Electronic Colloquium on Computational Complexity, Report No. 128 (2006)

On obtaining pseudorandomness from error-correcting codes*

Shankar Kalyanaraman
Department of Computer Science,
California Institute of Technology
Pasadena, CA 91125.
shankar@cs.caltech.edu

Christopher Umans
Department of Computer Science,
California Institute of Technology
Pasadena, CA 91125.
umans@cs.caltech.edu

October 5, 2006

Abstract

A number of recent results have constructed randomness extractors and pseudorandom generators (PRGs) directly from certain error-correcting codes. The underlying construction in these results amounts to picking a random index into the codeword and outputting m consecutive symbols (the codeword is obtained from the weak random source in the case of extractors, and from a hard function in the case of PRGs).

We study this construction applied to general cyclic error-correcting codes, with the goal of understanding what pseudorandom objects it can produce. We show that *every* cyclic code with sufficient distance yields extractors that fool all linear tests. Further, we show that *every* polynomial code with sufficient distance yields extractors that fool all low-degree prediction tests. These are the first results that apply to univariate (rather than multivariate) polynomial codes, hinting that Reed-Solomon codes may yield good randomness extractors.

Our proof technique gives rise to a systematic way of producing *unconditional* PRGs against restricted classes of tests. In particular, we obtain PRGs fooling all linear tests (which amounts to a construction of ϵ-biased spaces), and we obtain PRGs fooling all low-degree prediction tests.

1 Introduction

Two of the central objects in the area of derandomisation are *extractors* and *pseudorandom generators*. Extractors use a small number of truly random bits to transform "weak" random source into a nearly uniform one. Thus extractors allow the simulation of randomised procedures using only weak randomness (which, for example, may be available from a physical source). In addition to this original motivation, extractors have been used in numerous other settings including complexity theory [Sip88, NZ96, GZ97], algorithms [WZ93], hardness of approximation [Zuc96, Uma99, MU02], distributed protocols [Zuc97, RZ01], and coding theory [TSZ01]. For further discussion see Shaltiel's survey [Sha02]. Quite good constructions of extractors are known now (e.g., [RSW00], [SU05], [LRVW03]), but it remains an open problem to construct optimal extractors.

Pseudorandom generators (PRGs) use a small number of truly random bits to transform a hard function into a small set of strings (a *discrepancy set*) which cannot be distinguished from the uniform distribution by an efficient computational procedure. Thus PRGs prove "hardness vs. randomness" tradeoffs, which show

*This research was supported by NSF grant CCF-0346991 and by BSF grant 2004329.

ISSN 1433-8092

that randomised procedures may be simulated deterministically, under a suitable hardness assumption. A sequence of works has ultimately produced "optimal" PRG constructions [Uma03] that fool general randomised procedures. There is a substantial literature on PRGs that fool more restricted classes of tests, and in some instances *unconditional* constructions (not requiring access to a hard function) are available. For example ϵ-biased spaces [NN93, AGHP92] are PRGs that fool linear tests; other constructions fool affine tests [ABCR97], combinatorial rectangles (see the survey [Sri00]), and general space-bounded computation [Nis92, Nis94, NZ96, INW94, SZ99]. Recently, Bogdanov has constructed PRGs that fool low degree polynomial tests [Bog05].

There is a strong connection between these objects (extractors and PRGs) and *error-correcting codes*. For example, Trevisan's extractor construction [Tre01] uses at its core any good list-decodable code. Subsequent works [TSZS01, SU05] have constructed extractors directly from Reed-Müller codes (and in return, extractors have been used to construct good error-correcting codes in [TSZ01]). PRGs constructed in [STV01, Uma03] have at their core Reed-Müller codes, and it is well-known that ϵ-biased spaces are equivalent to codes with good distance.

In this paper we study a simple construction suited to any cyclic code. Specifically, given any q-ary cyclic error-correcting code $\mathcal{C} : \mathbb{F}_q^{\bar{k}} \to \mathbb{F}_q^{\bar{n}}$, and an additional parameter m, we define the function $f_{\mathcal{C},m} : \mathbb{F}_q^{\bar{k}} \times [\bar{n}] \to \mathbb{F}_q^m$ as follows:

$$f_{\mathcal{C},m}(x, y) = (\mathcal{C}(x)[y+1], \mathcal{C}(x)[y+2], \mathcal{C}(x)[y+3], \ldots, \mathcal{C}(x)[y+m]), \tag{1}$$

where the symbols of the code are indexed in the cyclic ordering. Our goal is to understand what derandomisation objects are produced by this construction. This construction already has a good "track record" — for certain specific kinds of codes the results of [SU05, Uma03] show that

- $f_{\mathcal{C},m}$ is a (k, ϵ)-extractor with $m = k^{1-\delta}$ when $\mathcal{C}$ is a Reed-Müller code with suitable parameters, and
- $f_{\mathcal{C},m}$ is an ϵ-PRG with $m = k^{\delta}$ when $\mathcal{C}$ is an "augmented" version of a Reed-Müller code with suitable parameters, and when we fix x to be the truth table of a function that cannot be computed by size k circuits.

We are interested in the following questions: Is $f_{\mathcal{C},m}$ a good extractor for *every* cyclic code $\mathcal{C}$ with sufficiently good distance? If so, what parameters does it achieve? What can be said about $f_{\mathcal{C},m}$ when $\mathcal{C}$ is a Reed-Solomon code? Is it a good extractor? Can it be used to produce PRGs against certain restricted classes of tests? We feel that studying the Reed-Solomon code question in particular may illuminate new ways of arguing about code-based extractor constructions (since the local-decodability of Reed-Müller codes that is so heavily relied on in [TSZS01, SU05] is not present in Reed-Solomon codes).

In general these seem to be difficult questions to resolve. In this paper we obtain some modest positive results. Our results are phrased in terms of "fooling" certain classes of tests. Using this terminology, extractors outputting m bits fool the class of all functions from $\{0,1\}^m$ to $\{0,1\}$, while PRGs fool the class of all functions from $\{0,1\}^m$ to $\{0,1\}$ with small circuits. The proofs for these constructions often transform these "distinguishing" tests into prediction tests (see Section 2 for formal definitions of distinguishers and predictors). In this paper we are concerned with prediction tests directly:

Definition 1.1. *A* **degree** d **prediction test** *is a degree-d polynomial $p : \mathbb{F}_q^m \to \mathbb{F}_q$ such that p can be expressed as*

$$p(x_1, \ldots, x_m) = x_i - p'(x_1, \ldots, x_{i-1})$$

for some i.

Theorem 1. *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary cyclic linear code with $1^{\bar{n}} \in \mathcal{C}$. For any k and $\rho > 0$, $f_{\mathcal{C},m}$ is a (k, ρ) q-ary extractor for the family of all linear prediction tests, provided that $\delta > 1 - \rho/2$, and $k > m \log q + \log(2/\rho)$.*

When $\mathcal{C}$ is further restricted to be a Reed-Müller code (importantly, including the univariate case, which are Reed-Solomon codes), we show:

Theorem 2. *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary Reed-Müller code with parameters ℓ, h. For any k and $\rho > 0$, $f_{\mathcal{C},m}$ is a (k, ρ) q-ary extractor for the family of all degree d prediction tests, provided that $\rho > 2dh/q$, and $k > m \log q + \log(2/\rho)$.*

Our proofs follow the so-called "reconstruction proof" methodology (see, e.g., [Tre01], [TSZS01], [SU05]). That is, we argue that if the distribution induced by $f_{\mathcal{C},m}$ has a "next-element" predictor of the appropriate type (linear or low-degree), then there is a fixed procedure that "reconstructs" many strings in the weak random source from short advice. This leads to a contradiction, as a source with high min-entropy cannot have many strings that have short descriptions.

Many extractor and PRG constructions employ this proof methodology. From one viewpoint the crucial step is transforming a next-element predictor that errs some fraction of the time into a next-element predictor that is *errorless* (here it becomes clear why error-correcting codes play an important role). From an errorless predictor the remainder of the argument is usually straightforward. From this perspective the main loss associated with the constructions of [SU05], that prevents them from being optimal constructions, is in the conversion from predictors that err to errorless predictors.

Our proofs of Theorem 1 and 2 are noteworthy in that they perform this transformation with *no loss*, for a *wide variety* of codes. Of course the price currently is that we only know how to use this argument to fool a restricted class of tests. Nevertheless, one motivation for exploring these questions, and this methodology in particular, is the possibility of exposing new "lighter-weight" proof techniques that may be useful in the quest to construct optimal extractors.

One consequence of our proof technique is that there is a systematic way to produce *unconditional* PRGs against restricted classes of tests from the above extractor constructions. For example, from the construction in Theorem 1, we obtain a PRG fooling linear prediction tests:

Theorem 3. *Let $\mathcal{C}$ be a systematic $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary cyclic linear code with $1^{\bar{n}} \in \mathcal{C}$. Let x be such that $\mathcal{C}(x)[1 \dots \bar{k}] = 0^{\bar{k}-1}1$. Then $\mathcal{S} = \{f_{\mathcal{C},\bar{k}-1}(x, y) : 1 \leq y \leq \bar{n}\}$ is a q-ary pseudorandom set that fools all linear prediction tests with success probability ρ, provided that $\rho \geq 1 - \delta$.*

By converting the q-ary pseudorandom sets into binary (using Theorem 5) we get ϵ-biased spaces of size $O(m\text{polylog}(m, 1/\epsilon)/\epsilon^3)$, which are comparable to those one can obtain using the well-known connection to error-correcting codes. By comparison, [NN93] gives a construction of size m/ϵ^c where $4 < c < 5$ while [AGHP92] provides a construction of size $(m/\epsilon)^2$.

Using the same idea, from the construction in Theorem 2, we obtain an unconditional PRG construction that fools low-degree prediction tests:

Theorem 4. *Let $\mathcal{C}$ be a systematic $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary cyclic Reed-Müller code with with parameters h, ℓ. Let x be such that $\mathcal{C}(x)[1 \dots \bar{k}] = 0^{\bar{k}-1}1$. Then $\mathcal{S} = \{f_{\mathcal{C},\bar{k}-1}(x, y) : 1 \leq y \leq \bar{n}\}$ is a q-ary ρ-pseudorandom set for the class of all degree d prediction tests, provided that $\rho \geq dh/q$.*

This construction may sound like it unconditionally derandomises polynomial identity testing. If "prediction tests" were replaced by "distinguishing tests" that would indeed be the case. Yao's Lemma [Yao82] shows how to convert any distinguishing test into a prediction test, but unfortunately it does not preserve low-degree-ness. However our result does derandomise polynomial identity testing for the restricted class of tests that can be phrased as degree d prediction tests.

Of course derandomising polynomial identity testing is a major open problem with significant consequences (see [KI04]). Several works have succeeded in derandomising polynomial identity testing for

restricted classes of polynomials ([DS05], [RS], [LV98]). Our result derandomises polynomial identity testing for degree d prediction tests; in fact it produces a stronger object, a *hitting set with density* $1-\rho$ (see the discussion following Definition 2.4). We don't know of any trivial constructions of hitting sets with density $1-\rho$ for even this simple class of polynomials, making it an interesting testbed for new techniques.

Two other works construct hitting sets with density $1-\rho$ for general classes of polynomials: Bogdanov [Bog05] constructs a hitting set of density $1-\rho$ against all m-variate polynomials of degree d, of size $m^{O(d\log(d/\rho))}$. Klivans and Spielman [KS01] construct hitting sets of density $1-\rho$ against all m-variate polynomials of degree d with M monomials, of size $O(mMd/\rho)$. Our construction has a much smaller size, md/ρ, against a particular subclass of m-variate polynomials of degree d (degree d prediction tests). For many settings of the parameters, this is an exponential improvement, albeit for a limited class of polynomials. It is in fact surprising that we obtain hitting sets of this size without an explicit constraint on the size of an arithmetic circuit computing the polynomial.

2 Preliminaries

Two distributions P and Q over a finite set S are said to be ϵ-close if their ℓ_1-distance given by $\sum_{x\in S}|P(x)-Q(x)|$ is at most 2ϵ or equivalently if $\max_{A\subseteq S}|P(A)-Q(A)|$ is at most ϵ. The min-entropy of a random variable X with distribution P on S is defined as $H_\infty(X)=\min_{x\in S}\log(1/P(x))$. We often use U_n as a uniformly distributed random variable.

Definition 2.1. *A* **distinguisher** with advantage ϵ *for a random variable $X=(X_1,X_2,\ldots,X_m)$ defined on $\mathbb{F}_q^m$ is a function $f:\mathbb{F}_q^m\to\mathbb{F}_q$ with the property that*

$$|Pr[f(X)=0]-Pr[f(U_m)=0]|\geq\epsilon$$

where U_m is uniformly distributed on $\mathbb{F}_q^m$.

Definition 2.2. *An i^{th}-element predictor with success probability ρ for a random variable $X=(X_1,X_2,\ldots,X_m)$ defined on $\mathbb{F}_q^m$ is a function $f:\mathbb{F}_q^{i-1}\to\mathbb{F}_q$ such that:*

$$Pr[f(X_1,\ldots,X_{i-1})=X_i]\geq\rho$$

If $\rho=1$ we say that f is errorless.

We will be concerned with linear and low-degree distinguishers and predictors. Note that a linear function f satisfies the identities (i) $f(\sum_{j=1}^k x_j)=\sum_{j=1}^k f(x_j)-(k-1)f(0)$ and (ii) $f(\alpha x)=\alpha f(x)-(\alpha-1)f(0)$ for any scalar α. A *homogeneous* linear function f has $f(0)=0$.

Definition 2.3. *A (k,ρ) q-ary extractor for a family of predictors $\mathcal{P}$ is a function $E:\{0,1\}^n\times\{0,1\}^t\to\mathbb{F}_q^m$ such that for every random variable X with $H_\infty(X)\geq k$, there is no i^{th}-element predictor $f\in\mathcal{P}$ for $E(X,U_t)$ with success probability ρ for any $i=1,\ldots m$.*

In our notation, the usual q-ary extractors (as defined in, e.g., [SU05]) are simply q-ary extractors for the family of all predictors. Rather than referring to PRGs directly we prefer to describe the set of strings they produce.

Definition 2.4. *A q-ary ρ-pseudorandom set for a family of predictors $\mathcal{P}$ is a multiset S such that there is no i-th element predictor $f\in\mathcal{P}$ with success probability ρ for the random variable induced by picking an element uniformly at random from S.*

In Bogdanov's terminology [Bog05], a ρ-pseudorandom set for a family of predictors $\mathcal{P}$ is called a *hitting set with density* $1-\rho$ for the class of degree d prediction tests[1]. In fact, it is a simple observation that a ρ-pseudorandom set S for the family of degree d predictors has the property that for every degree d prediction test g, the distribution $g(Z)$ is ρ-close to the distribution $g(X)$ in the max-norm (where Z is a random variable uniformly distributed on S, and X is a uniform random variable). One can also ask that $g(Z)$ and $g(X)$ be ρ-close in the ℓ_1 norm. This gives rise to genuinely a stronger object, termed a *pseudorandom generator of bias* ρ in [Bog05].

Definition 2.5. *An $[\bar{n}, \bar{k}, \bar{d}]$ q-***ary linear code** *is a subspace $\mathcal{C} \subseteq \mathbb{F}_q^{\bar{n}}$ for which the Hamming distance between every pair $x, y \in \mathcal{C}$ is at least $\bar{d}$.*

Given a string x, we will often use $C(x)$ to mean the x-th codeword in C (and all of the codes we consider come equipped with efficient ways to compute this encoding function). A code is *systematic* if the message appears as a prefix of every codeword.

Definition 2.6. *A code $\mathcal{C}$ is* **cyclic** *if it satisfies the following condition:*

$$(x_1, x_2, \ldots, x_{\bar{n}-1}, x_{\bar{n}}) \in \mathcal{C} \Rightarrow (x_{\bar{n}}, x_1, x_2, \ldots, x_{\bar{n}-1}) \in \mathcal{C}.$$

We always treat the indices into a cyclic code modulo $\bar{n}$.

A specific family of q-ary codes we will use are the Reed-Müller codes. The codewords of a *Reed-Müller code with parameters ℓ, h* are the evaluations of ℓ-variate polynomials of total degree at most h, at the points $\mathbb{F}_q^\ell \setminus \{0\}$. The special case of $\ell = 1$ gives the *Reed-Solomon codes*. All of these codes are cyclic (for an appropriate ordering of $\mathbb{F}_q^\ell \setminus \{0\}$) and linear.

Recall that the parity-check matrix H associated with an $[\bar{n}, \bar{k}, \bar{d}]$ q-ary linear code $\mathcal{C}$ satisfies $x \cdot H^T = 0$ for $x \in \mathcal{C}$. The following result is standard:

Proposition 2.1 (Singleton bound)**.** *For a linear $[\bar{n}, \bar{k}, \bar{d}]$ code $\mathcal{C}$, $n - k \geq d - 1$.*

3 Overview of the results

In this section we describe the high-level ideas behind our results, before giving the technical details and full proofs in the next section.

3.1 Extractors fooling linear tests

Let $\mathcal{C}$ be any cyclic code, and consider the function $f_{\mathcal{C},m}$ from (1). We show that for fixed x, if the distribution $f_{\mathcal{C},m}(x, y)$ with y chosen uniformly at random has a linear predictor p, then x has a short description. In this case p is a linear function for which:

$$p(\mathcal{C}(x)[y+1], \mathcal{C}(x)[y+2], \ldots, \mathcal{C}(x)[y+m-1]) = \mathcal{C}(x)[y+m] \tag{2}$$

with noticeable probability over the choice of y.

Our key observation is that if $\mathcal{C}$ has sufficiently good distance, then f must be *errorless*. To prove this we first select a subset S of those y for which (2) holds. If $\mathcal{C}$ has sufficiently good distance, then a given position r may be expressed as a linear combination ℓ of the values of $\mathcal{C}(x)$ at the positions S:

$$\mathcal{C}(x)[r] = \ell(\mathcal{C}(x)[y])_{y \in S}$$

[1] A *hitting set of density α for a family of functions $\mathcal{F}$* is a multiset $H \subseteq \mathbb{F}_q^m$ such that for every non-zero function $p \in \mathcal{F}$, $\Pr_{x \in H}[p(x) \neq 0] > \alpha$.

Since $\mathcal{C}$ is *cyclic*, this same equation holds for *every* cyclic shift; i.e., for all i:

$$\mathcal{C}(x)[r+i] = \ell(\mathcal{C}(x)[y+i])_{y \in S}$$

These equations together with (2), which holds for all $y \in S$, imply that (2) holds for r. Since r was arbitrary, we conclude that p is indeed errorless.

From here, it is easy to see that x may be described by $\mathcal{C}(x)[1 \ldots m-1]$, since we can use p to obtain $\mathcal{C}(x)[m]$, and again to obtain $\mathcal{C}(x)[m+1]$, and so on, until we have $\mathcal{C}(x)$ in its entirety. Finally decoding $\mathcal{C}(x)$ recovers x.

Note that "extractors that fool linear tests" are not meaningful in the usual setting of simulating randomised procedures using a weak random source. This is because if one is only trying to fool linear tests, one could use ϵ-biased spaces to do away with the randomness altogether. However, we believe that this setting is a good testbed for refining the "reconstruction proof" technique, and that it may be valuable to adapt it in the way we do here, to obtain an errorless predictor without relying on local-decodability of the underlying code. Additionally, our goal is to understand the construction in (1) in the most general setting possible, and the fact that an extractor object (albeit against a restricted class of tests) is produced from *any* cyclic code is a step toward that goal.

3.2 Extractors fooling low-degree tests

Now suppose further that $\mathcal{C}$ is a *polynomial* cyclic code; i.e., a Reed-Müller code, and we have the same setup except that the predictor p is now only *low degree*. That is, there is a function p of degree d for which:

$$p(\mathcal{C}(x)[y+1], \mathcal{C}(x)[y+2], \ldots, \mathcal{C}(x)[y+m-1]) = \mathcal{C}(x)[y+m] \tag{3}$$

with noticeable probability over the choice of y. The argument used for linear p breaks down, but a different argument works, relying on the fact that $\mathcal{C}(x)$ is now itself a low-degree polynomial. This means that there is a mapping between the index y and values for variables $y_1, y_2, \ldots, y_n$ for which $r_x(y_1, y_2, \ldots, y_n) \equiv \mathcal{C}(x)[y]$, where r_x is a low-degree polynomial depending on x. The fact that $\mathcal{C}$ is cyclic means that for all i there is a low-degree polynomial $r_{x,i}$ for which $r_{x,i}(y_1, y_2, \ldots, y_n) \equiv \mathcal{C}(x)[y+i]$.

Now, we observe that the left-hand side of (3) is a low-degree polynomial in $y_1, y_2, \ldots, y_n$, as is the right hand side. However, they agree with noticeable probability, so for an appropriate choice of parameters, they must be *equal*. This implies that p is an *errorless* predictor, since equation (3) holds for all y.

3.3 Unconditional PRGs fooling linear and low-degree tests

If for a given code $\mathcal{C}$, one can identify a fixed "good" x for which $f_{\mathcal{C},m}(x, \cdot)$ fools all *efficient* predictors, then $f_{\mathcal{C},m}(x, \cdot)$ generates a discrepancy set against all small circuits. It is standard that such an x yields a function that is not computable by small circuits, and thus in the absence of strong circuit lower bounds we can obtain (at best) a *conditional* construction. When the class of predictors is restricted in a different way, we can pursue the same strategy to produce a *pseudorandom set* against all predictors in this class.

One of the surprising side-effects of having transformations from a predictor to an errorless predictor like the ones we have is that it is easy to produce a "good" x, *unconditionally*. This is because we need only to find a codeword that cannot have an errorless predictor. In fact, any codeword beginning with 0^m1, will suffice. If such a codeword has an errorless predictor p, then that predictor must output 0 since

$$p(\mathcal{C}(x)[y+1], \mathcal{C}(x)[y+2], \ldots, \mathcal{C}(x)[y+m-1]) = \mathcal{C}(x)[y+m]$$

implies $p(0,0,0\ldots,0) = 0$ (when $y = 0$) and $p(0,0,0\ldots,0) = 1$ (when $y = 1$), a contradiction. This gives a simple construction of pseudorandom sets fooling all linear tests from any cyclic code with good

distance. We are also able to conclude that substrings of low-degree polynomials comprise a pseudorandom set that fools low-degree prediction tests, giving a derandomisation of polynomial identity testing for this restricted class of tests.

4 Proofs of main results

In this section, we shall provide formal proofs of our main theorems.

4.1 Extractors fooling linear tests

In this section, we present a construction for q-ary extractors that fool all linear prediction tests. We begin with a crucial property of linear codes:

Lemma 4.1. *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \bar{d}]$ q-ary linear code. Let $S = \{t_1, \ldots, t_m\} \subseteq \{1, 2, \ldots, n\}$ be a set of size at least $\bar{n} - \bar{d} + 1$, and pick $r \in \{1, 2, \ldots n\}$. Then there exists a homogeneous linear function $f : F_q^m \to \mathbb{F}_q$ such that for all x,*

$$\mathcal{C}(x)[r] = f(\mathcal{C}(x)[t_1], \ldots, \mathcal{C}(x)[t_m]).$$

Proof. For the case $r \in S$, this is trivial. We only therefore need to prove the statement for $u \in \overline{S}$. We choose a basis H of $\bar{n} - \bar{k}$ vectors in the dual space of $\mathcal{C}$ such that $H = [I_{\bar{n}-\bar{k}}|H']$ and for any codeword $\mathcal{C}(x)$ we rearrange its symbols to give $\mathcal{C}(x)'$ such that the first $|\bar{S}|$ symbols correspond to indices in $\bar{S}$. From Proposition 2.1 $|\bar{S}| = \delta\bar{n} - 1 \leq \bar{n} - \bar{k}$. By the properties of H, for $1 \leq i \leq |\bar{S}|$ $H_i \cdot \mathcal{C}(x)' = \sum_j H_{ij}\mathcal{C}(x)'[j] = 0$ where H_i is the i-th row vector in H. But for $j \in \bar{S}, H_{ij} = \delta_{ij}$ where $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $j \neq i$. Hence, $\mathcal{C}(x)[i]' = -\sum_{j \in S} H_{ij}\mathcal{C}(x)[j]'$ as claimed. □

We now prove that a "reasonably correct" linear predictor operating on a codeword in a suitable code must in fact be exactly correct.

Lemma 4.2. *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary cyclic linear code with $1^{\bar{n}} \in \mathcal{C}$, and fix x. Suppose p is a linear i^{th}-element predictor with success probability $\rho > (1 - \delta)$ for the random variable $f_{\mathcal{C},m}(x, y)$ induced by picking y uniformly from $\{1, 2, \ldots, \bar{n}\}$. Then, p is an errorless linear predictor.*

Proof. Define S to be the set of positions on which p is correct; i.e.,

$$S = \{s : p(\mathcal{C}(x)[s+1], \mathcal{C}(x)[s+2], \ldots, \mathcal{C}(x)[s+i-1]) = \mathcal{C}(x)[s+i]\}$$

We know that $|S| \geq (1 - \delta)\bar{n} + 1$. Now pick an arbitrary $r \in \{1, 2, \ldots, \bar{n}\}$, and let $f = \sum_{s \in S} \alpha_s z_s$ be the linear function guaranteed by Lemma 4.1. We have:

$$\begin{aligned} p(\mathcal{C}(x)[r+1], \ldots, \mathcal{C}(x)[r+i-1]) &= p\left(\sum_{s \in S} \alpha_s \mathcal{C}(x)[s+1], \ldots, \sum_{s \in S} \alpha_s \mathcal{C}(x)[s+i-1]\right) \\ &= \sum_{s \in S} \alpha_s p(\mathcal{C}(x)[s+1], \ldots, \mathcal{C}(x)[s+i-1]) + \left(1 - \sum_{s \in S} \alpha_s\right) p(0, \ldots 0) \\ &= \sum_{s \in S} \alpha_s p(\mathcal{C}(x)[s+1], \ldots, \mathcal{C}(x)[s+i-1]) = \sum_{s \in S} \alpha_s \mathcal{C}(x)[s+i] = \mathcal{C}(x)[r+i] \end{aligned}$$

where the second line follows from the fact that p is linear (using two properties of linear functions noted in Section 2), and the third line follows because $1^{\bar{n}} \in \mathcal{C}$ implies $(1 - \sum_{s \in S} \alpha_s) = 0$, and from the definition of S. □

We now prove our first main theorem, showing that $f_{\mathcal{C},m}$ for cyclic codes $\mathcal{C}$ is an extractor fooling q-ary linear tests.

Theorem 1 (restated). *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary cyclic linear code with $1^{\bar{n}} \in \mathcal{C}$. For any k and $\rho > 0$, $f_{\mathcal{C},m}$ is a (k, ρ) q-ary extractor for the family of all linear prediction tests, provided that $\delta > 1 - \rho/2$, and $k > m \log q + \log(2/\rho)$.*

Proof. Suppose $f_{\mathcal{C},m}$ is not an extractor with the parameters as claimed. Then there is some random variable X having distribution D, with min-entropy at least k, and for some i, a linear i^{th}-element predictor p satisfying

$$\Pr_{x \leftarrow D, y}[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] \geq \rho.$$

By an averaging argument

$$\Pr_{x \leftarrow D}[\Pr_{y}[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] \geq \rho/2] \geq \rho/2. \tag{4}$$

Now, for every x for which $\Pr_y[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] \geq \rho/2$, Lemma 4.2 implies that $\Pr_y[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] = 1$, since $\rho/2 > 1 - \delta$. Every such x can be described with $(i-1)$ elements of $\mathbb{F}_q$, by simply writing down $\mathcal{C}(x)[1 \ldots, i-1]$. From this, $p(\mathcal{C}(x)[1 \ldots, i-1]) = \mathcal{C}(x)[i]$, and then $p(\mathcal{C}(x)[2 \ldots, i] = \mathcal{C}(x)[i]$, $p(\mathcal{C}(x)[3 \ldots, i+1] = \mathcal{C}(x)[i+2]$, and so on until we obtain all of the symbols of $\mathcal{C}(x)$, which in turn determine x.

We can define a function $R : \mathbb{F}_q^{i-1} \to \mathbb{F}_q^{\bar{k}}$ that runs this procedure. Using equation (4) above, we get:

$$\Pr_{x \leftarrow D}[\exists a \in \mathbb{F}_q^{i-1} \text{ for which } R(a) = x] \geq \rho/2.$$

A given x is sampled with probability at most 2^{-k}, and so applying the union bound, the probability above is bounded above by $q^{i-1}2^{-k}$. Using the fact that $i \leq m$, we get a contradiction if $2^{m \log q - k} < \rho/2$, or equivalently $k > m \log q + \log(2/\rho)$. Our choice of k thus implies that $f_{\mathcal{C},m}$ must be the claimed extractor. □

To get a sense of the achievable extractor parameters here, we plug in a Reed-Müller code:

Corollary 4.3. *Fix n, k, and $\rho > 1/k^{O(1)}$. Let $\mathcal{C}$ be a Reed-Müller code with parameters $h = k$, $q = 2k/\rho$, and $\ell = \log n / \log k$. Then $f_{\mathcal{C},m}$ is a (k, ρ) q-ary extractor for the family of all linear prediction tests, with seed length $O(\log n)$ and output length $m \geq k/O(\log k)$.*

The constructions above can be modified to fool binary linear tests using the following theorem.

Theorem 5. *Let $X = (X_1, X_2, \ldots, X_m)$ be a random variable distributed on $\mathbb{F}_q^m$ that can be sampled using t random bits, and denote by $X_i(y)$ the value of the i-th random variable when sampling using y as the random bits. Let $\mathcal{C}$ be a systematic $[\bar{n}, \bar{k}, \delta\bar{n}]$ binary linear code with $1^{\bar{n}} \in \mathcal{C}$. Define B as follows:*

$$B(y,z) = (\mathcal{C}(X_1(y))[z], \mathcal{C}(X_2(y))[z], \ldots, \mathcal{C}(X_m(y))[z]).$$

If there is an i-th element linear predictor p with success probability $1/2 + \epsilon$ for the random variable $B(y,z)$ induced by picking y and z uniformly at random, then there exists an i-th element linear predictor p' with success probability $\epsilon/2$ for the random variable X, provided that $\delta > 1/2 - \epsilon/2$.

Proof. We have p for which

$$\Pr_{y,z}[p(\mathcal{C}(X_1(y))[z],\ldots,\mathcal{C}(X_{i-1}(y))[z]) = \mathcal{C}(X_i(y))[z]] \geq \frac{1}{2} + \epsilon.$$

By an averaging argument:

$$\Pr_{y}\left[\Pr_{z}[p(\mathcal{C}(X_1(y))[z],\ldots,\mathcal{C}(X_{i-1}(y))[z]) = \mathcal{C}(X_i(y))[z]] \geq \frac{1}{2} + \frac{\epsilon}{2}\right] \geq \frac{\epsilon}{2}.$$

Let us call a value y for which

$$\Pr_{z}[p(\mathcal{C}(X_1(y))[z],\ldots,\mathcal{C}(X_{i-1}(y))[z]) = \mathcal{C}(X_i(y))[z]] \geq \frac{1}{2} + \frac{\epsilon}{2}$$

holds "good."

Claim 6. *For good y,*

$$\Pr_{z}[p(\mathcal{C}(X_1(y))[z],\ldots,\mathcal{C}(X_{i-1}(y))[z]) = \mathcal{C}(X_i(y))[z]] = 1.$$

Proof (of Claim 6). The proof is similar to the proof of Lemma 4.2. Define S to be the set of positions on which p is correct; i.e.,

$$S = \{s : p(\mathcal{C}(X_1(y))[s],\ldots,\mathcal{C}(X_{i-1}(y))[s]) = \mathcal{C}(X_i(y))[s]\}$$

We know that $|S| \geq (1/2 + \epsilon/2)\bar{n} > (1-\delta)\bar{n}$. Now pick an arbitrary $r \in \{1, 2, \ldots, \bar{n}\}$, and let $f = \sum_{s\in S} \alpha_s z_s$ be the linear function guaranteed by Lemma 4.1. We have:

$$\begin{aligned}
p(\mathcal{C}(X_1(y))[r],\ldots,\mathcal{C}(X_{i-1}(y))[r]) &= p\left(\sum_{s\in S}\alpha_s\mathcal{C}(X_1(y))[s],\ldots,\sum_{s\in S}\alpha_s\mathcal{C}(X_{i-1}(y))[s]\right) \\
&= \sum_{x\in S}\alpha_s p(\mathcal{C}(X_1(y))[s],\ldots,\mathcal{C}(X_{i-1}(y))[s]) + \left(1-\sum_{s\in S}\alpha_s\right)p(0,\ldots 0) \\
= \sum_{s\in S}\alpha_s p(\mathcal{C}(X_1(y))[s],\ldots,\mathcal{C}(X_{i-1}(y))[s]) &= \sum_{s\in S}\alpha_s\mathcal{C}(X_i(y))[s] = \mathcal{C}(X_i(y))[r]
\end{aligned}$$

where the second line follows from the fact that p is linear (using two properties of linear functions noted in Section 2), and the third line follows because $1^{\bar{n}} \in \mathcal{C}$ implies $(1 - \sum_{s\in S}\alpha_s) = 0$, and from the definition of S. □

Let $p' : \mathbb{F}_q^{i-1} \to \mathbb{F}_q$ be the function given by:

$$\begin{aligned}
p'(x_1, x_2, \ldots, x_{i-1}) \quad = \quad & (p(\mathcal{C}(x_1)[1], \mathcal{C}(x_2)[1], \ldots, \mathcal{C}(x_{i-1})[1]), \\
& p(\mathcal{C}(x_1)[2], \mathcal{C}(x_2)[2], \ldots, \mathcal{C}(x_{i-1})[2]), \\
& \vdots \\
& p(\mathcal{C}(x_1)[\log q], \mathcal{C}(x_2)[\log q], \ldots, \mathcal{C}(x_{i-1})[\log q]))
\end{aligned}$$

We claim that p' is a q-ary linear predictor with success probability $\epsilon/2$ for the the random variable X defined on $\mathbb{F}_q^m$.

First, using Claim 6, observe that for good y

$$p'(X_1(y), X_2(y), \ldots, X_{i-1}(y)) = X_i(y).$$

Here we are relying on the fact that $\mathcal{C}$ is systematic, so the first $\log q$ bits contain the message, which in this case is an element of $\mathbb{F}_q$. Thus p' is the promised predictor with success probability at least $\epsilon/2$, since that is the probability of choosing a good y.

It only remains to verify that p' is indeed $\mathbb{F}_q$-linear. The j-th element output by p' is an $\mathbb{F}_2$-linear combination the j-th elements of the inputs to p', namely $x_1, x_2, \ldots, x_{i-1}$. In fact this linear combination is determined by p and it is the same for all $j = 1, 2, \ldots, \log q$, since $\mathcal{C}$ is systematic. Thus p' can be expressed as:

$$p'(\vec{x_1}, \vec{x_2}, \ldots, \vec{x}_{i-1}) = \sum_{j=1}^{i-1} \beta_j \vec{x_j} + \beta_0 \vec{1}.$$

where $\vec{1}$ is the all-ones vector in $\mathbb{F}_q$, and the β_js are the coefficients that give the linear function that is p (so $\beta_j \in \mathbb{F}_2$ for all j). □

4.2 Extractors fooling low-degree tests

We develop our result further to describe constructions of extractors that fool low-degree prediction tests. While the extractor constructions presented in the previous section can be derived in general from any cyclic, linear code the following constructions are obtained from Reed-Müller codes (including the special case of Reed-Solomon codes). Similar to the previous subsection, we present a lemma that says that a "reasonably good" low-degree predictor is an errorless low-degree predictor.

Lemma 4.4. *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \bar{d}]$ q-ary Reed-Müller code with parameters ℓ, h, and fix x. Suppose p is a degree d i-th element predictor with success probability $\rho > dh/q$ for the random variable $f_{\mathcal{C},m}(x, y)$ induced by picking y uniformly from $\{1, 2, \ldots, \bar{n}\}$. Then p is an errorless predictor.*

Proof. Note that $\mathcal{C}(x)$ is an ℓ-variate polynomial $r_x : \mathbb{F}_q^\ell \to \mathbb{F}_q$ with total degree h, and the symbols of $\mathcal{C}(x)$ are the evaluation of r_x at points $y = (y_1, \ldots, y_\ell) \in \mathbb{F}_q^\ell$. Since $\mathcal{C}$ is cyclic, for all i there is an ℓ-variate polynomial $r_{x,i} : \mathbb{F}_q^\ell \to \mathbb{F}_q$ with total degree h for which $r_{x,i}(y) \equiv \mathcal{C}(x)[y + i]$.

Now consider the i-variate polynomial $Q(z_1, \ldots, z_i) = z_i - p(z_1, \ldots, z_{i-1})$. Q has total degree at most d. The polynomial $Q'(y_1, \ldots, y_\ell) = Q(r_{x,1}, r_{x,2}, \ldots, r_{x,i})$ has total degree at most dh. Moreover, Q' vanishes on exactly those points $y = (y_1, \ldots, y_\ell) \in \mathbb{F}_q^\ell$ for which

$$p(\mathcal{C}(x)[y + 1], \ldots, \mathcal{C}(x)[y + i - 1]) = \mathcal{C}(x)[y + i]. \tag{5}$$

By the Schwartz-Zippel lemma [Sch80, Zip79], we know that a non-zero polynomial of total degree dh can vanish in at most a dh/q fraction of points in $\mathbb{F}_q^\ell$. Since $\rho > dh/q$, it must be that Q' is identically zero, which implies that (5) holds for all y. Thus p is errorless, as claimed. □

This gives us our second main theorem, showing that $f_{\mathcal{C},m}$ for Reed-Müller codes $\mathcal{C}$ is an extractor fooling low-degree tests.

Theorem 2 (restated). *Let $\mathcal{C}$ be an $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary Reed-Müller code with parameters ℓ, h. For any k and $\rho > 0$, $f_{\mathcal{C},m}$ is a (k, ρ) q-ary extractor for the family of all degree d prediction tests, provided that $\rho > 2dh/q$, and $k > m \log q + \log(2/\rho)$.*

Proof. Suppose $f_{\mathcal{C},m}$ is not an extractor with the parameters as claimed. Then there is some random variable X having distribution D, with min-entropy at least k, and for some i, a linear i^{th}-element predictor p satisfying

$$\Pr_{x \leftarrow D, y}[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] \geq \rho.$$

As in the proof of Theorem 1, using an averaging argument we assert that

$$\Pr_{x \leftarrow D}[\Pr_y[f_P(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] \geq \rho/2] \geq \rho/2. \tag{6}$$

Now, for every x for which $\Pr_y[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] \geq \rho/2$, Lemma 4.4 implies that $\Pr_y[p(f_{\mathcal{C},m}(x,y)_{1,\ldots,i-1}) = f_{\mathcal{C},m}(x,y)_i] = 1$, since $\rho/2 > dh/q$.

As in the proof of Theorem 1, these x are uniquely determined $\mathcal{C}(x)[1\ldots,i-1]$. Following the argument in that proof, we arrive at a contradiction if $k > m\log q + \log(2/\rho)$. Hence for our choice of k, $f_{\mathcal{C},m}$ is the claimed extractor. □

The following corollary plugs in Reed-Solomon codes, which correspond to $\ell = 1$ in Theorem 2.

Corollary 4.5. *Fix n, k, d and $\rho > 1/k^{O(1)}$. Let $\mathcal{C}$ be a q-ary Reed-Solomon code with parameters $q = 2dn/\rho$ and $h = n$. Then $f_{\mathcal{C},m}$ is a (k,ρ) q-ary extractor for the family of all degree d prediction tests, with seed length $O(\log n)$ and output length $m \geq k/O(\log dn)$.*

5 Pseudorandom sets for linear and low-degree tests

In this section we obtain unconditional PRGs by using a special feature of our proof methodology.

5.1 Pseudorandom sets for linear tests

Using Lemma 4.2 we can prove the following theorem giving a construction of a pseudorandom set for linear prediction tests:

Theorem 3 (restated). *Let $\mathcal{C}$ be a systematic $[\bar{n}, \bar{k}, \delta\bar{n}]$ q-ary cyclic linear code with $1^{\bar{n}} \in \mathcal{C}$. Let x be such that $\mathcal{C}(x)[1\ldots\bar{k}] = 0^{\bar{k}-1}1$. Then $\mathcal{S} = \{f_{\mathcal{C},\bar{k}-1}(x,y) : 1 \leq y \leq \bar{n}\}$ is a q-ary ρ-pseudorandom set for the class of all linear prediction tests, provided that $\rho > 1 - \delta$.*

Proof. The fact that $\mathcal{C}$ is systematic implies that there exists a codeword $\mathcal{C}(x)$ with the desired properties. Now suppose for the purpose of contradiction there exists an i^{th}-element linear predictor p for the uniform distribution on $\mathcal{S}$. Then by Lemma 4.2, p is an errorless linear predictor. In particular, $p(\mathcal{C}(x)[k-i],\ldots,\mathcal{C}(x)[k-2]) = \mathcal{C}(x)[k-1]$, and $p(\mathcal{C}(x)[k-i+1],\ldots,\mathcal{C}(x)[k-1]) = \mathcal{C}(x)[k]$. However by our choice of $\mathcal{C}(x)$,

$$p(\mathcal{C}(x)[k-i],\ldots,\mathcal{C}(x)[k-2]) = p(\mathcal{C}(x)[k-i+1],\ldots,\mathcal{C}(x)[k-1]) = p(0,\ldots,0)$$

and yet $0 = \mathcal{C}(x)[k-1] \neq \mathcal{C}(x)[k] = 1$, which gives a contradiction. □

Corollary 5.1. *Fix m, ρ. Let $\mathcal{C}$ be a systematic Reed-Solomon code with parameters h, q satisfying $q = h/\rho$. The set $\mathcal{S}$ described in Theorem 3 is a q-ary ρ-pseudorandom set in $\mathbb{F}_q^m$ of size h/ρ for the class of all linear prediction tests.*

Using the following proposition, we can fool linear *distinguishing* tests to give us Corollary 5.3.

Proposition 5.2. *Let $f : \mathbb{F}_q^m \to \mathbb{F}_q$ be a q-ary linear distinguisher for a distribution D with advantage ϵ. Then, there exists an i and a q-ary linear next-element predictor for D f' such that for a random variable x defined over $\mathbb{F}_q^m$,*

$$\Pr_{x \leftarrow D}[f'(x_1, \ldots, x_{i-1}) = x_i] \geq \frac{1}{q} + \frac{\epsilon}{q-1}$$

and for the case $\frac{1}{q} \leq \epsilon \leq 1 - \frac{1}{q}$, $\Pr_{x \leftarrow D}[f'(x_1, \ldots, x_{i-1}) = x_i] \geq \frac{1}{q} + \epsilon$.

Proof. Since f is a linear distinguisher, wlog we may assume that it is of the form $f(x_1, \ldots, x_m) = -x_m + \sum_{i=1}^{m-1} C_i x_i + C_0$. By definition,

$$\left| \Pr_{x \leftarrow D}[f(x_1, \ldots, x_m) = 0] - \Pr_{x}[f(x_1, \ldots, x_m) = 0] \right| \geq \epsilon$$

Note that $\Pr_x[f(x_1, \ldots, x_m) = 0] = 1/q$. Two cases arise: $\Pr_{x \leftarrow D}[f(x_1, \ldots, x_m) = 0] \geq 1/q + \epsilon$ or $\Pr_{x \leftarrow D}[f(x_1, \ldots, x_m) = 0] \leq 1/q - \epsilon$. In the former, $f'(x_1, \ldots, x_{m-1}) = C_0 + \sum_{i=1}^{m-1} C_i x_i$ is an m^{th}-element predictor with success probability $\frac{1}{q} + \epsilon$. In the latter, using a simple pigeonhole argument there exists some $v \in \mathbb{F}_q; v \neq 0$ for which

$$\Pr_{x \leftarrow D}[f(x_1, \ldots, x_m) = v] \geq \frac{1}{q-1} \cdot \left(1 - \frac{1}{q} + \epsilon\right) = \frac{1}{q} + \frac{\epsilon}{q-1}$$

Choosing $f'(x_1, \ldots, x_{m-1}) = C_0 + \sum_{i=1}^{m-1} C_i x_i$ where $C_0' = C_0 - v$ gets us an m^{th}-element predictor with success probability $\frac{1}{q} + \frac{\epsilon}{q-1}$. For the special case when $\frac{1}{q} \leq \epsilon \leq 1 - \frac{1}{q}$, we note that $\Pr_x[f(x_1, \ldots, x_m) = 0] = \frac{1}{q} \leq \epsilon$ and hence the distinguisher property implies that only the first case is possible. □

Corollary 5.3. *Let $\mathcal{C}$ and S be as defined above in Theorem 3. For every $v \in \mathbb{F}_q^m$,*

$$\left| \Pr_{s \in S}[s \cdot v = 0] - \Pr_{x}[x \cdot v = 0] \right| \leq \left(\rho - \frac{1}{q}\right)(q-1).$$

Pseudorandom sets for binary linear distinguishing tests are called ϵ-biased sample spaces. Using our constructions from above and combining them with good binary codes we can construct good ϵ-biased sample spaces.

Definition 5.1. *A multiset $\mathcal{T} \subseteq \{0,1\}^m$ is an ϵ-**biased sample space** if for every $\vec{v} \in \{0,1\}^m$*

$$\left| \Pr_{x \in T}[x \cdot v = 0] - \Pr_{x \in T}[x \cdot v = 1] \right| \leq \epsilon.$$

Theorem 7. *Let $\mathcal{C}_1$ be an $[\bar{n}_1, \bar{k}_1, \delta_1 \bar{n}_1]$ q-ary cyclic code, and $\mathcal{C}_2$ be an $[\bar{n}_2, \bar{k}_2 = \log q, \delta_2 \bar{n}_2]$ binary systematic code, and set $m = \bar{k}_1 - 1$. Define $\mathcal{S} = \{f_{\mathcal{C}_1, m}(x, y) : 1 \leq y \leq \bar{n_1}\}$ and define*

$$\mathcal{T} = \{(\mathcal{C}_2(s_1)[z], \mathcal{C}_2(s_2)[z], \ldots, \mathcal{C}_2(s_m)[z]) : (s_1, s_2, \ldots, s_m) \in S, z \in \{1, 2, \ldots, \bar{n}_2\}\}$$

The set $\mathcal{T}$ is a 4ϵ-biased sample space, provided $\delta_1 > 1 - \epsilon$, and $\delta_2 > 1/2 - \epsilon$.

Proof. Suppose otherwise. Then by definition, there exists a $\vec{v} \in \{0,1\}^m$ such that

$$\left| \Pr_{x \in T}[x \cdot v = 0] - \Pr_{x \in T}[x \cdot v = 1] \right| > 4\epsilon.$$

This means that

$$\Pr_{x\in T}[x\cdot\vec{v}=0]-\Pr_{x\in T}[x\cdot\vec{v}=1]>4\epsilon \text{ or } \Pr_{x\in T}[x\cdot\vec{v}=1]-\Pr_{x\in T}[x\cdot\vec{v}=0]>4\epsilon.$$

Combining this with

$$\Pr_{x\in T}[x\cdot\vec{v}=0]+\Pr_{x\in T}[x\cdot\vec{v}=1]=1,$$

we get

$$\Pr_{x\in T}[x\cdot\vec{v}=0]>1/2+2\epsilon \text{ or } \Pr_{x\in T}[x\cdot\vec{v}=0]<1/2-2\epsilon.$$

Note that $Pr_x[x\cdot\vec{v}=0]=1/2$ and therefore $\vec{v}$ describes a homogeneous linear distinguisher with advantage ϵ; i.e.

$$\left|\Pr_{x\in T}[x\cdot\vec{v}=0]-\Pr_{x}[x\cdot\vec{v}=0]\right|\geq 2\epsilon.$$

By Proposition 5.2, there exists a predictor p with success probability $1/2+2\epsilon$ for the random variables $Y=(Y_1,Y_2,\ldots,Y_m)$ induced by choosing an element of T uniformly at random. By Theorem 5 there exists a predictor p' with success probability ϵ for the random variable $X=(X_1,X_2,\ldots,X_m)$ induced by choosing an elements of S uniformly at random. But this contradicts Theorem 3, as it indicates that S does not fool all linear predictors with success probability $\epsilon>1-\delta_1$. □

By choosing appropriate codes for $\mathcal{C}_1$ and $\mathcal{C}_2$ we obtain Corollary 5.4 and, in particular by using a better binary code for $\mathcal{C}_2$ (Reed-Solomon concatenated with Hadamard) we obtain Corollary 5.5.

Corollary 5.4. *Fix m. Let $\mathcal{C}_1$ be a $[q,m+1,q-m]$ Reed-Solomon code with $q>m/\epsilon$ and let $\mathcal{C}_2$ be an $[q,\log q,q/2]$ binary Hadamard code. Then the set T defined above is an 4ϵ-biased sample space of size $O(m^2/\epsilon^2)$.*

Corollary 5.5. *Fix m. Let $\mathcal{C}_1$ be a $[q,m+1,q-m]$ Reed-Solomon code with $q>m/\epsilon$ and let $\mathcal{C}_2$ be an $[\bar{n}=O(\log^2 q/\epsilon^2),\log q,(1/2-\epsilon)\bar{n}]$ binary code. Then the set T defined above is an 4ϵ-biased sample space of size $O(m\text{polylog}(m,1/\epsilon)/\epsilon^3)$.*

5.2 Pseudorandom sets for low-degree tests

We extend the previous discussion to pseudorandom sets for low-degree tests derived from Reed-Müller codes.

Theorem 4 (restated). *Let $\mathcal{C}$ be a systematic $[\bar{n},\bar{k},\delta\bar{n}]$ q-ary cyclic Reed-Müller code with with parameters h,ℓ. Let x be such that $\mathcal{C}(x)[1\ldots\bar{k}]=0^{\bar{k}-1}1$. Then $\mathcal{S}=\{f_{\mathcal{C},\bar{k}-1}(x,y):1\leq y\leq\bar{n}\}$ is a q-ary ρ-pseudorandom set for the class of all degree d prediction tests, provided that $\rho\geq dh/q$.*

Proof. The proof is nearly identical to the proof of Theorem 3. The fact that $\mathcal{C}$ is systematic implies that there exists a codeword $C(x)$ with the desired properties. Now suppose for the purpose of contradiction there exists an i^{th}-element degree d predictor p for the uniform distribution on $\mathcal{S}$. Then by Lemma 4.4, p is an errorless predictor. In particular,

$$p(\mathcal{C}(x)[k-i],\ldots,\mathcal{C}(x)[k-2])=\mathcal{C}(x)[k-1],$$

and

$$p(\mathcal{C}(x)[k-i+1],\ldots,\mathcal{C}(x)[k-1])=\mathcal{C}(x)[k].$$

However by our choice of $\mathcal{C}(x)$,

$$p(\mathcal{C}(x)[k-i],\ldots,\mathcal{C}(x)[k-2]) = p(\mathcal{C}(x)[k-i+1],\ldots,\mathcal{C}(x)[k-1])$$

and yet

$$0 = \mathcal{C}(x)[k-1] \neq \mathcal{C}(x)[k] = 1,$$

which gives a contradiction. □

Corollary 5.6. *Fix m, ρ. Let $\mathcal{C}$ be a systematic Reed-Solomon code with parameters h, q satisfying $q = dh/\rho$. The set $\mathcal{S}$ described in Theorem 4 is a q-ary ρ-pseudorandom set in $\mathbb{F}_q^m$ of size hd/ρ for the class of all degree d prediction tests.*

Equivalently, we have an explicit construction of a hitting set with density $1-\rho$ against degree d prediction tests, with size md/ρ. As discussed in the introduction this is somewhat surprising. Even for this simple class of polynomials, there does not seem to be a trivial construction of a hitting set with density $1-\rho$, making Theorem 4 another example where the generic object $f_{C,m}$ yields a non-trivial pseudorandom construction.

6 Concluding remarks

There are many questions raised by these results. For example, is it possible to enlarge the class of tests fooled by the extractors and pseudorandom sets constructed from arbitrary cyclic linear codes? Similarly, is it possible to fool more general prediction tests using arbitrary polynomial codes? The results of [SU05] show that it is in the particular case of Reed-Müller codes (with certain parameters), but it is possible that something more general is true depending, e.g., only on the distance of the code.

We feel that one of the nicest questions of this type is the question of whether $f_{\mathcal{C},m}$ is a extractor (fooling all prediction tests), when $\mathcal{C}$ is a Reed-Solomon code.

Regarding pseudorandom sets for low-degree polynomials, we wonder if there is a nontrivial conversion of distinguishers to predictors (probably relying on the distinguisher being presented as a small arithmetic circuit) that preserves low-degree-ness. This would potentially lead to a non-trivial derandomisation of polynomial identity testing, because it would imply that the pseudorandom sets of Theorem 4 would in fact fool low-degree distinguishing tests with small circuits.

Acknowledgements. We thank Eli Ben-Sasson for helpful discussions and Andrej Bogdanov for sharing a draft of [Bog05] with us.

References

[ABCR97] A. E. Andreev, J. L. Baskakov, A. E. F. Clementi, and J. D. P. Rolim. Small random sets for affine spaces and better explicit lower bounds for branching programs. Technical Report TR04-053, ECCC, 1997.

[AGHP92] N. Alon, O. Goldreich, J. Hastad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Struct. Algorithms*, (3):289–304, 1992.

[Bog05] A. Bogdanov. Pseudorandom generators for low degree polynomials. In *Proceedings of STOC*, pages 21–30, 2005.

[DS05] Z. Dvir and A. Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. In *Proceedings of STOC*, pages 592–601, 2005.

[GZ97] O. Goldreich and D. Zuckerman. Another proof that BPP subseteq PH (and more). Technical Report TR97-045, ECCC, 1997.

[INW94] R. Impagliazzo, N. Nisan, and A. Wigderson. Pseudorandomness for network algorithms. In *Proceedings of STOC*, pages 356–364, 1994.

[KI04] V. Kabanets and R. Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. volume 13, pages 1–46, 2004.

[KS01] A. Klivans and D. Spielman. Randomness efficient identity testing of multivariate polynomials. In *Proceedings of STOC*, pages 216–223, 2001.

[LRVW03] C.-J. Lu, O. Reingold, S. Vadhan, and A. Wigderson. Extractors: optimal up to constant factors. In *Proceedings of STOC*, pages 602–611, 2003.

[LV98] D. Lewin and S. Vadhan. Checking polynomial identities over any field: Towards a derandomization? In *Proceedings of STOC*, pages 438–447, 1998.

[MU02] E. Mossel and C. Umans. On the complexity of approximating the VC dimension. *J. Comput. Syst. Sci.*, 65(4):660–671, 2002.

[Nis92] N. Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12:249–461, 1992.

[Nis94] N. Nisan. $\mathrm{RL} \subseteq \mathrm{SC}$. *Computational Complexity*, 4(1):1–11, 1994.

[NN93] J. Naor and M. Naor. Small-bias probability spaces: Efficient constructions and applications. *SICOMP*, 22(4):838–856, August 1993.

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

[RS] R. Raz and A. Shpilka. Deterministic polynomial identity testing in non-commutative models. In *CCC*.

[RSW00] O. Reingold, R. Shaltiel, and A. Wigderson. Extracting randomness via repeated condensing. In IEEE, editor, *FOCS*, pages 22–31, 2000.

[RZ01] A. Russell and D. Zuckerman. Perfect information leader election in $\log^* n + O(1)$ rounds. *J. Comput. Syst. Sci.*, 63(4):612–626, 2001.

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.

[Sha02] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of EATCS*, 77:67–95, June 2002. Columns: Computational Complexity.

[Sip88] M. Sipser. Expanders, randomness, or time versus space. *J. Comput. Syst. Sci.*, 36(3):379–383, 1988.

[Sri00] A. Srinivasan. Low-discrepancy sets for high-dimensional rectangles: a survey. *Bulletin of the EATCS*, 70:67–76, 2000.

[STV01] M. Sudan, L. Trevisan, and S. Vadhan. Pseudorandom generators without the XOR lemma. *J. Comput. Syst. Sci.*, 62(2):236–266, March 2001.

[SU05] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *J. ACM*, 52(2):172–216, 2005.

[SZ99] M. Saks and S. Zhou. BPSPACE(S) $\subseteq$ DSPACE($S^{3/2}$). *J. Comput. Syst. Sci.*, 58(2):376–403, 1999.

[Tre01] L. Trevisan. Extractors and pseudorandom generators. *J. ACM*, 48(4):860–879, July 2001.

[TSZ01] A. Ta-Shma and D. Zuckerman. Extractor codes. In ACM, editor, *Proceedings of STOC*, pages 193–199, 2001.

[TSZS01] A. Ta-Shma, D. Zuckerman, and S. Safra. Extractors from Reed-Muller codes. In *Proceedings of FOCS*, pages 638–647, 2001.

[Uma99] C. Umans. Hardness of approximating Σ_2^p minimization problems. In *Proceedings of FOCS*, pages 465–474, 1999.

[Uma03] C. Umans. Pseudo-random generators for all hardnesses. *J. Comput. Syst. Sci.*, 67(2):419–440, 2003.

[WZ93] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. In *Proceedings of STOC*, pages 245–251, 1993.

[Yao82] A. C. Yao. Theory and applications of trapdoor functions. In *Proceedings of the 23rd Symposium on Foundations of Computer Science (FOCS)*, pages 80–91. IEEE Computer Society Press, 1982.

[Zip79] R. E. Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of EUROSAM 79*, pages 216–226, 1979.

[Zuc96] D. Zuckerman. On unapproximable versions of NP -complete problems. *SICOMP*, 25(6):1293–1304, December 1996.

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

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/