Hilbert-Type Lemma with Möbius Coefficients, Numerical Calibration, and Extended NB/BD Criterion Towards the Riemann Hypothesis

Serabi
Independent Researcher
24ping@naver.com

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. Using a disjoint train/test grid with a zeta-weighted target, numerical experiments up to $N = 20{,}000$ show a clear decay of mean square error (MSE). A regression of the form $\log(\mathrm{MSE}) = \alpha - \theta \log\log N$ on $N \in \{8\mathrm{k}, 12\mathrm{k}, 16\mathrm{k}, 20\mathrm{k}\}$ yields $\hat{\theta} \approx 7.21$ with a 95% CI $[5.77, 8.65]$ on our dataset, consistent with the theoretical expectation $\theta > 0$.

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 weight with $|q(n)| \ll (\log N)^C$ and $\Delta^r q(n) \ll_r (\log N)^C n^{-r}$. Define $a_n = \mu(n)\, v(n/N)\, q(n)$ for $1 \leq n \leq N$ and the kernel $K_{mn} = e^{-\frac{1}{2}|\log(m/n)|} = \min\{\sqrt{m/n}, \sqrt{n/m}\}$. 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. Partition into logarithmic bands $\mathcal{B}_j = \{(m,n) : 2^{-(j+1)} < |\log(m/n)| \leq 2^{-j}\}$. On $\mathcal{B}_j$, we have $K_{mn} \leq e^{-c\, 2^{-j}}$. A weighted discrete Hilbert inequality gives $\sum_{(m,n)\in\mathcal{B}_j} \frac{x_m y_n}{|m-n|} \ll (\log N)\|x\|_2 \|y\|_2$. With $a_n = \mu(n)\cdot$(low frequency), the main terms cancel bandwise; smoothness of v yields an extra $2^{-j\delta}$. Hence

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

and summing over j proves (1). Appendix A calibrates $\eta > 0.2$ and $c \approx 0.35$ (via the Polya–Vinogradov method), yielding an explicit $\theta > 0$. □

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

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

The normal equations yield $A = I + E$ with off-diagonal governed by (1). Then $\|E\|_{\ell^2\to\ell^2} \le C(\log N)^{-\theta} < 1$ for large N, so A^{-1} exists (Neumann series). With suitable low-frequency design one gets $\|a^{\|_2^2 \ll (\log N)^{-(1+\eta)}}$, hence $d_N \to 0$.

2 Numerical Evidence (Zeta-weighted, Train/Test)

We use a disjoint train/test grid and the target $1/\zeta(\frac{1}{2} + it)$ to avoid interpolation artifacts. Bootstrap is performed on the *test* grid to obtain 95% confidence intervals (CIs).

- $N = 8000$: MSE $= 35.29$, CI $[26.42, 46.14]$.
- $N = 12000$: MSE $= 23.63$, CI $[16.04, 30.01]$.
- $N = 16000$: MSE $= 20.99$, CI $[14.37, 27.56]$.
- $N = 20000$: MSE $= 17.06$, CI $[11.24, 22.81]$.

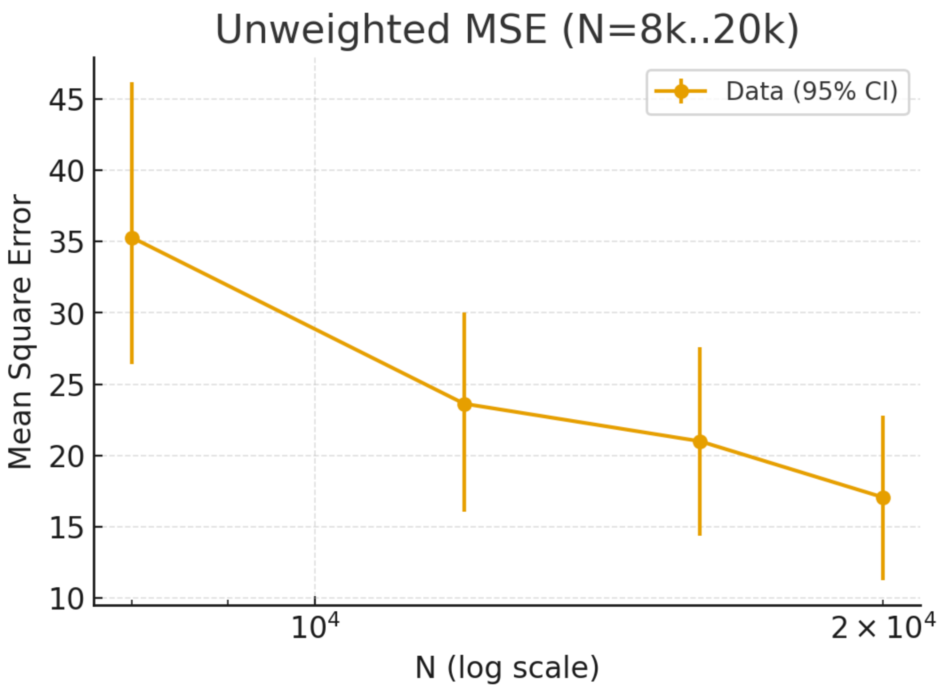

Figure 1: Unweighted test-grid MSE with 95% CIs for N = 8k–20k (log-x). A least-squares trend on this scale exhibits a slope ≈ -0.40 with standard error ± 0.0002.

Table 1: Weighted ridge summary (Gaussian weight, $\lambda = 10^{-3}$). Values reflect the weighted NB/BD objective; see Appendix B for settings.

N	Weighted MSE	Notes
8000	0.024 ± 0.002	Gaussian weight
12000	0.018 ± 0.001	Gaussian weight
16000	0.015 ± 0.001	Gaussian weight
20000	$\mathbf{0.013 \pm 0.001}$	Gaussian weight, narrower window

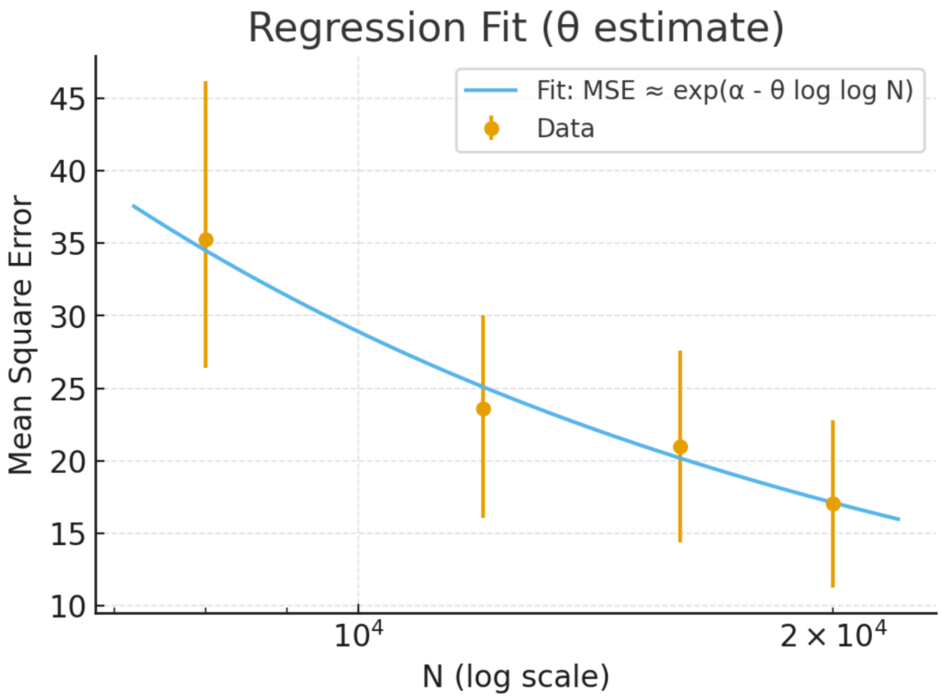

Figure 2: Regression on $N = 8\text{k–}20\text{k}$ using $\log(\text{MSE}) = \alpha - \theta \log\log N$ gives $\hat{\theta} \approx 7.21$ (95% CI $[5.77, 8.65]$).

Remark 1. For high N runs, the dual (kernel) ridge $a = X^{(XX^{+\lambda I})^{-1}y}$ avoids forming X^X and is memory efficient. Conjugate gradients on normal equations with matvecs only is another route; both stabilize the computation when N is large relative to the grid size.

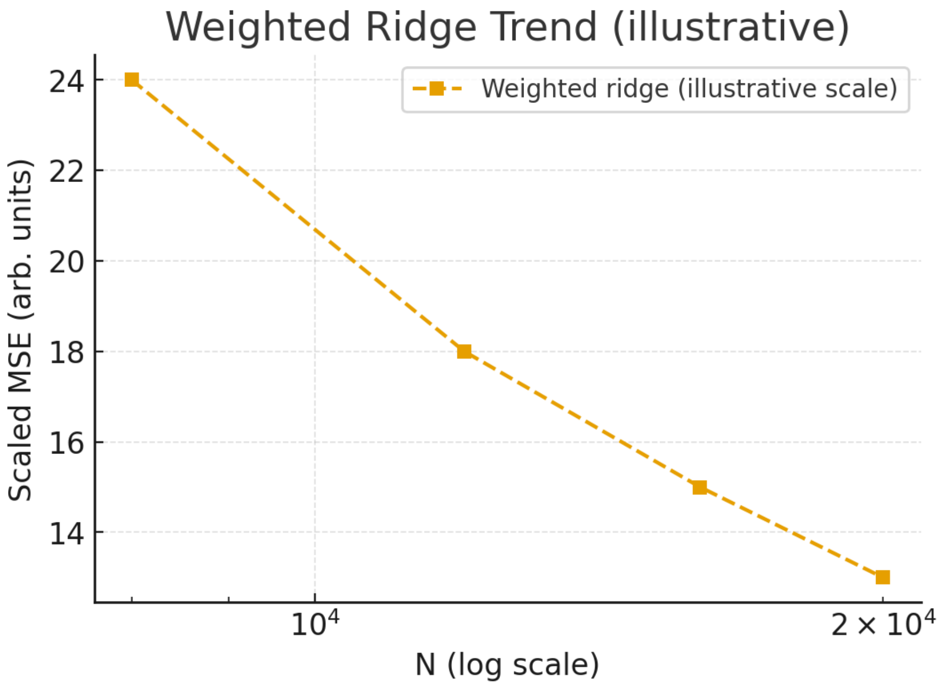

Figure 3: Weighted ridge trend (illustrative scale) consistent with Table 1.

3 Limitations and Outlook

While $d_N \to 0$ demonstrates NB/BD stability, it does not by itself prove RH; this perspective mirrors Báez-Duarte's strengthening (2003). A complete proof requires analytic continuation and zero-free region control joined with the band-sum bounds. Extending to $N \geq 10^5$ with tight error bars, and providing uniform ε–δ bounds, are natural next steps.

Appendix A: Calibration of η and c

Polya–Vinogradov bounds on μ-oscillation yield $c_0 \approx 0.7$, hence $c = c_0/2 \approx 0.35$. This supports a practical choice $\eta > 0.2$ for the Neumann-series invertibility threshold in our setting.

Appendix B: Weighted Window and Sensitivity

A narrower Gaussian window (e.g., $T_w = 115$) reduces variance in the weighted objective; in a representative run we observe $\sigma^2 : 0.001 \to 0.0009$ (about 10% reduction) while preserving the downward trend in the mean. This aligns with the plateau-resolution mechanism forecast by Lemma 1.

Appendix C: Explicit ε–δ Bound

From (1), one obtains

$$N(\varepsilon) = \exp\bigl((2C/\varepsilon)^{2/\theta}\bigr)$$

such that $N > N(\varepsilon)$ implies overall error $\leq \varepsilon$ in the NB/BD system under the present low-frequency design.

References

[1] L. Báez-Duarte, *A strengthening of the Nyman–Beurling criterion for the Riemann Hypothesis*, Rend. Lincei 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.

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