Local Decoding and Testing for Homomorphisms[⋆]

Elena Grigorescu, Swastik Kopparty, and Madhu Sudan

Massachusetts Institute of Technology, Cambridge, MA, USA.
{elena_g,swastik,madhu}@mit.edu

Abstract. Locally decodable codes (LDCs) have played a central role in many recent results in theoretical computer science. The role of finite fields, and in particular, low-degree polynomials over finite fields, in the construction of these objects is well studied. However the role of group homomorphisms in the construction of such codes is not as widely studied. Here we initiate a systematic study of local decoding of codes based on group homomorphisms. We give an efficient list decoder for the class of homomorphisms from any abelian group G to a fixed abelian group H. The running time of this algorithm is bounded by a polynomial in $\log |G|$ and an agreement parameter, where the degree of the polynomial depends on H. Central to this algorithmic result is a combinatorial result bounding the number of homomorphisms that have large agreement with any function from G to H. Our results give a new generalization of the classical work of Goldreich and Levin, and give new abstractions of the list decoder of Sudan, Trevisan and Vadhan. As a by-product we also derive a simple(r) proof of the local testability (beyond the Blum-Luby-Rubinfeld bounds) of homomorphisms mapping $\mathbb{Z}_p^n$ to $\mathbb{Z}_p$, first shown by M. Kiwi.

1 Introduction

Given a pair of finite groups $G = (G, +)$ and $H = (H, \cdot)$, the class of homomorphisms between G and H forms an "error-correcting code". Namely, for any two distinct homomorphisms $\phi, \psi : G \to H$, the fraction of elements $\alpha \in G$ such that $\phi(\alpha) = \psi(\alpha)$ is at most $1/2$. This observation has implicitly driven the quest for many "homomorphism testers" [3, 2, 8, 1, 13], which test to see if a function $f : G \to H$ given as an oracle is close to being a homomorphism. In this paper, we investigate the complementary "decoding" question: Given oracle access to a function $f : G \to H$ find all homomorphisms $\phi : G \to H$ that are close to f.

To define the questions we study more precisely, let $\mathrm{agree}(f, g)$ denote the agreement between $f, g : G \to H$, i.e., the quantity $\Pr_{x \leftarrow_U G}[f(x) =$

[⋆] Research supported in part by NSF Award CCR-0514915.

$g(x)]$. Let $\mathrm{Hom}(G, H) = \{\phi : G \to H \mid \phi(x+y) = \phi(x)\phi(y)\}$ denote the set of homomorphisms from G to H. We consider the *combinatorial* question: Given G, H and $\epsilon > 0$, what is the largest "list" of functions that can have ϵ-agreement with some fixed function, i.e, what is $\max_{f:G\to H} |\{\phi : G \to H | \phi \in \mathrm{Hom}(G, H), \mathrm{agree}(f, g) \geq \epsilon\}|$?

We also consider the algorithmic question: Given G, H, $\epsilon > 0$ and oracle access to a function $f : G \to H$, (implicitly) compute a list of all homomorphisms $\phi : G \to H$ that have agreement ϵ with f. (A formal definition of implicit decoding will be given later. For now, we may think of this as trying to compute the value of ϕ on a set of generators of G.) We refer to this as the "local decoding" problem for homomorphisms.

Local decoding of homomorphisms for the special case of $G = \mathbb{Z}_2^n$ and $H = \mathbb{Z}_2$ was the central technical problem considered in the seminal work of Goldreich and Levin [4]. They gave combinatorial bounds showing that for $\epsilon = \frac{1}{2} + \delta$, the list size is bounded by $\mathrm{poly}(1/\delta)$, and gave a local decoding algorithm with running time $\mathrm{poly}(n/\delta)$.

The work of Goldreich and Levin was previously abstracted as decoding the class of degree one n-variate polynomials over the field of two elements. This led Goldreich, Rubinfeld, and Sudan [5] to generalize the decoding algorithm to the case of degree one polynomials over any finite field. (In particular, this implies a decoding algorithm for homomorphisms from $G = \mathbb{Z}_p^n$ to $H = \mathbb{Z}_p$, that decodes from $\frac{1}{p} + \epsilon$ agreement and runs in time $\mathrm{poly}(n/\epsilon)$, where $\mathbb{Z}_p$ denotes the additive group of integers modulo a prime p.) Later Sudan, Trevisan, and Vadhan [11], generalized the earlier results to the case of higher degree polynomials over finite fields . This generalization, in turn led to some general reductions between worst-case complexity and average-case complexity.

Our work is motivated by the group-theoretic view of Goldreich and Levin, as an algorithm to decode group homomorphisms. While the group-theoretic view has been applied commonly to the complementary problem of "homomorphism testing", the decoding itself does not seem to have been examined formally before.

To motivate we start with a simple example.

Consider the case where $G = \mathbb{Z}_p^n$ and $H = \mathbb{Z}_p^m$. How many homomorphisms can have agreement $\frac{1}{p} + \delta$ with a fixed function $f : G \to H$? Most prior work in this setting used (versions) of the Johnson bound in coding theory. Unfortunately such a bound only works for agreement greater than $\frac{1}{\sqrt{p}}$ in this setting.[1] An ad-hoc counting argument gives a better bound on the list size of $\delta^{-O(m)}$. While better bounds ought to be possible, none are known, illustrating the need for further techniques. Our work exposes several such questions. It also sheds new light on some of the earlier algorithms.

Our results. Our results are restricted to the case of abelian groups G and H. Let $\Lambda = \Lambda_{G,H}$ denote the maximum possible agreement between two homomorphisms from G to H. Our main algorithmic result is an efficient algorithm, with running time $\mathrm{poly}(\log |G|, \frac{1}{\epsilon})$ to decode all homomorphisms with agreement $\Lambda + \epsilon$ with a function $f : G \to H$ given as an oracle, for any *fixed* group H. Note that in such a case the polynomial depends on H. See Theorem 2 for full details.

Crucial to our algorithmic result is a corresponding combinatorial one showing that there are at most $\mathrm{poly}(\frac{1}{\epsilon})$ homomorphisms with agreement $\Lambda_{G,H} + \epsilon$ with any function $f : G \to H$, for any fixed group H. Once again, the polynomial in the bound depends on H. See Theorem 1 for details.

Finally, we also include a new proof of a result of Kiwi [8] on testing homomorphisms from $\mathbb{Z}_p^n$ to $\mathbb{Z}_p$. This is not related to our main quests, but we include it since some of the techniques we use to decode homomorphisms yield a simple proof of this result. See Theorem 3.

Organization of this paper. In Section 2 we present basic terminology and our main results. In Section 3 we exploit the decomposition theorem for abelian groups to reduce the proofs of the main theorems to the special case of p-groups. In Section 4 we tackle the combinatorial problem of the list-size for p-groups. In Section 5 we consider the corresponding algorithmic problem. Section 6 analyzes a homomorphism tester for functions from $\mathbb{Z}_p^n$ to $\mathbb{Z}_p$ using some techniques of the previous sections.

2 Definitions and Main Results

Let G, H be abelian groups, and let $\mathrm{Hom}(G, H) = \{h : G \to H \mid h \text{ is a homomorphism}\}$. Note that $\mathrm{Hom}(G, H)$ forms a *code.* Indeed, if $f, g \in \mathrm{Hom}(G, H)$, then $G' = \{x \mid f(x) = g(x)\}$ is a subgroup of G. Since the largest subgroup of G has size at most $\frac{|G|}{2}$, it follows that f and g

[1] For those familiar with the application of the Johnson bound in the setting of $m = 1$, we point out that it relied crucially on the fact that the agreement of any pair of homomorphisms was $\frac{1}{|H|}$ which is no longer true when $m \neq 1$.

differ in at least $\frac{1}{2}$ of the domain.
For two functions $f, g : G \to H$, define

$$\text{agree}(f, g) = Pr_{x \leftarrow_U G}[f(x) = g(x)],$$

and

$$\Lambda_{G,H} = \max_{f,g \in \text{Hom}(G,H), f \neq g} \{\text{agree}(f,\ g)\}.$$

In the case when $\text{Hom}(G, H)$ contains only the trivial homomorphism we define $\Lambda_{G,H} = 0$.

The notions of decodability and local list decoders are standard in the context of error correcting codes. Below we formulate them for the case of group homomorphisms.

Definition 1. *[11] (List decodability) The code* $\text{Hom}(G, H)$ *is* (δ, l)*-list decodable if for every function* $f : G \to H$*, there exist at most* l *homomorphisms* $h \in \text{Hom}(G, H)$ *such that* $\text{agree}(f, h) \geq \delta$*.*

Definition 2. *[14](Local list decoding) A probabilistic oracle algorithm* $\mathcal{A}$ *is a* (δ, T) *local list decoder for* $\text{Hom}(G, H)$ *if given oracle access to any function* $f : G \to H$*, (notation* $\mathcal{A}^f$*), the following hold:*

1. *With probability* $\frac{3}{4}$ *over the random choices of* $\mathcal{A}^f$*,* $\mathcal{A}^f$ *outputs a list of probabilistic oracle machines* $M_1, \ldots, M_L$ *s.t., for any homomorphism* $h \in \text{Hom}(G, H)$ *with* $\text{agree}(f, h) \geq \delta$*,*

$$\exists j \in [L], \forall x, \ \ \Pr[M_j^f(x) = h(x)] \geq \frac{3}{4},$$

   where the probability is taken over the randomness of $M_j^f(x)$*.*
2. $\mathcal{A}$ *and each* M_j^f *run in time* T*.*

The model of computation with respect to groups is as follows. An abelian group G can be represented (see Sect. 3) by its cyclic decomposition $\mathbb{Z}_{p_1^{e_1}} \times \ldots \times \mathbb{Z}_{p_k^{e_k}}$, where p_i's are prime. An element of G is given by a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, with $\alpha_i \in \mathbb{Z}_{p_i^{e_i}}$.

Our main results are the list decodability and local list decodability of group homomorphism codes.

Theorem 1. *Let* H *be a fixed finite abelian group. Then for all finite abelian groups* G*,* $\text{Hom}(G, H)$ *is* $\left(\Lambda_{G,H} + \epsilon, \text{poly}_{|H|}(\frac{1}{\epsilon})\right)$ *list decodable.*

Remark: The exact polynomial bound on the list size that our proof gives, in general, depends on the structure of the groups in an intricate way, but can nevertheless be uniformly bounded by $O\left(\frac{1}{\epsilon^{4\log|H|}}|H|^5\right)$. Still, the precise bounds obtained by the proof are not optimal. For example, our proof gives that $\mathrm{Hom}(\mathbb{Z}_2^n, \mathbb{Z}_2^2)$ is $(\frac{1}{2}+\epsilon, O(\frac{1}{\epsilon^4}))$ list decodable, while it can be shown (via alternate means) that it is $(\frac{1}{2}+\epsilon, O(\frac{1}{\epsilon^2}))$ list decodable.

Theorem 2. *Let H be a fixed finite abelian group. Then for all finite abelian groups G there is a $(\Lambda_{G,H}+\epsilon, \mathrm{poly}_{|H|}(\log|G|, \frac{1}{\epsilon}))$ local list decoder for* $\mathrm{Hom}(G,H)$*.*

3 Decomposition and Reduction

We will embark on our quest by first decomposing the groups involved into slightly smaller but better-behaved groups. In this section we will see how these decompositions can be done and thereby reduce our main theorems to statements about list decoding on "p-groups". These statements will be proved in the following two sections by some Fourier analytic machinery and by generalizing the STV-style list decoders.

The structure theorem for finite abelian groups states that every abelian group G is of the form $\prod_{i=1}^{k}\mathbb{Z}_{p_i^{e_i}}$, where the p_i's are primes and the e_i's are positive integers. A *p-group* is a group of order p^r, for some positive integer r. The structure theorem implies that for any prime p, any finite abelian group G can be written as $G_p \times G'$, where G_p is a p-group and $\gcd(p,|G'|)=1$ (take $G_p = \prod_{p_i=p}\mathbb{Z}_{p_i^{e_i}}$). This decomposition will play a crucial role in what follows.

Remark 1. $\Lambda_{G,H}$ behaves well under decomposition of G and H:

1. If $\gcd(|G|,|H|)=1$ then $\mathrm{Hom}(G,H)$ contains only the trivial homomorphism and therefore, $\Lambda_{G,H}=0$.
2. Otherwise, let p be the smallest prime s.t. $p \mid \gcd(|G|,|H|)$. Then $\Lambda_{G,H}=\frac{1}{p}$.
   Indeed, it is enough to bound $\mathrm{agree}(h,\mathbf{0})$, for any nontrivial homomorphism $h: G \to H$. Let $d = |\mathrm{image}\ (h)|$ and note that $d \mid |H|$, since $\mathrm{image}(h)$ is a subgroup of H. Since $G/\ker(h) \cong\ \mathrm{image}(h)$, it follows that $|\ker(h)|/|G| = 1/d \le 1/p$, and thus $\Lambda_{G,H}\le\frac{1}{p}$.
   Finally, if $G=\mathbb{Z}_{p^t}\times G'$, and $H=\mathbb{Z}_{p^r}\times H'$, then the homomorphism $h: G\to H$ definde by $h(a,b)=(ap^{r-1},0)$ satisfies $\mathrm{agree}(h,\mathbf{0})=\frac{1}{p}$. Hence, $\Lambda_{G,H}=\frac{1}{p}$.
3. The above observations imply $\Lambda_{G_1\times G_2,H}=\max\{\Lambda_{G_1,H},\Lambda_{G_2,H}\}$ and $\Lambda_{G,H_1\times H_2}=\max\{\Lambda_{G,H_1},\Lambda_{G,H_2}\}$.

3.1 The decompositions $G \to H_1 \times H_2$ and $G_1 \times G_2 \to H$

The following two propositions (whose proofs are omitted from this version) say that list decoding questions for $Hom(G, H)$ can be reduced to list decoding questions on summands of G or H.

Proposition 1. *Let G, H_1, H_2 be abelian groups. Let $a_i = \Lambda_{G,H_i}$. Suppose for all $\epsilon > 0$, $\mathrm{Hom}(G, H_i)$ is $(a_i + \epsilon, \ell_i(\epsilon))$-list decodable, with $(a_i + \epsilon, T_i(\epsilon))$ local list decoders, for $i = 1, 2$. Then $\mathrm{Hom}(G, H_1 \times H_2)$ is $(\max\{a_1, a_2\} + \epsilon, \ell_1(\epsilon)\ell_2(\epsilon))$ list decodable and has a $(\max\{a_1, a_2\} + \epsilon, O\left((T_1(\epsilon)T_2(\epsilon))\right))$ local list decoder, for all $\epsilon > 0$.*

Proposition 2. *Let G_1, G_2, H be abelian groups. Let $a_i = \Lambda_{G_i,H}$. Suppose for all $\epsilon > 0$, $\mathrm{Hom}(G_i, H)$ is $(a_i + \epsilon, \ell_i(\epsilon))$-list decodable, with a $(a_i + \epsilon, T_i(\epsilon))$ local list decoder, for $i = 1, 2$. Then $\mathrm{Hom}(G_1 \times G_2, H)$ is $(\max\{a_1, a_2\}+\epsilon, O(\frac{1}{\epsilon^2}\ \ell_1(\epsilon)\ell_2(\epsilon)\ |H|^2))$ list decodable, and has a $(\max\{a_1, a_2\}+ \epsilon,\ O(\frac{|H|}{\epsilon^2}\ (T_1(\epsilon) + T_2(\epsilon)) + \ell_1(\epsilon)\ell_2(\epsilon)\ |H|^2)$ local list decoder, for all $\epsilon > 0$.*

3.2 Proof of the main theorems

Using the propositions proved in the previous section, our theorems will reduce to the main lemma given below. A proof is sketched in Section 4.

Lemma 1. *Let p be a fixed prime and $r > 0$ be a fixed integer. Then for any abelian p-group G, $\mathrm{Hom}(G, \mathbb{Z}_{p^r})$ is $\left(\frac{1}{p} + \epsilon, (2p)^{3r}\frac{1}{\epsilon^2}\right)$ list decodable.*

In Section 5, we shall use it to prove the corresponding algorithmic version.

Lemma 2. *Let p be a fixed prime and $r > 0$ be a fixed integer. Then for any abelian p-group G, $\mathrm{Hom}(G, \mathbb{Z}_{p^r})$ is $\left(\frac{1}{p} + \epsilon, \mathrm{poly}(\log|G|, \frac{1}{\epsilon})\right)$ locally list decodable.*

Proof (of Theorem 1). If $|G|, |H|$ are relatively prime then the result is obvious. Otherwise, let $p(= \frac{1}{\Lambda_{G,H}})$ be the smallest prime dividing both $|G|$ and $|H|$. Let $H = \prod_{i=1}^{r} \mathbb{Z}_{p_i^{\beta_i}}$. Let $i \in \{1, \ldots, r\}$. If $\gcd(p_i, |G|) = 1$, then $\mathrm{Hom}(G, \mathbb{Z}_{p_i^{\beta_i}})$ is $(\epsilon, 1)$ list decodable. Otherwise, write G as $G_{p_i} \times G'$, where G_{p_i} is a p_i-group and $\gcd(p_i, |G'|) = 1$. Then by Lemma 1 and Proposition 2, $\mathrm{Hom}(G, \mathbb{Z}_{p_i^{\beta_i}})$ is $\left(\frac{1}{p_i} + \epsilon, O(\frac{1}{\epsilon^4}(2p_i)^{3\beta_i}p^{2\beta_i})\right)$ list decodable, and hence is also $\left(\frac{1}{p} + \epsilon, \frac{1}{\epsilon^4}p_i^{5\beta_i}\right)$ list decodable (since if $p_i||G|$, then $p \le p_i$). Combining these for all $i \in \{1, \ldots, r\}$ by Proposition 1, $\mathrm{Hom}(G, H)$ is $\left(\frac{1}{p} + \epsilon, \prod_{p_i||G|} \frac{1}{\epsilon^4}(2p_i)^{5\beta_i}\right)$ list decodable, as required.

Proof (of Theorem 2). The proof of this theorem is directly analogous to the previous proof, using Lemma 2 instead of Lemma 1.

4 Combinatorial bounds for p-groups

In this section we will briefly touch upon how our main lemma (Lemma 1) is proved. Recall that we wish to obtain a combinatorial upper bound on the number of homomorphisms having agreement $\frac{1}{p}+\epsilon$ with a function $f : G \to \mathbb{Z}_{p^r}$, where G is a p-group. The starting point for our proof is the observation that $\mathbb{Z}_{p^r}$ is isomorphic to the multiplicative group μ_{p^r}, a subgroup of the complex numbers consisting of the p^rth roots of unity. This makes the tools of Fourier analysis available to us.

4.1 Sketch of the argument

In this version we only give a sketch of the proof at a very high level. We are given a function $f : G \to \mathbb{Z}_{p^r}$. We begin by giving a formula that expresses the agreement between our function and any given homomorphism in terms of Fourier coefficients of some functions related to f. This will imply that every homomorphism having high agreement with f "corresponds" to some large Fourier coefficient. Now Parseval's identity tells us that there can only be few large Fourier coefficients, and the end of the proof looks near. Unfortunately, it is possible that many distinct homomorphisms "correspond" to the same Fourier coefficients. Nevertheless, we will be able to bound the number of occurences of the above pathology in terms of the number of homomorphisms in $\mathrm{Hom}(G, \mathbb{Z}_{p^l})$ that have high agreement with a related function $f' : G \to \mathbb{Z}_{p^l}$, for some $l < r$. Thus, inducting on r, we will arrive at the result.

In the proof we use the following version of the Johnson bound, which is the base case for the induction, and is also useful in Section 6.

Proposition 3. *Let G be a p-group. Then*

1. $\mathrm{Hom}(G, \mu_p)$ *is* $(\frac{1}{p}+\epsilon, \frac{1}{\epsilon^2})$ *list decodable, for any* $\epsilon > 0$.
2. *Let* $f : G \to \mu_p$ *and* $\rho_t = \mathrm{agree}(f, \chi_t)$ *for* $\chi_t \in \mathrm{Hom}(G, \mu_p)$*, then*

$$\sum_{\chi_t \in \mathrm{Hom}(G,\mu_p)} \left(\rho_t - \frac{1}{p-1}(1-\rho_t)\right)^2 \leq 1.$$

5 Algorithmic results for p-groups

In this section we will turn our attention to the algorithmic decoding question suggested by the combinatorial results of the previous section. Here we will show Lemma 2 stated in Section 3.

Lemma 2. *Let p be a fixed prime and $r > 0$ be a fixed integer. Then for any abelian p-group G, $Hom(G, \mathbb{Z}_{p^r})$ is $\left(\frac{1}{p} + \epsilon, \mathrm{poly}(\log |G|, \frac{1}{\epsilon})\right)$ locally list decodable.*

We will provide an algorithm which, given access to a function $f : G \to \mathbb{Z}_{p^r}$, with G a p-group, outputs an implicit representation of the homomorphisms that agree in a $\frac{1}{p} + \epsilon$ with f. Intuitively, to get the value of such a homomorphism $h \in Hom(G, \mathbb{Z}_{p^r})$ at a point x, we restrict our attention to a random coset of a random subgroup of G that contains x. Provided that h restricted to this coset has agreement at least $\frac{1}{p} + \epsilon/2$ with f, we can deduce the value of $h(x)$. Along the way we prove a lemma that says that random cosets of a random subgroup of a p-group "sample well", which is shown using the second moment method.

5.1 Cosets of subgroups generated by enough elements sample well

Definition 3. *Let G be an abelian group, and let $z_1, \ldots, z_k \in G$. Define $S_{z_1,\ldots,z_k}$ to be the subgroup of G generated by $z_1, \ldots, z_k$.*

Before giving our decoding algorithms, we state a useful lemma (whose proof is omitted in this version).

Lemma 3. *Let G be an abelian p-group, let $A \subseteq G$, with $\mu = \frac{|A|}{|G|}$ and let $x, z_1, \ldots z_k \in G$ be picked uniformly at random. Then*

$$Pr_{x,z_1,\ldots,z_k}\left[\ \left|\frac{|A \cap (x + S_{z_1,\ldots,z_k})|}{|S_{z_1,\ldots,z_k}|} - \mu\right| > \epsilon\right] \leq \frac{1}{\epsilon^2 p^k}.$$

5.2 The generalized STV algorithm

We begin with a simple but useful observation [3]: homomorphisms have simple and efficient self-correctors, i.e., for $g : G \to H$, there is a randomized procedure $Corr^g : G \to H$ running in time $\mathrm{poly}(\log |G|)$ satisfying the following property

- *Self-corrector:* If $g : G \to H$ is such that there is some homomorphism $h : G \to H$ with agree$(g, h) > 7/8$, then with for all $x \in G$, $Corr^g(x) = h(x)$ with probability $> 3/4$.

Let $R_{x,z_1,\ldots,z_k}$ be the set $x + S_{z_1-x,\ldots,z_k-x}$, i.e., the "affine subspace" passing through $x, z_1, \ldots, z_k$. Let $r_{x,z_1,\ldots,z_k} : [T]^k \to (x + S_{z_1-x,\ldots,z_k-x})$ be the parametrization of $R_{x,z_1,\ldots,z_k}$ given by:

$$r_{x,z_1,\ldots,z_k}(\bar{\alpha}) = x + \sum_i \alpha_i (z_i - x).$$

For a function $g : G \to H$, define the restriction $g|_{R_{x,z_1,\ldots,z_k}} : [T]^k \to H$ by $g|_{R_{x,z_1,\ldots,z_k}}(\bar{\alpha}) = g(r_{x,z_1,\ldots,z_k}(\bar{\alpha}))$. Notice that when we restrict homomorphisms to a set of the form $R_{x,z_1,\ldots,z_k}$, we get an *affine homomorphism*, i.e., a function of the form $h + b$ where h is a homomorphism and $b \in H$.

```
The oracle $M^f_{z_1,\ldots,z_k,a_1,\ldots,a_k}(x)$:
For $b \in H$, define $h_b : [T]^k \to H$ by $h_b(\bar{\alpha}) = b + \sum \alpha_i (z_i - x)$.
1: For each $b$ in $H$, estimate (by random sampling)
$l_b = \text{agree}(f|_{R_{x,z_1,\ldots,z_k}}, h_b)$.
2: If there is exactly one $b$ with $l_b > \frac{1}{p} + \frac{\epsilon}{4}$ then output $b$, else fail.
```

```
The local list decoder:
Repeat $O(1)$ times:
1: Pick $z_1, \ldots, z_k \in G$ uniformly and independently at random, where $k = c_1 \log_p \frac{1}{\epsilon}$.
2: For each $(a_1, \ldots, a_k) \in H^k$, output $Corr^{M^f_{z_1,\ldots,z_k,a_1,\ldots,a_k}}$.
```

The analysis of the list-decoding algorithm is similar to that of [11] and we omit it in this version. It leads to the following lemma.

Lemma 4. *If h is a homomorphism s.t. agree$(h, f) \geq \frac{1}{p} + \epsilon$ then*

$$Pr_x[M^f_{z_1,\ldots,z_k,h(z_1),\ldots,h(z_k)}(x) = h(x)] \geq 7/8,$$

with probability $\frac{1}{2}$ over the choice of $z_1, \ldots, z_k \in G$.

Proof of Lemma 2

Let h be a homomorphism that agrees with f on a $\frac{1}{p} + \epsilon$ fraction of points. Consider the oracle $M^f_{z_1,\ldots,z_k,h(z_1),\ldots,h(z_k)}$ (where the a_i are "consistent" with h). By Lemma 4, $M^f_{z_1,\ldots,z_k,h(z_1),\ldots,h(z_k)}(x)$ is correct on at least

$\frac{15}{16} > \frac{7}{8}$ of the $x \in G$, and thus $Corr^{M^f_{z_1,\ldots,z_k,h(z_1),\ldots,h(z_k)}}$ computes h on all of G with probability at least $\frac{3}{4}$. It follows that each high-agreement homomorphism will appear w.h.p in the final list if the execution of the algorithm is repeated a constant number of times. This completes the proof of the lemma.

6 Homomorphism tester

In this section we will prove a result of Kiwi using techniques related to Section 4. The result says that the 3 query linearity tester given below for homomorphisms in $\mathrm{Hom}(\mathbb{Z}_p^n, \mu_p)$ has very good acceptance probability/maximum agreement trade-offs. In particular, its performance is far better than that of the BLR [3] test for $p > 2$.

Given $f : \mathbb{Z}_p^n \to \mu_p$.

We are analyzing the following linearity test:

- Pick $x, y \in \mathbb{Z}_p^n$, $\alpha, \beta \in \mathbb{Z}_p^*$ uniformly at random
- Accept if $f(\alpha x + \beta y) = f(x)^\alpha f(y)^\beta$, else reject.

Kiwi [8] analyzed this test to get the following theorem.

Theorem 3. *Suppose f passes the above test with probability δ, then f has agreement at least δ with some homomorphism in* $\mathrm{Hom}(\mathbb{Z}_p^n, \mu_p)$.

In fact, [8] proved a more general result for testing vector-space homomorphisms over any finite field $\mathbb{F}_q^n \to \mathbb{F}_q$, not necessarily over prime fields. His proof uses the MacWilliams identities and properties of the Krawchouk polynomials. Here we give a simple proof of the above theorem using elementary Fourier analysis. Our proof also generalizes to the case of vector-space homomorphisms (using Trace functions) though we don't include the proof in this version.

Proof. The proof will use Fourier analysis, and modeled along the general lines of the argument in [2] (i.e., expressing agreement and acceptance probabilities in terms of Fourier coefficients).

For $\eta \in \mu_p$, define $\mathcal{S}(\eta) = \mathbb{E}_{c \in \mathbb{Z}_p^*}[\eta^c]$. It is easily seen that

$$\mathcal{S}(\eta) = \begin{cases} 1, & \text{if } \eta = 1 \\ \frac{-1}{p-1}, & \text{otherwise} \end{cases}$$

Recall that every homomorphism from $\mathbb{Z}_p^n \to \mu_p$ is a character χ_t for some $t \in Z_p^n$, where $\chi_t(x) = e^{2\pi i (t \cdot x)/p}$. For $f : \mathbb{Z}_p^n \to \mathbb{C}$, the Fourier coefficient $\hat{f}(t)$ is defined to be $\mathbb{E}_{x\in\mathbb{Z}_p^n} f(x)\overline{\chi_t}(x)$. We will assume some familiarity with basic properties of characters and Fourier coefficients in this version of the paper.

For $t \in \mathbb{Z}_p^n$ let ρ_t be the agreement of f with χ_t. We shall prove that $\delta \leq \max_{t\in\mathbb{Z}_p^n} \rho_t$. This will prove the result.

We begin by finding an explicit formula for ρ_t in terms of the Fourier coefficients.

$$\rho_t - \frac{1}{p-1}(1-\rho_t) = \mathbb{E}_{x\in\mathbb{Z}_p^n}[S(f(x)\overline{\chi_t}(x))] = \mathbb{E}_{x\in\mathbb{Z}_p^n, c\in\mathbb{Z}_p^*}[f(x)^c\overline{\chi_t}(x)^c] \tag{1}$$

$$= \mathbb{E}_{c\in\mathbb{Z}_p^*}\mathbb{E}_{x\in\mathbb{Z}_p^n}[f(x)^c\overline{\chi_{ct}}(x)] = \mathbb{E}_{c\in F_p^*}[\widehat{f^c}(ct)] \tag{2}$$

We now find a similar formula for δ and perform some manipulations that allow us to relate it to our formula for ρ_t.

$$\delta - \frac{1}{p-1}(1-\delta) = \mathbb{E}_{x,y\in\mathbb{Z}_p^n}\mathbb{E}_{\alpha,\beta\in\mathbb{Z}_p^*}\left[S\left(f(x)^\alpha f(y)^\beta f(\alpha x+\beta y)^{-1}\right)\right] \tag{3}$$

$$= \mathbb{E}_{x,y\in\mathbb{Z}_p^n}\mathbb{E}_{\alpha,\beta\in\mathbb{Z}_p^*}\left[\mathbb{E}_{c\in\mathbb{Z}_p^*}[f(x)^{c\alpha}f(y)^{c\beta}f(\alpha x+\beta y)^{-c}]\right] \tag{4}$$

$$= p^n\mathbb{E}_{x,y,z}\mathbb{E}_{\alpha',\beta',\gamma'}\left[f(x)^{\alpha'}f(y)^{\beta'}f(z)^{\gamma'}\mathbf{1}(\alpha' x+\beta' y+\gamma' z=0)\right] \tag{5}$$

where we substituted $\alpha' = c\alpha, \beta' = c\beta, \gamma' = -c, z = \alpha x + \beta y$ (and one verifies that $z = \alpha x + \beta y$ is equivalent to $\alpha' x + \beta' y + \gamma' z = 0$). Note that since $\gamma' \in \mathbb{Z}_p^*$, the probability that a random $z \in \mathbb{Z}_p^n$ is such that $\alpha' x + \beta' y + \gamma' z = 0$ is $\frac{1}{p^n}$.

$$\begin{aligned}
(5) &= p^n\mathbb{E}_{x,y,z}\mathbb{E}_{\alpha',\beta',\gamma'}\left[f(x)^{\alpha'}f(y)^{\beta'}f(z)^{\gamma'}\mathbb{E}_{t\in\mathbb{Z}_p^n}[\overline{\chi_t}(\alpha' x+\beta' y+\gamma' z)]\right]\\
&= p^n\mathbb{E}_t\left[\mathbb{E}_{\alpha',\beta',\gamma'}\mathbb{E}_x\left[f(x)^{\alpha'}\overline{\chi_{\alpha' t}}(x)\right]\mathbb{E}_y\left[f(y)^{\beta'}\overline{\chi_{\beta' t}}(y)\right]\mathbb{E}_z\left[f(z)^{\gamma'}\overline{\chi_{\gamma' t}}(z)\right]\right]\\
&= \sum_t\left[\mathbb{E}_{\alpha',\beta',\gamma'}\left[\widehat{f^{\alpha'}}(\alpha' t)\widehat{f^{\beta'}}(\beta' t)\widehat{f^{\gamma'}}(\gamma' t)\right]\right]\\
&= \sum_t\left(\mathbb{E}_{\alpha'\in\mathbb{Z}_p^*}[\widehat{f^{\alpha'}}(\alpha' t)]\right)^3\\
&= \sum_t\left(\rho_t - \frac{1}{p-1}(1-\rho_t)\right)^3 \qquad \text{(By (2))}
\end{aligned}$$

Simplifying the last expression and using Proposition 3 we get $\delta \leq \max_t \rho_t$.

Acknowledgments

Thanks to Amir Shpilka for many valuable discussions.

References

1. Michael Ben-Or, Don Coppersmith, Michael Luby, Ronitt Rubinfeld, Non-Abelian Homomorphism Testing, and Distributions Close to their Self-Convolutions. RANDOM 2004.
2. Mihir Bellare and Don Coppersmith and Johan Håstad and Marcos Kiwi and Madhu Sudan. Linearity testing over characteristic two. *IEEE Transactions on Information Theory*, 42(6), 1781-1795, 1996.
3. Manuel Blum and Michael Luby and Ronitt Rubinfeld. Self-Testing/Correcting with Applications to Numerical Problems. *Journal of Computer and System Sciences*, 47(3), 549-595, 1993.
4. Oded Goldreich and Leonid Levin. A hard-core predicate for all one-way functions. *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, 25–32, 1989
5. Oded Goldreich and Ronitt Rubinfeld and Madhu Sudan. Learning polynomials with queries: The highly noisy case. *SIAM Journal on Discrete Mathematics,* 13(4):535-570, 2000.
6. Venkatesan Guruswami and Madhu Sudan. List decoding algorithms for certain concatenated codes. *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, 181-190, 2000.
7. Marcos Kiwi , Frédéric Magniez , Miklos Santha. Exact and approximate testing/correcting of algebraic functions: A survey. *Theoretical Aspects of Computer Science,* Teheran, Iran, Springer-Verlag, LNCS 2292, 30-83, 2002.
8. Marcos Kiwi. Testing and weight distributions of dual codes. *Theoretical Computer Science*, 299(1–3):81-106, 2003.
9. Eyal Kushilevitz and Yishay Mansour. Learning decision trees using the Fourier spectrum. SIAM Journal on Computing 22(6):1331-1348, 1993.
10. Dana Moshkovitz, Ran Raz. Sub-Constant Error Low Degree Test of Almost Linear Size, STOC 2006.
11. Madhu Sudan and Luca Trevisan and Salil Vadhan. Pseudorandom generators without the XOR lemma, *Proceedings of the 31st Annual ACM Symposium on Theory of Computing* 537-546, 1999.
12. Madhu Sudan. Algorithmic Introduction to Coding Theory. Lecture Notes, 2001.
13. Amir Shpilka and Avi Wigderson. Derandomizing Homomorphism Testing in General Groups. *Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 427-435, 2004.
14. L. Trevisan. Some Applications of Coding Theory in Computational Complexity. Survey Paper. *Quaderni di Matematica* 13:347-424, 2004