Communications in
Number Theory and Physics
Volume 11, Number 4, 837–877, 2017

Poisson distribution for gaps between sums of two squares and level spacings for toral point scatterers

Tristan Freiberg, Pär Kurlberg, and Lior Rosenzweig

We investigate the level spacing distribution for the quantum spectrum of the square billiard. Extending work of Connors–Keating, and Smilansky, we formulate an analog of the Hardy–Littlewood prime k-tuple conjecture for sums of two squares, and show that it implies that the spectral gaps, after removing degeneracies and rescaling, are Poisson distributed. Consequently, by work of Rudnick and Ueberschär, the level spacings of arithmetic toral point scatterers, in the weak coupling limit, are also Poisson distributed. We also give numerical evidence for the conjecture and its implications.

1 Introduction **837**

2 Discussion **842**

3 Notation **846**

4 Deducing Theorem 1.2 from Proposition 1.3 **847**

5 Preliminaries **850**

6 Proof of Proposition 1.3 **858**

References **874**

1. Introduction

According to the Berry–Tabor conjecture [2], the energy levels for generic integrable systems should be Poisson distributed in the semiclassical limit.

As noted by Connors and Keating [5], the square billiard, though integrable, is not generic: due to spectral degeneracies, the level spacing distribution tends to a δ-function at zero. However, if we remove the degeneracies and rescale so that the mean spacing is unity, numerics indicate Poisson spacings.

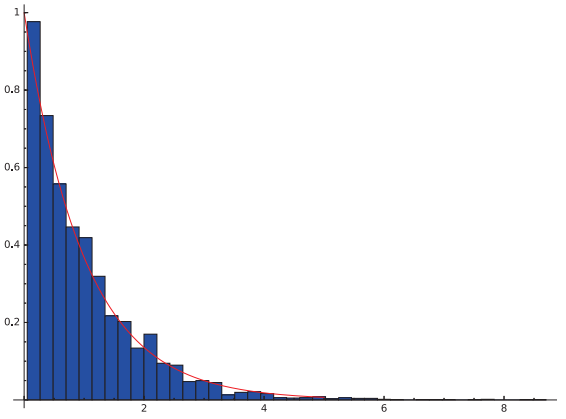

Figure 1.1: Rescaled gaps between consecutive energy levels in $[10^{99}, 10^{99} + 110000]$, after removing degeneracies. The rescaled gaps have mean one; without rescaling the mean gap is $19.42\cdots$. Number of gaps: 5663. We also plot the density function (red in color printout) $P(x) = \mathrm{e}^{-x}$, consistent with Poisson spacings.

The energy levels of the square billiard, say with side length 2π, are number theoretical in nature, and given by $a^2 + b^2$ for $a, b \in \mathbb{Z}$. After removing degeneracies and rescaling, we are led to study the nearest neighbor spacing distribution

$$\frac{1}{N(x)} \# \left\{ E_n \leqslant x : \frac{E_{n+1} - E_n}{x/N(x)} < \lambda \right\} \tag{1.1}$$

(as $x \to \infty$), where E_n denotes the nth smallest element of the set

$$\mathbb{E} := \{a^2 + b^2 : a, b \in \mathbb{Z}\}, \quad \text{and} \quad N(x) := \#\{E_n \leqslant x : E_n \in \mathbb{E}\}. \tag{1.2}$$

(In our setting, the leading order of the density of states is asymptotically equal to $C/\sqrt{\log x}$ as $x \to \infty$ [see (1.5)], and hence the spacing distribution of the unfolded levels $\left(CE_n/\sqrt{\log E_n}\right)_{n \geqslant 1}$ is asymptotically the same as that of the gaps in (1.1).)

Rather than studying the spacing distribution directly, we shall proceed by investigating *unordered* k-tuples of elements in $\mathbb{E}$. Thus, given $k \geqslant 1$ and $\boldsymbol{h} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$ with $\#\boldsymbol{h} = k$, consider the correlation function

$$R_k(\boldsymbol{h}; x) := \frac{1}{x} \sum_{n \leqslant x} \mathbf{1}_{\mathbb{E}}(n + h_1) \cdots \mathbf{1}_{\mathbb{E}}(n + h_k), \tag{1.3}$$

where $\mathbf{1}_{\mathbb{E}}$ denotes the indicator function of $\mathbb{E}$. If $\boldsymbol{h} = \{0\}$, this is the level density

$$R_1(x) := \frac{N(x)}{x}. \tag{1.4}$$

By a classical result of Landau [23],

$$R_1(x) \sim \frac{C}{\sqrt{\log x}} \quad (x \to \infty), \tag{1.5}$$

where $C > 0$ is an explicitly given constant (see (2.1)). To formulate an analog of (1.5) for $k > 1$ we need some further notation. Given a prime $p \not\equiv 1 \bmod 4$, define

$$\delta_{\boldsymbol{h}}(p) := \lim_{\alpha \to \infty} \frac{\#\{0 \leqslant a < p^\alpha : \forall h \in \boldsymbol{h}, a + h \equiv \square + \square \bmod p^\alpha\}}{p^\alpha}. \tag{1.6}$$

(That this limit exists is shown in Section 5, see Propositions 5.3 and 5.2.) Further, for $k \geqslant 1$ and a set $\boldsymbol{h} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$ with $\#\boldsymbol{h} = k$, we define the *singular series* for $\boldsymbol{h}$ by

$$\mathfrak{S}_{\boldsymbol{h}} := \prod_{p \not\equiv 1 \bmod 4} \frac{\delta_{\boldsymbol{h}}(p)}{\big(\delta_{\{0\}}(p)\big)^k}, \tag{1.7}$$

with $\delta_{\{0\}}(p)$ and $\delta_{\boldsymbol{h}}(p)$ as in (1.6). We note that $\delta_{\{0\}}(p) > 0$ for all $p \not\equiv 1 \bmod 4$, and that the product converges to a nonzero limit if $\delta_{\boldsymbol{h}}(p) > 0$ for all $p \not\equiv 1 \bmod 4$ (see Proposition 5.4). If $\delta_{\boldsymbol{h}}(p) = 0$ for some $p \not\equiv 1 \bmod 4$, we define $\mathfrak{S}_{\boldsymbol{h}}$ to be zero; it is easy to see that $R_k(\boldsymbol{h}; x) = 0$ for all x if $\mathfrak{S}_{\boldsymbol{h}} = 0$.

We can now formulate an analog of the Hardy–Littlewood prime k-tuple conjecture.

Conjecture 1.1. *Fix $k \geqslant 1$, and a set $\boldsymbol{h} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$ with $\#\boldsymbol{h} = k$. If $\mathfrak{S}_{\boldsymbol{h}} > 0$, then*

$$R_k(\boldsymbol{h}; x) \sim \mathfrak{S}_{\boldsymbol{h}} \big(R_1(x)\big)^k \quad (x \to \infty). \tag{1.8}$$

Our main result, Theorem 1.2 below, is conditional on the hypothesis that (1.8) holds on average. To be precise, let $\mathcal{E}_{\boldsymbol{h}}(x)$ be defined by the relation

$$R_k(\boldsymbol{h};x) =: \big(\mathfrak{S}_{\boldsymbol{h}} + \mathcal{E}_{\boldsymbol{h}}(x)\big)\big(R_1(x)\big)^k. \tag{1.9}$$

Further, let Δ^k be the region in $\mathbb{R}^k$ defined by

$$\Delta^k := \{(x_1,\dots,x_k) \in \mathbb{R}^k : 0 < x_1 < \cdots < x_k\}, \tag{1.10}$$

and, given $\mathscr{C} \subset \Delta^k$ and $y \in \mathbb{R}$, let $y\mathscr{C}$ be the dilation of $\mathscr{C}$ defined by

$$y\mathscr{C} := \{(yx_1,\dots,yx_k) : (x_1,\dots,x_k) \in \mathscr{C}\}.$$

Our hypothesis is that the error term $\mathcal{E}_{\boldsymbol{h}}(x)$ is small on average over dilates of certain bounded convex sets.

Hypothesis $(k,\mathscr{C},\boldsymbol{o})$. *Fix an integer $k \geqslant 1$, and a bounded convex set $\mathscr{C} \subset \Delta^k$. Set $\boldsymbol{o} := \emptyset$, or set $\boldsymbol{o} := \{0\}$. Let x and y be real parameters tending to infinity in such a way that $yR_1(x) \sim 1$. There exists a function $\varepsilon(x)$, with $\varepsilon(x) \to 0$ as $x \to \infty$, such that for x sufficiently large in terms of k and $\mathscr{C}$,*

$$\left|\sum_{(h_1,\dots,h_k)\in y\mathscr{C}\cap\mathbb{Z}^k} \mathcal{E}_{\boldsymbol{o}\cup\boldsymbol{h}}(x)\right| \leqslant \varepsilon(x) \sum_{(h_1,\dots,h_k)\in y\mathscr{C}\cap\mathbb{Z}^k} \mathfrak{S}_{\boldsymbol{o}\cup\boldsymbol{h}}, \tag{1.11}$$

where $\boldsymbol{h} = \{h_1,\dots,h_k\}$ in both summands.

Under the above hypothesis we find that the spacing distribution (1.1) is indeed Poissonian. Moreover, the distribution of the number of points in intervals of size comparable to the mean spacing is consistent with that of a Poisson process. (We remark that our hypothesis can be weakened slightly: see Section 4.)

Theorem 1.2. *Let x and y be real parameters tending to infinity in such a way that $yR_1(x) \sim 1$. Fix integers $m \geqslant 0$ and $r \geqslant 1$, and fix $\lambda, \lambda_1, \dots, \lambda_r \in \mathbb{R}^+$. Assume that Hypothesis $(k,\mathscr{C},\{0\}))$ (respectively, Hypothesis $(k,\mathscr{C},\emptyset)$ holds for all $k \geqslant 1$, and all bounded, convex sets $\mathscr{C} \subset \Delta^k$. Then (a) (respectively, (b)) holds.*

(*a*) *We have*

$$\frac{1}{N(x)}\#\{E_n \leqslant x : \forall j \leqslant r, E_{n+j} - E_{n+j-1} \leqslant \lambda_j y\} \sim \prod_{j=1}^{r} \int_0^{\lambda_j} \mathrm{e}^{-t}\,\mathrm{d}t \quad (x \to \infty). \tag{1.12}$$

(*b*) *We have*

$$\frac{1}{x}\#\{n \leqslant x : N(n+\lambda y) - N(n) = m\} \sim \mathrm{e}^{-\lambda}\frac{\lambda^m}{m!} \quad (x \to \infty). \tag{1.13}$$

In [28], Rudnick and Ueberschär considered the spectrum of "toral point scatterers", namely the Laplace operator, perturbed by a delta potential, on two dimensional tori. They showed (cf. [28, Corollary 1.3]) that the level spacings of the perturbed eigenvalues, in the weak coupling limit, have the same distribution as the level spacings of the unperturbed eigenvalues (after removing multiplicities.) An interesting consequence of Conjecture 1.1 (or, to be precise, Hypothesis $(k, \mathscr{C}, \{0\})$), is thus that the Berry–Tabor conjecture holds for toral point scatterers, in the weak coupling limit, for arithmetic tori of the form $\mathbb{R}^2/\mathbb{Z}^2$.

We remark that Gallagher [8] proved the analog of Theorem 1.2 (b) for primes. As in Gallagher's proof, a key technical result is that the singular series is of average order one, over certain geometric regions.

Proposition 1.3. *Fix an integer $k \geqslant 1$, and a bounded convex set $\mathscr{C} \subset \Delta^k$. Set $\boldsymbol{o} := \emptyset$, or set $\boldsymbol{o} := \{0\}$. As $y \to \infty$, we have*

$$\sum_{(h_1,\dots,h_k)\in y\mathscr{C}\cap\mathbb{Z}^k} \mathfrak{S}_{\boldsymbol{o}\cup\boldsymbol{h}} = y^k\Big(\mathrm{vol}(\mathscr{C}) + O\big(y^{-2/3+o(1)}\big)\Big), \tag{1.14}$$

where $\boldsymbol{h} = \{h_1, \dots, h_k\}$ in the summand, and vol *stands for volume in $\mathbb{R}^k$.*

We note that any qualitative error term in Proposition 1.3 is sufficient to deduce Theorem 1.2. (See Remark 6.6 at the end of Section 6 for a brief outline how Ford's [7] and Pintz's [26] simplification of Gallagher's arguments can be adapted to give a weaker error term in Proposition 1.3.)

Acknowledgements. We thank Z. Rudnick for stimulating discussions on the subject matter, D. Koukoulopoulos for his comments on an early version of the paper, and the anonymous referee for helpful comments. T. F. was

partially supported by a grant from the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine. P. K. and L. R. were partially supported by grants from the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine, and the Swedish Research Council (621-2011-5498).

2. Discussion

Connors and Keating [5] determined the singular series for shifted pairs of sums of two squares, giving a probabilistic derivation of Conjecture 1.1 in the special case $k = 2$, and found that it matched numerics quite well (to within 2%). Smilansky [30] then expressed the singular series for pairs as products of p-adic densities, showing that its mean value (over short intervals of shifts) is consistent with a Poisson distribution, and that the same is true for sums of two squares, on assuming a uniform version of Conjecture 1.1 for $k = 2$. Smilansky also gave the singular series for triples corresponding to the shifts $\boldsymbol{h} = \{0, 1, 2\}$.

As already mentioned, the analog of Theorem 1.2 (b) for primes is due to Gallagher, who in [8] showed that an appropriate form of the Hardy–Littlewood prime k-tuples conjecture implies the prime analog of (1.12). (That it implies the prime analog of (1.13) is mentioned in Hooley's survey article [13, p. 137].) To show that the singular series is one on average (i.e., the prime analog of Proposition 1.3), Gallagher uses combinatorial identities for Stirling numbers of the second kind. In [19], Kowalski developed an elegant probabilistic framework for evaluating averages of singular series. Rather than using combinatorial identities, Kowalski showed that a certain duality between k-th moments of m-tuples and m-th moments of k-tuples holds [19, Theorem 1]. That the k-th moment of 1-tuples is equal to one is more or less trivial, but, by duality, Kowalski obtains the non-trivial consequence that first moments of k-tuples is also equal to one. (Note that (1.14) can be viewed as a first moment of k-tuples when $\boldsymbol{o} = \emptyset$.)

Our approach originates with techniques developed in [20, 21], and further refined in [10, 22]. Loosely speaking, the singular series $\mathfrak{S}_{\boldsymbol{h}}$ is expanded into local factors of the form $1 + \epsilon_{\boldsymbol{h}}(p)$, and thus

$$\mathfrak{S}_{\boldsymbol{h}} = \prod_{p} (1 + \epsilon_{\boldsymbol{h}}(p)) = \sum_{\substack{d \geqslant 1 \\ \text{squarefree}}} \epsilon_{\boldsymbol{h}}(d),$$

where $\epsilon_{\boldsymbol{h}}(1) = 1$ and $\epsilon_{\boldsymbol{h}}(d) := \prod_{p|d} \epsilon_{\boldsymbol{h}}(p)$. Hence

$$\sum_{\boldsymbol{h}} \mathfrak{S}_{\boldsymbol{h}} = \sum_{\substack{d \geqslant 1 \\ \text{squarefree}}} \sum_{\boldsymbol{h}} \epsilon_{\boldsymbol{h}}(d),$$

and the main term is given by $d = 1$. For d large, $|\epsilon_{\boldsymbol{h}}(d)|$ can be shown to be small on average. For d small, we use that $\epsilon_{\boldsymbol{h}}(d)$ (approximately) only depends on $\boldsymbol{h}$ mod d, together with complete cancellation when summing over the *full* set of residues modulo d, i.e., $\sum_{\boldsymbol{h} \bmod d} \epsilon_{\boldsymbol{h}}(d) = 0$. This follows, via the Chinese remainder theorem, from local cancellations $\sum_{\boldsymbol{h} \bmod p} \epsilon_{\boldsymbol{h}}(p) = 0$, which in turn can be deduced from the following easily verifiable identity: given *any* subset $X_p \subseteq \mathbb{Z}/p\mathbb{Z}$, we have (see Lemma 6.3 (b) and its proof for more details):

$$\sum_{(h_1,h_2,\dots,h_k) \in (\mathbb{Z}/p\mathbb{Z})^k} \sum_{\substack{m \in \mathbb{Z}/p\mathbb{Z} \\ \forall i \leqslant k,\, m+h_i \in X_p}} 1 = \sum_{(h_1,h_2,\dots,h_k) \in (X_p)^k} 1$$

However, unlike the setup in [10, 20, 22], where the local error term $\epsilon_{\boldsymbol{h}}(p)$ is determined by $\boldsymbol{h}$ mod p, in the current setting, it is not determined by $\boldsymbol{h}$ mod p^α, for any fixed α. On the other hand, the function $\boldsymbol{h} \to \epsilon_{\boldsymbol{h}}(p)$ has nice p-adic regularity properties, allowing us to approximate $\epsilon_{\boldsymbol{h}}(p)$ by truncations $\epsilon_{\boldsymbol{h}}(p^\alpha)$, which do only depend on $\boldsymbol{h} \bmod p^\alpha$, and for which $\epsilon_{\boldsymbol{h}}(p) - \epsilon_{\boldsymbol{h}}(p^\alpha) \ll 1/p^{\alpha-1}$ for all α. Apart from making the argument more complicated, this also results in a weaker error term: if $\epsilon_{\boldsymbol{h}}(p)$ only depended on $\boldsymbol{h}$ mod p, in (1.14) we would get a relative error of size $y^{-1+o(1)}$, rather than $y^{-2/3+o(1)}$. We also note that David, Koukoulopoulos, and Smith [6], in studying statistics of elliptic curves, have developed quite general methods for finding asymptotics of weighted sums $\sum_{\boldsymbol{h}} w_{\boldsymbol{h}} \mathfrak{S}_{\boldsymbol{h}}$, provided that the local factors have p-adic regularity properties similar to those referred to above. In fact, Proposition 1.3, though with a weaker error term, can be deduced from [6, Theorem 4.2].

We finally remark that the corresponding question in the function field setting is better understood: Bary–Soroker and Fehm [1] have recently shown that the sums of two squares analog of the k-tuple conjecture holds in the large q-limit for the function field setting (e.g., replacing $\mathbb{Z}$ by $\mathbb{F}_q[T]$, and $\mathbb{Z}[i]$ by $\mathbb{F}_q[\sqrt{-T}]$).

2.1. Evidence towards Conjecture 1.1.

We begin by stating a qualitative version of Conjecture 1.1.

Conjecture 2.1. *Fix $k \geqslant 1$, and a set $\boldsymbol{h} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$ with $\#\boldsymbol{h} = k$. If $\mathfrak{S}_{\boldsymbol{h}} > 0$, then there exist infinitely many integers n such that $n + \boldsymbol{h} \subseteq \mathbb{E}$.*

We remark that whether or not $\mathfrak{S}_{\boldsymbol{h}} > 0$ can be determined by a finite computation: this follows from Propositions 5.2 and 5.3. Examples of sets $\boldsymbol{h}$ for which $\mathfrak{S}_{\boldsymbol{h}} = 0$ are $\{0, 1, 2, 3\}$ and $\{0, 1, 2, 4, 5, 8, 16, 21\}$: any translate of $\{0, 1, 2, 3\}$ contains an integer congruent to 3 modulo 4, and hence $\delta_{\boldsymbol{h}}(2) = 0$; any translate of $\{0, 1, 2, 4, 5, 8, 16, 21\}$ contains an integer congruent to 3 or 6 modulo 9, and hence $\delta_{\boldsymbol{h}}(3) = 0$.

It is possible to show that $\mathfrak{S}_{\boldsymbol{h}} > 0$ for *any* set $\boldsymbol{h}$ containing at most three integers. The question of whether, for any $h_1, h_2, h_3 \in \mathbb{Z}$, we have $n + \{h_1, h_2, h_3\} \subseteq \mathbb{E}$ for infinitely many n, was apparently raised by Littlewood, and was answered in the affirmative by Hooley [14], using the theory of ternary quadratic forms. Conjecture 2.1 remains open for $k \geqslant 4$.

For fixed $k \geqslant 1$, and $\boldsymbol{h} = \{h_1, \dots, h_k\}$ with $\#\boldsymbol{h} = k$, the upper bound

$$\sum_{n \leqslant x} \mathbf{1}_{\mathbb{E}}(n + h_1) \cdots \mathbf{1}_{\mathbb{E}}(n + h_k) \ll_k \frac{x}{(\log x)^{k/2}} \prod_{\substack{p \equiv 3 \bmod 4 \\ p \mid h_j - h_j \\ \text{some } i < j}} \left(1 + \frac{k}{p}\right),$$

can be deduced from Selberg's sieve (see [29]), which is of the correct order of magnitude, according to Conjecture 1.1. The special case $\boldsymbol{h} = \{0, 1\}$ is due to Rieger [27]; the special case $\boldsymbol{h} = \{0, 1, 2\}$ is due to Cochrane and Dressler [4]; the general case is due to Nowak [25].

Lower bounds are more subtle. For $k = 2$, Hooley [15] and Indlekofer [16] showed that, for any nonzero integer h,

$$\sum_{n \leqslant x} \mathbf{1}_{\mathbb{E}}(n) \mathbf{1}_{\mathbb{E}}(n + h) \gg \frac{x}{\log x} \prod_{\substack{p \mid h \\ p \equiv 3 \bmod 4}} \left(1 + \frac{1}{p}\right),$$

but we are not aware of any such bounds for $k \geqslant 3$.

Assuming that a certain analog of the Elliott–Halberstam conjecture holds for sums of two squares, it is possible to deduce, from a result of Iwaniec [17, Theorem 4], the asymptotic $\sum_{n \leqslant x} \mathbf{1}_{\mathbb{E}}(n) \mathbf{1}_{\mathbb{E}}(n + 1) \sim x/(2 \log x)$, as $x \to$

∞, in agreement with the aforementioned conjecture of Connors and Keating [5], and Conjecture 1.1. (See (2.3), and Figure 2.1 for a numerical comparison.) We remark that, on a slightly weaker formulation of an Elliott–Halberstam analogue for sums of two squares, Iwaniec [17, Corollary 2, (2.3)] gives $\sum_{n\leqslant x} \mathbf{1}_{\mathbb{E}}(n)\mathbf{1}_{\mathbb{E}}(n+1) \sim 3x/(8\log x)$, as $x \to \infty$. (Few details are given, so it is hard to pinpoint the discrepancy in the constants; possibly the contribution from those n with $n \equiv 0 \bmod 8$ is not taken into account.)

2.2. Numerical evidence

Using Propositions 5.2 (b), (c) and 5.3 (b), (c), we can give $\mathfrak{S}_{\boldsymbol{h}}$ explicitly, as in the following examples. Let us first record that the constant C in (1.5) is the Landau–Ramanujan constant, given by

$$C := \frac{1}{\sqrt{2}} \prod_{p\equiv 3 \bmod 4} \left(1 - \frac{1}{p^2}\right)^{-1/2} = 0.764223\ldots. \tag{2.1}$$

It is straightforward to verify that

$$\mathfrak{S}_{\{0,1\}} = \frac{1}{2C^2} = 0.856108\ldots. \tag{2.2}$$

If (1.8) holds with $\boldsymbol{h} = \{0,1\}$ then, by (1.5) and (2.2),

$$\begin{aligned} N(\{0,1\};x) := \sum_{n\leqslant x} \mathbf{1}_{\mathbb{E}}(n)\mathbf{1}_{\mathbb{E}}(n+1) &\sim \frac{x}{2C^2}\big(R_1(x)\big)^2 \\ &\sim \frac{x}{2\log x} \quad (x \to \infty). \end{aligned} \tag{2.3}$$

The agreement with numerics is quite good (to within 1%).

As the simplest example with $k = 3$, we verify that

$$\mathfrak{S}_{\{0,1,2\}} = \frac{A}{4C^2}, \quad A := \prod_{p\equiv 3 \bmod 4} \left(1 - \frac{2}{p(p-1)}\right),$$

so Conjecture 1.1 implies that

$$\begin{aligned} N(\{0,1,2\};x) := \sum_{n\leqslant x} \mathbf{1}_{\mathbb{E}}(n)\mathbf{1}_{\mathbb{E}}(n+1)\mathbf{1}_{\mathbb{E}}(n+2) &\sim \frac{Ax}{4C^2}\big(R_1(x)\big)^3 \\ &\sim \frac{ACx}{4(\log x)^{3/2}} \end{aligned}$$

as $x \to \infty$. The agreement between numerics and model is only to within 10%.

x	$N(\{0,1\};x)$	$x\mathfrak{S}_{\{0,1\}}(R_1(x))^2$	Ratio
1000000000	25927011	25690391.1	1.00921
2000000000	50042411	49603435.5	1.00885
3000000000	73560246	72930222.0	1.00864
4000000000	96705170	95891759.7	1.00848
5000000000	119584162	118589346.3	1.00839
6000000000	142253331	141080935.2	1.00831
7000000000	164749254	163403937.1	1.00823
8000000000	187100631	185584673.5	1.00817
9000000000	209327440	207642640.3	1.00811

Figure 2.1: Observed data vs prediction for $\boldsymbol{h} = \{0, 1\}$.

x	$N(\{0,1,2\};x)$	$x\mathfrak{S}_{\{0,1,2\}}(R_1(x))^3$	Ratio
1000000000	1490691	1362419.3	1.09415
2000000000	2818128	2584683.5	1.09032
3000000000	4093602	3762317.2	1.08805
4000000000	5338091	4912433.3	1.08665
5000000000	6560430	6042800.3	1.08566
6000000000	7764604	7157833.6	1.08477
7000000000	8954282	8260369.7	1.08400
8000000000	10132295	9352396.2	1.08339
9000000000	11299877	10435380.5	1.08284

Figure 2.2: Observed data vs prediction for $\boldsymbol{h} = \{0, 1, 2\}$.

3. Notation

We define the set of natural numbers as $\mathbb{N} := \{1, 2, \ldots\}$. The letter p stands for a prime, n for an integer. We let $\square + \square$ stand for a generic element of $\mathbb{E}$, possibly a different element each time. Thus, for instance, $a + h \equiv \square + \square \bmod p^\alpha$ denotes that $a + h \equiv E \bmod p^\alpha$ for some $E \in \mathbb{E}$. We view k as a fixed natural number, and $\boldsymbol{h}$ as a nonempty, finite set of integers, with $\#\boldsymbol{h} = k$ unless otherwise indicated. We let $n + \boldsymbol{h} := \{n + h : h \in \boldsymbol{h}\}$. For $n \in \mathbb{N}$, $\omega(n)$ denotes the number of distinct prime divisors of n, $\nu_p(n)$ the p-adic valuation of n. (We also define $\nu_p(0) := \infty$.) That $\nu_p(n) = \alpha$ may also

be denoted by $p^\alpha \parallel n$. The radical of n is $\mathrm{rad}(n) := \prod_{p|n} p$, not to be confused with the squarefree part of n, viz. $\mathrm{sf}(n) := \prod_{p\parallel n} p$. By the least residue of an integer a modulo n we mean the integer r such that $a \equiv r \bmod n$ and $0 \leqslant r < n$. When written in an exponent, $\alpha \bmod 2$ is to be interpreted as the least residue of α modulo 2: for instance, $p^{\alpha \bmod 2} = 1$ if α is even.

We view x as a real parameter tending to infinity. Expressions of the form $A \sim B$ denote that $A/B \to 1$ as $x \to \infty$. We also view y as real parameter tending to infinity, typically in such a way that $yR_1(x) \sim 1$, i.e. $y \sim x/N(x)$. We may assume that x and y are sufficiently large in terms of any fixed quantity. Expressions of the form $A = O(B)$, $A \ll B$ and $B \gg A$ all denote that $|A| \leqslant c|B|$, where c is some positive constant, throughout the domain of the quantity A. The constant c is to be regarded as independent of any parameter unless indicated otherwise by subscripts, as in $A = O_k(B)$ (c depends on k only), $A \ll_{k,\lambda} B$ (c depends on k and λ only), etc. By $o(1)$ we mean a quantity that tends to zero as $y \to \infty$.

4. Deducing Theorem 1.2 from Proposition 1.3

Given $\vec{\boldsymbol{\iota}} = (i_1, \dots, i_r) \in \mathbb{N}^r$ such that $i_1 + \cdots + i_r = k$, and $\vec{\boldsymbol{\lambda}} = (\lambda_1, \dots, \lambda_r) \in \mathbb{R}^r$, let

$$\Theta_{\vec{\boldsymbol{\iota}},\vec{\boldsymbol{\lambda}}} := \{(x_1, \dots, x_k) \in \Delta^k : x_{i_1+\cdots+i_j} - x_{i_1+\cdots+i_{j-1}} \leqslant \lambda_j, j = 1, \dots, r\}, \tag{4.1}$$

where for $j = 1$ we let $x_{i_1+i_{j-1}} = x_0 := 0$. In the case where $r = 1$ and $\vec{\boldsymbol{\lambda}} = (\lambda)$,

$$\Theta_{\vec{\boldsymbol{\iota}},\vec{\boldsymbol{\lambda}}} = \Theta_{k,\lambda} := \{(x_1, \dots, x_k) \in \mathbb{R}^k : 0 < x_1 < \cdots < x_k \leqslant \lambda\}. \tag{4.2}$$

The following proof shows that Theorem 1.2 (a) and (b) hold under slightly weaker hypotheses than the ones stated: for (a), it is enough to assume that Hypothesis $(k, \Theta_{\vec{\boldsymbol{\iota}},\vec{\boldsymbol{\lambda}}}, \{0\})$, where $\vec{\boldsymbol{\iota}} = (i_1, \dots, i_r)$ and $\vec{\boldsymbol{\lambda}} = (\lambda_1, \dots, \lambda_r)$, holds for all $k \geqslant r$, and all $\vec{\boldsymbol{\iota}} \in \mathbb{N}^r$ satisfying $i_1 + \cdots + i_r = k$; for (b), it is enough to assume that Hypothesis $(k, \Theta_{k,\lambda}, \emptyset)$ holds for all $k \geqslant 1$.

Deduction of Theorem 1.2. As this argument has appeared many times in the literature, we merely give an outline of it and provide references. (a) To ease notation, we let $\vec{\boldsymbol{\iota}} = (i_1, \dots, i_r)$, $\vec{\boldsymbol{h}} = (h_1, \dots, h_k)$, $\boldsymbol{h} = \{h_1, \dots, h_k\}$, and

$$N(\{0\} \cup \boldsymbol{h}; x) := \sum_{n \leqslant x} \mathbf{1}_{\mathbb{E}}(n)\mathbf{1}_{\mathbb{E}}(n + h_1) \cdots \mathbf{1}_{\mathbb{E}}(n + h_k).$$

Let $\ell \geqslant 0$ be an integer, arbitrarily large but fixed. An inclusion-exclusion argument (see [12], [20, Appendix A] or [18, Key Lemma 2.4.12]) shows that

$$\begin{aligned}(4.3)\qquad &\sum_{k=r}^{r+2\ell+1}(-1)^{k-r}\sum_{i_1+\cdots+i_r=k}\sum_{\vec{\boldsymbol{h}}\in y\Theta_{\vec{\iota},\vec{\lambda}}\cap\mathbb{Z}^k} N(\{0\}\cup\boldsymbol{h};x)\\ &\leqslant \sum_{\substack{E_n\leqslant x\\ E_{n+j}-E_{n+j-1}\leqslant\lambda_j y\\ j=1,\dots,r}} 1 \leqslant \sum_{k=r}^{r+2\ell}(-1)^{k-r}\sum_{i_1+\cdots+i_r=k}\sum_{\vec{\boldsymbol{h}}\in y\Theta_{\vec{\iota},\vec{\lambda}}\cap\mathbb{Z}^k} N(\{0\}\cup\boldsymbol{h};x),\end{aligned}$$

the sums over $i_1+\cdots+i_r=k$, here and below, being over all $\vec{\boldsymbol{\iota}}\in\mathbb{N}^r$ for which $i_1+\cdots+i_r=k$. We make the substitution (1.9), with $\{0\}\cup\boldsymbol{h}$ and $k+1$ in place of $\boldsymbol{h}$ and k; we apply Hypothesis $(k,\Theta_{\vec{\iota},\vec{\boldsymbol{\lambda}}},\{0\})$ for all k and $\vec{\boldsymbol{\iota}}$ satisfying $r\leqslant k\leqslant r+2\ell+1$ and $i_1+\cdots+i_r=k$; we use Proposition 1.3, and our assumption that $yR_1(x)\sim 1$, i.e. $y\sim x/N(x)$, as $x\to\infty$. Thus, we deduce from (4.3) that

$$(4.4)\qquad \sum_{k=r}^{r+2\ell+1}(-1)^{k-r}\sum_{i_1+\cdots+i_r=k}\operatorname{vol}(\Theta_{\vec{\iota},\vec{\lambda}}) \leqslant \liminf_{x\to\infty}\frac{1}{N(x)}\sum_{\substack{E_n\leqslant x\\ E_{n+j}-E_{n+j-1}\leqslant\lambda_j y\\ j=1,\dots,r}} 1,$$

and

$$(4.5)\qquad \limsup_{x\to\infty}\frac{1}{N(x)}\sum_{\substack{E_n\leqslant x\\ E_{n+j}-E_{n+j-1}\leqslant\lambda_j y\\ j=1,\dots,r}} 1 \leqslant \sum_{k=r}^{r+2\ell}(-1)^{k-r}\sum_{i_1+\cdots+i_r=k}\operatorname{vol}(\Theta_{\vec{\iota},\vec{\lambda}}).$$

Since $\operatorname{vol}(\Theta_{\vec{\iota},\vec{\lambda}})=\lambda_1^{i_1}\cdots\lambda_r^{i_r}/(i_1!\cdots i_r!)$, the sums on the left and right of (4.4) and (4.5) are truncations of the Taylor series for $(1-\mathrm{e}^{-\lambda_1})\cdots(1-\mathrm{e}^{-\lambda_r})$. We have chosen ℓ arbitrarily large, so we may conclude that (1.12) holds, provided Hypothesis $(k,\Theta_{\vec{\iota},\vec{\lambda}},\{0\})$ does whenever $k\geqslant r$ and $i_1+\cdots+i_r=k$.

(b) We use an argument of Gallagher [8], who proved an analogous result for primes. Let $\ell\geqslant 1$ be an integer, arbitrarily large but fixed. We have

$$\begin{aligned}
&\sum_{n\leqslant x}\big(N(n+\lambda y)-N(n)\big)^\ell\\
&=\sum_{n\leqslant x}\Big(\sum_{0<h\leqslant\lambda y}\mathbf{1}_{\mathbb{E}}(n+h)\Big)^\ell\\
&=\sum_{n\leqslant x}\sum_{0<h_1,\dots,h_\ell\leqslant\lambda y}\mathbf{1}_{\mathbb{E}}(n+h_1)\cdots\mathbf{1}_{\mathbb{E}}(n+h_\ell)\\
&=\sum_{k=1}^{\ell}\varrho(\ell,k)\sum_{0<h_1<\cdots<h_k\leqslant\lambda y}\ \sum_{n\leqslant x}\mathbf{1}_{\mathbb{E}}(n+h_1)\cdots\mathbf{1}_{\mathbb{E}}(n+h_k),
\end{aligned}$$

where $\varrho(\ell,k)$ denotes the number of maps from $\{1,\dots,\ell\}$ onto $\{1,\dots,k\}$. Thus,

$$\begin{aligned}
&\frac{1}{x}\sum_{n\leqslant x}\big(N(n+\lambda y)-N(n)\big)^\ell\\
&=\sum_{k=1}^{\ell}\left(\frac{N(x)}{x}\right)^k\varrho(\ell,k)\sum_{0<h_1<\cdots<h_k\leqslant\lambda y}\big(\mathfrak{S}_{\boldsymbol{h}}+\mathcal{E}_{\boldsymbol{h}}(x)\big),
\end{aligned}$$

with $\boldsymbol{h}=\{h_1,\dots,h_k\}$ in the last summand. To sum over $0<h_1<\cdots<h_k\leqslant\lambda y$ is to sum over $(h_1,\dots,h_k)\in y\Theta_{k,\lambda}\cap\mathbb{Z}^k$ (see (4.2)). If Hypothesis $(k,\Theta_{k,\lambda},\emptyset)$ holds then for some function $\varepsilon(x)$ with $\varepsilon(x)\to 0$ $(x\to\infty)$, we have

$$\sum_{0<h_1<\cdots<h_k\leqslant\lambda y}\big(\mathfrak{S}_{\boldsymbol{h}}+\mathcal{E}_{\boldsymbol{h}}(x)\big)=\big(1+O_{\lambda,k}(\varepsilon(x))\big)\sum_{0<h_1<\cdots<h_k\leqslant\lambda y}\mathfrak{S}_{\boldsymbol{h}}.$$

Applying Proposition 1.3 (noting that $\mathrm{vol}(\Theta_{k,\lambda})=\lambda^k/k!$), and our assumption that $yR_1(x)\sim 1$, i.e. $y\sim x/N(x)$, as $x\to\infty$, we see that if Hypothesis $(k,\Theta_{k,\lambda},\emptyset)$ holds for $1\leqslant k\leqslant\ell$, then

$$\frac{1}{x}\sum_{n\leqslant x}\big(N(n+\lambda y)-N(n)\big)^\ell\sim\sum_{k=1}^{\ell}\varrho(\ell,k)\frac{\lambda^k}{k!}\quad(x\to\infty).\tag{4.6}$$

Gallagher's calculation in [8, Section 3] shows that $\sum_{k=1}^{\ell}\varrho(\ell,k)\lambda^k/k!$ is the ℓth moment of the Poisson distribution with parameter λ, and that the corresponding moment generating function is entire. Since a Poisson distribution is determined by its moments, it follows (see [3, Section 30]) that for any

given $m \geqslant 0$, (1.13) holds as $x \to \infty$, provided Hypothesis $(k, \Theta_{k,\lambda}, \emptyset)$ holds for all $k \geqslant 1$. □

5. Preliminaries

A positive integer n is a sum of two squares if, and only if,

$$n = 2^{\beta_2} \prod_{p \equiv 1 \bmod 4} p^{\beta_p} \prod_{p \equiv 3 \bmod 4} p^{2\beta_p},$$

where β_2, β_p denote nonnegative integers. (See [11, Theorem 366].) In view of this and the next proposition, whose proof, being routine and elementary, is omitted, we have $\mathbb{E} = \bigcap_p S_p$, where $S_p = \bigcap_{\alpha \geqslant 1} \{n \in \mathbb{Z} : n \equiv \square + \square \bmod p^\alpha\}$. Further, as $S_p = \mathbb{Z}$ for primes $p \equiv 1 \bmod 4$, we may write $\mathbb{E} = \bigcap_{p \not\equiv 1 \bmod 4} S_p$.

Proposition 5.1. *Let $n \in \mathbb{Z}$. We have $n \in S_2$ if, and only if, either $n = 0$ or $n = 2^\beta m$ for some $\beta \geqslant 0$ and $m \equiv 1 \bmod 4$. For $p \equiv 3 \bmod 4$, we have $n \in S_p$ if, and only if, either $n = 0$ or $n = p^{2\beta} m$ for some $\beta \geqslant 0$ and $m \not\equiv 0 \bmod p$. For $p \equiv 1 \bmod 4$, we have $S_p = \mathbb{Z}$.*

Let us introduce some notation in order to state further results. Given a nonempty, finite set $\boldsymbol{h} \subseteq \mathbb{Z}$, let

$$\det(\boldsymbol{h}) := \prod_{\substack{h,h' \in \boldsymbol{h} \\ h > h'}} (h - h') > 0. \tag{5.1}$$

Note that if $p \leqslant k - 1$, where $k = \#\boldsymbol{h}$, then two elements of $\boldsymbol{h}$ occupy the same congruence class modulo p, so $p \mid \det(\boldsymbol{h})$. In other words, if $p \nmid \det(\boldsymbol{h})$ then $k \leqslant p$.

Let

$$\boldsymbol{h}_p := \{h' \in \boldsymbol{h} : -h' + \boldsymbol{h} \subseteq S_p\}. \tag{5.2}$$

Note that $\boldsymbol{h}_2$ contains at most one element, for if $h, h' \in \boldsymbol{h}_2$ then $\pm(h - h') \in S_2$, which by Proposition 5.1 holds only if $h - h' = 0$. Similarly, if $k = 1$ or $k = 2$, then $\#\boldsymbol{h}_2 = 1$. By Proposition 5.1, $\boldsymbol{h}_p$ for $p \equiv 3 \bmod 4$ consists precisely of those elements h' of $\boldsymbol{h}$ for which $2 \mid \nu_p(h - h')$ for every $h \in \boldsymbol{h}$ with $h \neq h'$. (Recall that $\nu_p(n)$ denotes the p-adic valuation of n.) For instance, if $p \nmid \det(\boldsymbol{h})$ then $\boldsymbol{h}_p = \boldsymbol{h}$.

Given $\alpha \geqslant 1$, let

$$T_{\boldsymbol{h}}(2^{\alpha+1}) := \left\{ 0 \leqslant a < 2^{\alpha+1} : a+\boldsymbol{h} \subseteq S_2 \text{ and } \max_{h\in\boldsymbol{h}} \nu_2(a+h) < \alpha \right\}. \tag{5.3}$$

By Proposition 5.1, this is the (possibly empty) set of least residues a modulo $2^{\alpha+1}$ such that, for each $h \in \boldsymbol{h}$, there is some $\beta \leqslant \alpha - 1$ and $m \equiv 1 \bmod 4$ such that $a + h = 2^\beta m$. Finally, for $p \equiv 3 \bmod 4$, let

$$T_{\boldsymbol{h}}(p^\alpha) := \left\{ 0 \leqslant a < p^\alpha : a + \boldsymbol{h} \subseteq S_p \text{ and } \max_{h\in\boldsymbol{h}} \nu_p(a+h) < \alpha \right\}. \tag{5.4}$$

This is the (possibly empty) set of least residues a modulo p^α such that, for each $h \in \boldsymbol{h}$, there exists $\beta \leqslant (\alpha-1)/2$ for which $p^{2\beta} \parallel a + h$. Note that, for $\alpha \geqslant 2$ and odd p, the difference between $T_{\boldsymbol{h}}(2^\alpha)$ and $T_{\boldsymbol{h}}(p^\alpha)$ is that $T_{\boldsymbol{h}}(2^\alpha)$ contains only integers a for which $\max_{h\in\boldsymbol{h}} \nu_2(a+h) \leqslant \alpha - 2$, whereas $T_{\boldsymbol{h}}(p^\alpha)$ contains a for which $\max_{h\in\boldsymbol{h}} \nu_p(a+h) \leqslant \alpha - 1$. As may be expected in view of Proposition 5.1, we will need to treat $p = 2$ as a separate case throughout.

Recall from (1.6) that $\delta_{\boldsymbol{h}}(p) := \lim_{\alpha\to\infty} \#S_{\boldsymbol{h}}(p^\alpha)/p^\alpha$, where

$$S_{\boldsymbol{h}}(p^\alpha) := \{0 \leqslant a < p^\alpha : \forall h \in \boldsymbol{h}, a + h \equiv \square + \square \bmod p^\alpha\}.$$

We have introduced $T_{\boldsymbol{h}}(p^\alpha)$ because it is more convenient than $S_{\boldsymbol{h}}(p^\alpha)$ to work with. It is not difficult to see that, for $p \not\equiv 1 \bmod 4$, $0 \leqslant \#S_{\boldsymbol{h}}(p^\alpha) - \#T_{\boldsymbol{h}}(p^\alpha) \leqslant 1$ once α is sufficiently large. (One may verify Proposition 5.1 by showing that $n \equiv \square + \square \bmod 2^\alpha$ if, and only if, $n \equiv 2^\beta m \bmod 2^\alpha$ for some $\beta \geqslant 0$ and odd m, and, for $p \equiv 3 \bmod 4$, that $n \equiv \square + \square \bmod p^\alpha$ if, and only if, $n \equiv p^{2\beta} m \bmod p^\alpha$ for some $\beta \geqslant 0$ and $m \not\equiv 0 \bmod p$.) Thus, the limit $\delta_{\boldsymbol{h}}(p)$ exists if, and only if, $\lim_{\alpha\to\infty} \#T_{\boldsymbol{h}}(p^\alpha)/p^\alpha$ exists, in which case the two limits are equal.

In the next two propositions, and throughout, we allow for the possibility that $k = 1$. In case $\boldsymbol{h} = \{h_1\}$, we define $\max_{i\neq j} \nu_p(h_i - h_j)$ to be zero (and $\det(\boldsymbol{h}) := 1$).

Proposition 5.2. *Let $\boldsymbol{h} = \{h_1, \dots, h_k\}$ be a set of $k \geqslant 1$ distinct integers.*

(a) The limits $\delta_{\boldsymbol{h}}(2)$ (see (1.6)) and $\lim_{\alpha\to\infty} \#T_{\boldsymbol{h}}(2^{\alpha+1})/2^{\alpha+1}$ exist, and are equal:

$$\delta_{\boldsymbol{h}}(2) = \lim_{\alpha\to\infty} \frac{\#T_{\boldsymbol{h}}(2^{\alpha+1})}{2^{\alpha+1}}. \tag{5.5}$$

Moreover, for all $\alpha \geqslant 1$, we have

$$\left| \frac{\#T_{\boldsymbol{h}}(2^{\alpha+1})}{2^{\alpha+1}} - \delta_{\boldsymbol{h}}(2) \right| \leqslant \frac{k}{2^{\alpha}}. \tag{5.6}$$

(*b*) *For any $\alpha \geqslant 2 + \max_{i \neq j} \nu_2(h_i - h_j)$, we have*

$$\delta_{\boldsymbol{h}}(2) = \frac{\#T_{\boldsymbol{h}}(2^{\alpha+1}) + \#\boldsymbol{h}_2}{2^{\alpha+1}}, \tag{5.7}$$

the right-hand side being constant for α in this range.

(*c*) *If $2 \nmid \det(\boldsymbol{h})$ (in which case $k \leqslant 2$), then $\delta_{\boldsymbol{h}}(2) = (1/2)^k$. As a special case, we record here that $\delta_{\{0\}}(2) = 1/2$.*

Proof. In essence, we use a Hensel-type argument: for $\alpha \geqslant 1$, the condition that $n \equiv \square + \square \bmod 2^{\alpha}$ can be lifted to $n \equiv \square + \square \bmod 2^{\alpha+1}$, unless $n = 2^{\alpha} m$ for some $m \equiv 3 \bmod 4$.

(a) As already noted, to show that $\delta_{\boldsymbol{h}}(2)$ and the right-hand side of (5.5) exist and are equal, it suffices to show that the right-hand side of (5.5) exists. Let $\alpha \geqslant 1$, and let $0 \leqslant b < 2^{\alpha+2}$, so $b = a + 2^{\alpha+1} q$, where $0 \leqslant a < 2^{\alpha+1}$ and either $q = 0$ or $q = 1$. Suppose that, for each i, there exists $\beta_i \leqslant \alpha - 1$ and $m_i \equiv \pm 1 \bmod 4$ such that $b + h_i = 2^{\beta_i} m_i$. Then, for each i, $a + h_i = 2^{\beta_i} m_i'$ and $a + 2^{\alpha+1} + h_i = 2^{\beta_i} m_i''$, where $m_i' \equiv m_i'' \equiv m_i \bmod 4$. Recalling Proposition 5.1 and definition (5.3), we see that the following statements are equivalent: (i) $b \in T_{\boldsymbol{h}}(2^{\alpha+2})$; (ii) both a and $a + 2^{\alpha+1}$ are in $T_{\boldsymbol{h}}(2^{\alpha+2})$; (iii) $a \in T_{\boldsymbol{h}}(2^{\alpha+1})$.

We have shown that we have a partition

$$T_{\boldsymbol{h}}(2^{\alpha+2}) = \{a, a + 2^{\alpha+1} : a \in T_{\boldsymbol{h}}(2^{\alpha+1})\} \cup U_{\boldsymbol{h}}(2^{\alpha+2}),$$

where

$$U_{\boldsymbol{h}}(2^{\alpha+2}) := \left\{ 0 \leqslant b < 2^{\alpha+2} : b + \boldsymbol{h} \subseteq S_2 \text{ and } \max_{h \in \boldsymbol{h}} \nu_2(b + h) = \alpha \right\}$$

is the set of elements b of $T_{\boldsymbol{h}}(2^{\alpha+2})$ for which $\nu_2(b + h_j) = \alpha$ for some $h_j \in \boldsymbol{h}$. Any element of $U_{\boldsymbol{h}}(2^{\alpha+2})$ is a least residue of $\pm 2^{\alpha} - h_j$ for some $h_j \in \boldsymbol{h}$, of which there are at most $2k$. We see that

$$\frac{\#T_{\boldsymbol{h}}(2^{\alpha+2})}{2^{\alpha+2}} - \frac{\#T_{\boldsymbol{h}}(2^{\alpha+1})}{2^{\alpha+1}} = \frac{\#U_{\boldsymbol{h}}(2^{\alpha+2})}{2^{\alpha+2}} \leqslant \frac{k}{2^{\alpha+1}}.$$

Consequently, for any β with $\beta \geqslant \alpha$, we have

$$0 \leqslant \frac{\#T_{\boldsymbol{h}}(2^{\beta+1})}{2^{\beta+1}} - \frac{\#T_{\boldsymbol{h}}(2^{\alpha+1})}{2^{\alpha+1}} = \sum_{r=1}^{\beta-\alpha} \frac{\#U_{\boldsymbol{h}}(2^{\alpha+r+1})}{2^{\alpha+r+1}} < \frac{k}{2^{\alpha}}.$$

It follows that the limit on the right-hand side of (5.5) exists, and that (5.6) holds for all $\alpha \geqslant 1$.

(b) Assume that $\alpha \geqslant 2 + \max_{i \neq j} \nu_2(h_i - h_j)$. Suppose that, for some j, there exists q such that $b + h_j = 2^{\alpha}(1 + 2q)$. We have $b + h_j \in S_2$ if, and only if, $2 \mid q$, equivalently, $b + h_j \equiv 2^{\alpha} \bmod 2^{\alpha+2}$. For $i \neq j$ we may write $h_i - h_j = 2^{\beta_{ij}} m_{ij}$ with $\beta_{ij} \leqslant \alpha - 2$ and $m_{ij} \equiv \pm 1 \bmod 4$. Thus,

$$b + h_i = 2^{\beta_{ij}}(m_{ij} + 2^{\alpha - \beta_{ij}}(1 + 2q))$$

is in S_2 if, and only if, $m_{ij} \equiv 1 \bmod 4$, equivalently, $h_i - h_j \in S_2$. By definition of $\boldsymbol{h}_2$, this holds for each $i \neq j$ if, and only if, $h_j \in \boldsymbol{h}_2$. We have shown that $b \in T_{\boldsymbol{h}}(2^{\alpha+2})$ and $\nu_2(b + h_j) = \alpha$ for some $h_j \in \boldsymbol{h}$ if, and only if, $\boldsymbol{h}_2$ is nonempty, h_j is the (necessarily unique) element of $\boldsymbol{h}_2$, and $b + h_j \equiv 2^{\alpha} \bmod 2^{\alpha+2}$. Thus,

$$U_{\boldsymbol{h}}(2^{\alpha+2}) = \{0 \leqslant b < 2^{\alpha+2} : \exists h' \in \boldsymbol{h}_2, b \equiv 2^{\alpha} - h' \bmod 2^{\alpha+2}\},$$

and $\#U_{\boldsymbol{h}}(2^{\alpha+2}) = \#\boldsymbol{h}_2$. Also, $\#T_{\boldsymbol{h}}(2^{\alpha+2}) = 2\#T_{\boldsymbol{h}}(2^{\alpha+1}) + \#\boldsymbol{h}_2$. Hence

$$\frac{\#T_{\boldsymbol{h}}(2^{\alpha+2}) + \#\boldsymbol{h}_2}{2^{\alpha+2}} = \frac{\#T_{\boldsymbol{h}}(2^{\alpha+1}) + \#\boldsymbol{h}_2}{2^{\alpha+1}}.$$

(c) Suppose $2 \nmid \det(\boldsymbol{h})$. If $k = 1$, i.e. if $\boldsymbol{h} = \{h_1\}$, then the elements of $T_{\boldsymbol{h}}(8)$ are precisely the least residues of $1 - h_1, 2 - h_1$ and $5 - h_1$ modulo 8. Also, $\boldsymbol{h}_2 = \boldsymbol{h}$. If $k = 2$, i.e. if $\boldsymbol{h} = \{h_1, h_2\}$, then either $h_2 - h_1 \equiv 1 \bmod 4$ or $h_1 - h_2 \equiv 1 \bmod 4$. Without loss of generality, suppose $h_2 - h_1 \equiv 1 \bmod 4$. Then the sole element of $T_{\boldsymbol{h}}(8)$ is the least residue of $h_2 - 2h_1$ modulo 8. Also, $\boldsymbol{h}_2 = \{h_1\}$. Therefore, by (b), $\delta_{\boldsymbol{h}}(2) = (1/2)^k$. □

For the next proposition, recall that $\alpha \bmod 2$, when written in an exponent, denotes the least residue of α modulo 2. For instance, $p^{\alpha \bmod 2} = 1$ if α is even.

Proposition 5.3. *Let $\boldsymbol{h} = \{h_1, \dots, h_k\}$ be a set of $k \geqslant 1$ distinct integers, and let p be a prime with $p \equiv 3 \bmod 4$.*

(*a*) *The limits* $\delta_{\boldsymbol{h}}(p)$ (*see* (1.6)) *and* $\lim_{\alpha\to\infty} \#T_{\boldsymbol{h}}(p^\alpha)/p^\alpha$ *exist, and are equal:*

$$\delta_{\boldsymbol{h}}(p) = \lim_{\alpha\to\infty} \frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha}. \tag{5.8}$$

Moreover, for all $\alpha \geqslant 1$, *we have*

$$\left|\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} - \delta_{\boldsymbol{h}}(p)\right| \leqslant \frac{k}{p^\alpha}\left(1+\frac{1}{p}\right)^{-1}\frac{1}{p^{\alpha \bmod 2}}. \tag{5.9}$$

(*b*) *For any* $\alpha \geqslant 1 + \max_{i\neq j} \nu_p(h_i - h_j)$, *we have*

$$\delta_{\boldsymbol{h}}(p) = \frac{1}{p^\alpha}\left(\#T_{\boldsymbol{h}}(p^\alpha) + \#\boldsymbol{h}_p\left(1+\frac{1}{p}\right)^{-1}\frac{1}{p^{\alpha \bmod 2}}\right), \tag{5.10}$$

the right-hand side being constant for α *in this range.*

(*c*) *We have*

$$\delta_{\boldsymbol{h}}(p) \geqslant \left(1+\frac{1}{p}\right)^{-1}\left(1-\frac{\min\{k-1,p\}}{p}\right), \tag{5.11}$$

with ***equality*** *attained if* $p \nmid \det(\boldsymbol{h})$ (*in which case* $k \leqslant p$). *As a special case, we record here that* $\delta_{\{0\}}(p) = (1+1/p)^{-1}$.

Proof. (a) As noted above the statement of Proposition 5.2, to show that $\delta_{\boldsymbol{h}}(p)$ and the right-hand side of (5.8) exist and are equal, it suffices to show that the right-hand side of (5.8) exists. Let $\alpha \geqslant 1$ and let $0 \leqslant b < p^{\alpha+1}$. Thus, $b = a + p^\alpha q$, where $0 \leqslant a < p^\alpha$ and $0 \leqslant q < p$. Suppose that, for each i, there exists $\beta_i \leqslant \alpha - 1$ and $m_i \not\equiv 0 \bmod p$ such that $b + h_i = p^{\beta_i} m_i$. Then, for each i and each q', $0 \leqslant q' < p$, we have $a + p^\alpha q' + h_i = p^{\beta_i} m_i'$, where $m_i' \equiv m_i \not\equiv 0 \bmod p$. Recalling Proposition 5.1 and definition (5.4), we see that the following are equivalent: (i) $b \in T_{\boldsymbol{h}}(p^{\alpha+1})$; (ii) $a + p^\alpha q' + h_i \in T_{\boldsymbol{h}}(p^{\alpha+1})$ for $0 \leqslant q' < p$; (iii) $a \in T_{\boldsymbol{h}}(p^\alpha)$.

We have shown that we have a partition

$$T_{\boldsymbol{h}}(p^{\alpha+1}) = \{a + p^\alpha q : a \in T_{\boldsymbol{h}}(p^\alpha), 0 \leqslant q < p\} \cup U_{\boldsymbol{h}}(p^{\alpha+1}),$$

where

$$U_{\boldsymbol{h}}(p^{\alpha+1}) := \left\{0 \leqslant b < p^{\alpha+1} : b + \boldsymbol{h} \subseteq S_p \text{ and } \max_{h\in\boldsymbol{h}} \nu_p(b+h) = \alpha\right\}$$

is the set of elements b of $T_{\boldsymbol{h}}(p^{\alpha+1})$ for which $\nu_p(b+h_j)=\alpha$ for some $h_j\in\boldsymbol{h}$. Plainly, $U_{\boldsymbol{h}}(p^{\alpha+1})$ is empty if α is odd. (If $b+\boldsymbol{h}\subseteq S_p$ then, by Proposition 5.1, $\nu_p(b+h_j)$ is even, and hence not equal to any odd α.) Also, any element of $U_{\boldsymbol{h}}(p^{\alpha+1})$ is a least residue of $p^\alpha q-h_j \bmod p^{\alpha+1}$, for some $0<q<p$ and $h_j\in\boldsymbol{h}$, of which there are at most $(p-1)k$. We see that

$$\frac{\#T_{\boldsymbol{h}}(p^{\alpha+1})}{p^{\alpha+1}}-\frac{\#T_{\boldsymbol{h}}(p^{\alpha})}{p^{\alpha}}=\frac{\#U_{\boldsymbol{h}}(p^{\alpha+1})}{p^{\alpha+1}}, \tag{5.12}$$

and that

$$0\leqslant\frac{\#U_{\boldsymbol{h}}(p^{\alpha+1})}{p^{\alpha+1}}\leqslant\Big(1-\frac{1}{p}\Big)\frac{k}{p^{\alpha}}, \tag{5.13}$$

with *equality* on the left if α is *odd*. Consequently, for any β with $\beta\geqslant\alpha$, we have

$$0\leqslant\frac{\#T_{\boldsymbol{h}}(p^{\beta})}{p^{\beta}}-\frac{\#T_{\boldsymbol{h}}(p^{\alpha})}{p^{\alpha}}=\sum_{r=1}^{\beta-\alpha}\frac{\#U_{\boldsymbol{h}}(p^{\alpha+r})}{p^{\alpha+r}}<\Big(1-\frac{1}{p}\Big)\frac{k}{p^{\alpha}}\sum_{\substack{r-1\geqslant 0\\ r-1\equiv\alpha \bmod 2}}\frac{1}{p^{r-1}}.$$

Since this last sum is equal to $1/(1-1/p^2)$ if α is even, and to $1/(p(1-1/p^2))$ if α is odd, we have

$$0\leqslant\frac{\#T_{\boldsymbol{h}}(p^{\beta})}{p^{\beta}}-\frac{\#T_{\boldsymbol{h}}(p^{\alpha})}{p^{\alpha}}<\frac{k}{p^{\alpha}}\Big(1+\frac{1}{p}\Big)^{-1}\frac{1}{p^{\alpha \bmod 2}}.$$

It follows that the limit on the right-hand side of (5.8) exists, and that (5.9) holds for all $\alpha\geqslant 1$.

(b) Let $0\leqslant b<p^{\alpha+1}$, and assume now that $\alpha\geqslant 1+\max_{i\neq j}\nu_p(h_i-h_j)$. Suppose that, for some j, we have $b+h_j=p^\alpha m_j$ for some $m_j\not\equiv 0 \bmod p$. We have $b+h_j\in S_p$ if, and only if, α is even. Let $i\neq j$. We may write $h_i-h_j=p^{\beta_{ij}}m_{ij}$ with $\beta_{ij}\leqslant\alpha-1$ and $m_{ij}\not\equiv 0 \bmod p$. Thus, $b+h_i=p^{\beta_{ij}}(m_{ij}+p^{\alpha-\beta_{ij}}m_j)$ is in S_p if, and only if, β_{ij} is even, equivalently, $h_i-h_j\in S_p$. By definition of $\boldsymbol{h}_p$, this holds for each $i\neq j$ if, and only if, $h_j\in\boldsymbol{h}_p$. In that case, for $0\leqslant q'<p$ with $q'\not\equiv -m_j \bmod p$, we have $b+p^\alpha q'+h_i\in S_p$ and $\nu_p(b+p^\alpha q'+h_i)=\beta_{ij}<\alpha$ for $i\neq j$; $b+p^\alpha q'+h_j\in S_p$ if, and only if, $b+h_j\in S_p$, and $\nu_p(b+p^\alpha q'+h_j)=\alpha$. For $q'\equiv -m_j \bmod p$, $\nu_p(b+p^\alpha q'+h_j)>\alpha$.

Thus, if $U_{\boldsymbol{h}}(p^{\alpha+1})\neq\emptyset$, then α is even and $\boldsymbol{h}_p\neq\emptyset$; and if $b\in U_{\boldsymbol{h}}(p^{\alpha+1})$, then the h_j for which $\nu_p(b+h_j)=\alpha$ is uniquely determined by b and must lie in $\boldsymbol{h}_p$. If α is even, then, writing $h_j=p^\alpha q_j+r_j$, with $0\leqslant r_j<p^\alpha$, we see

that

$$U_{\boldsymbol{h}}(p^{\alpha+1}) = \bigcup_{h_j \in \boldsymbol{h}_p} \{p^\alpha(q'+1) - r_j : 0 \leqslant q' < p, q' \not\equiv -q_j \bmod p\}.$$

Thus, $\#T_{\boldsymbol{h}}(p^{\alpha+1}) = p\#T_{\boldsymbol{h}}(p^\alpha)$ if α is odd, and $\#T_{\boldsymbol{h}}(p^{\alpha+1}) = p\#T_{\boldsymbol{h}}(p^\alpha) + (p-1)\#\boldsymbol{h}_p$ if α is even. Consequently, if α is odd, then

$$\frac{1}{p^{\alpha+1}}\left(\#T_{\boldsymbol{h}}(p^{\alpha+1}) + \#\boldsymbol{h}_p\frac{p}{p+1}\right) = \frac{1}{p^\alpha}\left(\#T_{\boldsymbol{h}}(p^\alpha) + \#\boldsymbol{h}_p\frac{1}{p+1}\right),$$

while if α is even, then

$$\frac{1}{p^{\alpha+1}}\left(\#T_{\boldsymbol{h}}(p^{\alpha+1}) + \#\boldsymbol{h}_p\frac{1}{p+1}\right) = \frac{1}{p^\alpha}\left(\#T_{\boldsymbol{h}}(p^\alpha) + \#\boldsymbol{h}_p\frac{p}{p+1}\right).$$

(c) Note that $T_{\boldsymbol{h}}(p) = \{0 \leqslant a < p : \forall i, a \not\equiv -h_i \bmod p\}$, so $\#T_{\boldsymbol{h}}(p) = p - \kappa$ where κ is the number of distinct congruence classes in $\{h_i \bmod p : h_i \in \boldsymbol{h}\}$. Thus, $\kappa = k$ if, and only if, $p \nmid \det(\boldsymbol{h})$. First, consider the case $p \mid \det(\boldsymbol{h})$, i.e. $\kappa \leqslant k-1$. As $\delta_{\boldsymbol{h}}(p) \geqslant 0$, (5.11) is trivial for $p \leqslant k-1$, so let us assume that $k \leqslant p$. The relation (5.12) shows that $\#T_{\boldsymbol{h}}(p^{\alpha+1})/p^{\alpha+1} \geqslant \#T_{\boldsymbol{h}}(p^\alpha)/p^\alpha$ for $\alpha \geqslant 1$, and hence

$$\delta_{\boldsymbol{h}}(p) \geqslant \frac{\#T_{\boldsymbol{h}}(p)}{p} \geqslant \frac{p-(k-1)}{p} > 1 - \frac{k}{p+1}.$$

The right-hand side of (5.11) is equal to $1 - k/(p+1)$ when $\min\{k-1, p\} = k-1$, as we are currently assuming. Next, consider the case $p \nmid \det(\boldsymbol{h})$, i.e. $\kappa = k$. In this case, we have $\#\boldsymbol{h} = \#\boldsymbol{h}_p = k$, and, by (5.10),

$$\delta_{\boldsymbol{h}}(p) = \frac{1}{p}\left(\#T_{\boldsymbol{h}}(p) + \#\boldsymbol{h}_p\left(1+\frac{1}{p}\right)^{-1}\frac{1}{p}\right) = \left(1+\frac{1}{p}\right)^{-1}\left(1 - \frac{k-1}{p}\right),$$

which is equal to the right-hand side of (5.11) (since $p \geqslant \kappa = k$). □

Notice that, for all $p \not\equiv 1 \bmod 4$, we have $0 \leqslant \delta_{\boldsymbol{h}}(p) \leqslant 1$, by definition. By the following proposition, the nonvanishing of

$$\mathfrak{S}_{\boldsymbol{h}} := \prod_{p \not\equiv 1 \bmod 4} \delta_{\{0\}}(p)^{-k}\delta_{\boldsymbol{h}}(p)$$

(the singular series for $\boldsymbol{h}$ [see (1.7)]), is equivalent to $\delta_{\boldsymbol{h}}(p) > 0$ for all $p \not\equiv 1 \bmod 4$.

Proposition 5.4. *Let $\boldsymbol{h} = \{h_1, \dots, h_k\}$ be a set of $k \geqslant 1$ distinct integers. We have*

$$\mathrm{e}^{-(k-1)} \leqslant \prod_{\substack{p \not\equiv 1 \bmod 4 \\ p \nmid \det(\boldsymbol{h})}} \delta_{\{0\}}(p)^{-k} \delta_{\boldsymbol{h}}(p) \leqslant 1, \tag{5.14}$$

and the product converges. Consequently,

$$\begin{aligned} &\frac{2^k \delta_{\boldsymbol{h}}(2)}{\mathrm{e}^{k-1}} \prod_{\substack{p \equiv 3 \bmod 4 \\ p \mid \det(\boldsymbol{h})}} \left(\left(1 + \frac{1}{p}\right)^k \delta_{\boldsymbol{h}}(p) \right) \\ &\qquad \leqslant \mathfrak{S}_{\boldsymbol{h}} \leqslant 2^k \delta_{\boldsymbol{h}}(2) \prod_{\substack{p \equiv 3 \bmod 4 \\ p \mid \det(\boldsymbol{h})}} \left(\left(1 + \frac{1}{p}\right)^k \delta_{\boldsymbol{h}}(p) \right). \end{aligned} \tag{5.15}$$

Proof. If $2 \nmid \det(\boldsymbol{h})$, then $k \leqslant 2$ and $\delta_{\{0\}}(2)^{-k} \delta_{\boldsymbol{h}}(2) = 1$ by Proposition 5.2 (c), so only the primes $p \equiv 3 \bmod 4$ have any bearing on the product in (5.14). Let $p \equiv 3 \bmod 4$, and suppose $p \nmid \det(\boldsymbol{h})$. By Proposition 5.3 (c), $k \leqslant p$ and

$$\delta_{\{0\}}(p)^{-k} \delta_{\boldsymbol{h}}(p) = \left(1 + \frac{1}{p}\right)^{k-1} \left(1 - \frac{k-1}{p}\right). \tag{5.16}$$

Thus, $\delta_{\{0\}}(p)^{-k} \delta_{\boldsymbol{h}}(p) = 1 + O_k(1/p^2)$, and consequently the product in (5.14) converges.

More precisely, from (5.16) we have, on the one hand,

$$\delta_{\{0\}}(p)^{-k} \delta_{\boldsymbol{h}}(p) = 1 - \sum_{j=2}^{k} \left\{ (k-1) \binom{k-1}{j-1} - \binom{k-1}{j} \right\} p^{-j} \leqslant 1,$$

with equality attained if $k = 1$, which gives the upper bound in (5.14), and also the lower bound for $k = 1$. On the other hand we have

$$\delta_{\{0\}}(p)^{-k} \delta_{\boldsymbol{h}}(p) \geqslant 1 - \frac{(k-1)^2}{p^2}.$$

For $k = 2$ we see that the product in (5.14) is at least $\prod_{p \equiv 3 \bmod 4} (1 - 1/p^2)$, which is equal to $1/(2C^2) = 0.856108\dots$ (with C being the Landau–Ramanujan constant; see (1.5)), and is greater than e^{-1}. For $k \geqslant 3$ we apply

the basic inequality $\log(1-x) \geqslant -x/(1-x)$ $(0 \leqslant x < 1)$ to the above, obtaining

$$\begin{aligned}\log \delta_{\{0\}}(p)^{-k}\delta_{\boldsymbol{h}}(p) &\geqslant -\frac{(k-1)^2}{p^2}\left(1-\frac{(k-1)^2}{p^2}\right)^{-1}\\ &\geqslant -\frac{(k-1)^2}{p^2}\left(1-\frac{(k-1)^2}{k^2}\right)^{-1}\end{aligned}$$

(since $k \leqslant p$). Noting that $-\sum_{p\nmid \det(\boldsymbol{h})} 1/p^2 \geqslant -\sum_{n\geqslant k} 1/n^2 \geqslant -1/(k-1)^2$, and that $-(1-(k-1)^2/k^2)^{-1} = -k^2/(2k-1) > -(k-1)$, then exponentiating, we see that product in (5.14) is greater than $\mathrm{e}^{-(k-1)}$. The inequalities in (5.15) follow upon recalling that $\delta_{\{0\}}(p) = (1+1/p)^k$ for $p \equiv 3 \bmod 4$ (see Proposition 5.3 (c)), and again that $\delta_{\{0\}}(2)^{-k}\delta_{\boldsymbol{h}}(2) = 1$ if $2 \nmid \det(\boldsymbol{h})$ (see Proposition 5.2 (c)). □

6. Proof of Proposition 1.3

We will make use of the following elementary bounds. Recall that, for $n \in \mathbb{N}$, $\omega(n) := \#\{p : p \mid n\}$, $\mathrm{rad}(n) := \prod_{p\mid n} p$, and $\mathrm{sf}(n) := \prod_{p\parallel n} p$.

Lemma 6.1. *Let*

$$\mathcal{N} := \{ab^2\,\mathrm{rad}(b) : a, b \in \mathbb{N}, (a,b) = 1, a \text{ squarefree}\}. \tag{6.1}$$

Fix any number $A \geqslant 1$. For $y \geqslant 1$ and integers $D \geqslant 1$, we have

$$\sum_{\substack{n\in\mathcal{N}\\ n>y}} A^{\omega(n)}\frac{(D,\mathrm{rad}(n))}{n\,\mathrm{sf}(n)} \ll_A (1+A)^{2\omega(D)}\frac{y^{O(1/\log\log 3y)}}{y^{2/3}}, \tag{6.2}$$

and

$$\sum_{\substack{n\in\mathcal{N}\\ n\leqslant y}} \frac{A^{\omega(n)}}{\mathrm{sf}(n)} \ll_A y^{1/3+O(1/\log\log 3y)}. \tag{6.3}$$

Proof. Let $y \geqslant 1$ and let $D \geqslant 1$. We claim that the following four bounds hold:

$$\sum_{\substack{n>y\\ \text{squarefree}}} A^{\omega(n)}\frac{(D,n)}{n^2} \ll_A (1+A)^{\omega(D)}\frac{y^{O(1/\log\log 3y)}}{y}; \tag{6.4}$$

$$\sum_{\substack{n\leqslant y\\ \text{squarefree}}} A^{\omega(n)}\frac{(D,n)}{n} \ll_A (1+A)^{\omega(D)} y^{O(1/\log\log 3y)}; \tag{6.5}$$

$$\sum_{n^2\operatorname{rad}(n)>y} \frac{A^{\omega(n)}(D,\operatorname{rad}(n))}{n^2\operatorname{rad}(n)} \ll_A (1+A)^{\omega(D)}\frac{y^{O(1/\log\log 3y)}}{y^{2/3}}; \tag{6.6}$$

and

$$\sum_{n^2\operatorname{rad}(n)\leqslant y} A^{\omega(n)} \ll_A y^{1/3+O(1/\log\log 3y)}. \tag{6.7}$$

Let us deduce (6.2) and (6.3). The left-hand side of (6.2) is at most

$$\sum_{\substack{a\leqslant y^{2/3}\\ \text{squarefree}}} A^{\omega(a)}\frac{(D,a)}{a^2} \sum_{b^2\operatorname{rad}(b)>y/a} \frac{A^{\omega(b)}(D,\operatorname{rad}(b))}{b^2\operatorname{rad}(b)}$$
$$+ \sum_{\substack{a>y^{2/3}\\ \text{squarefree}}} A^{\omega(a)}\frac{(D,a)}{a^2}\sum_{b\geqslant 1}\frac{A^{\omega(b)}}{b^2}.$$

By (6.5) and (6.6), the first double sum is

$$\ll_A (1+A)^{\omega(D)} y^{-2/3+o(1)} \sum_{\substack{a\leqslant y^{2/3}\\ \text{squarefree}}} A^{\omega(a)}\frac{(D,a)}{a^{4/3}}$$
$$\ll_A (1+A)^{2\omega(D)}\frac{y^{O(1/\log\log 3y)}}{y^{2/3}}.$$

By (6.4), and since $\sum_{b\geqslant 1}(A^{\omega(b)}/b^2) \ll_A 1$,

$$\sum_{\substack{a>y^{2/3}\\ \text{squarefree}}} A^{\omega(a)}\frac{(D,a)}{a^2}\sum_{b\geqslant 1}\frac{A^{\omega(b)}}{b^2} \ll_A (1+A)^{\omega(D)}\frac{y^{O(1/\log\log 3y)}}{y^{2/3}}.$$

Combining gives (6.2). The left-hand side of (6.3) is at most

$$\sum_{\substack{a\leqslant y\\ \text{squarefree}}} \frac{A^{\omega(a)}}{a} \sum_{b^2\operatorname{rad}(b)\leqslant y} A^{\omega(b)};$$

applying (6.5) and (6.7) gives (6.3).

We now prove our claim. For (6.4), we first consider the case $D = 1$. Note that

$$\sum_{\substack{n_1 \leqslant y \\ \text{squarefree}}} \frac{(A-1)^{\omega(n_1)}}{n_1} \leqslant \prod_{p \leqslant y} \left(1 + \frac{A-1}{p}\right) \leqslant \prod_{p \leqslant y} \left(1 + \frac{1}{p}\right)^{A-1} \ll_A (\log 3y)^{A-1}, \tag{6.8}$$

because $1 + 1/p < \mathrm{e}^{1/p}$ and $\sum_{p \leqslant y} 1/p = \log\log 3y + O(1)$ Mertens' theorem. Now,

$$\sum_{\substack{n > y \\ \text{squarefree}}} \frac{A^{\omega(n)}}{n^2} = \sum_{\substack{n > y \\ \text{squarefree}}} \frac{1}{n^2} \sum_{n_1 | n} (A-1)^{\omega(n_1)} \leqslant \sum_{\substack{n_1 \geqslant 1 \\ \text{squarefree}}} \frac{(A-1)^{\omega(n_1)}}{n_1^2} \sum_{\substack{m > y/n_1 \\ \text{squarefree}}} \frac{1}{m^2},$$

the inner sum being $O(n_1/y)$ for $n_1 \leqslant y$ and $O(1)$ for $n_1 > y$. Thus,

$$\sum_{\substack{n > y \\ \text{squarefree}}} \frac{A^{\omega(n)}}{n^2} \ll_A \frac{(\log 3y)^{A-1}}{y} + \sum_{\substack{n_1 > y \\ \text{squarefree}}} \frac{(A-1)^{\omega(n_1)}}{n_1^2}.$$

If $A \leqslant 2$, then this last sum is $O(1/y)$; otherwise, repeating the argument as many times as necessary gives

$$\sum_{\substack{n > y \\ \text{squarefree}}} \frac{A^{\omega(n)}}{n^2} \ll_A \frac{(\log 3y)^{A-1}}{y}.$$

It follows that, for any integer $d \geqslant 1$,

$$\sum_{\substack{n > y,\, d | n \\ \text{squarefree}}} \frac{A^{\omega(n)}}{n^2} \ll_A \frac{A^{\omega(d)}}{d} \cdot \frac{(\log 3y)^{A-1}}{y}.$$

For any integer $D \geqslant 1$, we trivially have $(D,n) \leqslant \sum_{d|D,\, d|n} d$, and hence

$$\sum_{\substack{n>y\\ \text{squarefree}}} A^{\omega(n)} \frac{(D,n)}{n^2} \leqslant \sum_{\substack{d|D\\ \text{squarefree}}} \sum_{\substack{n>y,\, d|n\\ \text{squarefree}}} \frac{A^{\omega(n)}}{n^2} \ll_A \frac{(\log 3y)^{A-1}}{y} \sum_{\substack{d|D\\ \text{squarefree}}} \frac{A^{\omega(d)}}{d}.$$

Since $\sum_{d|D,\,\text{squarefree}} A^{\omega(d)} = (1+A)^{\omega(D)}$ and $(\log 3y)^{A-1} \ll_A y^{O(1/\log\log 3y)}$, this gives (6.4). The bound (6.5) follows from (6.8) and $(D,n) \leqslant \sum_{d|D,\, d|n} d$.

For (6.6), we use the following ancillary bound. We have

$$\sum_{\substack{n>y\\ \mathrm{rad}(n)=m}} \frac{1}{n} \ll \frac{y^{O(1/\log\log 3y)}}{y}, \tag{6.9}$$

uniformly for integers squarefree integers $m \geqslant 1$. To establish (6.9), we use an estimate involving smooth numbers: for $y \geqslant z \geqslant 2$, let

$$\Psi(y,z) := \#\{n \leqslant y : p \mid n \Rightarrow p \leqslant z\}$$

denote the number of z-smooth positive integers $n \leqslant y$. The following can be found in [9, (1.19)]: for $y \geqslant z \geqslant 2$,

$$\log \Psi(y,z) = \left(\frac{\log y}{\log z}\right) g\left(\frac{z}{\log y}\right)\left(1 + O\left(\frac{1}{\log z} + \frac{1}{\log\log x}\right)\right), \tag{6.10}$$

where $g(w) = \log(1+w) + w\log(1+1/w) \leqslant w+1$ $(w>0)$. Noting that

$$\sum_{\substack{n\geqslant 1\\ \mathrm{rad}(n)=m}} \frac{1}{n^{1/2}} = \frac{1}{m^{1/2}} \sum_{\substack{n\geqslant 1\\ \mathrm{rad}(n)|m}} \frac{1}{n^{1/2}} = \frac{1}{m^{1/2}} \prod_{p|m}\left(\sum_{a\geqslant 0} \frac{1}{p^{a/2}}\right) = \prod_{p|m}\left(\frac{1}{p^{1/2}-1}\right),$$

we see that

$$\sum_{\substack{n>y^2\\ \mathrm{rad}(n)=m}} \frac{1}{n} \leqslant \sum_{\substack{n>y^2\\ \mathrm{rad}(n)=m}} \frac{1}{n}\left(\frac{n}{y^2}\right)^{1/2} \leqslant \frac{1}{y} \sum_{\substack{n\geqslant 1\\ \mathrm{rad}(n)=m}} \frac{1}{n^{1/2}} \ll \frac{1}{y}. \tag{6.11}$$

If $m > y^2$, then $\sum_{n>y,\,\mathrm{rad}(n)=m} 1/n = \sum_{n>y^2,\,\mathrm{rad}(n)=m} 1/n$, and we are done. Let us assume, then, that $y^2 \geqslant m$. Let $\ell_1, \dots, \ell_r$ denote the prime divisors of m, and let $p_1 = 2 < p_2 = 3 < \cdots < p_r$ denote the r smallest primes. Note that $\#\{(\alpha_1, \dots, \alpha_r) \in \mathbb{N}^r : \ell_1^{\alpha_1} \cdots \ell_r^{\alpha_r} \leqslant y^2\} \leqslant \#\{(\alpha_1, \dots, \alpha_r) \in \mathbb{N}^r : p_1^{\alpha_1} \cdots p_r^{\alpha_r} \leqslant y^2\}$, i.e. note that

$$\#\{n \leqslant y^2 : \mathrm{rad}(n) = m\} \leqslant \#\{n \leqslant y^2 : \mathrm{rad}(n) = p_1 \cdots p_r\}.$$

Since $y^2 \geqslant m \geqslant p_1 \cdots p_r$, we have $4 \log y^2 \geqslant 4 \log m \geqslant 4 \log(p_1 \cdots p_r) > p_r$ by one of Chebyshev's bounds for primes, so if $\mathrm{rad}(n) = p_1 \cdots p_r$, then n is y-smooth, where $y = 4 \log y^2$. Therefore,

$$\sum_{\substack{y<n\leqslant y^2 \\ \mathrm{rad}(n)=m}} \frac{1}{n} < \frac{1}{y} \sum_{\substack{n\leqslant y^2 \\ \mathrm{rad}(n)=m}} 1 \leqslant \frac{1}{y} \sum_{\substack{n\leqslant y^2 \\ \mathrm{rad}(n)=p_1\cdots p_r}} 1 \leqslant \frac{\Psi(y^2, 4\log y^2)}{y} \ll \frac{y^{O(1/\log\log 3y)}}{y}, \tag{6.12}$$

where the last bound follows, upon exponentiating, from (6.10). Combining (6.11) and (6.12) gives (6.9).

The left-hand side of (6.6) is at most

$$\sum_{\substack{m\leqslant y^{1/3} \\ \text{squarefree}}} \frac{A^{\omega(m)}(D,m)}{m} \sum_{\substack{n^2>y^{2/3} \\ \mathrm{rad}(n)=m}} \frac{1}{n^2} + \sum_{\substack{m>y^{1/3} \\ \text{squarefree}}} \frac{A^{\omega(m)}(D,m)}{m} \sum_{\substack{n\geqslant 1 \\ \mathrm{rad}(n)=m}} \frac{1}{n^2}.$$

By (6.5) and (6.9) (note that $1/n^2 < 1/(y^{1/3}n)$ when $n^2 > y^{2/3}$), we have

$$\sum_{\substack{m\leqslant y^{1/3} \\ \text{squarefree}}} \frac{A^{\omega(m)}(D,m)}{m} \sum_{\substack{n^2>y^{2/3} \\ \mathrm{rad}(n)=m}} \frac{1}{n^2} \ll_A (1+A)^{\omega(m)} \frac{y^{O(1/\log\log 3y)}}{y^{2/3}};$$

by (6.4) (note that $1/m^3 < 1/(y^{1/3}m^2)$ when $m > y^{1/3}$), and since

$$\sum_{\substack{n\geqslant 1 \\ \mathrm{rad}(n)=m}} \frac{1}{n^2} = \frac{1}{m^2} \sum_{\substack{n\geqslant 1 \\ \mathrm{rad}(n)\mid m}} \frac{1}{n^2} = \frac{1}{m^2} \prod_{p\mid m} \left(\sum_{a\geqslant 0} \frac{1}{p^{2a}} \right) \ll \frac{1}{m^2},$$

we have

$$\sum_{\substack{m>y^{1/3}\\ \text{squarefree}}} \frac{A^{\omega(m)}(D,m)}{m} \sum_{\substack{n\geqslant 1\\ \mathrm{rad}(n)=m}} \frac{1}{n^2} \ll \sum_{\substack{m>y^{1/3}\\ \text{squarefree}}} \frac{A^{\omega(m)}(D,m)}{m^3} \ll_A \frac{(1+A)^{\omega(D)}}{y^{2/3}}.$$

Combining gives (6.6).

For (6.7), we note that since $\mathrm{rad}(n)^3 \leqslant n^2 \,\mathrm{rad}(n)$ and $A^{\omega(n)} = A^{\omega(\mathrm{rad}(n))}$,

$$\sum_{n^2\,\mathrm{rad}(n)\leqslant y} A^{\omega(n)} \leqslant \sum_{\substack{a\leqslant y^{1/3}\\ \text{squarefree}}} A^{\omega(a)} \sum_{\substack{b^2\leqslant y\\ \mathrm{rad}(b)=a}} 1.$$

An argument similar to the one leading up to (6.12) shows that, uniformly for $a \leqslant y^{1/3}$, we have $\sum_{b^2\leqslant y,\,\mathrm{rad}(b)=a} 1 \ll y^{O(1/\log\log 3y)}$, and

$$\sum_{\substack{a\leqslant y^{1/3}\\ \text{squarefree}}} A^{\omega(a)} \leqslant y^{1/3} \sum_{\substack{a\leqslant y^{1/3}\\ \text{squarefree}}} \frac{A^{\omega(a)}}{a} \ll_A y^{1/3+O(1/\log\log 3y)}$$

by (6.5). Combining gives (6.7). □

To prove Proposition 1.3, we express $\mathfrak{S}_{\boldsymbol{h}}$ as a series. To this end, let us introduce some notation and establish some basic inequalities. Let a nonempty, finite set $\boldsymbol{h} \subseteq \mathbb{Z}$ be given, and let $k := \#\boldsymbol{h}$. Recall that $T_{\boldsymbol{h}}(2^\alpha)$ is defined (and nonempty when $\boldsymbol{h} = \{0\}$) for $\alpha \geqslant 2$, and for $p \equiv 3 \bmod 4$, $T_{\boldsymbol{h}}(p^\alpha)$ is defined (and nonempty when $\boldsymbol{h} = \{0\}$) for $\alpha \geqslant 1$. Let us set $T_{\boldsymbol{h}}(1) := \{1\}$ and $T_{\boldsymbol{h}}(2) := \{1, 2\}$ for completeness. For $p \not\equiv 1 \bmod 4$ and $\alpha \geqslant 1$, we may then define

$$\begin{aligned}\epsilon_{\boldsymbol{h}}(p^\alpha) := {} & \left(\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha}\right)^{-k}\left(\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha}\right)\\ & - \left(\frac{\#T_{\{0\}}(p^{\alpha-1})}{p^{\alpha-1}}\right)^{-k}\left(\frac{\#T_{\boldsymbol{h}}(p^{\alpha-1})}{p^{\alpha-1}}\right).\end{aligned} \tag{6.13}$$

Note that $\epsilon_{\boldsymbol{h}}(2^2) = 0$ by definition.

Lemma 6.2. *Let $\boldsymbol{h}$ be a nonempty, finite set of integers, and let $k := \#\boldsymbol{h}$.*

*(a) For $p \equiv 3 \bmod 4$ and **even** $\alpha \geqslant 2$, we have $\epsilon_{\boldsymbol{h}}(p^\alpha) = 0$.*

(*b*) *For* $p \not\equiv 1 \bmod 4$, *we have*

$$\epsilon_{\boldsymbol{h}}(p) \ll_k \frac{(\det(\boldsymbol{h}), p)}{p^2}. \tag{6.14}$$

(*c*) *For* $p \not\equiv 1 \bmod 4$ *and* $\alpha \geqslant 1$, *we have*

$$\epsilon_{\boldsymbol{h}}(p^\alpha) \ll_k \frac{(\det(\boldsymbol{h}), p)}{p^\alpha}. \tag{6.15}$$

(*d*) *For* $\beta \geqslant 1$, *we have*

$$\delta_{\{0\}}(2)^{-k}\delta_{\boldsymbol{h}}(2) = 1 + \sum_{\alpha=2}^{\beta} \epsilon_{\boldsymbol{h}}(2^\alpha) + O_k\bigg(\frac{1}{2^\beta}\bigg). \tag{6.16}$$

For $p \equiv 3 \bmod 4$ *and* $\beta \geqslant 1$, *we have*

$$\delta_{\{0\}}(p)^{-k}\delta_{\boldsymbol{h}}(p) = 1 + \sum_{\alpha=1}^{\beta} \epsilon_{\boldsymbol{h}}(p^{2\alpha-1}) + O_k\bigg(\frac{1}{p^{2\beta}}\bigg). \tag{6.17}$$

Proof. (a) Let $p \equiv 3 \bmod 4$ and let $\alpha \geqslant 1$. As can be seen from Proposition 5.3, (5.10) and part (c), we have

$$\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha} = \bigg(1 + \frac{1}{p}\bigg)^{-1}\bigg(1 - \frac{1}{p^{\alpha+\alpha \bmod 2}}\bigg). \tag{6.18}$$

For even α we therefore have

$$\epsilon_{\boldsymbol{h}}(p^\alpha) = \bigg(1 + \frac{1}{p}\bigg)^{k}\bigg(1 - \frac{1}{p^\alpha}\bigg)^{k}\bigg(\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} - \frac{\#T_{\boldsymbol{h}}(p^{\alpha-1})}{p^{\alpha-1}}\bigg),$$

and as we noted following (5.12) and (5.13),

$$\#T_{\boldsymbol{h}}(p^\alpha)/p^\alpha - \#T_{\boldsymbol{h}}(p^{\alpha-1})/p^{\alpha-1} = 0.$$

(b) Consider $p \equiv 3 \bmod 4$ (the case $p = 2$ is similar). Let $\alpha \geqslant 1$. Define $\eta_{\boldsymbol{h}}(p^\alpha)$ and $\kappa_{\boldsymbol{h}}(p)$ as the numbers given by the relations

$$\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} =: \delta_{\boldsymbol{h}}(p) + \eta_{\boldsymbol{h}}(p^\alpha) \quad \text{and} \quad \delta_{\boldsymbol{h}}(p) =: \bigg(1 + \frac{1}{p}\bigg)^{-1}\bigg(1 - \frac{\kappa_{\boldsymbol{h}}(p)}{p}\bigg). \tag{6.19}$$

Note that by Proposition 5.3, (5.9) and part (c), $|\eta_{\boldsymbol{h}}(p^\alpha)| < k/p^{\alpha+(\alpha \bmod 2)}$ and $\kappa_{\boldsymbol{h}}(p) \leqslant \min\{k-1, p\}$, with $\kappa_{\boldsymbol{h}}(p) = k-1$ if $p \nmid \det(\boldsymbol{h})$. Also, $\kappa_{\boldsymbol{h}}(p) \geqslant$

-1 (because $\delta_{\boldsymbol{h}}(p) \leqslant 1$). Since $\alpha + (\alpha \bmod 2) \geqslant 2$, we have

$$\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} = \left(1+\frac{1}{p}\right)^{-1}\left(1-\frac{\kappa_{\boldsymbol{h}}(p)}{p}+O\left(\frac{k}{p^2}\right)\right).$$

In the special case $\boldsymbol{h} = \{0\}$ we can take $\kappa_{\boldsymbol{h}}(p) = 0$. We therefore have

$$\begin{aligned}
&\left(\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha}\right)^{-k}\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha}\\
&= \left(1+\frac{1}{p}\right)^{k-1}\left(1-\frac{\kappa_{\boldsymbol{h}}(p)}{p}+O_k\left(\frac{1}{p^2}\right)\right)\\
&= \left(1+\frac{k-1}{p}+O_k\left(\frac{1}{p^2}\right)\right)\left(1-\frac{\kappa_{\boldsymbol{h}}(p)}{p}+O_k\left(\frac{1}{p^2}\right)\right)\\
&= 1+\frac{k-1-\kappa_{\boldsymbol{h}}(p)}{p}+O_k\left(\frac{1}{p^2}\right).
\end{aligned}$$

Writing $\xi_{\boldsymbol{h}}(p) \coloneqq k-1-\kappa_{\boldsymbol{h}}(p)$, we have

$$\left(\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha}\right)^{-k}\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} - 1 \ll_k \frac{\xi_{\boldsymbol{h}}(p)}{p}+\frac{1}{p^2}.$$

If $p \mid \det(\boldsymbol{h})$, then $\xi_{\boldsymbol{h}}(p)/p = \xi_{\boldsymbol{h}}(p)(\det(\boldsymbol{h}),p)/p^2$, and if $p \nmid \det(\boldsymbol{h})$ then, as already noted, $\kappa_{\boldsymbol{h}}(p) = k-1$, i.e. $\xi_{\boldsymbol{h}}(p) = 0$, so $\xi_{\boldsymbol{h}}(p)/p = \xi_{\boldsymbol{h}}(p)(\det(\boldsymbol{h}),p)/p^2$ in any case. Since, as already noted, $-1 \leqslant \kappa_{\boldsymbol{h}}(p) \leqslant k-1$, we have $0 \leqslant \xi_{\boldsymbol{h}}(p) \leqslant k$. Thus,

$$\left(\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha}\right)^{-k}\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} - 1 \ll_k \frac{(\det(\boldsymbol{h}),p)}{p^2}+\frac{1}{p^2} \ll \frac{(\det(\boldsymbol{h}),p)}{p^2}.$$

For $\alpha = 1$, the left-hand side is equal to $\epsilon_{\boldsymbol{h}}(p)$ (see (6.13)), so this gives (6.14).

(c) Consider $p \equiv 3 \bmod 4$ (the case $p = 2$ is similar). Let $\alpha \geqslant 1$. By (a) and (b), the result holds for $\alpha = 1$ and $\alpha \geqslant 2$ even, so we may assume that $\alpha \geqslant 3$ is odd. In that case, using (6.18) in the definition (6.13) of $\epsilon_{\boldsymbol{h}}(p^\alpha)$, we see that

$$\begin{aligned}
\epsilon_{\boldsymbol{h}}(p^\alpha) &= \left(1+\frac{1}{p}\right)^k\left\{\left(1-\frac{1}{p^{\alpha+1}}\right)^{-k}\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} - \left(1-\frac{1}{p^{\alpha-1}}\right)^{-k}\frac{\#T_{\boldsymbol{h}}(p^{\alpha-1})}{p^{\alpha-1}}\right\}\\
&= \left(1+\frac{1}{p}\right)^k\left\{\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} - \frac{\#T_{\boldsymbol{h}}(p^{\alpha-1})}{p^{\alpha-1}} + O_k\left(\frac{1}{p^{\alpha-1}}\right)\right\},
\end{aligned}$$

since, for any $\alpha \geqslant 1$, $(1-1/p^\alpha)^{-k} = 1 + O_k(1/p^\alpha)$ and $\#T_{\boldsymbol{h}}(p^\alpha)/p^\alpha = O(1)$. We deduce, from (5.12) and (5.13), that $\epsilon_{\boldsymbol{h}}(p^\alpha) \ll_k 1/p^{\alpha-1}$, which is (6.15) in the case $p \mid \det(\boldsymbol{h})$.

Now consider the case $p \nmid \det(\boldsymbol{h})$. Note that, by Proposition 5.3, (5.9) and part (c), we have, for any $\alpha \geqslant 1$,

$$\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} = \left(1+\frac{1}{p}\right)^{-1}\left(1-\frac{k-1}{p}-\frac{k}{p^{\alpha+\alpha \bmod 2}}\right).$$

In view of this and (the special case) (6.18), we have, for odd $\alpha \geqslant 3$,

$$\epsilon_{\boldsymbol{h}}(p^\alpha) = \left(1+\frac{1}{p}\right)^{k-1}\left\{\left(1-\frac{1}{p^{\alpha+1}}\right)^{-k}\left(1-\frac{k-1}{p}-\frac{k}{p^{\alpha+1}}\right) - \left(1-\frac{1}{p^{\alpha-1}}\right)^{-k}\left(1-\frac{k-1}{p}-\frac{k}{p^{\alpha-1}}\right)\right\}.$$

Since $(1-1/p^{\alpha+1})^{-k} = 1 + k/p^{\alpha+1} + O_k(1/p^{\alpha+2})$, we have

$$\left(1-\frac{1}{p^{\alpha+1}}\right)^{-k}\left(1-\frac{k-1}{p}-\frac{k}{p^{\alpha+1}}\right) = 1-\frac{k-1}{p}+O_k\left(\frac{1}{p^{\alpha+2}}\right);$$

similarly,

$$\left(1-\frac{1}{p^{\alpha-1}}\right)^{-k}\left(1-\frac{k-1}{p}-\frac{k}{p^{\alpha-1}}\right) = 1-\frac{k-1}{p}+O_k\left(\frac{1}{p^{\alpha}}\right).$$

Combining gives $\epsilon_{\boldsymbol{h}}(p^\alpha) \ll_k 1/p^\alpha$, i.e. (6.15), for odd $\alpha \geqslant 3$.

(d) Consider $p \equiv 3 \bmod 4$ (the case $p = 2$ is similar). Let $\beta \geqslant 1$. We have

$$1+\sum_{\alpha=1}^{\beta}\epsilon_{\boldsymbol{h}}(p^{2\alpha-1}) = 1+\sum_{\alpha=1}^{2\beta}\epsilon_{\boldsymbol{h}}(p^\alpha) = \left(\frac{\#T_{\{0\}}(p^{2\beta})}{p^{2\beta}}\right)^{-k}\left(\frac{\#T_{\boldsymbol{h}}(p^{2\beta})}{p^{2\beta}}\right),$$

because $\epsilon_{\boldsymbol{h}}(p^\alpha) = 0$ for α even (by (a)), and the middle sum telescopes. Now, Proposition 5.3 (c) gives $\delta_{\{0\}}(p)^{-k} = (1+1/p)^k$, and by definition of $\eta_{\boldsymbol{h}}(p^{2\beta})$ (see (6.19)), $\delta_{\boldsymbol{h}}(p) = \big(\#T_{\boldsymbol{h}}(p^{2\beta})/p^{2\beta}\big) - \eta_{\boldsymbol{h}}(p^{2\beta})$. With these substitutions, and (6.18), we verify that

$$\begin{aligned}&\delta_{\{0\}}(p)^{-k}\delta_{\boldsymbol{h}}(p) - \left(\frac{\#T_{\{0\}}(p^{2\beta})}{p^{2\beta}}\right)^{-k}\left(\frac{\#T_{\boldsymbol{h}}(p^{2\beta})}{p^{2\beta}}\right)\\ &= \frac{\#T_{\boldsymbol{h}}(p^{2\beta})}{p^{2\beta}}\left(1+\frac{1}{p}\right)^k\left(1-\left(1-\frac{1}{p^{2\beta}}\right)^{-k}-\eta_{\boldsymbol{h}}(p^{2\beta})\right).\end{aligned}$$

Now, $\#T_{\boldsymbol{h}}(p^{2\beta})/p^{2\beta} \leqslant 1$, $(1+1/p)^k \ll_k 1$, $(1-1/p^{2\beta})^{-k} = 1 + O_k(1/p^{2\beta})$, and as noted in (b), Proposition 5.3, (5.9) and part (c) show that $|\eta_{\boldsymbol{h}}(p^{2\beta})| < k/p^{2\beta}$. Combining gives (6.17). □

For $n \in \mathbb{N}$ such that $p \mid n$ implies $p \not\equiv 1 \bmod 4$, we extend (6.13) by defining

$$\epsilon_{\boldsymbol{h}}(n) := \prod_{p^\alpha \parallel n} \epsilon_{\boldsymbol{h}}(p^\alpha).$$

For such n, Lemma 6.2 (b) and (c) give

$$|\epsilon_{\boldsymbol{h}}(n)| \leqslant A_k^{\omega(n)} \frac{(\det(\boldsymbol{h}), \operatorname{rad}(n))}{n \operatorname{sf}(n)}, \tag{6.20}$$

provided A_k is sufficiently large in terms of k. Since $\epsilon_{\boldsymbol{h}}(2) = 0$ by definition, and by Lemma 6.2 (a), $\epsilon_{\boldsymbol{h}}(n) = 0$ if either $\nu_2(n) = 1$ or $\nu_p(n)$ is even (and nonzero) for some $p \equiv 3 \bmod 4$. Letting $\mathcal{N}_1 := \{n \in \mathcal{N} : p \mid n \Rightarrow p \not\equiv 1 \bmod 4\}$, where $\mathcal{N}$ is as in (6.1), we define

$$\mathcal{D} := \mathcal{N}_1 \cup \{2n : n \in \mathcal{N}_1, 2 \mid n\}. \tag{6.21}$$

Thus,

$$\mathcal{D} = \big\{2^\alpha p_1^{2\alpha_1 - 1} \cdots p_r^{2\alpha_r - 1} : \alpha \geqslant 0, \alpha \neq 1, r, \alpha_i \geqslant 1, p_i \equiv 3 \bmod 4 \ (i \leqslant r)\big\},$$

and $\epsilon_{\boldsymbol{h}}(n) = 0$ unless $n \in \mathcal{D}$. By definition (1.7) and Lemma 6.2 (d),

$$\begin{aligned} \mathfrak{S}_{\boldsymbol{h}} &= \bigg(1 + \sum_{\alpha \geqslant 2} \epsilon_{\boldsymbol{h}}(2^\alpha)\bigg) \prod_{p \not\equiv 1 \bmod 4} \bigg(1 + \sum_{\alpha \geqslant 1} \epsilon_{\boldsymbol{h}}(p^{2\alpha-1})\bigg) \\ &= 1 + \sum_{d \in \mathcal{D}} \epsilon_{\boldsymbol{h}}(d), \end{aligned} \tag{6.22}$$

the last sum being absolutely convergent in view of Lemma 6.1 and (6.20).

For the purposes of stating and proving the next lemma, we define

$$\begin{aligned} \epsilon_{\boldsymbol{h}}(p^\alpha; j) := &\left(\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha}\right)^{-j} \left(\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha}\right) \\ &- \left(\frac{\#T_{\{0\}}(p^{\alpha-1})}{p^{\alpha-1}}\right)^{-j} \left(\frac{\#T_{\boldsymbol{h}}(p^{\alpha-1})}{p^{\alpha-1}}\right), \end{aligned}$$

for $p \not\equiv 1 \bmod 4$, $\alpha \geqslant 1$, and $j \geqslant 1$; we then set $\epsilon_{\boldsymbol{h}}(n; j) := \prod_{p^\alpha \parallel n} \epsilon_{\boldsymbol{h}}(p^\alpha; j)$ for n composed of primes $p \not\equiv 1 \bmod 4$. Thus, $\epsilon_{\boldsymbol{h}}(n) = \epsilon_{\boldsymbol{h}}(n; j)$ when $j = \#\boldsymbol{h}$.

Lemma 6.3. *Set $\boldsymbol{o} := \emptyset$, or set $\boldsymbol{o} := \{0\}$. Let $n \geqslant 2$ be such that $p \mid n$ implies $p \not\equiv 1 \bmod 4$, and let $R_1, \dots, R_k$ be complete residue systems modulo n. We have*

$$\sum_{h_1 \in R_1} \cdots \sum_{h_k \in R_k} \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(n; \#\boldsymbol{o} + k) = 0,$$

where $\boldsymbol{h} = \{h_1, \dots, h_k\}$ in the summand. (Note that we may have $\#\boldsymbol{h} < k$ here.)

Proof. Let $p \not\equiv 1 \bmod 4$, $\alpha \geqslant 1$. Suppose $\boldsymbol{h} = \{h_1, \dots, h_k\}$ and $\boldsymbol{h}' = \{h'_1, \dots, h'_k\}$ satisfy $h_i \equiv h'_i \bmod p^\alpha$, and hence $h_i \equiv h'_i \bmod p^{\alpha-1}$ as well, for $i = 1, \dots, k$. For $p \equiv 3 \bmod 4$, it is clear from (5.4) that $\#T_{\boldsymbol{o}\cup\boldsymbol{h}}(p^\beta) = \#T_{\boldsymbol{o}\cup\boldsymbol{h}'}(p^\beta)$ for $\beta = \alpha$, and for $\beta = \alpha - 1$ as well. Thus, $\epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(p^\alpha; \#\boldsymbol{o} + k) = \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}'}(p^\alpha; \#\boldsymbol{o} + k)$. Similarly, we have $\epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(2^\alpha; \#\boldsymbol{o} + k) = \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}'}(2^\alpha; \#\boldsymbol{o} + k)$ (see (5.3)). Therefore, by the Chinese remainder theorem,

$$\sum_{h_1 \in R_1} \cdots \sum_{h_k \in R_k} \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(n; \#\boldsymbol{o} + k) = \prod_{p^\alpha \parallel n} \Bigg(\sum_{h_1 \in \mathbb{Z}_{p^\alpha}} \cdots \sum_{h_k \in \mathbb{Z}_{p^\alpha}} \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(p^\alpha; \#\boldsymbol{o} + k) \Bigg),$$

where $\boldsymbol{h} = \{h_1, \dots, h_k\}$ in both summands, and $\mathbb{Z}_{p^\alpha} := \{0, \dots, p^\alpha - 1\}$. It therefore suffices to show that

$$\sum_{h_1 \in \mathbb{Z}_{p^\alpha}} \cdots \sum_{h_k \in \mathbb{Z}_{p^\alpha}} \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(p^\alpha; \#\boldsymbol{o} + k) = 0 \tag{6.23}$$

for all $p \not\equiv 1 \bmod 4$ and $\alpha \geqslant 1$.

Consider the case $\boldsymbol{o} = \emptyset$. For $p \equiv 3 \bmod 4$ and $\alpha \geqslant 1$, we have

$$\sum_{h_1 \in \mathbb{Z}_{p^\alpha}} \cdots \sum_{h_k \in \mathbb{Z}_{p^\alpha}} \#T_{\boldsymbol{h}}(p^\alpha) = \sum_{a \in \mathbb{Z}_{p^\alpha}} \sum_{\substack{h_1 \in \mathbb{Z}_{p^\alpha} \\ a+h_1 \in S_p \\ \nu_p(a+h_1) < \alpha}} \cdots \sum_{\substack{h_k \in \mathbb{Z}_{p^\alpha} \\ a+h_k \in S_p \\ \nu_p(a+h_k) < \alpha}} 1,$$

as can be seen by applying the definition (5.4) of $T_{\boldsymbol{h}}(p^\alpha)$ and changing the order of summation. For $i = 1, \dots, k$, each sum over h_i on the right-hand side enumerates a translation of $T_{\{0\}}(p^\alpha)$, so the entire sum (i.e. the left-hand side) is equal to $p^\alpha (\#T_{\{0\}}(p^\alpha))^k$. Whence

$$\sum_{h_1 \in \mathbb{Z}_{p^\alpha}} \cdots \sum_{h_k \in \mathbb{Z}_{p^\alpha}} \left(\frac{\#T_{\{0\}}(p^\alpha)}{p^\alpha} \right)^{-k} \left(\frac{\#T_{\boldsymbol{h}}(p^\alpha)}{p^\alpha} \right) = p^{k\alpha}.$$

Since

$$\sum_{h_1\in\mathbb{Z}_{p^\alpha}}\cdots\sum_{h_k\in\mathbb{Z}_{p^\alpha}}\#T_{\boldsymbol{h}}(p^{\alpha-1})=p^k\sum_{h_1\in\mathbb{Z}_{p^{\alpha-1}}}\cdots\sum_{h_k\in\mathbb{Z}_{p^{\alpha-1}}}\#T_{\boldsymbol{h}}(p^{\alpha-1}),$$

we similarly have

$$\sum_{h_1\in\mathbb{Z}_{p^\alpha}}\cdots\sum_{h_k\in\mathbb{Z}_{p^\alpha}}\left(\frac{\#T_{\{0\}}(p^{\alpha-1})}{p^{\alpha-1}}\right)^{-k}\left(\frac{\#T_{\boldsymbol{h}}(p^{\alpha-1})}{p^{\alpha-1}}\right)=p^kp^{k(\alpha-1)}=p^{k\alpha}.$$

Subtracting gives (6.23) for $\alpha\geqslant 1$. In a similar fashion, we obtain (6.23) in the case $\boldsymbol{o}=\{0\}$. An analogous argument gives the same results for $p=2$. □

In the proof of Proposition 1.3, we also make use of basic lattice point counting arguments, as in the final two lemmas below.

Lemma 6.4. *Let $\mathcal{D}$ be as in* (6.21). *Set $\boldsymbol{o}:=\emptyset$, or set $\boldsymbol{o}:=\{0\}$. Fix an integer $k\geqslant 1$, and a number $M_k\geqslant 1$ that depends on k only. Also, fix $B\geqslant 1$. For $y\geqslant 1$, we have*

(6.24)
$$\sum_{\substack{d\in\mathcal{D}\\ d>y}}\frac{M_k^{\omega(d)}}{d\,\mathrm{sf}(d)}\sum_{0<h_1<\cdots<h_k\leqslant By}(\det(\boldsymbol{o}\cup\boldsymbol{h}),\mathrm{rad}(d))\ll_{k,B}y^{k-2/3+O(1/\log\log 3y)},$$

where $\boldsymbol{h}=\{h_1,\dots,h_k\}$ in the summand.

Proof. Let $y\geqslant 1$. Let us first show that, for any squarefree integer $c\geqslant 1$,

$$\sum_{\substack{0<h_1<\cdots<h_k\leqslant By\\ c\mid\det(\{0,h_1,\dots,h_k\})}}1\leqslant k^{2\omega(c)}\left(\frac{(By)^k}{c}+O_k\big((By)^{k-1}\big)\right).\tag{6.25}$$

Let $h_0=0,h_1,\dots,h_k$ be pairwise distinct integers, and suppose that c divides $\prod_{0\leqslant i<j\leqslant k}(h_i-h_j)$. Then, since c is squarefree, there exist pairwise coprime positive integers c_{ij} such that $c=\prod_{0\leqslant i<j\leqslant k}c_{ij}$ and $c_{ij}\mid h_i-h_j$, $0\leqslant i<j\leqslant k$. Therefore,

$$\sum_{\substack{0<h_1<\cdots<h_k\leqslant By\\ c\mid\det(\{h_0,h_1,\dots,h_k\})}}1\leqslant\sum_{c=c_{01}\cdots c_{(k-1)k}}\sum_{\substack{h_1\in I_{By}\\ 0\leqslant i<j\leqslant k-1\Rightarrow c_{ij}\mid h_i-h_j}}\sum_{h_2\in I_{By}}\cdots\sum_{h_{k-1}\in I_{By}}\sum_{\substack{h_k\in I_{By}\\ 0\leqslant i\leqslant k-1\Rightarrow c_{ik}\mid h_i-h_k}}1,$$

where on the right-hand side, the outermost sum is over all decompositions of c as a product of $\binom{k+1}{2}$ positive integers, and $I_{By} := (0, By]$.

Consider the decomposition $c = c_{01} \cdots c_{(k-1)k}$. Let us define $c_j := \prod_{i=0}^{j-1} c_{ij}$ for $j = 1, \dots, k$. Notice that $c = \prod_{j=1}^{k} c_j$. By the Chinese remainder theorem, the condition on h_k in the innermost sum above is equivalent to h_k being in some congruence class modulo c_k, uniquely determined by $h_0, h_1, \dots, h_{k-1}$. The sum is therefore equal to $By/c_k + O(1)$. Iterating this argument k times, we see that the inner sum over $h_1, \dots, h_k$ is equal to

$$\prod_{j=1}^{k} \left(\frac{By}{c_j} + O(1) \right) = \frac{(By)^k}{c} + O_k((By)^{k-1}).$$

The bound (6.25) follows by combining and noting that, since c is squarefree, the number of ways of writing c as a product of $\binom{k+1}{2}$ positive integers is $\binom{k+1}{2}^{\omega(c)}$, and that $\binom{k+1}{2} \leqslant k^2$.

For $\boldsymbol{h} = \{h_1, \dots, h_k\}$, with $h_1, \dots, h_k$ pairwise distinct, nonzero integers, and any $d \in \mathbb{N}$, we trivially have

$$(\det(\boldsymbol{o} \cup \boldsymbol{h}), \operatorname{rad}(d)) \leqslant \sum_{c \mid \det(\{0, h_1, \dots, h_k\}), \operatorname{rad}(d)} c.$$

If $h_1, \dots, h_k \leqslant By$ as well, then $p \mid c$ implies $p \leqslant By$. From this and (6.25), it follows that

$$\begin{aligned} \sum_{0 < h_1 < \cdots < h_k \leqslant By} (\det(\boldsymbol{o} \cup \boldsymbol{h}, \operatorname{rad}(d)) \ll_{k,B} y^k \sum_{c \mid \operatorname{rad}(d)} k^{2\omega(c)} \\ + y^{k-1} \sum_{\substack{c \mid \operatorname{rad}(d) \\ p \mid c \Rightarrow p \leqslant By}} c k^{2\omega(c)}, \end{aligned}$$

where $\boldsymbol{h} = \{h_1, \dots, h_k\}$ in the summand on the left. Now, for $c \mid \operatorname{rad}(d)$ we have $k^{2\omega(c)} \leqslant k^{2\omega(d)}$, and $\sum_{c \mid \operatorname{rad}(d)} 1 = 2^{\omega(d)}$. Applying these bounds to the left-hand side of (6.24), we see that it is

$$\ll_{k,B} y^k \sum_{\substack{d \in \mathcal{D} \\ d > y}} \frac{A_k^{\omega(d)}}{d \operatorname{sf}(d)} + y^{k-1} \sum_{d \in \mathcal{D}} \frac{A_k^{\omega(d)}}{d \operatorname{sf}(d)} \sum_{\substack{c \mid \operatorname{rad}(d) \\ p \mid c \Rightarrow p \leqslant By}} c, \tag{6.26}$$

where A_k, here and below, denotes a sufficiently large number depending on k, which may be a different number at each occurrence.

By definition (6.21) of $\mathcal{D}$, for every $d \in \mathcal{D}$, we have $d = n$ or $d = 2n$ for some $n \in \mathcal{N}$, where $\mathcal{N}$ is as in (6.1). Therefore, as a direct consequence of Lemma 6.1, we have

$$\sum_{\substack{d\in\mathcal{D}\\ d>y}} \frac{A_k^{\omega(d)}}{d\,\mathrm{sf}(d)} \ll_k \frac{y^{O(1/\log\log 3y)}}{y^{2/3}}. \tag{6.27}$$

More specifically, for every $d \in \mathcal{D}$, we have $d = ab^2\,\mathrm{rad}(b)$ or $d = 2ab^2\,\mathrm{rad}(b)$ for some uniquely determined $a, b \in \mathbb{N}$, where a is squarefree and $(a, b) = 1$. Furthermore, d is not exactly divisible by 2, and so we have $2 \nmid a$ in the case $d = ab^2\,\mathrm{rad}(b)$, while $2 \mid ab$ in the case $d = 2ab^2\,\mathrm{rad}(b)$. In either case, we have the following: $A_k^{\omega(d)} = A_k^{\omega(a)} A_k^{\omega(b)}$; $d\,\mathrm{sf}(d) = a^2b^2\,\mathrm{rad}(b)$ or $d\,\mathrm{sf}(d) = 2a^2b^2\,\mathrm{rad}(b)$; and $\mathrm{rad}(d) = a\,\mathrm{rad}(b)$. Thus, if $c \mid \mathrm{rad}(d)$, then $c = c_1c_2$, where $c_1 \mid a$ and $c_2 \mid \mathrm{rad}(b)$. Consequently,

$$\sum_{d\in\mathcal{D}} \frac{A_k^{\omega(d)}}{d\,\mathrm{sf}(d)} \sum_{\substack{c\mid \mathrm{rad}(d)\\ p\mid c\Rightarrow p\leqslant By}} c \ll \sum_{\substack{a\geqslant 1\\ \text{squarefree}}} \frac{A_k^{\omega(a)}}{a^2} \sum_{b\geqslant 1} \frac{A_k^{\omega(b)}}{b^2\,\mathrm{rad}(b)} \sum_{\substack{c_1\mid a\\ p\mid c_1\Rightarrow p\leqslant By}} c_1 \sum_{\substack{c_2\mid \mathrm{rad}(b)\\ p\mid c_2\Rightarrow p\leqslant By}} c_2.$$

Now,

$$\begin{aligned}\sum_{\substack{a\geqslant 1\\ \text{squarefree}}} \frac{A_k^{\omega(a)}}{a^2} \sum_{\substack{c_1\mid a\\ p\mid c_1\Rightarrow p\leqslant By}} c_1 &\leqslant \sum_{\substack{c_1\geqslant 1\\ \text{squarefree}\\ p\mid c_1\Rightarrow p\leqslant By}} \frac{A_k^{\omega(c_1)}}{c_1} \sum_{\substack{a_1\geqslant 1\\ \text{squarefree}}} \frac{A_k^{\omega(a_1)}}{a_1^2}\\ &\ll_k \sum_{\substack{c_1\geqslant 1\\ \text{squarefree}\\ p\mid c_1\Rightarrow p\leqslant By}} \frac{A_k^{\omega(c_1)}}{c_1};\end{aligned}$$

as can be seen by writing $a = a_1c_1$ and changing order of summation; also

$$\sum_{\substack{c_1\geqslant 1\\ \text{squarefree}\\ p\mid c_1\Rightarrow p\leqslant By}} \frac{A_k^{\omega(c_1)}}{c_1} \leqslant \prod_{p\leqslant By} \left(1 + \frac{A_k}{p}\right) \ll_{k,B} (\log 3y)^{A_k}.$$

(See (6.8).) Next, note that since $\sum_{c_2|\mathrm{rad}(b)} c_2 \leqslant \mathrm{rad}(b) \sum_{c_2|\mathrm{rad}(b)} 1 \leqslant 2^{\omega(b)} \mathrm{rad}(b)$,

$$\sum_{b\geqslant 1} \frac{A_k^{\omega(b)}}{b^2\,\mathrm{rad}(b)} \sum_{\substack{c_2|\mathrm{rad}(b)\\ p|c_2\Rightarrow p\leqslant By}} c_2 \leqslant \sum_{b\geqslant 1} \frac{A_k^{\omega(b)}}{b^2} \sum_{c_2|\mathrm{rad}(b)} 1 \leqslant \sum_{b\geqslant 1} \frac{A_k^{\omega(b)}}{b^2} \ll_k 1.$$

Combining all of this gives

$$\sum_{d\in\mathcal{D}} \frac{A_k^{\omega(d)}}{d\,\mathrm{sf}(d)} \sum_{\substack{c|\mathrm{rad}(d)\\ p|c\Rightarrow p\leqslant By}} c \ll_{k,B} (\log 3y)^{A_k}. \tag{6.28}$$

Finally, we obtain (6.24) by combining (6.26) with (6.27) and (6.28). □

Lemma 6.5. *Fix an integer $k \geqslant 1$, and a bounded convex set $\mathscr{C} \subset \mathbb{R}^k$. For $y \geqslant 1$ we have $\#(y\mathscr{C} \cap \mathbb{Z}^k) = y^k \mathrm{vol}(\mathscr{C}) + O_{k,\mathscr{C}}(y^{k-1})$.*

Proof. This is a special case of [24, pp. 128–129]. □

Proof of Proposition 1.3. Fix an integer $k \geqslant 1$, and a bounded convex set $\mathscr{C} \subset \Delta^k$, where $\Delta^k := \{(x_1, \dots, x_k) \in \mathbb{R}^k : 0 < x_1 < \cdots < x_k\}$ (see (1.10)). Set $\boldsymbol{o} := \emptyset$, or set $\boldsymbol{o} := \{0\}$. Let $y \geqslant 1$. To ease notation throughout, let $\mathcal{H} := y\mathscr{C} \cap \mathbb{Z}^k$, $\vec{\boldsymbol{h}} = (h_1, \dots, h_k)$, and $\boldsymbol{h} = \{h_1, \dots, h_k\}$. Note that $0 < h_1 < \cdots < h_k \ll_{\mathscr{C}} y$ for $\vec{\boldsymbol{h}} \in \mathcal{H}$. Also, let A_k stand for a sufficiently large number depending on k, which may be a different number at each occurrence.

In view of (6.22) we see, upon partitioning the sum over d and changing order of summation, that

$$\sum_{\vec{\boldsymbol{h}}\in\mathcal{H}} \mathfrak{S}_{\boldsymbol{o}\cup\boldsymbol{h}} = \sum_{\vec{\boldsymbol{h}}\in\mathcal{H}} 1 + \sum_{\substack{d\in\mathcal{D}\\ d\leqslant y}} \sum_{\vec{\boldsymbol{h}}\in\mathcal{H}} \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(d) + \sum_{\substack{d\in\mathcal{D}\\ d> y}} \sum_{\vec{\boldsymbol{h}}\in\mathcal{H}} \epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(d), \tag{6.29}$$

with $\mathcal{D}$ as defined in (6.21). By Lemma 6.5, we have

$$\sum_{\vec{\boldsymbol{h}}\in\mathcal{H}} 1 = y^k \mathrm{vol}(\mathscr{C}) + O_{k,\mathscr{C}}(y^{k-1}). \tag{6.30}$$

By (6.20) and Lemma 6.4, we have

$$\sum_{\substack{d\in\mathcal{D}\\ d>y}}\sum_{\vec{\boldsymbol{h}}\in\mathcal{H}}|\epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(d)| \leqslant \sum_{\substack{d\in\mathcal{D}\\ d>y}}\sum_{\vec{\boldsymbol{h}}\in\mathcal{H}} A_k^{\omega(d)}\frac{(\det(\boldsymbol{o}\cup\boldsymbol{h}),\operatorname{rad}(d))}{d\operatorname{sf}(d)} \ll_{k,\mathscr{C}} y^{k-1}\frac{y^{O(1/\log\log 3y)}}{y^{2/3}}. \tag{6.31}$$

Consider the middle sum on the right-hand side of (6.29). Let d be any element of $\mathcal{D}$ with $d\leqslant y$, and partition $\mathbb{R}^k$ into cubes

$$C_{d,\vec{\boldsymbol{t}}} := \{(x_1,\dots,x_k)\in\mathbb{R}^k : t_i d\leqslant x_i < (t_i+1)d, i=1,\dots,k\},$$

with $\vec{\boldsymbol{t}} := (t_1,\dots,t_k)$ running over $\mathbb{Z}^k$. Each $\vec{\boldsymbol{h}}\in\mathcal{H}$ is a point in a unique cube of this form: we call $\vec{\boldsymbol{h}}$ a *d-interior* point if this cube is entirely contained in $y\mathscr{C}$, and $\vec{\boldsymbol{h}}$ a *d-boundary* point if this cube has a nonempty intersection with the boundary of $y\mathscr{C}$. We partition $\mathcal{H}$ into d-interior points and d-boundary points. As $\vec{\boldsymbol{h}}$ runs over all d-interior points of $\mathcal{H}$, h_i $(i=1,\dots,k)$ runs over a pairwise disjoint union of complete residue systems modulo d, none of which contain 0. By Lemma 6.3 (we have $\#(\boldsymbol{o}\cup\boldsymbol{h})=\#\boldsymbol{o}+k$ for each $\vec{\boldsymbol{h}}\in\mathcal{H}$), it follows that

$$\sum_{\substack{d\in\mathcal{D}\\ d\leqslant y}}\sum_{\vec{\boldsymbol{h}}\in\mathcal{H}}\epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(d) = \sum_{\substack{d\in\mathcal{D}\\ d\leqslant y}}\sum_{\substack{\vec{\boldsymbol{h}}\in\mathcal{H}\\ d\text{-boundary}}}\epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(d). \tag{6.32}$$

By (6.20), and the aforementioned trivial bound for $(\det(\boldsymbol{o}\cup\boldsymbol{h}),\operatorname{rad}(d))$,

$$\begin{aligned}\sum_{\substack{d\in\mathcal{D}\\ d\leqslant y}}\sum_{\substack{\vec{\boldsymbol{h}}\in\mathcal{H}\\ d\text{-boundary}}}|\epsilon_{\boldsymbol{o}\cup\boldsymbol{h}}(d)| &\leqslant \sum_{\substack{d\in\mathcal{D}\\ d\leqslant y}}\frac{A_k^{\omega(d)}}{d\operatorname{sf}(d)}\sum_{\substack{\vec{\boldsymbol{h}}\in\mathcal{H}\\ d\text{-boundary}}}(\det(\boldsymbol{o}\cup\boldsymbol{h}),\operatorname{rad}(d))\\ &\leqslant \sum_{\substack{d\in\mathcal{D}\\ d\leqslant y}}\frac{A_k^{\omega(d)}}{d\operatorname{sf}(d)}\sum_{c\mid\operatorname{rad}(d)} c\sum_{\substack{\vec{\boldsymbol{h}}\in\mathcal{H}\\ d\text{-boundary}\\ c\mid\det(\{0,h_1,\dots,h_k\})}} 1.\end{aligned}$$

For each $d\in\mathcal{D}$ with $y/d\geqslant 1$, the proof of Lemma 6.5 (see [24, pp. 128–129]) shows that there are $\ll_{k,\mathscr{C}}(y/d)^{k-1}$ cubes $C_{d,\vec{\boldsymbol{t}}}$ that have a nonempty intersection with the boundary of $y\mathscr{C}$. For each such boundary cube $C_{d,\vec{\boldsymbol{t}}}$, the corresponding d-boundary points are all in $C_{d,\vec{\boldsymbol{t}}}\cap\mathbb{Z}^k$, which is a product of complete residue systems modulo d, and, given that $c\mid\operatorname{rad}(d)$ (and hence $c\mid d$),

the condition $c \mid \det(\{0, h_1, \dots, h_k\})$ is equivalent to $c \mid \det(\{0, h_1', \dots, h_k'\})$ when $h_i \equiv h_i' \bmod d$, $i = 1, \dots, k$.

If follows that, for $d \in \mathcal{D}$ with $d \leqslant y$, and for $c \mid \operatorname{rad}(d)$, we have

$$\sum_{\substack{\vec{\boldsymbol{h}} \in \mathcal{H} \\ d\text{-boundary} \\ c \mid \det(\{0, h_1, \dots, h_k\})}} 1 \ll_{k,\mathscr{C}} \frac{y^{k-1}}{d^{k-1}} \sum_{\substack{0 < h_1 < \dots < h_k \leqslant d \\ c \mid \det(\{0, h_1, \dots, h_k\})}} 1 \ll_k y^{k-1} d \left(\frac{A_k^{\omega(c)}}{c} \right)$$

by (6.25). Whence

$$\sum_{\substack{d \in \mathcal{D} \\ d \leqslant y}} \sum_{\substack{\vec{\boldsymbol{h}} \in \mathcal{H} \\ d\text{-boundary}}} |\epsilon_{\boldsymbol{o} \cup \boldsymbol{h}}(d)| \ll_{k,\mathscr{C}} y^{k-1} \sum_{\substack{d \in \mathcal{D} \\ d \leqslant y}} \frac{A_k^{\omega(d)}}{\operatorname{sf}(d)} \sum_{c \mid \operatorname{rad}(d)} A_k^{\omega(c)} \leqslant y^{k-1} \sum_{\substack{d \in \mathcal{D} \\ d \leqslant y}} \frac{A_k^{\omega(d)}}{\operatorname{sf}(d)},$$

since $\sum_{c \mid \operatorname{rad}(d)} A_k^{\omega(c)}$ is at most $A_k^{\omega(d)} \sum_{c \mid \operatorname{rad}(d)} 1 = (2A_k)^{\omega(d)}$. By (6.3), this last sum is $\ll_k y^{1/3 + O(1/\log\log 3y)}$. Combining, we obtain

$$\sum_{\substack{d \in \mathcal{D} \\ d \leqslant y}} \sum_{\vec{\boldsymbol{h}} \in \mathcal{H}} \epsilon_{\boldsymbol{o} \cup \boldsymbol{h}}(d) \ll_{k,\mathscr{C}} y^{k-1} y^{1/3 + O(1/\log\log 3y)}. \tag{6.33}$$

Combining (6.29) with (6.30), (6.31), and (6.33) gives (1.14). □

Remark 6.6. A simpler argument, though giving a much weaker error term, is to take $P = \prod_{p<y} p^{e_p}$ and $y = \frac{1}{3} \log t$ in Ford's argument [7] and choosing $(e_p)_{p<y}$ appropriately; together with the results in Section 5 we can then obtain a lower bound of the form

$$\sum_{(h_1, \dots, h_k) \in t\mathscr{C} \cap \mathbb{Z}^k} \mathfrak{S}_{\boldsymbol{o} \cup \boldsymbol{h}} \geqslant t^k (1 + o(1)), \quad t \to \infty.$$

Further, by removing the condition $\nu_p(a + h) < \alpha$ in the definition of $T_{\boldsymbol{h}}$, we obtain upper bounds for the p-adic densities, and a similar adaption of Ford's argument then gives a matching upper bound (up to lower order errors). □

References

[1] L. Bary-Soroker and A. Fehm, *Correlations of sums of two squares and other arithmetic functions in function fields*, preprint, `arXiv:1701.04092`, 2017.

[2] M. V. Berry and M. Tabor, *Level clustering in the regular spectrum*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **356** (1977), no. 1686, 375–394.

[3] P. Billingsley, *Probability and measure*, 3rd ed., John Wiley & Sons, New York, 1995.

[4] T. Cochrane and R. E. Dressler, *Consecutive triples of sums of two squares*, Arch. Math. (Basel) **49** (1987), no. 4, 301–304.

[5] R. D. Connors and J. P. Keating, *Two-point spectral correlations for the square billiard*, J. Phys. A **30** (1997), no. 6, 1817–1830.

[6] C. David, D. Koukoulopoulos, and E. Smith, *Sums of Euler products and statistics of elliptic curves*, Math. Ann. (2016), 1–68.

[7] K. Ford, *Simple proof of Gallagher's singular series sum estimate*, `arXiv:1108.3861`.

[8] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), no. 1, 4–9.

[9] A. Granville, *Smooth numbers: computational number theory and beyond*, in: Algorithmic number theory: lattices, number fields, curves and cryptography (Eds. J. P. Buhler and P. Stevenhagen), pp. 267–323, Math. Sci. Res. Inst. Publ. Vol. 44. Cambridge University Press, Cambridge, 2008.

[10] A. Granville and P. Kurlberg, *Poisson statistics via the Chinese remainder theorem*, Adv. Math. **218** (2008), no. 6, 2013–2042.

[11] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Clarendon Press, Oxford, 1938.

[12] C. Hooley, *On the difference between consecutive numbers prime to n. III*, Math. Z. **90** (1965), no. 5, 355–364.

[13] C. Hooley, *On the intervals between consecutive terms of sequences*, in: Proceedings of the Symposium in Pure Mathematics of the American Mathematical Society (Ed. H. G. Diamond), held at St. Louis University, pp. 129–140, St. Louis, MO, March 27–30, 1972. Proceedings of Symposia in Pure Mathematics, Vol. XXIV, Amer. Math. Soc., Providence, RI, 1973.

[14] C. Hooley, *On the intervals between numbers that are sums of two squares: II*, J. Number Theory. **5** (1973), no. 3, 215–217.

[15] C. Hooley, *On the intervals between numbers that are sums of two squares. III*, J. Reine Angew. Math. **267** (1974), 207–218.

[16] K.-H. Indlekofer, *Scharfe untere abschätzung für die anzahlfunktion der B-zwillinge*, Acta Arith. **26** (1974), no. 2, 207–212.

[17] H. Iwaniec, *The half-dimensional sieve*, Acta Arith. **29** (1976), no. 1, 69–95.

[18] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications, Vol. 45, American Mathematical Society, Providence, RI, 1999.

[19] E. Kowalski, *Averages of Euler products, distribution of singular series and the ubiquity of Poisson distribution*, Acta Arith. **148** (2011), no. 2, 153–187.

[20] P. Kurlberg and Z. Rudnick. *The distribution of spacings between quadratic residues*, Duke Math. J. **100** (1999), no. 2, 211–242.

[21] P. Kurlberg, *The distribution of spacings between quadratic residues. II*, Israel J. Math. **120(A)** (2000), 205–224.

[22] P. Kurlberg, *Poisson spacing statistics for value sets of polynomials*, Int. J. Number Theory **5** (2009), no. 3, 489–513.

[23] E. Landau, *Über die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindestzahl der zu ihrer additiven Zusammensetzung erforderlichen Quadrate*, Arch. der Math. u. Phys. (3) **13** (1908), 305–312.

[24] S. Lang, *Algebraic number theory*, 2nd ed., Graduate Texts in Mathematics, Vol. 110. Springer–Verlag, New York, 1994.

[25] W. G. Nowak, *On the distribution of M-tuples of B-numbers*, Publ. Inst. Math. (Beograd) (N. S.) **77** (2005), no. 91, 71–78.

[26] J. Pintz, *On the singular series in the prime k-tuple conjecture*, `arXiv:1004.1084`.

[27] G. J. Rieger, *Aufeinanderfolgende zahlen als summen von zwei quadraten*, Indag. Math. (Proceedings) **68** (1965), 208–220.

[28] Z. Rudnick and H. Ueberschär, *On the eigenvalue spacing distribution for a point scatterer on the flat torus*, Ann. Henri Poincaré **15** (2014), no. 1, 1–27.

[29] A. Selberg, *Remarks on multiplicative functions*, in: Number theory day (Proc. Conf., Rockefeller Univ., New York, 1976), pp. 232–241, Lecture Notes In Mathematics, Vol. 626, Springer, Berlin, 1977.

[30] Y. Smilansky, *Sums of two squares — pair correlation and distribution in short intervals*, Int. J. Number Theory **9** (2013), no. 7, 1687–1711.

Department of Pure Mathematics, University of Waterloo
Waterloo ON, Canada
E-mail address: tfreiberg@uwaterloo.ca

Department of Mathematics, KTH Royal Institute of Technology
Stockholm, Sweden
E-mail address: kurlberg@kth.se

Department of Mathematics, ORT Braude College
Karmiel, Israel
Current address:
Unit of Mathematics
AFEKA Tel-Aviv Academic College of Engineering
Tel Aviv, Israel
E-mail address: liorr@afeka.ac.il

Received January 30, 2017
Accepted April 5, 2017