RIGS-EEG: DOES REFERENCE-INVARIANT GRAPH-SPECTRAL RECONSTRUCTION IMPROVE MISSING EEG CHANNEL RECOVERY?

Anonymous authors
Paper under double-blind review

ABSTRACT

Whole-channel missingness removes every temporal observation from an electrode within the analysis window. Reconstruction must therefore borrow information from other electrodes, but that information is shaped at the same time by the common reference, spatial coverage, and observation quality. We propose RIGS-EEG to handle these coupled factors in sequence: it first separates the common reference from relative potentials, then builds an initial estimate with a reliability-weighted spatial kernel, and finally refines missing channels through five graph-time-frequency unrolled steps. A final hard projection preserves all observed samples exactly, while the entire model contains only 970 trainable parameters. Under a fixed mask and reliability weights that are invariant to a common shift, its output is equivariant to additive common-reference changes. We compare RIGS-EEG with 14 representative baselines on PhysioNet, SEED-IV, and MCD under 25%, 50%, and 75% whole-channel missingness. At 75% missingness, the mean NRMSEs are 0.0723, 0.1274, and 0.0922 on the three datasets, respectively. In the fixed-classifier evaluation on PhysioNet, balanced accuracy improves by 3.44-8.37 percentage points over corrupted inputs and Cohen's Kappa by 0.068-0.122; ablations further locate the roles of the spatial, spectral, and reliability components. The results show that explicitly assigning roles to reference, geometry, and spectral structure can recover missing EEG with very few learned parameters while retaining signal structure useful for recognition.

Keywords: EEG reconstruction; missing-channel imputation; graph-spectral unrolling; reliability weighting; reference equivariance

1 INTRODUCTION

If an electrode has no data for an entire window, where can its waveform come from? Scattered missing samples can often borrow information from neighboring time points of the same channel. A failed channel loses that whole timeline. Reconstruction is therefore no longer about drawing a smooth curve through a gap; it must infer the time-varying potential at the target location from other electrodes. Which electrodes should be trusted, under what reference their potential differences should be expressed, and whether the recovered signal remains useful to an existing recognizer are three sides of the same problem.

Existing EEG reconstruction methods span spatial interpolation, spatiotemporal correlation, low-rank completion, and more recent attention-, diffusion-, and masked-learning approaches (Perrin et al., 1987; 1989; Ramakrishnan et al., 2016; Bahador et al., 2021; Duan et al., 2021; Zhang et al., 2022; 2025; Liu et al., 2012; Chen et al., 2013; Tashiro et al., 2021; Chien et al., 2022; Mohammadi et al., 2024). Yet whole-channel loss still leaves three coupled questions: how the common reference is handled, how spatial information reaches a missing electrode, and how observed signals are preserved. RIGS-EEG is designed around these three decisions; detailed comparisons with prior methods are provided in Related Work and Appendix A.

The first difficulty comes from how EEG is measured. Recorded potentials are always relative to a reference; changing that reference shifts channel values without necessarily changing inter-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

electrode potential differences (Yao, 2001; Yao et al., 2019; Hu et al., 2019; Chella et al., 2016). If a reference is estimated from all channels after zero filling, even the number of missing electrodes can change the estimate. We instead estimate the common component only from visible electrodes, leave relative potentials to the reconstruction stage, and require one testable property: under the stated conditions, adding the same reference waveform to the input should produce the same global shift at the output.

The second difficulty is that visible electrodes do not contribute equally. A nearby electrode is not automatically reliable, and an unusually large-amplitude observation can influence several missing sites if propagation is left unconstrained. At high missing rates, a target electrode may even have no visible local neighbor. Simply increasing smoothness can suppress oscillation, but it can also erase genuine local differences. We therefore separate where an initial estimate comes from from how that estimate is refined: a broader spatial kernel first gathers available observations, then graph and spectral structure guide the updates. Reliability weights regulate propagation; they do not label high-amplitude signals as artifacts.

RIGS-EEG follows a compact, structured reconstruction pipeline that separates the common reference, builds a reliability-aware spatial initialization, and refines missing EEG through graph–time-frequency unrolling (Hammond et al., 2011; Shuman et al., 2013; Ortega et al., 2018; Gregor et al., 2010; Monga et al., 2021). The model contains only 970 trainable parameters, while geometry, transforms, and observation constraints remain explicit; the full computation is detailed in Fig. 1 and the Method section.

The main contributions are summarized as follows:

- We separate the common reference from channel-relative potentials and state the conditions under which reference equivariance holds.
- We organize reliability-aware spatial initialization, graph-time-frequency shrinkage, and observation consistency into a five-step unrolled network with only 970 trainable parameters.
- We evaluate this structured reconstruction on three datasets using frozen-classifier tests and component ablations under cross-subject and cross-session protocols.

2 Related Work

Spatial and spatiotemporal interpolation use electrode geometry and observed-channel correlations (Perrin et al., 1987; 1989; Ramakrishnan et al., 2016; Bahador et al., 2021). Low-rank and tensor completion exploit redundancy across channels and time (Duan et al., 2021; Liu et al., 2012; Chen et al., 2013). Graph methods make electrode relationships explicit (Cao et al., 2025a;b; Hammond et al., 2011; Shuman et al., 2013; Zhu et al., 2003), while masked representation learning and diffusion model richer dependencies (Zhang et al., 2025; Tashiro et al., 2021; Chien et al., 2022; Mohammadi et al., 2024). Our approach combines a dense geometric initialization with sparse graph-spectral refinement.

Reference choice changes recorded EEG potentials (Yao, 2001; Yao et al., 2019; Hu et al., 2019; Chella et al., 2016). Rather than estimating a physical infinity reference, we require the reconstruction to shift consistently under an additive common reference. Algorithm unrolling (Gregor et al., 2010; Monga et al., 2021) gives a finite sequence in which reference handling, reliability, spectral correction, and observation consistency can be specified separately. Extended comparisons are in Appendix B.

3 Methodology

Accordingly, RIGS-EEG organizes reconstruction as a structured sequence of common-reference separation, reliability-aware spatial initialization, and iterative graph-time-frequency refinement.

As shown in Fig. 1, RIGS-EEG first builds a reliability-aware spatial initialization in the relative-potential domain, then refines it through five independently parameterized graph-time-frequency updates. After the final step, the common reference is restored and an observation projection enforces

consistency with measured signals. The full reconstruction path uses only 970 trainable parameters, concentrated in the iterative correction and spectral-gain modules.

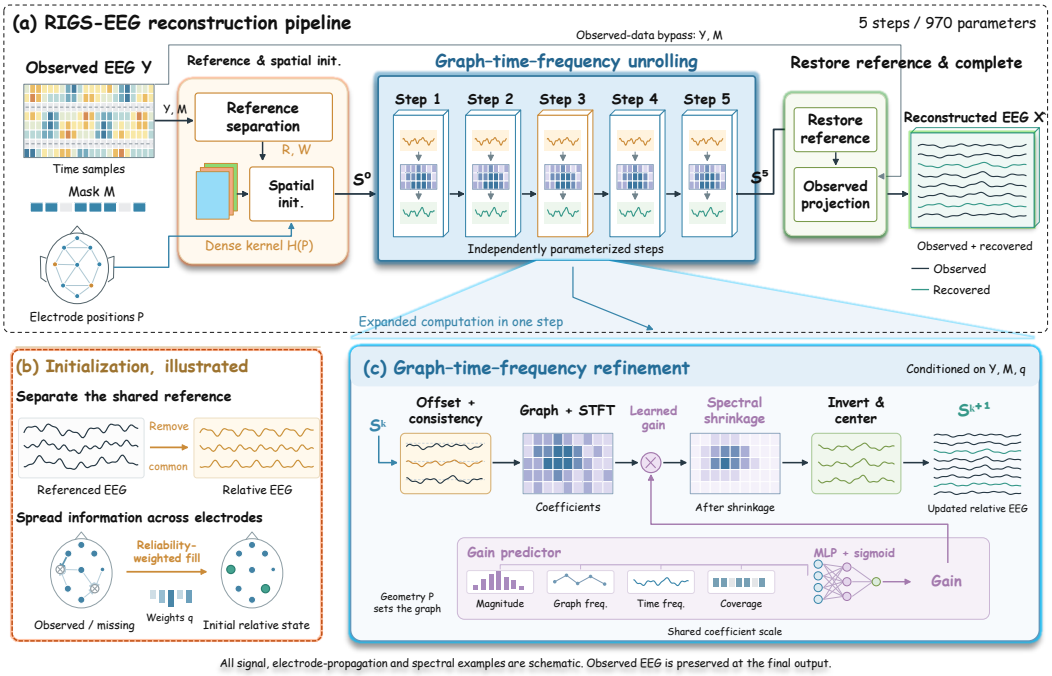

Figure 1: Overall architecture and key mechanisms of RIGS-EEG. (a) Reference-aware spatial initialization, five-step graph-time-frequency unrolling, reference restoration, and observation projection. (b) Common-reference separation and reliability-weighted spatial interpolation; gray hollow markers indicate missing electrodes and green markers indicate initialized potentials. (c) A single unrolled step performs offset estimation, weighted data consistency, graph-STFT transformation, conditional spectral shrinkage, inverse transformation, and channel centering. Waveforms, spatial propagation, and spectra are schematic, and the two spectral-coefficient panels share the same color scale.

3.1 Problem Definition and Shared Scaling

Let the complete EEG window be $X \in \mathbb{R}^{C\times T}$, the binary mask M indicate observed positions, and the electrode coordinates be $P \in \mathbb{R}^{C\times 3}$:

$$Y = M \odot X, \qquad \widehat{X} = F_\theta(Y, M, P). \tag{1}$$

During training, part of a complete window is artificially masked; at inference, only Y, M, P are provided, and data consistency is enforced at observed locations. All three datasets retain their native sampling rates, use a 1-45 Hz band-pass filter, and are segmented into 4 s windows with 50% overlap. Normalization statistics are estimated only from the training set, with one shared center and scale for all channels, so that per-channel standardization does not distort relative amplitudes.

3.2 Reliability-Weighted Reference Separation

A high-amplitude abnormal channel can spread erroneous information to multiple missing electrodes. To reduce this effect, we estimate channel reliability from the observed signal. Let a_c denote the RMS amplitude of a channel relative to the median reference, and a_{med} be the median across channel amplitudes. Then

$$q_c = clip\left[\left(\frac{\kappa \max\left(a_{med}, \varepsilon\right)}{\max\left(a_c, \varepsilon\right)}\right)^2, q_{\min}, 1\right], \qquad W_{ct} = M_{ct} q_c. \tag{2}$$

where $\kappa = 6$. This weight does not clip the observed waveform; it only reduces the contribution of high-amplitude channels to reference estimation and spatial propagation. PhysioNet uses $q_c = 1$, whereas SEED-IV and MCD enables Eq. (**??**).

We then estimate the common reference at each time point:

$$d_t = \sum_c W_{ct}, \qquad r_t = \frac{\sum_c W_{ct} Y_{ct}}{d_t}, \qquad R = M \odot \left(Y - \mathbf{1} r^\top\right). \tag{3}$$

Only actually observed samples contribute to the reference estimate, so placeholders at missing channels cannot pull the common component toward zero. The subsequent updates work mainly on relative potentials, while the common offset is handled through a separate path.

3.3 Reliability-Weighted Initialization on the Electrode Graph

A symmetric four-nearest-neighbor graph is built from three-dimensional electrode distances. Its Laplacian eigenvectors U and normalized eigenvalues $\overline{\lambda}$ define a fixed graph-spectral coordinate system (Hammond et al., 2011; Shuman et al., 2013; Ortega et al., 2018; Defferrard et al., 2016; Zhong et al., 2020; Hou et al., 2022).

When missingness is high, a local neighborhood may become entirely unavailable. We therefore introduce a dense Gaussian kernel H, giving

$$S^0 = \Pi\left(\left[H(W \odot R)\right] \oslash \left[HW\right]\right), \qquad \Pi(Z) = Z - \mathbf{1}\frac{\mathbf{1}^\top Z}{C}. \tag{4}$$

The dense kernel gives a missing electrode a usable initial state even when none of its local neighbors is visible (Narang et al., 2013a;b; Wang et al., 2015; Anis et al., 2016). The sparse graph used later then constrains the directions of refinement. Put simply, the first structure answers where usable information comes from, and the second how that information is organized.

3.4 Graph-Spectral Unrolled Reconstruction

Starting from S^0, the model unfolds $K = 5$ updates. At step k, it first estimates the common offset

$$b_t^k = \frac{\sum_c W_{ct}\left(Y_{ct} - S_{ct}^k\right)}{d_t}. \tag{5}$$

and performs the data-consistency update

$$\widetilde{S}^k = \Pi\left[S^k - \eta_k W \odot \left(S^k + \mathbf{1}\left(b^k\right)^\top - Y\right)\right], \qquad \eta_k = \sigma\left(\alpha_k\right). \tag{6}$$

In this way, common reference variation is mainly absorbed by b^k, while S^k focuses on relative structure across channels.

The updated state is then projected onto the graph basis and transformed with the STFT:

$$Z^k = STFT\left(U^\top \widetilde{S}^k\right), \qquad \rho = \frac{\sum_{c,t} W_{ct}}{CT}. \tag{7}$$

For each graph mode, temporal frequency, and frame, a lightweight MLP predicts a shrinkage gain

$$g_{jf\tau}^k = \sigma\left[MLP_k\left(\log\left(1 + \left|Z_{jf\tau}^k\right|\right), \overline{\lambda}_j, \overline{f}, \rho\right)\right], \tag{8}$$

and updates the state as

$$S^{k+1} = \Pi \left[UISTFT \left(g^k \odot Z^k \right) \right]. \tag{9}$$

The gain module changes only the magnitude of each complex coefficient while preserving its phase. Graph frequency, temporal frequency, coefficient magnitude, and observation coverage jointly determine the adjustment. Each step therefore behaves more like conditioned spectral shrinkage than unconstrained waveform generation. Each of the five unrolled steps contains 194 trainable parameters, yielding 970 in total; the parameter breakdown is provided in Appendix C.

After K steps, we re-estimate b^K and apply a hard observation projection to preserve all observed samples and fill only the missing positions:

$$\widehat{X} = M \odot Y + (1 - M) \odot \left(S^K + \mathbf{1} \left(b^K \right)^\top \right). \tag{10}$$

This output projection is used in all reported experiments. Reliability weights regulate reference estimation, spatial initialization, and internal data-consistency updates, while the final projection copies every observed value directly from the input. Only missing entries are replaced by the reconstruction.

The architecture is reference equivariant. If an arbitrary common reference waveform a is added to the input, then

$$F_\theta \left(M \odot \left(Y + \mathbf{1} a^\top \right), M, P \right) = F_\theta(Y, M, P) + \mathbf{1} a^\top. \tag{11}$$

In other words, changing the common reference produces the corresponding global shift at the output, while the recovered relative potentials remain unchanged. This statement assumes a fixed mask and reliability weights that are invariant to the common shift.

3.5 Training Objective and Configuration

The reconstruction loss is computed only at missing positions. SEED-IV uses channel-normalized RMSE:

$$\mathcal{L}_N = \frac{1}{|\mathcal{C}_\Omega|} \sum_{c \in \mathcal{C}_\Omega} \frac{\sqrt{|\Omega_c|^{-1} \sum_{t \in \Omega_c} \left(\widehat{X}_{ct} - X_{ct} \right)^2}}{\max \left(s_c, 10^{-8} \right)}, \tag{12}$$

where s_c is the dynamic range of the ground-truth channel. This normalization prevents high-amplitude channels from dominating training simply because of scale. PhysioNet uses MSE on missing positions together with spectral-magnitude and spatial-difference auxiliary losses.

Training uses 25% mixed missingness, sampled from random points, contiguous time segments, and whole-channel masks. Checkpoints are selected by validation NRMSE at 25% channel missingness and then evaluated at 25%, 50%, and 75%. AdamW is used with learning rate 3×10^{-4}, for at most 20 epochs, with early stopping after 5 epochs without validation improvement.

4 Experiments

The experiments follow three connected questions: can signals at missing electrodes be reconstructed accurately, can the recovered EEG still support a frozen recognizer, and which design choices produce the gains? PhysioNet, SEED-IV, and MCD span different recording tasks and split protocols. Three missing ratios probe progressively sparser observations, while the PhysioNet frozen-classifier study and dataset-specific ablations test downstream utility and component effects.

Table 1: MCD whole-channel reconstruction under different missing ratios. Full results are provided in Appendix B.

Method	25%	50%	75%
IDW	0.0872 ± 0.0177	0.0890 ± 0.0183	0.0944 ± 0.0197
HALRTC	0.0897 ± 0.0173	0.0907 ± 0.0176	0.0938 ± 0.0180
CT	0.0896 ± 0.0175	0.0904 ± 0.0179	0.0938 ± 0.0188
DMAE-EEG	0.0870 ± 0.0170	0.0890 ± 0.0162	0.1030 ± 0.0135
RIGS-EEG	0.0861 ± 0.0177	0.0883 ± 0.0184	0.0922 ± 0.0192

4.1 Datasets and Protocol

PhysioNet EEGMMIDB provides 64-channel motor execution and motor imagery recordings at 160 Hz; we use runs R03-R14 (Goldberger, 2000; Schalk et al., 2004). The 106 valid subjects are split by subject with seed 42 into 74 training, 10 validation, and 22 test subjects, and all runs from one subject stay in the same split. SEED-IV uses 62-channel raw time-domain EEG sampled at 200 Hz (Zheng et al., 2015; 2018). For its 15 subjects, sessions 1, 2, and 3 are used for training, validation, and testing, respectively, with 24 trials per session. PhysioNet therefore tests unseen subjects, whereas SEED-IV tests new sessions from subjects already seen during training; the two protocols target different forms of generalization.

MCD is an in-house, currently non-public EEG dataset collected by our research group, comprising reading, memory, and IQ recordings from 22 deaf and 23 normal-hearing participants. All participants provided informed consent, and data collection was conducted in accordance with applicable ethical requirements. Each stage contains 60 channels × 1600 samples × 64 trials at 200 Hz. We use a group-stratified split of 31 training, 4 validation, and 10 test subjects, keeping all three stages from the same participant in the same split. Reading and memory were recorded during continuous cap wear, whereas the cap was re-donned for the IQ stage; task and cap-placement changes therefore co-occur and cannot be separated in the present setup. Geometric input uses standard-template coordinates matched to the channel layout.

Data splitting is completed before preprocessing. Independent recordings or trials are then band-pass filtered at 1-45 Hz and segmented into 4 s windows with 50% overlap. All channels share one robust center and scale estimated only from the training set, and the same transform is used for validation and test data. Whole-channel loss is specified by an explicit binary mask, with paired test masks shared across methods under each evaluation condition. NRMSE is computed only at missing positions: RMSE is first obtained for each affected channel, normalized by the peak-to-peak range of the complete target window, and then aggregated across channels and windows. The target dynamic range is used only for scoring, never as a model input.

We compare against 14 methods: SS, IDW, KRIGING, STC, STDC, HALRTC, GLAE, CT, SAITS, CSDI, PATCHTST, GH, SRL-ITV, and DMAE-EEG. Learning-based baselines are built from the methods and public implementations described in their original papers, then trained and tested under the same data splits, missing masks, and evaluation protocol. For consistency, only the necessary data interfaces, input dimensions, and mask formats are adjusted; the core model structures and training objectives are left unchanged.

RIGS-EEG uses five unrolled steps. PhysioNet follows the v1 objective: MSE on missing positions plus spectral and spatial auxiliary losses. SEED-IV and MCD use the v2 normalized reconstruction loss with reliability weighting. AdamW is used with learning rate 3×10^{-4}, weight decay 10^{-4}, batch size 2, and at most 20 epochs; training stops after five validation epochs without improvement. One checkpoint is selected from validation for each dataset and used to evaluate all three missing ratios. Ablations follow the same dataset-specific training configuration.

Tables 2 and 4, together with Appendix Table 5, report mean $\pm$ sample standard deviation across subjects, corresponding to 22, 15, and 10 test subjects, respectively; MCD first averages the three stages equally within each subject. Reconstruction models use a single training seed, 42, so these standard deviations describe subject variability. Standard deviations for BA and Kappa come from five classifier/mask seeds, while reconstruction weights remain fixed across those repeated evaluations.

Table 2: Whole-channel reconstruction (NRMSE, mean ± subject SD). Lower is better.

Method	PhysioNet			SEED-IV		
	25%	50%	75%	25%	50%	75%
SS	0.0542 ± 0.0152	0.0609 ± 0.0168	0.0793 ± 0.0212	0.4178 ± 1.2595	0.4944 ± 1.5042	0.6736 ± 2.0729
IDW	0.0683 ± 0.0087	0.0705 ± 0.0088	0.0767 ± 0.0093	0.2864 ± 0.6454	0.2954 ± 0.6604	0.3050 ± 0.6551
KRIGING	0.0785 ± 0.0252	0.0894 ± 0.0291	0.0864 ± 0.0245	1.2404 ± 4.2701	1.3909 ± 4.7577	1.0064 ± 3.2866
STC	0.0924 ± 0.0093	0.0931 ± 0.0094	0.0954 ± 0.0096	0.3290 ± 0.7446	0.3337 ± 0.7471	0.3417 ± 0.7442
STDC	0.0895 ± 0.0090	0.0950 ± 0.0094	0.1032 ± 0.0101	0.3232 ± 0.6274	0.3534 ± 0.6945	0.3838 ± 0.6781
HALRTC	0.0890 ± 0.0091	0.0946 ± 0.0095	0.1027 ± 0.0102	0.3239 ± 0.6349	0.3538 ± 0.6758	0.3824 ± 0.6453
GLAE	0.0517 ± 0.0119	0.0590 ± 0.0110	0.0958 ± 0.0079	0.4156 ± 1.2777	0.3167 ± 0.8699	0.2315 ± 0.4539
CT	0.0548 ± 0.0154	0.0686 ± 0.0105	0.0826 ± 0.0169	0.1137 ± 0.0914	0.1324 ± 0.4899	0.1465 ± 0.1214
SAITS	0.0521 ± 0.0116	0.0631 ± 0.0114	0.1066 ± 0.0210	0.1387 ± 0.2387	0.1397 ± 0.1765	0.1561 ± 0.1055
CSDI	0.0655 ± 0.0104	0.0721 ± 0.0101	0.0883 ± 0.0113	0.1096 ± 0.0269	0.1170 ± 0.0289	0.1367 ± 0.0348
PATCHTST	0.0589 ± 0.0175	0.0692 ± 0.0265	0.0993 ± 0.0152	0.1146 ± 0.0365	0.1348 ± 0.0246	0.1573 ± 0.0401
GH	0.0544 ± 0.0115	0.0597 ± 0.0112	0.0788 ± 0.0109	0.2373 ± 0.5886	0.2639 ± 0.6536	0.2861 ± 0.6559
SRL-ITV	0.0555 ± 0.0094	0.0701 ± 0.0074	0.1167 ± 0.0067	0.0884 ± 0.0214	0.1013 ± 0.0386	0.1295 ± 0.0492
DMAE-EEG	0.0729 ± 0.0102	0.0757 ± 0.0103	0.0936 ± 0.0145	0.1639 ± 0.1125	0.1626 ± 0.1024	0.1774 ± 0.0702
RIGS-EEG	0.0509 ± 0.0120	0.0573 ± 0.0113	0.0723 ± 0.0104	0.0859 ± 0.0279	0.0991 ± 0.0338	0.1274 ± 0.0443

4.2 Whole-Channel Reconstruction

Table 2 reports the PhysioNet and SEED-IV methods and missing ratios; Appendix Table 5 reports MCD. RIGS-EEG has the lowest mean NRMSE in all nine conditions. At 75% missingness its means are 0.0723 on PhysioNet, 0.1274 on SEED-IV, and 0.0922 on MCD. The advantages over the strongest baseline are 5.74%, 1.62%, and 1.71%, respectively. Each dataset is trained separately; these numbers do not measure transfer to an unseen dataset.

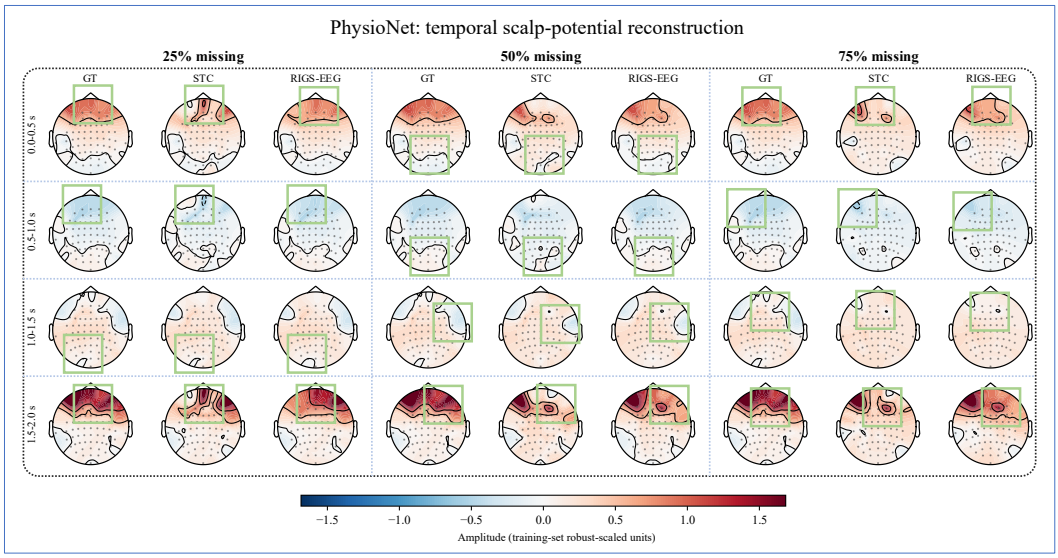

Figure 2: Scalp-potential reconstruction on PhysioNet at different channel-missing ratios. For the same representative EEG window, GT, STC, and RIGS-EEG are compared under 25%, 50%, and 75% whole-channel missingness. All panels use the same color scale, and amplitudes are shown in training-set shared robust-scaled units. Green boxes mark local regions for visual comparison.

Beyond average error, Figure 2 compares the spatial reconstruction on PhysioNet. The same representative window and a shared color scale are used throughout. As missingness rises from 25% to 75%, both STC and RIGS-EEG must recover the full scalp potential from progressively sparser observations; across several time slices, RIGS-EEG retains the main GT spatial pattern and local potential structure more closely.

Table 3: Frozen-classifier results on PhysioNet at 25% channel missingness (mean ± SD).

Method	Balanced accuracy (%)				Cohen's κ			
	MM-H	MM-HF	MI-H	MI-HF	MM-H	MM-HF	MI-H	MI-HF
Complete EEG	74.54 ± 1.55	49.36 ± 1.76	75.38 ± 0.32	54.89 ± 3.32	0.492 ± 0.031	0.326 ± 0.024	0.507 ± 0.006	0.397 ± 0.044
Corrupted EEG	68.07 ± 0.49	39.83 ± 1.71	71.57 ± 0.36	45.39 ± 1.69	0.362 ± 0.010	0.199 ± 0.023	0.431 ± 0.008	0.271 ± 0.022
SS	73.62 ± 1.06	47.94 ± 1.95	74.60 ± 0.76	53.01 ± 2.62	0.475 ± 0.021	0.307 ± 0.027	0.495 ± 0.015	0.370 ± 0.035
IDW	73.09 ± 1.67	47.06 ± 1.90	74.31 ± 0.58	51.35 ± 2.35	0.463 ± 0.034	0.294 ± 0.026	0.486 ± 0.012	0.348 ± 0.031
KRIGING	73.24 ± 0.93	47.12 ± 1.55	74.56 ± 1.21	52.85 ± 2.66	0.466 ± 0.019	0.295 ± 0.021	0.492 ± 0.024	0.368 ± 0.035
STC	72.06 ± 1.00	45.41 ± 1.84	73.58 ± 0.40	49.34 ± 2.11	0.440 ± 0.020	0.271 ± 0.025	0.470 ± 0.008	0.319 ± 0.028
STDC	72.30 ± 0.85	45.68 ± 1.25	73.60 ± 0.35	50.23 ± 1.98	0.457 ± 0.032	0.282 ± 0.036	0.484 ± 0.015	0.327 ± 0.024
HALRTC	72.54 ± 0.69	46.53 ± 1.01	73.89 ± 0.29	50.67 ± 1.22	0.460 ± 0.033	0.291 ± 0.037	0.489 ± 0.011	0.330 ± 0.042
GLAE	73.28 ± 1.26	47.74 ± 2.16	74.32 ± 0.49	52.78 ± 2.81	0.469 ± 0.025	0.307 ± 0.029	0.492 ± 0.010	0.371 ± 0.037
CT	73.77 ± 1.58	47.96 ± 2.03	74.18 ± 0.64	52.69 ± 2.41	0.476 ± 0.034	0.302 ± 0.029	0.491 ± 0.017	0.365 ± 0.039
SAITS	73.12 ± 1.43	47.83 ± 1.77	74.34 ± 0.59	52.91 ± 2.36	0.465 ± 0.029	0.305 ± 0.024	0.492 ± 0.012	0.370 ± 0.031
CSDI	72.38 ± 1.18	47.50 ± 2.21	74.50 ± 0.66	51.72 ± 2.67	0.453 ± 0.024	0.301 ± 0.030	0.490 ± 0.013	0.356 ± 0.035
PATCHTST	73.01 ± 1.93	47.21 ± 2.88	74.05 ± 0.73	52.44 ± 2.61	0.471 ± 0.082	0.290 ± 0.047	0.479 ± 0.020	0.361 ± 0.043
GH	73.80 ± 1.59	47.78 ± 2.02	74.31 ± 0.80	52.84 ± 2.44	0.479 ± 0.032	0.302 ± 0.028	0.489 ± 0.016	0.370 ± 0.032
SRL-ITV	72.85 ± 1.60	47.58 ± 2.16	74.15 ± 0.88	51.90 ± 2.37	0.461 ± 0.032	0.300 ± 0.029	0.486 ± 0.018	0.358 ± 0.031
DMAE-EEG	72.27 ± 2.29	46.37 ± 2.01	73.54 ± 0.68	51.16 ± 2.49	0.451 ± 0.046	0.287 ± 0.027	0.474 ± 0.014	0.349 ± 0.033
RIGS-EEG	74.15 ± 1.62	48.20 ± 1.91	75.01 ± 0.86	53.56 ± 2.43	0.484 ± 0.033	0.310 ± 0.026	0.499 ± 0.017	0.380 ± 0.032

4.3 Task Information Recovery

Lower numerical error is useful, but a more practical question follows: can the recovered signal still be used by the original recognizer? We train an EEGNet-style classifier (Lawhern et al., 2018) on complete training signals, select weights on the validation set, freeze the classifier, and then evaluate the same checkpoint on complete, corrupted, and reconstructed test inputs. MM-H and MI-H are left-versus-right fist binary classification for motor execution and motor imagery, respectively; MM-HF and MI-HF add both-fists and both-feet classes, yielding four-way classification. Windows are aligned to events so that they do not cross class boundaries. Complete EEG and Corrupted EEG provide the intact-input reference and the directly damaged-input control.

Across all four tasks, RIGS-EEG achieves the highest mean BA and Kappa among reconstruction methods. BA is 74.15%, 48.20%, 75.01%, and 53.56%, improving over corrupted inputs by 6.08, 8.37, 3.44, and 8.17 percentage points. Kappa increases by 0.122, 0.111, 0.068, and 0.109. Compared with the strongest reconstruction baseline for each task, BA improves by a further 0.35, 0.24, 0.41, and 0.55 points. Two of those margins are small, but the direction is consistent across all four tasks. Because the classifier remains frozen throughout, the differences come from the reconstructed input.

RIGS-EEG recovers 86.0%–94.0% of the BA loss caused by channel missingness, remaining only 0.37–1.33 percentage points below complete EEG. These results show that reconstruction preserves information useful to the frozen classifier.

4.4 Component Ablation

To trace the source of the gains, we retrain four variants. w/o reference removes common-reference handling; w/o graph removes the graph structure and replaces spatial initialization; w/o spectral skips spectral updates; and w/o reliability sets all reliability weights to 1. On PhysioNet, the graph and spectral variants also remove their corresponding auxiliary losses, so the comparison reflects the contribution of the full component.

The graph component has an especially large effect on both public datasets. At 25% missingness, removing it raises NRMSE from 0.0509 to 0.1092 on PhysioNet and from 0.0859 to 0.1957 on SEED-IV; at 75%, the corresponding changes are 0.0723 to 0.1114 and 0.1274 to 0.2051. Removing spectral updates also increases error at every missing ratio, reaching 0.0893 and 0.1610 at 75% on the two datasets. Spatial coverage and temporal-frequency constraints contribute different information when a whole channel is absent, and the ablation makes that division visible in the measured error.

The benefit of reference handling grows with the data condition. On SEED-IV at 75% missingness, the full model reduces NRMSE from 0.1361 for w/o reference to 0.1274, a 6.39% reduction; MCD shows the same direction at all three ratios. On PhysioNet at 25%, both variants are 0.0509, while

Table 4: Component ablation (NRMSE, mean ± subject SD).

Variant	PhysioNet			SEED-IV		
	25%	50%	75%	25%	50%	75%
Full model	0.0509 ± 0.0120	0.0573 ± 0.0113	0.0723 ± 0.0104	0.0859 ± 0.0279	0.0991 ± 0.0338	0.1274 ± 0.0443
w/o reference	0.0509 ± 0.0120	0.0576 ± 0.0112	0.0732 ± 0.0101	0.0862 ± 0.0206	0.0999 ± 0.0221	0.1361 ± 0.0261
w/o graph	0.1092 ± 0.0110	0.1096 ± 0.0110	0.1114 ± 0.0111	0.1957 ± 0.0676	0.1971 ± 0.0655	0.2051 ± 0.0719
w/o spectral	0.0857 ± 0.0090	0.0865 ± 0.0091	0.0893 ± 0.0095	0.1533 ± 0.0506	0.1559 ± 0.0518	0.1610 ± 0.0520
w/o reliability	–	–	–	0.2769 ± 0.7405	0.2925 ± 0.7499	0.3151 ± 0.7279

the full model has lower error at higher missingness. Explicit reference handling does not produce a large gain in every column, but its effect becomes easier to see when fewer electrodes remain.

SEED-IV also exposes the role of reliability weighting in reconstruction stability. Setting all weights to 1 raises mean NRMSE to 0.2769, 0.2925, and 0.3151, with subject standard deviations of 0.7405, 0.7499, and 0.7279. The full model reduces both the mean and the spread, showing that limiting the dominance of individual observations during spatial propagation is important on this dataset. On MCD, the graph-free and spectral-free variants are also worse than the full model at all three ratios. The components are not merely stacking small numerical gains; they control different parts of the information flow - source, update direction, and propagation weight.

5 Discussion and Limitations

Across the three datasets, the same reconstruction structure reaches low average error under different subjects, sessions, and recording stages, although the size of the advantage varies. The relative gap is more visible on PhysioNet at high missingness, while some SEED-IV and MCD settings show smaller mean differences. Ablations help explain this pattern: graph structure and spectral updates improve reconstruction fairly consistently, whereas reliability weighting has an especially strong effect on SEED-IV. The components therefore do not contribute the same amount to every recording; the model is best understood together with the observation conditions under which it operates.

The evaluation protocol also sets the boundary of these conclusions. PhysioNet and MCD are split by subject, whereas SEED-IV is split by session. In MCD, reading and memory are recorded during continuous cap wear, but the cap is donned again for the IQ stage; task and wearing condition therefore change together and cannot be disentangled here. MCD also uses template electrode coordinates rather than positions measured after each re-donning. These settings still cover meaningful acquisition variation, but the source of spatial mismatch would need a more targeted control experiment.

We use known artificial whole-channel masks, a single reconstruction-training seed, and subject standard deviations to describe individual variability. Stable mean rankings in the tables therefore do not characterize training-run uncertainty or real electrode failures in full, and no paired significance test is reported. Follow-up work can test multiple independent trainings, real failures, different reference remappings, and electrode positions measured at each session.

6 Conclusion

Whole-channel failure removes the target electrode's entire timeline, forcing reconstruction to rely on the remaining channels. RIGS-EEG addresses this with explicit common-reference separation, reliability-aware spatial initialization, five graph-time-frequency refinement steps, and observation consistency, using only 970 trainable parameters while satisfying reference equivariance under the stated conditions. Across nine whole-channel settings on PhysioNet, SEED-IV, and MCD, it achieves the lowest mean NRMSE reported in the tables, and the frozen-classifier study shows that most lost task performance can be recovered. Ablations further attribute the gains to spatial, spectral, reference, and reliability components. Together, these results show that assigning clear roles to reference, geometry, and spectral structure offers a compact and effective route to EEG reconstruction.

AI USE STATEMENT

Generative AI tools were used to assist with English translation and language editing, mathematical typesetting, and manuscript organization. They were not used to formulate research hypotheses or mathematical claims, develop the proposed methodology, design or conduct experiments, generate or process experimental data, implement the proposed method, interpret experimental results, or determine the scientific conclusions of this work. All AI-assisted content was reviewed and verified by the authors, who take full responsibility for the final manuscript.

REFERENCES

Shepard, Donald. "A two-dimensional interpolation function for irregularly-spaced data." Proceedings of the 1968 23rd ACM national conference. 1968.

Perrin F, Pernier J, Bertnard O, et al. Mapping of scalp potentials by surface spline interpolation[J]. Electroencephalography and clinical neurophysiology, 1987, 66(1): 75-81.

Perrin F, Pernier J, Bertrand O, et al. Spherical splines for scalp potential and current density mapping[J]. Electroencephalography and clinical neurophysiology, 1989, 72(2): 184-187.

Matheron G. " Principles of geostatistics", Economic Geology, 58, pp 1246-1266[J]. 1963.

Yao D. A method to standardize a reference of scalp EEG recordings to a point at infinity[J]. Physiological measurement, 2001, 22(4): 693-711.

Yao D, Qin Y, Hu S, et al. Which reference should we use for EEG and ERP practice?[J]. Brain topography, 2019, 32(4): 530-549.

Hu S, Yao D, Bringas-Vega M L, et al. The statistics of EEG unipolar references: derivations and properties[J]. Brain topography, 2019, 32(4): 696-703.

Chella F, Pizzella V, Zappasodi F, et al. Impact of the reference choice on scalp EEG connectivity estimation[J]. Journal of neural engineering, 2016, 13(3): 036016.

Bergmann T, Movshovich M, Shao Y, et al. Missing Data Gap Imputation Methods in Electroencephalogram (EEG) Signals: A Systematic Scoping Review[J]. Sensors, 2026, 26(8): 2431.

Ramakrishnan A G, Satyanarayana J V. Reconstruction of EEG from limited channel acquisition using estimated signal correlation[J]. Biomedical Signal Processing and Control, 2016, 27: 164-173.

Bahador N, Jokelainen J, Mustola S, et al. Reconstruction of missing channel in electroencephalogram using spatiotemporal correlation-based averaging[J]. Journal of Neural Engineering, 2021, 18(5): 056045.

Saba-Sadiya S, Alhanai T, Liu T, et al. EEG channel interpolation using deep encoder-decoder networks[C]//2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2020: 2432-2439.

El-Fiqi H, Kasmarik K, Bezerianos A, et al. Gate-layer autoencoders with application to incomplete EEG signal recovery[C]//2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019: 1-8.

Duan F, Jia H, Zhang Z W, et al. On the robustness of EEG tensor completion methods[J]. Science China Technological Sciences, 2021, 64(9): 1828-1842.

Zhang C, Han S, Zhang M. Single-channel EEG completion using Cascade Transformer[C]//2022 IEEE Biomedical Circuits and Systems Conference (BioCAS). IEEE, 2022: 605-609.

Cao Y, Wang J, Li Z, et al. Time-varying EEG signal reconstruction based on local graph signal smoothness[C]//ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2025: 1-5.

Cao Y, Li Z, Wang J, et al. Signal reconstruction from incomplete time-varying EEG via local and global graph learning[J]. Digital Signal Processing, 2025: 105647.

Li D, Zeng Z, Wang Z, et al. ESTformer: Transformer utilising spatiotemporal dependencies for electroencephalogram super-resolution[J]. Knowledge-Based Systems, 2025, 317: 113345.

Tang Y, Chen C, Chen D. Individual EEG super-resolution via ADMM-based coupled matrix decomposition towards long-term brain monitoring[J]. Biomedical Signal Processing and Control, 2025, 103: 107394.

Zhang Y, Yu Y, Li H, et al. DMAE-EEG: A pretraining framework for EEG spatiotemporal representation learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36(10): 17664-17678.

Liu J, Musialski P, Wonka P, et al. Tensor completion for estimating missing values in visual data[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(1): 208-220.

Chen Y L, Hsu C T, Liao H Y M. Simultaneous tensor decomposition and completion using factor priors[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 36(3): 577-591.

Che Z, Purushotham S, Cho K, et al. Recurrent neural networks for multivariate time series with missing values[J]. Scientific reports, 2018, 8(1): 6085.

Cao W, Wang D, Li J, et al. Brits: Bidirectional recurrent imputation for time series[J]. Advances in neural information processing systems, 2018, 31.

Yoon J, Jordon J, Schaar M. Gain: Missing data imputation using generative adversarial nets[C]//International conference on machine learning. PMLR, 2018: 5689-5698.

Tashiro Y, Song J, Song Y, et al. Csdi: Conditional score-based diffusion models for probabilistic time series imputation[J]. Advances in neural information processing systems, 2021, 34: 24804-24816.

Du W, Côté D, Liu Y. Saits: Self-attention-based imputation for time series[J]. Expert Systems with Applications, 2023, 219: 119619.

Nie Y, Nguyen N H, Sinthong P, et al. A time series is worth 64 words: Long-term forecasting with transformers[J]. arXiv preprint arXiv:2211.14730, 2022.

Du W, Wang J, Qian L, et al. Tsi-bench: Benchmarking time series imputation[J]. arXiv preprint arXiv:2406.12747, 2024.

Du W. PyPOTS: a Python toolbox for data mining on Partially-Observed Time Series[J]. arXiv preprint arXiv:2305.18811, 2023.

Hammond D K, Vandergheynst P, Gribonval R. Wavelets on graphs via spectral graph theory[J]. Applied and computational harmonic analysis, 2011, 30(2): 129-150.

Shuman D I, Narang S K, Frossard P, et al. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains[J]. IEEE signal processing magazine, 2013, 30(3): 83-98.

Narang S K, Gadde A, Ortega A. Signal processing techniques for interpolation in graph structured data[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013: 5445-5449.

Narang S K, Gadde A, Sanou E, et al. Localized iterative methods for interpolation in graph structured data[C]//2013 IEEE Global conference on signal and information processing. IEEE, 2013: 491-494.

Wang X, Liu P, Gu Y. Local-set-based graph signal reconstruction[J]. IEEE transactions on signal processing, 2015, 63(9): 2432-2444.

Anis A, Gadde A, Ortega A. Efficient sampling set selection for bandlimited graph signals using graph spectral proxies[J]. IEEE Transactions on Signal Processing, 2016, 64(14): 3775-3789.

Ortega A, Frossard P, Kovačević J, et al. Graph signal processing: Overview, challenges, and applications[J]. Proceedings of the IEEE, 2018, 106(5): 808-828.

Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering[J]. Advances in neural information processing systems, 2016, 29.

Zhu X, Ghahramani Z, Lafferty J D. Semi-supervised learning using gaussian fields and harmonic functions[C]//Proceedings of the 20th International conference on Machine learning (ICML-03). 2003: 912-919.

Zhong P, Wang D, Miao C. EEG-based emotion recognition using regularized graph neural networks[J]. IEEE Transactions on Affective Computing, 2020, 13(3): 1290-1301.

Hou Y, Jia S, Lun X, et al. GCNs-net: a graph convolutional neural network approach for decoding time-resolved eeg motor imagery signals[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 35(6): 7312-7323.

Gregor K, LeCun Y. Learning fast approximations of sparse coding[C]//Proceedings of the 27th international conference on international conference on machine learning. 2010: 399-406.

Monga V, Li Y, Eldar Y C. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing[J]. IEEE Signal Processing Magazine, 2021, 38(2): 18-44.

Chien H Y S, Goh H, Sandino C M, et al. Maeeg: Masked auto-encoder for eeg representation learning[J]. arXiv preprint arXiv:2211.02625, 2022.

Mohammadi Foumani N, Mackellar G, Ghane S, et al. Eeg2rep: enhancing self-supervised eeg representation through informative masked inputs[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2024: 5544-5555.

Lawhern V J, Solon A J, Waytowich N R, et al. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces[J]. Journal of neural engineering, 2018, 15(5): 056013.

Goldberger A L. Physiobank, physiotoolkkit, and physionet: components of a new research resource for complex physiologic signals[J]. circulation, 2000, 101(23): e215-e220.

Schalk G, McFarland D J, Hinterberger T, et al. BCI2000: a general-purpose brain-computer interface (BCI) system[J]. IEEE Transactions on biomedical engineering, 2004, 51(6): 1034-1043.

Zheng W L, Lu B L. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks[J]. IEEE Transactions on autonomous mental development, 2015, 7(3): 162-175.

Zheng W L, Liu W, Lu Y, et al. Emotionmeter: A multimodal framework for recognizing human emotions[J]. IEEE transactions on cybernetics, 2018, 49(3): 1110-1122.

A Additional Visualizations and Analysis

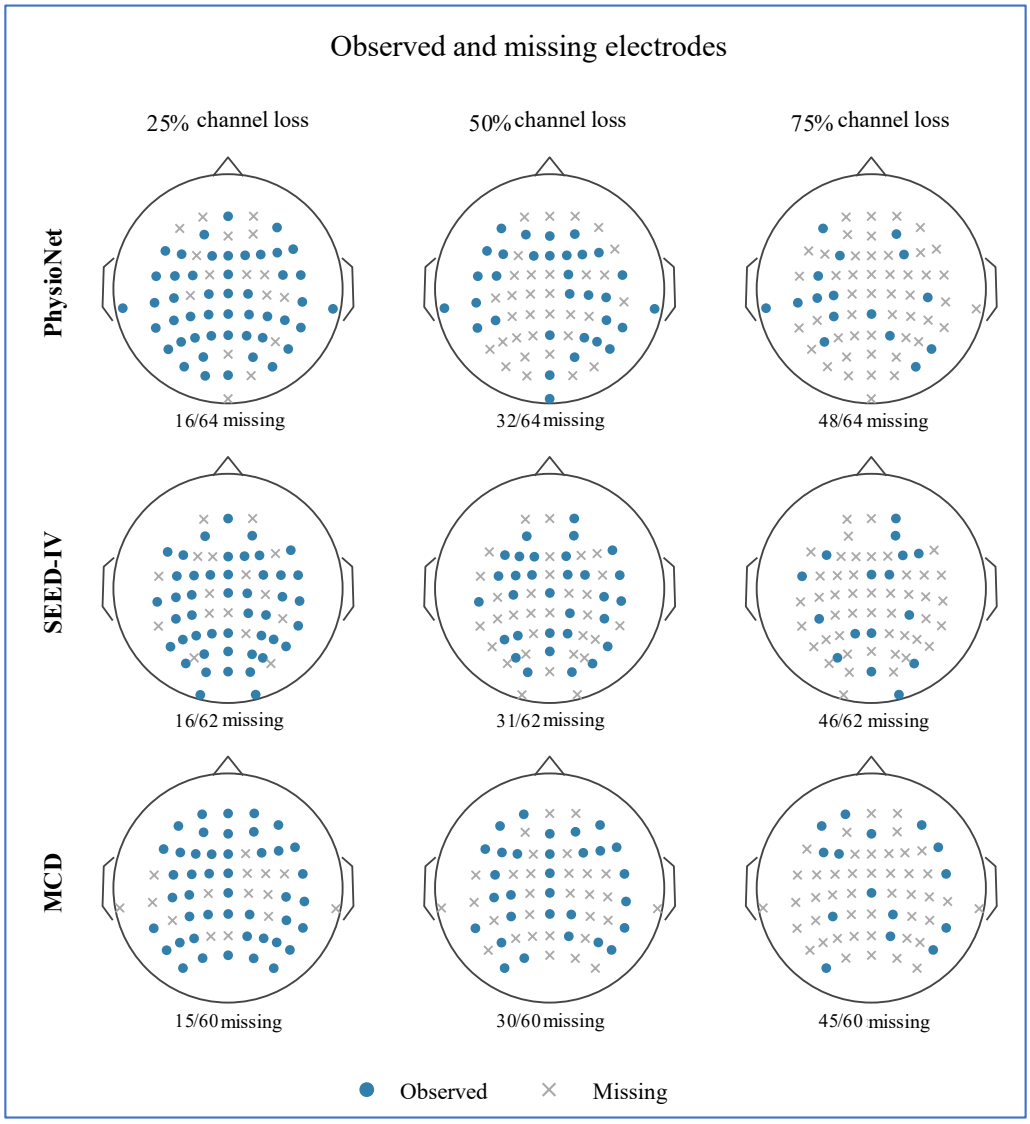

Figure 3: Observed and missing electrode distributions at different channel-missing ratios. Example masks are shown for PhysioNet, SEED-IV, and MCD at 25%, 50%, and 75% whole-channel missingness. Filled circles denote observed electrodes and crosses denote missing electrodes. The figure illustrates how spatial coverage changes as the missing ratio increases.

To make the loss of spatial coverage visible, Figure 3 shows example observed/missing electrode patterns for PhysioNet, SEED-IV, and MCD at 25%, 50%, and 75% whole-channel missingness. As missingness increases, the available spatial neighborhood becomes progressively sparser, making recovery based only on local interpolation more difficult.

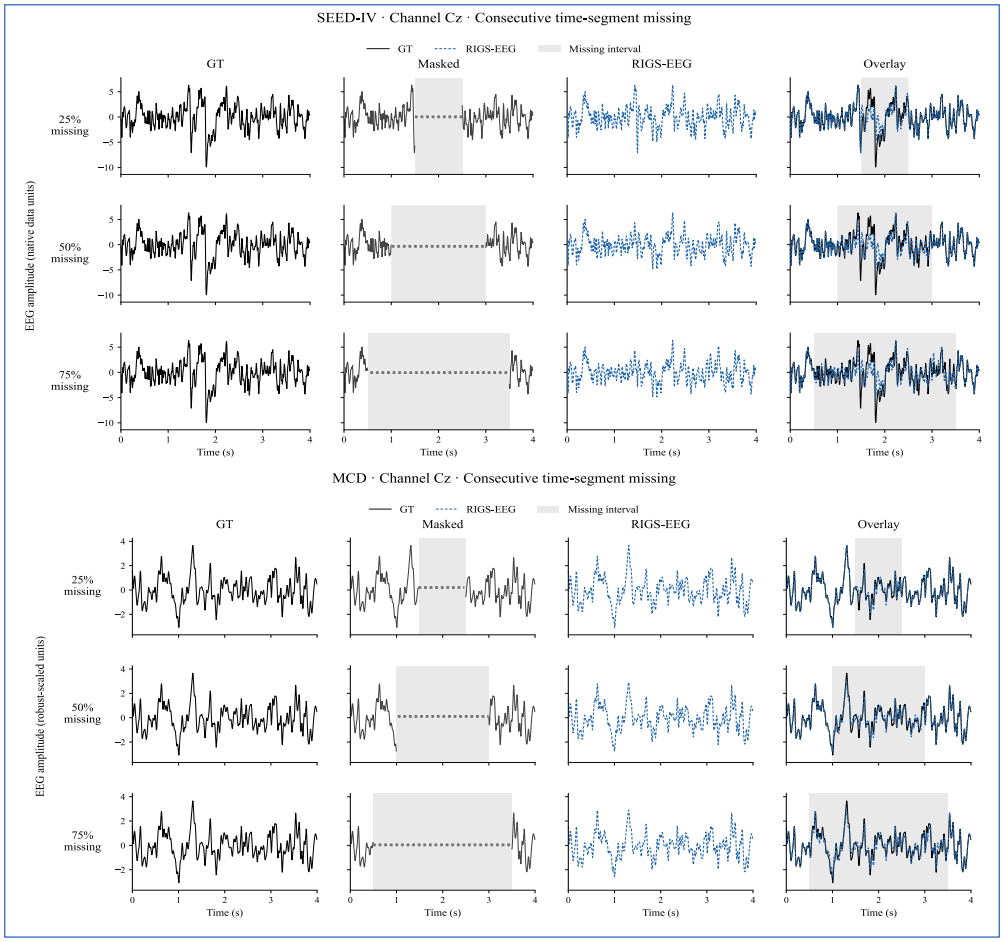

Figure 4: Single-channel examples under contiguous time-segment missingness on SEED-IV and MCD. For each dataset, a representative Cz channel is shown at 25%, 50%, and 75% missingness with GT, Masked, RIGS-EEG, and Overlay panels. Gray regions indicate missing intervals. This qualitative experiment examines local temporal waveform recovery and uses a different missingness protocol from the whole-channel setting in the main tables.

To look beyond spatial error at local temporal behavior, Fig. 4 shows waveform reconstruction on a representative Cz channel from SEED-IV and MCD under 25%, 50%, and 75% contiguous time-segment missingness. Gray shading marks the missing interval, and the overlay directly compares GT with the RIGS-EEG output. This figure uses contiguous temporal gaps only for qualitative inspection of waveform recovery; it does not correspond to the whole-channel missingness protocol used in the main tables.

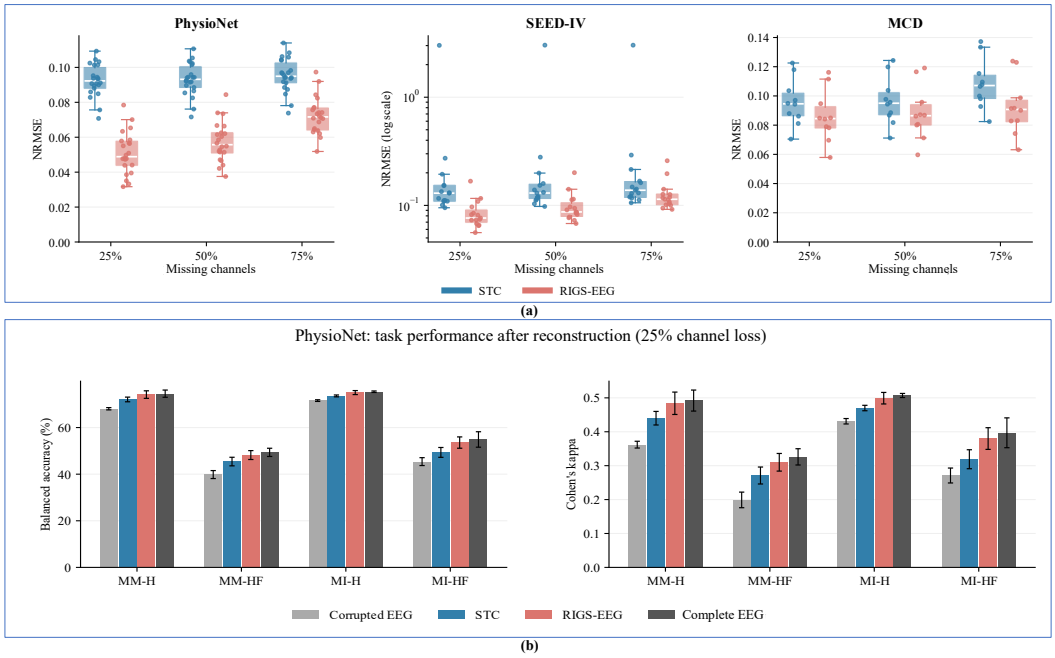

Figure 5: Reconstruction-error distributions and downstream task recovery. (a) Subject-level NRMSE distributions for STC and RIGS-EEG on PhysioNet, SEED-IV, and MCD under 25%, 50%, and 75% whole-channel missingness; SEED-IV uses a logarithmic y-axis to accommodate its wider range. (b) Balanced Accuracy and Cohen's Kappa from the frozen classifier on PhysioNet with 25% channel missingness, comparing Corrupted EEG, STC, RIGS-EEG, and Complete EEG. Error bars show the variability of the corresponding repeated evaluations.

Fig. 5(a) further shows subject-level NRMSE distributions for STC and RIGS-EEG across all three datasets and missing ratios, complementing the means and standard deviations in the tables. Fig. 5(b) summarizes the frozen-classifier results on PhysioNet with 25% channel missingness. Compared with directly corrupted EEG, RIGS-EEG reconstruction restores part of the BA and Cohen's Kappa across all four tasks and remains close to the complete-EEG reference.

Three pieces of evidence now form one chain: the full model has the lowest mean NRMSE in all nine reconstruction conditions, the four frozen PhysioNet classifiers recover 86.0%-94.0% of the BA lost to missingness, and the ablations link performance changes to spatial, spectral, reference, and reliability handling. The five-step network needs only 970 trainable parameters, suggesting that arranging computation around the structure of EEG measurement can matter more than simply adding learned capacity.

The scope is equally clear. The reconstruction model uses one training seed, the missing masks are artificial and known in advance, and the mean differences have not been tested with paired significance tests. Future evaluation should include real electrode failures, multiple independent training runs, explicit reference changes, and a design that separates task change from cap re-donning in MCD.

B Extended Related Work

B.1 Spatial Interpolation and Graph-Structured Reconstruction

SS, IDW, and KRIGING estimate missing potentials with spherical splines, distance weighting, and covariance structure, respectively (Shepard, 1968; Perrin et al., 1987; 1989; Matheron, 1963). The spherical-spline work of Perrin et al. established a classic geometric basis for scalp-potential interpolation (Perrin et al., 1987; 1989). These methods use electrode locations directly and require no large training set to learn a full mapping, but when the remaining electrodes cover the scalp

unevenly, their spatial assumptions can dominate the result. STC reconstructs missing EEG from inter-channel spatiotemporal correlations, offering another route that exploits statistical dependence across channels (Ramakrishnan et al., 2016; Bahador et al., 2021).

Graph-based methods write electrode relationships directly into the reconstruction process (Hammond et al., 2011; Shuman et al., 2013; Ortega et al., 2018; Defferrard et al., 2016). GH performs harmonic extension on a fixed graph (Zhu et al., 2003), while the original SRL-ITV jointly models local and global structure in incomplete time-varying EEG (Cao et al., 2025a;b). RIGS-EEG also keeps spatial structure explicit, but assigns different jobs to two structures: a dense kernel provides the initial state, a sparse graph drives later updates, and spectral gains control corrections across time. This avoids forcing initialization and frequency-domain refinement through the same adjacency pattern.

B.2 Low-Rank Completion and Autoencoder Reconstruction

HALRTC expresses multimodal low-rank priors through tensor nuclear norms (Liu et al., 2012), whereas STDC couples tensor decomposition with completion (Chen et al., 2013). Low-rank approaches can exploit redundancy across channels and time (Duan et al., 2021; Tang et al., 2025; Liu et al., 2012; Chen et al., 2013), but whole-channel loss places heavier demands on the cross-channel structure available from the remaining electrodes. GLAE uses gates to regulate how visible inputs contribute to missing-value recovery, while DMAE-EEG learns spatiotemporal EEG representations through regional masking and autoencoder pretraining (Saba-Sadiya et al., 2020; El-Fiqi et al., 2019; Zhang et al., 2025). RIGS-EEG takes a different route: learning is confined to explicit graph-time-frequency updates, while reference handling, geometry, and data consistency remain part of the reconstruction rule itself.

B.3 Attention Modeling and Generative Imputation

CT uses a cascade Transformer for EEG completion (Zhang et al., 2022; Li et al., 2025); SAITS learns multivariate missing values with diagonally masked self-attention (Du et al., 2023); and PATCHTST uses temporal patches and channel-independent modeling for long-sequence prediction (Nie et al., 2022; Du et al., 2024; Du, 2023). Together, these methods strengthen temporal modeling through progressive refinement, cross-variable dependency learning, and local patch representations (Che et al., 2018; Cao et al., 2018; Yoon et al., 2018; Tashiro et al., 2021; Du et al., 2023; Nie et al., 2022; Du et al., 2024; Du, 2023).

CSDI models the distribution of missing values through an observation-conditioned score diffusion process (Tashiro et al., 2021), providing a different path from deterministic reconstruction. RIGS-EEG instead uses a finite, deterministic sequence of graph-time-frequency updates in which electrode geometry, the common reference, and valid observations enter the computation explicitly. The point is not that one model family is universally better; the question is whether these EEG-specific structures translate into measurable gains when an entire channel is missing.

B.4 Reference Handling and Algorithm Unrolling

REST uses a head model to approximately transform EEG recordings toward an infinity reference (Yao, 2001; Yao et al., 2019; Hu et al., 2019). RIGS-EEG does not estimate that physical reference. It addresses a different, directly constrainable question: when an additive common reference changes, the reconstruction should shift with it while preserving relative channel structure. Algorithm unrolling writes iterative inference as a fixed-depth learnable process, a paradigm demonstrated by Gregor and LeCun for sparse coding (Gregor et al., 2010; Monga et al., 2021). We bring that idea to whole-channel EEG reconstruction, linking reference separation, reliability-aware spatial initialization, graph-time-frequency shrinkage, and observation consistency into one inference chain.

C MCD Reconstruction and Ablation

Table 5: MCD reconstruction and ablation (NRMSE, mean ± subject SD).

Method	25%	50%	75%
SS	0.1000 ± 0.0238	0.1114 ± 0.0276	0.1419 ± 0.0351
IDW	0.0872 ± 0.0177	0.0890 ± 0.0183	0.0944 ± 0.0197
KRIGING	0.1468 ± 0.0385	0.1658 ± 0.0444	0.1500 ± 0.0374
STC	0.0956 ± 0.0159	0.0964 ± 0.0163	0.1083 ± 0.0170
STDC	0.0907 ± 0.0175	0.0918 ± 0.0179	0.0942 ± 0.0187
HALRTC	0.0897 ± 0.0173	0.0907 ± 0.0176	0.0938 ± 0.0180
GLAE	0.0953 ± 0.0159	0.0942 ± 0.0105	0.1199 ± 0.0042
CT	0.0896 ± 0.0175	0.0904 ± 0.0179	0.0938 ± 0.0188
SAITS	0.0928 ± 0.0173	0.1017 ± 0.0186	0.1433 ± 0.0260
CSDI	0.1010 ± 0.0198	0.1301 ± 0.0219	0.2052 ± 0.0315
PATCHTST	0.0896 ± 0.0175	0.0905 ± 0.0180	0.0939 ± 0.0189
GH	0.0890 ± 0.0195	0.0919 ± 0.0202	0.0965 ± 0.0211
SRL-ITV	0.0878 ± 0.0144	0.0993 ± 0.0096	0.1388 ± 0.0028
DMAE-EEG	0.0870 ± 0.0170	0.0890 ± 0.0162	0.1030 ± 0.0135
RIGS-EEG	0.0861 ± 0.0177	0.0883 ± 0.0184	0.0922 ± 0.0192
w/o reference	0.0878 ± 0.0172	0.0899 ± 0.0177	0.0936 ± 0.0185
w/o graph	0.0904 ± 0.0163	0.0910 ± 0.0166	0.0929 ± 0.0176
w/o spectral	0.0896 ± 0.0176	0.0905 ± 0.0181	0.0928 ± 0.0190

D Derivations and Reproducibility Details

The appendix gives derivations and protocol details that are abbreviated in the main text. Notation follows the methodology section: mask value 1 denotes a visible position, electrode coordinates determine the spatial kernel and graph operators, and reliability weights are computed only from observed values.

Common reference and relative potentials. At each time point, the common component is the reliability-weighted average over visible electrodes. Missing positions have zero weight, so their placeholder values never enter the denominator. Suppose all visible electrodes are shifted by an arbitrary reference waveform while the mask, coordinates, and reliability weights remain fixed. The weighted numerator then gains "reference waveform x sum of weights," while the denominator is unchanged; their ratio therefore gains exactly that same reference waveform. Subtracting the estimated reference leaves the relative potentials unchanged point by point.

$$d_t = \sum_c W_{ct}, \qquad r_t = \frac{\sum_c W_{ct} Y_{ct}}{d_t}, \qquad R = M \odot \left(Y - \mathbf{1} r^\top\right).$$

This derivation also shows why the reliability weights must be invariant to a common shift. In v2, the median across observed channels is computed at each time point, and channel RMS is then measured relative to that median. Adding the same reference waveform moves both the observations and the median together, so their differences, and thus the weights, remain unchanged. Externally supplied weights must satisfy the same condition; otherwise reference equivariance does not follow directly.

Spatial initialization. The model builds a strictly positive dense Gaussian kernel from electrode coordinates and applies kernel averaging to reliability-weighted relative potentials. The numerator accumulates only observed electrodes, and the denominator accumulates the corresponding valid kernel weights. As long as at least one reliable observation exists at each time point, the denominator remains positive. The resulting vector is then centered across channels before entering graph-time-frequency unrolling. This explains why an initial estimate still exists when all four nearest neighbors of a target are missing, and why the dense kernel is not redundant with the later sparse graph.

$$S^0 = \Pi\left([H(W \odot R)] \oslash [HW]\right), \qquad \Pi(Z) = Z - \mathbf{1}\frac{\mathbf{1}^\top Z}{C}.$$

Stepwise update and equivariance. Assume that, at one step, the relative-potential estimate is the same before and after a common reference shift. The current offset is the reliability-weighted average of the observation minus that estimate. When the observations receive a common reference waveform, the offset gains the same waveform. The data-consistency residual, "estimate + offset - observation," is therefore unchanged, and so is the next relative-potential update.

$$b_t^k = \frac{\sum_c W_{ct}\left(Y_{ct} - S_{ct}^k\right)}{d_t},$$

Graph projection and the short-time Fourier transform operate only on relative potentials. Their graph-frequency, temporal-frequency, and observation-coverage features do not change under a common reference shift, so the spectral gains and inverse-transform result also remain unchanged. Repeating the argument through all five steps leaves the final relative estimate unchanged, while the common offset gains the input reference waveform and the output shifts by the same amount. This is the stepwise reason behind the equivariance equation in the methodology. The proof concerns additive common changes in the normalized model-input coordinate system; it does not claim equivariance to arbitrary linear re-referencing.

$$F_\theta\left(M \odot \left(Y + \mathbf{1}a^\top\right), M, P\right) = F_\theta(Y, M, P) + \mathbf{1}a^\top.$$

Scope of observation consistency. In all reported experiments, reliability weights are used only to modulate the internal reference estimation and spatial propagation of observed signals. The final output always applies a hard observation projection, so every measured position is preserved exactly and only missing entries are replaced by model estimates. Any optional interface allowing externally supplied reliability weights does not change this evaluation protocol and is not used to overwrite observed samples in the experiments reported here.

Training objective versus scoring. PhysioNet v1 uses mean squared error on missing positions as the main loss, with auxiliary spectral-magnitude and graph-edge-difference terms. SEED-IV and MCD v2 use channel-normalized RMSE for each affected channel. The normalization denominator is the dynamic range of the complete target channel and is used only for supervised training and offline scoring. At inference, the model never reads the true amplitude or dynamic range of a masked channel. Subject-level NRMSE standard deviations are computed after window aggregation and then subject aggregation; they should not be interpreted as variation across repeated model training runs.

Parameter count and computation. In each step, the gain MLP maps four input features to width 32 and then to one gain. The two linear layers contain 160 and 33 parameters, respectively, and one learnable step size brings the per-step total to 194. Parameters are not shared across the five steps, giving 970 trainable parameters in total. Graph eigendecomposition, dense-kernel computation, and the STFT at each step still consume time and memory; 970 describes trainable model size, not inference cost.

Protocol reproducibility details. PhysioNet is split by subject so that different runs from one person never cross splits. SEED-IV keeps the same subjects and uses sessions 1, 2, and 3 for training, validation, and testing. MCD keeps each participant's reading, memory, and IQ trials in the same split. Paired masks are used across methods within each evaluation condition, robust scaling is fitted only on training data, and model selection uses only validation data. Classification evaluation reuses a frozen classifier trained on complete signals and compares complete, corrupted, and reconstructed test inputs in paired fashion. These details define what the reported scores compare and prevent window- or subject-level information from crossing split boundaries.