Ergod. Th. & Dynam. Sys., (2026), **46**, 1250–1322 © The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

doi:10.1017/etds.2025.10264

Equidistribution for nilsequences along spheres over finite fields

WENBO SUN

Department of Mathematics, Virginia Tech, 225 Stanger Street, Blacksburg, VA 24061, USA
(e-mail: swenbo@vt.edu)

(*Received* 29 *October* 2024 *and accepted in revised form* 14 *November* 2025)

Abstract. In this paper, we prove a quantitative equidistribution theorem for polynomial sequences in a nilmanifold, where the average is taken along spheres instead of cubes. To be more precise, let $\Omega \subseteq \mathbb{Z}^d$ be the preimage of a sphere $\mathbb{F}_p^d$ under the natural embedding from $\mathbb{Z}^d$ to $\mathbb{F}_p^d$. We show that if a rational polynomial sequence $(g(n)\Gamma)_{n\in\Omega}$ is not equidistributed on a nilmanifold G/Γ, then there exists a non-trivial horizontal character η of G/Γ such that $\eta \circ g \bmod \mathbb{Z}$ vanishes on Ω.

Key words: number theory, classical ergodic theory, equidistribution, nilsequences
2020 Mathematics Subject Classification: 11T99, 37A99 (Primary)

Contents

1 Introduction 1250
2 Background material for polynomials 1255
3 Background material for nilmanifolds 1257
4 Background materials for quadratic forms 1264
5 The main equidistribution theorem 1273
6 Algebraic properties for quadratic forms in $\mathbb{F}_p$ 1276
7 Algebraic properties for quadratic forms in $\mathbb{Z}/p$ 1294
8 An example for the Heisenberg group 1303
9 Proof of the main equidistribution theorem 1305
Acknowledgements 1322
References 1322

1. *Introduction*

Nilmanifolds and nilsequences have been studied extensively in the last decade. They play a fundamental role in combinatorial number theory, ergodic theory, additive combinatorics,

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

and higher order Fourier analysis (see for example [3, 4, 7–9]). We begin with recalling some standard terminologies in this area.

Definition 1.1. (Filtered group) Let G be a group. An $\mathbb{N}$-*filtration* on G is a collection $G_{\mathbb{N}} = (G_i)_{i\in\mathbb{N}}$ of subgroups of G indexed by $\mathbb{N}$ such that the following holds:

(i) for all $i, j \in \mathbb{N}$ with $i \leq j$, we have that $G_i \supseteq G_j$;

(ii) for all $i, j \in \mathbb{N}$, we have $[G_i, G_j] \subseteq G_{i+j}$.

For $s \in \mathbb{N}$, we say that G is an ($\mathbb{N}$-*filtered) nilpotent group* of *degree* at most s (or of *degree* $\leq s$) with respect to some $\mathbb{N}$-filtration $(G_i)_{i\in\mathbb{N}}$ if G_i is trivial whenever $i > s$.

Definition 1.2. (Nilmanifold) Let Γ be a discrete and cocompact subgroup of a connected, simply connected nilpotent Lie group G with filtration $G_{\mathbb{N}} = (G_i)_{i\in\mathbb{N}}$ such that $\Gamma_i := \Gamma \cap G_i$ is a cocompact subgroup of G_i for all $i \in \mathbb{N}$. Then, we say that G/Γ is an ($\mathbb{N}$-*filtered) nilmanifold*, and we use $(G/\Gamma)_{\mathbb{N}}$ to denote the collection $(G_i/\Gamma_i)_{i\in\mathbb{N}}$ (which is called the $\mathbb{N}$-*filtration* of G/Γ). We say that G/Γ has degree $\leq s$ with respect to $(G/\Gamma)_{\mathbb{N}}$ if G has degree $\leq s$ with respect to $G_{\mathbb{N}}$.

Definition 1.3. (Polynomial sequences) Let $d, \in \mathbb{N}_+$ and G be a connected simply connected nilpotent Lie group and $(G_i)_{i\in\mathbb{N}}$ be an $\mathbb{N}$-filtration of G. Denote $\Delta_h g(n) := g(n+h)g(n)^{-1}$ for all $n, h \in G$. A map $g\colon \mathbb{Z}^d \to G$ is an ($\mathbb{N}$-*filtered) polynomial sequence if*

$$\Delta_{h_m} \cdots \Delta_{h_1} g(n) \in G_m$$

for all $m \in \mathbb{N}$ and $n, h_1, \ldots, h_m \in \mathbb{Z}^d$. The set of all $\mathbb{N}$-filtered polynomial sequences is denoted by $\mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$.

A fundamental question in higher order Fourier analysis is to study the equidistribution property of polynomial sequences. Let Ω be a non-empty subset of $\mathbb{Z}^d$ and G/Γ be an $\mathbb{N}$-filtered nilmanifold. A sequence $a\colon \Omega \to G/\Gamma$ is δ-*equidistributed* on G/Γ if for all Lipschitz function $F\colon G/\Gamma \to \mathbb{C}$ (with respect to the metric $d_{G/\Gamma}$ to be defined in §3), we have that

$$\limsup_{N\to\infty} \left| \frac{1}{|\Omega \cap [N]^d|} \sum_{n\in\Omega\cap[N]^d} F(a(n)) - \int_{G/\Gamma} F \, dm_{G/\Gamma} \right| \leq \delta \|F\|_{\mathrm{Lip}},$$

where $m_{G/\Gamma}$ is the Haar measure of G/Γ and the *Lipschitz norm* is defined as

$$\|F\|_{\mathrm{Lip}} := \sup_{x\in G/\Gamma} |F(x)| + \sup_{x,y\in G/\Gamma, x\neq y} \frac{|F(x) - F(y)|}{d_{G/\Gamma}(x, y)}.$$

It is known that the equidistribution property of polynomial sequences is connected to the *horizontal characters* of G/Γ, that is, group homomorphisms $\eta\colon G \to \mathbb{R}$ with $\eta(\Gamma) \subseteq \mathbb{Z}$. It was proved by Green and Tao [5, 6] that for any $\delta > 0$ and any polynomial sequence $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$, either $(g(n)\Gamma)_{n\in\mathbb{Z}^d}$ is δ-equidistributed on G/Γ or there exists a non-trivial horizontal character η of G/Γ whose complexity is comparable with the complexity of G/Γ (we will define complexity formally in §3) such that $\eta \circ g$ is a slowly varying function. This is a quantitative generalization of a result of Leibman [12].

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

It is natural to ask about the equidistribution property of polynomial sequences where the average is taken over some algebraic curves $\Omega \subseteq \mathbb{Z}^d$. In this paper, we initiate the study in this direction by showing a quantitative equidistribution property for sequences of the form $(g(n)\Gamma)_{n\in\Omega}$, where Ω is the preimage of a sphere over a finite field $\mathbb{F}_p^d$ under the natural embedding from $\mathbb{Z}^d$ to $\mathbb{F}_p^d$ (see [11, 14] for the study of related averages). The following is the main result of this paper (we refer the readers to §3 for definitions).

THEOREM 1.4. (Equidistribution for polynomial sequence along spheres) *Let* $0 < \delta < 1/2$, $C > 0$, $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ *with* $d \geq s + 13$, *and* $p \gg_{C,d} \delta^{-O_{C,d}(1)}$ *be a prime. Denote* $\Omega = \{n \in \mathbb{Z}^d : n \cdot n \equiv r \bmod p\mathbb{Z}\}$ *for some* $r \in \mathbb{Z}$ *(where* $n \cdot n$ *is the dot product of* n *and itself). Let* G/Γ *be an* s*-step* $\mathbb{N}$*-filtered nilmanifold of complexity at most* C *and* $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ *be a rational polynomial sequence. Then, either* $(g(n)\Gamma)_{n\in\Omega}$ *is* δ*-equidistributed on* G/Γ *or there exists a non-trivial horizontal character* η *of complexity at most* $O_{\delta,d}(1)$ *such that* $\eta \circ g \bmod \mathbb{Z}$ *is a constant on* Ω.

In fact, we will prove a stronger version of Theorem 1.4, which is Theorem 5.1 in this paper.

Remark 1.5. Note that in Theorem 1.4, it is not possible to require $\eta \circ g \bmod \mathbb{Z}$ to be constant on the entire space $\mathbb{Z}^d$. For example, let $G/\Gamma = \mathbb{R}/\mathbb{Z}$, $g(n) = (n \cdot n)/p$ for all $n \in \mathbb{Z}^d$. Since $g(n) \equiv r/p \bmod \mathbb{Z}$ for all $n \in \Omega$, it is not hard to see that $(g(n)\Gamma)_{n\in\Omega}$ is not δ-equidistributed on G/Γ. However, since any non-trivial horizontal character of G/Γ is of the form $\eta(x) := kx$ for some $k \in \mathbb{Z}\backslash\{0\}$, we have that $\eta \circ g(n) \bmod \mathbb{Z} = k(n \cdot n)/p \bmod \mathbb{Z}$ cannot be constant on $\mathbb{Z}$ (but is constant on Ω).

While our proof of Theorem 1.4 follows a framework similar to [5], they have many significant differences. The first is that to run the Green–Tao argument [5] in the spherical setting, we need to use a special property of hyperspheres called the *inheriting principle*, which says that many properties of hyperspheres are preserved under taking intersections (see §4 for a more precise statement). The second difference is that while in [5], one can easily obtain information of the 'leading coefficients' of the polynomial sequence g by applying the Cauchy–Schwarz inequality and induction hypothesis (at least when G is abelian), in our setting, only partial information can be obtained. This reduces the problem to solving some polynomial equations of the form (5.4). While the solution to (5.4) was immediate in the setting of [5], in our case, this is a difficulty question and it takes the bulk of the paper (§§6 and 7) to answer this question. We refer the readers to §5 for details.

It is natural to ask if a quantitative version of Theorem 1.4 holds for general (not necessarily rational) polynomial sequences. Nevertheless, the rational equidistribution theorem already has many applications in combinatorics and number theory. For example, as an immediate consequence of Theorem 1.4, we have the following immediate corollary for exponential sums along spheres, which is of independent interest in number theory.

COROLLARY 1.6. (Weyl's equidistribution theorem along spheres over finite field) *Let* $0 < \delta < 1/2$, $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ *with* $d \geq s + 13$ *and* p *be a prime. Let* $g : \mathbb{Z}^d \to \mathbb{Z}$ *be an*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

integer valued polynomial of degree s. Denote $\Omega = \{n \in \mathbb{F}_p^d : n \cdot n = r\}$ *for some* $r \in \mathbb{F}_p$. *If* $p \gg_d \delta^{-O_d(1)}$, *then either*

$$\left| \frac{1}{|\Omega|} \sum_{n \in \Omega} \exp(g \circ \tau(n)/p) \right| < \delta$$

or $g \circ \tau(n)/p \bmod \mathbb{Z}$ *is a constant on* Ω, *where* $\tau : \mathbb{F}_p \to \{0, \dots, p-1\}$ *is the natural bijective embedding. (By Corollary 7.3, we may further conclude that* $g(n) = (n \cdot n - \tau(r))g_1(n) + pg_2(n)$ *for some integer valued polynomials* g_1 *and* g_2.*)*

It is also worth noting that Theorem 1.4 is also the first step towards the geometric Ramsey conjecture over the finite fields. We refer the readers to [15–18] for details.

Motivated by [1, 5, 6], it is natural to ask whether Theorem 1.4 implies a factorization theorem, that is, whether one can decompose g as the product of a constant ϵ, a polynomial sequence g' being equidistributed on a subnilmanifold of G/Γ, and a polynomial sequence γ having strong periodic properties. An affirmative answer to this question will be provided in the forthcoming paper [18]. For example, the following is a special case of [18, Theorem 1.4] (see §3 for definitions).

THEOREM 1.7. *Let* $d \in \mathbb{N}_+, s \in \mathbb{N}$ *with* $d \gg_s 1$, $C > 0$, $\mathcal{F} : \mathbb{R}_+ \to \mathbb{R}_+$ *be a non-decreasing function with* $\mathcal{F}(n) \geq n$, *and* $p \gg_{C,d,\mathcal{F}} 1$ *be a prime. Denote* $\Omega = \{n \in \mathbb{Z}^d : n \cdot n \equiv r \bmod p\mathbb{Z}\}$ *for some* $r \in \mathbb{Z}$. *Let* G/Γ *be an s-step* $\mathbb{N}$*-filtered nilmanifold of complexity at most C and let* $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ *be a rational polynomial sequence on* Ω. *There exist some* $C \leq C' \leq O_{C,d,\mathcal{F}}(1)$, *a proper subgroup* G' *of* G *which is* C'*-rational relative to* $\mathcal{X}$, *and a factorization*

$$g(n) = \epsilon g'(n)\gamma(n) \quad \text{for all } n \in \mathbb{Z}^d$$

such that $\epsilon \in G$ *is of complexity* $O_{C'}(1)$, $g' \in \mathrm{poly}(\mathbb{Z}^d \to G'_{\mathbb{N}})$ *is rational, and* $(g'(n)\Gamma)_{n \in \Omega}$ *is* $\mathcal{F}(C')^{-1}$*-totally equidistributed on* G'/Γ', *where* $\Gamma' := G' \cap \Gamma$, *and that* $\gamma \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ *is such that* $\gamma(n + rm)^{-1}\gamma(n) \in \Gamma$ *for all* $m, n \in \mathbb{Z}^d$ *with* $n, n + rm \in \Omega$ *for some* $r \leq O_{C,d,s}(C')$.

1.1. *Organization of the paper.* We first provide the background material for polynomials, nilmanifolds, and quadratic forms in §§2, 3, and 4, respectively. Then, in §5, we state the main equidistribution result of the paper (that is, Theorem 5.1, which is a generalization of Theorem 1.4) and explain the outline of its proof. We also explain in §5 the main challenges in adapting the Green–Tao approach to the quadratic setting.

Sections 6, 7, and 9 are devoted to the proof of Theorem 5.1. In §6, we provide the solutions to some relevant algebraic equations for polynomials over $\mathbb{F}_p$. Among other results, we solve a polynomial equation of the form (6.25), which plays a central role in the proof of Theorem 1.4 for p-periodic polynomial sequences. Then, in §7, we extend the results in §6 to polynomials in $\mathbb{Q}$. Finally, in §9, we combine the results from §§6 and 7 with the approach of Green and Tao [5, 6] to complete the proof of Theorem 5.1.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

1.2. *Definitions and notation.*

Convention 1.8. Throughout this paper, we use $\tau\colon \mathbb{F}_p \to \{0, \dots, p-1\}$ to denote the natural bijective embedding, and use ι to denote the map from $\mathbb{Z}$ (or $\mathbb{Z}_K$ for any K divisible by p) to $\mathbb{F}_p$ given by $\iota(n) := \tau^{-1}(n \bmod p\mathbb{Z})$. We also use τ to denote the map from $\mathbb{F}_p^k$ to $\mathbb{Z}^k$ (or $\mathbb{Z}_K^k$) given by $\tau(x_1, \dots, x_k) := (\tau(x_1), \dots, \tau(x_k))$, and ι to denote the map from $\mathbb{Z}^k$ (or $\mathbb{Z}_K^k$) to $\mathbb{F}_p^k$ given by $\iota(x_1, \dots, x_k) := (\iota(x_1), \dots, \iota(x_k))$. When there is no confusion, we will not state the domain and range of τ and ι explicitly.

We may also extend the domain of ι to all the rational numbers of the form x/y with $(x, y) = 1$, $x \in \mathbb{Z}$, $y \in \mathbb{Z}\backslash p\mathbb{Z}$ by setting $\iota(x/y) := \iota(xy^*)$, where y^* is any integer with $yy^* \equiv 1 \bmod p\mathbb{Z}$.

Below is the notation we use in this paper.

- Let $\mathbb{N}, \mathbb{N}_+, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{R}_+, \mathbb{C}$ denote the set of non-negative integers, positive integers, integers, rational numbers, real numbers, positive real numbers, and complex numbers, respectively. Denote $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. Let $\mathbb{F}_p$ denote the finite field with p elements. Let $\mathbb{Z}_K$ denote the cyclic group with K elements, and let $1/K\mathbb{Z}$, $\mathbb{Z}/K$ denote the set of numbers of the from n/K for some $n \in \mathbb{Z}$.
- Throughout this paper, d is a fixed positive integer and p is a prime number.
- Throughout this paper, unless otherwise stated, all vectors are assumed to be horizontal (row) vectors.
- Let $\mathcal{C}$ be a collection of parameters and A, B, $c \in \mathbb{R}$. We write $A \gg_{\mathcal{C}} B$ or $A = O_{\mathcal{C}}(B)$ if $|A| \geq K|B|$ for some $K > 0$ depending only on the parameters in $\mathcal{C}$. In the above definitions, we allow the set $\mathcal{C}$ to be empty. In this case, K will be a universal constant.
- Let $[N]$ denote the set $\{0, \dots, N-1\}$.
- We use $\mathbf{0}$ to denote the zero vector in $\mathbb{F}_p^k$ or $\mathbb{R}^k$, where the underlying space will be clear from the context.
- For $i = (i_1, \dots, i_k) \in \mathbb{Z}^k$, denote $|i| := |i_1| + \cdots + |i_k|$. For $n = (n_1, \dots, n_k) \in \mathbb{Z}^k$ and $i = (i_1, \dots, i_k) \in \mathbb{N}^k$, denote $n^i := n_1^{i_1} \cdots n_k^{i_k}$ and $i! := i_1! \cdots i_k!$. For $n = (n_1, \dots, n_k) \in \mathbb{N}^k$ and $i = (i_1, \dots, i_k) \in \mathbb{N}^k$, denote $\binom{n}{i} := \binom{n_1}{i_1} \cdots \binom{n_k}{i_k}$.
- For any set F, let $F[x_1, \dots, x_k]$ denote the set of all polynomials in the variables $x_1, \dots, x_k$ whose coefficients are from F.
- For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer that is not larger than x and $\lceil x \rceil$ denote the smallest integer that is not smaller than x. Denote $\{x\} := x - \lfloor x \rfloor$.
- Let X be a finite set and $f\colon X \to \mathbb{C}$ be a function. Denote $\mathbb{E}_{x \in X} f(x) := 1/|X| \sum_{x \in X} f(x)$, the average of f on X.
- We say that a set $\Omega \subseteq \mathbb{Z}^k$ is *Q-periodic* if $\Omega = \Omega + Q\mathbb{Z}^k$.
- If $\Omega \subseteq \mathbb{Z}^k$ is a Q-periodic set and $f\colon \mathbb{Z}^k \to \mathbb{C}$ is such that $f(n) = f(n + Qm)$ for all $m, n \in \mathbb{Z}^k$ with $n, n + Qm \in \Omega$, then we denote $\mathbb{E}_{x \in \Omega} f(x) := \mathbb{E}_{x \in \Omega \cap [Q]^k} f(x)$.
- For $F = \mathbb{Z}^k$ or $\mathbb{F}_p^k$ and $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_k) \in F$, let $x \cdot y \in \mathbb{Z}$ or $\mathbb{F}_p$ denote the dot product given by $x \cdot y := x_1 y_1 + \cdots + x_k y_k$.
- Let $\exp\colon \mathbb{R} \to \mathbb{C}$ denote the function $\exp(x) := e^{2\pi i x}$.
- If G is a connected, simply connected Lie group, then we use $\log G$ to denote its Lie algebra. Let $\exp\colon \log G \to G$ be the exponential map and $\log\colon G \to \log G$ be the logarithm map. For $t \in \mathbb{R}$ and $g \in G$, denote $g^t := \exp(t \log g)$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

- If $f: H \to G$ is a function from an abelian group $H = (H, +)$ to some group $(G, \cdot)$, denote $\Delta_h f(n) := f(n+h) \cdot f(n)^{-1}$ for all $n, h \in H$.
- We write affine subspaces of $\mathbb{F}_p^d$ as $V + c$, where V is a subspace of $\mathbb{F}_p^d$ passing through $\mathbf{0}$, and $c \in \mathbb{F}_p^d$.

Let $D, D' \in \mathbb{N}_+$ and $C > 0$. Here are some basic notions of complexities.

- *Real and complex numbers:* a number $r \in \mathbb{R}$ is of *complexity* at most C if $r = a/b$ for some $a, b \in \mathbb{Z}$ with $-C \le a, b \le C$. If $r \notin \mathbb{Q}$, then we say that the complexity of r is infinity. A complex number is of *complexity* at most C if both its real and imaginary parts are of complexity at most C.
- *Vectors and matrices:* a vector or matrix is of *complexity* at most C if all of its entries are of complexity at most C.
- *Subspaces:* a subspace of $\mathbb{R}^D$ is of *complexity* at most C if it is the null space of a matrix of complexity at most C.
- *Linear transformations:* let $L: \mathbb{C}^D \to \mathbb{C}^{D'}$ be a linear transformation. Then, L is associated with a $D \times D'$ matrix A in $\mathbb{C}$. We say that L is of *complexity* at most C if A is of complexity at most C.
- *Lipschitz functions:* a Lipschitz function is of *complexity* at most C if its Lipschitz norm is at most C.

2. *Background material for polynomials*

Definition 2.1. (Polynomials in finite field) Let $\mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ be the collection of all functions $f: \mathbb{F}_p^d \to \mathbb{F}_p$ of the form

$$f(n_1, \dots, n_d) = \sum_{0 \le a_1, \dots, a_d \le p-1, a_i \in \mathbb{N}} C_{a_1,\dots,a_d} n_1^{a_1} \cdots n_d^{a_{d-1}} \tag{2.1}$$

for some $C_{a_1,\dots,a_d} \in \mathbb{F}_p^{d'}$. Let $f \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ be a function which is not constant zero. The *degree of* f (denoted as $\deg(f)$) is the largest $r \in \mathbb{N}$ such that $C_{a_1,\dots,a_d} \neq \mathbf{0}$ for some $a_1 + \cdots + a_d = r$. We say that f is *homogeneous of degree* r if $C_{a_1,\dots,a_d} \neq \mathbf{0}$ implies that $a_1 + \cdots + a_d = r$. We say that f is a *linear transformation* if f is homogeneous of degree 1.

Convention 2.2. For convenience, the *degree* of the constant zero function is allowed to take any integer value. For example, 0 can be regarded as a homogeneous polynomial of degree 10, as a linear transformation, or as a polynomial of degree -1. Here, we also adopt the convention that the only polynomial of negative degree is 0.

Definition 2.3. For any $R' \subseteq R \subseteq \mathbb{R}$ and $\Omega \subseteq \mathbb{Z}^d$, let $\mathrm{poly}(\mathbb{Z}^d \to R)$ be the collection of all polynomials in $\mathbb{Z}^d$ taking values in R, and $\mathrm{poly}(\Omega \to R|R')$ be the set of all $f \in \mathrm{poly}(\mathbb{Z}^d \to R)$ such that $f(n) \in R'$ for all $n \in \Omega$.

The following lemma is straightforward.

LEMMA 2.4. *Let $d \in \mathbb{N}_+$, p be a prime, and $f \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ be an integer valued polynomial of degree smaller than p. There exist integer valued polynomial*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$f_1 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ and integer coefficient polynomial $f_2 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$, both having degrees at most $\deg(f)$ such that $(1/p)f = f_1 + (1/p)f_2$.

Proof. Denote $s := \deg(f) < p$. By the multivariate polynomial interpolation, we may write

$$f(n) = \sum_{i \in \mathbb{N}^d, |i| \le s} \frac{a_i}{Q} n^i$$

for some $a_i \in \mathbb{Z}$ and $Q \in \mathbb{Z}$, $p \nmid Q$ (in fact, one can set $Q = (\deg(f)!\,)^d$). Let $Q^* \in \mathbb{Z}$ be such that $Q^* Q \equiv 1 \bmod p\mathbb{Z}$ and set

$$f_2(n) := \sum_{i \in \mathbb{N}^d, |i| \le s} Q^* a_i n^i.$$

Then, f_2 is an integer coefficient polynomial of degree at most s. Note that for all $n \in \mathbb{Z}^d$,

$$\frac{1}{p}(f(n) - f_2(n)) = \sum_{i \in \mathbb{N}^d, |i| \le s} \frac{a_i}{Q} \frac{1 - Q^* Q}{p} n^i,$$

the left-hand side of which belongs to $\mathbb{Z}/p$ and the right-hand side of which belongs to $\mathbb{Z}/Q$. So, $f_1(n) := (1/p)(f(n) - f_2(n))$ takes values in $(\mathbb{Z}/Q) \cap (\mathbb{Z}/p) = \mathbb{Z}$. □

There is a natural correspondence between polynomials taking values in $\mathbb{F}_p$ and polynomials taking values in $\mathbb{Z}/p$. Let $F \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ and $f \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ be polynomials of degree at most s for some $s < p$. If $F = \iota \circ (pf) \circ \tau$, then we say that F is *induced* by f and f is a *lifting* of F. (Note that since $pf \circ \tau$ takes values in $\mathbb{Z}^{d'}$, the coefficients of $pf \circ \tau$ belong to $\mathbb{Z}^{d'}/s!$, and thus $\iota \circ pf \circ \tau$ is well defined by Convention 1.8.) We say that f is a *regular lifting* of F if, in addition, f has the same degree as F and f has $\{0, (1/p), \ldots, (p-1)/p\}$-coefficients.

We provide some basic properties of lifting for later uses. Their proofs are simple and so we omit them.

LEMMA 2.5. (Basic properties of lifting) *Let $d, d', d'' \in \mathbb{N}_+$ and $p > \max\{d, d', d''\}$ be a prime.*

(i) *Every $f \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ of degree less than p induces $\iota \circ (pf) \circ \tau$, which is a polynomial of degree at most $\deg(f)$. Moreover, if f is homogeneous, then so is $\iota \circ (pf) \circ \tau$.*

(ii) *Every $F \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ of degree less than p admits a regular lifting. Moreover, if F is homogeneous, we may further require the lifting to be homogeneous.*

(iii) *If $f \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ is a lifting of some $F \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ with both f and F having degrees at most $p-1$, then for any $f' \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ of degree less than p, f' is a lifting of F if and only if $f - f'$ is integer valued.*

(iv) *If $f, f' \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ are liftings of $F, F' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$, respectively, with all of f, f', F, and F' having degrees at most $(p-1)/2$, then $f + f'$ is a lifting of $F + F'$ and pff' is a lifting of FF'.*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

(v) *If* $f \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ *is a lifting of some* $F \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ *with both* f *and* F *having degrees at most* $\sqrt{p}$*, and* $f \in \mathrm{poly}(\mathbb{Z}^{d'} \to (\mathbb{Z}/p)^{d''})$ *is a lifting of some* $F \in \mathrm{poly}(\mathbb{F}_p^{d'} \to \mathbb{F}_p^{d''})$ *with both* f *and* F *having degrees less than* $\sqrt{p}$*, then* $f' \circ (pf)$ *is a lifting of* $F' \circ F$.

(vi) *If* $f \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ *is a lifting of some* $F \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ *with both* f *and* F *having degrees less than* p*, then* $\tau \circ F \equiv (pf) \circ \tau \bmod \mathbb{Z}^{d'}$.

3. *Background material for nilmanifolds*

3.1. *Nilmanifolds, filtrations, and polynomial sequences.* We start with some basic definitions on nilmanifolds. Let G be a group and $g, h \in G$. Denote $[g, h] := g^{-1}h^{-1}gh$. For subgroups H, H' of G, let $[H, H']$ denote the group generated by $[h, h']$ for all $h \in H$ and $h' \in H'$.

Definition 3.1. (Filtered group) Let G be a group. An $\mathbb{N}$-*filtration* on G is a collection $G_\mathbb{N} = (G_i)_{i\in\mathbb{N}}$ of subgroups of G indexed by $\mathbb{N}$ with $G_0 = G$ such that the following hold:

(i) for all $i, j \in \mathbb{N}$ with $i \le j$, we have that $G_i \supseteq G_j$;

(ii) for all $i, j \in \mathbb{N}$, we have $[G_i, G_j] \subseteq G_{i+j}$.

For $s \in \mathbb{N}$, we say that G is an ($\mathbb{N}$-*filtered) nilpotent group* of *degree* at most s (or of *degree* $\le s$) with respect to some $\mathbb{N}$-filtration $(G_i)_{i\in\mathbb{N}}$ if G_i is trivial whenever $i > s$.

Example 3.2. The *lower central series* of a group G is the $\mathbb{N}$-filtration $(G^i)_{i\in\mathbb{N}}$ given by $G^0 := G$ and $G^{i+1} = [G^i, G]$ for all $i \in \mathbb{N}$. We say that G is a *nilpotent group* of step (or degree) at most s (with respect to the lower central series) if G^i is trivial for all $i \ge s + 1$.

Definition 3.3. (Nilmanifold) Let Γ be a discrete and cocompact subgroup of a connected, simply connected nilpotent Lie group G with filtration $G_\mathbb{N} = (G_i)_{i\in\mathbb{N}}$ such that $\Gamma_i := \Gamma \cap G_i$ is a cocompact subgroup of G_i for all $i \in \mathbb{N}$. (In some papers, such Γ_i is called a *rational* subgroup of G.) Then, we say that G/Γ is an ($\mathbb{N}$-*filtered) nilmanifold* and we use $(G/\Gamma)_\mathbb{N}$ to denote the collection $(G_i/\Gamma_i)_{i\in\mathbb{N}}$ (which is called the $\mathbb{N}$-*filtration* of G/Γ). We say that G/Γ has degree $\le s$ with respect to $(G/\Gamma)_\mathbb{N}$ if G has degree $\le s$ with respect to $G_\mathbb{N}$. (Unlike the convention in literature, for our convenience, we do not require $G_0 = G$ to be the same as G_1.)

We also need to define some special types of nilmanifolds.

Definition 3.4. (Sub-nilmanifold) Let G/Γ be an $\mathbb{N}$-filtered nilmanifold of degree $\le s$ with filtration $G_\mathbb{N}$ and H be a rational subgroup of G. Then, $H/(H \cap \Gamma)$ is also an $\mathbb{N}$-filtered nilmanifold of degree $\le s$ with the filtration $H_\mathbb{N}$ given by $H_i := G_i \cap H$ for all $i \in \mathbb{N}$ (see [8, Example 6.14]). We say that $H/(H \cap \Gamma)$ is a *sub-nilmanifold* of G/Γ, $H_\mathbb{N}$ (or $(H/(H \cap \Gamma))_\mathbb{N}$) is the filtration *induced by* $G_\mathbb{N}$ (or $(G/\Gamma)_\mathbb{N}$).

Definition 3.5. (Quotient nilmanifold) Let G/Γ be an $\mathbb{N}$-filtered nilmanifold of degree $\le s$ with filtration $G_\mathbb{N}$ and H be a normal subgroup of G. Then, $(G/H)/(\Gamma/(\Gamma \cap H))$ is

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

also an $\mathbb{N}$-filtered nilmanifold of most $\leq s$ with the filtration $(G/H)_{\mathbb{N}}$ given by $(G/H)_i := G_i/(H \cap G_i)$ for all $i \in \mathbb{N}$. We say that $(G/H)/(\Gamma/(\Gamma \cap H))$ is the *quotient nilmanifold* of G/Γ by H and that $(G/H)_{\mathbb{N}}$ is the filtration *induced by* $G_{\mathbb{N}}$.

Definition 3.6. (Product nilmanifold) Let G/Γ and G'/Γ' be $\mathbb{N}$-filtered nilmanifolds of degree $\leq s$ with filtration $G_{\mathbb{N}}$ and $G'_{\mathbb{N}}$. Then, $(G \times G')/(\Gamma \times \Gamma')$ is also an $\mathbb{N}$-filtered nilmanifold of most $\subseteq J$ with the filtration $((G \times G')/(\Gamma \times \Gamma'))_{\mathbb{N}}$ given by $((G \times G')/(\Gamma \times \Gamma'))_i := (G_i \times G'_i)/(\Gamma_i \times \Gamma'_i)$ for all $i \in \mathbb{N}$. We say that $(G \times G')/(\Gamma \times \Gamma')$ is the *product nilmanifold* of G/Γ and G'/Γ', and that $((G \times G')/(\Gamma \times \Gamma'))_{\mathbb{N}}$ is the filtration *induced by* $G_{\mathbb{N}}$ and $G'_{\mathbb{N}}$.

We remark that although we used the same terminology 'induce' in Definitions 3.4, 3.5, and 3.6, this will not cause any confusion in the paper as the meaning of 'induce' will be clear from the context.

Definition 3.7. (Filtered homomorphism) An ($\mathbb{N}$-*filtered) homomorphism* $\phi\colon G/\Gamma \to G'/\Gamma'$ between two $\mathbb{N}$-filtered nilmanifolds is a group homomorphism $\phi\colon G \to G'$ which maps Γ to Γ' and maps G_i to G'_i for all $i \in \mathbb{N}$.

Every nilmanifold has an explicit algebraic description by using the Mal'cev basis.

Definition 3.8. (Mal'cev basis) Let $s \in \mathbb{N}_+$, G/Γ be a nilmanifold of step at most s with the $\mathbb{N}$-filtration $(G_i)_{i\in\mathbb{N}}$. Let $\dim(G) = m$ and $\dim(G_i) = m_i$ for all $0 \leq i \leq s$. A basis $\mathcal{X} := \{X_1, \ldots, X_m\}$ for the Lie algebra $\log G$ of G (over $\mathbb{R}$) is a *Mal'cev basis* for G/Γ adapted to the filtration $G_{\mathbb{N}}$ if:

- for all $0 \leq j \leq m-1$, $\log H_j := \mathrm{Span}_{\mathbb{R}}\{\xi_{j+1}, \ldots, \xi_m\}$ is a Lie algebra ideal of $\log G$ and so $H_j := \exp(\log H_j)$ is a normal Lie subgroup of G;
- $G_i = H_{m-m_i}$ for all $0 \leq i \leq s$;
- the map $\psi^{-1}\colon \mathbb{R}^m \to G$ given by

$$\psi^{-1}(t_1, \ldots, t_m) = \exp(t_1 X_1) \cdots \exp(t_m X_m)$$

  is a bijection;
- $\Gamma = \psi^{-1}(\mathbb{Z}^m)$.

We call ψ the *Mal'cev coordinate map* with respect to the Mal'cev basis $\mathcal{X}$. If $g = \psi^{-1}(t_1, \ldots, t_m)$, we say that $(t_1, \ldots, t_m)$ are the *Mal'cev coordinates* of g with respect to $\mathcal{X}$.

We say that the Mal'cev basis $\mathcal{X}$ is *C-rational* (or of *complexity* at most C) if all the structure constants $c_{i,j,k}$ in the relations

$$[X_i, X_j] = \sum_k c_{i,j,k} X_k$$

are rational with complexity at most C.

By [5, Lemma A.14], for any $h \in G$, there is a unique way to write h as $h = \{h\}[h]$ such that $\psi(\{h\}) \in [0, 1)^m$ and $[h] \in \Gamma$. We adopt this notation throughout this paper.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

It is known that for every filtration $G_\bullet$ which is rational for Γ, there exists a Mal'cev basis adapted to it. See for example the discussion after [5, Definition 2.1].

We use the following quantities to describe the complexities of the objects defined above.

Definition 3.9. (Notions of complexities for nilmanifolds) Let G/Γ be a nilmanifold with an $\mathbb{N}$-filtration $G_\mathbb{N}$ and a Mal'cev basis $\mathcal{X} = \{X_1, \dots, X_D\}$ adapted to it. We say that G/Γ is of *complexity* at most C if the Mal'cev basis $\mathcal{X}$ is C-rational and $\dim(G) \le C$.

An element $g \in G$ is of *complexity* at most C (with respect to the Mal'cev coordinate map $\psi\colon G/\Gamma \to \mathbb{R}^m$) if $\psi(g) \in [-C, C]^m$.

Let G'/Γ' be a nilmanifold endowed with the Mal'cev basis $\mathcal{X}' = \{X'_1, \dots, X'_{D'}\}$, respectively. Let $\phi\colon G/\Gamma \to G'/\Gamma'$ be a filtered homomorphism. We say that ϕ is of *complexity* at most C if the map $X_i \to \sum_j a_{i,j} X'_j$ induced by ϕ is such that all $a_{i,j}$ are of complexity at most C.

Let $G' \subseteq G$ be a closed connected subgroup. We say that G' is *C-rational* (or of *complexity* at most C) relative to $\mathcal{X}$ if the Lie algebra $\log G$ has a basis consisting of linear combinations $\sum_i a_i X_i$ such that a_i are rational numbers of complexity at most C.

Convention 3.10. In the rest of the paper, all nilmanifolds are assumed to have a fixed filtration, Mal'cev basis, and a smooth Riemannian metric induced by the Mal'cev basis. Therefore, we will simply say that a nilmanifold, Lipschitz function, sub-nilmanifold, etc. are of complexity C without mentioning the reference filtration and Mal'cev basis.

Definition 3.11. (Polynomial sequences) Let $k \in \mathbb{N}_+$ and G be a connected simply connected nilpotent Lie group. Let $(G_i)_{i\in\mathbb{N}}$ be an $\mathbb{N}$-filtration of G. A map $g\colon \mathbb{Z}^k \to G$ is an *($\mathbb{N}$-filtered) polynomial sequence* if

$$\Delta_{h_m} \cdots \Delta_{h_1} g(n) \in G_m$$

for all $m \in \mathbb{N}$ and $n, h_1, \dots, h_m \in \mathbb{Z}^k$. (Recall that $\Delta_h g(n) := g(n+h)g(n)^{-1}$ for all $n, h \in H$.)

The set of all $\mathbb{N}$-filtered polynomial sequences is denoted by $\mathrm{poly}(\mathbb{Z}^k \to G_\mathbb{N})$.

By [8, Corollary B.4], $\mathrm{poly}(\mathbb{Z}^k \to G_\mathbb{N})$ is a group with respect to the pointwise multiplicative operation. We refer the readers to [8, Appendix B] for more properties for polynomial sequences.

3.2. *Different notions of polynomial sequences.* We start with some conventional definitions for polynomial sequences.

Definition 3.12. (Null, rational, and periodic polynomial sequences) Let $k \in \mathbb{N}_+$, G/Γ be a nilmanifold, $n_* \in \mathbb{Z}^k$, and $g \in \mathrm{poly}(\mathbb{Z}^k \to G_\mathbb{N})$. We say that g is:

- *null* if $g(n) \in \Gamma$ for all $n \in \mathbb{Z}^k$;
- *rational with base point n_** if $g(Qn + n_*) \in \Gamma$ for all $n \in \mathbb{Z}^k$ for some $Q \in \mathbb{N}_+$, in which case, we also say that g is *Q-rational with base point n_**. We simply say that g is *rational* or *Q-rational* when $n_* = \mathbf{0}$;

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

- *periodic* if $g(n+Qm)^{-1}g(n) \in \Gamma$ for all $m, n \in \mathbb{Z}^k$ for some $Q \in \mathbb{N}_+$, in which case, we also say that g is *Q-periodic*.

We refer the readers to [5, Appendix A] for properties of rational/periodic polynomial sequences.

In this paper, since we will work with polynomial sequences which are null, rational, or periodic on certain subsets of $\mathbb{Z}^k$, we introduce the following notation.

Definition 3.13. (Null, rational, and periodic polynomial sequences on subsets of $\mathbb{Z}^k$) Let $k, Q \in \mathbb{N}_+$, G/Γ be an $\mathbb{N}$-filtered nilmanifold, Ω be a subset of $\mathbb{Z}^k$, and $n_* \in \Omega$.

- We use $\mathrm{poly}(\Omega \to G_{\mathbb{N}}|\Gamma)$ to denote the set of all $g \in \mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$ that is *null on* Ω, meaning that $g(n) \in \Gamma$ for all $n \in \Omega$.
- We use $\mathrm{poly}_{\approx Q,n_*}(\Omega \to G_{\mathbb{N}}|\Gamma)$ to denote the set of all $g \in \mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$ that is *Q-rational on Ω with base point n_** meaning that $g(n_* + Qm) \in \Gamma$ for all $m \in \mathbb{Z}^k$ with $n_* + Qm \in \Omega$.
- We use $\mathrm{poly}_Q(\Omega \to G_{\mathbb{N}}|\Gamma)$ to denote the set of all $g \in \mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$ that is *Q-periodic on Ω*, meaning that $g(n+Qm)^{-1}g(n) \in \Gamma$ for all $m, n \in \mathbb{Z}^k$ with $n, n+Qm \in \Omega$.

We informally say that a polynomial sequence is *partially null* if it is null on a subset of (instead of the entire) $\mathbb{Z}^k$. We adopt similar conventions to rational and periodic sequences.

Example 3.14. The set $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}}|\Gamma)$ is simply the set of all Γ-valued (that is, null) polynomials. In other words, $\mathrm{poly}(\mathbb{Z}^k \to \Gamma_{\mathbb{N}}) = \mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}}|\Gamma)$. The set $\mathrm{poly}_{\approx Q,\mathbf{0}}(\mathbb{Z}^k \to G_{\mathbb{N}}|\Gamma)$ is the set of Q-rational polynomial sequences in $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$. The set $\mathrm{poly}_Q(\mathbb{Z}^k \to G_{\mathbb{N}}|\Gamma)$ is the set of Q-periodic polynomial sequences in $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$.

Example 3.15. Let $(\mathbb{R}/\mathbb{Z})_{\mathbb{N}}$ be the nilmanifold with the filtration given by $G_0 = \cdots = G_s = \mathbb{R}$ and $G_i = \{\mathrm{id}_G\}$ for $i > s$. Then, $\mathrm{poly}(\Omega \to \mathbb{R}_{\mathbb{N}}|\mathbb{Z})$ is the collection of all polynomials in $\mathrm{poly}(\Omega \to \mathbb{R}|\mathbb{Z})$ (see in Definition 2.3) of degree at most s.

Example 3.16. By [8, Corollary B.4], it is not hard to see that $\mathrm{poly}(\Omega \to G_{\mathbb{N}}|\Gamma)$ and $\mathrm{poly}_{\approx Q,n_*}(\mathbb{Z}^k \to G_{\mathbb{N}}|\Gamma)$ are groups. However, we caution the readers that $\mathrm{poly}_Q(\mathbb{Z}^k \to G_{\mathbb{N}}|\Gamma)$ is not necessarily a group. Let $k = 1$ and $G = (\mathbb{R}^3, *)$ be the Heisenberg group with the group action given by

$$(x, y, z) * (x', y', z') := (x + x', y + y', z + z' + xy')$$

with the lower central series as its $\mathbb{N}$-filtration, meaning that $G_0 = G_1 = G$, $G_2 = \{0\} \times \{0\} \times \mathbb{R}$, and $G_s = \{\mathbf{0}\}$ for $s \geq 3$. Indeed, let $d = k = 1$, $\Gamma = \mathbb{Z}^3$, $g_1, g_2 \in \Gamma$, and $f_i(n) = g_i^{n/Q}$ for all $n \in \mathbb{Z}$ and $i = 1, 2$. Then, $f_1, f_2 \in \mathrm{poly}_Q(\mathbb{Z} \to G_{\mathbb{N}}|\Gamma)$. However,

$$\begin{aligned} &f_2(n+Q)^{-1} f_1(n+Q)^{-1} f_1(n) f_2(n) \\ &\quad = g_2^{-1-(n/Q)} g_1^{-1} g_2^{n/Q} = g_2^{-1}[g_2^{n/Q}, g_1] g_1^{-1} = [g_2, g_1]^{n/Q} g_2^{-1} g_1^{-1}. \end{aligned}$$

In general, $[g_2, g_1]^{n/Q}$ does not always belong to Γ if we choose g_1 and g_2 properly. This means that $f_1 f_2$ does not need to belong to $\mathrm{poly}_p(\mathbb{Z} \to G_{\mathbb{N}}|\Gamma)$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

The next proposition shows that partially Q-rational sequences are also partially $O_{k,s}(Q^{O_{k,s}(1)})$-periodic, the later characterization of which is typically more useful in applications.

PROPOSITION 3.17. *Let $k, Q, s \in \mathbb{N}_+$, $p \gg_{k,Q,s} 1$ be a prime, $\Omega \subseteq \mathbb{Z}^k$ be a non-empty p-periodic set, $n_* \in \Omega$, G/Γ be a nilmanifold of step at most s and complexity at most Q, and $g \in \mathrm{poly}_{\approx Q,n_*}(\Omega \to G_{\mathbb{N}}|\Gamma)$. Then, there exists $Q' \in \mathbb{N}_+$ with $Q' \ll_{k,s} Q^{O_{k,s}(1)}$ such that g belongs to $\mathrm{poly}_{Q'}(\Omega \to G_{\mathbb{N}}|\Gamma)$.*

To prove Proposition 3.17, we need the following result.

PROPOSITION 3.18. *Let $k, s \in \mathbb{N}_+$, $\delta > 0$, and $P \in \mathrm{poly}(\mathbb{Z}^k \to \mathbb{Q})$ be a rational polynomial of degree at most s. Suppose that P is K-periodic for some $K \in \mathbb{N}_+$ and suppose that there exists a subset W of $[K]^k$ of cardinality at least δK^k such that $P(W) \subseteq \mathbb{Z}$. Then, the coefficients of P belong to $\mathbb{Z}/r$ for some $r \in \mathbb{N}_+$ with $r \ll_{k,s} \delta^{-O_{k,s}(1)}$.*

Proof. Let $F\colon \mathbb{T} \to [0,1]$ be a Lipschitz function supported on $(-\delta/10, \delta/10)$ with $F(0) = 1$ and with Lipschitz norm at most $100\delta^{-1}$. Then, $\mathbb{E}_{n\in\mathbb{Z}^k} F(P(n) \bmod \mathbb{Z}) \geq \delta$ and $\int_{\mathbb{T}} F\, dm \leq \delta/2$ (where m is the Lebesgue measure on $\mathbb{T}$). So,

$$\frac{|\mathbb{E}_{n\in\mathbb{Z}^k} F(P(n) \bmod \mathbb{Z}) - \int_{\mathbb{T}} F\, dm|}{\|F\|_{\mathrm{Lip}}} \geq \delta^2/200$$

and thus, the sequence $(P(n) \bmod \mathbb{Z})_{n\in\mathbb{Z}^k}$ is not $\delta^2/200$-equidistributed in $\mathbb{T}$. By [5, Theorem 8.6], there exist $k \in \mathbb{Z}\backslash\{0\}$ with $0 < |k| \ll_{k,s} \delta^{-O_{k,s}(1)}$ and $a \in \mathbb{R}$ such that $kP(n) \equiv a \bmod \mathbb{Z}$ for all $n \in \mathbb{Z}^k$. Since W is non-empty, we may take $a = 0$. So, by interpolation, the coefficients of P belong to $\mathbb{Z}/r$ for some $r \in \mathbb{N}_+$ with $r \ll_{k,s} \delta^{-O_{k,s}(1)}$. □

Proof of Proposition 3.17. Let $\psi = (\psi_1, \dots, \psi_m)\colon G \to \mathbb{R}^m$ be the Mal'cev coordinate map. Then, each $\psi_i(g(n))$ is a rational polynomial of degree at most s. For convenience, denote $f_i := \psi_i \circ g$. Since $p \gg_Q 1$, for any $x \in \Omega$, there exists $v_x, u_x \in \mathbb{Z}^k$ such that $n_* + Qv_x + pu_x = x$. Then, for all $y \in \mathbb{Z}^k$, we have that $n_* + Qv_x + pQy = n_* + Q(v_x + py) \in \Omega$ and thus, $f_i(n_* + Qv_x + pQy) \in \mathbb{Z}$. By interpolation, all the coefficients of the polynomial $f_i(n_* + Qv_x + pQ\cdot)$ belong to $\mathbb{Z}/(s!\,)^k$ and thus, all the coefficients of the polynomial $f_i(n_* + Qv_x + p\cdot)$ belong to $\mathbb{Z}/Q^s(s!\,)^k$. So, we have that

$$f_i(x) \in \mathbb{Z}/Q^s(s!\,)^k \quad \text{for all } x \in \Omega. \tag{3.1}$$

For $h = (h_1, \dots, h_m)$, $h' = (h'_1, \dots, h'_m) \in \mathbb{R}^m$, and $1 \leq i \leq m$, by [5, Lemma A.3], we may write $\psi_i(\psi^{-1}(h)\psi^{-1}(h'))$ as a polynomial of degree $O_{k,s}(1)$ with coefficients in $\mathbb{Z}/q_1$ for some $q_1 \ll_{k,s} Q^{O_{k,s}(1)}$, and this polynomial depends only on the first i variables of h and h'. Since $\psi_i(\psi^{-1}(h)\psi^{-1}(h)) = \mathbf{0}$, we may write $\psi_i(\psi^{-1}(h)\psi^{-1}(h')) = \psi_i(\psi^{-1}(h)\psi^{-1}(h')) - \psi_i(\psi^{-1}(h)\psi^{-1}(h))$ as $\sum_{i'=1}^{i}(h_{i'} - h'_{i'})P_{i,i'}(h, h')$ for some polynomials $P_{i,i'}$ of degree $O_{k,s}(1)$ with coefficients in $\mathbb{Z}/q_1$ for some $q_1 \in \mathbb{N}_+$ with $q_1 \ll_{k,s} Q^{O_{k,s}(1)}$. Therefore, we have that

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$$\psi_i(g(n'+qn)^{-1}g(n')) = \sum_{i'=1}^{i}(f_{i'}(n'+qn) - f_{i'}(n'))P_{i,i'}(\mathbf{f}(n'+qn), \mathbf{f}(n')) \tag{3.2}$$

for all $n, n' \in \mathbb{Z}^k$ and $q \in \mathbb{N}_+$, where $\mathbf{f}(n) := (f_1(n), \dots, f_i(n))$.

Since g is Q-rational with base point n_* on Ω, we have that $f_i(n_* + Qn) \in \mathbb{Z}$ whenever $n_* + Qn \in \Omega$. Since Ω is p-periodic, we have that $f_i(n_* + pQn) \in \mathbb{Z}$ for all $n \in \mathbb{Z}^k$ since Ω is p-periodic.

Since f_i is rational, it is also Q'-periodic for some $Q' \in \mathbb{N}$. We may assume without loss of generality that $pQ|Q'$. Note that the set W of $m \in [Q']^k$ for which $f_i(m) \in \mathbb{Z}$ contains vectors of the form $m = n_* + pQn$ for any $n \in \mathbb{Z}^k$ with $n_* + pQn \in [Q']^k$. So, the density of W in $[Q']^k$ is at least $(pQ)^{-k}$. So, by Proposition 3.18, the coefficients of the polynomial $f_i(n_* + pQ\cdot)$ belong to $\mathbb{Z}/rp^{O_{k,s}(1)}$ for some $r \ll_{k,s} Q^{O_{k,s}(1)}$. So, by interpolation, all the coefficients of the polynomial f_i belong to $\mathbb{Z}/Q^s rp^{O_{k,s}(1)}$.

Moreover, for any $x, y \in \mathbb{Z}^k$ and $Q' \in \mathbb{N}_+$, since all the coefficients of the polynomial f_i belong to $\mathbb{Z}/Q^s rp^{O_{k,s}(1)}$, we have that $f_i(x + Q'Q^s ry) - f_i(x) \in Q'\mathbb{Z}/p^{O_{k,s}(1)}$. If, in addition, $x, x + Q'Q^s ry \in \Omega$, then it follows from (3.1) that

$$f_i(x + Q'Q^s ry) - f_i(x) \in (\mathbb{Z}/Q^s(s!)^k) \cap (Q'\mathbb{Z}/p^{O_{k,s}(1)}) = Q'\mathbb{Z},$$

where the last equality follows from the fact that p is a prime with $p \gg_{k,Q,s} 1$. Therefore, it follows from (3.2) that there exists some $Q' \in \mathbb{N}_+$ with $Q' \ll_{k,s} Q^{O_{k,s}(1)}$ such that $\psi_i(g(n' + Q'Q^s rn)^{-1}g(n')) \in \mathbb{Z}$ for all $n, n' \in \mathbb{Z}^k$ with $n', n' + Q'Q^s rn \in \Omega$. This means that g is $Q'Q^s r$-periodic on Ω. This completes the proof. □

Since $\mathbb{Z}^k$ is p-periodic, as a consequence of Proposition 3.17, we recover [5, Lemma A.12].

COROLLARY 3.19. [5, Lemma A.12] *Let $k, s \in \mathbb{N}_+$ and G/Γ be a nilmanifold of step at most s. Then, every rational polynomial sequence in* $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$ *is periodic.*

3.3. *Type-I horizontal torus and character.* Horizontal torus and character are important notions which characterize the equidistribution properties of nilsequences. In the literature, there are at least three different types of horizontal torus and character. In this paper, as well as the series of work [15–17], we refer to them as *type-I*, *type-II*, and *type-III*. In this paper, we use the type-I horizontal torus and character defined next. (We will define the type-II horizontal torus and character in [16], and the type-III horizontal torus and character in [17].)

Definition 3.20. (Type-I horizontal torus and character) Let G/Γ be a nilmanifold endowed with a Mal'cev basis $\mathcal{X}$. The *type-I horizontal torus* of G/Γ is $G/[G, G]\Gamma$. A *type-I horizontal character* is a continuous homomorphism $\eta\colon G \to \mathbb{R}$ such that $\eta(\Gamma) \subseteq \mathbb{Z}$. When written in the coordinates relative to $\mathcal{X}$, we may write $\eta(g) = k \cdot \psi(g)$ for some unique $k = (k_1, \dots, k_m) \in \mathbb{Z}^m$, where $\psi\colon G \to \mathbb{R}^m$ is the coordinate map with respect to the Mal'cev basis $\mathcal{X}$. We call the quantity $\|\eta\| := |k| = |k_1| + \cdots + |k_m|$ the *complexity* of η (with respect to $\mathcal{X}$).

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

It is not hard to see that any type-I horizontal character mod $\mathbb{Z}$ vanishes on $[G, G]\Gamma$ and thus descends to a continuous homomorphism between the type-I horizontal torus $G/[G, G]\Gamma$ and $\mathbb{R}/\mathbb{Z}$. Moreover, $\eta \bmod \mathbb{Z}$ is a well-defined map from G/Γ to $\mathbb{R}/\mathbb{Z}$.

Type-I horizontal torus and character are used to characterize whether a nilsequence is equidistributed on a nilmanifold [5, 12]. We provide some properties for later use. The proof of the following lemma is similar to that of [5, Lemma 6.7], [8, Lemma B.9], and [1, Lemma 2.8]. We omit the details.

LEMMA 3.21. *Let $k \in \mathbb{N}_+$, $s \in \mathbb{N}$, and G be an $\mathbb{N}$-filtered group of degree at most s. A function $g\colon \mathbb{Z}^k \to G$ belongs to* $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$ *if and only if for all $i \in \mathbb{N}^k$ with $|i| \le s$, there exists $X_i \in \log(G_{|i|})$ such that*

$$g(n) = \prod_{i\in\mathbb{N}^k, |i|\le s} \exp\left(\binom{n}{i} X_i\right).$$

Moreover, if $g \in \mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$, then the choice of X_i is unique.

Furthermore, if G' is a subgroup of G and g takes values in G', then $X_i \in \log(G')$ for all $i \in \mathbb{N}^k$.

An equivalent way of saying that a function $g\colon \mathbb{Z}^k \to G$ belongs to $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$ is that

$$g(n) = \prod_{i\in\mathbb{N}^k, |i|\le s} g_i^{\binom{n}{i}} \tag{3.3}$$

for some $g_i \in G_{|i|}$ (with any fixed order in $\mathbb{N}^k$). We call g_i the (*ith) type-I Taylor coefficient* of g, and (3.3) the *type-I Taylor expansion* of g.

3.4. *Vertical torus and character.* The vertical torus is an important concept that allows us to do Fourier decompositions on nilmanifolds.

Definition 3.22. (Vertical torus and character) Let $s \in \mathbb{N}$, G/Γ be a nilmanifold with an $\mathbb{N}$-filtration $(G_i)_{i\in\mathbb{N}}$ of degree $\le s$ with a Mal'cev basis $\mathcal{X}$ adapted to it. Then, G_s lies in the center of G. The *vertical torus* of G/Γ is the set $G_s/(\Gamma \cap G_s)$, which is a connected compact abelian Lip group isomorphic to $\mathbb{T}^{\dim(G_s)}$. A *vertical character* of G/Γ (with respect to the filtration $(G_i)_{i\in\mathbb{N}}$) is a continuous homomorphism $\xi\colon G_s \to \mathbb{R}$ such that $\xi(\Gamma \cap G_s) \subseteq \mathbb{Z}$ (in particular, ξ descends to a continuous homomorphism from the vertical torus $G_s/(\Gamma \cap G_s)$ to $\mathbb{T}$). Let $\psi\colon G \to \mathbb{R}^m$ be the Mal'cev coordinate map and denote $m_s = \dim(G_s)$. Then, there exists $k \in \mathbb{Z}^{m_s}$ such that $\xi(g_s) = (\mathbf{0}, k) \cdot \psi(g_s)$ for all $g_s \in G_s$ (note that the first $\dim(G) - m_s$ coordinates of $\psi(g_s)$ are all zero). We call the quantity $\|\xi\| := |k|$ the *complexity* of ξ (with respect to $\mathcal{X}$).

The following lemma is useful in the study of equidistribution properties on nilmanifolds.

LEMMA 3.23. *Let Ω be a non-empty subset of $\mathbb{Z}^d$ and G/Γ be a nilmanifold with an $\mathbb{N}$-filtration $G_{\mathbb{N}}$ of degree s. Let $m_s = \dim(G_s)$ and $0 < \delta < 1/2$. If a sequence*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$a\colon \Omega \to G/\Gamma$ is not δ-equidistributed on G/Γ, then there exists a Lipschitz function $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{C})$ with a vertical character of complexity at most $O_{m_s}(\delta^{-O_{m_s}(1)})$ such that

$$\limsup_{N\to\infty} \left| \mathbb{E}_{n\in\Omega\cap[N]^d} F(a(n)) - \int_{G/\Gamma} F \, dm_{G/\Gamma} \right| \gg_{m_s} \delta^{O_{m_s}(1)} \|F\|_{\mathrm{Lip}},$$

where $m_{G/\Gamma}$ is the Haar measure of G/Γ.

The proof of Lemma 3.23 is a straightforward extension of [5, Lemma 3.7], we omit the details.

4. *Background materials for quadratic forms*

4.1. *Basic properties for quadratic forms.*

Definition 4.1. (Quadratic forms on $\mathbb{F}_p^d$) We say that a function $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ is a *quadratic form* if

$$M(n) = (nA)\cdot n + n\cdot u + v$$

for some $d\times d$ symmetric matrix A in $\mathbb{F}_p$, some $u \in \mathbb{F}_p^d$, and some $v \in \mathbb{F}_p$. We say that A is the matrix *associated to* M. (Sometimes in the literature, a quadratic form is defined to be a map $B\colon \mathbb{F}_p^d \times \mathbb{F}_p^d \to \mathbb{F}_p^d$ with $B(x,y) = (xA)\cdot y$ for some symmetric matrix A. We define quadratic forms differently for the convenience of this paper.)

We say that M is *pure* if $u = \mathbf{0}$. We say that M is *homogeneous* if $u = \mathbf{0}$ and $v = 0$. We say that M is *non-degenerate* if $\det(A) \neq 0$.

A quadratic form in which we are particularly interested is the one induced by the dot product, namely $M(n) := n\cdot n$, which is associated with the identity matrix.

Throughout this paper, we will crucially use the following philosophy.

Inheriting principle: the intersection of a hypersphere (that is, the set of zeros of a quadratic form) with a subspace or another hypersphere is a hypersphere inheriting properties of the original hypersphere.

One example of this principle is that the intersection inherits the rank of a quadratic form (Proposition 4.8). To make it precise, we need to define quadratic forms on affine subspaces of $\mathbb{F}_p^d$. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form. We use $\mathrm{rank}(M) := \mathrm{rank}(A)$ to denote the *rank* of M. The M being non-degenerate is equivalent of having rank equal to d. Let $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of dimension r. There exists a (not necessarily unique) bijective linear transformation $\phi\colon \mathbb{F}_p^r \to V$. We say that a function $M\colon V + c \to \mathbb{F}_p$ is a *quadratic form* if there exists a quadratic form $M'\colon \mathbb{F}_p^r \to \mathbb{F}_p$ and a bijective linear transformation $\phi\colon \mathbb{F}_p^r \to V$ such that $M(n) = M'(\phi^{-1}(n-c))$ for all $n \in V + c$ (or equivalently, $M'(m) = M(\phi(m) + c)$ for all $m \in \mathbb{F}_p^r$). We define the *rank* $\mathrm{rank}(M|_{V+c})$ of M restricted to $V + c$ as the rank of M'. Note that if $M(n) = M'(\phi^{-1}(n-c)) = M''(\phi'^{-1}(n-c))$ for some quadratic forms M', M'' associated to the matrices A', A'', then $A' = BA''B^T$, where B is the unique $r\times r$ invertible matrix such that $\phi^{-1}\circ\phi(m) = mB$ for all $m \in \mathbb{F}_p^r$. So, $\mathrm{rank}(M|_{V+c})$ is independent of the choice of M' and ϕ'.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

The following is a straightforward property for quadratic forms, which we record for later use.

LEMMA 4.2. *Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A, and $x, y, z \in \mathbb{F}_p^d$. Suppose that $M(x) = M(x+y) = M(x+z) = 0$ and that $p > 2$. Then, $M(x+y+z) = 0$ if and only if $(yA)\cdot z = 0$.*

Proof. The conclusion follows from the equality

$$M(x+y+z) - M(x+y) - M(x+z) + M(x) = 2(yA)\cdot z. \qquad \square$$

The next lemma says that using translations and linear transformations, any quadratic form can be transformed into a 'standard' form.

LEMMA 4.3. *Let $d \in \mathbb{N}_+$, $d' \in \mathbb{N}$, and $p > 2$ be a prime. For any quadratic form $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ of rank d', there exist $c, c', \lambda \in \mathbb{F}_p$ and $v \in \mathbb{F}_p^d$ with $c \neq 0$ and a $d \times d$ invertible matrix R in $\mathbb{F}_p$ such that writing $M'(n) := M(nR + v)$, we have that*

$$M'(n_1, \dots, n_d) = cn_1^2 + n_2^2 + \cdots + n_{d'}^2 + c' n_{d'+1} - \lambda. \tag{4.1}$$

Moreover, if M is homogeneous, then we may further require $c' = \lambda = 0$ in (4.1); if M is pure, then we may further require $v = \mathbf{0}$; if M is non-degenerate, then we may further require that $d' = d$.

Remark 4.4. Throughout this paper, whenever we write a quadratic form M' in the form (4.1), if $d' = d$, then we consider the term $c' n_{d'+1}$ as non-existing.

Proof. Similar to the real valued quadratic forms, by a suitable substitution $n \mapsto nR$ for some upper triangular invertible matrix R and some $v \in \mathbb{F}_p^d$, it suffices to consider the case when the matrix A associated to M is diagonal. Since $\operatorname{rank}(M) = d'$, we may assume without loss of generality that the first d' entry of the diagonal of A is non-zero. Let $c \in \{2, \dots, p-1\}$ be the smallest element such that $c \neq x^2$ for all $x \in \mathbb{F}_p$ (such an element obviously exists). Under a suitable change of variables $(n_1, \dots, n_d) \mapsto (c_1 n_1, \dots, c_{d'} n_{d'}, n_{d'+1}, \dots, n_d)$ for some $c_1, \dots, c_{d'} \in \mathbb{F}_p \backslash \{0\}$ and a change of the orders of $n_1, \dots, n_{d'}$, we may reduce the problem to the case when

$$M(n_1, \dots, n_d) = n_1^2 + \cdots + n_k^2 + cn_{k+1}^2 + \cdots + cn_{d'}^2 + a_1 n_1 + \cdots + a_d n_d - \lambda$$

for some $0 \le k \le d'$, where k is the number of non-zero diagonal entries in A that are quadratic residues. By a suitable substitution $n \mapsto n + m$ for some $m \in \mathbb{F}_p^d$, we may reduce the problem to the case when

$$M(n_1, \dots, n_d) = n_1^2 + \cdots + n_k^2 + cn_{k+1}^2 + \cdots + cn_{d'}^2 + a_{d'+1} n_{d'+1} + \cdots + a_d n_d - \lambda.$$

By another change of variables, we may reduce the problem to the case when

$$M(n_1, \dots, n_d) = n_1^2 + \cdots + n_k^2 + cn_{k+1}^2 + \cdots + cn_{d'}^2 + c' n_{d'+1} - \lambda.$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

By the minimality of c, we have that $a^2+1 \equiv c \bmod p\mathbb{Z}$ for some $1 \le a \le p-1$. If $k \le d-2$, then

$$cn_{k+1}^2 + cn_{k+2}^2 = (an_{k+1} + n_{k+2})^2 + (n_{k+1} - an_{k+2})^2.$$

So, if $k+2 \le d'$, then since $a^2+1 \not\equiv 0 \bmod p$, under another change of variables, we may use another change of variables to replace $cn_{k+1}^2 + cn_{k+2}^2$ by $n_{k+1}^2 + n_{k+2}^2$ in the expression of M. Inductively, we are reduced to the case $k = d'-1$ or d'. Therefore, (4.1) holds by switching the position of n_1 and $n_{d'}$.

The 'moreover' part follows immediately through the proof. □

4.2. *M-isotropic subspaces.* Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A and V be a subspace of $\mathbb{F}_p^d$. Let $V^{\perp M}$ denote the set of $\{n \in \mathbb{F}_p^d\colon (mA)\cdot$ n=0 for all $m \in V\}$.

PROPOSITION 4.5. *Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A and V be a subspace of $\mathbb{F}_p^d$.*

(i) *$V^{\perp M}$ is a subspace of $\mathbb{F}_p^d$, and $V \subseteq (V^{\perp M})^{\perp M}$. Moreover, $(V^{\perp M})^{\perp M} = V$ if M is non-degenerate.*

(ii) *Let $\phi\colon \mathbb{F}_p^d \to \mathbb{F}_p^d$ be a bijective linear transformation, $v \in \mathbb{F}_p^d$, and $c \in \mathbb{F}_p$. Let $M'\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be the quadratic form given by $M'(n) := M(\phi(n)+v)+c$. We have that* $\operatorname{rank}(M) = \operatorname{rank}(M')$ *and* $\phi(V^{\perp M'}) = (\phi(V))^{\perp M}$.

(iii) $\operatorname{rank}(M)$ *is equal to* $d - \dim((\mathbb{F}_p^d)^{\perp M})$.

Proof. Part (i) is easy to check by definition and the symmetry of A. For part (ii), assume that $\phi(x) = xB$ for some $d \times d$ invertible matrix B, then $\operatorname{rank}(M) = \operatorname{rank}(A) = \operatorname{rank}(BAB^T) = \operatorname{rank}(M \circ \phi)$. So, $\operatorname{rank}(M) = \operatorname{rank}(M \circ \phi) = \operatorname{rank}(M(\phi(n)+v))$ since shifting by v does not affect the rank of a quadratic form. The part $\phi(V^{\perp M'}) = (\phi(V))^{\perp M}$ is easy to check by definition.

We now prove part (iii). Let $r = \operatorname{rank}(M)$. Then, $1 \le r \le d$. By Lemma 4.3, we may assume that $M(n) = M'(\phi(n)+v)$ for some bijective linear transformation $\phi\colon \mathbb{F}_p^d \to \mathbb{F}_p^d$, $v \in \mathbb{F}_p^d$, $c, c', \lambda \in \mathbb{F}_p$, $c \ne 0$, and some quadratic form $M'\colon \mathbb{F}_p^d \to \mathbb{F}_p$ of the form

$$M'(n_1, \dots, n_d) = cn_1^2 + n_2^2 + \cdots + n_r^2 + c'n_{r+1} - \lambda.$$

It is clear that $\operatorname{rank}(M') = d - \dim((\mathbb{F}_p^d)^{\perp M'}) = r = \operatorname{rank}(M)$. By part (iii), $(\mathbb{F}_p^d)^{\perp M} = \phi^{-1}((\phi(\mathbb{F}_p^d))^{\perp M'}) = \phi^{-1}((\mathbb{F}_p^d)^{\perp M'})$. So, $\dim((\mathbb{F}_p^d)^{\perp M}) = \dim((\mathbb{F}_p^d)^{\perp M'})$ and thus, $\operatorname{rank}(M) = d - \dim((\mathbb{F}_p^d)^{\perp M})$. □

For $n \in \mathbb{R}^d$, the dot product induces the Euclidean norm $|n|^2 := n \cdot n$. However, in $\mathbb{F}_p^d$, the dot product (as well as general quadratic forms) does not induce a norm in $\mathbb{F}_p^d$, due to the existence of isotropic subspaces.

Definition 4.6. (M-isotropic subspaces) Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A. A subspace V of $\mathbb{F}_p^d$ is *M-isotropic* if $V \cap V^{\perp M} \ne \{\mathbf{0}\}$. We say that a tuple $(h_1, \dots, h_k)$ of vectors in $\mathbb{F}_p^d$ is *M-isotropic* if the span of $h_1, \dots, h_k$ is an

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

M-isotropic subspace. We say that a subspace or tuple of vectors is *non- M-isotropic* if it is not M-isotropic.

For example, if $M(n) = n \cdot n$ and $V = \mathrm{span}_{\mathbb{F}_{11}}\{(1, 1, 3), (1, 2, -1)\}$, then $(1, 1, 3) \in V^{\perp M}$ and so, V is M-isotropic. However, for $V' = \mathrm{span}_{\mathbb{F}_{11}}\{(1, 1, 3), (1, 1, 1)\}$, if $v \in V' \cap (V')^{\perp M}$, then since $v \in V$, we may write $v = x(1, 1, 3) + y(1, 1, 1)$ for some $x, y \in \mathbb{F}_{11}$. Since v is orthogonal to $(1, 1, 3)$ and $(1, 1, 1)$, we have that $0x + 5y = 5x + 3y = 0$, which implies that $x = y = 0$. So, V' is not M-isotropic.

Let $V = \mathrm{span}_{\mathbb{F}_p}\{h_1, \ldots, h_k\}$ be a subspace of $\mathbb{F}_p^d$ and A be the matrix associated to M. Then, V is M-isotropic if and only if the determinant of the matrix $((h_i A) \cdot h_j)_{1 \le i,j \le k}$ is 0. To see this, suppose that $v \in V \cap V^{\perp M}$. Since $v \in V$, we may write $v = \sum_{i=1}^k x_i h_i$ for some $x_i \in \mathbb{F}_p$. Then, $v \in V \cap V^{\perp M} \Leftrightarrow (vA) \cdot h_j = \sum_{i=1}^k x_i (h_i A) \cdot h_j = 0$ for $1 \le j \le k$, the right-hand side of which is a $k \times k$ linear equation system whose coefficient matrix is $((h_i A) \cdot h_j)_{1 \le i,j \le k}$.

The next lemma ensures the existence of non-M-isotropic subspaces.

LEMMA 4.7. *Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form. Then, for all $0 \le k \le \mathrm{rank}(M)$, there exists a non-$M$-isotropic subspace of $\mathbb{F}_p^d$ of dimension k.*

Proof. Since $\{\mathbf{0}\}$ is non-M-isotropic, it suffices to consider the case $k \ge 1$. Let $d' = \mathrm{rank}(M)$. By Lemma 4.3, under a change of variables, it suffices to consider the case when the associated matrix A of M is a diagonal matrix whose diagonal can be written as $(a_1, \ldots, a_{d'}, 0, \ldots, 0)$ for some $a_1, \ldots, a_{d'} \neq 0$. Let $e_1, \ldots, e_d$ denote the standard unit vectors. Clearly, for $1 \le k \le d'$, the matrix $((e_i A) \cdot e_j)_{1 \le i,j \le k}$ is the upper left $k \times k$ block of A, which is of determinant $a_1 \cdot \cdots \cdot a_k \neq 0$. This implies that $\mathrm{span}_{\mathbb{F}_p}\{e_1, \ldots, e_k\}$ is a non-M-isotropic subspace of $\mathbb{F}_p^d$ of dimension k. □

We summarize some basic properties of M-isotropic subspaces for later use.

PROPOSITION 4.8. *Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form and V be a subspace of $\mathbb{F}_p^d$ of co-dimension r, and $c \in \mathbb{F}_p^d$.*

(i) *We have* $\dim(V \cap V^{\perp M}) \le \min\{d - \mathrm{rank}(M) + r, d - r\}$.

(ii) *The rank of $M|_{V+c}$ is equal to $d - r - \dim(V \cap V^{\perp M})$ (that is,* $\dim(V) - \dim(V \cap V^{\perp M})$*).*

(iii) *The rank of $M|_{V+c}$ is at most $d - r$ and at least* $\mathrm{rank}(M) - 2r$.

(iv) *$M|_{V+c}$ is non-degenerate (that is,* $\mathrm{rank}(M|_{V+c}) = d - r$*) if and only if V is not an M-isotropic subspace.*

Proof. Part (i). Let $v_1, \ldots, v_{d-r}$ be a basis of V. Let $B = \begin{bmatrix} v_1 \\ \cdots \\ v_{d-r} \end{bmatrix}$. The dimension of $\dim(V^{\perp M})$ is equal to d minus the rank of the matrix BA. Since $\mathrm{rank}(BA) \ge \mathrm{rank}(B) + \mathrm{rank}(A) - d = \mathrm{rank}(M) - r$, we have that $\dim(V^{\perp M}) \le d - \mathrm{rank}(M) + r$. So, part (i) holds because $\dim(V \cap V^{\perp M}) \le \min\{\dim(V), \dim(V^{\perp M})\}$.

Part (ii). Let $\phi\colon \mathbb{F}_p^{d-r} \to V$ be a bijective linear transformation and denote $M'(m) := M(\phi(m) + c)$. Then, the rank of $M|_{V+c}$ is equal to the rank of M'. Assume that

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$\phi(m) = mB$ for some $(d-r) \times d$ matrix B of rank $d-r$. Then, the matrix associated to M' is BAB^T. Note that for all $m \in \mathbb{F}_p^r$,

$$\begin{aligned} m \in (\mathbb{F}_p^r)^{\perp M'} &\Leftrightarrow ((mB)A) \cdot (nB) = 0 \quad \text{for all } n \in \mathbb{F}_p^r \\ &\Leftrightarrow ((mB)A) \cdot v = 0 \quad \text{for all } v \in V \Leftrightarrow mB \in V \cap V^{\perp M}. \end{aligned}$$

By Proposition 4.5,

$$\operatorname{rank}(M') = d - r - \dim(\mathbb{F}_p^r)^{\perp M'} = d - r - \dim(V \cap V^{\perp M}).$$

This proves part (ii).

Parts (iii) and (iv) follows directly from parts (i) and (ii). □

An annoying issue when working with subspaces of $\mathbb{F}_p^d$ is that the restriction of a non-degenerate quadratic form M to an M-isotropic subspace V of $\mathbb{F}_p^d$ is not necessarily non-degenerate. Fortunately, due to Proposition 4.8(iii), the nullity of M does not increase too much when restricted to V, provided that the co-dimension of V is small. This is why Proposition 4.8 can be viewed as the inheriting principle for the ranks of quadratic forms.

The following is a consequence of Proposition 4.8, which will be used in [**15**].

LEMMA 4.9. *Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, V be a subspace of $\mathbb{F}_p^d$ of dimension r, and V' be a subspace of $\mathbb{F}_p^d$ of dimension r'. Suppose that* $\operatorname{rank}(M|_{V^{\perp M}}) = d - r$. *Then,* $\operatorname{rank}(M|_{(V+V')^{\perp M}}) \geq d - r - 2r'$.

Proof. Since $V^{\perp M}$ is of co-dimension r, by Proposition 4.8(ii), $\operatorname{rank}(M|_{V^{\perp M}}) = d - r$ implies that $\dim(V \cap V^{\perp M}) = 0$. Since $\dim(V + V') \leq r + r'$ and M is non-degenerate, we have that $\dim((V+V')^{\perp M}) \geq d - r - r'$. By Proposition 4.8(ii),

$$\begin{aligned} \operatorname{rank}(M|_{(V+V')^{\perp M}}) &\geq d - r - r' - \dim((V+V') \cap (V+V')^{\perp M}) \\ &\geq d - r - r' - \dim((V+V') \cap V^{\perp M}). \end{aligned}$$

So, it remains to show that $\dim((V+V') \cap V^{\perp M}) \leq r'$. It suffices to show that for any $v_1, \ldots, v_{r'+1} \in (V+V') \cap V^{\perp M}$, these vectors are linearly dependent. Since $v_1, \ldots, v_{r'+1} \in V + V'$, we may write $v_i = u_i + w_i$ for some $u_i \in V$ and $w_i \in V'$ for all $1 \leq i \leq r'+1$. Since $\dim(V') = r'$, there exist $c_1, \ldots, c_{r'+1}$ not all equal to 0 such that $c_1 w_1 + \cdots + c_{r'+1} w_{r'+1} = \mathbf{0}$. Therefore, we have that $c_1 v_1 + \cdots + c_{r'+1} v_{r'+1} = c_1 u_1 + \cdots + c_{r'+1} u_{r'+1} \in V$. Since $c_1 v_1 + \cdots + c_{r'+1} v_{r'+1} \in V^{\perp M}$ and $V \cap V^{\perp M} = \{\mathbf{0}\}$, we have that $c_1 v_1 + \cdots + c_{r'+1} v_{r'+1} = \mathbf{0}$. This means that $\dim((V+V') \cap V^{\perp M}) \leq r'$ and we are done. □

4.3. *Counting the zeros of quadratic forms.* For a polynomial $P \in \operatorname{poly}(\mathbb{F}_p^k \to \mathbb{F}_p)$, let $V(P)$ denote the set of $n \in \mathbb{F}_p^k$ such that $P(n) = 0$. For a family of polynomials $\mathcal{J} = \{P_1, \ldots, P_k\} \subseteq \operatorname{poly}(\mathbb{F}_p^k \to \mathbb{F}_p)$, denote $V(\mathcal{J}) := \bigcap_{i=1}^k V(P_i)$. The purpose of this section is to provide estimates on the cardinality of $V(\mathcal{J})$ for families of polynomials arising from quadratic forms.

We start with some basic estimates.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

LEMMA 4.10. (Schwartz–Zippel lemma) *Let $P \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ be of degree at most r. Then, $|V(P)| \leq O_{d,r}(p^{d-1})$ unless $P \equiv 0$.*

As an application of Lemma 4.10, we have the following lemma.

LEMMA 4.11. *Let $d, k \in \mathbb{N}_+$, p be a prime, and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form.*

(i) *The number of tuples $(h_1, \ldots, h_k) \in (\mathbb{F}_p^d)^k$ such that $h_1, \ldots, h_k$ are linearly dependent is at most $kp^{(d+1)(k-1)}$.*

(ii) *The number of M-isotropic tuples $(h_1, \ldots, h_k) \in (\mathbb{F}_p^d)^k$ is at most $O_{d,k}(p^{kd-1})$.*

Proof. Part (i). If $h_1, \ldots, h_k$ are linearly dependent, then there exists $1 \leq i \leq k$ such that $h_i = \sum_{1 \leq j \leq k, j \neq i} c_j h_j$ for some $c_j \in \mathbb{F}_p$. For each j, there are p choices of c_j and p^d choices of h_j. Since there are k choices of i, in total, there are at most $k \cdot (p \cdot p^d)^{k-1} = kp^{(d+1)(k-1)}$ possibilities.

Part (ii). By the discussion after Definition 4.6, $h_1, \ldots, h_k$ are M-isotropic if and only if the determinant of the matrix $((h_i A) \cdot h_j)_{1 \leq i,j \leq k}$ is zero. Viewing this matrix as a non-constant polynomial in $h_1, \ldots, h_k$, we get the conclusion from Lemma 4.10. □

A very useful application of Lemma 4.11 is as follows. Whenever we have a collection X of k-tuples $(h_1, \ldots, h_k) \in (\mathbb{F}_p^d)^k$ of positive density $\delta > 0$, if the dimension d is sufficiently large compared with k, and if p is sufficiently large compared with δ, d, k, then by passing to a subset of X, we may further require all the tuples in X to be linearly independent/non-M-isotropic with X still having a positive density.

The following lemma was proved in [**10**]. We provide its details for completeness.

LEMMA 4.12. *Let $d \in \mathbb{N}_+$ and p be a prime number. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form of rank r. Suppose that $r \geq 3$. Then,*

$$|V(M)| = p^{d-1}(1 + O(p^{-(r-2)/2})),$$

and for all $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{Z}^d$, we have that

$$\mathbb{E}_{n \in V(M)} \exp\left(\frac{\xi}{p} \cdot n\right) = \mathbf{1}_{\xi = \mathbf{0}} + O(p^{-(r-2)/2}). \tag{4.2}$$

Here, we slightly abuse the notation in $\xi/p \cdot n$ by identifying $n \in \mathbb{F}_p^d$ with its embedding $\tau(n) \in \mathbb{Z}^d$.

Proof. By Lemma 4.3, it is easy to see we only need to prove this lemma for the case when

$$M(n_1, \ldots, n_d) = a_1 n_1^2 + \cdots + a_r n_r^2 + a_{r+1} n_{r+1} + \cdots + a_d n_d - \lambda$$

for some $a_1, \ldots, a_r \in \mathbb{F}_p \backslash \{0\}$, $a_{r+1}, \ldots, a_d \in \mathbb{F}_p$ and $\lambda \in \mathbb{F}_p$. (Here, we do not need to use the full strength of Lemma 4.3.) Note that

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$$\sum_{n\in V(M)} \exp\left(\frac{\xi}{p}\cdot n\right) = \sum_{n\in\mathbb{F}_p^d} \mathbb{E}_{j\in\mathbb{F}_p} \exp\left(\frac{\xi}{p}\cdot n + \frac{j}{p}M(n)\right)$$

$$= p^{d-1}\mathbf{1}_{\xi=\mathbf{0}} + \frac{1}{p}\sum_{j\in\mathbb{F}_p\setminus\{0\}} \exp(-j\lambda) \sum_{n\in\mathbb{F}_p^d} \exp\left(\frac{\xi}{p}\cdot n + \frac{j}{p}\sum_{i=1}^{r} a_i n_i^2 + \frac{j}{p}\sum_{i=r+1}^{d} a_i n_i\right).$$

By a change of coordinate, we have that

$$\begin{aligned}
&\left|\frac{1}{p}\sum_{j\in\mathbb{F}_p\setminus\{0\}} \exp(-j\lambda) \sum_{n\in\mathbb{F}_p^d} \exp\left(\frac{\xi}{p}\cdot n + \frac{j}{p}\sum_{i=1}^{r} a_i n_i^2 + \frac{j}{p}\sum_{i=r+1}^{d} a_i n_i\right)\right| \\
&\quad\le \frac{1}{p}\sum_{j\in\mathbb{F}_p\setminus\{0\}} \left|\sum_{n\in\mathbb{F}_p^d} \exp\left(\frac{\xi}{p}\cdot n + \frac{j}{p}\sum_{i=1}^{r} a_i n_i^2 + \frac{j}{p}\sum_{i=r+1}^{d} a_i n_i\right)\right| \\
&\quad= \frac{1}{p}\sum_{j\in\mathbb{F}_p\setminus\{0\}} \left|\sum_{n\in\mathbb{F}_p^d} \exp\left(\frac{j}{p}\sum_{i=1}^{r} a_i n_i^2 + \frac{j}{p}\sum_{i=r+1}^{d} a_i n_i\right)\right| \\
&\quad= \frac{1}{p}\sum_{j\in\mathbb{F}_p\setminus\{0\}} \prod_{i=1}^{r}\left|\sum_{n\in\mathbb{F}_p} \exp(ja_i n^2/p)\right| \cdot \prod_{i=r+1}^{d}\left|\sum_{n\in\mathbb{F}_p} \exp(ja_i n/p)\right| \\
&\quad\le \frac{1}{p}\sum_{j\in\mathbb{F}_p\setminus\{0\}} p^{d-r}\prod_{i=1}^{r}\left|\sum_{n\in\mathbb{F}_p} \exp(ja_i n^2/p)\right|.
\end{aligned}$$

Since

$$\begin{aligned}
&\left|\sum_{n\in\mathbb{F}_p} \exp(ja_i n^2/p)\right|^2 \\
&\quad= \left|\sum_{n,h\in\mathbb{F}_p} \exp(ja_i(2hn+h^2)/p)\right| \le \sum_{h\in\mathbb{F}_p}\left|\sum_{n\in\mathbb{F}_p} \exp(ja_i(2hn+h^2)/p)\right| = p,
\end{aligned}$$

we have that

$$\sum_{n\in V(M)} \exp\left(\frac{\xi}{p}\cdot n\right) = p^{d-1}\mathbf{1}_{\xi=\mathbf{0}} + O(p^{d-(r/2)}). \tag{4.3}$$

Setting $\xi = \mathbf{0}$, we have that $|V(M)| = p^{d-1}(1 + O(p^{-(r-2)/2}))$. Dividing both sides of (4.3) by $|V(M)|$, we get (4.2). □

We remark that while the argument of Lemma 4.12 works for all $r \in \mathbb{N}$, the term $O(p^{-(r-2)/2})$ may be dominant if $r \le 2$. For example, if $M(n_1, n_2) = n_1^2 + an_2^2$ for some $a \in \mathbb{F}_p$ such that $x^2 \ne -a$ for all $x \in \mathbb{F}_p$, then $V(M)$ is an empty set.

The following is an application of Lemma 4.12. We say that t is a *quadratic residue* of $\mathbb{F}_p$ if there exists $n \in \mathbb{F}_p$ such that $n^2 = t$. It is clear that if $p > 2$, then the number of quadratic residues of $\mathbb{F}_p$ is $(p+1)/2$. We have the following lemma.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

LEMMA 4.13. *Let $d, r \in \mathbb{N}_+$, p be a prime number, and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form of rank r. Then, the number of $n \in \mathbb{F}_p^d$ such that $M(n)$ is a quadratic residue is $\frac{1}{2}p^d(1 + O(p^{-(r-2)/2}))$ if $r \geq 3$, and $\frac{1}{2}p^d(1 + O(p^{-(1/2)}))$ for $r = 2$.*

Proof. Let $S = \{s_1, \ldots, s_{(p+1)/2}\}$ be the set of quadratic residues of $\mathbb{F}_p$. We first consider the case $r \geq 3$. By Lemma 4.12, the number of $n \in \mathbb{F}_p^d$ such that $M(n) \in S$ is

$$\sum_{i=1}^{(p+1)/2} |V(M - s_i)| = \frac{p+1}{2} \cdot p^{d-1}(1 + O(p^{-(r-2)/2})) = \frac{1}{2}p^d(1 + O(p^{-(1/2)})).$$

We now consider the case $r = 2$. By [**13**, Theorem 5A on p. 52], the number of $n \in \mathbb{F}_p^2$ such that $M(n) = s_i$ is $p(1 + O(p^{-1/2}))$. So, the total number of $n \in \mathbb{F}_p^d$ such that $M(n)$ is a quadratic residue is $\frac{1}{2}p(p+1)(1 + O(p^{-1/2}))$. We are done. □

In the rest of this section, we provide some generalizations of Lemma 4.12. These results can be interpreted as the inheriting principle for the size of hyperspheres. First of all, as an immediate corollary of Lemma 4.12, we have the following estimate on the intersection of $V(M)$ with an affine subspace of $\mathbb{F}_p^d$.

COROLLARY 4.14. *Let $d \in \mathbb{N}_+$, $r \in \mathbb{N}$, and p be a prime number. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form and $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of co-dimension r. If either* $\operatorname{rank}(M|_{V+c}) \geq 3$ *or* $\operatorname{rank}(M) - 2r \geq 3$, *then*

$$|V(M) \cap (V + c)| = p^{d-r-1}(1 + O(p^{-(1/2)})).$$

Proof. If $\operatorname{rank}(M) - 2r \geq 3$, then $\operatorname{rank}(M|_{V+c}) \geq 3$ by Proposition 4.8(iii). So, in both cases, we have that $\operatorname{rank}(M|_{V+c}) \geq 3$. Let $\phi\colon \mathbb{F}_p^{d-r} \to V$ be a bijective linear transformation and denote $M'(m) := M(\phi(m) + c)$ for all $m \in \mathbb{F}_p^{d-r}$. Then, $\operatorname{rank}(M') \geq 3$. Note that $n \in V(M) \cap (V + c)$ if and only if $n = \phi(m) + c$ for some $m \in \mathbb{F}_p^{d-r}$ such that $M'(m) = 0$. So, the conclusion follows from Lemma 4.12. □

The following sets are studied extensively in this paper.

Definition 4.15. Let $r \in \mathbb{N}_+$, $h_1, \ldots, h_r \in \mathbb{F}_p^d$, and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form. Denote

$$V(M)^{h_1,\ldots,h_r} := \bigcap_{i=1}^{r} V(M(\cdot + h_i)) \cap V(M).$$

Lemma 4.12 can also be used to study the cardinality of $V(M)^{h_1,\ldots,h_r}$.

COROLLARY 4.16. *Let $d, r \in \mathbb{N}_+$ and p be a prime number. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form and $h_1, \ldots, h_r$ be linearly independent vectors. If $d - 2r \geq 3$, then*

$$|V(M)^{h_1,\ldots,h_r}| = p^{d-r-1}(1 + O(p^{-(1/2)})).$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Proof. Since $h_1, \dots, h_r$ are linearly independent and M is non-degenerate, it is not hard to see that $V(M)^{h_1,\dots,h_r}$ is the intersection of $V(M)$ with an affine subspace $V + c$ of $\mathbb{F}_p^d$ of co-dimension r. The conclusion follows from Corollary 4.14. □

We also need to study the following sets in this paper.

Definition 4.17. (Gowers set) Let Ω be a subset of $\mathbb{F}_p^d$ and $s \in \mathbb{N}$. Let $\square_s(\Omega)$ denote the set of $(n, h_1, \dots, h_s) \in (\mathbb{F}_p^d)^{s+1}$ such that $n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s \in \Omega$ for all $(\epsilon_1, \dots, \epsilon_s) \in \{0,1\}^s$. Here, we allow s to be 0, in which case, $\square_0(\Omega) = \Omega$. We say that $\square_s(\Omega)$ is the *sth Gowers set* of Ω.

It is worth noting that the Gowers set will also be used for the study of the spherical Gowers norms in [16]. We have the following description for Gowers sets.

LEMMA 4.18. *Let $s \in \mathbb{N}$, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A, and $V + c$ be an affine subspace of $\mathbb{F}_p^d$. For $n, h_1, \dots, h_s \in \mathbb{F}_p^d$, we have that $(n, h_1, \dots, h_s) \in \square_s(V(M) \cap (V + c))$ if and only if:*

- $n \in V + c$, $h_1, \dots, h_s \in V$;
- $n \in V(M)^{h_1,\dots,h_s}$;
- $(h_i A) \cdot h_j = 0$ *for all* $1 \le i, j \le s, i \ne j$.

In particular, let $\phi\colon \mathbb{F}_p^r \to V$ be any bijective linear transformation, $M'(m) := M(\phi(m) + c)$. We have that $(n, h_1, \dots, h_s) \in \square_s(V(M) \cap (V + c))$ if and only if $(n, h_1, \dots, h_s) = (\phi(n') + c, \phi(h_1'), \dots, \phi(h_s'))$ for some $(n', h_1', \dots, h_s') \in \square_s(V(M'))$.

Proof. If $(n, h_1, \dots, h_s) \in \square_s(V(M) \cap (V + c))$, then clearly, $n \in V(M)^{h_1,\dots,h_s}$. Since $n, n + h_1, \dots, n + h_s \in V + c$, we have that $h_1, \dots, h_s \in V$. Finally, the fact that $(h_i A) \cdot h_j = 0$ for all $1 \le i, j \le s, i \ne j$ follows from the definition of $\square_s(V(M) \cap (V + c))$ and Lemma 4.2.

Conversely, if $n \in V + c, h_1, \dots, h_s \in V$, then $n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s \in V + c$ for all $\epsilon_1, \dots, \epsilon_s \in \{0,1\}$. Since $n \in V(M)^{h_1,\dots,h_s}$, we have that $M(n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s) = 0$ whenever $|\epsilon| := \epsilon_1 + \cdots + \epsilon_s = 0$ or 1. Suppose we have shown that $M(n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s) = 0$ whenever $|\epsilon| := \epsilon_1 + \cdots + \epsilon_s \le i$ for some $1 \le i \le s - 1$, then it follows from Lemma 4.2 and the fact that $(h_i A) \cdot h_j = 0$ for all $1 \le i, j \le s, i \ne j$ that $M(n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s) = 0$ whenever $|\epsilon| := \epsilon_1 + \cdots + \epsilon_s \le i + 1$. In conclusion, we have that $(n, h_1, \dots, h_s) \in \square_s(V(M) \cap (V + c))$.

The proof of the 'in particular' part is left to the interested readers. □

We summarize some basic properties of Gowers sets.

PROPOSITION 4.19. (Some preliminary properties for Gowers sets) *Let $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ and p be a prime. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form with* $\operatorname{rank}(M) \ge s^2 + s + 3$. *Then:*

(i) *we have $|\square_s(V(M))| = p^{(s+1)d - ((s(s+1)/2)+1)}(1 + O_s(p^{-1/2}))$;*

(ii) *for all but at most $s(s+1)p^{d+s-\operatorname{rank}(M)}$ many $h_s \in \mathbb{F}_p^d$, the set of $(n, h_1, \dots, h_{s-1}) \in (\mathbb{F}_p^d)^s$ with $(n, h_1, \dots, h_s) \in \square_s(V(M))$ is at most $p^{ds - ((s(s+1)/2)+1)}$;*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

(iii) *if $s = 1$, then for any function $f : \square_1(V(M)) \to \mathbb{C}$ with norm bounded by 1, we have that*

$$\mathbb{E}_{(n,h)\in\square_1(V(M))} f(n,h) = \mathbb{E}_{h\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^h} f(n,h) + O(p^{-1/2}).$$

It is worth mentioning that Proposition 4.19(iii) can be interpreted as *Fubini's theorem* on $\square_1(V(M))$, which says the two-dimensional average of (n,h) over $\square_1(V(M))$ is approximately the same as the double average of n and h.

Proposition 4.19 can either be proved by direct computation plus the counting properties developed in §4, or it can be derived from a more general framework using [**18**, Example 5.4 and Theorem 5.11]. So, we only provide a sketch of the proof.

Sketch of the proof. Clearly, part (i) holds for $s = 0$ by Lemma 4.12. Suppose that part (i) holds for s and now we prove it for $s+1$. By Lemma 4.18, $(n, h_1, \dots, h_{s+1}) \in \square_{s+1}(V(M))$ if and only if $(n, h_1, \dots, h_s) \in \square_s(V(M))$ and h_{s+1} belongs to the set

$$W(n, h_1, \dots, h_s) := \{m \in \mathbb{F}_p^d : M(n+m) = (h_1A)\cdot m = \cdots = (h_sA)\cdot m = 0\}.$$

For most of the choices of $(n, h_1, \dots, h_s)$, $W(n, h_1, \dots, h_s)$ is the intersection of a hypersphere with a hyperplane of codimension s, and so is of cardinality approximately p^{d-s-1} by Corollary 4.14. So, by induction hypothesis,

$$\begin{aligned}|\square_{s+1}(V(M))| &\approx |\square_s(V(M))|\cdot p^{d-s-1} \approx p^{(s+1)d-((s(s+1)/2)+1)}\cdot p^{d-s-1}\\ &= p^{(s+2)d-(((s+1)(s+2)/2)+1)}.\end{aligned}$$

For part (ii), using Lemma 4.18, one can show that $(n, h_1, \dots, h_s) \in \square_s(V(M))$ if and only if $(n, h_1, \dots, h_{s-1}) \in \square_{s-1}(V(M))$, $n \in V + c$, and $h_1, \dots, h_{s-1} \in V$, where $V + c$ is the affine subspace of $\mathbb{F}_p^d$ consisting of all h with $M(h + h_s) = M(h)$, and V is the affine subspace of $\mathbb{F}_p^d$ consisting of all h with $(h_sA)\cdot h = 0$. This is equivalent of saying that $(n, h_1, \dots, h_{s-1}) \in \square_{s-1}(V(M) \cap (V + c))$. One can informally think of $V(M) \cap (V + c)$ as a hypersphere in a subspace of dimension $d - 1$ (for most of h_s), and so the number of such $(n, h_1, \dots, h_{s-1})$ by part (i) is approximately $p^{s(d-1)-(((s-1)s/2)+1)} = p^{sd-((s(s+1)/2)+1)}$.

Part (iii) can be derived by combining parts (i) and (ii). We omit the details. □

5. *The main equidistribution theorem*

With the help of the preparation in the previous sections, we are now ready to state the main result in this paper, which is a generalization of Theorem 1.4.

THEOREM 5.1. (Equidistribution theorem on $V(M)$) *Let $0 < \delta < 1/2$, $d, m \in \mathbb{N}_+$, $s, r \in \mathbb{N}$ with $d \geq r$, and p be a prime. Let $M : \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form and $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of co-dimension r. Suppose that* $\operatorname{rank}(M|_{V+c}) \geq s + 13$. *Let G/Γ be an s-step $\mathbb{N}$-filtered nilmanifold of dimension m, equipped with a $1/\delta$-rational Mal'cev basis $\mathcal{X}$, and let $g \in \operatorname{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ be a rational polynomial sequence. If $p \gg_{d,m} \delta^{-O_{d,m}(1)}$ and $(g(n)\Gamma)_{n\in\iota^{-1}(V(M)\cap(V+c))}$ is not δ-equidistributed on G/Γ, then there exists a non-trivial type-I horizontal character η with $0 < \|\eta\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ (independent of*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

k, M, and p) such that $\eta \circ g \bmod \mathbb{Z}$ is a constant on $\iota^{-1}(V(M) \cap (V + c))$. (It follows from Corollary 4.14 that $V(M) \cap (V + c)$ is non-empty since $d - r \geq s + 13$.)

To illustrate the idea for the proof of Theorem 5.1, we first present an outline of the proof of Corollary 1.6 (which is a special case of Theorem 5.1). Assume for simplicity that g is p-periodic. Suppose we have proven Corollary 1.6 for the case $s - 1$. Denote $\Omega = V(M)$. Assume on the contrary that

$$\left| \frac{1}{|\Omega|} \sum_{n \in \Omega} \exp(g \circ \tau(n)/p) \right| > \delta. \tag{5.1}$$

Taking squares on both sides and performing a change of variables, we have that

$$\begin{aligned} \delta^2 &< \frac{1}{|\Omega|^2} \sum_{n, n+h \in \Omega} \exp((g \circ \tau(n+h) - g \circ \tau(n))/p) \\ &= \frac{1}{|\Omega|^2} \sum_{h \in \mathbb{F}_p^d} \sum_{n \in \Omega \cap (\Omega - h)} \exp((g \circ \tau(n+h) - g \circ \tau(n))/p). \end{aligned}$$

Since (by the inheriting principle) the sets $\Omega \cap (\Omega - h)$, $h \in \mathbb{F}_p^d$ are hyperspheres whose cardinalities are comparable with each other for most of h, we have that

$$\frac{1}{p^d} \sum_{h \in \mathbb{F}_p^d} \left| \frac{1}{|\Omega \cap (\Omega - h)|} \sum_{n \in \Omega \cap (\Omega - h)} \exp((g \circ \tau(n+h) - g \circ \tau(n))/p) \right| \gg \delta^2$$

(this argument can be made precise by Proposition 4.19). By the pigeonhole principle, for 'many' $h \in \mathbb{F}_p^d$, we have that

$$\left| \frac{1}{|\Omega \cap (\Omega - h)|} \sum_{n \in \Omega \cap (\Omega - h)} \exp((g \circ \tau(n+h) - g \circ \tau(n))/p) \right| \gg \delta^2.$$

Informally, one can view $\Omega \cap (\Omega - h)$ as a sphere whose dimension is lower than Ω by 1. So, we may then apply the induction hypothesis to the degree $(s-1)$ polynomial sequence $g \circ \tau(n+h) - g \circ \tau(n)$ to conclude that

$$(g \circ \tau(n+h) - g \circ \tau(n))/p \bmod \mathbb{Z} \text{ is a constant on } \Omega \cap (\Omega - h) \quad \text{for 'many' } h \in \mathbb{F}_p^d. \tag{5.2}$$

(For the case when g is a polynomial sequence on a nilmanifold, the conclusion becomes $(\eta_h \circ g \circ \tau(n+h) - \eta_h \circ g \circ \tau(n))/p \bmod \mathbb{Z}$ is a constant on $\Omega \cap (\Omega - h)$ for 'many' $h \in \mathbb{F}_p^d$ for some horizontal character η_h of low complexity. The dependence of η_h on h can then be dropped by using the pigeonhole principle.)

For convenience, denote $Q := \iota \circ pg \circ \tau$, then $Q \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ is a polynomial of degree at most s. It then follows from (5.2) that

$$\Delta_{h_2} \Delta_{h_1} Q(n) = 0 \quad \text{for 'many'} (n, h_1, h_2) \in \square_2(\Omega), \tag{5.3}$$

where $\square_2(\Omega)$ is the second Gowers set of Ω defined in Definition 4.17.

We have now reduced Corollary 1.6 to a problem of solving the polynomial equation (5.3). If Ω were the entire space $\mathbb{F}_p^d$, then using some basic knowledge for polynomials,

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

it is not hard to show that (5.3) would imply that Q is of degree at most $s-1$, which would complete the proof of Corollary 1.6 by induction. When Ω is a sphere, the difficulty of the analysis of (5.3) increases significantly. Therefore, solving (5.3) (as well as its various generalizations) is a central topic in this paper and this is the main challenge in adapting the Green–Tao approach [**5**] to the quadratic case.

In Proposition 6.8, we show that (5.3) implies that Q can be written as $Q = MQ_1 + Q_2$ for some $Q_1, Q_2 \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(Q_2) \leq 1$. Then, (5.1) implies that

$$\left|\frac{1}{|\Omega|}\sum_{n\in\Omega}\exp\left(\frac{1}{p}\tau\circ Q(n)\right)\right| = \left|\frac{1}{|\Omega|}\sum_{n\in\Omega}\exp\left(\frac{1}{p}\tau\circ Q_2(n)\right)\right| > \delta.$$

Since Q_2 is linear, it forces $Q_2(n)/p \bmod Z$ to be a constant and therefore, $Q(n)/p \bmod \mathbb{Z}$ is a constant on Ω. This implies that $g\circ\tau(n)/p \bmod \mathbb{Z}$ is a constant on Ω, completing the (outline of the) proof of Corollary 1.6.

To prove Theorem 5.1, a more general version of Corollary 1.6, following the idea of Green and Tao [**5**, **6**], one can reduce the problem to solving the following generalized versions of (5.3) (see also (9.19)):

$$\Delta_{h_3}\Delta_{h_2}\Delta_{h_1}P(n) + \Delta_{h_2}\Delta_{h_1}Q(n) = 0 \quad \text{for 'many' } (n, h_1, h_2, h_3) \in \square_3(\Omega), \tag{5.4}$$

where $Q \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ comes from the composition of g with some type-I horizontal character of G/Γ (so $\deg(Q) \leq s$), and $P \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ is a polynomial of degree at most $s-1$ arising from the 'nonlinear' part of g.

The bulk of §6 is devoted to solving the following generalized version of (5.4) (see also (6.25)):

$$\Delta_{h_{s-1}}\cdots\Delta_{h_1}P(n) + \Delta_{h_s}\cdots\Delta_{h_1}Q(n) = 0 \quad \text{for 'many' } (n, h_1, \ldots, h_s) \in \square_s(\Omega). \tag{5.5}$$

(We study (5.5) for any $s \in \mathbb{N}$ not just for $s = 3$ for the purpose of future research.) We provide a solution to (5.5) in Proposition 6.6. In the same section, we also prove some other properties for polynomial equations in $\mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ for later use.

One can then combine the solution to (5.4) together with the machinery of Green and Tao [**5**, **6**] to prove Theorem 5.1 for the case when g is p-periodic. However, to deal with a partially p-periodic sequence g, we need to deal with the following even more general version of (5.5):

$$\Delta_{h_{s-1}}\cdots\Delta_{h_1}P(n) + \Delta_{h_s}\cdots\Delta_{h_1}Q(n) \in \mathbb{Z} \quad \text{for 'many' } (n, h_1, \ldots, h_s) \in \iota^{-1}(\square_s(\Omega)), \tag{5.6}$$

where $P, Q \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Q})$ are partially p-periodic polynomials on $\iota^{-1}(\Omega)$. The bulk of §7 is devoted to solving (5.6) (and we eventually solve (5.6) in Proposition 7.6). To solve (5.6), we develop an approach called the *p-expansion trick*, which allows us to lift the solutions to (5.6) from $\mathbb{Z}/p$-valued polynomials to $\mathbb{Z}/p^s$-valued polynomials. In the same section, we also use the same trick to extend other results in §6 to the partially p-periodic case. We remark that §§6 and 7 are the essential deviation of our approach from the work of Green and Tao [**5**].

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Finally, in §9, we combine the algebraic preparations from §§6 and 7 with the approach of Green and Tao [5, 6] to complete the proof of Theorem 5.1. While the outline of this section is similar to the approach of Green and Tao [5], there are a couple of highly non-trivial challenges to overcome to adopt this method, including the treatment of (9.22), and Propositions 9.8 and 9.9.

6. *Algebraic properties for quadratic forms in* $\mathbb{F}_p$

6.1. *Irreducible properties for quadratic forms.*

We begin with a dichotomy on the intersection of the sets of solutions of a polynomial and a quadratic form.

LEMMA 6.1. *Let $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ and p be a prime such that $p \gg_{d,s} 1$. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form of rank at least 3. For any $P \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most s, either $|V(P) \cap V(M)| \le O_{d,s}(p^{d-2})$ or $V(M) \subseteq V(P)$.*

Proof. Denote $r = \mathrm{rank}(M)$. We first assume that M and P are homogeneous. By Lemma 4.3, it suffices to consider the special case when

$$M(n) = c_1 n_1^2 + \cdots + c_r n_r^2$$

for some $c_1, \ldots, c_r \in \mathbb{F}_p \backslash \{0\}$. Let $U := V(P) \cap V(M)$. By the pigeonhole principle, there exists $1 \le i \le r$ such that the ith coordinate of at least $(|U| - p^{d-r})/r$ many $n \in U$ is non-zero. We may assume without loss of generality that $i = 1$. Write $n = (t, y)$, where $t \in \mathbb{F}_p$ and $y \in \mathbb{F}_p^{d-1}$. Since $\deg(P) = s \ll p$, by the long division algorithm, we may write

$$P(t, y) = M(t, y)R(t, y) + N_1(y)t + N_0(y) \tag{6.1}$$

for some $R \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$, $N_1, N_0 \in \mathrm{poly}(\mathbb{F}_p^{d-1} \to \mathbb{F}_p)$ such that the degrees of R, N_1, and N_0 are at most $s-2$, $s-1$, and s, respectively. Then,

$$P(t, y) = M(t, y) = N_1(y)t + N_0(y) = 0 \quad \text{for all } (t, y) \in U. \tag{6.2}$$

Note that if $M(t, y) = N_1(y)t + N_0(y) = 0$, then we must have that $t = -N_0(y)/N_1(y)$ (provided that $N_1(y) \neq 0$). Substituting this expression back to $M(t, y) = 0$, we have that $\Delta(y)N_1(y)^{-2} = 0$, where

$$\Delta(y) := M(0, y)N_1(y)^2 + c_1 N_0^2(y).$$

So, $\Delta(y)$ is an important polynomial in the study of (6.2). We first assume that $\Delta(y)$ is not constant 0. Let U' be the set of $(t, y) \in U$ such that $N_1(y) = 0$. Then, it follows from (6.2) that $N_0(y) = 0$ and so $\Delta(y) = 0$. Since we assumed that $\Delta(y)$ is not constant 0 and $\deg(\Delta(y)) \le 2s$, by Lemma 4.10, we have that $|U'| \le O_{d,s}(p^{d-2})$.

Suppose that $(t, y) \in U \backslash U'$ and $t \neq 0$. Then, $M(t, y) = N_1(y)t + N_0(y) = 0$ by (6.2). So, $t = -N_0(y)N_1(y)^{-1}$ and so, $M(0, y) = M(t, y) - c_1 t^2 = -c_1 t^2 = -c_1 N_0^2(y)N_1(y)^{-2}$, or equivalently, $\Delta(y) = 0$. Since $\Delta(y)$ is not constant 0, by Lemma 4.10, we have that $|(U \backslash U') \cap \{(t, y)\colon t \neq \mathbf{0}\}| \le O_{d,s}(p^{d-2})$. Since we assumed that $|U \cap \{(t, y)\colon t \neq \mathbf{0}\}| \ge (|U| - p^{d-r})/r$, we deduce that $|U| \le O_{d,s}(p^{d-2})$ since $r \ge 3$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

We now consider the case that $\Delta(y) \equiv 0$. Let W be the set of $y \in \mathbb{F}_p^{d-1}$ such that $-c_1 M(0, y)$ is not a square residue in $\mathbb{F}_p$. Since $\operatorname{rank}(M(0, \cdot)) \geq 2$, by Lemma 4.13, $|W| = \frac{1}{2}p^{d-1}(1 + O(p^{-1/2}))$. Since $\Delta(y) \equiv 0$ for all $y \in W$, we have that $N_1(y) = 0$. By Lemma 4.10, $|W| \leq O_s(p^{d-2})$ unless $N_1(y) \equiv 0$. If $p \gg_{d,s} 1$, then $|W| = \frac{1}{2}p^{d-1}(1 + O(p^{-1/2})) \gg_s p^{d-2}$ and so we must have that $N_1(y) \equiv 0$. Since $\Delta(y) \equiv 0$, we also have that $N_0(y) \equiv 0$. So, (6.1) implies that $P = MR$, which implies that $V(M) \subseteq V(P)$.

Finally, we consider the general case when M and P are not necessarily homogeneous. Then, there exist homogeneous polynomials $\tilde{M}, \tilde{P} \colon \mathbb{F}_p^{d+1} \to \mathbb{F}_p$ of degrees at most 2 and s, respectively, such that $M = \tilde{M}(\cdot, 1)$ and $P = \tilde{P}(\cdot, 1)$. Then, $\tilde{M}$ is a quadratic form of rank at least 3. By the previous case, either $|V(\tilde{P}) \cap V(\tilde{M})| \leq O_{d,s}(p^{d-1})$ or $V(\tilde{M}) \subseteq V(\tilde{P})$. In the former case, since

$$\begin{aligned} |V(\tilde{P}) \cap V(\tilde{M})| &\geq \sum_{i=1}^{p-1} |V(\tilde{P}(\cdot, i)) \cap V(\tilde{M}(\cdot, i))| \\ &= (p-1)|V(\tilde{P}(\cdot, 1)) \cap V(\tilde{M}(\cdot, 1))| \text{(since } \tilde{P} \text{ and } \tilde{M} \text{ are homogeneous)} \\ &= (p-1)|V(P) \cap V(M)|, \end{aligned} \tag{6.3}$$

we have that $|V(P) \cap V(M)| \leq O_{d,s}(p^{d-2})$. In the latter case, we have that $V(\tilde{M}(\cdot, 1)) \subseteq V(\tilde{P}(\cdot, 1))$, which implies that $V(M) \subseteq V(P)$. □

Let $P \colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a polynomial of degree s for some $s < p$ and $M \colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with matrix A. For the purpose of [18], we need a concrete algorithm to divide P by M. If the upper left entry of the matrix A is non-zero (that is, M contains a non-trivial n_1^2 term), we divide P by M as follows: whenever P has a term of the form $n_1^{i_1} \cdots n_d^{i_d}$ with $i_1 \geq 2$, by subtracting a suitable multiple of M, we may convert this term to several terms of the form $n_1^{i_1'} \cdots n_d^{i_d'}$ with $i_1' < i_1$. This eventually leads to a decomposition of the form

$$P(n) = M(n)Q(n) + (n)_1 R_1((n)_{2\sim d}) + R_0((n)_{2\sim d}), \tag{6.4}$$

where for $n \in \mathbb{F}_p^d$, we use $(n)_1$ to denote the first coordinate of n and $(n)_{2\sim d}$ to denote the last $d-1$ coordinate of n. We call this decomposition the *standard long division algorithm*.

However, the above decomposition is non-valid if M does not have the n_1^2 term (one example is $M(n_1, \ldots, n_d) = n_1 n_2 + n_3^2 + \cdots + n_d^2$). In this case, one has to do a change of variables before applying the standard long algorithm. Let B be a $d \times d$ non-degenerate matrix such that the upper left entry of the matrix BAB^T is non-zero. Then, we may use the standard long division algorithm to write

$$P(nB) = M(nB)Q(n) + n_1 R_1(n_{2\sim d}) + R_0(n_{2\sim d}).$$

So,

$$P(n) = M(n)Q(nB^{-1}) + (nB^{-1})_1 R_1((nB^{-1})_{2\sim d}) + R_0((nB^{-1})_{2\sim d}), \tag{6.5}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

where for $n \in \mathbb{F}_p^d$, we use n_1 to denote the first coordinate of n and $n_{2\sim d}$ to denote the last $d-1$ coordinates of n. We call this decomposition (6.5) of P the *B-standard long division algorithm* (in particular, the standard long division algorithm in (6.4) is the I-standard long division algorithm with I being the identity matrix).

An advantage of this algorithm is that all the coefficients of Q, R_1, and R_0 can be written in the form $F/c_{1,1}^s$ with F being a polynomial function (dependent on B) with respect to the coefficients of P and M of degree at most s, where $c_{1,1}$ is the coefficient of the n_1^2 term of $M(nB)$. This simple observation will be useful in [18].

We say that *M divides P with respect to the B-standard long division algorithm* if $R_1 = R_0 = 0$ in the B-standard long division algorithm (6.5). We have the following Hilbert Nullstellensatz type result for $V(M)$.

PROPOSITION 6.2. (Hilbert Nullstellensatz for $V(M)$) *Let $d \in \mathbb{N}_+$, $s \in \mathbb{N}$, $p \gg_{d,s} 1$ be a prime number, $P \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ be of degree at most s, and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form of rank at least 3 associated with the matrix A. If $V(M) \subseteq V(P)$, then for all $d \times d$ non-degenerate matrix B with the upper left entry of the matrix BAB^T being non-zero, M divides P with respect to the B-standard long division algorithm. In particular, $P = MR$ for some $R \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most $s-2$.*

It is tempting to prove Proposition 6.2 by using Hilbert Nullstellensatz. Let $F_i(n_1, \dots, n_d) := n_i^p - n_i$ for all $1 \le i \le d$. By using Hilbert Nullstellensatz in the finite field (see [2]), if $V(M) \subseteq V(P)$, then one can show that

$$P = MR + \sum_{i=1}^{d} F_i R_i$$

for some polynomials $R, R_1, \dots, R_d$. Unfortunately, this is not helpful for Proposition 6.2, since there is no restriction on the degree of R, and it is unclear how to remove the polynomials $R_1, \dots, R_d$. So, we have to prove Proposition 6.2 with a different method relying on the long division algorithm.

Proof of Proposition 6.2. Denote $r = \mathrm{rank}(M)$. Under a change of variables, it suffices to show that for any $P \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most s and any quadratic form, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ of rank at least 3 associated with the matrix A whose upper left entry is non-zero. If $V(M) \subseteq V(P)$, then M divides P with respect to the standard long division algorithm.

We first assume that M and P are homogeneous. Write $n = (t, y)$, where $t \in \mathbb{F}_p$ and $y \in \mathbb{F}_p^{d-1}$. Since the coefficient of the t^2 term of $M(t, y)$ is non-zero, we may apply the standard long division algorithm to write

$$P(t, y) = M(t, y)R(t, y) + N_1(y)t + N_0(y) \tag{6.6}$$

for some $R, N_1, N_0 \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degrees at most $s-2$, $s-1$, and s, respectively. Our goal is to show that $N_0, N_1 \equiv 0$. By Lemma 4.3 and a change of variables, it suffices to show that if (6.6) holds for

$$M(n) = c_1 n_1^2 + \cdots + c_r n_r^2$$

for some $c_1, \dots, c_r \in \mathbb{F}_p \backslash \{0\}$, and that $V(M) \subseteq V(P)$, then $N_0, N_1 \equiv 0$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Let

$$\Delta(y) := M(0, y)N_1(y)^2 + c_1 N_0^2(y).$$

If $\Delta(y)$ is not constant 0, then by the proof of Lemma 6.1, the number of $n \in \mathbb{F}_p^d$ such that $M(n) = P(n) = 0$ is at most $O_{d,s}(p^{d-2})$. However, since $V(M) \subseteq V(P)$, by Lemma 4.12, the number of $n \in \mathbb{F}_p^d$ such that $M(n) = P(n) = 0$ is $p^{d-1}(1 + O(p^{-(r-2/2)}))$, which is a contradiction.

So, we must have that $\Delta(y) \equiv 0$. Following the proof of Lemma 6.1, we have that $N_0(y) = N_0(y) \equiv 0$ and so $P = MR$. Since the degree of R is at most $s - 2$, we are done.

We now consider the general case when M and P are not necessarily homogeneous. Then, there exist homogeneous polynomials $\tilde{M}, \tilde{P} \colon \mathbb{F}_p^{d+1} \to \mathbb{F}_p$ of degrees at most 2 and s, respectively, such that $M = \tilde{M}(\cdot, 1)$ and $P = \tilde{P}(\cdot, 1)$. Since $|V(P) \cap V(M)| = |V(M)| = p^{d-1}(1 + O(p^{-1/2}))$ by Lemma 4.12, it follows from (6.3) that $|V(\tilde{P}) \cap V(\tilde{M})| = p^d(1 + O(p^{-1/2}))$. By Lemma 6.1, we have that $V(\tilde{M}) \subseteq V(\tilde{P})$. So, by the previous case, we have that $\tilde{M}$ divides $\tilde{P}$ with respect to the standard long division algorithm. Since the n_1^2 term of M is non-zero, it is not hard to see that M also divides P with respect the standard long division algorithm. □

6.2. *A special anti-derivative property.* For a polynomial

$$Q(n_1, \dots, n_d) = \sum_{0 \le a_1, \dots, a_d \le p-1} C_{a_1,\dots,a_d} n_1^{a_1} \cdots n_d^{a_d},$$

denote the ith directional formal derivative of Q by

$$\partial_i Q(n_1, \dots, n_d) := \sum_{0 \le a_1, \dots, a_d \le p-1} a_i C_{a_1,\dots,a_d} n_1^{a_1} \cdots n_i^{a_i - 1} \cdots n_d^{a_d}.$$

In many scenarios (such as in (5.5)), we need to recover information on Q from the information on its derivatives $\partial_1 Q, \dots, \partial_d Q$. In particular, a question we need to understand is: what is a characterization of the property that Q is a constant in $V(M)$ in terms of $\partial_1 Q, \dots, \partial_d Q$? If Q is constant zero in $V(M)$, then by Proposition 6.2, we have that $Q = MQ'$ for some polynomial Q'. Then, it is not hard to see that

$$(\partial_j M)(\partial_i Q) - (\partial_i M)(\partial_j Q) = M((\partial_j M)(\partial_i Q') - (\partial_i M)(\partial_j Q')),$$

that is, $\partial_j M \partial_i Q$ is equal to $\partial_i M \partial_j Q$ on $V(M)$. We now show that the converse also holds.

PROPOSITION 6.3. *Let $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ with $d \ge 3$, p be a prime, and $M \colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form of rank at least 3. Let $Q \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ be a polynomial of degree at most s with $Q(\mathbf{0}) = 0$ such that*

$$\partial_1 M(n) \partial_i Q(n) = \partial_i M(n) \partial_1 Q(n)$$

for all $1 \le i \le d$ and $n \in V(M)$. If $p \gg_{d,s} 1$, then $Q = MQ'$ for some $Q' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most $s - 2$.

To deduce Proposition 6.3, we consider its following variation.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

PROPOSITION 6.4. *Let $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ with $d \geq 3$, p be a prime, and $M : \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form of rank at least 3. Let $H_1, \ldots, H_d \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ be polynomials of degree at most s such that*

$$\partial_i M(n) H_j(n) = \partial_j M(n) H_i(n) \tag{6.7}$$

for all $1 \leq i, j \leq d$ and $n \in V(M)$. If $p \gg_{d,s} 1$, then there exist $G, F_1, \ldots, F_d \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(G) \leq s-1$ and $\deg(F_i) \leq s-2$ such that

$$H_i = (\partial_i M)G + MF_i \quad \text{for all } 1 \leq i \leq d.$$

Proof of Proposition 6.3 assuming Proposition 6.4. Our strategy is to apply Proposition 6.4 repeatedly. For $\ell \geq 1$, we say that Property ℓ holds if there exist $W, R_1, \ldots, R_d \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(W) \leq s-2$, $\deg(R_i) \leq s - 2\ell - 1$ such that

$$\partial_i(Q - MW) = M^\ell R_i \tag{6.8}$$

for all $1 \leq i \leq d$. Setting $N_i = \partial_i Q$, it follows from Proposition 6.4 that there exist $F, R_1, \ldots, R_d \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(F) \leq s-2$, $\deg(R_i) \leq s-3$ such that

$$\partial_i Q = (\partial_i M)F + MR_i.$$

This implies that Property 1 holds since for all $1 \leq i \leq d$, we have that

$$\partial_i(Q - MF) = M(R_i - \partial_i F).$$

Now, suppose that Property-ℓ holds for some $1 \leq \ell < p$. Since $\deg(\partial_i Q') \leq p-1$, for all $1 \leq j \leq d$, it follows from (6.8) that

$$\partial_j \partial_i (Q - MW) = \partial_j (M^{\ell-1} R_i) = M^{\ell-2}((\ell-1)(\partial_j M)R_i + M\partial_j R_i). \tag{6.9}$$

By symmetry,

$$\partial_i \partial_j (Q - MW) = M^{\ell-2}((\ell-1)(\partial_i M)R_j + M\partial_i R_j). \tag{6.10}$$

Combining (6.9) and (6.10), we have that for all $1 \leq i, j \leq d$, there exists a polynomial $N_{i,j} \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ such that

$$(\partial_i M)R_j - (\partial_j M)R_i = MN_{i,j}.$$

By Proposition 6.4, we may write

$$R_i = (\partial_i M)F' + MR_i'$$

for some $F', R_1', \ldots, R_d' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(F') \leq s - 2\ell - 2$, $\deg(R_i) \leq s - 2\ell - 3$. Substituting this back to (6.8), we have that

$$\partial_i(Q - M(W + F')) = M^{\ell+1}(R_i' - \partial_i F')$$

for all $1 \leq i \leq d$. So Property $(\ell+1)$ holds.

So, we may inductively conclude that Property ℓ holds when $\ell = [(s+1)/2]$, in which case, the polynomials H_i in (6.8) are constant 0 and so all the formal derivatives of $Q - MW$ are zero for some polynomial W of degree at most $s-2$. Since $Q(\mathbf{0}) = 0$, we must have that $Q = MW$ and we are done. □

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

The rest of the section is devoted to the proof of Proposition 6.4. We first provide an outline of the proof. For convenience, assume that $M(n) = n \cdot n$. Assume first that (6.7) holds for all $n \in \mathbb{F}_p^d$ (not just for $n \in V(M)$). Then, (6.7) implies that $n_i H_j(n)$ is divisible by n_j and so $H_j(n)$ is divisible by n_j (by choosing $i \neq j$). So, we may write $H_j(n) = 2n_j G_j(n)$ for some polynomial G_j. Substituting this back to (6.7), we have that all of G_j are equal to the same G. So, $H_j(n) = 2n_j G(n) = \partial_i M(n) G(n)$ and we are done.

For the general case, our goal is to show that each H_i can be written as the sum of a multiple of $\partial_i M(n)$ (or n_i) and a multiple of M.

Step 1: factorization over n_i. By Proposition 6.2, for all $1 \leq i, j \leq d$, there exist some polynomial $N_{i,j}$ such that

$$n_i H_j(n) - n_j H_i(n) = M(n) N_{i,j}(n) \quad \text{for all } n \in \mathbb{F}_p^d. \tag{6.11}$$

In this case, $n_j H_i(n)$ is no longer divisible by n_i because of the appearance of the term $M N_{i,j}$. However, if we decompose $N_{i,j}$ as

$$N_{i,j}(n) = n_i N_{1,i,j}(n) + n_j N_{2,i,j}(n) + n_i n_j N_{3,i,j}(n),$$

we can deduce from (6.11) that $n_i(H_j - M N_{1,i,j})$ is divisible by n_j and thus $H_j - M N_{1,i,j}$ is divisible by n_j. So, there exists some polynomial $G_{i,j}$ such that

$$H_j(n) = 2n_j G_{i,j}(n) + M(n) N_{1,i,j}(n) = \partial_j M(n) G_{i,j}(n) + M(n) N_{1,i,j}(n). \tag{6.12}$$

Step 2: removing the dependence of $N_{1,i,j}$ on i. Let $E_{i,j}$ denote the left-hand side of (6.11). Then, we have that $E_{i,j} + E_{j,i} = E_{i,j} + E_{j,k} + E_{k,i} = 0$. These symmetric properties provide extra information for $N_{i,j}$, and one can use them to show that the polynomials $N_{1,i,j}$ and $G_{i,j}$ in (6.12) are independent of i. So,

$$H_j(n) = 2n_j G_j(n) + M(n) N_j(n) \tag{6.13}$$

for some polynomials G_j and N_j. See the proof for details of this step.

Step 3: removing the dependence on j. Our final step is to show that the polynomial G_j in (6.13) can be chosen to be independent of j. Substituting (6.13) back to (6.11), we have that $n_i n_j (G_i - G_j)$ is divisible by M, which implies that $G_i - G_j$ is divisible by M. We may then set $G_i = G_j$ for all $1 \leq i, j \leq d$ by absorbing the error term in $M(n) N_j(n)$.

We now provide the rigorous proof of Proposition 6.4.

Proof of Proposition 6.4. For convenience, write $n = (n_1, \ldots, n_d) \in \mathbb{F}_p^d$. By Lemma 4.3 and a change of variables, we may assume without loss of generality that

$$M(n) = c n_1^2 + n_2^2 + \cdots + n_{d'}^2 + c' n_{d'+1} - \lambda$$

for some $3 \leq d' \leq d$, and $\lambda, c, c' \in \mathbb{F}_p$ with $c \neq 0$. Let A be the (diagonal) matrix associated to M with a_i being the (i, i)th entry of A. Then, $a_1 = c \neq 0$, $a_2 = \cdots = a_{d'} = 1$, and $a_{d'+1} = \cdots = a_d = 0$. First-time readers are advised to focus on the case when $M(n) = n \cdot n$, $c = 1$, and $d' = d$, which already captures the most essential part of the proof.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Step 1: factorization over $\partial_i M$. By assumption and Proposition 6.2, we have that for all $1 \le i, j \le d$, there exists a polynomial $N_{i,j} \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ such that

$$(\partial_i M)H_j - (\partial_j M)H_i = MN_{i,j}. \tag{6.14}$$

By Proposition 6.2, $\deg(N_{i,j}) \le s-1$. For all $1 \le i, j \le d', i \ne j$, there exist unique polynomials $N_{1,i,j}, N_{2,i,j}, N_{3,i,j} \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(N_{1,i,j}), \deg(N_{2,i,j}) \le s-2$ and $\deg(N_{3,i,j}) \le s-3$ such that

$$N_{i,j} = (\partial_i M)N_{1,i,j} + (\partial_j M)N_{2,i,j} + \tfrac{1}{2}(\partial_i M)(\partial_j M)N_{3,i,j}, \tag{6.15}$$

where $N_{1,i,j}$ is independent of n_j and $N_{2,i,j}$ is independent of n_i. We could derive from this a conclusion in analogous to (6.12). However, this is unnecessary and thus, we now directly move to Step 2.

Step 2: removing the dependence of $N_{1,i,j}$ on i. We first use the property $E_{i,j} + E_{j,i} = 0$ to show some anti-symmetric properties on the indices i and j in $N_{1,i,j}$, $N_{2,i,j}$ and $N_{3,i,j}$ (i.e., (6.18)). Since (6.14) implies that

$$a_i n_i H_j(n) - a_j n_j H_i(n) = M(n)N_{i,j}(n)/2 \tag{6.16}$$

and thus

$$a_j n_j H_i(n) - a_i n_i H_j(n) = M(n)N_{j,i}(n)/2 \tag{6.17}$$

by symmetry, it follows from (6.16) and (6.17) that $N_{i,j} + N_{j,i} = 0$. So,

$$n_i(N_{1,i,j}(n) + N_{2,j,i}(n)) + n_j(N_{1,j,i}(n) + N_{2,i,j}(n)) + n_i n_j(N_{3,i,j}(n) + N_{3,j,i}(n)) = 0.$$

Since $N_{1,i,j} + N_{2,j,i}$ is independent of n_j and $N_{1,j,i}(n) + N_{2,i,j}(n)$ is independent of n_i, we have that

$$N_{1,i,j} + N_{2,j,i} = N_{3,i,j} + N_{3,j,i} = 0. \tag{6.18}$$

Combining (6.15), (6.16), and (6.18), we have that

$$a_i n_i H_j(n) - a_j n_j H_i(n) = M(n)(n_i N_{1,i,j}(n) - n_j N_{1,j,i}(n) + n_i n_j N_{3,i,j}) \tag{6.19}$$

for all distinct $1 \le i, j \le d'$.

Next, we use the property $E_{i,j} + E_{j,k} + E_{k,i} = 0$ to show that $N_{1,i,j}$ is independent of i. For all $1 \le k \le d', k \ne i, j$, (6.19) implies that

$$a_i n_k n_i H_j(n) - a_j n_j n_k H_i(n) = M(n)(n_k n_i N_{1,i,j}(n) - n_j n_k N_{1,j,i}(n) + n_i n_j n_k N_{3,i,j}). \tag{6.20}$$

By symmetry,

$$a_j n_i n_j H_k(n) - a_k n_k n_i H_j(n) = M(n)(n_i n_j N_{1,j,k}(n) - n_k n_i N_{1,k,j}(n) + n_i n_j n_k N_{3,j,k}) \tag{6.21}$$

and

$$a_k n_j n_k H_i(n) - a_i n_i n_j H_k(n) = M(n)(n_j n_k N_{1,k,i}(n) - n_i n_j N_{1,i,k}(n) + n_i n_j n_k N_{3,k,i}). \tag{6.22}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

To remove the term $N_{3,i,j}$, we consider (6.20) $\times a_k$ + (6.21) $\times a_i$ + (6.22) $\times a_j$ and get

$$n_k n_i (a_k N_{1,i,j} - a_i N_{1,k,j}) + n_i n_j (a_i N_{1,j,k} - a_j N_{1,i,k}) + n_j n_k (a_j N_{1,k,i} - a_k N_{1,j,i})$$

is divisible by $n_i n_j n_k$. In particular, $n_k n_i (a_k N_{1,i,j}(n) - a_i N_{1,k,j}(n))$ is divisible by n_j. Since $N_{1,i,j} - N_{1,k,j}$ is independent from n_j, we must have that $a_k N_{1,i,j} = a_i N_{1,k,j}$ for all distinct $1 \le i, j, k \le d'$. This implies that $N_{1,i,j} = a_i N_j$ for some $N_j \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most $s-2$ for all $1 \le i \le d'$, $i \ne j$. This completes the second step.

Step 3: removing the dependence on j. Let $H_j' = H_j - M N_j$. Substituting the conclusion that $N_{1,i,j} = a_i N_j$ back to (6.19), we have that for all $1 \le i, j \le d'$, $i \ne j$,

$$a_i n_i H_j'(n) - a_j n_j H_i'(n) = n_i n_j M(n) N_{3,i,j}. \tag{6.23}$$

Setting $j = 1$ and $i = 2$, (6.23) implies that

$$H_1'(n) = 2cn_1 G(n) = \partial_1 M(n) G(n) \tag{6.24}$$

for some polynomial G of degree at most $s-1$.

By (6.14) and (6.24), for all $1 \le j \le d$,

$$(\partial_1 M) H_j = M N_{1,j} + (\partial_j M) H_1 = M(N_{1,j} + (\partial_j M) N_1) + (\partial_j M)(\partial_1 M) G,$$

which implies that

$$2cn_1 H_j'(n) = M(n)(N_{1,j}(n) + \partial_j M(n) N_1(n)).$$

It is not hard to see that $N_{1,j}(n) + \partial_j M(n) N_1(n) = 2cn_1 F_j(n)$ for some polynomial F_j of degree at most $s-2$. Therefore,

$$H_j - (\partial_j M) G = M F_j.$$

We are done. □

6.3. *Solution to a special polynomial equation.* In this section, we study the solutions to polynomial equations of the form

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P(n) + \Delta_{h_s} \cdots \Delta_{h_1} Q(n) = 0 \quad \text{for all/many } (n, h_1, \ldots, h_s) \in \square_s(V(M)) \tag{6.25}$$

for some quadratic form $M : \mathbb{F}_p^d \to \mathbb{F}_p$, which will be used crucially later in the paper. We begin with a preparatory result.

PROPOSITION 6.5. *Let $\delta > 0$, $d, d', s \in \mathbb{N}$ with $d \ge d' \ge 4$, $p \gg_{d,s} \delta^{-O_{d,s}(1)}$ be a prime number, $P, Q \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ be polynomials of degrees at most s and $s+1$, respectively, with $Q(\mathbf{0}) = 0$, $c \in \mathbb{F}_p \backslash \{0\}$, and $H \subseteq \mathbb{F}_p^d$ with $|H| > \delta p^d$. Let $M : \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form given by*

$$M(n_1, \ldots, n_d) := cn_1^2 + n_2^2 + \cdots + n_{d'}^2.$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Suppose that for all $n \in \mathbb{F}_p^d$ and $h = (h_1, \dots, h_d) \in H$ with $M(n) = (nA) \cdot h = 0$ (where A is the matrix associated to M), we have that

$$P(n) + \sum_{i=1}^{d} h_i \partial_i Q(n) = 0. \tag{6.26}$$

Then, there exist $P', Q' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degrees at most $s-2$ and $s-1$, respectively, such that $P = MP'$ and $Q = MQ'$.

We first explain why Proportion 6.5 is useful to study (6.25). For convenience, assume that $s = 1$, M is homogeneous, $\deg(P) = k$, and $\deg(Q) = k+1$. Intuitively, if (6.25) holds for all $(n, h) \in \square_1(V(M))$, then by the spirit of Hilbert Nullstenllensatz, we may expect to write $P(n) + \Delta_h Q(n)$ as a linear combination of $M(n)$ and $M(n+h) - M(n)$. Considering h as a fixed vector, the leading terms of $P(n) + \Delta_h Q(n)$ must be a linear combination of the leading terms of $M(n)$ and $M(n+h) - M(n)$. It is not hard to see that the leading terms of $P(n) + \Delta_h Q(n)$ are

$$P + h \cdot \nabla Q = P + \sum_{i=1}^{d} h_i \partial_i Q, \tag{6.27}$$

where $h = (h_1, \dots, h_d)$ and $\nabla Q = (\partial_1 Q, \dots, \partial_d Q)$. However, the leading terms of $M(n)$ and $M(n+h) - M(n)$ are $M(n)$ and $(nA) \cdot h$, respectively. Therefore, we have that (6.27) is zero if $M(n) = (nA) \cdot h = 0$. This is where Proposition 6.5 becomes useful.

We now explain the outline of the proof of Proposition 6.5. Fix $n \in \mathbb{F}_p^d$. The lower bound $|H| > \delta p^d$ ensures that (6.26) holds for many h with $(nA) \cdot h = 0$. By linearity, we must have that (6.26) holds for all $h \in \mathbb{F}_p^d$ with $(nA) \cdot h = 0$. This will force $P(n) = 0$ and $(\partial_1 Q(n), \dots, \partial_d Q(n))$ to be parallel to $2nA = (\partial_1 M(n), \dots, \partial_d M(n))$. So, $\partial_i Q(n)/\partial_i M(n)$ is independent of the choice of i. If this is the case for many $n \in V(M)$, then it follows from Lemma 6.1 and Proposition 6.3 that P and Q are divisible by M.

Proof of Proposition 6.5. Throughout the proof, all the constants will depend implicitly on d and s. Since $d' \geq 3$, by Lemma 4.12, the set of $h \in \mathbb{F}_p^d$ with $M(h) = 0$ is of cardinality $O_{d,s}(p^{d-1})$. Passing to a subset if necessary, we may assume without loss of generality that $M(h) \neq 0$ for all $h \in H$. Let A be the (diagonal) matrix associated to M and a_i be the (i,i)th entry of A. To carry out the strategy described above, our first task is to identify those $n \in V(M)$ for which there are many $h \in \mathbb{F}_p^d$ with $(nA) \cdot h = 0$. Denote

$$V_h := \{n \in \mathbb{F}_p^d \colon (nA) \cdot h = 0\} = (\mathrm{span}_{\mathbb{F}_p}\{h\})^{\perp M}.$$

Note that for most of the choices of h, the set V_h is of cardinality $\gg p^{d-2}$. So, there are in total $\gg_\delta p^{2d-2}$ number of pairs of (n, h) for which (6.26) holds. Since $V(M)$ is of cardinality $\gg p^{d-1}$, on average, there are $\gg_\delta p^{d-1}$ many 'good' $n \in V(M)$ for which (6.26) holds for at least $\gg_\delta p^{d-1}$ many $h \in H$. Our first goal is to identify those 'good' n (as well as those h for which (6.26) holds for the pair (n, h)).

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Let Z be the set of $n \in \mathbb{F}_p^d$ with $nA = \mathbf{0}$. Then, $|Z| = p^{d-d'} \leq p^{d-4}$. We say that $(n, h) \in \mathbb{F}_p^d \times \mathbb{F}_p^d$ is a *good pair* if

$$(nA) \cdot h = M(n) = 0, \quad n \in \mathbb{F}_p^d \backslash Z \quad \text{and} \quad h \in H.$$

For each $h \in H$, V_h is a subspace of $\mathbb{F}_p^d$ of co-dimension 1 and $V_h \cap (V_h)^{\perp M} = \{\mathbf{0}\}$ (since $M(h) \neq 0$). By Proposition 4.8, $\operatorname{rank}(M|_{V_h}) \geq d' - 1 \geq 3$. By Corollary 4.14 and the fact that $|Z| \leq p^{d-4}$, there exist $p^{d-2}(1 + O(p^{-1/2}))$ many $n \in \mathbb{F}_p^d$ such that (n, h) is a good pair. So, there are in total at least $\delta p^{2d-2}(1 + O(p^{-1/2}))$ many good pairs. However, for each $n \in \mathbb{F}_p^d \backslash Z$ with $M(n) = 0$, since $nA \neq \mathbf{0}$, there exist p^{d-1} many $h \in \mathbb{F}_p^d$ such that $(nA) \cdot h = 0$. By Lemma 6.1, the number of $n \in \mathbb{F}_p^d \backslash Z$ with $M(n) = 0$ is $p^{d-1}(1 + O(p^{-1/2}))$. From this, it is easy to see that there exists a subset $X \subseteq \mathbb{F}_p^d \backslash Z$ with

$$|X| > \frac{\delta}{2} p^{d-1}(1 + O(p^{-1/2}))$$

such that $M(n) = 0$ for all $n \in X$, and there exists a subset $H_n \subseteq H$ with

$$|H_n| > \frac{\delta}{2} p^{d-1}(1 + O(p^{-1/2}))$$

such that (n, h) is a good pair for all $h \in H_n$.

Our next step is to show that for any $n \in X$, (6.26) holds not only for $h \in H_n$, but also for all $h \in V_n$. To see this, we may view the left-hand side of (6.26) as the dot product $(P(n), \partial_1 Q(n), \ldots, \partial_d Q(n)) \cdot (1, h)$. If the dot product is zero for all $h \in H_n$, then by linearity and the largeness of H_n (this is where we need the lower bound $|H| > \delta p^d$ to ensure the largeness of H_n), $(P(n), \partial_1 Q(n), \ldots, \partial_d Q(n))$ is orthogonal to W_n, the subspace of $\mathbb{F}_p^{d+1}$ spanned by $(1, h)$, $h \in H_n$.

We claim that $W_n = \mathbb{F}_p \times V_n$. Clearly, $W_n \subseteq \mathbb{F}_p \times V_n$, so it suffices to show that W_n is of dimension d. Suppose in contrast that W_n is of dimension at most $d - 1$. Let $v_i = (v_{i,0}, \ldots, v_{i,d}) \in \mathbb{F}_p^{d+1}$, $1 \leq i \leq L$ be a basis of W_n. If $(1, h) = a_1 v_1 + \cdots + a_L v_L$, we must have that $a_1 v_{i,0} + \cdots + a_L v_{L,0} = 1$. So, there are at most $p^{L-1} \leq p^{d-2}$ many $(1, h)$ belonging to W_n, meaning that $|H_n| \leq p^{d-2}$, which is a contradiction since $|H_n| > (\delta/2) p^{d-1}(1 + O(p^{-1/2}))$ and $p \gg \delta^{-O(1)}$. So, $W_n = \mathbb{F}_p \times V_n$. This proves the claim.

We are now ready to complete the proof of the proposition. Since $P(n) + \sum_{i=1}^d h_i Q_i(n) = 0$ for all $n \in X$ and $h \in H_n$, we have that

$$(P(n), Q_1(n), \ldots, Q_d(n)) \cdot (1, h) = 0$$

for all $h \in H_n$. By linearity,

$$(P(n), Q_1(n), \ldots, Q_d(n)) \cdot v = 0$$

for all $v \in W_n$. Since by the claim $W_n = \mathbb{F}_p \times V_n$, we have that $P(n) = 0$ and that

$$(nA) \cdot h = 0 \Rightarrow (Q_1(n), \ldots, Q_d(n)) \cdot h = 0$$

for all $h \in \mathbb{F}_p^d$ and $n \in X$. This implies that the vector nA is parallel to $(Q_1(n), \ldots, Q_d(n))$, and so we may write $(Q_1(n), \ldots, Q_d(n)) = F(n)(nA) = (cn_1 F(n), n_2 F(n), \ldots,$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$n_d F(n))$ for some $F(n) \in \mathbb{F}_p$ for all $n \in X$. In other words, for all $1 \le i, j \le d$ and $n \in X$,

$$\partial_j M(n) Q_i(n) = 2c_j a_j Q_i(n) = 2c_j c_i a_j a_i = \partial_i M(n) Q_j(n),$$

where $c_1 = c$ and $c_2 = \cdots = c_d = 1$.

Finally, since $|X| > (\delta/2)p^{d-1}(1 + O(p^{-1/2}))$, by Lemma 6.1, for all $n \in \mathbb{F}_p^d$, $M(n) = 0$ implies that $P(n) = 0$ and $\partial_j M(n) Q_i(n) = \partial_i M(n) Q_j(n)$ for all $1 \le i, j \le d$. The conclusion then follows from Proposition 6.3. □

We are now able provide a solution to (6.25).

PROPOSITION 6.6. *Let $d, s \in \mathbb{N}_+$, $k \in \mathbb{N}$, $\delta > 0$, $p \gg_{d,k} \delta^{-O_{d,k}(1)}$ be a prime, and $M : \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form. Let W be a subset of $\square_s(V(M))$ such that either:*

- *$W = \square_s(V(M))$ and $\operatorname{rank}(M) \ge s + 3$; or*
- *$|W| \ge \delta |\square_s(V(M))|$ and $\operatorname{rank}(M) \ge s^2 + s + 3$.*

Let $P, Q \in \operatorname{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P) \le k - 1$ and $\deg(Q) \le k$ be such that for all $(n, h_1, \dots, h_s) \in W$, we have that

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P(n) + \Delta_{h_s} \cdots \Delta_{h_1} Q(n) = 0 \tag{6.28}$$

(where $\Delta_{h_{s-1}} \cdots \Delta_{h_1} P(n)$ is understood as $P(n)$ when $s = 1$). Then,

$$P = M P_1 + P_2 \quad \text{and} \quad Q = M Q_1 + Q_2$$

for some $P_1, P_2, Q_1, Q_2 \in \operatorname{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P_1) \le k - 3$, $\deg(P_2) \le s - 2$, $\deg(Q_1) \le k - 2$, $\deg(Q_2) \le s - 1$.

Therefore, Proposition 6.6 provides solutions to (6.25) in two cases. The first case is when (6.25) holds for all $(n, h_1, \dots, h_s) \in \square_s(V(M))$, the second case is when (6.25) holds for a dense subset of $(n, h_1, \dots, h_s) \in \square_s(V(M))$ (in the latter case, we need to impose a stronger lower bound for the dimension d).

We remark that Proposition 6.6 is trivial if $k \le s - 1$. We briefly explain the outline of the proof. Assume for convenience that $W = \square_s(V(M))$. We prove it by induction.

Step 1. Fix $h_s \in \mathbb{F}_p^d$ and write $h = h_s$ for convenience. Note that

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P + \Delta_h \cdots \Delta_{h_1} Q = \Delta_{h_{s-1}} \cdots \Delta_{h_1} (P + \Delta_h Q)$$

and that $(n, h_1, \dots, h_{s-1}, h) \in \square_s(V(M)) \Leftrightarrow (n, h_1, \dots, h_{s-1}) \in \square_{s-1}(V(M)^h)$. By the inheriting principle, we may informally think of $V(M)^h$ as a lower degree hypersphere (which is the set of zeros of a lower dimensional quadratic form M_h and a subspace V_h of $\mathbb{F}_p^d$) and apply the induction hypothesis to conclude that the restriction of $P_h + \Delta_h Q_h$ on V_h can be written as $M_h R_h + R_h'$ for some lower degree polynomial R_h', where P_h and Q_h are the 'restrictions' of P and Q on V_h (this can be made precise by (6.31)).

Step 2. For simplicity, assume that all of M, P, Q are homogeneous. Then, R_h' vanishes and so the conclusion of Step 1 implies that $P(n) + \Delta_h Q(n) = 0$ for all $n \in V(M)^h$, that is, whenever $(n, h) \in \square_1(V(M))$. Since P, Q are homogeneous, similar to (6.27), the

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

degree $k-1$ terms of $P(n)+Q(n+h)-Q(n)$ are equal to

$$P+h\cdot\nabla Q = P+h_s\cdot\nabla Q = P+\sum_{i=1}^{d} h_{s,i}\partial_i Q,$$

where $h=h_s=(h_{s,1},\dots,h_{s,d})$. Since (6.28) holds for all $(n,h_1)\in\square_1(V(M))$, we may invoke Proposition 6.5 to the degree $k-1$ terms of the left-hand side of (6.28) to conclude that P and Q are divisible by M.

Proof of Proposition 6.6. Throughout the proof, we assume that $p\gg_{d,k}\delta^{-O_{d,k}(1)}$. Let A be the matrix associated with M and denote $d'=\operatorname{rank}(M)$. By Lemmas 4.3 and 4.18, under a change of variables, it suffices to consider the case when A is a diagonal matrix with diagonal $(c,1,\dots,1,0,\dots,0)$ for some $c\in\mathbb{F}_p\backslash\{0\}$, where there are $d-d'$ zeros on the diagonal.

To carry out Step 1 of the outline, it is convenient to work on the set $V(M)^h-h/2$ instead of $V(M)^h$. So, we need some definitions to describe such a set (this will make the notation of the proof slightly different from that used in the outline). For $h\in\mathbb{F}_p^d$, let

$$U_h := V(M)^h-h/2=\{n\in\mathbb{F}_p^d\colon M(n+h/2)=M(n-h/2)\}.$$

Then, we may write $U_h=V_h+u_h$ for some $u_h\in\mathbb{F}_p^d$, where

$$V_h:=\{n\in\mathbb{F}_p^d\colon (nA)\cdot h=0\}.$$

Let $L_h\colon\mathbb{F}_p^{d-1}\to V_h$ be any bijective linear transformation and denote

$$M_h(m):=M(L_h(m)+u_h-h/2).$$

By the definition of U_h, we have that $M(L_h(m)+u_h-h/2)=M(L_h(m)+u_h+h/2)$ for all $m\in\mathbb{F}_p^{d-1}$. From this, it is not hard to check that

$$\begin{aligned}(m,m_1,\dots,m_{s-1})\in\square_{s-1}(M_h)\Leftrightarrow(L_h(m)\\ +u_h-h/2,L_h(m_1),\dots,L_h(m_{s-1}),h)\in\square_s(M).\end{aligned}\tag{6.29}$$

So, the assumption and (6.29) imply that

$$\Delta_{m_{s-1}}\cdots\Delta_{m_1}(P_h(m)+Q_h(m))=0\tag{6.30}$$

for all $(m,m_1,\dots,m_{s-1})\in\square_{s-1}(M_h)$ with $(L_h(m)+u_h-h/2,L_h(m_1),\dots,L_h(m_{s-1}),h)\in W$, where

$$P_h:=P(L_h(\cdot)+u_h-h/2)\quad\text{and}\quad Q_h:=Q(L_h(\cdot)+u_h+h/2)-Q(L_h(\cdot)+u_h-h/2).$$

We remark that (6.30) is also valid when $s=1$, in which case, (6.30) is understood as

$$P_h(m)+Q_h(m)=0.\tag{6.31}$$

We introduce some further notation to describe the set of $(m,m_1,\dots,m_{s-1})$ satisfying (6.29). For all $h\in\mathbb{F}_p^d$, let W_h be the set of $(n,h_1,\dots,h_{s-1})\in(\mathbb{F}_p^d)^s$ such that $(n,h_1,\dots,h_{s-1},h)\in W$ (that is, we decompose W into fibers W_h with respect to the last

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

variable). We now parameterize W_h using $m, m_1, \dots, m_{s-1}$. For all $h \in \mathbb{F}_p^d$, let W_h' be the set of $(m, m_1, \dots, m_{s-1})$ such that

$$(L_h(m) + u_h - h/2, L_h(m_1), \dots, L_h(m_{s-1})) \in W_h.$$

Since $(n, h_1, \dots, h_s) \in \square_s(V(M))$ for all $(n, h_1, \dots, h_{s-1}) \in W_{h_s}$, it follows from (6.29) that

$$W_h' \subseteq \square_{s-1}(V(M_h)) \quad \text{and} \quad |W_h'| = |W_h| \quad \text{for all } h \in \mathbb{F}_p^d. \tag{6.32}$$

Therefore, we have that (6.30) holds for all $h \in \mathbb{F}_p^d$ and $(m, m_1, \dots, m_{s-1}) \in W_h'$.

Recall that Step 1 is to fix h and use the induction hypothesis or Proposition 6.2 on (6.30) to obtain a description for $P_h + Q_h$. So, we need to choose h carefully so that the dimension of V_h is $d-1$, the rank of M_h is $d'-1$, and the set W_h' is large. Let Z be the set of $h \in \mathbb{F}_p^d$ such that $\dim(V_h) \neq d-1$ or $\dim(V_h \cap V_h^{\perp M}) \neq d-d'$. Then, by Proposition 4.8(ii), for all $h \in \mathbb{F}_p^d \backslash Z$, we have that $\operatorname{rank}(M_h) = d'-1$. We now show that Z is a small set.

CLAIM 1. *We have that* $|Z| \le O_d(p^{d-1})$.

Proof of Claim 1. Let Z' be the set of $h \in \mathbb{F}_p^d$ such that either $(hA) \cdot h = 0$ or at least one of the coordinates of h is 0. Since $\operatorname{rank}(A) = d' \ge 3$, by Lemma 4.12, it is not hard to see that $|Z'| = O_d(p^{d-1})$. Now, fix any $h = (h_1, \dots, h_d) \in \mathbb{F}_p^d \backslash Z'$. Then, since $hA \neq \mathbf{0}$, we have that $\dim(V_h) = d-1$. Note that

$$\{-h_i e_1 + c h_1 e_i : 2 \le i \le d'\} \cup \{e_j : d'+1 \le j \le d\}$$

is a basis of V_h, where e_i is the ith standard unit vector. From this, it is not hard to compute that

$$(h_1, \dots, h_{d'}, 0, \dots, 0), e_{d'+1}, \dots, e_d$$

is a basis of $(V_h)^{\perp M}$. Therefore, $\dim((V_h)^{\perp M}) = d - d' + 1$. Finally, it is not hard to compute by induction that

$$\det \begin{bmatrix} -h_2 & ch_1 & & & \\ -h_3 & & ch_1 & & \\ \cdots & & & \cdots & \\ -h_{d'} & & & & ch_1 \\ h_1 & \cdots & \cdots & \cdots & h_{d'} \end{bmatrix} = (-1)^{d'+1}(ch_1)^{d'-2}((hA) \cdot h) \neq 0.$$

This implies that $V_h + (V_h)^{\perp M} = \mathbb{F}_p^d$. So,

$$\begin{aligned} \dim(V_h \cap V_h^{\perp M}) &= \dim(V_h) + \dim(V_h^{\perp M}) - \dim(V_h + (V_h)^{\perp M}) \\ &= (d-1) + (d-d'+1) - d = d - d'. \end{aligned}$$

Therefore, $Z \subseteq Z'$ and thus, $|Z| \le O_d(p^{d-1})$. This proves Claim 1. □

We still need to remove those h for which $|W_h'|$ is small. This is done by the next claim.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

CLAIM 2. *There exists $W_* \subseteq \mathbb{F}_p^d$ of cardinality at least $\delta p^d/4$ such that for all $h \in W_* \backslash Z$,*

$$|W_h'| = |W_h| \geq O(\delta)|\square_{s-1}(V(M_h))|. \tag{6.33}$$

Proof of Claim 2. If $|W| \geq \delta|\square_s(V(M))|$ and $\operatorname{rank}(M) \geq s^2 + s + 3$, then by Proposition 4.19, we have that

$$\sum_{h \in \mathbb{F}_p^d} |W_h| = |W| > \delta|\square_s(V(M))| = \delta p^{(s+1)d-((s(s+1)/2)+1)}(1 + O_s(p^{-1/2})).$$

Moreover, for all but at most $s(s+1)p^{d+s-\operatorname{rank}(M)}$ many $h_s \in \mathbb{F}_p^d$, we have that $|W_{h_s}| \leq p^{ds-((s(s+1)/2)+1)}$. Also, the sum of $|W_{h_s}|$ over those at most $s(s+1)p^{d+s-d'}$ exceptionally $h_s \in \mathbb{F}_p^d$ is at most $s(s+1)p^{(s+1)d+s-d'}$. It then follows from the pigeonhole principle that there exists $W_* \subseteq \mathbb{F}_p^d \backslash Z$ of cardinality at least $\delta p^d/4$ such that $|W_h| \geq \delta p^{ds-((s(s+1)/2)+1)}/4$ for all $h_s \in W_*$. Since $\operatorname{rank}(M_h) \geq (s-1)^2 + (s-1) + 3$, by Proposition 4.19 and (6.32), we have that

$$|W_h'| = |W_h| \geq \delta p^{ds-((s(s+1)/2)+1)}/4 = \delta p^{(d-1)s-((s(s-1)/2)+1)}/4 = O(\delta)|\square_{s-1}(V(M_h))|$$

for all $h \in W_* \backslash Z$.

However, if $W = \square_s(V(M))$ and $\operatorname{rank}(M) \geq s + 3$, then for all $h \in \mathbb{F}_p^d \backslash Z$, it follows from (6.29) that $W_h' = \square_{s-1}(V(M_h))$. This completes the proof of Claim 2. □

Step 2 is summarized by the following claim.

CLAIM 3. *Let $t \in \mathbb{N}$. Suppose that for all $h \in W_* \backslash Z$, we may write*

$$P_h + Q_h = M_h R_h + R_h' \tag{6.34}$$

for some $R_h, R_h' \in \text{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(R_h) \leq k-3$ and $\deg(R_h') \leq t-1$. Then,

$$P = MP' + P'' \quad \text{and} \quad Q = MQ' + Q''$$

for some $P', P'', Q', Q'' \in \text{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P') \leq k-3$, $\deg(P'') \leq t-1$, $\deg(Q') \leq k-2$, and $\deg(Q'') \leq t$.

We first show that Claim 3 implies the desired conclusion by following the two steps described in the outline. We prove by induction on s. If $s = 1$, then for all $h \in W_* \backslash Z$, since $\operatorname{rank}(M_h) \geq 3$ and (6.31) holds for all $(m, m_1, \ldots, m_{s-1}) \in W_h'$, it follows from Proposition 6.2 that

$$P_h + Q_h = M_h R_h$$

for some $R_h \in \text{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(R_h) \leq k-3$. By Claim 3 (setting $t = 0$),

$$P = MP_1 \quad \text{and} \quad Q = MQ_1 + C$$

for some $C \in \mathbb{F}_p$ and $P_1, Q_1 \in \text{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P_1) \leq k-3$, $\deg(Q_1) \leq k-2$.

Suppose we have proved Proposition 6.6 for $s-1$ for some $s \geq 2$. We now prove it for s. For all $h \in W_* \backslash Z$, since $\operatorname{rank}(M_h) = \operatorname{rank}(M) - 1$, $\deg(P_h + \Delta_h Q_h) \leq k-1$, and since

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

(6.30) holds for all $(m, m_1, \ldots, m_{s-1}) \in W_h'$, by the induction hypothesis, we may write

$$P_h + Q_h = M_h R_h + R_h'$$

for some $R_h, R_h' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(R_h) \le k-3$ and $\deg(R_h') \le s-2$ for all $h \in W_* \backslash Z$. By Claim 3 (setting $t = s+1$),

$$P = MP' + P'', \quad Q = MQ' + Q''$$

for some $P', P'', Q', Q'' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P') \le k-3$, $\deg(P'') \le s-2$, $\deg(Q') \le k-2$, and $\deg(Q'') \le s-1$. This completes the induction step and we are done.

It remains to prove Claim 3.

Proof of Claim 3. The idea is to apply Proposition 6.5 to the leading terms of the polynomials. Unlike the special case discussed in the outline, here, the polynomials M, P, and Q are not homogeneous. To overcome this difficulty, each time we apply Proposition 6.5 to the leading terms, we may factor out a multiple of M to lower the degrees of P and Q. Then, we apply Proposition 6.5 again to the leading terms of the new P and Q, until $k \le s-1$. So, we prove Claim 3 by induction on k.

If $\max\{\deg(P), \deg(Q)-1\} \le s-1$, then Claim 3 follows by setting $P = P''$, $Q = Q''$, and $P' = Q' = 0$. Assume now that Claim 3 holds when $\max\{\deg(P), \deg(Q)-1\} \le k'-1$ for some $s \le k' \le k-1$ and we prove that Claim 3 holds when $\max\{\deg(P), \deg(Q)-1\} \le k'$. Note that (6.34) implies that

$$\begin{aligned} &P(L_h(m) + u_h - h/2) + Q(L_h(m) + u_h + h/2) - Q(L_h(m) + u_h - h/2) \\ &\quad = M(L_h(m) + u_h - h/2) R_h(m) + R_h'(m) \end{aligned} \tag{6.35}$$

for all $h \in W_* \backslash Z$. Since $\deg(R_h') \le k'-1$, it is not hard to see that $\deg(R_h) \le k'-2$. Let $\tilde{P}, \tilde{Q}, \tilde{M}, \tilde{R}_h$ be the degree $k', k'+1, 2, k'-2$ terms of P, Q, M, R_h, respectively. Write $h = (h_1, \ldots, h_d)$. Comparing the degree k' terms of (6.35) and noticing that the degree k' terms of $Q(L_h(m) + u_h + h/2) - Q(L_h(m) + u_h - h/2)$ are $\sum_{i=1}^d h_i \partial_i \tilde{Q}(L_h(m))$, we have that

$$\tilde{P}(L_h(m)) + \sum_{i=1}^d h_i \partial_i \tilde{Q}(L_h(m)) = \tilde{M}(L_h(m)) \tilde{R}_h(m).$$

Therefore, for all $n \in \mathbb{F}_p^d$ and $h \in W_* \backslash Z$ with $(nA) \cdot h = M(n) = 0$, there exists $m \in \mathbb{F}_p^{d-1}$ with $L_h(m) = n$ and thus,

$$\tilde{P}(n) + \sum_{i=1}^d h_i \partial_i \tilde{Q}(n) = \tilde{M}(n) \tilde{R}_h(m) = 0.$$

Since $|Z| \le O_d(p^{d-1})$ by Claim 1, we have that $|W_* \backslash Z| \ge p^d/8$. Since $\mathrm{rank}(M) \ge s+3 \ge 4$, by Proposition 6.5, we have that $\tilde{P} = \tilde{M} P'$ and $\tilde{Q} = \tilde{M} Q'$ for some $P', Q' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P') \le k'-2 \le k-3$ and $\deg(Q') \le k'-1 \le k-2$. Since $M - \tilde{M}$ is of degree at most 1, we may write

$$P = MP' + P'', \quad Q = MQ' + Q''$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

for some $P'', Q'' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P'') \le k'-1$ and $\deg(Q'') \le k'$. Writing $P_h'' := P''(L_h(\cdot)+u_h-h/2)$ and $Q_h'' := Q''(L_h(\cdot)+u_h+h/2) - Q''(L_h(\cdot)+u_h-h/2)$, it follows from (6.34) and the identity

$$M_h = M(L_h(\cdot)+u_h-h/2) = M(L_h(\cdot)+u_h+h/2)$$

that for all $h \in W_*\backslash Z$, we may write

$$P_h'' + Q_h'' = M_h R_h'' + R_h''' \tag{6.36}$$

for some $R_h'', R_h'' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(R_h'') \le k-3$ and $\deg(R_h''') \le s-1$. Since $\max\{\deg(P''), \deg(Q'')-1\} \le k'-1$, by induction hypothesis, (6.36) implies that

$$P'' = MP''' + P'''', \quad Q'' = MQ''' + Q''''$$

for some $P''', P'''', Q''', Q'''' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(P''') \le k-3$, $\deg(P'''') \le s-1$, $\deg(Q''') \le k-2$, and $\deg(Q'''') \le s$. Then,

$$P = M(P'+P''') + P'''' \quad \text{and} \quad Q = M(Q'+Q''') + Q''''.$$

This completes the induction step and proves Claim 3. □

This therefore completes the proof of Proposition 6.6. □

6.4. *Intrinsic definitions for polynomials on $V(M)$.* Let $M: \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form. Recall that we define a function $f: V(M) \to \mathbb{F}_p$ to be a polynomial if f is the restriction to $V(M)$ of some polynomial $f': \mathbb{F}_p^d \to \mathbb{F}_p$ defined in $\mathbb{F}_p^d$. It is natural to ask if there is an intrinsic way to define polynomials on $V(M)$ (that is, without looking at points outside $V(M)$). A natural way to do so is to use anti-derivative properties of polynomials. Let $g \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$. It is not hard to see that g is of degree at most $s-1$ if and only if $\Delta_{h_s}\cdots\Delta_{h_1} g(n) = 0$ for all $n, h_1, \ldots, h_s \in \mathbb{F}_p^d$. This observation provides us with a promising alternative way to define polynomials and leads to the following question.

Conjecture 6.7. (Intrinsic definition for polynomials) Let $g: \mathbb{F}_p^d \to \mathbb{F}_p$ be a function, $d, s \in \mathbb{N}_+$, and $p \gg_d 1$ be a prime. Let $M: \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form with $\mathrm{rank}(M) \gg_s 1$. Then, the following are equivalent:

(i) $\Delta_{h_s}\cdots\Delta_{h_1} g(n) = 0$ for all $(n, h_1, \ldots, h_s) \in \square_s(V(M))$;

(ii) there exists a polynomial $g' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most $s-1$ such that $g(n) = g'(n)$ for all $n \in V(M)$.

If Conjecture 6.7 holds, then we may use part (i) of Conjecture 6.7 as an intrinsic definition for polynomials on $V(M)$. In this paper, we were unable to prove Conjecture 6.7. Instead, we prove the following special case of Conjecture 6.7, which is good enough for the purpose of this paper.

PROPOSITION 6.8. *Let $d, s \in \mathbb{N}_+$, $p \gg_d 1$ be a prime, and $M: \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A of rank at least $s+3$. Then, for any $g \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(g) \le s$, the following are equivalent:*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

(i) *for all* $(n, h_1, \dots, h_s) \in \square_s(V(M))$, *we have that* $\Delta_{h_s} \cdots \Delta_{h_1} g(n) = 0$;
(ii) *we have*

$$g = Mg_1 + g_2$$

for some $g_1, g_2 \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *with* $\deg(g_1) \le s-2$, $\deg(g_2) \le s-1$;
(i) *we have*

$$g(n) = ((nA)\cdot n)g_1(n) + g_2(n)$$

for some $g_1, g_2 \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *with* $\deg(g_1) \le s-2$, $\deg(g_2) \le s-1$.

Proof. We first show that (ii) $\Rightarrow$ (i). For all $(n, h_1, \dots, h_s) \in \square_s(V(M))$, note that

$$\Delta_{h_s} \cdots \Delta_{h_1} g(n) = \Delta_{h_s} \cdots \Delta_{h_1}(M(n)g_1(n)).$$

Since $M(n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s) = 0$ for all $\epsilon_1, \dots, \epsilon_s \in \{0, 1\}$, we have $\Delta_{h_s} \cdots \Delta_{h_1}(M(n)g_1(n)) = 0$ and we are done.

The part (i) $\Rightarrow$ (ii) follows from Proposition 6.6 (by setting $k = s$, $P \equiv 0$, and $Q = g$). Finally, (ii) $\Leftrightarrow$ (iii), since the degree of the polynomial $M(n) - (nA)\cdot n$ is at most 1. □

As a corollary of Proposition 6.8, we have the following.

COROLLARY 6.9. *Let* $d, k, s \in \mathbb{N}_+$, $p \gg_d 1$ *be a prime, and* $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ *be a non-degenerate quadratic form associated with the matrix A. Let* $m_1, \dots, m_k \in \mathbb{F}_p^d$ *be linearly independent and non-M-isotropic vectors. If* $d \ge k + s + 3$, *then for any polynomial* $g \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *of degree at most s, the following are equivalent:*

(i) *for all* $(n, h_1, \dots, h_s) \in \square_s(V(M)^{m_1,\dots,m_k})$, *we have that* $\Delta_{h_s} \cdots \Delta_{h_1} g(n) = 0$;
(ii) *we have*

$$g(n) = M(n)g_0(n) + \sum_{i=1}^{k}(M(n+m_i) - M(n))g_i(n) + g'(n)$$

for some homogeneous $g_0, \dots, g_k \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *with* $\deg(g_0) = s-2$ *and* $\deg(g_1) = \cdots = \deg(g_k) = s-1$, *and some* $g' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *of degree at most* $s-1$ *(recall that* $g_0, \dots, g_k$ *are allowed to be constant 0 by Convention 2.2);*
(iii) *we have*

$$g(n) = ((nA)\cdot n)g_0(n) + \sum_{i=1}^{k}((nA)\cdot m_i)g_i(n) + g'(n)$$

for some homogeneous $g_0, \dots, g_k \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *with* $\deg(g_0) = s-2$ *and* $\deg(g_1) = \cdots = \deg(g_k) = s-1$, *and some* $g' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ *of degree at most* $s-1$.

Proof. Let V be the span of $m_1, \dots, m_k$. Since $m_1, \dots, m_k$ are linearly independent and non-M-isotropic, and since A is invertible, we have that $V^{\perp M}$ is a subspace of $\mathbb{F}_p^d$ of dimension $d-k$ with $V \cap V^{\perp M} = \{\mathbf{0}\}$. Let

$$W := \{n \in \mathbb{F}_p^d \colon 2(nA)\cdot m_i + (m_i A)\cdot m_i = 0,\ 1 \le i \le k\}.$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Since $m_1, \dots, m_k$ are linearly independent and A is invertible, we may write $W = V^{\perp M} + w$ for some $w \in \mathbb{F}_p^d$.

Let $\phi\colon \mathbb{F}_p^{d-k} \to V^{\perp M}$ be any bijective linear transformation. Denote $M'(n') := M(\phi(n') + w)$ and $g'(n') := g(\phi(n') + w)$ for all $n' \in \mathbb{F}_p^{d-k}$. By Proposition 4.8(ii), M' is a non-degenerate quadratic form and thus, $\operatorname{rank}(M') = d - k \geq s + 3$. By Lemma 4.18, we have that $(n, h_1, \dots, h_s) \in \square_s(V(M)^{m_1,\dots,m_k})$ if and only if there exists $(n', h_1', \dots, h_s') \in \square_s(V(M'))$ with $(n, h_1, \dots, h_s) = (\phi(n') + w, \phi(h_1'), \dots, \phi(h_s'))$.

We may then apply Proposition 6.8 to the polynomial g' and the quadratic form M', and then apply the change of variables ϕ to get to the conclusions. We leave the details to the interested readers. □

6.5. *A parallel matrix proposition.* Let A and B be $\mathbb{F}_p$-valued $d \times d$ matrices. If B is a multiple of A, then clearly $nB = \mathbf{0}$ implies that $nA = \mathbf{0}$. It is natural to ask if the converse is true, namely, if for many $n \in \mathbb{F}_p^d$, $nB = \mathbf{0}$ implies that $nA = \mathbf{0}$, then is it true that B must be a multiple of A. The next proposition answers this question.

PROPOSITION 6.10. *Let $d \in \mathbb{N}_+$, $\delta > 0$, $p \gg_d \delta^{-O_d(1)}$ be a prime, and A be a $\mathbb{F}_p$-valued $d \times d$ matrix of rank at least 3. Let B be an $\mathbb{F}_p$-valued $d \times d$ matrix and $v \in \mathbb{F}_p^d$. Let W be a subset of $\mathbb{F}_p^d$ of cardinality at least δp^d. Suppose that for all $w \in W$ and $n \in \mathbb{F}_p^d$ with $(nA) \cdot w = 0$, we have that*

$$(nB + v) \cdot w = 0.$$

Then, $v = \mathbf{0}$ and $B = cA$ for some $c \in \mathbb{F}_p$.

Proof. Throughout the proof, we assume that $p \gg_d \delta^{-O_d(1)}$. Since $\operatorname{rank}(A) \geq 1$, the set of $w \in \mathbb{F}_p^d$ with $wA = \mathbf{0}$ is of cardinality at most p^{d-1}. So, shrinking δ to $\delta/2$ if necessary, we may assume without loss of generality that $wA \neq \mathbf{0}$ for all $w \in W$. Fix any $w \in W$. Pick any $n \in \mathbb{F}_p^d$ with $(nA) \cdot w = 0$, then also $(2nA) \cdot w = 0$. By assumption, we have $(nB + v) \cdot w = (2nB + v) \cdot w = 0$, which implies that $v \cdot w = 0$. Moreover, we have $(nA) \cdot w = 0 \Rightarrow (nB) \cdot w = 0$, or equivalently, $(wA) \cdot n = 0 \Rightarrow (wB) \cdot n = 0$ for all $n \in \mathbb{F}_p^d$. So, wA is parallel to wB. Since $wA \neq \mathbf{0}$, we have that $wB = c_w wA$ for some $c_w \in \mathbb{F}_p$.

Since W is of cardinality at least δp^d, by Lemma 4.11, there exist linearly independent $w_1, \dots, w_d \in W$. So, $v \cdot w_i = 0$ for $1 \leq i \leq d$. This implies that $v = \mathbf{0}$.

For $z \in \mathbb{F}_p^d$, let $M_z(n) := (nA) \cdot (n + z)$. Then,

$$\begin{aligned} \sum_{z \in \mathbb{F}_p^d} |V(M_z) \cap W| &= \sum_{z \in \mathbb{F}_p^d} |\{w \in W\colon (wA) \cdot (w + z) = 0\}| \\ &= \sum_{w \in W} |\{z \in \mathbb{F}_p^d\colon (wA) \cdot (w + z) = 0\}|. \end{aligned} \tag{6.37}$$

Since $wA \neq \mathbf{0}$, we have $|\{z \in \mathbb{F}_p^d\colon (wA) \cdot (w + z) = 0\}| = p^{d-1}$. So, (6.37) is at least $|W| \cdot p^{d-1} \geq \delta p^{2d-1}$. By the pigeonhole principle, there exists $z \in \mathbb{F}_p^d$ such that $|V(M_z) \cap W| \gg \delta p^{d-1}$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Since $wB = c_w wA$ for some $c_w \in \mathbb{F}_p$ for all $w \in W$, we have that

$$(wB) \cdot (w+z) = c_w(wA) \cdot (w+z) = 0$$

for all $w \in V(M_z) \cap W$. Since $|V(M_z) \cap W| \gg \delta p^{d-1}$, by Lemma 6.1 and Proposition 6.2, we may write

$$(wB) \cdot (w+z) = r_z M_z(w) = r_z(wA) \cdot (w+z)$$

for some $r_z \in \mathbb{F}_p$ for all $w \in \mathbb{F}_p^d$. Viewing both sides as polynomials in the variable w and comparing all of their degree 2 terms, we must have that $B = r_z A$. We are done. □

7. *Algebraic properties for quadratic forms in $\mathbb{Z}/p$*

In this section, we extend the results in §6 to $\mathbb{Q}$-valued polynomials.

7.1. *Quadratic forms in $\mathbb{Z}/p$.*

By using the correspondence between polynomials in $\mathbb{F}_p^d$ and in $\mathbb{Z}^d$, we may lift all the definitions pertaining to quadratic forms in $\mathbb{F}_p^d$ to the $\mathbb{Z}^d$ setting in the natural way.

Definition 7.1. (Quadratic forms in $\mathbb{Z}^d$) We say that a function $M\colon \mathbb{Z}^d \to \mathbb{Z}/p$ is a *quadratic form* if

$$M(n) = \frac{1}{p}((nA) \cdot n + n \cdot u + v)$$

for some $d \times d$ symmetric matrix A in $\mathbb{Z}$, some $u \in \mathbb{Z}^d$, and some $v \in \mathbb{Z}$. We say that A is the matrix *associated to* M.

It is more rigorous to name M as a *p-quadratic form*. However, since the quantity p is always clear in this paper, we simply call M to be a quadratic form for short. Whenever we define a quadratic form in this paper, we will write either $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ or $M\colon \mathbb{Z}^d \to \mathbb{Z}/p$ to specify the domain and range of M, so that the term quadratic form will not cause confusion.

By Lemma 2.5, any quadratic form $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$ associated with the matrix $\tilde{A}$ induces a quadratic form $M := \iota \circ (p\tilde{M}) \circ \tau\colon \mathbb{F}_p^d \to \mathbb{F}_p$ associated with the matrix $\iota(\tilde{A})$. Conversely, any quadratic form $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ associated with the matrix A admits a regular lifting $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$, which is a quadratic form associated with the matrix $\tau(A)$.

For a quadratic form $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$, we say that $\tilde{M}$ is *pure/homogeneous/p-non-degenerate* if the quadratic form $M := \iota \circ (p\tilde{M}) \circ \tau$ induced by $\tilde{M}$ is pure/homogeneous/non-degenerate. The *p-rank* of $\tilde{M}$, denoted by $\operatorname{rank}_p(\tilde{M})$, is defined to be the rank of M.

We say that $h_1, \dots, h_k \in \mathbb{Z}^d$ are *p-linearly independent* if for all $c_1, \dots, c_k \in \mathbb{Z}/p$, $c_1h_1 + \cdots + c_kh_k \in \mathbb{Z}$ implies that $c_1, \dots, c_k \in \mathbb{Z}$, or equivalently, if $\iota(h_1), \dots, \iota(h_k)$ are linearly independent.

We can also lift the definitions of sets of zeros and Gowers sets to the $\mathbb{Z}/p$ setting.

- For a polynomial $P \in \operatorname{poly}(\mathbb{Z}^k \to \mathbb{R})$, let $V_p(P)$ denote the set of $n \in \mathbb{Z}^k$ such that $P(n + pm) \in \mathbb{Z}$ for all $m \in \mathbb{Z}^k$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

- For $h_1, \dots, h_t \in \mathbb{Z}^k$, let $V_p(P)^{h_1,\dots,h_t}$ denote the set of $n \in \mathbb{Z}^k$ such that $P(n + pm)$, $P(n + h_1 + pm), \dots, P(n + h_t + pm) \in \mathbb{Z}$ for all $m \in \mathbb{Z}^k$.
- For $\Omega \subseteq \mathbb{Z}^d$ and $s \in \mathbb{N}$, let $\square_{p,s}(\Omega)$ denote the set of $(n, h_1, \dots, h_s) \in (\mathbb{Z}^d)^{s+1}$ such that $n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s \in \Omega + p\mathbb{Z}^d$ for all $\epsilon_1, \dots, \epsilon_s \in \{0, 1\}$. We say that $\square_{p,s}(\Omega)$ is the *sth p-Gowers set* of Ω.

The following lemma is straightforward.

LEMMA 7.2. *For any p-periodic set $\Omega \subseteq \mathbb{Z}^k$ (recall the definition in §1.2) and $K \in \mathbb{N}_+$, we have that*

$$\frac{1}{p^k}|\Omega \cap [p]^k| = \frac{1}{(pK)^k}|\Omega \cap [pK]^k|.$$

Using the connections between $\mathbb{F}_p$-valued polynomials and $\mathbb{Z}/p$-valued polynomials, one can easily extend many results in §6 to the $\mathbb{Z}/p$-setting. The following is an example.

COROLLARY 7.3. (Lifting of Proposition 6.2) *Let $d \in \mathbb{N}_+$, $s \in \mathbb{N}$, $p \gg_{d,s} 1$ be a prime number, $P \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z}/p)$ be a polynomial of degree at most s, and $M : \mathbb{Z}^d \to \mathbb{Z}/p$ be a quadratic form of p-rank at least 3. Then, either $|V_p(M) \cap V_p(P) \cap [p]^d| \le O_{d,s}(p^{d-2})$ or $V_p(M) \subseteq V_p(P)$.*

Moreover, if $V_p(M) \subseteq V_p(P)$, then $P = MP_1 + P_0$ for some $P_0 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ of degree at most s, and some integer coefficient polynomial $P_1 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ of degree at most $s - 2$.

Proof. Let $M' : \mathbb{F}_p^d \to \mathbb{F}_p$ be the quadratic form of rank at least 3 induced by M, $P' : \mathbb{F}_p^d \to \mathbb{F}_p$ be the polynomial of degree at most s induced by P (whose existence is given by Lemma 2.5(i)). By Proposition 6.2, either $|V(M') \cap V(P')| \le O_{d,s}(p^{d-2})$ or $V(M') \subseteq V(P')$. By the definition of lifting, the form case implies that $|V_p(M) \cap V_p(P) \cap [p]^d| = |V(M') \cap V(P')| \le O_{d,s}(p^{d-2})$, and the later case implies that $V_p(M) \subseteq V_p(P)$.

We now assume that $V_p(M) \subseteq V_p(P)$. By the definition of lifting, we have that $V(M') \subseteq V(P')$. By Proposition 6.2, $P' = M'Q'$ for some $Q' \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ of degree at most $s - 2$. Let $Q \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z}/p)$ be a regular lifting of Q' of degree at most $s - 2$ (whose existence is given by Lemma 2.5(ii)). By Lemma 2.5(iv) and (v), $P - pMQ$ is a lifting of $P' - M'Q' \equiv 0$. Then, by Lemma 2.5(iii), we have that $P - pMQ$ is integer valued and is of degree at most s. By Lemma 2.4, we may write $Q = pQ_1 + Q_2$ for some $\mathbb{Z}$-valued polynomial Q_1 and some integer coefficient polynomial Q_2 both having degrees at most $s - 2$. Then, $P = MQ_2 + ((P - pMQ) + pMQ_1)$ with $(P - pMQ) + pMQ_1$ being an integer valued polynomial of degree at most s. We are done. □

We remark that for factorization results, the $\mathbb{Z}/p$ setting differs from the $\mathbb{F}_p$ setting in the appearances of integer valued polynomials. For example, the factorization in Corollary 7.3 is of the form $P = MP_1 + P_0$, while the factorization in Proposition 6.2 is of the form $P = MP_1$ (where P_1 is the R in the statement of Proposition 6.2). This extra integer valued polynomial P_0 arises when passing from a factorization in the $\mathbb{F}_p$ setting to that in the $\mathbb{Z}/p$ setting. The same comment applies to all the results in this section.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

The approach in the proof of Corollary 7.3 can be adapted to lift many other results from the $\mathbb{F}_p$-setting to the $\mathbb{Z}/p$-setting. Below are the liftings of some results from §6 that will be used in the future. We leave the proofs to the interested readers.

COROLLARY 7.4. (Lifting of a special case of Proposition 6.6) *Let $d, s \in \mathbb{N}_+$, $k \in \mathbb{N}$, $\delta > 0$, $p \gg_{d,k} \delta^{-O_{d,k}(1)}$ be a prime, $M\colon \mathbb{Z}^d \to \mathbb{Z}/p$ be a quadratic form of p-rank at least s^2+s+3, and W be a subset of $\square_{p,s}(V(M)) \cap ([p]^d)^{s+1}$ of cardinality at least $\delta|\square_{p,s}(V(M)) \cap ([p]^d)^{s+1}|$. Let $P, Q \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z}/p)$ with $\deg(P) \le k-1$ and $\deg(Q) \le k$. Suppose that for all $(n, h_1, \dots, h_s) \in W$, we have that*

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P(n) + \Delta_{h_s} \cdots \Delta_{h_1} Q(n) \in \mathbb{Z},$$

(where $\Delta_{h_s} \cdots \Delta_{h_2} P(n)$ is understood as $P(n)$ when $s = 1$). Then,

$$P = MP_1 + P_2 + P_3 \quad \text{and} \quad Q = MQ_1 + Q_2 + Q_3 \tag{7.1}$$

for some integer coefficient polynomials P_1 and Q_1 of degrees at most $k-3$ and $k-2$, respectively, some integer valued polynomials P_2 and Q_2 of degrees at most $k-1$ and k, respectively, and some $\mathbb{Z}/p$-valued polynomials P_3 and Q_3 of degrees at most $s-2$ and $s-1$, respectively.

COROLLARY 7.5. (Lifting of Proposition 6.10) *Let $d \in \mathbb{N}_+$, $\delta > 0$, $p \gg_d \delta^{-O_d(1)}$ be a prime, and A be a $\mathbb{Z}$-valued $d \times d$ matrix of p-rank at least 3. Let B be a $\mathbb{Z}$-valued $d \times d$ matrix and $v \in \mathbb{Z}^d$. Let W be a subset of $[p]^d$ of cardinality at least δp^d. Suppose that for all $w \in W$ and $n \in \mathbb{Z}^d$ with $(nA) \cdot w \in p\mathbb{Z}$, we have that*

$$(nB + v) \cdot w \in p\mathbb{Z}.$$

Then, $v \in p\mathbb{Z}$ and $B = cA + pB_0$ for some $c \in \mathbb{Z}$ and some $d \times d$ integer valued matrix B_0.

7.2. *Passing from p-periodic polynomials to rational ones.* In this section, we extend Corollary 7.4 from $\mathbb{Z}/p$-valued polynomials to rational polynomials.

PROPOSITION 7.6. *Let $d, K, s \in \mathbb{N}_+$, $k \in \mathbb{N}$, $\delta > 0$, $p \gg_{d,k} \delta^{-O_{d,k}(1)}$ be a prime, $M\colon \mathbb{Z}^d \to \mathbb{Z}/p$ be a quadratic form of p-rank at least s^2+s+3, and W be a subset of $\square_{p,s}(V(M)) \cap ([p^2K]^d)^{s+1}$ of cardinality at least $\delta|\square_{p,s}(V(M)) \cap ([p^2K]^d)^{s+1}|$. Let $P, Q \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Q})$ with $\deg(P) \le k-1$ and $\deg(Q) \le k$. Suppose that*

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P(n) + \Delta_{h_s} \cdots \Delta_{h_1} Q(n) \in \mathbb{Z} \quad \text{for all } (n, h_1, \dots, h_s) \in W + p^2K(\mathbb{Z}^d)^{s+1} \tag{7.2}$$

(where $\Delta_{h_s} \cdots \Delta_{h_2} P(n)$ is understood as $P(n)$ when $s = 1$). Then,

$$qP = P_1 + P_2 \quad \text{and} \quad qQ = Q_1 + Q_2$$

for some integer $q \in \mathbb{N}_+$ with $q \ll_{d,k} \delta^{-O_{d,k}(1)}$, some $P_1, Q_1 \in \mathrm{poly}(V_p(M) \to \mathbb{R}|\mathbb{Z})$ with $\deg(P_1) \le k-1$, $\deg(Q_1) \le k$, and some $P_2, Q_2 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Q})$ with $\deg(P_2) \le s-2$, $\deg(Q_2) \le s-1$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

In applications, K will be a large number. We also remark that all the constants in Proposition 7.6 are independent of s since the statement holds trivially when $s \geq k$.

Let $f = P$ or Q. Our strategy is to decompose f as $qf = \sum_{i=0}^{k}(f_i/p^i)$ for some $k \in \mathbb{N}$ and $q \in \mathbb{Z}$, $p \nmid q$, where each f_i is a 'good' polynomial, meaning that it behaves very similarly to a $\mathbb{Z}/p$-valued polynomial. One can then prove that periodicity conditions on f descend to similar conditions on f_k, leading to the conclusion that f_k/p is 'good'. We may then absorb the term f_k/p^k into $(f_{k-1})/(p^{k-1})$ and reduce k by 1. Inductively, we have $k = 0$, which implies that qf itself is 'good'. For future references, we refer to this approach as the *p-expansion trick*.

Proof of Proposition 7.6. Throughout the proof, we assume that $p \gg_{d,k} \delta^{-O_{d,k}(1)}$.

Step 1: setup for the induction. For convenience, denote $d' := \mathrm{rank}_p(M) \geq s^2 + s + 3$. We assume that $s \leq k+1$ since otherwise, one can simply take $P = P_2$, $Q = Q_2$, $q = 1$, and $P_1 = Q_1 = 0$. Since P, Q take values in $\mathbb{Q}$ and are of degree at most k, by the multivariate polynomial interpolation, there exists $q \in \mathbb{N}$, $p \nmid q$, and $t_0 \in \mathbb{N}$ such that

$$qP = \sum_{j=0}^{t_0} \frac{P_j'}{p^j} \quad \text{and} \quad qQ = \sum_{j=0}^{t_0} \frac{Q_j'}{p^j}$$

for some integer coefficient polynomials $P_j', Q_j' \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(P_j') \leq k-1$ and $\deg(Q_j') \leq k$. We say that a polynomial $f : \mathbb{Z}^d \to \mathbb{Z}$ is *good* if

$$f = \sum_{j=0}^{\lfloor \deg(f)/2 \rfloor} M^j f_j$$

for some integer coefficient polynomial $f_j : \mathbb{Z}^d \to \mathbb{Z}$ of degree at most $\deg(f) - 2j$. A good polynomial not only takes integer values in $V_p(M)$, but also admits a nice decomposition in terms of M. Clearly, $P_0', \ldots, P_{t_0}', Q_0', \ldots, Q_{t_0}'$ are good polynomials.

Let $t \in \mathbb{N}$ be the smallest integer such that we can write

$$q'P = \sum_{j=0}^{t} \frac{P_j''}{p^j} + P' \quad \text{and} \quad q'Q = \sum_{j=0}^{t} \frac{Q_j''}{p^j} + Q'$$

for some $q' \in \mathbb{N}$, $p \nmid q'$, some good polynomials $P_j'', Q_j'' \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(P_j'') \leq k-1$, $\deg(Q_j'') \leq k$, and some $P', Q' \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Q})$ with $\deg(P') \leq s-2$ and $\deg(Q') \leq s-1$. Obviously, such t exists and is at most t_0. Our first goal is to show that $t = 0$.

Suppose in contrast that $t > 0$. We consider the last terms and write

$$P_t'' = \sum_{j=0}^{\lfloor (k-1)/2 \rfloor} M^j P_j \quad \text{and} \quad Q_t'' = \sum_{j=0}^{\lfloor k/2 \rfloor} M^j Q_j$$

for some integer coefficient polynomials $P_j, Q_j \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(P_j) \leq k-1-2j$ and $\deg(Q_j) \leq k-2j$. We say that a polynomial $g \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ is *s-simple* if

$$q'g = pg_0 + pMg_1 + g_2$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

for some $q' \in \mathbb{N}_+$, $q' \le O_{k,s}(1)$, some $g_0 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(g_0) \le \deg(g)$, some integer coefficient polynomials $g_1 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(g_0) \le \deg(g) - 2$, and some $g_2 \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Q})$ with $\deg(g_2) \le s - 1$. Roughly speaking, if g is s-simple, then $Q = q'g/p$ is a $\mathbb{Z}/p$-valued polynomial that admits a decomposition similar to (7.1). If we can show that all of $P_0, \dots, P_{\lfloor (k-1)/2 \rfloor}$ are $(s-1)$-simple and all of $Q_0, \dots, Q_{\lfloor k/2 \rfloor}$ are s-simple, then enlarging the constant q' if necessary, we may absorb the term P_t''/p^t into $(P_{t-1}'')/(p^{t-1}) + P'$, and the term Q_t''/p^t by $(Q_{t-1}'')/(p^{t-1}) + Q'$, which contradicts the minimality of t.

In the rest of the proof, we say that g is *simple* instead of s-simple if the value of s is clear from the context. By induction, it then suffices to show that if all of $P_{i+1}, \dots, P_{\lfloor k/2 \rfloor}, Q_{i+1}, \dots, Q_{\lfloor k/2 \rfloor}$ are simple for some $0 \le i \le \lfloor k/2 \rfloor$, then P_i and Q_i are also simple (for the special case $i = \lfloor k/2 \rfloor$, we show that P_i and Q_i are simple without induction hypothesis).

Since $P_{i+1}, \dots, P_{\lfloor k/2 \rfloor}, Q_{i+1}, \dots, Q_{\lfloor k/2 \rfloor}$ are simple, enlarging the constant q' and absorbing the term $1/p^t \sum_{j=i+1}^{\lfloor k/2 \rfloor} M^j P_j$ into $(P_{t-1}'')/(p^{t-1}) + P'$, and the term $(1/p^t) \sum_{j=i+1}^{\lfloor k/2 \rfloor} M^j Q_j$ by $(Q_{t-1}'')/(p^{t-1}) + Q'$ if necessary, we may assume without loss of generality that

$$P_t'' = \sum_{j=0}^{i} M^j P_j \quad \text{and} \quad Q_t'' = \sum_{j=0}^{i} M^j Q_j.$$

Step 2: the isolation trick. We first provide an outline ti show that P_i, Q_i are simple. For simplicity, we assume for the moment that (7.2) holds for $(n, h_1, \dots, h_s) \in \square_{p,s}(V(M))$. By the construction of P_t'' and Q_t'', and by multiplying (7.2) by p^t, we have

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P_t''(n) + \Delta_{h_s} \cdots \Delta_{h_1} Q_t''(n) \in p\mathbb{Z} \quad \text{for all } (n, h_1, \dots, h_s) \in \square_{p,s}(V(M)). \tag{7.3}$$

If $i = 0$, then we may apply Corollary 7.4 directly to show that P_0 and Q_0 are simple. When $i > 0$, our strategy is to replace $(n, h_1, \dots, h_s)$ by $(n + pm_0, h_1 + pm_1, \dots, h_s + pm_s)$ and use the periodicity property of P_t'' and Q_t'' to isolate P_i and Q_i for different i (for convenience, we call this the *isolation trick*). To illustrate this, consider the case when $i = 1$, $s = 1$, and $M(n) = 1/p(n \cdot n)$. Then, (7.3) implies that

$$\begin{aligned} &P_0(n + pm_0) + (MP_1)(n + pm_0) + \Delta_{h_1 + pm_1} Q_0(n + pm_0) \\ &\quad + \Delta_{h_1 + pm_1}(MQ_1)(n + pm_0) \in p\mathbb{Z} \end{aligned} \tag{7.4}$$

for all $m_0, m_1 \in \mathbb{Z}^d$. Since $M(n + pm) \equiv M(n) + 2m \cdot n \bmod p\mathbb{Z}$ and $R(n + pm) \equiv R(n) \bmod p\mathbb{Z}$ for all integer coefficient polynomial R, one can show that the expression in (7.4) modulo $p\mathbb{Z}$ is equal to $P_0(n) + (MP_1)(n) + \Delta_{h_1} Q_0(n) + \Delta_{h_1}(MQ_1)(n) + 2(m_0 \cdot n)P_1(n)$ plus

$$2(n + h_1) \cdot (m_0 + m_1)Q_1(n + h_1) - 2n \cdot m_0 Q_1(n) + 2n \cdot m_0 P_1(n).$$

Since m_0 and m_1 are arbitrary, this implies that $Q_1(n + h_1)$, $Q_1(n) - p_1(n) \in p\mathbb{Z}$ whenever $n, n + h_1 \in V_p(M)$ and $n, n + h \neq \mathbf{0}$, which implies that P_1 and Q_1 are simple.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

We now carry out this trick to the general setting. To apply the isolation trick, we need to restrict ourself to the set of $(n, h_1, \dots, h_s)$ where there is sufficient freedom in choosing $m_0, \dots, m_s$. By the pigeonhole principle, there exists a subset W_0 of $\square_{p,s}(V(M)) \cap ([p]^d)^{s+1}$ of cardinality at least $\delta|\square_{p,s}(V(M)) \cap ([p]^d)^{s+1}|/2$ and for each $(n, h_1, \dots, h_s) \in W_0$, a set $W_0(n, h_1, \dots, h_s) \subseteq ([pK]^d)^{s+1}$ of cardinality at least $\delta(pK)^{d(s+1)}/2$ such that

$$(n + pm_0, h_1 + pm_1, \dots, h_s + pm_s) \in W$$

for all $(n, h_1, \dots, h_s) \in W_0$ and $(m_0, \dots, m_s) \in W_0(n, h_1, \dots, h_s)$. Now, fix any $(n, h_1, \dots, h_s) \in W_0$ and $(m_0, \dots, m_s) \in W_0(n, h_1, \dots, h_s)$. We have

$$\Delta_{h_{s-1}+pm_{s-1}} \cdots \Delta_{h_1+pm_1} P(n) + \Delta_{h_s+pm_s} \cdots \Delta_{h_1+pm_1} Q(n) \in \mathbb{Z}$$

by assumption. So,

$$\Delta_{h_{s-1}+pm_{s-1}} \cdots \Delta_{h_1+pm_1} P_t''(n) + \Delta_{h_s+pm_s} \cdots \Delta_{h_1+pm_1} Q_t''(n) \in p\mathbb{Z}. \tag{7.5}$$

We now show that (7.5) in fact holds for all $m_0, \dots, m_s$ (and thus, we are in good shape to use the isolation trick). Assume that

$$M(n) = \frac{1}{p}((nA) \cdot n + u \cdot n + v)$$

for some $d \times d$ symmetric integer coefficient matrix A, some $u \in \mathbb{Z}^d$, and some $v \in \mathbb{Z}$. Then, for all $m, n \in V_p(M)$, we have

$$M(n + pm) \equiv M(n) + ((2(nA) + u) \cdot m) \bmod p\mathbb{Z}. \tag{7.6}$$

For any $r \in \mathbb{N}_+ \epsilon = (\epsilon_1, \dots, \epsilon_r) \in \{0, 1\}^r$ and $a = (a_1, \dots, a_r) \in (\mathbb{Z}^d)^r$, denote

$$\epsilon \cdot a := \epsilon_1 a_1 + \cdots + \epsilon_r a_r \in \mathbb{Z}^d.$$

Since P_j and Q_j have integer coefficients, it is not hard to see from (7.5) and (7.6) that for any $(n, h_1, \dots, h_s) \in W_0$ and $(m_0, \dots, m_s) \in W_0(n, h_1, \dots, h_s)$, we have

$$\begin{aligned}
&\frac{1}{p} \sum_{j=0}^{i} \sum_{\epsilon_* \in \{0,1\}^{s-1}} (-1)^{|\epsilon_*|} (M(n' + \epsilon_* \cdot h_*) \\
&\qquad + (2(n' + \epsilon_* \cdot h_*)A + u) \cdot (\epsilon_* \cdot m_*))^j P_j(n' + \epsilon_* \cdot h_*) \\
&+ \frac{1}{p} \sum_{j=0}^{i} \sum_{\epsilon \in \{0,1\}^s} (-1)^{|\epsilon|} (M(n' + \epsilon \cdot h) \\
&\qquad + (2(n' + \epsilon \cdot h)A + u) \cdot (\epsilon \cdot m))^j Q_j(n' + \epsilon \cdot h) \in \mathbb{Z},
\end{aligned} \tag{7.7}$$

where $n' = n + pm_0$, $h = (h_1, \dots, h_s)$, $h_* = (h_1, \dots, h_{s-1})$, $m = (m_1, \dots, m_s)$, and $m_* = (m_1, \dots, m_{s-1})$, and the terms ϵ_*, h_*, m_* are regarded as non-existing if $s = 1$. Since P_i and Q_i have integer coefficients, the residue class mod $\mathbb{Z}$ of the left-hand side of (7.7) remains unchanged if we replace m_i by $m_i \bmod p\mathbb{Z}$ for $0 \le i \le s$. In other words, for fixed $(n, h_1, \dots, h_s) \in W_0$, the left-hand side of (7.7) is a p-periodic polynomial in the variables $(m_0, \dots, m_s)$. Since $W_0(n, h_1, \dots, h_s) \bmod p(\mathbb{Z}^d)^{s+1}$ contains at least

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$\delta p^{d(s+1)}/2$ residue classes mod $p\mathbb{Z}$, it follows from Lemma 4.10 that (7.7) (and thus (7.5)) holds for all $m_0, \dots, m_s \in \mathbb{Z}^d$ and for all $(n, h_1, \dots, h_s) \in W_0$ (and thus, for all $(n, h_1, \dots, h_s) \in W_0 + p(\mathbb{Z}^d)^{s+1}$).

We now apply the isolation trick to (7.7) to deduce information on P_i and Q_i. We first consider the case when $2nA + u, h_1A, \dots, h_sA$ are p-linearly independent (this corresponds to the condition where $n, n + h \neq \mathbf{0}$ in the outlined example). It is not hard to see that for any $1 \leq j \leq i$ and $c_j \in \mathbb{Z}^d$, there exists $m_j \in \mathbb{Z}^d$ such that

$$(2nA + u) \cdot m_j, (h_{j'}A) \cdot m_j, (h_{i+j}A) \cdot m_j - c_j \in p\mathbb{Z}$$

for all $1 \leq j' \leq s$, $j' \neq i + j$. For $i + 1 \leq j \leq s$, let $m_j = \mathbf{0}$.

The generalization of the isolation trick to the case $i \geq 2$ requires a lot of computations, which is what we will do now. Then, for all $0 \leq j \leq i$ and $\epsilon_* \in \{0, 1\}^{s-1}$, we have that

$$\begin{aligned}
&(M(n + \epsilon_* \cdot h_*) + (2(n + \epsilon_* \cdot h_*)A + u) \cdot (\epsilon_* \cdot m_*))^j P_j(n + \epsilon_* \cdot h_*) \\
&\quad = (M(n + \epsilon_* \cdot h_*) + (2(n + \epsilon' \cdot h' + \epsilon''_* \cdot h''_*)A + u) \cdot (\epsilon' \cdot m'))^j P_j(n + \epsilon_* \cdot h_*) \\
&\quad \equiv \left(M(n + \epsilon_* \cdot h_*) + 2\sum_{\ell=1}^{i} \epsilon_\ell \epsilon_{i+\ell} c_\ell\right)^j P_j(n + \epsilon_* \cdot h_*) \bmod p\mathbb{Z}, \qquad (7.8)
\end{aligned}$$

where $h' := (h_1, \dots, h_i)$, $h''_* := (h_{i+1}, \dots, h_{s-1})$, $\epsilon' := (\epsilon_1, \dots, \epsilon_i)$, and $\epsilon''_* := (\epsilon_{i+1}, \dots, \epsilon_{s-1})$. Similarly, for all $0 \leq j \leq i$ and $\epsilon \in \{0, 1\}^s$, we have that

$$\begin{aligned}
&(M(n + \epsilon \cdot h) + (2(n + \epsilon \cdot h)A + u) \cdot (\epsilon \cdot m))^j Q_j(n + \epsilon \cdot h) \\
&\quad \equiv \left(M(n + \epsilon \cdot h) + 2\sum_{\ell=1}^{i} \epsilon_\ell \epsilon_{i+\ell} c_\ell\right)^j Q_j(n + \epsilon \cdot h) \bmod p\mathbb{Z}. \qquad (7.9)
\end{aligned}$$

Combining (7.7), (7.8), and (7.9), we have that

$$\begin{aligned}
&\sum_{j=0}^{i} \sum_{\epsilon_* \in \{0,1\}^{s-1}} (-1)^{|\epsilon_*|} \left(M(n + \epsilon_* \cdot h_*) + 2\sum_{\ell=1}^{i} \epsilon_\ell \epsilon_{i+\ell} c_\ell\right)^j P_j(n + \epsilon_* \cdot h_*) \\
&\quad + \sum_{j=0}^{i} \sum_{\epsilon \in \{0,1\}^s} (-1)^{|\epsilon|} \left(M(n + \epsilon \cdot h) + 2\sum_{\ell=1}^{i} \epsilon_\ell \epsilon_{i+\ell} c_\ell\right)^j Q_j(n + \epsilon \cdot h) \in p\mathbb{Z} \qquad (7.10)
\end{aligned}$$

for all $c_1, \dots, c_j \in \mathbb{Z}^d$. Viewing (7.10) as a polynomial of $c_1, \dots, c_i$ and considering the coefficient of the term $c_1 \cdot \dots \cdot c_i$, by interpolation, we have that

$$\begin{aligned}
&q'' \sum_{\epsilon'''_* \in \{0,1\}^{s-2i-1}} (-1)^{|\epsilon'''_*|} P_i(n + h_1 + \dots + h_{2i} + \epsilon'''_* \cdot h'''_*) \\
&\quad + q'' \sum_{\epsilon''' \in \{0,1\}^{s-2i}} (-1)^{|\epsilon'''|} Q_i(n + h_1 + \dots + h_{2i} + \epsilon''' \cdot h''') \in p\mathbb{Z}
\end{aligned}$$

for some $q'' \in \mathbb{N}_+$, $q'' \leq O_{d,k}(1)$, where $h''' = (h_{2i+1}, \dots, h_s)$, $h'''_* = (h_{2i+1}, \dots, h_{s-1})$, $\epsilon'''_* = (\epsilon_{2i+1}, \dots, \epsilon_{s-1})$, and $\epsilon''' = (\epsilon_{2i+1}, \dots, \epsilon_s)$, and the sum

$$\sum_{\epsilon \in \{0,1\}^r} (-1)^{|\epsilon|} a(n + \epsilon \cdot h)$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

is considered as $a(n)$ if $r = 0$, and as 0 if $r < 0$. In conclusion, we have that

$$q''(\Delta_{h_{s-1}} \cdots \Delta_{h_{2i+1}} P_i(n') + \Delta_{h_s} \cdots \Delta_{h_{2i+1}} Q_i(n')) \in p\mathbb{Z} \tag{7.11}$$

with $n' = n + h_1 + \cdots + h_{2i}$ for all $(n, h_1, \ldots, h_s) \in W_0 + p(\mathbb{Z}^d)^{s+1}$ such that $2nA + u, h_1A, \ldots, h_sA$ are p-linearly independent.

Step 3: imposing linear independent conditions. The problem is now reduced to a point where Corollary 7.4 is applicable, as long as $2nA + u, h_1A, \ldots, h_sA$ are rarely p-linearly dependent. It is convenient to work in the $\mathbb{F}_p^d$-setting in Step 3. Let $\tilde{M}\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be the non-degenerate quadratic form induced by M, and W' be the set of $(n', h_{2i+1}, \ldots, h_s) \in (\mathbb{F}_p^d)^{s-2i+1}$ such that $n' = n + h_1 + \cdots + h_{2i}$ for some $(n, h_1, \ldots, h_s) \in \iota(W_0)$ such that $2n\iota(A) + \iota(u), h_1\iota(A), \ldots, h_s\iota(A)$ are linearly independent. Then, it is not hard to see that (7.11) holds for all $(n', h_{2i+1}, \ldots, h_s) \in \iota^{-1}(W')$. If we can show that $|W'| \geq \delta|\square_{s-2i}(V(\tilde{M}))|/8$, then it follows from (7.11) and Corollary 7.4 that P_i and Q_i are simple.

We now estimate the cardinality of W'. Since $d' \geq s^2 + s + 3$, it follows from Proposition 4.19 that

$$|\square_j(V(\tilde{M}))| = p^{(j+1)d-((j(j+1)/2)+1)}(1 + O_j(p^{-1/2})) \tag{7.12}$$

for all $0 \leq j \leq s$. For convenience, denote $\mathbf{x} := (n', h_{2i+1}, \ldots, h_s)$. We say that $\mathbf{x}$ is *good* if $2n'\iota(A) + \iota(u), h_{2i+1}\iota(A), \ldots, h_s\iota(A)$ are linearly independent. Let $W'_{\mathbf{x}}$ denote the set of $(h_1, \ldots, h_{2i}) \in (\mathbb{F}_p^d)^{2i}$ such that $(n' - h_1 - \cdots - h_{2i}, h_1, \ldots, h_s) \in \iota(W)$. Then,

$$\sum_{\mathbf{x} \in \square_{s-2i}(V(\tilde{M}))} |W'_{\mathbf{x}}| = |W_0| \geq \delta|\square_s(V(\tilde{M}))|/2 = \frac{1}{2}\delta p^{(s+1)d-((s(s+1)/2)+1)}(1 + O_s(p^{-1/2})). \tag{7.13}$$

Therefore, informally speaking, by (7.12) and (7.13), on average, the cardinality of $W'_{\mathbf{x}}$ is at least $\delta p^{2id-(2s-2i+1)i}/2$.

Note that for good $\mathbf{x}$, $(h_1, \ldots, h_{2i}) \in W'_{\mathbf{x}}$ only if $(n' - h_1 - \cdots - h_{2i}, h_1, \ldots, h_s) \in \square_s(V(\tilde{M}))$. Using Lemma 4.18, this is equivalent of saying that:

- $\tilde{M}(n') - \tilde{M}(n' - h_j) = (2n'\iota(A) + \iota(u)) \cdot h_j = 0$ for all $1 \leq j \leq 2i$;
- $(h_j\iota(A)) \cdot h_{j'} = 0$ for all $1 \leq j, j' \leq s$ with at least one of j, j' belonging to $\{1, \ldots, 2i\}$.

There are in total $(2s - 2i + 1)i \leq (s + 1/2)^2/2$ independent equations above in total. Since $d \geq s^2 + s + 3$, by repeatedly using Corollary 4.16 and the fact that $\mathbf{x}$ is good, it is not difficult to show that

$$|W'_{\mathbf{x}}| \leq p^{2id-(2s-2i+1)i}(1 + O_s(p^{-1/2})). \tag{7.14}$$

We leave the proof of (7.14) to the interested readers. (Alternatively, (7.14) can also be deduced as a special case of [**18**, Theorem 5.11] (in the terminology introduced in [**18**], $W'_{\mathbf{x}}$ is a subset of an $\tilde{M}$-set of total co-dimension $(2s - 2i + 1)i$).)

Let $V := \{n\iota(A)\colon n \in \mathbb{F}_p^d\} \subseteq \mathbb{F}_p^d$. Then, V is a subspace of $\mathbb{F}_p^d$ of dimension d'. Let V' be the set of $m, m_{2i+1}, \ldots, m_s \in V$ such that $2m + \iota(u), m_{2i+1}, \ldots, m_s$ are linearly dependent. Similar to the proof of Lemma 4.11(i), it is not hard to see

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

the $|V'| \le sp^{(d+1)(s-2i)}$. However, for each $m \in V$, the number of $n \in \mathbb{F}_p^d$ such that $n\iota(A) = m$ is at most $p^{d-d'}$. Therefore, the number of $\mathbf{x} \in (\mathbb{F}_p^d)^{s-2i+1}$ which are not good is at most $p^{d-d'}|V'| \le sp^{(s-2i+1)d+(s-2i)-d'}$. So,

$$\sum_{\mathbf{x}\in(\mathbb{F}_p^d)^{s-2i+1},\ \mathbf{x} \text{ is not good}} |W'_{\mathbf{x}}| \le sp^{(s-2i+1)d+(s-2i)-d'} \cdot p^{2id} = sp^{(s+1)d+(s-2i)-d'}. \quad (7.15)$$

Since $(s+1)d + (s-2i) - d' < (s+1)d - ((s(s+1)/2)+1)$, it follows from (7.12), (7.13), (7.14), (7.15), and the pigeonhole principle that there exists a subset $W'' \subseteq \square_{s-2i}(V(\tilde{M}))$ of cardinality at least $\delta|\square_{s-2i}(V(\tilde{M}))|/20$ such that $|W'_{\mathbf{x}}| \ge \delta p^{2id-(2s-2i+1)i}/20$ for all $\mathbf{x} \in W''$. In particular, $W'_{\mathbf{x}}$ is non-empty for all $\mathbf{x} \in W''$. Therefore, W'' is a subset of W' and thus, $|W'| \ge |W''| \ge \delta|\square_{s-2i}(V(\tilde{M}))|/20$.

Step 4: reducing the complexity of the denominator. In conclusion, we have shown by induction that $t = 0$ and, therefore, we may write

$$qP = P'_0 + P' \quad \text{and} \quad qQ = Q'_0 + Q'$$

for some good polynomials P'_0, Q'_0 of degrees at most $k-1$ and k, respectively, and some P', $Q' \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Q})$ of with $\deg(P') \le s-2$ and $\deg(Q') \le s-1$. Our final goal is to show that one can further require q to be small.

Let $q^* \in \mathbb{N}_+$ be such that $qq^* \equiv 1 \bmod p^k\mathbb{Z}$. Then, $P = (1-qq^*/q)P'_0 + q^*P'_0 + 1/qP'$. Denote $P_0 := (1-qq^*)P'_0$. Since P'_0 is good, we have that P_0 has coefficients in $\mathbb{Z}$ and that $\Delta_{h_{s-1}} \cdots \Delta_{h_1}(q^*P'_0)(n) \in \mathbb{Z}$ for all $(n, h_1, \dots, h_s) \in \square_{p,s}(V(M))$. Similarly, $Q_0 := (1-qq^*)Q'_0$ also has coefficients in $\mathbb{Z}$, and $\Delta_{h_s} \cdots \Delta_{h_1}(q^*Q'_0)(n) \in \mathbb{Z}$ for all $(n, h_1, \dots, h_s) \in \square_{p,s}(V(M))$. So, we have that

$$\Delta_{h_{s-1}} \cdots \Delta_{h_1} P_0(n) + \Delta_{h_s} \cdots \Delta_{h_1} Q_0(n) \in q\mathbb{Z} \quad (7.16)$$

for all $(n, h_1, \dots, h_s) \in W + ([pK']^d)^{s+1}$, where for convenience, we denote $K' := pK$.

Our strategy is to fix $(n, h_1, \dots, h_s)$, apply (7.16) for $(n+pm_0, h_1+pm_1, \dots, h_s+pm_s)$, and then apply Proposition 3.18 as a polynomial in the variables $m_0, \dots, m_s$. By the pigeonhole principle, there exists some $(n, h_1, \dots, h_s) \in \square_{p,s}(V(M)) \cap ([p]^d)^{s+1}$ and a subset $J \subseteq ([K']^d)^{s+1}$ of cardinality at least $\delta K'^{d(s+1)}$ such that $(n+pm_0, h_1+pm_1, \dots, h_s+pm_s) \in W$ for all $(m_0, \dots, m_s) \in J$. Write

$$\begin{aligned} G(m_0, m_1, \dots, m_s) &:= \Delta_{h_{s-1}+pm_{s-1}} \cdots \Delta_{h_1+pm_1} P_0(n+pm_0) \\ &\quad + \Delta_{h_s+pm_s} \cdots \Delta_{h_1+pm_1} Q_0(n+pm_0). \end{aligned}$$

Then, G is a rational polynomial of degree at most k. By assumption, $(1/q)G(\mathbf{m}) \in \mathbb{Z}$ for all $\mathbf{m} \in J + ([K']^d)^{s+1}$. So, by Proposition 3.18, there exists $r \in \mathbb{N}_+$ with $r \ll_{d,k,s} \delta^{-O_{d,k,s}(1)}$ such that all the coefficients of $(1/q)G$ are in $\mathbb{Z}/r$. Write

$$\begin{aligned} G'(m_0, m_1, \dots, m_s) &:= \frac{1}{q} G\left(\frac{1}{p}m_0 - n, \frac{1}{p}m_1 - h_1, \dots, \frac{1}{p}m_s - h_s\right) \\ &= \Delta_{m_{s-1}} \cdots \Delta_{m_1}\left(\frac{1}{q}P_0\right)(m_0) + \Delta_{m_s} \cdots \Delta_{m_1}\left(\frac{1}{q}Q_0\right)(m_0). \end{aligned}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Then, the coefficients of G' are in $\mathbb{Z}/rp^k$. However, since P_0 and Q_0 are integer coefficient polynomials, the coefficients of G' are in $\mathbb{Z}/q$ and are thus in $(\mathbb{Z}/rp^k) \cap (\mathbb{Z}/q) \subseteq \mathbb{Z}/r$.

Now, assume that

$$\frac{1}{q}P_0(n) := \sum_{i\in\mathbb{N}^d,|i|\le k} a_i n^i \quad \text{and} \quad \frac{1}{q}Q_0(n) := \sum_{i\in\mathbb{N}^d,|i|\le k} b_i n^i$$

for some $a_i, b_i \in \mathbb{Q}$. Fix any $i = (i_1, \dots, i_d) \in \mathbb{N}^d$ with $|i| \ge s-1$. Let $(m_1, \dots, m_{|i|+1}) \in (\mathbb{Z}^d)^{|i|}$ be any vector such that $m_s = \mathbf{0}$ and such that the number of $m_{i'}$, $1 \le i' \le |i|+1$, $i' \ne s$ which is equal to the jth standard unit vector is equal to i_j for all $1 \le j \le d$. Then, it is not hard to see that

$$\Delta_{m_{|i|+1}} \cdots \Delta_{m_{s+1}} G'(\mathbf{0}, m_1, \dots, m_s) = i!\, a_i.$$

Since the coefficients of G' are in $\mathbb{Z}/r$, we have that $a_i \in \mathbb{Z}/r'$ for all $i \in \mathbb{N}^d$ with $|i| \ge s-1$, where $r' := (k!)^d r$.

Now, fix any $i = (i_1, \dots, i_d) \in \mathbb{N}^d$ with $|i| \ge s$. Let $(m_1, \dots, m_{|i|}) \in (\mathbb{Z}^d)^{|i|}$ be any vector such that the number of $m_{i'}$, $1 \le i' \le |i|$ which is equal to the jth standard unit vector is equal to i_j for all $1 \le j \le d$. Then, it is not hard to see that

$$\Delta_{m_{|i|}} \cdots \Delta_{m_{s+1}} G'(\mathbf{0}, m_1, \dots, m_s) = i!\, b_i + (i - m_s)!\, a_{i-m_s}.$$

Since the coefficients of G' are in $\mathbb{Z}/r$ and since $a_{i-m_s} \in \mathbb{Z}/r'$, we have that $b_i \in \mathbb{Z}/r'$ for all $i \in \mathbb{N}^d$ with $|i| \ge s$. So, we may write P_0 as $(1/q)P_0 = P_1 + P_2$ for some polynomial P_1 with coefficients in $\mathbb{Z}/r'$ having at most the same degree as P and some P_2 of degree at most $s-2$. So,

$$r'P = r'(q^* P_0' + P_1) + \left(r'P_2 + \frac{1}{q}P'\right),$$

where $r'(q^* P_0' + P_1)$ take integer values in $V_p(M)$. A similar decomposition apples to Q and so we are done. □

8. *An example for the Heisenberg group*

Before proving Theorem 5.1, we use a sample case to illustrate the idea of the proof. The Heisenberg group is the set $G = \mathbb{R}^3$ with the group action given by

$$(a, b, c) * (a', b', c') := (a + a', b + b', c + c' + ab') \quad \text{for all } a, b, c, a', b', c' \in \mathbb{R}.$$

This is a 2-step nilmanifold with respect to the lower central series $G_{\mathbb{N}}$ (that is, $G_0 = G_1 = G$, $G_2 = \{0\}^2 \times \mathbb{R}$, and $G_i = \{(0, 0, 0)\}$ for all $i \ge 3$). Let $\Gamma = \mathbb{Z}^3$. The Lie algebra of G has a Mal'cev basis $\mathcal{X} = \{X_1, X_2, X_3\}$ with the Mal'cev coordinate map $\psi^{-1} : \mathbb{R}^3 \to G$ given by

$$\psi^{-1}(t_1, t_2, t_3) := \exp(t_1 X_1) \exp(t_2 X_2) \exp(t_3 X_3) = (t_1, t_2, t_1 t_2 + t_3).$$

In this section, we provide an outline for the proof of the following simple case of Theorem 5.1 for linear polynomial sequences.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

PROPOSITION 8.1. *Let $0 < \delta < 1/2$, $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ with $d \gg_s 1$, and p be a prime. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be the quadratic form given by $M(n) = n \cdot n$. Let $G_{\mathbb{N}}$ be the Heisenberg group with the lower central series filtration and $g\colon \mathbb{Z}^d \to G$ be of the form $g(n_1, \dots, n_d) = a_1^{n_1} \cdot \dots \cdot a_d^{n_d}$ for some $a_1, \dots, a_d \in G$ commuting with each other with $a_i^p \in \Gamma$. If $p \gg_{\delta,d} 1$ and $(g(n)\Gamma)_{n\in\iota^{-1}(V(M))}$ is not δ-equidistributed on G/Γ, then there exists a non-trivial type-I horizontal character η with $0 < \|\eta\| \ll_{\delta,d} 1$ such that $\eta \circ g \mod \mathbb{Z}$ is a constant on $V_p(M)$.*

Sketch of the proof of Proposition 8.1. Step 1: construction of a nilsequence in a joining system. By Lemma 3.23, if $(g(n)\Gamma)_{n\in\iota^{-1}(V(M))}$ is not δ-equidistributed on G/Γ, then there exist $\xi \in \mathbb{Z}$ with $|\xi| \ll \delta^{-O(1)}$ and a function $F\colon G/\Gamma \to \mathbb{C}$ with $\|F\|_{\mathrm{Lip}} \le 1$ and vertical frequency ξ such that

$$\left|\mathbb{E}_{n\in\iota^{-1}(V(M))} F(g(n)\Gamma) - \int_{G/\Gamma} F \, dm_{G/\Gamma}\right| \gg_d \delta^{O_d(1)},$$

where $m_{G/\Gamma}$ is the Haar measure of G/Γ. If $\xi = 0$, then we may invoke the discussion in [5, after (7.1)] to reduce the problem to the 1-step case, which can be addressed by Lemma 9.6. So, from now on, we assume that $\xi \neq 0$. Then, it follows that

$$|\mathbb{E}_{n\in\iota^{-1}(V(M))} F(g(n)\Gamma)| \gg_d \delta^{O_d(1)}.$$

For $h = (h_1, \dots, h_d) \in \mathbb{Z}^d$ and $a = (a_1, \dots, a_d) \in G^d$, denote $a^h := a_1^{h_1} \cdot \dots \cdot a_d^{h_d}$. Denote

$$\tilde{F}_h(x, y) := F(\{a^h\}x)\overline{F}(y) \quad \text{for all } x, y \in G/\Gamma.$$

Using the (multi-dimensional version of) Green–Tao argument in [5, §5], we may generalize the outline of the proof of Corollary 1.6 described in §5 to non-abelian nilmanifolds and conclude that

$$|\mathbb{E}_{n\in\iota^{-1}(V(M)^{\iota(h)})} F_h^{\square}(\tilde{a}_h^n \Gamma^{\square})| \gg_d \delta^{O_d(1)},$$

where $G^{\square} := \{(g, g') \in G^2\colon g^{-1}g' \in G_2\}$, $\Gamma^{\square} := \{(g, g') \in \Gamma^2\colon g^{-1}g' \in G_2\}$, $F_h^{\square}$ is the restriction of $\tilde{F}_h$ on $G^{\square}$, and

$$\tilde{a}_h^n := (\{a^h\}^{-1}a_1^{n_1}\{a^h\}, \dots, \{a^h\}^{-1}a_d^{n_d}\{a^h\}, a_1^{n_1}, \dots, a_d^{n_d}).$$

It is not hard to see that $\tilde{F}_h$ is invariant under $[G^{\square}, G^{\square}]$ and, thus, one can use the tools for 1-step nilmanifold (such as Lemma 9.6) to conclude that there exists a non-trivial type-I horizontal character $\eta_h\colon G^{\square} \to \mathbb{R}$ with $0 < \|\eta_h\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $P_h(n) := \eta_h \circ \tilde{a}_h^n \mod \mathbb{Z}$ is a constant on $V_p(M)^h$. By the pigeonhole principle, we may assume without loss of generality that η_h equals the same η for many h. Writing $P_h(n) = n \cdot (\eta \circ \tilde{a}_h^1) := n \cdot v_h\colon \mathbb{Z}^d \to \mathbb{R}$ for some $v_h \in \mathbb{Z}^d/p$, we have that

$$P_h \mod \mathbb{Z} \text{ is a constant on } \iota^{-1}(V(M)^{\iota(h)}) = \iota^{-1}(V(M) \cap \{n \in \mathbb{F}_p^d\colon 2n \cdot h + h \cdot h = 0\}). \tag{8.1}$$

Step 2: solving polynomial equations. For any linear polynomial $F\colon \mathbb{F}_p^d \to \mathbb{F}_p$, if it vanishes on $V(M)$, then it must vanish everywhere (this is because F is of the form

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

MR by Proposition 6.2, which forces $F = R \equiv 0$ by comparing the degrees of F and MR). Similarly, if a linear polynomial F vanishes on the intersection of $V(M)$ and the affine subspace $\{n \in \mathbb{F}_p^d : 2n \cdot h + h \cdot h = 0\}$, then the inheriting principle implies that F must vanish on this entire affine subspace. Translating this observation into the $\mathbb{Z}/p$-setting, the linearity of P_h and (8.1) implies that $P_h \bmod \mathbb{Z}$ must be a constant on $\iota^{-1}(\{n \in \mathbb{F}_p^d : 2n \cdot h + h \cdot h = 0\}) = \{n \in \mathbb{Z}^d : 2n \cdot h + h \cdot h \in p\mathbb{Z}\})$. This implies that the vector v_h is parallel to $(1/p)h \bmod \mathbb{Z}^d$. Unlike in the case of [5], here, we cannot conclude that $v_h \in p\mathbb{Z}^d$.

Step 3: analyzing the coefficients. Suppose that $\psi(a_i) = (\gamma_{i,1}, \gamma_{i,2}, *)$. Write $\gamma_i := (\gamma_{i,1}, \gamma_{i,2})$ and $\zeta_i := (-\gamma_{i,2}, \gamma_{i,1})$. Following the Green–Tao argument in [5, §5], one can conclude that the ith coordinate of v_h is equal to $k_1 \cdot \gamma_i + k_2\zeta_i \cdot \{\sum_{i=1}^d h_i\gamma_i\}$ for some $k_1 \in \mathbb{Z}^2$ and $k_2 \in \mathbb{Z}$ of complexities $\ll_{d,m} \delta^{-O_{d,m}(1)}$ which are not all zero (this is a multi-dimensional version of [5, equation before (5.7)]). Since v_h is parallel to $(1/p)h \bmod \mathbb{Z}^d$, one can conclude that

$$\begin{aligned}\left(k_1 \cdot \gamma_1 + k_2\zeta_1 \cdot \left\{\sum_{i=1}^d h_i\gamma_i\right\}, \ldots, k_1 \cdot \gamma_d + k_2\zeta_d \cdot \left\{\sum_{i=1}^d h_i\gamma_i\right\}\right)\\ - c_h h \in \mathbb{Z}^d \quad \text{for some } c_h \in \mathbb{Z}/p,\end{aligned} \tag{8.2}$$

whose general case is (9.26). The expression (8.2) differs from [6, Lemma 4.3] because of the appearance of the term $c_h h$. However, its generalization (to be proved in Proposition 9.8) allows us to deduce that there exists $k \in \mathbb{Z}^2$ with $0 < |k| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $k \cdot \gamma_i \in \mathbb{Z}$ for all $1 \le i \le d$. Let η' be the type-I horizontal character induced by k, then $\eta' \circ g \bmod \mathbb{Z}$ is a constant on $\mathbb{Z}^d$ and thus on $V_p(M)$.

Step 4: completion of the proof. The proof of Proposition 8.1 is already completed in Step 3. However, it is worth mentioning that for general polynomial sequences, due to the appearance of nonlinear terms, one still has to run an argument similar to [5, discussion after Claim 7.7] to complete the proof. □

9. *Proof of the main equidistribution theorem*

9.1. *Some preliminary reductions.* We prove Theorem 5.1 in this section. We start with a technical lemma on the connections between subspaces of $\mathbb{F}_p^d$ and those of $\mathbb{Z}^d$. Let $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of dimension d' and $L : \mathbb{F}_p^{d'} \to V$ be a bijective linear transformation. Let $\tilde{V}$ be the $\mathbb{Z}$-span of $\tau(L(e_1)), \ldots, \tau(L(e_{d'}))$ and $\tilde{L} : \mathbb{Z}^{d'} \to \tilde{V}$ be the bijective linear transformation given by

$$\tilde{L}(n_1, \ldots, n_{d'}) := n_1\tau(L(e_1)) + \cdots + n_{d'}\tau(L(e_{d'})).$$

In other words, $\tilde{L}/p$ is the regular lifting of L. Then, it is clear that $\iota \circ \tilde{L} = L \circ \iota$ and $\tilde{L} \circ \tau \equiv \tau \circ L \bmod p\mathbb{Z}^d$.

Let B be a subset of $V + c$. It is clear that there is a bijection between $L^{-1}(B - c)$ and B given by $z \mapsto L(z) + c$. However, the map $z \mapsto \tilde{L}(z) + \tau(c)$ is not necessarily a bijection between $\iota^{-1}(L^{-1}(B - c))$ and $\iota^{-1}(B)$. To connect the sets $\iota^{-1}(L^{-1}(B - c))$ and $\iota^{-1}(B)$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

by $\tilde{L}$, we define a map $\Phi_{L,c}\colon \mathbb{Z}^{d'} \times \mathbb{Z}^d \to \mathbb{Z}^d$ as

$$\Phi_{L,c}(z, w) := \tau(c) + \tilde{L}(z) + pw.$$

LEMMA 9.1. *For all $K, K' \in \mathbb{N}_+$, the map $\Phi_{L,c} \bmod pK\mathbb{Z}^d$ is a K'^d-to-1 map between $(\iota^{-1}(L^{-1}(B-c)) \cap [pK']^{d'}) \times [K]^d$ and $\iota^{-1}(B) \cap [pK]^d$.*

Proof. For convenience, denote $\Phi := \Phi_{L,c}$. Let $(z, w) \in \iota^{-1}(L^{-1}(B-c)) \times \mathbb{Z}^d$. Then,

$$\iota \circ \Phi(z, w) = \iota(\tau(c) + \tilde{L}(z) + pw) = c + \iota \circ \tilde{L}(z) = c + L \circ \iota(z) \in B.$$

So, the range of $\Phi \bmod pK\mathbb{Z}^d$ is contained in $\iota^{-1}(B) \cap [pK]^d$.

Now, for all $n \in \iota^{-1}(B)$, we may pick some $z' \in L^{-1}(B-c)$ such that $\iota(n) = L(z') + c$. Let $z = \tau(z')$. Then, z belongs to $\iota^{-1}(L^{-1}(B-c)) \cap [p]^{d'}$ and

$$\tilde{L}(z) + \tau(c) = \tilde{L} \circ \tau(z') + \tau(c) \equiv \tau \circ L(z') + \tau(c) = \tau \circ \iota(n) \equiv n \bmod pk\mathbb{Z}^d.$$

So, there exists $w \in \mathbb{Z}^d$ such that $n = \tilde{L}(z) + \tau(c) + pw$. Therefore, $\Phi(z, w \bmod K\mathbb{Z}^d) \equiv n \bmod pk\mathbb{Z}^d$. So, the map $\Phi \bmod pK\mathbb{Z}^d$ is surjective.

Finally, for $(z, w), (z', w') \in (\iota^{-1}(L^{-1}(B-c)) \cap [pK']^{d'}) \times [K]^d$, we have that $\Phi(z, w) \equiv \Phi(z', w') \bmod pK\mathbb{Z}^d$ if and only if $\tilde{L}(z - z') + p(w - w') \in pK\mathbb{Z}^d$. If this is the case, then $\tilde{L}(z - z') \in p\mathbb{Z}^d$ and thus, $\mathbf{0} = \iota \circ \tilde{L}(z - z') = L \circ \iota(z - z')$. Since L is injective, this implies that $\iota(z - z') = \mathbf{0}$ and thus, $z \equiv z' \bmod p\mathbb{Z}^{d'}$. Now, for any such $z - z'$ and any fixed w, there is a unique $w' \in [K]^d$ such that $(1/p)\tilde{L}(z - z') + (w - w') \in K\mathbb{Z}^d$. Since for each $z \in \mathbb{Z}^{d'}$, the number of $z' \in [pK']^d$ for which $z \equiv z' \bmod p\mathbb{Z}^{d'}$ is equal to K'^d. So, $\Phi \bmod pK\mathbb{Z}^d$ is a K'^d-to-1 map. □

Before proving Theorem 5.1, we start with some simplifications. Suppose first that the conclusion holds for $V + c = \mathbb{F}_p^d$. For the general case, let $L\colon \mathbb{F}_p^{d-r} \to V$ be a bijective linear transformation. Let $M'\colon \mathbb{F}_p^{d-r} \to \mathbb{F}_p$ be the quadratic map given by $M'(z) := M(L(z) + c)$. Then, it is clear that $L^{-1}((V(M) \cap (V + c)) - c) = V(M')$.

Let $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ be rational. By Corollary 3.19, we may assume that g is pK-periodic for some $K \in \mathbb{N}_+$. Then, for any function $F\colon G/\Gamma \to \mathbb{C}$, it follows from Lemma 9.1 (setting $B = V(M) \cap (V + c)$ and $K' = K$) that

$$\begin{aligned}\mathbb{E}_{n \in \iota^{-1}(V(M) \cap (V+c))} F(g(n)\Gamma) &= \mathbb{E}_{n \in \iota^{-1}(V(M) \cap (V+c)) \cap [pK]^d} F(g(n)\Gamma) \\ &= \mathbb{E}_{z \in \iota^{-1}(V(M')) \cap [pK]^{d-r}} \mathbb{E}_{w \in [K]^d} F(g(\tau(c) + \tilde{L}(z) + pw)\Gamma) \\ &= \mathbb{E}_{(z,w) \in \iota^{-1}(V(M'')) \cap [pK]^{2d-r}} F(g(\tau(c) + \tilde{L}(z) + pw)\Gamma),\end{aligned}$$

where $M''\colon \mathbb{F}_p^{2d-r} \to \mathbb{F}_p$ is the quadratic form given by $M''(z, w) := M'(z)$ and $\tilde{L}/p$ is the regular lifting of L. It is clear that all of M, M', and M'' have the same rank, and that the map $(z, w) \mapsto g(\tau(c) + \tilde{L}(z) + pw)$ is rational. So, by assumption, if $(g(n)\Gamma)_{n \in \iota^{-1}(V(M) \cap (V+c))}$ is not δ-equidistributed on G/Γ, then there exists a non-trivial type-I horizontal character η with $0 < \|\eta\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $\eta \circ g(\tau(c) + \tilde{L}(z) + pw) \bmod \mathbb{Z}$ is a constant for all $z \in \iota^{-1}(V(M'))$ and $w \in \mathbb{Z}^d$. By Lemma 9.1, we have that $\eta \circ g(n) \bmod \mathbb{Z}$ is a constant for all $n \in \iota^{-1}(V(M) \cap (V + c))$. So, in the rest of §9, we may make the following assumption.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Assumption 9.2. Assume $V + c = \mathbb{F}_p^d$, $r = 0$, and $\operatorname{rank}(M) \geq s + 13$.

Denote $d' = \operatorname{rank}(M)$. Suppose that we have already proved Theorem 5.1 for all quadratic forms of the form

$$M_{c,c',\lambda}(n_1, \ldots, n_d) := cn_1^2 + n_2^2 + \cdots + n_{d'}^2 + c'n_{d'+1} - \lambda \tag{9.1}$$

for some $c, c', \lambda \in \mathbb{F}_p$, $c \neq 0$. Let M be an arbitrary quadratic form of rank d'. By Lemma 4.3, there exist $c, c', \lambda \in \mathbb{F}_p$ and $v \in \mathbb{F}_p^d$ with $c \neq 0$ and a $d \times d$ invertible matrix R in $\mathbb{F}_p$ such that $M(n) = M_{c,c',\lambda}(nR + v)$.

For convenience, denote $M' := M_{c,c',\lambda}$. Then, $R(V(M) + vR^{-1}) = V(M')$. Let $g \in \operatorname{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ be rational. By Corollary 3.19, we may assume that g is pK-periodic for some $K \in \mathbb{N}_+$. Then, for any function $F\colon G/\Gamma \to \mathbb{C}$, it follows from Lemma 9.1 (setting $B = V(M)$ and $K' = K$) that

$$\begin{aligned}
&\mathbb{E}_{n \in \iota^{-1}(V(M))} F(g(n)\Gamma) = \mathbb{E}_{n \in \iota^{-1}(V(M)) \cap [pK]^d} F(g(n)\Gamma) \\
&\quad = \mathbb{E}_{z \in \iota^{-1}(V(M')) \cap [pK]^d} \mathbb{E}_{w \in [K]^d} F(g(\tau(-vR^{-1}) + \tilde{R^{-1}}(z) + pw)\Gamma) \\
&\quad = \mathbb{E}_{(z,w) \in \iota^{-1}(V(M'')) \cap [pK]^{2d}} F(g(\tau(-vR^{-1}) + \tilde{R^{-1}}(z) + pw)\Gamma),
\end{aligned}$$

where $M''\colon \mathbb{F}_p^{2d} \to \mathbb{F}_p$ is the quadratic form given by $M''(z, w) := M'(z)$ and $\tilde{R^{-1}}/p$ is the regular lifting of R^{-1}. It is clear that all of M, M', and M'' have the same rank, and that M'' is also of the form (9.1) (for the same d' but with dimension increased from d to $2d$). Moreover, the map $(z, w) \mapsto g(\tau(-vR^{-1}) + \tilde{R}^{-1}(z) + pw)$ is rational. So, by assumption, if $(g(n)\Gamma)_{n \in \iota^{-1}(V(M))}$ is not δ-equidistributed on G/Γ, then there exists a non-trivial type-I horizontal character η with $0 < \|\eta\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $\eta \circ g(\tau(-vR^{-1}) + \tilde{R^{-1}}(z) + pw) \bmod \mathbb{Z}$ is a constant for all $z \in \iota^{-1}(V(M'))$ and $w \in \mathbb{Z}^d$. By Lemma 9.1, we have that $\eta \circ g(n) \bmod \mathbb{Z}$ is a constant for all $n \in \iota^{-1}(V(M))$. So, in the rest of §9, we may make the following assumption.

Assumption 9.3. Assume $M = M_{c,c',\lambda}$ is given by (9.1).

Let $\psi\colon G \to \mathbb{R}^m$ denote the Mal'cev coordinate map associated to $\mathcal{X}$. Now, suppose that the conclusion holds under the additional assumption that $g(\mathbf{0}) = \mathrm{id}_G$, the *identity element* of G. For the general case, write $g(\mathbf{0}) = \{g(\mathbf{0})\}[g(\mathbf{0})]$. Let $g'(n) := \{g(\mathbf{0})\}^{-1} g(n) g(\mathbf{0})^{-1} \{g(\mathbf{0})\}$. Then, g' is rational by the Baker–Campbell–Hausdorff formula and $g(n)\Gamma = \{g(\mathbf{0})\}g'(n)\Gamma$ for all $n \in \mathbb{Z}^d$. Since $\|F(\{g(\mathbf{0})\}\cdot)\|_{\mathrm{Lip}} \leq O(1)\|F\|_{\mathrm{Lip}}$ by [5, Lemma A.5], if $(g(n)\Gamma)_{n \in \iota^{-1}(V(M))}$ is not δ-equidistributed on G/Γ, then $(g'(n)\Gamma)_{n \in \iota^{-1}(V(M))}$ is not $O(\delta)$-equidistributed on G/Γ. Since $g'(\mathbf{0}) = \mathrm{id}_G$, by assumption, there exists a non-trivial type-I horizontal character η with $0 < \|\eta\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that for all $n \in \iota^{-1}(V(M))$, $\eta \circ g'(n) \bmod \mathbb{Z}$ is a constant, and thus, $\eta \circ g(n) \bmod \mathbb{Z}$ is a constant. So, in the rest of §9, we may make the following assumption.

Assumption 9.4. Assume $g(\mathbf{0}) = \mathrm{id}_G$. As a result, we may further assume that $G = G_0 = G_1$.

Now, suppose that the conclusion holds under the additional assumption that $\psi(g(e_i))$ belongs to $[0, 1)^m$ for all $1 \leq i \leq d$, where e_i is the ith standard unit vector in $\mathbb{Z}^d$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

For the general case, write $g(e_i) = \{g(e_i)\}[g(e_i)]$ and let $g'(n) := g(n)\prod_{i=1}^{d}[g(e_i)]^{-n_i}$. Then, $\psi(g'(e_i)) \in [0,1)^m$ and g' is rational by the Baker–Campbell–Hausdorff formula. Moreover, for any type-I horizontal character η, since η vanishes on Γ, $\eta \circ g'$ differs from $\eta \circ g$ by a constant. This enables us to prove the general case of Theorem 5.1. So, in the rest of §9, we may make the following assumption.

Assumption 9.5. Assume $\psi(g(e_i)) \in [0,1)^m$ for all $1 \le i \le d$.

It is helpful to introduce some further convention of notation. Throughout this section, we assume that $p \gg_{d,m} \delta^{-O_{d,m}(1)}$. Denote

$$M_c(n_1, \dots, n_d) = cn_1^2 + n_2^2 + \cdots + n_{d'}^2,$$

which is the degree two terms of M, and let $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$ be a regular lifting of M and $\tilde{M}_c\colon \mathbb{Z}^d \to \mathbb{Z}/p$ be an homogeneous regular lifting of M_c (whose existence is guaranteed by Lemma 2.5). Then, $M(n) = \iota \circ (p\tilde{M}) \circ \tau(n)$ and $M_c(n) = \iota \circ (p\tilde{M}_c) \circ \tau(n)$ for all $n \in \mathbb{F}_p^d$. It is clear that $\iota^{-1}(V(M)) = V_p(\tilde{M})$.

Let A be the $d \times d\mathbb{F}_p$-valued matrix associated to M and M_c. Then, $\tau(A)$ is the $d \times d$ integer valued matrix associated to $\tilde{M}$ and $\tilde{M}_c$. With a slight abuse of notation, we identify the matrix A both as the $\mathbb{F}_p$-valued matrix and as the $[p]$-valued matrix $\tau(A)$.

Let $(G/\Gamma)_{\mathbb{N}}$ be a nilmanifold with filtration $(G_i)_{i\in\mathbb{N}}$. Denote $m := \dim(G)$, $m_i := \dim(G_i)$, $m_{ab} := \dim(G) - \dim([G,G])$ (the dimension of the Type-I horizontal torus of $(G/\Gamma)_{\mathbb{N}}$), $m_{\mathrm{lin}} := m - m_2$ (the *linear degree* of $(G/\Gamma)_{\mathbb{N}}$), and $m_* := m_{ab} - m_{\mathrm{lin}}$ (the *nonlinear degree* of $(G/\Gamma)_{\mathbb{N}}$).

9.2. *The case $s = 1$.* We first prove Theorem 5.1 for the case $s = 1$ (the case $s = 0$ is trivial). We make use of the following corollary of Lemma 4.12.

LEMMA 9.6. *Let $d, K \in \mathbb{N}_+$ and p be a prime number. Let $\tilde{M}\colon \mathbb{Z} \to \mathbb{Z}/p$ be a quadratic form of p-rank r. Suppose that $r \ge 3$. Then, for all $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{Z}^d \setminus pK\mathbb{Z}^d$, we have that*

$$\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} \exp\left(\frac{\xi}{pK}\cdot n\right) = O(p^{-(r-2)/2}).$$

Proof. Note that

$$\begin{aligned}\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} \exp\left(\frac{\xi}{pK}\cdot n\right) &= \mathbb{E}_{n\in V_p(\tilde{M})\cap[p]^d}\mathbb{E}_{m\in[K]^d} \exp\left(\frac{\xi}{pK}\cdot(n+pm)\right)\\ &= \mathbb{E}_{n\in V_p(\tilde{M})\cap[p]^d}\mathbb{E}_{m\in[K]^d} \exp\left(\frac{\xi}{pK}\cdot n\right)\mathbb{E}_{m\in[K]^d} \exp\left(\frac{\xi}{K}\cdot m\right).\end{aligned}$$

Since $\xi \in \mathbb{Z}^d$, it is not hard to see that $\mathbb{E}_{m\in[K]^d}\exp(\xi/K\cdot m) = 0$ unless $\xi = K\xi'$ for some $\xi' \in \mathbb{Z}^d$, in which case,

$$\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} \exp\left(\frac{\xi}{pK}\cdot n\right) = \mathbb{E}_{n\in V_p(\tilde{M})\cap[p]^d} \exp\left(\frac{\xi'}{p}\cdot n\right)$$

and thus, the conclusion follows from Lemma 4.12. □

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Since $s = 1$, we have $m_* = m$ and $G/\Gamma = \mathbb{R}^m/\mathbb{Z}^m = \mathbb{T}^m$. We may write

$$g(n) = (b_1 \cdot n, \dots, b_m \cdot n) + (v_1, \dots, v_m)$$

for some $K \in \mathbb{N}_+$, $b_i \in \mathbb{Z}^d/pK$, and $v_i \in \mathbb{R}$. Clearly, g is pK-periodic. If $(g(n)\Gamma)_{n\in V_p(\tilde{M})\cap[pK]^d}$ is not δ-equidistributed, then by Lemma 3.23, there exists a vertical frequency $\xi = (\xi_1, \dots, \xi_m) \in \mathbb{Z}^m$ with $|\xi| \ll_m \delta^{-O_m(1)}$ such that

$$\left|\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} \exp(\xi \cdot g(n)) - \int_{\mathbb{T}^m} \exp(\xi \cdot x)\, dm_{\mathbb{T}^m}(x)\right| \gg_m \delta^{O_m(1)}, \tag{9.2}$$

where $m_{\mathbb{T}^m}$ is the Haar measure on $\mathbb{T}^m$. So, we have that $\xi \neq \mathbf{0}$ as otherwise, the left-hand side of (9.2) is zero. Thus,

$$|\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} \exp(\xi \cdot g(n))| \gg_m \delta^{O_m(1)}.$$

Note that $\xi \cdot g(n) = (\xi B) \cdot n + \xi \cdot v$, where B is the $\mathbb{Z}/pK$-valued $m \times d$ matrix whose ith row is b_i. If $\xi B \notin \mathbb{Z}^d$, then since $\operatorname{rank}_p(M) \geq 3$, Lemma 9.6 implies that

$$|\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} \exp((\xi B) \cdot n)| = O(p^{-(d-2)/2}),$$

which is impossible if $p \gg_m \delta^{-O_m(1)}$. So, we must have that $\xi B \in \mathbb{Z}^d$, which implies that $\xi \circ g \bmod \mathbb{Z}$ is a constant. This finishes the proof.

9.3. *Construction of a nilsequence in a joining system.* We now assume that $s \geq 2$ (and thus, $d' \geq s + 13$) and that Theorem 5.1 holds for smaller s (and any m_*), and for the same value of s and a smaller value of m_*. The readers are advised to familiarize themselves with the machinery in [**5**, §8] before reading this section in detail. We follow the outline of the proof of Proposition 8.1 in §8. Following Step 1 in §8, we first run the argument of [**5**] to reduce the problem to a polynomial equation (9.18). By Corollary 3.19, we may assume that g is pK-periodic for some $K \in \mathbb{N}_+$. By Lemma 3.23, if $(g(n)\Gamma)_{n\in V_p(\tilde{M})\cap[pK]^d}$ is not δ-equidistributed on G/Γ, then there exist a vertical frequency $\xi \in \mathbb{Z}^{m_s}$ with $\|\xi\| \ll_m \delta^{-O_m(1)}$, and a function $F\colon G/\Gamma \to \mathbb{C}$ with $\|F\|_{\mathrm{Lip}} \leq 1$ and vertical frequency ξ such that

$$\left|\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} F(g(n)\Gamma) - \int_{G/\Gamma} F\, dm_{G/\Gamma}\right| \gg_m \delta^{O_m(1)}, \tag{9.3}$$

where $m_{G/\Gamma}$ is the Haar measure of G/Γ. If $\xi = \mathbf{0}$, then we may invoke the discussion in [**5**, after (7.1)] to reduce the problem to the $s - 1$ degree case and then use the induction hypothesis to finish the proof. So, from now on, we assume that $\xi \neq \mathbf{0}$.

Since F is of vertical frequency ξ, (9.3) implies that

$$|\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} F(g(n)\Gamma)| \gg_m \delta^{O_m(1)}. \tag{9.4}$$

By taking squares on both sides of (9.4), we have that

$$\begin{aligned}\delta^{O_m(1)} &\ll_m \mathbb{E}_{n,m\in V_p(\tilde{M})\cap[pK]^d} F(g(m)\Gamma)\overline{F}(g(n)\Gamma)\\ &= \mathbb{E}_{(n,h)\in\square_{p,1}(V_p(\tilde{M}))\cap[pK]^{2d}} F(g(n+h)\Gamma)\overline{F}(g(n)\Gamma),\end{aligned}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

where we used Lemma 7.2 and the fact that $\square_{p,1}(V_p(\tilde{M}))$ is a p-periodic set. By Lemma 4.12, $|\square_{p,1}(V_p(\tilde{M})) \cap [p]^d| = |V_p(\tilde{M}) \cap [p]^d|^2 = p^{2d-2}(1 + O(p^{-1/2}))$. So, this implies that

$$\frac{1}{|V_p(\tilde{M}) \cap [pK]^d|^2} \sum_{h \in V_p(\tilde{M}) \cap [pK]^d} \sum_{n \in V_p(\tilde{M})^h \cap [pK]^d} F(g(n+h)\Gamma)\overline{F}(g(n)\Gamma) \gg_m \delta^{O_m(1)}. \tag{9.5}$$

Since $d \geq 5$, it follows from Proposition 4.19 that

$$\mathbb{E}_{h \in [pK]^d} |\mathbb{E}_{n \in V_p(\tilde{M})^h \cap [pK]^d} F(g(n+h)\Gamma)\overline{F}(g(n)\Gamma)| \gg_m \delta^{O_m(1)}.$$

By the pigeonhole principle, there exists a subset J_1 of $[pK]^d$ of cardinality $\gg_m \delta^{O_m(1)}(pK)^d$ such that

$$|\mathbb{E}_{n \in V_p(\tilde{M})^h \cap [pK]^d} F(g(n+h)\Gamma)\overline{F(g(n)\Gamma)}| \gg_m \delta^{O_m(1)} \tag{9.6}$$

for all $h \in J_1$. Recall that A is the $d \times d\mathbb{Z}$-valued matrix associated to M. Since $d' \geq 3$, by Lemmas 4.12 and 7.2, passing to a subset if necessary, we may assume that

$$(hA) \cdot h, h_1, \ldots, h_d \notin p\mathbb{Z} \quad \text{for all } h = (h_1, \ldots, h_d) \in J_1. \tag{9.7}$$

It is helpful to view $V_p(\tilde{M})^h$ as the preimage of ι of the intersection of the hyperspheres $V(M)$ and $V(M(\cdot + h))$. We will use the inheriting principle to apply the induction hypothesis on these sub-hyperspheres. Recall that $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ is the quadratic form induced by $\tilde{M}$ and $V_p(\tilde{M})^h = \iota^{-1}(V(M)^{\iota(h)})$. For $h \in J_1$, note that $V(M)^{\iota(h)} = V(M) \cap U_h$, where $U_h := \{n \in \mathbb{F}_p^d\colon M(n + \iota(h)) = M(n)\}$. Then, we may write $U_h = V_h + u_h$ for some $u_h \in \mathbb{F}_p^d$, where $V_h := \{n \in \mathbb{F}_p^d\colon (nA) \cdot \iota(h) = 0\}$. It follows from (9.6) that V_h is a subspace of $\mathbb{F}_p^d$ of co-dimension 1. Let $L_h\colon \mathbb{F}_p^{d-1} \to V_h$ be any bijective linear transformation, and let $M_h\colon \mathbb{F}_p^{d-1} \to \mathbb{F}_p$ be the quadratic form given by $M_h(z) := M(L_h(z) + u_h)$ and let $\tilde{M}_h\colon \mathbb{Z}^{d-1} \to \mathbb{Z}/p$ be the quadratic form induced by M_h. Since $(hA) \cdot h \notin p\mathbb{Z}$, we have that

$$\begin{aligned} \operatorname{rank}_p(\tilde{M}_h) &= \operatorname{rank}(M_h) = \operatorname{rank}(M|_{\operatorname{span}_{\mathbb{F}_p}\{\iota(h)\}^{\perp_M}}) \\ &= \operatorname{rank}(M) - 1 \geq (s-1) + 13 \quad \text{for all } h \in J_1. \end{aligned} \tag{9.8}$$

Let $\tilde{L}_h/p$ be the regular lifting of L_h. Then,

$$(hA) \cdot \tilde{L}_h(z) \in p\mathbb{Z} \quad \text{for all } z \in \mathbb{Z}^{d-1}. \tag{9.9}$$

Since $V(M)^{\iota(h)} = V(M) \cap (V_h + u_h)$, it is not hard to see that $L_h^{-1}(V(M)^{\iota(h)} - u_h) = V(M_h)$ and $V_p(\tilde{M}_h) = \iota^{-1}(V(M_h))$. Since g is pK-periodic, by (9.6) and Lemma 9.1 (setting $B = V(M)^{\iota(h)}$ and $K' = K$), the average

$$|\mathbb{E}_{(z,m) \in (V_p(\tilde{M}_h) \cap [pK]^{d-1}) \times [pK]^d} F(g(pm + \tilde{L}_h(z) + \tau(u_h) + h)\Gamma)\overline{F(g(pm + \tilde{L}_h(z) + \tau(u_h))\Gamma)}| \tag{9.10}$$

is at least $\gg_m \delta^{O_m(1)}$ for all $h \in J_1$.

We decompose g as $g = g_{\text{nlin}} g_{\text{lin}}$, where

$$g_{\text{lin}}(n_1, \ldots, n_d) := g(e_1)^{n_1} \cdots g(e_d)^{n_d}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

is the *linear component* of g and

$$g_{\mathrm{nlin}}(n) := g(n)g_{\mathrm{lin}}(n)^{-1}$$

is the *nonlinear component* of g, where e_i is the ith standard unit vector in $\mathbb{Z}^d$. Since $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$, we have that $g_{\mathrm{lin}}, g_{\mathrm{nlin}} \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ by [8, Corollary B.4].

Note that $g_{\mathrm{nlin}}(\mathbf{0}) = g_{\mathrm{nlin}}(e_1) = \cdots = g_{\mathrm{nlin}}(e_d) = \mathrm{id}_G$, and g_{nlin} takes values in G_2 by the Baker–Campbell–Hausdorff formula. For all $h \in \mathbb{Z}^d$, let $g_h \colon \mathbb{Z}^d \to G^2$ be the sequence

$$g_h(n) := (\{g_{\mathrm{lin}}(h)\}^{-1}g(n+h)[g_{\mathrm{lin}}(h)]^{-1}, g(n)), \tag{9.11}$$

and $F_h \colon G^2/\Gamma^2 \to \mathbb{C}$ be the function

$$F_h(x, y) := F(\{g_{\mathrm{lin}}(h)\}x)\overline{F}(y).$$

Then, (9.10) implies that

$$|\mathbb{E}_{(z,m)\in(V_p(\tilde{M}_h)\cap[pK]^{d-1})\times[pK]^d} F_h(g_h(pm + \tilde{L}_h(z) + \tau(u_h))\Gamma^2)| \gg_m \delta^{O_m(1)} \tag{9.12}$$

for all $h \in J_1$.

For groups G and H with $H < G$, denote

$$G \times_H G := \{(g, h) \in G \times G \colon g^{-1}h \in H\}.$$

Note that $G \times_H G$ is a group as long as $[G, H] \subseteq H$. It is not hard to check that g_h takes values in $G^\square := G \times_{G_2} G$ for all $h \in \mathbb{Z}^d$ (see the argument in [5, after (7.6)]). So (9.12) implies that

$$|\mathbb{E}_{(z,m)\in(V_p(\tilde{M}_h)\cap[pK]^{d-1})\times[pK]^d} F_h^\square(g_h^\square(pm + \tilde{L}_h(z) + \tau(u_h))\Gamma^\square)| \gg_m \delta^{O_m(1)} \tag{9.13}$$

for all $h \in J_1$, where $F_h^\square$, $g_h^\square$, and $\Gamma^\square$ are the restrictions of F_h, g_h, and Γ^2 to $G^\square$, respectively. By repeating the argument in [5, after (7.7)], $F_h^\square$ is invariant under $G_s^\triangle := \{(g_s, g_s) \colon g_s \in G_s\}$. Thus, $F_h^\square$ induces a function $\overline{F_h^\square}$ on $\overline{G^\square} := G^\square/G_s^\triangle$. By (9.13), we have that

$$|\mathbb{E}_{(z,m)\in(V_p(\tilde{M}_h)\cap[pK]^{d-1})\times[pK]^d} \overline{F_h^\square}(\overline{g_h^\square}(pm + \tilde{L}_h(z) + \tau(u_h))\overline{\Gamma^\square})| \gg_m \delta^{O_m(1)}, \tag{9.14}$$

where $\overline{\Gamma^\square} := \Gamma^\square/(\Gamma^\square \cap G_k^\triangle)$ and $\overline{g_h^\square}$ is the projection of $g_h^\square$ to $\overline{G^\square}$.

By [6, Proposition 4.2] and [5, Lemma 7.4], we have that:

- the groups $G^\square$ and $\overline{G^\square}$ are connected, simply connected nilpotent Lie groups of degree $s-1$ with respect to the $\mathbb{N}$-filtrations given by $(G^\square)_i := G_i \times_{G_{i+1}} G_i$ and $(\overline{G^\square})_i := (G_i \times_{G_{i+1}} G_i)/G_k^\triangle$, respectively;
- there exist $O_{m,s}(\delta^{-O_{m,s}(1)})$-rational Mal'cev basis $\mathcal{X}^\square$ for $G^\square/\Gamma^\square$, and $\overline{\mathcal{X}^\square}$ for $\overline{G^\square}/\overline{\Gamma^\square}$, adapted to the filtrations $(G^\square)_\mathbb{N}$ and $(\overline{G^\square})_\mathbb{N}$, respectively;
- the Mal'cev coordinate map $\psi_{\mathcal{X}^\square}(x, x')$ of $\mathcal{X}^\square$ is a polynomial of degree $O_{m,s}(1)$ with rational coefficients of complexity $O_{m,s}(\delta^{-O_{m,s}(1)})$ in the coordinates $\psi(x)$ and $\psi(x')$;
- under the metrics induced by $\mathcal{X}^\square$ and $\overline{\mathcal{X}^\square}$, we have that $\|F_h^\square\|_{\mathrm{Lip}}, \|\overline{F_h^\square}\|_{\mathrm{Lip}} \ll_{m,s} (\delta^{-O_{m,s}(1)})$;

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

- $g_h^{\square}$ belongs to $\mathrm{poly}(\mathbb{Z}^d \to (G^{\square})_{\mathbb{N}})$ and $\overline{g_h^{\square}}$ belongs to $\mathrm{poly}(\mathbb{Z}^d \to (\overline{G^{\square}})_{\mathbb{N}})$.

Since g is rational and pK-periodic, it is not hard to see that:

- $g_h^{\square}$ and $\overline{g_h^{\square}}$ are rational and pK-periodic.

Let $\tilde{M}_h' \colon \mathbb{Z}^{2d-1} \to \mathbb{Z}/p$ be the quadratic form given by $\tilde{M}_h'(z, m) := \tilde{M}_h(z)$. It is clear from (9.8) that $\mathrm{rank}_p(\tilde{M}_h') = \mathrm{rank}_p(\tilde{M}_h) \geq (s-1)+13$ for all $h \in J_1$. So, by (9.14) and the induction hypothesis, for all $h \in J_1$, there exist a type-I horizontal character $\overline{\eta}_h \colon \overline{G^{\square}} \to \mathbb{R}$ with $0 < \|\overline{\eta}_h\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ and a constant $a_h \in \mathbb{R}$ such that for all $m \in \mathbb{Z}^d$ and $z \in V_p(\tilde{M}_h)$,

$$\overline{\eta}_h \circ \overline{g_h^{\square}}(pm + \tilde{L}_h(z) + \tau(u_h)) \equiv a_h \bmod \mathbb{Z}.$$

By the pigeonhole principle, there exists a subset J_2 of J_1 with $|J_2| \gg_{d,m} \delta^{O_{d,m}(1)}(pK)^d$ such that $\overline{\eta}_h$ is equal to a same $\overline{\eta}$ for all $h \in J_2$. Writing $\eta_0 \colon G^{\square} \to \mathbb{R}$ to be the type-I horizontal character $\eta_0(x) = \overline{\eta}(\overline{x})$, where $\overline{x}$ is the projection of x, we have that η_0 is a type-I horizontal character with $0 < \|\eta_0\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ and that

$$\eta_0 \circ g_h^{\square}(pm + \tilde{L}_h(z) + \tau(u_h)) \equiv a_h \bmod \mathbb{Z} \quad \text{for all } h \in J_2, m \in \mathbb{Z}^d, z \in V_p(\tilde{M}_h). \tag{9.15}$$

We apply Lemma 9.1 to simplify this condition. Denote $\Phi_{L_h,u_h}(z, m) := pm + \tilde{L}_h(z) + \tau(u_h)$. Applying Lemma 9.1 with $d' = d-1$, $K' = 1$, $V = V_h$, $c = u_h$, and $B = V(M)^{\iota(h)}$, we have that $\Phi_{L_h,u_h} \bmod p\mathbb{Z}^d$ is a bijection between $\iota^{-1}((L_h^{-1}(V(M)^{\iota(h)} - u_h)) \cap [p]^{d-1}) \times [K]^d$ and $\iota^{-1}(V(M)^{\iota(h)}) \cap [pK]^d$, which is between $V_p(\tilde{M}_h) \times [K]^d$ and $V_p(\tilde{M})^h \cap [pK]^d$ (recall the discussion between (9.9) and (9.10)). Since $g_h^{\square}$ is pK-periodic, by (9.15),

$$\eta_0 \circ g_h^{\square}(n) \equiv a_h \bmod \mathbb{Z} \quad \text{for all } h \in J_2, n \in V_p(\tilde{M})^h. \tag{9.16}$$

Finally, since g is pK-periodic, $(g_{h+pKm}^{\square})^{-1} g_h^{\square}$ take values in Γ^2 for all $h \in J_2$ and $m \in \mathbb{Z}^d$. So $\eta_0 \circ g_h^{\square} = \eta_0 \circ g_{h+pKm}^{\square} \bmod \mathbb{Z}$. Therefore, we may upgrade (9.16) to conclude that

$$\eta_0 \circ g_h^{\square}(n) \equiv a_h \bmod \mathbb{Z} \quad \text{for all } h \in J_2 + pK\mathbb{Z}^d, n \in V_p(\tilde{M})^h \tag{9.17}$$

for some constant a_h.

Denote $\eta_1 \colon G \to \mathbb{R}$, $\eta_1(g) := \eta_0(g, g)$ and $\eta_2 \colon G_2 \to \mathbb{R}$, $\eta_2(g) = \eta_0(g, \mathrm{id}_G)$, we have that $\eta_0(g', g) = \eta_1(g) + \eta_2(g'g^{-1})$ for all $(g', g) \in G^{\square}$. Similar to [5, Lemma 7.5], η_1 is a type-I horizontal character on G, and η_2 is a type-I horizontal character on G_2 which annihilates $[G, G_2]$. Moreover, $\|\eta_1\|, \|\eta_2\| \ll_{d,m} \delta^{-O_{d,m}(1)}$.

Similar to the approach [6, between (4.8) and (4.10)] (along with some notation changes), one can compute that (we leave the details to the interested readers)

$$Q(n+h) - Q(n) + P(n) + \sigma(h) \cdot n - a_h \equiv 0 \bmod \mathbb{Z} \tag{9.18}$$

for all $h \in J_2 + pK\mathbb{Z}^d$ and $n \in V_p(\tilde{M})^h$, where

$$P(n) := \eta_1(g(n)),\ Q(n) := \eta_2(g_{\mathrm{nlin}}(n)),\ \sigma(h) := \omega_h + h\Lambda,$$
$$\omega_h := (\eta_2([g(e_1), \{g_{\mathrm{lin}}(h)\}]), \ldots, \eta_2([g(e_d), \{g_{\mathrm{lin}}(h)\}])) \in \mathbb{R}^d,$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

and $\Lambda = (\lambda_{i,j})_{1\le i,j\le d}$ is the $d \times d$ matrix in $\mathbb{R}$ with $\lambda_{i,j} = \eta_2([g(e_i), g(e_j)])$ if $i < j$ and $\lambda_{i,j} = 0$ otherwise. Since g is rational, it is clear that all the entries of Λ are in $\mathbb{Q}$.

9.4. *Solving polynomial equations.* Following Step 2 in §8, we now solve (9.18). Note that (9.18) is more complicated than the equation in the Heisenberg case (8.1) due to the appearance of the nonlinear terms, and than [**6**, (4.10)] since (9.18) only holds on some hyperspheres. As a result, this step is the most challenging part in adapting the Green–Tao approach, which is also the reason why the efforts in §§6 and 7 are needed.

Since g is rational and is of step at most s, we have that P and Q are of degree at most s and take values in $\mathbb{Q}$. Replacing n in (9.18) by n, $n + h_2$, $n + h_3$, $n + h_2 + h_3$, respectively, replacing h in (9.18) by h_1, and taking the second-order differences, we have that

$$\Delta_{h_3}\Delta_{h_2}\Delta_{h_1} Q(n) + \Delta_{h_2}\Delta_{h_1} P(n) \in \mathbb{Z} \tag{9.19}$$

for all $h_3 \in J_2 + pK\mathbb{Z}^d$ and $(n, h_1, h_2) \in \square_{p,2}(V_p(\tilde{M})^{h_3})$. By Proposition 4.19 (and a change of variables), it is not hard to compute that the number of $(n, h_1, h_2, h_3) \in \square_{p,3}(V_p(\tilde{M})) \cap ([pK]^d)^4$ with $h_3 \in J_2$ is $\gg_{d,m} \delta^{-O_{d,m}(1)}|\square_{p,3}(V_p(\tilde{M})) \cap ([pK]^d)^4|$. Since $d' \ge s + 13 \ge 3^2 + 3 + 3$, it follows from (9.19) and Proposition 7.6 (if K is not divisible by p, then we may replace the set J_2 by $(J_2 + pK\mathbb{Z}^d) \cap [p^2K]^d$ to apply Proposition 7.6) that

$$Q = \frac{1}{q}Q_1 + Q_0 \quad \text{and} \quad P = \frac{1}{q}P_1 + P_0 \tag{9.20}$$

for some $q \in \mathbb{N}_+$ with $q \ll_d \delta^{-O_d(1)}$, some polynomials $Q_1, P_1 \in \mathrm{poly}(V_p(\tilde{M}) \to \mathbb{R}|\mathbb{Z})$ with $\deg(Q_1) \le \deg(Q)$, $\deg(P_1) \le \deg(P)$, and some $Q_0, P_0 \colon \mathbb{Z}^d \to \mathbb{R}$ such that

$$Q_0(n) = (n\Omega)\cdot n + v \cdot n + v_0 \quad \text{and} \quad P_0(n) = u \cdot n + u_0 \tag{9.21}$$

for some $v, u \in \mathbb{R}^d$, $v_0, u_0 \in \mathbb{R}$, and symmetric $d \times d\mathbb{Q}$-valued matrix Ω. (If P and Q were $\mathbb{Z}/p$-valued polynomials, then we could apply Corollary 7.4 instead of Proposition 7.6 and simplify the proof. However, this is not the case in general. This is the reason why we need to improve Corollary 7.4 to the more general version Proposition 7.6.)

9.5. *Analyzing the low degree terms.* Following Step 3 in §8, our next goal is to derive better descriptions for the lower degree polynomials Q_0 and P_0 by substituting (9.20) and (9.21) back to (9.18). For all $h \in J_2 + pK\mathbb{Z}^d$ and $n \in V_p(\tilde{M})^h$, it follows from (9.20) that

$$Q(n) - Q_0(n) \equiv Q(n+h) - Q_0(n+h) \equiv P(n) - P_0(n) \equiv 0 \mod \frac{1}{q}\mathbb{Z}.$$

So (9.18) implies that

$$t_h \cdot n \equiv b_h \bmod \mathbb{Z} \quad \text{for all } h \in J_2 + pK\mathbb{Z}^d \quad \text{and} \quad n \in V_p(\tilde{M})^h \tag{9.22}$$

for some $b_h \in \mathbb{R}$, where

$$t_h := q(2h\Omega + u + \sigma(h)). \tag{9.23}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Note that the conditions $n \in V_p(\tilde{M})^h$ remain unchanged if we replace n by $n + pn'$ for any $n' \in \mathbb{Z}^d$. So, the condition $t_h \cdot n \equiv b_h \bmod \mathbb{Z}$ is also unaffected if we change n to $n + pn'$. Therefore, we must have that $t_h \in \mathbb{Z}^d/p$, $b_h \in \mathbb{Z}/p$ for all $h \in J_2 + pK\mathbb{Z}^d$.

This is another place where our approach deviates from that of [5] (in which case, one can easily conclude that $t_h \in \mathbb{Z}^d$). Recall that the ι-image of the set $V_p(\tilde{M})^h$ is the intersection of the hypersphere $V(M)$ and an affine subspace of co-dimension at most 1 which is orthogonal to hA. Therefore, we may expect to deduce from (9.22) that the ι image of t_h is parallel to hA. To be more precise, we now show that

$$t_h - c_h(hA) \in \mathbb{Z}^d \quad \text{for some } c_h \in \mathbb{Z}/p \text{ for all } h \in J_2 + pK\mathbb{Z}^d. \tag{9.24}$$

Note that (9.22) implies that the $\iota(t_h)$ is orthogonal to the $\iota(V_p(\tilde{M})^h - V_p(\tilde{M})^h)$. So, (9.24) follows if one can show that the span of $\iota(V_p(\tilde{M})^h - V_p(\tilde{M})^h)$ is the orthogonal complement of $\iota(hA)$. To do so, we fix $h \in J_2 + pK\mathbb{Z}^d$ and identify $V_p(\tilde{M})^h$ by $V_p(\tilde{M}_h)$ via the linear transformation $\tilde{L}_h$. The desired spanning property can be restated as follows.

CLAIM. *There exist $w_0, \dots, w_{d-1} \in V_p(\tilde{M}_h)$ such that $w_1 - w_0, \dots, w_{d-1} - w_0$ are p-linearly independent.*

Proof of Claim. Recall that $M_h \colon \mathbb{F}_p^d \to \mathbb{F}_p$ is the quadratic form induced by $\tilde{M}_h$ and by (9.8) $\operatorname{rank}(M_h) \geq 3$. Translating everything in the $\mathbb{F}_p^d$-setting, it suffices to show that there exist $v_0, \dots, v_{d-1} \in V(M_h)$ such that $v_1 - v_0, \dots, v_{d-1} - v_0$ are linearly independent. Fix any $v_0 \in V(M_h)$ (whose existence is given by Lemma 4.12). Suppose that for some $0 \leq k \leq d-2$, we have chosen $v_1, \dots, v_k \in V(M_h)$ such that $v_1 - v_0, \dots, v_k - v_0$ are linearly independent. If $k \leq d-3$, then the number of $v \in \mathbb{F}_p^d$ with $v_1 - v_0, \dots, v_k - v_0, v - v_0$ being linearly dependent is at most $p^k \leq p^{d-3}$, while $|V(M_h)| = p^{d-2}(1 + O(p^{-1/2}))$. If $p \gg_d 1$, then there exists $v_{k+1} \in V(M_h)$ such that $v_1 - v_0, \dots, v_{k+1} - v_0$ are linearly independent. If $k = d-2$, then let $Q(v)$ denote the determinant of the $(d-1) \times (d-1)$ matrix with $v_1 - v_0, \dots, v_{d-2} - v_0, v - v_0$ being the row vectors. Then, Q is a non-constant degree 1 polynomial in v. So, Q cannot be written in the form $Q = Q' M_h$ for some $Q' \in \operatorname{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ with $\deg(Q') \leq \deg(Q)$. By Proposition 6.2, there exists $v_{d-1} \in V(M_h) \backslash V(Q)$. So, $v_1 - v_0, \dots, v_{d-1} - v_0$ are linearly dependent. This completes the proof of the claim. □

We now continue the proof of (9.24). Let $w_0, \dots, w_{d-1} \in V_p(\tilde{M}_h)$ be given by the claim. Then, for all $1 \leq i \leq d-1$, it follows from (9.22) that

$$t_h \cdot (\tilde{L}_h(w_i - w_0)) = t_h \cdot ((\tilde{L}_h(w_i) + u_h) - (\tilde{L}_h(w_0) + u_h)) \equiv b_h - b_h = 0 \bmod \mathbb{Z}.$$

Since $w_1 - w_0, \dots, w_{d-1} - w_0$ are p-linearly independent, by linearity, there exists $q_h \in \mathbb{N}$, $p \nmid q_h$ such that $q_h t_h \cdot \tilde{L}_h(e_i) \in \mathbb{Z}$ for all $1 \leq i \leq d-1$. Since $t_h \in \mathbb{Z}^d/p$, we have

$$t_h \cdot \tilde{L}_h(e_i) \in \frac{1}{q_h}\mathbb{Z} \cap \frac{1}{p}\mathbb{Z} = \mathbb{Z}. \tag{9.25}$$

Recall that $h_1 \notin p\mathbb{Z}$ since $h \in J_2 + pK\mathbb{Z}^d$. Since $t_h \in \mathbb{Z}^d/p$, we may write $t_h = c_h(hA) + (0, t_h')$ for some $c_h \in \mathbb{Z}/p$ and $t_h' \in \mathbb{Z}^{d-1}/p$. Since $(hA) \cdot \tilde{L}_h(e_i) \in p\mathbb{Z}$ for all $1 \leq i \leq d-1$ by (9.9), it follows from (9.25) that $t_h' B \in \mathbb{Z}^{d-1}$, where B is the

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$(d-1)\times(d-1)$ matrix whose column vectors are transposes of $\tilde{L}_h(e_1),\dots,\tilde{L}_h(e_{d-1})$ with the first coefficient removed.

If $\tilde{L}_h(e_1),\dots,\tilde{L}_h(e_{d-1}),(1,0,\dots,0)$ are p-linearly dependent, then the fact that $(hA)\cdot\tilde{L}_h(e_i)\in p\mathbb{Z}$, $1\le i\le d-1$ implies that $(hA)\cdot\tilde{L}_h(1,0,\dots,0)=\tau(c)h_1\in p\mathbb{Z}$, which is a contradiction. Therefore, $\tilde{L}_h(e_1),\dots,\tilde{L}_h(e_{d-1}),(1,0,\dots,0)$ are p-linearly independent and thus, $\det(B)\notin p\mathbb{Z}$. So, $t_h'\in(1/q_h')\mathbb{Z}^{d-1}$ for some $q_h'\in\mathbb{N}$, $p\nmid q_h'$. Since $t_h'\in(1/p)\mathbb{Z}^{d-1}$, we have that $t_h'\in\mathbb{Z}^{d-1}$. This completes the proof of (9.24).

Next, we need to take a closer look at (9.24) by unpacking the definition of $\sigma(h)$. Since the map $(b,c)\to\eta_2([b,c])$ is bilinear, the map $c\to\eta_2([g(e_i),c])$ is a homomorphism. So, there exists $\xi=(\xi_1,\dots,\xi_d)\in(\mathbb{R}^m)^d$ such that

$$\eta_2([g(e_i),x])=\xi_i\cdot\psi(x)\bmod\ \mathbb{Z}$$

for all $1\le i\le d$ and $x\in G$. Similar to [5, discussion after (7.15)], all but the first m_{lin} coefficients of ξ_i are non-zero, and $|\xi_i|\ll_{d,m}\delta^{-O_{d,m}(1)}$. Denote $\gamma_i:=\psi(g(e_i))$ and $\gamma:=(\gamma_1,\dots,\gamma_d)\in(\mathbb{R}^m)^d$. Since g is rational, it is clear that $\xi,\gamma\in(\mathbb{Q}^m)^d$.

Denote $\gamma(h):=\sum_{j=1}^d\gamma_j h_j\in\mathbb{Q}^m$. Let $\{\gamma(h)\}\in[0,1)^m$ be the coordinate-wise fractional part of $\gamma(h)$. Then,

$$\sigma(h)=h\Lambda+\xi*\{\gamma(h)\},$$

where $*$ is the coordinate-wise product of vectors in $\mathbb{R}^m$. So for all $h\in J_2+pK\mathbb{Z}^d$, it follows from (9.23) and (9.24) that

$$q(hB+u+\xi*\{\gamma(h)\})-c_h(hA)\in\mathbb{Z}^d\quad\text{for some } c_h\in\mathbb{Z}/p,\tag{9.26}$$

where $B:=2\Omega+\Lambda$.

We need the following lemma to study (9.26).

LEMMA 9.7. *Let $d,m,N\in\mathbb{N}_+$, p be a prime, $\delta>0,\beta,\alpha_1,\dots,\alpha_d\in\mathbb{R}$, and $\xi,\gamma_0,\gamma_1,\dots,\gamma_d\in\mathbb{Q}^m$. Let V be a subset of $\{-N,\dots,N\}^d$ of cardinality at least $\delta(2N+1)^d$ such that*

$$\beta+\sum_{i=1}^d\alpha_i h_i+\xi\cdot\left\{\gamma_0+\sum_{i=1}^d\gamma_i h_i\right\}\in\mathbb{Z}$$

for all $h=(h_1,\dots,h_d)\in V$. If $N\gg_{\delta,d,m,p}1$ (we caution the readers that N is dependent on p. However, this will not cause any trouble for the application of this lemma) then at least one of the following holds:

(i) *there exists $r\in\mathbb{Z}$ with $0<r\ll_{d,m}\delta^{-O_{d,m}(1)}$ such that $r\xi\in\mathbb{Z}^m$;*

(ii) *there exists $k\in\mathbb{Z}^m$ with $0<|k|\ll_{d,m}\delta^{-O_{d,m}(1)}$ such that $k\cdot\gamma_i\in\mathbb{Z}$ for all $1\le i\le d$.*

Proof. Lemma 9.7 generalizes [6, Lemma 4.3] in two aspects. The first is that Lemma 9.7 allows the appearance of the term γ_0. The second is that we do not impose any bound for the complexity of ξ. The strategy of the proof is to reduce the problem to the case when γ_0 is zero and ξ is bounded in the unit box, and then apply [6, Lemma 4.3].

Since

$$\left\{\gamma_0\right\}+\left\{\sum_{i=1}^d\gamma_i h_i\right\}-\left\{\gamma_0+\sum_{i=1}^d\gamma_i h_i\right\}\tag{9.27}$$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

takes at most 2^m different values, passing to a subset of V if necessary, we may assume without loss of generality that (9.27) take the same value for all $h \in V$. Absorbing constant terms into β if necessary, we may assume without loss of generality that $\gamma_0 = \mathbf{0}$.

Since

$$a\{b\} = \{a\}\{b\} + \lfloor a \rfloor b - \lfloor a \rfloor \lfloor b \rfloor$$

for all $a, b \in \mathbb{R}$, we have that

$$\beta + \sum_{i=1}^{d} \alpha_i h_i + \xi \cdot \left\{ \sum_{i=1}^{d} \gamma_i h_i \right\} \equiv \beta + \sum_{i=1}^{d} (\alpha_i + \lfloor \xi \rfloor \cdot \gamma_i) h_i + \{\xi\} \cdot \left\{ \sum_{i=1}^{d} \gamma_i h_i \right\} \bmod \mathbb{Z}.$$

So, it suffices to prove Lemma 9.7 for the case $\xi \in [0, 1)^m$. By [6, Lemma 4.3], either there exists $r \in \mathbb{Z}$, $0 < r \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $\|r\xi_i\|_{\mathbb{R}/\mathbb{Z}} \ll_{d,m} \delta^{-O_{d,m}(1)}/N$ for all $1 \le i \le d$ (where $\xi = (\xi_1, \dots, \xi_m)$), or there exists $k \in \mathbb{Z}^m$ with $0 < |k| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $\|k \cdot \gamma_i\|_{\mathbb{R}/\mathbb{Z}} \ll_{d,m} \delta^{-O_{d,m}(1)}/N$ for all $1 \le i \le d$.

Since $\xi, \gamma_1, \dots, \gamma_d \in \mathbb{Q}^m$, by choosing N to be sufficiently large depending on δ, d, m, and p, we have that either $r\xi_i \in \mathbb{Z}$ for all $1 \le i \le d$ or $k \cdot \gamma_i \in \mathbb{Z}$ for all $1 \le i \le d$. □

With the help of Lemma 9.26, we may conclude from (9.26) the following proposition.

PROPOSITION 9.8. *We have that either there exists $r \in \mathbb{N}_+$ with $r \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $r\xi \in (\mathbb{Z}^m)^d$, or there exists $k \in \mathbb{Z}^{m_{\mathrm{lin}}} \times \{0\}^{m-m_{\mathrm{lin}}}$ with $0 < |k| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $k \cdot \gamma_i \in \mathbb{Z}$ for all $1 \le i \le d$.*

Proof. We remark that we cannot apply the Green–Tao approach (or Lemma 9.26) directly due to the appearance of the term $c_h(hA)$ in (9.26). To overcome this difficulty, we get rid of this term by multiplying both sides of (9.26) by vectors which are orthogonal to hA.

Let $N \gg_{\delta,d,m,p} 1$ be chosen later. For $w \in \mathbb{F}_p^d$, let $V_w := \{n \in \mathbb{F}_p^d \colon (nA) \cdot w = 0\}$ and $Z_w = J_2 \cap \iota^{-1}(V_w)$. For $w' \in \iota^{-1}(w)$ and $h \in Z_w + pK\mathbb{Z}^d$, note that $(hA) \cdot w' \equiv \tau((\iota(h)A) \cdot w) = 0 \bmod p\mathbb{Z}$. Taking the dot product of both sides of (9.26) with w', we have that

$$hB \cdot qw' + u \cdot qw' + \left(\sum_{i=1}^{d} qw_i' \xi_i \right) \cdot \{\gamma(h)\} \in \mathbb{Z} \tag{9.28}$$

for all $w \in \mathbb{F}_p^d$, $w' \in \iota^{-1}(w)$ and $h \in Z_w + pK\mathbb{Z}^d$. To apply Lemma 9.7, we need to identify those w for which the set Z_w is large.

Since

$$\sum_{w \in \mathbb{F}_p^d, wA \neq \mathbf{0}} |Z_w| \ge p^{d-1}|J_2| \gg_{d,m} \delta^{O_{d,m}(1)} p^{d-1}(pK)^d,$$

$|Z_w| \le (pK)^d$, and

$$|\{w \in \mathbb{F}_p^d, wA \neq \mathbf{0}\}| \ge p^d - p^{d-d'} \ge p^d - p^{d-2},$$

by the pigeonhole principle, there exists a subset W of $\{w \in \mathbb{F}_p^d, wA \neq \mathbf{0}\}$ of cardinality $\gg_{d,m} \delta^{O_{d,m}(1)} p^d$ such that $|Z_w| \gg_{d,m} \delta^{O_{d,m}(1)} p^{d-1} K^d$ for all $w \in W$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

To apply Lemma 9.7, we also need to rewrite (9.28) by identifying V_w as $\mathbb{F}_p^{d-1}$ via linear transformations. Since $wA \neq \mathbf{0}$, V_w is a subspace of $\mathbb{F}_p^d$ of co-dimension 1. Let $L_w\colon \mathbb{F}_p^{d-1} \to V_w$ be any bijective linear transformation and let $\tilde{L}_w/p$ be the regular lifting of L_w. Since all of B, ξ_i, γ are rational, there exists $K' \in \mathbb{N}_+$ such that for any $h, h' \in \mathbb{Z}^d$, $\{\gamma(h)\}$ remains unchanged if we replace h by $h + pK'h'$. By Lemma 9.1, the map $\Phi\colon \mathbb{Z}^{d-1} \times \mathbb{Z}^d \to \mathbb{Z}^d$ given by

$$\Phi(z, m) := \tilde{L}_w(z) + pm$$

mod $pK\mathbb{Z}^d$ is a K'^{d-1}-to-1 map between $[pK']^{d-1} \times [K]^d$ and $\iota^{-1}(V_w) \cap [pK]^d$. Since $|Z_w| = |\iota^{-1}(V_w) \cap J_2| \gg_{d,m} \delta^{O_{d,m}(1)} p^{d-1} K^d$, we have that $|\Phi^{-1}(Z_w) \cap ([pK']^{d-1} \times [K]^d)| \gg_{d,m} \delta^{O_{d,m}(1)} (pK')^{d-1} K^d$. So, there exists some $m_w \in [K]^d$ such that the set

$$I_w := \{z \in [pK']^{d-1} \colon \Phi(z, m_w) \in Z_w + pK\mathbb{Z}^d\}$$

is of cardinality $\gg_{d,m} \delta^{O_{d,m}(1)} (pK')^{d-1}$. So, it follows from (9.28) that

$$\tilde{L}_w(z)B \cdot qw' + u + (pm_wB) \cdot qw' + \left(\sum_{i=1}^d qw_i'\xi_i\right) \cdot \{\gamma(\tilde{L}_w(z) + pm_w)\} \in \mathbb{Z} \quad (9.29)$$

for all $z \in I_w$. By the choice of K', we have that (9.29) also holds for all $z \in I_w + pK'\mathbb{Z}^{d-1}$.

We are now ready to apply Lemma 9.7. Note that

$$\liminf_{N\to\infty} \frac{|(I_w + pK'\mathbb{Z}^{d-1}) \cap \{-N, \dots, N\}^{d-1}|}{(2N+1)^{d-1}} = \frac{|I_w|}{(pK')^{d-1}} \gg_{d,m} \delta^{O_{d,m}(1)}.$$

Recall that the last $m - m_{\mathrm{lin}}$ coefficients of ξ_i are zero and so the last $m - m_{\mathrm{lin}}$ coefficients of ξ_i and γ_i does not affect the expression in (9.29). If $N \gg_{\delta,d,m,p} 1$, then we may apply Lemma 9.7 to (9.29) to conclude that either there exists $r_w \in \mathbb{Z}$, $0 < r_w \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $r_w q \sum_{i=1}^d w_i'\xi_i \in \mathbb{Z}^m$ for all $w' \in \iota^{-1}(w)$, or there exists $k_w \in \mathbb{Z}^{m_{\mathrm{lin}}} \times \{0\}^{m-m_{\mathrm{lin}}}$, $0 < |k_w| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $k_w \cdot (\gamma \circ \tilde{L}_w(z)) \in \mathbb{Z}$ for all $z \in \mathbb{Z}^{d-1}$ (or equivalently, $k_w \cdot \gamma(h) \in \mathbb{Z}$ for all $h \in V_w$).

By the pigeonhole principle, there exists a subset W' of W of cardinality $\gg_{d,m} \delta^{O_{d,m}(1)} p^d$ such that either there exists $r \in \mathbb{Z}$ with $0 < r \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $rq \sum_{i=1}^d w_i'\xi_i \in \mathbb{Z}^m$ for all $w' \in \iota^{-1}(W')$, or there exists $k \in \mathbb{Z}^m$ with $0 < |k| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $k \cdot \gamma(h) \in \mathbb{Z}$ for all $h \in \tilde{V}_w := \tilde{L}_w(\mathbb{Z}^{d-1}) = \mathrm{span}_{\mathbb{Z}}\{\tau(L_w(e_1)), \dots, \tau(L_w(e_{d-1}))\}$.

If it is the former case, then since $\iota^{-1}(W') = \tau(W') + p\mathbb{Z}^d$, we must have that $rq\xi_i \in \mathbb{Z}^m/p$ for all $1 \le i \le d$. Since $p \gg_{d,m} \delta^{-O_{d,m}(1)}$, there exist $w_1, \dots, w_d \in W'$ such that the determinant Q of the matrix $(w_{i,j})_{1\le i,j\le d}$ is not divisible by p, where $w_j := (w_{j,1}, \dots, w_{j,d})$. Then, the fact that $rq \sum_{i=1}^d \tau(w_{j,i})\xi_i \in \mathbb{Z}^m$ for $1 \le j \le d$ implies that $rq\xi_i \in \mathbb{Z}^m/Q$. So, $rq\xi_i \in (1/Q)\mathbb{Z}^m \cap (1/p)\mathbb{Z}^m = \mathbb{Z}^m$ and thus, $rq\xi \in (\mathbb{Z}^m)^d$ (recall that $q \ll_d \delta^{-O_d(1)}$).

If it is the latter case, then it is clear that there exist two $w, w' \in W'$ such that $\tilde{V}_w \neq \tilde{V}_{w'}$ and that $\theta \cdot \gamma(h) \in \mathbb{Z}$ for all $h \in \tilde{V}_w \cup \tilde{V}_w$. By linearity, $k \cdot \gamma(h) \in \mathbb{Z}$ for all $h \in \tilde{V}_w + \tilde{V}_{w'}$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Since $\tilde{V}_w$ and $\tilde{V}_{w'}$ are distinct subspaces of $\mathbb{Z}^d$ of co-dimension 1 and $d \geq 2$, we have that $\tilde{V}_w + \tilde{V}_{w'} = \mathbb{Z}^d$. Equivalently, this means that $k \cdot \gamma_i \in \mathbb{Z}$ for all $1 \leq i \leq d$. □

9.6. *Completion of the proof.* As is mentioned in Step 4 in §8, our final task is to use Proposition 9.8 to complete the proof of Theorem 5.1. Except for Proposition 9.9, the outline of the proof is similar to the argument in [**5**, after Claim 7.7 to the end of §7]. Suppose first that there exists $k \in \mathbb{Z}^{m_{\text{lin}}} \times \{0\}^{m-m_{\text{lin}}}$ with $0 < |k| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $k \cdot \gamma_i \in \mathbb{Z}$ for all $1 \leq i \leq d$. Consider the map $\eta\colon G \to \mathbb{R}/\mathbb{Z}$ defined by

$$\eta(x) := k \cdot \psi(x).$$

Then, η is a type-I horizontal character and is of complexity at most $O_{d,m}(\delta^{-O_{d,m}(1)})$. Moreover, for all $n_1, \ldots, n_d \in \mathbb{Z}$, we have that

$$\eta(g(n_1, \ldots, n_d)) = \eta(g(e_1)^{n_1} \cdots g(e_d)^{n_d}) = \sum_{i=1}^{d} n_i k \cdot \gamma_i \equiv 0 \bmod \mathbb{Z}.$$

This completes the proof of Theorem 5.1.

We now consider the case $r\xi \in (\mathbb{Z}^m)^d$ for some $r \in \mathbb{N}_+$ with $r \ll_{d,m} \delta^{-O_{d,m}(1)}$. For $1 \leq j \leq m$, let $\alpha_j\colon G \to \mathbb{R}$ be given by $\alpha_j(x) := \eta_2([x, \exp(X_j)])$. Note that α_j is a type-I horizontal character, annihilates G_2, and $\|\alpha_j\| \ll_{d,m} \delta^{-O_{d,m}(1)}$. If α_j is non-trivial for some $1 \leq j \leq m$, then for all $n_1, \ldots, n_d \in \mathbb{Z}$, we have that

$$r\alpha_j \circ g(n_1, \ldots, n_d) = r\alpha_j(g(e_1)^{n_1} \cdots g(e_d)^{n_d}) = \sum_{i=1}^{d} n_i r\xi_i \cdot \psi(X_j) \equiv 0 \bmod \mathbb{Z},$$

and we again prove Theorem 5.1.

Now, suppose that all of α_j are trivial. In this case, η_2 annihilates $[G, G]$. So, we have that Λ is the zero matrix and that ξ is the zero vector. Therefore, $\sigma(h) = 0$. It follows from (9.24) that

$$t_h - c_h(hA) = q(2h\Omega + u) - c_h(hA) \in \mathbb{Z}^d \tag{9.30}$$

for some $c_h \in \mathbb{Z}/p$ for all $h \in J_2 + pK\mathbb{Z}^d$. This implies that $t_h = q(2h\Omega + u) \in \mathbb{Z}^d/p$ for all $h \in J_2 + pK\mathbb{Z}^d$, which further implies that $q\Omega$ and qu takes $\mathbb{Z}/p$ coefficients. So, for all $h \in J_2 + pK\mathbb{Z}^d$ and $n \in \mathbb{Z}^d$ with $(nA) \cdot w \in p\mathbb{Z}^d$, (9.30) implies that $q(2h\Omega + u) \cdot n \in \mathbb{Z}$. Since J_2 contains $\gg_{d,m} \delta^{O_{d,m}(1)} p^d$ residue classes mod $p\mathbb{Z}$, it follows from Corollary 7.5 that $qu \in \mathbb{Z}$ and $2q\Omega = rA + B'$ for some $r \in \mathbb{Z}/p$ and matrix B' with integer entries. It then follows from (9.20) and (9.21) that (replacing q by $2q$ if necessary)

$$\eta_2(g_{\text{nlin}}(n)) = Q(n) = \frac{1}{q} Q_2(n) + T(n) \tag{9.31}$$

for some $Q_2 \in \text{poly}(V_p(\tilde{M}) \to \mathbb{R}|\mathbb{Z})$ and some polynomial $T\colon \mathbb{Z}^d \to \mathbb{R}$ of degree at most 1.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

If η_2 is trivial, then $Q \equiv 0$ and Ω can be taken to be the zero matrix. It follows from (9.30) that

$$qu - c_h(hA) \in \mathbb{Z}^d$$

for some $c_h \in \mathbb{Z}/p$ for all $h \in J_2$. Note that if $qu \notin \mathbb{Z}^d$, then the number of such h is at most $K^d \cdot p^{d+1-\mathrm{rank}_p(A)} \leq K^d p^{d-1} < |J_2|$, which is a contradiction. So, we must have that $qu \in \mathbb{Z}^d$. So, it follows from (9.20) that if $M(n) \in \mathbb{Z}$, then

$$q\eta_1(g(n)) = qP(n) = P_1(n) + (qu) \cdot n + qu_0 \in \mathbb{Z} + qu_0.$$

Since $\eta_0(g', g) = \eta_1(g) + \eta_2(g'g^{-1})$ is non-trivial and η_2 is trivial, we have that η_1 is non-trivial. Since $q \ll_d \delta^{-O_d(1)}$, this completes the proof of Theorem 5.1 by setting $\eta = q\eta_1$.

So, from now on, we assume that η_2 is not trivial. If $m_* = 0$, then η_2 vanishes on G_2 and so it is trivial, which is a contradiction. We now assume that $m_* \geq 1$. Since η_2 annihilates $[G, G]$, we may extend η_2 to a genuine type-I horizontal character $\tilde{\eta}_2$ on G in an arbitrary way such that $\tilde{\eta}_2|_{G_2} = \eta_2$, and that the complexity of $\tilde{\eta}_2$ is comparable to η_2. With a slight abuse of notation, we denote $\tilde{\eta}_2$ by η_2 as well and treat η_2 as a genuine type-I horizontal character on G. Suppose that $\eta_2 \colon G \to \mathbb{R}$ is given by $\eta_2(x) = k \cdot \psi(x)$ for some $k = (k_1, \ldots, k_m) \in \mathbb{Z}^m$ with $|k| \ll_{d,m} \delta^{-O_{d,m}(1)}$. Since η_2 annihilates $[G, G]$, we have that $k_i = 0$ for all $i > m_{ab}$.

Denote $G'_0 = G'_1 = G$ and $G'_i = G_i \cap \ker(\eta_2)$ for $i \geq 2$. By [**5**, Lemma 7.8], $G'_{\mathbb{N}} = (G'_i)_{i \in \mathbb{N}}$ is a filtration of G with degree at most s and nonlinearity degree $m'_* \leq m_* - 1$. Each G'_i is closed, connected, and $O_{d,m}(\delta^{-O_{d,m}(1)})$-rational with respect to the Mal'cev basis $\mathcal{X}$ on G/Γ adapted to $G_{\mathbb{N}}$. We need a special factorization proposition.

PROPOSITION 9.9. *Fix any $n_* \in V_p(\tilde{M})$. There exists $r \in \mathbb{N}_+$ with $r \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that we may write $g = g'\gamma$ for some rational $g' \in \mathrm{poly}(\mathbb{Z}^d \to G'_{\mathbb{N}})$ and some $\gamma \in \mathrm{poly}_{\approx r, n_*}(V_p(\tilde{M}) \to G_{\mathbb{N}}|\Gamma)$.*

We first complete the proof of Theorem 5.1 assuming Proposition 9.9. Let the notation be the same as Proposition 9.9. By Proposition 3.17, there exists $r' \in \mathbb{N}_+$ with $r' \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $\gamma \in \mathrm{poly}_{r'}(V_p(\tilde{M}) \to G_{\mathbb{N}}|\Gamma)$. For all $m \in [r']^d$, there exists $\tilde{m} \in \mathbb{Z}^d$ with $\tilde{m} - m \in r'\mathbb{Z}^d$ and $\tilde{m} - n_* \in p\mathbb{Z}^d$. Since $n_* \in V_p(\tilde{M})$, we have $\tilde{m} \in V_p(\tilde{M})$. Since $\gamma \in \mathrm{poly}_{r'}(V_p(\tilde{M}) \to G_{\mathbb{N}})$, we have that

$$\gamma(\tilde{m})\Gamma = \gamma(r'n + m)\Gamma \quad \text{for all } n \in \mathbb{Z}^d \quad \text{with } r'n + m \in V_p(\tilde{M}). \tag{9.32}$$

Denote

$$g'_m(n) := \{\gamma(\tilde{m})\}^{-1} g'(r'n + m)\{\gamma(\tilde{m})\}.$$

Since g' and γ are rational, by Corollary 3.19, there exists $K' \in \mathbb{N}_+$ which is divisible by pK such that g' and g'_m, $m \in [r']^d$ are all K'-periodic. For $m \in \mathbb{Z}^d$, let $M_m \colon \mathbb{Z}^d \to \mathbb{Z}/p$ be the quadratic map given by $\tilde{M}_m(n) := \tilde{M}(r'n + m)$. Then, since $r \ll_{d,m} \delta^{-O_{d,m}(1)}$, $\tilde{M}_m$ and $\tilde{M}$ have the same p-rank. Note that for all $n \in V_p(\tilde{M}_m)$, we have that $r'n + m \in$

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

$V_p(\tilde{M})$. Then, it follows from (9.32) and (9.4) that

$$\begin{aligned}\delta^{O_{d,m}(1)} &\ll_{d,m} |\mathbb{E}_{n\in V_p(\tilde{M})\cap[pK]^d} F(g(n)\Gamma)| = |\mathbb{E}_{n\in V_p(\tilde{M})\cap[r'K']^d} F(g(n)\Gamma)|\\ &= |\mathbb{E}_{m\in[r']^d}\mathbb{E}_{n\in V_p(\tilde{M}_m)\cap[K']^d} F(g(r'n+m)\Gamma)| + O(p^{-1/2})\\ &= |\mathbb{E}_{m\in[r']^d}\mathbb{E}_{n\in V_p(\tilde{M}_m)\cap[K']^d} F(g'(r'n+m)\gamma(r'n+m)\Gamma)| + O(p^{-1/2})\\ &= |\mathbb{E}_{m\in[r']^d}\mathbb{E}_{n\in V_p(\tilde{M}_m)\cap[K']^d} F(g'(r'n+m)\gamma(\tilde{m})\Gamma)| + O(p^{-1/2})\\ &= |\mathbb{E}_{m\in[r']^d}\mathbb{E}_{n\in V_p(\tilde{M}_m)\cap[K']^d} F(\{\gamma(\tilde{m})\}g'_m(n)\Gamma)| + O(p^{-1/2}).\end{aligned}$$

So by the pigeonhole principle, there exists some $m \in [r']^d$ such that

$$|\mathbb{E}_{n\in V_p(\tilde{M}_m)\cap[K']^d} F'_m(g'_m(n)\Gamma)| \gg_{d,m} \delta^{O_{d,m}(1)},$$

where $F'_m := F(\{\gamma(\tilde{m})\}\cdot)$. By the construction of K', g'_m is a rational and K'-periodic polynomial in $\mathrm{poly}(\mathbb{Z}^d \to G'_{\mathbb{N}})$ and $\|F'\|_{\mathrm{Lip}} = O(1)$. We may now invoke the induction hypothesis to deduce that there exists a non-trivial type-I horizontal character η of complexity at most $O_{d,m}(\delta^{-O_{d,m}(1)})$ such that $\eta \circ g'_m \bmod \mathbb{Z}$ is a constant on $V_p(\tilde{M}_m)$. Since

$$\begin{aligned}\eta \circ g'_m(n) &= \eta \circ g'(r'n+m)\\ &= \eta \circ g(r'n+m) - \eta \circ \gamma(r'n+m) \equiv \eta \circ g(r'n+m) - \eta \circ \gamma(m) \bmod \mathbb{Z}\end{aligned}$$

for all $n \in V_p(\tilde{M}_m)$, we have that $\eta \circ g \bmod \mathbb{Z}$ is a constant on $V_p(\tilde{M})$. This finishes the proof of Theorem 5.1.

So, it remains to prove Proposition 9.9. Proposition 9.9 can be viewed as an analog of [5, Lemma 7.9], but its proof deviates significantly from [5, Lemma 7.9].

Proof of Proposition 9.9. Let $0 \le s' \le s$ be the largest integer such that at least one of the entries $k_{m-m_{s'+1}}, \dots, k_{m-m_{s'+1}}$ is non-zero (where we denote $m_0 := 0$). Since η_2 is non-trivial, such an s' always exists and $s' \ge 2$. Denote $m' := m - m_{s'+1}$. Then, we may write $k = (k', \mathbf{0})$ for some $k' \in \mathbb{Z}^{m'}$.

Assume that

$$\psi(g(n)) = \sum_{i\in\mathbb{N}^d, 0\le|i|\le s} w_i \binom{n}{i}$$

for some $w_i \in \{0\}^{m-m_{|i|}} \times \mathbb{R}^{m_{|i|}}$. Since $Q(\mathbf{0}) = Q(e_i) = 0$, we may write

$$\psi(g_{\mathrm{nlin}}(n)) = \sum_{m\in\mathbb{N}^d, 2\le|i|\le s} t_i \binom{n}{i}$$

for some $t_i \in \{0\}^{m-m_{|i|}} \times \mathbb{R}^{m_{|i|}}$. Write $w_i = (w'_i, w''_i)$, $t_i = (t'_i, t''_i)$ for some $t'_i, w'_i \in \mathbb{R}^{m'}$ and $t''_i, w''_i \in \mathbb{R}^{m-m'}$. Since the first $m - m_{|i|}$ entries of t_i are zero and the last $m_{s'+1}$ entries of k are zero, we have that $k \cdot t_i = 0$ if $|i| \ge s'+1$. By (9.31),

$$\frac{1}{q}Q_2(n) + T(n) = \eta_2(g_{\mathrm{nlin}}(n)) = \sum_{i\in\mathbb{N}^d, 2\le|i|\le s} (k\cdot t_i)\binom{n}{i} = \sum_{i\in\mathbb{N}^d, 2\le|i|\le s'} (k'\cdot t'_i)\binom{n}{i}.$$

In particular, we have that $\deg(Q_2) \le s'$.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

Since k is of complexity $\ll_{d,m} \delta^{-O_{d,m}(1)}$, there exists $q' \in \mathbb{N}_+$, $q' \ll_{d,m} \delta^{-O_{d,m}(1)}$, and $u = (u_1, \dots, u_m) \in \mathbb{Z}^m$ whose only non-zero entries are $u_{m-m_{s'}+1}, \dots, u_{m-m_{s'+1}}$ such that $k \cdot u = q'$. Denote

$$\gamma(n) := \prod_{i=m-m_{s'+1}}^{m-m_{s'}+1} \exp\left(\frac{1}{qq'} u_i (Q_2(n) - Q_2(n_*)) X_i\right).$$

Since $\deg(Q_2) \le s'$, it follows from [8, Corollary B.4] that $\gamma \in \mathrm{poly}(\mathbb{Z}^d \to (G_2)_{\mathbb{N}})$.

We now show that there exists $r \in \mathbb{N}_+$ with $r \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that γ belongs to $\mathrm{poly}_{\approx r,n_*}(V_p(\tilde{M}) \to G_{\mathbb{N}}|\Gamma)$ and takes values in G_2. In fact, it suffices to show that $Q_2(n_* + rn) - Q_2(n_*) \in qq'\mathbb{Z}$ for all $n \in \mathbb{Z}^d$ with $n_* + rn \in V_p(\tilde{M})$. By interpolation, all the coefficients of the polynomial Q_2 belongs to $\mathbb{Z}/p^{s'}q''$ for some $s' = O_{d,s}(1)$ and $q'' \in \mathbb{N}_+$ with $q'' \le O_{d,s}(1)$. So, for all $n \in \mathbb{Z}^d$, we have that $Q_2(n_* + rn) - Q_2(n_*) \in qq'\mathbb{Z}/p^{s'}$, where $r := qq'q''$. However, if $n_* + rn, n_* \in V_p(\tilde{M})$, then $Q_2(n_* + rn) - Q_2(n_*) \in \mathbb{Z}$. Since qq' is not divisible by p, this implies that $Q_2(n_* + rn) - Q_2(n_*) \in qq'\mathbb{Z}$. So, $\gamma \in \mathrm{poly}_{\approx r,n_*}(V_p(\tilde{M}) \to G_{\mathbb{N}}|\Gamma)$ and takes values in G_2.

Writing $h := g\gamma^{-1}$, we have that $h \in \mathrm{poly}(\mathbb{Z}^d \to (G_2)_{\mathbb{N}}|(\Gamma \cap G_2))$. Note that for all $n, h_1, h_2 \in \mathbb{Z}^d$,

$$\begin{aligned}\Delta_{h_2}\Delta_{h_1}(\eta_2 \circ g(n)) &= \eta_2 \circ (\Delta_{h_2}\Delta_{h_1} g)(n) = \eta_2 \circ (\Delta_{h_2}\Delta_{h_1} g_{\mathrm{nlin}})(n) \\ &= \Delta_{h_2}\Delta_{h_1}(\eta_2 \circ g_{\mathrm{nlin}})(n) = \Delta_{h_2}\Delta_{h_1}\left(\frac{1}{q}Q_2(n) + T(n)\right) = \Delta_{h_2}\Delta_{h_1}\left(\frac{1}{q}Q_2(n)\right).\end{aligned} \tag{9.33}$$

So, (9.33) implies that

$$\Delta_{h_2}\Delta_{h_1}(\eta_2 \circ h(n)) = \Delta_{h_2}\Delta_{h_1}(\eta_2 \circ g(n)) - \Delta_{h_2}\Delta_{h_1}(\eta_2 \circ \gamma(n)) = 0. \tag{9.34}$$

Set $g' := g\gamma^{-1} = g_{\mathrm{nlin}} g_{\mathrm{lin}} \gamma^{-1}$. It is clear that $g' \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$. Moreover, by (9.34),

$$\Delta_{h_2}\Delta_{h_1}(\eta_2 \circ g')(n) = \eta_2(\Delta_{h_2}\Delta_{h_1} g'(n)) = 0$$

for all $n, h_1, h_2 \in \mathbb{Z}^d$. So, it is not hard to see that g' belongs to $\mathrm{poly}(\mathbb{Z}^d \to G'_{\mathbb{N}})$. Finally, since $g' = g\gamma^{-1}$, it follows from the Baker–Campbell–Hausdorff formula that g' is rational. □

We end this paper with an immediate consequence of Theorem 5.1, which will be used in [16].

COROLLARY 9.10. *Let $0 < \delta < 1/2$, $d, k, m \in \mathbb{N}_+$, $s, r \in \mathbb{N}$ with $d \ge r$, and $p \gg_{d,m} \delta^{-O_{d,m}(1)}$ be a prime. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form and $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of co-dimension r. Suppose that $\mathrm{rank}(M|_{V+c}) \ge s + 13$. Let G/Γ be an s-step $\mathbb{N}$-filtered nilmanifold of dimension m, equipped with a $1/\delta$-rational Mal'cev basis $\mathcal{X}$, and that $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ be a rational polynomial sequence. Let $\gamma \in G$ be an element of complexity at most 1, and let $g' \in \mathrm{poly}(\mathbb{Z}^d \to G'_{\mathbb{N}})$ be the map given by $g'(n) := \gamma^{-1} g(n) \gamma$ for all $n \in \mathbb{Z}^d$, where $G'_i := \gamma^{-1} G_i \gamma$ for all $i \in \mathbb{N}$. Denote $\Gamma' := \gamma^{-1}\Gamma\gamma$. If $(g(n)\Gamma)_{n \in \iota^{-1}(V(M) \cap (V+c))}$ is not δ-equidistributed on*

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press

G/Γ, then $(g'(n)\Gamma')_{n\in\iota^{-1}(V(M)\cap(V+c))}$ is not $C^{-1}\delta^C$-equidistributed on G'/Γ' for some $C = C(d,m) \geq 1$.

Proof. By Theorem 5.1, there exists a non-trivial type-I horizontal character η with $0 < \|\eta\| \ll_{d,m} \delta^{-O_{d,m}(1)}$ such that $\eta \circ g' = \eta \circ g \bmod \mathbb{Z}$ is a constant on $\iota^{-1}(V(M) \cap (V+c))$. By [5, Lemma A.13], G'/Γ' is an s-step $\mathbb{N}$-filtered nilmanifold of dimension m, equipped with a $1/(\delta^{O(1)})$-rational Mal'cev basis. Since $|\mathbb{E}_{n\in\iota^{-1}(V(M)\cap(V+c))} \exp(\eta(g'(n)))| = 1$, this implies that $(g'(n)\Gamma')_{n\in\iota^{-1}(V(M)\cap(V+c))}$ is not $C^{-1}\delta^C$-equidistributed on G'/Γ' for some $C = C(d,m) \geq 1$. □

Acknowledgements. We thank James Leng for helpful discussions. We also thank the anonymous referee, whose comments greatly increased the presentation of the paper. This paper was partially supported by the NSF Grant DMS-2247331.

References

[1] P. Candela and O. Sisask. Convergence results for systems of linear forms on cyclic groups, and periodic nilsequences. *SIAM J. Discrete Math.* **28**(2) (2014), 786–810.
[2] S. R. Ghorpade. A note on Nullstellensatz over finite fields. *Contemp. Math.* **738** (2019), 23–32.
[3] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)* **167** (2008), 481–547.
[4] B. Green and T. Tao. Linear equations in the primes. *Ann. of Math. (2)* **171** (2010), 1753–1850.
[5] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math. (2)* **175**(2) (2012), 465–540.
[6] B. Green and T. Tao. On the quantitative distribution of polynomial nilsequences - erratum. *Ann. of Math. (2)* **179**(3) (2014), 1175–1183.
[7] B. Green, T. Tao and T. Ziegler. An inverse theorem for the Gowers U^4-norm. *Glasg. Math. J.* **53** (2011), 1–50.
[8] B. Green, T. Tao and T. Ziegler. An inverse theorem for the Gowers U^{s+1}-norm. *Ann. of Math. (2)* **176**(2) (2012), 1231–1372.
[9] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)* **161**(1) (2005), 397–488.
[10] A. Iosevich and M. Rudnev. Erdös distance problem in vector spaces over finite fields. *Trans. Amer. Math. Soc.* **359**(12) (2007), 6127–6142.
[11] B. Kra, N. Shah and W. Sun. Equidistribution of dilated curves on nilmanifolds. *J. Lond. Math. Soc. (2)* **98**(3) (2018), 708–732.
[12] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifolds. *Ergod. Th. & Dynam. Sys.* **25**(1) (2005), 201–213.
[13] W. Schmidt. *Equations over Finite Fields, an Elementary Approach (Lecture Notes in Mathematics, 536)*. Springer-Verlag, Berlin–Heidelberg–New York, 1976.
[14] W. Sun. Weak ergodic averages over dilated measures. *Ergod. Th. & Dynam. Sys.* **41**(2) (2021), 606–621.
[15] W. Sun. Spherical higher order Fourier analysis over finite fields II: additive combinatorics for shifted ideals. *Preprint*, 2024, arXiv:2312.06650.
[16] W. Sun. Spherical higher order Fourier analysis over finite fields III: a spherical Gowers inverse theorem. *Preprint*, 2024, arXiv:2312.06636.
[17] W. Sun. Spherical higher order Fourier analysis over finite fields IV: an application to the geometric Ramsey conjecture. *Preprint*, 2024, arXiv:2312.06649.
[18] W. Sun. Leibman dichotomy for nilsequences along spheres and beyond. *Preprint*. Available on https://sites.google.com/view/wenbosunmath/.

https://doi.org/10.1017/etds.2025.10264 Published online by Cambridge University Press