OPEN ACCESS

University of BRISTOL

Peer reviewed version

Link to published version (if available):
10.1093/imrn/rnw096

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

A TRANSFERENCE APPROACH TO A ROTH-TYPE THEOREM IN THE SQUARES

T. D. BROWNING AND S. M. PRENDIVILLE

ABSTRACT. We show that any subset of the squares of positive relative upper density contains non-trivial solutions to a translation-invariant linear equation in five or more variables, with explicit quantitative bounds. As a consequence, we establish the partition regularity of any diagonal quadric in five or more variables whose coefficients sum to zero. Unlike previous approaches, which are limited to equations in seven or more variables, we employ transference technology of Green to import bounds from the linear setting.

CONTENTS

1. INTRODUCTION

Initiated by Roth [Rot53, Rot54], there has been much work (see [Blo12, SchSi]) on bounding the density of subsets of integers lacking solutions to a single linear equation

$$c_1 y_1 + \cdots + c_s y_s = 0, \tag{1.1}$$

with integer coefficients. It is customary to impose the condition that $c_1 + \cdots + c_s = 0$ in order to avoid the existence of a congruence class lacking solutions. Our knowledge is more formative for the analogous non-linear equation

$$c_1 x_1^2 + \cdots + c_s x_s^2 = 0, \tag{1.2}$$

where the aim is to show that sufficiently dense subsets of integers have generic solutions. Given a union K of k proper subspaces of the hyperplane (1.1), our notion of non-generic solutions is captured by K-triviality, where a solution $(x_1, \ldots, x_s)$ is said to be K-*trivial* if $(x_1^2, \ldots, x_s^2) \in K$. For instance, one could regard solutions with all x_i equal as K-trivial, by taking K to be the diagonal subspace

$$K := \{(y, \ldots, y) : y \in \mathbb{Q}\}, \tag{1.3}$$

for which $k = 1$. More generally, taking

$$K := \bigcup_{i \neq j} \{\mathbf{y} \in \mathbb{Q}^s : y_i = y_j \text{ and } \mathbf{c} \cdot \mathbf{y} = 0\},$$

we have $k = \binom{s}{2}$ and the K-nontrivial solutions correspond to those in which all x_i are distinct.

All work on equation (1.2) (see [Smi09, Ke14, Keil15, Hen15]) is based on an adaptation of the density increment approach to Roth's theorem, and is at present limited to equations in at least $s \geqslant 7$ variables and with at least two positive and two negative coefficients. Adapting the transference technology of Green [Gre05], we offer an alternative approach to this problem, an approach which is applicable to a wider class of equations, in particular those in $s \geqslant 5$ variables and with no sign restriction on the coefficients.

Theorem 1.1. *Let $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$ with $c_1 + \cdots + c_s = 0$ and $s \geqslant 5$. Let K denote a union of at most k proper subspaces of the hyperplane (1.1) each containing the diagonal subspace (1.3). If $A \subset [X]$ is such that the only solutions to*

$$c_1 x_1^2 + \cdots + c_s x_s^2 = 0 \qquad (x_i \in A), \tag{1.4}$$

are K-trivial, then for any $\varepsilon > 0$ we have

$$|A| \ll_{\mathbf{c}, k, \varepsilon} \frac{X}{(\log \log \log X)^{\frac{s}{2} - 1 - \varepsilon}}.$$

Let $c_1, \ldots, c_s$ be as in the statement of Theorem 1.1. Then this result implies that any subset $A \subset \mathbb{N}$, with positive upper density

$$\limsup_{X \to \infty} \frac{|A \cap [X]|}{X} > 0, \tag{1.5}$$

must contain generic solutions to (1.4).

A Diophantine equation $f(x_1, \ldots, x_s) = 0$ is called *partition regular* if, given any partition of $\mathbb{N}$ into finitely many colour classes, at least one class contains a solution to the equation in which all x_i are distinct. An old problem of Erdős and Graham [EG80] asks if the Pythagorean equation $x^2 + y^2 = z^2$ is partition regular. Graham [Gra08] has offered \$250 for a resolution of this conjecture. Similarly, it is unknown if the equation $x^2 + y^2 = 2z^2$ is partition regular (see [FrHo14]). Given an extra two variables, we can answer questions of the latter type.

Corollary 1.2. *Let* $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$ *with* $c_1 + \cdots + c_s = 0$ *and* $s \geqslant 5$. *Then the equation* $c_1 x_1^2 + \cdots + c_s x_s^2 = 0$ *is partition regular.*

This follows from Theorem 1.1, since at least one colour class must have positive upper density (1.5).

Define the *Rado number* (see [GRS90, p.103]) of the equation (1.4) to be the smallest positive integer $R_{\mathbf{c}}(r)$ such that any r-colouring of the interval $\{1, 2, \ldots, R_{\mathbf{c}}(r)\}$ results in at least one monochromatic tuple $(x_1, \ldots, x_s)$ satisfying (1.4) with all x_i distinct. Since any r-colouring of $[X]$ results in a colour class of size at least X/r, one can use the estimate given in Theorem 1.1 to deduce the following result.

Corollary 1.3. *Let* $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$ *with* $c_1 + \cdots + c_s = 0$ *and* $s \geqslant 5$. *Then there exists a constant* $C_{\mathbf{c}} > 0$, *depending only on* $c_1, \ldots, c_s$, *such that*

$$R_{\mathbf{c}}(r) \leqslant \exp \exp \exp(C_{\mathbf{c}} r)$$

The study of density bounds for sets lacking solutions to (1.4) was initiated by Smith [Smi09] who proceeded by considering instead the system of equations

$$
\begin{aligned}
c_1 x_1^2 + \cdots + c_s x_s^2 &= 0, \\
c_1 x_1 + \cdots + c_s x_s &= 0.
\end{aligned}
\tag{1.6}
$$

The advantage of this system is that it is translation-invariant, in that the set of solutions remains unchanged under a shift from the diagonal subspace (1.3). This allows one to follow Roth's original approach to the equation (1.1), using an L^p-estimate (mean value) to deduce that if a set A lacks solutions to (1.6), then there is an exponential sum associated to A which is large. It follows that A has a density increment on a long progression of the form $a + q \cdot [N']$, and so (by translation-invariance) there exists a denser set $A' \subset [N']$ lacking solutions to (1.6). By iterating this argument, a bound can eventually be extracted on the density of the original set A.

Smith [Smi09] accomplished this programme for equations in 9 or more variables, obtaining a bound of the form $|A| \ll X(\log \log X)^{-c}$ for some small $c > 0$. This small exponent resulted from a use of quadratic Fourier analysis, a tool that Keil [Ke14] avoided, obtaining a bound of the form $X(\log \log X)^{-1/15}$. Moreover, Keil incorporated a restriction L^p-estimate in order to handle equations in seven or more variables. Subsequently, Henriot [Hen15] has improved the cardinality bound to $X(\log X)^{-c}$ for $s \geqslant 7$ and an exponent $c > 0$ that depends on the choice of coefficients $c_1, \ldots, c_s$.

In all these approaches, one requires that there are at least two positive and two negative coefficients in order to guarantee that the system (1.6) has a non-diagonal real solution. Furthermore, $s \geqslant 7$ is a natural limit of the analytic method, as for $s \leqslant 6$ the system may possess more solutions than that predicted by the usual probabilistic heuristic. However, if one is only interested in sets of integers lacking solutions to the top equation (1.4), then neither of these conditions should be necessary.

Analysing the single equation (1.4) removes translation-invariance, making a density increment approach more problematic. Instead we implement the transference principle devised in Green [Gre05] to import results from the dense linear setting to the sparse linear setting, viewing a dense subset of the squares as a sparse subset of the integers. Although Roth's theorem only holds for dense subsets of integers, we show that an appropriately normalised indicator function of the squares has sufficient pseudorandomness properties that any dense subset of squares can be modelled by a dense subset of integers. Applying Roth's theorem in the dense setting, we are then able to transfer back to the sparse setting to bound the size of our set of squares. We describe this method further in §2.

It is worth remarking that Gowers' quantitative version of Szemerédi's theorem [Gow01] yields bounds for sets lacking solutions to very many homogeneous Diophantine equations, an observation the second author learnt from Trevor Wooley. (A variant of this idea was put to use by Brüdern et al [BDLW10].) For the sake of completeness, we include a proof of the folklore argument.

Proposition 1.4. *Let $F \in \mathbb{Z}[x_1, \ldots, x_s]$ be a homogeneous polynomial with integer coefficients and let K denote a finite union of proper subspaces of $\mathbb{Q}^s$, one of which is equal to the diagonal subspace (1.3). Suppose that the solution set $F(x_1, \ldots, x_s) = 0$ contains a two-dimensional rational subspace which contains the diagonal (1.3) and which is not contained in K. Then any set $A \subset [X]$ which satisfies the implication*

$$F(\mathbf{x}) = 0 \quad (\mathbf{x} \in A^s) \qquad \Longrightarrow \qquad \mathbf{x} \in K,$$

has cardinality

$$|A| \ll \frac{X}{(\log \log X)^c},$$

for some positive constant $c = c(F) > 0$.

Proof. Let V denote a two-dimensional rational subspace of the solution set $F(\mathbf{x}) = 0$, which contains the diagonal (1.3) and is not contained in K. Since $(1, \ldots, 1) \in V \cap K$ and $\dim V = 2$, there exists $\mathbf{x} \in V \setminus K$ which is linearly independent of $(1, \ldots, 1)$. Clearing denominators and shifting along the diagonal subspace by a sufficiently large integer, we may assume that $\mathbf{x} \in ([0, k) \cap \mathbb{Z})^s$ for some $k \in \mathbb{N}$.

Notice that for any $a, q \in \mathbb{Z}$ the tuple $\mathbf{y} = (a + qx_1, \ldots, a + qx_s)$ lies in V, so that $F(\mathbf{y}) = 0$. Suppose now that A contains an arithmetic progression $a, a + q, \ldots, a + (k-1)q$ of length k. Then it follows that A^s contains $\mathbf{y}$, which is also a solution to $F = 0$. Moreover, this solution is not contained in K, since otherwise

$$\mathbf{x} = \frac{\mathbf{y} - (a, \ldots, a)}{q} \in K.$$

Thus A cannot contain an arithmetic progression of length k, so that $|A| \ll X (\log \log X)^{-c_k}$, for some $c_k > 0$, by Gowers's bounds in Szemerédi's theorem [Gow01]. $\qquad\square$

This argument covers those equations of the form (1.4) tackled by Smith, Keil and Henriot, namely those in which $c_1 + \cdots + c_s = 0$ for $s \geqslant 7$, and where there are at least two positive and two negative coefficients. Indeed, an application of the Hardy–Littlewood method (as carried out by Keil [Ke14]) gives an integer vector $\mathbf{x} \in \mathbb{Z}^s \setminus K$ satisfying the system (1.6) under these conditions. Expanding the square, one can check that for any $\lambda, \mu \in \mathbb{Q}$ the vector $(\lambda + \mu x_1, \ldots, \lambda + \mu x_s)$ also satisfies (1.6). In particular, the solution set of (1.4) contains the two dimensional subspace generated by the diagonal and $\mathbf{x}$, a subspace which is not contained in K.

Unlike the bounds obtained via Proposition 1.4, the advantage of the bounds of Smith and Keil is that their logarithmic exponent is *uniform* in the choice of coefficients, whilst Henriot succeeds in replacing the double logarithm derived from Gowers's bounds by a single logarithm.

Note that one cannot use Proposition 1.4 to treat all equations covered by Theorem 1.1. For instance, the convex equation

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 4x_5^2$$

is of the form (1.4) but is not satisfied by a two dimensional subspace containing the diagonal. Should the solution set of this equation contain such a set, then this subspace can be written in the form

$$\{(\lambda, \ldots, \lambda) + \mu \mathbf{x} : \lambda, \mu \in \mathbb{Q}\}. \tag{1.7}$$

Expanding the squares in the equation $\sum_i (\lambda + \mu x_i)^2 = 4(\lambda + \mu x_5)^2$ reveals that $\mathbf{x}$ also satisfies the system (1.6). Substituting the linear equation into the quadratic equation shows that

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 4(x_1 + x_2 + x_3 + x_4)^2,$$

which forces $\mathbf{x}$ to be an element of the diagonal, contradicting the fact that (1.7) is two dimensional.

Finally we remark that in all likelihood our methods can be adapted to handle diagonal equations of degree k in s variables, provided that one has a restriction estimate at exponent $p < s$ for the appropriate exponential sum. This gives an alternative approach to recent results of Henriot [Hen15], albeit with much weaker density bounds. Whereas Henriot's methods rely on a restriction estimate associated to the Vinogradov system

$$x_1^k + \cdots + x_s^k = y_1^k + \cdots + y_s^k$$
$$\vdots$$
$$x_1 + \cdots + x_s = y_1 + \cdots + y_s,$$

the transfer approach of this paper would merely require a mean value estimate for the top equation in the system. For small values of k, it appears that the transference approach may allow one to improve the number of variables required in Henriot's work. For instance, an eighth moment estimate of Vaughan [Vau86] should yield the partition regularity of a diagonal cubic equation in 9 or more variables, whereas the work of Henriot [Hen15] requires at least 13 variables when specified to this situation.

2. The structure of our argument

Definition 2.1 (Majorant). A *majorant on* $[N]$ is a non-negative function $\nu : \mathbb{Z} \to [0, \infty)$ with support contained in the interval $[N]$.

The next definition, and much of this section, follows the exposition of Tao [Tao12, §1.7].

Definition 2.2 (**c**-pseudorandom). Let $\mathbf{c} = (c_1, \ldots, c_s) \in \mathbb{Z}^s$. We say that a majorant ν is **c**-*pseudorandom* if for any $\delta > 0$ there exists $c(\delta) > 0$ such that for any $0 \leqslant f \leqslant \nu$ we have the implication

$$\sum_x f(x) \geqslant \delta \sum_x \nu(x) \quad \Longrightarrow \quad \sum_{\mathbf{c} \cdot \mathbf{x} = 0} \prod_{i=1}^s f(x_i) \geqslant c(\delta) \sum_{\mathbf{c} \cdot \mathbf{x} = 0} \prod_{i=1}^s \nu(x_i).$$

The indicator function of an interval provides an important example of a **c**-pseudorandom majorant, as can be seen from the following result of Bloom [Blo12] (building on work of Sanders [San11]).

Bloom's theorem ([Blo12]). *Let* $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$ *with* $s \geqslant 3$ *and* $c_1 + \cdots + c_s = 0$. *Then the indicator function* $1_{[N]}$ *is* **c**-*pseudorandom with quantitative dependence*

$$c(\delta) \gg_{\mathbf{c}} \exp\left(-C/\delta^{\frac{1}{s-2-\varepsilon}}\right)$$

for some absolute constant $C = C(s, \varepsilon)$ *and any* $\varepsilon > 0$.

Our aim is to show that there exists a function related to the indicator function of the squares which is also **c**-pseudorandom, at least when $s \geqslant 5$. There are three conditions which, taken together, are sufficient to guarantee that a majorant is **c**-pseudorandom. We discuss these in the remainder of this section.

Define the L^p-norm of a function $f : \mathbb{Z} \to \mathbb{C}$ with respect to counting measure, so for example

$$\|f\|_1 := \sum_n |f(n)|.$$

Provided that this L^1-norm is finite, we define the Fourier transform of f at $\alpha \in \mathbb{T} := \mathbb{R}/\mathbb{Z}$ by

$$\hat{f}(\alpha) := \sum_n f(n)e(\alpha n). \tag{2.1}$$

For functions on $\mathbb{T}$, all L^p-norms are taken with respect to the Haar probability measure, so that $\|\hat{f}\|_2 = \|f\|_2$.

Definition 2.3 (Fourier decay)**.** We say that a majorant ν has *Fourier decay of level θ* if

$$\left\|\hat{\nu} - \hat{1}_{[N]}\right\|_\infty \leqslant \theta N.$$

The relevance of this definition is demonstrated in work of Green [Gre05]. The idea is that one can approximate an unbounded function by a bounded function, provided the unbounded function has a majorant with sufficient Fourier decay. For a proof of the following form of the bounded approximation lemma we refer the reader to the second author's survey [Pre15].

Lemma 2.4 (Bounded approximation)**.** *Suppose that the majorant ν has Fourier decay of level θ. Then for any $0 \leqslant f \leqslant \nu$ there exists a bounded function $0 \leqslant g \leqslant 1_{[N]}$ such that*

$$\left\|\hat{f} - \hat{g}\right\|_\infty \ll \frac{N}{\log(1/\theta)^{3/2}}.$$

The advantage of a bounded approximant is that it is amenable to an application of Bloom's theorem, and should therefore count many solutions to (1.1). We wish to approximate the characteristic function of a set of squares weighted by an unbounded majorant. Given a sufficiently good approximation, we might therefore hope that the number of solutions to (1.1) counted by our unbounded function is similar to the count for the bounded function. This is indeed the case, provided that our majorant possesses an additional property.

Definition 2.5 (Restriction at p)**.** We say that a majorant ν on $[N]$ satisfies a *restriction estimate at exponent p* if

$$\sup_{|\phi| \leqslant \nu} \int_{\mathbb{T}} \left|\hat{\phi}(\alpha)\right|^p \mathrm{d}\alpha \ll_p \|\nu\|_1^p N^{-1}.$$

The utility of this property for our purposes is demonstrated in the following lemma, which we prove in §3.

Lemma 2.6 (Restriction implies configuration control)**.** *Suppose that ν is a majorant with $\|\nu\|_1 \ll N$ and which satisfies a restriction estimate at some exponent $2 \leqslant p < s$. Let $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$. Then for functions $f, g : \mathbb{Z} \to \mathbb{C}$ with $|f| \leqslant \nu$ and $|g| \leqslant 1_{[N]}$ we have*

$$\left|\sum_{\mathbf{c}\cdot\mathbf{x}=0} \left(\prod_{i=1}^s f(x_i) - \prod_{i=1}^s g(x_i)\right)\right| \ll_{p,s} N^{s-1} \left(N^{-1}\|\hat{f} - \hat{g}\|_\infty\right)^{1-\{p\}}.$$

Finally, we need to ensure that our majorant does not only count K-trivial solutions to (1.1).

Definition 2.7 (Saves η on the K-trivial solutions). Let $\eta > 0$ and let K denote a finite union of k proper subspaces of the hyperplane (1.1). We say that the majorant ν *saves η on the K-trivial solutions* if

$$\sum_{\mathbf{x} \in K} \prod_{i=1}^{s} \nu(x_i) \ll_{s,k} \|\nu\|_1^s N^{-1-\eta}.$$

With these definitions in hand we are able to prove the following.

Proposition 2.8 (Sufficient majorant conditions). *Let $c_i \in \mathbb{Z} \setminus \{0\}$ with $c_1 + \cdots + c_s = 0$ and $s \geqslant 3$. Let K denote a finite union of k proper subspaces of the hyperplane (1.1). Suppose that ν is a majorant satisfying*

- *the restriction estimate at some exponent $p \in [s - 1, s)$;*
- *Fourier decay of level $\theta \leqslant 1$;*
- *the K-trivial estimate with saving η.*

Then any set $A \subset \operatorname{supp}(\nu)$ containing only K-trivial solutions to (1.1) satisfies

$$\sum_n 1_A(n)\nu(n) \ll_{\mathbf{c},p,k,\eta,\varepsilon} \frac{N}{\min\{\log\log(1/\theta),\ \log N\}^{s-2-\varepsilon}}. \qquad (2.2)$$

Note that in applications we do not expect to be able to take θ any smaller that N^{-c}, in which case the minimum occurring in the right hand side of (2.2) is $\log\log(1/\theta)$.

The statistic

$$\sum_n 1_A(n)\nu(n)$$

is not necessarily the same as the quantity $|A|$, to which Theorem 1.1 pertains. However, in most applications there is simple polynomial dependence between these quantities. We discuss this in the context of our application in Lemma 4.1.

Proof of Proposition 2.8. Define $f := 1_A \nu$ and set

$$\delta := N^{-1} \sum_n f(n).$$

We wish to determine a lower bound on δ in terms of $s, p, k, \eta, \mathbf{c}$ and θ which guarantees that A contains at least one K-nontrivial solution to (1.1). By the bounded approximation lemma, there exists a function $0 \leqslant g \leqslant 1_{[N]}$ such that

$$\|\hat{f} - \hat{g}\|_\infty \ll \frac{N}{\log(1/\theta)^{3/2}}.$$

Since the L^1-norm of a non-negative function is equal to the sup-norm of its Fourier transform, we have

$$\sum_n g(n) = \|\hat{g}\|_\infty \geqslant \|\hat{f}\|_\infty - \|\hat{f} - \hat{g}\|_\infty = \sum_n f(n) - \|\hat{f} - \hat{g}\|_\infty$$

$$= \delta N + O\left(\frac{N}{\log(1/\theta)^{3/2}}\right).$$

Thus we may deduce that

$$\sum_n g(n) \geqslant (\delta/2)N,$$

provided we assume

$$\delta \geqslant \frac{C}{\log(1/\theta)^{3/2}}, \tag{2.3}$$

for some absolute constant C. Bloom's theorem then implies that there exists $c(\delta/2) > 0$ such that

$$\sum_{\mathbf{c}\cdot\mathbf{x}=0} \prod_{i=1}^s g(x_i) \gg_{\mathbf{c}} c(\delta/2)N^{s-1}.$$

Since our majorant has Fourier decay of level $\leqslant 1$, we have $\|\nu\|_1 = \|\hat{\nu}\|_\infty \leqslant \|\hat{\nu} - \hat{1}_{[N]}\|_\infty + \|\hat{1}_{[N]}\|_\infty \ll N$. Thus Lemma 2.6 yields

$$\sum_{\mathbf{c}\cdot\mathbf{x}=0} \prod_{i=1}^s f(x_i) = \sum_{\mathbf{c}\cdot\mathbf{x}=0} \prod_{i=1}^s g(x_i) + O_{s,p}\left(\frac{N^{s-1}}{\log(1/\theta)^{\frac{3(s-p)}{2}}}\right).$$

Hence, if we ensure that

$$c(\delta/2) \geqslant \frac{C_{s,p,\mathbf{c}}}{\log(1/\theta)^{\frac{3(s-p)}{2}}} \tag{2.4}$$

for some absolute constant $C_{s,p,\mathbf{c}}$, we may conclude that

$$\sum_{\mathbf{c}\cdot\mathbf{x}=0} \prod_{i=1}^s f(x_i) \gg_{\mathbf{c}} c(\delta/2)N^{s-1}.$$

Yet if A contains only K-trivial solutions to (1.1), then the fact that ν saves η on the K-trivial solutions gives

$$\sum_{\mathbf{c}\cdot\mathbf{x}=0} \prod_{i=1}^s f(x_i) \leqslant \sum_{\mathbf{x}\in K} \prod_{i=1}^s \nu(x_i) \ll_{s,k} N^{s-1-\eta},$$

again using $\|\nu\|_1 \ll N$. Hence A must contain a K-nontrivial solution on ensuring that

$$c(\delta/2) \geqslant C_{s,k,\mathbf{c}}N^{-\eta} \tag{2.5}$$

for some absolute constant $C_{s,k,\mathbf{c}}$.

Employing Bloom's lower bound for $c(\delta/2)$, our conditions (2.3), (2.4) and (2.5) reduce to ensuring that for any $\varepsilon > 0$ there exists an absolute constant $C_{s,p,k,\eta,\mathbf{c},\varepsilon}$ such that

$$\delta \geqslant C_{s,p,k,\eta,\mathbf{c},\varepsilon} \max\left\{ \frac{1}{\log(1/\theta)^{3/2}}, \frac{1}{\big(\log\log(1/\theta)\big)^{s-2-\varepsilon}}, \frac{1}{(\log N)^{s-2-\varepsilon}} \right\}.$$

The result is now obvious. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Heuristically, our strategy is to show that the following weighted indicator function of the squares is a suitable majorant:

$$\nu(n) := \begin{cases} 2\sqrt{n} & \text{if } n = x^2 \text{ for some } x \in [X], \\ 0 & \text{otherwise.} \end{cases} \qquad (2.6)$$

Taking $N := X^2$ we see that ν is supported on the interval $[N]$ with $\|\nu\|_1 \sim N$. It follows from work of Bourgain [Bou89] that this majorant satisfies a restriction estimate at any exponent $p > 4$. Moreover, a standard point-counting argument (see §7) shows that this majorant saves at least $1/2$ (essentially) on the K-trivial solutions. However, the majorant does not exhibit sufficient Fourier decay. Indeed, it is well-known from the classical circle method that there are values of α for which $\big|\hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha)\big| \gg N$. (For example, take $\alpha = a/q$ with $2 < q \ll 1$.) These problematic values of α arise from the major arcs, and are a consequence of the fact that the squares are not equidistributed in congruence classes to small moduli.

We overcome this difficulty by implementing the W-trick, as found in Green [Gre05]. Our W-tricked version of the majorant (2.6) is found in §4. Although this allows us to verify the existence of non-trivial Fourier decay in §5, the restriction estimate no longer follows directly from Bourgain's work. We show how to modify his proof in §6. Finally, in §7 we show that the W-tricked majorant saves $1/5$ on the K-trivial solutions and in §8 we bring everything together to prove Theorem 1.1 via Proposition 2.8.

3. Configuration control

The aim of this section is to establish Lemma 2.6. Let $c_1, \ldots, c_s \in \mathbb{Z} \backslash \{0\}$ and suppose that ν is a majorant with $\|\nu\|_1 \ll N$ and which satisfies a restriction estimate at exponent $2 \leqslant p < s$. Then we need to show that for any functions $f, g : \mathbb{Z} \to \mathbb{C}$ with $|f| \leqslant \nu$ and $|g| \leqslant 1_{[N]}$ we have

$$\left| \sum_{\mathbf{c}\cdot\mathbf{x}=0} \left(\prod_{i=1}^s f(x_i) - \prod_{i=1}^s g(x_i) \right) \right| \ll_{p,s} N^{s-1} \left(N^{-1}\|\hat{f} - \hat{g}\|_\infty \right)^{1-\{p\}}.$$

On recording the telescoping identity

$$\prod_{i=1}^s f(x_i) - \prod_{i=1}^s g(x_i) = \sum_{j=1}^s \big(f(x_j) - g(x_j)\big) \prod_{i<j} f(x_i) \prod_{i>j} g(x_i),$$

we see that Lemma 2.6 is implied by the following result.

Lemma 3.1 (Configuration control). *Let $c_1, \ldots, c_s \in \mathbb{Z} \setminus \{0\}$. Suppose that ν is a majorant with $\|\nu\|_1 \ll N$ and which satisfies a restriction estimate at some exponent $2 \leqslant p < s$. Then for any functions $f_1, \ldots, f_s : \mathbb{Z} \to \mathbb{C}$, each satisfying $|f_i| \leqslant \nu$ or $|f_i| \leqslant 1_{[N]}$, we have*

$$\left| \sum_{\mathbf{c} \cdot \mathbf{x} = 0} \prod_{i=1}^{s} f_i(x_i) \right| \ll_p N^{s-1} \left(N^{-1} \min_i \|\hat{f}_i\|_\infty \right)^{1 - \{p\}}. \tag{3.1}$$

Proof. Notice that if $|f_i| \leqslant 1_{[N]}$ then, by Parseval and our assumption that $p \geqslant 2$, we have the bound

$$\|\hat{f}_i\|_p \leqslant \|f_i\|_2^{2/p} \|f_i\|_1^{(p-2)/p} \leqslant N^{1 - \frac{1}{p}}.$$

Similarly, if $|f_i| \leqslant \nu$ then our restriction estimate assumption and the bound $\|\nu\|_1 \ll N$ gives that $\|\hat{f}_i\|_p \ll_p N^{1 - \frac{1}{p}}$.

Let $t = \lfloor p \rfloor + 1$. Using orthogonality and the trivial estimate $|\hat{f}_i| \ll N$, we have

$$\sum_{\mathbf{c} \cdot \mathbf{x} = 0} \prod_{i=1}^{s} f_i(x_i) = \int_{\mathbb{T}} \hat{f}_1(c_1 \alpha) \cdots \hat{f}_s(c_s \alpha) \mathrm{d}\alpha \ll N^{s-t} \int_{\mathbb{T}} |\hat{f}_1(c_1 \alpha) \cdots \hat{f}_t(c_t \alpha)| \mathrm{d}\alpha.$$

By Hölder's inequality

$$\int_{\mathbb{T}} |\hat{f}_1(c_1 \alpha) \cdots \hat{f}_t(c_t \alpha)| \mathrm{d}\alpha \leqslant \|\hat{f}_1\|_\infty^{t-p} \|\hat{f}_1\|_p^{1-(t-p)} \prod_{1 < i \leqslant t} \|\hat{f}_i\|_p$$

$$\ll_{t,p} \|\hat{f}_1\|_\infty^{t-p} N^{p-1}$$

$$= \left(\|\hat{f}_1\|_\infty N^{-1} \right)^{t-p} N^{t-1}.$$

The lemma now follows with $i = 1$ on the right-hand side of (3.1). The general case follows on re-ordering the indices of the f_i. $\qquad \square$

4. The W-trick

To prove Theorem 1.1 we need to construct a majorant supported on the squares which satisfies the conditions of Proposition 2.8. We give the initial construction in this section. The idea is that although (2.6) may not possess sufficient Fourier decay, by making an appropriate choice of integers W and b, the Fourier transform of the function

$$n \mapsto \nu(Wn + b)$$

should exhibit the required cancellation.

Set

$$W := 8 \prod_{2 < p \leqslant w} p, \qquad \text{where} \qquad w := \sqrt{\log X}. \tag{4.1}$$

By the prime number theorem, there exists an absolute constant C such that

$$W \leqslant \exp(C\sqrt{\log X}) \ll_\varepsilon X^\varepsilon, \tag{4.2}$$

for any $\varepsilon > 0$.

Let b_1 be a w-smooth number and let $b_2 \in [W]$ with $(b_2, W) = 1$. Notice that every positive square lies in one and only one of the arithmetic progressions

$$b_1^2(W \cdot \mathbb{Z} - b_2).$$

This motivates the definition of our W-tricked majorant function. Set

$$\sigma(b_2) := \# \left\{ z \in [W] : z^2 + b_2 \equiv 0 \mod W \right\}.$$

Then we define the function $\nu_{\mathbf{b}} : \mathbb{Z} \to [0, \infty)$ by

$$\nu_{\mathbf{b}}(n) := \begin{cases} \frac{2\sqrt{Wn-b_2}}{\sigma(b_2)} & \text{if } b_1^2(Wn - b_2) = x^2 \text{ for some } x \in [X], \\ 0 & \text{otherwise.} \end{cases} \qquad (4.3)$$

Note that $\sigma(b_2) \neq 0$ whenever $n \in \mathbb{Z}$ is such that $b_1^2(Wn - b_2) = x^2$ for some $x \in [X]$. Indeed, in this case, $-b_2$ must be a square modulo W, which is equivalent to requiring $-b_2$ to be a square modulo 8 and a square modulo every odd prime $p \leqslant w$. There are 4 possible values of $z \mod 8$ such that $z^2 \equiv -b_2 \mod 8$ and 2 values of $z \mod p$ such that $z^2 \equiv -b_2 \mod p$ for any odd prime prime $p \leqslant w$. Hence the Chinese remainder theorem yields $\sigma(b_2) = 2 \cdot 2^{\pi(w)}$. Furthermore, on writing

$$N_{\mathbf{b}} := \left\lfloor \frac{X^2}{b_1^2 W} \right\rfloor + 1, \qquad (4.4)$$

one can check that $\nu_{\mathbf{b}}$ is supported on $[N_{\mathbf{b}}]$ and that

$$\sum_n \nu_{\mathbf{b}}(n) = N_{\mathbf{b}} + O(1 + Xb_1^{-1}) = N_{\mathbf{b}} + O(\sqrt{WN_{\mathbf{b}}}). \qquad (4.5)$$

Given a set of integers A, let

$$A_{\mathbf{b}} := \left\{ n \in \mathbb{Z} : b_1^2(Wn - b_2) = x^2 \text{ for some } x \in A \right\}.$$

We proceed by establishing the following result.

Lemma 4.1. *There exists an absolute constant C such that the following holds for any $\delta \geqslant C(\log X)^{-1}$. Let $A \subset [X]$ with $|A| = \delta X$. Then there exists a w-smooth number $b_1 \leqslant X^{1/2}$ and there exists $b_2 \in [W]$ with $(b_2, W) = 1$ such that*

$$\sum_n 1_{A_{\mathbf{b}}}(n) \nu_{\mathbf{b}}(n) \gg \delta^2 N_{\mathbf{b}}.$$

Consideration of the set $A = [\delta X]$ shows that the lower bound in this estimate is optimal.

Proof of Lemma 4.1. It follows from Rankin's trick that the number of integers in $[X]$ divisible by a w-smooth number exceeding $X^{1/2}$ is at most

$$\sum_{\substack{b_1 > X^{1/2} \\ b_1 \text{ } w\text{-smooth}}} \frac{X}{b_1} \ll \sum_{b_1 \text{ } w\text{-smooth}} \frac{X}{b_1} \left(\frac{b_1}{X^{1/2}} \right)^{1/2} \ll X^{3/4} \sum_{b_1 \text{ } w\text{-smooth}} \frac{1}{b_1^{1/2}}$$

$$\ll X^{3/4} \prod_{p < w} \left(1 + \frac{1}{p^{1/2}} \right)$$

$$\ll X^{3/4+\varepsilon},$$

for any $\varepsilon > 0$. It follows that

$$\delta X = \sum_{\substack{b_1 \leqslant X^{1/2} \\ b_1 \text{ is } w\text{-smooth}}} \sum_{\substack{b_2 \in [W] \\ (b_2, W) = 1}} |A_{\mathbf{b}}| + O_\varepsilon \left(X^{3/4+\varepsilon} \right).$$

Our assumption that $\delta \geqslant C (\log X)^{-1}$ therefore ensures that

$$\delta X \ll \sum_{\substack{b_1 \leqslant X^{1/2} \\ b_1 \text{ is } w\text{-smooth}}} \sum_{\substack{b_2 \in [W] \\ (b_2, W) = 1}} |A_{\mathbf{b}}|.$$

Similarly,

$$\sum_{\substack{b_1 \leqslant X^{1/2} \\ b_1 \text{ is } w\text{-smooth}}} \sum_{\substack{b_2 \in [W] \\ (b_2, W) = 1}} |[X]_{\mathbf{b}}| \ll X,$$

so that

$$\delta \sum_{\substack{b_1 \leqslant X^{1/2} \\ b_1 \text{ is } w\text{-smooth}}} \sum_{\substack{b_2 \in [W] \\ (b_2, W) = 1}} |[X]_{\mathbf{b}}| \ll \sum_{\substack{b_1 \leqslant X^{1/2} \\ b_1 \text{ is } w\text{-smooth}}} \sum_{\substack{b_2 \in [W] \\ (b_2, W) = 1}} |A_{\mathbf{b}}|.$$

Applying the pigeon-hole principle, we conclude that there exists a w-smooth number $b_1 \leqslant X^{1/2}$ and $b_2 \in [W]$ with $(b_2, W) = 1$ such that

$$\delta |[X]_{\mathbf{b}}| \ll |A_{\mathbf{b}}|.$$

For each $z \in [W]$ with $z^2 + b_2 \equiv 0 \bmod W$, write $A_{\mathbf{b}, z}$ for the set of positive integers $x \equiv z \bmod W$ such that $x^2 = Wn - b_2$ for some $n \in A_{\mathbf{b}}$. Together these sets satisfy

$$\sum_{z^2 + b_2 \equiv 0 \bmod W} |A_{\mathbf{b}, z}| = |A_{\mathbf{b}}| \gg \delta |[X]_{\mathbf{b}}|.$$

For any $T > 0$ we claim that there are at most $T + 1$ values of $x \in A_{\mathbf{b}, z}$ satisfying $x \leqslant WT$. This follows on noting that elements of $A_{\mathbf{b}, z}$ take the shape $z + Wy$ for suitable y. But there are clearly at most $T - z/W + 1$ choices of y satisfying $y \leqslant T - z/W$, from which the claim follows. Applying

the claim with $T = \frac{1}{2}|A_{\mathbf{b},z}|$, we deduce that for at least $\frac{1}{2}|A_{\mathbf{b},z}| - 1$ values of $x \in A_{\mathbf{b},z}$ we have $x > \frac{1}{2}W|A_{\mathbf{b},z}|$. It follows from (4.3) that

$$
\sum_n 1_{A_{\mathbf{b}}}(n)\nu_{\mathbf{b}}(n) = \frac{1}{\sigma(b_2)} \sum_{z^2+b_2 \equiv 0 \bmod W} \sum_{x \in A_{\mathbf{b},z}} 2x
$$

$$
\geqslant \frac{2}{\sigma(b_2)} \sum_{z^2+b_2 \equiv 0 \bmod W} (\tfrac{1}{2}|A_{\mathbf{b},z}| - 1)\tfrac{1}{2}W|A_{\mathbf{b},z}|.
$$

An application of the Cauchy–Schwarz inequality now shows that the right hand side is

$$
> \frac{W}{2}\left(\left(\frac{1}{\sigma(b_2)} \sum_{z^2+b_2 \equiv 0 \bmod W} |A_{\mathbf{b},z}| \right)^2 - 2\frac{1}{\sigma(b_2)} \sum_{z^2+b_2 \equiv 0 \bmod W} |A_{\mathbf{b},z}| \right)
$$

$$
= \frac{W}{2}\left(\left(\frac{|A_{\mathbf{b}}|}{\sigma(b_2)} \right)^2 - 2\frac{|A_{\mathbf{b}}|}{\sigma(b_2)} \right).
$$

Recalling that $b_1 \leqslant X^{1/2}$ and $W \ll_\varepsilon X^\varepsilon$, provided we take C sufficiently large in our lower bound $\delta \geqslant C(\log X)^{-1}$, we may proceed under the assumption that $|A_{\mathbf{b}}| \geqslant 4\sigma(b_2)$. But then it follows that

$$
\left(\left(\frac{|A_{\mathbf{b}}|}{\sigma(b_2)} \right)^2 - 2\frac{|A_{\mathbf{b}}|}{\sigma(b_2)} \right) \geqslant \frac{1}{2}\left(\frac{|A_{\mathbf{b}}|}{\sigma(b_2)} \right)^2.
$$

Since we have

$$
|[X]_{\mathbf{b}}| = \sigma(b_2)\left(\frac{X}{b_1 W} + O(1) \right),
$$

it finally follows that

$$
\sum_n 1_{A_{\mathbf{b}}}(n)\nu_{\mathbf{b}}(n) \gg \delta^2\left(\frac{X^2}{b_1^2 W} \right) \gg \delta^2 N_{\mathbf{b}},
$$

as required to complete the proof of the lemma. $\qquad \square$

5. Exponential sum estimates

Henceforth we fix a choice of $\mathbf{b}$ afforded by Lemma 4.1 and denote both $\nu_{\mathbf{b}}$ and $N_{\mathbf{b}}$ by ν and N, respectively. For $z \in [W]$ such that $z^2 + b_2 \equiv 0 \bmod W$, write

$$
S_q(a,z) := \sum_{r=1}^q e\left(\frac{a(Wr^2 + 2zr + \frac{z^2+b_2}{W})}{q} \right) \quad \text{and} \quad I(\beta) := \int_0^N e(\beta t)\mathrm{d}t.
$$

Bearing these quantities in mind, we may now establish the following result.

Lemma 5.1 (Major arc asymptotic)**.** *Suppose that $\|q\alpha\| = |q\alpha - a|$ for $q, a \in \mathbb{Z}$ with $q > 0$. Then*

$$\hat{\nu}(\alpha) = \frac{1}{\sigma(b_2)} \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} S_q(a, z) I(\alpha - \tfrac{a}{q}) + O\left(\sqrt{NW}(q + N\|q\alpha\|)\right).$$

Proof. Recalling (2.1) and (4.3), we break the Fourier transform into congruence classes modulo W to give

$$\hat{\nu}(\alpha) = \frac{1}{\sigma(b_2)} \sum_{\substack{x \leqslant Xb_1^{-1} \\ x^2 + b_2 \equiv 0 \bmod W}} 2x \, e\big(\alpha(x^2 + b_2)/W\big)$$

$$= \frac{e(\alpha b_2/W)}{\sigma(b_2)} \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} \sum_{z + Wy \leqslant Xb_1^{-1}} 2(z + Wy) \, e\big(\alpha(z + Wy)^2/W\big).$$

We now write $\alpha = \frac{a}{q} + \beta$ and break the sum over y into residue classes modulo q. This gives

$$\sum_{z + Wy \leqslant Xb_1^{-1}} 2(z + Wy) \, e\big(\alpha(z + Wy)^2/W\big)$$

$$= \sum_{r=1}^{q} \sum_{\substack{z + Wy \leqslant Xb_1^{-1} \\ y \equiv r \bmod q}} 2(z + Wy) \, e\big((\tfrac{a}{q} + \beta)(z + Wy)^2/W\big)$$

$$= \sum_{r=1}^{q} e\big(a(z + Wr)^2/(qW)\big) \sum_{z + Wr + Wqx \leqslant Xb_1^{-1}} \phi(x),$$

where $\phi(x) := 2(z + Wr + Wqx) \, e\big(\beta(z + Wr + Wqx)^2/W\big)$.

We shall use Euler–Maclaurin to estimate the inner sum over x (in the form [Vau97, Eq. (4.8)], for example). Taking $Y_0 = (Xb_1^{-1} - (z + Wr))/(Wq)$ and $X_0 = -(z + Wr)/(Wq)$, this yields

$$\sum_{X_0 < x \leqslant Y_0} \phi(x) = \int_{X_0}^{Y_0} \phi(t) \mathrm{d}t - [\phi(t)\psi(t)]_{X_0}^{Y_0} + \int_{X_0}^{Y_0} \phi'(t)\psi(t)\mathrm{d}t,$$

where $\psi(t) = \{t\} - \frac{1}{2} = O(1)$. Next we make the change of variables $u = (z + Wr + Wqt)^2/W$, under which the first term becomes

$$\int_{X_0}^{Y_0} \phi(t)\mathrm{d}t = q^{-1} \int_0^{X^2/(b_1^2 W)} e(\beta t)\mathrm{d}t = q^{-1}(I(\beta) + O(1)).$$

Next, on noting that $|\phi(t)| \leqslant 2Xb_1^{-1}$ on $(X_0, Y_0]$ and

$$\sup_{t \in (X_0, Y_0]} |\phi'(t)| \ll W(q + N\|q\alpha\|),$$

we find that the remaining two terms are

$$\ll \frac{X}{b_1} + \frac{X}{b_1 W q} W(q + N\|q\alpha\|) \ll \frac{\sqrt{NW}}{q}(q + N\|q\alpha\|).$$

Inserting this into our previous estimate, we conclude that

$$\hat{\nu}(\alpha) = \frac{e(\alpha b_2/W)}{\sigma(b_2)} \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} I(\alpha - \tfrac{a}{q}) \sum_{r=1}^{q} e\big(a(z + Wr)^2/(qW)\big)$$
$$+ O\left(\sqrt{NW}(q + N\|q\alpha\|)\right).$$

Now the inner exponential sum over r is equal to $S_q(a, z)e(-ab_2/(qW))$. The lemma therefore follows on noting that

$$e(\alpha b_2/W)e(-ab_2/(qW))I(\alpha - \tfrac{a}{q}) = \int_0^N e\left((\alpha - \tfrac{a}{q})(t + \tfrac{b_2}{W})\right) dt$$
$$= I(\alpha - \tfrac{a}{q}) + O(1),$$

since $b_2/W \in [0, 1]$. $\qquad\square$

We need the following basic estimates for the quantities featuring in Lemma 5.1.

Lemma 5.2. *For $(a, q) = 1$ we have*

$$|S_q(a, z)| \leqslant 2\sqrt{q} \quad and \quad I(\beta) \leqslant \min\left\{N, \frac{1}{\|\beta\|}\right\}.$$

Proof. The estimate for $I(\beta)$ is standard and so we focus on $S_q(a, z)$. Let $h = (q, W)$ and put $q = hq'$ and $W = hW'$, with $(q', W') = 1$. We then write $r = r_1 + q'r_2$ for r_1 running modulo q' and r_2 running modulo h, finding that

$$S_q(a, z) = e\left(\frac{a(z^2 + b_2)}{qW}\right) \sum_{r=1}^{q} e\left(\frac{a(Wr^2 + 2zr)}{q}\right)$$
$$= e\left(\frac{a(z^2 + b_2)}{qW}\right) \sum_{r_1=1}^{q'} e\left(\frac{a(W'r_1^2 + 2zr_1/h)}{q'}\right) \sum_{r_2=1}^{h} e\left(\frac{2azr_2}{h}\right).$$

The inner sum vanishes unless $h \mid 2az$. But this happens if and only if $h \mid 2$, since $z^2 + b_2 \equiv 0 \bmod W$ and $(a, q) = (b_2, W) = 1$. Hence $S_q(a, z) = 0$ if $h \nmid 2$ and

$$S_q(a, z) = h\, e\left(\frac{a(z^2 + b_2)}{qW}\right) \sum_{r_1=1}^{q'} e\left(\frac{a(W'r_1^2 + 2zr_1/h)}{q'}\right). \tag{5.1}$$

if $h \mid 2$. Assuming that $h \mid 2$, we see that the desired bound for $S_q(a, z)$ is an immediate consequence of the basic estimate

$$\left| \sum_{r=1}^{s} e\left(\frac{br^2 + cr}{s} \right) \right| \leqslant \sqrt{2s},$$

which is valid for any positive integer s and any integers b, c with b coprime to s. The latter estimate follows on invoking the squaring and differencing approach found in the standard analysis of Weyl sums. $\square$

Building on the proof of this result we may establish the following.

Lemma 5.3. *If $q > 1$ and q is w-smooth then*

$$\sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} S_q(a, z) = 0.$$

Proof. Suppose that q is w-smooth and bigger than 1. We saw in the proof of the previous result that $S_q(a, z) = 0$ unless $(q, W) \mid 2$. This implies that $q = 2^h$ for some $h \geqslant 1$, since q is w-smooth. Since $8 \mid W$ we must have $\min\{h, 3\} = 2$, whence $q = 2$ and $a = 1$. Substituting these values into (5.1), we deduce that

$$S_2(1, z) = 2e\left(\frac{z^2 + b_2}{2W} \right).$$

Let $\tilde{W}$ be the odd integer such that $W = 8\tilde{W}$. We now introduce the sum over z, writing $z = u + 4\tilde{W}v$ for u running modulo $4\tilde{W}$ and v running modulo 2. In particular, we have

$$e\left(\frac{z^2 + b_2}{2W} \right) = e\left(\frac{u^2 + b_2}{2W} \right) e\left(\frac{uv}{2} \right)$$

and $z^2 + b_2 \equiv 0 \bmod W$ if and only if $u^2 + b_2 \equiv 0 \bmod W$. We must have $2 \nmid u$ since $z^2 + b_2 \equiv 0 \bmod W$ and $(b_2, W) = 1$. But then

$$\sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} S_q(a, z) = \sum_{\substack{u \bmod 4\tilde{W} \\ u^2 + b_2 \equiv 0 \bmod W}} e\left(\frac{u^2 + b_2}{2W} \right) \sum_{v \bmod 2} e\left(\frac{uv}{2} \right) = 0,$$

as required. $\square$

Recalling (4.5), we have the trivial estimate $\hat{\nu}(\alpha) \ll N$. Our next result is a standard Weyl bound, showing how one can beat the trivial estimate when α is not well approximated by a rational with small denominator.

Lemma 5.4 (Weyl bound). *Suppose that $b_1 W \leqslant X$ and $\|q\alpha\| = |q\alpha - a|$ for coprime $a, q \in \mathbb{Z}$ with $q > 0$. Then*

$$|\hat{\nu}(\alpha)| \ll N\sqrt{W \log N}$$
$$\times \left((WN)^{-1/2} + \|q\alpha\| + qN^{-1} + \min\left\{ q^{-1}, (\|q\alpha\| N)^{-1} \right\} \right)^{1/2}.$$

Proof. As in the proof of Lemma 5.1, we have

$$|\hat{\nu}(\alpha)| \leqslant \frac{1}{\sigma(b_2)} \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} \left| \sum_{z + Wy \leqslant Xb_1^{-1}} 2(z + Wy)\, e\big(\alpha(Wy^2 + 2zy)\big) \right|.$$

By partial summation and the assumption that $b_1 W \leqslant X$, one sees that

$$\left| \sum_{z + Wy \leqslant Xb_1^{-1}} 2(z + Wy)\, e\big(\alpha(Wy^2 + 2zy)\big) \right|$$

$$\ll Xb_1^{-1} \sup_{Y \leqslant X/(b_1 W)} \left| \sum_{0 \leqslant y \leqslant Y} e\big(\alpha(Wy^2 + 2zy)\big) \right|. \quad (5.2)$$

Let us denote temporarily by $S(Y)$ the inner sum in (5.2). This is an exponential sum over a polynomial with unbounded height and we require a version of Weyl's inequality applicable to this situation. A standard Weyl differencing argument gives that $S(Y)$ is

$$\ll Y (\log Y)^{1/2} \left(Y^{-1} + \|q\alpha\|\, W + qY^{-2} + \min\left\{ Wq^{-1}, (\|q\alpha\|\, Y^2)^{-1} \right\} \right)^{1/2}.$$

(cf. [BP14, Lemma 3.3] with $(d, n, \sigma) = (2, 1, 0)$ and $(P, H) = (Y, W)$ in the notation therein.) One can check that this bound for $S(Y)$ is increasing in Y, and so it is maximised when $Y = X/(b_1 W)$. Since $\sqrt{N/W} \ll X/(b_1 W) \ll \sqrt{N/W}$, we conclude that the right hand side of (5.2) is

$$\ll Xb_1^{-1} \cdot \sqrt{N \log N \left((NW)^{-1/2} + \|q\alpha\| + qN^{-1} + \min\left\{ q^{-1}, (\|q\alpha\|\, N)^{-1} \right\} \right)}.$$

The statement of the lemma is now obvious. $\qquad\square$

Note that $b_1 W \leqslant X$ since $b_1 \leqslant X^{1/2}$ and $W \ll_\varepsilon X^\varepsilon$ by (4.2), and so the hypothesis of the preceding result is met. We end this section by showing that our majorant function has suitable Fourier decay.

Lemma 5.5. *We have the estimate*

$$\|\hat{\nu} - \hat{1}_{[N]}\|_\infty \ll \frac{N}{\sqrt{w}}$$

Proof. Let $\tau = \frac{1}{100}$. We define a set of major arcs $\mathfrak{M}$ to be the union of all intervals

$$\mathfrak{M}(r, b) = \{\alpha \in \mathbb{T} : |\alpha - b/r| \leqslant N^{-1+\tau}\}$$

for coprime integers b, r such that $0 \leqslant b < r \leqslant N^\tau$. Now let $\alpha \notin \mathfrak{M}$. By Dirichlet's approximation theorem we can find coprime integers a, q such that $q \leqslant N^\tau$ and $|\alpha - \frac{a}{q}| \leqslant q^{-1} N^{-\tau}$. Since α is not on the major arcs we must have $|q\alpha - a| > qN^{-1+\tau}$. It therefore follows from Lemma 5.4 that

$$\hat{\nu}(\alpha) \ll N^{1-\tau/2} \sqrt{W \log N}. \quad (5.3)$$

Since

$$\hat{1}_{[N]}(\alpha) = \sum_{1 \leqslant n \leqslant N} e(\alpha n) \ll \frac{q}{\|q\alpha\|} \ll N^{1-\tau},$$

and $\sqrt{W \log N} \ll N^{\tau/4}$ by (4.2), we conclude that

$$\left| \hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha) \right| \ll N^{1-\tau/4}$$

for any $\alpha \notin \mathfrak{M}$.

Next let $\alpha \in \mathfrak{M}(q, a)$ for coprime $a, q \in \mathbb{Z}$ with $0 \leqslant a < q \leqslant N^\tau$. According to Lemma 5.1 we have

$$\hat{\nu}(\alpha) = \frac{1}{\sigma(b_2)} \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} S_q(a, z) I(\alpha - \tfrac{a}{q}) + O\left(N^{1/2+3\tau}\right), \qquad (5.4)$$

since $q + N \|q\alpha\| \ll q + N \cdot qN^{-1+\tau} \ll qN^\tau \ll N^{2\tau}$ and $W \leqslant N^\tau$. In a similar fashion (see [Vau97, Lemma 2.7], for example), one finds that

$$\hat{1}_{[N]}(\alpha) = \begin{cases} O(N^{2\tau}) & \text{if } q > 1, \\ I(\alpha) + O(N^{2\tau}) & \text{if } q = 1. \end{cases}$$

The main term in (5.4) has absolute value

$$\leqslant \frac{N}{\sigma(b_2)} \left| \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} S_q(a, z) \right|,$$

since $|I(\beta)| \leqslant N$. Lemma 5.3 implies that this vanishes unless either $q = 1$ or q is not w-smooth. In the latter case we may conclude that $q > w$, whence an application of Lemma 5.2 shows that the main term has absolute value at most $2N/\sqrt{q} \leqslant 2N/\sqrt{w}$. It follows that

$$\left| \hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha) \right| \ll \frac{N}{\sqrt{w}} + N^{1/2+3\tau} + N^{2\tau} \ll \frac{N}{\sqrt{w}}$$

for any $\alpha \in \bigcup_{q \neq 1} \bigcup_{a \bmod q} \mathfrak{M}(q, a)$.

Finally, when $q = 1$ and $\alpha \in \mathfrak{M}(1, 0)$, we deduce from (5.4) that

$$\hat{\nu}(\alpha) = I(\alpha) + O\left(N^{1/2+3\tau}\right),$$

since $S_1(0, z) = 1$. Hence

$$\left| \hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha) \right| \ll N^{1/2+3\tau}$$

for any $\alpha \in \mathfrak{M}(1, 0)$. This completes the proof of the lemma. $\qquad \square$

6. The restriction estimate

We require an analogue of Bourgain's restriction estimate [Bou89] for the W-tricked quadratic exponential sum. The proof of the following result closely follows the argument in [Bou89, §4].

Proposition 6.1 (Restriction bound). *For any $p > 4$ there exists an absolute constant C_p such that for any function $\phi : \mathbb{Z} \to \mathbb{C}$ satisfying $|\phi| \leqslant \nu$ we have the bound*

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^p \mathrm{d}\alpha \leqslant C_p N^{p-1}.$$

The following result is concerned with calculating the fourth moment of $\hat{\phi}$, since it plays an important role in the proof of Proposition 6.1.

Lemma 6.2. *There exists an absolute constant C such that for any function $\phi : \mathbb{Z} \to \mathbb{C}$ satisfying $|\phi| \leqslant \nu$ we have the bound*

$$\int_{\mathbb{T}} |\hat{\phi}(\alpha)|^4 \mathrm{d}\alpha \leqslant N^{3+C/\log\log N}$$

Proof. The left hand side is equal to

$$\|\hat{\phi}\|_4^4 = \left\| \sum_{m,n} \phi(m)\bar{\phi}(n) e\left(\alpha(m-n)\right) \right\|_2^2 \leqslant \sum_{|k|\leqslant N} \left| \sum_{\substack{m,n \\ m-n=k}} \phi(m)\bar{\phi}(n) \right|^2.$$

The contribution to the sum over k from $k = 0$ is

$$\leqslant \left(\sum_n |\phi(n)|^2 \right)^2 \leqslant \left(\sum_n \nu(n)^2 \right)^2 \ll \left(Xb_1^{-1}N \right)^2 \ll WN^3,$$

by (4.3), (4.4) and (4.5). To handle the contribution from $k \neq 0$, the definition of ν makes it clear that

$$\sum_{\substack{m,n \\ m-n=k}} \phi(m)\bar{\phi}(n) = \sum_{\substack{x,y\leqslant X/b_1 \\ x^2\equiv y^2\equiv -b_2 \bmod W \\ x^2-y^2=Wk}} \phi\left(\frac{x^2+b_2}{W} \right) \bar{\phi}\left(\frac{y^2+b_2}{W} \right)$$

For fixed non-zero k there are clearly at most $d(k)$ choices for x, y. Hence Cauchy–Schwarz gives

$$\|\hat{\phi}\|_4^4 \leqslant \sum_{0<|k|\leqslant N} d(k) \sum_{\substack{x,y\leqslant X/b_1 \\ x^2\equiv y^2\equiv -b_2 \bmod W \\ x^2-y^2=Wk}} \nu\left(\frac{x^2+b_2}{W} \right)^2 \nu\left(\frac{y^2+b_2}{W} \right)^2 + O(WN^3)$$

The standard estimate for the divisor function yields $d(k) \leqslant N^{O(1/\log\log N)}$, for $0 < |k| \leqslant N$. Hence

$$\|\hat{\phi}\|_4^4 \leqslant N^{O(1/\log\log N)} \left(\sum_{\substack{x \leqslant X/b_1 \\ x^2+b_2 \equiv 0 \bmod W}} \nu\left(\frac{x^2-b_2}{W}\right)^2 \right)^2 + O(WN^3)$$

$$\leqslant N^{O(1/\log\log N)} \left(X b_1^{-1} N \right)^2 + O(WN^3)$$

$$\leqslant W N^{3+O(1/\log\log N)},$$

as before. Finally, since $X/b_1 \geqslant X^{1/2}$, it follows from (4.2) that W is at most $N^{O(1/\log\log N)}$. This concludes the proof of the lemma. $\square$

Define the region

$$\mathscr{R}_\delta = \left\{ \alpha \in \mathbb{T} : |\hat{\phi}(\alpha)| > \delta N \right\},$$

for any $\delta \in (0,1)$. Proposition 6.1 is a consequence of the following result.

Lemma 6.3. *For any $\delta \in (0,1)$ and any $\varepsilon > 0$ there exists a constant C_ε depending only ε such that*

$$\mathrm{meas}(\mathscr{R}_\delta) \leqslant \frac{C_\varepsilon}{\delta^{4+\varepsilon} N}.$$

Although standard, let us spell out how Lemma 6.3 implies Proposition 6.1. Breaking into dyadic intervals it follows from Lemma 6.3 that

$$\|\hat{\phi}\|_p^p \leqslant \sum_{j \geqslant 0} \left(\frac{N}{2^{j-1}}\right)^p \mathrm{meas}\left\{ \alpha \in \mathbb{T} : \frac{N}{2^j} < |\hat{\phi}(\alpha)| \leqslant \frac{N}{2^{j-1}} \right\}$$

$$\leqslant C_\varepsilon 2^p N^{p-1} \sum_{j \geqslant 0} (2^{4+\varepsilon-p})^j.$$

Taking $\varepsilon = \varepsilon(p) > 0$ sufficiently small we can ensure that this geometric series converges, which thereby gives Proposition 6.1.

We begin the proof of Lemma 6.3 with a cruder upper bound based on our fourth moment estimate. Thus Lemma 6.2 implies that there exists an absolute constant $C > 0$ such that

$$(\delta N)^4 \mathrm{meas}(\mathscr{R}_\delta) \leqslant \|\hat{\phi}\|_4^4 \leqslant N^{3+C/\log\log N},$$

whence

$$\mathrm{meas}(\mathscr{R}_\delta) \leqslant \frac{N^{C/\log\log N}}{\delta^4 N}.$$

In particular, Lemma 6.3 follows from this bound unless

$$\delta > N^{-C(\varepsilon \log\log N)^{-1}}, \tag{6.1}$$

as we henceforth assume.

As in the proof of Lemma 5.5, let $\tau = \frac{1}{100}$. To each fraction a/q, with $0 \leqslant a < q \leqslant N^\tau$ and $(a, q) = 1$, we associate the major arc

$$\mathfrak{M}(q, a) = \{\alpha \in \mathbb{T} : |\alpha - a/q| \leqslant N^{-1+\tau}\}.$$

We let $\mathfrak{M}$ denote the union of all major arcs. As in (5.3) we have

$$\hat{\nu}(\alpha) \ll N^{1-\tau/4} \quad \text{if } \alpha \notin \mathfrak{M}. \tag{6.2}$$

On the major arcs we can conclude from (5.4) and Lemma 5.2 that

$$|\hat{\nu}(\alpha)| \leqslant \frac{1}{\sigma(b_2)} \sum_{\substack{z \in [W] \\ z^2 + b_2 \equiv 0 \bmod W}} q^{-1} |S_q(a, z) I(\alpha - \tfrac{a}{q})| + O(N^{1/2+2\tau}) \tag{6.3}$$

$$\leqslant \frac{2}{\sqrt{q}} \min\left\{N, \|\alpha - \tfrac{a}{q}\|^{-1}\right\} + O(N^{1/2+2\tau})$$

if $\alpha \in \mathfrak{M}(q, a)$ for some a, q satisfying $0 \leqslant a < q \leqslant N^\tau$ and $(a, q) = 1$.

Returning to our analysis of $\hat{\phi}$, let $\theta_1, \ldots, \theta_R$ be $1/N$-spaced points in $\mathbb{T}$ such that $|\hat{\phi}(\theta_r)| > \delta N$, for $1 \leqslant r \leqslant R$. In order to prove Lemma 6.3 it will suffice to show that

$$R \leqslant \frac{C_\varepsilon}{\delta^{4+\varepsilon}}. \tag{6.4}$$

To see this, let $R \geqslant 1$ be the largest integer such that there exists a sequence of $1/N$-spaced points $\theta_1, \ldots, \theta_R$ in $\mathscr{R}_\delta$. Then every $\theta \in \mathscr{R}_\delta$ necessarily satisfies $\|\theta - \theta_r\| < 1/N$ for some r. This implies that the measure of $\mathscr{R}_\delta$ does not exceed $2R/N$, as required.

To prove (6.4), we let $a_n \in \mathbb{C}$ be such that $|a_n| \leqslant 1$ and $\phi(n) = a_n \nu(n)$ for integers $1 \leqslant n \leqslant N$. Furthermore, let $c_1, \ldots, c_R \in \mathbb{C}$ be such that $|c_r| = 1$ and

$$c_r \hat{\phi}(\theta_r) = |\hat{\phi}(\theta_r)|$$

for $1 \leqslant r \leqslant R$. Then it follows from the Cauchy–Schwarz inequality that

$$\delta^2 N^2 R^2 \leqslant \left(\sum_{1 \leqslant r \leqslant R} |\hat{\phi}(\theta_r)|\right)^2$$

$$= \left(\sum_{1 \leqslant r \leqslant R} c_r \sum_n a_n \nu(n) e(n\theta_r)\right)^2$$

$$\ll N \sum_n \nu(n) \left|\sum_{1 \leqslant r \leqslant R} c_r e(n\theta_r)\right|^2,$$

so that

$$\delta^2 N R^2 \ll \sum_{1 \leqslant r, r' \leqslant R} |\hat{\nu}(\theta_r - \theta_{r'})|.$$

Now let $\gamma > 2$ be fixed. It follows from Hölder's inequality that

$$\delta^{2\gamma} N^{\gamma} R^2 \ll \sum_{1 \leqslant r, r' \leqslant R} |\hat{\nu}(\theta_r - \theta_{r'})|^{\gamma} .$$

Let us put $\theta = \theta_r - \theta_{r'}$ for given $r \neq r'$. It follows from (6.1) and (6.2) that $|\hat{\nu}(\theta)|$ is negligible when $\theta \notin \mathfrak{M}$. When $\theta \in \mathfrak{M}$ we use the major arc approximation recorded in (6.3) to conclude that

$$\hat{\nu}(\theta) \ll \frac{N}{\sqrt{q}}(1 + N\|\theta - \tfrac{a}{q}\|)^{-1} + O(N^{1/2 + 2\tau}),$$

for some a, q satisfying $0 \leqslant a < q \leqslant N^{\tau}$ and $(a, q) = 1$. Now let $Q = \delta^{-5}$. If $q \geqslant Q$ then we simply take the upper bound

$$\frac{N}{\sqrt{q}}(1 + N\|\theta - \tfrac{a}{q}\|)^{-1} \ll \delta^{5/2} N.$$

Recalling that $\delta < 1$ satisfies the lower bound in (6.1), this leads to the conclusion that

$$\delta^{2\gamma} R^2 \ll \sum_{q \leqslant Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} \sum_{1 \leqslant r, r' \leqslant R} \frac{q^{-\gamma/2}}{(1 + N\|\theta_r - \theta_{r'} - \tfrac{a}{q}\|)^{\gamma}}.$$

Hence

$$\delta^{2\gamma} R^2 \ll \sum_{1 \leqslant r, r' \leqslant R} G(\theta_r - \theta_{r'}),$$

where

$$G(\alpha) := \sum_{q \leqslant Q} \sum_{\substack{a \bmod q \\ (a,q)=1}} q^{-\gamma/2}(1 + N|\sin(\alpha - \tfrac{a}{q})|)^{-\gamma}$$

and we recall that $\theta_1, \ldots, \theta_R$ is a sequence of $1/N$-spaced points in $\mathbb{T}$.

   We have finally arrived at exactly the same expression as that found in [Bou89, Eq. (4.16)], but with N^2 replaced by N. Bourgain's argument therefore directly applies and leads us to the desired bound (6.4) for R.

7. The K-trivial count

   The purpose of this section is to show that $\nu = \nu_{\mathbf{b}}$ saves something over the trivial estimate when counting solutions to (1.1) contained in a proper subspace.

Lemma 7.1. *Let $\varepsilon > 0$ and let $K \subset \mathbb{Q}^s$ denote a finite union of at most k proper subspaces of the hyperplane*

$$c_1 y_1 + \cdots + c_s y_s = 0.$$

Then the number of $\mathbf{x} \in [X]^s$ with $(x_1^2, \ldots, x_s^2) \in K$ is $O_{k,s,\varepsilon}(X^{s - \frac{5}{2} + \varepsilon})$.

Proof. The set K is a union of at most k subspaces of the form

$$\{\mathbf{y} \in \mathbb{Q}^s : \mathbf{c} \cdot \mathbf{y} = 0 \text{ and } \mathbf{d} \cdot \mathbf{y} = 0\},$$

where $\mathbf{d} \in \mathbb{Q}^s$ is a fixed s-tuple which is not proportional to $\mathbf{c}$. Our task is therefore to count solutions $\mathbf{x} \in [X]^s$ to the simultaneous quadrics

$$c_1 x_1^2 + \cdots + c_s x_s^2 = d_1 x_1^2 + \cdots + d_s x_s^2 = 0.$$

This pair of equations defines a projective subvariety $V \subset \mathbb{P}_{\mathbb{Q}}^{s-1}$ of codimension 2. Let Λ be any k-plane contained in V. Then Λ is contained in the non-singular quadric hypersurface $c_1 x_1^2 + \cdots + c_s x_s^2 = 0$, in $\mathbb{P}_{\mathbb{Q}}^{s-1}$. It follows that $k \leqslant (s-2)/2$ (see Harris [Har13, Chap. 22], for example). Since $s \geqslant 5$ we conclude that V does not contain any linear component of maximal dimension. Thus we may apply work of Pila [Pil95] to each of the irreducible components of V, concluding that the number of solutions $\mathbf{x} \in [X]^s$ to the equations defining V is $O_{s,\varepsilon}(X^{s-3+\frac{1}{2}+\varepsilon})$, for any $\varepsilon > 0$. This completes the proof of the lemma. $\square$

Corollary 7.2. *Let $K \subset \mathbb{Q}^s$ denote a finite union of at most k proper subspaces of the hyperplane (1.1). Then, in the terminology of Definition 2.7, the majorant $\nu = \nu_{\mathbf{b}}$ saves $1/5$ on the K-trivial solutions.*

Proof. We wish to estimate $\sum_{\mathbf{n} \in K} \prod_{i=1}^s \nu(n_i)$. If $\mathbf{n}$ is counted by this sum, then on setting $x_i = \sqrt{W n_i - b_2}$, we obtain integers $x \in [X/b_1]$ such that $(x_1^2, \ldots, x_s^2) \in K$. Combining the crude estimate $\nu(n_i) \ll X/b_1$ with Lemma 7.1, we deduce that

$$\sum_{\mathbf{x} \in K} \prod_{i=1}^s \nu(x_i) \ll_\varepsilon (X b_1^{-1})^{2s-\frac{5}{2}+\varepsilon} \ll_\varepsilon (NW)^{s-\frac{5}{4}+\frac{\varepsilon}{2}} \ll_\varepsilon N^{s-\frac{5}{4}+\varepsilon},$$

since $W^{s-\frac{5}{4}+\frac{\varepsilon}{2}} \ll_\varepsilon N^{\frac{\varepsilon}{2}}$ by (4.2). This completes the proof. $\square$

8. Proof of Theorem 1.1

Let $A \subset [X]$ with $|A| = \delta X$ and such that A contains only K-trivial solutions to (1.4). We may assume that $\delta \geqslant C(\log X)^{-1}$ where C is the absolute constant appearing in Lemma 4.1, otherwise the proof of Theorem 1.1 is complete. Under this assumption, Lemma 4.1 yields the existence of a w-smooth number $b_1 \leqslant X^{1/2}$ and $b_2 \in [W]$ with $(b_2, W) = 1$ such that

$$\sum_n 1_{A_{\mathbf{b}}}(n) \nu_{\mathbf{b}}(n) \gg \delta^2 N_{\mathbf{b}},$$

where $A_{\mathbf{b}} = \{n \in \mathbb{Z} : b_1^2(Wn - b_2) = x^2 \text{ for some } x \in A\}$ and $N_{\mathbf{b}}$ is given by (4.4). Given such a choice of $\mathbf{b}$, we have established that

- $\nu_{\mathbf{b}}$ satisfies a restriction estimate at any exponent $p > 4$ (Prop. 6.1);
- $\nu_{\mathbf{b}}$ has Fourier decay of level $\theta \ll 1/\sqrt{w}$ (Lemma 5.5);
- $\nu_{\mathbf{b}}$ saves $1/5$ on the K-trivial solutions (Cor. 7.2).

Incorporating these facts into Proposition 2.8, we deduce that

$$\sum_n 1_{A_{\mathbf{b}}}(n)\nu_{\mathbf{b}}(n) \ll_{\mathbf{c},k,\varepsilon} \frac{N_{\mathbf{b}}}{\min\{\log\log w,\ \log N_{\mathbf{b}}\}^{s-2-\varepsilon}}.$$

Recalling (4.1), (4.2) and (4.4), we finally deduce that

$$\delta \ll_{\mathbf{c},k,\varepsilon} (\log\log\log X)^{-(\frac{s}{2}-1-\varepsilon)}.$$

This completes the proof of Theorem 1.1.

REFERENCES

[Blo12] T. F. Bloom, Translation invariant equations and the method of Sanders. *Bull. Lond. Math. Soc.* **44** (2012), 1050–1067.

[Bou89] J. Bourgain, On $\Lambda(p)$-subsets of squares. *Israel J. Math.* **67** (1989), 291–311.

[BP14] T. D. Browning and S. M. Prendiville, Improvements in Birch's theorem on forms in many variables. *J. reine angew. Math.*, to appear.

[BDLW10] J. Brüdern, R. Dietmann, J. Y. Liu, T. D. Wooley, A Birch–Goldbach theorem. *Archiv der Mathematik* **94** (2010), 53–58.

[EG80] P. Erdős and R. Graham, Old and new problems and results in combinatorial number theory. *Monographies de L'Enseignement Mathématique* vol. **28**, Université de Genève, Geneva, 1980.

[FrHo14] N. Frantzikinakis and B. Host, Higher order Fourier analysis of multiplicative functions and applications. *Submitted*, 2014. (`arXiv:1403.0945`)

[Gow01] W. T. Gowers, A new proof of Szemerédi's Theorem. *GAFA* **11** (2001), 465–588.

[Gra08] R. Graham, Old and new problems in Ramsey theory. *Horizons of combinatorics*, 105118, Bolyai Soc. Math. Stud. **17**, Springer, Berlin, 2008.

[GRS90] R. Graham, B. Rothschild and J. H. Spencer, *Ramsey Theory*. Wiley, 1990.

[Gre05] B. J. Green, Roth's theorem in the primes. *Ann. of Math.* **161** (2005), 1609–1636.

[Har13] J. Harris, *Algebraic geometry: a first course*. Springer, 2013.

[Hen15] K. Henriot, Logarithmic bounds for translation-invariant equations in squares. *Int. Math. Research Not.* **2015** (2015), 12540–12562.

[Hen15] _____, Additive equations in dense variables, and truncated restriction estimates. *Submitted*, 2015. (`arXiv:1508.05923`)

[Ke14] E. Keil, On a diagonal quadric in dense variables. *Glasg. Math. J.* **56** (2014), 601–628.

[Keil15] _____, Some refinements for translation invariant quadratic forms in dense sets. *Submitted*, 2015. (`arXiv:1408.1535`)

[Pil95] J. Pila, Density of integral and rational points on varieties. *Astérisque* **228** (1995), 183–187.

[Pre15] S. M. Prendiville, Four variants of the Fourier-analytic transference principle. *Submitted*, 2015 (`arXiv:1509.09200`)

[Rot53] K. F. Roth, On certain sets of integers. *J. London Math. Soc.* **28** (1953), 104–109.

[Rot54] _____, On certain sets of integers. II. *J. London Math. Soc.* **29** (1954), 20–26.

[San11] T. Sanders, On Roth's theorem on progressions. *Ann. of Math.* **174** (2011), 619– 636.

[SchSi] T. Schoen and O. Sisask, Roth's theorem for four variables and additive structures in sums of sparse sets. *Forum of Mathematics, Sigma*, to appear.

[Smi09]   M. L. Smith, On solution-free sets for simultaneous quadratic and linear equations. *J. Lond. Math. Soc.* **79** (2009), 273–293.

[Tao12]   T. Tao, *Higher order Fourier analysis.* Graduate studies in mathematics, vol. **142**, American Mathematical Society, 2012.

[Vau86]   R. C. Vaughan, On Waring's problem for cubes. *J. reine angew. Math.* **365** (1986), 122–170.

[Vau97]   ———, *The Hardy-Littlewood method*, 2nd ed., Cambridge Tracts in Mathematics, vol. **125**, Cambridge University Press, Cambridge, 1997.

School of Mathematics, University of Bristol, Bristol, BS8 1TW
E-mail address: t.d.browning@bristol.ac.uk

Mathematical Institute, University of Oxford, Oxford, OX2 6GG
E-mail address: sean.prendiville@maths.ox.ac.uk