Four squares of primes and 165 powers of 2

by

JIANYA LIU and GUANGSHI LÜ (Jinan)

1. Introduction. It was shown by Linnik [12], [13] that each large even integer N is a sum of two primes and a bounded number of powers of 2,

$$(1.1) \qquad N = p_1 + p_2 + 2^{\nu_1} + 2^{\nu_2} + \cdots + 2^{\nu_k},$$

where (and throughout) p and ν, with or without subscripts, denote a prime number and a positive integer respectively. Later Gallagher [2] established a stronger result by a different method. An explicit value for the number k of powers of 2 was firstly established by the first author, Liu and Wang [17], who found that $k = 54000$ is acceptable. The original value for the number k was subsequently improved by Li [10], Wang [22], and Li [11]. Recently Heath-Brown and Puchta [6] applied a rather different approach to this problem and showed that $k = 13$ is acceptable.

In 1938, Hua [7] proved that each large integer congruent to $5 \pmod{24}$ can be written as a sum of five squares of primes. In view of this result and Lagrange's theorem of four squares, it seems reasonable to conjecture that each large integer $n \equiv 4 \pmod{24}$ is a sum of four squares of primes,

$$(1.2) \qquad n = p_1^2 + p_2^2 + p_3^2 + p_4^2.$$

Motivated by this conjecture and the above works of Linnik and Gallagher, it is proved in [18] that every large even integer N can be written as a sum of four squares of primes and powers of 2,

$$(1.3) \qquad N = p_1^2 + p_2^2 + p_3^2 + p_4^2 + 2^{\nu_1} + 2^{\nu_2} + \cdots + 2^{\nu_k}.$$

And in [15], it is showed that $k = 8330$ is acceptable in (1.3).

In this paper we sharpen this result considerably by establishing the following theorem.

THEOREM 1.1. *Every large even integer is the sum of four squares of primes and* 165 *powers of* 2.

Our Theorem 1.1 implies that there is a set $\Im$ of integers $n \leq x$ of cardinality only $O(\log^{165} x)$, such that every large even integer $N \leq x$ can be written as $N = p_1^2 + p_2^2 + p_3^2 + p_4^2 + n$, with p_1, p_2, p_3, p_4 being primes and $n \in \Im$. Thus our result can be compared with other approximations to the conjecture (1.2). In [4], Greaves gave a lower bound for the number of representations of an integer as a sum of two squares of integers and two squares of primes. Later Shields [21], Plaksin [19], and Koval'chik [9] obtained, among other things, an asymptotic formula for the number of representations of an integer as a sum of two squares of integers and two squares of primes. Brüdern and Fouvry [1] proved that every large $n \equiv 4$ $(\mathrm{mod}\, 24)$ is a sum of four squares of integers with each of their prime factors greater than $n^{1/68.86}$. Very recently the first author [14] proved that, with at most $O(N^{2/5+\varepsilon})$ exceptions, all positive integers $n \equiv 4 \ (\mathrm{mod}\, 24)$ not exceeding N can be written as (1.2).

Notation. As usual, $\varphi(n)$ and $\mu(n)$ stand for the function of Euler and Möbius respectively. N is a large integer, and $L = \log_2 N$. If there is no ambiguity, we express $\frac{a}{b} + \theta$ as $a/b + \theta$ or $\theta + a/b$. The same convention will be applied for quotients. The letter ε denotes a positive constant which is arbitrarily small.

2. Outline of the method. Here we give an outline for the proof of Theorem 1.1. In order to apply the circle method, we set

(2.1) $$P = N^{1/5-\varepsilon}, \quad Q = N/PL^{14}.$$

By Dirichlet's lemma on rational approximation, each $\alpha \in [1/Q, 1 + 1/Q]$ may be written in the form

(2.2) $$\alpha = a/q + \lambda, \quad |\lambda| \leq 1/(qQ),$$

for some integers a, q with $1 \leq a \leq q \leq Q$ and $(a, q) = 1$. We denote by $\mathcal{M}(a, q)$ the set of α satisfying (2.2), and define the *major arcs* $\mathcal{M}$ and the *minor arcs* $C(\mathcal{M})$ as follows:

(2.3) $$\mathcal{M} = \bigcup_{\substack{q \leq P}} \bigcup_{\substack{a=1 \\ (a,q)=1}}^{q} \mathcal{M}(a, q), \quad C(\mathcal{M}) = \left[\frac{1}{Q}, 1 + \frac{1}{Q}\right] \backslash \mathcal{M}.$$

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

Let

(2.4) $$T(\alpha) = \sum_{p^2 \leq N} (\log p)e(p^2\alpha), \quad G(\alpha) = \sum_{2^\nu \leq N} e(2^\nu\alpha) = \sum_{\nu \leq L} e(2^\nu\alpha)$$

and

(2.5) $$r_k(N) = \sum_{N = p_1^2 + \cdots + p_4^2 + 2^{\nu_1} + \cdots + 2^{\nu_k}} (\log p_1) \cdots (\log p_4).$$

Then $r_k(N)$ can be written as

$$(2.6) \qquad r_k(N) = \int_0^1 T^4(\alpha) G^k(\alpha) e(-N\alpha) \, d\alpha$$

$$= \left\{ \int_{\mathcal{M}} + \int_{C(\mathcal{M})} \right\} T^4(\alpha) G^k(\alpha) e(-N\alpha) \, d\alpha.$$

To handle the integral on the major arcs, we quote the following lemma.

LEMMA 2.1. *Let $\mathcal{M}$ be as in (2.3) with P determined by (2.1). Then for $2 \le n \le N$, we have*

$$(2.7) \qquad \int_{\mathcal{M}} T^4(\alpha) e(-n\alpha) \, d\alpha = \frac{\pi^2}{16} \mathfrak{S}(n) n + O\left(\frac{N}{\log N} \right).$$

Here $\mathfrak{S}(n)$ is defined in (5.2), and satisfies $\mathfrak{S}(n) \gg 1$ for $n \equiv 4 \pmod{24}$.

This is Theorem 1.2 in [14]. The asymptotic formula (2.7) was previously established in [16] for the smaller $\mathcal{M}$ with $P = N^{2/15-\varepsilon}$ in (2.1).

A crucial step in bounding the contributions of the minor arcs is an upper bound for the number of solutions of the equation

$$(2.8) \qquad n = p_1^2 + p_2^2 - p_3^2 - p_4^2, \quad |n| \le N, \, p_j^2 \le N.$$

LEMMA 2.2. *Let $n \ne 0$ be an integer with $n \equiv 0 \pmod{24}$, and $r_-(n)$ the number of representations of n in the form (2.8). Then we have*

$$(2.9) \qquad r_-(n) \le c_1 \mathfrak{S}_-(n) \frac{\pi^2}{16} \frac{N}{\log^4 N}$$

with $c_1 \le (1+\varepsilon)^6 \cdot 101 \cdot 44^4$ and

$$(2.10) \quad \mathfrak{S}_-(n) = \left(2 - \frac{1}{2^{\beta_0-1}} - \frac{1}{2^{\beta_0}} \right) \prod_{\substack{p^\beta \| n \\ p \ge 3, \, \beta \ge 0}} \left(1 + \frac{1}{p} - \frac{1}{p^{\beta+1}} - \frac{1}{p^{\beta+2}} \right)$$

where β_0 satisfies $2^{\beta_0} \| n$ and the singular series $\mathfrak{S}_-(n)$ is defined by (3.4), (3.5) and (3.7).

This lemma improves Theorem 2 of [15].

On the minor arcs, we also need estimates for the measure of the set

$$\mathcal{E}_\lambda = \{ \alpha \in (0,1] : |G(\alpha)| \ge \lambda L \}.$$

The following lemma is due to Heath-Brown and Puchta [6].

LEMMA 2.3. *Let*

$$G_h(\alpha) = \sum_{0 \le n \le h-1} e(\alpha 2^n)$$

and

$$F(\xi, h) = \frac{1}{2^h} \sum_{r=0}^{2^h-1} \exp\left\{\xi \operatorname{Re}\left(G_h\left(\frac{r}{2^h}\right)\right)\right\}.$$

Then

$$\operatorname{meas}(\mathcal{E}_\lambda) \le N^{-E(\lambda)},$$

where

$$E(\lambda) = \frac{\xi\lambda}{\log 2} - \frac{\log F(\xi, h)}{h \log 2} - \frac{\varepsilon}{\log 2}$$

holds true for any $h \in N$, *any* $\xi > 0$ *and* $\varepsilon > 0$.

On the minor arcs, the new result of Ren [20] (see Lemma 5.4 below) on exponential sums over primes will also be applied.

The proof of Lemma 2.2 follows the same lines as that of Theorem 2 in [15], so we only give an outline in §3. In §4, we estimate an integral. In the last section, we complete the proof of Theorem 1.1.

3. Proof of Lemma 2.2. In this section, boldface symbols denote 4-dimensional vectors, for example $\mathbf{d} = (d_1, d_2, d_3, d_4)$. The letter $\mathbf{e}$ is reserved for $(1, 1, 1, 1)$. Also we define $|\mathbf{d}| = \max |d_j|$ and $\widetilde{\mathbf{d}} = d_1 d_2 d_3 d_4$. For a vector $\mathbf{d}$, we write $\mu(\mathbf{d})$ for $\mu(d_1)\mu(d_2)\mu(d_3)\mu(d_4)$.

In order to establish Lemma 2.2, we want to sieve the set

$$(3.1) \qquad \mathcal{A} = \{(x_1, \ldots, x_4) \in \mathbb{N}^4 : x_1^2 + x_2^2 - x_3^2 - x_4^2 = n,$$
$$1 \le |n| \le N, 1 \le x_j^2 \le N\}.$$

To this end, we require information concerning the distribution of the sequence $\mathcal{A}$ in arithmetic progressions.

Let $\mathbf{x} \equiv 0 \pmod{\mathbf{d}}$ denote the simultaneous condition $x_j \equiv 0 \pmod{d_j}$ for $j = 1, \ldots, 4$. We need an asymptotic formula for the cardinality of

$$\mathcal{A}_{\mathbf{d}} = \{\mathbf{x} \in \mathcal{A} : \mathbf{x} \equiv 0 \pmod{\mathbf{d}}\},$$

i.e. the number of solutions of the equation

$$(3.2) \quad d_1^2 x_1^2 + d_2^2 x_2^2 - d_3^2 x_3^2 - d_4^2 x_4^2 = n, \quad 1 \le |n| \le N, \ 1 \le d_j^2 x_j^2 \le N.$$

The expected main term of $|\mathcal{A}_{\mathbf{d}}|$ is

$$(3.3) \qquad \frac{1}{\widetilde{\mathbf{d}}} \frac{\pi}{16} \mathfrak{S}(n, \mathbf{d}) \mathfrak{I}\left(\frac{n}{N}\right) N,$$

where

$$(3.4) \qquad \mathfrak{S}(n, \mathbf{d}) = \sum_{q=1}^{\infty} q^{-4} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} e\left(\frac{an}{q}\right) S(q, ad_1^2) S(q, ad_2^2)$$
$$\times S(q, -ad_3^2) S(q, -ad_4^2)$$

with

$$(3.5) \qquad S(q, a) = \sum_{m=1}^{q} e\left(\frac{am^2}{q}\right),$$

and where

$$(3.6) \qquad \Im(x) = 2 \int_{\max(0,-x)}^{\min(1,1-x)} v^{-1/2}(1 - x - v)^{1/2}\, dv.$$

Clearly, for $1 \leq |n| \leq N$ we have $0 \leq \Im(n/N) \leq \pi$. The singular series in Lemma 2.2 is defined by

$$(3.7) \qquad \mathfrak{S}_-(n) = \mathfrak{S}(n, \mathbf{e}).$$

Also, we define

$$(3.8) \qquad \omega(\mathbf{d}) = \omega(\mathbf{d}, n) = \mathfrak{S}(n, \mathbf{d})/\mathfrak{S}_-(n),$$

so that (3.3) becomes

$$\frac{\omega(\mathbf{d})}{\widetilde{\mathbf{d}}} \frac{\pi}{16} \mathfrak{S}_-(n)\Im\left(\frac{n}{N}\right)N.$$

The difference between $|\mathcal{A}_{\mathbf{d}}|$ and its main term expected above has been estimated on average in Lemma 9.1 of [18], which is a minor modification of Theorem 3 in Brüdern and Fouvry [1].

LEMMA 3.1. *Let $\omega(\mathbf{d})$ be as in (3.8), and let $D = N^{1/22 - 2\varepsilon}$. Define $R(n, N, \mathbf{d})$ by*

$$(3.9) \qquad |\mathcal{A}_{\mathbf{d}}| = \frac{\omega(\mathbf{d})}{\widetilde{\mathbf{d}}} \frac{\pi}{16} \mathfrak{S}_-(n)\Im\left(\frac{n}{N}\right)N + R(n, N, \mathbf{d}),$$

where $\mathfrak{S}_-(n)$ and $\Im(n/N)$ are as in (3.7) and (3.6) respectively. Then for arbitrary $A > 0$, we have

$$\sum_{|\mathbf{d}| \leq D} \mu^2(\mathbf{d})|R(n, N, \mathbf{d})| \ll NL^{-A}.$$

We suppose throughout that $\mu^2(\mathbf{d}) = 1$. For $u, v = 0, 1, 2$, let $\mathbf{e}_{u,v}(p)$ denote the following 4-dimensional vector:

$$(\underbrace{p, \ldots, p}_{u}, \underbrace{1, \ldots, 1}_{2-u}, \underbrace{p, \ldots, p}_{v}, \underbrace{1, \ldots, 1}_{2-v}).$$

Clearly, $\mathbf{e}_{0,0}(p) = \mathbf{e}$. It has been proved in [18] that $\omega(\mathbf{d})$ has the decomposition

$$(3.10) \qquad \omega(\mathbf{d}) = \prod_{p^{u+v} \| \widetilde{\mathbf{d}}} \omega_{u,v}(p).$$

Trivially, we have $\omega_{0,0}(p) = 1$. For $1 \leq u + v \leq 4$, the values of $\omega_{u,v}(p)$ are given in Lemmas 8.1 and 8.2 of [18].

LEMMA 3.2. *With $\omega_{u,v}(p)$ defined as in (3.10), put*

$$(3.11) \quad \Omega(p) = 2\omega_{1,0}(p) + 2\omega_{0,1}(p) - \frac{\omega_{2,0}(p)}{p} - \frac{\omega_{0,2}(p)}{p} - \frac{4\omega_{1,1}(p)}{p}$$
$$+ \frac{2\omega_{2,1}(p)}{p^2} + \frac{2\omega_{1,2}(p)}{p^2} - \frac{\omega_{2,2}(p)}{p^3},$$

and

$$W(z) = \prod_{2 < p \leq z} \left(1 - \frac{\Omega(p)}{p} \right).$$

Then

$$\log^{-4} z \ll W(z) \leq c_2 e^{-4\gamma} \log^{-4} z$$

with $c_2 \leq 101$, where γ denotes the Euler constant.

Proof. This lemma has been proved in [15] with the value $c_2 = \pi^{24}/2^{32} = 198.901\ldots$, so we only have to show that the value $c_2 = 101$ is acceptable. To this end, we should estimate $1 - \Omega(p)/p$ for all $p \geq 3$. We distinguish two cases according as $p \mid n$ or not. For convenience we write $x = 1/p$.

Suppose first that $p^\beta \| n$ with $p \geq 3$ and $\beta \geq 1$. Then by (3.11) and Lemma 8.2 in [18],

$$\frac{\Omega(p)}{p} \geq \begin{cases} \dfrac{4x - 11x^2 + 9x^3 - x^{\beta+1} - x^{\beta+2}}{1 + x - x^{\beta+1} - x^{\beta+2}} & \text{if } p \equiv 1 \ (\mathrm{mod}\,4), \\[2mm] \dfrac{4x - 7x^2 + 5x^3 - x^{\beta+1} - x^{\beta+2}}{1 + x - x^{\beta+1} - x^{\beta+2}} & \text{if } p \equiv 3 \ (\mathrm{mod}\,4), \end{cases}$$

and consequently,

$$(3.12) \quad 1 - \frac{\Omega(p)}{p} \leq \begin{cases} (1-x)^4 \dfrac{9x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} & \text{if } p \equiv 1 \ (\mathrm{mod}\,4), \\[2mm] (1-x)^4 \dfrac{5x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} & \text{if } p \equiv 3 \ (\mathrm{mod}\,4). \end{cases}$$

One can easily see that for $p \equiv 3 \ (\mathrm{mod}\,4)$ and $p \geq 7$,

$$\frac{5x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} \leq (1-x^2)^{-11}$$

and for $p \equiv 1 \ (\mathrm{mod}\,4)$ and $p \geq 29$,

$$\frac{9x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} \leq (1-x^2)^{-11}.$$

Thus

$$(3.13) \quad 1 - \frac{\Omega(p)}{p} \leq \begin{cases} (1-x)^4(1-x^2)^{-11} & \text{if } p \equiv 1 \ (\mathrm{mod}\,4), \ p \geq 29, \\ (1-x)^4(1-x^2)^{-11} & \text{if } p \equiv 3 \ (\mathrm{mod}\,4), \ p \geq 7. \end{cases}$$

Now we consider the second case $p \nmid n$. For these p, we must have $p > 3$, since the other case will contradict $n \equiv 0 \pmod{24}$. Thus by (3.11) and Lemma 8.1 of [18], we have

$$\frac{\Omega(p)}{p} \geq \frac{4}{p+1} - \frac{2}{p(p-1)} - \frac{4}{p(p+1)} = \frac{4p^2 - 10p + 2}{p(p^2-1)},$$

and consequently

$$1 - \frac{\Omega(p)}{p} \leq \frac{1 - 4x + 9x^2 - 2x^3}{1 - x^2}.$$

It is easily seen that for $x < 1/3$, we also have

$$(3.14) \qquad 1 - \frac{\Omega(p)}{p} \leq \begin{cases} (1-x)^4 \dfrac{9x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} & \text{if } p \equiv 1 \pmod 4, \\[4mm] (1-x)^4 \dfrac{5x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} & \text{if } p \equiv 3 \pmod 4 \end{cases}$$

and hence the bound (3.13) still holds true in this case.

From (3.12)–(3.14), we conclude that

$$\prod_{3 \leq p \leq z} \left(1 - \frac{\Omega(p)}{p}\right)$$

$$\leq \prod_{3 \leq p \leq z} \left(1 - \frac{1}{p}\right)^4 \prod_{p \geq 3} \left(1 - \frac{1}{p^2}\right)^{-11} \prod_{\substack{p=5,13,17, \\ 29,37,41}} \frac{(9x^2 - 2x + 1)(1-x^2)^{11}}{(1-x^2)^2(1-x)^2}$$

$$\times \prod_{\substack{p=3,7,11, \\ 19,23}} \frac{5x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} (1-x^2)^{11}$$

$$\leq 2^4 \prod_{2 \leq p \leq z} \left(1 - \frac{1}{p}\right)^4 \left(1 - \frac{1}{2^2}\right)^{11} \prod_{p \geq 2} \left(1 - \frac{1}{p^2}\right)^{-11}$$

$$\times \prod_{\substack{p=5,13,17, \\ 29,37,41}} \frac{(9x^2 - 2x + 1)(1-x^2)^{11}}{(1-x^2)^2(1-x)^2} \prod_{\substack{p=3,7,11, \\ 19,23}} \frac{(5x^2 - 2x + 1)(1-x^2)^{11}}{(1-x^2)^2(1-x)^2}$$

$$\leq \frac{2^4 e^{-4\gamma}}{\log^4 z} \left(\frac{3}{4}\right)^{11} \zeta^{11}(2) \prod_{\substack{p=5,13,17, \\ 29,37,41}} \frac{9x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} (1-x^2)^{11}$$

$$\times \prod_{\substack{p=3,7,11, \\ 19,23}} \frac{5x^2 - 2x + 1}{(1-x^2)^2(1-x)^2} (1-x^2)^{11}.$$

Thus we can show that $c_2 < 101$ by computations and the well known fact $\zeta(2) = \pi^2/6$.

Proof of Lemma 2.2. Let $P(z_0, z) = \prod_{z_0 \le p < z} p$ and define $T(\mathcal{A}, z_0)$ to be the set of all $\mathbf{x} \in \mathcal{A}$ such that $p \mid \widetilde{\mathbf{x}} \Rightarrow p = 2$ or $p > z_0$. Then

$$S(\mathcal{A}, z) = \sum_{\mathbf{x} \in T(\mathcal{A}, z_0)} \prod_{j=1}^{4} (\mu * \mathbf{1})((x_j, P(z_0, z))),$$

where

$$(\mu * \mathbf{1})((x_j, P(z_0, z))) = \sum_{d \mid (x_j, P(z_0, z))} \mu(d).$$

We fix $z = N^{1/44 - \varepsilon}$ and $z_0 = \log^{20} N$. Following the lines of [15], we can show that

$$S(\mathcal{A}, z) \le W(z_0) 44^4 e^{4\gamma} (1 + \varepsilon)^4 \frac{\log^4 z_0}{\log^4 N} \frac{\pi}{16} \mathfrak{S}_-(n) \mathfrak{I}\left(\frac{n}{N}\right) N + O\left(\frac{N}{\log^5 N}\right)$$

$$\le 101 \cdot 44^4 (1 + \varepsilon)^5 \cdot \mathfrak{S}_-(n) \frac{\pi^2}{16} \frac{N}{\log^4 N},$$

where we have used Lemmas 3.1 and 3.2. Thus we have

$$r_-(n) \le S(\mathcal{A}, z) + N^{2/3} \le 101 \cdot 44^4 (1 + \varepsilon)^6 \mathfrak{S}_-(n) \frac{\pi^2}{16} \frac{N}{\log^4 N}.$$

4. Estimation of an integral. In this section we shall estimate the integral $\int_0^1 |T(\alpha)G(\alpha)|^4 \, d\alpha$. We have

LEMMA 4.1. *Let* $T(\alpha)$ *and* $G(\alpha)$ *be as in* (2.4). *Then*

$$\int_0^1 |T(\alpha)G(\alpha)|^4 \, d\alpha \le c_3 \frac{\pi^2}{16} NL^4,$$

where

$$c_3 \le \left(\frac{44^4 \cdot 101 \cdot 43}{25 \cdot 3} + \frac{2^3}{\pi^2} \log^2 2\right)(1 + \varepsilon)^9.$$

To show this, we need

LEMMA 4.2. *For odd* q, *let* $\varrho(q)$ *be the smallest positive integer* ϱ *such that* $2^\varrho \equiv 1 \pmod{q}$. *Then the series* $\sum_{q=1, 2 \nmid q}^{\infty} \mu^2/q\varrho(q)$ *is convergent, and its value* c_4 *satisfies* $c_4 < 43/25$.

Proof. See Lemma 5.2 in [15].

Proof of Lemma 4.1. One easily sees that

$$(4.1) \qquad \int_0^1 |T(\alpha)G(\alpha)|^4 \, d\alpha \le (\log \sqrt{N})^4 Z(N),$$

where $Z(N)$ denotes the number of solutions of the equation

(4.2) $$p_1^2 + p_2^2 - p_3^2 - p_4^2 = 2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4}$$

with

(4.3) $$p_j \leq N^{1/2}, \quad m_j \leq L.$$

Now we distinguish between two cases.

CASE 1. In this case we treat the contribution from those (m_1, m_2, m_3, m_4) such that

(4.4) $$2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4} \neq 0.$$

Let (m_1, m_2, m_3, m_4) be fixed and satisfy, in addition to (4.4), $2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4} \equiv 0 \pmod{24}$. For these (m_1, m_2, m_3, m_4), one trivially has $|2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4}| \leq 2N$, so one deduces from Lemma 2.2 that

$$|\{(p_1, p_2, p_3, p_4) : p_j \text{ satisfies (4.2) and (4.3)}\}|$$
$$\leq c_1 \mathfrak{S}_-(2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4}) \pi^2 2N/(16 \log^4(2N))$$
$$\leq 4c_1 g(2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4}) \pi^2 N/(16 \log^4 N),$$

where $g(h) = \prod_{p|h,\, p \geq 3}(1 + 1/p)$. On the other hand, for (m_1, m_2, m_3, m_4) satisfying (4.4) and $2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4} \not\equiv 0 \pmod{24}$, we have

$$|\{(p_1, p_2, p_3, p_4) : p_j \text{ satisfies (4.2) and (4.3)}\}| \ll N^{1/2+\varepsilon}.$$

Therefore $Z_1(N)$, the number of solutions of (4.2) with p_j, m_j satisfying both (4.3) and (4.4), can be estimated as

$$Z_1(N) \leq 4c_1(1 + \varepsilon) \frac{\pi^2}{16} \frac{N}{\log^4 N} \sum_{1 \leq m_1, \ldots, m_4 \leq L} g(2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4}).$$

Denote by Σ the sum above. Noting that $g(h) = g(-h)$ for $h \neq 0$, and that

$$\sum_{1 \leq m_1, \ldots, m_4 \leq L} 1 \leq 4(1 + \varepsilon) \sum_{\substack{1 \leq m_4 \leq m_1, m_2, m_3 \leq L \\ 3m_4 < m_1 + m_2 + m_3}} 1,$$

we get

$$\Sigma \leq 4(1 + \varepsilon) \sum_{\substack{1 \leq m_4 \leq m_1, m_2, m_3 \leq L \\ 3m_4 < m_1 + m_2 + m_3}} g(2^{m_1} + 2^{m_2} - 2^{m_3} - 2^{m_4}).$$

Here the condition $3m_4 < m_1 + m_2 + m_3$ in the above sum guarantees (4.4).

For a fixed integral vector (h_1, h_2, h_3) with $1 \leq h_j \leq L$, we have

$$|\{(m_1, m_2, m_3, m_4) : 1 \leq m_j \leq L, m_1 - m_4 = h_1, m_2 - m_4 = h_2, m_3 - m_4 = h_3\}|$$
$$\leq \min(L - h_1, L - h_2).$$

Thus,

(4.5) $\Sigma \leq 4(1+\varepsilon)$

$$\times \sum_{\substack{1 \leq m_4 \leq m_1, m_2, m_3 \leq L \\ 3m_4 < m_1 + m_2 + m_3}} g(2^{m_4}) g(2^{m_1-m_4} + 2^{m_2-m_4} - 2^{m_3-m_4} - 1)$$

$$\leq 4(1+\varepsilon) \sum_{\substack{0 \leq h_1, h_2, h_3 \leq L \\ h_1 + h_2 + h_3 > 0}} \min(L - h_1, L - h_2) g(2^{h_1} + 2^{h_2} - 2^{h_3} - 1).$$

Here the condition $h_1 + h_2 + h_3 > 0$ indicates that h_1, h_2, h_3 cannot vanish at the same time. Obviously, there are at most $O(L^2)$ terms in the last sum such that one or two of h_1, h_2, h_3 vanish, and the total contribution of these terms to Σ is $\ll L^3 \log \log N \ll L^3 \log L$, on using the elementary bound $g(d) \ll \log \log d$. Hence (4.5) becomes

$$\Sigma \leq 4(1+\varepsilon) \sum_{1 \leq h_1, h_2, h_3 \leq L} \min(L - h_1, L - h_2) g(2^{h_1} + 2^{h_2} - 2^{h_3} - 1) + O(L^3 \log L).$$

Since the positions of h_1, h_2 are symmetric, one deduces further that

(4.6) $$\Sigma \leq 8(1+\varepsilon) \sum_{1 \leq h_3 \leq L} \sum_{1 \leq h_1 \leq L} (L - h_1) \sum_{1 \leq h_2 \leq h_1} g(2^{h_1} + 2^{h_2} - 2^{h_3} - 1)$$
$$+ O(L^3 \log L).$$

Following the similar arguments in [15], we can use Lemma 4.2 to show that for $H \gg 1$,

$$\sum_{1 \leq h \leq H} g(2^h - t) \leq c_4 H$$

uniformly for all possible odd numbers t with $|t| \leq N$. Thus we obtain

$$\sum_{1 \leq h_3 \leq L} \sum_{1 \leq h_1 \leq L} (L - h_1) \sum_{1 \leq h_2 \leq h_1} g(2^{h_1} + 2^{h_2} - 2^{h_3} - 1) \leq \frac{c_4}{6} L^4 + O(L^{3+\varepsilon}).$$

Inserting this into (4.6), we get $\Sigma \leq \frac{4}{3} c_4 (1+\varepsilon)^2 L^4$, and consequently,

(4.7) $$Z_1(N) \leq \frac{c_1 c_4}{3} (1+\varepsilon)^3 \frac{\pi^2 N L^4}{\log^4 N}.$$

CASE 2. It has been proved in [15] that the upper bound of $Z_2(N)$, the number of solutions of (4.2) with p_j, m_j satisfying (4.3) but not (4.4), satisfies

(4.8) $$Z_2(N) \leq 8(1+\varepsilon) \frac{N L^2}{\log^2 N}.$$

We can now conclude from (4.7) and (4.8) that

$$Z(N) = Z_1(N) + Z_2(N) \leq \left(\frac{c_1 c_4}{3} + \frac{8}{\pi^2} \log^2 2 \right) (1+\varepsilon)^3 \frac{\pi^2 N L^4}{\log^4 N},$$

which in combination with (4.1) gives Lemma 4.1.

5. Proof of Theorem 1.1. We need the following four lemmas.

LEMMA 5.1 [15, Lemma 6.1]. *Let* $\Xi(N,k) = \{n \geq 2 : n = N - 2^{\nu_1} - \cdots - 2^{\nu_k}\}$, *with* $k \geq 2$. *Then for* $N \equiv 4 \pmod 8$,

$$\sum_{\substack{n \in \Xi(N,k) \\ n \equiv 4 \,(\mathrm{mod}\,24)}} n \geq \frac{1}{4}(1 - \varepsilon)NL^k.$$

LEMMA 5.2. *Let*

$$(5.1) \quad C(q,a) = \sum_{\substack{m=1 \\ (m,q)=1}}^{q} e\left(\frac{am^2}{q}\right), \quad B(n,q) = \sum_{\substack{a=1 \\ (a,q)=1}}^{q} C^4(q,a)e\left(-\frac{an}{q}\right),$$

and

$$(5.2) \quad A(n,q) = \frac{B(n,q)}{\varphi^4(q)}, \quad \mathfrak{S}(n) = \sum_{q=1}^{\infty} A(n,q).$$

Then for $n \equiv 4 \pmod{24}$, *one has*

$$c_5 < \mathfrak{S}(n) \ll (\log\log n)^{11}$$

with $c_5 = 4.952$, *while for* $n \not\equiv 4 \pmod{24}$, *one has* $\mathfrak{S}(n) = 0$.

Proof. This is Proposition 4.3 in [18] except for the value of c_5. It has been shown in [18] that

$$(5.3) \quad \mathfrak{S}(n) = \{1 + A(n,2) + A(n,2^2) + A(n,2^3)\} \prod_{p \geq 3} \{1 + A(n,p)\},$$

where $A(n,q)$ is defined as in (5.2). It has also been proved in Lemma 4.2 of [18] that when $n \equiv 4 \pmod{24}$,

$$(5.4) \quad 1 + A(n,2) + A(n,2^2) + A(n,2^3) = 8, \quad 1 + A(n,3) = 3.$$

Therefore to estimate $\mathfrak{S}(n)$ it remains to compute $1 + A(n,p)$ for $p \geq 5$.

We will use the notation

$$G(\chi,n) = \sum_{m=1}^{q} \chi(m)e\left(\frac{nm}{q}\right), \quad c_q(n) = \sum_{\substack{m=1 \\ (m,q)=1}}^{q} e\left(\frac{nm}{q}\right),$$

where $c_q(n)$ is the Ramanujan sum. We will also use the notation $S(q,a)$ introduced in (3.5). By Theorem 7.5.4 in [8], for $p \geq 5$ we have

$$C(p,a) = S(p,a) - 1 = \chi(a)S(p,1) - 1,$$

where χ is the Legendre symbol $\left(\frac{a}{p}\right)$. Inserting this into (5.2), one sees that

$$B(n,p) = S^4(p,1)c_p(-n) - 4S^3(p,1)G(\chi,-n) + 6S^2(p,1)c_p(-n)$$
$$- 4S(p,1)G(\chi,-n) + c_p(-n).$$

Using the well known formulae (see Theorems 7.5.5 and 7.5.8 in [8])

$$S(p,1) = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod 4, \\ i\sqrt{p} & \text{if } p \equiv 3 \pmod 4, \end{cases}$$

and

$$|G(\chi,n)| = \begin{cases} \sqrt{p} & \text{if } p \nmid n, \\ 0 & \text{if } p \mid n, \end{cases} \qquad c_p(n) = \begin{cases} -1 & \text{if } p \nmid n, \\ p-1 & \text{if } p \mid n, \end{cases}$$

one obtains

$$B(n,p) \geq \begin{cases} -5p^2 + 2p - 1 & \text{if } p \nmid n, \ p \equiv 3 \pmod 4, \\ -5p^2 - 10p - 1 & \text{if } p \nmid n, \ p \equiv 1 \pmod 4, \\ (p-1)(p^2 - 6p + 1) & \text{if } p \mid n. \end{cases}$$

Hence by (5.2), we have

$$\prod_{p \geq 5} \{1 + A(n,p)\}$$

$$\geq \prod_{\substack{p \equiv 1 \pmod 4 \\ p \geq 5, \, p \nmid n}} \left(1 - \frac{5p^2 + 10p + 1}{(p-1)^4}\right) \prod_{\substack{p \equiv 3 \pmod 4 \\ p \geq 7, \, p \nmid n}} \left(1 - \frac{5p^2 - 2p + 1}{(p-1)^4}\right)$$

$$\times \prod_{\substack{p \geq 5 \\ p \mid n}} \left(1 + \frac{p^2 - 6p + 1}{(p-1)^3}\right)$$

$$> \prod_{\substack{p \equiv 1 \pmod 4 \\ p \geq 5}} \left(1 - \frac{5p^2 + 10p + 1}{(p-1)^4}\right) \prod_{\substack{p \equiv 3 \pmod 4 \\ p \geq 7}} \left(1 - \frac{5p^2 - 2p + 1}{(p-1)^4}\right).$$

To estimate the products above, we apply the elementary inequality

$$(1+x)^a < 1 + ax - \frac{a(a-1)}{2}x^2 \quad \text{if } a > 2, \ -1 < x < 0.$$

It is easy to see that for $p \geq 23$ and $p \equiv 1 \pmod 4$,

$$1 - \frac{5p^2 + 10p + 1}{(p-1)^4} \geq \left(1 - \frac{1}{(p-1)^2}\right)^6,$$

while for $p \geq 11$ and $p \equiv 3 \pmod 4$,

$$1 - \frac{5p^2 - 2p + 1}{(p-1)^4} \geq \left(1 - \frac{1}{(p-1)^2}\right)^6.$$

Thus we have

$$\prod_{p \geq 5}\{1 + A(n,p)\}$$

$$\geq \prod_{p=5,13,17}\left(1 - \frac{5p^2 + 10p + 1}{(p-1)^4}\right)\prod_{p=7}\left(1 - \frac{5p^2 - 2p + 1}{(p-1)^4}\right)$$

$$\times \prod_{\substack{p \equiv 1 \,(\mathrm{mod}\,4)\\ p \geq 23}}\left(1 - \frac{1}{(p-1)^2}\right)^6 \prod_{\substack{p \equiv 3 \,(\mathrm{mod}\,4)\\ p \geq 11}}\left(1 - \frac{1}{(p-1)^2}\right)^6$$

$$= \prod_{p=5,13,17}\left(1 - \frac{5p^2 + 10p + 1}{(p-1)^4}\right)\prod_{p=7}\left(1 - \frac{5p^2 - 2p + 1}{(p-1)^4}\right)$$

$$\times \prod_{\substack{p=3,5,\\ 7,13,17}}\left(1 - \frac{1}{(p-1)^2}\right)^{-6}\prod_{p \geq 3}\left(1 - \frac{1}{(p-1)^2}\right)^6$$

$$\geq 2.494447 \cdot (0.6601)^6 > 0.2063634,$$

where we have used $\prod_{p \geq 3}(1 - (p-1)^{-2}) = 0.6601\ldots$ (see [5]). This in combination with (5.3) and (5.4) ensures that one can take $c_5 = 4.952$. The proof is complete.

In order to apply Lemma 2.3 in this paper, we need to find an optimal λ such that $E(\lambda) > 3/4$. Thus we have to compute

$$F(\xi, h) = \frac{1}{2^h}\sum_{r=0}^{2^h - 1}\exp\left\{\xi \cdot \sum_{i=1}^{h}\cos\left(\frac{2\pi r}{2^i}\right)\right\}$$

and optimize for ξ. Use Mathematica 4.1 on a PC and the following procedure:

a=N[Sum[Cos[2πr/2^i], {i, 1, 22}]];

b=Apply[Plus, Table[Exp[$\xi*$a], {r, 0, $2^{22} - 1$}]];

(Log[b/2^{22}]/22/Log[2] + 3/4)*Log[2]/ξ.

We can take $\xi = 1.21$, $h = 22$ in Lemma 2.3 to get

LEMMA 5.3. *Let $E(\lambda)$ be as in Lemma 2.3. Then*

$$E(0.887167) > 3/4 + 10^{-10}.$$

LEMMA 5.4. *Let $T(\alpha)$ be as in (2.4) and $\alpha = a/q + \lambda$ subject to $(a,q) = 1$ and $\lambda \in \mathbb{R}$. Then*

$$T(\alpha) \ll N^{1/4+\varepsilon}\sqrt{q(1 + |\lambda|N)} + N^{2/5+\varepsilon} + \frac{N^{1/2+\varepsilon}}{\sqrt{q(1 + |\lambda|N)}}.$$

Proof. This is a special case of Theorem 1.1 in [20].

Now we prove the main result of this paper.

Proof of Theorem 1.1. We distinguish two cases according as $N \equiv 4$ $(\mathrm{mod}\, 8)$ or not.

CASE 1. Suppose $N \equiv 4\,(\mathrm{mod}\, 8)$. Let $\mathcal{E}_\lambda$ be as in Lemma 2.3 and $\mathcal{M}$ as in (2.3) with P, Q determined by (2.1). Then (2.6) becomes

$$(5.5) \quad r_k(N) = \int_0^1 T^4(\alpha) G^k(\alpha) e(-N\alpha)\, d\alpha = \int_{\mathcal{M}} + \int_{C(\mathcal{M}) \cap \mathcal{E}_\lambda} + \int_{C(\mathcal{M}) \cap C(\mathcal{E}_\lambda)} .$$

Introducing the notation $\Xi(N, k)$ and then applying Lemma 2.1, we see that the first integral on the right-hand side of (5.5) is

$$(5.6) \quad \sum_{n \in \Xi(N,k)} \int_{\mathcal{M}} T^4(\alpha) e(-n\alpha)\, d\alpha = \frac{\pi^2}{16} \sum_{n \in \Xi(N,k)} \mathfrak{S}(n) n + O(NL^{k-1})$$

$$\geq c_5 \frac{\pi^2}{16} \Big\{ \sum_{\substack{n \in \Xi(N,k) \\ n \equiv 4\,(\mathrm{mod}\, 24)}} n \Big\} + O(NL^{k-1}) \geq \frac{4.952}{4}\, (1-\varepsilon)\, \frac{\pi^2}{16}\, NL^k,$$

where in the last two inequalities we have used Lemmas 5.2 and 5.1 respectively.

By Dirichlet's lemma on rational approximations each real number $\alpha \in C(\mathcal{M})$ can be written as $\alpha = a/q + \lambda$, $(a, q) = 1$, with

$$1 \leq q \leq Q_0 = N^{3/4}, \quad |\lambda| \leq 1/(qQ_0).$$

We let $\mathcal{N}$ be the set of $\alpha \in C(\mathcal{M})$ satisfying $\alpha = a/q + \lambda$, $(a, q) = 1$, such that

$$P_0 = N^{1/4} < q \leq Q_0, \quad |\lambda| \leq 1/(qQ_0).$$

On $\mathcal{N}$, we apply Ghosh's result in [3], which states that

$$(5.7) \quad \max_{\alpha \in \mathcal{N}} |T(\alpha)| \ll N^{1/2+\varepsilon} P_0^{-1/4} + N^{7/16+\varepsilon} + N^{1/4+\varepsilon} Q_0^{1/4} \ll N^{1/2-1/16+\varepsilon}.$$

Let $\mathcal{J}$ be the complement of $\mathcal{N}$ in $C(\mathcal{M})$, so that $C(\mathcal{M}) = \mathcal{J} \cup \mathcal{N}$. For $\alpha \in \mathcal{J}$, we have either

$$P < q \leq P_0, \quad |\lambda| \leq 1/(qQ_0),$$

or

$$q \leq P, \quad 1/(qQ) < |\lambda| \leq 1/(qQ_0).$$

In either case, we have

$$N^{1/10-\varepsilon} \ll \sqrt{q(1 + |\lambda| N)} \ll N^{1/8}.$$

Therefore, Lemma 5.4 gives

$$(5.8) \quad \max_{\alpha \in \mathcal{J}} |T(\alpha)| \ll N^{2/5+\varepsilon}.$$

Collecting (5.7) and (5.8), we have

$$\max_{\alpha \in C(\mathcal{M})} |T(\alpha)| \ll N^{1/2 - 1/16 + \varepsilon}.$$

Thus the second integral in (5.5) satisfies

(5.9) $$\int_{C(\mathcal{M}) \cap \mathcal{E}_\lambda} \ll N^{-E(\lambda)} N^{2 - 1/4 + 4\varepsilon} L^k \ll N L^{k-1},$$

where we have used Lemma 5.3.

On using the definition of $\mathcal{E}_\lambda$ and Lemma 2.2, the last integral in (5.5) can be estimated as

(5.10) $$\int_{C(\mathcal{M}) \cap C(\mathcal{E}_\lambda)} \leq \{\lambda L\}^{k-4} \int_0^1 |T(\alpha) G(\alpha)|^4 \, d\alpha \leq c_3 \lambda^{k-4} \frac{\pi^2}{16} N L^k.$$

Inserting (5.6), (5.9), and (5.10) into (5.5), we get

(5.11) $$r_k(N) \geq \frac{\pi^2}{16} N L^k \left(\frac{4.952}{4} - c_3 \lambda^{k-4} \right) (1 - \varepsilon)^2,$$

when $k \geq 4$ and $N \geq N_{k,\varepsilon}$. Also when $k \geq 163$ and $\varepsilon = 10^{-8}$, one has $(4.952/4 - c_3 \lambda^{k-4})(1 - \varepsilon)^2 > 1/90$. Consequently, if $k \geq 163$ and $N \geq N_k$, then (5.11) becomes

(5.12) $$r_k(N) \geq N L^k / 200.$$

It therefore follows from (5.12) that for any $k \geq 163$, every large even integer $N \geq N_k$ with $N \equiv 4 \pmod{8}$ can be expressed in the form of (1.3).

CASE 2. Now suppose N is even but $N \not\equiv 4 \pmod{8}$. Since for any even integer N there exist $\mu_1, \mu_2 \in \{1, 2, 3\}$ such that $N - 2^{\mu_1} - 2^{\mu_2} \equiv 4 \pmod{8}$, we deduce from Case 1 that if $k \geq 165$ then every even integer $N \geq N_k + 16$ can be written in the form of (1.3), and

$$r_k(N) \geq N L^{k-2} / 200.$$

This completes the proof of Theorem 1.1.

References

[1] J. Brüdern and E. Fouvry, *Lagrange's four squares theorem with almost prime variables*, J. Reine Angew. Math. 454 (1994), 59–96.

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

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

[4] G. Greaves, *On the representation of a number in the form $x^2 + y^2 + p^2 + q^2$ where p and q are odd primes*, Acta Arith. 29 (1976), 257–274.

[5] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, London, 1974.

[6] D. R. Heath-Brown and J.-C. Puchta, *Integers represented as a sum of primes and powers of two*, Asian J. Math. 6 (2002), 535–565.

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

[8] —, *Introduction to Number Theory*, Science Press, Beijing, 1957 (in Chinese); English version: Springer, Berlin, 1982.

[9] F. B. Koval'chik, *Some analogies of the Hardy–Littlewood equation*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. 116 (1982), 86–95, 163 (in Russian).

[10] H. Z. Li, *The number of powers of 2 in a representation of large even integers by sums of such powers and of two primes*, Acta Arith. 92 (2000), 229–237.

[11] —, *The number of powers of 2 in a representation of large even integers by sums of such powers and of two primes II*, ibid. 96 (2001), 369–379.

[12] Yu. V. Linnik, *Prime numbers and powers of two*, Trudy Mat. Inst. Steklov. 38 (1951), 152–169 (in Russian).

[13] —, *Addition of prime numbers with powers of one and the same number*, Mat. Sb. (N.S.) 32 (1953), 3–60 (in Russian).

[14] J. Y. Liu, *On Lagrange's theorem with prime variables*, Quart. J. Math. Oxford 54 (2003), 453–462.

[15] J. Y. Liu and M. C. Liu, *Representation of even integers as sums of squares of primes and powers of 2*, J. Number Theory 83 (2000), 202–225.

[16] —, —, *The exceptional set in four prime squares problem*, Illinois J. Math. 44 (2000), 272–293.

[17] J. Y. Liu, M. C. Liu, and T. Z. Wang, *The number of powers of 2 in a representation of large even integers* (II), Sci. China Ser. A 41 (1998), 1255–1271.

[18] J. Y. Liu, M. C. Liu, and T. Zhan, *Squares of primes and powers of two*, Monatsh. Math. 128 (1999), 283–313.

[19] V. A. Plaksin, *An asymptotic formula for the number of solutions of a nonlinear equation with prime numbers*, Math. USSR-Izv. 18 (1982), 275–348.

[20] X. M. Ren, *On exponential sums over primes and application in the Waring–Goldbach problem*, to appear.

[21] P. Shields, *Some applications of sieve methods in number theory*, thesis, Univ. of Wales, 1979.

[22] T. Z. Wang, *On Linnik's almost Goldbach theorem*, Sci. China Ser. A 42 (1999), 1155–1172.

Department of Mathematics
Shandong University
Jinan, Shandong 250100
P.R. China
E-mail: jyliu@sdu.edu.cn
        gslv@sina.com