Math. Z. (2018) 289:51–69
https://doi.org/10.1007/s00209-017-1943-7

Mathematische Zeitschrift

Monochromatic sums of squares

Gyan Prakash[1] · D. S. Ramana[1] · O. Ramaré[2]

Received: 11 September 2016 / Accepted: 6 June 2017 / Published online: 12 October 2017

Abstract For any integer $K \geq 1$ let $s(K)$ be the smallest integer such that in any colouring of the set of squares of the integers in K colours every large enough integer can be written as a sum of no more than $s(K)$ squares, all of the same colour. A problem proposed by Sárközy asks for optimal bounds for $s(K)$ in terms of K. It is known by a result of Hegyvári and Hennecart that $s(K) \geq K \exp\left(\frac{(\log 2+o(1))\log K}{\log\log K}\right)$. In this article we show that $s(K) \leq K \exp\left(\frac{(3\log 2+o(1))\log K}{\log\log K}\right)$. This improves on the bound $s(K) \ll_\epsilon K^{2+\epsilon}$, which is the best available upper bound for $s(K)$.

Keywords Monochromatic · Squares · Circle method

Mathematics Subject Classification Primary 11N36; Secondary 11P99

1 Introduction

For any integer $K \geq 1$, a colouring in K colours of the set $\mathfrak{Q}$ of the squares of the integers is a partition of $\mathfrak{Q}$ into K disjoint subsets. Each subset of $\mathfrak{Q}$ in such a partition is called a colour of the colouring. Let $s(K)$, for any integer $K \geq 1$, be the smallest integer such that given any colouring of $\mathfrak{Q}$ in K colours, every sufficiently large integer is expressible as a

✉ D. S. Ramana
suri@hri.res.in

Gyan Prakash
gyan@hri.res.in

O. Ramaré
ramare@math.univ-lille1.fr

1 Harish-Chandra Research Institute (HBNI), Jhunsi, Allahabad 211 019, India

2 CNRS/Institut de Mathématiques de Marseille, Aix Marseille Université, Centrale Marseille, 12 M UMR 7373, 13453 Marseille, France

sum of at most $s(K)$ squares, all of the same colour. Then Sárközy remarks on page 29 of [10] that it is easily seen that $s(K)$ is finite for each integer $K \geq 1$ and, in Problem 40 of the list of problems in [10], Sárközy asks for bounds, in terms of K, for $s(K)$ as well as the corresponding integer in the analogous problem for the set of prime numbers.

Our present contribution towards the solution of Sárközy's problem for the squares is the following theorem.

Theorem 1.1 *For any integer $K \geq 2$ we have $s(K) \leq K \exp\left(\frac{(3\log 2+o(1))\log K}{\log\log K}\right)$.*

Here $o(1) \ll \frac{\log\log\log K}{\log\log K}$ for all large enough K. This improves on the bounds $s(K) \ll (K\log K)^5$ given by Theorem 1, page 318 of Hegyvári and Hennecart [4] and $s(K) \ll_\epsilon K^{2+\epsilon}$ given subsequently by Theorem 1.1, page 18 of Akhilesh and Ramana [1]. Moreover, our upper bound for $s(K)$ compares fairly well with the lower bound

$$s(K) \geq K \exp\left(\frac{(\log 2 + o(1))\log K}{\log\log K}\right) \tag{1}$$

for all $K \geq 2$ provided by Theorem 2, page 319 of [4].

For the convenience of the reader we summarise here the proof of the lower bound (1) from [4]. For any integer $m \geq 1$, let U_m be the product of the first m prime numbers. We partition the squares coprime to U_m by the classes they belong to in $\mathbf{Z}/U_m\mathbf{Z}$ and partition the remaining squares by their smallest divisor from the set of primes dividing U_m. This defines a colouring of $\mathfrak{Q}$. The number of colours in this colouring is $K_m = m + b_m$, where b_m is the number of invertible square classes in $\mathbf{Z}/U_m\mathbf{Z}$. It is then verified that at least U_m summands are required to represent any given squarefree multiple of U_m as a sum of squares, all of the same colour with respect to this colouring of $\mathfrak{Q}$. This implies that $s(K_m) \geq U_m$ for all $m \geq 1$. The lower bound (1) results on applying this conclusion to m such that $K_m \leq K < K_{m+1}$ for a given integer $K \geq 1$ and using standard estimates on the distribution of prime numbers to express U_m in terms of K.

We now turn to the proof of Theorem 1.1. As with Ramana and Ramaré [7], which treats Sárközy's problem for the set of primes, and [1], our proof of Theorem 1.1 ultimately relies on the elegant principle underlying the argument in [4] for the upper bound $s(K) \ll (K\log K)^5$. We paraphrase this principle in Lemma 1.2 below with the aid of the following notation.

For any subset S of the integers and any integer $m \geq 1$, we write $E_m(S)$ for the number of tuples $(x_1, x_2, \ldots, x_{2m})$ in S^{2m} satisfying

$$x_1 + x_2 + x_3 + \cdots + x_m = x_{m+1} + x_{m+2} + \cdots + x_{2m}. \tag{2}$$

Lemma 1.2 *Let N, L and m be integers and D a real number satisfying the conditions $L \geq N \geq 2D(mD+1)$, $D \geq 1$ and $m \geq 2$. If S is a subset of the integers in the interval $(N, N+L]$ such that*

$$E_m(S) \;\leq\; \frac{|S|^{2m} D}{L} \tag{3}$$

and if S contains an integer that is not divisible by any prime number $p \leq \lceil mD \rceil$ then every integer $n \geq (2\lceil mD\rceil + 1)m(N+L)$ is a sum of no more than $\frac{n}{N}$ elements of S.

This lemma is a consequence of a well-known finite addition theorem, also due to Sárközy. We use this theorem in the form provided by Lev [5]. Deferring the detailed proof of Lemma 1.2 to Sect. 4.1, let us describe how this lemma applies to Sárközy's problem. For an integer $K \geq 1$, let $\mathcal{B}$ be the set of squares of integers that are not divisible by any prime

$p \leq B$, where B is a fixed but large power of K, say $B = K^{13}$. For a given integer $N \geq 1$, let $\mathcal{B}(N)$ denote $\mathcal{B} \cap (N, 4N]$. It is then readily verified that there is a $C > 0$ such that $|\mathcal{B}(N)| \geq \frac{N^{\frac{1}{2}}}{C \log K} \geq K$ when N is large enough. Suppose now that $\cup_{1 \leq i \leq K} \mathfrak{Q}_i$ is a partition of the set $\mathfrak{Q}$ into K disjoint subsets. Then for some i in $[1, K]$ the set $\mathfrak{Q}_i \cap \mathcal{B}(N)$ contains at least $\frac{|\mathcal{B}(N)|}{K}$ of the elements of $\mathcal{B}(N)$. Thus if we set

$$S = \mathfrak{Q}_i \cap (N, 4N], \tag{4}$$

then S is a subset of the squares in the interval $(N, 4N]$ satisfying $|S| \geq N^{\frac{1}{2}}/A$, with $A = CK \log K \geq 1$. As we verify later (see (57)), it follows from the classical bounds for the number of representations of integers as sums of five squares that for any subset S of the squares in $(N, 4N]$ satisfying $|S| \geq N^{\frac{1}{2}}/A$ for some $A \geq 1$ we have

$$E_5(S) \ll |S|^5 N^{\frac{3}{2}} \ll \frac{|S|^{10} A^5}{N}. \tag{5}$$

Therefore the bound (3) holds with $m = 5$ and $L = 3N$ and $D = C_1 A^5$, for some $C_1 \geq 1$. Since $\lceil 5D \rceil \leq K^{13}$ when K is large enough and since S contains elements of $\mathcal{B}$, the set S satisfies the conditions of Lemma 1.2. We then conclude that every integer $n \geq (200D+60)N$ is a sum of no more than $\frac{n}{N}$ elements of S. In particular, every integer in the interval $I(N) = ((200D + 60)N, (200D + 61)N]$ is a sum of at most $C_2(K \log K)^5$ squares all belonging to S and hence to $\mathfrak{Q}_i$, for some $C_2 > 0$. Thus when N is large enough, every integer in $I(N)$ is the sum of no more than $C_2(K \log K)^5$ squares, all of the same colour. Note, of course, that the colour may vary with N. Nevertheless, since $I(N)$ meets $I(N + 1)$ for all large enough N, we deduce that $s(K) \ll (K \log K)^5$, as given by [4].

In the remainder of this article we shall show that the argument of the preceding paragraph can be improved to yield Theorem 1.1 essentially by taking S in (4) to be $\mathfrak{Q}_i \cap \mathcal{B}(N)$ rather than $\mathfrak{Q}_i \cap (N, 4N]$. This is on account of the following theorem, suggested by [7] and the recent work of Browning and Prendiville [2].

Theorem 1.3 *Let $A \geq e^{e^2}$ and $l \geq 12$ be real numbers. Then for all sufficiently large integers N, depending only on A and ℓ, and any subset S of the squares in the interval $(N, 4N]$ with $|S| \geq \frac{N^{\frac{1}{2}}}{A}$ and such that no integer in S is divisible by a prime $p \leq A^l$ we have*

$$E_6(S) \leq \frac{|S|^{11}}{N^{\frac{1}{2}}} \exp\left(\frac{(3 \log 2 + o_l(1)) \log A}{\log \log A}\right), \tag{6}$$

where $o_\ell(1) \ll_\ell \frac{\log \log \log A}{\log \log A}$.

Let us note that the bound (6) does not necessarily hold if we assume only that S is a subset of the squares in the interval $(N, 4N]$ satisfying $|S| \geq \frac{N^{\frac{1}{2}}}{A}$ for some $A \geq 1$. For instance, we may take $S = \{A^2 n^2 \mid M < n \leq 2M\}$ where M and A are integers ≥ 1. Then with $N = A^2 M^2$ we have $S \subseteq (N, 4N]$ and $|S| = M = \frac{N^{\frac{1}{2}}}{A}$. A classical application of the circle method now shows that for some $C > 0$ we have $E_6(S) \sim CM^{10}$ as $M \to +\infty$, so that $E_6(S) \gg |S|^{10}$, contradicting (6) when A and M are sufficiently large.

We prove Theorem 1.3 in Sect. 3. Our basic strategy for proving this theorem is similar to that in [7] and goes back to the method of Ramaré and Ruzsa [8]. More precisely, we set $U = \prod_{p \le A^\ell} p$ and first show that

$$E_6(S) \le \frac{5\tau(U)}{2N^{\frac{1}{2}}} |\{x \in S^{11} | f(x) \text{ an invertible square mod } 4U\}| + O\left(\frac{|S|^{11}}{AN^{\frac{1}{2}}}\right), \qquad (7)$$

where $f(x)$ denotes $x_1 + x_2 + \cdots + x_6 - x_7 - \cdots - x_{11}$ for any $x = (x_1, x_2, \ldots, x_{11}) \in S^{11}$ and $\tau(U)$ is the number of divisors of U. We obtain (7) by an application of the circle method following [2]. We then complete the proof of Theorem 1.3 by estimating

$$|\{x \in S^{11} | f(x) \text{ an invertible square mod } 4U\}| \qquad (8)$$

using Theorem 2.1 of Sect. 2, which treats a more general problem. In Sect. 4, our concluding section, we finally detail the path from Theorems 1.3 to 1.1.

Throughout this article we use $e(z)$ to denote $e^{2\pi i z}$, for any complex number z and write $e_p(z)$ for $e^{\frac{2\pi i z}{p}}$ when p is a prime number. Further, all constants implied by the symbols $\ll$ and $\gg$ are absolute except when dependencies are indicated, either in words or by subscripts to these symbols. The Fourier transform $\widehat{f}$ of an integrable function f on $\mathbf{R}$ is defined by $\widehat{f}(u) = \int_{\mathbf{R}} f(t)e(-ut)dt$. Finally, the notations $[a, b]$, $(a, b]$ etc. will denote intervals in $\mathbf{Z}$, rather than in $\mathbf{R}$, with end points a, b, unless otherwise specified.

2 The local problem

The main result of this section is Theorem 2.1. We shall suppose that $A \ge e^{e^2}$ and $l \ge 2$ and let

$$U = \prod_{p \le w} p, \quad \text{where} \quad w = A^l. \qquad (9)$$

In addition, we let $\mathcal{Z}$ be a subset of the integers satisfying the conditions

$$|\mathcal{Z}| \ge \frac{M}{A} \quad \text{and} \quad |\{z \in \mathcal{Z} | z \equiv a \bmod U\}| \le \frac{BM}{U}, \qquad (10)$$

for all classes a in $\mathbf{Z}/U\mathbf{Z}$ and some $B > 0$ and $M \ge 1$, real numbers. As before, $\tau(U) = 2^{\pi(w)}$ is the number of divisors of U. Also, we denote by $\mathbf{c} = \{c(i)\}_{i \in I}$ a given finite sequence of integers and finally we let $R_U(\mathcal{Z}, \mathbf{c})$ denote the set of triples (x, y, i) in $\mathcal{Z} \times \mathcal{Z} \times I$ such that $x^2 + y^2 + c(i)$ reduces to an invertible square modulo U.

Theorem 2.1 *With notation as above and supposing also that $A^\ell \ge 4BA \ge 4e^{e^2}$ we have*

$$|R_U(\mathcal{Z}, \mathbf{c})| \le \frac{|\mathcal{Z}|^2 |I|}{\tau(U)} \exp\left(\frac{(3 \log 2 + o_{\ell,B}(1)) \log BA}{\log\log BA}\right), \qquad (11)$$

where $o_{\ell,B}(1) \ll_\ell \frac{\log\log\log BA}{\log\log BA}$.

We prove Theorem 2.1 in Sect. 2.4. We do this by using the optimisation principle given by Lemma 2.7 to pass to a problem in $\mathbf{Z}/U\mathbf{Z}$, dealt with by Theorem 2.6. By means of a pair of applications Hölder's inequality and the Chinese Remainder Theorem we reduce the proof of Theorem 2.6 to the solution of a problem in $\mathbf{Z}/p\mathbf{Z}$ for a given prime $p|U$. This problem is treated by Proposition 2.2 of the following subsection. Theorem 2.6 is the analogue of

Proposition 2.3 of [7] in our context. However, the argument we use for Theorem 2.6 is both conceptually simpler and more efficient than the argument leading to Proposition 2.3 in [7], even if the first few steps in both cases are similar. In fact, and as will be shown in another paper, our proof of Theorem 2.6 can be adapted to improve the conclusion of the cited proposition from [7] and hence also that of the main result of [7].

2.1 A sum over Z/pZ

Throughout this subsection p shall denote fixed prime number, G_p the ring $\mathbf{Z}/p\mathbf{Z}$ and c a given element of G_p. Also, $\lambda_p(x)$ shall denote the Legendre symbol $(\frac{x}{p})$, for any x in G_p. Furthermore, for any (x, y) in G_p^2 we set $\delta_p(x, y) = \lambda_p(x^2 + y^2 + c)$ and $\epsilon_p(x, y) = 1 + \delta_p(x, y)$.

We endow G_p, and likewise G_p^t for any integer $t \geq 1$, with their uniform probability measures and write $\mathbb{E}_x$ and $\mathbb{E}_{x_1,x_2,\ldots,x_t}$ in place of $\frac{1}{p}\sum_{x\in G_p}$ and $\frac{1}{p^t}\sum_{x_1,x_2,\ldots,x_t\in G_p}$ respectively. When t is fixed, we will use $\mathbf{x} = (x_1, x_2, \ldots, x_t)$ for elements of G_p^t and abbreviate $\mathbb{E}_{x_1,x_2,\ldots,x_t}$ further to $\mathbb{E}_{\mathbf{x}}$. Also, we will use these notations in the same sense with other letters in place of x. Finally, we define $\mathcal{E}_p(k, t)$ for any integer k with $1 \leq k \leq t$ by

$$\mathcal{E}_p(k,t) = \mathbb{E}_{y_1,y_2,\ldots,y_t}\mathbb{E}_{x_1,x_2,\ldots,x_t}\prod_{\substack{1\leq i\leq t,\\ 1\leq j\leq k.}}\epsilon_p(x_i,y_j) = \mathbb{E}_{\mathbf{y}}\mathbb{E}_{\mathbf{x}}\prod_{\substack{1\leq i\leq t,\\ 1\leq j\leq k.}}\epsilon_p(x_i,y_j). \tag{12}$$

Proposition 2.2 *For any integers t, k satisfying $t \geq 2$ and $1 \leq k \leq \frac{t}{2}$ we have*

$$\mathcal{E}_p(k,t) \leq \exp\left(\frac{8kt^4 2^t}{p}\right). \tag{13}$$

We shall prove (13) for a given integer $t \geq 2$ and all integers k satisfying $1 \leq k \leq \frac{t}{2}$ by induction on k starting from $k = 1$, using Proposition 2.3 and Lemma 2.4 below. Let us note that the trivial upper bound 2^{kt} for $\mathcal{E}_p(k, t)$ implies (13) when $p \leq 8t^3 2^t$. This allows us, in particular, to assume that $p > 2$.

Proposition 2.3 *Let t be an integer ≥ 1 and J be a non-empty subset of $\{1, 2, \ldots t\}$. Further, let $\mathcal{B}(J)$ be the subset of G_p^t consisting of $\mathbf{y} = (y_1, y_2, \ldots, y_t)$ in G_p^t such that either $y_j^2 = y_k^2$ for some distinct j, k in J or $y_j^2 = -c$ for some j in J. Then we have*

(i) $|\mathbb{E}_x \prod_{j\in J}\delta_p(x, y_j)| < \frac{2|J|}{\sqrt{p}}$ *when* $\mathbf{y} = (y_1, y_2, \ldots, y_t) \notin \mathcal{B}(J)$ *and*

(ii) $\left|\mathbb{E}_{\mathbf{y}}\mathbb{E}_x \prod_{j\in J}\delta_p(x, y_j)\right| \leq \frac{2}{p}$.

Proof The bound (i) is a consequence of the Weil bounds for character sums. Indeed, it follows from Theorem 2C on page 43 of [11] that

$$|\mathbb{E}_x\prod_{j\in J}\delta_p(x,y_j)| = \frac{1}{p}\left|\sum_{x\in G_p}\lambda_p\left(\prod_{j\in J}(x^2+y_j^2+c)\right)\right| \leq \frac{2|J|-1}{\sqrt{p}}, \tag{14}$$

when the polynomial $f(X) = \prod_{j\in J}(X^2 + y_j^2 + c)$ is not a square in $\overline{\mathbf{F}}_p[X]$, where $\overline{\mathbf{F}}_p$ is an algebraic closure of $\mathbf{F}_p$. Since this condition holds for $\mathbf{y} = (y_1, y_2, \ldots, y_t) \notin \mathcal{B}(J)$, we have (i).

To verify (ii), we begin by recalling that for all $x \in G_p$ we have the classical identity

$$\gamma_p\lambda_p(x) = \sum_{a\in G_p}\lambda_p(a)e_p(ax) \tag{15}$$

where γ_p, the Gauss sum to modulus p, is the right hand side of the above relation evaluated at $x = 1$. If for any $a \in G_p$ we set $\ell(a) = 0$ when $a = 0$ and $\ell(a) = \frac{\gamma_p}{p}$ when $a \neq 0$ then it is easily seen from (15) that

$$\lambda_p(a)\mathbb{E}_y e_p(ay^2) = \ell(a) \quad \text{for all } \ a \text{ in } G_p. \tag{16}$$

On combining (15) and (16) we deduce that for any $b \in G_p$ we have

$$\mathbb{E}_y \lambda_p(y^2 + b) = \frac{1}{\gamma_p} \sum_{a \in G_p} \lambda_p(a)\mathbb{E}_y e_p(ay^2) e_p(ab) = \frac{\mu(b)}{p}, \tag{17}$$

where we have set $\mu(b) = \sum_{a \in G_p^*} e(ab)$, for any $b \in G_p$, with G_p^* denoting the set of non-zero elements of G_p. Thus $\mu(b)$ is $p - 1$ when $b = 0$ and is -1 when $b \neq 0$. By means of (17) we then have that

$$\mathbb{E}_x \mathbb{E}_{\mathbf{y}} \prod_{j \in J} \delta_p(x, y_j) = \mathbb{E}_x \prod_{j \in J} \mathbb{E}_{y_j} \lambda_p(x^2 + y_j^2 + c) = \frac{1}{p^m} \mathbb{E}_x \, \mu(x^2 + c)^m, \tag{18}$$

where $m = |J|$. On using the values of $\mu(b)$ given above to evaluate the last term in each of the cases $c = 0, -c$ is non-zero square and $-c$ is not a square we finally get

$$\left| \mathbb{E}_{\mathbf{y}} \mathbb{E}_x \prod_{j \in J} \delta_p(x, y_j) \right| = \left| \mathbb{E}_x \mathbb{E}_{\mathbf{y}} \prod_{j \in J} \delta_p(x, y_j) \right| \leq \frac{2p^m}{p^{m+1}} = \frac{2}{p}. \tag{19}$$

□

The following is the well-known Hoeffding's lemma from elementary probability theory.

Lemma 2.4 *Let Z be a real valued random variable on a probability space satisfying $a \leq Z \leq b$, for real numbers $a \leq b$. Then for any real s we have*

$$\mathbb{E} \exp(sZ) \leq \exp(s\mathbb{E}Z) \exp\left(\frac{s^2(b-a)^2}{8} \right). \tag{20}$$

Proof Replacing Z with $Z - \mathbb{E}Z$, we may suppose that $\mathbb{E}Z = 0$. Then (20) is easily deduced from the convexity of the function $r \mapsto \exp(sr)$ on the interval $[a, b]$. The details may be found in the proof of Lemma 5.1, page 64 of [3], for example. □

Proof of Proposition 2.2 Let t be an integer ≥ 2. We begin by noting that

$$\begin{aligned} \mathcal{E}_p(1, t) = \mathbb{E}_{\mathbf{x}} \mathbb{E}_{y_1} \prod_{1 \leq i \leq t} \big(1 + \delta_p(x_i, y_1)\big) &\leq 1 \\ &+ \sum_{\substack{J \subseteq \{1,2,\ldots,t\}, \\ J \neq \emptyset.}} \left| \mathbb{E}_{\mathbf{x}} \mathbb{E}_{y_1} \prod_{i \in J} \delta_p(x_i, y_1) \right|, \end{aligned} \tag{21}$$

on expanding the product over $1 \leq i \leq t$ and using the triangle inequality. From the bound (ii) of Proposition 2.3 applied to each summand in the sum over J in (21) we then obtain

$$\mathcal{E}_p(1, t) \leq 1 + \frac{2 \cdot 2^t}{p} \leq \exp\left(\frac{2^{t+1}}{p} \right), \tag{22}$$

which verifies (13) for $k = 1$. Suppose now that $t \geq 4$ and that (13) holds for $k - 1$, where k is an integer satisfying $2 \leq k \leq \frac{t}{2}$, and let us verify it for k. We recall the definition of $\mathcal{B}(J)$

from Proposition 2.3 and set $\mathcal{B} = \mathcal{B}(J)$ with $J = \{1, 2, \ldots, k\}$. Then on writing $\mathcal{B}'$ for the complement of $\mathcal{B}$ in G_p^t we have

$$\mathcal{E}_p(k,t) = \mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}}(\mathbf{y}) \mathbb{E}_{\mathbf{x}} \prod_{\substack{1 \le i \le t, \\ j \in J.}} \epsilon_p(x_i, y_j) + \mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}'}(\mathbf{y}) \mathbb{E}_{\mathbf{x}} \prod_{\substack{1 \le i \le t, \\ j \in J.}} \epsilon_p(x_i, y_j). \tag{23}$$

Let us estimate the first term on the right hand side of (23). To this end, we set $\alpha_l(\mathbf{y}) = 1$ for any $l \in J$ and any $\mathbf{y}$ in $(y_1, y_2, \ldots, y_t) \in G_p^t$ if either $y_l^2 = y_j^2$ for some $j \in J$ distinct from l or if $y_l^2 = -c$ and set $\alpha_l(\mathbf{y}) = 0$ otherwise. Then for all $\mathbf{y} \in G_p^t$ we have $1_{\mathcal{B}}(\mathbf{y}) \le \sum_{l \in J} \alpha_l(\mathbf{y})$ and consequently

$$\mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}}(\mathbf{y}) \mathbb{E}_{\mathbf{x}} \prod_{\substack{1 \le i \le t, \\ j \in J.}} \epsilon_p(x_i, y_j) \le \sum_{l \in J} \mathbb{E}_{\mathbf{y}} \mathbb{E}_{\mathbf{x}} \alpha_l(\mathbf{y}) \prod_{\substack{1 \le i \le t, \\ j \in J.}} \epsilon_p(x_i, y_j). \tag{24}$$

For any $l \in J$ let us write $\mathbb{E}_{\hat{\mathbf{y}}_l}$ for $\mathbb{E}_{y_1, y_2, \ldots, y_t}$ with the variable y_l dropped. Then the trivial bound $\prod_{1 \le i \le t} \epsilon_p(x_i, y_l) \le 2^t$ shows that the right hand side of (24) does not exceed

$$2^t \sum_{l \in J} \mathbb{E}_{\hat{\mathbf{y}}_l} \mathbb{E}_{\mathbf{x}} \, \mathbb{E}_{y_l} \alpha_l(\mathbf{y}) \prod_{\substack{1 \le i \le t, \\ j \in J, \, j \ne l.}} \epsilon_p(x_i, y_j). \tag{25}$$

For any $l \in J$ we have $\mathbb{E}_{y_l} \alpha_l(\mathbf{y}) \le \frac{2k}{p}$ and $\mathbb{E}_{\hat{\mathbf{y}}_l} \mathbb{E}_{\mathbf{x}} \prod_{\substack{1 \le i \le t, \\ j \in J, \, j \ne l.}} \epsilon_p(x_i, y_j) = \mathcal{E}_p(k-1, t)$. Since $|J| = k$, it follows that (25) does not exceed $\frac{2^{t+1}k^2}{p} \mathcal{E}_p(k-1, t)$. We then conclude from (24) that

$$\mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}}(\mathbf{y}) \mathbb{E}_{\mathbf{x}} \prod_{\substack{1 \le i \le t, \\ j \in J.}} \epsilon_p(x_i, y_j) \le \frac{2^{t+1}k^2}{p} \mathcal{E}_p(k-1, t). \tag{26}$$

Turning now to the second term on the right hand side of (23), we define the random variable X on G_p^t by

$$X(\mathbf{y}) = \left(\mathbb{E}_x \prod_{j \in J} \epsilon_p(x, y_j) \right) - 1 = \sum_{\substack{I \subseteq J, \\ I \ne \emptyset.}} \mathbb{E}_x \prod_{j \in I} \delta_p(x, y_j). \tag{27}$$

and set $Z = 1_{\mathcal{B}'} X$. Then we have

$$\mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}'}(\mathbf{y}) \mathbb{E}_{\mathbf{x}} \prod_{\substack{1 \le i \le t, \\ j \in J.}} \epsilon_p(x_i, y_j) = \mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}'} (1 + X)^t \le \mathbb{E}_{\mathbf{y}} \exp(tZ), \tag{28}$$

since $1_{\mathcal{B}'}(1 + X) \le \exp(1_{\mathcal{B}'} X)$ and also $0 \le 1 + X$ from (27). We apply Lemma 2.4 to estimate the last term of (28). Let us first note from (27) that for any $\mathbf{y} \in G_p^t$ we have

$$|Z(\mathbf{y})| = 1_{\mathcal{B}'}(\mathbf{y}) |X(\mathbf{y})| \le 1_{\mathcal{B}'}(\mathbf{y}) \sum_{\substack{I \subseteq J, \\ I \ne \emptyset}} \left| \mathbb{E}_x \prod_{j \in I} \delta_p(x, y_j) \right| \le \frac{2k \cdot 2^k}{\sqrt{p}}, \tag{29}$$

by the triangle inequality and (i) of Proposition 2.3 applied to each summand in the sum over I. Further, we have $Z \leq X + 1_{\mathcal{B}}$ since $0 \leq 1_{\mathcal{B}}(1 + X)$. It follows that

$$\mathbb{E}_{\mathbf{y}} Z \leq \sum_{\substack{I \subseteq J, \\ I \neq \emptyset}} \left| \mathbb{E}_{\mathbf{y}} \mathbb{E}_x \prod_{j \in I} \delta_p(x, y_j) \right| + \mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}} \leq \frac{2^{k+1} + 2k^2}{p}, \tag{30}$$

on now using (ii) of Proposition 2.3 to bound each summand in the sum over I and remarking that $\mathbb{E}_{\mathbf{y}} 1_{\mathcal{B}}(\mathbf{y}) \leq \sum_{l \in J} \mathbb{E}_{y_l} \alpha_l(\mathbf{y}) \leq \frac{2k^2}{p}$. From (30), (29) and Lemma 2.4 we then conclude that

$$\mathbb{E}_{\mathbf{y}} \exp(tZ) \leq \exp\left(\frac{(2^{k+1} + 2k^2)t + 2k^2 t^2 4^k}{p} \right) \leq \exp\left(\frac{4t^4 2^t}{p} \right), \tag{31}$$

by means of the inequalities $2^{k+1} + 2k^2 \leq 2^{t+1}$ and $2k^2 t^2 4^k \leq 2t^4 2^t$, valid since $t \geq 4$ and $k \leq \frac{t}{2}$. This relation taken together with (28), (26) and (23) gives

$$\mathcal{E}_p(k, t) \leq \frac{2^{t+1} k^2}{p} \mathcal{E}_p(k-1, t) + \exp\left(\frac{4t^4 2^t}{p} \right). \tag{32}$$

By the induction hypothesis (13) holds for $k - 1$ and consequently we deduce from (32) that

$$\mathcal{E}_p(k, t) \leq \left(\frac{2^{t+1} k^2}{p} + 1 \right) \exp\left(\frac{(2(k-1) + 1) 4t^4 2^t}{p} \right). \tag{33}$$

Using again the inequality $1 + s \leq \exp(s)$ and noting that $2^{t+1} k^2 \leq 4t^4 2^t$ we then conclude from (33) that (13) holds for k, completing the induction step. □

Remark 2.5 It is perhaps the case that the conclusion of Proposition 2.2 holds for all $t \geq 2$ and all k satisfying $1 \leq k \leq t$. A proof of this assertion would allow us to replace $3 \log 2$ with $2 \log 2$ in (11) and, as a consequence, in Theorem 1.1 as well.

2.2 The problem modulo U

Let, as above, $l \geq 2$ and $A \geq e^{e^2}$ be real numbers and $U = \prod_{p \leq w} p$, where $w = A^{\ell}$. Suppose further that $\mathcal{X}$ and $\mathcal{Y}$ are subsets of $\mathbf{Z}/U\mathbf{Z}$ of density at least $\frac{1}{A}$. That is,

$$|\mathcal{X}| \quad \text{and} \quad |\mathcal{Y}| \geq \frac{U}{A}. \tag{34}$$

For a given element c of $\mathbf{Z}/U\mathbf{Z}$, let $T_c(\mathcal{X}, \mathcal{Y})$ denote the set of pairs $(x, y) \in \mathcal{X} \times \mathcal{Y}$ such that $x^2 + y^2 + c$ is an invertible square in $\mathbf{Z}/U\mathbf{Z}$.

Theorem 2.6 *For all l, A, U, $\mathcal{X}$, $\mathcal{Y}$ and c as above, we have*

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \frac{|\mathcal{X}||\mathcal{Y}|}{\tau(U)} \exp\left(\frac{\left(3 \log 2 + O_l\left(\frac{\log\log\log A}{\log\log A}\right)\right) \log A}{\log\log A} \right). \tag{35}$$

Proof We shall write G for the ring $\mathbf{Z}/U\mathbf{Z}$ and continue to use G_p for $\mathbf{Z}/p\mathbf{Z}$. Also, for any x in G and $p|U$ we denote the canonical image of x in $\mathbf{Z}/p\mathbf{Z}$ by x_p and, to be consistent with the notation of preceding subsection, write $\lambda_p(x)$ for the Legendre symbol $(\frac{x_p}{p})$. Then we have that

$$|T_c(\mathcal{X},\mathcal{Y})| \leq \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\prod_{p|U}\left(\frac{1+\lambda_p(x^2+y^2+c)}{2}\right), \tag{36}$$

since $0 \leq 1+\lambda_p(x^2+y^2+c) \leq 2$ for any pair (x, y) in $\mathcal{X}\times\mathcal{Y}$, with equality in the upper bound for every prime $p|U$ when x^2+y^2+c is an invertible square in G. On extending the definitions of δ_p and ϵ_p from Sect. 2.2 by setting $\delta_p(x, y) = \lambda_p(x^2+y^2+c)$ and $\epsilon_p(x, y) = 1+\delta_p(x, y)$ for any (x, y) in G^2 and $p|U$, we may rewrite (36) as

$$|T_c(\mathcal{X},\mathcal{Y})| \leq \frac{1}{\tau(U)}\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\prod_{p|U}\epsilon_p(x, y). \tag{37}$$

Let $t \geq 2$ be an even integer. Then an interchange of summations followed by an application of Hölder's inequality to exponent t to the right hand side of (37) gives

$$|T_c(\mathcal{X},\mathcal{Y})| \leq \frac{|\mathcal{Y}|^{1-\frac{1}{t}}}{\tau(U)}\left(\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}}\prod_{p|U}\epsilon_p(x, y)\right)^t\right)^{\frac{1}{t}}. \tag{38}$$

To bound the sum over $y \in \mathcal{Y}$ on the right hand side of the inequality above, we first expand the summand in this sum and extend the summation to all $y \in G$. By this we see that

$$\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}}\prod_{p|U}\epsilon_p(x, y)\right)^t \leq \sum_{y\in G}\sum_{(x_1,x_2,\ldots,x_t)\in\mathcal{X}^t}\prod_{1\leq i\leq t}\prod_{p|U}\epsilon_p(x_i, y). \tag{39}$$

Interchanging the summations over G and $\mathcal{X}^t$ on the right hand side of the above relation and applying Hölder's inequality again, this time to exponent $\frac{t}{2}$, we deduce that the right hand side of (39) does not exceed

$$|\mathcal{X}|^{t-2}\left(\sum_{(x_1,x_2,\ldots,x_t)\in\mathcal{X}^t}\left(\sum_{y\in G}\prod_{1\leq i\leq t}\prod_{p|U}\epsilon_p(x_i, y)\right)^{\frac{t}{2}}\right)^{\frac{2}{t}}. \tag{40}$$

Finally, on expanding the summand in the sum over $\mathcal{X}^t$ in (40) and extending the summation to all of G^t we conclude using (39) and (38) and a rearrangement of terms that

$$|T_c(\mathcal{X},\mathcal{Y})| \leq \frac{|\mathcal{X}||\mathcal{Y}|}{\tau(U)}\left(\frac{U^3}{|\mathcal{X}|^2|\mathcal{Y}|}\right)^{\frac{1}{t}}\mathcal{E}\left(\frac{t}{2}, t\right)^{\frac{2}{t^2}}, \tag{41}$$

where, for any integer k with $1 \leq k \leq t$, we have set

$$\mathcal{E}(k, t) = \frac{1}{U^{2t}}\sum_{(y_1,y_2,\ldots,y_t)\in G^t}\sum_{(x_1,x_2,\ldots,x_t)\in G^t}\prod_{p|W}\prod_{\substack{1\leq i\leq t,\\ 1\leq j\leq k.}}\epsilon_p(x_i, y_j). \tag{42}$$

The Chinese Remainder Theorem gives $G = \prod_{p|U} G_p$. Moreover, for all $p|U$ and (x, y) in G^2 we have $\epsilon_p(x, y) = \epsilon_p(x_p, y_p)$. It follows that $\mathcal{E}(k, t) = \prod_{p|U}\mathcal{E}_p(k, t)$, where $\mathcal{E}_p(k, t)$ is as defined by (12). Using (13) with $k = \frac{t}{2}$, valid on account of Proposition 2.2, and recalling that $U = \prod_{p \le A^\ell} p$ we then obtain

$$\mathcal{E}(k,t) = \prod_{p|U} \mathcal{E}_p(k,t) \le \exp\left(4t^5 2^t \sum_{p \le A^\ell} \frac{1}{p}\right). \tag{43}$$

From (3.20) on page 70 of [9] we see that $\sum_{p \le A^\ell} \frac{1}{p} \le (\log 2\ell) \log\log A$, since $A \ge 4$ and $\ell \ge 2$. On combining this remark with (43), (34) and (41) we then conclude that for any even integer $t \ge 2$ we have

$$|T_c(\mathcal{X},\mathcal{Y})| \le \frac{|\mathcal{X}||\mathcal{Y}|}{\tau(U)} \exp\left(\frac{3\log A}{t} + 8(\log 2\ell) t^3 2^t \log\log A\right). \tag{44}$$

Let us now set $v \log 2 = \log\left(\frac{\log A}{(\log\log A)^6}\right)$ and suppose that $A_0 \ge e^{e^2}$ is such that we have $\frac{\log\log A}{\log\log\log A} \ge 12$ and $v \ge 2$ for all $A > A_0$. For such A we take t in (44) to be an even integer satisfying $v \le t \le v+2$. Also, with $\kappa = \frac{6\log\log\log A}{\log\log A}$ we have $\kappa \le \frac{1}{2}$ and $v = \frac{(1-\kappa)\log\log A}{\log 2}$. Thus $\frac{1}{t} \le \frac{1}{v} \le \frac{(\log 2)(1+2\kappa)}{\log\log A}$ and $t^3 2^t \le 32 v^3 2^v \le \frac{32\log A}{(\log 2)^3 (\log\log A)^3}$. Substituting these inequalities in (44) we obtain (35) for $A > A_0$. To obtain (35) for $e^{e^2} \le A \le A_0$ it suffices to take $t = 2$ in (44). □

2.3 An optimisation principle

This subsection summarises Subsection 2.3 of [7]. Suppose that $n \ge 1$ is an integer and let P and H be real numbers > 0. Further, assume that the subset $\mathcal{K}$ of points $x = (x_1, x_2, \ldots, x_n)$ in $\mathbf{R}^n$ satisfying the conditions

$$\sum_{1 \le i \le n} x_i = P \quad \text{and} \quad 0 \le x_i \le H \quad \text{for all } i. \tag{45}$$

is not empty. Then $\mathcal{K}$ is a non-empty, compact and convex subset of $\mathbf{R}^n$ and we have the following standard fact.

Lemma 2.7 *If $f : \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ a bilinear form with real coefficients then*

(i) *There are extreme points x^* and y^* of $\mathcal{K}$ so that $f(x,y) \le f(x^*, y^*)$ for all $x, y \in \mathcal{K}$.*
(ii) *If $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ is an extreme point of $\mathcal{K}$ then $x_i^* = 0$ or $x_i^* = H$ for all i excepting at most one. Thus if m is the number of i such $x_i^* \ne 0$ then $mH \ge P > (m-1)H$.*

Proof See the proof of Proposition 2.2 of [7], for example.

2.4 Proof of Theorem 2.1

Let a, b be any elements of $\mathbf{Z}/U\mathbf{Z}$. For any i in I we set $\alpha_i(a,b) = 1$ if $a^2 + b^2 + c(i)$ is an invertible square in $\mathbf{Z}/U\mathbf{Z}$ and 0 otherwise. Further, we write $m(a)$ for the number of z in $\mathcal{Z}$ such that $z \equiv a \bmod U$. Then if $\tilde{\mathcal{Z}}$ denotes the image of $\mathcal{Z}$ in $\mathbf{Z}/U\mathbf{Z}$ we have

$$|R_U(\mathcal{Z}, \mathbf{c})| = \sum_{i \in I} \sum_{(a,b) \in \tilde{\mathcal{Z}}^2} \alpha_i(a,b)\, m(a) m(b). \tag{46}$$

Moreover, we have

$$\sum_{a\in\tilde{\mathcal{Z}}} m(a) = |\mathcal{Z}| \quad \text{and} \quad 0 \le m(a) \le H, \tag{47}$$

with $H = \frac{BM}{U}$, on account of the second assumption in (10). Let us bound the inner sum on the right hand side of (46) for a given i in I. By means of Lemma 2.7 and (47) we obtain

$$\sum_{(a,b)\in\tilde{\mathcal{Z}}^2} \alpha_i(a,b)\, m(a)m(b) \le \sum_{(a,b)\in\tilde{Z}^2} \alpha_i(a,b)\, x_a^* y_b^*, \tag{48}$$

for some x_a^* and y_b^*, with a and b varying over $\tilde{\mathcal{Z}}$, satisfying the following condition. All the x_a^*, and similarly all the y_b^*, are either equal to 0 or to H excepting at most one, which must lie in $(0, H)$. Let $\mathcal{X}$ and $\mathcal{Y}$ be, respectively, the subsets of $\tilde{\mathcal{Z}}$ for which $x_a^* \ne 0$ and $y_b^* \ne 0$. Then we have from (ii) of Lemma 2.7 that $|\mathcal{X}|H \ge |\mathcal{Z}| > (|\mathcal{X}|-1)H$. From the first condition in (10) we then get $|\mathcal{X}| \ge \frac{|\mathcal{Z}|}{H} \ge \frac{U}{AB} \ge 2$, since $U \ge \frac{A^\ell}{2} \ge 2AB$. This gives $H \le \frac{|\mathcal{Z}|}{|\mathcal{X}|-1} \le \frac{2|\mathcal{Z}|}{|\mathcal{X}|}$. The same inequalities hold with $|\mathcal{X}|$ replaced by $|\mathcal{Y}|$. It follows that $H^2 \le \frac{4|\mathcal{Z}|^2}{|\mathcal{X}||\mathcal{Y}|}$. Further, with $T_{c(i)}(\mathcal{X},\mathcal{Y})$ as in Sect. 2.2 we have $\sum_{(a,b)\in\mathcal{X}\times\mathcal{Y}} \alpha_i(a,b) = T_{c(i)}(\mathcal{X},\mathcal{Y})$. Since $\alpha_i(a,b) \ge 0$ for all (a,b), we then deduce that

$$\sum_{(a,b)\in\tilde{Z}^2} \alpha_i(a,b)\, x_a^* y_b^* \le H^2 \sum_{(a,b)\in\mathcal{X}\times\mathcal{Y}} \alpha_i(a,b) \le \frac{4|T_{c(i)}(\mathcal{X},\mathcal{Y})||\mathcal{Z}|^2}{|\mathcal{X}||\mathcal{Y}|}. \tag{49}$$

Combining this with (48), (46) and the bound supplied by (35) for $|T_{c(i)}(\mathcal{X},\mathcal{Y})|$, applicable since $AB \ge e^{e^2}$, we conclude that (11) holds. □

3 An application of the circle method

We prove Theorem 1.3 in this section. As stated in Sect. 1, our first step will be to prove the inequality (7). This is carried out in Sects. 3.1 through 3.3 starting from the preliminaries given below. We then complete the proof of Theorem 1.3 in Sect. 3.4 by applying Theorem 2.1 to estimate (8).

We suppose that $A \ge e^{e^2}$ and $l \ge 12$ are real numbers and assume that N is a sufficiently large integer depending only on A and l, its actual size varying to suit our requirements at various stages of the argument. We set

$$U = \prod_{p\le w} p \quad \text{and} \quad W = 2U, \quad \text{where} \quad w = A^l. \tag{50}$$

Also, we set $\alpha(t) = 1 - \left|\frac{2t}{5N}\right|$ when $|t| \le \frac{5N}{2}$ and 0 for all other $t \in \mathbf{R}$ and set $\beta(t) = \alpha(t - \frac{5N}{2})$. Thus $\beta(t) \ge 0$ for all t in $\mathbf{R}$ and $\beta(t) \ge \frac{2}{5}$ when $t \in [N, 4N]$. Further, S will denote a given subset of the squares in $(N, 4N]$ satisfying the hypotheses of Theorem 1.3. Finally, for all $t \in \mathbf{R}$ we set

$$\psi(t) = \sum_{\substack{0\le r<W,\\ (r,W)=1.}} \sum_{n\equiv r \bmod W} 2n\beta(n^2)e(n^2 t) \tag{51}$$

and $\widehat{S}(t) = \sum_{x\in S} e(xt)$. Then by analogy with (3.1) of [7] we observe that

$$\frac{4}{5}\sqrt{N}E_6(S) \le \int_0^1 \widehat{S}(t)^6\widehat{S}(-t)^5\psi(-t)\,dt. \tag{52}$$

Indeed, $E_6(S)$ is the same as the number of $x \in S^{11}$ such that $f(x) \in S$, with $f(x)$ as in (7). For any such x if $f(x) = n^2$ then $\frac{4}{5}\sqrt{N} \le 2n\beta(n^2)$ and n is invertible modulo W. This remark implies (52) by positivity of β and orthogonality.

We shall apply the circle method to estimate the integral on the right hand side of (52). To this end, we set $L = (\log N)^2$, $Q = W^2A^{12}$, $M = \frac{N}{L}$ and, for any integers a and q satisfying

$$0 \le a \le q \le Q \quad \text{and} \quad (a,q) = 1, \tag{53}$$

we call the interval $[\frac{a}{q} - \frac{1}{M}, \frac{a}{q} + \frac{1}{M})$ the major arc $\mathfrak{M}(\frac{a}{q})$. It is easily checked that distinct major arcs are in fact disjoint when $M > 2Q^2$, which holds when N is sufficiently large depending only on A and l. We denote by $\mathfrak{M}$ the union of the family of major arcs $\mathfrak{M}(\frac{a}{q})$. Each interval in the complement of $\mathfrak{M}$ in $[0, 1)$ is called a minor arc. We denote the union of the minor arcs by $\mathfrak{m}$.

We have

$$\int_0^1 \widehat{S}(t)^6\widehat{S}(-t)^5\psi(-t)\,dt = \int_{-\frac{1}{M}}^{1-\frac{1}{M}} \widehat{S}(t)^6\widehat{S}(-t)^5\psi(-t)\,dt \tag{54}$$

by the periodicity of the integrand. From the definitions given above it is easily seen that the interval $[-\frac{1}{M}, 1 - \frac{1}{M})$ is the union of $\mathfrak{m}$ and $\mathfrak{M}\backslash[1 - \frac{1}{M}, 1 + \frac{1}{M})$. Since distinct major arcs are disjoint, it then follows that the right hand side of (54) is the same as

$$\sum_{1\le q\le Q}\ \sum_{\substack{0\le a<q,\\ (a,q)=1.}} \int_{\mathfrak{M}(\frac{a}{q})} \widehat{S}(t)^6\widehat{S}(-t)^5\psi(-t)\,dt + \int_{\mathfrak{m}} \widehat{S}(t)^6\widehat{S}(-t)^5\psi(-t)\,dt. \tag{55}$$

We shall presently estimate each of the two terms in (55). We begin by observing that

$$\int_0^1 |\widehat{S}(t)|^{11}\,dt \ll |S|^9A^3. \tag{56}$$

In effect, the integral in (56) does not exceed $|S|E_5(S)$. Thus (56) follows from $|S| \ge N^{\frac{1}{2}}/A$ and

$$E_5(S) = \sum_{1\le n} R_5^2(n) = \sum_{1\le n\le 20N} R_5^2(n) \ll N^{\frac{3}{2}} \sum_{n\ge 1} R_5(n) = |S|^5N^{\frac{3}{2}}, \tag{57}$$

where $R_5(n)$ denotes the number of representations of an integer n as a sum of five elements of S. To verify (57) we note that $R_5(n) = 0$ when $n > 20N$ and $R_5(n) \le r_5(n)$, the number of representations of n as a sum of five squares of natural numbers, and recall that $r_5(n) \ll n^{\frac{3}{2}}$, by a standard application of the circle method. As a consequence of (56) we have

$$\sum_{1\le q\le Q}\ \sum_{\substack{0\le a<q,\\ (a,q)=1.}} \int_{\mathfrak{M}(\frac{a}{q})} |\widehat{S}(t)|^{11}dt \le \int_{-\frac{1}{M}}^{1-\frac{1}{M}} |\widehat{S}(t)|^{11}dt \ll |S|^9A^3. \tag{58}$$

3.1 The minor arc contribution

Here we bound the second term in (55). Let us first verify that for all $t \in \mathfrak{m}$ we have

$$|\psi(t)| \ll \frac{N}{A^6}, \tag{59}$$

when N is large enough, depending only on A and l. Indeed, for any real t Dirichlet's approximation theorem gives a rational number $\frac{a}{q}$ satisfying $|t - \frac{a}{q}| \leq \frac{1}{qM}$ together with $1 \leq q \leq M$ and $(a,q) = 1$. When t is in $\mathfrak{m}$ we see that $\frac{a}{q}$ is in $[0, 1]$ since $\mathfrak{m} \subseteq [\frac{1}{M}, 1 - \frac{1}{M})$. Consequently, we also have $0 \leq a \leq q$. Since, however, t is not in $\mathfrak{M}$, we must then have $Q < q$ on account (53). We then conclude using $q^2 \leq qM$ that for each t in $\mathfrak{m}$ there are integers a and $q \neq 0$ with $(a,q) = 1$ satisfying

$$|t - \frac{a}{q}| \leq \frac{1}{q^2} \quad \text{and} \quad Q < q \leq M. \tag{60}$$

Next, for a given class r modulo W, we temporarily let $a(n) = 1$ when $n \equiv r \bmod W$ and $a(n) = 0$ otherwise. Then on setting $P = \sqrt{5N}$ and $T(u) = \sum_{0 \leq n \leq u} a(n) e\left(n^2 t\right)$ for a given t in $\mathfrak{m}$ and integrating by parts we get

$$\sum_{n \equiv r \bmod W} 2n\beta(n^2) e(n^2 t) = \int_0^P 2u\beta(u^2) dT(u) \ll \sqrt{N} \sup_{0 \leq u \leq P} |T(u)|, \tag{61}$$

since $u \mapsto 2u\beta(u^2)$ is monotonic on each of the intervals on $[0, \frac{P}{\sqrt{2}}]$ and $(\frac{P}{\sqrt{2}}, P]$. By means of the classical Weyl squaring and differencing argument, given, for example, on page 17 of [6], and remarking that for any n, $a(n)a(n+h)$ is $a(n)$ when $W|h$ and is 0 otherwise, we obtain

$$|T(u)|^2 \leq \sum_{0 \leq n \leq u} a(n) + \sum_{\substack{1 \leq |h| \leq u, \\ W|h.}} \left| \sum_{n \in I(h)} a(n) e(2htn) \right|, \tag{62}$$

for all u, $0 \leq u \leq P$, where $I(h)$ is an interval of length $u - |h| \leq P$. If N is large enough so that $P \geq W$, the first term on the right hand side of (62) is $\leq \frac{2P}{W}$. Also, on using (2), page 40 of [6] to bound the sum over $n \in I(h)$ in (62) we get

$$\sum_{\substack{1 \leq |h| \leq u, \\ W|h.}} \left| \sum_{n \in I(h)} a(n) e(2htn) \right| \ll \sum_{\substack{1 \leq |k| \leq 2PW, \\ 2W^2|k.}} \inf\left(\frac{P}{W}, \frac{1}{\|kt\|}\right). \tag{63}$$

We estimate the right hand side of (63) ignoring the condition $2W^2|k$ and applying (9), page 41 of [6] with $\frac{a}{q}$ as in (60). This together with (62) gives

$$|T(u)|^2 \ll \frac{P^2}{q} + PW \log q + \frac{P}{W} + q \log q \ll \frac{P^2}{Q} \ll \frac{N}{Q}, \tag{64}$$

for all u, $0 \leq u \leq P$. Combining (64) with (61), (51) and applying the triangle inequality we obtain (59). From (59) and (56) we then conclude that for all N large enough, depending only on A and ℓ, we have

$$\int_{\mathfrak{m}} |\widehat{S}(t)|^{11}\, |\psi(t)|\ dt \ll \frac{N}{A^6} \int_0^1 |\widehat{S}(t)|^{11}\ dt\ \ll \frac{N|S|^9}{A^3} \ll \frac{|S|^{11}}{A}, \tag{65}$$

since $|S| \geq N^{\frac{1}{2}}/A$. It now follows that

$$\int_{\mathfrak{m}} \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t)\, dt\ \ll\ \frac{|S|^{11}}{A}. \tag{66}$$

3.2 The function ψ on a major arc

For any integers a, q and r, with $q > 0$, we set $G_r(a, q) = \sum_{0\leq m<q} e\left(\frac{a(r+mW)^2}{q}\right)$.

Lemma 3.1 *Let a and q be any integers satisfying* (53). *Then for all t in the major arc* $\mathfrak{M}(\frac{a}{q})$ *we have*

$$\psi(t) = \frac{1}{qW} \sum_{\substack{0\leq r<W,\\ (r,W)=1.}} G_r(a, q)\ \widehat{\widehat{\beta}}\left(t - \frac{a}{q}\right) + O(Q\phi(W)\sqrt{N}(\log N)^2). \tag{67}$$

Proof Let $\theta = t - \frac{a}{q}$ and $\eta(u) = 2u\beta(u^2)e(u^2\theta)$ for any real u. Then we have

$$\sum_{n\equiv r \bmod W} 2n\beta(n^2)e(n^2t) = \sum_j \eta(r + jW)e\left(\frac{a(r + jW)^2}{q}\right). \tag{68}$$

We split j on the right hand side of (68) into arithmetical progressions modulo q and sum both sides over r to get

$$\psi(t) = \sum_{\substack{0\leq r<W,\\ (r,W)=1.}} \sum_{0\leq m<q} e\left(\frac{a(r + mW)^2}{q}\right) \sum_k \eta(r + (m + kq)W). \tag{69}$$

Let $\varphi(u) = \eta(r + (m + uq)W)$ for all real u. Then φ is a continuous compactly supported function on $\mathbf{R}$. Its support is the union of two disjoint intervals on the interior of each of which φ is differentiable. Applying the Euler–Mclaurin formula to φ on each of these intervals and adding the results we obtain

$$\sum_k \varphi(k) = \int \varphi(u)du + O\left(\sup_u |\varphi(u)| + \int |\varphi'(u)|du\right). \tag{70}$$

The left hand side of (70) is the same as the sum over k in (69). From the definitions of η, β we have $\sup_u |\varphi(u)| \ll \sqrt{N}$. By means of the change of variable $(r + (m + uq)W)^2 \mapsto u$ we see that $\int \varphi(u)du = \frac{1}{qW}\widehat{\widehat{\beta}}(\theta)$. Finally, the change of variable $(r + (m + uq)W) \mapsto u$ gives $\int |\varphi'(u)|du = \int |\eta'(u)|du$. From the definition of $\eta(u)$ we have

$$\eta'(u) = 2\beta(u^2)e(u^2\theta) + 4u^2\beta'(u^2)e(u^2\theta) + 8\pi i u^2\theta\beta(u^2)e(u^2\theta), \tag{71}$$

which gives $|\eta'(u)| \ll \frac{N}{M}$, since $0 \leq \beta(u^2) \leq 1$, $|\beta'(u^2)| \ll \frac{1}{N}$, $|\theta| \leq \frac{1}{M}$ and $u^2 \ll N$ for u in the support of η'. Since the measure of this support is $\sqrt{5N}$, we conclude on recalling the definition of M that $\int |\varphi'(u)|du \ll \sqrt{N}(\log N)^2$. The preceding remarks together with (70) and (69) and the triangle inequality yield (67). □

The following lemma gives the key parts of Lemmas 5.2 and 5.3 of [2] in our context. The conclusions of this lemma explain the utility of the condition $(r, W) = 1$ in the definition (51) of $\psi(t)$.

Lemma 3.2 *Let a and q be integers satisfying (53) and r any integer coprime to W. Then we have*

(i) $G_r(a,q) = 0$ *unless* $q|2W$ *or there is a prime* $p > w$ *such that* $p|q$.

(ii) $\frac{1}{q}|G_r(a,q)| \leq \sqrt{\frac{2}{w}}$ *when* q *does not divide* $2W$.

Proof Following [2], we first note that for any integers c_0, c_1, c_2 and $d > 0$, if $P(X)$ is the quadratic polynomial $c_0X^2 + c_1X + c_2$ and $d_1, d_2 > 0$ are integers such that $d = d_1d_2$ and $d_2|c_0$ then

$$\sum_{0\leq m<d} e\left(\frac{P(m)}{d}\right) = \sum_{0\leq m_1<d_1} e\left(\frac{P(m_1)}{d}\right)\sum_{0\leq m_2<d_2} e\left(\frac{c_1m_2}{d_2}\right). \tag{72}$$

This is verified by remarking that the map $(m_1, m_2) \mapsto m_1 + m_2d_1$ is a bijection from $[0, d_1) \times [0, d_2)$ to $[0, d)$ and that if $m = m_1 + m_2d_1$, then $\frac{P(m)}{d} - \frac{P(m_1)}{d} - \frac{c_1m_2}{d_2} \in \mathbf{Z}$, because $d|c_0d_1$. Now the sum over m_2 on the right hand side of (72) is 0 unless $d_2|c_1$. Therefore the sum on the left hand side of (72) is also 0 unless $d_2|c_1$. Using this with $P(X) = a(r + WX)^2 = aW^2X^2 + 2WarX + ar^2$, $d_1 = \frac{q}{(q,W^2)}$ and $d_2 = (q, W^2)$, we deduce that for any integer r coprime to W we have $G_r(a,q) = 0$ unless $(q, W^2)|2W$, since ar is coprime to (q, W^2). If for any integer m and prime p, we write $v_p(m)$ for the exponent of p in the prime factorisation of m, then the condition $(q, W^2)|2W$ is equivalent to $\inf(v_p(q), 2v_p(W)) \leq v_p(2W)$ for all primes $p|2W$. From the definition of W in (50) we have $2v_p(W) > v_p(2W)$ for all primes $p|2W$. Consequently, $G_r(a,q) = 0$ unless $v_p(q) \leq v_p(2W)$ for all primes $p|2W$, which is the same as (i).

In light of the preceding paragraph, we may verify (ii) supposing that $(q, W^2)|2W$ and $\frac{q}{(q,W^2)} > w$. Let us set $Q(X) = \frac{aW^2}{(q,W^2)}X^2 + \frac{2War}{(q,W^2)}X$. Then with $P(X)$, d_1 and d_2 as above we obtain from (72) that

$$\frac{1}{q}|G_r(a,q)| = \frac{d_2}{q}\left|\sum_{0\leq m_1<d_1} e\left(\frac{Q(m_1)}{d_1}\right)\right| \leq \sqrt{\frac{2(q,W^2)}{q}} \leq \sqrt{\frac{2}{w}}, \tag{73}$$

on remarking that $\left|\sum_{0\leq m_1<d_1} e\left(\frac{Q(m_1)}{d_1}\right)\right| \leq \sqrt{2d_1}$, by the classical quadratic Weyl bound, applicable since the leading coefficient of $Q(X)$ and d_1 are coprime. □

3.3 The major arc contribution

In this subsection we complete the proof of (7). Let us first dispose of the first term in (55), which we denote here by T. Also, we shall write T_1 for

$$\sum_{\substack{0\leq r<W,\\(r,W)=1.}}\sum_{1\leq q\leq Q}\frac{1}{qW}\sum_{\substack{0\leq a<q,\\(a,q)=1.}} G_r(-a,q)\int_{\mathfrak{M}(\frac{a}{q})}\widehat{\beta}\left(t-\frac{a}{q}\right)\widehat{S}(t)^6\widehat{S}(-t)^5\,dt. \tag{74}$$

Then by substituting the complex conjugate of right hand side of (67) for $\psi(-t) = \overline{\psi}(t)$ in T and using the triangle inequality together with (58) we deduce that

$$T - T_1 \ll QW\sqrt{N}(\log N)^2\int_0^1|\widehat{S}(t)|^{11}dt \ll A^3QW|S|^9\sqrt{N}(\log N)^2. \tag{75}$$

If we now set

$$T(W)=\sum_{\substack{0\le r<W,\\(r,W)=1.}}\sum_{q|2W}\frac{1}{qW}\sum_{\substack{0\le a<q,\\(a,q)=1.}}G_r(-a,q)\int_{\mathfrak{M}(\frac{a}{q})}\widehat{\beta}\left(t-\frac{a}{q}\right)\widehat{S}(t)^6\widehat{S}(-t)^5\,dt. \quad (76)$$

then by (ii) of Lemma 3.2 combined with the triangle inequality and (58) we get

$$T_1-T(W)\ll\frac{\phi(W)\|\widehat{\beta}\|_\infty|S|^9A^3}{W\sqrt{w}}\ll\frac{A^3|S|^9N}{\sqrt{w}}, \quad (77)$$

since $\|\widehat{\beta}\|_\infty=\sup_{t\in\mathbf{R}}|\widehat{\beta}(t)|\le\frac{5N}{2}$. From (77), (75) and on recalling that $|S|\ge\frac{\sqrt{N}}{A}$ and $w=A^l\ge A^{12}$ we conclude that

$$T=T(W)+O\left(\frac{|S|^{11}}{A}\right), \quad (78)$$

when N is sufficiently large, depending only on A and l. Let us now estimate $T(W)$. When $q|2W$ we have $(r+mW)^2\equiv r^2$ modulo q for all integers m, since $2W|W^2$. Therefore we have $G_r(-a,q)=qe\left(-\frac{ar^2}{q}\right)$ when $q|2W$, for all $0\le a<q$. Furthermore, since $r\mapsto r+W$ is a bijection from the integers coprime to $2W$ in $[0,W)$ to those in $(W,2W]$ coprime to $2W$, we obtain

$$\frac{1}{qW}\sum_{\substack{0\le r<W,\\(r,W)=1.}}G_r(-a,q)=\frac{1}{2W}\sum_{\substack{0\le r<2W,\\(r,2W)=1.}}e\left(-\frac{ar^2}{q}\right) \quad (79)$$

for any $q|2W$ and all $0\le a<q$. Also, we have $\widehat{S}(t)^6\widehat{S}(-t)^5=\sum_{x\in S^{11}}e\,(f(x)t)$, with $f(x)$ as in (7). By means of the change of variable $t-\frac{a}{q}\mapsto t$ in the integrals in (76) we then see that

$$T(W)=\frac{1}{2W}\sum_{\substack{0\le r<2W,\\(r,2W)=1.}}\sum_{q|2W}\sum_{\substack{0\le a<q,\\(a,q)=1.}}\int_{-\frac{1}{M}}^{\frac{1}{M}}\widehat{\beta}\,(t)\sum_{x\in S^{11}}e\,(tf(x))\,e\left(\frac{a(f(x)-r^2)}{q}\right)dt. \quad (80)$$

Finally, on interchanging summations and remarking that

$$\frac{1}{2W}\sum_{q|2W}\sum_{\substack{0\le a<q,\\(a,q)=1.}}e\left(\frac{a(f(x)-r^2)}{q}\right)=\frac{1}{2W}\sum_{0\le a<2W}e\left(\frac{a(f(x)-r^2)}{2W}\right) \quad (81)$$

we conclude that the right hand side of (80) is the same as the left hand side of

$$\sum_{\substack{0\le r<2W,\\(r,2W)=1.}}\sum_{\substack{x\in S^{11},\\f(x)\equiv r^2\bmod 2\mathrm{W}}}\int_{-\frac{1}{M}}^{\frac{1}{M}}\widehat{\beta}(t)e\,(tf(x))\,dt\le\sum_{\substack{0\le r<2W,\\(r,2W)=1.}}\sum_{\substack{x\in S^{11},\\f(x)\equiv r^2\bmod 2\mathrm{W}}}1, \quad (82)$$

where we have used $|\int_{-\frac{1}{M}}^{\frac{1}{M}}\widehat{\beta}(t)e(tf(x))dt|\le\int_{\mathbf{R}}\widehat{\alpha}(t)dt=1$, since $|\widehat{\beta}(t)|=\widehat{\alpha}(t)$ for all $t\in\mathbf{R}$. For each class b in $\mathbf{Z}/2W\mathbf{Z}$, the number of r in $[0,2W)$ coprime to $2W$ and such that $r^2\equiv b$ modulo $2W$ is $2\tau(U)$. Then it follows from (82) and (80) that

$$T(W)=2\tau(U)|\{x\in S^{11}\mid f(x)\ \text{ an invertible square mod } 2W\}|. \quad (83)$$

Since $W=2U$ we obtain (7) on combining (83) with (78), (66) and recalling that (55) is the same as the integral in (52).

3.4 Proof of Theorem 1.3 completed

It remains only to bound (8) using Theorem 2.1. Let $\mathcal{Z}$ be the set of integers $n > 0$ such that $n^2 \in S$. The set $\mathcal{Z}$ is contained in $[M, 2M)$ with $M = \sqrt{N}$ and satisfies $|\mathcal{Z}| \geq \frac{M}{A}$ and $|\{z \in \mathcal{Z} | z \equiv a \bmod U\}| \leq \frac{MB}{U}$ with $B = 2$, when N is sufficiently large depending on A and l. Finally, let $I = S^9$ and for any $x = (x_1, x_2, \dots, x_9) \in S^9$ we set $c(x) = x_1 + \cdots + x_4 - x_5 - \cdots - x_9$. Then with $R_U(\mathcal{Z}, \mathbf{c})$ as in Theorem 2.1 we have that

$$|\{x \in S^{11} \mid f(x) \ \text{ an invertible square modulo } 2W\}| \leq |R_U(\mathcal{Z}, \mathbf{c})|, \tag{84}$$

since $U | 2W$. On combining the bound for $|R_U(\mathcal{Z}, \mathbf{c})|$ given by Theorem 2.1 with (84) and (7) we finally obtain (6), as required.

4 Monochromatic representation

Here we deduce Theorem 1.1 from Theorem 1.3. We first take up Lemma 1.2.

4.1 Proof of Lemma 1.2

A standard application of the Cauchy–Schwarz inequality gives $|mS|\, E_m(S) \geq |S|^{2m}$. Using (3) and $L \geq 2(mD + 1)D$ we then obtain

$$|mS| \geq \frac{L}{D} \geq \frac{mL}{k+1} + 2 \tag{85}$$

for any $k \geq mD$. We take k to be the integer $\lceil mD \rceil$. Since the set mS contained in the interval $(mN, mN + mL]$, its translate $mS - mN$ is contained in $[1, mL]$ and satisfies $(k+1)(|mS - mN| - 2) + 1 \geq mL$ on account of (85). Then by means of Theorem 2′, page 129 of [5] applied to the set $mS - mN$ we conclude that there are integers h, d and e with $1 \leq h \leq 2k + 1$ and $1 \leq d \leq k$ such that hmS contains the arithmetical progression

$$\mathcal{A} = hmN + \{(e+1)d, (e+2)d, \dots, (e+mL)d\}, \tag{86}$$

of mL terms and to the modulus d. Since $hmS \subseteq (hmN, hm(N + L)]$, each a in $\mathcal{A}$ satisfies

$$hmN < a \leq hm(N + L) \leq (2\lceil mD \rceil + 1)m(N + L). \tag{87}$$

Since $1 \leq d \leq \lceil mD \rceil$, there is an integer x in S coprime to d. Also, we have $x \leq mL$ since $x \leq N + L$, $L \geq N$ and $m \geq 2$. Therefore the number of terms in the arithmetical progression $\mathcal{A}$ is at least x and its modulus d is coprime to x. Consequently, $\mathcal{A}$ contains a complete system of residue classes modulo x and every integer n can be written as $n = a + rx$ with a in $\mathcal{A}$ and $r \in \mathbf{Z}$. For any integer $n \geq (2\lceil mD \rceil + 1)m(N + L)$ we have from $N \leq x$ and the lower bound for a in (87) that

$$0 \leq r = \frac{n-a}{x} \leq \frac{n}{N} - hm. \tag{88}$$

Since each $a \in \mathcal{A}$ is a sum of hm elements of S, the conclusion of the lemma now follows.

4.2 Proof of Theorem 1.1

Since $s(K)$ is increasing with K, it suffices to prove Theorem 1.1 for all K sufficiently large. For such a K, let $\cup_{1 \leq i \leq K} \mathfrak{Q}_i$ be a partition of the set of squares $\mathfrak{Q}$ into K disjoint subsets.

As in Sect. 1, let $\mathcal{B}$ be the set of squares of integers that are not divisible by any prime $p \leq B$, where $B = K^{13}$, and let $\mathcal{B}(N)$ denote $\mathcal{B} \cap (N, 4N]$, for a given integer $N \geq 1$. Then for all $N \geq N_0$, with N_0 depending only on K, we have by the principle of inclusion and exclusion and Mertens' formula as given by (3.27), page 70 of [9] that

$$|\mathcal{B}(N)| \geq N^{\frac{1}{2}} \prod_{p \leq B} \left(1 - \frac{1}{p}\right) - 2^B \geq \frac{N^{\frac{1}{2}}}{4 \log B} - 2^B \geq \frac{N^{\frac{1}{2}}}{100 \log K} \geq e^{e^2} K. \tag{89}$$

Let N be an integer $\geq N_0$. There is an i, $1 \leq i \leq K$, such that $\mathfrak{Q}_i \cap \mathcal{B}(N)$ contains at least $\frac{|\mathcal{B}(N)|}{K}$ of the elements of $\mathcal{B}(N)$. For such an i we set $S = \mathfrak{Q}_i \cap \mathcal{B}(N)$. Then S is a set of squares in $(N, 4N]$ with $|S| \geq \frac{N^{\frac{1}{2}}}{A}$, where $A = 100K \log K \geq e^{e^2}$ and no integer in S is divisible by a prime $p \leq A^{12}$, since $A^{12} \leq B$ when K is sufficiently large. It now follows from Theorem 1.3 that (3) holds with $m = 6$, $L = 3N$ and $D = A \exp\left(\frac{(3 \log 2 + o(1)) \log A}{\log \log A}\right)$. Since S contains an element of $\mathcal{B}$ and since $\lceil 6D \rceil \leq B$ when K is large enough, we may apply Lemma 1.2 to S to deduce that every integer $n \geq (288D + 72)N$ is a sum of no more than $\frac{n}{N}$ elements of S. In particular, there is a $C_1 > 0$ such that every integer $I(N) = ((288D+72)N, (288D+73)N]$ is a sum of at most $C_1 D$ squares all belonging to S and therefore to $\mathfrak{Q}_i$. Thus for all large enough N, every integer in the interval $I(N)$ can be expressed as a sum of no more than $C_1 D$ squares all of the same colour. On remarking that the interval $I(N)$ meets $I(N+1)$ for all large enough N, we obtain that $s(K) \leq C_1 D$. This yields the conclusion of Theorem 1.1 since $A = 100K \log K$ and therefore $C_1 D \leq K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$.

Acknowledgements We are grateful to Professor J. Brüdern for insisting to us that $s(K)$ ought to be essentially of order K. Our best thanks are due to Professors R. Balasubramanian, T.D. Browning and J. Oesterlé for their encouragement and a number of useful suggestions. We are obliged to Mr. K. Mallesham for going through various drafts of this article carefully and pointing out several errors. We sincerely thank the referee for the time spent on this article and for comments provided. This work was carried out under the CEFIPRA project 5401-1.

References

1. Akhilesh, P., Ramana, D.S.: A chromatic version of Lagrange's four squares theorem. Monatsh Math. **176**(1), 17–29 (2015)
2. Browning, T.D., Prendiville, S.M.: A transference approach to a Roth-type theorem in the squares. Int. Math. Res. Not. **2017**(7), 2219–2248 (2017)
3. Dubashi, D.P., Panconesi, A.: Concentration of Measure for the Analysis of Randomised Algorithms. Cambridge University Press, Cambridge (2009)
4. Hegyvári, N., Hennecart, F.: On monochromatic sums of squares and primes. J. Number Theory **124**, 314–324 (2007)
5. Lev, V.F.: Optimal representation by sumsets and subset sums. J. Number Theory **62**, 127–143 (1997)
6. Montgomery, H.L.: Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis. Regional Conference Series in Mathematics, vol. 84, p 220. CBMS, AMS, Providence, RI (1994)
7. Ramana, D.S., Ramaré, O.: Additive energy of dense sets of primes and monochromatic sums. Isr. J. Math. **199**, 955–974 (2014)
8. Ramaré, O., Ruzsa, I.Z.: Additive properties of dense subsets of sifted sequences. Journal de théorie des nombres de Bordeaux **13**(2), 559–581 (2001)
9. Rosser, B., Schoenfeld, I.: Approximate formulas for some functions of prime numbers. Ill. J. Math. **6**(1), 64–94 (1962)

10. Sárközy, A.: Unsolved problems in number theory. Period. Math. Hung. **42**, 17–35 (2001)
11. Schmidt, W.M.: Equations over Finite Fields: An Elementary Approach. Lecture Notes in Mathematics, vol. 536. Springer, Berlin, Heidelberg (1976)