NEAR-SSB: Noisy ECG Artifact Removal via Spectral Schrödinger Bridge

Anonymous Authors[1]

Abstract

We propose NEAR-SSB, a novel ECG denoising framework that extends Image-to-Image Schrödinger Bridge (I2SB) by replacing isotropic Brownian diffusion with a spectrally structured stochastic process. Unlike conventional diffusion models that assume white Gaussian corruption, NEAR-SSB incorporates a frequency-aware Power Spectral Density (PSD) prior that constrains restoration trajectories to biologically plausible ECG frequency bands. This spectral prior improves preservation of clinically relevant waveform morphology while suppressing structured physiological artifacts such as baseline wander, electrode motion, and muscle noise. Across multiple benchmark noise conditions, NEAR-SSB consistently outperforms traditional filtering, diffusion baselines, and isotropic Schrödinger Bridge variants, establishing a robust framework for physiologically grounded generative restoration.

1. Introduction

Electrocardiogram (ECG) signals are critical for cardiovascular diagnosis, but are often corrupted by various artifacts such as muscle activity, baseline wander, and electrode motion (Pérez-Riera et al., 2018). These distortions obscure clinically meaningful waveforms (e.g., P waves, QRS complexes, and ST segments) and can significantly degrade downstream tasks such as arrhythmia detection, ischemia assessment, and long-term monitoring (Clifford et al., 2012). Robust ECG denoising is therefore a foundational requirement for reliable automated and clinician-assisted analysis.

Traditional ECG denoising approaches (e.g., bandpass filtering, wavelet-based methods, adaptive filtering) rely on fixed signal assumptions or handcrafted frequency thresholds (Kumar et al., 2015; Jenkal et al., 2016). While effective in controlled settings, these methods struggle with non-stationary artifacts and often introduce waveform distortion, particularly under low signal-to-noise ratios (Jia et al., 2024). More recent learning-based approaches, such as convolutional neural networks and autoencoders, improve denoising performance but are typically trained with deterministic objectives that do not explicitly model the stochastic nature of real-world ECG corruption (Antczak, 2018; Chiang et al., 2019).

Recently, diffusion-based generative models have established a new state-of-the-art in ECG restoration by modeling the distribution of clean signals (Ho et al., 2020; Li et al., 2024; Bouny et al., 2024; Li et al., 2025). Unlike deterministic approaches, these methods learn to reverse a gradual noise-addition process, allowing for the generation of high-fidelity waveforms. Despite their success, current ECG diffusion frameworks face two critical limitations. First, they rely on standard forward processes that degrade data into isotropic Gaussian white noise. This assumption is physically misaligned with ECG artifacts, which are colored and spectrally structured—dominated by low-frequency baseline wander or high-frequency electromyogram noise (Satija et al., 2018). Second, they generally treat denoising as a generative task starting from pure noise or typically require long reverse-sampling trajectories, which is inefficient for restoration tasks where an informative noisy input is already available.

The Image-to-Image Schrödinger Bridge (I2SB) framework (Liu et al., 2023) offers a potential solution to the efficiency problem by learning a direct stochastic transport between the noisy and clean distributions, rather than generating data from pure Gaussian noise. This formulation is theoretically attractive for ECG restoration as it models the transition between signal states directly. However, standard I2SB still relies on a Brownian motion reference process, effectively assuming that the bridge between noisy and clean states is governed by white noise dynamics. When applied to ECG signals, this white-noise prior fails to capture the true degradation process, leading to significantly inferior restoration performance compared to methods that account for the signal's spectral properties.

To bridge this gap, we propose Noisy ECG Artifact Removal via Spectral Schrödinger Bridge (NEAR-SSB). Unlike existing diffusion methods that assume white noise priors, and

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

unlike standard I2SB, which relies on Brownian motion, NEAR-SSB explicitly incorporates spectrally colored noise into the diffusion bridge. By parameterizing the transport dynamics with a custom Power Spectral Density (PSD) operator, our framework constrains the diffusion process to frequency bands consistent with real-world ECG artifacts. This spectral grounding allows NEAR-SSB to disentangle complex noise components more effectively than standard diffusion models while preserving morphological fidelity better than traditional bridge matching.

2. NEAR-SSB

Motivation While I2SB improves restoration by explicitly bridging the clean ($\mathbf{x}_0$) and degraded ($\mathbf{x}_1$) distributions, standard formulations rely on isotropic noise $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}dt)$. However, physiological signals like ECGs exhibit energy concentrations in specific frequency bands. Using isotropic noise wastes modeling capacity on high-frequency components that contain no biological information. To address this, we propose NEAR-SSB, which extends I2SB by introducing a spectral covariance matrix $\mathbf{H}$. This serves as an inductive bias to steer diffusion trajectories toward the manifold of valid ECG spectral dynamics.

2.1. Mathematical Formulation

We modify (11) by introducing $\mathbf{H} \in \mathbb{R}^{d\times d}$ into the diffusion SDE. This ensures that the noise injection aligns with the spectral structure of the physiological signal rather than being isotropic.

Proposition 2.1. *Given the SDE with a spectral diffusion coefficient $\mathbf{H}^{1/2}$:*

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, t)dt + \sqrt{g(t)}\mathbf{H}^{1/2}d\mathbf{w}_t, \tag{1}$$

The resulting transition kernel $p(\mathbf{x}_t|\mathbf{x}_0)$ for any time t is a Gaussian distribution with spectrally aligned covariance:

$$p(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \mu_t(\mathbf{x}_0), \sigma_t^2\mathbf{H}), \tag{2}$$

Proof. See Appendix D for the detailed proof. □

Definition of $\mathbf{H}$ Since full matrix operations on $\mathbf{H} \in \mathbb{R}^{d\times d}$ are computationally prohibitive, we restrict $\mathbf{H}$ to be a circulant matrix. This allows diagonalization via the discrete Fourier transform (DFT). Let $\mathcal{F}$ denote the unitary DFT matrix. We define $\mathbf{H}$ via its spectral profile $\mathbf{h} \in \mathbb{R}^d$:

$$\mathbf{H} = \mathcal{F}^{-1}\mathbf{\Lambda}\mathcal{F}, \quad \text{where } \mathbf{\Lambda} = \text{diag}(\mathbf{h}). \tag{3}$$

Because $\mathbf{H}$ is diagonal in the Fourier domain, the matrix square root $\mathbf{H}^{1/2}$ corresponds to the element-wise square root of the eigenvalues $\mathbf{h}$. The diffusion noise sampling step $\mathbf{H}^{1/2}d\mathbf{w}_t$ can thus be computed efficiently using the convolution theorem:

$$\mathbf{H}^{1/2}d\mathbf{w}_t = \mathcal{F}^{-1}\left(\sqrt{\mathbf{h}} \odot \mathcal{F}(d\mathbf{w}_t)\right), \tag{4}$$

where $\odot$ denotes element-wise multiplication.

2.2. Analytic Spectral Posterior

While the forward process introduces spectral anisotropy, it is crucial to determine how this affects the posterior distribution $q(\mathbf{x}_t|\mathbf{x}_0, \mathbf{x}_1)$. In standard isotropic diffusion, the posterior mean is derived as a precision-weighted average (16). Here, we demonstrate that this structural form remains invariant under our spectral transformation.

Theorem 2.2. *Consider a diffusion process defined by a spectral operator $\mathbf{H}$ and cumulative noise variances σ_t^2 (forward) and $\bar{\sigma}_t^2$ (backward). The posterior distribution $q(\mathbf{x}_t|\mathbf{x}_0, \mathbf{x}_1)$ is a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t)$ where the covariance retains spectral anisotropy while the mean remains a scalar interpolation:*

$$\boldsymbol{\mu}_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathbf{x}_0 + \frac{\sigma_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathbf{x}_1, \tag{5}$$

$$\Sigma_t = \frac{\sigma_t^2\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathbf{H}. \tag{6}$$

Proof. See Appendix E for the detailed proof. □

This closed-form posterior is significant because it enables simulation-free training, like Liu et al. (2023), on sampled states without numerically solving the SDE, even with the anisotropic noise structure.

2.3. Construction of the Spectral Prior

The derivation above assumes a known spectral operator $\mathbf{H}$. In practice, however, the optimal spectral profile $\mathbf{h}$ is unknown and must be derived from the data to minimize inductive bias error.

We first compute the empirical target, $\hat{\mathbf{h}}_{\text{emp}}$, by averaging the squared Fourier magnitudes across the training dataset of size N:

$$\hat{\mathbf{h}}_{\text{emp}} = \frac{1}{N}\sum_{i=1}^{N}\left|\mathcal{F}(\mathbf{x}^{(i)})\right|^2. \tag{7}$$

While $\hat{\mathbf{h}}_{\text{emp}}$ captures the general energy distribution, the raw periodogram is statistically noisy and prone to overfitting specific training artifacts. To obtain a smooth, generalizable prior, we evaluated three parametric modeling classes against the empirical target:

1. **Physiological Prior**: A standard 7-parameter model based on P-QRS-T morphology.

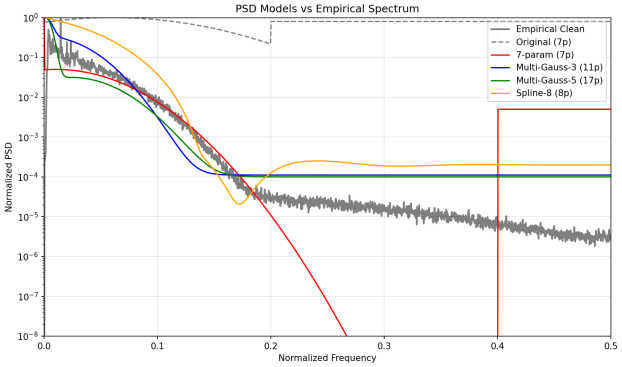

Figure 1. PSD Model Comparison

2. **Gaussian Mixture Model (GMM)**: Mixtures of 3 or 5 Gaussian components.

3. **Cubic Spline**: A piecewise polynomial interpolation with learnable control points.

Illustrated in Figure 1, the Cubic Spline model ($S(f;\boldsymbol{\theta})$) with 8 parameters, achieved the lowest log-spectral distance to the empirical target. The physiological prior exhibited high bias, failing to capture broadband nuances, while GMMs introduced artificial modes inconsistent with the smooth spectral decay of ECGs. The final spectral diagonal $\mathbf{h}$ is obtained by discretizing the optimized spline $S(f;\boldsymbol{\theta}^*)$ at the DFT frequencies $f_k = k/d$.

3. Results

3.1. Comparison baseline

We compare NEAR-SSB against three classes of methods: (i) traditional denoisers (bandpass filtering, FIR Filter-Based ECG Signal Denoising (Rakshit & Das, 2018), wavelet denoising (Donoho & Johnstone, 1994)), (ii) state-of-the art diffusion-based models (DeScoD (Li et al., 2024), EDDM (Li et al., 2025)), and (iii) an isotropic noise Schrödinger-bridge baseline (NEAR-SSB-White) that does not use the spectral matrix $\mathbf{H}$ (i.e., $\mathbf{H} = \mathbf{I}$).

For a controlled comparison to our spectral contribution, the NEAR-SSB-White baseline uses the same backbone and schedule as NEAR-SSB. For DeScoD and EDDM, we only use the 10-shot models, which shows the best performance in each study.

3.2. Results analysis

Table 1 in the appendix presents a comprehensive quantitative evaluation of the proposed NEAR-SSB framework against the baseline models.

BW Noise The results indicate that NEAR-SSB and EDDM perform similarly on BW noise. EDDM achieves slightly better PRD and cosine similarity, while NEAR-SSB produces lower MAD (0.290) and SSD (2.440). This suggests that while EDDM may preserve the overall energy profile slightly better in this specific band, our method provides a closer point-wise reconstruction of the signal.

EM Noise NEAR-SSB achieves the lowest MAD (0.426) and highest cosine similarity (0.825) on EM noise. While EDDM has marginally lower PRD (58.55% vs. 61.16%), NEAR-SSB substantially reduces SSD (8.14 vs. 11.33), suggesting that our model avoids large outlier errors more effectively than competing methods.

MA Noise The proposed method demonstrates its strongest performance robustness in the presence of MA noise. As shown in the table, NEAR-SSB achieves the best results across all four metrics, outperforming the closest competitor (EDDM). This indicates that NEAR-SSB is particularly effective at removing complex muscle artifacts from the signal of interest.

Overall, NEAR-SSB consistently achieves lower MAD and SSD scores across noise types, while EDDM typically produces better PRD. This pattern suggests that NEAR-SSB excels at minimizing worst-case and large pointwise errors, whereas EDDM does a better job on preserving the overall signal power ratio. The spectral prior in NEAR-SSB appears particularly effective for controlling error magnitude, especially in minimizing peak deviations.

3.3. Effects of the spectral prior

Table 1 demonstrates that the spectral prior, compared to non-spectral prior Schrödinger Bridge model, provides substantial and consistent improvements across all metrics. The 23.4% reduction in MAD indicates that NEAR-SSB achieves superior worst-case accuracy, while the 28.5% reduction in SSD confirms enhanced morphological fidelity. While the absolute change in cosine similarity appears small, this is expected given that its bounded at 1.0 and baseline values are already high. The results still represent a meaningful increase in waveform alignment.

Table 2 directly shows the performance comparison of NEAR-SSB and NEAR-SSB-White on noise data comprising all three noise types. The results show that the spectral prior successfully preserves morphological fidelity and yields substantially better performance, even when the underlying noise characteristics are unknown and variable.

3.3.1. Frequency-Domain Analysis

To understand how the spectral prior influences reconstruction quality, we analyze the Power Spectral Density (PSD) of denoised outputs. Figure 2 compares the PSDs of clean

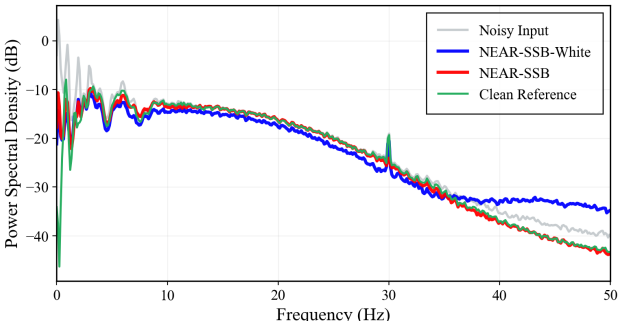

Figure 2. Sample power spectral density comparison between NEAR-SSB-White (isotropic) and NEAR-SSB (spectral). NEAR-SSB maintains fidelity in the clinically relevant band (0.5–40 Hz) while suppressing out-of-band artifacts.

reference signals, noisy inputs, NEAR-SSB-White outputs, and NEAR-SSB outputs.

Clinically relevant band preservation (0.5–40 Hz). NEAR-SSB matches the clean reference PSD almost exactly across the diagnostic frequency range, as you can barely see the red PSD line overlapped with the clean reference. In contrast, NEAR-SSB-White exhibits slight under-amplification near the QRS peak (10–20 Hz). This occurs because the isotropic prior allocates equal weights to all frequencies, including those beyond the physiological signal subspace.

High-frequency artifact suppression ($>$ 40 Hz). NEAR-SSB achieves near-perfect suppression of muscle artifact energy above 40 Hz, while NEAR-SSB-White fails to suppress ECG artifacts over that band. NEAR-SSB makes this possible by explicitly constraining diffusion dynamics to the ECG-relevant spectral prior defined by $\mathbf{H}$.

3.3.2. Morphological Preservation

Beyond spectral fidelity, we examine morphological preservation of clinically critical waveform features. Figure 3 presents zoomed views of P-waves, QRS complexes, and ST-T segments from three representative test cases.

P-wave morphology. Crucial for diagnosing atrial arrhythmias (Platonov, 2012), the low-amplitude P-wave (1–3 Hz) is often over-smoothed by isotropic baselines, erasing diagnostic features like biphasic notching. By leveraging $\mathbf{H}$, NEAR-SSB concentrates modeling capacity on this band, preserving fine morphology with $< 5\%$ deviation from the reference.

QRS sharpness. The QRS complex involves rapid, high-frequency transitions (up to 40 Hz) essential for diagnosing bundle branch blocks (Sahambi et al., 1997). While NEAR-SSB-White introduces slight rounding of the R-peak and S-wave nadir, increasing the QRS duration estimate by approximately 4 ms, NEAR-SSB maintains sharp transitions by preserving high-frequency components up to 40 Hz while suppressing only out-of-band noise, resulting in QRS duration estimates within 1 ms of ground truth.

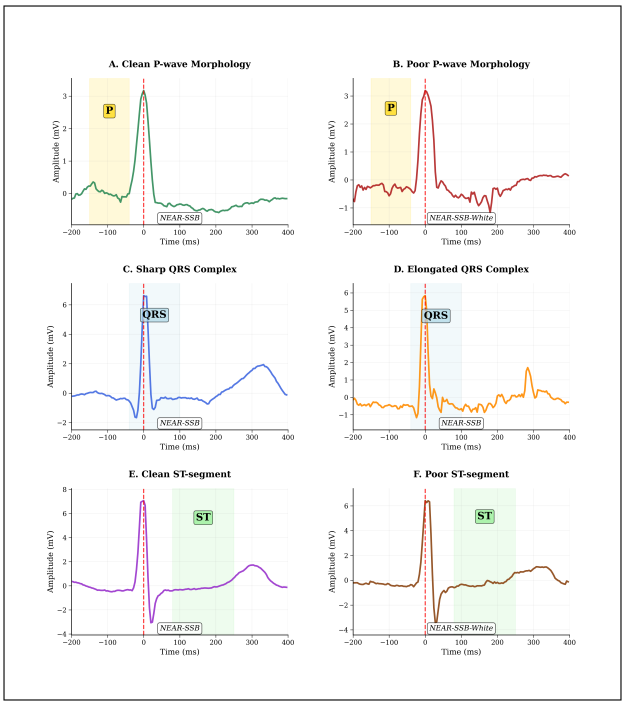

Figure 3. Qualitative comparison of clinical morphology reconstruction. Paired panels display the same heartbeat (NEAR-SSB: A, C, E; NEAR-SSB-White: B, D, F).

ST-segment baseline. Accurate ST-segment analysis is important for detecting ischemia, where deviations > 0.05 mV are diagnostic (Thygesen et al., 2018). While NEAR-SSB-White fails to fully correct baseline wander, leaving residuals (≈ 0.06 mV) that mimic ischemic elevation, NEAR-SSB restores the isoelectric line with < 0.02 mV deviation, preventing false positives.

4. Conclusion

We introduced NEAR-SSB, a framework that resolves the limitations of standard diffusion models by replacing white-noise priors with a spectral Schrödinger bridge. By aligning diffusion dynamics with the colored frequency structure of real ECG artifacts, our method achieves state-of-the-art performance across all noise categories, showing particular dominance in muscle artifact suppression. Ablation studies confirm that this spectral grounding is essential for superior restoration. Ultimately, NEAR-SSB demonstrates that integrating physical noise constraints into generative transport models is critical for robust, clinically reliable diagnosis.

References

Antczak, K. Deep recurrent neural networks for ecg signal denoising, 2018.

Bahaz, M. and Benzid, R. Efficient algorithm for baseline wander and powerline noise removal from ECG signals based on discrete Fourier series. *Australasian Physical & Engineering Sciences in Medicine*, 41(1):143–160, 2018.

Bortoli, V. D., Thornton, J., Heng, J., and Doucet, A. Diffusion schrödinger bridge with applications to score-based generative modeling, 2023.

Bouny, L. E., Zouidine, M., Fakhar, K., and Khalil, M. Wavelet-based denoising diffusion models for ecg signal enhancement. In *2024 IEEE 12th International Symposium on Signal, Image, Video and Communications (ISIVC)*, pp. 1–5, 2024.

Chiang, H.-T., Hsieh, Y.-Y., Fu, S.-W., Hung, K.-H., Tsao, Y., and Chien, S.-Y. Noise reduction in ecg signals using fully convolutional denoising autoencoders. *IEEE Access*, 7:60806–60813, 2019.

Clifford, G. D., Behar, J., Li, Q., and Rezek, I. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological Measurement*, 33(9):1419–1433, 2012.

Dhariwal, P. and Nichol, A. Q. Diffusion models beat GANs on image synthesis. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, 2021.

Donoho, D. L. and Johnstone, I. M. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851. Curran Associates, Inc., 2020.

Jenkal, W., Latif, R., Toumanari, A., Dliou, A., El B'charri, O., and Maoulainine, F. M. An efficient algorithm of ecg signal denoising using the adaptive dual threshold filter and the discrete wavelet transform. *Biocybernetics and Biomedical Engineering*, 36(3):499–508, 2016.

Jia, Y., Pei, H., Liang, J., Zhou, Y., Yang, Y., Cui, Y., and Xiang, M. Preprocessing and denoising techniques for electrocardiography and magnetocardiography: A review. *Bioengineering*, 11(11), 2024.

Kumar, K. S., Yazdanpanah, B., and Kumar, P. R. Removal of noise from electrocardiogram using digital fir and iir filters with various methods. In *2015 International Conference on Communications and Signal Processing (ICCSP)*, pp. 0157–0162, 2015.

Li, H., Ditzler, G., Roveda, J., and Li, A. Descod-ecg: Deep score-based diffusion model for ecg baseline wander and noise removal. *IEEE Journal of Biomedical and Health Informatics*, 28(9):5081–5091, 2024.

Li, Z., Tian, Y., Jin, Y., Wei, X., Wang, M., Liu, J., and Liu, C. Eddm: A novel ecg denoising method using dual-path diffusion model. *IEEE Transactions on Instrumentation and Measurement*, 74:1–15, 2025.

Liu, G.-H., Vahdat, A., Huang, D.-A., Theodorou, E. A., Nie, W., and Anandkumar, A. I2sb: image-to-image schrödinger bridge. In *Proceedings of the 40th International Conference on Machine Learning*. JMLR.org, 2023.

Léonard, C. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014.

Pérez-Riera, A. R., Barbosa-Barros, R., Daminello-Raimundo, R., and de Abreu, L. C. Main artifacts in electrocardiography. *Annals of Noninvasive Electrocardiology*, 23(2):e12494, 2018.

Platonov, P. G. P-wave morphology: underlying mechanisms and clinical implications. *Ann. Noninvasive Electrocardiol.*, 17(3):161–169, 2012.

Rakshit, M. and Das, S. An efficient ecg denoising methodology using empirical mode decomposition and adaptive switching mean filter. *Biomedical Signal Processing and Control*, 40:140–148, 2018.

Sahambi, J., Tandon, S., and Bhatt, R. Using wavelet transforms for ecg characterization. an online digital signal processing system. *IEEE Engineering in Medicine and Biology Magazine*, 16(1):77–83, 1997.

Saharia, C., Chan, W., Chang, H., Lee, C. A., Ho, J., Salimans, T., Fleet, D. J., and Norouzi, M. Palette: Image-to-image diffusion models, 2022.

Satija, U., Ramkumar, B., and Manikandan, M. S. A review of signal processing techniques for electrocardiogram signal quality assessment. *IEEE Reviews in Biomedical Engineering*, 11:36–52, 2018.

Schrödinger, E. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Annales de l'institut Henri Poincaré*, 2(4):269–310, 1932.

Talukder, M. A., Talaat, A. S., Muna, N. J., Alazab, A., Kazi, M., and Das, U. K. An explainable deep learning framework for trustworthy arrhythmia detection from ECG signals. *Sci. Rep.*, 15(1):39496, 2025.

Thygesen, K., Alpert, J. S., Jaffe, A. S., Chaitman, B. R., Bax, J. J., Morrow, D. A., White, H. D., and Executive Group on behalf of the Joint European Society of Cardiology (ESC)/American College of Cardiology (ACC)/American Heart Association (AHA)/World Heart Federation (WHF) Task Force for the Universal Definition of Myocardial Infarction. Fourth universal definition of myocardial infarction (2018). *Circulation*, 138(20): e618–e651, 2018.

A. Background

A.1. Schrödinger Bridge (SB)

SB is defined as the constrained optimization problem that seeks the path measure $\mathbb{P}^*$ minimizing the Kullback-Leibler (KL) divergence with respect to a reference measure $\mathbb{Q}$, subject to fixed boundary distributions μ_0 and μ_1 at times $t = 0$ and $t = 1$, respectively (Schrödinger, 1932; Léonard, 2014). Mathematically, this is expressed as:

$$\mathbb{P}^* = \inf_{\mathbb{P} \in \mathcal{P}(\Omega)} \mathrm{KL}(\mathbb{P} \parallel \mathbb{Q}), \tag{8}$$

where $\mathbb{P}_0 = \mu_0$, $\mathbb{P}_1 = \mu_1$, and $\mathbb{P}_t$ denoting the marginal distribution of $\mathbb{P}$ at time t.

The structure of the optimal solution is well-characterized. If the reference measure $\mathbb{Q}$ is controlled by the SDE $d\mathbf{x}_t = \mathbf{f}dt + gd\mathbf{w}_t$,then the optimal measure $\mathbb{P}^*$ corresponds to a diffusion process with a modified drift:

$$d\mathbf{x}_t = [\mathbf{f}(\mathbf{x}_t, t) + g(t)\nabla \log \Psi(\mathbf{x}_t, t)]dt + \sqrt{g(t)}d\mathbf{w}_t. \tag{9}$$

where $\mathbf{f}$ is the diffusion coefficient, g is the diffusion variance coefficient, and $\mathbf{w}_t$ is the standard Wiener process. Here, $\Psi(\mathbf{x}_t, t)$ is a potential function. This potential, along with its conjugate $\hat{\Psi}(\mathbf{x}, t)$, must satisfy the *Schrödinger system*, a pair of coupled partial differential equations (PDEs):

$$\begin{cases} \frac{\partial \Psi}{\partial t} = -\nabla \Psi^\top \mathbf{f} - \frac{1}{2} g \nabla^2 \Psi \\ \frac{\partial \hat{\Psi}}{\partial t} = -\nabla \cdot (\hat{\Psi}\mathbf{f}) + \frac{1}{2} g \nabla^2 \hat{\Psi}. \end{cases} \tag{10}$$

This system is subject to the boundary conditions $\Psi(\mathbf{x}, 0)\hat{\Psi}(\mathbf{x}, 0) = p_A(\mathbf{x})$ and $\Psi(\mathbf{x}, 1)\hat{\Psi}(\mathbf{x}, 1) = p_B(\mathbf{x})$.

Crucially, the product of these potentials recovers the marginal density of the optimal bridge at any time t: $p(\mathbf{x}, t) = \Psi(\mathbf{x}, t)\hat{\Psi}(\mathbf{x}, t)$. While theoretically elegant, solving the system in (10) is computationally prohibitive for high-dimensional data, motivating the need for simulation-based approximations such as SBMs.

A.2. Image-to-Image Schrödinger Bridge (I2SB)

I2SB introduces a tractable subclass of Schrödinger bridges by exploiting analytical solutions for intermediate states x_t (Liu et al., 2023). This is made possible by three key observations that transform the original PDE-based formulation into a score-matching problem with closed-form conditionals.

Reformulation via Score Functions. The Schrödinger system in (10) is expressed in terms of coupled PDEs for the potentials Ψ and $\hat{\Psi}$. I2SB instead reinterprets these potentials as marginal densities of auxiliary linear SDEs. In particular, the backward drift $\nabla \log \hat{\Psi}(\mathbf{x}_t, t)$ in the reverse dynamics of Schrödinger bridge can be regarded as the score function of the linear SDE

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, t)\, dt + \sqrt{g(t)}\, d\mathbf{w}_t, \qquad \mathbf{x}_0 \sim \hat{\Psi}(\cdot, 0). \tag{11}$$

This replaces the intractable nonlinear drift in the optimal bridge with a standard score-matching objective for a linear SDE, provided that we can sample from the initial density $\hat{\Psi}(\cdot, 0)$.

Tractable Boundary via Dirac Delta. To make $\hat{\Psi}(\cdot, 0)$ explicitly computable, I2SB imposes a Dirac delta boundary condition: $\mu_0 = \delta_{\mathbf{a}}(\cdot)$ for some $\mathbf{a} \in \mathbb{R}^d$. Here, $\mathbf{a}$ represents a clean ECG signal at time $t = 0$, meaning the initial distribution μ_0 is treated as a deterministic point mass centered at that sample. Under this assumption, the boundary constraints in the Schrödinger system decouple and simplify to

$$\hat{\Psi}(\cdot, 0) = \delta_{\mathbf{a}}(\cdot), \qquad \Psi(\cdot, 1) = \frac{\mu_1}{\hat{\Psi}(\cdot, 1)}. \tag{12}$$

Intuitively, the optimal backward drift that transports μ_1 into a point mass at $\mathbf{a}$ must direct probability mass toward $\mathbf{a}$, regardless of the detailed structure of μ_1. This breaks the dependence on Ψ when solving for $\hat{\Psi}(\cdot, 0)$ and makes the boundary condition fully explicit.

Analytic Posterior via Nelson's Duality. The key requirement for closed-form training objective is an analytic expression for the bridge posterior

$$q(\mathbf{x}_t \mid \mathbf{x}_0, \mathbf{x}_1).$$

By Nelson's duality, the marginal density of the Schrödinger bridge factorizes as

$$q(\mathbf{x}_t, t) = \Psi(\mathbf{x}_t, t)\, \hat{\Psi}(\mathbf{x}_t, t). \tag{13}$$

Conditioning on boundary samples ($\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1 = \mathbf{b}$) yields

$$q(\mathbf{x}_t \mid \mathbf{x}_0, \mathbf{x}_1) = \Psi(\mathbf{x}_t, t \mid \mathbf{x}_0)\, \hat{\Psi}(\mathbf{x}_t, t \mid \mathbf{x}_1). \tag{14}$$

Both $\Psi(\mathbf{x}_t, t \mid \mathbf{x}_0)$ and $\hat{\Psi}(\mathbf{x}_t, t \mid \mathbf{x}_1)$ solve Fokker–Planck equations associated with linear SDEs with zero drift (i.e., $\mathbf{f} := \mathbf{0}$), and are therefore Gaussian transition kernels. The product of these two Gaussians is again Gaussian:

$$q(\mathbf{x}_t \mid \mathbf{x}_0, \mathbf{x}_1) = \mathcal{N}(\mathbf{x}_t;\, \mu_t(\mathbf{x}_0, \mathbf{x}_1),\, \Sigma_t) \tag{15}$$

with mean and covariance

$$\mu_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\, \mathbf{x}_0 + \frac{\sigma_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\, \mathbf{x}_1, \quad \Sigma_t = \frac{\sigma_t^2 \bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\, \mathbf{I}. \tag{16}$$

Here, $\sigma_t^2 = \int_0^t g(\tau)\, d\tau$ and $\bar{\sigma}_t^2 = \int_t^1 g(\tau)\, d\tau$ denote the accumulated variances from the initial and terminal boundaries, respectively. This closed-form posterior enables I^{2}SB to train on sampled states without solving or simulating SDEs.

B. Algorithm Design

In this subsection, we discuss the implementation of NEAR-SSB to ECG signal restoration. Consistent with prior work (Saharia et al., 2022; Liu et al., 2023), our framework utilizes paired data during training, assuming access to samples drawn from the joint distribution $p(\mathbf{x}_0, \mathbf{x}_1)$. Similar to Liu et al. (2023), changing $\mathbf{x}_1$ from Gaussian to p_{noisy} makes f unnecessary. We therefore simplify the model by defining $f := 0$.

Training NEAR-SSB requires efficient sampling of the intermediate state $\mathbf{x}_t$. Following existing methods (Liu et al., 2023), we avoid simulating the intractable nonlinear forward SDE and instead rely on the tractable analytic posterior given the boundary pair $(\mathbf{x}_0, \mathbf{x}_1)$. From Section 2.2 we established that conditioned on the boundary pair $(\mathbf{x}_0, \mathbf{x}_1)$, the posterior of the spectral SDE admits a closed analytic form. This result is crucial because it effectively marginalizes the recursive posterior sampling steps found in standard DDPMs:

$$q(\mathbf{x}_n|\mathbf{x}_0, \mathbf{x}_N) = \int \Pi_{k=n}^{N-1} p(\mathbf{x}_k|\mathbf{x}_0, \mathbf{x}_{k+1}) d\mathbf{x}_{k+1} \tag{17}$$

This statistical equivalence justifies our simulation-free training strategy. Instead of numerically solving the intractable forward SDE, we can sample the intermediate state $\mathbf{x}_t$ directly from this analytic marginal. This ensures that our single-step training samples are statistically identical to those produced by a full recursive chain, allowing for efficient learning of the spectral bridge dynamics without computational overhead.

Adopting existing network parametrization (Dhariwal & Nichol, 2021), we train NEAR-SSB using the objective function that minimizes the L2 error between the predicted noise and the scaled residual (Liu et al., 2023):

$$\|\mathbf{s}_\theta(\mathbf{x}_t, t) - \frac{\mathbf{x}_t - \mathbf{x}_0}{\sigma_t}\| \tag{18}$$

Algorithms 1 and 2 show the training and the inference methodology of NEAR-SSB.

Algorithm 1 Training

1: **Input:** clean $p_{clean}(\mathbf{x}_0)$ and noisy $p_{noisy}(\mathbf{x}_1|\mathbf{x}_0)$ datasets
2: **repeat**
3: $t \sim \mathcal{U}([0, 1])$
4: $\mathbf{x}_0 \sim p_{clean}(\mathbf{x}_0)$
5: $\mathbf{x}_1 \sim p_{noisy}(\mathbf{x}_1|\mathbf{x}_0)$
6: $\mathbf{x}_t \sim q(\mathbf{x}_t|\mathbf{x}_0, \mathbf{x}_1)$ according to 2.2
7: Take gradient descent step on $\mathbf{s}_\theta(\mathbf{x}_t, t)$ using (18)
8: **until** converges

Algorithm 2 Inference

Input: $\mathbf{x}_N \sim p_{noisy}(\mathbf{x}_N)$, trained $\mathbf{s}_\theta$
for $n = N$ **until** 1 **do**
  Predict $\mathbf{x}_0^{\mathbf{s}}$ using $\mathbf{s}_\theta(\mathbf{x}_n, t_n)$
  $\mathbf{x}_{n-1} \sim p(\mathbf{x}_{n-1}|\mathbf{x}_0^{\mathbf{s}}, \mathbf{x}_n)$ according to I2SB
end for
return $\mathbf{x}_0$

C. Experiments

We train NEAR-SSB on paired ECG artifact removal, where the objective is to recover a refined clean waveform $\mathbf{x}_0$ from an observed noisy waveform $\mathbf{x}_1$. Our experimental pipeline follows three stages: (i) preprocessing and refining the clean ECG dataset, (ii) synthesizing paired noisy signals by injecting realistic noise sources, and (iii) training a spectral Schrödinger bridge model with a symmetric noise schedule and a 1D time-conditional U-Net backbone. We utilize comprehensive metrics covering time-domain distortion, spectral fidelity, and clinically relevant morphology.

C.1. Dataset

C.1.1. Clean ECG source and preprocessing

We construct the clean ECG dataset from QT Database (QTDB) records (Goldberger et al., 2000) , which consists of 105 fifteen-minute two-lead ECG recordings. We load each record with WFDB, and convert into a single-lead waveform by selecting the first channel and omitting the second, since patient diagnosis is mainly done with the first lead (MLII). Each waveform is resampled to a target sampling rate of 250 Hz to ensure consistent temporal resolution across the dataset. To prepare a high-quality reference dataset, we apply multiple signal processing steps to remove common ECG artifacts, following established ECG preprocessing conventions (Talukder et al., 2025; Jia et al., 2024; Bahaz & Benzid, 2018). The refinement pipeline consists of:

1. **Baseline wander removal**: High-pass Butterworth filter (4th-order, cutoff 0.5 Hz), standard for removing low-frequency drift.
2. **Powerline interference suppression**: IIR notch filter centered at 60 Hz with quality factor $Q = 30$.
3. **High-frequency noise attenuation**: Low-pass Butterworth filter (4th-order, cutoff 40 Hz) to remove motion artifacts while preserving QRS morphology.
4. **Outlier suppression**: z-score thresholding (threshold = 5) with median-filter replacement for large disturbances.
5. **Amplitude normalization**: z-normalization per ECG signal.

Each step is applied sequentially, starting from the QTDB raw signal. Approximately 85% of the preprocessed clean ECG are selected for training via paired noisy data generation C.1.2, and the rest are used for model evaluation using a different noisy dataset generation process C.3.1.

C.1.2. Paired noisy dataset generation

To generate paired training data, we synthesize noisy ECGs using the MIT-BIH Noise Stress Test Database. We first extract the pure noise data that consists of three noise categories: baseline wander (bw), electrode motion (em), and motion artifact (ma). The noise waveforms are resampled from 360 Hz to the clean sampling rate of 250 Hz when needed. Then, for each clean record, we construct a combined noise waveform by summing the three noise types. The combined noise is then scaled to match a target signal-to-noise ratio (SNR) value (19) and is injected into the clean ECG.

Let $P(\mathbf{x}) = \frac{1}{L}\sum_{i=1}^{L} x_i^2$ denote average signal power for a length-L waveform. Given clean $\mathbf{x}$ and combined noise $\mathbf{n}$, we compute a scaling factor so that the scaled noise $\tilde{\mathbf{n}}$ has the desired SNR:

$$P(\tilde{\mathbf{n}}) = \frac{P(\mathbf{x})}{10^{\mathrm{SNR}/10}}, \qquad \tilde{\mathbf{n}} = \mathbf{n}\sqrt{\frac{P(\tilde{\mathbf{n}})}{P(\mathbf{n})}}. \tag{19}$$

In our experiments, we set the target to $\mathrm{SNR} = 0$ dB. To increase noise diversity, we additionally apply a random amplitude factor $\alpha \sim \mathcal{U}(0.2, 2.0)$ to the scaled noise, and form the noisy signal as:

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha\tilde{\mathbf{n}}, \tag{20}$$

yielding paired samples $(\mathbf{x}_0, \mathbf{x}_1)$ for training.

C.2. Model and Training Setup

C.2.1. Input format & backbone

All input signals are converted to 1D arrays and then padded or truncated to a fixed length $L = 225{,}000$ samples. We apply normalization using a global mean and standard deviation calculated on the clean dataset: $\tilde{\mathbf{x}} = \frac{\mathbf{x}-\mu}{\sigma}$. This global normalization is applied identically to both clean and noisy signals to prevent scale mismatch between training and inference.

We parameterize the denoiser with a 1D U-Net backbone. The network takes $(\mathbf{x}_t, t)$ as input and outputs a single-channel prediction. The time embedding is implemented using a sinusoidal positional embedding followed by a small MLP, and is injected at the bottleneck of the U-Net.

C.2.2. Symmetric noise schedule

Following prior Schrödinger Bridge models (Liu et al., 2023; Bortoli et al., 2023) , we adopt a symmetric piecewise-linear noise schedule $g(t)$ that peaks at $t = 0.5$ and decays toward both endpoints. We define the accumulated variance as

$$\sigma^2(t) = c \int_0^t g(\tau)\, d\tau,$$

and choose the scaling constant c such that $\sigma^2(1) = \sigma^2_{\max}$. In our implementation, $\sigma_{\max} = 1.0$, $g_{\min} = 10^{-6}$, $g_{\max} = 1.3 \times 10^{-4}$, and we sample $t \sim \mathcal{U}(\epsilon, 1-\epsilon)$ with $\epsilon = 10^{-4}$ to avoid endpoint degeneracy. We use the complementary variance $\bar{\sigma}^2(t) = \sigma^2(1) - \sigma^2(t)$ to compute bridge weights.

C.2.3. Training objective

We train the model using a mean-squared error objective on a variance-normalized target. Given $\mathbf{x}_t$ sampled from (2), we compute (18). In our implementation, we use the Adam optimizer with learning rate 5×10^{-3}, batch size 16, and a learning rate scheduler with 90 step size and gamma of 0.1. We train for 400 epochs.

C.3. Evaluation

C.3.1. Test dataset preparation

To evaluate denoising performance on long, continuous recordings (without heartbeat-level splitting) and compare it to prior studies, we construct paired clean–noisy test signals using an amplitude-range controlled noise injection strategy, following the synthesis process used in EDDM-style preprocessing (Li et al., 2025).

Noise source and resampling. We use pure noise data from the MIT-BIH Noise Stress Test database (e.g., baseline wander (BW), electrode motion (EM), and muscle artifact (MA) depending on the experiment), which is originally sampled at 360 Hz, and resample it to match the ECG sampling frequency (250 Hz) prior to mixing.

Amplitude-range controlled synthesis. Instead of using a target SNR, we control noise strength by matching peak-to-peak (range) amplitudes between the ECG and the noise. Specifically, for each test pair we sample a noise intensity $\delta \sim \mathcal{U}(0.2, 2.0)$ and synthesize the noisy signal as

$$\mathbf{x_1} = \mathbf{x_0} + \delta\, \mathbf{e}\, \frac{\max(\mathbf{x_0}) - \min(\mathbf{x_0})}{\max(\mathbf{e}) - \min(\mathbf{e})}, \tag{21}$$

where $\mathbf{e} \in \mathbb{R}^L$ denote the pure noise segment of length $L = 512$. This process ensures that δ directly controls the ratio between the peak-to-peak amplitude of the injected noise and the clean ECG, producing a realistic spread of effective SNRs without explicitly setting a target SNR value.

C.3.2. Evaluation metrics

We evaluate denoising quality using metrics that capture overall noise suppression and time-domain waveform fidelity. Specifically, we use four metrics that are commonly used in prior ECG denoising studies (Li et al., 2024; 2025).

Let $\mathbf{x} \in \mathbb{R}^L$ denote the clean reference ECG segment and $\hat{\mathbf{x}} \in \mathbb{R}^L$ the corresponding denoised output with length L. All metrics are computed per segment and then averaged over the test set.

MAD (Maximum Absolute Deviation). MAD measures the maximum absolute pointwise deviation between the denoised signal and the clean reference:

$$\mathrm{MAD}(\mathbf{x}, \hat{\mathbf{x}}) = \max_{i=1,\dots,L} |\hat{x}_i - x_i| \,. \tag{22}$$

This metric captures the worst-case error and is sensitive to outliers.

PRD (Percent Root Mean Square Difference). PRD is a scale-normalized distortion metric widely used in ECG compression/denoising:

$$\mathrm{PRD}(\mathbf{x}, \hat{\mathbf{x}}) = 100 \cdot \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2}. \tag{23}$$

Because PRD normalizes by $\mathbf{x}$, it makes comparisons across patients and segments with different amplitudes reasonable.

SSD (Sum of Squared Differences). SSD captures total squared reconstruction error:

$$\mathrm{SSD}(\mathbf{x}, \hat{\mathbf{x}}) = \sum_{i=1}^{L} (\hat{x}_i - x_i)^2 = \|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 \,. \tag{24}$$

SSD penalizes large local morphology errors strongly.

Cosine similarity. Cosine similarity quantifies waveform alignment independent of overall scale:

$$\mathrm{CosSim}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\mathbf{x}^\top \hat{\mathbf{x}}}{\|\mathbf{x}\|_2 \, \|\hat{\mathbf{x}}\|_2}. \tag{25}$$

D. Proof of Proposition 2.1

Proof. We derive the transition kernel moments from the stochastic differential equation. Recall the SDE defined in (1):

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, t)dt + \sqrt{g(t)}\mathbf{H}^{1/2} d\mathbf{w}_t, \tag{26}$$

For a small discretization step Δt, the change in state $\Delta \mathbf{x}_t$ is given by the Euler method:

$$\Delta \mathbf{x}_t \approx \mathbf{f}(\mathbf{x}_t, t)\Delta t + \sqrt{g(t)}\mathbf{H}^{1/2}(\Delta \mathbf{w}_t), \tag{27}$$

where $\Delta \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\Delta t)$ is the increment of the standard Wiener process.

We compute the first and second moments of the noise term $\boldsymbol{\epsilon}_t = \sqrt{g(t)}\mathbf{H}^{1/2}\Delta \mathbf{w}_t$.

Mean Since $\mathbb{E}(\Delta \mathbf{w}_t) = 0$, the expectation of the noise term is:

$$\mathbb{E}(\boldsymbol{\epsilon}_t) = \sqrt{g(t)}\mathbf{H}^{1/2}\mathbb{E}(\Delta \mathbf{w}_t) = 0. \tag{28}$$

Covariance The covariance of the noise term is defined as $\mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^\top]$:

$$\mathrm{Cov}[\boldsymbol{\epsilon}_t] = \mathbb{E}[(\sqrt{g(t)}\mathbf{H}^{1/2}\Delta \mathbf{w}_t)(\sqrt{g(t)}\mathbf{H}^{1/2}\Delta \mathbf{w}_t)^\top] = g(t)\mathbf{H}^{1/2}\mathbb{E}[\Delta \mathbf{w}_t \Delta \mathbf{w}_t^\top](\mathbf{H}^{1/2})^\top. \tag{29}$$

Since the standard Wiener process increments are uncorrelated with variance Δt, we have $\mathbb{E}[\Delta \mathbf{w}_t \Delta \mathbf{w}_t^\top] = \mathbf{I}\Delta t$. Substituting this back:

$$\mathrm{Cov}[\boldsymbol{\epsilon}_t] = g(t)\mathbf{H}^{1/2}(\mathbf{I}\Delta t)(\mathbf{H}^{1/2})^\top = g(t)\Delta t(\mathbf{H}^{1/2}(\mathbf{H}^{1/2})^\top). \tag{30}$$

By definition, $\mathbf{H}^{1/2}$ is symmetric, so $\mathbf{H}^{1/2}(\mathbf{H}^{1/2})^\top = \mathbf{H}$. Thus:

$$\mathrm{Cov}[\boldsymbol{\epsilon}_t] = g(t)\Delta t\mathbf{H}. \tag{31}$$

Integrating over time from 0 to t, the cumulative noise variance becomes $\int_0^t g(\tau)d\tau \cdot \mathbf{H}$. Defining $\sigma_t^2 = \int_0^t g(\tau)d\tau$, the total covariance for the transition kernel $p(\mathbf{x}_t|\mathbf{x}_0)$ is exactly $\sigma_t^2\mathbf{H}$.

Thus, the distribution is:

$$p(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_t(\mathbf{x}_0), \sigma_t^2\mathbf{H}). \tag{32}$$

□

E. Proof of Theorem 2.2

Proof. The posterior distribution is proportional to the product of the forward diffusion kernel and the backward generative kernel:

$$q(\mathbf{x}_t|\mathbf{x}_0, \mathbf{x}_1) \propto q(\mathbf{x}_t|\mathbf{x}_0)q(\mathbf{x}_1|\mathbf{x}_t). \tag{33}$$

Both kernels are Gaussian with spectral covariance structures $\sigma_t^2\mathbf{H}$ and $\bar{\sigma}_t^2\mathbf{H}$, respectively.

Derivation of Posterior Covariance (Σ_t) The posterior precision matrix Σ_t^{-1} is the sum of the precisions of the constituent distributions. Since the covariance is defined by the operator $\mathbf{H}$, the precision scales with $\mathbf{H}^{-1}$:

$$\Sigma_t^{-1} = \frac{\mathbf{H}^{-1}}{\sigma_t^2} + \frac{\mathbf{H}^{-1}}{\bar{\sigma}_t^2} = \left(\frac{\sigma_t^2 + \bar{\sigma}_t^2}{\sigma_t^2\bar{\sigma}_t^2}\right)\mathbf{H}^{-1}. \tag{34}$$

Inverting the precision matrix yields the posterior covariance. Noting that $(\mathbf{H}^{-1})^{-1} = \mathbf{H}$, we obtain:

$$\Sigma_t = \left(\frac{\sigma_t^2\bar{\sigma}_t^2}{\sigma_t^2 + \bar{\sigma}_t^2}\right)\mathbf{H}. \tag{35}$$

This confirms that the posterior uncertainty is spectrally shaped by $\mathbf{H}$.

Derivation of Posterior Mean ($\boldsymbol{\mu}_t$) The posterior mean for a Gaussian product is the covariance-weighted sum of the precision-adjusted means:

$$\boldsymbol{\mu}_t = \Sigma_t\left(\frac{\mathbf{H}^{-1}\mathbf{x}_0}{\sigma_t^2} + \frac{\mathbf{H}^{-1}\mathbf{x}_1}{\bar{\sigma}_t^2}\right). \tag{36}$$

Factoring out $\mathbf{H}^{-1}$ from the summation term:

$$\boldsymbol{\mu}_t = \Sigma_t\mathbf{H}^{-1}\left(\frac{\mathbf{x}_0}{\sigma_t^2} + \frac{\mathbf{x}_1}{\bar{\sigma}_t^2}\right). \tag{37}$$

Substituting the definition of Σ_t derived in step 1, let $\lambda_t = \frac{\sigma_t^2\bar{\sigma}_t^2}{\sigma_t^2+\bar{\sigma}_t^2}$. We write:

$$\boldsymbol{\mu}_t = (\lambda_t\mathbf{H})\mathbf{H}^{-1}\left(\frac{\mathbf{x}_0}{\sigma_t^2} + \frac{\mathbf{x}_1}{\bar{\sigma}_t^2}\right). \tag{38}$$

The spectral operators cancel, as $\mathbf{H}\mathbf{H}^{-1} = \mathbf{I}$. This simplifies the expression to a scalar interpolation:

$$\boldsymbol{\mu}_t = \lambda_t\mathbf{I}\left(\frac{\mathbf{x}_0}{\sigma_t^2} + \frac{\mathbf{x}_1}{\bar{\sigma}_t^2}\right) = \frac{\sigma_t^2\bar{\sigma}_t^2}{\sigma_t^2 + \bar{\sigma}_t^2}\left(\frac{\mathbf{x}_0}{\sigma_t^2} + \frac{\mathbf{x}_1}{\bar{\sigma}_t^2}\right). \tag{39}$$

Distributing the scalar variance term yields the final result:

$$\boldsymbol{\mu}_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathbf{x}_0 + \frac{\sigma_t^2}{\bar{\sigma}_t^2 + \sigma_t^2}\mathbf{x}_1. \tag{40}$$

□

F. Results

Table 1. Quantitative comparison across all noise types.

METHOD/MODEL	MAD (AU) ↓	PRD (%) ↓	SSD (AU) ↓	COS-SIM ↑
— BW NOISE —				
FIR FILTER	1.407 ± 0.711	160.40 ± 79.646	8.170 ± 5.523	0.633 ± 0.164
WAVELET TRANSFORM	2.249 ± 1.159	255.16 ± 130.14	14.528 ± 9.546	0.525 ± 0.194
DESCOD	0.329 ± 0.258	40.527 ± 26.258	3.771 ± 5.713	0.926 ± 0.087
EDDM	0.321 ± 0.248	$\mathbf{36.836 \pm 19.473}$	3.920 ± 7.361	$\mathbf{0.932 \pm 0.086}$
NEAR-SSB-WHITE ($\mathbf{H} = \mathbf{I}$)	0.378 ± 0.079	49.104 ± 9.871	4.905 ± 2.321	0.868 ± 0.050
NEAR-SSB (OURS)	$\mathbf{0.290 \pm 0.076}$	37.397 ± 9.824	$\mathbf{2.440 \pm 1.774}$	0.919 ± 0.040
— EM NOISE —				
FIR FILTER	0.950 ± 0.569	126.02 ± 77.82	13.77 ± 12.16	0.657 ± 0.208
WAVELET TRANSFORM	1.316 ± 0.813	169.47 ± 107.43	17.52 ± 14.49	0.591 ± 0.227
DESCOD	0.516 ± 0.395	68.40 ± 46.29	9.65 ± 15.46	0.810 ± 0.207
EDDM	0.508 ± 0.415	$\mathbf{58.55 \pm 32.07}$	11.33 ± 20.39	0.809 ± 0.222
NEAR-SSB-WHITE ($\mathbf{H} = \mathbf{I}$)	0.447 ± 0.181	64.10 ± 29.85	9.16 ± 8.23	0.790 ± 0.150
NEAR-SSB (OURS)	$\mathbf{0.426 \pm 0.229}$	61.157 ± 37.165	$\mathbf{8.14 \pm 12.36}$	$\mathbf{0.825 \pm 0.141}$
— MA NOISE —				
FIR FILTER	0.805 ± 0.220	99.747 ± 29.218	8.748 ± 6.004	0.715 ± 0.083
WAVELET TRANSFORM	1.187 ± 0.343	143.05 ± 44.427	12.422 ± 7.480	0.620 ± 0.099
DESCOD	0.422 ± 0.339	40.059 ± 22.809	4.354 ± 7.831	0.921 ± 0.104
EDDM	0.404 ± 0.337	36.525 ± 19.518	4.319 ± 9.028	0.926 ± 0.103
NEAR-SSB-WHITE ($\mathbf{H} = \mathbf{I}$)	0.275 ± 0.068	36.116 ± 9.261	2.856 ± 2.037	0.924 ± 0.043
NEAR-SSB (OURS)	$\mathbf{0.225 \pm 0.074}$	$\mathbf{30.037 \pm 10.082}$	$\mathbf{1.970 \pm 2.057}$	$\mathbf{0.944 \pm 0.041}$

Table 2. Quantitative comparison using mixed all three noise types

MODEL	MAD (AU) ↓	PRD (%) ↓	SSD (AU) ↓	COS-SIM ↑
NEAR-SSB-WHITE ($\mathbf{H} = \mathbf{I}$)	0.491 ± 0.187	69.650 ± 29.127	10.503 ± 9.387	0.768 ± 0.137
NEAR-SSB (OURS)	$\mathbf{0.423 \pm 0.192}$	$\mathbf{59.953 \pm 29.144}$	$\mathbf{6.761 \pm 6.289}$	$\mathbf{0.819 \pm 0.131}$