Article

Not peer-reviewed version

A Centered Geometric Framework for the Distinct-Prime Goldbach Problem

Frank Vega *

Posted Date: 2 June 2026

doi: 10.20944/preprints202511.0701.v7

Keywords: Goldbach conjecture; geometric construction; semiprimes; prime distribution; Dusart's theorem; computational evidence

Article

A Centered Geometric Framework for the Distinct-Prime Goldbach Problem

Frank Vega

Information Physics Institute, 840 W 67th St, Hialeah, FL 33012, USA; vega.frank@gmail.com

Abstract

We develop a centered geometric and combinatorial framework for the distinct-prime Goldbach problem—the assertion that every even integer $2N \geq 8$ is the sum of two *distinct* primes. The rigorous content is an exact geometric reformulation: for each $N \geq 4$, the problem is equivalent to finding an integer $M \in [1, N-3]$ such that $N - M$ and $N + M$ are both prime, or equivalently such that the L-shaped region between nested squares of side lengths N and M has area $(N-M)(N+M)$ with both factors prime. We define the centered sets $C_N = \{M : N - M \text{ is prime}\}$ and $D_N = \{M : N + M \text{ is prime}\}$ inside $\{1, \ldots, N-3\}$; then $C_N \cap D_N \neq \varnothing$ is exactly the distinct-prime Goldbach assertion for $2N$. We also study the diagnostic set E_N of admissible half-differences arising from straddling prime pairs; its elements correspond to Goldbach partitions of nearby even numbers. A decomposition by upper primes expresses $|E_N|$ in terms of the half-differences generated by each prime $Q \in (N, 2N)$ minus explicit collision sets. Under a fixed upper-prime collision hypothesis, explicit estimates of Rosser–Schoenfeld and Dusart imply positivity of the statistic $H(N) = \log^2(2N) - ((N-3) - |E_N|)$ for all sufficiently large N, and finite computation verifies the remaining tested range. Positivity of $H(N)$ gives a pigeonhole result $C_N \cap E_N \neq \varnothing$. The distinct-prime Goldbach conjecture is not proved here; the remaining gap is a centering theorem, equivalently a uniform lower bound for $R(2N) = |C_N \cap D_N|$.

Keywords: Goldbach conjecture; geometric construction; semiprimes; prime distribution; Dusart's theorem; computational evidence

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 develops a **centered geometric and combinatorial framework** for studying 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 nested squares whose L-shaped difference region has area $(N-M)(N+M)$ with both factors prime. The framework combines three elements. First, a *geometric equivalence* (Section 3) establishes that the variant Goldbach conjecture is equivalent to the existence of $M \in [1, N-3]$ such that $N - M$ and $N + M$ are both prime. Second, an exact *centered intersection criterion* defines $C_N = \{M : N - M \text{ is prime}\}$ and $D_N = \{M : N + M \text{ is prime}\}$ inside the admissible interval; the condition $C_N \cap D_N \neq \varnothing$ is precisely the desired partition. Third, a half-difference set E_N is studied as a computational diagnostic for partitions of nearby even numbers by straddling prime pairs. The centered set D_N records primality of the specific upper endpoint $N + M$. The diagnostic set E_N records half-differences of prime pairs that straddle N, independently of whether a pair is centered at N. This

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2026 doi:10.20944/preprints202511.0701.v7

separation keeps the exact Goldbach criterion and the exploratory prime-pair statistic conceptually distinct. For the diagnostic set E_N, we consider the statistic

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

which measures how close E_N is to filling the admissible interval. The computational evidence in Section 4 is presented as heuristic support for the framework, together with direct verification of partitions in the tested range. The main results are conditional in two places. First, the derivation of $H(N) > 0$ for all large N requires a fixed upper-prime collision hypothesis controlling repeated half-differences. Second, even $H(N) > 0$ yields only $C_N \cap E_N \neq \varnothing$ by pigeonhole; a complete proof of the distinct-prime Goldbach conjecture still requires a centering theorem that upgrades this to $C_N \cap D_N \neq \varnothing$. The remainder of this paper is organised as follows. Section 2 collects the analytic number theory prerequisites used in the quantitative discussion. Section 3 develops the geometric framework and proves the centered set-intersection criterion. Section 4 presents computational evidence. Section 5 records the rigorous consequences, the remaining conditional statement, and the gap that must be closed for a complete proof. Section 6 discusses significance and future directions.

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 [3] and Rosser–Schoenfeld [4].

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 1 ([3, 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 1 is derived from the following bound on consecutive primes.

Proposition 2 ([3, 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 1 in [3] proceeds by using Proposition 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 also require explicit bounds on $\pi(x)$, the number of primes not exceeding x.

Proposition 3 ([3, Théorème 1.10, p. 36])**.** *The following inequalities hold:*

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

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

(3) $\pi(x) \geq \dfrac{x}{\ln x - 1}$ *for all $x \geq 5393$.*

(4) $\dfrac{x}{\ln x}\left(1 + \dfrac{1}{\ln x} + \dfrac{1.8}{\ln^2 x}\right) \leq \pi(x)$ *for all $x \geq 32299$.*

Throughout this paper, log denotes the natural logarithm (i.e., log ≡ ln).

doi:10.20944/preprints202511.0701.v7

2.3. Primes in Dyadic Intervals

We shall also use the following classical explicit form of Bertrand's postulate due to Rosser and Schoenfeld [4].

Proposition 4 ([4, Corollaries 3.8–3.9]). *For every real number $x \geq 41/2$,*

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

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 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. □

doi:10.20944/preprints202511.0701.v7

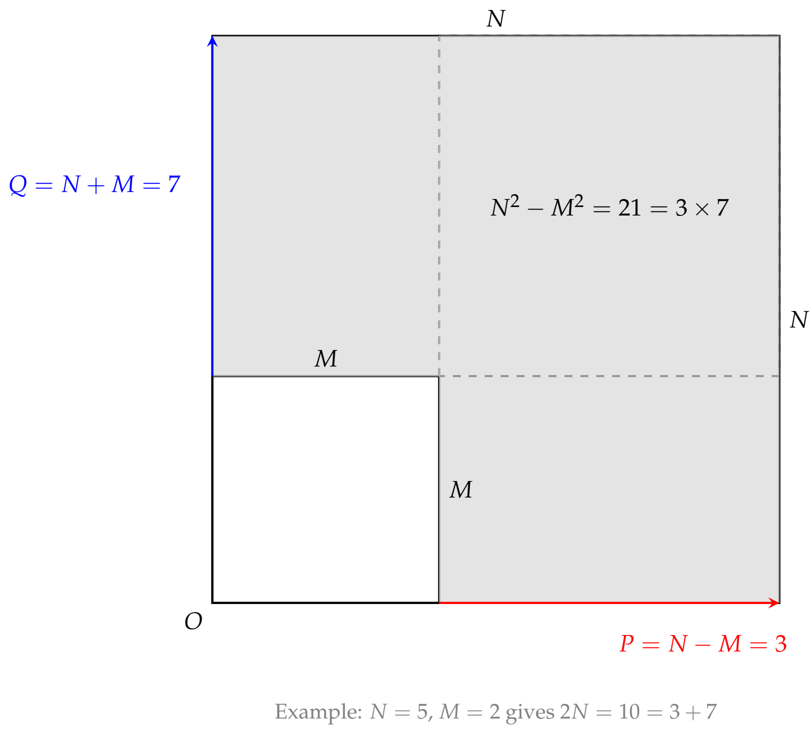

Figure 1. The geometric construction for $N = 5$, $M = 2$: The L-shaped region has area $N^2 - M^2 = 25 - 4 = 21 = 3 \times 7$, a semiprime. The factors $P = 3$ and $Q = 7$ are both prime and sum to $2N = 10$, providing the Goldbach partition $10 = 3 + 7$.

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 5 (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 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

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

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2026 doi:10.20944/preprints202511.0701.v7

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 $C_N \cap D_N$.

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

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

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. Counting E_N by Upper Primes

The preceding interpretation leads to a concrete counting decomposition. Let

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

and let

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

If $2N-3$ is prime, then $2N = 3 + (2N-3)$ is already a centered Goldbach partition with $M = N-3$. Thus the endpoint $2N-3$ is included in U_N and is also a trivial-success endpoint for the centered problem. 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

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

Proposition 4 therefore gives, for $N \ge 41/2$,

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

For each $Q_j \in U_N$, define

$$\mathcal{M}(Q_j) = \left\{ \frac{Q_j - P}{2} \;\middle|\; P \in L_N \right\}.$$

Because $3 \le P < N < Q_j \le 2N-3$, every element of $\mathcal{M}(Q_j)$ lies in A_N, and the map $P \mapsto (Q_j - P)/2$ is injective. Hence

$$|\mathcal{M}(Q_j)| = a_N$$

for every j.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2026 doi:10.20944/preprints202511.0701.v7

Let

$$E_N^- = \bigcup_{j=1}^{r} \mathcal{M}(Q_j) \subseteq E_N.$$

When $\mathcal{M}(Q_j)$ is added after $\mathcal{M}(Q_1), \ldots, \mathcal{M}(Q_{j-1})$, the excluded values are exactly those produced by lower primes $P \in L_N$ for which

$$\frac{Q_j - P}{2} = \frac{Q_i - P'}{2}$$

for some $i < j$ and some $P' \in L_N$. Equivalently,

$$P' - P = Q_i - Q_j > 0.$$

Define the collision set

$$\mathcal{X}_j = \{P \in L_N \mid P + Q_i - Q_j \in L_N \text{ for some } i < j\}.$$

Then

$$|E_N^-| = a_N + \sum_{j=2}^{r} (a_N - |\mathcal{X}_j|) = ra_N - \sum_{j=2}^{r} |\mathcal{X}_j|.$$

Consequently, since $E_N^- \subseteq E_N$, the inequality $H(N) > 0$ follows from the collision bound

$$\sum_{j=2}^{r} |\mathcal{X}_j| < ra_N - (N-3) + \log^2(2N).$$

This is the precise form of the exclusion problem: each new upper prime contributes a_N possible half-differences, minus those already created by earlier upper primes through equal lower-prime gaps.

For a fixed upper prime Q_j, define the set of upper gaps already present above it by

$$\Gamma_j = \{Q_i - Q_j \mid 1 \le i < j\}.$$

Then

$$|\Gamma_j| \le j - 1.$$

Thus the number of *distinct* upper-prime gap sizes that can collide with values generated by Q_j is at most $j - 1$. The actual number of repeated M-values contributed by Q_j satisfies the sharper bound

$$|\mathcal{X}_j| \le \sum_{d \in \Gamma_j} \rho_L(d),$$

where $\rho_L(d)$ counts lower-prime pairs with difference d. The following hypothesis asserts that, in the sequential construction, the total number of repeated M-values contributed by Q_j is bounded by the number of upper primes already considered:

$$|\mathcal{X}_j| \le j - 1 \qquad (2 \le j \le r).$$

This assertion is stronger than the bound $|\Gamma_j| \le j - 1$: the latter counts only distinct upper-gap sizes, whereas a single gap size d may have several lower-prime realizations $P' - P = d$.

Hypothesis 1 (Fixed upper-prime collision hypothesis)**.** *For every $N \ge 3275$ and every $2 \le j \le r$,*

$$|\mathcal{X}_j| \le j - 1 \qquad (2 \le j \le r).$$

 doi:10.20944/preprints202511.0701.v7

Under Hypothesis 1,

$$\sum_{j=2}^{r} |\mathcal{X}_j| \le \sum_{j=2}^{r} (j-1) = \frac{r(r-1)}{2}.$$

Therefore $H(N) > 0$ follows whenever

$$\frac{r(r-1)}{2} < ra_N - (N-3) + \log^2(2N).$$

Proposition 6 (Rosser–Schoenfeld support for $H(N) > 0$)**.** *Assume Hypothesis 1. Then $H(N) > 0$ for every $N \ge 3275$.*

Proof. Hypothesis 1 gives

$$|E_N| \ge ra_N - \frac{r(r-1)}{2} = r\left(a_N - \frac{r-1}{2}\right).$$

By Proposition 4,

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

By Proposition 3(1), for $N \ge 3275$,

$$a_N = \pi(N-1) - 1 \ge \frac{N-1}{\log(N-1)}\left(1 + \frac{1}{\log(N-1)}\right) - 1.$$

Combining these explicit inequalities yields

$$|E_N| > \left(\frac{3N}{5\log N} - 1\right)\left[\frac{N-1}{\log(N-1)}\left(1 + \frac{1}{\log(N-1)}\right) - 1 - \frac{1}{2}\left(\frac{7N}{5\log N} - 1\right)\right].$$

Let $\Phi(x)$ denote the right-hand side of the preceding displayed inequality, with N replaced by the real variable x, minus

$$(x-3) - \log^2(2x).$$

An elementary differentiation check gives $\Phi'(x) > 0$ for $x \ge 3275$, and

$$\Phi(3275) > 38088.$$

Hence $\Phi(N) > 0$ for every integer $N \ge 3275$, so

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

which is equivalent to $H(N) > 0$. □

Corollary 1 (Conditional positivity of $H(N)$)**.** *If Hypothesis 1 holds, and the finite range $4 \le N \le 3274$ is verified directly, then $H(N) > 0$ for every $N \ge 4$.*

Proof. The range $N \ge 3275$ is Proposition 6. The remaining range is finite. □

This exclusion problem can be restated in terms of prime-difference sets. Let

$$\Delta_L = \{P' - P > 0 \mid P, P' \in L_N\}, \qquad \Delta_U = \{Q_i - Q_j > 0 \mid Q_i, Q_j \in U_N,\ i < j\}.$$

Then collisions can occur only for gap sizes in $\Delta_L \cap \Delta_U$. The number of distinct lower-prime gaps is bounded by

$$|\Delta_L| \le \binom{a_N}{2} < a_N^2,$$

and the number of distinct upper-prime gaps is bounded by

$$|\Delta_U| \le \binom{r}{2} < r^2 \le (\pi(2N) - \pi(N))^2 < \left(\frac{7N}{5\log N}\right)^2 \qquad (N \ge 41/2),$$

Thus the number of distinct gap sizes that can produce collisions is at most

$$|\Delta_L \cap \Delta_U| \le \min\left\{a_N^2,\ (\pi(2N) - \pi(N))^2\right\}.$$

Under the Rosser–Schoenfeld upper bound, the second term is $< (7N/(5\log N))^2$.

For multiplicities, define

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

The total number of collision witnesses is bounded by the additive-energy-type quantity

$$\mathcal{C}_N := \sum_{d>0} \rho_L(d)\rho_U(d).$$

Since every excluded value in the sequential construction has at least one such witness,

$$\sum_{j=2}^{r} |\mathcal{X}_j| \le \mathcal{C}_N.$$

Consequently, a sufficient condition for $H(N) > 0$ is

$$\mathcal{C}_N < ra_N - (N-3) + \log^2(2N).$$

The crude universal bound

$$\mathcal{C}_N \le \left(\sum_{d>0} \rho_L(d)\right)\left(\sum_{d>0} \rho_U(d)\right) \le \binom{a_N}{2}\binom{r}{2}$$

is generally too coarse to prove positivity of $H(N)$, but it isolates the needed refinement: one must exploit the distribution of prime gaps, not merely the number of available lower and upper primes.

3.8. 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

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

Intuitively, $H(N) > 0$ means that the half-difference diagnostic set E_N is “almost full”, with fewer than $\log^2(2N)$ admissible values missing. This is a useful diagnostic for the distribution of straddling prime-pair differences.

Hypothesis 2 (Half-Difference Density Hypothesis)**.** *For every integer $N \ge 3275$, we have $H(N) > 0$.*

Hypothesis 3 (Centered Prime-Correlation Hypothesis)**.** *For every integer $N \ge 3275$, the centered prime-pair count*

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

is positive.

Remark 1 (Status of the hypotheses)**.** *Hypothesis 1 is a sufficient condition for Hypothesis 2 by Proposition 6. Hypothesis 2, in turn, is sufficient for the pigeonhole conclusion $C_N \cap E_N \neq \emptyset$ by Proposition 8. Neither statement is by itself sufficient for the centered Goldbach partition; that requires Hypothesis 3.*

Theorem 2 (Conditional Main Result)**.** *If Hypothesis 3 holds, and the remaining cases $4 \leq N \leq 3274$ are verified directly, then every even integer $2N \geq 8$ is the sum of two distinct primes.*

4. Computational Evidence

We computed $|E_N|$ and $H(N)$ for all $N \in [4, 2^{14}]$ using Python 3.12 with the Gmpy2 library [5]. The results are summarised in Table 1. Three features of these data merit attention. First, $H(N) > 0$ for *every* $N \in [4, 2^{14}]$, providing empirical support for Hypothesis 2. Second, the minimum value of $H(N)$ in each successive dyadic interval $[2^m, 2^{m+1}]$ strictly increases with m, suggesting that the positivity margin for the half-difference diagnostic set widens as N grows. Third, the N-values at which the minima are attained tend to be primes or near-primes, consistent with the expectation that $|E_N|$ is smallest when primes near N are sparse.

Table 1. Minimum $H(N)$ values in dyadic intervals $[2^m, 2^{m+1}]$. Note that $H(N) > 0$ for all tested values, and the minima strictly increase with m.

Interval (m)	Range $[2^m, 2^{m+1}]$	N achieving min	Min $H(N)$
2	$[4, 8]$	5	4.301898
3	$[8, 16]$	9	7.354249
4	$[16, 32]$	19	10.232033
5	$[32, 64]$	61	14.078618
6	$[64, 128]$	73	17.836335
7	$[128, 256]$	151	20.608977
8	$[256, 512]$	269	23.537165
9	$[512, 1024]$	541	28.812111
10	$[1024, 2048]$	1327	33.154668
11	$[2048, 4096]$	2161	35.081569
12	$[4096, 8192]$	7069	42.329014
13	$[8192, 16384]$	14138	44.057758

5. Rigorous Consequences and Conditional Statement

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 7 (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 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 2 (Baseline half-difference coverage)**.** *For every integer $N \geq 3275$,*

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

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2026 doi:10.20944/preprints202511.0701.v7

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 7 therefore gives the stated lower bound. □

Corollary 2 is much weaker than Hypothesis 2, which asks that E_N be almost full. The centered proof target is the stronger prime-correlation condition $R(2N) = |C_N \cap D_N| > 0$.

Proposition 8 (A pigeonhole consequence of $H(N) > 0$)**.** *For $N \geq 3275$, 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, $|C_N| = \pi(N-1) - 1$. By Proposition 3*(1)*,

$$|C_N| \geq \frac{N-1}{\log(N-1)}\left(1 + \frac{1}{\log(N-1)}\right) - 1 > \log^2(2N)$$

for every $N \geq 3275$; the last comparison is elementary and is already true at $N = 3275$, after which the left side grows on the order of $N/\log N$ while the right side grows on the order of $\log^2 N$. Hence C_N cannot be contained in $A_N \setminus E_N$, and so $C_N \cap E_N$ is nonempty. □

Proposition 8 is the natural pigeonhole consequence of the statistic $H(N)$. It says that at least one lower-centered prime $N - M$ has the same half-difference as a straddling prime pair. To obtain a Goldbach partition of $2N$, this must be strengthened to $C_N \cap D_N \neq \emptyset$.

5.2. The Gap to a Complete Proof

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

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

A complete proof within this framework would follow from

$$R(2N) > 0 \qquad (N \geq 3275).$$

Equivalently, one must prove that the prime set below N and its reflection across N intersect at an admissible odd prime. Separate density bounds for primes below and above N, and high coverage of the diagnostic set E_N, are not enough unless they are supplemented by a centering theorem that controls this reflected correlation.

In the pigeonhole route suggested by Proposition 8, the needed centering theorem can be stated as

$$H(N) > 0 \quad \Longrightarrow \quad C_N \cap E_N \cap D_N \neq \emptyset$$

for all sufficiently large N. Equivalently, one must show that among the lower-centered primes $N - M$ whose half-differences occur in nearby straddling partitions, at least one has the centered partner $N + M$ prime.

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 $N \geq 3275$. A weaker but sufficient target is simply a proven inequality $R(2N) \geq 1$ for all $N \geq 3275$.

5.3. Proof of Theorem 2

Proof. Assume Hypothesis 3. For every $N \geq 3275$, Proposition 5 immediately gives a representation of $2N$ as the sum of two distinct primes.

It remains only to check $4 \leq N \leq 3274$. The first few 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 $13 \leq N \leq 3274$, direct computation verifies at least one centered intersection value $M \in C_N \cap D_N$. Thus the theorem follows conditionally from Hypothesis 3 and finite verification. □

Remark 2 (Scope of the conditional theorem)**.** *Hypothesis 3 is a centered form of the distinct-prime Goldbach assertion for large N. Therefore Theorem 2 is a precise conditional statement, not an unconditional proof. The diagnostic density hypothesis for E_N remains heuristic unless it is supplemented by a centering theorem for $R(2N)$.*

Remark 3 (Computational verification)**.** *Our implementation verified the existence of centered Goldbach partitions for all even integers up to $2 \times 2^{14} = 32{,}768$. The same computations also produced the values of $H(N)$ reported in Table 1.*

6. Conclusion

We have developed a centered geometric and combinatorial framework for the distinct-prime Goldbach conjecture. The rigorous part of the framework is the exact equivalence between a distinct-prime Goldbach partition of $2N$, a nested-square semiprime configuration, and the centered intersection condition $C_N \cap D_N \neq \emptyset$. The diagnostic half-difference set E_N provides additional numerical information about nearby prime-pair differences.

Summary of Main Results

Theorem 1 establishes the geometric equivalence, and Proposition 5 gives the exact centered set-intersection criterion. Proposition 7 and Corollary 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.

Status of the Conjecture

The distinct-prime Goldbach conjecture is not proved in this paper. The paper proves an exact centered reformulation and several conditional implications. In particular, Hypothesis 1 would imply $H(N) > 0$ for all large N, and $H(N) > 0$ implies by the pigeonhole principle that $C_N \cap E_N$ is nonempty. The final step needed for a proof is stronger: one must show that for some such admissible M, the centered upper endpoint $N + M$ is prime, i.e.,

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

Equivalently, one must prove the positivity of $R(2N) = |C_N \cap D_N|$ for all sufficiently large N.

Key Insights

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 appears positive in the tested range. The upper-prime decomposition in Section 3.7 reduces a conditional proof of $H(N) > 0$ to bounding the collision sets $\mathcal{X}_j$. A complete Goldbach proof using this framework must then prove a centering theorem strong enough to guarantee $|C_N \cap D_N| \geq 1$ for every sufficiently large N.

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. Can Hypothesis 1, or a weaker average version of it, be proved sharply enough to imply $H(N) > 0$? 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 $H(N)$? 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?

Acknowledgments: The author is sincerely grateful to Iris, Marilin, Sonia, Yoselin, Arelis, Anissa, Liuva, Yudit, Gretel, Gema, and Blaquier, as well as Israel, Arderi, Juan Carlos, Yamil, Alejandro, Aroldo, Yary, Reinaldo, Alex, Emmanuel, and Michael for their constant support. Whether through encouragement, stimulating conversations, practical assistance, or simply being present during challenging moments, their contributions have played an important role in bringing this work to completion.

References

1. Goldbach, C. Lettre XLIII. In *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle*; Fuss, P.H., Ed.; Imperial Academy of Sciences: St. Petersburg, 1843; Vol. 1, pp. 125–129. (letter to Leonhard Euler) (in German).
2. Oliveira e Silva, T.; Herzog, S.; Pardi, S. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$. *Mathematics of Computation* **2014**, *83*, 2033–2060. https://doi.org/10.1090/S0025-5718-2013-02787-1.
3. Dusart, P. Autour de la fonction qui compte le nombre de nombres premiers. https://www.unilim.fr/pages_perso/pierre.dusart/Documents/T1998_01.pdf, 1998. Accessed: 2026-04-11.
4. Rosser, J.B.; Schoenfeld, L. Approximate formulas for some functions of prime numbers. *Illinois Journal of Mathematics* **1962**, *6*, 64–94. https://doi.org/10.1215/ijm/1255631807.
5. Vega, F. Experimental Results on Goldbach's Conjecture. https://github.com/frankvegadelgado/goldbach, 2026. Accessed: 2026-04-11.