

LEMKE OLIVER AND SOUNDARARAJAN BIAS FOR CONSECUTIVE SUMS OF TWO SQUARES

CHANTAL DAVID, LUCILE DEVIN, JUNGBAE NAM AND JEREMY SCHLITT

ABSTRACT. In a surprising recent work, Lemke Oliver and Soundararajan noticed how experimental data exhibits erratic distributions for consecutive pairs of primes in arithmetic progressions, and proposed a heuristic model based on the Hardy–Littlewood conjectures containing a large secondary term, which fits the data very well. In this paper, we study consecutive pairs of sums of squares in arithmetic progressions, and develop a similar heuristic model based on the Hardy–Littlewood conjecture for sums of squares, which also explain the biases in the experimental data. In the process, we prove several results related to averages of the Hardy–Littlewood constant in the context of sums of two squares.

1. INTRODUCTION

We study in this paper the distribution of consecutive sums of two squares in arithmetic progressions. Our work is inspired by a recent paper of Lemke Oliver and Soundararajan [LOS16] who proposed a heuristic model based on the Hardy–Littlewood conjecture for the distribution of consecutive primes in arithmetic progressions.

Roughly speaking, it is expected that numbers described by reasonable multiplicative constraints should be well-distributed, in short intervals and in arithmetic progressions. The case of prime numbers is of course well-studied, and this philosophy was also tested for numbers expressible as sums of two squares, as well as square-free numbers¹. Gallagher [Gal76] proved that the distribution of primes of size up to x in intervals of size $\log x$ has a Poisson spacing distribution, assuming some explicit form of the Hardy–Littlewood conjecture. This was generalized to primes in arithmetic progressions by Granville [Gra89] and to sums of two squares by Freiberg, Kurlberg and Rosenzweig [FKR17], for intervals of size $\sqrt{\log x}/K$, which is the correct analogue to Gallagher’s result in view of (1). For primes in larger intervals, Montgomery and Soundararajan [MS04] showed that the spacings exhibit a normal distribution around the mean, assuming again some explicit form of the Hardy–Littlewood conjecture. We prove in this paper a weaker version of their results (Theorem 3.4) for the case of sums of two squares which is needed to study the distribution of successive sums of two squares in arithmetic progressions. We speculate that the full analogue of their results can be obtained for sums of two squares, but we did not pursue it as Theorem 3.4 is sufficient for our application. Some unexpected irregularities in the distribution of primes in short intervals were discovered by Maier [Mai85], and it was shown by Balog and Wooley [BW00] that sums of two squares exhibit the same irregularities. Sums of two squares in short intervals were also studied over function fields of a finite field \mathbb{F}_q , where many results which are inaccessible over number fields can be proven when the size of the finite field \mathbb{F}_q grows [BSW16, BBF18, BF19, GR20].

We first fix some notations. We denote by

$$\mathbf{E} = \{a^2 + b^2 : a, b \in \mathbb{Z}\} = \{E_n : n \in \mathbb{N}\}$$

the set of sums of two squares (enumerated in increasing order), such that E_n is the n th number that can be written as a sum of two squares. Let $\mathbf{1}_{\mathbf{E}}$ the indicator function of this set. By a classical result of Landau,

$$(1) \quad \sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) \sim K \frac{x}{\sqrt{\log x}},$$

2010 *Mathematics Subject Classification.* 11N25, 11N69, 11P55.

Key words and phrases. Sums of two squares, Hardy–Littlewood conjecture, Selberg–Delange method.

¹In the case of square-free numbers, the Hardy–Littlewood conjecture is a theorem [Mir49], and the analogue of [LOS16] has been proved recently by Mennema [Men17].

where K is the constant defined by (5). The distribution of sum of two squares in arithmetic progressions exhibit different behavior depending on the modulus q of the progression, and we restrict in this paper to the case where q is a prime number such that $q \equiv 1 \pmod{4}$. In that case, the sums of squares are equidistributed in all the residue classes $a \pmod{q}$, including the class $a \equiv 0 \pmod{q}$ (see Theorem 2.2), but unlike the case of the primes, there is a large secondary term depending if the residue class $a \equiv 0 \pmod{q}$ or not (see Theorem 2.4).

We consider in this paper the following question, which was studied by Lemke Oliver and Soundararajan for primes [LOS16]. Fix a prime number $q \equiv 1 \pmod{4}$, and integers a, b . What is the distribution of

$$N(x; q, (a, b)) := \#\{E_n \leq x : E_n \equiv a \pmod{q}, E_{n+1} \equiv b \pmod{q}\} ?$$

Using a model based on randomness, we expect successive sums of two squares to be well-distributed in arithmetic progressions, and each of the q^2 pairs of classes (a, b) to contain the same proportion (asymptotically) of sums of two squares, with possibly a bias towards the pairs (a, b) where $ab \equiv 0 \pmod{q}$ in view of Theorem 2.4. However, the numerical data of Table 1 (for $q = 5$ and $x = 10^{12}$) shows a lot of fluctuation, and in particular an unexpected large bias against the classes (a, a) including $(0, 0)$. Interestingly, this bias goes in the opposite direction of the bias for sums of squares in arithmetic progressions: there are “more” sums of two squares congruent to $0 \pmod{q}$, but there are “less” consecutive sums of two squares congruent to $(0, 0) \pmod{q}$.

a	b	$N(10^{12}; 5, (a, b))$	a	b	$N(10^{12}; 5, (a, b))$	a	b	$N(10^{12}; 5, (a, b))$
0	0	4 108 407 474	2	0	8 049 996 586	4	0	7 155 732 959
	1	7 153 121 164		1	5 516 037 772		1	5 356 545 210
	2	5 604 312 560		2	3 754 593 831		2	7 730 855 281
	3	8 054 714 831		3	6 837 553 372		3	5 497 266 920
	4	5 780 373 060		4	5 350 735 550		4	3 768 530 444
1	0	5 777 315 850	3	0	5 609 476 219			
	1	3 765 205 659		1	7 718 021 263			
	2	6 870 009 299		2	5 549 146 140			
	3	5 354 226 097		3	3 765 159 558			
	4	7 742 174 162		4	6 867 117 598			

TABLE 1. $N(x; q, (a, b))$ for $q = 5$ and $x = 10^{12}$. The average of $N(x; q, (a, b))$ is 5 949 465 154.

Estimates for the consecutive sums of squares (or consecutive primes) in arithmetic progressions is a very difficult question, and few results are known. For consecutive primes in arithmetic progressions, it was conjectured by Chowla that there are infinitively many primes p_n such that $p_{n+i-1} \equiv a \pmod{q}$ for $1 \leq i \leq r$, for any $(a, q) = 1$ and $r \geq 2$. This was proven by Shiu [Shi00]. Recent progress in sieve theory have led to a new proof of Shiu’s result [BFT15], and Maynard has proven that the number of such primes is $\gg \pi(x)$ [May16]. It would be interesting to see if those recent progresses could be applied to get lower bounds for the number of successive sums of two squares E_n such that $E_{n+i-1} \equiv a \pmod{q}$ for $1 \leq i \leq r$, for any a and $r \geq 2$, but this question was not addressed yet in the literature.

We propose in this paper a heuristic model predicting an asymptotic for $N(x; q, (a, b))$, based on the heuristic of Lemke Oliver and Soundararajan [LOS16] for the case of primes, and exhibiting a similar bias.

Conjecture 1.1. Fix a prime $q \equiv 1 \pmod{4}$, and $J \geq 1$. Then, for any $a \in \mathbb{N}$, we have

$$N(x; q, (a, a)) = \frac{K}{q^2} \frac{x}{\sqrt{\log x}} \left(1 - \frac{\sqrt{2}\phi(q)}{\pi} \frac{\sqrt{\log \log x}}{\sqrt{\log x}} + \frac{1}{\sqrt{\log x}} \sum_{j=1}^J C_j (\log \log x)^{\frac{1}{2}-j} \right) + O\left(\frac{x}{\log x (\log \log x)^{J+\frac{1}{2}}}\right),$$

for some explicit constants C_j depending only on q . For $a, b \in \mathbb{N}$ with $a \not\equiv b \pmod{q}$, we have

$$N(x; q, (a, b)) = \frac{K}{q^2} \frac{x}{\sqrt{\log x}} \left(1 + \frac{\sqrt{2} \sqrt{\log \log x}}{\pi \sqrt{\log x}} + \frac{C_{a,b}}{\sqrt{\log x}} - \frac{1}{\phi(q) \sqrt{\log x}} \sum_{j=1}^J C_j (\log \log x)^{\frac{1}{2}-j} \right) + O\left(\frac{x}{\log x (\log \log x)^{J+\frac{1}{2}}}\right),$$

with

$$(2) \quad C_{a,b} := \frac{1}{2K} \frac{q}{\phi(q)} \sum_{\chi \neq \chi_0} \bar{\chi}(b-a) C_{q,\chi},$$

where the sum is over the non-principal Dirichlet characters modulo q and $C_{q,\chi}$ is defined in (28). The value of C_1 is given in Conjecture 4.3.

Our heuristic model leading to Conjecture 1.1 follows very closely [LOS16], and as such it is based on the Hardy–Littlewood conjectures for sums of squares, which are stated in Section 3. Our exposition for that section, and many of the results used for the properties of the (conjectural) Hardy–Littlewood constants for sums of squares follow from [FKR17]. Fix $k \geq 1$ and $\{d_1, \dots, d_k\} \subseteq \mathbb{Z}$. We denote $\mathfrak{S}(\{d_1, \dots, d_k\})$ the Hardy–Littlewood constants for k -tuples of sums of two squares defined in Section 3. As the results of [LOS16], our conjecture follows from an average of the Hardy–Littlewood constants, which is one of the main results of our paper.

Theorem 1.2. *Let $q \equiv 1 \pmod{4}$ be a prime. For each Dirichlet character $\chi \neq \chi_0 \pmod{q}$, let $C_{q,\chi}$ be defined by (28). Then, for any $J \geq 1$, and $v \neq 0 \pmod{q}$, we have*

$$\begin{aligned} \sum_{h \geq 1} \mathfrak{S}(\{0, h\}) e^{-h/H} &= H - \frac{2}{K\pi} \sqrt{\log H} + \sum_{j=1}^J c(j) (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right) \\ \sum_{\substack{h \geq 1 \\ h \equiv 0 \pmod{q}}} \mathfrak{S}(\{0, h\}) e^{-h/H} &= \frac{H}{q} - \frac{2}{K\pi} \sqrt{\log H} + \sum_{j=1}^J c_0(j) (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right) \\ \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} \mathfrak{S}(\{0, h\}) e^{-h/H} &= \frac{H}{q} + \frac{1}{2K^2 \phi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \bar{\chi}(v) C_{q,\chi} + \sum_{j=1}^J c_1(j) (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right). \end{aligned}$$

The constants $c(j)$ are absolute while the constants $c_0(j), c_1(j)$ depend only on q , they can all be explicitly computed, in particular the values for $j = 1$ are given in (34) and (38). Moreover they satisfy the relation

$$(3) \quad c_0(j) + \phi(q) c_1(j) = c(j), \quad j \geq 1.$$

By using Theorem 3.4, which is the analogue of the work [MS04] for sums of two squares, we need only to compute a weighted average of the constants $\mathfrak{S}(\{0, h\})$ associated to 2-tuples, while [FKR17] compute a more general average of the constants $\mathfrak{S}(\{h_1, \dots, h_k\})$ associated to k -tuples. Since the Hardy–Littlewood constants $\mathfrak{S}(\{0, h\})$ can be described explicitly with a simple formula from the work of Connors and Keating [CK97], this allows us to get a very precise result exhibiting a small secondary term which gives the bias. A similar average of the constants $\mathfrak{S}(\{0, h\})$ was computed by Smilansky [Smi13], and we also use some of his results. Moreover, the techniques developed in this paper yield a more precise form of the averages considered in [Smi13] and [FKR17].

Proposition 1.3. *Assume the Generalized Riemann Hypothesis. For $\varepsilon > 0$ and $k \geq 2$, we have*

$$\sum_{\substack{1 \leq d_1, \dots, d_k \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, \dots, d_k\}) = H^k + \frac{k(k-1)H^{k-1}}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^{\sigma-1} + F(\sigma)H^{\sigma-1} \log H}{|\sigma-1|^{1/2}} d\sigma + O_{k,\varepsilon}(H^{k-\frac{3}{2}+\varepsilon}),$$

where $F(s) = \zeta(s-1)M(s-1)[(s-1)\zeta(s)]^{1/2} s^{-1}$, with $M(s)$ as defined by (37).

Finally, the heuristic leading to Conjecture 1.1 can be generalized to predict an asymptotic for r successive sums of two squares in arithmetic progressions.

Conjecture 1.4. Fix a prime $q \equiv 1 \pmod{4}$, $r \geq 2$ and $\mathbf{a} = (a_1, \dots, a_r) \in \mathbb{N}^r$. Let

$$N(x; q, \mathbf{a}) := \#\{E_n \leq x : E_{n+i-1} \equiv a_i \pmod{q}\}.$$

We have

$$\begin{aligned} N(x; q, \mathbf{a}) &= \frac{x}{q^r} \frac{K}{\sqrt{\log x}} \left(1 + C_{-1}(\mathbf{a}) \frac{(\log \log x)^{\frac{1}{2}}}{(\log x)^{\frac{1}{2}}} + \frac{C_0(\mathbf{a})}{(\log x)^{\frac{1}{2}}} + \frac{C_1(\mathbf{a})}{(\log \log x)^{\frac{1}{2}} (\log x)^{\frac{1}{2}}} \right) \\ &\quad + O(x(\log \log x)^{-\frac{3}{2}} (\log x)^{-1}), \end{aligned}$$

where

$$\begin{aligned} C_{-1}(\mathbf{a}) &= \frac{q\sqrt{2}}{\pi} \sum_{i=1}^{r-1} \left(\frac{1}{q} - \delta(a_{i+1} \equiv a_i) \right) \\ C_0(\mathbf{a}) &= \sum_{\substack{1 \leq i \leq r-1 \\ a_i \not\equiv a_{i+1} \pmod{q}}} C_{a_i, a_{i+1}} \\ C_1(\mathbf{a}) &= -\frac{qC_1}{\phi(q)} \sum_{i=1}^{r-1} \left(\frac{1}{q} - \delta(a_{i+1} \equiv a_i) \right) + \frac{q\sqrt{2}}{\sqrt{\pi}} \sum_{k=1}^{r-2} \sum_{i=1}^{r-1-k} \frac{\frac{1}{q} - \delta(a_{i+k+1} \equiv a_i)}{k}, \end{aligned}$$

and the constants $C_{a_i, a_{i+1}}$ are defined by (2), and the value of C_1 is given in Conjecture 4.3.

The structure of the paper is as follows: we review in Section 2 the basic properties of sums of two squares, including the secondary terms for the counting function of sums of two squares in arithmetic progressions, which surprisingly we did not find in the literature. We discuss the Hardy–Littlewood conjectures for sums of two squares in Section 3. We present the heuristic model leading to Conjecture 1.1 in Section 4, following Lemke Oliver and Soundararajan [LOS16]; in particular, we explain how the heuristic reduces Conjecture 1.1 to an average of Hardy–Littlewood constants (Theorem 1.2), which we prove in Section 5 using the Selberg–Delange method.

We prove Theorem 3.4, which is an analogue of the main result of Montgomery and Soundararajan [MS04] mentioned above used to justify our heuristic, in Section 6. We then use this result in Section 7 to prove Proposition 1.3 which improve the average results of [FKR17] and [Smi13]. Finally, we explain how to deduce Conjecture 1.4 from our heuristic in Section 8, and we present numerical data in Section 9.

ACKNOWLEDGEMENTS

We thank the organizers of the MOBIUS ANT who invited us to present a preliminary version of the results contained in this paper in a friendly environment with subsequent discussions that led to several improvements in our exposition. We are particularly grateful to Dimitris Koukoulopoulos for constructive discussions, and for pointing out some relevant material from his book [Kou19], and the post of Lucia on Mathoverflow [Luc13].

Our results and conjectures are supported by several numerical data, that were computed using SageMath [S⁺09], PARI/GP [PARI] and Mathematica [MATH]. The authors thank the Centre de Recherches Mathématiques (CRM) in the Université de Montréal for offering their clusters for some of the numerical computations.

The first author was supported by a NSERC Discovery Grant and a FQRNT Team Grant; the second author was supported by the grant KAW 2019.0517 from the Knut and Alice Wallenberg Foundation; the fourth author was supported by a Concordia CUSRA.

2. SUMS OF TWO SQUARES IN ARITHMETIC PROGRESSIONS

By a classical result of Landau [Lan08], we have

$$(4) \quad \sum_{n \leq x} \mathbf{1}_E(n) \sim K \frac{x}{\sqrt{\log x}},$$

where

$$(5) \quad K = \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \pmod{4}} (1 - p^{-2})^{-\frac{1}{2}}$$

is the Landau–Ramanujan constant. We remark that, unlike the prime number theorem and contrarily to a claim of Ramanujan (see e.g. [CM99]), the asymptotic above gives only the main term, and there is no simple integral similar to $\text{li}(x)$ which approximates well the number of sums of two squares up to x . This is caused by the fact that the generating series for sums of two squares has an essential singularity at $s = 1$, its contribution is evaluated by the Selberg–Delange method which gives (4). It is possible to iterate the Selberg–Delange method to write, for any $J \geq 1$,

$$(6) \quad \sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) = Kx \left(\sum_{j=0}^J \frac{c_j}{(\log x)^{1/2+j}} \right) + O\left(x(\log x)^{-3/2-J}\right).$$

Explicit values for the constants c_j can be found in the literature for $c_0 = 1$ [Lan08], $c_1 = 0.581948659 \dots$ [Sta28, Sta29, Sha64, CLM21] and up to c_{15} in [EG18]. It is possible to get an expression for the number of sums of two squares smaller than x with a better error term, but one loses the simplicity of the formula above as a sum of descending powers of \log . We state this result in the next theorem, that we will prove in Section 7. A similar expression for the number of sums of two squares exhibiting squareroot cancellation under the GRH can be found in [GR20, Theorem B.1], inspired by the work of [Ram76]. Such an expression is also suggested in a note of Tenenbaum [Ten15, page 291].

Theorem 2.1. *Let $0 < \varepsilon < 1/2$. There exists a constant $c > 0$ such that*

$$\sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) = \frac{1}{\pi} \int_{1/2+\varepsilon}^1 G(\sigma) \frac{x^\sigma}{\sigma|\sigma-1|^{1/2}} d\sigma + O\left(x \exp\left(-c\sqrt{\log x}\right)\right),$$

where $G(s) = (\zeta(s)(s-1))^{1/2} L(s, \chi_4)^{1/2} (1-2^{-s})^{-1/2} \prod_{p \equiv 3 \pmod{4}} (1-p^{-2s})^{-1/2}$ and χ_4 is the non-trivial Dirichlet character modulo 4, so $G(s)$ is an analytic function for $\text{Re}(s) > 1/2 + \varepsilon$. If we assume the Riemann Hypothesis for $\zeta(s)$ and $L(s, \chi_4)$, we can replace the error term by $O(x^{1/2+\varepsilon})$.

Even if it is more precise (see Table 2), this formula gives somehow less insight on the behaviour of the secondary terms and we come back to the Selberg–Delange method when separating the sums of two squares into congruence classes.

x	Actual	(4)	(6)	Theorem 2.1	(4)	(6)	Theorem 2.1
10^9	173 229 059	167 877 068	172 591 375	173 226 354	1.0319	1.0037	1.00001562
10^{10}	1 637 624 157	1 592 621 708	1 632 873 166	1 637 616 416	1.0283	1.0029	1.00000473
10^{11}	15 570 512 745	15 185 052 177	15 533 945 443	15 570 488 969	1.0254	1.0024	1.00000153
10^{12}	148 736 628 859	145 385 805 874	148 447 838 016	148 736 563 568	1.0230	1.0019	1.00000044

TABLE 2. Comparison of the experimental data for the number of sums of two squares up to x with the asymptotic of (4), the asymptotic of (6) with the first two terms and the integral of Theorem 2.1. The three rightmost columns are the percentage errors. Notice that the error for the integral approximation of Theorem 2.1 agrees with the error term under the Riemann Hypothesis.

Let us now consider the distribution of sums of two squares in arithmetic progression modulo q . For $a \in \mathbb{N}$, following the notations introduced in Section 1, let us denote

$$N(x; q, a) := \#\{E_n \leq x : E_n \equiv a \pmod{q}\}.$$

The case $q \equiv 1 \pmod{4}$ is a prime is particularly simple, and we restrict to that case. We refer the reader to [Rie65, Satz 1] (see also [BW00, Lemma 2.1]) for the general case.

Theorem 2.2. [Rie65, Satz 1] *Let $q \equiv 1 \pmod{4}$ be a prime. Then, for $a \in \mathbb{Z}/q\mathbb{Z}$,*

$$N(x; q, a) := \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mathbf{1}_{\mathbf{E}}(n) \sim \frac{K}{q} \frac{x}{\sqrt{\log x}}.$$

If one compares the above theorem with experimental data for $N(x; q, a)$ as shown in Table 3, there is a discrepancy, and the experimental data shows an excess for $a \equiv 0 \pmod{q}$ compared to the other classes modulo q . This is caused by secondary terms that depends on the class a , which do not seem to appear in the literature, and we compute the first such term in Theorem 2.4 below. The proof uses the Selberg–Delange method which evaluates the contribution of essential singularities by using Hankel’s formula, replacing Cauchy’s residue theorem for this case. We state below the version of the method needed for the proof of Theorem 2.4, and we refer the reader to [Ten15, Chapter II.5] and [Kou19, Chapter 13], and to Section 5 for more details.

Theorem 2.3. [Kou19, Theorem 13.2] *Let $f(n)$ be a multiplicative function with generating function $F(s) = \sum_{n \geq 1} f(n)n^{-s}$. Suppose there exists $\kappa \in \mathbb{C}$ be such that for x large enough*

$$\sum_{p \leq x} f(p) \log p = \kappa x + O_A(x/(\log x)^A),$$

for each fixed $A > 0$, and such that $|f(n)| \leq \tau_k(n)$ for some $k \in \mathbb{N}$, where τ_k is the k -th divisor function. For $j \geq 0$, let \tilde{c}_j be the Taylor coefficients about 1 of the function $(s-1)^\kappa F(s)/s$. Then, for any $J \in \mathbb{N}$, and x large enough, we have

$$\sum_{n \leq x} f(n) = x \sum_{j=0}^J \tilde{c}_j \frac{(\log x)^{\kappa-j-1}}{\Gamma(\kappa-j)} + O\left(\frac{x}{(\log x)^{J+2-\operatorname{Re}(\kappa)}}\right).$$

Theorem 2.4. *Let $q \equiv 1 \pmod{4}$ be a prime, and let K and c_1 be as defined above. Then,*

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mathbf{1}_{\mathbf{E}}(n) = \frac{K}{q} x \sum_{j=0}^J \frac{c_{j,a}}{(\log x)^{1/2+j}} + O\left(\frac{x}{(\log x)^{J+3/2}}\right),$$

where

$$(7) \quad c_{0,a} = c_0 = 1 \quad \text{and} \quad c_{1,a} := \begin{cases} c_1 + \frac{\log q}{2} & \text{if } a \equiv 0 \pmod{q} \\ c_1 - \frac{\log q}{2(q-1)} & \text{otherwise.} \end{cases}.$$

We refer the reader to Table 3 for the comparison between the numerical data and Theorem 2.4.

q	a	$N(x; q, a)$	Main term	Main and secondary terms
0		30 700 929 089	29 077 161 174	30 536 403 581
1		29 508 931 067		29 477 858 608
5	2	29 508 917 111		
	3	29 508 920 778		
	4	29 508 930 814		

TABLE 3. Comparison of the experimental data for $N(x; q, a)$ and the asymptotic of Theorem 2.4 using only the main term, or the main term and the first secondary term for $q = 5$ and $x = 10^{12}$. The average of $N(x; q, a)$ is $\approx 29\,747\,325\,771$.

Proof. Let $F(s) := \sum_{n \geq 1} \mathbf{1}_{\mathbf{E}}(n)n^{-s}$ be the generating series for sums of two squares. Using the well-known fact that n is a sum of two squares if and only if $v_p(n)$ is even for all primes $p \equiv 3 \pmod{4}$, it is easy to see that for $\operatorname{Re}(s) > 1$,

$$\begin{aligned} F^2(s) &= \prod_{p \not\equiv 3 \pmod{4}} \left(1 - \frac{1}{p^s}\right)^{-2} \prod_{p \equiv 3 \pmod{4}} \left(1 - \frac{1}{p^{2s}}\right)^{-2} \\ &= \zeta(s)L(s, \chi_4) \left(1 - \frac{1}{2^s}\right)^{-1} \prod_{p \equiv 3(4)} \left(1 - \frac{1}{p^{2s}}\right)^{-1} \end{aligned}$$

where χ_4 is the non-principal Dirichlet character modulo 4. Landau [Lan08] also showed that in a neighborhood of $s = 1$,

$$\frac{F(s)}{s^2} = \sum_{\ell \geq 0} ia_\ell (1-s)^{\ell-1/2},$$

with $a_0 = K\sqrt{\pi}$ and $a_1 = a_0(2c_1 + 1)$ [Sha64]. Applying Theorem 2.3 with $\kappa = 1/2$, we get (6), using the values a_0, a_1 to get explicit values for the first two Taylor coefficients of $(s-1)^{1/2}F(s)/s$.

To introduce the congruence condition, we write for $a \not\equiv 0 \pmod{q}$,

$$(8) \quad N(x; q, a) = \frac{1}{q-1} \sum_{\chi \pmod{q}} \bar{\chi}(a) \sum_{n \leq x} \chi(n) \mathbf{1}_{\mathbf{E}}(n),$$

and we denote the generating function of $f_\chi(n) = \chi(n) \mathbf{1}_{\mathbf{E}}(n)$ by $F_\chi(s) := \sum_{n \geq 1} \chi(n) \mathbf{1}_{\mathbf{E}}(n)n^{-s}$. For χ_0 the principal character modulo q and $\chi \neq \chi_0$, we have for $\operatorname{Re}(s) > 1$,

$$(9) \quad \begin{aligned} F_\chi^2(s) &= L(s, \chi)L(s, \chi_4\chi) \left(1 - \frac{\chi(2)}{2^s}\right)^{-1} \prod_{p \equiv 3(4)} \left(1 - \frac{\chi^2(p)}{p^{2s}}\right)^{-1} \\ F_{\chi_0}^2(s) &= \left(1 - \frac{1}{q^s}\right)^2 F^2(s). \end{aligned}$$

For $\chi \neq \chi_0$, since $F_\chi(s)$ is analytic for $\operatorname{Re}(s) > 1/2$, we have for any $\varepsilon > 0$ that

$$\sum_{n \leq x} \chi(n) \mathbf{1}_{\mathbf{E}}(n) = O\left(x^{1/2+\varepsilon}\right),$$

and the theorem will follow by evaluating $\sum_{n \leq x} \chi_0(n) \mathbf{1}_{\mathbf{E}}(n)$ with the Selberg–Delange method. Let \tilde{b}_j be the Taylor coefficients of $(s-1)^{1/2}F_{\chi_0}(s)/s$ around $s = 1$, and \tilde{c}_j are the Taylor coefficients of $(s-1)^{1/2}F(s)/s$ around $s = 1$. From (9), it is easy to compute

$$\begin{aligned} \tilde{b}_0 &= (1 - q^{-1}) \tilde{c}_0 = (1 - q^{-1}) K\sqrt{\pi} \\ \tilde{b}_1 &= (1 - q^{-1}) \tilde{c}_1 + \frac{\log q}{q} \tilde{c}_0 = K\sqrt{\pi} \left(\frac{\log q}{q} - 2c_1(1 - q^{-1}) \right). \end{aligned}$$

Applying Theorem 2.3 with $\kappa = 1/2$, to estimate the sum $\sum_{n \leq x} \chi_0(n) \mathbf{1}_{\mathbf{E}}(n)$, and replacing in (8), we get the statement of the theorem when $a \not\equiv 0 \pmod{q}$, with

$$\frac{K}{q} c_{0,a} = \frac{\tilde{b}_0}{(q-1)\Gamma(1/2)}, \quad \frac{K}{q} c_{1,a} = \frac{\tilde{b}_1}{(q-1)\Gamma(-1/2)}.$$

For $a \equiv 0 \pmod{q}$, we use the above and (6) to obtain

$$\begin{aligned} N(x; q, 0) &= \sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) - \sum_{a \not\equiv 0 \pmod{q}} N(x; q, a) \\ &= \frac{K}{q} x \left(\frac{1}{(\log x)^{1/2}} + \left(c_1 + \frac{\log q}{2} \right) \frac{1}{(\log x)^{3/2}} + \sum_{j=2}^J \frac{c_j - (q-1)c_{j,1}}{(\log x)^{1/2+j}} \right) + O\left(x(\log x)^{-3/2-J}\right), \end{aligned}$$

which completes the proof. \square

3. HARDY–LITTLEWOOD CONJECTURES IN ARITHMETIC PROGRESSIONS FOR SUM OF TWO SQUARES

We state in this section the analogue of the Hardy–Littlewood prime k -tuple conjectures for the case of sums of two squares, following [FKR17]. We also state new bounds on the average of the Hardy–Littlewood constant in this context that are useful in our heuristic for Conjecture 1.1, but are also interesting in themselves as they are related to the distribution of gaps between sums of two squares.

For $k \geq 1$, let $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$, and

$$R_k(\mathcal{H}; x) := \frac{1}{x} \sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n + h_1) \dots \mathbf{1}_{\mathbf{E}}(n + h_k).$$

In the case $\mathcal{H} = \{0\}$, we have

$$R_1(x) := R_1(\{0\}; x) = \frac{1}{x} \sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) \sim \frac{K}{\sqrt{\log x}}.$$

The philosophy of the Hardy–Littlewood conjecture is that the events $\mathbf{1}_{\mathbf{E}}(n + h_i)$ are “independent”, and the probability that $n + h_i$ are simultaneously sums of two squares for $1 \leq i \leq k$ is the product of the probabilities, which is (ignoring the small differences between $\log n$ or $\log n + h_i$)

$$\left(\frac{K}{\sqrt{\log n}} \right)^k.$$

Of course, the events are not really independent, so we adjust by considering the probabilities that $n + h_i$ are sums of two squares modulo p versus the probability that k independent integers are sums of two squares modulo p . To do so, for each prime p , we define

$$\delta_{\mathcal{H}}(p) = \lim_{\alpha \rightarrow \infty} \frac{\#\{0 \leq a < p^\alpha : \forall h \in \mathcal{H}, a + h \equiv \square + \square \pmod{p^\alpha}\}}{p^\alpha}.$$

Since $\delta_{\mathcal{H}}(p) = 1$ for $p \equiv 1 \pmod{4}$ (see e.g. [FKR17, Proposition 5.1]), we define the singular series for $\mathcal{H} = \{h_1, \dots, h_k\}$ by

$$(10) \quad \mathfrak{S}(\mathcal{H}) := \prod_{p \not\equiv 1 \pmod{4}} \frac{\delta_{\mathcal{H}}(p)}{(\delta_{\{0\}}(p))^k}.$$

It is proven in [FKR17] that the limit defining $\delta_{\mathcal{H}}(p)$ exists, and the Euler product converges to a non-zero limit provided that $\delta_{\mathcal{H}}(p) > 0$ for all $p \not\equiv 1 \pmod{4}$.

Conjecture 3.1. [FKR17, Conjecture 1.1] Fix $k \geq 1$, and $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$. If $\mathfrak{S}(\mathcal{H}) > 0$, then

$$R_k(\mathcal{H}; x) \sim \mathfrak{S}(\mathcal{H}) (R_1(x))^k \sim \mathfrak{S}(\mathcal{H}) \left(\frac{K}{\sqrt{\log x}} \right)^k$$

This conjecture is still open, but it is known that $\sum_n \mathbf{1}_{\mathbf{E}}(n + h_1) \dots \mathbf{1}_{\mathbf{E}}(n + h_k)$ is infinite for $k = 2, 3$ by the work of Hooley [Hoo71, Hoo73].

It is not straightforward to give a simple formula for the singular series $\mathfrak{S}(\mathcal{H})$ for a given set \mathcal{H} (see Section 6), except the trivial cases $\mathfrak{S}(\emptyset) = \mathfrak{S}(\{h\}) = 1$. For $\mathcal{H} = \{0, h\}$, Connors and Keating [CK97] computed

$$(11) \quad \mathfrak{S}(\{0, h\}) = \frac{1}{2K^2} W_2(h) \prod_{\substack{p \equiv 3 \pmod{4} \\ p|h}} \frac{1 - p^{-v_p(h)-1}}{1 - p^{-1}},$$

where

$$W_2(h) = \begin{cases} 1 & \text{if } 2 \nmid h \\ 2 - 3 \cdot 2^{-v_2(h)} & \text{otherwise,} \end{cases}$$

and v_p is the p -adic valuation.

Notice that it means that $\mathfrak{S}(\mathcal{H}) > 0$ when $k = 2$. This can also be proven for $k = 3$, but for general k , we can find sets \mathcal{H} such that $\mathfrak{S}(\mathcal{H}) = 0$. It is easy to see that $\sum_n \mathbf{1}_{\mathbf{E}}(n + h_1) \dots \mathbf{1}_{\mathbf{E}}(n + h_k)$ is finite when $\mathfrak{S}(\mathcal{H}) = 0$.

We now state a slight generalization of the Hardy–Littlewood conjecture where n is restricted to an arithmetic progression modulo q .

Conjecture 3.2. (Hardy–Littlewood for sums of two squares in arithmetic progressions) Fix $k \geq 1$, and $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$. Let $q \equiv 1 \pmod{4}$ be a prime, and $a \in \mathbb{Z}$. If $\mathfrak{S}(\mathcal{H}) > 0$, then

$$\begin{aligned} R_k(\mathcal{H}; x, q, a) &:= \frac{1}{x} \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mathbf{1}_{\mathbf{E}}(n + h_1) \dots \mathbf{1}_{\mathbf{E}}(n + h_k) \\ &\sim \frac{\mathfrak{S}(\mathcal{H})}{q} \left(\frac{K}{\sqrt{\log x}} \right)^k. \end{aligned}$$

We remark that unlike the generalized Hardy–Littlewood conjecture of [LOS16], we do not need to adjust the local factors at the prime numbers dividing q in $\mathfrak{S}(\mathcal{H})$ since we fixed q to be prime with $q \equiv 1 \pmod{4}$, and this prime does not appear in the Euler product (10) defining $\mathfrak{S}(\mathcal{H})$.

In Conjectures 3.1 and 3.2, we used $K/\sqrt{\log n}$ for the probability that n is a sum of two squares. As the secondary term for this probability depends on the residue class modulo q from Theorem 2.4, we get more precise results by using this second term to refine the probability in Conjecture 3.2. We state that in the conjecture below, and we used it to illustrate the fit with the numerical data in Table 4, but not in the rest of the paper while getting in the heuristic model leading to Conjecture 1.1 and Conjecture 1.4 (as those secondary terms would be smaller than some error terms occurring in the heuristic).

Conjecture 3.3. (Refined Hardy–Littlewood in arithmetic progressions) Fix $k \geq 1$, and $\mathcal{H} = \{h_1, \dots, h_k\} \subseteq \mathbb{Z}$. Let $q \equiv 1 \pmod{4}$ be a prime, and $a \in \mathbb{Z}$. If $\mathfrak{S}(\mathcal{H}) > 0$, then

$$R_k(\mathcal{H}; x, q, a) = \frac{\mathfrak{S}(\mathcal{H})}{q} K^k \left(\frac{1}{(\log x)^{k/2}} + \frac{1}{(\log x)^{k/2+1}} \sum_{h \in \mathcal{H}} c_{1, h+a} + O\left(\frac{1}{(\log x)^{k/2+2}}\right) \right),$$

where $c_{1, h}$ is defined by (7).

a	h	$xR_k(\mathcal{H}; x, q, a)$	Main term	Main and secondary term	Err ₁	Err ₂
0	1	3 906 419 030	3 619 120 683	3 850 620 130	1.0794	1.0145
1	1	3 751 339 794	3 619 120 683	3 718 867 172	1.0365	1.0087
1	2	1 925 818 092	1 809 560 341	1 859 433 586	1.0642	1.0357
0	5	4 062 607 000	3 619 120 682	3 982 373 088	1.1225	1.0201

TABLE 4. Numerical data versus Conjecture 3.3 for $\mathcal{H} = \{0, h\}$, $x = 10^{12}$, $q = 5$. The third column shows the numerical data, the 4-th and 5-th columns show the product of x and the prediction of Conjecture 3.3 with the main term, and with the main and first secondary term respectively. The last two columns show their percentage errors, respectively.

Finally, we need an equivalent form of Conjecture 3.2, inspired by the work of Montgomery and Soundararajan [MS04] for the case of primes, namely

$$(12) \quad \frac{1}{x} \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \prod_{h \in \mathcal{H}} \left(\mathbf{1}_{\mathbf{E}}(n + h) - \frac{K}{\sqrt{\log n}} \right) \sim \frac{\mathfrak{S}_0(\mathcal{H})}{q} \left(\frac{K}{\sqrt{\log x}} \right)^{|\mathcal{H}|}.$$

Assuming that Conjecture 3.2 holds, we get relations between the constants $\mathfrak{S}_0(\mathcal{H})$ and $\mathfrak{S}(\mathcal{H})$, and it is easy to see that

$$\begin{aligned} \mathfrak{S}_0(\emptyset) &= 1 \\ \mathfrak{S}_0(\{h\}) &= 0 \\ \mathfrak{S}_0(\{h_1, h_2\}) &= \mathfrak{S}(\{h_1, h_2\}) - 1, \end{aligned}$$

and that for a general set \mathcal{H} ,

$$(13) \quad \mathfrak{S}_0(\mathcal{H}) = \sum_{\mathcal{T} \subseteq \mathcal{H}} (-1)^{|\mathcal{H} \setminus \mathcal{T}|} \mathfrak{S}(\mathcal{T}).$$

Mirroring [MS04], we prove in Section 6 the following result, which is critical to justify our heuristic.

Theorem 3.4. *Let $\mathfrak{S}_0(\mathcal{H})$ the constants defined by (13). Then, for any $k \geq 1$ and $\varepsilon > 0$, we have*

$$\sum_{\substack{\mathcal{H} \subseteq [1, h] \\ |\mathcal{H}| = k}} \mathfrak{S}_0(\mathcal{H}) \ll_{k, \varepsilon} h^{\frac{k}{2} + \varepsilon}.$$

Note that our result is weaker than the result of Montgomery and Soundararajan who computed an asymptotic for the average of Theorem 3.4 in the case of primes [MS04, Theorem 2]. We did not pursue that as Theorem 3.4 is sufficient for our application. We observe that, similar upper bounds are given by Aryan [Ary15a, Ary15b] in the general context of k -tuples of reduced residues.

4. HEURISTIC FOR THE CONJECTURE

We now develop a heuristic leading to Conjecture 1.1 following [LOS16]. Let $q \equiv 1 \pmod{4}$ be a prime, and we recall that

$$N(x; q, (a, b)) = \#\{E_n \leq x : E_n \equiv a \pmod{q}, E_{n+1} \equiv b \pmod{q}\}.$$

We first write

$$(14) \quad N(x; q, (a, b)) = \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \sum_{\substack{h > 0 \\ h \equiv b - a \pmod{q}}} \mathbf{1}_{\mathbf{E}}(n) \mathbf{1}_{\mathbf{E}}(n+h) \prod_{t=1}^{h-1} (1 - \mathbf{1}_{\mathbf{E}}(n+t)).$$

We introduce the notation

$$\tilde{\mathbf{1}}_{\mathbf{E}}(n) = \mathbf{1}_{\mathbf{E}}(n) - \frac{K}{\sqrt{\log n}},$$

and for each fixed h in (14), we study the sum

$$S_h := \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \left(\frac{K}{\sqrt{\log n}} + \tilde{\mathbf{1}}_{\mathbf{E}}(n) \right) \left(\frac{K}{\sqrt{\log(n+h)}} + \tilde{\mathbf{1}}_{\mathbf{E}}(n+h) \right) \prod_{0 < t < h} \left(1 - \frac{K}{\sqrt{\log(n+t)}} - \tilde{\mathbf{1}}_{\mathbf{E}}(n+t) \right).$$

If we ignore the small differences among $\sqrt{\log n}$, $\sqrt{\log(n+h)}$, and $\sqrt{\log(n+t)}$ and we expand out the product, we get

$$S_h = \sum_{\mathcal{A} \subset \{0, h\}} \sum_{\mathcal{T} \subset [1, h-1]} (-1)^{|\mathcal{T}|} \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \left(\frac{K}{\sqrt{\log n}} \right)^{2-|\mathcal{A}|} \prod_{\substack{t \in [1, h-1] \\ t \notin \mathcal{T}}} \left(1 - \frac{K}{\sqrt{\log n}} \right) \prod_{t \in \mathcal{A} \cup \mathcal{T}} \tilde{\mathbf{1}}_{\mathbf{E}}(n+t)$$

Finally, denoting

$$\alpha(n) = 1 - \frac{K}{\sqrt{\log n}},$$

and using (12), we conjecture that

$$\begin{aligned} S_h &= \sum_{\mathcal{A} \subset \{0, h\}} \sum_{\mathcal{T} \subset [1, h-1]} (-1)^{|\mathcal{T}|} \sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \left(\frac{K}{\sqrt{\log(n)}} \right)^{2-|\mathcal{A}|} \alpha(n)^{h-1-|\mathcal{T}|} \prod_{t \in \mathcal{A} \cup \mathcal{T}} \tilde{\mathbf{1}}_{\mathbf{E}}(n+t) \\ &\sim \frac{x}{q} \sum_{\mathcal{A} \subset \{0, h\}} \sum_{\mathcal{T} \subset [1, h-1]} (-1)^{|\mathcal{T}|} \mathfrak{S}_0(\mathcal{A} \cup \mathcal{T}) \left(\frac{K}{\sqrt{\log x}} \right)^{2+|\mathcal{T}|} \alpha(x)^{h-1-|\mathcal{T}|}. \end{aligned}$$

We emphasize that this is a heuristic argument: in obtaining this expression for S_h , we have not paid attention to the error terms in (12), in particular on the dependency on the size of the sets $\mathcal{A} \cup \mathcal{T}$ and on h .

Summing S_h over all $h \equiv b - a \pmod{q}$, this gives the conjectural estimate

$$(15) \quad N(x; q, (a, b)) \sim \frac{x}{q} \alpha(x)^{-1} \left(\frac{K}{\sqrt{\log x}} \right)^2 \mathcal{D}(a, b; x),$$

where

$$(16) \quad \mathcal{D}(a, b; x) = \sum_{\substack{h > 0 \\ h \equiv b-a \pmod{q}}} \sum_{\mathcal{A} \subset \{0, h\}} \sum_{\mathcal{T} \subset [1, h-1]} (-1)^{|\mathcal{T}|} \mathfrak{S}_0(\mathcal{A} \cup \mathcal{T}) \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^{|\mathcal{T}|} \alpha(x)^h.$$

In order to evaluate (16), we will use the following notations. Let

$$(17) \quad \alpha(x)^h = \left(1 - \frac{K}{\sqrt{\log x}} \right)^h = e^{-h/H} \iff H = -\frac{1}{\log \alpha(x)},$$

which implies that

$$\begin{aligned} H &= \frac{\sqrt{\log x}}{K} - \frac{1}{2} + O\left((\log x)^{-1/2}\right) \\ \log H &= \frac{1}{2} \log \log x - \log K + O\left((\log x)^{-1/2}\right). \end{aligned}$$

4.1. Discarding the singular series involving larger sets. We approximate $\mathcal{D}(a, b; x)$ by discarding all the singular series where $\mathcal{A} \cup \mathcal{T}$ has more than 2 elements, which is justified by Theorem 3.4. We separate in 3 cases, depending on the possible choices for the set $\mathcal{A} \subseteq \{0, h\}$. We use the notation defined in (17) for H , and the bound $\sum_{\substack{h > 0 \\ h \equiv v \pmod{q}}} \alpha(x)^h h^\ell \ll_\ell H^{\ell+1}$ for any $\ell \geq 0$, and $v \in \mathbb{Z}$.

If $\mathcal{A} = \emptyset$, then for $k \geq 3$, we deduce from Theorem 3.4 that

$$\begin{aligned} \sum_{\substack{h > 0 \\ h \equiv b-a \pmod{q}}} \sum_{\substack{\mathcal{T} \subset [1, h-1] \\ |\mathcal{T}|=k}} \mathfrak{S}_0(\mathcal{T}) \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^k \alpha(x)^h &\ll_k \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^k \sum_{\substack{h > 0 \\ h \equiv b-a \pmod{q}}} h^{\frac{k}{2} + \varepsilon} \alpha(x)^h \\ &\ll_k \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^k H^{1 + \frac{k}{2} + \varepsilon} \ll_k (\log x)^{-\frac{k}{4} + \frac{1}{2} + \varepsilon}. \end{aligned}$$

If $\mathcal{A} = \{h\}$ and $|\mathcal{A} \cup \mathcal{T}| \geq 3$, we have for $k \geq 2$

$$\begin{aligned} \sum_{\substack{h > 0 \\ h \equiv b-a \pmod{q}}} \sum_{\substack{\mathcal{T} \subset [1, h-1] \\ |\mathcal{T}|=k}} \mathfrak{S}_0(\mathcal{T} \cup \{h\}) \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^k \alpha(x)^h \\ \approx_k \frac{1}{q} \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^k \sum_{\substack{\mathcal{D} \subset [1, H] \\ |\mathcal{D}|=k+1}} \mathfrak{S}_0(\mathcal{D}) \ll_k \frac{1}{q} \left(\frac{K}{\alpha(x) \sqrt{\log x}} \right)^k H^{\frac{k+1}{2} + \varepsilon} \ll_k \frac{1}{q} (\log x)^{-\frac{k}{4} + \frac{1}{4} + \varepsilon}, \end{aligned}$$

where we are approximating the sum over h and \mathcal{T} of the first line as the sum over all \mathcal{D} of size $k+1$ contained in $[1, H]$, which we then bound by Theorem 3.4. We obtain the same bound for $\mathcal{A} = \{0\}$ using the fact that \mathfrak{S}_0 is invariant by translation.

Finally, in the case $\mathcal{A} = \{0, h\}$, we introduce an extra average. Since \mathfrak{S}_0 is translation invariant, we have

$$\sum_{s \geq 1} \mathfrak{S}_0(\{s, t_1 + s, \dots, t_k + s, h + s\}) e^{-s/H} = \mathfrak{S}_0(\{0, t_1, \dots, t_k, h\}) \sum_{s \geq 1} e^{-s/H} \approx \mathfrak{S}_0(\{0, t_1, \dots, t_k, h\}) H,$$

and using this, we get for $k \geq 1$

$$\begin{aligned}
& \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right)^k \sum_{\substack{h>0 \\ h \equiv b-a \pmod{q}}} \sum_{\substack{\mathcal{T} \subset [1, h-1] \\ |\mathcal{T}|=k}} \mathfrak{S}_0(\mathcal{T} \cup \{0, h\}) \alpha(x)^h \\
& \approx \frac{1}{qH} \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right)^k \sum_{s \geq 1} \sum_{h \geq 1} \sum_{0 < t_1 < \dots < t_k < h} \mathfrak{S}_0(\{s, t_1 + s, \dots, t_k + s, h + s\}) e^{-(s+h)/H} \\
& \approx \frac{1}{qH} \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right)^k \sum_{0 < s < t'_1 < \dots < t'_k < h' < 2H} \mathfrak{S}_0(\{s, t'_1, \dots, t'_k, h'\}) \\
& \ll_k \frac{1}{q} \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right)^k (2H)^{-1 + \frac{k+2}{2} + \varepsilon} \ll_k \frac{1}{q} (\log x)^{-\frac{k}{4} + \varepsilon}.
\end{aligned}$$

Discarding all the singular series where $\mathcal{A} \cup \mathcal{T}$ has more than 2 elements from (16), and working again heuristically by ignoring the dependence on $|\mathcal{A} \cup \mathcal{T}|$ in the error terms, we are led to the model

$$\mathcal{D}(a, b; x) = (\mathcal{D}_0 + \mathcal{D}_1 + \mathcal{D}_2)(a, b; x) + O_\varepsilon((\log x)^{-\frac{1}{4} + \varepsilon}),$$

where

$$\begin{aligned}
\mathcal{D}_0(a, b; x) &= \sum_{\substack{h>0 \\ h \equiv b-a \pmod{q}}} \left(1 + \mathfrak{S}_0(\{0, h\}) \right) \alpha(x)^h \\
\mathcal{D}_1(a, b; x) &= - \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right) \sum_{\substack{h>0 \\ h \equiv b-a \pmod{q}}} \sum_{t \in [1, h-1]} \left(\mathfrak{S}_0(\{0, t\}) + \mathfrak{S}_0(\{t, h\}) \right) \alpha(x)^h \\
\mathcal{D}_2(a, b; x) &= \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right)^2 \sum_{\substack{h>0 \\ h \equiv b-a \pmod{q}}} \sum_{1 \leq t_1 < t_2 < h} \left(\mathfrak{S}_0(\{t_1, t_2\}) \right) \alpha(x)^h.
\end{aligned}$$

Replacing in (15), we then conjecture that up to error term of order $x(\log x)^{-\frac{5}{4} + \varepsilon}$, we have

$$(18) \quad N(x; q, (a, b)) \sim \frac{x}{q} \alpha(x)^{-1} \left(\frac{K}{\sqrt{\log x}} \right)^2 (\mathcal{D}_0 + \mathcal{D}_1 + \mathcal{D}_2)(a, b; x).$$

4.2. Evaluation of the sums of singular series involving sets of size 2. In order to evaluate (18), we first evaluate the simple exponential sums. We will use the notation

$$f(v; q) := \begin{cases} -\frac{1}{2} & v = 0 \\ \frac{q-2v}{2q} & 1 \leq v \leq q-1 \end{cases}$$

which gives

$$\begin{aligned}
E(H) &:= \sum_{h>0} e^{-h/H} = H - \frac{1}{2} + O(H^{-1}) = \frac{\sqrt{\log x}}{K} - 1 + O((\log x)^{-1/2}) \\
E(q, v; H) &:= \sum_{\substack{h>0 \\ h \equiv v \pmod{q}}} e^{-h/H} = \frac{H}{q} + f(v; q) + O(H^{-1}) = \frac{\sqrt{\log x}}{Kq} + f(v; q) - \frac{1}{2q} + O((\log x)^{-1/2}).
\end{aligned}$$

Let

$$\begin{aligned}
(19) \quad S(q, v; H) &:= \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} \mathfrak{S}(\{0, h\}) e^{-h/H} \\
S_0(q, v; H) &:= \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} \mathfrak{S}_0(\{0, h\}) e^{-h/H} = \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} (\mathfrak{S}(\{0, h\}) - 1) e^{-h/H}.
\end{aligned}$$

and

$$\begin{aligned} S(H) &:= \sum_{h \geq 1} \mathfrak{S}(\{0, h\}) e^{-h/H} = \sum_{v \pmod{q}} S(q, v; H) \\ S_0(H) &:= \sum_{h \geq 1} \mathfrak{S}_0(\{0, h\}) e^{-h/H} = \sum_{v \pmod{q}} S_0(q, v; H). \end{aligned}$$

We then have

$$(20) \quad \begin{aligned} S_0(q, v; H) &= S(q, v; H) - \frac{H}{q} - f(v; q) + O(H^{-1}) \\ S_0(H) &= S(H) - H + \frac{1}{2} + O(H^{-1}). \end{aligned}$$

Using Theorem 1.2, we evaluate $\mathcal{D}_0(a, b; x)$, $\mathcal{D}_1(a, b; x)$ and $\mathcal{D}_2(a, b; x)$.

Proposition 4.1. *Let $q \equiv 1 \pmod{4}$ be a prime. For $j \geq 1$, let $c(j)$ be the constants from Theorem 1.2. Then,*

$$\begin{aligned} &\mathcal{D}_0(a, b; x) + \mathcal{D}_1(a, b; x) + \mathcal{D}_2(a, b; x) \\ &= S(q, b-a; H) + \frac{2}{qK\pi} (\log H)^{1/2} - \frac{1}{2q} - \frac{1}{q} \sum_{j=1}^J c(j) (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J} + \frac{\sqrt{\log H}}{\sqrt{\log x}}\right) \end{aligned}$$

where we use the change of variables (17). We remark that the error term $(\log H)^{-1/2-J}$ is the largest one, for any value of J .

Proof. First, notice that $\mathcal{D}_0(a, b; x) = S(q, b-a; H)$. For $\mathcal{D}_1(a, b; x)$, we first compute

$$(21) \quad \begin{aligned} \sum_{\substack{h \geq 2 \\ h \equiv b-a \pmod{q}}} \sum_{1 \leq t \leq h-1} \mathfrak{S}_0(\{0, t\}) e^{-h/H} &= \sum_{t \geq 1} \mathfrak{S}_0(\{0, t\}) e^{-t/H} \sum_{\substack{h \geq 1 \\ h \equiv b-a-t \pmod{q}}} e^{-h/H} \\ &= \left(\frac{H}{q} + O(1)\right) S_0(H), \end{aligned}$$

and

$$-\left(\frac{K}{\alpha(x)\sqrt{\log x}}\right) \sum_{\substack{h > 0 \\ h \equiv b-a \pmod{q}}} \sum_{1 \leq t \leq h-1} \mathfrak{S}_0(\{0, t\}) e^{-h/H} = \left(-\frac{1}{q} + O\left(\frac{1}{\sqrt{\log x}}\right)\right) S_0(H).$$

We get a similar estimate for the second sum in $\mathcal{D}_1(a, b; y)$ involving $\mathfrak{S}_0(\{t, h\})$ by making a change of variable to replace it by $\mathfrak{S}_0(\{0, r\})$ with $r = h - t$, which gives

$$\mathcal{D}_1(a, b; x) = \left(-\frac{2}{q} + O\left((\log x)^{-1/2}\right)\right) S_0(H).$$

Similarly, for $\mathcal{D}_2(a, b; x)$, we first compute

$$(22) \quad \begin{aligned} &\sum_{\substack{h \geq 3 \\ h \equiv b-a \pmod{q}}} \sum_{1 \leq t_1 < t_2 < h} \mathfrak{S}_0(\{t_1, t_2\}) e^{-h/H} \\ &= \sum_{1 \leq t_1 < t_2} \mathfrak{S}_0(\{0, t_2 - t_1\}) \sum_{\substack{h \equiv b-a \pmod{q} \\ h \geq t_2+1}} e^{-h/H} = \sum_{r \geq 1} \mathfrak{S}_0(\{0, r\}) \sum_{t_2 \geq r+1} e^{-t_2/H} \sum_{\substack{h' \geq 1 \\ h' \equiv b-a-t_2 \pmod{q}}} e^{-h'/H} \\ &= \sum_{r \geq 1} \mathfrak{S}_0(\{0, r\}) e^{-r/H} \sum_{t'_2 \geq 1} e^{-t'_2/H} \sum_{\substack{h' \geq 1 \\ h' \equiv b-a-t'_2-r \pmod{q}}} e^{-h'/H} = \left(\frac{H^2}{q} + O(H)\right) S_0(H), \end{aligned}$$

and replacing in the definition of $\mathcal{D}_2(a, b; x)$, we have

$$\mathcal{D}_2(a, b; x) = \left(\frac{1}{q} + O\left((\log x)^{-1/2}\right) \right) S_0(H).$$

Using Theorem 1.2 and (20) to evaluate $S_0(H)$, this completes the proof. \square

One can be more precise regarding the dependence on the congruence classes by separating the sum over t in (21) and the sum over r in (22) in congruence classes modulo q . In particular, the following refinement of Proposition 4.1 will be used for numerical testing. The proof follows directly from the proof of Proposition 4.1, and we omit it.

Proposition 4.2. *Let $q \equiv 1 \pmod{4}$ be a prime. Then*

$$\begin{aligned} & \mathcal{D}_0(a, b; x) + \mathcal{D}_1(a, b; x) + \mathcal{D}_2(a, b; x) \\ &= E(q, b-a; H) + S_0(q, b-a; H) - 2 \frac{K}{\alpha(x)\sqrt{\log x}} \sum_{c \pmod{q}} S_0(q, b-a-c; H) E(q, c; H) \\ &+ \left(\frac{K}{\alpha(x)\sqrt{\log x}} \right)^2 \sum_{c, d \pmod{q}} S_0(q, b-a-c-d; H) E(q, c; H) E(q, d; H). \end{aligned}$$

4.3. Completing the heuristic. We now deduce Conjecture 1.1, by replacing Theorem 1.2 and Proposition 4.1 in (18). If $a \equiv b \pmod{q}$, we have

$$\begin{aligned} N(x; q, (a, a)) &= \frac{xK^2}{q \log x} \left(1 + \frac{K}{\sqrt{\log x}} + O\left(\frac{1}{\log x}\right) \right) \left[\frac{\sqrt{\log x}}{Kq} - \frac{2(q-1)}{qK\pi} (\log H)^{1/2} - \frac{1}{q} \right. \\ &\quad \left. + \sum_{j=1}^J \left(c_0(j) - \frac{c(j)}{q} \right) (\log H)^{1/2-j} + O\left((\log H)^{-J-1/2}\right) \right] \\ (23) \quad &= \frac{Kx}{q^2 \sqrt{\log x}} \left[1 + \frac{1}{\sqrt{\log x}} \sum_{j=0}^J b_0(j) (\log H)^{1/2-j} + O\left(\frac{1}{\sqrt{\log x}(\log H)^{J+1/2}}\right) \right] \end{aligned}$$

where $b_0(0) = -2(q-1)/\pi$ and $b_0(j) = K(qc_0(j) - c(j))$ for $j \geq 1$.

If $a \not\equiv b \pmod{q}$, we have

$$\begin{aligned} N(x; q, (a, b)) &= \frac{xK^2}{q \log x} \left(1 + \frac{K}{\sqrt{\log x}} + O\left(\frac{1}{\log x}\right) \right) \left[\frac{\sqrt{\log x}}{Kq} + \frac{2}{qK\pi} (\log H)^{1/2} - \frac{1}{q} \right. \\ &\quad \left. + \frac{1}{2K^2\phi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \chi(v)^{-1} C_{q,\chi} + \sum_{j=1}^J \left(c_1(j) - \frac{c(j)}{q} \right) (\log H)^{1/2-j} + O\left((\log H)^{-J-1/2}\right) \right] \\ (24) \quad &= \frac{Kx}{q^2 \sqrt{\log x}} \left[1 + C_{a,b} + \frac{1}{\sqrt{\log x}} \sum_{j=0}^J b_1(j) (\log H)^{1/2-j} + O\left((\log H)^{-J-1/2}\right) \right] \\ &= \frac{Kx}{q^2 \sqrt{\log x}} \left[1 + C_{a,b} - \frac{1}{\phi(q)} \frac{1}{\sqrt{\log x}} \sum_{j=0}^J b_0(j) (\log H)^{1/2-j} + O\left((\log H)^{-J-1/2}\right) \right], \end{aligned}$$

where $C_{a,b} = \frac{q}{2K\phi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \chi(v)^{-1} C_{q,\chi}$, $b_1(0) = 2/\pi$ and $b_1(j) = K(qc_1(j) - c(j))$ for $j \geq 1$. For the

last line, we used (3) which gives $b_1(j) = -\frac{b_0(j)}{\phi(q)}$, for $j \geq 0$.

To deduce Conjecture 1.1 and obtain the explicit expressions for the constants C_j for $0 \leq j \leq J$, from the above expressions (23) and (24), we approximate $(\log H)^{1/2-j}$ for $0 \leq j \leq J$, where H is given by (17). We illustrate the process below for $J = 1$.

Using the approximations

$$\begin{aligned} (\log H)^{1/2} &= \frac{1}{\sqrt{2}} \sqrt{\log \log x} - \frac{\log K}{\sqrt{2}} \frac{1}{\sqrt{\log \log x}} + O\left((\log \log x)^{-3/2}\right), \\ (\log H)^{-1/2} &= \frac{\sqrt{2}}{\sqrt{\log \log x}} + O\left((\log \log x)^{-3/2}\right), \end{aligned}$$

we obtain

$$(25) \quad \sum_{j=0}^1 b_0(j) (\log H)^{1/2-j} = -\frac{\sqrt{2}(q-1)}{\pi} \sqrt{\log \log x} + \left(\frac{\sqrt{2}(q-1) \log K}{\pi} + \sqrt{2}b_0(1) \right) (\log \log x)^{-1/2} + O\left((\log \log x)^{-3/2}\right)$$

Using the values of $c(1)$ and $c_0(1)$ given by (34) and (38), we have

$$b_0(1) = K(qc_0(1) - c(1)) = \frac{\phi(q)}{\pi} \left(\frac{\omega + \gamma}{2} + \frac{q}{\phi(q)} \log q \right),$$

where γ is the Euler-Mascheroni constant and ω is defined in Lemma 5.2. Replacing in (25) and then in (23), we get Conjecture 4.3 below which is the special case of Conjecture 1.1 for $J = 1$. The case $a \not\equiv b \pmod{q}$ follows from multiplying the corresponding term by $\frac{-1}{\phi(q)}$ in (24). The general case of Conjecture 1.1 follows similarly by using approximations for $(\log H)^{-1/2-j}$ as above for $1 \leq j \leq J$.

Conjecture 4.3. Fix $q \equiv 1 \pmod{4}$. Then,

$$N(x, q, (a, a)) \sim \frac{x}{q^2} \frac{K}{\sqrt{\log x}} \left(1 - \frac{\sqrt{2}\phi(q)}{\pi} \frac{(\log \log x)^{1/2}}{(\log x)^{1/2}} + \frac{C_1}{(\log x)^{1/2}(\log \log x)^{1/2}} \right)$$

up to an error term of $O\left(\frac{x}{\log x(\log \log x)^{3/2}}\right)$, and with

$$C_1 = \frac{\sqrt{2}\phi(q)}{\pi} \left(\log K + \frac{\omega + \gamma}{2} \right) + \frac{\sqrt{2}q \log q}{\pi},$$

where γ is the Euler-Mascheroni constant and ω is defined in Lemma 5.2.

For $a \not\equiv b \pmod{q}$,

$$N(x, q, (a, b)) = \frac{x}{q^2} \frac{K}{\sqrt{\log x}} \left(1 + \frac{\sqrt{2}}{\pi} \frac{\sqrt{\log \log x}}{\sqrt{\log x}} + \frac{C_{a,b}}{\sqrt{\log x}} - \frac{C_1}{\phi(q)(\log x)^{1/2}(\log \log x)^{1/2}} \right)$$

up to an error term of $O\left(\frac{x}{\log x(\log \log x)^{3/2}}\right)$, and with

$$C_{a,b} := \frac{1}{2K} \frac{q}{\phi(q)} \sum_{\chi \neq \chi_0} \bar{\chi}(b-a) C_{q,\chi}$$

where the sum is over the non-principal Dirichlet characters modulo q and $C_{q,\chi}$ is defined in (28).

4.4. Another formulation of Theorem 1.2. Finally, as we did in Theorem 2.1 for the counting function of the number of sums of two squares up to x , we now state a different version of Theorem 1.2 with a very good error term by using an integral form for the main term. We prove this Proposition in Section 7, and we use it for numerical testing in Section 9.

Proposition 4.4. Fix $\varepsilon > 0$. There exists $c > 0$ such that for $v \not\equiv 0 \pmod{q}$

$$S(q, v, H) = \frac{H}{q} + \frac{1}{2K^2\phi(q)} \sum_{\chi \neq \chi_0} \chi(v)^{-1} C_{q,\chi} + \frac{1}{2\pi K^2\phi(q)} \int_{1/2+\varepsilon}^1 \frac{F_{\chi_0}(\sigma) H^{\sigma-1}}{|\sigma-1|^{1/2}} d\sigma + O\left(\exp(-c\sqrt{\log H})\right),$$

and

$$S(q, 0, H) = \frac{H}{q} + \frac{1}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^{\sigma-1} + F(\sigma)H^{\sigma-1}(\log H - A_q(\sigma)/2)}{|\sigma - 1|^{1/2}} d\sigma + O\left(\exp(-c\sqrt{\log H})\right)$$

where $F(s) = \zeta(s-1)M(s-1)[(s-1)\zeta(s)]^{1/2}\Gamma(s)$ and $F_{\chi_0}(s) = \frac{1-q^{-(s-1)}}{s-1}F(s)$, with $M(s)$ as defined by (37) and $A_q(s) = \frac{1-q^{-(s-1)}}{s-1}$. Assuming the Riemann Hypothesis, we can replace the error terms above by $O(H^{-1/2+\varepsilon})$.

Observe that close to $\frac{1}{2}$ one has $F'(\sigma) \asymp (\sigma - \frac{1}{2})^{-\frac{5}{4}}$, so the error term in the formula for $S(q, 0, H)$ depends strongly on ε .

5. PROOF OF THEOREM 1.2

Proof. As in [Smi13], we define $a(h) = 2K^2\mathfrak{S}(\{0, h\})$. Then using (11), we see that $a(h)$ is a multiplicative function of h with

$$a(p^k) = \begin{cases} 1 & \text{for } p \equiv 1 \pmod{4}, \\ 2 - \frac{3}{2^k} & \text{for } p = 2, k \geq 1, \\ \frac{1 - p^{-(k+1)}}{1 - p^{-1}} & \text{for } p \equiv 3 \pmod{4}. \end{cases}$$

Using Mellin Inversion, we have

$$2K^2S(H) = \sum_{h \geq 1} a(h)e^{-h/H} = \frac{1}{2\pi i} \int_{(2)} D(s)H^s\Gamma(s)ds,$$

where $D(s) = \sum_{h \geq 1} a(h)h^{-s}$. Similarly, for χ a character modulo q ,

$$(26) \quad 2K^2S(H, \chi) := \sum_{h \geq 1} a(h)\chi(h)e^{-h/H} = \frac{1}{2\pi i} \int_{(2)} D_\chi(s)H^s\Gamma(s)ds,$$

where $D_\chi(s) = \sum_{h \geq 1} a(h)\chi(h)h^{-s}$.

In order to compute $S(H)$ and $S(H, \chi)$, we move the contour integral and pick up the contributions of the singularities of the integrand. So, first, we need to understand the analytic properties of the generating series $D(s)$ and $D_\chi(s)$. Using the formulas for $a(p^k)$ above, we have

$$D_\chi(s) = R_\chi(s)P_\chi(s)Q_\chi(s),$$

where

$$\begin{aligned} R_\chi(s) &= 1 + 2\left(\frac{\chi(2)2^{-s}}{1 - \chi(2)2^{-s}}\right) - 3\left(\frac{\chi(2)2^{-(s+1)}}{1 - \chi(2)2^{-(s+1)}}\right) \\ P_\chi(s) &= \prod_{p \equiv 1 \pmod{4}} (1 - \chi(p)p^{-s})^{-1} \\ Q_\chi(s) &= \prod_{p \equiv 3 \pmod{4}} (1 - \chi(p)p^{-s})^{-1}(1 - \chi(p)p^{-(s+1)})^{-1}. \end{aligned}$$

This can be rewritten as

$$D_\chi(s) = L(s, \chi)(1 - \chi(2)2^{-s})R_\chi(s)Q_{1, \chi}(s) = L(s, \chi)L(s+1, \chi)^{\frac{1}{2}}M_\chi(s)$$

where

$$\begin{aligned}
Q_{1,\chi}(s) &= \prod_{p \equiv 3 \pmod{4}} (1 - \chi(p)p^{-(s+1)})^{-1} \\
&= \left(\frac{L(s+1, \chi)}{L(s+1, \chi \cdot \chi_4)} \right)^{\frac{1}{2}} (1 - \chi(2)2^{-(s+1)})^{\frac{1}{2}} \prod_{p \equiv 3 \pmod{4}} (1 - \chi(p)^2 p^{-2(s+1)})^{-\frac{1}{2}}, \\
M_\chi(s) &= \left(1 + \chi(2)2^{-s} - 3 \left(\frac{\chi(2)2^{-(s+1)}(1 - \chi(2)2^{-s})}{1 - \chi(2)2^{-(s+1)}} \right) \right) L(s+1, \chi \cdot \chi_4)^{-\frac{1}{2}} \\
&\quad \times (1 - \chi(2)2^{-(s+1)})^{\frac{1}{2}} \prod_{p \equiv 3 \pmod{4}} (1 - \chi(p)^2 p^{-2(s+1)})^{-\frac{1}{2}} \\
&= (1 - \chi(2)2^{-s} + \chi(4)2^{-2s}) L(s+1, \chi \cdot \chi_4)^{-\frac{1}{2}} \\
&\quad \times (1 - \chi(2)2^{-(s+1)})^{-\frac{1}{2}} \prod_{p \equiv 3 \pmod{4}} (1 - \chi(p)^2 p^{-2(s+1)})^{-\frac{1}{2}},
\end{aligned}$$

where χ_4 is the primitive character modulo 4. The formula for $Q_{1,\chi}(s)$ follows from developing the identity $(1 - \chi(p)^2 p^{-2(s+1)}) = (1 - \chi(p)p^{-(s+1)})(1 + \chi(p)p^{-(s+1)})$. Since $\chi \neq \chi_4$, the function M_χ is holomorphic in the half plane $\operatorname{Re}(s) \geq 0$, and we can push this limit a bit further to the left depending on the zero-free region of $L(s+1, \chi \cdot \chi_4)$ (up to $\operatorname{Re}(s) > -\frac{1}{2}$ under Riemann Hypothesis).

In the case χ is a non-principal character, $L(s, \chi)$ is entire on the complex plane, and $L(s+1, \chi)^{\frac{1}{2}}$ is holomorphic in a region containing the half-plane $\operatorname{Re}(s) \geq 0$, where $L(s+1, \chi)$ does not vanish. As such, there is no pole or singularity in the integrand at $s = 1$. If we shift the line of integration of (26) to the left of the line $\operatorname{Re}(s) = 0$ using the standard zero-free region and estimates for L -functions, we obtain (for some $c > 0$ and any $\varepsilon > 0$)

$$(27) \quad S(H, \chi) = \frac{C_{q,\chi}}{2K^2} + \begin{cases} O(H^{-1/2+\varepsilon}) & \text{under GRH} \\ O(\exp(-c\sqrt{\log H})) & \text{otherwise,} \end{cases}$$

where the constant term comes from the contribution of pole of order 1 from $\Gamma(s)$ at $s = 0$, and

$$(28) \quad \begin{aligned} C_{q,\chi} &= D_\chi(0) = L(0, \chi)L(1, \chi)^{\frac{1}{2}}M_\chi(0) \\ &= L(0, \chi)L(1, \chi)^{\frac{1}{2}}L(1, \chi \cdot \chi_4)^{-\frac{1}{2}}(1 - \chi(2) + \chi(4))(1 - \chi(2)2^{-1})^{-\frac{1}{2}} \prod_{p \equiv 3 \pmod{4}} (1 - \chi(p)^2 p^{-2})^{-\frac{1}{2}}. \end{aligned}$$

Note that $C_{q,\chi} \neq 0$ only when $\chi(-1) = -1$.

If $\chi = \chi_0$, $D_{\chi_0}(s)$ has a simple pole at $s = 1$ with residue $2K^2\phi(q)/q$, and no other singularities for $\operatorname{Re}(s) > 0$, and we move the integral to $\operatorname{Re}(s) = \varepsilon > 0$. This gives

$$(29) \quad \begin{aligned} \sum_{h \geq 1} a(h)\chi_0(q)e^{-h/H} &= \frac{1}{2\pi i} \int_{(2)} D_{\chi_0}(s)H^s\Gamma(s)ds \\ &= 2\frac{\phi(q)}{q}K^2H + \frac{1}{2\pi i} \int_{(\varepsilon)} D_{\chi_0}(s)H^s\Gamma(s)ds. \end{aligned}$$

Similarly, we have that

$$(30) \quad \sum_{h \geq 1} a(h)e^{-h/H} = \frac{1}{2\pi i} \int_{(2)} D(s)H^s\Gamma(s)ds = 2K^2H + \frac{1}{2\pi i} \int_{(\varepsilon)} D(s)H^s\Gamma(s)ds,$$

where

$$D(s) = \zeta(s)(1 - 2^{-s})R(s)Q_1(s) = \zeta(s)\zeta(s+1)^{\frac{1}{2}}M(s),$$

and the functions R, Q_1, M are obtained by taking $\chi \equiv 1$ in the previous definitions.

To account for the contribution of the singularity of $D_{\chi_0}(s)$ and $D(s)$ as $s = 0$ to the integrals (29) and (30), we use again the Selberg–Delange method. Since we are evaluating a Mellin transform, we

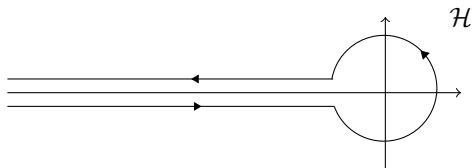


FIGURE 1. Hankel's Contour

cannot use directly [Kou19, Theorem 13.2] as in Section 2, but we are following the same standard steps. We first approximate the line of integration $\Re(s) = \varepsilon$ by the truncated segment from $\varepsilon - iT$ to $\varepsilon + iT$, which we then deform to a truncated Hankel's contour. This is possible since there are no residue inside this contour. We then replace this contour by the infinite Hankel's contour \mathcal{H} of Figure 1 with a very good error term, which allows us to use Theorem 5.1 to compute the contribution of the singularity of the generating functions (for each term of the Taylor series). We refer the reader to [Kou19, Chapter 13] and [Ten15, Chapter 5] for more details. The contributions to $S(H)$ and $S(H, \chi_0)$ will be different in magnitude, because the singularities of $\zeta(s)\zeta(s+1)^{1/2}$ and $L(s, \chi_0)L(s+1, \chi_0)^{1/2}$ at $s = 0$ are different, since $L(0, \chi_0) = 0$, but $\zeta(0) \neq 0$.

Theorem 5.1 (Hankel's formula [Ten15] Theorem 0.17 p.179). *Fix any $r > 0$, and let \mathcal{H} be the Hankel's contour, which is the path consisting of the circle $|s| = r$ excluding the point $s = -r$, and of the half-line $(-\infty, -r]$ covered twice, with respective arguments π and $-\pi$. Then for any complex number z , we have*

$$\frac{1}{2\pi i} \int_{\mathcal{H}} s^{-z} e^s ds = \frac{1}{\Gamma(z)}.$$

We first work with $D(s) = \zeta(s)\zeta(s+1)^{\frac{1}{2}}M(s)$. The function $M(s)$ is analytic around $s = 0$, and

$$M(0) = 2KL(1, \chi_4)^{-\frac{1}{2}} = \frac{4K}{\sqrt{\pi}}.$$

Then, $s^{\frac{3}{2}}D(s)\Gamma(s)$ is analytic and non-zero around $s = 0$ with Taylor series $\sum_{n \geq 0} c_n s^n$, and we write

$$D(s)\Gamma(s) = \frac{a(3/2)}{s^{3/2}} + \frac{a(1/2)}{s^{1/2}} + a(-1/2)s^{1/2} + \dots$$

We now compute the contribution to the integral (30) for each term of the series above using Theorem 5.1. For every term $a(z)/s^z$ of the Taylor series above (where $z = 3/2, 1/2, -1/2, \dots$), we have

$$\begin{aligned} \frac{1}{2\pi i} \int_{\mathcal{H}} a(z)s^{-z}H^s ds &= \frac{a(z)}{2\pi i} \int_{\mathcal{H}} s^{-z}e^{s \log H} ds \\ &= \frac{a(z)(\log H)^{z-1}}{2\pi i} \int_{\mathcal{H}} t^{-z}e^t dt = \frac{a(z)(\log H)^{z-1}}{\Gamma(z)}, \end{aligned}$$

where we used the change of variables $t = s \log H$. This gives, for any integer $N \geq 1$,

$$\frac{1}{2\pi i} \int_{(\varepsilon)} D(s)H^s \Gamma(s) ds = \sum_{n=0}^N \frac{a(3/2-n)(\log H)^{1/2-n}}{\Gamma(3/2-n)} + O\left((\log H)^{1/2-N-1}\right).$$

Replacing in (30), this gives

$$(31) \quad S(H) = H + \sum_{j=0}^J c(j)(\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right),$$

with

$$c(j) = \frac{a(3/2-j)}{2K^2\Gamma(3/2-j)}, \quad j \geq 0.$$

To complete the proof of Theorem 1.2, we now compute the values of $c(0)$ and $c(1)$. Using the expansions² around $s = 0$

$$\sqrt{\zeta(s+1)} = \frac{1}{s^{1/2}} \left(1 + \frac{\gamma}{2}s + O(s^2) \right), \quad \Gamma(s) = \frac{1}{s} - \gamma + O(s),$$

we have

$$(32) \quad D(s)\Gamma(s) = \frac{1}{s^{3/2}} \zeta(s)M(s) \left(1 - \frac{\gamma}{2}s + O(s^2) \right),$$

which gives

$$a(3/2) = \zeta(0)M(0) = -\frac{2K}{\sqrt{\pi}} \implies c(0) = \frac{a(3/2)}{2K^2\Gamma(3/2)} = -\frac{2}{K\pi}.$$

To get the value of $c(1)$, we need the first 2 terms of the Taylor series around $s = 0$ of the analytic function

$$(33) \quad Z(s) = \zeta(s)M(s) = Z(0) + Z'(0)s + O(s^2).$$

Replacing in (32), and using Lemma 5.2 for the value of $Z'(0)$, we have

$$(34) \quad a(1/2) = \left(Z'(0) - \frac{\gamma}{2}Z(0) \right) = \frac{K(\omega + \gamma)}{\sqrt{\pi}} \implies c(1) = \frac{a(1/2)}{2K^2\Gamma(1/2)} = \frac{\omega + \gamma}{2\pi K},$$

where γ is the Euler-Mascheroni constant and ω is defined in Lemma 5.2.

We now turn to the secondary term for the sum

$$(35) \quad \begin{aligned} \sum_{h \geq 1} a(h)\chi_0(h)e^{-h/H} &= \frac{1}{2\pi i} \int_{(2)} D_{\chi_0}(s)H^s\Gamma(s)ds \\ &= 2\frac{\phi(q)}{q}K^2H + \frac{1}{2\pi i} \int_{(\varepsilon)} D_{\chi_0}(s)H^s\Gamma(s)ds \end{aligned}$$

which is similar to the above replacing $D(s)$ with D_{χ_0} , where χ_0 is the principal character modulo q . We have

$$D_{\chi_0}(s) = L(s, \chi_0)L(s+1, \chi_0)^{\frac{1}{2}}M_{\chi_0}(s),$$

where

$$M_{\chi_0}(s) = (1 - 2^{-s} + 2^{-2s})L(s+1, \chi_0 \cdot \chi_4)^{-\frac{1}{2}}(1 - 2^{-(s+1)})^{-\frac{1}{2}} \prod_{p \equiv 3 \pmod{4}} (1 - p^{-2(s+1)})^{-\frac{1}{2}}$$

since $\chi_0(p) = 1$ for each $p \nmid q$ and $q \equiv 1 \pmod{4}$. We remark that $M_{\chi_0}(s) = (1 - q^{-(s+1)})^{-\frac{1}{2}}M(s)$, which implies that

$$D_{\chi_0}(s) = L(s, \chi_0)\zeta(s+1)^{\frac{1}{2}}M(s) = (1 - q^{-s})D(s).$$

Then writing $(1 - q^{-s}) = (\log q)s + O(s^2)$, we notice that $s^{1/2}D_{\chi_0}(s)\Gamma(s)$ is analytic and non-zero around $s = 0$. Indeed, $L(s, \chi_0)$ has a simple zero at $s = 0$ which cancels the pole of $\Gamma(s)$. Around $s = 0$, we write

$$D_{\chi_0}(s)\Gamma(s) = \frac{b(1/2)}{s^{1/2}} + b(-1/2)s^{1/2} + b(-3/2)s^{3/2} + \dots,$$

and working as above this gives

$$\frac{1}{2\pi i} \int_{(\varepsilon)} D_{\chi_0}(s)H^s\Gamma(s)ds = \sum_{n=0}^N \frac{b(1/2-n)(\log H)^{-1/2-n}}{\Gamma(1/2-n)} + O\left((\log H)^{-1/2-N-1}\right),$$

²We used the Laurent expansion of $\zeta(s)$ at $s = 1$

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=0}^{\infty} (-1)^n \frac{\gamma_n}{n!} (s-1)^n, \quad \text{with } \gamma_n := \lim_{N \rightarrow \infty} \left(\sum_{1 \leq k \leq N} \frac{\log^n k}{k} - \int_1^N \frac{\log^n t}{t} dt \right), \quad \text{and } \gamma := \gamma_0.$$

replacing in (35), we have

$$(36) \quad S(H, \chi_0) = \frac{\phi(q)}{q} H + \sum_{j=1}^J c(j, \chi_0) (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right).$$

Using the expansion of $D(s)\Gamma(s)$ above, we have

$$b(1/2) = a(3/2) \log q = -\frac{2K}{\sqrt{\pi}} \log q \implies c(1, \chi_0) = \frac{b(1/2)}{2K^2\Gamma(1/2)} = -\frac{1}{K\pi} \log q.$$

We now complete the proof of Theorem 1.2. Using (27) and (36) and the orthogonality relations, we have for $v \neq 0$,

$$\begin{aligned} S(q, v, H) &= \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} \mathfrak{S}(\{0, h\}) e^{-h/H} = \frac{1}{\phi(q)} \sum_{\chi} \chi(v)^{-1} S(H, \chi) \\ &= \frac{1}{\phi(q)} S(H, \chi_0) + \frac{1}{2K^2\phi(q)} \sum_{\chi \neq \chi_0} \chi(v)^{-1} C_{q, \chi} + O(\exp(-c\sqrt{\log H})) \\ &= \frac{H}{q} + \frac{1}{2K^2\phi(q)} \sum_{\chi \neq \chi_0} \chi(v)^{-1} C_{q, \chi} + \sum_{j=1}^J \frac{c(j, \chi_0)}{\phi(q)} (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right). \end{aligned}$$

For $v = 0$, we use (31) and the above to get

$$\begin{aligned} S(q, 0, H) &= S(H) - \sum_{v \in (\mathbb{Z}/q\mathbb{Z})^*} S(q, v, H) \\ &= \frac{H}{q} - \frac{2}{K\pi} \sqrt{\log H} + \sum_{j=1}^J (c(j) - c(j, \chi_0)) (\log H)^{1/2-j} + O\left((\log H)^{-1/2-J}\right), \end{aligned}$$

where we used the fact that

$$(37) \quad \sum_{v \in (\mathbb{Z}/q\mathbb{Z})^*} \sum_{\chi \neq \chi_0} \chi(v)^{-1} C_{q, \chi} = 0$$

by the orthogonality relations. This completes the proof of the proposition, with $c_1(j) = c(j, \chi_0)/\phi(q)$ and $c_0(j) = c(j) - c(j, \chi_0)$ for $j \geq 1$, from which the relation $c_0(j) + \phi(q)c_1(j) = c(j)$ easily follows. From the values $c(1)$ and $c(1, \chi_0)$ computed above, we have

$$(38) \quad c_1(1) = -\frac{\log q}{K\phi(q)\pi} \quad \text{and} \quad c_0(1) = \frac{1}{K\pi} \left(\frac{\omega + \gamma}{2} + \log q \right),$$

where γ is the Euler-Mascheroni constant and ω is defined in Lemma 5.2. □

Lemma 5.2. *Let $Z(s)$ be the function defined by (33). Then,*

$$Z'(0) = \frac{K}{\sqrt{\pi}} \omega \approx -0.3851314513 \dots$$

where

$$\omega = \log \frac{2}{\pi^2} + \frac{L'(1, \chi_4)}{L(1, \chi_4)} + 2 \sum_{p \equiv 3(4)} \frac{\log p}{p^2 - 1}.$$

Proof. Firstly, observe $M(s)$ in (33) and rewrite it as

$$Z(s) = \zeta(s)M(s) = \zeta(s)A(s)B(s)$$

where

$$\begin{aligned} A(s) &= 1 - 2^{-s} + 2^{-2s}, \\ B(s) &= \left(L(s+1, \chi_4) (1 - 2^{-(s+1)}) \prod_{p \equiv 3(4)} \left(1 - \frac{1}{p^{2(s+1)}} \right) \right)^{-1/2}. \end{aligned}$$

Then, we have

$$\frac{Z'}{Z}(0) = \frac{\zeta'}{\zeta}(0) + \frac{A'}{A}(0) + \frac{B'}{B}(0),$$

Hence, we need to compute ζ, A, A', B and B' at $s = 0$. Indeed, the following special values for ζ are well-known:

$$\zeta(0) = -\frac{1}{2}, \quad \zeta'(0) = -\frac{\log 2\pi}{2}.$$

Moreover, for $A(s)$, we have $A(0) = 1, A'(0) = -\log 2$. We may use the recursive formula for $B(s)$ to obtain $B(0) = 4K/\sqrt{\pi}$ and

$$\frac{B'}{B}(0) = -\frac{1}{2} \left(\frac{L'(1, \chi_4)}{L(1, \chi_4)} + \log 2 + 2 \sum_{p \equiv 3(4)} \frac{\log p}{p^2 - 1} \right) = -\frac{1}{2} (\log 2 + \alpha_1 + \beta_1)$$

where we denote

$$\alpha_1 = \frac{L'(1, \chi_4)}{L(1, \chi_4)} = 0.2456096036\dots, \quad \beta_1 = 2 \sum_{p \equiv 3(4)} \frac{\log p}{p^2 - 1} = 0.4574727064\dots$$

One can compute the value of α_1 by $L(1, \chi_4) = \pi/4$ and $L'(1, \chi_4) = 0.192901331574902\dots$ \square

6. PROOF OF THEOREM 3.4

In the heuristic leading to Conjecture 1.1, we used Theorem 3.4 to justify that the terms involving a sum of singular series for sets with three or more elements contribute to the error term. Theorem 3.4 is an analogue of [MS04, Theorem 2] of Montgomery and Soundararajan adapted from primes to sums of two squares. We now prove Theorem 3.4, following closely the argument developed in [MS04], without giving all the details but insisting on the points that are different in the case of the sums of two squares. To help with the comparison, we stay close to the notation used in loc. cit. so we may use notation that differs from the rest of the paper, which should not cause trouble to the reader as this section is relatively independent from the rest of the paper. We use the standard notation $e(x) = e^{2i\pi x}$.

6.1. The singular series. The first step in the proof is to write the singular series as an actual series (and not a Euler product), the way it was introduced by Hardy and Littlewood (see [MS04, Lemma 3]). We begin with giving a new expression for the local factors of the singular series. Let $\mathcal{D} = \{d_1, \dots, d_k\} \subseteq \mathbb{Z}$. We recall that for any $p \not\equiv 1 \pmod{4}$, we have

$$\delta_{\mathcal{D}}(p) = \lim_{\alpha \rightarrow \infty} \frac{\#\{0 \leq a < p^\alpha : \forall d \in \mathcal{D}, a + d \equiv \square + \square \pmod{p^\alpha}\}}{p^\alpha},$$

and the singular series is defined by

$$\mathfrak{S}(\mathcal{D}) := \prod_{p \not\equiv 1 \pmod{4}} \frac{\delta_{\mathcal{D}}(p)}{(\delta_{\{0\}}(p))^k}.$$

Lemma 6.1. *Let $\mathcal{D} = \{d_1, \dots, d_k\} \subseteq \mathbb{Z}$ be a set with k elements. For any prime number $p \not\equiv 1 \pmod{4}$, one has*

$$\frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} = \sum_{q_1, \dots, q_k | p^\infty} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k),$$

where for any $q_1, \dots, q_k \in \mathbb{N}$,

$$A_{\mathcal{D}}(q_1, \dots, q_k) = \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} e \left(\sum_{i=1}^k \frac{a_i d_i}{q_i} \right) \prod_{i=1}^k C(q_i, a_i),$$

with

$$C(q, a) = \begin{cases} 1 & \text{if } q \text{ is odd,} \\ 0 & \text{if } 2 \mid q \text{ but } 4 \nmid q, \\ 2e(-a/4) & \text{if } 4 \mid q. \end{cases}$$

and λ_2 is the multiplicative function defined on the prime powers by

$$\lambda_2(p^m) = \begin{cases} (-1)^m & \text{if } p \text{ is odd,} \\ 1 & \text{if } p = 2. \end{cases}$$

Proof. Let $p \equiv 3 \pmod{4}$ be a prime number. For $\mathcal{D} = \{d_1, \dots, d_k\} \subseteq \mathbb{Z}$ a set with k elements, we deduce from [FKR17, Proposition 5.1, Proposition 5.3(a) and (5.4)] that

$$\delta_{\mathcal{D}}(p) = \lim_{\alpha \rightarrow \infty} p^{-\alpha} \sum_{x=1}^{p^\alpha} \prod_{i=1}^k \mathbf{1}_{S_{p,\alpha}}(x + d_i)$$

where $\mathbf{1}_{S_{p,\alpha}}$ is the characteristic function of the set $S_{p,\alpha} = \{p^{2\beta}m : 0 \leq \beta < \frac{\alpha}{2}, m \not\equiv 0 \pmod{p}\}$. In particular, for α even, following the idea of the proof of [MV86, Lemma 2], we write that

$$\begin{aligned} \mathbf{1}_{S_{p,\alpha}}(x) &= \sum_{\beta=0}^{\frac{\alpha}{2}-1} \sum_{s|p} \frac{\mu(s)}{p^{2\beta}s} \sum_{a=1}^{p^{2\beta}s} e\left(\frac{ax}{p^{2\beta}s}\right) \\ &= \sum_{\beta=0}^{\frac{\alpha}{2}-1} \frac{1}{p^{2\beta}} \left\{ \left(1 - \frac{1}{p}\right) \sum_{r|p^{2\beta}} \sum_{a \in (\mathbb{Z}/r\mathbb{Z})^*} e\left(\frac{ax}{r}\right) - \frac{1}{p} \sum_{a \in (\mathbb{Z}/p^{2\beta+1}\mathbb{Z})^*} e\left(\frac{ax}{p^{2\beta+1}}\right) \right\} \\ &= \sum_{\gamma=0}^{\alpha-2} \left\{ \sum_{\beta=\lceil \frac{\gamma}{2} \rceil}^{\frac{\alpha}{2}-1} \frac{1}{p^{2\beta}} \left(1 - \frac{1}{p}\right) \sum_{a \in (\mathbb{Z}/p^\gamma\mathbb{Z})^*} e\left(\frac{ax}{p^\gamma}\right) \right\} - \sum_{\beta=0}^{\frac{\alpha}{2}-1} \frac{1}{p^{2\beta+1}} \sum_{a \in (\mathbb{Z}/p^{2\beta+1}\mathbb{Z})^*} e\left(\frac{ax}{p^{2\beta+1}}\right) \\ &= \sum_{\gamma=0}^{\alpha-1} \frac{(-p)^{-\gamma} - p^{-\alpha}}{1 + \frac{1}{p}} \sum_{a \in (\mathbb{Z}/p^\gamma\mathbb{Z})^*} e\left(\frac{ax}{p^\gamma}\right), \end{aligned}$$

where we used the fact that α is even in the last line. Thus, we have,

$$\frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} = \lim_{\substack{\alpha \rightarrow \infty \\ \alpha \text{ even}}} p^{-\alpha} \sum_{x=1}^{p^\alpha} \prod_{i=1}^k \left(\sum_{\gamma_i=0}^{\alpha-1} ((-p)^{-\gamma_i} - p^{-\alpha}) \sum_{a_i \in (\mathbb{Z}/p^{\gamma_i}\mathbb{Z})^*} e\left(\frac{(x+d_i)a_i}{p^{\gamma_i}}\right) \right),$$

since $\delta_{\{0\}}(p) = (1 + \frac{1}{p})^{-1}$ [FKR17, Proposition 5.3(c)]. We swap the sums and begin with the sum over x , as $\gamma_1, \dots, \gamma_k \leq \alpha - 1$, we have

$$p^{-\alpha} \sum_{x=1}^{p^\alpha} e\left(x \sum_{i=1}^k \frac{a_i}{p^{\gamma_i}}\right) = \begin{cases} 1 & \text{if } \sum_{i=1}^k \frac{a_i}{p^{\gamma_i}} \in \mathbb{Z} \\ 0 & \text{otherwise.} \end{cases}$$

This yields

$$\frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} = \lim_{\substack{\alpha \rightarrow \infty \\ \alpha \text{ even}}} \sum_{\gamma_1=0}^{\alpha-1} \cdots \sum_{\gamma_k=0}^{\alpha-1} \prod_{i=1}^k ((-p)^{-\gamma_i} - p^{-\alpha}) A_{\mathcal{D}}(p^{\gamma_1}, \dots, p^{\gamma_k}).$$

We obtain the formula announced in the Lemma for $p \equiv 3 \pmod{4}$ by taking the limit $\alpha \rightarrow \infty$, and using the bound $|A_{\mathcal{D}}(q_1, \dots, q_k)| \leq \frac{q_1 \cdots q_k}{[q_1, \dots, q_k]}$ (see (46)).

The proof is similar for $p = 2$. By [FKR17, Proposition 5.2(a) and (5.3)], for $\alpha \geq 2$ we can take $S_{2,\alpha} = \{2^\beta m : 0 \leq \beta < \alpha - 1, m \equiv 1 \pmod{4}\}$, and [FKR17, Proposition 5.2(c)] gives $\delta_{\{0\}}(2) = \frac{1}{2}$.

We write that

$$\begin{aligned}
\mathbf{1}_{S_{2,\alpha}}(x) &= \sum_{\beta=0}^{\alpha-2} 2^{-\beta-2} \sum_{a=1}^{2^\beta} e\left(\frac{ax}{2^\beta}\right) \sum_{t=1}^4 e\left(\left(\frac{x}{2^\beta} - 1\right)\frac{t}{4}\right) \\
&= \sum_{\beta=0}^{\alpha-2} 2^{-\beta-2} \sum_{r|2^{\beta+2}} \sum_{b \in (\mathbb{Z}/r\mathbb{Z})^*} e\left(\frac{(x-2^\beta)b}{r}\right) \\
&= \sum_{\gamma=0}^{\alpha} \sum_{b \in (\mathbb{Z}/2^\gamma\mathbb{Z})^*} e\left(\frac{xb}{2^\gamma}\right) \sum_{\beta=\max\{0,\gamma-2\}}^{\alpha-2} 2^{-\beta-2} e(-2^{\beta-\gamma}b) \\
&= \sum_{\gamma=0}^{\alpha} \sum_{b \in (\mathbb{Z}/2^\gamma\mathbb{Z})^*} e\left(\frac{xb}{2^\gamma}\right) (C(2^\gamma, b)2^{-\gamma-1} - 2^{-\alpha}),
\end{aligned}$$

note that in the sum we always have $(b, 2) = 1$ so $1 + e(\frac{-b}{2}) = 0$. Thus, we have

$$\begin{aligned}
\frac{\delta_{\mathcal{D}}(2)}{\delta_{\{0\}}(2)^k} &= \lim_{\alpha \rightarrow \infty} 2^{-\alpha+k} \sum_{x=1}^{2^\alpha} \prod_{i=1}^k \mathbf{1}_{S_{2,\alpha}}(x + d_i) \\
&= \lim_{\alpha \rightarrow \infty} 2^{-\alpha} \sum_{x=1}^{2^\alpha} \prod_{i=1}^k \left(\sum_{\gamma_i=0}^{\alpha} \sum_{b_i \in (\mathbb{Z}/2^{\gamma_i}\mathbb{Z})^*} e\left(\frac{(x+d_i)b_i}{2^{\gamma_i}}\right) (C(2^{\gamma_i}, b_i)2^{-\gamma_i} - 2^{-\alpha+1}) \right)
\end{aligned}$$

Exchanging the sums and computing the sum over x first, this yields

$$\frac{\delta_{\mathcal{D}}(2)}{\delta_{\{0\}}(2)^k} = \lim_{\alpha \rightarrow \infty} \sum_{\gamma_1=0}^{\alpha} \cdots \sum_{\gamma_k=0}^{\alpha} \prod_{i=1}^k 2^{-\gamma_i} A_{\mathcal{D}}(2^{\gamma_1}, \dots, 2^{\gamma_k}),$$

which gives the formula announced in the Lemma for $p = 2$. \square

We now give the analogue of [MS04, (44) and (45)]. The main difference between the case of primes and the case of sum of two squares is that the local probabilities $\delta_{\mathcal{D}}(p)$ at each prime p involve all powers of p , and then the sum over q_1, \dots, q_k in Lemma 6.1 runs over all integers (and not only square-free integers). We then approximate $\mathfrak{S}(\mathcal{D})$ by taking all integers supported on primes $p \leq y$ and appearing with power at most N , for the appropriate values of y and N .

Lemma 6.2. *Let $\mathcal{D} \subseteq \mathbb{N} \cap [1, h]$ be a set with k elements. Let $y > h$, $N \geq 4 \log y$, and $P_y := \prod_{\substack{p \leq y \\ p \not\equiv 1 \pmod{4}}} p$.*

Then,

$$(39) \quad \mathfrak{S}(\mathcal{D}) = \sum_{q_1, \dots, q_k | P_y^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_k(y^{-1}(\log y)^{k-1})$$

$$(40) \quad \text{and} \quad \mathfrak{S}_0(\mathcal{D}) = \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_k(y^{-1}(\log y)^{k-1}),$$

where $A_{\mathcal{D}}(q_1, \dots, q_k)$ is defined in Lemma 6.1.

Proof. First, it follows from the Chinese Remainder Theorem that for $q_1, \dots, q_k, q'_1, \dots, q'_k \in \mathbb{N}$ satisfying $(\prod_{i=1}^k q_i, \prod_{i=1}^k q'_i) = 1$, one has $A_{\mathcal{D}}(q_1, \dots, q_k) A_{\mathcal{D}}(q'_1, \dots, q'_k) = A_{\mathcal{D}}(q_1 q'_1, \dots, q_k q'_k)$. Since $y > h \geq \max \mathcal{D}$, from [FKR17, Proposition 5.3.(c)] we deduce

$$\prod_{p > y} \frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} = \prod_{p > y} (1 + \frac{1}{p})^{k-1} (1 - \frac{k-1}{p}) = \prod_{p > y} (1 + O_k(p^{-2})) = 1 + O_k((y \log y)^{-1}).$$

By definition we have $\delta_{\mathcal{D}}(p) \leq 1$ for all prime number p , thus

$$\prod_{p \leq y} \frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} \leq 2^k \prod_{\substack{p \leq y \\ p \equiv 3 \pmod{4}}} \left(1 + \frac{1}{p}\right)^k \ll_k (\log y)^k,$$

which gives

$$\mathfrak{S}(\mathcal{D}) = \prod_{p \leq y} \frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} + O(y^{-1}(\log y)^{k-1}).$$

Using Lemma 6.1 and the bound $|A_{\mathcal{D}}(q_1, \dots, q_k)| \leq 2^{\frac{k}{2}} \frac{q_1 \dots q_k}{[q_1, \dots, q_k]}$ (see (46)), we have

$$\begin{aligned} \frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} &= \sum_{q_1, \dots, q_k | p^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_k \left(\sum_{p^{N+1} | q_1 | p^\infty} \sum_{q_2, \dots, q_k | p^\infty} \frac{1}{[q_1, \dots, q_k]} \right) \\ &= \sum_{q_1, \dots, q_k | p^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_k \left(\sum_{n=N+1}^{\infty} \frac{(n+1)^{k-1}}{p^n} \right) \\ &= \sum_{q_1, \dots, q_k | p^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_k(p^{-N-1}(N+2)^{k-1}). \end{aligned}$$

Moreover, using again the bound (46), we have

$$\begin{aligned} \prod_{p \leq y} \sum_{q_1, \dots, q_k | p^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) &\ll_k \sum_{q_1, \dots, q_k | P_y^N} \prod_{i=1}^k \frac{|\lambda_2(q_i)|}{q_i} \frac{q_1 \dots q_k}{[q_1, \dots, q_k]} \\ &\leq \sum_{q_1, \dots, q_k | P_y^N} \frac{1}{[q_1, \dots, q_k]} \\ (41) \quad &\ll_k \prod_{p \leq y} \sum_{n=0}^N \frac{(n+1)^k}{p^n} \leq \prod_{p \leq y} \left(1 + \frac{C_k}{p}\right) \ll_\varepsilon y^\varepsilon \end{aligned}$$

for some constant $C_k > 0$, for any $\varepsilon > 0$. Finally,

$$\begin{aligned} \prod_{p \leq y} \frac{\delta_{\mathcal{D}}(p)}{\delta_{\{0\}}(p)^k} &= \sum_{q_1, \dots, q_k | P_y^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_{k,\varepsilon} \left(y^\varepsilon \sum_{\substack{q | P_y \\ q \neq 1}} q^{-N-1} N^{(k-1)\omega(q)} \right) \\ &= \sum_{q_1, \dots, q_k | P_y^N} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} A_{\mathcal{D}}(q_1, \dots, q_k) + O_{k,\varepsilon} \left(y^\varepsilon 2^{-N} N^{k-1} \right). \end{aligned}$$

Choosing $\frac{\log y}{\log 2}(1 + \varepsilon) < N$ gives (39). We deduce (40) from (39) using the formula

$$\mathfrak{S}_0(\mathcal{D}) = \sum_{\mathcal{T} \subseteq \mathcal{D}} (-1)^{|\mathcal{D} \setminus \mathcal{T}|} \mathfrak{S}(\mathcal{T})$$

and the relation $A_{\{d_1, \dots, d_k\}}(q_1, \dots, q_{k-1}, 1) = A_{\{d_1, \dots, d_{k-1}\}}(q_1, \dots, q_{k-1})$. \square

In particular, taking $y = h^{k+1}$ in (40), one has

$$(42) \quad \sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}|=k}} \mathfrak{S}_0(\mathcal{D}) = \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} \sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}|=k}} A_{\mathcal{D}}(q_1, \dots, q_k) + o_k(1).$$

6.2. An easier version of the main term. To continue with notation similar to [MS04], we define

$$(43) \quad V_k(y, N, h) = \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} \sum_{1 \leq d_1, \dots, d_k \leq h} A_{(d_1, \dots, d_k)}(q_1, \dots, q_k),$$

where we remark that the difference with the main term above is that d_1, \dots, d_k do not have to be distinct. Let us introduce some other useful notations and results from [MS04]. For $\alpha \in \mathbb{R}$, we denote

$$(44) \quad E_h(\alpha) = \sum_{d=1}^h e(\alpha d) \quad \text{and} \quad F_h(\alpha) = \min(h, \|\alpha\|^{-1}),$$

where $\|\cdot\|$ is the distance to the nearest integer, so that we have $|E_h(\alpha)| \leq F_h(\alpha)$. We have (see [MS04, (54)])

$$(45) \quad \sum_{a=1}^{q-1} F_h\left(\frac{a}{q}\right)^2 \ll q \min(q, h).$$

We will also use the following result from the work of Montgomery and Vaughan [MV84] and which is an analogue of [MS04, Lemma 1] that applies to the case of non necessarily square-free numbers.

Lemma 6.3 (Theorem 1 of [MV84]). *Let $k \geq 2$ be an integer and for $1 \leq i \leq k$, let $q_i \in \mathbb{N}$ and G_i be a 1-periodic complex valued function. Then, we have*

$$\left| \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=1}^k G_i\left(\frac{a_i}{q_i}\right) \right| \leq \frac{1}{[q_1, \dots, q_k]} \prod_{i=1}^k (q_i \sum_{\substack{1 \leq a_i \leq q_i \\ (q_i, a_i) = 1}} |G_i\left(\frac{a_i}{q_i}\right)|^2)^{\frac{1}{2}}.$$

In particular we deduce the bound for $A_{\mathcal{D}}(q_1, \dots, q_k)$ that we used in the proofs of Lemma 6.1 and 6.2:

$$(46) \quad |A_{\mathcal{D}}(q_1, \dots, q_k)| \leq \frac{1}{[q_1, \dots, q_k]} \prod_{i=1}^k (q_i \sum_{\substack{1 \leq a_i \leq q_i \\ (q_i, a_i) = 1}} |C(q_i, a_i)|^2)^{\frac{1}{2}}.$$

We also have a bound for $V_k(y, N, h)$.

Corollary 6.4. *For any $h, y, N > 0$, one has $V_k(y, N, h) \ll_{k, \varepsilon} h^{\frac{k}{2}} y^\varepsilon$.*

Proof. Recall that we defined

$$V_k(y, N, h) = \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \prod_{i=1}^k \frac{\lambda_2(q_i)}{q_i} \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=1}^k E_h\left(\frac{a_i d_i}{q_i}\right) C(q_i, a_i),$$

where $C(q, a) = 1$ for odd q and $|C(q, a)| \leq 2$ in general. We use (44) to write

$$|V_k(y, N, h)| \leq \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \prod_{i=1}^k \frac{2}{q_i} \sum_{\substack{a_1, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=1}^k F_h\left(\frac{a_i}{q_i}\right).$$

Then Lemma 6.3, (45) and the bound in (41) yield

$$\begin{aligned} |V_k(y, N, h)| &\leq \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \frac{2^k}{[q_1, \dots, q_k]} \prod_{i=1}^k \frac{1}{q_i^{\frac{1}{2}}} \left(\sum_{\substack{1 \leq a_i \leq q_i \\ (q_i, a_i) = 1}} F_h\left(\frac{a_i}{q_i}\right)^2 \right)^{\frac{1}{2}} \\ &\leq \sum_{\substack{q_1, \dots, q_k | P_y^N \\ q_i > 1}} \frac{2^k}{[q_1, \dots, q_k]} h^{\frac{k}{2}} \ll_{k, \varepsilon} h^{\frac{k}{2}} y^\varepsilon, \end{aligned}$$

which is the bound announced. \square

6.3. The main estimate. We now prove the analogue of [MS04, (60)], writing $\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}| = k}} \mathfrak{S}_0(\mathcal{D})$ in terms of $V_k(y, N, h)$. Again, the idea of the proof is very similar to the work of Montgomery and Soundararajan, except that we deal with a wider summation (namely a sum over all integers instead of a sum over square-free integers).

Lemma 6.5. *For any $h > k \in \mathbb{N}$, let $y = h^{k+1}$ and $N \geq 4 \log y$. One has*

$$\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}| = k}} \mathfrak{S}_0(\mathcal{D}) = \sum_{j=0}^{k/2} \binom{k}{2j} \frac{(2j)!}{j!2^j} \left(-h \sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2} \right)^j V_{k-2j}(y, N, h) + O_{k, \varepsilon}(h^{\frac{k-1}{2}} y^\varepsilon),$$

where $V_k(y, N, h)$ is defined in (43), and

$$C(d) = \begin{cases} 1 & \text{if } d \text{ is odd} \\ 0 & \text{if } 2 \mid q, 4 \nmid d \\ 4 & \text{if } 4 \mid d. \end{cases}$$

Proof. Following the arguments of [MS04], we can prove the analogue of [MS04, (52)] in our context, which is

$$(47) \quad \sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}| = k}} A_{\mathcal{D}}(q_1, \dots, q_k) = \sum_{\mathcal{P} = \{\mathcal{S}_1, \dots, \mathcal{S}_M\}} w(\mathcal{P}) \sum_{\substack{1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=1}^k C(q_i, a_i) \prod_{m=1}^M \sum_{d_m=1}^h e\left(\sum_{i \in \mathcal{S}_m} \frac{a_i}{q_i} d_m\right),$$

where the first sum is over partitions $\mathcal{P} = \{\mathcal{S}_1, \dots, \mathcal{S}_M\}$ of $\{1, \dots, k\}$, and $w(\mathcal{P})$ is defined in [MS04, p. 17].

In the case of a partition \mathcal{P} containing at least one part of size ≥ 3 , write $\mathcal{N}_1 = \bigcup_{|\mathcal{S}_m|=1} \mathcal{S}_m$, $\mathcal{N}_2 = \{1, \dots, k\} \setminus \mathcal{N}_1$ and $m_2 = |\{1 \leq m \leq M : |\mathcal{S}_m| \geq 2\}|$. Using (44) and $|C(q, x)| \leq 2$, we have

$$\left| \sum_{\substack{1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=1}^k C(q_i, a_i) \prod_{m=1}^M \sum_{d_m=1}^h e\left(\sum_{i \in \mathcal{S}_m} \frac{a_i}{q_i} d_m\right) \right| \leq 2^k h^{m_2} \sum_{\substack{1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i \in \mathcal{N}_1} F_h\left(\frac{a_i}{q_i}\right).$$

Then we apply Lemma 6.3 and the bound (45) to obtain that the sum above is

$$\begin{aligned} &\leq \frac{2^k h^{m_2}}{[q_1, \dots, q_k]} \prod_{i \in \mathcal{N}_1} \left(q_i \sum_{\substack{1 \leq a_i \leq q_i \\ (q_i, a_i) = 1}} (F_h\left(\frac{a_i}{q_i}\right))^2 \right)^{\frac{1}{2}} \prod_{i \in \mathcal{N}_2} \left(q_i \sum_{\substack{1 \leq a_i \leq q_i \\ (q_i, a_i) = 1}} 1^2 \right)^{\frac{1}{2}} \\ &\leq 2^k h^{\frac{k-1}{2}} \frac{q_1 \cdots q_k}{[q_1, \dots, q_k]}, \end{aligned}$$

where we used $\frac{1}{2}|\mathcal{N}_1| + m_2 \leq \frac{k-1}{2}$ when the partition \mathcal{P} contains at least one part of size ≥ 3 . Replacing this bound in (47) and then in (42), we sum over $1 < q_1, \dots, q_k | P_y^N$ as in (42) and use the bound (41) to

obtain that the contribution of the partitions containing at least one part of size ≥ 3 in $\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}|=k}} \mathfrak{S}_0(\mathcal{D})$ is at most $O_{k, \varepsilon}(h^{\frac{k-1}{2} + \varepsilon})$.

We now turn our attention to partitions of $\{1, \dots, k\}$ with sets of size at most 2. The combinatorics leading to [MS04, (56)] work similarly and give

$$\begin{aligned}
\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}|=k}} \mathfrak{S}_0(\mathcal{D}) &= \sum_{0 \leq j \leq \frac{k}{2}} (-1)^j \binom{k}{2j} \frac{(2j)!}{j! 2^j} \sum_{r_1, \dots, r_j | P_y^N} \sum_{\substack{b_1, \dots, b_j \\ 1 \leq b_i \leq r_i, (r_i, b_i)=1}} \prod_{i=1}^j H\left(\frac{b_i}{r_i}\right) \\
(48) \quad &\times \sum_{\substack{q_{2j+1}, \dots, q_k | P_y^N \\ q_i > 1}} \sum_{\substack{a_{2j+1}, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i)=1 \\ \sum_{i=1}^j \frac{b_i}{r_i} + \sum_{i=2j+1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=2j+1}^k \frac{\lambda_2(q_i) C(q_i, a_i)}{q_i} \sum_{d_i=1}^h e\left(\frac{a_i d_i}{q_i}\right) \\
&+ O_k\left(h^{\frac{k-1}{2} + \varepsilon}\right)
\end{aligned}$$

where

$$H\left(\frac{b}{r}\right) = \sum_{\substack{q_1, q_2 | P_y^N \\ q_i > 1}} \sum_{\substack{a_1, a_2 \\ 1 \leq a_i \leq q_i, (q_i, a_i)=1 \\ \frac{a_1}{q_1} + \frac{a_2}{q_2} \in \frac{b}{r} + \mathbb{Z}}} \frac{\lambda_2(q_1) C(q_1, a_1) \lambda_2(q_2) C(q_2, a_2)}{q_1 q_2} \sum_{d=1}^h e\left(\frac{b d}{r}\right).$$

In particular, we have

$$H(1) = \sum_{1 < q | P_y^N} \sum_{\substack{1 \leq a \leq q \\ (q, a)=1}} \frac{C(q, a) C(q, q-a)}{q^2} h = h \sum_{1 < q | P_y^N} \frac{C(q) \phi(q)}{q^2},$$

and the contribution of the terms with all $r_i = 1$ in the sum above is

$$\sum_{0 \leq j \leq \frac{k}{2}} \binom{k}{2j} \frac{(2j)!}{j! 2^j} \left(-h \sum_{1 < q | P_y^N} \frac{C(q) \phi(q)}{q^2}\right)^j V_{k-2j}(y, N, h).$$

We now show that the contribution to (48) of the terms where not all r_i are 1 can be absorbed in the error term. Let ℓ be the number of i 's for which $r_i > 1$. For any $\ell > 0$, and any $r_1, \dots, r_\ell, q_{2j+1}, \dots, q_k > 1$ up to re-ordering and applying Lemma 6.3, we have

$$\begin{aligned}
\sum_{\substack{b_1, \dots, b_\ell \\ 1 \leq b_i \leq r_i, (r_i, b_i)=1}} \prod_{i=1}^{\ell} H\left(\frac{b_i}{r_i}\right) \sum_{\substack{a_{2j+1}, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i)=1 \\ \sum_{i=1}^j \frac{b_i}{r_i} + \sum_{i=2j+1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=2j+1}^k \frac{\lambda_2(q_i) C(q_i, a_i)}{q_i} \sum_{d_i=1}^h e\left(\frac{a_i d_i}{q_i}\right) \\
(49) \quad &\ll_k \frac{1}{[r_1, \dots, r_\ell, q_{2j+1}, \dots, q_k]} \prod_{i=1}^{\ell} \left(r_i \sum_{1 \leq b \leq r_i, (r_i, b)=1} \left|H\left(\frac{b}{r_i}\right)\right|^2\right)^{\frac{1}{2}} \\
&\times \prod_{i=2j+1}^k \left(\frac{1}{q_i} \sum_{1 \leq a \leq q_i, (q_i, a)=1} \left|F_h\left(\frac{a}{q_i}\right)\right|^2\right)^{\frac{1}{2}}.
\end{aligned}$$

To obtain a bound for $|H(\frac{b}{r})|$ we proceed similarly to [MS04] which gives

$$\begin{aligned}
H\left(\frac{b}{r}\right) &\ll F_h\left(\frac{b}{r}\right) \sum_{\substack{s_1, s_2 | P_y^N \\ [s_1, s_2] = r}} \sum_{\substack{c_1, c_2 \\ 1 \leq c_i \leq s_i, (s_i, c_i) = 1 \\ \frac{c_1}{s_1} + \frac{c_2}{s_2} \in \frac{b}{r} + \mathbb{Z}}} \frac{1}{s_1 s_2} \sum_{\substack{t | P_y^N \\ (t, r) = 1}} \frac{\phi(t)}{t^2} \\
&\ll F_h\left(\frac{b}{r}\right) \frac{1}{\phi(r)} \sum_{\substack{s_1, s_2 | P_y^N \\ [s_1, s_2] = r}} \frac{\phi(s_1)\phi(s_2)}{s_1 s_2} \prod_{\substack{p \leq y \\ p \nmid r}} \left(1 + \frac{1}{1 - p^{-1}} \sum_{n=1}^N p^{-n}\right) \\
&\ll F_h\left(\frac{b}{r}\right) \frac{1}{r} \prod_{p|r} (1 - p^{-1})^2 (1 + v_p(r) - v_p(r)p^{-1}) \prod_{\substack{p \leq y \\ p \nmid r}} (1 + p^{-1}) \\
&\ll F_h\left(\frac{b}{r}\right) \frac{d(r)}{r} \log y,
\end{aligned}$$

where $d(r)$ is the number of divisors of r . Using (45), this gives

$$\sum_{\substack{1 \leq b \leq r \\ (r, b) = 1}} |H(\frac{b}{r})|^2 \leq \min(r, h) \frac{d(r)^2}{r} (\log y)^2.$$

Using this bound and (45) in (49) summed over all $r_1, \dots, r_\ell, q_{2j+1}, \dots, q_k > 1$ divisors of P_y^N , we obtain

$$\begin{aligned}
&\sum_{\substack{r_1, \dots, r_\ell | P_y^N \\ r_i > 1}} \sum_{\substack{b_1, \dots, b_\ell \\ 1 \leq b_i \leq r_i, (r_i, b_i) = 1}} \prod_{i=1}^{\ell} H\left(\frac{b_i}{r_i}\right) \sum_{\substack{q_{2j+1}, \dots, q_k | P_y^N \\ q_i > 1}} \sum_{\substack{a_{2j+1}, \dots, a_k \\ 1 \leq a_i \leq q_i, (q_i, a_i) = 1 \\ \sum_{i=1}^j \frac{y_i}{r_i} + \sum_{i=2j+1}^k \frac{a_i}{q_i} \in \mathbb{Z}}} \prod_{i=2j+1}^k \frac{\lambda_2(q_i) C(q_i, a_i)}{q_i} \sum_{d_i=1}^h e\left(\frac{a_i}{q_i} d_i\right) \\
&\ll_k \sum_{\substack{r_1, \dots, r_\ell | P_y^N \\ r_i > 1}} \sum_{\substack{q_{2j+1}, \dots, q_k | P_y^N \\ q_i > 1}} \frac{1}{[r_1, \dots, r_\ell, q_{2j+1}, \dots, q_k]} \prod_{i=1}^{\ell} \left(h^{\frac{1}{2}} d(r_i) \log y\right) h^{\frac{k-2j}{2}} \\
&\ll_k h^{\frac{k+\ell-2j}{2}} (\log y)^\ell \sum_{m | P_y^N} \frac{1}{m} \left(\sum_{r|m} d(r)\right)^\ell \left(\sum_{q|m} 1\right)^{k-2j} \\
&\ll_k h^{\frac{k+\ell-2j}{2}} (\log y)^\ell \prod_{p | P_y} \left(1 + \sum_{n=1}^N (n+1)^{k+\ell-2j} \left(\frac{n+2}{2}\right)^\ell p^{-n}\right) \ll_{k, \ell, j, \varepsilon} h^{\frac{k+\ell-2j}{2}} y^\varepsilon.
\end{aligned}$$

Finally, summing the contribution for each $\ell \geq 0$ yields

$$\begin{aligned}
\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}| = k}} \mathfrak{S}_0(\mathcal{D}) &= \sum_{j=0}^{k/2} \binom{k}{2j} \frac{(2j)!}{j! 2^j} \left(-h \sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2}\right)^j V_{k-2j}(y, N, h) \\
&\quad + \sum_{j=0}^{k/2} \binom{k}{2j} \frac{(2j)!}{j! 2^j} \sum_{\ell=1}^j \binom{j}{\ell} \left(h \sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2}\right)^{j-\ell} O\left(h^{\frac{k+\ell-2j}{2}} y^\varepsilon\right) + O_k\left(h^{(k-1+\varepsilon)/2}\right),
\end{aligned}$$

and using $\sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2} \ll_\varepsilon y^\varepsilon$, we deduce

$$\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}| = k}} \mathfrak{S}_0(\mathcal{D}) = \sum_{j=0}^{k/2} \binom{k}{2j} \frac{(2j)!}{j! 2^j} \left(-h \sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2}\right)^j V_{k-2j}(y, N, h) + O_{k, \varepsilon}\left(h^{\frac{k-1}{2}} y^\varepsilon\right).$$

which completes the proof of Lemma 6.5. \square

The proof of Theorem 3.4 is now relatively straightforward. Lemma 6.5 gives for any $h > k \in \mathbb{N}$, and $N \geq 4(k+1) \log h$ that

$$\sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}|=k}} \mathfrak{S}_0(\mathcal{D}) = \sum_{j=0}^{k/2} \binom{k}{2j} \frac{(2j)!}{j!2^j} \left(-h \sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2} \right)^j V_{k-2j}(h^{k+1}, N, h) + O_{k,\varepsilon}(h^{\frac{k-1}{2}+\varepsilon}).$$

Then the bound from Corollary 6.4 yields

$$\begin{aligned} \sum_{\substack{\mathcal{D} \subseteq [1, h] \\ |\mathcal{D}|=k}} \mathfrak{S}_0(\mathcal{D}) &\ll_{k,\varepsilon} \sum_{j=0}^{k/2} \binom{k}{2j} \frac{(2j)!}{j!2^j} \left(h \sum_{1 < d | P_y^N} \frac{C(d)\phi(d)}{d^2} \right)^j h^{\frac{k-2j}{2}+\varepsilon} + O_{k,\varepsilon}(h^{\frac{k-1}{2}+\varepsilon}) \\ &\ll_k h^{\frac{k}{2}+\varepsilon} \left(\prod_{p < y} 1 - \frac{1}{p} \right)^{\frac{-k}{2}} \ll_{k,\varepsilon} h^{\frac{k}{2}+2\varepsilon}, \end{aligned}$$

which finishes the proof of Theorem 3.4.

7. INTEGRAL FORM AND IMPROVED ERROR TERMS

Using the methods of Section 5 and Theorem 3.4, we can obtain a more precise form of the averages of the Hardy–Littlewood constants for sums of two squares of [Smi13, Theorem 1.1] and [FKR17, Proposition 1.3] (in a special case) by exhibiting a secondary term. In order to see the secondary term, we need to express the results of Section 5 differently, as a closed-form expression which contains implicitly all the descending powers of $\log h$. We first prove that we can write such an asymptotic for the number of sums of two squares, with a square-root cancellation error term under the Riemann Hypothesis (Theorem 2.1). The argument for the proof of Theorem 2.1 is essentially due to Selberg and known to experts, it appeared as a mathoverflow post [Luc13], and an exercise in the book of Koukoulopoulos [Kou19, Exercise 13.7]. Note also the observation of Tenenbaum [Ten15, page 291] as well as the independent analogue result of Gorodetsky and Rodgers [GR20, Theorem B.1] inspired by [Ram76]. With the same techniques, we then prove Proposition 7.1, which exhibits the secondary term for the average of the Hardy–Littlewood constants for 2-tuples of sums of two squares. The general case is Proposition 1.3 and it follows by using Theorem 3.4 to show that the average over k -tuples reduces to the average over 2-tuples.

Proof of Theorem 2.1: We first assume the Riemann Hypothesis. Using Perron’s formula, we have for any $\delta > 0$

$$(50) \quad \sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) = \int_{1+\delta-iT}^{1+\delta+iT} F(s) \frac{x^s}{s} ds + O\left(\frac{x^{1+\delta} \log x}{T}\right),$$

where $F(s) = \zeta^{1/2}(s)L(s, \chi_4)^{1/2}(1-2^{-s})^{-1/2} \prod_{p \equiv 3 \pmod{4}} (1-p^{-2s})^{-1/2}$ as seen in the proof of Theorem 2.4. The above path integral is part of a contour which enclose a region of analyticity of the integrand, which is the usual contour going from $1+\delta-iT$ to $1+\delta+iT$ then to $1/2+\varepsilon+iT$ then to $1/2+\varepsilon-iT$ and then back to $1+\delta-iT$ with a slit along the real axis between $1/2+\varepsilon$ and 1, with a line just above the real axis from $1/2+\varepsilon$ to 1, and a line just below the real axis from 1 to $1/2+\varepsilon$. More precisely, for any $\varepsilon, \eta > 0$ and for $0 < \kappa < \delta$, we define line segments L_j , $j = 1, 2, \dots, 7$ as in Figure 2.

Together with the line segment $1+\delta+iT \rightarrow 1+\delta-iT$ of the integral (50), this gives the closed contour of Figure 2, which enclose a region of analyticity of the function $F(s) = G(s)(s-1)^{-1/2}$ since we are assuming the Riemann Hypothesis and $\zeta^{1/2}(s)(s-1)^{1/2}$, $L(s, \chi_4)^{1/2}(1-2^{-s})^{-1/2} \prod_{p \equiv 3 \pmod{4}} (1-p^{-2s})^{-1/2}$ are analytic for $\operatorname{Re}(s) > 1/2+\varepsilon$. Then, using Cauchy’s theorem, we have

$$\int_{1+\delta-iT}^{1+\delta+iT} F(s) \frac{x^s}{s} ds = \sum_{j=1}^7 \int_{L_j} F(s) \frac{x^s}{s} ds.$$

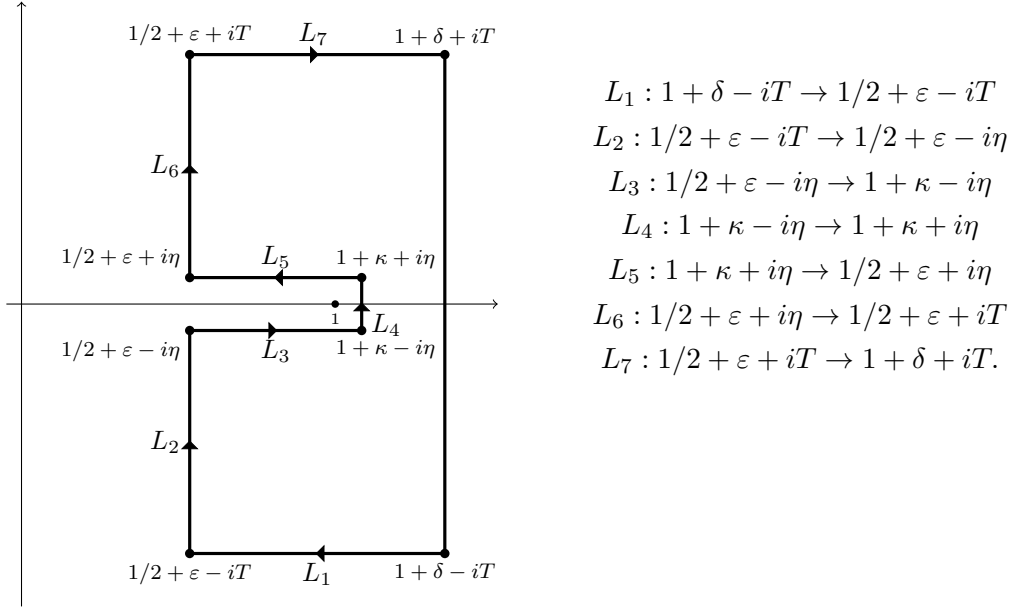


FIGURE 2. The contour used in the proof of Theorem 2.4.

The contribution coming from L_1, L_2, L_4, L_6, L_7 are bounded by the classical estimates, where we use the Lindelöf Hypothesis to bound $|\zeta^{1/2}(\sigma + it)|, |L^{1/2}(\sigma + it, \chi_4)| \ll_\sigma |t|^{\varepsilon_1}$ for $1/2 < \sigma < 1$ and $\varepsilon_1 > 0$. For the horizontal integral over L_1 , we have

$$\int_{L_1} F(s) \frac{x^s}{s} ds \ll \int_{1/2+\varepsilon}^{1+\delta} \frac{x^\sigma}{T^{1-2\varepsilon_1}} d\sigma = O\left(\frac{x^{1+\delta}}{T^{1-2\varepsilon_1}}\right),$$

where we also used the fact that the Euler product $(1 - 2^{-s})^{-1/2} \prod_{p \equiv 3 \pmod{4}} (1 - p^{-2s})^{-1/2}$ is absolutely bounded for $\operatorname{Re}(s) > 1/2 + \varepsilon$. We get the same bound for \int_{L_7} .

For the vertical integral over L_2 , we have

$$\int_{L_2} F(s) \frac{x^s}{s} ds \ll \int_\eta^T \frac{x^{1/2+\varepsilon}}{(t + \frac{1}{2})^{1-2\varepsilon_1}} dt = O\left(x^{1/2+\varepsilon} T^{2\varepsilon_1}\right),$$

which also hold for \int_{L_6} . Finally, we have

$$\int_{L_4} F(s) \frac{x^s}{s} ds \ll \eta x^{1+\kappa},$$

and choosing $T = x^{1/2}$ and $\eta < x^{-\frac{1}{2}-\kappa}$, this gives

$$\int_{1+\delta-iT}^{1+\delta+iT} F(s) \frac{x^s}{s} ds = \lim_{\eta \rightarrow 0^+} \left(\int_{1/2+\varepsilon-i\eta}^{1+\kappa-i\eta} - \int_{1/2+\varepsilon+i\eta}^{1+\kappa+i\eta} \right) F(s) \frac{x^s}{s} ds + O\left(x^{1/2+\varepsilon}\right).$$

Note that κ can be arbitrarily small, and choosing for example $\kappa = x^{-2}$, we have

$$\lim_{\eta \rightarrow 0^+} \left(\int_{1/2+\varepsilon-i\eta}^{1+\kappa-i\eta} - \int_{1/2+\varepsilon+i\eta}^{1+\kappa+i\eta} \right) F(s) \frac{x^s}{s} ds = \lim_{\eta \rightarrow 0^+} \left(\int_{1/2+\varepsilon-i\eta}^{1-i\eta} - \int_{1/2+\varepsilon+i\eta}^{1+i\eta} \right) F(s) \frac{x^s}{s} ds + O(1).$$

Putting everything together, we have

$$\sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) = \frac{1}{2\pi i} \int_{1/2+\varepsilon}^1 G(\sigma) \frac{x^\sigma}{\sigma} \lim_{\eta \rightarrow 0^+} \left((\sigma - i\eta - 1)^{-1/2} - (\sigma + i\eta - 1)^{-1/2} \right) d\sigma + O\left(x^{1/2+\varepsilon}\right),$$

where $G(s) = (s - 1)^{1/2} F(s)$.

We use the fact that when $\sigma \in (0, 1)$, $\log(\sigma \pm i\eta - 1) \sim \log|\sigma - 1| \pm i\pi$ as $\eta \rightarrow 0^+$. Writing $(\sigma \pm i\eta - 1)^{-1/2} = \exp(-\frac{1}{2}\log(\sigma \pm i\eta - 1))$, we see that $(\sigma \pm i\eta - 1)^{-1/2} \sim \mp i|\sigma - 1|^{-1/2}$, and we have

$$\lim_{\eta \rightarrow 0^+} \left((\sigma - i\eta - 1)^{-1/2} - (\sigma + i\eta - 1)^{-1/2} \right) = 2i|\sigma - 1|^{-1/2}.$$

Replacing above, this proves the theorem under the Riemann Hypothesis. Unconditionally, we start from (50), and we use a similar contour, but with $1/2 + \varepsilon$ replaced by $1 - c/\sqrt{\log x}$, where c is small enough to insure that the contour does not contain any zeroes of $\zeta(s)$ or $L(s, \chi_4)$. Working as above, we get

$$\sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) = \frac{1}{\pi} \int_{1 - \frac{c}{\sqrt{\log x}}}^1 \frac{x^\sigma}{\sigma} G(\sigma) |\sigma - 1|^{-1/2} d\sigma + O\left(\frac{x^{1+\delta}}{T^{1-2\varepsilon_1}} + x \exp(-c\sqrt{\log x}) T^{2\varepsilon_1}\right),$$

and choosing $\delta = 1/\log x$ and $T = \exp(c\sqrt{\log x})$, we get

$$\sum_{n \leq x} \mathbf{1}_{\mathbf{E}}(n) = \frac{1}{\pi} \int_{1 - c/\sqrt{\log x}}^1 \frac{x^\sigma}{\sigma} G(\sigma) |\sigma - 1|^{-1/2} d\sigma + O\left(x \exp(-c_0\sqrt{\log x})\right),$$

for some $c_0 > 0$. Finally, we have

$$\int_{\frac{1}{2} + \varepsilon}^{1 - c/\sqrt{\log x}} \frac{x^\sigma}{\sigma} G(\sigma) |\sigma - 1|^{-1/2} d\sigma \ll \int_{\frac{1}{2} + \varepsilon}^{1 - c/\sqrt{\log x}} \frac{x^\sigma}{\sigma |\sigma - 1|^{1/2}} d\sigma \ll x^{1 - c/\sqrt{\log x}}$$

which shows the unconditional result. \square

The following proposition is a more precise version of [Smi13, Theorem 1.1] who showed that

$$\sum_{\substack{1 \leq d_1, d_2 \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, d_k\}) = H^2 + O(H^{1+\varepsilon}).$$

We remark that our normalization differs from [Smi13] for the singular series.

Proposition 7.1. *Fix $\varepsilon > 0$. There exists $c > 0$ such that*

$$\sum_{\substack{1 \leq d_1, d_2 \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, d_k\}) = H^2 + \frac{2}{\pi K^2} \int_{1/2 + \varepsilon}^1 \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma - 1|^{1/2}} d\sigma + O\left(H \exp(-c\sqrt{\log H})\right)$$

where $F(s) = \zeta(s-1)M(s-1)[(s-1)\zeta(s)]^{1/2} s^{-1}$, with $M(s)$ as defined by (37). Assuming the Riemann Hypothesis, we can replace the error term by $O(H^{1/2+\varepsilon})$.

Proof. As in [Smi13, § 2.3], we have

$$\begin{aligned} \sum_{\substack{1 \leq d_1, d_2 \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, d_k\}) &= 2 \sum_{1 \leq d < H} \mathfrak{S}(\{0, d\})(H - d) \\ &= \frac{1}{K^2} \frac{1}{2i\pi} \int_{(2)} \frac{D(s)}{s(s+1)} H^{s+1} ds, \end{aligned}$$

where $D(s) = \zeta(s)\zeta(s+1)^{\frac{1}{2}}M(s)$ as defined in the beginning of section 5. As [Smi13], we compute the main term, coming from the pole of $D(s)$ at $s = 1$, which gives

$$\sum_{\substack{1 \leq d_1, d_2 \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, d_k\}) = H^2 + \frac{1}{K^2} \frac{1}{2i\pi} \int_{(\varepsilon)} \frac{D(s)}{s(s+1)} H^{s+1} ds.$$

We first assume the Riemann hypothesis and we evaluate the integral

$$\frac{1}{2i\pi} \int_{(\varepsilon)} \frac{D(s)}{s(s+1)} H^{s+1} ds = \frac{1}{2i\pi} \int_{(1+\varepsilon)} \frac{F(s)}{(s-1)^{3/2}} H^s ds$$

where $F(s) = \zeta(s-1)M(s-1)[(s-1)\zeta(s)]^{1/2}s^{-1}$ is analytic for $\operatorname{Re}(s) > 1/2 + \varepsilon$. We begin with an integration by part to obtain

$$\begin{aligned} \int_{(1+\varepsilon)} \frac{F(s)}{(s-1)^{3/2}} H^s ds &= \lim_{T \rightarrow \infty} [-2F(s)H^s(s-1)^{-1/2}]_{1+\varepsilon-iT}^{1+\varepsilon+iT} + 2 \int_{(1+\varepsilon)} \frac{F'(s)H^s + F(s)H^s \log H}{(s-1)^{1/2}} ds \\ &= 2 \int_{(1+\varepsilon)} \frac{F'(s)H^s + F(s)H^s \log H}{(s-1)^{1/2}} ds. \end{aligned}$$

To evaluate the last integral, we first approximate the line integral by the segment from $1 + \varepsilon - iT$ to $1 + \varepsilon + iT$, and use the contour of Figure 2. Working as in the proof of Theorem 2.1, we get

$$\begin{aligned} &\frac{2}{2\pi i} \int_{(1+\varepsilon)} \frac{F'(s)H^s + F(s)H^s \log H}{(s-1)^{1/2}} ds \\ &= \frac{1}{\pi i} \int_{1/2+\varepsilon}^1 (F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H) \lim_{\eta \rightarrow 0^+} \left((\sigma-1-i\eta)^{-1/2} - (\sigma-1+i\eta)^{-1/2} \right) d\sigma + O\left(H^{1/2+\varepsilon}\right) \\ &= \frac{2}{\pi} \int_{1/2+\varepsilon}^1 (F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H) |\sigma-1|^{-1/2} d\sigma + O\left(H^{1/2+\varepsilon}\right). \end{aligned}$$

Replacing above, this gives (under the Riemann Hypothesis)

$$\frac{1}{K^2} \frac{1}{2\pi i} \int_{(\varepsilon)} \frac{D(s)}{s(s+1)} H^{s+1} ds = \frac{2}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma + O\left(H^{1/2+\varepsilon}\right).$$

To do a proof without the Riemann Hypothesis, we proceed as in the proof of Theorem 2.1, and we get

$$\frac{2}{2\pi i} \int_{(1+\varepsilon)} \frac{F'(s)H^s + F(s)H^s \log H}{(s-1)^{1/2}} ds = \frac{2}{\pi} \int_{1-c_1/\sqrt{\log H}}^1 \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma + O\left(H \exp\left(-c\sqrt{\log H}\right)\right).$$

To conclude the proof, we show that

$$\int_{1-c_1/\sqrt{\log H}}^1 \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma = \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma + O\left(H \exp\left(-c\sqrt{\log H}\right)\right).$$

This follows from the fact that ζ does not vanish on $[\frac{1}{2} + \varepsilon, 1]$, so F and F' are defined and continuous on $[\frac{1}{2} + \varepsilon, 1]$, in particular, they are uniformly bounded. We have

$$\begin{aligned} \int_{\frac{1}{2}}^{1-c_1/\sqrt{\log H}} \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma &\ll_F \int_{\frac{1}{2}}^{1-c_1/\sqrt{\log H}} \frac{H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma \\ &\ll H^{1-c_1/\sqrt{\log H}} (\log H)^{\frac{5}{4}} \ll_c H \exp\left(-c\sqrt{\log H}\right) \end{aligned}$$

for any $c < c_1$. \square

We can now prove Proposition 1.3. We observe that it is a more precise version of (a particular case of) [FKR17, Proposition 1.3] who showed that

$$\sum_{\substack{1 \leq d_1, \dots, d_k \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, \dots, d_k\}) = H^k + O\left(H^{k-2/3+o(1)}\right).$$

Proof of Proposition 1.3. Note that the cases $k = 0$ or 1 are easy. We have $\mathfrak{S}(\emptyset) = \mathfrak{S}(\{d\}) = 1$, so we obtain 1 and H respectively, without error term. The case $k = 2$ is proven in Proposition 7.1. Similarly to [MS04, (17)], we have

$$\begin{aligned} \sum_{\substack{1 \leq d_1, \dots, d_k \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, \dots, d_k\}) &= \sum_{r=0}^k \binom{k}{r} \frac{(H-r)!}{(H-k)!} \sum_{\substack{1 \leq d_1, \dots, d_r \leq H \\ \text{distinct}}} \mathfrak{S}_0(\{d_1, \dots, d_r\}) \\ &= \frac{H!}{(H-k)!} + \binom{k}{2} \frac{(H-2)!}{(H-k)!} \sum_{\substack{1 \leq d_1, d_2 \leq H \\ \text{distinct}}} \mathfrak{S}_0(\{d_1, d_2\}) + O\left(H^{k-\frac{3}{2}+\varepsilon}\right), \end{aligned}$$

where we used the decomposition $\mathfrak{S}(\mathcal{H}) = \sum_{\mathcal{T} \subseteq \mathcal{H}} \mathfrak{S}_0(\mathcal{T})$, the fact that $\mathfrak{S}_0(\{d\}) = 0$, and the bound from Theorem 3.4 as soon as the size of the set is larger than 2. Using the estimates

$$\begin{aligned} \frac{H!}{(H-k)!} &= H(H-1)\dots(H-k+1) = H^k + H^{k-1} \sum_{i=1}^{k-1} (-i) + O_k(H^{k-2}) \\ &= H^k - H^{k-1} \frac{k(k-1)}{2} + O_k(H^{k-2}), \\ \frac{(H-2)!}{(H-k)!} &= H^{k-2} + O_k(H^{k-3}), \end{aligned}$$

and Proposition 7.1, this gives

$$\begin{aligned} &\sum_{\substack{1 \leq d_1, \dots, d_k \leq H \\ \text{distinct}}} \mathfrak{S}(\{d_1, \dots, d_k\}) \\ &= H^k - \frac{k(k-1)}{2} H^{k-1} + \binom{k}{2} H^{k-2} \sum_{\substack{1 \leq d_1, d_2 \leq H \\ \text{distinct}}} (\mathfrak{S}(\{d_1, d_2\}) - 1) + O(H^{k-\frac{3}{2}+\varepsilon}) \\ &= H^k - \frac{k(k-1)}{2} H^{k-1} + \binom{k}{2} H^{k-2} \left(\frac{2}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^\sigma + F(\sigma)H^\sigma \log H}{|\sigma-1|^{1/2}} d\sigma + H \right) + O(H^{k-\frac{3}{2}+\varepsilon}) \\ &= H^k + k(k-1) \frac{H^{k-1}}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^{\sigma-1} + F(\sigma)H^{\sigma-1} \log H}{|\sigma-1|^{1/2}} d\sigma + O(H^{k-\frac{3}{2}+\varepsilon}). \end{aligned}$$

□

We conclude this section by proving Proposition 4.4. The proof is similar to the other proofs of this section, and we skip the details.

Proof of Proposition 4.4: Starting from (30), we write

$$S(H) = H + \frac{1}{2K^2} \frac{1}{2\pi i} \int_{(1+\varepsilon)} \frac{F(s)}{(s-1)^{\frac{3}{2}}} H^{s-1} ds,$$

where $F(s) = \zeta(s-1)M(s-1)[(s-1)\zeta(s)]^{1/2} \Gamma(s)$, with $M(s)$ as defined by (37). Proceeding as in the proof of Proposition 7.1, with an integration by part before moving the contour of integration gives the following

$$S(H) = H + \frac{1}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma)H^{\sigma-1} + F(\sigma)H^{\sigma-1} \log H}{|\sigma-1|^{1/2}} d\sigma + O\left(\exp(-c\sqrt{\log H})\right).$$

Similarly, using (29) and without the integration by part, we have

$$\begin{aligned} S(H, \chi_0) &= \frac{\phi(q)}{q} H + \frac{1}{2K^2} \frac{1}{2\pi i} \int_{(1+\varepsilon)} \frac{F_{\chi_0}(s)}{(s-1)^{\frac{1}{2}}} H^{s-1} ds \\ &= \frac{\phi(q)}{q} H + \frac{1}{2\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F_{\chi_0}(\sigma)H^{\sigma-1}}{|\sigma-1|^{1/2}} d\sigma + O\left(\exp(-c\sqrt{\log H})\right), \end{aligned}$$

where

$$\begin{aligned} F_{\chi_0}(s) &= L(s-1, \chi_0) \Gamma(s-1) M_{\chi_0}(s-1) [(s-1)L(s, \chi_0)]^{1/2} \\ &= \frac{1 - q^{-(s-1)}}{s-1} F(s) =: A_q(s) F(s) \end{aligned}$$

where we used $M_{\chi_0}(s) = (1 - q^{-(s+1)})^{-1/2} M(s)$. Assuming the Riemann Hypothesis, we can replace the error term by $O(H^{-1/2+\varepsilon})$. Then, we obtain the expressions in Proposition 4.4 by using the orthogonality of characters and expression (27) for the contribution of non-trivial characters as in the proof of

Theorem 1.2. For $v \not\equiv 0 \pmod{q}$, we have

$$\begin{aligned} S(q, v, H) &\sim \frac{1}{2K^2\phi(q)} \sum_{\chi \neq \chi_0} \chi(v)^{-1} C_{q,\chi} + \frac{1}{\phi(q)} S(H, \chi_0) \\ &\sim \frac{H}{q} + \frac{1}{2K^2\phi(q)} \sum_{\chi \neq \chi_0} \chi(v)^{-1} C_{q,\chi} + \frac{1}{2\pi K^2\phi(q)} \int_{1/2+\varepsilon}^1 \frac{F_{\chi_0}(\sigma) H^{\sigma-1}}{|\sigma-1|^{1/2}} d\sigma \end{aligned}$$

and

$$\begin{aligned} S(q, 0, H) &\sim S(H) - \frac{\phi(q)}{q} H - \frac{1}{2\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F_{\chi_0}(\sigma) H^{\sigma-1}}{|\sigma-1|^{1/2}} d\sigma \\ &\sim \frac{H}{q} + \frac{1}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma) H^{\sigma-1} + F(\sigma) H^{\sigma-1} \log H}{|\sigma-1|^{1/2}} d\sigma - \frac{1}{2\pi K^2} \int_{1/2+\varepsilon}^1 \frac{A_q(\sigma) F(\sigma) H^{\sigma-1}}{|\sigma-1|^{1/2}} d\sigma \\ &\sim \frac{H}{q} + \frac{1}{\pi K^2} \int_{1/2+\varepsilon}^1 \frac{F'(\sigma) H^{\sigma-1} + F(\sigma) H^{\sigma-1} (\log H - A_q(\sigma)/2)}{|\sigma-1|^{1/2}} d\sigma. \end{aligned}$$

□

8. HEURISTIC IN THE CASE OF r -UPLETS

As in [LOS16], the essence for the general conjecture in the case of the distribution of r consecutive sums of two squares is really in the particular case $r = 2$ that we explained in more details. In this section we present the heuristic for Conjecture 1.4 with highlights on the differences from the case $r = 2$, for this we follow again the exposition of [LOS16]. Let $r \geq 3$, $q \equiv 1 \pmod{4}$ and $\mathbf{a} = (a_1, \dots, a_r) \in \mathbb{N}^r$ be fixed. We write

$$\begin{aligned} N(x; q, \mathbf{a}) &= \sum_{\substack{n \leq x \\ n \equiv a_1 \pmod{q}}} \sum_{\substack{h_2, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \mathbf{1}_{\mathbf{E}}(n) \prod_{i=2}^r \mathbf{1}_{\mathbf{E}}(n + h_2 + \dots + h_i) \\ &\quad \times \prod_{t=1}^{h_i-1} (1 - \mathbf{1}_{\mathbf{E}}(n + h_2 + \dots + h_{i-1} + t)). \end{aligned}$$

As in Section 4, we use the notation $\tilde{\mathbf{1}}_{\mathbf{E}}(n) = \mathbf{1}_{\mathbf{E}}(n) - \frac{K}{\sqrt{\log n}}$, approximate all the $\log(n+t)$ by $\log x$, expand out the products and apply the Hardy–Littlewood Conjecture (12) in our context, neglecting the terms corresponding to products over more than 3 terms thanks to Theorem 3.4. Thus, heuristically, up to error of size $x(\log x)^{-\frac{r}{2}-\frac{1}{4}+\varepsilon}$, we obtain

$$N(x; q, \mathbf{a}) \sim \frac{x}{q} \left(\frac{K}{\sqrt{\log x}} \right)^r \alpha(x)^{-r+1} (\mathcal{D}_0(\mathbf{a}, x) + \mathcal{D}_1(\mathbf{a}, x) + \mathcal{D}_2(\mathbf{a}, x)),$$

where $\alpha(x) = 1 - \frac{K}{\sqrt{\log x}}$ and

$$\mathcal{D}_0(\mathbf{a}, x) = \sum_{\substack{h_2, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \left(1 + \sum_{1 \leq i < j \leq r} \mathfrak{S}_0(\{0, h_{i+1} + \dots + h_j\}) \right) \alpha(x)^{h_2 + \dots + h_r}$$

$$\mathcal{D}_1(\mathbf{a}, x) = -\frac{K}{\alpha(x)\sqrt{\log x}} \sum_{\substack{h_2, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \sum_{i=1}^r \sum_{j=2}^r \sum_{t=1}^{h_j-1} \mathfrak{S}_0(\{h_2 + \dots + h_i, h_2 + \dots + h_{j-1} + t\}) \alpha(x)^{h_2 + \dots + h_r}$$

$$\mathcal{D}_2(\mathbf{a}, x) = \frac{K^2}{\alpha(x)^2 \log x} \sum_{\substack{h_2, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \sum_{2 \leq i < j \leq r} \sum_{t_1=1}^{h_i-1} \sum_{\substack{t_2=1 \\ t_2 > t_1 \text{ if } i=j}}^{h_j-1} \mathfrak{S}_0(\{t_1, h_i + \dots + h_{j-1} + t_2\}) \alpha(x)^{h_2 + \dots + h_r}.$$

Let us begin with studying $\mathcal{D}_0(\mathbf{a}, x)$ in more details. As in Section 4, we write $H = -\frac{1}{\log \alpha(x)} \iff \alpha(x)^h = e(-h/H)$. The contribution of the 1 to $\mathcal{D}_0(\mathbf{a}, x)$ gives

$$(51) \quad \sum_{\substack{h_2, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} e^{-(h_2 + \dots + h_r)/H} = \prod_{\ell=2}^r \left(\frac{H}{q} + f(a_\ell - a_{\ell-1}; q) + O(H^{-1}) \right) \\ = \left(\frac{H}{q} \right)^{r-1} + \left(\frac{H}{q} \right)^{r-2} \sum_{\ell=2}^r f(a_\ell - a_{\ell-1}; q) + O(H^{r-3}).$$

For the contribution of $\sum_{1 \leq i < j \leq r}$ to $\mathcal{D}_0(\mathbf{a}, x)$, we first make a change of variables by writing $j = i + k$, and we exchange the order of summation, which gives

$$\sum_{\substack{1 \leq i \leq r-1 \\ 1 \leq k \leq r-i}} \left(\sum_{\substack{h_2, \dots, h_i, h_{i+k+1}, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} e^{-(h_2 + \dots + h_i + h_{i+k+1} + h_r)/H} \right) \\ \times \left(\sum_{\substack{h_{i+1}, \dots, h_{i+k} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \mathfrak{S}_0(\{0, h_{i+1} + \dots + h_{i+k}\}) e^{-(h_{i+1} + \dots + h_{i+k})/H} \right).$$

For each fixed i, k , the second factor in the inner sum above is

$$(52) \quad \sum_{\substack{h > 0 \\ h \equiv a_{i+k} - a_i \pmod{q}}} \mathfrak{S}_0(\{0, h\}) e^{-h/H} \sum_{\substack{h_{i+1}, \dots, h_{i+k} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q} \\ h_{i+1} + \dots + h_{i+k} = h}} 1 \\ = \frac{1}{(k-1)! q^{k-1}} \sum_{\substack{h > 0 \\ h \equiv a_{i+k} - a_i \pmod{q}}} \mathfrak{S}_0(\{0, h\}) e^{-h/H} (h^{k-1} + O(h^{k-2})).$$

We need some notation, generalizing the functions defined in Section 4.2. For $v, k \in \mathbb{N}$, let

$$S^{(k)}(q, v, H) := \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} \mathfrak{S}(\{0, h\}) h^k e^{-h/H} \\ S_0^{(k)}(q, v, H) := \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} \mathfrak{S}_0(\{0, h\}) h^k e^{-h/H} \\ S^{(k)}(H) := \sum_{h \geq 1} \mathfrak{S}(\{0, h\}) h^k e^{-h/H} \\ S_0^{(k)}(H) := \sum_{h \geq 1} \mathfrak{S}_0(\{0, h\}) h^k e^{-h/H}.$$

Note that $S^{(0)}(q, v, H) = S(q, v, H)$ as defined in (19). Moreover, we have

$$S_0^{(k)}(H) = S^{(k)}(H) - \sum_{h \geq 1} h^k e^{-h/H} = S^{(k)}(H) - k! H^{k+1} + O(H^{k-1})$$

$$\text{and} \quad S_0^{(k)}(q, v, H) = S^{(k)}(q, v, H) - \sum_{\substack{h \geq 1 \\ h \equiv v \pmod{q}}} h^k e^{-h/H} = S^{(k)}(q, v, H) - \frac{k!}{q} H^{k+1} + O(H^{k-1}).$$

Proposition 8.1. *Let $q \equiv 1 \pmod{4}$ be a prime. For any $k \geq 1$, we have*

$$S^{(k)}(H) = k! H^{k+1} - \frac{(k-1)!}{K\sqrt{\pi}} H^k (\log H)^{-\frac{1}{2}} + O(H^k (\log H)^{-\frac{3}{2}}).$$

and

$$S^{(k)}(q, v, H) = \begin{cases} \frac{k!}{q} H^{k+1} + O(H^k (\log H)^{-\frac{3}{2}}) & \text{if } v \not\equiv 0 \pmod{q} \\ \frac{k!}{q} H^{k+1} - \frac{(k-1)!}{K\sqrt{\pi}} H^k (\log H)^{-\frac{1}{2}} + O(H^k (\log H)^{-\frac{3}{2}}) & \text{if } v \equiv 0 \pmod{q}. \end{cases}$$

We observe that the secondary term is relatively smaller in the case $k \geq 1$ than in the case $k = 0$ (which is Theorem 1.2). This is due to the fact that the order of the singularity is smaller when $k \geq 1$ as the poles of the functions $\zeta(k+1+s)$ and $\Gamma(s)$ do not coincide. Note also that, similarly to Theorem 1.2, one could develop the secondary term using a sum of descending powers of $\log H$ with explicit coefficients. We chose not to do so in this statement as we are mostly interested in the direction of the bias in the distribution of consecutive sums of two squares in arithmetic progressions.

Proof. The proof is similar to the proof of Theorem 1.2, and we just give a sketch. The main idea is to approximate the sums $S^{(k)}(H)$ and $S^{(k)}(q, v, H)$ via contour integration of the shifted functions $D(s)$ and $D_\chi(s-k)$ (for χ a character modulo q) respectively, where the functions $D(s)$ and D_χ are as defined in Section 5. For $k \geq 1$ and $\chi \neq \chi_0$, the function $\Gamma(s)D_\chi(s-k)$ is analytic on a zero free region containing the line $\operatorname{Re}(s) = k$, thus, we have

$$\sum_{h \geq 1} 2K^2 \mathfrak{S}(\{0, h\}) \chi(h) h^k e^{-h/H} = O(H^k e^{-c\sqrt{\log H}}).$$

For $S^{(k)}(H)$, the function $\Gamma(s)D(s-k)$ has a simple pole at $s = k+1$ with residue $2K^2\Gamma(k+1)$ and an essential singularity at $s = k$ of the shape $(s-k)^{-\frac{1}{2}}$. We deduce that

$$\sum_{h \geq 1} 2K^2 \mathfrak{S}(\{0, h\}) h^k e^{-h/H} = H^{k+1} 2K^2 \Gamma(k+1) - 2\Gamma(k) \frac{K}{\sqrt{\pi}} H^k (\log H)^{-\frac{1}{2}} + O(H^k (\log H)^{-\frac{3}{2}}),$$

which gives

$$S^{(k)}(H) = H^{k+1} \Gamma(k+1) - \Gamma(k) \frac{1}{K\sqrt{\pi}} H^k (\log H)^{-\frac{1}{2}} + O(H^k (\log H)^{-\frac{3}{2}}).$$

In the case $\chi = \chi_0$, the function $\Gamma(s)D_{\chi_0}(s-k)$ has a simple pole at $s = k+1$ with residue $2K^2\Gamma(k+1)\frac{\phi(q)}{q}$ and an essential singularity at $s = k$ of the shape $(s-k)^{\frac{1}{2}}$. We deduce

$$\sum_{h \geq 1} 2K^2 \mathfrak{S}(\{0, h\}) \chi_0(h) h^k e^{-h/H} = H^{k+1} 2K^2 \Gamma(k+1) \frac{\phi(q)}{q} + O(H^k (\log H)^{-\frac{3}{2}}).$$

Finally, we obtain the expressions in the statement of Lemma 8.1 using the orthogonality relations in the case $v \not\equiv 0 \pmod{q}$, and the case $v \equiv 0 \pmod{q}$ is then deduced by subtracting the contributions of all non-zero v 's to $S^{(k)}(H)$. \square

Using Lemma 8.1 and (51), (52), we get

$$\begin{aligned} \mathcal{D}_0(\mathbf{a}, x) &= \left(\frac{H}{q}\right)^{r-1} + \left(\frac{H}{q}\right)^{r-2} \sum_{i=2}^r f(a_i - a_{i-1}; q) \\ &\quad + \sum_{i=1}^{r-1} \sum_{k=1}^{r-i} \left(\frac{H}{q}\right)^{r-k-1} \frac{1}{(k-1)! q^{k-1}} S_0^{(k-1)}(q, a_{i+k} - a_i, H) + O(H^{r-3}) \\ &= \left(\frac{H}{q}\right)^{r-1} + \left(\frac{H}{q}\right)^{r-2} \sum_{i=1}^{r-1} \left(S_0(q, a_{i+1} - a_i, H) + f(a_{i+1} - a_i; q) \right) \\ &\quad + \left(\frac{H}{q}\right)^{r-2} \frac{(\log H)^{-\frac{1}{2}}}{K\sqrt{\pi}(k-1)} \sum_{\substack{1 \leq i, j \leq r \\ j > i+1}} \delta(a_j \equiv a_i) + O(H^{r-2} (\log H)^{-\frac{3}{2}}). \end{aligned}$$

Let us now study $\mathcal{D}_1(\mathbf{a}, x)$. We first write

$$(53) \quad \mathcal{D}_1(\mathbf{a}, x) = -\frac{K}{\alpha(x)\sqrt{\log x}} \sum_{\substack{h_2, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \left(\sum_{\substack{2 \leq j \leq r \\ 2 \leq i \leq j}} \sum_{t=1}^{h_j-1} \mathfrak{S}_0(\{0, h_i + \dots + h_{j-1} + t\}) e^{-(h_2 + \dots + h_r)/H} \right. \\ \left. + \sum_{\substack{2 \leq j \leq r \\ j \leq i \leq r}} \sum_{t=1}^{h_j-1} \mathfrak{S}_0(\{h_j + \dots + h_i, t\}) e^{-(h_2 + \dots + h_r)/H} \right)$$

We focus on the first inner sum of (53). Exchanging the order of summation, for each fixed i and $j = i + k \geq i$, we have

$$\sum_{\substack{h_i, \dots, h_{i+k} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \sum_{t=1}^{h_{i+k}-1} \mathfrak{S}_0(\{0, h_i + \dots + h_{i+k-1} + t\}) e^{-(h_i + \dots + h_{i+k})/H} \\ \times \sum_{\substack{h_2, \dots, h_{i-1}, h_{i+k+1}, \dots, h_r > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} e^{-(h_2 + \dots + h_{i-1} + h_{i+k+1} + \dots + h_r)/H}$$

The second sum of the above is evaluated by (51), and

$$\sum_{\substack{h_i, \dots, h_{i+k} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \sum_{t=1}^{h_{i+k}-1} \mathfrak{S}_0(\{0, h_i + \dots + h_{i+k-1} + t\}) e^{-(h_i + \dots + h_{i+k})/H} \\ = \sum_{u>0} \mathfrak{S}_0(\{0, u\}) \sum_{\substack{h>u \\ h \equiv a_{i+k} - a_{i-1} \pmod{q}}} e^{-h/H} \sum_{\substack{h_i, \dots, h_{i+k-1} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q} \\ h_i + \dots + h_{i+k-1} < u}} \sum_{\substack{h_{i+k} > 0 \\ h_i + \dots + h_{i+k} = h}} 1 \\ = \sum_{u>0} \mathfrak{S}_0(\{0, u\}) e^{-u/H} \sum_{\substack{h'>0 \\ h' \equiv a_{i+k} - a_{i-1} - u \pmod{q}}} e^{-h'/H} \sum_{\substack{h_i, \dots, h_{i+k-1} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q} \\ h_i + \dots + h_{i+k-1} < u}} 1 \\ = \sum_{u>0} \mathfrak{S}_0(\{0, u\}) e^{-u/H} \left(\frac{1}{k!} \left(\frac{u}{q} \right)^k + O(u^{k-1}) \right) \left(\frac{H}{q} + O(1) \right) \\ = \frac{H}{k! q^{k+1}} S_0^{(k)}(H) + O(H^{k+\varepsilon})$$

We get a similar estimate for the second inner sum of (53) involving $\mathfrak{S}_0(\{h_j + \dots + h_i, t\})$ by making a change of variable to replace it by $\mathfrak{S}_0(\{0, r + h_{j+1} + \dots + h_i\})$ with $r = h_j - t$, and we obtain

$$\mathcal{D}_1(\mathbf{a}, x) = -2 \frac{K}{\alpha(x)\sqrt{\log x}} \sum_{i=2}^r \sum_{k=0}^{r-i} \left(\frac{H}{q} \right)^{r-2-k} \frac{H}{q^{k+1} k!} S_0^{(k)}(H) + O(H^{r-3+\varepsilon}) \\ = -2 \frac{K}{\alpha(x)\sqrt{\log x}} \frac{H^{r-1}}{q^{r-1}} \left((r-1) S_0(H) - \frac{(\log H)^{-\frac{1}{2}}}{K\sqrt{\pi}} \sum_{k=1}^{r-2} \frac{(r-1-k)}{k} \right) + O(H^{r-2} (\log H)^{-\frac{3}{2}})$$

The same ideas are used to estimate $\mathcal{D}_2(\mathbf{a}, x)$. Let i and $j = i + k \geq i$ be fixed, and let us study the sum in $\mathcal{D}_2(\mathbf{a}, x)$. In the case $i = j$, this is

$$\sum_{\substack{h_i > 0 \\ h_i \equiv a_i - a_{i-1} \pmod{q}}} \sum_{1 \leq t_1 < t_2 \leq h_i - 1} \mathfrak{S}_0(\{t_1, t_2\}) e^{-h_i/H} = \left(\frac{H^2}{q} + O(H) \right) S_0(H)$$

as we already saw in the case $r = 2$. In the case $k \geq 1$, we have

$$\begin{aligned}
& \sum_{\substack{h_i, \dots, h_{i+k} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q}}} \sum_{t_1=1}^{h_{i-1} h_{i+k} - 1} \sum_{t_2=1}^{h_{i+k} - 1} \mathfrak{S}_0(\{t_1, h_i + \dots + h_{i+k-1} + t_2\}) e^{-(h_i + \dots + h_{i+k})/H} \\
&= \sum_{1 \leq t_1 < t_2} \mathfrak{S}_0(\{0, t_2' - t_1\}) \sum_{h > t_2} e^{-h/H} \sum_{\substack{h_i, \dots, h_{i+k-1} > 0 \\ h_\ell \equiv a_\ell - a_{\ell-1} \pmod{q} \\ t_1 < h_i + \dots + h_{i+k-1} < t_2'}} \sum_{\substack{h_{i+k} > 0 \\ h_i + \dots + h_{i+k} = h}} 1 \\
&= \sum_{u > 0} \mathfrak{S}_0(\{0, u\}) \sum_{t_2' > u} \sum_{h > t_2'} e^{-h/H} \left(\frac{1}{k!} \left(\frac{u}{q} \right)^k + O(u^{k-1}) \right) \\
&= \sum_{u > 0} \mathfrak{S}_0(\{0, u\}) e^{-u/H} \left(\frac{1}{k!} \left(\frac{u}{q} \right)^k + O(u^{k-1}) \right) \left(\frac{H^2}{q} + O(H) \right) \\
&= \frac{H^2}{k! q^{k+1}} S_0^{(k)}(H) + O(H^{k+1+\varepsilon}).
\end{aligned}$$

We deduce that

$$\begin{aligned}
\mathcal{D}_2(\mathbf{a}, x) &= \frac{K^2}{\alpha(x)^2 \log x} \sum_{i=2}^r \sum_{k=0}^{r-i} \left(\frac{H}{q} \right)^{r-2-k} \frac{H^2}{k! q^{k+1}} S_0^{(k)}(H) + O(H^{r-3+\varepsilon}) \\
&= \frac{K^2}{\alpha(x)^2 \log x} \frac{H^r}{q^{r-1}} \left((r-1) S_0(H) - \frac{(\log H)^{-\frac{1}{2}}}{K \sqrt{\pi}} \sum_{k=1}^{r-2} \frac{(r-1-k)}{k} \right) + O(H^{r-2} (\log H)^{-\frac{3}{2}}).
\end{aligned}$$

Wrapping up, we obtain

$$\begin{aligned}
N(x; q, \mathbf{a}) &= \frac{x}{q} \left(\frac{K}{\sqrt{\log x}} \right)^r \alpha(x)^{-r+1} \left(\left(\frac{H}{q} \right)^{r-1} + \left(\frac{H}{q} \right)^{r-2} \sum_{i=1}^{r-1} (\mathcal{D}_0(a_i, a_{i+1}, x) - \frac{H}{q} + \mathcal{D}_1(a_i, a_{i+1}, x) + \mathcal{D}_2(a_i, a_{i+1}, x)) \right. \\
&\quad \left. - \left(\frac{H}{q} \right)^{r-2} \frac{(\log H)^{-\frac{1}{2}}}{K \sqrt{\pi}} \left(\sum_{k=1}^{r-2} \sum_{i=1}^{r-1-k} \frac{\delta(a_{i+k+1} \equiv a_i) - \frac{1}{q}}{k} \right) + O(H^{r-2} (\log H)^{-\frac{3}{2}}) \right).
\end{aligned}$$

Then using $H = \frac{\sqrt{\log x}}{K} - \frac{1}{2} + O((\log x)^{-\frac{1}{2}})$, and the estimates for $\mathcal{D}_i(a, b, x)$, $i = 0, 1, 2$ from Theorem 1.2 we obtain Conjecture 1.4.

9. NUMERICAL DATA

We present in this section some numerical data testing the approximation of Conjecture 1.1 for $N(x; q, (a, b))$. One of the challenges of the numerical testing is the change of scale introduced by the change of variable (17), which gives $H = \sqrt{\log x}/K$. The actual value of $N(x; q, (a, b))$ were obtained by using SageMath [S⁺09] on about 20 CPU cores in a Linux cluster for a couple of months, which allow us to take $x = 10^{12}$. But then, $H \approx 6.356$ in Theorem 1.2, which is very small even for this large value of x .

There are some technical methods for computing the Euler products, whenever they converge, and their derivatives with enough precision, and we used the following equality, which gives us a faster convergence:

$$\prod_{p \equiv 3 \pmod{4}} (1 - p^{-2s}) = \prod_{1 \leq j \leq J} \left(\frac{L(2^j s, \chi_4)}{\zeta(2^j s) (1 - 2^{-2^j s})} \right)^{1/2^j} \prod_{p \equiv 3 \pmod{4}} (1 - p^{-2^{J+1} s})^{1/2^J}.$$

Note that the rightmost hand side product converges much faster than the left hand side one. Also, its derivatives can be computed by taking the derivatives of the right hand side instead so that one might obtain some recursive formula.

We present in Table 5 some numerical data for Conjecture 1.1, for $q = 5$ and $x = 10^{12}$. There are 25 cases for $N(x; q, (a, b))$ in Table 5, but the conjectural asymptotic of Conjecture 1.1 only depends on $b - a \pmod{q}$, and there are then unavoidable fluctuations in the data for various pairs (a, b) with the same value of $b - a \pmod{5}$. The fit between the numerical data and the conjecture is slightly better

when $b - a \not\equiv 0 \pmod{5}$. The numerical data is also influenced by the bias of Theorem 2.4, which is of smaller magnitude than the bias of Conjecture 1.1 but in the opposite direction, and the data when $a = b = 0$ in particular shows the influence of both biases. We have used several asymptotic approximations of our conjecture in Table 5. We used Conjecture 1.1 as such with $J = 1$ (the column labelled “Conjecture 1.1”), and we also used the more complicated expression of Proposition 4.2 for $\mathcal{D}_0(a, b; x) + \mathcal{D}_1(a, b; x) + \mathcal{D}_2(a, b; x)$ in (18), where we evaluate the exponential sums $E(q, v; H)$ exactly for each residue class (recall that $H = \sqrt{\log x}/K \approx 6.356$ when $x = 10^{12}$). We then replaced $S_0(q, v; H)$ in that expression by the approximation of Theorem 1.2 with $J = 1$ (the column labelled “Theorem 1.2”), and by the actual numerical value of $S_0(q, v; H)$ (the column labelled “ $S_0(q, v; H)$ ”).

We also present some numerical data for Theorem 1.2 in Table 6 and 7 for larger values of H . We tested the asymptotic of Theorem 1.2 for $J = 1, 2, 3$ and the integral formula of Proposition 4.4 for various values of H . For $H \approx 6.356$, larger values of J or the integral formula of Proposition 4.4 are not approximating well $S(q, v; H)$, but one can see the fit for larger values of H . The values of the constants $c_0(2), c_0(3), c_1(2), c_1(3)$ can be computed by taking more terms in the Taylor expansions of the proof of Theorem 1.2, similarly to the computations of $c_0(1), c(1)$ in Section 5. We did not include those computations (which are lengthy but straightforward and not very interesting) in the paper. The numerical values are

$$\begin{aligned} c_0(1) &\approx 0.604541230, & c_0(2) &\approx 0.696827721, & c_0(3) &\approx 1.185903185 \\ c_1(1) &\approx -0.167588374, & c_1(2) &\approx -0.054190676, & c_1(3) &\approx -0.328019051. \end{aligned}$$

a	b	$N(x; q, (a, b))$	$S_0(q, v; H)$	Theorem 1.2	Conjecture 1.1	Error1	Error2	Error3
0	0	$4\,108 \cdot 10^6$	$3\,585 \cdot 10^6$	$3\,219 \cdot 10^6$	$3\,919 \cdot 10^6$	1.1461	1.2763	1.0483
	1	$7\,153 \cdot 10^6$	$6\,949 \cdot 10^6$	$6\,904 \cdot 10^6$	$6\,841 \cdot 10^6$	1.0294	1.0360	1.0457
	2	$5\,604 \cdot 10^6$	$5\,430 \cdot 10^6$	$5\,493 \cdot 10^6$	$5\,426 \cdot 10^6$	1.0320	1.0203	1.0329
	3	$8\,055 \cdot 10^6$	$7\,487 \cdot 10^6$	$7\,858 \cdot 10^6$	$7\,153 \cdot 10^6$	1.0759	1.0250	1.1261
	4	$5\,780 \cdot 10^6$	$5\,626 \cdot 10^6$	$5\,603 \cdot 10^6$	$5\,738 \cdot 10^6$	1.0274	1.0317	1.0073
1	0	$5\,777 \cdot 10^6$	$5\,626 \cdot 10^6$	$5\,603 \cdot 10^6$	$5\,738 \cdot 10^6$	1.0269	1.0312	1.0068
	1	$3\,765 \cdot 10^6$	$3\,585 \cdot 10^6$	$3\,219 \cdot 10^6$	$3\,919 \cdot 10^6$	1.0503	1.1697	0.9607
	2	$6\,870 \cdot 10^6$	$6\,949 \cdot 10^6$	$6\,904 \cdot 10^6$	$6\,841 \cdot 10^6$	0.9886	0.9950	1.0043
	3	$5\,354 \cdot 10^6$	$5\,430 \cdot 10^6$	$5\,493 \cdot 10^6$	$5\,426 \cdot 10^6$	0.9860	0.9747	0.9868
	4	$7\,742 \cdot 10^6$	$7\,487 \cdot 10^6$	$7\,858 \cdot 10^6$	$7\,153 \cdot 10^6$	1.0341	0.9853	1.0824
2	0	$8\,050 \cdot 10^6$	$7\,487 \cdot 10^6$	$7\,858 \cdot 10^6$	$7\,153 \cdot 10^6$	1.0752	1.0244	1.1254
	1	$5\,516 \cdot 10^6$	$5\,626 \cdot 10^6$	$5\,603 \cdot 10^6$	$5\,738 \cdot 10^6$	0.9804	0.9845	0.9613
	2	$3\,755 \cdot 10^6$	$3\,585 \cdot 10^6$	$3\,219 \cdot 10^6$	$3\,919 \cdot 10^6$	1.0474	1.1664	0.9580
	3	$6\,838 \cdot 10^6$	$6\,949 \cdot 10^6$	$6\,904 \cdot 10^6$	$6\,841 \cdot 10^6$	0.9840	0.9903	0.9996
	4	$5\,351 \cdot 10^6$	$5\,430 \cdot 10^6$	$5\,493 \cdot 10^6$	$5\,426 \cdot 10^6$	0.9853	0.9741	0.9861
3	0	$5\,609 \cdot 10^6$	$5\,430 \cdot 10^6$	$5\,493 \cdot 10^6$	$5\,426 \cdot 10^6$	1.0330	1.0212	1.0338
	1	$7\,718 \cdot 10^6$	$7\,487 \cdot 10^6$	$7\,858 \cdot 10^6$	$7\,153 \cdot 10^6$	1.0309	0.9822	1.0790
	2	$5\,549 \cdot 10^6$	$5\,626 \cdot 10^6$	$5\,603 \cdot 10^6$	$5\,738 \cdot 10^6$	0.9863	0.9904	0.9670
	3	$3\,765 \cdot 10^6$	$3\,585 \cdot 10^6$	$3\,219 \cdot 10^6$	$3\,919 \cdot 10^6$	1.0503	1.1697	0.9607
	4	$6\,867 \cdot 10^6$	$6\,949 \cdot 10^6$	$6\,904 \cdot 10^6$	$6\,841 \cdot 10^6$	0.9882	0.9946	1.0039
4	0	$7\,156 \cdot 10^6$	$6\,949 \cdot 10^6$	$6\,904 \cdot 10^6$	$6\,841 \cdot 10^6$	1.0298	1.0364	1.0461
	1	$5\,357 \cdot 10^6$	$5\,430 \cdot 10^6$	$5\,493 \cdot 10^6$	$5\,426 \cdot 10^6$	0.9864	0.9752	0.9872
	2	$7\,731 \cdot 10^6$	$7\,487 \cdot 10^6$	$7\,858 \cdot 10^6$	$7\,153 \cdot 10^6$	1.0326	0.9838	1.0808
	3	$5\,497 \cdot 10^6$	$5\,626 \cdot 10^6$	$5\,603 \cdot 10^6$	$5\,738 \cdot 10^6$	0.9771	0.9812	0.9580
	4	$3\,769 \cdot 10^6$	$3\,585 \cdot 10^6$	$3\,219 \cdot 10^6$	$3\,919 \cdot 10^6$	1.0512	1.1707	0.9615

TABLE 5. The experimental value of $N(x; q, (a, b))$ versus several estimates for Conjecture 1.1 with $J = 1$ for $q = 5$ and $x = 10^{12}$. We used Conjecture 1.1 as such with $J = 1$ (the column labelled “Conjecture 1.1”), and we also used the more complicated expression of Proposition 4.2 for $\mathcal{D}_0(a, b; x) + \mathcal{D}_1(a, b; x) + \mathcal{D}_2(a, b; x)$ in (18), where we evaluate the exponential sums $E(q, v; H)$ exactly for each residue class (recall that $H = \sqrt{\log x}/K \approx 6.356$ when $x = 10^{12}$). We then replaced $S_0(q, v; H)$ in that expression by the approximation of Theorem 1.2 with $J = 1$ (the column labelled “Theorem 1.2”), and by the actual numerical value of $S_0(q, v; H)$ (the column labelled “ $S_0(q, v; H)$ ”). Error1, Error2, Error3 are the percentage errors for the 4th, 5th and 6th columns, respectively.

H	$S(q, 0; H) - H/q$	Prop. 4.4	$J = 1$	$J = 2$	$J = 3$	Prop. 4.4	$J = 1$	$J = 2$	$J = 3$
6.356	-0.6093	-0.0087	-0.6889	-0.4122	-0.1577	70.3362	0.8843	1.4779	3.8630
16	-0.8852	-0.5540	-1.0240	-0.8731	-0.7804	1.5980	0.8645	1.0139	1.1343
10^2	-1.3968	1.2847	-1.5059	-1.4354	-1.4094	1.0862	0.9275	0.9731	0.9910
10^4	-2.2932	-2.2839	-2.3289	-2.3040	-2.2994	1.0041	0.9846	0.9953	0.9973
10^6	-2.9169	-2.9162	-2.9337	-2.9201	-2.9184	1.0002	0.9943	0.9989	0.9995

TABLE 6. The numerical value of $S(q, 0; H) - H/q$ for $q = 5$ and various values of H versus the asymptotic of Proposition 4.4 and Theorem 1.2 for $J = 1, 2, 3$. The last 4 columns are the percentage errors.

H	$S(q, 3; H) - H/q$	Prop. 4.4	$J = 1$	$J = 2$	$J = 3$	Prop. 4.4	$J = 1$	$J = 2$	$J = 3$
6.356	0.0327	0.0728	0.0811	0.0596	-0.0108	0.4485	0.4029	0.5485	-3.0166
16	0.0788	0.0919	0.1036	0.0919	0.0663	0.8575	0.7609	0.8581	1.1900
10^2	0.1120	0.1171	0.1262	0.1207	0.1135	0.9565	0.8875	0.9278	0.9868
10^4	0.1456	0.1461	0.1490	0.1471	0.1458	0.9966	0.9770	0.9899	0.9986
10^6	0.15813	0.15819	0.1592	0.1581	0.1577	0.9997	0.9935	1.0001	1.0030

TABLE 7. The numerical value of $S(q, 3; H) - H/q$ for $q = 5$ and various values of H versus the asymptotic of Proposition 4.4 and Theorem 1.2 for $J = 1, 2, 3$. The last 4 columns are the percentage errors.

REFERENCES

- [Ary15a] F. Aryan, *The distribution of k -tuples of reduced residues*. *Mathematika* 61 (2015), no. 1, 72–88.
- [Ary15b] F. Aryan, *Distribution of squares modulo a composite number*. *Int. Math. Res. Not. IMRN* (2015), no. 23, 12405–12431
- [BBF18] E. Bank, L. Bary-Soroker, and A. Fehm, *Sums of two squares in short intervals in polynomial rings over finite fields*. *Am. J. Math.* 140 (2018), no. 4, 1113–1131.
- [BF19] L. Bary-Soroker and A. Fehm, *Correlations of sums of two squares and other arithmetic functions in function fields*. *Int. Math. Res. Not.*, Volume 2019, Issue 14 (2019), 4469–4515.
- [BFT15] W. D. Banks, T. Freiberg and C. L. Turnage-Butterbaugh, *Consecutive primes in tuples*. *Acta Arith.* 167 (2015), no. 3, 261–266.
- [BSW16] L. Bary-Soroker, Y. Smilansky, A. Wolf, *On the Function Field Analogue of Landau’s Theorem on Sums of Squares*. *Finite Fields Appl.* 39 (2016), 195–215.
- [BW00] A. Balog, and T. D. Wooley, Trevor, *Sums of two squares in short intervals*. *Canad. J. Math.* 52 (2000), no. 4, 673–694.
- [CK97] R. D. Connors and J. P. Keating, *Two-Point Spectral Correlations for the Square Billiard*. *J. Phys. A* 30 (1997), no. 6, 1817–1830.
- [CLM21] A. Ciolan, A. Languasco and P. Moree, *Landau and Ramanujan approximations for divisor sums and coefficients of cusp forms*. (2021), in preparation.
- [CM99] J. Cazarán and P. Moree, *On a claim of Ramanujan in his first letter to Hardy*. *Exposition. Math.* 17 (1999), no. 4, 289–311.
- [EG18] I. Eriksson and L. Gustafsson, *On the Asymptotic Behaviour of Sums of Two Squares*, Degree project, KTH Royal Institute of Technology, School of Engineering Sciences (2018), 31 pp.
- [FKR17] T. Freiberg, P. Kurlberg and L. Rosenzweig, *Poisson distribution for gaps between sums of two squares and level spacings for toral point scatterers*. *Commun. Number Theory Phys.* 11 (2017), no. 4, 837–877.
- [Gal76] P. X. Gallagher, *On the distribution of primes in short intervals*. *Mathematika* 23 (1976), no. 1, 4–9.
- [GR20] O. Gorodetsky and B. Rodgers, *The variance of the number of sums of two squares in $\mathbb{F}_q[T]$ in short intervals*. *Journal of the AMS*, to appear. <https://arxiv.org/pdf/1810.06002.pdf>.
- [Gra89] A. Granville, *Least primes in arithmetic progressions*. *Théorie des nombres*, Proceedings of the International Conference held at the Université Laval, Quebec, Quebec, July 5–18, 1987, Edited by Jean-Marie De Koninck and Claude Levesque. Walter de Gruyter & Co., Berlin, (1989), 306–321.
- [Hoo71] C. Hooley, *On the intervals between numbers that are sums of two squares*. *Acta Math.* 127 (1971), 279–297.
- [Hoo73] C. Hooley, *On the intervals between numbers that are sums of two squares. II*. *J. Number Theory* 5 (1973), 215–217.
- [Kou19] D. Koukoulopoulos. *The distribution of prime numbers*. Graduate Studies in Mathematics, 203. American Mathematical Society, Providence, RI, [2019], xii + 356 pp.
- [Lan08] E. Landau *Über die einteilung der positiven ganzen zahlen in vier klassen nach der mindestzahl der zu ihrer additiven zusammensetzung erforderlichen quadrate*. *Arch. Math u. Phys.* (3) 13 (1908), 305–312.
- [LOS16] R. J. Lemke Oliver and K. Soundararajan, *Unexpected biases in the distribution of consecutive primes*. *Proc. Natl. Acad. Sci. USA* 113 (2016), no. 31, E4446–E4454.
- [Luc13] Lucia (<https://mathoverflow.net/users/38624/lucia>), Asymptotic density of k -almost primes, URL (version: 2013-10-31): <https://mathoverflow.net/q/146469>.
- [Mai85] H. Maier, *Primes in short intervals*. *Michigan Math. J.* 32 (1985), no. 2, 221–225.
- [MATH] Mathematica, Version 12.0, Wolfram Research Inc., 2020, <https://www.wolfram.com/mathematica/>.
- [May16] J. Maynard. *Dense clusters of primes in subsets*. *Compos. Math.* 152 (2016), no. 7, 1517–1554.
- [Men17] I. Mennema, *The Distribution of Consecutive Square-Free Numbers*, Master Thesis, Department of Mathematics Leiden University (2017), 31 pp.
- [Mir49] L. Mirsky, *On the frequency of pairs of square-free numbers with a given difference*. *Bull. Amer. Math. Soc.* 55 (1949), 936–939.

- [MS04] H. L. Montgomery and K. Soundararajan, *Primes in Short Intervals*. Comm. Math. Phys. 252 (2004), no. 1–3, 589–617.
- [MV84] H. L. Montgomery and R. C. Vaughan, *A basic inequality*. Proceedings of the Congress on Number Theory (Spanish) (Zarauz, 1984), 163–175, Univ. País Vasco-Euskal Herriko Unib., Bilbao, 1989.
- [MV86] H. L. Montgomery and R. C. Vaughan, *On the Distribution of Reduced Residues*. Ann. of Math. (2) 123 (1986), no. 2, 311–333.
- [PARI] The PARI Group, PARI/GP version 2.11.1, Univ. Bordeaux, 2019, <http://pari.math.u-bordeaux.fr/>.
- [Ram76] K. Ramachandra, *Some problems of analytic number theory*. Acta Arithmetica 31 (1976), 313–324.
- [Rie65] Rieger, G. J. *Über die Anzahl der als Summe von zwei Quadraten darstellbaren und in einer primen Restklasse gelegenen Zahlen unterhalb einer positiven Schranke. II.* (German) J. Reine Angew. Math. 217 (1965), 200–216.
- [S+09] *SageMath, the Sage Mathematics Software System (Version 8.6.rc0)*, The Sage Developers, 2019, <https://www.sagemath.org>.
- [Sha64] D. Shanks, *The Second-Order Term in the Asymptotic Expansion of $B(x)$* . Math. Comp. 18 (1964), 75–86.
- [Shi00] D. K. L. Shiu, *Strings of congruent primes*. J. London Math. Soc. (2) 61 (2000), no. 2, 359–373.
- [Smi13] Y. Smilansky. *Sums of two squares — pair correlation and distribution in short intervals*. Int. J. Number Theory 9 (2013), no. 7, 1687–1711.
- [Sta28] G. K. Stanley, *Two Assertions made by Ramanujan*. J. London Math. Soc. 3 (1928), no. 3, 232–237.
- [Sta29] G. K. Stanley, *Corrigenda: Two Assertions made by Ramanujan*. J. London Math. Soc. 4 (1929), no. 1, 32.
- [Ten15] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*. Third edition. Translated from the 2008 French edition by Patrick D. F. Ion. Graduate Studies in Mathematics, 163. American Mathematical Society, Providence, RI, 2015. xxiv+629 pp.

DEPARTMENT OF MATHEMATICS & STATISTICS, CONCORDIA UNIVERSITY, 1455 DE MAISONNEUVE BLVD. WEST, MONTRÉAL, QUÉBEC, H3G 1M8, CANADA
Email address: chantal.david@concordia.ca

DEPARTMENT OF MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND THE UNIVERSITY OF GOTHENBURG, SE-412 96 GOTHENBURG, SWEDEN
Email address: devin@chalmers.se

DEPARTMENT OF MATHEMATICS & STATISTICS, CONCORDIA UNIVERSITY, 1455 DE MAISONNEUVE BLVD. WEST, MONTRÉAL, QUÉBEC, H3G 1M8, CANADA
Email address: jungbae.nam@concordia.ca

DEPARTMENT OF MATHEMATICS & STATISTICS, CONCORDIA UNIVERSITY, 1455 DE MAISONNEUVE BLVD. WEST, MONTRÉAL, QUÉBEC, H3G 1M8, CANADA
Email address: jeremy.schlitt@mail.concordia.ca