ON THE LINEAR COMPLEXITY OF SUBSETS OF $\mathbb{F}_p^n$ BOUNDED VC_2-DIMENSION

H. SHEATS AND C. TERRY

ABSTRACT. Previous work of the second author and Wolf showed that given a set $A \subseteq \mathbb{F}_p^n$ of bounded VC_2-dimension, there is a high rank quadratic factor $\mathcal{B}$ of bounded complexity such that A is approximately equal to a union of atoms of $\mathcal{B}$. The main ingredients in that proof were a counting lemma for a local version of the U^3-norm, along with a quadratic arithmetic regularity lemma of Green and Tao. This approach yielded bounds of tower type bounds on the linear and quadratic complexities. It was later shown by the same authors that the quadratic complexity can be improved to $\log_p(\mathrm{poly}(\epsilon^{-1}))$, however that proof provided no improvement on the linear component. In this paper we prove that the bound on the linear complexity can be improved to a triple exponential in the case of linear rank functions, and a quadruple exponential for polynomial rank functions of higher degree.

Our strategy is based on the one developed by Gishboliner, Wigderson, and Shapira to prove the analogous result in the hypergraph setting. Step 1 is to prove a "cylinder" version of the quadratic arithmetic regularity lemma, which says that given a set $A \subseteq G = \mathbb{F}_p^n$, there is a partition of G into atoms of (possibly distinct) quadratic factors of high rank and bounded complexity, so that most atoms in the partition are uniform with respect to the set A, in the sense of a certain local U^3 norm. Step 2 is to show that if A has bounded VC_2-dimension, then it has density near 0 or 1 on all atoms which are uniform in the sense of Step 1. We then obtain the structure theorem for A by taking a high rank, common refinement of the factors appearing in the previous steps. Step 1 relies on a recent local version of the U^3 inverse theorem due to Prendiville, and is necessarily phrased in terms of a local U^3 norm implicit in that paper. On the other hand, Step 2 relies on a counting lemma for a different local U^3 due to Terry and Wolf, which we prove here is approximately the same as the local U^3 norm used in Step 1.

1. INTRODUCTION

There is now a large body of work studying the behavior of bounds in Szemerédi's graph regularity lemma and its generalizations. We will focus this introduction on the parts of the story most relevant to the current paper, namely arithmetic regularity and strong hypergraph regularity, and refer the interested reader to [16, 21, 30, 41, 44, 45, 51, 56] for more background.

1.1. **Arithmetic regularity.**

Analogues of the graph regularity lemma for subsets of groups were first developed by Green in [36]. In $\mathbb{F}_p^n$, this regularity lemma says that for

The second author was partially supported by NSF CAREER Award DMS-2115518 and a Sloan Research Fellowship.

any subset $A \subseteq \mathbb{F}_p^n$, there exists a subgroup $H \leq \mathbb{F}_p^n$ of bounded index, so that most cosets of H are "ϵ-regular" with respect to A, where "ϵ-regular" is defined in terms of the Gowers U^2-norm, localized to the coset in question. The bounds in the graph regularity lemma are known to be necessarily of tower type (see [21, 31]), and Green proved the same is true of the bounds in this arithmetic analogue [36].

By the mid 2010's, several papers had appeared showing that in the setting of graphs, various combinatorial restrictions yield improved versions of the regularity lemma (see e.g. [2, 20, 22, 42, 43]). The first result along these lines in the arithmetic setting is due to the second author and Wolf, who showed that k-stable subsets of $\mathbb{F}_p^n$ admit ϵ-regular subgroups with index $\exp_p(\epsilon^{-O_k(1)})$ and no irregular cosets. Several more papers on "tame" arithmetic regularity lemmas have since appeared, including extensions and refinements in the stable case [11, 13, 64], in addition to results in the more general setting of bounded VC-dimension [3, 14, 15, 54].[1] These results are all naturally stated as structure theorems for the set in question. In $\mathbb{F}_p^n$, they say the set A is well approximated by a union of cosets of a bounded index subgroup, with the distinctions among them boiling down to differences in the meaning of "well approximated."

1.2. **Hypergraph regularity.** Several variants of the regularity lemma exist for hypergraphs (see [47] for a survey of the 3-uniform case). The version relevant to this paper is the "strong" form of the regularity lemma for 3-uniform hypergraphs, as developed by subsets of the authors Gowers, Frankl, Kohayakawa, Nagle, Rödl, Skokan, and Schacht [25, 33, 34, 46, 49, 50]. For results related to the analogous problem for *weak* regularity, we refer the reader to [1, 9, 10, 20, 22, 23, 26, 60, 61].

A strong regular decomposition for a 3-uniform hypergraph consists of a pair $\mathcal{P} = (\mathcal{P}_{vert}, \mathcal{P}_{pairs})$, where $\mathcal{P}_{vert}$ partitions the vertex set, and $\mathcal{P}_{pairs}$ partitions the pairs of vertices. This gives rise to two complexity parameters, t for the vertex partition, and s for the pairs partition. The statement of the regularity lemma contains a constant error parameter $\epsilon > 0$ in addition to an "error function" $\psi : \mathbb{N} \to (0, 1]$. In practice, the sets in $\mathcal{P}_{pairs}$ are required to be $\psi(s)$-regular as graph relations, where s is the pairs complexity. In essentially all applications, it suffices to take ψ a sufficiently fast growing polynomial in s^{-1}. The first major result on the bounds for strong hypergraph regularity was due to Moshkovitz and Shapira [45], who showed that even for polynomial choices of ψ, the vertex partition must grow at a wowzer rate, matching the corresponding upper bound from the proofs of the regularity lemma.[2] This opened up the field to a new problem: which combinatorial restrictions give rise to better bounds in the strong regularity lemma for 3-uniform hyergraphs?

The first result along these lines is due to the second author, who showed 3-uniform hypergraphs of bounded VC_2-dimension (a higher arity analogue of VC-dimension)[3] admit

[1] While the setting of bounded VC-dimension is more general than the stable setting, the structure theorems obtained in the stable setting are stronger.

[2] In fact their theorems are much more general, giving lower bounds for k-uniform hypergraphs for all $k \geq 3$.

[3] See Definition 2.2.

polynomial bounds on the pairs partition. A systematic investigation of these problems in the 3-uniform case was later undertaken by the second author in [62,63], with [62] studying the growth of $|\mathcal{P}_{vert}|$ and [63] studying the growth of $|\mathcal{P}_{pairs}|$. This work left open several interesting questions. The two most important problems related to the vertex partition were subsequently resolved by Gishboliner, Shapira, and Wigderson in [26] and [27]. As the latter is important for this paper, we explain its contribution in more detail below.

When studying the bounds in the strong hypergraph regularity lemma for 3-uniform hypergraphs, a choice must be made for what ψ are allowed. In [62, 63], the choice was made to study the problem where arbitrary ψ are allowed. Interestingly, only the jump to the fastest growth rate was sensitive to this choice. In the case of the vertex partition, combining [62] with [26][4] implies that when arbitrary ψ are allowed, there is a jump all the way from exponential to wowzer growth, characterized by whether or not a property is "close" to finite slicewise VC-dimension.[5] This suggests that in the regime where arbitrary ψ are allowed, VC_2-dimension does not characterize a jump. This was surprising, as VC_2-dimension characterizes many other dichotomies for 3-uniform hypergraphs, including in the growth of the pairs partition [58, 59, 63], and in the labeled enumeration problem [58]. On the other hand, the work in [62] does not address the question of what happens with the vertex partition between the exponential and wowzer ranges, in the regime where only reasonably slow ψ are allowed. This is a problem of great interest, as such ψ are the ones used in applications.

Gishboliner, Shapira, and Wigderson resolved this problem by showing that after modestly restricting the growth rate of ψ, an additional jump appears between exponential and wowzer, characterized by VC_2-dimension. One direction of this is implicit in prior work of Moshkovitz and Shapira (i.e. that unbounded VC_2-dimension implies wowzer growth, even with polynomial ψ). The contribution of [27] is the other direction. Specifically, they prove that when the VC_2-dimension is bounded and ψ is sub-tower, there is a double tower bound on the vertex partition. We outline their proof in broad strokes below, as it guides our strategy in this paper:

Step 1: Prove a Duke-Leffman-Rodl "cylinder" style hypergraph regularity lemma, which admits better bounds (namely tower) when the ψ is assumed to be polynomial.

Step 2: When the hypergraph has bounded VC_2-dimension, deduce it has density near 0 or 1 on all regular triads from Step 1. This is a straightforward corollary of Gowers' counting lemma, which crucially requires only polynomially growing ψ (see [33]).

Step 3: Take the common refinement of all the components arising from the cylinder decompositions from Step 1.

[4]The work of [62] uses an earlier paper, [60], to prove a jump from double exponential to wowzer. This can be improved to a jump from single exponential to wowzer by replacing the use of [60] there with [26].
[5]The jump to wowzer in [62] uses the fact that the strong graph regularity lemma requires wowzer bounds (see [16, 51]). A similar result was recently proved for the strong linear arithmetic regularity lemma by Gladkova [28].

Step 4: Apply Szemerédi's regularity lemma to regularize the resulting pairs partition. As densities close to 0 and 1 are mostly preserved under refinements, the resulting decomposition still has the property that most triads have density near 0 or 1. As triads with density near 0 or 1 are always regular, this yields a regular decomposition.

Because Step 1 yields a tower bound, and Step 4 applies the graph regularity lemma one additional time, this strategy yields a double tower bound.

1.3. **Quadratic arithmetic regularity and main results.** There exists a hierarchy of arithmetic regularity lemmas in rough correspondence with the strong regularity lemmas for k-uniform hypergraphs. The arithmetic analogue of the strong regularity lemma for 3-uniform hypergraphs is the *quadratic arithmetic regularity lemma*. In $\mathbb{F}_p^n$, a regularity lemma of this type comes equipped with *quadratic factor*, consisting of a *linear* component and a *quadratic* component (see Subsection 2.2 for detailed definitions). This gives rise to two complexity parameters, ℓ for the linear component, and q for the quadratic component. The quadratic arithmetic regularity lemma making the analogy to hypergraph regularity most explicit is a formulation from [67], which says that given any $A \subseteq \mathbb{F}_p^n$, there is a high rank quadratic factor $\mathcal{B}$ of bounded complexity so that on most atoms of $\mathcal{B}$, the set A is uniform in the sense of a local U^3-norm defined there, which we will denote by $\|\cdot\|_{U^3(d)}^{TW}$ (see Subsection 3.2 for more details).

Theorem 1.1 (Proposition 4.2 in [67][6])**.** *For all odd primes p, all $\delta \in (0,1)$, and all growth functions[7] ρ, there is a constant $M(p,\delta,\rho)$ such that the following holds.*

For all $A \subseteq \mathbb{F}_p^n$, there are integers ℓ, q and a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $\rho(\ell + q)$ such that

(1) $\ell, q \leq M$

(2) for at least a $(1-\delta)$-fraction of the tuples $d \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$, we have

$$\|1_A - \alpha_{B(\Sigma(d))}\|_{U^3(d)}^{TW} < \delta,$$

where $\alpha_{B(\Sigma(d))}$ is the density of A on the atom $B(\Sigma(d))$.

In [67], Theorem 1.1 is deduced from a quadratic arithmetic regularity lemma of Green and Tao from [38]. As a result, when the rank function ρ is polynomial, the proof of Theorem 1.1 in [67] give a tower type bound for M (see Section 5 for details).

The formulation of Theorem 1.1 allows us to make a precise analogy between quadratic arithmetic regularity and hypergraph regularity. Specifically, Theorem 1.1 implies that a special version of the hypergraph regularity lemma holds for the 3-uniform hypergraph $(\mathbb{F}_p^n, \{xyz : x+y+z \in A\})$, where $\mathcal{P}_{vert}$ consists of the atoms of a quadratic factor $\mathcal{B}$, and $\mathcal{P}_{pairs}$ consists of the fibres of the associated bilinear form (see Appendix B of [65]). The vertex complexity is then $p^{\ell+q}$, the pairs complexity is p^q, and the error function ψ corresponds to the so-called *rank function* ρ. It is important to note that this tells us such hypergraphs admit regularity lemmas with better bounds on the vertex complexity than is

[6]This result first appeared in early versions of [65], and was later moved to its own paper [67].

[7]A *growth function* is an increasing function from $\mathbb{R}$ into $\mathbb{R}^{\geq 0}$.

possible in general. In particular, the bound $p^{\ell+q}$ in Theorem 1.1 is tower type, while [45] showed that in general, there exist 3-uniform hypergraphs requiring a vertex partition of wowzer size. On the other hand, recent results of Gladkova [28] show one cannot do better than tower bounds on the linear complexity in a version of Theorem 1.1 stated for functions (rather than sets).

Using Theorem 1.1 and a counting lemma for the local U^3-norm appearing there, the second author and Wolf proved a structure theorem for sets of bounded VC_2-dimension, showing they are approximated by unions of quadratic atoms [67]. This result relied on Theorem 1.1, and thus yielded tower bounds for both ℓ and q (when the rank function is polynomial). The second author and Wolf later showed [68] that the bound on the quadratic complexity of the factor can be drastically improved, yielding the following theorem.

Theorem 1.2 (Theorem 1.3 in [68])**.** *For all odd primes p, all integers $k \geq 1$, all growth functions ρ, and all sufficiently small $\delta \in (0,1)$, there is a constant $C = C(p,k,\rho,\delta)$ so that the following holds for all sufficiently large n.*

For all $A \subseteq G = \mathbb{F}_p^n$ with VC_2-dimension at most k, there exists a quadratic factor $\mathcal{B}$ on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $\rho(\ell+q)$ such that

(1) $0 \leq \ell \leq C$,
(2) $0 \leq q \leq \log_p(\delta^{-k+o(1)})$, and
(3) there is a union Y of atoms of $\mathcal{B}$, such that $|A \Delta Y| \leq \delta |G|$.

As explained above, Appendix B of [65] shows p^q is the correct analogue of the pairs complexity in Theorem 1.2. With this in mind, we see that the upper bound $\log_p(\delta^{-k+o(1)})$ translates to a bound of $\delta^{-k+o(1)}$ for the pairs complexity of the corresponding hypergraph decomposition. This matches what one would expect in the bounded VC_2-dimension setting from the analogous hypergraph results (see [59, 63]).

The proof of Theorem 1.2 proceeds by first approximating the set A with a quadratic factor using the structure theorem of [67], then building a new factor with the same linear component, but much smaller quadratic component. Thus, while this strategy improves the quadratic complexity, it leaves the linear complexity unchanged (namely bounded by a tower when ρ is polynomial).

The goal of this paper is to improve the bound on the linear component in Theorem 1.2. Our result can be seen as an arithmetic analogue of [27]. However, as the general bound on the linear complexity in Theorem 1.1 is tower type already, our analogous theorem must attain a sub-tower bound. Specifically, we show that when the rank function is polynomial, one can improve the bound on ℓ in Theorem 1.2 to a quadruple exponential, and when ρ is linear (as suffices for most applications), one can obtain a triple exponential. The machinery of [68] allows us to obtain this improvement on the linear complexity simultaneously with the improvement on the quadratic complexity from Theorem 1.2, yielding the following main result.

Theorem 1.3. *For all odd primes p, all integers $k \geq 1$, all polynomial growth functions ρ, and all sufficiently small $\delta \in (0,1)$, the following holds for all sufficiently large n.*

For all $A \subseteq G = \mathbb{F}_p^n$ with VC_2-dimension at most k, there exists a quadratic factor $\mathcal{B}$ on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $\rho(\ell + q)$ such that

(1) $0 \leq \ell \leq \exp_p(\exp_p(\exp_p(\exp_p(O_{k,p,\rho}(\delta^{-O_{k,p}(1)})))))$,
(2) $0 \leq q \leq \log_p(\delta^{-k-o(1)})$, *and*
(3) there is a union Y of atoms of $\mathcal{B}$, such that $|A \Delta Y| \leq \delta |G|$.

Moreover, if ρ has degree 1, then one can ensure $\ell \leq \exp_p(\exp_p(\exp_p(O_{k,p,\rho}(\delta^{-O_{k,p}(1)}))))$.

Our overall strategy is based on the one used in [27] by Gishboliner, Shapira, and Wigderson to prove the hypergraph analogue. We give a rough outline of the arithmetic version of this argument for a set $A \subseteq \mathbb{F}_p^n$ of bounded VC_2-dimension, after which we discuss the additional ingredients needed to make this strategy work.

Step 1: Prove a "cylinder" version of the quadratic arithmetic regularity lemma with double exponential bounds for linear rank functions. This theorem will partition the group into atoms of (possibly distinct) high rank factors, so that the set A is appropriately uniform on most atoms in the partition.

Step 2: Show that since A has bounded VC_2-dimension, it has density near 0 or 1 on all of the uniform atoms from Step 1.

Step 3: Take the common refinement of all the factors used to build the cylinder partition from Step 1.

Step 4: Take a high rank refinement of the resulting factor. As refinements mostly preserve densities near 0 or 1, this will result in a factor where A has density near 0 or 1 on most atoms. From this we deduce that A looks approximately like the union of the "density near 1" atoms.

There are several technical hurtles related to Steps 1 and 2 which must be overcome to make this strategy work. We are able to resolve these difficulties, in part due to several recent results from the literature.

First, proving a cylinder version of the quadratic arithmetic regularity lemma requires a localized version of the U^3 inverse theorem which was proven recently by Prendiville [48] (see Subsection 3.1 for more details). Using this result, we accomplish Step 1, where the resulting notion of "uniform on most atoms" is defined in terms of a local version of the U^3-norm implicit in Prendiville's work. We give a slightly informal statement of this cylinder regularity lemma below and refer the reader to Section 6 for details. While a result of this type in for the U^2 norm was proven by Fox, Tidor, and Zhao in [24], this is to our knowledge the first such result for the U^3 norm.

Theorem 1.4 ("Cylinder" quadratic arithmetic regularity lemma)**.** *Fix p an odd prime, ρ a polynomial growth function, and $\delta \in (0,1)$. For all $A \subseteq G = \mathbb{F}_p^n$, there exists a partition $\mathcal{P}$ of G such that for each $P \in \mathcal{P}$ there is a quadratic factor $\mathcal{B}_P = (\mathcal{L}_P, \mathcal{Q}_P)$ of complexity (ℓ_P, q_P) such that the following hold.*

(1) For all $P \in \mathcal{P}$, P is an atom of $\mathcal{B}_P$, $\mathcal{B}_P$ has rank at least $\rho(\ell_P + q_P)$, and

$$\ell_P \leq \exp_p(\exp_p(\exp_p(O_{\rho,p}(\delta^{-1})))) \text{ and } q_P \leq O_p(\delta^{-O_p(1)});$$

(2) $|\bigcup_{P\in\mathcal{J}} P| \geq (1-\delta)|G|$, where $\mathcal{J}$ is the set of $P \in \mathcal{P}$ which are δ-uniform with respect to A, in terms of the local U^3 norm from [48].

Moreover, if ρ has degree 1, one can ensure that for all $P \in \mathcal{P}$, $\ell_P \leq \exp_p(\exp_p(O_{\rho,p}(\delta^{-1})))$.

Theorem 1.4 is proven via an energy increment argument which iteratively applies Prendiville's local U^3 inverse theorem. Clearly the bounds generated by such an argument will depend on the bounds in Prendiville's result. Luckily for us, these are polynomial, a fact which in turn depends on the recent resolution of the polynomial Freiman Rusza conjecture in $\mathbb{F}_p^n$ [35]. As Prendiville's inverse theorem requires a rank assumption, the proof of Theorem 1.4 interweaves steps in which we apply the inverse theorem and steps in which we perform operations to make the factors have high rank. It requires some delicacy to analyze the bounds resulting from such a process in enough detail to obtain the triple exponential bound. This extra care is not required in the hypergraph case [27], because there one need only obtain a tower bound at this step.

Turning to Step 2, one runs into another hurtle, namely proving an appropriately general counting lemma for the local U^3-norm in the conclusion of Theorem 1.4. One has no choice over which local norm is used in Step 1, as it must match the hypotheses of Prendiville's local inverse theorem. For this reason, we denote this local norm by $\|\cdot\|^P_{U^3(b)}$, where the "P" stands for Prendiville. On the other hand, a *different* local U^3 norm, which we denote by $\|\cdot\|^{TW}_{U^3(d)}$, was defined in [67], where an accompanying counting lemma was proved. In fact, exactly the statement needed for Step 2 was proved in [67], except for the local norm $\|\cdot\|^{TW}_{U^3(d)}$ rather than the local norm $\|\cdot\|^P_{U^3(d)}$. To make Step 2 work, we show these norms are in fact approximately the same (see Subsection 3.3). A consequence of this is that [67] implies a counting lemma for $\|\cdot\|^P_{U^3(b)}$, which may be of general interest (see Propositions 3.19 and D.11 in [67]).

We end this introduction with remarks on lower bounds. First, it was shown in [68] that there exist sets requiring q to grow at a rate of at least a power of δ^{-1} in Theorem 1.1, showing Theorem 1.2 does not hold in general. Regarding the linear component, a recent result of Gladkova shows the version of Theorem 1.1 for *functions* requires tower type lower bounds on on the linear complexity. It seems likely this can be turned into an example of a subset of $\mathbb{F}_p^n$ requiring tower sized linear complexity in Theorem 1.1 (showing the bound on ℓ in Theorem 1.3 cannot hold in general). However, we do not undertake this task in this paper. We believe the bound on q in Theorem 1.3 is roughly optimal, however, it seems possible the bound for ℓ can be improved further. It remains open what happens when faster rank functions are allowed. In the setting of 3-uniform hypergraphs, it was shown in [62] that there exist 3-uniform hypergraphs of bounded VC_2-dimension requiring wowzer lower bounds on the vertex partition for fast growing ψ. Gladkova has proved arithmetic analagues of several ingredients in that argument, however it remains open whether that strategy can be fully executed in the arithmetic setting (see Section 4.1 of [29]).

1.4. **Acknowledgements.** The second author thanks Julia Wolf for many helpful discussions about topics related to this paper, and especially for clarifying the bounds in the quadratic arithmetic regularity lemma of [37].

1.5. **Notation.** Throughout the rest of this paper, p is an odd prime. By convention, the natural numbers start at 0. For an integer $n \geq 1$, we let $[n] = \{1, \dots, n\}$.

Given a set X, $\mathbb{E}_{x \in X}$ means $\frac{1}{|X|}\sum_{x \in X}$. Given a finite index set $I = \{\alpha_1, \dots, \alpha_t\}$, and sets X_i for $i \in I$, we write $\mathbb{E}_{\substack{x_i \in X_i \\ i \in I}}$ to mean $\mathbb{E}_{x_1 \in X_{\alpha_1}} \dots \mathbb{E}_{x_t \in X_{\alpha_t}}$.

Given real numbers r, s, we write $r = s \pm \epsilon$ to mean $|r - s| \leq \epsilon$. Given real valued functions f, g, h on the same domains, we write $f \gg g$ to mean there exists a constant C such that $f(x) \geq Cg(x)$ for all x, and we write $f = h(1 \pm O(g))$ to mean there is a constant C so that for all x, $|f(x) - h(x)| \leq Ch(x)g(x)$. Given a parameter θ, we write $f \gg_\theta g$ and/or $f = h(1 \pm O_\theta(g))$ to denote that the relevant constant C depends on the parameter θ. Throughout the paper, we will refer to the following notion of a *growth function*.

Definition 1.5 (Growth functions)**.** A *growth function* is any increasing function $\rho : \mathbb{R} \to \mathbb{R}^{\geq 0}$. We say a growth function is *polynomial* if there exists a constant $C > 0$ and an integer $d \geq 1$ so that for all $x \geq 1$, $\rho(x) \leq Cx^d$. The minimal d for which this holds is then called the *degree* of ρ.

1.6. **Outline.** In Section 2, we introduce the definition of the VC_2-dimension of a subset in a group, the definitions of linear and quadratic factors, and state several lemmas for computing sizes of sets defined by such factors. We also cover notation needed to translate the results of [48] into the terminology used in this paper. In Section 3, we introduce the localized inverse theorem of Prendiville [48], and introduce both the local U^3 norm implicit in that paper, as well as the one defined by the second author and Wolf in [67]. Also in this section, we prove these local norms are approximately equal, allowing us to translate a crucial theorem about sets of bounded VC_2-dimension from [67] into a statement involving the local norm from [48]. In Section 4 we set out definitions and prove lemmas which we will need to analyze the bounds in our cylinder regularity lemma, Theorem 6.1. In Section 5, as a warm up to Theorem 6.1, we use Prendiville's local inverse theorem to give an energy increment proof of a quadratic arithmetic regularity lemma. In Section 6 we prove our cylinder style quadratic arithmetic regularity lemma, Theorem 6.1. Finally, in Section 7, we prove our main theorem.

2. Preliminaries

In this section we cover the definition of VC_2-dimension, the higher arity version of VC-dimension which is the topic of this paper. We then introduce preliminaries related to linear and quadratic factors.

2.1. VC_2**-dimension.** There are several ways to generalize the notion of VC-dimension in graphs to the setting of hypergraphs. This paper focuses on an analogue for 3-uniform hypergraphs called VC_2-dimension. For more papers studying this notion, we refer the reader to the literature [4–7, 19, 27, 40, 52, 58, 59, 62, 63, 65–68]. For related papers on other

higher arity generalizations of VC-dimension, we refer the reader to [7, 8, 10, 17, 18, 23, 26, 60, 61, 66], for example. This paper is interested in specifically subsets of groups, and we will restrict our attention to this context in what follows. For context, we begin by defining the VC-dimension of a subset of a finite group.

Definition 2.1. Given an integer $k \geq 1$, a finite group G, and a subset $A \subseteq G$, we say A *has VC-dimension at least* k if there exist $a_1, \ldots, a_k \in G$ and $\{b_S : S \subseteq [k]\} \subseteq G$ so that for all $i \in [k]$ and $S \subseteq [k]$,

$$a_i \cdot b_S \in A \text{ if and only if } i \in S.$$

The *VC-dimension of* A is then defined to be

$$\mathrm{VC}(A) = \max\Big(\{0\} \cup \{k \in \mathbb{Z}^{\geq 1} : A \text{ has VC-dimension at least } k\}\Big).$$

As mentioned in the introduction, subsets of groups of bounded VC-dimension have been studied in [3, 14, 15, 54]. We now define the VC_2-dimension of a subset of a group.

Definition 2.2. Given an integer $k \geq 1$, a finite group G, and a subset $A \subseteq G$, we say A *has VC-dimension at least* k if there exist $a_1, \ldots, a_k, b_1, \ldots, b_k \in G$ and $\{c_S : S \subseteq [k] \times [k]\} \subseteq G$ so that for all $i, j \in [k]$ and $S \subseteq [k]$,

$$a_i \cdot b_j \cdot c_S \in A \text{ if and only if } (i, j) \in S.$$

The VC_2*-dimension of* A is then defined to be

$$\mathrm{VC}_2(A) = \max\Big(\{0\} \cup \{k \in \mathbb{Z}^{\geq 1} : A \text{ has } \mathrm{VC}_2\text{-dimension at least } k\}\Big).$$

Subsets of elementary abelian p-groups of bounded VC_2-dimension have been previously studied in [65, 66, 68]. For more history on these subjects, we refer the reader back to the introduction.

2.2. **Linear and quadratic factors.** In this section we introduce notation for the linear and quadratic factors used in this paper. Our setup is largely based on [37], and is designed to be consistent with [65, 67, 68]. We then cover background required to translate results from [48] into our setup. Finally, we cover several lemmas which approximately compute the sizes of sets arising from quadratic factors.

We begin by defining linear factors and linear atoms. We define these in two cases, based on whether or not the complexity is trivial.

Definition 2.3 (Linear factors and atoms)**.** Given an integer $\ell \geq 1$, a *linear factor on* $\mathbb{F}_p^n$ *of complexity* ℓ is a set $\mathcal{L} = \{r_1, \ldots, r_\ell\} \subseteq \mathbb{F}_p^n$ of linearly independent vectors. Given $a = (a_1, \ldots, a_\ell) \in \mathbb{F}_p^\ell$, define

$$L(a) = \{x \in \mathbb{F}_p^n : \text{ for each } i \in [\ell], x \cdot r_i = a_i\}.$$

The sets of form $L(a)$ for $a \in \mathbb{F}_p^\ell$ are called *atoms of* $\mathcal{L}$, and we refer to the tuple a as the *label* for the atom $L(a)$. We let $\mathrm{At}(\mathcal{L})$ denote the set of atoms of $\mathcal{L}$.

The *linear factor on* $\mathbb{F}_p^n$ *of complexity* 0 is $\mathcal{L} = \emptyset$. It has a unique atom, $L(0) := \mathbb{F}_p^n$, where the label 0 is the unique element of the 0-dimensional vector space $\mathbb{F}_p^0$. We then let $\mathrm{At}(\mathcal{L}) = \{\mathbb{F}_p^n\} = \{L(a) : a \in \mathbb{F}_p^0\}$.

The reader may notice that our notation for the atoms of a factor depends on a fixed enumeration of the vectors in $\mathcal{L}$. For this reason it would be more correct to define factors as tuples of vectors rather than sets of vectors. In practice, no confusion will arise from this minor abuse of notation.

Every linear factor comes with an associated linear form, whose fibres make up its atoms.

Definition 2.4 (Linear form associated to a linear factor)**.** If $\ell \geq 1$ is an integer, and $\mathcal{L} = \{r_1, \ldots, r_\ell\}$ is a linear factor on $G = \mathbb{F}_p^n$ of complexity ℓ, define $\beta_{\mathcal{L}} : G \to \mathbb{F}_p^\ell$ by setting

$$\beta_{\mathcal{L}}(x) = (x \cdot r_1, \ldots, x \cdot r_\ell),$$

for each $x \in G$.

If $\mathcal{L}$ is the linear factor on $G = \mathbb{F}_p^n$ of complexity 0, define $\beta_{\mathcal{L}} : G \to \mathbb{F}_p^0$ by setting $\beta_{\mathcal{L}}(x) = 0$ for all $x \in G$.

In the notation of Definition 2.4, it is clear $\beta_{\mathcal{L}}$ is a linear form, and for each $a \in \mathbb{F}_p^\ell$, $L(a) = \beta_{\mathcal{L}}^{-1}(a)$. It will be useful to know that any linear map from $\mathbb{F}_p^n$ into $\mathbb{F}_p$ comes from such a map in the following sense.

Fact 2.5. *Any linear map $\phi : \mathbb{F}_p^n \to \mathbb{F}_p$ has the form $\phi(x) = \beta_{\mathcal{L}}(x) + c$ for some linear factor $\mathcal{L}$ on $\mathbb{F}_p^n$, and some constant $c \in \mathbb{F}_p$.*

Proof. Let $e_1, \ldots, e_n$ denote the standard basis vectors in $\mathbb{F}_p^n$. Setting $\mathcal{L} = \{\phi(e_1), \ldots, \phi(e_n)\}$ and $c = \phi(0)$, it is easy to check that for all $x \in \mathbb{F}_p^n$, $\phi(x) = \beta_{\mathcal{L}}(x) + c$, as desired. □

We now define purely quadratic factors. We again do this in two cases, based on whether or not the complexity is trivial.

Definition 2.6 (Purely quadratic factors)**.** Given an integer $q \geq 1$, a *purely quadratic factor on* $\mathbb{F}_p^n$ *of complexity* q is a set $\mathcal{Q} = \{M_1, \ldots, M_q\}$ of pairwise distinct symmetric $n \times n$ matrices with entries from $\mathbb{F}_p$. Given $a = (a_1, \ldots, a_q) \in \mathbb{F}_p^q$, define

$$Q(a) = \{x \in \mathbb{F}_p^n : \text{ for each } i \in [q], x^T M_i x = a_i\}.$$

Sets of the form $Q(a)$ are called *atoms of* $\mathcal{Q}$, and we refer to the tuple $a \in \mathbb{F}_p^q$ as the *label* for the atom $Q(a)$. We let $\mathrm{At}(\mathcal{Q})$ denote the set of atoms of $\mathcal{Q}$.

The *purely quadratic factor on* $\mathbb{F}_p^n$ *of complexity* 0 is $\mathcal{Q} = \emptyset$. It has a unique atom, $Q(0) := \mathbb{F}_p^n$, where the label 0 is the unique element of the 0-dimensional vector space $\mathbb{F}_p^0$. We then let $\mathrm{At}(\mathcal{Q}) = \{\mathbb{F}_p^n\} = \{Q(a) : a \in \mathbb{F}_p^0\}$.

Every purely quadratic factor comes with a corresponding bilinear form.

Definition 2.7 (Bilinear forms associated to purely quadratic factors)**.** Let $q \geq 1$ be an integer, and let $\mathcal{Q} = \{M_1, \ldots, M_q\}$ be a purely quadratic factor on $G = \mathbb{F}_p^n$ of complexity q. Define $\beta_{\mathcal{Q}} : G^2 \to \mathbb{F}_p^q$ by setting

$$\beta_{\mathcal{Q}}(x, y) = (x^T M_1 y, \ldots, x^T M_q y),$$

for each $(x, y) \in G^2$.

If $\mathcal{Q}$ is the purely quadratic factor on $G = \mathbb{F}_p^n$ of complexity 0, define $\beta_{\mathcal{Q}} : G^2 \to \mathbb{F}_p^0$ by setting $\beta_{\mathcal{Q}}(x, y) = 0$ for all $(x, y) \in G^2$.

In the notation of Definition 2.7, it is not difficult to see that $\beta_{\mathcal{Q}}$ is a symmetric, bilinear form. The fibres, $\beta_{\mathcal{Q}}^{-1}(b) = \{(x, y) \in G^2 : \beta_{\mathcal{Q}}(x, y) = b\}$, will play an important role in Section 3. The atoms of $\mathcal{Q}$ can also be written in terms of $\beta_{\mathcal{Q}}$. Specifically, for every $b \in \mathbb{F}_p^q$, $Q(b) = \{x \in G : \beta_{\mathcal{Q}}(x, x) = b\}$.

Purely quadratic factors come with the following notion of rank.

Definition 2.8 (Rank of a purely quadratic factor)**.** Suppose $q \geq 0$ and $\mathcal{Q}$ is a purely quadratic factor on $\mathbb{F}_p^n$ of complexity q. We define the *rank of* $\mathcal{Q}$ in cases below.

(1) If $q = 0$, then $\mathcal{Q}$ has rank n.
(2) If $q > 0$ and $\mathcal{Q} = \{M_1, \ldots, M_q\}$, then the rank of $\mathcal{Q}$ is the minimal rank of a non-trivial linear combination of the matrices in $\mathcal{Q}$. In other words,

$$\min\{\mathrm{rk}(\lambda_1 M_1 + \ldots + \lambda_q M_q) : \lambda_1, \ldots, \lambda_q \in \mathbb{F}_p \text{ not all zero}\}.$$

We now define quadratic factors, which are obtained by combining linear and purely quadratic factors.

Definition 2.9 (Quadratic factors)**.** Given integers $\ell, q \geq 0$, a *quadratic factor on* $\mathbb{F}_p^n$ *of complexity* (ℓ, q) is a pair $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ where $\mathcal{L}$ is a linear factor on $\mathbb{F}_p^n$ of complexity ℓ and $\mathcal{Q}$ is a purely quadratic factor on $\mathbb{F}_p^n$ of complexity q. An *atom of* $\mathcal{B}$ is a set of the form

$$B(a, b) = L(a) \cap Q(b),$$

for some $(a, b) \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$. We call the pair $(a, b) \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ the *label* of the atom $B(a, b)$. We let $\mathrm{At}(\mathcal{B})$ denote the set of atoms of $\mathcal{B}$.

Given a quadratic factor $\mathcal{B}$, it will be useful to have a function taking a group element x to the label of the atom of $\mathcal{B}$ containing x. This function is naturally built from the the linear and bilinear forms $\beta_{\mathcal{L}}$ and $\beta_{\mathcal{Q}}$.

Definition 2.10. Suppose $\ell, q \geq 0$ are integers and $\mathcal{B}$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q). Define $\beta_{\mathcal{B}} : G \to \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ by setting

$$\beta_{\mathcal{B}}(x) = (\beta_{\mathcal{L}}(x), \beta_{\mathcal{Q}}(x, x)),$$

for all $x \in G$.

In the notation of Definition 2.10, it is easy to check we have $x \in B(\beta_{\mathcal{B}}(x))$ for all $x \in G$. We next extend the notion of rank to quadratic factors in the obvious way.

Definition 2.11 (Rank of a quadratic factor)**.** Suppose $\ell, q \geq 0$ are integers and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q). The *rank of* $\mathcal{B}$ is the rank of $\mathcal{Q}$, as defined in Definition 2.8.

When a factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ has high rank, sets involving the fibres of the maps $\beta_{\mathcal{B}}$ and $\beta_{\mathcal{Q}}$ have predicable sizes. We will use several statements to this effect throughout the paper. For the convenience of the reader, we collect these and related lemmas in their own subsection at the end of this section (see Subsection 2.3).

Given two quadratic factors $\mathcal{B}$ and $\mathcal{B}'$ on $\mathbb{F}_p^n$, we say $\mathcal{B}'$ *refines* $\mathcal{B}$, denoted $\mathcal{B}' \preceq \mathcal{B}$, if the partition $\mathrm{At}(\mathcal{B}')$ of $\mathbb{F}_p^n$ refines the partition $\mathrm{At}(\mathcal{B})$ of $\mathbb{F}_p^n$. The following lemma says that given a quadratic factor, one can find a high rank refinement of bounded complexity (see Lemma 3.11 in [37]).

Lemma 2.12 (Rank Lemma)**.** *Let ρ be a growth function. Suppose $\ell, q \geq 0$ are integers, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q). There there exist $0 \leq q' \leq q$ and*

$$\ell \leq \ell' \leq C = C(\ell, q, \rho),$$

and a quadratic factor $\mathcal{B}' \preceq \mathcal{B}$ of complexity (ℓ', q') and rank at least $\rho(\ell' + q')$.

The shape of the bound $C(\ell, q, \rho)$ in Lemma 2.12 is important for this paper's main results, and will be considered in detail in Section 4.

The setup outlined above differs slightly from that used in Prendiville's paper [48], the main result of which is crucial for us. We next set out enough definitions from [48] to state the main result from [48], and then translate it into our setup.

Definition 2.13. A *quadratic polynomial on $G = \mathbb{F}_p^n$* is a map from $\mathbb{F}_p^n$ to $\mathbb{F}_p$ of the form $\psi(x) = \beta(x, x) + \phi(x) + c$, where $\beta : G^2 \to \mathbb{F}_p$ is a symmetric bilinear from, $\phi : G \to \mathbb{F}_p$ is a linear form, and $c \in \mathbb{F}_p$ is a constant.

As observed in [48], it is not difficult to show that any quadratic polynomial can be built in the following way.

Fact 2.14. *Any quadratic polynomial $\psi : G = \mathbb{F}_p^n \to \mathbb{F}_p$ has the form $x^T M x + r \cdot x + c$ for some symmetric $n \times n$ matrix M with entries from $\mathbb{F}_p$, some $r \in \mathbb{F}_p^n$, and some $c \in \mathbb{F}_p$.*

Proof. Say $\psi(x) = \beta(x, x) + \phi(x) + c$. Let $e_1, \dots, e_n$ denote the standard basis vectors in $\mathbb{F}_p^n$. Then let M be the matrix with ij-th entry $\beta(e_i, e_j)$, and let r be the vector $(\phi(e_1), \dots, \phi(e_n))$. It is easy to check $\psi(x) = x^T M x + r \cdot x + c$, and that M is symmetric since β is. □

In light of the fact above, we will from here on assume all quadratic polynomials have the form appearing in Fact 2.14. The results in [48] are stated in terms of preimages of maps built from tuples of quadratic polynomials. We now give some definitions related to such tuples.

Definition 2.15. Let $d \geq 1$ be an integer, and for each $1 \leq i \leq d$, let ψ_i be a quadratic polynomial of the form $\psi_i(x) = x^T M_i x + r_i \cdot x + c_i$. The *rank* of the tuple $\Psi = (\psi_1, \dots, \psi_d)$ is the rank of the purely quadratic factor $\{M_1, \dots, M_d\}$ (in the sense of Definition 2.8).

Given $b = (b_1, \dots, b_d) \in \mathbb{F}_p^d$, let

$$\Psi^{-1}(b) = \{x \in G : \psi_i(x) = b_i \text{ for each } i \in [d]\}.$$

We can now translate between this notation and ours.

Fact 2.16. *Suppose $\ell, q \geq 0$ are integers satisfying $\ell + q \geq 1$, $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor of complexity (ℓ, q), and $(a, b) \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$. Then there is a tuple $\Psi = (\psi_1, \dots, \psi_{\ell+q})$ of quadratic polynomials, whose rank is equal to the rank of $\mathcal{B}$, such that $\Psi^{-1}(0) = B(a, b)$.*

Proof. Suppose first ℓ and q are both at least 1. Let $\mathcal{L} = \{r_1, \dots, r_\ell\}$, $\mathcal{Q} = \{M_1, \dots, M_q\}$, and fix $a = (a_1, \dots, a_\ell) \in \mathbb{F}_p^\ell$ and $b = (b_1, \dots, b_q) \in \mathbb{F}_p^q$.

For each $1 \leq i \leq \ell$, set $\phi_i(x) = r_i \cdot x - a_i$, and for each $1 \leq j \leq q$, set $\psi_{\ell+j}(x) = x^T M_i x - b_i$. By construction, $\Psi^{-1}(0) = B(a, b)$, and the rank of Q is the rank of $\mathcal{Q}$, which is by definition the rank of $\mathcal{B}$. Thus we are done in the case where both ℓ and q are positive.

If one of ℓ or q is 0, simply repeat the same argument omitting either the linear or quadratic maps. □

2.3. **Sizes of atoms and related sets.** In this subsection we collect several lemmas about sizes of sets related to quadratic factors. The first such lemma tells us that atoms of a high rank factors all have about the same size (see Lemma 4.2 in [37]).

Lemma 2.17. *Suppose $\ell, q \geq 0$ are integers, $\tau \in \mathbb{R}^{\geq 0}$, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least τ. Given any $e \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$,*

$$|B(e)| = (1 \pm O(p^{\ell+q-\tau/2}))p^{n-\ell-q}.$$

The next lemma tells us the approximate size of fibres of the map $\beta_{\mathcal{Q}}$ from Definition 2.7.

Lemma 2.18. *Suppose $\ell, q \geq 0$ are integers, $\tau \in \mathbb{R}^{\geq 0}$, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least τ. Given any $d \in \mathbb{F}_p^q$,*

$$|\beta_{\mathcal{Q}}^{-1}(d)| = (1 \pm O(p^{q-\tau}))p^{2n-q}.$$

We will repeatedly use the above size estimates in conjunction with a very general statement from [67] (Lemma 2.20 below). It uses the following normalized indicator functions.

Notation 2.19. Given a group G and $\Gamma \subseteq G \times G$, define $\mu_\Gamma : G \times G \to \mathbb{R}$ by setting $\mu_\Gamma(x, y) = \frac{|G|^2}{|\Gamma|} 1_\Gamma(x, y)$ for all $(x, y) \in G^2$.

The following is a slightly more general version of Lemma A.2 in [67]. It is straightforward to see this more general statement follows from exactly the same proof as that of Lemma A.2 in [67] by simply adding the appropriate superscripts throughout.

Lemma 2.20 (Lemma A.2 of [67])**.** *Let $\tau \geq 0$ be a real, let $m \geq 1$ and $\ell, q \geq 0$ be integers, and suppose $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least τ. For each $i \in [m]$, let $a_1^i, a_2^i, a_3^i \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ and for each $1 \leq i, j \leq m$, let $b_{12}^{ij}, b_{13}^{ij}, b_{23}^{ij} \in \mathbb{F}_p^q$. Then for all sets $I, J, K \subseteq [m]$,*

$$\mathbb{E}_{\substack{x_i \in B(a_1^i) \\ i \in I}} \mathbb{E}_{\substack{y_j \in B(a_2^i) \\ j \in J}} \mathbb{E}_{\substack{z_k \in B(a_3^i) \\ k \in K}} \prod_{i \in I, j \in J} \mu_{\beta_{\mathcal{Q}}^{-1}(b_{12}^{ij})}(x_i, y_j) \prod_{i \in I, k \in K} \mu_{\beta_{\mathcal{Q}}^{-1}(b_{13}^{ij})}(x_i, z_k) \prod_{j \in J, k \in K} \mu_{\beta_{\mathcal{Q}}^{-1}(b_{23}^{ij})}(y_j, z_k)$$
$$= 1 \pm O_m(p^{(\ell+q)(|I|+|J|+|K|)+q(|I||J|+|I||K|+|J||K|)-\tau/2}).$$

The same holds whenever any instance of $\mu_{\beta_{\mathcal{Q}}^{-1}(b_{uv}^{ij})}$ is replaced by the constant function 1.

Finally, at a certain juncture in the paper, we will be working with a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$, and need to consider an auxiliary factor of the form $(\mathcal{L}', \mathcal{Q})$, where

$$\mathcal{L}' = \mathcal{L} \cup \{Mw_1, Mw_2, Mw_3, Mw_4 : M \in \mathcal{Q}\},$$

for some group elements w_1, w_2, w_3, w_4. When $\mathcal{L}'$ is a linearly independent set, the pair $(\mathcal{L}', \mathcal{Q})$ is indeed a quadratic factor in the sense of Definition 2.9. In this case, we can estimate sizes of its atoms and related sets using the above lemmas. On the other hand, for some choices of group elements $w_1, \dots, w_4$, the resulting $\mathcal{L}'$ will not be linearly independent. We will need to know the overall number of such "bad" choices for $w_1, \dots, w_4$ is small. This is the purpose of the following lemma.

Lemma 2.21. *Let $r \geq 0$ be a real, let $\ell, q \geq 0$ be integers, and let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least r. Then*

$$|\{(w_1, w_2, w_3, w_4) \in G^4 : \mathcal{L} \cup \{Mw_1, Mw_2, Mw_3, Mw_4 : M \in \mathcal{Q}\} \text{ is not linearly independent}\}|$$

is at most $14p^{4n+\ell+4q-r}$.

The proof of Lemma 2.21 is straightforward and appears in the Appendix.

3. Local U^3 norms and the local inverse theorem

This section accomplishes three objectives. First, in Subsection 3.1, we state Prendiville's localized version of the U^3 inverse theorem in terms of a local U^3 norm implicit in that paper [48]. Second, Subsection 3.2 reviews the definition of a different local U^3 norm from [67], and states a crucial fact, also proved in [67], that a set A of bounded VC_2-dimension must have density near 0 or 1 on any atom which is uniform with respect to A in this sense. Finally, in Subsection 3.3, we prove the local norms from [48] and [67] are approximately equal. This allows us to deduce the necessary corollary about sets of bounded VC_2-dimension in terms of the local norm from [48], as is required for the proof of our main theorem.

3.1. Prendiville's inverse theorem for the U^3-norm localized to a quadratic atom.

In this section we state a localized version of the inverse theorem for the U^3-norm[8] due to Prendiville [48]. Throughout the section we deal with functions $f : \mathbb{F}_p^n \to [-1, 1]$, as opposed to the more general functions $f : \mathbb{F}_p^n \to \mathbb{C}$ considered in Prendiville's paper. We do this merely to ease notation and because this suffices for our purposes. We do not attempt to fully contextualize the definition of the Gowers norms or the statement of the inverse theorem. We refer the reader to the literature for more extensive background [31, 32, 35, 38, 48, 57].

We begin by defining the usual (global) versions of the Gowers U^2 and U^3 norms.

Definition 3.1 (U^2 and U^3-norms)**.** Suppose $G = \mathbb{F}_p^n$.

[8]Throughout this paper, we use the term "norm" in a somewhat colloquial manner, as several of the objects we define yield only semi-norms.

(1) Given a function $f : G \to [-1, 1]$, the *U^2-norm of f*, denoted $\|f\|_{U^2}$, is defined by the equation

$$\|f\|_{U^2}^4 = \sum_{x,h_2,h_2 \in G} f(x)f(x+h_1)f(x+h_2)f(x+h_1+h_2).$$

(2) Given a function $f : G \to [-1, 1]$, the *U^3-norm of f*, denoted $\|f\|_{U^3}$, is defined by the equation

$$\|f\|_{U^3}^8 = \sum_{h \in G} \|\Delta_h f\|_{U^2}^4,$$

where $\Delta_h f$ is the function defined by setting $\Delta_h f(x) = f(x+h)$.

It is well known the above define norms on the space of functions from G into $[-1, 1]$ (see e.g. Section 11 of [57]). The following notation will be convenient for dealing with the U^3 norm.

Notation 3.2. Given $f : G \to [-1, 1]$ and $(x, h_1, h_2, h_3) \in G^4$, let

$$\begin{aligned}\pi_f(x, h_1, h_2, h_3) =& f(x)f(x+h_1)f(x+h_2)f(x+h_3)\cdot \\ & f(x+h_1+h_2)f(x+h_1+h_3)f(x+h_2+h_3)f(x+h_1+h_2+h_3).\end{aligned}$$

This notation allows us to rewrite the U^3 norm as follows.

Observation 3.3. *Let $G = \mathbb{F}_p^n$. Given a function $f : G \to [-1, 1]$,*

$$\|f\|_{U^3}^8 = \sum_{(x,h_1,h_2,h_3) \in G^4} \pi_f(x, h_1, h_2, h_3).$$

The inverse theorem for the U^3 norm says a function with large U^3 has some correlation with a quadratic phase function. For the version stated below, see Corollary 1.2 of [35].

Theorem 3.4. *[Global U^3 inverse theorem in $\mathbb{F}_p^n$ with polynomial bounds [35]] Let $\delta \in (0, 1)$, and suppose $f : G = \mathbb{F}_p^n \to [-1, 1]$ satisfies $\|f\|_{U^3} \geq \delta p^{4n}$. Then there exists a quadratic polynomial $q : G \to \mathbb{F}_p$ such that*

$$\left| \sum_{x \in G} f(x) e^{2\pi i q(x)/p} \right| \gg_p \delta^{O_p(1)} p^n.$$

In [48], Prendiville proved a version of this inverse theorem localized to a single quadratic atom. This theorem will play a crucial role in our main result.

Theorem 3.5. *[Local U^3 inverse theorem with polynomial bounds [48]] Let $\delta \in (0, 1)$, let $d \geq 1$ be an integer, and let $\Psi = (\psi_1, \ldots, \psi_d)$ be a d-tuple of quadratic polynomials on G. Suppose $f : G \to [-1, 1]$ is a function with support contained in $B = \Psi^{-1}(0)$ satisfying $\|f\|_{U^3} \geq \delta \|1_B\|_{U^3}$. Then one of the following hold.*

- *there exists a quadratic polynomial $\psi : G \to \mathbb{F}_p$ such that*

$$\left| \sum_{x \in B} f(x) e^{2\pi i \psi(x)/p} \right| \gg_p \delta^{O_p(1)} |B|;$$

- Ψ *has rank at most* $O_p(d + \log(2/\delta))$.

Our next goal is to rephrase Theorem 3.5 in terms of the quadratic factors defined in Subsection 2.2, and in terms of a local U^3 norm.

First, we define a version of the U^3 norm localized to a quadratic atom. We decorate this norm with a P to denote it is implicit in Prendiville's paper, and to distinguish it from a different local U^3-norm (due to the second author and Wolf) which we will also use in this paper (see Definition 3.16).

Definition 3.6 (Local U^3-norm: Prendiville version)**.** Suppose $\ell, q \geq 0$ are integers, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor of complexity (ℓ, q) on $G = \mathbb{F}_p^n$. Given a function $f : G \to [-1, 1]$ and a tuple $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$, define

$$\|f\|_{U^3(b)}^P = \begin{cases} \frac{\|f \cdot 1_{B(b)}\|_{U^3}}{\|1_{B(b)}\|_{U^3}} & \text{if } \|f \cdot 1_{B(b)}\|_{U^3} \neq 0, \text{ and} \\ 0 & \text{if } \|f \cdot 1_{B(b)}\|_{U^3} = 0. \end{cases}$$

We note that in Definition 3.6, we have suppressed the dependence on the factor $\mathcal{B}$ in the notation $\|\cdot\|_{U^3(b)}^P$. When we use this notation, the factor in question will be clear from context. We now restate Theorem 3.5.

Corollary 3.7. *There exist a constant $C = C(p) > 0$ so that the following holds. Let $\delta \in (0, 1)$ and let $n \geq 1$ and $\ell, q \geq 0$ be integers.*

Suppose $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q), $f : \mathbb{F}_p^n \to [-1, 1]$ is a function, and $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$. If $\|f\|_{U^3(b)}^P \geq \delta$, then one of the following hold.

- *there exists a quadratic polynomial $\psi : G \to \mathbb{F}_p$ such that*

$$\left| \sum_{x \in B(b)} f(x) e^{2\pi i \psi(x)/p} \right| \geq C^{-1} \delta^C |B(b)|;$$

- *$\mathcal{B}$ has rank at most $C(\ell + q + \log(2/\delta))$.*

Proof. Let $C = C(p)$ be sufficiently large based on the $\gg_p$ and $O_p(1)$ appearing in Theorems 3.4 and 3.5. Fix δ, n, ℓ, q, and $\mathcal{B}$ as in the hypotheses. Suppose $f : G = \mathbb{F}_p^n \to [-1, 1]$ and $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ are such that $\|f\|_{U^3(b)}^P \geq \delta$.

Suppose first $\ell = q = 0$. In this case, $b = (0, 0) \in \mathbb{F}_p^0 \times \mathbb{F}_p^0$, $B(b) = \mathbb{F}_p^n$, and the hypothesis $\|f\|_{U^3(b)}^P \geq \delta$ translates into saying the global U^3 norm is large, specifically, $\|f\|_{U^3} \geq \delta p^{4n}$. The desired conclusion then follows immediately from the global U^3 inverse theorem, Theorem 3.4.

Suppose now $\min\{\ell, q\} \geq 1$. By Fact 2.16, there is a tuple of quadratic polynomials $\Psi = (\psi_1, \ldots, \psi_{\ell+q})$ with the same rank as $\mathcal{B}$, such that $\Psi^{-1}(0) = B(b)$. By Definition 3.6, the hypothesis $\|f\|_{U^3(b)}^P \geq \delta$ means $\|f \cdot 1_{B(b)}\|_{U^3} \geq \delta \|1_{B(b)}\|_{U^3}$. Consequently, applying Theorem 3.5 to Ψ and the function $f \cdot 1_{B(b)}$, we have that either there exists a quadratic

polynomial $\psi : G \to \mathbb{F}_p$ such that

$$\left| \sum_{x \in B(b)} f(x) e^{2\pi i \psi(x)/p} \right| \geq C^{-1} \delta^C |B(b)|,$$

or Ψ has rank at most $C(\ell + q + \log(2/\delta))$. Since the rank of Ψ is the rank of $\mathcal{B}$, we are done. □

We will end this subsection with some observations about Definition 3.6 that we will need in Subsection 3.3. For these remarks, it will be useful to have the following definition for the set of 4-tuples which yield a non-zero term in the sum appearing in $\|1_{B(b)}\|_{U^3}^8$.

Definition 3.8. Suppose $\mathcal{B}$ is a quadratic factor on $G = \mathbb{F}_p^n$ and $B \in \mathrm{At}(\mathcal{B})$. Define

$$\Omega_B = \Big\{ (x, h_1, h_2, h_3) \in G^4 : \\ 1_B(x) \Big(\prod_{i=1}^{3} 1_B(x + h_i) \Big) \Big(\prod_{i \neq j \in [3]} 1_B(x + h_i + h_j) \Big) 1_B(x + h_1 + h_2 + h_3) \neq 0 \Big\}.$$

Observe that in the notation of Definitions 3.8 and Definition 3.1, $|\Omega_B| = \|1_B\|_{U^3}^8$. Using this observation, it is easy to show $\|f\|_{U^3(b)}^P$ from Definition 3.6 is always bounded by 1.

Observation 3.9. *Suppose $\ell, q \geq 0$ are integers, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor of complexity (ℓ, q) on $G = \mathbb{F}_p^n$. For any $f : G \to [-1, 1]$ and $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$, $\|f\|_{U^3(b)}^P \leq 1$.*

Proof. Since f is 1-bounded, it is easy to see

$$\|f \cdot 1_{B(b)}\|_{U^3}^8 = \sum_{(x, h_1, h_2, h_3) \in \Omega_{B(b)}} \pi_f(x, h_1, h_2, h_3) \leq |\Omega_{B(b)}| = \|1_{B(b)}\|_{U^3}^8.$$

Thus, $\|f \cdot 1_{B(b)}\|_{U^3} \leq \|1_{B(b)}\|_{U^3}$, so by definition, $\|f\|_{U^3(b)}^P \leq 1$. □

The fact that $|\Omega_B| = \|1_B\|_{U^3}^8$ also allows us to rewrite Definition 3.6 as follows (when the factor in question has sufficiently high rank).

Observation 3.10. *There exists a constant $C = C(p) > 0$ so that the following holds. Suppose $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $C(\ell + q)$. For any function $f : \mathbb{F}_p^n \to [-1, 1]$ and $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$, we have*

$$\Big(\|f\|_{U^3(b)}^P \Big)^8 = \frac{\|f \cdot 1_{B(b)}\|_{U^3}^8}{|\Omega_{B(b)}|}.$$

Observation 3.10 follows immediately from the definitions of $\Omega_{B(b)}$ and $\|\cdot\|_{U^3(b)}^P$, and the fact that when $\mathcal{B}$ has sufficiently large rank, every atom of $\mathcal{B}$ will be nonempty (by Lemma 2.17).

In light of Observation 3.10, we will need to estimate the size of $\Omega_{B(b)}$ in order to fully understand $\|f\|_{U^3(b)}^P$. For this we will use the following lemma.

Lemma 3.11. *Suppose $\ell, q \geq 0$ are integers, $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q), and $B \in \mathrm{At}(\mathcal{B})$. Then for any $(x, h_1, h_2, h_3) \in G^4$, the following are equivalent.*

(1) $(x, h_1, h_2, h_3) \in \Omega_B$,
(2) the following hold: $x \in B$, $h_i \in L(0)$ for all $i \in [3]$, $2\beta_{\mathcal{Q}}(x, h_i) = -\beta_{\mathcal{Q}}(h_i, h_i)$ for all $i \in [3]$, and $\beta_{\mathcal{Q}}(h_i, h_j) = 0$ for all $i \neq j \in [3]$.

The proof of Lemma 3.11 is straightforward and appears in the appendix. Using Lemma 3.11, we can now estimate the size of the sets of the form Ω_B.

Lemma 3.12. *There is a constant $C > 0$ so that the following holds. Suppose $\ell, q \geq 0$ are integers and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $C(\ell + q + \log_p(\epsilon))$. Then for all $B \in \mathrm{At}(\mathcal{B})$,*

$$|\Omega_B| = (1 \pm \epsilon)p^{4n-4\ell-7q}.$$

Proof. Let C be sufficiently large. By Lemma 3.11,

$$|\Omega_B| = \sum_{b_1,b_2,b_3 \in \mathbb{F}_p^q} \sum_{h_1 \in B(0,b_1)} \sum_{h_2 \in B(0,b_2)} \sum_{h_3 \in B(0,b_3)} \sum_{x \in B} 1_{\beta_{\mathcal{Q}}^{-1}(-b_1)}(x, h_1) 1_{\beta_{\mathcal{Q}}^{-1}(-b_2)}(x, h_2) 1_{\beta_{\mathcal{Q}}^{-1}(-b_3)}(x, h_3) 1_{\beta_{\mathcal{Q}}^{-1}(0)}(h_1, h_2) 1_{\beta_{\mathcal{Q}}^{-1}(0)}(h_1, h_3) 1_{\beta_{\mathcal{Q}}^{-1}(0)}(h_2, h_3).$$

For any $b_1, b_2, b_3 \in \mathbb{F}_p^q$, we have by Lemma 2.20,

$$\sum_{h_1 \in B(0,b_1)} \sum_{h_2 \in B(0,b_2)} \sum_{h_3 \in B(0,b_3)} \sum_{x \in B} 1_{\beta_{\mathcal{Q}}^{-1}(-b_1)}(x, h_1) 1_{\beta_{\mathcal{Q}}^{-1}(-b_2)}(x, h_2) 1_{\beta_{\mathcal{Q}}^{-1}(-b_3)}(x, h_3) 1_{\beta_{\mathcal{Q}}^{-1}(0)}(h_1, h_2) 1_{\beta_{\mathcal{Q}}^{-1}(0)}(h_1, h_3) 1_{\beta_{\mathcal{Q}}^{-1}(0)}(h_2, h_3)$$
$$= (1 \pm \epsilon)p^{4n-4\ell-10q}.$$

Combining these equations yields that $|\Omega_B| = \sum_{b_1,b_2,b_3 \in \mathbb{F}_p^q} (1 \pm \epsilon)p^{4n-4\ell-10q} = (1 \pm \epsilon)p^{4n-4\ell-7q}$. □

3.2. **The local U^3-norms of [63].** In this subsection we introduce local U^3 norms defined by the second author and Wolf in [63], after which we state a crucial result, also from [63], which gives a sufficient condition for a set of bounded VC_2-dimension to have density near 0 or 1 on an atom of a quadratic factor in terms of this local norm (see Theorem 1.3). Before we define these local norms, we require several definitions.

First, we introduce notation aimed at computing, given a triple $(x, y, z) \in G^3$, which atom of a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ contains the sum $x + y + z$. It turns out that where $x + y + z$ lands depends on the labels associated to x, y, and z (i.e. $\beta_{\mathcal{B}}(x)$, $\beta_{\mathcal{B}}(y)$, and $\beta_{\mathcal{B}}(z)$ from Definition 2.10), as well as the values of the bilinear form on the pairs (i.e. $\beta_{\mathcal{Q}}(x, y)$, $\beta_{\mathcal{Q}}(x, y)$ and $\beta_{\mathcal{Q}}(y, z)$ from Definition 2.7). This is the motivation for the following notation, which considers the set of triples carrying a fixed tuple of these labels.

Definition 3.13. Suppose $\ell, q \geq 0$ are integers and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q). Given tuples

$$d_1 = (d_a, d_b, d_c) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \text{ and } d_2 = (d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^q)^3,$$

define

$$K_{1,1,1}(d_1, d_2) = \{(x, y, z) \in G^3 : \beta_{\mathcal{B}}(x) = d_a, \beta_{\mathcal{B}}(y) = d_b, \beta_{\mathcal{B}}(z) = d_c, \text{ and } \beta_{\mathcal{Q}}(x, y) = d_{ab}, \beta_{\mathcal{Q}}(x, z) = d_{ac}, \beta_{\mathcal{Q}}(y, z) = d_{bc}\}.$$

The notation $K_{1,1,1}(d_1, d_2)$ is meant to remind the reader that $K_{1,1,1}(d_1, d_2)$ consists of copies of $K_{1,1,1}$ (the complete tripartite graphs with parts of size 1) in an auxiliary graph which depends on the tuple (d_1, d_2). It is clear that the sets of the form $K_{1,1,1}(d_1, d_2)$ partition G^3, and it turns out that triples in the same piece of this partition always sum into the same atom. To make this precise, we will use the following notation.

Definition 3.14. Suppose $\ell, q \geq 0$ are integers and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor of complexity (ℓ, q). Given tuples

$$d_1 = (d_a, d_b, d_c) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \text{ and } d_2 = (d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^q)^3,$$

define

$$\Sigma(d_1, d_2) = d_a + d_b + d_c + 2(0d_{ab}) + 2(0d_{ac}) + 2(0d_{bc}),$$

where for each $xy \in \{ab, ac, bc\}$, $0d_{xy}$ denotes the tuple obtained by adjoining ℓ many 0's to the left of d_{xy}.

Observation 3.15. *In the notation of Definition 3.14, we have that $(x, y, z) \in K_{1,1,1}(d_1, d_2)$ implies $x + y + z \in B(\Sigma(d_1, d_2))$.*

We now define the local norm from [63], which we decorate with a TW to distinguish it from Definition 3.6. We note that unlike Definition 3.6, this norm is not localized using the label of a single specific atom, but instead a tuple of labels $d = (d_1, d_2)$ as in Definition 3.13 (see Notation 2.19 for the definition of the normalized indicator functions used below).

Definition 3.16 (Local U^3 norm: Terry-Wolf version)**.** Suppose $\ell, q \geq 0$ are integers, $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q), and

$$d_1 = (d_a, d_b, d_c) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \text{ and } d_2 = (d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^q)^3.$$

Then for $d = (d_1, d_2)$ and $f : G \to [-1, 1]$, define $\|f\|_{U^3(d)}^{TW}$ by the following equation.

$$\left(\|f\|_{U^3(d)}^{TW}\right)^8 = \mathbb{E}_{x_0, x_1 \in B(d_a)} \mathbb{E}_{y_0, y_1 \in B(d_b)} \mathbb{E}_{z_0, z_1 \in B(d_c)} \prod_{(i,j) \in [2]^2} \mu_{\beta_{\mathcal{Q}}^{-1}(d_{ab})}(x_i, y_j) \mu_{\beta_{\mathcal{Q}}^{-1}(d_{ac})}(x_i, z_j) \mu_{\beta_{\mathcal{Q}}^{-1}(d_{bc})}(y_i, z_j) \prod_{(i,j,k) \in [2]^3} f(x_i + y_j + z_k).$$

The second author and Wolf proved a U^3 regularity lemma in terms of these local norms (Proposition 4.2 of [65]), in addition to a corresponding counting lemma (Propositions 3.19

and D.11 in [67]).[9] For the purposes of this paper, the following corollary of said counting lemma is key. Specifically, Theorem 3.17 below gives a sufficient condition for when a set of bounded VC_2-dimension has density near 0 or 1 on a quadratic atom, in terms of this type of local U^3 norm.

Theorem 3.17 (Proposition 4.3 of [67])**.** *For all integers $k \geq 1$ there is a constant $C > 0$ so that the following holds. Suppose $\epsilon \in (0,1)$, $\ell, q \geq 0$ are integers[10], and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) and rank[11] at least $C(\ell + q + \log_p(\epsilon^{-1}))$. Assume $A \subseteq \mathbb{F}_p^n$ satisfies $\mathrm{VC}_2(A) \leq k$.*

Let $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$, and let $\alpha_{B(b)}$ denote the density of A on the atom $B(b)$. If there exists some $d \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^6$ such that $\Sigma(d) = b$ and $\|1_A - \alpha_{B(b)}\|_{U^3(d)}^{TW} < (\epsilon/4)^{m^2 2^{m^2}}$, then $\alpha_{B(b)} \in [0, \epsilon) \cup (1 - \epsilon, 1]$.

To use Theorem 3.17, we will need to relate the norm of Definition 3.16 with that from Definition 3.6. We do this in the next subsection. Preparing for that, it will be convenient to have an expression for $\|f\|_{U^3(d)}^{TW}$ which hides the normalization. For this we will us the following notation.

Definition 3.18. Let $\ell, q \geq 0$ be integers, and let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q). Given tuples

$$d_1 = (d_a, d_b, d_c) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3, \text{ and } (d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^q)^3,$$

define

$$\begin{aligned} K_{2,2,2}(d_1, d_2) = \{&(x_1, x_2, y_1, y_2, z_1, z_2) \in G^6 : \\ &\text{for each } i \in [2],\ \beta_{\mathcal{B}}(x_i) = d_a,\ \beta_{\mathcal{B}}(y_i) = d_b,\ \beta_{\mathcal{B}}(z_i) = d_c, \text{ and} \\ &\text{for each } (i,j) \in [2]^2,\ \beta_{\mathcal{Q}}(x_i, y_j) = d_{ab},\ \beta_{\mathcal{Q}}(x_i, z_j) = d_{ac},\ \beta_{\mathcal{Q}}(y_i, z_j) = d_{bc}\}. \end{aligned}$$

We use the notation $K_{2,2,2}(d_1, d_2)$ to indicate this consists of copies of $K_{2,2,2}$ (the complete tripartite graph with parts of size 2) in a certain auxiliary graph depending on the tuple (d_1, d_2). In the case where the factor $\mathcal{B}$ has high rank, we can express $\|\cdot\|_{U^3(d)}^{TW}$ as approximately a sum over tuples from $K_{2,2,2}(d_1, d_2)$.

Proposition 3.19. *There is a constant $C = C(p)$ so that the following holds. Suppose $\epsilon \in (0,1)$, $\ell, q \geq 0$ are integers, $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) of rank at least $C(\ell + q + \log_p(\epsilon^{-1}))$, and*

$$d_1 = (d_a, d_b, d_c) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \text{ and } d_2 = (d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^q)^3.$$

[9]These results were originally proved in [67], but were recently separated into their own paper [67] with reworked proofs and a new emphasis on the local norms.

[10]The ℓ and q in [67] are implicitly assumed to be at least 1, but it is not difficult to see the proofs extend to the cases where one or both are 0.

[11]An inspection of the proofs in [67] shows such a rank assumption suffices.

Then for $d = (d_1, d_2)$ and $f : G \to [-1, 1]$,

$$\left(\|f\|^{TW}_{U^3(d)}\right)^8 = (1 \pm \epsilon)p^{-6n+6\ell+18q} \sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k).$$

Proof. By definition of $\|f\|^{TW}_{U^3(d)}$ and $K_{2,2,2}(d)$,

$$\left(\|f\|^{TW}_{U^3(d)}\right)^8 = \left(\frac{1}{|B(d_a)||B(d_b)||B(d_c)|}\right)^2 \left(\frac{1}{|\beta_{\mathcal{Q}}^{-1}(d_{ab})|\beta_{\mathcal{Q}}^{-1}(d_{ac})||\beta_{\mathcal{Q}}^{-1}(d_{bc})|}\right)^4 \sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k).$$

Combining this with Lemmas 2.17 and 2.18, it is not difficult to see that if $\mathcal{B}$ has rank at least $C(\ell + q + \log_p(\epsilon^{-1}))$ for a sufficiently large C, the above will imply

$$\left(\|f\|^{TW}_{U^3(d)}\right)^8 = (1 \pm \epsilon)p^{-6n+6\ell+18q} \sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k).$$

□

3.3. **Relating the local U^3-norms.** The goal of this subsection is to show that when the factor in question has high rank, the local norm $\|\cdot\|^P_{U^3(b)}$ of Definition 3.6 is approximately equal to the local norm $\|\cdot\|^{TW}_{U^3(d)}$ of Definition 3.16, for any tuple d satisfying $\Sigma(d) = b$ (see Definition 3.14). Specifically, we will prove the following.

Theorem 3.20. *There is a constant $C > 0$ so that the following holds. Let $\epsilon \in (0, 1)$, let $\ell, q \geq 0$ be integers, and let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $C(\ell + q + \log_p(\epsilon^{-1}))$.*

For any $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ and $d \in (\mathbb{F}_p^{\ell+q})^3 \times (\mathbb{F}_p^q)^3$ such that $\Sigma(d) = b$, and any function $f : G \to [-1, 1]$,

$$\left(\|f\|^{TW}_{U^3(d)}\right)^8 = \left(\|f\|^P_{U^3(b)}\right)^8 \pm \epsilon.$$

We start by reviewing a well known reformulation of the global U^3 norm, which uses a change of variables to rewrite the norm as a sum over 6-tuples, rather than 4-tuples (Fact 3.23 below). We first set notation for a change of variables function, which takes a 6-tuple of group elements to a 4-tuple.

Definition 3.21. Let $G = \mathbb{F}_p^n$. Define $\Psi : G^6 \to G^4$ by setting

$$\Psi(x_1, x_2, y_1, y_2, z_1, z_2) = (x_1 + y_1 + z_1, x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

Observation 3.22. *The function Ψ in Definition 3.21 is a surjection, all of whose fibres have size $|G|^2$. Moreover, for any function $f : G \to [-1, 1]$ and any $(x_1, x_2, y_1, y_2, z_1, z_2) \in G^6$, if $\Psi(x_1, x_2, y_1, y_2, z_1, z_2) = (x, h_1, h_2, h_3)$, then the following holds (recall Notation 3.2).*

$$\prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k) = \pi_f(x, h_1, h_2, h_3).$$

Proof. Fix $(x, h_1, h_2, h_3) \in G^4$. Choose any $(x_1, y_1, z_1) \in G^3$ satisfying $x = x_1 + y_1 + z_1$, and let $x_2 = x_1 + h_1$, $y_2 = y_1 + h_2$, and $z_2 = z_1 + h_3$ (there are $|G|^2$ ways to do this). It is an easy exercise to check that $(x_1, x_2, y_1, y_2, z_1, z_2) \in \Psi^{-1}(x, h_1, h_2, h_3)$, and moreover, different choices of such (x_1, y_1, z_1) give rise to distinct elements of $\Psi^{-1}(x, h_1, h_2, h_3)$. On the other hand, by definition, any element of $\Psi^{-1}(x, h_1, h_2, h_3)$ arises in this way. This shows $|\Psi^{-1}(x, h_1, h_2, h_3)| = |G|^2$. The moreover statement regarding products is obvious from the definition of Ψ. □

Observation 3.22 immediately implies the following well known reformulation of the global U^3 norm from Definition 3.1.

Fact 3.23. *For any function $f : G = \mathbb{F}_p^n \to [-1, 1]$,*

$$\|f\|_{U^3}^8 = p^{-2n} \sum_{x_0, x_1 \in G} \sum_{y_0, y_1 \in G} \sum_{z_0, z_1 \in G} \prod_{(i,j,k) \in [2]^3} f(x_i + y_j + z_k).$$

The change of variable formula Ψ will play a crucial role in relating the local norms of Definitions 3.6 and 3.16. Specifically, we will need to analyze the behavior of Ψ on subsets of G^6 of the form $K_{2,2,2}(d)$ (see Definition 3.18). Our first observation in this direction is basically immediate from the relevant definitions.

Observation 3.24. *Let $\ell, q \geq 0$ be integers, and let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q). Given $d \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$, $K_{2,2,2}(d) \subseteq \Psi^{-1}(\Omega_{B(\Sigma(d))})$.*

Proof. Fix $(x_1, x_2, y_1, y_2, z_1, z_2) \in K_{2,2,2}(d)$. By Observation 3.15, we have that for each $(i, j, k) \in [2]^3$, $x_i + y_j + z_k \in B(\Sigma(d))$. Combining this with Observation 3.22, we have

$$1 = \prod_{(i,j,k) \in [2]^3} 1_{B(\Sigma(d))}(x_i + y_j + z_k) = \pi_{1_{B(\Sigma(d))}}(\Psi(x_1, x_2, y_1, y_2, z_1, z_2)).$$

Thus by definition, $\Psi(x_1, x_2, y_1, y_2, z_1, z_2) \in \Omega_{B(\Sigma(d))}$, i.e. $(x_1, x_2, y_1, y_2, z_1, z_2) \in \Psi^{-1}(\Omega_{B(\Sigma(d))})$. □

The above observation shows $K_{2,2,2}(d)$ is always covered by $\Psi^{-1}(\Omega_{B(\Sigma(d))})$, which in turn can be written as a union of fibres

$$\Psi^{-1}(\Omega_{B(\Sigma(d))}) = \bigcup_{z \in \Omega_{B(\Sigma(d))}} \Psi^{-1}(z).$$

Our next task involves understanding how the set $K_{2,2,2}(d)$ is distributed across the individual fibres in this union. Our first step in this direction is to give an algebraic description of the intersection of $K_{2,2,2}(d)$ with such a fibre. We note the proof is straightforward, although a bit tedious.

Proposition 3.25. *There is a constant $C > 0$ so that the following holds. Let $\ell, q \geq 0$ be integers, and let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a factor on $G = \mathbb{F}_p^n$ complexity (ℓ, q) and rank at least $C(\ell + q + \log_p(\epsilon))$.*

Fix $e \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ and $d = (d_a, d_b, d_c, d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$ such that $\Sigma(d) = e$, and let $B = B(e)$. Given $(w, h_a, h_b, h_c) \in \Omega_B$, we have

$$\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)$$
$$= \{(x, x + h_a, y, y + h_b, w - x - y, w - x - y + h_c) : x \in X, y \in Y, \text{ and } \beta_{\mathcal{Q}}(x, y) = d_{ab}\},$$

where

$$X = \{x \in B(d_a) : \beta_{\mathcal{Q}}(x, h_b) = \beta_{\mathcal{Q}}(x, h_c) = 0, 2\beta_{\mathcal{Q}}(x, h_a) = -\beta_{\mathcal{Q}}(h_a, h_a), \beta_{\mathcal{Q}}(x, w) = d_a + d_{ac} + d_{ab}\},$$

and

$$Y = \{y \in B(d_b) : \beta_{\mathcal{Q}}(y, h_a) = \beta_{\mathcal{Q}}(y, h_c) = 0, 2\beta_{\mathcal{Q}}(y, h_b) = -\beta_{\mathcal{Q}}(h_b, h_b), \beta_{\mathcal{Q}}(y, w) = d_b + d_{bc} + d_{ab}\}.$$

Proof. We first show the $\subseteq$ direction of the desired equality. Fix an element $(x_1, x_2, y_1, y_2, z_1, z_2)$ of $\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)$. Since $\Psi(x_1, x_2, y_1, y_2, z_1, z_2) = (w, h_a, h_b, h_c)$, we know $w = x_1 + y_1 + z_1$, $x_2 = x_1 + h_a$, $y_2 = y_1 + h_b$, and $z_2 = z_1 + h_c$. These equalities imply

$$(x_1, x_2, y_1, y_2, z_1, z_2) = (x_1, x_1 + h_a, y_1, y_1 + h_b, w - x_1 - y_1, w - x_1 - y_1 + h_c).$$

Since $(x_1, x_2, y_1, y_2, z_1, z_2) \in K_{2,2,2}(d)$, we already know $\beta_{\mathcal{Q}}(x_1, y_1) = d_{ab}$. Thus, it suffices to show $x \in X$ and $y \in Y$. We first show $x \in X$. Since $(x_1, x_2, y_1, y_2, z_1, z_2) \in K_{2,2,2}(d)$, we know $x_1 \in B(d_a)$. We also know from this that $x_2 \in B(d_a)$. Since $x_2 = x_1 + h_2$, this implies

$$d_a = \beta_{\mathcal{Q}}(x_1 + h_a, x_1 + h_a) = d_a + 2\beta_{\mathcal{Q}}(x_1, h_a) + \beta_{\mathcal{Q}}(h_a, h_a).$$

This implies $2\beta_{\mathcal{Q}}(x_1, h_a) = -\beta_{\mathcal{Q}}(h_a, h_a)$. On the other hand, we know $d_{ab} = \beta_{\mathcal{Q}}(x_1, y_2) = \beta_{\mathcal{Q}}(x_1, y_1)$, and $y_2 = y_1 + h_b$. These imply

$$d_{ab} = \beta_{\mathcal{Q}}(x_1, y_2) = \beta_{\mathcal{Q}}(x_1, y_1 + h_b) = d_{ab} + \beta_{\mathcal{Q}}(x_1, h_b),$$

which implies $\beta_{\mathcal{Q}}(x_1, h_b) = 0$. Similarly, we have

$$d_{ac} = \beta_{\mathcal{Q}}(x_1, z_2) = \beta_{\mathcal{Q}}(x_1, z_1 + h_c) = d_{ab} + \beta_{\mathcal{Q}}(x_1, h_c),$$

which implies $\beta_{\mathcal{Q}}(x_1, h_c) = 0$. Since $w = x_1 + y_1 + z_1$, we have

$$\beta_{\mathcal{Q}}(x_1, w) = \beta_{\mathcal{Q}}(x_1, x_1 + y_1 + z_1) = d_a + d_{ab} + d_{ac},$$

where the last equality uses the assumption $(x_1, x_2, y_1, y_2, z_1, z_2) \in K_{2,2,2}(d)$. This finishes our verification that $x \in X$. A symmetric argument shows $y \in Y$.

We now prove the $\supseteq$ direction of the desired inequality. To this end, fix a tuple $(x_1, x_2, y_1, y_2, z_1, z_2) \in G^6$ such that

$$(x_1, x_2, y_1, y_2, z_1, z_2) = (x_1, x_1 + h_a, y_1, y_1 + h_b, w - x_1 - y_1, w - x_1 - y_1 + h_c), \tag{3.1}$$

and such that $x_1 \in X$, $y_1 \in Y$, and $\beta_{\mathcal{Q}}(x_1, y_1) = d_{ab}$. Note (3.1) implies $x_1 + y_1 + z_1 = w$, $h_a = x_2 - x_1$, $h_b = y_2 - y_1$, and $h_c = z_2 - z_1$. Combining with the definition of Ψ, we have

$$\Psi(x_1, x_2, y_1, y_2, z_1, z_2) = (x_1 + y_1 + z_1, x_2 - x_1, y_2 - y_1, z_2 - z_1) = (w, h_a, h_b, h_c).$$

Thus, $(x_1, x_2, y_1, y_2, z_1, z_2) \in \Psi^{-1}(w, h_a, h_b, h_c)$. It now suffices to show $(x_1, x_2, y_1, y_2, z_1, z_2) \in K_{2,2,2}(d)$.

Since $x_1 \in X$ and $y_1 \in Y$, we know $x_1 \in B(d_a)$ and $y_1 \in B(d_b)$. We also know by assumption $\beta_{\mathcal{Q}}(x_1, y_1) = d_{ab}$. Note

$$\beta_{\mathcal{Q}}(x_2, x_2) = \beta_{\mathcal{Q}}(x_1 + h_a, x_1 + h_a) = d_a + 2\beta_{\mathcal{Q}}(x_1, h_a) + \beta_{\mathcal{Q}}(h_a, h_a) = d_a,$$

where the last equality uses that $x_1 \in X$. An identical argument shows $\beta_{\mathcal{Q}}(y_2, y_2) = d_b$. We also have

$$\beta_{\mathcal{Q}}(x_1, y_2) = \beta_{\mathcal{Q}}(x_1, y_1 + h_b) = d_{ab} + \beta_{\mathcal{Q}}(x_1, h_b) = d_{ab},$$

where the second equality uses that $x_1 \in X$. A symmetric argument shows $\beta_{\mathcal{Q}}(y_1, x_2) = d_{ab}$. We have now checked the constraints from the definition of $K_{2,2,2}(d)$ involving only the x_i's and y_j's.

We next deal with the constraints involving z_1 and z_2. Note $w = x_1 + y_1 + z_1$ and $w \in B(e)$ implies

$$\begin{aligned}
\beta_{\mathcal{Q}}(z_1, z_1) &= \beta_{\mathcal{Q}}(w - x_1 - y_1, w - x_1 - y_1) \\
&= e + d_a + d_c - 2\beta_{\mathcal{Q}}(x_1, y_1) - 2\beta_{\mathcal{Q}}(x_1, w) - 2\beta_{\mathcal{Q}}(y_1, w) \\
&= e + d_a + d_c + 2d_{ab} - 2(d_a + d_{ac} + d_{ab}) - 2(d_c + d_{bc} + d_{ab}) \\
&= e - d_a - d_c - 2d_{ab} - 2d_{ac} - 2d_{bc} \\
&= d_b.
\end{aligned}$$

where the second equality uses that $x_1 \in X$ and $y_1 \in Y$, and the last equality uses $\Sigma(d) = e$. We also have, from (3.1) and $x_1 \in X$ and $y_1 \in Y$ that

$$\beta_{\mathcal{Q}}(x_1, z_1) = \beta_{\mathcal{Q}}(x_1, w - x_1 - y_1) = \beta_{\mathcal{Q}}(x_1, w) - d_a - d_{ab} = d_{ac}. \tag{3.2}$$

A symmetric argument shows $\beta_{\mathcal{Q}}(y_1, z_1) = d_{ac}$. Observe (3.1) and $x_1 \in X$ imply

$$\beta_{\mathcal{Q}}(x_1, z_2) = \beta_{\mathcal{Q}}(x_1, z_1 + h_c) = \beta_{\mathcal{Q}}(x_1, z_1) + \beta_{\mathcal{Q}}(x_1, h_c) = d_{ac}.$$

A symmetric argument shows $\beta_{\mathcal{Q}}(y_1, z_2) = d_{bc}$. Similarly (3.1) and (3.2) imply

$$\begin{aligned}
\beta_{\mathcal{Q}}(x_2, z_1) = \beta_{\mathcal{Q}}(x_1 + h_a, z_1) = d_{ac} + \beta_{\mathcal{Q}}(h_a, z_1) &= d_{ac} + \beta_{\mathcal{Q}}(h_a, w - x_1 - y_1) \\
&= d_{ac} + \beta_{\mathcal{Q}}(h_a, w) - \beta_{\mathcal{Q}}(x_1, h_a) - \beta_{\mathcal{Q}}(y_1, h_a) \\
&= d_{ac},
\end{aligned}$$

where the last equality uses that $\beta_{\mathcal{Q}}(y_1, h_a) = 0$ (since $y_1 \in Y$), that $\beta_{\mathcal{Q}}(x_1, h_a) = -2\beta_{\mathcal{Q}}(h_a, h_a)$ (since $x_1 \in X$), and that $\beta_{\mathcal{Q}}(h_a, w) = -2\beta_{\mathcal{Q}}(h_a, h_a)$ (by Lemma 3.11). A symmetric argument shows $\beta_{\mathcal{Q}}(y_2, z_1) = d_{bc}$. Finally, (3.1), (3.2), and $x_1 \in X$ imply

$$\beta_{\mathcal{Q}}(x_1, z_2) = \beta_{\mathcal{Q}}(x_1, z_1 + h_c) = \beta_{\mathcal{Q}}(x_1, z_1) + \beta_{\mathcal{Q}}(x_1, h_c) = d_{ac}.$$

A symmetric argument shows $\beta_{\mathcal{Q}}(y_1, z_2) = d_{bc}$. This finishes the proof. □

We can now give sufficiently good size estimate on the intersection of $K_{2,2,2}(d)$ with the fibres of the form $\Psi^{-1}(z)$ for z in $\Omega_{B(\Sigma(d))}$ (in the case where $\mathcal{B}$ is a high rank factor). In particular, we give a very rough upper bound on the size of any such intersection, and then show that when z satisfies some weak independence conditions, we can compute the size of the intersection almost exactly.

Proposition 3.26. *There is a constant $C > 0$ so that the following holds. Let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a factor on $G = \mathbb{F}_p^n$ complexity (ℓ, q) and rank at least $C(\ell + q + \log_p(\epsilon))$.*

Fix $e \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ and $d = (d_a, d_b, d_c, d_{ab}, d_{ac}, d_{bc}) \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$ such that $\Sigma(d) = e$, and set $B = B(e)$. Then for all $(w, h_a, h_b, h_c) \in \Omega_B$,

$$|\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)| \leq (1 + \epsilon)p^{2n-2\ell-3q}.$$

Moreover, if $\mathcal{L} \cup \{Mw, Mh_a, Mh_b, Mh_c : M \in \mathcal{Q}\}$ is a linearly independent set, then

$$|\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)| = (1 \pm \epsilon)p^{2n-2\ell-11q}.$$

Proof. Let $C > 0$ be a sufficiently large constant. Fix ℓ, q, $\mathcal{B}$, e, d and $B = B(e)$ as in the hypotheses. Fix $(w, h_a, h_b, h_c) \in \Omega_B$. By Proposition 3.25,

$$\begin{aligned}&\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)\\ &= \{(x, x + h_a, y, y + h_b, w - x - y, w - x - y + h_c) : x \in X, y \in Y, \text{ and } \beta_{\mathcal{Q}}(x, y) = d_{ab}\},\end{aligned}$$

where

$$X = \{x \in B(d_a) : \beta_{\mathcal{Q}}(x, h_b) = \beta_{\mathcal{Q}}(x, h_c) = 0, 2\beta_{\mathcal{Q}}(x, h_a) = -\beta_{\mathcal{Q}}(h_a, h_a), \beta_{\mathcal{Q}}(x, w) = d_a + d_{ac} + d_{ab}\},$$

and

$$Y = \{y \in B(d_b) : \beta_{\mathcal{Q}}(y, h_a) = \beta_{\mathcal{Q}}(y, h_c) = 0, 2\beta_{\mathcal{Q}}(y, h_b) = -\beta_{\mathcal{Q}}(h_b, h_b), \beta_{\mathcal{Q}}(y, w) = d_b + d_{bc} + d_{ab}\}.$$

Thus,

$$|\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)| = |\{(x, y) \in X \times Y : \beta_{\mathcal{Q}}(x, y) = d_{ab}\}|.$$

In particular,

$$|\Psi^{-1}(w, h_a, h_b, h_c) \cap K_{2,2,2}(d)| \leq |\{(x, y) \in B(d_a) \times B(d_b) : \beta_{\mathcal{Q}}(x, y) = d_{ab}\}| \leq (1+\epsilon)p^{2n-2\ell-3q},$$

where the last inequality is by Lemmas 2.17, 2.18, and 2.20 (and because $\mathcal{B}$ has high rank).

Suppose now that, moreover, $\mathcal{L} \cup \{Mw, Mh_a, Mh_b, Mh_c : M \in \mathcal{Q}\}$ is a linearly independent set. Let $\mathcal{B}' = (\mathcal{L}', \mathcal{Q}')$ where

$$\mathcal{L}' = \mathcal{L} \cup \{Mw, Mh_a, Mh_b, Mh_c : M \in \mathcal{Q}\} \text{ and } \mathcal{Q}' = \mathcal{Q}.$$

As $\mathcal{L}'$ is linearly independent, this is a factor on $\mathbb{F}_p^n$ of complexity $(\ell', q') = (\ell + 4q, q)$ with the same rank as $\mathcal{B}$. It is not hard to see that both X and Y are atoms of this factor $\mathcal{B}'$, and thus we are trying to compute the size of a set of the form

$$|\beta_{\mathcal{Q}}^{-1}(d_{ac}) \cap (B'(u) \times B'(v))|.$$

Using Lemmas 2.17, 2.18, and 2.20, and our rank assumption, we see this set has size

$$p^{-2n-2(\ell'-q')-q'}(1 \pm \epsilon) = (1 \pm \epsilon)p^{2n-2(\ell+4q+q)-q} = (1 \pm \epsilon)p^{2n-2\ell-11q}.$$

□

We now have the necessary ingredients to prove the main result of this section, Theorem 3.20, which says the local U^3 norms of Definitions 3.6 and 3.16 are approximately the same.

Proof of Theorem 3.20. The reader may wish to review Definitions 3.1, 3.6, and 3.16 as all three "U^3 norms" make appearances in this proof. Let $C > 0$ be a sufficiently large

constant. Fix $\epsilon \in (0,1)$, integers $\ell, q \geq 0$, and a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $C(\ell + q + \log_p(\epsilon))$. Without loss of generality (after possibly replacing ϵ), assume $0 < \epsilon < 1/11$. Suppose $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$ and $d \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$ are such that $\Sigma(d) = b$. Let $f : G \to [-1,1]$ be a function. Recall that by Lemma 3.12 (and because $\mathcal{B}$ has sufficiently high rank), we have

$$|\Omega_{\mathcal{B}}| = (1 \pm \epsilon^2)p^{4n-4\ell-7q}. \tag{3.3}$$

We will work with a partition of $\Omega_{\mathcal{B}}$ consisting of the following two sets.

$\Omega_{good} = \{(x, h_1, h_2, h_3) \in \Omega_B : \mathcal{L} \cup \{xM, h_1M, h_2M, h_3M\}$ is a linearly independent set$\}$ and $\Omega_{bad} = \Omega_B \setminus \Omega_{good}$.

In light of Observations 3.22 and 3.24, we can write the following (see Notation 3.2 as well).

$$\begin{aligned}
&\sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k) \\
&= \sum_{(x,h_1,h_2,h_3)\in\Omega_B} \Big(\sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in\Psi^{-1}(x,h_1,h_2,h_3)\cap K_{2,2,2}(d)} \pi_f(x,h_1,h_2,h_3) \Big) \\
&= \sum_{(x,h_1,h_2,h_3)\in\Omega_B} |\Psi^{-1}(x,h_1,h_2,h_3) \cap K_{2,2,2}(d)| \pi_f(x,h_1,h_2,h_3).
\end{aligned}$$

Combining with Definition 3.1, we have

$$\begin{aligned}
&\Big| \Big(\sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k) \Big) - \Big(p^{2n-2\ell-11q} \|f \cdot 1_B\|_{U^3}^8 \Big) \Big| \\
&= \Big| \Big(\sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i + y_j + z_k) \Big) \\
&- \Big(p^{2n-2\ell-11q} \sum_{(x,h_1,h_2,h_3)\in\Omega_B} \pi_f(x,h_1,h_2,h_3) \Big) \Big| \\
&= \Big| \sum_{(x,h_1,h_2,h_3)\in\Omega_B} \Big(|\Psi^{-1}(x,h_1,h_2,h_3) \cap K_{2,2,2}(d)| - p^{2n-2\ell-11q} \Big) \pi_f(x,h_1,h_2,h_3) \Big| \\
&\leq \Big| \sum_{(x,h_1,h_2,h_3)\in\Omega_{good}} \Big(|\Psi^{-1}(x,h_1,h_2,h_3) \cap K_{2,2,2}(d)| - p^{2n-2\ell-11q} \Big) \pi_f(x,h_1,h_2,h_3) \Big| \\
&+ \Big| \sum_{(x,h_1,h_2,h_3)\in\Omega_{bad}} \Big(|\Psi^{-1}(x,h_1,h_2,h_3) \cap K_{2,2,2}(d)| - p^{2n-2\ell-11q} \Big) \pi_f(x,h_1,h_2,h_3) \Big|,
\end{aligned}$$

where the last inequality is by the triangle inequality and because $\Omega_{\mathcal{B}} = \Omega_{good} \sqcup \Omega_{bad}$ by definition. We consider the two terms in the final sum above separately. First, by applying the triangle inequality, using the fact f is 1-bounded, and applying Proposition 3.26 (with

error parameter ϵ^2), we have

$$\begin{aligned}
&\Big|\sum_{(x,h_1,h_2,h_3)\in\Omega_{good}}\Big(|\Psi^{-1}(x,h_1,h_2,h_3)\cap K_{2,2,2}(d)|-p^{2n-2\ell-11q}\Big)\pi_f(x,h_1,h_2,h_3)\Big|\\
&\leq\sum_{(x,h_1,h_2,h_3)\in\Omega_{good}}\Big||\Psi^{-1}(x,h_1,h_2,h_3)\cap K_{2,2,2}(d)|-p^{2n-2\ell-11q}\Big|\\
&\leq\epsilon^2p^{2n-2\ell-11q}|\Omega_{good}|\\
&\leq\epsilon^2p^{2n-2\ell-11q}|\Omega_B|,
\end{aligned}$$

where the last inequality is because $\Omega_{good}\subseteq\Omega_B$. Turning to the second sum, we have by the triangle inequality, because f is 1-bounded, and by Proposition 3.26 (again with error parameter ϵ^2), that

$$\begin{aligned}
&\Big|\sum_{(x,h_1,h_2,h_3)\in\Omega_{bad}}\Big(|\Psi^{-1}(x,h_1,h_2,h_3)\cap K_{2,2,2}(d)|-p^{2n-2\ell-11q}\Big)\pi_f(x,h_1,h_2,h_3)\Big|\\
&\leq\sum_{(x,h_1,h_2,h_3)\in\Omega_{bad}}\Big(|\Psi^{-1}(x,h_1,h_2,h_3)\cap K_{2,2,2}(d)|+p^{2n-2\ell-11q}\Big)\\
&\leq|\Omega_{bad}|((1+\epsilon^2)p^{2n-2\ell-3q}+p^{2n-2\ell-11q})\\
&\leq2|\Omega_{bad}|(1+\epsilon^2)p^{2n-2\ell-3q}\\
&\leq28(1+\epsilon^2)p^{6n-\ell+q-r}\\
&\leq\epsilon^2p^{2n-2\ell-11q}|\Omega_B|,
\end{aligned}$$

where the second to last inequality uses Lemma 2.21, and the last uses (3.3) and the fact $\mathrm{rk}(\mathcal{B})$ is sufficiently large. We can now conclude

$$\begin{aligned}
&\Big|\Big(\sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)}\prod_{(i,j,k)\in[2]^3}f(x_i+y_j+z_k)\Big)-\Big(p^{2n-2\ell-11q}\Big(\|f\cdot1_B\|_{U^3}\Big)^8\Big)\Big|\\
&\qquad\leq2\epsilon^2p^{2n-2\ell-11q}|\Omega_B|.
\end{aligned}\tag{3.4}$$

By Proposition 3.19 (and since $\mathrm{rk}(\mathcal{B})$ is sufficiently large), we have

$$\Big(\|f\|^{TW}_{U^3(d)}\Big)^8=(1\pm\epsilon^2)p^{-6n+6\ell+18q}\sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)}\prod_{(i,j,k)\in[2]^3}f(x_i+y_j+z_k).\tag{3.5}$$

Combining (3.4) and (3.5), we have

$$\begin{aligned}
&\Big(\|f\|_{U^3(d)}^{TW}\Big)^8 \\
&= (1\pm\epsilon^2)p^{-6n+6\ell+18q} \sum_{(x_1,x_2,y_1,y_2,z_1,z_2)\in K_{2,2,2}(d)} \prod_{(i,j,k)\in[2]^3} f(x_i+y_j+z_k) \\
&= (1\pm\epsilon^2)p^{-6n+6\ell+18q}\Big(p^{2n-2\ell-11q}\Big(\|f\cdot 1_B\|_{U^3}\Big)^8 \pm 2\epsilon^2 p^{2n-2\ell-11q}|\Omega_B|\Big) \\
&= (1\pm\epsilon^2)p^{-4n+4\ell+7q}\Big(\Big(\|f\cdot 1_B\|_{U^3}\Big)^8 \pm 2\epsilon^2|\Omega_B|\Big) \\
&= (1\pm\epsilon^2)p^{-4n+4\ell+7q}|\Omega_B|\Big(\Big(\|f\|_{U^3(e)}^{P}\Big)^8 + 2\epsilon^2\Big) \\
&= (1\pm\epsilon^2)^2\Big(\Big(\|f\|_{U^3(e)}^{P}\Big)^8 + 2\epsilon^2\Big),
\end{aligned}$$

where the last inequality is by (3.3). Since $\epsilon < 1/11$, and since $(\|f\|_{U^3(e)}^{P})^8 \le 1$ (see Observation 3.9), this yields the desired statement $(\|f\|_{U^3(d)}^{TW})^8 = (\|f\|_{U^3(e)}^{P})^8 \pm \epsilon$. □

We can now deduce the key corollary of Theorems 3.17 and 3.20 needed for the proof of our main theorem.

Corollary 3.27. *For all integers $k \ge 1$, there exists a constant $K > 0$ such that the following holds. Suppose $0 < \epsilon < 1/11$, $\ell, q \ge 0$, are integers, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $C(\ell + q + \log_p(\epsilon))$.*

Suppose $A \subseteq \mathbb{F}_p^n$ satisfies $\mathrm{VC}_2(A) \le k$, $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$, and $\|1_A - \alpha_{B(b)}\|_{U^3(b)}^{P} < (\epsilon/4)^{m^2 2^{m^2}} 2^{-1}$, where $\alpha_{B(b)}$ denotes the density of A on the atom $B(b)$. Then $\alpha_{B(b)} \in [0,\epsilon) \cup (1-\epsilon, 1]$.

Proof. Let $C > 0$ be as in Theorem 3.20 and let $K = 2k^2 2^{k^2} C$. Fix $0 < \epsilon < 1/11$, integers $\ell, q \ge 0$, a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $K(\ell + q + \log_p(\epsilon))$, and a label $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$. Assume $A \subseteq \mathbb{F}_p^n$ has VC_2-dimension at most k and satisfies

$$\|1_A - \alpha_{B(b)}\|_{U^3(b)}^{P} < (\epsilon/4)^{m^2 2^{m^2}} 2^{-1}, \tag{3.6}$$

where $\alpha_{B(b)}$ is the density of A on the atom $B(b)$. Fix any tuple $d \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$ satifying $\Sigma(d) = b$. Our choice of K implies $\mathcal{B}$ has rank at least

$$C(\ell + q + \log_p((\epsilon/4)^{m^2 2^{m^2}} 2^{-1})),$$

and thus, Theorem 3.20 and (3.6) imply

$$\|1_A - \alpha_{B(b)}\|_{U^3(d)}^{TW} \le \|1_A - \alpha_{B(b)}\|_{U^3(b)}^{P} + (\epsilon/4)^{m^2 2^{m^2}} 2^{-1} < (\epsilon/4)^{m^2 2^{m^2}}.$$

By Theorem 3.17, we have $\alpha_{B(b)} \in [0,\epsilon) \cup (1-\epsilon, 1]$. □

4. Matrix deletion and addition

In this section we examine the bounds generated by performing sequences of operations on factors. We begin by considering the rank lemma (Lemma 2.12) in detail, which is proved by iteratively applying a matrix deletion operation. We then discuss more complicated procedures, which interweave matrix deletion and addition steps.

Those familiar with the proof of Lemma 2.12 will know it is proved by iteratively deleting a matrix involved in a low rank linear combination, then adding enough vectors to the linear component to recover the deleted information. It will be useful to have the following definition for when one factor arises from another by a single such deletion.

Definition 4.1. Suppose ρ is a growth function, $\ell \geq 0$ and $q \geq 1$ are integers, and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q) with $\mathcal{Q} = \{M_1, \ldots, M_q\}$.

A *ρ-matrix deletion of $\mathcal{B}$* is a quadratic factor $\mathcal{B}' = (\mathcal{L}', \mathcal{Q}')$ on $\mathbb{F}_p^n$ such that for some non-trivial linear combination $U = \lambda_1 M_1 + \ldots + \lambda_q M_q$, the following hold.

(1) $\mathrm{rk}(U) < \rho(\ell + q)$,
(2) there is some $i \in [q]$ such that $\lambda_i \neq 0$ and $\mathcal{Q}' = \mathcal{Q} \setminus \{M_i\}$,
(3) $\mathcal{L}'$ is a minimal set of vectors containing $\mathcal{L}$ and spanning $\ker(U)^\perp$.

If $\mathcal{B}'$ is a ρ-matrix deletion of $\mathcal{B}$, we obtain the following information about the complexity of $\mathcal{B}'$ by definition.

Observation 4.2. *Suppose ρ is a growth function, $\ell \geq 0$ and $q \geq 1$ are integers, and $\mathcal{B}$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q). Suppose $\mathcal{B}'$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ', q'), and assume $\mathcal{B}'$ is a ρ-matrix deletion of $\mathcal{B}$. Then*

$$q' = q - 1 \text{ and } \ell' < \ell + \rho(\ell + q).$$

While it is easy to understand the bounds resulting from a single matrix deletion, we also need to understand what happens to the bounds after a sequence of matrix deletions. Towards this goal, we define functions to help us bound the complexity after performing i many ρ-matrix deletions.

Definition 4.3. Suppose ρ is a growth function. We define functions $\tau_i^\rho : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ for all $i \in \mathbb{N}$ inductively as follows.

- Define τ_0^ρ by setting $\tau_0^\rho(x, y) = x$ for all $(x, y) \in \mathbb{Z} \times \mathbb{Z}$.
- Given $i \geq 0$, define τ_{i+1}^ρ in terms of τ_i^ρ by setting
$$\tau_{i+1}^\rho(x, y) = \tau_i^\rho(x, y) + \rho(\tau_i^\rho(x, y) + y - i),$$
for all $(x, y) \in \mathbb{Z} \times \mathbb{Z}$.

Definition 4.3 is designed to bound the complexity of a factor obtained at the end of a chain of ρ-matrix deletions. This is the content of the following lemma.

Lemma 4.4. *Let ρ be a growth function, and let $m \geq 0$ be an integer. Suppose that for each $0 \leq i \leq m$, $\mathcal{B}_i = (\mathcal{L}_i, \mathcal{Q}_i)$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ_i, q_i), and assume that for each $0 \leq j \leq m - 1$, $\mathcal{B}_{j+1}$ is ρ-matrix deletion of $\mathcal{B}_j$. Then*

$$|\mathcal{L}_m| \leq \tau_m^\rho(\ell_0, q_0)$$

Proof. This is immediate from Definition 4.3 and Observation 4.2. □

As a corollary, we can state the following more quantitative version of Lemma 2.12.

Lemma 4.5. *Let ρ be a growth function, let $\ell, q \geq 0$ be integers, and let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ be a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q). Then there exists a quadratic factor $\mathcal{B}' \preceq \mathcal{B}$ of complexity (ℓ', q') and rank at least $\rho(\ell' + q')$ for some integers ℓ', q' satisfying*

$$\ell \leq \ell' \leq \tau_q^\rho(\ell, q) \text{ and } 0 \leq q' \leq q.$$

Proof. This is immediate from Lemma 4.4 and the proof of Lemma 2.12 from [37], which obtains $\mathcal{B}'$ at the end of a sequence of at most q many ρ-matrix deletions. □

To analyze the bounds in our proofs, we will need the following lemma, which bounds $\tau_i^\rho(x, y)$ for polynomial ρ and certain inputs x and y.

Lemma 4.6. *Suppose $C > 1$ is a real and $k \geq 1$ is an integer, and ρ is a growth function satisfying $\rho(x) \geq x$ and $\rho(x) \leq Cx^k$ for all $x \geq 1$. Then for all integers $i \geq 0$ and all $(x, y) \in \mathbb{N} \times \mathbb{N}^{\geq i}$,*

$$\tau_i^\rho(x, y) \leq 2^{ik^i} C^{ik^i} (x + y)^{k^i}.$$

Proof. We prove this by induction on $i \geq 0$.
Base Case: Suppose first $i = 0$. By definition of τ_0 and since $C > 1$, we have that for any $(x, y) \in \mathbb{N} \times \mathbb{N}$,

$$\tau_0^\rho(x, y) = x \leq x + y = 2^{ik^i} C^{ik^i} (x + y)^{k^i}.$$

Induction Step: Suppose $i \geq 0$, and assume by induction we have shown that for all $(x, y) \in \mathbb{N} \times \mathbb{N}^{\geq i}$, $\tau_i^\rho(x, y) \leq 2^{ik^i} C^{ik^i} (x + y)^{k^i}$. Fix $(x, y) \in \mathbb{N} \times \mathbb{N}^{\geq i+1}$. By definition of τ_{i+1}^ρ, our induction hypothesis, and since ρ is increasing,

$$\begin{aligned}
\tau_{i+1}^\rho(x, y) &= \tau_i^\rho(x, y) + \rho(\tau_i^\rho(x, y) + y - i - 1) \\
&\leq 2^{ik^i} C^{ik^i} (x + y)^{k^i} + \rho(2^{ik^i} C^{k^i} (x + y)^{k^i} + y - i - 1) \\
&\leq 2\rho(2^{ik^i} C^{ik^i} (x + y)^{k^i} + y - i - 1) \\
&\leq 2\rho(2^{ik^i + 1} C^{ik^i} (x + y)^{k^i}),
\end{aligned}$$

where the second inequality is because $\rho(z) \geq z$ for all $z \geq 1$, and the third is because ρ is increasing (this step uses our assumption that $y \geq i + 1$). Using that $\rho(z) \leq Cz^k$ for all $z \in \mathbb{R}^{\geq 0}$, this is at most

$$2C(2^{ik^i + 1} C^{ik^i} (x + y)^{k^i})^k = 2^{ik^{i+1} + k + 1} C^{ik^{i+1} + 1} (x + y)^{k^{i+1}} \leq 2^{(i+1)k^{i+1}} C^{(i+1)k^{i+1}} (x + y)^{k^{i+1}}.$$

This finishes the proof. □

With these tools in hand, we can easily upper bound the linear complexity in Lemma 2.12 when the growth function ρ is bounded above by a polynomial.

Lemma 4.7. *For any polynomial growth function ρ of degree $k \geq 1$ satisfying $\rho(x) \geq x$ for all $x \geq 1$, there is a constant $C > 0$ so that the following holds. Suppose $\ell, q \geq 0$ are integers and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $\mathbb{F}_p^n$ of complexity (ℓ, q). There there exist integers ℓ', q' satisfying $0 \leq q' \leq q$ and*

$$\ell \leq \ell' \leq 2^{qk^q} C^{qk^q} (\ell + q)^{k^q},$$

and a quadratic factor $\mathcal{B}' \preceq \mathcal{B}$ of complexity (ℓ', q') and rank at least $\rho(\ell' + q')$.

Proof. Choose C sufficiently large so that for all $x \geq 1$, $p(x) \leq Cx^k$. The stated bound is then immediate from Lemmas 4.5 and 4.6. □

Lemma 4.7 gives us a good estimate on the bounds resulting from performing many matrix deletions in a row (as occurs in the proof of Lemma 4.5). It also suffices for computing bounds in proofs which alternate Lemma 4.5 with steps that add new linear and quadratic terms to a factor, as long as this alternation follows a simple pattern. An example of this occurs in our energy increment proof of the (non-cylinder) quadratic arithmetic regularity lemma given in Section 5.

However, analyzing the bounds produced by the proof of our cylinder quadratic arithmetic regularity lemma (Theorem 6.1) will be more subtle. The individual steps in this proof will take one of two forms: a single ρ-matrix deletion (we call this a *deletion step*), or the addition of at most one new linear term and at most one new quadratic term to a factor (we call this an *addition step*). We will need to analyze the bounds resulting from a sequence of such steps, performed in an order we have very little information about.

To make our analysis of such a process precise, we will use sequences of -1's and 1's to keep track of the pattern of addition and deletion steps performed, with -1 representing a deletion step, and 1 representing an addition step. For this reason, we will need several definitions and notational conventions related to sequences of -1's and 1's, which we will refer to as *binary strings*.

Notation 4.8 (Binary strings)**.**

- Given an integer $m \geq 1$, a *binary string of length m* is a tuple $\sigma \in \{-1, 1\}^m$.
- By convention, $\{-1, 1\}^0$ contains a unique element, called the *empty string*, which we denote by $<>$.
- Given an integer $m \geq 0$, let

$$\{-1, 1\}^{\leq m} = \bigcup_{i=0}^{m} \{-1, 1\}^i.$$

- Given integers $m, m' \geq 1$, a sequence $\sigma = (\sigma_1, \ldots, \sigma_m) \in \{-1, 1\}^m$, and a sequence $\sigma' = (\sigma'_1, \ldots, \sigma'_{m'}) \in \{-1, 1\}^{m'}$, we let $\sigma \wedge \sigma'$ denote the element of $\{-1, 1\}^{m+m'}$ obtained by appending σ' to the end of σ, i.e.

$$\sigma \wedge \sigma' := (\sigma_1, \ldots, \sigma_m, \sigma'_1, \ldots, \sigma'_{m'}).$$

- For integers $m_1 \geq m_2 \geq 1$ and $\sigma = (\sigma_1, \ldots, \sigma_{m_1}) \in \{-1,1\}^{m_1}$, define the *restriction of σ to $[m_2]$* to be
$$\sigma|_{[m_2]} = (\sigma_1, \ldots, \sigma_{m_2}) \in \{-1,1\}^{m_2}.$$
- By convention, for any $m \geq 0$ and $\sigma \in \{-1,1\}^m$, $<> \wedge \sigma = \sigma \wedge <> = \sigma$ and $\sigma|_{[0]} = <>$.

We will use the following simple observation several times.

Observation 4.9. *For any integer $m > 0$ and $\sigma \in \{-1,1\}^m$, there is $\sigma' \in \{-1,1\}^{m-1}$ and $u \in \{-1,1\}$ such that $\sigma = \sigma' \wedge u$.*

In our proofs, it will be important to track the number of 1's appearing in a binary string, as well as the discrepancy between the number of 1's and the number of -1's. For this we will use the following notation.

Definition 4.10. Given $m \geq 1$ and $\sigma = (\sigma_1, \ldots, \sigma_m) \in \{-1,1\}^m$, define
$$|\sigma| = |\{i \in [m] : \sigma_i = 1\}|,$$
and define the *discrepancy* of σ to be
$$\mathrm{disc}(\sigma) = \sum_{i=1}^{m} \sigma_i.$$
By convention, we set $| <> | = 0$ and $\mathrm{disc}(<>) = 0$.

Our next definition will help us understand quadratic factors which arise from performing a sequence of matrix addition and deletion steps, starting with the trivial factor. Informally, a (ρ, σ)-chain will be a sequence of factors, each of which is obtained from the preceding factor by performing either a deletion or addition step, and where the pattern of steps is dictated by the binary string σ. It will be convenient later to also have a definition for a (ρ, σ)-chain when σ is the empty string. Such a chain will consist of the trivial factor.

Definition 4.11. Suppose ρ is a growth function.

(1) A $(\rho, <>)$-*chain on* $\mathbb{F}_p^n$ is simply the trivial factor $\mathcal{B}_0 = (\mathcal{L}_0, \mathcal{Q}_0)$ where $\mathcal{L}_0 = \mathcal{Q}_0 = \emptyset$.
(2) Given an integer $m \geq 1$ and $\sigma = (\sigma_1, \ldots, \sigma_m) \in \{-1,1\}^m$, a (ρ, σ)-*chain on* $\mathbb{F}_p^n$ is a chain of quadratic factors on $\mathbb{F}_p^n$,
$$\mathcal{B}_0 = (\mathcal{L}_0, \mathcal{Q}_0) \preceq \cdots \preceq \mathcal{B}_m = (\mathcal{L}_m, \mathcal{Q}_m),$$
such that $\mathcal{L}_0 = \mathcal{Q}_0 = \emptyset$, and such that for each $0 \leq i \leq m-1$, one of the following hold.
   - $\sigma_i = -1$ and $\mathcal{B}_{i+1}$ is a ρ-matrix deletion of $\mathcal{B}_i$,
   - $\sigma_i = 1$ and $\mathcal{L}_i \subseteq \mathcal{L}_{i+1}$, $\mathcal{Q}_i \subseteq \mathcal{Q}_{i+1}$, $|\mathcal{L}_{i+1}| \leq |\mathcal{L}_i| + 1$, and $|\mathcal{Q}_{i+1}| \leq |\mathcal{Q}_i| + 1$.

We now state Definition 4.12, whose goal is to define a pair of real numbers, denoted f_σ^ρ, to serve as a bound on the complexity of a factor obtained at the end of a (ρ, σ)-chain.

Definition 4.12. Suppose ρ is a growth function. For each integer $m \geq 0$ and $\sigma \in \{-1,1\}^m$, we define a value $f_\sigma^\rho \in \mathbb{R} \times \mathbb{R}$ by induction as follows.

(1) Define $f^\rho_{<>} = (0,0)$.
(2) Suppose $m > 0$ and we have defined $f^\rho_\sigma \in \mathbb{R} \times \mathbb{R}$ for each $\sigma \in \{-1,1\}^m$. Then for each $\sigma \in \{-1,1\}^m$, define $f^\rho_{\sigma\wedge 1}$ and $f^\rho_{\sigma\wedge -1}$ as follows, in terms of $(a,b) = f^\rho_\sigma$:

$$f^\rho_{\sigma\wedge 1} = (a+1, b+1) \text{ and } f^\rho_{\sigma\wedge -1} = (a + \rho(a+b), b-1).$$

Our next lemma shows Definition 4.12 can be used to bound complexities of factors arising from (ρ,σ)-chains. It moreover shows the discrepancy of σ can be used to bound the quadratic complexity of the final factor. We note this will implicitly tell us that only certain σ give rise to (ρ,σ)-chains, as not all σ have non-negative discrepancy.

Lemma 4.13. *Let ρ be a growth function, let $m \geq 0$ be an integer, and let $\sigma \in \{-1,1\}^m$. Suppose*

$$\mathcal{B}_0 = (\mathcal{L}_0, \mathcal{Q}_0) \preceq \cdots \preceq \mathcal{B}_m = (\mathcal{L}_m, \mathcal{Q}_m)$$

is a (ρ,σ)-chain on $\mathbb{F}_p^n$ and for each $i \in \{0,\ldots,m\}$, (ℓ_i, q_i) denotes the complexity of $\mathcal{B}_i$. Then the following hold.

(a) For all $0 \leq i \leq m$, $0 \leq q_i \leq \operatorname{disc}(\sigma|_{[i]})$,
(b) For all $0 \leq i \leq m$, if $(a_i, b_i) = f^\rho_{\sigma|_{[i]}}$, then $\ell_i \leq a_i$ and $q_i \leq b_i$.

Proof. Fix a growth function ρ. We show the desired conclusion holds for this ρ and all $\sigma \in \{-1,1\}^m$ by induction on $m \geq 0$.
Base Case: If $m = 0$, the desired conclusions hold trivially.
Induction Step: Suppose now $m > 0$, and assume by induction the claim holds for all $0 \leq m' < m$. Fix $\sigma \in \{-1,1\}^m$ and assume

$$\mathcal{B}_0 = (\mathcal{L}_0, \mathcal{Q}_0) \preceq \cdots \preceq \mathcal{B}_m = (\mathcal{L}_m, \mathcal{Q}_m)$$

is a (ρ,σ)-chain on $\mathbb{F}_p^n$. For each $0 \leq i \leq m$, let (ℓ_i, q_i) be the complexity of $\mathcal{B}_i$ and let $(a_i, b_i) = f^\rho_{\sigma|_{[i]}}$. By Observation 4.9, there exists some $\sigma' \in \{-1,1\}^{m-1}$ and $u \in \{-1,1\}$ such that $\sigma = \sigma' \wedge u$. Given $0 \leq i \leq m-1$, since $\sigma'|_{[i]} = \sigma|_{[i]}$, the induction hypothesis implies that $0 \leq q_i \leq \operatorname{disc}(\sigma|_{[i]})$, and $\ell_i \leq a_i$ and $q_i \leq b_i$. We deal with the $i = m$ case using separate arguments based on whether $u = -1$ or $u = 1$.

Suppose first $u = -1$. Then by definition of a (ρ,σ)-chain, we must have that

$$q_m = q_{m-1} - 1 \text{ and } \ell_m \leq \ell_{m-1} + \rho(\ell_{m-1} + q_{m-1}),$$
$$b_m = b_{m-1} - 1 \text{ and } a_m = a_{m-1} + \rho(a_{m-1} + b_{m-1}).$$

Combining these inequalities with $b_{m-1} \leq q_{m-1}$ and $\ell_{m-1} \leq a_{m-1}$ (which hold by induction), and the fact that ρ is increasing, we have

$$q_m = q_{m-1} - 1 \leq b_{m-1} - 1 = b_m \text{ and } \ell_m \leq \ell_{m-1} + \rho(\ell_{m-1} + q_{m-1}) \leq a_{m-1} + \rho(a_{m-1} + b_{m-1}) = a_m.$$

Since (ℓ_m, q_m) is the complexity of $\mathcal{B}_m$, we obviously have $q_m \geq 0$. Since $\sigma = \sigma' \wedge -1$, $\operatorname{disc}(\sigma) = \operatorname{disc}(\sigma') - 1$. By induction, we have $0 \leq q_{m-1} \leq \operatorname{disc}(\sigma')$. Thus,

$$0 \leq q_m = q_{m-1} - 1 \leq \operatorname{disc}(\sigma') - 1 = \operatorname{disc}(\sigma).$$

This finishes the case $u = -1$.

Suppose now $u = 1$. Then by definition of a (ρ, σ)-chain, and Definition 4.12, we must have that

$$q_m \leq q_{m-1} + 1 \text{ and } \ell_m \leq \ell_{m-1} + 1,$$
$$b_m = b_{m-1} + 1 \text{ and } a_m = a_{m-1} + 1.$$

Combining these inequalities with $b_{m-1} \leq q_{m-1}$ and $\ell_{m-1} \leq a_{m-1}$ (which hold by induction), we have

$$q_m \leq q_{m-1} + 1 \leq b_{m-1} + 1 = b_m \text{ and } \ell_m \leq \ell_{m-1} + 1 \leq a_{m-1} + 1 = a_m.$$

Again, we must have $0 \leq q_m$ as (ℓ_m, q_m) is the complexity of the factor $\mathcal{B}_m$. On the other hand, since $\sigma = \sigma' \wedge 1$, $\text{disc}(\sigma) = \text{disc}(\sigma') + 1$. By induction, we have $0 \leq q_{m-1} \leq \text{disc}(\sigma')$. Thus,

$$0 \leq q_m = q_{m-1} + 1 \leq \text{disc}(\sigma') + 1 = \text{disc}(\sigma).$$

This finishes the case $u = 1$. □

In light of Lemma 4.13, we need to obtain bounds on the coordinates of f^ρ_σ from Definition 4.12. We accomplish this through a series of lemmas. Our first lemma in this direction, Lemma 4.14 below, says that if σ_1 and σ_2 are two binary sequences of the same length such that $f^\rho_{\sigma_1}$ is (coordinatewise) bounded by $f^\rho_{\sigma_2}$, then this remains true after appending a binary string μ to both σ_1 and σ_2. This statement appears technical at first, but the idea is that when the sequence of matrix deletions and additions dictated by σ_2 generates a factor with worse bounds than the one generated by σ_1, then this cannot be reversed by applying identical deletion and addition steps to the two resulting factors.

Lemma 4.14. *Suppose ρ is a growth function, $t \geq 0$ is an integer, and $\sigma_1, \sigma_2 \in \{-1, 1\}^t$ are such that $a_1 \leq a_2$ and $b_1 \leq b_2$ where $(a_1, b_1) = f^\rho_{\sigma_1}$ and $(a_2, b_2) = f^\rho_{\sigma_2}$.*

Then for all integers $s \geq 0$ and all $\mu \in \{-1, 1\}^s$, if $f^\rho_{\sigma_1 \wedge \mu} = (c_1, d_1)$ and $f^\rho_{\sigma_2 \wedge \mu} = (c_2, d_2)$, then $c_1 \leq c_2$ and $d_1 \leq d_2$. Moreover, if $b_1 = b_2$ then $d_1 = d_2$.

Proof. Fix an integer $t \geq 0$ and $\sigma_1, \sigma_2 \in \{-1, 1\}^t$ such that $a_1 \leq a_2$ and $b_1 \leq b_2$ where $(a_1, b_1) = f^\rho_{\sigma_1}$ and $(a_2, b_2) = f^\rho_{\sigma_2}$. We prove the claim holds for this σ_1 and σ_2 and all $\mu \in \{-1, 1\}^s$ by induction on $s \geq 0$.

Base Case: If $s = 0$ there is nothing to show (since $\sigma_1 \wedge <> = \sigma_1$ and $\sigma_2 \wedge <> = \sigma_2$).

Induction Step: Suppose $s \geq 0$, and assume by induction the claim holds for s. Fix $\mu \in \{-1, 1\}^{s+1}$, and let $(c_1, d_1) = f^\rho_{\sigma_1 \wedge \mu}$ and $(c_2, d_2) = f^\rho_{\sigma_2 \wedge \mu}$. By Observation 4.9, μ can be written as $\mu' \wedge v$ for some $\mu' \in \{-1, 1\}^s$ and $v \in \{-1, 1\}$. Let $(c_1', d_1') = f^\rho_{\sigma_1 \wedge \mu'}$ and $(c_2', d_2') = f^\rho_{\sigma_2 \wedge \mu'}$. By the induction hypothesis, $c_1' \leq c_2'$ and $d_1' \leq d_2'$. We now deal proceed in cases based on whether $v = 1$ or $v = -1$.

Suppose first $v = 1$. Then

$$(c_1, d_1) = f^\rho_{\sigma_1 \wedge \mu' \wedge v} = (c_1' + 1, d_1' + 1) \text{ and } (c_2, d_2) = f^\rho_{\sigma_2 \wedge \mu' \wedge v} = (c_2' + 1, d_2' + 1).$$

Combining with the induction hypothesis, this implies $c_1 = c_1' + 1 \leq c_2' + 1 = c_2$ and $d_1 = d_1' + 1 \leq d_2' + 1 = d_2$, as desired. If we also knew that $b_1 = b_2$ holds, then the induction hypothesis implies $d_1' = d_2'$, and consequently, $d_1 = d_1' + 1 = d_2' + 1 = d_2$. This finishes the $v = 1$ case.

Suppose now $v = -1$. Then

$(c_1, d_1) = f^\rho_{\sigma_1 \wedge \mu' \wedge v} = (c'_1 + \rho(c'_1 + d'_1), d'_1 - 1)$ and $(c_2, d_2) = f^\rho_{\sigma_2 \wedge \mu' \wedge v} = (c'_2 + \rho(c'_2 + d'_2), d'_2 - 1)$. Combining this with the fact ρ is increasing and the induction hypothesis, we have

$$c_1 = c'_1 + \rho(c'_1 + d'_1) \leq c'_2 + \rho(c'_2 + d'_2) = c_2.$$

This also implies (again using the induction hypothesis) that $d_1 = d'_1 - 1 \leq d'_2 - 1 = d_2$. We have now shown $c_1 \leq c_2$ and $d_1 \leq d_2$. Suppose that $b_1 = b_2$ was also true. Then, by the induction hypothesis, $d'_1 = d'_2$. Consequently, $d_1 = d'_1 - 1 = d'_2 - 1 = d_2$. This finishes case $v = -1$, and thus the proof. □

Using Lemma 4.14, we now prove that given a binary string σ, altering σ by replacing an instance of $(-1, 1)$ with $(1, -1)$ does not change the second coordinate of f^ρ_σ, and can only make the first coordinate go up.

Lemma 4.15. *Let ρ be a growth function, and let $m \geq 2$ be an integer. Suppose that $\sigma = (\sigma_1, \ldots, \sigma_m) \in \{-1, 1\}^m$ is such that for some $1 \leq i \leq m - 1$, $\sigma_i = -1$ and $\sigma_{i+1} = 1$. Let $\phi = (\phi_1, \ldots, \phi_m) \in \{-1, 1\}^m$ be defined by setting $\phi_i = 1$, $\phi_{i+1} = -1$, and $\phi_j = \sigma_j$ for all $j \in [m] \setminus \{i, i + 1\}$.*

Suppose $f^\rho_\sigma = (a_\sigma, b_\sigma)$ and $f^\rho_\phi = (a_\phi, b_\phi)$. Then $a_\sigma \leq a_\phi$ and $b_\sigma = b_\phi$.

Proof. Let $a, b \in \mathbb{R}$ be such that $f^\rho_{\sigma|_{[i-1]}} = f^\rho_{\phi|_{[i-1]}} = (a, b)$ (these values agree by definition of ϕ). Since $\sigma_i = 0$ and $\phi_i = 1$, we have $f^\rho_{\sigma|_{[i]}} = (a + \rho(a + b), b - 1)$ and $f^\rho_{\phi|_{[i]}} = (a + 1, b + 1)$. Further, since $\sigma_{i+1} = 1$ and $\phi_{i+1} = 0$, we have

$$f^\rho_{\sigma|_{[i+1]}} = (a + \rho(a + b) + 1, b) \text{ and } f^\rho_{\phi|_{[i+1]}} = (a + \rho(a + b + 2) + 1, b).$$

Since ρ is increasing, $a + \rho(a + b) + 1 \leq a + \rho(a + b + 2) + 1$. By Lemma 4.14, we can conclude $a_\sigma \leq a_\phi$ and $b_\sigma = b_\phi$.

□

By iterating Lemma 4.15, one can show that, among the sequences with a given number of 1's appearing, the coordinates of f^ρ_σ are maximized when σ begins with all 1's and ends with all -1's.

Lemma 4.16. *Let ρ be a growth function, let $m \geq 1$ and $m \geq k \geq 0$ be integers, and let $\sigma \in \{-1, 1\}^m$ satisfy $|\sigma| = k$ (see Definition 4.10).*

Define $\theta = (\theta_1, \ldots, \theta_m) \in \{-1, 1\}^m$ by setting $\theta_i = 1$ if $1 \leq i \leq k$ and $\theta_i = -1$ if $k + 1 \leq i \leq m$.

If $f^\rho_\sigma = (a_\sigma, b_\sigma)$ and $f^\rho_\theta = (a_\theta, b_\theta)$, then $a_\sigma \leq a_\theta$ and $b_\sigma = b_\theta$.

Proof. This follows from Lemma 4.15 and the fact that θ can be obtained from σ by a sequence of transpositions, each of which switches an instance of $(-1, 1)$ to $(1, -1)$. □

We are almost ready to state the main result needed for the analysis of our bounds in Theorem 6.1. While the preceding three lemmas considered arbitrary binary sequences, in what follows we need only consider those σ which can actually give rise to (ρ, σ)-chains.

As mentioned above, by Lemma 4.13, we can restrict our attention to σ with non-negative discrepancy.

We now compute bounds for f^ρ_σ when σ has non-negative discrepancy. The idea is that, by Lemma 4.16, it suffices to compute bounds for the case where σ starts with all 1's then switches to all -1's. Since σ has non-negative discrepancy, the number of -1's at the end is at most the number of 1's at the start. This means f^ρ_σ roughly computes the bounds arising from $|\sigma|$-many ρ-matrix deletions applied to some factor of with complexity $(|\sigma|, |\sigma|)$. This allows us to apply Lemma 4.6, which is designed to compute bounds in exactly this scenario.

Lemma 4.17. *Let $d \geq 1$ be an integer, let $C > 1$ be a real, and let ρ be a growth function satisfying $\rho(x) \geq x$ and $\rho(x) \leq Cx^d$ for all $x \geq 1$.*

Let $m \geq 1$ and $m \geq k \geq 0$ be integers, and let $\sigma \in \{-1, 1\}^m$ satisfy $|\sigma| = k$ and $\operatorname{disc}(\sigma) \geq 0$. *If $(a, b) = f^\rho_\sigma$, then*

$$0 \leq b = 2k - m \leq k \text{ and}$$
$$0 \leq a \leq \tau^\rho_{m-k}(k, k) \leq (2C)^{(m-k)d^{m-k}}(2k)^{d^{m-k}}.$$

Proof. We first observe that since $\operatorname{disc}(\sigma) \geq 0$, $k \geq m/2$ and $m - k \leq k$. This implies $\tau^\rho_{m-k}(k, k)$ is defined and can be bounded using Lemma 4.6.

As in Lemma 4.16, define $\theta = (\theta_1, \ldots, \theta_m) \in \{-1, 1\}^m$ by setting $\theta_i = 1$ for all $1 \leq i \leq k$ and $\theta_i = -1$ for all $k + 1 \leq i \leq m$. Say $(a_\theta, b_\theta) = f^\rho_\theta$. By Lemma 4.16, $a \leq a_\theta$ and $b = b_\theta$.

Since θ begins with k many 1's, $f^\rho_{\theta|_{[k]}} = (k, k)$. Then, since the remaining $m - k$ coordinates of θ are -1's, we have $f^\rho_\theta = (\tau^\rho_{m-k}(k, k), 2k - m)$. We can immediately conclude $b = b_\theta = 2k - m \geq 0$, and, by Lemma 4.6,

$$a \leq a_\theta = \tau^\rho_{m-k}(k, k) \leq (2C)^{(m-k)d^{m-k}}(2k)^{d^{m-k}}.$$

□

We can finally combine our lemmas to bound the complexity of a factor arising at the end of (ρ, σ)-chain.

Corollary 4.18. *Let $d \geq 1$ be an integer, let $C > 1$ be a real, and Let ρ be a growth function satisfying $\rho(x) \geq x$ and $\rho(x) \leq Cx^d$ for all $x \geq 1$.*

Let $m \geq 1$ and $m \geq k \geq 0$ be integers, and let $\sigma \in \{-1, 1\}^m$ satisfy $|\sigma| = k$. Assume

$$\mathcal{B}_0 = (\mathcal{L}_0, \mathcal{Q}_0) \preceq \cdots \preceq \mathcal{B}_m = (\mathcal{L}_m, \mathcal{Q}_m)$$

is a (ρ, σ)-chain on $\mathbb{F}_p^n$. Then the following hold, where (ℓ, q) denotes the complexity of $\mathcal{B}_m$.

$$0 \leq q = 2k - m \leq k \text{ and}$$
$$0 \leq \ell \leq \tau^\rho_{m-k}(k, k) \leq (2C)^{(m-k)d^{m-k}}(2k)^{d^{m-k}}.$$

Proof. By Lemma 4.13(a), $\operatorname{disc}(\sigma) \geq 0$. The conclusion now follows from Lemma 4.13(b) and Lemma 4.17. □

5. Warm up: quadratic arithmetic regularity lemma

In this section, we reprove a known quadratic regularity lemma using an energy increment argument and the local inverse theorem (Corollary 3.7), arriving at a conclusion stated in terms of the local U^3 norm from Definition 3.6 (see Theorem 5.3). The fact that such a theorem holds will be obvious to the reader familiar with arithmetic regularity lemmas and [48], and in light of Theorem 3.20, is also equivalent to a formulation appearing in [67]. We reprove this theorem here because such a proof has not yet appeared explicitly in the literature (to our knowledge), and it is closely related to our proof of Theorem 6.1. We also clarify the relationship between this result and other quadratic arithmetic regularity lemmas in the literature, and discuss the difference in the bounds in Theorem 5.3 versus Theorem 6.1.

For context, we begin by stating a U^3 arithmetic regularity lemma due to Green and Tao. We will use the following notation.

Notation 5.1. Given a partition $\mathcal{P}$ of a group G and a function $f : G \to [-1, 1]$, define $\mathbb{E}(f|\mathcal{P})$ by setting $\mathbb{E}(f|\mathcal{P})(x) = \mathbb{E}_{y\in P} f(y)$, where P is the element of $\mathcal{P}$ containing x.

When $\mathcal{P} = \mathrm{At}(\mathcal{B})$ for some quadratic factor $\mathcal{B}$ on $\mathbb{F}_p^n$, we write $\mathbb{E}(f|\mathcal{B})$ to mean $\mathbb{E}(f|\mathrm{At}(\mathcal{B}))$, where we recall $\mathrm{At}(\mathcal{B})$ denotes the partition of $\mathbb{F}_p^n$ into the atoms of $\mathcal{B}$.

Theorem 5.2 (Proposition 3.9 in [37])**.** *Let ρ, ω be growth functions with $\omega(x) > 0$ for all $x \in \mathbb{R}$, and let $\epsilon \in (0, 1)$. There exists a constant $M = M(p, \rho, \omega, \epsilon)$ such that the following holds. For any function $f : G = \mathbb{F}_p^n \to [-1, 1]$, there are integers $0 \le \ell, q \le M$ and a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ of complexity (ℓ, q) such that*

$$f = f_1 + f_2 + f_3,$$

where $f_1 = \mathbb{E}(f|\mathcal{B})$, $\|f_2\|_2 \le \epsilon$, and $\|f\|_{U^3} < 1/\omega(\ell + q)$.

This result is proved using the global U^3 inverse theorem of [38] (Theorem 3.4) and several energy incrementing arguments. When the rank function ρ is a polynomial and the function ω is exponential (the case of interest to us), one can check the proof of Theorem 5.2 in [37] gives a tower type bound. Bounds of this form were recently shown to be necessary by Gladkova in [28].

In [67], Theorem 5.2 was used by the second author and Wolf to prove the following quadratic arithmetic regularity lemma, phrased in terms of the local U^3 norm of Definition 3.16.

Theorem 5.3. *There is a constant $K > 0$ so that the following holds. Let $\delta \in (0, 1)$ and let ρ be a growth function satisfying $\rho(x) > K(x + \log_p(\delta^{-1}))$. For all $A \subseteq \mathbb{F}_p^n$, there is a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $\rho(\ell + q)$ such that*

(1) $\ell, q \le M(p, \delta, \rho, K)$

(2) for at least $(1 - \delta)$-fraction of the choice of labels $d \in (\mathbb{F}_p^\ell \times \mathbb{F}_p^q)^3 \times (\mathbb{F}_p^q)^3$, we have

$$\|1_A - \alpha_{B(\Sigma(d))}\|_{U^3(d)}^{TW} < \delta,$$

where $\alpha_{B(\Sigma(d))}$ is the density of A on the atom $B(\Sigma(d))$ (see Definitions 3.14 and 3.16).

The proof of Theorem 5.3 takes as its starting point Theorem 5.2 with an exponential choice of ω, and thus obtains tower type bounds when the rank function ρ is a polynomial. In light of Theorem 3.20, Theorem 5.3 is equivalent to the analogous statement in terms of the local U^3 norm from Definition 3.6.

Theorem 5.4. *There is a constant $K > 0$ so that the following holds. Let $\delta \in (0,1)$ and let ρ be a growth function satisfying $\rho(x) > K(x + \log_p(\delta^{-1}))$. For all $A \subseteq \mathbb{F}_p^n$, there is a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ on $\mathbb{F}_p^n$ of complexity (ℓ, q) and rank at least $\rho(\ell + q)$ such that*

(1) $\ell, q \leq M(p, \delta, \rho, K)$
(2) for at least $(1-\delta)$-fraction of the choice of labels $b \in \mathbb{F}_p^\ell \times \mathbb{F}_p^q$, we have

$$\|1_A - \alpha_{B(b)}\|_{U^3(b)}^P < \delta,$$

where $\alpha_{B(b)}$ is the density of A on the atom $B(b)$ (see Definition 3.6).

The goal of this section is to give another proof of Theorem 5.4 using a straightforward energy incrementing argument and Prendiville's local inverse theorem, Corollary 3.7. The bounds we obtain will still be tower. It seems likely these are the correct bounds. Indeed, Theorem 5.4 implies the conclusion of Theorem 5.2 holds for the function 1_A and the factor $\mathcal{B}$, with an exponential choice of ω, and it is likely Theorem 5.2 requires tower type bounds in this case. The reason this seems likely is that it is known to be the case in the setting of functions (rather than sets) by work of Gladkova [28], and one can often (with some work) turn a function example into a set example.

We will use the same index function as [37] to track energy increments.

Definition 5.5. For $A \subseteq G = \mathbb{F}_p^n$ and a partition $\mathcal{P}$ of G, define

$$\mathrm{ind}(A, \mathcal{P}) = \frac{1}{p^n} \sum_{P \in \mathcal{P}} \left(\frac{|A \cap P|}{|P|} \right)^2 |P|.$$

We will frequently be applying the above definition in the case where $\mathcal{P}$ arises as the set of atoms of a quadratic factor. In this case, the following notation will be more convenient.

Notation 5.6. Suppose $A \subseteq \mathbb{F}_p^n$ and $\mathcal{B}$ is a quadratic factor on $\mathbb{F}_p^n$. We write $\mathrm{ind}(A, \mathcal{B})$ to mean $\mathrm{ind}(A, \mathrm{At}(\mathcal{B}))$, where we recall $\mathrm{At}(\mathcal{B})$ is the partition of $\mathbb{F}_p^n$ into the atoms of $\mathcal{B}$.

We will use the following standard fact about this type of index function (see page 21 of [37]). Specifically, it helps us understand how the index function goes up under refinements.

Fact 5.7 (Pythagoras's theorem)**.** *Suppose $A \subseteq \mathbb{F}_p^n$ and $\mathcal{P}' \preceq \mathcal{P}$ are partitions of $\mathbb{F}_p^n$. Then*

$$\mathrm{ind}(A, \mathcal{P}') = \mathrm{ind}(A, \mathcal{P}) + \frac{1}{p^n} \sum_{P \in \mathcal{P}} \sum_{\{P' \in \mathcal{P}' : P' \subseteq P\}} (\alpha_P - \alpha_{P'})^2 |P'|,$$

where α_X denotes the density of A on a set $X \subseteq \mathbb{F}_p^n$.

We now ready to reprove Theorem 5.4.

Proof of Theorem 5.4. Let $K > 0$ be sufficiently large. Fix $\delta \in (0,1)$ and a growth function ρ satisfying $\rho(x) \geq K(x+\log_p(\delta^{-1}))$. Let $C = C(p)$ be as in Corollary 3.7. Assume $n \geq 1$ and $A \subseteq G = \mathbb{F}_p^n$. Let $\alpha = |A|/|G|$, and given a set $X \subseteq G$, let $\alpha_X = |A \cap X|/|X|$.

We inductively construct a sequence of quadratic factors $\mathcal{B}_i = (\mathcal{L}_i, \mathcal{Q}_i)$ of complexity (ℓ_i, q_i) as follows.

Step 0: Let $\mathcal{B}_0 = (\mathcal{L}_0, \mathcal{Q}_0)$ be the trivial factor and let $\sigma(0)$ be the empty string. If $\|1_A - \alpha\|_{U^3}^P < \delta$, then let $\mathcal{B} = \mathcal{B}_0$ and end the proof. Otherwise, go to the next step.

Step $i+1$: Suppose now $i \geq 0$, and assume by induction we have constructed integers $\ell_i, q_i \geq 0$ and a quadratic factor $\mathcal{B}_i = (\mathcal{L}_i, \mathcal{Q}_i)$ on G of complexity (ℓ_i, q_i) and rank at least $\rho(\ell_i + q_i)$ such that $\mathrm{ind}(A, \mathcal{B}_i) \geq iC^{-2}\delta^{2C+2}$ and such that $|\bigcup_{B \in \mathcal{J}_i} B| > \delta|G|$, where

$$\mathcal{J}_i := \{B(b) \in \mathrm{At}(\mathcal{B}_i) : \|1_A - \alpha_{B(b)}\|_{U^3(b)}^P > \delta\}.$$

For clarity, in the definition of $\mathcal{J}_i$, the norms $\|1_A - \alpha_{B(b)}\|_{U^3(b)}^P$ are all computed relative to the factor $\mathcal{B}_i$. By Corollary 3.7, for each $B \in \mathcal{J}_i$ there is a quadratic polynomial $q_B(x) = x^T M_B x + r_B^T x + c_B$ such that

$$\left|\sum_{x \in B} (1_A(x) - \alpha_B) e^{2\pi i q_B(x)/p}\right| > C^{-1}\delta^C |B|. \tag{5.1}$$

Define $\mathcal{Q}_i' = \mathcal{Q}_i \cup \{M_B : B \in \mathcal{J}_i\}$, let $\mathcal{L}_i'$ be a minimal linearly independent set containing $\mathcal{L}_i$ and spanning $\{r_B : B \in \mathcal{J}_i\}$, and set $\mathcal{B}_i' = (\mathcal{L}_i', \mathcal{Q}_i')$. Let (ℓ_i', q_i') be the complexity of $\mathcal{B}_i'$. Clearly

$$\ell_i' = |\mathcal{L}_i'| \leq \ell_i + |\mathcal{J}_i| \leq \ell_i + p^{\ell_i + q_i} \text{ and } q_i' = |\mathcal{Q}_i'| \leq q_i + |\mathcal{J}_i| \leq q_i + p^{\ell_i + q_i}.$$

Fix $B \in \mathrm{At}(\mathcal{B}_i)$. Clearly $q_B(x)$ is constant on all atoms B' of $\mathcal{B}_i'$. Given $B' \in \mathrm{At}(\mathcal{B}_i')$, let $q_B(B')$ denote this constant value. Since $|\bigcup_{B \in \mathcal{J}_i} B| > \delta|G|$ and by (5.1), we have that

$$\begin{aligned}
\delta^{C+1}C^{-1}|G| \leq \sum_{B \in \mathcal{J}_i} \delta^C C^{-1}|B| &\leq \sum_{B \in \mathcal{J}_i} \left|\sum_{x \in B} (1_A(x) - \alpha_B) e^{2\pi i q_B(x)/p}\right| \\
&= \sum_{B \in \mathcal{J}_i} \left|\sum_{B' \in \mathrm{At}(\mathcal{B}_i')} \sum_{x \in B'} (1_A(x) - \alpha_B) 1_B(x) e^{2\pi i q_B(x)/p}\right| \\
&= \sum_{B \in \mathcal{J}_i} \left|\sum_{B' \in \mathrm{At}(\mathcal{B}_i')} e^{2\pi i q_B(B')/p} \sum_{x \in B'} (1_A(x) - \alpha_B) 1_B(x)\right| \\
&= \sum_{B \in \mathcal{J}_i} \left|\sum_{\{B' \in \mathrm{At}(\mathcal{B}_i') : B' \subseteq B\}} e^{2\pi i q_B(B')/p} (\alpha_{B'} - \alpha_B)|B'|\right| \\
&\leq \sum_{B \in \mathcal{J}_i} \sum_{\{B' \in \mathrm{At}(\mathcal{B}_i') : B' \subseteq B\}} |B'||\alpha_{B'} - \alpha_B|, \qquad (5.2)
\end{aligned}$$

where the last inequality is by the triangle inequality. By Fact 5.7, we have

$$\begin{aligned}\mathrm{ind}(A,\mathcal{B}_i') - \mathrm{ind}(A,\mathcal{B}_i) &= \frac{1}{p^n}\sum_{B\in\mathrm{At}(\mathcal{B}_i)}\sum_{\{B'\in\mathrm{At}(\mathcal{B}_i'):B'\subseteq B\}}(\alpha_B-\alpha_{B'})^2|B'| \\ &= \mathbb{E}_{x\in G}|\mathbb{E}(1_A|\mathcal{B}_i')(x)-\mathbb{E}(1_A|\mathcal{B}_i)(x)|^2 \text{(see Notation 5.1).}\end{aligned}$$

By Jensen's inequality, this is at least

$$\begin{aligned}\left(\mathbb{E}_{x\in G}|\mathbb{E}(1_A|\mathcal{B}_i')(x)-\mathbb{E}(1_A|\mathcal{B}_i)(x)|\right)^2 &= \left(\frac{1}{p^n}\sum_{B\in\mathrm{At}(\mathcal{B}_i)}\sum_{\{B'\in\mathrm{At}(\mathcal{B}_i'):B'\subseteq B\}}|\alpha_B-\alpha_{B'}||B'|\right)^2 \\ &\geq \left(\frac{1}{p^n}\sum_{B\in\mathcal{J}_i}\sum_{\{B'\in\mathrm{At}(\mathcal{B}_i'):B'\subseteq B\}}|\alpha_B-\alpha_{B'}||B'|\right)^2 \\ &\geq \left(\delta^{C+1}C^{-1}\right)^2 \\ &= \delta^{2C+2}C^{-2},\end{aligned}$$

where the third inequality is by (5.2). Combining, we have that

$$\mathrm{ind}(A,\mathcal{B}_i') - \mathrm{ind}(A,\mathcal{B}_i) \geq \delta^{2C+2}C^{-2}.$$

Apply Lemma 4.5 to obtain a factor $\mathcal{B}_{i+1} = (\mathcal{L}_{i+1},\mathcal{Q}_{i+1})$ refining $\mathcal{B}_i'$, of complexity (ℓ_{i+1},q_{i+1}) and rank at least $\rho(\ell_{i+1}+q_{i+1})$, for some

$$q_{i+1} \leq q_i' \leq q_i + p^{\ell_i+q_i}$$

and

$$\ell_{i+1} \leq \tau^{\rho}_{q_i'}(\ell_i',q_i') \leq \tau_{q_i+p^{\ell_i+q_i}}(\ell_i+p^{\ell_i+q_i},q_i+p^{\ell_i+q_i}).$$

By the Fact 5.7, the above, and our induction hypothesis, we have

$$\mathrm{ind}(A,\mathcal{B}_{i+1}) \geq \mathrm{ind}(A,\mathcal{B}_i') \geq \mathrm{ind}(A,\mathcal{B}_i) + \delta^{2C+2}C^{-2} \geq (i+1)C^{-2}\delta^{2C+2}.$$

Let

$$\mathcal{J}_{i+1} = \{B(b)\in\mathrm{At}(\mathcal{B}_{i+1}) : \|1_A-\alpha_B\|^P_{U^3(b)} \geq \delta\},$$

where the norms in $\mathcal{J}_{i+1}$ are computed relative to the factor $\mathcal{B}_{i+1}$. If $|\bigcup_{B\in\mathcal{J}_{i+1}}B| \leq \delta|G|$, let $\mathcal{B}' = \mathcal{B}_{i+1}$ and end the proof. Otherwise, go to the next step.

Clearly we may repeat this process at most some $t \leq C^2\delta^{-2C-2}$ many times. □

We now consider the bound generated by the proof above, in the case where ρ is polynomial. Clearly the inductive process described in the proof ends at some stage $t \leq \mathrm{poly}(\delta^{-1})$ because of the increase in the index at each step. We can see from the proof that the sum ℓ_t+q_t is bounded by $\phi^{(t)}(0)$, where $\phi(x) = \tau^{\rho}_{x+p^x}(x+p^x,x+p^x)$. When the rank function ρ is a polynomial, Lemma 4.6 implies

$$\tau^{\rho}_{x+p^x}(x+p^x,x+p^x) \leq \exp_p(\exp_p(\exp_p(O(x)))),$$

in which case $\phi^{(t)}(0)$ is clearly bounded by a tower-type function in a power of δ^{-1}.

Our proof of the "cylinder" version of this result (Theorem 6.1) will differ from the proof of Theorem 5.3 in several crucial ways, with the goal of increasing the complexity at most polynomially at every step (rather than exponentially as occurs above). First, we will allow our partition to use atoms of distinct factors. Then, each time we apply the localized inverse theorem to an atom, we use refine only the factor involved with that specific atom. This will allow us to avoid taking the common refinements at each step, as occurs above. Second, we avoid applying Lemma 2.12 at any stage, instead choosing to interweave individual matrix deletions with applications of the inverse theorem. A detailed analysis of the bounds generated by such a process will show they are similar to a single application of Lemma 2.12.

6. "Cylinder" quadratic arithmetic regularity

In this section we prove our "cylinder" version of the quadratic arithmetic regularity lemma. This result, Theorem 6.1 below, partitions $\mathbb{F}_p^n$ into atoms of (possibly distinct) high rank quadratic factors, so that most of the group is covered by atoms which are uniform with respect to A in the sense of Definition 3.6. We avoid in this section the decorated notation set out in Definition 3.6, instead working with the definition itself. We do this as we will be working with partitions $\mathcal{P}$, one part of which we will frequently denote with a P.

Theorem 6.1. *For all $\delta \in (0,1)$ and all growth functions ρ, there exists a constant $M = M(p,\delta,\rho)$ so that the following holds. For all $A \subseteq \mathbb{F}_p^n = G$, then there is a partition $\mathcal{P}$ of G such that for each $P \in \mathcal{P}$ there is a quadratic factor $\mathcal{B}_P = (\mathcal{L}_P, \mathcal{Q}_P)$ of complexity (ℓ_P, q_P) and rank at least $\rho(\ell_P + q_P)$ such that $\ell_P, q_P \leq M$, such that $P \in \mathrm{At}(\mathcal{B}_P)$, and such that $|\bigcup_{P\in\mathcal{J}} P| \leq \delta|G|$, where*

$$\mathcal{J} = \{P \in \mathcal{P} : \|(1_A - \alpha_P)1_P\|_{U^3} > \delta\|1_P\|_{U^3}, \text{ where } \alpha_P \text{ denotes the density of } A \text{ on } P\}.$$

Moreover, if ρ is polynomial of degree $d > 1$, then for all $P \in \mathcal{P}$,

$$q_P \leq O_p(\delta^{-O_p(1)}) \text{ and } \ell_P \leq \exp_p(\exp_p(\exp_p(O_{\rho,p}(\delta^{-O_p(1)})))).$$

Moreover, if ρ is polynomial of degree 1, then for all $P \in \mathcal{P}$,

$$q_P \leq O_p(\delta^{-O_p(1)}) \text{ and } \ell_P \leq \exp_p(\exp_p(O_{\rho,p}(\delta^{-O_p(1)}))).$$

Before we proceed to the formal proof, we give an informal outline of the strategy, which consists of an algorithm generating at each step a partition of the group.

The process begins by defining $\mathcal{P}_0$ to be the trivial partition of G. At step $i+1$, we will be given a partition $\mathcal{P}_i$ of G in which every $P \in \mathcal{P}_i$ is an atom of a quadratic factor, $\mathcal{B}_P$. If $\mathcal{P}_i$ has all the desired properties, we end the construction. Otherwise, we do one of two types of induction step to generate a new partition $\mathcal{P}_{i+1}$:

Inductive step type 1: If it is the case that all factors $\mathcal{B}_P$ have high rank (as P ranges over the elements of $\mathcal{P}_i$), we can conclude that many atoms fail to be uniform with respect to A, since otherwise the algorithm would have already halted. For each $P \in \mathcal{P}_i$ that is already sufficiently uniform with respect to A, we leave P unchanged. For each $P \in \mathcal{P}_i$ that

fails to be sufficiently uniform with respect to A, we apply the inverse theorem localized to P (see Corollary 3.7), to obtain a quadratic polynomial correlating with $1_A - \alpha_P$ on P. From this quadratic polynomial, we generate a refinement $\mathcal{B}'_P \preceq \mathcal{B}_P$ which adds at most one new linear term and at most one new quadratic term to $\mathcal{B}_P$. We then delete the set P from the partition, and in its place, we add the atoms of $\mathcal{B}'_P$ which union to P. We check that after doing this on all the non-uniform P, the index of the partition has gone up (see Definition 5.5), and then we go the next step.

Inductive step type -1: If it is not the case that all the factors $\mathcal{B}_P$ have high rank (as P ranges over the elements of $\mathcal{P}_i$), we do the following. For each $P \in \mathcal{P}$ such that $\mathcal{B}_P$ has high rank, we leave P unchanged. For each $P \in \mathcal{P}$ such that $\mathcal{B}_P$ has low rank, we preform one ρ-matrix deletion (see Definition 4.1) to obtain a new factor $\mathcal{B}'_P$. We then delete P from the partition, and in its place, we add the atoms of $\mathcal{B}'_P$ which union to P. We then go to the next step.

We keep track of how this process unfolds by associating a binary string $\sigma(P)$ (see Notation 4.8) to each set P in our partitions as follows.

- Suppose at step $i+1$ we perform an inductive step of type 1. If $P \in \mathcal{P}_i$ was left unchanged in $\mathcal{P}_{i+1}$, we leave $\sigma(P)$ unchanged. On the other hand, if $P \in \mathcal{P}_i$ got deleted and replaced with subatoms in $\mathcal{P}_{i+1}$, then for each $P' \in \mathcal{P}_{i+1}$ with $P' \subseteq P$, we define $\sigma(P') = \sigma(P) \wedge 1$. By construction, $\mathcal{B}_{P'}$ is obtained by adding at most one linear and at most one quadratic constraint to $\mathcal{B}_P$ in this case.
- Suppose at step $i+1$ we perform an inductive step of type -1. If $P \in \mathcal{P}_i$ was left unchanged in $\mathcal{P}_{i+1}$, we leave $\sigma(P)$ unchanged. On the other hand, if $P \in \mathcal{P}_i$ got deleted and replaced with subatoms in $\mathcal{P}_{i+1}$, then for each $P' \in \mathcal{P}_{i+1}$ with $P' \subseteq P$, we define $\sigma(P') = \sigma(P) \wedge -1$. By construction, $\mathcal{B}_{P'}$ is a ρ-matrix deletion of $\mathcal{B}_P$ in this case.

At the end of the process, we will obtain a partition $\mathcal{P}$ of G such that each $P \in \mathcal{P}$ is associated to a binary string $\sigma(P)$ and a quadratic factor $\mathcal{B}_P$, such that P is an atom of $\mathcal{B}_P$, and such that $\mathcal{B}_P$ is the final factor in a $(\rho, \sigma(P))$-chain (see Definition 4.11). We will see that each binary string $\sigma(P)$ cannot contain too many 1's, as each new 1 corresponded to an increase in the index. We can then bound the complexity of the factors $\mathcal{B}_P$ using Corollary 4.18. Indeed, Corollary 4.18 was precisely designed to handle (ρ, σ)-chains where σ has boundedly many 1's. This ends our outline of the proof of Theorem 6.1.

Before proceeding to the formal proof, we prove a lemma allowing us to perform "Inductive steps of type 1" quickly.

Lemma 6.2. *Let $C = C(p)$ be as in Corollary 3.7. Let $\delta \in (0,1)$, let $A \subseteq \mathbb{F}_p^n = G$, and let $\mathcal{P}$ be a partition of G such that for each $P \in \mathcal{P}$ there is a quadratic factor $\mathcal{B}_P = (\mathcal{L}_P, \mathcal{Q}_P)$ of complexity at most (ℓ_P, q_P) and rank at least $C(\ell_P + q_P + \log_p(\delta^{-1}))$ such that $P \in \mathrm{At}(\mathcal{B}_P)$. Suppose $|\bigcup_{P \in \mathcal{J}} P| > \delta |G|$ where*

$$\mathcal{J} = \{P \in \mathcal{P} : \|(1_A - \alpha_P)1_P\|_{U^3} > \delta \|1_P\|_{U^3}, \text{ where } \alpha_P \text{ denotes the density of } A \text{ on } P\}.$$

Then there is a partition $\mathcal{P}'$ refining $\mathcal{P}$ such that $\mathrm{ind}(A,\mathcal{P}') \geq \mathrm{ind}(A,\mathcal{P}) + C^2\delta^{2C+2}$*, and such that for each $P' \in \mathcal{P}'$ one of the following hold.*

- $P' \in \mathcal{P} \setminus \mathcal{J}$,
- $P' \subseteq P$ *for some* $P \in \mathcal{J}$*, and morevoer,* P' *is an atom of a factor* $\mathcal{B}'_P = (\mathcal{L}'_P, \mathcal{Q}'_P) \preceq \mathcal{B}_P$ *of complexity* (ℓ'_P, q'_P) *where* $\mathcal{L}_P \subseteq \mathcal{L}'_P$, $\mathcal{Q}_P \subseteq \mathcal{Q}'_P$, $\ell'_P \leq \ell_P + 1$, *and* $q'_P \leq q_P + 1$.

Proof. Let $C = C(p)$ be as in Corollary 3.7. Assume $\delta, A, \mathcal{P}, \mathcal{J}$ are as in the hypotheses. By Corollary 3.7, for each $P \in \mathcal{J}$ there is a quadratic polynomial $q_P(x) = x^T M_P x + r_P^T x + c_P$ such that

$$\left|\sum_{x\in G}(1_A - \alpha_P)1_P e^{2\pi i q(x)/p}\right| \geq C^{-1}\delta^C|P|. \tag{6.1}$$

For each $P \in \mathcal{P}$, define a quadratic factor $\mathcal{B}'_P = (\mathcal{L}'_P, \mathcal{Q}'_P)$, where $\mathcal{L}'_P$ is a minimal set of vectors containing $\mathcal{L}_P$ and spanning r_P, and $\mathcal{Q}'_P = \mathcal{Q}_P \cup \{M_P\}$. Note that for each $P \in \mathcal{P}$, $q_P(x)$ is constant on all atoms B' of $\mathcal{B}'_P$, so given $B' \in \mathrm{At}(\mathcal{B}'_P)$, we can let $q_P(B')$ denote this constant value. We then have, since $|\bigcup_{P\in\mathcal{J}} P| \geq \delta|G|$, that

$$\begin{aligned}
C^{-1}\delta^{C+1}|G| \leq C^{-1}\delta^C \sum_{P\in\mathcal{J}}|P| &= \sum_{P\in\mathcal{J}} C^{-1}\delta^C|P| \\
&\leq \sum_{P\in\mathcal{J}}\left|\sum_{x\in P}(1_A(x)-\alpha_P)e^{2\pi i q_P(x)/p}\right| \\
&= \sum_{P\in\mathcal{J}}\left|\sum_{B'\in\mathrm{At}(\mathcal{B}'_P)}\sum_{x\in B'}(1_A(x)-\alpha_P)1_P(x)e^{2\pi i q_P(x)/p}\right| \\
&= \sum_{P\in\mathcal{J}}\left|\sum_{B'\in\mathrm{At}(\mathcal{B}'_P)} e^{2\pi i q_P(B')/p}\sum_{x\in B'}(1_A(x)-\alpha_P)1_P(x)\right| \\
&= \sum_{P\in\mathcal{J}}\left|\sum_{\{B'\in\mathrm{At}(\mathcal{B}'_P):B'\subseteq P\}} e^{2\pi i q_B(B')/p}(\alpha_{B'}-\alpha_P)|B'|\right| \\
&\leq \sum_{P\in\mathcal{J}}\sum_{\{B'\in\mathrm{At}(\mathcal{B}'_P):B'\subseteq P\}} |B'||\alpha_{B'}-\alpha_P|,
\end{aligned} \tag{6.2}$$

where second inequality is by (6.1), and the last inequality is by the triangle inequality. Now, define $\mathcal{P}'$ as follows:

$$\mathcal{P}' = \{P : P \in \mathcal{P}\setminus\mathcal{J}\} \cup \bigcup_{P\in\mathcal{J}}\{B \in \mathrm{At}(\mathcal{B}'_P) : B \subset P\}.$$

We estimate the difference between $\mathrm{ind}(A,\mathcal{P})$ and $\mathrm{ind}(A,\mathcal{P}')$. By Pythagoras theorem (Fact 5.7), we have the following (see also Notation 5.1)

$$\begin{aligned}\mathrm{ind}(A,\mathcal{P})-\mathrm{ind}(A,\mathcal{P}') &= \frac{1}{p^n}\sum_{P\in\mathcal{P}}\sum_{\{P'\in \mathrm{At}(\mathcal{B}'_P):P'\subseteq P\}}(\alpha_P-\alpha_{P'})^2|P'| \\ &= \mathbb{E}_{x\in G}|\mathbb{E}(1_A|\mathcal{P}')(x)-\mathbb{E}(1_A|\mathcal{P})(x)|^2.\end{aligned}$$

By Jensen's inequality, this is at least

$$\begin{aligned}\left(\mathbb{E}_{x\in G}|\mathbb{E}(1_A|\mathcal{P}')(x)-\mathbb{E}(1_A|\mathcal{P})(x)|\right)^2 &= \frac{1}{p^{2n}}\left(\sum_{P\in\mathcal{P}}\sum_{\{P'\in \mathrm{At}(\mathcal{B}'_P):P'\subseteq P\}}|\alpha_P-\alpha_{P'}||P'|\right)^2 \\ &\geq \frac{1}{p^{2n}}\left(\sum_{P\in\mathcal{J}}\sum_{\{P'\in \mathrm{At}(\mathcal{B}'_P):P'\subseteq P\}}|\alpha_P-\alpha_{P'}||P'|\right)^2 \\ &\geq \frac{1}{p^{2n}}\left(C^{-1}\delta^{C+1}|G|\right)^2 \\ &= C^{-2}\delta^{2C+2},\end{aligned}$$

where the second inequality is by (6.2).

□

We are now ready to prove Theorem 6.1. The reader may wish to review Notation 4.8 and Definition 4.10 before reading the proof.

Proof of Theorem 6.1. Let $C=C(p)$ be as in Corollary 3.7. Fix $\delta\in(0,1)$ and a growth function ρ. After replacing ρ if necessary, we assume without loss of generality that $\rho(x)\geq C(x+\log_p(\delta^{-1}))$ for all $x\geq 1$. Fix a set $A\subseteq G=\mathbb{F}_p^n$. Throughout the proof, for a set $X\subseteq G$, α_X denotes the density of A on the set X.

We describe an inductive process, where at step i we either obtain the desired partition $\mathcal{P}$, or we generate a partition $\mathcal{P}_i$ of our group, and to each $P\in\mathcal{P}_i$, we associate a binary string $\sigma(P)\in\{-1,1\}^{\leq i}$. At each step i in the process, we will ensure that for each $P\in G$, $\mathrm{ind}(A,\mathcal{P}_i)\geq|\sigma(P)|C^2\delta^{-2C-2}$, and that P is an atom of a factor $\mathcal{B}_P$ which arises as the last element of a $(\rho,\sigma(P))$-chain.

Base Case: Let $\mathcal{P}_0=\{G\}$ be the trivial partition, and define $\sigma(G)=<>$ (the empty string). Trivially, we have $\mathrm{ind}(A,\mathcal{P}_0)\geq 0=|\sigma(G)|C^2\delta^{-2C-2}$. Moreover, G is an atom of the trivial factor, which is by definition the last (and first) element of any $(\rho,<>)$-chain.
Induction Step: Suppose $i\geq 0$, and assume by induction we have a partition $\mathcal{P}_i$ of G, and for each $P\in\mathcal{P}_i$ a sequence $\sigma(P)\in\{-1,1\}^{\leq i}$ such that P is an atom of a quadratic factor $\mathcal{B}_P$ with complexity (ℓ_P,q_P), such that $\mathcal{B}_P$ is the last element in a $(\rho,\sigma(P))$-chain on $\mathbb{F}_p^n$, and such that $\mathrm{ind}(A,\mathcal{P}_i)\geq|\sigma(P)|C^2\delta^{-2C-2}$. Define two subsets of $\mathcal{P}_i$ as follows.

$$\begin{aligned}\mathcal{J}_i &= \{P\in\mathcal{P}:\|(1_A-\alpha_P)1_P\|_{U^3}\geq\delta\|1_P\|_{U^3}\}\text{ and} \\ \mathcal{S}_i &= \{P\in\mathcal{P}:\mathrm{rk}(\mathcal{B}_P)<\rho(\ell_P+q_P)\}.\end{aligned}$$

If $|\bigcup_{P\in\mathcal{J}_i} P_i| < \delta|G|$ and $\mathcal{S}_i = \emptyset$, let $\mathcal{P} = \mathcal{P}_i$ and end the construction. Otherwise we proceed in two cases.

Inductive step type -1: Suppose first that $\mathcal{S}_i \neq \emptyset$. Fix $P \in \mathcal{S}_i$. Then there are $\lambda_1, \ldots, \lambda_{q_P}$ not all zero such that

$$\mathrm{rk}(\lambda_1 M_1 + \ldots + \lambda_{q_P} M_{q_p}) < \rho(\ell_P + q_P).$$

Without loss of generality, suppose $\lambda_{q_P} \neq 0$. Set $U := \lambda_1 M_1 + \ldots + \lambda_{q_P} M_{q_p}$, and let $\mathcal{L}'_P$ be a minimal set of vectors containing $\mathcal{L}_P$ and spanning $\ker(U)^\perp$, and let $\mathcal{Q}'_P = \mathcal{Q}_P \setminus \{M_{q_P}\}$. Then $\mathcal{B}'_P := (\mathcal{L}'_P, \mathcal{Q}'_P)$ is a refinement of $\mathcal{B}_P$ of complexity $(\ell_{P'}, q_{P'})$, where $q_{P'} = q_P - 1$ and $\ell_{P'} \leq \ell_P + \rho(\ell_P + q_P)$. By the inductive hypothesis, $\mathcal{B}_P$ is the final factor appearing in a $(\rho, \sigma(P))$-chain. Thus, $\mathcal{B}'_P$ is the final factor appearing in a $(\rho, \sigma(P) \wedge -1)$-chain. For each $P' \in \mathrm{At}(\mathcal{B}'_P)$ satisfying $P' \subseteq P$, set $\sigma(P') = \sigma(P) \wedge -1$, set $\mathcal{B}_{P'} = \mathcal{B}'_P$, and let $(\ell_{P'}, q_{P'})$ be the complexity of $\mathcal{B}_{P'}$. We note that, for such P', $|\sigma(P')| = |\sigma(P)|$, and thus our induction hypotheses imply

$$\mathrm{ind}(A, \mathcal{P}_{i+1}) \geq \mathrm{ind}(A, \mathcal{P}_i) \geq |\sigma(P)| C^2 \delta^{-2C-2} = |\sigma(P')| C^2 \delta^{-2C-2},$$

where the first inequality is because $\mathcal{P}_{i+1}$ refines $\mathcal{P}_i$, and thus has index at least that of $\mathcal{P}_i$ by Fact 5.7. We then define our new partition to be

$$\mathcal{P}_{i+1} = (\mathcal{P}_i \setminus \mathcal{S}_i) \cup \bigcup_{P\in\mathcal{S}_i} \{P' \in \mathrm{At}(\mathcal{B}'_P) : P' \subseteq P\}.$$

By construction, the inductive hypotheses, and the fact $\mathrm{ind}(A, \mathcal{P}_{i+1}) \geq \mathrm{ind}(A, \mathcal{P}_i)$, we can conclude that for all $P \in \mathcal{P}_{i+1}$, P is an atom of a quadratic factor $\mathcal{B}_P$ with complexity (ℓ_P, q_P) that arises as the last element in a $(\rho, \sigma(P))$-chain, and moreover, that $\mathrm{ind}(A, \mathcal{P}_{i+1}) \geq |\sigma(P)| C^2 \delta^{-2C-2}$.

Inductive step type 1: Suppose now $\mathcal{S}_i = \emptyset$. Then we must have $|\bigcup_{P\in\mathcal{J}_i} P| > \delta|G|$, so we can apply Lemma 6.2 to obtain a partition $\mathcal{P}_{i+1} \preceq \mathcal{P}_i$ such that the following hold.

(1) $\mathrm{ind}(A, \mathcal{P}') \geq \mathrm{ind}(A, \mathcal{P}) + C^{-2}\delta^{2C+2}$,
(2) for each $P' \in \mathcal{P}'_{i+1}$, either $P' \in \mathcal{P}_i \setminus \mathcal{J}_i$, or $P' \subseteq P$ for some $P \in \mathcal{J}_i$, and in this case, P' is an atom of a factor $\mathcal{B}_{P'} = (\mathcal{L}_{P'}, \mathcal{Q}_{P'}) \preceq \mathcal{B}_P$ of complexity $(\ell_{P'}, q_{P'})$ satisfying $\mathcal{L}_P \subseteq \mathcal{L}_{P'}$, $\mathcal{Q}_P \subseteq \mathcal{Q}_{P'}$, $\ell_{P'} \leq \ell_P + 1$, and $q_{P'} \leq q_P + 1$.

For each $P \in \mathcal{J}_i$, and $P' \in \mathcal{P}_{i+1}$ with $P' \subseteq P$, define $\sigma(P') = \sigma(P) \wedge 1$. Recall that by the induction hypothesis, $\mathcal{B}_P$ is the final factor in a $(\rho, \sigma(P))$-chain, and thus, by construction, $\mathcal{B}_{P'}$ is the final factor in a $(\rho, \sigma(P'))$-chain. Moreover, by induction, by (1) above, and since $|\sigma(P')| = |\sigma(P)| + 1$, we have

$$\mathrm{ind}(A, \mathcal{P}_{i+1}) \geq \mathrm{ind}(A, \mathcal{P}) + C^{-2}\delta^{2C+2} \geq |\sigma(P)| C^{-2}\delta^{2C+2} + C^{-2}\delta^{2C+2} = |\sigma(P')| C^{-2}\delta^{2C+2}.$$

Thus, by construction, the inductive hypotheses, and the fact $\mathrm{ind}(A, \mathcal{P}_{i+1}) \geq \mathrm{ind}(A, \mathcal{P}_i)$, we can conclude that for all $P \in \mathcal{P}_{i+1}$, P is an atom of a quadratic factor $\mathcal{B}_P$ with complexity (ℓ_P, q_P) that arises as the last element in a $(\rho, \sigma(P))$-chain, and moreover, that $\mathrm{ind}(A, \mathcal{P}_{i+1}) \geq |\sigma(P)| C^2 \delta^{-2C-2}$.

This completes our description of step $i + 1$ of the construction.

We claim this process halts after at some finite number of steps t. First, we observe that at every step i, and for every $P \in \mathcal{P}_i$, we have $\mathrm{ind}(A, \mathcal{P}_i) \geq |\sigma(P)|C^{-2}\delta^{2C+2}$, and thus, $\sigma(P)$ can contain at most $C^2\delta^{-2C-2}$ many 1's. On the other hand, by construction, there also exists a $(\rho, \sigma(P))$-chain on $\mathbb{F}_p^n$. By Lemma 4.13(a), this implies we must have $\mathrm{disc}(\sigma(P)) \geq 0$, so $\sigma(P)$ contains at most $C^2\delta^{-2C-2}$ many -1's (since it has at most that many 1's). Consequently, no $\sigma(P)$ can have length longer than $2C^2\delta^{-2C-2}$. Since each step in the process adds -1 or 1 to some $\sigma(P)$, there can be only finitely many steps total (by Kőnig's lemma).

Thus, at the end of the process, we arrive at a partition $\mathcal{P} := \mathcal{P}_t$, and for each $P \in \mathcal{P}$, a binary string $\sigma(P) \in \{-1, 1\}^{\leq t}$ with the following properties.

(i) $|\bigcup_{P\in\mathcal{J}} P| < \delta|G|$ where $\mathcal{J} = \{P \in \mathcal{P} : \|(1_A - \alpha_P)1_P\|_{U^3} \geq \delta\|1_P\|_{U^3}\}$,
(ii) for each $P \in \mathcal{P}$, P is an atom of a factor $\mathcal{B}_P$ of complexity (ℓ_P, q_P) and rank at least $\rho(\ell_P + q_P)$,
(iii) for each $P \in \mathcal{P}$, $\mathcal{B}_P$ is the final factor in a $(\rho, \sigma(P))$-chain on $\mathbb{F}_p^n$,
(iv) for each $P \in \mathcal{P}$, $\mathrm{ind}(A, \mathcal{P}) \geq |\sigma(P)|C^{-2}\delta^{2C+2}$.

Clearly item (iv) implies that for all $P \in \mathcal{P}$, $|\sigma(P)| \leq C^2\delta^{-2C-C}$. Combining this with (iii), we see that Corollary 4.18 implies

$$0 \leq q_P \leq 2C^2\delta^{-2C-2} \text{ and}$$
$$0 \leq \ell_P \leq (2K)^{2C^2\delta^{-2C-2}d^{2C^2\delta^{-2C-2}}}(4C^2\delta^{-2C-2})^{d^{2C^2\delta^{-2C-2}}}.$$

Clearly the bound for q_P above has the form $O_p(\delta^{-O_p(1)})$. If $d = 1$, the bound for ℓ_P above has the form $\exp_p(\exp_p(O_{p,\rho}(\delta^{-O_p(1)})))$, while if $d > 1$, then the bound for ℓ_P above has the form $\exp_p(\exp_p(\exp_p(O_{p,\rho}(\delta^{-O_p(1)}))))$. This finishes the proof. □

7. Putting it all together

We can now put things together to prove the main theorem. It will be convenient to use the following terminology from [65,67] for an "almost ϵ-homogeneous partition" relative to a distinguished subset of a group.

Definition 7.1 (Almost homogeneous partitions)**.** Let G be a finite group and let $A \subseteq G$. We say a partition $\mathcal{P}$ of G is *almost ϵ-homogeneous with respect to A* if $|\bigcup_{P\in\Sigma} P| \geq (1-\epsilon)|G|$, where

$$\Sigma = \{P \in \mathcal{P} : |A \cap P| \geq (1-\epsilon)|P| \text{ or } |A \cap P| \leq \epsilon|P|\}.$$

We will use the following simple averaging fact (see Fact 4.25 in [65] for a proof).

Fact 7.2. *Suppose $A \subseteq G = \mathbb{F}_p^n$. If $\mathcal{P}$ is a partition of G which is almost ϵ-homogeneous with respect to A and $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then $\mathcal{P}'$ is almost $2\sqrt{\epsilon}$-homogeneous with respect to A.*

For convenience, we state a simplified version of Theorem 6.1 which holds for linear growth functions, as this is the version we will need. We use the letter $\mathcal{R}$ to denote partitions to avoid confusion with the notation from Definition 3.6.

Theorem 7.3. *For all $K > 0$ there exists $M = M(K, p)$ so that the following holds. For all $A \subseteq G = \mathbb{F}_p^n$, there exists a partition $\mathcal{R}$ of G such that for each $R \in \mathcal{R}$ there is a quadratic factor $\mathcal{B}_R = (\mathcal{L}_R, \mathcal{Q}_R)$ of complexity (ℓ_R, q_R) and a label $b_R \in \mathbb{F}_p^{\ell_R} \times \mathbb{F}_p^{q_R}$ such that the following hold.*

(1) $R = B_R(b_R) \in \mathrm{At}(\mathcal{B}_R)$, $\mathcal{B}_R$ has rank at least $K(\ell_R + q_R + \log_p(\delta^{-1}))$, and

$$\ell_R + q_R \leq p^{\delta^{-M}};$$

(2) $|\bigcup_{R \in \mathcal{J}} R| \leq \delta |G|$, where $\mathcal{J}$ is the set of $R \in \mathcal{R}$ such that $\|1_A - \alpha_R\|_{U^3(b_R)}^P \geq \delta$, where α_R denotes the density of A on R, and the norm $\|1_A - \alpha_R\|_{U^3(b_R)}^P$ is computed relative to the factor $\mathcal{B}_R$.

Proof. This is an immediate corollary of Theorem 6.1 in the case of a linear growth function (i.e. $d = 1$) and an inspection of Definition 3.6. □

Finally, we will use the following result, proved in [68], which says that given a set of bounded VC_2-dimension, any high rank quadratic factor which is almost homogeneous with respect to A can be replaced with a (possibly different) factor which is still almost homogeneous with respect to A, which has the same linear component and rank as the original factor, and which has small quadratic complexity.[12]

Theorem 7.4 (Theorem 5.4 of [68])**.** *For all integers $k \geq 0$, there exists constant $K = K(k, p) > 0$ and $\delta_0 = \delta_0(k) > 0$ so that for all $0 < \delta < \delta_0$ and all $0 < \epsilon \leq (\delta/120)^{k+2}$, the following holds for all sufficiently large n.*

Suppose $A \subseteq G = \mathbb{F}_p^n$ has VC_2-dimension at most k, and suppose there exist integers $\ell, q \geq 0$ and a quadratic factor $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ on G of complexity (ℓ, q) and rank at least $K(\ell + q + \log_p(\delta^{-1}))$ which is almost ϵ-homogeneous with respect to A. Then there exist an integer $q' \geq 0$ and a quadratic factor $\mathcal{B}' = (\mathcal{L}, \mathcal{Q}')$ on $\mathbb{F}_p^n$ of complexity (ℓ, q') with $\mathrm{rk}(\mathcal{B}') = \mathrm{rk}(\mathcal{B})$ such that

(1) $q' \leq \log_p(\delta^{-k - o_{k,p}(1)})$;

(2) there exists a union Y satisfying $|A \Delta Y| \leq 16\delta^{1 - o_{k,p}(1)}|G|$, where $o_{k,p}(1)$ tends to 0 as δ tends to 0 at a rate depending on k and p.

We now prove the main theorem.

Proof of Theorem 1.3. Fix an integer $k \geq 1$ and let $\delta_0 = \delta_0(k)$ be from Theorem 7.4. Fix any $0 < \delta < \delta_o(k)$ and set $\epsilon = (\delta/120)^{k+2}$. To ease notation, let $\mu = (\epsilon^2/8)^{k^2 2^{k^2}} 2^{-1}$. Choose $K > 0$ sufficiently large compared to k and p, and let $M = M(K, p)$ be as in Theorem 7.3. Let ρ be a polynomial growth function. After possibly replacing ρ, we may assume without loss of generality that $\rho(x) \geq x$ for all $x \geq 1$. Let $C > 1$ be such that for all $x \geq 1$, $\rho(x) \leq Cx^d$.

[12]The statement below is slightly more detailed than that appearing in [68], but follows directly from its proof there.

Fix $A \subseteq G = \mathbb{F}_p^n$ a subset with VC_2-dimension at most k. Given $X \subseteq G$, we let α_X denote the density of A on the set X, i.e. $\alpha_X = |A \cap X|/|X|$. By Theorem 6.1 there exists a partition $\mathcal{R}$ of G so that the following hold.

(1) for each $R \in \mathcal{R}$, there are integers $\ell_R, q_R \geq 0$, a quadratic factor $\mathcal{B}_R = (\mathcal{L}_R, \mathcal{Q}_R)$ on $\mathbb{F}_p^n$ of complexity (ℓ_R, q_R), and $b_R \in \mathbb{F}_p^{\ell_R} \times \mathbb{F}_p^{q_R}$ such that
   (a) $\mathcal{B}_P$ has rank at least $K(\ell_R + q_R + \log_p(\mu^{-1}))$;
   (b) $\ell_R + q_R \leq p^{\mu^{-M}}$;
   (c) $R = B_R(b_R) \in \mathrm{At}(\mathcal{B}_R)$.

(2) The union $|\bigcup_{R\in\mathcal{J}} R|$ has size at most $\mu|G|$, where $\mathcal{J}$ is the set of $R \in \mathcal{R}$ satisfying $\|1_A - \alpha_R\|^P_{U^3(b_R)} \geq \mu$ (where the norm is computed relative to the factor $\mathcal{B}_R$).

We now make a few observations about this partition. First, we observe that by (a), (c), and Lemma 2.17, we have that for all $R \in \mathcal{R}$,

$$|R| \geq (1-\mu)p^{n-\ell_R-q_R} \geq (1-\mu)p^{n-p^{\mu^{-M}}}.$$

Consequently,

$$|\mathcal{R}| \leq \frac{|G|}{(1-\mu)p^{n-p^{\mu^{-M}}}} = p^{p^{\mu^{-M}}}(1-\mu)^{-1}.$$

Further, we observe that (a), the definition of $\mathcal{J}$, and Corollary 3.27 imply that for all $R \in \mathcal{R} \setminus \mathcal{J}$, $\alpha_R \in [0, \frac{\epsilon^2}{2}) \cup (1 - \frac{\epsilon^2}{2}, 1]$. Thus, combining with (2) above, we have that $\mathcal{R}$ is almost $\epsilon^2/2$-homogeneous with respect to A (recall $\mu < \epsilon^2/2$ by definition).

We now define an initial quadratic factor by simply combining the individual factors generated by the $\mathcal{B}_R$ as R ranges over the elements in $\mathcal{R} \setminus \mathcal{J}$. Specifically, we define $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ where $\mathcal{Q} = \bigcup_{R\in\mathcal{R}\setminus\mathcal{J}} \mathcal{Q}_R$, and $\mathcal{L}$ is a minimal set of vectors spanning $\bigcup_{R\in\mathcal{J}} \mathcal{L}_R$. By construction, $\mathcal{B}$ has complexity (ℓ, q) for some integers $\ell, q \geq 0$ satisfying

$$\ell \leq \sum_{R\in\mathcal{J}} \ell_R \leq |\mathcal{R}|p^{\mu^{-M}} \leq p^{p^{\mu^{-M}}}(1-\mu)^{-1}p^{\mu^{-M}} = (1-\mu)^{-1}p^{p^{\mu^{-M}}+\mu^{-M}},$$

and similarly,

$$q \leq \sum_{R\in\mathcal{J}} q_R \leq |\mathcal{R}|p^{\mu^{-M}} \leq (1-\mu)^{-1}p^{p^{\mu^{-M}}+\mu^{-M}}.$$

It is not difficult to see from this, and our definition of μ, that we can bound both ℓ and q by $\exp_p(\exp_p(O_{k,p}(\delta^{-O_{k,p}(1)}))$.

Apply Lemma 2.12 to $\mathcal{B}$ to obtain $\mathcal{B}' = (\mathcal{L}', \mathcal{Q}') \preceq \mathcal{B}$ of complexity (ℓ', q') and rank at least $\rho(\ell' + q')$ for some integers ℓ', q' satisfying $0 \leq q' \leq q$ and

$$\ell' \leq \tau_q^\rho(\ell, q) \leq 2^{qd^q}C^{qd^q}(\ell+q)^{d^q},$$

where the second inequality is by Lemma 4.6. Note that if $d = 1$, this implies

$$\ell' \leq 2^q C^q(\ell+q) \leq \exp_p(\exp_p(\exp_p(O_{k,p,C}(\delta^{-O_{k,p}(1)})))),$$

while if $d > 1$, this implies

$$\ell' \leq \exp_p(\exp_p(\exp_p(\exp_p(O_{k,p,C,d}(\delta^{-O_{k,p}(1)}))))).$$

Define

$$\Sigma_0 = \{B \in \mathrm{At}(\mathcal{B}') : \alpha_B \leq \epsilon\} \text{ and } \Sigma_1 = \{B \in \mathrm{At}(\mathcal{B}') : \alpha_B \geq (1-\epsilon)\}.$$

Since $\mathrm{At}(\mathcal{B}')$ refines $\mathcal{R}$, Fact 7.2 implies $|\bigcup_{B\in\Sigma_0\cup\Sigma_1} B| \geq (1-\epsilon)|G|$. Thus, we see that $\mathcal{B}'$ is almost ϵ-homogeneous with respect to A. By Theorem 7.4, there exist an integer $q'' \geq 0$ and a quadratic factor $\mathcal{B}'' = (\mathcal{L}', \mathcal{Q}'')$ on $\mathbb{F}_p^n$ of complexity (ℓ', q'') with $\mathrm{rk}(\mathcal{B}'') = \mathrm{rk}(\mathcal{B}')$ such that

(1) $q'' \leq \log_p(\delta^{-k-o_{k,p}(1)})$;
(2) there exists a union Y of atoms of $\mathcal{B}''$ satisfying $|A\Delta Y| \leq 16\delta^{1-o_{k,p}(1)}|G|$.

This finishes the proof. □

8. Appendix

In this appendix, we collect the proofs of some straightforward lemmas used earlier in the paper. We begin by proving Lemma 3.11.

Lemma 8.1. *Suppose $\ell, q \geq 0$ are integers and $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ is a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q), and $B \in \mathrm{At}(\mathcal{B})$. Then for any $(x, h_1, h_2, h_3) \in G^4$, the following are equivalent.*

(1) $(x, h_1, h_2, h_3) \in \Omega_B$,
(2) the following hold: $x \in B$, $h_i \in L(0)$ for all $i \in [3]$, $2\beta_{\mathcal{Q}}(x, h_i) = -\beta_{\mathcal{Q}}(h_i, h_i)$ for all $i \in [3]$, and $\beta_{\mathcal{Q}}(h_i, h_j) = 0$ for all $i \neq j \in [3]$.

Proof. Suppose first $(x, h_1, h_2, h_3) \in \Omega_B$. Since $1_B(x) \neq 0$, $x \in B$. For each $i \in [3]$, since $1_B(x + h_i) \neq 0$, $x + h_i \in B$, which, along with $x \in B$ implies we must have $h_i \in L(0)$. Further,

$$\beta_{\mathcal{Q}}(x, x) = \beta_{\mathcal{B}}(x + h_i, x + h_i) = \beta_{\mathcal{Q}}(x, x) + 2\beta_{\mathcal{Q}}(x, h_i) + \beta_{\mathcal{Q}}(h_i, h_i).$$

This immediately implies $2\beta_{\mathcal{Q}}(x, h_i) = -\beta_{\mathcal{Q}}(h_i, h_i)$. Finally, given $i \neq j \in [3]$, since $1_B(x) \neq 0$ and $1_B(x + h_i + h_j) \neq 0$, we have

$$\begin{aligned}\beta_{\mathcal{Q}}(x, x) &= \beta_{\mathcal{Q}}(x + h_i + h_j, x + h_i + h_j)\\ &= \beta_{\mathcal{Q}}(x, x) + 2\beta_{\mathcal{Q}}(x, h_i) + 2\beta_{\mathcal{Q}}(x, h_j) + 2\beta_{\mathcal{Q}}(h_i, h_j) + \beta_{\mathcal{Q}}(h_i, h_i) + 2\beta_{\mathcal{Q}}(h_j, h_j)\\ &= \beta_{\mathcal{Q}}(x, x) + 2\beta_{\mathcal{Q}}(h_i, h_j),\end{aligned}$$

where the last equality is because we have already shown $2\beta_{\mathcal{Q}}(x, h_i) = -\beta_{\mathcal{Q}}(h_i, h_i)$ and $2\beta_{\mathcal{Q}}(x, h_j) = -\beta_{\mathcal{Q}}(h_j, h_j)$. The displayed equation above then implies $\beta_{\mathcal{Q}}(h_i, h_j) = 0$. This finishes the proof that (1) implies (2).

Conversely, assume $(x, h_1, h_2, h_3) \in G^4$ satisfies (2). Since $x \in B$, $1_B(x) \neq 0$. For each $i \in [3]$, $h_i \in L(0)$ tells us then that

$$\beta_{\mathcal{L}}(x) = \beta_{\mathcal{L}}(x + h_1).$$

Further, since $2\beta_{\mathcal{Q}}(x, h_i) = -\beta_{\mathcal{Q}}(h_i, h_i)$,

$$\beta_{\mathcal{Q}}(x + h_i, x + h_i) = \beta_{\mathcal{Q}}(x, x),$$

and thus, we can conclude $x + h_i \in B$ holds as well. Fix now $i \neq j \in [3]$. Again, since $h_i, h_j \in L(0)$, we know

$$\beta_{\mathcal{L}}(x) = \beta_{\mathcal{L}}(x + h_i + h_j).$$

Using that $2\beta_{\mathcal{Q}}(x, h_i) = -\beta_{\mathcal{Q}}(h_i, h_i)$, $2\beta_{\mathcal{Q}}(x, h_j) = -\beta_{\mathcal{Q}}(h_j, h_j)$, and $\beta_{\mathcal{Q}}(h_i, h_j) = 0$, we have

$$\beta_{\mathcal{Q}}(x + h_i + h_j, x + h_i + h_j) = \beta_{\mathcal{Q}}(x, x),$$

and consequently, we can conclude $1_B(x+h_i+h_j) \neq 0$. Finally, the fact $1_B(x+h_1+h_2+h_3) \neq 0$ follows since $h_1, h_2, h_3 \in L(0)$ implies

$$\beta_{\mathcal{L}}(x) = \beta_{\mathcal{L}}(x + h_1 + h_2 + h_3),$$

and, since $\beta_{\mathcal{Q}}(h_1, h_2) = \beta_{\mathcal{Q}}(h_1, h_3) = \beta_{\mathcal{Q}}(h_2, h_3) = 0$,

$$\begin{aligned}\beta_{\mathcal{Q}}(x + h_1 + h_2 + h_3, x + h_1 + h_2 + h_3) &= \beta_{\mathcal{Q}}(x + h_1 + h_2, x + h_1 + h_2) + 2\beta_{\mathcal{Q}}(x, h_3) + 2\beta_{\mathcal{Q}}(h_3, h_3) \\ &= \beta_{\mathcal{Q}}(x, x),\end{aligned}$$

where the last equality is because we have already shown $\beta_{\mathcal{Q}}(x + h_1 + h_2, x + h_1 + h_2) = \beta_{\mathcal{Q}}(x, x)$, and since by assumption, $2\beta_{\mathcal{Q}}(x, h_3) + 2\beta_{\mathcal{Q}}(h_3, h_3) = 0$. We can now conclude $x + h_1 + h_2 + h_3 \in B$, finishing the proof. □

As a lemma towards Lemma 2.21, we prove the following.

Lemma 8.2. *Let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ a quadratic factor on $G = \mathbb{F}_p^n$ of complexity (ℓ, q) and rank r. Suppose $k \geq 0$ and $S \subseteq G$ is a sst of size k such that $\mathcal{L} \cup \{Mw : M \in \mathcal{Q}, w \in S\}$ is linearly independent. Then*

$$|\{x \in G : \mathcal{L} \cup \{Mw : M \in \mathcal{Q}, w \in S\} \cup \{Mx : M \in \mathcal{Q}\} \text{ is not linearly independent}\}| \leq p^{n+\ell+(k+1)q-r}.$$

Proof. Observe,

$$\begin{aligned}&|\{x \in G : \mathcal{L} \cup \{Mw : M \in \mathcal{Q}, w \in S\} \cup \{Mw : M \in \mathcal{Q}\} \text{ is not linearly independent}\}| \\ &\leq \sum_{v \in \mathrm{Span}(\mathcal{L} \cup \{Mw : M \in \mathcal{Q}, w \in S\}) \setminus \{0\}} \sum_{M' \in \mathcal{Q}} |\{x \in G : M'x = v\}| \\ &\leq \sum_{v \in \mathrm{Span}(\mathcal{L} \cup \{Mw : M \in \mathcal{Q}, w \in S\}) \setminus \{0\}} \sum_{M' \in \mathcal{Q}} p^{n-r} \\ &\leq p^{n+\ell+(k+1)q-r}.\end{aligned}$$

□

We can now prove Lemma 2.21, which we restate here for the convenience of the reader.

Lemma 8.3. *Let $\mathcal{B} = (\mathcal{L}, \mathcal{Q})$ a high rank factor of complexity (ℓ, q) and rank r. Then*

$$|\{(w_1, w_2, w_3, w_4) \in G^4 : \mathcal{L} \cup \{Mw_1, Mw_2, Mw_3, Mw_4 : M \in \mathcal{Q}\} \text{ is not linearly independent}\}|$$

is at most $14p^{4n+\ell+4q-r}$.

Proof. To ease notation, let

$\Sigma_i = \{(w_1, \ldots, w_i) \in G^i : \mathcal{L} \cup \{Mw_1, Mw_2, Mw_3, Mw_4 : M \in \mathcal{Q}\} \text{ is linearly independent}\}$.

Observe

$$\begin{aligned}
&|\{(w_1, w_2, w_3, w_4) \in G^4 : \mathcal{L} \cup \{Mw_1, Mw_2, Mw_3, Mw_4 : M \in \mathcal{Q}\} \text{ is not linearly independent}\}| \\
&\leq 4|G|^3|\{w \in G : \mathcal{L} \cup \{Mw : M \in \mathcal{Q}\} \text{ not linearly independent}\}| \\
&+ 6 \sum_{w \in \Sigma_1} |G|^2|\{x \in G : \mathcal{L} \cup \{Mw, Mx : M \in \mathcal{Q}\} \text{ not linearly independent}\}| \\
&+ 4 \sum_{(w_1, w_2) \in \Sigma_2} |G||\{x \in G : \mathcal{L} \cup \{Mw_1, Mw_2, Mx : M \in \mathcal{Q}\} \text{ not linearly independent}\}| \\
&\leq 4|G|^3 p^{n+\ell+q-r} + 6|G|^2 p^{n+\ell+3q-r} + 4|G| p^{n+\ell+4q-r} \\
&\leq 14 p^{4n+\ell+4q-r},
\end{aligned}$$

where the second to last inequality is by Lemma 8.2. □

References

[1] Nate Ackerman, Cameron Freer, Rehana Patel, *Stable regularity for relational structures*, arXiv:1712.09305 (2017). 2

[2] Noga Alon, Eldar Fischer, and Ilan Newman, *Efficient testing of bipartite graphs for forbidden induced subgraphs*, SIAM Journal on Computing **37** (2007), no. 3, pp. 959–976. 2

[3] Noga Alon, Jacob Fox, and Yufei Zhao, *Efficient arithmetic regularity and removal lemmas for induced bipartite patterns*, Discrete Analysis (2019), no. 3, 14pp. 2, 9

[4] Artem Chernikov, Daniel Palacin, and Kota Takeuchi, *On n-dependence*, Notre Dame Journal of Formal Logic **60** (2019), no. 2, 195–214. 8

[5] Artem Chernikov and Nadja Hempel, *On n-dependent groups and fields II*, Forum of Mathematics, Sigma **E38** (2021), no. 9, 1–51. 8

[6] ———, *On n-dependent groups and fields III. Multilinear forms and invariant connected components*, arXiv:2412.19921 (2024) 8

[7] Artem Chernikov and Henry Towsner, *Hypergraph regularity and higher arity VC-dimension*, arXiv:2010.00726 (2020). 8, 9

[8] ———, *Higher-arity PAC learning, VC dimension and packing lemma*, arXiv:2510.02420 (2025) 9

[9] Artem Chernikov and Sergei Starchenko, *Regularity lemma for distal structures*, J. Eur. Math. Soc. **20** (2018) pp.2437–2466. 2

[10] Artem Chernikov and Sergei Starchenko, *Definable regularity lemmas for NIP hypergraphs*, The Quarterly Journal of Mathematics **72**, no. 4, (2021), pp. 1401–1433. 2, 9

[11] Gabriel Conant, *Quantitative structure of stable sets in arbitrary finite groups*, Proceedings of the American Mathematical Society **149** (2021), no. 9, pp. 4015–4028. 2

[12] Gabriel Conant and Anand Pillay, *Pseudofinite groups and VC-dimension*, Journal of Mathematical Logic **21** (2021), no. 2, 2150009, 23pp.

[13] Gabriel Conant, Anand Pillay, and Caroline Terry, *A group version of stable regularity*, Mathematical Proceedings of the Cambridge Philosophical Society **168**, no. 2 (2020), pp. 405–413. 2

[14] ———, *Structure and regularity for subsets of groups with finite* VC*-dimension*, Journal of the European Mathematical Society **24** (2022), no. 2, pp. 583–621. 2, 9

[15] Gabriel Conant and Caroline Terry, *Stabilizers and NIP arithmetic regularity*, arXiv:2509.04271 (2025). 2, 9

[16] David Conlon and Jacob Fox, *Bounds for graph regularity and removal lemmas*, Geometric and Functional Analysis **22** (2012), pp. 1191–1256. 1, 3
[17] Leonardo Coregliano and Maryanthe Malliaris, *High-arity PAC learning via exchangeability*, arXiv:2402.14294 (2024). 9
[18] ———, *A packing lemma for VCN_k-dimension and learning high-dimensional data*, arXiv:2505.15688 (2025). 9
[19] ——— *Sample completion, structured correlation, and Netflix problems*, arXiv:2509.20404 (2025). 8
[20] J. Fox, M. Gromov, V. Lafforgue, A. Naor, and J. Pach, *Overlap properties of geometric expanders*, J. Reine Angew. Math. (Crelle's Journal) **671** (2012), pp.49–83. 2
[21] Jacob Fox, László Miklós Lovász, *A tight bound for Szemerédi's regularity lemma*, Combinatorica **37** (2017), no. 5, pp.911–951. 1, 2
[22] Jacob Fox, Jànos Pach, and Andrew Suk, *A Polynomial Regularity Lemma for Semialgebraic Hypergraphs and Its Applications in Geometry and Property Testing*, SIAM Journal on Computing **45** (2016) pp.2199–2223. 2
[23] ———, *Erdős-Hajnal Conjecture for Graphs with Bounded VC-dimension*, Discrete and Computational Geometry **61** (2019), no. 4, pp.809–829. 2, 9
[24] Jacob Fox, Jonathan Tidor, and Yufei Zhao, *Induced arithmetic removal: complexity 1 patterns over finite fields*, Israel Journal of Mathematics **248** (2022), pp.1–38. 6
[25] Peter Frankl and Vojtěch Rödl, *Extremal problems on set systems*, Random Structures and Algorithms **20** (2002), no. 2, pp. 131–164. 2
[26] Lior Gishboliner, Asaf Shapira, and Yuval Wigderson, *Is it easy to regularize a hypergraph with easy links?*, arXiv:2506.15582 (2025). 2, 3, 9
[27] Lior Gishboliner, Asaf Shapira, and Yuval Wigderson, *Regularity for hypergraphs with bounded* VC_2*-dimension*, arXiv:2508.09969 (2025). 3, 5, 6, 7, 8
[28] Val Gladkova, *A note on lower bounds for arithmetic regularity partitions*, arXiv:2510.15532 (2025). 3, 5, 37, 38
[29] Val Gladkova, *Arithmetic regularity lemmas and applications*, Apollo – University of Cambridge Repository, 2025. https://doi.org//10.17863/CAM.116416 7
[30] Timothy Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geometric and Functional Analysis **7** (1997), no. 2, pp. 322–337. 1
[31] ———, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geometric and Functional Analysis **8** (1998), no. 3 pp. 529–551. 2, 14
[32] ———, *A new proof of Szemerédi's theorem*, Geometric and Functional Analysis **11** (2001), no. 3 pp. 465–588. 14
[33] ———, *Quasirandomness, counting and regularity for 3-uniform hypergraphs*, Combinatorics, Probability and Computing **15** (2006), no. 1–2 pp. 143–184. 2, 3
[34] ———, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Annals of Mathematics. Second Series **166** (2006), no. 3 pp. 897–946. 2
[35] Timothy Gowers, Ben Green, Freddie Manners, and Terence Tao, *Marton's conjecture in abelian groups with bounded torsion*, arXiv:2404.02244 (2024). 7, 14, 15
[36] Ben Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geometric and Functional Analysis (2005) **15**, no. 2, pp. 340–376. 1, 2
[37] ———, *Montréal notes on quadratic Fourier analysis*, Additive combinatorics (Montréal 2006, ed. Granville et al.), CRM Proceedings **43**, 69–102, AMS, 2007. 8, 9, 12, 13, 30, 37, 38
[38] Ben Green and Terrance Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proceedings of the Edinburgh Mathematical Society **51**, no. 1, (2008), pp. 73–153 4, 14, 37
[39] ———, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An Irregular Mind, Bolyai Soc. Math. Stud., no. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 261–334.
[40] Nadja Hempel, *On n-dependent groups and fields*, Mathematical Logic Quarterly **62** (2016), no. 3, 215–224. 8

[41] Janós Komlós and Miklós Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics, Paul Erdos is eighty, Bolyai Soc. Math. Stud., no. 2, János Bolyai Math. Soc., Budapest, 1996 pp. 295–352,. 1

[42] Laszlo Lovasz and Balazs Szegedy, *Regularity partitions and the topology of graphons*, An Irregular Mind, Bolyai Soc. Math. Stud., no. 21, Janos Bolyai Math. Soc., Budapest, 2010, pp. 415–446. 2

[43] Maryanthe Malliaris and Saharon Shelah, *Regularity lemmas for stable graphs*, Transactions of the American Mathematical Society **366** (2014), no. 3, pp. 1551–1585. 2

[44] Guy Moshkovitz and Asaf Shapira, *A short proof of Gowers' lower bound for the regularity lemma*, Combinatorica **36** (2016), pp. 187–194. 1

[45] ———, *A Tight Bound for Hypergaph Regularity*, Geometric and Functional Analysis **29** (2019), no. 5 pp. 1531–1578. 1, 2, 5

[46] Brendan Nagle and Vojtěch Rödl and Mathias Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Structures and Algorithms **28** (2006), no. 2, pp. 113–179. 2

[47] Brendan Nagle, Annika Poerschke, Vojtěch Rödl, and Matthias Schacht, *Hypergraph regularity and quasi-randomness*, SIAM J Discrete Mathematics (2013), pp. 227–235. 2

[48] Sean Prendiville, *An inverse theorem for the Gowers U^3-norm relative to quadratic level sets*, arXiv:2409.07962 (2024). 6, 7, 8, 9, 12, 14, 15, 37

[49] V. Rödl, B. Nagle, J. Skokan, M. Schacht, and Y. Kohayakawa, *The hypergraph regularity method and its applications*, Proceedings of the National Academy of Sciences **102** (2005), no. 23, pp. 8109–8113. 2

[50] Vojtěch Rödl and Jozef Skokan, *Counting subgraphs in quasi-random 4-uniform hypergraphs*, Random Structures Algorithms **26** (2005), no. 1-2, pp. 160–203. 2

[51] Subrahmanyam Kalyanasundaram and Asaf Shapira, *A Wowzer-type lower bound for the strong regularity lemma*, Proceedings of the London Mathematical Society **106** (2013), no. 3, vol. 106, pp.621–649. 1, 3

[52] Saharon Shelah, *Strongly dependent theories*, Israel Journal of Mathematics **204** (2014), no. 1, pp. 1–83. 8

[53] ———, *Definable groups for dependent and 2-dependent theories*, Sarajevo Journal of Mathematics **25** (2017), no. 13, pp. 3–25.

[54] Olof Sisask, *Convolutions of sets with bounded VC-dimension are uniformly continuous*, Discrete Analysis (2021), no. 1, 25pp. 2, 9

[55] Endre Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** no. 1 (1975), pp. 199–245.

[56] ———, *Regular partitions of graphs*, Proc Colloque Inter CNRS **260**, CNRS, Paris, 1978, 399–401. 1

[57] Terence Tao and Van Vu, *Additive combinatorics*, Cambridge University Press, vol. 105, Cambridge University Press, 2010. 14, 15

[58] Caroline Terry, *VC_l-dimension and the jump to the fastest speed of a hereditary L-property*, Proc. Amer. Math. Soc. **146** (2018), no. 7, pp. 3111–3126. 3, 8

[59] ———, *An improved bound for regular decompositions of 3-uniform hypergraphs of bounded VC_2-dimension*, Model Theory **2** (2023), no. 2, pp. 325–356. 3, 5, 8

[60] ———, *Growth of regular partitions 1: improved bounds for small slicewise VC-dimension*, arXiv:2404.01274 (2024). 2, 3, 9

[61] ———, *Growth of regular partitions 2: weak regularity*, arXiv:2404.01293 (2024). 2, 9

[62] ———, *Growth of regular partitions 3: strong regularity and the vertex partition*, arXiv:2404.02024 (2024). 3, 7, 8

[63] ———, *Growth of regular partitions 4: strong regularity and the pairs partition*, 2404.02030 (2024). 3, 5, 8, 18, 19

[64] Caroline Terry and Julia Wolf, *Quantitative structure of stable sets in finite abelian groups*, Transactions of the American Mathematical Society **373** (2020), no. 6, pp.3885–3903. 2

[65] ———, *Higher-order generalizations of stability and arithmetic regularity*, arXiv:2111.01739 (2025). 4, 5, 8, 9, 19, 46
[66] ———, *Irregular triads in 3-uniform hypergraphs*, To appear, Memoirs of the American Mathematical Society, arXiv:2111.01737 (2021). 8, 9
[67] ———, *The structure of subsets of $\mathbb{F}_p^n$ of bounded VC_2-dimension*, arXiv:2510.12867 (2025). 4, 5, 7, 8, 9, 13, 14, 20, 37, 46
[68] ———, *On the quadratic complexity of subsets of $\mathbb{F}_p^n$ of bounded VC_2-dimension*, arXiv:2510.12767 (2025). 5, 7, 8, 9, 47

Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL, USA
Email address: hshea3@uic.edu
Email address: caterry@uic.edu