

BOUSSOLE-Pmin: Predicting the First Goldbach Prime Pair from the Boundary of Very Large Even Integers

Bouchaib Bahbouhi
Independent Researcher, France

Abstract. For an even integer E , $P_{\min}(E)$ is defined as the smallest prime p for which $E-p$ is also prime. This paper studies $P_{\min}$ as the first Goldbach synchronization encountered from the small-prime boundary. A longitudinal BOUSSOLE series contains 100 studied scales of $E = 10^N + 2$ and reaches $E = 10^{4000} + 2$. A separate frozen predictive model was tested on 2000 new even integers $E = 10^N + c$, with $N = 20, 40, \dots, 400$ and $c = 1002, 1004, \dots, 1200$. The empirical scale $D(E) = \log(E) \log \log(E)$ was used with $Q_{50} = 0.258D$, $Q_{90} = 0.790D$ and $Q_{95} = 1.088D$. Observed coverage was 51.5%, 79.4% and 85.4%, respectively. The median $P_{\min}/D$ was 0.24893, close to the frozen median coefficient, while the upper tail was much heavier than predicted. BOUSSOLE therefore uses prediction as a search compass, not as a stopping rule. Finally, Goldbach is reformulated as the finiteness of $P_{\min}(E)$, and the future analytical target is uniform positivity of a boundary counting function $R(E, H)$. The computational evidence is not presented as a proof of Goldbach's conjecture.

Keywords: Goldbach conjecture; $P_{\min}$; minimal Goldbach partition; prime pairs; computational number theory; sieve; prediction

1. Introduction

The binary Goldbach conjecture states that every even integer greater than 2 is the sum of two primes. It remains open despite strong analytic results and extensive computation. Hardy and Littlewood developed the classical heuristic framework for Goldbach representations [1]. Chen proved that every sufficiently large even integer is the sum of a prime and a number with at most two prime factors [2]. Montgomery and Vaughan obtained major exceptional-set results [3]. Oliveira e Silva, Herzog and Pardi verified the even Goldbach conjecture for all even integers up to 4×10^{18} and also studied minimal Goldbach partitions [4].

The present work studies a simple boundary quantity. For each even E , we ask for the smallest prime p whose complement $E-p$ is also prime. This quantity is called $P_{\min}(E)$. The contribution is not the invention of the smallest component of a Goldbach partition; minimal Goldbach partitions have already been studied [4]. The contribution here is to use $P_{\min}$ as the central variable of a longitudinal, predictive and adaptive computational framework.

Two datasets are kept separate. The first is a 100-level longitudinal series $E = 10^N + 2$ extending to $N = 4000$. The second is a 2000-case validation cohort built from new integers $E = 10^N + c$. The predictive coefficients were frozen before validation and were not recalibrated during the experiment.

2. $P_{\min}$ and the FIRST Method

2.1 Definition

For an even integer E greater than 2, $P_{\min}(E)$ is the smallest prime p such that $q = E-p$ is prime. When it exists, $E = P_{\min}(E) + q$ is a Goldbach representation. The word smallest is essential: a successful pair is not a $P_{\min}$ result unless every smaller prime candidate has already been eliminated.

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

The ordered procedure is called FIRST. Prime candidates p are examined in increasing order. For each p , $q = E-p$ is computed. Cheap modular tests and small-prime sieving reject most composite q values. Strong probable-prime or primality testing is reserved for survivors. FIRST stops at the first successful q only when the preceding prime candidates have been exhausted.

2.2 Fifty longitudinal examples

For the family $E = 10^N + 2$, the first 50 longitudinal observations are listed compactly below as $N:P_{\min}$:

10:59, 20:13, 30:13, 40:19, 50:59, 60:163, 70:449, 80:13, 90:281, 100:1289.

110:71, 120:709, 130:43, 140:5, 150:593, 160:523, 170:631, 180:3221, 190:1451, 200:191.

210:173, 220:743, 230:1723, 240:2039, 250:2531, 260:5651, 270:863, 280:541, 290:4583, 300:71.

310:173, 320:181, 330:853, 340:599, 350:2099, 360:10891, 370:2011, 380:149, 390:71, 400:809.

410:61, 420:1201, 430:5653, 440:2593, 450:191, 460:2551, 470:701, 480:2029, 490:3631, 500:1039.

These observations show immediately that $P_{\min}$ is not monotone. For example, $P_{\min}(10^{290} + 2) = 4583$, whereas $P_{\min}(10^{300} + 2) = 71$. Likewise, $P_{\min}(10^{360} + 2) = 10891$ and $P_{\min}(10^{380} + 2) = 149$.

2.3 Successive even integers

A local experiment keeps the magnitude essentially fixed and changes E by 2. Around 10^{20} , the first eleven studied offsets 1002, 1004, ..., 1022 produced $P_{\min}$ values 41, 11, 13, 47, 17, 19, 53, 23, 277, 59 and 29. This shows that the FIRST distance can fluctuate strongly even when E changes only slightly. Some neighboring integers also share the same complementary prime q ; this is useful as a q -anchor diagnostic, but it does not replace the ordered FIRST audit.

3. Predictive Scale and Validation Protocol

The exact FIRST search answers what $P_{\min}$ is. A separate predictive layer asks how far the search is likely to travel. Motivated by prime-density considerations and by the reference data, the empirical characteristic scale retained for testing was:

$$D(E) = \log(E) \times \log \log(E), \text{ with natural logarithms.}$$

The normalized variable is $Z(E) = P_{\min}(E)/D(E)$. Before the independent validation began, three coefficients were frozen:

$$Q50 = 0.258D, Q90 = 0.790D, Q95 = 1.088D.$$

These labels were intended as predictive coverage levels, not mathematical bounds. The validation cohort used 20 values $N = 20, 40, \dots, 400$ and, at each level, 100 even offsets $c = 1002, 1004, \dots, 1200$. This produced exactly 2000 new even integers. No coefficient was changed after the search began.

For every E , D , $Q50$, $Q90$ and $Q95$ were fixed before comparison with the observed $P_{\min}$. Coverage indicators recorded whether $P_{\min}$ fell below each threshold. The experiment was designed to test the frozen model, including its failures.

4. Results

4.1 Independent 2000-case validation

The observed coverage rates were 51.5% for $Q50$, 79.4% for $Q90$ and 85.4% for $Q95$. Thus the central prediction was close to its nominal 50% target, while the upper predictive levels were under-calibrated. A total of 292 of the 2000 observations, or 14.6%, exceeded the frozen $Q95$ threshold.

Across the cohort, mean $P_{\min}$ was 1553.895 and median $P_{\min}$ was 607. The observed range was 3 to 29917. The largest absolute $P_{\min}$ occurred at $E = 10^{360} + 1198$, where $P_{\min} = 29917$ and $q = 10^{360} - 28719$.

For $Z = P_{\min}/D$, the mean was 0.54417 and the median was 0.24893. The empirical 90th, 95th and 99th percentiles were approximately 1.48979, 2.01189 and 4.40848. The largest observed Z was approximately 13.62671, at $E = 10^{80} + 1054$ with $P_{\min} = 13093$.

The closeness of the independent median 0.24893 to the frozen coefficient 0.258 supports D as a useful central scale in this cohort. The much larger upper quantiles show that the distribution has a heavy tail relative to the original prediction. Accordingly, $Q90$ and $Q95$ must not be treated as upper bounds.

4.2 Long-range observations to $10^{4000} + 2$

The separate longitudinal series extends far beyond the validation range. Selected observed values are: $N=800$, $P_{\min}=30763$; 900, 3469; 1000, 6569; 1100, 7211; 1200, 4129; 1300, 18041; 1400, 18251; 1500, 1283; 1600, 6491; 1700, 2281; 1800, 89; 1900, 28879; 2000, 25951; 2250, 9539; 2500, 9941; 2750, 3719; 3000, 26693; 3250, 25639; 3500, 29333; 3750, 16139; and 4000, 37643.

For $N=2250$ and $N=2500$, the minimality claim retains the audit qualification recorded in the project. At $N=4000$ the observed pair is $p = 37643$ and $q = 10^{4000} - 37641$; the complementary q was externally reported as prime, while the $P_{\min}$ interpretation also requires completeness of the preceding FIRST audit. These qualifications are kept explicit.

5. Final BOUSSOLE- $P_{\min}$ Algorithm

The final algorithm separates exact search, prediction and adaptive continuation.

Step 1 - Audit. Verify that E is even and compute inexpensive residue information.

Step 2 - Predict. Compute $D = \log(E) \log \log(E)$, then $Q50$, $Q90$ and $Q95$.

Step 3 - Generate in order. Examine prime p values in strictly increasing order.

Step 4 - Eliminate cheaply. For $q = E - p$, use congruences and small-prime sieving to reject composites before expensive testing.

Step 5 - Test survivors. Apply a strong probable-prime or primality test to the remaining q values.

Step 6 - FIRST. The first successful q determines $P_{\min}$ only if all smaller prime p values have been eliminated.

Step 7 - Adapt. If no hit occurs by Q_{50} , continue to Q_{90} ; then to Q_{95} . If Q_{95} fails, continue beyond it, for example in progressively enlarged windows. Q_{95} is never a stopping rule.

Step 8 - Diagnose deserts. Record destructive factors, survivor depth and local q -anchor behavior for unusually long FIRST searches.

Step 9 - Verify status. Distinguish composite, probable-prime and certified-prime outcomes, and verify $p+q=E$ exactly.

The compact logic is: $E \rightarrow$ audit $\rightarrow$ predictive windows $\rightarrow$ increasing prime $p \rightarrow$ modular elimination $\rightarrow$ sieve $\rightarrow$ strong test of $q \rightarrow$ FIRST $\rightarrow$ verification $\rightarrow P_{\min}$.

6. Discussion and a Route Toward Goldbach

The computational results do not prove Goldbach's conjecture and do not prove a universal bound $P_{\min}(E) = O(\log(E) \log \log(E))$. The validation itself demonstrates why such a claim would be unjustified: the upper tail is much heavier than the initial model predicted.

The relation between $P_{\min}$ and Goldbach is nevertheless exact. For an even integer E greater than 2, $P_{\min}(E)$ exists if and only if E has a Goldbach representation. Thus Goldbach can be restated as: $P_{\min}(E)$ is finite for every even E greater than 2.

To turn this reformulation into an analytical target, let $R(E, H)$ denote the number of primes $p \leq H$ for which $E - p$ is also prime. Then:

$$P_{\min}(E) \leq H \text{ if and only if } R(E, H) > 0.$$

A future proof route is therefore to find an explicit function $H(E)$ for which $R(E, H(E))$ is uniformly positive for every sufficiently large even E . The empirical scale $\log(E) \log \log(E)$ need not be the first theoretical target; a proof may initially require a much larger boundary window.

The main obstruction remains simultaneous primality. Chen's theorem reaches prime plus prime-or-semiprime [2], but Goldbach requires prime plus prime. Sieve methods efficiently control many small-factor failures, yet the parity barrier prevents a straightforward promotion of every deep survivor to a prime. The $P_{\min}$ viewpoint does not remove this difficulty; it localizes it into a boundary non-avoidance problem.

A useful proof program is therefore: first control the small-factor coverage of $E - p$; then understand the dispersed large-factor regime; and finally establish a positive lower bound for $R(E, H)$ uniformly in E . The computational FIRST deserts provide concrete examples of the configurations that such a proof must control.

The largest present experiments are limited by computation, not by the definition of $P_{\min}$. The algorithm can be stated for arbitrarily large E , but this does not mean that its empirical predictive calibration has been proved at arbitrary scale. Practical scalability and theoretical validity are different questions.

7. Conclusion

BOUSSOLE- $P_{\min}$ studies the first Goldbach synchronization from the small-prime boundary. A 100-level longitudinal series reaches $10^4 000 + 2$, while an independent 2000-case cohort tests a frozen predictive model. The central scale $D = \log(E) \log \log(E)$ gave a median prediction that transferred well to the new cohort, but the upper tail was substantially heavier than expected.

The final method therefore uses prediction only as a compass. FIRST remains the exact ordered search, and an exceeded predictive window causes adaptive continuation rather than termination. The theoretical value of the framework is the reduction of the future proof target to uniform positivity of $R(E, H)$. The transition from frequent computational success to a uniform theorem for every even E remains open.

Declaration of AI Assistance

OpenAI ChatGPT was used as an AI-assisted research and writing tool for computational scripting, organization, language editing and manuscript preparation. The human author directed the research program, supplied and reviewed the experimental record, evaluated the mathematical interpretations, and takes responsibility for the final content. The

article is therefore intended for an archive that accepts AI-assisted scholarly work.

Data Availability

The complete supplementary record prepared for this study contains the 100 longitudinal BOUSSOLE cases and the 2000 validation cases, with each studied number and its observed Pmin. The validation calculations used to report the aggregate coverage values are retained in the project data files.

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 Mathematica* 44 (1923), 1-70. DOI: 10.1007/BF02403921.
- [2] Chen, J.R. On the Representation of a Large Even Integer as the Sum of a Prime and the Product of at Most Two Primes. *Scientia Sinica* 16 (1973), 157-176.
- [3] Montgomery, H.L.; Vaughan, R.C. The Exceptional Set in Goldbach's Problem. *Acta Arithmetica* 27 (1975), 353-370. DOI: 10.4064/aa-27-1-353-370.
- [4] 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} . *Mathematics of Computation* 83 (2014), 2033-2060. DOI: 10.1090/S0025-5718-2013-02787-1.
- [5] 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. DOI: 10.20944/preprints202609.1348.v1.

Appendix A. Numerical Examples of Pmin

The entries below give only the studied even number E and its observed Pmin. This appendix is included because no separate supplementary file is assumed.

A.1. Longitudinal series: all 100 studied cases

Number	Pmin	Number	Pmin	Number	Pmin	Number	Pmin
10^10 + 2	59	10^260 + 2	5651	10^510 + 2	2699	10^760 + 2	4603
10^20 + 2	13	10^270 + 2	863	10^520 + 2	3109	10^770 + 2	5981
10^30 + 2	13	10^280 + 2	541	10^530 + 2	3191	10^780 + 2	1451
10^40 + 2	19	10^290 + 2	4583	10^540 + 2	571	10^790 + 2	5581
10^50 + 2	59	10^300 + 2	71	10^550 + 2	4201	10^800 + 2	30763
10^60 + 2	163	10^310 + 2	173	10^560 + 2	373	10^900 + 2	3469
10^70 + 2	449	10^320 + 2	181	10^570 + 2	4663	10^1000 + 2	6569
10^80 + 2	13	10^330 + 2	853	10^580 + 2	2081	10^1100 + 2	7211
10^90 + 2	281	10^340 + 2	599	10^590 + 2	3559	10^1200 + 2	4129
10^100 + 2	1289	10^350 + 2	2099	10^600 + 2	3359	10^1300 + 2	18041
10^110 + 2	71	10^360 + 2	10891	10^610 + 2	5113	10^1400 + 2	18251
10^120 + 2	709	10^370 + 2	2011	10^620 + 2	11173	10^1500 + 2	1283
10^130 + 2	43	10^380 + 2	149	10^630 + 2	4093	10^1600 + 2	6491
10^140 + 2	5	10^390 + 2	71	10^640 + 2	1913	10^1700 + 2	2281
10^150 + 2	593	10^400 + 2	809	10^650 + 2	1193	10^1800 + 2	89
10^160 + 2	523	10^410 + 2	61	10^660 + 2	3691	10^1900 + 2	28879
10^170 + 2	631	10^420 + 2	1201	10^670 + 2	79	10^2000 + 2	25951
10^180 + 2	3221	10^430 + 2	5653	10^680 + 2	1613	10^2250 + 2	9539
10^190 + 2	1451	10^440 + 2	2593	10^690 + 2	139	10^2500 + 2	9941
10^200 + 2	191	10^450 + 2	191	10^700 + 2	4603	10^2750 + 2	3719
10^210 + 2	173	10^460 + 2	2551	10^710 + 2	8819	10^3000 + 2	26693
10^220 + 2	743	10^470 + 2	701	10^720 + 2	16889	10^3250 + 2	25639
10^230 + 2	1723	10^480 + 2	2029	10^730 + 2	1583	10^3500 + 2	29333
10^240 + 2	2039	10^490 + 2	3631	10^740 + 2	3643	10^3750 + 2	16139
10^250 + 2	2531	10^500 + 2	1039	10^750 + 2	2029	10^4000 + 2	37643

Appendix A. Numerical Examples of Pmin (continued)

A.2. Independent validation cohort: representative examples (1/2)

Number	Pmin	Number	Pmin	Number	Pmin	Number	Pmin
10^20 + 1002	41	10^60 + 1104	593	10^120 + 1006	739	10^160 + 1108	2377
10^20 + 1022	29	10^60 + 1124	251	10^120 + 1026	3	10^160 + 1128	479
10^20 + 1042	241	10^60 + 1144	271	10^120 + 1046	23	10^160 + 1148	499
10^20 + 1062	101	10^60 + 1164	653	10^120 + 1066	43	10^160 + 1170	521
10^20 + 1082	11	10^60 + 1184	311	10^120 + 1086	509	10^160 + 1190	541
10^20 + 1102	31	10^80 + 1004	1031	10^120 + 1106	83	10^180 + 1010	41
10^20 + 1122	3	10^80 + 1024	1051	10^120 + 1128	1049	10^180 + 1030	61
10^20 + 1142	23	10^80 + 1044	233	10^120 + 1148	227	10^180 + 1050	5879
10^20 + 1162	43	10^80 + 1064	3163	10^120 + 1168	97	10^180 + 1070	101
10^20 + 1182	41	10^80 + 1086	5237	10^120 + 1188	1109	10^180 + 1090	379
10^40 + 1002	11	10^80 + 1106	17	10^140 + 1008	2309	10^180 + 1110	29
10^40 + 1022	31	10^80 + 1126	1153	10^140 + 1028	59	10^180 + 1130	419
10^40 + 1044	53	10^80 + 1146	233	10^140 + 1048	79	10^180 + 1150	181
10^40 + 1064	73	10^80 + 1166	8563	10^140 + 1068	2111	10^180 + 1170	89
10^40 + 1084	163	10^80 + 1186	97	10^140 + 1088	179	10^180 + 1190	109
10^40 + 1104	113	10^100 + 1006	739	10^140 + 1108	139	10^200 + 1010	211
10^40 + 1124	17	10^100 + 1026	1823	10^140 + 1128	6197	10^200 + 1030	673
10^40 + 1144	37	10^100 + 1046	97	10^140 + 1148	179	10^200 + 1050	251
10^40 + 1164	173	10^100 + 1066	3313	10^140 + 1168	199	10^200 + 1070	271
10^40 + 1184	193	10^100 + 1086	137	10^140 + 1188	5351	10^200 + 1090	463
10^60 + 1004	131	10^100 + 1106	157	10^160 + 1008	359	10^200 + 1110	311
10^60 + 1024	151	10^100 + 1126	859	10^160 + 1028	379	10^200 + 1130	331
10^60 + 1044	593	10^100 + 1146	197	10^160 + 1048	5737	10^200 + 1150	523
10^60 + 1064	5	10^100 + 1166	3413	10^160 + 1068	419	10^200 + 1170	4241
10^60 + 1084	211	10^100 + 1186	919	10^160 + 1088	439	10^200 + 1190	563

Appendix A. Numerical Examples of Pmin (continued)

A.2. Independent validation cohort: representative examples (2/2)

Number	Pmin	Number	Pmin	Number	Pmin	Number	Pmin
10^220 + 1012	19	10^260 + 1114	3361	10^320 + 1016	163	10^360 + 1118	1571
10^220 + 1032	149	10^260 + 1134	17	10^320 + 1036	337	10^360 + 1138	16519
10^220 + 1052	59	10^260 + 1154	37	10^320 + 1056	9833	10^360 + 1158	12101
10^220 + 1072	79	10^260 + 1174	6823	10^320 + 1076	223	10^360 + 1178	2441
10^220 + 1092	1931	10^260 + 1194	4337	10^320 + 1096	397	10^360 + 1198	29917
10^220 + 1112	229	10^280 + 1014	1553	10^320 + 1116	263	10^380 + 1018	631
10^220 + 1132	139	10^280 + 1034	797	10^320 + 1136	283	10^380 + 1038	1979
10^220 + 1152	269	10^280 + 1054	4933	10^320 + 1156	457	10^380 + 1058	1999
10^220 + 1172	179	10^280 + 1074	1061	10^320 + 1176	4073	10^380 + 1078	691
10^220 + 1192	199	10^280 + 1094	857	10^320 + 1196	1229	10^380 + 1098	2039
10^240 + 1012	1129	10^280 + 1114	877	10^340 + 1016	1613	10^380 + 1118	3931
10^240 + 1032	5981	10^280 + 1134	2273	10^340 + 1036	10267	10^380 + 1138	751
10^240 + 1052	3089	10^280 + 1154	1693	10^340 + 1056	24023	10^380 + 1158	2099
10^240 + 1072	2341	10^280 + 1174	937	10^340 + 1076	2339	10^380 + 1180	1327
10^240 + 1092	6359	10^280 + 1194	1181	10^340 + 1096	1693	10^380 + 1200	2141
10^240 + 1112	1229	10^300 + 1014	683	10^340 + 1116	8387	10^400 + 1020	1721
10^240 + 1132	1249	10^300 + 1034	311	10^340 + 1138	6829	10^400 + 1040	971
10^240 + 1152	6101	10^300 + 1054	3	10^340 + 1158	5	10^400 + 1060	991
10^240 + 1172	1289	10^300 + 1074	23	10^340 + 1178	1867	10^400 + 1080	3797
10^240 + 1192	3229	10^300 + 1096	373	10^340 + 1198	10429	10^400 + 1100	1031
10^260 + 1012	3259	10^300 + 1116	16547	10^360 + 1018	1471	10^400 + 1120	1051
10^260 + 1032	983	10^300 + 1136	1483	10^360 + 1038	509	10^400 + 1140	4817
10^260 + 1054	3301	10^300 + 1156	433	10^360 + 1058	1511	10^400 + 1160	1091
10^260 + 1074	947	10^300 + 1176	1523	10^360 + 1078	1531	10^400 + 1180	1693
10^260 + 1094	967	10^300 + 1196	809	10^360 + 1098	569	10^400 + 1200	1901