

THE MIDPOINT FORMULATION OF GOLDBACH’S CONJECTURE: RESIDUE CONSTRAINTS ON SYMMETRIC PRIME OFFSETS

CHAIYA TANTISUKAROM*

8 October 2026

Abstract

We reformulate Goldbach’s Conjecture in terms of the midpoint $C = N/2$ and a symmetric offset x , so that $N = (C - x) + (C + x)$. For prime pairs exceeding 3, the offset x is confined to explicit residue classes modulo 6 determined by $C \bmod 6$; this is an elementary mod-6 wheel restriction, and we state exactly how the restricted search relates to the conjecture. We show that shielding a center $C = P_n$ by the first n primes guarantees prime pairs only for $n \leq 3$, and we explain why this is a limitation of trial division rather than a demonstration of the sieve parity barrier. Under Hardy–Littlewood heuristics we derive the expected number of valid offsets and a Poisson-type model for the smallest one, and compare both with an exact computation over every center $C \in [10^5, 2 \cdot 10^5]$. The counting prediction is accurate to within 1%. The leading-order prediction for the mean smallest offset is too large by 11–34%; we trace this to the point offset $x = 0$ (the case C prime) and to the regularity of the sifted set of candidates, and a model incorporating both agrees with the data to within 2%. We also confirm by computation that every center $5 \leq C \leq 10^6$ admits a valid offset, and we report the smallest offset at the centers $C = 6 \cdot 10^j$ for $0 \leq j \leq 23$, which reaches heights far beyond the exhaustive range. The note is expository; no progress toward a proof is claimed.

1 The Midpoint Framework

Goldbach’s Conjecture asserts that every even integer $N \geq 4$ is a sum of two primes. The change of variables used here, $N = (C - x) + (C + x)$, is elementary, and the note makes no claim of new theorems on Goldbach’s problem. Its purpose is to record precisely what the midpoint coordinates make visible (a residue-class restriction on the search, and a clean heuristic for the size of the smallest offset) and to test those heuristics against exact data.

Writing $C = N/2$, any representation $N = p_a + p_b$ with $p_a \leq p_b$ can be written uniquely as

$$p_a = C - x, \quad p_b = C + x, \quad x = \frac{1}{2}(p_b - p_a) \in \mathbb{Z}_{\geq 0}.$$

Define, for integers $C \geq 2$,

$$X(C) = \{x \in \mathbb{Z}_{\geq 0} : C - x \in \mathbb{P}, C + x \in \mathbb{P}\},$$

*C142, drchaiya@gmail.com

and, for $C \geq 5$, the subset of offsets whose pair consists of primes exceeding 3,

$$X^*(C) = \{x \in X(C) : C - x > 3\}.$$

The reason for X^* is that primes greater than 3 are coprime to 6, which yields clean congruence conditions on x (Lemma 1). Representations involving the primes 2 and 3 are handled separately.

Small cases

The three smallest cases are $4 = 2 + 2$, $6 = 3 + 3$ and $8 = 3 + 5$. For $N > 4$ no representation can use the prime 2, since $N - 2$ would be even and larger than 2. Hence for $N \geq 6$ the smaller summand is at least 3, and the only representations excluded from $X^*(C)$ are those of the form $N = 3 + (N - 3)$. From now on we work with $N \geq 10$; the map $N \mapsto C = N/2$ then identifies the even integers $N \geq 10$ with the integers $C \geq 5$, which we sort by their residue $C \bmod 6$.

Lemma 1 (Offset congruences). *Let $C \geq 5$ and let $x \geq 0$ be an integer. Then $\gcd(C - x, 6) = \gcd(C + x, 6) = 1$ if and only if both of the following hold:*

- (i) $x \not\equiv C \pmod{2}$;
- (ii) $3 \mid x$ if $3 \nmid C$, and $3 \nmid x$ if $3 \mid C$.

Combining (i) and (ii) by the Chinese Remainder Theorem gives the admissible classes of $x \bmod 6$ listed in Table 1. In particular every $x \in X^*(C)$ lies in these classes.

Proof. Parity. Since $(C + x) - (C - x) = 2x$ is even, $C - x$ and $C + x$ have the same parity, and both are odd if and only if $C + x$ is odd, i.e. $x \not\equiv C \pmod{2}$.

Modulo 3. If $3 \mid C$, then $C \pm x \equiv \pm x \pmod{3}$, so neither term is divisible by 3 iff $3 \nmid x$. If $3 \nmid C$, then $3 \mid C - x$ iff $x \equiv C \pmod{3}$, and $3 \mid C + x$ iff $x \equiv -C \pmod{3}$. Because $C \not\equiv 0$, the residues C and $-C$ are the two nonzero residues modulo 3, so avoiding both is equivalent to $x \equiv 0 \pmod{3}$.

Combination. The six rows of Table 1 follow by applying (i) and (ii) to each class of $C \bmod 6$. For example, for $C \equiv 2 \pmod{6}$ we have C even, so x is odd, and $C \equiv 2 \pmod{3}$, so $3 \nmid x$; hence $x \equiv 3 \pmod{6}$. $\square$

Table 1: Admissible offset classes for pairs $C \pm x$ coprime to 6.

$C \bmod 6$	Admissible $x \pmod{6}$	Proportion of all x
0	1, 5	1/3
1	0	1/6
2	3	1/6
3	2, 4	1/3
4	3	1/6
5	0	1/6

Corollary 2. *Only a proportion 1/3 of the offsets x is admissible when $3 \mid C$, and only 1/6 when $3 \nmid C$.*

Remark 3. Lemma 1 is a standard mod-6 wheel sieve written in midpoint coordinates. It reduces the number of candidate offsets that must be tested but contains no information about whether a candidate pair is prime. Note also that the admissible values still need $x \leq C - 5$ so that $C - x \geq 5$.

2 Relation to Goldbach's Conjecture

Proposition 4 (Equivalence and one-sided reformulation). *(i) Goldbach's Conjecture holds for all even $N \geq 4$ if and only if $X(C) \neq \emptyset$ for every integer $C \geq 2$.*

(ii) If $X^(C) \neq \emptyset$ for every $C \geq 5$, then Goldbach's Conjecture holds for all even $N \geq 4$.*

(iii) Conversely, Goldbach's Conjecture implies $X^(C) \neq \emptyset$ for every $C \geq 5$ if and only if no even $N \geq 10$ has $N = 3 + (N - 3)$ as its only representation as a sum of two primes.*

Proof. (i) A representation $N = p_a + p_b$ with $p_a \leq p_b$ corresponds to the offset $x = (p_b - p_a)/2 \in X(C)$, with $C = N/2$, and conversely.

(ii) For $N = 4, 6, 8$ use $2 + 2, 3 + 3, 3 + 5$. For even $N \geq 10$ we have $C \geq 5$, and any $x \in X^*(C)$ gives $N = (C - x) + (C + x)$.

(iii) By the discussion of small cases, for $N \geq 10$ the only representations missing from $X^*(C)$ are those with smaller summand equal to 3. $\square$

Part (iii) means the restricted formulation is formally at least as strong as Goldbach's Conjecture. Only even N for which $N - 3$ is prime can be exceptional, and such an N would need $N - p$ composite for every prime $5 \leq p \leq N/2$; heuristically no such N exists, and the difference between the two formulations is expected to be immaterial.

Direct computation (a sieve of Eratosthenes followed by a scan over x) confirms $X^*(C) \neq \emptyset$ for all $5 \leq C \leq 10^6$, i.e. for all even $10 \leq N \leq 2 \cdot 10^6$. Writing $x_{\min}^*(C) = \min X^*(C)$, the largest value over this range is $x_{\min}^* = 1515$, attained at $C = 936118$ (class $C \equiv 4 \pmod{6}$). For comparison, the heuristic of Section 4 predicts a maximum of order $\ln^3 C_0 / (2\Pi_2) \approx 2.0 \times 10^3$ for $C_0 = 10^6$, the same order as the observed value.

3 Small-Prime Shielding and the Parity Barrier

Let $P_n = p_1 p_2 \cdots p_n$ be the n -th primorial.

Proposition 5 (What shielding by small primes gives). *Let $C = P_n$ and let x satisfy $\gcd(x, P_n) = 1$ and $1 \leq x \leq P_n - 2$. Then every prime factor of $P_n \pm x$ exceeds p_n . If, in addition, $P_n + x < p_{n+1}^2$, then both $P_n - x$ and $P_n + x$ are prime.*

Proof. For $p_i \leq p_n$ we have $\gcd(P_n \pm x, p_i) = \gcd(x, p_i) = 1$, which gives the first claim. A composite integer m has a prime factor at most $\sqrt{m}$. If $1 < m < p_{n+1}^2$ has no prime factor $\leq p_n$, then m is prime. Here $P_n - x \geq 2$ and $P_n + x < p_{n+1}^2$ give the second claim. $\square$

Coprimality to P_n requires no bound on x ; the bound $P_n + x < p_{n+1}^2$ is needed only to upgrade “no small prime factor” to “prime”. The next lemma shows that this bound can be met only for small n .

Lemma 6 (Primorials outgrow squares). *For every $n \geq 4$ we have $P_n > p_{n+1}^2$. Consequently, for $n \geq 4$ the condition $P_n + x < p_{n+1}^2$ fails for every $x \geq 0$.*

Proof. For $n = 4$: $P_4 = 210 > 121 = p_5^2$. Suppose $P_n > p_{n+1}^2$ for some $n \geq 4$. By Bertrand's postulate $p_{n+2} < 2p_{n+1}$, and $p_{n+1} \geq 11 > 4$, so

$$P_{n+1} = P_n p_{n+1} > p_{n+1}^3 > 4p_{n+1}^2 > p_{n+2}^2.$$

The second statement follows because $P_n + x \geq P_n$. $\square$

Example 1 ($n = 3$: shielding suffices). For $P_3 = 30$ and $p_4^2 = 49$, the condition is $x \leq 18$ with $\gcd(x, 30) = 1$. Each of $x = 1, 7, 11, 13, 17$ produces a prime pair: $(29, 31)$, $(23, 37)$, $(19, 41)$, $(17, 43)$, $(13, 47)$.

Example 2 ($n = 4$: the sufficient condition fails). By Lemma 6 the second hypothesis of Proposition 5 cannot hold for $C = P_4 = 210$, so shielding no longer *guarantees* a prime pair; it may still produce one. Take $x = 37$, which is coprime to 210:

$$\begin{aligned} 210 - 37 &= 173 \quad (\text{prime}), \\ 210 + 37 &= 247 = 13 \times 19 \quad (\text{composite}). \end{aligned}$$

Although 247 has no prime factor ≤ 7 , it is composite. By contrast, $x = 13$ (also coprime to 210) gives the prime pair $(197, 223)$. Neither outcome contradicts Proposition 5: the second hypothesis is needed for a guarantee, not for primality.

Remark 7 (Relation to the parity barrier). Selberg's parity phenomenon [3] states that sieve methods based on information about small prime divisors cannot separate integers with an even number of prime factors from those with an odd number. This is a deeper obstruction than the failure seen above, which merely reflects that trial division by the primes up to p_n certifies primality only below p_{n+1}^2 . The integer $247 = 13 \times 19$ is a semiprime that survives sieving by 2, 3, 5, 7; it is the kind of object that low-level sieving cannot exclude, but a single example does not exhibit the barrier, which concerns what sieve weights can detect even when sieving to a high level. In the Goldbach setting the barrier is the reason sieve methods yield results such as Chen's theorem [4] (every large even N is $p + P_2$) but not $p + p'$; see [6].

4 Heuristic Estimates and Computation

4.1 Expected number of offsets

Let $\Pi_2 = \prod_{p>2} (1 - (p-1)^{-2}) \approx 0.6601618$ denote the twin prime constant, and for even N let

$$\mathfrak{S}(N) = 2\Pi_2 \prod_{\substack{p|N \\ p>2}} \frac{p-1}{p-2}.$$

Conjecture 8 (Hardy–Littlewood [1]). *As $C \rightarrow \infty$,*

$$|X(C)| \sim \frac{1}{2} \mathfrak{S}(2C) \int_2^{2C-2} \frac{dt}{\ln t \ln(2C-t)} \sim \Pi_2 \prod_{\substack{p|C \\ p>2}} \frac{p-1}{p-2} \cdot \frac{2C}{\ln^2(2C)}. \quad (1)$$

The factor $\frac{1}{2}$ accounts for ordered representations counting each pair twice (as x and $-x$). The single point $x = 0$ contributes at most 1 to the count and is negligible there, although it matters for the smallest offset (Section 4.2). The two asymptotic forms agree to leading order, but the integral form is markedly more accurate at finite height: the ratio of the integral to $2C/\ln^2(2C)$ is about 1.05 at $N = 10^{18}$. The extra weight in the integral comes from large offsets, where the logarithms $\ln t$ and $\ln(N - t)$ are smaller; at small offsets the local pair density is $\mathfrak{S}/\ln^2 C$ to leading order.

The factor $(p - 1)/(p - 2)$ at $p = 3$ equals 2, which is the counting-level counterpart of Corollary 2: centers divisible by 3 have about twice as many offsets as those that are not.

4.2 The smallest offset

Conjecture 9 (Poisson heuristic for $x_{\min}$, leading order). *Set $\lambda(C) = \mathfrak{S}(2C)/\ln^2 C$. For $1 \ll X \ll C$, the expected number of $x \leq X$ in $X(C)$ is approximately λX . Modeling the valid offsets as a Poisson process in the spirit of Cramér’s random model [2],*

$$\Pr[x_{\min}(C) > X] \approx e^{-\lambda X}, \quad \mathbb{E} x_{\min}(C) \approx \frac{\ln^2 C}{\mathfrak{S}(2C)}.$$

Consequently, over all centers $C \leq C_0$, the largest value of $x_{\min}(C)$ is of order $\lambda_{\min}^{-1} \ln C_0$, where $\lambda_{\min} = 2\Pi_2/\ln^2 C_0$ corresponds to the smallest singular series; centers with $\mathfrak{S}(2C)$ close to $2\Pi_2$ (for instance $C = 2p$ with p prime) are plentiful. This is a typical size, about $\ln^3 C_0/(2\Pi_2)$, not an upper bound.

The growth of $x_{\min}$ like $\ln^2 C$ is illustrated at heights up to $6 \cdot 10^{23}$ in Figure 1 (Section 4.4). Three caveats apply to the model.

1. *Not a uniform bound.* $x_{\min} = O(\ln^2 C)$ is not uniform under this model: for any fixed K , the event $x_{\min} > K \ln^2 C$ has probability about $e^{-K\mathfrak{S}}$, not $o(1)$, so $O(\ln^2 C)$ describes only the typical center.
2. *The point $x = 0$.* Offset 0 lies in $X(C)$ exactly when C is prime. For $C \equiv \pm 1 \pmod{6}$ this happens with probability about $3/\ln C$ (the density of primes among integers coprime to 6), and then $x_{\min} = 0$; for all other $C \geq 5$ the center is composite and $x = 0$ is impossible. This is a point mass at zero that the uniform-intensity Poisson model does not see, and it lowers $\mathbb{E} x_{\min}$ by the factor $1 - 3/\ln C$ for $C \equiv \pm 1 \pmod{6}$ (about 25% at $C \approx 1.5 \times 10^5$, about 7% at $C \approx 5 \times 10^{17}$).
3. *Regularity of the sifted set.* The candidates for x are the survivors of a sieve (first the lattice of Lemma 1, then sieving by 5, 7, 11, ...), and survivors of a sieve are far more evenly spaced than the points of a Poisson process. A first hit among regularly spaced candidates occurs earlier than in the Poisson model, so the mean and median of $x_{\min}$ are smaller than the leading-order values. This changes the *offset* of the distribution but not the exponential rate λ of its tail, so the extremal estimate above is unaffected at this level.

Example 3 (Predictions at $N = 10^{18}$). Let $N = 10^{18}$, so $C = 5 \times 10^{17} = 6.833333333333333 + 2$, a center of class $C \equiv 2 \pmod{6}$.

- By Lemma 1, admissible offsets satisfy $x \equiv 3 \pmod{6}$ (odd and divisible by 3). The center is even, so $x = 0$ is not available.
- The only odd prime dividing 10^{18} is 5, so $\mathfrak{S}(10^{18}) = 2\Pi_2 \cdot \frac{4}{3} = \frac{8}{3}\Pi_2 \approx 1.7604$.
- With $N/\ln^2 N \approx 5.8213 \times 10^{14}$, the leading-order prediction is

$$|X(5 \times 10^{17})| \approx \frac{1}{2}\mathfrak{S} \frac{N}{\ln^2 N} \approx 5.12 \times 10^{14},$$

and the integral form of Conjecture 8 gives approximately 5.39×10^{14} .

- Here $\ln^2 C \approx 1660.8$, so Conjecture 9 predicts $\mathbb{E} x_{\min} \approx 9.4 \times 10^2$ with median about 6.5×10^2 , i.e. $x_{\min}/C \approx 1.9 \times 10^{-15}$. The largest $x_{\min}$ over all centers up to C is predicted to be a few times 10^4 (about 4×10^4 at $\mathfrak{S} = 1.76$, and up to about 5×10^4 if $\mathfrak{S}$ is as small as $2\Pi_2$).

These are predictions, not data. By caveat 3 and the computation of Section 4.3, the leading-order mean and median are probably too large by roughly the 10% seen at $C \approx 10^5$; our data do not determine the size of this correction at 10^{18} . The value $N = 10^{18}$ lies within the range covered by the published verification [8], but that computation records minimal Goldbach partitions (smallest prime) rather than symmetric offsets, so it contains no data on $x_{\min}$. We have not computed $x_{\min}$ at this particular center; Figure 1 gives exact values at the neighboring centers $6 \cdot 10^{17}$ and $6 \cdot 10^{18}$ (class 0, hence with a different singular series), namely $x_{\min} = 157$ and $x_{\min} = 23$, against leading-order means of about 476 and 531 respectively.

4.3 Direct computation at moderate height

To test Conjectures 8 and 9 at an accessible height we computed $X(C)$ exactly, by a sieve of Eratosthenes and a scan over x , for *every* center $C \in [10^5, 2 \cdot 10^5]$ (100 001 centers, about 16 667 per class of $C \pmod{6}$). Each reported mean is therefore an exact average over its class, not a sample estimate, and every deviation from a prediction is a deviation of the model rather than sampling noise. The count $|X(C)|$ includes all prime pairs (including any with the prime 3), and $x_{\min}$ is the least element of $X(C)$ (so it equals 0 when C is prime). Predictions are computed center by center and then averaged.

Table 2: Mean $|X(C)|$ over all centers in each class, $C \in [10^5, 2 \cdot 10^5]$. “Pred.” is the integral form of Conjecture 8.

$C \pmod{6}$	Centers	Obs.	Pred.	Obs./Pred.
0	16 667	3359.4	3375.0	0.995
1	16 667	1673.4	1687.5	0.992
2	16 667	1686.7	1687.5	1.000
3	16 666	3359.5	3375.1	0.995
4	16 667	1673.3	1687.5	0.992
5	16 667	1686.8	1687.5	1.000

The counts agree with the prediction to within 1% in every class, and the classes with $3 \mid C$ ($C \equiv 0, 3 \pmod{6}$) have about twice as many offsets, as predicted.

For $x_{\min}$ we compare the data with the leading-order value $\ln^2 C/\mathfrak{S}(2C)$ of Conjecture 9 and with two refined models. Both treat $x = 0$ separately from $x \geq 1$. For a sieve level y , let $\nu_C(2) = 1$, $\nu_C(p) = 1$ for odd $p \mid C$, and $\nu_C(p) = 2$ for all other primes p , and put

$$A_y(C) = \{x \geq 1 : p \nmid C^2 - x^2 \text{ for all primes } p \leq y\}, \quad \delta_y(C) = \prod_{p \leq y} \left(1 - \frac{\nu_C(p)}{p}\right).$$

$A_y(C)$ is the set of offsets whose pair $C \pm x$ has no prime factor $\leq y$, and $\delta_y(C)$ is its density. *Model B_y* : each $x \in A_y(C)$ lies in $X(C)$ independently with probability $q_y(C) = \lambda(C)/\delta_y(C)$, which preserves the overall intensity λ ; and $x = 0$ lies in $X(C)$ with probability $\pi_y(C) = (\ln C \prod_{p \leq y} (1 - 1/p))^{-1}$ if C has no prime factor $\leq y$, and 0 otherwise. If $x_1 < x_2 < \dots$ enumerate $A_y(C)$, then

$$\mathbb{E} x_{\min} = (1 - \pi_y) \sum_{k \geq 1} x_k q_y (1 - q_y)^{k-1}.$$

Model A is B_3 : its candidate set is exactly the admissible lattice of Lemma 1, and $\pi_3 = 3/\ln C$ for $C \equiv \pm 1 \pmod{6}$, so it incorporates caveat 2 and the lattice structure but no further sieving. *Model B* is B_{97} , with sieving by all primes up to 97 (survivor lists truncated at $x \leq 2000$; the neglected tail has probability below 10^{-7}). The tail of $x_{\min}$ in every B_y decays at the same rate $e^{-\lambda x}$ as in Conjecture 9.

Table 3: Mean $x_{\min}$ over all centers in each class, $C \in [10^5, 2 \cdot 10^5]$. “Lead.” is $\ln^2 C/\mathfrak{S}(2C)$; Models A and B are defined in the text.

$C \bmod 6$	Obs.	Lead.	A	Obs./A	B	Obs./B
0	42.2	48.1	46.6	0.905	42.2	1.001
1	63.8	96.3	72.0	0.887	64.2	0.995
2	85.9	96.3	93.3	0.921	85.2	1.008
3	42.9	48.1	46.7	0.920	42.9	1.002
4	85.1	96.3	93.3	0.912	84.6	1.006
5	66.2	96.3	72.0	0.919	64.9	1.019

The results are as follows.

- The leading-order value overestimates the observed mean by 11–34%. The largest deficits occur in the classes $C \equiv \pm 1 \pmod{6}$ (observed/leading 0.66 and 0.69).
- The point $x = 0$ accounts for most of that excess deficit. In classes 1 and 5 we find $x_{\min} = 0$ for 25.2% and 25.1% of centers, against a predicted average of $3/\ln C$ of 25.2%. Because of this point mass the distribution of $x_{\min}$ is not close to exponential in these classes. Model A, which includes it, raises the ratio in these classes from about 0.67 to 0.89–0.92.
- Model A still overestimates by 8–11% in every class, with almost the same ratio in each. This is the signature of caveat 3: sieving by primes beyond 3 makes the candidate set more regular, and Model B, which thins the survivors of a sieve to level 97, agrees with the data to within 2% in every class. Moving from level 3 to level 97 moves the ratio from about 0.9 to about 1.0.

Model B has a free parameter y and is a heuristic description, not a derivation; its agreement at one height supports the structure of the explanation (the point $x = 0$ plus a regular sifted set) but does not test the large- C behaviour. The counting prediction and the tail rate λ are the parts of the heuristic that this computation supports directly.

4.4 Smallest offsets at the centers $C = 6 \cdot 10^j$

The exhaustive computation above is limited to moderate height. To see the behaviour of $x_{\min}$ far beyond it, Figure 1 plots $x_{\min}(C)$ for the family of centers $C = 6 \cdot 10^j$, $0 \leq j \leq 23$ (all of class 0 mod 6), against the leading-order prediction $\ln^2 C / \mathfrak{S}(2C)$ of Conjecture 9. For each center we scanned $x = 0, 1, 2, \dots$ and tested $C \pm x$ for primality with the deterministic Miller–Rabin test to the thirteen bases $2, 3, \dots, 41$, which is exact for all integers below $3.3 \cdot 10^{24}$ and therefore covers every number tested. For $j \geq 1$ the odd primes dividing C are 3 and 5, so $\mathfrak{S}(2C) = 2\Pi_2 \cdot 2 \cdot \frac{4}{3} = \frac{16}{3}\Pi_2 \approx 3.52$; for $j = 0$, $\mathfrak{S}(12) = 4\Pi_2$.

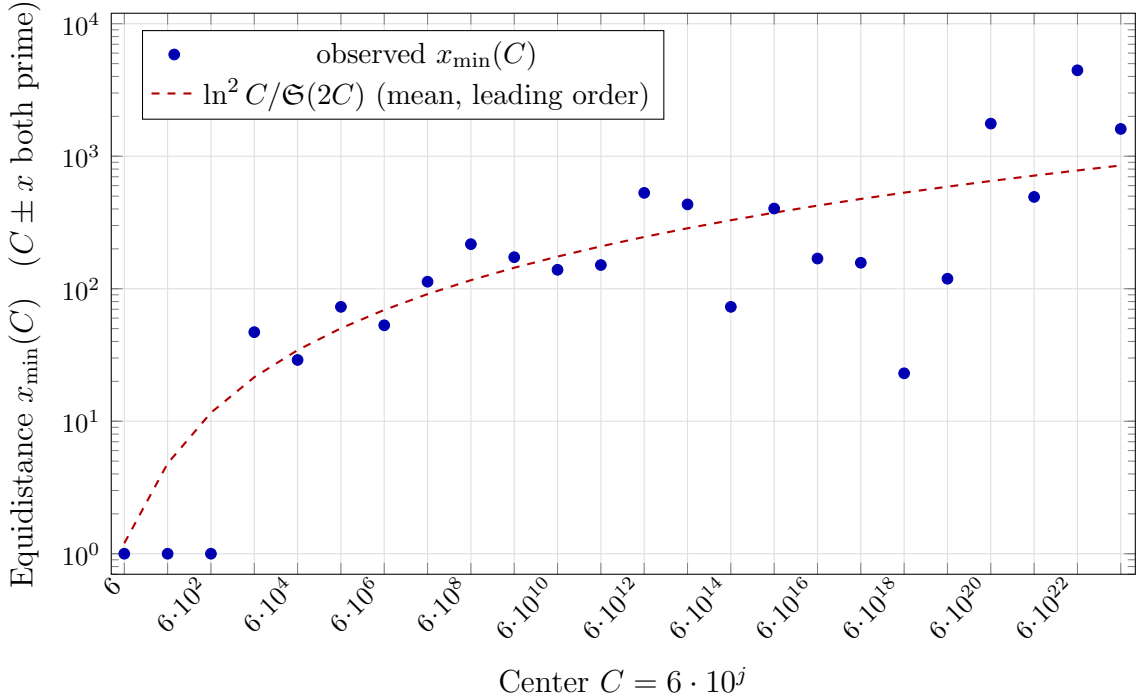


Figure 1: Smallest symmetric offset $x_{\min}(C)$ at the centers $C = 6 \cdot 10^j$, $0 \leq j \leq 23$, against the leading-order prediction $\ln^2 C / \mathfrak{S}(2C)$ for its mean (Conjecture 9). Each dot is a single center, so the scatter about the dashed line is that of a single draw from a roughly exponential distribution rather than of a mean. Both axes are logarithmic; unlabeled ticks mark the odd powers of ten.

Every smallest offset in Figure 1 lies in the classes $x \equiv 1, 5 \pmod{6}$ required by Lemma 1 for $C \equiv 0 \pmod{6}$ (the first values are 1, 1, 1, 47, 29, 73, $\dots$), and $x = 0$ never occurs since C is composite. The dashed curve rises from about 1 to about 850 over the 24 decades shown, and the observed points follow it with the large scatter expected of a single exponential-type variable. The three smallest centers, $j \leq 2$, have $x_{\min} = 1$ (the twin-prime pairs (5, 7), (59, 61), (599, 601)); they are too small for the asymptotic prediction to be meaningful and are excluded from the following summary. For the

remaining 21 centers, $3 \leq j \leq 23$, write $r = x_{\min}/(\ln^2 C/\mathfrak{S}(2C))$; under the model, r is approximately exponentially distributed with mean 1 and median $\ln 2 \approx 0.69$.

- The observed ratios have mean 1.33 and median 1.07, with 10 points below the dashed line and 11 above. The mean is driven by one outlier; the standard error of a mean of 21 unit-mean exponentials is 0.22.
- The largest ratio is 5.70, at $C = 6 \cdot 10^{22}$ ($x_{\min} = 4453$). This has probability $e^{-5.70} \approx 0.3\%$ for a single center, and about 7% that at least one of the 21 centers is this extreme, so it is unusual but not alarming. The smallest ratio is 0.043, at $C = 6 \cdot 10^{18}$ ($x_{\min} = 23$); this has probability about 4% for a single center and about 60% that at least one of 21 is this small.
- Fifteen of the 21 ratios exceed the model median 0.69. Treated as a post hoc test this is a mild excess (one-sided binomial $p \approx 0.04$). We read nothing into it, but note that the refined models of Section 4.3 predict a shift in the opposite direction, toward *smaller* values than the leading-order estimate.

With 21 single-draw points the figure cannot test the constant in the prediction or detect a correction of order 10%; it illustrates the order of magnitude $\ln^2 C$ and the heavy exponential-type scatter. The quantitative test remains the exhaustive computation of Section 4.3.

5 Conclusion

Writing $N = (C - x) + (C + x)$ turns the search for a Goldbach representation into a search over symmetric offsets, and for pairs exceeding 3 the offsets are confined to the residue classes of Table 1, a mod-6 wheel restriction that removes 2/3 or 5/6 of candidates depending on whether $3 \mid C$. Proposition 4 makes the relation to Goldbach’s Conjecture precise. Proposition 5 and Lemma 6 show that small-prime shielding of $C = P_n$ guarantees prime pairs only for $n \leq 3$; beyond that it no longer guarantees primes, although many shielded offsets still give prime pairs. This is reminiscent of, but should not be mistaken for, the sieve parity barrier. The Hardy–Littlewood heuristic predicts the number of offsets, and an exact computation over $C \in [10^5, 2 \cdot 10^5]$ confirms the counts to within 1% (Table 2). The same heuristic predicts the tail of the smallest offset; its leading-order mean is too large at moderate height, and the discrepancy is explained quantitatively by the point offset $x = 0$ and the regularity of the sifted candidate set (Table 3). At the centers $6 \cdot 10^j$ up to $j = 23$ the smallest offsets follow the predicted $\ln^2 C$ growth with the expected scatter (Figure 1). None of this approaches a proof. The notable unconditional results remain Chen’s theorem [4], the density-zero exceptional set [5], Helfgott’s proof of the ternary conjecture [7], and the published verification of the binary conjecture up to $4 \cdot 10^{18}$ [8].

References

- [1] G. H. Hardy and J. E. Littlewood, Some problems of “Partitio Numerorum”: III. On the expression of a number as a sum of primes, *Acta Math.* **44** (1923), 1–70.

- [2] H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, *Acta Arith.* **2** (1936), 23–46.
- [3] A. Selberg, On elementary methods in prime number theory and their limitations, *Proc. 11th Scandinavian Math. Congress, Trondheim, 1949*, Johan Grundt Tanums Forlag, Oslo, 1952, 13–22.
- [4] J.-R. Chen, On the representation of a large even integer as the sum of a prime and the product of at most two primes, *Sci. Sinica* **16** (1973), 157–176.
- [5] H. L. Montgomery and R. C. Vaughan, The exceptional set in Goldbach’s problem, *Acta Arith.* **27** (1975), 353–370.
- [6] J. Friedlander and H. Iwaniec, *Opera de Cribro*, AMS Colloquium Publications **57**, 2010.
- [7] H. A. Helfgott, The ternary Goldbach conjecture is true, arXiv:1312.7748 (2013).
- [8] T. Oliveira e Silva, S. Herzog and S. Pardi, Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$, *Math. Comp.* **83** (2014), 2033–2060.