

A square-root error term for the Hardy-Littlewood twin prime count, with persistent sign oscillation

Ahmet Mat

Hypothesis · Mathematics · Analytic number theory

Signed 1 October 2026

Papex PXH-RDEJ26WS3AFF · version 1

Keywords: twin primes; Hardy-Littlewood conjecture; prime counting function; error term; sign changes; computational number theory

Claim

Let $\pi_2(x)$ count primes $p \leq x$ with $p+2$ also prime, let $C_2 = 0.66016181584686957\dots$ be the twin prime constant, $\text{li}_2(x) = \int_2^x dt/(\log t)^2$, and define the Hardy-Littlewood error $E(x) = \pi_2(x) - 2C_2\text{li}_2(x)$. Conjectured, in three parts: (H1) for every $\epsilon > 0$, $E(x) = O(x^{1/2+\epsilon})$ as $x \rightarrow \infty$; (H2) $E(x)$ changes sign infinitely often, and the number of sign changes in $[10^k, 10^{k+1})$ is unbounded in k , so the oscillation does not die out; (H3) the bound in H1 cannot be sharpened by a logarithmic factor of this size, i.e. $\limsup |E(x)| \log^2(x)/\sqrt{x} = \infty$. This is a conjecture about the ERROR TERM of the Hardy-Littlewood asymptotic. It is logically downstream of the twin prime conjecture and presupposes it; it does NOT prove, and makes no attempt to prove, that there are infinitely many twin primes.

Why it matters

H1 is the twin prime analogue of the Riemann-Hypothesis-strength error bound for $\pi(x)$, so pinning its truth or falsity locates twin primes within the same error-term framework that governs the primes themselves. H2 is the part with independent interest: for the single primes, $\pi(x) - \text{li}(x)$ is observed to be of one sign throughout every computed range, and Littlewood's theorem that it nevertheless changes sign infinitely often is famous precisely because computation does not see it. The twin prime error behaves oppositely in the computed range - it oscillates freely and increasingly, 51 sign changes over 10^4 to 10^{10} with the per-decade count rising 0, 0, 3, 13, 14, 21. If that contrast is real rather than an artefact of a short baseline, the twin prime error term is not simply a scaled copy of the prime error term, and a correct model of it has to explain the difference. H3 says where the truth sits: between $\sqrt{x}/\log^2(x)$ and $\sqrt{x}$.

What would falsify this?

H1 is falsified by exhibiting x with $|E(x)| > x^{0.55}$, or by showing $|E(x)|/\sqrt{x}$ grows without bound; the observed decade maxima FALL monotonically $0.1594 \rightarrow 0.0637$ across 10^4 - 10^{10} , so a sustained rise over three consecutive decades falsifies. H2 is falsified by showing $E(x)$ is eventually of one sign, or that the per-decade sign-change count is bounded; observed counts RISE 0, 0, 3, 13, 14, 21, so decay toward zero over 10^{10} - 10^{13} falsifies. H3 is falsified by showing $|E(x)| \log^2(x)/\sqrt{x}$ is bounded; observed decade maxima RISE 18.69 $\rightarrow$ 28.78, so flattening to a finite asymptote falsifies. Each part is independently falsifiable; H1 can hold while H2 and H3 fail.

Methodology

Segmented odd-only sieve of Eratosthenes over $[1, 10^{10}]$ in index space (segment 2^{26} , base primes to 10^5), recording every p with p and $p+2$ both prime, with segment-boundary pairs handled explicitly; 27,412,679 pairs found. $\pi_2(x)$ by binary search over the stored pair list. $\text{li}_2(x)$ by adaptive quadrature split at decade boundaries (epsabs 1e-10, epsrel 1e-12). E evaluated on 3,000 log-spaced points over $10^4 \leq x \leq 10^{10}$. Validation: the segmented sieve reproduces an independent single-array sieve to 10^9 exactly, and the decade counts 205; 1,224; 8,169; 58,980; 440,312; 3,424,506; 27,412,679 agree with the standard published values. Next step is an independent reimplementation and extension to 10^{12} - 10^{13} .

Expected outcomes

Extending the sieve to 10^{12} - 10^{13} is expected to show $|E|/\sqrt{x}$ continuing to fall below 0.06, $|E| \cdot \log^2(x)/\sqrt{x}$ continuing to rise past 29, and the per-decade sign-change count continuing to grow roughly linearly in the decade index. A result contradicting any of these three is the informative outcome and is the reason the predictions are stated numerically here before the computation is run.

Research setup

System or population Prime pairs $(p, p+2)$ with $p \leq x$. Computationally exhaustive over $10^4 \leq x \leq 10^{10}$; the conjecture is asymptotic in x .

Variables Independent: x . Dependent: $\pi_2(x)$ (exact integer count); $E(x) = \pi_2(x) - 2 \cdot C_2 \cdot \text{li}_2(x)$; normalisations $E(x)/\sqrt{x}$ and $E(x) \cdot \log^2(x)/\sqrt{x}$; per-decade sign-change count.

Parameters $C_2 = 0.6601618158468695739278121100145557784326233602847$ (twin prime constant, fixed). Sieve bound 10^{10} . Segment size 2^{26} . 3,000 log-spaced evaluation points. Quadrature epsabs 1e-10, epsrel 1e-12.

Observer or reference frame *Not applicable* — Pure mathematics: the objects are integers and a deterministic counting function, so there is no observer or reference frame.

Assumptions and scope Conditional on the Hardy-Littlewood twin prime asymptotic: the subtracted main term $2 \cdot C_2 \cdot \text{li}_2(x)$ presupposes it, and all three parts are vacuous unless there are infinitely many twin primes. Nothing in this record establishes either the asymptotic or the infinitude. Computational evidence spans six decades only; every asymptotic statement is extrapolation.

Units *Not applicable* — All quantities are dimensionless: integer counts and ratios of counts to real-valued analytic functions.

Tolerance or uncertainty Sieve counts are exact integers (zero tolerance). Numerical quadrature for $\text{li}_2(x)$ carries absolute error below 1e-9, which is 3-12 orders of magnitude below $|E(x)|$ across the evaluated range and therefore cannot affect the sign of E or the normalisation maxima.

Research mode Mixed

Known limitations Six decades is a short baseline for an asymptotic claim. Littlewood's theorem on sign changes of $\pi(x) - \text{li}(x)$ is the standing warning that computed ranges can mislead about exactly this kind of question, and is why H2 is a conjecture rather than a reading of the data. The first two decades (10^4 - 10^6) show zero sign changes, so H2 rests on four decades. li_2 is evaluated numerically, not in closed form. No model of the distribution of E is proposed and no comparison against random-matrix predictions is attempted. The computation is single-implementation; it has not been independently reproduced.

Files

- Twin prime error term: write-up and computed data (x up to 1e10) — Supplement · PXA-NZHPWK45MPZT v1 · 459.2 KB

The PDF states the conjecture and the evidence behind it; the CSV is the exact numerical output the figure and all quoted maxima are read from, so a reader can recheck every number without rerunning the sieve.

- twin_prime_error_data.csv · CSV · 271.2 KB · 0x2884...3628
- twin_prime_hypothesis.pdf · PDF · 188.0 KB · 0xba5d...eed

Preparation

Method Agent-prepared

Application operon-gateway 0.1.0 (self-reported)

Model claude-opus-5 (self-reported)

Origin Every field above: Agent-researched, human-accepted

References

- [1] G. H. Hardy, J. E. Littlewood. Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes. *Acta Mathematica*, 1923. <https://doi.org/10.1007/bf02403921>
Builds on
- [2] Richard P. Brent. Irregularities in the distribution of primes and twin primes. *Mathematics of Computation*, 1975. <https://doi.org/10.1090/s0025-5718-1975-0369287-1>
Compares with
- [3] Yitang Zhang. Bounded gaps between primes. *Annals of Mathematics*, 2014. <https://doi.org/10.4007/annals.2014.179.3.7>
Cites
- [4] James Maynard. Small gaps between primes. *Annals of Mathematics*, 2015. <https://doi.org/10.4007/annals.2015.181.1.7>
Cites
- [5] D. H. J. Polymath. Variants of the Selberg sieve, and bounded intervals containing many primes. 2014. arXiv:1407.4897
Cites

Generated by Papex from the signed record (research-record/v7). The signature covers the record's content, not this PDF file. Check it at <https://papex.org/verify/PXH-RDEJ26WS3AFF?version=1>. Cite as <https://papex.org/id/PXH-RDEJ26WS3AFF/v/1>.