Pacific Journal of Mathematics

INTEGRAL SOLUTIONS TO SYSTEMS OF DIAGONAL EQUATIONS

NICK ROME AND SHUNTARO YAMAGISHI

Volume 340 No. 1 January 2026

PACIFIC JOURNAL OF MATHEMATICS
Vol. 340, No. 1, 2026
https://doi.org/10.2140/pjm.2026.340.179

INTEGRAL SOLUTIONS TO SYSTEMS OF DIAGONAL EQUATIONS

NICK ROME AND SHUNTARO YAMAGISHI

We obtain an asymptotic formula for the number of integral solutions to a system of diagonal equations. We obtain an asymptotic formula for the number of solutions with variables restricted to smooth numbers as well. We improve the required number of variables compared to previous results by incorporating recent progress on Waring's problem and the resolution of the main conjecture in Vinogradov's mean value theorem.

1. Introduction

Consider the system of equations defined by

$$\begin{aligned} m_{1,1}x_1^d + \cdots + m_{1,n}x_n^d &= \mu_1, \\ &\vdots \\ m_{R,1}x_1^d + \cdots + m_{R,n}x_n^d &= \mu_R, \end{aligned} \tag{1-1}$$

which we write as $M\boldsymbol{x}^d = \boldsymbol{\mu}$, where $M = [m_{i,j}]_{\substack{1\leqslant i\leqslant R\\ 1\leqslant j\leqslant n}}$ is the coefficient matrix with integer entries and

$$\boldsymbol{x}^d = \begin{bmatrix} x_1^d \\ \vdots \\ x_n^d \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_R \end{bmatrix} \in \mathbb{Z}^R.$$

The system of diagonal equations (1-1) with $\boldsymbol{\mu} = \mathbf{0}$ was first studied by Davenport and Lewis, who established the following.

Theorem 1.1 [11, Lemma 32]. *Let $d \geqslant 3$ and $\boldsymbol{\mu} = \mathbf{0}$. Suppose that all n variables occur explicitly in the equations (1-1). Suppose that any linear combination, not identically zero, of the R rows of M contains more than $(2H + 3d - 1)R$ nonzero entries, where $H = \lfloor 3d \log Rd \rfloor$. Suppose the equations (1-1) have a nonsingular solution in every p-adic field, and further, if d is even, a real nonsingular solution. Then the equations (1-1) have infinitely many solutions in integers.*

MSC2020: primary 11P55; secondary 11D45, 11D72, 11P05.
Keywords: Hardy–Littlewood method, circle method, diagonal equation, Waring's problem, exponential sum.

In fact, they obtained an asymptotic formula for the number of solutions. Their main results [11, Theorems 1 and 2] are consequences of this theorem and require

$$n \geqslant \begin{cases} \lfloor 9R^2 d \log(3Rd) \rfloor & \text{if } d \text{ is odd,} \\ \lfloor 48R^2 d^3 \log(3Rd^2) \rfloor & \text{if } d \geqslant 4 \text{ is even,} \end{cases}$$

for the conclusions to hold. By incorporating the breakthrough on Waring's problem by Vaughan [17], Brüdern and Cook [6] improved the number of variables required to

$$n > n_0(d)R,$$

where $n_0(d) = 2d(\log d + O(\log\log d))$, under a suitable "rank condition" on the coefficient matrix M. They also obtained an asymptotic formula for the number of solutions but with variables restricted to smooth numbers, which in turn provided a lower bound for the number of solutions in positive integers.

Since the release of these two papers, there has been great progress regarding Waring's problem (for example, by Wooley [22; 23] and more recently by Wooley and Brüdern [9]) and also the resolution of the main conjecture in Vinogradov's mean value theorem (see the works by Bourgain, Demeter and Guth [2] and by Wooley [25; 27]). The purpose of this paper is to incorporate these recent progress to improve the required number of variables in both the setting of solutions in positive integers as in [11] and in the smooth numbers as in [6].

There have been a number of results regarding pairs of diagonal equations, in which the improvements have been achieved by making use of various developments in the theory of smooth Weyl sums. For example, the work of Parsell [14] on pairs of equations of small degrees, Parsell and Wooley [15] on pairs of quintic equations, and Brüdern and Wooley [7] on pairs of cubic equations. For larger systems of diagonal equations, there are the papers of Brüdern and Wooley [8] on systems of cubic equations, and of Brandes and Parsell [3] and Brandes and Wooley [4] on systems of equations involving different degrees. These works assume that the system is "highly nonsingular" which is to say that any $R \times R$ submatrix of the coefficient matrix is invertible. Our work is instead in line with [11] and [6] which hold for systems of diagonal equations of the same (arbitrary) degree with slightly less restrictive conditions on the underlying coefficient matrix.

For $X \geqslant 1$ and $\mathfrak{B} \subseteq \mathbb{N}$, we introduce the counting function

$$N(\mathfrak{B}; X) = \#\{\boldsymbol{x} \in (\mathfrak{B} \cap [1, X])^n : M\boldsymbol{x}^d = \boldsymbol{\mu}\}.$$

Instead of restricting the singularity of the variety defined by our system of equations, as in the work of Birch [1], we will require a condition on how well the underlying coefficient matrix can be partitioned.

Definition 1.2. For an $R \times n$ matrix A with $n \geqslant R$, we define $\Psi(A)$ to be the largest integer $\mathfrak{T}$ such that there exists

$$\{\mathfrak{D}_1, \ldots, \mathfrak{D}_{\mathfrak{T}}\},$$

where each $\mathfrak{D}_i$ is a linearly independent set of R columns of A and $\mathfrak{D}_i \cap \mathfrak{D}_j = \varnothing$ if $i \neq j$.

Remark 1.3. There are at least two ways to obtain lower bounds for $\Psi(M)$: by studying the ranks of submatrices (thanks to a result of Low, Pitman and Wolff [13, Lemma 1]) as in [12], or by algorithmically enumerating sets of R linearly independent columns as in [16]. In these two papers, lower bounds of the form a constant times $\frac{n}{R}$ were obtained for coefficient matrices related to $n \times n$ magic squares.

The following are the main results of this paper.

Theorem 1.4. *Let $d \geqslant 2$ and $T_{\text{int}}(d)$ be as recorded in Table 1. If $\Psi(M) \geqslant T_{\text{int}}(d)+1$, then there exists $\gamma > 0$ for which*

$$N(\mathbb{N}; X) = \mathfrak{S}\mathfrak{I}X^{n-dR} + O(X^{n-dR-\gamma}),$$

where $\mathfrak{S}$ is the singular series defined in (4-1) and $\mathfrak{I}$ is the singular integral defined in (4-5). We remark that $T_{\text{int}}(d) \leqslant \min\{2^d, d(d+1)\}$ for all $d \geqslant 2$.

Given $1 \leqslant Z \leqslant X$, we denote the Z-smooth numbers by

$$\mathscr{A}(X, Z) = \{x \in [1, X] \cap \mathbb{Z} : \text{prime } p | x \text{ implies } p \leqslant Z\}.$$

Theorem 1.5. *Let $d \geqslant 5$ and $T_{\text{smo}}(d)$ be as recorded in Table 2. If $\Psi(M) \geqslant T_{\text{smo}}(d)+1$, then for $\eta > 0$ sufficiently small, there exists $\gamma > 0$ such that*

$$N(\mathscr{A}(X, X^{\eta}); X) = \varrho\,(1/\eta)^n\, \mathfrak{S}\mathfrak{I}X^{n-dR} + O(X^{n-dR}(\log X)^{-\gamma}),$$

where $\mathfrak{S}$ is the singular series defined in (4-1), $\mathfrak{I}$ is the singular integral defined in (4-5) and ϱ is Dickman's function. We remark that $T_{\text{smo}}(d) \leqslant \lceil d(\log d + 4.20032)\rceil$ for all $d \geqslant 5$.

Remark 1.6. Instead of the condition $\Psi(M) \geqslant T_{\text{int}}(d)+1$ in Theorem 1.4, we may assume that there exists an $R \times (RT_{\text{int}}(d)+1)$ submatrix of M with the property that after removing any one of its columns it still contains $T_{\text{int}}(d)$ pairwise disjoint $R \times R$ invertible submatrices; the same holds for Theorem 1.5 with $T_{\text{smo}}(d)$ in place of $T_{\text{int}}(d)$. This is essentially the hypothesis assumed in [6], but for simplicity we assume the former condition; to assume the latter condition, one needs to slightly modify the proof of Proposition 3.3.

Remark 1.7. The proofs of Theorems 1.4 and 1.5 show that one may take γ to be any number in the intervals

$$\left(0, \frac{\lambda(d)}{(2R+4)(R+1)}\right) \quad \text{and} \quad \left(0, \frac{1}{2d(2R+4)(R+1)}\right],$$

respectively, where $\lambda(d)$ is defined in Lemma 2.1.

An immediate corollary is a lower bound for $N(\mathbb{N}; X)$ which requires a smaller value of $\Psi(M)$ than in Theorem 1.5.

Corollary 1.8. *Let $d \geqslant 3$ and suppose $\Psi(M) \geqslant T_{\mathrm{smo}}(d) + 1$. Then for $\boldsymbol{\mu} = \mathbf{0}$ such that $\mathfrak{S}\mathfrak{I} > 0$, we have*

$$N(\mathbb{N}; X) \gg X^{n-dR}.$$

We note that for a fixed choice of $\boldsymbol{\mu}$, by standard arguments, $\mathfrak{S} > 0$ if the equations (1-1) have a nonsingular solution in the ring of p-adic integers for every prime p, and $\mathfrak{I} > 0$ if the equations (1-1) have a nonsingular solution in $(\mathbb{R}_{>0})^n$.

The function $\theta(d) \in \{1, 2\}$ that appears in Table 1 is defined in (2-1). The values of $T_{\mathrm{int}}(d)$ are described in Lemma 2.2, and they correspond to the smallest known number of variables s required to produce an asymptotic formula for the number of representations of any sufficiently large natural number as a sum of s d-th powers (compare with [27, Corollary 14.7] for larger powers and [24, Theorem 4.1] for intermediate powers).

For $d \geqslant 13$, the values of $T_{\mathrm{smo}}(d)$ correspond to the best known values of $G(d)$, the least number of variables required to represent every sufficiently large natural number as a sum of d-th powers. Note the distinction that in this problem one asks only for the existence of a solution, and not the asymptotic formula for the number of solutions. For smaller values of d, $T_{\mathrm{smo}}(d)$ is slightly larger than the best known values of $G(d)$, which are $G(7) \leqslant 31$, $G(8) \leqslant 39$, $G(9) \leqslant 47$, $G(10) \leqslant 55$, $G(11) \leqslant 63$ and $G(12) \leqslant 72$, as found in [26]); these values are obtained by

d	2	3	4	5	6	7	8	9	$\geqslant 10$
$T_{\mathrm{int}}(d)$	4	8	15	23	34	47	61	78	$d^2 - d + 2\lfloor\sqrt{2d+2}\rfloor - \theta(d)$

Table 1. Values of T_{int} (see Theorem 1.4).

d	5	6	7	8	9	10	11	12
$T_{\mathrm{smo}}(d)$	19	25	33	41	49	57	65	73
d	13	14	15	16	17	18	19	$\geqslant 20$
$T_{\mathrm{smo}}(d)$	81	89	97	105	113	121	129	$\lceil d(\log d + 4.20032)\rceil$

Table 2. Values of T_{smo} (see Theorem 1.5).

considering solutions to the underlying Diophantine equations for which only some of the variables are restricted to the smooth numbers.

Notation. We use the standard abbreviations $e(z) = e^{2\pi iz}$ and $e_q(z) = e^{2\pi iz/q}$. Given a vector $\boldsymbol{a} = (a_1, \dots, a_R) \in \mathbb{Z}^R$, by $0 \leqslant \boldsymbol{a} \leqslant q$ we mean $0 \leqslant a_i \leqslant q$ for each $1 \leqslant i \leqslant R$. We also let $|\boldsymbol{\gamma}| = \max_{1 \leqslant i \leqslant R} |\gamma_i|$ for any $\boldsymbol{\gamma} \in \mathbb{R}^R$.

2. Preliminaries

Weyl sums. In this section, we collect two results that are key to proving Theorem 1.4. Both are consequences of the resolution of the main conjecture in Vinogradov's mean value theorem (by Bourgain, Demeter and Guth [2] and by Wooley [25; 27]).

Lemma 2.1. *Let $d \geqslant 2$. Let $\alpha \in \mathbb{R}$ and suppose that there exist $q \in \mathbb{N}$ and $a \in \mathbb{Z}$ with* $\gcd(q, a) = 1$ *such that $|\alpha - a/q| \leqslant q^{-2}$ and $q \leqslant X^d$. We define*

$$\lambda(d) = \begin{cases} \dfrac{1}{2^{d-1}} & \text{if } 2 \leqslant d \leqslant 5, \\ \dfrac{1}{d(d-1)} & \text{otherwise.} \end{cases}$$

Then

$$\Bigl|\sum_{1 \leqslant x \leqslant X} e(\alpha x^d)\Bigr| \ll X^{1+\varepsilon}(q^{-1} + X^{-1} + qX^{-d})^{\lambda(d)},$$

for any $\varepsilon > 0$.

Proof. The bound for $2 \leqslant d \leqslant 5$ is the classic Weyl's inequality [18, Lemma 2.4]. The other estimate for larger d is a consequence of the resolution of the main conjecture in Vinogradov's mean value theorem (cf. [5, Lemma 2.4]). □

Let us define

$$\theta(d) = \begin{cases} 1 & \text{if } 2d+2 \geqslant \lfloor\sqrt{2d+2}\rfloor^2 + \lfloor\sqrt{2d+2}\rfloor, \\ 2 & \text{if } 2d+2 < \lfloor\sqrt{2d+2}\rfloor^2 + \lfloor\sqrt{2d+2}\rfloor. \end{cases} \tag{2-1}$$

Lemma 2.2. *Let $d \geqslant 2$ and let s be a real number greater than or equal to*

$$T_{\text{int}}(d) = \min\Biggl\{2^d,\, d^2-d+2\lfloor\sqrt{2d+2}\rfloor-\theta(d),\, d^2+1-\max_{\substack{1\leqslant j\leqslant d-1\\ 2^j\leqslant d^2}}\Biggl\lceil\frac{dj-2^j}{d+1-j}\Biggr\rceil\Biggr\}.$$

Then

$$\int_0^1 \Bigl|\sum_{1\leqslant x\leqslant X} e(\alpha x^d)\Bigr|^s \mathrm{d}\alpha \ll X^{s-d+\varepsilon},$$

for any $\varepsilon > 0$.

Proof. The first bound, $s \geqslant 2^d$, is the classical version of Hua's lemma [18, Lemma 2.5], while the other two bounds are consequences of the resolution of the main conjecture in Vinogradov's mean value theorem. The second bound can be found in [27, Corollary 14.7], on noting that the bound for s_0 in the statement is given by

$$s_0 \leqslant \lfloor s_0 \rfloor + 1 \leqslant d^2 - d + 2\lfloor \sqrt{2d+2} \rfloor - \theta(d),$$

as explained in the proof. The third bound essentially follows from [24, Theorem 4.1]; it can be seen in the proof that the integral over the minor arcs satisfies $\ll X^{s-d+\varepsilon}$, while over the major arcs the same estimate follows by combining familiar estimates from the major arc analysis in the theory of Waring's problem (see [18, Section 4]). It can be verified that the values of $T_{\mathrm{int}}(d)$ are precisely as in Table 1 (see the paragraph following [27, Corollary 14.7] and the proof of [24, Theorem 4.1]). □

Smooth Weyl sums. In this section, we record some key estimates regarding the smooth Weyl sums needed to prove Theorem 1.5. Let $d \geqslant 3$. We let

$$f(\alpha; X, Z) = \sum_{x \in \mathscr{A}(X,Z)} e(\alpha x^d).$$

We first need two estimates from [6]. We begin with [6, Lemma 3] which is obtained by combining [17, Theorem 1.8] and [19, Lemma 7.2].

Lemma 2.3 [6, Lemma 3]. *Let $d \geqslant 3$ and $\varepsilon > 0$ be sufficiently small. Suppose $\eta > 0$ is sufficiently small. Then there exists $\gamma = \gamma(d) > 0$ such that given $\alpha \in [0, 1]$ one of the following two alternatives holds*:

(i) $|f(\alpha; X, X^\eta)| < X^{1-\gamma}$.

(ii) *There exist* $0 \leqslant a \leqslant q$, $\gcd(q, a) = 1$ *such that*

$$f(\alpha; X, X^\eta) \ll q^\varepsilon X (q + X^d |q\alpha - a|)^{-1/(2d)} (\log X)^3.$$

The following is [6, Lemma 4], which is a special case of [19, Lemma 8.5].

Lemma 2.4 [6, Lemma 4]. *Let $d \geqslant 3$. Suppose $\eta > 0$ is sufficiently small. Let $A_0 > 0$. Suppose* $\gcd(q, a) = 1$, $1 \leqslant q \leqslant (\log X)^{A_0}$ *and* $|q\alpha - a| \leqslant (\log X)^{A_0} X^{-d}$. *Then*

$$f(\alpha; X, X^\eta) \ll X q^\varepsilon (q + X^d |q\alpha - a|)^{-1/d},$$

for any $\varepsilon > 0$.

Given a real parameter $\mathfrak{L} \geqslant 1$, we define

$$\mathfrak{N}_{\mathfrak{L}} = \bigcup_{1 \leqslant q \leqslant \mathfrak{L}} \bigcup_{\substack{0 \leqslant a \leqslant q \\ \gcd(q,a)=1}} \{\theta \in [0, 1] : |q\theta - a| < \mathfrak{L} X^{-d}\}.$$

We make use of the previous two lemmas to prove the following.

Lemma 2.5. *Let $\delta > 0$, $A = 2d\delta$ and $\mathfrak{L} = (\log X)^A$. Suppose $\eta > 0$ is sufficiently small. If*

$$|f(\alpha; X, X^\eta)| > X(\log X)^{-\delta}$$

holds for $X \geqslant 1$ sufficiently large, then

$$\alpha \in \mathfrak{N}_{\mathfrak{L}}.$$

Proof. Since we are in alternative (ii) of Lemma 2.3, it follows that

$$X(\log X)^{-\delta} < Cq^\varepsilon X(q + X^d|q\alpha - a|)^{-1/(2d)}(\log X)^3,$$

for $\varepsilon > 0$ sufficiently small and some $C > 0$, which in turn implies

$$q^{1/(2d)} < Cq^\varepsilon(\log X)^{\delta+3} \quad \text{and} \quad (X^d|q\alpha - a|)^{1/(2d)} < Cq^\varepsilon(\log X)^{\delta+3}.$$

Therefore, by setting $A_0 = (\delta+3)4d$, we obtain $1 \leqslant q < (\log X)^{4d(\delta+3)}$, $\gcd(q, a) = 1$ and $|q\alpha - a| < (\log X)^{A_0}X^{-d}$. It then follows from Lemma 2.4 that

$$X(\log X)^{-\delta} < C_1q^\varepsilon X(q + X^d|q\alpha - a|)^{-1/d},$$

for some $C_1 = C_1(d, \delta, \varepsilon) > 0$, which in turn implies

$$q^{1/d} < C_1q^\varepsilon(\log X)^\delta \quad \text{and} \quad (X^d|q\alpha - a|)^{1/d} < C_1q^\varepsilon(\log X)^\delta.$$

Therefore, for $\mathfrak{L} = (\log X)^A$ with $A = 2d\delta$, it follows that $\alpha \in \mathfrak{N}_{\mathfrak{L}}$ as desired. □

Finally, we have the following mean value estimate from [9].

Lemma 2.6. *Let $d \geqslant 5$ and s be an integer such that $s \geqslant T_{\text{smo}}(d)$ as recorded in Table 2. Let $\eta > 0$ be sufficiently small and $1 \leqslant Z \leqslant X^\eta$. Then*

$$\int_0^1 |f(\alpha; X, Z)|^s \mathrm{d}\alpha \ll X^{s-d}.$$

Proof. A real number Δ_s is referred to as an admissible exponent (for d) if it has the property that, whenever $\varepsilon > 0$ and η is a positive number sufficiently small in terms of ε, d and s, then whenever $1 \leqslant Z \leqslant X^\eta$ and X is sufficiently large, one has

$$\int_0^1 |f(\alpha; X, Z)|^s \mathrm{d}\alpha \ll X^{s-d+\Delta_s+\varepsilon}.$$

Let us introduce the number

$$\tau(d) = \max_{w \in \mathbb{N}} \frac{d - 2\Delta_{2w}}{4w^2}.$$

Suppose that s is a real number with $s \geqslant 2$, and that the exponents Δ_u are admissible for $2 \leqslant u \leqslant s$. We define

$$\Delta_s^* = \min_{0 \leqslant t \leqslant s-2} (\Delta_{s-t} - t\tau(d)),$$

and refer to Δ_s^* as an admissible exponent for minor arcs. Let $d \geqslant 3$, $s \geqslant 2d+3$ and let Δ_s^* be an admissible exponent for minor arcs with $\Delta_s^* < 0$. Then applying [9, Theorem 6.1] with $Q = 1$ provides the bound

$$\int_0^1 |f(\alpha; X, Z)|^s \mathrm{d}\alpha \ll X^{s-d}.$$

We now follow the argument in the proof of [9, Theorem 6.2]. We assume that we have available an admissible exponent Δ_u for each positive number u (which we know we may assume as explained in [9, Section 2], and also see [9, (7.1)] for further information regarding Δ_u when u is even and $d \geqslant 4$). When $d \geqslant 4$, we define

$$G_0(d) = \min_{v \geqslant 2} \left(v + \frac{\Delta_v}{\tau(d)} \right). \tag{2-2}$$

Suppose that $d \geqslant 4$ and $s \geqslant \max\{\lfloor G_0(d) \rfloor + 1, 2d+3\}$. Then there exists a positive number v with $v \geqslant 2$ and an admissible exponent Δ_v for which the exponent Δ_s^* is admissible for minor arcs, where

$$\Delta_s^* = \Delta_v - (s-v)\tau(d) = -\tau(d)(s - G_0(d)) < 0.$$

For $d \geqslant 14$, the value of $T_{\mathrm{smo}}(d)$ is precisely the value of $\lfloor G_0(d) \rfloor + 1$ found in the proofs of [9, Theorems 1.1 and 1.3], which can be seen to be greater than $2d+3$. For smaller d, we follow the proof of [9, Theorem 8.1] and compute $G_0(d)$ using the expression

$$T(d) = \frac{4w^2}{d - 2\Delta_{2w}}$$

for a suitably chosen value of w. Since $\tau(d) \geqslant T(d)^{-1}$, we clearly have

$$G_0(d) \leqslant v' + \Delta_{v'} T(d)$$

for any choice of $v' \geqslant 2$. We use the values of w and the corresponding admissible exponents Δ_{2w} recorded in Vaughan–Wooley [21, §9–15]. Here, the exponents λ_w of [21] are related to Δ_{2w} via the formula $\Delta_{2w} = \lambda_w - 2w + d$. Table 3 shows the chosen values of w and v used to compute $2v + \Delta_{2v} T(d)$.

For $d = 5$ and 6, the necessary data come instead from the appendix of [20] and we choose the values shown in Table 4. One readily observes that

$$T_{\mathrm{smo}}(d) = \lfloor 2v + \Delta_{2v} T(d) \rfloor + 1 \geqslant \lfloor G_0(d) \rfloor + 1$$

for the listed values of d. □

d	w	Δ_{2w}	$T(d)$	v	Δ_{2v}	$2v+\Delta_{2v}T(d)$
7	6	2.0143820	48.46467935	16	0.0105382	32.51073048
8	7	2.3105992	58.00873304	19	0.0473193	40.74493264
9	8	2.6039271	67.50795289	22	0.0727119	48.90863152
10	9	2.8945712	76.94394605	25	0.0895832	56.89288491
11	10	3.1849727	86.39206976	28	0.1020502	64.81632800
12	11	3.4700805	95.65521749	31	0.1118679	72.70074830
13	12	3.7557170	104.94544480	35	0.1010835	80.60825287

Table 3

d	w	Δ_{2w}	$T(d)$	v	Δ_{2v}	$2v+\Delta_{2v}T(d)$
5	4	1.4386563	30.15045927	8	0.0773627	18.33252094
6	5	1.7246965	39.20635362	12	0.0000000	24.00000000

Table 4

3. The Hardy–Littlewood circle method

Let $\mathfrak{B} = \mathbb{N}$ or $\mathscr{A}(X, X^\eta)$. Throughout the remainder of the paper, unless stated otherwise, we assume $d \geqslant 2$ if $\mathfrak{B} = \mathbb{N}$, and $d \geqslant 3$ if $\mathfrak{B} = \mathscr{A}(X, X^\eta)$. Our main tool to study $N(\mathfrak{B}; X)$ is the Hardy–Littlewood circle method and the key input are the estimates regarding the associated exponential sums. In contrast to the exposition in [6], we find it more natural to index our exponential sums by the columns of the corresponding coefficient matrix. For $\boldsymbol{\theta} \in [0, 1]^R$ and $\boldsymbol{c} \in \mathrm{Col}(M)$, we introduce the exponential sum

$$S_{\boldsymbol{c}}(\boldsymbol{\theta}) = S_{\boldsymbol{c}}(\mathfrak{B}; \boldsymbol{\theta}) = \sum_{x \in \mathfrak{B} \cap [1, X]} e(\boldsymbol{c} \cdot \boldsymbol{\theta} x^d).$$

Then

$$N(\mathfrak{B}; X) = \int_{[0,1]^R} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} S_{\boldsymbol{c}}(\mathfrak{B}; \boldsymbol{\theta}) \cdot e\Bigl(-\sum_{i=1}^{R} \mu_i \theta_i\Bigr) \, \mathrm{d}\boldsymbol{\theta}. \tag{3-1}$$

We set

$$\mathfrak{L} = \begin{cases} X^\delta & \text{if } \mathfrak{B} = \mathbb{N}, \\ (\log X)^A & \text{if } \mathfrak{B} = \mathscr{A}(X, X^\eta), \end{cases} \tag{3-2}$$

where $\delta, A > 0$ are to be chosen in due course. We define the major arcs

$$\mathfrak{M}_{\mathfrak{L}} = \bigcup_{1 \leqslant q \leqslant \mathfrak{L}} \bigcup_{\substack{\boldsymbol{a} \in \mathbb{Z}^R \\ 0 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(q, \boldsymbol{a}) = 1}} \{\boldsymbol{\theta} \in [0, 1]^R : |q\theta_i - a_i| < \mathfrak{L} X^{-d} \ (1 \leqslant i \leqslant R)\},$$

and the minor arcs

$$\mathfrak{m}_{\mathfrak{L}} = [0, 1]^R \setminus \mathfrak{M}_{\mathfrak{L}}.$$

From here on out, we will use the following notation for simplicity.

Definition 3.1. We let T be a natural number such that $\Psi(M) \geqslant T$.

The minor arc estimate. The following lemma allows us to understand when a phase of the form $\boldsymbol{c}\cdot\boldsymbol{\theta}$ belongs to $[0, 1]\backslash\mathfrak{N}_{\mathfrak{L}}$. Given a set of vectors $\mathfrak{D} = \{\boldsymbol{c}_1, \dots, \boldsymbol{c}_R\}$, we denote by $M(\mathfrak{D}) = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_R]$ the matrix with these vectors as its columns.

Lemma 3.2. *Let $\mathfrak{D} = \{\boldsymbol{c}_1, \dots, \boldsymbol{c}_R\} \subseteq \mathrm{Col}(M)$ be a set of R linearly independent vectors. Suppose $X \geqslant 1$ is sufficiently large. If $\boldsymbol{c}_i \cdot \boldsymbol{\theta} \in \mathfrak{N}_{\mathfrak{L}^{1/(R+1)}}$ for all $1 \leqslant i \leqslant R$, then $\boldsymbol{\theta} \in \mathfrak{M}_{\mathfrak{L}}$.*

Proof. We have

$$\begin{bmatrix} q_1\boldsymbol{c}_1\cdot\boldsymbol{\theta} \\ \vdots \\ q_R\boldsymbol{c}_R\cdot\boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} a_1 + E_1 \\ \vdots \\ a_R + E_R \end{bmatrix}$$

for some $1 \leqslant q_i \leqslant \mathfrak{L}^{1/(R+1)}$ and $1 \leqslant a_i \leqslant q_i$ such that $\gcd(a_i, q_i) = 1$ and $|E_i| < \mathfrak{L}^{1/(R+1)} X^{-d}$ for each $1 \leqslant i \leqslant R$. Then

$$q_1 \cdots q_R M(\mathfrak{D})^t \boldsymbol{\theta} = \begin{bmatrix} q_1 \cdots q_R (a_1 + E_1)/q_1 \\ \vdots \\ q_1 \cdots q_R (a_R + E_R)/q_R \end{bmatrix},$$

and the result follows by multiplying both sides of the equation by the inverse of $M(\mathfrak{D})^t$ on the left and simplifying the resulting equation. □

We are now ready to bound the contribution from the minor arcs.

Proposition 3.3. *Suppose that*

$$T \geqslant \begin{cases} T_{\mathrm{int}}(d) + 1 & \text{if } \mathfrak{B} = \mathbb{N}, \\ T_{\mathrm{smo}}(d) + 1 & \text{if } \mathfrak{B} = \mathscr{A}(X, X^{\eta}). \end{cases}$$

Suppose $\eta > 0$ is sufficiently small. Then, we may choose $\delta, A > 0$ such that there exists $\gamma > 0$ satisfying

$$\int_{\mathfrak{m}_{\mathfrak{L}}} \prod_{\boldsymbol{c}\in\mathrm{Col}(M)} |S_{\boldsymbol{c}}(\mathfrak{B}; \boldsymbol{\theta})| \, \mathrm{d}\boldsymbol{\theta} \ll X^{n-dR} \mathfrak{L}^{-\gamma}.$$

Proof. Let

$$\mathfrak{D}_1, \dots, \mathfrak{D}_T$$

be pairwise disjoint sets of R linearly independent columns of M. We begin by applying Lemma 3.2 with $\mathfrak{D}_T = \{\boldsymbol{c}_1, \dots, \boldsymbol{c}_R\}$. Given $\boldsymbol{\theta} \in \mathfrak{m}_{\mathfrak{L}}$, it follows from

Lemma 3.2 that there exists $1 \leqslant i \leqslant R$ such that $\boldsymbol{c}_i \cdot \boldsymbol{\theta} \notin \mathfrak{N}_{\mathfrak{L}'}$ with $\mathfrak{L}' = \mathfrak{L}^{1/(R+1)}$. Extracting the contribution from this column, we have the bound

$$\int_{\mathfrak{m}_{\mathfrak{L}}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} |S_{\boldsymbol{c}}(\boldsymbol{\theta})| \,\mathrm{d}\boldsymbol{\theta} \leqslant X^{R-1} \sup_{\alpha \in [0,1] \setminus \mathfrak{N}_{\mathfrak{L}'}} \Bigg| \sum_{x \in \mathfrak{B} \cap [1,X]} e(\alpha x^d) \Bigg| \int_{[0,1]^R} \prod_{\boldsymbol{c} \in \mathrm{Col}(M) \setminus \mathfrak{D}_T} |S_{\boldsymbol{c}}(\boldsymbol{\theta})| \,\mathrm{d}\boldsymbol{\theta}.$$

Bounding the contribution from any column which does not belong to $\mathfrak{D}_1, \ldots, \mathfrak{D}_{T-1}$ trivially gives a bound for the integral over the minor arcs of

$$X^{n-(T-1)R-1} \sup_{\alpha \in [0,1] \setminus \mathfrak{N}_{\mathfrak{L}'}} \Bigg| \sum_{x \in \mathfrak{B} \cap [1,X]} e(\alpha x^d) \Bigg| \int_{[0,1]^R} \prod_{\ell=1}^{T-1} \prod_{\boldsymbol{c} \in \mathfrak{D}_\ell} |S_{\boldsymbol{c}}(\boldsymbol{\theta})| \,\mathrm{d}\boldsymbol{\theta}.$$

Applying Hölder's inequality, this is bounded by

$$X^{n-(T-1)R-1} \sup_{\alpha \in [0,1] \setminus \mathfrak{N}_{\mathfrak{L}'}} \Bigg| \sum_{x \in \mathfrak{B} \cap [1,X]} e(\alpha x^d) \Bigg| \prod_{\ell=1}^{T-1} \left(\int_{[0,1]^R} \prod_{\boldsymbol{c} \in \mathfrak{D}_\ell} |S_{\boldsymbol{c}}(\boldsymbol{\theta})|^{T-1} \,\mathrm{d}\boldsymbol{\theta} \right)^{1/(T-1)}.$$

Since the columns in $\mathfrak{D}_\ell$ are linearly independent, by a linear change of variables we obtain

$$\int_{[0,1]^R} \prod_{\boldsymbol{c} \in \mathfrak{D}_\ell} |S_{\boldsymbol{c}}(\boldsymbol{\theta})|^{T-1} \,\mathrm{d}\boldsymbol{\theta} \ll \prod_{i=1}^{R} \int_0^1 \Bigg| \sum_{x \in \mathfrak{B} \cap [1,X]} e(\lambda_i x^d) \Bigg|^{T-1} \mathrm{d}\lambda_i,$$

for each $1 \leqslant \ell \leqslant T-1$. We may now apply the bounds from Lemmas 2.1 and 2.2 or from Lemmas 2.5 and 2.6, depending on $\mathfrak{B}$, to conclude the proof. □

Major arc analysis. We define

$$\mathfrak{M}_{\mathfrak{L}}^{+} = \bigcup_{1 \leqslant q \leqslant \mathfrak{L}} \bigcup_{\substack{\boldsymbol{a} \in \mathbb{Z}^R \\ 0 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(q,\boldsymbol{a})=1}} \left\{ \boldsymbol{\theta} \in [0,1]^R : |q\theta_i - a_i| < q\mathfrak{L}X^{-d} \ (1 \leqslant i \leqslant R) \right\},$$

which clearly satisfies $\mathfrak{M}_{\mathfrak{L}} \subseteq \mathfrak{M}_{\mathfrak{L}}^{+}$. For any $q \in \mathbb{N}$, $a \in \mathbb{Z}$ and $\beta \in \mathbb{R}$, we introduce the standard notation

$$S(q,a) = \sum_{1 \leqslant x \leqslant q} e_q(ax^d) \quad \text{and} \quad I(\beta) = \int_0^1 e(\beta \xi^d) \,\mathrm{d}\xi.$$

Lemma 3.4. *Suppose that $q \in \mathbb{N}$, $a \in \mathbb{Z}$ and $\beta = \alpha - a/q$. Then*

$$\sum_{1 \leqslant x \leqslant X} e(\alpha x^d) = Xq^{-1} S(q,a) I(X^d \beta) + O\left(\frac{q}{\gcd(q,a)} (1 + X^d |\beta|) \right).$$

Proof. The statement with the additional hypothesis $\gcd(q,a)=1$ follows from [18, Theorem 4.1]. Suppose $\gcd(q,a)=g$ and let $q_0=q/g$ and $a_0=a/g$. Then

$$\begin{aligned} q^{-1}S(q,a) &= q^{-1}\sum_{1\leqslant x\leqslant q} e_q(ax^d) = q^{-1}\sum_{1\leqslant x\leqslant q} e_{q_0}(a_0x^d) \\ &= q^{-1}g\sum_{1\leqslant x\leqslant q_0} e_{q_0}(a_0x^d) = q_0^{-1}S(q_0,a_0). \end{aligned}$$

Therefore, we see that we may remove the coprimality condition. □

For the smooth Weyl sums we have the following.

Lemma 3.5. *Suppose that* $1\leqslant q\leqslant Z$, $a\in\mathbb{Z}$ *and* $\beta=\alpha-a/q$. *Then*

$$f(\alpha;X,Z)=q^{-1}S(q,a)w(\beta)+O\left(\frac{qX}{\gcd(q,a)\log X}(1+X^d|\beta|)\right),$$

where

$$w(\beta)=\sum_{Z^d<m\leqslant X^d}\frac{1}{d}m^{1/d-1}\varrho\left(\frac{\log m}{d\log Z}\right)e(\beta m)$$

and ϱ *is Dickman's function* (see [17, p. 53], for example).

Proof. The statement with the additional hypothesis $\gcd(q,a)=1$ is precisely [17, Lemma 5.4]. The coprimality condition may be removed in the same way as in the proof of Lemma 3.4. □

Lemma 3.6. *Let* $|\beta|<\mathfrak{L}X^{-d}$ *and* w *be as in Lemma 3.5. Then*

$$w(\beta)=\varrho\left(\frac{d\log X}{d\log Z}\right)XI(X^d\beta)+O\left(\frac{X}{\log Z}+Z\right).$$

Proof. Let us denote

$$P(y)=\sum_{Z^d<m\leqslant y}\frac{1}{d}m^{1/d-1}e(\beta m).$$

Then, by summation by parts, it follows that

$$\begin{aligned} w(\beta) &= \sum_{Z^d<m\leqslant X^d}\frac{1}{d}m^{1/d-1}e(\beta m)\,\varrho\left(\frac{\log m}{d\log Z}\right) \\ &= P(X^d)\,\varrho\left(\frac{d\log X}{d\log Z}\right)+O\left(1+\int_{Z^d}^{X^d}|P(y)|\frac{1}{y\log Z}\mathrm{d}y\right). \end{aligned}$$

Since $|P(y)|\ll y^{1/d}$, we have

$$\int_{Z^d}^{X^d}|P(y)|\frac{1}{y\log Z}dy\ll\frac{1}{\log Z}\int_{Z^d}^{X^d}y^{1/d-1}\mathrm{d}y\ll\frac{X}{\log Z}.$$

Therefore, we obtain

$$w(\beta)=\varrho\left(\frac{d\log X}{d\log Z}\right)\sum_{1\leqslant m\leqslant X^d}\frac{1}{d}m^{1/d-1}e(\beta m)+O\left(\frac{X}{\log Z}+Z\right).$$

By the mean value theorem, we obtain

$$\begin{aligned}\frac{1}{d}\sum_{1\leqslant m\leqslant X^d} & m^{1/d-1}e(\beta m)\\ &=\frac{1}{d}\int_1^{X^d}x^{1/d-1}e(\beta x)\,\mathrm{d}x+O\left(1+\sum_{1\leqslant m\leqslant X^d}m^{1/d-1}(m^{-1}+|\beta|)\right)\\ &=\int_0^X e(\beta t^d)\,\mathrm{d}t+O(1)=X\int_0^1 e(X^d\beta y^d)\,\mathrm{d}y+O(1)\\ &=XI(X^d\beta)+O(1).\end{aligned}$$

□

Let us now combine the above three lemmas in the following convenient manner.

Lemma 3.7. *Let $\eta>0$ be sufficiently small and*

$$C_{\mathfrak{B}}=\begin{cases}1 & \text{if } \mathfrak{B}=\mathbb{N},\\ \varrho(1/\eta) & \text{if } \mathfrak{B}=\mathscr{A}(X,X^\eta).\end{cases}\tag{3-3}$$

Let $\delta, A>0$ be sufficiently small. Suppose that $0\leqslant a\leqslant q\leqslant\mathfrak{L}$, $\beta=\alpha-a/q$ and $|\beta|<\mathfrak{L}X^{-d}$. Then

$$\left|\sum_{x\in\mathfrak{B}\cap[1,X]}e(\alpha x^d)-C_{\mathfrak{B}}Xq^{-1}S(q,a)I(X^d\beta)\right|\ll\begin{cases}\mathfrak{L}^2 & \text{if } \mathfrak{B}=\mathbb{N},\\ \dfrac{X\mathfrak{L}^2}{\log X} & \text{if } \mathfrak{B}=\mathscr{A}(X,X^\eta).\end{cases}$$

We define the truncated singular series

$$\mathfrak{S}(B)=\sum_{1\leqslant q\leqslant B}q^{-n}\sum_{\substack{1\leqslant\boldsymbol{a}\leqslant q\\ \gcd(q,\boldsymbol{a})=1}}\prod_{\boldsymbol{c}\in\mathrm{Col}(M)}S(q,\boldsymbol{a}\cdot\boldsymbol{c})\cdot e_q\left(-\sum_{i=1}^R\mu_i a_i\right)$$

for any $B\geqslant 1$, and the truncated singular integral

$$\mathfrak{J}(B)=\int_{|\boldsymbol{\gamma}|<B}\prod_{\boldsymbol{c}\in\mathrm{Col}(M)}I(\boldsymbol{\gamma}\cdot\boldsymbol{c})\cdot e\left(-\frac{1}{X^d}\sum_{i=1}^R\mu_i\gamma_i\right)\mathrm{d}\boldsymbol{\gamma}$$

for any $B>0$.

Proposition 3.8. *Let $\eta>0$ be sufficiently small and $C_{\mathfrak{B}}$ as in (3-3). Then*

$$\int_{\mathfrak{M}^+_{\mathfrak{L}}}\prod_{\boldsymbol{c}\in\mathrm{Col}(M)}S_{\boldsymbol{c}}(\boldsymbol{\theta})\cdot e\left(-\sum_{i=1}^R\mu_i\theta_i\right)\mathrm{d}\boldsymbol{\theta}=C_{\mathfrak{B}}^nX^{n-dR}\mathfrak{S}(\mathfrak{L})\mathfrak{J}(\mathfrak{L})+O(X^{n-dR}\mathfrak{L}^{-1}).$$

Proof. First, if $\boldsymbol{\theta} \in \mathfrak{M}_{\mathfrak{L}}^{+}$ then there exist $0 \leqslant \boldsymbol{a} \leqslant q$ such that $\gcd(q, \boldsymbol{a}) = 1$ and

$$\left| \boldsymbol{c} \cdot \boldsymbol{\theta} - \frac{\boldsymbol{c} \cdot \boldsymbol{a}}{q} \right| < C \mathfrak{L} X^{-d},$$

where $C > 0$ is a constant depending only on $\boldsymbol{c}$; therefore, $\boldsymbol{c} \cdot \boldsymbol{\theta}$, reduced modulo 1, satisfies the hypotheses of Lemma 3.7 with $\boldsymbol{c} \cdot \boldsymbol{\gamma}$ and $C\mathfrak{L}$ in place of β and $\mathfrak{L}$, respectively. Thus we may apply Lemma 3.7 to $S_{\boldsymbol{c}}(\boldsymbol{\theta})$ for any $\boldsymbol{c} \in \mathrm{Col}(M)$ and $\boldsymbol{\theta} \in \mathfrak{M}_{\mathfrak{L}}^{+}$. The measure of $\mathfrak{M}_{\mathfrak{L}}^{+}$ is at most $\mathfrak{L}^{2R+1} X^{-dR}$ and thus integrating the error term coming from applying Lemma 3.7 to $\prod_{\boldsymbol{c} \in \mathrm{Col}(M)} S_{\boldsymbol{c}}(\boldsymbol{\theta})$ gives a total error of size

$$\begin{cases} O(X^{n-dR-1} \mathfrak{L}^{2R+3}) & \text{if } \mathfrak{B} = \mathbb{N}, \\ O\left(X^{n-dR} \dfrac{\mathfrak{L}^{2R+3}}{\log X} \right) & \text{if } \mathfrak{B} = \mathscr{A}(X, X^{\eta}). \end{cases}$$

In the former case, the saving is $X^{1-(2R+3)\delta}$, which is greater than X^{δ} when $\delta \leqslant 1/(2R+4)$. Similarly, in the latter case, one saves $(\log X)^{1-A(2R+3)}$ which is again sufficient under the condition that $A \leqslant 1/(2R+4)$.

As a result we have

$$\begin{aligned}
&\int_{\mathfrak{M}_{\mathfrak{L}}^{+}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} S_{\boldsymbol{c}}(\boldsymbol{\theta}) \cdot e\left(-\sum_{i=1}^{R} \mu_i \theta_i \right) \mathrm{d}\boldsymbol{\theta} \\
&= C_{\mathfrak{B}}^{n} X^{n} \sum_{1 \leqslant q \leqslant \mathfrak{L}} q^{-n} \sum_{\substack{1 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(q, \boldsymbol{a}) = 1}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} S_{\boldsymbol{c}}(\boldsymbol{a}/q) \cdot e_q\left(-\sum_{i=1}^{R} \mu_i a_i \right) \\
&\qquad \times \int_{|\boldsymbol{\gamma}| < \mathfrak{L} X^{-d}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} I(X^{d} \boldsymbol{c} \cdot \boldsymbol{\gamma}) \cdot e\left(-\sum_{i=1}^{R} \mu_i \gamma_i \right) \mathrm{d}\boldsymbol{\gamma} + O(X^{n-dR} \mathfrak{L}^{-1}) \\
&= C_{\mathfrak{B}}^{n} X^{n} \mathfrak{S}(\mathfrak{L}) \int_{|\boldsymbol{\gamma}| < \mathfrak{L} X^{-d}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} I(X^{d} \boldsymbol{c} \cdot \boldsymbol{\gamma}) \cdot e\left(-\sum_{i=1}^{R} \mu_i \gamma_i \right) \mathrm{d}\boldsymbol{\gamma} + O(X^{n-dR} \mathfrak{L}^{-1}) \\
&= C_{\mathfrak{B}}^{n} X^{n-dR} \mathfrak{S}(\mathfrak{L}) \int_{|\boldsymbol{\gamma}| < \mathfrak{L}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} I(\boldsymbol{c} \cdot \boldsymbol{\gamma}) \cdot e\left(-\frac{1}{X^{d}} \sum_{i=1}^{R} \mu_i \gamma_i \right) \mathrm{d}\boldsymbol{\gamma} + O(X^{n-dR} \mathfrak{L}^{-1}) \\
&= C_{\mathfrak{B}}^{n} X^{n-dR} \mathfrak{S}(\mathfrak{L}) \mathfrak{J}(\mathfrak{L}) + O(X^{n-dR} \mathfrak{L}^{-1}),
\end{aligned}$$

which completes the claim. □

4. Singular series and singular integral

Let us denote

$$A(q) = q^{-n} \sum_{\substack{1 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(q, \boldsymbol{a}) = 1}} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} S(q, \boldsymbol{a} \cdot \boldsymbol{c}) \cdot e_q\left(-\sum_{i=1}^{R} \mu_i a_i \right).$$

We define the singular series as

$$\mathfrak{S} = \sum_{q=1}^{\infty} A(q) = \lim_{B\to\infty} \mathfrak{S}(B). \tag{4-1}$$

In the following lemma, we bound the quantity $A(q)$ in order to show that the singular series does indeed converge absolutely.

Lemma 4.1. *Suppose $T > d$ and let $q \in \mathbb{N}$. Then*

$$A(q) \ll q^{-T/d+1+\varepsilon},$$

for any $\varepsilon > 0$.

Proof. By [10, Lemma 6.4], we have

$$\begin{aligned} |S(q, \boldsymbol{a}\cdot\boldsymbol{c})| &= \gcd(q, \boldsymbol{a}\cdot\boldsymbol{c}) \left| S\left(\frac{q}{\gcd(q, \boldsymbol{a}\cdot\boldsymbol{c})}, \frac{\boldsymbol{a}}{\gcd(q, \boldsymbol{a}\cdot\boldsymbol{c})}\right)\right| \\ &\ll \gcd(q, \boldsymbol{a}\cdot\boldsymbol{c}) \left(\frac{q}{\gcd(q, \boldsymbol{a}\cdot\boldsymbol{c})}\right)^{1-1/d}. \end{aligned}$$

We know that there exist pairwise disjoint sets $\mathfrak{D}_1, \dots, \mathfrak{D}_T$ of R linearly independent columns of M. Applying Hölder's inequality, it follows that

$$\begin{aligned} |A(q)| &\leqslant q^{-TR} \sum_{\substack{1\leqslant \boldsymbol{a}\leqslant q \\ \gcd(q,\boldsymbol{a})=1}} \prod_{\ell=1}^{T} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} |S(q, \boldsymbol{a}\cdot\boldsymbol{c})| \\ &\ll q^{-TR} \prod_{\ell=1}^{T} \biggl(\sum_{\substack{1\leqslant \boldsymbol{a}\leqslant q \\ \gcd(q,\boldsymbol{a})=1}} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} |S(q, \boldsymbol{a}\cdot\boldsymbol{c})|^T \biggr)^{1/T} \\ &\ll q^{-TR/d} \prod_{\ell=1}^{T} \biggl(\sum_{\substack{1\leqslant \boldsymbol{a}\leqslant q \\ \gcd(q,\boldsymbol{a})=1}} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} \gcd(q, \boldsymbol{a}\cdot\boldsymbol{c})^{T/d} \biggr)^{1/T}. \end{aligned} \tag{4-2}$$

Let us suppose for the time being that we have

$$\sum_{\substack{1\leqslant \boldsymbol{a}\leqslant q \\ \gcd(q,\boldsymbol{a})=1}} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} \gcd(q, \boldsymbol{a}\cdot\boldsymbol{c})^{T/d} \ll q^{(R-1)T/d+1+\varepsilon}, \tag{4-3}$$

for every $1 \leqslant \ell \leqslant T$ and any $\varepsilon > 0$. Then, by substituting this estimate into (4-2), we obtain

$$|A(q)| \ll q^{-TR/d} q^{(R-1)T/d+1+\varepsilon} = q^{-T/d+1+\varepsilon},$$

as desired.

We now prove the estimate (4-3). Write $\mathfrak{D}_\ell = \{\boldsymbol{c}_1, \dots, \boldsymbol{c}_R\}$. First we have

$$\sum_{\substack{1 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(\boldsymbol{a},q)=1}} \prod_{\boldsymbol{c} \in \mathfrak{D}_\ell} \gcd(q, \boldsymbol{a}\cdot\boldsymbol{c})^{T/d} \tag{4-4}$$

$$\leqslant \sum_{d_1|q,\dots,d_R|q} (d_1 \cdots d_R)^{T/d} \, \# \left\{ \begin{array}{c} 1 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(\boldsymbol{a}, q) = 1 \end{array} : d_i \,|\, \boldsymbol{a}\cdot\boldsymbol{c}_i \quad (1 \leqslant i \leqslant R) \right\}.$$

Let $\tilde{d} = \gcd(d_1, \dots, d_R)$. From $d_i \,|\, \boldsymbol{a}\cdot\boldsymbol{c}_i$ $(1 \leqslant i \leqslant R)$, it follows that

$$M(\mathfrak{D}_\ell)^t \boldsymbol{a} \equiv \boldsymbol{0} \pmod{\tilde{d}}.$$

By multiplying both sides of this congruence relation by the adjugate of $M(\mathfrak{D}_\ell)^t$, we obtain

$$\det M(\mathfrak{D}_\ell) a_i \equiv 0 \pmod{\tilde{d}}$$

for all $1 \leqslant i \leqslant R$. Since $\gcd(\boldsymbol{a}, q) = \gcd(\boldsymbol{a}, \tilde{d}) = 1$, it follows that

$$\tilde{d} \,|\, \det M(\mathfrak{D}_\ell).$$

Thus we obtain that (4-4) is

$$\ll \sum_{\substack{d_1|q,\dots,d_R|q \\ \gcd(d_1,\dots,d_R) \ll 1}} (d_1 \cdots d_R)^{T/d} \, \# \left\{ \begin{array}{c} 1 \leqslant \boldsymbol{a} \leqslant q \\ \gcd(\boldsymbol{a}, q) = 1 \end{array} : d_i \,|\, \boldsymbol{a}\cdot\boldsymbol{c}_i \quad (1 \leqslant i \leqslant R) \right\}$$

$$\ll \sum_{\substack{d_1|q,\dots,d_R|q \\ \gcd(d_1,\dots,d_R) \ll 1}} (d_1 \cdots d_R)^{T/d} \frac{q^R}{d_1 \cdots d_R}$$

$$\ll q^R \sum_{\substack{\tilde{d}|q \\ \tilde{d} \ll 1}} \sum_{\substack{v_1|q/\tilde{d},\dots,v_R|q/\tilde{d} \\ \gcd(v_1,\dots,v_R)=1}} (\tilde{d}^R v_1 \cdots v_R)^{T/d-1}.$$

Since $\gcd(v_1, \dots, v_R) = 1$, we may deduce from $v_1|q/\tilde{d}, \dots, v_R|q/\tilde{d}$ that

$$v_1 \cdots v_R \,|\, (q/\tilde{d})^{R-1}.$$

Therefore, the final expression above is

$$\ll q^R \sum_{\substack{\tilde{d}|q \\ \tilde{d} \ll 1}} \sum_{w|(q/\tilde{d})^{R-1}} w^{T/d-1} \, \#\big\{(v_1, \dots, v_R) \in \mathbb{N}^R : w = v_1 \cdots v_R\big\}$$

$$\ll q^R \sum_{\substack{\tilde{d}|q \\ \tilde{d} \ll 1}} \left(\frac{q}{\tilde{d}}\right)^{(R-1)(T/d-1)+\varepsilon} \ll q^{R+(R-1)(T/d-1)+\varepsilon} = q^{(R-1)T/d+1+\varepsilon},$$

for any $\varepsilon > 0$. □

Using this lemma we may extend the truncated singular series.

Lemma 4.2. *Suppose $T > 2d$ and let $\varepsilon > 0$ be sufficiently small. Then*

$$\mathfrak{S} = \mathfrak{S}(B) + O(B^{2-T/d+\varepsilon})$$

for any $B \geqslant 1$. In fact,

$$\mathfrak{S} = \prod_{p \text{ prime}} \chi(p),$$

where

$$\chi(p) = 1 + \sum_{k=1}^{\infty} A(p^k).$$

Proof. The statement is obtained from Lemma 4.1 by writing

$$|\mathfrak{S} - \mathfrak{S}(B)| \leqslant \sum_{q>B} |A(q)| \ll \sum_{q>B} q^{-T/d+1+\varepsilon} \ll B^{2-T/d+\varepsilon}.$$

Since $A(q_1 q_2) = A(q_1)A(q_2)$ for any coprime positive integers q_1 and q_2, we also have

$$\mathfrak{S} = \prod_{p \text{ prime}} \chi(p)$$

as desired. □

Similarly, we define the singular integral as

$$\mathfrak{I} = \int_{\mathbb{R}^R} \prod_{\boldsymbol{c} \in \mathrm{Col}(M)} I(\boldsymbol{\gamma} \cdot \boldsymbol{c}) \cdot e\biggl(-\frac{1}{X^d} \sum_{i=1}^{R} \mu_i \gamma_i\biggr) \,\mathrm{d}\boldsymbol{\gamma} = \lim_{B \to \infty} \mathfrak{I}(B). \tag{4-5}$$

We may also extend the truncated singular integral.

Lemma 4.3. *Suppose $T > d$. Then*

$$\mathfrak{I}(B) = \mathfrak{I} + O(B^{1-T/d})$$

for any $B > 1$.

Proof. We begin with the bound

$$I(\boldsymbol{\gamma} \cdot \boldsymbol{c}) = \int_0^1 e(\boldsymbol{\gamma} \cdot \boldsymbol{c}\xi^d) \,\mathrm{d}\xi \ll \min\{1, |\boldsymbol{\gamma} \cdot \boldsymbol{c}|^{-1/d}\}, \tag{4-6}$$

which for instance can be found in [10, p. 21] or [18, Lemma 2.8]. We know that there exist pairwise disjoint sets $\mathfrak{D}_1, \ldots, \mathfrak{D}_T$ of R linearly independent columns of M. It then follows by Hölder's inequality that

$$|\mathfrak{I}-\mathfrak{I}(B)| \leqslant \int_{|\boldsymbol{\gamma}|\geqslant B} \prod_{\boldsymbol{c}\in\mathrm{Col}(M)} \min\{1, |\boldsymbol{\gamma}\cdot\boldsymbol{c}|^{-1/d}\}\,\mathrm{d}\boldsymbol{\gamma} \tag{4-7}$$

$$\leqslant \int_{|\boldsymbol{\gamma}|\geqslant B} \prod_{\ell=1}^{T} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} \min\{1, |\boldsymbol{\gamma}\cdot\boldsymbol{c}|^{-1/d}\}\,\mathrm{d}\boldsymbol{\gamma}$$

$$\leqslant \prod_{\ell=1}^{T}\biggl(\int_{|\boldsymbol{\gamma}|\geqslant B} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} \min\{1, |\boldsymbol{\gamma}\cdot\boldsymbol{c}|^{-1/d}\}^T\mathrm{d}\boldsymbol{\gamma}\biggr)^{1/T}.$$

By the change of variable $\tilde{\boldsymbol{\gamma}} = M(\mathfrak{D}_\ell)^t\boldsymbol{\gamma}$, we obtain

$$\int_{|\boldsymbol{\gamma}|\geqslant B} \prod_{\boldsymbol{c}\in\mathfrak{D}_\ell} \min\{1, |\boldsymbol{\gamma}\cdot\boldsymbol{c}|^{-1/d}\}^T\mathrm{d}\boldsymbol{\gamma}$$

$$\ll \int_{|\tilde{\boldsymbol{\gamma}}|\gg B} \min\{1, |\tilde{\gamma}_1|^{-1/d}\}^T \cdots \min\{1, |\tilde{\gamma}_R|^{-1/d}\}^T\mathrm{d}\tilde{\boldsymbol{\gamma}}$$

$$\ll \int_{\substack{\tilde{\gamma}_R>\cdots>\tilde{\gamma}_1\geqslant 0\\ \tilde{\gamma}_R\gg B}} \min\{1, |\tilde{\gamma}_1|^{-1/d}\}^T \cdots \min\{1, |\tilde{\gamma}_R|^{-1/d}\}^T\mathrm{d}\tilde{\boldsymbol{\gamma}}$$

$$\ll B^{1-T/d}$$

for each $1 \leqslant \ell \leqslant T$. On substituting this estimate into (4-7), it follows that

$$|\mathfrak{I}-\mathfrak{I}(B)| \ll B^{1-T/d}. \qquad \square$$

We may now conclude the proof of our main results.

Proof of Theorems 1.4 and 1.5. Recall our starting point for the circle method (3-1) and that $\mathfrak{M}_{\mathfrak{L}} \subseteq \mathfrak{M}_{\mathfrak{L}}^{+}$. On combining Propositions 3.3 and 3.8, we have

$$N(\mathfrak{B}; X) = C_{\mathfrak{B}}^{n} X^{n-dR}\mathfrak{S}(\mathfrak{L})\mathfrak{I}(\mathfrak{L}) + O(X^{n-dR}\mathfrak{L}^{-\gamma}), \tag{4-8}$$

for some $\gamma > 0$. Lastly, we obtain from Lemmas 4.2 and 4.3 that

$$\mathfrak{S}(\mathfrak{L})\mathfrak{I}(\mathfrak{L}) = \mathfrak{S}\mathfrak{I} + O(\mathfrak{L}^{2-T/d+\varepsilon} + \mathfrak{L}^{1-T/d}),$$

for any $\varepsilon > 0$ sufficiently small. These two equations together give the desired asymptotic formula. $\square$

Remark 4.4. To see the allowable value of γ given in Remark 1.7, observe that the dominant error term in (4-8) comes from Proposition 3.3. Here, we have replaced $\mathfrak{L}$ by $\mathfrak{L}^{1/(R+1)}$ after an application of Lemma 3.2 and then saved a power $\lambda(d)$ or $1/(2d)$ of this via Lemmas 2.1 and 2.5, respectively. Finally, one recalls from the proof of Proposition 3.8 that $\delta, A \leqslant 1/(2R+4)$.

Acknowledgements

Nick Rome was supported by FWF project ESP 441-NBL and Shuntaro Yamagishi by an FWF grant (DOI 10.55776/P32428). The authors are grateful to Jörg Brüdern for helpful discussions regarding his paper [6] and to Trevor Wooley for numerous helpful comments on an earlier version of this paper and on the state of the art in Waring's problem. We also thank Daniel Flores for pointing out an error in the proof of Lemma 4.1 in an earlier version of this paper.

References

[1] B. J. Birch, "Forms in many variables", *Proc. Roy. Soc. London Ser. A* **265** (1961/62), 245–263.

[2] J. Bourgain, C. Demeter, and L. Guth, "Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three", *Ann. of Math.* (2) **184**:2 (2016), 633–682.

[3] J. Brandes and S. T. Parsell, "Simultaneous additive equations: repeated and differing degrees", *Canad. J. Math.* **69**:2 (2017), 258–283.

[4] J. Brandes and T. D. Wooley, "Optimal mean value estimates beyond Vinogradov's mean value theorem", *Acta Arith.* **200**:2 (2021), 149–182.

[5] T. D. Browning and S. Yamagishi, "Arithmetic of higher-dimensional orbifolds and a mixed Waring problem", *Math. Z.* **299**:1-2 (2021), 1071–1101.

[6] J. Brüdern and R. J. Cook, "On simultaneous diagonal equations and inequalities", *Acta Arith.* **62**:2 (1992), 125–149.

[7] J. Brüdern and T. D. Wooley, "The Hasse principle for pairs of diagonal cubic forms", *Ann. of Math.* (2) **166**:3 (2007), 865–895.

[8] J. Brüdern and T. D. Wooley, "The Hasse principle for systems of diagonal cubic forms", *Math. Ann.* **364**:3-4 (2016), 1255–1274.

[9] J. Brüdern and T. D. Wooley, "On Waring's problem for larger powers", *J. Reine Angew. Math.* **805** (2023), 115–142.

[10] H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, 2nd ed., Cambridge University Press, 2005.

[11] H. Davenport and D. J. Lewis, "Simultaneous equations of additive type", *Philos. Trans. Roy. Soc. London Ser. A* **264** (1969), 557–595.

[12] D. Flores, "A circle method approach to K-multimagic squares", (2024). arXiv 2406.08161

[13] L. Low, J. Pitman, and A. Wolff, "Simultaneous diagonal congruences", *J. Number Theory* **29**:1 (1988), 31–59.

[14] S. T. Parsell, "Pairs of additive equations of small degree", *Acta Arith.* **104**:4 (2002), 345–402.

[15] S. T. Parsell and T. D. Wooley, "On pairs of diagonal quintic forms", *Compositio Math.* **131**:1 (2002), 61–96.

[16] N. Rome and S. Yamagishi, "On the existence of magic squares of powers", preprint, 2024. arXiv 2406.09364

[17] R. C. Vaughan, "A new iterative method in Waring's problem", *Acta Math.* **162** (1989), 1–71.

[18] R. C. Vaughan, *The Hardy–Littlewood method*, 2nd ed., Cambridge Tracts in Mathematics **125**, Cambridge University Press, 1997.

[19] R. C. Vaughan and T. Wooley, "On Waring's problem: some refinements", *Proc. London Math. Soc.* (3) **63** (1991), 35–68.

[20] R. C. Vaughan and T. D. Wooley, "Further improvements in Waring's problem", *Acta Math.* **174**:2 (1995), 147–240. MR

[21] R. C. Vaughan and T. D. Wooley, "Further improvements in Waring's problem, IV: Higher powers", *Acta Arith.* **94**:3 (2000), 203–285.

[22] T. D. Wooley, "Large improvements in Waring's problem", *Ann. of Math.* (2) **135**:1 (1992), 131–164.

[23] T. D. Wooley, "New estimates for smooth Weyl sums", *J. London Math. Soc.* (2) **51**:1 (1995), 1–13.

[24] T. D. Wooley, "The asymptotic formula in Waring's problem", *Int. Math. Res. Not.* **2012**:7 (2012), 1485–1504.

[25] T. D. Wooley, "The cubic case of the main conjecture in Vinogradov's mean value theorem", *Adv. Math* **294** (2016), 532–561.

[26] T. D. Wooley, "On Waring's problem for intermediate powers", *Acta Arith.* **176**:3 (2016), 241–247.

[27] T. D. Wooley, "Nested efficient congruencing and relatives of Vinogradov's mean value theorem", *Proc. Lond. Math. Soc.* (3) **118**:4 (2019), 942–1016.

Received September 18, 2024. Revised September 16, 2025.

NICK ROME
INSTITUTE OF ANALYSIS AND NUMBER THEORY
TU GRAZ
GRAZ
AUSTRIA
rome@tugraz.at

SHUNTARO YAMAGISHI
IST AUSTRIA
KLOSTERNEUBURG
AUSTRIA
shuntaro.yamagishi@ist.ac.at