Discrepancy Sets and Pseudorandom Generators for Combinatorial Rectangles*

Roy Armoni† Michael Saks‡ Avi Wigderson§ Shiyu Zhou¶

Abstract

A common subproblem of DNF approximate counting and derandomizing RL is the discrepancy problem for combinatorial rectangles. We explicitly construct a $poly(n)$-size sample space that approximates the volume of any combinatorial rectangle in $[n]^n$ to within $o(1)$ error (improving on the constructions of [EGLNV92]). The construction extends the techniques of [LLSZ95] for the analogous hitting set problem, most notably via discrepancy preserving reductions.

1 Introduction

In a general discrepancy problem, we are given a family of sets and want to construct a small sample space that approximates the volume of an arbitrary set in the family. This problem is closely related to other important issues in combinatorial constructions such as the problem of constructing small sample spaces that approximate the independent distributions on many multivalued random variables [KW84, Lub85, ABI86, CG89, NN90, AGHP90, EGLNV92, Sch92, KM93, KK94], and the problem of constructing pseudorandom generators for space bounded computation [Nis90, NZ93, INW94, AW96].

More precisely, one can define the following notion of a discrepancy set: Let U be a set and $\mathcal{F}$ be a family of subsets of U. A multiset S of U is said to be an ϵ-discrepancy set for $\mathcal{F}$ if for each element $A \in \mathcal{F}$, the difference between the fraction of points in U that belong to A and the fraction of points in S that belong to A is within ϵ. In this paper, motivated by the issue of the construction of space bounded pseudorandom generators, we investigate the problem of constructing small discrepancy sets for a special class of sets called combinatorial rectangles: for positive integers m and d, a *combinatorial rectangle* of type (m, d) (or an (m, d)-*rectangle*) is a subset of $[m]^d$ of the form $R = R_1 \times R_2 \times \ldots \times R_d$, where each $R_i \subseteq [m]$. We use $\mathcal{R}(m, d)$ to denote the family of all (m, d)-rectangles.

It is easy to show that if we select $O(md/\epsilon)$ points from $[m]^d$ uniformly at random, then the resulting set is almost surely an ϵ-discrepancy set for $\mathcal{R}(m, d)$. On the other hand, the problem of finding an explicit construction for such a set of size polynomial in m, d and ϵ^{-1}, is still open.

The discrepancy set problem for rectangles was first studied in the context of number theory and real analysis, where the family of sets considered was the family of geometric rectangles in which each "side" R_i of the rectangle is an interval. [BC87] and [Nie] are good references for this material, which contain sharp existential bounds mainly for small dimensions. The general problem of explicit constructions in high dimensions, and for combinatorial rather than geometric rectangles, was first formulated in [EGLNV92]. Their motivation was approximating independent multivalued distributions. They gave a poly size construction for the geometric case, and two quasi-polynomial constructions for the general case: of size $(md/\epsilon)^{\log d}$ (based on Nisan's bounded-space generator [Nis90]) and size $(md/\epsilon)^{\log 1/\epsilon}$ (based on k-wise independence). These become polynomial size (respectively) if either the dimension d or the error ϵ are bounded.

Another source of explicit constructions of small sample spaces for this problem comes from observing that (non)membership in a rectangle can be checked by a DNF formula of size $O(md)$. Thus small sample spaces which approximate such circuits are good discrepancy sets with the same error. Nisan's

*The second and the fourth authors were supported in part by NSF grant CCR-9215293. The third author was partially supported by a Wolfson Research Award, administered by the Israeli Academy of Sciences and Humanities, and by a grant from the Sloan Foundation. All four authors were supported in part by DIMACS (Center for Discrete Mathematics & Theoretical Computer Science), through NSF grant NSF-STC91-19999 and by the New Jersey Commission on Science and Technology.

†Computer Science Institute, The Hebrew University, Jerusalem, Israel. E-mail: *aroy@cs.huji.ac.il*.

‡Department of Mathematics, Rutgers University, New Brunswick, NJ 08854, USA. E-mail: *saks@math.rutgers.edu*.

§Computer Science Institute, The Hebrew University, Jerusalem, Israel. Currently on leave at the Institute for Advanced Study and Princeton University. E-mail: *avi@math.ias.edu*.

¶Department of Computer Science, Rutgers University, New Brunswick, NJ 08854, USA. E-mail: *szhou@cs.rutgers.edu*.

constant-depth generator [Nis91] (and the improvements in [LVW93]) again give (completely different) quasi-polynomial constructions.

Our main result is an explicit construction of a sample space for this problem of size polynomial in m, d and $\epsilon^{\log \epsilon}$. In other words it is polynomial as long as $1/\epsilon < exp(\sqrt{\log md})$. While reducing the error significantly below the constant bound that follows from previous constructions, it falls short of the natural goal of $1/\epsilon = (md)^{O(1)}$. Our construction is different than previous ones for the discrepancy problem, and follows rather closely the constructions for the related hitting set problem for rectangles (where all we need is that every rectangle of volume at least ϵ is hit by at least one point in the small sample space). Thus, the hitting set problem is the one-sided error version of the discrepancy problem. For this simpler problem a fully polynomial solution, namely an explicit set of size $(m \log d/\epsilon)^{O(1)}$, was given by [LLSZ95]. It is achieved by a sequence of "hitting-preserving" reductions, which we generalize in this paper to "discrepancy-preserving" reductions. These simplify and reduce (in turn) various parameters of the problem without affecting too much the volume of the sets and the size of the sample needed. Naturally, while in the (one-sided) hitting set problem it suffices to control a lower bound on the volume, we need to keep tight upper and lower bounds throughout these reductions.

Our description (and motivation) of the constructions comes from the perspective of constructing pseudorandom generators, which we explain below. Let l, m, d be positive integers. An (l, m, d)-*generator* G is a function that maps $\{0,1\}^l$ to $[m]^d$. G is said to ϵ-*fool* the (m, d)-rectangles if the multiset image of G in $[m]^d$, which is of size 2^l, is an ϵ-discrepancy set for $\mathcal{R}(m, d)$. Thus to construct a small discrepancy set for $\mathcal{R}(m, d)$, it suffices to construct an (l, m, d)-generator that ϵ-fools the (m, d)-rectangles with small l.

An (m, d)-rectangle can be visualized as a width-2 read-once branching program of length d over alphabet $[m]$ in the natural way. Since in general any pseudorandom generator for non-uniform space bounded computation in the finite state machine model [Nis90] can fool width-$m^{\Omega(1)}$ read-once branching programs over alphabet $[m]$, it fools all (m, d)-rectangles. In particular, Nisan's pseudorandom generator [Nis90] gives an (l, m, d)-generator with $l = \log d(2 \log m + \log d + \log \epsilon^{-1})$ that ϵ-fools the (m, d)-rectangles, and the Impagliazzo-Nisan-Wigderson generator [INW94] gives one with $l = \log m + 2 \log d(\log d + \log \epsilon^{-1})$. Both of these generators are *efficient* in the sense that they are computable in polynomial time (polynomial in the length of the output) and linear space (linear in the length of the input). Nevertheless, these fall short of the natural lower bound of $l = O(\log m + \log d + \log \epsilon^{-1})$. Our new construction, in this language, is an efficient generator with $l = O(\log m + \log d + \log^2 \epsilon^{-1})$. It is interesting to observe that achieving a bound $l = \log m + O(\log \epsilon^{-1}) + f(d)$ for an arbitrary function f will result in a $o((\log n)^2)$-bit generator which fools all constant-width read-once branching programs - one of the main challenges in derandomizing space-bounded computation and a major motivation for our interest in the discrepancy problem.

The rest of the paper is organized as follows. In Section 2, we provide basic notation and definitions; moreover, we formalize the reduction framework in terms of the compositions of function reductions which helps us to clarify certain subtleties in the generator construction. An overview of the construction is given in Section 3.1 and the details of the construction are given in the later sections.

2 Preliminaries

2.1 Basic Notation

For a set U, we let 2^U denote the family of subsets of U and let $\mathcal{M}(U)$ denote the family of multi-subsets of U. With respect to a fixed order of the elements in U, we identify each $S \in \mathcal{M}(U)$ with a nonnegative integer vector indexed by U such that for any element $u \in U$, the entry indexed by u is the number of appearances of u in S. For example, if $U = \{a, b, c\}$ whose elements are in alphabetical order and $S = \{a, c, c, c\} \in \mathcal{M}(U)$, then we have $U = (1, 1, 1)$ and $S = (1, 0, 3)$. Clearly, the inner product $< S, U >$ is the cardinality of S, which we denote by $|S|$.

Let U and V be sets. We say that a matrix is a $U \times V$ matrix if the rows and columns of the matrix are indexed by the elements of U and V, respectively.

For typographical simplicity, we will not specify whether a vector is a row or a column vector in the case that this is easily seen from the context.

All integers are positive unless otherwise specified. If m is an integer, we use $[m]$ to denote the set of integers $\{1, 2, \ldots, m\}$.

2.2 Discrepancy Sets and Reductions

Let U be a set. For a subset $A \subseteq U$, the *volume* of A (in U), denoted $vol(A)$, is defined to be the fraction of elements in U that lie in A, i.e.,

$$vol(A) = vol^U(A) = \frac{< A, U >}{< U, U >}.$$

Let $S \in \mathcal{M}(U)$. The *discrepancy* of A with respect to S (in U) is defined to be

$$disc_S(A) = disc_S^U(A) = |\frac{< A, S >}{< U, S >} - vol(A)|.$$

For $\mathcal{A} \subseteq 2^U$, we define $disc_S(\mathcal{A}) = \max_{A \in \mathcal{A}} disc_S(A)$. We say that S is an *ϵ-discrepancy set* for $\mathcal{A}$ if $disc_S(\mathcal{A}) \leq \epsilon$.

Remark: Here we emphasize the facts that volume is defined only on sets but not on multisets, while discrepancy is defined only on sets but with respect to multisets.

Let U and V be sets. A *reduction Λ between U and V* is a $U \times V$ nonnegative integral matrix. The *cost* of the reduction, denoted $cost(\Lambda)$, is defined to be the maximum column sum of the matrix; the *image* of the reduction, denoted $image(\Lambda)$, is defined to be ΛV. (Here we emphasize that $image(\Lambda)$ is a multiset.) Clearly, $|image(\Lambda)| \leq cost(\Lambda)|V|$. It is often convenient for us to view such a reduction as a bipartite multigraph on U and V such that there are k edges connecting vertex $u \in U$ and vertex $v \in V$ if and only if the (u, v)-th entry of the reduction is k. The cost of the reduction is thus the maximum degree of any vertex in V, and the image of the reduction is the set of neighbors of V in U counting multiplicity.

Let Λ be a reduction between U and V. Then it is clear that for any $A \in \mathcal{M}(U)$ and any $B \in \mathcal{M}(V)$, we have $A\Lambda \in \mathcal{M}(V)$ and $\Lambda B \in \mathcal{M}(U)$.

Suppose $\mathcal{A} \subseteq 2^U$ and $\mathcal{B} \subseteq 2^V$. A reduction Λ between U and V is said to be *$(\mathcal{A}, \mathcal{B}, \delta)$-discrepancy preserving* if for any $S \in \mathcal{M}(V)$, $disc_{\Lambda S}(\mathcal{A}) \leq disc_S(\mathcal{B}) + \delta$. That is, S is an ϵ-discrepancy set for $\mathcal{B}$ implies that ΛS is an $(\epsilon + \delta)$-discrepancy set for $\mathcal{A}$. Therefore, intuitively, such a reduction reduces the problem of finding a discrepancy set for family $\mathcal{A}$ to the problem of finding a discrepancy set for family $\mathcal{B}$.

Proposition 2.1 *Let U and V be sets and let $\mathcal{A} \subseteq 2^U$. For an arbitrary $\mathcal{B} \subseteq 2^V$, suppose Λ is a reduction between U and V that is $(\mathcal{A}, \mathcal{B}, \delta)$-discrepancy preserving, then $image(\Lambda)$ is a δ-discrepancy set for $\mathcal{A}$.*

This is because $disc_V(B) = 0$ for any $B \subseteq V$.

2.3 Function Reductions

We will be dealing with reductions specified by function families, which we call *function reductions*. One remark on notation: in the case that a function family is a singleton set $\{f\}$, we may use f for simplicity.

Any function f that maps V to U specifies a reduction Λ_f between U and V in a natural way: for $u \in U$ and $v \in V$, $\Lambda_f(u, v) = 1$ if and only if $f(v) = u$. We note that for any $X \subseteq U$, $X\Lambda_f$ is the subset $f^{-1}(X)$ of V and in particular, we have $U\Lambda_f = V$. For a family $\mathcal{F}$ of functions mapping V to U, the function reduction $\Lambda_{\mathcal{F}}$ between U and V specified by $\mathcal{F}$ is defined to be the sum of Λ_f over $f \in \mathcal{F}$, i.e., $\Lambda_{\mathcal{F}} = \sum_{f \in \mathcal{F}} \Lambda_f$. The *image* of $\mathcal{F}$, $image(\mathcal{F})$, is defined to be $image(\Lambda_{\mathcal{F}}) = \Lambda_{\mathcal{F}} V$. It is clear that $cost(\Lambda_{\mathcal{F}}) = |\mathcal{F}|$ and $|image(\mathcal{F})| = |\mathcal{F}||V|$.

Let $\mathcal{A} \subseteq 2^U$ and $\mathcal{B} \subseteq 2^V$. $\mathcal{F}$ is said to be *$(\mathcal{A}, \mathcal{B}, \delta)$-good* if for each $A \in \mathcal{A}$ the following hold:

- for any $f \in \mathcal{F}$, $A\Lambda_f \in \mathcal{B}$, and
- $|E_{f \in \mathcal{F}}[vol(A\Lambda_f)] - vol(A)| \leq \delta$, where the expectation is over a randomly chosen $f \in \mathcal{F}$.

Lemma 2.1 *Let U and V be sets. Suppose $\mathcal{F}$ is a family of functions mapping V to U. For $\mathcal{A} \subseteq 2^U$ and $\mathcal{B} \subseteq 2^V$, if $\mathcal{F}$ is $(\mathcal{A}, \mathcal{B}, \delta)$-good, then the function reduction $\Lambda_{\mathcal{F}}$ is $(\mathcal{A}, \mathcal{B}, \delta)$-discrepancy preserving and, consequently, $image(\mathcal{F})$ is a δ-discrepancy set for $\mathcal{A}$.*

Proof: Fix any $S \in \mathcal{M}(V)$. We want to show that for any $A \in \mathcal{A}$, $disc_{\Lambda_{\mathcal{F}} S}(A) \leq disc_S(\mathcal{B}) + \delta$.

$$\begin{aligned}
disc_{\Lambda_{\mathcal{F}} S}(A) &= |\frac{< A, \Lambda_{\mathcal{F}} S >}{< U, \Lambda_{\mathcal{F}} S >} - vol(A)| \\
&= |\frac{\sum_{f \in \mathcal{F}} < A\Lambda_f, S >}{\sum_{f \in \mathcal{F}} < U\Lambda_f, S >} - vol(A)| \\
&= |\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} (\frac{< A\Lambda_f, S >}{< V, S >} - vol(A))| \\
&\leq \frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} |\frac{< A\Lambda_f, S >}{< V, S >} - vol(A\Lambda_f)| + \\
&\quad |\frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} vol(A\Lambda_f) - vol(A)| \\
&\leq \max_{B \in \mathcal{B}} disc_S(B) + \\
&\quad |E_{f \in \mathcal{F}}[vol(A\Lambda_f)] - vol(A)| \\
&\leq disc_S(\mathcal{B}) + \delta.
\end{aligned}$$

Now Proposition 2.1 concludes the proof since $image(\mathcal{F})$ is $\Lambda_{\mathcal{F}} V$ by definition. □

Lemma 2.1 suggests that in order to construct a small-sized ϵ-discrepancy set for $\mathcal{A} \subseteq 2^U$, it suffices to construct a family $\mathcal{F}$ of functions mapping V to U for some V that is $(\mathcal{A}, \mathcal{B}, \epsilon)$-good for some $\mathcal{B} \subseteq 2^V$, such that both $|\mathcal{F}|$ and $|V|$ are small (thus $|image(\mathcal{F})|$ is small).

2.4 The Composition of Families of Functions

Let $V_1, V_2, \ldots, V_k$ be sets, and let $\mathcal{F}_i, 1 \leq i \leq k-1$, be a sequence of families of functions with each $\mathcal{F}_i$ mapping V_{i+1} to V_i. The composition of $\mathcal{F}_i$, denoted $\mathcal{F}^{(k-1)} = \bigodot_{i=1}^{k-1} \mathcal{F}_i$, is defined to be the family of all functions of the form $f_1 \circ f_2 \circ \ldots \circ f_{k-1}$, where $f_i \in \mathcal{F}_i$ for each i and $\circ$ denotes the function composition. The following fact can be easily proved by induction:

Proposition 2.2 *Let $V_1, V_2, \ldots, V_k$ be sets and for each $1 \leq i \leq k-1$, let $\mathcal{F}_i$ be a family of functions mapping V_{i+1} to V_i. Then the composition $\mathcal{F}^{(k-1)}$ is a family of functions mapping V_k to V_1 of size $\Pi_{i=1}^{k-1}|\mathcal{F}_i|$ such that $\Lambda_{\mathcal{F}^{(k-1)}} = \Pi_{i=1}^{k-1}\Lambda_{\mathcal{F}_i}$. In words, the function reduction specified by the composition is the product of the function reduction specified by each single family in the composition.*

The next lemma will be useful for our generator construction.

Lemma 2.2 *Let $V_1, V_2, \ldots, V_k$ be sets and let $\mathcal{A}_i \subseteq 2^{V_i}$ for $1 \leq i \leq k$. Suppose for each $1 \leq i \leq k-1$, $\mathcal{F}_i$ is a family of functions mapping V_{i+1} to V_i that is $(\mathcal{A}_i, \mathcal{A}_{i+1}, \delta_i)$-good. Then the composition $\mathcal{F}^{(k-1)}$ is $(\mathcal{A}_1, \mathcal{A}_k, \delta)$-good where $\delta = \sum_{i=1}^{k-1} \delta_i$.*

Proof: We prove by induction on k that $\mathcal{F}^{(k-1)}$ is $(\mathcal{A}_1, \mathcal{A}_k, \delta)$-good. The case where $k = 2$ is trivial. Assume that it holds for $k-1$ and we show for k.

Fix any $A \in \mathcal{A}_1$. We first need to show that for any $f \in \mathcal{F}^{(k-1)}$, $A\Lambda_f \in \mathcal{A}_k$. Let $f = f_1 \circ f_2 \circ \ldots \circ f_{k-1} \in \mathcal{F}^{(k-1)}$ be arbitrary. Then it follows from Proposition 2.2 that $\Lambda_f = \Pi_{i=1}^{k-1}\Lambda_{f_i}$. By the induction hypothesis, we have $A\Pi_{i=1}^{k-2}\Lambda_{f_i} \in \mathcal{A}_{k-1}$. Since $\mathcal{F}_{k-1}$ is $(\mathcal{A}_{k-1}, \mathcal{A}_k, \delta_{k-1})$-good, by definition, $(A\Pi_{i=1}^{k-2}\Lambda_{f_i})\Lambda_{f_{k-1}} \in \mathcal{A}_k$. It remains to show that $|E_{f \in \mathcal{F}^{(k-1)}}[vol(A\Lambda_f)] - vol(A)| \leq \delta$.

$$
\begin{aligned}
& |E_{f=f_1\circ\ldots\circ f_{k-1}\in\mathcal{F}^{(k-1)}}[vol(A\Lambda_f)] - vol(A)| \\
= \; & |E_{f_1\circ\ldots\circ f_{k-2}\in\mathcal{F}^{(k-2)}} E_{f_{k-1}\in\mathcal{F}_{k-1}} \\
& [vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i}\Lambda_{f_{k-1}})] - vol(A)| \\
= \; & |E_{f_1\circ\ldots\circ f_{k-2}\in\mathcal{F}^{(k-2)}}[\ (E_{f_{k-1}\in\mathcal{F}_{k-1}} \\
& [vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i}\Lambda_{f_{k-1}})] - vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i})) \\
& + (vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i}) - vol(A))\ ]| \\
\leq \; & E_{f_1\circ\ldots\circ f_{k-2}\in\mathcal{F}^{(k-2)}}[\ |(E_{f_{k-1}\in\mathcal{F}_{k-1}} \\
& [vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i}\Lambda_{f_{k-1}})] - vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i})|\ ] \\
& + |E_{f_1\circ\ldots\circ f_{k-2}\in\mathcal{F}^{(k-2)}}[vol(A\Pi_{i=1}^{k-2}\Lambda_{f_i})] - vol(A)| \\
\leq \; & \delta_{k-1} + \sum_{i=1}^{k-2} \delta_i = \delta
\end{aligned}
$$

where the first term of the last inequality holds since $A\Pi_{i=1}^{k-2}\Lambda_{f_i} \in \mathcal{A}_{k-1}$ by induction, and $\mathcal{F}_{k-1}$ is $(\mathcal{A}_{k-1}, \mathcal{A}_k, \delta_{k-1})$-good; the second term holds because $\mathcal{F}^{(k-2)}$ is $(\mathcal{A}_1, \mathcal{A}_{k-2}, \sum_{i=1}^{k-2} \delta_i)$-good by induction, and $\Lambda_{f_1\circ\ldots\circ f_{k-2}} = \Pi_{i=1}^{k-2}\Lambda_i$. □

2.5 Efficiency in Function Computation

For the purposes of constructing pseudorandom generators, we review some facts about the efficiency in the computation of functions.

We say that a function f is in $TS(t(n), s(n))$ if on input of length n, f is computable in time $t^{O(1)}(n)$ and in space $O(s(n))$; and we say that a family of functions is in $TS(t(n), s(n))$ if each function in the family is so.

A family $\mathcal{F}$ of functions is said to be *indexable* if $\mathcal{F}$ can be *identified* with $[|\mathcal{F}|]$ in the sense that each function $f \in \mathcal{F}$ can be indexed by an integer in $[|\mathcal{F}|]$ (or a bit-sequence of length $\lceil \log |\mathcal{F}| \rceil$) so that if f is in $TS(t(n), s(n))$, then given its index, the computation of f can be simulated in $TS(t(n), s(n))$ as well.

It is not difficult to see the following:

Proposition 2.3 *Let k be fixed and let $V_1, V_2, \ldots, V_k$ be sets. Suppose for each $1 \leq i \leq k-1$, $\mathcal{F}_i$ is a family of functions mapping V_{i+1} to V_i. If $\mathcal{F}_i$ is in $TS(t_i(n), s_i(n))$ for each i, then $\bigodot_{i=1}^{k-1} \mathcal{F}_i$ is in $TS(\Pi_{i=1}^{k-1} t_i(n), \sum_{i=1}^{k-1} s_i(n))$. Moreover, if each $\mathcal{F}_i$ is indexable, so is $\bigodot_{i=1}^{k-1} \mathcal{F}_i$.*

Given an indexable family $\mathcal{F}$ of functions mapping V to U, the *unification* of $\mathcal{F}$, denoted $G_{\mathcal{F}}$, is defined to be a function mapping $[|\mathcal{F}|] \times V$ to U such that: on input $(\alpha, v) \in [|\mathcal{F}|] \times V$, $G_{\mathcal{F}}$ takes the function $f \in \mathcal{F}$ indexed by α, simulates the computation of $f(v)$ and outputs the result. It follows immediately from the construction that:

Proposition 2.4 *Let $\mathcal{F}$ be an indexable family of functions. Then image$(G_{\mathcal{F}})$ = image$(\mathcal{F})$, and if $\mathcal{F}$ is in $TS(t(n), s(n))$, so is $G_{\mathcal{F}}$.*

2.6 k-wise Independent Hash Function Family

Let a, b be integers. A family H of functions mapping $[a]$ to $[b]$ is said to be a *k-wise independent hash function family* if for any $u_1, u_2, \ldots, u_k \in [a]$ such that $u_i \neq u_j$ for $1 \leq i < j \leq k$, and any $v_1, v_2, \ldots, v_k \in [b]$,

$$Pr_{h \in H}[h(u_i) = v_i \text{ for } 1 \leq i \leq k] = 1/b^k.$$

(A pairwise independent hash function family is usually called a *universal hash function family* [CW79].)

It is easy to check that a k-wise independent hash function family is a $(k-1)$-wise independent hash function family. We will need the following well-known fact:

Theorem 2.1 *Let k be fixed. Then for any a, b that are integer powers of 2, there is an explicit construction of a k-wise independent hash function family mapping $[a]$ to $[b]$ of size $(\max(a,b))^k$. Moreover, the family is indexable and is in $TS(\log ab, \log ab)$.*

2.7 Combinatorial Rectangles, Discrepancy Sets and Pseudorandom Generators

For integers m and d, let $U = [m]^d$. A *combinatorial rectangle* of type (m,d) (or an (m,d)-*rectangle*) is a subset of U of the form $R = R_1 \times R_2 \times \ldots \times R_d$, where each $R_i \subseteq [m]$. By definition, the volume of R is thus

$$vol(R) = |R|/|U| = \Pi_{i=1}^{d} |R_i|/m^d.$$

R is said to be PIP, which stands for *pairwise independent projections*, if for any $1 \leq i < j \leq d$, $\frac{|R_i \cap R_j|}{m} = \frac{|R_i|}{m} \cdot \frac{|R_j|}{m}$. We use $\mathcal{R}(m,d)$ to denote the family of all (m,d)-rectangles, and we use PIP-$\mathcal{R}(m,d)$ to denote the family of all PIP (m,d)-rectangles.

Let l, m, d be positive integers. An (l,m,d)-*generator* G is a function that maps $\{0,1\}^l$ to $[m]^d$. (Note that the output length of G is $d\lceil \log m \rceil$.) We call l the *input-length* of G. G is said to ϵ-*fool* the (m,d)-rectangles if G is $(\mathcal{R}(m,d), 2^{\{0,1\}^l}, \epsilon)$-good, or equivalently, $image(G)$ is an ϵ-discrepancy set for $\mathcal{R}(m,d)$; we call such a generator G *efficient* if it is in $TS(d \log m/\epsilon, l)$, i.e., if it is computable in time polynomial in the length of the output over ϵ, and in space linear to the length of the input.

We can see now to efficiently construct a small-sized ϵ-discrepancy set for $\mathcal{R}(m,d)$, where by *efficient construction* we mean that the construction time is polynomial in the size of the output set, it suffices to construct an efficient (l,m,d)-generator that ϵ-fools the (m,d)-rectangles with small l. In the next section we will present such a construction.

2.8 The INW Generator for Path Networks

Our construction will make use of a pseudorandom generator introduced in [INW94] which applies to the following communication model.

Suppose there are d processors $p_1, \ldots, p_d$ connected by a path. Each processor p_i receives an input $x_i \in [m]$. They then follow some communication protocol Π in which each processor can send messages to adjacent processors (where the protocol specifies the messages sent by each p_i depending on its input x_i and the messages it has received so far). Eventually the protocol terminates with processor p_d either in an "accept" or a "reject" state. We will call such a protocol an (m,d)-protocol. The *accepting set* $ACC(\Pi)$ of the protocol is the set of inputs $(x_1, \ldots, x_d) \in [m]^d$ which cause p_d to accept. The complexity of the protocol is the maximum over all the inputs $(x_1, \ldots, x_d)$ and processors p_i of the number of bits sent by p_i on input $(x_1, \ldots, x_d)$.

An (l,m,d)-generator G is said to ϵ-fool an (m,d)-protocol Π if

$$|Pr_{y \in \{0,1\}^l}[G(y) \in ACC(\Pi)] - \frac{|ACC(\Pi)|}{m^d}| \leq \epsilon.$$

The following theorem is a restatement of a result in [INW94].

Theorem 2.2 *For each positive integer m, d, c and any $0 < \epsilon \leq 1$, there is an explicit construction of an efficient (l,m,d)-generator that ϵ-fools all (m,d)-protocols of complexity at most c with $l = O(\log m + \log d(c + \log d + \log \epsilon^{-1}))$.*

3 The Generator Construction

In this section we present the construction of our pseudorandom generator for combinatorial rectangles and prove the following:

Theorem 3.1 *Let m, d be positive integers and let $0 < \epsilon \leq 1$. Then for some $l = O(\log m + \log d + \log^2 \epsilon^{-1})$, there is an explicit construction of an efficient (l,m,d)-generator that ϵ-fools the (m,d)-rectangles. Consequently, there is an efficient construction of an ϵ-discrepancy set for (m,d)-rectangles of size polynomial in m, d and $\epsilon^{\log \epsilon}$.*

In particular, in the case that $\epsilon = 2^{-O(\sqrt{\log md})}$, the size of the discrepancy set in our construction is polynomial in m and d.

3.1 The Overview of the Construction

For this discussion, let us fix integers m, d and a real $0 < \epsilon \leq 1$. We want to construct an efficient (l,m,d)-generator that ϵ-fools the (m,d)-rectangles with $l = O(\log m + \log d + \log^2 \epsilon^{-1})$. The starting point of our construction is the pseudorandom generator for communication networks of [INW94].

Any (m,d)-rectangle $R = R_1 \times \ldots \times R_d$ can be naturally visualized as an accepting set of an (m,d)-protocol Π of complexity 1 in the following way: Let $p_1, \ldots, p_d$ be d processors in a path network as defined in Section 2.8. On input $x = (x_1, \ldots, x_d) \in [m]^d$ to the network, for each $1 \leq i \leq d$ the processor p_i receives the i-th coordinate $x_i \in [m]$ and sends 1 bit to p_{i+1} such that, it sends 1 if and only if it receives a 1 from p_{i-1} and at the same time $x_i \in R_i$, where we assume

that p_1 always gets 1 from an imaginary p_0 and the bit sent by p_d is the output of the protocol. So p_d accepts x if and only if $x \in R$. That is, $R = ACC(\Pi)$. Also it is clear that the complexity of the protocol is 1. Now by Theorem 2.2 we have:

Corollary 3.1 *Let m, d be integers, $0 < \epsilon \leq 1$, and let $l = O(\log m + \log d(\log d + \log \epsilon^{-1})$. Then there is an explicit construction of an efficient (l, m, d)-generator G^* that is $(\mathcal{R}(m, d), 2^{\{0,1\}^l}, \epsilon)$-good.*

Remark: With a more careful analysis for the special case of dealing with (m, d)-protocols of complexity 1, we can strengthen the above result to have $l = \lceil \log m \rceil + 2\lceil \log d \rceil (\lceil \log d \rceil + \lceil \log \epsilon^{-1} \rceil)$.

With respect to what we need, the shortcomings of G^* are that the dependence of l on $\log d$ is not linear and that the dependence of l on d and ϵ^{-1} is multiplicative but not additive. On the other hand, if we apply generator G^* to (m', d')-rectangles for some m', d' where d' depends polynomially *only* on ϵ, then the input-length we need for G^* in this case is $O(\log m' + \log^2 \epsilon^{-1})$. Intuitively, what this observation suggests is that if we could first construct a function family $\mathcal{F}^*$ of "small" size that reduces the problem for (m, d)-rectangles to the problem for (m', d')-rectangles where m' is polynomial in m, d, ϵ^{-1} (thus $\log m'$ is linear in $\log m, \log d$ and $\log \epsilon^{-1}$) and, importantly, d' is polynomial in ϵ^{-1}, then G^* for the latter problem would have short input-length $O(\log m + \log d + \log^2 \epsilon^{-1})$ and so, by composing $\mathcal{F}^*$ and G^*, we could obtain a "small"-sized family $\mathcal{F}$ of functions with short input-length. Therefore, the unification of $\mathcal{F}$ would provide a desired generator.

More precisely, what we will do for our construction is the following: For some $m' = poly(m, d, \epsilon^{-1})$ and $d' = poly(\epsilon^{-1})$, we first construct a family $\mathcal{F}^*$ of functions mapping $[m']^{d'}$ to $[m]^d$ that is $(\mathcal{R}(m, d), \mathcal{R}(m', d'), 2\epsilon/3)$-good, where the size of $\mathcal{F}^*$ is 2^s for some $s = O(\log d + \log \epsilon^{-1})$. Furthermore, $\mathcal{F}^*$ is indexable and is in $TS(d \log m/\epsilon, \log m + \log d + \log \epsilon^{-1})$. Then we let G^* be the (l', m', d')-generator given by Corollary 3.1 such that G^* is $(\mathcal{R}(m', d'), 2^{\{0,1\}^{l'}}, \epsilon/3)$-good. Thus $l' = O(\log m + \log d + \log^2 \epsilon^{-1})$ and moreover, G^* is in $TS(d' \log m'/\epsilon, l')$ because of its efficiency. Define $\mathcal{F} = \mathcal{F}^* \circ G^*$. By Lemma 2.2, $\mathcal{F}$ is a family of (l', m, d)-generators of size 2^s that is $(\mathcal{R}(m, d), 2^{\{0,1\}^{l'}}, \epsilon)$-good. Moreover, by Proposition 2.3 $\mathcal{F}$ is indexable and is in $TS(d \log m/\epsilon, l')$. Now Proposition 2.4 tells us that $G_{\mathcal{F}}$ is an efficient (l, m, d)-generator that ϵ-fools the (m, d)-rectangles, where $l = l' + s = O(\log m + \log d + \log^2 \epsilon^{-1})$. Thus to accomplish our generator construction as stated in Theorem 3.1, it suffices to build the family $\mathcal{F}^*$.

The family $\mathcal{F}^*$ in our construction is a composition of a sequence of three families of functions $\mathcal{F}_i, 0 \leq i \leq 2$, with each $\mathcal{F}_i$ mapping V_{i+1} to V_i for some V_i. Each function family in the sequence specifies a function reduction that reduces one construction problem to another one with simpler structure. The reduction sequence mainly follows the idea in [LLSZ95] where the problem for general rectangles is first reduced to the problem for PIP rectangles, and then is further reduced to the problem for rectangles whose dimension depends polynomially only on ϵ. One difference between our construction and the one in [LLSZ95] is in the dimension reduction. In [LLSZ95], the error introduced by this reduction can be bounded only from above, which is sufficient for their purposes of constructing (one-sided) hitting sets, but is inadequate for our purposes of (two-sided) discrepancy set construction. One technical contribution of our work is that in the dimension reduction, we reduce the dimension to polynomial in ϵ^{-1} while keeping the error bounded small from both sides. The details of the constructions will be given in the next few subsections. The properties that this sequence $\mathcal{F}_i$ satisfies are summarized below.

For $0 \leq i \leq 3$, each V_i is of the form $[m_i]^{d_i}$ for some m_i, d_i such that, with possible exceptions on $m_0 = m$ and $d_0 = d$, all the other m_i and d_i are integer powers of 2. (Note that the m', d' in the above description are now m_3 and d_3, respectively.)

Family $\mathcal{F}_0$: As a preliminary for the next two constructions, the purpose of this function family is to reduce the problem for (m, d)-rectangles where m, d are arbitrary to the problem for (m', d')-rectangles where m', d' are integer powers of 2.

For $m_1 = O(m_0 d_0/\epsilon)$ and $d_1 = O(d_0)$, $\mathcal{F}_0$ is a family of one single function from V_1 to V_0 that is $(\mathcal{R}(m, d), \mathcal{R}(m_1, d_1), \frac{\epsilon}{3})$-good. Moreover, it is in $TS(d_1 \log m_1, \log m_1 + \log d)$. We will call $\mathcal{F}_0$ the *preliminary reduction function family.*

Family $\mathcal{F}_1$: To accomplish the dimension reduction, it is desirable to deal with PIP rectangles and not general ones. Function family $\mathcal{F}_1$ reduces the problem for general rectangles to the problem for PIP rectangles.

For $m_2 = (\max(m_1, d_1))^2$ and $d_2 = d_1$, $\mathcal{F}_1$ is a family of one single function from V_2 to V_1 that is $(\mathcal{R}(m_1, d_1), PIP\text{-}\mathcal{R}(m_2, d_2), 0)$-good. Moreover, it is in $TS(d_2 \log m_2, \log m_2)$. We will call $\mathcal{F}_1$ the PIP *reduction function family.*

Family $\mathcal{F}_2$: This is the major component of our construction which specifies a function reduction that reduces the PIP rectangle problem to the problem for rectangles whose demension depends polynomially only on ϵ^{-1}.

For $m_3 = m_2$ and $d_3 = O(\epsilon^{-1} ln^2 \epsilon^{-1})$, $\mathcal{F}_2$ is a family of functions mapping V_3 to V_2 that is $(PIP\text{-}\mathcal{R}(m_2, d_2), \mathcal{R}(m_3, d_3), \frac{\epsilon}{3})$-good. The size of $\mathcal{F}_2$ is $(\max(d_2, d_3))^3$. Moreover, $\mathcal{F}_2$ is in $TS((d_2 + d_3)\log m_3, \log(d_2 d_3))$. We will call $\mathcal{F}_2$ the *dimension reduction function family.*

Furthermore, each function family $\mathcal{F}_i$ is indexable.

It is clear from the parameters chosen above that $m' = m_3 = poly(m, d, \epsilon^{-1}), d' = d_3 = poly(\epsilon^{-1})$, and since $2^s = \Pi_{i=0}^{2} |\mathcal{F}_i|$, $s = \log(\max(d_2, d_3))^3 = O(\log d + \log \epsilon^{-1})$. Moreover, by Lemma 2.2, $\mathcal{F}^* = \bigodot_{i=0}^{2} \mathcal{F}_i$ is a family of functions mapping $[m']^{d'}$ to $[m]^d$ that is $(\mathcal{R}(m, d), \mathcal{R}(m', d'), 2\epsilon/3)$-good; and by Proposition 2.3, $\mathcal{F}^*$ is indexable and is in $TS(d \log m/\epsilon, \log m + \log d + \log \epsilon^{-1})$. Thus $\mathcal{F}^*$ is what we needed.

In the rest of this section, we present the constructions of the sequence of function families described above. For the clarity of our presentation, we will first give the constructions of $\mathcal{F}_1, \mathcal{F}_2$, assuming that m_1 and d_1 are integer powers of 2. We will then justify this assumption later in Section 3.4 by presenting the construction of family $\mathcal{F}_0$.

3.2 PIP Reduction Function Family

Let $V_1 = [m_1]^{d_1}$ where both m_1 and d_1 are assumed to be integer powers of 2, and let $V_2 = [m_2]^{d_2}$ where $m_2 = (\max(m_1, d_1))^2$ and $d_2 = d_1$. We construct a family $\mathcal{F}_1$ of functions mapping V_2 to V_1 that is $(\mathcal{R}(m_1, d_1), PIP\text{-}\mathcal{R}(m_2, d_2), 0)$-good. That is, the function family $\mathcal{F}_1$ specifies a function reduction, which we call PIP reduction, that reduces the problem of finding a discrepancy set for (m_1, d_1)-rectangles in V_1 to the problem of finding a discrepancy set for PIP (m_2, d_2)-rectangles in V_2. We want $\mathcal{F}_1$ to be indexable and in $TS(d_2 \log m_2, \log m_2)$ as well.

The construction given here follows from the construction of the PIP-reduction in [LLSZ95]. We present the construction for completeness.

Let H be a pairwise independent hash function family mapping $[d_2]$ to $[m_1]$ of size $(\max(m_1, d_2))^2 = m_2$ obtained by Theorem 2.1. We identify H with $[m_2]$. Define a function $f : V_2 \to V_1$ as follows

$$f(h_1, \ldots, h_{d_2}) = (h_1(1), \ldots, h_{d_2}(d_2)).$$

Let $\mathcal{F}_1 = \{f\}$. Then $\mathcal{F}_1$ is trivially indexable. It is not difficult to check that $\mathcal{F}_1$ is in $TS(d_2 \log m_2, \log m_2)$. Thus it remains to show that $\mathcal{F}_1$ is $(\mathcal{R}(m_1, d_1), PIP\text{-}\mathcal{R}(m_2, d_2), 0)$-good.

Let $R = R_1 \times \ldots \times R_{d_1} \in \mathcal{R}(m_1, d_1)$. For $1 \le i \le d_2 = d_1$, define $R'_i = \{h \in H = [m_2] : h(i) \in R_i\}$. Then $R\Lambda_f = R'_1 \times \ldots \times R'_{d_2} = R'$ is an (m_2, d_2)-rectangle in V_2.

By the definition of H, $\frac{|R'_i|}{m_2} = \Pr_{h \in H}[h(i) \in R_i] = \frac{|R_i|}{m_1}$, which implies that $vol(R') = vol(R)$. Hence, $|E_{f \in \mathcal{F}_1}[vol(R')] - vol(R)| = 0$.

To complete the proof, we now show that R' is PIP. For any $1 \le i < j \le d_2$,

$$\begin{aligned} \frac{|R'_i \cap R'_j|}{m_2} &= \Pr_{h \in H}[h(i) \in R_i \text{ and } h(j) \in R_j] \\ &= \frac{|R_i|}{m_1}\frac{|R_j|}{m_1} \\ &= \frac{|R'_i|}{m_2}\frac{|R'_j|}{m_2}. \end{aligned}$$

3.3 Dimension Reduction Function Family

The purpose of function family $\mathcal{F}_2$ is to specify a function reduction, which we call dimension reduction, that reduces the problem of finding a discrepancy set for PIP (m_2, d_2)-rectangles in V_2 to the problem of finding a discrepancy set for rectangles whose dimension depends only on ϵ.

Let $V_3 = [m_3]^{d_3}$ where $m_3 = m_2$ and $d_3 = 2^{\lceil \log(8\epsilon^{-1} \ln^2(8\epsilon^{-1}))\rceil}$ (thus $d_3 \ge \frac{8 \ln^2 \frac{8}{\epsilon}}{\epsilon}$). Let H be a 3-wise independent hash function family mapping $[d_2]$ to $[d_3]$ of size $(\max(d_2, d_3))^3$ given by Theorem 2.1. For each $h \in H$, we define a function $f_h : V_3 \to V_2$ as follows:

$$f_h(p_1, \ldots, p_{d_3}) = (p_{h(1)}, \ldots, p_{h(d_2)}).$$

Let $\mathcal{F}_2 = \{f_h : h \in H\}$. Since H is indexable and is in $TS(\log d_2 d_3, \log d_2 d_3)$ by Theorem 2.1, $\mathcal{F}_2$ is indexable and is in $TS((d_2 + d_3)\log m_3, \log d_2 d_3)$. We want to show that $\mathcal{F}_2$ is $(PIP\text{-}\mathcal{R}(m_2, d_2), \mathcal{R}(m_3, d_3), \frac{\epsilon}{3})$-good.

Let m, d be integers and $R = R_1 \times R_2 \times \ldots \times R_d \in \mathcal{R}(m, d)$. For the proof we need the following notations. For $1 \le i \le d$, let $\beta_i = \frac{|R_i|}{m}$, $\delta_i = 1 - \beta_i$ and for a subset $S \subseteq [d]$, we denote

$$T_S = \cap_{i \in S} R_i, \quad \gamma(S) = \frac{|T_S|}{m_3}, \quad \pi(S) = \prod_{i \in S} \beta_j,$$

$$\nu(S) = \sum_{i \in S} \delta_i, \quad \mu(S) = \sum_{i,j \in S, i<j} \delta_i \delta_j, \text{ and}$$

$$\tau(S) = \sum_{i,j,k \in S, i<j<k} \delta_i \delta_j \delta_k.$$

Fix any $R = R_1 \times \ldots \times R_{d_2} \in PIP\text{-}\mathcal{R}(m_2, d_2)$. It is easy to verify that for every $h \in H$, $R\Lambda_{f_h} = T_{h^{-1}(1)} \times \ldots \times T_{h^{-1}(d_3)} = R^{(h)} \in \mathcal{R}(m_3, d_3)$. We are left to show that $|E_{f\in\mathcal{F}_2}[vol(R\Lambda_f)] - vol(R)| = |E_{h\in H}[vol(R^{(h)})] - vol(R)| \leq \frac{\epsilon}{3}$.

We want to estimate $|E_{h\in H}[vol(R^{(h)})] - vol(R)|$. The main lemma we will need is the following:

Lemma 3.1 *Let m, d, d' be integers and H be a 3-wise independent hash function family mapping $[d]$ to $[d']$. Then for any $R \in PIP\text{-}\mathcal{R}(m, d)$ and for every $h \in H$,*

$$\begin{aligned} & vol(R) - \sum_{q\in[d']} \mu(h^{-1}(q)) \\ \leq\ & vol(R^{(h)}) \\ \leq\ & vol(R) + \sum_{q\in[d']} \tau(h^{-1}(q)). \end{aligned}$$

Let us first assume that the lemma holds and see that the estimation follows. By the lemma, we have that

$$\begin{aligned} & |E_{h\in H}[vol(R^{(h)})] - vol(R)| \\ \leq\ & E_{h\in H}[\sum_{q\in d_3} \mu(h^{-1}(q))] + E_{h\in H}[\sum_{q\in d_3} \tau(h^{-1}(q))]. \end{aligned}$$

For the estimation, we consider two cases depending on the volume of R. The first case is when $vol(R) \geq \frac{\epsilon}{8}$. Then $\frac{\epsilon}{8} \leq \prod_{i\in[d_2]} \beta_i = \prod_{i\in[d_2]}(1-\delta_i) \leq \exp(-\sum_{i\in[d_2]} \delta_i)$, which implies that $\sum_{i\in[d_2]} \delta_i \leq \ln \frac{8}{\epsilon}$.

$$\begin{aligned} & E_{h\in H}[\sum_{q\in[d_3]} \tau(h^{-1}(q))] \\ =\ & \sum_{q\in[d_3]} E_{h\in H}[\sum_{i,j,k\in h^{-1}(q), i<j<k} \delta_i\delta_j\delta_k] \\ =\ & \sum_{q\in[d_3]} \sum_{i,j,k\in[d_2], i<j<k} \Pr_{h\in H}[h(i)=h(j)=h(k)=q]\delta_i\delta_j\delta_k \\ =\ & \sum_{q\in[d_3]} (\frac{1}{d_3})^3 \sum_{i,j,k\in[d_2], i<j<k} \delta_i\delta_j\delta_k \\ \leq\ & (\frac{1}{d_3})^2 (\sum_{i\in[d_2]} \delta_i)^3 \\ \leq\ & \frac{\epsilon^2}{64\ln^4(\frac{8}{\epsilon})} \ln^3(\frac{8}{\epsilon}) \\ \leq\ & \frac{\epsilon}{64} \end{aligned}$$

where the third equality is by the choice of H as a 3-wise independent hash function family. Similarly, one can show that $E_{h\in H}[\sum_{q\in d_3} \mu(h^{-1}(q))] \leq \frac{\epsilon}{8}$. Therefore $|E_{h\in H}[vol(R^{(h)})] - vol(R)| \leq \frac{\epsilon}{3}$ in this case.

In the second case we have $vol(R) < \frac{\epsilon}{8}$. We let R' be another PIP (m_2, d_2)-rectangle containing R such that $\frac{\epsilon}{8} \leq vol(R') \leq \frac{\epsilon}{4}$. (The existence of such an R' is not difficult to justify.) Since R' contains R, then for every $h \in H$, $R'^{(h)} = R'\Lambda_{f_h}$ contains $R^{(h)}$, and thus $E_{h\in H}[vol(R^{(h)})] \leq E_{h\in H}[vol(R'^{(h)})]$. By the right inequality in Lemma 3.1 and the above analysis, we have that

$$\begin{aligned} E_{h\in H}[vol(R'^{(h)})] &\leq vol(R) + E_{h\in H}[\sum_{q\in[d_3]} \tau(h^{-1}(q))] \\ &\leq \frac{\epsilon}{4} + \frac{\epsilon}{64} \\ &< \frac{\epsilon}{3}. \end{aligned}$$

Since volumes are always nonnegative, $|E_{h\in H}[vol(R^{(h)})] - vol(R)| \leq \frac{\epsilon}{3}$ in this case as well.

Now it remains to prove Lemma 3.1.

Proof of Lemma 3.1:

Let $R = R_1 \times R_2 \times \ldots \times R_d$. We first show two preliminary facts.

Proposition 3.1 *For any $S \subseteq [d]$,*

$$\pi(S) - \mu(S) \leq \gamma(S) \leq \pi(S) + \tau(S).$$

Proof: Let $S \subseteq [d]$ be arbitrary. By definition, $\gamma(S) = \frac{|\cap_{i\in S} R_i|}{m} = 1 - \frac{|\cup_{i\in S} \bar{R}_i|}{m}$. Applying inclusion-exclusion, we have that

$$\gamma(S) \leq 1 - \sum_{i\in S} \frac{|\bar{R}_i|}{m} + \sum_{i,j\in S, i<j} \frac{|\bar{R}_i \cap \bar{R}_j|}{m} = 1 - \nu(S) + \mu(S),$$

where the equality follows from the fact that R is PIP, and that

$$\gamma(S) \geq 1 - \sum_{i\in S} \frac{|\bar{R}_i|}{m} = 1 - \nu(S).$$

Since it is easy to prove by induction on the size of S that

$$1 - \nu(S) + \mu(S) - \tau(S) \leq \pi(S) \leq 1 - \nu(S) + \mu(S),$$

combining this with the above facts about $\gamma(S)$ we draw the conclusion. □

Proposition 3.2 *Let x, y, z be nonnegative real numbers such that $x \leq 1$ and $0 \leq x - y \leq z \leq x + y$. Let x', y', z' be nonnegative real numbers such that $0 \leq x' - y' \leq z' \leq x' + y' \leq 1$. Then $xx' - y' - y \leq zz' \leq xx' + y' + y$.*

Proof: By assumptions, $(x-y)(x'-y') \leq zz' \leq (x+y)(x'+z')$. Also we have that $(x-y)(x'-y') \geq x(x'-y')-y \geq xx'-y'-y$ and that $(x+y)(x'+y') \leq x(x'+y')+y \leq xx'+y'+y$. The proof is complete. □

Now we prove by induction the following inequalities: for each $1 \leq t \leq d'$,

$$\begin{aligned} & \prod_{q=1}^{t} \pi(h^{-1}(q)) - \sum_{q=1}^{t} \mu(h^{-1}(q)) \\ \leq & \prod_{q=1}^{t} \gamma(h^{-1}(q)) \\ \leq & \prod_{q=1}^{t} \pi(h^{-1}(q)) + \sum_{q=1}^{t} \tau(h^{-1}(q)). \end{aligned}$$

The case where $t = 1$ follows from Proposition 3.1. For the inductive step, it is easy to check that Proposition 3.2 establishes the inequalities.

Finally, the lemma follows by substituting d' for t in the above inequalities. □

3.4 Preliminary Reduction Function Family

We justify the assumption that m_1 and d_1 are integer powers of 2.

We have $V_0 = [m]^d$ where m, d are arbitrary integers. For some $m_1 = O(md/\epsilon)$ and $d_1 = O(d)$ that are integer powers of 2, we will construct a function family $\mathcal{F}_0$ that is in $TS(d_1 \log m_1, \log m_1 + \log d)$ and is $(\mathcal{R}(m,d), \mathcal{R}(m_1,d_1), \frac{\epsilon}{3})$-good.

The function family we shall construct contains a single function $f_{m,d}$ which is obtained by composing two other functions: f_m and f_d. Let $W = [m_1]^d$. f_m is a function from W to V_0 that is in $TS(d \log m_1, \log m_1)$ and is $(\mathcal{R}(m,d), \mathcal{R}(m_1,d), \frac{\epsilon}{3})$-good, and f_d is a function from V_1 to W that is in $TS(d_1 \log m_1, \log d)$ and is $(\mathcal{R}(m_1,d), \mathcal{R}(m_1,d_1), 0)$-good. Thus $f_{m,d} = f_m \circ f_d$ is as needed.

3.4.1 The Function f_m

Let $c_0 = \lceil \frac{8d}{\epsilon} \rceil$ and let $m_1 = 2^{\lceil \log c_0 m \rceil}$. Define the function $f_m : W \to V_0$ as follows: $f(p_1, \ldots, p_d) = (q_1, \ldots, q_d)$ if and only if for each $1 \leq i \leq d$, $p_i \equiv q_i \pmod m$. Clearly, f_m is in $TS(d \log m_1, \log m)$. We want to show that f_m is $(\mathcal{R}(m,d), \mathcal{R}(m_1,d), \frac{\epsilon}{3})$-good.

Fix any $R = R_1 \times \ldots \times R_d \in \mathcal{R}(m,d)$. Let $R'_i = \{p \in [m_1] : (p \pmod m) \in R_i\}$. Then $R\Lambda_f = R'_1 \times \ldots \times R'_{d_2} = R' \in \mathcal{R}(m_1,d)$. All that remains is to show that $|vol(R') - vol(R)| \leq \frac{\epsilon}{4}$.

Let c be the unique integer such that $cm \leq m_1 < (c+1)m$. Then it is clear that $c \geq c_0$. Since $\lfloor \frac{m_1}{m} \rfloor |R_i| \leq |R'_i| \leq \lceil \frac{m_1}{m} \rceil |R_i|$, we have that

$$(\frac{cm}{m_1})^d vol(R) \leq vol(R') \leq (\frac{(c+1)m}{m_1})^d vol(R).$$

Now

$$(\frac{cm}{m_1})^d \geq (\frac{c}{c+1})^d \geq (1 - \frac{\epsilon}{8d})^d \geq (\frac{1}{4})^{\frac{\epsilon}{8}} \geq 1 - \frac{\epsilon}{4},$$

which implies that $(1 - \frac{\epsilon}{4}) vol(R) \leq vol(R')$, and

$$(\frac{(c+1)m}{m_1})^d \leq (\frac{c+1}{c})^d \leq (1 + \frac{\epsilon}{8d})^d \leq 2^{\frac{\epsilon}{8}} \leq 1 + \frac{\epsilon}{8},$$

which implies that $vol(R') \leq (1 + \frac{\epsilon}{8}) vol(R)$. Therefore $|vol(R') - vol(R)| \leq \frac{\epsilon}{3}$ and this concludes the proof.

3.4.2 The Function f_d

Let $d_1 = 2^{\lceil \log d \rceil}$ and let $V_1 = [m_1]^{d_1}$. Define the function $f_d : V_1 \to W$ as follows: $f_d(q_1, \ldots, q_{d_1}) = (q_1, \ldots, q_d)$. That is, f_d simply outputs the first d coordinates of the input and discard the rest. Obviously, f_d is in $TS(d_1 \log m_1, \log d)$.

To see that f_d is $(\mathcal{R}(m_1,d), \mathcal{R}(m_1,d_1), 0)$-good, we observe that for any combinatorial rectangle $R = R_1 \times \ldots \times R_d \in \mathcal{R}(m_1,d)$,

$$R\Lambda_{f_d} = R_1 \times \ldots \times R_d \times [m_1]^{d_1 - d} \in \mathcal{R}(m_1,d_1)$$

and $R\Lambda_{f_d}$ has the same volume as R has.

References

[ABI86] N. Alon, L. Babai, A. Itai, *A fast and simple randomized parallel algorithm for the Maximal Independent Set Problem*, J. Algorithms 7, pp 567-583, 1986.

[AGHP90] N. Alon, O, Goldreich, J. Hastad, R. Peralta, *Simple constructions of almost k-wise independent random variables*, Random Structures and Algorithms 3(3), pp 289-303, 1992.

[AW96] R. Armoni and A. Wigderson, *Pseudo-randomness for space-bounded computations*, Unpublished manuscript, 1996.

[BC87] J. Beck and W. Chen, *Irregularities of distribution*, Cambridge University Press, 1987.

[CW79] L. Carter and M. Wegman, *Universal hash functions*, J. Comp. and Syst. Sci., 18(2):143-154, 1979.

[CG89] B. Chor and O. Goldreich, *On the power of two-point based sampling*, J. of Complexity, vol 5, 1989, pp 96-106.

[EGLNV92] G. Even, O. Goldreich, M. Luby, N. Nisan, and B. Velicković, *Approximations of general independent distributions*, Proceedings of the 24th Annual ACM Symposium on Theory of Computing, pages 10–16. ACM, 1992.

[INW94] R. Impagliazzo, N. Nisan, and A. Wigderson, *Pseudorandomness for network algorithms* ACM Symposium on Theory of Computing (STOC), 1994.

[KK94] D. Karger and D. Koller, *(De)randomized construction of small sample spaces in NC*, 35th IEEE FOCS, pp 252-263, 1994.

[KW84] R. Karp and A. Wigderson, *A fast parallel algorithm for the maximal independent set problem*, Proceedings of the 16th Annual ACM Symposium on Theory of Computing, 1984.

[KM93] D. Koller and N. Megiddo, *Finding small sample spaces satisfying given constraints*, Proceedings of the 25th Annual ACM Symposium on Theory of Computing, pp 268-277, 1993.

[LLSZ95] N. Linial, M. Luby, M. Saks, and D. Zuckerman, *Efficient construction of a small hitting set for combinatorial rectangles in high dimension*, 1995.

[Lub85] M. Luby, *A simple parallel algorithm for the maximal independent set problem*, SIAM J. Comput. 15(4), pp 1036-1053, 1986.

[LV91] M. Luby, B. Velickovic, *On deterministic approximation of DNF*, 23rd ACM STOC, pp 430-438, 1991.

[LVW93] M. Luby, B. Velickovic and A. Wigderson, *Deterministic Approximate Counting of Depth–2 Circuits*, Proc. of the 2nd ISTCS (Israeli Symposium on Theoretical Computer Science), pp. 18–24, 1993.

[NN90] J. Naor, M. Naor, *Small-bias probability spaces: efficient constructions and applications*, SIAM J. Comput. 22(4), pp 838-856, 1990.

[Nie] H. Niederreiter, *Constructions of low-discrepancy point sets and sequences*, manuscript and lecture notes.

[Nis91] N. Nisan *Pseudo-random bits for constant depth circuits*, Combinatorica 11 (1), 63-70, 1991

[Nis90] N. Nisan, *Pseudorandom generators for space-bounded computation*, Proc. 22nd ACM Symposium on Theory of Computing, 1990, pp. 204-212.

[NZ93] N. Nisan and D. Zuckerman, *More Deterministic Simulation in Logspace*, Proc. ACM Symposium on Theory of Computing, 1993, pp. 235–244.

[Sch92] L. Schulman, *Sample spaces uniform on neighborhoods*, 24th ACM STOC, pp 17-25, 1992.