

# Gradual Small-Bias Sample Spaces

Avraham Ben-Aroya\*      Gil Cohen†

## Abstract

A  $(k, \varepsilon)$ -biased sample space is a distribution over  $\{0, 1\}^n$  that  $\varepsilon$ -fools every nonempty linear test of size at most  $k$ . Since they were introduced by Naor and Naor (SIAM J. Computing, 1993), these sample spaces have become a central notion in theoretical computer science with a variety of applications. When constructing such spaces, one usually attempts to minimize the seed length as a function of  $n$ ,  $k$  and  $\varepsilon$ . Given such a construction, if we reverse the roles and consider a fixed seed length, then the smaller we pick  $k$ , the better the bound on the bias  $\varepsilon$  becomes. However, once the space is constructed we have a *single* bound on the bias of all tests of size at most  $k$ .

In this work we initiate the study of a new pseudorandom object, which we call a *gradual*  $(k, \varepsilon)$ -biased sample space. Roughly speaking, this is a sample space that  $\varepsilon$ -fools linear tests of size *exactly*  $k$  and moreover, the bound on the bias for linear tests of size  $i \leq k$  decays as  $i$  gets smaller. We show how to construct gradual  $(k, \varepsilon)$ -biased sample spaces of size comparable to the (non-gradual) spaces constructed by Alon *et al.* (Random Structures and Algorithms, 1992), and prove a lower bound on their size. Our construction is based on the lossless expanders of Guruswami *et al.* (J. ACM, 2009), combined with the Quadratic Character Construction of Alon *et al.* (Random Structures and Algorithms, 1992)

---

\*Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: [avraham.ben-aroya@weizmann.ac.il](mailto:avraham.ben-aroya@weizmann.ac.il).

†Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: [gil.cohen@weizmann.ac.il](mailto:gil.cohen@weizmann.ac.il). Research supported by Israel Science Foundation (ISF) grant.

# 1 Introduction

An  $\varepsilon$ -biased sample space  $S$  over  $\{0, 1\}^n$  is a sample space with the following property: for every nonempty  $T \subseteq [n]$ , the random variable  $s_T \triangleq \bigoplus_{i \in T} s_i$ , where  $s$  is sampled from  $S$ , has bias at most  $\varepsilon$ . In other words, a sample space is  $\varepsilon$ -biased if it  $\varepsilon$ -fools every nontrivial linear test. When it is not desired or not important to specify  $\varepsilon$ , one usually refers to such a sample space as a *small-bias sample space*.

The notion of a small-bias sample space was introduced in the seminal paper of Naor and Naor [NN93] and has become a fundamental notion in theoretical computer science, with a variety of applications [Nao92, BNS92, NN93, HPS93, AR94, MW02, BSSVW03, VW08, Vio09, GMR<sup>+</sup>12].

Several explicit constructions of small-bias sample spaces that attempt to minimize the sample space size in terms of  $n$  and  $\varepsilon$  are known [AGHP92, ABN<sup>+</sup>92, NN93, BT09]. These constructions give incomparable sizes. Unfortunately, all known constructions fall short from achieving sample spaces of size  $O(n/\varepsilon^2)$ , which are guaranteed to exist by a simple probabilistic argument. Another research direction, which this work falls into, studies variations and generalizations of small-bias sample spaces [AIK<sup>+</sup>90, RSW93, EGL<sup>+</sup>92, AM95, MST06, Shp06].

A relaxation of the notion of a small-bias sample space requires only that *small* linear tests will be fooled. Formally, a  $(k, \varepsilon)$ -biased sample space is a sample space  $S$  over  $\{0, 1\}^n$  such that for every nonempty  $T \subseteq [n]$  of size at most  $k$ , the random variable  $s_T$  has bias at most  $\varepsilon$ , where again  $s$  is sampled from  $S$ . The advantage of this relaxed notion is that fooling only small tests, rather than every nontrivial test, can be achieved by much smaller sample spaces. The original motivation for studying  $(k, \varepsilon)$ -biased sample spaces was to obtain almost  $k$ -wise independent random variables. However,  $(k, \varepsilon)$ -biased sample spaces had proved to be useful in their own right, and found several applications [SZ94, Raz05, CRS12].

Naor and Naor [NN93] gave a general method for constructing  $(k, \varepsilon)$ -biased sample spaces from  $\varepsilon$ -biased sample spaces. Their method yields  $(k, \varepsilon)$ -biased sample spaces that are exponentially smaller in terms of  $n$  than what is possible for  $\varepsilon$ -biased sample spaces. In terms of seed-length, they showed that a seed of length  $O(\log k + \log \log n + \log \varepsilon^{-1})$  is sufficient in order to fool tests of size  $k$ , while it is known that a seed of length  $\Omega(\log n + \log \varepsilon^{-1})$  is necessary in order to fool every nontrivial linear test (see, e.g. [AGHP92, Alo09]).

**Gradual small-bias sample spaces.** Consider two pairs  $(k_1, \varepsilon_1)$  and  $(k_2, \varepsilon_2)$  such that

$$s = \log k_1 + \log \varepsilon_1^{-1} = \log k_2 + \log \varepsilon_2^{-1}.$$

Potentially, one could hope that a seed of length  $O(s + \log \log n)$  would be sufficient to  $\varepsilon_1$ -fool tests of size  $k_1$  and *simultaneously* to  $\varepsilon_2$ -fool tests of size  $k_2$ . In other words, we are considering a  $(k, \varepsilon)$ -biased sample space that has the following property: for tests of size  $t < k$ , the “spare”  $\log k - \log t$  bits of the seed are utilized to reduce the bias. In this paper we initiate the study of such sample spaces, which have a better bound on the bias for smaller tests.

**Definition 1.1.** A sample space  $S$  over  $\{0, 1\}^n$  is called gradual  $(k, \varepsilon)$ -biased if for every nonempty  $T \subseteq [n]$  of size at most  $k$ ,

$$\left| \mathbb{E}_{s \sim S} [(-1)^{\sum_{i \in T} s_i}] \right| \leq \varepsilon \cdot \frac{|T|}{k}.$$

A few words about the definition are in order. First, note that when  $T$  is of size exactly  $k$ , the bound on the bias is  $\varepsilon$ , i.e., a gradual  $(k, \varepsilon)$ -biased sample space is, in particular,  $(k, \varepsilon)$ -biased. One may consider a more general definition, which allows an arbitrary decaying function  $\phi : \mathbb{N} \rightarrow [0, 1]$  as the bound on the bias (say,  $\phi(|T|) = \varepsilon \cdot (|T|/k)^d$  for some parameter  $d$ ). We choose this function to be  $\varepsilon \cdot |T|/k$  in the definition and discuss a more general definition in Section 4.

## 1.1 Motivation

Besides being a natural generalization of small-bias sample spaces, the study of gradual small-bias sample spaces was partially motivated by an approach for constructing rigid matrices, introduced recently by Alon and Cohen [AC12], which we now briefly describe. Let  $\rho \in (0, 1)$  be an arbitrary constant. For a subspace  $U \in \mathbb{F}_2^n$ , define the polynomial

$$p_U(x) = \frac{1}{W_\rho(U)} \cdot \sum_{u \in U} \rho^{|u|} \cdot (-1)^{\langle u, x \rangle},$$

where  $W_\rho(U) = \sum_{u \in U} \rho^{|u|}$ . We note that this is indeed a polynomial when working over  $\{1, -1\}$  rather than over  $\{0, 1\}$ . Moreover, it can be shown to be positive.

In [AC12] it is shown that the problem of constructing rigid matrices is reduced to the problem of constructing a small set  $S \subset \mathbb{F}_2^n$  such that for every subspace  $U$  of dimension  $n/2$ , the quantity  $\mathbb{E}_{s \sim S} [p_U(s)]$  is small. Using this reduction, the authors concluded that any small-bias sample space induces a rigid set that matches the best known construction [APY09].

Roughly speaking, proving this conclusion (that is, showing that  $\mathbb{E}_{s \sim S} [p_U(s)]$  is small where  $S$  is a small bias sample space) amounts to the following calculation. If  $S$  is an  $\varepsilon$ -biased sample space, then, for a suitable  $\rho$ ,

$$\begin{aligned} \mathbb{E}_{s \sim S} [p_U(s)] &= \frac{1}{W_\rho(U)} \cdot \sum_{u \in U} \rho^{|u|} \cdot \mathbb{E}_{s \sim S} [(-1)^{\langle u, s \rangle}] \\ &\leq \frac{1}{W_\rho(U)} + \frac{1}{W_\rho(U)} \cdot \sum_{u \in U \setminus \{0\}} \rho^{|u|} \cdot \varepsilon \\ &< \frac{1}{W_\rho(U)} + \varepsilon = O(\varepsilon). \end{aligned}$$

That is, by using a small-bias sample space  $S$ , one can bound the expectation that appears in each summand (except for  $u = 0$ ) by  $\varepsilon$ . However, this does not take advantage of the coefficient  $\rho^{|u|}$ . We note that when  $|u|$  is large this coefficient is small, and thus we can afford  $\mathbb{E}_{s \sim S} [(-1)^{\langle u, s \rangle}]$  to be large for such  $u$ . That is, it is enough that the desired sample space  $S$  will have a weaker promise on the bias for heavy linear tests. In other words, one is better off taking  $S$  to be a gradual small-bias sample space.

Replacing  $S$  with an appropriate gradual small-bias sample space results in an improvement to the rigidity problem for a special family of subspaces, in which the dual subspace contains many low-weight vectors.

More generally, it is known, and easy to see, that a small-bias sample space is a pseudorandom generator for classes of functions that have small Fourier  $\ell_1$ -norm (and in fact, this is close to characterize the classes of functions that are fooled by small-bias sample spaces [Baz09, DETT10]). The same argument shows that a gradual small-bias sample space is a pseudorandom generator for classes of functions that have a more relaxed condition on their Fourier coefficients. Namely, the quantity that should be small is not  $\ell_1(\widehat{f}) = \sum_x |\widehat{f}(x)|$ , but rather  $\sum_x |\widehat{f}(x)| \cdot \phi(|x|)$  for some function  $\phi : \mathbb{N} \rightarrow [0, 1]$  that determines the gradual behavior of the sample space.

Another neat perspective on gradual small-bias sample spaces is observed by looking at their Fourier spectrum. A (non gradual) small-bias sample space  $S$  induces a function  $f : \{0, 1\}^n \rightarrow \mathbb{R}$  with small support and bounded Fourier coefficients. If  $S$  is also gradual, a stronger property holds, namely, the bound on the Fourier coefficients is stronger for coefficients in the lower levels.

## 1.2 Main Result

The following theorem is our main result:

**Theorem 1.2.** *For any integers  $n$  and  $k \leq n$ , for any  $\varepsilon > 0$ , and for any constant  $\delta > 0$ <sup>1</sup> there exists an explicit construction of a gradual  $(k, \varepsilon)$  sample space of size*

$$m = O_\delta \left( \left( \frac{k}{\varepsilon} \right)^{2+\delta} + \left( \frac{\log n}{\log k} \right)^{2+4/\delta} k^{1+\delta} \right),$$

where the  $O_\delta$  hides a multiplicative constant that depends only on  $\delta$ .

Obviously, one can find a value for  $\delta$  that minimizes  $m$  as a function of  $n, k$  and  $\varepsilon$ . However, when no assumptions are made on the relations between  $n, k$  and  $\varepsilon$ , the expression one would get is cumbersome and non-informative. Moreover, when conducting such minimization one can no longer ignore the multiplicative dependency in  $\delta$  that is hidden under the big  $O_\delta$  notation. We therefore choose to specify our bound in the more readable way presented above. Nevertheless, in the following table we consider three natural ranges for  $k$  in terms of  $n$ , and for those we give the minimum value of  $m$  with respect to  $\delta$ . We also make a comparison with the size of the (non-gradual)  $(k, \varepsilon)$ -biased sample space from [AGHP92], which equals to  $m_{\text{AGHP}} = (k\varepsilon^{-1} \log n)^2$ . The comparison is meant to show that using our construction of gradual small-bias sample space, one does not pay much more in the sample space size for having a decaying bound on the bias.<sup>2</sup>

<sup>1</sup>In fact, the construction works without assuming  $\delta$  is constant, and this assumption appears only to simplify the presentation of the theorem. See Theorem 3.1 for a more general statement.

<sup>2</sup>A central building block in our construction is an unbalanced expander. Unfortunately, even the state of the art explicit construction of unbalanced expanders [GUV09] are somewhat far from optimal. When analyzing our construction of gradual small-bias sample spaces using optimal unbalanced expanders, the resulting sample space would have size matching that of the (non-gradual) construction of [AGHP92].

Range of $k$	Sample space size	
	[AGHP92]	Theorem 1.2
$k = n^\gamma$ , for any constant $\gamma \leq 1$	$(k/\varepsilon)^{2+o(1)}$	$O_\delta((k/\varepsilon)^{2+\delta})$ for any constant $\delta > 0$
$k = \log^c n$ , for any constant $c \geq 6$ , and $\varepsilon = \Omega(1)$	$(\log n)^{2(c+1)}$	$(\log n)^{2(c+1)+3}$
$k = O(1)$ and $\varepsilon = \log^{-c} n$ , for any constant $c \geq 1/2$	$(\log n)^{2(c+1)}$	$O_c((\log n)^{2(c+1)})$

### 1.3 Informal Description of the Construction

Our high-level strategy is composed of two steps:

1. Obtaining a gradual  $(n, \varepsilon)$ -biased sample space, over  $\{0, 1\}^n$ .
2. Transforming it into a gradual  $(k, \varepsilon)$ -biased sample space with a shorter seed.

A similar approach was used by [NN93] to construct (non-gradual)  $(k, \varepsilon)$ -biased sample spaces. We now elaborate on each of the steps.

#### 1.3.1 Gradual small-bias sample spaces from quadratic characters

For the first step, we use the Quadratic Character Construction of small-bias sample spaces of [AGHP92], which we now describe.

Let  $q$  be an odd prime power. Denote by  $\mathbb{F}_q$  the finite field with  $q$  elements. The quadratic character  $\chi : \mathbb{F}_q \rightarrow \{-1, 0, 1\}$  is defined as

$$\chi(x) = \begin{cases} 0, & x = 0; \\ 1, & \exists y \in \mathbb{F}_q \setminus \{0\} \text{ such that } x = y^2; \\ -1, & \text{otherwise.} \end{cases}$$

The sample space in this construction consists of  $q$  strings, in correspondence with the elements of  $\mathbb{F}_q$ . A string in the sample space is composed of  $n$  bits, which are indexed by elements from some arbitrarily chosen set  $I \subseteq \mathbb{F}_q$  of size  $|I| = n$ . For  $i \in I$  and  $x \in \mathbb{F}_q$ , the  $i^{\text{th}}$  bit of the  $x^{\text{th}}$  string is given by  $\chi(x + i)$ .<sup>3</sup>

The bias of this construction for linear tests of size  $k$  is essentially the expectation of  $\chi$  over the image of some degree  $k$  polynomial. Weil's Theorem (see Theorem 2.1) bounds precisely expectations of this form. Moreover, the bound this theorem provides is linear in  $k$ , the degree of the aforementioned polynomial. This implies a better bound for smaller tests. It follows that this space is indeed a gradual  $(n, \varepsilon)$ -biased sample space.

<sup>3</sup>Formally, the support of the sample space should be  $\{0, 1\}^n$ . This minor technicality is resolved by mapping  $\pm 1$  to  $\{0, 1\}$ , and 0 arbitrarily.

### 1.3.2 Shortening the seed length

In order to obtain a  $(k, \varepsilon)$ -biased sample space clearly it suffices to construct an  $(n, \varepsilon)$ -biased sample space (since being  $(n, \varepsilon)$ -biased implies being  $(k, \varepsilon)$ -biased). We now describe a cleverer way to transform  $(n, \varepsilon)$ -biased sample spaces into  $(k, \varepsilon)$ -biased ones. This transformation is due to [NN93].

Let  $U_n$  denote the uniform distribution over  $\{0, 1\}^n$ . We say that a linear transformation  $T : \{0, 1\}^n \rightarrow \{0, 1\}^N$  generates a  $k$ -wise independent space if the  $N$  random variables  $(Z_i)_{i=1}^N$  defined by

$$Z_i = T(U_n)_i$$

are  $k$ -wise independent. Suppose  $S$  is an  $(n, \varepsilon)$ -biased sample space over  $\{0, 1\}^n$  and suppose that  $T : \{0, 1\}^n \rightarrow \{0, 1\}^N$  generates a  $k$ -wise independent space. Then, in [NN93] it is shown that the sample space over  $\{0, 1\}^N$  defined by  $T(S)$  is  $(k, \varepsilon)$ -biased. The advantage of this transformation is that it allows  $N$  to be significantly larger than  $n$ , thus shortening the seed length as a function of the output length.

Similar to this approach, we also suggest a general way to transform a gradual  $(n, \varepsilon)$ -biased sample space into a gradual  $(k, \varepsilon)$ -biased one. We mention that the idea is to use a linear transformation  $T : \{0, 1\}^n \rightarrow \{0, 1\}^N$ , which generates a  $k$ -wise independent space, but which is also *sparse*, in the sense that each output bit depends only on a small number of input bits. We claim that in this case, provided that  $S$  is a gradual  $(n, \varepsilon)$ -biased sample space,  $T(S)$  is a gradual  $(k, \varepsilon)$ -biased sample space.

Let us sketch the proof idea. Suppose  $T$  is such a transformation, in which each output bit is a sum of at most  $\ell$  input bits, and suppose  $S$  is a gradual  $(n, \varepsilon)$ -biased sample space. To sample from the new sample space, we first sample  $s$  from  $S$  and then output  $T(s)$ . Consider a linear test  $A(\cdot)$  of size  $r \leq k$ , applied to  $T(s)$ . By the sparsity of  $T$ , it follows that  $A(T(s))$  is a sum of at most  $\ell \cdot r$  bits of  $s$ . Moreover, since  $T$  generates a  $k$ -wise independent space, this sum is not empty. Thus, the bias of  $A(T(s))$  is the bias of some test of weight  $\ell \cdot r$  in the sample space  $S$ , and the claim follows. It might be useful to note that this transformation works even if  $S$  is a gradual  $(\ell \cdot r, \varepsilon)$ -biased sample space (as apposed to  $S$  being a gradual  $(n, \varepsilon)$ -biased sample space).

We present an explicit construction of such a transformation  $T$ , based on expanders (see Section 2.2). The construction that we use is essentially the parity-check matrix of the codes of Sipser and Spielman [SS96] when combined with the unbalanced expanders of [GUV09]. The idea of using unbalanced expanders to preserve sparsity in different context has been used in the literature [Shp06, Vio12].

## 1.4 Organization

In Section 2 we state some preliminary definitions and results that we need. In Section 3 we present a construction of a gradual small-bias sample space and prove Theorem 1.2. In Section 4 we study a more general definition of gradual small-bias sample spaces. In particular, we address the problem of achieving a stronger decay in the bound on the bias, and prove a lower bound on the size of such sample spaces. Section 5 contains concluding remarks and some open problems.

## 2 Preliminaries

All logarithms in this paper are in base 2. For a natural number  $n$  we define  $[n] = \{1, 2, \dots, n\}$ .

### 2.1 Quadratic Characters

We denote by  $\chi_q$  the quadratic character over  $\mathbb{F}_q$ . When the field is understood from the context, we omit the subscript and simply denote this character by  $\chi$ . We use a special case of Weil's Theorem regarding character sums (see e.g., [Sch76]).

**Theorem 2.1** (Weil's Theorem). *Let  $q$  be an odd prime power. Let  $f \in \mathbb{F}_q[x]$  be a degree  $d$  polynomial. Assume that  $f(x) \neq c \cdot g(x)^2$  for any  $c \in \mathbb{F}_q, g \in \mathbb{F}_q[x]$ . Then,*

$$\left| \sum_{x \in \mathbb{F}_q} \chi(f(x)) \right| \leq (d-1)\sqrt{q}.$$

### 2.2 Expanders and Codes

We associate a bipartite graph  $G = (L, R, E)$  with  $|L|$  left-vertices,  $|R|$  right-vertices and left-degree  $d$  with the adjacency function  $G : L \times [d] \rightarrow R$ , where  $G(x, i) = y$  if and only if  $y$  is the  $i$ th neighbor of  $x$ . For a set of left-vertices  $A \subseteq L$  we denote by  $G(A)$  the set of neighbors of  $A$ .

**Definition 2.2.** *A bipartite graph  $G : L \times [d] \rightarrow R$  is a  $k$ -unique-neighbor expander if for any nonempty subset  $A \subseteq L$  of size at most  $k$ , there exists some  $y \in R$  that is adjacent to exactly one vertex in  $A$ .*

**Definition 2.3.** *A bipartite graph  $G : L \times [d] \rightarrow R$  is a  $(\leq k, \alpha)$  expander if for any subset  $A \subseteq L$  of size at most  $k$ ,*

$$|G(A)| \geq \alpha \cdot |A|.$$

We will need the well known fact that a graph whose expansion is greater than half of the degree is also a unique-neighbor expander.

**Fact 2.4.** *If  $G : L \times [d] \rightarrow R$  is a  $(\leq k, \alpha)$  expander for  $\alpha > d/2$  then  $G$  is a  $k$ -unique-neighbor expander.*

**Proof:** Consider a nonempty set of left-vertices  $A \subseteq L$  of size at most  $k$ . The number of outgoing edges from  $A$  is  $|A| \cdot d$ . Suppose that each vertex in  $G(A)$  is adjacent to at least two vertices in  $A$ . This implies,  $|A| \cdot d \geq 2 \cdot |G(A)|$ , which contradicts the fact that  $G$  is a  $(\leq k, \alpha)$  expander for  $\alpha > d/2$ .  $\square$

We will make use of the following expanders, constructed by [GUV09].

**Theorem 2.5** ([GUV09, Theorem 3.2]). *Let  $q$  be a prime power.<sup>4</sup> For every integers  $\ell, r, h \geq 1$  there exists an explicit construction of a graph  $G : [q^\ell] \times [q] \rightarrow [q^{r+1}]$  which is an  $(\leq h^r, q - (\ell - 1)(h - 1)r)$  expander. In particular,  $G$  is an  $h^r$ -unique-neighbor expander when  $q > (\ell - 1)(h - 1)r/2$ .*

In essence our construction uses the error-correcting code whose parity check matrix is defined by the above graph (as in [SS96]). In comparison, the (non-gradual)  $(k, \varepsilon)$ -biased sample spaces of [NN93, AGHP92] use the BCH code. Unlike the BCH code, the expander code is a low-density parity-check code, and this property plays a crucial role in our construction.

### 3 The Construction

In this section we describe our construction of a gradual  $(k, \varepsilon)$ -biased sample space, and prove Theorem 1.2. For simplicity we combine the two conceptual steps that appear in the informal description of the construction (Section 1.3).

Let  $r \geq 2$  be an integer.<sup>5</sup> Let  $q$  be an odd prime power to be determined later. Set  $\ell = \lceil \frac{\log n}{\log q} \rceil$ . For the construction, we assume that we have a bipartite graph  $G = (L, R, E)$  which is a  $k$ -unique-neighbor expander with  $|L| = q^\ell$ ,  $|R| = q^{r+1}$ , and left-degree  $q$ . By our choice of  $\ell$  we have  $|L| \geq n$ . Fix an arbitrary subset  $L'$  of  $L$  such that  $|L'| = n$ . Set  $m = q^{r+1}$  and identify  $R$  with the finite field  $\mathbb{F}_m$ . For every vertex  $v \in L'$  define the polynomial  $p_v(x) \in \mathbb{F}_m[x]$  by

$$p_v(x) = \prod_{w : (v,w) \in E} (x - w).$$

We now describe the sample space  $S$  over  $\{0, 1\}^n$ .<sup>6</sup> Each element in  $S$  corresponds to a field element in  $\mathbb{F}_m$ , that is,  $S = \{s_x : x \in \mathbb{F}_m\}$ . The string  $s_x$  is indexed by elements from the set  $L'$ . In particular, for every  $x \in \mathbb{F}_m$  and  $v \in L'$ , we define

$$(s_x)_v = \begin{cases} \frac{1 - \chi_m(p_v(x))}{2}, & p_v(x) \neq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (3.1)$$

The following theorem readily implies Theorem 1.2 by setting  $\delta = 4/(r - 1)$ .

**Theorem 3.1.** *For every integers  $n, k, r$  such that  $n \geq k$  and  $r \geq 2$ , and for any  $\varepsilon > 0$ , there is a way to choose  $q$  such that the construction defined above is an explicit gradual  $(k, \varepsilon)$  sample space over  $\{0, 1\}^n$  with size*

$$m \leq \max \left\{ (10r^2)^{r+1} \left( \frac{\log n}{\log k} \right)^{r+1} k^{1+1/r}, 3^{r+1} \left( \frac{2k}{\varepsilon} \right)^{2+4/(r-1)} \right\}.$$

<sup>4</sup>For this construction to be explicit, the characteristic of  $\mathbb{F}_q$  should be small. In our construction we take it to be 3.

<sup>5</sup>The parameter  $r$  is related to the parameter  $\delta$  that appears in Theorem 1.2. In particular  $r = 1 + 4/\delta$ .

<sup>6</sup>In fact, we define  $S$  as a multi-set. The sample space is induced in the natural way, namely, to sample from the sample space, one sample an element  $s \in S$  with probability proportional to the multiplicity of  $s$  in  $S$ .



To prove Theorem 3.1 we prove the following two claims.

**Claim 3.2.** *If*

$$q \geq \left(\frac{2k}{\varepsilon}\right)^{2/(r-1)}$$

*then the sample space defined above is gradual  $(k, \varepsilon)$ -biased.*

**Claim 3.3.** *If*

$$q \geq 3.3 \cdot \frac{\log n}{\log k} \cdot k^{1/r} r^2,$$

*then we have an explicit construction of the  $k$ -unique-neighbor expander graph  $G = (L, R, E)$  required by the above construction.*

Before proving the two claims we derive Theorem 3.1 from them. By choosing

$$q \geq \max \left\{ 3.3 \cdot \frac{\log n}{\log k} \cdot k^{1/r} r^2, \left(\frac{2k}{\varepsilon}\right)^{2/(r-1)} \right\}, \quad (3.2)$$

Claim 3.3 assures us that we can obtain the graph  $G$  that we need in the construction. Having this graph, Claim 3.2 guarantees that the above sample space is gradual  $(k, \varepsilon)$ -biased. Certainly one can efficiently find a choice for  $q = 3^z$  which is at most three times the right hand side of Equation (3.2).<sup>7</sup> As  $m = q^{r+1}$  we get the following upper bound on  $m$ , the sample space size

$$m \leq \max \left\{ (10r^2)^{r+1} \left(\frac{\log n}{\log k}\right)^{r+1} k^{1+1/r}, 3^{r+1} \left(\frac{2k}{\varepsilon}\right)^{2+4/(r-1)} \right\},$$

hence Theorem 3.1 follows.

**Proof of Claim 3.2:** Let  $T \subseteq L'$  be a non-empty set of size at most  $k$ . Define

$$p_T(x) = \prod_{v \in T} p_v(x).$$

Since  $p_T(x)$  is defined as a product of  $|T|$  polynomials, each of degree at most  $q$ , we have that  $\deg(p_T(x)) \leq q \cdot |T|$ . Moreover, we claim that  $p_T(x)$  has a simple root. Indeed,  $T$  is a nonempty set of size at most  $k$  of  $L' \subseteq L$ . By our assumption,  $G$  is a  $k$ -unique-neighbor expander, and so there exists a vertex  $w \in R$  with exactly one neighbor,  $v$ , in  $T$ . This implies that  $w$  is a simple root of  $p_v(x)$ , while for every  $u \in T \setminus \{v\}$ ,  $p_u(w) \neq 0$ . Hence, by the definition of  $p_T(x)$  we have that  $w$  is a simple root of  $p_T(x)$ . Now, the bias of the linear test defined by  $T$  is

$$\sum_{x \in \mathbb{F}_m} (-1)^{\sum_{v \in T} (s_x)_v} = \sum_{x \in \mathbb{F}_m} \prod_{v \in T} (-1)^{(s_x)_v}. \quad (3.3)$$

---

<sup>7</sup>Observe also that this solves the minor issue regarding the need for small characteristic for the explicitness requirements of Theorem 2.5.

Suppose  $x$  is not a root of  $p_T(x)$ . Then the value such an  $x$  contributes to the sum in Equation (3.3) is

$$\prod_{v \in T} (-1)^{(s_x)v} = \prod_{v \in T} \chi_m(p_v(x)) = \chi_m \left( \prod_{v \in T} p_v(x) \right) = \chi_m(p_T(x)),$$

where the middle equality follows from the fact that  $\chi$  is a multiplicative homomorphism. As  $p_T(x)$  has at most  $\deg(p_T) \leq q \cdot |T|$  roots, we have that

$$\left| \sum_{x \in \mathbb{F}_m} (-1)^{\sum_{v \in T} (s_x)v} \right| \leq \left| \sum_{x \in \mathbb{F}_m} \chi_m(p_T(x)) \right| + q \cdot |T|$$

Since  $p_T(x)$  has a simple root,  $p_T(x)$  is not of the form  $c \cdot g(x)^2$  for any  $c \in \mathbb{F}_m$  and  $g \in \mathbb{F}_m[x]$ . Therefore, we can apply Weil's Theorem (Theorem 2.1) to get

$$\left| \sum_{x \in \mathbb{F}_m} \chi_m(p_T(x)) \right| < q \cdot |T| \cdot \sqrt{m}.$$

Hence,

$$\frac{1}{m} \left| \sum_{x \in \mathbb{F}_m} (-1)^{\sum_{v \in T} (s_x)v} \right| \leq \frac{2q \cdot |T|}{\sqrt{m}} = 2|T| \cdot q^{(1-r)/2}.$$

To get a bound of at most  $\varepsilon$  on the bias for tests of size exactly  $k$ , we require that

$$2k \cdot q^{(1-r)/2} \leq \varepsilon,$$

or

$$q \geq \left( \frac{2k}{\varepsilon} \right)^{2/(r-1)}. \quad (3.4)$$

□

**Proof of Claim 3.3:** We use the expanders from Theorem 2.5 with  $h = \lceil k^{1/r} \rceil$ . If

$$q - (\ell - 1)(h - 1)r \geq 0.51q$$

then  $G$  is a  $k$ -unique-neighbor expander. By the definition of  $\ell$ , for the above equation to hold, it is enough to require

$$q \log q \geq 2.05 \cdot \log n \cdot k^{1/r} r. \quad (3.5)$$

We use the following simple claim that can be easily verified.

**Claim 3.4.** *For every  $x, y > 1$ , if*

$$x \geq 1.6 \cdot \frac{y}{\log y}$$

*then  $x \log x \geq y$ .*

By the Claim 3.4, for equation (3.5) to hold, it is enough to require that

$$q \geq 1.6 \cdot \frac{2.05 \cdot \log n \cdot k^{1/r} r}{\log(k^{1/r})} = 3.28 \cdot \frac{\log n}{\log k} \cdot k^{1/r} r^2,$$

which concludes the proof. □

## 4 Non-Linear Bias Decay

The definition of a gradual  $(k, \varepsilon)$ -biased sample space that appears in the introduction requires a bound of the form  $\varepsilon \cdot |T|/k$  on the bias for any nonempty set  $T$  of size at most  $k$ . The construction we suggest in this paper indeed has such linear decay. However, one may consider a more general definition where the decay exponent is a non-negative real parameter  $d$ .

**Definition 4.1.** *A sample space  $S$  over  $\{0, 1\}^n$  is called gradual  $(k, d, \varepsilon)$ -biased if for every nonempty  $T \subseteq [n]$  of size at most  $k$ ,*

$$\left| \mathbb{E}_{s \sim S} [(-1)^{\sum_{i \in T} s_i}] \right| \leq \varepsilon \cdot \left( \frac{|T|}{k} \right)^d.$$

We call  $d$  the decay exponent.

A straightforward probabilistic argument shows that a random sample space  $S$  over  $\{0, 1\}^n$  of size  $m = O(k^{2d} \cdot \varepsilon^{-2} \cdot \log n)$  is, with high probability, a  $(k, d, \varepsilon)$ -biased sample space. We start this section by proving an almost matching lower bound on the size of  $(k, d, \varepsilon)$ -biased sample spaces (Theorem 4.3 below). We then turn to present two simple methods that transform a gradual  $(k, d, \varepsilon)$ -biased sample space to a gradual  $(k, d', \varepsilon)$ -biased sample space for  $d' > d$ . These methods, together with the construction for the case  $d = 1$  (Theorem 1.2) yields constructions with larger decay exponents (Corollary 4.4).

### 4.1 A Lower Bound

To prove a lower bound on the size of a gradual small-bias sample space we will use the following known lower bound on the size of (non-gradual)  $(k, \varepsilon)$ -biased sample spaces.

**Theorem 4.2** ([AAK<sup>+</sup>07, Alo09]). *Let  $S$  be a  $(k, \varepsilon)$ -biased sample space over  $\{0, 1\}^n$  of size  $m$ . If  $\varepsilon \geq \binom{n}{k/2}^{-1/2}$  then*

$$m \geq \Omega \left( \frac{k \log(n/k)}{\varepsilon^2 \cdot \log 1/\varepsilon} \right).$$

We now state and prove a lower bound for the size of gradual  $(k, d, \varepsilon)$  sample spaces.

**Theorem 4.3.** *Let  $S$  be a gradual  $(k, d, \varepsilon)$ -biased sample space over  $\{0, 1\}^n$  of size  $m$ . If  $d \leq k/\log k$  and  $\varepsilon > (d \log k/n)^{O(d \log k)}$  then*

$$m \geq \Omega \left( \frac{\log n}{\varepsilon^2 \cdot \log 1/\varepsilon} \cdot \left( \frac{k}{d \cdot \log k} \right)^{2d} \right).$$

**Proof:** Let  $S$  be a gradual  $(k, d, \varepsilon)$ -biased sample space over  $\{0, 1\}^n$  of size  $m$ . Then, in particular,  $S$  is a  $(k', \varepsilon')$ -biased sample space with  $k' = d \log k$  and  $\varepsilon' = \varepsilon \cdot \left( \frac{d \log k}{k} \right)^d$ . As  $\varepsilon' \geq \binom{n}{k'/2}^{-1/2}$  we can use Theorem 4.2 to deduce that

$$m \geq \Omega \left( \frac{k' \log(n/k')}{(\varepsilon')^2 \cdot \log 1/\varepsilon'} \right) \geq \Omega \left( \left( \frac{k}{d \cdot \log k} \right)^{2d} \cdot \frac{1}{\varepsilon^2 \cdot \log(1/\varepsilon)} \cdot \frac{\log k \cdot \log \left( \frac{n}{d \log k} \right)}{\log \left( \frac{k}{d \log k} \right)} \right).$$

Since we assume that  $d \leq k/\log k$ , and since  $k \leq n$ , we have that

$$\frac{\log k \cdot \log\left(\frac{n}{d \log k}\right)}{\log\left(\frac{k}{d \log k}\right)} \geq \log n,$$

thus we have the desired lower bound on  $m$ .  $\square$

## 4.2 Amplifying the Decay Exponent

We now present two simple methods to amplify the decay exponent of a given sample  $S$ . Of course, one must pay in the sample space size in order to get a stronger decay. The two methods we suggest give incomparable sample space sizes. One is better than the other depending on how the size of  $S$  depends on  $n$ ,  $k$  and  $\varepsilon$ . Starting with our construction, the two methods give roughly the same sample space size.

The first method is based on the following trivial observation: every gradual  $(k, d, \varepsilon/k)$ -biased sample space  $S$  on  $n$  variables is a gradual  $(k, d+1, \varepsilon)$ -biased sample space on  $n$  variables. Indeed, for every nonempty  $T \subseteq [n]$  of size at most  $k$ ,

$$\left| \mathbb{E}_{s \sim S} [(-1)^{\sum_{i \in T} s_i}] \right| \leq \frac{\varepsilon}{k} \cdot \left(\frac{|T|}{k}\right)^d \leq \varepsilon \cdot \left(\frac{|T|}{k}\right)^{d+1}.$$

That is, choosing a smaller error to begin with, will result in a larger decay exponent. This observation, together with Theorem 1.2 immediately implies the following corollary.

**Corollary 4.4.** *For any integers  $n, k, d$ , such that  $k \leq n$ , for any  $\varepsilon > 0$ , and for any constant  $\delta > 0$ , there exists an explicit construction of a gradual  $(k, d, \varepsilon)$  sample space of size*

$$m = O_\delta \left( \left(\frac{k^d}{\varepsilon}\right)^{2+\delta} + \left(\frac{\log n}{\log k}\right)^{2+4/\delta} k^{1+\delta} \right).$$

For constructions where the dependency in  $\varepsilon^{-1}$  is small, and this is the case in our construction, the above method is quite effective. However, for a construction that suffers a worse dependency on  $\varepsilon^{-1}$ , the following method, which doubles the decay exponent, would be preferred: given a gradual small-bias sample space  $S$ , use the sample space  $S + S$ .<sup>8</sup> More formally,

**Claim 4.5.** *Let  $S$  be a gradual  $(k, d, \sqrt{\varepsilon})$ -biased sample space. Then  $S + S$  is a gradual  $(k, 2d, \varepsilon)$ -biased sample space.*

**Proof:** Let  $T \subseteq [n]$  be a non-empty set of size at most  $k$ . Then,

$$\left| \mathbb{E}_{s \sim S+S} [(-1)^{\sum_{i \in T} s_i}] \right| = \left( \mathbb{E}_{s \sim S} [(-1)^{\sum_{i \in T} s_i}] \right)^2 \leq \left( \sqrt{\varepsilon} \cdot \left(\frac{|T|}{k}\right)^d \right)^2 = \varepsilon \cdot \left(\frac{|T|}{k}\right)^{2d}.$$

$\square$

---

<sup>8</sup>The sample space  $S + S$  is defined by sampling  $s_1$  and  $s_2$ , independently, from  $S$  and then outputting  $s_1 \oplus s_2$ .

## 5 Concluding Remarks and Open Problems

Our method for transforming a gradual  $\varepsilon$ -biased sample space into a gradual  $(k, \varepsilon)$ -biased sample space uses, as a black box, the unbalanced expanders of [GUV09]. Hence, improved constructions of unbalanced expanders, or low-density parity-check codes in general, may lead to improved constructions of gradual  $(k, \varepsilon)$ -biased sample spaces. Indeed, our general method has the potential to generate very good gradual  $(k, \varepsilon)$ -biased sample spaces from the Quadratic Characters Construction given better constructions of unbalanced expanders. Specifically, using the unbalanced expanders given by the probabilistic construction (see, e.g., [GUV09]), our method yields a gradual  $(k, \varepsilon)$ -biased sample space of size  $O((k\varepsilon^{-1} \log n)^2)$ . This is as good as the non-gradual  $(k, \varepsilon)$ -biased sample space of [AGHP92]. It would therefore be interesting to construct a gradual small-bias sample space that matches the parameters of the non-gradual sample space of [AGHP92].

For non-gradual small-bias sample spaces there are a few explicit constructions with incomparable size [AGHP92, ABN<sup>+</sup>92, NN93, BT09]. Finding an explicit construction of a gradual small-bias sample space with better (or incomparable) size to ours is therefore a natural research goal. One possible route towards this goal is to construct a gradual small-bias sample space that has an incomparable size with that of the Quadratic Character Construction. We are not aware of such a construction in the literature.

The original motivation for studying (non-gradual)  $(k, \varepsilon)$ -biased sample spaces was to construct a sample space  $S$  that is almost  $k$ -wise independent. Using a gradual  $(k, \varepsilon)$ -biased sample space instead of the non-gradual one improves the size of  $S$  by a mere multiplicative constant factor. Nevertheless, we hope that applications that exploit the gradual bound on the bias would be found.

## Acknowledgements

The second author would like to thank his advisor Ran Raz for his continuous support and encouragement.

## References

- [AAK<sup>+</sup>07] N. Alon, A. Andoni, T. Kaufman, K. Matulef, R. Rubinfeld, and N. Xie. Testing  $k$ -wise and almost  $k$ -wise independence. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 496–505. ACM, 2007.
- [ABN<sup>+</sup>92] N. Alon, J. Bruck, J. Naor, M. Naor, and R. Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. *IEEE Transactions on Information Theory*, 38:509–516, 1992.
- [AC12] N. Alon and G. Cohen. On rigid matrices and  $U$ -polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 2012.

- [AGHP92] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple construction of almost  $k$ -wise independent random variables. *Random Structures and Algorithms*, 3(3):289–304, 1992.
- [AIK<sup>+</sup>90] M. Ajtai, H. Iwaniec, J. Komlos, J. Pintz, and E. Szemerédi. Construction of a thin set with small Fourier coefficients. *Bulletin of the London Mathematical Society*, 22:583–590, 1990.
- [Alo09] N. Alon. Perturbed identity matrices have high rank: proof and applications. *Combinatorics, Probability and Computing*, 18(1-2):3–15, 2009.
- [AM95] N. Alon and Y. Mansour.  $\epsilon$ -discrepancy sets and their application for interpolation of sparse polynomials. *Information Processing Letters*, 54(6):337–342, 1995.
- [APY09] N. Alon, R. Panigrahy, and S. Yekhanin. Deterministic approximation algorithms for the nearest codeword problem. In *APPROX-RANDOM*, pages 339–351, 2009.
- [AR94] N. Alon and Y. Roichman. Random Cayley graphs and expanders. *Random Structures and Algorithms*, 5(2):271–285, 1994.
- [Baz09] L.M.J. Bazzi. Polylogarithmic independence can fool DNF formulas. *SIAM Journal on Computing*, 38(6):2220–2272, 2009.
- [BNS92] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols, pseudo-random generators for logspace, and time-space trade-offs. *J. of Computer and System Sciences*, 45(2):204–232, 1992.
- [BSSVW03] E. Ben-Sasson, M. Sudan, S. Vadhan, and A. Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. In *Proceedings of the 35th STOC*, pages 612–621, 2003.
- [BT09] A. Ben-Aroya and A. Ta-Shma. Constructing small-bias sets from algebraic-geometric codes. In *Proceedings of the 50th annual IEEE symposium on foundations of computer science (FOCS)*, 2009.
- [CRS12] G. Cohen, R. Raz, and G. Segev. Non-malleable extractors with short seeds and applications to privacy amplification. In *Proceedings of the 27rd Annual CCC*, 2012.
- [DETT10] A. De, O. Etesami, L. Trevisan, and M. Tulsiani. Improved pseudorandom generators for depth 2 circuits. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 504–517, 2010.
- [EGL<sup>+</sup>92] G. Even, O. Goldreich, M. Luby, N. Nisan, and B. Velickovic. Approximations of general independent distributions. In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 10–16, 1992.

- [GMR<sup>+</sup>12] P. Gopalan, R. Meka, O. Reingold, L. Trevisan, and S. Vadhan. Better pseudorandom generators from milder pseudorandom restrictions. *arXiv preprint arXiv:1210.0049*, 2012.
- [GUV09] V. Guruswami, C. Umans, and S. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *J. ACM*, 56(4):1–34, 2009.
- [HPS93] J. Håstad, S. Phillips, and S. Safra. A well-characterized approximation problem. *Information Processing Letters*, 47(6):301–305, 1993.
- [MST06] E. Mossel, A. Shpilka, and L. Trevisan. On  $\varepsilon$ -biased generators in  $\text{NC}^0$ . *Random Structures and Algorithms*, 29(1):56–81, 2006.
- [MW02] R. Meshulam and A. Wigderson. Expanders from symmetric codes. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 669–677, 2002.
- [Nao92] M. Naor. Constructing Ramsey graphs from small probability spaces. Technical report, IBM Research Report RJ 8810, 1992.
- [NN93] J. Naor and M. Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM J. on Computing*, 22(4):838–856, 1993.
- [Raz05] R. Raz. Extractors with weak random seeds. In *Proceedings of the 37th Annual STOC*, pages 11–20, 2005.
- [RSW93] A. A. Razborov, E. Szemerédi, and A. Wigderson. Constructing small sets that are uniform in arithmetic progressions. *Combinatorics, Probability & Computing*, 2:513–518, 1993.
- [Sch76] W. M. Schmidt. *Equations over Finite Fields: An elementary approach*. Springer-Verlag, 1976.
- [Shp06] A. Shpilka. Constructions of low-degree and error-correcting  $\varepsilon$ -biased generators. In *21st Annual IEEE Conference on Computational Complexity*, pages 33–45, 2006.
- [SS96] M. Sipser and D. Spielman. Expander codes. *Information Theory, IEEE Transactions on*, 42(6):1710–1722, 1996.
- [SZ94] A. Srinivasan and D. Zuckerman. Computing with very weak random sources. In *35th Annual Symposium on Foundations of Computer Science*, pages 264–275. IEEE, 1994.
- [Vio09] E. Viola. The sum of  $d$  small-bias generators fools polynomials of degree  $d$ . *Computational Complexity*, 18(2):209–217, 2009.
- [Vio12] E. Viola. The complexity of distributions. *SIAM Journal on Computing*, 41(1):191–218, 2012.

- [VW08] E. Viola and A. Wigderson. Norms, XOR lemmas, and lower bounds for  $GF(2)$  polynomials and multiparty protocols. *Theory of Computing*, (4):137–168, 2008.