Nested Squares and Distinct-Prime Goldbach Partitions

Frank Vega

Information Physics Institute, 840 W 67th St, Hialeah, FL 33012, USA

Abstract

The distinct-prime Goldbach conjecture asserts that for every $N \geq 4$ there is $M \in A_N = [1, N-3]$ with $N - M$ and $N + M$ both prime; geometrically, nested squares of side lengths N and M have an L-shaped difference of semiprime area $(N - M)(N + M)$. With C_N and D_N the sets of $M \in A_N$ for which $N - M$, respectively $N + M$, is prime, the conjecture for $2N$ is $C_N \cap D_N \neq \varnothing$. We embed this centered condition in a straddling one: $C_N \cap D_N$ lies in the set E_N of half-differences $(Q - P)/2 \in A_N$ of primes $2 < P < N < Q < 2N$. From Mikawa's theorem on prime pairs with almost all differences we deduce $|A_N \setminus E_N| \ll_D N/(\log N)^D$ for every $D > 0$, and with explicit constants we prove $|E_N| > N/760$ for every $N \geq 3275$. Consequently, for all large N and all but $O_D(N/(\log N)^D)$ primes $P < N$, the number $2N - P$ becomes prime after adding the difference $P' - P$ of two primes $P' < N$; Goldbach's partition is the case $P' = P$. In contrast, $|C_N \cap D_N| < (N-3) - \log^2(2N)$ for every $N \geq 34$, and covering all but $\log^2(2N)$ values of A_N for all large N would imply the Cramér-type bound $p' - p < 2\log^2(2p + 2) + 2$. The method also shows that $2N$ is a sum of two distinct primes for more than $X/9600$ integers $N \leq X$, for every $X \geq 10^6$. Computations up to $N = 10^6$ compare the number of distinct-prime partitions with the Hardy–Littlewood prediction and measure the missing values of E_N forced by prime gaps. The conjecture itself is not proved here.

Keywords: Goldbach conjecture, differences of primes, additive energy, explicit sieve bounds, prime gaps, Cramér's conjecture, computational evidence
2020 MSC: 11P32, 51M15, 11A25

1. Introduction

The Goldbach conjecture, proposed in 1742, asserts that every even integer greater than 2 can be expressed as the sum of two prime numbers [1]. Despite centuries of effort and computational verification up to 4×10^{18} [2], the conjecture remains unproven. This paper studies a natural variant: every even integer $2N \geq 8$ is the sum of two *distinct* primes. Our variant requires the two primes to be distinct, thus excluding the trivial representations $4 = 2 + 2$ and $6 = 3 + 3$. This restriction emerges naturally from a geometric reformulation. For $N \geq 4$ (so $2N \geq 8$), finding a Goldbach partition with distinct primes is equivalent to finding $M \geq 1$ with $N - M$ and $N + M$ both prime; this symmetric form of the problem is known [3]. Geometrically, it asks for nested squares whose L-shaped difference region has area $(N - M)(N + M)$ with both factors prime. With $A_N = [1, N-3]$, let

$$C_N = \{M \in A_N : N - M \text{ is prime}\}, \qquad D_N = \{M \in A_N : N + M \text{ is prime}\}.$$

The distinct-prime problem for $2N$ is $C_N \cap D_N \neq \varnothing$, and the number of distinct-prime partitions of $2N$ is $R(2N) = |C_N \cap D_N|$ (Proposition 3.2). Our approach embeds this centered condition in the straddling condition defining the half-difference set

$$E_N = \left\{ \frac{Q - P}{2} \in A_N \;\middle|\; P, Q \text{ prime},\ 2 < P < N < Q < 2N \right\} :$$

an element M of $C_N \cap D_N$ is the half-difference of the pair $(N-M, N+M)$ centered at N, so $C_N \cap D_N \subseteq E_N$ (Lemma 3.3), while membership in E_N only requires some prime pair that straddles N. We prove that the straddling condition holds

Email address: vega.frank@gmail.com (Frank Vega)

for almost all M, derive from this a form of Goldbach's conjecture up to a difference of two primes, and isolate the centering step. Our results are the following.

(1) Coverage of A_N by straddling pairs. $|A_N \setminus E_N| \ll_D N/(\log N)^D$ for every $D > 0$ (Proposition 3.11), which we deduce from Mikawa's theorem on the Hardy–Littlewood asymptotic for prime pairs with almost all differences [4]. With explicit constants, $|E_N| > N/760$ for every $N \geq 3275$ (Theorem 3.9); the proof combines the Cauchy–Schwarz inequality, an exact identity for the additive energy of the primes below and above N (Lemma 3.7), Siebert's explicit sieve bound [5], and the estimates of Dusart [6] and of Rosser and Schoenfeld [7].

(2) Goldbach partitions up to a difference of two primes. For all large N, $|C_N \cap E_N| = \pi(N-1) - 1 - O_D(N/(\log N)^D)$; equivalently, for all but $O_D(N/(\log N)^D)$ primes $3 \leq P < N$ there are primes $2 < P' < N < Q' < 2N$ with $Q' = (2N - P) + (P' - P)$ (Corollary 3.12). A distinct-prime Goldbach partition $2N = P + Q'$ is the case $P' = P$. With explicit constants, $C_N \cap E_N \neq \varnothing$ for every $N \geq 356$ with $H(N) > 0$, where $H(N) = \log^2(2N) - |A_N \setminus E_N|$ (Proposition 5.3).

(3) The centering step. In contrast with E_N, $|C_N \cap D_N| < (N-3) - \log^2(2N)$ for every $N \geq 34$ (Proposition 5.6). The prime gap around N and the prime gap below $2N$ force missing values of E_N (Lemmas 3.4 and 3.5); consequently $H(N) > 0$ for all large N would imply $p' - p < 2\log^2(2p+2) + 2$ for consecutive primes $p < p'$ (Theorem 5.4), and the theorem of Ford, Green, Konyagin, Maynard and Tao on long prime gaps [8] shows that $|A_N \setminus E_N|$ is unbounded (Corollary 5.5).

(4) An explicit positive proportion of Goldbach numbers. The method of (1) shows that $2N$ is the sum of two distinct primes for more than $X/9600$ integers $N \leq X$, for every $X \geq 10^6$ (Theorem 3.13).

(5) Computation. For every $N \leq 10^6$ we compute E_N and $R(2N)$: $H(N) > 0$, $|E_N| \geq 0.9906\,N$ for $N \geq 3275$, the ratios of $R(2N)$ to the Hardy–Littlewood prediction, and the share of missing values of E_N forced by prime gaps, which decreases to 0.2698 on $(2^{19}, 2^{20}]$ (Section 4).

The symmetric reformulation [3], its geometric form (Section 3), and the verified range (Proposition 5.7, which combines the verification of Oliveira e Silva, Herzog and Pardi [2] with our computation) are the setting of these results. By Proposition 5.7, the distinct-prime Goldbach conjecture is equivalent to $C_N \cap D_N \neq \varnothing$ for every $N > 2 \cdot 10^{18}$, and it is not proved here. The remainder of this paper is organised as follows. Section 2 collects the analytic number theory prerequisites used in the quantitative discussion. Section 3 presents the geometric setting and the set E_N, and proves Theorems 3.9 and 3.13. Section 4 presents computational evidence. Section 5 records the rigorous consequences, the open centered problem, and the small cases. Section 6 summarises the results and lists open questions.

2. Preliminaries: Prime Distribution Results

In this section we collect the results from analytic number theory that are used in the quantitative discussion. We use explicit estimates of Dusart [6] and Rosser–Schoenfeld [7].

2.1. Primes in Short Intervals

The following result guarantees the existence of at least one prime in every sufficiently short interval. It is the principal tool for controlling the density of primes in $(N, 2N)$.

Proposition 2.1 ([6, Théorème 1.9, p. 35])**.** *For every real number $x \geq 3275$, there exists a prime p satisfying*

$$x < p \leq x\left(1 + \frac{1}{2\ln^2 x}\right).$$

Proposition 2.1 is derived from the following bound on consecutive primes.

Proposition 2.2 ([6, Proposition 1.10, p. 34])**.** *For $k \geq 463$ (equivalently, $p_k \geq 3299$), the consecutive primes p_k and p_{k+1} satisfy*

$$p_{k+1} \leq p_k\left(1 + \frac{1}{2\ln^2 p_k}\right).$$

The proof of Proposition 2.1 in [6] proceeds by using Proposition 2.2 for all $x \geq p_{463} = 3299$ and then verifying the claim computationally for $3275 \leq x < 3299$.

2.2. Bounds on the Prime-Counting Function

We use the following bounds of Dusart [6] and of Rosser and Schoenfeld [7] for $\pi(x)$, the number of primes not exceeding x.

Proposition 2.3 ([6, Théorème 1.10, p. 36])**.** *The following inequalities hold:*

(1) $\dfrac{x}{\ln x}\left(1+\dfrac{1}{\ln x}\right) \le \pi(x)$ *for all* $x \ge 599$.

(2) $\pi(x) \le \dfrac{x}{\ln x}\left(1+\dfrac{1.2762}{\ln x}\right)$ *for all* $x > 1$.

Proposition 2.4 ([7, Corollary 1, (3.5)])**.** *For every real number* $x \ge 17$*,*

$$\pi(x) > \frac{x}{\ln x}.$$

Throughout this paper, log denotes the natural logarithm (i.e., $\log \equiv \ln$).

2.3. Primes in Dyadic Intervals

We shall also use the following classical explicit form of Bertrand's postulate due to Rosser and Schoenfeld [7].

Proposition 2.5 ([7, Corollaries 3.8–3.9])**.** *For every real number* $x \ge 41/2$*,*

$$\frac{3x}{5\log x} < \pi(2x) - \pi(x) < \frac{7x}{5\log x}.$$

2.4. An Elementary Inequality

Lemma 2.6. *For every real number* $x \ge 356$*,*

$$\frac{x}{\log x + 2} - 2 > \log^2(2x).$$

Proof. Let $f(x) = \frac{x}{\log x+2} - 2 - \log^2(2x)$ for $x > 1$, and write $L = \log x$. Then

$$f'(x) = \frac{L+1}{(L+2)^2} - \frac{2(L+\log 2)}{x},$$

so $f'(x) > 0$ if and only if $x(L+1) > 2(L+\log 2)(L+2)^2$. Since $\log 2 < 1$, we have $2(L+\log 2)(L+2)^2 < 2(L+1)(L+2)^2$, and therefore $f'(x) > 0$ whenever $x > 2(L+2)^2$, that is, whenever

$$h(L) := L - \log 2 - 2\log(L+2) > 0.$$

For $L > 0$ we have $h'(L) = 1 - \frac{2}{L+2} > 0$, so h is increasing on $(0,\infty)$. At $L = \log 356$ we have $2(L+2)^2 < 124.1 < 356 = e^L$, so $h(\log 356) > 0$. Hence $h(L) > 0$ for every $L \ge \log 356$, and f is increasing on $[356,\infty)$. Finally,

$$\frac{356}{\log 356 + 2} - 2 > 43.206 \quad \text{and} \quad \log^2(712) < 43.140,$$

so $f(356) > 0$, and therefore $f(x) > 0$ for every $x \ge 356$. □

3. Geometric Construction and Reformulation

We now develop the geometric framework used in the reformulation.

3.1. Nested Squares and Semiprime Areas

Consider a square S_N with integer side length $N \geq 4$, having area N^2. Inside S_N, inscribe a smaller square S_M with side length M, where $1 \leq M \leq N-3$, sharing the bottom-left corner with S_N. The region between S_N and S_M forms an L-shaped annulus with area

$$N^2 - M^2 = (N-M)(N+M).$$

Define $P = N - M$ and $Q = N + M$. The bounds on M translate to constraints on P and Q: $M \geq 1$ gives $P \leq N-1$ and $Q \geq N+1$, while $M \leq N-3$ gives $P \geq 3$. Thus $3 \leq P \leq N-1$ and $Q \geq N+1$, with $P < Q$ since $M \geq 1$.

3.2. Connection to Goldbach Partitions

The sum and difference of P and Q are:

$$P + Q = (N-M) + (N+M) = 2N \geq 8,$$
$$Q - P = (N+M) - (N-M) = 2M.$$

Since both the sum and difference are even, P and Q have the same parity. For both to be prime with $P \geq 3$, they must both be odd primes, hence distinct. The area $N^2 - M^2 = P \cdot Q$ is a semiprime (product of exactly two primes) if and only if both P and Q are prime.

3.3. The Geometric Equivalence

Theorem 3.1 (Geometric Goldbach Variant)**.** *The following are equivalent for all $N \geq 4$:*

(i) *The even integer $2N$ can be written as the sum of two distinct primes.*

(ii) *There exists $M \in [1, N-3]$ such that $P = N - M$ and $Q = N + M$ are both prime.*

(iii) *The L-shaped region between squares S_N and S_M (sharing a corner) has area $(N-M)(N+M)$ that is a semiprime for some $M \in [1, N-3]$.*

Proof. **(i) ⇒ (ii):** If $2N = p + q$ with distinct primes $p < q$, set $M = (q-p)/2$. Since p and q are distinct odd primes (as $2N \geq 8$), both $(q+p)/2 = N$ and $(q-p)/2 = M$ are positive integers. We have $P = N - M = p$ and $Q = N + M = q$, both prime. To verify $M \in [1, N-3]$: distinctness gives $q - p \geq 2$, so $M \geq 1$; and $M \leq N-3$ is equivalent to $p \geq 3$, which holds since p is an odd prime. **(ii) ⇒ (iii):** Immediate, as $N^2 - M^2 = (N-M)(N+M)$ with both factors prime. **(iii) ⇒ (i):** If $N^2 - M^2 = (N-M)(N+M)$ is a semiprime and $M \in [1, N-3]$, then both factors $N - M$ and $N + M$ are greater than 1. Hence the two displayed factors must be prime, and their sum is $2N$. Since $M \geq 1$, they are distinct. □

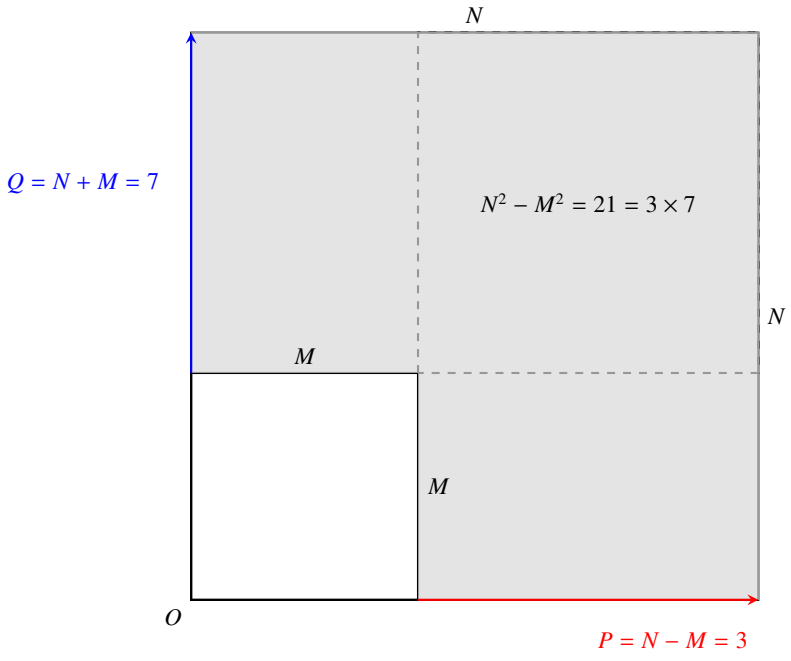


Figure 1: The geometric construction for $N = 5$, $M = 2$: The L-shaped region has area $N^2 - M^2 = 25 - 4 = 21 = 3 \times 7$, a semiprime. The factors $P = 3$ and $Q = 7$ are both prime and sum to $2N = 10$, providing the Goldbach partition $10 = 3 + 7$.

3.4. Reformulation as a Centered Set Intersection Problem

Let

$$A_N = \{1, 2, \ldots, N - 3\}$$

be the admissible range for M. For each $N \geq 4$, define two centered subsets of A_N:

- **Lower-prime candidate set:** $C_N = \{M \in A_N \mid N - M \text{ is prime}\}$.
- **Upper-prime candidate set:** $D_N = \{M \in A_N \mid N + M \text{ is prime}\}$.

Both centered sets lie in a single parity class. Indeed, if $N - M$ or $N + M$ is an odd prime, then M has parity opposite to N. Thus

$$C_N, D_N \subseteq B_N := \{M \in A_N \mid M \not\equiv N \pmod 2\}.$$

Proposition 3.2 (Centered intersection criterion)**.** *The distinct-prime Goldbach conjecture holds for the even integer $2N$ ($N \geq 4$) if and only if $C_N \cap D_N \neq \emptyset$.*

Proof. ($\Rightarrow$) Suppose $2N = p + q$ with distinct primes $p < q$. Set $M = (q - p)/2$. As in the proof of Theorem 3.1, $M \in A_N$, $p = N - M$, and $q = N + M$. Hence $M \in C_N \cap D_N$.

($\Leftarrow$) Suppose $M \in C_N \cap D_N$. Then $N - M$ is prime by the definition of C_N, and $N + M$ is prime by the definition of D_N. Their sum is $2N$, and they are distinct because $M \geq 1$. □

3.5. The Half-Difference Diagnostic Set

It is useful computationally to compare the centered sets with a set of half-differences arising from straddling prime pairs. Define

$$E_N = \left\{ \frac{Q - P}{2} \in A_N \;\middle|\; 2 < P < N < Q < 2N,\ P, Q \text{ prime} \right\}.$$

The restriction $\frac{Q-P}{2} \in A_N$ is part of the definition, so every listed element of E_N lies in $\{1, \ldots, N-3\}$. The set E_N measures the availability of prime-pair differences near N, while D_N measures primality of the centered upper endpoint $N + M$. The proof target remains the centered intersection $C_N \cap D_N$.

Lemma 3.3. *For every $N \geq 4$,*

$$C_N \cap D_N \subseteq E_N.$$

Consequently, $C_N \cap E_N \cap D_N = C_N \cap D_N$.

Proof. Let $M \in C_N \cap D_N$, and set $P = N - M$ and $Q = N + M$. Both are prime, and $M \in A_N$ gives $P \geq 3 > 2$, $P < N < Q$ and $Q \leq 2N - 3 < 2N$. Since $M = \frac{Q-P}{2} \in A_N$, we get $M \in E_N$. The second statement follows by intersecting both sides with C_N. □

Lemma 3.4 (Gap obstruction)**.** *Let $N \geq 4$, let p^- be the largest prime below N, and let p^+ be the smallest prime above N. If $M \in E_N$, then $2M \geq p^+ - p^-$. Hence every $M \in A_N$ with $M < (p^+ - p^-)/2$ lies in $A_N \setminus E_N$.*

Proof. If $M = (Q - P)/2$ with primes $P < N < Q$, then $P \leq p^-$ and $Q \geq p^+$, so $2M = Q - P \geq p^+ - p^-$. □

Lemma 3.5 (Upper gap obstruction)**.** *Let $N \geq 4$, and let q be the largest prime below $2N$. If $M \in E_N$, then $2M \leq q - 3$. Hence every $M \in A_N$ with $M > (q-3)/2$ lies in $A_N \setminus E_N$; if $q \leq 2N - 3$, there are $(2N - 3 - q)/2$ such values.*

Proof. If $M = (Q - P)/2$ with primes $2 < P$ and $Q < 2N$, then $P \geq 3$ and $Q \leq q$, so $2M \leq q - 3$. The values $M \in A_N$ with $M > (q-3)/2$ are $M = (q-3)/2 + 1, \ldots, N - 3$, and there are $N - 3 - (q-3)/2 = (2N - 3 - q)/2$ of them. □

3.6. Nearby Even Numbers Represented by E_N

Each element of E_N arises from a Goldbach partition of a nearby even number. If $M \in E_N$, then there exist odd primes P, Q such that

$$2 < P < N < Q < 2N, \qquad M = \frac{Q - P}{2},$$

and the even number

$$S = P + Q$$

has a Goldbach partition with half-difference M. Conversely, any Goldbach partition $S = P + Q$ satisfying these inequalities and $\frac{Q-P}{2} \in A_N$ contributes its half-difference to E_N.

The possible sums S lie in the even part of the interval

$$N + 4 \leq S \leq 3N - 2,$$

with endpoint adjustments forced by parity and by the primality of P and Q. If one excludes the endpoint possibility $Q = 2N - 1$, the upper endpoint becomes $3N - 4$ when N is even. The centered Goldbach problem for $2N$ is the special case $S = 2N$, where the partition has the form

$$2N = (N - M) + (N + M).$$

3.7. Lower and Upper Primes

Let

$$L_N = \{P \mid 3 \leq P < N,\ P \text{ prime}\}, \qquad a_N = |L_N| = \pi(N-1) - 1,$$

and let

$$U_N = \{Q \mid N < Q \leq 2N - 3,\ Q \text{ prime}\} = \{Q_1 > Q_2 > \cdots > Q_r\}.$$

If $2N - 3$ is prime, then $2N = 3 + (2N - 3)$ is already a centered Goldbach partition with $M = N - 3$. For estimates of r, the only possible prime counted by $\pi(2N) - \pi(N)$ but not by $\pi(2N - 3) - \pi(N)$ is $2N - 1$. Hence

$$\pi(2N) - \pi(N) - 1 \leq r \leq \pi(2N) - \pi(N).$$

Proposition 2.5 therefore gives, for $N \geq 41/2$,

$$\frac{3N}{5\log N} - 1 < r < \frac{7N}{5\log N}.$$

For $P \in L_N$ and $Q \in U_N$ we have $3 \leq P < N < Q \leq 2N-3$, so $1 \leq (Q-P)/2 \leq N-3$; hence $(Q-P)/2 \in E_N$ for every pair $(P,Q) \in L_N \times U_N$. For $d > 0$ define

$$\rho_L(d) = \#\{(P,P') \in L_N^2 \mid P' - P = d\}, \qquad \rho_U(d) = \#\{(i,j) \mid i < j,\ Q_i - Q_j = d\},$$

and the additive energy

$$C_N = \sum_{d>0} \rho_L(d)\rho_U(d),$$

which counts the quadruples (P, P', Q_i, Q_j) with $P, P' \in L_N$, $i < j$ and $P' - P = Q_i - Q_j$.

3.8. A Linear Lower Bound for the Half-Difference Set

The additive energy C_N of Section 3.7 controls $|E_N|$ from below, through the Cauchy–Schwarz inequality. We use the following explicit sieve bound of Siebert, stated in [5] for primes p with $ap + b$ prime, $2 \mid ab$; we take $a = 1$ and $b = h$.

Proposition 3.6 ([5])**.** *Let*

$$\omega = \prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right), \qquad \kappa(h) = \prod_{\substack{p \mid h \\ p>2}} \frac{p-1}{p-2}.$$

For every real number $x > 1$ and every even integer $h \neq 0$,

$$\#\{p \leq x \mid p \text{ and } p+h \text{ are prime}\} \leq 16\,\omega\,\kappa(h)\,\frac{x}{\log^2 x}.$$

For $M \in A_N$ define

$$\nu_N(M) = \#\{(P,Q) \in L_N \times U_N \mid Q - P = 2M\}.$$

Every pair $(P,Q) \in L_N \times U_N$ gives $(Q-P)/2 \in E_N$, as shown in Section 3.7. Hence ν_N vanishes outside E_N, and

$$\sum_{M \in E_N} \nu_N(M) = a_N\, r.$$

Lemma 3.7 (Energy identity)**.** *For every $N \geq 4$,*

$$\sum_{M \in A_N} \nu_N(M)^2 = a_N\, r + 2\,C_N.$$

Proof. The left-hand side counts the quadruples $(P,Q,P',Q') \in L_N \times U_N \times L_N \times U_N$ with $Q - P = Q' - P'$, that is, with $P' - P = Q' - Q =: h$. The value $h = 0$ gives $P = P'$ and $Q = Q'$, hence $a_N r$ quadruples. For $h > 0$ the number of quadruples is $\rho_L(h)\rho_U(h)$, and exchanging (P,Q) with (P',Q') maps the quadruples with difference h bijectively onto those with difference $-h$. □

Lemma 3.8 (Energy bound)**.** *Let*

$$G = \prod_{p>2}\left(1 + \frac{2p-3}{p(p-2)^2}\right).$$

For every $N \geq 4$,

$$C_N \leq 256\,\omega^2 G\,\frac{N^3}{\log^4 N}.$$

Moreover, $\omega < 0.6614$ and $G < 2.658$.

Proof. All primes in $L_N \cup U_N$ are odd, so $\rho_L(h) = 0$ unless h is even with $2 \le h \le N-4$. For such h, Proposition 3.6 with $x = N$ and with $x = 2N$ gives

$$\rho_L(h) \le 16\,\omega\,\kappa(h)\frac{N}{\log^2 N}, \qquad \rho_U(h) \le 16\,\omega\,\kappa(h)\frac{2N}{\log^2(2N)} \le 32\,\omega\,\kappa(h)\frac{N}{\log^2 N},$$

since every pair counted by $\rho_U(h)$ consists of primes $Q < Q + h \le 2N$. Hence

$$C_N \le 512\,\omega^2\frac{N^2}{\log^4 N}\sum_{1\le k\le N/2}\kappa(2k)^2.$$

Since κ depends only on the odd prime divisors, $\kappa(2k) = \kappa(k)$. The function κ^2 satisfies $\kappa(k)^2 = \sum_{d\mid k} g(d)$, where g is the multiplicative function supported on odd squarefree integers with

$$g(p) = \left(\frac{p-1}{p-2}\right)^2 - 1 = \frac{2p-3}{(p-2)^2} \qquad (p > 2).$$

Since $g \ge 0$, for every $K \ge 1$

$$\sum_{k\le K}\kappa(k)^2 = \sum_{d\le K} g(d)\left\lfloor\frac{K}{d}\right\rfloor \le K\sum_{d\ge 1}\frac{g(d)}{d} = K\,G.$$

Taking $K = N/2$ gives the bound for C_N.

Every factor of ω lies in $(0,1)$, so ω is at most the product of its factors with $3 \le p \le 97$, which is less than 0.66138. For G, the product of its factors with $3 \le p \le 97$ is less than 2.63013. For $p \ge 101$ we have $\frac{2p-3}{p(p-2)^2} < \frac{2}{(p-2)^2}$ and $p-2$ is odd, so, using $\log(1+y) \le y$,

$$\begin{aligned}\log\prod_{p\ge 101}\left(1+\frac{2p-3}{p(p-2)^2}\right) &< 2\sum_{\substack{m\ge 99\\ m\text{ odd}}}\frac{1}{m^2} \le 2\left(\frac{1}{99^2}+\frac{1}{2}\int_{99}^{\infty}\frac{dt}{t^2}\right)\\ &= 2\left(\frac{1}{9801}+\frac{1}{198}\right) < 0.01031.\end{aligned}$$

Hence $G < 2.63013\,e^{0.01031} < 2.658$. □

Theorem 3.9 (Linear lower bound)**.** *For every integer $N \ge 3275$,*

$$|E_N| > \frac{N}{760}.$$

Proof. Write $L = \log N$ and $\lambda = \log 2$. By the Cauchy–Schwarz inequality and Lemma 3.7,

$$(a_N r)^2 = \Big(\sum_{M\in E_N}\nu_N(M)\Big)^2 \le |E_N|\sum_{M\in A_N}\nu_N(M)^2 = |E_N|\,(a_N r + 2C_N).$$

Lower bound for $a_N r$. By Proposition 2.4 with $x = N-1 \ge 17$, $a_N = \pi(N-1) - 1 > \frac{N-1}{L} - 1$, so

$$a_N > \frac{N}{L}\,\mathcal{A}(N), \qquad \mathcal{A}(N) = 1 - \frac{1+L}{N}.$$

As shown in Section 3.7, $r \ge \pi(2N) - \pi(N) - 1$. Proposition 2.3*(1)* with $x = 2N \ge 599$ and Proposition 2.3*(2)* with $x = N$ give

$$r \ge \frac{2N}{L+\lambda}\left(1+\frac{1}{L+\lambda}\right) - \frac{N}{L}\left(1+\frac{1.2762}{L}\right) - 1 = \frac{N}{L}\,\mathcal{B}(N),$$

where, using $\frac{2L}{L+\lambda} = 2 - \frac{2\lambda}{L+\lambda}$,

$$\mathcal{B}(N) = 1 - \frac{2\lambda}{L+\lambda} + \frac{2L}{(L+\lambda)^2} - \frac{1.2762}{L} - \frac{L}{N}.$$

Both $\mathcal{A}$ and $\mathcal{B}$ are increasing in N for $N \geq 3275$. For $\mathcal{A}$ this holds because $(1+\log x)/x$ is decreasing for $x > 1$. For $\mathcal{B}$, the term $-L/N$ is increasing for $N > e$, and the derivative with respect to L of the remaining terms is

$$\begin{aligned}\frac{2\lambda(\lambda+1)-(2-2\lambda)L}{(L+\lambda)^3}+\frac{1.2762}{L^2} &> -\frac{(2-2\lambda)L}{L^3}+\frac{1.2762}{L^2}\\ &= \frac{1.2762-(2-2\lambda)}{L^2} > 0,\end{aligned}$$

since $2-2\lambda < 0.614$; here the first numerator is negative for $L \geq \log 3275$. At $N = 3275$, $\mathcal{A} > 0.99722$ and $\mathcal{B} > 0.89174$. Hence, for every $N \geq 3275$,

$$a_N r > \alpha \frac{N^2}{L^2}, \qquad \alpha = 0.8892.$$

Conclusion. The function $t \mapsto t^2/(t+K)$ is increasing for $t > 0$, and decreasing in $K > 0$. With $X = \alpha N^2/L^2$ and, by Lemma 3.8, $2C_N \leq \overline{K} = \beta N^3/L^4$, where $\beta = 512 \cdot 0.6614^2 \cdot 2.658 < 595.33$, we obtain

$$|E_N| \geq \frac{(a_N r)^2}{a_N r + 2C_N} > \frac{X^2}{X+\overline{K}} = \frac{\alpha^2 N}{\beta(1+\delta_N)}, \qquad \delta_N = \frac{X}{\overline{K}} = \frac{\alpha L^2}{\beta N}.$$

The function $t \mapsto \log^2 t/t$ is decreasing for $t > e^2$, so $\delta_N \leq \delta_{3275} < 3 \cdot 10^{-5}$. Therefore

$$|E_N| > \frac{0.8892^2}{595.33 \cdot 1.00003} N > 0.001328\, N > \frac{N}{760}.$$

□

By definition, $M \in E_N$ if and only if there is a prime p with $\max\{2, N-2M\} < p < N$ such that $p+2M$ is a prime smaller than $2N$. For a fixed shift $2k$, the number of primes $p \leq y$ with $p+2k$ prime is conjectured by Hardy and Littlewood to be asymptotic to $2\omega\kappa(2k)\, y/\log^2 y$, and this asymptotic is known for almost all shifts. Wolke [9] proved it for almost all $k \leq x$ and all $y \in [2x, x^{8/5-\varepsilon}]$, improving earlier results of Lavrik [10]. Mikawa [4] extended the range of y; his theorem is stated below in the form that we use. Perelli and Pintz [11] proved the asymptotic for all but $O(H(\log N)^{-B})$ shifts $2k$ with k in intervals of length H, $N^{1/2+\varepsilon} \leq H \leq N/4$. Matomäki, Radziwiłł and Tao [12] proved that $\sum_{X<n\leq 2X} \Lambda(n)\Lambda(n+h)$ has the Hardy–Littlewood asymptotic for all but $O(H(\log X)^{-A})$ values of h with $|h-h_0| \leq H$, for $X^{8/33+\varepsilon} \leq H \leq X^{1-\varepsilon}$ and $0 \leq h_0 \leq X^{1-\varepsilon}$. These results are stated with unspecified constants.

Let

$$\Psi(y, 2k) = \sum_{2k<n\leq y} \Lambda(n)\Lambda(n-2k),$$

where Λ is the von Mangoldt function, and let $\mathfrak{S}(2k) = 2\omega\kappa(2k)$ be the singular series, so that $\mathfrak{S}(2k) \geq 2\omega$.

Proposition 3.10 ([4])**.** *Let $\varepsilon, A, B > 0$. For all sufficiently large x and every y with $2x \leq y \leq x^{3-\varepsilon}$, all but $O(x(\log x)^{-A})$ integers k with $1 \leq k \leq x$ satisfy*

$$\Psi(y, 2k) = \mathfrak{S}(2k)(y-2k) + O\left(\frac{y}{(\log y)^B}\right),$$

where the implied constants depend only on ε, A and B.

We use Proposition 3.10 only for $2x \leq y \leq 4x$.

Proposition 3.11 (Almost all half-differences)**.** *For every $D > 0$,*

$$|A_N \setminus E_N| \ll_D \frac{N}{(\log N)^D}.$$

In particular, $|E_N| = N - o(N)$.

Proof. Write $L = \log N$ and $a = \lceil N/L^D \rceil$, and apply Proposition 3.10 with $\varepsilon = 1/2$, $A = D$ and $B = D+1$. The terms of $\Psi(y, 2k)$ with $y \le 2N$ in which n or $n-2k$ is a prime power p^j with $j \ge 2$ contribute $O(\sqrt{N}L^3)$, since there are $O(\sqrt{N}L)$ such prime powers up to $2N$ and each term is at most $L^2 + O(L)$.

The range $a < k \le N/2$. Take $x = N/2$ and the two values $y_1 = N$ and $y_2 = N + 2a$, which satisfy $2x \le y_1 < y_2 \le 4x$. Outside a set of $O(NL^{-D})$ integers $k \le x$, Proposition 3.10 applies to both values, and

$$\begin{aligned}\Psi(y_2, 2k) - \Psi(y_1, 2k) &= \sum_{N < n \le N+2a} \Lambda(n)\Lambda(n-2k) \\ &= 2a\, \mathfrak{S}(2k) + O\Big(\frac{N}{L^{D+1}}\Big) \ge 4\omega a - O\Big(\frac{N}{L^{D+1}}\Big),\end{aligned}$$

which exceeds the prime-power contribution $O(\sqrt{N}L^3)$ for large N. Hence there is an integer n with $N < n \le N + 2a$ such that $Q = n$ and $P = n - 2k$ are both prime. Then $N < Q < 2N$, and $P \le N + 2a - 2k < N$ because $k > a$; moreover P is odd, because Q is odd and $2k$ is even, so $P > 2$. Thus $k \in E_N$.

The range $N/2 < k \le N - 3 - a$. Take $x = N - 1$ and $y = 2N - 1$, so that $2x \le y \le 4x$. Outside a set of $O(NL^{-D})$ integers $k \le x$,

$$\Psi(2N-1, 2k) = \mathfrak{S}(2k)(2N - 1 - 2k) + O\Big(\frac{N}{L^{D+1}}\Big) \ge 2\omega(2a+5) - O\Big(\frac{N}{L^{D+1}}\Big),$$

which exceeds $O(\sqrt{N}L^3)$ for large N. Hence there is an integer n with $2k < n \le 2N - 1$ such that $Q = n$ and $P = n - 2k$ are both prime. Then $N < 2k < Q < 2N$ and $P \le 2N - 1 - 2k < N$, and $P > 2$ as before. Thus $k \in E_N$.

Conclusion. The remaining integers $k \in A_N$ satisfy $k \le a$ or $k > N - 3 - a$, and there are at most $2a$ of them. Hence $|A_N \setminus E_N| \le 2a + O(NL^{-D}) \ll_D N/L^D$. □

Corollary 3.12 (Goldbach partitions up to a difference of two primes)**.** *For every $D > 0$,*

$$|C_N \cap E_N| = \pi(N-1) - 1 + O_D\Big(\frac{N}{(\log N)^D}\Big).$$

Equivalently, for all but $O_D(N/(\log N)^D)$ primes P with $3 \le P < N$ there are primes P' and Q' with $2 < P' < N < Q' < 2N$ and

$$Q' = (2N - P) + (P' - P).$$

In particular, $C_N \cap E_N \ne \emptyset$ for all sufficiently large N. The case $P' = P$ is a distinct-prime Goldbach partition $2N = P + Q'$.

Proof. The map $M \mapsto N - M$ is a bijection from C_N onto the set of primes P with $3 \le P \le N-1$, so $|C_N| = \pi(N-1) - 1$. Since $|C_N| - |A_N \setminus E_N| \le |C_N \cap E_N| \le |C_N|$, Proposition 3.11 gives the first statement. For a prime $3 \le P < N$, the value $M = N - P$ lies in A_N, and $M \in E_N$ if and only if $2(N-P) = Q' - P'$ for primes $2 < P' < N < Q' < 2N$, that is, $Q' = (2N - P) + (P' - P)$. Finally, $\pi(N-1) - 1 > (N-1)/\log N - 1$ by Proposition 2.4, which exceeds the error term for large N. □

Proposition 3.11 is deduced from Proposition 3.10 and gives no explicit value of N from which a stated bound holds. Theorem 3.9 gives the explicit bound $|E_N| > N/760$ for every $N \ge 3275$.

3.9. A Positive Proportion of Centered Partitions

The same energy method applies to the centered count

$$R(2N) = |C_N \cap D_N| = \#\{p \mid 3 \le p < N,\ p \text{ and } 2N - p \text{ prime}\}.$$

Theorem 3.13 (Positive proportion)**.** *For every real number $X \ge 10^6$,*

$$\#\{N \in \mathbb{Z} \mid 4 \le N \le X,\ R(2N) \ge 1\} > \frac{X}{9600}.$$

Proof. Write $L = \log X$ and $s = X/L$, and let S be the set of odd primes not exceeding $2X$. Let

$$T = \sum_{4 \le N \le X} R(2N), \qquad V = \sum_{4 \le N \le X} R(2N)^2.$$

By the Cauchy–Schwarz inequality, the number of N with $R(2N) \ge 1$ is at least T^2/V.

Lower bound for T. Every pair of odd primes $p < q \le X$ gives $N = (p+q)/2$ with $4 \le N < X$ and $M = N - p \in C_N \cap D_N$, and distinct pairs give distinct pairs (N, p) counted by $R(2N)$. Hence, by Proposition 2.4,

$$T \ge \binom{\pi(X) - 1}{2} > \frac{(s-2)^2}{2}.$$

Upper bound for V. The square $R(2N)^2$ counts pairs of representations $2N = p + q = p' + q'$ with $p < q$ and $p' < q'$, so V is at most the number of quadruples $(p, q, p', q') \in S^4$ with $p + q = p' + q'$, that is, with $p' - p = q - q' =: h$. With $\rho_S(h) = \#\{(p, p') \in S^2 \mid p' - p = h\}$, this number is $\sum_h \rho_S(h)^2$. The value $h = 0$ contributes $|S|^2 \le \pi(2X)^2$. For even $h \ne 0$ with $|h| \le 2X$, Proposition 3.6 with $x = 2X$ gives

$$\rho_S(h) \le 16\,\omega\,\kappa(h)\frac{2X}{\log^2(2X)} \le 32\,\omega\,\kappa(h)\frac{X}{L^2},$$

and $\rho_S(h) = 0$ for odd h and for $|h| > 2X$. As in the proof of Lemma 3.8, $\sum_{2 \le h \le 2X,\, h \text{ even}} \kappa(h)^2 \le XG$. Hence

$$V \le \pi(2X)^2 + 2 \cdot 1024\,\omega^2 \frac{X^2}{L^4}\,XG = \pi(2X)^2 + K\frac{X^3}{L^4}, \qquad K = 2048\,\omega^2 G.$$

By Proposition 2.3*(2)*, and since $y \mapsto \frac{1}{y}(1 + \frac{1.2762}{y})$ is decreasing, $\pi(2X) \le \frac{2X}{L}(1 + \frac{1.2762}{L})$.

Conclusion. Since $X^3/L^4 = s^4/X$,

$$\frac{T^2}{V} > \frac{(s-2)^4/4}{4s^2(1 + \frac{1.2762}{L})^2 + Ks^4/X} = \frac{X\,(1 - 2/s)^4}{4(K + e(X))},$$

where $e(X) = \frac{4L^2}{X}(1 + \frac{1.2762}{L})^2$. For $X \ge 10^6$, s is increasing and $e(X)$ is decreasing, so $(1 - 2/s)^4 > 0.99988$ and $e(X) < 0.00092$. By Lemma 3.8, $K < 2048 \cdot 0.6614^2 \cdot 2.658 < 2381.3$. Therefore

$$\frac{T^2}{V} > \frac{0.99988}{4 \cdot 2381.30092}\,X > \frac{X}{9600}.$$

□

Montgomery and Vaughan [13] proved that, for some $\delta > 0$ and all sufficiently large X, at most $X^{1-\delta}$ even integers $n \le X$ are not sums of two primes. Theorem 3.13 is weaker in its conclusion, but it is explicit and holds for every $X \ge 10^6$. By Proposition 5.7, its conclusion is trivial for $X \le 2 \cdot 10^{18}$.

3.10. A Half-Difference Coverage Statistic

For the diagnostic set E_N, define

$$H(N) = \log^2(2N) - ((N-3) - |E_N|).$$

Rearranging, $H(N) > 0$ is equivalent to

$$|E_N| > (N-3) - \log^2(2N).$$

Thus $H(N) > 0$ holds exactly when fewer than $\log^2(2N)$ admissible values are missing from E_N.

Hypothesis 3.14 (Half-Difference Density Hypothesis)**.** *For every integer $N > 10^6$, we have $H(N) > 0$.*

Hypothesis 3.15 (Centered Prime-Correlation Hypothesis)**.** *For every integer $N > 2 \cdot 10^{18}$, the centered prime-pair count*

$$R(2N) := |C_N \cap D_N|$$

is positive.

Remark 3.16 (Status of the hypotheses)**.** By Theorem 5.4 and the computation of Section 4, Hypothesis 3.14 implies the Cramér-type bound $p' - p < 2\log^2(2p+2) + 2$ for all consecutive primes $p < p'$ with $p \geq 3$. Hypothesis 3.14 is sufficient for the pigeonhole conclusion $C_N \cap E_N \neq \emptyset$ by Proposition 5.3. It is not by itself sufficient for the centered Goldbach partition; that requires Hypothesis 3.15. By Propositions 3.2 and 5.7, Hypothesis 3.15 is equivalent to the distinct-prime Goldbach conjecture; it is an open problem, and it is not used to derive any result of this paper.

4. Computational Evidence

The computations use two scripts of the repository [14]: `experiment.py`, which uses only the Python 3.12 standard library, and `large_range.py`, which also uses NumPy. The scripts were written with Claude (Claude Opus 5.5) and run by the author; the running times below are machine dependent.

The script `experiment.py` computes $|E_N|$, $H(N)$, $|C_N|$, $R(2N) = |C_N \cap D_N|$ and the sets $(C_N \cap E_N) \setminus D_N$ for all $N \in [4, 2^{14}]$, with a sieve of Eratosthenes up to 2^{15} and bit-mask shifts that form the half-differences of E_N; the run took 18.7 s. The script `large_range.py` computes $|E_N|$ and $R(2N)$ for all $N \in [4, 10^6]$, in 184.4 s. It maintains the number of pairs (P, Q) with each difference $Q - P$ as N increases, and it obtains $R(2N)$ from one fast Fourier transform of the indicator function of the odd primes. On $[4, 2^{14}]$ the two scripts give the same minima of $H(N)$. The tested ranges are split into the disjoint dyadic intervals

$$I_2 = [4, 8], \qquad I_m = (2^m, 2^{m+1}] \quad (m \geq 3).$$

Table 1 gives, for $2 \leq m \leq 19$, the minimum of $H(N)$ on I_m, the largest exceptional set $A_N \setminus E_N$ on I_m, and the minimum of $|E_N|/N$ on I_m. The data show three features. First, $H(N) > 0$ for *every* $N \in [4, 10^6]$, providing empirical support for Hypothesis 3.14. Second, the minimum of $H(N)$ on I_m strictly increases for $2 \leq m \leq 16$; the minimum on I_{17} is smaller than the minimum on I_{16}, and the minima increase again from I_{17} to I_{19}. Third, the largest exceptional set on I_{19} has 140 elements, at $N = 668364$, where $\log^2(2N) > 198$.

Table 1: Minimum $H(N)$, largest exceptional set $A_N \setminus E_N$ (with the N attaining it), and minimum $|E_N|/N$ on the dyadic intervals I_m. Note that $H(N) > 0$ for all tested values.

m	**Range** I_m	N **at min** H	**Min** $H(N)$	**Max** $\lvert A_N \setminus E_N \rvert$ **(at** N**)**	**Min** $\lvert E_N \rvert / N$
2	$[4, 8]$	5	4.301898	2 (7)	0.200000
3	$(2^3, 2^4]$	9	7.354249	2 (11)	0.545455
4	$(2^4, 2^5]$	19	10.232033	6 (29)	0.684211
5	$(2^5, 2^6]$	61	14.078618	9 (61)	0.803279
6	$(2^6, 2^7]$	73	17.836335	10 (113)	0.863014
7	$(2^7, 2^8]$	151	20.608977	13 (165)	0.900662
8	$(2^8, 2^9]$	269	23.537165	16 (269)	0.929368
9	$(2^9, 2^{10}]$	541	28.812111	24 (986)	0.957486
10	$(2^{10}, 2^{11}]$	1327	33.154668	29 (1327)	0.975377
11	$(2^{11}, 2^{12}]$	2161	35.081569	35 (2161)	0.982416
12	$(2^{12}, 2^{13}]$	7069	42.329014	49 (7069)	0.991389
13	$(2^{13}, 2^{14}]$	14138	44.057758	61 (14138)	0.994665
14	$(2^{14}, 2^{15}]$	17863	46.906582	64 (29443)	0.996305
15	$(2^{15}, 2^{16}]$	43331	55.271687	74 (43331)	0.998006
16	$(2^{16}, 2^{17}]$	66239	58.102490	90 (107525)	0.998732
17	$(2^{17}, 2^{18}]$	206699	54.240920	113 (206699)	0.999315
18	$(2^{18}, 2^{19}]$	484853	55.019283	135 (484853)	0.999576
19	$(2^{19}, 2^{20}]$	668364	58.971771	140 (668364)	0.999749

4.1. *Centered Partitions and the Pigeonhole Count*

The run of `experiment.py` gives the following results for every $N \in [4, 2^{14}]$.

(1) $R(2N) \geq 1$. Hence every even integer $2N$ with $8 \leq 2N \leq 2^{15}$ is the sum of two distinct primes, and, by Lemma 3.3, $C_N \cap E_N \supseteq C_N \cap D_N \neq \emptyset$.

(2) The pigeonhole count $|C_N| > |A_N \setminus E_N|$ holds for every N except $N = 5$ and $N = 7$. For these two values, $C_5 \cap E_5 = \{2\}$ and $C_7 \cap E_7 = \{4\}$.

(3) $(C_N \cap E_N) \setminus D_N \neq \emptyset$ for 16,377 of the 16,381 values of N, namely for every N except $N \in \{4, 5, 7, 8\}$. The first elements are $M = 3$ for $N = 6$, $M = 6$ for $N = 9$, $M = 5$ for $N = 10$, $M = 4$ for $N = 11$ and $M = 9$ for $N = 12$; for instance, $3 \in C_6 \cap E_6$, while $6 + 3 = 9$ is composite.

(4) Among the integers $2 \leq N \leq 10^6$, the inequality of Lemma 2.6 fails at $N = 355$ and holds for every $N \geq 356$; the bound 356 of Lemma 2.6 is therefore the least possible integer threshold.

For comparison with $H(N)$, the minima of the centered statistic $H_R(N)$ of Section 5.3 on the intervals I_m are given in Table 2. The minimum is positive on I_2 and I_3, and negative on every interval I_m with $4 \leq m \leq 13$; by Proposition 5.6, $H_R(N) < 0$ for every $N \geq 34$.

Table 2: Minimum $H_R(N)$ values on the dyadic intervals I_m.

Interval (m)	**Range** I_m	N **achieving min**	**Min** $H_R(N)$
2	$[4, 8]$	7	3.96
3	$(2^3, 2^4]$	16	1.01
4	$(2^4, 2^5]$	31	−8.97
5	$(2^5, 2^6]$	64	−34.46
6	$(2^6, 2^7]$	128	−86.25
7	$(2^7, 2^8]$	256	−203.08
8	$(2^8, 2^9]$	511	−441.98
9	$(2^9, 2^{10}]$	1024	−937.87
10	$(2^{10}, 2^{11}]$	2047	−1923.82
11	$(2^{11}, 2^{12}]$	4096	−3935.80
12	$(2^{12}, 2^{13}]$	8191	−7953.83
13	$(2^{13}, 2^{14}]$	16369	−16029.92

4.2. *The Exceptional Set and the Linear Bound*

The minimum of $|E_N|/N$ over $3275 \leq N \leq 10^6$ is 0.990603; Theorem 3.9 gives $|E_N|/N > 1/760$ for every $N \geq 3275$, and Proposition 3.11 gives $|E_N|/N \to 1$. For each N in Table 1 at which the minimum of $H(N)$ is attained, the script lists the missing values $A_N \setminus E_N$. For $N \leq 2161$ the complete lists show that every missing value satisfies $M \leq 50$ or $N - 3 - M \leq 14$. For instance, for $N = 1327$ the missing values are $M = 1, \ldots, 19, 22, 25, 28, 34, 43, 50, 1316, 1319, 1323$ and 1324. The primes around 1327 are $p^- = 1321$ and $p^+ = 1361$, so Lemma 3.4 forces exactly the missing values $M \leq 19$ below $(p^+ - p^-)/2 = 20$. Likewise, the gaps $p^+ - p^- = 26$ around $N = 2161$ and $p^+ - p^- = 24$ around $N = 541$ force the missing values $M = 1, \ldots, 12$ and $M = 1, \ldots, 11$, and these are the initial runs of missing values for these two N. At the other end, Lemma 3.5 forces the final runs. The largest primes below $2N$ for $N = 269, 541$ and 2161 are $q = 523, 1069$ and 4297, which force $M = 261, \ldots, 266$, $M = 534, \ldots, 538$ and $M = 2148, \ldots, 2158$. For $N = 269$ and $N = 541$ these are all the missing values near $N - 3$; for $N = 2161$ the only other missing value with $N - 3 - M \leq 14$ is $M = 2144$.

4.3. *Comparison with the Hardy–Littlewood Prediction*

Let

$$\mathrm{HL}(2N) = \omega \prod_{\substack{p \mid N \\ p > 2}} \frac{p-1}{p-2} \int_2^{2N-2} \frac{dt}{\log t \, \log(2N - t)},$$

with ω as in Proposition 3.6. The script `large_range.py` computes the integral by 64-point Gauss–Legendre quadrature and ω from the primes up to $2 \cdot 10^6$, which gives $\omega \approx 0.6601618372$. Table 3 gives, on each interval I_m, the minimum of $R(2N)$ and the minimum, maximum and mean of $R(2N)/\mathrm{HL}(2N)$. Since $R(2N) \geq 1$ for every $N \in [4, 10^6]$, every even integer $2N$ with $8 \leq 2N \leq 2 \cdot 10^6$ is the sum of two distinct primes. On I_{19} the ratio lies between 0.9657 and 1.0296, with mean 0.9982.

Table 3: Minimum of $R(2N)$ (with the N attaining it) and the minimum, maximum and mean of $R(2N)/\mathrm{HL}(2N)$ on the dyadic intervals I_m.

m	**Range** I_m	**Min** $R(2N)$ (N)	**Min ratio**	**Max ratio**	**Mean ratio**
2	$[4, 8]$	1 (4)	0.2497	0.8705	0.5112
3	$(2^3, 2^4]$	2 (9)	0.3711	0.6800	0.5565
4	$(2^4, 2^5]$	1 (19)	0.2806	1.1702	0.6706
5	$(2^5, 2^6]$	2 (34)	0.4264	1.0380	0.7407
6	$(2^6, 2^7]$	4 (76)	0.5738	1.0863	0.7970
7	$(2^7, 2^8]$	6 (139)	0.5280	1.0909	0.8530
8	$(2^8, 2^9]$	9 (271)	0.6300	1.1322	0.8890
9	$(2^9, 2^{10}]$	16 (556)	0.6960	1.1621	0.9258
10	$(2^{10}, 2^{11}]$	26 (1039)	0.7159	1.1448	0.9463
11	$(2^{11}, 2^{12}]$	45 (2104)	0.7833	1.1361	0.9624
12	$(2^{12}, 2^{13}]$	75 (4276)	0.7754	1.1131	0.9750
13	$(2^{13}, 2^{14}]$	131 (8311)	0.8673	1.1129	0.9820
14	$(2^{14}, 2^{15}]$	222 (16519)	0.8807	1.0906	0.9887
15	$(2^{15}, 2^{16}]$	406 (32789)	0.9093	1.0711	0.9918
16	$(2^{16}, 2^{17}]$	707 (65551)	0.9181	1.0570	0.9946
17	$(2^{17}, 2^{18}]$	1279 (132274)	0.9424	1.0529	0.9962
18	$(2^{18}, 2^{19}]$	2291 (262436)	0.9536	1.0385	0.9975
19	$(2^{19}, 2^{20}]$	4130 (525166)	0.9657	1.0296	0.9982

4.4. Missing Values Forced by Prime Gaps

For each N, let $F(N)$ be the number of $M \in A_N$ that Lemma 3.4 or Lemma 3.5 places in $A_N \setminus E_N$, and let $U(N) = |A_N \setminus E_N| - F(N)$ be the number of the remaining missing values. The script `large_range.py` computes $F(N)$ and $U(N)$ for every $N \in [4, 10^6]$ from the primes around N and the largest prime below $2N$, and confirms $F(N) \leq |A_N \setminus E_N|$ for every such N. Table 4 gives, on each interval I_m, the largest values of $|A_N \setminus E_N|$, $F(N)$ and $U(N)$, the mean of $U(N)$, and the share $\sum F(N)/\sum |A_N \setminus E_N|$ of the missing values that the two lemmas force.

The share of forced values strictly decreases with m, from 1 on I_2 to 0.2698 on I_{19}, while the mean of $U(N)$ increases from 0 to 41.545. On I_6 and on every I_m with $8 \leq m \leq 14$, every N has at least one missing value that the two lemmas do not force, and on every I_m with $15 \leq m \leq 19$ the share of N with $U(N) = 0$ is below 0.00005. No $N \leq 10^6$ has $F(N) \geq \log^2(2N)$ or $U(N) > \log^2(2N)$, and the largest value of $U(N)/\log^2(2N)$ over $1000 \leq N \leq 10^6$ is 0.6263, at $N = 484857$.

Table 4: Largest exceptional set, largest number $F(N)$ of forced missing values, largest number $U(N)$ of the remaining missing values (with the N attaining it), mean of $U(N)$, and share of forced missing values on the dyadic intervals I_m.

m	I_m	**Max** $\lvert A_N \setminus E_N\rvert$	**Max** F	**Max** U (N)	**Mean** U	**Share**
2	$[4, 8]$	2	2	0 (4)	0.000	1.0000
3	$(2^3, 2^4]$	2	2	1 (12)	0.125	0.9091
4	$(2^4, 2^5]$	6	4	4 (30)	0.812	0.7111
5	$(2^5, 2^6]$	9	7	5 (60)	1.375	0.6562
6	$(2^6, 2^7]$	10	9	6 (128)	2.125	0.6344
7	$(2^7, 2^8]$	13	11	9 (166)	3.555	0.5221
8	$(2^8, 2^9]$	16	14	9 (311)	4.648	0.4868
9	$(2^9, 2^{10}]$	24	17	15 (702)	7.020	0.4439
10	$(2^{10}, 2^{11}]$	29	21	20 (1373)	8.673	0.4420
11	$(2^{11}, 2^{12}]$	35	25	26 (3083)	10.956	0.4092
12	$(2^{12}, 2^{13}]$	49	27	35 (7072)	14.507	0.3706
13	$(2^{13}, 2^{14}]$	61	50	42 (14249)	16.576	0.3619
14	$(2^{14}, 2^{15}]$	64	45	51 (24683)	20.686	0.3350
15	$(2^{15}, 2^{16}]$	74	52	57 (63199)	23.915	0.3218
16	$(2^{16}, 2^{17}]$	90	53	69 (107449)	27.847	0.3072
17	$(2^{17}, 2^{18}]$	113	69	90 (250433)	32.279	0.2922
18	$(2^{18}, 2^{19}]$	135	83	119 (484857)	36.733	0.2803
19	$(2^{19}, 2^{20}]$	140	93	115 (854722)	41.545	0.2698

5. Rigorous Consequences and the Open Centered Problem

5.1. What Follows from the Short-Interval Estimate

The results of Dusart imply useful lower bounds for primes above N, and hence for the diagnostic set E_N.

Proposition 5.1 (Primes in $(N, 2N)$)**.** *For every integer $N \geq 3275$, the interval $(N, 2N)$ contains at least $\lfloor \log^2 N \rfloor - 1$ primes.*

Proof. Starting with $x_0 = N$, form consecutive intervals

$$(x_j,\ x_j(1 + 1/(2\log^2 x_j))], \qquad j = 0, 1, 2, \ldots,$$

until the first index $m + 1$ for which $x_{m+1} \geq 2N$. By Proposition 2.1, each of the first m intervals contains a prime lying in $(N, 2N)$. Since $x_j \leq 2N$ and $\log x_j \geq \log N$ throughout the construction, each interval has length at most $N/\log^2 N$. Reaching total length N therefore requires at least $\log^2 N$ such interval lengths. Thus $m \geq \log^2 N - 1$, and the number of primes in $(N, 2N)$ is at least $\lfloor \log^2 N \rfloor - 1$. □

Corollary 5.2 (Baseline half-difference coverage)**.** *For every integer $N \geq 3275$,*

$$|E_N| \geq \lfloor \log^2 N \rfloor - 2.$$

Proof. Let Q be a prime in $(N, 2N)$. If $Q \leq 2N - 3$, then with $P = 3$ the value $M = (Q - 3)/2$ lies in A_N and belongs to E_N. Distinct primes Q give distinct values of M. The only possible prime in $(N, 2N)$ excluded by the admissible range in this construction is $Q = 2N - 1$, which would give $M = N - 2$. Proposition 5.1 therefore gives the stated lower bound. □

Theorem 3.9 gives the linear bound $|E_N| > N/760$ for the same range of N, while Hypothesis 3.14 asks for $|E_N| > (N-3) - \log^2(2N)$. The centered proof target is the stronger prime-correlation condition $R(2N) = |C_N \cap D_N| > 0$.

Proposition 5.3 (A pigeonhole consequence of $H(N) > 0$)**.** *For $N \geq 356$, if $H(N) > 0$, then*

$$C_N \cap E_N \neq \emptyset.$$

Proof. The condition $H(N) > 0$ gives

$$|A_N \setminus E_N| < \log^2(2N).$$

On the other hand, $|C_N| = \pi(N-1) - 1 \geq \pi(N) - 2$. By Proposition 2.4, since $N \geq 356 > 17$,

$$\pi(N) > \frac{N}{\log N} > \frac{N}{\log N + 2},$$

and by Lemma 2.6,

$$|C_N| > \frac{N}{\log N + 2} - 2 > \log^2(2N).$$

Hence C_N cannot be contained in $A_N \setminus E_N$, and so $C_N \cap E_N$ is nonempty. □

Proposition 5.3 gives $C_N \cap E_N \neq \emptyset$ with an explicit range of N under the condition $H(N) > 0$, which the computation of Section 4 verifies for $4 \leq N \leq 10^6$. Corollary 3.12 gives the same conclusion for all sufficiently large N without this condition. In both cases, at least one prime $N - M$ below N has its half-difference M realized by a straddling prime pair; a distinct-prime Goldbach partition of $2N$ requires a prime $N - M$ for which the centered pair $(N - M, N + M)$ itself is such a pair.

5.2. Prime Gaps and the Coverage Statistic

Lemma 3.4 turns every long prime gap into many missing values of E_N. This gives a lower bound for the exceptional set along a subsequence, and it shows what positivity of $H(N)$ would imply for prime gaps.

Theorem 5.4 (Coverage implies a Cramér-type bound)**.** *Let $N_0 \geq 4$, and suppose that $H(N) > 0$ for every integer $N \geq N_0$. Then every pair of consecutive primes $p < p'$ with $p \geq \max\{3, N_0 - 1\}$ satisfies*

$$p' - p < 2\log^2(2p + 2) + 2.$$

Proof. Let $N = p + 1$, so that $N \geq N_0$ and $N \geq 4$. The largest prime below N is p, and the smallest prime above N is p', because $p' \geq p + 2 > N$. Let $g = p' - p$, which is even. By Lemma 3.4, the values $M = 1, \ldots, g/2 - 1$ lie in $A_N \setminus E_N$; they lie in A_N because Bertrand's postulate gives $p' < 2p$, hence $g < p = N - 1$ and $g/2 - 1 \leq N - 3$. Thus $|A_N \setminus E_N| \geq g/2 - 1$. Since $H(N) > 0$ is equivalent to $|A_N \setminus E_N| < \log^2(2N)$, we get $g/2 - 1 < \log^2(2p + 2)$. □

Cramér [15] conjectured that $p' - p = O(\log^2 p)$ for consecutive primes $p < p'$, while the best unconditional bound is $p' - p \leq p^{0.525}$ for all sufficiently large p [16]. Since the computation of Section 4 gives $H(N) > 0$ for every $4 \leq N \leq 10^6$, Theorem 5.4 shows that Hypothesis 3.14 implies $p' - p < 2\log^2(2p + 2) + 2$ for all consecutive primes $p < p'$ with $p \geq 3$, which is a form of Cramér's conjecture with an explicit constant.

In the other direction, long prime gaps give large exceptional sets. Ford, Green, Konyagin, Maynard and Tao [8] proved that there is an absolute constant $c_0 > 0$ such that, for all sufficiently large X,

$$\max_{p_n \leq X}(p_{n+1} - p_n) \geq c_0\, F(X), \qquad F(X) = \frac{\log X \, \log\log X \, \log\log\log\log X}{\log\log\log X},$$

where p_n denotes the n-th prime.

Corollary 5.5 (Large exceptional sets)**.** *There is an absolute constant $c > 0$ such that $|A_N \setminus E_N| \geq c\, F(N)$ for infinitely many integers N. In particular, $|A_N \setminus E_N|$ is unbounded.*

Proof. For large X, choose consecutive primes $p_n < p_{n+1}$ with $p_n \leq X$ and $p_{n+1} - p_n \geq c_0 F(X)$, and let $N = p_n + 1 \leq X + 1$. As in the proof of Theorem 5.4, $|A_N \setminus E_N| \geq (p_{n+1} - p_n)/2 - 1 \geq c_0 F(X)/2 - 1$. By Bertrand's postulate, $p_n > p_{n+1} - p_n \geq c_0 F(X)$, so $N \to \infty$ as $X \to \infty$, and infinitely many N occur. Since F is increasing for large arguments, $N \leq X + 1$ gives $F(N) \leq F(X + 1)$, and $F(X + 1) \sim F(X)$. Hence $|A_N \setminus E_N| \geq c\, F(N)$ for every fixed $c < c_0/2$ once X is large. □

Since $F(N) = o(\log^2 N)$, Corollary 5.5 does not contradict Hypothesis 3.14.

5.3. The Centered Statistic

For the centered sets, the analogue of $H(N)$ is negative for every $N \geq 34$. Let

$$H_R(N) = \log^2(2N) - ((N-3) - R(2N)),$$

obtained from $H(N)$ by replacing E_N with $C_N \cap D_N$.

Proposition 5.6 (Negativity of the centered statistic)**.** *For every integer $N \geq 34$, $H_R(N) < 0$.*

Proof. Write $L = \log N$. Since $R(2N) \leq |C_N| = \pi(N-1) - 1 < \pi(N)$, Proposition 2.3 gives $H_R(N) < -f(N)$, where

$$f(x) = x - 3 - \frac{x}{\log x}\left(1 + \frac{1.2762}{\log x}\right) - \log^2(2x).$$

For $x \geq 8$, write $\ell = \log x$. Then

$$f'(x) = 1 - \frac{\ell - 1}{\ell^2} - 1.2762\,\frac{\ell - 2}{\ell^3} - \frac{2\log(2x)}{x} \geq 1 - \frac{1}{4} - \frac{1.2762}{27} - \frac{2\log 16}{8} > 0,$$

because $(\ell-1)/\ell^2 \leq 1/4$ and $(\ell-2)/\ell^3 \leq 1/27$ for $\ell > 0$, and $2\log(2x)/x$ is decreasing for $x \geq 8$. Hence f is increasing on $[8, \infty)$. At $x = 34$,

$$\frac{34}{\log 34}\left(1 + \frac{1.2762}{\log 34}\right) < 13.132 \quad \text{and} \quad \log^2 68 < 17.805,$$

so $f(34) > 34 - 3 - 13.132 - 17.805 > 0.06$. Therefore $f(N) > 0$ and $H_R(N) < 0$ for every $N \geq 34$. □

5.4. The Gap to a Complete Proof

The missing ingredient is a uniform lower bound for the centered binary prime correlation

$$R(2N) = |C_N \cap D_N| = \#\{p < N \mid p \text{ and } 2N - p \text{ are prime}\}.$$

By Proposition 5.7, a complete proof would follow from

$$R(2N) > 0 \qquad (N > 2 \cdot 10^{18}).$$

Equivalently, the set of primes below N and its reflection across N must intersect. By Corollary 3.12 and Lemma 3.3, for all large N the set $C_N \cap E_N$ has $\pi(N-1) - 1 - O_D(N/(\log N)^D)$ elements, and $C_N \cap D_N = C_N \cap E_N \cap D_N$ is the subset of those M for which $N + M$ is prime, that is, for which the centered pair $(N - M, N + M)$ realizes the half-difference M. The conjecture for $2N$ asks for one such M. By Proposition 5.6, $|C_N \cap D_N| < (N-3) - \log^2(2N)$ for every $N \geq 34$, whereas $|E_N| = N - o(N)$ by Proposition 3.11.

A strong sufficient form would be an explicit Hardy–Littlewood-type lower bound

$$R(2N) \geq c\,\frac{N}{\log^2 N} \prod_{\substack{p \mid N \\ p > 2}} \frac{p-1}{p-2}$$

with a positive constant c and an error term small enough to keep the right-hand side positive for every $N > 2 \cdot 10^{18}$. A weaker but sufficient target is simply a proven inequality $R(2N) \geq 1$ for all $N > 2 \cdot 10^{18}$.

5.5. The Verified Range

Proposition 5.7 (Verified range)**.** *For every integer N with $4 \leq N \leq 2 \cdot 10^{18}$, $R(2N) \geq 1$.*

Proof. Oliveira e Silva, Herzog and Pardi verified the even Goldbach conjecture for all even integers up to $4 \cdot 10^{18}$ [2], and the least prime of a Goldbach partition of every such integer is at most 9781 [17]. Hence $2N = p + q$ with primes $p \leq 9781$ and $q = 2N - p$. If $N > 9781$, then $p < N < q$, and $p \geq 3$ because q would otherwise be an even number larger than 2; hence $M = N - p \in C_N \cap D_N$, and $R(2N) \geq 1$. If $4 \leq N \leq 9781$, the computation of Section 4 gives $R(2N) \geq 1$. □

Consequently, the distinct-prime Goldbach conjecture is equivalent to $R(2N) \geq 1$ for every $N > 2 \cdot 10^{18}$.

5.6. Small Cases

The first cases illustrate the centered sets explicitly:

- $N = 4$ ($2N = 8$): $C_4 = \{1\}$ and $D_4 = \{1\}$, giving $8 = 3 + 5$.
- $N = 5$ ($2N = 10$): $C_5 = \{2\}$ and $D_5 = \{2\}$, giving $10 = 3 + 7$.
- $N = 6$ ($2N = 12$): $C_6 = \{3, 1\}$ and $D_6 = \{1\}$, giving $12 = 5 + 7$.
- $N = 7$ ($2N = 14$): $C_7 = \{4, 2\}$ and $D_7 = \{4\}$, giving $14 = 3 + 11$.
- $N = 8$ ($2N = 16$): $C_8 = \{5, 3, 1\}$ and $D_8 = \{3, 5\}$, giving $16 = 3 + 13$.
- $N = 9$ ($2N = 18$): $C_9 = \{6, 4, 2\}$ and $D_9 = \{2, 4\}$, giving $18 = 5 + 13$.
- $N = 10$ ($2N = 20$): $C_{10} = \{7, 5, 3\}$ and $D_{10} = \{1, 3, 7\}$, giving $20 = 3 + 17$.
- $N = 11$ ($2N = 22$): $C_{11} = \{8, 6, 4\}$ and $D_{11} = \{2, 6, 8\}$, giving $22 = 3 + 19$.
- $N = 12$ ($2N = 24$): $C_{12} = \{9, 7, 5, 1\}$ and $D_{12} = \{1, 5, 7\}$, giving $24 = 5 + 19$.

For $13 \le N \le 2^{14}$, the computation of Section 4 verifies at least one centered intersection value $M \in C_N \cap D_N$.

Remark 5.8 (Computational verification)**.** Our implementation verified the existence of centered Goldbach partitions for all even integers $2N$ with $8 \le 2N \le 2 \cdot 10^6$. The same computations produced the values reported in Tables 1, 2 and 3.

6. Conclusion

We have embedded the distinct-prime Goldbach condition $C_N \cap D_N \neq \emptyset$, which asks for a prime pair centered at N, in the condition defining the half-difference set E_N, which asks only for a prime pair straddling N. The equivalence between a distinct-prime Goldbach partition of $2N$, a nested-square semiprime configuration, and the condition $C_N \cap D_N \neq \emptyset$ is the classical symmetric form of the problem. The straddling condition holds for almost all M, and as a consequence $2N$ has Goldbach partitions up to a difference of two primes; the centering step remains open.

Summary of Main Results

Proposition 5.7 gives $R(2N) \ge 1$ for every $4 \le N \le 2 \cdot 10^{18}$, and Theorem 3.13 proves unconditionally that $R(2N) \ge 1$ for more than $X/9600$ integers $N \le X$, for every $X \ge 10^6$. Theorem 3.9 proves unconditionally that $|E_N| > N/760$ for every $N \ge 3275$, Proposition 3.11, deduced from a theorem of Mikawa, gives $|A_N \setminus E_N| \ll_D N/(\log N)^D$, and Corollary 3.12 shows that, for all large N and all but $O_D(N/(\log N)^D)$ primes $P < N$, the number $2N - P$ becomes prime after adding the difference $P' - P$ of two primes below N. Theorem 3.1 establishes the geometric equivalence, and Proposition 3.2 gives the exact centered set-intersection criterion. Proposition 5.1 and Corollary 5.2 record what follows rigorously from Dusart's short-interval theorem: the interval $(N, 2N)$ contains many primes, and E_N has a baseline supply of admissible half-differences. Lemma 2.6, proved by differentiation, and Proposition 2.4 give Proposition 5.3 for every $N \ge 356$, Lemma 3.3 shows $C_N \cap D_N \subseteq E_N$, Lemmas 3.4 and 3.5 show that the prime gaps around N and below $2N$ force missing values of E_N. Theorem 5.4 shows that $H(N) > 0$ for all large N would imply a Cramér-type bound for prime gaps, Corollary 5.5 shows that $|A_N \setminus E_N|$ is unbounded, and Proposition 5.6 shows $H_R(N) < 0$ for every $N \ge 34$. The computation of Section 4 verifies $R(2N) \ge 1$ and $H(N) > 0$ for $4 \le N \le 10^6$, gives $|E_N| \ge 0.9906\, N$ for $3275 \le N \le 10^6$, and shows $(C_N \cap E_N) \setminus D_N \neq \emptyset$ for every $N \le 2^{14}$ except $N \in \{4, 5, 7, 8\}$.

Status of the Conjecture

The distinct-prime Goldbach conjecture is not proved in this paper. The paper proves that $C_N \cap E_N$ contains almost all of C_N for all large N (Corollary 3.12), and that it is nonempty for every $N \ge 356$ with $H(N) > 0$ (Proposition 5.3). The step needed for a proof is the centering step: one must show that for some such M, the centered upper endpoint $N + M$ is prime, i.e.,

$$C_N \cap D_N \neq \emptyset.$$

Equivalently, by Proposition 5.7, one must prove the positivity of $R(2N) = |C_N \cap D_N|$ for every $N > 2 \cdot 10^{18}$.

The Centered Gap

The proof gap is the centered correlation estimate $R(2N) > 0$. The statistic

$$H(N) = \log^2(2N) - ((N-3) - |E_N|)$$

measures half-difference coverage and is positive for every $N \le 10^6$, while the centered statistic $H_R(N)$ is negative for every $N \ge 34$ (Proposition 5.6). The prime-gap obstructions of Lemmas 3.4 and 3.5 force 26.98% of the missing values of E_N on $(2^{19}, 2^{20}]$ (Table 4). A proof of the distinct-prime Goldbach conjecture along these lines requires $|C_N \cap D_N| \ge 1$ for every $N > 2 \cdot 10^{18}$.

Relation to the Classical Goldbach Conjecture

The framework addresses the variant requiring *distinct* primes, thus excluding $4 = 2 + 2$ and $6 = 3 + 3$. Extending the geometric construction to allow $P = Q$ would require handling the case $M = 0$, which lies outside the present admissible interval.

Open Questions

Several natural questions remain. By Proposition 3.11, $|E_N| = N - o(N)$, but with no explicit range of N. For which explicit $c < 1$ and N_0 can one prove $|E_N| \ge cN$ for every $N \ge N_0$? Theorem 3.9 gives $c = 1/760$ and $N_0 = 3275$, and the computation gives $|E_N| \ge 0.9906\,N$ for $3275 \le N \le 10^6$. Which arithmetic structure accounts for the missing values of E_N that Lemmas 3.4 and 3.5 do not force? Can one prove an explicit lower bound for $R(2N)$, either directly or through a centering theorem for prime-pair differences? What is the exact asymptotic behaviour of $H(N)$, and does $|A_N \setminus E_N|$ exceed $\log^2(2N)$ for infinitely many N? By Theorem 5.4, the latter holds if $p' - p \ge 2\log^2(2p+2) + 2$ for infinitely many consecutive primes $p < p'$. Can the geometric framework accommodate $M = 0$, thereby addressing the full classical Goldbach conjecture? Finally, can similar geometric reformulations illuminate other additive problems, such as the ternary Goldbach conjecture or Waring's problem?

Statements and Declarations

Declaration of interests

The author declares no competing interests.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of generative AI in scientific writing

The author used Claude Opus 5.5 to write the experiment scripts `experiment.py` and `large_range.py`, and to improve the readability of this manuscript; the author ran all the experiments. After using this tool, the author reviewed and edited the code and the content as needed and takes full responsibility for the content of the publication.

Data availability

Code and data: https://github.com/frankvegadelgado/goldbach; the computations of Section 4 are reproduced by running the scripts `experiment.py` and `large_range.py` of the repository, which write their output to log files.

References

[1] C. Goldbach, Lettre XLIII, in: P. H. Fuss (Ed.), Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle, Vol. 1, Imperial Academy of Sciences, St. Petersburg, 1843, pp. 125–129, (letter to Leonhard Euler) (in German).

[2] T. Oliveira e Silva, S. Herzog, S. Pardi, Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$, Mathematics of Computation 83 (288) (2014) 2033–2060. doi:10.1090/S0025-5718-2013-02787-1.

[3] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, sequence A082467: Least $k > 0$ such that $n - k$ and $n + k$ are both primes, https://oeis.org/A082467, accessed: 2026-11-03 (2003).

[4] H. Mikawa, On prime twins, Tsukuba Journal of Mathematics 15 (1) (1991) 19–29. doi:10.21099/tkbjm/1496161564.

[5] H. Siebert, Montgomery's weighted sieve for dimension two, Monatshefte für Mathematik 82 (4) (1976) 327–336. doi:10.1007/BF01540603.

[6] P. Dusart, Autour de la fonction qui compte le nombre de nombres premiers, https://www.unilim.fr/pages_perso/pierre.dusart/Documents/T1998_01.pdf, accessed: 2026-11-03 (1998).

[7] J. B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois Journal of Mathematics 6 (1) (1962) 64–94. doi:10.1215/ijm/1255631807.

[8] K. Ford, B. Green, S. Konyagin, J. Maynard, T. Tao, Long gaps between primes, Journal of the American Mathematical Society 31 (1) (2018) 65–105. doi:10.1090/jams/876.

[9] D. Wolke, Über das Primzahl-Zwillingsproblem, Mathematische Annalen 283 (1989) 529–537. doi:10.1007/BF01442852.

[10] A. F. Lavrik, The number of k-twin primes lying on an interval of a given length, Doklady Akademii Nauk SSSR 136 (2) (1961) 281–283. URL https://mathnet.ru/eng/dan24497

[11] A. Perelli, J. Pintz, On the exceptional set for the $2k$-twin primes problem, Compositio Mathematica 82 (3) (1992) 355–372. URL http://www.numdam.org/item?id=CM_1992__82_3_355_0

[12] K. Matomäki, M. Radziwiłł, T. Tao, Correlations of the von Mangoldt and higher divisor functions I. Long shift ranges, Proceedings of the London Mathematical Society 118 (2019) 284–350, preprint at https://arxiv.org/abs/1707.01315.

[13] H. L. Montgomery, R. C. Vaughan, The exceptional set in Goldbach's problem, Acta Arithmetica 27 (1975) 353–370. doi:10.4064/aa-27-1-353-370.

[14] F. Vega, Experimental Results on Goldbach's Conjecture, https://github.com/frankvegadelgado/goldbach, accessed: 2026-11-03 (2026).

[15] H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, Acta Arithmetica 2 (1) (1936) 23–46. doi:10.4064/aa-2-1-23-46.

[16] R. C. Baker, G. Harman, J. Pintz, The difference between consecutive primes, II, Proceedings of the London Mathematical Society 83 (3) (2001) 532–562. doi:10.1112/plms/83.3.532.

[17] T. Oliveira e Silva, Goldbach conjecture verification, https://sweet.ua.pt/tos/goldbach.html, accessed: 2026-11-03.