Mathematics and Computer Science
2024, Vol. 9, No. 5, pp. 96-113
https://doi.org/10.11648/j.mcs.20240905.12

Research Article

On the Binary Goldbach Conjecture: Analysis and Alternate Formulations Using Projection, Optimization, Hybrid Factorization, Prime Symmetry and Analytic Approximation

Ioannis Papadakis*

Independent Researcher, Charlotte, United States of America

Abstract

An analysis, based on different mathematical approaches, of the binary Goldbach conjecture −which states that every even integer $s \geq 6$ is the sum of two odd primes, called Goldbach primes− is presented. Each approach leads to a different reformulation of this conjecture, thus contributing unique insights into the structure, properties and distribution of prime numbers. The above-mentioned reformulations are based on the following distinct, interrelated and complementary approaches: projection, optimization, hybrid prime factorization, prime symmetry and analytic approximation. Additionally, it is shown that prime factorization is an optimal projection operation on the set of integers; that Goldbach pairs correspond to solutions of an optimization problem; that hybrid prime factorization can be used to generate Goldbach primes; that prime symmetry, a powerful property of Goldbach primes, can be used to validate the binary Goldbach conjecture in short intervals, and to determine the rules that govern the "algebraic evolution" of Goldbach pairs, as the value of s increases; and that analytic approximation, using translational and rotational shifts of smooth functions, leads to a useful approximation of a primality test function and the prime counting function $\pi(s)$. The paper's findings support the broader hypothesis that prime numbers, by virtue of their optimality in representing, additively and multiplicatively, any measurable quantity in the universe, supported by the Fundamental Theorem of Arithmetic and the binary Goldbach conjecture, may be a viable alternative to the exclusive use of binary logic, as a means of achieving additional computational efficiencies of scale in the future.

Keywords

Primes, Goldbach Conjecture, Projection, Optimization, Factorization, Prime Symmetry, Analytic Approximation

1. Introduction

The binary Goldbach conjecture (bGC), also called even or strong Goldbach conjecture, was originally proposed by mathematician Christian Goldbach in 1742. It states that every even integer ≥ 6 can be represented as sum of two primes. The bGC implies the so-called tertiary, or weak, Goldbach conjecture (tGC), which states that every odd number larger than 5 can be written as a sum of three primes. In the 280 years since the conjecture was proposed, it has inspired a large body of mathematical work, including significant inroads in proving and validating it. The basic concepts and a review of advances on bGC/tGC, with extensive bibliography, can be found in [1-4].

*Corresponding author: ypap@primactaglobal.com (Ioannis Papadakis)

Received: 27 October 2024; **Accepted:** 12 November 2024; **Published:** 29 November 2024

The work on bGC/tGC has generally proceeded on two tracks: numerical validation and mathematical proof. In 2013, the bGC has been computationally validated up to $4\cdot 10^{18}$, and in 1975, H. Montgomery and R. C. Vaughan improved the results of earlier research by proving that the set of even integers that are not the sum of two primes has density zero [5, 6].

More progress has been made in proving the tGC. In 2013, H. A. Helfgott submitted proof of the tGC, and it was accepted in 2015 for publication in the Annals of Mathematics Studies [7]. The manuscript is close to being finalized for publication, with some sections available online [8].

Initially, some of the conditional proofs of the tGC, such as the proof by Hardy and Littlewood (1923), were contingent on the generalized Riemann hypothesis; in this case, this dependence was later eliminated by M. Vinogradov. Researchers continue to publish proofs of the bGC, contingent on the Riemann hypothesis (RH). In 2023, Cully-Hugill and Dudek proved that under the RH, there exists at least one Goldbach number in the open interval $(\mathrm{x}, \mathrm{x}+9696\cdot\log^2\mathrm{x})$, for all $\mathrm{x}\geq 2$, i.e. that there exists a pair of odd primes adding up to a number in that interval [9-11, 12]. Despite the hard work of many mathematicians, over more than two centuries, no widely accepted, unconditional proof of the bGC exists today.

This paper is organized in five sections. In Section 2, an alternative formulation of the bGC, based on projection, is introduced. In Section 3, formulations of the bGC using optimization methods are discussed, extending prior research results in this area [13]. Section 4 describes a formulation based on the concept of Hybrid Prime Factorization and its application to the algebraic representation of primes and composites, as it applies to the bGC [14, 15]. The approach in Section 5 is based on a fundamental symmetry property of Goldbach primes: for any even number $\mathrm{s}\geq 6$, Goldbach primes are symmetric, i.e. they are equidistant from the midpoint $\mathrm{s}/2$. In Section 5.2 prime symmetry is used in a proof of the bGC for s in a short interval, and Section 5.3 describes the algebraic evolution of Goldbach and non-Goldbach pairs, over larger intervals. In Section 6, translational and rotational shifts of smooth functions are used to reformulate the bGC, and to derive useful representations and approximations of a primality test function and of the prime-counting function $\pi(\mathrm{s})$. Conclusions and a discussion of findings are included in Section 7.

2. Projection

From the Fundamental Theorem of Arithmetic (FTA), it follows that for every integer $\mathrm{n}\in\mathbb{Z}, \mathrm{n}\geq 2$, there is a unique vector $\bar{\mathrm{a}}$ of non-negative integers such that

$$\mathrm{a}^{\mathrm{T}}\log(\bar{\mathrm{p}})=\log(\mathrm{n}) \tag{1}$$

where, on the left-hand side of expression (1), the natural logarithm, $\log(\cdot)$, is understood to apply elementwise on the prime vector $\bar{\mathrm{p}}$ with dimensionality $\pi(n)\times 1$, where $\pi(\cdot)$ denotes the prime counting function.

Since (1) applies to $2\leq \mathrm{n}\leq \mathrm{N}$, it can be expressed in matrix form

$$\mathrm{A}\cdot\log(\bar{\mathrm{p}})=\log(\bar{\mathrm{n}}) \tag{2}$$

where $\mathrm{A}\in\mathbb{Z}^{(N-1)\times\pi(\mathrm{N})}$, and the vector $\bar{\mathrm{n}}\in\mathbb{Z}^{(N-1)\times 1}$ is defined as

$$\bar{\mathrm{n}}^{\mathrm{T}}=[2\ 3\ 4\ 5\ 6\ \dots\ (\mathrm{N}-1)\ \mathrm{N}]. \tag{3}$$

From the FTA, it follows that matrix A is unique; it also has full column rank, since each prime element in $\log(\bar{\mathrm{n}})$ corresponds to a unique unit row of A. Given that the elements of $\bar{\mathrm{p}}$ are integers, expression (2) can be rewritten as

$$\mathrm{P}\cdot\log(\bar{\mathrm{n}})=\log(\bar{\mathrm{n}}) \tag{4}$$

where the columns of the square matrix $\mathrm{P}\in\mathbb{Z}^{(N-1)\times(N-1)}$ match those of matrix A, if the corresponding element of $\bar{\mathrm{n}}$ is prime, otherwise they are equal to a $\bar{0}$ vector.

Example 1. For $\mathrm{N}=6$, after the elementwise prime factorization of $\bar{\mathrm{n}}$, the unique solution, described by (2), is given by

$$\begin{bmatrix}1&0&0\\0&1&0\\2&0&0\\0&0&1\\1&1&0\end{bmatrix}\begin{bmatrix}\log(2)\\\log(3)\\\log(5)\end{bmatrix}=\begin{bmatrix}\log(2)\\\log(3)\\\log(4)\\\log(5)\\\log(6)\end{bmatrix} \tag{5}$$

where a unique unit row is associated with each one of the primes in $\log(\bar{\mathrm{n}})$: 2 (row 1), 3 (row 2) and 5 (row 4). Expression (5) is equivalent to

$$\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\2&0&0&0&0\\0&0&0&1&0\\1&1&0&0&0\end{bmatrix}\begin{bmatrix}\log(2)\\\log(3)\\\log(4)\\\log(5)\\\log(6)\end{bmatrix}=\begin{bmatrix}\log(2)\\\log(3)\\\log(4)\\\log(5)\\\log(6)\end{bmatrix}. \tag{6}$$

From (4) and (6)

$$\mathrm{P}=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\2&0&0&0&0\\0&0&0&1&0\\1&1&0&0&0\end{bmatrix} \tag{7}$$

and since $\mathrm{P}^2=\mathrm{P}$, it follows that $\mathrm{P}\neq\mathrm{I}$, where I denotes the identity matrix, is a non-trivial projection operator in $\mathbb{Z}^5$. Since the largest prime factor of a number n is never greater than n/2, it follows that P is a lower triangular matrix and its rank is the sum of its diagonal elements. This sum equals the number of primes $\le$ n, thus, rank(P) = π(n). For the matrix P in Example 1, rank(P) = π(5) = 3. The FTA guarantees that rank(P) is the smallest possible, hence P is the unique solution to the elementwise factorization of $\bar{n}$ that satisfies (4).

If P is determined through prime factorization, as in the previous example for N = 6, it represents a minimal rank and dimensionality, non-trivial solution to (4), in the form of a projection operator in the vector space $\mathbb{Z}^{(N-1)}$. This general result is summarized in the next proposition.

Proposition 1. If P in expression (4) is generated via prime factorization, the rank of P is the smallest possible and rank(P) = π(N); P is a unique, integer, lower unitriangular and idempotent matrix, and represents a projection operator in $\mathbb{Z}^{(N-1)}$.

Proof. The existence, uniqueness, lower unitriangular form and non-triviality of $P \in \mathbb{Z}^{(N-1)\times(N-1)}$ follows from the FTA. The rank of P is the number of nonzero elements on the main diagonal, or equivalently, due to its unitriangular form, the sum of all elements on its main diagonal; therefore, the rank of P is π(n), which implies that P has an eigenvalue of 1, with multiplicity π(n). P is the unique integer solution, in a rank-minimization sense, to the elementwise (prime) factorization of $\bar{n}$, subject to (4). Observe that $P = I$, where $I \in \mathbb{Z}^{(N-1)\times(N-1)}$ denotes the identity matrix, is a trivial, but suboptimal solution to this rank-minimization factorization problem. Solutions where P corresponds to a non-prime factorization of elements of $\bar{n}$, e.g. representing 18 as $2 \cdot 9$, instead of $2 \cdot 3^2$, are also suboptimal.

Multiplication of both sides of (4) with P yields

$$P^2 \cdot \log(\bar{n}) = P \cdot \log(\bar{n}) \tag{8}$$

and by using (4), the right-hand side of the above expression equals $\log(\bar{n})$. Therefore, (8) may be written as

$$P^2 \cdot \log(\bar{n}) = \log(\bar{n}). \tag{9}$$

Using standard results from linear algebra, it follows that rank(P^2) $\le$ rank(P) [16]. Since the FTA guarantees that P is unique and of minimum rank

$$P^2 = P. \tag{10}$$

Therefore, P is idempotent, and a projection operator in the vector space $\mathbb{Z}^{(N-1)}$.

Remark 1. Given the one-to-one relationship between the matrices A and P, defined by (2) and (4) respectively, it follows that the placement information of primes within $\bar{n}$ is preserved. The prime vector $\bar{p}$, as the unique solution of (2), can be recovered from $\bar{n}$ via the Moore-Penrose pseudoinverse, given by

$$\bar{p} = \exp[(A^T A)^{-1} A^T \log(\bar{n})] \tag{11}$$

where exp($\cdot$) and log($\cdot$) apply elementwise to the corresponding vector elements [17]. Since $A^T A$ is full rank and therefore invertible, it is straightforward to show that if $\bar{p}$ satisfies (2), it also satisfies (11).

Example 2. In Example 1, the matrix A is given by

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

so that

$$(A^T A)^{-1} A^T = \begin{bmatrix} 0.182 & -0.09 & 0.364 & 0 & 0.091 \\ -0.09 & 0.545 & -0.18 & 0 & 0.455 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

and hence, $\bar{p}$ also satisfies (11), since

$$\begin{bmatrix} \log(2) \\ \log(3) \\ \log(5) \end{bmatrix} = \begin{bmatrix} 0.182 & -0.09 & 0.364 & 0 & 0.091 \\ -0.09 & 0.545 & -0.18 & 0 & 0.455 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \log(2) \\ \log(3) \\ \log(4) \\ \log(5) \\ \log(6) \end{bmatrix}.$$

Remark 2. One of the consequences of P being a projection matrix is that the Rank Theorem may be applied, to uniquely represent any vector $\bar{x} \in \mathbb{R}^{(N-1)}$, as follows [18]:

$$\bar{x} = \bar{x}_1 + \bar{x}_2 \tag{12}$$

where the component vectors $\bar{x}_1 \in \mathbb{R}^{(N-1)}$, $\bar{x}_2 \in \mathbb{R}^{(N-1)}$ are unique, given by

$$\bar{x}_1 = P \cdot \bar{x} \tag{13}$$

$$\bar{x}_2 = \bar{x} - \bar{x}_1. \tag{14}$$

From (10) and (13)-(14), it follows that $P \cdot \bar{x}_1 = \bar{x}_1$ and $P \cdot \bar{x}_2 = \bar{0}$, i.e. $\bar{x}_1$ is in the projection space (range) of P and $\bar{x}_2$ is in the null space of P.

As proved in prior research work, prime factorization is the unique solution to an integer minimization problem, where a number is factorially represented with the minimum possible factor sum, including any factors with multiplicity >1 [13]. Therefore, the prime factorization of $\bar{n}$, expressed in matrix form by (2), implies that A is such that each element of the vector $A \cdot \bar{n}$ attains its minimum value. The same is true for matrix P, since the prime factor of a prime is the prime itself. Since P has a zero-column for each non-prime element of $\bar{n}$, the vectors $A \cdot \bar{p}$ and $P \cdot \bar{n}$ are identical, as shown in the following example.

Example 3. From (5)

$$A \cdot \bar{p} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 5 \end{bmatrix} \quad (15)$$

and from (6)

$$P \cdot \bar{n} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 5 \end{bmatrix}. \quad (16)$$

Each element of the resulting vector equals the total prime factor sum of the corresponding number, including any prime factors with multiplicity greater than 1, as in (15)-(16) for $4 = 2^2$.

The vector $P \cdot \bar{n}$ includes the total prime factor sums (including any prime factors with multiplicity greater than 1), corresponding to the factorization of every number in the interval [2, N]. This observation provides a connection with the bGC: if every even number $s \geq 6$ is a sum of two odd primes, then there exists at least one integer $\hat{s} = f(s)$ whose prime factor sum equals s. The set of all $\hat{s}$ is the union of the subsets, $\mathbb{Z}_k$, where each subset containing all $\hat{s}$ with k prime factors, including prime factors with multiplicity greater than 1, if any. Since the smallest prime is 2, it follows that

$$1 \leq k \leq \left\lfloor \frac{s}{2} \right\rfloor$$

where $\lfloor \cdot \rfloor$ denotes integer part. Since the bGC is characterized by Goldbach pairs, the subset $\mathbb{Z}_2$ is of special interest; any $\hat{s} \in \mathbb{Z}_2$ is subsequently referred to as a second-order Goldbach coprime of s.

Definition 1. An integer, $\hat{s}$, is called a Goldbach coprime of order k, of an even number $s \geq 6$, if the prime factorization of $\hat{s}$ has k prime factors and the total sum of all prime factors, including any prime factors with multiplicity > 1, is equal to s.

Remark 3. Let $\hat{s}$ be a Goldbach coprime of order k, corresponding to an even number $s \geq 6$. From the Arithmetic Mean-Geometric Mean inequality, it follows that $\hat{s} \leq \hat{s}_{max}(k, s)$, where $\hat{s}_{max}(k, s) = (s/k)^k$. By using elementary calculus, it can be shown that the global maximum, $e^{s/e}$, of $f(x) = (s/x)^x$ with respect to x, occurs at $x = s/e$, where $e = 2.7183$ ... is Euler's number. Hence, an upper bound of any Goldbach coprime of s, of any order, is $3^{s/3}$. Another way to establish this upper bound, is by showing that, for $x > 3$, the derivative of f(x) is negative, therefore $f(3) > f(x)$. For values of n less than 3, it can be shown that $f(3) > f(2) > f(1)$, for $s \geq 8$. If $s = 6$, n takes 3 values, $n = 1$, $n = 2$ or $n = 3$, and since $(6/2)^2 > (6/3)^3 > (6/1)^1$, f(n) satisfies $(s/n)^n \leq 3^{s/3}$, for $s \geq 6$ and $n \geq 1$.

Example 4. The Goldbach coprimes of $s = 10$, grouped by their order k, are listed below:

$$k = 2{:}\ 25, 21 \quad (17)$$

$$k = 4{:}\ 24 \quad (18)$$

$$k = 5{:}\ 32 \quad (19)$$

The Goldbach coprimes of order 2, for an even $s \geq 6$, are given by

$$\hat{s}_2 = m \cdot (s - m) \quad (20)$$

where $m \in [2, \frac{s}{2} - 2]$. Because the bGC only refers to odd primes, Goldbach coprimes with even values are not relevant and thus, the scope of (20) may be narrowed, to include only odd Goldbach coprimes. Hence, $m \in \left[3, \frac{s}{2} - 3\right]$ is odd, and the two odd Goldbach coprimes of 10, both of order 2, are: 25 (m = 5) and 21 (m = 3). From these two Goldbach coprimes, the Goldbach pairs of 10 can be recovered through prime factorization: $25 \longrightarrow (5, 5)$, $21 \longrightarrow (3, 7), (7, 3)$.

For any even $s \geq 6$, the dimensionality of $\bar{n}$ may be increased, so that $P \cdot \bar{n}$ includes every odd Goldbach coprime of order 2. If bGC is assumed to be true, then every Goldbach pair of s is encoded in the rows of matrix P. More specifically, Goldbach pairs correspond to those rows of P that satisfy

$$\textstyle\sum_{j=1}^{M} P_{ij} = 2. \quad (21)$$

Expressing m in terms of its distance, M, from the midpoint s/2

$$m = \frac{s}{2} - M \quad (22)$$

$$s - m = \frac{s}{2} + M \quad (23)$$

where $M \in [0, \frac{s}{2} - 3]$ is odd, if s/2 is even, and vice versa. Substituting (22)-(23) into (20) gives

$$\hat{s}_2 = \left(\frac{s}{2} - M\right) \cdot \left(\frac{s}{2} + M\right) = \frac{s^2}{4} - M^2 \quad (24)$$

and therefore, $\hat{s}_2$ is maximized for $M = 0$, if s/2 is odd and for $M = 1$ if s/2 is even.

Example 5. From (24), the maximum second-order odd Goldbach coprimes of $s = 10$, $s = 12$ and $s = 18$ are: 100/4 = 25, 144/4 – 1 = 35 and 324/4 = 81, respectively.

From the above analysis, it follows that if $\bar{n}$ is extended to the maximum odd Goldbach coprime of s, the projection of $\bar{n}$, i.e. the vector $P \cdot \bar{n}$, includes all possible second-order Goldbach pairs of s. This result is formalized in the next proposition.

Proposition 2. For any even number $s \geq 6$, using the same notation and definitions as in (1)-(4), where $N = s^2/4$, then every Goldbach pair of s is represented by the projection vector $P \cdot \bar{n}$.

Proof. From the derivation of (4) with $N = s^2/4$, it follows that P represents the prime factorization of all numbers in $[2, s^2/4]$. From the FTA, the bGC and (24), it follows that the product of every Goldbach pair of s is a unique number, no greater than $s^2/4$. Thus, the vector $\bar{n}$, with $N = s^2/4$, contains all possible such numbers and therefore encodes all Goldbach pairs of s.

For $n = 2$, the upper bound $s^2/4$, generated by (24) and $M = 0$, is smaller than the global upper bound, $3^{s/3}$, as discussed in Remark 3. This can also be directly verified through elementary algebra for $s \geq 6$.

Remark 4. Only those elements of the projection vector $P \cdot \bar{n}$ that are equal to s are candidates for possible Goldbach pairs. Of those, as implied by (21), Goldbach pairs correspond to the primes associated only with those elements for which the corresponding row sum of the projection matrix P is 2.

Example 6. For $s = 10$, $N = s^2/4 = 25$, and the dimensionality of P is 24x24. Due to space limitations, it is impractical to display P here in full. Since every Goldbach coprime of order 2 is odd, P can be reduced by half without loss of information, by omitting matrix columns that correspond to even numbers. The resulting 12x12 linear system is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ 9 \\ 11 \\ 13 \\ 15 \\ 17 \\ 19 \\ \mathbf{21} \\ 23 \\ \mathbf{25} \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ 6 \\ 11 \\ 13 \\ 8 \\ 17 \\ 19 \\ \mathbf{10} \\ 23 \\ \mathbf{10} \end{bmatrix}.$$

where the projection vector includes two elements equal to $s = 10$, shown in bold typeface. These elements correspond to the two Goldbach coprimes, 21 and 25, of $s = 10$, also shown in bold. The relevant equations from the above system are

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

and since both row sums are equal to 2, the two Goldbach pairs are (3, 7) and (5, 5). The conjugate Goldbach pair (7, 3) is implied, since multiplication is commutative.

The bGC may be reformulated, in terms of the projection approach described in this section, by expressing Proposition 2 as a conjecture:

Conjecture 1 (Goldbach Pair Projection). For any even number $s \geq 6$, using the same notation and definitions as in (1)-(4), with $N = s^2/4$, the projection vector $P \cdot \bar{n}$ contains at least one Goldbach pair of s.

Remark 5. In the 12 x12 linear system of Example 6, the Goldbach coprime with the smallest projection value is 21; it corresponds to the Goldbach pair (3, 7) and is therefore of order 2. This observation connects the projection approach, discussed in this section, to the minimization approach described in Section 3.

3. Optimization

An alternative formulation of the bGC may be obtained through an optimization approach, described in this section. Consider the class of functions defined by

$$\hat{s} = \hat{s}(p_i, p_j) = p_i \cdot {p_j}^n \tag{25}$$

where p_i, p_j odd primes and n is such that

$$p_i + n \cdot p_j = s \tag{26}$$

where $s \geq 6$ is even. From (25)-(26), it follows that $\hat{s}$ is a Goldbach coprime of order $n + 1$. By constraining

$$n = \frac{s - p_i}{p_j} \in \mathbb{Z} \tag{27}$$

i.e. $n = 1$, or $n > 1$ and p_j to be a prime factor of $s - p_i$, (25) becomes

$$\hat{s} = \hat{s}(p_i, p_j) = p_i \cdot p_j^{\frac{s - p_i}{p_j}}. \tag{28}$$

It can be proved that the minimization of $\hat{s}$, with respect to primes p_i, p_j, results in a Goldbach pair, in the form of a Goldbach coprime of order 2, as stated in the following proposition.

Proposition 3. If the bGC is true, then for any even number $s \geq 6$, a Goldbach pair of primes, p_i, p_j, such that

$$p_i + p_j = s \tag{29}$$

exists and is the solution of the minimization of $\hat{s}$, given by (28), with respect to primes p_i, p_j, subject to (27) and

$$p_i \in [3, \tfrac{s}{2}] \tag{30}$$

$$p_j \in [3, s - p_i]. \tag{31}$$

Proof. If $\frac{s - p_i}{p_j} = 1$, then $\hat{s} = p_i \cdot p_j$ and $p_i + p_j = s$. Therefore, $\hat{s}$ is a Goldbach coprime of order 2. In Section 2, it was shown that all such coprimes satisfy

$$\hat{s} = p_i \cdot p_j = \left(\frac{s}{2} - M\right) \cdot \left(\frac{s}{2} + M\right) = \frac{s^2}{4} - M^2 \tag{32}$$

where M is a non-negative integer. Thus, the value of $\hat{s}$ for any Goldbach coprimes of order 2 does not exceed $\frac{s^2}{4}$.

For the case $n > 1$, since $n \in \mathbb{Z}$ and n is odd, it follows that $n \geq 3$, thus, from (26)

$$p_j \leq \frac{s-p_i}{3}. \tag{33}$$

Next step is to show that, for any even number $s \geq 18$

$$\hat{s}(p_i, p_j) \geq \frac{s^2}{4} \tag{34}$$

which implies that the minimum solution never occurs at a higher order Goldbach coprime of s, i.e. for p_i, p_j such that (26) holds for $n \geq 3$. From (34), given that the minimum solution has a value $< \frac{s^2}{4}$, it must occur at one of the second-order Goldbach primes of s, whose existence is ensured by the assumption that the bGC is true.

Given (28), the inequality (34) can be written as

$$p_j^{\frac{s-p_i}{p_j}} \geq \frac{s^2}{4p_i} \tag{35}$$

and after taking natural logarithms on both sides

$$\frac{s-p_i}{p_j} \cdot \log(p_j) \geq \log\left(\frac{s^2}{4p_i}\right) \tag{36}$$

or equivalently

$$\frac{\log(p_j)}{p_j} \geq \frac{1}{s-p_i} \cdot \log\left(\frac{s^2}{4p_i}\right). \tag{37}$$

Since log(x)/x is strictly decreasing for $x \geq 3$, it follows that

$$\frac{\log(p_j)}{p_j} \geq \frac{\log\left(\frac{s-p_i}{3}\right)}{\frac{s-p_i}{3}} \tag{38}$$

and, therefore, it suffices to show that

$$\frac{\log\left(\frac{s-p_i}{3}\right)}{\frac{s-p_i}{3}} \geq \frac{1}{s-p_i} \cdot \log\left(\frac{s^2}{4p_i}\right) \tag{39}$$

or

$$\log\left[\left(\frac{s-p_i}{3}\right)^3\right] \geq \log\left(\frac{s^2}{4p_i}\right). \tag{40}$$

Since $p_i \leq s/2$, it suffices to have

$$\left(\frac{s}{6}\right)^3 \geq \frac{s^2}{4p_i} \tag{41}$$

and, given that $1/3 \geq 1/p_i$, to show that

$$\left(\frac{s}{6}\right)^3 \geq \frac{s^2}{12} \tag{42}$$

which holds for $s \geq 18$. The numbers 6, 8 and 10 have no higher order Goldbach coprimes, hence the proposition applies trivially. For $s = 12$, the proposition holds, since there is only one second and one fourth-order Goldbach coprime, i.e. $5 \cdot 7 < 3 \cdot 3^3$. Similarly, if $s = 14$, there is only one Goldbach coprime of order >2, i.e. $5 \cdot 3^3$, which exceeds the maximum value, $\frac{s^2}{4} = 7^2$, of the second-order Goldbach coprimes, i.e. $11 \cdot 3 < 7^2 < 5 \cdot 3^3$. For $s = 16$, $3 \cdot 13 < 5 \cdot 11 < 7 \cdot 3^3$. Thus, the proposition holds for $s \geq 6$.

Remark 6. The set of feasible p_j may be further reduced, by observing that, for any given p_i, if $\bar{p}_j > \bar{\bar{p}}_j$ are two prime factors of $s - p_i$, with corresponding Goldbach coprimes $\bar{s}, \bar{\bar{s}}$, respectively, $\bar{s} > \bar{\bar{s}}$, since $\bar{p}_j > \bar{\bar{p}}_j$, and $\log(x)/x$ is strictly decreasing for $x \geq 3$, so that

$$\frac{\log(\bar{p}_j)}{\bar{p}_j} < \frac{\log(\bar{\bar{p}}_j)}{\bar{\bar{p}}_j}. \tag{43}$$

From (43), after some elementary algebra,

$$\log(p_i) + \frac{(s-p_i)}{\bar{p}_j}\log(\bar{p}_j) < \log(p_i) + \frac{(s-p_i)}{\bar{\bar{p}}_j}\log(\bar{\bar{p}}_j)$$

or equivalently

$$\log\left(p_i \cdot \bar{p}_j^{\frac{s-p_i}{\bar{p}_j}}\right) < \log\left(p_i \cdot \bar{\bar{p}}_j^{\frac{s-p_i}{\bar{\bar{p}}_j}}\right).$$

Hence, if there are multiple prime factors of $s - p_i$, only the largest one needs to be considered in the minimization problem of Proposition 3.

Remark 7. Since $\hat{s}$, given by (28), can be very large, taking logarithms of both sides helps in the graphing and computational minimization of $\hat{s}$. Thus, the minimization of $\hat{s}$ can be replaced by that of $\log(\hat{s})$, given by

$$\log(\hat{s}) = \log(p_i) + \frac{s-p_i}{p_j} \cdot \log(p_j). \tag{44}$$

Example 7. For $s = 126$, the values of $\log(\hat{s})$ that correspond to Goldbach coprimes of second-order are shown in Table 1. Coprimes of higher order occur for $p_j < 67$, and result in higher $\log(\hat{s})$ values. Four such instances are:

$$(p_i, p_j) = (3, 3) \longrightarrow \log(\hat{s}) = 46.14$$

$$(p_i, p_j) = (3, 41) \longrightarrow \log(\hat{s}) = 12.24$$

$$(p_i, p_j) = (31, 5) \longrightarrow \log(\hat{s}) = 34.01$$

$$(p_i, p_j) = (61, 13) \longrightarrow \log(\hat{s}) = 16.94.$$

Table 1. *Values of $log(\hat{s})$ for $s = 126$, $p_i \in [3, 61]$ (rows) and $p_j \in [67, 113]$ (columns), corresponding to the second-order Goldbach coprimes of s.*

	67	**71**	**73**	**79**	**83**	**89**	**97**	**101**	**103**	**107**	**109**	**113**
3												
5												
7												
11												
13												7.3
17											7.5	
19										7.6		
23									7.8			
29							7.9					
31												
37						8.1						
41												
43					8.18							
47				8.22								
53			8.26									
59	8.282											
61												

The minimum value of $\log(\hat{s})$, from Table 1, corresponds to the Goldbach pair (13, 113) and is equal to

$$\log(\hat{s}_{min}) = \log(13) + \frac{126-13}{113} \cdot \log(113) = 7.3 \tag{45}$$

All 10 values, in Table 1, are less than $\log(126^2/4) = 8.286$, each corresponding to a Goldbach pair of $s = 126$.

The bGC may be reformulated in terms of the optimization approach described in this section, by expressing Proposition 3 as a conjecture:

Conjecture 2 (Goldbach Pair Optimization). For any even $s \geq 6$, the solution to the minimization of (28) or, equivalently, (44), subject to (27) and (30)-(31), corresponds to a Goldbach pair of s.

4. Hybrid Prime Factorization (HPF)

In this section, an alternative formulation of the bGC is discussed, using the concept of HPF [14, 15]. Consider the first n primes, $p_1, p_2, \ldots, p_n$, in ascending order, where $n \geq 2$. An HPF is an algebraic expression of the form

$$\mathrm{HPF} = p_1^{a_1} \cdot p_2^{a_2} \cdot \ldots \cdot p_n^{a_n} \pm p_1^{b_1} \cdot p_2^{b_2} \cdot \ldots \cdot p_n^{b_n} \tag{46}$$

where the integers a_i, b_i satisfy:

$$a_i, b_i \geq 0 \tag{47}$$

$$a_i \cdot b_i = 0 \tag{48}$$

$$a_i + b_i \geq 1 \tag{49}$$

$$\sum_{i=1}^{n} a_i \geq 1 \tag{50}$$

$$\sum_{i=1}^{n} b_i \geq 1 \tag{51}$$

$$\mathrm{HPF} > 1 \tag{52}$$

for i = 1, 2, …, n. Expression (46) is called an HPF of order n.

An HPF of order n is the sum or difference of two products of the first n primes, where no common factor exists between the product terms, and no prime is excluded.

For any even $s \geq 6$, if there exists an HPF up to dimensionality $m = \pi(s/2)$, such that

$$\mathrm{HPF} = \frac{s}{2} \pm p_1^{b_1} \cdot p_2^{b_2} \cdot \ldots \cdot p_m^{b_m} < p_{m+1}^2 \tag{53}$$

then expression (46) represents a Goldbach pair. For $s = 20$, the HPF of order 3, given by $2 \cdot 5 \pm 3 < 7^2$ represents the Goldbach pair (7, 13) [14].

Unlike the approaches described in the previous two sections, the HPF approach is computationally more complex and algebraically restrictive. Expression (53) is cumbersome to test numerically, especially for large values of s. Alternative methods that extend the scope of the applicability of (53), to more efficiently determine the Goldbach pairs of an even number s, can be found in [14].

The HPF expression is a useful tool for algebraically representing, and searching for, larger primes and coprimes, as reported in [15]. In cases where (53) is feasible, bGC follows; hence, (53) is a sufficient condition for the bGC.

5. Prime Symmetry

Goldbach primes are symmetric with respect to the midpoint s/2. In this section, this symmetry property is used to prove the validity of the bGC within a small interval and analyze how Goldbach pairs evolve as the value of s increases.

5.1. Preliminaries

For any Goldbach pair, $(p_i,\ p_j)$, of an even $s \geq 6$, the relationship

$$\frac{s}{2} - p_i = p_j - \frac{s}{2} \tag{54}$$

is a direct consequence of the bGC, equivalent to

$$p_i + p_j = s. \tag{55}$$

Without loss of generality, let M_i be the distance from the midpoint, $s/2$, i.e.

$$M_i = \left|\frac{s}{2} - p_i\right| = \left|p_j - \frac{s}{2}\right| \tag{56}$$

and let

$$p_i \leq p_j \tag{57}$$

so that (56) becomes

$$M_i = \frac{s}{2} - p_i = p_j - \frac{s}{2} \tag{58}$$

with

$$p_i \in [3, s/2] \tag{59}$$

$$p_j \in [s/2\,, s-3] \tag{60}$$

$$M_i \in [0, \frac{s}{2} - 3]. \tag{61}$$

A consequence of (58) is that, if $s/2$ is odd, then M_i is even and vice versa.

Let m_1, m_2, n be the number of primes in the left closed half-interval, $[3, s/2]$, the right semi-open half-interval, $(s/2\,, s-3]$, and $[3, s-3]$ respectively. Hence,

$$n = \pi(s-3) - 1 = m_1 + m_2 \tag{62}$$

and

$$m_1 + m_2 = \pi(s-3) - 1 \tag{63}$$

For each of the m_1 primes in $[3, s/2]$, define the (possibly non-prime) odd reciprocal, q_i, of p_i, as

$$q_i = s - p_i. \tag{64}$$

Each of the pairs $(p_i,\ q_i)$, $i = 1, \ldots, m_1$ is either of type p-p, if q_i is prime, or type p-q, if q_i is non-prime. Similarly, each of the pairs $(q_j,\ p_j)$, $j = 1, \ldots, m_2$, is either of type p-p, if $q_j = s - p_j$ is prime, or type p-q, if q_j is non-prime.

5.2. Validation of the bGC

This reflective correspondence between primes and odd composites in the two half-intervals can be used to construct a limited-scope proof of the bGC.

Assuming that the bGC is false, then the number of p-p pairs is zero. Therefore, the number of odd composites in $[9, s-3]$ is the sum of three contributions: (1) the reciprocal odd composites to the primes in the left half-interval, (2) the reciprocal odd composites to the primes in the right half-interval, and (3) the pairs of odd composites corresponding to the remaining values of M. A contradiction argument can be constructed, if the number of composites in this third subset is negative, i.e. if the contributions from the first two subsets exceed those of all feasible values of M.

In the special case when s/2 is prime, this approach offers an additional benefit, albeit in a narrow interval: for each $s \in [10, 90]$ where s/2 is prime, there is at least one Goldbach pair with different primes. This stronger version of bGC, denoted here as bGC$^+$, is equivalent to the bGC, if $s/2$ is not prime; if $s/2$ is prime, it implies the existence of at least two Goldbach pairs. The following proposition summarizes this result.

Proposition 4. The bGC$^+$ is true for every even $s \in [10, 90]$.

Proof. The proof is based on the prime-prime and prime-composite relationships described in this section. It shows how a capacity-based argument, combined with prime symmetry leads to proof of a stronger version than the bGC, in the interval $s \in [10, 90]$.

Assume that the bGC$^+$ is false. The three cases: $s/2$ even,

$s/2$ odd non-prime, and $s/2$ prime, are considered separately below.

(i) If $s/2$ is even, let M be such that

$$M \in [0, \tfrac{s}{2} - 3] \text{ and } M\text{: odd.} \tag{65}$$

The number of odd values of M is

$$\#\{M\text{: odd}\} = \frac{\frac{s}{2} - 3 + 1}{2} = \frac{s}{4} - 1 \tag{66}$$

For each of those values of M, there are three possibilities for the pair $\left(\frac{s}{2} - M, \frac{s}{2} + M\right)$, spanning every odd number in $[3, s-3]$:

Each prime in $[3, s/2]$ has an odd reciprocal composite in $(s/2, s-3]$, generated by a non-zero value of M; this contributes m_1 odd composites in $(s/2, s-3]$.

Each prime in $(s/2, s-3]$ has an odd reciprocal composite in $[9, s/2]$, since no odd composite reciprocal < 9 exists; this contributes m_2 odd composites in $[9, s/2]$.

Each of the remaining odd values of M, if any, generates a pair of composites, (q_1, q_2), where $q_1 \in [9, s/2]$ and $q_2 \in (s/2, s-3]$. Since the values of M that generate the composites in the previous two scenarios do not overlap, there are C remaining odd values of M, where

$$C(s) = \frac{s}{4} - 1 - (m_1 + m_2). \tag{67}$$

Using (63), expression (67) can be rewritten as

$$C = \frac{s}{4} - \pi(s-3). \tag{68}$$

To avoid contradiction, (68) should be non-negative. This is not the case for small s, e.g. for $s = 20$, $s/2$ is even, $\pi(17) = 7$ and (63) gives $C = -2$. The Prime Counting Theorem guarantees the existence of some s^*, such that $C \geq 0$, for $s \geq s^*$. Using elementary computations, it can be shown that $s^* = 120$. Figure 1 displays the graph of $C(s)$. For $s \leq 92$, (68) generates a negative value for C, a contradiction.

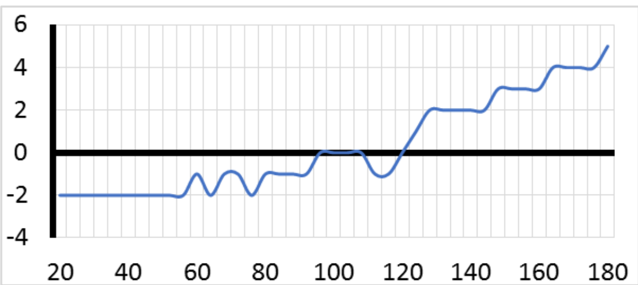

Figure 1. *Smooth line graph of* $C(s) = {}^{s}/_{4} - \pi(s-3)$ *for* $s/2$ *even.* $C(s) < 0$ *for* $s \leq 92$, $s = 112$, *and* $s = 116$.

(ii) If $s/2$ is odd and non-prime, the range of M is

$$M \in [0, \tfrac{s}{2} - 3] \text{ and } M\text{: even} \tag{69}$$

and therefore, the number of values M takes is

$$\#\{M\text{: even}\} = \frac{\frac{s}{2} - 3}{2} + 1 = \frac{s}{4} - \frac{1}{2}. \tag{70}$$

As in the previous case, the primes of each half-interval can be mirrored to an equal number of odd composites, a total of $m_1 + m_2$ odd composites. The number of remaining values, C_e, of M, are

$$C_e(s) = \frac{s}{4} - \frac{1}{2} - (m_1 + m_2) = \frac{s}{4} + \frac{1}{2} - \pi(s-3). \tag{71}$$

As in the previous case, contradiction is avoided only if (71) represents a non-negative quantity. Using similar computations as before, (71) results in $C_e < 0$ for $s \leq 86$, as shown in Figure 2.

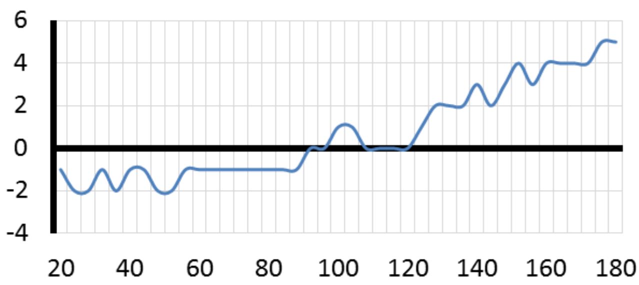

Figure 2. *Smooth line graph of* $C_e(s) = {}^{s}/_{4} + {}^{1}/_{2} - \pi(s-3)$ *for* $s/2$ *odd and non-prime.* $C_e(s) < 0$ *for* $s \leq 86$.

(iii) If $s/2$ is prime, the range of M is

$$M \in [0, \tfrac{s}{2} - 3] \text{ and } M\text{: even.} \tag{72}$$

and therefore, the number of values M takes is

$$\#\{M\text{: even}\} = \frac{\frac{s}{2} - 3}{2} + 1 = \frac{s}{4} - \frac{1}{2}. \tag{73}$$

In this case, there is a trivial Goldbach pair solution, generated by (58) with $M = 0$, i.e. $p_i = p_j = s/2$. Therefore, the bCG is true. Under the assumption that bGC^{+} is false, the remaining nonzero values, $\check{C}_e$, of M, after all primes have been mapped to odd composites is

$$\check{C}_e(s) = \frac{s}{4} - \frac{1}{2} - (m_1 + m_2) \tag{74}$$

or

$$\check{C}_e(s) = \frac{s}{4} + \frac{1}{2} - \pi(s-3). \quad (75)$$

Using similar computations as before, $\check{C}_e(s)$ is negative for $s < 86$. In this case, the bGC+ is stronger than the bGC, and the contradiction implies that there is at least one Goldbach pair, other than (s/2, s/2), for every even $s \in [10, 86]$, with $s/2$ prime. There are 12 primes in [5, 43] and, therefore, it is straightforward to validate that the bGC+ is true for each one: $10 = 3 + 7$, $14 = 3 + 11$, $22 = 3 + 19, \ldots, 38 = 7 + 41, \ldots, 86 = 19 + 67$. Hence, the proposition is true.

The next section extends the above results, by including additional information about the distribution of Goldbach and non-Goldbach pairs of s, between the two half-intervals.

5.3. Encoding the Distribution of Goldbach and Non-Goldbach Pairs

The ideas related to the counts of Goldbach and non-Goldbach pairs in the two half-intervals, described in the previous section, were used to validate a stronger version of the bGC for $s \in [10, 90]$. For $s > 86$, except for $s = 112$ and $s = 116$, as observed from Figure 1, $C(s)$, given by (67), and $C_e(s)$, given by (71), are non-negative. Therefore, the admissible (odd or even) values of M are not exhausted in pairs that include one prime, and hence, no counting contradiction exists.

Subsequent analysis of the bGC should therefore include additional information, such as the distribution of Goldbach and non-Goldbach pairs in the two half-intervals. The formulation approach described in this section is based on a distributional encoding of all such pairs of an even number s. It is shown that such a representation describes the algebraic evolution of Goldbach pairs, over the next $s/2$ values of s, i.e. a portion of the Goldbach pairs of every even $\tilde{s} \in [s + 2, 3s/2)$ may be predicted and algebraically calculated from the distributional encoding of s.

For any even number $s \geq 6$, if $q_i \in [3, \frac{s}{2}]$, $q_j \in [\frac{s}{2}, s-3]$ are odd numbers such that $q_i + q_j = s$, the proposed distributional encoding assigns a value to each pair (q_i, q_j), based on type, as shown in the table below.

Table 2. Encoded values of Goldbach and non-Goldbach pairs.

q_i	q_j	Encoded Value
prime	prime	0
prime	composite	1
composite	prime	2
composite	composite	3

The following example describes the proposed distributional encoding method.

Example 8. For $s = 20$ and $s = 30$, the two half-intervals, corresponding pairs and associated encoded values are shown in Tables 3 and 4 respectively. From Table 3, it follows that the 4-digit distributional encoding of $s = 20$ is $c_d(20) = 0102$.

Table 3. The distributional encoding of $s = 20$ is $c_d(20) = 0102$.

#	q_i	q_j	Value
1	3	17	0
2	5	15	1
3	7	13	0
4	9	11	2

From Table 4 below, it follows that the 7-digit distributional encoding of $s = 30$ is $c_d(30) = 1103003$.

Table 4. For $s = 30$, $c_d(30) = 1103003$.

#	q_i	q_j	Value
1	3	27	1
2	5	25	1
3	7	23	0
4	9	21	3
5	11	19	0
6	13	17	0
7	15	15	3

The distributional encoding, $c_d(s)$, represents the count and relative distribution of Goldbach and non-Goldbach pairs of s. It therefore includes more information than pair-counts, and as such, it can be a useful tool for the analysis of the bGC, as shown below.

Given an even $s \geq 6$ and its distributional encoding, $c_d(s)$, the pair-count information described in Section 5.2 can be recovered, as follows. For s/2 even,

$$\#c_d(s) = k_0 + k_1 + k_2 + k_3 = \frac{s}{4} - 1 \quad (76)$$

$$k_0 + k_1 = m_1 \quad (77)$$

$$k_0 + k_2 = m_2 \quad (78)$$

where $\#c_d(s)$ is the length of $c_d(s)$, i.e. the number of its digits, and k_0, k_1, k_2, k_3 are the number of digits equal to 0, 1, 2 and 3 respectively. Similar expressions may be derived for the other two cases: s/2 odd and non-prime, and s/2 prime.

From the above description of $c_d(s)$, a mathematically equivalent statement to the bGC is that "for any even number $s \geq 6$, its distributional encoding, $c_d(s)$, includes at least one 0". In the previous two examples, $c_d(20)$, $c_d(30)$ have 2 and 3 zeroes respectively. If s/2 is odd, the encodings of s and s+2 have identical lengths, as is the case for $c_d(18) = 1003$ and $c_d(20) = 0102$. This implies that by using the length of the encoding, s can be recovered with an error of ±2, i.e. if $\#c_d(s) = 4$, then $s = 18$ or $s = 20$. In general,

$$s = 3 + 4 \cdot [\#c_d(s)] \pm 1. \tag{79}$$

The inability to recover s from $\#c_d(s)$ does not impede the analysis of how pairs evolve, as the value of s increases.

The distributional information encoded in $c_d(s)$, combined with the uncertainty about the primeness of $s - 1$, and the assumption that $c_d(s)$ includes at least one zero, for any even $s \geq 6$, are used to model the evolution of $c_d(s)$ digits for larger values of s, in the next section.

5.4. Modeling the Evolution of $c_d(s)$

The process by which the digits that comprise $c_d(s)$ evolve to generate portions of $c_d(s+2)$, $c_d(s+4)$ etc., is described below. Consider the example $s^* = 30$, shown in Table 4 of the previous section, and $s^* + 2 = 32$, shown in Table 5 below.

Table 5. Half-intervals and distributional encoding of $s = 32$.

#	q_i	q_j	Value
1	3	29	0
2	5	27	1
3	7	25	1
4	9	23	2
5	11	21	1
6	13	19	0
7	15	17	2

The q_i column in Tables 4 and 5 remains unchanged, while the entries of column q_j, in Table 4, have moved down by one row in Table 5. The increase in the value of s results in a new entry, $29 = s^* - 1$, at the top of column q_j in Table 5. Since it is impossible to positively predict if the new entry is prime or not, based on the knowledge of the pairs in Table 4, it is reasonable to expect that the distributional encoding of s^* includes the necessary information to ensure the presence of at least one 0 in $c_d(s^* + 2)$, i.e. to ensure the continued validity of the bGC for $s^* + 2$. Hence, the uncertainty about the primality of $s^* - 1$, combined with the assumption that the bGC holds for $s^* + 2$, suggests making the hypothesis that the information ensuring the presence of at least one zero in $c_d(s^* + 2)$ is embedded in $c_d(s^*)$. To describe this in more detail, it is necessary to first identify what code-strings, in $c_d(s^*)$, ensure that a 0 is propagated, or created, in $c_d(s^* + 2)$. This is shown in Table 6 below.

Table 6. The 4 propagation substrings in $c_d(s^*)$, ensuring the presence of a Goldbach pair in $s^* + 2$, i.e. the presence of a 0 digit in $c_d(s^* + 2)$.

#	$c_d(s^*)$ substring	$c_d(s^* + 2)$ substring
1	00	x0
2	01	x0
3	20	x0
4	21	x0

The first three substrings propagate an existing Goldbach pair, while the last one creates a new one. Table 6 is the result of two observations: (1) either all, or all but the last one of the entries of column q_i, corresponding to the half-intervals of s^* and $s^* + 2$, are identical; and, (2) the entries of column q_j shift down by one row. Therefore, if $c_d(s^*)$ contains the substrings 00, 01, 20 or 21, it follows from Table 6 that $c_d(s^* + 2)$ also includes the substring x0, where "x" denotes 0, 1, 2 or 3. Thus, the presence of any propagation substring in $c_d(s^*)$ ensures the presence of at least one Goldbach pair for $s^* + 2$, i.e. the validity of the bGC for $s^* + 2$ is embedded in the distributional information contained in $c_d(s^*)$, regardless of the primality or not of $s^* - 1$.

In all previous encoding examples, there is at least one propagation substring present in the distributional encoding of s, shown here in **bold typeface**: $c_d(20) = 0\mathbf{10}2$, $c_d(30) = 1\mathbf{10}300\mathbf{3}$, $c_d(32) = \mathbf{0112102}$.

From the above, it is concluded that the universal validity of the bGC, combined with the uncertainty of the primality of $s^* - 1$ results in the presence of a propagating substring in $c_d(s^*)$. Since this uncertainty also applies to subsequent entries, it follows that at least one of the 0 digits in $c_d(s^* + n)$, where

$$2 \leq n \leq \max(q_i) + \max(q_j) \tag{80}$$

and n is even, is generated by the evolution of the half-interval columnal entries. This implies that as $s^* + n$ increases, up to its maximum value, the evolution of Goldbach pairs can be predicted from the distributional information encoded in $c_d(s^*)$. Since $q_i \in [3, \frac{s^*}{2}]$ and $q_j \in [\frac{s^*}{2}, s^* - 3]$, from (80) it follows that

$$2 \leq n \text{ (even)} \leq \frac{3s^*}{2} - 3 \tag{81}$$

which determines the length of the evolution window for $s^* + n$.

The evolution of pairs described above can be quantified, based on considerations similar to those of Table 6, as shown in Table 7 below.

Table 7. *The 4 m-order propagation substrings in $c_d(s^*)$, ensuring the presence of a Goldbach pair in $s^* + 2(n+1)$, i.e. the presence of a 0 digit in $c_d\big(s^* + 2(n+1)\big)$.*

#	$c_d(s^*)$ substring	$c_d\big(s^* + 2(n+1)\big)$ substring
1	$0X_10$	x0
2	$0X_11$	x0
3	$2X_10$	x0
4	$2X_11$	x0

where X_1 is any substring of length $n \geq 0$, and x is 0, 1, 2 or 3. Table 7 is a generalization of Table 6 (n = 0); if $c_d(s^*)$ contains a substring of length $n \geq 0$, between digits of any of the 4 propagation substrings in Table 7, then $c_d\big(s^* + 2(n+1)\big)$ includes at least one Goldbach pair. For example, if $c_d(s^*)$ contains the substring 030, then $c_d(s^* + 4)$ has at least one 0. The propagation logic for $n \geq 1$ is similar to that which underlies Table 6.

Because of the continuous expansion and simultaneous folding of the two half-intervals, as the value of s increases, the evolution of the last pair, i.e. the last digit of $c_d(\cdot)$ follows a special rule, described in Table 8 below. The logic is similar to the previous two tables, e.g. if $s^*/2$ is even and the last digit of $c_d(s^*)$ is 0 or 2, then the last entry, $\tilde{q} = s^*/2 + 1$, of column q_j, is a prime number, and therefore the last digit of $c_d(s^* + 2)$ is 0, since it corresponds to the prime-prime pair $(\tilde{q}, \tilde{q})$.

Table 8. *Evolution of the last digit of $c_d(s^* + 2)$, given the last digit of $c_d(s^*)$ for s/2 = even.*

s/2 = even		
#	**$c_d(s^*)$ last digit**	**$c_d(s^* + 2)$ last digit**
1	0	0
2	1	3
3	2	0
4	3	3

Table 9. *Evolution of the last digit of $c_d(s^* + 2)$, given the last two digits of $c_d(s^*)$, for s/2 = odd.*

s/2 = odd		
#	**$c_d(s^*)$ ends in**	**$c_d(s^* + 2)$ last digit**
1	00, 20	0
2	10, 30	1
3	03, 23	2
4	13, 33	3

The following table describes the rules by which any two digit substrings of $c_d(s^*)$ determine the evolution of a digit in $c_d(s^* + 2)$. This ruleset is complete, i.e. it includes every two-digit substring combination.

Table 10. *Evolution of a $c_d(s^*)$ substring to a $c_d(s^* + 2)$ digit.*

#	$c_d(s^*)$ substring	$c_d(s^* + 2)$ digit
1	00,01,20,21	0
2	10,11,30,31	1
3	02,03,22,23	2
4	12,13,32,33	3

Table 6 corresponds to the first row of Table 10; the next 3 rows of Table 10 are obtained using similar logic.

Tables 8–10 describe a complete set of rules for the evolution of a rolling portion of $c_d(s)$, from $c_d(s^*)$, where $s^* + 2 \leq s \leq \frac{3s^*}{2} - 3$. Within this range of s, the distributional information in $c_d(s^*)$ determines the evolution of pairs in a portion of $c_d(s)$, ensuring the continued presence of at least one Goldbach pair, regardless of the primality (or not) of the new entries, given by $s^* - 1$, $s^* + 1$, $s^* + 3$ etc., inserted at the top of column q_j.

Example 9. The evolution of pairs of $s^* = 20$, in Table 11 below, as s varies through the evolution window, i.e. for $s = 22, 24, 26$, is shown in Tables 12-14.

Table 11. Pairs and $c_d(s^*)$ digits for $s^* = 20$.

$s^* = 20$			
#	q_i	q_j	Value
1	**3**	**17**	0
2	**5**	**15**	1
3	**7**	**13**	0
4	**9**	**11**	2

Table 12. Pairs and $c_d(s)$ digits for $s = 22$.

$s = 22 = s^* + 2$			
#	q_i	q_j	Value
1	**3**	19	0
2	**5**	**17**	**0**
3	**7**	**15**	**1**
4	**9**	**13**	**2**
5	**11**	**11**	**0**

Table 13. Pairs and $c_d(s)$ digits for $s = 24$.

$s = 24 = s^* + 4$			
#	q_i	q_j	Value
1	**3**	21	1
2	**5**	19	0
3	**7**	**17**	**0**
4	**9**	**15**	**3**
5	**11**	**13**	**0**

Table 14. Pairs and $c_d(s)$ digits for $s = 26$.

$s = 26 = s^* + 6$			
#	q_i	q_j	Value
1	**3**	23	0
2	**5**	21	1
3	**7**	19	0
4	**9**	**17**	**2**
5	**11**	**15**	**1**
6	**13**	**13**	**0**

In Tables 11-14, the pairs of s = 20 are marked in **bold blue** and the associated digits of $c_d(s)$ in **bold**. Tables 12-14 show the evolution of those pairs, and the associated $c_d(s)$ digits, as the value of s increases from 22 to 26, respectively. This evolution follows the rules described in Tables 8−10. The existence of at least one Goldbach pair for s = 22 and s = 24 is guaranteed, since $c_d(20)$ includes the substrings 01 and 010. From Tables 8 and 9, the last digits of $c_d(22)$, $c_d(24)$ evolve in such a way that are both equal to 0. Hence, the presence of at least one 0 digit in $c_d(22)$, $c_d(24)$ and $c_d(26)$ is a direct consequence of how $c_d(20)$ evolves, as the value of s increases.

The last 4 digits of $c_d(22)$ evolve from $c_d(20)$ as follows: each two digit substring in $c_d(20)$ generates one digit of $c_d(22)$ based on Table 10, i.e. 01⟶0, 10⟶1, 02⟶2, and the last digit of $c_d(20)$ generates the last digit of $c_d(22)$, 2⟶0, from Table 8, so that the evolved (in bold) portion of $c_d(22)$ is x0120, where x is the digit associated with the pair (3,19), which corresponds to the new entry, i.e. the number 19, at the top of column q_j. Since it is not possible to know, a priori, if each new entry is prime or not, the assumption that the bGC holds for $s = 22$, implies that at least one of the last 4 digits of $c_d(22)$ is zero.

A similar argument holds for the last 3 digits of $c_d(24)$, i.e. 01⟶0, 12⟶3, and the last digit of $c_d(24)$ is generated from the last two-digit substring of $c_d(22)$, based on Table 9, is 20⟶0, so that the evolved portion of $c_d(24)$ is xx030. Finally, the presence of a zero in $c_d(26)$ is guaranteed, as shown in Table 7, since the last digit of the evolved portion of $c_d(24)$ is 0.

The boundary pair (9, 17), in table 14, corresponds to entries in both columns of Tables 11-13. However, for s = 28, the boundary pair becomes (11, 17), corresponding to entries only from the second column of Table 11.

The evolution window of s^*, expressed by (80)-(81), does not extend beyond $\max(q_i) + \max(q_j)$, because if

$$s > \max(q_i) + \max(q_j)$$

then each evolved pair, $(\tilde{q}_i, \tilde{q}_j)$, of s is generated from the folded entries of the q_j column of s^*, i.e. $\tilde{q}_{i,j} \geq s^*/2$, and not from each column of $s^*/2$, i.e. one entry from each half-interval. This can be observed for s = 18, shown in Table 15

Table 15. Pairs and $c_d(s)$ digits for $s = 18$.

s = 18			
#	q_i	q_j	Value
1	**3**	**15**	**1**
2	**5**	**13**	**0**
3	**7**	**11**	**0**
4	**9**	**9**	**3**

by examining the evolution of its entries outside of the evolution window, e.g. for $s > \max(q_i) + \max(q_j) = 24$. Consider s = 30, shown in Table 16, where only one evolved pair, (15, 15), originates in the right half-interval column of s = 18.

Table 16. Pairs and $c_d(s)$ digits for $s = 30$.

s = 30			
#	q_i	q_j	Value
1	**3**	27	1
2	**5**	25	1
3	**7**	23	0
4	**9**	21	3
5	**11**	19	0
6	**13**	17	0
7	**15**	**15**	**3**

For s = 30, the one remaining evolved pair is not of type 0, hence the distributional information embedded in $c_d(18)$ does not suffice to guarantee at least one 0 digit in the evolved portion of $c_d(30)$.

In summary, the analysis in this section supports a novel hypothesis, expressed below, as a conjecture.

Conjecture 3 (Goldbach Pair Evolution). The existence of at least one 0 digit in $c_d(s + n)$, where

$$2 \leq n \leq \frac{3s^*}{2} - 3 \tag{82}$$

and n is even, is encoded in the distributional information, $c_d(s)$, of s and is expressed in the evolution of its half-interval pairs, according to the propagation rules described in Tables 8−10.

The above conjecture was programmatically implemented using Microsoft's Visual Basic for Applications (VBA) and tested on a personal computer system, for all even $s \in [2, 2000]$. A downloadable list of primes, available online, was used [19]. The cumulative runtime for all evolution scenarios was approximately 40 hours.

6. Analytic Approximation

Prior research has established a link between expressing the prime-counting function, $\pi(s)$, and the bGC analytically [13]. The quantity

$$\hat{T}(n) = 1 - \delta\left[\prod_{i=2}^{\lfloor n/2 \rfloor}\left((n/i) - \lfloor n/i \rfloor\right)\right] \tag{83}$$

where $n \geq 4$, $\delta(\cdot)$ denotes the Kronecker delta function, i.e. $\delta(0) = 1$ and $\delta(x) = 0$ for $x \neq 0$, and $\lfloor x \rfloor$ is the integer part of x, is an algebraic representation of a primality test function, i.e. $\hat{T}(n) = 1$ if n is prime and $\hat{T}(n) = 0$ otherwise.

For any even $s \geq 6$, an algebraic representation of the bGC is

$$D(s) = 0 \tag{84}$$

where

$$D(s) = \prod_{i=3}^{s/2}[P(i) + P(s - i)] \tag{85}$$

$$P(n) = 1 - \hat{T}(n) \tag{86}$$

since it is necessary and sufficient to have $P(i)$ and $P(s - i)$ equal to zero, for some i, i.e. $P(i) + P(s - i) = 0$ [13].

Similarly, an algebraic representation of the prime-counting function may be generated from (83), as follows

$$\pi(s) = 2 + \sum_{j=4}^{s} \hat{T}(j) \tag{87}$$

where $s \geq 4$ is even.

The algebraic expressions (84) and (87) are nonlinear, non-smooth representations of the bGC and $\pi(s)$ respectively. In Sections 6.1 and 6.2, transitional and rotational shifts are used to generate smooth approximations of $\hat{T}(n)$.

6.1. Translational Shifts

The Kronecker delta and integer part operators can be approximated using the un-normalized Gaussian distribution, given by

$$G_\sigma(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right). \tag{88}$$

As the value of $\sigma > 0$ decreases, $G_\sigma(\cdot)$ approximates the Kronecker delta function, i.e.

$$\delta(x) \cong \exp(-\frac{x^2}{2\sigma^2}) \text{ for } \sigma \ll x. \quad (89)$$

The integer part operation is usually approximated by a series of translationally-shifted Heaviside (unit step) functions, where each function is approximated by integrating a shifted instance of the normalized Gaussian density function, $\widehat{G}_\sigma(x)$, given by

$$\widehat{G}_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{x^2}{2\sigma^2}) \quad (90)$$

as follows

$$\lfloor x\rfloor \cong \sum_{1\le k\le x}\int_0^x \widehat{G}_\sigma(y-k)\,dy \text{ for } \sigma \ll x, k. \quad (91)$$

or equivalently

$$\lfloor x\rfloor \cong \sum_{1\le k\le x}\Phi_\sigma(x,k) \text{ for } \sigma \ll x. \quad (92)$$

where

$$\Phi_\sigma(x,k) = \int_{-\infty}^x \widehat{G}_\sigma(y-k)\,dy \quad (93)$$

is the shifted cumulative function.

An approximation of $\widehat{T}(n)$ is obtained, by using (89) and (92), in (83), for $\sigma \ll n$, by writing

$$\widehat{T}(n) \cong 1 - G_\sigma\left(\prod_{i=2}^{\lfloor n/2\rfloor}((n/i) - \lfloor n/i\rfloor)\right). \quad (94)$$

The integer-part operation, $\lfloor n/2\rfloor$, can be omitted, as shown below

$$\widehat{T}(n) \cong 1 - G_\sigma\left(\prod_{2\le i\le\frac{n}{2}} F(n,i)\right) \quad (95)$$

where

$$F(n,i) = (n/i) - \lfloor n/i\rfloor \quad (96)$$

includes an integer-part operation that needs to be approximated. By using (92), the above quantity is approximated by

$$F(n,i) \cong (n/i) - \sum_{1\le k\le n/i}\Phi_\sigma(n/i,k). \quad (97)$$

Alternatively, the shifted Heavyside step function can be approximated by

$$\widehat{\Phi}_\lambda(x,k) = \frac{1}{1+\exp[-(x-k)/\lambda]} \text{ for } \lambda \ll x, k. \quad (98)$$

which does not require integration. The parameters σ and λ do not have to be independent for the approximating functions to work, and therefore it can be assumed that λ = σ, so that an alternative approximation for $F(n,i)$ is

$$F(n,i) \cong (n/i) - \sum_{1\le k\le n/i}\widehat{\Phi}_\sigma(n/i,k) \quad (99)$$

which can be evaluated without performing any integration.

The type of approximation described above generates uncontrollable rounding errors, irrespectively of the value of σ. The root cause of this problem is that any such approximation of $F(n,i)$ results in a numerical error if $n/i \in \mathbb{Z}$. For example, if n = 6 and i = 3, from (99) it follows that

$$F(6,2) \cong 3 - \sum_{1\le k\le 3}\widehat{\Phi}_\sigma(3,k). \quad (100)$$

From (98), by using a large enough λ = σ, the values of $\widehat{\Phi}_\sigma(3,1)$ and $\widehat{\Phi}_\sigma(3,2)$ are very close to 1, but $\widehat{\Phi}_\sigma(3,3) = 1/2$, independently of σ, generating an approximation error of 0.5. A similar issue arises if, instead of (98), expression (92) is used to approximate the integer part operation, since (90) is symmetric around 0, i.e. $\Phi_\sigma(k,k) = 1/2$.

The numerical errors described above can be circumvented by observing that the argument of the delta function, in (83), may be approximated by a sum of shifted delta functions. Since $(n/i) - \lfloor n/i\rfloor$ equals 0, if $n/i \in \mathbb{Z}$, and some value in the open interval (0, 1), if $n/i \notin \mathbb{Z}$, it follows that $\widehat{T}(n)$ may be expressed as a nested function of a shifted series of $\delta(\cdot)$, approximated by $\widetilde{G}(x)$, as shown below.

Consider the expression

$$\widetilde{T}(n) \simeq \sum_{2\le i\le n/2}\widetilde{G}\left(\widetilde{G}(n/i)\right) \quad (101)$$

where

$$\widetilde{G}(x) = \sum_{1\le k\le \max(1,x/2)} G_\sigma(x-k) \quad (102)$$

and $\sigma \ll n$. From (101), if n is prime, every n/i takes a non-integer value, therefore $\widetilde{G}(n/i)$ and $\widetilde{T}(n)$ are close to 0. If n is not prime, at least one of the values for n/i is an integer > 1 and therefore $\widetilde{G}(n/i)$ is close to 1, and $\widetilde{T}(n)$ obtains some nonzero value, explained below.

Expression (101) may be used in conjunction with (84)-(86) to test the validity of the bGC. The following graph shows how such an approximation performs. Figure 3 shows the graph of $\widetilde{T}(j) + \widetilde{T}(s-j)$, for $s = 30$, which equals 0, if (j, s – j) is a Goldbach pair, otherwise its value is equal to the number of factors of j and s – j, respectively, that are less than j and s – j. The value of $1/(2\sigma^2) = 100$.

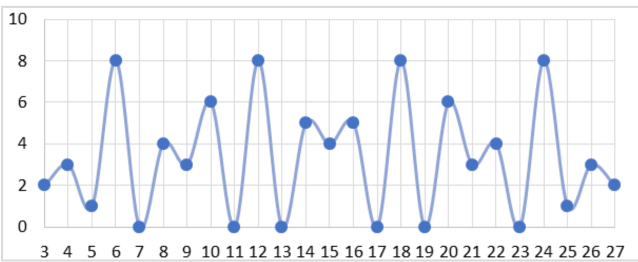

Figure 3. *Smooth line graph of the approximation of $\widetilde{T}(j) + \widetilde{T}(s-j)$ for s = 30 and j = 3,…,27.*

As expected, from the discussion in Section 5, the graph of Figure 3 is symmetric with respect to the mid-point, s/2 = 15, and touches the horizontal axis, i.e. has a value of zero, only on those (symmetric) prime values, for which Goldbach pairs exist, i.e. 7, 11, 13 and 17, 19, 23. Hence, the number of Goldbach pairs is equal to the number of times the graph touches the x axis to the left of the midpoint. Equivalently, given an even $s \geq 6$, the roots of $f(j) = \tilde{T}(j) + \tilde{T}(s-j)$ are symmetric with respect to s/2 and correspond to the Goldbach primes of s.

The nonzero values of the graph in Figure 3 are interpreted as follows: if $f(m)$ is nonzero, its value corresponds to the sum of factors of m and s – m, that are smaller than m and s – m, respectively. For example, $f(12) = 8$, since 12 has 4 factors (2, 3, 4 and 6), and $18 = 30 - 12$ has 4 factors (2, 3, 6 and 9). For m = 3, $f(3) = 2$, and since 3 is prime, it has no factors less than its value; 27 has 2 factors: 3 and 9. Also, since $f(j)$ is symmetric about s/2, it follows that $f(m) = f(s-m)$, for $m \in [3, s-3]$.

A similar approach can be used to generate an approximation of the prime counting function, $\pi(n)$, for $n \geq 6$. Since $\tilde{T}(n)$, given by (101), is close to zero, if n is prime, and non-zero otherwise, it follows that $\pi(n)$ may be approximated by $\hat{\pi}(n)$, where

$$\pi(n) \cong \hat{\pi}(n) = 2 + \sum_{j=4}^{n} G_\sigma\left(\tilde{T}(j)\right) \qquad (103)$$

or, more efficiently, through the recursion

$$\hat{\pi}(n+1) = \hat{\pi}(n) + G_\sigma\left(\tilde{T}(n+1)\right). \qquad (104)$$

The graph in Figure 4 shows a comparison of the above approximation to $\pi(n)$, for $n \in [10, 100]$. To further reduce numerical errors induced in the approximations (103)-(104) for higher values of n, the value of σ decreases, as the value of n increases, i.e. $1/(2\sigma^2) = 5 \cdot n$, generating a maximum error of $|\pi(n) - \hat{\pi}(n)| < 4 \cdot 10^{-12}$. For higher values of n, σ can be further optimized, to result in low approximation errors without numerical overflows.

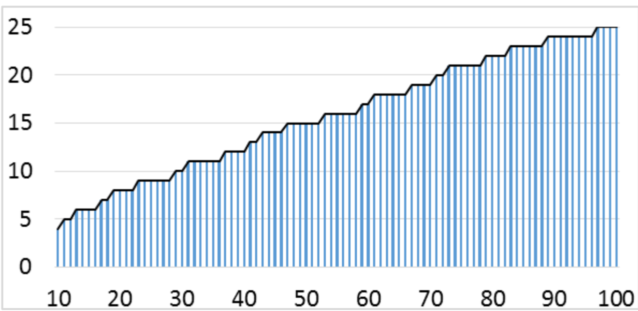

Figure 4. *Graph of the approximation, $\hat{\pi}(n)$, of the prime-counting function, $\pi(n)$, for $n \in [10, 100]$. The line graph corresponds to the values of $\hat{\pi}(n)$ and the bars to $\pi(n)$. The approximation error is* $< 4 \cdot 10^{-12}$.

6.2. Rotational Shifts

Rotational-shift approximations have some advantages compared to those based on translational-shift ones. They are not as prone to numerical errors due to overflow as translational-shift approximations. They allow for additional functional representations, such as trigonometric functions and Fourier series. Rotational-shift approximations also enable a comparison to the harmonic approximation of the prime counting function, $\pi(n)$, generated by adding the spectral contributions of the non-trivial zeroes of the Riemann zeta function [20].

Consider the expression

$$\hat{F}(n) = \prod_{2 \leq i \leq n/2} \left[\mu_w \cdot \sin\left(\frac{n}{i} \cdot \pi\right)\right] \qquad (105)$$

where $\mu_w > 0$ is a scaling parameter, described below, and $n \geq 4$. If n is prime, then $\hat{F}(n) \neq 0$, otherwise $\hat{F}(n) = 0$. By using (105), the prime counting function $\pi(n)$ can be expressed by

$$\pi(n) = \sum_{2 \leq k \leq n} \left[1 - \delta\left(\hat{F}(k)\right)\right] \qquad (106)$$

and approximated by

$$\pi(n) \cong \tilde{\pi}(n) = \sum_{2 \leq k \leq n} \left[1 - G_\sigma\left(\hat{F}(k)\right)\right] \qquad (107)$$

where use of a small value for $\sigma > 0$ lowers the numerical errors resulting from approximating $\delta(\cdot)$ with $G_\sigma(\cdot)$.

If $\mu_w = 1$ and n is prime, as the value of n increases, the product $\left|\prod_{2 \leq i \leq n/2} \left[\sin\left(\frac{n}{i} \cdot \pi\right)\right]\right| \ll 1$, and (105) leads to numerical errors. This is circumvented by using $\mu_w > 1$, so that the contribution of a prime, n, to $\hat{F}(n)$ is amplified. From numerical simulations, for the approximations to work, the value of μ_w should be $\mu_w > \log(\frac{n}{i})$. If μ_w takes larger values, i.e. $\mu_w = n/i$, the values of $\hat{F}(n)$, for larger primes, are also high, leading to numerical errors. In simulation tests, the scaling factor $\mu_w = \left[\log(\frac{n}{i})\right]^2$ was found to perform satisfactorily, i.e. (105) generates consistent separation between primes and non-primes, with insignificant numerical errors.

In Figure 5, $\tilde{\pi}(n)$, given by (107), with $1/(2\sigma^2) = 50$, is compared to the prime counting function, $\pi(n)$, for $n \in [10, 100]$.

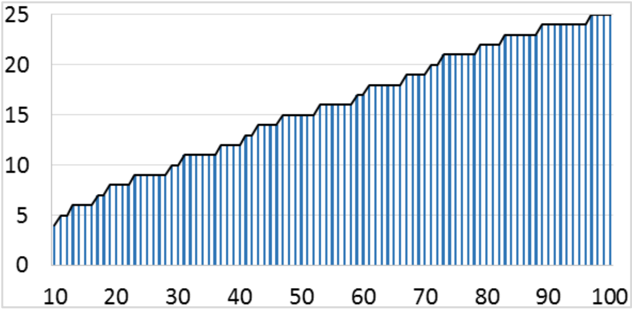

Figure 5. *Graph of the approximation,* $\hat{\pi}(n)$*, of the prime-counting function,* $\pi(n)$*, for* $n \in [10, 100]$*. The line graph corresponds to the values of* $\hat{\pi}(n)$*, and the bars to* $\pi(n)$*. The approximation error is effectively zero, i.e.* $< 10^{-30}$.

7. Conclusions

In the preceding five sections, an equal number of approaches, distinct but interrelated, have been used to discuss, analyze, and reformulate the bGC. From the union of characteristic insights into the bGC, offered by each, a number of broader conclusions may be drawn.

The projection approach, confirms and enhances the links between primes and optimization, reported in earlier research. It shows that, as discussed in Section 2, the set of primes, $\mathbb{P} \subset \mathbb{Z}$, is a unique quantum projection of $\mathbb{Z}$, in a dual sense: $\mathbb{P}$ has the smallest possible cardinality and the lowest possible member values with which to algebraically represent, most efficiently, any member of $\mathbb{Z}$. This unique combination of quantum economy and efficiency may be the most profound characteristic of primes; this has not been systematically explored from a wider range of mathematical and scientific perspectives.

Integer optimization and number theory are closely related areas of mathematics. They are linked over different analytical approaches, as supported by findings in this paper and prior research. The fundamental mathematical properties of primes, related to arithmetic representational efficiency, expressed by the FTA (multiplicative representation) and the bGC (additive representation), is connected to unique solutions of linear and nonlinear integer optimization problems.

The bGC is a critical part of the larger optimal representation problem. Effectively, the bGC states that any positive integer, except 0 and 1, may be algebraically expressed by using the least possible number of operations, i.e. at most two addition operations, performed among the members of a relatively sparse subset, $\mathbb{P} \subset \mathbb{Z}$, having the fewest and smallest members. This unique and universal, additive and multiplicative, representational efficiency of primes renders the bGC and the FTA as fundamental blueprints of mathematics. Viewed from this perspective, the approaches described in this paper, help highlight, connect and analyze fundamental aspects of the bGC and its powerful, ubiquitous and far-reaching properties, not only into the field of number theory and mathematics, but also to other scientific areas, where the efficiency of number representation is of critical importance, such as cryptography and computer science.

The prime symmetry of Goldbach primes is especially interesting, since primes, in general, are not distributed using repeating or symmetrical patterns. This paper, and in particular, Section 5, contributes to a body of research focused on gaining a deeper understanding of how Goldbach prime symmetry affects the values, distribution and relationships of the members of this unique, sparse subset of integers, the prime numbers. Ultimately, such an understanding informs how mathematically quantifiable and scientifically measurable concepts with universal scope, such as symmetry/asymmetry and order/disorder are perceived, modeled and analyzed, both scientifically and philosophically.

The fundamental building blocks of computer science and computer technology have been Boolean algebra, binary logic and high-density, integrated, digital microprocessor architecture, which evolved from rudimentary analog, transistor-based processors. This digital revolution, and the associated rapid technological development it brought with it, is based on a mathematical representation and logic system that exclusively uses 0 and 1, two non-negative integers that are not members of $\mathbb{P}$. Binary computer systems have greatly advanced our technological capabilities in every single area of science and technology. However, their computational inefficiency, as evidenced by their high-energy consumption requirements, is starting to become a concern, especially with the development of large data centers, data mining and processing facilities, and rapid AI technology advancements. Transitioning out of an exclusively binary, linear computational architecture, and more into a prime-based, quantum computing architecture, might present a more viable pathway for achieving computational efficiencies of scale in the future. Such a capability is needed to efficiently and effectively tackle vastly more complex problems, such as climate change, the nature of space, gravity, dark matter and energy, quantum physics and artificial intelligence.

Abbreviations

AI	Artificial Intelligence
bGC	Binary Goldbach Conjecture
bGC^+	Strong Binary Goldbach Conjecture
FTA	Fundamental Theorem of Arithmetic
HPF	Hybrid Prime Factorization
tGC	Tertiary Goldbach Conjecture
RH	Riemann Hypothesis
VBA	Visual Basic for Applications

Author Contributions

Ioannis Papadakis is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press New York NY, USA; 2008 (6^{th} edition), p. 23.

[2] E. W. Weisstein, Goldbach Conjecture. MathWorld-A Wolfram Web Resource. [Online]. Available https://mathworld.wolfram.com/GoldbachConjecture.html

[3] C. K. Caldwell, The Prime Pages Goldbach's conjecture. [Online]. Available https://t5k.org/glossary/page.php?sort=GoldbachConjecture

[4] C. K. Caldwell, The Prime Pages Prime Conjectures and Open Questions. [Online]. Available https://t5k.org/notes/conjectures/

[5] T. Oliveira e Silva, S. Herzog and S. Pardi, Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$. Mathematics of Computation, 83 (2014), 2033-2060, November 2013. Available https://www.ams.org/journals/mcom/2014-83-288/S0025-5718-2013-02787-1/S0025-5718-2013-02787-1.pdf

[6] H. Montgomery and R. Vaughan, The exceptional set of Goldbach's problem. Acta Arithmetica 27.1 (1975) 353-370. http://eudml.org/doc/205349

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

[8] H. A. Helfgott, *The ternary Goldbach Problem*, Ann. of Math. Studies (in publication). Available https://webusers.imj-prg.fr/~harald.helfgott/anglais/publications.html

[9] J. Derbyshire, *Prime Obsession Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*; Plume Washington, DC, USA; 2004.

[10] H. M. Edwards, *Riemann's Zeta Function*, New York NY, USA Academic Press, 1964 (Dover edition reprint, 2001), pp. 1-54.

[11] B. Mazur and W. Stein, *Prime Numbers and the Riemann Hypothesis*, UK Cambridge University Press, 2017 (3^{rd} printing).

[12] M. Cully-Hugill and A. W. Dudek, An explicit mean-value estimate for the Prime Number Theorem in intervals. Journal of the Australian Mathematical Society (2023), pp. 1–15. https://doi.org/10.1017/S1446788723000113

[13] I. N. M. Papadakis, On the Universal Encoding Optimality of Primes. Mathematics 2021, 9 (24), 3155. https://doi.org/10.3390/math9243155

[14] I. N. M. Papadakis, Algebraic Representation of Primes by Hybrid Factorization. Math. Comput. Sci. 2024, 9(1), 12-25. https://doi.org/10.11648/j.mcs.20240901.12

[15] I. Papadakis, Representation and Generation of Prime and Coprime Numbers by Using Structured Algebraic Sums. Math. Comput. Sci. 2024, 9(3), 57-63. https://doi.org/10.11648/j.mcs.20240903.12

[16] W. K. Nicholson, *Linear Algebra with Applications*, Lyryx Open Edition, p. 306 (Corollary 5.4.4). [Online]. Available https://lyryx.com/wp-content/uploads/2018/01/Nicholson-OpenLAWA-2018A.pdf

[17] Wikipedia, Overdetermined system. [Online]. Available https://en.wikipedia.org/w/index.php?title=Overdetermined_system&oldid=1235933306

[18] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill New York, NY, USA; 1976 (3^{rd} edition), p. 228. [Online]. Available: https://www.gbv.de/dms/ilmenau/toc/254172466.PDF

[19] C. K. Caldwell, The Prime Pages The first fifty million primes. [Online]. Available https://t5k.org/lists/small/millions/

[20] H. Riesel and G. Göhl, Some Calculations Related to Riemann's Prime Number Theorem. Mathematics of Computation 1970, vol. 24, no. 112, 969-983. https://doi.org/10.2307/2004630

Research Field

Ioannis Papadakis: Number Theory, Goldbach Conjecture, Optimization, Information Theory, Adaptive and Self-organizing Systems