

Polynomial-Rail and Random-Sequence Synchronisation in Goldbach Prime-Pair Searches: A Reproducible Experimental Framework with Local Modular Prediction

Bouchaib Bahbouhi
Independent Researcher, France
Independent preprint

Abstract

This article develops a reproducible experimental framework for the search and quantitative study of Goldbach prime pairs. For a fixed even integer E , one member of the pair is generated either by a structured quadratic polynomial rail or by a prespecified random sequence of indices, and the complementary member is imposed exactly by $q = E - p$. The principal polynomial family considered here is $f_C(x) = x^2 + x + C$ with C in $\{11, 17, 41\}$, corresponding to discriminants $-43, -67$, and -163 . The central experimental event is bilateral synchronisation: both p and q satisfy the stated primality criterion at the same sampled index. We formalise four states, P/P , P/C , C/P , and C/C ; distinguish probable primality from proved or externally certified primality; derive exact local obstruction counts modulo odd primes; and define finite local scores that can be computed before expensive primality tests. The method is explicitly designed for exact reproduction and independent replication: the pseudo-random generator, seed, domain, sequence length, duplicate policy, sieve depth, software version, and row-level outcomes are frozen and archived. The article also relates the new framework to the minimal Goldbach partition $P_{\min}(E)$, which records the first prime p for which $E-p$ is prime. The complete empirical records for $P_{\min}$ and P/P experiments are not reproduced in the manuscript; they are supplied in a separate supplementary data file containing the row-level observations and audit fields. Two mathematical appendices are reserved exclusively for the synchronisation phenomenon and the proposed prediction method. No claim of a proof of the binary Goldbach conjecture is made.

Keywords: Goldbach conjecture; Goldbach prime pairs; $P_{\min}$; Euler polynomial; quadratic polynomial rails; random sequences; bilateral synchronisation; modular obstruction; local factors; primality testing; computational number theory; reproducibility.

1. Introduction

The binary Goldbach conjecture states that every even integer greater than 2 can be represented as a sum of two primes. It remains one of the best-known open problems in additive number theory. The classical analytic framework originates in the Hardy-Littlewood circle method and its singular-series heuristics [1,2].

Vinogradov proved the corresponding three-prime theorem for sufficiently large odd integers [3], and Helfgott later completed the ternary Goldbach theorem [4,5]. Chen established that every sufficiently large even integer is the sum of a prime and an integer with at most two prime factors [6]. These achievements are profound, but they do not resolve the binary conjecture for every even integer.

Computation supplies a complementary perspective. Oliveira e Silva, Herzog, and Pardi verified the even Goldbach conjecture for every even integer up to 4×10^{18} and simultaneously studied minimal Goldbach partitions [7]. Such work demonstrates the value of treating the first successful representation as a mathematical statistic rather than merely as a certificate of existence. In the present programme this statistic is

denoted $P_{\min}(E)$: the least prime p for which $E-p$ is prime, whenever the search has been audited exhaustively from the boundary.

The present article studies a second, structurally different search mechanism. Instead of restricting attention to consecutive small primes, it samples candidates from polynomial rails and from reproducible random sequences. The motivation is not that a polynomial can generate primes universally - it cannot - but that some quadratic forms display strong local prime-producing behaviour over substantial ranges. Euler's n^2+n+41 is the classical example, and the related forms with constants 17 and 11 correspond to the distinguished discriminants -67 and -43 [8-11].

The key idea is bilateral. A candidate p is never evaluated in isolation: as soon as p is generated, $q=E-p$ is fixed. A successful event therefore requires simultaneous primality on the two sides of a rigid additive identity. We call this a synchronisation event. The word is used operationally rather than metaphysically: it means that two dependent primality conditions are simultaneously satisfied at one predeclared sampled position.

This formulation allows the experimental programme to ask questions that are not visible in a conventional prime-density calculation. How long is a sequence before the first P/P event? How do P/C and C/P states alternate before P/P? Do different polynomial rails exhibit complementary local obstructions? Can an arithmetic score computed from congruence data before primality testing predict which rail or which search window is more likely to synchronise early? These are falsifiable questions and can be investigated without claiming a proof of Goldbach.

A second objective is methodological. Large computational claims are fragile if the realised random sequence, stopping rule, software version, or primality status is not preserved. Consequently, reproducibility is built into the definition of the experiment. Aggregate counts are considered publication-grade only when the underlying row-level observations exist in a materialised data file. The supplementary data file accompanying this work is intended to contain all empirical $P_{\min}$ and P/P observations used for tables, validation, and replication.

The article is organised as follows. Sections 2-4 define the objects and the random-sequence protocol. Sections 5-8 describe local modular filtering, primality status, experimental design, and the relationship between $P_{\min}$ and synchronisation. Sections 9-12 discuss interpretation, validation, limitations, and the relation to the classical literature. Tables 1-6 summarise the fixed definitions and audit requirements. Appendix 1 gives the mathematical description of the synchronisation phenomenon. Appendix 2 gives the mathematical prediction framework. The appendices intentionally contain no additional empirical narrative.

2. Definitions and Notation

2.1. Goldbach pairs and $P_{\min}$

$$E = p + q, \quad p, q \text{ prime}, \quad E \text{ even.}$$

For an even integer E , define the boundary statistic

$$P_{\min}(E) = \min \{ p \text{ prime} : E - p \text{ is prime} \}.$$

When the search over primes p is performed in strictly increasing order and every earlier candidate has been audited, $P_{\min}(E)$ is the first Goldbach synchronisation encountered from the lower boundary. If earlier candidates have not been exhaustively closed, the observed p is a Goldbach hit but not a certified $P_{\min}$ value.

2.2. Polynomial rails

$$f_C(x) = x^2 + x + C, \quad C \text{ in } \{11, 17, 41\}.$$

$$p_C(x) = f_C(x), \quad q_C(E, x) = E - f_C(x).$$

The discriminant of f_C is $\Delta_C = 1 - 4C$. Thus $C=11,17,41$ gives $\Delta_C=-43,-67,-163$, respectively. These values motivate the selected rails historically and arithmetically, but the present work does not assert that they are universally optimal.

2.3. Four-state classification

At each sampled index, the pair is classified as P/P, P/C, C/P, or C/C. In computational logs, P is further annotated as PRP or certified/proved P whenever this distinction matters. A P/P event is the target synchronisation event.

3. Random-Sequence Polynomial Synchronisation Method

3.1. Principle of the sequence experiment

For each fixed E , the experiment begins by freezing an admissible domain I for x , a sequence length L , a pseudo-random number generator, a seed, and a duplicate policy. These choices are made before primality outcomes are inspected. The generator produces $x_1, \dots, x_L$. The realised sequence is then held fixed for all rail comparisons.

At sampled position i and rail C , the algorithm computes $p_{i,C} = f_C(x_i)$ and $q_{i,C} = E - p_{i,C}$. Boundary-invalid values are rejected by a rule fixed in advance. The remaining pair is passed through inexpensive arithmetic filters and only then through stronger primality tests. The stopping rule is also fixed in advance: a run may stop at the first P/P event when the endpoint is first-hit distance, or it may continue through all L positions when the endpoint is P/P frequency.

The same realised x -sequence must be used for $C=11,17,41$ when the purpose is a paired comparison. Otherwise, a difference in observed yield could arise from different samples rather than from the arithmetic of the rails. Independent replication is a separate operation: it changes the seed whilst leaving E , I , L , rails, filters, and stopping rule unchanged.

The method is deliberately different from post-selection. One must not search many random sequences, retain only a successful sequence, and then report it as if it had been prespecified. Every attempted run belongs in the audit file. This is particularly important when studying unusually long P/P-free sequences or unusually high P/P counts.

The sequence need not be statistically uniform for the method to be defined, but its law must be stated. Uniform sampling without replacement is the simplest design for a bounded domain. Sampling with replacement is also legitimate if duplicates are retained as part of the design. Quasi-random or stratified sequences may be studied in future work, but they should not be mixed with the baseline random-sequence cohort without explicit labelling.

3.2. Empirical synchronisation statistics

$$N_{PP}(E,C,L) = \sum_{i=1}^L 1_P(p_{i,C}) 1_P(q_{i,C}).$$

$$S_{\hat{C}}(E,L) = N_{PP}(E,C,L) / L.$$

$$T_C(E) = \min \{ i : p_{i,C} \text{ and } q_{i,C} \text{ are both prime} \}.$$

N_{PP} measures the number of simultaneous events, $S_{\hat{C}}$ is the observed rate per sampled position, and T_C is the first-hit position when a hit exists. For the union of rails, duplicate Goldbach pairs should be deduplicated when counting distinct representations, while first-hit statistics preserve the realised order.

3.3. Reproduction versus replication

Reproduction means re-running the identical experiment from the same E , generator, seed, domain, sequence length, software, and parameters. Replication means repeating the same design with a new seed or on a

prespecified new E. The distinction is essential. Reproduction checks deterministic integrity; replication tests whether the phenomenon persists beyond one realised sequence.

4. Computational Pipeline

4.1. Cheap filters first

Candidate generation is followed by a hierarchy of filters. Parity and obvious congruence restrictions are applied first. For each small odd prime r , an index can be rejected if p_i, C is divisible by r or q_i, C is divisible by r , apart from the exceptional case in which the candidate itself equals r . This stage is exact and inexpensive.

The purpose of presieving is not merely computational speed. The pattern of rejected residue classes is itself an observable. It describes how a rail interacts with the fixed target E and gives a local arithmetic signature that can be compared with later P/P outcomes.

After presieving, surviving candidates may be subjected to Miller-Rabin, Lucas/BPSW-style probable-prime tests, or other strong PRP procedures [13-15,23]. Only the small subset requiring publication-grade certification needs more expensive proof-oriented methods. This cascade prevents the interpretation of a cheap filter as a proof of primality.

4.2. Exact identity audit

Every final pair must satisfy $p+q=E$ exactly. For integers represented compactly as 10^N+a , the compact expression is retained in the audit file because it reduces transcription error. Complete decimal expansions may be supplied separately when an external service requires them.

4.3. Status vocabulary

The manuscript uses C for composite, PRP for a strong probable prime not yet promoted to proved status, and P for a prime regarded as proved or externally certified according to the recorded verification. The supplementary file should record the software or service, date, and returned status whenever external verification is used.

5. Local Modular Structure

5.1. Obstruction count

Let r be an odd prime. Define $nu_r(E, C)$ as the number of residue classes x modulo r for which at least one member of the pair is divisible by r :

$$nu_r(E, C) = \# \{ x \bmod r : f_C(x)[E - f_C(x)] = 0 \bmod r \}.$$

Let $\Delta_C = 1 - 4C$. The congruence $f_C(x) = 0 \bmod r$ has $1 + (\Delta_C/r)$ roots, where $(./r)$ is the Legendre symbol. Likewise, $f_C(x) = E \bmod r$ has $1 + ((\Delta_C + 4E)/r)$ roots. If r does not divide E , the two root sets are disjoint, giving

$$nu_r(E, C) = 2 + (\Delta_C/r) + ((\Delta_C + 4E)/r), \quad r \text{ does not divide } E.$$

If r divides E , the two congruences coincide, giving

$$nu_r(E, C) = 1 + (\Delta_C/r), \quad r \text{ divides } E.$$

These identities are exact elementary statements. They are among the strongest rigorous components of the present framework because they require no probabilistic model.

5.2. Finite local survival score

$$S_C(E; z) = \text{product}_{\{3 \leq r \leq z, r \text{ prime}\}} (1 - nu_r(E, C)/r).$$

$S_C(E;z)$ measures the fraction of residue classes that would survive the listed local obstructions under a multiplicative independence model across distinct primes. It is used as a finite score, not as a theorem giving the P/P probability. The prediction method in Appendix 2 normalises and compares such scores rather than interpreting them as exact probabilities.

6. Experimental Design

6.1. Fixed-E rail comparison

For one E, generate a single sequence and apply it unchanged to all rails. Record the number of positions rejected at each sieve stage, the four-state counts, the first P/P position, and the total number of distinct P/P pairs. This is the cleanest experiment for identifying rail-specific differences.

6.2. Independent-seed replication

Repeat the same fixed-E experiment with several independent seeds. The empirical distribution of first-hit positions and P/P counts can then be studied without changing the underlying target. Long sequences are especially informative when the objective is to measure the tail of the first-hit distribution rather than merely to find one pair.

6.3. Cohorts across E

A cohort experiment freezes the design and applies it to many even integers. The E-values themselves must be prespecified or generated by a declared rule. A cohort may study increasing decimal scale, local offsets E_0+2k , or a family such as 10^N+c . Coverage is the fraction of targets for which at least one P/P event is found within the fixed computational horizon.

6.4. Negative and structural controls

Controls are required to distinguish a Goldbach-specific effect from the intrinsic prime density of the polynomial. Suitable controls include alternative quadratics, altered complements that do not preserve the same additive geometry, or shuffled associations between p-values and E-values. Controls must use the same sequence length and testing criteria as the principal experiment.

7. P_{min} and Boundary Synchronisation

P_{min}(E) and random-sequence synchronisation measure different aspects of the same additive constraint. P_{min} is ordered and boundary-based: candidates p are tested from the smallest prime upwards. The random-sequence method is sampling-based: it deliberately breaks the monotone order in order to probe the candidate space through a fixed stochastic design.

The two methods should therefore not be conflated. P_{min} is especially useful for studying the first Goldbach partition and for comparison with the minimal-partition literature [7]. Random-sequence P/P counts are useful for studying multiplicity, persistence, and rail complementarity. A target may have a small P_{min} yet exhibit a long random first-hit distance under a particular rail, or conversely.

The supplementary data file is intended to keep both families of observations. P_{min} rows should include E, the observed or certified P_{min}, $q=E-P_{\min}$, the audit status of all smaller candidates when minimality is claimed, and the primality status of q. Random-sequence rows should include E, seed, position, x, C, p, q, sieve status, primality status, and run identifier.

For publication, a distinction is made between a Goldbach hit and a minimal Goldbach partition. A P/P pair found by a polynomial rail is a valid representation if both members are prime, but it is not automatically

$P_{\min}$. This prevents the structured search from accidentally being described as an exhaustive boundary search.

8. Supplementary Empirical Data File

All empirical $P_{\min}$ and P/P data supporting this article are to be supplied in a separate supplementary data file. The main manuscript intentionally contains only definitions, methods, summary tables, mathematical derivations, and interpretation. The supplementary file is the authoritative source for row-level computational observations.

The $P_{\min}$ component should contain at least: target identifier; exact E or compact exact representation; decimal size; candidate prime p ; complementary q ; primality status; whether minimality is exhaustive or provisional; number of candidates examined; sieve parameters; software; and independent verification notes. Where a prediction window is used, the predicted threshold must be recorded before the search outcome.

The P/P component should contain at least: run identifier; E ; rail C ; PRNG; seed; domain; L ; sampled position; x ; p ; q ; modular rejection information; PRP/P/C status for each side; first-hit indicator; and distinct-pair identifier. The complete realised sequence or an equivalent lossless representation should be retained.

This separation between article and data file serves two purposes. First, it keeps a long manuscript readable. Secondly, it allows every numerical summary to be regenerated programmatically. If a table in the manuscript changes after an audit, the underlying rows remain unchanged and the revision can be traced.

9. Results and Empirical Interpretation

The experimental programme motivating this framework has produced repeated P/P events on Euler-type rails, repeated P/P events in long random-sequence experiments, and a growing collection of boundary $P_{\min}$ observations. These observations motivated the present formalisation of synchronisation and the insistence on row-level materialisation.

The strongest result that can be stated independently of the supplementary empirical file is the exact local congruence formula of Section 5. It demonstrates that the quality of a polynomial rail depends jointly on its discriminant Δ_C and on the target-dependent shifted discriminant $\Delta_C + 4E$. Consequently, rail behaviour is not expected to be constant across E .

The empirical programme further suggests that using several rails can be more robust than relying on a single quadratic. Such complementarity is plausible because the Legendre-symbol patterns entering $\nu_r(E, C)$ differ with C . Nevertheless, any quantitative claim about gain from the union must be calculated from the supplementary file after deduplication and must include unsuccessful runs.

The same caution applies to very large examples. A compact expression for q can be exact, but the status of q depends on the recorded primality test. The article therefore avoids presenting a list of large primes without the associated audit trail. Selected examples may be highlighted in future revisions once the supplementary file and independent certification records are frozen.

No finite computational success rate is interpreted as a proof of Goldbach. The scientific claim is narrower: the framework creates reproducible observables for the transition from locally admissible candidates to simultaneous primality, and it provides a mathematical local score that can be tested prospectively.

10. Statistical Interpretation of the Synchronisation Phenomenon

A naive model would treat primality of p and q as approximately independent events of probabilities $1/\log p$ and $1/\log q$. The Goldbach constraint and the polynomial rail invalidate literal independence because

congruence conditions couple the two sides. The local obstruction count is therefore introduced before any global density heuristic.

After conditioning on survival of small moduli, a residual heuristic can be formulated in terms of inverse logarithms. This is analogous in spirit to the singular-series correction in the Hardy-Littlewood framework [1,2], although the sampling measure is different because p is restricted to polynomial values rather than all integers or all primes in an interval.

The first-hit statistic is naturally a waiting-time variable. If the conditional success probability per sampled position were constant and independent, T_C would be geometric. Real data need not satisfy those assumptions. Deviations from a geometric survival curve are therefore informative: they may indicate changing candidate size, local correlations, duplicate structure, or heterogeneity in the rail.

For a fixed E and finite sequence, the empirical survival function can be defined as the fraction of replicated runs with no P/P event by position k . Comparing this function between rails is more informative than comparing only mean first-hit positions, especially when the distribution has a heavy tail.

When the objective is prediction, the crucial rule is prospective freezing. A score, threshold, or ordering must be computed before the P/P search whose outcome will be used to evaluate it. This is the same discipline applied to the $P_{\min}$ prediction programme: calibration and validation cohorts must remain separate.

11. Relation to the Classical Literature

The Hardy-Littlewood conjectural formula for Goldbach representations predicts a main term modified by local arithmetic factors [1,2]. The present work shares the principle that local congruence information matters, but it changes the candidate measure by imposing a polynomial rail. For this reason, the classical Goldbach singular series cannot simply be copied as an exact predictor of rail performance.

The minimal-partition computations of Oliveira e Silva, Herzog, and Pardi are particularly relevant because $P_{\min}$ is a boundary statistic and because their large verification demonstrates the value of recording first occurrences, not only existence [7]. Their computation verifies every even integer through 4×10^{18} , whereas the present very-large-integer experiments sample sparse, explicitly defined families and do not constitute an exhaustive interval verification.

Chen's theorem and modern Chen-prime work show how close sieve theory can approach binary prime-prime representation while confronting the parity barrier [6,17,25]. The present computational filters encounter an analogous practical distinction: sieving can make q rough, but roughness alone does not certify primality. This is why the pipeline keeps PRP, P , and composite states separate.

Recent work on the exceptional set of the Goldbach problem continues to refine major-arc analysis and the size of sets not covered by known methods [18]. Such results underscore the gap between strong average statements and a theorem valid for every even E . The synchronisation framework should be judged within that context: it proposes experimental structure and exact local identities, not a replacement for the missing uniform analytic estimate.

The Euler quadratics are used because of their classical prime-producing behaviour and their connection with discriminants -43, -67, and -163 [8-11]. The historical significance of these discriminants does not imply that the three rails are uniquely privileged for Goldbach. An expanded study should compare them with randomly chosen and deliberately unfavourable control polynomials.

12. Prediction as a Prospective Experiment

The proposed prediction problem is deliberately modest: before strong primality testing, rank candidate rails, blocks, or search windows according to information available from E and cheap modular computations. A

successful predictor should reduce the expected computational effort or improve the probability of finding P/P within a fixed budget.

A predictor can combine a local survival score, logarithmic prime-density terms, and a penalty for repeated or structurally redundant candidates. The coefficients, if any, must be calibrated on a training set and then frozen. Evaluation is performed on a disjoint validation set. The relevant endpoints include first-hit rank, P/P coverage at fixed L, and computational cost.

The P_min programme supplies an independent prediction setting. There the target is not the best polynomial rail but a boundary threshold $H(E)$ within which the first Goldbach prime is expected. Empirical scales such as $\log(E)$ $\log \log(E)$ can be investigated, but they are heuristic unless supported by a theorem. A prediction interval is not a proof that P_min lies inside it.

Appendix 2 gives the mathematical form of the local predictor and shows how it can be normalised. The appendix intentionally avoids empirical claims. All calibration constants and validation outcomes belong to the supplementary data file or to the main Results section once audited.

13. Reproducibility, Audit, and Falsification

A fixed experimental design must have a visible failure state. For a random-sequence run, failure means no P/P event within the prespecified L positions. Increasing L after observing failure creates a new run and must not erase the original failure. For a P_min prediction, failure means that the certified first hit lies outside the frozen prediction window.

Every attempt must be retained. This rule prevents survivorship bias, which is particularly dangerous when many seeds, rails, or offsets are tried. A striking successful run has little inferential value if unsuccessful attempts have disappeared from the record.

Independent primality checks are desirable for selected large candidates. However, external services use different status labels and may update their databases. The supplementary file should therefore store the date and the exact returned status, and the manuscript should not silently convert a probable-prime label into a proof.

Code versioning is equally important. Changes in the presieve, random-number generator, or candidate ordering can change first-hit statistics even if the mathematical definition is unchanged. The publication package should therefore contain a versioned script or pseudocode sufficient to regenerate the source rows.

A final audit should regenerate Tables 1-6 from definitions or source data, verify every compact identity $p+q=E$, confirm that no figure is required for interpretation, and check that Appendix 1 and Appendix 2 contain only mathematical explanation as requested.

14. Limitations

The principal limitation is logical: computational evidence cannot prove an infinite universal statement. Even experiments at thousands of decimal digits sample an infinitesimal subset of the even integers. The method may discover patterns that motivate a theorem, but the theorem requires an independent proof.

The second limitation is selection. Polynomial rails deliberately enrich a subset of integers and therefore change the sampling distribution. Observed P/P frequencies cannot be compared directly with frequencies from uniform integer sampling without correcting for the different candidate measures.

The third limitation is computational. At very large sizes, probable-prime tests are feasible on many survivors, whereas proof-oriented certification can be much more expensive. Consequently, large cohorts may naturally contain a mixture of certified P and PRP statuses. These categories must remain separate.

The fourth limitation is model risk. Local modular factors capture small-prime obstructions exactly, but they do not by themselves control the distribution of large prime factors or guarantee independence across positions. A strong empirical correlation between a local score and P/P yield would therefore remain a predictive observation until an analytic theorem is established.

Finally, the chosen rails are not exhaustive. Other quadratic polynomials, higher-degree forms, linear prime sequences, or hybrid candidate generators may perform differently. The present article establishes a framework in which such alternatives can be compared fairly.

15. Conclusions

This work introduces a reproducible framework for studying Goldbach prime-pair synchronisation through structured polynomial rails and fixed random sequences. Its defining feature is bilateral construction: p is generated by a declared rule and $q=E-p$ is imposed exactly, so success is a simultaneous event rather than a one-sided prime-generation result.

Three levels of information are kept separate. Exact arithmetic includes the Goldbach identity and the local obstruction formula $\text{nu}_r(E,C)$. Computational evidence includes $P_{\min}$ observations, P/P counts, and first-hit distances. Prediction consists of prospective scores or windows that must be frozen before the corresponding search. This separation is essential for a credible research programme.

The accompanying supplementary data file is intended to contain all empirical $P_{\min}$ and P/P source rows. The manuscript itself contains only six tables and no figures. Appendix 1 isolates the mathematics of synchronisation; Appendix 2 isolates the mathematics of prediction. This structure is designed to make later empirical expansion possible without changing the core definitions.

The immediate theoretical question is whether the finite local information can be connected to a rigorous lower bound for the number of simultaneous prime values in a useful range. Until such a result is obtained, the synchronisation score remains a heuristic predictor and the binary Goldbach conjecture remains open.

16. Tables

Table 1. Polynomial rails and discriminants

C	Polynomial $f_C(x)$	Discriminant Δ_C
11	$x^2 + x + 11$	-43
17	$x^2 + x + 17$	-67
41	$x^2 + x + 41$	-163

The three rails are evaluated under identical sequence conditions in paired comparisons.

Table 2. Four-state classification

State	p-side	q-side	Interpretation
P/P	Prime	Prime	Goldbach synchronisation
P/C	Prime	Composite	One-sided success
C/P	Composite	Prime	Complementary one-sided success
C/C	Composite	Composite	No prime side

PRP is retained as a separate computational status in the source data and is not silently promoted to P.

Table 3. Random-sequence reproducibility fields

Field	Required record
Target	Exact E
Generator	PRNG and implementation
Seed	Exact seed
Domain	Admissible interval for x
Length	L
Duplicate policy	Retain or remove, fixed in advance
Rails	Ordered list of C values
Presieve	Prime bound and rules
Primality tests	Algorithms, parameters, software version
Stopping rule	First hit or full length

These fields are sufficient to distinguish exact reproduction from independent-seed replication.

Table 4. Principal synchronisation statistics

Quantity	Definition	Use
N_PP	Number of P/P events in L positions	Multiplicity
S_hat	N_{PP}/L	Empirical rate
T_C	First P/P position	Waiting distance
Coverage	Targets with ≥ 1 P/P / targets tested	Cohort endpoint
$nu_r(E,C)$	Locally excluded residue classes mod r	Exact local obstruction
$S_C(E;z)$	Finite product of local survival factors	Prospective local score

No finite score is interpreted as a proof of existence.

Table 5. P_min supplementary-data fields

Field	Meaning
E	Exact even target
p	Observed first prime candidate
q	E-p
Minimality status	Exhaustive or provisional
Candidates tested	Boundary search depth
q status	C, PRP, or P
Prediction window	Frozen threshold, if used

Verification	Software/service and date
--------------	---------------------------

The complete empirical P_min records are supplied in the separate supplementary data file.

Table 6. P/P supplementary-data fields

Field	Meaning
Run ID	Unique experiment identifier
E	Exact target
C	Polynomial rail
Seed	Random-sequence seed
i, x_i	Position and sampled index
p, q	Exact complementary pair
Sieve outcome	First rejecting modulus or survivor
p status	C/PRP/P
q status	C/PRP/P
First-hit flag	Whether row closes the first P/P search

The complete empirical P/P records are supplied in the separate supplementary data file.

Appendix 1. Mathematical Description of the Synchronisation Phenomenon

Let E be fixed and even, and let $f(x)=x^2+x+C$. Define the two primality indicator functions $A_{E,C}(x)=1_P(f(x))$ and $B_{E,C}(x)=1_P(E-f(x))$. The synchronisation indicator is their product $Z_{E,C}(x)=A_{E,C}(x)B_{E,C}(x)$. A P/P event occurs exactly when $Z_{E,C}(x)=1$.

For a finite deterministic set X of sampled indices, the synchronisation count is $G_C(E;X)=\sum_{x \in X} Z_{E,C}(x)$. For an ordered sequence $X_L=(x_1, \dots, x_L)$, the first synchronisation index is $T_C(E;X_L)=\min\{i: Z_{E,C}(x_i)=1\}$, with $T=\infty$ if no event occurs.

The modular structure is obtained by replacing primality with local admissibility. For an odd prime r, define $D_r(E,C)=\{x \bmod r: f_C(x)=0 \bmod r \text{ or } f_C(x)=E \bmod r\}$. Then $\text{nu}_r(E,C)=|D_r(E,C)|$. The discriminants of the two quadratic congruences are Δ_C and Δ_C+4E . Hence, for r not dividing E, their solution sets are disjoint and $\text{nu}_r(E,C)=2+(\Delta_C/r)+((\Delta_C+4E)/r)$. If r divides E, the equations coincide and $\text{nu}_r(E,C)=1+(\Delta_C/r)$.

For a finite prime set R, the Chinese remainder theorem gives an exact admissible residue proportion when the conditions are combined modulo the square-free product $M=\text{product}_{r \in R} r$. Because the restrictions are coordinatewise modulo distinct primes, the number of admissible classes modulo M is $\text{product}_{r \in R} (r - \text{nu}_r(E,C))$. Therefore the exact local admissible proportion is $\text{product}_{r \in R} (1 - \text{nu}_r(E,C)/r)$. This statement concerns avoidance of the selected small prime divisors; it does not state that the surviving values are prime.

The distinction between local admissibility and primality is the mathematical core of the synchronisation phenomenon. A position may survive every $r \leq z$ and still be C/C, P/C, or C/P because one or both numbers possess larger factors. Thus the transition from local bilateral admissibility to P/P is precisely the part not settled by the elementary congruence calculation.

If one introduces heuristic prime densities after local conditioning, a natural schematic expression for the expected contribution of an index x is proportional to $W_{E,C}(x;z)/[\log f_C(x) \log(E-f_C(x))]$, where W is a

correction derived from the local factors. This is a heuristic bridge to the Hardy-Littlewood philosophy, not a proven asymptotic formula for the polynomially sampled problem.

For multiple rails C in a finite set C_{set} , define $Z_{\text{union}}(x) = 1 - \max_{C \in C_{\text{set}}} (1 - Z_{E,C}(x))$. The union count is not generally the sum of individual counts because two rails can produce the same pair or can synchronise at the same sampled position. Inclusion-exclusion or direct deduplication is therefore required for exact union statistics.

The waiting-time formulation follows directly from the ordered sequence. If hypothetical independent success probabilities π_i were assigned to positions, then $P(T > k) = \prod_{i \leq k} (1 - \pi_i)$. This identity is exact only within the hypothetical independent Bernoulli model. Empirically observed departures from such a curve quantify dependence or heterogeneity; they are not anomalies to be removed.

P_{min} fits the same formal language with a different ordered candidate set. Let $p_1 < p_2 < \dots$ be the primes. Define $Z_{E(j)} = 1 - P(E = p_j)$. Then $P_{\text{min}}(E) = p_T$ where $T = \min\{j : Z_{E(j)} = 1\}$. Thus boundary P_{min} and polynomial random-sequence synchronisation are two waiting-time constructions over different candidate orderings.

Appendix 2. Mathematical Form of the Prediction Method

The prediction problem is to assign, before strong primality testing, a score to a rail, block, or candidate position using only E , the candidate-generation rule, and inexpensive arithmetic information. Let R_z denote the odd primes $r \leq z$. Define the exact finite local factor $L_C(E; z) = \prod_{r \in R_z} (1 - \nu_r(E, C)/r)$.

Because L_C is affected by the baseline density of residue classes, a normalised score can be formed relative to a reference model. One convenient family is $Q_C(E; z) = L_C(E; z)/B(z)$, where $B(z)$ is a fixed baseline independent of C for the comparison under consideration. The choice of B changes scale but not rail ranking when it is common to all rails.

Candidate size enters through logarithmic density. For a sampled position x , define $H_{E,C}(x) = Q_C(E; z) / [\log f_C(x) \log(E - f_C(x))]$ whenever both arguments exceed 1. A block score over a prespecified set X is $H_{E,C}(X) = \sum_{x \in X} H_{E,C}(x)$. This is a heuristic expected-yield score; it is not asserted to equal the true expected number of P/P events.

For rail selection, the prospective rule is $C_{\text{star}}(E) = \arg\max_C H_{E,C}(X)$. Its validation is purely empirical: compute C_{star} without looking at strong primality outcomes, run the fixed experiment, and compare first-hit rank or P/P count with the competing rails. Re-estimating the rule after seeing the validation outcomes invalidates that validation cohort.

For first-hit prediction within one rail, define cumulative score $A_C(k) = \sum_{i=1}^k H_{E,C}(x_i)$. Under a Poisson-style heuristic, a schematic survival approximation is $P(T_C > k) \approx \exp(-\alpha A_C(k))$, where α is a calibration constant estimated only on training data. Quantile predictions solve $\alpha A_C(k_{\text{tau}}) = -\log(1 - \text{tau})$. This equation defines a prospective prediction mechanism but is not a theorem about prime values.

For P_{min} , let p_j be the j -th candidate prime in increasing order and define $h_E(p_j)$ as a locally corrected heuristic weight proportional to $1/\log(E - p_j)$, with the appropriate Goldbach local factor. The cumulative boundary intensity is $A_E(H) = \sum_{p \leq H} h_E(p)$. A heuristic quantile H_{tau} can again be defined by $\alpha A_E(H_{\text{tau}}) = -\log(1 - \text{tau})$. Empirical scales such as $\log(E) \log \log(E)$ may be used to parameterise H , but no uniform polylogarithmic bound is claimed.

A prediction model is considered successful only prospectively. Let Y be an endpoint such as first-hit rank or the indicator of at least one P/P event within L . Training data determine α or any additional coefficients. Validation data are then evaluated without modification. Coverage, calibration, and ranking performance are reported together with failures.

The local score can also be used diagnostically. If two rails have similar logarithmic candidate sizes but substantially different $L_C(E;z)$, the model predicts a difference caused by small-prime obstructions. If observed performance contradicts that ordering repeatedly on validation data, the local model is incomplete and must be rejected or extended rather than retrospectively reinterpreted.

The mathematical status of the prediction method is therefore explicit. The congruence counts and finite admissible products are exact. The conversion of admissibility into prime-pair probability is heuristic. Calibration constants are empirical. A future theorem would require a rigorous estimate connecting the weighted simultaneous-prime count to these local factors with an error term strong enough for the intended range of E and X .

Author Contributions

Conceptualisation, methodology, investigation, experimental programme, and mathematical formulation: Bouchaib Bahbouhi. Computational and editorial assistance was used under the author's direction. The author remains responsible for verification of the mathematical statements, empirical data, and final submitted manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

A separate supplementary data file accompanies the manuscript and is intended to contain the complete empirical $P_{\min}$ and P/P row-level data used in the study, together with reproducibility fields, sequence identifiers, primality status, and audit information. The main article does not reproduce the full empirical dataset.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Hardy, G.H.; Littlewood, J.E. Some problems of 'Partitio Numerorum'; III: On the expression of a number as a sum of primes. *Acta Math.* 1923, 44, 1-70.
2. Vaughan, R.C. *The Hardy-Littlewood Method*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1997.
3. Vinogradov, I.M. Representation of an odd number as a sum of three primes. *Dokl. Akad. Nauk SSSR* 1937, 15, 291-294.

4. Helfgott, H.A. Major arcs for Goldbach's problem. arXiv 2013, arXiv:1305.2897.
5. Helfgott, H.A. The ternary Goldbach problem. In Proceedings of the International Congress of Mathematicians; Seoul, Korea, 2014; Vol. II, pp. 391-418.
6. Chen, J.R. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. Sci. Sin. 1973, 16, 157-176.
7. Oliveira e Silva, T.; Herzog, S.; Pardi, S. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to 4×10^{18} . Math. Comp. 2014, 83, 2033-2060. <https://doi.org/10.1090/S0025-5718-2013-02787-1>.
8. Heffernan, R.; Lord, N.; MacHale, D. Euler's prime-producing polynomial revisited. Math. Gaz. 2024, 108. <https://doi.org/10.1017/mag.2024.11>.
9. Rabinowitsch, G. Eindeutigkeit der Zerlegung in Primzahlfaktoren in quadratischen Zahlkörpern. Proc. Fifth Int. Congress Math. 1913, 1, 418-421.
10. Hardy, G.H.; Wright, E.M. An Introduction to the Theory of Numbers, 6th ed.; Oxford University Press: Oxford, UK, 2008.
11. Cox, D.A. Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication, 2nd ed.; Wiley: Hoboken, NJ, USA, 2013.
12. Ribenboim, P. The New Book of Prime Number Records; Springer: New York, NY, USA, 1996.
13. Miller, G.L. Riemann's hypothesis and tests for primality. J. Comput. Syst. Sci. 1976, 13, 300-317.
14. Rabin, M.O. Probabilistic algorithm for testing primality. J. Number Theory 1980, 12, 128-138.
15. Baillie, R.; Wagstaff, S.S. Lucas pseudoprimes. Math. Comp. 1980, 35, 1391-1417.
16. Montgomery, H.L.; Vaughan, R.C. The exceptional set in Goldbach's problem. Acta Arith. 1975, 27, 353-370.
17. Iwaniec, H.; Kowalski, E. Analytic Number Theory; American Mathematical Society: Providence, RI, USA, 2004.
18. Bhowmik, G.; Grimmelt, L. The exceptional set of the Goldbach problem. Anal. Math. 2026. <https://doi.org/10.1007/s10476-026-00191-z>.
19. Helfgott, H.A.; Platt, D.J. Numerical verification of the ternary Goldbach conjecture up to 8.875×10^{30} . Exp. Math. 2013, 22, 406-409.
20. Davenport, H. Multiplicative Number Theory, 3rd ed.; revised by Montgomery, H.L.; Springer: New York, NY, USA, 2000.
21. Nathanson, M.B. Additive Number Theory: The Classical Bases; Springer: New York, NY, USA, 1996.
22. Granville, A. Harald Cramer and the distribution of prime numbers. Scand. Actuar. J. 1995, 1995, 12-28.
23. Crandall, R.; Pomerance, C. Prime Numbers: A Computational Perspective, 2nd ed.; Springer: New York, NY, USA, 2005.
24. Bahbouhi, B. Bahbouhi.b.goldbach.2026: An Adaptive Multilevel Algorithm for Constructing Goldbach Prime Pairs in Very Large Even Integers. Preprints 2026, 202609.1348.v1. <https://doi.org/10.20944/preprints202609.1348.v1>.
25. Grimmelt, L.; Teravainen, J. The Exceptional Set in Goldbach's Problem with two Chen Primes. arXiv 2025, arXiv:2508.16400; revised 2026.
26. Bombieri, E. On the large sieve. Mathematika 1965, 12, 201-225.

27. Vinogradov, A.I. The density hypothesis for Dirichlet L-series. *Izv. Akad. Nauk SSSR Ser. Mat.* 1965, 29, 903-934.
28. Selberg, A. On elementary methods in prime-number theory and their limitations. In *Collected Papers*; Springer: Berlin, Germany, 1989; Vol. I.
29. Halberstam, H.; Richert, H.-E. *Sieve Methods*; Academic Press: London, UK, 1974.
30. Tenenbaum, G. *Introduction to Analytic and Probabilistic Number Theory*, 3rd ed.; American Mathematical Society: Providence, RI, USA, 2015.

AI-Assistance Disclosure

AI-assisted tools were used in the preparation, editing, formatting, and computational discussion of this manuscript under the author's direction. The author remains responsible for the mathematical claims, numerical data, verification status, interpretation, and final submitted text.

Supplementary Material

A separate supplementary file contains the preserved numerical examples and empirical records used in the $P_{\min}$ and P/P studies, including status and audit notes where available.