OPTIMAL MEAN VALUE ESTIMATES BEYOND VINOGRADOV'S MEAN VALUE THEOREM

JULIA BRANDES AND TREVOR D. WOOLEY

Abstract. We establish improved mean value estimates associated with the number of integer solutions of certain systems of diagonal equations, in some instances attaining the sharpest conjectured conclusions. This is the first occasion on which bounds of this quality have been attained for Diophantine systems not of Vinogradov type. As a consequence of this progress, whenever $u \geqslant 3v$ we obtain the Hasse principle for systems consisting of v cubic and u quadratic diagonal equations in $6v + 4u + 1$ variables, thus attaining the convexity barrier for this problem.

1. Introduction

In recent years, our understanding of systems of diagonal equations and their associated mean values has advanced rapidly. Whilst only a few years ago, such mean values had been comprehensively understood only in the most basic cases, the resolution of the main conjecture associated with Vinogradov's mean value theorem by the second author [13, 14] and Bourgain, Demeter and Guth [1] has transformed the landscape. It now seems feasible to address the challenge of establishing similarly strong results for a much wider class of cognate problems.

In this memoir, we make progress towards, and in certain cases attain, the convexity barrier for a family of mean values associated with systems of equations that fail to be translation-dilation invariant and thus lie outside the scope of the efficient congruencing and ℓ^2-decoupling methods developed by the second author [13, 14] and Bourgain, Demeter and Guth [1]. The most accessible of our results addresses systems of cubic and quadratic diagonal equations. Let $\mathcal{N}_{s,v,u}(X)$ denote the number of integral solutions $\mathbf{x} \in [-X, X]^s$ of the system of equations

$$\begin{aligned} c_{i,1}^{(3)} x_1^3 + \ldots + c_{i,s}^{(3)} x_s^3 &= 0 \qquad (1 \leqslant i \leqslant v) \\ c_{j,1}^{(2)} x_1^2 + \ldots + c_{j,s}^{(2)} x_s^2 &= 0 \qquad (1 \leqslant j \leqslant u), \end{aligned} \tag{1.1}$$

consisting of u quadratic and v cubic equations of diagonal shape. Here and throughout we assume the coefficients $c_{i,j}^{(k)}$ of such systems to be integral. It is clear that the presence of coefficients in such systems necessitates some kind of non-singularity condition, lest the equations interact in some non-generic way. We refer to an $r \times s$ matrix C as *highly non-singular* if $s \geqslant r$ and any collection of r distinct columns of C forms a non-singular matrix.

2010 *Mathematics Subject Classification.* 11L15, 11D45, 11L07, 11P55.
Key words and phrases. Exponential sums, Hardy–Littlewood method.

Our first result shows that $\mathcal{N}_{s,v,u}(X)$ satisfies the anticipated asymptotic formula for all sets of coefficients in general position, provided that $s \geqslant 6v + 4u + 1$ and $u \geqslant 3v$. This achieves the conjectured convexity barrier.

Theorem 1.1. *Suppose that $u \geqslant 3v$ and that $s \geqslant 6v + 4u + 1$. Assume further that the coefficient matrices*

$$C^{(2)} = (c^{(2)}_{i,j})_{\substack{1\leqslant i\leqslant u\\ 1\leqslant j\leqslant s}} \quad \text{and} \quad C^{(3)} = (c^{(3)}_{i,j})_{\substack{1\leqslant i\leqslant v\\ 1\leqslant j\leqslant s}}$$

are highly non-singular. Then there exist constants $\mathcal{C} \geqslant 0$ and $\delta > 0$ such that

$$\mathcal{N}_{s,v,u}(X) = \mathcal{C}X^{s-3v-2u} + O(X^{s-3v-2u-\delta}). \tag{1.2}$$

Moreover, if the system (1.1) *has non-singular real and p-adic solutions for all primes p, then $\mathcal{C} > 0$.*

In general, asymptotic formulæ like the one supplied by (1.2) are expected to hold whenever the number of variables exceeds twice the total degree of the system. However, thus far the validity of such an asymptotic formula has been proved only in a few isolated instances. Arguably the first non-trivial case in which this convexity barrier was achieved occurs in work of Cook [7, 8] concerning pairs and triples of diagonal quadratic equations. Recent work of Brüdern and the second author [5, 6] obtains asymptotic lower bounds at the convexity limit for systems of diagonal cubic forms. In the case of mixed systems of cubic and quadratic equations, work of the second author underlying [12, Theorem 1.2] achieves the convexity limit in the case $u = v = 1$ with $s \geqslant 11$ relating to systems consisting of one cubic and one quadratic diagonal equation. Most recently, investigations of the first author joint with Parsell [3, Theorem 1.4] establish an asymptotic formula tantamount to (1.2) for systems of v cubic and u quadratic diagonal equations, though under the more restrictive hypothesis that $s \geqslant \lfloor 20v/3\rfloor + 4u + 1$, thus missing the convexity barrier whenever $v \geqslant 2$. In subsequent work [2], the first author proved that an asymptotic formula of the shape (1.2) holds when $v \geqslant 2u$ and $s \geqslant 6v + \lfloor 14u/3\rfloor + 1$, which misses the convexity barrier when $u \geqslant 2$. Thus, Theorem 1.1 provides the first instance where bounds of the expected quality have been achieved for systems of v cubic and u quadratic equations in settings where both u and v exceed 1.

Theorem 1.1 is in fact a special case of our more general Theorem 1.5 below. Both of these results rest on our new estimates for certain mean values of Vinogradov type. In their most general form, such mean values encode the number of integral solutions of systems of the general shape

$$c^{(l)}_{i,1}(x_1^l - y_1^l) + \ldots + c^{(l)}_{i,s}(x_s^l - y_s^l) = 0 \qquad (1 \leqslant i \leqslant r_l,\ 1 \leqslant l \leqslant k), \tag{1.3}$$

in which $r_1, \ldots, r_k$ are non-negative integers and the coefficients $c^{(l)}_{i,j}$ are integers. When all of the coefficient matrices

$$C^{(l)} = \big(c^{(l)}_{i,j}\big)_{\substack{1\leqslant i\leqslant r_l\\ 1\leqslant j\leqslant s}}$$

are highly non-singular, then the main conjecture states that the number of integral solutions $\mathbf{x}, \mathbf{y} \in [-X, X]^s$ of the system (1.3) should be at most of order $X^{s+\varepsilon} + X^{2s-K}$, for any $\varepsilon > 0$, where $K = r_1 + 2r_2 + \ldots + kr_k$ denotes the system's total degree. A corresponding lower bound, with $\varepsilon = 0$, is provided by an argument akin to that delivering [11, equation (7.4)]. Systems of the shape (1.3) have previously been studied by the first author together with Parsell [3], where it was shown that the main conjecture for such systems holds when $r_l \geqslant r_{l+1}$ for all $1 \leqslant l \leqslant k-1$. In the latter circumstances, the system (1.3) can be viewed as a superposition of Vinogradov systems of various degrees (see Theorem 2.1 and Corollary 2.2 in that paper). In wider generality, bounds of the strength of those described in Theorems 1.1 and 1.5 were known hitherto only for systems of quadratic equations and systems of Vinogradov type, as well as superpositions of these two special classes of systems.

The goal of the work at hand is to enlarge the range of systems of type (1.3) for which the main conjecture is known to hold. When the coefficient matrices $C^{(l)}$ are highly non-singular for $2 \leqslant l \leqslant k$, we denote by $I^{v,u}_{s,k}(X) = I^{v,u}_{s,k}(X; C^{(2)}, \ldots, C^{(k)})$ the number of integral solutions $\mathbf{x}, \mathbf{y} \in [-X, X]^s$ of the system (1.3), where

$$r_k = v, \qquad r_{k-1} = \ldots = r_2 = u, \qquad r_1 = 0.$$

Write further

$$K = \tfrac{1}{2}k(k-1)u + kv - u, \tag{1.4}$$

so that K denotes the total degree of the system.

In order to describe our new results concerning the mean value $I^{v,u}_{s,k}(X)$, we need to consider certain auxiliary systems of equations. Let $l \geqslant 2$ be an integer and write $\sigma = \frac{1}{2}l(l+1)$. Then, given a positive number X, we denote by $M^*_l(X)$ the number of integer tuples $\mathbf{x}, \mathbf{y} \in [-X, X]^{\sigma-2}$ and $\mathbf{z}, \mathbf{h} \in [-X, X]^2$ satisfying

$$\sum_{i=1}^{\sigma-2} (x_i^j - y_i^j) + j(z_1^{j-1}h_1 + z_2^{j-1}h_2) = 0 \qquad (1 \leqslant j \leqslant l). \tag{1.5}$$

The main conjecture for systems of the shape (1.5) claims that

$$M^*_l(X) \ll X^{\frac{1}{2}l(l+1)+\varepsilon}. \tag{1.6}$$

Our first main result is as follows.

Theorem 1.2. *Suppose that $k \geqslant 3$, $v \geqslant 1$ and $u \geqslant 2v$ are integers with $u|kv$, and assume* (1.6) *for $l = k-1$. Then for any $s \geqslant u$ and any $\varepsilon > 0$ we have*

$$I^{v,u}_{s,k}(X) \ll X^{\varepsilon}(X^s + X^{2s-K}).$$

By combining the ideas of the proof of Theorem 1.2 with those underlying [3, Theorem 2.1], we can extend our results to cover also superpositions of systems of equations of the kind considered in Theorem 1.2. Fix a collection of degrees $k_1 > k_2 > \ldots > k_n \geqslant 3$ with associated multiplicities $v_1, \ldots, v_n \in \mathbb{N}$. Moreover, fix a tuple $u_0, u_1, \ldots, u_n$ of non-negative integers with $u_i \geqslant v_i$ for $1 \leqslant i \leqslant n$, set $k = k_1$, and define $w_0 = 0$ and $w_i = u_1 + \ldots + u_i$ for $1 \leqslant i \leqslant n$. Now define the parameter r_l by putting

$$r_1 = 0, \qquad r_2 = w_n + u_0, \qquad r_l = \begin{cases} w_n & \text{when } 3 \leqslant l < k_n, \\ w_j & \text{when } k_{j+1} < l < k_j, \\ w_{j-1} + v_j & \text{when } l = k_j \text{ for some } j. \end{cases} \tag{1.7}$$

We denote by

$$I_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) = I_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X; C^{(2)}, \ldots, C^{(k)}) \tag{1.8}$$

the number of integer solutions $\mathbf{x}, \mathbf{y} \in [-X, X]^s$ of the system (1.3) with $\mathbf{r}$ defined as in (1.7). These systems can be viewed as superpositions of systems of the shape considered in Theorem 1.2 with parameters (k_j, v_j, u_j), together with u_0 additional quadratic equations. Here, the total degree is given by

$$K = \sum_{l=2}^{k} l r_l = \sum_{j=1}^{n} K_j + 2u_0, \tag{1.9}$$

where, in accordance with (1.4), we write

$$K_j = \tfrac{1}{2} k_j(k_j - 1)u_j + k_j v_j - u_j.$$

In this notation, we have the following generalisation of Theorem 1.2.

Theorem 1.3. *Let u_0 be a non-negative integer. Suppose that $k_1 > \ldots > k_n \geqslant 3$, and let $u_1, \ldots, u_n$ as well as $v_1, \ldots, v_n$ be natural numbers with $u_j \geqslant 2v_j$ and $u_j | k_j v_j$ for $1 \leqslant j \leqslant n$. Also, assume (1.6) for all degrees $l = k_j - 1$ with $1 \leqslant j \leqslant n$. Then for $s \geqslant r_2$ and any $\varepsilon > 0$ we have*

$$I_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) \ll X^{\varepsilon}(X^s + X^{2s-K}).$$

We also have an alternative, unconditional formulation of this result, which is given in Theorem 3.3 below.

To illustrate the strength of our results in Theorems 1.2 and 1.3, we discuss in more detail some of the most relevant special cases. Among the systems of diagonal equations not of Vinogradov type, the most well-studied ones are systems of cubic equations and systems of cubic and quadratic equations, such as we considered in our motivating example in Theorem 1.1. Regarding such systems, it is immediate from work of the second author [12, Theorem 1.1] that for every $\varepsilon > 0$ one has $I_{5,3}^{1,1}(X) \ll X^{31/6+\varepsilon}$, and this bound implies via [3, Theorem 2.1] that $I_{3+2u,3}^{1,u}(X) \ll X^{3+2u+1/6+\varepsilon}$ for all $u \geqslant 1$. Theorem 1.3 now allows us to improve this result.

Corollary 1.4. *Suppose that $v \geqslant 1$ and $s \geqslant u \geqslant 3v$. Then for any $\varepsilon > 0$ we have*

$$I_{s,3}^{v,u}(X) \ll X^{\varepsilon}(X^s + X^{2s-3v-2u}).$$

This follows from Theorem 1.3 in combination with Lemma 2.1 below. Corollary 1.4 represents only the second occasion, after the second author's successful treatment of the cubic case of Vinogradov's mean value theorem [13], that the convexity barrier has been attained for a system of diagonal equations involving cubic equations. In particular, we now have the main conjecture for mean values that correspond to systems consisting of one cubic and three quadratic diagonal equations. This is the main new input that enables us to prove Theorem 1.1.

Our results complement older ones that can be obtained by other means. On the one hand, it follows from Theorem 2.1 and Corollary 2.2 of the first author's work with

Parsell [3] in combination with Vinogradov's mean value theorem [1, Theorem 1.1] that the conclusion of Theorem 1.3 holds unconditionally in the range

$$s \geqslant \sum_{j=1}^{n} \left(v_j \frac{k_j(k_j+1)}{2} + (u_j - v_j)\frac{k_j(k_j-1)}{2} \right) + 2u_0 = K + w_n.$$

On the other hand, for small s the second author's result [14, Corollary 1.2] can be combined with the arguments of [3, Theorem 2.1] to establish the conclusion of Theorem 1.3 unconditionally in the range

$$\begin{aligned} s &\leqslant \sum_{j=1}^{n} \left(v_j \frac{k_j(k_j-1)}{2} + (u_j - v_j)\frac{(k_j-1)(k_j-2)}{2} \right) + 2u_0 \\ &= K - \sum_{j=1}^{n} ((k_j-2)u_j + v_j). \end{aligned}$$

Mean value estimates like those of Theorems 1.2 and 1.3 have long been employed to establish asymptotic formulæ for the number of solutions of simultaneous diagonal equations. For $\mathbf{r}$ as in (1.7) and highly non-singular coefficient matrices $C^{(l)}$ $(2 \leqslant l \leqslant k)$, denote by $N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X)$ the number of integral solutions of the system of equations

$$c_{i,1}^{(l)} x_1^l + \ldots + c_{i,s}^{(l)} x_s^l = 0 \qquad (1 \leqslant i \leqslant r_l,\ 2 \leqslant l \leqslant k) \tag{1.10}$$

with $|x_j| \leqslant X$ for $1 \leqslant j \leqslant s$. It is well known that, if s is sufficiently large in terms of $\mathbf{k}$, $\mathbf{v}$ and $\mathbf{u}$, there is an asymptotic formula of the shape

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) = (\mathcal{C} + o(1))X^{s-K}, \tag{1.11}$$

where $\mathcal{C}$ is a non-negative constant encoding the local solubility data for the system (1.10). The relevant question is how large s has to be for an asymptotic formula like that of (1.11) to hold. Theorem 1.1 of [3] provides a bound for s that is somewhat unwieldy, but can likely be reduced to

$$s \geqslant \sum_{j=1}^{n} \big(v_j k_j(k_j+1) + (u_j - v_j)k_j(k_j-1) \big) + 4u_0 + 1 = 2K + 2w_n + 1$$

by accounting for our revised treatment of the major arcs described in §§5–6 below. On the other hand, unless fundamentally new methods become available that avoid the use of mean values, we cannot expect to be able to establish such asymptotic formulæ when $s \leqslant 2K$. Thanks to our new mean value estimates in Theorem 1.3, we are now able to make progress towards this theoretical barrier.

Theorem 1.5. *Let u_0 be a non-negative integer. Suppose that $k_1 > \ldots > k_n \geqslant 3$, and let $u_1, \ldots, u_n$ as well as $v_1, \ldots, v_n$ be natural numbers with $u_j \geqslant 2v_j$ and $u_j | k_j v_j$ for $1 \leqslant j \leqslant n$. Also, assume (1.6) for all degrees $l = k_j - 1$ with $1 \leqslant j \leqslant n$. Then for $s \geqslant 2K + 1$ the asymptotic formula (1.11) holds with $\mathcal{C} \geqslant 0$. If, furthermore, the system (1.10) has non-singular solutions in $\mathbb{R}$ as well as in the fields $\mathbb{Q}_p$ for all p, then the constant $\mathcal{C}$ is positive.*

Again, we refer to Theorem 4.1 below for an unconditional version of this result. Moreover, we note that in Lemma 2.1 below it is shown that the bound (1.6) holds for $l = 2$, and thus Theorem 1.1 can be deduced as a special case of Theorem 1.5, corresponding to the parameters $k = 3$ and $u_0 = u - 3v$.

The proofs of our results rest on an idea that played a crucial role in the second author's work on pairs of quadratic and cubic diagonal equations [12], and which has been explored further in the authors' recent work on incomplete Vinogradov systems [4]. In these papers, the missing linear equation is artificially added in, which makes it possible to exploit the strong bounds on Vinogradov's mean value theorem. By taking advantage of the translation-dilation invariance of the newly completed Vinogradov systems, we then relate these systems to the auxiliary mean values $M_l^*(X)$ introduced above. Whilst our understanding of these auxiliary mean values remains unsatisfactory for general degree, the quantity $M_2^*(X)$ may be comprehensively understood in terms of quadratic Vinogradov systems. This observation plays a pivotal role in our argument, and it is the main reason why we attain the convexity barrier in Theorem 1.1 and Corollary 1.4.

Notation. Throughout, the letters s, u, v, and k, as well as the entries of the vectors $\mathbf{k}$, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{r}$, will denote non-negative integers. The letter ε will be used to denote an arbitrary, but sufficiently small positive number, and we adopt the convention that whenever it appears in a statement, we assert that the statement holds for all sufficiently small $\varepsilon > 0$. We take X to be a large positive number which, just like the implicit constants in the notations of Landau and Vinogradov, is permitted to depend at most on s, $\mathbf{k}$, $\mathbf{v}$, $\mathbf{u}$, the coefficient matrices $C^{(l)}$ $(2 \leqslant l \leqslant k)$, and ε. We employ the non-standard notation that when $G : [0,1)^n \to \mathbb{C}$ is integrable for some $n \in \mathbb{N}$, then

$$\oint G(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \int_{[0,1)^n} G(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}.$$

Here and elsewhere, we use vector notation liberally in a manner that is easily discerned from the context. In particular, when $\mathbf{b}$ denotes the integer tuple $(b_1, \ldots, b_n)$, we write $(q, \mathbf{b}) = \gcd(q, b_1, \ldots, b_n)$.

Acknowledgements. Both authors thank the Fields Institute in Toronto for excellent working conditions and support that made this work possible during the Thematic Program on Unlikely Intersections, Heights, and Efficient Congruencing. This work was further facilitated by subsequent visits of the first author to the University of Bristol, and of the second author to the University of Waterloo. The authors gratefully acknowledge the hospitality of both institutions. The work of both authors was supported by the National Science Foundation under Grant No. DMS-1440140 while they were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester. The first author's work was supported in part by Starting Grant 2017-05110 from Vetenskapsrådet. The second author's work was supported by a European Research Council Advanced Grant under the European Union's Horizon 2020 research and innovation programme via grant agreement No. 695223, and in the final stages by the National Science Foundation via Grant No. DMS-1854398 and DMS-2001549.

The authors are also very grateful to Scott Parsell for pointing out an oversight in an earlier version of this paper, which necessitated a fundamental re-write of parts of the argument.

2. Preliminaries and preparatory steps

Our goal in this and the next section is the proof of Theorem 1.3. Before delving to the core of the argument, we pause to introduce some notation and establish a mean value estimate that will be of use in our subsequent discussion. For $2 \leqslant l \leqslant k$ we define the exponential sum $K_l(\boldsymbol{\alpha}; Z, H)$ by putting

$$K_l(\boldsymbol{\alpha}; Z, H) = \sum_{|h| \leqslant H} \sum_{|z| \leqslant Z} e(h\alpha^{(1)} + 2hz\alpha^{(2)} + \ldots + lhz^{l-1}\alpha^{(l)}), \tag{2.1}$$

and we write

$$f_l(\boldsymbol{\alpha}; X) = \sum_{|x| \leqslant X} e(\alpha^{(1)}x + \alpha^{(2)}x^2 + \ldots + \alpha^{(l)}x^l).$$

Then, with the standard notation associated with Vinogradov's mean value theorem in mind, we put

$$J_{s,l}(X) = \oint |f_l(\boldsymbol{\alpha}; X)|^{2s} \, \mathrm{d}\boldsymbol{\alpha}.$$

We note that the main conjecture associated with Vinogradov's mean value theorem is now known to hold for all degrees. This is classical when $l = 2$, it is a consequence of work of the second author [13] for degree $l = 3$, and for degrees exceeding three it follows from the work of Bourgain, Demeter and Guth, and of the second author (see [1, Theorem 1.1] and [14, Corollary 1.3]). Thus, for all $\sigma > 0$ one has

$$J_{\sigma,l}(X) \ll X^{\varepsilon}(X^{\sigma} + X^{2\sigma - l(l+1)/2}). \tag{2.2}$$

For future reference, we record the trivial inequality

$$|a_1 \cdots a_n| \leqslant |a_1|^n + \ldots + |a_n|^n, \tag{2.3}$$

which is valid for all $a_1, \ldots, a_n \in \mathbb{C}$.

We begin by bounding the mean value $M_2^*(X)$.

Lemma 2.1. *Let X, Z and H be large real numbers. Then one has*

$$\oint |f_2(\boldsymbol{\alpha}; X) K_2(\boldsymbol{\alpha}; Z, H)|^2 \, \mathrm{d}\boldsymbol{\alpha} \ll (XHZ)^{\varepsilon}(HZ^2 + XZ^2 + XZH). \tag{2.4}$$

Proof. Upon considering the underlying system of equations, we see that the mean value on the left hand side of (2.4) is given by the number of integer solutions of the system of equations

$$\begin{aligned} x_1^2 - x_2^2 &= 2(h_1 z_1 - h_2 z_2) \\ x_1 - x_2 &= h_1 - h_2, \end{aligned} \tag{2.5}$$

with $|x_i| \leqslant X$, $|h_i| \leqslant H$ and $|z_i| \leqslant Z$ for $i = 1, 2$. The second of these equations permits the substitution $h_2 = h_1 - x_1 + x_2$ into the first, whence

$$(x_1 - x_2)(x_1 + x_2 - 2z_2) = 2h_1(z_1 - z_2).$$

Suppose first that $h_1(z_1 - z_2)$ is non-zero. Then for each of the $O(HZ^2)$ possible choices for h_1, z_1 and z_2 fixing the latter integer in such a manner, an elementary divisor function estimate shows there to be $O((HZ)^\varepsilon)$ possible choices for the integers $x_1 - x_2$ and $x_1 + x_2 - 2z_2$, and hence also for x_1 and x_2. These choices also fix $h_2 = h_1 - x_1 + x_2$, so we see that there are $O(H^{1+\varepsilon}Z^{2+\varepsilon})$ solutions of this first type. Meanwhile, if $h_1(z_1 - z_2) = 0$, then $h_1 = 0$ or $z_1 = z_2$, and at the same time either $x_1 = x_2$ or $x_1 = 2z_2 - x_2$. In any case, therefore, each of the $O(XZ)$ possible choices for z_2 and x_2 determine x_1 and either h_1 or z_1. Since there are $O(Z + H)$ possible choices left by this constraint for the latter, and h_2 is again fixed by these choices just as before, we find that there are $O(XZ(Z + H))$ solutions of this second type. The conclusion of the lemma follows by summing the contributions from both types of solutions. □

Upon taking $X = H = Z$ in Lemma 2.1, we conclude that $M_2^*(X) \ll X^{3+\varepsilon}$, which establishes (1.6) for $l = 2$. We remark also that the system (2.5) can be interpreted as being of Vinogradov shape of degree two by means of the substitution $h_i = u_i - v_i$ and $z_i = u_i + v_i$ for $i = 1, 2$. Viewed in this way, Lemma 2.1 amounts to no more than a rephrasing of the classical elementary proof of the quadratic case in Vinogradov's mean value theorem.

We now initiate the proof of Theorem 1.3, assuming the hypotheses of its statement. For $l \geqslant 2$, let

$$g_l(\boldsymbol{\alpha}; X) = \sum_{|x| \leqslant X} e(\alpha^{(2)}x^2 + \alpha^{(3)}x^3 + \ldots + \alpha^{(l)}x^l).$$

Define $\boldsymbol{\alpha}^{(l)} = (\alpha_i^{(l)})_{1 \leqslant i \leqslant r_l}$ for $2 \leqslant l \leqslant k$. When $1 \leqslant i \leqslant r_2$, write $\boldsymbol{\alpha}_i = (\alpha_i^{(l)})$, where l runs over all values for which $r_l \geqslant i$. We then put

$$\gamma_j^{(l)} = \sum_{i=1}^{r_l} c_{i,j}^{(l)} \alpha_i^{(l)} \qquad (1 \leqslant j \leqslant s). \tag{2.6}$$

Also, set $\boldsymbol{\gamma}_j = (\gamma_j^{(l)})_{2 \leqslant l \leqslant k}$ for $1 \leqslant j \leqslant s$ and $\boldsymbol{\gamma}^{(l)} = (\gamma_j^{(l)})_{1 \leqslant j \leqslant s}$ for $2 \leqslant l \leqslant k$, and put $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_s) = (\boldsymbol{\gamma}^{(2)}, \ldots, \boldsymbol{\gamma}^{(k)})^T$. Then by orthogonality we have

$$I_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) = \oint \prod_{j=1}^{s} |g_k(\boldsymbol{\gamma}_j; X)|^2 \, \mathrm{d}\boldsymbol{\alpha}.$$

Set $t_0 = 2$, and for a set of positive integers $t_1, \ldots, t_n$ to be fixed later take

$$s_0 = t_0 u_0 + t_1 u_1 + \ldots + t_n u_n. \tag{2.7}$$

Thus, on recalling (1.7), we see in particular that

$$s_0 \geqslant u_0 + u_1 + \ldots + u_n = r_2.$$

Further, let $\mathcal{I}$ denote the set of all integral r_2-tuples

$$(j_{m,1},\ldots,j_{m,u_m})_{0\leqslant m\leqslant n}$$

with pairwise distinct entries $j_{m,h}\in\{1,\ldots,s\}$, and put

$$\mathcal{G}^{\mathbf{v},\mathbf{u}}_{\mathbf{t},\mathbf{k}}(X)=\max_{\mathbf{j}\in\mathcal{I}}\oint\prod_{m=0}^{n}\prod_{h=1}^{u_m}|g_k(\boldsymbol{\gamma}_{j_{m,h}};X)|^{2s_0/r_2}\,\mathrm{d}\boldsymbol{\alpha}. \tag{2.8}$$

We can bound $I^{\mathbf{v},\mathbf{u}}_{s,\mathbf{k}}(X)$ in terms of $\mathcal{G}^{\mathbf{v},\mathbf{u}}_{\mathbf{t},\mathbf{k}}(X)$. In particular, this will allow us to concentrate on the case when $s=s_0$.

Lemma 2.2. *For any fixed choice of the positive integers $t_1,\ldots,t_n$, we have the bounds*

$$I^{\mathbf{v},\mathbf{u}}_{s,\mathbf{k}}(X)\ll\begin{cases}(\mathcal{G}^{\mathbf{v},\mathbf{u}}_{\mathbf{t},\mathbf{k}}(X))^{s/s_0} & \text{when } r_2\leqslant s\leqslant s_0,\\ X^{2s-2s_0}\mathcal{G}^{\mathbf{v},\mathbf{u}}_{\mathbf{t},\mathbf{k}}(X) & \text{when } s>s_0.\end{cases}$$

Proof. When $s>s_0$, the trivial bound $g_k(\boldsymbol{\gamma}_j;X)=O(X)$ delivers the estimate

$$I^{\mathbf{v},\mathbf{u}}_{s,\mathbf{k}}(X)\ll X^{2s-2s_0}\oint\prod_{j=1}^{s_0}|g_k(\boldsymbol{\gamma}_j;X)|^2\,\mathrm{d}\boldsymbol{\alpha},$$

and the conclusion of the lemma follows in this case from (2.3). Suppose now that $r_2\leqslant s\leqslant s_0$. Then from (2.3) and an application of Hölder's inequality, we find that

$$\begin{aligned}I^{\mathbf{v},\mathbf{u}}_{s,\mathbf{k}}(X)&\ll\max_{\mathbf{j}\in\mathcal{I}}\oint\prod_{m=0}^{n}\prod_{h=1}^{u_m}|g_k(\boldsymbol{\gamma}_{j_{m,h}};X)|^{2s/r_2}\,\mathrm{d}\boldsymbol{\alpha}\\&\ll\left(\max_{\mathbf{j}\in\mathcal{I}}\oint\prod_{m=0}^{n}\prod_{h=1}^{u_m}|g_k(\boldsymbol{\gamma}_{j_{m,h}};X)|^{2s_0/r_2}\,\mathrm{d}\boldsymbol{\alpha}\right)^{s/s_0}.\end{aligned}$$

Thus the lemma is established in both cases. □

Suppose that the maximum in (2.8) is assumed at the tuple $\mathbf{j}\in\mathcal{I}$, which we consider fixed for the remainder of the analysis. For $2\leqslant l\leqslant k$ and $1\leqslant i\leqslant r_l$, set $d^{(l)}_{i,w_{m-1}+h}=c^{(l)}_{i,j_{m,h}}$ when $1\leqslant h\leqslant u_m$ and $1\leqslant m\leqslant n$, and likewise $d^{(l)}_{i,w_n+h}=c^{(l)}_{i,j_{0,h}}$ when $1\leqslant h\leqslant u_0$. We then have the coefficient matrices

$$D^{(l)}=\big(d^{(l)}_{i,j}\big)_{\substack{1\leqslant i\leqslant r_l\\1\leqslant j\leqslant r_2}}\quad(2\leqslant l\leqslant k).$$

We define $\delta^{(l)}_i$ via the relations $\boldsymbol{\delta}^{(l)}=(D^{(l)})^T\boldsymbol{\alpha}^{(l)}$ for $2\leqslant l\leqslant k$, and put $\boldsymbol{\delta}_j=(\delta^{(l)}_j)_{2\leqslant l\leqslant k}$ for $1\leqslant j\leqslant r_2$. Here, we employ notational conventions analogous to those described in the sequel to (2.6).

Write

$$G^{v_j,u_j}_{t_j,k_j}(X)=\oint\prod_{h=1}^{u_j}|g_{k_j}(\boldsymbol{\delta}_{w_{j-1}+h};X)|^{2t_j}\,\mathrm{d}\boldsymbol{\alpha}\qquad(1\leqslant j\leqslant n).$$

Thus, in the case $n=1$ and $u_0=0$, we have $\mathcal{G}^{v,u}_{\mathbf{t},k}(X)=G^{v,u}_{t,k}(X)$.

Lemma 2.3. *One has*

$$\mathcal{G}_{\mathbf{t},\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) \ll X^{2u_0+\varepsilon} \prod_{j=1}^{n} G_{t_j,k_j}^{v_j,u_j}(X).$$

Proof. Recall (2.7) and (2.8). For temporary notational convenience, we put $u_{n+1} = u_0$. Then, after possibly relabelling indices, we see from (2.3) that

$$\mathcal{G}_{\mathbf{t},\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) \ll \oint \prod_{j=1}^{n+1} \prod_{h=1}^{u_j} |g_k(\boldsymbol{\delta}_{w_{j-1}+h}; X)|^{2t_j} \, \mathrm{d}\boldsymbol{\alpha}. \tag{2.9}$$

The desired conclusion now follows by essentially the same argument as in [3, Theorem 2.1]. Recall that the coefficient matrices $C^{(l)}$ are highly non-singular. Consequently, the matrices $D^{(l)}$ underlying the mean value in (2.8) inherit that property. Upon considering the underlying Diophantine equations and applying elementary row operations, we may thus assume without loss of generality that the first $r_l \times r_l$ submatrix of each matrix $D^{(l)}$ is diagonal.

Recall the definition of the parameter r_l from (1.7). Since $\mathbf{j} \in \mathcal{I}$ has r_2 entries, we see that the matrix $D^{(2)}$ is of square format and hence diagonal. Thus we have $\delta_i^{(2)} = d_{i,i}^{(2)} \alpha_i^{(2)}$ for $1 \leqslant i \leqslant w_n + u_0$. In particular, the entries of $\boldsymbol{\delta}_i$ with $1 \leqslant i \leqslant w_n$ are independent of all the variables $\alpha_{w_n+i}^{(2)}$ with $1 \leqslant i \leqslant u_0$. We may therefore interpret $\boldsymbol{\alpha}$ as the ordered pair $(\boldsymbol{\alpha}_{n+1}^{\dagger}, \boldsymbol{\alpha}_{n+1}^{*})$ with $\boldsymbol{\alpha}_{n+1}^{\dagger} = (\boldsymbol{\alpha}_i)_{1 \leqslant i \leqslant w_n}$ and $\boldsymbol{\alpha}_{n+1}^{*} = (\alpha_{w_n+i}^{(2)})_{1 \leqslant i \leqslant u_0}$. In this notation we can write

$$\oint \prod_{j=1}^{n+1} \prod_{h=1}^{u_j} |g_k(\boldsymbol{\delta}_{w_{j-1}+h}; X)|^{2t_j} \, \mathrm{d}\boldsymbol{\alpha} = \oint \mathcal{F}_{n+1}(\boldsymbol{\alpha}_{n+1}^{\dagger}) G_{n+1}(\boldsymbol{\alpha}_{n+1}^{\dagger}; X) \, \mathrm{d}\boldsymbol{\alpha}_{n+1}^{\dagger}, \tag{2.10}$$

where

$$\mathcal{F}_{n+1}(\boldsymbol{\alpha}_{n+1}^{\dagger}) = \prod_{m=1}^{n} \prod_{i=1}^{u_m} |g_k(\boldsymbol{\delta}_{w_{m-1}+i}; X)|^{2t_m}$$

and

$$G_{n+1}(\boldsymbol{\alpha}_{n+1}^{\dagger}; X) = \oint \prod_{i=1}^{u_0} |g_k(\boldsymbol{\delta}_{w_n+i}; X)|^4 \, \mathrm{d}\boldsymbol{\alpha}_{n+1}^{*}.$$

The latter mean value counts integer solutions $\mathbf{x}, \mathbf{y} \in [-X, X]^{2u_0}$ of the system

$$x_{i1}^2 + x_{i2}^2 - y_{i1}^2 - y_{i2}^2 = 0 \qquad (1 \leqslant i \leqslant u_0),$$

where each solution is counted with a unimodular weight depending on $\boldsymbol{\alpha}_{n+1}^{\dagger}$. It then follows from the triangle inequality and Hua's lemma that

$$G_{n+1}(\boldsymbol{\alpha}_{n+1}^{\dagger}; X) \leqslant G_{n+1}(\mathbf{0}; X) \ll X^{2u_0+\varepsilon}. \tag{2.11}$$

We now iterate this procedure for $j = n, n-1, \dots, 1$. For each index j, we see from (1.7) that $r_l > w_{j-1}$ only when $l \leqslant k_j$. Moreover, we have $r_l \geqslant w_j$ if $l \leqslant k_j - 1$, and $r_l = w_{j-1} + v_j$ if $l = k_j$. Since we had arranged for the first $r_l \times r_l$ submatrices of each $D^{(l)}$ to be diagonal, it follows that the entries of $\boldsymbol{\delta}_i$ with $1 \leqslant i \leqslant w_{j-1}$ are independent of all the variables $\alpha^{(l)}_{w_{j-1}+i}$ with $1\leqslant i\leqslant u_j$ and $2\leqslant l\leqslant k_j-1$, and also of all $\alpha^{(k_j)}_{w_{j-1}+i}$ with $1\leqslant i\leqslant v_j$. Together, these latter groups of variables form the vectors $\boldsymbol{\alpha}_i$ with $w_{j-1}+1\leqslant i\leqslant w_j$. Hence by a similar argument to that encountered before, we can write $\boldsymbol{\alpha}^\dagger_{j+1}=(\boldsymbol{\alpha}^\dagger_j,\boldsymbol{\alpha}^*_j)$, where $\boldsymbol{\alpha}^\dagger_j=(\boldsymbol{\alpha}_i)_{1\leqslant i\leqslant w_{j-1}}$ and $\boldsymbol{\alpha}^*_j=(\boldsymbol{\alpha}_{w_{j-1}+i})_{1\leqslant i\leqslant u_j}$, noting in particular that the vector $\boldsymbol{\alpha}^\dagger_1$ is empty. For $2\leqslant j\leqslant n$ put

$$\mathcal{F}_j(\boldsymbol{\alpha}^\dagger_j)=\prod_{m=1}^{j-1}\prod_{i=1}^{u_m}|g_k(\boldsymbol{\delta}_{w_{m-1}+i};X)|^{2t_m},$$

and take $\mathcal{F}_1(\boldsymbol{\alpha}^\dagger_1)=1$. Also, let

$$G_j(\boldsymbol{\alpha}^\dagger_j;X)=\oint\prod_{i=1}^{u_j}|g_k(\boldsymbol{\delta}_{w_{j-1}+i};X)|^{2t_j}\,\mathrm{d}\boldsymbol{\alpha}^*_j\qquad(1\leqslant j\leqslant n).$$

Note that this mean value counts integer solutions $|\mathbf{x}|,|\mathbf{y}|\leqslant X$ to the system

$$\begin{aligned}\sum_{i=1}^{u_j}d^{(k_j)}_{w_{j-1}+i,m}\sum_{h=1}^{t_j}(x^{k_j}_{i,h}-y^{k_j}_{i,h})&=0\qquad(1\leqslant m\leqslant v_j),\\ \sum_{i=1}^{u_j}d^{(l)}_{w_{j-1}+i,m}\sum_{h=1}^{t_j}(x^{l}_{i,h}-y^{l}_{i,h})&=0\qquad(1\leqslant m\leqslant u_j,\ 2\leqslant l\leqslant k_j-1),\end{aligned}$$

where each solution is counted with a unimodular weight depending on $\boldsymbol{\alpha}^\dagger_j$. An application of the triangle inequality shows that $G_j(\boldsymbol{\alpha}^\dagger_j;X)\leqslant G_j(\mathbf{0};X)=G^{v_j,u_j}_{t_j,k_j}(X)$. We thus deduce that for $1\leqslant j\leqslant n$ we have

$$\oint\mathcal{F}_{j+1}(\boldsymbol{\alpha}^\dagger_{j+1})\,\mathrm{d}\boldsymbol{\alpha}^\dagger_{j+1}=\oint\mathcal{F}_j(\boldsymbol{\alpha}^\dagger_j)G_j(\boldsymbol{\alpha}^\dagger_j;X)\,\mathrm{d}\boldsymbol{\alpha}^\dagger_j\ll G^{v_j,u_j}_{t_j,k_j}(X)\oint\mathcal{F}_j(\boldsymbol{\alpha}^\dagger_j)\,\mathrm{d}\boldsymbol{\alpha}^\dagger_j,$$

and upon iterating we find that

$$\oint\mathcal{F}_{n+1}(\boldsymbol{\alpha}^\dagger_{n+1})\,\mathrm{d}\boldsymbol{\alpha}^\dagger_{n+1}\ll\prod_{j=1}^n G^{v_j,u_j}_{t_j,k_j}(X).$$

The conclusion of the lemma follows upon combining this bound with (2.9), (2.10) and (2.11). □

3. The underlying mean value

From Lemma 2.3 it is clear that the desired bound $\mathcal{G}^{\mathbf{v},\mathbf{u}}_{\mathbf{t},\mathbf{k}}(X)\ll X^{K+\varepsilon}$ will follow if we can show that $G^{v_j,u_j}_{t_j,k_j}(X)\ll X^{K_j+\varepsilon}$ for $j=1,\dots,n$. We thus proceed to establish the latter bound. In the discussion of Lemmata 3.1 and 3.2 that follows, it is expedient to drop all mention of the indices j with $1\leqslant j\leqslant n$. Note also that in this situation, we have $r_k=v$ and $r_l=u$ for $2\leqslant l\leqslant k-1$. We introduce variables $\boldsymbol{\alpha}^{(1)}\in[0,1)^u$ and define $D^{(1)}$ to be the $u\times u$ identity matrix. Set further $\boldsymbol{\delta}^{(1)}=\boldsymbol{\alpha}^{(1)}$, and extend our previous notational conventions surrounding the vector $\boldsymbol{\delta}$ so as to incorporate $\boldsymbol{\delta}^{(1)}$ in the natural manner.

Next, we define

$$H_{t,k}^{v,u}(X)=\oint\prod_{j=1}^{u}|f_k(\boldsymbol{\delta}_j;2X)|^{2t}K_k(-\boldsymbol{\delta}_j;X,2tX)\,\mathrm{d}\boldsymbol{\alpha}. \tag{3.1}$$

We begin by establishing the bound contained in the following lemma.

Lemma 3.1. *One has* $G_{t,k}^{v,u}(X)\ll X^{-u}H_{t,k}^{v,u}(X)$.

Proof. Define ω_l to be 1 when $l=1$, and 0 otherwise. We decompose the set $\{1,\ldots,tu\}$ into the blocks $\mathcal{B}_m=\{(m-1)t+1,\ldots,mt\}$ for $1\leqslant m\leqslant u$. The mean value $G_{t,k}^{v,u}(X)$ counts the number of integral solutions of the system of equations

$$\sum_{m=1}^{u}d_{j,m}^{(l)}\xi_m^{(l)}=\omega_l h_j\qquad(1\leqslant j\leqslant r_l,\ 1\leqslant l\leqslant k), \tag{3.2}$$

where

$$\xi_m^{(l)}=\sum_{i\in\mathcal{B}_m}(x_i^l-y_i^l)\qquad(1\leqslant m\leqslant u,\ 1\leqslant l\leqslant k),$$

with $-X\leqslant x_i,y_i\leqslant X$ for $1\leqslant i\leqslant tu$ and $|h_j|\leqslant 2tX$ for $1\leqslant j\leqslant u$. Observe that in our current situation all coefficient matrices $D^{(l)}$ with $1\leqslant l\leqslant k-1$ are of format $u\times u$. Just as in the proof of Lemma 2.3, we can therefore assume without loss of generality that the coefficients $d_{j,m}^{(l)}$ with $1\leqslant l\leqslant k-1$ vanish except when $j=m$. Also, note that the constraints on the expressions $\xi_m^{(1)}$ for $(1\leqslant m\leqslant u)$ imposed by the linear equations in (3.2) are void, since the ranges for the new variables h_j automatically accommodate all possible values for $\xi_m^{(1)}$within (3.2).

We now consider the effect of shifting every variable with index in a given block $\mathcal{B}_m$ by an integer z_m with $|z_m|\leqslant X$. By the binomial theorem, for any family of shifts $\mathbf{z}$, one finds that $(\mathbf{x},\mathbf{y})$ is a solution of (3.2) if and only if it is also a solution of the system

$$\sum_{m=1}^{u}d_{j,m}^{(l)}\zeta_m^{(l)}=\sum_{m=1}^{u}d_{j,m}^{(l)}lh_mz_m^{l-1}\qquad(1\leqslant j\leqslant r_l,\ 1\leqslant l\leqslant k),$$

where

$$\zeta_m^{(l)}=\sum_{i\in\mathcal{B}_m}\big((x_i+z_m)^l-(y_i+z_m)^l\big)\qquad(1\leqslant m\leqslant u,\ 1\leqslant l\leqslant k).$$

Thus, for each fixed integer u-tuple $\mathbf{z}$ with $|z_m|\leqslant X$ $(1\leqslant m\leqslant u)$, the mean value $G_{t,k}^{v,u}(X)$ is bounded above by the number of integral solutions of the system

$$\sum_{m=1}^{u}d_{j,m}^{(l)}\sum_{i\in\mathcal{B}_m}(v_i^l-w_i^l)=\sum_{m=1}^{u}d_{j,m}^{(l)}lh_mz_m^{l-1}\qquad(1\leqslant j\leqslant r_l,\ 1\leqslant l\leqslant k)$$

with $|\mathbf{v}|,|\mathbf{w}| \leqslant 2X$ and $|\mathbf{h}| \leqslant 2tX$. On applying orthogonality and averaging over all possible choices for $\mathbf{z}$, we therefore infer that

$$G_{t,k}^{v,u}(X) \ll X^{-u} \sum_{|\mathbf{z}|\leqslant X} \oint \prod_{m=1}^{u} |f_k(\boldsymbol{\delta}_m;2X)|^{2t} \mathfrak{k}(-\boldsymbol{\delta}_m;z_m)\,\mathrm{d}\boldsymbol{\alpha},$$

where

$$\mathfrak{k}(\boldsymbol{\alpha};z) = \sum_{|h|\leqslant 2tX} e(h\alpha^{(1)} + 2hz\alpha^{(2)} + \ldots + khz^{k-1}\alpha^{(k)}).$$

The proof of the lemma is completed by reference to (2.1) and (3.1). □

We can now turn to the task of estimating $H_{t,k}^{v,u}(X)$. We will do this in somewhat wider generality than is required for the proofs of Theorems 1.3 and 1.5. This will allow us to prove the unconditional results adumbrated in the introduction.

When $l \geqslant 2$ is an integer and $\sigma \geqslant m \geqslant 0$, denote by $M_{l,\sigma,m}(X)$ the mean value

$$M_{l,\sigma,m}(X) = \oint |f_l(\boldsymbol{\alpha};X)|^{2\sigma-2m} K_l(\boldsymbol{\alpha};X,X)^m\,\mathrm{d}\boldsymbol{\alpha}. \tag{3.3}$$

When σ and m are integers, this mean value counts the number of integer tuples $\mathbf{x},\mathbf{y} \in [-X,X]^{\sigma-m}$ and $\mathbf{z},\mathbf{h} \in [-X,X]^m$ satisfying

$$\sum_{i=1}^{\sigma-m}(x_i^j - y_i^j) + \sum_{i=1}^{m} jh_i z_i^{j-1} = 0 \qquad (1 \leqslant j \leqslant l).$$

In particular, we have $M_l^*(X) = M_{l,\frac{1}{2}l(l+1),2}(X)$. The main conjecture for mean values of the shape (3.3) states that for all $\sigma \geqslant m$ one should have

$$M_{l,\sigma,m}(X) \ll X^{\varepsilon}(X^{\sigma} + X^{2\sigma-l(l+1)/2}). \tag{3.4}$$

Note that the case when $m=0$ corresponds to Vinogradov's mean value theorem, and in this case the bound (3.4) is known (see equation (2.2) above).

Suppose that σ and m are integers with $2 \leqslant m \leqslant \sigma$ and

$$m|(\sigma - \tfrac{1}{2}k(k-1)). \tag{3.5}$$

We now choose

$$t = K/u + (\sigma - \tfrac{1}{2}k(k-1))/m, \tag{3.6}$$

so that at the critical point $\sigma = \frac{1}{2}k(k-1)$ we have $tu = K$. Note also that by (3.5) as well as the definition of K in (1.4), the quantity t is indeed an integer whenever $u|kv$.

Lemma 3.2. *Let σ and m be integers with $2 \leqslant m \leqslant \sigma$ and satisfying the conditions $2|m$ and $m|(\sigma - \frac{1}{2}k(k-1))$. Assume also that $u \geqslant \frac{m}{m-1}v$ and $u|kv$. Then we have*

$$H_{t,k}^{v,u}(X) \ll X^{K+u+\varepsilon}\left(\frac{M_{k-1,\sigma,m}(2tX)}{X^{k(k-1)/2}}\right)^{u/m}.$$

Proof. Set

$$\mathfrak{G}_1(\boldsymbol{\delta})=\prod_{j=1}^{v}|f_k(\boldsymbol{\delta}_j;2X)|^{k(k+1)} \tag{3.7}$$

and

$$\mathfrak{G}_2(\boldsymbol{\delta})=\prod_{j=v+1}^{u}\left|f_k(\boldsymbol{\delta}_j;2X)^{2t-k(k+1)v/u}K_k(\boldsymbol{\delta}_j;X,2tX)\right|^{u/(u-v)}.$$

Then it follows from (3.1) via (2.3) that, after possibly relabelling variables, we have

$$H_{t,k}^{v,u}(X)\ll\oint\mathfrak{G}_1(\boldsymbol{\delta})\mathfrak{G}_2(\boldsymbol{\delta})\,\mathrm{d}\boldsymbol{\alpha}. \tag{3.8}$$

Recall now that we had arranged for the coefficient matrices $D^{(1)},\dots,D^{(k-1)}$ to be diagonal. Consequently, the variables $\boldsymbol{\delta}_1,\dots,\boldsymbol{\delta}_v$ are independent of those $\alpha_i^{(l)}$ having $1\leqslant l\leqslant k-1$ and $v+1\leqslant i\leqslant u$. Then, by setting $\boldsymbol{\eta}_1=(\boldsymbol{\alpha}_i)_{1\leqslant i\leqslant v}$ and $\boldsymbol{\eta}_2=(\boldsymbol{\alpha}_i)_{v+1\leqslant i\leqslant u}$, it follows that $\boldsymbol{\eta}_1$ fully determines $\boldsymbol{\delta}_1,\dots,\boldsymbol{\delta}_v$, and $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ together completely determine all entries of $\boldsymbol{\delta}$. On recalling (3.7), we may thus rewrite the integral on the right hand side of (3.8) to obtain the bound

$$H_{t,k}^{v,u}(X)\ll\oint\mathfrak{H}_1(\boldsymbol{\eta}_1)\mathfrak{H}_2(\boldsymbol{\eta}_1)\,\mathrm{d}\boldsymbol{\eta}_1, \tag{3.9}$$

where $\mathfrak{H}_1(\boldsymbol{\eta}_1)=\mathfrak{G}_1(\boldsymbol{\delta})$ and

$$\mathfrak{H}_2(\boldsymbol{\eta}_1)=\oint\mathfrak{G}_2(\boldsymbol{\delta})\,\mathrm{d}\boldsymbol{\eta}_2.$$

Define

$$U_1(\boldsymbol{\eta}_1)=\oint\prod_{j=v+1}^{u}|f_k(\boldsymbol{\delta}_j;2X)|^{k(k-1)}\,\mathrm{d}\boldsymbol{\eta}_2$$

and

$$U_2(\boldsymbol{\eta}_1)=\oint\prod_{j=v+1}^{u}|f_k(\boldsymbol{\delta}_j;2X)^{2\sigma-2m}K_k(\boldsymbol{\delta}_j;X,2tX)^m|\,\mathrm{d}\boldsymbol{\eta}_2.$$

Also, write

$$\omega=\frac{u}{m(u-v)},$$

and note that, as a consequence of (1.4) and (3.6), one has

$$\begin{aligned}(2\sigma-2m)\omega+k(k-1)(1-\omega)&=\omega(2mt-2Km/u-2m)+k(k-1)\\&=\frac{2tu-(k(k-1)u+2kv-2u)-2u}{u-v}+k(k-1)\\&=(2t-k(k+1)v/u)\,\frac{u}{u-v}.\end{aligned}$$

Then, since $m\geqslant 2$ and $u\geqslant\frac{m}{m-1}v$, it follows via Hölder's inequality that

$$\mathfrak{H}_2(\boldsymbol{\eta}_1)\ll U_1(\boldsymbol{\eta}_1)^{1-\omega}U_2(\boldsymbol{\eta}_1)^{\omega}. \tag{3.10}$$

Since m is an even integer, it follows by standard orthogonality considerations that $U_1(\boldsymbol{\eta}_1)$ and $U_2(\boldsymbol{\eta}_1)$ count solutions to their respective associated systems of equations with degrees $1,\ldots,k-1$, with each solution being counted with a unimodular weight depending on $\boldsymbol{\eta}_1$. It thus follows from the triangle inequality that $U_i(\boldsymbol{\eta}_1)\leqslant U_i(\mathbf{0})$ for $i=1,2$. Using the fact that the coefficient matrices $D^{(l)}$ with $1\leqslant l\leqslant k-1$ are all diagonal, and recalling (2.2), we thus discern that

$$U_1(\boldsymbol{\eta}_1)\ll \oint \prod_{j=v+1}^{u} |f_{k-1}(\boldsymbol{\alpha}_j;2X)|^{k(k-1)}\,\mathrm{d}\boldsymbol{\eta}_2 \ll X^{\frac{1}{2}k(k-1)(u-v)+\varepsilon}. \tag{3.11}$$

By an analogous chain of reasoning, we derive from the definition (3.3) of $M_{l,\sigma,m}(X)$ and a consideration of the underlying system of equations the corresponding bound

$$\begin{aligned}U_2(\boldsymbol{\eta}_1)&\ll \oint \prod_{j=v+1}^{u} |f_{k-1}(\boldsymbol{\alpha}_j;2X)^{2\sigma-2m}K_{k-1}(\boldsymbol{\alpha}_j;X,2tX)^m|\,\mathrm{d}\boldsymbol{\eta}_2\\ &\ll \big(M_{k-1,\sigma,m}(2tX)\big)^{u-v}.\end{aligned} \tag{3.12}$$

Thus, from (3.10), (3.11) and (3.12) we have

$$\begin{aligned}\mathfrak{H}_2(\boldsymbol{\eta}_1)&\ll X^{\frac{1}{2}k(k-1)(u-v)(1-\omega)+\varepsilon}M_{k-1,\sigma,m}(2tX)^{(u-v)\omega}\\ &\ll X^{\frac{1}{2}k(k-1)(u-v)+\varepsilon}\left(\frac{M_{k-1,\sigma,m}(2tX)}{X^{k(k-1)/2}}\right)^{u/m}.\end{aligned}$$

At this stage in our argument, we discern from (3.9) that

$$H_{t,k}^{v,u}(X)\ll X^{\frac{1}{2}k(k-1)(u-v)+\varepsilon}\left(\frac{M_{k-1,\sigma,m}(2tX)}{X^{k(k-1)/2}}\right)^{u/m}\oint \mathfrak{H}_1(\boldsymbol{\eta}_1)\,\mathrm{d}\boldsymbol{\eta}_1. \tag{3.13}$$

Recall that $\mathfrak{H}_1(\boldsymbol{\eta}_1)=\mathfrak{G}_1(\boldsymbol{\delta})$, where $\mathfrak{G}_1(\boldsymbol{\delta})$ is defined by (3.7). Since the first $v\times v$ minors of the coefficient matrices $D^{(l)}$ for $1\leqslant l\leqslant k$ are now diagonal, we deduce from (2.2) that

$$\oint \mathfrak{H}_1(\boldsymbol{\eta}_1)\,\mathrm{d}\boldsymbol{\eta}_1 \ll (J_{k(k+1)/2,k}(2X))^v \ll X^{\frac{1}{2}k(k+1)v+\varepsilon}. \tag{3.14}$$

Finally, on substituting (3.14) into (3.13) and recalling (1.4), we conclude that

$$H_{t,k}^{v,u}(X)\ll X^{K+u+\varepsilon}\left(\frac{M_{k-1,\sigma,m}(2tX)}{X^{k(k-1)/2}}\right)^{u/m}.$$

This completes the proof of the lemma. □

We now resume the practice of appending the suffix j to the parameters k, u, v and K that we temporarily abandoned during the discussion of Lemmata 3.1 and 3.2. We assume, moreover, that σ_j and m_j are integers with $2\leqslant m_j\leqslant \sigma_j$ and

$$m_j|(\sigma_j-\tfrac{1}{2}k_j(k_j-1)).$$

In accordance with (3.6), we now fix the parameters t_j by taking

$$t_j=K_j/u_j+(\sigma_j-\tfrac{1}{2}k_j(k_j-1))/m_j\quad (1\leqslant j\leqslant n). \tag{3.15}$$

Hence, whenever $u_j|k_jv_j$, the quantity t_j is an integer. With these natural numbers t_j defined thus, we recall the definition of s_0 from (2.7). We are now equipped to provide an unconditional version of Theorem 1.3

Theorem 3.3. *Suppose that $k_1 > \ldots > k_n \geqslant 3$. Assume further that $\mathbf{u}, \mathbf{v}, \mathbf{m}, \boldsymbol{\sigma} \in \mathbb{N}^n$ satisfy the relations*

$$2|m_j, \quad 2 \leqslant m_j \leqslant \sigma_j, \quad u_j \geqslant \frac{m_j}{m_j - 1} v_j, \quad u_j|k_jv_j \quad \textit{and} \quad m_j|(\sigma_j - \tfrac{1}{2}k_j(k_j-1))$$

for $1 \leqslant j \leqslant n$. Let u_0 be a non-negative integer. Then for any $\varepsilon > 0$, one has

$$I_{s_0,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) \ll X^{K+\varepsilon} \prod_{j=1}^n \left(\frac{M_{k_j-1,\sigma_j,m_j}(2t_jX)}{X^{k_j(k_j-1)/2}} \right)^{u_j/m_j}.$$

Proof. We apply Lemma 2.2 with $s = s_0$, followed by Lemmata 2.3, 3.1 and 3.2. This shows that

$$I_{s_0,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) \ll X^{2u_0+\varepsilon} \prod_{j=1}^n X^{K_j} \left(\frac{M_{k_j-1,\sigma_j,m_j}(2t_jX)}{X^{k_j(k_j-1)/2}} \right)^{u_j/m_j},$$

and the proof is complete upon reference to (1.9). □

We can now complete the proof of Theorem 1.3. To this end, we choose $m_j = 2$ and $\sigma_j = \frac{1}{2}k_j(k_j-1)$ for $1 \leqslant j \leqslant n$. With this choice of parameters the hypotheses of Theorem 3.3 are satisfied whenever $\mathbf{k}, \mathbf{v}$ and $\mathbf{u}$ are in accordance with the conditions of Theorem 1.3, and moreover the conjectural bound $M_{k_j-1,\sigma_j,2}(2t_jX) \ll X^{\sigma_j+\varepsilon}$ is then tantamount to both (1.6) and (3.4). Thus, in the case $s = s_0$ the desired conclusion is an immediate consequence of the conclusion of Theorem 3.3, and for general values of s it follows in like manner upon utilising the additional flexibility offered by Lemma 2.2.

4. The Hardy-Littlewood method

We can now initiate the derivation of Theorem 1.5 from the mean value estimate of Theorem 1.3. We shall prove the following rather more general result.

Theorem 4.1. *Suppose that $k_1 > \ldots > k_n \geqslant 3$. Suppose further that $\mathbf{u}, \mathbf{v}, \mathbf{m}, \boldsymbol{\sigma} \in \mathbb{N}^n$ lie in the respective ranges*

$$2 \leqslant m_j \leqslant \sigma_j, \quad \sigma_j \geqslant \tfrac{1}{2}k_j(k_j-1) \quad \textit{and} \quad u_j \geqslant \frac{m_j}{m_j-1}v_j,$$

and satisfy the divisibility conditions

$$2|m_j, \quad u_j|k_jv_j \quad \textit{and} \quad m_j|(\sigma_j - \tfrac{1}{2}k_j(k_j-1)),$$

for $1 \leqslant j \leqslant n$. Assume, moreover, that

$$M_{k_j-1,\sigma_j,m_j}(X) \ll X^{2\sigma_j - \frac{1}{2}k_j(k_j-1)+\varepsilon} \quad (1 \leqslant j \leqslant n). \tag{4.1}$$

Let u_0 be a non-negative integer, put $t_0 = 2$ and define t_j via (3.15) for $1 \leqslant j \leqslant n$. Set $s_0 = t_0u_0 + \ldots + t_nu_n$, suppose that $s \geqslant 2s_0 + 1$ and write $K = 2u_0 + K_1 + \ldots + K_n$ for the total degree of the system as usual. Then the asymptotic formula

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) = (\mathcal{C} + o(1))X^{s-K}$$

holds with $\mathcal{C} \geqslant 0$. If, furthermore, the system (1.10) has non-singular solutions in $\mathbb{R}$ as well as in the fields $\mathbb{Q}_p$ for all p, then the constant $\mathcal{C}$ is positive.

Note that Theorem 1.5 follows from the special case of Theorem 4.1 in which $m_j = 2$ and $\sigma_j = \frac{1}{2}k_j(k_j - 1)$ for $1 \leqslant j \leqslant n$.

We make use of the notation introduced in §2, and recall in particular (2.6) and its sequel. From now on we will set $r = r_2 + \ldots + r_k$ and $w = u_0 + \ldots + u_n$, so that $w = r_2 = w_n + u_0$. Also, we will assume throughout that $\mathbf{k}$, $\mathbf{v}$, $\mathbf{u}$, $\boldsymbol{\sigma}$, $\mathbf{m}$ satisfy the hypotheses of the statement of Theorem 4.1.

When $\mathfrak{B} \subseteq [0, 1)^r$ is a measurable set, put

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X; \mathfrak{B}) = \int_{\mathfrak{B}} \prod_{j=1}^{s} g_k(\boldsymbol{\gamma}_j; X) \,\mathrm{d}\boldsymbol{\alpha}. \tag{4.2}$$

Our Hardy–Littlewood dissection is defined as follows. When Y and Q are parameters with $1 \leqslant Q \leqslant Y$, we take the major arcs $\mathfrak{M}_Y = \mathfrak{M}_Y(Q)$ to be the union of the boxes

$$\mathfrak{M}_Y(q, \mathbf{a}) = \left\{ \boldsymbol{\alpha} \in [0, 1)^r : |\alpha_j^{(l)} - a_j^{(l)}/q| \leqslant QY^{-l} \quad (1 \leqslant j \leqslant r_l,\, 2 \leqslant l \leqslant k) \right\}, \tag{4.3}$$

with $0 \leqslant \mathbf{a} \leqslant q \leqslant Q$ and $(q, \mathbf{a}) = 1$. The corresponding set of minor arcs $\mathfrak{m}_Y = \mathfrak{m}_Y(Q)$ is defined by putting $\mathfrak{m}_Y(Q) = [0, 1)^r \setminus \mathfrak{M}_Y(Q)$. Unless indicated otherwise, we fix $Y = X$ and $Q = X^{1/(6r)}$, and abbreviate $\mathfrak{M}_X$ to $\mathfrak{M}$ and $\mathfrak{m}_X$ to $\mathfrak{m}$.

We require certain auxiliary functions in order to analyse the contribution of the major arcs $N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X; \mathfrak{M})$. Write

$$S_k(q, \mathbf{a}) = \sum_{x=1}^{q} e((a^{(2)}x^2 + \ldots + a^{(k)}x^k)/q),$$

and recall that the argument of [11, Theorem 7.1] gives

$$S_k(q, \mathbf{a}) \ll (q, \mathbf{a})^{1/k} q^{1-1/k+\varepsilon}. \tag{4.4}$$

Further, set

$$v_k(\boldsymbol{\beta}; X) = \int_{-X}^{X} e(\beta^{(2)}z^2 + \ldots + \beta^{(k)}z^k) \,\mathrm{d}z,$$

and recall from the arguments of [11, Theorem 7.3] the estimate

$$v_k(\boldsymbol{\beta}; X) \ll X \left(1 + |\beta^{(2)}|X^2 + \ldots + |\beta^{(k)}|X^k\right)^{-1/k}. \tag{4.5}$$

We put

$$\Lambda_j^{(l)} = \sum_{i=1}^{r_l} c_{i,j}^{(l)} a_i^{(l)} \quad \text{and} \quad \vartheta_j^{(l)} = \sum_{i=1}^{r_l} c_{i,j}^{(l)} \beta_i^{(l)} \qquad (1 \leqslant j \leqslant s,\, 2 \leqslant l \leqslant k). \tag{4.6}$$

Following the same convention regarding vector notation as we applied for $\boldsymbol{\gamma}$ in (2.6) and its sequel, we have $\boldsymbol{\vartheta} = \boldsymbol{\gamma} - \boldsymbol{\Lambda}/q$. Then as a consequence of [11, Theorem 7.2], we find that when $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta} \in \mathfrak{M}$, one has

$$g_k(\boldsymbol{\gamma}_j; X) = q^{-1} S_k(q, \boldsymbol{\Lambda}_j) v_k(\boldsymbol{\vartheta}_j; X) + O(Q^2). \tag{4.7}$$

Finally, define

$$\mathfrak{S}(Q) = \sum_{q \leqslant Q} \sum_{\substack{1 \leqslant \mathbf{a} \leqslant q \\ (q, \mathbf{a}) = 1}} \prod_{j=1}^{s} q^{-1} S_k(q, \boldsymbol{\Lambda}_j) \tag{4.8}$$

and

$$\mathfrak{J}_X(Q) = \int_{\mathcal{I}(X,Q)} \prod_{j=1}^{s} v_k(\boldsymbol{\vartheta}_j; X) \, \mathrm{d}\boldsymbol{\beta},$$

where

$$\mathcal{I}(X, Q) = \prod_{l=2}^{k} [-QX^{-l}, QX^{-l}]^{r_l}.$$

The preliminary conclusion of our major arcs analysis is summarised in the following lemma.

Lemma 4.2. *There is a positive number* ω *for which*

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X; \mathfrak{M}) = X^{s-K} \mathfrak{S}(Q) \mathfrak{J}_1(Q) + O(X^{s-K-\omega}).$$

Proof. Since $\mathrm{vol}(\mathfrak{M}) \ll Q^{2r+1} X^{-K}$, it follows from (4.7) that

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X; \mathfrak{M}) = \mathfrak{S}(Q) \mathfrak{J}_X(Q) + O(X^{s-K-1} Q^{2r+3}). \tag{4.9}$$

Furthermore, by a change of variables we see that $\mathfrak{J}_X(Q) = X^{s-K} \mathfrak{J}_1(Q)$. The conclusion of the lemma therefore follows from our choice $Q = X^{1/(6r)}$. □

In order to address the contribution of the minor arcs, we need the following Weyl-type estimate.

Lemma 4.3. *Suppose that* $\boldsymbol{\alpha} \in \mathfrak{m}$. *There exists* $\tau > 0$ *such that for each* w*-tuple* $(j_1, \ldots, j_w)$ *of distinct indices there exists an index* i *with* $1 \leqslant i \leqslant w$ *for which one has*

$$|g_k(\boldsymbol{\gamma}_{j_i}; X)| \leqslant X Q^{-\tau}.$$

Proof. This is the content of [3, Lemma 3.1]. Note that the minor arcs in our setting are a subset of the minor arcs defined in the context of that lemma. □

We now complete the analysis of the minor arcs for Theorem 4.1.

Lemma 4.4. *Assume the hypotheses of Theorem 4.1. Then there is a positive number* ω *for which* $N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X; \mathfrak{m}) \ll X^{s-K-\omega}$.

Proof. Given a measurable set $\mathfrak{B} \subseteq [0,1)^r$, we write

$$N^*(X;\mathfrak{B}) = \int_{\mathfrak{B}} \prod_{j=1}^{2s_0+1} |g_k(\boldsymbol{\gamma}_j;X)| \, \mathrm{d}\boldsymbol{\alpha}.$$

We begin by estimating the last $s-(2s_0+1)$ exponential sums in the product (4.2) trivially, so that

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X;\mathfrak{m}) \ll X^{s-(2s_0+1)} N^*(X;\mathfrak{m}). \tag{4.10}$$

For $1 \leqslant i \leqslant w$ and $\tau > 0$ sufficiently small, let $\mathfrak{m}^{(i)}$ denote the set of $\boldsymbol{\alpha} \in [0,1)^r$ for which $|g_k(\boldsymbol{\gamma}_i;X)| \leqslant XQ^{-\tau}$. In view of (2.3), we can identify a subset of indices $\mathcal{J}_i \subseteq \{1,\ldots,2s_0+1\} \setminus \{i\}$ with $\mathrm{card}(\mathcal{J}_i) = s_0$ for which

$$N^*(X;\mathfrak{m}^{(i)}) \ll XQ^{-\tau} \oint \prod_{j\in\mathcal{J}_i} |g_k(\boldsymbol{\gamma}_j;X)|^2 \, \mathrm{d}\boldsymbol{\alpha}.$$

Write $C_i^{(l)}$ for the submatrix of $C^{(l)}$ having columns indexed by $\mathcal{J}_i$. The condition that the coefficient matrices $C^{(l)}$ be highly non-singular implies that the submatrices $C_i^{(l)}$ of $C^{(l)}$ are also highly non-singular. Thus, by orthogonality, we see from the definition (1.8) of the mean value $I_{s_0,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X)$ that

$$N^*(X;\mathfrak{m}^{(i)}) \ll XQ^{-\tau} I_{s_0,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X;C_i^{(2)},\ldots,C_i^{(k)}).$$

Consider a fixed $\boldsymbol{\alpha} \in \mathfrak{m}$. If τ has been chosen sufficiently small, Lemma 4.3 ensures that we can find an index j with $1 \leqslant j \leqslant w$ such that $\boldsymbol{\alpha} \in \mathfrak{m}^{(j)}$. Thus we see that we have the inclusion $\mathfrak{m} \subseteq \mathfrak{m}^{(1)} \cup \cdots \cup \mathfrak{m}^{(w)}$, whence

$$N^*(X;\mathfrak{m}) \ll XQ^{-\tau} \max_{1\leqslant i\leqslant w} I_{s_0,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X;C_i^{(2)},\ldots,C_i^{(k)}). \tag{4.11}$$

Now recall that $Q = X^{1/(6r)}$. Note also that the hypotheses of Theorem 4.1 under which we are currently working permit the assumption of those of Theorem 3.3. Thus, upon combining the estimate (4.11) with Theorem 3.3, inserting (4.1) and recalling (3.15), we obtain the bound

$$\begin{aligned} N^*(X;\mathfrak{m}) &\ll XQ^{-\tau} X^{K+\varepsilon} \prod_{j=1}^{n} \left(X^{2\sigma_j - k_j(k_j-1)} \right)^{u_j/m_j} \\ &\ll XQ^{-\tau} X^{K+\varepsilon} \prod_{j=1}^{n} X^{2t_j u_j - 2K_j} \\ &\ll X^{2s_0+1-K} Q^{-\tau/2}. \end{aligned}$$

By substituting this estimate into (4.10), we obtain the conclusion of the lemma. □

Upon combining the results of Lemmata 4.2 and 4.4, we infer that for some $\omega > 0$ one has the asymptotic formula

$$N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) = X^{s-K} \mathfrak{S}(Q) \mathfrak{J}_1(Q) + O(X^{s-K-\omega}). \tag{4.12}$$

This completes our analysis of the minor arcs.

5. Initial considerations for the singular series

It remains to show that the singular series $\mathfrak{S}(Q)$ and singular integral $\mathfrak{J}_1(Q)$ converge as Q tends to infinity. We now put

$$\ell_j = \begin{cases} k_i & \text{when } w_{i-1}+1 \leqslant j \leqslant w_{i-1}+v_i, \\ k_i - 1 & \text{when } w_{i-1}+v_i+1 \leqslant j \leqslant w_i, \\ 2 & \text{when } w_n+1 \leqslant j \leqslant w. \end{cases}$$

In this notation, the system under consideration can be viewed as a superposition of w Vinogradov systems with respective degrees ℓ_j, all missing the linear slice, and thus it follows from the definition (1.9) that the total degree of this system is

$$K = \sum_{j=1}^{w} \big(\tfrac{1}{2}\ell_j(\ell_j+1)-1\big).$$

Throughout this and the next section, we work under the assumption that $s \geqslant 2K+1$.

We first attend to the singular series. Put

$$A(q) = q^{-s} \sum_{\substack{1\leqslant \mathbf{a} \leqslant q \\ (q,\mathbf{a})=1}} \prod_{j=1}^{s} S_k(q, \mathbf{\Lambda}_j). \tag{5.1}$$

By applying (2.3), we find that for some choice of distinct indices $j_1, \ldots, j_w \in \{1, \ldots, s\}$ we have the asymptotic bound

$$A(q) \ll q^{2K-s} \max_{\substack{1\leqslant \mathbf{a} \leqslant q \\ (q,\mathbf{a})=1}} \left(\prod_{i=1}^{w} |S_k(q, \mathbf{\Lambda}_{j_i})| \right)^{(s-2K)/w} A_1(q), \tag{5.2}$$

where

$$A_1(q) = q^{-2K} \sum_{\substack{1\leqslant \mathbf{a} \leqslant q \\ (q,\mathbf{a})=1}} \prod_{i=1}^{w} |S_k(q, \mathbf{\Lambda}_{j_i})|^{\ell_i(\ell_i+1)-2}.$$

Note that both $A(q)$ and $A_1(q)$ are multiplicative in q. For this reason, the key to understanding the singular series is to maintain good control over the multiplicative quantity

$$B_1(q) = \sum_{d|q} A_1(d) \tag{5.3}$$

as q runs over the prime powers.

Define τ_j by setting $\tau_j = \frac{1}{2}\ell_j(\ell_j+1)-1$ for $1 \leqslant j \leqslant w$, and write $T_j = \tau_1 + \ldots + \tau_j$, so that $T_w = K$. For consistency we also set $T_0 = 0$. Now, adopting a notation similar to that of Section 2, when $2 \leqslant l \leqslant k$ we write $D^{(l)}$ for the submatrices

$$\big(d_{i,h}^{(l)}\big)_{\substack{1\leqslant i \leqslant r_l \\ 1\leqslant h \leqslant w}} = \big(c_{i,j_h}^{(l)}\big)_{\substack{1\leqslant i \leqslant r_l \\ 1\leqslant h \leqslant w}}$$

of the coefficient matrices $C^{(l)}$ consisting of the columns indexed by $j_1, \ldots, j_w$. Note that the hypothesis that each $C^{(l)}$ is highly non-singular ensures that the same is true for each $D^{(l)}$. For $1 \leqslant h \leqslant w$ and $2 \leqslant l \leqslant k$ we set $\Delta_h^{(l)} = \Lambda_{j_h}^{(l)}$, and we employ the same conventions regarding vector notation as in (4.6) and also (2.6) and its sequel. Thus, we write $\mathbf{\Delta}_j = (\Delta_j^{(l)})_{2\leqslant l\leqslant k}$ and $\mathbf{\Delta}^{(l)} = (\Delta_j^{(l)})_{1\leqslant j\leqslant w}$, so that

$$\mathbf{\Delta}^{(l)} = (D^{(l)})^T \mathbf{a}^{(l)} \qquad (2 \leqslant l \leqslant k). \tag{5.4}$$

In this notation, it follows from standard orthogonality relations that

$$q^{2K-r} B_1(q) = q^{-r} \sum_{1\leqslant \mathbf{a}\leqslant q} \prod_{j=1}^{w} |S_k(q, \mathbf{\Delta}_j)|^{2\tau_j}$$

counts the number of solutions $\mathbf{x}, \mathbf{y} \in (\mathbb{Z}/q\mathbb{Z})^K$ of the system of congruences

$$\sum_{j=1}^{w} d_{i,j}^{(l)} \Biggl(\sum_{h=T_{j-1}+1}^{T_j} (x_h^l - y_h^l) \Biggr) \equiv 0 \pmod q, \tag{5.5}$$

where $1 \leqslant i \leqslant r_l$ and $2 \leqslant l \leqslant k$.

Our first goal is to apply a procedure inspired by the proof of Theorem 2.1 in [3] in order to disentangle the congruences in (5.5). This will enable us to replace the sum $B_1(q)$ by a related expression in which for all indices j the degree k in the exponential sum $S_k(q, \mathbf{\Delta}_j)$ is replaced by ℓ_j. Since ℓ_j is typically smaller than k, we will reap the rewards of this preparatory step when the reduced degrees allow us to exert greater control on the size of the exponential sums in question.

Given a $(k-1)$-tuple of variables $\xi^{(2)}, \ldots, \xi^{(k)}$, we adopt the convention that $\boldsymbol{\xi}^{[l]} = (\xi^{(2)}, \ldots, \xi^{(l)})$ for $2 \leqslant l \leqslant k$. Also, when $\mathbf{d} = (d_2, \ldots, d_k)$ is a coefficient vector, we abbreviate the vector $(d_2\xi^{(2)}, \ldots, d_k\xi^{(k)})$ to $\mathbf{d}\boldsymbol{\xi}$, and we appropriate the notation $\mathbf{d}^{[l]}$ and $(\mathbf{d}\boldsymbol{\xi})^{[l]}$ to denote the corresponding subvectors whose entries are indexed by $2 \leqslant i \leqslant l$. The following observation will play a part in our ensuing arguments.

Lemma 5.1. *Let l, q and t be natural numbers, with $2 \leqslant l \leqslant k-1$. Suppose that $d_2, \ldots, d_k$ and $c_2, \ldots, c_k$ are fixed integers, and put*

$$\Gamma_q(\mathbf{d}^{[l]}) = \prod_{j=2}^{l} (q, d_j).$$

Then for any fixed integers $a^{(l+1)}, \ldots, a^{(k)}$ we have

$$\sum_{1\leqslant \mathbf{a}^{[l]}\leqslant q} |S_k(q, \mathbf{d}\mathbf{a} + \mathbf{c})|^{2t} \leqslant \Gamma_q(\mathbf{d}^{[l]}) \sum_{1\leqslant \mathbf{a}^{[l]}\leqslant q} |S_l(q, \mathbf{a}^{[l]})|^{2t}.$$

Proof. By standard orthogonality relations, the sum

$$T = q^{1-l} \sum_{1\leqslant \mathbf{a}^{[l]}\leqslant q} |S_k(q, \mathbf{d}\mathbf{a} + \mathbf{c})|^{2t} \tag{5.6}$$

counts solutions $\mathbf{x}, \mathbf{y} \in (\mathbb{Z}/q\mathbb{Z})^t$ of the system of congruences

$$d_j \sum_{i=1}^{t} (x_i^j - y_i^j) \equiv 0 \pmod q \qquad (2 \leqslant j \leqslant l), \tag{5.7}$$

where each solution is counted with a unimodular weight depending on the inert variables $a^{(l+1)},\dots,a^{(k)}$, together with the coefficients $\mathbf{d}$ and $\mathbf{c}$. Thus, by the triangle inequality, one finds that

$$T \leqslant q^{1-l} \sum_{1\leqslant \mathbf{a}^{[l]}\leqslant q} |S_l(q,(\mathbf{d}\mathbf{a})^{[l]})|^{2t}.$$

We therefore discern that T is bounded above by the number of solutions of (5.7) counted without weights, and hence by the number of solutions $\mathbf{x},\mathbf{y}\in (\mathbb{Z}/q\mathbb{Z})^t$ of the system of congruences

$$\sum_{i=1}^t (x_i^j - y_i^j) \equiv 0 \pmod{q/(q,d_j)} \qquad (2\leqslant j\leqslant l).$$

We interpret the latter as the number of solutions of the system

$$\sum_{i=1}^t (x_i^j - y_i^j) \equiv \frac{e_j q}{(q,d_j)} \pmod{q} \qquad (2\leqslant j\leqslant l),$$

with $\mathbf{x},\mathbf{y}\in (\mathbb{Z}/q\mathbb{Z})^t$ and $1\leqslant e_j\leqslant (q,d_j)$ for $2\leqslant j\leqslant l$. Thus, by orthogonality and the triangle inequality, one sees that

$$\begin{aligned} T &\leqslant \sum_{\substack{1\leqslant e_j\leqslant (q,d_j)\\ (2\leqslant j\leqslant l)}} q^{1-l} \sum_{1\leqslant \mathbf{a}^{[l]}\leqslant q} |S_l(q,\mathbf{a}^{[l]})|^{2t} e\left(-\sum_{j=2}^l \frac{e_j a^{(j)}}{(q,d_j)}\right) \\ &\leqslant \Gamma_q(\mathbf{d}^{[l]}) q^{1-l} \sum_{1\leqslant \mathbf{a}^{[l]}\leqslant q} |S_l(q,\mathbf{a}^{[l]})|^{2t}. \end{aligned}$$

The conclusion of the lemma is now immediate from (5.6). □

We now define

$$B_1^*(q) = q^{-2K} \sum_{1\leqslant \mathbf{a}\leqslant q} \prod_{j=1}^w |S_{\ell_j}(q,\mathbf{a}_j^{[\ell_j]})|^{2\tau_j}.$$

The crucial bound for our analysis of the singular series is contained in the following lemma.

Lemma 5.2. *Let q be a natural number, and suppose that the matrices $D^{(l)}$ are all highly non-singular. Then there exists a finite set of primes $\Omega(D)$ and a natural number $\mathcal{R}(q)=\mathcal{R}(q,D)$, both depending at most on the coefficient matrices $D^{(l)}$ and in the latter case also q, with the property that*

$$B_1(q) \leqslant \mathcal{R}(q) B_1^*(q).$$

The constant $\mathcal{R}(q)$ is bounded above uniformly in q, and one can take $\mathcal{R}(q)=1$ whenever $(q,p)=1$ for all $p\in\Omega(D)$.

Proof. Recall that $q^{2K-r}B_1(q)$ counts the number of solutions $\mathbf{x},\mathbf{y}\in(\mathbb{Z}/q\mathbb{Z})^K$ of the system of congruences (5.5) for $1\leqslant j\leqslant r_l$ and $2\leqslant l\leqslant k$. Since $B_1(q)$ is a multiplicative function of q, it is apparent that it suffices to establish the conclusion of the lemma in the special case in which q is a prime power, say $q=p^h$ for a given prime p. By applying suitable elementary row operations within the coefficient matrices $D^{(l)}$ for $2\leqslant l\leqslant k$ that are invertible over $\mathbb{Z}/p^h\mathbb{Z}$, we may suppose without loss of generality that each coefficient matrix $D^{(l)}$ is in upper row echelon form. This operation corresponds to taking appropriate linear combinations of the congruences comprising (5.5). Here, we stress that the property that each $D^{(l)}$ is highly non-singular implies that the first $r_l\times r_l$ submatrix of $D^{(l)}$ is now upper triangular. We denote this matrix by $D_0^{(l)}$. Note that the power of p dividing the diagonal entries of $D_0^{(l)}$ depends only on the first $r_l\times r_l$ submatrices of the original coefficient matrices $D^{(l)}$. In particular, by defining $\Omega(D)$ to be the set of all primes dividing any of the determinants of the latter submatrices, we ensure that when $p\notin\Omega(D)$, then none of the diagonal entries of $D_0^{(l)}$ is divisible by p.

We now employ an inductive argument in order to successively reduce the degrees of the exponential sums occurring within the mean value

$$B_1(p^h)=p^{-2Kh}\sum_{1\leqslant\mathbf{a}\leqslant p^h}\prod_{j=1}^{w}|S_k(p^h,\boldsymbol{\Delta}_j)|^{2\tau_j}.$$

Observe that, as a result of our preparatory manipulations, the $r_\ell\times r_\ell$ coefficient matrics $D^{(l)}$ with $2\leqslant l\leqslant \ell_w$ are upper triangular. Thus, the only exponential sum within the above formula for $B_1(p^h)$ that depends on $\mathbf{a}_w^{[\ell_w]}$ is the one involving $\boldsymbol{\Delta}_w$. In order to save clutter, we temporarily drop the modulus p^h in our exponential sums $S_k(p^h,\boldsymbol{\Delta}_j)$. We may thus write

$$B_1(p^h)=p^{-2Kh}\sum_{\substack{1\leqslant\mathbf{a}_j^{[\ell_j]}\leqslant p^h\\(1\leqslant j\leqslant w-1)}}\left(\prod_{j=1}^{w-1}|S_k(\boldsymbol{\Delta}_j)|^{2\tau_j}\right)\sum_{1\leqslant\mathbf{a}_w^{[\ell_w]}\leqslant p^h}|S_k(\boldsymbol{\Delta}_w)|^{2\tau_w}.$$

The inner sum is of the shape considered in Lemma 5.1 with $l=\ell_w$. On writing $\mathbf{d}_j=(d_{j,j}^{(l)})_{2\leqslant l\leqslant\ell_j}$ $(1\leqslant j\leqslant w)$, we thus obtain the bound

$$B_1(p^h)\leqslant p^{-2Kh}\Gamma_{p^h}(\mathbf{d}_w)\sum_{\substack{1\leqslant\mathbf{a}_j^{[\ell_j]}\leqslant p^h\\(1\leqslant j\leqslant w-1)}}\left(\prod_{j=1}^{w-1}|S_k(\boldsymbol{\Delta}_j)|^{2\tau_j}\right)\sum_{1\leqslant\mathbf{a}_w^{[\ell_w]}\leqslant p^h}|S_{\ell_w}(\mathbf{a}_w^{[\ell_w]})|^{2\tau_w}.$$

Now suppose that for some index j with $1\leqslant j\leqslant w-1$ we have the bound

$$B_1(p^h)\leqslant p^{-2Kh}\Upsilon_j\prod_{i=j+1}^{w}\Gamma_{p^h}(\mathbf{d}_i)\sum_{1\leqslant\mathbf{a}_i^{[\ell_i]}\leqslant p^h}|S_{\ell_i}(\mathbf{a}_i^{[\ell_i]})|^{2\tau_i},\tag{5.8}$$

where

$$\Upsilon_j = \sum_{\substack{1\leqslant \mathbf{a}_i^{[\ell_i]}\leqslant p^h\\ (1\leqslant i\leqslant j)}} \left(\prod_{i=1}^{j} |S_k(\boldsymbol{\Delta}_i)|^{2\tau_i}\right).$$

Again, since we may assume all coefficient matrices $D^{(l)}$ to be in upper row echelon form, the only exponential sum within the mean value defining Υ_j that depends on the vector $\mathbf{a}_j^{[\ell_j]}$ is the one involving $\boldsymbol{\Delta}_j$. Thus, as in the case $j=w$ considered above, we may isolate the exponential sum indexed by j and apply Lemma 5.1. As a result, we find that

$$\begin{aligned}\Upsilon_j &= \sum_{\substack{1\leqslant \mathbf{a}_i^{[\ell_i]}\leqslant p^h\\ (1\leqslant i\leqslant j-1)}} \left(\prod_{i=1}^{j-1} |S_k(\boldsymbol{\Delta}_i)|^{2\tau_i}\right) \sum_{1\leqslant \mathbf{a}_j^{[\ell_j]}\leqslant p^h} |S_k(\boldsymbol{\Delta}_j)|^{2\tau_j}\\ &\leqslant \Gamma_{p^h}(\mathbf{d}_j) \sum_{1\leqslant \mathbf{a}_j^{[\ell_j]}\leqslant p^h} |S_{\ell_j}(\mathbf{a}_j^{[\ell_j]})|^{2\tau_j} \sum_{\substack{1\leqslant \mathbf{a}_i^{[\ell_i]}\leqslant p^h\\ (1\leqslant i\leqslant j-1)}} \prod_{i=1}^{j-1} |S_k(\boldsymbol{\Delta}_i)|^{2\tau_i}.\end{aligned}$$

Inserting this bound into (5.8) reproduces (5.8) with j replaced by $j-1$. We may clearly iterate, and after w steps we find that

$$B_1(p^h) \leqslant p^{-2Kh} \prod_{j=1}^{w} \Gamma_{p^h}(\mathbf{d}_j) \sum_{1\leqslant \mathbf{a}_j^{[\ell_j]}\leqslant p^h} |S_{\ell_j}(\mathbf{a}_j^{[\ell_j]})|^{2\tau_j}.$$

Clearly, the vectors $\mathbf{a}_j^{[\ell_j]}$ with $1\leqslant j\leqslant w$ together list the coordinates of $\mathbf{a}$. Since $B_1(q)$ is multiplicative, the assertion of the lemma is now confirmed upon taking $\mathcal{R}(q)$ to be the multiplicative function defined via the formula

$$\mathcal{R}(p^h) = \prod_{j=1}^{w} \Gamma_{p^h}(\mathbf{d}_j).$$

In particular, we note that $\mathcal{R}(p^h)$ depends at most on the coefficient matrices $D^{(l)}$, and one has $\mathcal{R}(p^h)=1$ whenever $p\notin\Omega(D)$. □

6. Conclusion of the major arcs analysis

With Lemma 5.2 we are now equipped to engage with our goal of showing that the singular series $\mathfrak{S} = \lim_{Q\to\infty} \mathfrak{S}(Q)$ converges absolutely. In this context, for each prime number p we define the p-adic factor

$$\chi_p = \sum_{h=0}^{\infty} A(p^h). \tag{6.1}$$

Lemma 6.1. *Suppose that the coefficient matrices $C^{(l)}$ associated with the system* (1.10) *are highly non-singular, and that $r_l \geqslant r_{l+1}$ for $2 \leqslant l \leqslant k-1$. Also, assume that $s \geqslant 2K+1$. Then the p-adic densities χ_p exist, the singular series $\mathfrak{S}$ is absolutely convergent, and $\mathfrak{S} = \prod_p \chi_p$. In particular, one has $\mathfrak{S}(Q) = \mathfrak{S} + O(Q^{-\delta})$ for some $\delta > 0$. Moreover, if the system* (1.10) *has a non-singular p-adic solution for all primes p, then $\mathfrak{S} \gg 1$.*

Proof. On recalling (4.8) and (5.1), we see that $\mathfrak{S}(Q) = \sum_{1 \leqslant q \leqslant Q} A(q)$, and so the estimation of the quantity $A(q)$ is our central focus. The multiplicativity of $A(q)$ allows us to restrict our attention to the cases where q is a prime power. Set $\chi_p(H) = \sum_{h=0}^{H} A(p^h)$ and $L_p(Q) = \lfloor \log Q / \log p \rfloor$. If the product

$$\prod_{p \leqslant Q} \chi_p(L_p(Q))$$

converges absolutely as $Q \to \infty$, then so does $\mathfrak{S}(Q)$ with the same limit. In such circumstances, one has $\mathfrak{S} = \prod_p \chi_p$. It is therefore sufficient to show that for all primes p the limit

$$\chi_p = \lim_{H \to \infty} \chi_p(H)$$

exists, and moreover that there exists a positive number δ having the property that $\chi_p = 1 + O(p^{-1-\delta})$ for all but at most a finite set of primes p.

On recalling (5.2), we find from (4.4) that

$$\begin{aligned} A(p^h) &\ll (p^h)^{2K-s} \max_{\substack{1 \leqslant \mathbf{a} \leqslant p^h \\ (\mathbf{a},p)=1}} \left(\prod_{j=1}^{w} |S_k(p^h, \boldsymbol{\Delta}_j)| \right)^{(s-2K)/w} A_1(p^h) \\ &\ll \max_{\substack{1 \leqslant \mathbf{a} \leqslant p^h \\ (\mathbf{a},p)=1}} \left(\prod_{j=1}^{w} (p^h)^{-1/k+\varepsilon} (p^h, \boldsymbol{\Delta}_j)^{1/k} \right)^{(s-2K)/w} A_1(p^h). \end{aligned}$$

The invertibility of the coordinate transform (5.4) implies that when $(\mathbf{a}, p^h) = 1$, then there is at least one index j with $1 \leqslant j \leqslant w$ such that $(p^h, \boldsymbol{\Delta}_j) \ll 1$, with an implied constant depending at most on the coefficient matrices $C^{(l)}$. Since $s - 2K \geqslant 1$ and ε may be taken arbitrarily small, we deduce that there is a positive number c_1, depending at most on the coefficient matrices $C^{(l)}$, having the property that

$$A(p^h) \leqslant c_1 p^{-h/(2kw)} A_1(p^h). \tag{6.2}$$

We now wish to apply Lemma 5.2. To this end, we first recall (5.3) and observe that a summation by parts yields the relation

$$\sum_{h=0}^{L} p^{-\frac{h}{2kw}} A_1(p^h) = p^{-\frac{L}{2kw}} B_1(p^L) + \sum_{h=0}^{L-1} \left(p^{-\frac{h}{2kw}} - p^{-\frac{h+1}{2kw}} \right) B_1(p^h). \tag{6.3}$$

Since all coefficients on the right hand side are positive, and also both $B_1(p^h)$ and $B_1^*(p^h)$ are non-negative for all non-negative integers h, it follows from Lemma 5.2 that we may majorise the right hand side of (6.3) by replacing $B_1(p^h)$ with $\mathcal{R}(p^h) B_1^*(p^h)$ for

$0 \leqslant h \leqslant L$. Set $\mathcal{R}_p = \max_{h\geqslant 0} \mathcal{R}(p^h)$, noting that this maximum exists as $\mathcal{R}(p^h)$ is an integer which is bounded uniformly for all non-negative integers h. Also, in analogy to the definition of $B_1^*(p^h)$, we put

$$A_1^*(p^h) = p^{-2Kh} \sum_{\substack{1\leqslant \mathbf{a}\leqslant p^h\\ (\mathbf{a},p)=1}} \prod_{j=1}^{w} |S_{\ell_j}(p^h, \mathbf{a}_j^{[\ell_j]})|^{2\tau_j}.$$

Thus, another summation by parts shows that the right hand side of (6.3) is no larger than

$$\mathcal{R}_p \left(p^{-\frac{L}{2kw}} B_1^*(p^L) + \sum_{h=0}^{L-1} \left(p^{-\frac{h}{2kw}} - p^{-\frac{h+1}{2kw}} \right) B_1^*(p^h) \right) = \mathcal{R}_p \sum_{h=0}^{L} p^{-\frac{h}{2kw}} A_1^*(p^h).$$

We have therefore established the bound

$$\sum_{h=0}^{L} p^{-\frac{h}{2kw}} A_1(p^h) \leqslant \mathcal{R}_p \sum_{h=0}^{L} p^{-\frac{h}{2kw}} A_1^*(p^h). \tag{6.4}$$

Since $2\tau_j = \ell_j(\ell_j+1) - 2 \geqslant \ell_j^2$ for all j, we can infer further from (4.4) that there exists a positive number c_2, depending at most on ε, such that

$$A_1^*(p^h) \leqslant c_2 p^{h\varepsilon} \sum_{\substack{1\leqslant \mathbf{a}\leqslant p^h\\ (\mathbf{a},p)=1}} \prod_{j=1}^{w} \left(p^{-h/\ell_j} (p^h, \mathbf{a}_j^{[\ell_j]})^{1/\ell_j} \right)^{2\tau_j} \leqslant c_2 p^{h\varepsilon} \sum_{\substack{1\leqslant \mathbf{a}\leqslant p^h\\ (\mathbf{a},p)=1}} \prod_{j=1}^{w} p^{-h\ell_j} (p^h, \mathbf{a}_j^{[\ell_j]})^{\ell_j}.$$

For a fixed vector $\mathbf{e} \in \mathbb{Z}_{\geqslant 0}^w$ denote by $\Xi(p^h, \mathbf{e})$ the number of vectors $\mathbf{a} \in \mathbb{Z}^r$ satisfying $1 \leqslant \mathbf{a} \leqslant p^h$ and $(p^h, \mathbf{a}_j^{[\ell_j]}) = p^{e_j}$ for $1 \leqslant j \leqslant w$. Then one has

$$A_1^*(p^h) \leqslant c_2 p^{h\varepsilon} \sum_{\mathbf{e}} \Xi(p^h, \mathbf{e}) \prod_{j=1}^{w} p^{(e_j - h)\ell_j},$$

where the sum is over all vectors $\mathbf{e} \in \mathbb{Z}^w$ satisfying $0 \leqslant e_j \leqslant h$ and having the property that $e_j = 0$ for at least one index j. For any fixed j, the number of choices for $\mathbf{a}_j^{[\ell_j]} \in \mathbb{Z}^{\ell_j - 1}$ having $1 \leqslant \mathbf{a}_j^{[\ell_j]} \leqslant p^h$ and $(p^h, \mathbf{a}_j^{[\ell_j]}) = p^{e_j}$ is at most $p^{(h-e_j)(\ell_j-1)}$. It follows that

$$\Xi(p^h, \mathbf{e}) \leqslant \prod_{j=1}^{w} p^{(h-e_j)(\ell_j-1)},$$

and hence

$$A_1^*(p^h) \leqslant c_2 p^{h\varepsilon} \sum_{\substack{0\leqslant \mathbf{e}\leqslant h\\ e_1\cdots e_w=0}} \prod_{j=1}^{w} p^{e_j - h} \leqslant c_2 w p^{h\varepsilon - h} \left(\sum_{e=0}^{h} p^{e-h} \right)^{w-1} \leqslant c_2 w 2^w (p^h)^{-1+\varepsilon}. \tag{6.5}$$

On recalling (6.2), (6.4) and (6.5) we find that

$$\begin{aligned}\sum_{h=1}^\infty |A(p^h)| &\leqslant c_1\Bigl(-1+\sum_{h=0}^\infty p^{-h/(2kw)}A_1(p^h)\Bigr)\\ &\leqslant c_1\Bigl(-1+\mathcal{R}_p\sum_{h=0}^\infty p^{-h/(2kw)}A_1^*(p^h)\Bigr)\\ &\leqslant c_1(\mathcal{R}_p-1)+w2^wc_1c_2\mathcal{R}_p\sum_{h=1}^\infty p^{-h(1+1/(2kw)-\varepsilon)}\ll 1.\end{aligned}$$

It follows that the p-adic density χ_p defined in (6.1) exists. In particular, whenever $\mathcal{R}_p=1$ we have

$$\sum_{h=1}^\infty |A(p^h)|\leqslant c_3p^{-1-1/(3kw)} \tag{6.6}$$

for some positive number c_3 depending at most on the coefficient matrices $C^{(l)}$. On recalling the conclusion of Lemma 5.2, one sees that $\mathcal{R}_p=1$ for all primes p with $p\notin\Omega(D)$, and thus

$$\prod_p\Bigl(\sum_{h=0}^\infty |A(p^h)|\Bigr)\ll\prod_p\bigl(1+p^{-1-1/(3kw)}\bigr)^{c_3}<\zeta(1+1/(3kw))^{c_3}.$$

Hence, the singular series $\mathfrak{S}$ converges absolutely and one has $\mathfrak{S}=\prod_p\chi_p$.

Furthermore, a standard argument yields

$$\chi_p=\lim_{h\to\infty}p^{-h(s-r)}M(p^h),$$

where $M(q)$ denotes the number of solutions $\mathbf{x}\in(\mathbb{Z}/q\mathbb{Z})^s$ of the congruences

$$c_{j,1}^{(l)}x_1^l+\ldots+c_{j,s}^{(l)}x_s^l\equiv 0\pmod q\qquad(1\leqslant j\leqslant r_l,\ 2\leqslant l\leqslant k),$$

corresponding to the equations (1.10). Using again the observation that $\mathcal{R}_p=1$ for all sufficiently large primes p, we discern from (6.6) that there exists an integer p_0 with the property that

$$1/2\leqslant\prod_{\substack{p>p_0\\ p\text{ prime}}}\chi_p\leqslant 3/2.$$

For the remaining finite set of primes, a standard application of Hensel's lemma shows that $\chi_p>0$ whenever the system (1.10) possesses a non-singular solution in $\mathbb{Q}_p$. We thus conclude that under the hypotheses of the lemma we have $\mathfrak{S}\gg 1$ as claimed. $\square$

We next demonstrate the existence of the limit

$$\chi_\infty=\lim_{Q\to\infty}\mathfrak{J}_1(Q).$$

With this goal in mind, when W is a positive real number, we introduce the auxiliary mean value

$$\mathfrak{J}_1^*(W)=\int_{[-W,W]^r}\prod_{j=1}^s |v_k(\boldsymbol{\vartheta}_j;1)|\,\mathrm{d}\boldsymbol{\beta}.$$

Lemma 6.2. *Under the hypotheses of Theorem 4.1, there is a positive number δ for which one has $\mathfrak{J}_1^*(2Q)-\mathfrak{J}_1^*(Q)\ll Q^{-\delta}$, and hence the limit χ_∞ exists. In particular, one has*

$$\mathfrak{J}_1(Q)=\chi_\infty+O(Q^{-\delta}).$$

Furthermore, if the system (1.10) *has a non-singular solution inside the real unit cube* $(-1,1)^s$, *then the singular integral χ_∞ is positive.*

Proof. The first part of the proof is inspired by a singular series argument of Heath-Brown and Skorobogatov (see [9, pages 173 and 174]). Recall that

$$s_0=t_0u_0+\ldots+t_nu_n,$$

where the integers t_j are defined by means of (3.15). Thus, the hypotheses of Theorem 4.1 imply that $s\geqslant 2s_0+1\geqslant 2K+1$. Let $\mathcal{J}$ denote the set of s_0-element subsets $\{j_1,\ldots,j_{s_0}\}$ of $\{1,\ldots,s\}$. When $J\in\mathcal{J}$, define

$$\mathscr{S}_J(Q)=\sum_{1\leqslant q\leqslant Q}\sum_{\substack{1\leqslant \mathbf{a}\leqslant q\\ (q,\mathbf{a})=1}} q^{-2s_0}\prod_{j\in J}|S_k(q,\boldsymbol{\Lambda}_j)|^2 \tag{6.7}$$

and

$$\mathscr{J}_J(Q)=\int_{[-Q,Q]^r}\prod_{j\in J}|v_k(\boldsymbol{\vartheta}_j;1)|^2\,\mathrm{d}\boldsymbol{\beta}.$$

Set $Y=Q^{6r}$, and define the major arcs $\mathfrak{M}_Y(Q)$ via (4.3). By making the necessary modifications to our initial analysis of the major arcs, we see from (4.9) that for any $J\in\mathcal{J}$ one has

$$\int_{\mathfrak{M}_Y(Q)}\prod_{j\in J}|g_k(\boldsymbol{\gamma}_j;Y)|^2\,\mathrm{d}\boldsymbol{\alpha}=Y^{2s_0-K}\mathscr{S}_J(Q)\mathscr{J}_J(Q)+O(Y^{2s_0-K}Q^{-1}). \tag{6.8}$$

Note that we have $S_k(1,\mathbf{1})=1$ for the term corresponding to $q=1$ in (6.7). Since all other summands are non-negative, it follows that for any $Q\geqslant 1$ and any $J\in\mathcal{J}$, one has

$$\mathscr{S}_J(Q)\geqslant 1. \tag{6.9}$$

On the other hand, for $Y=Q^{6r}$ the major arcs $\mathfrak{M}_Y(Q)$ are disjoint, and we conclude from Theorem 3.3 that under the hypotheses of Theorem 4.1 we have

$$\int_{\mathfrak{M}_Y(Q)}\prod_{j\in J}|g_k(\boldsymbol{\gamma}_j;Y)|^2\,\mathrm{d}\boldsymbol{\alpha}\ll I_{s_0,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(Y)\ll Y^{2s_0-K+\varepsilon}.$$

In combination with (6.8) and (6.9) it follows that

$$\max_{J\in\mathcal{J}}\mathscr{J}_J(Q)\ll Y^{\varepsilon}. \tag{6.10}$$

Since Y is a power of Q, we discern from (4.5) and (6.10) via (2.3) that for any $Q>1$ we have

$$|\mathfrak{J}_1^*(2Q)-\mathfrak{J}_1^*(Q)| \ll \left(\sup_{|\boldsymbol{\beta}|>Q}\prod_{j=1}^s |v_k(\boldsymbol{\vartheta}_j;1)|\right)^{1-2s_0/s}\int_{[-2Q,2Q]^r}\prod_{j=1}^s |v_k(\boldsymbol{\vartheta}_j;1)|^{2s_0/s}\,\mathrm{d}\boldsymbol{\beta}$$
$$\ll Q^{-1/(ks)}\max_{J\in\mathcal{J}}\mathscr{J}_J(2Q)\ll Q^{-1/(ks)+\varepsilon}.$$

Here, we exploited the fact that, since the coefficient matrices $C^{(l)}$ are highly non-singular, the condition $|\boldsymbol{\beta}|>Q$ implies that $|\boldsymbol{\vartheta}_j|\gg Q$ for some index j with $1\leqslant j\leqslant s$. This implies the first statement of the lemma. In particular, the singular integral χ_∞ converges absolutely.

In order to establish the second claim, we follow an argument of Schmidt [10]. When $T\geqslant 1$, define

$$w_T(y)=\begin{cases}T(1-T|y|), & \text{when } |y|\leqslant T^{-1},\\ 0, & \text{otherwise,}\end{cases}$$

and recall that

$$w_T(y)=\int_{-\infty}^{\infty} e(\beta y)\left(\frac{\sin(\pi\beta/T)}{\pi\beta/T}\right)^2\mathrm{d}\beta, \tag{6.11}$$

where the integral converges absolutely. Set

$$\Phi_j^{(l)}(\mathbf{x})=c_{j,1}^{(l)}x_1^l+\ldots+c_{j,s}^{(l)}x_s^l \qquad (1\leqslant j\leqslant r_l,\ 2\leqslant l\leqslant k),$$

and put

$$W_T=\int_{[-1,1]^s}\prod_{l=2}^k\prod_{j=1}^{r_l} w_T(\Phi_j^{(l)}(\mathbf{z}))\,\mathrm{d}\mathbf{z}.$$

We adapt the argument of §11 in Schmidt's work [10] to show that $W_T\to\chi_\infty$ as $T\to\infty$.

Set

$$\psi_T(\boldsymbol{\beta})=\prod_{l=2}^k\prod_{j=1}^{r_l}\left(\frac{\sin(\pi\beta_j^{(l)}/T)}{\pi\beta_j^{(l)}/T}\right)^2.$$

Then in light of (6.11) a change of the order of integration shows that

$$W_T=\int_{\mathbb{R}^r}\left(\prod_{i=1}^s v_k(\boldsymbol{\vartheta}_i;1)\right)\psi_T(\boldsymbol{\beta})\,\mathrm{d}\boldsymbol{\beta},$$

and hence

$$W_T-\chi_\infty=\int_{\mathbb{R}^r}\left(\prod_{i=1}^s v_k(\boldsymbol{\vartheta}_i;1)\right)(\psi_T(\boldsymbol{\beta})-1)\,\mathrm{d}\boldsymbol{\beta}. \tag{6.12}$$

In order to analyse the integral on the right hand side of (6.12), it is convenient to consider two domains separately. Write $U_1 = [-\sqrt{T}, \sqrt{T}]^r$, and set $U_2 = \mathbb{R}^r \setminus U_1$. From the power series expansion of ψ_T we find that

$$0 \leqslant 1 - \psi_T(\boldsymbol{\beta}) \ll \min\left\{1, \sum_{l=2}^{k} \sum_{j=1}^{r_l} (|\beta_j^{(l)}|/T)^2\right\},$$

whence we discern that the domain U_1 contributes at most

$$\sup_{\boldsymbol{\beta}\in U_1} |1 - \psi_T(\boldsymbol{\beta})| \int_{\mathbb{R}^r} \prod_{i=1}^{s} |v_k(\boldsymbol{\vartheta}_i; 1)| \,\mathrm{d}\boldsymbol{\beta} \ll T^{-1}.$$

Note that in the last step we used our previous insight that the singular integral converges absolutely. Meanwhile, the contribution from U_2 is bounded above by

$$\sum_{i=1}^{\infty} |\mathfrak{J}_1^*(2^i\sqrt{T}) - \mathfrak{J}_1^*(2^{i-1}\sqrt{T})| \ll \sum_{i=1}^{\infty} (2^i\sqrt{T})^{-\delta} \ll T^{-\delta/2},$$

for some positive number δ with $\delta < 1$, where again we took advantage of our earlier findings. Thus we infer from (6.12) that

$$|W_T - \chi_\infty| \ll T^{-\delta/2} \tag{6.13}$$

for all $T \geqslant 1$, and hence W_T does indeed converge to χ_∞, as claimed.

Suppose now that the system (1.10) has a non-singular solution inside $(-1,1)^s$. Then it follows from the implicit function theorem that the real manifold described by the equations in (1.10) has positive $(s-r)$-dimensional volume inside $(-1,1)^s$. In such circumstances, Lemma 2 of Schmidt [10] shows that $W_T \gg 1$ uniformly in T. We therefore deduce from (6.13) that χ_∞ is indeed positive, confirming the second claim of the lemma. □

Upon combining (4.12) with Lemmata 6.1 and 6.2, we conclude that

$$\begin{aligned} N_{s,\mathbf{k}}^{\mathbf{v},\mathbf{u}}(X) &= X^{s-K}\left(\mathfrak{S} + O(Q^{-\delta})\right)\left(\chi_\infty + O(Q^{-\delta})\right) + O(X^{s-K-\omega}) \\ &= (\mathcal{C} + o(1))X^{s-K}, \end{aligned}$$

where $\mathcal{C} = \chi_\infty \prod_p \chi_p$. Moreover, the constant $\mathcal{C}$ is positive whenever the system (1.10) possesses non-singular solutions in all local fields. This confirms the asymptotic formula (1.11), and completes our proof of Theorem 4.1.

References

[1] 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** (2016), no. 2, 633–682.

[2] J. Brandes, *The Hasse principle for systems of quadratic and cubic diagonal equations*, Q. J. Math. **68** (2017), no. 3, 831–850.

[3] J. Brandes and S. T. Parsell, *Simultaneous additive equations: repeated and differing degrees*, Canad. J. Math. **69** (2017), no. 2, 258–283.

[4] J. Brandes and T. D. Wooley, *Vinogradov systems with a slice off*, Mathematika **63** (2017), no. 3, 797–817.

[5] J. Brüdern and T. D. Wooley, *The Hasse principle for pairs of diagonal cubic forms*, Ann. of Math. (2) **166** (2007), no. 3, 865–895.
[6] J. Brüdern and T. D. Wooley, *The Hasse principle for systems of diagonal cubic forms*, Math. Ann. **364** (2016), no. 3-4, 1255–1274.
[7] R. J. Cook, *Simultaneous quadratic equations*, J. London Math. Soc. (2) **4** (1971), 319–326.
[8] R. J. Cook, *Simultaneous quadratic equations II*, Acta Arith. **25** (1973/74), 1–5.
[9] R. Heath-Brown and A. Skorobogatov, *Rational solutions of certain equations involving norms*, Acta Math. **189** (2002), no. 2, 161–177.
[10] W. M. Schmidt, *Simultaneous rational zeros of quadratic forms*, in: Seminar on Number Theory (Paris 1980-81), Progr. Math. 22, Birkhäuser, Boston, MA, 1982, 281–307.
[11] R. C. Vaughan, *The Hardy–Littlewood Method*. Second ed., Cambridge University Press, Cambridge, 1997.
[12] T. D. Wooley, *Rational solutions of pairs of diagonal equations, one cubic and one quadratic*, Proc. London Math. Soc. (3) **110** (2015), no. 2, 325–356.
[13] T. D. Wooley, *The cubic case of the main conjecture in Vinogradov's mean value theorem*, Adv. Math. **294** (2016), 532–561.
[14] T. D. Wooley, *Nested efficient congruencing and relatives of Vinogradov's mean value theorem*, Proc. London Math. Soc. **118** (2019), no. 4, 942–1016.

Mathematical Sciences, University of Gothenburg and Chalmers Institute of Technology, 412 96 Göteborg, Sweden

E-mail address: brjulia@chalmers.se

Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA

E-mail address: twooley@purdue.edu