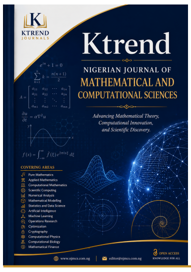


Ktrend - Nigerian Journal of Mathematical and Computational Sciences

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

Computational Visualization of Symmetric Prime Pair Distributions Around Even Integers

Ifeyinwa Eunice Daniel[1], Olufemi Johnson Ogunsola[2]

[1]Department of Mathematics and Computer Science, Spiritan University Nneochi, Abia State, Nigeria
[2]Department of Mathematics, Federal University of Agriculture, Abeokuta, Nigeria
Corresponding author: ifeyinwa.daniel@spiritanuniversity.edu.ng

Abstract. *This paper develops a substantially expanded computational and visual study of symmetric prime-pair distributions around even integers. For an even integer* $2n$*, a symmetric prime pair* (p_1, p_2) *satisfies* $p_1 + p_2 = 2n$*, and the nearest symmetric distance is* $d_n = \min\{k \geq 0 : n - k \text{ and } n + k \text{ are prime}\}$*. The normalized quantity* $R_n = d_n/n$*, introduced by Daniel and Ogunsola, measures the relative displacement of the nearest Goldbach pair from the midpoint. An exhaustive deterministic computation is carried out for every midpoint* $2 \leq n \leq 500{,}000$*, corresponding to all* $499{,}999$ *even integers* $4 \leq 2n \leq 10^6$*. The analysis combines exact enumeration, record-value analysis, logarithmic binning, empirical quantiles, frequency distributions and asymptotic heuristics. The largest observed ratio is* $R_{22} = 9/22 \approx 0.4090909$*, attained at* $2n = 44$*, and no later value in the tested range exceeds it. Scale-dependent means, medians and upper quantiles decline sharply, while the running maximum stabilizes at* $9/22$*. Mathematical propositions establish elementary structural properties of* d_n *and* R_n*, and a Hardy–Littlewood/Cramér-type heuristic predicts a typical scale* d_n *of polylogarithmic order and hence* $R_n \to 0$ *heuristically. The findings provide stronger finite-range evidence for boundedness and relative decay, but they neither prove Goldbach's conjecture nor establish a universal upper bound.*

Keywords: symmetric prime pairs; Goldbach representations; computational number theory; prime distribution; relative distance ratio; asymptotic decay; data visualization.

1. Introduction

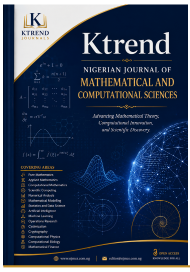
The additive structure of prime numbers is one of the central themes of analytic and computational number theory. Goldbach's conjecture asserts that every even integer greater than 2 is a sum of two primes. Although the full binary conjecture remains unresolved, major theoretical and computational advances have shaped its modern study, including the circle method of Hardy and Littlewood [2], Vinogradov-type results, Chen's theorem [3], numerical verification over very large ranges [6], and Helfgott's proof of the ternary Goldbach conjecture [4, 5].

For a fixed even integer $2n$, every Goldbach representation

$$2n = p_1 + p_2$$

may be interpreted geometrically as two primes placed symmetrically about the midpoint n, because

$$n - p_1 = p_2 - n.$$



Ktrend - Nigerian Journal of Mathematical and Computational Sciences

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

This midpoint formulation converts an additive representation into a local distance problem. Daniel and Ogunsola [1] formalized this viewpoint by introducing the nearest symmetric prime distance d_n and the normalized relative distance ratio $R_n = d_n/n$. Their computations for selected even integers up to 10^6 suggested that R_n is bounded and becomes small as n grows.

The present paper extends that foundational work in four directions. First, it replaces selected sampling by exhaustive enumeration of every midpoint in the interval. Second, it gives a more explicit mathematical foundation for the distance function and its relation to Goldbach representations. Third, it introduces record, quantile, bin-wise and frequency-based visualizations that distinguish local irregularity from large-scale decay. Fourth, it places the observed behavior in the context of the Prime Number Theorem, the Hardy–Littlewood prime-pair heuristic, Cramér's random model and known work on prime gaps [11, 12, 14, 15].

The contribution is computational rather than a proof of an infinite theorem. In particular, finding a symmetric prime pair for all even integers up to 10^6 is a finite verification only. Likewise, the stabilization of the running maximum at $9/22$ is strong numerical evidence within the tested range, not a proof that $9/22$ is a universal maximum.

2. Literature Review and Mathematical Context

Goldbach's conjecture has motivated a large body of work in additive prime theory. The Hardy–Littlewood circle method led to a conjectural asymptotic formula for the number of representations of a large even integer as a sum of two primes [2, 7]. The formula contains the singular series, which captures arithmetic restrictions arising from residue classes. Modern expositions of the conjecture and its history are provided by Wang [9] and Vaughan [8].

On the computational side, extensive verification has established the binary Goldbach conjecture over enormous finite intervals. Oliveira e Silva, Herzog and Pardi [6] verified it through 4×10^{18}, demonstrating the power of optimized sieving and distributed computation. Such verification does not settle the conjecture, but it supplies important evidence and tests the behavior of representation counts.

The distance from a midpoint to nearby primes is also connected with prime-gap phenomena. The Prime Number Theorem [16, 17] predicts average prime density near x of approximately $1/\log x$. Cramér's probabilistic model [11] treats primality events as approximately random with this density, while Granville [12] explains why arithmetic correlations require refinements. Results on bounded gaps between primes by Zhang [13], Maynard [14] and Tao [15] show that small gaps occur infinitely often, although these theorems address a different problem from symmetric Goldbach pairs.

Computational visualization is particularly useful in this setting because individual prime events are irregular, whereas normalized statistics may display stable scale-dependent behavior. The present work therefore emphasizes not only individual values of R_n, but also empirical distributions, running records and logarithmically binned summaries. This distinction prevents a misleading interpretation of "decay": the sequence R_n is not monotone, but its typical and upper-tail magnitudes may decrease with n.

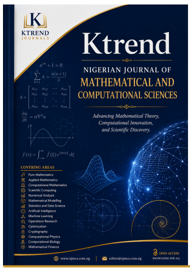
3. Definitions and Structural Results

Definition 3.1 (Symmetric prime pair). *Let $n \geq 2$. A pair of primes (p_1, p_2) is called a symmetric prime pair about n if*

$$p_1 + p_2 = 2n.$$

Equivalently, there exists an integer $k \geq 0$ such that $p_1 = n - k$ and $p_2 = n + k$.

Definition 3.2 (Nearest symmetric distance). *Whenever at least one symmetric prime pair exists*



Ktrend - Nigerian Journal of Mathematical and Computational Sciences

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

about n, define

$$d_n = \min\{k \in \mathbb{N}_0 : n-k \text{ and } n+k \text{ are prime}\}.$$

The associated nearest pair is $(n-d_n, n+d_n)$.

Definition 3.3 (Relative distance ratio)**.** *The normalized nearest symmetric distance is*

$$R_n = \frac{d_n}{n}.$$

Proposition 3.4 (Equivalence with a Goldbach representation)**.** *For $n \geq 2$, d_n exists if and only if the even integer $2n$ has a Goldbach representation.*

Proof. If d_n exists, then $n-d_n$ and $n+d_n$ are prime and their sum is $2n$, so $2n$ has a Goldbach representation. Conversely, if $2n = p_1 + p_2$ with primes $p_1 \leq p_2$, then $n = (p_1+p_2)/2$ and $k = (p_2 - p_1)/2$ is a nonnegative integer satisfying $p_1 = n-k$ and $p_2 = n+k$. The finite nonempty set of such distances has a least element, namely d_n. □

Proposition 3.5 (Basic bounds)**.** *Whenever d_n exists,*

$$0 \leq d_n \leq n-2, \qquad 0 \leq R_n \leq 1 - \frac{2}{n} < 1.$$

Proof. Nonnegativity is immediate. If (p_1, p_2) is any Goldbach pair with $p_1 \leq p_2$, then $p_1 \geq 2$ and $d_n = n - p_1 \leq n-2$. Division by n yields the bound for R_n. □

Proposition 3.6 (Zero ratio criterion)**.** *For $n \geq 2$, $R_n = 0$ if and only if n is prime.*

Proof. If n is prime, then (n, n) is a symmetric prime pair, so $d_n = 0$. Conversely, if $R_n = 0$, then $d_n = 0$, and the defining pair is (n, n); hence n is prime. □

Lemma 3.7 (Parity restriction)**.** *If $n > 3$ and $d_n > 0$, then d_n has parity opposite to that of n, except when one member of the pair equals 2.*

Proof. For a nontrivial pair in which both primes exceed 2, both $n-d_n$ and $n+d_n$ are odd. Therefore n and d_n have opposite parity. The exceptional case occurs when one prime is 2. □

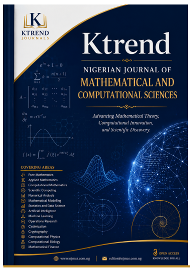
Definition 3.8 (Record envelope)**.** *For $N \geq 2$, define the running record*

$$M(N) = \max_{2 \leq n \leq N} R_n.$$

Then $M(N)$ is nondecreasing and piecewise constant.

Proposition 3.9 (Interpretation of a stabilized record)**.** *If $M(N) = c$ for every $N_0 \leq N \leq X$, then $R_n \leq c$ for all $n \leq X$. This gives a finite-range bound only and does not imply $R_n \leq c$ for $n > X$.*

Proof. The first assertion follows directly from the definition of the maximum. The second follows because no information about indices beyond X is contained in the finite computation. □



Ktrend - Nigerian Journal of Mathematical and Computational Sciences

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

4. Computational Methodology

4.1. Exhaustive search

A Boolean sieve of Eratosthenes is constructed up to 10^6. For each midpoint $n = 2, 3, \ldots, 500{,}000$, the algorithm tests $k = 0, 1, 2, \ldots$ until both $n-k$ and $n+k$ are marked prime. Because the largest tested endpoint $n+k$ cannot exceed $2n \leq 10^6$ for a valid pair, the sieve range is sufficient.

Algorithm 1 Nearest symmetric prime-pair search

```
Require: Upper even bound X = 10^6
  Construct primality table P[0:X] by the sieve of Eratosthenes
  for n = 2 to X/2 do
      k ← 0
      while not (P[n − k] = true and P[n + k] = true) do
          k ← k + 1
      end while
      d_n ← k, R_n ← k/n
      store (2n, n, n − k, n + k, d_n, R_n)
  end for
```

4.2. Complexity and reproducibility

The sieve requires $O(X \log\log X)$ time and $O(X)$ memory. If $K_n = d_n$ denotes the number of search steps for midpoint n, the second stage requires $O(\sum_{n \leq X/2}(K_n + 1))$ Boolean lookups. The observed smallness of most d_n makes the search efficient in practice. All calculations use exact integer arithmetic for the primality tests and distances; floating-point arithmetic is used only to report ratios and descriptive statistics.

The experiment contains all 499,999 midpoints in the specified interval, not a random sample. Summary statistics are therefore exact descriptive statistics for the finite dataset. Their interpretation remains empirical with respect to the infinite sequence.

4.3. Visualization and summary statistics

Four main visual summaries are used:

(i) a scatter plot of positive R_n values on logarithmic axes;
(ii) the running maximum $M(N)$;
(iii) logarithmically binned means, medians and 90th percentiles;
(iv) a frequency plot of positive distances $d_n \leq 150$.

The bin-wise statistics separate the shrinking scale of typical values from non-monotone local fluctuations.

5. Numerical Results

The exhaustive computation found at least one symmetric prime pair for every even integer $4 \leq 2n \leq 10^6$. This is consistent with much stronger published finite verification results [6], but it is not a proof of Goldbach's conjecture.

Among the 499,999 tested midpoints, 41,538 satisfy $R_n = 0$, exactly corresponding to prime midpoints. The overall mean ratio is approximately 0.000635079, while the median is approximately 0.000204479. The largest observed ratio is

$$R_{22} = \frac{9}{22} \approx 0.4090909,$$

corresponding to $2n = 44$ and the nearest pair $(13, 31)$. This agrees with the foundational computation

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

of Daniel and Ogunsola [1].

Table 1: Scale-dependent summaries of the exhaustive dataset.

Range of $2n$	Count	Mean	Median	90th percentile	Maximum
4–100	49	0.117236	0.088235	0.285000	0.409091
102–1,000	450	0.043037	0.028186	0.105993	0.258621
1,002–10,000	4,500	0.010968	0.006454	0.026643	0.165399
10,002–100,000	45,000	0.002051	0.001140	0.004995	0.058468
100,002–1,000,000	450,000	0.000335	0.000181	0.000815	0.010479

Table 1 shows a pronounced scale effect. From the first to the last range, the mean decreases by more than two orders of magnitude, the median falls from about 8.82×10^{-2} to 1.81×10^{-4}, and the 90th percentile falls from 0.285 to approximately 8.15×10^{-4}. The maximum within each successive range also decreases after the early record at 44.

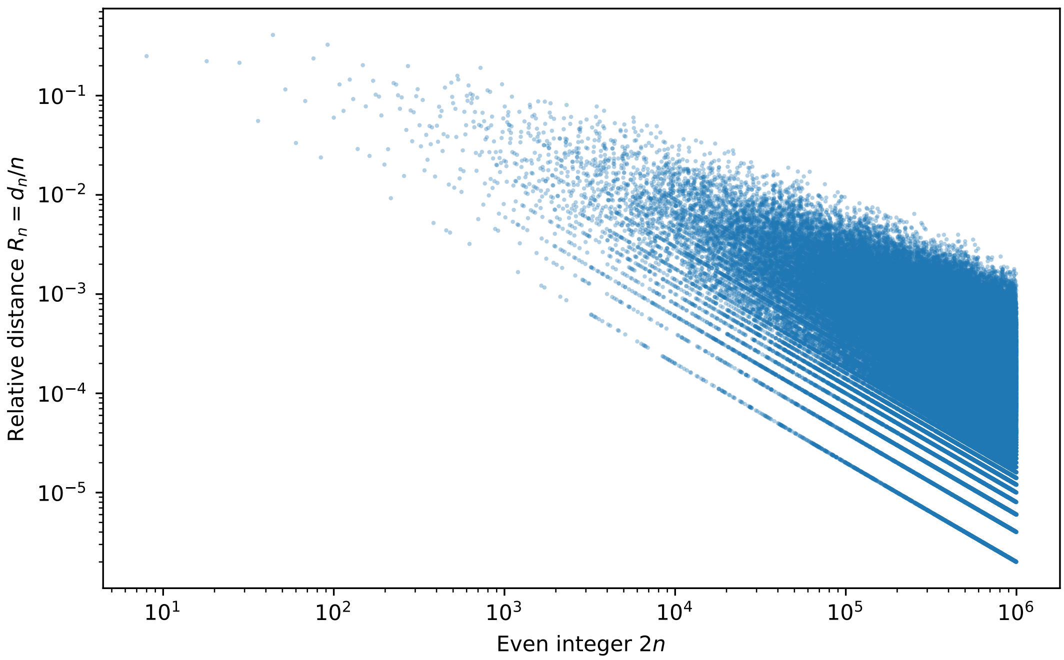


Figure 1: Scatter visualization of positive relative distance ratios R_n for even integers up to 10^6. Logarithmic scaling reveals large early ratios and a dense cloud of much smaller values at larger scales. Zero values are excluded from the logarithmic vertical axis.

Figure 1 demonstrates that individual ratios fluctuate irregularly, as expected from the local irregularity of primes. Nevertheless, the upper envelope and the main concentration of the point cloud decline as the even integers increase.

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

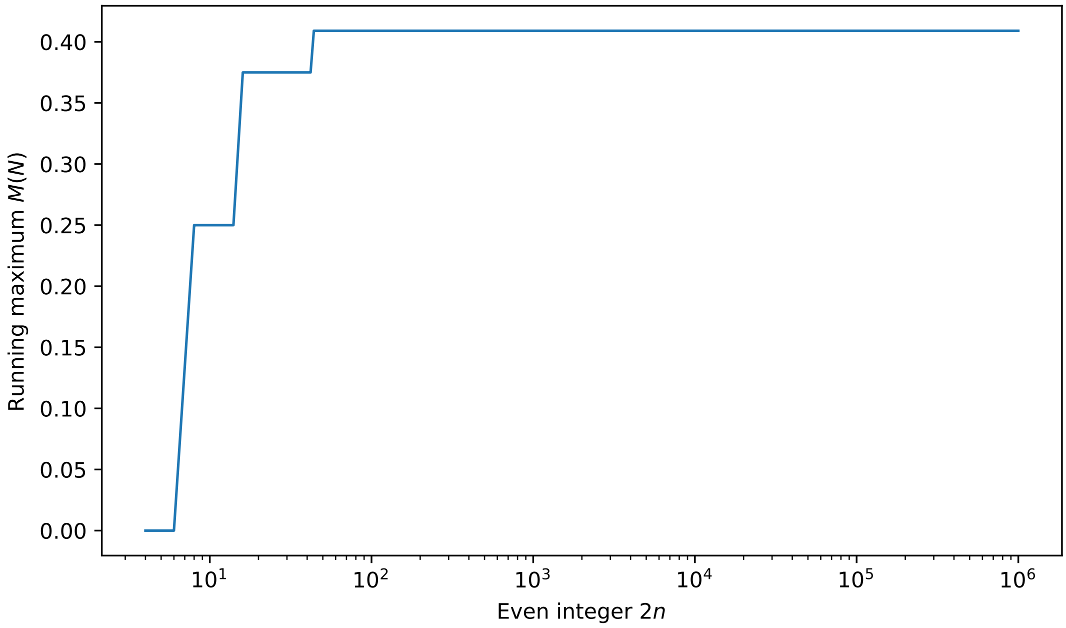


Figure 2: Running maximum $M(N)$ of R_n. The record reaches $9/22$ at $2n = 44$ and is not exceeded through 10^6.

The running maximum in Figure 2 rises through a few early records and then stabilizes at $9/22$. Thus,

$$R_n \leq \frac{9}{22} \qquad (2 \leq n \leq 500{,}000)$$

is an exact statement for the computed interval.

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

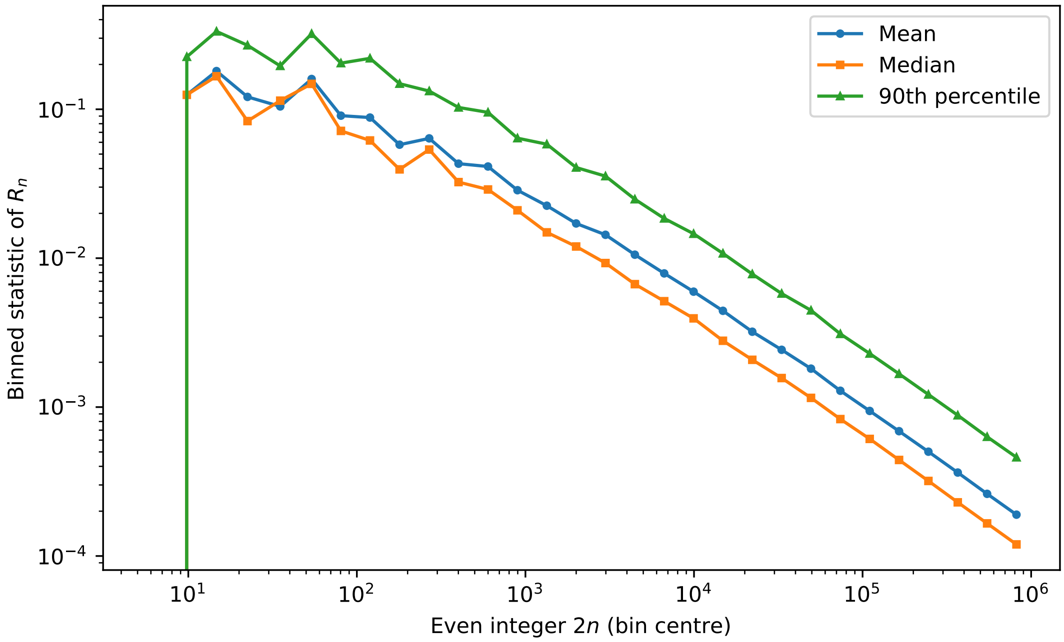


Figure 3: Logarithmically binned mean, median and 90th percentile of R_n. All three scale-dependent statistics decline strongly with increasing $2n$.

Figure 3 gives a clearer representation of decay than the raw scatter plot. The median lies below the mean in most bins, indicating a right-skewed distribution in which occasional larger distances raise the average. The decline of the 90th percentile shows that the upper tail also contracts relative to n.

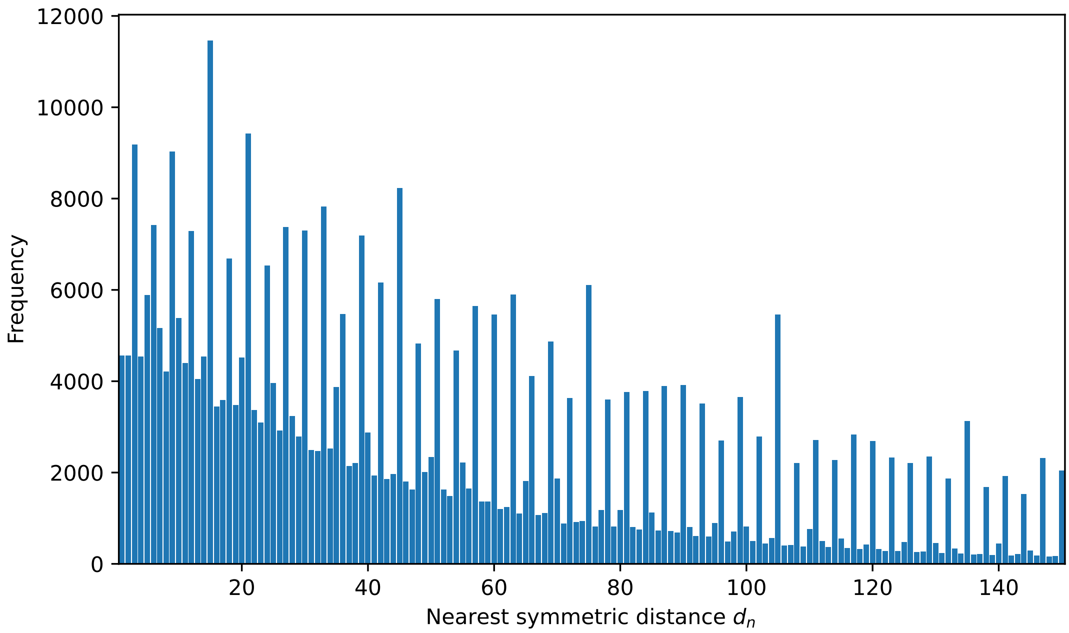


Figure 4: Frequency distribution of positive nearest symmetric distances $d_n \leq 150$. Small distances dominate, while larger local excursions occur with lower frequency.

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

The frequency plot in Figure 4 is compatible with the expectation that nearby symmetric pairs are common. Its oscillatory structure also reflects parity and residue-class restrictions rather than a smooth continuous law.

6. Asymptotic Heuristic

Near a large number n, the Prime Number Theorem suggests that an integer is prime with probability roughly $1/\log n$ [16, 17]. If the events that $n-k$ and $n+k$ are prime were independent, their simultaneous probability would be approximately

$$\frac{1}{(\log n)^2}.$$

A geometric waiting-time argument would then suggest a typical distance of order

$$d_n \asymp (\log n)^2,$$

and consequently

$$R_n = \frac{d_n}{n} \asymp \frac{(\log n)^2}{n} \longrightarrow 0.$$

The independence assumption is only a first approximation. The Hardy–Littlewood prime-pair framework introduces a singular-series correction depending on the arithmetic structure of $2n$ and the offset k [2, 18]. In heuristic form, the local chance of simultaneous primality can be represented as

$$\mathbb{P}(n-k \text{ and } n+k \text{ prime}) \approx \frac{\mathfrak{S}(n,k)}{(\log n)^2},$$

where $\mathfrak{S}(n,k)$ is an arithmetic correction factor. Provided this factor does not force systematically linear waiting distances, the normalization by n still strongly favors $R_n \to 0$.

Conjecture 6.1 (Relative decay)**.** *For every n for which a symmetric prime pair exists, let $R_n = d_n/n$. Then*

$$\lim_{n\to\infty} R_n = 0.$$

Equivalently, $d_n = o(n)$.

Conjecture 6.2 (Observed extremal bound)**.** *For all $n \geq 2$ for which d_n exists,*

$$R_n \leq \frac{9}{22},$$

with equality at $n = 22$.

Both conjectures are supported by the present finite computation and by the earlier paper [1]; neither is proved here.

7. Discussion

The principal empirical conclusion is not that R_n decreases monotonically. It does not. Rather, three different types of evidence indicate relative decay: the point cloud narrows on logarithmic scales, bin-wise central and upper-tail statistics decrease, and later interval maxima are far below the early global record. These features are mutually reinforcing and are more informative than a single graph.

The normalization $R_n = d_n/n$ is essential. The unnormalized distance d_n may increase along subsequences even while R_n decreases, because division by the growing midpoint changes the scale. Hence a growing absolute gap is compatible with shrinking relative displacement.

Ktrend - Nigerian Journal of Mathematical and Computational Sciences

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

The record result at $2n = 44$ is striking. It shows that every tested even integer above 44 has a nearest symmetric Goldbach pair lying proportionally closer to its midpoint than the pair $(13, 31)$ lies to 22. However, finite stabilization cannot rule out a later exceptional value. A proof of the proposed universal bound would require new control over the location of Goldbach representations.

The results also demonstrate why visual summaries should be paired with mathematical caution. Logarithmic graphics can reveal scale behavior, but no finite plot proves an asymptotic limit. Likewise, a fitted trend line would be model-dependent and could obscure arithmetic oscillations. The current analysis therefore emphasizes exact finite statements, descriptive summaries and clearly labeled heuristics.

8. Limitations and Reproducibility

The upper limit 10^6 is modest compared with the largest computational verification ranges for Goldbach's conjecture [6]. The purpose here is not to compete with those verifications, but to study the complete distribution of a normalized midpoint statistic and to provide interpretable visual evidence.

The logarithmic bins are descriptive and depend on the chosen boundaries. The histogram is truncated at $d_n \leq 150$ to display the most frequent distances clearly. The analysis also focuses only on the nearest pair. The second-, third- and higher-order symmetric pairs may have different scaling behavior and require separate order-statistic analysis.

For reproducibility, the computation should report the programming language, version, sieve implementation, exact upper bound, hardware and output checksum in any archived software supplement. The pseudocode in Algorithm 1 specifies the mathematical procedure independently of programming language.

9. Open Problems and Future Research

The computations motivate the following questions:

1. Can the finite bound $R_n \leq 9/22$ be proved for all n admitting a Goldbach representation?
2. What is the correct average order of d_n? Is it comparable with $(\log n)^2$, or does the singular series produce a different leading scale?
3. Does a suitably normalized version of d_n possess a limiting distribution?
4. How do the second-nearest and third-nearest symmetric distances behave?
5. What are the record values of d_n and R_n beyond 10^6?
6. How do residue classes of n modulo small primorials affect the empirical distribution of d_n?
7. Can rigorous results on almost-all Goldbach representations yield nontrivial bounds for R_n on a density-one subset?

10. Conclusion

This paper upgrades the relative-distance framework of Daniel and Ogunsola [1] into an exhaustive computational and visualization study of all even integers up to 10^6. The mathematical foundations clarify the equivalence between the existence of d_n and a Goldbach representation, establish basic bounds and explain the significance of the running record. The empirical analysis shows that local fluctuations persist, but means, medians, upper quantiles and range-specific maxima decline strongly with scale. The global record remains $R_{22} = 9/22$ at $2n = 44$ throughout the tested interval.

The evidence is consistent with the heuristic scale $d_n \asymp (\log n)^2$ and the conjectural decay $R_n \to 0$. Nevertheless, the conclusions remain finite and computational. Extending the range, studying higher-order symmetric pairs and connecting empirical distributions with Hardy–Littlewood singular-series predictions are natural directions for future work.

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

Funding

The authors received no specific funding for this work.

Conflict of Interest

The authors declare no conflict of interest.

Data and Code Availability

The finite dataset is reproducible from Algorithm 1 using a deterministic sieve of Eratosthenes. The authors may archive the implementation and generated data in a public repository alongside the final publication.

References

[1] I. E. Daniel and O. J. Ogunsola, "Bounded Relative Distances of Symmetric Prime Pairs Around Even Integers," *Ktrend - African Journal of Mathematics, Statistics and Computer Science*, vol. 1, no. 1, 2026, doi:10.5281/zenodo.20748565.

[2] 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 Mathematica*, vol. 44, pp. 1–70, 1923.

[3] 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," *Scientia Sinica*, vol. 16, pp. 157–176, 1973.

[4] H. A. Helfgott, "Major arcs for Goldbach's theorem," arXiv:1305.2897, 2013.

[5] H. A. Helfgott, "The ternary Goldbach conjecture is true," arXiv:1312.7748, revised 2015.

[6] T. Oliveira e Silva, S. Herzog and S. Pardi, "Empirical verification of the even Goldbach conjecture and computation of prime gaps up to 4×10^{18}," *Mathematics of Computation*, vol. 83, no. 288, pp. 2033–2060, 2014.

[7] R. C. Vaughan, *The Hardy–Littlewood Method*, 2nd ed., Cambridge University Press, 1997.

[8] R. C. Vaughan, "Goldbach's conjectures: A historical perspective," in *Open Problems in Mathematics*, Springer, 2016, pp. 479–520.

[9] Y. Wang, *The Goldbach Conjecture*, World Scientific, Singapore, 2002.

[10] D. M. Burton, *Elementary Number Theory*, 7th ed., McGraw-Hill, 2010.

[11] H. Cramér, "On the order of magnitude of the difference between consecutive prime numbers," *Acta Arithmetica*, vol. 2, pp. 23–46, 1936.

[12] A. Granville, "Harald Cramér and the distribution of prime numbers," *Scandinavian Actuarial Journal*, vol. 1995, no. 1, pp. 12–28, 1995.

[13] Y. Zhang, "Bounded gaps between primes," *Annals of Mathematics*, vol. 179, no. 3, pp. 1121–1174, 2014.

[14] J. Maynard, "Small gaps between primes," *Annals of Mathematics*, vol. 181, no. 1, pp. 383–413, 2015.

[15] T. Tao, "Polymath8b: Bounded intervals with many primes, after Maynard," 2014, available online.

ISSN: 3141-643X
DOI: 10.5281/zenodo.21591621
Vol. 1, Issue 2(2026)

[16] J. Hadamard, "Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques," *Bulletin de la Société Mathématique de France*, vol. 24, pp. 199–220, 1896.

[17] C. J. de la Vallée Poussin, "Recherches analytiques sur la théorie des nombres premiers," *Annales de la Société Scientifique de Bruxelles*, vol. 20, pp. 183–256, 1896.

[18] C. Tafula, "An elementary heuristic for Hardy–Littlewood extended Goldbach's conjecture," *São Paulo Journal of Mathematical Sciences*, vol. 14, pp. 391–405, 2020.

[19] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge University Press, 2007.

[20] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer, 1976.

[21] M. B. Nathanson, *Additive Number Theory: The Classical Bases*, Springer, 1996.

[22] P. Ribenboim, *The Little Book of Bigger Primes*, 2nd ed., Springer, 2004.

[23] C. Pomerance, "The prime graph," *Mathematics of Computation*, vol. 33, no. 145, pp. 399–408, 1979.

[24] B. Y. Wang and X. Wang, "Symmetric distribution of primes and their gaps," *Advances in Pure Mathematics*, vol. 12, pp. 447–456, 2022.