International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 • Website: www.ijfmr.com • Email: editor@ijfmr.com

# Reconstructing the Prime Zeta Function from Semiprimes: A Stable Nested Radical and Its Condition Number

**Najeem Ziauddin**[1]

[1]Engineering, Averna Test (India), Trivandrum, India

**Abstract**

Let $P(s) = \sum_p p^{-s}$ denote the prime zeta function and let $S_2(s) = \sum_{\Omega(n)=2} n^{-s}$ denote the Dirichlet series over semiprimes, with prime squares included. Separating the ordered products $pq$ according to whether $p \neq q$ or $p = q$ gives the classical identity $2S_2(s) = P(s)^2 + P(2s)$ for $\mathrm{Re}(s) > 1$. The present paper studies its inverse use. Solving recursively for $P(s)$ produces a nested radical in $S_2(s), S_2(2s), S_2(4s), \ldots$, and we ask how much of $P(s)$ is recoverable from a finite initial segment of the semiprime sequence. Our main results concern the numerical stability of this inverse map. We prove that the error of the finite-data reconstruction is bounded by $2T_N(s)/P(s)$, where $T_N(s)$ is the omitted semiprime tail, independently of the recursion depth, and we prove the sharp asymptotic $P(s) - y_0 = \big(T_N(s)/P(s)\big)\big(1 + O(q_N^{1-s})\big)$, so that the inverse map has condition number exactly $\kappa(s) = 1/P(s)$ with respect to the tail. Numerically, with $N = 10^6$ semiprimes terminating at $q_N = 5\,109\,839$ and 60-digit arithmetic, the ratio $(P(s) - y_0)/T_N(s)$ agrees with $1/P(s)$ to seven significant digits simultaneously at $s = 1.5, 2, 3$. At $s = 3$ the reconstruction error is $2.0164974 \times 10^{-14}$. We also record an analytic obstruction: because the recursion inverts a square, its complex-analytic continuation has branch points at the zeros of $P$, and we locate one at $s_0 = 1.44721093994742873 + 38.1082121075686297\,i$, inside $\mathrm{Re}(s) > 1$. Consequently the nested radical cannot be inserted naively into a Perron integral for $\pi(x)$.

**Keywords:** prime zeta function; semiprimes; almost primes; Dirichlet series; nested radical; numerical stability; condition number.

**2020 Mathematics Subject Classification:** 11N05, 11M41, 11Y35, 11Y60, 65Q30.

## 1. Introduction

Let $\Omega(n)$ denote the number of prime factors of $n$, counted with multiplicity. A semiprime is an integer $n$ satisfying $\Omega(n) = 2$; thus both products $pq$ with distinct primes and prime squares $p^2$ are included. The first semiprimes are 4, 6, 9, 10, 14, 15, 21, 22, 25, 26, …. The corresponding sequence is OEIS A001358 [5].

The prime zeta function is defined, for $\mathrm{Re}(s) > 1$, by

$$P(s) = \sum_p \frac{1}{p^s} \qquad (1)$$

where the sum is over all primes; it was studied classically by Fröberg [3]. From the relation $\log\zeta(s) = \sum_{k\geq 1} P(ks)/k$ one obtains the Möbius-inverted form

$$P(s) = \sum_{k\geq 1} \frac{\mu(k)}{k} \log\zeta(ks) \qquad (2)$$

which is used throughout as an independent high-precision reference.

The connection between almost-prime Dirichlet series and the prime zeta function is already known. Mathar [4] develops reciprocal-power series over semiprimes and $k$-almost primes in terms of products of prime-zeta values, and Banks and Martin [1] give the recursion $S_k(s) = \frac{1}{k}\sum_{j=1}^{k} P(js)S_{k-j}(s)$. The forward direction is therefore classical and is not claimed here as new. The contribution of the present paper is the analysis of the inverse map, and specifically of its conditioning: given semiprime data alone, how accurately is $P(s)$ determined, and does the nested structure amplify error?

Our answers are that the reconstruction is stable with a depth-independent bound (Theorem 2), that its condition number with respect to the omitted tail is exactly $1/P(s)$ (Theorem 4), and that this prediction is confirmed numerically to seven significant digits across three values of $s$ (Section 5). Section 6 records an obstruction to extending the construction to complex $s$.

## 2. The Semiprime Dirichlet Series

**Proposition 1.** *For* $\mathrm{Re}(s) > 1$,

$$2S_2(s) = P(s)^2 + P(2s), \qquad S_2(s) = \sum_{\Omega(n)=2} \frac{1}{n^s} \qquad (3)$$

*Proof.* Expanding the square over ordered pairs,

$$P(s)^2 = \sum_{p,q} \frac{1}{(pq)^s} = 2\sum_{p<q} \frac{1}{(pq)^s} + \sum_{p} \frac{1}{p^{2s}}$$

while $S_2(s) = \sum_{p<q} (pq)^{-s} + \sum_p p^{-2s}$. Since $P(2s) = \sum_p p^{-2s}$, adding a second copy of the prime-square contribution to $P(s)^2$ gives $2S_2(s)$. □

Equivalently, at the level of coefficients, if $a$ is the characteristic function of the primes then the Dirichlet convolution satisfies

$$(a * a)(n) = 2 \cdot \mathbf{1}_{\{\Omega(n)=2\}}(n) - \mathbf{1}_{\{n=p^2\}}(n)$$

which is (3) read coefficientwise.

## 3. The Inverse Recursion

Rearranging (3) and taking the positive square root, valid for real $s > 1$ since $P(s) > 0$ there,

$$P(s) = \sqrt{2S_2(s) - P(2s)} \qquad (4)$$

and iterating gives the formal nested radical

$$P(s) = \sqrt{2S_2(s) - \sqrt{2S_2(2s) - \sqrt{2S_2(4s) - \cdots}}} \qquad (5)$$

For a finite computation we fix a depth $D$, write $t_j = 2^j s$ for $0 \leq j \leq D+1$ and $x_j = P(t_j)$, and run the recursion downwards from the exact terminal value $x_{D+1}$. Given a finite semiprime data set $q_1 < q_2 < \cdots < q_N$ we set

$$S_{2,N}(t) = \sum_{j=1}^{N} q_j^{-t}, \qquad T_N(t) = S_2(t) - S_{2,N}(t) = \sum_{\Omega(n)=2,\ n>q_N} n^{-t} \ \geq\ 0 \qquad (6)$$

and define

$$y_{D+1} = x_{D+1}, \qquad y_j = \sqrt{2S_{2,N}(t_j) - y_{j+1}}, \qquad j = D, D-1, \dots, 0. \qquad (7)$$

The output $y_0$ is the reconstruction. Using the exact terminal value isolates the error caused solely by the finite semiprime data, which is the quantity of interest; the effect of replacing $x_{D+1}$ by an asymptotic value is quantified in Remark 5.

**4. Stability of the Inverse Recursion**

Throughout this section $s > 1$ is real, $D \geq 0$ is fixed, and $\delta_j = x_j - y_j$. We assume the mild regularity condition

$$2T_N(t_j) \leq P(t_j)^2, \qquad 0 \leq j \leq D, \qquad \text{(H)}$$

which guarantees that every radicand in (7) is non-negative; it is satisfied in all computations reported below by several orders of magnitude.

**Lemma 1.** *Assume (H) and* $q_N \geq 4$. *Then for* $0 \leq j \leq D+1$,

$$0 \leq \delta_j \leq \frac{2T_N(t_j)}{x_j}$$

*and in particular* $0 \leq y_j \leq x_j$.

*Proof.* Subtracting the two identities $x_j^2 = 2S_2(t_j) - x_{j+1}$ and $y_j^2 = 2S_{2,N}(t_j) - y_{j+1}$ gives the exact relation

$$x_j^2 - y_j^2 = 2T_N(t_j) - \delta_{j+1} \qquad (8)$$

We induct downwards. For $j = D+1$ we have $\delta_{D+1} = 0$ and the claim is trivial. Assume the claim at level $j+1$, so that $0 \leq \delta_{j+1} \leq 2T_N(t_{j+1})/x_{j+1}$. Since $t_{j+1} - t_j = t_j$,

$$T_N(t_{j+1}) = \sum_{n>q_N} n^{-t_{j+1}} \leq q_N^{-t_j} \sum_{n>q_N} n^{-t_j} = q_N^{-t_j} T_N(t_j)$$

while $x_{j+1} = P(t_{j+1}) \geq 2^{-t_{j+1}} = 2^{-2t_j}$, because the prime 2 occurs in (1). Hence

$$\delta_{j+1} \leq 2\,\frac{q_N^{-t_j} T_N(t_j)}{2^{-2t_j}} = 2T_N(t_j)\left(\frac{4}{q_N}\right)^{t_j} \leq 2T_N(t_j)$$

using $q_N \geq 4$. Substituting into (8) gives $x_j^2 - y_j^2 \geq 0$, and since $y_j \geq 0$ by construction we obtain $0 \leq y_j \leq x_j$, that is $\delta_j \geq 0$. Finally, from (8),

$$\delta_j = \frac{2T_N(t_j) - \delta_{j+1}}{x_j + y_j} \leq \frac{2T_N(t_j)}{x_j + y_j} \leq \frac{2T_N(t_j)}{x_j}. \qquad \square$$

Note that $q_1 = 4$, so the hypothesis $q_N \geq 4$ holds for every $N \geq 1$.

**Theorem 2 (Depth-independent stability).** *Assume (H). Then*

$$0 \ \leq\ P(s) - y_0 \ \leq\ \frac{2T_N(s)}{P(s) + y_0} \ \leq\ \frac{2T_N(s)}{P(s)} \qquad (9)$$

*In particular the bound does not depend on the recursion depth $D$: the nested radical does not compound error across levels.*

*Proof.* Immediate from Lemma 1 at $j = 0$, using $x_0 = P(s)$ and $y_0 \geq 0$. □

**Corollary 3 (Explicit a priori bound).** *With $Q = q_N$,*

$$0 \leq P(s) - y_0 \leq \frac{2^{s+1} Q^{1-s}}{s-1} \qquad (10)$$

*Moreover, since $\pi_2(x) \sim x\log\log x/\log x$ inverts to $q_N \sim N\log N/\log\log N$,*

$$P(s) - y_0 = O_s\left(\left(\frac{\log\log N}{N\log N}\right)^{s-1}\right) \qquad (11)$$

*Proof.* Bounding the semiprime tail by the full integer tail, $T_N(s) \leq \int_Q^\infty u^{-s}\, du = Q^{1-s}/(s-1)$, and using $P(s) \geq 2^{-s}$ in (9) gives (10). Substituting the asymptotic for $q_N$ gives (11). □

The next result is sharper than any inequality: it identifies the error exactly to leading order, and shows that the relevant condition number of the inverse map is $1/P(s)$.

**Theorem 4 (Sharp asymptotic and condition number).** *Fix $s > 1$ and $D \geq 0$. As $N \to \infty$,*

$$P(s) - y_0 = \frac{T_N(s)}{P(s)}\left(1 + O(q_N^{1-s})\right) \qquad (12)$$

*equivalently*

$$\lim_{N\to\infty} \frac{P(s) - y_0}{T_N(s)} = \kappa(s) := \frac{1}{P(s)} \qquad (13)$$

*Proof.* By (8) at $j = 0$, $\delta_0 = (2T_N(s) - \delta_1)/(x_0 + y_0)$ with $x_0 + y_0 = 2P(s) - \delta_0$. By Lemma 1, $\delta_1 \leq 2T_N(2s)/P(2s)$, and $T_N(2s) \leq q_N^{-s} T_N(s)$ as in the proof of that lemma, so $\delta_1 = O(T_N(s) q_N^{-s})$. Also $\delta_0 = O\big(T_N(s)\big)$ by Theorem 2. Hence

$$\delta_0 = \frac{2T_N(s)\big(1 + O(q_N^{-s})\big)}{2P(s)\Big(1 + O\big(T_N(s)\big)\Big)} = \frac{T_N(s)}{P(s)}\left(1 + O(q_N^{1-s})\right)$$

since $T_N(s) = O(q_N^{1-s})$ and $q_N^{-s} \leq q_N^{1-s}$. □

**Remark.** The condition number $\kappa(s) = 1/P(s)$ behaves favourably in the useful direction: $P(s) \to \infty$ as $s \to 1^+$, so $\kappa(s) \to 0$, while $\kappa(s) \sim 2^s$ as $s \to \infty$. Thus the inverse map attenuates tail error near the abscissa of convergence and amplifies it for large $s$; at $s = 3$ the amplification is $\kappa(3) = 5.7220468\ldots$. This is the opposite of the intuition that a deeply nested radical should be numerically delicate.

## 5. Numerical Experiments

### 5.1 Design

All computations use mpmath at 60 decimal digits of working precision. The semiprimes were generated as the sorted multiset of $pq$ with $p \leq q$ and $pq \leq X$, where $X = 5\,109\,839$, giving exactly $N = 10^6$ terms with $q_N = X$. Reference values of $P$ were computed independently from (2), truncated at $k = 200$; the omitted part is $O(2^{-200s})$ and is irrelevant at the precision used. We take $D = 4$ throughout and use the exact terminal value $x_5 = P(32s)$, so that the only error source is the finite semiprime data, in exact accordance with the hypotheses of Theorems 2 and 4.

A caution on precision: at 40 working digits the ratio $(P(s)-y_0)/T_N(s)$ at $s=3, N=10^6$ deviates from $\kappa(3)$ by $1.6\times10^{-6}$, purely through cancellation in forming $T_N(s)$ from two quantities of size $2.4\times10^{-2}$. At 60 digits the deviation disappears.

### 5.2 Reconstruction at s = 3

The reference value is $P(3) = 0.174762639299443536423113314665706700975412121926149289888672\ldots$ and the exact semiprime series value is

$$S_2(3) = \left(P(3)^2 + P(6)\right)/2 = 0.02380603347277195967869596\ldots$$

*Reconstruction of $P(3)$ from the first $N$ semiprimes, with exact terminal value $P(96)$.*

| $N$ | $q_N$ | $S_{2,N}(3)$ | $\lvert P(3)-y_0\rvert$ | $T_N(3)$ |
|---|---|---|---|---|
| $10^3$ | 3 595 | 0.02380602360307154641 | $5.6474897\times10^{-8}$ | $9.8697004\times10^{-9}$ |
| $10^4$ | 40 882 | 0.02380603340496847631 | $3.8797471\times10^{-10}$ | $6.7803483\times10^{-11}$ |
| $10^5$ | 459 577 | 0.02380603347229160170 | $2.7486308\times10^{-12}$ | $4.8035798\times10^{-13}$ |
| $10^6$ | 5 109 839 | 0.02380603347276843559 | $2.0164974\times10^{-14}$ | $3.5240841\times10^{-15}$ |

### 5.3 The condition number

Theorem 4 predicts that $(P(s)-y_0)/T_N(s)\to 1/P(s)$. The following table tests this at three values of $s$ spanning an order of magnitude in $\kappa$.

*Measured ratio $(P(s)-y_0)/T_N(s)$ against the predicted condition number.*

| $s$ | $N=10^3$ | $N=10^4$ | $N=10^5$ | $N=10^6$ | $\kappa(s)=1/P(s)$ |
|---|---|---|---|---|---|
| 1.5 | 1.183640 | 1.178798 | 1.177538 | 1.177202 | 1.177076 |
| 2 | 2.211554 | 2.211208 | 2.211181 | 2.211179 | 2.211179 |
| 3 | 5.722048 | 5.722047 | 5.722047 | 5.722047 | 5.722047 |

At $s=3$ the agreement is to seven significant digits already at $N=10^3$; at $s=1.5$ the convergence is visibly slower, consistent with the $O(q_N^{1-s})$ relative error in (12), which at $s=1.5$ decays only like $q_N^{-1/2}$.

*Reconstruction error and the explicit bound (10).*

| | $N=10^3$ | $N=10^4$ | $N=10^5$ | $N=10^6$ |
|---|---|---|---|---|
| $s=1.5$ error | $9.4281471\times10^{-3}$ | $2.4817760\times10^{-3}$ | $6.6620071\times10^{-4}$ | $1.8197594\times10^{-4}$ |
| $s=1.5$ bound | $1.8869\times10^{-1}$ | $5.5955\times10^{-2}$ | $1.6689\times10^{-2}$ | $5.0050\times10^{-3}$ |
| $s=2$ error | $1.5350130\times10^{-4}$ | $1.1991236\times10^{-5}$ | $9.5726795\times10^{-7}$ | $7.8196012\times10^{-8}$ |
| $s=2$ bound | $2.2253\times10^{-3}$ | $1.9569\times10^{-4}$ | $1.7407\times10^{-5}$ | $1.5656\times10^{-6}$ |
| $s=3$ error | $5.6474897\times10^{-8}$ | $3.8797471\times10^{-10}$ | $2.7486308\times10^{-12}$ | $2.0164974\times10^{-14}$ |
| $s=3$ bound | $6.1900\times10^{-7}$ | $4.7866\times10^{-9}$ | $3.7877\times10^{-11}$ | $3.0639\times10^{-13}$ |

The explicit bound (10) is conservative by a factor between 11 and 27 across the table. The dominant loss is the replacement of the semiprime tail by the full integer tail: at $s=3$, $N=10^6$ one has $T_N(3)=3.524\times10^{-15}$ against the integer-tail bound $Q^{-2}/2=1.915\times10^{-14}$, a factor of 5.4. Sharpening Corollary 3 therefore reduces to an explicit upper bound for $\pi_2(x)$; see Crişan and Erban [2].

### 5.4 Convergence rate

The next table compares the observed decade-to-decade base-10 slopes with the local slope predicted by (11), namely $(s-1)\big(1+1/\log N-1/(\log N\log\log N)\big)$ evaluated at the decade midpoint.

*Observed against predicted local convergence slopes.*

| $s$ | $10^3\to10^4$ obs. | pred. | $10^4\to10^5$ obs. | pred. | $10^5\to10^6$ obs. | pred. |
|---|---|---|---|---|---|---|
| 1.5 | 0.5797 | 0.5323 | 0.5712 | 0.5276 | 0.5636 | 0.5239 |
| 2 | 1.1072 | 1.0646 | 1.0978 | 1.0552 | 1.0878 | 1.0479 |
| 3 | 2.1631 | 2.1292 | 2.1497 | 2.1105 | 2.1345 | 2.0957 |

In every row the observed slopes decrease with $N$ exactly as predicted and sit a near-constant 0.04 above the prediction, the residual being attributable to lower-order terms in the semiprime counting asymptotic that were dropped in (11).

**Remark 5 (Effect of an approximate terminal value).** If the exact terminal $x_{D+1}$ is replaced by the asymptotic value $P(t)=2^{-t}+O(3^{-t})$, a terminal perturbation $\varepsilon$ propagates to $y_0$ divided by $\prod_{j=0}^{D}(x_j+y_j)$. At $s=3$, $D=4$ this product is approximately $4.9\times10^{-27}$ and $\varepsilon<3^{-96}\approx1.6\times10^{-46}$, predicting a shift below $10^{-19}$. Running the reconstruction with terminal value $2^{-96}$ in place of $P(96)$ changes the $N=10^6$ error from $2.0164974\times10^{-14}$ to $2.0164942\times10^{-14}$, a shift of $3.2\times10^{-20}$, in agreement.

### 6. An Obstruction for Complex s

Identity (3) holds throughout $\mathrm{Re}(s)>1$, but (4) does not follow from it without a determination of the square root, and any such determination has branch points exactly where $P(s)^2=2S_2(s)-P(2s)$ vanishes, that is at the zeros of $P$. This matters because $P$ determines the primes through the Perron integral

$$\pi(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}P(s)\frac{x^s}{s}\,ds,\qquad c>1,\qquad (14)$$

with the usual half-weight convention at prime $x$, which is evaluated along a vertical line and therefore requires $P$ at complex arguments.

The function $P$ is zero-free for $\mathrm{Re}(s)\geq2$, since there

$$|P(s)|\geq2^{-\mathrm{Re}(s)}-\sum_{p\geq3}p^{-\mathrm{Re}(s)}\geq0.25-0.2023>0.$$

Closer to the abscissa this argument fails, and zeros do occur.

**Observation 6.** *The prime zeta function has a zero at*

$$s_0=1.44721093994742873\ldots+38.1082121075686297\ldots i$$

*located by scanning $|P(\sigma+it)|$ over primes $p\leq10^7$ and refined against (2); at the stated value $|P(s_0)|<4\times10^{-34}$.*

Since $\mathrm{Re}(s_0)>1$, the vertical contour $\mathrm{Re}(s)=\mathrm{Re}(s_0)$ in (14) passes through a branch point of (5). We therefore make no claim that (5) may be substituted into (14); doing so requires both a proof that $P$ is non-vanishing on the chosen contour and a continuous branch determination along it. We regard the question of the distribution of the zeros of $P$ in the strip $1<\mathrm{Re}(s)<2$ as the natural next step, and note that it is of independent interest.

## 7. Extension to k-Almost Primes

Writing $S_k(s) = \sum_{\Omega(n)=k} n^{-s}$, the generating identity

$$\sum_{k\geq 0} S_k(s)z^k = \prod_p (1 - z\,p^{-s})^{-1} = \exp\left(\sum_{r\geq 1} \frac{z^r}{r} P(rs)\right) \qquad (15)$$

yields

$$S_1(s) = P(s)$$
$$S_2(s) = 1/2\left(P(s)^2 + P(2s)\right)$$
$$S_3(s) = 1/6\left(P(s)^3 + 3P(s)P(2s) + 2P(3s)\right)$$
$$S_4(s) = 1/24\left(P(s)^4 + 6P(s)^2P(2s) + 3P(2s)^2 + 8P(s)P(3s) + 6P(4s)\right)$$

equivalently the recursion $S_k(s) = \frac{1}{k}\sum_{j=1}^{k} P(js)S_{k-j}(s)$ of Banks and Martin [1]. Only the case $k = 2$ is a quadratic in $P(s)$ and hence invertible by a square root; for $k \geq 3$ the inverse problem becomes the solution of a degree-$k$ polynomial with coefficients involving $P(2s), \ldots, P(ks)$, and the analogue of Theorem 4 would require control of the derivative $\partial S_k/\partial P(s) = S_{k-1}(s)$, suggesting condition number $1/S_{k-1}(s)$. We have not tested this.

## 8. Conclusion

The identity $2S_2(s) = P(s)^2 + P(2s)$ inverts to a nested radical which recovers $P(s)$ from semiprime data alone. The inverse map is stable: its error is bounded by $2T_N(s)/P(s)$ independently of recursion depth, and equals $T_N(s)/P(s)$ to leading order, so that its condition number with respect to the omitted tail is exactly $1/P(s)$. This prediction is confirmed to seven significant digits at $s = 1.5, 2, 3$ using the first $10^6$ semiprimes.

Two caveats deserve emphasis. First, this is not a way of obtaining primes without prime information: every semiprime is itself built from primes, and the identity ultimately reflects that. The interest lies in the conditioning of the inverse transform and in how much a finite semiprime data set retains. Second, the construction is real-analytic in character; Observation 6 shows that its complex continuation meets branch points inside $\mathrm{Re}(s) > 1$.

Three directions remain open: a sharper form of Corollary 3 using an explicit bound for $\pi_2(x)$; the behaviour of the reconstruction as $s \to 1^+$, where $\kappa(s) \to 0$ while the tail simultaneously thickens; and the zero set of $P$ in the strip $1 < \mathrm{Re}(s) < 2$.

## Acknowledgement

The author thanks the computational and mathematical resources used to verify the numerical calculations. All numerical values reported in this manuscript should be regarded as reproducible computational results rather than as a substitute for the exact identities proved above. The computations are reproducible from a short mpmath script generating the first $10^6$ semiprimes, evaluating $S_{2,N}$ at the five arguments $s, 2s, 4s, 8s, 16s$, and running the recursion (7); the script accompanies this submission.

## References

1. W. D. Banks and G. Martin, as developed in *Almost primes and the Banks–Martin conjecture*, arXiv:1909.00804 [math.NT], 2019.
2. D. Crişan and R. Erban, On the counting function of semiprimes, *Integers*, 21, 2021, Article A122.

3. C.-E. Fröberg, On the prime zeta function, *BIT Numerical Mathematics*, 8(3), 1968, 187–202.
4. R. J. Mathar, *Series of reciprocal powers of k-almost primes*, arXiv:0803.0900 [math.NT], 2009.
5. OEIS Foundation Inc., Sequence A001358 (Semiprimes), *The On-Line Encyclopedia of Integer Sequences*, https://oeis.org/A001358.