COUNTING MONOCHROMATIC SOLUTIONS TO DIAGONAL DIOPHANTINE EQUATIONS

SEAN PRENDIVILLE

Abstract. Given a finite colouring of the positive integers, we count monochromatic solutions to a variety of Diophantine equations, each of which can be written by setting a diagonal quadratic form equal to a linear form. As a consequence, we determine an algebraic criterion for when such equations are partition regular. Our methods involve discrete harmonic analysis and require a number of 'mixed' restriction estimates, which may be of independent interest.

Contents

1. Introduction

A substantial portion of Ramsey theory concerns properties which persist under finite partitions, such as the property of solving a pre-determined Diophantine equation.

Definition 1.1 (Partition regular)**.** Given a polynomial $P \in \mathbb{Z}[x_1, \dots, x_s]$ we say that the equation $P(x_1, \dots, x_s) = 0$ is *partition regular* if for any finite partition of the positive integers $\mathbb{N} = C_1 \cup \cdots \cup C_r$ there exists C_j and infinitely many $(x_1, \dots, x_s) \in C_j^s$ such that $P(x_1, \dots, x_s) = 0$. One may think of a partition into r parts as a colouring with r colours, in which case we call $(x_1, \dots, x_s)$ a *monochromatic solution*.

Rado [Rad33] completely characterised which linear forms $P = a_1x_1 + \cdots + a_sx_s$ are partition regular: it is both necessary and sufficient that there exists $I \neq \emptyset$ such that $\sum_{i\in I} a_i = 0$. There are few results for non-linear Diophantine equations. For instance, it is a longstanding problem of Erdős and Graham [Gra07, Gra08] to determine whether the Pythagorean equation is partition regular.

Perhaps the first truly non-linear result is due to Bergelson [Ber96], asserting partition regularity of the equation $x - y = z^2$. We prove a counting version of Bergelson's theorem, which is prototypical of the results of this paper.

Theorem 1.2. *For any r-colouring $C_1 \cup \cdots \cup C_r = \{1, 2, \dots, N\}$ there exists a colour class C_j such that*[1]

$$\sum_{x-y=z^2} 1_{C_j}(x)1_{C_j}(y)1_{C_j}(z) \gg_r N^{3/2^r}(1 - o_r(1)). \tag{1.1}$$

The lower bound in (1.1) is far from the total number of solutions to the equation $x - y = z^2$ in the interval $[N] = \{1, 2, \dots, N\}$, which is of order $N^{3/2}$. However, the order of magnitude in (1.1) is optimal, as can be seen from the colouring

$$C_1 := (N^{1/2}, N], \quad \dots \quad , \quad C_{r-1} := (N^{1/2^{r-1}}, N^{1/2^{r-2}}], \quad C_r := [N^{1/2^{r-1}}]. \tag{1.2}$$

Proposition 1.3. *There exists an r-colouring of $[N]$ with at most $O(N^{3/2^r})$ monochromatic solutions to the equation $x - y = z^2$.*

The argument underlying Theorem 1.2 utilises the Fourier-analytic regularity lemma of Green [Gre05], which has a convenient formulation due to Green and Tao [GT10]. The robustness of the regularity lemma allows us to prove the following generalisation of Theorem 1.2.

Theorem 1.4 (Linear counting theorem)**.** *Let $a_1, \dots, a_s, b_1, \dots, b_t \in \mathbb{Z} \setminus \{0\}$ and suppose that there exists $I \neq \emptyset$ such that $\sum_{i \in I} a_i = 0$. For any r-colouring $C_1 \cup \cdots \cup C_r = [N]$ there exists a colour class C_j such that*

$$\sum_{a_1x_1+\cdots+a_sx_s=b_1y_1^2+\cdots+b_ty_t^2} 1_{C_j}(x_1)\cdots 1_{C_j}(x_s)1_{C_j}(y_1)\cdots 1_{C_j}(y_t) \gg_r N^{(|I|+s+t-2)/2^r}(1 - o_r(1)). \tag{1.3}$$

(Here we have suppressed the dependence of implicit constants on a_i and b_j.)

Turning to equations without linear terms, Chow, Lindqvist and the author [CLP] have classified partition regular diagonal equations

$$a_1x_1^k + \cdots + a_sx_s^k = 0. \tag{1.4}$$

This classification is subject to the caveat[2] that the number of variables s is sufficiently large in terms of the degree k, but is otherwise identical to Rado's criterion: it is both necessary and sufficient that there exists $I \neq \emptyset$ such that $\sum_{i \in I} a_i = 0$. For squares ($k = 2$) we require $s \geqslant 5$ at present. The methods of [CLP] do not yield a lower bound on the number of monochromatic solutions to (1.4), though it was conjectured [CLP, §3.1] that such a result should be true. The original motivation for the present paper is to settle this conjecture affirmatively.

Theorem 1.5. *Let $a_1, \dots, a_s \in \mathbb{Z} \setminus \{0\}$ with $s \geqslant 5$. Suppose that there exists $I \neq \emptyset$ such that $\sum_{i \in I} a_i = 0$. Then for any for any r-colouring $[N] = C_1 \cup \cdots \cup C_r$ there exists $C \in \{C_1, \dots, C_r\}$ such that*

$$\sum_{a_1x_1^2+\cdots+a_sx_s^2=0} 1_C(x_1)\cdots 1_C(x_s) \gg_r N^{s-2}(1 - o_r(1)).$$

---

[1] For our conventions regarding asymptotic notation, see §1.5.

[2] The Fermat cubic illustrates that some such caveat is necessary.

(Here we have suppressed the dependence of implicit constants on the coefficients a_i.)

A standard application of the circle method (see [Vau97]) shows that the bound in Corollary 1.5 is optimal.

Proposition 1.6. *For any $a_1, \ldots, a_s \in \mathbb{Z} \setminus \{0\}$ with $s \geqslant 5$ we have the upper bound*

$$\sum_{a_1x_1^2+\cdots+a_sx_s^2=0} 1_{[N]}(x_1)\cdots 1_{[N]}(x_s) \ll N^{s-2}.$$

Remark (Higher degree diagonal equations)**.** The methods of this paper also yield a counting result for higher degree diagonal equations of the form (1.4). We restrict our attention to squares for a simpler exposition.

One consequence of celebrated work of Moreira [Mor17] is partition regularity of the equation

$$a_1x_1^2 + \cdots + a_sx_s^2 = x_0, \tag{1.5}$$

under the assumption that

$$a_1 + \cdots + a_s = 0. \tag{1.6}$$

Moreira's methods are inductive, and locate a monochromatic solution arising from a special two-parameter subvariety. To obtain a counting result for (1.5) by modifying these methods seems implausible. Using an alternative approach, we obtain a counting result for (1.5) and in addition are able to substantially relax the assumption (1.6) on the coefficients. The price we pay for this strengthening is that we must assume the quadratic form has sufficiently many variables.

Theorem 1.7 (Quadratic counting theorem)**.** *Let $a_1, \ldots, a_s \in \mathbb{Z}\setminus\{0\}$ and $b_1, \ldots, b_t \in \mathbb{Z} \setminus \{0\}$ with $s \geqslant 3$ and $s + t \geqslant 5$. Suppose that there exists $I \neq \emptyset$ such that $\sum_{i\in I} a_i = 0$. Then for any r-colouring $[N] = C_1 \cup \cdots \cup C_r$ there exists $C \in \{C_1, \ldots, C_r\}$ such that*

$$\sum_{a_1x_1^2+\cdots+a_sx_s^2=b_1y_1+\cdots+b_ty_t} \prod_i 1_C(x_i) \prod_j 1_C(y_j) \gg_r N^{s+t-2}(1 - o_r(1)).$$

(Here we have suppressed the dependence of implicit constants on a_i and b_j.)

All equations so far considered have the form

$$a_1x_1^2 + \cdots + a_sx_s^2 = b_1y_1 + \cdots + b_ty_t, \tag{1.7}$$

where the a_i and b_j are non-zero integers. Another equation of this type, $x+y = z^2$, has received attention from Green–Lindqvist [GL19] and Pach [Pac18]. They demonstrate that $x + y = z^2$ has infinitely many monochromatic solutions in any 2-colouring, but that there is a 3-colouring with no monochromatic solutions beyond $(x, y, z) = (2, 2, 2)$. With this in mind, it is natural to ask the following.

Question 1.8. *Let $a_1, \ldots, a_s, b_1, \ldots, b_t \in \mathbb{Z} \setminus \{0\}$ with $s, t \geqslant 1$. When is the equation (1.7) partition regular?*

Ideally we would like an algebraic characterisation comparable to that of [Rad33] and [CLP], a criterion which can be easily checked by a computer. A necessary condition is provided in work of Di Nasso and Luperi Baglini [DNLB18, Theorem 3.10].

Proposition 1.9 (Di Nasso and Luperi Baglini). *If the equation* (1.7) *is partition regular, then there exists $I \neq \emptyset$ such that either $\sum_{i\in I} a_i = 0$ or $\sum_{i\in I} b_i = 0$.*

We are able to show that this condition is sufficient in all but one case.

Theorem 1.10 (Linear–quadratic partition regularity). *Let $a_1, \dots, a_s, b_1, \dots, b_t \in \mathbb{Z} \setminus \{0\}$ with $s, t \geqslant 1$. Suppose that* (1.7) *does not take the form*

$$a(x_1^2 - x_2^2) = by^2 + cz \tag{1.8}$$

for some non-zero integers a, b, c. Then (1.7) *is partition regular if and only if there exists $I \neq \emptyset$ such that $\sum_{i\in I} a_i = 0$ or $\sum_{i\in I} b_i = 0$.*

This almost resolves [DNLB18, Open Problem 1] when the question is restricted to the family of Diophantine equations given by (1.7). Our lack of knowledge regarding (1.8) is an artefact of our methods. We believe that Di Nasso and Luperi Baglini's criterion is the correct characterisation.

Conjecture 1.11. *For any non-zero integers a, b, c, the equation* (1.8) *is partition regular.*

As evidence towards this conjecture, we prove that a special case of (1.8) is partition regular conditional on the following notorious problem of Hindman.

Conjecture 1.12 (Hindman). *In any finite colouring of $\mathbb{N}$ there is a monochromatic configuration of the form $\{x, y, x + y, xy\}$.*

Theorem 1.13. *If Hindman's conjecture is true, then the equation*

$$x_1^2 - x_2^2 = y^2 + z \tag{1.9}$$

is partition regular.

1.1. **Mixed restriction estimates.** The main tools used in proving our results are the Hardy–Littlewood circle method, the abelian arithmetic regularity lemma and the Fourier analytic transference principle. All three of these tools are part of discrete harmonic analysis, and key to their success are so-called *discrete restriction estimates*[3].

Colourings such as (1.2), when combined with the inhomogeneity of the equation (1.7), force us to count solutions to equations in certain 'skewed' regions, where some variables are constrained to much smaller intervals than is typical in the circle method. This necessitates the development of some novel 'mixed' restriction estimates (see Lemma 6.1), such as the following.

Theorem 1.14 (Mixed restriction). *Let W be a positive integer and $p > 2$. Then either $N \ll_p W^{O_p(1)}$ or, for any $f, g : \mathbb{Z} \to \mathbb{C}$ with $|f|, |g| \leqslant 1_{[N]}$ we have*

$$\int_{\mathbb{T}} \left| \sum_{N/2 < x \leqslant N} f(x) e(W\alpha x^2) \sum_{N/2 < y \leqslant N} g(y) e(\alpha y) \right|^p \mathrm{d}\alpha \ll_p N^{2p-2} W^{-1}. \tag{1.10}$$

We note that

$$\int_{\mathbb{T}} \left| \sum_{N/2 < x \leqslant N} f(x) e(W\alpha x^2) \right|^{2p} \mathrm{d}\alpha \ll_p N^{2p-2}$$

[3]See the introduction to [HH18] for motivation and history.

and

$$\int_{\mathbb{T}} \left| \sum_{N/2<y\leqslant N} g(y)e(\alpha y) \right|^{2p} \mathrm{d}\alpha \ll_p N^{2p-1}.$$

Hence the obvious application of the Cauchy–Schwarz inequality does not deliver a bound as strong as (1.10).

In addition to (1.10), we require three further mixed restriction estimates, and to prove all four simultaneously we abstract an approach of Bourgain [Bou89]. Hence, in §4 we prove a general restriction estimate for exponential sums obeying certain hypotheses and in §5 we verify that each of our four mixed exponential sums satisfy these hypotheses.

1.2. **The utility of counting results.** In the study of partition regularity it is often desirable to delineate between 'trivial' and 'non-trivial' solutions to an equation, as some equations possess monochromatic solutions for uninteresting reasons. For instance $x + y = z^2$ has the solution $(2, 2, 2)$, whilst $x + y = 2z$ is always solved by the diagonal (x, x, x). One commonly encountered choice of non-triviality is a solution in which all variables are distinct, but the precise notion may depend on the application. A counting result allows one to ensure the existence of monochromatic solutions avoiding *any* sparse subset of solutions. This implies that there are monochromatic solutions of 'generic type', i.e. not lying on a proper Zariski closed subset. For if all monochromatic solutions took this form then counting arguments would likely give a power saving in the number of monochromatic solutions when compared with the total number of solutions.

Frankl, Graham and Rödl [FGR88] pioneered the counting of monochromatic solutions to systems of *linear* equations, obtaining lower bounds of the correct order of magnitude for all such partition regular systems. The non-linear theory is much less developed, mainly due to our lack of knowledge regarding when such equations are partition regular. The author hopes this paper encourages the development of further non-linear counting results.

1.3. **Organisation of this paper.** We sketch some of the ideas behind our methods in §2. In §3 we use the arithmetic regularity lemma to prove that dense sets of integers contain certain polynomial configurations, from which all of our counting results are ultimately derived. We derive Theorem 1.4 from the results of §3 in §8.1.

We devote §4–7 to modifying the results of §3 to apply to dense sets of *squares*, instead of just dense sets of integers. In §4 we generalise an approach of Bourgain [Bou89] to prove a general restriction estimate for exponential sums obeying certain hypotheses and in §5 we verify these hypotheses for the exponential sums of relevance. In §6 we use these restriction estimates to show how the Fourier transform of a set completely determines the number of solutions it contains to the equations we are interested in.

All of our counting results are derived from density results in §8. Finally in §9 we adapt an argument of Moreira to establish partition regularity of equations of the form (1.7) which are not covered by our counting theorems. This allows us to combine all previous results to deduce our partition regularity criteria (Theorem 1.10).

1.4. **Acknowledgements.** The author thanks Sofia Lindqvist for the arguments of §9, and Sam Chow for the idea of using Lemma 6.3.

1.5. **Notation.**

Standard conventions. We use $\mathbb{N}$ to denote the positive integers. For a real number $X \geqslant 1$, write $[X] = \{1, 2, \dots, \lfloor X \rfloor\}$. A complex-valued function is said to be *1-bounded* if the modulus of the function does not exceed 1.

We use counting measure on $\mathbb{Z}$, so that for $f, g : \mathbb{Z} \to \mathbb{C}$, we have

$$\|f\|_{L^p} := \Big(\sum_x |f(x)|^p\Big)^{\frac{1}{p}}, \ \langle f, g\rangle := \sum_x f(x)\overline{g(x)}, \text{ and } (f*g)(x) := \sum_y f(y)g(x-y).$$

Any sum of the form $\sum_x$ is to be interpreted as a sum over $\mathbb{Z}$. The *support* of f is the set $\mathrm{supp}(f) := \{x \in \mathbb{Z} : f(x) \neq 0\}$. We write $\|f\|_\infty$ for $\sup_x |f(x)|$.

For a finite set S and function $f : S \to \mathbb{C}$, denote the average of f over S by

$$\mathbb{E}_{s\in S} f(s) := \frac{1}{|S|}\sum_{s\in S} f(s).$$

We use Haar probability measure on $\mathbb{T} := \mathbb{R}/\mathbb{Z}$, so that for integrable $F, G : \mathbb{T} \to \mathbb{C}$, we have

$$\|F\|_{L^p} := \Big(\int_{\mathbb{T}} |F(\alpha)|^p d\alpha\Big)^{\frac{1}{p}} = \Big(\int_0^1 |F(\alpha)|^p d\alpha\Big)^{\frac{1}{p}},$$

$$\langle F, G\rangle := \int_{\mathbb{T}} F(\alpha)\overline{G(\alpha)}\mathrm{d}\alpha, \quad \text{and} \quad (F * G)(\alpha) := \int_{\mathbb{T}} F(\alpha - \beta)G(\beta)\mathrm{d}\beta.$$

We write $\|\alpha\|_{\mathbb{T}}$ for the distance from $\alpha \in \mathbb{R}$ to the nearest integer $\min_{n\in\mathbb{Z}} |\alpha - n|$. This remains well-defined on $\mathbb{T}$.

Definition 1.15 (Fourier transform)**.** For $f : \mathbb{Z}^d \to \mathbb{C}$ with finite support define $\hat{f} : \mathbb{T}^d \to \mathbb{C}$ by

$$\hat{f}(\alpha) := \sum_{n\in\mathbb{Z}^d} f(n)e(\alpha \cdot n).$$

Here $e(\beta)$ stands for $e^{2\pi i\beta}$. We sometimes write $e_q(a)$ for $e(a/q)$.

Given integrable $F : \mathbb{T}^d \to \mathbb{C}$ write

$$\hat{F}(n) := \int_{\mathbb{T}^d} F(\alpha)e(-\alpha \cdot n)\mathrm{d}\alpha.$$

Definition 1.16 (Smooth/rough numbers)**.** We say that an integer n is w-smooth if all of its prime divisors are at most w. We say n is w-rough if all of its prime divisors are at least w.

Asymptotic notation. For a complex-valued function f and positive-valued function g, write $f \ll g$ or $f = O(g)$ if there exists a constant C such that $|f(x)| \leq Cg(x)$ for all x. We write $f = \Omega(g)$ if $f \gg g$. The notation $f \asymp g$ means that $f \ll g$ and $f \gg g$.

We write $f = o(g)$ if for any $\varepsilon > 0$ there exists $X \in \mathbb{R}$ such that for all $x \geqslant X$ we have $|f(x)| \leqslant \varepsilon g(x)$.

Local conventions. The following are idiosyncratic to this paper, and may not be adhered to elsewhere.

Definition 1.17 (Quadratic Fourier transform)**.** Given $f : \mathbb{Z} \to \mathbb{C}$ with finite support, define the *quadratic Fourier transform* by

$$\tilde{f}(\alpha) := \sum_x f(x)e(\alpha x^2).$$

Definition 1.18 (Non-singular linear form)**.** Let $c_1, \dots, c_s \in \mathbb{Z}$. We call a polynomial of the form

$$L(x_1, \dots, x_s) = c_1x_1 + \dots c_sx_s$$

a *linear form*. We say the linear form is *non-singular* if $c_i \neq 0$ for all i. If $x = (x_1, \dots, x_s) \in \mathbb{Z}^s$, then it will be convenient to use the shorthand

$$L(x^2) := L(x_1^2, \dots, x_s^2).$$

Remark (Dependence of implicit constants on linear forms)**.** A number of results in the remainder of the paper concern three non-singular linear forms L_1, L_2, L_3. Throughout we suppress dependence of implicit constants on the number of variables and the coefficients of the L_i. One may think of all data associated to the L_i as being $O(1)$.

2. A sketch of our methods

As with the author's previous two papers on partition regularity [CLP, CP], we first exhibit the method underlying our results with a proof of Schur's theorem.

2.1. The regularity approach to Schur's theorem.

Theorem 2.1 (Schur)**.** *For any r-colouring $C_1 \cup \dots \cup C_r = \{1, 2, \dots, N\}$ there exists a colour class C_j and $x, y, z \in C_j$ such that $x + y = z$.*

We sketch a proof of this using the Fourier-analytic regularity lemma (Lemma 3.3) originating in [Gre05]. The take-away of the regularity lemma is that we can find a Bohr set

$$B := \{x \in [N] : \|\alpha_i x\|_{\mathbb{T}} \leqslant \eta \text{ for } i = 1, \dots, d\} \tag{2.1}$$

such that each colour class C_j is approximately invariant under shifts by B, so that

$$1_{C_j}(x + y) \approx 1_{C_j}(x). \tag{2.2}$$

We have been deliberately vague about the nature of the approximation in (2.2). There is an important trade-off to keep in mind: the closer one wishes the approximation (2.2), the smaller the resulting Bohr set (2.1). The nature of the approximation (2.2) allows us to conclude that for any colour classes C_i and C_j we have

$$\sum_{x\in[N]} \sum_{y\in B} 1_{C_i}(x)1_{C_j}(y)1_{C_i}(x + y) \approx \sum_{x\in[N]} 1_{C_i}(x)^2 \sum_{y\in B} 1_{C_j}(y). \tag{2.3}$$

Using Cauchy–Schwarz the right-hand side of (2.3) is at least

$$N^{-1}|C_i|^2|C_j \cap B|.$$

By the pigeon-hole principle there exists a colour class C_j with $|C_j \cap B| \geqslant |B|/r$ and hence for all i we have

$$\sum_{x\in[N]} 1_{C_i}(x)^2 \sum_{y\in B} 1_{C_j}(y) \geqslant \frac{|C_i|^2|B|}{rN}. \tag{2.4}$$

The obvious strategy is to now take $i := j$ in (2.3) and (2.4), to yield

$$\sum_{x\in[N]}\sum_{y\in B} 1_{C_j}(x)1_{C_j}(y)1_{C_j}(x+y) \approx \sum_{x\in[N]} 1_{C_j}(x)^2 \sum_{y\in B} 1_{C_j}(y) \geqslant \frac{|C_j|^2|B|}{rN}. \tag{2.5}$$

The drawback with this approach is that the error term in (2.3) is of the form $\varepsilon N|B|$. Hence in order to use (2.5) to deduce the existence of a monochromatic solution to $x + y = z$, we need the lower bound in (2.5) to be of order $N|B|$. This may not happen: imagine the situation in which the colour class C_j is equal to the Bohr set B (for the purposes of this sketch, $|B|$ should be thought of as $o(N)$). The problem we have encountered is that the colour class C_j which is good for the regularity lemma (as it has large intersection with the Bohr set B) may not be a dense colour class (which we need for the lower bound in (2.4) to be useful).

Our solution to this problem is twofold. By adapting the regularity argument outlined above, we first prove an asymmetric version of Schur's theorem.

Theorem 2.2 (Asymmetric Schur)**.** *Let $\delta > 0$ and $A_1, \dots, A_s \subset [N]$ each with $|A_i| \geqslant \delta N$. Then for any colouring $[N] = C_1 \cup \dots \cup C_r$ there exists a colour class C_j such that for any A_i we have*

$$\sum_{x+y=z} 1_{A_i}(x)1_{C_j}(y)1_{A_i}(z) \gg_{\delta,r,s} N^2(1 - o_{\delta,r,s}(1)).$$

Next, in order to deduce Schur's theorem from this asymmetric version, we 'cleave' colour classes into those which are dense and those which are sparse. Fix a growth function $\mathcal{F}$. A combinatorial argument allows us to find a density $1/M$ with $M = O_{r,\mathcal{F}}(1)$ such that for every colour class C_i one of the following holds:

- either C_i is $1/M$ dense, in that $|C_i| \geqslant N/M$;
- or C_i is $1/\mathcal{F}(M)$ sparse, in that $|C_i| < N/\mathcal{F}(M)$.

We have 'cleaved', in that we have found a threshold parameter M such that each colour class is either extremely dense in terms of M, or extremely sparse in terms of M, there are no intermediate colour classes.

Having cleaved, we apply our asymmetric Schur theorem, taking the sets A_i to be those colour classes which are $1/M$ dense. This yields a colour class C_j such that for any $1/M$ dense colour class C_i we have

$$\sum_{x+y=z} 1_{C_i}(x)1_{C_j}(y)1_{C_i}(z) \gg_{M,r} N^2.$$

We would like to take $i = j$ in the above, but we can only do this if C_j is $1/M$ dense. Let us see why this is so. A counting argument shows that

$$|C_j|N \geqslant \sum_{x+y=z} 1_{C_i}(x)1_{C_j}(y)1_{C_i}(z).$$

Hence

$$|C_j| \gg_{M,r} N. \tag{2.6}$$

Provided we have chosen our growth function $\mathcal{F}$ so that the implicit constant in (2.6) is larger than $1/\mathcal{F}(M)$, we deduce that C_j is not $1/\mathcal{F}(M)$ sparse, hence it must be $1/M$ dense, by cleaving.

2.2. **Adapting this to Bergelson's theorem.** Using quadratic Bohr sets in place of Bohr sets, it is relatively simple to adapt the regularity argument underlying Theorem 2.2 to prove the following.

Theorem 2.3 (Asymmetric Bergelson)**.** *Let $\delta > 0$ and $A_1, \dots, A_s \subset [N]$ each with $|A_i| \geqslant \delta N$. Then for any colouring $[N^{1/2}] = C_1 \cup \cdots \cup C_r$ there exists a colour class C_j such that for any A_i we have*

$$\sum_{x-y=z^2} 1_{A_i}(x) 1_{A_i}(y) 1_{C_j}(z) \gg_{\delta,r,s} N^{3/2}(1 - o_{\delta,r,s}(1)). \tag{2.7}$$

The problem now is how to cleave? Notice that (2.7) counts $z \in C_j \cap [N^{1/2}]$, and the density/sparsity of C_j on the interval $[N^{1/2}]$ may be independent of the density/sparsity of C_j on $[N]$ (see the colouring (1.2)). To overcome this we find $M = O_{r,\mathcal{F}}(1)$ and scales $X_1, \dots, X_r \geqslant X^2$ such that C_i is $1/M$ dense on $[X_i]$ if it is C_i is $1/\mathcal{F}(M)$ dense on $[X]$. Averaging, there is a translate $a_i + [X^2]$ such that if C_i is $1/\mathcal{F}(M)$ dense on $[X]$ then C_i is $1/M$ dense on $a_i + [X^2]$. We then take

$$A_i := \left\{x \in [X^2] : a_i + x \in C_i\right\}$$

in Theorem 2.3, and apply a similar argument to that given for Schur's theorem.

We note that key to the success of this strategy is the translation invariance of the linear form $x - y$, in that

$$(x + a) - (y + a) = z \quad \text{iff} \quad x - y = z.$$

This is a property enjoyed by any linear form whose coefficients sum to zero. Unfortunately, the same is not true of a quadratic form whose coefficients sum to zero. Overcoming this is the subject of the next subsection

2.3. **The W-trick for squares and linearisation.** To prove Theorem 1.7, when the coefficients of the quadratic form satisfy Rado's criterion, we combine our 'cleaving' strategy with the following asymmetric density-colouring result.

Theorem 2.4 (Quadratic density–colouring result)**.** *Let $\delta > 0$ and let r be a positive integer. For any sets of integers $A_1, \dots, A_s \subset [N]$ each satisfying $|A_i| \geqslant \delta N$ and for any r-colouring $B_1 \cup \cdots \cup B_r = [N]$ there exists $B \in \{B_1, \dots, B_r\}$ such that for all $A \in \{A_1, \dots, A_r\}$ we have*

$$\sum_{x_1^2 - x_2^2 = y^2 + z_1 + z_2} 1_A(x_1) 1_A(x_2) 1_B(y) 1_B(z_1) 1_B(z_2) \gg_{\delta,r,s} N^3(1 - o_{\delta,r,s}(1)).$$

This is a representative special case of Theorem 7.1, which we have stated for simplicity. Using a Fourier analytic transference principle (see [Pre17]), we deduce Theorem 2.4 from a *linear* density–colouring result, where we have removed the squares from the x_i variables.

Lemma 2.5 (Linear density–colouring result)**.** *Let $\delta > 0$ and let r be a positive integer. For any sets of integers $A_1, \dots, A_s \subset [N^2]$ each satisfying $|A_i| \geqslant \delta N^2$ and* *for any r-colouring $B_1 \cup \dots \cup B_r = [N]$ there exists $B \in \{B_1, \dots, B_r\}$ such that for all $A \in \{A_1, \dots, A_r\}$ we have*

$$\sum_{x_1 - x_2 = y^2 + z_1 + z_2} 1_A(x_1) 1_A(x_2) 1_B(y) 1_B(z_1) 1_B(z_2) \gg_{\delta,r,s} N^5(1 - o_{\delta,r,s}(1)).$$

This is superficially similar to the strategy employed in [CLP], but without the presence of the strongly structured 'homogeneous sets' (more properly termed multiplicatively syndetic sets, see [Cha]). The lack of such structure presents additional obstacles too technical to discuss here. We refer the interested reader to §7.

3. A linear density result

The aim of this section is to count solutions to equations of the form (1.7) when certain linear variables are constrained to dense sets, and the remaining variables are constrained to a colouring. We eventually use this density result to derive both our linear counting result (Theorem 1.4) and our quadratic counting result (Theorem 1.7). Before stating this we remind the reader of our conventions (Definition 1.18) regarding linear forms.

Theorem 3.1 (Linear density result)**.** *Let L_1, L_2, L_3 denote non-singular linear forms, each in s_i variables with $s_1 \geqslant 2$ and $s_1 + s_2 \geqslant 3$ (we allow for $s_2 = 0$ or $s_3 = 0$). Suppose that $L_1(1, \dots, 1) = 0$. For any $\delta > 0$ and positive integer r, there exists $\eta \gg_{r,\delta} 1$ such that for any positive integers W and N, either $N \ll_{\delta,r,W} 1$ or the following holds. Suppose that $W = 1$ or $s_3 > 0$. Then for any sets $A_1, \dots, A_r \subset [N]$ with $|A_i| \geqslant \delta N$ for all i, and any r-colouring $C_1 \cup \dots \cup C_r = [\eta(N/W)^{1/2}, (N/W)^{1/2}]$, there exists C_j such that for all A_i we have*

$$\sum_{L_1(x) = W L_2(y^2) + L_3(z)} 1_{A_i}(x_1) \cdots 1_{A_i}(x_{s_1}) 1_{C_j}(y_1) \cdots 1_{C_j}(y_{s_2}) 1_{C_j}(z_1) \cdots 1_{C_j}(z_{s_3})$$
$$\geqslant \eta N^{s_1 + \frac{1}{2}(s_2+s_3) - 1} W^{-\frac{1}{2}(s_2+s_3)}.$$

We prove Theorem 3.1 using Fourier analysis and the arithmetic regularity lemma. To state the regularity lemma we require the following.

Definition 3.2 (Lipschitz constant on $\mathbb{T}^d$)**.** We say that $F : \mathbb{T}^d \to \mathbb{C}$ is M-Lipschitz if for any $\alpha, \beta \in \mathbb{T}^d$ we have

$$|F(\alpha) - F(\beta)| \leqslant M \min_{1 \leqslant i \leqslant d} \|\alpha_i - \beta_i\|_{\mathbb{T}}.$$

Lemma 3.3 (Arithmetic regularity)**.** *Let $\varepsilon > 0$ and let $\mathcal{F} : \mathbb{N} \to \mathbb{N}$. For any functions $f_i : [N] \to [0,1]$ with $i = 1, \dots, r$ there exists $M \ll_{\varepsilon,\mathcal{F},r} 1$ and decompositions*

$$f_i = f_i^{\mathrm{str}} + f_i^{\mathrm{sml}} + f_i^{\mathrm{unf}} \qquad (1 \leqslant i \leqslant r),$$

with the following properties.

(Str). *There exist $d \leqslant M$ and $\theta \in \mathbb{T}^d$, such that for each i there is an M-Lipschitz function $F_i : \mathbb{T}^d \to [0,1]$ with $f_i^{\mathrm{str}}(x) = F_i(\theta x)$ for all $x \in [N]$.*

(Sml). *$f_i^{\mathrm{sml}} : [N] \to [-1,1]$ with*

$$\sum_{x \in [N]} f_i^{\mathrm{sml}}(x) = 0, \quad \left\|f_i^{\mathrm{sml}}\right\|_1 \leqslant \varepsilon \left\|1_{[N]}\right\|_1 \quad \text{and} \quad f_i^{\mathrm{str}} + f_i^{\mathrm{sml}} \geqslant 0.$$

(Unf). $f_i^{\mathrm{unf}} : [N] \to [-1,1]$ *with*

$$\sum_{x\in[N]} f_i^{\mathrm{unf}}(x) = 0 \quad \textit{and} \quad \left\|\hat{f}_i^{\mathrm{unf}}\right\|_\infty \leqslant \left\|\hat{1}_{[N]}\right\|_\infty / \mathcal{F}(M).$$

Proof. This can be proved by following the arguments of [GT10] or [Tao12]. □

To prove Theorem 3.1 we apply the regularity lemma to decompose each 1_{A_i} into a structured, small and uniform part. We eventually show that the small and uniform parts do not contribute substantially to our count of solutions. It is therefore necessary to show that the structured part has a large contribution.

Lemma 3.4 (Structured counting lemma)**.** *Let L_1, L_2, L_3 be linear forms, each in s_i variables. Given an M-Lipschitz function $F : \mathbb{T}^d \to [0,1]$ and $\theta \in \mathbb{T}^d$, define $f : \mathbb{Z} \to [0,1]$ by*

$$f(x) := \begin{cases} F(\theta x), & x \in [N]; \\ 0, & x \notin [N]. \end{cases}$$

For fixed $0 < \eta \leqslant 1/2$, define the Bohr set

$$B_1 := \left\{x \in [\eta N] : \|\theta_i x\|_{\mathbb{T}} \leqslant \eta \textit{ for all } i\right\}.$$

For integers $c, W \geqslant 1$ let B_2' denote a subset of the quadratic Bohr set

$$B_2 := \left\{x \in [\eta(N/W)^{1/2}] : c \mid x \textit{ and } \left\|\theta_i W x^2/c\right\|_{\mathbb{T}}, \left\|\theta_i x/c\right\|_{\mathbb{T}} \leqslant \eta \textit{ for all } i\right\}.$$

Then

$$\sum_{x\in[N]} \sum_{\substack{d_i \in B_1\\ y_j, z_k \in B_2'}} f(x) f(x+cd_1)\cdots f(x+cd_{s_1}) f\left(x + L_1(d) + \tfrac{WL_2(y^2)+L_3(z)}{c}\right)$$
$$\geqslant |B_1|^{s_1}|B_2'|^{s_2+s_3} \sum_{x\in[N]} f(x)^{s_1+1} - O_c(\eta M |B_1|^{s_1}|B_2'|^{s_2+s_3} N).$$

Proof. Suppressing dependence on L_i, there exists a constant $C \ll_c 1$ such that if $d_i \in B_1$, $y_j, z_k \in B_2$ and

$$C\eta N \leqslant x \leqslant N - C\eta N \tag{3.1}$$

then

$$x + cd_1,\ \ldots,\ x + cd_{s_1},\ x + L_1(d) + \tfrac{WL_2(y^2)+L_3(z)}{c} \in [N].$$

Restricting our summation over x to (3.1) introduces an error of $O_c(\eta|B_1|^{s_1}|B_2'|^{s_2+s_3}N)$. On restricting in this manner, each term in our summation satisfies $f(x+cd_i) = f(x) + O_c(\eta M)$ and

$$f\left(x + L_1(d) + \tfrac{WL_2(y^2)+L_3(z)}{c}\right) = f(x) + O(\eta M).$$

The result follows. □

Lemma 3.5 (L^1-control)**.** *Let L_1, L_2, L_3 be linear forms, each in s_i variables and let S_1, S_2 be finite sets of integers, with every element of S_2 divisible by c. For any*

1-bounded functions $f_i : \mathbb{Z} \to \mathbb{C}$ with support in $[N]$ we have the estimate

$$\left| \sum_{x\in\mathbb{Z}} \sum_{\substack{d_i \in S_1\\ y_j, z_k \in S_2}} f_0(x) f_1(x+cd_1)\cdots f_{s_1}(x+cd_{s_1}) f_{-1}\left(x + L_1(d) + \tfrac{WL_2(y^2)+L_3(z)}{c}\right)\right| \leqslant \min_i \|f_i\|_{L^1[N]} |S_1|^{s_1} |S_2|^{s_2+s_3}.$$

Proof. This is clear from the triangle inequality for $i = 0$. The same argument applies for other values of i on changing variables. □

Combining our structured count with L^1-control, we can prove a version of the structured counting lemma which allows for small perturbations in the L^1-norm. The proof follows from the standard telescoping identity

$$\prod_i g(x_i) - \prod_i f(x_i) = \sum_i (g(x_i) - f(x_i)) \prod_{j<i} g(x_j) \prod_{k>i} f(x_k)$$

Corollary 3.6. *Let L_1, L_2, L_3 be linear forms, each in s_i variables. Given an M-Lipschitz function $F : \mathbb{T}^d \to [0, 1]$ and $\theta \in \mathbb{T}^d$, define $f : [N] \to [0, 1]$ by*

$$f(x) := F(\theta x), \qquad (x \in [N]).$$

For fixed $0 < \varepsilon \leqslant 1/2$, define the Bohr set

$$B_1 := \{x \in [\varepsilon N] : \|\theta_i x\|_{\mathbb{T}} \leqslant \varepsilon/M \text{ for all } i\}.$$

For integers $c, W \geqslant 1$ let B_2' denote a subset of the quadratic Bohr set

$$B_2 := \left\{x \in [\varepsilon (N/W)^{1/2}] : c \mid x \text{ and } \left\|\theta_i W x^2/c\right\|_{\mathbb{T}}, \left\|\theta_i x/c\right\|_{\mathbb{T}} \leqslant \varepsilon/M \text{ for all } i\right\}.$$

Then for any function $g : [N] \to [-1, 1]$ with $\|f - g\|_{L^1[N]} \leqslant \varepsilon N$ we have

$$\sum_{x\in[N]} \sum_{\substack{d_i \in B_1\\ y_j, z_k \in B_2'}} g(x) g(x+cd_1)\cdots g(x+cd_{s_1}) g\left(x + L_1(d) + \tfrac{WL_2(y^2)+L_3(z)}{c}\right) \geqslant |B_1|^{s_1} |B_2'|^{s_2+s_3} \sum_{x\in[N]} f(x)^{s_1+2} - O_c(\varepsilon |B_1|^{s_1} |B_2'|^{s_2+s_3} N).$$

The uniform part of the decomposition afforded by the regularity lemma (Lemma 3.3) has small Fourier coefficients. The next lemma shows that such functions make negligible contribution to the count in Theorem 3.1.

Lemma 3.7 (Fourier control)**.** *Let L_1, L_2, L_3 denote non-singular linear forms, each in s_i variables with $s_1 \geqslant 2$ and $s_1 + s_2 \geqslant 3$. Let W be a positive integer and suppose that $W = 1$ or $s_3 \geqslant 1$. Then for any positive integer $N \geqslant W^3$, 1-bounded functions $f_1, \dots, f_{s_1} : [N] \to \mathbb{C}$ and set $B \subset [(N/W)^{1/2}]$ we have the estimate*

$$\left| \sum_{L_1(x) = WL_2(y^2)+L_3(z)} f_1(x_1)\cdots f_{s_1}(x_{s_1}) 1_B(y_1)\cdots 1_B(y_{s_2}) 1_B(z_1)\cdots 1_B(z_{s_3})\right| \ll N^{s_1 + \frac12(s_2+s_3)-1} W^{-\frac12(s_2+s_3)} \min_i \left(\frac{\|\hat{f}_i\|_\infty}{N}\right)^{1/3}. \quad (3.2)$$

Proof. Write

$$\tilde{1}_B(\alpha) := \sum_{x\in B} e(\alpha x^2)$$

and

$$L_i(x) = c_1^{(i)}x_1 + \cdots + c_{s_i}^{(i)}x_{s_i}. \tag{3.3}$$

The orthogonality relations give the identity

$$\sum_{L_1(x)=WL_2(y^2)+L_3(z)} f_1(x_1)\cdots f_{s_1}(x_{s_1})1_B(y_1)\cdots 1_B(y_{s_2})1_B(z_1)\cdots 1_B(z_{s_3})$$
$$= \int_{\mathbb{T}} \prod_i \hat{f}_i\left(c_i^{(1)}\alpha\right)\prod_j \tilde{1}_B\left(Wc_j^{(2)}\alpha\right)\prod_k \hat{1}_B\left(c_k^{(3)}\alpha\right)\mathrm{d}\alpha. \tag{3.4}$$

Let us first suppose that $s_1 \geqslant 3$. Fix distinct integers $i, j, k \in [s_1]$. In this case we may estimate all exponential sums involving 1_B trivially, then employ Parseval to bound (3.4) by

$$N^{s_1-3}(N/W)^{\frac{1}{2}(s_2+s_3)}\int_{\mathbb{T}}\left|\hat{f}_i\left(c_i^{(1)}\alpha\right)\right|\left|\hat{f}_j\left(c_j^{(1)}\alpha\right)\right|\left|\hat{f}_k\left(c_k^{(1)}\alpha\right)\right|\mathrm{d}\alpha$$
$$\leqslant N^{s_1-3}(N/W)^{\frac{1}{2}(s_2+s_3)}\|\hat{f}_i\|_\infty\|\hat{f}_j\|_2\|\hat{f}_k\|_2 \leqslant N^{s_1-2}(N/W)^{\frac{1}{2}(s_2+s_3)}\|\hat{f}_i\|_\infty.$$

Henceforth we assume that $s_1 = 2$, and so $s_2 \geqslant 1$ (since $s_1 + s_2 \geqslant 3$). Let us deal with the case in which $W = 1$ and $s_3 = 0$. Then the orthogonality relations show our count equals

$$\sum_{L_1(x)=L_2(y^2)} f_1(x_1)f_2(x_2)1_B(y_1)\cdots 1_B(y_{s_2})$$
$$= \int_{\mathbb{T}} \hat{f}_1\left(c_1^{(1)}\alpha\right)\hat{f}_2\left(c_2^{(1)}\alpha\right)\prod_j \tilde{1}_B\left(c_j^{(2)}\alpha\right)\mathrm{d}\alpha.$$

By Hölder's inequality and the trivial bound on exponential sums, the Fourier integral is at most

$$N^{\frac{1}{2}(s_2-1)}\|\hat{f}_1\|_\infty^{1/3}\|\hat{f}_1\|_2^{2/3}\|\hat{f}_2\|_2\|\tilde{1}_B\|_6.$$

By Parseval $\|\hat{f}_i\|_2 \leqslant N^{1/2}$, and by (say) Bourgain's restriction estimate [Bou89] we have $\|\tilde{1}_B\|_6 \ll N^{1/3}$ (more elementary proofs exist for the latter). The estimate (3.2) now follows in this case.

Next let us deal with the case that W is arbitrary, in which case we may assume that $s_3 \geqslant 1$, in addition to our assumptions that $s_1 = 2$ and $s_2 \geqslant 1$. As above, Hölder's inequality allows us to bound (3.4) by

$$(N/W)^{\frac{1}{2}(s_2+s_3-2)}\|\hat{f}_1\|_\infty^{1/3}\|\hat{f}_1\|_2^{2/3}\|\hat{f}_2\|_2\|\tilde{1}_B(c_1^{(2)}W\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_6.$$

Hence it suffices to prove that for non-zero integers c and c' we have the estimate

$$\|\tilde{1}_B(cW\alpha)\hat{1}_B(c'\alpha)\|_6 \ll_{c,c'} N^{5/6}W^{-1}. \tag{3.5}$$

By orthogonality, the sixth power of the norm in (3.5) is bounded above by the number of solutions to the equation

$$cW(y_1^2+y_2^2+y_3^2-y_4^2-y_5^2-y_6^2)=c'(z_1+z_2+z_3-z_4-z_5-z_6), \qquad \big(y_i,z_j\in[(N/W)^{1/2}]\big). \quad (3.6)$$

Suppose that both sides of (3.6) are equal to Wm for some $m\in\mathbb{Z}$. Then the size constraints on the right-hand side force $|m|\ll_{c'}(N/W)^{1/2}W^{-1}$. Hence there are at most $O_{c'}((N/W)^{1/2}W^{-1}+1)$ choices for m, and given this choice there are at most $(N/W)^{\frac{5}{2}}$ choices for $(z_1,\dots,z_6)$. Furthermore, by orthogonality the number of choices for $(y_1,\dots,y_6)$ is at most

$$\int_{\mathbb{T}}\left|\tilde{1}_{[(N/W)^{1/2}]}(c\alpha)\right|^6 e(\alpha m)\mathrm{d}\alpha \ll \int_{\mathbb{T}}\left|\tilde{1}_{[(N/W)^{1/2}]}(\alpha)\right|^6\mathrm{d}\alpha \ll (N/W)^2,$$

the latter following from (say) Bourgain's restriction estimate [Bou89] (again, more elementary proofs exist). The required estimate (3.5) follows. □

Lemma 3.8 (Quadratic Bohr set bound)**.** *Let $0<\eta\leqslant 1/2$ and $\alpha,\beta\in\mathbb{T}^d$. Then for any positive integer N, either $N\ll_{\eta,d}1$ or*

$$\#\left\{x\in[N]:\left\|\alpha_i x^2\right\|_{\mathbb{T}},\left\|\beta_i x\right\|_{\mathbb{T}}\leqslant\eta \text{ for all } i\right\}\gg_{\eta,d}N.$$

Proof. This follows from Tao [Tao12, Ex.1.1.23]. □

We now begin our proof of Theorem 3.1 in earnest. Some of our summations become cleaner if we view functions $f:[N]\to\mathbb{C}$ as functions $f:\mathbb{Z}\to\mathbb{C}$ which are equal to zero outside of $[N]$. We first prove Theorem 3.1 under the assumption that we have a colouring $C_1\cup\dots\cup C_r=[(N/W)^{1/2}]$ of the full interval. We deduce the stated version subsequently.

We apply Lemma 3.3 to the indicator functions of the sets A_i for $i=1,\dots,r$. The precise values of ε and $\mathcal{F}$ in our use of this lemma are to be determined. In this way we obtain $M\ll_{\varepsilon,\mathcal{F},r}1$ and decompositions

$$1_{A_i}=f_i^{\mathrm{str}}+f_i^{\mathrm{sml}}+f_i^{\mathrm{unf}} \qquad (1\leqslant i\leqslant r)$$

which satisfy the conclusions of the arithmetic regularity lemma. In particular, there exists $d\leqslant M$ and $\theta\in\mathbb{T}^d$, such that for each i there is an M-Lipschitz function $F_i:\mathbb{T}^d\to[0,1]$ with $f_i^{\mathrm{str}}(x)=F_i(\theta x)$ for all $x\in[N]$. Since f_i^{str} has the same mean as 1_{A_i}, we have

$$\sum_{x\in[N]}f_i^{\mathrm{str}}(x)^{s_1}\geqslant\frac{1}{N^{s_1-1}}\left(\sum_{x\in[N]}f_i^{\mathrm{str}}(x)\right)^{s_1}\geqslant\delta^{s_1}N.$$

Write

$$L_1(x)=c_1x_1+\dots+c_{s_1}x_{s_1},$$

and set

$$\tilde{L}_1(d_3,\dots,d_{s_1})=-\left(c_3d_3+\dots+c_{s_1}d_{s_1}\right).$$

Taking B_1 and B_2 as in Corollary 3.6, with $c := c_1$, we deduce that for $g_i := f_i^{\mathrm{str}} + f_i^{\mathrm{sml}}$ and $B_2' \subset B_2$ we have

$$\sum_{x\in[N]} \sum_{\substack{d_i\in B_1\\ y_j,z_k\in B_2'}} g_i(x)g_i(x+c_1d_3)\cdots g_i(x+c_1d_{s_1})g_i\left(x+\tilde{L}_1(d)+\tfrac{WL_2(y^2)+L_3(z)}{c_1}\right)$$
$$\geqslant \delta^{s_1}|B_1|^{s_1-2}|B_2'|^{s_2+s_3}N - O(\varepsilon|B_1|^{s_1-2}|B_2'|^{s_2+s_3}N).$$

Hence we may take ε satisfying $\varepsilon^{-1} \ll \delta^{-s_1}$ and ensure that

$$\sum_{x\in[N]} \sum_{\substack{d_i\in B_1\\ y_j,z_k\in B_2'}} g_i(x)g_i(x+c_1d_1)\cdots g_i(x+c_1d_{s_1})g_i\left(x+\tilde{L}_1(d)+\tfrac{WL_2(y^2)+L_3(z)}{c_1}\right)$$
$$\gg \delta^{s_1}|B_1|^{s_1-2}|B_2'|^{s_2+s_3}N.$$

By the pigeon-hole principle, there exists a colour class C_j satisfying

$$|C_j \cap B_2| \geqslant |B_2|/r. \tag{3.7}$$

On setting $B_2' := C_j \cap B_2$ and employing Lemma 3.8, we deduce that

$$\sum_{x\in[N]} \sum_{\substack{d_i\in B_1\\ y_j,z_k\in B_2'}} g_i(x)g_i(x+c_1d_1)\cdots g_i(x+c_1d_{s_1})g_i\left(x+\tilde{L}_1(d)+\tfrac{WL_2(y^2)+L_3(z)}{c_1}\right)$$
$$\gg_{\delta,r,M} N^{s_1-1}(N/W)^{\frac{1}{2}(s_2+s_3)}.$$

Notice that under the assumption that $c_1 \mid y_j$ and $c_1 \mid z_k$, we obtain an integer solution to the equation $L_1(x) = WL_2(y^2) + L_3(z)$ on setting

$$x_1 := x + \tilde{L}_1(d) + \tfrac{WL_2(y^2)+L_3(z)}{c_1},\ x_2 := x,$$
$$x_3 := x + c_1d_3,\ \ldots\ ,\ x_{s_1} := x + c_1d_{s_3}.$$

Using the non-negativity of $g_i := f_i^{\mathrm{str}} + f_i^{\mathrm{sml}}$, we deduce that for $C = C_j$ satisfying (3.7) we have

$$\sum_{\substack{L_1(x)=WL_2(y^2)+L_3(z)\\ y_j,z_k\in C\cap[(N/W)^{1/2}]}} g_i(x_1)\cdots g_i(x_{s_1}) \gg_{\delta,r,M} N^{s_1-1}(N/W)^{\frac{1}{2}(s_2+s_3)}. \tag{3.8}$$

Employing Lemma 3.7 and a telescoping identity then gives

$$\sum_{\substack{L_1(x)=WL_2(y^2)+L_3(z)\\ y_j,z_k\in C\cap[(N/W)^{1/2}]}} 1_{A_i}(x_1)\cdots 1_{A_i}(x_{s_1}) = \sum_{\substack{L_1(x)=WL_2(y^2)+L_3(z)\\ y_j,z_k\in C\cap[(N/W)^{1/2}]}} g_i(x_1)\cdots g_i(x_{s_1})$$
$$+ O\left(N^{s_1-1}(N/W)^{\frac{1}{2}(s_2+s_3)}\mathcal{F}(M)^{-1/3}\right)$$

Hence taking $\mathcal{F}(M)$ sufficiently large in terms of the implicit constant in (3.8), we conclude that for the colour class $C = C_j$ satisfying (3.7) we have

$$\sum_{\substack{L_1(x)=WL_2(y^2)+L_3(z)\\ y_j,z_k\in C\cap[(N/W)^{1/2}]}} 1_{A_i}(x_1)\cdots 1_{A_i}(x_{s_1}) \gg_{\delta,r} N^{s_1-1}(N/W)^{\frac{1}{2}(s_2+s_3)}. \tag{3.9}$$

This completes the proof of Theorem 3.1 under the assumption that our colouring is of the full interval $C_1 \cup \cdots \cup C_r = [(N/W)^{1/2}]$. Notice that if $s_2 + s_3 = 0$ then this vacuously implies the stated version of Theorem 3.1. Let us therefore assume that $s_2 + s_3 > 0$. Let η denote the implicit constant in (3.9) divided through by $2(s_2 + s_3)$. Since the inverse image of the linear form L_1 has size at most $N^{s_1 - 1}$ in $[N]^{s_1}$, the number of solutions to the equation $L_1(x) = WL_2(y^2) + L_3(z)$ with $x_i \in [N]$, $y_j, z_k \in [(N/W)^{1/2}]$ and either $y_j \leqslant \eta(N/W)^{1/2}$ for some j or $z_k \leqslant \eta(N/W)^{1/2}$ for some k is at most

$$(s_2 + s_3)\eta N^{s_1 - 1}(N/W)^{\frac{1}{2}(s_2+s_3)}.$$

It follows that the bound (3.9) remains valid under the assumption that $C_1 \cup \cdots \cup C_r = [\eta(N/W)^{1/2}, (N/W)^{1/2}]$, as required to prove Theorem 3.1 in full generality.

4. An abstract restriction estimate

To prove the quadratic counting theorem (Theorem 1.7) we would like to prove an analogue of the Fourier control lemma (Lemma 3.7), which was key to our proof of the linear counting theorem (Theorem 1.4). The main ingredients in our proof of the Fourier control lemma were Hölder's inequality and estimates for the L^p-norm of certain exponential sums. In order to prove an analogous result for our quadratic counting theorem (see Lemma 6.2) we require four distinct L^p-estimates, each of which involves the product of two distinct exponential sums (see Lemma 6.1). We term these *mixed restriction estimates*. To avoid repetition, we begin by proving an abstract restriction estimate, then verify that the four exponential sums of relevance satisfy the hypotheses of this theorem.

In the following $[-N, N]$ denotes an interval of integers.

Definition 4.1 (Major arc hypothesis)**.** We say that $\nu : [-N, N] \to [0, \infty)$ satisfes a *major arc hypothesis with constant* K if for all $1 \leqslant a \leqslant q \leqslant Q$ with $\mathrm{hcf}(a, q) = 1$ and $\left\|\alpha - \frac{a}{q}\right\|_{\mathbb{T}} \leqslant Q/N$ we have

$$\frac{|\hat{\nu}(\alpha)|}{\|\nu\|_1} \leqslant Kq^{-1} \max\left\{1, \left\|\alpha - \tfrac{a}{q}\right\|_{\mathbb{T}} N\right\}^{-1} + Q^{O(1)}N^{-\Omega(1)}. \tag{4.1}$$

If $K = O(1)$ then we simply say that ν satisfes a *major arc hypothesis.*

Definition 4.2 (Minor arc hypothesis)**.** We say that $\nu : [-N, N] \to [0, \infty)$ satisfies a *minor arc hypothesis* if for any $\delta > 0$ we have the implication

$$|\hat{\nu}(\alpha)| \geqslant \delta \|\nu\|_1 \quad \implies \quad \left[\exists q \ll \delta^{-O(1)} \text{ such that } \|q\alpha\|_{\mathbb{T}} \ll \delta^{-O(1)}/N\right].$$

Here it is implicitly understood that q is a positive integer.

Definition 4.3 (Hua-type hypothesis)**.** We say that $\nu : [-N, N] \to [0, \infty)$ satisfies the *Hua-type hypothesis with exponent* ε if we have the bound

$$\|\nu\|_2 \leqslant \|\nu\|_1 N^{\varepsilon - 1}.$$

Theorem 4.4 (Restriction estimate)**.** *Let* $\nu : [-N, N] \to [0, \infty)$ *satisfy the major and minor arc hypotheses, the major arc hypothesis with constant* K*. Given* $p > 2$ *there exists*[4] $\varepsilon = \Omega_p(1)$ *such that if* ν *satisfies the Hua-type hypothesis with*

[4]The value of ε depends on the implicit constants in the major and minor arc hypotheses. However, these are absolute constants in our applications.

exponent ε, then for any 1-bounded function $\phi : [-N, N] \to \mathbb{C}$ we have

$$\int_{\mathbb{T}} \left|\widehat{\phi\nu}(\alpha)\right|^p \mathrm{d}\alpha \ll_{K,p} \|\nu\|_1^p N^{-1}.$$

Our proof of Theorem 4.4 follows Bourgain's distributional approach [Bou89], which has a nice exposition due to Henriot and Hughes [HH18].

Lemma 4.5 (Distributional estimate)**.** *Let $\nu : [-N, N] \to [0, \infty)$ satisfy the major and minor arc hypotheses, the major arch hypothesis with constant K. Given*[5] *$0 < \delta \leqslant 1/2$ and a 1-bounded function $\phi : [-N, N] \to \mathbb{C}$, define*

$$E_\delta(\phi, \nu) := \left\{\alpha \in \mathbb{T} : \left|\widehat{\phi\nu}(\alpha)\right| > \delta \|\nu\|_1\right\}. \tag{4.2}$$

Then for any $\varepsilon > 0$, either[6] *$N \leqslant \delta^{-O_\varepsilon(1)}$ or*

$$\mathrm{meas}\big(E_\delta(\phi, \nu)\big) \ll_{K,\varepsilon} N^{-1}\delta^{-2-\varepsilon}.$$

Proof of Theorem 4.4 given Lemma 4.5. Let E_δ be as in (4.2) with $0 < \delta \leqslant 1/2$. By Lemma 4.5 with $\varepsilon = \frac{p}{2} - 1$, either $N \leqslant \delta^{-O_p(1)}$ or

$$\mathrm{meas}(E_\delta) \ll_{K,p} N^{-1}\delta^{-1-\frac{p}{2}}.$$

It follows that there exists $\Delta \leqslant N^{-\Omega_p(1)}$ such that for any $\delta \in (\Delta, 1/2]$ we have

$$\mathrm{meas}(E_\delta) \ll_{K,p} N^{-1}\delta^{-1-\frac{p}{2}}. \tag{4.3}$$

By dyadic decomposition

$$\int_{\mathbb{T}} \left|\widehat{\phi\nu}(\alpha)\right|^p \mathrm{d}\alpha \leqslant \int_{\mathbb{T}\setminus E_\Delta} \left|\widehat{\phi\nu}(\alpha)\right|^p \mathrm{d}\alpha + \sum_{1 \leqslant j < \log_2(1/\Delta)} \int_{E_{2^{-j}}\setminus E_{2^{1-j}}} \left|\widehat{\phi\nu}(\alpha)\right|^p \mathrm{d}\alpha.$$

Since $\Delta \leqslant N^{-\Omega_p(1)}$, we can take $\varepsilon = \varepsilon(p)$ sufficiently small in our Hua-type hypothesis (Definition 4.3) to deduce that

$$\begin{aligned}\int_{\mathbb{T}\setminus E_\Delta} \left|\widehat{\phi\nu}(\alpha)\right|^p \mathrm{d}\alpha \leqslant (\Delta \|\nu_1\|_1)^{p-2} \int_{\mathbb{T}} \left|\widehat{\phi\nu}(\alpha)\right|^2 \mathrm{d}\alpha &\leqslant \Delta^{p-2} \|\nu\|_1^p N^{\varepsilon-1} \\ &\leqslant \|\nu\|_1^p N^{-1}.\end{aligned}$$

By (4.3) we have

$$\begin{aligned}\sum_{1 \leqslant j < \log_2(1/\Delta)} \int_{E_{2^{-j}}\setminus E_{2^{1-j}}} \left|\widehat{\phi\nu}(\alpha)\right|^p \mathrm{d}\alpha &\leqslant \|\nu\|_1^p \sum_{1 \leqslant j < \log_2(1/\Delta)} 2^{(1-j)p}\mathrm{meas}(E_{2^{-j}}) \\ &\ll_{K,p} \|\nu\|_1^p N^{-1} \sum_{j=1}^{\infty} 2^{(1-j)p} 2^{j(1+\frac{p}{2})}.\end{aligned}$$

The latter sum converges to an absolute constant of order $O_p(1)$ since $p > 2$. □

Our proof of Lemma 4.5 utilises the following divisor bound.

[5] Our assumption that $\delta \leqslant 1/2$ is a convenience which allows us to replace bounds of the form $O(\delta^{-O(1)})$ with $\delta^{-O(1)}$.

[6] The careful reader will observe that the implicit constants in our conclusion depend on the implicit constants in our major/minor arc hypotheses.

Lemma 4.6. *Let*

$$d(n,Q) := \sum_{\substack{1\leqslant q\leqslant Q\\ q\mid n}} 1. \tag{4.4}$$

Then for any integer $B \geqslant 1$ and any real $X \geqslant 1$ we have

$$\sum_{|n|\leqslant X} d(n,Q)^B \ll_{\varepsilon,B} Q^B + Q^{\varepsilon} X.$$

Proof. We follow Bourgain [Bou89, p.307]:

$$\sum_{|n|\leqslant X} d(n,Q)^B = \sum_{1\leqslant q_1,\dots,q_B\leqslant Q}\ \sum_{\substack{|n|\leqslant X\\ q_i\mid n}} 1 \ll \sum_{1\leqslant q_1,\dots,q_B\leqslant Q}\left(1+\frac{X}{[q_1,\dots,q_B]}\right) \leqslant$$

$$Q^B + X\sum_{1\leqslant q\leqslant Q^B}\frac{d(q)^B}{q} \ll_{\varepsilon,B} Q^B + X\sum_{1\leqslant q\leqslant Q^B}\frac{1}{q^{1-\frac{\varepsilon}{B}}} \ll_{\varepsilon,B} Q^B + Q^{\varepsilon}X.$$

□

Proof of Lemma 4.5. Write $E_\delta := E_\delta(\phi,\nu)$ and

$$\epsilon(\alpha) := \begin{cases} \frac{\widehat{\phi\nu}(\alpha)}{|\widehat{\phi\nu}(\alpha)|}, & \text{if } \widehat{\phi\nu}(\alpha)\neq 0;\\ 0, & \text{otherwise.}\end{cases}$$

Then

$$\begin{aligned}\delta\,\|\nu\|_1 \operatorname{meas}(E_\delta) &\leqslant \int_{E_\delta}\left|\widehat{\phi\nu}(\alpha)\right|\mathrm{d}\alpha = \int_{E_\delta}\widehat{\phi\nu}(\alpha)\overline{\epsilon(\alpha)}\mathrm{d}\alpha\\ &= \sum_x \phi(x)\sqrt{\nu(x)}\sqrt{\nu(x)}\int_{E_\delta} e(\alpha x)\overline{\epsilon(\alpha)}\mathrm{d}\alpha\\ &\leqslant \|\nu\|_1^{1/2}\left(\sum_x \nu(x)\left|\int_{E_\delta} e(\alpha x)\overline{\epsilon(\alpha)}\mathrm{d}\alpha\right|^2\right)^{1/2}.\end{aligned}$$

Expanding absolute values, then using linearity of integration and the triangle inequality, we have

$$\sum_x \nu(x)\left|\int_{E_\delta} e(\alpha x)\overline{\epsilon(\alpha)}\mathrm{d}\alpha\right|^2 \leqslant \int_{E_\delta}\int_{E_\delta}|\hat{\nu}(\alpha_1-\alpha_2)|\mathrm{d}\alpha_1\mathrm{d}\alpha_2.$$

Hence we deduce the Tomas-Stein inequality

$$\delta^2\,\|\nu\|_1 \operatorname{meas}(E_\delta)^2 \leqslant \int_{E_\delta}\int_{E_\delta}|\hat{\nu}(\alpha_1-\alpha_2)|\mathrm{d}\alpha_1\mathrm{d}\alpha_2.$$

Consider the Fejér kernel

$$F_N(\alpha) := N^{-1}\left|\hat{1}_{[N]}(\alpha)\right|^2 = \sum_n \left(1-\tfrac{|n|}{N}\right)_+ e(\alpha n),$$

a trigonometric polynomial of degree $N-1$ which is also a probability measure on $\mathbb{T}$. Set

$$\psi_N(\alpha) := \left(1+e(\alpha N)+e(-\alpha N)\right)F_N(\alpha).$$

For $|n| \leqslant N$ one can check that

$$\hat{\psi}_N(n) = \left(1 - \tfrac{|n|}{N}\right)_+ + \left(1 - \tfrac{|n-N|}{N}\right)_+ + \left(1 - \tfrac{|n+N|}{N}\right)_+ = 1.$$

One can also check that if $n \in \mathbb{Z} \setminus [-N, N]$ then

$$\int_{\mathbb{T}} \hat{\nu}(\alpha) e(-\alpha n) \mathrm{d}\alpha = 0.$$

Therefore the Fourier coefficients of $\alpha \mapsto \hat{\nu}(\alpha)$ agree with those of the convolution

$$\hat{\nu} * \psi_N : \alpha \mapsto \int_{\mathbb{T}} \hat{\nu}(\alpha - \beta)\psi_N(\beta)\mathrm{d}\beta.$$

By Fejér's theorem [Kat04, Theorem 3.1] these functions must be identical and we deduce that

$$\begin{aligned}\delta^2 \|\nu\|_1 \operatorname{meas}(E_\delta)^2 &\leqslant \int_{\mathbb{T}^3} |\hat{\nu}(\alpha_1 - \alpha_2 - \beta)\psi_N(\beta)| 1_{E_\delta}(\alpha_1) 1_{E_\delta}(\alpha_2) \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \mathrm{d}\beta \\ &\ll \int_{\mathbb{T}^3} |\hat{\nu}(\alpha_1 - \alpha_2 - \beta)| F_N(\beta) 1_{E_\delta}(\alpha_1) 1_{E_\delta}(\alpha_2) \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \mathrm{d}\beta.\end{aligned}$$

Let $Q \geqslant 1$ (to be determined) and write

$$\mathfrak{M} := \bigcup_{\substack{1 \leqslant a \leqslant q \leqslant Q \\ \mathrm{hcf}(a,q)=1}} \left\{ \alpha \in \mathbb{T} : \left\| \alpha - \tfrac{a}{q} \right\|_{\mathbb{T}} \leqslant Q/N \right\}. \tag{4.5}$$

Our minor arc hypothesis (Definition 4.2) shows that if $\alpha \in \mathbb{T} \setminus \mathfrak{M}$ then $|\hat{\nu}(\alpha)| \ll Q^{-\Omega(1)} \|\nu\|_1$. Hence

$$\int_{\alpha_1 - \alpha_2 - \beta \in \mathbb{T} \setminus \mathfrak{M}} |\hat{\nu}(\alpha_1 - \alpha_2 - \beta)| F_N(\beta) 1_{E_\delta}(\alpha_1) 1_{E_\delta}(\alpha_2) \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \mathrm{d}\beta \\ \ll Q^{-\Omega(1)} \|\nu\|_1 \operatorname{meas}(E_\delta)^2.$$

It follows that either $Q \leqslant \delta^{-O(1)}$ or

$$\delta^2 \|\nu\|_1 \operatorname{meas}(E_\delta)^2 \ll \int_{\alpha_1 - \alpha_2 - \beta \in \mathfrak{M}} |\hat{\nu}(\alpha_1 - \alpha_2 - \beta)| F_N(\beta) 1_{E_\delta}(\alpha_1) 1_{E_\delta}(\alpha_2) \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \mathrm{d}\beta.$$

Letting $C = O(1)$ denote a sufficiently large absolute constant, set

$$Q := \delta^{-C} \leqslant \delta^{-O(1)}. \tag{4.6}$$

Then, for any $p \geqslant 1$, Hölder's inequality yields

$$\delta^{2p} \|\nu\|_1^p \operatorname{meas}(E_\delta)^2 \ll \int_{\alpha_1 - \alpha_2 - \beta \in \mathfrak{M}} |\hat{\nu}(\alpha_1 - \alpha_2 - \beta)|^p F_N(\beta) 1_{E_\delta}(\alpha_1) 1_{E_\delta}(\alpha_2) \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \mathrm{d}\beta.$$

Our major arc hypothesis (Definition 4.1) implies that for $p \in [1, 2]$ we have the bound

$$\frac{1_{\mathfrak{M}}(\alpha) |\hat{\nu}(\alpha)|^p}{\|\nu\|_1^p} \ll_K \sum_{1 \leqslant a \leqslant q \leqslant Q} q^{-1} \max\left\{1, \left\|\alpha - \tfrac{a}{q}\right\|_{\mathbb{T}} N\right\}^{-p} + Q^{O(1)} N^{-\Omega(1)}.$$

Hence for $p \in [1,2]$ we deduce that either $N \leqslant \delta^{-O(1)}$ or

$$\delta^{2p}\mathrm{meas}(E_\delta)^2 \ll_K \sum_{1\leqslant a\leqslant q\leqslant Q} q^{-1} \int_{E_\delta\times E_\delta\times\mathbb{T}} \max\left\{1, \left\|\alpha_1-\alpha_2-\beta-\tfrac{a}{q}\right\|_{\mathbb{T}} N\right\}^{-p} F_N(\beta). \quad (4.7)$$

Set

$$\mu_{Q,p}(\alpha) := \sum_{1\leqslant a\leqslant q\leqslant Q} q^{-1} \max\left\{1, \left\|\alpha-\tfrac{a}{q}\right\|_{\mathbb{T}} N\right\}^{-p}.$$

Then we can re-write (4.7) as

$$\delta^{2p}\mathrm{meas}(E_\delta)^2 \ll_K \int_{E_\delta}\int_{E_\delta} \mu_{Q,p} * F_N(\alpha_1-\alpha_2)\mathrm{d}\alpha_1\mathrm{d}\alpha_2. \quad (4.8)$$

Using the following normalisation for inner products

$$\langle f, g\rangle := \int_{\mathbb{T}} f(\alpha)\overline{g(\alpha)}\mathrm{d}\alpha,$$

the inequality (4.8) implies that

$$\delta^{2p}\mathrm{meas}(E_\delta)^2 \ll_K \left\langle \mu_{Q,p} * F_N, 1_{E_\delta} * 1_{-E_\delta}\right\rangle.$$

By Parseval's theorem [Kat04, Theorem 5.5(d)] and the convolution identity [Kat04, Theorem 1.7] we have

$$\left\langle \mu_{Q,p} * F_N, 1_{E_\delta} * 1_{-E_\delta}\right\rangle = \left\langle \hat{\mu}_{Q,p}\hat{F}_N, \left|\hat{1}_{E_\delta}\right|^2\right\rangle \leqslant \sum_{|n|<N} |\hat{\mu}_{Q,p}(n)| \left|\hat{1}_{E_\delta}(n)\right|^2.$$

Hence any $p \in [1,2]$ yields the estimate

$$\delta^{2p}\mathrm{meas}(E_\delta)^2 \ll_K \sum_{|n|<N} |\hat{\mu}_{Q,p}(n)| \left|\hat{1}_{E_\delta}(n)\right|^2.$$

Recalling the definition (4.4) of $d(n,Q)$, a change of variables shows that for any $p > 1$ we have

$$|\hat{\mu}_{Q,p}(n)| \leqslant d(n,Q)\int_{\mathbb{T}} \max\left\{1, \|\alpha\|_{\mathbb{T}} N\right\}^{-p}\mathrm{d}\alpha \ll \frac{d(n,Q)}{(p-1)N}. \quad (4.9)$$

Hence, for any $B \geqslant 1$ and $p \in (1,2]$, Hölder's inequality gives

$$\delta^{2p}\mathrm{meas}(E_\delta)^2 \ll_K \frac{1}{(p-1)N}\left(\sum_{|n|<N} d(n,Q)^B\right)^{1/B}\left(\sum_n \left|\hat{1}_{E_\delta}(n)\right|^{2B/(B-1)}\right)^{1-\frac{1}{B}}.$$

Applying Parseval again, together with Lemma 4.6, we conclude that for any $p \in (1,2]$ and any integer $B \geqslant 1$ we have

$$\delta^{2p}\mathrm{meas}(E_\delta)^2 \ll_{K,B} \frac{1}{(p-1)N}(Q^B + QN)^{1/B}\mathrm{meas}(E_\delta)^{1+\frac{1}{B}}.$$

Set $B := 1 + \lceil 1/\varepsilon\rceil$ and $p := 1 + B^{-1}$. Recalling our choice (4.6) of Q, either $Q^B \leqslant QN$ or $N \leqslant \delta^{-O_\varepsilon(1)}$. In the former case we have

$$\mathrm{meas}(E_\delta) \ll_{K,\varepsilon} N^{-1}Q^{\frac{1}{B-1}}\delta^{-\frac{2pB}{B-1}} \leqslant N^{-1}\delta^{-2-O(\varepsilon)}.$$

The result follows on rescaling ε to absorb the $O(1)$ constant in the exponent. $\square$

5. Exponential sum estimates

The purpose of this section is to prove various bounds on exponential sums which are needed to verify the hypotheses required for our mixed restriction estimates (see Lemma 6.1). Three out of four of these mixed restriction estimates involve the following majorant, which also plays a prominent role in [BP17] and [CLP].

Definition 5.1 (Majorant for squares). For fixed $\xi \in [W]$, with W even, define $\nu = \nu_{W,\xi} : [N] \to [0, \infty)$ by

$$\nu(n) := \begin{cases} Wx + \xi, & \text{if } n = \frac{(Wx+\xi)^2 - \xi^2}{2W} \text{ for some } x \in \mathbb{Z}; \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

Before estimating the Fourier transform of ν we recall Weyl's inequality for squares.

Lemma 5.2 (Weyl's inequality). *Let I be an interval of at most N integers and let $\alpha, \beta \in \mathbb{T}$. Suppose that*

$$\left| \sum_{n \in I} e(\alpha n^2 + \beta n) \right| \geqslant \delta N.$$

Then there exists $q \ll \delta^{-O(1)}$ such that $\|q\alpha\|_{\mathbb{T}} \ll \delta^{-O(1)}/N^2$.

Proof. See Green–Tao [GT08, Lemma A.11]. □

Corollary 5.3 (Coarse minor arc estimate for ν). *Let W be an even positive integer, $\xi \in [W]$ and define $\nu = \nu_{W,\xi}$ as in (5.1). Suppose that*

$$\left|\hat{\nu}(\alpha)\right| \geqslant \delta N.$$

Then either $N \ll W$ or there exists $1 \leqslant a \leqslant q \ll \delta^{-O(1)} W$ such that $\mathrm{hcf}(a, q) = 1$ *and $\left\|\alpha - \frac{a}{q}\right\|_{\mathbb{T}} \ll \delta^{-O(1)}/N$.*

Proof. Summation by parts gives that

$$\sum_{x < n \leqslant y} f(n) a_n = f(y) \sum_{x < n \leqslant y} a_n - \int_x^y f'(t) \left(\sum_{x < n \leqslant t} a_n \right) \mathrm{d}t$$

Hence

$$\begin{aligned} \hat{\nu}(\alpha) &= \sum_{0 < \frac{1}{2}Wx^2 + \xi x \leqslant N} (Wx + \xi) e\left(\alpha\left(\tfrac{1}{2}Wx^2 + \xi x\right)\right) \\ &= \left(\sqrt{2WN + \xi^2}\right) \sum_{0 < \frac{1}{2}Wx^2 + \xi x \leqslant N} e\left(\alpha\left(\tfrac{1}{2}Wx^2 + \xi x\right)\right) \\ &\qquad - W \int_0^{\frac{\sqrt{2WN+\xi^2} - \xi^2}{W}} \sum_{0 < x \leqslant t} e\left(\alpha\left(\tfrac{1}{2}Wx^2 + \xi x\right)\right) \mathrm{d}t. \end{aligned}$$

Since $\xi \in [W]$, one can check that either $N \ll W$ or the interval $\{x : 0 < \frac{1}{2}Wx^2 + \xi x \leqslant N\}$ has length of order $\asymp \sqrt{N/W}$. It therefore follows that if

$|\hat{\nu}(\alpha)| \geqslant \delta N$ then there exists $t \ll \sqrt{N/W}$ such that

$$\left|\sum_{0<\frac{1}{2}Wx^2+\xi x\leqslant t} e\left(\alpha\left(\tfrac{1}{2}Wx^2+\xi x\right)\right)\right| \gg \delta\sqrt{N/W}$$

Applying Weyl's inequality, there exists $q_0 \ll \delta^{-O(1)}$ such that $\left\|q_0\alpha\frac{1}{2}W\right\|_{\mathbb{T}} \ll \delta^{-O(1)}W/N$. Setting $q := \frac{1}{2}Wq_0$ then yields the result. □

Lemma 5.4 (Major arc asymptotic for ν)**.** *Let W be an even positive integer, $\xi \in [W]$ and define $\nu = \nu_{W,\xi}$ as in (5.1). Suppose that $\|q\alpha\|_{\mathbb{T}} = |q\alpha - a|$ for some $q, a \in \mathbb{Z}$ with $q > 0$. Then either $N \ll W$ or*

$$\hat{\nu}(\alpha) = \mathbb{E}_{r\in[q]}e_q\left(a\left(\tfrac{1}{2}Wr^2+\xi r\right)\right)\int_0^N e\left(\left(\alpha-\tfrac{a}{q}\right)t\right)\mathrm{d}t \\ + O\left(\left(\sqrt{WN}+Wq\right)\left(q+\|q\alpha\|_{\mathbb{T}}N\right)\right).$$

Proof. Writing $\beta := \alpha - \frac{a}{q}$ and summing over congruence classes mod q, we have

$$\hat{\nu}(\alpha) = \sum_{r=1}^{q} e_q\left(a\left(\tfrac{1}{2}Wr^2+\xi r\right)\right)\sum_{\substack{0<\frac{1}{2}Wx^2+\xi x\leqslant N\\ x\equiv r \bmod q}}(Wx+\xi)e\left(\beta\left(\tfrac{1}{2}Wx^2+\xi x\right)\right). \quad (5.2)$$

Comparing the inner sum with an integral, as in [Tao, Ex11], we have

$$\sum_{\substack{0<\frac{1}{2}Wx^2+\xi x\leqslant N\\ x\equiv r \bmod q}}(Wx+\xi)e\left(\beta\left(\tfrac{1}{2}Wx^2+\xi x\right)\right) = \\ q^{-1}\int_0^N e(\beta t)\,\mathrm{d}t + O\left(\left(\sqrt{WN}+Wq\right)(1+N|\beta|)\right).$$

Substituting this into (5.2) gives the result. □

Lemma 5.5 (Local Weyl estimate)**.** *For any integers a, b, q with q positive we have*

$$\left|\mathbb{E}_{r\in[q]}e_q\left(ar^2+br\right)\right| \ll \mathrm{hcf}(a,q)^{1/2}q^{-1/2}.$$

Proof. Let $q_0 := q/\mathrm{hcf}(2a,q)$. Squaring and Weyl differencing gives

$$\left|\mathbb{E}_{r\in[q]}e_q\left(ar^2+br\right)\right|^2 \leqslant \mathbb{E}_{h\in[q]}\left|\mathbb{E}_{r\in[q]}e_q(a2hr)\right| \\ = q^{-1}\#\{h\in[q] : 2ah \equiv 0 \bmod q\} = q^{-1}\#\{h\in[q] : h \equiv 0 \bmod q_0\} \\ = q^{-1}\mathrm{hcf}(2a,q) \ll q^{-1}\mathrm{hcf}(a,q). \quad \square$$

Lemma 5.6. *Let W be an even positive integer and let $\xi \in [W]$ with $\mathrm{hcf}(\xi, W) = 1$. Then for any positive integers a and q with $\mathrm{hcf}(a,q) = 1$ we have*

$$\mathbb{E}_{r\in[q]}e_q\left(a\left(\tfrac{1}{2}Wr^2+\xi r\right)\right) \ll \begin{cases} 0, & \text{if } \mathrm{hcf}(q, \frac{1}{2}W) > 1; \\ q^{-1/2}, & \text{otherwise.}\end{cases}$$

Proof. Write $q = q_0 q_1$ where $q_1 = \mathrm{hcf}(\frac{1}{2}W, q)$. Writing $r = r_0 + q_0 r_1$ where $r_0 \in [q_0]$ and $r_1 \in [q_1]$ we have

$$\left|\mathbb{E}_{r\in[q]} e_q\left(a\left(\tfrac{1}{2}Wr^2 + \xi r\right)\right)\right| = \left|\mathbb{E}_{r_0\in[q_0]} e_q\left(a\tfrac{1}{2}Wr_0^2\right)\mathbb{E}_{r_1\in[q_1]} e_q\left(a\left(\xi(r_0+q_0r_1)\right)\right)\right| \leqslant \left|\mathbb{E}_{r_1\in[q_1]} e_{q_1}\left(a\xi r_1\right)\right| = 1_{q_1|\xi}. \quad (5.3)$$

The estimate now follows if $q_1 = \mathrm{hcf}(\frac{1}{2}W, q) > 1$, since in this case $q_1 \nmid \xi$ because $\mathrm{hcf}(\xi, W) = 1$. The case when $\mathrm{hcf}(\frac{1}{2}W, q) = 1$ follows from Lemma 5.5. □

Lemma 5.7. *Let W be an even positive integer and let $\xi \in [W]$ with* $\mathrm{hcf}(\xi, W) = 1$*. Then for any positive integers a and q we have*

$$\mathbb{E}_{r\in[q]} e_q\left(a\left(\tfrac{1}{2}Wr^2 + \xi r\right)\right) \ll \begin{cases} 0 & \text{if } \frac{1}{2}W \text{ and } \frac{q}{(a,q)} \text{ are not coprime;} \\ (a,q)^{1/2}q^{-1/2} & \text{otherwise.} \end{cases}$$

Proof. Write $q = q_0 q_1$ and $a = a_0 q_1$ where $q_1 = \mathrm{hcf}(a, q)$. Summing over residues mod q_0 we have

$$\left|\mathbb{E}_{r\in[q]} e_q\left(a\left(\tfrac{1}{2}Wr^2 + \xi r\right)\right)\right| = \left|\mathbb{E}_{r\in[q_0]} e_{q_0}\left(a_0\left(\tfrac{1}{2}Wr^2 + \xi r\right)\right)\right|.$$

The result follows on applying Lemma 5.3. □

Lemma 5.8 (Refined minor arc estimate for ν)**.** *Let W be an even positive integer and $\xi \in [W]$ with* $\mathrm{hcf}(\xi, W) = 1$*. Define $\nu = \nu_{W,\xi}$ as in (5.1). Suppose that*

$$\left|\hat{\nu}(\alpha)\right| \geqslant \delta N.$$

Then either $N \ll W^{O(1)}$ or there exists $q \ll \delta^{-O(1)}$ such that $\|q\alpha\|_{\mathbb{T}} \ll \delta^{-O(1)}/N$. In particular, either $N \ll W^{O(1)}$ or ν satisfies the minor arc hypothesis (Definition 4.2).

Proof. Applying Corollay 5.3, there exists $1 \leqslant a \leqslant q \ll W\delta^{-O(1)}$ for which $\mathrm{hcf}(a, q) = 1$ and $\left\|\alpha - \frac{a}{q}\right\|_{\mathbb{T}} \ll \delta^{-O(1)}/N$. By Lemma 5.4, either $N \ll (W/\delta)^{O(1)}$ or

$$\left|\mathbb{E}_{r\in[q]} e_q\left(a\left(\tfrac{1}{2}Wr^2 + \xi r\right)\right)\right| \gg \delta.$$

Applying Lemma 5.6, we deduce that $q \ll \delta^{-2}$. Finally we note that $N \ll (W/\delta)^{O(1)}$ implies that either $N \ll W^{O(1)}$ or $N \ll \delta^{-O(1)}$, and the conclusion of the minor arc hypothesis is trivial if the latter holds. □

Lemma 5.9 (Linear exponential estimates)**.** *Let $I \subset \mathbb{R}$ be an interval and $\beta \in \mathbb{R}$. Then*

$$\int_I e(\beta t)\mathrm{d}t \ll \min\left\{\mathrm{meas}(I), |\beta|^{-1}\right\}.$$

If $\alpha \in \mathbb{T}$ then

$$\sum_{x\in I} e(\alpha x) \ll \min\left\{\mathrm{meas}(I) + 1, \|\alpha\|_{\mathbb{T}}^{-1}\right\}.$$

Furthermore

$$\sum_{x\in I} e(\beta x) - \int_I e(\beta t)\mathrm{d}t \ll 1 + |\beta|\mathrm{meas}(I).$$

Proof. The first estimate follows from integration, the second from summing the geometric series, the third from approximating a sum by an integral as in [Tao, Ex11]. □

Lemma 5.10 (Fourier decay)**.** *Suppose that W is divisible by $2\prod_{p\leqslant w} p$ and that* $\mathrm{hcf}(\xi, W) = 1$*. Define $\nu = \nu_{W,\xi}$ as in (5.1). Then either $N \ll W^{O(1)}$ or*

$$\left\|\hat{\nu} - \hat{1}_{[N]}\right\|_{\infty} \ll w^{-1/2}N.$$

Proof. First suppose that $|\hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha)| \geqslant \delta N$. Then by the triangle inequality, either $|\hat{\nu}(\alpha)| \gg \delta N$ or $|\hat{1}_{[N]}(\alpha)| \gg \delta N$. In the latter case, Lemma 5.9 gives that $\|\alpha\|_{\mathbb{T}} \ll \delta^{-1}/N$. We claim that a similar conclusion holds under the assumption that $|\hat{\nu}(\alpha)| \gg \delta N$.

To establish the claim we first repeat the argument of Lemma 5.8 to conclude that either $N \ll (W/\delta)^{O(1)}$ or there exists $1 \leqslant a \leqslant q \ll \delta^{-2}$ with $\mathrm{hcf}(a, q) = 1$ such that $\left\|\alpha - \frac{a}{q}\right\|_{\mathbb{T}} \ll \delta^{-O(1)}/N$ and

$$\left|\mathbb{E}_{r\in[q]} e_q\left(a\left(\tfrac{1}{2}Wr^2 + \xi r\right)\right) \int_0^N e\left(\left(\alpha - \tfrac{a}{q}\right)t\right) \mathrm{d}t\right| \gg \delta N. \tag{5.4}$$

Applying Lemma 5.6, we deduce that $\mathrm{hcf}(q, \frac{1}{2}W) = 1$. Since we are assuming that $\frac{1}{2}W$ is divisible by all primes $p \leqslant w$, we conclude that $q > w$ or $q = 1$. If $q = 1$ then we may bound the integral in (5.4) using Lemma 5.9 to deduce that $\|\alpha\|_{\mathbb{T}} \ll \delta^{-1}/N$, as claimed.

We may therefore conclude that the assumption $|\hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha)| \geqslant \delta N$ implies that either $N \ll (W/\delta)^{O(1)}$, or $w \ll \delta^{-2}$ or

$$\|\alpha\|_{\mathbb{T}} \ll \delta^{-1}/N. \tag{5.5}$$

Supposing that (5.5) holds, if we substitute the approximations given by Lemma 5.9 and Lemma 5.4 into the inequality $|\hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha)| \geqslant \delta N$, then again we deduce that $N \ll (W/\delta)^{O(1)}$.

To summarise: if $|\hat{\nu}(\alpha) - \hat{1}_{[N]}(\alpha)| \geqslant \delta N$ then either $N \ll (W/\delta)^{O(1)}$ or $w \ll \delta^{-2}$. The lemma is complete on taking $\delta := Cw^{-1/2}$ for a sufficiently large absolute constant C, and on observing that $w \leqslant 2\prod_{p\leqslant w} p \leqslant W$ (by Bertrand's postulate, for instance). □

Lemma 5.11 (Quadratic major arc asymptotic)**.** *Let W be a positive integer, $\eta \in (0, 1/2]$ and define the interval*

$$I := \left[\eta(N/W)^{1/2}, (N/W)^{1/2}\right].$$

Suppose that $\|q\alpha\| = |q\alpha - a|$ for some $q, a \in \mathbb{Z}$ with $q > 0$. Then either $N \ll W^{O(1)}$ or

$$\sum_{x\in I} e(\alpha W x^2) = W^{-1/2}\mathbb{E}_{r\in[q]} e_q\left(aWr^2\right) \int_{\eta\sqrt{N}}^{\sqrt{N}} e\left(\beta t^2\right) \mathrm{d}t + O\left(q + \|q\alpha\|_{\mathbb{T}} N\right). \tag{5.6}$$

Proof. Let $\alpha \in \mathfrak{M}(a, q)$ and let β denote the least absolute real in the congruence class $\alpha - \frac{a}{q} \pmod 1$. Summing over residues mod q, we have

$$\sum_{x\in I} e(\alpha W x^2) = \sum_{r=1}^{q} e_q\left(aWr^2\right) \sum_{\substack{x\in I\\ x\equiv r \bmod q}} e\left(\beta W x^2\right). \tag{5.7}$$

Comparing the inner sum with an integral as in [Tao, Ex11] gives

$$\sum_{\substack{x\in I\\ x\equiv r \bmod q}} e\left(\beta W x^2\right) = q^{-1}W^{-1/2}\int_{\eta\sqrt{N}}^{\sqrt{N}} e\left(\beta t^2\right)\mathrm{d}t + O\left(1+|\beta|N\right).$$

Substituting this into (5.7) gives (5.6). □

Lemma 5.12 (Quadratic exponential integral bound)**.** *For $\beta\in\mathbb{R}$ we have*

$$\left|\int_{\eta N}^{N} e(\beta t^2)\mathrm{d}t\right| \ll N\max\left\{1,\eta|\beta|N^2\right\}^{-1}.$$

Proof. Let us show that for $\beta>0$ we have

$$\left|\int_{\eta N}^{N} e(\beta t^2)\mathrm{d}t\right| \ll \frac{1}{\eta\beta N}.$$

The claimed bound then follows on incorporating the trivial estimate of N, and utilising conjugation to deal with $\beta<0$. By a change of variables

$$\int_{\eta N}^{N} e(\beta t^2)\mathrm{d}t = \beta^{-1/2}\int_{\eta^2\beta N^2}^{\beta N^2}\frac{e(v)}{2v^{1/2}}\mathrm{d}v.$$

Integrating by parts shows that for $0<x\leqslant y$ we have

$$\int_x^y \frac{e(t)}{t^{1/2}}\mathrm{d}t \ll x^{-1/2}.$$

□

Lemma 5.13 (Major arc hypotheses)**.** *Let W_1 and W_2 be w-smooth positive integers such that W_1 is divisible by $2\prod_{p\leqslant w}p$. Given $\xi\in[W_1]$ with $\mathrm{hcf}(\xi,W_1)=1$, define $\nu=\nu_{W_1,\xi}:[N]\to[0,\infty)$ as in (5.1). Given $\eta\in(0,1/2]$, define the interval*

$$I := \left[\eta(N/W_2)^{1/2},(N/W_2)^{1/2}\right].$$

Fix non-zero integers $b_1,b_2=O(1)$ and write $B:=|b_1|+|b_2|$. Consider the following four majorants, mapping each $n\in[-BN,BN]$ to one of

$$\sum_{b_1x+b_2y=n}\nu(x)\nu(y),\qquad \sum_{b_1x+b_2W_2y^2=n}\nu(x)1_I(y),$$

$$\sum_{b_1x+b_2y=n}\nu(x)1_I(y),\qquad \sum_{b_1W_2x^2+b_2y=n}1_I(x)1_I(y). \tag{5.8}$$

Then either $N\ll(W_1W_2)^{O(1)}$ or all four majorants satisfy the major arc hypothesis (Definition 4.1), the latter with constant $\eta^{-O(1)}$.

Proof. Let $1\leqslant a\leqslant q\leqslant Q$ with $\mathrm{hcf}(a,q)=1$ and $\left\|\alpha-\frac{a}{q}\right\|_{\mathbb{T}}\leqslant Q/(BN)$. We may choose $\alpha\in\mathbb{R}$ so that $\left|\alpha-\frac{a}{q}\right|=\left\|\alpha-\frac{a}{q}\right\|_{\mathbb{T}}$. Our task is to bound the Fourier transform of our majorant at α.

The first majorant in (5.8) has Fourier transform $\hat{\nu}(b_1\alpha)\hat{\nu}(b_2\alpha)$. We claim that this is bounded in magnitude by

$$\ll q^{-1}N^2\max\left\{1,\left|\alpha-\tfrac{a}{q}\right|N\right\}^{-1}+N^{3/2}W_1^{1/2}Q^2.$$

As $1 \leqslant B \ll 1$, the major arc hypothesis for this majorant follows, since it has L^1-norm $\gg N^2$ (unless $N \ll W_1^{O(1)}$). To establish the claim it suffices to show that for any non-zero integer $b = O(1)$ we have the bound

$$|\hat{\nu}(b\alpha)| \ll q^{-1/2} N \max\left\{1, \left|\alpha - \tfrac{a}{q}\right| N\right\}^{-1} + N^{1/2} W_1 Q^2. \tag{5.9}$$

This follows from Lemmas 5.4, 5.7 and 5.9.

Next we turn to the second majorant in (5.8). Employing Lemma 5.4 and Lemma 5.11, either $N \ll (W_1W_2)^{O(1)}$ or this majorant has Fourier transform bounded in magnitude by

$$W_2^{-1/2} \left|\mathbb{E}_{r_1\in[q]} e_q\left(b_1 a\left(\tfrac12 W_1 r_1^2 + \xi r_1\right)\right)\right| \left|\mathbb{E}_{r_2\in[q]} e_q\left(b_2 a W_2 r_2^2\right)\right| \times \\ \left|\int_1^N e\left(\left(\alpha - \tfrac{a}{q}\right) b_1 t_1\right) \mathrm{d}t_1\right| \left|\int_{\eta\sqrt{N}}^{\sqrt{N}} e\left(\left(\alpha - \tfrac{a}{q}\right) b_2 t_2^2\right) \mathrm{d}t_2\right| + O\left(N^{3/2} W_1 Q^2\right).$$

By Lemma 5.9, the first integral is at most $N \max\{1, \left|\alpha - \frac{a}{q}\right| N\}^{-1}$. Applying the trivial bound to the second integral, it suffices to prove the bound

$$\left|\mathbb{E}_{r_1\in[q]} e_q\left(b_1 a\left(\tfrac12 W_1 r_1^2 + \xi r_1\right)\right)\right| \left|\mathbb{E}_{r_2\in[q]} e_q\left(b_2 a W_2 r_2^2\right)\right| \ll q^{-1}. \tag{5.10}$$

By Lemma 5.7, the left-hand side of (5.10) is zero if $\frac12 W_1$ and $q/(b_1, q)$ are not coprime. We may therefore assume that they are coprime. Since W_2 is w-smooth and $\frac12 W_1$ is divisible by the primorial $\prod_{p\leqslant w} p$, we must have $\mathrm{hcf}(W_2, q/(b_1, q)) = 1$, and so $\mathrm{hcf}(b_2W_2, q) \leqslant b_1 b_2$. Hence Lemmas 5.5 and 5.7 combine to give the bound

$$\left|\mathbb{E}_{r_1\in[q]} e_q\left(b_1 a\left(\tfrac12 W_1 r_1^2 + \xi r_1\right)\right)\right| \left|\mathbb{E}_{r_2\in[q]} e_q\left(b_2 a W_2 r_2\right)\right| \ll_{b_1,b_2} q^{-1}.$$

The major arc bound (4.1) follows with $K = O(1)$.

We simultaneously analyse the third and fourth majorants in (5.12). Under the assumption of our rational approximation to α, we have the lower bound

$$q^{-1} \max\{1, \|\alpha - \tfrac{a}{q}\|_{\mathbb{T}} N\}^{-1} \geqslant Q^{-2}$$

Hence using the trivial bound on the quadratic exponential sum, and Lemma 5.9 on the linear exponential sum, we obtain the major arc bound (4.1) unless

$$\|b_2\alpha\|_{\mathbb{T}} \leqslant Q^2 W_2^{1/2} N^{-1/2}.$$

In this situation the triangle inequality implies that

$$\|b_2 a/q\|_{\mathbb{T}} \leqslant Q^2 W_2^{1/2} N^{-1/2} + QN^{-1}. \tag{5.11}$$

Observe that if $N \ll Q^{O(1)}$ then (4.1) follows trivially. Assuming that this is not the case, and that it is not the case that $N \ll W_2^{O(1)}$, we deduce from (5.11) that $\|b_2 a/q\|_{\mathbb{T}} < 1/q$. The only way this can happen is if $q \mid b_2$. It therefore suffices to assume that $q \mid b_2$, so that $q = O(1)$.

In the case of the third majorant, Lemma 5.4 and the trivial bound for the linear sum together give an upper bound of the form

$$(N/W_2)^{1/2} \left|\int_0^N e\left(b_1\left(\alpha - \tfrac{a}{q}\right) t\right) \mathrm{d}t\right| + O(W_1 N Q^2) \\ \ll N^{3/2} W_2^{-1/2} \max\left\{1, \left|b_1\left(\alpha - \tfrac{a}{q}\right)\right| N\right\}^{-1} + W_1 N Q^2.$$

This yields the major arc bound (4.1) with $K = O(1)$.

In the case of the fourth majorant, Lemmas 5.11 and 5.12 combine (with the trivial bound for the linear sum) to give an upper bound of the form

$$(N/W_2)\eta^{-1}\max\left\{1, \left|b_1\left(\alpha - \tfrac{a}{q}\right)\right| N\right\}^{-1} + N^{1/2}Q^2$$

From this we obtain the major arc bound (4.1) with $K = \eta^{-O(1)}$. □

Lemma 5.14 (Minor arc hypotheses)**.** *Let W_1 be an even positive integer and $\xi \in [W_1]$ with $\mathrm{hcf}(\xi, W_1) = 1$. Define $\nu = \nu_{W_1,\xi} : [N] \to [0,\infty)$ as in (5.1). Let W_2 be a positive integer, $\eta \in (0, 1/2]$ and define the interval*

$$I := \left[\eta(N/W_2)^{1/2}, (N/W_2)^{1/2}\right].$$

Fix non-zero integers $b_1, b_2 = O(1)$ and write $B := |b_1| + |b_2|$. Consider the following four majorants, mapping each $n \in [-BN, BN]$ to one of

$$\sum_{b_1x+b_2y=n} \nu(x)\nu(y), \qquad \sum_{b_1x+b_2W_2y^2=n} \nu(x)1_I(y),$$
$$\sum_{b_1x+b_2y=n} \nu(x)1_I(y), \qquad \sum_{b_1W_2x^2+b_2y=n} 1_I(x)1_I(y). \tag{5.12}$$

Then either $N \ll (W_1W_2)^{O(1)}$ or all four majorants satisfy the minor arc hypothesis (Definition 4.2).

Proof. By the convolution identity, the Fourier transform of each of the first three majorants is bounded in magnitude by $|\hat{\nu}(b_1\alpha)||I|$. The result then follows for these majorants using Lemma 5.8 and the fact that $0 < |b_1| \ll 1$.

Letting ν_2 denote the fourth majorant, suppose that $|\hat{\nu}_2(\alpha)| \geqslant \delta \|\nu_2\|_1$. We have $\|\nu_2\|_1 \gg N/W_2$, unless $N \ll W_2$. Hence by the convolution identity

$$\left|\sum_{x\in I} e(b_1\alpha W_2 x^2) \sum_{y\in I} e(b_2\alpha y)\right| \gg \delta N/W_2.$$

Thus both of the following estimates hold

$$\left|\sum_{x\in I} e(b_1\alpha W_2x^2)\right| \gg \delta\sqrt{N/W_2} \qquad \text{and} \qquad \left|\sum_{y\in I} e(b_2\alpha y)\right| \gg \delta\sqrt{N/W_2}. \tag{5.13}$$

Applying Weyl's inequality (Lemma 5.2) to the first sum in (5.13), we deduce the existence of $q_0 \ll \delta^{-O(1)}$ such that $\|q_0b_1W_2\alpha\|_{\mathbb{T}} \ll \delta^{-O(1)}W_2/N$. Dividing through by $q_0b_1W_2$ and cancelling common factors, it follows that there exist integers $1 \leqslant a \leqslant q \ll W_2\delta^{-O(1)}$ with $\mathrm{hcf}(a, q) = 1$ and such that $\left\|\alpha - \frac{a}{q}\right\|_{\mathbb{T}} \ll \delta^{-O(1)}/N$. We claim that $q \leqslant |b_2|$, hence completing our proof.

Applying the linear exponential sum estimate (Lemma 5.9) to the second sum in (5.13), we deduce that $\|b_2\alpha\|_{\mathbb{T}} \ll \delta^{-1}\sqrt{W_2/N}$, hence by the triangle inequality

$$\left\|\tfrac{b_2a}{q}\right\|_{\mathbb{T}} \ll \left\|\alpha - \tfrac{a}{q}\right\|_{\mathbb{T}} + \|b_2\alpha\|_{\mathbb{T}} \ll \frac{\delta^{-1}W_2^{1/2}}{N^{1/2}} + \frac{\delta^{-O(1)}}{N}.$$

If $q \nmid b_2$ then $\|b_2a/q\|_{\mathbb{T}} \geqslant 1/q$ and so either $q \gg \delta^{-1}\sqrt{N/W_2}$ or $q \gg \delta^{O(1)}N$. Each of these conclusions contradict our bound of $q \ll W_2\delta^{-O(1)}$, unless $N \ll (W_2/\delta)^{O(1)}$. The latter implies that $N \ll W_2^{O(1)}$ or $N \ll \delta^{-O(1)}$. If $N \ll \delta^{-O(1)}$ then the conclusion of the minor arc hypothesis is trivial. We may therefore assume that $q \mid b_2$, which certainly implies that $q \leqslant |b_2|$ (as required). □

Lemma 5.15 (Hua-type hypotheses)**.** *Let W_1 be an even positive integer and $\xi \in [W_1]$. Define $\nu = \nu_{W_1,\xi} : [N] \to [0,\infty)$ as in (5.1). Let W_2 be a positive integer, $\eta \in (0, 1/2]$ and define the interval*

$$I := \left[\eta(N/W_2)^{1/2}, (N/W_2)^{1/2}\right].$$

Fix non-zero integers $b_1, b_2 = O(1)$ and write $B := |b_1| + |b_2|$. Consider the following four majorants, mapping each $n \in [-BN, BN]$ to one of

$$\sum_{b_1x+b_2y=n} \nu(x)\nu(y), \qquad \sum_{b_1x+b_2W_2y^2=n} \nu(x)1_I(y),$$
$$\sum_{b_1x+b_2y=n} \nu(x)1_I(y), \qquad \sum_{b_1W_2x^2+b_2y=n} 1_I(x)1_I(y). \tag{5.14}$$

Then either $N \ll (W_1W_2)^{O(1/\varepsilon)}$ or all four majorants satisfy the Hua-type hypothesis (Definition 4.3) with exponent ε.

Proof. We observe that either $N \ll W_1^{O(1)}$ or we have the following estimates

$$\|\nu\|_1 \asymp N \quad \text{and} \quad \|\nu\|_\infty \ll \sqrt{NW_1} \tag{5.15}$$

We also observe the standard divisor-type estimate: for $n \in \mathbb{Z} \setminus \{0\}$ we have

$$\sum_{x^2-y^2=n} 1 \ll_\varepsilon |n|^\varepsilon. \tag{5.16}$$

We begin with the first majorant in (5.14). In this case, the square of the L^2-norm is equal to the count

$$\sum_{b_1n_1+b_2n_2=b_1n_3+b_2n_4} \nu(n_1)\nu(n_2)\nu(n_3)\nu(n_4).$$

The diagonal contribution to this count, when $n_1 = n_3$, is at most

$$\left(\sum_n \nu(n)^2\right)^2 \ll \|\nu\|_\infty^2 \|\nu\|_1^2 \ll W_1N^3 \asymp W_1 \|\nu\|_1^4 N^{-1}.$$

Using Cauchy–Schwarz and the divisor-bound (5.16), the non-diagonal count is given by

$$\sum_{0<|n|<BN} \left(\sum_{b_1(n_1-n_3)=n} \nu(n_1)\nu(n_3)\right)\left(\sum_{b_2(n_4-n_2)=n} \nu(n_2)\nu(n_4)\right)$$
$$\ll (W_1N)^2 \sum_{0<|n|<BN} \left(\sum_{(W_1x+\xi)^2-(W_1y+\xi)^2=2W_1n} 1\right)^2$$
$$\ll_\varepsilon (W_1N)^2N(W_1N)^\varepsilon \ll W_1^{O(1)} \|\nu\|_1^4 N^{\varepsilon-1}.$$

We conclude that for any $\varepsilon > 0$ our majorant's second moment has an upper bound of the form

$$O_\varepsilon\left(W_1^{O(1)} \|\nu\|_1^4 N^{\varepsilon-1}\right).$$

On assuming that it is not the case that $N \ll_\varepsilon W_1^{O(1/\varepsilon)}$, this can be replaced by an upper bound of the form

$$\|\nu\|_1^4 N^{2\varepsilon-1}.$$

This establishes the lemma for the first majorant.

We turn now to the second majorant in (5.14). In this case the square of the L^2-norm is equal to

$$\sum_{b_1n_1+b_2W_2y_1^2=b_1n_2+b_2W_2y_2^2} \nu(n_1)\nu(n_2)1_I(y_1)1_I(y_2).$$

Provided that $N \geqslant 2W_2$, the diagonal contribution to this count is at most

$$\sum_n \nu(n)^2 \sum_y 1_I(y) \ll \|\nu\|_\infty \|\nu\|_1 |I| \ll W_1 N^2 W_2^{-1/2} \ll W_1 W_2^{1/2} \|\nu\|_1^2 |I|^2 N^{-1}. \tag{5.17}$$

Using the divisor-bound (5.16), the non-diagonal count is given by

$$\sum_{0<|n|<BN} \sum_{b_1(n_1-n_2)=n} \nu(n_1)\nu(n_2) \sum_{b_2W_2(y_1^2-y_2^2)=n} 1_I(y_1)1_I(y_2) \ll_\varepsilon (W_1W_2)^{O(1)} N^{2+\varepsilon}$$
$$\ll (W_1W_2)^{O(1)} \|\nu\|_1^2 |I|^2 N^{\varepsilon-1}.$$

Using a similar argument to before, this establishes the result for the second majorant.

For the third majorant in (5.14), the diagonal contribution is the same as that in (5.17). The non-diagonal count is given by

$$\sum_{0<|n|<BN} \sum_{b_1(n_1-n_2)=n} \nu(n_1)\nu(n_2) \sum_{b_2(y_1-y_2)=n} 1_I(y_1)1_I(y_2).$$

Notice that if $y_1, y_2 \in I$ then $|y_1 - y_2| < (N/W_2)^{1/2}$. Hence the non-diagonal count is in fact equal to

$$\sum_{0<|n|<B(N/W_2)^{1/2}} \sum_{b_1(n_1-n_2)=n} \nu(n_1)\nu(n_2) \sum_{b_2(y_1-y_2)=n} 1_I(y_1)1_I(y_2)$$
$$\ll_\varepsilon (N/W_2)^{1/2} N W_1 N^\varepsilon (N/W_2)^{1/2} \ll (W_1W_2)^{O(1)} \|\nu\|_1^2 |I|^2 N^{\varepsilon-1}.$$

This establishes the result for the third majorant.

For the fourth majorant in (5.14), the diagonal contribution is given by

$$\left(\sum_y 1_I(y)\right)^2 = |I|^2 = W_2^{O(1)} |I|^4 N^{-1}.$$

The non-diagonal count is given by

$$\sum_{0<|n|<B(N/W_2)^{1/2}} \sum_{b_1W_2(x_1^2-x_2^2)=n} 1_I(y_1)1_I(y_2) \sum_{b_2(y_1-y_2)=n} 1_I(y_1)1_I(y_2) \ll_\varepsilon (N/W_2)^{1+\varepsilon}$$
$$\ll W_2^{O(1)} |I|^4 N^{-1}.$$

This establishes the result for the fourth majorant. □

6. Controlling the counting operator

The purpose of this section is to prove an analogue of the Fourier control lemma (Lemma 3.7) for the counting operator encountered in the quadratic counting theorem (Theorem 1.7).

Before embarking on this section the reader may wish to recall the definition of $\nu = \nu_{W,\xi}$ (Definition 5.1), as well as our notation for the Fourier transform (Definition 1.15) and quadratic Fourier transform (Definition 1.17).

Lemma 6.1 (Mixed restriction estimates)**.** *Let W_1 and W_2 be w-smooth positive integers such that W_1 is divisible by $2\prod_{p\leqslant w} p$. Given $\xi \in [W_1]$ with $\mathrm{hcf}(\xi, W_1) = 1$, define $\nu = \nu_{W_1,\xi} : [N] \to [0,\infty)$ as in (5.1). Given $\eta \in (0, 1/2]$, define the interval*

$$I := \left[\eta(N/W_2)^{1/2}, (N/W_2)^{1/2}\right].$$

Let $p > 2$ and fix non-zero integers $b_1, b_2 = O(1)$. Then either $N \ll_p (W_1W_2)^{O_p(1)}$ or, for any $f : [N] \to \mathbb{C}$ with $|f| \leqslant 1_{[N]} + \nu$ and any $B \subset [(N/W_2)^{1/2}]$, we have

$$\int_{\mathbb{T}} \left|\hat{f}(b_1\alpha)\hat{1}_B(b_2\alpha)\right|^p \mathrm{d}\alpha,\ \int_{\mathbb{T}} \left|\hat{f}(b_1\alpha)\tilde{1}_B(b_2W_2\alpha)\right|^p \mathrm{d}\alpha \ll_p N^{\frac{3p}{2}-1}W_2^{-\frac{p}{2}},$$

whilst

$$\int_{\mathbb{T}} \left|\tilde{1}_B(b_1W_2\alpha)\hat{1}_B(b_2\alpha)\right|^p \mathrm{d}\alpha \ll_{p,\eta} N^{p-1}W_2^{-p} \quad \text{and} \quad \int_{\mathbb{T}} \left|\hat{f}(\alpha)\right|^{2p} \mathrm{d}\alpha \ll_p N^{2p-1}.$$

Proof. Let us first suppose that $|f| \leqslant \nu$. Then the bounds follow from the abstract restriction estimate (Lemma 4.4) in conjunction with the verification of the major/minor/Hua-type hypotheses (Lemmas 5.14, 5.13, 5.15).

Next let us suppose that $|f| \leqslant 1_{[N]}$. We estimate $\left|\hat{1}_B\right|^p$ and $\left|\tilde{1}_B\right|^p$ using the trivial bound of $\ll (N/W_2)^{p/2}$. We estimate $\left|\hat{f}\right|^{p-2}$ using the trivial bound of N^{p-2}, and $\left|\hat{f}\right|^{2p-2}$ with N^{2p-2}. Finally we employ Parseval to give the bound

$$\int_{\mathbb{T}} \left|\hat{f}(b_1\alpha)\right|^2 \mathrm{d}\alpha = \sum_n |f(n)|^2 \leqslant N.$$

Combining these inequalities gives the claimed bounds.

Finally, we assume the general bound $|f| \leqslant 1_{[N]} + \nu$. Write $f = \theta|f|$, where $|\theta(n)| \leqslant 1$ for all n. Put $f_1 := \theta \min\left\{|f|, 1_{[N]}\right\}$ and $f_2 := f - f_1$. Then $f = f_1 + f_2$ with $|f_1| \leqslant 1_{[N]}$ and $|f_2| \leqslant \nu$. Applying the triangle inequality, the estimates now follow from our previous arguments. □

Lemma 6.2 (Fourier control)**.** *For each $i = 1, 2, 3$, let L_i denote a non-singular linear form in s_i variables with $s_1 \geqslant 2$, $s_1 + s_2 \geqslant 3$ and $s_1 + s_2 + s_3 \geqslant 5$ (we allow for $s_2 = 0$ or $s_3 = 0$). Let W_1 and W_2 be w-smooth positive integers such that W_1 is divisible by $2\prod_{p\leqslant w} p$. Given $\xi \in [W_1]$ with $\mathrm{hcf}(\xi, W_1) = 1$, define $\nu = \nu_{W_1,\xi} : [N] \to [0,\infty)$ as in (5.1). Given $\eta \in (0, 1/2]$, define the interval*

$$I := \left[\eta(N/W_2)^{1/2}, (N/W_2)^{1/2}\right].$$

Suppose that either $W_2 = 1$ or $s_3 > 0$. Then either $N \ll (W_1W_2)^{O(1)}$ or for any $f_1, \dots, f_{s_1} : \mathbb{Z} \to \mathbb{C}$, each satisfying $|f_i| \leqslant 1_{[N]} + \nu$, and any $B \subset I$ we have

$$\left| \sum_{L_1(x)=W_2L_2(y^2)+L_3(z)} f_1(x_1)\cdots f_{s_1}(x_{s_1})1_B(y_1)\cdots 1_B(y_{s_2})1_B(z_1)\cdots 1_{B_j}(z_{s_3}) \right|$$
$$\ll_\eta N^{s_1+\frac{1}{2}(s_2+s_3)-1} W_2^{-\frac{1}{2}(s_2+s_3)} \min_i \left(\frac{\|\hat{f}_i\|_\infty}{N} \right)^{1/10}.$$

Proof. Write
$$L_i(x) = c_1^{(i)}x_1 + \cdots + c_{s_i}^{(i)}x_{s_i}.$$

Case 1: $s_3 = 0$.

In this case our assumptions imply that $W_2 = 1$. The orthogonality relations then show that our counting operator is equal to

$$\sum_{L_1(x)=L_2(y^2)} f_1(x_1)\cdots f_{s_1}(x_{s_1})1_B(y_1)\cdots 1_B(y_{s_2})$$
$$= \int_{\mathbb{T}} \prod_i \hat{f}_i\left(c_i^{(1)}\alpha\right) \prod_j \tilde{1}_B\left(c_j^{(2)}\alpha\right) \mathrm{d}\alpha.$$

We apply Hölder's inequality to bound the Fourier integral by
$$\|\tilde{1}_B\|_{s_1+s_2}^{s_2} \|\hat{f}_i\|_\infty^{0.1} \|\hat{f}_i\|_{0.9(s_1+s_2)}^{0.9} \prod_{j\neq i} \|\hat{f}_j\|_{s_1+s_2}. \tag{6.1}$$

Since $s_1 + s_2 \geqslant 5$, Bourgain's restriction estimate [Bou89] gives that $\|\tilde{1}_B\|_{s_1+s_2} \ll N^{\frac{1}{2}-\frac{1}{s_1+s_2}}$. Since $0.9(s_1 + s_2) > 4$, Lemma 6.1 gives that
$$\|\hat{f}_i\|_{0.9(s_1+s_2)}^{0.9} \ll_\eta N^{0.9-\frac{1}{s_1+s_2}} \quad \text{and} \quad \|\hat{f}_j\|_{s_1+s_2} \ll_\eta N^{1-\frac{1}{s_1+s_2}}.$$

The claimed bound follows on incorporating these estimates into (6.1).

Case 2: $s_3 \geqslant 1$.

In this case we must assume that W_2 is arbitrary. The orthogonality relations show our counting operator equals

$$\sum_{L_1(x)=W_2L_2(y^2)+L_3(z)} f_1(x_1)\cdots f_{s_1}(x_{s_1})1_B(y_1)\cdots 1_B(y_{s_2})1_B(z_1)\cdots 1_B(z_{s_3})$$
$$= \int_{\mathbb{T}} \prod_i \hat{f}_i\left(c_i^{(1)}\alpha\right) \prod_j \tilde{1}_B\left(W_2c_j^{(2)}\alpha\right) \prod_k \hat{1}_B\left(c_k^{(3)}\alpha\right) \mathrm{d}\alpha. \tag{6.2}$$

We break into further subcases. Notice that our assumptions that $s_1 \geqslant 2$, $s_1+s_2 \geqslant 3$ and $s_1 + s_2 + s_3 \geqslant 5$ imply that we are in one of the following five situations.

Case 2a: $s_1 \geqslant 4$, $s_3 \geqslant 1$.

Fix distinct $i, j \in \{1, \dots, s_1\}$. Applying the trivial estimate to $\tilde{1}_B$ and all but one copy of $\hat{1}_B$, Hölder's inequality shows that the Fourier integral (6.2) is at most

$$(N/W_2)^{\frac{1}{2}(s_2+s_3-1)} \|\hat{f}_i\|_\infty^{0.1} \|\hat{f}_i\|_{0.9(s_1+1)}^{0.9} \|\hat{f}_j(c_j^{(1)}\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_{(s_1+1)/2} \prod_{k\notin\{i,j\}} \|\hat{f}_k\|_{(s_1+1)}.$$

Since $0.9(s_1+1) \geqslant 4.5$ and $(s_1+1)/2 \geqslant 2.5$, Lemma 6.1 gives that

$$\|\hat{f}_i\|_{0.9(s_1+1)}^{0.9} \ll_\eta N^{0.9-\frac{1}{s_1+1}}, \quad \|\hat{f}_k\|_{s_1+1} \ll_\eta N^{1-\frac{1}{s_1+1}},$$
$$\|\hat{f}_j(c_j^{(1)}\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_{(s_1+1)/2} \ll_\eta N^{\frac{3}{2}-\frac{2}{s_1+1}} W_2^{-1/2}.$$

The claimed bound follows.

Case 2b: $s_1 = 3$, $s_2 = 0$, $s_3 \geqslant 2$.

Let $\{i, j, k\} = \{1, 2, 3\}$. Applying the trivial estimate to all but two copies of $\hat{1}_B$, Hölder's inequality shows that the Fourier integral (6.2) is at most

$$(N/W_2)^{\frac{1}{2}(s_3-2)} \|\hat{f}_i\|_\infty^{0.1} \|\hat{f}_i\|_{4.5}^{0.9} \|\hat{f}_j(c_j^{(1)}\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_{2.5} \|\hat{f}_k(c_k^{(1)}\alpha)\hat{1}_B(c_2^{(3)}\alpha)\|_{2.5}.$$

The claimed bound follows again on employing Lemma 6.1.

Case 2c: $s_1 = 3$, $s_2 \geqslant 1$, $s_3 \geqslant 1$.

Let $\{i, j, k\} = \{1, 2, 3\}$. Applying the trivial estimate to all but one copy of $\tilde{1}_B$ and all but one copy of $\hat{1}_B$, Hölder's inequality shows that the Fourier integral (6.2) is at most

$$(N/W_2)^{\frac{1}{2}(s_2+s_3-2)} \|\hat{f}_i\|_\infty^{0.1} \|\hat{f}_i\|_{4.5}^{0.9} \|\hat{f}_j(c_j^{(1)}\alpha)\tilde{1}_B(W_2 c_1^{(2)}\alpha)\|_{2.5} \|\hat{f}_k(c_k^{(1)}\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_{2.5}.$$

The claimed bound follows from Lemma 6.1.

Case 2d: $s_1 = 2$, $s_2 = 1$, $s_3 \geqslant 2$.

Let $\{i, j\} = \{1, 2\}$. We apply the trivial estimate to all but two copies of $\hat{1}_B$. Hölder's inequality then shows that the Fourier integral (6.2) is at most

$$(N/W_2)^{\frac{1}{2}(s_3-2)} \|\hat{f}_i\|_\infty^{0.1} \|\hat{f}_i\|_{4.5}^{0.9} \|\tilde{1}_B(W_2 c_1^{(2)}\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_{2.5} \|\hat{f}_j(c_j^{(1)}\alpha)\hat{1}_B(c_2^{(3)}\alpha)\|_{2.5}.$$

The claimed bound then follows from Lemma 6.1.

Case 2e: $s_1 = 2$, $s_2 \geqslant 2$, $s_3 \geqslant 1$.

Let $\{i, j\} = \{1, 2\}$. We apply the trivial estimate to all but two copies of $\tilde{1}_B$ and all but one copy of $\hat{1}_B$. Hölder's inequality then shows that the Fourier integral (6.2) is at most

$$(N/W_2)^{\frac{1}{2}(s_3-2)} \|\hat{f}_i\|_\infty^{0.1} \|\hat{f}_i\|_{4.5}^{0.9} \|\tilde{1}_B(W_2 c_1^{(2)}\alpha)\hat{1}_B(c_1^{(3)}\alpha)\|_{2.5} \|\hat{f}_j(c_j^{(1)}\alpha)\tilde{1}_B(W_2 c_2^{(2)}\alpha)\|_{2.5}.$$

The claimed bound follows from Lemma 6.1. □

Lemma 6.3. *Let $a_1, a_2, b_1, b_2 \in \mathbb{Z} \setminus \{0\}$ and let $\eta \in (0, 1)$. Given functions $f : (\eta N, N] \to [-1, 1]$ and $g : [N] \to [-1, 1]$ we have the bound*

$$\int_{\mathbb{T}} \left|\tilde{f}(a_1\alpha)\tilde{f}(a_2\alpha)\hat{g}(b_1\alpha)\hat{g}(b_2\alpha)\right| \mathrm{d}\alpha \ll_{b_1,b_2} \eta^{-1} N^2.$$

Proof. By Cauchy–Schwarz it suffices to bound an integral of the form

$$\int_{\mathbb{T}} \left|\tilde{f}(a\alpha)\hat{g}(b\alpha)\right|^2 \mathrm{d}\alpha$$

for some non-zero integers a, b. By orthogonality, this is at most the number of solutions to the equation

$$a(x_1^2 - x_2^2) = b(y_1 - y_2), \qquad (x_i \in (\eta N, N],\ y_j \in [N]) . \tag{6.3}$$

The diagonal contribution (when $y_1 = y_2$) yields at most N^2 solutions. Fix distinct $y_1, y_2 \in [N]$. Then any solution (x_1, x_2) to (6.3) satisfies

$$|x_1 - x_2| = \frac{|b||y_1 - y_2|}{|a|(x_1 + x_2)} \leqslant b\eta^{-1}.$$

The estimate follows. □

Lemma 6.4 (L^1 control). *Let $a_1, \dots, a_r \in \mathbb{Z} \setminus \{0\}$, $b_1, \dots, b_s \in \mathbb{Z} \setminus \{0\}$ and $c_1, \dots, c_t \in \mathbb{Z} \setminus \{0\}$. Suppose that*

$$r \geqslant 2, \quad r + s \geqslant 3, \quad s + t \geqslant 1, \quad r + s + t \geqslant 5.$$

Then for any $B \subset [N]$ and $\eta \in (0, 1)$ we have

$$\sum_{\sum_i a_i x_i^2 = \sum_j b_j y_j^2 + \sum_k c_k z_k} \prod_i 1_{(\eta N, N]}(x_i) \prod_j 1_B(y_j) \prod_k 1_B(z_k)$$
$$\ll_{c_i} \eta^{-O(1)} N^{r+s+t-2} \left(\frac{|B|}{N}\right)^{1/2}. \tag{6.4}$$

Proof. The left-hand side of (6.4) can be written as the Fourier integral

$$\int_{\mathbb{T}} \prod_i \tilde{1}_{(\eta N, N]}(a_i\alpha) \prod_j \tilde{1}_B(b_j\alpha) \prod_k \hat{1}_B(c_k\alpha) \mathrm{d}\alpha.$$

If $r + s \geqslant 5$ then (6.4) follows from extracting $|B|^{1/2}$ from the Fourier integral, then applying Hölder's inequality and the estimates

$$\int_{\mathbb{T}} \Bigg|\sum_{x \in (\eta N, N]} e\left(\alpha x^2\right)\Bigg|^{4.5} \mathrm{d}\alpha, \ \int_{\mathbb{T}} \Bigg|\sum_{x \in B} e\left(\alpha x^2\right)\Bigg|^{4.5} \mathrm{d}\alpha \ll N^{2.5}.$$

These bounds are a consequence of [Bou89].

Let us therefore suppose that $r + s \leqslant 4$, in which case we must have $t \geqslant 1$. We divide into two cases.

<u>**Case 1:** $t \geqslant 2$**:**</u>

Since $r + s \geqslant 3$, our Fourier integral contains at least three quadratic exponential sums, at least two of which are equal to $\tilde{1}_{(\eta N, N]}$ (since $r \geqslant 2$). Employing the bounds $1_{(\eta N, N]} \leqslant 1_{[N]}$ or $1_B \leqslant 1_{[N]}$ on the physical side, we may assume that our third quadratic exponential sum is equal to $\tilde{1}_{[N]}$. Then using the orthogonality relations and Hölder's inequality, we can bound the left-hand side of (6.4) by

$$N^{r+s+t-5}\|\hat{1}_B\|_\infty^{\frac12}\left(\int_{\mathbb{T}}\left|\tilde{1}_{(\eta N,N]}(a_1\alpha)\hat{1}_B(c_1\alpha)\right|^2\mathrm{d}\alpha\right)^{1/4}$$
$$\left(\int_{\mathbb{T}}\left|\tilde{1}_{(\eta N,N]}(a_2\alpha)\hat{1}_B(c_2\alpha)\right|^2\mathrm{d}\alpha\right)^{1/2}\left(\int_{\mathbb{T}}\left|\tilde{1}_{(\eta N,N]}(\alpha)\right|^6\mathrm{d}\alpha\right)^{1/12}\left(\int_{\mathbb{T}}\left|\tilde{1}_{[N]}(\alpha)\right|^6\mathrm{d}\alpha\right)^{1/6}.$$

The estimate now follows from Lemma 6.3 and Bourgain's restriction estimate [Bou89].

Case 2: $t = 1$:

In this case our Fourier integral contains at least four quadratic exponential sums, at least one of which equals $\tilde{1}_{(\eta N,N]}$. Proceeding as in Case 1, the left-hand side of (6.4) can be bounded by

$$N^{r+s+t-5}\|\hat{1}_B\|_\infty^{\frac12}\left(\int_{\mathbb{T}}\left|\tilde{1}_{(\eta N,N]}(a_1\alpha)\hat{1}_B(c_1\alpha)\right|^2\mathrm{d}\alpha\right)^{1/4}$$
$$\left(\int_{\mathbb{T}}\left|\tilde{1}_{(\eta N,N]}(\alpha)\right|^{14/3}\mathrm{d}\alpha\right)^{3/28}\left(\int_{\mathbb{T}}\left|\tilde{1}_{[N]}(\alpha)\right|^{14/3}\mathrm{d}\alpha\right)^{9/14}.$$

Again the estimate follows from Lemma 6.3 and Bourgain's restriction estimate [Bou89]. □

7. A quadratic density result

The purpose of this section is to prove the following.

Theorem 7.1 (Density–colouring result)**.** *For each $i = 1, 2, 3$, let L_i denote a non-singular linear form in s_i variables with $s_1 \geqslant 2$, $s_1 + s_2 \geqslant 3$ and $s_1 + s_2 + s_3 \geqslant 5$ (we allow for $s_2 = 0$ or $s_3 = 0$). Suppose that $L_1(1, \dots, 1) = 0$. Let $\delta > 0$ and let r be a positive integer. Then either $N \ll_{\delta,r} 1$ or the following holds. For any sets of integers $A_1, \dots, A_r \subset [N]$ each satisfying $|A_i| \geqslant \delta N$ and for any r-colouring $B_1 \cup \dots \cup B_r = [N]$ there exists $B \in \{B_1, \dots, B_r\}$ such that for all $A \in \{A_1, \dots, A_r\}$ we have*

$$\sum_{L_1(x^2)=L_2(y^2)+L_3(z)}\prod_i 1_A(x_i)\prod_j 1_B(y_j)\prod_k 1_B(z_k) \gg_{\delta,r} N^{s_1+s_2+s_3-2}.$$

Let $\mathcal{R}_w(N)$ denote the set of w-rough numbers in $[N]$, that is those integers all of whose prime divisors exceed w. We have the following disjoint partition

$$[N] = \bigcup_{\zeta \text{ is } w\text{-smooth}} \zeta\cdot\mathcal{R}_w(N/\zeta).$$

For each i we would like to find ζ_i which is not too large and satisfies

$$|A_i \cap (\zeta_i\cdot\mathcal{R}_w(N/\zeta_i))| \geqslant \tfrac{\delta}{2}\left|\mathcal{R}_w(N/\zeta_i)\right|. \tag{7.1}$$

By [CLP, Lemma A.3] there are at most $10^w N M^{-1/2}$ elements of $[N]$ divisible by a w-smooth number greater than M. It follows that for each A_i there exists a w-smooth number ζ_i satisfying

$$\zeta_i \ll \delta^{-O(1)}\exp\left(O(w)\right)$$

and such that (7.1) holds.

Define
$$W := 4\zeta_1^2 \cdots \zeta_r^2 \prod_{p \leqslant w} p \qquad \text{and} \qquad W_i := \frac{W}{2\zeta_i^2}. \tag{7.2}$$

Since W_i is w-smooth and divisible by the primorial $\prod_{p\leqslant w} p$, we can partition $\mathcal{R}_w(N/\zeta_i)$ into congruence classes
$$\mathcal{R}_w(N/\zeta_i) \cap (\xi \bmod W_i) = (W_i \cdot \mathbb{Z} + \xi) \cap [N/\zeta_i], \qquad (\xi \in (\mathbb{Z}/W_i\mathbb{Z})^\times).$$

By the pigeon-hole principle, there exists $\xi_i \in (\mathbb{Z}/W_i\mathbb{Z})^\times$ such that
$$|A_i \cap (\zeta_i \cdot ((W_i \cdot \mathbb{Z} + \xi_i) \cap [N/\zeta_i]))| \geqslant \tfrac{\delta}{2} \left|(W_i \cdot \mathbb{Z} + \xi_i) \cap [N/\zeta_i]\right|.$$

It follows that there exists a set A_i' of integers such that for every $x \in A_i'$ we have $\zeta_i(W_i x + \xi_i) \in A_i$, and moreover we can ensure that
$$A_i' \subset \left(\frac{\delta N}{4\zeta_i W_i}, \frac{N - \zeta_i \xi_i}{\zeta_i W_i}\right] \quad \text{and} \quad |A_i'| \geqslant \frac{\delta N}{4\zeta_i W_i} - O(1). \tag{7.3}$$

We define a colouring of $\left[\frac{N}{W}\right]$ by setting
$$B_j' := \{x \in \mathbb{N} : Wx \in B_j\}.$$

It follows that
$$\sum_{L_1(x^2) = L_2(y^2) + L_3(z)} \prod_l 1_{A_i}(x_l) \prod_m 1_{B_j}(y_m) \prod_n 1_{B_j}(z_n) \geqslant \sum_{L_1(\frac{1}{2}W_i x^2 + \xi_i x) = W L_2(y^2) + L_3(z)} \prod_l 1_{A_i'}(x_l) \prod_m 1_{B_j'}(y_m) \prod_n 1_{B_j'}(z_n). \tag{7.4}$$

Set
$$X := \frac{N^2}{W},$$
and let $\nu_i := \nu_{W_i, \xi_i} : [X] \to [0, \infty)$ be as in (5.1). The containment in (7.3) ensures that for every $x \in A_i'$ we have
$$\frac{\delta N}{\zeta_i} \ll \nu_i(\tfrac{1}{2}W_i x^2 + \xi_i x) \leqslant \frac{N}{\zeta_i}. \tag{7.5}$$

Define
$$f_i(n) := \begin{cases} \nu_i(n) & \text{if } n = \frac{1}{2}W_i x^2 + \xi_i x \text{ for some } x \in A_i', \\ 0 & \text{otherwise.} \end{cases}$$

Then we have that
$$\sum_{L_1(\frac{1}{2}W_i x^2 + \xi_i x) = W L_2(y^2) + L_3(z)} \prod_l 1_{A_i'}(x_l) \prod_m 1_{B_j'}(y_m) \prod_n 1_{B_j'}(z_n) \geqslant \left(\frac{\zeta_i}{N}\right)^{s_1} \sum_{L_1(n) = W L_2(y^2) + L_3(z)} \prod_l f_i(n_l) \prod_m 1_{B_j'}(y_m) \prod_n 1_{B_j'}(z_n). \tag{7.6}$$

Notice that (7.3) and (7.5) give
$$\sum_{n \in [X]} f_i(n) \gg \frac{\delta N}{\zeta_i} \left(\frac{\delta N}{\zeta_i W_i} - O(1)\right),$$

so that either $N \ll_{\delta,r,w} 1$ or

$$\sum_{n\in[X]} f_i(n) \gg \delta^2 X.$$

Using Lemma 5.10 and the dense model lemma recorded in [Pre17, Theorem 5.1], there exists $0 \leqslant g_i \leqslant 1_{[X]}$ satisfying

$$\left\|\hat{f}_i - \hat{g}_i\right\|_\infty \ll (\log w)^{-3/2} X. \tag{7.7}$$

It follows that either $w \ll_\delta 1$ or, on comparing Fourier coefficients at 0, we deduce that $\sum_{x\in[X]} g_i(x) \gg \delta^2 X$. Thresholding, define

$$\tilde{A}_i := \left\{x \in [X] : g_i(x) \geqslant c\delta^2\right\},$$

with c a small positive absolute constant. The popularity principle [TV06, Ex.1.1.4] shows that $\tilde{A}_i \gg \delta^2 X$. Hence by Theorem 3.1 there exists $\eta \gg_{\delta,r} 1$ and there exists

$$\tilde{B}_j := B'_j \cap [\eta N/W, N/W]$$

such that either $N \ll_{\delta,r,w} 1$ or for each $i = 1, \ldots, r$ we have

$$\sum_{L_1(n)=WL_2(y^2)+L_3(z)} \prod_l 1_{\tilde{A}_i}(n_l) \prod_m 1_{\tilde{B}_j}(y_m) \prod_n 1_{\tilde{B}_j}(z_n) \geqslant \eta X^{s_1+\frac{1}{2}(s_2+s_3)-1} W^{-\frac{1}{2}(s_2+s_3)}.$$

Using our lower bound for g_i on $\tilde{A}_i$ we deduce that

$$\sum_{L_1(n)=WL_2(y^2)+L_3(z)} \prod_l g_i(n_l) \prod_m 1_{\tilde{B}_j}(y_m) \prod_n 1_{\tilde{B}_j}(z_n) \gg_{\delta,r} X^{s_1+\frac{1}{2}(s_2+s_3)-1} W^{-\frac{1}{2}(s_2+s_3)}.$$

By a telescoping identity there exist functions $h_1, \ldots, h_{s_1} \in \{f_i, g_i, f_i - g_i\}$, at least one of which is equal to $f_i - g_i$, such that

$$\left|\sum_{L_1(n)=WL_2(y^2)+L_3(z)} \left(\prod_l g_i(n_l) - \prod_l f_i(n_l)\right) \prod_m 1_{\tilde{B}_j}(y_m) \prod_n 1_{\tilde{B}_j}(z_n)\right| \ll \left|\sum_{L_1(n)=WL_2(y^2)+L_3(z)} \prod_l h_l(n_l) \prod_m 1_{\tilde{B}_j}(y_m) \prod_n 1_{\tilde{B}_j}(z_n)\right|.$$

By Lemma 6.2 and (7.7), either $N \ll_{\delta,r,w} 1$ or the latter quantity is at most

$$\ll_{\delta,r} \left(\frac{\left\|\hat{f}_i - \hat{g}_i\right\|_\infty}{X}\right)^{\frac{1}{10}} X^{s_1+\frac{1}{2}(s_2+s_3)-1} W^{-\frac{1}{2}(s_2+s_3)} \ll X^{s_1+\frac{1}{2}(s_2+s_3)-1} W^{-\frac{1}{2}(s_2+s_3)} \log^{-3/20} w.$$

It follows that either $w \ll_{\delta,r} 1$ or that

$$\sum_{L_1(n)=WL_2(y^2)+L_3(z)} \prod_l f_i(n_l) \prod_m 1_{\tilde{B}_j}(y_m) \prod_n 1_{\tilde{B}_j}(z_n) \gg_{\delta,r} X^{s_1+\frac{1}{2}(s_2+s_3)-1} W^{-\frac{1}{2}(s_2+s_3)}.$$

Taking w sufficiently large in terms of δ and r, we deduce that either $N \ll_{\delta,r} 1$ or, on recalling (7.4) and (7.6), we have

$$\sum_{L_1(x^2)=L_2(y^2)+L_3(z)} \prod_l 1_{A_i}(x_l) \prod_m 1_{B_j}(y_m) \prod_n 1_{B_j}(z_n) \gg_{\delta,r} \left(\frac{\zeta_i}{N}\right)^{s_1} X^{s_1+\frac{1}{2}(s_2+s_3)-1} W^{-\frac{1}{2}(s_2+s_3)} \gg_{\delta,r} N^{s_1+s_2+s_3-2}.$$

This completes the proof of Theorem 7.1.

8. Deduction of colouring results from density results

8.1. **When the linear form satisfies Rado's criterion.** The purpose of this section is to prove the following strengthening of Theorem 1.4. To streamline notation, we suppress the dependence of implicit constants on the coefficients a_i and b_j.

Theorem 8.1. *Let $a_1, \dots, a_s, b_1, \dots, b_t \in \mathbb{Z} \setminus \{0\}$ with $s, t \geqslant 1$ and suppose that there exists $S \neq \emptyset$ such that $\sum_{i \in S} a_i = 0$. For any positive integers r and N, either $N \ll_r 1$ or for any colouring $C_1 \cup \dots \cup C_r = [N]$ there exists $1 \leqslant n \leqslant r$, a colour class C_j and an interval I of length $N^{1/2^{n-1}}$ such that on setting $M := N^{1/2^n}$ we have*

$$\sum_{a_1x_1+\dots+a_sx_s=b_1y_1^2+\dots+b_ty_t^2} \prod_{i\in S} 1_{C_j\cap I}(x_i) \prod_{i\notin S} 1_{C_j\cap[M]}(x_i) \prod_{i=1}^{t} 1_{C_j\cap[M]}(y_i) \gg_r M^{|S|+s+t-2}. \tag{8.1}$$

The utility of this result over Theorem 1.4 is that it can be used to show that non-trivial monochromatic solutions exist, given any sensible notion of 'trivial'. For if the only monochromatic solutions to our equation are trivial, then the left-hand side of (8.1) should[7] have order $o(M^{|S|+s+t-2})$, which yields a contradiction if N is sufficiently large in terms of r.

Proof of Theorem 8.1. Re-labelling variables, we can write our equation in the form

$$L_1(x) = L_2(y^2) + L_3(z),$$

where the L_i are non-singular linear forms in s_i variables satisfying $s_1 + s_3 = s$, $s_1 = |S|$, $s_2 = t \geqslant 1$ and $L_1(1, \dots, 1) = 0$. In particular, the latter ensures that $s_1 \geqslant 2$, and so $s_1 + s_2 \geqslant 3$. It follows that the conditions of Theorem 3.1 are met with $W = 1$. Let $\eta(\delta, r)$ denote the parameter appearing in this theorem. A little thought shows that this quantity is increasing with δ and $1/r$, and redefining if necessary, we may assume that $\eta(\delta, r) \leqslant \min\{\delta, r^{-1}\}$. Set

$$\delta_n := \begin{cases} 1/r & \text{when } n = 0; \\ \frac{1}{2}\eta(\frac{1}{2}\delta_{n-1}, r) & \text{otherwise.} \end{cases} \tag{8.2}$$

[7]For instance, any algebraic notion of 'trivial' is likely to deliver a power saving in this estimate.

Let us say that a colour class C_i is *good at scale* n if

$$\left|C_i \cap \left(N^{1/2^{n+1}}, N^{1/2^n}\right]\right| \geqslant \delta_n \left|\mathbb{Z} \cap \left(N^{1/2^{n+1}}, N^{1/2^n}\right]\right|.$$

We claim that there exists $1 \leqslant n \leqslant r$ such that if any C_i is good at scale n then it is also good at scale $m = m(i)$ for some $0 \leqslant m < n$.

If the claim does not hold, then on defining

$$S_n := \{i \in [r] : C_i \text{ is good at scale } n\},$$

we have a chain of strictly increasing subsets

$$\emptyset \neq S_0 \subsetneq (S_0 \cup S_1) \subsetneq \cdots \subsetneq (S_0 \cup \cdots \cup S_r),$$

the last of which must have size at least $r+1$. This contradicts the fact that every element in this chain is a subset of $\{1, 2, \ldots, r\}$.

Given n satisfying our claim, each colour class C_i satisfies the implication

$$\left|C_i \cap \left(N^{1/2^{n+1}}, N^{1/2^n}\right]\right| \geqslant \delta_n \left|\mathbb{Z} \cap \left(N^{1/2^{n+1}}, N^{1/2^n}\right]\right| \quad \Longrightarrow$$
$$\exists m = m(i) < n \text{ with } \left|C_i \cap \left(N^{1/2^{m+1}}, N^{1/2^m}\right]\right| \geqslant \delta_m \left|\mathbb{Z} \cap \left(N^{1/2^{m+1}}, N^{1/2^m}\right]\right|. \tag{8.3}$$

Fixing $i \in S_n$, let $m(i) = m$ be such that $m < n$ and $i \in S_m$. We can partition $(N^{1/2^{m+1}}, N^{1/2^m}]$ into consecutive half-open intervals of integers, all of cardinality at most $N^{1/2^{n-1}}$. In this manner, provided that N is sufficiently large in terms of r, the pigeonhole-principle yields an interval of integers I_i satisfying

$$N^{1/2^{n-1}} \geqslant |I_i| \geqslant |I_i \cap C_i| \geqslant \tfrac{1}{2}\delta_m N^{1/2^{n-1}} \geqslant \tfrac{1}{2}\delta_{n-1} N^{1/2^{n-1}}.$$

Letting $t_i + 1$ denote the smallest integer in I_i, define the set

$$A_i := \left\{x \in [N^{1/2^{n-1}}] : x + t_i \in C_i\right\}.$$

Then $A_i \subset [N^{1/2^{n-1}}]$ and $|A_i| \geqslant \frac{1}{2}\delta_{n-1} N^{1/2^{n-1}}$ for all $i \in S_n$.

Notice that Theorem 3.1 remains valid if there are less than r sets A_i of density δ (simply define new sets A_i to all equal A_1). Applying this result, we deduce that there exists $\tilde{C}_j := C_j \cap \left[N^{1/2^n}\right]$ such that for all A_i with $i \in S_n$ we have

$$\sum_{L_1(x) = L_2(y^2) + L_3(z)} 1_{A_i}(x_1) \cdots 1_{A_i}(x_{s_1}) 1_{\tilde{C}_j}(y_1) \ldots 1_{\tilde{C}_j}(y_{s_2}) 1_{\tilde{C}_j}(z_1) \ldots 1_{\tilde{C}_j}(z_{s_3})$$
$$\geqslant \eta(\tfrac{1}{2}\delta_{n-1}, r) N^{(2s_1+s_2+s_3-2)/2^n}. \tag{8.4}$$

Since $s_1, s_2 \geqslant 1$ we have the estimate

$$\sum_{L_1(x) = L_2(y^2) + L_3(z)} 1_{A_i}(x_1) \cdots 1_{A_i}(x_{s_1}) 1_{\tilde{C}_j}(y_1) \ldots 1_{\tilde{C}_j}(y_{s_2}) 1_{\tilde{C}_j}(z_1) \ldots 1_{\tilde{C}_j}(z_{s_3})$$
$$\leqslant |\tilde{C}_j| N^{(2s_1+s_2+s_3-3)/2^n}.$$

Therefore

$$|C_j \cap (N^{1/2^{n+1}}, N^{1/2^n}]| \geqslant \eta(\tfrac{1}{2}\delta_{n-1}, r) N^{1/2^n} - N^{1/2^{n+1}}.$$

Hence, provided that N is sufficiently large in terms of r, we have

$$|C_j \cap (N^{1/2^{n+1}}, N^{1/2^n}]| \geqslant \tfrac{1}{2}\eta(\tfrac{1}{2}\delta_{n-1}, r) |\mathbb{Z} \cap (N^{1/2^{n+1}}, N^{1/2^n}]|.$$

As $\delta_n = \frac{1}{2}\eta(\frac{1}{2}\delta_{n-1}, r)$, we conclude that $j \in S_n$, so we may take $i := j$ in (8.4), completing the proof of the theorem. $\square$

8.2. **When the quadratic form satisfies Rado's criterion.** The purpose of this subsection is to prove Theorem 1.7. Again, we suppress dependence of implicit constants on the coefficients a_i, b_j and the number of variables s, t.

Proof of Theorem 1.7. Re-labelling variables, we can write our equation in the form

$$L_1(x^2) = L_2(y^2) + L_3(z),$$

where the L_i are non-singular linear forms in s_i variables satisfying $s_1 + s_2 = s \geqslant 3$, $s_1 = |I|$, $s_3 = t$ and $L_1(1, \dots, 1) = 0$. In particular, the latter ensures that $s_1 \geqslant 2$. We note that we may assume that $s_2 + s_3 \geqslant 1$, for otherwise Theorem 7.1 implies that for any $A \subset [N]$ with $|A| \geqslant \delta N$ we have

$$\sum_{L_1(x^2)=0} \prod_l 1_A(x_l) \gg_\delta N^{s_1-2}.$$

This yields Theorem 1.7 since every r-colouring has a colour class of density at least $1/r$.

Under the assumption that $s_2 + s_3 \geqslant 1$, let $C = O(1)$ denote the implicit constant appearing in Lemma 6.4, so that for any $B \subset [N]$ and $\eta \in (0, 1)$ we have the bound

$$\sum_{L_1(x^2)=L_2(y^2)+L_3(z)} \prod_l 1_{(\eta N, N]}(x_l) \prod_m 1_B(y_m) \prod_n 1_B(z_n) \leqslant C\eta^{-C} N^{s_1+s_2+s_3-2}(|B|/N)^{1/2}. \quad (8.5)$$

Let $c_0(\delta, r)$ denote the implicit constant occurring in the conclusion of Theorem 7.1. Clearly this quantity is increasing with δ and r^{-1}, and we may assume that $c_0(\delta, r) \leqslant \min\{\delta, r^{-1}\}$.

Set

$$\delta_n := \begin{cases} 1/r & \text{if } n = 1; \\ \left(\frac{c_0(\delta_{n-1}/2, r)}{C(\delta_{n-1}/2)^{-C}}\right)^2 & \text{otherwise.} \end{cases}$$

Define

$$\epsilon_n(i) := \begin{cases} 1 & \text{if } |C_i| \geqslant \delta_n N; \\ 0 & \text{otherwise.} \end{cases}$$

Since $\delta_{n+1} \leqslant \delta_n$, the sequence $\epsilon_n = (\epsilon_n(1), \dots, \epsilon_n(r)) \in \{0,1\}^r \setminus \{0\}$ is monotone increasing in each coordinate as n increases. It follows that this sequence cannot be strictly increasing if it has length at least $r + 1$. Hence there exists $1 \leqslant n \leqslant r$ for which $\epsilon_n = \epsilon_{n+1}$. In particular, for any i we have the implication

$$|C_i| \geqslant \delta_{n+1} N \quad \Longrightarrow \quad |C_i| \geqslant \delta_n N. \quad (8.6)$$

For each C_i satisfying $|C_i| \geqslant \delta_n N$ we have

$$|C_i \cap (\tfrac{1}{2}\delta_n N, N]| \geqslant \tfrac{1}{2}\delta_n N.$$

Notice that Theorem 7.1 remains valid if there are less than r sets A_i of density δ (simply define new sets A_i to all equal A_1). We may therefore apply Theorem 7.1, taking our dense sets to be those $C_i \cap (\frac{1}{2}\delta_n N, N]$ for which $|C_i| \geqslant \delta_n N$. We thereby deduce that there exists C_j such that for all C_i satisfying $|C_i| \geqslant \delta_n N$ we have

$$\sum_{L_1(x^2)=L_2(y^2)+L_3(z)} \prod_l 1_{C_i\cap(\frac{1}{2}\delta_n N,N]}(x_l) \prod_m 1_{C_j}(y_m) \prod_n 1_{C_j}(z_n) \geqslant c_0(\tfrac{1}{2}\delta_n, r)N^{s_1+s_2+s_3-2}. \quad (8.7)$$

Applying (8.5), we conclude that

$$C(\delta_n/2)^{-C}\left(|C_j|/N\right)^{1/2} \geqslant c_0(\delta_n/2, r).$$

By our construction of the sequence δ_n it follows that $|C_j| \geqslant \delta_{n+1}N$, hence by (8.6) we conclude that $|C_j| \geqslant \delta_n N$. We may therefore take $i := j$ in (8.7), completing the proof of the theorem. □

9. The Moreira–Lindqvist argument

In this section we complete our characterisation of when equation (1.7) is partition regular (Theorem 1.10). The methods we employ to prove Theorem 1.7 do not, at present, succeed for all of the equations covered by Theorem 1.10. We begin this section by adapting an argument of Moreira to cover those equations of the form (1.7) for which the quadratic coefficients sum to zero, but for whom the number of variables is not sufficient for us to employ Theorem 1.7. The adaptation of Moreira's argument was explained to the author by Sofia Lindqvist. We begin by using this argument to prove Theorem 1.13, where the idea is perhaps more transparent.

Proof of Theorem 1.13. We first observe that Hindman's conjecture (Conjecture 1.12) implies the existence of infinitely many monochromatic tuples of the form $(x, y, x + y, xy)$. For given a finite list of such tuples, all monochromatic under the same colour, one can introduce finitely many new colours each attached to the x appearing in a tuple. Re-applying Hindman's conjecture, one obtains a monochromatic configuration under this new colouring, and since the new colour classes introduced are all singletons (and the configuration is not), the configuration is monochromatic under the original colouring (and distinct from each tuple in the list).

Given an r-colouring $c : \mathbb{N} \to [r]$ define a new colouring $\tilde{c}$ by giving all odd numbers the colour $r + 1$ and, if n is even, then it receives the colour $c(n/2)$. Assuming Conjecture 1.12, there exists infinitely many $\tilde{c}$-monochromatic tuples of the form $(x,\ y,\ x + y,\ xy)$. Since all elements of this tuple share the same parity, we deduce that every element is even. It follows that

$$(x/2,\ y/2,\ (x + y)/2,\ xy/2)$$

consists of integers which are monochromatic under c. Finally, we observe that

$$\left(\frac{x+y}{2}\right)^2 - \left(\frac{x}{2}\right)^2 = \left(\frac{y}{2}\right)^2 + \frac{xy}{2}.$$ □

Theorem 9.1. *Let $a_1, \ldots, a_s, b_1, \ldots, b_t \in \mathbb{Z} \setminus \{0\}$ with $s, t \geqslant 1$ and*

$$a_1 + \cdots + a_s = 0. \quad (9.1)$$

Then in any finite colouring of $\mathbb{N}$ *there are infinitely many tuples* $(x_1, \dots, x_s, y_1, \dots, y_t)$ *which are monochromatic and which solve the equation* (1.7)*.*

Proof. We closely follow the proof of [Mor17, Corollary 1.7]. As is shown in [Mor17, §6], there are integers $u_1, \dots, u_s$ not all of which are zero and which satisfy

$$a_1 u_1^2 + \cdots + a_s u_s^2 = 0. \tag{9.2}$$

We claim that we may assume that $a_1u_1 + \cdots + a_s u_s > 0$. If $a_1u_1 + \cdots + a_s u_s < 0$ then we reverse the sign of all the u_i. If $a_1u_1 + \cdots + a_s u_s = 0$ then reversing the sign of a single non-zero u_i gives $a_1u_1 + \cdots + a_s u_s \neq 0$ and we proceed as before.

Let $v_1, \dots, v_t$ denote integers satisfying

$$b_1 v_1 + \cdots + b_t v_t = 0.$$

For instance, one could take $b_i = 0$ for all i, but this is a poor choice if one wishes to generate a monochromatic solution to (1.7) in which all variables are distinct.

Set

$$a := 2(a_1u_1 + \cdots + a_s u_s) \quad \text{and} \quad b := b_1 + \cdots + b_t. \tag{9.3}$$

Given a colouring $c : \mathbb{N} \to [r]$ define

$$\tilde{c}(n) := \begin{cases} c(bn/a) & \text{if } a \mid n, \\ r + (n \bmod a) & \text{otherwise.} \end{cases}$$

Then $\tilde{c}$ is a finite colouring of $\mathbb{N}$. Applying [Mor17, Theorem 1.4], there exists infinitely many tuples (x, y, z) giving rise to a $\tilde{c}$-monchromatic configuration of the form

$$x, \quad x+y, \quad x+u_1y, \quad \dots \quad , \quad x+u_sy, \quad xy, \quad xy+v_1z, \quad \dots \quad , \quad xy+v_tz. \tag{9.4}$$

Since $x \equiv x + y \pmod a$, we must have that $y \equiv 0 \pmod a$. Since $x \equiv xy \pmod a$, it follows that all the elements of (9.4) are divisible by a, and that the configuration

$$\frac{b(x+u_1y)}{a}, \quad \dots \quad , \quad \frac{b(x+u_sy)}{a}, \quad \frac{b(xy+v_1z)}{a}, \quad \dots \quad , \quad \frac{b(xy+v_tz)}{a} \tag{9.5}$$

is monochromatic under c.

Setting

$$x_i := \frac{b(x+u_iy)}{a} \quad \text{and} \quad y_j := \frac{b(xy+v_jz)}{a}$$

we obtain a monochromatic solution to the equation (1.7). □

With this in hand, we are able to complete our proof of Theorem 1.10. Since Proposition 1.9 establishes the necessity of Di Nasso and Luperi Baglini's criterion, we need only show that the criterion is sufficient for partition regularity. In other words, we wish to show that if $a_1, \dots, a_s, b_1, \dots, b_t \in \mathbb{Z} \setminus \{0\}$ with $s, t \geqslant 1$ and one of the following holds

(1) there exists $I \neq \emptyset$ with $\sum_{i \in I} a_i = 0$;
(2) there exists $I \neq \emptyset$ with $\sum_{i \in I} b_i = 0$.

then the equation

$$a_1x_1^2 + \cdots + a_sx_s^2 = b_1y_1 + \cdots + b_ty_t \tag{9.6}$$

is partition regular. According to our formulation of Theorem 1.10, we may assume that (9.6) does not take the form

$$a(x_1^2 - x_2^2) = by + cz \tag{9.7}$$

for some non-zero integers a, b, c.

Let us first suppose that we are in situation (2). Applying Theorem 1.4 we obtain infinitely many monochromatic solutions by letting $N \to \infty$.

Next let us suppose that we are in situation (1). If $s \geqslant 3$ and $s + t \geqslant 5$ then we may employ Theorem 1.7. Hence we may assume that either $s < 3$ or $s + t < 5$. Supposing that $s < 3$, condition (1) implies that $2 \geqslant s \geqslant |I| \geqslant 2$, so that $I = \{1, 2\} = [s]$. This situation is covered by Theorem 9.1

Finally let us suppose that $s \geqslant 3$ and $s + t < 5$. Since $t \geqslant 1$, we must have $s = 3$ and $t = 1$. If $I = \{1, 2, 3\} = [s]$ then we are in the situation covered by Theorem 9.1. We may therefore assume that $|I| = 2$, $s = 3$ and $t = 1$. Hence our equation can be written in the form (9.7), a case we do not have to deal with. This completes our proof of Theorem 1.10.

References

[Ber96] V. Bergelson. Ergodic Ramsey theory—an update. In *Ergodic theory of* $\mathbf{Z}^d$ *actions (Warwick, 1993–1994)*, volume 228 of *London Math. Soc. Lecture Note Ser.*, pages 1–61. Cambridge Univ. Press, Cambridge, 1996. ↑2

[Bou89] J. Bourgain. On $\Lambda(p)$-subsets of squares. *Israel J. Math.*, 67(3):291–311, 1989. ↑5, ↑13, ↑14, ↑17, ↑18, ↑31, ↑33, ↑34

[BP17] T. D. Browning and S. Prendiville. A transference approach to a Roth-type theorem in the squares. *Int. Math. Res. Not.*, (7):2219–2248, 2017. ↑21

[Cha] J. Chapman. Partition regularity and multiplicatively syndetic sets. *Acta Arith. (to appear)*. ↑10

[CLP] S. Chow, S. Lindqvist, and S. Prendiville. Rado's criterion over squares and higher powers. *J. Eur. Math. Soc. (JEMS) (to appear)*. ↑2, ↑3, ↑7, ↑10, ↑21, ↑34

[CP] J. Chapman and S. Prendiville. On the Ramsey number of the Brauer configuration. *Bull. Lond. Math. Soc. (to appear)*. ↑7

[DNLB18] M. Di Nasso and L. Luperi Baglini. Ramsey properties of nonlinear Diophantine equations. *Adv. Math.*, 324:84–117, 2018. ↑3, ↑4

[FGR88] P. Frankl, R. L. Graham, and V. Rödl. Quantitative theorems for regular systems of equations. *J. Combin. Theory Ser. A*, 47(2):246–261, 1988. ↑5

[GL19] B. Green and S. Lindqvist. Monochromatic solutions to $x + y = z^2$. *Canad. J. Math.*, 71(3):579–605, 2019. ↑3

[Gra07] R. Graham. Some of my favorite problems in Ramsey theory. In *Combinatorial number theory*, pages 229–236. de Gruyter, Berlin, 2007. ↑1

[Gra08] R. Graham. Old and new problems and results in Ramsey theory. In *Horizons of combinatorics*, volume 17 of *Bolyai Soc. Math. Stud.*, pages 105–118. Springer, Berlin, 2008. ↑1

[Gre05] B. Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geom. Funct. Anal.*, 15(2):340–376, 2005. ↑2, ↑7

[GT08] B. Green and T. Tao. Quadratic uniformity of the Möbius function. *Ann. Inst. Fourier (Grenoble)*, 58(6):1863–1935, 2008. ↑21

[GT10] B. Green and T. Tao. An arithmetic regularity lemma, an associated counting lemma, and applications. In *An irregular mind*, volume 21 of *Bolyai Soc. Math. Stud.*, pages 261–334. János Bolyai Math. Soc., Budapest, 2010. ↑2, ↑11

[HH18] K. Henriot and K. Hughes. Restriction estimates of ε-removal type for k-th powers and paraboloids. *Math. Ann.*, 372(3-4):963–998, 2018. ↑4, ↑17

[Kat04] Y. Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004. ↑19, ↑20

[Mor17] J. Moreira. Monochromatic sums and products in $\mathbb{N}$. *Ann. of Math. (2)*, 185(3):1069–1090, 2017. ↑3, ↑41

[Pac18] P. P. Pach. Monochromatic solutions to $x + y = z^2$ in the interval $[N, cN^4]$. *Bull. Lond. Math. Soc.*, 50(6):1113–1116, 2018. ↑3

[Pre17] S. Prendiville. Four variants of the Fourier-analytic transference principle. *Online J. Anal. Comb.*, (12):25, 2017. ↑9, ↑36

[Rad33] R. Rado. Studien zur Kombinatorik. *Math. Z.*, 36(1):424–470, 1933. ↑1, ↑3

[Tao] T. Tao. 254A, Notes 1: Elementary multiplicative number theory. Available at https://bit.ly/38I25rM. ↑22, ↑23, ↑25

[Tao12] T. Tao. *Higher order Fourier analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. ↑11, ↑14

[TV06] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. ↑36

[Vau97] R. C. Vaughan. *The Hardy-Littlewood method*, volume 125 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 1997. ↑3

Department of Mathematics and Statistics, Lancaster University, UK
E-mail address: s.prendiville@lancaster.ac.uk