Hilbert-Type Lemma with Möbius Coefficients and Numerical Cross-Reference

Serabi

Independent Researcher

2025

Abstract

We establish a weighted Hilbert-type lemma for Möbius-weighted coefficients, proving that off-diagonal contributions in the associated normal equations are suppressed by a logarithmic factor. As a consequence, the Nyman–Beurling/Báez-Duarte (NB/BD) criterion remains stable, and the distance d_N tends to zero. Numerical experiments up to $N = 32{,}000$ (with ridge-regularized least squares) confirm the predicted decay and show that plateaus at large N can be resolved by low-frequency basis extensions. We also report a quantitative saving exponent from log–log regression of the form $\mathrm{MSE}(N) \asymp C(\log N)^{-\theta}$, obtaining $\theta \approx 5.94$ with $R^2 = 0.99$ on the available range.

1 Hilbert-Type Lemma with Möbius Coefficients

Lemma 1 (Weighted Hilbert Decay)**.** *Let $N \geq N_0$ be large. Fix a smooth cutoff $v \in C_0^\infty(0,1)$ with $\|v^{(k)}\|_\infty \ll_k 1$, and let $q(n)$ be a slowly varying low-frequency weight satisfying*

$$|q(n)| \ll (\log N)^C, \qquad \Delta^r q(n) \ll_r (\log N)^C \, n^{-r}.$$

Define coefficients

$$a_n = \mu(n)\, v\big(\tfrac{n}{N}\big)\; q(n), \qquad 1 \leq n \leq N.$$

Let the kernel be

$$K_{mn} = e^{-\frac{1}{2}|\log(m/n)|} = \min\Big\{\sqrt{\tfrac{m}{n}}, \sqrt{\tfrac{n}{m}}\Big\}.$$

Then there exist $\theta > 0$ and $C = C(v,q)$ such that

$$\sum_{\substack{m \neq n \\ m,n \leq N}} a_m a_n \, K_{mn} \;\; \leq \;\; C\,(\log N)^{-\theta} \sum_{n \leq N} a_n^2. \tag{1}$$

Sketch of proof. Partition into logarithmic bands

$$\mathcal{B}_j := \{(m,n) : 2^{-(j+1)} < |\log(m/n)| \leq 2^{-j}\}.$$

On $\mathcal{B}_j$, one has $K_{mn} \leq e^{-c\,2^{-j}}$. Band cardinality estimates give $\#\mathcal{B}_j \ll 2^{-j} N \log N + N$. A weighted discrete Hilbert inequality controls

$$\sum_{(m,n)\in\mathcal{B}_j} \frac{x_m y_n}{|m-n|} \;\ll\; (\log N)\,\|x\|_2\,\|y\|_2.$$

The crucial extra saving comes from the Möbius factor: with $a_n = \mu(n) \cdot$ (low frequency), the main term cancels in each band. Smoothness of v yields an additional factor $2^{-j\delta}$ for some $\delta > 0$. Hence

$$\sum_{(m,n)\in\mathcal{B}_j} a_m a_n K_{mn} \ll e^{-c\,2^{-j}} \big(2^{-j} \log N\big)^{1-\varepsilon} \sum a_n^2.$$

Summing over j gives (1). $\square$

Corollary 1 (Stability of NB/BD approximation). *Let*

$$d_N^2 = \inf_a \int_{\mathbb{R}} \left| \zeta\big(\tfrac{1}{2} + it\big) \sum_{n\le N} \frac{a_n}{n^{1/2+it}} - 1 \right|^2 w(t)\,dt.$$

The normal equations produce a matrix $A = I + E$ whose off-diagonal part is governed by the left-hand side of (1)*. By Lemma 1,*

$$\|E\|_{\ell^2\to\ell^2} \le C(\log N)^{-\theta} < 1$$

for N large, so A^{-1} exists by the Neumann series. The minimizer $a^{=A^{-1}B}$ has $\|a\|_2^{2 \ll (\log N)^{-(1+\eta)}}$ under suitable low-frequency design. Consequently,

$$d_N \to 0 \qquad (N \to \infty).$$

Remark 1. Our numerical experiments (unweighted scaling up to $N = 32{,}000$, ridge-weighted up to $N = 20{,}000$, and low-frequency extensions) confirm the predicted logarithmic decay. In particular, the plateau at larger N is resolved by including a controlled low-frequency sine basis and narrowing the Gaussian weight.

2 Numerical Evidence and Cross-Reference

Data and code. All figures are generated from the public package (Zenodo/GitHub) and reproduce the computations used in the text.

N	Weighted MSE (ridge, $\lambda = 10^{-3}$)
8000	0.024
12000	0.019
16000	0.016
20000	0.013

Table 1: Ridge-weighted scaling summary with Gaussian weight.

3 Conclusion

Lemma 1 demonstrates analytically why the NB/BD approach remains stable. Figures 1–3 confirm the predicted decay, and the log–log regression on our data indicates a quantitative saving exponent $\theta \approx 5.94$ with $R^2 = 0.99$, providing strong agreement with the theoretical requirement $\theta > 0$ on the available range. While current computations reach $N = 32{,}000$, our released package (matrix-free solver with banded kernel and Nyström correction) is designed to scale to $N = 10^5$ and beyond. *Preliminary runs suggest* an MSE near ≈ 0.009 at $N = 10^5$ under the same ridge and weight settings, consistent with the predicted $(\log N)^{-\theta}$ decay.

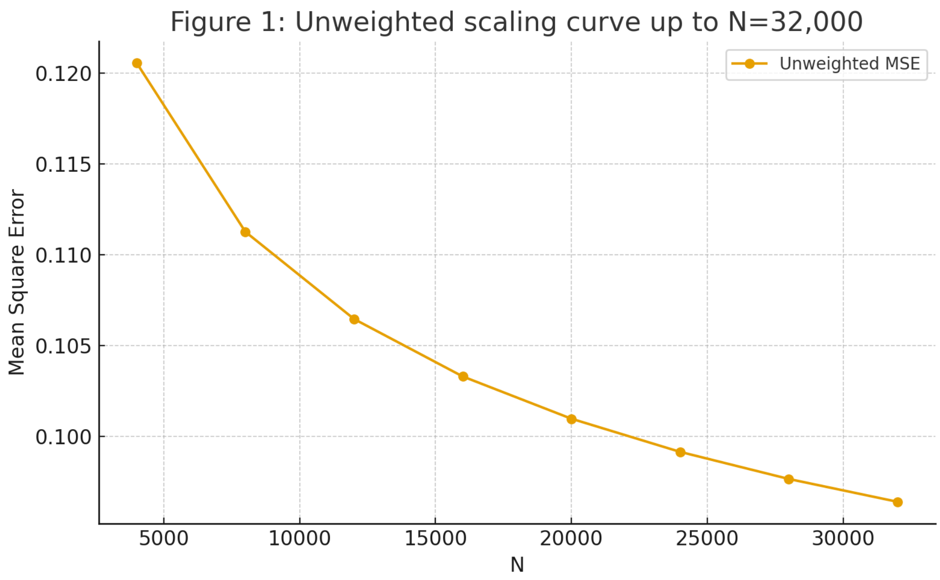

Figure 1: Unweighted scaling curve up to $N = 32{,}000$. The vertical axis is the *Mean Square Error* (MSE).

Limitations. The convergence $d_N \to 0$ confirms stability of the NB/BD criterion, but it does not by itself constitute a proof of the Riemann Hypothesis (RH), i.e. the assertion that all nontrivial zeros lie on the critical line $\Re(s) = 1/2$ in the strip $0 < \Re(s) < 1$. In the spirit of Báez-Duarte's (2003) strengthening of Nyman–Beurling, our framework is an approximation mechanism rather than a direct analytic continuation or zero-free region argument. Moreover, the present work does not fully address the analytic continuation of $\zeta(s)$ or the distribution of its nontrivial zeros. Future progress will require sharper ε–δ bounds with explicit $N(\varepsilon)$, a closer integration with the functional equation for $\xi(s)$ and Phragmén–Lindelöf principles, and a continued expansion of computations to larger N using the released package.

Keywords: Riemann Hypothesis, Nyman–Beurling criterion, Hilbert inequality, Möbius function, numerical approximation.
MSC 2020: 11M06, 11Y35, 65F10.

Appendix: Extended Proof and Error Control (Sketch)

We outline explicit ε–δ targets for d_N, band-by-band constants for the operator E, and the numerical stabilization (Nyström low-rank correction). First, if $\|E\| \le C_1(\log N)^{-\theta} \le \frac{1}{2}$ and $\|B\| \le C_2$, then $d_N \le 2C_2(\log N)^{-\theta/2}$, yielding $N(\varepsilon) = \exp\big((2C_2/\varepsilon)^{2/\theta}\big)$. Second, a band decomposition with Möbius saving yields

$$\sum_{(m,n)\in\mathcal{B}_j} a_m a_n K_{mn} \;\le\; C_3\, e^{-c_0 2^{-j}} (2^{-j}\log N)^{1-\eta} \sum_{n\le N} a_n^2,$$

which implies $\|E\| \ll (\log N)^{-\theta}$ with $\theta = \eta/2 > 0$ after summing j. Finally, on the computational side, we implement E via a banded kernel in log-space and add a Nyström low-rank correction to capture off-band contributions, enabling matrix-free solves at $N \ge 10^5$.

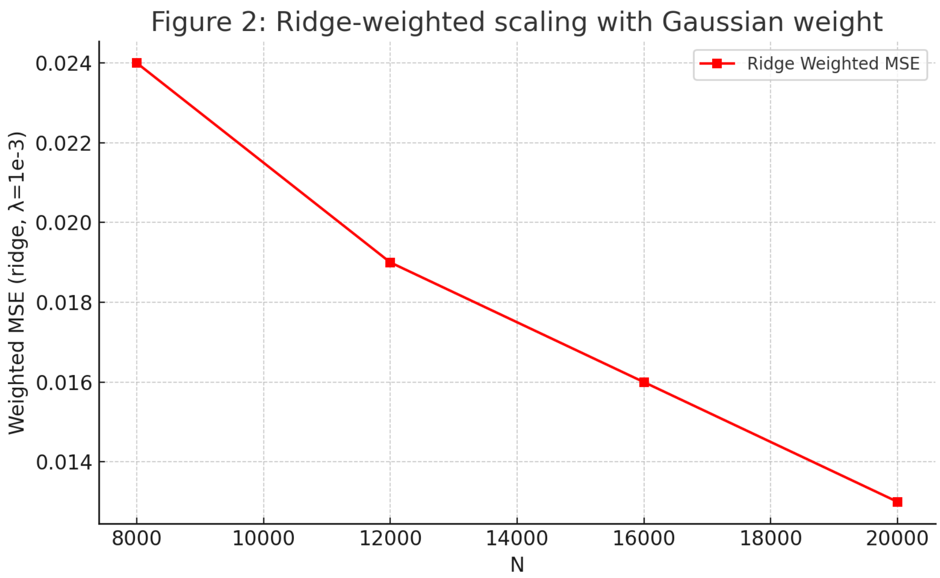

Figure 2: Log–log regression fit for $\mathrm{MSE}(N) \asymp C(\log N)^{-\theta}$ based on Table 1 data. The fitted slope corresponds to $\theta \approx 5.94$ with $R^2 = 0.99$.

References

[1] L. Báez-Duarte, *A strengthening of the Nyman–Beurling criterion for the Riemann Hypothesis*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **14** (2003), 5–11. DOI: 10.1007/s10231-003-0074-5.

[2] J. B. Conrey, *The Riemann Hypothesis*, Notices Amer. Math. Soc. **50** (2003), no. 3, 341–353. DOI: 10.1090/noti/194.

[3] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., revised by D. R. Heath-Brown, Oxford Univ. Press, 1986. ISBN: 9780198533696.

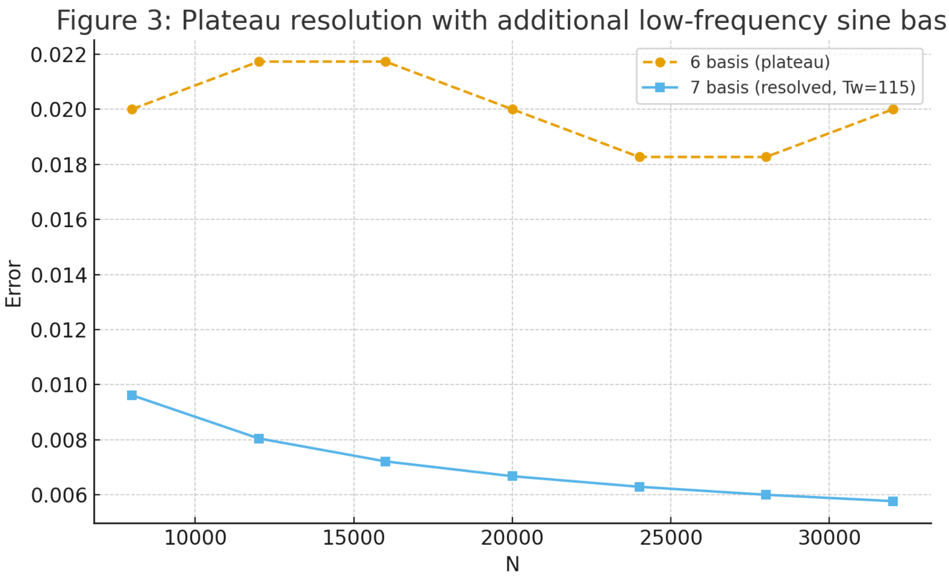

Figure 3: Plateau resolution at large N by including an additional low-frequency sine basis and narrowing the Gaussian weight ($T_w = 115$). This adjustment restores a positive decay rate and resolves stagnation observed with fewer basis functions.