List Decoding Quotient Reed-Muller Codes

Omri Gotlib
Department of Computer Science, Bar-Ilan University, Ramat-Gan, Israel

Tali Kaufman
Department of Computer Science, Bar-Ilan University, Ramat-Gan, Israel

Shachar Lovett
Department of Computer Science and Engineering, UC San Diego, CA, USA

Abstract

Reed-Muller codes consist of evaluations of n-variate polynomials over a finite field $\mathbb{F}$ with degree at most d. Much like every linear code, Reed-Muller codes can be characterized by constraints, where a codeword is valid if and only if it satisfies all *degree-d* constraints.

For a subset $\tilde{X} \subseteq \mathbb{F}^n$, we introduce the notion of *$\tilde{X}$-quotient* Reed-Muller code. A function $F : \tilde{X} \to \mathbb{F}$ is a valid codeword in the quotient code if it satisfies all the constraints of degree-d polynomials *lying in* $\tilde{X}$. This gives rise to a novel phenomenon: a quotient codeword may have *many* extensions to original codewords. This weakens the connection between original codewords and quotient codewords which introduces a richer range of behaviors along with substantial new challenges.

Our goal is to answer the following question: what properties of $\tilde{X}$ will imply that the quotient code inherits its distance and list-decoding radius from the original code?
We address this question using techniques developed by Bhowmick and Lovett [8], identifying key properties of $\mathbb{F}^n$ used in their proof and extending them to general subsets $\tilde{X} \subseteq \mathbb{F}^n$. By introducing a new tool, we overcome the novel challenge in analyzing the quotient code that arises from the weak connection between original and quotient codewords. This enables us to apply known results from additive combinatorics and algebraic geometry [34, 35, 37] to show that when $\tilde{X}$ is a *high rank variety*, $\tilde{X}$-quotient Reed-Muller codes inherit the distance and list-decoding parameters from the original Reed-Muller codes.

2012 ACM Subject Classification Theory of computation → Error-correcting codes; Mathematics of computing → Coding theory; Theory of computation → Pseudorandomness and derandomization

Keywords and phrases Reed-Muller Codes, Quotient Code, Quotient Reed-Muller Code, List Decoding, High Rank Variety, High-Order Fourier Analysis, Error-Correcting Codes

Digital Object Identifier 10.4230/LIPIcs.CCC.2025.1

Related Version *Full Version*: `https://arxiv.org/abs/2502.15650` [25]

Funding *Tali Kaufman*: Supported by ISF.
Shachar Lovett: supported by NSF award 2425349 and a Simons investigator award.

1 Introduction

Let $\mathbb{F}$ be a finite field, $n \in \mathbb{N}$, and let $\tilde{X} \subseteq \mathbb{F}^n$ be a subset[1] . We begin by introducing a new definition applicable to any linear code over $\mathbb{F}$: the *$\tilde{X}$-quotient code*. We then illustrate this novel definition using Reed-Muller codes, and present a property of $\tilde{X}$ which we use to show that $\tilde{X}$-quotient Reed-Muller code *inherits its distance and list decoding radius* from the original Reed-Muller code. Finally, leveraging known results from additive combinatorics and algebraic geometry, we establish as a corollary that this inheritance holds when $\tilde{X}$ is a *high-rank variety*.

[1] As a convention, we use $\tilde{\square}$ to denote properties of the subset, and thus also the subset itself.

The Quotient Code

Let $\mathfrak{C}$ be a linear code over $\mathbb{F}$. Each codeword of $\mathfrak{C}$ can be described as a function $F : \mathbb{F}^n \to \mathbb{F}$ that is in the span of the columns of the code's *generator matrix*. An equivalent way to describe $\mathfrak{C}$ is using a *parity check matrix*, where a function F is a codeword if and only if it satisfies the constraints represented by parity-check matrix. Each such constraint can be thought of as a requirement over a few inputs of F from $\mathbb{F}^n$: the requirement that their weighted sum will equal 0.

The first novel definition we introduce is the definition of the $\tilde{X}$*-induced* code:

▶ **Definition 1** (The $\tilde{X}$-Induced Code). *We define the* $\tilde{X}$-induced code $\mathfrak{C}_{\tilde{X}}$ *to be the set of all functions* $f : \tilde{X} \to \mathbb{F}$ [2] *that satisfy all the constraints of* $\mathfrak{C}$ that lie in $\tilde{X}$*, i.e. constraints that are supported only on points from* $\tilde{X}$ [3] *.*

Let us briefly describe the connection between codewords in $\mathbb{F}^n$ and $\tilde{X}$-induced codewords. One can easily verify that each original codeword *restricted* to $\tilde{X}$ is a valid codeword in the induced code. This is because each original codeword satisfies all the constraints in $\mathbb{F}^n$ by definition, and the constraints that words need to satisfy to be considered induced-codewords are only a subset of those constraints.
We call an extension of an $\tilde{X}$-induced codeword $f : \tilde{X} \to \mathbb{F}$ to valid codeword in the original code (extending its domain to $\mathbb{F}^n$), a *lift* of f. When each induced codeword has a unique lift, there is a natural 1-to-1 correspondence between the original and induced codeword. This becomes substantially more interesting for subsets $\tilde{X}$ in which induced codewords have *multiple* lifts. This non-uniqueness weakens the connection between the original codewords and induced codewords, and leads to a richer range of phenomena (and interesting new challenges).

We also note that the other direction is not always true: For a general subset $\tilde{X}$, there might be an induced codeword (a valid codeword in the induced code) that *cannot be lifted* to a valid codeword in $\mathbb{F}^n$. We are interested to better understand $\mathfrak{C}_{\tilde{X}}$ using $\mathfrak{C}$ and vice-versa, and therefore we introduce a new notion, which is the notion of the $\tilde{X}$*-quotient code*:

▶ **Definition 2** (The $\tilde{X}$-Quotient Code). *Let* $\mathfrak{C}$ *be a linear code, and let* $\mathfrak{C}_{\tilde{X}}$ *be the* $\tilde{X}$*-induced code of* $\mathfrak{C}$*. We say* $\mathfrak{C}_{\tilde{X}}$ *is a* $\tilde{X}$-quotient code *if every quotient codeword* $f \in \mathfrak{C}_{\tilde{X}}$ *has a lift to* $\mathbb{F}^n$*.*

In the case described above, we also say that $\tilde{X}$ is a *lift-enabler* for $\mathfrak{C}$ and that the code $\mathfrak{C}$ is a *covering code* for the code $\mathfrak{C}_{\tilde{X}}$.
The novelty of this definition is that it captures subsets in which *there is* a correspondence between codewords in $\tilde{X}$ and in $\mathbb{F}^n$, and the correspondence may be *1-to-many*.

Importance of Definition

This timely definition extends a fundamental and useful concept previously introduced for graphs and complexes – namely, the notion of a *covering graph* or alternatively, the *quotient graph*. This concept gained an increasing prominence in theoretical computer science, where it was recently employed to construct *high dimensional expanders* [18, 6] and achieve improved

[2] By convention, we use uppercase letters to denote functions with domain $\mathbb{F}^n$ and lowercase letters to denote functions with domain $\tilde{X}$.
[3] We note that this definition is in fact a property of the *constraints* of the code, such as other well-studied desired code properties (e.g. local testability).

local testing results [26, 19, 3], where the latter also played a crucial role in constructions of PCPs. Consequently, the study of covering spaces for graphs has found usages in theoretical computer science and specifically in development of PCPs with enhanced properties. We believe our question, which explores the analogous question for codes, will similarly lead to meaningful applications in theoretical computer science.

In addition to that, the question of *puncturing* of codes has caught much attention recently, in a line of work [14, 2, 12, 13], followed by the resolution of the GM-MDS conjecture [39, 45]. Where the question of puncturing is focused exclusively on the case where the lift is *unique*, the study of quotient codes also tackles subsets $\tilde{X} \subseteq \mathbb{F}^n$ where the lift is *not unique*. Notably, in the unique-lift case there are well-established lower-bounds for the size of $\tilde{X}$ such as [20, Theorem 1.1]. In contrast, the size of $\tilde{X}$ in quotient codes may be *much smaller* than its lower-bound in punctured code (for example in Reed-Muller codes), suggesting the potential for new insights and improved results.

Our Question

Our goal is to answer the following question: what properties of $\tilde{X}$ will imply that the quotient code inherits its distance and list-decoding radius from the original code?

This question is analogous to the study of quotients of expander graphs – just as not all quotients of an expander necessarily preserve expansion, not all subsets $\tilde{X}$ necessarily yield a well-behaved quotient code. Understanding the conditions under which expansion is preserved has been a fundamental problem in the study of expanders, and similarly, identifying the conditions under which a quotient code retains key properties of the original code is a central challenge in our work. Given this parallel, we believe our question may have broader implications for future research in both coding theory and theoretical computer science.

We answer this question in the context of *Reed-Muller codes*. Notably, our approach does *not only* address the case of where there are multiple lifts, but also introduces a novel framework for analyzing unique-lift (puncturing) setting when the field size is constant-a scenario that is typically considered more challenging.

Reed-Muller Codes

Let $\mathbb{F}$ be a finite field, and let n, d be integers. We focus on prime fields ($\mathbb{F} = \mathbb{F}_p$) and assume this setting unless explicitly stating otherwise. This assumption also applies to all fields considered in the works we reference [4] .

Each codeword in Reed-Muller code $RM_{\mathbb{F},\mathbb{F}^n}(d)$, is defined by a polynomial over $\mathbb{F}$ in n variables with total degree $\leq d$ [5] . The message that one wishes to encode is represented in the code as a polynomial $P : \mathbb{F}^n \to \mathbb{F}$, whose coefficients are the different message characters. The encoding of the message is a vector of the different evaluation of P over *all* possible points in $\mathbb{F}^n$.

Alternatively, one can describe Reed-Muller codes using a set of local constraints. A function $F : \mathbb{F}^n \to \mathbb{F}$ is a polynomial of degree $\leq d$ if and only if the (alternating) sum of each possible *cube*, which is a set of points of the form $\left\{x + \sum_{i \in S} y_i\right\}_{S \subseteq [d+1]}$ for $x, y_1, ..., y_{d+1} \in \mathbb{F}^n$, equals 0. Thus each cube represents a constraint, and we refer to the set of all cubes *the set of constraints of degree-d polynomials*. See Section 2.2 for more information in this regard.

[4] We believe that our techniques may extend to non-prime fields as well, but we do not pursue this direction in the current work.

[5] We focus on the regime where $d, |\mathbb{F}|$ are considered constants and n is considered very large.

Next, we present our notations for the induced Reed-Muller code:

▶ **Notation 3** (The $\tilde{X}$-Induced Reed-Muller Code)**.** *We say a function $F : \tilde{X} \to \mathbb{F}$ is a* polynomial of degree $\leq d$ *in $\tilde{X}$ if it satisfies all the constraints of degree-d polynomials* that lie in $\tilde{X}$ [6] .
We denote the $\tilde{X}$-induced Reed-Muller code:

$$RM_{\mathbb{F},\tilde{X}}(d) = \left\{p : \tilde{X} \to \mathbb{F} \middle| p \text{ is a polynomial of degree } \leq d \text{ in } \tilde{X}\right\}$$

Properties of Induced Reed-Muller Codes

A study of Ziegler and Kazhdan [34, 35, 36] shows that if $\tilde{X}$ is a *high rank variety* [7] , then $\tilde{X}$ is a *lift-enabler* for $RM_{\mathbb{F},\mathbb{F}^n}(d)$. In other words, the authors showed that the $\tilde{X}$-induced Reed-Muller code is in fact a *$\tilde{X}$-quotient Reed-Muller code*. We rely on this property of $\tilde{X}$ as a black-box. See Section 3 for more details in this regard.

An additional property of $\tilde{X} \subseteq \mathbb{F}^n$ we rely on is the connection between *algebraic structure* and *random behavior (equidistribution)* of polynomials in $\tilde{X}$.
For $\mathbb{F}^n$, this connection is a well-studied result [28, 33, 9]. It lies in the heart of many results in higher-order Fourier analysis, and specifically was used in [8] to analyze the list decoding radius of Reed-Muller code in $\mathbb{F}^n$.
The equivalent of this relation for subsets $\tilde{X} \subseteq \mathbb{F}^n$ was studied in [37, 27]. These works captured the measure of algebraic-structure in $\tilde{X}$ by a definition called *relative rank*, and captured the lack of random behavior in $\tilde{X}$ by a definition called *relative bias*. We note that for subsets, the definition of algebraic structure of a polynomial in $\tilde{X}$ considers the algebraic structure of *all its possible* lifts. It was shown in [37] that when $\tilde{X}$ is a high-rank variety, high relative rank implies low relative bias [8] .
We use this property as a black box as well. When a subset $\tilde{X} \subseteq \mathbb{F}^n$ has such property for polynomials of degree $\leq d$, we say that it has the *d-relative rank-bias property*. See Section 4 for more details.

Our Results

Next, let us present our main theorem more concretely. Our work focuses on the regime where $d < |\mathbb{F}|$. Denote the *minimum normalized distance of* $RM_{\mathbb{F},\mathbb{F}^n}(d)$ by $\delta_{\mathbb{F},\mathbb{F}^n}(d)$, shorthand by $\delta_{\mathbb{F}}(d)$. We have:

$$\delta_{\mathbb{F}}(d) = 1 - d/|\mathbb{F}|$$

Moreover, we define the *list decoding count* of $RM_{\mathbb{F},\mathbb{F}^n}(d)$ by:

$$\ell_{\mathbb{F},\mathbb{F}^n}(d,\tau) \coloneqq \max_{F:\mathbb{F}^n\to\mathbb{F}} \left|\{P \in Poly_{\leq d}(\mathbb{F}^n \to \mathbb{F}) | dist\,(P,F) \leq \tau\}\right|$$

Let $LDR_{\mathbb{F},\mathbb{F}^n}(d)$ be the *list decoding radius* of $RM_{\mathbb{F},\mathbb{F}^n}(d)$, which is the maximum τ for which $\ell_{\mathbb{F},\mathbb{F}^n}(d,\tau-\epsilon)$ is bounded by a *constant* depending only on $\epsilon, |\mathbb{F}|, d$.

---

[6] That is, the set of all cubes that their points are in $\tilde{X}$.
[7] Under some conditions we describe later.
[8] Note that even though Gowers and Karam [27] also acheived a similar relation for a type of subsets, the definition of rank they used is slightly different than the standard definition of rank. While this difference may seem unharmful at first, it is, to our knowledge, does not allow to do a *regularization* process (note that a generalization of this process is the heart of our proof).

In the paper [8] it was shown that for constant field size and degree, the list decoding radius *reaches the distance of the code*, as conjectured earlier by [24] [9] .

For $\tilde{X}$, we define the $\tilde{X}$-list decoding count:

$$\ell_{\mathbb{F},\tilde{X}}(d,\tau) \coloneqq \max_{F:\tilde{X}\to\mathbb{F}} \left|\left\{P \in Poly_{\leq d}(\tilde{X}\to\mathbb{F}) \middle| dist\left(P,F\right) \leq \tau\right\}\right|$$

We denote the distance parameters of $\tilde{X} \subseteq \mathbb{F}^n$ by $\delta_{\mathbb{F},\tilde{X}}(d)$ and $LDR_{\mathbb{F},\tilde{X}}(d)$ respectively,

We next present our main theorem, which establishes that the *list decoding radius* of the quotient Reed-Muller code is *at least as good* as the that of the original code:

▶ **Theorem** (List Decoding Quotient Reed-Muller Code). [10] *Let $\mathbb{F}$ be a finite field of constant size, let $d \in \mathbb{N}$ be a constant such that $d < |\mathbb{F}|$, and let $n \in \mathbb{N}$ be an integer.*
Let $\tilde{X} \subseteq \mathbb{F}^n$ be a subset that is a lift-enabler for $RM_{\mathbb{F},\mathbb{F}^n}(d)$ and has the d-relative rank-bias property.
Then, $RM_{\mathbb{F},\tilde{X}}(d)$ inherits its list decoding radius *from $RM_{\mathbb{F},\mathbb{F}^n}(d)$, i.e:*

$$LDR_{\mathbb{F},\tilde{X}}(d) \geq LDR_{\mathbb{F},\mathbb{F}^n}(d)$$

In addition, we also achieve a (simpler) result regarding the *distance* of the quotient Reed-Muller code (Theorem 68): Under the conditions described above, $RM_{\mathbb{F},\tilde{X}}(d)$ also inherits its *distance* from $RM_{\mathbb{F},\mathbb{F}^n}(d)$, i.e $\delta_{\mathbb{F},\tilde{X}}(d) \geq \delta_{\mathbb{F},\mathbb{F}^n}(d)$ [11] .

As a corollary, using results studied in [34, 35, 37] regarding high-rank varieties, we obtain the following:

▶ **Corollary** (List Decoding Quotient Reed-Muller Code: High Rank Variety). *Let $\tilde{X} \subseteq \mathbb{F}^n$ be a* high rank variety*, that is, $\tilde{X}$ is the set of common zeros of a collection of polynomials $\tilde{\mathcal{L}} = (L_1, ..., L_{\tilde{c}})$ that is of* high rank [12] [13] *, i.e. $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right) = \{x | \forall i : L_i(x) = 0\}$.*
Then, $RM_{\mathbb{F},\tilde{X}}(d)$ inherits its distance parameters from $RM_{\mathbb{F},\mathbb{F}^n}(d)$, i.e:
1. $\delta_{\mathbb{F},\tilde{X}}(d) \geq \delta_{\mathbb{F},\mathbb{F}^n}(d)$.
2. $LDR_{\mathbb{F},\tilde{X}}(d) \geq LDR_{\mathbb{F},\mathbb{F}^n}(d)$.

▶ **Example 4.** Let $d, n' \in \mathbb{N}$, and denote $n = d \cdot n'$. Define $L_n : \mathbb{F}^n \to \mathbb{F}$ to be:

$$L_n(x_1, ..., x_n) \coloneqq \sum_{i=1}^{n'} \prod_{j=1}^{d} x_{(i-1)\cdot d+j}$$

It was shown [35, Theorem 1.9] that $rank\left(L_n\right) \to \infty$ as $n \to \infty$. Thus for sufficiently large n, the variety $\tilde{X} \coloneqq Z\left(L_n\right) = \{x | L_n(x) = 0\} \subseteq \mathbb{F}^n$ satisfies the necessary conditions so that $\tilde{X}$-quotient Reed-Muller code inherits its distance parameters from the original Reed-Muller code [14] .

---

[9] Note that it is known that $LDR_{\mathbb{F},\mathbb{F}^n}(d) \leq \delta_{\mathbb{F}}(d)$, and therefore, in a sense, their result is *optimal in* $\mathbb{F}^n$ assuming $d, |\mathbb{F}|$ are considered as constants.
[10] Informal, for formal see Theorem 76.
[11] Our techniques also show that also the other direction is true, which yields an *equality* in the distance of the two codes.
[12] To be more precise, the greater or equals in the decoding parameters wirtten in the theorem are not exact, but they are true up to some ϵ that depend on the rank of the collection. Thus the higher the rank of the collection is, the more similar the quotient Reed-Muller code and the original Reed-Muller code in terms of distance and list-decoding radius.
[13] We also note that for this result some assumptions are needed regarding the field size or the degree of the polynomials in the collection.
[14] More accuratly, it was shown that its *schmidt rank* goes to infinity as n goes to infinity. This is a sufficient condition for applying Theorem 76.

Main Technical Challenge

We achieve these results by combining the two black-box properties of subsets $\tilde{X} \subseteq \mathbb{F}^n$ we presented. Analysis of the polynomials in $\tilde{X}$ raises a new challenge, as previous techniques that were used to analyze low-degree polynomials, both regarding $\mathbb{F}^n$ [28] and regarding subsets $\tilde{X}$ [37], were focused on maintaining the behavior of polynomials *within* the relevant domain, without maintaining a connection between polynomials over $\tilde{X}$ and their extensions to $\mathbb{F}^n$.

The novelty of our new technique is that it uses a similar approach to analyze polynomials in $\tilde{X}$ as commonly used in $\mathbb{F}^n$, *while simultaneously maintaining a connection* between polynomials in $\tilde{X}$ to polynomials in $\mathbb{F}^n$. This connection allows us to deduce that polynomials in $\tilde{X}$ behave similarly to polynomials in $\mathbb{F}^n$. Informally, given a question about a polynomial defined over $\tilde{X}$, our technique enables us to identify a suitable lift of the polynomial to a polynomial over $\mathbb{F}^n$, and answer the question using properties of that lift. Crucially, the appropriate lift depends on the nature of the question, meaning that no single canonical lift suffices for all purposes.

Next we describe this challenge in more detail.

Analyses of polynomials in $\mathbb{F}^n$ were commonly based on the structure-randomness connection of polynomials in $\mathbb{F}^n$. A central tool that leverages this property of $\mathbb{F}^n$ is the *regularization process* [28], which transforms any collection of polynomials into a new one that:

1. Is equidistributed in $\mathbb{F}^n$.
2. Captures the same functions as the original collection.

This capturing is formalized via measurability: a function $F : \mathbb{F}^n \to \mathbb{F}$ is said to be *measurable* with respect to a collection $\mathcal{P} = (P_1, ..., P_c)$ if it can be expressed as a function of the polynomials in $\mathcal{P}$. The regularization process guarantees that any function measurable by the original collection remains measurable by the new one.

Equidistribution is obtained by enforcing the collection to be of *high rank*, which informally means the polynomial has extremly low algebraic structure, that is, it cannot be approximated or predicted by a small number of lower-degree polynomials. In $\mathbb{F}^n$, this structural condition implies low *bias* in $\mathbb{F}^n$, where a low-biased polynomial is equidistributed.

To generalize this process to $\tilde{X} \subseteq \mathbb{F}^n$, while being able to deduce properties of polynomials in $\tilde{X}$ using polynomials in $\mathbb{F}^n$, we aim to construct a new collection of polynomials that has additional requirements:

3. Is equidistributed *restricted to* $\tilde{X}$.
4. Captures the same functions as the original collection *restricted to* $\tilde{X}$.

The requirement of the conditions *together* is unique to our setting and introduces a key difficulty. Equidistribution over $\tilde{X}$ requires the collection to have high *relative rank*, which requires one to eliminate structure not just from a polynomial, but from all of its $\tilde{X}$*-equivalent polynomials* – those over $\mathbb{F}^n$ that agree with it on $\tilde{X}$ and have the same degree bound [15] . Ensuring this while maintaining measurability in $\mathbb{F}^n$ is nontrivial.

To achieve equidistribution in $\mathbb{F}^n$, the regularization process replaces structured polynomials by a small collection of high-rank [16] polynomials that capture them. Over $\tilde{X}$, avoiding structure in all $\tilde{X}$-equivalents may require replacing a polynomial with polynomials that capture a different – but $\tilde{X}$-equivalent – function. This risks breaking measurability in $\mathbb{F}^n$ and thus losing the connection to the polynomials in $\mathbb{F}^n$.

[15] This is the same as considering all lifts of the polynomial $P|_{\tilde{X}}$, assuming such lift exist.
[16] More accuratly, the replacement is done recursively unitl the collection is of high rank.

In summary, our main technical challenge is to achieve equidistribution over $\tilde{X}$ while preserving measurability over both $\tilde{X}$ and $\mathbb{F}^n$, despite the need to eliminate structure across *all* $\tilde{X}$-equivalent polynomials.

Introducing New Tools

We overcome this challenge by presenting a new definition that relaxes the notion of *measurable* we required for functions in $\mathbb{F}^n$, which we call $\tilde{X}$*-measurable*. This enables us to describe a relaxed version of the regularization process, in which we require that every function in $\mathbb{F}^n$ that was $\tilde{X}$-measurable by the old collection will still be $\tilde{X}$-measurable by the new collection. In contrast to the original regularization process, which mandated that functions that were measurable by the old collection will be measurable by the collection, this relaxed definition only requires such functions to be $\tilde{X}$*-measurable* by the new collection.

Even though we no longer need to capture all previously captured functions in $\mathbb{F}^n$, it is important that the new relaxed-definition is strict enough to keep the connection between polynomials in $\tilde{X}$ and in $\mathbb{F}^n$. Therefore, maintaining the $\tilde{X}$-measurable functions throughout the regularization process cannot be done trivially, and this is handled in a procedure we call *the* $\tilde{X}$*-relative regularization process* which is a stronger-version of the regularization process that is used in $\mathbb{F}^n$. This new definition and procedure are thoroughly described in Section 5.

We note that these new definition and procedure are a novel contribution of this work, and we believe they can be useful in future research of the quotient Reed-Muller code.

1.1 Comparison to Related Work

In [8] the authors studied the list decoding radius of Reed Muller codes $\mathbb{F}^n$. They proved that, for prime fields, the list decoding radius *reaches the distance of the code*, as conjectured earlier by [24] [17] [18] . Formally, they showed the following theorem:

▶ **Theorem 5** ([8, Theorem 1]). *Let $\mathbb{F}$ be a prime field. Let $\epsilon > 0$ and $d, n \in \mathbb{N}$. There exists a constant* [19] *$c := c(|\mathbb{F}|, d, \epsilon)$ such that:*

$$\ell_{\mathbb{F},\mathbb{F}^n}(d, \delta_{\mathbb{F}}(d) - \epsilon) \leq c$$

Our work gives new tools for analyzing polynomials in $\tilde{X} \subseteq \mathbb{F}^n$, which we later use to follow their line of proof and show an equivalent result *in* $\tilde{X}$.

We next present related work regarding the study of polynomial codes in subsets $\tilde{X} \subseteq \mathbb{F}^n$. Before presenting them specifically, we note that our work has a *fundamental difference* than that of the previous study of polynomials in subsets. Most works which studied polynomials over subsets $\tilde{X} \subseteq \mathbb{F}^n$ were focused on subsets in which every polynomial has a *unique* lift. This ensures that there is a 1-to-1 correspondence between polynomials in $\tilde{X}$ and in $\mathbb{F}^n$ and therefore allows easier connection between polynomials in $\tilde{X}$ and in $\mathbb{F}^n$.
We note that our work is non-trivial even in this case: it extracts the properties of $\mathbb{F}^n$ that were used in [8], in a way they can be used to analyze quotient Reed-Muller codes. However, as described earlier, our work addresses an additional substantial challenge which arise when the lift is *not* unique. Thus our work is only comparable to other works in the unique-lift case, which is the less-challenging case we address.

[17] Note that it is known that $LDR_{\mathbb{F},\mathbb{F}^n}(d) \leq \delta_{\mathbb{F}}(d)$, and therefore, in a sense, their result is *optimal in* $\mathbb{F}^n$ assuming $d, |\mathbb{F}|$ are considered as constants.
[18] We also note that their work also apply to the regime $d \geq |\mathbb{F}|$.
[19] It is important to note that c is *independent of* n.

The first line of work we mention is this regard is the study of hitting sets for low degree polynomials [40, 15, 31], and a stronger variant of it which is the study of pseudorandom-generators against low degree polynomials [10, 11, 38, 44, 16, 17, 22] Both definitions capture subsets [20] $\tilde{X} \subseteq \mathbb{F}^n$ such that every polynomial over $\mathbb{F}^n$ has a non-negligible distance from 0 *when restricted to* $\tilde{X}$. This requirement implicitly implies that every low degree polynomial over $\tilde{X}$ has at most a *single* lift.

Another line of work worth mentioning in this regard is [21, 30], which studied *puncturing of Reed-Muller codes*. This line of work studied the construction of sets $\tilde{X} \subseteq \mathbb{F}^n$, such that puncturing Reed-Muller codes over $\tilde{X}$, that is, taking every original codeword and *restricting* it to $\tilde{X}$, will yield a good error-correction code. To perform their analysis, it was important that every polynomial in $\tilde{X}$ has at most a single lift, and therefore their work was focused on subsets where there is a unique lift.

The papers [14, 2, 12] also studied similar questions. This line of work is followed by the resolution of the *GM-MDS conjecture*, which was proved by [39, 45].
We note that these works were focused on the regime where the field is *large*. More specifically, they require that the field is *large in respect of* n, i.e. $\Omega(n)$. We emphasize that our work is focused on *constant fields*. Moreover, their results were regarding *random* puncturing, while our result makes an *explicit* puncturing.

We also note that most studies presented above also achieved results regarding the *rate* of the punctured code. This property of the code can be analyzed naturally when each polynomial over $\tilde{X}$ has a *unique* lift, as such assumption implies that the number of polynomials remains the same in $\tilde{X}$ as of in $\mathbb{F}^n$. In contrast, our work does not rely on such a uniqueness assumption, and therefore does not address the rate of the resulting code. As our work does *not* assume such uniqueness, the rate of the code we consider is not analyzed in our work. Nonetheless, we note that the *Hilbert function* of a subset $\tilde{X} \subseteq \mathbb{F}^n$ corresponds to the rate of the $\tilde{X}$-quotient Reed-Muller code, and that some great progress has been made in analyzing this function [5, 4, 41, 23, 1].

1.2 Proof Overview

Our main technical contribution is a generalization of the regularization process to subsets $\tilde{X} \subseteq \mathbb{F}^n$, which we call the *relative regularization process*. This tool addresses the core difficulty of non-unique lift, and relies on a new notion we introduce: $\tilde{X}$-measurability.

Measurablity and Regularization in $\mathbb{F}^n$

Measurablility is a mathematical-analysis notion, which was first used in a similar context in [29]. It is defined as follows:

▶ **Definition 6** (Measurable). *Let* [21] $\mathcal{P} = (P_1, ..., P_c)$ *be a collection of polynomials of degree* $\leq d$. *A function* $F : \mathbb{F}^n \to \mathbb{F}$ *is* measurable in respect of $\mathcal{P}$ *if it can be determined by the values of* $P_1, ..., P_c$*:*

$$F(x) = \Gamma_F(P_1(x), ..., P_c(x))$$

for some function $\Gamma_F : \mathbb{F}^c \to \mathbb{F}$ [22] .

[20] Sometimes this subset is allowed to be a *multiset*.
[21] In this context we think of c as a small (constant for example).
[22] One can think of this definition as a generalization of linear span: the collection *spans* the function, where Γ is some notion of a span.

Intuitively, $\mathcal{P}$ captures the information required to compute F. If $\mathcal{P}$ is a high-rank collection, the tuple $(P_1(x), \ldots, P_c(x))$ is equidistributed over $x \in \mathbb{F}^n$ as $\mathbb{F}^n$ has the property that high-rank implies equidistribution. This allows one to analyze F through the simpler function Γ_F.

The regularization process is a fundamental tool presented in [28] that constructs a high-rank collection $\mathcal{P}'$ refining $\mathcal{P}$, meaning that all $\mathcal{P}$-measurable functions remain measurable with respect to $\mathcal{P}'$.

$\tilde{X}$-Relative Rank

We remind the reader that rank is a notion that measures the algebraic structure of polynomials, where high rank implies extremly low algebraic structure. In addition, $\tilde{X}$-relative rank is a notion that measures the algebraic structure of a polynomial *in a subset* $\tilde{X} \subseteq \mathbb{F}^n$, by considering the structure of all of its $\tilde{X}$-equivalent polynomials. This notion was presented by [27, 37], and is used to achieve equidistribution in $\tilde{X}$ assuming $\tilde{X}$ has relative rank-bias property. It is defined as follows:

▶ **Definition 7** (Relative Rank, informal. See definition 48)**.** *Let $\tilde{X} \subseteq \mathbb{F}^n$ be a subset, let $d \in \mathbb{N}$, and let $P : \mathbb{F}^n \to \mathbb{F}$ be a polynomial of degree $= d$. The $\tilde{X}$-relative rank of P is defined as follows:*

$$rank_{\tilde{X}}(P) := \min\left\{rank\left(P - \overline{P}\right) \middle| \overline{P} \in Poly_{\leq d}(\mathbb{F}^n \to \mathbb{F}), \overline{P}|_{\tilde{X}} \equiv 0\right\}$$

1.2.1 $\tilde{X}$-Measurablity and The $\tilde{X}$-Relative Regularization Process

In this subsection we discuss the generalization of the regularization process to subsets $\tilde{X} \subseteq \mathbb{F}^n$ using the equivalent of rank-bias relation in $\tilde{X}$. We name this tool *the relative regularization process*.

Practically, we use this tool to show that given a specific question in mind, every $p : \tilde{X} \to \mathbb{F}$ has some polynomial $P : \mathbb{F}^n \to \mathbb{F}$ that behave "similarly" in respect to this question. This allows us to pull properties of P to better understand p. The perfect candidate for such P is a *lift* of p.

In order to use P to deduce properties of p, we use the well-studied properties of polynomials in $\mathbb{F}^n$ to acheive properties of P, and relate these to properties of p. More specifically, assume that p and P are measurable in respect of a collection of polynomials $\mathcal{P}$ (each in its domain). Our strategy is to use P to deduce properties of Γ_P, and then use the properties of Γ_P to deduce properties of p.

Now let us describe the extra challenge. We start by following the ideas of the regularization process we described for $\mathbb{F}^n$. Assuming the collection is not a collection of $\tilde{X}$-relative high rank, then there must exist a polynomial in the collection that has low *relative* rank, which we denote by $P^\star$ [23] . Note that in relative rank, this does not necessarily mean that $P^\star$ is of low rank, but that there exists another $\tilde{X}$-equivalent polynomial that has a low rank. Thus, even if we remove the low-rank $\tilde{X}$-equivalent polynomial and add to the collection all the polynomials that decomposed it, we cannot require that every function that was measurable by the old collection will still be measurable by the new collection: even the polynomial we removed is not necessarily measurable by the new collection!

[23] More precisely, some linear combination of polynomials has low relative rank.

To allow such regularization process to still apply, we note that while P might not be measurable in respect of the new collection, a $\tilde{X}$-equivalent polynomial of P *is* measurable with respect of it. Therefore, we relax the notion of being measurable to being $\tilde{X}$*-measurable*. We say a function F is $\tilde{X}$-measurable in respect of $\mathcal{P}$ if it can be determined by the polynomials of $\mathcal{P}$ *up to a valid* $\tilde{X}$*-remainder*. We first describe an incomplete definition, then present the challenge that rises with it, and finally present its resolution.

▶ **Definition 8** ($\tilde{X}$-measurable, Incomplete Definition)**.** [24] *We say a function F is* $\tilde{X}$-measurable in respect of $\mathcal{P} = (P_1, ..., P_c)$ *if there exists a function* $\Gamma : \mathbb{F}^c \to \mathbb{F}$ *and a* $\tilde{X}$*-remainder, i.e. a function* $\overline{F} : \mathbb{F}^n \to \mathbb{F}$ *with* $\overline{F}|_{\tilde{X}} \equiv 0$ *such that:*

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma(P_1(a), ..., P_c(a)) + \overline{F}(a)$$

Previous works analyzing polynomials in $\mathbb{F}^n$ were able to deduce two things from F being measurable by a high-rank collection $\mathcal{P}$. The first, is that the structure of Γ is similar to the structure of F: for example, if one is a polynomial of bounded degree, so is the other. The second, is that a random input of Γ behave similarly to a random input of F: that is, the output distribution of Γ (over its inputs $\mathbb{F}^c$) is close to the output distribution of F (over its inputs $\mathbb{F}^n$).
To study polynomials in $\tilde{X}$, we wish to connect p to P (which is a lift of p). Thus, we think of $F = P$, and require two similar things. Firstly, we want the structure of Γ to be similar to the structure of F (in this case, P), which we understand as F is a polynomial in $\mathbb{F}^n$. Secondly, we want a random input of Γ to behave similarly to a random input of p, as p is the polynomial we wish to understand. The latter is easily achieved using the fact high $\tilde{X}$-relative rank implies equidistribution in $\tilde{X}$. The former, however, might be damaged by the remainder as we defined it: we can only learn the structure of Γ using the structure of $F - \overline{F}$ using the equality $F - \overline{F} = \Gamma(P_1, ..., P_c)$. However, the structure of $F - \overline{F}$ can be very different from the structure of F, as we did not require any structure of the $\tilde{X}$-remainder $\overline{F}$. Thus, we can not deduce the structure of Γ via the structure of F using the incomplete definition described above.

To handle this issue, we add one more requirement regarding the $\tilde{X}$-remainder, which ensures that the structure of F can be understood via the structure of Γ:

$$\deg(F - \overline{F}) \leq \deg(F)$$

If the $\tilde{X}$-remainder also has this property, we say it is a *valid* $\tilde{X}$-remainder for F. This can be summarized by the following (complete) definition:

▶ **Definition 9** ($\tilde{X}$-measurable)**.** *We say a function F is* $\tilde{X}$-measurable in respect of $\mathcal{P} = (P_1, ..., P_c)$ *if there exists a function* $\Gamma : \mathbb{F}^c \to \mathbb{F}$ *and a* valid $\tilde{X}$*-remainder, i.e. a function* $\overline{F} : \mathbb{F}^n \to \mathbb{F}$ *with* $\overline{F}|_{\tilde{X}} \equiv 0$ *and* $\deg(F - \overline{F}) \leq \deg(F)$ *such that:*

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma(P_1(a), ..., P_c(a)) + \overline{F}(a)$$

We use this new definition the following way: Instead of using F to understand Γ, we use $F - \overline{F}$ to do so. We choose $F - \overline{F}$ as it has the same structure as F, but it is "closer" to the function Γ as $F - \overline{F} = \Gamma(P_1, ..., P_c)$ [25] . Finally, as Γ behaves similarly to p for random inputs, we can use Γ to deduce properties regarding p.

[24] This incomplete definition lacks the requirement of the *validity* of the $\tilde{X}$-remainder
[25] One can think of this step as "taking the right $\tilde{X}$-equivalent" in respect of $\mathcal{P}$.

With this in hand, let us finish describing the relative-regularization process. The requirement on the validity of the $\tilde{X}$-remainder raises a new challenge in the $\tilde{X}$-relative regularization process: we need to somehow control the structure of the $\tilde{X}$-remainder, even though this "error" is substituted in Γ each time we wish to replace a polynomial in our collection. We address this challenge using a Lemma proved in [9] called the "faithful composition lemma", which allows us to deduce strong properties regarding the structure of Γ given the collection was of a high (regular) rank in the first place. Therefore, we add to each step of the relative-regularization process a (regular) regularization, which ensures Γ is very structured. This strong structure of Γ is later used to control the error and deduce it is in the form of a valid $\tilde{X}$-remainder. For the exact details, see Theorem 64. We conclude this by informally stating our main technical theorem, which is the relative regularization process we just described:

▶ **Theorem 10** (Relative Regularization Process, Informal, See Theorem 64)**.** *Let $r, d \in \mathbb{N}$ be integers that represents a requested rank and degree respectively, and let $P_1, ..., P_c$ be a collection of polynomials of degree $\leq d$. Then, there is another collection $P'_1, ..., P'_{c'}$ of polynomials of degree $\leq d$, such that:*

1. *Every function that is $\tilde{X}$-measurable in respect to the first collection is also $\tilde{X}$-measurable in respect to the new collection.*
2. *The new collection is of $\tilde{X}$-relative rank $\geq r$.*
3. *The new collection is of bounded size, i.e. $c' \leq C_{r,d,c}$.*

1.2.2 List Decoding in $\tilde{X}$ via $\tilde{X}$-Relative Regularization

In this subsection, we demonstrate how to use the relative regularization process to achieve our main theorem: analysis of the list decoding radius of $RM_{\mathbb{F},\tilde{X}}(d)$.

We follow the line of proof of [8], but this time, we are interested in bounding the amount of polynomials *in* $\tilde{X}$ around every function *in* $\tilde{X}$. More specifically, we wish to show that there is a constant number of words that are $(\delta_{\mathbb{F}}(d) - \epsilon)$-close to any fixed function in $\tilde{X}$.

Let $f : \tilde{X} \to \mathbb{F}$ be a received word. First, we apply a lemma proved in [8, Corollary 3.3]. The lemma shows that there is a constant-sized (depending on ϵ) collection of polynomials in $\tilde{X}$, denoted by $\mathfrak{h}$, such that the distance of f to *any* polynomial can be approximated by the distance of f to some function that is measurable by $\mathfrak{h}$ in $\tilde{X}$. This means that instead of bounding the number of polynomials in the radius of f, one can bound the number of polynomials in the radius of some function measurable by $\mathfrak{h}$. Thereby, every polynomial-specific measurable function can be thought of as a *low complexity proxy* for f in respect to the polynomial.

Next, we lift each polynomial from $\mathfrak{h}$ and apply the *relative regularization process*. This yields a new collection of polynomials in $\mathbb{F}^n$ that is constant sized and randomly-behaving (in $\mathbb{F}^n$). Denote this new collection by $\mathcal{H}'$ [26] . Thereby, the question of list decoding is reduced to the following question: We have a specific constant-sized randomly-behaving collection of polynomials $\mathcal{H}' = \{H'_1, ..., H'_{c'}\}$ that was constructed using the function f. We need to bound the amount of polynomials in $\tilde{X}$ that are $(\delta_d(\mathbb{F}) - \epsilon/2)$-close to be measurable by this collection in $\tilde{X}$. Note that the randomly-behaving property was achieved using the *relative rank-bias property* of $\tilde{X}$. Additionally, we note the collection $\mathcal{H}'$ is a collection of polynomials in $\mathbb{F}^n$ which we obtained by using the *lift-enabler property* of $\tilde{X}$.

[26] We use the same notations as the original proof for clearannce.

From there (and similarly to the analysis in $\mathbb{F}^n$), the strategy is to show that polynomials that are that close to being measurable by the randomly-behaving collection $\mathcal{H}'$, are in fact *measurable* by it. This will bound the number of such polynomials by the amount of possible functions that are measurable by $\mathcal{H}'$, which is constant as the collection is of constant size.

Let $p : \tilde{X} \to \mathbb{F}$ be a polynomial of degree $\leq d$, and consider a lift of it $P : \mathbb{F}^n \to \mathbb{F}$. Consider the collection $\mathcal{H}' \cup \{P\}$. Surely, P is measurable by this collection in $\mathbb{F}^n$. Applying $\tilde{X}$-relative-regularization to this collection yields a new collection $\mathcal{H}''$ that is equidistributed in $\tilde{X}$, such that every $\tilde{X}$-measurable function by the old collection is $\tilde{X}$-measurable by the new collection. By a reason we have not explained in this brief explanation, we can ensure this collection is of the form $\mathcal{H}'' = \mathcal{H}' \cup \{H''_1, ..., H''_{c''}\}$.

As P was $\tilde{X}$-measurable by $\mathcal{H}' \cup \{P\}$ (it was even measurable), P is $\tilde{X}$-measurable by the new collection $\mathcal{H}''$: That is, P is measurable by $\mathcal{H}''$ up to a *valid* remainder, denoted by $\overline{P}$. This means there exists $\Phi : \mathbb{F}^{c'+c''} \to \mathbb{F}$ such that:

$$\forall a \in \mathbb{F}^n : P(a) = \Phi(H'_1(a), ..., H'_{c'}(a), H''_1(a), ..., H''_{c''}(a))) + \overline{P}(a)$$

In $\mathbb{F}^n$, the proof would follow by studying the structure of the function Φ and use it to induce that Φ does not depend on its last c'' variables. This implies that P is measurable by the original collection $\mathcal{H}'$ which concludes the proof [27] .

More accurately, the analysis in $\mathbb{F}^n$ used the fact that substituting *randomly behaving* polynomials in Φ yields a structured function [28] . This is used to show that Φ as a function by itself, with inputs from $\mathbb{F}^{c'+c''}$, is a very structured function. The strong structure of Φ, with the fact that Φ (with inputs substitued to be the functions of $\mathcal{H}''$) is close to the function f, are then combined to deduce that Φ does not depend on its last c'' variables.

This paradigm can not be extended effortlessly to our case. In $\tilde{X}$, deducing that Φ is very structured requires a one-more major step. This is because we do *not* know any correspondence in the behavior of Φ (which we want to understand) with the behavior of P (which we know is structured). We only know there is a correspondence between Φ to another function $P - \overline{P}$, which apriori we do not know is structured!

Fortunately, the relative regularization process (Theorem 64) mandates that the remainder of the measurement is *valid*. That is, if P was structured (a polynomial of degree $\leq d$), then so does $P - \overline{P}$. This is *crucial*, as it allows us to use the relation between Φ and $P - \overline{P}$ to deduce that Φ is structured, and continue the original outline of the proof of [8]. For more details in this regard, see Theorem 76.

1.3 Organization

In Section 2 we present some basic notations and conventions, and define the preliminaries we have regarding high-order Fourier analysis in $\mathbb{F}^n$: polynomials, rank and regularization. We later generalize each component we presented in Section 2 to study polynomials in $\mathbb{F}^n$ to also study polynomials in $\tilde{X}$: in Section 3 we present the set of polynomials in $\tilde{X}$ and present the *lift-enabler property*; in Section 4 we present the *$\tilde{X}$-relative rank-bias property*; and in Section 5 we present the $\tilde{X}$-measurable notion, and our main tool, which is the *$\tilde{X}$-relative regularization process*. Next, we present two applications regarding the distance parameters of Reed-Muller codes in $\tilde{X}$: In Section 6 we prove the inheritance of the *distance* of the code; and in Section 7 we prove the inheritance of the *list decoding distance* of it (which is much more involved).

---

[27] Note that in $\mathbb{F}^n$ there is no remainder, so the equation above (with the last c'' variables as constants) implies measurability by $\mathcal{H}'$.

[28] In our notations, this structured function is P, which is a polynomial of degree $\leq d$ and thus structured

2 Preliminaries

2.1 Basic Definitions and Notations

We denote by $\mathbb{N}$ the set set of integers, i.e. natural numbers (excluding 0). For an integer k we denote $[k] \coloneqq \{1, 2, ..., k\}$. We use $y = x \pm \epsilon$ to denote $y \in [x - \epsilon, x + \epsilon]$, and similarly $y = x \mp \lambda$ to denote $y \in [x + \lambda, x - \lambda]$ (usually when $\lambda < 0$).
Fix a prime field $\mathbb{F} = \mathbb{F}_p$. Denote by $|\cdot|$ the natural map of $\mathbb{F}$ to $\{1, ..., p-1\} \in \mathbb{N}$. We denote the character from $\mathbb{F}$ by $e[x] \coloneqq e^{2\pi i \cdot |x|}$

Generally speaking and unless stated otherwise, we use the following conventions: We use $n \in \mathbb{N}$ to denote the number of variables in Reed-Muller code. We use d to denote a degree (typically the degree of the polynomials in our code), and $\tilde{X}$ to denote the subset of $\mathbb{F}^n$ we work in i.e. $\tilde{X} \subseteq \mathbb{F}^n$. Properties of the subset $\tilde{X}$ will usually be denoted with $\tilde{\square}$. We use F, G, H to denote general functions with domain $\mathbb{F}^n$, and f, g, h to denote functions with domain $\tilde{X}$. We use $\mathfrak{F}, \mathfrak{G}, \mathfrak{H}$ and $\mathfrak{f}, \mathfrak{g}, \mathfrak{h}$ respectively to denote sets of such functions. Similarly, we use P, Q, H to denote polynomials with domain $\mathbb{F}^n$, and p, q, h polynomials with domain $\tilde{X}$ (polynomials as defined in Section 3). We use $\mathcal{P}, \mathcal{Q}, \mathcal{H}$ and $\mathfrak{p}, \mathfrak{q}, \mathfrak{h}$ respectively to denote sets of such polynomials.

2.2 Polynomials in $\mathbb{F}^n$

We start by presenting a standard definition for a polynomial over a finite field.

▶ **Definition 11** (Polynomial: Global Definition). *Let $d \in \mathbb{N}$ be a constant. A function $P : \mathbb{F}^n \to \mathbb{F}$ is called a* polynomial of degree $\leq d$ *if it is of the following form:*

$$P(x_1, ..., x_n) = \sum_{0 \leq d_1, ..., d_n : \sum_{i=1}^n d_i \leq d} c_{d_1, ..., d_n} \prod_{i=1}^n x_i^{d_i}$$

We denote the set of all polynomials of degree $\leq d$ by $Poly_{\leq d}(\mathbb{F}^n \to \mathbb{F})$. The value d in the definition above is called the global degree *of the function P, shorthand by the* degree *of P, and it is denoted by $\deg(P) = d$.*
Additionally, the set of all polynomials from $\mathbb{F}^n$ to $\mathbb{F}$ of degree $\leq d$ is denoted by:

$$Poly_{\leq d}(\mathbb{F}^n \to \mathbb{F})$$

▶ Note 12. Note that it is a folklore that every function $F : \mathbb{F}^n \to \mathbb{F}$ is a polynomial function, that is, can be written in the representation stated above for *some* degree. This follows from the representation of *Dirac delta function* over $\mathbb{F}^n$ as a polynomial:

$$\mathbb{1}_{\vec{0}}(x_1, ..., x_n) \coloneqq \begin{cases} 1 & x_1 = \cdots = x_n = 0 \\ 0 & \text{otherwise} \end{cases} = \prod_{i=1}^n (1 - (x_i)^{|\mathbb{F}|-1})$$

Therefore the definition of total degree is meaningful for *all* functions $F : \mathbb{F}^n \to \mathbb{F}$.

Next, we present a known equivalent definition for a polynomial using derivatives. To do so, we first define a derivative in the case of finite fields.

▶ **Definition 13** (Derivative). *Given a function $F : \mathbb{F}^n \to \mathbb{F}$ and $a \in \mathbb{F}^n$, we define the derivative of F in direction a as a function $D_a F : \mathbb{F}^n \to \mathbb{F}$ defined as follows:*

$$D_a F(x) \coloneqq F(x + a) - F(x)$$

▶ **Lemma 14.** *Let $d \in \mathbb{N}$. A function $F : \mathbb{F}^n \to \mathbb{F}$ is a polynomial of degree $\leq d$ if and only if $D_a F$ is a polynomial of degree $\leq d-1$ for all $a \in \mathbb{F}^n$.*

This leads us to a natural definition of a *degree* of a function using derivatives.

▶ **Definition 15** (Local Degree)**.** *For a function $F : \mathbb{F}^n \to \mathbb{F}$, we define its* local degree*, to be the least integer $d \in \mathbb{N}$ such that for all $a_1, ..., a_{d+1}, x \in \mathbb{F}^n$:*

$$D_{a_{d+1}} ... D_{a_1} F(x) = 0$$

In $\mathbb{F}^n$, the two definitions of degree coincide, and we get a single definition of a degree:

▶ **Lemma 16** (Equivalance of definitions of a degree)**.** *Let $F : \mathbb{F}^n \to \mathbb{F}$ be a function, and let $d \in \mathbb{N}$ be an integer. Then, the global degree of a F* equals *its local degree.*

▶ Remark 17. We sometimes refer to the requirement that the local degree of a function is $\leq d$, as the *local criteria* of degree $\leq d$ polynomials.

2.3 Rank-bias in $\mathbb{F}^n$

We start by defining the notion of *bias*, which is a measure of how the function is far from being equidistributed (see Appendix A for the exact details).

▶ **Definition 18** (Bias)**.** *Let $F : \mathbb{F}^n \to \mathbb{F}$. The* bias *of the function F is defined in the following way:*

$$bias(F) \coloneqq 1/\left|\mathbb{F}^n\right| \cdot \sum_{x \in \mathbb{F}^n} e\left[F(x)\right]$$

Moreover, for a subset $\tilde{X} \subseteq \mathbb{F}^n$, we define the bias of F in $\tilde{X}$ *to be:*

$$bias_{\tilde{X}}(F) \coloneqq 1/\left|\tilde{X}\right| \cdot \sum_{x \in \tilde{X}} e\left[F(x)\right]$$

Next, we present a standard definition of rank of a polynomial, which is a notion that measures how *structured* is the function. Note that low rank implies the polynomial is highly structured. Formally we have the following definition:

▶ **Definition 19** (Rank of a Polynomial)**.** *Given a constant $d \in \mathbb{N}$ and a polynomial P, the* d-rank *of P, denoted as $rank_d(P)$ is defined to be the smallest integer r such that P can be computed given r polynomials of degree $< d$. In other wards, we say $rank_d(P) = r$ if r is the smallest integer such that there exists r polynomials $Q_1, ..., Q_r \in Poly_{\leq d-1}(\mathbb{F}^n \to \mathbb{F})$ and a function $\Gamma : \mathbb{F}^n \to \mathbb{F}$ such that:*

$$P(x) = \Gamma\left(Q_1(x), ..., Q_r(x)\right)$$

If $d = 1$, then 1-rank is defined to be ∞ if P is non constant, and 0 otherwise.
Moreover, for a polynomial P of degree $\deg(P) = d$ we denote $rank(P) \coloneqq rank_d(P)$.
We call such function Γ a decomposition *or a* computation *of P using lower-degree polynomials.*

Let us now define a factor. Note that we focus our discussion to factors in $\mathbb{F}^n$, but define the basic definitions over a general set U so they will apply for factors over a general sets. This is necessary as we will later use them also for other sets such as $\tilde{X} \subseteq \mathbb{F}^n$.

▶ **Definition 20** (Factor). *Let $U \subseteq \mathbb{F}^n$ be a set. Let $F_1, ..., F_c : U \to \mathbb{F}$ be a collection of functions. A* factor *defined by $\mathcal{B}_{F_1,...,F_c}$ over U is the map:*

$$\mathcal{B}_{F_1,...,F_c}(u) \to (F_1(u), ..., F_c(u))$$

By an abuse of notation, we also use $\mathcal{B}$ to denote the partition of the set U defined by the map. We call each subset in the partition is an atom*:*

$$\{u \in U | F_1(u) = b_1, ..., F_c(u) = b_c\}$$

for all $b_1, ..., b_c \in \mathbb{F}$. By an abuse of notation, $\mathcal{B}$ sometimes refers to the set of all atoms (which is a partition of U).

▶ **Notation.** *Let $F_1, ..., F_c : U \to \mathbb{F}$ be a collection of functions. For a factor $\mathcal{B} := \mathcal{B}_{F_1,...,F_c}$, we denote by $|\mathcal{B}|$ the amount of functions that define it, i.e. $|\mathcal{B}| = c$. Moreover, we denote $\|\mathcal{B}\| := |\mathbb{F}|^c$, which is the maximal amount of (possibly empty) atoms.*

▶ **Definition 21** (Polynomial Factor). *We say a factor $\mathcal{B}$ over $\mathbb{F}^n$ is a* polynomial factor *if it is defined by a collection of polynomials $P_1, ..., P_c : \mathbb{F}^n \to \mathbb{F}$, i.e. $\mathcal{B} = \mathcal{B}_{P_1,...,P_c}$. The degree of the factor, denote as $\deg(\mathcal{B})$ is the maximal degree of the polynomials $P_1, ..., P_c$.*

▶ Note. We emphasize that every function $F : \mathbb{F}^n \to \mathbb{F}$ is a polynomial function for some degree, thus the phrase "polynomial factor" is used to emphasize that there is a degree bound on the functions defining the factor. Also note that the notion of degree (and polynomial) are defined only for functions over $\mathbb{F}^n$, therefore this definition is well-defined only for $U = \mathbb{F}^n$.

▶ **Definition 22** (Rank of a Factor). *Let $\mathcal{P}$ be a collection of polynomials $P_1, ..., P_c : \mathbb{F}^n \to \mathbb{F}$. The rank of the polynomial collection is defined as:*

$$rank\,(\mathcal{P}) := \min \left\{ rank_d \left(\sum_{i=1}^{c} \lambda_i P_i \right) \middle| 0 \neq \vec{\lambda} \in \mathbb{F}^c, d = \max_{i \in [c]} \deg(\lambda_i P_i) \right\}$$

For a factor $\mathcal{B}$ defined by a collection of polynomials $\mathcal{P}$, we define its rank to be the rank of the collection of polynomials defining it. For a non-decreasing function $r : \mathbb{N} \to \mathbb{N}$, a factor $\mathcal{B}$ is called r-regular if its rank is at least $r(|\mathcal{B}|)$.

▶ Note. Note that in the definition above, the rank of each linear combination is calculated as the d-rank, where d is the maximal degree of a polynomial that participates in the linear combination non-trivially. This is crucial as it ensures that a high rank factor do not have linear dependence in the largest-degree homogenous component of any of its polynomials.

We now present a fundamental property of high rank polynomials, that was first proved by [28] when $d < |\mathbb{F}|$, later extended to general fields by [33], and further extended also to large fields by [9]. This property of high rank polynomials is that they have low bias:

▶ **Theorem 23** (Rank-bias in $\mathbb{F}^n$). *Let $\mathbb{F}$ be a finite field. Let $\epsilon > 0$ and $d \in \mathbb{N}$. There exists $r_{23} := r_{23}(\mathbb{F}, d, \epsilon)$, such that for every degree-d polynomial $P : \mathbb{F}^n \to \mathbb{F}$: if $rank\,(P) \geq r_{23}$ then $bias(P) < \epsilon$.*

▶ Remark 24. This property implies that a collection of polynomials that have high rank is *equidistributed*. See Appendix A for more details in this regard.

2.4 Regularization in $\mathbb{F}^n$

In this subsection we define the regularization process in $\mathbb{F}^n$. Before doing so, let us present some definitions in this regard. Note that we define the basic definitions over a general set U so they will apply for factors over a general sets, as this is necessary as we will later use them also for other sets such as $\tilde{X} \subseteq \mathbb{F}^n$.

▶ **Definition 25** (Measureable). *Let U be a set, and let $A \subseteq U$. Let $\mathfrak{F} = \{F_1, ..., F_c\}$ be a collection of functions $F_i : U \to \mathbb{F}$. We say a function $G : U \to \mathbb{F}$ is* measurable in respect of $\mathfrak{F}$ in A*, shorthand by* $\mathfrak{F}$-measurable in A*, if there exists a function $\Gamma : \mathbb{F}^c \to \mathbb{F}$ such that:*

$$\forall a \in A : g(a) = \Gamma(F_1(a), ..., F_c(a))$$

When discussing the factor over A defined by $\mathcal{B} = \mathcal{B}_{F_1,...,F_c}$, we also say G is measurable in resepct of $\mathcal{B}$*. The function Γ will be denoted as the* measurement function *of G in respect of $\mathfrak{F}$. Additionally, when $A = U$, we sometimes omit the specification of the domain, and say G is measurable in respect of $\mathfrak{F}$.*
Note that in this paper, we usually think of $U = \mathbb{F}^n$, and A is either $\mathbb{F}^n$ or $\tilde{X} \subseteq \mathbb{F}^n$.

▶ Remark 26. If G is $\mathfrak{F}$-measurable in A, then every value of G in A can be determined by the values of $F_1, ..., F_c$. In other words, the function G is constant inside every atom of the factor defined by $\mathfrak{F}$.

▶ **Definition 27** (Syntactic Refinement). *Let $\mathcal{B}$ and $\mathcal{B}'$ be polynomial factors over $U \subseteq \mathbb{F}^n$. We say a factor $\mathcal{B}'$ is a* syntactic refinement *of the factor $\mathcal{B}$, if the collection of functions defining $\mathcal{B}$ is a subset of the set of functions defining $\mathcal{B}'$. We denote this property of $\mathcal{B}'$ by $\mathcal{B}' \succeq_{syn} \mathcal{B}$.*

We now present a standard generalized definition of refinement, where we only require the atoms induced by the refined factors are sub-atoms of those that are induced by the original factor. Note that in this refinement, we allow the refined factor to include completely different polynomials than the original factor.

▶ **Definition 28** (Semantic Refinement). *Let $\mathcal{B}$ and $\mathcal{B}'$ be polynomial factors on U defined by $\mathcal{P}$ and $\mathcal{P}'$ respectively. We say the factor $\mathcal{B}'$ is a* semantic refinement *of the factor $\mathcal{B}$ in $A \subseteq U$, if $x, y \in A$ with $\mathcal{B}'(x) = \mathcal{B}'(y)$ implies that $\mathcal{B}(x) = \mathcal{B}(y)$. We denote this property of $\mathcal{B}'$ by $\mathcal{B}' \succeq_{sem|A} \mathcal{B}$. When $A = U$, we sometimes omit A from the syntax and denote it with $\mathcal{B}' \succeq_{sem} \mathcal{B}$*
Note that $\mathcal{B}' \succeq_{syn} \mathcal{B}$ implies $\mathcal{B}' \succeq_{sem|A} \mathcal{B}$ for every $A \subseteq U$.

▶ Remark 29. A handy property of semantic refinement is that if $F : A \to \mathbb{F}$ is $\mathcal{P}$-measurable, then it is also $\mathcal{P}'$-measurable in A. Moreover, the other direction is also true: If every $\mathcal{P}$-measurable function $F : A \to \mathbb{F}$ in A is also $\mathcal{P}'$-measurable in A, then $\mathcal{B}' \succeq_{\text{sem}|A} \mathcal{B}$.

Next, we recall a lemma that was presented in [7, Theorem 4.1], that allows us, given a polynomial that is measurable by a high rank factor in $\mathbb{F}^n$, to replace the polynomials in the measurement function to any collection of polynomials with smaller or equal degree, and preserve the degree of the original polynomial. Note that we state the lemma under the constraint that $d < |\mathbb{F}|$, but it is also valid for when $d \geq |\mathbb{F}|$ with proper generalization of definitions to this case (See [7, Theorem 4.1] for the exact statement).

▶ **Lemma 30** (Preserving Degree in $\mathbb{F}^n$). *Let $d > 0$ an integer such that $d < |\mathbb{F}|$, and let $P_1, ..., P_c : \mathbb{F}^n \to \mathbb{F}$ be polynomials of degree at most d, that form a factor of rank $\geq r^{30}(\mathbb{F}, d, c)$. Assume that for $\Gamma : \mathbb{F}^c \to \mathbb{F}$, the function $\gamma : \mathbb{F}^n \to \mathbb{F}$ defined as $\gamma(a) := \Gamma(P_1(a), ..., P_c(a))$ is of $\deg(\gamma) = d'$.*
Then, for every collection of polynomials $Q_1, ..., Q_c : \mathbb{F}^n \to \mathbb{F}$ that satisfy $\deg(Q_i) \leq \deg(P_i)$, the function γ' defined as $\gamma'(a) = \Gamma(Q_1(a), ..., Q_c(a))$ is a polynomial of $\deg(\gamma') \leq d'$.

Next, we restate a useful lemma from [9, Lemma 4.17] that shows that under the conditions above, Γ is as a low-degree polynomial (with even stronger conditions). Formally, they showed:

▶ **Lemma 31** (Faithful Composition). *In the case discussed above, the structure of Γ is as follows:*

$$\Gamma(z_1,...,z_c) = \sum_{\alpha\in[p-1]^c} C_\alpha \cdot \prod_{i=1}^{c} z_i^{\alpha_i}$$

where $C_\alpha = 0$ whenever $\sum_{i=1}^{c}(\alpha_i \cdot \deg(P_i)) > d'$.
In other words, this means that Γ as a function $\Gamma : \mathbb{F}^c \to \mathbb{F}$, is a polynomial of degree $\leq d'$, even when substituting its i-th input by any polynomial of degree $\leq \deg(P_i)$.

Next, we restate the regularization process, that was first presented by [28, Lemma 2.3], with the second part of the lemma presented in [28, Lemma 9.3] (a statement the combines the two can be found [32, Lemma 7.29]).
We begin with a definition:

▶ **Definition 32.** *Let $\mathcal{P} = (P_1,...,P_c)$ be a collection of polynomials of degree $\leq d$ that defines a factor $\mathcal{B}$. Define $M(\mathcal{B}) := (M_d,...,M_1) \in \mathbb{N}^d$, where M_i denotes the number of polynomials in $\mathcal{P}$ that have degree exactly i. Thus, $\sum_{i=1}^{d} M_i = c$.*
We define the lexicographical order on $\mathbb{N}^d$ where $M > M'$ if and only if $M_i > M'_i$ for some $1 \leq i \leq d$, and $M_j = M'_j$ for all $j > i$.

The regularization process shows that every factor have a high-rank factor that semantically refines it, without increasing the size of the factor too much (its new size is independent of n).

▶ **Lemma 33** (Regularization in $\mathbb{F}^n$). *Let $r : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function and let $d \in \mathbb{N}$. There exists $C^{33}_{r,d} : \mathbb{N} \to \mathbb{N}$ such that the following holds: Let $\mathcal{B}$ be a factor on $\mathbb{F}^n$ defined by polynomials $\mathcal{P} = (P_1,...,P_c)$ where for all $i \in [c]$: $P_i : \mathbb{F}^n \to \mathbb{F}$ and $\deg(P_i) \leq d$ Then, there is an r-regular factor $\mathcal{B}'$ defined by polynomials $\mathcal{Q} = (Q_1,...,Q_{c'})$ where for all $i \in [c']$: $Q_i : \mathbb{F}^n \to \mathbb{F}$ such that $\mathcal{B}' \succeq_{sem} \mathcal{B}$, $M(\mathcal{B}') \leq M(\mathcal{B})$ and $c' \leq C^{33}_{r,d}(c)$.*
Moreover, if $\mathcal{B} \succeq_{syn} \bar{\mathcal{B}}$ for some polynomial factor $\bar{\mathcal{B}}$ with rank at least $r(c') + c' + 1$, then we can require that $\mathcal{B}' \succeq_{syn} \bar{\mathcal{B}}$.

▶ Note. Note that in the definitions above implicitly assume there are no constants in the collections of polynomials we discuss (no polynomials of degree 0). This is a valid assumption as we are interested in the set of functions that are measurable in respect to the collections, and this property is unaffected by constant polynomials in the collection. Therefore, we can always assume there are no such polynomials in any collection we consider in this context.

3 Polynomials in $\tilde{X}$

In this section we wish to generalize the definition of degree-d polynomials for functions $f : \tilde{X} \to \mathbb{F}$. Note that we wish to define it using a property of f that is intrinsic to $\tilde{X}$: given a function $f : \tilde{X} \to \mathbb{F}$, we wish be able to determine its degree only using values of $\tilde{X}$, without considering any value outside of $\tilde{X}$ (such as values of $\mathbb{F}^n \setminus \tilde{X}$).
To define such property, we generalize the local definition of a degree that is defined for polynomials in $\mathbb{F}^n$. We remind the reader that in $\mathbb{F}^n$, we said a function over $\mathbb{F}^n$ is a polynomial of degree $\leq d$ if and only if its $(d+1)$-derivative in every direction is $\equiv 0$. Thus,

in order to determine the $(d+1)$-derivative of a function in directions $y_1, ..., y_{d+1}$, one needs to evaluate the function over all the points of the cube generated by $x, y_1, ..., y_{d+1}$, which is the set of points $\left\{x + \sum_{i \in S} y_i\right\}_{S \subseteq [d+1]}$. This raises a challnge in extending this definition for functions defined over $\tilde{X} \subseteq \mathbb{F}^n$: depending on $\tilde{X}$, the function $f : \tilde{X} \to \mathbb{F}$ is not be defined to all points in all the cubes of $\mathbb{F}^n$, because some of those points do not lie in $\tilde{X}$.
Therefore, to generalize the definition of a polynomial to $\tilde{X}$, we start by giving the formal definition and notation of the set of cubes in $\tilde{X}$:

▶ **Definition 34** (Cubes). *Let $k \in \mathbb{N}$ be an integer and let $x, y_1, ..., y_k \in \mathbb{F}^n$. We define the cube $(x|y_1, ...y_k)$ as follows:*

$$(x|y_1, ..., y_k) \coloneqq \left\{ x + \sum_{i \in S} y_i \right\}_{S \subseteq [k]}$$

We refer to x as the offset *of the cube, and $y_1, ..., y_k$ as the* directions *of the cube.*
Moreover, let $\tilde{X} \subseteq \mathbb{F}^n$ be a subset. We define the set of cubes of $\tilde{X}$ of size k *as follows:*

$$C_k(\tilde{X}) \coloneqq \left\{ (x|y_1, ..., y_k) \middle| \forall S \subseteq [k] : (x + \sum_{i \in S} y_i) \in \tilde{X} \right\}$$

Using this definition, we can define a polynomial of degree $\leq d$ for subsets of $\mathbb{F}^n$:

▶ **Definition 35** (Polynomials in $\tilde{X}$). *Let $d \in \mathbb{N}$ be an integer, and let $\tilde{X} \subseteq \mathbb{F}^n$. We say the degree of a function $f : \tilde{X} \to \mathbb{F}$ is d if d is the smallest integer such that f vanishes over all cubes of size $(d+1)$, i.e:*

$$\forall (x|y_1, ..., y_{d+1}) \in C_{d+1}(\tilde{X}) : D_{y_{d+1}} ... D_{y_1} p(x) = 0$$

A function over $\tilde{X}$ of degree $\leq d$ is also called a polynomial of degree $\leq d$ over $\tilde{X}$. *We sometimes also refer to such functions as polynomials* in $\tilde{X}$*, and use the two interchangeably. We denote the set of polynomials of degree $\leq d$ over $\tilde{X}$ by $Poly_{\leq d}(\tilde{X} \to \mathbb{F})$.*

▶ Note. For $\tilde{X} = \mathbb{F}^n$, the definition above coincides with the local definition of polynomials.

3.1 Lifting Polynomials

Our goal to achieve good properties for polynomials over $\tilde{X}$. To do so, we wish to connect the desired properties of polynomials defined over $\tilde{X}$, to properties known for polynomials over $\mathbb{F}^n$. Following such strategy raises a question: given a polynomial $p : \tilde{X} \to \mathbb{F}$, which polynomial over $\mathbb{F}^n$ should we consider to deduce properties of p? To find such a polynomial over $\mathbb{F}^n$, it would have been useful that all polynomials over $\tilde{X}$ actually "came from" polynomials over $\mathbb{F}^n$. More formally, it would have been useful that all polynomials $p : \tilde{X} \to \mathbb{F}$ would be equal to a restriction of some polynomial $P : \mathbb{F}^n \to \mathbb{F}$ of degree $\leq d$, to the set $\tilde{X}$. This would give us a "good candidate" (or candidates) to polynomials over $\mathbb{F}^n$, that using their known properties, we could achieve the properties we desire for polynomials over $\tilde{X}$.
Generally speaking, the existence of such polynomial $P : \mathbb{F}^n \to \mathbb{F}$ is not trivial by itself, and it mapy depend on the polynomial p and the set $\tilde{X}$. In this subsection, we discuss sets $\tilde{X} \subseteq \mathbb{F}^n$ that have this property for every polynomial $p : \tilde{X} \to \mathbb{F}$. Before formulating the notion above, we start by a simple remark:

▶ Remark 36. By the local criteria for $\mathbb{F}^n$, we have that a restriction of a polynomial of degree $\leq d$ over $\mathbb{F}^n$ to $\tilde{X}$ is a polynomial of degree $\leq d$ over $\tilde{X}$. Therefore, the other direction is true: every restriction of a polynomial over $\mathbb{F}^n$ to $\tilde{X}$ is a polynomial over $\tilde{X}$.

Next, let us define subsets $\tilde{X} \subseteq \mathbb{F}^n$ that have the desired property, which we call *d-lift-enabler variety.*

▶ **Definition 37** (d-lift-enabler Subset)**.** *Let $\mathbb{F}$ be a field, and $n > 0$ be an integer. For an integer $d > 0$, we say a subset $\tilde{X} \subseteq \mathbb{F}^n$ is* d-lift-enabler *if for every $d' \leq d$, for every polynomial $p \in Poly_{d'}(\tilde{X} \to \mathbb{F})$ there exist a polynomial $\widehat{p} \in Poly_{d'}(\mathbb{F}^n \to \mathbb{F})$ such that $p|_{\tilde{X}} = \widehat{p}|_{\tilde{X}}$.*

▶ Remark 38. Using the local criterion of polynomials and the fact that that $C_{d+1}(\tilde{X}) \subseteq C_{d+1}(\mathbb{F}^n)$, one can see that for a polynomial $p : \tilde{X} \to \mathbb{F}$ with $\deg(p) = d$, every extension $P : \mathbb{F}^n \to \mathbb{F}$ with $p = P|_{\tilde{X}}$ holds the bound $\deg(\widehat{p}) \geq d$. The other direction is not true in the general case, but it is specifically promised when the variety is d-lift-enabler.

This definition naturally raises the following definition:

▶ **Definition 39** (The Lift Operator)**.** *Let $d \in \mathbb{N}$ be an integer. Let $\tilde{X} \subseteq \mathbb{F}^n$ be a d-lift-enabler subset. We define* the d-lift operator *to be an operator $\widehat{\square} : Poly_{\leq d}(\tilde{X} \to \mathbb{F}) \to Poly_{\leq d}(\mathbb{F}^n \to \mathbb{F})$ the following way:*
Let $d' \leq d$. Given a polynomial $p : \tilde{X} \to \mathbb{F}$ of degree d', the operator $\widehat{\square}$ returns a polynomial $\widehat{p} : \mathbb{F}^n \to \mathbb{F}$ of degree d' such that $p = \widehat{p}|_{\tilde{X}}$. Note that we did not require the lift to be unique. Thus, in case there are multiple valid lifts for a polynomial $p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$, the lift operator picks a single (consistent) one of them. Moreover, the lift always exists because the subset $\tilde{X}$ is d-lift-enabler.
In addition, for a collection $\mathfrak{p} = (p_1, ..., p_c)$ of polynomials $p_i \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$, we denote $\widehat{\mathfrak{p}} \coloneqq (\widehat{p_1}, ..., \widehat{p_c})$

In the following subsections, we give example to two concrete sets $\tilde{X} \subseteq \mathbb{F}^n$ that are d-lift-enablers. Before doing so, we define an algebraic variety:

▶ **Definition 40** (Algebraic Variety)**.** *For a collection of functions $\mathfrak{F} \coloneqq \{F_1, ...F_c\}$ such that $F_i : \mathbb{F}^n \to \mathbb{F}$, we denote $Z(\mathfrak{F}) \coloneqq \{x \in \mathbb{F}^n | \forall i : F_i(x) = 0\}$.*
If the collection is a collection of polynomials, we call $Z(\mathfrak{F})$ an algebraic variety*, shorthand by* variety*.*
The degree of the variety which is a complete intersection is the product of the degrees of the polynomials in the collection that defines it.

3.2 High Rank Varieties of High Minimal Degree

We now present a theorem proved in [34, Corollary 1.10], that shows that high rank varieties are d-lift-enabler when the polynomials defining the variety are of degree $> d$:

▶ **Theorem 41.** *Let $\mathbb{F}$ be a finite field, and let $\tilde{d}, \tilde{c} > 0$ representing parameters of a variety. Let $d < \tilde{d}$ a positive integer representing a degree of a polynomial which we wish to lift. There exists $\bar{r} = \bar{r}(\mathbb{F}, \tilde{d}, \tilde{c}) > 0$ such that for all $n \in \mathbb{N}$, any variety $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right) \subseteq \mathbb{F}^n$ for $\tilde{\mathcal{L}} = (L_1, ..., L_{\tilde{c}})$ which is a complete intersection such that rank $\left(\tilde{\mathcal{L}}\right) > \bar{r}$ [29] , degree $\deg(\tilde{\mathcal{L}}) = \tilde{d}$, with all defining polynomials of degree $\deg(L_i) > d$, it holds that $\tilde{X}$ is a d-lift-enabler subset.*

▶ Remark 42. Under the conditions stated above, it was proved in [34] that the lift is in fact *unique*. Formally, if $p : \tilde{X} \to \mathbb{F}$ is a polynomial of degree $\leq d$, then there exists a *unique* polynomial $P : \mathbb{F}^n \to \mathbb{F}$ such that $P|_{\tilde{X}} \equiv p$.

[29] The definition of rank used in thier proof is slightly different than our definition of rank. This is addressed in Appendix B.

3.3 High Rank Varieties on a Large Field

In this subsection, we recall a theorem proved by [35, Theorem 1.7] regarding high rank varieties that are defined on "large" fields. We note that the fields are large in respect of the degree d one wish to lift, but still does not depend on n.

Next we define a weakly polynomial, which generalizes our definition of a polynomial in $\tilde{X}$, that was used in [35, Definition 1.1]:

▶ **Definition 43.** *Let $\tilde{X} \subseteq \mathbb{F}^n$ be a set. We say a function $F : \tilde{X} \to \mathbb{F}$ is a* weakly polynomial of degree $\leq d$ *if for any affine subspace $L \subseteq \tilde{X}$, the restriction $F|_L$ is a polynomial of degree $\leq d$.*

▶ Remark 44. By the local criteria of a polynomial, it is easy to see that every $P \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ is a weakly polynomial of degree $\leq d$.

And now, we can present the lifting theorem for large fields, as proved in [36, Theorem 2.17].

▶ **Theorem 45** ([36, Theorem 2.17])**.** *Let $d, \tilde{d} \in \mathbb{N}$, and let $\mathbb{F}$ be a finite field such that $|\mathbb{F}| > d \cdot \tilde{d}$. There exists $r_{45} = r_{45}(\tilde{d}, d)$ such that for any variety $\tilde{X} \subseteq \mathbb{F}^n$ of degree $\leq \tilde{d}$ which is a complete intersection, defined by a collection of polynomials with rank [30] $\geq r_{45}$, have the following property: Every weakly polynomial function $p : \tilde{X} \to \mathbb{F}$ of degree $\leq d$ can be lifted to a polynomial function $P : \mathbb{F}^n \to \mathbb{F}$ of degree $\leq d$.*

▶ Note. Note that we stated the theorem above to finite fields, but it is also valid for infinite algebraically closed fields.

The theorem above implies the following corollary:

▶ **Corollary 46.** *Let $d, \tilde{d} \in \mathbb{N}$, and let $\mathbb{F}$ be a finite field such that $|\mathbb{F}| > d \cdot \tilde{d}$. There exists $r_{45} = r_{45}(\tilde{d}, d)$ such that for any variety $\tilde{X} \subseteq \mathbb{F}^n$ of degree $\tilde{d}$ and rank $\geq r_{45}$ is a d-lift-enabler.*

4 Relative Rank-Bias Property

In this section, we generalize the relation between rank and bias that is known for $\mathbb{F}^n$ also for $\tilde{X} \subseteq \mathbb{F}^n$. Specifically, in Theorem 23, it was shown that high rank factors have low bias in $\mathbb{F}^n$. We wish to define an alternative definition of rank for $\tilde{X} \subseteq \mathbb{F}^n$, called *$\tilde{X}$-relative rank*, such that high $\tilde{X}$-relative rank implies low bias in $\tilde{X}$. This type of relation (and definition) was shown previously to a few sets; in [37, Theorem 1.8] for sets $\tilde{X} = Z(\mathcal{Q})$ where $\mathcal{Q}$ is a collection of polynomials of high rank; and in [27, Theorem 1.4] for sets $\tilde{X} = S^n$ for $S \subset \mathbb{F}$.

To understand this notion, we first introduce a simple example that demonstrates the need for a different definition of rank to achieve equidistribution properties in subsets of $\mathbb{F}^n$.

▶ **Example 47.** Let $\tilde{X} = \{x \in \mathbb{F}^n | x_1 = 0\}$. Define $P : \mathbb{F}^n \to \mathbb{F}$ by $P(x) \coloneqq x_1$.
In $\mathbb{F}^n$, P has rank ∞ as it can not be decomposed polynomials of degree < 1 (constants). Additionally, it is perfectly equidistributed. This is the simplest example of the rank-bias relation in $\mathbb{F}^n$.
However, when restricting P to $\tilde{X}$, we get $P|_{\tilde{X}} \equiv 0$. As 0 is a constant function, it is the *least* equidistributed possible in $\tilde{X}$. Therefore, we see that the way we defined rank in $\mathbb{F}^n$ does not imply the desired equidistribution in $\tilde{X}$: we found a polynomial with high rank (infinity) that has a very high bias *in* $\tilde{X}$ (the maximal).

[30] The definition of rank used in thier proof is slightly different than our definition of rank. This is addressed in Appendix B.

The reason the definition of rank in $\mathbb{F}^n$ fails to capture equidistribution even on subsets that are really similar to $\mathbb{F}^n$ (isomorphic to $\mathbb{F}^k$), is because of the following reason: Even though our polynomial P does *not* have a decomposition to a few lower-degree polynomials by itself, there *exists* a $\tilde{X}$*-equivalent* polynomial that has such structured decomposition. Here, by $\tilde{X}$-equivalent we mean a polynomial in $\mathbb{F}^n$ that is bounded by the same degree bound, and is equal to P in $\tilde{X}$. In the example described above, this equivalent polynomial is the constant function 0, and its decomposition is the trivial one (any function decomposes a constant function). An alternative perspective which we use throughout this paper to $\tilde{X}$-equivalence is that both polynomials are equal up to a *valid* $\tilde{X}$*-remainder*: a bounded degree polynomial that is $\equiv 0$ in $\tilde{X}$.

Generally speaking, high $\tilde{X}$-relative rank may not imply low bias in $\tilde{X}$. Therefore, this structure-randomness relation is not true for a general subset $\tilde{X} \subseteq \mathbb{F}^n$, but is a *property* of the subset $\tilde{X}$. Thus, we say that a subset has the *relative rank-bias property* if this relation holds, i.e. if high $\tilde{X}$-relative rank implies equidistribution in $\tilde{X}$.

Let us now formally define our definition for relative rank, inspired by the two different definitions of relative rank presented in [37, Definition 1.6] and in [27, Definition 1.3]:

▶ **Definition 48** (Relative Rank of a Polynomial)**.** *Let $\tilde{X} \subseteq \mathbb{F}^n$ and let $d \in \mathbb{N}$. For an integer $d \in \mathbb{N}$ and a polynomial $P : \mathbb{F}^n \to \mathbb{F}$, we define its* d-relative rank *in respect of $\tilde{X}$ as:*

$$rank_{d,\tilde{X}}(P) \coloneqq \min\left\{rank_d\left(P - \overline{P}\right) \middle| \overline{P} \in Poly_{\leq \deg(P)}(\mathbb{F}^n \to \mathbb{F}), \overline{P}|_{\tilde{X}} \equiv 0\right\}$$

For a polynomial P of degree $\deg(P) = d$ we denote $rank_{\tilde{X}}(P) \coloneqq rank_{d,\tilde{X}}(P)$.

▶ **Definition 49** ($\tilde{X}$-equivalent and $\tilde{X}$-remainder)**.** *Moreover, we say a polynomial is $\tilde{X}$-*equivalent *to P if its restriction to $\tilde{X}$ is $\equiv P|_{\tilde{X}}$. We say it is* valid $\tilde{X}$-equivalent to P *if it is $\tilde{X}$-equivalent to P and it is of the* same *degree of P.*
*Similarly, we say a polynomial is $\tilde{X}$-*remainder *if its restriction to $\tilde{X}$ is $\equiv 0$. We say it is* valid $\tilde{X}$-remainder P *if it is $\tilde{X}$-remainder and it is of degree* smaller or equal *of the degree of P.*
We typically denote such polynomial as $\overline{P}$.

In other words, the d-relative rank of a polynomial P is the smallest d-rank of all valid $\tilde{X}$-equivalents of P.

▶ Note 50. Note that [27, Definition 1.3] defines rank in a substantially different way than our definition, and consequentially our results will not apply to the sets they presented. One of the main differences in the definition of rank occurs for $d = 1$. In the definition we use for rank, the rank of every (non-constant) degree-1 polynomial is ∞, where in the definition used in [27] it is a finite number (which is possibly very small). This difference is crucial, as for example, it makes regularization according to their definition not-trivially possible, where it is known to be possible when rank is defined by the definition we use (Lemma 33).

▶ **Definition 51** (Relative Rank of a Factor)**.** *Let $\tilde{X} \subseteq \mathbb{F}^n$. Let $\mathcal{P}$ be a set of polynomials $\mathcal{P} = \{P_1, ..., P_c\}$. The rank of the polynomial set $\mathcal{P}$ relative to the subset $\tilde{X}$ is defined as:*

$$rank_{\tilde{X}}(\mathcal{P}) \coloneqq \min\left\{rank_{d,\tilde{X}}\left(\sum_{i=1}^{c} \lambda_i P_i\right) \middle| 0 \neq \vec{\lambda} \in \mathbb{F}^c, d = \max_{i\in[c]} \deg(\lambda_i P_i)\right\}$$

For a factor $\mathcal{B}$ defined by a collection of polynomials, we define its relative rank relative to $\tilde{X}$ to be the relative rank of the collection of polynomials defining it, relative to the set $\tilde{X}$.
For a non-decreasing function $r : \mathbb{N} \to \mathbb{N}$, a factor $\mathcal{B}$ is called r-$\tilde{X}$-regular if its relative rank in respect to $\tilde{X}$ is at least $r(|\mathcal{B}|)$.

4.1 Relative Rank-Bias Property

▶ **Definition 52** (Relative Rank-Bias property). *Let $\mathbb{F}$ be a finite field, and let $d \in \mathbb{N}$ be an integer. Let $\tilde{r} : \mathbb{R}^+ \to \mathbb{N}$ be a function that represents the rank-bias relation for a fixed $d, \mathbb{F}$. We say a set $\tilde{X} \subseteq \mathbb{F}^n$ has the* $(\tilde{r}, \mathbb{F}, d)$-relative rank-bias property *if for every $\epsilon > 0$, for every polynomial P of degree $\leq d$ with $rank_{\tilde{X}}(P) \geq \tilde{r}(\epsilon)$ we have:*

$$bias_{\tilde{X}}(P) < \epsilon$$

As an immediate corollary of Theorem 23 that shows that high rank implies low bias, we have that $\tilde{X} = \mathbb{F}^n$ has the relative rank-bias property.

▶ **Corollary 53** ($\mathbb{F}^n$ has the relative rank-bias property). *For every finite field $\mathbb{F}$ and $d \in \mathbb{N}$, let $\tilde{r} : \mathbb{R}^+ \to \mathbb{N}$ defined as $\tilde{r}(\epsilon) \coloneqq r_{23}(\mathbb{F}, d, \epsilon)$. Then, we have that the set $\tilde{X} = \mathbb{F}^n$ has the $(\tilde{r}, \mathbb{F}, d)$-relative rank-bias property.*

Proof. This is a simple usage of Theorem 23: Note that when $\tilde{X} = \mathbb{F}^n$, we have that $rank_{\tilde{X}}(P) = rank(P)$. Now, if P is a polynomial of degree $\leq d$ and $rank_{\tilde{X}}(P) = rank(P) \geq \tilde{r}(\epsilon) = \tilde{r}_{23}(\mathbb{F}, d, \epsilon)$, then:

$$bias_{x \in \mathbb{F}^n}(P(x)) < \epsilon$$ ◀

4.2 Limited-Relative Rank-Bias Property

Sometimes, however, we can not request $\tilde{X}$ to be such that high relative rank implies low bias for every $\epsilon > 0$, but only for $\epsilon' \geq \epsilon$ for some constant $\epsilon > 0$. This leads to defining the *limited relative rank-bias property*, which will be used to discuss such sets $\tilde{X} \subseteq \mathbb{F}^n$.
As we will later see, this definition raises naturally where $\tilde{X}$ is a high rank variety, in which for the relative rank-bias property to hold for some $\epsilon > 0$, the rank of the variety should be greater than a value that is dependent of ϵ. Thus, to have the relative rank-bias property for a high rank variety but without requiring an infinitely large rank, we must limit the relative rank-bias property for $\epsilon' \geq \epsilon$ We formulate the definition of this property as follows:

▶ **Definition 54** (Limited Relative Rank-bias property). *Let $\mathbb{F}$ be a finite field, let $d \in \mathbb{N}$ be an integer, and let $\epsilon > 0$ be a constant. Let $\tilde{r} : [\epsilon, \infty] \to \mathbb{N}$ be a function that represents the limited-relative-rank-bias relation.*
We say a set $\tilde{X} \subseteq \mathbb{F}^n$ has the $(\tilde{r}, \mathbb{F}, d, \epsilon)$-limited-relative-rank-bias property *if for every $\epsilon' \geq \epsilon$, for every polynomial P of degree $\leq d$ with $rank_{\tilde{X}}(P) \geq \tilde{r}(\epsilon')$ we have:*

$$bias_{\tilde{X}}(P) < \epsilon'$$

As a convention, we denote by $\tilde{\epsilon}$ the ϵ such that the limited-relative-rank-bias property holds for $\tilde{X}$.

4.2.1 High Rank Varieties

In this subsection, we are discussing specifically $\tilde{X} \subseteq \mathbb{F}^n$ that are in the form $\tilde{X} = Z(\mathcal{Q})$ for a set of polynomials $\mathcal{Q}$ that form a high rank factor. Let us present some known results of the relative rank-bias relation for high rank varieites: In the scenario when we are working relative to $\tilde{X}$, the equivalent for Theorem 23 is also known when we assume $d < char(\mathbb{F})$, as shown in [37, Theorem 1.8]:

▶ **Theorem 55** (High relative rank implies low bias in high rank varieties)**.** *Let $\mathbb{F}$ be a finite field and let $0 \leq d < char(\mathbb{F})$. Let $\epsilon > 0$ be a constant, and let $\tilde{c} \in \mathbb{N}$. There exist $\bar{r}^{55} = \bar{r}^{55}(\mathbb{F}, d, \tilde{c}, \epsilon)$ and $r^{55} = r^{55}(\mathbb{F}, d, \epsilon)$ such that the following holds:*
Let $\tilde{\mathcal{L}} = (L_1, ..., L_{\tilde{c}})$ be a collection of polynomials of degrees $\leq d$ with rank $\left(\tilde{\mathcal{L}}\right) \geq \bar{r}^{55}$ and let P be a polynomial of degree $\leq d$.
Then, if $rank_{\tilde{X}}(P) \geq r^{55}$, we have:

$$bias_{\tilde{X}}(P) < \epsilon$$

▶ Note. Note that the original statements in [37] are stated for a different definition of rank, noted as *schmidt rank*. In the appendix B we compare the two different definitions, and show that our definition of rank is comprehensive enough in a sense that a polynomial with high rank also has high schmidt rank. Additionally, we show that for a given $r \in \mathbb{N}$, the lower bound of rank required for a polynomial to be of schmidt rank $\geq r$, is only $c \cdot r$ for some constant $c \in \mathbb{N}$.

▶ Remark 56. Note that in the original statement of theorem 55 as stated in [37, Theorem 1.8], there are good bounds on the rank needed for $\tilde{\mathcal{L}}$ and P for the theorem to hold.
Specifically, there exist constants $A(d), B(d)$ such that for an error $\epsilon = |\mathbb{F}^{-s}|$, if $\bar{r}^{55} = A(\tilde{c} + s)^B$ and $r^{55} = A(1 + s)^B$, then we have:

$$bias_{Z\left(\tilde{\mathcal{L}}\right)}(P) < |\mathbb{F}|^{-s}$$

In our proof, it is enough that the bounds on r and $\bar{r}$ are independent of n, thus we omit the exact bounds stated above and use the statement as stated in Theroem 55.

▶ Remark 57. Note that both $r^{55}(\mathbb{F}, d, \epsilon)$ and $\bar{r}^{55}(\mathbb{F}, d, \tilde{c}, \epsilon)$ are decreasing when ϵ is increasing. This means for example, that for all $\epsilon' \geq \epsilon$, a variety that satisfies the theorem's rank condition for ϵ also satisfies the theorem's rank condition for ϵ'. Therefore, a polynomial with rank $\geq r^{55}(\mathbb{F}, d, \epsilon)$ will have a bias $< \epsilon'$.

As a corollary of Theorem 55 and Remark 88, we have that high rank varieties has the limited-relative-rank-bias property. Formally, we have:

▶ **Corollary 58** (High Rank Varieties Have the Limited-Relative Rank-Bias Property)**.** *Let $\mathbb{F}$ be a finite field, and let $\tilde{d} \in \mathbb{N}$ such that $0 < \tilde{d} < |\mathbb{F}|$. Let $\tilde{\epsilon} > 0$ be a constant which represents the desired relative rank-bias limit. There exists $\tilde{r}_{58} : [\tilde{\epsilon}, \infty] \to \mathbb{N}$ with $\tilde{r}_{58} \coloneqq \tilde{r}_{58}(\mathbb{F}, \tilde{d})$ such that the following holds:*
Let $\tilde{c} \in \mathbb{N}$ be an integer. There exists $\bar{r}_{58} \coloneqq \bar{r}_{58}(\mathbb{F}, \tilde{d}, \tilde{c}, \tilde{\epsilon})$ such that for every $\tilde{\mathcal{L}} = (L_1, ..., L_{\tilde{c}})$ collection of polynomials with rank $\left(\tilde{\mathcal{L}}\right) \geq \bar{r}$, defining a variety $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right)$ of degree $\leq \tilde{d}$ which is a complete intersection, we have:
The variety $\tilde{X}$ has the $(\tilde{r}_{58}, \mathbb{F}, \tilde{d}, \tilde{\epsilon})$-limited-relative-rank-bias property.

Proof. Let $\mathbb{F}$ be a finite field, and let $\tilde{d} \in \mathbb{N}$ such that $0 < \tilde{d} < |\mathbb{F}|$. Let $\tilde{\epsilon} > 0$. We choose:

$$\tilde{r}_{58}(\epsilon) \coloneqq r_{55}(\mathbb{F}, \tilde{d}, \epsilon)$$

Note that for every ϵ in its domain, $\tilde{r}_{58}$ does not depend on $\tilde{\epsilon}$. Let $\tilde{c} \in \mathbb{N}$. Now, we choose:

$$\bar{r}_{58}(\mathbb{F}, \tilde{d}, \tilde{c}, \tilde{\epsilon}) \coloneqq \bar{r}_{55}(\mathbb{F}, \tilde{d}, \tilde{c}, \tilde{\epsilon})$$

Using Theorem 55 that shows high rank implies low bias in $\tilde{X}$, and the assumption that $\epsilon \geq \tilde{\epsilon}$ (specifically Remark 57) concludes the proof. ◀

5 Regularization Relative to $\tilde{X}$

In this section, we generalize the definitions and statements regarding factors and regularization in $\mathbb{F}^n$, to their corresponding definitions and statements to relative rank in respect of $\tilde{X} \subseteq \mathbb{F}^n$.
Note that in oppose to the previous chapter that we discussed a general U and $A \subseteq U$, in this chapter we discuss only $U = A = \mathbb{F}^n$. This is done for clearance and to avoid defining definitions we will not use in our main proof.

▶ **Definition 59** (Measurable Relative to $\tilde{X}$). *Let $\mathfrak{F} = \{F_1, ..., F_c\}$ be a set of functions $F_i : \mathbb{F}^n \to \mathbb{F}$. We say a function $G : \mathbb{F}^n \to \mathbb{F}$ is* measurable in respect of $\mathfrak{F}$ relative to $\tilde{X}$*, or* $\tilde{X}$-relative $\mathfrak{F}$-measurable*, if there exists a function $\overline{G} : \mathbb{F}^n \to \mathbb{F}$ with $\overline{G}|_{\tilde{X}} \equiv 0$ and a function $\Gamma : \mathbb{F}^c \to \mathbb{F}$ such that:*

$$\forall a \in \mathbb{F}^n : G(a) = \Gamma(F_1(a), ..., F_c(a)) + \overline{G}(a)$$

And:

$$\deg(G - \overline{G}) \leq \deg(G)$$

We sometimes refer to Γ as the $\tilde{X}$-relative measurement function.

▶ Note. Note that if $\deg(G - \overline{G}) \leq \deg(G)$ as discussed above, then the same bound also bounds the degree of the remainder, i.e. $\deg(\overline{G}) \leq \deg(G)$. Therefore $\overline{G}$ is a valid $\tilde{X}$-remainder of G. Moreover, this requirement is equivalent to the definition above, as if $\deg(\overline{G}) \leq \deg(G)$, then we also have $\deg(G - \overline{G}) \leq \deg(G)$.

▶ Note. Also note that without the bound on the degree of the remainder, being measurable relative to $\tilde{X}$ is in fact equivalent for being a measurable in $A = \tilde{X}$. This is true because under these conditions, the remainder $\overline{G}$ has no constraints but $\overline{G}|_{\tilde{X}} \equiv 0$, thus the condition left on the measurement is just being a measurement to G in $\tilde{X}$.

▶ Remark 60. If G is a function that it is $\mathfrak{F}$-measurable relative to $\tilde{X}$, then every value of G can be determined by the values of $F_1, ..., F_c$ up to a remainder $\overline{G}$ of degree $\leq d$. Thus, perhaps we do not know that the function G is constant inside every atom of $\mathfrak{F}$ as in a regular semantic refinement, but we do know that there exists a function $(G - \overline{G})$ that equals to G on $\tilde{X}$, is constant on every atom of $\mathfrak{F}$ and it is a function with a bounded degree i.e. $\deg(G - \overline{G}) \leq \deg(G)$.

Next, we present a new type of refinement, which is a relaxation of semantic refinement. This relaxation will allow us to discuss the corresponding claim of the polynomial regularity lemma (Lemma 33) for relative rank (instead of rank).

▶ **Definition 61** (Semantic Refinement Relative to $\tilde{X}$). *Let $\mathcal{B}$ and $\mathcal{B}'$ be polynomial factors on $\mathbb{F}^n$, defined by sets of polynomials $\mathcal{P}, \mathcal{P}'$ respectively, and let $d \in \mathbb{N}$. We say a factor $\mathcal{B}'$ is a* semantic refinement relative to $\tilde{X}$ *of the factor $\mathcal{B}$, or* $\tilde{X}$-relative semantic refinement*, if the following holds: Every function $F : \mathbb{F}^n \to \mathbb{F}$ that is $\mathcal{P}$-measurable relative to $\tilde{X}$, is also $\mathcal{P}'$-measurable relative to $\tilde{X}$. If the definition above holds, we denote $\mathcal{B}' \succeq^{\tilde{X}}_{sem} \mathcal{B}$.*

▶ Note. It is easy to see that this relation is transitive, i.e. if $\mathcal{B}' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}$ and $\mathcal{B}'' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}'$, then $\mathcal{B}'' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}$.

▶ Remark 62. In $\tilde{X}$, semantic refinements relative to $\tilde{X}$ behave the same as regular semantic refinements in the perspective of being measurable: every function that is $\mathcal{P}$-measurable in $\tilde{X}$ is also $\mathcal{P}'$-measurable in $\tilde{X}$. However, the two definitions behave differently in the

perspective of being measurable in $\mathbb{F}^n$. Specifically, in relative semantic refinements, if G is a $\mathcal{P}$-measurable function it is not necessarily $\mathcal{P}'$-measurable. However, it is measurable up to a remainder $\overline{G}$ of degree $\leq \deg(G)$ such that $\overline{G}|_{\tilde{X}} \equiv 0$.

▶ **Corollary 63.** *If $\mathcal{B}' \succeq^{\tilde{X}}_{sem} \mathcal{B}$, then in $\tilde{X}$ it is a regular semantic refinement, i.e. $\mathcal{B}' \succeq_{sem|\tilde{X}} \mathcal{B}$.*

Next, we present a new regularization process that allows us to increase the *relative* rank of a factor without increasing the size of the factor too much (independent of n). This regularization process generalizes the regularization process in $\mathbb{F}^n$, that we stated in Lemma 33. We call this type of regularization process a *relative-regularization process* relative to $\tilde{X}$, shorthand by $\tilde{X}$-regularization For a specific function r, we will sometimes call applying this lemma a *r-$\tilde{X}$-regularization*. Note that to allow such a relative-regularization process to hold, we must use the relaxed definition of semantic refinement that is presented above.

▶ **Theorem 64.** *Let $r : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function and let $d \in \mathbb{N}$. There exists $C^{64}_{r,d} : \mathbb{N} \to \mathbb{N}$ such that the following holds: Let $\mathcal{B}$ be a factor defined by polynomials $\mathcal{P} = (P_1, ..., P_c)$ where for all $i \in [c]$: $P_i : \mathbb{F}^n \to \mathbb{F}$ and $\deg(P_i) \leq d$. Then, there is an r-$\tilde{X}$-regular factor $\mathcal{B}'$ defined by polynomials $\mathcal{P}' = (P'_1, ..., P'_{c'})$ where for all $i \in [c]$: $P'_i : \mathbb{F}^n \to \mathbb{F}$ and $\deg(P'_i) \leq d$ such that $\mathcal{B}' \succeq^{\tilde{X}}_{sem} \mathcal{B}$ and $c' \leq C^{64}_{r,d}(c)$.*
Moreover, if $\mathcal{B} \succeq_{syn} \bar{\mathcal{B}}$ for some polynomial factor $\bar{\mathcal{B}}$ with relative rank of at least $r(c') + c' + 1$ and rank of at least $r_{30}(\mathbb{F}, d, c') + c' + 1$, then we can require that $\mathcal{B}' \succeq_{syn} \bar{\mathcal{B}}$.

Proof. We follow the lines of the proof given by [32][Lemma 7.29], but here, we wish to increase the *relative* rank of the factor instead of its rank. We present an iterative process, which will eventually lead us to a factor of size c' with relative rank higher than $r(c')$, that is a semantic refinement relative to $\tilde{X}$. Let $d \in \mathbb{N}$, and let $\mathcal{B}$ be a polynomial factor defined by $\mathcal{P} = (P_1, ..., P_c)$ such that $P_i : \mathbb{F}^n \to \mathbb{F}$ of degree $\leq d$. We remind the reader definition 32, where we defined $M(\mathcal{B}) \coloneqq (M_d, ..., M_1) \in \mathbb{N}^d$, where M_i denotes the number of polynomials in $\mathcal{P}$ that have degree exactly i, and the lexicographical order on $\mathbb{N}^d$ where $M > M'$ if and only if $M_i > M'_i$ for some $1 \leq i \leq d$, and $M_j = M'_j$ for all $j > i$. This proof will be by transfinite induction on M under the lexicographical order. Next we describe a step of the regularization process.
Let $\mathcal{B}$ be a polynomial factor defined by $\mathcal{P} = (P_1, ..., P_c)$. Note that this is an abuse of notations: the factor $\mathcal{B}$ and the set $\mathcal{P}$ refer to the original factor in the first step, and also to the current factor in the middle of the relative-regularization process. If $\mathcal{B}$ is r-$\tilde{X}$-regular, then we are done. Otherwise, we change $\mathcal{B}$ as follows: First, we denote $r^{\mathbb{F},d}_{30}(c) \coloneqq r_{30}(\mathbb{F}, d, c)$, and we r_{30}-regularize $\mathcal{P}$ using lemma 33 to get a set of polynomials $\mathcal{P}_1 = (P^1_1, ..., P^1_{c_1})$ of degree $\leq d$, which defines a factor $\mathcal{B}_1$ and has a rank $\geq r^{\mathbb{F},d}_{30}(c_1)$. Note that $M(\mathcal{B}_1) \leq M(\mathcal{B})$. Then, again, if somehow $\mathcal{B}_1$ is now r-$\tilde{X}$-regular, we are done.
Otherwise, by definition, there exists some linear combination of the polynomials in $\mathcal{P}_1$ that has $d^\star$-relative rank less than $r(c_1)$, where $d^\star$ is the maximal degree that participates in the linear combination. Let $\vec{P}(x) = \sum_{i=0}^{c_1} \lambda_i P^1_i(x)$ where $\vec{0} \neq \vec{\lambda} \in \mathbb{F}^{c_1}$, be the linear combination with $rank_{d^\star,\tilde{X}}\left(\vec{P}\right) \leq r(c_1)$ where $d^\star \coloneqq \max_{i\in[c_1]} \deg(\lambda_i P^1_i)$. By definition of relative rank, there exists $\overline{P} \in Poly_{\leq \deg(\vec{P})}(\mathbb{F}^n \to \mathbb{F})$ with $\overline{P}|_{\tilde{X}} \equiv 0$ such that $rank_{d^\star}\left(\vec{P} - \overline{P}\right) \leq r(c_1)$. Note that $\deg(\overline{P}) \leq d^\star$. By definition of $d^\star$-rank, we have that we can decompose $\vec{P} - \overline{P}$ as a function of $r(c_1)$ polynomials of degree $\leq d^\star - 1$. In other words, there exist a measurement function $\vec{\Gamma} : \mathbb{F}^{r(c_1)} \to \mathbb{F}$ and polynomials $Q_1, ..., Q_{r(c_1)}$ with $\deg(Q_i) \leq d^\star - 1$ such that:

$$\forall a \in \mathbb{F}^n : \vec{P}(a) - \overline{P}(a) = \vec{\Gamma}\left(Q_1(a), ..., Q_{r(c_1)}(a)\right)$$

Now, let $\mathcal{P}^\star \subseteq \mathcal{P}_1$ be the set of all such maximal-degree polynomials, and let $i^\star$ be chosen such that $P^1_{i^\star} \in \mathcal{P}^\star$. Note that the set $\mathcal{P}^\star$ is non empty, as by definition, $d^\star$ is the maximal degree of polynomial in the expression $\sum_{i=1}^{c_1} \lambda_i P^1_i$ such that $\lambda_i \neq 0$.
For the next step, define the polynomial factor $\mathcal{B}_2$ be the polynomial factor defined by the set:

$$\mathcal{P}_2 \coloneqq \mathcal{P}_1 \setminus \left\{P^1_{i^\star}\right\} \cup \left\{Q_1, ..., Q_{r(c_1)}\right\}$$

Finally, the factor $\mathcal{B}_2$ will be the factor returned from the relative-regularization step.
It is easy to see that if the process above halts, we get a r-$\tilde{X}$-regular factor. Now, we prove the first part of the lemma by showing the following claims:

▷ Claim 65. The factor generated from the regularization above is of bounded size: a bound that may depend on r, d, c, but does not depend on n. Formally, we claim that there exists $C^{64}_{r,d} : \mathbb{N} \to \mathbb{N}$ such that we have $c' \leq C^{64}_{r,d}(c)$.

Proof. It is enough to prove the following:

1. In each step, the amount of polynomials there are in $\mathcal{P}_1, \mathcal{P}_2$ are bounded by a bound that depend only on r, d, c (independent of n).
2. The number of steps of the relative-regularization process is also bounded by a bound that depends only on r, d, c (independent of n).

The combination of these two will obtain the desired bound of the amount of polynomials in the last-step regularized factor, which is $C^{64}_{r,d}(c)$. Note that the bound on the last-step relative-regularized factor in not simply the multiplication of the two bounds, but a recursively-substitution of the bound in 1, a bounded amount of times (bounded by the bound in 2).
For 1, we first notice that the number of polynomials in the regular regularization process is bounded, specifically we have $\left|\mathcal{P}^1\right| = c_1 \leq C^{33}_{r_{30},d}(c)$. Moreover, the polynomial factor $\mathcal{B}_2$ is generated by adding at most $r(c_1)$ polynomials to the factor, and thus we have $|\mathcal{P}_2| \leq c_1 + r(c_1)$ which is also bounded by substituting the bound on c_1.
For 2, we use the transfinite induction on M we mentioned earlier to show that the process must halt after a bounded number of steps. Formally, we show that there exist M' which depends only on $M(\mathcal{B})$ such that $M(\mathcal{B}_2) \leq M' < M(\mathcal{B})$. This will bound the number of steps by a value that depend only on $M(\mathcal{B})$, which depends only on r, d, c. To do so, we first notice that the regular regularization does not increase the value of M, i.e. $M(\mathcal{B}_1) \leq M(\mathcal{B})$. Thus, we can focus on the second part of the relative-regularization. In this part, we replace a single degree $d^\star$ polynomial by at most $r(c_1)$ polynomials of degree $\leq d^\star - 1$. Therefore, by choosing $M' \coloneqq (M_d, ..., M_{d^\star+1}, M_{d^\star} - 1, M_{d^\star-1} + r(c_1), ..., M_1 + r(c_1))$ we get that $M(\mathcal{B}_2) \leq M' < M(\mathcal{B}_1) \leq M(\mathcal{B})$, which concludes 2. ◁

▷ Claim 66. The factor generated from the regularization above is a $\tilde{X}$-relative semantic refinement of the original factor, i.e. $\mathcal{B}' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}$.

Proof. It is enough to show that in each step, the factors generated by the relative-regularization process are semantic refinements relative to $\tilde{X}$ of the previous step's factor. Specifically, we show $\mathcal{B}_2 \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}_1 \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}$ and the claim will follow from transitivity of relative semantic refinements.
We start by proving $\mathcal{B}_1 \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}$. Let $F : \mathbb{F}^n \to \mathbb{F}$ be a function that is $\mathcal{P}$-measurable relative to $\tilde{X}$. We denote $d_F \coloneqq \deg(F)$. By definition, there exists $\Gamma : \mathbb{F}^c \to \mathbb{F}$, $\overline{F} : \mathbb{F}^n \to \mathbb{F}$ where $\deg(\overline{F}), \deg(F - \overline{F}) \leq d_F$ and $\overline{F}|_{\tilde{X}} \equiv 0$, such that:

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma(P_1(a), ..., P_c(a)) + \overline{F}(a)$$

Clearly, the function $\Gamma(P_1(a),...,P_c(a))$ is $\mathcal{P}$-measurable in $\mathbb{F}^n$, and because we have $\mathcal{B} \succeq_{\text{sem}} \mathcal{B}_1$, it is also $\mathcal{P}_1$-measurable in $\mathbb{F}^n$. Thus there exists $\Gamma_1 : \mathbb{F}^{c_1} \to \mathbb{F}$ such that:

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma_1(P_1^1(a),...,P_{c_1}^1(a)) + \overline{F}(a)$$

And therefore we have $\mathcal{B}_1 \succeq_{\text{sem}}^{\tilde{X}} \mathcal{B}$.
Now, we prove $\mathcal{B}_2 \succeq_{\text{sem}}^{\tilde{X}} \mathcal{B}_1$. Let $F : \mathbb{F}^n \to \mathbb{F}$ be a function that is $\mathcal{P}_1$-measurable relative to $\tilde{X}$. Again, we denote $d_F \coloneqq \deg(F)$, and by definition there exists $\Gamma_1 : \mathbb{F}^{c_1} \to \mathbb{F}$, $\overline{F_1} : \mathbb{F}^n \to \mathbb{F}$ where $\deg(F - \overline{F_1}) \leq d_F$ and $\overline{F_1}|_{\tilde{X}} \equiv 0$, such that:

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma_1(P_1^1(a),...,P_{c_1}^1(a)) + \overline{F}_1(a) \tag{1}$$

Note that we also have $\deg(\overline{F_1}) \leq d_F$. We will refer this equation, and its simplifications we do throughout the proof, as *the $\mathcal{P}_1$-decomposition of F*.
We wish to show that there exists $\Gamma_2 : \mathbb{F}^{c_2} \to \mathbb{F}$ and $\overline{F}_2 : \mathbb{F}^n \to \mathbb{F}$ where $\deg(F - \overline{F}_2) \leq d_F$ and $\overline{F}_2|_{\tilde{X}} \equiv 0$, such that:

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma_2\left(P_1^1(a),...P_{i^\star-1}^1(a), P_{i^\star+1}^1(a),...,P_c^1(a), Q_1(a),...,Q_{r(c_1)}(a)\right) + \overline{F}_2(a)$$

We will do so using the $\mathcal{P}_1$-decomposition of F. Note that showing $\deg(F - \overline{F_2}) \leq d_F$ is equivalent of showing $\deg(\overline{F}_2) \leq d_F$.
First, by the way we built $\mathcal{P}_2$, using the same notations in the regularization step, we have:

$$\forall a \in \mathbb{F}^n : P_{i^\star}^1(a) = \vec{\Gamma}\left(Q_1(a),...,Q_{r(c_1)}(a)\right) + \overline{P}(a) - \sum_{i \neq i^\star} \lambda_i P_i^1(a)$$

Next, we substitute the value of $P_{i^\star}^1$ in the $\mathcal{P}_1$-decomposition of F (1), and get another decomposition of F that does not depend on $P_{i^\star}^1$. Specifically we have:

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma_1\left(P_1^1(a),...,\left(\vec{\Gamma}\left(Q_1(a),...,Q_{r(c_1)}(a)\right) + \overline{P}(a) - \sum_{i \neq i^\star} \lambda_i P_i^1(a)\right),...,P_{c_1}^1(a)\right) \tag{2}$$

$$+ \overline{F}_2(a) \tag{3}$$

We wish to use the equation above to show that F is $\mathcal{P}_2$-measurable relative to $\tilde{X}$. However, in order to show that the equation above is in the desired structure that proves that F is $\mathcal{P}_2$-measurable, the expression inside Γ_1 must not depend on $\overline{P}$ because $\overline{P} \notin \mathcal{P}_2$. Note that this is enough as the rest of the polynomials in the expression above are in $\mathcal{P}_2$, and therefore without $\overline{P}$ the expression is $\mathcal{P}_2$-measurable.
To do so, we start by simplifying some of the notations. We denote:

$$\vec{P}_2(a) \coloneqq \vec{\Gamma}\left(Q_1(a),...,Q_{r(c_1)}(a)\right) - \sum_{i \neq i^\star} \lambda_i P_i^1(a)$$

This is the part of the sum that decomposes $P_{i^\star}^1(a)$ that is $\mathcal{P}_2$-measurable, thus the following equality applies:

$$P_{i^\star}^1(a) = \vec{P}_2(a) + \overline{P}(a)$$

where $\deg(\vec{P}_2), \deg(\overline{P}) \leq d^\star$. Using this notation, we write the $\mathcal{P}_1$-decomposition of F (2), and get:

$$\forall a \in \mathbb{F}^n : F(a) = \Gamma_1\left(P_1^1(a),...,\left(\vec{P}_2(a) - \overline{P}(a)\right),...,P_{c_1}^1(a))\right) + \overline{F}_1(a) \tag{4}$$

Now, we use the following key observation: $rank\,(\mathcal{P}_1) \geq r_{30}(\mathbb{F}, d, c_1)$, and as $\deg(\Gamma_1(P_1^1, ..., P_{c_1}^1)) \leq d_F$ we can use Lemma 30 to achieve that Γ_1 is a polynomial of the form:

$$\Gamma_1(z_1, ..., z_{c_1}) = \sum_{\alpha \in [p-1]^{c_1}} C_\alpha \cdot \prod_{i=1}^{c_1} z_i^{\alpha_i} \tag{$\star$}$$

where $C_\alpha = 0$ whenever $\sum_{i=1}^{c_1}(\alpha_i \cdot \deg(P_i^1)) > d_F$.
Next, we substitute the polynomial structure of Γ_1 $(\star)$ in the $\mathcal{P}_1$-decomposition of F (4), and observe what happens to each summand monomial with non-zero coefficients of Γ_1 in the expression after the substitution.
We will show that each such monomial is either $\mathcal{P}_2$-measurable, or a sum of a $\mathcal{P}_2$-measurable function with a valid $\tilde{X}$-remainder, i.e. a polynomial of degree $\leq d_F$ that is $\equiv 0$ in $\tilde{X}$. Note that if this is true for each monomial, every linear combination of such monomials is also a sum of $\mathcal{P}_2$-measurable function with a valid $\tilde{X}$-remainder. Thus, this will also be true for the entire decomposition of F, as it is a linear combination of such monomials summed with a valid remainder $\overline{F}_1$. This will conclude the proof.
Let $\alpha = (\alpha_1, ..., \alpha_{c_1})$ be a vector of degrees that represents such a monomial. If $\alpha_{i^\star} = 0$, then the monomial is in the form:

$$\prod_{i \in [c_1]} {P_i}^{\alpha_i} = \prod_{i \in [c_1] \setminus \{i^\star\}} {P_i}^{\alpha_i}$$

and therefore it is clearly $\mathcal{P}_2$-measurable as all the polynomials in the expression above are in $\mathcal{P}_2$.
Next, if $\alpha_{i^\star} \neq 0$, then the monomial is in the form:

$$\prod_{i \in [c_1]} {P_i}^{\alpha_i} = (\vec{P_2} + \overline{P})^{\alpha_{i^\star}} \cdot \left(\prod_{i \in [c_1] \setminus \{i^\star\}} {P_i}^{\alpha_i} \right) \tag{5}$$

where $\sum_{i \in [c_1]}(\alpha_i \cdot \deg(P_i^1)) \leq d_F$. As $\deg(\vec{P_2} + \overline{P}) = \deg(P_{i^\star}) = d^\star$, we have:

$$\deg\left(\prod_{i \in [c_1] \setminus \{i^\star\}} {P_i}^{\alpha_i} \right) = \sum_{i \in [c_1] \setminus i^\star} (\alpha_i \cdot \deg(P_i^1)) \leq d_F - \alpha_{i^\star} \cdot d^\star$$

Now, we open the left brackets in (5), i.e. $(\vec{P_2} + \overline{P})^{\alpha_{i^\star}}$. This enables us to separate the monomial to the part that only depend on $\vec{P_2}$ summed with a polynomial with bounded degree multiplied by $\overline{P}$ (and therefore a valid remainder). To be more specific, the monomial is in the form:

$$(\vec{P_2} + \overline{P})^{\alpha_{i^\star}} = {\vec{P_2}}^{\alpha_{i^\star}} + \overline{P_\alpha}$$

for some polynomial $\overline{P_\alpha}$ such that:
1. $\overline{P_\alpha}$ is of degree $\deg(\overline{P_\alpha}) \leq \max\left\{\deg(\vec{P_2}), \deg(\overline{P})\right\} \cdot \alpha_{i^\star} \leq \alpha_{i^\star} \cdot d^\star$
2. $\overline{P_\alpha}$ is a multiple of $\overline{P}$, and therefore $\overline{P_\alpha}|_{\tilde{X}} \equiv 0$
Therefore, by substituting the left brackets back to the equation (5) and as $\vec{P_2}$ and P_i for $i \neq i^\star$ are $\mathcal{P}_2$-measurable, one can see that the monomial is a sum of a $\mathcal{P}_2$-measurable polynomial with a valid remainder. Specifically, the remainder $\equiv 0$ in $\tilde{X}$, and its degree is $\leq \alpha_{i^\star} \cdot d^\star + d_F - \alpha_{i^\star} \cdot d^\star = d_F$. This concludes the proof of the claim. ◁

Now, it remains to prove the second part of the Theorem 64.

▷ Claim 67. If $\mathcal{B} \succeq_{\text{syn}} \bar{\mathcal{B}}$ for some polynomial factor $\bar{\mathcal{B}}$ with relative rank of at least $r(c') + c' + 1$ and rank of at least $r_{30}(\mathbb{F}, d, c') + c' + 1$, then we can require that $\mathcal{B}' \succeq_{\text{syn}} \bar{\mathcal{B}}$.

Proof. We will show claim step-by-step. We denote by $\mathcal{P}, \bar{\mathcal{P}}, \mathcal{P}_1, \mathcal{P}_2$ the polynomial sets that generate the factors $\mathcal{B}, \bar{\mathcal{B}}, \mathcal{B}_1, \mathcal{B}_2$. Note that $\mathcal{B}_1, \mathcal{B}_2$ are the factors in the current step of the regularization process, and thus change in each step of the proof. We show that in each step, if $\mathcal{B} \succeq_{\text{syn}} \bar{\mathcal{B}}$ for some polynomial factor $\bar{\mathcal{B}}$ with relative rank of at least $r(c') + c' + 1$ and rank of at least $r_{30}(\mathbb{F}, d, c') + c' + 1$, then we can require that $\mathcal{B}_1 \succeq_{\text{syn}} \bar{\mathcal{B}}$, and also that $\mathcal{B}_2 \succeq_{\text{syn}} \bar{\mathcal{B}}$. For the first part, we have $\mathcal{B}_1 \succeq_{\text{syn}} \bar{\mathcal{B}}$ by a simple usage of the second part of lemma 33, as:

$$rank\left(\bar{\mathcal{P}}\right) > r_{30}(\mathbb{F}, d, c') + c' + 1 \geq r_{30}(\mathbb{F}, d, c_1) + c_1 + 1$$

Now we prove the second part. We show that in the current regularization step, we could replace $P^1_{i^\star} \in \mathcal{P}_1$ such that $P^1_{i^\star} \notin \bar{\mathcal{P}}$. Note that this is possible whenever $\mathcal{P}^\star \cap \bar{\mathcal{P}} \neq \emptyset$ as the choice of $i^\star$ is arbitrary in polynomials which are in $\mathcal{P}^\star$.
Assume that is not possible and the factor $\mathcal{P}_1$ is still not r-$\tilde{X}$-regular. Then, we have a linear combination $\vec{P}(x) := \sum_{i=0}^{c_1} \lambda_i P^1_i(x)$ with $rank_{d^\star, \tilde{X}}\left(\vec{P}\right) \leq r(c_1)$ where $d^\star = \max_{i \in [c_1]} \deg(\lambda_i P^1_i)$. We denote by $I^\star \subseteq [c_1]$ the set of indexes of such maximal-degree polynomials. By this notation, our assumption states that for all $i \in I^\star$ we have $P^1_i \in \bar{\mathcal{P}}$. Additionally, note that for all $i \notin I^\star$ we have $\deg(P^1_i) < d^\star$. Therefore, as the linear combination is of $d^\star$-relative rank $\leq r(c_1)$, there exists a polynomial $\overline{P}$ of degree $\leq \deg(\vec{P}) \leq d^\star$ with $\overline{P}|_{\tilde{X}} \equiv 0$ such that $rank_{d^\star}\left(\vec{P} - \overline{P}\right) \leq r(c_1)$. In other words, there exist a measurement function $\vec{\Gamma} : \mathbb{F}^{r(c_1)} \to \mathbb{F}$ and polynomials $Q_1, ..., Q_{r(c_1)}$ with $\deg(Q_i) \leq d^\star$ such that:

$$\forall a \in \mathbb{F}^n : \vec{P}(a) - \overline{P}(a) = \vec{\Gamma}\left(Q_1(a), ..., Q_{r(c_1)}(a)\right)$$

By a simple calculation we have:

$$\forall a \in \mathbb{F}^n : \sum_{i \in I^\star} P^1_i(a) - \overline{P}(a) = \vec{\Gamma}\left(Q_1(a), ..., Q_{r(c_1)}(a)\right) + \sum_{i \notin I^\star} P^1_i(a)$$

and by this we found a linear combination of polynomials in $\bar{\mathcal{P}}$ with maximal degree $d^\star$, that has $d^\star$-relative-rank $\leq r(c_1) + c_1 + 1$. This is a contradiction to our assumptions on $\bar{\mathcal{B}}$, which completes the proof of the claim. ◁

This completes the proof of the lemma. ◀

6 Radius of Reed-Muller over $\tilde{X}$

We recall that the normalized distances of Reed-Muller codes over $\mathbb{F}^n$ and over $\tilde{X}$ are denoted by $\delta_{\mathbb{F}}(d)$ and $\delta_{\mathbb{F}, \tilde{X}}(d)$ respectively. We present a theorem that shows that Reed-Muller codes over a subset $\tilde{X} \subseteq \mathbb{F}^n$ that is d-lift-enabler and has the (limited) relative rank-bias property, has (approximately) an *equal* normalized distance as Reed-Muller codes over $\mathbb{F}^n$.

▶ **Theorem 68.** *There exist a function $\epsilon_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ such that the following holds: Let $\mathbb{F}$ be a finite field, and let $d \in \mathbb{N}$ be an integer that represents a degree. Let $\tilde{\epsilon} > 0$, and let $\tilde{r} : [\tilde{\epsilon}, \infty] \to \mathbb{N}$ be a limited-relative-rank-bias function.*
Let $\tilde{X} \subseteq \mathbb{F}^n$ be a set with the following properties

1. *$\tilde{X}$ is d-lift-enabler with a lift operator $\widehat{\square}$.*
2. *$\tilde{X}$ has the $(\tilde{r}, \mathbb{F}, d, \tilde{\epsilon})$-relative-rank-bias property.*
Then, for $\epsilon_1 \coloneqq \epsilon_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ we have that for all $n \in \mathbb{N}$:

$$\delta_{\mathbb{F},\tilde{X}}(d) \geq \ \delta_{\mathbb{F}}(d) - \epsilon_1$$

Proof. We wish to do a reduction of our question regarding the radius of Reed-Muller in $\tilde{X}$ to the same question about Reed-Muller in $\mathbb{F}^n$. Let $\mathbb{F}$ be a finite field, and let $d \in \mathbb{N}$ be an integer that represents a degree. Let $\tilde{\epsilon} > 0$, and let $\tilde{r} : [\tilde{\epsilon}, \infty] \to \mathbb{N}$ be a limited-rank-relative-bias function. Let $\epsilon_1 \coloneqq \epsilon_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ be a function we will specify later. Let $\tilde{X} \subseteq \mathbb{F}^n$ be a set with the properties defined above.
Moreover, let $\epsilon > \epsilon_1$ be some positive value. We will show that:

$$\delta_{\mathbb{F},\tilde{X}}(d) > \ \delta_{\mathbb{F}}(d) - \epsilon$$

This will be enough as if the above holds for every $\epsilon > \epsilon_1$, we get that in fact $\delta_{\mathbb{F},\tilde{X}}(d) \geq \delta_{\mathbb{F}}(d) - \epsilon_1$.
For start, we note a simple observation: as Reed-Muller over $\tilde{X}$ is a linear code, we have

$$\delta_{\mathbb{F},\tilde{X}}(d) = \min\left\{ \Pr_{x\in\tilde{X}}\left[p(x) \neq 0\right] \middle| p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F}) \right\}$$

Now, let $p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ be a polynomial over $\tilde{X}$, and denote $d_p \coloneqq \deg(p)$. We wish to lower-bound the value of $\Pr_{x\in\tilde{X}}\left[p(x) \neq 0\right]$. To do so, we will equivalently upper-bound the value of $\Pr_{x\in\tilde{X}}\left[p(x) = 0\right]$. Precisely, to complete the proof all we need to show is:

$$\Pr_{x\in\tilde{X}}\left[p(x) = 0\right] \leq 1 - \delta_{\mathbb{F}}(d) + \epsilon$$

Now we begin the proof itself. First, we lift the polynomial p and get a polynomial $\widehat{p} : \mathbb{F}^n \to \mathbb{F}$ such that $\widehat{p}|_{\tilde{X}} \equiv p$ and $\deg(\widehat{p}) = d_p$. Next, denote by $\mathcal{B}_{\widehat{p}}$ the factor defined by the set of single polynomial $\mathcal{P} = \{\widehat{p}\}$. Trivially, the polynomial $\widehat{p}$ is measurable in respect of $\mathcal{P}$.
We define the rank function:

$$r(m) \coloneqq \max\left\{ \tilde{r}\left(\frac{\epsilon/2}{|\mathbb{F}|^m}\right), r_{23}\left(\mathbb{F}, d, \frac{\epsilon/2}{|\mathbb{F}|^m}\right) \right\}$$

Then, we r-$\tilde{X}$-regularize $\mathcal{P}$ using Lemma 64. This gives us a r-$\tilde{X}$-regular factor $\mathcal{B}'$, which is defined by a set of polynomials $\mathcal{P}' \coloneqq \{P'_1, ..., P'_{c'}\}$ of degree $\leq d$ such that $\mathcal{B}' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}_{\widehat{p}}$ with $rank_{\tilde{X}}(\mathcal{P}') \geq r$ and with bounded amount of polynomials defining it i.e,$c' \leq C^{64}_{r,d}(1)$. Therefore, from definition we have that $\widehat{p}$ is $\mathcal{P}'$-measurable relative to $\tilde{X}$. Thus, there exists a measurement function $\Gamma : \mathbb{F}^{c'} \to \mathbb{F}$ and a remainder $\overline{\Gamma} : \mathbb{F}^n \to \mathbb{F}$ with $\overline{\Gamma}|_{\tilde{X}} \equiv 0$ and degree bounded by d_p, such that:

$$\forall a \in \mathbb{F}^n : \widehat{p}(a) = \Gamma(P'_1(a), ..., P'_{c'}(a)) + \overline{\Gamma}(a)$$

Next, we denote $P' \coloneqq \widehat{p} - \overline{\Gamma}$. By definition of remainder function, we have that $P'|_{\tilde{X}} \equiv p$. Additionally, note that P' is a polynomial over $\mathbb{F}^n$ of degree $\deg(P') = d_p \leq d$, and hence by the definition of $\delta_{\mathbb{F}}(d)$:

$$\Pr_{a\in\mathbb{F}^n}\left[P'(a) = 0\right] \leq 1 - \delta_{\mathbb{F}}(d) \tag{6}$$

For the next step, we claim that P' equals 0 in $\mathbb{F}^n$ approximately with the same probability it equals 0 in $\tilde{X}$. Note that this is the heart of the proof: it allows use properties known in $\mathbb{F}^n$ to new properties in $\tilde{X}$. This is formulated as follows:

▷ Claim 69. We have:

$$\left| \Pr_{a\in\mathbb{F}^n}[P'(a)=0] - \Pr_{x\in\tilde{X}}[P'(x)=0] \right| \leq \epsilon$$

Proof. Denote $S \coloneqq \mathbb{F}^{c'}$, and for all $s \in S$, denote:

$$p_1(s) \coloneqq \Pr_{a\in\mathbb{F}^n}[(P'_1(a),...,P'_{c'}(a)) = s]$$

As of our choice of r, we have $rank(\mathcal{P}') \geq r_{23}\left(\mathbb{F}, d, \frac{\epsilon/2}{|\mathbb{F}|^{c'}}\right)$. By combining Theorem 23 with Lemma 81, we have that p_1 is $(\epsilon/2\,|S|)$-equidistributed, i.e:

$$p_1(s) = \frac{1 \pm \epsilon/2}{|S|}$$

Similarly, denote:

$$p_2(s) \coloneqq \Pr_{x\in\tilde{X}}[(P'_1(x),...,P'_{c'}(x)) = s]$$

As of our choice of r, we have $rank_{\tilde{X}}(\mathcal{P}') \geq \tilde{r}(\epsilon/2\,|S|)$. Now, we wish to use the relative rank-bias relation with Lemma 81 to conclude similarly that p_2 is $(\epsilon/2\,|S|)$-equidistributed, i.e:

$$p_2(s) = \frac{1 \pm \epsilon/2}{|S|}$$

However, in order to so, we must first ensure that $(\epsilon/2\,|S|) \geq \tilde{\epsilon}$. This is done by choosing a correct ϵ_1, and formulated in the following claim:

▷ Claim 70. One can choose $\epsilon_1 \coloneqq \epsilon_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ such that if $\epsilon \geq \epsilon_1$ we have that $\epsilon/2\,|S| \geq \epsilon_1$.

Proof. We need that:

$$\frac{\epsilon}{2\,|\mathbb{F}|^{c'}} \geq \tilde{\epsilon}$$

As $c' \leq C^{64}_{r,d}(1)$, for the term above to hold it is enough that the following will be true:

$$\epsilon \geq \tilde{\epsilon} \cdot 2\,|\mathbb{F}|^{C^{64}_{r,d}(1)}$$

and as r and thus also $C^{64}_{r,d}(1)$ are independent of n, we can pick $\epsilon_1 = \epsilon_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ and get what we aimed for. ◁

Now, under that assumption of ϵ_1 written above, we have that p_2 is $(\epsilon/2\,|S|)$-equidistributed. This allows us to use the similar distributions of $\mathcal{P}'$ in $\mathbb{F}^n$ and in $\tilde{X}$ to conclude that P' behaves similar in $\mathbb{F}^n$ and in $\tilde{X}$:

$$\begin{aligned}\Pr_{a\in\mathbb{F}^n}[P'(a)=0] &= \sum_{s\in S} p_1(s)\cdot 1_{\Gamma(s)=0} \\ &= \sum_{s\in S} p_2(s)\cdot 1_{\Gamma(s)=0} \pm \epsilon \\ &= \Pr_{x\in\tilde{X}}[P'(x)=0] \pm \epsilon\end{aligned}$$

which concludes the proof of the claim. ◁

Finally, as $P'|_{\tilde{X}} \equiv p$, we have that $\Pr_{x\in\tilde{X}}[P'(x)=0] = \Pr_{x\in\tilde{X}}[p(x)=0]$. Thus, the claim above combining with (6) shows that the probability we wished to bound is bounded as we aimed for:

$$\Pr_{x\in\tilde{X}}[p(x)=0] \leq 1-\delta_{\mathbb{F}}(d)+\epsilon$$

This concludes the proof of the theorem. ◀

▶ Remark 71. Under the same conditions, the distance of Reed-Muller codes in $\tilde{X}$ is also bounded *from above* by the distance of Reed-Muller codes in $\mathbb{F}^n$, and we have:

$$\delta_{\mathbb{F},\tilde{X}}(d) \leq \delta_{\mathbb{F}}(d)+\epsilon_1$$

Proof. Let $P:\mathbb{F}^n\to\mathbb{F}$ be the polynomial in $\mathbb{F}^n$ with the *smallest* distance from 0 as possible, that is $\delta_{\mathbb{F}}(d)$. Denote $p \coloneqq P|_{\tilde{X}}$. Note that p is a polynomial in $\tilde{X}$. Now repeat the proof using these two polynomials, and by Claim 69, we have that a random input of P yields 0 (approximately) the same as a random input of p yields 0. Thus as we have $\Pr_{x\in\mathbb{F}^n}[P(x)=0] = 1-\delta_{\mathbb{F}}(d)$ we also get:

$$\Pr_{x\in\tilde{X}}[p(x)=0] \geq 1-\delta_{\mathbb{F}}(d)-\epsilon_1$$

This bounds *from above* the distance of Reed-Muller code in $\tilde{X}$ and we have:

$$\delta_{\mathbb{F},\tilde{X}}(d) \leq \delta_{\mathbb{F}}(d)+\epsilon_1$$ ◀

▶ **Corollary 72.** *If we assume $\tilde{X}$ has the limited-relative rank-bias property to* any extent *(or just the relative rank-bias property), then the theorem above proves an exact equality $\delta_{\mathbb{F},\tilde{X}}(d) = \delta_{\mathbb{F}}(d)$.*

7 List Decoding Reed Muller Over $\tilde{X}$

In this section, we prove our main theorem: we prove the list decoding radius of Reed-Muller codes in $\tilde{X}$ is *at least* the list decoding radius of Reed-Muller codes in $\mathbb{F}^n$, assuming $\tilde{X}$ is lift-enabler and has the relative rank-bias property. We start by presenting formally the list decoding radius in $\tilde{X}$.

▶ **Definition 73** (List Decoding in $\tilde{X}$)**.** *Let $\mathbb{F}$ be a finite field. Let $d,n\in\mathbb{N}$, and let $\tilde{X}\subseteq\mathbb{F}^n$. We define the Reed-Muller list-decoding count in $\tilde{X}$ at distance τ as follows:*

$$\ell_{\mathbb{F},\tilde{X}}(d,\tau) \coloneqq \max_{F:\tilde{X}\to\mathbb{F}} \left|\left\{P\in Poly_{\leq d}(\tilde{X}\to\mathbb{F}) \middle| dist(P,F)\leq\tau\right\}\right|$$

Additionally, we define $LDR_{\mathbb{F},\tilde{X}}(d)$ to be the list decoding radius*, which is the maximum τ for which $\ell_{\mathbb{F},\tilde{X}}(d,\tau-\epsilon)$ is bounded by a* constant *depending only on $\epsilon,|\mathbb{F}|,d$.*

We recall that it was shown in [8, Theorem 1] that the list decoding radius of Reed Muller is $\delta_{\mathbb{F}}(d)$. To be more precise, it was shown that for every $\epsilon>0$, the list-decoding count is constant (independent of n) in distance $\tau=\delta_{\mathbb{F}}(d)-\epsilon$. Formally, they have shown the following theorem:

▶ **Theorem 74** (List Decoding RM in $\mathbb{F}^n$)**.** *There exists a function $c(\mathbb{F},d,\epsilon)$ such that the following holds: Let $\mathbb{F}$ be a finite field, let $\epsilon>0$, and let $d,n\in\mathbb{N}$. Then, we have:*

$$\ell_{\mathbb{F},\mathbb{F}^n}(d,\delta_{\mathbb{F}}(d)-\epsilon) \leq c(\mathbb{F},d,\epsilon)$$

Additionally, we recall a lemma that was presented in [8, Corollary 3.3], and was used in the analysis of the list decoding radius of Reed-Muller codes in $\mathbb{F}^n$:

▶ **Lemma 75** (Low Complexity Approximation). *[8, Corollary 3.3] Let $G : A \to B$, and let $\epsilon > 0$. Let $\mathfrak{F} \subseteq B^A$ be a collection of functions from A to B. Then there exists $c \leq 1/\epsilon^2$ functions $F_1, ..., F_c \in \mathfrak{F}$ such that for every $F \in \mathfrak{F}$, there is a function $\Gamma_F : B^c \to B$ such that:*

$$\Pr_{x \in A}[\Gamma_F(F_1(x), ..., F_c(x)) = F(x)] \geq \Pr_{x \in A}[G(x) = F(x)] - \epsilon$$

The lemma shows that G can "estimated" by a only a few functions from $\mathfrak{F}$. Note that the estimation is close to G compared to every $F \in \mathfrak{F}$ and not necessarily close to G itself.

Finally, we present our main theorem, which shows that under assumptions on the subset $\tilde{X} \subseteq \mathbb{F}^n$, the list decoding radius of polynomials in $\tilde{X}$ will be similar to the list decoding radius in $\mathbb{F}^n$.
In more details (and informally), we show that if $\tilde{X} \subseteq \mathbb{F}^n$ is lift-enabler, has the limited-relative-rank-bias-property, the list-decoding count is constant (independent of n) for every valid ϵ in distance $\tau = \delta_{\mathbb{F}}(d) - \epsilon$. Note that not every $\epsilon > 0$ will be valid: the valid values of ϵ will depend on the limitations of the rank-bias property. Formally, we show the following:

▶ **Theorem 76** (List Decoding RM in $\tilde{X}$). *There exist functions $c_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ and $c_2(\mathbb{F}, d, \tilde{r}, \epsilon)$ such that the following holds: Let $\mathbb{F}$ be a finite field, and let $d \in \mathbb{N}$ be an integer that represents a degree. Let $\tilde{\epsilon} > 0$, and let $\tilde{r} : [\tilde{\epsilon}, \infty] \to \mathbb{N}$ be a limited-relative-rank-bias function. Let $\tilde{X} \subseteq \mathbb{F}^n$ be a set with the following properties*
1. *$\tilde{X}$ is d-lift-enabler with a lift operator $\widehat{\square}$.*
2. *$\tilde{X}$ has the $(\tilde{r}, \mathbb{F}, d, \tilde{\epsilon})$-relative-rank-bias property.*

Then, for every $\epsilon \geq c_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ it holds:

$$\ell_{\mathbb{F}, \tilde{X}}(d, \delta_{\mathbb{F}, \mathbb{F}^n}(d) - \epsilon) \leq c_2(\mathbb{F}, d, \tilde{r}, \epsilon)$$

Proof. We follow the lines of the proof of [8, Theorem 1]. Let $\mathbb{F}$ be a finite field, and let $d \in \mathbb{N}$ be an integer that represents a degree. Let $\tilde{\epsilon} > 0$, and let $\tilde{r} : [\tilde{\epsilon}, \infty] \to \mathbb{N}$ be a limited-rank-relative-bias function. Let $c_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ be a function we will specify later. Let $\tilde{X} \subseteq \mathbb{F}^n$ be a set with the properties defined above.
Finally, let $\epsilon \geq c_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ for c_1 that we will specify later, and let $f : \tilde{X} \to \mathbb{F}$ be a received word. We wish to bound the amount of polynomials in $Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ that are $(\delta_{\mathbb{F}}(d) - \epsilon)$-close to f.
Apply Lemma 75 with $A = \tilde{X}$, $B = \mathbb{F}$, $G = f$, $\mathfrak{F} = Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ and approximation parameter $\epsilon/2$ to obtain $\mathfrak{h} \subset Poly_{\leq d}(\tilde{X} \to \mathbb{F})$, defined by $\mathfrak{h} = (h_1, ..., h_c)$ where $c \leq 4/\epsilon^2$, such that for every $p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ there is a function $\Gamma_p : \mathbb{F}^c \to \mathbb{F}$ that approximates f in $\tilde{X}$ relative to $Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ i.e.:

$$\forall p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F}) : \Pr_{x \in \tilde{X}}[\Gamma_p(h_1(x), ..., h_c(x)) = p(x)] \geq \Pr_{x \in \tilde{X}}[f(x) = p(x)] - \epsilon/2$$

Let $r_1, r_2 : \mathbb{N} \to \mathbb{N}$ be two non-decreasing functions that represents rank that we will specify later. For r_1, we will require that for all $m \geq 1$:

$$r_1(m) \geq \max \left\{ \begin{matrix} r_2(C^{64}_{r_2,d}(m+1)) + C^{64}_{r_2,d}(m+1) + 1, \\ r_2(C^{64}_{r_{30},d}(m+1)) + C^{64}_{r_{30},d}(m+1) + 1 \end{matrix} \right\}$$

Note that in the expression above, we denote $r_{30} : \mathbb{N} \to \mathbb{N}$, as follows: $r_{30}(c) \coloneqq r_{30}(\mathbb{F}, d, c)$. The reason we chose this r_1, is that by our choice of r_1 we can use the second part of Lemma 64. Specifically, if we start with r_1-$\tilde{X}$-regular factor and we r_2-$\tilde{X}$-regularize it, we get that the r_2-$\tilde{X}$-regular factor that we received is a syntactic refinement of the r_1-$\tilde{X}$-regular factor we started with.

As a first step, we lift the polynomial factor to get $\mathcal{H} \coloneqq \widehat{\mathfrak{h}}$. Note that because $\forall x \in \tilde{X} : \widehat{h_i}(x) = h_i(x)$, for all $p \in F$ we have:

$$\Pr_{x \in \tilde{X}} \left[\Gamma_p(\widehat{h_1}(x), ..., \widehat{h_c}(x)) = p(x) \right] \geq \Pr_{x \in \tilde{X}} \left[f(x) = p(x) \right] - \epsilon/2$$

Next, we r_1-$\tilde{X}$-regularize the factor $\mathcal{B}$ generated by the collection $\mathcal{H}$ by Theorem 64. This gives us a r_1-$\tilde{X}$-regular factor $\mathcal{B}'$, which is defined by a set of polynomials $\mathcal{H}' \coloneqq (H'_1, ..., H'_{c'})$ of degree $\leq d$ such that $\mathcal{B}' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}$, with $rank_{\tilde{X}}(\mathcal{H}') \geq r(c')$ and with bounded amount of polynomials defining it i.e. $c' \leq C^{64}_{r_1,d}(c)$. We apply Corollary 63 and get that $\mathcal{B}' \succeq_{\text{sem}|\tilde{X}} \mathcal{B}$. We then use the fact that $\Gamma_p(\widehat{h_1}(x), ..., \widehat{h_c}(x))$ is measurable in respect of $\mathcal{H}$ *in* $\tilde{X}$, and deduce we have a similar approximation of p using $\mathcal{H}'$ as the approximation of p using $\mathcal{H}$. Formally, there exists a function $\Gamma'_p : \mathbb{F}^{c'} \to \mathbb{F}$ such that:

$$\Pr_{x \in \tilde{X}} \left[\Gamma'_p(H'_1(x)), ..., H'_{c'}(x)) = p(x) \right] \geq \Pr_{x \in \tilde{X}} \left[f(x) = p(x) \right] - \epsilon/2$$

Now we recall that we wished to bound the amount of polynomials $p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ such that $\Pr_{x \in \tilde{X}} [f(x) \neq p(x)] \leq \delta_{\mathbb{F}}(d) - \epsilon$. Let $p \in Poly_{\leq d}(\tilde{X} \to \mathbb{F})$ be a polynomial as we just described. We will show that such p is measurable with respect to $\mathcal{H}'$ in $\tilde{X}$. This will upper bound the amount of possible polynomials p by the amount of possible different $\Gamma'_p : \mathbb{F}^{c'} \to \mathbb{F}$, which is $|\mathbb{F}|^{\|\mathcal{B}'\|} = p^{(p^{c'})}$, and thus $c_2(\mathbb{F}, d, \tilde{r}, \epsilon) \leq p^{(p^{c'})}$.

By our choice of c' we have that $c' \leq C^{64}_{r_1,d}(4/\epsilon^2)$, and thus c_2 is bounded by a function of $(\mathbb{F}, d, r_1, \epsilon)$. Note that we have not yet specified the value of r_1, because it is determined by the choice of r_2 that we will later define its exact values. The important thing about our future choice of r_2 is that the value of r_2 must be independent of n, but can depend on $(\mathbb{F}, d, \tilde{r}, \epsilon)$. This will conclude the proof.

Now, consider a lift of p, i.e. $P \coloneqq \widehat{p}$. Note that by the definition of lift $\forall x \in \tilde{X} : P(x) = p(x)$. We will show that P is measurable in respect of $\mathcal{H}'$ in $\tilde{X}$.

We consider the factor $\mathcal{B}_P$ that is generated by $\mathcal{H}_P \coloneqq \mathcal{H}' \cup \{P\}$. By using Theorem 64, we can r_2-$\tilde{X}$-regularize it and get the polynomial factor $\mathcal{B}''$ that relative-refines $\mathcal{B}_P$. We denote the set of polynomials in the factor as $\mathcal{H}''$.

Next, notice that the factor $\mathcal{B}''$ is a r_2-regular factor, therefore by our choice of r_1 and the second part of Theorem 64, we in fact have $\mathcal{B}'' \succeq_{\text{syn}} \mathcal{B}'$. This is true because by our choice of r_1:

$$rank_{\tilde{X}}(\mathcal{H}') \geq r_1(c') \geq r_2(C^{64}_{r_2,d}(c'+1)) + C^{64}_{r_2,d}(c'+1) + 1 \geq r_2(|\mathcal{B}''|) + |\mathcal{B}''| + 1$$

And as rank is always bigger than relative rank, we also have:

$$rank(\mathcal{H}') \geq r_1(c') \geq r_2(C^{64}_{r_{30},d}(c'+1)) + C^{64}_{r_{30},d}(c'+1) + 1$$

Thus, the polynomials defining $\mathcal{B}''$ are in the form $\mathcal{H}'' \coloneqq \mathcal{H}' \cup \{H''_1, ..., H''_{c''}\}$. Note that as promised in Theorem 64, we have $|\mathcal{H}''| = c' + c'' \leq C^{64_{r_2,d}}(c')$.

Additionally, by the way we built $\mathcal{H}_P$, the function P is measurable in respect of it. Therefore, as $\mathcal{B}'' \succeq^{\tilde{X}}_{\text{sem}} \mathcal{B}_P$, we have that P is $\mathcal{H}''$-measurable relative to $\tilde{X}$. In other words, there exists

$\Phi : \mathbb{F}^{c'+c''} \to \mathbb{F}$ and $\overline{P} : \mathbb{F}^n \to \mathbb{F}$ with $\deg(\overline{P}), \deg(P - \overline{P}) \leq \deg(P) \leq d$ and $\overline{P}|_{\tilde{X}} \equiv 0$ such that:

$$\forall a \in \mathbb{F}^n : P(a) = \Phi(H'_1(a), ..., H'_{c'}(a), H''_1(a), ..., H''_{c''}(a))) + \overline{P}(a)$$

And specifically in $\tilde{X}$ we have:

$$\forall x \in \tilde{X} : P(x) = \Phi(H'_1(x), ..., H'_{c'}(x), H''_1(x), ..., H''_{c''}(x)))$$

Denote $P' \coloneqq P - \overline{P}$. We will show the polynomial P' does not depend on its last c'' variables, and thus Φ does not depend on its last c'' variables. This will imply that P is measurable in respect of $\mathcal{H}'$ in $\tilde{X}$, which will conclude the proof.
Now, we choose r_2 to be such that:

$$r_2(m) \geq \max\left\{\tilde{r}\left(\frac{\epsilon/4}{|\mathbb{F}|^m}\right), r_{23}\left(\frac{\epsilon/4}{|\mathbb{F}|^m}\right), r_{30}(m)\right\}$$

Note that in the expression above we are discussing fixed field and degree, i.e. $\mathbb{F}, d$. Therefore we denote $r_{30} : \mathbb{N} \to \mathbb{N}$ as $r_{30}(c) \coloneqq r_{30}(\mathbb{F}, d, c)$ and $r_{23} : \mathbb{N} \to \mathbb{N}$ as $r_{23}(\epsilon) \coloneqq r_{23}(\mathbb{F}, d, \epsilon)$.
Next, we show that even if we change the polynomials in the factor to have a disjoint set of inputs in $\mathbb{F}^n$, we still obtain a polynomial in the same degree, which have an approximation close to the approximation we had in $\tilde{X}$. Note that after this step, the proof becomes very similar to the proof of list decoding Reed Muller in $\mathbb{F}^n$ [8, Theorem 1]: we omit the dependence of $\tilde{X}$ and get the same approximation by functions of multiple variables, as we had in $\mathbb{F}^n$. This is done by the following lemma:

▶ **Lemma 77.** *Let $\left\{a^i, b^j\right\}, i \in [c'], j \in [c'']$ be pairwise disjoint sets of n variables each. Let $n' \coloneqq n(c' + c'')$. Let $\vec{P} : \mathbb{F}^{n'} \to \mathbb{F}$ and $\vec{f} : \mathbb{F}^{n'} \to \mathbb{F}$ be functions of n' variables defined as follows:*

$$\vec{P'}(\vec{a}) \coloneqq \Phi\left(H'_1(a^1), ..., H'_{c'}(a^{c'}), H''_1(b^1), ..., H''_{c''}(b^{c''})\right)$$

and:

$$\vec{f}(\vec{a}) \coloneqq \Gamma'_p(H'_1(a^1)), ..., H'_{c'}(a^{c'}))$$

Note that $\vec{f}$ is a function that receives n' variables, and ignores its last c'' variables.
Then:

1. *The degree of $\vec{P'}$ remains bounded, i.e. $\deg(\vec{P'}) \leq d$.*
2. *The approximation of $\vec{f}$ to $\vec{P'}$ in $\mathbb{F}^{n'}$ is close to the approximation of Γ'_p to p in $\tilde{X}$. Specifically, we show:*

$$\left|\Pr_{\vec{a} \in \mathbb{F}^{n'}}\left[\vec{f}(\vec{a}) = \vec{P'}(\vec{a})\right] - \Pr_{x \in \tilde{X}}\left[\Gamma'_p(H'_1(x)), ..., H'_{c'}(x)) = p(x)\right]\right| \leq \epsilon/4$$

Proof. We start by proving the first part of the lemma: bounding the degree of $\vec{P'}$ by d. First, we recall that $P' = P - \overline{P}$ where $\overline{P}$ is a valid remainder. Specifically, we have $\deg(P') = \deg(P - \overline{P}) \leq \deg(P) \leq d$. In addition, by the way we built Φ we have:

$$\forall a \in \mathbb{F}^n : P'(a) = \Phi(H'_1(a), ..., H'_{c'}(a), H''_1(a), ..., H''_{c''}(a)))$$

Thus the function above is of degree $\leq d$. Moreover, we have:

$$rank\left(\mathcal{H}''\right) \geq rank_{\tilde{X}}\left(\mathcal{H}''\right) \geq r_2(|\mathcal{H}''|) \geq r_{30}(|\mathcal{H}''|)$$

Therefore we can use Lemma 30 to get that $\deg(\vec{P'}) \leq \deg(P') \leq d$. Note that in order to use the lemma formally, we had to extend the input space of P' to be of n' variables (and make it depend only on the first n variables as it used to). Because lemma 30 require bounds independent of n, this is done smoothly.
Now we move to the second part of the lemma: bounding the approximation of $\vec{f}$ to $\vec{P'}$. Denote $S \coloneqq \mathbb{F}^{c'+c''}$, and for each $s \in S$ denote:

$$p_1(s) \coloneqq \Pr_{x \in \tilde{X}} \left[(H'_1(x), ..., H'_{c'}(x), H''_1(x), ..., H''_{c''}(x)) = s \right]$$

and as of our choice of r_2, we have $rank\left(\mathcal{H}''\right) \geq \tilde{r}(\epsilon/8\left|S\right|)$. Therefore, if we require that the relative rank-bias relation holds for $\epsilon/8\left|S\right|$, we can use Lemma 81 with $A = \tilde{X}$ to get that p_1 is $(\epsilon/8\left|S\right|)$-almost uniform, i.e:

$$p_1(s) = \frac{1 \pm \epsilon/8}{|S|}$$

We show that this can be done in the following claim by choosing a proper c_1:

▷ Claim 78. One can choose $c_1 \coloneqq c_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ such that if $\epsilon \geq c_1$ we have that $\epsilon/8\left|S\right| \geq c_1$.

Proof. This is done by using the bound we already know. We need that:

$$\tilde{\epsilon} \leq \frac{\epsilon}{8\left|\mathbb{F}\right|^{c'+c''}}$$

As $c' + c'' \leq C^{64}_{r_2,d}(c')$, for the term above to hold it is enough that the following will be true:

$$\epsilon \geq \tilde{\epsilon} \cdot 8\left|\mathbb{F}\right|^{C^{64}_{r_2,d}(c')}$$

and as r_2, c' and thus also $C^{64}_{r_2,d}(c')$ are independent of n, we can pick $c_1 = c_1(\mathbb{F}, d, \tilde{r}, \tilde{\epsilon})$ and get what we aimed for. ◁

Thus, we can assume that p_1 is $(\epsilon/8\left|S\right|)$-almost uniform. Now, let:

$$p_2(s) \coloneqq \Pr_{\vec{a} \in \mathbb{F}^{n'}} \left[\left(H'_1(a^1), ..., H'_{c'}(a^{c'}), H''_1(b^1), ..., H''_{c''}(b^{c''}) \right) = s \right]$$

Note that the rank of $\vec{\mathcal{H}}'' = \left\{ H'_1(a^1), ..., H'_{c'}(a^{c'}), H''_1(b^1), ..., H''_{c''}(b^{c''}) \right\}$, as a factor defined over $\mathbb{F}^{n'}$, can not be lower than the rank of $\mathcal{H}''$ and thus we have $rank\left(\vec{\mathcal{H}}''\right) \geq r_{23}\left(\frac{\epsilon/8}{|\mathbb{F}|^m}\right)$.
By using Theorem 23, which shows the rank-bias relation for $\mathbb{F}^{n'}$, we can similarly use Lemma 81 with $A = \mathbb{F}^{n'}$ to get that p_2 is also $(\epsilon/8\left|S\right|)$-almost-uniform, i.e:

$$p_2(s) = \frac{1 \pm \epsilon/8}{|S|}$$

Now, we show the approximations are the same. Denote by s' the restriction of s to its first c' coordinates, and consider the approximation:

$$\begin{aligned}
\Pr_{\vec{a} \in \mathbb{F}^{n'}} \left[\vec{f}(\vec{a}) = \vec{P'}(\vec{a}) \right] = & \\
&= \sum_{s \in S} p_2(s) \cdot 1_{\Phi(s) = \Gamma'_P(s')} \\
&= \sum_{s \in S} p_1(s) \cdot 1_{\Phi(s) = \Gamma'_P(s')} \pm \epsilon/4 \\
&= \Pr_{x \in \tilde{X}} \left[\Gamma'_p(H'_1(x)), ..., H'_{c'}(x)) = p(x) \right] \pm \epsilon/4
\end{aligned}$$

This completes the proof the lemma. ◀

The proof is followed by the same methods used in [8]. We repeat if for completeness. We next restate a lemma proved in [8, Claim 4.2], which is a varaiant of the Schwartz-Zippel lemma [43, 46]:

▶ **Lemma 79.** *Let d, n_1, $n_2 \in \mathbb{N}$ be integers. Let $P_1 \in Poly_{\leq d}(\mathbb{F}^{n_1+n_2} \to \mathbb{F})$, and let $F_1 : \mathbb{F}^{n_1} \to \mathbb{F}$ be a function. Assume the polynomial is $\delta_{\mathbb{F}}(d)$-close to the function, i.e:*

$$\Pr_{x_1,\ldots,x_{n_1+n_2}\in\mathbb{F}}[P_1(x_1,...,x_{n_1+n_2}) = F_1(x_1,...,x_n)] > 1 - \delta_{\mathbb{F}}(d)$$

Then, P_1 does not depend on $x_{n_1+1},...,x_{n_1+n_2}$.

Now, apply Lemma 79 to $P_1 = \vec{P'}$, $F_1 = \vec{f}$, $n_1 = nc'$, $n_2 = nc''$. We obtain that $\vec{P'}$ does not depend on its last c'' variables, and thus by denoting $C_i \coloneqq H''_i(0)$ for $i \in [c'']$ we have:

$$\vec{P'}(\vec{a}) = \Phi\left(H'_1(a^1),...,H'_{c'}(a^{c'}),C_1,...,C_{c''}\right)$$

Now, for every $a \in \mathbb{F}^n$, if we substitute a in the i-th component of $\vec{a}$ for every $i \in [c']$ in the equation above, we get the following is true:

$$P'(a) = \Phi\left(H'_1(a),...,H'_{c'}(a),C_1,...,C_{c''}\right)$$

Hence P' does not depend on its last c'' variables. As explained earlier, this implies that P is measurable in respect of $\mathcal{H}'$ in $\tilde{X}$. This completes the proof of the theorem. ◀

References

1. Omar Alrabiah, Jesse Goodman, Jonathan Mosheiff, and João Ribeiro. Low-degree polynomials are good extractors, 2025. doi:10.48550/arXiv.2405.10297.
2. Omar Alrabiah, Venkatesan Guruswami, and Ray Li. Randomly punctured reed–solomon codes achieve list-decoding capacity over linear-sized fields, 2024. doi:10.48550/arXiv.2304.09445.
3. Mitali Bafna and Dor Minzer. Characterizing direct product testing via coboundary expansion, 2024. doi:10.48550/arXiv.2308.09668.
4. Paul Beame, Shayan Oveis Gharan, and Xin Yang. On the bias of reed-muller codes over odd prime fields, 2018. arXiv:1806.06973.
5. Ido Ben-Eliezer, Rani Hod, and Shachar Lovett. Random low degree polynomials are hard to approximate. In *Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, APPROX '09 / RANDOM '09, pages 366–377, Berlin, Heidelberg, 2009. Springer-Verlag. doi:10.1007/978-3-642-03685-9_28.
6. Inbar Ben Yaacov, Yotam Dikstein, and Gal Maor. Sparse high dimensional expanders via local lifts, 2024. doi:10.48550/arXiv.2405.19191.
7. Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, Pooya Hatami, and Shachar Lovett. Every locally characterized affine-invariant property is testable, 2013. arXiv:1212.3849.
8. Abhishek Bhowmick and Shachar Lovett. List decoding reed-muller codes over small fields, 2014. arXiv:1407.3433.
9. 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. arXiv:1506.02047.
10. Andrej Bogdanov. Pseudorandom generators for low degree polynomials. In *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '05, pages 21–30, New York, NY, USA, 2005. Association for Computing Machinery. doi:10.1145/1060590.1060594.
11. Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 41–51, 2007. doi:10.1109/FOCS.2007.42.

12 Joshua Brakensiek, Manik Dhar, and Sivakanth Gopi. Generalized gm-mds: Polynomial codes are higher order mds, 2024. doi:10.48550/arXiv.2310.12888.

13 Joshua Brakensiek, Manik Dhar, Sivakanth Gopi, and Zihan Zhang. Ag codes achieve list-decoding capacity over constant-sized fields, 2024. doi:10.48550/arXiv.2310.12898.

14 Joshua Brakensiek, Sivakanth Gopi, and Visu Makam. Generic reed-solomon codes achieve list-decoding capacity, 2024. arXiv:2206.05256.

15 Nader Bshouty. Testers and their applications. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 327–352, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2554797.2554828.

16 Gil Cohen and Amnon Ta-Shma. Pseudorandom generators for low degree polynomials from algebraic geometry codes. *Electron. Colloquium Comput. Complex.*, TR13, 2013. URL: https://api.semanticscholar.org/CorpusID:13155686.

17 Harm Derksen and Emanuele Viola. Fooling polynomials using invariant theory. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 399–406. IEEE, 2022. doi:10.1109/FOCS54457.2022.00045.

18 Yotam Dikstein. New high dimensional expanders from covers, 2022. doi:10.48550/arXiv.2211.13568.

19 Yotam Dikstein and Irit Dinur. Agreement theorems for high dimensional expanders in the small soundness regime: the role of covers, 2024. doi:10.48550/arXiv.2308.09582.

20 Dean Doron, Amnon Ta-Shma, and Roei Tell. On hitting-set generators for polynomials that vanish rarely. *Comput. Complex.*, 31(2):16, 2022. doi:10.1007/S00037-022-00229-2.

21 Zeev Dvir and Amir Shpilka. Noisy interpolating sets for low degree polynomials. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 140–148, 2008. doi:10.1109/CCC.2008.14.

22 Ashish Dwivedi, Zeyu Guo, and Ben Lee Volk. Optimal pseudorandom generators for low-degree polynomials over moderately large fields, 2024. doi:10.48550/arXiv.2402.11915.

23 Alexander Golovnev, Zeyu Guo, Pooya Hatami, Satyajeet Nagargoje, and Chao Yan. Hilbert functions and low-degree randomness extractors, 2024. doi:10.48550/arXiv.2405.10277.

24 Parikshit Gopalan, Adam R. Klivans, and David Zuckerman. List-decoding reed-muller codes over small fields. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 265–274, New York, NY, USA, 2008. Association for Computing Machinery. doi:10.1145/1374376.1374417.

25 Omri Gotlib, Tali Kaufman, and Shachar Lovett. List decoding quotient reed-muller codes, 2025. doi:10.48550/arXiv.2502.15650.

26 Roy Gotlib and Tali Kaufman. List agreement expansion from coboundary expansion, 2022. doi:10.48550/arXiv.2210.15714.

27 W. T. Gowers and Thomas Karam. Equidistribution of high-rank polynomials with variables restricted to subsets of $\mathbb{F}_p$, 2022. arXiv:2209.04932.

28 Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the gowers norms, 2007. arXiv:0711.3191.

29 Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions, 2007. arXiv:math/0404188.

30 Venkatesan Guruswami, Lingfei Jin, and Chaoping Xing. Efficiently list-decodable punctured reed-muller codes, 2017. arXiv:1508.00603.

31 Venkatesan Guruswami and Chaoping Xing. Hitting sets for low-degree polynomials with optimal density. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 161–168, June 2014. doi:10.1109/CCC.2014.24.

32 Hamed Hatami, Pooya Hatami, and Shachar Lovett. *Higher-order Fourier Analysis and Applications*. Now Foundation and Trends, January 2019. doi:10.1561/9781680835939.

33 Tali Kaufman and Shachar Lovett. Worst case to average case reductions for polynomials, 2008. arXiv:0806.4535.

34 David Kazhdan and Tamar Ziegler. Polynomial functions as splines, 2018. arXiv:1712.09047.

35 David Kazhdan and Tamar Ziegler. Extending weakly polynomial functions from high rank varieties, 2019. arXiv:1808.09439.

36 David Kazhdan and Tamar Ziegler. Properties of high rank subvarieties of affine spaces, 2020. arXiv:1902.00767.

37 Amichai Lampert and Tamar Ziegler. Relative rank and regularization, 2021. arXiv:2106.03933.

38 Shachar Lovett. Unconditional pseudorandom generators for low degree polynomials. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 557–562, New York, NY, USA, 2008. Association for Computing Machinery. doi:10.1145/1374376.1374455.

39 Shachar Lovett. MDS matrices over small fields: A proof of the GM-MDS conjecture. *CoRR*, abs/1803.02523, 2018. arXiv:1803.02523.

40 Chi-Jen Lu. Hitting set generators for sparse polynomials over any finite fields. In *2012 IEEE 27th Conference on Computational Complexity*, pages 280–286, 2012. doi:10.1109/CCC.2012.20.

41 Shay Moran and Cyrus Rashtchian. Shattered sets and the hilbert function, 2020. arXiv:1511.08245.

42 Wolfgang M. Schmidt. The density of integer points on homogeneous varieties. *Acta Mathematica*, 154(3-4):243–296, 1985. doi:10.1007/BF02392473.

43 J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, October 1980. doi:10.1145/322217.322225.

44 Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 124–127, 2008. doi:10.1109/CCC.2008.16.

45 Hikmet Yildiz and Babak Hassibi. Optimum linear codes with support constraints over small fields. *CoRR*, abs/1803.03752, 2018. arXiv:1803.03752.

46 Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation*, 1979. URL: https://api.semanticscholar.org/CorpusID:15629042.

A Equidistribution of Functions

Assume we have a collection of functions $(F_1, ..., F_c)$, where $F_i : A \to \mathbb{F}$ for some finite set A. We are interested in showing that the functions are equidistributed, which means that their values behave close to independent random variables. We begin by formulating this definition:

▶ **Definition 80** (Equidistribution of Functions). *Given $\epsilon > 0$ and $A \subseteq \mathbb{F}^n$, we say a collection of functions $\mathfrak{F} = (F_1, ..., F_c)$ where $F_i : A \to \mathbb{F}$ is ϵ-equidistributed in A if for all $\vec{\alpha} = (\alpha_1, ..., \alpha_c) \in \mathbb{F}^c$ we have:*

$$\Pr_{x \in A}[(F_1(x), ..., F_c(x)) = \vec{\alpha}] = \frac{1}{|\mathbb{F}|^c} \pm \epsilon$$

The following is a standard lemma that shows that if every linear combination of a collection of functions has low bias, the collection is equidistributed. We repeat the steps of the proof of [32, Lemma 7.24], but here, we think of A as any finite set (and not particularly $\mathbb{F}^n$):

▶ **Lemma 81.** *Let $\epsilon > 0$, and let A be a finite set. Let $\mathfrak{F} = (F_1, ..., F_c)$ be a collection of functions defined over A, i.e. $F_i : A \to \mathbb{F}$. Assume each linear combination of the collection has low bias, i.e. for each $\lambda = (\lambda_1, ..., \lambda_c) \in \mathbb{F}^c$ such that $\lambda \neq \vec{0}$ we have:*

$$bias_{x \in A}(\sum_{i=1}^{c} \lambda_i F_i) < \epsilon$$

Then, the collection $\mathfrak{F}$ is ϵ-equidistributed over A.
In particular, for $\epsilon < \frac{1}{|\mathbb{F}|^c}$, the lemma shows that each atom of $\mathfrak{F}$ is not empty i.e. for all $\vec{\alpha}$ there is some $x \in A$ such that $(F_1(x), ..., F_c(x)) = \vec{\alpha}$.

Proof. We wish to show that for each $\vec{\alpha} \in \mathbb{F}^c$ we have:

$$\Pr_{x \in A}\left[(F_1(x), ..., F_c(x)) = \vec{\alpha}\right] = \frac{1}{|\mathbb{F}|} \pm \epsilon$$

We express the fraction of inputs that are in the atom $\vec{\alpha}$ the following way:

$$\Pr_{x \in A}\left[(F_1(x), ..., F_c(x))\right] = \mathbb{E}_{x \in A}\left[\prod_{i=1}^{c} 1_{[F_i(x)=\alpha_i]}\right]$$

We use the fact that for every $0 \neq x \in \mathbb{F}$, we have $\sum_{\lambda=0}^{p-1} e\left[\lambda x\right] = 0$, and if $x = 0$ we have $\sum_{\lambda=0}^{p-1} e\left[\lambda x\right] = p$. Therefore, the expression above equals:

$$= \mathbb{E}_{x \in A}\left[\prod_{i=1}^{c}\left(\frac{1}{p} \cdot \sum_{\lambda_i=0}^{p-1} e\left[\lambda_i(F_i(x) - \alpha_i)\right]\right)\right] = \frac{1}{p^c} \cdot \mathbb{E}_{x \in A}\left[\prod_{i=1}^{c}\sum_{\lambda_i=0}^{p-1} e\left[\lambda_i(F_i(x) - \alpha_i)\right]\right]$$

By the definition of character functions, we have that $e\left[a+b\right] = e\left[a\right] \cdot e\left[b\right]$, and therefore the expression above equals:

$$\frac{1}{p^c} \cdot \sum_{(\lambda_1,...,\lambda_c) \in \prod_{i=1}^{c}[0,p-1]}\left(\mathbb{E}_{x \in A}\left[e\left[\sum_{i=0}^{c}\lambda_i(F_i(x) - \alpha_i)\right]\right]\right)$$

Now, we use the fact that:

$$bias_{x \in A}(\sum_{i=1}^{c}(\lambda_i(F_i(x) - \alpha_i)) = bias_{x \in A}(\sum_{i=1}^{c}(\lambda_i F_i(x))) < \epsilon$$

and get that:

$$\Pr_{x \in A}\left[(F_1(x), ..., F_c(x)) = \vec{\alpha}\right] = \frac{1}{p^c} \cdot \left(1 \pm \epsilon \prod_{i=1}^{c} p\right) = \frac{1}{|\mathbb{F}|^c} \pm \epsilon$$ ◀

B Comparing Ranks

In this section, we compare the definition of rank we used in this paper to another definition of rank used implicitly throughout this paper. This comparison is crucial, as there is no universally accepted definition of rank; different theorems presented throughout this paper employ distinct definitions. We demonstrate that our definition is sufficiently comprehensive, in that a polynomial (or a factor) classified as having high rank according to our criteria also exhibits high rank according to the second implicitly-used definition. While in many cases the comparison may appear straightforward, we include it for the sake of completeness.
Specifically, we compare our definition of rank with the definition established in [37]. The paper [37] extended the original definition of rank that was presented in [42], to include also the concept of relative rank. It is important to note that this definition is specifically defined to subsets $\tilde{X} \subseteq \mathbb{F}^n$ that can be expressed as sets in the form $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right)$ for some set of polynomials $\tilde{\mathcal{L}}$, and not to a general set $\tilde{X} \subseteq \mathbb{F}^n$.
First, we present a useful notation that is used in the definition presented in [37]:

▶ **Notation** (Largest Degree Homogenous Part). *For a polynomial P of degree d, we denote by $\mathrm{h}(P)$ its degree-d homogenous component. In other words, $\mathrm{h}(P)$ is the sum of all the monomials of P of degree exactly d. For a set of polynomials $\mathcal{P} = \{P_1, ..., P_c\}$, we define $\mathrm{h}(\mathcal{P}) \coloneqq \{\mathrm{h}(P_i) | i = 1, ..., c\}$.*

Next, we present the exact definition of rank for a polynomial:

▶ **Definition 82** (Schmidt Rank of a Polynomial). *The schmidt rank of a homogenous polynomial $P : \mathbb{F}^n \to \mathbb{F}$, noted as $schmrank(P)$, is the minimal r such that there exist $(Q_i, H_i)_{i\in[r]}$ with $\deg Q_i, \deg H_i < \deg P$ such that:*

$$P(x) = \sum_{i=1}^{r} (Q(x) \cdot H(x))$$

For a general polynomial P of degree d, we set its rank to be the rank of its degree-d homogenous component, i.e. $schmrank(P) \coloneqq schmrank(\mathrm{h}(P))$.

▶ Remark 83 (High rank implies high schmidt rank). If $rank(P) \geq 2 \cdot r + 1$ for some constant $r \in \mathbb{N}$, then $schmrank(P) \geq r$.

Proof. For homogenous polynomial P, assume $schmrank(P) < r$. Then, there exist $r' < r$ such that there exist $(Q_i, H_i)_{i=1}^{r'}$ with $\deg Q_i, \deg H_i < \deg P$ such that:

$$P(x) = \sum_{i=1}^{r'} (Q(x) \cdot H(x))$$

Then we can choose $\Gamma : \mathbb{F}^{2r'} \to \mathbb{F}$ to be a sum of multiples of each two consecutive variables to get that $P(x) = \Gamma(Q_1(x), H_1(x), ..., Q_{r'}(x), H_{r'}(x))$, where the polynomials are from a degree $< \deg(P)$. This means that $rank(P) \leq 2r' < 2r$ as we requested.
If we do not assume P is homogenous, by adding $P - \mathrm{h}(P)$ as an input to Γ, one can create a $\Gamma' : \mathbb{F}^{2r'+1} \to \mathbb{F}$ which equals to P when substituting the inputs with some polynomials with degree $< \deg P$, which concludes the proof in a similar way. ◀

Next, we present the definition of Schmidt rank of a factor as defined in [37].

▶ **Definition 84** (Schmidt Rank of a Factor). *For a factor of homogenous polynomials $\mathcal{P} = (P_1, ..., P_c)$, the schmidt rank of the factor is defined as:*

$$schmrank(\mathcal{P}) \coloneqq \min\left(schmrank\left(\sum_{i=1}^{c} \lambda_i P_i\right) \middle| 0 \neq (\lambda_1, ..., \lambda_c) \in \mathbb{F}^c \right)$$

Similarly, for a factor of general polynomials $\mathcal{P}$, we set its rank to be the rank of its matching homogenous-factor, i.e. $schmrank(\mathcal{P}) \coloneqq schmrank(\mathrm{h}(\mathcal{P}))$ For a factor $\mathcal{B}$ generated by $\mathcal{P}$, we define $schmrank(\mathcal{B}) \coloneqq schmrank(\mathcal{P})$.

To establish the equivalence of this definition with the one employed throughout the paper, we must first acknowledge two key distinctions between the definitions. The first distinction is that this definition focuses on the largest-degree homogeneous components of the polynomials involved in the factor, rather than considering linear combinations of polynomials from the factor. The second distinction pertains to the treatment of d in the computation of d-rank of each linear combination. This definition uses the degree of the linear combination directly to calculate the rank that participates in the minimum, in contrast to our definition which uses $\max_i \deg(\lambda_i P_i)$. Despite these differences, we will demonstrate that both definitions ultimately yield a similar rank assessment, thereby affirming their equivalence.

▶ Remark 85 (High Rank Implies High Schmidt Rank for Factors). Let $\mathcal{P} = (P_1, ..., P_c)$ be a set of polynomials and let $r \in \mathbb{N}$ be a positive integer, i.e. $r > 0$. If $rank(\mathcal{P}) \geq 2 \cdot r + 1$, then $schmrank(\mathcal{P}) \geq r$.

Proof. Assume that $schmrank(\mathcal{P}) \leq r$ for $r > 0$. We will show that $rank(\mathcal{P}) \leq 2r + 1$. By definition, there exists a linear combination of polynomials in $\mathrm{h}(\mathcal{P})$ with rank $\leq r$. In other words, there exists $\vec{0} \neq \lambda \in \mathbb{F}^c$ such that $schmrank(\sum_{i=1}^{c} \lambda_i \mathrm{h}(P_i)) \leq r$. Denote $\vec{P_h} \coloneqq \sum_{i=1}^{c} \lambda_i \mathrm{h}(P_i)$. As was shown in a previous remark, a rank of a polynomial is smaller than its schmidt rank up to a constant factor, thus $rank\left(\vec{P_h}\right) \leq 2r + 1$ (see Remark 83). Next, we denote $\vec{P} \coloneqq \sum_{i=1}^{c} \lambda_i P_i$, and $d_M \coloneqq \max_{i \in [c]} \lambda_i P_i$. Note that $\deg(\vec{P}) \leq d_M$. We wish to show that $rank_{d_M}\left(\vec{P}\right) \leq 2r + 1$. First, we observe that the d_M-degree homogenous component of $\vec{P_h}$ equals the d_M-degree homogenous component of $\vec{P}$. This is true because every highest-degree component of polynomials in the linear combination that generated $\vec{P}$, also exists in the linear combination that generates $\vec{P_h}$. In particular, all homogenous components of degree d_M exists in both linear combinations $\vec{P_h}$ and $\vec{P}$. Therefore, if the degree of $\vec{P}$ equals d_M, we have that $rank_{d_M}\left(\vec{P}\right) = rank\left(\vec{P}\right) \geq 2r + 1$. Otherwise, if $\deg(\vec{P}) < d_M$, then $rank_{d_M}\left(\vec{P}\right) = 1 \leq 2r + 1$. This completes the proof. ◀

▶ Note. In the case discussed above, if $\deg(\vec{P}) < d_M$, then $schmrank(\mathcal{P}) = 0$.

Proof. Assume that $\deg(\vec{P}) < d_M$. Therefore, the degree of the linear combination $\vec{P} = \sum_{i=1}^{c} \lambda_i P_i$ is strictly smaller than the degree of at least one of the polynomials participating in it. Denote by $\vec{\lambda}^\star$ the sub-combination of $\vec{\lambda}$ that consists only the polynomials that participated in $\vec{P}$ that are of degree $= d_M$. Trivially, $\vec{\lambda}^\star \neq \vec{0}$. Additionally, we have $\deg(\sum_{i=1}^{c} \lambda_i^\star P_i) < d_M$. Now, we use the following observation: the linear combination above, when summing only the homogenous components of each polynomial, equals 0, i.e. $\sum_{i=1}^{c} \lambda_i^\star \mathrm{h}(P_i) \equiv 0$. By this, we found a linear combination of $\mathrm{h}(\mathcal{P})$ that is $\equiv 0$. Thus by definition, we have $schmrank(\mathcal{P}) = 0$. ◀

▶ Note. This shows that if we compare only the differences in the definition of rank of a factor, i.e. the focus on linear combinations of the largest-degree homogenous components in contrast to the use of the maximal degree d-rank, the two definitions for a rank of a factor are equal up to ± 1 (in case we use the same definition of rank for a single polynomial). To avoid confusion, we omit the exact definitions and respective proof.

We now present the definition of relative rank as stated in [37, Definition 1.6]: We remind the reader that this definition is specifically defined to subsets $\tilde{X} \subseteq \mathbb{F}^n$ that can be expressed by $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right)$ for some set of polynomials $\tilde{\mathcal{L}}$, and not to a general set $\tilde{X} \subseteq \mathbb{F}^n$.

▶ **Definition 86** (Relative Schmidt Rank of a Polynomial)**.** *The relative schmidt rank of a homogeneous polynomial P relative to a collection of homogeneous polynomials $\tilde{\mathcal{L}} = (L_1, \ldots, L_{\tilde{c}})$ is*

$$shcmrank_{\tilde{\mathcal{L}}}(P) \coloneqq \min\left\{ schmrank\left(P + \sum_{i=1}^{\tilde{c}} R_i L_i\right) \middle| \deg(L_i) + \deg(R_i) \leq \deg(P), \forall i \in [\tilde{c}] \right\}$$

Note that whenever $\deg L_i > \deg P$, this implies $R_i = 0$.
For general polynomial P and general collection of polynomials $\tilde{\mathcal{L}}$, we define the schmidt rank of the former in respect to the latter by the relative rank of their largest-degree homogenous component, i.e. $shcmrank_{\tilde{\mathcal{L}}}(P) \coloneqq shcmrank_{\mathrm{h}(\tilde{\mathcal{L}})}(\mathrm{h}(P))$.

▶ Remark 87 (High Relative Rank ⇒ High Relative Schmidt Rank). Let P and $\tilde{\mathcal{L}} = \{L_1, ..., L_{\tilde{c}}\}$ be polynomials, and let $\tilde{X} \subseteq \mathbb{F}^n$ be defined as $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right)$.

If $rank_{\tilde{X}}(P) \geq 2 \cdot r + 2$ for some constant $r \in \mathbb{N}$, then $shcmrank_{\tilde{\mathcal{L}}}(P) \geq r$.

Proof. Let P and $L_1, ..., L_{\tilde{c}}$ be polynomials. Assume that $shcmrank_{\tilde{\mathcal{L}}}(P) \leq r$. Then, there exists $R_1, ..., R_{\tilde{c}}$ with $\deg(L_i) + \deg(R_i) \leq \deg(P)$ for all $i \in [\tilde{c}]$ such that:

$$schmrank\left(\mathrm{h}(P) + \sum_{i=1}^{\tilde{c}} R_i \mathrm{h}(L_i)\right) \leq r$$

Denote $\overline{P_h} \coloneqq \sum_{i=1}^{\tilde{c}} R_i \mathrm{h}(L_i)$. As we have shown earlier, a rank of a polynomial is smaller than its schmidt rank up to a constant factor (See Remark 83). Thus:

$$rank\left(\mathrm{h}(P) + \overline{P_h}\right) \leq 2 \cdot schmrank\left(\mathrm{h}(P) + \overline{P_h}\right) + 1 \leq 2 \cdot shcmrank_{\tilde{X}}(P) + 1 = 2r + 1$$

Next, we denote the respective remainder polynomial for the non-homogenous analogue, i.e. $\overline{P} \coloneqq \sum_{i=1}^{\tilde{c}} R_i L_i$. By observing the highest degree homogenous component of each summand, one can see that $\mathrm{h}(P + \overline{P}) = \mathrm{h}(\mathrm{h}(P) + \overline{P_h})$. Therefore, by adding to the decomposition the non higest-degree-homogenous-component, one can see that:

$$rank\left(P + \overline{P}\right) \leq rank\left(\mathrm{h}(P) + \overline{P_h}\right) + 1 \leq 2r + 2$$

This completes the proof as $rank_{\tilde{X}}(P) \leq rank\left(P + \overline{P}\right) \leq 2r + 2$. ◀

▶ Remark 88 (Relative Schmidt Rank over Varieties of High Degree). If the polynomials defining the variety $\tilde{\mathcal{L}} = (L_1, ..., L_{\tilde{c}})$ are of degree $> \deg(P)$, then, $shcmrank_{\tilde{\mathcal{L}}}(P) = schmrank(P)$. This is true because in this case, in the calculation of the minimum in the definition of relative schmidt rank, we must have $R_i = 1$ for all $i \in [\tilde{c}]$ and therefore the minimum above is simply $rank(P)$.
Note that a similar statement holds for factors aswell. If $\mathcal{P} = (P_1, ..., P_c)$ is a factor of degree d, then if all the polynomials in $\tilde{\mathcal{L}}$ have degree $> d$, then the statement above is also true i.e. $shcmrank_{\tilde{\mathcal{L}}}(\mathcal{P}) = schmrank(\mathcal{P})$. This is true because for every linear combination of $\mathcal{P}$ has degree $\leq d$ and therefore its relative schimdt rank equals its rank.

Finally, we present the extension of the definition of relative rank for polynomials factors:

▶ **Definition 89** (Relative Schmidt Rank of a Factor). *The relative rank of a set of homogenous polynomials $\mathcal{P} = \{P_1, ..., P_c\}$ relative to another collection of polynomials $\tilde{\mathcal{L}} = \{L_1, ..., L_{\tilde{c}}\}$ is defined as:*

$$shcmrank_{\tilde{\mathcal{L}}}(\mathcal{P}) \coloneqq \min\left\{ shcmrank_{\tilde{\mathcal{L}}}\left(\sum_{i=1}^{c} \lambda_i P_i\right) \middle| \vec{0} \neq (\lambda_1, ..., \lambda_c) \in \mathbb{F}^c \right\}$$

If $\mathcal{P}$ is a general collection of polynomials, then $shcmrank_{\tilde{\mathcal{L}}}(\mathcal{P}) \coloneqq shcmrank_{\tilde{\mathcal{L}}}(\mathrm{h}(\mathcal{P}))$.
For a factor $\mathcal{B}$ generated by a set of polynomials $\mathcal{P}$, we define its schmidt rank relative to $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right)$ to be $shcmrank_{\tilde{X}}(\mathcal{B}) \coloneqq shcmrank_{\tilde{\mathcal{L}}}(\mathcal{P})$.

▶ Remark 90. Let $\mathcal{P} = \{P_1, ..., P_c\}$ and $\tilde{\mathcal{L}} = \{L_1, ..., L_{\tilde{c}}\}$ be sets of polynomials, and let $\tilde{X} \subseteq \mathbb{F}^n$ be defined as $\tilde{X} = Z\left(\tilde{\mathcal{L}}\right)$. Additionally, let $r \in \mathbb{N}$ such that $r > 0$. If $rank_{\tilde{X}}(\mathcal{P}) \geq 2 \cdot r + 2$ for some constant $r \in \mathbb{N}$, then $shcmrank_{\tilde{\mathcal{L}}}(P) \geq r$.

Proof. Assume that $shcmrank_{\tilde{\mathcal{L}}}(\mathcal{P}) \leq r$. We will show that $rank_{\tilde{X}}(\mathcal{P}) \leq 2r+2$. Let $\vec{0} \neq \vec{\lambda} \in \mathbb{F}^c$ be some vector of coefficients. Let $\vec{P} \coloneqq \sum_{i=1}^{c} \lambda_i P_i$ and $\vec{P_h} \coloneqq \sum_{i=1}^{c} \lambda_i \mathrm{h}(P_i)$ be the linear combinations of polynomials in $\mathcal{P}$ and $\mathrm{h}(\mathcal{P})$ with coefficients $\vec{\lambda}$ respectively, and let $d_M \coloneqq \max_{i\in[c]} \deg(\lambda_i P_i)$. Additionally, denote $\hat{r} \coloneqq shcmrank_{\tilde{\mathcal{L}}}\left(\vec{P_h}\right) \leq r$. It is enough to show that $rank_{d_M,\tilde{X}}\left(\vec{P}\right) \leq 2\hat{r}+2$, If $\deg(\vec{P}) < d_M$, then $rank_{d_M,\tilde{X}}\left(\vec{P}\right) = 1 \leq 2r+2$. Otherwise, if $\deg(\vec{P}) = d_M$, then the remark follows from Remark 87 as:

$$rank_{d_M,\tilde{X}}\left(\vec{P}\right) = rank_{\tilde{X}}\left(\vec{P}\right) \leq 2 \cdot shcmrank_{\tilde{\mathcal{L}}}\left(\vec{P}\right) + 2$$

Where:

$$shcmrank_{\tilde{\mathcal{L}}}\left(\vec{P}\right) shcmrank_{\tilde{\mathcal{L}}}\left(\mathrm{h}(\vec{P})\right) = shcmrank_{\tilde{\mathcal{L}}}\left(\mathrm{h}(\vec{P_h})\right) = shcmrank_{\tilde{\mathcal{L}}}\left(\vec{P_h}\right) = \hat{r} \quad \blacktriangleleft$$