COUNTERACTING DETRIMENTAL PROPAGATION ACROSS SPACE AND TIME IN EEG

Anonymous authors
Paper under double-blind review

ABSTRACT

Electroencephalography (EEG) enables non-invasive, real-time recording of neural activity, with broad applications in neurological diagnosis, monitoring and rehabilitation. Yet, reliable EEG analysis is hindered by task-irrelevant signals that propagate across electrodes, accumulate over time, and become entangled with predictive cues. Existing attempts focus on the outcome level by regularizing resulting representations or structures, leaving the propagation process largely unaddressed. By the time they intervene, detrimental signals may already have spread and become entangled with predictive cues, making post-hoc removal intrinsically difficult and risking collateral damage to task-relevant neural signatures. In this paper, we propose CounterField, which recasts the mitigation of task-irrelevant signals as counteraction against their propagation from a graph stochastic partial differential equation (SPDE) perspective. CounterField structures a spectral response against detrimental propagation, enforcing coherent counteraction while preserving predictive flow. We extensively evaluate CounterField across diverse EEG tasks, where consistent gains demonstrate the effectiveness of propagation-level counteraction in learning compact yet predictive spatiotemporal dynamics.

1 INTRODUCTION

Electroencephalography (EEG) provides a non-invasive, real-time measure of brain activity through electrical signals recorded from scalp electrodes, enabling a broad spectrum of applications ranging from neurological diagnosis Kotoge et al. (2026) and emotion recognition Liu et al. (2024) to motor imagery classification Wei et al. (2025a), sleep staging Wang et al. (2024b), and braincomputer interfaces Pan et al. (2022).

To translate raw EEG recordings into meaningful signatures, conventional approaches organize multichannel signals or spectrograms as regular tensors or sequences, modeling temporal and spectral patterns with CNNs Raghu et al. (2020), RNNs Wang et al. (2023), and Transformers Phan et al. (2022), while recent foundation models learn transferable representations through large-scale pretraining Wei et al. (2025b); Jiang et al. (2024; 2025). Yet, these paradigms provide limited characterization of the relational organization of multichannel EEG. Motivated by evidence that neural activity involves coordinated interactions across functionally related brain regions Hawkins & Ahmad (2016); Rubinov & Sporns (2010), graph-based approaches explicitly model inter-electrode relations through predefined topologies Yi et al. (2023), hierarchical organization Covert et al. (2019), or learned structures Tang et al. (2021), with encouraging performance suggesting that such relational organization constitutes an informative dimension of EEG dynamics.

Despite these advances, a fundamental challenge remains unresolved: EEG captures a noisy and incomplete mixture of multiple latent neural sources, allowing detrimental components to spread through inter-electrode interactions and accumulate over time, thereby becoming intertwined with predictive dynamics. Reliable analysis therefore requires more than expressive capacity; it necessitates suppressing such contamination while preserving task-relevant signatures. This conceptually connects to a broad family of compact and denoising graph methods, which suppress redundant information through hand-crafted structural criteria Fang et al. (2019b); Ketkar et al. (2005), learned sparsification Zhang et al. (2025), pooling Wang et al. (2024c), or condensation Luo et al. (2025), and information-theoretic regularization such as GIB Wu et al. (2020); Yu et al. (2021); Hoang & Lee (2025); Wang & Fan (2024); Ding et al. (2024).

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.

From an information-flow perspective, the above formulations primarily impose regulation on static outcomes rather than on the propagation process itself. Yet, EEG dynamics are inherently structured and evolving, where both task-relevant and task-irrelevant signals diffuse, mix, and accumulate through rounds of interactions. By the time these strategies intervene, the contamination may have already dispersed across the graph in a spatially heterogeneous manner and become deeply entangled with useful cues. Such entanglement not only makes post-hoc removal fundamentally difficult, but also risks causing collateral damage to task-relevant dynamics.

In this paper, we argue that the key to mitigating such detrimental effects in EEG dynamics lies in the propagation process. This motivates us to revisit EEG evolution from the perspective of stochastic dynamical systems. Specifically, we propose CounterField, which characterizes the spatial-temporal transition as predictive flow entangled with task-irrelevant propagation, and extend the stochastic term into a Matérn spectral operator that structures a counteractive field against detrimental dynamics between EEG states. This formulation enables compactness to act directly within the propagation of spatio-temporal dynamics, rather than post-hoc on the representational or structural readouts, counteracting task-irrelevant signals before they spread across electrodes and accumulate over time.

CounterField comprises three complementary components: spectral structure learning, counteractive refinement and sufficiency alignment. Spectral structure learning defines the spectral geometry for the Matérn kernel that structures CounterField, providing a reference grounded in the frequency-dependent organization of EEG. Counteractive refinement shapes a spectral response against task-irrelevant propagation and enforces coherent counteraction via a parsimony objective. Sufficiency alignment safeguards predictive flow by constraining counteraction to preserve downstream predictive contribution, preventing aggressive refinement from compromising task-relevant dynamics.

Our main contributions are summarized as follows: (i) We propose CounterField, which recasts the mitigation of detrimental components in EEG dynamics as learning to counteract task-irrelevant propagation through a graph-SPDE perspective; (ii) We establish a principled formulation for CounterField by structuring counteraction with EEG spectral characteristics, together with parsimony and sufficiency objectives that drive it against task-irrelevant nuisances without compromising downstream predictive contribution; (iii) We conduct extensive experiments across diverse EEG benchmarks, and consistent improvements over state-of-the-art methods validate both the motivation and formulation of CounterField.

2 METHOD

Our goal is to capture predictive EEG dynamics while counteracting the propagation of task-irrelevant signals. To this end, we first infer inter-electrode structures from the complex EEG spectrum, providing a geometric prior for subsequent counteraction. On this basis, we formulate state transitions from a graph SPDE perspective and introduce counteraction along their propagation.

2.1 BACKGROUND

EEG Analysis. Existing EEG methods primarily model temporal and spectral patterns by translating EEG signals into generic representations or tokens for subsequent neural processing, while several graph-based approaches introduce inter-electrode interactions through predefined topologies, hierarchical organization, or learned relations. Yet, reliable EEG analysis remains challenged by task-irrelevant signals that propagate across electrodes, accumulate over time, and become entangled with predictive contents. Existing paradigms focus on developing more expressive or transferable networks but leave this fundamental challenge unaddressed. In this paper, we revisit the propagation process and formulate CounterField to inhibit detrimental signals along EEG dynamics.

Compact Graph Representation Learning. Existing efforts are mainly devoted to structural or representational pruning through sparsification, condensation, pooling, or information-theoretic compression. However, they remain confined to a long-standing outcome-level perspective, refining resulting structures or representations rather than the information propagation process that produces them and thereby allowing detrimental signals to spread and accumulate across nodes and over time before refinement takes effect. In contrast, CounterField reframes compactness as propagation-level counteraction, directly intervening along graph dynamics.

Graph Diffusion and Stochastic Propagation. Graph diffusion provides a continuous-time view of information propagation over graph structures, which can be extended to stochastic dynamics as:

$$d\mathbf{H}(t) = \underbrace{f(\mathbf{H}(t))dt}_{\textit{Drift term}} + \underbrace{g(\mathbf{H}(t))d\mathbf{W}(t)}_{\textit{Stochastic term}}. \tag{1}$$

where the drift term $f(\mathbf{H}(t))dt$ characterizes deterministic propagation, while $g(\mathbf{H}(t))d\mathbf{W}(t)$ introduces structured stochastic perturbations. Beyond its conventional role in uncertainty modeling, this stochastic formulation provides a principled means of structuring perturbations over graph geometry. We build on this perspective by turning structured perturbation into counteraction against detrimental propagation.

2.2 Spectral Structure Learning

Given an EEG clip recorded from N electrodes, we divide it into T temporal snapshots. Applying the short-time Fourier transform (STFT) to each snapshot yields M retained frequency components, where the complex response of electrode i at frequency m is:

$$Z_{i,m}^t = \mathcal{A}_{i,m}^t \exp(\mathrm{i}\phi_{i,m}^t). \tag{2}$$

Here, $\mathcal{A}_{i,m}^t$ and $\phi_{i,m}^t$ denote the amplitude and phase of the m-th frequency component, respectively, and i is the imaginary unit. On this basis, we represent each electrode i at snapshot t as:

$$x_i^t = [\log(\mathcal{A}_i^t + \epsilon) \,\|\, \cos\phi_i^t \,\|\, \sin\phi_i^t], \quad H_t = [x_1^t, \ldots, x_N^t]^\top, \tag{3}$$

where ϵ ensures numerical stability, x_i^t encodes the amplitude and phase information across retained frequencies, and H_t stacks the spectral representations of all electrodes at snapshot t.

The geometry inferred from each snapshot shapes the spectral structure of subsequent counteraction, thereby providing an inductive bias for suppressing task-irrelevant variations while preserving predictive dynamics. Therefore, we derive inter-electrode structure from the inherent frequency organization of EEG, complementing generic feature proximity with multiple oscillatory cues.

To this end, we first characterize frequency-wise co-activation for each electrode pair (i, j) via:

$$a_{ij}^t(m) = \exp\left[-\eta_m \left(\log \mathcal{A}_{i,m}^t - \log \mathcal{A}_{j,m}^t\right)^2\right], \tag{4}$$

in which $\eta_m > 0$ controls the frequency-specific sensitivity to amplitude discrepancy. Beyond co-activation, we introduce phase compatibility to capture coordinated timing between the corresponding oscillations:

$$p_{ij}^t(m) = \exp\left\{\kappa_m \left[\cos(\phi_{i,m}^t - \phi_{j,m}^t - \mu_m) - 1\right]\right\}. \tag{5}$$

Here, μ_m denotes the learnable preferred phase offset and $\kappa_m > 0$ controls phase selectivity. We then require credible structural evidence to be jointly supported by these complementary oscillatory cues, and integrate them through harmonic fusion:

$$e_{ij}^t(m) = \frac{2a_{ij}^t(m)p_{ij}^t(m)}{a_{ij}^t(m) + p_{ij}^t(m) + \epsilon}. \tag{6}$$

This yields frequency-specific structural evidence $e_{ij}^t(m)$, where strong support emerges only when spectral co-activation is corroborated by compatible phase behavior. On this basis, we learn latent nonlinear coupling patterns as:

$$r_{ij}^t = \Phi_\theta\left(x_i^t, x_j^t\right) + \beta\, e_{ij}^t, \quad e_{ij}^t = [e_{ij}^t(1), \ldots, e_{ij}^t(M)], \tag{7}$$

where Φ_θ denotes a parameterized kernel and β controls the contribution of spectral evidence. This formulation complements flexible nonlinear coupling with explicit oscillatory evidence, allowing structure inference to remain expressive while grounded in EEG spectral characteristics.

Further, a crucial characteristic of EEG is the temporal continuity of predictive dynamics. That is, ambiguous cues in an isolated snapshot can become informative when interpreted in light of preceding states, which in turn provide context for those emerging subsequently. Therefore, we instantiate this idea through a lightweight update:

$$q^t = r^t + \gamma c^{t-1}, \quad c^t = (1 - z^t)c^{t-1} + z^t[r^t - \tau]_+, \tag{8}$$

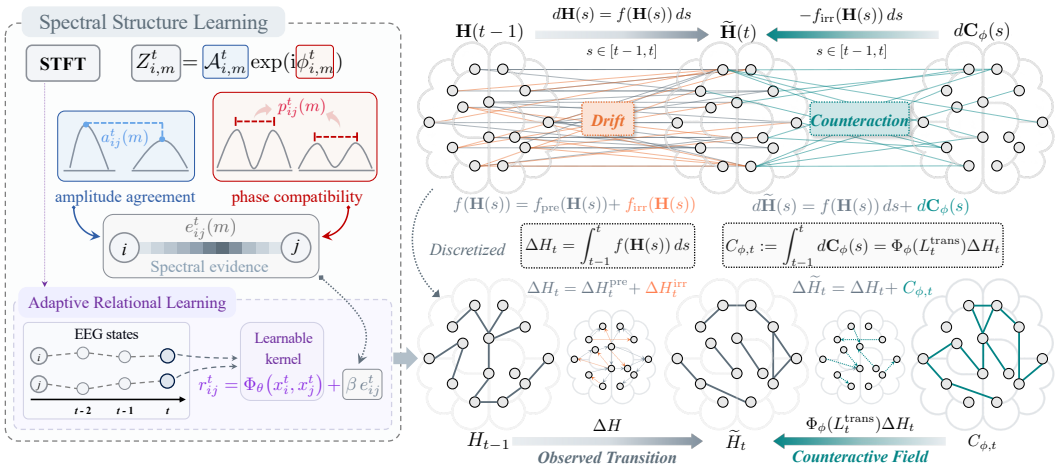


Figure 1: Overview of the proposed formulation.

in which c^{t-1} denotes the relational context inherited from preceding states, γ controls its contribution, and z^t is a learnable update coefficient. The threshold τ prevents weak relations from being persistently accumulated in the temporal context.

Building on the above formulation, we construct each EEG snapshot adjacency through:

$$A_{ij}^t = \sum_{b=1}^{B} \omega_b \left(\frac{1}{|\mathcal{B}_b|} \sum_{\mathcal{B}_b} q_{ij}^t \right). \tag{9}$$

This hierarchical aggregation preserves the spectral organization of EEG in a lightweight and flexible manner, where $\mathcal{B}_b$ denotes the frequency components within the b-th canonical EEG band and ω_b controls its contribution.

2.3 Counteractive Compact Learning

Our goal is to learn compact yet predictive spatiotemporal dynamics from EEG graphs by mitigating the propagation of task-irrelevant nuisances while preserving predictive information flow. To this end, we begin with the observed transition:

$$\Delta H_t := H_t - H_{t-1} = \Delta H_t^{\text{pre}} + \Delta H_t^{\text{irr}}, \tag{10}$$

where ΔH_t^{pre} denotes predictive dynamics, while ΔH_t^{irr} collects task-irrelevant variations. However, both of them are latent and entangled in the observed transition, making direct estimation of ΔH_t^{irr} inherently ill-posed. To address this, we interpret Eq. (10) through a graph-SPDE perspective and characterize the transition as a graph-structured propagation field[1]:

$$d\mathbf{H}(t) = f(\mathbf{H}(t))\, dt \tag{11}$$

Following Eq. (27), the drift field $f(\mathbf{H}(t))\, dt$ captures the propagation of EEG states, where the predictive component $f_{\text{pre}}(\mathbf{H}(t))$ and task-irrelevant nuisances $f_{\text{irr}}(\mathbf{H}(t))$ are entangled:

$$f(\mathbf{H}(t)) = f_{\text{pre}}(\mathbf{H}(t)) + f_{\text{irr}}(\mathbf{H}(t)). \tag{12}$$

Over the transition from $t-1$ to t, the latter contributes:

$$\Delta H_t^{\text{irr}} = \int_{t-1}^{t} f_{\text{irr}}(\mathbf{H}(s))\, ds. \tag{13}$$

This suggests the removal of ΔH_t^{irr} amounts to counteracting the accumulated effect of f_{irr} along the transition, guiding the intervention to act on the propagation process itself rather than cleaning up the resulting state in a post-hoc manner. We therefore introduce a counteractive field $\mathbf{C}_\phi$:

$$d\widetilde{\mathbf{H}}(t) = f(\mathbf{H}(t))\, dt + d\mathbf{C}_\phi(t), \quad d\mathbf{C}_\phi(t) \approx -f_{\text{irr}}(\mathbf{H}(t))\, dt \tag{14}$$

[1]Bold symbols denote continuous-time processes, while their regular counterparts denote discrete snapshot-level quantities.

such that, ideally, $d\widetilde{\mathbf{H}}(t) \approx f_{\text{pre}}(\mathbf{H}(t))dt$. This requires a structured parameterization of $d\mathbf{C}_\phi(t)$ over the graph geometry of EEG propagation. To this end, we first present the following definition:

Definition 1 (Graph Gaussian Random Field). *Given an undirected graph $G = (V, E)$ with graph Laplacian $L = U\Lambda U^\top$, a graph Gaussian random field is defined as $\mathbf{f} \sim \mathcal{N}(0, \mathbf{C})$, where $\mathbf{C} = \Phi(L) = U\Phi(\Lambda)U^\top$, and Φ is a spectral function of the graph Laplacian that specifies the covariance of the field along its graph spectral modes.*

By applying a spectral function to the graph Laplacian, the above definition provides a general principle for constructing structured fields. However, Eq. (14) necessitates intervention along the propagation between consecutive states, rather than on either state in isolation. We therefore structure $\mathbf{C}_\phi$ over the structural transition between consecutive EEG states:

$$A_t^{\text{trans}} = |A_t - A_{t-1}|, \qquad L_t^{\text{trans}} = D_t^{\text{trans}} - A_t^{\text{trans}}. \tag{15}$$

Here, the absolute variation preserves the magnitude of structural changes while yielding a valid Laplacian L_t^{trans} for the transition. On this basis, we define the counteractive field as:

$$C_{\phi,t} := \int_{t-1}^{t} d\mathbf{C}_\phi(s) = \Phi_\phi(L_t^{\text{trans}})\Delta H_t, \tag{16}$$

where $L_t^{\text{trans}} = U_t\Lambda_t U_t^\top$, and the corresponding spectral operator is parameterized as:

$$\Phi_\phi(L_t^{\text{trans}}) = U_t h_\phi(\Lambda_t)U_t^\top, \quad h_\phi(\lambda) = m_\phi(\lambda)\left(\frac{2\nu}{\kappa^2} + \lambda\right)^{-\nu}. \tag{17}$$

Here, the Matérn term encodes the spectral regularity induced by the transition geometry, with ν and κ controlling its smoothness and scale, respectively. The signed modulator $m_\phi(\lambda)$ adapts this structural prior into a task-driven counteractive response over ΔH_t. On this basis, the resulting field is applied to the observed transition, giving:

$$\Delta\widetilde{H}_t = \Delta H_t + C_{\phi,t}, \quad \widetilde{H}_t = H_{t-1} + \Delta\widetilde{H}_t, \tag{18}$$

where $\Delta\widetilde{H}_t$ denotes the refined transition from $t-1$ to t, and $\widetilde{H}_t$ represents the resulting state with task-irrelevant propagation counteracted along $\Delta\widetilde{H}_t$. Interestingly, Eq. (18) also admits the following algebraically equivalent views:

$$\underbrace{H_{t-1} + \Delta\widetilde{H}_t}_{\text{transition view}} \equiv \underbrace{H_t + C_{\phi,t}}_{\text{residual view}}. \tag{19}$$

In particular, the transition view describes evolution from the preceding state with task-irrelevant propagation counteracted along the transition, while the residual view equivalently expresses this intervention as a counteractive residual on the current raw state. This equivalence highlights the key distinction of our formulation: counteraction targets the propagation underlying the observed transition rather than its resulting state alone.

Next, we impose a parsimony objective on $\widetilde{H}_t$ to drive $C_{\phi,t}$ against $f_{\text{irr}}(\mathbf{H}(t))$. Formally, we have:

$$\mathcal{L}_{\text{par}} = \frac{1}{T-1}\sum_{t=2}^{T}\left[\mathcal{V}(\widetilde{H}_t) - \mathcal{V}(H_t) + \delta\right]_+, \quad \mathcal{V}(H) = \log\det(I + \alpha HH^\top), \tag{20}$$

where $\mathcal{V}(H)$ measures the dispersion of H, with $\alpha > 0$ controlling its scale and $\delta > 0$ preventing identity solution. This objective serves as a tractable proxy to encourage $C_{\phi,t}$ to contract redundant variations without prescribing which spectral components should be inhibited[2]. However, parsimony is not our only objective; here, we need to maintain the predictive flow intact. This calls for an additional constraint from the sufficiency side to preserve predictive information during counteraction.

Discussion. Conceptually, Eq. (20) aligns with the minimal-sufficiency principle of compact graph learning, while shifting regularization from resulting readouts to the propagation process. This

[2]Under a Gaussian approximation, the log-determinant of the covariance is proportional to differential entropy up to an additive constant Rasmussen & Williams (2006), providing an information-theoretic interpretation of representation volume.

distinction is closely tied to the relational nature of graph data, and becomes particularly important for spatio-temporal signals such as EEG, where redundant and misleading information becomes spatially entangled through propagation and accumulates over time. Relative to Lin et al. (2024); Xu & Markovich (2025); Borovitskiy et al. (2021), our formulation builds on the graph diffusion view and extends its role from capturing structured propagation to actively counteracting detrimental propagation, thereby providing a process-level perspective on compact graph learning.

2.4 Spatiotemporal Predictive Sufficiency

To prevent the predictive flows from being compromised, sufficiency should be reconsidered across both spatial and temporal dimensions. This departs from the static setting, where a compact readout Z can be directly evaluated against the task target Y, the predictive contribution of each H_t in spatio-temporal graph learning is realized jointly through downstream temporal processing. Imposing state-wise sufficiency on every $\widetilde{H}_t$ would instead require each intermediate state to be independently evaluated against Y, mischaracterizing its role in the overall dynamics.

To address this, we formulate sufficiency with respect to the downstream task: counteraction at an intermediate state is considered aggressive if its estimated effect induces an excessive increase in the final task objective. Formally, we define:

$$\mathcal{L}_{\mathrm{suf}} = \left[\langle \mathrm{sg}(g_t), C_{\phi,t}\rangle - \epsilon_s\right]_+, \quad g_t := \nabla_{H_t}\mathcal{L}_{\mathrm{task}}, \tag{21}$$

where g_t denotes the gradient of the task objective $\mathcal{L}_{\mathrm{task}}$ back-propagated to H_t, $\mathrm{sg}(\cdot)$ is the stop-gradient operator, and ϵ_s is an optional tolerance for local predictive degradation.

In particular, Eq. (21) can be intuitively interpreted through the first-order variation of $\mathcal{L}_{\mathrm{task}}$, *i.e.*, $\mathcal{L}_{\mathrm{task}}(H_t + C_{\phi,t}) - \mathcal{L}_{\mathrm{task}}(H_t) \approx \langle g_t, C_{\phi,t}\rangle$, where $\langle g_t, C_{\phi,t}\rangle$ quantifies the first-order change in the final task objective induced by counteraction at H_t. From this perspective, Eq. (21) functions only when the degradation exceeds ϵ_s, while allowing neutral or beneficial counteraction to proceed.

Together, the parsimony and sufficiency objectives banish the spatial spread and temporal accumulation of task-irrelevant signals while preventing aggressive counteraction of task-relevant information, yielding a tractable objective for compact yet predictive spatiotemporal representations.

2.5 Practical Implementation

To avoid the eigendecomposition required by Eq. (17), we follow Borovitskiy et al. (2021); Defferrard et al. (2016) and approximate the resulting spectral operator as:

$$C_{\phi,t} \approx \sum_{k=0}^{K} c_k(\phi)\, T_k(\widetilde{L}_t^{\mathrm{trans}})\Delta H_t, \quad \widetilde{L}_t^{\mathrm{trans}} = \frac{2L_t^{\mathrm{trans}}}{\lambda_{\max}} - I. \tag{22}$$

Here, $T_k(\cdot)$ denotes the k-th degree Chebyshev polynomial, and $c_k(\phi)$ is the coefficient induced by the spectral response h_ϕ. This reduces the computation from explicit eigendecomposition to $O(K|E|)$ sparse graph operations.

Finally, the counteractive formulation is jointly optimized with the downstream task through the overall objective:

$$\mathcal{L} = \mathcal{L}_{\mathrm{task}} + \lambda_{\mathrm{par}}\mathcal{L}_{\mathrm{par}} + \lambda_{\mathrm{suf}}\mathcal{L}_{\mathrm{suf}}. \tag{23}$$

The overall framework is illustrated in Fig. 1. SpecSL first translates the EEG spectra into dynamic inter-electrode structures, providing the geometric prior for the subsequent counteraction. On this basis, CCL (right panel) characterizes the formation of the refined state from two complementary views: (i) continuously, the counteractive field $\mathbf{C}_\phi$ neutralize task-irrelevant part in drift along the propagation process; (ii) discretely, the transition geometry structures $C_{\phi,t}$ through a Matérn kernel, enabling targeted counteraction against $\Delta H_t^{\mathrm{irr}}$ to refine the observed transition.

3 Experiments

3.1 Experimental Setup

Benchmarks and Evaluation Protocols. CounterField is evaluated on four EEG tasks: seizure diagnosis on **TUSZ** Shah et al. (2018) (134,625 clips, 19 channels, resampled to 200 Hz); emotion

Table 1: Performance comparison on TUSZ dataset on seizure detection and prediction. Best and second-best results are shown in bold and underlined, respectively, with the baseline shaded.

	Detection				Prediction			
	12 s		60 s		12 s		60 s	
Method	**AUROC**	**F1**	**AUROC**	**F1**	**AUROC**	**F1**	**AUROC**	**F1**
BIOT Yang et al. (2023a)	$0.729_{\pm 0.013}$	$0.320_{\pm 0.017}$	$0.655_{\pm 0.028}$	$0.286_{\pm 0.011}$	$0.577_{\pm 0.016}$	$0.424_{\pm 0.018}$	$0.571_{\pm 0.005}$	$0.391_{\pm 0.010}$
LaBraM Jiang et al. (2024)	$0.824_{\pm 0.006}$	$0.477_{\pm 0.010}$	$0.798_{\pm 0.006}$	$0.469_{\pm 0.009}$	$0.668_{\pm 0.003}$	$\underline{0.480}_{\pm 0.008}$	$\mathbf{0.674}_{\pm \mathbf{0.017}}$	$\underline{0.418}_{\pm 0.024}$
EEGPT Wang et al. (2024a)	$0.805_{\pm 0.012}$	$0.416_{\pm 0.014}$	$0.748_{\pm 0.006}$	$0.412_{\pm 0.015}$	$\underline{0.676}_{\pm 0.002}$	$0.464_{\pm 0.012}$	$0.608_{\pm 0.017}$	$0.394_{\pm 0.022}$
EvolveGCN Pareja et al. (2020)	$0.760_{\pm 0.005}$	$0.344_{\pm 0.011}$	$0.670_{\pm 0.023}$	$0.343_{\pm 0.018}$	$0.621_{\pm 0.010}$	$0.436_{\pm 0.011}$	$0.538_{\pm 0.023}$	$0.347_{\pm 0.023}$
DCRNN Tang et al. (2022)	$0.822_{\pm 0.013}$	$0.414_{\pm 0.036}$	$0.808_{\pm 0.014}$	$0.434_{\pm 0.025}$	$0.640_{\pm 0.021}$	$0.405_{\pm 0.025}$	$0.607_{\pm 0.033}$	$0.397_{\pm 0.028}$
GRA4MER Tang et al. (2023)	$0.836_{\pm 0.005}$	$0.416_{\pm 0.023}$	$0.779_{\pm 0.020}$	$0.446_{\pm 0.014}$	$0.630_{\pm 0.018}$	$0.436_{\pm 0.025}$	$0.643_{\pm 0.026}$	$0.353_{\pm 0.032}$
GRU-GCN Gao & Ribeiro (2022)	$0.863_{\pm 0.011}$	$0.512_{\pm 0.015}$	$0.819_{\pm 0.017}$	$0.442_{\pm 0.016}$	$0.659_{\pm 0.023}$	$0.451_{\pm 0.013}$	$0.603_{\pm 0.035}$	$0.398_{\pm 0.017}$
EvoBrain Kotoge et al. (2026)	$\underline{0.877}_{\pm 0.005}$	$\underline{0.539}_{\pm 0.009}$	$\underline{0.865}_{\pm 0.009}$	$\underline{0.483}_{\pm 0.006}$	$0.675_{\pm 0.015}$	$0.470_{\pm 0.003}$	$0.651_{\pm 0.023}$	$0.401_{\pm 0.023}$
Ours	$\mathbf{0.897}_{\pm 0.007}$	$\mathbf{0.556}_{\pm 0.004}$	$\mathbf{0.881}_{\pm 0.011}$	$\mathbf{0.495}_{\pm 0.005}$	$\mathbf{0.685}_{\pm 0.016}$	$\mathbf{0.482}_{\pm 0.004}$	$\underline{0.673}_{\pm 0.021}$	$\mathbf{0.419}_{\pm 0.018}$

Table 2: Performance comparison on FACED and ISRU. Best and second-best results are shown in bold and underlined, respectively, with the baseline shaded.

	FACED, 9-class			ISRUC, 5-class		
Method	**Balanced Acc.**	**Cohen's Kappa**	**Weighted F1**	**Balanced Acc.**	**Cohen's Kappa**	**Weighted F1**
SPaRCNet Jing et al. (2023)	$0.4673_{\pm 0.0155}$	$0.3978_{\pm 0.0289}$	$0.4729_{\pm 0.0133}$	$0.7487_{\pm 0.0075}$	$0.7097_{\pm 0.0132}$	$0.7624_{\pm 0.0092}$
ContraWR Yang et al. (2023b)	$0.4887_{\pm 0.0078}$	$0.4231_{\pm 0.0151}$	$0.4884_{\pm 0.0074}$	$0.7402_{\pm 0.0126}$	$0.7178_{\pm 0.0156}$	$0.7610_{\pm 0.0137}$
CNN-Trans Peh et al. (2022)	$0.4697_{\pm 0.0132}$	$0.4017_{\pm 0.0168}$	$0.4720_{\pm 0.0125}$	$0.7363_{\pm 0.0087}$	$0.7129_{\pm 0.0121}$	$0.7719_{\pm 0.0105}$
FFCL Li et al. (2022)	$0.4673_{\pm 0.0158}$	$0.3987_{\pm 0.0383}$	$0.4699_{\pm 0.0145}$	$0.7277_{\pm 0.0182}$	$0.7016_{\pm 0.0291}$	$0.7614_{\pm 0.0197}$
ST-Trans Song et al. (2021)	$0.4810_{\pm 0.0079}$	$0.4137_{\pm 0.0133}$	$0.4795_{\pm 0.0096}$	$0.7381_{\pm 0.0205}$	$0.7013_{\pm 0.0352}$	$0.7681_{\pm 0.0175}$
BIOT Yang et al. (2023a)	$0.5118_{\pm 0.0118}$	$0.4476_{\pm 0.0254}$	$0.5136_{\pm 0.0112}$	$0.7527_{\pm 0.0121}$	$0.7192_{\pm 0.0231}$	$0.7790_{\pm 0.0146}$
LaBraM-B Jiang et al. (2024)	$0.5273_{\pm 0.0107}$	$0.4698_{\pm 0.0188}$	$0.5288_{\pm 0.0102}$	$0.7633_{\pm 0.0102}$	$0.7231_{\pm 0.0182}$	$0.7810_{\pm 0.0133}$
CBraMod Wang et al. (2025)	$\underline{0.5509}_{\pm 0.0089}$	$\underline{0.5041}_{\pm 0.0122}$	$\underline{0.5618}_{\pm 0.0093}$	$\underline{0.7865}_{\pm 0.0110}$	$\underline{0.7442}_{\pm 0.0152}$	$\underline{0.8011}_{\pm 0.0099}$
Ours	$\mathbf{0.5731}_{\pm 0.0092}$	$\mathbf{0.5206}_{\pm 0.0102}$	$\mathbf{0.5776}_{\pm 0.0114}$	$\mathbf{0.7931}_{\pm 0.0097}$	$\mathbf{0.7583}_{\pm 0.0155}$	$\mathbf{0.8012}_{\pm 0.0092}$

recognition on **FACED** Chen et al. (2023) (10,332 10-second samples, 32 channels, 250 Hz, nine classes); sleep staging on **ISRUC** Khalighi et al. (2016) (89,240 30-second samples, 6 channels, 200 Hz, five stages); and imagined speech classification on **BCIC2020-3** Jeong et al. (2022) (6,000 3-second samples, 64 channels, 256 Hz, five classes). We report AUROC, F1 score (F1), balanced accuracy (BA), Cohens kappa (C-Kappa), and weighted F1 (W-F1).

Implementation Details. For seizure diagnosis, we adopt a Mamba-based baseline Kotoge et al. (2026), while a Transformer-based baseline Wang et al. (2025) is used for the remaining applications. For all datasets, we follow their official data splits and the corresponding baseline protocols. We set $\lambda_{\text{par}} = 1$, $\lambda_{\text{suf}} = 2.5$, and the polynomial order $K = 3$, while α and δ are automatically calibrated from training statistics.[3] We train all models using AdamW with a learning rate of 5e^{-4} and weight decay of 5e^{-2}, with a batch size of 128. The maximum number of training epochs is set to 100 for TUSZ and 50 for the other datasets, with early stopping using a patience of 20 epochs. All experiments are repeated over five random seeds on NVIDIA A6000 and RTX 5090 GPUs.

3.2 Comparison with State-of-the-arts

We compare CounterField against representative state-of-the-arts across the evaluated EEG applications. As reported on Tab. 1-3, we have the following observations:

(1) Consistent gains across different backbone architectures and tasks demonstrate the broad effectiveness of counteracting detrimental propagation along spatiotemporal dynamics.
(2) CounterField further exhibits advantages over recent foundation models across diverse EEG applications, suggesting that counteracting task-irrelevant propagation provides gains beyond increasingly expressive representation learning.
(3) Larger gains are observed on datasets with richer inter-electrode relational structure, where increased complexity may introduce more nuisances and intensify their spread. Such intractability in turn amplifies the benefit of counteraction in distilling predictive signatures.

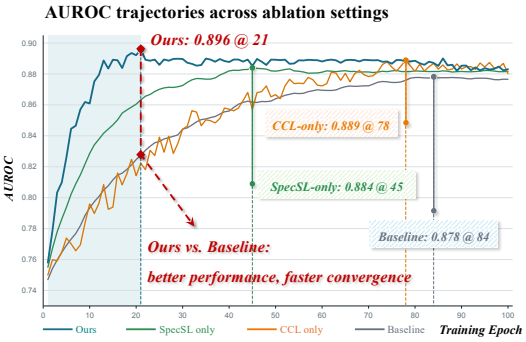
[3]The volume coefficient α is set to the inverse of the median node-state energy on the training split, whereas δ is obtained by scaling the median raw representation volume with a fixed relative compression target. Both of them are computed once and held fixed thereafter.

Table 3: Performance comparison on BCIC2020-3, with baseline shaded.

Method	BCIC2020-3, 5-class		
	BA	C-Kappa	W-F1
EEGNet Lawhern et al. (2018)	$0.441_{\pm 0.010}$	$0.302_{\pm 0.012}$	$0.441_{\pm 0.010}$
SPaRCNet Jing et al. (2023)	$0.443_{\pm 0.016}$	$0.303_{\pm 0.023}$	$0.442_{\pm 0.011}$
ContraWR Yang et al. (2023b)	$0.426_{\pm 0.016}$	$0.308_{\pm 0.022}$	$0.441_{\pm 0.018}$
CNN-Trans Peh et al. (2022)	$0.453_{\pm 0.009}$	$0.317_{\pm 0.012}$	$0.451_{\pm 0.013}$
FFCL Li et al. (2022)	$0.468_{\pm 0.020}$	$0.330_{\pm 0.036}$	$0.469_{\pm 0.021}$
ST-Trans Song et al. (2021)	$0.413_{\pm 0.012}$	$0.294_{\pm 0.016}$	$0.425_{\pm 0.014}$
BIOT Yang et al. (2023a)	$0.492_{\pm 0.009}$	$0.365_{\pm 0.018}$	$0.492_{\pm 0.008}$
LaBraM-B Jiang et al. (2024)	$0.506_{\pm 0.016}$	$0.380_{\pm 0.024}$	$0.505_{\pm 0.021}$
CBraMod Wang et al. (2025)	$0.537_{\pm 0.011}$	$0.422_{\pm 0.016}$	$0.538_{\pm 0.010}$
Ours	$\mathbf{0.562}_{\pm 0.009}$	$\mathbf{0.451}_{\pm 0.008}$	$\mathbf{0.564}_{\pm 0.011}$

Table 4: Balanced accuracy obtained with different methods on ISRUC, FACED, and BCIC2020.

Method	ISRUC	FACED	BCIC2020-3
Baseline	0.7685	0.5509	0.5373
GD/RS	0.7488	0.5281	0.5131
GC	0.7621	0.5512	0.5294
GIB	0.7524	0.5527	0.5312
DGC	0.7711	0.5503	0.5279
DGIB	0.7674	0.5534	0.5406
CounterField	**0.7931**	**0.5643**	**0.5624**



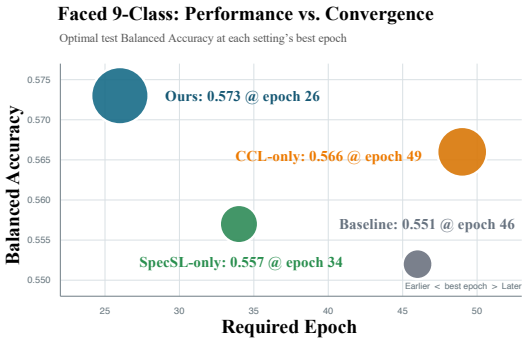


Figure 2: AUROC trajectories and Performance-convergence across ablation settings.

3.3 Ablation Study

As shown in Figs. 2, we disentangle the contributions of SpecSL and CCL by investigating their predictive gains and learning dynamics through training trajectories and required training epochs.

(1) Compared with the baseline, CounterField consistently achieves superior performance at substantially earlier training stages (*e.g.,* 0.896 AUROC at epoch 21 vs. 0.878 at epoch 84), demonstrating joint improvements in predictive performance and learning efficiency.
(2) The comparison between the full model and CCL-only setting highlights the accelerating and stabilizing effects of SpecSL. We conjecture this is because its structured spectral geometry better aligns with EEG dynamics, reducing unnecessary exploration while enabling more targeted counteraction against task-irrelevant patterns.
(3) In contrast, CCL continually challenges the current solution through counteractive refinement, driving the model toward stronger predictive states but with increased fluctuations.
(4) Together, the distinct yet complementary roles of SpecSL and CCL enable CounterField to jointly achieve superior performance, faster convergence, and stable learning.

3.4 Understanding Counteraction Dynamics

To further elucidate the mechanism of CounterField, we investigate how the parsimony and sufficiency objectives regulate counteraction. As shown in Fig. 3 (a) & (b), the top and bottom panels report predictive performance and counteraction degree[4], respectively. We fix one coefficient ($\lambda_{\text{suf}} = 2.5$ or $\lambda_{\text{par}} = 1.0$) and vary the other accordingly. Based on Fig. 3 (a), we show:

(1) Unconstrained regime ($\lambda_{\text{suf}} = 0$)**.** CounterField exhibits the strongest counteraction but underperforms the unrefined baseline, indicating that unaligned counteraction may compromise predictive information. **(2) Calibrated regime** ($0 < \lambda_{\text{suf}} \leq 2.5$)**.** Increasing λ_{suf} progressively moderates counteraction while improving performance, showing how sufficiency calibrates refinement against predictive preservation. **(3) Over-conservative regime** ($\lambda_{\text{suf}} > 2.5$)**.** Further strengthening suf-

[4]Counteraction is quantified by the relative reduction in transition energy: $C_\Delta(t) = 1 - \frac{|\Delta H_t + C_{\phi,t}|_{F^2}}{|\Delta H_t|_{F^2}}$, where positive values indicate net cancellation and negative values indicate amplification.

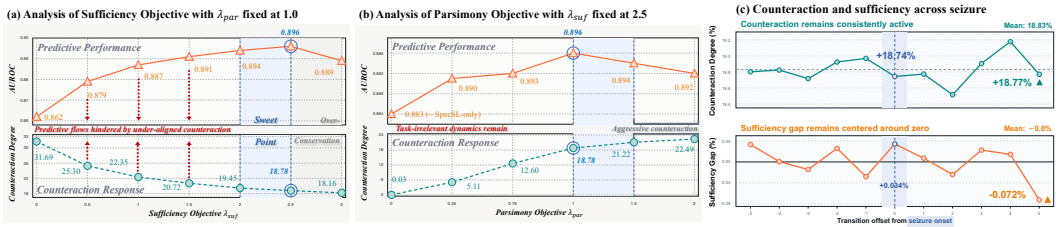


Figure 3: Counteraction behavior under objective regulation and across seizure onset.

ficiency suppresses counteraction with no additional performance gain, suggesting that excessive preservation limits effective refinement.

We then fix λ_{suf} and vary λ_{par} to examine how parsimony regulates counteraction strength.

(1) Beneficial regime ($0 \leq \lambda_{\text{par}} \leq 1$)**.** Counteraction and performance increase together, indicating that moderate refinement is beneficial under the sufficiency constraint. **(2) Balanced regime** ($1 < \lambda_{\text{par}} < 1.5$)**.** CounterField maintains near-optimal performance as counteraction continues to increase, marking a favorable balance between refinement and predictive preservation. **(3) Aggressive regime** ($\lambda_{\text{par}} \geq 1.5$)**.** Further counteraction yields diminishing returns and eventually degrades performance, suggesting that excessive refinement may compromise predictive information.

Next, we move beyond parameter-wise analysis and unfold the identified optimum ($\lambda_{\text{suf}} = 2.5$ or $\lambda_{\text{par}} = 1.0$) along the temporal axis. As shown in Fig. 3(c), we track counteraction degree and the sufficiency gap[5] over five states before and after seizure onset. Key observations are as follows:

(1) **Counteraction keeps persistently active** throughout the onset neighborhood, with an average degree of 18.83%, confirming the near-zero sufficiency gap is not achieved through trivial intervention; (2) Despite this substantial intervention, **the sufficiency gap remains tightly centered around zero**, with a maximum positive deviation of only 0.034%, indicating that CounterField substantially reshapes transition dynamics while almost perfectly preserving predictive sufficiency.

Is Outcome-Level Compactness Enough? As shown in Tab. 4, we compare CounterField with representative compact graph learning strategies, including graph density (GD) Fang et al. (2019a), recurring substructure (RS) Yan & Han (2002), graph condensation (GC) Jin et al. (2022), graph information bottleneck (GIB) Wu et al. (2020), and their dynamic variants D-GC Chen et al. (2025) and D-GIB Fang et al. (2025). Key observations are as follows:

(1) Outcome-level compactness does not reliably translate into predictive gains. GD/RS consistently degrades performance, while GC and GIB yield at most marginal improvements over the baseline.
(2) Incorporating additional temporal graph context, as in DGC and DGIB, partially alleviates this limitation, but the resulting gains remain modest and inconsistent across datasets.
(3) In contrast, CounterField consistently achieves substantially larger gains across all three datasets. This contrast suggests that the key lies not merely in how compactness is formulated, but in where intervention occurs, namely, along the propagation process rather than on its isolated outcomes.

4 Conclusion

In this work, we investigate task-irrelevant signal propagation in EEG and argue that mitigating its detrimental effects requires intervention along spatiotemporal dynamics rather than refining the resulting representations or relational structures. To this end, we propose CounterField, a propagation-level framework that builds on graph SPDEs to formulate counteraction over EEG transitions while preserving predictive information. Extensive experiments demonstrate consistent improvements across different tasks and architectures, while further analyses reveal sustained counteraction with negligible loss of predictive sufficiency. These findings support a shift from refining propagation outcomes toward counteracting detrimental effects along the propagation process itself.

[5]Defined as $G_{\text{suf}} = (\overline{\mathcal{L}}_{\text{task}}^{\text{ref}} - \overline{\mathcal{L}}_{\text{task}}^{\text{raw}})/\overline{\mathcal{L}}_{\text{task}}^{\text{raw}}$, where positive, zero, and negative values indicate an increase, no change, and a decrease in the task objective, respectively. Results are averaged over 600 test samples.

5 AI USE STATEMETN

Generative AI tools were used solely to improve the readability of the manuscript and identify typographical errors. We have not used generative AI tools for research, ideation, methodology, or experiments. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Viacheslav Borovitskiy, Iskander Azangulov, Alexander Terenin, Peter Mostowsky, Marc Deisenroth, and Nicolas Durrande. Matérn gaussian processes on graphs. In *International Conference on Artificial Intelligence and Statistics*, pp. 2593–2601, 2021.

Ben Chamberlain, James Rowbottom, Maria I Gorinova, Michael Bronstein, Stefan Webb, and Emanuele Rossi. Grand: Graph neural diffusion. In *International conference on machine learning*, pp. 1407–1418. PMLR, 2021.

Dong Chen, Shuai Zheng, Yeyu Yan, Muhao Xu, Zhenfeng Zhu, Yao Zhao, and Kunlun He. Dynamic graph condensation. *CoRR*, abs/2506.13099, 2025.

Jingjing Chen, Xiaobin Wang, Chen Huang, Xin Hu, Xinke Shen, and Dan Zhang. A large finer-grained affective computing eeg dataset. *Scientific Data*, 10(1):740, 2023.

Ian C. Covert, Balu Krishnan, Imad Najm, Jiening Zhan, Matthew Shore, John Hixson, and Ming Jack Po. Temporal graph convolutional networks for automatic seizure detection. In *Proceedings of the Machine Learning for Healthcare Conference*, volume 106, pp. 160–180, 2019.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29, 2016.

Shifei Ding, Wei Du, Ling Ding, Lili Guo, and Jian Zhang. Learning efficient and robust multi-agent communication via graph information bottleneck. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 17346–17353, 2024.

Moshe Eliasof, Eldad Haber, and Eran Treister. Pde-gcn: Novel architectures for graph neural networks motivated by partial differential equations. *Advances in neural information processing systems*, 34:3836–3849, 2021.

Hui Fang, Haishuai Wang, Yang Gao, Yonggang Zhang, Jiajun Bu, Bo Han, and Hui Lin. Insgnn: Interpretable spatio-temporal graph neural networks via information bottleneck. *Inf. Fusion*, 119:102997, 2025.

Yixiang Fang, Kaiqiang Yu, Reynold Cheng, Laks V. S. Lakshmanan, and Xuemin Lin. Efficient algorithms for densest subgraph discovery. *Proc. VLDB Endow.*, 12(11):1719–1732, 2019a.

Yixiang Fang, Kaiqiang Yu, Reynold Cheng, Laks VS Lakshmanan, and Xuemin Lin. Efficient algorithms for densest subgraph discovery. *arXiv preprint arXiv:1906.00341*, 2019b.

Jianfei Gao and Bruno Ribeiro. On the equivalence between temporal and static equivariant graph representations. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning*, volume 162, pp. 7052–7076, 2022.

Jeff Hawkins and Subutai Ahmad. Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Frontiers in neural circuits*, 10:23, 2016.

Van Thuy Hoang and O-Joun Lee. Pre-training graph neural networks on molecules by using subgraph-conditioned graph information bottleneck. *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 17204–17213, 2025.

Ji-Hoon Jeong, Jeong-Hyun Cho, Young-Eun Lee, Seo-Hyun Lee, Gi-Hwan Shin, Young-Seok Kweon, José del R. Millán, Klaus-Robert Müller, and Seong-Whan Lee. 2020 international braincomputer interface competition: A review. *Frontiers in Human Neuroscience*, Volume 16 - 2022, 2022.

Wei-Bang Jiang, Liming Zhao, and Bao-liang Lu. Large brain model for learning generic representations with tremendous eeg data in bci. In *International Conference on Learning Representations*, pp. 16405–16426, 2024.

Weibang Jiang, Yansen Wang, Bao-Liang Lu, and Dongsheng Li. Neurolm: A universal multi-task foundation model for bridging the gap between language and EEG signals. In *International Conference on Learning Representations*, 2025.

Wei Jin, Lingxiao Zhao, Shichang Zhang, Yozen Liu, Jiliang Tang, and Neil Shah. Graph condensation for graph neural networks. In *International Conference on Learning Representations*, 2022.

Jin Jing, Wendong Ge, Shenda Hong, Marta Bento Fernandes, Zhen Lin, Chaoqi Yang, Sungtae An, Aaron F Struck, Aline Herlopian, Ioannis Karakis, et al. Development of expert-level classification of seizures and rhythmic and periodic patterns during eeg interpretation. *Neurology*, 100(17):e1750–e1762, 2023.

Nikhil S Ketkar, Lawrence B Holder, and Diane J Cook. Subdue: Compression-based frequent pattern discovery in graph data. In *Proceedings of the 1st international workshop on open source data mining: frequent pattern mining implementations*, pp. 71–76, 2005.

Sirvan Khalighi, Teresa Sousa, José Moutinho dos Santos, and Urbano Nunes. Isruc-sleep: A comprehensive public dataset for sleep researchers. *Comput. Methods Programs Biomed.*, 124:180–192, 2016.

Rikuto Kotoge, Zheng Chen, Tasuku Kimura, Yasuko Matsubara, Takufumi Yanagisawa, Haruhiko Kishima, and Yasushi Sakurai. Evobrain: Dynamic multi-channel eeg graph modeling for time-evolving brain networks. In *Advances in Neural Information Processing Systems*, volume 38, pp. 144514–144543, 2026.

Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 15(5):056013, 2018.

Hongli Li, Man Ding, Ronghua Zhang, and Chunbo Xiu. Motor imagery eeg classification algorithm based on cnn-lstm feature fusion network. *Biomedical signal processing and control*, 72:103342, 2022.

Katharina Limbeck, Lydia Mezrag, Guy Wolf, and Bastian Rieck. Geometry-aware edge pooling for graph neural networks. In *Advances in Neural Information Processing Systems*, 2025.

Xixun Lin, Wenxiao Zhang, Fengzhao Shi, Chuan Zhou, Lixin Zou, Xiangyu Zhao, Dawei Yin, Shirui Pan, and Yanan Cao. Graph neural stochastic diffusion for estimating uncertainty in node classification. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 30457–30478, 2024.

Fang Liu, Pei Yang, Yezhi Shu, Niqi Liu, Jenny Sheng, Junwen Luo, Xiaoan Wang, and Yong-Jin Liu. Emotion recognition from few-channel EEG signals by integrating deep feature aggregation and transfer learning. *IEEE Trans. Affect. Comput.*, 15(3):1315–1330, 2024.

Jiayi Luo, Qingyun Sun, Beining Yang, Haonan Yuan, Xingcheng Fu, Yanbiao Ma, Jianxin Li, and Philip S. Yu. Robust graph condensation via classification complexity mitigation. In *Advances in Neural Information Processing Systems*, 2025.

Yue-Ting Pan, Jing-Lun Chou, and Chun-Shu Wei. Matt: A manifold attention network for eeg decoding. In *Advances in Neural Information Processing Systems*, volume 35, pp. 31116–31129, 2022.

Aldo Pareja, Giacomo Domeniconi, Jie Chen, Tengfei Ma, Toyotaro Suzumura, Hiroki Kanezashi, Tim Kaler, Tao B. Schardl, and Charles E. Leiserson. Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In *Association for the Advancement of Artificial Intelligence*, pp. 5363–5370, 2020.

Wei Yan Peh, Yuanyuan Yao, and Justin Dauwels. Transformer convolutional neural networks for automated artifact detection in scalp eeg. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3599–3602. IEEE, 2022.

Huy Phan, Kaare B. Mikkelsen, Oliver Y. Chén, Philipp Koch, Alfred Mertins, and Maarten De Vos. Sleeptransformer: Automatic sleep staging with interpretability and uncertainty quantification. *IEEE Trans. Biomed. Eng.*, 69(8):2456–2467, 2022.

Shivarudhrappa Raghu, Natarajan Sriraam, Yasin Temel, Shyam Vasudeva Rao, and Pieter Leonard Kubben. EEG based multi-class seizure type classification using convolutional neural network and transfer learning. *Neural Networks*, 124:202–212, 2020.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, 2006.

Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069, 2010.

Vinit Shah, Eva Von Weltin, Silvia Lopez de Diego, James Riley McHugh, Lillian Veloso, Meysam Golmohammadi, Iyad Obeid, and Joseph Picone. The temple university hospital seizure detection corpus. *Frontiers Neuroinformatics*, 12:83, 2018.

Yonghao Song, Xueyu Jia, Lie Yang, and Longhan Xie. Transformer-based spatial-temporal feature learning for eeg decoding. *arXiv preprint arXiv:2106.11170*, 2021.

Siyi Tang, Jared A Dunnmon, Khaled Saab, Xuan Zhang, Qianying Huang, Florian Dubost, Daniel L Rubin, and Christopher Lee-Messer. Self-supervised graph neural networks for improved electroencephalographic seizure analysis. *arXiv preprint arXiv:2104.08336*, 2021.

Siyi Tang, Jared Dunnmon, Khaled Kamal Saab, Xuan Zhang, Qianying Huang, Florian Dubost, Daniel Rubin, and Christopher Lee-Messer. Self-supervised graph neural networks for improved electroencephalographic seizure analysis. In *International Conference on Learning Representations*, 2022.

Siyi Tang, Jared A. Dunnmon, Liangqiong Qu, Khaled Kamal Saab, Tina Baykaner, Christopher Lee-Messer, and Daniel L. Rubin. Modeling multivariate biosignals with graph neural networks and structured state space models. In *Conference on Health, Inference, and Learning*, volume 209, pp. 50–71, 2023.

Guangyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. EEGPT: pretrained transformer for universal and reliable representation of EEG signals. In *Advances in Neural Information Processing Systems*, 2024a.

Jialing Wang, Shiwei Cheng, Jieming Tian, and Yuefan Gao. A 2d CNN-LSTM hybrid algorithm using time series segments of EEG data for motor imagery classification. *Biomed. Signal Process. Control.*, 83:104627, 2023.

Jihong Wang, Yuxin Bai, Chunqiang Zhu, Hao Qian, Ziqi Liu, and Minnan Luo. Dynamic graph representation learning with disentangled information bottleneck. *Neural Networks*, 194:108056, 2026.

Jiquan Wang, Sha Zhao, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. Generalizable sleep staging via multi-level domain alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 265–273, 2024b.

Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. Cbramod: A criss-cross brain foundation model for EEG decoding. In *International Conference on Learning Representations*. OpenReview.net, 2025.

Yan Wang, Wenju Hou, Nan Sheng, Ziqi Zhao, Jialin Liu, Lan Huang, and Juexin Wang. Graph pooling in graph neural networks: methods and their applications in omics studies. *Artificial Intelligence Review*, 57 (11):294, 2024c.

Zixiao Wang and Jicong Fan. Graph classification via reference distribution learning: Theory and practice. In *Advances in Neural Information Processing Systems*, 2024.

Kangkang Wei, Weiqi Luo, Shunquan Tan, and Ji-Wu Huang. Ctnet: A convolutional transformer network for color image steganalysis. *J. Comput. Sci. Technol.*, 40(2):413–427, 2025a.

Xinxu Wei, Kanhao Zhao, Yong Jiao, Hua Xie, Lifang He, and Yu Zhang. Pre-training graph contrastive masked autoencoders are strong distillers for EEG. In *International Conference on Machine Learning*, volume 267, 2025b.

Tailin Wu, Hongyu Ren, Pan Li, and Jure Leskovec. Graph information bottleneck. In *Advances in Neural Information Processing Systems*, 2020.

Fred Xu and Thomas Markovich. Uncertainty estimation on graphs with structure informed stochastic partial differential equations. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Xifeng Yan and Jiawei Han. gspan: Graph-based substructure pattern mining. In *Proceedings of the International Conference on Data Mining*, pp. 721–724, 2002.

Chaoqi Yang, M. Brandon Westover, and Jimeng Sun. BIOT: biosignal transformer for cross-data learning in the wild. In *Advances in Neural Information Processing Systems*, 2023a.

Chaoqi Yang, Cao Xiao, M Brandon Westover, and Jimeng Sun. Self-supervised electroencephalogram representation learning for automatic sleep staging: model development and evaluation study. *JMIR AI*, 2(1): e46769, 2023b.

Ke Yi, Yansen Wang, Kan Ren, and Dongsheng Li. Learning topology-agnostic eeg representations with geometry-aware modeling. *Advances in Neural Information Processing Systems*, 36:53875–53891, 2023.

Junchi Yu, Tingyang Xu, Yu Rong, Yatao Bian, Junzhou Huang, and Ran He. Graph information bottleneck for subgraph recognition. In *Proceedings of the International Conference on Learning Representations*, 2021.

Guibin Zhang, Xiangguo Sun, Yanwei Yue, Chonghe Jiang, Kun Wang, Tianlong Chen, and Shirui Pan. Graph sparsification via mixture of graphs. In *International Conference on Learning Representations*, 2025.

A Preliminary and Related Work

A.1 EEG Analysis

EEG analysis spans diverse tasks, including seizure detection Kotoge et al. (2026), emotion recognition Liu et al. (2024), motor imagery classification Wei et al. (2025a), and sleep staging Wang et al. (2024b). A common paradigm organizes EEG signals or spectrograms as regular tensors or sequences, modeling temporal and spectral patterns with CNNs Raghu et al. (2020), RNNs Wang et al. (2023), or Transformers Phan et al. (2022). More recently, large-scale pretraining has enabled foundation models to learn transferable EEG representations across diverse datasets Wei et al. (2025b); Jiang et al. (2024; 2025). Despite their different focuses, these paradigms provide limited characterization of the relational organization underlying multichannel EEG, which constitutes an important source of discriminative information given the coordinated interactions across functionally related brain regions Hawkins & Ahmad (2016); Rubinov & Sporns (2010). Accordingly, graph-based approaches explicitly model inter-electrode structure through predefined topologies Yi et al. (2023), hierarchical organization Covert et al. (2019), or learned relations Tang et al. (2021), showing promising gains in prediction and interpretability.

A.2 Compact Graph Representation Learning

Compact graph representation learning aims to preserve task-relevant information and structures while discarding redundant content, promising improved generalization, robustness, and interpretability beyond empirical performance. To achieve this, traditional methods employ hand-crafted criteria, such as graph density Fang et al. (2019b) or recurring substructures Ketkar et al. (2005), to determine which components to retain. More recent approaches learn compact graphs in a data-driven manner, including graph sparsification Zhang et al. (2025), pooling Wang et al. (2024c), and condensation Luo et al. (2025). Despite their empirical effectiveness, existing analyses Limbeck et al. (2025) suggest that these methods often rely on task-specific heuristics and may discard structurally important information due to the lack of a unified characterization of redundancy.

To complement this, another line of research adopts the information bottleneck (IB) principle Yu et al. (2021); Hoang & Lee (2025); Wang & Fan (2024); Ding et al. (2024) to learn a minimal sufficient representation or subgraph Z: $\mathcal{L}_{\mathrm{IB}} := I(X;Z) - \beta I(Y;Z)$, where β balances compression against predictive sufficiency. However, faithfully instantiating IB on graphs remains challenging, as existing estimation paradigms are mainly developed under Euclidean *i.i.d.* settings and are difficult to align with non-Euclidean topology Wu et al. (2020).

Beyond static graphs, recent efforts have extended GIB Fang et al. (2025); Wang et al. (2026) and graph condensation Chen et al. (2025) to dynamic settings, where information not only propagates across nodes but also accumulates along the time axis.

Despite their advances in compact graph learning, the above formulations share a fundamental limitation: they refine structural or representational outcomes rather than the propagation that produces them. Consequently, detrimental signals may already be broadcast and become entangled with predictive cues when such refinement takes effect, making their removal prone to collateral damage.

A.3 Graph Diffusion and Stochastic Propagation

Graph diffusion provides a continuous-time perspective on information propagation over graphs Eliasof et al. (2021); Chamberlain et al. (2021), typically describing the evolution of node representations $\mathbf{H}(t)$ at propagation depth t through:

$$\frac{\partial \mathbf{H}(t)}{\partial t} = \mathrm{div}[\mathcal{G}_\theta(\mathbf{H}(t), t)\nabla \mathbf{H}(t)], \tag{24}$$

where ∇ and div denote the graph gradient and divergence operators, respectively, and $\mathcal{G}_\theta$ parameterizes the diffusivity over the graph. Building on this foundation, subsequent works Lin et al. (2024); Xu & Markovich (2025) further extend the deterministic graph diffusion to stochastic propagation for uncertainty estimation:

$$d\mathbf{H}(t) = \underbrace{f(\mathbf{H}(t))dt}_{\textit{Drift term}} + \underbrace{g(\mathbf{H}(t))d\mathbf{W}(t)}_{\textit{Stochastic term}}. \tag{25}$$

Here, the drift term models deterministic graph diffusion, while the stochastic term serves as a structured perturbation carrier, primarily used for uncertainty modeling in existing works. More importantly, this formulation provides a principled means of structuring perturbations over graph geometry through tractable spectral priors. Building on this perspective, we further shift compactness from cleaning up contaminated outcomes to inhibiting the propagation of detrimental components.

A.4 Graph Diffusion and Stochastic Propagation

Graph diffusion provides a continuous-time perspective on information propagation over graphs Eliasof et al. (2021); Chamberlain et al. (2021), typically describing the evolution of node representations $\mathbf{H}(t)$ at propagation depth t through:

$$\frac{\partial \mathbf{H}(t)}{\partial t} = \mathrm{div}[\mathcal{G}_\theta(\mathbf{H}(t), t)\nabla \mathbf{H}(t)], \tag{26}$$

where ∇ and div denote the graph gradient and divergence operators, respectively, and $\mathcal{G}_\theta$ parameterizes the diffusivity over the graph. Building on this foundation, subsequent works Lin et al. (2024); Xu & Markovich (2025) further extend the deterministic graph diffusion to stochastic propagation for uncertainty estimation:

$$d\mathbf{H}(t) = \underbrace{f(\mathbf{H}(t))dt}_{\textit{Drift term}} + \underbrace{g(\mathbf{H}(t))d\mathbf{W}(t)}_{\textit{Stochastic term}}. \tag{27}$$

Here, the drift term models deterministic graph diffusion, while the stochastic term serves as a structured perturbation carrier, primarily used for uncertainty modeling in existing works. More importantly, this formulation provides a principled means of structuring perturbations over graph geometry through tractable spectral priors. Building on this perspective, we further shift compactness from cleaning up contaminated outcomes to inhibiting the propagation of detrimental components.