SQUARES OF PRIMES AND POWERS OF 2, II

Jianya Liu, Ming-Chit Liu and Tao Zhan

Abstract. We prove that the density of integers $\equiv 2(\mathrm{mod} 24)$, which can be represented as the sum of two squares of primes and k powers of 4, tends to 1 as $k \to \infty$ in the sequence $k \equiv 0(\mathrm{mod} 3)$. Consequently, there exists a positive integer k_0 such that every large integer $\equiv 4(\mathrm{mod} 24)$ is the sum of four squares of primes and k_0 powers of 4.

1991 Mathematics Subject Classification. 11P32, 11P05, 11P55.

1. Introduction

In this article we continue the study in our previous paper [LLZ] on the representation of even integers as the sum of squares of primes and powers of 2. The main result of [LLZ] includes the following

Theorem 0. *Let $\Lambda(m)$ be the von Mangoldt function, and*

$$\rho_k(N) = \sum_{N=m_1^2+...+m_4^2+2^{\nu_1}+...+2^{\nu_k}} \Lambda(m_1)\cdots\Lambda(m_4), \tag{1.1}$$

where m_j and ν_j denote positive integers. Then for $k \geq 4$ there exists a positive integer N_k depending on k only, such that for each $N \geq N_k$ with $N \equiv 4(\mathrm{mod} 8)$,

$$\rho_k(N) \gg N \log_2^k N \left\{1 + O\left(\frac{1}{k}\right)\right\}, \tag{1.2}$$

where the constants implied by the $\gg$ and O-symbols are absolute.

It therefore follows that there is an absolute positive integer k, such that each large even integer N can be represented as

$$N = p_1^2 + p_2^2 + p_3^2 + p_4^2 + 2^{\nu_1} + 2^{\nu_2} + ... + 2^{\nu_k}. \tag{1.3}$$

Here and throughout, p_j or p denotes a prime. Very recently, the first two authors [LL] have proved that $k = 8330$ is acceptable in (1.3). Theorem 0 is closely related to two well-known results in the additive theory of prime numbers: (a) The "almost Goldbach" theorem of Linnik-Gallagher [L1] [L2] [G] on the representation of even integers as the sum of two primes and a bounded number of powers of 2; (b) The theorem of Hua [H1] on the representation of $N \equiv 5(\mathrm{mod} 24)$ as the sum of five squares of primes. For brief history and recent results in these directions, the reader may refer to [LLZ].

The purpose of the present paper is to establish the asymptotic formula in our Theorem 2 instead of the lower bound in (1.2). To obtain this asymptotic formula, we need essentially our Theorem 1 below which is a parallel result to the following works on an Euler problem. Our Theorems 1 and 2 form an extension of Gallagher's results in [G] for the "almost Goldbach" problem.

In a letter to Goldbach, Euler asked, and later answered by himself negatively, the problem of representing each positive odd integer n as the sum of a prime and a power of 2, namely $n = p + 2^\nu$. Romanoff [Ro] showed in 1934 that a positive proportion of the positive odd integers can be written in this way. And Gallagher's result in [G] states that the density of odd integers which may be written in the form

$$n = p_1 + 2^{\nu_1} + ... + 2^{\nu_k} \tag{1.4}$$

tends to 1 as $k \to \infty$.

Analogous to (1.4), we shall consider the representation

$$n = p_1^2 + p_2^2 + 4^{\nu_1} + ... + 4^{\nu_k}, \tag{1.5}$$

and prove the following results.

Theorem 1. *Let*

$$r'_k(n) = \sum_{\substack{n=m_1^2+m_2^2+4^{\nu_1}+...+4^{\nu_k}\\ \nu_j\geq 2}} \Lambda(m_1)\Lambda(m_2). \tag{1.6}$$

Then for any positive integer $k \equiv 0(\bmod 3)$, *there is a* $N_k > 0$ *depending on* k *only, such that for each* $N \geq N_k$,

$$\sum_{\substack{n\leq N\\ n\equiv 2(\bmod 24)}} (r'_k(n) - 6\pi \log_4^k N)^2 \ll \frac{\log^2 k}{k} N \log_4^{2k} N. \tag{1.7}$$

It follows that the density of integers $n \equiv 2(\bmod 24)$, which may be written in the form (1.5), tends to 1 as $k \to \infty$ in the sequence $k \equiv 0(\text{mod}3)$. Now let

$$r''_k(N) = \sum_{\substack{N=m_1^2+...+m_4^2+4^{\nu_1}+...+4^{\nu_k}\\ \nu_j\geq 2}} \Lambda(m_1)\cdots\Lambda(m_4). \tag{1.8}$$

If we consider the number $r''_k(N)$ instead of $\rho_k(N)$ in (1.1), then we can get the following

Theorem 2. *For any positive integer* $k \equiv 0(\text{mod}3)$, *there is* $N_k > 0$ *depending on* k *only, such that for each integer* N *with* $N \equiv 4(\text{mod}24)$ *and* $N \geq N_k$,

$$r''_k(N) = \frac{3}{2}\pi^2 N \log_4^k N \left\{1 + O\left(\frac{\log^2 k}{k}\right)\right\}. \tag{1.9}$$

Thus, each sufficiently large integer N with $N \equiv 4(\text{mod}24)$ can be written as

$$N = p_1^2 + p_2^2 + p_3^2 + p_4^2 + 4^{\nu_1} + 4^{\nu_2} + ... + 4^{\nu_k}. \tag{1.10}$$

One sees that (1.9) forms a more desirable result than (1.2).

2. Outline of the method

The estimate in (1.7) will be an indispensable tool in the establishment of the asymptotic formula in (1.9). In the proof of our Theorem 1, we have to adapt Gallagher's method (Proof of Theorem 1 in [G], p.139) which is a modification of the well-known Linnik dispersion method. One of the vital steps in the method is the cancellation of all main terms of the three sums in (5.9). Therefore the right numerical value of the coefficient of each main term becomes very important in the cancellation. This delicate part will be achieved by Lemmas 5.2 and 5.3. In particular, the O-term in (5.1) will give the required bound in Theorem 1 after the cancellation.

Lemma 5.2 will take care of the right numerical value $3/2(= (3 \times 8)/16)$ in the first sum of (5.9) (see the inequality next to (5.9)), which comes from Proposition 2.1 (for the 16) and Lemma 3.1(1)(2) (for the 24) via Lemma 4.5. It is interesting to note that Lemma 3.1 implies the requirement $n \equiv 0(\bmod 3 \times 8)$ and that, as a consequence, we need the 4-adic expression in (1.10) with all $\nu_j \geq 2$. This can be seen in the $4^\nu \equiv 16(\text{mod}24)$ for all $\nu \geq 2$ appearing in the beginning of the proof of Lemma 4.5.

Lemma 5.3 will provide with the right numerical value $1/4$ in the second sum of (5.9). It is interesting to note that the log factor $(\log \sqrt{N})^2$ in the formula between (5.6) and (5.7) gives exactly the constant $1/4$, and that in the argument below (5.8) if $3|k$ then $16k \equiv 0(\mathrm{mod} 24)$, and then we get through.

For the proof of Lemma 5.2 we need the circle method to obtain Proposition 2.1 below and some sieve methods to prove Lemma 5.1. These two results will be used to deduce (5.3) and (5.5) respectively for the establishment of (5.1).

Let $e(\alpha) = \exp(i2\pi\alpha)$ and for large integer $N > 0$ let

$$T(\alpha) = \sum_{m^2 \le N} \Lambda(m) e(m^2\alpha), \qquad G(\alpha) = \sum_{4^2 \le 4^\nu \le N} e(4^\nu \alpha). \tag{2.1}$$

Since

$$T^2(\alpha) G^k(\alpha) = \sum_{n \le N} r_k'(n) + \sum_{n > N},$$

we have

$$\sum_{n \le N} (r_k'(n))^2 \le \int_0^1 |T^2(\alpha) G^k(\alpha)| d\alpha. \tag{2.2}$$

Now it suffices to get the required upper bound as in Theorem 1 for this integral.

In order to apply the circle method, we set

$$P = N^\theta, \qquad Q = N^{1-2\theta}, \tag{2.3}$$

where θ is a constant satisfying $0 < \theta \le 1/25$. By Dirichlet's lemma on rational approximations, each $\alpha \in [1/Q, 1 + 1/Q]$ may be written in the form

$$\alpha = \frac{a}{q} + \lambda, \qquad |\lambda| \le \frac{1}{qQ} \tag{2.4}$$

for some integers a, q with $1 \le a \le q \le Q$ and $(a, q) = 1$. We denote by $\mathfrak{M}(a, q)$ the set of α satisfying (2.4), and define the major arcs $\mathfrak{M}$ and the minor arcs $C(\mathfrak{M})$ as follows:

$$\mathfrak{M} = \bigcup_{q \le P} \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathfrak{M}(a, q), \qquad C(\mathfrak{M}) = \left[\frac{1}{Q}, 1 + \frac{1}{Q}\right] \backslash \mathfrak{M}. \tag{2.5}$$

It follows from $2P \le Q$ that the major arcs $\mathfrak{M}(a, q)$ are mutually disjoint.

Now the integral in (2.2) takes the form

$$\left\{ \int_{\mathfrak{M}} + \int_{C(\mathfrak{M})} \right\} |T^2(\alpha) G^k(\alpha)|^2 d\alpha. \tag{2.6}$$

Since in our Lemma 5.2 which is essential to the proof of Theorem 1, we obtain the upper bound in (5.1), we need to add absolute value signs in (2.6). This is the only difference between (2.6) and [LLZ,(2.5)]. So following the same arguments as in [LLZ,§3-§5] we can obtain the following Proposition 2.1 which is parallel to [LLZ,Proposition 2.1]. The absolute value signs in (2.7) evoke the integral (2.8) in (2.7) instead of an additional π factor as in [LLZ,(2.6)].

Proposition 2.1. *Let $1 \le n \le N$. Then for $0 < \theta \le 1/25$ in (2.3), we have*

$$\int_{\mathfrak{M}} |T(\alpha)|^4 e(n\alpha) d\alpha = \frac{\pi}{16} \mathfrak{S}(n) \mathfrak{I}\left(\frac{n}{N}\right) N + O\left(\frac{N}{\log N}\right), \tag{2.7}$$

where $\mathfrak{S}(n)$ *is the singular series defined as in (3.2), and satisfies* $\mathfrak{S}(n) \gg 1$ *for* $n \equiv 0(\mathrm{mod} 24)$. *And* $\mathfrak{I}(n/N)$ *is definded as*

$$\mathfrak{I}\left(\frac{n}{N}\right) = 2\int_{\max(0,-n/N)}^{\min(1,1-n/N)} v^{-1/2}\left(1-\frac{n}{N}-v\right)^{1/2} dv, \tag{2.8}$$

and satisfies $0 \le \mathfrak{I}(n/N) \le \pi$ *for* $|n| \le N^2$.

3. Estimates related to the singular series

We need some more notation. As usual, $\varphi(n)$ and $\mu(n)$, stand for the function of Euler and Möbius respectively, $d(n)$ the divisor function. We use χ mod q and χ^0 mod q to denote a Dirichlet character and the principal character modulo q. Define

$$C(\chi, a) = \sum_{h=1}^{q} \bar{\chi}(h) e\left(\frac{ah^2}{q}\right), \qquad C(q,a) = C(\chi^0, a). \tag{3.1}$$

If $\chi_1, ..., \chi_4$ are characters mod$q$, then we write

$$B(n,q,\chi_1,...,\chi_4) = \sum_{\substack{a=1\\(a,q)=1}}^{q} e\left(\frac{an}{q}\right) C(\chi_1,a)C(\chi_2,a)\overline{C(\chi_3,a)C(\chi_4,a)},$$

$$B(n,q) = B(n,q,\chi_1^0,...,\chi_4^0) = \sum_{\substack{a=1\\(a,q)=1}}^{q} |C(q,a)|^4 e\left(\frac{an}{q}\right),$$

and

$$A(n,q) = \frac{B(n,q)}{\varphi^4(q)}, \qquad \mathfrak{S}(n) = \sum_{q=1}^{\infty} A(n,q). \tag{3.2}$$

Lemma 3.1 *Let* $A(n,q)$ *be defined as in (3.2). Then*

(1) For $p = 2$, *one has*

$$1 + A(n,2) + A(n,2^2) + A(n,2^3) = \begin{cases} 0, & \text{if } \; n \not\equiv 0(\mathrm{mod} 8), \\ 8, & \text{if } \; n \equiv 0(\mathrm{mod} 8). \end{cases}$$

(2) For $p = 3$, *one has*

$$1 + A(n,3) = \begin{cases} 0, & \text{if } \; n \not\equiv 0(\mathrm{mod} 3), \\ 3, & \text{if } \; n \equiv 0(\mathrm{mod} 3). \end{cases}$$

(3) For $p \ge 5$, *one has*

$$1 + A(n,p) > 0.$$

Proof. The proof is similar to that of Lemmas 13 and 14 in Hua [H1], so we only prove part (1). By the method of [H1], Lemmas 13, the quantity $1 + A(n,2) + A(n,2^2) + A(n,2^3)$ is equal to $2^3 M/\varphi^4(2^3)$, where M is the number of incongruent solutions of the equation

$$m_1^2 + m_2^2 - m_3^2 - m_4^2 \equiv n(\mathrm{mod} 8), \qquad 2 \nmid m_1 m_2 m_3 m_4. \tag{3.3}$$

Since $m^2 \equiv 1(\mathrm{mod} 8)$ for odd m, (3.3) has no solution unless $n \equiv 0(\mathrm{mod} 8)$. Clearly $M = 4^4$ if $n \equiv 0(\mathrm{mod} 8)$. This proves part (1) and hence the lemma.

Now we can give some properties of the singular series $\mathfrak{S}(n)$. Since these are not covered by [H2], Lemma 8.10, we state them as a proposition.

Proposition 3.2. *The singular series* $\mathfrak{S}(n)$ *is absolutely convergent. For* $n \equiv 0(\mathrm{mod} 24)$, *one has* $1 \ll \mathfrak{S}(n) \ll (\log\log n)^{11}$, *while for* $n \not\equiv 0(\mathrm{mod} 24)$, *one has* $\mathfrak{S}(n) = 0$.

Proof. This can be proved similarly as [LLZ,Proposition 4.3]. The only difference is that in [LLZ] $n \equiv 4$ or $n \not\equiv 4 \pmod{24}$ was assumed. This does not cause essential changes in the proof.

Lemma 3.3. *(1) Let*

$$\gamma = \begin{cases} 3, & \text{if } p = 2, \\ 1, & \text{if } p \geq 3. \end{cases}$$

If $t > \gamma$, then $C(p^t, a) = 0$, and consequently, $A(n, p^t) = 0$.

(2) For odd q,

$$|C(q, a)|^2 \leq \mu^2(q)\sigma(q),$$

where $\sigma(q)$ is the sum of all the divisors of q.

Proof. (1) The proof is similar to that of Hua [H1], Lemma 4.

(2) Let p be an odd prime. By (3.1) and Hua [H3,Theorem 7.5.4]

$$C(p, a) = \sum_{m=1}^{p} e(am^2/p) - 1 = \left(\frac{a}{p}\right) \sum_{m=1}^{p} e(m^2/p) - 1$$

where $(\frac{a}{p})$ is the Legendre symbol. Then apply [H3,Theorem 7.5.5] we get

$$|C(q, a)|^2 \leq p + 1.$$

If q is not square-free, then $|C(q, a)| = 0$ by part (1), and the lemma holds trivially. Now suppose q is square-free. It follows that

$$|C(q, a)|^2 = \prod_{p|q} |C(p, a)|^2 \leq \prod_{p|q} (p + 1) = \sigma(q).$$

The lemma is proved.

4. Lemmas concerning the sequence $4, 4^2, 4^3, ...$

Let $G(\alpha)$ be defined as in (2.1) and $L = \log_4 N$. The following Lemmas 4.1 and 4.2 are quoted from [G], Lemmas 2 and 3.

Lemma 4.1. *The set $\mathfrak{E}$ of $\alpha \in (0, 1]$ for which $|G(\alpha)| \geq (1 - \eta)L$ has measure $\ll N^{\Theta - 1}$, where $\Theta = c_1 \eta \log(e/\eta)$ and c_1 is a positive constant.*

For odd q, we denote by $\varrho(q)$ the least positive integer ϱ for which $4^\varrho \equiv 1 (\mathrm{mod} q)$.

Lemma 4.2. *If α is a rational number with denominator q satisfying $(q, 6) = 1$, and if $1 < \varrho(q) \leq L$, then*

$$|G(\alpha)| \leq \left(1 - \frac{c_2}{\varrho(q)}\right) L,$$

where c_2 is a positive constant.

Lemma 4.3. *Denote by $r_{k,k}(n)$ the number of representations of n in the form*

$$n = 4^{\nu_1} + ... + 4^{\nu_k} - (4^{\mu_1} + ... + 4^{\mu_k})$$

with $1 \leq \nu_i$, $\mu_j \leq L$. Then for $k \geq 3$,

$$r_{k,k}(0) \leq 2(k + 1)! L^{2k-3}, \tag{4.1}$$

and for odd q,

$$\sum_{q|n} r_{k,k}(n) \ll L^{2k-1} \left(1 + \frac{L}{\varrho(q)}\right). \tag{4.2}$$

Proof. The second estimate (4.2) is Lemma 5 of [G].

To prove (4.1), we note that $r_{k,k}(0)$ is the number of solutions of the equation $4^{\nu_1}+...+4^{\nu_k}=4^{\mu_1}+...+4^{\mu_k}$. Fixing $\mu_1,...,\mu_k$ arbitrarily, we have

$$r_{k,k}(0)\le L^k\max_{m\le kN} r_k(m),$$

where $r_k(m)$ is the number of solutions of

$$m=4^{\nu_1}+...+4^{\nu_k}. \tag{4.3}$$

Let $s_k(m)$ be the number of solutions of (4.3) with the restriction

$$\nu_1\le\nu_2\le...\le\nu_k.$$

Then

$$r_k(m)\le k!s_k(m).$$

Obviously, $s_k(m)=s_k^*(m)+s_k^{**}(m)$, where $s_k^*(m),s_k^{**}(m)$ denote respectively the number of solutions of (4.3) with

(*) $\nu_1<\nu_2<...<\nu_k$,

or

(**) there exists $1\le j\le k-1$, such that $\nu_j=\nu_{j+1}$.

Since there is only one 4-adic representation for m, we have $s_k^*(m)\le 1$. To bound $s_k^{**}(m)$, we note that $s_k^{**}(m)\le kt_k(m)$, where $t_k(m)$ is the number of solutions of

$$m=4^{\nu}+4^{\nu}+4^{\nu_3}+...+4^{\nu_k},\qquad \nu_3\le...\le\nu_k. \tag{4.4}$$

For $k\ge 4$, we fix $\nu,\nu_3,\nu_4,...,\nu_{k-2}$ in (4.4) arbitrarily, then we get $t_k(m)\le L^{k-3}$, on noting that there is at most 1 solution for $n=4^{\ell_1}+4^{\ell_2}$ with $\ell_1\le\ell_2$. In the special case $k=3$, the estimate $t_3(m)\le 2$ is obvious. Thus,

$$r_k(m)\le k!(1+2kL^{k-3})\le 2(k+1)!L^{k-3},$$

and (4.1) follows.

Lemma 4.4. *Let $\sigma(q)$ be defined as in Lemma 3.3. Then for $x\ge 2$ we have*

$$\sum_{\varrho(q)\le x}\frac{\mu^2(q)\sigma^2(q)}{\varphi^3(q)}\ll\log x,\qquad \sum_{\varrho(q)\le x}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}\ll\log x.$$

Proof. Let $X=\prod_{\varrho\le x}(4^{\varrho}-1)$. Then $q|X$ whenever $\varrho(q)\le x$, and clearly $2\nmid X, X\le 4^{x^2}$. It therefore follows that

$$\sum_{\substack{\varrho(q)\le x\\ 2\nmid q}}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}\le\sum_{\substack{q|X\\ 2\nmid q}}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}=\frac{1}{19}\prod_{p|2X}\frac{p}{(p-1)}\prod_{p}\left(1+\frac{5p^2-2p+1}{p(p-1)^3}\right).$$

The last infinite product is convergent. Thus,

$$\sum_{\substack{\varrho(q)\le x\\ 2\nmid q}}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}\ll\prod_{p|2X}\frac{p}{(p-1)}=\frac{2X}{\varphi(2X)}\ll\log\log(2X)\ll\log x.$$

This proves the second inequality in the lemma. The proof for the first one is similar.

The following lemma shows that although $\mathfrak{S}(n)\ll 1$ is not true in general, it is true on average.

Lemma 4.5. *Let $r_{k,k}(n)$ and $\mathfrak{S}(n)$ be as in Lemma 4.3 and (3.2) respectively. Then there exists $N_k > 0$ depending on k only, such that when $N \geq N_k$ we have*

$$\sum_{n \neq 0} r_{k,k}(n)\mathfrak{S}(n) = 24L^{2k}\left\{1 + O\left(\frac{\log^2 k}{k}\right)\right\}.$$

Proof. Note that for $n = 4^{\nu_1} + ... + 4^{\nu_k} - 4^{\mu_1} - ... - 4^{\mu_k}$ with $2 \leq \nu_j, \mu_j \leq L$, we always have $n \equiv 0(\bmod 24)$ as $4^{\nu} \equiv 16(\bmod 24)$ for $\nu \geq 2$. For such n, Lemma 3.1 and a minor modification of the argument as in [LLZ,(4.13) and the equality for $\mathfrak{S}(n)$ below (4.13)] give

$$\begin{aligned} \mathfrak{S}(n) &= \{1 + A(n,2) + A(n,2^2) + A(n,2^3)\}\{1 + A(n,3)\}\prod_{p \geq 5}\{1 + A(n,p)\} = 3 \times 8 \sum_{\substack{q=1 \\ (q,6)=1}}^{\infty} A(n,q) \\ &= 24 \sum_{\substack{q \leq R \\ (q,6)=1}} A(n,q) + O\{R^{-1+\varepsilon}d(n)\}, \end{aligned}$$

where $R \geq 1$ is a parameter to be specified later. We can also suppose $|n| \leq k4^L = kN$, since otherwise $r_{k,k}(n) = 0$. Thus, by (3.2) and (4.2) with $q = 1$,

$$\sum_{n \neq 0} r_{k,k}(n)\mathfrak{S}(n) = 24 \sum_{\substack{q \leq R \\ (q,6)=1}} \frac{1}{\varphi^4(q)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |C(q,a)|^4 \sum_{n \neq 0} r_{k,k}(n) e\left(\frac{an}{q}\right) + O(R^{-1+\varepsilon}L^{2k}(kN)^{\varepsilon}),$$

where we have used the estimate $d(n) \ll |n|^{\varepsilon} \ll (kN)^{\varepsilon}$ with arbitrarily small $\varepsilon > 0$. If there is the term corresponding to $n = 0$ in the above sum $\sum_{\substack{q \leq R \\ (q,6)=1}}$, this term would contribute

$$\ll \sum_{q \leq R} \frac{1}{\varphi^4(q)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |C(q,a)|^4 r_{k,k}(0) \ll (k+1)!L^{2k-3} \sum_{q \leq R} \frac{\mu^2(q)\sigma^2(q)}{\varphi^3(q)} \ll (k+1)!L^{2k-3}\log R$$

by (4.1), Lemma 3.3, and Lemma 4.4. With the term $n = 0$ put in and the inner sum simplified, one has

$$\begin{aligned} \sum_{n \neq 0} r_{k,k}(n)\mathfrak{S}(n) &= 24 \sum_{\substack{q \leq R \\ (q,6)=1}} \frac{1}{\varphi^4(q)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |C(q,a)|^4 \left|G\left(\frac{a}{q}\right)\right|^{2k} \\ &\quad + O((k+1)!L^{2k-3}\log R) + O(R^{-1+\varepsilon}L^{2k}(kN)^{\varepsilon}). \end{aligned} \tag{4.5}$$

In the double sum in (4.5) the term $q = 1$ contributes L^{2k}. We spilt the remaining sum according to the size of $\varrho(q)$:

$$\sum_{\substack{q \leq R \\ (q,6)=1}} \frac{1}{\varphi^4(q)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |C(q,a)|^4 \left|G\left(\frac{a}{q}\right)\right|^{2k} = L^{2k} + \left\{\sum_{\substack{2 \leq q \leq R \\ (q,6)=1 \\ \varrho(q) \leq E}} + \sum_{\substack{2 \leq q \leq R \\ (q,6)=1 \\ \varrho(q) > E}}\right\} \frac{1}{\varphi^4(q)} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} |C(q,a)|^4 \left|G\left(\frac{a}{q}\right)\right|^{2k}, \tag{4.6}$$

where E is a parameter satisfying $2 \leq E \leq L$. By Lemmas 4.2, 3.3(2) and 4.4, the first sum on the right-hand side of (4.6) can be estimated as

$$\sum_{\substack{2 \leq q \leq R \\ (q,6)=1 \\ \varrho(q) \leq E}} \ll L^{2k}\left(1 - \frac{c_2}{E}\right)^{2k} \sum_{\substack{2 \leq q \leq R \\ \varrho(q) \leq E}} \frac{\mu^2(q)\sigma^2(q)}{\varphi^3(q)} \ll L^{2k}\left(1 - \frac{c_2}{E}\right)^{2k} \log E.$$

To treat the last sum in (4.6), one appeals to Lemma 4.3, which gives

$$\sum_{\substack{a=1\\(a,q)=1}}^{q}\left|G\left(\frac{a}{q}\right)\right|^{2k}\le\sum_{a=1}^{q}\left|G\left(\frac{a}{q}\right)\right|^{2k}=q\sum_{q|n}r_{k,k}(n)\ll qL^{2k-1}\left(1+\frac{L}{\varrho(q)}\right),$$

and consequently, by Lemma 3.3(2),

$$\sum_{\substack{2\le q\le R\\(q,6)=1\\\varrho(q)>E}}\ll L^{2k}\sum_{\substack{q\le R\\\varrho(q)>E}}\frac{\mu^2(q)\sigma^2(q)}{\varphi^4(q)}\frac{q}{\varrho(q)}+L^{2k-1}\sum_{q\le R}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}.$$

The second sum on the right-hand side above is $\ll \log R$ by Lemma 4.4. Using partial summation and Lemma 4.4, the first sum is

$$\ll\sum_{m>E}\frac{1}{m}\sum_{\varrho(q)=m}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}=\int_E^\infty\frac{1}{t^2}\left(\sum_{\varrho(q)\le t}\frac{\mu^2(q)\sigma^2(q)q}{\varphi^4(q)}\right)dt\ll\int_E^\infty\frac{\log t}{t^2}dt\ll\frac{\log E}{E}.$$

Hence,

$$\sum_{\substack{q\le R\\(q,6)=1\\\varrho(q)>E}}\ll\frac{\log E}{E}L^{2k}+L^{2k-1}\log R.$$

Summing up the above estimates, we conclude that

$$\sum_{n\ne0}r_{k,k}(n)\mathfrak{S}(n)-24L^{2k}\ll L^{2k}\left\{\left(1-\frac{c_6}{E}\right)^{2k}\log E+\frac{\log E}{E}+\frac{\log R}{L}+R^{-1+\varepsilon}(kN)^{\varepsilon}\right\}.$$

Take $E=2c_6k/\log k$ and $R=N^{1/k}$, then the right-hand side is $\ll L^{2k}k^{-1}\log^2 k$ if $N\ge N_k$. This proves Lemma 4.5.

5. Proofs of Theorems 1 and 2

Lemma 5.1. *Let $T(\alpha)$ and $G(\alpha)$ be as in (2.1). Then*

$$\int_0^1|T(\alpha)G(\alpha)|^4d\alpha\ll NL^4.$$

Proof. This can be proved similarly as [LLZ,Lemma 6.1]. In the proof we need [LLZ,Proposition 2.2] which was obtained by a modification of the sieve methods of Brüdern and Fouvry [BF].

The theorems stated in §1 depend on the two mean-value estimates for $r_k'(n)$ given in the following Lemmas 5.2 and 5.3.

Lemma 5.2. *Let $r_k'(n)$ be as in Theorem 1. Then there is $N_k>0$ depending on k only, such that when $N\ge N_k$ we have*

$$\sum_{n\le N}(r_k'(n))^2\le\frac{3}{2}\pi^2NL^{2k}\left\{1+O\left(\frac{\log^2 k}{k}\right)\right\}.\tag{5.1}$$

Lemma 5.3. *Let $r_k'(n)$ be as in Theorem 1 with $3|k$. Then as $N\to\infty$,*

$$\sum_{\substack{n\le N\\n\equiv2(\bmod 24)}}r_k'(n)\sim\frac{1}{4}\pi NL^k.$$

Proof of Lemma 5.2. By (2.2) and (2.6), we have

$$\sum_{n\le N}(r'_k(n))^2 \le \int_0^1 |T^2(\alpha)G^k(\alpha)|^2 d\alpha = \int_{\mathfrak{M}} + \int_{C(\mathfrak{M})\cap\mathfrak{E}} + \int_{C(\mathfrak{M})\cap C(\mathfrak{E})} \tag{5.2}$$

where $\mathfrak{E}$ is defined as in Lemma 4.1 and $\mathfrak{M}$ as in (2.5) with θ satisfying $0 < \theta \le 1/25$.

Now we estimate the three integrals in (5.2) respectively. As given in [LLZ,between (6.9) and (6.10)] the number of solutions $Z_0(N)$ of the equation

$$p_1^2 + p_2^2 = p_3^2 + p_4^2$$

satisfies

$$Z_0(N) \ll N \log^{-2} N.$$

We have

$$\int_{\mathfrak{M}} = \sum_m r_{k,k}(m) \int_{\mathfrak{M}} |T(\alpha)|^4 e(m\alpha) d\alpha.$$

In the above formula one can suppose $|m| \le kN$, since otherwise $r_{k,k}(m) = 0$. Then by (2.1), we have

$$\int_0^1 |T(\alpha)|^4 d\alpha \ll Z_0(N) L^4 \ll NL^2.$$

By this and (4.1), the term $m = 0$ contributes

$$\ll r_{k,k}(0) \int_0^1 |T(\alpha)|^4 d\alpha \ll (k+1)! NL^{2k-1}.$$

And by Proposition 2.1 and (4.2) with $q = 1$, the other terms contribute

$$\sum_{m\ne 0} r_{k,k}(m) \left\{ \frac{\pi}{16} \mathfrak{S}(m) \mathfrak{I}\left(\frac{m}{N}\right) N + O\left(\frac{N}{\log N}\right)\right\} \le \frac{\pi^2}{16} N \sum_{m\ne 0} r_{k,k}(m)\mathfrak{S}(m) + O(NL^{2k-1}).$$

Applying Lemma 4.5 to the above sum $\sum\limits_{m\ne 0}$, we get, for $N \ge N_k$,

$$\int_{\mathfrak{M}} \le \frac{24}{16}\pi^2 NL^{2k} \left\{1 + O\left(\frac{\log^2 k}{k}\right)\right\}. \tag{5.3}$$

To estimate the second integral in (5.2), one notes that each $\alpha \in C(\mathfrak{M})$ can be written as (2.4) for some $P < q \le Q$ and $1 \le a \le q$ with $(q, a) = 1$. We now apply Theorem 2 of Ghosh [Gh], which states that, if $|\alpha - a/q| \le q^{-2}$, then

$$\sum_{m\le x} \Lambda(m) e(m^2\alpha) \ll x^{1+\varepsilon} \left(\frac{1}{q} + \frac{1}{x^{1/2}} + \frac{q}{x^2}\right)^{1/4}.$$

Hence we have

$$\sup_{\alpha\in C(\mathfrak{M})} |T(\alpha)| \ll N^{1/2-\theta/5}.$$

Now we take $\theta = 2c_1\eta \log(e/\eta) = 2\Theta$ and η sufficiently small such that $0 < \theta \le 1/25$. Then by Lemma 4.1, the second integral in (5.2) satisfies

$$\int_{C(\mathfrak{M})\cap\mathfrak{E}} \ll N^{\Theta-1} N^{2-4\theta/5} L^{2k} \ll NL^{2k-1}. \tag{5.4}$$

Setting $\eta^{-1} = 2k\log^{-1}k$ and then using Lemmas 4.1 and 5.1, the last integral in (5.2) can be estimated as

$$\int_{C(\mathfrak{M})\cap C(\mathfrak{E})} \ll \{(1-\eta)L\}^{2k-4}\int_0^1 |T(\alpha)G(\alpha)|^4 d\alpha \ll (1-\eta)^{2k-4}NL^{2k} \ll \frac{\log^2 k}{k}NL^{2k}. \tag{5.5}$$

With our choice of η, we have $\theta \ll k^{-1}\log^2 k$, so we do have $\theta \le 1/25$ for sufficiently large k, say $k \ge k_0$. Combining (5.3), (5.4) and (5.5), we get (5.1) for $k \ge k_0$ and $N \ge N_k$. For k with $3 \le k < k_0$, (5.1) follows directly from Lemma 5.1. The proof of Lemma 5.2 is complete.

Proof of Lemma 5.3. The proof is much more easier than that of Lemma 5.2. In particular it does not need Proposition 2.1 or Lemma 5.1.

We have, by (1.6),

$$\sum_{n\le N} r_k'(n) = \sum_{m_1^2+m_2^2\le N} f(n-m_1^2-m_2^2)\Lambda(m_1)\Lambda(m_2),$$

where $f(t)$ is the number of integral vectors $(\nu_1, ..., \nu_k)$ with $2 \le \nu_j \le L$ for which $4^{\nu_1}+...+4^{\nu_k} \le t$. Since $f(t)$ is non-negative and increasing,

$$\begin{aligned} f(N/L)\sum_{m_1^2+m_2^2\le N-N/L}\Lambda(m_1)\Lambda(m_2) &\le \sum_{m_1^2+m_2^2\le N} f(n-m_1^2-m_2^2)\Lambda(m_1)\Lambda(m_2) \\ &\le f(N)\sum_{m_1^2+m_2^2\le N}\Lambda(m_1)\Lambda(m_2). \end{aligned} \tag{5.6}$$

By the method of Rieger [R], Satz 4 (see also the proof of Shields [S], Theorem 2), we have

$$\sum_{m_1^2+m_2^2\le N}\Lambda(m_1)\Lambda(m_2) = \pi(\log\sqrt{N})^2\frac{N}{\log^2 N}\left\{1+O\left(\frac{1}{\log^{1/2}N}\right)\right\} = \frac{1}{4}\pi N\left\{1+O\left(\frac{1}{\log^{1/2}N}\right)\right\}.$$

Also, one easily sees $\log_4^k(N/k) \le f(N) \le \log_4^k N$, from which it follows that $f(N) \sim L^k$ as $N \to \infty$. Inserting these estimates into (5.6), one gets as $N \to \infty$,

$$\sum_{n\le N} r_k'(n) \sim \frac{1}{4}\pi NL^k. \tag{5.7}$$

Now we estimate the difference

$$\sum_{n\le N} r_k'(n) - \sum_{\substack{n\le N\\ n\equiv 2(\bmod 24)}} r_k'(n) = \sum_{*}(\log p_1)(\log p_2),$$

where $*$ indicates that the summation is over the set of $p_j \ge 2, \ell_j \ge 1, \nu_j \ge 2$ such that

$$p_1^{2\ell_1}+p_2^{2\ell_2}+4^{\nu_1}+...+4^{\nu_k}\begin{cases} \le & N; \\ \not\equiv & 2(\bmod 24). \end{cases} \tag{5.8}$$

Since $4^\nu \equiv 16(\bmod 24)$ for all $\nu \ge 2$, we have in (5.8) that

$$4^{\nu_1}+...+4^{\nu_k} \equiv 16k \equiv 0(\bmod 24)$$

on noting that $3|k$; consequently (5.8) implies

$$p_1^{2\ell_1}+p_2^{2\ell_2} \not\equiv 2(\bmod 24).$$

The above congruence holds only if at least one of the $p_j \le 3$, since $p^{2\ell} \equiv 1 (\bmod 24)$ for all $p \ge 5$ and all $\ell \ge 1$. Going back to (5.8), we see that

$$\sum_{*} (\log p_1)(\log p_2) \ll N^{1/2} L^{k+3},$$

which in combination with (5.7) gives Lemma 5.3.

Now we give

Proof of Theorem 1. We follow Gallagher [G] closely. Let $R = 6\pi L^k$. Then the left-hand side of (1.7) can be written as

$$\sum_{\substack{n \le N \\ n \equiv 2 (\bmod 24)}} \left(r'_k(n) - R\right)^2 = \sum_{\substack{n \le N \\ n \equiv 2 (\bmod 24)}} \left(r'_k(n)\right)^2 - 2R \sum_{\substack{n \le N \\ n \equiv 2 (\bmod 24)}} r'_k(n) + R^2 \sum_{\substack{n \le N \\ n \equiv 2 (\bmod 24)}} 1. \tag{5.9}$$

Clearly, the last sum in (5.9) is $\sim N/24$ as $N \to \infty$. The other two sums on the right-hand side of (5.9) can be estimated by Lemmas 5.2 and 5.3 respectively, which gives

$$\begin{aligned} \sum_{\substack{n \le N \\ n \equiv 2 (\bmod 24)}} \left(r'_k(n) - R\right)^2 \quad &\le \quad \frac{3}{2}\pi^2 N L^{2k} \left\{1 + O\left(\frac{\log^2 k}{k}\right)\right\} - \frac{2 \times 6}{4}\pi^2 N L^{2k}\{1 + o(1)\} \\ &+ \frac{6^2}{24}\pi^2 N L^{2k}\{1 + o(1)\} \ll \frac{\log^2 k}{k} N L^{2k}. \end{aligned}$$

The desired result (1.7) now follows. This completes the proof of Theorem 1.

Proof of Theorem 2. Since $3|k$, we have $k = 6k'$ or $k = 6k' + 3$. Let $i = 3k'$, and $j = i$ or $j = i + 3$ according as $k = 6k'$ or $k = 6k' + 3$. Then in either case $k = i + j$ with $3|i$ and $3|j$. Thus by (1.8) and (1.6) for $N \equiv 4 (\bmod 24)$,

$$\begin{aligned} r''_k(N) = \sum_{m+n=N} r'_i(m) r'_j(n) \quad &= \quad \sum_{\substack{m+n=N \\ m,n \equiv 2 (\bmod 24)}} r'_i(m) r'_j(n) + O\left\{(\max_m r'_i(m)) \sum_{\substack{n \le N \\ n \not\equiv 2 (\bmod 24)}} r'_j(n)\right\} \\ &= \quad \sum_{\substack{m+n=N \\ m,n \equiv 2 (\bmod 24)}} r'_i(m) r'_j(n) + O(N^{1/2+\varepsilon}), \end{aligned} \tag{5.10}$$

by an argument similar to that after (5.7). We put

$$r'_i(m) = 6\pi L^i + s_i(m),$$

for $m \equiv 2 (\bmod 24)$, and define $s_j(n)$ similarly. Then the sum in (5.10) is

$$\begin{aligned} &36\pi^2 L^k \sum_{\substack{m+n=N \\ m,n \equiv 2 (\bmod 24)}} 1 + 6\pi \left\{ L^i \sum_{\substack{m+n=N \\ m,n \equiv 2 (\bmod 24)}} s_j(n) + L^j \sum_{\substack{m+n=N \\ m,n \equiv 2 (\bmod 24)}} s_i(m) \right\} \\ &\quad + \sum_{\substack{m+n=N \\ m,n \equiv 2 (\bmod 24)}} s_i(m) s_j(n). \end{aligned} \tag{5.11}$$

By Cauchy's inequality and Theorem 1, the last sum in (5.11) is

$$\ll \left\{\frac{\log^2 i}{i} N L^{2i}\right\}^{1/2} \left\{\frac{\log^2 j}{j} N L^{2j}\right\}^{1/2} \ll \frac{\log^2 k}{k} N L^k$$

if $N \ge N_k$. By Lemma 5.3,

$$\sum_{\substack{m+n=N \\ m,n\equiv 2(\bmod 24)}} s_\ell(m) = \sum_{\substack{m\le N \\ m\equiv 2(\bmod 24)}} s_\ell(m) = o(NL^\ell)$$

for $\ell = i, j$. The first sum in (5.11) is $\sim N/24$ as $N \to \infty$. Summing up these estimates, we get (1.9) hence Theorem 2.

Acknowledgement. The research is partially supported by Hong Kong Government RGC research grant. In addition, the first and the third authors are partially supported by The National Natural Science Foundation.

References

[BF] J. Brüdern and E. Fouvry, Lagrange's four squares theorem with almost prime variables, J. reine angew. Math., 454(1994), 59-96.

[G] P.X. Gallagher, Primes and powers of 2, Invent. Math. 29(1975), 125-142.

[Gh] A. Ghosh, The distribution of αp^2 modulo 1, Proc. London Math. Soc. (Ser.3), 42(1981), 252-269.

[H1] L.K. Hua, Some results in the additive prime number theory, Quart. J. Math. (Oxford) 9(1938), 68-80.

[H2] L.K. Hua, *Additive theory of prime numbers* (in Chinese), Science press, 1957; English version, Amer. Math. Soc. 1965.

[H3] L.K. Hua, *Introduction to number theory,* Science press, 1957; English version, Springer, 1982.

[L1] Yu.V. Linnik, Prime numbers and powers of two, Trudy Mat. Inst. Steklov, 38(1951), 151-169.

[L2] Yu.V. Linnik, Addition of prime numbers and powers of one and the same number, Mat. Sb. (N.S.), 32(1953), 3-60.

[LL] J.Y. Liu, M.C. Liu, Representation of even integers as squares of primes and powers of 2, to appear in J. Number Theory.

[LLZ] J.Y. Liu, M.C. Liu and T. Zhan, Squares of primes and powers of 2, Mh. Math., 128(1999), 283-313.

[R] G.J. Rieger, Über die Summe aus einem Quadrat und einem Primzahlquadrat, J. reine angew. Math. 251(1968), 89-100.

[Ro] N.P. Romanoff,Über einige Sätze der additiven Zahlentheorie, Math. Ann. 57(1934), 668-678.

[S] P. Shields, Representation of integers as a sum of two h-th powers of primes, Proc. London Math. Soc. (Ser.3), 38(1979), 369-384.

Jianya Liu
Department of Mathematics
Shandong University
Jinan, Shandong 250100
P. R. China
e-mail:jyliu@sdu.edu.cn

Ming-Chit Liu
Department of Mathematics
The University of Hong Kong
Pokfulam, Hong Kong

current address

P.O.Box 625
Alhambra, CA 91802
U.S.A.
e-mail:matmcliu@yahoo.com

Tao Zhan
Department of Mathematics
Shandong University
Jinan, Shandong 250100
P. R. China
e-mail:zhantao@sdu.edu.cn