Quest Journals
Journal of Research in Applied Mathematics
Volume 12 ~ Issue 11 (November 2026) pp: 01-18
ISSN (Online): 2394-0743 ISSN (Print): 2394-0735
www.questjournals.org

Research Paper

A Computational Analysis of Goldbach Representations: Frequency, Normalization, and Empirical Patterns

Ernesto Cervantes López
[1]*General Directorate of Integration Analysis and Research,*
National Institute of Statistics and Geography,
Aguascalientes, México.

ABSTRACT : *The strong Goldbach conjecture states that every even integer greater than two can be expressed as the sum of two primes. Despite substantial theoretical progress and extensive computational verification, general proof of this conjecture remains unknown. This paper develops a reproducible computational framework for studying the frequency of Goldbach representations and their variation with the magnitude of even integers. We define a representation-counting function for unordered prime pairs and introduce a normalized function based on the principal scale* $N/(\log N)^2$*. The empirical analysis covers 49,998 even integers between 6 and 100,000, using a computational dataset that records the number of representations and the corresponding prime pairs. The total number of recorded representations is 25,366,982. The mean number of representations is approximately 507.36, while the median is 456. For the normalized function, the mean is approximately 1.232 and the median is approximately 1.014. The results illustrate the increasing overall scale of the representation count and the relative fluctuations remaining after normalization. The analysis is interpreted in the context of the Hardy-Littlewood heuristic and previous computational work. The distinction between finite computational verification, empirical regularities, and mathematical proof is maintained throughout. The study does not claim to prove the strong Goldbach conjecture; instead, it provides a transparent framework for examining its numerical structure and motivating further research.*

KEYWORDS: *Goldbach conjecture; prime numbers; additive number theory; computational number theory; Hardy-Littlewood heuristic*

*Received 02 Nov., 2026; **R**evised 11 Nov., 2026; **A**ccepted 13 Nov., 2026 © The author(s) 2026. **P**ublished with open access at **www.questjournals.org***

I. INTRODUCTION

The distribution of prime numbers constitutes a fundamental pillar of number theory. Extending beyond the examination of individual primes, additive number theory explores the representation of integers as sums drawn from sets of arithmetical significance. Among the most celebrated challenges within this domain is the strong Goldbach conjecture, which postulates that every even integer exceeding two can be decomposed into the sum of two primes.

The Goldbach conjecture has a long mathematical history, originating in the correspondence between Christian Goldbach and Leonhard Euler in 1742. Its binary formulation states that every even integer greater than two can be expressed as the sum of two prime numbers. Traditionally associated with this eighteenth-century correspondence, the conjecture remains one of the fundamental open problems in number theory, as no general proof of its binary formulation has yet been established [16].

It is important to distinguish the strong, or binary, Goldbach conjecture from the weak, or ternary, Goldbach conjecture, which states that every odd integer greater than five can be expressed as the sum of three prime numbers. While the latter was successfully proven by Helfgott [8], with the result announced in 2013 and developed in a subsequent 2014 publication, the strong conjecture remains unresolved [7],[16].

Beyond its historical significance, the Goldbach conjecture has stimulated the development of important techniques in analytic number theory, particularly the circle method and sieve methods. The analytical investigation of additive representations establishes a fundamental connection with the distribution of prime numbers. In 1923, Hardy and Littlewood [5] formulated influential conjectural expressions to estimate the number of integers that can be represented as sums of primes. Their approach highlights the interplay between local arithmetic constraints and prime density in determining the number of such representations. A comprehensive mathematical framework for understanding these relationships is provided by standard works in analytic number theory [1],[9],[16].

In addition to theoretical developments, computational verification provides a complementary perspective on the conjecture. For example, Oliveira e Silva [13], reported the computational verification of the binary Goldbach conjecture for every even integer up to 4×10^{18}. Although this achievement demonstrates the considerable range of modern computational methods, finite verification and a general mathematical proof remain logically distinct. Consequently, despite substantial theoretical and computational progress, the binary Goldbach conjecture continues to represent an open problem in mathematics.

The purpose of this paper is to investigate the numerical behavior of Goldbach representation counts over a finite interval. Specifically, we define a counting function for unordered prime pairs, examine its descriptive statistics, introduce a normalized transformation, and evaluate the resulting empirical patterns. The analysis uses a dataset covering even integers from 6 through 100,000.

The paper makes three contributions. First, it formalizes the representation count used in the computational analysis. Second, it documents the distribution of the observed counts and their normalized counterparts. Third, it establishes a reproducible distinction between verified finite results, heuristic interpretations, and questions that remain mathematically unresolved.

The objective is not to establish the truth of the strong Goldbach conjecture for all even integers. Rather, the study provides an empirical framework for analyzing representation frequencies and assessing the extent to which a simple normalization clarifies their numerical variation.

II. MATHEMATICAL FRAMEWORK

Although the fundamental theorem of arithmetic guarantees the unique factorization of every integer greater than one into prime factors, the distribution of prime numbers exhibits irregularities that make it difficult to establish general results concerning their additive relationships. In this context, Goldbach's strong conjecture is a central problem because of its simple formulation and the difficulty of obtaining a general proof.

Let N be an even integer greater than two. The conjecture states that there exist prime numbers p and q such that

$$N = p + q,\ p, q \in \mathbb{P} \qquad (1)$$

where $\mathbb{P}$ denotes the set of prime numbers. This is a universal statement: it must hold for every even integer in the specified domain. Goldbach-type problems constitute a fundamental part of additive number theory, which investigates the representation of integers as sums of elements belonging to specified numerical sets, particularly prime numbers. The classical formulation of Goldbach's conjecture and its relationship with other additive problems provide an important theoretical foundation for understanding the arithmetic structure examined in this study open [6].

To quantify these representations, define the function $G(N)$ as the number of pairs of prime numbers whose sum equals N, considering only pairs satisfying $p \leq q$. This function provides a mathematical framework for examining the number of prime representations of even integers and studying their behavior through computational and analytical approaches.

Consequently, verifying a finite number of cases, however large, does not constitute a mathematical proof of the conjecture. Nevertheless, computational verification provides a valuable complementary approach, as it makes it possible to investigate the structure of prime representations, formulate auxiliary hypotheses, and

test predictions derived from mathematical models. These findings must, however, be interpreted with caution, since computational verification alone does not establish a formal mathematical demonstration.

Formally,

$$G(N) = \#\left\{p \in \mathbb{P}: p \leq \frac{N}{2}, N - p \in \mathbb{P}\right\} \qquad (2)$$

where $\#$ denotes the cardinality of a set. This definition prevents the same representation from being counted twice, since the pairs (p, q) and (q, p) correspond to the same unordered decomposition.

The function defined in (2) transforms the original statement into a counting problem and makes it possible to investigate quantitative properties that are not captured by simply checking whether a pair exists. Relevant questions include the growth of $G(N)$, its variation across different even integers, the frequency of small values, and the relationship between numerical observations and estimates from analytic number theory. These questions must be formulated so that experimental observations are not confused with universally established results.

Recent developments in the analytic study of the binary Goldbach problem have focused not only on the representation of integers as sums of primes but also on estimating the exceptional set of integers for which such representations cannot be established. In particular, Bhowmik and Grimmelt [2] review advances based on the Hardy–Littlewood circle method and derive new explicit results concerning the major arcs, contributing to the understanding of the distribution of Goldbach representations.

Another recent contribution concerns the distribution of Goldbach numbers within short intervals. Assuming the Riemann Hypothesis (RH), Chirre and Hagen [2] prove that every interval $(x,\ x + 123\log^2 x]$ for $x \geq 2$, contains at least one even integer that can be expressed as the sum of two primes. In other words, there exists an even integer n such that $x < n \leq x + 123\log^2 x$ and $n = p_1 + p_2$, where p_1 and p_2 are prime numbers. This result provides an important perspective on the local distribution of Goldbach numbers and complements broader investigations of their global counting behavior.

This perspective is fundamentally based on the following theoretical reference [5], who developed analytical methods for studying the representation of numbers as sums of primes and formulated asymptotic predictions related to such problems. In general terms, their approach suggests that the number of Goldbach representations of large even integers depends on factors associated with the magnitude of the integer and local arithmetic restrictions.

A heuristic approximation of the order of magnitude relates this number to $\frac{N}{(\ln N)^2}$, although a more precise formulation requires an arithmetic singular-series factor. This prediction should not be interpreted as an inequality proved for every even integer. Its value lies in providing a theoretical reference against which data can be compared and research questions developed.

Computational evidence also plays an important role in the study of the problem. The result of Oliveira e Silva [13] provides extensive evidence supporting the conjecture within the verified range, but it does not eliminate the need for mathematical proof valid for all even integers greater than two. Their work also demonstrates that data on the number of Goldbach representations can be used to examine predictions concerning the distribution of prime numbers. Accordingly, an additional computational study should clearly define its research question and potential contribution rather than treating the mere extension of a table of values as an original result.

Building on these precedents, this paper proposes an exploratory study of the function $G(N)$, aimed at describing its behavior over an explicitly defined computational interval and examining its relationship with growth estimates suggested by analytic number theory. The proposed procedure includes generating prime numbers using a sieve, calculating the representations of each even integer, independently verifying the results, and statistically analyzing the representation function. Reproducibility requires specifying the interval studied, the algorithm employed, the counting conventions, and the limitations of the procedure. It is also necessary to distinguish calculated results from interpretations derived from them, since a pattern observed in a finite dataset does not, by itself, establish an asymptotic law.

In particular, the following normalization is proposed

$$H(N) = \frac{G(N)(\ln N)^2}{N} \qquad (3)$$

investigate whether this transformation facilitates comparisons between the numbers of representations corresponding to different even integers. The expression is motivated by the growth scale suggested by heuristic approximations, but it does not presuppose that $H(N)$ converges, remains bounded, or exhibits uniform behavior. Its usefulness will be assessed through the data and comparison with relevant theoretical predictions. Any identified regularity must also be subjected to additional tests and compared with existing results before being considered a novel contribution.

Accordingly, the main objective of this research is to establish a verifiable mathematical and computational foundation for analyzing Goldbach representations, identify patterns that may lead to more specific questions, and determine which observations admit rigorous justification. The scope of the study is deliberately limited. The study's potential contributions lie in developing a reproducible procedure, providing a quantitative description of the behavior of $G(N)$ within the selected domain, and identifying possible avenues for future research based on the actual results obtained.

III. MATERIALS AND METHODS

2.1. Research design

Additive number theory provides a systematic framework for studying the representation of integers as sums of elements from specified numerical sets, particularly prime numbers. Within this framework, Goldbach's conjecture represents a central problem, connecting the distribution of prime numbers, additive representations, sieve methods, and asymptotic analysis. These mathematical foundations are relevant to the formulation and interpretation of the Goldbach counting function examined in this study [12].

This research project adopts a quantitative, computational, and exploratory approach to the study of Goldbach representations. This methodology is appropriate because Goldbach's strong conjecture makes a universal claim about even integers, whereas computational experimentation allows for the systematic investigation of a finite domain and the formulation of questions suitable for subsequent mathematical research. Accordingly, the independent variable is defined as the even integer N, while the principal response variable is the Goldbach counting function $G(N)$, defined as

$$G(N) = \sum_{\substack{p \in \mathbb{P} \\ 3 \le p \le N/2}} \mathbf{1}_{\mathbb{P}}(N-p) \qquad (4)$$

where $\mathbf{1}_{\mathbb{P}}(m)$ denotes the indicator function, which equals 1 when m is prime and 0 otherwise. This function counts the distinct unordered pairs of odd prime numbers whose sum equals N.

The summation is restricted to odd primes because the only even prime is 2, and, for even integers $N > 4$, a representation involving 2 would require the other summand to be an even integer greater than 2, which cannot be prime. The case $N = 4$, whose only representation is $2 + 2$, is therefore treated separately when examining the complete formulation of the conjecture.

The proposed pilot interval includes all even numbers between 6 and 100,000. This interval represents an initial methodological choice rather than a claim that the computational experiment has already been performed. To identify the prime numbers required for the analysis, the study will employ the Sieve of Eratosthenes, a classical and computationally efficient algorithm for generating all primes up to a prescribed upper bound [4].

To construct the list of primes up to a limit M, a sequence of integers greater than or equal to two is initialized, the smallest integer not yet eliminated is selected successively, and its multiples are removed starting from its square. The procedure terminates when the square of the selected prime exceeds M. The integers that remain are precisely the primes not exceeding M.

To calculate the Goldbach representations of each even integer N, it is sufficient to generate primes up to $N/2$. The primality of the complement $N - p$ can be checked using a primality table extending to the largest

required value. This strategy avoids repeating divisibility tests for each candidate pair and facilitates reproducibility.

The correctness of the procedure follows from the fact that every composite integer $m > 1$ has at least one prime divisor less than or equal to $\sqrt{m}$. Consequently, if no prime in that range divides m, then m is prime.

2.2. Procedure for calculating representations

For each even integer N in the selected interval, the algorithm will examine odd primes p satisfying $p \leq N/2$. At each iteration, the complement $q = N - p$ will be calculated and tested for primality. When both conditions are satisfied, the representation (p, q) will be recorded for N, and the counter $G(N)$ will be incremented by one.

The restriction $p \leq N/2$ ensures that each pair is counted only once, since $p + q$ and $q + p$ represent the same unordered decomposition.

The output will consist of a table containing at least the following variables: even integer N, number of representations $G(N)$, first recorded representation, and number of representations per unit of the analyzed interval. To facilitate independent verification, the identified pairs will be retained, and each pair will be checked to ensure that both components are prime and that their sum equals N. The total number of representations will also be checked against the number of recorded pairs, without duplication. These controls will help identify implementation errors and ensure internal consistency.

2.3. Derived variables and descriptive analysis

The arithmetic structure of integers is closely related to the distribution of their prime divisors. In analytic number theory, multiplicative functions and estimates concerning prime divisors provide essential tools for investigating arithmetic properties and asymptotic behavior. These concepts are particularly relevant when examining the arithmetic factors that influence the distribution of prime-related representations [15].

In addition to the Goldbach counting function $G(N)$, the normalized variable

$$H(N) = \frac{G(N)(\ln N)^2}{N} \qquad (5)$$

will be examined, where ln denotes the natural logarithm. This transformation will serve as an exploratory tool for comparing representation counts across different magnitudes of N. Its selection is motivated by heuristic approximations in analytic number theory, particularly the Hardy-Littlewood conjectural framework, which relates the number of representations to a scale of order $\frac{N}{(\ln N)^2}$, modified by additional arithmetic factors [5]. Consequently, $H(N)$ will not be treated as a known theoretical constant, nor will convergence to a particular limit be assumed.

The descriptive statistical analysis will include the mean, median, standard deviation, minimum, and maximum values of $G(N)$, together with the identification of even integers having the fewest representations within the studied domain. The distribution of $H(N)$ and its variation across consecutive subintervals will also be examined to investigate possible patterns and fluctuations in the normalized representation counts.

Furthermore, graphical representations of $G(N)$ against N and $H(N)$ against N will be constructed to identify trends, fluctuations, and possible deviations from the heuristic scale predicted by analytic number theory. Where appropriate, interval-based groupings will be employed to reduce the influence of individual fluctuations and facilitate comparisons across different portions of the domain.

III. RESULTS

3.1. Result verification and reproducibility

Verification of the computational results will be conducted through independent checks of selected even integers, as well as direct inspection of the recorded prime representations. For each test case, every pair must simultaneously satisfy the primality condition and the identity $p + q = N$. Additionally, selected small values of N will be verified manually, since their prime representations can be listed without the use of a computer program. These procedures provide an additional means of checking the accuracy of the computational results [4].

To ensure the reproducibility of the study, the methodological record will include documentation of the algorithm, execution parameters, data format, and instructions required to reproduce the calculations. This documentation will allow the computational procedure and its results to be independently examined and verified.

Finally, computational results will be interpreted strictly as verifications within a finite domain. This distinction is essential because checking a large number of cases does not establish that the Goldbach conjecture holds for every even integer greater than two. The computational work of Oliveira e Silva [13] provides a relevant precedent for understanding both the scope of large-scale computational verifications and the importance of documenting their limits. Therefore, reproducibility will be regarded as a fundamental condition for establishing the methodological reliability of this study.

The interpretation of the results distinguishes between three levels of evidence

- Properties established through mathematical arguments (see annexes).
- Exact computational results obtained within the selected interval.
- Empirical regularities suggested by statistical analysis.

Identifying a trend in $G(N)$ or $H(N)$ will not be considered an original contribution in itself. To determine whether a result is novel, observations and potential propositions must be compared with specialized literature, particularly analytical work on additive representations of prime numbers. The objective of this stage is therefore to produce a verifiable dataset and characterize the behavior of the representation function within the selected interval.

3.2. General characterization of Goldbach representations

The computational experiments involved studying how often even integers can be expressed as the sum of two prime numbers and analyzing this behavior using a representation function and a normalized transformation. The set considered was $\mathcal{N} = \{N \in 2\mathbb{Z}: 6 \leq N \leq 100{,}000\}$ which contains 49,998 even integers. For each $N \in \mathcal{N}$, the number of unordered representations was calculated as $G(N) = \#\{(p, q) \in \mathbb{P}^2 : p \leq q,\ p + q = N\}$.

The condition $p \leq q$ prevents the same representation from being counted twice by exchanging its summands.

The summary file records a total of 25,366,982 representations. The descriptive statistics of the function $G(N)$ are presented in the next Table.

Statistic	Result
Even integers analyzed	49,998
Total representations	25,366,982
Mean of $G(N)$	507.359934

Median of $G(N)$	456
Population standard deviation	332.595719
Minimum of $G(N)$	1
Maximum of $G(N)$	2,168
Mean of $H(N)$	1.23159341
Median of $H(N)$	1.01437891
Population standard deviation of $H(N)$	0.47734977

Table 1. Descriptive statistics of the Goldbach representation function

The statistics were calculated using the full set of even integers within the studied interval. The standard deviation was calculated using a divisor of 49,998. The $H(N)$ values were obtained using the specified normalization.

The mean of $G(N)$ exceeds its median by approximately 51.36 representations. Along with a standard deviation of 332.60, this difference indicates significant variation in the number of prime decompositions. Consequently, the number of representations cannot be adequately described by a single average value, and variation across the interval must also be examined.

The maximum observed value was $G(99{,}330) = 2{,}168$, this local maximum should not be confused with the value at the upper endpoint of the interval; for $N = 100{,}000$, the file records $G(N) = 810$. These differences demonstrate that $G(N)$ exhibits substantial local fluctuations and is not strictly increasing.

3.3. Behavior of the function $G(N)$

Recent research has continued to investigate the analytical structure of Goldbach's counting function, particularly the behavior of secondary terms in its asymptotic formulas. Matsumoto and Suzuki examine the limiting distributions of these secondary terms through the introduction of M-functions and establish connections with the Riemann zeta-function [11]. These developments provide additional theoretical context for interpreting the fluctuations observed in computational studies of Goldbach representations.

In our situation, the first graph presents the Goldbach counting function $G(N)$ as a function of the even integer N. It illustrates a general increase in the number of representations as N becomes larger, accompanied by noticeable fluctuations between consecutive even integers. This overall pattern is consistent with the expectation that larger even integers tend to admit more representations as sums of two primes.

However, the observed behavior is not strictly monotonic. Individual values exhibit local maxima and considerable declines, indicating that the number of representations does not increase uniformly with the magnitude of N. Therefore, the graph should not be interpreted as evidence of a simple deterministic relationship between $G(N)$ and N.

The observed dispersion is particularly relevant because it demonstrates that two even integers of similar magnitude may have substantially different numbers of prime representations. Consequently, the analysis must distinguish between the overall growth of the counting function and its local fluctuations. The latter may reflect differences in the arithmetic properties of the integers involved, as well as the influence of the factors governing the distribution of prime representations.

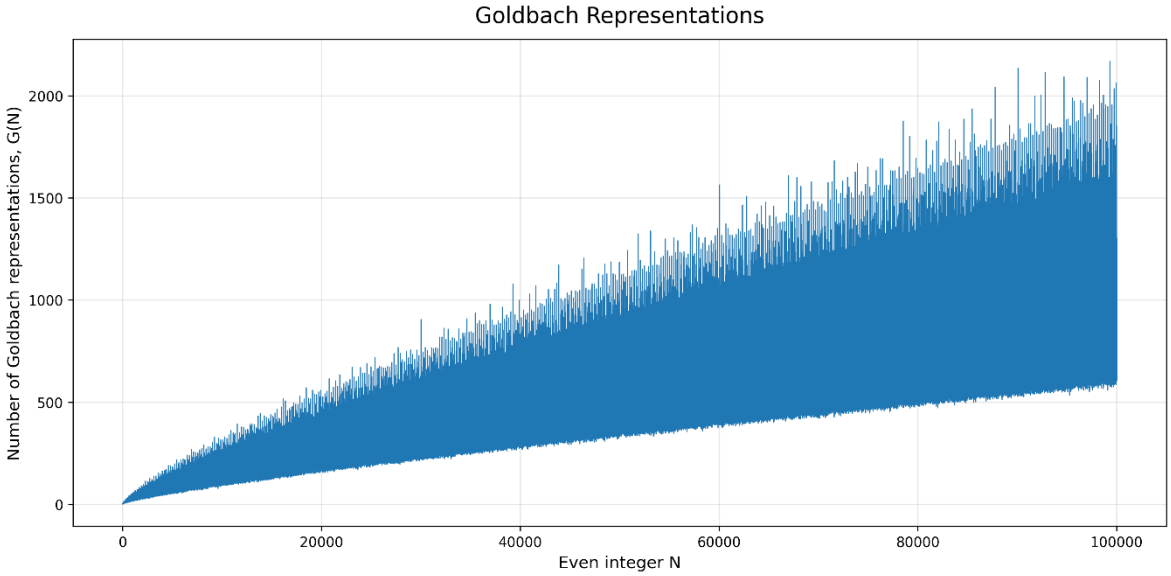


Figure 1. Number of unordered Goldbach representations, $G(N)$, for even integers N from integers N from 6 to $100{,}000$**.** Each value counts the distinct prime pairs $p \leq q$ satisfying $p + q = N$. The overall increase reflects the growing number of Goldbach representations, while the fluctuations reveal substantial variation across even integers.

The graph displays the number of unordered representations of each even integer as the sum of two prime numbers. A general increasing trend is observed over the analyzed interval, accompanied by substantial oscillations. The behavior is not strictly monotonic: the maximum occurs at $N = 99{,}330$, whereas the even integer $N = 100{,}000$ has 810 representations.

3.4. Analysis of the normalized function $H(N)$

To study the number of representations relative to a reference growth scale, the function was defined as (5) and that we can manipulate and that remains

$$H(N) = \frac{G(N)}{N/(\ln N)^2}.$$

This transformation makes it possible to compare the observed number of representations with the scale $N/(\ln N)^2$, associated with the heuristic approximation of Hardy and Littlewood [5].

The mean of the normalized function was 1.23159341, its median was 1.01437891, and its population standard deviation was 0.47734977. The difference between the mean and median indicates that relatively high values influence the average, although the dispersion of $H(N)$ is numerically smaller than that observed for $G(N)$. This comparison must be interpreted cautiously because the two functions have different scales and meanings.

The second graph shows that normalization makes it possible to observe variations in $G(N)$ relative to the selected scale without allowing absolute growth to dominate the visual representation to the same extent. Local oscillations and extreme values nevertheless persist. The transformation therefore facilitates exploratory analysis of fluctuations but does not eliminate them.

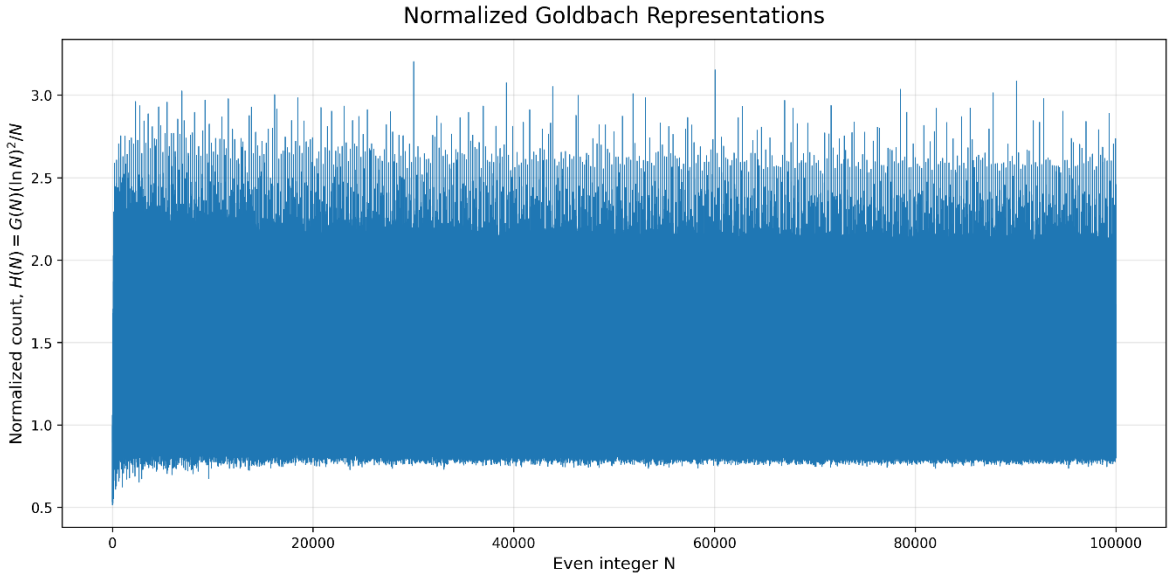


Figure 2. Normalized Goldbach representation count, $H(N) = G(N)\frac{(\ln N)^2}{N}$, for even integers N from 6 to $100{,}000$**.** The normalization accounts for the leading-order growth rate suggested by prime-density heuristics and facilitates the examination of fluctuations in the representation count. The remaining variability illustrates that normalization does not eliminate arithmetic fluctuations[1].

The consistency of the results was examined using two complementary files. The first file contains N, $G(N)$ and $H(N)$, as well as the verification indicator. The second file contains the pairs of prime numbers used to construct the representations.

The sum of the individual counts is agreed with the reported total:

$$\sum_{N \in \mathcal{N}} G(N) = 25{,}366{,}982.$$

The pair file also contains 25,366,982 records, and the counts for each even integer reproduce the corresponding values of $G(N)$ in the summary. The consistency review further confirms that the recorded pairs satisfy $p + q = N$, that their components are prime, and that the convention $p \leq q$ is used to prevent duplication through the exchange of summands.

These checks provide evidence of internal consistency between the files and the computational counts. However, agreement between the number of rows and the sum of $G(N)$ does not, by itself, prove exhaustiveness. Establishing that all representations have been found requires the search procedure to examine systematically every admissible candidate.

The verification indicator is true for all 49,998 even integers analyzed. This result confirms that the program classified every record as verified; interpreting it as exhaustive verification depends on the procedure used to generate that indicator.

3.5. Cases with the fewest representations

We can see that the first values of the function illustrate the local behavior of the count. For $N = 6{,}8{,}10{,}12{,}14{,}16$, the recorded values are, respectively,

[1] Normalization does not imply that $H(N)$ must converge to a constant. The complete Hardy-Littlewood approximation incorporates an arithmetic factor that depends on N; therefore, a rigorous theoretical comparison must account for that factor.

$G(N) = (1, 1, 2, 1, 2, 2)$.

The corresponding representations are

$6 = 3 + 3,$
$8 = 3 + 5,$
$10 = 3 + 7 = 5 + 5,$
$12 = 5 + 7,$
$14 = 3 + 11 = 7 + 7,$
$16 = 3 + 13 = 5 + 11.$

These examples show that the representation function does not increase at every increment of two. They also demonstrate that the counting convention includes central representations $N = 2p$ when $N/2$ is prime and counts each unordered pair exactly once. The results are consistent with Goldbach's strong conjecture over the studied interval, in the sense that every even integer analyzed has at least one recorded representation. The mathematical conclusion nevertheless depends on the exhaustiveness of the enumeration and the validity of the primality tests used.

IV. Discussion

Recent developments in sieve theory have provided new explicit and asymptotic results relevant to classical problems in additive number theory, including Goldbach-type conjectures. These methods offer complementary approaches to studying representations involving prime numbers and establishing explicit bounds for related arithmetic problems [10].

4.1. Interpretation of the Growth in Representations

The results indicate that the Goldbach counting function $G(N)$ exhibits an overall increasing trend within the studied interval, although its values fluctuate considerably. This observation is consistent with the perspective of analytic number theory, in which the frequency of Goldbach representations is closely related to the distribution of prime numbers. Nevertheless, the observed trend should not be interpreted as a uniform growth law or as proof of the Goldbach conjecture.

To examine this behavior further, the normalized function $H(N)$, defined previously, allows the representation counts to be compared relative to the heuristic scale predicted by analytic number theory. However, this normalization does not, by itself, constitute a complete theoretical prediction or establish a particular asymptotic relationship.

Therefore, the principal contribution of this analysis is not to establish that $G(N)$ follows a specific asymptotic law, but rather to demonstrate how individual representation counts, descriptive statistical measures, and normalization techniques can be combined to characterize the behavior and fluctuations of Goldbach representations within a finite computational domain.

4.2. Mathematical significance of the normalization

The interpretation of (5) must be distinguished between two questions. The first is whether normalization produces a visually more stable series than the original function. The graphical results suggest that it does, in the sense that variations are observed on a different scale. The second is whether $H(N)$ converges, remains bounded, or follows a limiting distribution. This second question cannot be resolved through visual inspection or the mean of a finite interval.

A rigorous comparison with Hardy-Littlewood would require defining the complete theoretical arithmetic factor $C(N)$, using the same convention for counting representations, and comparing $H(N)$ with the resulting prediction. One possible measure would be to examine the residuals $E(N) = H(N) - C(N)$ provided that $C(N)$ is defined consistently with the asymptotic formula and the adopted counting convention. It would also be useful to examine the mean, dispersion, and extreme values of these residuals across successive

subintervals. This extension would help distinguish fluctuations associated with the arithmetic factor from discrepancies between the data and the heuristic approximation.

4.3. Computational consistency and verification of the conjecture

Computational and analytical investigations of Goldbach-type problems are closely connected to the broader development of additive number theory. In particular, the Hardy-Littlewood circle method and Vinogradov's theorem provide essential theoretical tools for studying representations of integers as sums of prime numbers. These methods also establish important connections between the ternary Goldbach problem and the broader analytical framework underlying additive prime number theory [14].

To ensure the reliability and reproducibility of the computational experiment, correspondence between the summary file and the file containing the individual prime pairs verified. Internal consistency assessed by comparing the sum of the values in the summary file with the total number of rows in the pair file, It was also verified whether the counts for the individual representations could be reproduced from the recorded data.

Further mathematical validation is provided by the theorems presented in the appendices. According to the symmetry theorem, if $r(N)$ denotes the number of unordered representations and $R(N)$ denotes the number of ordered representations, then $R(N) = 2r(N) - \varepsilon(N)$ where $\varepsilon(N) = 1$ when $\frac{N}{2}$ is prime and 0 otherwise, provided that $N > 4$. Additionally, the theorem concerning the parity of the summands establishes that, for every even integer $N > 4$, each Goldbach representation consists of two odd prime numbers. These properties can be used to validate the algorithm's counting logic and reduce the search space.

However, these mathematical properties must be distinguished from a proof of computational exhaustiveness. A program may satisfy the expected counting identities and still omit valid prime pairs if its search procedure does not examine every possible candidate. Therefore, verifying internal consistency and establishing that the algorithm has exhaustively examined the relevant search space are two separate requirements.

Finally, the computational work of Oliveira e Silva [13] demonstrates the importance of combining efficient algorithms, independent computational checks, and reproducible documentation when investigating Goldbach-type problems. Nevertheless, the interval examined in the present study is considerably smaller, and its findings must therefore be interpreted within the limits of the selected computational domain.

4.4. Comparison with the Hardy-Littlewood Prediction

A more informative assessment of the empirical representation counts requires comparison with a heuristic that accounts for the arithmetic structure of each even integer. Hardy and Littlewood [5] developed a conjectural asymptotic framework for representations of integers as sums of primes. In this framework, the expected number of representations depends on both a principal logarithmic scale and an arithmetic factor determined by the prime divisors of the integer.

Let C_2 denote the twin-prime constant

$$C_2 = \prod_{\substack{p \text{ prime} \\ p>2}} \left(1 - \frac{1}{(p-1)^2}\right) \approx 0.6601618158. \qquad (6)$$

For an even integer N, define the singular-series factor

$$\mathfrak{S}(N) = \prod_{\substack{p \mid N \\ p>2}} \frac{p-1}{p-2} \qquad (7)$$

where the product is taken over the distinct odd prime divisors of N. An empty product is defined to equal one.

Under the Hardy-Littlewood heuristic, the leading-order prediction for the number of unordered representations is

$$\hat{G}_{\mathrm{HL}}(N) = C_2 \mathfrak{S}(N) \frac{N}{(\log N)^2} \qquad (8)$$

This expression corresponds to the convention used in the present study, in which each pair is counted once under the restriction $p \le q$. The corresponding leading-order prediction for ordered representations has twice the main term, apart from the negligible diagonal correction associated with $p = q = \frac{N}{2}$.

For an additional finite-range comparison, define the logarithmic-density integral

$$I(N) = \int_2^{N-2} \frac{dx}{\log x \, \log(N-x)} \qquad (9)$$

The resulting integral-based prediction is

$$\tilde{G}_{\mathrm{HL}}(N) = C_2 \mathfrak{S}(N) I(N). \qquad (10)$$

Equation (10) is a heuristic refinement of the principal density approximation. It should not be interpreted as an exact identity or as a proven finite-N version of the Hardy-Littlewood conjecture.

Both predictions were evaluated for the 49,998 even integers in the dataset. The leading-term prediction produced a mean predicted count of 415.240, compared with an observed mean of 507.360. Its mean absolute error was 92.120, its root mean squared error was 109.086, and its mean absolute percentage error was 23.25 %.

The integral-based prediction produced a mean predicted count of 513.361. Its mean absolute error was 9.332, its root mean squared error was 12.207, and its mean absolute percentage error was 2.5 %. The correlation between the observed and predicted counts was 0.999517, while the mean observed-to-predicted ratio was 0.983489.

These results indicate that the integral-based expression provides a closer numerical approximation to the observed counts over the interval examined. The improvement is consistent with the possibility that retaining the variation in the logarithmic density across the summation range offers a useful finite-range approximation. However, the comparison does not establish that the integral-based prediction is universally superior or that its empirical error remains stable beyond the analyzed interval.

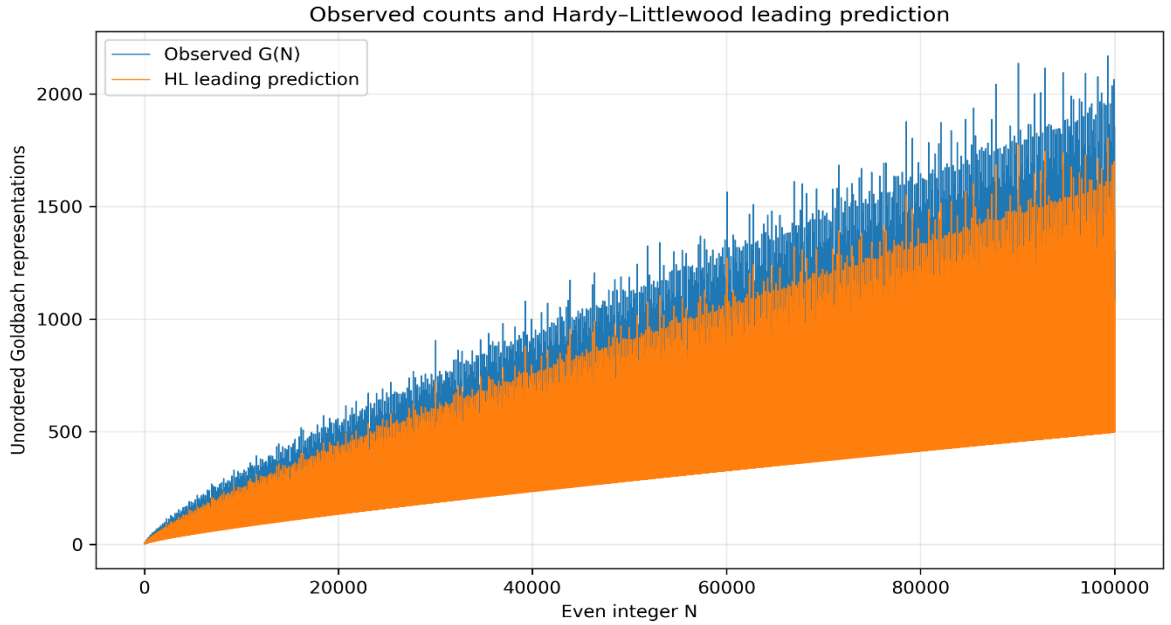


Figure 3. Observed Goldbach representations and the leading Hardy-Littlewood prediction**.** The figure compares the observed number of unordered Goldbach representations, $G(N)$, with the leading Hardy-Littlewood approximation

$$\hat{G}_{\mathrm{HL}}(N) = C_2 \mathfrak{S}(N) \frac{N}{(\ln N)^2}$$

for even integers N in the range $6 \le N \le 100{,}000$. Here, C_2 denotes the twin-prime constant and $\mathfrak{S}(N)$ is the arithmetic singular series accounting for the contribution of the distinct odd prime divisors of N.

The comparison illustrates the model's ability to capture the broad growth pattern of Goldbach representations, while the systematic differences between observed and predicted values indicate that this leading-order approximation does not fully reproduce the finite-range behavior of the data.

The next figure compares the empirical counting function $G(N)$ with the integral-based approximation

$\tilde{G}_{\mathrm{HL}}(N) = C_2 \mathfrak{S}(N) I(N)$

where

$$I(N) = \int_2^{N-2} \frac{1}{\ln x \, \ln(N-x)} \, dx .$$

The integral incorporates the variation in the estimated density of prime numbers across the admissible summands, rather than relying exclusively on the leading-order expression $N/(\ln N)^2$.

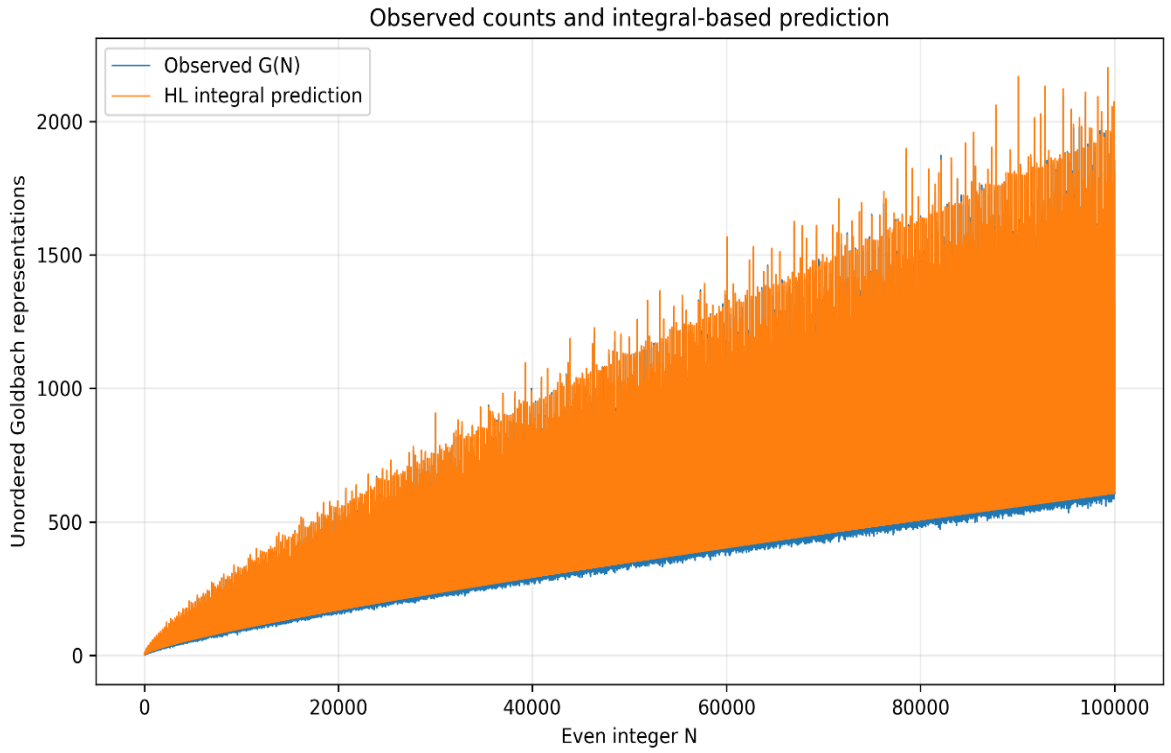


Figure 4. Observed Goldbach representations and the integral refinement of the Hardy-Littlewood prediction.

Over the examined range, the integral refinement provides a closer empirical fit to the observed counts. This improvement should be interpreted as a finite-range numerical finding, not as a proof of the Hardy-Littlewood prediction or of Goldbach's conjecture.

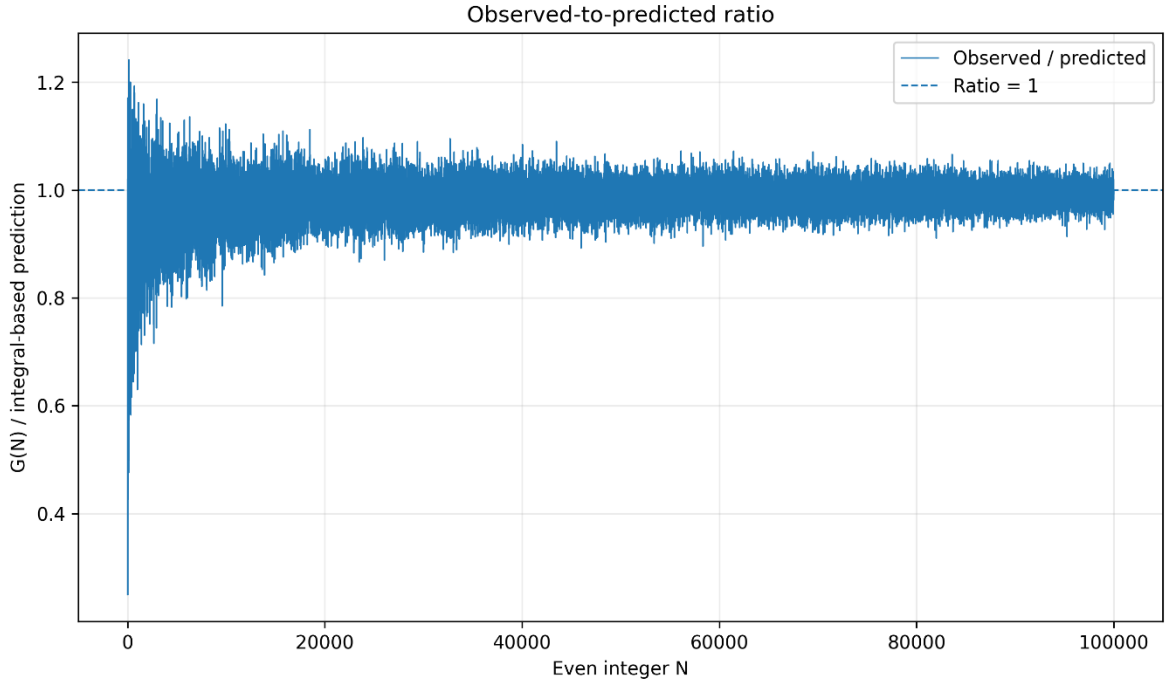


Figure 5. Ratio of observed to predicted Goldbach representations.

The figure displays the ratio

$$\frac{G(N)}{\tilde{G}_{\mathrm{HL}}(N)}$$

for even integers N between 6 and $100{,}000$, where $\tilde{G}_{\mathrm{HL}}(N)$ is the integral-based Hardy-Littlewood approximation. A ratio equal to one indicates exact agreement between the observed count and the model prediction; values above one indicate underprediction, whereas values below one indicates overprediction. The distribution of this ratio provides a scale-independent assessment of the approximation's accuracy and reveals the remaining fluctuations around the predicted level. These deviations are consistent with the irregular behavior of prime representations at finite values of N and should not, by themselves, be interpreted as evidence for or against the asymptotic conjecture

Together, these figures distinguish the broad growth pattern from the remaining local deviations. The numerical agreement should be interpreted cautiously. The arithmetic factor accounts for the prime divisors of N, but the finite-range comparison alone cannot establish an asymptotic theorem. A more extensive assessment would examine errors across subintervals, investigate the largest residuals, and evaluate the stability of the comparison over larger ranges.

V. CONCLUSIONS

This study developed a computational procedure for examining representations of even integers as sums of two prime numbers and the results show an overall increasing trend in the number of representations, accompanied by considerable local fluctuations. The difference between the mean and median, together with the observed dispersion, confirms that the behavior of $G(N)$ cannot be described by a strictly increasing function or a single representative value.

Remember that the normalized function $H(N) = \frac{G(N)(\ln N)^2}{N}$ made it possible to study the counts relative to the scale $N/(\ln N)^2$. These statistics, together with the normalized graph, provide a complementary

description of the fluctuations. However, they do not establish convergence or independently identify an asymptotic constant.

The correspondence between the summary and pair files provides evidence of internal consistency in the counting procedure. Moreover, the mathematical properties of symmetry and parity formulated in the appendices provide validation checks for the computational procedure.

Overall, the experiment provides a reproducible methodology for the quantitative study of Goldbach representations over a finite interval. By combining the enumeration of prime number pairs, statistical analysis and mathematical normalization, while clearly distinguishing between proven identities and observed regularities, it makes a valuable contribution.

Future research should compare $H(N)$ with the Hardy-Littlewood complete arithmetic function, investigate fluctuations in subintervals, evaluate the sensitivity of the results to the selected range and formally document the guarantees of exhaustiveness provided by the algorithm.

It should be noted that the computational verification carried out here is limited to even integers within the selected interval. Consequently, the results do not constitute a proof of the strong Goldbach conjecture. Nevertheless, this study provides an empirical basis for formulating new questions about the distribution of representations and their relationship to predictions from analytic number theory.

Appendix A. Symmetry and Cardinality of Goldbach Representations

A.1. Definitions and notation

Let n be an even integer and let $\mathbb{P}$ denote the set of prime numbers. The set of unordered Goldbach representations of n is defined as $G_u(n) = \{\{p,q\}: pq \in \mathbb{P}, p+q=n\}$.

The corresponding unordered representation function is defined by $g_u(n) = | G_u(n) |$, where $| G_u(n) |$ denotes the cardinality of the set. Thus, each pair of prime numbers is counted exactly once, independently of the order of its components.

Similarly, the set of ordered Goldbach representations is defined as $G_o(n) = \{(p,q) \in \mathbb{P}^2: p+q=n\}$, with corresponding counting function $g_o(n) = | G_o(n) |$. In this case, the ordered pairs (p,q) and (q,p) are considered distinct whenever $p \neq q$.

Finally, we introduce the indicator function of the existence of a central representation, $\delta(n) = \begin{cases} 1 & \text{if } n/2 \in \mathbb{P} \\ 0 & \text{otherwise} \end{cases}$

This function indicates whether n admits a central Goldbach representation, that is, a representation in which both prime summands are equal.

A.2. Symmetry theorem for representations

Theorem A.1. For every even integer $n \geq 4$, the ordered and unordered Goldbach representation functions satisfy the identity $g_o(n) = 2g_u(n) - \delta(n)$.

Proof. Consider the mapping $\varphi: G_o(n) \longrightarrow G_u(n)$ defined by $\varphi\big((p,q)\big) = \{p,q\}$. This mapping associates each ordered representation with its corresponding unordered representation. It is surjective because every unordered representation has at least one corresponding ordered representation.

For any element $\{p,q\} \in G_u(n)$, there are two possible cases. If $p \neq q$, the unordered representation corresponds to exactly two ordered representations, namely (p,q) and (q,p). Consequently, the fiber of φ over this representation has cardinality two.

On the other hand, if $p = q$, the corresponding unordered representation has only one ordered representation, since exchanging its components does not produce a distinct ordered pair. In this case, the fiber has cardinality one.

A central representation exists if and only if $n/2$ is prime. Therefore, the number of unordered representations with distinct prime summands is $g_u(n) - \delta(n)$, while the number of central representations is $\delta(n)$.

Since the fibers of φ partition the set $G_o(n)$, summing their cardinalities yields $g_o(n) = 2\big(g_u(n) - \delta(n)\big) + \delta(n) = 2g_u(n) - \delta(n)$. Therefore, $g_o(n) = 2g_u(n) - \delta(n)$ which proves the theorem.

Corollary A.1.1. Parity of the ordered representation function. For every even integer $n \geq 4$, $g_o(n) \equiv \delta(n) \pmod 2$. In particular, the ordered representation function is odd if and only if $n/2$ is prime.

Proof. The result follows directly from Theorem A.1. Since $g_o(n) = 2g_u(n) - \delta(n)$, and $2g_u(n)$ is even, reduction modulo two gives $g_o(n) \equiv -\delta(n) \equiv \delta(n) \pmod 2$. This proves the corollary.

A.3. Computational consequence

Theorem A.1 establishes a necessary consistency condition for validating two independent computational procedures that enumerate ordered and unordered Goldbach representations. For every even integer $n \geq 4$, the computational results must satisfy $g_o(n) = 2g_u(n) - \delta(n)$ where $g_o(n)$ and $g_u(n)$ denote the values obtained through the respective enumeration procedures.

This identity provides a direct means of detecting inconsistencies between the two counts. Nevertheless, satisfying this condition does not, by itself, establish that either enumeration is exhaustive, since both procedures could potentially omit the same admissible representations. Consequently, this consistency check must be complemented by explicit primality verification, verification of the defining condition $p + q = n$, and an exhaustive search procedure that guarantees that no admissible pairs have been omitted.

It is important to emphasize that Theorem A.1 is a combinatorial identity concerning the symmetry and cardinality of Goldbach representations. It does not constitute proof of Goldbach's strong conjecture, nor does it establish that $g_u(n) \geq 1$ for every even integer $n \geq 4$.

Appendix B. Parity Restrictions on Goldbach Summands

B.1. Fundamental lemma on the parity of prime numbers

Lemma B.1. The only even prime number is 2. Consequently, every prime number $p \neq 2$ is odd.

Proof. Every even integer greater than two can be expressed in the form $m = 2k$, where $k \geq 2$. Therefore, such an integer has a divisor other than one and itself and cannot be prime. Since 2 is prime, it follows that it is the only even prime number. Thus, every other prime number must be odd.

B.2. Theorem on the parity of the summands

Theorem B.1. Let $n \geq 4$ be an even integer. If $n$admits a representation as the sum of two prime numbers, $n = p + q$, then both summands are odd. Equivalently, $p \equiv q \equiv 1 \pmod 2$.

Proof. Suppose that n is even and that there exist prime numbers p and q satisfying $n = p + q$. By Lemma B.1, each summand must either be equal to 2 or be an odd prime number. Suppose, for the sake of contradiction, that one of the summands is equal to 2. Without loss of generality, let $p = 2$. The representation then becomes $q = n - 2$.

Since n is even, $n - 2$ is also even. Moreover, because $n \geq 4$, we have $n - 2 \geq 2$. If $n > 4$, this implies that q is an even integer greater than two, contradicting Lemma B.1, since q is assumed to be prime.

For $n = 4$, the representation is precisely $4 = 2 + 2$, which is an exceptional case. Therefore, the statement of the theorem requires the additional restriction $n > 4$ if both summands are to be odd.

Hence, for every even integer $n > 4$ admitting a Goldbach representation, neither summand can equal 2. Consequently, both summands must be odd primes, and therefore $p \equiv q \equiv 1 \pmod 2$. This proves the theorem for every even integer $n > 4$.

B.3. Corollary: reduction of the search space

Corollary B.1. For every even integer $n > 4$, the set of unordered Goldbach representations satisfies $G_u(n) = \{\{p, q\}: p, q \in \mathbb{P}, pq \geq 3, p + q = n\}$. Consequently, Goldbach representations of even integers greater than four can be searched for exclusively among odd prime numbers.

Proof. By Theorem B.1, every prime summand in a Goldbach representation of an even integer $n > 4$ must be odd. Since the smallest odd prime is 3, every representation must satisfy $pq \geq 3$. Furthermore, the condition $p \leq q$ can be imposed when enumerating unordered representations, ensuring that each admissible pair is counted exactly once.

Therefore, it is sufficient to examine odd primes $p$satisfying $3 \leq p \leq n/2$ and verify whether the complementary integer $n - p$ is also prime. This proves the corollary.

B.4. Implications for the computational algorithm

Theorem B.1 provides a direct reduction of the search space when enumerating Goldbach representations. For every even integer $n > 4$, it is sufficient to consider odd primes $p$satisfying $3 \leq p \leq n/2$ and determine whether the complementary integer $q = n - p$ is also prime.

The resulting enumeration procedure can be expressed formally as $g_u(n) = \sum_{\substack{3 \leq p \leq n/2 \\ p \in \mathbb{P}}} 1_{\mathbb{P}}(n - p),\ n > 4$ where $1_{\mathbb{P}}(m)$ denotes the primality indicator function, defined by $1_{\mathbb{P}}(m) = \begin{cases} 1 & m \in \mathbb{P} \\ 0 & m \notin \mathbb{P} \end{cases}$

This expression counts every unordered Goldbach representation exactly once because the restriction $p \leq n/2$ prevents the same pair from being counted in both possible orders. In particular, when $n/2$ is prime, the central representation is included exactly once.

The formula requires special treatment of the smallest even integers. For $n = 4$, the only Goldbach representation is $4 = 2 + 2$, which is excluded by the restriction to odd primes. For $n = 2$, no representation as the sum of two prime numbers exists, since the smallest possible sum of two primes is $2 + 2 = 4$. Thus, the computational formula above is valid for even integers $n > 4$, while these exceptional cases must be handled separately.

Finally, the parity restriction established in this appendix provides a mathematically justified reduction of the computational search space. However, it constitutes a necessary condition on the prime summands rather than a sufficient condition for the existence of a Goldbach representation. Consequently, neither Theorem B.1 nor Corollary B.1 proves Goldbach's strong conjecture.

REFERENCES

[1]. Apostol, T. M. (1976). *Introduction to analytic number theory*. Springer.

[2]. Bhowmik, G., & Grimmelt, L. (2026). The exceptional set of the Goldbach problem. *Analysis Mathematica*. https://doi.org/10.1007/s10476-026-00191-z

[3]. Chirre, A., & Hagen, M. V. (2026). On Goldbach numbers in short intervals. *Research in Number Theory, 12*, Article 82. https://doi.org/10.1007/s40993-026-00767-6

[4]. Crandall, R., & Pomerance, C. (2005). *Prime numbers: A computational perspective* (2nd ed.). Springer. https://doi.org/10.1007/0-387-28979-8

[5]. Hardy, G. H., & Littlewood, J. E. (1923). Some problems of Partitio numerorum; III: On the expression of a number as a sum of primes. *Acta Mathematica, 44*, 1–70. https://doi.org/10.1007/BF02403921

[6]. Hardy, G. H., & Wright, E. M. (2008). *An introduction to the theory of numbers* (6th ed.). Oxford University Press. https://doi.org/10.1093/oso/9780199219858.001.0001

[7]. Helfgott, H. A. (2014). The ternary Goldbach problem. *Snapshots of Modern Mathematics from Oberwolfach, 2014*(3). https://doi.org/10.14760/snap-2014-003-en

[8]. Helfgott, H. A. (2013). The ternary Goldbach conjecture is true. *arXiv*. https://arxiv.org/abs/1312.7748

[9]. Iwaniec, H., & Kowalski, E. (2004). *Analytic number theory*. American Mathematical Society.

[10]. Johnston, D. R. (2026). New explicit and asymptotic results in sieve theory. *Bulletin of the Australian Mathematical Society, 113*(2), 372–374. https://doi.org/10.1017/S0004972725100828

[11]. Matsumoto, K., & Suzuki, M. (2026). M-functions and screw functions: Applications to Goldbach's problem and zeros of the Riemann zeta-function. *Journal of Number Theory, 280*, 918–946. https://doi.org/10.1016/j.jnt.2025.09.013

[12]. Nathanson, M. B. (1996). *Additive number theory: The classical bases*. Springer. https://doi.org/10.1007/978-1-4757-3845-2
[13]. Oliveira e Silva, T., Herzog, S., & Pardi, S. (2014). Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$. *Mathematics of Computation, 83*(288), 2033–2060. https://doi.org/10.1090/S0025-5718-2013-02787-1
[14]. Rassias, M. Th. (Ed.). (2017). *Goldbach's problem: Selected topics*. Springer.
[15]. Tenenbaum, G. (2015). *Introduction to analytic and probabilistic number theory* (3rd ed., P. Ion, Trans.). American Mathematical Society.
[16]. Vaughan, R. C. (2016). Goldbach's conjectures: A historical perspective. In J. Nash Jr. & M. Th. Rassias (Eds.), *Open problems in mathematics* (pp. 479–520). Springer.