PAUCITY PROBLEMS AND SOME RELATIVES OF VINOGRADOV'S MEAN VALUE THEOREM

TREVOR D. WOOLEY

Abstract. When $k \geqslant 4$ and $0 \leqslant d \leqslant (k-2)/4$, we consider the system of Diophantine equations

$$x_1^j + \ldots + x_k^j = y_1^j + \ldots + y_k^j \quad (1 \leqslant j \leqslant k,\, j \neq k-d).$$

We show that in this cousin of a Vinogradov system, there is a paucity of non-diagonal positive integral solutions. Our quantitative estimates are particularly sharp when $d = o(k^{1/4})$.

1. Introduction

Recent progress on Vinogradov's mean value theorem has resolved the main conjecture in the subject. Thus, writing $J_{s,k}(X)$ for the number of integral solutions of the system of equations

$$x_1^j + \ldots + x_s^j = y_1^j + \ldots + y_s^j \quad (1 \leqslant j \leqslant k), \tag{1.1}$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$, it is now known that whenever $\varepsilon > 0$, one has

$$J_{s,k}(X) \ll X^{s+\varepsilon} + X^{2s-k(k+1)/2} \tag{1.2}$$

(see [1] or [13, 14]). Denote by $T_s(X)$ the number of s-tuples $\mathbf{x}$ and $\mathbf{y}$ in which $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$, and $(x_1, \ldots, x_s)$ is a permutation of $(y_1, \ldots, y_s)$. Thus $T_s(X) = s!X^s + O(X^{s-1})$. A conjecture going beyond the main conjecture (1.2) asserts that when $1 \leqslant s < \frac{1}{2}k(k+1)$, one should have

$$J_{s,k}(X) = T_s(X) + o(X^s). \tag{1.3}$$

This conclusion is essentially trivial for $1 \leqslant s \leqslant k$, in which circumstances one has the definitive statement $J_{s,k}(X) = T_s(X)$. When $s \geqslant k+2$, meanwhile, the conclusion (1.3) is at present far beyond our grasp. This leaves the special case $s = k+1$. Here, one has the asymptotic relation

$$J_{k+1,k}(X) = T_{k+1}(X) + O(X^{\sqrt{4k+5}}) \tag{1.4}$$

established in joint work of the author with Vaughan [10, Theorem 1]. An analogous conclusion is available when the equation of degree $k-1$ in the system (1.1) is removed, but in no other close relative of Vinogradov's mean value theorem has such a conclusion been obtained hitherto. Our purpose in this paper is to derive estimates of strength paralleling (1.4) in systems of the shape (1.1) in which a large degree equation is removed.

2010 *Mathematics Subject Classification.* 11D45, 11P05.

Key words and phrases. Paucity, Vinogradov's mean value theorem.

In order to describe our conclusions, we must introduce some notation. When $k \geqslant 2$ and $0 \leqslant d < k$, we denote by $I_{k,d}(X)$ the number of integral solutions of the system of equations

$$x_1^j + \ldots + x_k^j = y_1^j + \ldots + y_k^j \quad (1 \leqslant j \leqslant k,\, j \neq k - d), \tag{1.5}$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant k)$. Also, when $k \geqslant 3$ and $d \geqslant 0$, we define the exponent

$$\gamma_{k,d} = \min_{2 \leqslant r \leqslant k} \left(r + \frac{k}{r} + \sum_{l=1}^{r} \max\{d - l + 1, 0\} \right). \tag{1.6}$$

Theorem 1.1. *Suppose that $k \geqslant 3$ and $0 \leqslant d < k/2$. Then, for each $\varepsilon > 0$, one has*

$$I_{k,d}(X) - T_k(X) \ll X^{\gamma_{k,d}+\varepsilon}.$$

When k is large and d is small compared to k, the conclusion of this theorem provides strikingly powerful paucity estimates.

Corollary 1.2. *Suppose that $d \leqslant \sqrt{k}$. Then*

$$I_{k,d}(X) - T_k(X) \ll X^{\sqrt{4k+1}+d(d+1)/2}.$$

In particular, when $d = o(k^{1/4})$, one has

$$I_{k,d}(X) = T_k(X) + O(X^{(2+o(1))\sqrt{k}}).$$

Although for larger values of d our paucity estimates become weaker, they remain non-trivial whenever $d < (k-2)/4$.

Corollary 1.3. *Provided that $d \geqslant 1$ and $k \geqslant 4d + 3$, one has*

$$I_{k,d}(X) = k!X^k + O(X^{k-1/2}).$$

Moreover, when $1 \leqslant d \leqslant k/4$, one has

$$I_{k,d}(X) - T_k(X) \ll X^{\sqrt{4k(d+1)+(d+1)^2}},$$

so that whenever η is small and positive, and $1 \leqslant d \leqslant \eta^2 k$, then

$$I_{k,d}(X) = T_k(X) + O(X^{3\eta k}).$$

Previous work on this problem is confined to the two cases considered by Hua [4, Lemmata 5.2 and 5.4]. Thus, the asymptotic formula (1.4) derived by the author jointly with Vaughan [10, Theorem 1] is tantamount to the case $d = 0$ of Theorem 1.1. Meanwhile, it follows from [10, Theorem 2] that

$$I_{k,1}(X) = T_k(X) + O(X^{\gamma_{k,1}-1+\varepsilon}),$$

and the error term here is slightly sharper than that provided by the case $d = 1$ of Theorem 1.1. The conclusion of Theorem 1.1 is new whenever $d \geqslant 2$. It would be interesting to derive analogues of Theorem 1.1 in which more than one equation is removed from the Vinogradov system (1.1), or indeed to derive analogues in which the number of variables is increased and yet one is able nonetheless to confirm the paucity of non-diagonal solutions. We have more to say on such matters in §5 of this paper. For now, we confine ourselves to remarking that when many, or even most, lower degree equations are removed, then approaches based on the determinant method are available. Consider, for example, natural numbers $d_1, \ldots, d_k$ with $1 \leqslant d_1 < d_2 < \ldots < d_k$ and $d_k \geqslant 2s-1$. Also, denote by $M_{\mathbf{d},s}(X)$ the number of integral solutions of the system of equations

$$x_1^{d_j} + \ldots + x_s^{d_j} = y_1^{d_j} + \ldots + y_s^{d_j} \quad (1 \leqslant j \leqslant k),$$

with $1 \leqslant x_i, y_i \leqslant X$ $(1 \leqslant i \leqslant s)$. Then it follows from [7, Theorem 5.2] that whenever $d_1 \cdots d_k \geqslant (2s-k)^{4s-2k}$, one has

$$M_{\mathbf{d},s}(X) = s!X^s + O(X^{s-1/2}).$$

The proof of Theorem 1.1, in common with our earlier treatment in [10] of the Vinogradov system (1.1), is based on the application of multiplicative polynomial identities amongst variables in pursuit of parametrisations that these days would be described as being of torsorial type. The key innovation of [10] was to relate not merely two product polynomials, but instead $r \geqslant 2$ such polynomials, leading to a decomposition of the variables into $(k+1)^r$ parameters. Large numbers of these parameters may be determined via divisor function estimates, and thereby one obtains powerful bounds for the difference $J_{k+1,k}(X) - T_{k+1}(X)$. In the present situation, the polynomial identities are more novel, and sacrifices must be made in order to bring an analogous plan to fruition. Nonetheless, when $d < k/2$, the kind of multiplicative relations of [10] may still be derived in a useful form.

This paper is organised as follows. We begin in §2 of this paper by deriving the polynomial identities required for our subsequent analysis. In §3 we refine this infrastructure so that appropriate multiplicative relations are obtained involving few auxiliary variables. A complication for us here is the problem of bounding the number of choices for these auxiliary variables, since they are of no advantage to us in the ensuing analysis of multiplicative relations. In §4, we exploit the multiplicative relations by extracting common divisors between tuples of variables, following the path laid down in our earlier work [10] joint with Vaughan. This leads to the proof of Theorem 1.1. Finally, in §5, we discuss the corollaries to Theorem 1.1 and consider also refinements and potential generalisations of our main results.

Our basic parameter is X, a sufficiently large positive number. Whenever ε appears in a statement, either implicitly or explicitly, we assert that the statement holds for each $\varepsilon > 0$. In this paper, implicit constants in Vinogradov's notation $\ll$ and $\gg$ may depend on ε, k, and s. We make frequent use of vector notation in the form $\mathbf{x} = (x_1, \ldots, x_r)$. Here, the dimension r depends on the course of the argument. We also write $(a_1, \ldots, a_s)$ for the greatest common divisor of the integers $a_1, \ldots, a_s$. Any ambiguity between ordered s-tuples and corresponding greatest common divisors will be easily resolved by context. Finally, as usual, we write $e(z)$ for $e^{2\pi iz}$.

Acknowledgements: The author's work is supported by NSF grant DMS-2001549 and the Focused Research Group grant DMS-1854398. The author is grateful to the referee of this paper for their time and attention.

2. Polynomial identities

We begin by introducing the power sum polynomials

$$s_j(\mathbf{z}) = z_1^j + \ldots + z_k^j \quad (1 \leqslant j \leqslant k).$$

On recalling (1.5), we see that $I_{k,d}(X)$ counts the number of integral solutions of the system of equations

$$\left.\begin{aligned} s_j(\mathbf{x}) &= s_j(\mathbf{y}) \quad (1 \leqslant j \leqslant k,\, j \neq k-d) \\ s_{k-d}(\mathbf{x}) &= s_{k-d}(\mathbf{y}) + h, \end{aligned}\right\} \tag{2.1}$$

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$ and $|h| \leqslant kX^{k-d}$. Our first task is to reinterpret this system in terms of elementary symmetric polynomials, so that our first multiplicative relations may be extracted.

The elementary symmetric polynomials $\sigma_j(\mathbf{z}) \in \mathbb{Z}[z_1, \ldots, z_k]$ may be defined by means of the generating function identity

$$1 + \sum_{j=1}^k \sigma_j(\mathbf{z})(-t)^j = \prod_{i=1}^k (1 - tz_i).$$

Since

$$\sum_{i=1}^k \log(1 - tz_i) = -\sum_{j=1}^\infty s_j(\mathbf{z})\frac{t^j}{j},$$

we deduce that

$$1 + \sum_{j=1}^k \sigma_j(\mathbf{z})(-t)^j = \exp\biggl(-\sum_{j=1}^\infty s_j(\mathbf{z})\frac{t^j}{j}\biggr).$$

When $n \geqslant 1$, the formula

$$\sigma_n(\mathbf{z}) = (-1)^n \sum_{\substack{m_1+2m_2+\ldots+nm_n=n \\ m_i \geqslant 0}} \prod_{i=1}^n \frac{(-s_i(\mathbf{z}))^{m_i}}{i^{m_i} m_i!} \tag{2.2}$$

then follows via an application of Faà di Bruno's formula. By convention, we put $\sigma_0(\mathbf{z}) = 1$. We refer the reader to [5, equation (2.14′)] for a self-contained account of the relation (2.2).

Suppose now that $0 \leqslant d < k/2$, and that the integers $\mathbf{x}, \mathbf{y}, h$ satisfy (2.1). When $1 \leqslant n < k - d$, it follows from (2.2) that

$$\sigma_n(\mathbf{x}) = (-1)^n \sum_{\substack{m_1+2m_2+\ldots+nm_n=n \\ m_i \geqslant 0}} \prod_{i=1}^n \frac{(-s_i(\mathbf{y}))^{m_i}}{i^{m_i} m_i!} = \sigma_n(\mathbf{y}). \tag{2.3}$$

When $k - d \leqslant n \leqslant k$, on the other hand, we instead obtain the relation

$$\sigma_n(\mathbf{x}) = (-1)^n \sum_{\substack{m_1+2m_2+\ldots+nm_n=n \\ m_i \geqslant 0}} \frac{(-s_{k-d}(\mathbf{y}) - h)^{m_{k-d}}}{(k-d)^{m_{k-d}} m_{k-d}!} \prod_{\substack{1 \leqslant i \leqslant n \\ i \neq k-d}} \frac{(-s_i(\mathbf{y}))^{m_i}}{i^{m_i} m_i!}.$$

Since $d < k/2$, the summation condition on $\mathbf{m}$ ensures that $m_{k-d} \in \{0,1\}$. Thus, by isolating the term in which $m_{k-d} = 1$, we see that

$$\sigma_n(\mathbf{x}) = \sigma_n(\mathbf{y}) + h\psi_n(\mathbf{y}), \tag{2.4}$$

where by (2.2),

$$\begin{aligned}\psi_n(\mathbf{y}) &= \frac{(-1)^{n+1}}{k-d} \sum_{\substack{m_1+2m_2+\ldots+(n-k+d)m_{n-k+d}=n-k+d\\ m_i\geqslant 0}} \prod_{i=1}^{n-k+d} \frac{(-s_i(\mathbf{y}))^{m_i}}{i^{m_i} m_i!}\\ &= \frac{(-1)^{k-d+1}}{k-d}\sigma_{n-k+d}(\mathbf{y}).\end{aligned}$$

We deduce from (2.3) and (2.4) that

$$\begin{aligned}\prod_{i=1}^k (t-x_i) - \prod_{i=1}^k (t-y_i) &= (-1)^k \sum_{n=0}^k (\sigma_n(\mathbf{x}) - \sigma_n(\mathbf{y}))(-t)^{k-n}\\ &= (-1)^{d-1}\frac{h}{k-d}\sum_{m=0}^d \sigma_m(\mathbf{y})(-t)^{d-m}. \end{aligned} \tag{2.5}$$

Define the polynomial

$$\tau_d(\mathbf{y};w) = (-1)^{d-1}\sum_{m=0}^d \sigma_m(\mathbf{y})(-w)^{d-m}. \tag{2.6}$$

Then we deduce from (2.5) that for $1 \leqslant j \leqslant k$, one has the relation

$$(k-d)\prod_{i=1}^k (y_j - x_i) = \tau_d(\mathbf{y};y_j)h. \tag{2.7}$$

By comparing the relation (2.7) with $j = s$ and $j = t$ for two distinct indices s and t satisfying $1 \leqslant s < t \leqslant k$, it is apparent that

$$\tau_d(\mathbf{y};y_t)\prod_{i=1}^k (y_s - x_i) = \tau_d(\mathbf{y};y_s)\prod_{i=1}^k (y_t - x_i). \tag{2.8}$$

Furthermore, by applying the relations (2.3), we see that $\sigma_m(\mathbf{y}) = \sigma_m(\mathbf{x})$ for $1 \leqslant m \leqslant d$, and thus it is a consequence of (2.6) that

$$\tau_d(\mathbf{y};y_j) = \tau_d(\mathbf{x};y_j) \quad (1 \leqslant j \leqslant k). \tag{2.9}$$

We therefore deduce from (2.8) that for $1 \leqslant s < t \leqslant k$, one has

$$\tau_d(\mathbf{x};y_t)\prod_{i=1}^k (y_s - x_i) = \tau_d(\mathbf{x};y_s)\prod_{i=1}^k (y_t - x_i). \tag{2.10}$$

These are the multiplicative relations that provide the foundation for our analysis. One additional detail shall detain us temporarily, however, for to be useful we must ensure that all of the factors on left and right hand sides of (2.8) and (2.10) are non-zero.

Suppose temporarily that there are indices l and m with $1 \leqslant l, m \leqslant k$ for which $x_l = y_m$. By relabelling variables, if necessary, we may suppose that $l = m = k$, and then it follows from (2.1) that

$$x_1^j + \ldots + x_{k-1}^j = y_1^j + \ldots + y_{k-1}^j \quad (1 \leqslant j \leqslant k,\ j \neq k-d).$$

There are $k-1$ equations here in $k-1$ pairs of variables x_i, y_i, and thus it follows from [9] that $(x_1, \ldots, x_{k-1})$ is a permutation of $(y_1, \ldots, y_{k-1})$. We may therefore conclude that in the situation contemplated at the beginning of this paragraph, the solution $\mathbf{x}, \mathbf{y}$ of (2.1) is counted by $T_k(X)$, with $(x_1, \ldots, x_k)$ a permutation of $(y_1, \ldots, y_k)$. In particular, in any solution $\mathbf{x}, \mathbf{y}$ of (2.1) counted by $I_{k,d}(X) - T_k(X)$, it follows that $x_l = y_m$ for no indices l and m satisfying $1 \leqslant l, m \leqslant k$. In view of (2.7) and (2.9), such solutions also satisfy the conditions

$$h \neq 0 \quad \text{and} \quad \tau_d(\mathbf{y}; y_j) = \tau_d(\mathbf{x}; y_j) \neq 0 \quad (1 \leqslant j \leqslant k). \tag{2.11}$$

We summarise the deliberations of this section in the form of a lemma.

Lemma 2.1. *Suppose that* $\mathbf{x}, \mathbf{y}$ *is a solution of the Diophantine system* (2.1) *counted by* $I_{k,d}(X) - T_k(X)$*. Then the relations* (2.8), (2.10) *and* (2.11) *hold.*

3. Reduction to efficient multiplicative relations

We seek to estimate the number $I_{k,d}(X) - T_k(X)$ of solutions of the system (2.1), with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$ and $|h| \leqslant kX^{k-d}$, for which $(x_1, \ldots, x_k)$ is not a permutation of $(y_1, \ldots, y_k)$. We divide these solutions into two types according to a parameter r with $1 < r \leqslant k$. Let $V_{1,r}(X)$ denote the number of such solutions in which there are fewer than r distinct values amongst $x_1, \ldots, x_k$, and likewise fewer than r distinct values amongst $y_1, \ldots, y_k$. Also, let $V_{2,r}(X)$ denote the corresponding number of solutions in which there are either at least r distinct values amongst $x_1, \ldots, x_k$, or at least r distinct values amongst $y_1, \ldots, y_k$. Then one has

$$I_{k,d}(X) - T_k(X) = V_{1,r}(X) + V_{2,r}(X). \tag{3.1}$$

The solutions counted by $V_{1,r}(X)$ are easily handled via an expedient argument of circle method flavour.

Lemma 3.1. *One has* $V_{1,r}(X) \ll X^{r-1}$*.*

Proof. It is convenient to introduce the exponential sum

$$f(\boldsymbol{\alpha}) = \sum_{1 \leqslant x \leqslant X} e\Biggl(\sum_{\substack{1 \leqslant j \leqslant k \\ j \neq k-d}} \alpha_j x^j \Biggr).$$

In a typical solution $\mathbf{x}, \mathbf{y}$ of (2.1) counted by $V_{1,r}(X)$, we may relabel indices in such a manner that $x_j \in \{x_1, \ldots, x_{r-1}\}$ for $1 \leqslant j \leqslant k$, and likewise $y_j \in \{y_1, \ldots, y_{r-1}\}$ for $1 \leqslant j \leqslant k$. On absorbing combinatorial factors into the constant implicit in the notation of Vinogradov, therefore, we discern via orthogonality that there are integers a_i, b_i $(1 \leqslant i \leqslant r-1)$, with $1 \leqslant a_i, b_i \leqslant k$, for which one has

$$V_{1,r}(X) \ll \int_{[0,1)^{k-1}} \left(\prod_{i=1}^{r-1} f(a_i \boldsymbol{\alpha}) f(-b_i \boldsymbol{\alpha}) \right) \mathrm{d}\boldsymbol{\alpha}.$$

An application of Hölder's inequality shows that

$$V_{1,r}(X) \ll \prod_{i=1}^{r-1} I(a_i)^{1/(2r-2)} I(b_i)^{1/(2r-2)},$$

where we write

$$I(c) = \int_{[0,1)^{k-1}} |f(c\boldsymbol{\alpha})|^{2r-2} \, \mathrm{d}\boldsymbol{\alpha}.$$

Thus, by making a change of variables, we discern that

$$V_{1,r}(X) \ll \int_{[0,1)^{k-1}} |f(\boldsymbol{\alpha})|^{2r-2} \, \mathrm{d}\boldsymbol{\alpha}.$$

By orthogonality, the latter mean value counts the integral solutions of the system

$$x_1^j + \ldots + x_{r-1}^j = y_1^j + \ldots + y_{r-1}^j \quad (1 \leqslant j \leqslant k,\ j \neq k-d),$$

with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$. Since the number of equations here is $k-1$, and the number of pairs of variables is $r-1 \leqslant k-1$, it follows from [9] that $(x_1, \ldots, x_{r-1})$ is a permutation of $(y_1, \ldots, y_{r-1})$, and hence we deduce that

$$V_{1,r}(X) \ll T_{r-1}(X) \sim (r-1)! X^{r-1}.$$

This establishes the upper bound claimed in the statement of the lemma. □

We next consider the solutions $\mathbf{x}, \mathbf{y}, h$ of the system (2.1) counted by $V_{2,r}(X)$. Here, by taking advantage of the symmetry between $\mathbf{x}$ and $\mathbf{y}$, and if necessary relabelling indices, we may suppose that $y_1, \ldots, y_r$ are distinct. Suppose temporarily that the integers y_t and $x_i - y_t$ have been determined for $1 \leqslant i \leqslant k$ and $1 \leqslant t \leqslant r$. It follows that y_t and x_i are determined for $1 \leqslant i \leqslant k$ and $1 \leqslant t \leqslant r$, and hence also that the coefficients $\sigma_m(\mathbf{x})$ of the polynomial $\tau_d(\mathbf{x}; w)$ are fixed for $0 \leqslant m \leqslant d$. The integers y_s for $r < s \leqslant k$ may consequently be determined from the polynomial equations (2.10) with $t = 1$. Here, it is useful to observe that with y_1 and $x_1, \ldots, x_k$ already fixed, and all the factors on the left and right hand side of (2.10) non-zero, the equation (2.10) becomes a polynomial in the single variable y_s. On the left hand side one has a polynomial of degree k, whilst on the right hand side the polynomial has degree $d = \deg_y(\tau_d(\mathbf{x}; y)) < k$. Thus y_s is determined by a polynomial of degree k to which there are at most k solutions. Given fixed choices for y_t and $x_i - y_t$ for $1 \leqslant i \leqslant k$ and $1 \leqslant t \leqslant r$, therefore, there are $O(1)$ possible choices for $y_{r+1}, \ldots, y_k$.

Let $M_r(X; \mathbf{y})$ denote the number of integral solutions $\mathbf{x}$ of the system of equations (2.10) $(1 \leqslant s < t \leqslant r)$, satisfying $1 \leqslant \mathbf{x} \leqslant X$, wherein $\mathbf{y} =$

$(y_1,\dots,y_r)$ is fixed with $1\leqslant \mathbf{y}\leqslant X$ and satisfies (2.11). Then it follows from the above discussion in combination with Lemma 2.1 that

$$V_{2,r}(X)\ll X^r\max_{\mathbf{y}} M_r(X;\mathbf{y}), \tag{3.2}$$

in which the maximum is taken over distinct $y_1,\dots,y_r$ with $1\leqslant \mathbf{y}\leqslant X$.

Consider fixed values of $y_1,\dots,y_r$ with $1\leqslant y_i\leqslant X$ $(1\leqslant i\leqslant r)$. We write $N_r(X;\mathbf{y})$ for the number of r-tuples

$$(\tau_d(y_1,\dots,y_k;y_1),\dots,\tau_d(y_1,\dots,y_k;y_r)), \tag{3.3}$$

with $1\leqslant y_j\leqslant X$ $(r<j\leqslant k)$. It is apparent from (2.6) and (2.11) that in each such r-tuple, one has

$$1\leqslant |\tau_d(\mathbf{y};y_j)|\ll X^d, \tag{3.4}$$

and thus a trivial estimate yields the bound

$$N_r(X;\mathbf{y})\ll X^{rd}. \tag{3.5}$$

On the other hand, we may consider the number of d-tuples

$$(\sigma_1(y_1,\dots,y_k),\dots,\sigma_d(y_1,\dots,y_k)),$$

with $1\leqslant y_j\leqslant X$ $(1\leqslant j\leqslant k)$. Since $|\sigma_m(\mathbf{y})|\ll X^m$ $(1\leqslant m\leqslant d)$, the number of such d-tuples is plainly $O(X^{d(d+1)/2})$. Recall that $\sigma_0(\mathbf{y})=1$. Then for each fixed choice of this d-tuple, and for each fixed index j, it follows from (2.6) that the value of $\tau_d(y_1,\dots,y_k;y_j)$ is determined. We therefore infer that

$$N_r(X;\mathbf{y})\ll X^{d(d+1)/2}. \tag{3.6}$$

These simple estimates are already sufficient for many purposes. However, by working harder, one may obtain an estimate that is oftentimes superior to both (3.5) and (3.6). This we establish in Lemma 3.3 below. For the time being we choose not to interrupt our main narrative, and instead explain how bounds for $N_r(X;\mathbf{y})$ may be applied to estimate $V_{2,r}(X)$.

When $1\leqslant j\leqslant r$, we substitute

$$u_{0j}=\tau_d(\mathbf{x};y_j)^{-1}\prod_{i=1}^r \tau_d(\mathbf{x};y_i). \tag{3.7}$$

Observe that there are at most $N_r(X;\mathbf{y})$ distinct values for the integral r-tuple $(u_{01},\dots,u_{0r})$. Moreover, in any such r-tuple it follows from (3.4) that $1\leqslant |u_{0j}|\ll X^{d(r-1)}$. There is consequently a positive integer $C=C(k)$ with the property that, in any solution $\mathbf{x},\mathbf{y}$ counted by $M_r(X;\mathbf{y})$, one has $1\leqslant |u_{0j}|\leqslant CX^{d(r-1)}$.

Next we substitute

$$u_{ij}=x_i-y_j\quad (1\leqslant i\leqslant k,\, 1\leqslant j\leqslant r).$$

Then from (2.10) we see that $M_r(X;\mathbf{y})$ is bounded above by the number of integral solutions of the system

$$\prod_{i_1=0}^k u_{i_1 1}=\prod_{i_2=0}^k u_{i_2 2}=\ldots=\prod_{i_r=0}^k u_{i_r r}, \tag{3.8}$$

with

$$y_1 + u_{i1} = y_2 + u_{i2} = \ldots = y_r + u_{ir} \quad (1 \leqslant i \leqslant k), \tag{3.9}$$

$$1 \leqslant |u_{ij}| \leqslant X \quad (1 \leqslant i \leqslant k,\, 1 \leqslant j \leqslant r), \tag{3.10}$$

and with u_{0j} given by (3.7) for $1 \leqslant j \leqslant r$. Denote by $W(X; \mathbf{y}, \mathbf{u}_0)$ the number of integral solutions of the system (3.8) subject to (3.9) and (3.10). Then on recalling (3.2), we may summarise our deliberations thus far concerning $V_{2,r}(X)$ as follows.

Lemma 3.2. *One has*

$$V_{2,r}(X) \ll X^r \max_{\mathbf{y}} \left(N_r(X; \mathbf{y}) \max_{\mathbf{u}_0} W(X; \mathbf{y}, \mathbf{u}_0) \right),$$

where the maximum with respect to $\mathbf{y} = (y_1, \ldots, y_r)$ *is taken over* $y_1, \ldots, y_r$ *distinct with* $1 \leqslant y_j \leqslant X$ $(1 \leqslant j \leqslant r)$*, and the maximum over* r*-tuples* $\mathbf{u}_0 = (u_{01}, \ldots, u_{0r})$ *is taken over*

$$1 \leqslant |u_{0j}| \leqslant CX^{d(r-1)} \quad (1 \leqslant j \leqslant r).$$

Before fulfilling our commitment to establish an estimate for $N_r(X; \mathbf{y})$ sharper than the pedestrian bounds already obtained, we introduce the exponent

$$\theta_{d,r} = \sum_{l=1}^{r} \max\{d - l + 1, 0\}. \tag{3.11}$$

Lemma 3.3. *Let* d *and* r *be non-negative integers and let* $C \geqslant 1$ *be fixed. Also, let*

$$\mathcal{A}_d = \{(a_0, a_1, \ldots, a_d) \in \mathbb{Z}^{d+1} : |a_l| \leqslant CX^{d-l} \ (0 \leqslant l \leqslant d)\}.$$

Finally, when $\mathbf{a} \in \mathcal{A}_d$*, define*

$$f_{\mathbf{a}}(t) = a_0 + a_1 t + \ldots + a_d t^d.$$

Suppose that $y_1, \ldots, y_r$ *are fixed integers with* $1 \leqslant y_i \leqslant X$ $(1 \leqslant i \leqslant r)$*. Then one has*

$$\text{card}\{f_{\mathbf{a}}(y_i) : \mathbf{a} \in \mathcal{A}_d \text{ and } 1 \leqslant i \leqslant r\} \ll X^{\theta_{d,r}}.$$

Proof. We proceed by induction on d. Note first that when $d = 0$, the polynomials $f_{\mathbf{a}}(t)$ are necessarily constant with $|a_0| \leqslant C$, and thus

$$\text{card}\{f_{\mathbf{a}}(y_i) : \mathbf{a} \in \mathcal{A}_0 \text{ and } 1 \leqslant i \leqslant r\} \leqslant (2C+1)^r \ll 1.$$

Since $\theta_{0,r} = 0$, the conclusion of the lemma follows for $d = 0$. Observe also that when $r = 0$ the conclusion of the lemma is trivial, for then one has $\theta_{d,0} = 0$ and the set of values in question is empty.

Having established the base of the induction, we proceed under the assumption that the conclusion of the lemma holds whenever $d < D$, for some integer D with $D \geqslant 1$. In view of the discussion of the previous paragraph, we may now restrict attention to the situation with $d = D \geqslant 1$ and $r \geqslant 1$. Since $1 \leqslant y_r \leqslant X$ and y_r is fixed, we see that whenever $\mathbf{a} \in \mathcal{A}_D$ one has

$$|f_{\mathbf{a}}(y_r)| \leqslant |a_0| + |a_1|y_r + \ldots + |a_D|y_r^D \leqslant (D+1)CX^D. \tag{3.12}$$

Put

$$g_{\mathbf{a}}(y_r,t)=\frac{f_{\mathbf{a}}(y_r)-f_{\mathbf{a}}(t)}{y_r-t},\tag{3.13}$$

so that

$$g_{\mathbf{a}}(y_r,t)=\sum_{l=1}^{D}a_l(t^{l-1}+t^{l-2}y_r+\ldots+y_r^{l-1}).$$

Then one sees that whenever $\mathbf{a}\in\mathcal{A}_D$, one may write

$$g_{\mathbf{a}}(y_r,t)=F_{\mathbf{b}}(t),\tag{3.14}$$

where

$$F_{\mathbf{b}}(t)=b_0+b_1t+\ldots+b_{D-1}t^{D-1},$$

and, for $0\leqslant l\leqslant D-1$, one has

$$|b_l|\leqslant|a_{l+1}|+|a_{l+2}|y_r+\ldots+|a_D|y_r^{D-l-1}\leqslant CDX^{D-l-1}.$$

Put

$$\mathcal{B}_{D-1}=\{(b_0,b_1,\ldots,b_{D-1})\in\mathbb{Z}^D:|b_l|\leqslant CDX^{D-1-l}\ (0\leqslant l\leqslant D-1)\}.$$

Then the inductive hypothesis for $d=D-1$ implies that

$$\text{card}\{F_{\mathbf{b}}(y_i):\mathbf{b}\in\mathcal{B}_{D-1}\text{ and }1\leqslant i\leqslant r-1\}\ll X^{\theta_{D-1,r-1}}.\tag{3.15}$$

On recalling (3.13) and (3.14), we see that

$$f_{\mathbf{a}}(y_i)=f_{\mathbf{a}}(y_r)-(y_r-y_i)F_{\mathbf{b}}(y_i)\quad(1\leqslant i\leqslant r-1).$$

The values of y_i $(1\leqslant i\leqslant r-1)$ are fixed, and by (3.15) there are $O(X^{\theta_{D-1,r-1}})$ possible choices for $F_{\mathbf{b}}(y_i)$ $(1\leqslant i\leqslant r-1)$. Then for each fixed choice of $f_{\mathbf{a}}(y_r)$, there are $O(X^{\theta_{D-1,r-1}})$ choices available for $f_{\mathbf{a}}(y_i)$ $(1\leqslant i\leqslant r-1)$. We therefore deduce from (3.12) that

$$\text{card}\{f_{\mathbf{a}}(y_i):\mathbf{a}\in\mathcal{A}_D\text{ and }1\leqslant i\leqslant r\}\ll X^D\cdot X^{\theta_{D-1,r-1}}.$$

Since, from (3.11), one has

$$\begin{aligned}\theta_{D-1,r-1}+D&=D+\sum_{l=1}^{r-1}\max\{(D-1)-l+1,0\}\\&=\sum_{l=1}^{r}\max\{D-l+1,0\}=\theta_{D,r},\end{aligned}$$

we find that

$$\text{card}\{f_{\mathbf{a}}(y_i):\mathbf{a}\in\mathcal{A}_D\text{ and }1\leqslant i\leqslant r\}\ll X^{\theta_{D,r}}.$$

The inductive hypothesis therefore follows for $d=D$ and all values of r. The conclusion of the lemma consequently follows by induction. □

On recalling (2.6), a brief perusal of (3.3) and the definition of $N_r(X;\mathbf{y})$ leads from Lemma 3.3 to the estimate $N_r(X;\mathbf{y})\ll X^{\theta_{d,r}}$. We may therefore conclude this section with the following upper bound for $I_{k,d}(X)-T_k(X)$.

Lemma 3.4. *One has*

$$I_{k,d}(X) - T_k(X) \ll X^{r-1} + X^{r+\theta_{d,r}} \max_{\mathbf{y},\mathbf{u}_0} W(X;\mathbf{y},\mathbf{u}_0),$$

where the maximum is taken over distinct $y_1,\ldots,y_r$ *with* $1 \leqslant y_j \leqslant X$ *and over* $1 \leqslant |u_{0j}| \leqslant CX^{d(r-1)}$ $(1 \leqslant j \leqslant r)$.

Proof. It follows from Lemma 3.2 together with the bound for $N_r(X;\mathbf{y})$ just obtained that

$$V_{2,r}(X) \ll X^{r+\theta_{d,r}} \max_{\mathbf{y},\mathbf{u}_0} W(X;\mathbf{y},\mathbf{u}_0).$$

The conclusion of the lemma is obtained by substituting this estimate together with that supplied by Lemma 3.1 into (3.1). □

4. Exploiting multiplicative relations

Our goal in this section is to estimate the quantity $W(X;\mathbf{y},\mathbf{u}_0)$ that counts solutions of the multiplicative equations (3.8) equipped with their ancillary conditions (3.9) and (3.10). For this purpose, we follow closely the trail first adopted in our work with Vaughan [10, §2].

Lemma 4.1. *Suppose that* $y_1,\ldots,y_r$ *are distinct integers with* $1 \leqslant \mathbf{y} \leqslant X$, *and that* u_{0j} $(1 \leqslant j \leqslant r)$ *are integers with* $1 \leqslant |u_{0j}| \leqslant CX^{d(r-1)}$. *Then one has* $W(X;\mathbf{y},\mathbf{u}_0) \ll X^{k/r+\varepsilon}$.

Proof. We begin with a notational device from [10, §2]. Let $\mathcal{I}$ denote the set of indices $\mathbf{i} = (i_1,\ldots,i_r)$ with $0 \leqslant i_m \leqslant k$ $(1 \leqslant m \leqslant r)$. Define the map $\varphi : \mathcal{I} \to [0,(k+1)^r) \cap \mathbb{Z}$ by putting

$$\varphi(\mathbf{i}) = \sum_{m=1}^{r} i_m (k+1)^{m-1}.$$

The map φ is bijective, and we may define the *successor* $\mathbf{i}+1$ of the index $\mathbf{i}$ by means of the relation

$$\mathbf{i}+1 = \varphi^{-1}(\varphi(\mathbf{i})+1).$$

We then define $\mathbf{i}+h$ inductively via the formula $\mathbf{i}+(h+1) = (\mathbf{i}+h)+1$. Finally, when $\mathbf{i} \in \mathcal{I}$, we write $\mathcal{J}(\mathbf{i})$ for the set of indices $\mathbf{j} \in \mathcal{I}$ having the property that, for some $h \in \mathbb{N}$, one has $\mathbf{j}+h = \mathbf{i}$. Thus, the set $\mathcal{J}(\mathbf{i})$ is the set of all precursors of $\mathbf{i}$, in the natural sense.

Equipped with this notation, we now explain how systematically to extract common factors between the variables in the system of equations (3.8). Put

$$\alpha_{\mathbf{0}} = (u_{01}, u_{02}, \ldots, u_{0r}),$$

noting that by hypothesis, this integer is fixed. Suppose at stage $\mathbf{i}$ that $\alpha_{\mathbf{j}}$ has been defined for all $\mathbf{j} \in \mathcal{J}(\mathbf{i})$. We then define

$$\alpha_{\mathbf{i}} = \left(\frac{u_{i_1 1}}{\beta_{\mathbf{i}}^{(1)}}, \frac{u_{i_2 2}}{\beta_{\mathbf{i}}^{(2)}}, \ldots, \frac{u_{i_r r}}{\beta_{\mathbf{i}}^{(r)}} \right),$$

in which we write

$$\beta_{\mathbf{i}}^{(m)} = \prod_{\substack{\mathbf{j}\in\mathcal{J}(\mathbf{i})\\ j_m=i_m}} \alpha_{\mathbf{j}}.$$

As is usual, the empty product is interpreted to be 1. As a means of preserving intuition concerning the numerous variables generated in this way, we write

$$\widetilde{\alpha}_{lm}^{\pm} = \pm \prod_{\substack{\mathbf{j}\in\mathcal{I}\\ j_m=l}} \alpha_{\mathbf{j}} \quad (0\leqslant l\leqslant k,\ 1\leqslant m\leqslant r).$$

Then, much as in [10, §2], it follows that when $0\leqslant l\leqslant k$ and $1\leqslant m\leqslant r$, for some choice of the sign $\pm$, one has $u_{lm}=\widetilde{\alpha}_{lm}^{\pm}$. Note here that the ambiguity in the sign of u_{lm} relative to $|\widetilde{\alpha}_{lm}^{\pm}|$ is a feature overlooked in the treatment of [10], though the ensuing argument requires no significant modification to be brought to play in order that the same conclusion be obtained. At worst, an additional factor $2^{r(k+1)}$ would need to be absorbed into the constants implicit in Vinogradov's notation.

With this notation in hand, it follows from its definition that $W(X;\mathbf{y},\mathbf{u}_0)$ is bounded above by the number $\Omega_r(X;\mathbf{y},\mathbf{u}_0)$ of solutions of the system

$$y_1+\widetilde{\alpha}_{i1}^{\pm} = y_2+\widetilde{\alpha}_{i2}^{\pm} = \ldots = y_r+\widetilde{\alpha}_{ir}^{\pm} \quad (1\leqslant i\leqslant k), \tag{4.1}$$

with

$$1\leqslant |\widetilde{\alpha}_{ij}^{\pm}|\leqslant X \quad (1\leqslant i\leqslant k,\ 1\leqslant j\leqslant r). \tag{4.2}$$

Notice here that $\widetilde{\alpha}_{0m}^{\pm}=u_{0m}$. Thus, it follows from a divisor function estimate that when the integers u_{0m} are fixed with

$$1\leqslant |u_{0m}|\leqslant CX^{d(r-1)} \quad (1\leqslant m\leqslant r),$$

then there are $O(X^{\varepsilon})$ possible choices for the variables $\alpha_{\mathbf{i}}$ having the property that $i_m=0$ for some index m with $1\leqslant m\leqslant r$.

Having carefully prepared the notational infrastructure to make comparison with [10, §§2 and 3] transparent, we may now follow the argument of the latter mutatis mutandis. When $1\leqslant p\leqslant r$, we write

$$B_p=\prod_{\mathbf{i}} \alpha_{\mathbf{i}}, \tag{4.3}$$

where the product is taken over all $\mathbf{i}\in\mathcal{I}$ with $i_l>i_p$ $(l\ne p)$, and $i_l>0$ $(1\leqslant l\leqslant r)$. Thus, in view of (4.2), one has

$$\prod_{p=1}^r B_p \leqslant \prod_{\substack{\mathbf{i}\in\mathcal{I}\\ i_l>0\,(1\leqslant l\leqslant r)}} \alpha_{\mathbf{i}} \leqslant \prod_{i=1}^k |\widetilde{\alpha}_{i1}^{\pm}| \leqslant X^k,$$

and so in any solution $\boldsymbol{\alpha}^{\pm}$ of (4.1) counted by $\Omega_r(X;\mathbf{y},\mathbf{u}_0)$, there exists an index p with $1\leqslant p\leqslant r$ such that

$$1\leqslant B_p\leqslant X^{k/r}. \tag{4.4}$$

By relabelling variables, we consequently deduce that

$$\Omega_r(X;\mathbf{y},\mathbf{u}_0) \ll \Upsilon_r(X;\mathbf{y},\mathbf{u}_0),$$

where $\Upsilon_r(X;\mathbf{y},\mathbf{u}_0)$ denotes the number of integral solutions of the system

$$\widetilde{\alpha}_{i1}^{\pm}-\widetilde{\alpha}_{ij}^{\pm}=L_j \quad (1\leqslant i\leqslant k,\, 2\leqslant j\leqslant r), \tag{4.5}$$

with $L_j=y_j-y_1$ $(2\leqslant j\leqslant r)$, and with the integral tuples $\alpha_{\mathbf{i}}$ satisfying (4.2) together with the inequality

$$1\leqslant B_1\leqslant X^{k/r}. \tag{4.6}$$

We emphasise here that, when $y_1,\ldots,y_r$ are distinct, then $L_j\neq 0$ $(2\leqslant j\leqslant r)$.

We now proceed under the assumption that $y_1,\ldots,y_r$ are fixed and distinct, whence the integers L_j $(2\leqslant j\leqslant r)$ are fixed and non-zero. It follows just as in the final paragraphs of [10, §2] that, when the variables $\alpha_{\mathbf{i}}$, with $\mathbf{i}\in\mathcal{I}$ satisfying $i_l>i_1$ $(2\leqslant l\leqslant r)$, are fixed, then there are $O(X^\varepsilon)$ possible choices for the tuples $\alpha_{\mathbf{i}}$ satisfying (4.2) and (4.5). Here we make use of the fact that the variables $\alpha_{\mathbf{i}}$, in which $i_m=0$ for some index m with $1\leqslant m\leqslant r$, may be considered fixed with the potential loss of a factor $O(X^\varepsilon)$ in the resulting estimates. By making use of standard estimates for the divisor function, however, we find from (4.6) and the definition (4.3) that there are $O(X^{k/r+\varepsilon})$ possible choices for the variables $\alpha_{\mathbf{i}}$ with $\mathbf{i}\in\mathcal{I}$ satisfying $i_l>i_1$ $(2\leqslant l\leqslant r)$. We therefore infer that $\Upsilon_r(X;\mathbf{y},\mathbf{u}_0)\ll X^{k/r+\varepsilon}$, whence $\Omega_r(X;\mathbf{y},\mathbf{u}_0)\ll X^{k/r+\varepsilon}$, and finally $W(X;\mathbf{y},\mathbf{u}_0)\ll X^{k/r+\varepsilon}$. This completes the proof of the lemma. □

The proof of Theorem 1.1 is now at hand. By applying Lemma 3.4 in combination with Lemma 4.1, we obtain the upper bound

$$I_{k,d}(X)-T_k(X)\ll X^{r+\theta_{d,r}}\cdot X^{k/r+\varepsilon}.$$

By minimising the right hand side over $2\leqslant r\leqslant k$, a comparison of (1.6) and (3.11) now confirms that this estimate delivers the one claimed in the statement of Theorem 1.1.

5. Corollaries and refinements

We complete our discussion of incomplete Vinogradov systems by first deriving the corollaries to Theorem 1.1 presented in the introduction, and then considering refinements to the main strategy.

The proof of Corollary 1.2. Suppose that $d\leqslant\sqrt{k}$ and take r to be the integer closest to $\sqrt{k}$. Thus $d\leqslant r$ and we find from (1.6) that

$$\gamma_{k,d}\leqslant r+k/r+d(d+1)/2<\sqrt{4k+1}+d(d+1)/2.$$

An application of Theorem 1.1 therefore leads us to the asymptotic formula

$$I_{k,d}(X)=T_k(X)+O(X^{\sqrt{4k+1}+d(d+1)/2}),$$

confirming the first claim of the corollary. In particular, when $d=o(k^{1/4})$, we discern that

$$\sqrt{4k+1}+d(d+1)/2\leqslant\sqrt{4k+1}+o(k^{1/2})=(2+o(1))\sqrt{k},$$

and so the final claim of the corollary follows. □

The proof of Corollary 1.3. Suppose that $d \geqslant 1$ and $k \geqslant 4d+3$. In this situation, by reference to (1.6) with $r=2$, we find that

$$\gamma_{k,d} \leqslant 2 + \tfrac{1}{2}k + 2d - 1 = \tfrac{1}{2}(k+4d+2) \leqslant k - \tfrac{1}{2}.$$

Consequently, it follows from Theorem 1.1 that $I_{k,d}(X) - T_k(X) \ll X^{k-1/2}$, so that the first claim of the corollary follows.

Next by considering (1.6) with r taken to be the integer closest to $\sqrt{k/(d+1)}$, we find that

$$\gamma_{k,d} \leqslant rd + (r + k/r) \leqslant (d+1)\sqrt{4k/(d+1)+1}.$$

In this instance, Theorem 1.1 supplies the asymptotic formula

$$I_{k,d}(X) = T_k(X) + O(X^{\sqrt{4k(d+1)+(d+1)^2}}),$$

which establishes the second claim of the corollary.

Finally, when η is small and positive, and $1 \leqslant d \leqslant \eta^2 k$, one finds that

$$\gamma_{k,d} \leqslant \sqrt{4\eta^2k^2 + \eta^4k^2 + (4+2\eta^2)k + 1} < 3\eta k.$$

The final estimate of the corollary follows, and this completes the proof. $\square$

Some refinement is possible within the argument applied in the proof of Theorem 1.1 for smaller values of k. Thus, an argument analogous to that discussed in the final paragraph of [10, §2] shows that the bound $1 \leqslant B_p \leqslant X^{k/r}$ of equation (4.4) may be replaced by the corresponding bound

$$1 \leqslant B_p \leqslant X^{\omega(k,r)},$$

where we write

$$\omega(k,r) = k^{1-r}\sum_{i=1}^{k-1} i^{r-1}.$$

In order to justify this assertion, denote by $\mathcal{I}^+$ the set of indices $\mathbf{i} \in \mathcal{I}$ such that $i_l > 0$ $(1 \leqslant l \leqslant r)$, and let $\mathcal{I}^*$ denote the corresponding set of indices subject to the additional condition that for some index p with $1 \leqslant p \leqslant r$, one has $i_l > i_p$ whenever $l \neq p$. Then, just as in [10, §2], one has $\text{card}(\mathcal{I}^+) = k^r$ and $\text{card}(\mathcal{I}^*) = r\psi_r(k)$, where

$$\psi_r(k) = \sum_{i=1}^{k-1} i^{r-1} < k^r/r.$$

In the situation of the proof of Lemma 4.1 in §4, the variables $\alpha_{\mathbf{i}}$ with $i_l = 0$ for some index l with $1 \leqslant l \leqslant r$ are already determined via a divisor function estimate. By permuting and relabelling indices i_l, for each fixed index l, as necessary, the argument of the proof can be adapted to show that $W(X;\mathbf{y},\mathbf{u}_0) \ll Y_r(X)$, where $Y_r(X)$ denotes the number of solutions $\boldsymbol{\alpha}^{\pm}$ as before, but subject to the additional condition

$$\prod_{\mathbf{i}\in\mathcal{I}^*} \alpha_{\mathbf{i}} \leqslant \Biggl(\prod_{\mathbf{i}\in\mathcal{I}^+} \alpha_{\mathbf{i}}\Biggr)^{\text{card}(\mathcal{I}^*)/\text{card}(\mathcal{I}^+)}.$$

Then

$$\prod_{p=1}^{r} B_p \leqslant \prod_{\mathbf{i}\in\mathcal{I}^*} \alpha_{\mathbf{i}} \leqslant (X^k)^{r\psi_r(k)/k^r}.$$

Consequently, in any solution $\boldsymbol{\alpha}^{\pm}$ of (4.1) counted by $\Omega_r(X;\mathbf{y};\mathbf{u}_0)$, there exists an index p with $1 \leqslant p \leqslant r$ such that

$$1 \leqslant B_p \leqslant X^{\psi_r(k)/k^{r-1}} = X^{\omega(k,r)}.$$

By pursuing the same argument as in our earlier treatment, mutatis mutandis, we now derive the upper bound

$$I_{k,d}(X) - T_k(X) \ll X^{\gamma'_{k,d}+\varepsilon},$$

where

$$\gamma'_{k,d} = \min_{2\leqslant r\leqslant k}\left(r + \omega(k,r) + \sum_{l=1}^{r} \max\{d-l+1,0\}\right).$$

We conclude from these deliberations that Theorem 1.1 and the first conclusion of Corollary 1.3 may be refined as follows.

Theorem 5.1. *Suppose that $k \geqslant 3$ and $0 \leqslant d < k/2$. Then, for each $\varepsilon > 0$, one has*

$$I_{k,d}(X) - T_k(X) \ll X^{\gamma'_{k,d}+\varepsilon},$$

where

$$\gamma'_{k,d} = \min_{2\leqslant r\leqslant k}\left(r + k^{1-r}\sum_{i=1}^{k-1} i^{r-1} + \sum_{l=1}^{r} \max\{d-l+1,0\}\right).$$

In particular, provided that $d \geqslant 1$ and $k \geqslant 4d+2$, one has

$$I_{k,d}(X) = k!X^k + O(X^{k-1/2}).$$

Proof. The proof of the first conclusion has already been outlined. As for the second, by taking $r = 2$ we discern that

$$\gamma'_{k,d} \leqslant 2 + \tfrac{1}{2}(k-1) + 2d - 1.$$

Thus, provided that $k > 4d+1$, one finds that $\gamma'_{k,d} \leqslant k - 1/2$, and hence the final conclusion of the theorem follows from the first. □

Energetic readers will find a smorgasbord of problems to investigate allied to those examined in this paper. We mention three in order to encourage work on these topics.

We begin by noting that the conclusions of Theorem 1.1 establish the paucity of non-diagonal solutions in the system (1.5) when d is smaller than about $k/4$. In principle, the methods employed remain useful when $d < k/2$. However, when $d > k/2$ the analogue of the identity (2.4) that would be obtained would contain terms involving h^2, or even larger powers of h, and this precludes the possibility of eliminating all of the terms involving h in any useful manner. A simple test case would be the situation with $d = k-1$, wherein the system (1.5) assumes the shape

$$x_1^j + \ldots + x_k^j = y_1^j + \ldots + y_k^j \quad (2 \leqslant j \leqslant k).$$

When $k = 3$ an affine slicing approach has been employed in [12] to resolve the associated paucity problem. It would be interesting to address this problem when $k \geqslant 4$.

The focus of this paper has been on the situation in which one slice is removed from a Vinogradov system. When more than one slice is removed, two or more auxiliary variables $h_1, h_2, \ldots$ take the place of the single variable h in the identity (2.4), and this seems to pose serious problems for our methods. A simple test case in this context would address the system of equations

$$x_1^j + \ldots + x_k^j = y_1^j + \ldots + y_k^j \quad (j \in \{1, 2, \ldots, k-2, k+1\}),$$

with $k \geqslant 3$. Here, the situation with $k = 3$ has been successfully addressed by a number of authors (see [3, 8] and [6, Corollary 0.3]), but little seems to be known for $k \geqslant 4$. Much more is known when the omitted slices are carefully chosen so that the resulting systems assume a special shape. Most obviously, one could consider systems of the shape

$$x_1^{tj} + \ldots + x_k^{tj} = y_1^{tj} + \ldots + y_k^{tj} \quad (1 \leqslant j \leqslant k-1).$$

By specialising variables, one finds from [10, Theorem 1] that the number of non-diagonal solutions of this system with $1 \leqslant \mathbf{x}, \mathbf{y} \leqslant X$ is $O(X^{t\sqrt{4k+1}})$, and this is $o(T_k(X))$ provided only that the integer t is smaller than $\frac{1}{2}\sqrt{k} - 1$. Moreover, the ingenious work of Brüdern and Robert [2] shows that when $k \geqslant 4$, there is a paucity of non-diagonal solutions to systems of the shape

$$x_1^{2j-1} + \ldots + x_k^{2j-1} = y_1^{2j-1} + \ldots + y_k^{2j-1} \quad (1 \leqslant j \leqslant k-1),$$

wherein all of the even degree slices are omitted. A strategy for systems having arbitrary exponents can be extracted from [11], though the work there misses a paucity estimate by a factor $(\log X)^A$, for a suitable $A > 0$.

We remark finally that the system of equations (1.5) central to Theorem 1.1 has the property that there are $k - 1$ equations and k pairs of variables x_i, y_i. No paucity result is available when the number of pairs of variables exceeds k. The simplest challenge in this direction would be to establish that when $k \geqslant 3$, one has

$$J_{k+2,k}(X) = T_{k+2}(X) + o(X^{k+2}).$$

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. Brüdern and O. Robert, *A paucity estimate related to Newton sums of odd degree*, Mathematika **58** (2012), no. 2, 225–235.

[3] G. Greaves, *Some Diophantine equations with almost all solutions trivial*, Mathematika **44** (1997), no. 1, 14–36.

[4] L.-K. Hua, *Additive theory of prime numbers*, American Math. Soc., Providence, RI, 1965.

[5] I. G. MacDonald, *Symmetric functions and Hall polynomials*, Second Edition, Oxford Mathematical Monographs, Oxford University Press, 1995.

[6] P. Salberger, *Rational points of bounded height on threefolds*, Analytic Number Theory, Clay Math. Proc. **7** (2007), pp. 207–216, Amer. Math. Soc., Providence, RI.

[7] P. Salberger and T. D. Wooley, *Rational points on complete intersections of higher degree, and mean values of Weyl sums*, J. London Math. Soc. (2) **82** (2010), no. 2, 317–342.

[8] C. M. Skinner and T. D. Wooley, *On the paucity of non-diagonal solutions in certain diagonal Diophantine systems*, Quart. J. Math. Oxford (2) **48** (1997), 255–277.

[9] J. Steinig, *On some rules of Laguerre's, and systems of equal sums of like powers*, Rend. Mat. (6) **4** (1971), 629–644.

[10] R. C. Vaughan and T. D. Wooley, *A special case of Vinogradov's mean value theorem*, Acta Arith. **79** (1997), no. 3, 193–204.

[11] T. D. Wooley, *A note on symmetric diagonal equations*, Number Theory with an emphasis on the Markoff spectrum (Provo, UT, 1991), Editors: A. D. Pollington and W. Moran, Dekker, New York, 1993, pp 317-321.

[12] T. D. Wooley, *An affine slicing approach to certain paucity problems*, Analytic Number Theory: Proceedings of a Conference in Honor of Heini Halberstam (B.C. Berndt, H. G. Diamond and A. J. Hildebrand, eds.), vol. 2, 1996, pp. 803–815, Prog. Math. **139**, Birkhäuser, Boston.

[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. (3) **118** (2019), no. 4, 942–1016.

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

E-mail address: twooley@purdue.edu