Electronic Colloquium on Computational Complexity, Report No. 185 (2015)

Lower bounds for constant query affine-invariant LCCs and LTCs

Arnab Bhattacharyya*
Indian Institute of Science
arnabb@csa.iisc.ernet.in

Sivakanth Gopi†
Princeton University
sgopi@cs.princeton.edu

Abstract

Affine-invariant codes are codes whose coordinates form a vector space over a finite field and which are invariant under affine transformations of the coordinate space. They form a natural, well-studied class of codes; they include popular codes such as Reed-Muller and Reed-Solomon. A particularly appealing feature of affine-invariant codes is that they seem well-suited to admit local correctors and testers.

In this work, we give lower bounds on the length of locally correctable and locally testable affine-invariant codes with constant query complexity. We show that if a code $\mathcal{C} \subset \Sigma^{\mathbb{K}^n}$ is an r-query locally correctable code (LCC), where $\mathbb{K}$ is a finite field and Σ is a finite alphabet, then the number of codewords in $\mathcal{C}$ is at most $\exp(O_{\mathbb{K},r,|\Sigma|}(n^{r-1}))$. Also, we show that if $\mathcal{C} \subset \Sigma^{\mathbb{K}^n}$ is an r-query locally testable code (LTC), then the number of codewords in $\mathcal{C}$ is at most $\exp(O_{\mathbb{K},r,|\Sigma|}(n^{r-2}))$. The dependence on n in these bounds is tight for constant-query LCCs/LTCs, since Guo, Kopparty and Sudan (ITCS '13) construct affine-invariant codes via lifting that have the same asymptotic tradeoffs. Note that our result holds for non-linear codes, whereas previously, Ben-Sasson and Sudan (RANDOM '11) assumed linearity to derive similar results.

Our analysis uses higher-order Fourier analysis. In particular, we show that the codewords corresponding to an affine-invariant LCC/LTC must be far from each other with respect to Gowers norm of an appropriate order. This then allows us to bound the number of codewords, using known decomposition theorems which approximate any bounded function in terms of a finite number of low-degree non-classical polynomials, upto a small error in the Gowers norm.

*Research partially supported by a DST Ramanujan Fellowship.
†Research partially supported by NSF grants CCF-1523816 and CCF-1217416

ISSN 1433-8092

1 Introduction

Error-correcting codes which admit local algorithms are of significant interest in theoretical computer science. A code is called a locally correctable code (LCC) if there is a randomized algorithm that, given an index i and a received word w close to a codeword c in Hamming distance, outputs c_i by querying only a few positions of w. A code is called a locally testable code (LTC) if there is a randomized algorithm that, given a received word w, determines whether w is in the code or whether w is far in Hamming distance from every codeword, based on queries to a small number of locations of w. The number of positions of the received word queried is called the query complexity of the LCC or LTC.

The notions of local correctability and local testability have a long history in computer science by now. Also called "self-correction", the idea of local correction originated in works by Lipton [Lip90] and by Blum and Kannan [BK95] on program checkers. LCCs are closely related to locally decodable codes (LDCs), where the goal is to recover a symbol of the underlying message when given a corrupted codeword, using a small number of queries [KT00]. LDCs and LCCs have found applications in private information retrieval schemes [CKGS98, BIW07] and derandomization [STV99]. See [Yek11] for a detailed survey on LDCs and LCCs. Research on LTCs implicitly started with Blum, Luby, and Rubinfeld's seminal discovery [BLR93] that the Hadamard code is an LTC with query complexity 3; they were first formally defined by Goldreich and Sudan in [GS06]. LTCs have been used (implicitly and explicitly) in many contexts, most notably in the construction of PCP's [AS98, ALM+98, Din07].

In spite of the wide interest in them, some basic questions about LCCs and LTCs remain unanswered. We restrict ourselves throughout to the setting where the query complexity is a constant (independent of the length of the code) and consider the tradeoff between query complexity and code length. The current best constant-query LCCs have exponential length, while the current best constant-query LTCs have near-linear length but they are quite complicated [BS08, Din07, Mei09, Vid15]. Getting subexponential length LCCs or linear length LTCs with constant query complexity are major open problems in the area.

Intuitively, for LCCs and LTCs with constant query complexity, there must be a lot of redundancy in the code, since every symbol of the codeword must satisfy local constraints with most other symbols in the codeword. A systematic way to generate redundancy is to make sure that the code has a large group of *invariances*[*]. Formally, given a code $\mathcal{C} \subset \Sigma^N$ of length N over alphabet Σ, a codeword $c \in \mathcal{C}$ can be naturally viewed as a function $c : [N] \to \Sigma$. Then, we say that $\mathcal{C}$ is invariant under a set $G \subset \{[N] \to [N]\}$[†] if for every $\pi \in G$ and codeword $c \in \mathcal{C}$, $c \circ \pi$ also describes a codeword $c' \in \mathcal{C}$. Now, the key observation is that if for every codeword $c \in \mathcal{C}$, if there is a constraint among $c(i_1), \ldots, c(i_k)$ for some $i_1, \ldots, i_k \in [N]$, then for every $c \in \mathcal{C}$, there must also be a constraint among $c(\pi(i_1)), \ldots, c(\pi(i_k))$ for any π in the invariance set G, since $c \circ \pi$ is itself another codeword. Hence if G is large, the presence of one local constraint immediately implies presence of many and suggests the possibility of local algorithms for the code. This connection between invariance and correctability/testability was first explicitly examined by Kaufman and Sudan [KS08]. One is then motivated to understand more clearly the possibilities and limitations of local correctors/testers for codes possessing natural symmetries.

We focus on affine-invariant codes, for which the domain $[N]$ is an n-dimensional vector space $\mathbb{K}^n$ over a finite field $\mathbb{K}$ and the code $\mathcal{C} \subset \{\mathbb{K}^n \to \Sigma\}$ is invariant under affine transformations $A : \mathbb{K}^n \to \mathbb{K}^n$. Affine invariance is a very natural symmetry for "algebraic codes" and has long been studied in coding theory [KLP67]. The study of affine-invariant LCCs and LTCs was initiated in [KS08] and has been investigated in several follow-up works [BS11, Guo13, BRS12, GSVW15]. The hope is that because affine-invariant codes have a large group of invariance and, at the same time, are conducive to non-trivial algebraic constructions, they may contain a code that improves current constructions of LCCs or LTCs.

The current best parameters for constant-query affine-invariant LCCs and LTCs are achieved by the lifted codes of Guo, Kopparty and Sudan [GKS13]. They construct an affine-invariant code $\mathcal{F} \subset \{\mathbb{F}_{2^\ell}^n \to \mathbb{F}_2\}$ with $\exp(\Theta(n^{r-2}))$ codewords that is an $(r-1)$-query LCC and an r-query LTC, where $r = 2^\ell$. The $\Theta(\cdot)$ notation hides factors that depend on r but not n. For LCCs, the same asymptotic tradeoff between query complexity

[*] A quite different way to generate redundancy is through *tensoring*; see [BS04]. Invariances and tensoring are essentially the only two "generic" reasons known to cause local correctability/testability.

[†] $\{A \to B\}$ and B^A denote the set of all functions from A to B.

and code length is achieved by the Reed-Muller code. For every $r \geq 2$, the Reed-Muller code of order $r-1$ (i.e., polynomials over $\mathbb{F}_q$ on n variables of total degree $\leq r-1$ with $q > r$) is an affine-invariant r-query LCC with $\exp(\Theta(n^{r-1}))$ codewords. In fact, even if we drop the affine-invariance requirement, Reed-Muller codes and the construction of [GKS13] achieve the best known codeword length for constant query LCCs[‡].

In this work, we show that the parameters for the lifted codes of [GKS13] are, in fact, tight for affine-invariant LCCs/LTCs in $\{\mathbb{K}^n \to \Sigma\}$ for any fixed finite field $\mathbb{K}$ and any fixed finite alphabet Σ.

Theorem 1 (Main Result, informal)**.**

(i) Let $\mathcal{C} \subset \{\mathbb{K}^n \to \Sigma\}$ be an r-query affine-invariant LCC. Then $|\mathcal{C}| \leq \exp\left(O_{\mathbb{K},r,|\Sigma|}(n^{r-1})\right)$.

(ii) Let $\mathcal{C} \subset \{\mathbb{K}^n \to \Sigma\}$ be an r-query affine-invariant LTC. Then $|\mathcal{C}| \leq \exp\left(O_{\mathbb{K},r,|\Sigma|}(n^{r-2})\right)$.

1.1 Related Work

Ben-Sasson and Sudan in [BS11] obtained a similar result as Theorem 1, when the code is assumed to be linear, i.e., when the codewords form a vector space. They showed that if $\mathcal{C} \subset \{\mathbb{K}^n \to \mathbb{F}\}$ is an $(r-1)$-query locally correctable or r-query locally testable *linear*, affine-invariant code, where $\mathbb{K}$ and $\mathbb{F}$ are finite fields of characteristic $p > 0$ with $\mathbb{K}$ an extension of $\mathbb{F}$, then the dimension of $\mathcal{C}$ as a vector space over $\mathbb{F}$ is at most $(n \log_p |\mathbb{K}|)^{r-2}$. When $\mathbb{K}$ is fixed (as in [GKS13]'s construction of constant query LCCs/LTCs), the result of [BS11] is a very special case of our Theorem 1. On the other hand, [BS11]'s result also applies when the size of $\mathbb{K}$ is growing (as long as $\mathbb{K}$ extends $\mathbb{F}$), whereas ours does not.

There are several works which study lower bounds for constant query LCCs [KT00, GKST02, DS07, KdW03, BDYW11, BDSS11, Woo12, DSW14]. For general (non-affine-invariant) LCCs, tight lower bounds are known only for 2-query LCCs. Kerendis and deWolf [KdW03] prove that if $\mathcal{C} \subset \{\{0,1\}^n \to \Sigma\}$ is a 2-query LCC[§], then $|\mathcal{C}| \leq \exp(O(n|\Sigma|^5))$. This is tight for constant Σ and achieved by the Hadamard code. For r-query LCCs where $r > 2$, the lower bounds known are much weaker. The best known bounds, due to [KdW03, Woo07], show that if $\mathcal{C} \subset \{\{0,1\}^n \to \{0,1\}\}$ is an r-query LCC, then

$$|\mathcal{C}| \leq \exp\left(2^{n/(1+1/(\lceil r/2 \rceil+1))+o(n)}\right).$$

Higher-order Fourier analysis was applied to other problems in coding theory in [BL15b, TW14].

1.2 Proof Overview

Our arguments are based on standard techniques from higher-order Fourier analysis [Tao12], but they are new in this context. We show that if an affine-invariant code is an r-query LCC, then its codewords are far from each other in the U^r-norm, the *Gowers norm of order r*. Similarly, we show that the codewords of an affine-invariant r-query LTC are far from each other in the U^{r-1}-norm. Therefore, we can upper bound the number of LCC/LTC codewords in terms of the size of a net that is fine enough with respect to the Gowers norm of an appropriate order. We bound the size of such a net by explicitly constructing one using a standard decomposition theorem (analogous to Szemerédi's regularity lemma): any bounded function $f : \mathbb{K}^n \to \mathbb{C}$ can be approximated, upto a small error in the Gowers norm, by a composition of a bounded number of low-degree non-classical polynomials [TZ12].

The way we argue that two codewords f and g of an r-query LCC are far in the Gowers norm is that if $\|f-g\|_{U^r} < \epsilon$, then for small enough ϵ (with respect to r, $|\Sigma|$ and correctness probability), the local corrector when applied to f can act as if it is applied to g. The argument is, briefly, as follows. On the one hand, the codewords f and g must be far in Hamming distance, because the definition of LCC implies that there is a unique codeword close to any string. So, with constant probability over choice of $y \in \mathbb{K}^n$, the local

[‡]In contrast, there exist non-affine-invariant LTCs of constant query complexity and inverse polylogarithmic rate. This corresponds to an LTC with $\exp(N/\text{polylog}(N))$ codewords, where N is the code length, while the affine-invariant LTC of [GKS13] and Reed-Muller codes have $\exp(\text{polylog}(N))$ codewords for constant query complexity.

[§]Their lower bound also holds for the weaker notion of locally decodable dode (LDC)

corrector's guess for $f(y)$ must differ from $g(y)$. On the other hand, we can lower bound by a constant the probability of the event that the corrector outputs $g(y)$ when it queries coordinates of f, because f and g are close in the $\|\cdot\|_{U^r}$ norm. This last calculation uses the affine invariance of the code and the *generalized von Neumann inequality*, which bounds by $\|f_0\|_{U^k}$ the expectation over $z_1, \ldots, z_m \in \mathbb{K}^n$ of the product $\prod_{i=0}^{k} f_i(\mathcal{L}_i(z_1, \ldots, z_m))$, where the $\mathcal{L}_i$'s are arbitrary linear forms so that no two are linearly dependent and $f_i : \mathbb{K}^n \to \mathbb{C}$ are arbitrary functions with $|f_i| \leq 1$.

The argument for r-query LTCs is similar. Suppose f and g are close in the $\|\cdot\|_{U^{r-1}}$ norm. Consider the random function H such that for every x independently, $H(x)$ equals $f(x)$ with probability $1/2$ and $g(x)$ with probability $1/2$. H itself is far from a codeword with high probability. But we show that since the local tester accepts f, it will also accept $H \circ \ell$ for a random invertible affine map $\ell : \mathbb{K}^n \to \mathbb{K}^n$ with good probability. This implies that with good probability, $H \circ \ell$ is close to a codeword and by affine-invariance, H itself is close to a codeword which gives a contradiction. To draw this conclusion, we again use the generalized von Neumann inequality as well as a hybrid argument.

Organization. Section 2 contains preliminaries that lay the foundations of our analysis. Section 3 proves the first part of our main result about LCCs, while Section 4 proves the second part about LTCs.

2 Preliminaries

2.1 Error-correcting codes

Let $\mathcal{X}$ be a finite set called the set of coordinates and Σ be an other finite set called the alphabet. Let $\Sigma^{\mathcal{X}}$ denote the set of all functions from $\mathcal{X} \to \Sigma$. A subset $\mathcal{C} \subset \Sigma^{\mathcal{X}}$ is called a code and its elements are called *codewords*.

Definition 1 (Hamming distance)**.** *Given $f, g \in \Sigma^{\mathcal{X}}$, we define the normalized Hamming distance between f and g is defined as $\Delta(f, g) := \Pr_{x \in \mathcal{X}}[f(x) \neq g(x)]$ where x is uniformly chosen from $\mathcal{X}$. For a code $\mathcal{C} \subset \Sigma^{\mathcal{X}}$, we define the minimum distance of $\mathcal{C}$ as $\min_{f,g \in \mathcal{C}, f \neq g} \Delta(f, g)$.*

Let $\blacktriangle_\Sigma = \{q : \Sigma \to \mathbb{R}_{\geq 0} : \sum_{i \in \Sigma} q(i) = 1\}$ denote the probability simplex on Σ. We embed Σ into $\blacktriangle_\Sigma$ by sending $i \in \Sigma$ to e_i which is the i^{th} coordinate vector in $\mathbb{R}^\Sigma$. This also lets us extend functions $f : \mathcal{X} \to \Sigma$ to $\hat{f} : \mathcal{X} \to \blacktriangle_\Sigma$ using the embedding. We call $\hat{f}$ the simplex extension of f. Now given $f, g \in \Sigma^{\mathcal{X}}$, we can write the Hamming distance between them as

$$\Delta(f, g) = 1 - \Pr_{x \in \mathcal{X}}[f(x) = g(x)] = 1 - \mathbb{E}_{x \in \mathcal{X}} \langle \hat{f}, \hat{g} \rangle$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^\Sigma$.

Definition 2 (Affine invariance)**.** *Let $\mathcal{X}$ be a finite dimensional vector space over some finite field $\mathbb{K}$, then $\mathcal{C} \subset \Sigma^{\mathcal{X}}$ is called affine invariant if for every $f \in \mathcal{C}$ and every invertible affine map $\ell : \mathcal{X} \to \mathcal{X}$, $f \circ \ell \in \mathcal{C}$.*

Locally correctable and testable codes are defined formally in Sections 3 and 4 respectively.

2.2 Higher order Fourier analysis

Fix a finite field $\mathbb{F}_p$ of prime order p, and let $\mathbb{K} = \mathbb{F}_q$ where $q = p^t$ for a positive integer t. $\mathbb{K}$ is then a vector space of dimension t over $\mathbb{F}_p$. We denote by $\mathrm{Tr} : \mathbb{K} \to \mathbb{F}_p$ the *trace function*:

$$\mathrm{Tr}(x) = x + x^p + x^{p^2} + \cdots + x^{p^{t-1}}.$$

Also, we use $|\cdot|$ to denote the obvious map from $\mathbb{F}_p$ to $\{0, 1, \ldots, p-1\}$.

Given functions $f, g : \mathbb{K}^n \to \mathbb{C}$, we define their inner product as $\langle f, g \rangle = \mathbb{E}_x[\overline{f(x)} g(x)]$ where x is chosen uniformly from $\mathbb{K}^n$. We define $\|\cdot\|_p$-norm on such functions as $\|f\|_p = \mathbb{E}_x[|f(x)|^p]^{1/p}$. We say a function $f : \mathbb{K}^n \to \mathbb{C}$ is *bounded* if $|f| \leq 1$. Let $\mathbb{T}$ denote the circle group $\mathbb{R}/\mathbb{Z}$ and $e : \mathbb{T} \to \mathbb{C}$ be the map given by $e(x) = \exp(2\pi i x)$.

Definition 3 (Non-classical Polynomials). *A* non-classical polynomial of degree $< d$ *is a function* $f : \mathbb{K}^n \to \mathbb{T}$ *if*

$$\forall h_1, h_2 \cdots, h_d \in \mathbb{K}^n \quad D_{h_1} D_{h_2} \cdots D_{h_d} f = 0$$

where D_h *is the difference operator defined as* $D_h f(x) = f(x+h) - f(x)$. *For such an* f, *the function* $e(f)$ *is called a* non-classical phase polynomial of degree $< d$.

Let $\alpha_1, \cdots, \alpha_t \in \mathbb{K}$ be a basis for $\mathbb{K}$ when viewed as a vector space over $\mathbb{F}_p$. It is known [TZ12, BB15] that non-classical polynomials of degree $\leq d$ are exactly those functions $P : \mathbb{K}^n \to \mathbb{T}$ which have the following form:

$$P(x_1, \ldots, x_n) = \theta + \sum_{k \geq 0} \sum_{\substack{0 \leq d_{i,j} < p \ \forall i \in [n], j \in [t]; \\ 0 < \sum_{i=1}^{n} \sum_{j=1}^{t} d_{i,j} \leq d - k(p-1)}} \frac{c_{d_{1,1},\ldots,d_{n,t},k} \prod_{i=1}^{n} \prod_{j=1}^{t} |\mathrm{Tr}(\alpha_j x_i)|^{d_{i,j}}}{p^{k+1}} \pmod 1 \qquad (1)$$

for some $c_{d_{1,1},\ldots,d_{n,t},k} \in \{0, 1, \cdots, p-1\}$ and $\theta \in \mathbb{T}$. Next, we define the Gowers norm for arbitrary functions $f : \mathbb{K}^n \to \mathbb{C}$.

Definition 4 (Gowers uniformity norm [Gow01]). *For a function* $f : \mathbb{K}^n \to \mathbb{C}$, *the* Gowers norm of order r, *denoted by* $\| \cdot \|_{U^r}$, *is defined as*

$$\|f\|_{U^r} = (\mathbb{E}_{x, h_1, \cdots, h_r \in \mathbb{K}^n} [\Delta_{h_1} \Delta_{h_2} \cdots \Delta_{h_r} f(x)])^{1/2^r}$$

where Δ_h *is the multiplicative difference operator defined as* $\Delta_h f(x) = f(x+h)\overline{f(x)}$.

The Gowers norm is an actual norm when $r \geq 2$. It also satisfies a useful monotonicity property: for any function $f : \mathbb{K}^n \to \mathbb{C}$,

$$|\mathbb{E}[f(x)]| = \|f\|_{U^1} \leq \|f\|_{U^2} \leq \cdots \leq \|f\|_{U^r} \leq \cdots \leq \|f\|_\infty.$$

See [Tao12] for more on Gowers norm. Observe that if $f : \mathbb{K}^n \to \mathbb{C}$ is a non-classical phase polynomial of degree $< r$ then $\|f\|_{U^r} = 1$. The inverse Gowers theorem is a partial converse to this. It shows that the Gowers norm of order r of a function is in direct correspondence with its correlation with non-classical phase polynomials of degree $< r$. In particular:

Lemma 1 (Inverse Gowers theorem [TZ12]). *For any bounded* $f : \mathbb{K}^n \to \mathbb{C}$, *if* $\|f\|_{U^r} > \delta$ *then there exists a non-classical polynomial* P *of degree* $< r$ *such that*

$$|\langle f, e(P) \rangle| \geq c(\delta, \mathbb{K}, r)$$

where $c(\delta, \mathbb{K}, r)$ *is a constant depending only on* $\delta, \mathbb{K}, r$.

A linear form on m variables is a vector $\mathcal{L} = (w_1, \cdots, w_m) \in \mathbb{K}^m$ that is interpreted as a function $\mathcal{L} : (\mathbb{K}^n)^m \to \mathbb{K}^n$ via the map $(x_1, \cdots, x_m) \mapsto \sum_{i=1}^{m} w_i x_i$. A key reason that the Gowers norm is useful in applications is that if a function has small Gowers norm of the appropriate order, then it behaves pseudorandomly in a certain way with respect to linear forms.

Lemma 2 (Generalized von Neumann inequality (Exercise 1.3.23 in [Tao12])). *Let* $f_0, f_1, f_2, \cdots, f_k : \mathbb{K}^n \to \mathbb{C}$ *be bounded functions and let* $\mathcal{L} = \{\mathcal{L}_0, \mathcal{L}_1, \cdots, \mathcal{L}_k\}$ *be a system of* $k+1$ *linear forms in* m *variables such that no form is a multiple of another. Then*

$$|\mathbb{E}_{z_1, \cdots, z_m \in \mathbb{K}^n} [\prod_{i=0}^{k} f_i(\mathcal{L}_i(z_1, \cdots, z_m))]| \leq \min_{0 \leq i \leq k} \|f_i\|_{U^k}$$

See Appendix A for proof.

2.3 A net for Gowers norm

The goal of this section is to establish the following claim.

Theorem 2 (ϵ-net for U^r norm)**.** *The metric induced by the $\|\cdot\|_{U^r}$ norm on the space of all bounded functions $\{f : \mathbb{K}^n \to \mathbb{C}\}$ has an ϵ-net of size $\exp(O_{\epsilon,\mathbb{K},r}(n^{r-1}))$.*

For the proof, we need the following definitions.

Definition 5 (Polynomial factors)**.** *A polynomial factor $\mathcal{B}$ is a sequence of non-classical polynomials $P_1, ..., P_k : \mathbb{K}^n \to \mathbb{T}$. We also identify it with the function $\mathcal{B} : \mathbb{K}^n \to \mathbb{T}^k$ mapping $x \mapsto (P_1(x), ..., P_k(x))$. The partition induced by $\mathcal{B}$ is the partition of $\mathbb{K}^n$ given by $\{\mathcal{B}^{-1}(y) : y \in \mathbb{T}^k\}$. The complexity of $\mathcal{B}$ is the number of defining polynomials, $|\mathcal{B}| = k$. The degree of $\mathcal{B}$ is the maximum degree among its defining polynomials $P_1, \cdots, P_k$. A function $f : \mathbb{K}^n \to \mathbb{C}$ is called $\mathcal{B}$-measurable if it is constant in each cell of the partition induced by $\mathcal{B}$ or equivalently f can be written as a $\tau(P_1, \cdots, P_k)$ for some function $\tau : \mathbb{T}^k \to \mathbb{C}$.*

Definition 6 (Conditional expectations)**.** *Given a polynomial factor $\mathcal{B}$, the conditional expectation of $f : \mathbb{K}^n \to \mathbb{C}$ over $\mathcal{B}$, denoted by $\mathbb{E}[f|\mathcal{B}]$, is the $\mathcal{B}$-measurable function defined by*

$$\mathbb{E}[f|\mathcal{B}](x) = \mathbb{E}_{y \in \mathcal{B}^{-1}(\mathcal{B}(x))}[f(y)].$$

Definition 7 (Factor refinement)**.** *Given two polynomial factors $\mathcal{B}, \mathcal{B}'$, we say $\mathcal{B}'$ is a refinement of $\mathcal{B}$, denoted by $\mathcal{B}' \preceq \mathcal{B}$, if every cell in the partition induced by $\mathcal{B}'$ is contained in some cell in the partition induced by $\mathcal{B}$.*

The definition of refinement immediately implies:

Lemma 3 (Pythagoras theorem)**.** *Let $\mathcal{B}, \mathcal{B}'$ be polynomial factors such that $\mathcal{B}' \preceq \mathcal{B}$, then for any function $f : \mathbb{K}^n \to \mathbb{C}$,*

$$\|\mathbb{E}[f|\mathcal{B}']\|_2^2 = \|\mathbb{E}[f|\mathcal{B}]\|_2^2 + \|\mathbb{E}[f|\mathcal{B}'] - \mathbb{E}[f|\mathcal{B}]\|_2^2.$$

The next claim shows that any bounded function is "close" to being measurable by a polynomial factor of bounded complexity. Precisely:

Lemma 4 (Decomposition Theorem)**.** *Any bounded $f : \mathbb{K}^n \to \mathbb{C}$ can be approximated in $\|\cdot\|_{U^r}$ by a function of a small number of degree $< r$ non-classical polynomials i.e. for any $\epsilon > 0$, there exists non-classical polynomials $P_1, P_2, \cdots, P_k$ of degree $< r$ with $P_i(\bar{0}) = 0\ \forall i$ and a bounded function $\tau : \mathbb{T}^k \to \mathbb{C}$ such that*

$$\|f - \tau(P_1, P_2, \cdots, P_k)\|_{U^r} \leq \epsilon$$

where $k = k(\epsilon, \mathbb{K}, r)$ is a constant depending only on $\epsilon, \mathbb{K}, r$.

Proof. The proof is similar to the proof of the Quadratic Koopman-von Neumann decompostion which is Prop 3.7 in [Gre06] but using the full Inverse Gowers Theorem (Lemma 1) and similar claims are implicit elsewhere, but for completeness, we give the proof.

The main idea is to approximate the function f using its conditional expectation over a suitable polynomial factor $\mathcal{B}$ of degree $< r$. We will start with the trivial factor $\mathcal{B}_0 = (1)$ and iteratively construct more refined partitions $\mathcal{B}_i \preceq \mathcal{B}_{i-1}$ until we find a factor $\mathcal{B}_k$ which satisfies $\|f - \mathbb{E}[f|\mathcal{B}_k]\|_{U^r} \leq \epsilon$. To bound the number of iterations needed to achieve this, we will show that the energy $\|\mathbb{E}[f|\mathcal{B}_i]\|_2^2$ which is bounded above by 1, increases by a fixed constant in every step.

Suppose that after step $i-1$, we still have $\|f - \mathbb{E}[f|\mathcal{B}_{i-1}]\|_{U^r} > \epsilon$. Let $g = f - \mathbb{E}[f|\mathcal{B}_{i-1}]$, then by the inverse Gowers theorem (Lemma 1), we have some non-classical polynomial P_i of degree $< r$ such that $|\langle g, e(P_i)\rangle| \geq \kappa = c(\epsilon, p, r)$. We can assume that $P_i(\bar{0}) = 0$. Refine the factor $\mathcal{B}_{i-1}$ by adding the polynomial P_i to obtain $\mathcal{B}_i \preceq \mathcal{B}_{i-1}$. Now consider the energy increment,

$$\|\mathbb{E}[f|\mathcal{B}_i]\|_2^2 - \|\mathbb{E}[f|\mathcal{B}_{i-1}]\|_2^2 = \|\mathbb{E}[f|\mathcal{B}_i] - \mathbb{E}[f|\mathcal{B}_{i-1}]\|_2^2 = \|\mathbb{E}[g|\mathcal{B}_i]\|_2^2$$

where we used the Pythagoras theorem(Lemma 3) and the fact that $\mathbb{E}\big[\mathbb{E}[f|\mathcal{B}_{i-1}]\big|\mathcal{B}_i\big] = \mathbb{E}[f|\mathcal{B}_{i-1}]$ since $\mathcal{B}_i \preceq \mathcal{B}_{i-1}$. So

$$\begin{aligned}\kappa^2 \le |\mathbb{E}[g\cdot e(P_i)]|^2 &= \big|\mathbb{E}\big[\mathbb{E}[g\cdot e(P_i)|\mathcal{B}_i]\big]\big|^2 = \big|\mathbb{E}\big[e(P_i)\mathbb{E}[g|\mathcal{B}_i]\big]\big|^2\\ &\le \|\mathbb{E}[g|\mathcal{B}_i]\|_1^2 \le \|\mathbb{E}[g|\mathcal{B}_i]\|_2^2 = \|\mathbb{E}[f|\mathcal{B}_i]\|_2^2 - \|\mathbb{E}[f|\mathcal{B}_{i-1}]\|_2^2.\end{aligned}$$

Thus the energy increases by κ^2 every step. But since the energy is bounded above by 1, the process should end in a finite number of steps $k \le \frac{1}{\kappa^2}$. So $\|f - \mathbb{E}[f|\mathcal{B}_k]\|_{U^r} \le \epsilon$, but since $\mathbb{E}[f|\mathcal{B}_k]$ is $\mathcal{B}_k$-measurable, we can write $\mathbb{E}[f|\mathcal{B}_k] = \tau(P_1,\cdots,P_k)$ for some function τ with $|\tau| = |\mathbb{E}[f|\mathcal{B}_k]| \le |f| \le 1$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. Recall that $\mathbb{K}$ is an extension field of dimension t over a prime field $\mathbb{F}_p$. The ϵ-net will be the set $\mathcal{N}$ of all functions of the form $\tau(P_1,\cdots,P_k)$ where $P_1,\cdots,P_k$ are degree $< r$ non-classical polynomials with zero constant terms, $\tau : \mathbb{T}^k \to \mathbb{C}$ is a bounded function and $k = k(\epsilon,p,r)$ is the constant given by Lemma 4. But we will not include all possible bounded $\tau : \mathbb{T}^k \to \mathbb{C}$. Firstly by Equation 1, $P_1,\cdots,P_k$ take values only in $\frac{1}{p^r}\mathbb{Z}/\mathbb{Z}$. Next we will discretize the set $\{z \in \mathbb{C} : |z| \le 1\}$ into the ϵ-lattice i.e. we will only consider maps $\tau : (\frac{1}{p^r}\mathbb{Z}/\mathbb{Z})^k \to \{z \in \mathbb{C} : |z| \le 1\} \cap \epsilon(\mathbb{Z} + i\mathbb{Z})$. The number of such maps is bounded by $(4/\epsilon^2)^{p^{rk}}$.

By Equation 1, a non-classical polynomial of degree $< r$ in n variables with zero constant term can be represented by $\le \binom{nt+r-1}{r-1} r$ coefficients in $\{0,1,\cdots,p-1\}$. So the number of such non-classical polynomials is bounded by $\exp\big(O_{r,\mathbb{K}}(n^{r-1})\big)$. Combining both the bounds,

$$|\mathcal{N}| \le \exp\big(O_{r,\mathbb{K}}(n^{r-1})\big)^k \cdot (4/\epsilon^2)^{p^{rk}} = \exp\big(O_{\epsilon,\mathbb{K},r}(n^{r-1})\big).$$

We will now prove that $\mathcal{N}$ is a 3ϵ-net. Given any $f : \mathbb{K}^n \to [-1,1]$, using Lemma 4, there is a function $\tau(P_1,\cdots,P_k)$ such that

$$\|f - \tau(P_1,P_2,\cdots,P_k)\|_{U^r} \le \epsilon.$$

If we consider the $\tilde{\tau} \in \mathcal{N}$ by rounding values real and imaginary parts of τ to the nearest multiple of ϵ, we get

$$\begin{aligned}\|f - \tilde{\tau}(P_1,P_2,\cdots,P_k)\|_{U^r} &\le \|f - \tau(P_1,P_2,\cdots,P_k)\|_{U^r} + \|\tau(P_1,P_2,\cdots,P_k) - \tilde{\tau}(P_1,P_2,\cdots,P_k)\|_{U^r}\\ &\le \epsilon + \|\tau(P_1,P_2,\cdots,P_k) - \tilde{\tau}(P_1,P_2,\cdots,P_k)\|_\infty \le 3\epsilon.\end{aligned}$$

□

3 Locally Correctable Codes

We begin by defining locally correctable codes formally. Note that the definition below differs from the conventional one in terms of a local correction algorithm and adversarial errors (see, for instance, [Yek11]); however, our definition is certainly weaker. Therefore, this makes our lower bounds stronger.

Definition 8 (Locally Correctable Code (LCC))**.** *An (r,δ,τ) LCC is a code $\mathcal{C} \subset \Sigma^{\mathcal{X}}$ with the following property:*
For each $x \in \mathcal{X}$ there is a distribution $\mathcal{M}_x$ over r-tuples of distinct[¶] coordinates such that whenever $\tilde{f} \in \Sigma^{\mathcal{X}}$ is δ-close to some codeword $f \in \mathcal{C}$ in Hamming distance,

$$\Pr_{(y_1,\cdots,y_r)\sim\mathcal{M}_x}[\mathcal{D}_{x,y_1,\cdots,y_r}(\tilde{f}(y_1),\tilde{f}(y_2),\cdots,\tilde{f}(y_r)) = f(x)] \ge 1-\tau$$

where $\mathcal{D}_{x,y_1,\cdots,y_r} : \Sigma^r \to \Sigma$, called the decoding operator, depends only on $x, y_1,\cdots,y_r$.
If furthermore $\mathcal{X}$ is a vector space and $\mathcal{C}$ is affine invariant then we call it an affine invariant LCC.

[¶]WLOG we can assume the tuples have distinct coordinates by adding dummy coordinates and modifying the decoding functions $\mathcal{D}_{x,y_1,\cdots,y_r}$

Remark 1. *Let $|\Sigma| = m$, WLOG we can assume that $\Sigma = \{1, 2, \cdots, m\}$. Then we can extend functions $f : \mathcal{X} \to \Sigma$ to $\hat{f} : \mathcal{X} \to \blacktriangle_m$. The decoding operators $\mathcal{D} : \Sigma^r \to \Sigma$ can also be extended to $\widehat{\mathcal{D}} : \blacktriangle_m^r \to \blacktriangle_m$ as follows: For $z_1, \cdots, z_r \in \blacktriangle_m$ define*

$$\widehat{\mathcal{D}}(z_1, \cdots, z_r) = \sum_{1 \le \ell_1, \cdots, \ell_r \le m} e_{\mathcal{D}(\ell_1, \cdots, \ell_r)} (z_1)_{\ell_1} \cdots (z_r)_{\ell_r}$$

where e_j stands for the j^{th} coordinate vector in $\mathbb{R}^m$ and $(z_j)_\ell$ is the ℓ^{th} coordinate of the vector z_j. Now we can rewrite the decoding condition as:

$$\mathbb{E}_{(y_1, \cdots, y_r) \sim \mathcal{M}_x} [\left\langle \hat{f}(x), \widehat{\mathcal{D}}_{x, y_1, \cdots, y_r}(\hat{f}(y_1), \hat{f}(y_2), \cdots, \hat{f}(y_r)) \right\rangle] \ge 1 - \tau.$$

First, we make the observation that any LCC must have good minimum distance.

Lemma 5. *Let $\mathcal{C} \subset \Sigma^{\mathcal{X}}$ be an (r, δ, τ) LCC with $\tau < 1/2$, then the minimum distance of $\mathcal{C}$ is at least 2δ.*

Proof. Let $f, g \in \mathcal{C}$ be two distinct codewords such that $\Delta(f, g) < 2\delta$. Let h be the midpoint of f and g i.e. h is δ-close to both f and g. Let $x \in \mathcal{X}$ be such that $f(x) \neq g(x)$. By the LCC property,

$$\Pr_{(y_1, \cdots, y_r) \sim \mathcal{M}_x} [f(x) = \mathcal{D}_{x, y_1, \cdots, y_r}(h(y_1), \cdots, h(y_r))] \ge 1 - \tau$$

$$\Pr_{(y_1, \cdots, y_r) \sim \mathcal{M}_x} [g(x) = \mathcal{D}_{x, y_1, \cdots, y_r}(h(y_1), \cdots, h(y_r))] \ge 1 - \tau.$$

This is a contradiction when $\tau < \frac{1}{2}$. Therefore every two codewords must be at least 2δ apart. □

Now, we are ready to prove our main result of this section.

Theorem 3 (Lower bound for LCCs)**.** *Let $\mathcal{C} \subset \Sigma^{\mathbb{K}^n}$ be an (r, δ, τ) affine-invariant LCC where $\tau < \frac{2\delta}{3}$. Then $|\mathcal{C}| \le \exp\left(O_{\delta, \mathbb{K}, r, |\Sigma|}(n^{r-1})\right)$.*

Proof. Let $|\Sigma| = m$. Let $\mathcal{N}$ be an $\epsilon/2$-net for the space of all bounded functions $\{f : \mathbb{K}^n \to \mathbb{C}\}$ with the metric induced by $\|\cdot\|_{U^r}$-norm where $\epsilon = \frac{2\delta}{3m^r}$. Given a bounded $f : \mathbb{K}^n \to \mathbb{C}$, define

$$\phi(f) := \text{argmin}_{h \in \mathcal{N}} \|f - h\|_{U^r}$$

(break ties arbitrarily). Since $\mathcal{N}$ is an $\epsilon/2$ net, we have $\|f - \phi(f)\|_{U^r} \le \epsilon/2$. Define $\Psi : \mathcal{C} \to \mathcal{N}^m$ as

$$\Psi(f) := (\phi(\hat{f}_1), \cdots, \phi(\hat{f}_m))$$

where $\hat{f}_i : \mathbb{K}^n \to \mathbb{R}_{\ge 0}$ is the i^{th} coordinate function of the simplex extension $\hat{f} : \mathbb{K}^n \to \blacktriangle_m$ of f. We claim that Ψ is one-one which implies that $|\mathcal{C}| \le |\mathcal{N}|^m$. Now using Theorem 2, the required bound follows. Suppose that Ψ is not one-one. Let $f, g \in \mathcal{C}$ be two distinct codewords such that $\Psi(f) = \Psi(g)$. This implies that

$$\forall\ i \in [m]\ \|\hat{f}_i - \hat{g}_i\|_{U^r} \le \|\hat{f}_i - \phi(\hat{f}_i)\|_{U^r} + \|\hat{g}_i - \phi(\hat{g}_i)\|_{U^r} \le \epsilon.$$

By affine invariance of $\mathcal{C}$, $f \circ \ell \in \mathcal{C}$ for all invertible affine maps $\ell : \mathbb{K}^n \to \mathbb{K}^n$. So by the local correction property,

$$\Pr_{\ell, y_0, (y_1, \cdots, y_r) \sim \mathcal{M}_{y_0}} [f \circ \ell(y_0) = \mathcal{D}_{y_0, y_1, \cdots, y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] \ge 1 - \tau$$

where ℓ ranges uniformly over all invertible affine maps from $\mathbb{K}^n \to \mathbb{K}^n$ and y_0 ranges uniformly over $\mathbb{K}^n$. Now consider the following difference:

$$\begin{aligned}
&\Pr_{\ell, y_0, (y_1, \cdots, y_r) \sim \mathcal{M}_{y_0}} [f \circ \ell(y_0) = \mathcal{D}_{y_0, y_1, \cdots, y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] \\
&\qquad - \Pr_{\ell, y_0, (y_1, \cdots, y_r) \sim \mathcal{M}_{y_0}} [g \circ \ell(y_0) = \mathcal{D}_{y_1, \cdots, y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] \\
&= \mathbb{E}_\ell \mathbb{E}_{y_0} \mathbb{E}_{(y_1, \cdots, y_r) \sim \mathcal{M}_{y_0}} \Big[\left\langle \hat{f} \circ \ell(y_0), \widehat{\mathcal{D}}_{y_0, y_1, \cdots, y_r}(\hat{f} \circ \ell(y_1), \cdots, \hat{f} \circ \ell(y_r)) \right\rangle \\
&\qquad\qquad - \left\langle \hat{g} \circ \ell(y_0), \widehat{\mathcal{D}}_{y_1, \cdots, y_r}(\hat{f} \circ \ell(y_1), \cdots, \hat{f} \circ \ell(y_r)) \right\rangle \Big] \\
&= \mathbb{E}_{y_0} \mathbb{E}_{(y_1, \cdots, y_r) \sim \mathcal{M}_{y_0}} \left[\mathbb{E}_\ell \left[\left\langle \hat{f} \circ \ell(y_0) - \hat{g} \circ \ell(y_0), \widehat{\mathcal{D}}_{y_0, y_1, \cdots, y_r}(\hat{f} \circ \ell(y_1), \cdots, \hat{f} \circ \ell(y_r)) \right\rangle \right] \right].
\end{aligned}$$

Now we fix $y_0, y_1, \cdots, y_r$ and show that inner expectation is small for each tuple $(y_0, y_1, \cdots, y_r)$. Let us denote $\mathcal{D} = \mathcal{D}_{y_0,y_1,\cdots,y_r}$ for brevity. Let $t = \text{rank}(y_0, y_1, \cdots, y_r)$, thus there exist independent vectors $v_1, \cdots, v_t \in \mathbb{K}^n$ such that for every $0 \le i \le r$, $y_i = \sum_{j=1}^{t} \lambda_{ij} v_j$ for some fixed $\lambda_{ij} \in \mathbb{K}$. The action of a random invertible affine map ℓ can be approximated by sampling $z_0, z_1, \cdots, z_t \in \mathbb{K}^n$ uniformly and mapping $y_i \mapsto z_0 + \sum_{j=1}^{t} \lambda_{ij} z_j$ since with probability $1 - o_n(1)$, $z_1, \cdots, z_t$ will be independent. Therefore,

$$
\begin{aligned}
&\mathbb{E}_\ell \left[\left\langle \hat{f} \circ \ell(y_0) - \hat{g} \circ \ell(y_0), \widehat{\mathcal{D}}_{y_0,y_1,\cdots,y_r}(\hat{f} \circ \ell(y_1), \cdots, \hat{f} \circ \ell(y_r)) \right\rangle \right] \\
&=_{o_n(1)} \mathbb{E}_{z_0,z_1,\cdots,z_t \in \mathbb{K}^n} \left[\left\langle (\hat{f} - \hat{g})(z_0 + \sum_{j=1}^{t} \lambda_{0j} z_j), \widehat{\mathcal{D}} \left(\hat{f}(z_0 + \sum_{j=1}^{t} \lambda_{1j} z_j), \cdots, \hat{f}(z_0 + \sum_{j=1}^{t} \lambda_{rj} z_j) \right) \right\rangle \right] \\
&= \mathbb{E}_{z_0,z_1,\cdots,z_t \in \mathbb{K}^n} \left[\left\langle (\hat{f} - \hat{g})(z_0 + \sum_{j=1}^{t} \lambda_{0j} z_j), \left(\sum_{1 \le \ell_1,\cdots,\ell_r \le m} e_{\mathcal{D}(\ell_1,\cdots,\ell_r)} \prod_{i=1}^{r} \hat{f}_{\ell_i}(z_0 + \sum_{j=1}^{t} \lambda_{ij} z_j) \right) \right\rangle \right] \\
&= \mathbb{E}_{z_0,z_1,\cdots,z_t \in \mathbb{K}^n} \left[\left(\sum_{1 \le \ell_1,\cdots,\ell_r \le m} (\hat{f} - \hat{g})_{\mathcal{D}(\ell_1,\cdots,\ell_r)}(z_0 + \sum_{j=1}^{t} \lambda_{0j} z_j) \cdot \prod_{i=1}^{r} \hat{f}_{\ell_i}(z_0 + \sum_{j=1}^{t} \lambda_{ij} z_j) \right) \right] \\
&\le \left(\sum_{0 \le \ell_1,\cdots,\ell_r \le m-1} \|(\hat{f} - \hat{g})_{\mathcal{D}(\ell_1,\cdots,\ell_r)}\|_{U^r} \right) \le m^r \epsilon
\end{aligned}
$$

where the first inequality is obtained by applying generalized von Neumann inequality (Lemma 2) to each term. Therefore

$$
\begin{aligned}
&\Pr_{\ell,y_0,(y_1,\cdots,y_r) \sim \mathcal{M}_{y_0}} [g \circ \ell(y_0) = \mathcal{D}_{y_1,\cdots,y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] \\
&\ge \Pr_{\ell,y_0,(y_1,\cdots,y_r) \sim \mathcal{M}_{y_0}} [f \circ \ell(y_0) = \mathcal{D}_{y_1,\cdots,y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] - m^r \epsilon \ge 1 - \tau - 2\delta/3.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\Pr_{\ell,y_0,(y_1,\cdots,y_r) \sim \mathcal{M}_{y_0}} [g \circ \ell(y_0) = \mathcal{D}_{y_1,\cdots,y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] \\
&\le \Pr_{\ell,y_0,(y_1,\cdots,y_r) \sim \mathcal{M}_{y_0}} [g \circ \ell(y_0) = f \circ \ell(y_0)] + \Pr_{\ell,y_0,(y_1,\cdots,y_r) \sim \mathcal{M}_{y_0}} [f \circ \ell(y_0) \ne \mathcal{D}_{y_1,\cdots,y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] \\
&\le \Pr_x[f(x) = g(x)] + \tau \le 1 - 2\delta + \tau \qquad \text{(By Lemma 5)}
\end{aligned}
$$

This is a contradiction when $\tau < \frac{2\delta}{3}$.

□

4 Locally Testable Codes

We start by defining locally testable codes in a formulation convenient for our use.

Definition 9 (Locally Testable Code (LTC))**.** *An (r, δ, τ) LTC is a code $\mathcal{C} \subset \Sigma^{\mathcal{X}}$ with minimum distance at least δ and the following property:*
There is a distribution $\mathcal{M}$ over r-tuples of distinct[‖] coordinates such that for each codeword $f \in C$,

$$
\Pr_{(y_1,\cdots,y_r) \sim \mathcal{M}} [\mathcal{D}_{y_1,\cdots,y_r}(f(y_1), f(y_2), \cdots, f(y_r)) = 1] \ge 3/4
$$

[‖]WLOG we can assume the tuples have distinct coordinates by adding dummy coordinates and modifying the decoding functions $\mathcal{D}_{y_1,\cdots,y_r}$

and for every $g \in \Sigma^{\mathcal{X}}$ which is τ-far away from every codeword,

$$\Pr_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(g(y_1), f(y_2),\cdots,f(y_r)) = 1] \leq 1/4$$

where $\mathcal{D}_{y_1,\cdots,y_r} : \Sigma^r \to \{0,1\}$, called the testing operator, depends only on $y_1,\cdots,y_r$.
If furthermore $\mathcal{X}$ is a vector space and $\mathcal{C}$ is affine-invariant then we call it an affine invariant LTC.

Remark 2. *Let $|\Sigma| = m$, WLOG we can assume that $\Sigma = \{1,2,\cdots,m\}$. We can extend $f : \mathcal{X} \to \Sigma$ to $\hat{f} : \mathcal{X} \to \blacktriangle_m$. The testing operator $\mathcal{D} : \Sigma^r \to \{0,1\}$ can also be extended to $\widehat{\mathcal{D}} : \blacktriangle_m^r \to [0,1]$ as follows: For $z_1,\cdots,z_r \in \blacktriangle_m$ define*

$$\widehat{\mathcal{D}}(z_1,\cdots,z_r) = \sum_{1\leq\ell_1,\cdots,\ell_r\leq m} \mathcal{D}(\ell_1,\cdots,\ell_r)(z_1)_{\ell_1}\cdots(z_r)_{\ell_r}. \tag{2}$$

Now we can rewrite the probability in terms of expectation as:

$$\Pr_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(f(y_1),\cdots,f(y_r)) = 1] = \mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{f}\circ\ell(y_1),\cdots,\hat{f}\circ\ell(y_r))]$$

We are now ready to prove the main result of this section.

Theorem 4 (Lower bound for LTC's)**.** *Let $\mathcal{C} \subset \Sigma^{\mathbb{K}^n}$ be an $(r,\delta,\delta/3)$ affine invariant LTC, then $|\mathcal{C}| \leq \exp\left(O_{\delta,\mathbb{K},r,|\Sigma|}(n^{r-2})\right)$.*

Proof. Let $|\Sigma| = m$. The proof is very similar to that of Theorem 3. Let $\mathcal{N}$ be an $\epsilon/2$-net for the space of all bounded functions $\{f : \mathbb{K}^n \to \mathbb{C}\}$ with the metric induced by $\|\cdot\|_{U^{r-1}}$-norm where $\epsilon = 1/2rm^r$. Define $\Psi : \mathcal{C} \to \mathcal{N}^m$ as in the proof of Theorem 3, it is enough to show that Ψ is one-one. Suppose that Ψ is not one-one. Then there exists $f,g \in \mathcal{C}$ which are distinct such that $\Psi(f) = \Psi(g)$. This implies that

$$\forall\ i \in [m]\ \|\hat{f}_i - \hat{g}_i\|_{U^{r-1}} \leq \epsilon.$$

By affine invariance of $\mathcal{C}$, $f\circ\ell \in \mathcal{C}$ for all invertible affine maps $\ell : \mathbb{K}^n \to \mathbb{K}^n$. So

$$\mathbb{E}_\ell\mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(f\circ\ell(y_1), f\circ\ell(y_2),\cdots,f\circ\ell(y_r))] \geq 3/4$$

where ℓ ranges over all invertible affine maps from $\mathbb{K}^n \to \mathbb{K}^n$. Let $H \in \Sigma^{\mathcal{X}}$ be a random word where for each coordinate $x \in \mathcal{X}$ independently,

$$H(x) = \begin{cases} f(x) & w.p.\ 1/2 \\ g(x) & w.p.\ 1/2 \end{cases}$$

Define $\hat{h} : \mathcal{X} \to \blacktriangle_m$ as $\hat{h}(x) = \mathbb{E}_H[\widehat{H}(x)] = \frac{\hat{f}(x)+\hat{g}(x)}{2}$ where $\hat{f},\hat{g}$ are the simplex extensions of the original f,g. So $\forall\ i \in [m]\ \|\hat{f}_i - \hat{h}_i\|_{U^{r-1}} = \|\hat{f}_i - \hat{g}_i\|_{U^{r-1}}/2 \leq \epsilon/2$. We will now show that the test accepts $H\circ\ell$ with good probability when ℓ is a random invertible affine map from $\mathbb{K}^n \to \mathbb{K}^n$.

$$\begin{aligned}
&\mathbb{E}_H\mathbb{E}_\ell\mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(f\circ\ell(y_1),\cdots,f\circ\ell(y_r)) - \mathcal{D}_{y_1,\cdots,y_r}(H\circ\ell(y_1),\cdots,H\circ\ell(y_r))] \\
&= \mathbb{E}_H\mathbb{E}_\ell\mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{f}\circ\ell(y_1),\cdots,\hat{f}\circ\ell(y_r)) - \widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\widehat{H}\circ\ell(y_1),\cdots,\widehat{H}\circ\ell(y_r))] \\
&= \mathbb{E}_\ell\mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{f}\circ\ell(y_1),\cdots,\hat{f}\circ\ell(y_r)) - \widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{h}\circ\ell(y_1),\cdots,\hat{h}\circ\ell(y_r))] \\
&\qquad \text{(by using the multilinear expansion of } \widehat{\mathcal{D}}_{y_1,\cdots,y_r} \text{(Equation 2) and taking expectation over } H) \\
&= \mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}\left[\mathbb{E}_\ell\left[\widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{f}\circ\ell(y_1),\cdots,\hat{f}\circ\ell(y_r)) - \widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{h}\circ\ell(y_1),\cdots,\hat{h}\circ\ell(y_r))\right]\right]
\end{aligned}$$

Now we fix $y_1,\cdots,y_r$ and show that inner expectation is small for each tuple $(y_1,\cdots,y_r)$. Let us denote $\mathcal{D} = \mathcal{D}_{y_1,\cdots,y_r}$ for brevity. Let $t = \text{rank}(y_1,\cdots,y_r)$, thus there exist independent vectors $v_1,\cdots,v_t \in \mathbb{K}^n$ such that for every $1 \leq i \leq r$, $y_i = \sum_{j=1}^t \lambda_{ij}v_j$ for some fixed $\lambda_{ij} \in \mathbb{K}$. The action of a random invertible affine

map ℓ can be approximated by sampling $z_0, z_1, \cdots, z_t \in \mathbb{K}^n$ uniformly and mapping $y_i \mapsto z_0 + \sum_{j=1}^{t} \lambda_{ij} z_j$ since with probability $1 - o_n(1)$, $z_1, \cdots, z_t$ will be independent. Therefore,

$$\begin{aligned}
&\mathbb{E}_\ell \left[\widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{f} \circ \ell(y_1), \cdots, \hat{f} \circ \ell(y_r)) - \widehat{\mathcal{D}}_{y_1,\cdots,y_r}(\hat{h} \circ \ell(y_1), \cdots, \hat{h} \circ \ell(y_r)) \right] \\
&=_{o_n(1)} \mathbb{E}_{z_0,\cdots,z_t \in \mathbb{K}^n} \left[\widehat{\mathcal{D}}(\hat{f}(z_0 + \sum_{j=1}^{t} \lambda_{1j} z_j), \cdots, \hat{f}(z_0 + \sum_{j=1}^{t} \lambda_{rj} z_j)) - \mathcal{D}(\hat{h}(z_0 + \sum_{j=1}^{t} \lambda_{1j} z_j), \cdots, \hat{h}(z_0 + \sum_{j=1}^{t} \lambda_{rj} z_j)) \right] \\
&= \mathbb{E}_{z_0,z_1,\cdots,z_t \in \mathbb{K}^n} \left[\sum_{1 \le \ell_1,\cdots,\ell_r \le m} \mathcal{D}(\ell_1, \cdots, \ell_r) \left(\prod_{i=1}^{r} \hat{f}_{\ell_i}(z_0 + \sum_{j=1}^{t} \lambda_{ij} z_j) - \prod_{i=1}^{r} \hat{h}_{\ell_i}(z_0 + \sum_{j=1}^{t} \lambda_{ij} z_j) \right) \right] \\
&\le r \cdot m^r \cdot \frac{\epsilon}{2} = \frac{1}{4}
\end{aligned}$$

where the last line is obtained by forming hybrids i.e. writing

$$\hat{f}_{\ell_1} \cdot \hat{f}_{\ell_2} \cdots \hat{f}_{\ell_r} - \hat{h}_{\ell_1} \cdot \hat{h}_{\ell_2} \cdots \hat{h}_{\ell_r} = (\hat{f}_{\ell_1} - \hat{h}_{\ell_1}) \cdot \hat{f}_{\ell_2} \cdots \hat{f}_{\ell_r} + \hat{h}_{\ell_1} \cdot (\hat{f}_{\ell_2} - \hat{h}_{\ell_2}) \cdots \hat{f}_{\ell_r} + \cdots + \hat{h}_{\ell_1} \cdot \hat{h}_{\ell_2} \cdots (\hat{f}_{\ell_r} - \hat{h}_{\ell_r})$$

and using Lemma 2 for each term. Therefore

$$\begin{aligned}
&\mathbb{E}_H \mathbb{E}_\ell \mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(H \circ \ell(y_1), \cdots, H \circ \ell(y_r))] \\
&\ge \mathbb{E}_\ell \mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(f \circ \ell(y_1), \cdots, f \circ \ell(y_r))] - \frac{1}{4} \ge \frac{3}{4} - \frac{1}{4} = \frac{1}{2}.
\end{aligned}$$

By Markov inequality,

$$\begin{aligned}
\frac{1}{4} &\le \Pr_H \left[\mathbb{E}_\ell \mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(H \circ \ell(y_1), \cdots, H \circ \ell(y_r))] \ge \frac{1}{3} \right] \\
&\le \Pr_H \left[\exists \ell \;\; \mathbb{E}_{(y_1,\cdots,y_r)\sim\mathcal{M}}[\mathcal{D}_{y_1,\cdots,y_r}(H \circ \ell(y_1), \cdots, H \circ \ell(y_r))] \ge \frac{1}{3} \right] \\
&\le \Pr_H \left[\exists \ell \; \Delta(H \circ \ell, \mathcal{C})] \le \frac{\delta}{3} \right] && \text{(by the soundness of the tester)} \\
&= \Pr_H \left[\Delta(H, \mathcal{C})] \le \frac{\delta}{3} \right] && \text{(since } \ell \text{ is invertible and } \mathcal{C} \text{ is affine invariant)}
\end{aligned}$$

Let $\mathcal{H} = \mathrm{Supp}(H)$ be the set of words between f and g i.e. the set of all words $e \in \Sigma^{\mathbb{K}^n}$ such that $e(x) = f(x)$ or $e(x) = g(x)$ for all $x \in \mathbb{K}^n$. We have $|\mathcal{H}| = 2^{\Delta(f,g)n}$. Since the distribution of H is uniform in $\mathcal{H}$, we proved that at least $\frac{1}{4}$ fraction of words in $\mathcal{H}$ contain a codeword in their $\delta/3$ neighborhood, let $\mathcal{H}' \subset \mathcal{H}$ denote this subset. Therefore the $\delta/6$ neighborhoods around the points in $\mathcal{H}'$ must be disjoint or else two distinct codewords will be $< \delta$ close to each other. The number of words in $\mathcal{H}$ which lie in a Hamming ball of radius $\delta/6$ around a point of $\mathcal{H}'$ is

$$\sum_{i=0}^{\delta n/6} \binom{\Delta(f,g)n}{i} \ge 2^{H(\delta/6\Delta(f,g))\Delta(f,g)n - o(n)} \ge 2^{H(\delta/6)\Delta(f,g)n - o(n)}$$

where $H(\cdot)$ is the binary entropy function. By a packing argument, we can upper bound the size of $\mathcal{H}'$ as

$$|\mathcal{H}'| \le \frac{2^{\Delta(f,g)n}}{2^{H(\delta/6)\Delta(f,g)n - o(n)}} = o(|\mathcal{H}|).$$

This contradicts the fact that $|\mathcal{H}'| \ge |\mathcal{H}|/4$.

□

5 Concluding Remarks

In this work, we proved tight lower bounds for constant query affine-invariant LCCs and LTCs when the number of queries r, underlying field $\mathbb{K}$ and the alphabet Σ are constant. However the constants in the bounds we obtain are of Ackermann-type in $r, |\mathbb{K}|, |\Sigma|$ because of the use of higher-order Fourier analysis. Improving the dependence on these parameters is an open problem which might require new ideas. In a recent work, Bhowmick and Lovett [BL15a] obtain a "bias implies low rank" theorem for polynomials over growing fields. This might be a first step towards proving a variant of the inverse Gowers theorem (Lemma 1) for growing field size, which could then be used to make our lower bounds extend to the case of growing field size.

We also remark that our lower bounds work for any LCC or LTC where the queries are obtained as fixed linear combinations of uniformly chosen points from $\mathbb{K}^n$. Affine-invariant codes are a natural class of local codes where this is true. Relaxing these conditions to get lower bounds for a more general class of LCCs or LTCs is an open problem.

Acknowledgements

We thank Madhu Sudan for helpful pointers to previous work. The second author would like to thank his advisor, Zeev Dvir, for his guidance and encouragement.

References

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998. 1

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM (JACM)*, 45(1):70–122, 1998. 1

[BB15] Arnab Bhattacharyya and Abhishek Bhowmick. Using higher-order Fourier analysis over general fields. *arXiv preprint arXiv:1505.00619*, 2015. 4

[BDSS11] Arnab Bhattacharyya, Zeev Dvir, Amir Shpilka, and Shubhangi Saraf. Tight lower bounds for 2-query LCCs over finite fields. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 638–647. IEEE, 2011. 2

[BDYW11] Boaz Barak, Zeev Dvir, Amir Yehudayoff, and Avi Wigderson. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 519–528. ACM, 2011. 2

[BIW07] Omer Barkol, Yuval Ishai, and Enav Weinreb. On locally decodable codes, self-correctable codes, and t-private PIR. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 311–325. Springer, 2007. 1

[BK95] Manuel Blum and Sampath Kannan. Designing programs that check their work. *Journal of the ACM (JACM)*, 42(1):269–291, 1995. 1

[BL15a] Abhishek Bhowmick and Shachar Lovett. Bias vs structure of polynomials in large fields, and applications in effective algebraic geometry and coding theory. *CoRR*, abs/1506.02047, 2015. 11

[BL15b] Abhishek Bhowmick and Shachar Lovett. The list decoding radius of Reed-Muller codes over small fields. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 277–285, 2015. 2

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47(3):549–595, 1993. 1

[BRS12] Eli Ben-Sasson, Noga Ron-Zewi, and Madhu Sudan. Sparse affine-invariant linear codes are locally testable. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 561–570. IEEE, 2012. 1

[BS04] Eli Ben-Sasson and Madhu Sudan. Robust locally testable codes and products of codes. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 286–297, 2004. 1

[BS08] Eli Ben-Sasson and Madhu Sudan. Short PCPs with polylog query complexity. *SIAM Journal on Computing*, 38(2):551–607, 2008. 1

[BS11] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 412–423. Springer, 2011. 1, 2

[CKGS98] Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *Journal of the ACM (JACM)*, 45(6):965–981, 1998. 1

[Din07] Irit Dinur. The PCP theorem by gap amplification. *Journal of the ACM (JACM)*, 54(3):12, 2007. 1

[DS07] Zeev Dvir and Amir Shpilka. Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits. *SIAM Journal on Computing*, 36(5):1404–1434, 2007. 2

[DSW14] Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Breaking the quadratic barrier for 3-LCC's over the reals. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 784–793. ACM, 2014. 2

[GKS13] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 529–540. ACM, 2013. 1, 2

[GKST02] Oded Goldreich, Howard Karloff, Leonard J Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. In *Computational Complexity, 2002. Proceedings. 17th IEEE Annual Conference on*, pages 143–151. IEEE, 2002. 2

[Gow01] William T Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11(3):465–588, 2001. 4

[Gre06] Ben Green. Montreal lecture notes on quadratic Fourier analysis. *arXiv preprint math/0604089*, 2006. 5

[GS06] Oded Goldreich and Madhu Sudan. Locally testable codes and PCPs of almost-linear length. *Journal of the ACM*, 53(4):558 – 655, July 2006. 1

[GSVW15] Venkatesan Guruswami, Madhu Sudan, Ameya Velingker, and Carol Wang. Limitations on testable affine-invariant codes in the high-rate regime. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1312–1325. SIAM, 2015. 1

[Guo13] Alan Xinyu Guo. *Some closure features of locally testable affine-invariant properties*. PhD thesis, Massachusetts Institute of Technology, 2013. 1

[KdW03] Iordanis Kerenidis and Ronald de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 106–115. ACM, 2003. 2

[KLP67] T. Kasami, S. Lin, and W.W. Peterson. Some results on cyclic codes which are invariant under the affine group and their applications. *Information and Control*, 11(5–6):475–496, 1967. 1

[KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 403–412. ACM, 2008. 1

[KT00] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 80–86. ACM, 2000. 1, 2

[Lip90] Richard J Lipton. Efficient checking of computations. In *STACS 90*, pages 207–215. Springer, 1990. 1

[Mei09] Or Meir. Combinatorial construction of locally testable codes. *SIAM J. Comput.*, 39(2):491–544, 2009. 1

[STV99] Madhu Sudan, Luca Trevisan, and Salil Vadhan. Pseudorandom generators without the XOR lemma. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 537–546. ACM, 1999. 1

[Tao12] Terence Tao. *Higher order Fourier analysis*, volume 142. American Mathematical Soc., 2012. 2, 4, 13

[TW14] Madhur Tulsiani and Julia Wolf. Quadratic Goldreich-Levin theorems. *SIAM Journal on Computing*, 43(2):730–766, 2014. 2

[TZ12] Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Annals of Combinatorics*, 16(1):121–188, 2012. 2, 4

[Vid15] Michael Viderman. Explicit strong LTCs with inverse poly-log rate and constant soundness. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:20, 2015. 1

[Woo07] David Woodruff. New lower bounds for general locally decodable codes. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 14, 2007. 2

[Woo12] David P Woodruff. A quadratic lower bound for three-query linear locally decodable codes over any field. *Journal of Computer Science and Technology*, 27(4):678–686, 2012. 2

[Yek11] Sergey Yekhanin. Locally decodable codes. In *Computer Science–Theory and Applications*, pages 289–290. Springer, 2011. 1, 6

A Proof of generalized von Neumann inequality (Lemma 2)

Since the lemma is not stated in the form we want in [Tao12], we will include a proof here for completeness. To prove Lemma 2, we need the following lemma first.

Lemma 6 (Exercise 1.3.22 in [Tao12])**.** *Let $f : \mathbb{K}^n \to \mathbb{C}$ be a function, and for each $1 \leq i \leq k$, let $g_i : (\mathbb{K}^n)^k \to \mathbb{C}$ be a bounded function which is independent of the i^{th} coordinate of $(\mathbb{K}^n)^k$. Then,*

$$|\mathbb{E}_{x_1,\cdots,x_k \in \mathbb{K}^n}[f(x_1 + x_2 + \cdots + x_k) \prod_{i=1}^{k} g_i(x_1, \cdots, x_k)]| \leq \|f\|_{U^k}$$

Proof. The proof is by induction on k and using Cauchy-Schwarz inequality repeatedly. The case $k = 1$ is true by definition of $\|\cdot\|_{U^1}$.

$$\begin{aligned}
&\left|\mathbb{E}_{x_1,\cdots,x_k\in\mathbb{K}^n}\left[f(x_1+x_2+\cdots+x_k)\prod_{i=1}^{k}g_i(x_1,\cdots,x_k)\right]\right| \\
&= \left|\mathbb{E}_{x_2,\cdots,x_k}\left[g_1(x_1,\cdots,x_k)\mathbb{E}_{x_1}\left[f(x_1+x_2+\cdots+x_k)\prod_{i=2}^{k}g_i(x_1,\cdots,x_k)\right]\right]\right| \\
&\qquad\qquad \text{(since } g_1 \text{ doesn't depend on } x_1\text{)} \\
&\leq \left|\mathbb{E}_{x_2,\cdots,x_k}\left[\mathbb{E}_{x_1'}\left[f(x_1'+x_2+\cdots+x_k)\prod_{i=2}^{k}g_i(x_1',x_2,\cdots,x_k)\right]\mathbb{E}_{x_1}\left[\bar{f}(x_1+x_2+\cdots+x_k)\prod_{i=2}^{k}\bar{g}_i(x_1,x_2,\cdots,x_k)\right]\right]\right|^{1/2} \\
&\qquad\qquad \text{(By Cauchy-Schwarz inequality and the fact that } |g_1|\leq 1\text{)} \\
&= \left|\mathbb{E}_{x_1,h_1}\left[\mathbb{E}_{x_2,\cdots,x_k}\left[\Delta_{h_1}f(x_1+x_2+\cdots+x_k)\prod_{i=2}^{k}g_i(x_1+h_1,x_2,\cdots,x_k)\bar{g}_i(x_1,x_2,\cdots,x_k)\right]\right]\right|^{1/2} \\
&\qquad\qquad \text{(By substituting } x_1' = x_1+h_1\text{)} \\
&\leq \left|\mathbb{E}_{x_1,h_1}\left[\mathbb{E}_{h_2,\cdots,h_k,z}\left[\Delta_{h_k}\cdots\Delta_{h_1}f(x_1+z)\right]^{1/2^{k-1}}\right]\right|^{1/2} \\
&\qquad\qquad \text{(By induction hypothesis and the definition of Gowers norm)} \\
&\leq \left|\mathbb{E}_{x_1,h_1,h_2,\cdots,h_k,z}\left[\Delta_{h_k}\cdots\Delta_{h_1}f(x_1+z)\right]\right|^{1/2^k} \qquad \text{(By Jensen's inequality)} \\
&= \left|\mathbb{E}_{h_1,h_2,\cdots,h_k,z}\left[\Delta_{h_k}\cdots\Delta_{h_1}f(z)\right]\right|^{1/2^k} = \|f\|_{U^k}
\end{aligned}$$

□

Proof of Lemma 2. By symmetry, it is enough to show that

$$\left|\mathbb{E}_{z_1,\cdots,z_m\in\mathbb{K}^n}[f_0(\mathcal{L}_0(z_1,\cdots,z_m))\prod_{i=1}^{k}f_i(\mathcal{L}_i(z_1,\cdots,z_m))]\right| \leq \|f_0\|_{U^k}.$$

We will make a linear change of variables so that we can use Lemma 6 to get the required bound. For each $1 \leq i \leq k$, since $\mathcal{L}_0$ is not a multiple of $\mathcal{L}_i$, there exists a vector $v_i \in \mathbb{K}^m$ such that $\mathcal{L}_0(v_i) = 1$ and $\mathcal{L}_i(v_i) = 0$. Now we make the following change of variables: $(z_1,\cdots,z_m) \to (x_1,\cdots,x_m) + \sum_{i=1}^{k} y_i v_i^T$ where $x_1,\cdots,x_m$ and $y_1,\cdots,y_k$ are the new variables which range over $\mathbb{K}^n$.

$$\begin{aligned}
&\left|\mathbb{E}_{z_1,\cdots,z_m\in\mathbb{K}^n}[f_0(\mathcal{L}_0(z_1,\cdots,z_m))\prod_{i=1}^{k}f_i(\mathcal{L}_i(z_1,\cdots,z_m))]\right| \\
&= \left|\mathbb{E}_{x_1,\cdots,x_m,y_1,\cdots,y_k\in\mathbb{K}^n}\left[f_0\left(\mathcal{L}_0(x_1,\cdots,x_m)+\sum_{j\in[k]}y_j\right)\prod_{i\in[k]}f_i\left(\mathcal{L}_i(x_1,\cdots,x_m)+\sum_{j\in[k]\setminus\{i\}}y_j\mathcal{L}_i(v_j)\right)\right]\right| \\
&\qquad\qquad \text{(By change of variables and linearity of } \mathcal{L}_i\text{)} \\
&\leq \mathbb{E}_{x_1,\cdots,x_m\in\mathbb{K}^n}\left[\left|\mathbb{E}_{y_1,\cdots,y_k\in\mathbb{K}^n}\left[f_0\left(\mathcal{L}_0(x_1,\cdots,x_m)+\sum_{j\in[k]}y_j\right)\prod_{i\in[k]}f_i\left(\mathcal{L}_i(x_1,\cdots,x_m)+\sum_{j\in[k]\setminus\{i\}}y_j\mathcal{L}_i(v_j)\right)\right]\right|\right] \\
&\leq \|f_0\|_{U^k} \qquad \text{(By Lemma 6)}
\end{aligned}$$

□

ECCC ISSN 1433-8092
http://eccc.hpi-web.de