CORRELATION ESTIMATES FOR SUMS OF THREE CUBES

JÖRG BRÜDERN AND TREVOR D. WOOLEY

Abstract. We establish estimates for linear correlation sums involving sums of three positive integral cubes. Under appropriate conditions, the underlying methods permit us to establish the solubility of systems of homogeneous linear equations in sums of three positive cubes whenever these systems have more than twice as many variables as equations.

1. Introduction

We shall be concerned in this memoir with the number $\rho(n)$ of ways the natural number n can be written as the sum of three positive integral cubes. Our principal goal is to provide upper bounds for linear correlation sums involving $\rho(n)$ and certain of its relatives. As an application of the underlying methods, we consider the solubility of systems of homogeneous linear equations in sums of three positive integral cubes. Provided that the system is in general position, and it has a solution in positive integers, we are able to show that it is soluble in sums of three positive cubes whenever the number of variables exceeds twice the number of equations.

Some notation is required before we may introduce the family of higher correlation sums that are central to our focus. Let s and r be natural numbers with $s \geqslant r$, and consider an $r \times s$ integral matrix $A = (a_{ij})$. We associate with A the collection of linear forms

$$\Lambda_j(\boldsymbol{\alpha}) = \sum_{i=1}^{r} a_{ij}\alpha_i \quad (1 \leqslant j \leqslant s), \tag{1.1}$$

and its positive cone

$$\mathcal{P} = \{\boldsymbol{\alpha} \in \mathbb{R}^r : \alpha_i > 0 \ (1 \leqslant i \leqslant r) \text{ and } \Lambda_j(\boldsymbol{\alpha}) > 0 \ (1 \leqslant j \leqslant s)\}.$$

Note that $\mathcal{P}$ is open, and hence its truncation $\mathcal{P}(N) = \mathcal{P} \cap [1, N]^r$ has measure $\gg N^r$ whenever $\mathcal{P}$ is non-empty. Given an s-tuple $\mathbf{h}$ of non-negative integers, we may now define the sum $\Xi_s(N) = \Xi_s(N; A; \mathbf{h})$ by putting

$$\Xi_s(N; A; \mathbf{h}) = \sum_{\mathbf{n} \in \mathcal{P}(N)} \rho(\Lambda_1(\mathbf{n}) + h_1) \cdots \rho(\Lambda_s(\mathbf{n}) + h_s). \tag{1.2}$$

We refer to the coefficient matrix A as being *highly non-singular* if all collections of at most r of its s columns are linearly independent.

2010 *Mathematics Subject Classification.* 11D72, 11P55, 11E76.

Key words and phrases. Sums of cubes, Hardy-Littlewood method.

The authors thank the Hausdorff Institute in Bonn, the Heilbronn Institute in Bristol, and CIRM in Luminy, for excellent working conditions conducive to writing this paper.

Theorem 1.1. *Let $A \in \mathbb{Z}^{r\times 2r}$ be highly non-singular, and let $h_i \in \mathbb{N}\cup\{0\}$ $(1 \leqslant i \leqslant 2r)$. Then $\Xi_{2r}(N;A;\mathbf{h}) \ll N^{r+1/6+\varepsilon}$, where the constant implict in Vinogradov's notation depends at most on A and $\varepsilon > 0$.*

Classical approaches to the simplest correlation sum proceed via Cauchy's inequality. Thus, by utilising Hua's lemma (see [17, Lemma 2.5]), one obtains

$$\sum_{n\leqslant N}\rho(n)\rho(n+h) \ll \sum_{n\leqslant N+h}\rho(n)^2 \ll N^{7/6+\varepsilon}. \tag{1.3}$$

This traditional argument is easily generalised to handle the sum $\Xi_{2r}(N)$. Writing $m_j = \Lambda_j(\mathbf{n}) + h_j$ for the sake of brevity, Cauchy's inequality yields

$$\Xi_{2r}(N) \leqslant \prod_{j\in\{0,1\}}\biggl(\sum_{\mathbf{n}\in\mathcal{P}(N)}\rho(m_{jr+1})^2\cdots\rho(m_{jr+r})^2\biggr)^{1/2}.$$

Since $\Lambda_1,\Lambda_2,\ldots,\Lambda_r$ are linearly independent, one may sum over the values $m_1,m_2,\ldots,m_r$ as if these were independent variables. Thus, by symmetry, it follows as a consequence of the second inequality of (1.3) that there is a number $C = C(A) \geqslant 1$ such that

$$\Xi_{2r}(N) \leqslant \biggl(\sum_{n\leqslant CN}\rho(n)^2\biggr)^r \ll N^{7r/6+\varepsilon}. \tag{1.4}$$

The bound (1.4) is certainly part of the folklore in the area, and constitutes the state of the art hitherto. It is widely believed that the upper bound $N^{7/6+\varepsilon}$ in (1.3) may be replaced by N, and indeed the slightly weaker estimate $N^{1+\varepsilon}$ has been established by Hooley [11] and Heath-Brown [10] based on speculative hypotheses concerning the distribution of the zeros of certain Hasse-Weil L-functions. Accepting one or other of these estimates as a working hypothesis, one finds that $\Xi_{2r}(N) \ll N^{r+\varepsilon}$, or even $\Xi_{2r}(N) \ll N^r$. For certain coefficient matrices A, readers will have little difficulty in convincing themselves that the lower bound $\Xi_{2r}(N) \gg N^r$ is to be expected. Although the bound on $\Xi_{2r}(N)$ presented in Theorem 1.1 does not improve on the classical estimate (1.4) when $r = 1$, for all larger values of r it is substantially sharper.

For applications to problems of Waring's type, mollified versions of $\rho(n)$ have been utilised since the invention by Hardy and Littlewood [9] of diminishing ranges. Most modern innovations within this circle of ideas involve the use of sets of smooth numbers having positive density. Thus, given $\eta > 0$, let $\rho_\eta(n)$ denote the number of integral solutions of the equation $n = x^3+y^3+z^3$, subject to the condition that none of the prime divisors of yz exceed $n^{\eta/3}$. Then it follows from [19, 20] that for each $\varepsilon > 0$, there is a positive number η such that

$$\sum_{1\leqslant n\leqslant N}\rho_\eta(n)^2 \ll N^{1+\xi+\varepsilon}, \tag{1.5}$$

where $\xi = (\sqrt{2833}-43)/123 < 1/12$. Define $\Xi_{s,\eta}(N;A;\mathbf{h})$ as in (1.2), but with ρ_η in place of ρ throughout.

Theorem 1.2. *Let $A \in \mathbb{Z}^{r\times 2r}$ be highly non-singular, and let $h_i \in \mathbb{N} \cup \{0\}$ $(1 \leqslant i \leqslant 2r)$. Then for each $\varepsilon > 0$, there is a number $\eta > 0$ such that*

$$\Xi_{2r,\eta}(N; A; \mathbf{h}) \ll N^{r+\xi+\varepsilon}.$$

The constant implict in Vinogradov's notation depends at most on A, ε and η.

We turn now to systems of linear equations in sums of three cubes. Let $C \in \mathbb{Z}^{r\times s}$ be highly non-singular, and suppose that the system

$$\sum_{j=1}^{s} c_{ij} n_j = 0 \quad (1 \leqslant i \leqslant r) \tag{1.6}$$

has a solution in positive integers $n_1, \ldots, n_s$. Denote by $\Upsilon(N)$ the number of solutions of the system (1.6) with $n_j \leqslant N$ in which n_j is a sum of three positive integral cubes. We emphasise that $\Upsilon(N)$ counts solutions without weighting them for the number of representations as the sum of three cubes.

Theorem 1.3. *Let $C \in \mathbb{Z}^{r\times s}$ be highly non-singular, and suppose that* (1.6) *has a solution $\mathbf{n} \in (0, \infty)^s$. Then whenever $s > 2r$ and $\varepsilon > 0$, one has*

$$\Upsilon(N) \gg N^{s(1-2\xi)-r-\varepsilon}.$$

Were sums of three positive integral cubes to have positive density in the natural numbers, then one imagines that a suitable enhancement of the methods of Gowers [8] ought to deliver the stronger conclusion $\Upsilon(N) \gg N^{s-r}$ for $s \geqslant r + 2$. However, there seems to be no prospect of any such density result at present, and so one is forced to contemplate the possibility that the number of positive integers $n \leqslant N$, representable as the sum of three positive integral cubes, may be as small as $N^{1-\xi}$. In such circumstances, even the lower bound $\Upsilon(N) \geqslant 1$ is highly non-trivial. Indeed, in cases where s is close to $2r+1$, such a conclusion is established for the first time within this paper. When sums of three cubes are replaced by sums of two squares, on the other hand, the value set comes very close to achieving positive density, and the methods of Gowers are in play. In this setting, the work of Matthiesen [13, 14] comes within a factor N^{ε} of achieving the natural analogue of the above lower bound.

Subject to appropriate additional hypotheses, a conclusion similar to that of Theorem 1.3 may be obtained for the analogue of $\Upsilon(N)$ in which (1.6) is replaced by an inhomogeneous system of linear equations. Note also that Balog and Brüdern [1] consider systems of linear equations in sums of three cubes of special type. In the case of a single equation, their work more efficiently removes the multiplicity inherent in $\rho_\eta(n)$, and establishes a superior bound in this case for $\Upsilon(N)$.

The conclusions of this paper depend on a new mean value estimate that is of independent interest. In §2 we examine systems of equations in which the coefficient matrices are of linked block type, and establish an auxiliary bound for their number of solutions. This prepares the way for the proof of the central estimate, Theorem 3.4, in §3, accomplished by a novel *complification* argument in which mean values are bounded by blowing up the number of equations so as to apply the powerful estimates of the previous section. We then establish the correlation estimates of Theorems 1.1 and 1.2 in §4, and finish in §5 by applying the Hardy-Littlewood method to prove Theorem 1.3.

Our basic parameter is P, a sufficiently large positive number. In this paper, implicit constants in Vinogradov's notation $\ll$ and $\gg$ may depend on s, r and ε, as well as ambient coefficients. Whenever ε appears in a statement, either implicitly or explicitly, we assert that the statement holds for each $\varepsilon > 0$. We employ the convention that whenever $G : [0,1)^k \to \mathbb{C}$ is integrable, then

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

Here and elsewhere, we use vector notation in the natural way. Finally, we write $e(z)$ for $e^{2\pi iz}$.

The authors are very grateful to a referee for identifying obscurities in the original version of this paper. In the current version, the treatment of the central mean value estimate in §2, though somewhat longer, is both more explicit and considerably simpler in detail.

2. Auxiliary equations

In this section we establish near-optimal mean value estimates for certain products of cubic Weyl sums. The formal coefficient matrices associated with these exponential sums have repeated columns, with multiplicities 2 and 4, and so would appear to be rather special. However, it transpires that this structure enables us to accommodate systems of cubic equations quite far from being in general position, and thus our principal conclusions are more flexible than the corresponding estimates of our earlier works [3, 6].

We begin by describing the matrices important in our arguments. For $0 \leqslant l \leqslant n$, consider natural numbers r_l, s_l and $r_l \times s_l$ matrices C_l having non-zero columns. Let $\mathrm{diag}(C_0, C_1, \ldots, C_n)$ be the conventional diagonal block matrix with the upper left hand corner of C_l sited at (i_l, j_l). For $0 \leqslant l \leqslant n$, append a row to the bottom of the matrix C_l, giving an $(r_l+1) \times s_l$ matrix C_l'. Next, consider the matrix $C^\dagger$ obtained from $\mathrm{diag}(C_0, C_1, \ldots, C_n)$ by replacing C_l by C_l' for $0 \leqslant l \leqslant n$, with the upper left hand corner of C_l' still sited at (i_l, j_l). We refer to this new matrix $C^\dagger$ as being a *linked block matrix*. It has additional entries by comparison to $\mathrm{diag}(C_0, C_1, \ldots, C_n)$, with the property that adjacent blocks are glued together by a shared row sited at index i_l, for $1 \leqslant l \leqslant n$.

We next describe the special linked block matrices relevant to our discussion. Let I_k denote the $k \times k$ identity matrix, and write $\mathbf{0}$ for the zero row vector with k components. We introduce the block matrices

$$I_k^* = \begin{pmatrix} I_k \\ \mathbf{0} \end{pmatrix} \quad \text{and} \quad I_k^+ = \begin{pmatrix} \mathbf{0} \\ I_k \\ \mathbf{0} \end{pmatrix}.$$

When $n \geqslant 0$, $r \geqslant t \geqslant 2$ and $\omega \in \{0, 1\}$, we consider fixed positive integers λ_l, and matrices M_l of format

$$\begin{cases} t \times (t+\omega), & \text{when } l = 0, \\ r \times r, & \text{when } 1 \leqslant l \leqslant n, \end{cases}$$

having the property that every one of their square minors is non-singular. For ease of reference, we think of M_l as the block matrix $(\mathbf{m}_l, B_l')$, where $\mathbf{m}_l$ denotes the first column of M_l. Associated with each of these matrices, we consider the block matrices

$$A_l' = \begin{cases} (\lambda_l I_{t-1}^*, \mathbf{m}_0), & \text{when } l = 0, \\ (\lambda_l I_{r-2}^+, \mathbf{m}_l), & \text{when } 1 \leqslant l \leqslant n. \end{cases}$$

Viewing the matrices A_l' and B_l' as examples of the matrices C_l' introduced in the previous paragraph, we form the linked block matrices $A^\dagger$ and $B^\dagger$. We refer to the block matrix $D = (A^\dagger, B^\dagger)$ as an *auxiliary matrix of type* $(n, r, t)_\omega$, and write $D = (d_{ij})$. Put

$$R = n(r-1) + t \quad \text{and} \quad S = 2R - 1 + \omega. \tag{2.1}$$

Then we see that $A^\dagger$ and $B^\dagger$ have respective formats $R \times R$ and $R \times (R-1+\omega)$, whilst D has format $R \times S$.

To illustrate this definition, we note that all the square minors of the matrix

$$\begin{pmatrix} 7 & 5 & 6 & 3 \\ 7 & 1 & 4 & 8 \\ 9 & 4 & 5 & 7 \\ 6 & 3 & 3 & 8 \end{pmatrix}$$

are non-singular, as the reader may care to verify, and hence[1]

$$\left(\begin{array}{ccccccccccccccccccccccccc}
8 & & & 7 & & & & & & & & & & 5 & 6 & 3 & & & & & & & & & \\
 & 8 & & 7 & & & & & & & & & & 1 & 4 & 8 & & & & & & & & & \\
 & & 8 & 9 & & & & & & & & & & 4 & 5 & 7 & & & & & & & & & \\
 & & & 6 & & & 7 & & & & & & & 3 & 3 & 8 & 5 & 6 & 3 & & & & & & \\
 & & & & 8 & & 7 & & & & & & & & & & 1 & 4 & 8 & & & & & & \\
 & & & & & 8 & 9 & & & & & & & & & & 4 & 5 & 7 & & & & & & \\
 & & & & & & 6 & & & 7 & & & & & & & 3 & 3 & 8 & 5 & 6 & 3 & & & \\
 & & & & & & & 8 & & 7 & & & & & & & & & & 1 & 4 & 8 & & & \\
 & & & & & & & & 8 & 9 & & & & & & & & & & 4 & 5 & 7 & & & \\
 & & & & & & & & & 6 & & & 7 & & & & & & & 3 & 3 & 8 & 5 & 6 & 3 \\
 & & & & & & & & & & 8 & & 7 & & & & & & & & & & 1 & 4 & 8 \\
 & & & & & & & & & & & 8 & 9 & & & & & & & & & & 4 & 5 & 7 \\
 & & & & & & & & & & & & 6 & & & & & & & & & & 3 & 3 & 8
\end{array}\right)$$

is an auxiliary matrix of type $(3, 4, 4)_0$. Were one to delete the first row and column of this matrix, the result would be an auxiliary matrix of type $(3, 4, 3)_1$.

Next, consider an integral auxiliary matrix $D = (d_{ij})$ of type $(n, r, t)_\omega$, define R and S as in (2.1), and define the linear forms

$$\gamma_j = \sum_{i=1}^{R} d_{ij} \alpha_i \quad (1 \leqslant j \leqslant S).$$

[1] We adopt the convention that zero entries in a matrix are left blank.

Introducing the Weyl sum
$$f(\alpha)=\sum_{1\leqslant x\leqslant P}e(\alpha x^3),$$
we define the mean value
$$I_\omega(P;D)=\oint |f(\gamma_1)\cdots f(\gamma_R)|^2|f(\gamma_{R+1})\cdots f(\gamma_S)|^4\,\mathrm{d}\boldsymbol{\alpha}. \tag{2.2}$$
Here, we use the suffix ω merely as an aide-memoire in keeping track of the type of the matrix D. We note in this context that by considering the underlying Diophantine system, one finds that $I_\omega(P;D)$ is unchanged by elementary row operations on D, and so in the discussion to come we may always pass to a convenient matrix row equivalent to D.

Before announcing our pivotal mean value estimate, we recall that Hua's lemma (see [17, Lemma 2.5]) shows that, for each natural number c, one has
$$\int_0^1 |f(c\theta)|^{2\nu}\,\mathrm{d}\theta \ll P^{\nu+\varepsilon}\quad (\nu=1,2). \tag{2.3}$$

Lemma 2.1. *Let D be an integral auxiliary matrix of type $(n,r,t)_\omega$ with $r\geqslant 3$. Then $I_\omega(P;D)\ll P^{3R-2+3\omega+\varepsilon}$.*

Proof. Throughout, we assume the nomenclature for the infrastructure of the matrix D introduced in the preamble to this lemma. We proceed by induction on $R\geqslant 2$. Since it is supposed that $t\geqslant 2$, it follows from (2.1) that when $R=2$, then $n=0$ and $t=2$. In such circumstances, one has
$$I_0(P;D)=\int_0^1\int_0^1 |f(\gamma_1)^2f(\gamma_2)^2f(\gamma_3)^4|\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2.$$
Observe that $\gamma_1=\lambda\alpha_1$ and, since all minors of M_0 are non-singular, it follows that γ_3 is linearly independent of both γ_1 and γ_2. Hence, by applying Schwarz's inequality in combination with (2.3) and a change of variables, one obtains
$$I_0(P;D)\ll\left(\int_0^1 |f(\theta)|^4\,\mathrm{d}\theta\right)^2\ll P^{4+\varepsilon}=P^{3R-2+\varepsilon}.$$
Meanwhile,
$$I_1(P;D)=\int_0^1\int_0^1 |f(\gamma_1)^2f(\gamma_2)^2f(\gamma_3)^4f(\gamma_4)^4|\,\mathrm{d}\alpha_1\,\mathrm{d}\alpha_2.$$
Since $\gamma_1=\lambda\alpha_1$ and all minors of M_0 are non-singular, we may employ the trivial estimate $|f(\gamma_2)|=O(P)$ in combination with a change of variables to deduce that there are fixed positive integers a, b, c and d for which
$$I_1(P;D)\ll P^2\int_0^1\int_0^1 |f(a\theta_1)^4f(b\theta_2)^4f(c\theta_1+d\theta_2)^2|\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2.$$
Consequently, by a pedestrian generalisation of [4, Theorem 1] (see especially equations (6) and (7) therein), one finds that
$$I_1(P;D)\ll P^2(P^{5+\varepsilon})=P^{3R+1+\varepsilon}.$$

We have thus confirmed the conclusion of the lemma when $R = 2$.

Suppose next that $R \geqslant 3$, and that the conclusion of the lemma holds for all auxiliary matrices D having fewer than R rows. We divide our discussion into cases according to the value of ω.

Case I: $\omega = 0$.
We first consider the situation in which the integral auxiliary matrix D has R rows and $\omega = 0$. By orthogonality, one sees that $I_0(P; D)$ counts the number of integral solutions of the system

$$\sum_{j=1}^{R} d_{ij}(x_{j1}^3 - x_{j2}^3) + \sum_{j=R+1}^{S} d_{ij}(x_{j1}^3 + x_{j2}^3 - x_{j3}^3 - x_{j4}^3) = 0 \quad (1 \leqslant i \leqslant R), \tag{2.4}$$

with $1 \leqslant x_{jl} \leqslant P$ for each j and l. Let T_0 denote the number of these solutions in which $x_{j1} = x_{j2}$ for $1 \leqslant j \leqslant t-1$, and let T_j denote the corresponding number where instead $x_{j1} \neq x_{j2}$. Then

$$I_0(P; D) \leqslant T_0 + T_1 + \ldots + T_{t-1}. \tag{2.5}$$

An inspection of (2.4) reveals that

$$T_0 \ll P^{t-1} J_0, \tag{2.6}$$

where J_0 counts the number of integral solutions of the system

$$\sum_{j=t}^{R} d_{ij}(x_{j1}^3 - x_{j2}^3) + \sum_{j=R+1}^{S} d_{ij}(x_{j1}^3 + x_{j2}^3 - x_{j3}^3 - x_{j4}^3) = 0 \quad (1 \leqslant i \leqslant R), \tag{2.7}$$

with $1 \leqslant x_{jl} \leqslant P$ for each j and l. We observe that the equations in (2.7) with $1 \leqslant i \leqslant t-1$ involve only the variables x_{jl} with $j = t$ and $R+1 \leqslant j \leqslant R+t-1$. The coefficient matrix associated with these equations and variables is the matrix M_0^* obtained from M_0 by deleting its final row. By taking appropriate linear combinations of the first $t-1$ equations of (2.7), corresponding to elementary row operations on M_0^*, we may therefore replace the equations in (2.7) with $1 \leqslant i \leqslant t-1$ by the new equations

$$u_i(x_{t,1}^3 - x_{t,2}^3) + v_i(x_{R+i,1}^3 + x_{R+i,2}^3 - x_{R+i,3}^3 - x_{R+i,4}^3) = 0 \quad (1 \leqslant i \leqslant t-1), \tag{2.8}$$

in which u_i and $v_i \neq 0$ $(1 \leqslant i \leqslant t-1)$ are suitable integers. Put $\tau = t + r - 1$. Then adding appropriate multiples of the equations (2.8) to the equation in (2.7) with $i = t$, one finds that the latter equation may be replaced by

$$u_t(x_{t1}^3 - x_{t2}^3) + d_{t\tau}(x_{\tau 1}^3 - x_{\tau 2}^3) + \sum_{j=R+t}^{S} d_{tj}(x_{j1}^3 + x_{j2}^3 - x_{j3}^3 - x_{j4}^3) = 0, \tag{2.9}$$

for a suitable rational number u_t. The coefficient matrix M_0^+ associated with these t new equations (2.8) and (2.9), and variables x_{tl} and x_{jl} $(R+1 \leqslant j \leqslant R+t-1)$ has been obtained from M_0 by a succession of elementary row operations, and hence is non-singular. Since $\det(M_0^+) = (-1)^{t-1} u_t v_1 \cdots v_{t-1}$, we therefore see that $u_t \neq 0$.

We now investigate the number N_0 of integral solutions of the system of equations defined by (2.9) and the equations of (2.7) for which $t+1 \leqslant i \leqslant R$, with $1 \leqslant x_{jl} \leqslant P$ for each j and l. When $n = 0$, the whole system reduces to the single equation

$$u_t(x_{t1}^3 - x_{t2}^3) = 0,$$

so that $N_0 \ll P$. Otherwise, when $n \geqslant 1$, we observe that, by taking appropriate non-zero integral multiples of the equations, there is no loss of generality in assuming that $u_t = d_{ii}$ $(t+1 \leqslant i \leqslant t+r-2)$. In this way, one finds that $N_0 = I_0(P; D_1)$, where D_1 is an auxiliary matrix of type $(n-1, r, r)_0$ having $(n-1)(r-1)+r$ rows. Consequently, our inductive hypothesis shows that

$$I_0(P; D_1) \ll P^{3((n-1)(r-1)+r)-2+\varepsilon} = P^{3(R-t)+1+\varepsilon}.$$

Then in both cases, we have $N_0 \ll P^{3(R-t)+1+\varepsilon}$.

Now consider any fixed solution counted by N_0, and consider the number N_1 of solutions x_{jl} $(R+1 \leqslant j \leqslant R+t-1$ and $1 \leqslant l \leqslant 4)$ satisfying the equations (2.8). Since the variables x_{t1} and x_{t2} are fixed, it follows from orthogonality via the triangle inequality that

$$N_1 \ll \prod_{i=1}^{t-1} \left(\int_0^1 |f(v_i\theta_i)|^4 \, \mathrm{d}\theta_i \right).$$

Then we conclude from (2.3) that $N_1 \ll (P^{2+\varepsilon})^{t-1}$, and hence

$$J_0 \ll (P^{2+\varepsilon})^{t-1} N_0 \ll P^{3(R-t)+2(t-1)+1+t\varepsilon}.$$

On substituting this estimate into (2.6), we obtain the bound

$$T_0 \ll P^{3(R-t)+3(t-1)+1+\varepsilon} = P^{3R-2+\varepsilon}. \tag{2.10}$$

We next turn to the problem of bounding T_j for $1 \leqslant j \leqslant t-1$. We restrict attention in the first instance to the case $j = 1$, since, as will become transparent as our argument unfolds, the same method applies also for the remaining values of j. Write

$$T(h) = \oint |f(\gamma_2) \cdots f(\gamma_R)|^2 |f(\gamma_{R+1}) \cdots f(\gamma_S)|^4 e(\gamma_1 h) \, \mathrm{d}\boldsymbol{\alpha}. \tag{2.11}$$

Then we find by orthogonality that

$$T_1 = \sum_{h \in \mathbb{Z} \setminus \{0\}} c_h T(h),$$

where c_h denotes the number of integral solutions of $d_{11}(x^3 - y^3) = h$, with $1 \leqslant x, y \leqslant P$. An elementary divisor function estimate shows that $c_h = O(|h|^\varepsilon)$ when $h \neq 0$. Since $c_h = 0$ for $|h| > P^4$, one deduces from (2.11) and a consideration of the underlying Diophantine system that

$$T_1 \ll P^\varepsilon \sum_{h \in \mathbb{Z} \setminus \{0\}} T(h). \tag{2.12}$$

The sum over h on the right hand side here is bounded above by the number N_2 of solutions of the system (2.4) with $2 \leqslant i \leqslant R$ and $x_{11} = x_{12} = 0$. When

$t \geqslant 3$, one sees that $N_2 = I_1(P; D_2)$, where D_2 is the auxiliary matrix of type $(n, r, t-1)_1$ obtained from D by deleting its first row and column. Since D_2 has $R-1$ rows, it follows from the inductive hypothesis that

$$T_1 \ll P^{\varepsilon} I_1(P; D_2) \ll P^{3(R-1)+1+2\varepsilon}.$$

We therefore conclude from (2.5) via (2.10) that when $t \geqslant 3$, one has

$$I_0(P; D) \ll P^{3R-2+\varepsilon},$$

thereby confirming the inductive hypothesis for D.

It remains to handle the situation in which $t = 2$. Note that since we have assumed $R \geqslant 3$, it follows that $n \geqslant 1$. For the sake of concision, we abbreviate $(\alpha_2, \ldots, \alpha_R)$ to $\boldsymbol{\alpha}'$, and then define $\gamma_j'(\boldsymbol{\alpha}') = \gamma_j(0, \alpha_2, \ldots, \alpha_R)$. We put

$$\mathfrak{F}(\alpha_2) = \oint |f(\gamma_3') \cdots f(\gamma_R')|^2 |f(\gamma_{R+2}') \cdots f(\gamma_S')|^4 \, \mathrm{d}(\alpha_3, \ldots, \alpha_R),$$

and observe that, by orthogonality, one has

$$N_2 = \int_0^1 |f(d_{2,2}\alpha_2)^2 f(d_{R+1,2}\alpha_2)^4| \mathfrak{F}(\alpha_2) \, \mathrm{d}\alpha_2.$$

We apply the Hardy-Littlewood method to estimate N_2. Denote by $\mathfrak{M}$ the union of the intervals

$$\mathfrak{M}(q, a) = \{\alpha \in [0, 1) : |q\alpha - a| \leqslant P^{-9/4}\}, \tag{2.13}$$

with $0 \leqslant a \leqslant q \leqslant P^{3/4}$ and $(a, q) = 1$, and put $\mathfrak{m} = [0, 1) \setminus \mathfrak{M}$. Let c be a fixed non-zero integer. Then, as a special case of [2, Lemma 3.4], or as a consequence of the methods of [17, Chapter 4]), one has

$$\int_{\mathfrak{M}} |f(c\theta)|^4 \, \mathrm{d}\theta \ll P^{1+\varepsilon}. \tag{2.14}$$

In addition, an enhanced version of Weyl's inequality (see [15, Lemma 1]) shows that

$$\sup_{\theta \in \mathfrak{m}} |f(c\theta)| \ll P^{3/4+\varepsilon}. \tag{2.15}$$

On the one hand, it follows from (2.15) that

$$\int_{\mathfrak{m}} |f(d_{2,2}\alpha_2)^2 f(d_{R+1,2}\alpha_2)^4| \mathfrak{F}(\alpha_2) \, \mathrm{d}\alpha_2 \ll P^{3+\varepsilon} \int_0^1 |f(d_{2,2}\alpha_2)^2| \mathfrak{F}(\alpha_2) \, \mathrm{d}\alpha_2.$$

Put $\tau = r + 1$. Then, by orthogonality, the integral on the right hand side counts the number of integral solutions of the system of equations given by

$$d_{22}(x_{21}^3 - x_{22}^3) + d_{2\tau}(x_{j1}^3 - x_{j2}^3) + \sum_{j=R+2}^{S} d_{2j}(x_{j1}^3 + x_{j2}^3 - x_{j3}^3 - x_{j4}^3) = 0$$

and

$$\sum_{j=3}^{R} d_{ij}(x_{j1}^3 - x_{j2}^3) + \sum_{j=R+2}^{S} d_{ij}(x_{j1}^3 + x_{j2}^3 - x_{j3}^3 - x_{j4}^3) = 0 \quad (3 \leqslant i \leqslant R), \tag{2.16}$$

with $1 \leqslant x_{jl} \leqslant P$ for each j and l. By taking appropriate non-zero integral multiples of these equations, there is no loss of generality in assuming that $d_{22} = d_{ii}$ $(3 \leqslant i \leqslant r)$. The coefficient matrix D_3 associated with these equations and variables arises from D by deleting its first row, and the first and $(R+1)$-st column, and can be seen to be an auxiliary matrix of type $(n-1, r, r)_0$ having $R-1$ rows. It therefore follows from the inductive hypothesis that

$$\int_{\mathfrak{m}} |f(d_{2,2}\alpha)^2 f(d_{R+1,2}\alpha)^4| \mathfrak{F}(\alpha) \, \mathrm{d}\alpha \ll P^{3+\varepsilon} I_0(P; D_3) \\ \ll P^{3+\varepsilon}(P^{3(R-1)-2+\varepsilon}). \tag{2.17}$$

We next consider the corresponding major arc contribution. By orthogonality, the mean value $\mathfrak{F}(\alpha)$ is bounded above by the number of solutions of an associated Diophantine system, in which each solution is counted with a unimodular weight depending on α. Thus we have $\mathfrak{F}(\alpha) \leqslant \mathfrak{F}(0)$. Consequently, it follows from (2.14) via the trivial estimate $|f(d_{2,2}\alpha)| = O(P)$ that

$$\int_{\mathfrak{M}} |f(d_{2,2}\alpha)^2 f(d_{R+1,2}\alpha)^4| \mathfrak{F}(\alpha) \, \mathrm{d}\alpha \ll \mathfrak{F}(0) P^2 \int_{\mathfrak{M}} |f(d_{R+1,2}\alpha)|^4 \, \mathrm{d}\alpha \\ \ll \mathfrak{F}(0) P^{3+\varepsilon}.$$

By orthogonality, the mean value $\mathfrak{F}(0)$ counts the number of integral solutions of the system (2.16). The coefficient matrix D_4 associated with these equations and variables arises from D by deleting its first two rows, and columns 1, 2 and $R+1$, and can be seen to be an auxiliary matrix of type $(n-1, r, r-1)_1$ having $R-2$ rows. It therefore follows from the inductive hypothesis that

$$\int_{\mathfrak{M}} |f(d_{2,2}\alpha)^2 f(d_{R+1,2}\alpha)^4| \mathfrak{F}(\alpha) \, \mathrm{d}\alpha \ll P^{3+\varepsilon} I_1(P; D_4) \\ \ll P^{3+\varepsilon}(P^{3(R-2)+1+\varepsilon}). \tag{2.18}$$

On combining (2.12), (2.17) and (2.18), we conclude that when $t = 2$ one has $T_1 \ll P^\varepsilon N_2 \ll P^{3R-2+2\varepsilon}$. We therefore deduce from (2.5) and (2.10) that $I_0(P; D) \ll P^{3R-2+\varepsilon}$, confirming the inductive hypothesis for D when $t = 2$.

Case II: $\omega = 1$.

We now turn to the situation in which the integral auxiliary matrix D has R rows and $\omega = 1$. Observe that $\gamma_{R+j}(\boldsymbol{\alpha})$ depends only on $\boldsymbol{\alpha}^* = (\alpha_1, \ldots, \alpha_t)$ for $1 \leqslant j \leqslant t$. When $1 \leqslant j \leqslant t$, we define $\mathfrak{B}_j^*$ to be the set of t-tuples $\boldsymbol{\alpha}^* \in [0,1)^t$ for which $\gamma_{R+j}(\boldsymbol{\alpha}^*, \mathbf{0}) \in \mathfrak{m} + \mathbb{Z}$, and we define $\mathfrak{B}_0^*$ to be the complementary set of t-tuples $\boldsymbol{\alpha}^* \in [0,1)^t$ for which $\gamma_{R+j}(\boldsymbol{\alpha}^*, \mathbf{0}) \notin \mathfrak{m} + \mathbb{Z}$ $(1 \leqslant j \leqslant t)$. We then put $\mathfrak{B}_j = \mathfrak{B}_j^* \times [0,1)^{R-t}$ $(0 \leqslant j \leqslant t)$. Thus $[0,1)^R \subseteq \mathfrak{B}_0 \cup \mathfrak{B}_1 \cup \ldots \cup \mathfrak{B}_t$. When $\mathfrak{B} \subseteq [0,1)^R$, we write

$$I(\mathfrak{B}) = \int_{\mathfrak{B}} |f(\gamma_1) \cdots f(\gamma_R)|^2 |f(\gamma_{R+1}) \cdots f(\gamma_S)|^4 \, \mathrm{d}\boldsymbol{\alpha}.$$

Then it follows from (2.2) that

$$I_1(P; D) \leqslant I(\mathfrak{B}_0) + I(\mathfrak{B}_1) + \ldots + I(\mathfrak{B}_t). \tag{2.19}$$

We begin by estimating $I(\mathfrak{B}_1)$. It follows from (2.15) that

$$\sup_{\boldsymbol{\alpha}\in\mathfrak{B}_1} |f(\gamma_{R+1})| \leqslant \sup_{\gamma\in\mathfrak{m}} |f(\gamma)| \ll P^{3/4+\varepsilon},$$

and hence

$$I(\mathfrak{B}_1) \ll P^{3+\varepsilon} \oint |f(\gamma_1)\cdots f(\gamma_R)|^2 |f(\gamma_{R+2})\cdots f(\gamma_S)|^4 \,\mathrm{d}\boldsymbol{\alpha}.$$

The integral on the right hand side counts the number N_3 of integral solutions of the system (2.4) with $1 \leqslant x_{jl} \leqslant P$ for each $j \neq R+1$ and l, but with $x_{R+1,l} = 0$. Thus $N_3 = I(P; D_5)$, where D_5 is the matrix obtained from D by deleting its $(R+1)$-st column. Note that deleting a column from a matrix, all of whose square minors are non-singular, does not change the latter property. Hence D_5 is an auxiliary matrix of type $(n, r, t)_0$ having R rows. It therefore follows from the inductive hypothesis that

$$I(\mathfrak{B}_1) \ll P^{3+\varepsilon} I_0(P; D_5) \ll P^{3+\varepsilon}(P^{3R-2+\varepsilon}). \tag{2.20}$$

As indicated earlier, a symmetrical argument shows that $I(\mathfrak{B}_j)$ is bounded in the same manner for $2 \leqslant j \leqslant t$.

We finish by estimating $I(\mathfrak{B}_0)$. Note that whenever $\boldsymbol{\alpha} \in \mathfrak{B}_0$, then $\gamma_{R+j} \in \mathfrak{M} + \mathbb{Z}$ for $1 \leqslant j \leqslant t$. We put

$$\mathfrak{G}(\boldsymbol{\alpha}^*) = \oint |f(\gamma_{t+1})\cdots f(\gamma_R)|^2 |f(\gamma_{R+t+1})\cdots f(\gamma_S)|^4 \,\mathrm{d}(\alpha_{t+1}, \dots, \alpha_R),$$

and apply the trivial estimate $|f(\gamma_j)| \leqslant P$ $(1 \leqslant j \leqslant t)$. Then one finds that

$$I(\mathfrak{B}_0) \ll P^{2t} \int_{\mathfrak{B}_0^*} |f(\gamma_{R+1})\cdots f(\gamma_{R+t})|^4 \mathfrak{G}(\boldsymbol{\alpha}^*) \,\mathrm{d}\boldsymbol{\alpha}^*. \tag{2.21}$$

Observe that by orthogonality, and an argument paralleling that in the discussion following (2.17), one has $\mathfrak{G}(\boldsymbol{\alpha}^*) \leqslant \mathfrak{G}(\mathbf{0})$. Also, one sees that $\mathfrak{G}(\mathbf{0})$ counts the number of integral solutions of the system (2.4) for $t+1 \leqslant i \leqslant R$, with $1 \leqslant x_{jl} \leqslant P$ for $t+1 \leqslant j \leqslant R$ and $R+t+1 \leqslant j \leqslant S$, and with the remaining variables 0. Thus $\mathfrak{G}(\mathbf{0}) = I_1(P; D_6)$, where D_6 is the matrix obtained from D by deleting its first t rows and columns j with $1 \leqslant j \leqslant t$ and $R+1 \leqslant j \leqslant R+t$. Hence D_6 is an auxiliary matrix of type $(n-1, r, r-1)_1$ having $R-t$ rows. It therefore follows from the inductive hypothesis that $\mathfrak{G}(\mathbf{0}) \ll P^{3(R-t)+1+\varepsilon}$. By substituting this estimate into (2.21) and making an appropriate change of variables justified by the non-singularity of the matrix B_0', it follows that

$$I(\mathfrak{B}_0) \ll P^{2t}\mathfrak{G}(\mathbf{0}) \int_{\mathfrak{M}^t} |f(\theta_1)\cdots f(\theta_t)|^4 \,\mathrm{d}\boldsymbol{\theta}.$$

An application of (2.14) therefore yields

$$I(\mathfrak{B}_0) \ll P^{3R-t+1+\varepsilon}(P^{1+\varepsilon})^t \ll P^{3R+1+(t+1)\varepsilon}.$$

In combination with (2.20), and its generalisations estimating $I(\mathfrak{B}_j)$ for $2 \leqslant j \leqslant t$, we conclude from (2.19) that $I_1(P; D) \ll P^{3R+1+\varepsilon}$. This confirms the inductive hypothesis for D when $\omega = 1$, completing the proof of the lemma. $\square$

By a modification of the argument of the proof of Lemma 2.1, one may handle also the case $r = 2$. However, we are able to establish all of the conclusions recorded in the introduction without appealing to this special case.

3. Complification

We now employ a recursive *complification* argument, in which, at each step, mean values associated with R equations are estimated in terms of a mean value associated with $2R-1$ equations. In this way, we are able to apply the estimates supplied by Lemma 2.1 to obtain powerful estimates for suitable mixed moments of order $2R$ of generating functions associated with sums of three cubes. We begin with a lemma concerning highly non-singular matrices.

Lemma 3.1. *Let $A = (A_1, A_2)$ be a block matrix in which A_1 and A_2 are each of format $r \times r$. Then A is highly non-singular if and only if A_1 and A_2 are non-singular, and all square minors of $A_1^{-1}A_2$ are non-singular.*

Proof. The non-singularity condition on A_1 and A_2 is immediate from the definition of what it means to be highly non-singular. Thus, by applying elementary row operations, it suffices to consider the situation with $A = (I_r, A_1^{-1}A_2)$. The matrix A is highly non-singular if and only if all collections of r of its columns are linearly independent. Given any $l \times l$ minor M of $A_1^{-1}A_2$, inhabiting the columns $\mathbf{v}_1, \ldots, \mathbf{v}_l$, say, one can select a complementary set of columns $\mathbf{e}_1, \ldots, \mathbf{e}_{r-l}$ from I_r in such a manner that

$$\det(M) = \pm \det(\mathbf{v}_1, \ldots, \mathbf{v}_l, \mathbf{e}_1, \ldots, \mathbf{e}_{r-l}).$$

Then all collections of r of the columns of A are linearly independent if and only if $\det(M) \neq 0$ for all square minors M of $A_1^{-1}A_2$, as claimed. $\Box$

We next prepare the cast of generating functions needed to describe the complification process. With the needs of §5 in mind, we proceed in slightly greater generality than demanded by the proofs of Theorems 1.1 and 1.2. Let $\sigma \in [0, 1)$. When $2 \leqslant Z \leqslant P$, we put

$$\mathcal{A}(P, Z) = \{n \in [1, P] \cap \mathbb{Z} : p \text{ prime and } p|n \Rightarrow p \leqslant Z\},$$

and introduce the exponential sums

$$f_0(\alpha) = \sum_{\sigma P < x \leqslant P} e(\alpha x^3) \quad \text{and} \quad g(\alpha) = \sum_{\substack{\sigma P < x \leqslant P \\ x \in \mathcal{A}(P, P^\eta)}} e(\alpha x^3). \tag{3.1}$$

We then take $\phi_1(\alpha) = f_0(\alpha)^2$ and $\phi_2(\alpha) = g(\alpha)^2$, and write

$$F_l(\alpha) = f_0(\alpha)\phi_l(\alpha) \quad \text{and} \quad \Phi_l(\alpha) = f_0(\alpha)^2\phi_l(\alpha) \quad (l = 1, 2). \tag{3.2}$$

Finally, for the sake of convenience, we put

$$\nu_1 = \tfrac{1}{2} \quad \text{and} \quad \nu_2 = 3\xi = (\sqrt{2833} - 43)/41. \tag{3.3}$$

Lemma 3.2. *When $\eta > 0$ is sufficiently small, one has*

$$\int_0^1 |F_l(\alpha)|^2 \, \mathrm{d}\alpha \ll P^{3+\nu_l+\varepsilon} \quad (l = 1, 2).$$

Proof. When $l = 1$, this is an immediate consequence of Hua's lemma (see [17, Lemma 2.5]) in combination with Schwarz's inequality. When $l = 2$, meanwhile, this follows from [20, Theorem 1.2] by considering the underlying Diophantine equations. $\square$

Next, let n and r be non-negative integers with $r \geqslant 2$, and write $R = n(r-1)$. Let $\Lambda = (\lambda_{i,j})$ be an integral $(R+1)\times(2R+2)$ matrix, write $\boldsymbol{\lambda}_j$ for the column vector $(\lambda_{i,j})_{1\leqslant i\leqslant R+1}$, and define $\boldsymbol{\lambda}_j^\flat$ to be the column vector $(\lambda_{R+2-i,j})_{1\leqslant i\leqslant R+1}$ in which the entries of $\boldsymbol{\lambda}_j$ are flipped upside-down. Also, let

$$\beta_j(\boldsymbol{\alpha}) = \sum_{i=1}^{R+1} \lambda_{i,j}\alpha_i \quad (0 \leqslant j \leqslant 2R+1). \tag{3.4}$$

We say that the matrix Λ is *adjuvant of type* (n,r) when the column vectors $\boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_R$ and $\boldsymbol{\lambda}_{R+2}, \boldsymbol{\lambda}_{R+3}, \ldots, \boldsymbol{\lambda}_{2R+1}$, respectively, may be permuted to form matrices $A^\dagger$ and $B^\dagger$ having the property that the block matrix $(A^\dagger, B^\dagger)$ is auxiliary of type $(n-1, r, r)_0$, and also the same property holds for the respective column vectors $\boldsymbol{\lambda}_{R+1}^\flat, \boldsymbol{\lambda}_R^\flat, \ldots, \boldsymbol{\lambda}_1^\flat$ and $\boldsymbol{\lambda}_{2R+1}^\flat, \boldsymbol{\lambda}_{2R}^\flat, \ldots, \boldsymbol{\lambda}_{R+2}^\flat$. We also adopt the convention that

$$\boldsymbol{\phi}_{a,b}^{(l)}(\boldsymbol{\beta}) = \prod_{j=a}^{b} \phi_l(\beta_j) \quad \text{and} \quad \boldsymbol{\Phi}_{a,b}^{(l)}(\boldsymbol{\beta}) = \prod_{j=a}^{b} \Phi_l(\beta_j).$$

We then introduce the mean value

$$J_l(P;\Lambda) = \oint |F_l(\beta_0)\boldsymbol{\phi}_{1,R}^{(l)}(\boldsymbol{\beta})F_l(\beta_{R+1})\boldsymbol{\Phi}_{R+2,2R+1}^{(l)}(\boldsymbol{\beta})|\,\mathrm{d}\boldsymbol{\alpha}. \tag{3.5}$$

Finally, we fix $\eta > 0$ to be sufficiently small in the context of Lemma 3.2.

Lemma 3.3. *Suppose that Λ is an integral adjuvant matrix of type (n,r). Then there exists an integral adjuvant matrix Λ^* of type $(2n,r)$ for which*

$$J_l(P;\Lambda) \ll (P^{3+\nu_l+\varepsilon})^{1/2} J_l(P;\Lambda^*)^{1/2} \quad (l = 1,2).$$

Proof. We fix $l \in \{1,2\}$, and for the sake of concision suppress mention of l in our notation. Define the linear forms β_j as in (3.4). Since the matrix Λ is adjuvant, one may suppose that $\beta_{R+1} = \lambda_{R+1,R+1}\alpha_{R+1}$ with $\lambda_{R+1,R+1} \neq 0$. Define

$$T(P;\Lambda) = \int_0^1 \left(\oint |F(\beta_0)\boldsymbol{\phi}_{1,R}(\boldsymbol{\beta})\boldsymbol{\Phi}_{R+2,2R+1}(\boldsymbol{\beta})|\,\mathrm{d}\widehat{\boldsymbol{\alpha}}_R\right)^2 \mathrm{d}\alpha_{R+1},$$

where $\widehat{\boldsymbol{\alpha}}_R$ denotes $(\alpha_1, \ldots, \alpha_R)$. Then Schwarz's inequality conveys us from (3.5) to the bound

$$J(P;\Lambda) \leqslant \left(\int_0^1 |F(\beta_{R+1})|^2 \,\mathrm{d}\alpha_{R+1}\right)^{1/2} T(P;\Lambda)^{1/2}. \tag{3.6}$$

Define $\beta_j^* = \beta_j^*(\widehat{\boldsymbol{\alpha}}_{2R+1})$ by

$$\beta_j^* = \begin{cases} \beta_j(\alpha_1, \ldots, \alpha_{R+1}), & \text{when } 0 \leqslant j \leqslant R, \\ \beta_{2R+1-j}(\alpha_{2R+1}, \ldots, \alpha_{R+1}), & \text{when } R+1 \leqslant j \leqslant 2R+1, \\ \beta_{j-R}(\alpha_1, \ldots, \alpha_{R+1}), & \text{when } 2R+2 \leqslant j \leqslant 3R+1, \\ \beta_{5R+3-j}(\alpha_{2R+1}, \ldots, \alpha_{R+1}), & \text{when } 3R+2 \leqslant j \leqslant 4R+1. \end{cases}$$

Then, by expanding the square inside the outermost integration, we see that

$$T(P;\Lambda) = \oint |F(\beta_0^*)\boldsymbol{\phi}_{1,2R}(\boldsymbol{\beta}^*)F(\beta_{2R+1}^*)\boldsymbol{\Phi}_{2R+2,4R+1}(\boldsymbol{\beta}^*)| \,\mathrm{d}\widehat{\boldsymbol{\alpha}}_{2R+1}.$$

The integral $(2R+1) \times (4R+2)$ matrix $\Lambda^* = (\lambda_{ij}^*)$ defining the linear forms $\beta_0^*, \ldots, \beta_{4R+1}^*$ is adjuvant of type $(2n, r)$.

Write Λ^* in block form (A^*, B^*) with A^* and B^* having $2R+2$ and $2R$ columns, respectively. It may be illuminating to note that one may permute the columns of the matrix B^* to form a linked block matrix $(B^*)^\dagger$ built from two blocks, with upper left hand block B and lower right hand block $B^\flat$, in which $B^\flat$ denotes the matrix B rotated through $180°$. Likewise, one sees that the columns of the matrix A^* may be permuted to form a linked block matrix $(A^*)^\dagger$ built in similar manner, but with upper left hand block A_1, where A_1 denotes the matrix A with final column deleted, and with lower right hand block $A_1^\flat$, in the sense described.

Thus we conclude that $T(P;\Lambda) = J(P;\Lambda^*)$. The conclusion of the lemma therefore follows from (3.6), since Lemma 3.2 supplies the estimate

$$\int_0^1 |F(\beta_{R+1})|^2 \,\mathrm{d}\alpha_{R+1} \ll P^{3+\nu+\varepsilon}.$$

$\square$

Consider an $r \times 2r$ integral matrix $C = (c_{ij})$, write $\mathbf{c}_j$ for the column vector $(c_{ij})_{1 \leqslant i \leqslant r}$, and put

$$\gamma_j = \sum_{i=1}^r c_{ij}\alpha_i \quad (0 \leqslant j \leqslant 2r-1).$$

Also, write

$$K_l(P;C) = \oint |F_l(\gamma_0) \cdots F_l(\gamma_{2r-1})| \,\mathrm{d}\boldsymbol{\alpha} \quad (l = 1, 2). \tag{3.7}$$

We divide the proof of the next theorem according to whether $r \geqslant 3$ or $r = 2$.

Theorem 3.4. *Suppose that $r \geqslant 2$, and that the $r \times 2r$ integral matrix C is highly non-singular. Then $K_l(P;C) \ll P^{3r+\nu_l+\varepsilon}$ $(l = 1, 2)$.*

Proof when $r \geqslant 3$. We again suppress mention of l in our notation within this proof. We begin by applying Schwarz's inequality to $K(P;C)$, showing that

$$K(P;C) \leqslant K^{(1)}(P;C)^{1/2} K^{(2)}(P;C)^{1/2}, \tag{3.8}$$

where

$$K^{(1)}(P;C)=\oint |F(\gamma_0)\boldsymbol{\phi}_{1,r-2}(\boldsymbol{\gamma})F(\gamma_{r-1})\phi(\gamma_r)\boldsymbol{\Phi}_{r+1,2r-1}(\boldsymbol{\gamma})|\,\mathrm{d}\boldsymbol{\alpha}$$

and

$$K^{(2)}(P;C)=\oint |F(\gamma_0)\boldsymbol{\phi}_{r+1,2r-2}(\boldsymbol{\gamma})F(\gamma_{r-1})\phi(\gamma_{2r-1})\Phi(\gamma_r)\boldsymbol{\Phi}_{1,r-2}(\boldsymbol{\gamma})|\,\mathrm{d}\boldsymbol{\alpha}.$$

The coefficient matrix associated with the linear forms

$$\gamma_0,\gamma_{r+1},\dots,\gamma_{2r-2},\gamma_{r-1},\gamma_{2r-1},\gamma_r,\gamma_1,\dots,\gamma_{r-2}$$

is obtained by permuting the columns of C, and hence is highly non-singular. We may therefore confine our attention to $K^{(1)}(P;C)$.

Write C in block form (A,B), where both A and B are $r\times r$ integral matrices, noting that the highly non-singular property of C ensures that both A and B are non-singular. Note also that Lemma 3.1 shows that every square minor of $A^{-1}B$ is non-singular. It is convenient to put

$$\boldsymbol{\gamma}^{(1)}=(\gamma_0,\dots,\gamma_{r-1})^T\quad\text{and}\quad\boldsymbol{\gamma}^{(2)}=(\gamma_r,\dots,\gamma_{2r-1})^T.$$

We then have

$$\boldsymbol{\gamma}^{(1)}=A^T\boldsymbol{\alpha}\quad\text{and}\quad\boldsymbol{\gamma}^{(2)}=B^T\boldsymbol{\alpha}.$$

Let $\Delta=|\det A|$. We substitute $\boldsymbol{\theta}=\Delta^{-1}A^T\boldsymbol{\alpha}$, so that

$$\boldsymbol{\gamma}^{(1)}=\Delta\boldsymbol{\theta}\quad\text{and}\quad\boldsymbol{\gamma}^{(2)}=\Delta(A^{-1}B)^T\boldsymbol{\theta},$$

and then define the linear forms $\beta_j(\theta_1,\dots,\theta_r)\in\mathbb{Z}[\boldsymbol{\theta}]$ by means of the relation $\boldsymbol{\beta}=\Delta(A^{-1}B)^T\boldsymbol{\theta}$, in which $\boldsymbol{\beta}=(\beta_{r+1},\dots,\beta_{2r})^T$. Since the underlying exponential sums are periodic with period 1, we may apply the transformation formula to conclude that

$$K^{(1)}(P;C)=\oint |F(\Delta\theta_1)\boldsymbol{\phi}_{2,r-1}(\Delta\boldsymbol{\theta})F(\Delta\theta_r)\phi(\beta_{r+1})\boldsymbol{\Phi}_{r+2,2r}(\boldsymbol{\beta})|\,\mathrm{d}\boldsymbol{\theta}.$$

The matrix of coefficients of the linear forms defining this mean value, namely

$$\Delta\theta_1,\dots,\Delta\theta_r,\beta_{r+1}(\boldsymbol{\theta}),\dots,\beta_{2r}(\boldsymbol{\theta}),$$

is given by $(\Delta I_r,\Delta(A^{-1}B)^T)$, which, in view of Lemma 3.1, is highly non-singular. In particular, all the square minors of $\Delta(A^{-1}B)^T$ are non-singular.

Define

$$T(P;C)=\int_0^1\left(\oint |F(\Delta\theta_1)\boldsymbol{\phi}_{2,r-1}(\Delta\boldsymbol{\theta})\phi(\beta_{r+1})\boldsymbol{\Phi}_{r+2,2r}(\boldsymbol{\beta})|\,\mathrm{d}\widehat{\boldsymbol{\theta}}_{r-1}\right)^2\mathrm{d}\theta_r.$$

Then by Schwarz's inequality, one finds that

$$K^{(1)}(P;C)\leqslant\left(\int_0^1|F(\Delta\theta_r)|^2\,\mathrm{d}\theta_r\right)^{1/2}T(P;C)^{1/2}.\tag{3.9}$$

By expanding the square in the definition of $T(P;C)$, we see that

$$T(P;C)=\oint |F(\beta_0^*)\boldsymbol{\phi}_{1,2r-2}(\boldsymbol{\beta}^*)F(\beta_{2r-1}^*)\boldsymbol{\Phi}_{2r,4r-3}(\boldsymbol{\beta}^*)|\,\mathrm{d}\widehat{\boldsymbol{\theta}}_{2r-1},$$

where $\beta_j^* = \beta_j^*(\boldsymbol{\theta})$ is defined by

$$\beta_j^* = \begin{cases} \Delta\theta_{j+1}, & \text{when } 0 \leqslant j \leqslant 2r-3 \text{ and } j \neq r-1, \\ \beta_{r+1}(\theta_1, \dots, \theta_r), & \text{when } j = r-1, \\ \beta_{r+1}(\theta_{2r-1}, \dots, \theta_r), & \text{when } j = 2r-2, \\ \Delta\theta_{2r-1}, & \text{when } j = 2r-1, \\ \beta_{j-r+2}(\theta_1, \dots, \theta_r), & \text{when } 2r \leqslant j \leqslant 3r-2, \\ \beta_{5r-1-j}(\theta_{2r-1}, \dots, \theta_r), & \text{when } 3r-1 \leqslant j \leqslant 4r-3. \end{cases}$$

It is apparent that the matrix of coefficients C' of the linear forms

$$\beta_0^*(\boldsymbol{\theta}), \dots, \boldsymbol{\beta}_{4r-3}^*(\boldsymbol{\theta})$$

is an integral adjuvant matrix of type $(2, r)$. Thus, in the notation introduced in (3.5), we see that $T(P; C) = J(P; C')$. By virtue of the conclusion of Lemma 3.2, we therefore infer from (3.8) and (3.9) that there exist integral adjuvant matrices $C^{(1)}$ and $C^{(2)}$ of type $(2, r)$ for which

$$\begin{aligned} K(P; C) &\ll (P^{3+\nu+\varepsilon})^{1/2} J(P; C^{(1)})^{1/4} J(P; C^{(2)})^{1/4} \\ &\ll (P^{3+\nu+\varepsilon})^{1/2} \max_{\iota=1,2} J(P; C^{(\iota)})^{1/2}. \end{aligned} \tag{3.10}$$

By symmetry, there is no loss of generality in supposing that the maximum on the right hand side occurs with $\iota = 1$.

We put $B_1 = C^{(1)}$, and show by induction that for each natural number m, there exists an integral adjuvant matrix B_m of type $(2^m, r)$ for which

$$K(P; C) \ll (P^{3+\nu+\varepsilon})^{1-2^{-m}} J(P; B_m)^{2^{-m}}. \tag{3.11}$$

This bound holds when $m = 1$ as a trivial consequence of (3.10). Suppose then that the estimate (3.11) holds for $1 \leqslant m \leqslant M$. By applying Lemma 3.3, we see that there exists an integral adjuvant matrix B_{M+1} of type $(2^{M+1}, r)$ with

$$J(P; B_M) \ll (P^{3+\nu+\varepsilon})^{1/2} J(P; B_{M+1})^{1/2}.$$

Substituting this estimate into the case $m = M$ of (3.11), one confirms that the bound (3.11) holds with $m = M + 1$. The bound (3.11) consequently follows for all m by induction.

We now apply the bound just established. Let δ be any small positive number, and choose m large enough that $2^{1-m}(2-\nu) < \delta$. We have shown that an integral adjuvant matrix $B_m = (b_{ij})$ of type $(2^m, r)$ exists for which (3.11) holds. The matrix B_m is of format $(R+1) \times (2R+2)$, where $R = 2^m(r-1)$. In view of (3.5), together with the trivial estimates $|f(\alpha)| \leqslant P$ and $|g(\alpha)| \leqslant P$, we find that

$$J(P; B_m) \ll P^4 \oint |\boldsymbol{\phi}_{0,R}(\boldsymbol{\beta}) \boldsymbol{\Phi}_{R+2,2R+1}(\boldsymbol{\beta})| \, \mathrm{d}\boldsymbol{\alpha}.$$

The matrix of coefficients associated with a suitable permutation of the linear forms

$$\beta_0(\boldsymbol{\alpha}), \beta_1(\boldsymbol{\alpha}), \dots, \beta_R(\boldsymbol{\alpha}), \beta_{R+2}(\boldsymbol{\alpha}), \dots, \beta_{2R+1}(\boldsymbol{\alpha}),$$

is auxiliary of type $(2^m-1,r,r)_0$. By orthogonality, a consideration of the underlying Diophantine equations shows that $J(P;B_m) \ll P^4 I_0(P;B_m)$, and hence we deduce from Lemma 2.1 that

$$J(P;B_m) \ll P^4\left(P^{3(R+1)-2+\varepsilon}\right) = P^{3(2^m(r-1))+5+\varepsilon}.$$

By substituting the estimate just obtained into (3.11), we conclude that

$$\begin{aligned} K(P;C) &\ll \left(P^{3+\nu+\varepsilon}\right)^{1-2^{-m}}\left(P^{3(2^m(r-1))+5+\varepsilon}\right)^{2^{-m}} \\ &= P^{3r+\nu+(2-\nu)2^{-m}+\varepsilon}. \end{aligned}$$

In view of our assumed upper bound $2^{1-m}(2-\nu)<\delta$, one therefore finds that for each $\varepsilon'>0$, one has $K(P;C) \ll P^{3r+\nu+\delta/2+\varepsilon'}$. The conclusion of the theorem follows on taking $\delta=\varepsilon$ and $\varepsilon'=\frac{1}{2}\varepsilon$. $\square$

Proof when $r=2$. An application of the elementary inequality $|z_1\cdots z_n| \leqslant |z_1|^n+\ldots+|z_n|^n$ yields

$$F_l(\gamma_0)\cdots F_l(\gamma_3) \ll \sum_{0\leqslant a<b<c\leqslant 3} |F_l(\gamma_a)F_l(\gamma_b)F_l(\gamma_c)|^{4/3},$$

and hence there exist integers a, b and c with $0\leqslant a<b<c\leqslant 3$ for which

$$K_l(P;C) \ll \oint |F_l(\gamma_a)F_l(\gamma_b)F_l(\gamma_c)|^{4/3}\,\mathrm{d}\boldsymbol{\alpha}. \tag{3.12}$$

It is convenient to define

$$\Omega_h(P;C) = \oint |h(\gamma_a)h(\gamma_b)h(\gamma_c)|^4\,\mathrm{d}\boldsymbol{\alpha},$$

with h taken to be either f_0 or g. It is immediate from [12, Theorem 1.8] that

$$\Omega_g(P;C) \ll P^{6+(6\nu_2-1)/4+\varepsilon}. \tag{3.13}$$

The argument of the proof of the latter theorem also readily yields the estimate $\Omega_{f_0}(P;C) \ll P^{6+\nu_1+\varepsilon}$. In order to see this, one observes that the bound

$$\int_0^1 |f_0(\alpha)|^6\,\mathrm{d}\alpha \ll P^{3+\nu_1+\varepsilon},$$

stemming from Hua's lemma (see [17, Lemma 2.5]), can be substituted for the bound

$$\int_0^1 |f_0(\alpha)^2 g(\alpha)^4|\,\mathrm{d}\alpha \ll P^{3+\nu_2+\varepsilon}$$

underlying the proof of [12, Theorem 1.8]. In this way, one finds as in [12, equation (4.10)] that

$$\Omega_{f_0}(P;C) \ll P^{23/4+3\nu_1/2+\varepsilon} = P^{6+\nu_1+\varepsilon}.$$

In the above notation, when $l=1$, we now infer from the bound (3.12) that

$$K_1(P;C) \ll \Omega_{f_0}(P;C) \ll P^{6+\nu_1+\varepsilon}.$$

Also, applying Hölder's inequality to (3.12), we obtain via (3.13) the estimate

$$K_2(P;C) \ll \Omega_{f_0}(P;C)^{1/3}\Omega_g(P;C)^{2/3} \ll P^{6+\nu_2+\varepsilon},$$

on observing that

$$\frac{1}{3}\nu_1 + \frac{2}{3}\left(\frac{6\nu_2 - 1}{4}\right) = \nu_2.$$

This completes the proof of Theorem 3.4 in the case $r = 2$. $\square$

4. Correlation estimates

We apply Theorem 3.4 in this section to provide estimates for the correlation sums $\Xi_{s,\eta}(N; A; \mathbf{h})$. By reference to (1.1) and (1.2), we see that when $A \in \mathbb{Z}^{r\times 2r}$ is a highly non-singular matrix, and $h_i \in \mathbb{N} \cup \{0\}$, then $\Xi_{2r,\eta}(N; A; \mathbf{h})$ counts the number of integral solutions of the system

$$X_j = \Lambda_j(\mathbf{n}) \quad (1 \leqslant j \leqslant 2r), \tag{4.1}$$

with $\mathbf{n} \in \mathcal{P}(N)$, in which $X_j = x_j^3 + y_j^3 + z_j^3 - h_j$ and $x_j, y_j, z_j \in \mathbb{N}$, and none of the prime divisors of $y_j z_j$ exceed $(X_j + h_j)^{\eta/3}$. Since $X_j + h_j$ is no larger than CN, for a suitable positive constant C depending at most on the coefficients of the Λ_j, one sees that x_j, y_j, z_j are each bounded above by $P = (CN)^{1/3}$.

The system (4.1) may be written in the shape $A^T\mathbf{n} = \mathbf{X}$. It is convenient to consider a block matrix decomposition of A, say $A = (A_1, A_2)$ with A_1 and A_2 each $r \times r$ matrices, and also to write

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix},$$

with $\mathbf{X}_1$ and $\mathbf{X}_2$ each r-dimensional column vectors. Thus $\mathbf{X}_i = A_i^T\mathbf{n}$ for $i = 1, 2$. Since A is highly non-singular, the matrices A_1 and A_2 are necessarily invertible, and we deduce that

$$(A_1^{-1})^T\mathbf{X}_1 = \mathbf{n} = (A_2^{-1})^T\mathbf{X}_2.$$

Thus we find that $B'\mathbf{X} = \mathbf{0}$, where

$$B' = \left((A_1^{-1})^T, -(A_2^{-1})^T\right).$$

By applying Lemma 3.1, one sees that the matrix B' is highly non-singular if and only if $(A_1^{-1})^T$ and $(A_2^{-1})^T$ are non-singular, and all the square minors of $A_1^T(A_2^{-1})^T = (A_2^{-1}A_1)^T$ are non-singular. The non-singularity of $(A_1^{-1})^T$ and $(A_2^{-1})^T$ is immediate from that of A_1 and A_2. Likewise, the non-singularity of the square minors of $(A_2^{-1}A_1)^T$ is equivalent to the non-singularity of the square minors of $A_2^{-1}A_1$, which is a consequence of the highly non-singular nature of the block matrix (A_2, A_1), again by Lemma 3.1. We hence conclude that B' is highly non-singular. Finally, we take λ to be the least natural number with the property that $\lambda B'$ has integral entries, and define the matrix $B = (b_{ij})$ by putting $B = \lambda B'$.

At this point, we have established that $\Xi_{2r,\eta}(N; A; \mathbf{h})$ is bounded above by the number of solutions of the system of equations

$$\sum_{j=1}^{2r} b_{ij}(x_j^3 + y_j^3 + z_j^3) = H_i \quad (1 \leqslant i \leqslant r),$$

with $1 \leqslant x_j \leqslant P$ and $y_j, z_j \in \mathcal{A}(P, P^\eta)$ $(1 \leqslant j \leqslant 2r)$, in which

$$H_i = \sum_{j=1}^{2r} b_{ij} h_j.$$

Define

$$\beta_j = \sum_{i=1}^{r} b_{ij} \alpha_i \quad (1 \leqslant j \leqslant 2r).$$

Making use of the notation (3.2) with σ taken implicitly to be 0, it therefore follows from orthogonality that

$$\Xi_{2r,\eta}(N; A; \mathbf{h}) \leqslant \oint F_l(\beta_1) \cdots F_l(\beta_{2r}) e(-\boldsymbol{\alpha} \cdot \mathbf{H}) \, \mathrm{d}\boldsymbol{\alpha} \quad (l = 1, 2).$$

We note here that one should view η as being 1 in the case $l = 1$, and when $l = 2$ view η as being a positive number sufficiently small in terms of ε. An application of the triangle inequality in conjunction with Theorem 3.4 consequently reveals that $\Xi_{2r,\eta}(N; A; \mathbf{h}) \ll P^{3r+\nu_l+\varepsilon}$ $(l = 1, 2)$. Theorems 1.1 and 1.2 follow by reference to (3.3), since one has $P = O(N^{1/3})$.

5. Systems of linear equations

We turn now to the proof of Theorem 1.3. Let $C = (c_{ij})$ denote an integral $r \times s$ highly non-singular matrix with $r \geqslant 2$ and $s \geqslant 2r + 1$. We define

$$\gamma_j(\boldsymbol{\alpha}) = \sum_{i=1}^{r} c_{ij} \alpha_i \quad (1 \leqslant j \leqslant s).$$

Let N be a large positive number, and put $P = \frac{1}{2} N^{1/3}$. Let η be a positive number sufficiently small in the context of Lemma 3.2, and let σ be a positive number sufficiently small in terms of C and η. Recalling (3.1), we put $\mathfrak{f}(\alpha) = f_0(\alpha)$ and $\mathfrak{g}(\alpha) = g(\alpha)$, and for the sake of concision write $\mathfrak{g}_j = \mathfrak{g}(\gamma_j(\boldsymbol{\alpha}))$ and $\mathfrak{f}_j = \mathfrak{f}(\gamma_j(\boldsymbol{\alpha}))$. When $\mathfrak{B} \subseteq [0,1)^r$ is measurable, we then define

$$\mathcal{N}(P; \mathfrak{B}) = \int_{\mathfrak{B}} \prod_{j=1}^{s} \mathfrak{f}_j \mathfrak{g}_j^2 \, \mathrm{d}\boldsymbol{\alpha}.$$

By orthogonality, it follows from this definition that $\mathcal{N}(P; [0,1)^r)$ counts the number of integral solutions of the system

$$\sum_{j=1}^{s} c_{ij}(x_j^3 + y_j^3 + z_j^3) = 0 \quad (1 \leqslant i \leqslant r), \tag{5.1}$$

with $\sigma P < x_j, y_j, z_j \leqslant P$ and $y_j, z_j \in \mathcal{A}(P, P^\eta)$ $(1 \leqslant j \leqslant s)$. Hence we find that $\mathcal{N}(P; [0,1)^r)$ counts the solutions of the system (1.6) with each solution $\mathbf{n}$ counted with weight $\rho_\eta(n_1; P) \cdots \rho_\eta(n_s; P)$, in which $\rho_\eta(n; P)$ denotes the number of integral solutions of the equation $n = x^3 + y^3 + z^3$, with $\sigma P < x, y, z \leqslant P$ and $y, z \in \mathcal{A}(P, P^\eta)$. We aim to show that $\mathcal{N}(P; [0,1)^r) \gg (P^3)^{s-r}$.

In pursuit of the above objective, we apply the Hardy-Littlewood method. Write $L = \log\log P$, denote by $\mathfrak{N}$ the union of the intervals

$$\mathfrak{N}(q,a) = \{\alpha \in [0,1) : |q\alpha - a| \leqslant LP^{-3}\},$$

with $0 \leqslant a \leqslant q \leqslant L$ and $(a,q)=1$, and put $\mathfrak{n} = [0,1)\backslash\mathfrak{N}$. Finally, we introduce a multi-dimensional set of arcs. Let $Q = L^{10r}$, and define the narrow set of major arcs $\mathfrak{P}$ to be the union of the boxes

$$\mathfrak{P}(q,\mathbf{a}) = \{\boldsymbol{\alpha} \in [0,1)^r : |\alpha_i - a_i/q| \leqslant QP^{-3}\ (1 \leqslant i \leqslant r)\},$$

with $0 \leqslant a_i \leqslant q \leqslant Q\ (1 \leqslant i \leqslant r)$ and $(a_1,\ldots,a_r,q)=1$.

Lemma 5.1. *One has* $\mathcal{N}(P;\mathfrak{P}) \gg P^{3s-3r}$.

Proof. We begin by defining the auxiliary functions

$$S(q,a) = \sum_{r=1}^{q} e(ar^3/q) \quad \text{and} \quad v(\beta) = \int_{\sigma P}^{P} e(\beta\gamma^3)\,\mathrm{d}\gamma.$$

For $1 \leqslant j \leqslant s$, put $S_j(q,\mathbf{a}) = S(q,\gamma_j(\mathbf{a}))$ and $v_j(\boldsymbol{\beta}) = v(\gamma_j(\boldsymbol{\beta}))$, and define

$$A(q) = \sum_{\substack{a_1=1\\ (q,a_1,\ldots,a_r)=1}}^{q} \cdots \sum_{a_r=1}^{q} q^{-3s} \prod_{j=1}^{s} S_j(q,\mathbf{a})^3 \quad \text{and} \quad V(\boldsymbol{\beta}) = \prod_{j=1}^{s} v_j(\boldsymbol{\beta})^3. \tag{5.2}$$

Finally, write $\mathcal{B}(X)$ for $[-XP^{-3}, XP^{-3}]^r$, and define

$$\mathfrak{J}(X) = \int_{\mathcal{B}(X)} V(\boldsymbol{\beta})\,\mathrm{d}\boldsymbol{\beta} \quad \text{and} \quad \mathfrak{S}(X) = \sum_{1 \leqslant q \leqslant X} A(q).$$

We prove first that there exists a positive constant $\mathcal{C}$ with the property that

$$\mathcal{N}(P;\mathfrak{P}) - \mathcal{C}\mathfrak{S}(Q)\mathfrak{J}(Q) \ll P^{3s-3r}L^{-1}. \tag{5.3}$$

It follows from [18, Lemma 8.5] (see also [16, Lemma 5.4]) that there exists a positive constant $c = c(\eta)$ such that whenever $\boldsymbol{\alpha} \in \mathfrak{P}(q,\mathbf{a}) \subseteq \mathfrak{P}$, then

$$\mathfrak{g}(\gamma_j(\boldsymbol{\alpha})) - cq^{-1}S_j(q,\mathbf{a})v_j(\boldsymbol{\alpha} - \mathbf{a}/q) \ll P(\log P)^{-1/2}.$$

Under the same constraints on $\boldsymbol{\alpha}$, one finds from [17, Theorem 4.1] that

$$\mathfrak{f}(\gamma_j(\boldsymbol{\alpha})) - q^{-1}S_j(q,\mathbf{a})v_j(\boldsymbol{\alpha} - \mathbf{a}/q) \ll \log P.$$

Thus, whenever $\boldsymbol{\alpha} \in \mathfrak{P}(q,\mathbf{a}) \subseteq \mathfrak{P}$, one has

$$\prod_{j=1}^{s} \mathfrak{f}_j\mathfrak{g}_j^2 - c^{2s}q^{-3s}\prod_{j=1}^{s} S_j(q,\mathbf{a})^3 v_j(\boldsymbol{\alpha} - \mathbf{a}/q)^3 \ll P^{3s}(\log P)^{-1/2}.$$

The measure of the major arcs $\mathfrak{P}$ is $O(Q^{2r+1}P^{-3r})$, so that on integrating over $\mathfrak{P}$, we confirm the relation (5.3) with $\mathcal{C} = c^{2s}$.

We next discuss the singular integral $\mathfrak{J}(Q)$. By applying an argument paralleling that of [6] leading to equation (4.4) of that paper, one finds that

$$\mathfrak{J}(Q) \gg P^{3s-3r}. \tag{5.4}$$

Here, we make use of the hypothesis that the system (1.6) has a solution $\mathbf{n} \in (0,\infty)^s$, and hence also one with $\mathbf{n} \in (0,1)^s$. Thus, on taking σ sufficiently small, we ensure that a non-singular solution $\mathbf{n}$ of (1.6) exists with $\mathbf{n} \in (2\sigma, 1)^s$.

We turn now to the singular series $\mathfrak{S}(Q)$. It follows from [17, Theorem 4.2] that whenever $(q,a)=1$, one has $S(q,a) \ll q^{2/3}$. Given a summand $\mathbf{a}$ in the formula for $A(q)$ provided in (5.2), write $h_j = (q, \gamma_j(\mathbf{a}))$. Then we find that

$$A(q) \ll q^{-s} \sum_{\substack{a_1=1 \\ (q,a_1,\dots,a_r)=1}}^{q} \cdots \sum_{a_r=1}^{q} h_1 \cdots h_s.$$

By hypothesis, we have $s \geqslant 2r+1$. The proof of [7, Lemma 23] is therefore easily modified to show that

$$A(q) \ll \sum_{\substack{u_1 | q \\ (u_1,\dots,u_r) \ll 1}} \dots \sum_{u_r | q} q^{r-s} (u_1 \cdots u_r)^{s/r-1} \ll q^{r-s+(s-r)(1-1/r)+\varepsilon}.$$

Thus, the series $\mathfrak{S} = \lim_{X\to\infty} \mathfrak{S}(X)$ is absolutely convergent and

$$\mathfrak{S} - \mathfrak{S}(Q) \ll \sum_{q>Q} q^{-1-1/(2r)} \ll Q^{-1/(2r)} \ll L^{-1}.$$

We observe in the next step that the system (5.1) has a non-singular p-adic solution. For on taking $(x_j, y_j, z_j) = (1,-1,0)$ for each j, we solve (5.1) with the Jacobian determinant

$$\det(3c_{ij}x_j^2)_{1\leqslant i,j\leqslant r} = 3^r \det(c_{ij})_{1\leqslant i,j\leqslant r}$$

non-zero, since the first r columns of C are linearly independent. A modification of the proof of [7, Lemma 31] therefore shows that $\mathfrak{S} > 0$, whence $\mathfrak{S}(Q) = \mathfrak{S} + O(L^{-1}) > 0$. The proof of the lemma is completed by recalling (5.4) and substituting into (5.3) to obtain the lower bound

$$\mathcal{N}(P;\mathfrak{P}) \gg P^{3s-3r} + O(P^{3s-3r}L^{-1}).$$

□

Recall the definition of the major arcs $\mathfrak{M}(q,a)$ and their union $\mathfrak{M}$ from (2.13). In order to prune a wide set of major arcs down to the narrow set $\mathfrak{P}$ just considered, we introduce the auxiliary sets of arcs

$$\mathfrak{M}_j = \{\boldsymbol{\alpha} \in [0,1)^r : \gamma_j(\boldsymbol{\alpha}) \in \mathfrak{M} + \mathbb{Z}\},$$

and we put $\mathfrak{V} = \mathfrak{M}_1 \cap \ldots \cap \mathfrak{M}_s$. In addition, we define $\mathfrak{m}_j = [0,1)^r \setminus \mathfrak{M}_j$ $(1 \leqslant j \leqslant s)$, and write $\mathfrak{v} = [0,1)^r \setminus \mathfrak{V}$. Thus $\mathfrak{v} \subseteq \mathfrak{m}_1 \cup \ldots \cup \mathfrak{m}_s$. We begin with an auxiliary lemma.

Lemma 5.2. *Let δ be a fixed positive number. Then one has*

$$\int_{\mathfrak{M}} |\mathfrak{f}(\theta)\mathfrak{g}(\theta)^2|^{2+\delta}\, \mathrm{d}\theta \ll P^{3+3\delta}$$

and

$$\int_{\mathfrak{M}\setminus\mathfrak{N}} |\mathfrak{f}(\theta)\mathfrak{g}(\theta)^2|^{2+\delta}\,\mathrm{d}\theta \ll P^{3+3\delta}L^{-\delta/6}.$$

Proof. On applying a special case of [5, Lemma 9], we obtain the bound

$$\int_{\mathfrak{M}} |\mathfrak{f}(\theta)|^{2+\delta}|\mathfrak{g}(\theta)|^2\,\mathrm{d}\theta \ll P^{1+\delta},$$

and so the first conclusion follows on making use of a trivial estimate for $\mathfrak{g}(\theta)$. For the second inequality, one observes that the methods of [17, Chapter 4] show that

$$\sup_{\alpha\in\mathfrak{M}\setminus\mathfrak{N}} |\mathfrak{f}(\theta)| \ll PL^{-1/3}.$$

Thus, on making use also of a trivial estimate for $\mathfrak{g}(\theta)$, one obtains in like manner the bound

$$\begin{aligned}\int_{\mathfrak{M}\setminus\mathfrak{N}} |\mathfrak{f}(\theta)\mathfrak{g}(\theta)^2|^{2+\delta}\,\mathrm{d}\theta &\ll (PL^{-1/3})^{\delta/2}P^{2+2\delta}\int_{\mathfrak{M}\setminus\mathfrak{N}} |\mathfrak{f}(\theta)|^{2+\delta/2}|\mathfrak{g}(\theta)|^2\,\mathrm{d}\theta\\ &\ll (PL^{-1/3})^{\delta/2}P^{3+5\delta/2}.\end{aligned}$$

This completes the proof of the lemma. □

Lemma 5.3. *One has* $\mathcal{N}(P;\mathfrak{V}\setminus\mathfrak{P}) \ll P^{3s-3r}(\log L)^{-1}$.

Proof. Let $\boldsymbol{\alpha} \in \mathfrak{V}\setminus\mathfrak{P}$, and suppose temporarily that $\gamma_{j_m} \in \mathfrak{N}+\mathbb{Z}$ for r distinct indices $j_m \in [1,s]$. For each m there is a natural number $q_m \leqslant L$ having the property that $\|q_m\gamma_{j_m}\| \leqslant LP^{-3}$. With $q = q_1\cdots q_r$, one has $q \leqslant L^r$ and $\|q\gamma_{j_m}\| \leqslant L^rP^{-3}$. Next eliminating between $\gamma_{j_1},\ldots,\gamma_{j_r}$ in order to isolate $\alpha_1,\ldots,\alpha_r$, one finds that there is a positive integer κ, depending at most on (c_{ij}), such that $\|\kappa q\alpha_l\| \leqslant L^{r+1}P^{-3}$ $(1\leqslant l\leqslant r)$. Since $\kappa q \leqslant L^{r+1}$, it follows that $\boldsymbol{\alpha}\in\mathfrak{P}$, yielding a contradiction to our hypothesis that $\boldsymbol{\alpha}\in\mathfrak{V}\setminus\mathfrak{P}$. Thus $\gamma_\nu(\boldsymbol{\alpha}) \in \mathfrak{n}+\mathbb{Z}$ for at least $s-r\geqslant r+1$ of the suffices ν with $1\leqslant\nu\leqslant s$. Let $\mathcal{H}$ denote the set of all r element subsets of $\{1,2,\ldots,s\}$, and put $H=\mathrm{card}(\mathcal{H})$. Then by Hölder's inequality, we find that

$$\mathcal{N}(P;\mathfrak{V}\setminus\mathfrak{P}) \leqslant \prod_{\boldsymbol{\nu}\in\mathcal{H}} I(\boldsymbol{\nu})^{1/H}, \tag{5.5}$$

where

$$I(\boldsymbol{\nu}) = \int_{\mathfrak{V}\setminus\mathfrak{P}} \prod_{j=1}^{r} |\mathfrak{f}_{\nu_j}\mathfrak{g}_{\nu_j}^2|^{s/r}\,\mathrm{d}\boldsymbol{\alpha}.$$

When $\boldsymbol{\nu}\in\mathcal{H}$, one finds by a change of variable that

$$I(\boldsymbol{\nu}) \leqslant \int_{\mathfrak{M}^r} \prod_{j=1}^{r} |\mathfrak{f}(\beta_j)\mathfrak{g}(\beta_j)^2|^{s/r}\,\mathrm{d}\boldsymbol{\beta},$$

so that Lemma 5.2 shows that $I(\boldsymbol{\nu}) \ll P^{3s-3r}$. Further, since there exists some $\boldsymbol{\nu}\in\mathcal{H}$ such that $\gamma_{\nu_j}(\boldsymbol{\alpha}) \in \mathfrak{n}+\mathbb{Z}$ for $1\leqslant j\leqslant r$, one finds for this subset that one has the bound

$$I(\boldsymbol{\nu}) \leqslant \int_{(\mathfrak{M}\setminus\mathfrak{N})^r} \prod_{j=1}^{r} |\mathfrak{f}(\beta_j)\mathfrak{g}(\beta_j)^2|^{s/r}\,\mathrm{d}\boldsymbol{\beta} \ll P^{3s-3r}L^{-1/6}.$$

Thus we conclude from (5.5) that

$$\mathcal{N}(P;\mathfrak{V}\setminus\mathfrak{P}) \ll P^{3s-3r}L^{-1/(6H)},$$

and the conclusion of the lemma follows. □

Our final task in the application of the Hardy-Littlewood method is the analysis of the minor arcs $\mathfrak{v}$.

Lemma 5.4. *There is a positive number δ such that $\mathcal{N}(P;\mathfrak{v}) \ll P^{3s-3r-\delta}$.*

Proof. Since $\mathfrak{v} \subseteq \mathfrak{m}_1 \cup \ldots \cup \mathfrak{m}_s$, the conclusion of the lemma follows by showing that $\mathcal{N}(P;\mathfrak{m}_j) \ll P^{3s-3r-\delta}$ for $1 \leqslant j \leqslant s$. By symmetry, moreover, we may restrict attention to the case $j = s$. Suppose then that $\gamma_s(\boldsymbol{\alpha}) \in \mathfrak{m}+\mathbb{Z}$. Observe that the matrix C is highly non-singular, and thus the matrix C', in which the final $s-2r$ columns of C are deleted, is also highly non-singular. Then it follows from (3.7) and Theorem 3.4 that

$$\oint \prod_{j=1}^{2r} |\mathfrak{f}_j\mathfrak{g}_j^2|\,\mathrm{d}\boldsymbol{\alpha} \ll P^{3r+\nu_2+\varepsilon}.$$

Observe that by Weyl's inequality (see [15, Lemma 1]), one has

$$\sup_{\gamma_s(\boldsymbol{\alpha})\in\mathfrak{m}+\mathbb{Z}} |\mathfrak{f}(\gamma_s(\boldsymbol{\alpha}))| \ll P^{3/4+\varepsilon}.$$

Hence, by employing trivial estimates for $\mathfrak{f}_j$ and $\mathfrak{g}_j$ as necessary, one obtains the bound

$$\begin{aligned}\mathcal{N}(P;\mathfrak{m}_s) &\leqslant P^{3s-6r-1}\Bigl(\sup_{\gamma_s(\boldsymbol{\alpha})\in\mathfrak{m}+\mathbb{Z}} |\mathfrak{f}(\gamma_s(\boldsymbol{\alpha}))|\Bigr) \oint \prod_{j=1}^{2r} |\mathfrak{f}_j\mathfrak{g}_j^2|\,\mathrm{d}\boldsymbol{\alpha}\\ &\ll P^{3s-3r+\nu_2-1/4+\varepsilon}.\end{aligned}$$

From (3.3), we have $\nu_2 < 1/4$, and so the conclusion of the lemma now follows. □

By combining the conclusions of Lemmata 5.1, 5.3 and 5.4, we conclude that

$$\mathcal{N}(P) = \mathcal{N}(P;\mathfrak{P}) + \mathcal{N}(P;\mathfrak{V}\setminus\mathfrak{P}) + \mathcal{N}(P;\mathfrak{v}) \gg P^{3s-3r}. \tag{5.6}$$

Our final task is to remove the multiplicity of representations implicit in the definition of $\rho_\eta(n;P)$. Note that $\rho_\eta(n;P) \leqslant \rho_\eta(n)$ for each $n \in \mathbb{N}$. It is useful to introduce the set

$$\mathcal{S}_\theta(N) = \{1 \leqslant n \leqslant N : \rho_\eta(n) > N^\theta\}.$$

Lemma 5.5. *One has*

$$\sum_{n\in\mathcal{S}_\theta(N)} \rho_\eta(n) \ll N^{1+\xi-\theta+\varepsilon}.$$

Proof. In view of (1.5), one has

$$\sum_{n\in\mathcal{S}_\theta(N)} \rho_\eta(n) < N^{-\theta} \sum_{n\in\mathcal{S}_\theta(N)} \rho_\eta(n)^2 \ll N^{1+\xi-\theta+\varepsilon},$$

and the conclusion of the lemma follows. □

Let δ be a positive number, and consider the number Y_1 of solutions of the system (5.1) in which one has $\rho_\eta(x_j^3+y_j^3+z_j^3) > N^{2\xi+\delta}$ for some index j with $1\leqslant j\leqslant s$. Without loss of generality, one may assume that $j=s$. Then by orthogonality, one has

$$Y_1 \ll \sum_{n\in\mathcal{S}_{2\xi+\delta}(N)} \rho_\eta(n) \oint \biggl(\prod_{j=1}^{s-1} \mathfrak{f}_j\mathfrak{g}_j^2\biggr) e(n\gamma_s(\boldsymbol{\alpha}))\,\mathrm{d}\boldsymbol{\alpha}.$$

By the triangle inequality, Theorem 3.4 and Lemma 5.5, we thus deduce that

$$\begin{aligned} Y_1 &\ll N^{s-r-1+\xi+\varepsilon} \sum_{n\in\mathcal{S}_{2\xi+\delta}(N)} \rho_\eta(n) \\ &\ll N^{s-r+2\xi-(2\xi+\delta)+2\varepsilon} \ll N^{s-r-\delta/2}. \end{aligned}$$

Let Y_0 denote the contribution to $\mathcal{N}(P)$ arising from those solutions of (5.1) in which $\rho_\eta(x_j^3+y_j^3+z_j^3)\leqslant N^{2\xi+\delta}$ for all j. Then it follows from (5.6) that

$$Y_0 \gg N^{s-r} + O(N^{s-r-\delta/2}) \gg N^{s-r}.$$

Since Y_0 counts solutions $\mathbf{n}$ of (1.6), with each solution counted with weight at most $\rho_\eta(n_1)\cdots\rho_\eta(n_s)\leqslant (N^{2\xi+\delta})^s$, we conclude that $\Upsilon(N)\gg N^{s-r}(N^{2\xi+\delta})^{-s}$. As δ may be chosen arbitrarily small, though positive, this completes the proof of Theorem 1.3.

References

[1] A. Balog and J. Brüdern, *Sums of three cubes in three linked three-progressions*, J. reine angew. Math. **466** (1995), 45–85.

[2] J. Brüdern, K. Kawada and T. D. Wooley, *Additive representation in thin sequences, I: Waring's problem for cubes*, Ann. Sci. École Norm. Sup. (4) **34** (2001), no. 4, 471–501.

[3] J. Brüdern and T. D. Wooley, *Hua's lemma and simultaneous diagonal equations*, Bull. London Math. Soc. **34** (2002), no. 3, 279–283.

[4] J. Brüdern and T. D. Wooley, *The paucity problem for certain pairs of diagonal equations*, Quart. J. Math. **54** (2003), no. 1, 41–48.

[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*, submitted; arXiv:1304:5165.

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

[8] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.

[9] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum' (VI): Further researches in Waring's problem*, Math. Z. **23** (1925), no. 1, 1–37.

[10] D. R. Heath-Brown, *The circle method and diagonal cubic forms*, Phil. Trans. Roy. Soc. London Ser. A **356** (1998), no. 1738, 673–699.

[11] C. Hooley, *On Hypothesis K^* in Waring's problem*, Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), pp. 175–185, Cambridge University Press, 1997.

[12] K. Kawada and T. D. Wooley, *Davenport's method and slim exceptional sets: the asymptotic formulae in Waring's problem*, Mathematika **56** (2010), no. 2, 305–321.

[13] L. Matthiesen, *Linear correlations amongst numbers represented by positive definite binary quadratic forms*, Acta Arith. **154** (2012), no. 3, 235–306.

[14] L. Matthiesen, *Correlations of representation functions of binary quadratic forms*, Acta Arith. **158** (2013), no. 3, 245–252.

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

[16] R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), no. 1–2, 1–71.

[17] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd edition, Cambridge University Press, Cambridge, 1997.

[18] T. D. Wooley, *On simultaneous additive equations, II*, J. reine angew. Math. **419** (1991), 141–198.

[19] T. D. Wooley, *Breaking classical convexity in Waring's problem: sums of cubes and quasi-diagonal behaviour*, Invent. Math. **122** (1995), no. 3, 421–451.

[20] T. D. Wooley, *Sums of three cubes*, Mathematika **47** (2000), no. 1–2, 53–61.

Mathematisches Institut, Bunsenstrasse 3–5, D-37073 Göttingen, Germany
E-mail address: bruedern@uni-math.gwdg.de

School of Mathematics, University of Bristol, University Walk, Clifton, Bristol BS8 1TW, United Kingdom
E-mail address: matdw@bristol.ac.uk