^{\frac{1}{k-2}} = \left(\frac{n \cdot \beta_{(k)}}{n^{\frac{k}{2}} \cdot \alpha_{(2)}^{k/2}}\right)^{\frac{1}{k-2}} = \frac{1}{\sqrt{n}} \cdot \left(\frac{\beta_{(k)}}{\alpha_{(2)}^{k/2}}\right)^{\frac{1}{k-2}} \le \frac{1}{\sqrt{n}} \cdot c(k)^{\frac{k}{(k-2)}}.$$

Now, let us define $Z = (X_1 + \ldots + X_n)/\sigma$. Note that Z lies on a lattice on a lattice with span $1/\sigma$. Let us call this lattice $\mathcal{L}$. In this setting, we have the following theorem.

Theorem 7. *Let $X_1, \ldots, X_n$ be as defined above. Finally, let us define $I = \frac{\sqrt{n}}{c(3s)^{\frac{3s}{3s-2}}}$. Then, for any $z \in \mathcal{L}$,*

$$\left| \Pr[Z = z] - \frac{1}{2\pi\sigma} \cdot \int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z} \cdot e^{-\frac{\xi^2}{2}} \cdot \left(1 + \sum_{\nu=1}^{s-3} P_\nu(i\xi)\right) d\xi \right| \leq \frac{s^{O(s)} \cdot c(3s)^{\frac{3s^2}{s-2}}}{n^{\frac{s-2}{2}}} + e^{-\frac{I^2}{6}} + s^{O(s)} \cdot e^{-\frac{I^2}{4}} + \sup_{|\zeta| \in \left[\frac{\alpha_{(2)}}{\beta_{(3)}}, \pi\right]} \left|\widehat{X}(\zeta)\right|^n$$

Proof. We simply apply Theorem 6 and evaluate the error terms. First, we evaluate η_{low} to get

$$\eta_{\mathsf{low}} = \frac{s^{O(s)} \cdot c(3s)^{\frac{3s^2}{s-2}}}{n^{\frac{s-2}{2}}}.$$

Next, we note that $I = \frac{\sqrt{n}}{c(3s)^{\frac{3s}{3s-2}}}$. Thus, we get

$$\eta_{\mathsf{med}} = e^{-\frac{I^2}{6}} + s^{O(s)} \cdot e^{-\frac{I^2}{4}}.$$

Finally, noting that

$$\sup_{|\zeta| \in [\sigma^2/\beta_3, \pi]} \left| \prod_{i=1}^{n} \widehat{X_i}(\zeta) \right| = \sup_{|\zeta| \in \left[\frac{\alpha_{(2)}}{\beta_{(3)}}, \pi\right]} \left|\widehat{X}(\zeta)\right|^n,$$

implies that $\eta_{\mathsf{high}} = \sup_{|\zeta| \in \left[\frac{\alpha_{(2)}}{\beta_{(3)}}, \pi\right]} \left|\widehat{X}(\zeta)\right|^n + s^{O(s)} \cdot e^{-\frac{I^2}{4}}$, which finishes the proof. □

To see the utility of the above theorem, consider the case when $c(s)$ is upper bounded by some polynomial function of s (this is indeed true for a large class of random variables). In this case, for $s = \tilde{o}(\sqrt{n})$, the first three error terms contribute an error of $n^{-(s-2)/2}$. The next lemma shows that the remaining error term has an exponential decay in n.

Lemma 20. *Let X be a random variable supported on a lattice with maximal span 1. For any $c > 0$, $\sup_{\pi \geq |\zeta| \geq c} |\widehat{X}(\zeta)| < 1$.*

Proof. Note that if X is supported on the lattice $\mathcal{L}$ (with maximal span 1),

$$\widehat{X}(\zeta) = \sum_{\nu=-\infty}^{\infty} \Pr[X = \nu + a] \cdot e^{i \cdot (\nu + a) \cdot \zeta}.$$

Note that if $|\widehat{X}(\zeta)| = 1$, then $e^{i \cdot (\nu + a) \cdot \zeta}$ must be the same for all points ν such that $\Pr[X = \nu + a] \neq 0$. Let A_ν be the set of all such ν. Then, we immediately get that for $\nu_1, \nu_2 \in A_\nu$, $e^{i \cdot (\nu_1 - \nu_2) \cdot \zeta} = 1$. Consider the set $A_Z = \{\nu_1 - \nu_2 : \nu_1, \nu_2 \in A_\nu\}$. Since the maximal span of the lattice is 1, it means that there is some integral linear combination of the elements in A_ν which is 1. However, that means that $e^{1 \cdot \zeta} = 1$ which contradicts our assumption on ζ. □

Thus, the above lemma shows that $\sup_{|\zeta| \in \left[\frac{\alpha_{(2)}}{\beta_{(3)}}, \pi\right]} \left|\widehat{X}(\zeta)\right|^n$ is inverse exponential in n. Of course, the exact exponent in the exponential decay (which is critical) can be significantly different depending up on the particular random variable involved. As we will show, in our main application (even for the non i.i.d. case), the last error term is indeed negligible.

4 Derandomization for combinatorial shapes

We begin by recalling the main theorem of this section.

Theorem. *There is a polynomial time computable PRG $G_{csum} : \{0,1\}^{t_{\mathsf{csum}}} \to [m]^n$ which ϵ-fools $\mathsf{Csum}(m,n)$. Here $t_{\mathsf{csum}} = O(\log m + \log^{3/2}(n/\epsilon))$.*

The proof of this theorem proceeds by proving the following two results.

Claim 21. *Let $G_{csum,poly} : \{0,1\}^{t_{\mathsf{csum,poly}}} \to [m]^n$ be a polynomial time computable PRG which ϵ-fools $\mathsf{Csum}(m,n)$ for $m = O((n/\epsilon)^3)$ with $t_{\mathsf{csum,poly}} = O(\log^{3/2}(n/\epsilon))$. Then, there is a polynomial time computable PRG $G_{csum} : \{0,1\}^{t_{\mathsf{csum}}} \to [m]^n$ which ϵ-fools $\mathsf{Csum}(m,n)$ where $t_{\mathsf{csum}} = O(\log m + \log^{3/2}(n/\epsilon))$.*

Theorem 8. *There is a polynomial computable PRG $G_{csum,poly} : \{0,1\}^{t_{\mathsf{csum,poly}}} \to [m]^n$ which ϵ-fools $\mathsf{Csum}(m,n)$ for $m = O((n/\epsilon)^3)$ with $t_{\mathsf{csum,poly}} = O(\log^{3/2}(n/\epsilon))$.*

Clearly, by combining Claim 21 with Theorem 8, we get the main theorem. Thus, we focus on proving Claim 21 and Theorem 8.

We can associate a n-tuple of independent Bernoulli random variables $(Y_{f,1}, \ldots, Y_{f,n})$ where $Y_{f,i} \sim f_i(y_i)$ where $y_i \sim U_m$. We adopt the convention that given a subset $B \subseteq [n]$, we use $Z_{f,B} = \sum_{i \in B} f_i(y_i)$. We use Z_f to denote the random variable when $B = [n]$.

Proof of Claim 21. Assume that $m \gg (n/\epsilon)^3$ (else we are already done). Set $m_0 = A_1 \cdot (n/\epsilon)^3$ where A_1 is a sufficiently large constant. Further, also assume that A_1 is a prime. For simplicity, we assume that m is a power of m_0. Let $p_0 = m_0/m$ and set $\delta = \epsilon/100$. Let $\mathcal{H}_{2,m,p_0}$ be the family of p_0-biased pairwise independent hash functions instantiated using Lemma 43. Note that for any $h \sim \mathcal{H}_{2,m,p_0}$, $|h^{-1}(1)| = m_0$. Let $G_{csum,poly} : \{0,1\}^{t_{\mathsf{csum,poly}}} \to [m_0]^n$ δ-fool the class $\mathsf{Csum}(m_0, n)$ (instantiated using Theorem 8. Then, we let $G_{csum} : \mathcal{H}_{2,m,p_0} \times \{0,1\}^{t_{\mathsf{csum,poly}}} \to [m]^n$ as follows:

$$G_{csum}(h, x) = y' \text{ where } y'_{h^{-1}(1)} = G_{csum,poly}(x).$$

Note that the seed required to sample and element of $\mathcal{H}_{2,m,p_0}$ is $O(\log m)$ and thus combining with the bound on $t_{\mathsf{csum,poly}}$ from Theorem 8, we get the stated bound on t_{csum}. The bound on the quality of G_{csum} remains to be proven. Toward this, we have the following claim.

Claim 22. *Fix any $i \in [n]$.*

$$\Pr_{h \in \mathcal{H}_{2,m,p_0}} \left[\| f_i(U_{h^{-1}(1)}) - f_i(U_{[m]}) \|_1 \le \frac{\epsilon}{8n} \right] \le \frac{\epsilon}{16n}.$$

Proof. Consider $h \sim \mathcal{H}_{2,m,p_0}$. We define random variables $\{A_{j,h}\}_{j=0,1}$ where $A_{j,h} = |h^{-1}(1) \cap f_i^{-1}(j)|$. Also, let $\{p'_j\}_{j=0,1}$ be defined as $p'_j = \Pr_{x \in U_{[m]}}[f_i(x) = j]$. Note that for our purposes, it suffices to show with probability $1 - \epsilon/(8n)$,

$$\left| p'_0 - \frac{|A_{0,h}|}{m_0} \right| \le \frac{\epsilon}{16n} \text{ and } \left| p'_1 - \frac{|A_{1,h}|}{m_0} \right| \le \frac{\epsilon}{16n}. \tag{23}$$

Further, note that showing any one of the above implies the other one as well. Assume without loss of generality, $1/2 \le p'_0 \le 1$. Let us define an indicator random variable $X^{(j,h)}$ as follows:

$$X^{(j,h)} = 1 \text{ if and only if } f_i(j) = 0 \text{ and } h(j) = 1.$$

Note that $|A_{0,h}| = \sum X^{(j,h)}$. Further, because $h \sim \mathcal{H}_{2,m,p_0}$, $\{X^{(j,h)}\}$ are pairwise independent random variables. Thus, $\mathbf{E}[|A_{0,h}|] = m_0 \cdot p_0'$ and $\mathrm{Var}(|A_{0,h}|) \leq m_0 \cdot p_0'$. By using Lemma 36,

$$\Pr_{h \in \mathcal{H}_{2,m,p_0}}\left[\left||A_{0,h}| - m_0 \cdot p_0'\right| \leq 4 \cdot \sqrt{\frac{n \cdot m_0 \cdot p_0'}{\epsilon}}\right] \geq 1 - \frac{\epsilon}{16n}.$$

Since $m_0 = A(n/\epsilon)^3$, by making A sufficiently large, we get that with probability $1 - \epsilon/(16n)$ for $h \sim \mathcal{H}_{2,m,p_0}$, (23) holds. This finishes the proof. □

Let $(x_1, \ldots, x_n) \sim U^n_{h^{-1}(1)}$ and $(y_1, \ldots, y_n) \sim U^n_{[m]}$. Let us define $h \sim \mathcal{H}_{2,m,p_0}$ to be *good* if

$$\left\|\sum_i f_i(x_i) - \sum_i f_i(y_i)\right\|_1 \leq \frac{\epsilon}{8}.$$

By using Claim 22, we get that $h \sim \mathcal{H}_{2,m,p_0}$ is good with probability at least $1 - \epsilon/16$. Consider such a h. Fix a canonical permutation $\pi_h : h^{-1}(1) \to [m_0]$. Define the distribution $D_h' = G_{csum}(h, U_{t_{\mathsf{csum,poly}}})$ and $D' = G_{csum,poly}(U_{t_{\mathsf{csum,poly}}})$. Further, we also define the function $g_h : [m_0]^n \to [n]$ as $g_h(z_1, \ldots, z_n) = \sum f_i(\pi_h^{-1}(z_i))$. Let $(y_1', \ldots, y_n') \sim D_h'$ and $(z_1', \ldots, z_n') \sim D'$

$$\sum_i f_i(y_i') = \sum_i g_{h,i}(z_i'). \tag{24}$$

In the above, the equality denotes the equality of the two distributions. Let $(z_1, \ldots, z_n) \sim U^n_m$. Since $g_h \in \mathsf{Csum}(m_0, n)$, hence

$$\left\|\sum_i g_{h,i}(z_i') - \sum_i g_{h,i}(z_i)\right\|_1 \leq \delta. \tag{25}$$

Finally, since h is good, we can reinterpret the guarantee as

$$\left\|\sum_i g_{h,i}(z_i) - \sum_i f_i(y_i)\right\|_1 \leq \frac{\epsilon}{8}.$$

Combining this with (24) and (25), we get that if h is good, then

$$\left\|\sum_i f_i(y_i') - \sum_i f_i(y_i)\right\|_1 \leq \frac{\epsilon}{8} + \delta.$$

Since h is good with probability $1 - \epsilon/16$, we get that G_{csum} is ϵ-fools the class $\mathsf{Csum}(m, n)$. □

4.1 Derandomizing combinatorial sums when $m = \mathsf{poly}(n/\epsilon)$

We assume that $\epsilon = n^{-\Omega(1)}$. We also assume that $\epsilon = n^{-o(\log n)}$. This is because Theorem 1 guarantees derandomization of $\mathsf{Csum}(m, n)$ with seed length $O(\log m + \log^2 n + \log n \cdot \log(1/\epsilon))$. Thus, for $\epsilon = n^{-\Omega(\log n)}$, the guarantee of Theorem 8 is guaranteed by Theorem 1. We partition $\mathsf{Csum}(m, n)$ into two classes depending upon $\mathrm{Var}(Z_f)$. Let $A_1 > 0$ to be a sufficiently large constant. We let $\delta = 2^{-A_1 \cdot \sqrt{\log(1/\epsilon)}}$. Next, we define

$$\mathsf{Csum}_{\mathsf{low}}(m, n) = \{f \in \mathsf{Csum}(m, n) : \mathrm{Var}(Z_f) \leq \delta^{-1}\} \qquad \text{and}$$
$$\mathsf{Csum}_{\mathsf{high}}(m, n) = \{f \in \mathsf{Csum}(m, n) : \mathrm{Var}(Z_f) > \delta^{-1}\}.$$

Using Claim 45, we get that fooling $\mathsf{Csum}_{\mathsf{low}}(m, n)$ and $\mathsf{Csum}_{\mathsf{high}}(m, n)$ individually is sufficient to fool $\mathsf{Csum}(m, n)$.

5 Fooling small combinatorial sums

For this section, the strategy is as follows. We first describe a PRG $G_{\mathsf{low-easy}}$ which ϵ-fools $\mathsf{Csum}_{\mathsf{low}}(m,n)$ albeit with a poor seed length. The seed length is then reduced to the target length by a straightforward application of Theorem 1. Choose $t = \delta^{-10}$ and $k = c\sqrt{\log(1/\epsilon)}$ for some sufficiently large constant c. Let $\mathcal{H}_{k,n,t}$ be the family of hash functions from Lemma 41. For convenience, assume that for all $i \in [t]$, the size of $|h^{-1}(i)|$ is the same for all $i \in [t]$. Let $G_{3k+2,n/t,m} : \{0,1\}^{t_2} \to [m]^{n/t}$ be the $2+3k$-wise independent generator from Lemma 42. We define PRG $G_{\mathsf{low-easy}} : \mathcal{H}_{k,n,t} \times (\{0,1\}^{t_2})^t \to [m]^n$ defined as follows:

$$G_{\mathsf{low-easy}}(h, x_1, \ldots, x_t) = y' \text{ where } y'_{h^{-1}(i)} = G_{3k+2,n/t,m}(x_i).$$

The main lemma showing the quality of this PRG is the following:

Lemma 23. $G_{\mathsf{low-easy}}$ *is an* ϵ*-PRG for* $\mathsf{Csum}_{\mathsf{low}}(m,n)$.

Toward proving this lemma, we first partition $[n]$ into two sets: L and H which are defined as

- $L = \{i \in [n] : \mathrm{Var}(Y_{f,i}) \le \delta^5\}$.
- $H = \{i \in [n] : \mathrm{Var}(Y_{f,i}) > \delta^5\}$.

Note that $|H| \le O(1/\delta^6)$. Next, we define the random variable $\mathsf{Count}_{H,i}$ as follows:

$$\mathsf{Count}_{H,i} = \{j \in [n] : h(j) = i \text{ and } j \in H\}$$

The next claim controls the maximum of $\mathsf{Count}_{H,i}$.

Claim 24.

$$\Pr\left[\max_{i\in[t]} \mathsf{Count}_{H,i} \ge k\right] \le \delta^{3k}.$$

Proof. Choose any particular $i \in [t]$. We first bound the probability $\Pr[\mathsf{Count}_{H,i} \le k]$ and then apply a union bound. For every $\ell \in H$, we define an indicator random variable $\mathbf{I}_\ell$ which is 1 if and only if $h(\ell) = i$. Note that the random variables $\{\mathbf{I}_\ell\}_{\ell\in H}$ are k-wise independent and $\Pr[\mathbf{I}_\ell = 1] = 1/t = \delta^{10}$ for all $\ell \in H$.

$$\begin{aligned}\Pr\left[\mathsf{Count}_{H,i} \ge k\right] &\le \sum_{(u_1,\ldots,u_k)\in H} \Pr_{h\in\mathcal{H}_{k,n,t}}\left[\wedge_{j=1}^k \mathbf{I}_{u_j} = 1\right] \\ &\le |H|^k \cdot \frac{1}{t^k} \le \delta^{4k}.\end{aligned}$$

Thus, applying a union bound, we get that $\Pr\left[\max_{i\in[t]} \mathsf{Count}_{H,i} \ge k\right] \le t \cdot \delta^{4k}$. Noting that $k \ge 10$, we get the claim. □

We next define the quantity $\mathsf{Var}_{L,i}$ as follows:

$$\mathsf{Var}_{L,i} = \mathrm{Var}\bigg(Z_{f,h^{-1}(i)\cap L}\bigg)$$

The next claim controls the value of $\mathsf{Var}_{L,i}$ in the "buckets" defined by the map h.

Claim 25.

$$\Pr_{h\in\mathcal{H}_{k,n,t}}\left[\max_{i\in t} \mathsf{Var}_{L,i} > k\cdot\delta\right] \le \delta^{3k}.$$

Proof. We fix an $i \in [t]$ and the bound the probability that $\mathsf{Var}_{L,i} > k \cdot \delta$. Towards this, note that

$$\mathsf{Var}_{L,i} = \sum_{j \in L} \mathrm{Var}(Y_{f,j}) \cdot \mathbf{I}_j$$

where $\mathbf{I}_j$ is an indicator random variable which is 1 if and only if $h(j) = i$. Observe that for $1 \leq j \leq n$, $\mathbf{E}[\mathbf{I}_j] = \delta^{10}$. To bound the growth of $\mathsf{Var}_{L,i}$, we consider bounding its k^{th} moment. Toward this, we consider the expansion of the variable $\mathsf{Var}_{L,i}^k$.

$$\mathsf{Var}_{L,i}^k = \left(\sum_{j \in L} \mathrm{Var}(Y_{f,j}) \cdot \mathbf{I}_j \right)^k.$$

Note that in the formal expansion of the term on the right, we will get monomials in the variables $\{\mathbf{I}_j\}_{j \in L}$. Toward bounding the expectation of $\mathsf{Var}_{L,i}^k$, we define the notion of signature of a monomial. Consider a monomial $\prod_{j \in L} (\mathrm{Var}(Y_{f,j}) \cdot \mathbf{I}_j)^{a_j}$. Let $\overline{a} = (a_1, \ldots, a_L) \in \mathbb{Z}^L$. Note that for any such monomial, $\sum_{j \in L} a_j = k$ and for all $j \in L$, $a_j \geq 0$. As a corollary, $\overline{a}$ has at most k non-zero entries. The signature of the monomial is given by arranging the vector $\overline{a}$ in decreasing order and truncating the zero part of the vector. Consider any signature vector $\sigma = (v_1, \ldots, v_r)$ where $\sum v_i = k$, $v_1 \geq \ldots \geq v_r > 0$. Let us define S_σ as

$$S_\sigma = \mathbf{E}\Big[\sum_{\overline{a}:\mathsf{sign}(\overline{a})=\sigma} \prod_{j \in L} (\mathrm{Var}(Y_{f,j}) \cdot \mathbf{I}_j)^{a_j} \Big]$$

$$\begin{aligned}
S_\sigma &\leq \binom{k}{v_1 \; \ldots \; v_r} \prod_{i=1}^{r} \Big(\sum_{j \in L} \mathbf{E}\left[(\mathrm{Var}(Y_{f,j}) \cdot \mathbf{I}_j)^{v_i} \right] \Big), \\
&\leq \binom{k}{v_1 \; \ldots \; v_r} \prod_{i=1}^{r} \Big(\sum_{j \in L} \mathbf{E}\left[\mathrm{Var}(Y_{f,j})^{v_i} \cdot \mathbf{I}_j \right] \Big), \\
&\leq \binom{k}{v_1 \; \ldots \; v_r} \prod_{i=1}^{r} \delta^{5(v_i-1)} \cdot \Big(\sum_{j \in L} \mathbf{E}\left[\mathrm{Var}(Y_{f,j}) \cdot \mathbf{I}_j \right] \Big), \text{ (using } \mathrm{Var}(Y_{f,i}) \leq \delta^5) \\
&\leq \binom{k}{v_1 \; \ldots \; v_r} \prod_{i=1}^{r} \delta^{5(v_i-1)} \cdot \delta^9 = \binom{k}{v_1 \; \ldots \; v_r} \cdot \delta^{5k+4r} \leq 2^k \cdot \delta^{5k+4}.
\end{aligned}$$

Next, we note that the total number of possible signatures is at most 2^k, and hence we get,

$$\mathsf{Var}_{L,i}^k = \left(\sum_{j \in L} \mathrm{Var}(Y_{f,j}) \cdot Z_j \right)^k \leq 4^k \cdot \delta^{5k+4}.$$

Noting that $\mathsf{Var}_{L,i}$ is a non-negative random variable, applying Markov's inequality, we get that

$$\Pr_{h \in \mathcal{H}_{k,n,t}} [\mathsf{Var}_{L,i} > k \cdot \delta] \leq \frac{4^k \cdot \delta^{5k+4}}{k^k \cdot \delta^k}.$$

$$\Pr_{h \in \mathcal{H}_{k,n,t}} [\max_{i \in [t]} \mathsf{Var}_{L,i} > k \cdot \delta] \leq \frac{4^k \cdot \delta^{5k-6}}{k^k \cdot \delta^k} \leq \delta^{3k},$$

where the last inequality uses that $k \geq 10$. □

We next define a $h \in \mathcal{H}_{k,n,t}$ to be *good* if

- $\max_{i \in [t]} \mathsf{Count}_{H,i} < k$.
- $\max_{i \in [t]} \mathsf{Var}_{L,i} < k \cdot \delta$.

Using Claim 24 and Claim 25, we get that $h \in \mathcal{H}_{k,n,t}$ is *good* with probability $1 - 2 \cdot \delta^{3k}$. We next recall the following useful lemma from [GMRZ13].

Lemma 26. *Let $\{X_i\}_{i=1}^n$ and $\{Y_i\}_{i=1}^n$ be (jointly) independent Bernoulli random variables and let $X = \sum X_i$ and $Y = \sum Y_i$. Let $\{X_i'\}_{i=1}^n$ and $\{Y_i'\}_{i=1}^n$ be (jointly) $2k+2$-wise independent random variables such that for all $i \in [n]$, $\mathbf{E}[X_i] = \mathbf{E}[X_i']$ and likewise, $\mathbf{E}[Y_i] = \mathbf{E}[Y_i']$. Let $X' = \sum X_i$ and $Y' = \sum Y_i$. If $\mathbf{E}[X] \le \eta$ and $\mathbf{E}[Y] \le \eta$, then, $\|(X,Y) - (X',Y')\|_1 \le e^k \cdot \eta^k$.*

Following [GMRZ13], we can get the following claim.

Claim 27. *Let $\{X_i\}_{i=1}^n$ be independent Bernoulli random variables and $\{X_i'\}_{i=1}^n$ be a set of k-wise independent Bernoulli random variables such that $\mathbf{E}[X_i] = \mathbf{E}[X_i']$. If $X = \sum X_i$ and $X' = \sum X_i'$ such that* $\mathrm{Var}(X) \le \eta$, *then $\|X - X'\|_1 \le (2 \cdot e)^k \cdot \eta^k$.*

Proof. Let us define $A = \{i \in [n] : \mathbf{E}[X_i] \le 1/2\}$. We observe that

$$2 \cdot \mathrm{Var}\bigg(\sum_{i \in A} X_i\bigg) \ge \mathbf{E}\bigg[\sum_{i \in A} X_i\bigg] \quad \text{and} \quad 2 \cdot \mathrm{Var}\bigg(\sum_{i \in \overline{A}} X_i\bigg) \ge \mathbf{E}\bigg[\sum_{i \in \overline{A}} (1 - X_i)\bigg].$$

By Lemma 26,

$$\left\| \bigg(\sum_{i \in A} X_i, \sum_{i \in \overline{A}} (1 - X_i)\bigg) - \bigg(\sum_{i \in A} X_i', \sum_{i \in \overline{A}} (1 - X_i')\bigg) \right\|_1 \le (2e)^k \cdot \eta^k.$$

Thus, we get that

$$\left\| \sum_{i=1}^n X_i, \sum_{i=1}^n X_i' \right\|_1 \le (2e)^k \cdot \eta^k.$$

□

The next claim extends the above result by allowing a small number of arbitrary Boolean random variables.

Claim 28. *Let $\{X_i\}_{i=1}^n$ be independent Bernoulli random variables and $\{X_i'\}_{i=1}^n$ be a set of $k+k'$-wise independent Bernoulli random variables such that for $i \in [n]$, $\mathbf{E}[X_i] = \mathbf{E}[X_i']$. Let $X = \sum X_i$, $X' = \sum X_i'$ and for $S \subseteq [n]$, define $X_S = \sum_{i \in S} X_i$. If $|S| = n - k'$ and* $\mathrm{Var}(X_S) \le \eta$, *then $\|X - X'\|_1 \le (2e)^k \cdot \eta^k$.*

Proof. Without loss of generality, assume that $S = \{k'+1, \ldots, n\}$. For any $y \in \{0,1\}^{k'}$, observe that the conditional distribution

$$(X_{k'+1}', \ldots, X_n') | X_1' = y_1, \ldots, X_{k'}' = y_{k'}$$

is k-wise independent and for all $i \in [k'+1 \ \ldots \ n]$, the marginal of X_i' in the conditional distribution is the same as the marginal of X_i. Applying Claim 27, we get that for any $y \in \{0,1\}^{k'}$

$$\left\| \sum_{i=k'+1}^n X_i - \sum_{i=k'+1}^n X_i' | X_1' = y_1, \ldots, X_{k'}' = y_{k'} \right\|_1 \le (2e)^k \cdot \eta^k.$$

Thus,

$$\left\| \sum_{i=1}^n X_i | X_1 = y_1, \ldots, X_{k'} = y_{k'} - \sum_{i=1}^n X_i' | X_1' = y_1, \ldots, X_{k'}' = y_{k'} \right\|_1 \le (2e)^k \cdot \eta^k.$$

Since the distribution on $\{X_i'\}_{i=1}^{k'}$ is k'-wise independent, we get the claim. □

Proof of Lemma 23: Fix a *good* h and $i \in [t]$. Let $x \in U_{t_2}$, $y'_{h^{-1}(i)} \sim G_{3k+2,n/t,m}(x)$ and $y_{h^{-1}(i)} \sim U_{[m]^{n/t}}$. Applying Claim 28 along with Claim 24 and Claim 25,

$$\left\| \sum_{j \in h^{-1}(i)} f_j(y_j) - \sum_{j \in h^{-1}(i)} f_j(y'_j) \right\|_1 \leq (2 \cdot e)^k \cdot (k \cdot \delta)^k. \tag{26}$$

Using the fact that the distributions of $(y_{h^{-1}(1)}, \ldots, y_{h^{-1}(t)})$ are mutually independent as are the distributions of $(y'_{h^{-1}(1)}, \ldots, y'_{h^{-1}(t)})$, we get that for any good h,

$$\left\| \sum_{j \in [n]} f_j(y_j) - \sum_{j \in [n]} f_j(y'_j) \right\| \leq (2 \cdot e)^k \cdot (k \cdot \delta)^k \cdot t. \tag{27}$$

Since the total fraction of *good* h is at least $1 - 2 \cdot \delta^{3k}$,

$$\left\| \sum_{j \in [n]} f_j(y'_j) - \sum_{j \in [n]} f_j(y_j) \right\| \leq (2 \cdot e)^k \cdot (k \cdot \delta)^k \cdot t + 2 \cdot \delta^{3k} \leq \epsilon. \tag{28}$$

□

The only problem with $G_{\mathsf{low-easy}}$ is the prohibitively large requirement in terms of seed length. However, observe that for any fixed h, the function $f_h : (\{0,1\}^{t_2})^t \to \mathbb{Z}$ given by

$$f(x_1, \ldots, x_t) = \sum_{i \in [n]} f_i(y'_i) \quad \text{where for } j \in [t],\ y'_{h^{-1}(j)} = G_{3k+2,n/t,m}(x_i).$$

Thus, f_h can be computed by a (n, t_2, t)-ROBP and hence we can use Theorem 1 to get a PRG for f_h. The seed length requirement will be $O(t_2 + (\log(n/\epsilon)) \cdot \log t) = O(\log^{3/2}(1/\epsilon))$ (where we use that $\log(1/\epsilon) = \Omega(\log n)$. Thus, we have the following theorem.

Theorem 9. *There is a polynomial time computable PRG $G_{\mathsf{low}} : \{0,1\}^{t_{\mathsf{low}}} \to [m]^n$ which ϵ-fools $\mathsf{Csum}_{\mathsf{low}}(m,n)$. Here $t_{\mathsf{low}} = O(\log m + \log^{3/2}(n/\epsilon))$.*

6 Fooling large combinatorial sums

As before, our strategy is to first describe a PRG $G_{\mathsf{high-easy}}$ which fools $\mathsf{Csum}_{\mathsf{high}}(m,n)$ with a poor seed length. We then achieve the desired seed length by applying Theorem 1. We now define the construction of the PRG $G_{\mathsf{high-easy}}$. We let $t = \delta^{-1/2}$ and $k = c\sqrt{\log(1/\epsilon)}$ for a sufficiently large constant c. Further, all the calculations are made assuming n is sufficiently large (compared to all the other constants). In particular, we will assume that $\delta^{k/8} \leq \epsilon$. Further, we will also assume that $k \geq 10$. Let $\mathcal{H}_{4k,n,t}$ be the family of hash functions from Theorem 42. As before, for convenience, we assume that for any $h \in \mathcal{H}_{4k,n,t}$ and $i \in [t]$, $|h^{-1}(i)| = n/t$. Let $G_{4k,n/t,m} : \{0,1\}^{t_2} \to [m]^{n/t}$ be the $4k$-wise independent generator from Theorem 42. Let $G^{(cs)} : \{0,1\}^{t_3} \to [m]^{n/t}$ be the PRG obtained from [GMRZ13] with ϵ set to a sufficiently small constant. The PRG $G_{\mathsf{high\text{-}easy}} : \mathcal{H}_{k,n,t} \times (\{0,1\}^{t_2})^t \times (\{0,1\}^{t_3})^t \to [m]^n$ is described as

$$G_{\mathsf{high\text{-}easy}}(h, x_1, \ldots, x_t, z_1, \ldots, z_t) = y' \text{ where } y'_{h^{-1}(i)} = G_{k,n/t,m}(x_i) \oplus_m G^{(cs)}(z_i)$$

We also define $G_{\mathsf{uniform}} : \mathcal{H}_{k,n,t} \times ([m]^{n/t})^t \to [m]^n$ as

$$G_{\mathsf{uniform}}(h, x_1, \ldots, x_t) = y \text{ where } y_{h^{-1}(i)} = x_i.$$

Note that when the input is a uniform random element from the domain of G_{uniform}, then the output distribution of G_{uniform} is identical to the uniform distribution on $[m]^n$. However, using this way to sample from $U_{[m]^n}$ allows for an easier comparison with the output of $G_{\mathsf{high\text{-}easy}}$. Also, for notational brevity, for this section, unless mentioned otherwise, y and y' are shorthands for

$$y \sim G_{\mathsf{high\text{-}easy}}(h, x_1, \ldots, x_t, z_1, \ldots, z_t), \quad y \sim G_{\mathsf{uniform}}(h, x_1, \ldots, x_t).$$

The main claim concerning the quality of $G_{\mathsf{high\text{-}easy}}$ is the following.

Claim 29.

$$\Big\| \sum_{j \in [n]} f_j(y_j) - \sum_{j \in [n]} f_j(y'_j) \Big\|_1 \leq \epsilon.$$

Toward proving Claim 29, we make a few more definitions. Let $\mu_j = \mathbf{E}_{x \in U_{[m]}}[f_j(x)]$. Also, for $j \in [n]$, define $f'_j : [m] \to \mathbb{R}$ as $f'_j(x) = f_j(x) - \mu_j$. Note that for any j, y_j and y'_j are uniform on $[m]$. Thus, $f'_j(y_j)$ and $f'_j(y'_j)$ are centered random variables. Further,

$$\Big\| \sum_{j \in [n]} f_j(y_j) - \sum_{j \in [n]} f_j(y'_j) \Big\|_1 = \Big\| \sum_{j \in [n]} f'_j(y_j) - \sum_{j \in [n]} f'_j(y'_j) \Big\|_1.$$

Thus to prove Claim 29, it suffices to show that

$$\Big\| \sum_{j \in [n]} f'_j(y_j) - \sum_{j \in [n]} f'_j(y'_j) \Big\|_1 \leq \epsilon.$$

Further, for any set $B \subseteq [n]$, define the random variables U_B and U'_B as follows:

$$U_B = \sum_{j \in B} f'_j(y_j), \quad U'_B = \sum_{j \in B} f'_j(y'_j).$$

Thus, proving Claim 29 is equivalent to showing that

$$\|U_{[n]} - U'_{[n]}\|_1 \leq \epsilon.$$

Further, observe that for any $j \in [n]$, $\mathrm{Var}(f'_j(y_j)) = \mathrm{Var}(f_j(y_j))$. We next have the following claim.

Claim 30. *Let $\mathsf{Var\text{-}tot} = \sum_{j \in [n]} \mathrm{Var}(f_j(y_j))$. Then,*

$$\Pr_{h \in \mathcal{H}_{k,n,t}} \left[\max_{i \in [t]} \left| \mathrm{Var}(U_{h^{-1}(i)}) - \mu \right| \leq \frac{\mu}{2} \right] \leq \delta^{k/4},$$

where $\mu = \mathsf{Var\text{-}tot}/t$.

Proof. Consider a fixed $i \in [t]$. Then, note that $\mathrm{Var}(U_{h^{-1}(i)}) = \sum_{j \in [n]} \mathrm{Var}(f_j(y_j)) \cdot \mathbf{I}_j$ where $\mathbf{I}_j$ is an indicator random variable which is 1 if and only if $h(j) = i$. Note that the random variable $\{\mathbf{I}_j\}_{j \in [n]}$ and hence $\{\mathrm{Var}(f_j(y_j)) \cdot \mathbf{I}_j\}_{j \in [n]}$ are $4k$-wise independent. We observe that using linearity of expectation, $\mathbf{E}[\mathrm{Var}(U_{h^{-1}(i)})] = \mu$. By applying the bounds from Theorem 37, we get

$$\Pr_{h \in \mathcal{H}_{k,n,t}} \left[\left| \mathrm{Var}(U_{h^{-1}(i)}) - \mu \right| \leq \frac{\mu}{2} \right] \leq \left(4 \cdot \frac{4k\mu + 16k^2}{\mu^2} \right)^{2k}$$

Note that $\mu = \mathsf{Var\text{-}tot}/t \geq \delta^{-1/2}$. As $k = O(\log(1/\delta))$, we get that the quantity on the right hand side can be bounded by $\delta^{k/2}$. By taking an union bound over all $i \in [t]$ and $t = \delta^{-1/2}$, we get the stated bound. □

Define $h \sim \mathcal{H}_{k,n,t}$ to be *good* if

$$\max_{i \in [t]} \left| \mathrm{Var}(U_{h^{-1}(i)}) - \mu \right| \leq \frac{\mu}{2}.$$

We next make several simple but important claims. We leave some of the proofs as they are straightforward to verify.

Claim 31. *For any fixed h, the random variables $\{U_{h^{-1}(i)}\}_{i \in [t]}$ and $\{U'_{h^{-1}(i)}\}_{i \in [t]}$ are independent.*

Claim 32. *For any fixed h, the following holds:*

- *The random variables $U_{[n]}$ and $U'_{[n]}$ are supported on a lattice $\mathcal{L}$ of span 1 with $2n+1$ support points.*
- *For any $\ell \leq 4k$ and $i \in [t]$, $\mathbf{E}[U'^{\ell}_{h^{-1}(i)}] = \mathbf{E}[U^{\ell}_{h^{-1}(i)}]$.*
- *As a consequence, for any $\ell \leq 4k$, $\mathbf{E}[U'^{\ell}_{[n]}] = \mathbf{E}[U^{\ell}_{[n]}]$.*

Based on Claim 31 and Claim 32, we use the following notation. Roughly, the convention we will stick to is to use a notion with a prime for random variables generated using the output of the $G_{\textsf{high-easy}}$ and use unprimed notation for the analogous random variables using the output of $G_{\textsf{uniform}}$.

- For $0 \leq \ell \leq 4k$ and $i \in [t]$, $\alpha_{\ell,i,(h)} = \mathbf{E}[U^{\ell}_{h^{-1}(i)}] = \mathbf{E}[U'^{\ell}_{h^{-1}(i)}]$.
- For $0 \leq \ell \leq 4k$, $\alpha_{\ell,(h)} = \mathbf{E}[U^{\ell}_{[n]}] = \mathbf{E}[U'^{\ell}_{[n]}]$.
- For $0 \leq \ell \leq 4k$ and $i \in [t]$, $\beta_{\ell,i,(h)} = \mathbf{E}[|U_{h^{-1}(i)}|^{\ell}]$ and $\beta'_{\ell,i,(h)} = \mathbf{E}[U'^{\ell}_{h^{-1}(i)}]$. If ℓ is even, then note that $\alpha_{\ell,i,(h)} = \beta_{\ell,i,(h)} = \beta'_{\ell,i,(h)}$.
- For any $0 \leq \ell \leq 4k$, $\beta_{\ell,(h)} = \sum_{i \in [t]} \mathbf{E}[U^{\ell}_{h^{-1}(i)}]$ and $\beta'_{\ell,(h)} = \sum_{i \in [t]} \mathbf{E}[U'^{\ell}_{h^{-1}(i)}]$.
- For any $i \in [t]$, we let $\sigma^2_{i,(h)} = \mathrm{Var}(U_{h^{-1}(i)}) = \mathrm{Var}(U'_{h^{-1}(i)})$. Likewise, $\sigma^2 = \mathrm{Var}(U_{[n]}) = \mathrm{Var}(U'_{[n]})$.

Claim 33. *Fix any h. For any $0 \leq |\xi| \leq \pi$ and $i \in [t]$,*

$$\left| \widehat{U_{h^{-1}(i)}}(\xi) \right| \leq \left(-\frac{9 \cdot \sigma_i^2 \cdot \xi^2}{200} \right).$$

Proof of Claim 29: We start by fixing a good h. Note that by Claim 30, h is good with probability at least $1 - \delta^{k/4} \geq 1 - \epsilon^2$. Thus, to prove the main claim, it suffices to show that for a good h, $\|U_{[n]} - U'_{[n]}\| \leq \epsilon/2$. For any fixed h, note that $U_{[n]} = \sum_{i \in [t]} U_{h^{-1}(i)}$ and $U'_{[n]} = \sum_{i \in [t]} U'_{h^{-1}(i)}$. Hence, using Claim 31 and Claim 32, both of them are sums of independent lattice valued random variables and are supported on a lattice $\mathcal{L}$ with $2n+1$ support points. Thus, it suffices to show that for every z such that $\sigma \cdot z \in \mathcal{L}$,

$$\left| \Pr\left[\frac{U_{[n]}}{\sigma} = z \right] - \Pr\left[\frac{U'_{[n]}}{\sigma} = z \right] \right| \leq \frac{\epsilon}{2n}.$$

To bound the difference, we apply Theorem 5 (provided the hypothesis for applying the theorem is satisfied). Let us choose s to be the largest even integer such that $3s \leq 4k$. Let us define

$$L_{\ell,(h)} = \left(\frac{\beta_{\ell,(h)}}{\sigma^{\ell}} \right)^{\frac{1}{\ell-2}}, \quad L'_{\ell,(h)} = \left(\frac{\beta'_{\ell,(h)}}{\sigma^{\ell}} \right)^{\frac{1}{\ell-2}}.$$

$$I_{(h)} = \frac{1}{C}\min\left\{\min_{i\in[t]}\frac{\sigma}{\sigma_{i,(h)}}, L_{3s,(h)}^{-1}\right\},\ I'_{(h)} = \frac{1}{C}\min\left\{\min_{i\in[t]}\frac{\sigma}{\sigma_{i,(h)}}, L_{3s,(h)}^{\prime -1}\right\}.$$

With this, we also have the following: If ℓ is even, $L_{\ell,(h)} = L'_{\ell,(h)}$. Also, $I_{(h)} = I'_{(h)}$. We next bound $\beta_{\ell,(h)}$ for an even integer ℓ.

$$\begin{aligned}\beta'_{\ell,(h)} = \beta_{\ell,(h)} = \sum_{i=1}^{t}\mathbf{E}\Big[U^{\ell}_{h^{-1}(i)}\Big] \quad &\leq \quad \sum_{i=1}^{t}\ell^{\ell}\Big(\mathbf{E}[U^2_{h^{-1}(i)}]\Big)^{\ell/2} + \sum_{i=1}^{t}\ell^{\ell}\Big(\sum_{j\in h^{-1}(i)}\mathbf{E}[f'_j(y_j)^{\ell}]\Big)\\ &\leq \quad \sum_{i=1}^{t}\ell^{\ell}\Big(\mathbf{E}[U^2_{h^{-1}(i)}]\Big)^{\ell/2} + \sum_{i=1}^{t}\ell^{\ell}\Big(\sum_{j\in h^{-1}(i)}\mathbf{E}[f'_j(y_j)^{2}]\Big)\\ &\leq \quad \sum_{i=1}^{t}\ell^{\ell}\Big(\frac{3\cdot\sigma^2}{2\cdot t}\Big)^{\ell/2} + \ell^{\ell}\sigma^2 \leq \ell^{\ell}\cdot 2^{\ell/2}\cdot\frac{\sigma^{\ell}}{t^{(\ell-2)/2}}. \qquad (29)\end{aligned}$$

The first inequality uses Theorem 38, the second inequality uses that f'_j is supported in $[-1,1]$ and the third inequality uses that h is good. Using this, we get

$$L'_{\ell,(h)} = L_{\ell,(h)} = \left(\frac{\beta_{\ell,(h)}}{\sigma^{\ell}}\right)^{\frac{1}{\ell-2}} \leq \frac{32\cdot\ell}{\sqrt{t}}. \qquad (30)$$

We also let $\{P_{\nu}(i\xi)\}$ be the family of polynomials appearing in Theorem 5 defined with respect to the random variables $\{U_{h^{-1}(i)}\}_{i\in[t]}$. Likewise, we let $\{P'_{\nu}(i\xi)\}$ be the analogous polynomials defined with respect to the random variables $\{U'_{h^{-1}(i)}\}_{i\in[t]}$. Note that for $1 \leq \nu \leq 4k$, the polynomial $P_{\nu}(i\xi)$ is identical to $P'_{\nu}(i\xi)$. We next see that

$$I_{(h)} = I'_{(h)} = \frac{1}{C}\min\left\{\min_{i\in[t]}\frac{\sigma}{\sigma_{i,(h)}}, L^{-1}_{3s,(h)}\right\} \geq \frac{1}{C}\min\left\{\frac{\sqrt{t}}{32\cdot\ell}, \frac{2\sqrt{t}}{3}\right\} = \frac{\sqrt{t}}{32\cdot C\cdot\ell} \geq t^{1/4}. \qquad (31)$$

Thus, applying Theorem 5, we get that for any $\sigma\cdot z\in\mathcal{L}$,

$$\left|\Pr\left[\frac{U_{[n]}}{\sigma} = z\right] - \frac{1}{2\pi\sigma}\cdot\int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z}\cdot e^{-\frac{\xi^2}{2}}\cdot\left(1+\sum_{\nu=1}^{s-3}P_{\nu}(i\xi)\right)d\xi\right| \leq \eta_{\mathsf{low}} + \eta_{\mathsf{med}} + \eta_{\mathsf{high}}$$

and

$$\left|\Pr\left[\frac{U'_{[n]}}{\sigma} = z\right] - \frac{1}{2\pi\sigma}\cdot\int_{-\pi\sigma}^{\pi\sigma} e^{-i\xi z}\cdot e^{-\frac{\xi^2}{2}}\cdot\left(1+\sum_{\nu=1}^{s-3}P'_{\nu}(i\xi)\right)d\xi\right| \leq \eta'_{\mathsf{low}} + \eta'_{\mathsf{med}} + \eta'_{\mathsf{high}}$$

where

$$\eta_{\mathsf{low}} = \eta'_{\mathsf{low}} = s^{O(s)}\cdot L^{s-2}_{3s,(h)} = s^{O(s)}\cdot L'^{s-2}_{3s,(h)},$$

$$\eta_{\mathsf{med}} = \eta'_{\mathsf{med}} = e^{-\frac{I^2_{(h)}}{6}} + s^{O(s)}\cdot e^{-\frac{I^2_{(h)}}{4}} = e^{-\frac{I'^2_{(h)}}{6}} + s^{O(s)}\cdot e^{-\frac{I'^2_{(h)}}{4}},$$

$$\eta_{\mathsf{high}} = \sup_{|\zeta|\in[\sigma^2/\beta_{3,(h)},\pi]}\left|\prod_{i=1}^{t}\widehat{U_{h^{-1}(i)}}(\zeta)\right| + s^{O(s)}\cdot e^{-\frac{I^2_{(h)}}{4}} \text{ and}$$

$$\eta'_{\mathsf{high}} = \sup_{|\zeta|\in[\sigma^2/\beta'_{3,(h)},\pi]}\left|\prod_{i=1}^{t}\widehat{U'_{h^{-1}(i)}}(\zeta)\right| + s^{O(s)}\cdot e^{-\frac{I'^2_{(h)}}{4}}.$$

Applying the triangle inequality, we get

$$\left|\Pr\left[\frac{U_{[n]}}{\sigma} = z\right] - \Pr\left[\frac{U'_{[n]}}{\sigma} = z\right]\right| \leq 2\cdot\eta_{\mathsf{low}} + 2\cdot\eta_{\mathsf{med}} + \eta_{\mathsf{high}} + \eta'_{\mathsf{high}}.$$

We now upper bound the quantities appearing on the right hand side. All the calculations are done assuming the constant c is sufficiently large, n is sufficiently large compared to c and $\epsilon \le n^{-2}$. First, plugging in the values for k and t, it is easy to see that

$$s^{O(s)} \cdot L^{s-2}_{3s,(h)} \le \frac{k^{O(k)}}{t^{k/2}} \le \frac{\epsilon}{32n}; \quad s^{O(s)} \cdot e^{-\frac{I^2_{(h)}}{4}} \le k^{O(k)} \cdot e^{-\sqrt{t}/4} \le \frac{\epsilon}{32n}; \quad e^{-\frac{I'^2_{(h)}}{6}} \le e^{-\sqrt{t}/6} \le \frac{\epsilon}{32n}. \tag{32}$$

$$\begin{aligned}
\sup_{|\zeta| \in \left[\frac{\sigma^2}{\beta_{3,(h)}}, \pi\right]} \left| \prod_{i=1}^{t} \widehat{U_{h^{-1}(i)}}(\zeta) \right| &\le \sup_{|\zeta| \in \left[\frac{1}{\sigma \cdot L_{4,(h)}}, \pi\right]} \left| \prod_{i=1}^{t} \widehat{U_{h^{-1}(i)}}(\zeta) \right| \text{ (uses Claim 8)} \\
&\le \exp\left(\frac{-9 \cdot \sum_{i=1}^{t} \sigma_i^2}{200 \cdot \sigma^2 \cdot L^2_{4,(h)}} \right) \text{ (uses Claim 33)} \\
&= \exp\left(\frac{-9}{200 \cdot L^2_{4,(h)}} \right) \text{ (uses (30))} \\
&= e^{-\Omega(t)} \le \frac{\epsilon}{32n}. && (33)
\end{aligned}$$

Thus, it remains to bound $\sup_{|\zeta| \in [\sigma^2/\beta'_{3,(h)}, \pi]} \left| \prod_{i=1}^{t} \widehat{U'_{h^{-1}(i)}}(\zeta) \right|$. Towards this,

$$\sup_{|\zeta| \in \left[\frac{\sigma^2}{\beta'_{3,(h)}}, \pi\right]} \left| \prod_{i=1}^{n} \widehat{U'_{h^{-1}(i)}}(\zeta) \right| \le \sup_{|\zeta| \in \left[\frac{1}{\sigma \cdot L'_{4,(h)}}, \pi\right]} \left| \prod_{i=1}^{t} \widehat{U'_{h^{-1}(i)}}(\zeta) \right| \text{ (uses Claim 8).}$$

Further, we have that for all $i \in [t]$,

$$\begin{aligned}
\sup_{|\zeta| \in \left[\frac{1}{\sigma \cdot L'_{4,(h)}}, \pi\right]} \left| \widehat{U_{h^{-1}(i)}}(\zeta) \right| &\le \exp\left(\frac{-9 \cdot \sigma_i^2}{200 \cdot \sigma^2 \cdot L'^2_{4,(h)}} \right) \text{ (uses Claim 33)} \\
&\le (1 - \Omega(1) \text{ (uses (30)) and that } h \text{ is good.)}
\end{aligned}$$

Thus, by using Claim 45 and Claim 40, we have that for all $i \in [t]$,

$$\sup_{|\zeta| \in \left[\frac{1}{\sigma \cdot L'_{4,(h)}}, \pi\right]} \left| \widehat{U_{h^{-1}(i)}}(\zeta) \right| \le (1 - \Omega(1)). \tag{34}$$

Thus,

$$\sup_{|\zeta| \in \left[\frac{\sigma^2}{\beta'_{3,(h)}}, \pi\right]} \left| \prod_{i=1}^{t} \widehat{U'_{h^{-1}(i)}}(\zeta) \right| \le e^{-\Omega(t)} \le \frac{\epsilon}{32n}.$$

Putting (32), (33) and (34), we get that

$$\left| \Pr\left[\frac{U_{[n]}}{\sigma} = z \right] - \Pr\left[\frac{U'_{[n]}}{\sigma} = z \right] \right| \le \frac{\epsilon}{2n}$$

which concludes the proof.

□

Acknowledgments

I am grateful to Maksym Radziwill, Rocco Servedio and Avi Wigderson for several critical discussions on this project throughout the past year. Without these, the paper would not have existed. Rocco Servedio made significant contributions at the initial stages of this project but very generously declined to be a co-author. Besides these, the conversations I had with Zeev Dvir, Edinah Gnang and Ofer Zeitouni were of great help.

References

[ASWZ96] Roy Armoni, Michael E. Saks, Avi Wigderson, and Shiyu Zhou. Discrepancy Sets and Pseudorandom Generators for Combinatorial Rectangles. In *Proc. 37th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 412–421, 1996.

[Baz07] Louay Bazzi. Polylogarithmic independence can fool DNF formulas. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007)*, pages 63–73. IEEE Computer Society, 2007.

[BČ02] Andrew D. Barbour and Vydas Čekanvičius. Total variation asymptotics for sums of independent integer random variables. *Annals of Probability*, 30:509–545, 2002.

[BHW94] James G. Booth, Peter Hall, and Andrew T. A. Wood. On the validity of Edgeworth and saddlepoint approximations. *Journal of Multivariate Analysis*, 51:121–138, 1994.

[BR86] Rabindra N. Bhattacharya and Ramaswamy R. Rao. *Normal approximation and asymptotic expansions.* Robert E. Krieger Publishing Company, 1986.

[BR94] Mihir Bellare and John Rompel. Randomness-efficient oblivious sampling. In *35th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 276–287, 1994.

[BRRY10] Mark Braverman, Anup Rao, Ran Raz, and Amir Yehudayoff. Pseudorandom Generators for Regular Branching Programs. In *Proc. 51st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 40–47, 2010.

[BV10] Joshua Brody and Elad Verbin. The Coin Problem and Pseudorandomness for Branching Programs. In *Proc. 51st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 30–39, 2010.

[CGS10] Louis H.Y. Chen, Larry Goldstein, and Qi-Man Shao. Normal approximations by Stein's method. *Springer*, 2010.

[Cra28] Harald Cramér. On the composition of elementary errors, I. *Skand. Aktuarietidskr*, 1:13–74, 1928.

[CRV11] Kai-Min Chung, Omer Reingold, and Salil P. Vadhan. S-T connectivity on digraphs with a known stationary distribution. *ACM Transactions on Algorithms*, 7(3):30, 2011.

[De11] Anindya De. Pseudorandomness for Permutation and Regular Branching Programs. In *IEEE Conference on Computational Complexity*, pages 221–231, 2011.

[EGL$^+$92] Guy Even, Oded Goldreich, Michael Luby, Noam Nisan, and Boban Velickovic. Approximations of General Independent Distributions. In *Proc. 24th Annual ACM Symposium on Theory of Computing (STOC)*, pages 10–16, 1992.

[Ess45] C.-G. Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Mathematica*, 77(1):1–125, 1945.

[Fel68] W. Feller. *An introduction to probability theory and its applications.* John Wiley & Sons, 1968.

[Gar07] David J. H. Garling. *Inequalities: A Journey into Linear Analysis.* Cambridge University Press, 2007.

[GMR$^+$12] Parikshit Gopalan, Raghu Meka, Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Better Pseudorandom Generators from Milder Pseudorandom Restrictions. In *Proc. 53rd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 120–129, 2012.

[GMRZ13] Parikshit Gopalan, Raghu Meka, Omer Reingold, and David Zuckerman. Pseudorandom generators for combinatorial shapes. *SIAM Journal of Computing*, 42:1051–1076, 2013.

[IKW02] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. In search of an easy witness: Exponential time vs. Probabilistic polynomial time. *Journal of Computer and System Sciences*, 65(4):672–694, 2002.

[IL71] Ildar A. Ibragimov and Yuri V. Linnik. *Independent and Stationary Sequences of Random Variables.* Woolters Noordhoff, Groningen, 1971.

[INW94] Russell Impagliazzo, Noam Nisan, and Avi Wigderson. Pseudorandomness for network algorithms. In *Proc. 26th Annual ACM Symposium on Theory of Computing (STOC)*, pages 356–364, 1994.

[IS06] R. Ibragimov and Sh. Sharakhmetov. The exact constant in the rosenthal inequality for random variables with mean zero. *Theory of Probability and its Applications*, 46(1):127–132, 2006.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds. *Computational Complexity*, 13(1-2):1–46, 2004.

[KNP11] Michal Koucký, Prajakta Nimbhorkar, and Pavel Pudlák. Pseudorandom generators for group products. In *Proc. 43rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 263–272, 2011.

[LRTV09] Shachar Lovett, Omer Reingold, Luca Trevisan, and Salil Vadhan. Pseudorandom bit generators that fool modular sums. In *13th International Workshop on Randomization and Computation*, pages 615–630, 2009.

[Lu02] Chi-Jen Lu. Improved Pseudorandom Generators for Combinatorial Rectangles. *Combinatorica*, 22(3):417–434, 2002.

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

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM J. on Comput.*, 22(4):838–856, 1993. Earlier version in STOC'90.

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

[Pet75] Valentin Petrov. *Sums of Independent Random Variables.* Springer, 1975.

[Rei08] Omer Reingold. Undirected connectivity in log-space. *Journal of the ACM*, 55(4), 2008. Preliminary version in STOC'05.

[RSV13] Omer Reingold, Thomas Steinke, and Salil P. Vadhan. Pseudorandomness for Regular Branching Programs via Fourier Analysis. In *17th International Workshop on Randomization and Computation*, pages 655–670, 2013.

[RTV06] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Pseudorandom walks on regular digraphs and the RL vs. L problem. In *Proc. 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 457–466, 2006.

[Sav70] Walter J. Savitch. Relationships Between Nondeterministic and Deterministic Tape Complexities. *J. Comput. Syst. Sci.*, 4(2):177–192, 1970.

[Ste12] Thomas Steinke. Pseudorandomness for Permutation Branching Programs Without the Group Theory. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:83, 2012.

[SVW14] Thomas Steinke, Salil P. Vadhan, and Andrew Wan. Pseudorandomness and Fourier Growth Bounds for Width 3 Branching Programs. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:76, 2014.

[SZ95] Michael E. Saks and Shiyu Zhou. $\mathsf{RSPACE(S)} \subseteq \mathsf{DSPACE(S^{3/2})}$. In *Proc. 36th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 344–353, 1995.

[Wat13] Thomas Watson. Pseudorandom generators for combinatorial checkerboards. *Computational Complexity*, 22(4):727–769, 2013.

A Useful analytic inequalities

The following inequalities can found in [Gar07].

Lemma 34. Monotonicity of norms: *Let $1 \leq p \leq q < \infty$ and $f \in L_p \cap L_q$. Then,*

$$\|f\|_p \leq \|f\|_q$$

.

Lemma 35. Lyupanov's inequality: *Let $1 \leq p, q < \infty$ and $0 \leq \lambda \leq 1$. Let $r = \lambda p + (1-\lambda)q$. Let $f \in L_p \cap L_q$. Then,*

$$\|f\|_r^r \leq \|f\|_p^{\lambda p} \cdot \|f\|_q^{(1-\lambda)q}.$$

Lemma 36. Chebyshev's inequality: *For any random variable X, with $\mathbf{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$, for any $t \geq 0$,*

$$\Pr[|X - \mu| \leq t \cdot \sigma] \leq \frac{1}{t^2}.$$

The next couple of inequalities are concentration bounds on sums of random variables.

Lemma 37. Bellare-Rompel [BR94]: *Let $X_1, \ldots, X_n$ be k-wise independent random variables supported in $[0,1]$. Let $X = X_1 + \ldots + X_n$. If $\mu = \mathbf{E}[X]$, then*

$$\Pr[|X - \mu| > a] \leq 8 \cdot \left(\frac{k\mu + k^2}{a^2}\right)^{k/2}$$

Lemma 38. Rosenthal's inequality [IS06]: *Let $X_1, \ldots, X_n$ be independent random variables such that* $\mathbf{E}[X_i] = 0$*. If $Z = (X_1 + \ldots + X_n)$, then*

$$\mathbf{E}[|Z|^{2m}] \leq (2m)^{2m} \cdot \max \left\{ \sum_{i=1}^{n} \mathbf{E}[|X_i|^{2m}], \left(\sum_{i=1}^{n} \mathbf{E}[X_i^2] \right)^{m/2} \right\}.$$

The next lemma gives a lower bound on the decay of Fourier spectrum for a sum of independent Bernoulli variables.

Lemma 39. *Let $X_1, \ldots, X_n$ be independent Bernoulli random variables where* $\Pr[X_i = 1] = p_i$ *(for $i = 1, \ldots, n$). Let $Z = X_1 + \ldots + X_n$ and σ be the variance of Z. Then, for $|\xi| \leq \pi$,*

$$|\widehat{Z}(\xi)| \leq \exp\left(-\frac{9 \cdot \sigma^2 \cdot \xi^2}{200} \right).$$

Proof.

$$\widehat{Z}(\xi) = \prod_{i=1}^{n} \widehat{X_i}(\xi) = \prod_{i=1}^{n} (1 - p_i + p_i \cdot e^{i\xi}).$$

Thus,

$$|\widehat{Z}(\xi)|^2 = \prod_{i=1}^{n} |\widehat{X_i}^2(\xi)| = \prod_{i=1}^{n} (1 - p_i(1-p_i)\sin^2(\xi/2)) \leq \prod_{i=1}^{n} \left(1 - \frac{9 \cdot p_i(1-p_i)\xi^2}{100} \right).$$

The last inequality uses the elementary fact that for $-\pi/2 \leq \xi \leq \pi/2$, $|\sin(\xi)| \geq 0.6 \cdot \xi$. Further, using that for $x \geq 0$, $e^{-x} \geq 1 - x$, we get

$$\prod_{i=1}^{n} \left(1 - \frac{9 \cdot p_i(1-p_i)\xi^2}{100} \right) \leq \prod_{i=1}^{n} \exp\left(-\frac{9 \cdot p_i(1-p_i)\xi^2}{100} \right) = \exp\left(-\frac{9 \cdot \sigma^2 \cdot \xi^2}{100} \right).$$

This finishes the proof. □

Lemma 40. *Let X and Y be two random variables supported* $\mathbb{R}$*. Then, for any $\xi \in \mathbb{R}$,*

$$\left| \mathbf{E}_{x \in X}[e^{i \cdot \xi \cdot x}] - \mathbf{E}_{y \in Y}[e^{i \cdot \xi \cdot y}] \right| \leq \|X - Y\|_1.$$

Proof. Recall that for any function $f \in L_1(\mathbb{R})$, $\|f\|_1 \geq \|\widehat{f}\|_\infty$. Viewing X and Y as a map from $\mathbb{R}$ to $[0, 1]$ and putting $f = X - Y$, we get the lemma. □

B Useful existing constructs in pseudorandomness

Lemma 41. ***k-wise independent hash function:*** *A family of functions $\mathcal{H}_{k,n,t} = \{h : [n] \to [t]\}$ is said to be k-wise independent if for all $i_1, \ldots, i_k \in [n]$ and $j_1, \ldots, j_k \in [t]$*

$$\Pr_{h \in \mathcal{H}_k}[h(i_1) = j_1 \ \wedge \ \ldots \ \wedge h(i_k) = j_k] = \frac{1}{t^k}.$$

Further, elements from such a family $\mathcal{H}_k$ can be sampled efficiently using $O(k(\log t + \log n))$ bits.

Lemma 42. ***k-wise independent PRG:*** *$G_{k,n,m} : \{0,1\}^{t_{k,m,n}} \to [m]^n$ is said to be a k-wise independent generator if for every $i_1, \ldots, i_k \in [n]$ and $j_1, \ldots, j_k \in [m]$,*

$$\Pr_{x \in U_{t_{k,m,n}}}[G_{k,n,m}(x)_{i_1} = j_1 \ \wedge \ \ldots \ \wedge G_{k,n,m}(x)_{i_k} = j_k] = \frac{1}{m^k}.$$

Further, an efficiently computable $G_{k,n,m}$ exists with $t_{k,m,n} = O(k \log m + k \log n)$ bits.

Lemma 43. ***Biased pairwise independence*** *A family of functions $\mathcal{H}_{2,n,p} = \{h : [n_1] \to \{-1,1\}\}$ is said to be pairwise independent with bias $p \in [0,1]$ if for any $i_1, i_2 \in [n]$,*

$$\Pr_{h \in \mathcal{H}_{2,n,p}}[h(i_1) = 1 \ \wedge h(i_2) = 1] = p^2 \text{ and} \Pr_{h \in \mathcal{H}}[h(i_1) = 1] = p.$$

Further, if $q = n \cdot p$ is a prime number, then elements from $\mathcal{H}_{2,n,p}$ can be sampled with $O(\log n)$ bits.

C Some invariance results for PRGs

The following simple results show that PRGs for certain classes remain invariant under shifting by a fixed string.

Claim 44. *Let $G : \{0,1\}^{t_{k,m,n}} \to [m]^n$ be a k-wise independent generator. Then, $G'_z : \{0,1\}^{t_{k,m,n}} \to [m]^n$ defined as $G'_z(x) = G(x) \oplus_m z$ is also a k-wise independent generator.*

The proof for the above is left as exercise for the reader.

Claim 45. *Let $G : \{0,1\}^t \to [m]^n$ be an ϵ-PRG for $\mathsf{Csum}(m,n)$. Then, for $z \in [m]^n$, $G'_z : \{0,1\}^{t_{k,m,n}} \to [m]^n$ defined as $G'_z(x) = G(x) \oplus_m z$ also ϵ-fools $\mathsf{Csum}(m,n)$. The same also holds for PRGs for the classes $\mathsf{Csum}_{\mathsf{low}}(m,n)$ and $\mathsf{Csum}_{\mathsf{high}}(m,n)$.*

Proof. Let $f \in \mathsf{Csum}(m,n)$ be specified by the tuple $(f_1, \ldots, f_n)$.

$$\|f(G'_z(U_t)) - f(U_{[m]^n})\|_1 = \|f(G'_z(U_t)) - f(U_{[m]^n} \oplus_m z)\|_1 = \|g(G(U_t)) - g(U_{[m]^n})\|_1,$$

where $g = (g_1, \ldots, g_n)$ and for $1 \le i \le n$, $g_i(x_i) = f_i(x_i \oplus_m z_i)$. Since $g \in \mathsf{Csum}(m,n)$, hence we get that G'_z is an ϵ-PRG for $\mathsf{Csum}(m,n)$. Further, since $\mathrm{Var}(Z_g) = \mathrm{Var}(Z_f)$, we have the same for $\mathsf{Csum}_{\mathsf{low}}(m,n)$ and $\mathsf{Csum}_{\mathsf{high}}(m,n)$. □