Vibe-Spike: An Energy-Preserving EEG Foundation Model Through the Landscape of Neural Coherence

Jun Ye[1†] **Yuting Li**[1†] **Tingting Dan**[2*] **Guorong Wu**[2*]
[1]Bio-X Institutes, Key Laboratory for the Genetics of Development and Neuropsychiatric Disorders, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, China
[2]Departments of Computer Science and Psychiatry, UNC Chapel Hill, USA
{Tingting_Dan,grwu}@med.unc.edu

Abstract

Current foundation models of electroencephalography (EEG) sought to characterize the fluctuations in neural signals, yet this objective stands fundamentally at odds with the intrinsic physiological mechanisms governing brain activity. As with their LLM predecessors, these models pack millions of neurons, even those contributing negligibly small weights, yielding a computational footprint incompatible with resource-constrained portable brain-computer interfaces (BCIs). To overcome these challenges, we propose *Vibe-Spike*, an energy-preserving and biologically inspired EEG foundation model operating in the regime of oscillatory neural synchronization. Drawing on principles from cognitive neuroscience, our foundation model is designed to recover the masked EEG signals through the transient spikes of spatially coupled neurons, from which self-organized cortical synchronization (aka. *brain rhythms*) emerges to reciprocally modulate ongoing neural firings. Under the hood, we introduce a macro-micro learning mechanism that bridges two scales of neural computation: lower layers capture neural oscillations via a spiking neural network (SNN), while upper layers model oscillatory synchronization across large neuronal populations, with cross-scale interactions learned end-to-end. Taken together, our *Vibe-Spike* sheds new light on learning intrinsic representation of functional neural fluctuations. We pre-train the model to reconstruct segmented EEG signals directly from the synchronized spiking representations using a purely self-supervised objective on a diverse corpus of **75** public datasets. To demonstrate its universal representational power, we conduct massive-scale evaluations across ten heterogeneous downstream datasets, including clinical pathology, sleep staging, and cognitive assessment. *Vibe-Spike* not only facilitates deployment on portable edge devices through its lightweight architecture and energy-preserving inference, but also provides profound neuroscience insights. By further capturing pre- and post-synaptic firing, *Vibe-Spike* unlocks a new pathway for characterizing effective interactions and the oscillatory synchronization that underpins cognition.

1 Introduction

Electroencephalography (EEG) serves as a paramount non-invasive window into the human brain, underpinning critical applications in brain-computer interfaces (BCIs), continuous sleep monitoring, and neurological disorder diagnosis. Despite significant strides driven by deep representation learning, the development of universal, highly generalizable BCI systems remains fundamentally bottlenecked by the inherently low signal-to-noise ratio, extreme non-stationarity, and severe inter-subject variability of EEG signals.

*Corresponding authors.
†Equal contribution. Code: github.com/acmlab/Vibe-Spike.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

To tackle these real-world challenges, existing EEG decoding methodologies have evolved along two distinct trajectories. The first encompasses task-specific, non-foundation models designed with complex representational architectures, such as EEGNet [Lawhern et al., 2018], EEGConformer [Song et al., 2023], ST-Transformer [Song et al., 2021], and CNN-Transformer [Peh et al., 2022] for spatiotemporal feature extraction, alongside contrastive learning paradigms like SPaRCNet [Jing et al., 2023], ContraWR [Yang et al., 2023b], and FFCL [Li et al., 2022]. While effective within closed-domain datasets, these models severely lack the capacity to generalize across unseen subjects or novel clinical scenarios. To shatter this barrier, a recent wave of large language model (LLM)-inspired EEG foundation models has emerged. Pioneering frameworks, including BIOT [Yang et al., 2023a], LaBraM [Jiang et al., 2024], CBraMod [Wang et al., 2025], and REVE [Ouahidi et al., 2026], leverage massive-scale self-supervised pre-training on unlabeled data to establish new performance baselines, demonstrating unprecedented cross-dataset representational power.

Nevertheless, current LLM-inspired foundation models are fundamentally crippled by two critical limitations. First, their heavy reliance on massive Transformer-based architectures incurs immense computational overhead and energy consumption, limiting their deployment on low-power, portable edge BCI devices. Second, and perhaps more crucially, these purely data-driven models treat EEG merely as generic time-series data. By blindly fitting superficial signal appearances, they operate as "black boxes" that completely ignore the intrinsic physiological mechanisms governing brain activity. This lack of biological plausibility not only obscures structural interpretability but also hinders these models from capturing the authentic dynamic nature of the human cortex.

To bridge the chasm between deep representation learning and biological reality, we propose *Vibe-Spike*, an energy-preserving and biologically inspired EEG foundation model grounded in human brain dynamics, as conceptually illustrated in Fig. 1. Rather than merely fitting signal waveforms, our architecture elegantly couples two neural learning mechanisms depicted in the figure. At the microscopic level, a spiking neural network (SNN) is integrated to model intrinsic, energy-efficient *sparse firing* patterns, which, as indicated by the green arrow in Fig. 1, can be directly harnessed for low-power edge BCIs. At the macroscopic level, a neural synchronization network coordinates phase modulation across distinct brain regions by capturing *oscillatory synchronization* (orange arrow), where the well-studied Kuramoto model [Acebrón et al., 2005, Miyato et al., 2024] is employed to shape the phase dynamics. Through this seamless macro-micro coupling mechanism, *Vibe-Spike* yields an intrinsic representation of functional neural fluctuations, enabling the downstream tasks of cognitive assessment, emotion recognition, and disease diagnosis listed in Fig. 1. During pre-training, the model is optimized to reconstruct segmented EEG signals directly from the synchronized spiking representations.

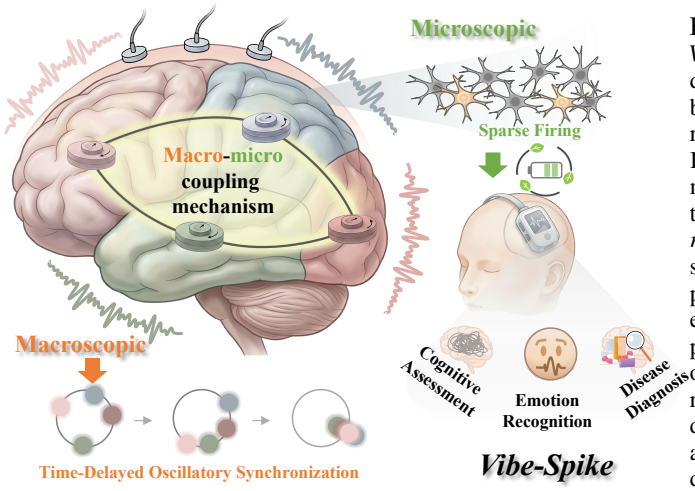

Figure 1: Conceptual framework of *Vibe-Spike*. Our proposed EEG foundation model bridges macroscopic brain dynamics and microscopic neural behaviors. *Bottom-left*. A Kuramoto-driven network captures macroscopic oscillatory synchronization across distinct brain regions. *Top-right.* A SNN models microscopic sparse firing patterns, ensuring energy-preserving computations for portable edge BCIs. *Bottom-right*. By coupling these macro-micro dynamics, our model learns biologically inspired representations to empower diverse downstream tasks, including cognitive assessment, emotion recognition, and disease diagnosis.

The main contributions of this work are summarized as follows:

• **A biologically inspired foundation model:** We introduce *Vibe-Spike*, pioneering the integration of macroscopic oscillatory synchronization with microscopic spiking dynamics. It transcends the traditional data-driven paradigm by offering a physics-informed, neuroscience-inspired architecture for universal EEG representation learning.

• **Lightweight architecture and energy-preserving inference:** Unlike traditional foundation models that suffer from massive parameter bloat, *Vibe-Spike* features a lightweight parameter footprint. By leveraging the sparse, event-driven nature of its SNN backbone, our model minimizes active energy consumption, making it highly amenable to deployment on neuromorphic and portable edge BCIs.

• **Novel methodology for effective interactions:** By structurally tracking the incidence of pre- and post-synaptic neuron firings, *Vibe-Spike* introduces a new paradigm to uncover the effective interactions and neural oscillatory synchronization underlying cognitive states.

• **Massive-scale empirical validation:** To show its universal representational power, we conduct extensive evaluations across ten heterogeneous downstream datasets, including clinical pathology, sleep staging, and cognitive assessment. *Vibe-Spike* not only achieves competitive performance against computationally heavy foundation models but also sets a new standard for structural interpretability.

2 Method

The design of *Vibe-Spike* is underpinned by the brain-inspired learning principle of macro-micro architecture [Dan and Wu, 2026], aiming to bridge the gap between abstract signal processing and intrinsic neural dynamics. Unlike previous EEG foundation models that treat brain signals as generic time-series data, our framework treats the presence of fluctuations in EEG as a manifestation of self-organized cortical synchronization. Specifically, we conceptualize the representation learning process as a dual-layer information exchange: a microscopic layer capturing energy-efficient neural firing via SNNs, and a macroscopic layer modeling rhythmic coherence through Kuramoto-driven oscillatory dynamics. This biologically inspired coupling ensures that the learned embeddings are not only computationally efficient for edge-device deployment but also structurally interpretable in the regime of cognitive neuroscience.

The network architecture of *Vibe-Spike* is conceptualized as an end-to-end generative pipeline, as illustrated in Fig. 2. The workflow includes four primary stages: (1) *Unified Spatial Alignment (SA)* for resolving heterogeneous montages, (2) *Time-Frequency Tokenization (TFT)* for spectral feature extraction, (3) *Neural Coherence Encoding (NCE)* to capture intrinsic oscillatory dynamics, and (4) *Generative Spectral Recovery (GSR)* as the self-supervised objective.

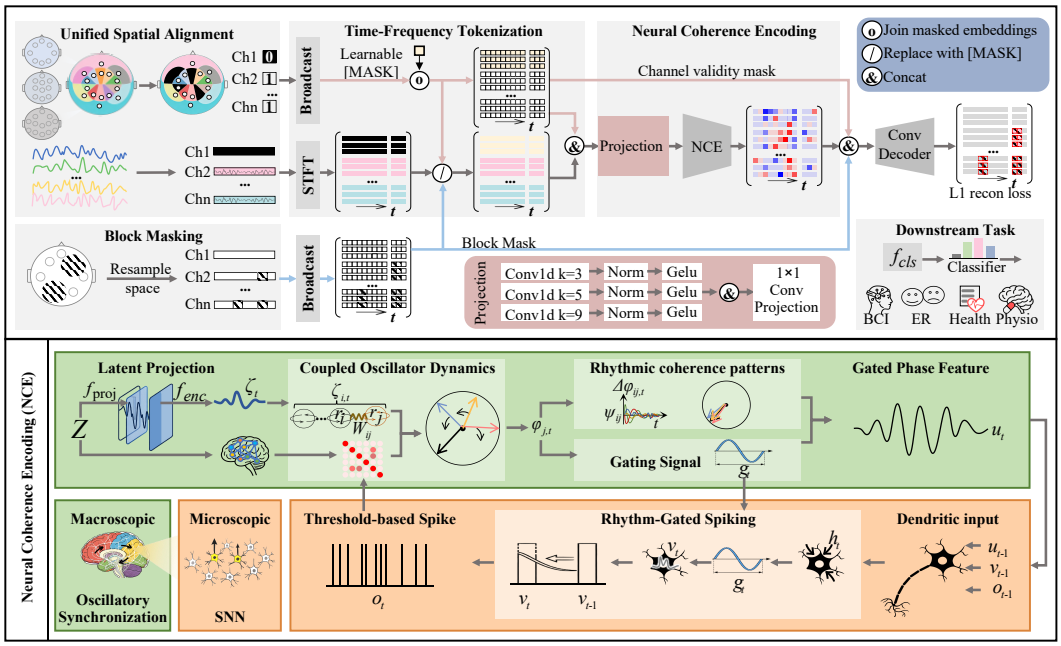

Figure 2: The overall network architecture of our *Vibe-Spike*.

2.1 Unified Spatial Alignment (SA)

To handle the heterogeneity of electrode layouts across diverse EEG datasets, we first establish a canonical target space $\mathcal{P}_{tgt} = \{p^i_{tgt}\}_{i=1}^{C_{tgt}}$ defined on a standardized head shell. Given an input EEG segment $X \in \mathbb{R}^{C_{src} \times T}$ with source coordinates P_{src}, the SA module performs shell-projected resampling. Specifically, each source electrode p^i_{src} is projected onto the target shell via radial normalization, and subsequently assigned to its nearest target electrode p^j_{tgt} via a Voronoi-like partition. The aligned signal X_{tgt} and its corresponding channel-validity mask $\mathbf{m}_c \in \{0,1\}^{C_{tgt}}$ are obtained via the shell-projection function $\mathcal{R}_{shell}$:

$$X_{tgt}, \mathbf{m}_c = \mathcal{R}_{shell}(X, P_{src}; \mathcal{P}_{tgt}) \tag{1}$$

Specifically, the temporal features for the j-th target channel are formulated element-wise as:

$$(X_{tgt})_{j,:} = \begin{cases} x_{i^*_j,:}, & \text{if target channel } j \text{ is covered,} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

where i^*_j denotes the closest source electrode to target channel j. This normalization ensures that all subsequent operations reside in a unified spatial manifold regardless of the original device's configuration.

2.2 Time-Frequency Tokenization (TFT)

Unlike models that operate on raw 1D waveforms, *Vibe-Spike* treats EEG as a compositional spectral entity. The aligned signal X_{tgt} is transformed via short-time Fourier transform (STFT) into a complex-valued spectrogram $S \in \mathbb{R}^{C_{tgt} \times 2 \times F \times T'}$, where F and T' denote the frequency bins and temporal frames, respectively. We denote the real and imaginary components as S^{re} and S^{im}, respectively, i.e., $S = [S^{\text{re}}, S^{\text{im}}]$ along the complex-component dimension. To facilitate self-supervised representation learning, we implement a joint spatio-temporal block masking strategy. Let $M_b \in \{0,1\}^{C_{tgt} \times 1 \times 1 \times T'}$ be a block mask. We combine it with the channel-validity mask M_c (broadcast from $\mathbf{m}_c$) to define the input "tokens":

$$S_{in} = (1 - M_b) \odot (S + (1 - M_c) \odot E_c) \tag{3}$$

where E_c is a learnable channel-missing embedding. This process converts the continuous signal into a masked representation $H_{in} = \text{Concat}(\text{Flatten}(S_{in}), \text{Flatten}(M_c))$, which serves as the discrete-like input for the neural backbone.

2.3 Neural Coherence Encoding (NCE)

NCE functions as the core engine that bridges macroscopic oscillations with microscopic spiking dynamics through three hierarchical steps:

1. Latent projection and structural prior: The input H_{in} is first projected into a compact latent space $Z = \Phi_{stem}(H_{in}) \in \mathbb{R}^{d \times T'}$ using a multi-scale convolutional stem. We derive a functional connectivity graph $W \in \mathbb{R}^{d \times d}$ by computing the temporal correlations among latent dimensions:

$$W = (1 - \rho_W)\tilde{W} + \rho_W I_d + \epsilon, \quad \tilde{W}_{ij} = \text{Softmax}(\max(\text{Corr}(W_i, W_j), 0)) \tag{4}$$

where I_d is the identity matrix to preserve self-connections, $\rho_W \in [0,1]$ is a residual weight for the identity connection. ϵ is a small constant used for numerical stability.

2. Coupled oscillator dynamics: Each latent dimension is modeled as an oscillator whose phase φ_t evolves according to the sensory input ζ_t (derived from Z via two deep neural networks $\zeta_t = f_{\text{enc}}(f_{\text{proj}}(Z_{:,t})) \in \mathbb{R}^{d \times D}$, see Appendix A for architecture details) and connectivity-aware Kuramoto dynamics:

$$\dot{\varphi}_{i,t} = r_i + \lambda \sum_{j=1}^{d} W_{ij} \sin(\varphi_{j,t} - \varphi_{i,t} - \psi_{ij}) + \eta_i \odot \sin(\zeta_{i,t} - \varphi_{i,t}) \tag{5}$$

Notably, ψ_{ij} denotes the phase lag that enforces partial synchronization, reflecting the intrinsic feedback from the microscopic spiking neuron dynamics described below. This formulation enables the model to capture the rhythmic coherence patterns essential for neural interpretability. Furthermore, we introduce a time delay τ to generate a rhythm signal $g_t = \frac{1}{2}(1 + \sin(\bar{\varphi}_{t-\tau}))$, where $\bar{\varphi}_{t-\tau}$ is

obtained by averaging $\varphi_{t-\tau}$ along the internal oscillator dimension D. The resulting g_t acts as a macroscopic gating mechanism to modulate the subsequent neuronal spiking activity.

Remark 1. Unlike static feature extractors that treat latent variables as independent vectors, Eq. (5) models the *temporal coordination* of brain states. By incorporating the structural prior W and the sensory drive strength η_i, the model ensures that the latent oscillators do not just "fit" the data, but rather "synchronize" with it. This allows the model to maintain a stable, rhythmically organized representation even in the presence of noise or missing segments in the raw EEG signal.

3. Rhythm-gated spiking: The synchronized phases are converted into sparse spike trains o_t via an SNN [Yan et al., 2025]. We use a rhythm gate $\widetilde{g}_t$ (derived from g_t: $\widetilde{g}_t = \sigma\left(f_{\text{gate}}(g_t)\right)$, where f_{gate} is a learnable projection layer) to modulate the Leaky Integrate-and-Fire (LIF) dynamics:

$$h_t = \kappa v_{t-1} + (1-\kappa)u_t - \vartheta o_{t-1}, \quad v_t = \widetilde{g}_t \odot h_t + (1-\widetilde{g}_t) \odot v_{t-1}, \quad o_t = \widetilde{g}_t \odot \mathcal{H}(v_t - \vartheta) \tag{6}$$

where $u_t = f_u(\sin(\varphi_t) \odot g_t)$ denotes the synaptic input, f_u is a learnable projection layer. The parameter κ controls the temporal leakage, while ϑ denotes the firing threshold that determines spike generation, and $\mathcal{H}(\cdot)$ represents the Heaviside step function. Together, this formulation ensures that the microscopic spiking activity is coupled with the underlying macroscopic brain rhythms.

Remark 2. The rhythm gate $\widetilde{g}_t$ functions as a phase-locked modulator that selectively enables membrane potential integration and spike triggering. By halting the membrane state v_t update during low-excitability phases, it restricts active computation to specific temporal windows. This temporal sparsity not only enforces a rigorous coupling between microscopic spiking and the macroscopic rhythmic context, but also underpins the energy-preserving nature of the model. A more detailed explanation of the NCE module is provided in Appendix A.

2.4 Generative Spectral Recovery (GSR)

The final stage completes the generative loop by recovering the original spectral distribution from the sparse, synchronized spikes o. A spiking readout layer Φ_{read} maps the spikes to a continuous representation: $Y = \Phi_{read}(o)$. A lightweight convolutional decoder Φ_{dec} maps the encoded representation back to the real component of the STFT space: $\hat{S}^{\text{re}} = \Phi_{dec}(\text{Concat}(Y, M_b, M_c))$. The GSR objective is defined by minimizing the reconstruction error within the masked valid regions:

$$\mathcal{L}_{GSR} = \frac{\left\| M_b \odot M_c \odot \left(\hat{S}^{\text{re}} - S^{\text{re}}\right) \right\|_1}{\left\| M_b \odot M_c \right\|_1 + \varepsilon}. \tag{7}$$

where ε is a small constant to prevent division by zero. By forcing the model to "recover" the spectrum through the lens of neural coherence, *Vibe-Spike* learns a robust, biologically-informed foundation representation of EEG signals.

3 Experiments

3.1 Pretraining

Pretraining dataset. To build a robust foundation for neural representation, we perform large-scale self-supervised pre-training on the open-access subset of the **REVE corpus** [Ouahidi et al., 2026]. This subset encompasses approximately **4.8 TB** of raw EEG recordings, representing one of the largest publicly available EEG aggregates. The corpus is characterized by its high *heterogeneity*, integrating data from **75** public sources with diverse electrode montages, sampling rates, and clinical/cognitive paradigms. The full list of constituent datasets and their original citations is provided in Appendix B.

We follow the REVE [Ouahidi et al., 2026] curation pipeline: recordings shorter than 10s or those overlapping with our downstream evaluation sets are excluded to prevent data leakage. The released REVE data are already sampled at 200 Hz, band-pass filtered (0.5–99.5 Hz), normalized using session-level Z-score statistics. Values exceeding 15 standard deviations are clipped.

Implementation details. *Vibe-Spike* is trained via `Accelerate` on 8× NVIDIA RTX 5090 GPUs (∼24h convergence). *Optimization:* AdamW ($lr = 3.0 \times 10^{-4}$, $\beta = (0.9, 0.99)$, weight decay 0.01) with a 5k-step warmup-stable-decay schedule (min $\eta = 10^{-5}$). *Objective:* Self-supervised reconstruction with a 0.25 spatio-temporal block masking ratio. *Input:* The released 200 Hz signals are downsampled to 100 Hz and processed via STFT as described in Sec. 2.2. To further bridge the

gap between inconsistent channel layouts (e.g., 10-20 system vs. high-density caps), we apply the shell-projected montage resampling (see Sec. 2.1). This maps sparse source signals onto a unified 3D manifold, ensuring the foundation model learns a geometry-invariant representation. Additional implementation details are provided in Appendix C.

Table 1: Overview of downstream EEG tasks and datasets.

Task	Dataset	Rate	# Ch.	Duration	# Samples	Label
Sleep staging	ISRUC	200 Hz	6	30 s	89,240	5-class
Sleep staging	HMC	256 Hz	4	30 s	137,243	5-class
Mental disorder diagnosis	Mumtaz2016	256 Hz	19	5 s	7,143	2-class
Seizure detection	CHB-MIT	256 Hz	16	10 s	326,993	2-class
AD diagnosis	ADFTD-RS	500 Hz	19	30 s	1,586	2-class
Vigilance detection	SADT	500 Hz	32	5 s	27,192	2-class
Vigilance detection	SEED-VIG	200 Hz	17	8 s	20,355	2-class
Stress detection	MentalArithmetic	500 Hz	20	5 s	1,707	2-class
Event type classification	TUEV	250 Hz	16	5 s	112,491	6-class
Abnormal detection	TUAB	250 Hz	16	10 s	409,455	2-class

3.2 Downstream Tasks

Downstream tasks and datasets. To evaluate the proposed *Vibe-Spike* across diverse EEG analysis scenarios, we use multiple downstream datasets, including ISRUC [Khalighi et al., 2016] and HMC [Alvarez-Estevez and Rijsman, 2022] for sleep staging, Mumtaz2016 [Mumtaz, 2016] for mental disorder diagnosis, CHB-MIT [Guttag, 2010] for seizure detection, ADFTD-RS [Miltiadous et al., 2023] for Alzheimer's Disease (AD) diagnosis, SADT [Cao et al., 2019] and SEED-VIG [Zheng and Lu, 2017] for vigilance detection, MentalArithmetic [Zyma et al., 2019] for stress detection, and TUEV and TUAB [Obeid and Picone, 2016] for event type classification and abnormal EEG detection, respectively. The selected tasks and datasets are summarized in Table 1. For all downstream datasets, EEG signals are resampled to 100 Hz before being fed into the model. Details are in Appendix D.

For datasets included in prior EEG foundation model benchmarks, we follow the reported train/validation/test splits and preprocessing protocols [Wang et al., 2025, Jiang et al., 2024, Yang et al., 2023a, Ouahidi et al., 2026]. For newly added datasets, we adopt the same splitting principle and use each baseline model with its recommended task-specific configuration. All inputs are resampled to 100 Hz, whereas baseline models are evaluated at their prescribed sampling rates.

Baselines. We compare the proposed *Vibe-Spike* with both non-foundation and foundation model baselines. For non-foundation models, we include representative EEG decoding architectures, including EEGNet [Lawhern et al., 2018], EEGConformer [Song et al., 2023], SPaRCNet [Jing et al., 2023], ContraWR [Yang et al., 2023b], CNN-Transformer [Peh et al., 2022], FFCL [Li et al., 2022], and ST-Transformer [Song et al., 2021]. For foundation model baselines, we compare with BIOT [Yang et al., 2023a], LaBraM [Jiang et al., 2024], CBraMod [Wang et al., 2025], and REVE [Ouahidi et al., 2026]. These baselines cover both task-specific EEG decoding models and recent large-scale pre-trained EEG foundation models.

Metrics. Following CBraMod [Wang et al., 2025] and REVE [Ouahidi et al., 2026], we adopt balanced accuracy, AUC-PR, and AUROC for binary classification tasks, and balanced accuracy, Cohen's kappa, and weighted F1 for multi-class classification tasks.

4 Results and Discussion

We evaluate our *Vibe-Spike* from four perspectives: (1) downstream task performance, (2) structured component ablations, (3) mechanistic interpretability analysis, and (4) model robustness and pretraining dynamics. The following sections summarize the quantitative comparisons, qualitative analyses, and robustness studies.

4.1 Downstream Task Performance

Table 2 presents the balanced accuracy across ten diverse EEG classification tasks. The results highlight several key advantages of the proposed *Vibe-Spike*.

Table 2: Balanced accuracy ($\pm$ std) of different methods across downstream EEG classification tasks.

Method	**Params**	**TUAB**	**TUEV**	**ISRUC**	**HMC**	**Mumtaz2016**
EEGNet	0.003M	0.7642 ± 0.0036	0.3876 ± 0.0143	0.7154 ± 0.0121	0.6856 ± 0.0064	0.9232 ± 0.0104
EEGConformer	0.55M	0.7758 ± 0.0049	0.4074 ± 0.0164	0.7400 ± 0.0133	0.7210 ± 0.0102	0.9308 ± 0.0117
SPaRCNet	0.79M	0.7896 ± 0.0018	0.4161 ± 0.0262	0.7487 ± 0.0075	0.4242 ± 0.0541	0.9316 ± 0.0095
ContraWR	1.6M	0.7746 ± 0.0041	0.4384 ± 0.0349	0.7402 ± 0.0126	0.4242 ± 0.0541	0.9195 ± 0.0115
CNN-Transformer	3.2M	0.7777 ± 0.0022	0.4087 ± 0.0161	0.7363 ± 0.0087	0.6573 ± 0.0141	0.9305 ± 0.0068
FFCL	2.4M	0.7848 ± 0.0038	0.3979 ± 0.0104	0.7277 ± 0.0182	0.4427 ± 0.0702	0.9314 ± 0.0038
ST-Transformer	3.5M	0.7966 ± 0.0023	0.3984 ± 0.0228	0.7381 ± 0.0205	0.2559 ± 0.0141	0.9135 ± 0.0103
BIOT	3.2M	0.7959 ± 0.0057	0.5281 ± 0.0225	0.7527 ± 0.0121	0.6862 ± 0.0041	0.9358 ± 0.0052
LaBraM	5.8M	0.8140 ± 0.0019	0.6409 ± 0.0065	0.7633 ± 0.0102	0.7286 ± 0.0101	0.9409 ± 0.0079
CBraMod	4.0M	0.8289 ± 0.0022	0.6671 ± 0.0107	0.7865 ± 0.0110	0.7335 ± 0.0047	0.9560 ± 0.0056
REVE	69M	$\mathbf{0.8315}^{**} \pm 0.0014$	$\mathbf{0.6759} \pm 0.0229$	0.7819 ± 0.0078	0.7401 ± 0.0075	0.9644 ± 0.0097
Vibe-Spike	0.41M	0.8048 ± 0.0014	0.5385 ± 0.0148	$\mathbf{0.7977}^{*} \pm 0.0086$	$\mathbf{0.7848}^{***} \pm 0.0118$	$\mathbf{0.9694} \pm 0.0068$
Method	Params	**CHB-MIT**	**ADFTD-RS**	**SADT**	**SEED-VIG**	**MentalArith.**
EEGNet	0.003M	0.5658 ± 0.0106	0.6885 ± 0.0374	0.6890 ± 0.0036	0.6843 ± 0.0480	0.6770 ± 0.0116
EEGConformer	0.55M	0.5976 ± 0.0141	0.7250 ± 0.0186	0.7301 ± 0.0106	0.7262 ± 0.0476	0.6805 ± 0.0123
SPaRCNet	0.79M	0.5876 ± 0.0191	0.7280 ± 0.0167	0.7558 ± 0.0121	0.7265 ± 0.0391	0.6879 ± 0.0107
ContraWR	1.6M	0.6344 ± 0.0002	0.6885 ± 0.0155	0.7013 ± 0.0069	0.7223 ± 0.0265	0.6631 ± 0.0097
CNN-Transformer	3.2M	0.6389 ± 0.0067	0.7335 ± 0.0238	0.7561 ± 0.0032	0.6956 ± 0.0308	0.6779 ± 0.0268
FFCL	2.4M	0.6262 ± 0.0104	0.7127 ± 0.0208	0.7289 ± 0.0043	0.7177 ± 0.0434	0.6798 ± 0.0142
ST-Transformer	3.5M	0.5915 ± 0.0195	0.7144 ± 0.0145	0.7098 ± 0.0067	0.6721 ± 0.0348	0.6631 ± 0.0173
BIOT	3.2M	0.7068 ± 0.0457	0.7330 ± 0.0141	0.7322 ± 0.0106	0.7242 ± 0.0364	0.6875 ± 0.0186
LaBraM	5.8M	0.7075 ± 0.0358	0.6962 ± 0.0184	0.7036 ± 0.0057	0.6630 ± 0.0593	0.6909 ± 0.0125
CBraMod	4.0M	0.7398 ± 0.0284	0.7564 ± 0.0193	0.7494 ± 0.0055	0.6488 ± 0.0331	0.7256 ± 0.0132
REVE	69M	0.6919 ± 0.0193	0.7750 ± 0.0385	0.6869 ± 0.0196	0.6569 ± 0.1096	$\mathbf{0.7660}^{*} \pm 0.0355$
Vibe-Spike	0.41M	$\mathbf{0.8323}^{***} \pm 0.0278$	$\mathbf{0.8664}^{***} \pm 0.0270$	$\mathbf{0.7568} \pm 0.0020$	$\mathbf{0.7348} \pm 0.0106$	$\underline{0.7277 \pm 0.0315}$

Note. Bold indicates best balanced accuracy. Significance ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$) is computed via Welch's t-test. Statistics assume $N = 10$ runs unless prior baseline studies specify otherwise. See Appendix E for details.

Results. *Vibe-Spike* achieves state-of-the-art (SOTA) or competitive performance with only **0.41M** parameters. This is significantly more efficient than existing EEG foundation models; for instance, it outperforms or matches the performance of REVE (69M) and LaBraM (5.8M) while being approximately **168**$\times$ and **14**$\times$ smaller, respectively. This demonstrates that biologically-inspired inductive biases (rhythmic spiking) can be more effective than raw scaling for neural signal representation. Appendix G reports energy estimates and hardware measurements.

Remark. *Vibe-Spike* achieves the highest balanced accuracy on 7 out of 10 tasks, with particularly substantial gains in clinical scenarios: *Seizure & AD Diagnosis:* On CHB-MIT and ADFTD-RS, *Vibe-Spike* outperforms the second-best models by absolute margins of **9.25%** and **9.14%**, respectively ($p < 0.001$). *Sleep Staging:* It sets new benchmarks on ISRUC and HMC, outperforming large-scale models like CBraMod and REVE. These clinical tasks are inherently dependent on the precise detection of transient, rhythmic electrophysiological biomarkers. For instance, accurate sleep staging on ISRUC and HMC relies on decoding localized multi-scale oscillatory events, such as sleep spindles (11–16 Hz) and K-complexes [Berry et al., 2012]. Similarly, seizure detection (CHB-MIT) requires identifying sudden, high-amplitude epileptiform discharges (e.g., spikes and sharp waves) that abruptly disrupt background rhythms [Fisher et al., 2014], while AD diagnosis (ADFTD-RS) is strongly associated with the "slowing" of spontaneous EEG rhythms and a significant reduction in signal complexity and synchrony [Jeong, 2004].

The consistent superiority of *Vibe-Spike* validates its alignment with clinical EEG characteristics. Unlike standard neural networks that often over-smooth transient signals, our event-driven spiking backbone explicitly captures abrupt neuro-dynamic shifts. Combined with the GSR pre-training, *Vibe-Spike* effectively encodes critical neural oscillations, yielding a discriminative, biologically-informed representation for clinical diagnostics.

Robust generalization. Despite its compact size, *Vibe-Spike* exhibits robust generalization across heterogeneous paradigms, from mental disorder diagnosis (Mumtaz2016) to vigilance detection (SEED-VIG), consistently ranking as the top-tier performer. Such results indicate the versatility of our unified spatial alignment and generative pre-training framework in handling diverse electrode montages and sampling rates.

4.2 Structured Component Ablations

We examine the contributions of the main architectural components and GSR pretraining on five EEG datasets. Table 3 presents the ablation results.

Table 3: Balanced accuracy (mean ± std) of *Vibe-Spike* variants across five EEG datasets. All variants except the full model are trained from scratch; Δ denotes the change in average balanced accuracy (pp) relative to w/o GSR pretraining.

Model variant	CHB-MIT	ISRUC	HMC	SADT	Mumtaz2016	Avg. (Δ, pp)
Full model (GSR pretrained)	**0.8323**** ± 0.0278	**0.7977***** ± 0.0086	**0.7848***** ± 0.0118	**0.7568***** ± 0.0020	**0.9694***** ± 0.0068	**0.8282** (+4.03)
w/o GSR pretraining	0.7816 ± 0.0334	0.7734 ± 0.0061	0.7546 ± 0.0157	0.7049 ± 0.0086	0.9248 ± 0.0121	0.7879 (ref.)
w/o spatial alignment	0.7783 ± 0.0391	0.7740 ± 0.0058	0.7505 ± 0.0194	0.7027 ± 0.0108	0.9215 ± 0.0155	0.7854 (−0.25)
SA → position-based mapping	0.7345 ± 0.0362	0.7329 ± 0.0169	0.7138 ± 0.0218	0.6610 ± 0.0132	0.8686 ± 0.0177	0.7422 (−4.57)
w/o STFT tokenization	0.6814 ± 0.0362	0.2000 ± 0.0000	0.3795 ± 0.0198	0.5126 ± 0.0263	0.6008 ± 0.0187	0.4749 (−31.30)
STFT → temporal convolution	0.7629 ± 0.0349	0.7661 ± 0.0134	0.7374 ± 0.0171	0.6818 ± 0.0101	0.9052 ± 0.0135	0.7707 (−1.72)
STFT → temporal patches	0.7591 ± 0.0357	0.7697 ± 0.0132	0.7340 ± 0.0176	0.6860 ± 0.0098	0.9036 ± 0.0139	0.7705 (−1.74)
w/o Kuramoto coupling	0.7690 ± 0.0260	0.7059 ± 0.0138	0.6175 ± 0.0232	0.6857 ± 0.0140	0.8781 ± 0.0023	0.7312 (−5.66)
w/o rhythm gate	0.7704 ± 0.0341	0.7731 ± 0.0128	0.7519 ± 0.0163	0.7153 ± 0.0091	0.9147 ± 0.0128	0.7851 (−0.28)
w/o SNN	0.7496 ± 0.0366	0.7495 ± 0.0165	0.7196 ± 0.0219	0.6649 ± 0.0124	0.8923 ± 0.0158	0.7552 (−3.27)
Temporal stride $s_{\text{temp}} = 16$	0.7040 ± 0.0300	0.6420 ± 0.0180	0.4046 ± 0.0144	0.5395 ± 0.0193	0.6267 ± 0.0125	0.5834 (−20.45)
STFT + classifier	0.6250 ± 0.0260	0.2380 ± 0.0120	0.2009 ± 0.0000	0.5278 ± 0.0000	0.4985 ± 0.0310	0.4180 (−36.98)
Raw + classifier	0.5740 ± 0.0240	0.2110 ± 0.0080	0.2041 ± 0.0023	0.5129 ± 0.0000	0.5352 ± 0.0081	0.4074 (−38.04)

GSR pretraining. GSR pretraining increases average BA from 0.7879 to 0.8282 (+4.03 pp), with gains across all five datasets. The architecture and downstream protocol remain unchanged, isolating the contribution of reconstruction pretraining.

Unified spatial alignment. SA accommodates heterogeneous electrode layouts during pretraining. It maps channels to a canonical montage using electrode coordinates and masks missing channels, without learnable mixing. Removing SA reduces average BA by only 0.25 pp under dataset-specific training, where montages are already fixed. Position-based mapping causes a larger decrease of 4.57 pp. This loss may reflect cross-channel mixing that obscures localized spatial patterns.

Time-frequency tokenization and temporal resolution. STFT converts raw EEG into shorter token sequences while retaining spectral information. This reduces the sequential updates required by the oscillator and membrane dynamics. Removing STFT reduces average BA by 31.30 pp. Coarse temporal processing ($s_{\text{temp}} = 16$) also causes a substantial decrease of 20.45 pp. Temporal convolution and temporal patches yield smaller decreases of 1.72 and 1.74 pp, respectively. These alternatives shorten the sequence but may discard transients or reduce temporal resolution.

Coupled dynamics and rhythmic gating. Removing Kuramoto coupling or the SNN decreases average BA by 5.66 and 3.27 pp, respectively. The coupling ablation produces particularly large decreases on HMC (13.71 pp) and ISRUC (6.75 pp). In contrast, removing the rhythm gate decreases average BA by only 0.28 pp but increases the reported spike-rate statistic on every dataset. The mean of the five dataset-specific relative increases is 99.84%. These results support distinct roles for the components: coupling and spiking dynamics contribute to predictive performance, whereas rhythmic gating has a stronger effect on the level of spiking activity.

4.3 Mechanistic Interpretability of Rhythmic Spikes

To demystify the "black box" nature of *Vibe-Spike*, Fig. 3 visualizes the learned rhythmic dynamics on TUAB (abnormality detection) and CHB-MIT (seizure detection). More results are in Figs. 6 - 13.

Temporal evidence accumulation. The temporal firing-rate curves (Fig. 3, left) reveal that *Vibe-Spike* does not rely on static features but performs *dynamic evidence accumulation*. For seizure samples, we observe sharp, rhythmic peaks in the latent firing rates that precisely coincide with the onset of ictal discharges. These spiking activities are integrated into "evidence trajectories," where the model's confidence progressively diverges as it encounters more pathologically relevant oscillations, closely mirroring the neural basis of decision-making and the clinical process of sequential EEG scoring [Gold and Shadlen, 2007, Tatum, 2021].

Spatiotemporal activation on scalp manifolds. The topographic maps (Fig. 3, center) show that the model learns *physiologically grounded spatial priors*. **Seizure localization (CHB-MIT):** The activation is concentrated in the *temporal and frontal electrodes (e.g., T3, T4, F7, F8)*, which are clinically established as classic focal points for epileptiform discharges in focal epilepsies [Noachtar and Rémi, 2009]. The channel-time activation maps further show that the model selectively "attends" to rhythmic spikes within these specific leads while suppressing background noise in occipital regions. **Pathological shifts (TUAB):** In contrast to the sparse, low-amplitude spikes in normal EEG, abnormal samples trigger widespread, high-intensity rhythmic responses across the frontal

cortex. This suggests the model captures the "spectral slowing" (e.g., increased delta/theta power) and disrupted synchrony that characterize general brain dysfunction [Niedermeyer and da Silva, 2005].

Latent spike manifolds. The magnified latent spike maps reveal that *Vibe-Spike* encodes EEG into sparse, temporally organized "bursts." Different clinical classes activate distinct subsets of latent spiking units, forming a *discriminative neural code*. This indicates that our GSR pre-training forces the model to internalize a generative manifold of neural microstates [Michel and Koenig, 2018], where specific spike sequences serve as reliable biomarkers for downstream diagnostics.

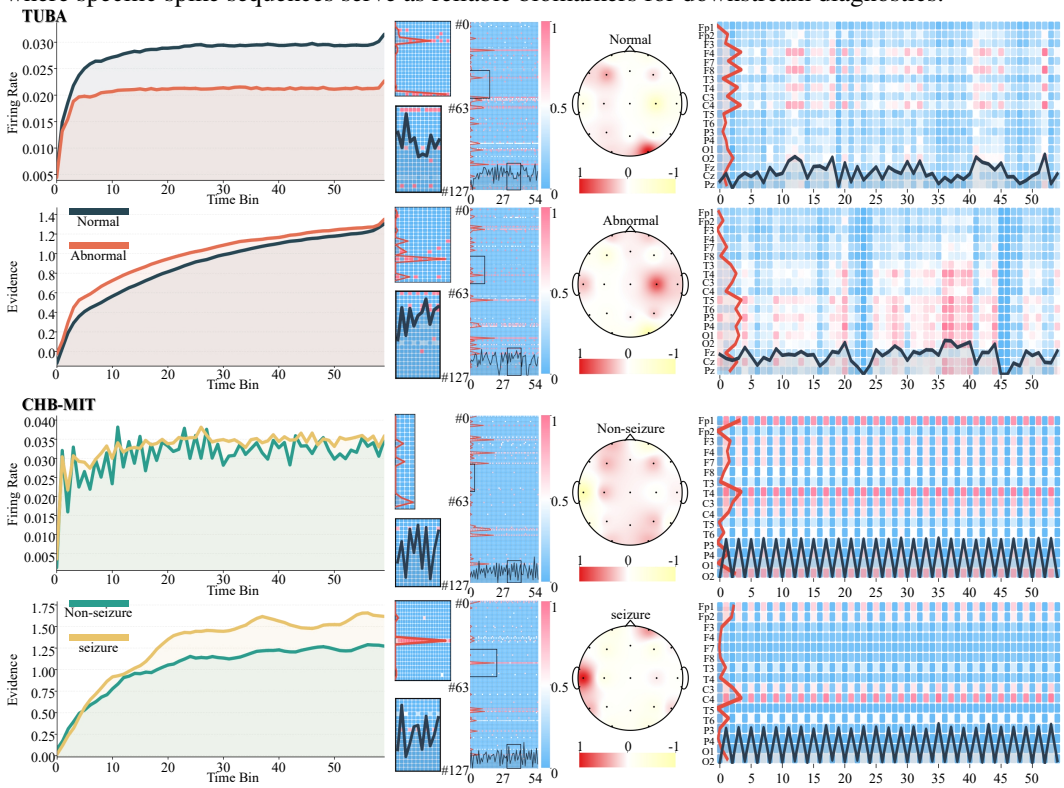

Figure 3: Spatiotemporal dynamics and mechanistic interpretability of *Vibe-Spike*. The upper and lower panels detail the TUAB (abnormality) and CHB-MIT (seizure) tasks, respectively. From left to right: (1) class-wise temporal firing and evidence accumulation trajectories; (2) representative latent spike manifolds with magnified temporal bursts; (3) scalp topographies of discriminative target-space activations; and (4) channel-time activation maps (with marginal curves denoting aggregate spatial/temporal profiles). The visualizations show that *Vibe-Spike* effectively localizes physiologically relevant clinical biomarkers. More results are shown in Appendix F.

4.4 Model Robustness and Pretraining Dynamics

To validate the structural stability and optimization efficiency of *Vibe-Spike*, we perform the tailored analyses on the SADT dataset (Figs. 4 and 5). More results are shown in Appendix H.

Architectural robustness. We evaluate the sensitivity of the NCE module by sweeping key hyperparameters: coupling strength λ, phase-delay τ, oscillator dimension D, and graph diagonal ρ_W (models are trained from scratch to isolate architectural effects). As shown in Fig. 4, performance metrics remain highly stable across all one-dimensional perturbations. Importantly, this robustness stems from the intrinsic properties of our coupled oscillatory dynamics: rather than relying on precise, dataset-specific scalar values, the networked oscillators naturally gravitate toward stable synchronization manifolds. This topological stability allows the architecture to absorb parameter variances without

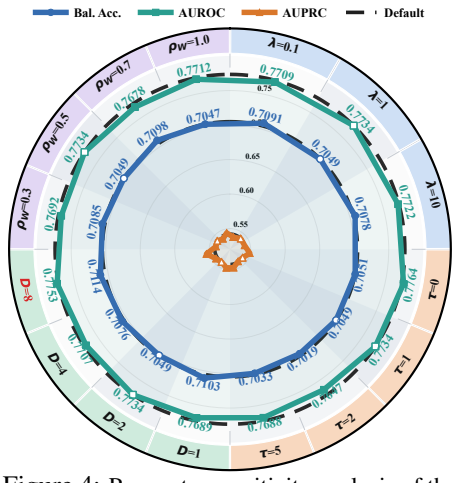

Figure 4: Parameter sensitivity analysis of the NCE module on the SADT dataset.

severe performance degradation, ensuring reliable deployment across diverse clinical EEG paradigms without brittle hyperparameter tuning. Details in Appendix H.

Pretraining and spiking dynamics. Fig. 5 contrasts the fine-tuning trajectories of pretrained versus scratch-trained models. Pretraining not only accelerates convergence and elevates final accuracy but fundamentally alters internal network dynamics. Deep SNNs notoriously suffer from the "dead neuron" (vanishing spike) problem during early optimization, which is clearly evidenced by the scratch baseline remaining virtually silent (near-zero spike rate) during early epochs. In contrast, our GSR module pretraining overcomes this by forcing the network to reconstruct continuous spectral features. This objective initializes the membrane potentials and synaptic weights into an actively firing, physiologically valid regime. Finally, the pretrained *Vibe-Spike* maintains robust spiking activity from the very first epoch, proving that GSR establishes a highly reusable synchronization prior for rapid downstream adaptation.

4.5 Discussion

The strength of *Vibe-Spike* mainly comes from its biologically inspired dynamical constraints, which are well aligned with the intrinsic properties of EEG signals. By modeling rhythmic dynamics and spiking responses, the model can efficiently capture temporal oscillations, transient neural activities, and spatially structured electrophysiological patterns. This design contributes to its decent performance across a wide range of clinical and cognitive EEG tasks.

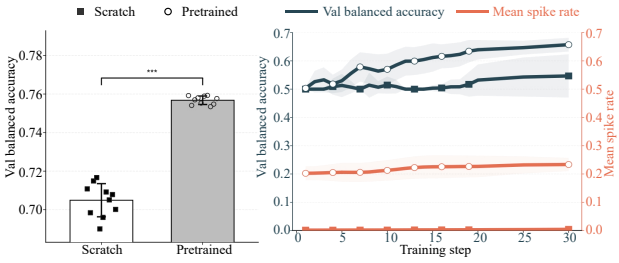

Figure 5: Pretraining improves convergence, final balanced accuracy, and spiking activity during fine-tuning.

Limitations. However, this inductive bias can also be a double-edged sword. Although rhythmic and spatial dynamical priors are beneficial for many EEG tasks, they may be less suitable for tasks whose discriminative information is not primarily expressed through rhythmic temporal patterns or physiologically meaningful spatial variations. For example, as shown in the Appendix J.1, the model is less effective on motor imagery tasks. One possible reason is that motor imagery classification often relies on learning task-specific spatial projections and transformed feature spaces, rather than directly exploiting explicit biological priors. In such cases, a strongly constrained dynamical backbone may limit the flexibility needed to discover optimal discriminative representations. Appendix J.2 presents a failure analysis.

Future work. Future work will explore more adaptive architectures that retain the advantages of rhythmic spiking dynamics while improving flexibility for tasks with different feature characteristics. One promising direction is to integrate learnable task-adaptive projection modules or hybrid representation pathways into the dynamical backbone, allowing the model to combine biologically informed priors with data-driven feature transformations. This may help overcome the current limitation and further improve the generalization of *Vibe-Spike* across broader EEG paradigms.

5 Conclusion

In this work, we embarked on *Vibe-Spike*, a biologically-inspired foundation model for EEG representation learning. By integrating a *macro-micro coupling mechanism* within a rhythmic spiking backbone and a *generative spectral recovery* pre-training objective, *Vibe-Spike* bridges the gap between large-scale foundation modeling and the intrinsic multi-scale dynamics of the human brain. This architecture effectively internalizes the interplay between macroscopic neural oscillations and microscopic event-driven spiking, enabling the model to capture complex neural coherence. Our experiments across ten diverse datasets demonstrate that *Vibe-Spike* achieves state-of-the-art or competitive performance in clinical diagnostics and cognitive monitoring. Remarkably, it achieves this with only *0.41M parameters*, representing a *168*$\times$ reduction in complexity compared to massive Transformers like REVE, while significantly outperforming them in tasks sensitive to transient neurodynamics, such as seizure detection and sleep staging. Mechanistic interpretability indicates that the learned spiking manifolds are *physiologically grounded*, accurately localizing clinically relevant biomarkers on the scalp manifold. Our work highlights that appropriate inductive biases are far more efficient and effective for neural signals than pure parameter scaling.

Acknowledgments and Disclosure of Funding

This work was partially supported by the National Institutes of Health (AG091653, AG068399, AG084375) and the Foundation of Hope.

Competing Interests

The authors declare no competing interests.

References

Juan A Acebrón, Luis L Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of modern physics*, 77(1):137–185, 2005.

Blanca Aguado-Lopez, Ana F. Palenciano, Jose M. G. Penalver, Paloma Diaz-Gutierrez, David Lopez-Garcia, Chiara Avancini, Luis F. Ciria, and Maria Ruz. Competition eeg, 2024. URL https://openneuro.org/datasets/ds005089/versions/1.0.1.

J. J. Allen, H. L. Urry, S. K. Hitt, and J. A. Coan. The stability of resting frontal electroencephalographic asymmetry in depression. *Psychophysiology*, 41(2):269–280, 2004. doi: 10.1111/j.1469-8986.2003.00149.x.

Diego Alvarez-Estevez and Roselyne Rijsman. Haaglanden Medisch Centrum sleep staging database. *PhysioNet*, March 2022. doi: 10.13026/t79q-fr32. URL https://doi.org/10.13026/t79q-fr32. Version 1.1.

American Academy of Sleep Medicine. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine, Darien, IL, version 3 edition, 2023. URL https://aasm.org/clinical-resources/scoring-manual/. Digital publication. Available to AASM active members at https://aasm.org/clinical-resources/scoring-manual/.

Edilberto Amorim, Wei-Long Zheng, Jong Woo Lee, Susan Herman, Mohammad Ghassemi, Adithya Sivaraju, Nicolas Gaspard, Jeannette Hofmeijer, Michel J A M van Putten, Matthew Reyna, Gari Clifford, and Brandon Westover. I-CARE: International Cardiac Arrest REsearch consortium Database. *PhysioNet*, June 2023. doi: 10.13026/avek-0p97. URL https://doi.org/10.13026/avek-0p97. Version 2.0.

Carlos Valle Araya, Carolina Mendez-Orellana, and Maria Rodriguez-Fernandez. Large spanish eeg, 2024. URL https://openneuro.org/datasets/ds004279/versions/1.1.2.

Imad J. Bajwa1, Andre S. Nilsen1, René Skukies1,3, Arnfinn Aamodt1, Gernot Ernst2, Johan F. Storm1, and Bjørn E. Juel1,2. A repeated awakening study exploring the capacity of complexity measures to capture dreaming during propofol sedation, 2025. URL https://openneuro.org/datasets/ds005620/versions/1.0.0.

Alexandre Barachant. *Commande robuste d'un effecteur par une interface cerveau machine EEG asynchrone*. PhD thesis, Université de Grenoble, 2012.

Cemre Baykan and Alexander C. Schütz. Eeg responses to the number of objects in partially occluded and uncovered scenes, 2024. URL https://openneuro.org/datasets/ds005586/versions/2.0.0.

Benjamin Lowe (Ben.Lowe@Mq.Edu.Au), Jonathan Robinson (Jonathan.Robinson@Monash.Edu), Naohide Yamamoto (Naohide.Yamamoto@Qut.Edu.Au), Hinze Hogendoorn (Hinze.Hogendoorn@Qut.Edu.Au), and Patrick Johnston (Dr.Pat.Johnston@Icloud.Com). Visual attribute-specific contextual trajectory paradigm, 2023. URL https://openneuro.org/datasets/ds004603/versions/1.1.0.

Richard B Berry, Rita Brooks, Charlene E Gamaldo, Susan M Harding, Carole Marcus, Bradley V Vaughn, et al. The aasm manual for the scoring of sleep and associated events. *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine*, 176(2012):7, 2012.

Ole Bialas, Emily Teoh, Andrew Anderson, and Edmund Lalor. Invariant encoding of phonemes in neural responses to continuous speech, 2023. URL https://openneuro.org/datasets/ds004408/versions/1.0.8.

Jiahang Cao, Ziqing Wang, Hanzhong Guo, Hao Cheng, Qiang Zhang, and Renjing Xu. Spiking denoising diffusion probabilistic models. In *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 4900–4909, 2024. doi: 10.1109/WACV57701.2024.00484.

Zehong Cao, Michael Chuang, J.T. King, and Chin-Teng Lin. Multi-channel EEG recordings during a sustained-attention driving task (raw dataset). 2 2019. doi: 10.6084/m9.figshare.6427334.v5. URL https://figshare.com/articles/dataset/Multi-channel_EEG_recordings_during_a_sustained-attention_driving_task/6427334.

James F Cavanagh and Trevor C J Jackson. Mood manipulation and pst, experiment 1, 2022. URL https://openneuro.org/datasets/ds004315/versions/1.0.0.

Luis Alberto Barradas Chacón and Selina C. Wriessnegger. Toonfaces, 2023. URL https://openneuro.org/datasets/ds004324/versions/1.0.0.

Hohyun Cho, Minkyu Ahn, Sangtae Ahn, Moonyoung Kwon, and Sung Chan Jun. Eeg datasets for motor imagery brain–computer interface. *GigaScience*, 6(7):gix034, 2017.

Chuqin Xiang, Xinrui Fan, Duo Bai, Ke Lv, and Xu Lei. A resting-state eeg dataset for sleep deprivation, 2024. URL https://openneuro.org/datasets/ds004902/versions/1.0.8.

Jose Cordoba-Silva, Rafael Maya, Mario Valderrama, Luis Felipe Giraldo, William Betancourt-Zapata, Andrés Salgado-Vascob, Juliana Marín-Sánchez, Viviana Gómez-Ortega, and Mark Ettenberger. Dataset of electrophysiological signals (eeg, ecg, emg) during music therapy with adult burn patients in the intensive care unit., 2023. URL https://openneuro.org/datasets/ds004840/versions/1.0.1.

Ian Daly, Nicoletta Nicolaou, Duncan Williams, Faustina Hwang, Alexis Kirke, Eduardo Miranda, and Slawomir J. Nasuto. An eeg dataset recorded during affective music listening, 2024. URL https://openneuro.org/datasets/ds002721/versions/1.0.3.

Tingting Dan and Guorong Wu. From cortical synchronous rhythm to brain inspired learning mechanism: An oscillatory spiking neural network with time-delayed coordination, 2026. URL https://arxiv.org/abs/2605.01656.

Arnaud Delorme. Go-nogo categorization and detection task, 2021. URL https://openneuro.org/datasets/ds002680/versions/1.2.0.

Arnaud Delorme and Claire Braboszcz. Meditation vs thinking task, 2021. URL https://openneuro.org/datasets/ds003969/versions/1.0.0.

Arnaud Delorme and Tracy Brandmeyer. Eeg meditation study, 2024. URL https://openneuro.org/datasets/ds001787/versions/1.1.1.

Pablo Rodríguez-San Esteban, Ana B. Chica, and José A. González-López. Neural representation of consciously seen and unseen information, 2024. URL https://openneuro.org/datasets/ds005273/versions/1.0.0.

Robert S Fisher, Carlos Acevedo, Alexis Arzimanoglou, Alicia Bogacz, J Helen Cross, Christian E Elger, Jerome Engel Jr, Lars Forsgren, Jacqueline A French, Mike Glynn, et al. Ilae official report: a practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482, 2014.

Lukas Gehrke, Sezen Akman, Albert Chen, Pedro Lopes, and Klaus Gramann. Prediction error, 2024. URL https://openneuro.org/datasets/ds003846/versions/2.0.2.

Joshua I Gold and Michael N Shadlen. The neural basis of decision making. *Annu. Rev. Neurosci.*, 30(1):535–574, 2007.

Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *circulation*, 101(23):e215–e220, 2000.

Tijl Grootswagers, Ivy Zhou, Amanda Robinson, Martin Hebart, and Thomas Carlson. Human electroencephalography recordings from 50 subjects for 22,248 images from 1,854 object concepts, 2022. URL https://openneuro.org/datasets/ds003825/versions/1.2.0.

Tijl Grootswagers, Amanda Robinson, Sofia Shatek, and Thomas Carlson. Eeg-attention-rsvp-exp1, 2023a. URL https://openneuro.org/datasets/ds004816/versions/1.0.1.

Tijl Grootswagers, Amanda Robinson, Sofia Shatek, and Thomas Carlson. Eeg-attention-rsvp-exp2, 2023b. URL https://openneuro.org/datasets/ds004817/versions/1.0.1.

Tijl Grootswagers, Amanda Robinson, Sofia Shatek, and Thomas Carlson. Features-eeg, 2024. URL https://openneuro.org/datasets/ds004357/versions/1.0.1.

John Guttag. CHB-MIT Scalp EEG Database. *PhysioNet*, June 2010. doi: 10.13026/C2K01R. URL https://doi.org/10.13026/C2K01R. Version 1.0.0.

C. Halkiopoulos, E. Gkintoni, and B. Boutsinas. Mapping the digital mind: A meta-analysis of EEG biomarkers in cognition, emotion, and mental health. *Brain Sciences*, 16(4):368, 3 2026. doi: 10.3390/brainsci16040368.

Cameron D. Hassall, Yan Yan, and Laurence T. Hunt. Drum trainer, 2022a. URL https://openneuro.org/datasets/ds004152/versions/1.1.2.

Cameron D. Hassall, Yan Yan, and Laurence T. Hunt. Continuous feedback processing, 2022b. URL https://openneuro.org/datasets/ds004262/versions/1.0.0.

Cameron D. Hassall, Yan Yan, and Laurence T. Hunt. Steer the ship, 2022c. URL https://openneuro.org/datasets/ds004264/versions/1.1.0.

Cameron D. Hassall, Laurence T. Hunt, and Clay B. Holroyd. Average task value, 2024. URL https://openneuro.org/datasets/ds004147/versions/1.0.2.

Christoffer Hatlestad-Hall, Trine Waage Rygvold, and Stein Andersson. Srm resting-state eeg, 2022. URL https://openneuro.org/datasets/ds003775/versions/1.2.1.

Marleen Haupt, Monika Graumann, Santani Teng, Carina Kaltenbach, and Radoslaw M. Cichy. Braille letters - eeg, 2024. URL https://openneuro.org/datasets/ds004951/versions/1.0.0.

Jason Helbing, Dejan Draschkow, and Melissa L.-H. Võ. Search superiority recollection familiarity, 2024. URL https://openneuro.org/datasets/ds005189/versions/1.0.1.

J. B. Henriques and R. J. Davidson. Left frontal hypoactivation in depression. *Journal of Abnormal Psychology*, 100(4):535–545, 1991. doi: 10.1037/0021-843x.100.4.535.

L. J. Hirsch, S. M. LaRoche, N. Gaspard, E. Gerard, A. Svoronos, S. T. Herman, and J. Claassen. American clinical neurophysiology society's standardized critical care EEG terminology: 2012 version. *Journal of Clinical Neurophysiology*, 30(1):1–27, 2013. doi: 10.1097/WNP.0b013e3182784729.

Mark Horowitz. Computing's energy problem (and what we can do about it). In *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pages 10–14, 2014. doi: 10.1109/ISSCC.2014.6757323.

Jaeseung Jeong. Eeg dynamics in patients with alzheimer's disease. *Clinical neurophysiology*, 115 (7):1490–1505, 2004.

Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous EEG data in BCI. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=QzTpTRVtrP.

Jin Jing, Wendong Ge, Shenda Hong, Marta Bento Fernandes, Zhen Lin, Chaoqi Yang, Sungtae An, Aaron F. Struck, Aline Herlopian, Ioannis Karakis, Jonathan J. Halford, Marcus C. Ng, Emily L. Johnson, Brian L. Appavu, Rani A. Sarkis, Gamaleldin Osman, Peter W. Kaplan, Monica B. Dhakar, Lakshman Arcot Jayagopal, Zubeda Sheikh, Olga Taraschenko, Sarah Schmitt, Hiba A. Haider, Jennifer A. Kim, Christa B. Swisher, Nicolas Gaspard, Mackenzie C. Cervenka, Andres A. Rodriguez Ruiz, Jong Woo Lee, Mohammad Tabaeizadeh, Emily J. Gilmore, Kristy Nordstrom, Ji Yeoun Yoo, Manisha G. Holmes, Susan T. Herman, Jennifer A. Williams, Jay Pathmanathan, Fábio A. Nascimento, Ziwei Fan, Samaneh Nasiri, Mouhsin M. Shafi, Sydney S. Cash, Daniel B. Hoch, Andrew J. Cole, Eric S. Rosenthal, Sahar F. Zafar, Jimeng Sun, and M. Brandon Westover. Development of expert-level classification of seizures and rhythmic and periodic patterns during eeg interpretation. *Neurology*, 100(17):e1750–e1762, 2023. doi: 10.1212/WNL.0000000000207127.

Michael J. Kahana and Joseph H. Rudoler. Penn electrophysiology of encoding and retrieval study (peers), 2023. URL https://openneuro.org/datasets/ds004395/versions/2.0.0.

Emmanuel K Kalunga, Sylvain Chevallier, and Quentin Barthélemy. Using riemannian geometry for ssvep-based brain computer interface. *arXiv preprint arXiv:1501.03227*, 2015.

Sirvan Khalighi, Teresa Sousa, José Santos, and Urbano Nunes. Isruc-sleep: A comprehensive public dataset for sleep researchers. *Computer Methods and Programs in Biomedicine*, 124, 11 2016. doi: 10.1016/j.cmpb.2015.10.013.

Louis Korczowski, Martine Cederhout, Anton Andreev, Grégoire Cattan, Pedro Luiz Coelho Rodrigues, Violette Gautheret, and Marco Congedo. *Brain Invaders calibration-less P300-based BCI with modulation of flash duration Dataset (bi2015a)*. PhD thesis, GIPSA-lab, 2019a.

Louis Korczowski, Martine Cederhout, Anton Andreev, Grégoire Cattan, Pedro Luiz Coelho Rodrigues, Violette Gautheret, and Marco Congedo. *Brain invaders cooperative versus competitive: multi-user P300-based brain-computer interface dataset (BI2015b)*. PhD thesis, GIPSA-lab, 2019b.

Louis Korczowski, Ekaterina Ostaschenko, Anton Andreev, Grégoire Cattan, Pedro Luiz Coelho Rodrigues, Violette Gautheret, and Marco Congedo. *Brain Invaders calibration-less P300-based BCI using dry EEG electrodes Dataset (bi2014a)*. PhD thesis, GIPSA-lab, 2019c.

Louis Korczowski, Ekaterina Ostaschenko, Anton Andreev, Grégoire Cattan, Pedro Luiz Coelho Rodrigues, Violette Gautheret, and Marco Congedo. *Brain invaders solo versus collaboration: Multi-user P300-based brain-computer interface dataset (bi2014b)*. PhD thesis, GIPSA-lab, 2019d.

A. J. Krause, E. B. Simon, B. A. Mander, S. M. Greer, J. M. Saletin, A. N. Goldstein-Piekarski, and M. P. Walker. The sleep-deprived human brain. *Nature Reviews Neuroscience*, 18(7):404–418, 2017. doi: 10.1038/nrn.2017.55.

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, jul 2018. doi: 10.1088/1741-2552/aace8c. URL https://doi.org/10.1088/1741-2552/aace8c.

Min-Ho Lee, O-Yeon Kwon, Yong-Jeong Kim, Hong-Kyung Kim, Young-Eun Lee, John Williamson, Siamac Fazli, and Seong-Whan Lee. Eeg dataset and openbmi toolbox for three bci paradigms: An investigation into bci illiteracy. *GigaScience*, 8(5):giz002, 2019a.

Min-Ho Lee, O-Yeon Kwon, Yong-Jeong Kim, Hong-Kyung Kim, Young-Eun Lee, John Williamson, Siamac Fazli, and Seong-Whan Lee. Eeg dataset and openbmi toolbox for three bci paradigms: An investigation into bci illiteracy. *GigaScience*, 8(5):giz002, 2019b.

Min-Ho Lee, O-Yeon Kwon, Yong-Jeong Kim, Hong-Kyung Kim, Young-Eun Lee, John Williamson, Siamac Fazli, and Seong-Whan Lee. Eeg dataset and openbmi toolbox for three bci paradigms: An investigation into bci illiteracy. *GigaScience*, 8(5):giz002, 2019c.

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. ISSN 1746-8094. doi: https://doi.org/10.1016/j.bspc.2021.103342. URL https://www.sciencedirect.com/science/article/pii/S1746809421009393.

Haijie Liu, Penghu Wei, Haochong Wang, Xiaodong Lv, Wei Duan, Meijie Li, Yan Zhao, Qingmei Wang, Xinyuan Chen, Gaige Shi, et al. An eeg motor imagery dataset for brain computer interface in acute stroke patients. *Scientific Data*, 11(1):131, 2024.

Y. Luo, M. Tang, and X. Fan. Meta analysis of resting frontal alpha asymmetry as a biomarker of depression. *npj Mental Health Research*, 4(1):2, 2025. doi: 10.1038/s44184-025-00117-x.

Dominique Makowski, An-Shu Te, Stephanie Kirk, and Zi Liang Ngoi. Fakefaceemo data, 2023. URL https://openneuro.org/datasets/ds004582/versions/1.0.0.

Donia Metwalli, Eslam Ahmed, Antony Emil, Yousef A. Radwan, Mariam Barakat, Anas Ahmed, Amro Omar, and Sahar Selim. Areeg: Arabic inner speech eeg dataset, 2025. URL https://openneuro.org/datasets/ds005262/versions/1.0.1.

Christoph M Michel and Thomas Koenig. Eeg microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage*, 180:577–593, 2018.

Andreas Miltiadous, Katerina D. Tzimourta, Theodora Afrantou, Panagiotis Ioannidis, Nikolaos Grigoriadis, Dimitrios G. Tsalikakis, Pantelis Angelidis, Markos G. Tsipouras, Euripidis Glavas, Nikolaos Giannakeas, and Alexandros T. Tzallas. A dataset of scalp eeg recordings of alzheimer's disease, frontotemporal dementia and healthy subjects from routine eeg. *Data*, 8(6), 2023. ISSN 2306-5729. doi: 10.3390/data8060095. URL https://www.mdpi.com/2306-5729/8/6/95.

Takeru Miyato, Sindy Löwe, Andreas Geiger, and Max Welling. Artificial kuramoto oscillatory neurons. *arXiv preprint arXiv:2410.13821*, 2024.

Denise Moerel, Tijl Grootswagers, Amanda Robinson, Sophia Shatek, Alexandra Woolgar, Thomas Carlson, and Anina Rich. The time-course of feature-based attention effects dissociated from temporal expectation and target-related processes, 2022. URL https://openneuro.org/datasets/ds004043/versions/1.1.0.

D. V. Moretti, D. Paternicò, G. Binetti, O. Zanetti, and G. B. Frisoni. Eeg markers are associated to gray matter changes in thalamus and basal ganglia in subjects with mild cognitive impairment. *NeuroImage*, 60(1):489–496, 2012. doi: 10.1016/j.neuroimage.2011.11.086.

F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz. Seizure prediction: the long and winding road. *Brain*, 130(Pt 2):314–333, 2 2007. doi: 10.1093/brain/awl241.

Wajid Mumtaz. MDD Patients and Healthy Controls EEG Data (New). 11 2016. doi: 10.6084/m9.figshare.4244171.v2. URL https://figshare.com/articles/dataset/EEG_Data_New/4244171.

Ernst Niedermeyer and FH Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

Soheyl Noachtar and Jan Rémi. The role of eeg in epilepsy: a critical review. *Epilepsy & Behavior*, 15(1):22–33, 2009.

Iyad Obeid and Joseph Picone. The temple university hospital eeg data corpus. *Frontiers in Neuroscience*, Volume 10 - 2016, 2016. ISSN 1662-453X. doi: 10.3389/fnins.2016.00196. URL https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2016.00196.

Patrick Ofner, Andreas Schwarz, Joana Pereira, and Gernot R Müller-Putz. Upper limb movements can be decoded from the time-domain of low-frequency eeg. *PloS one*, 12(8):e0182578, 2017.

Julie Onton and Scott Makeig. Imagined emotion study, 2022. URL https://openneuro.org/datasets/ds003004/versions/1.1.1.

Yassine El Ouahidi, Jonathan Lys, Philipp Thölke, Nicolas Farrugia, Bastien Pasdeloup, Vincent Gripon, Karim Jerbi, and Giulia Lioi. REVE: A foundation model for EEG - adapting to any setup with large-scale pretraining on 25,000 subjects. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2026.

{Wei Yan} Peh, Yuanyuan Yao, and Justin Dauwels. Transformer convolutional neural networks for automated artifact detection in scalp eeg. In *Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 3599–3602, United States, 2022. IEEE. ISBN 978-1-7281-2783-5. doi: 10.1109/EMBC48229.2022.9871916. Green Open Access added to TU Delft Institutional Repository 'You share, we take care!' - Taverne project https://www.openaccess.nl/en/you-share-we-take-care Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.; 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) ; Conference date: 11-07-2022 Through 15-07-2022.

Maria J. Ribeiro and Miguel Castelo-Branco. Eeg, ecg and pupil data from young and older adults: rest and auditory cued reaction time tasks, 2021. URL `https://openneuro.org/datasets/ds003690/versions/1.0.0`.

Alexander P. Rockhill, Nicko Jackson, Jobi George, Adam Aron, and Nicole C. Swann. Uc san diego resting state eeg data from patients with parkinson's disease, 2022. URL `https://openneuro.org/datasets/ds002778/versions/1.0.5`.

Joseph H. Rudoler, Matthew R. Dougherty, Brandon S. Katerman, James P. Bruska, Woohyeuk Chang, David J. Halpern, Nicholas B. Diamond, and Michael J. Kahana. Spatial memory and non-invasive closed-loop stimulus timing, 2023. URL `https://openneuro.org/datasets/ds004706/versions/1.0.0`.

Gerwin Schalk, Dennis J McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R Wolpaw. Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Transactions on biomedical engineering*, 51(6):1034–1043, 2004.

Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, 38(11):5391–5420, 2017.

Tong Shan, Madeline S. Cappelloni, and Ross K. Maddox. Music and speech elicit similar subcortical responses in human listeners, 2024. URL `https://openneuro.org/datasets/ds004356/versions/2.2.1`.

Sophia M. Shatek, Amanda K. Robinson, Tijl Grootswagers, and Thomas A. Carlson. Capacity for movement is an organisational principle in object representations: Eeg data from experiment 2, 2022. URL `https://openneuro.org/datasets/ds003887/versions/1.2.3`.

Sophia M. Shatek, Amanda K. Robinson, Tijl Grootswagers, and Thomas A. Carlson. Capacity for movement is an organisational principle in object representations: Eeg data from experiment 1, 2023. URL `https://openneuro.org/datasets/ds003885/versions/1.0.8`.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 1, 2024a. URL `https://openneuro.org/datasets/ds005505/versions/1.0.1`.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 4, 2024b. URL `https://openneuro.org/datasets/ds005508/versions/1.0.1`.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 6, 2024c. URL `https://openneuro.org/datasets/ds005510/versions/1.0.1`.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 2, 2025a. URL https://openneuro.org/datasets/ds005506/versions/1.0.1.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 3, 2025b. URL https://openneuro.org/datasets/ds005507/versions/1.0.1.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 5, 2025c. URL https://openneuro.org/datasets/ds005509/versions/1.0.1.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 7, 2025d. URL https://openneuro.org/datasets/ds005511/versions/1.0.1.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 8, 2025e. URL https://openneuro.org/datasets/ds005512/versions/1.0.1.

Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. Healthy brain network (hbn) eeg - release 9, 2025f. URL https://openneuro.org/datasets/ds005514/versions/1.0.1.

Elizabeth M. Siefert, Sindhuja Uppuluri, Jianing Mu, Marlie C. Tandoc, James W. Antony, and Anna C. Schapiro. Siefert2024, 2024. URL https://openneuro.org/datasets/ds005121/versions/1.0.2.

Yonghao Song, Xueyu Jia, Lie Yang, and Longhan Xie. Transformer-based Spatial-Temporal Feature Learning for EEG Decoding. *arXiv e-prints*, art. arXiv:2106.11170, June 2021. doi: 10.48550/arXiv.2106.11170.

Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2023. doi: 10.1109/TNSRE.2022.3230250.

Jan Sosulski and Michael Tangermann. Spatial filters for auditory evoked potentials transfer between different experimental conditions. In *GBCIC*, 2019.

K. Sugimoto, H. Kurashiki, Y. Xu, M. Takemi, and K. Amano. Electroencephalographic biomarkers of relaxation: A systematic review and meta-analysis. *Journal of Cognitive Neuroscience*, 38(3): 598–628, 3 2026.

Tasos Papastylianou, Rodrigo Ramele, Luca Citi, Caterina Cinel, and Riccardo Poli. Pes - pandemic emergency scenario, 2023. URL https://openneuro.org/datasets/ds004477/versions/1.0.2.

William O Tatum. *Handbook of EEG interpretation*. Springer Publishing Company, 2021.

Jack E. Taylor, Rasmus Sinn, Cosimo Iaia, and Christian J. Fiebach. Alphabetic decision task, 2024. URL https://openneuro.org/datasets/ds005594/versions/1.0.3.

M. Thomas, H. Sing, G. Belenky, H. Holcomb, H. Mayberg, R. Dannals, H. Wagner, D. Thorne, K. Popp, L. Rowland, A. Welsh, S. Balwinski, and D. Redmond. Neural basis of alertness and cognitive performance impairments during sleepiness. i. effects of 24 h of sleep deprivation on waking human regional brain activity. *Journal of Sleep Research*, 9(4):335–352, 2000. doi: 10.1046/j.1365-2869.2000.00225.x.

J. A. Urigüen and B. Garcia-Zapirain. Eeg artifact removal-state-of-the-art and guidelines. *Journal of Neural Engineering*, 12(3):031001, 6 2015. doi: 10.1088/1741-2560/12/3/031001.

J. Veillette, S. Heald, B. Wittenbrink, and H. Nusbaum. eeg-neuroforecasting, 2022. URL https://openneuro.org/datasets/ds004284/versions/1.0.0.

John Veillette, Pedro Lopes, and Howard Nusbaum. Illusion of agency over electrically-actuated movements, 2023. URL https://openneuro.org/datasets/ds004561/versions/1.0.0.

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 *The Thirteenth International Conference on Learning Representations*, 2025.

Weilong Li and Jiaxin Zhao. Perceiveimagine, 2024. URL https://openneuro.org/datasets/ds005697/versions/1.0.2.

M. B. Westover, M. M. Shafi, M. T. Bianchi, L. M. V. R. Moura, D. M. O'Rourke, E. S. Rosenthal, C. J. Chu, S. Donovan, D. B. Hoch, R. D. Kilbride, A. J. Cole, and S. S. Cash. The probability of seizures during EEG monitoring in critically ill adults. *Clinical Neurophysiology*, 126(3):463–471, 3 2015. doi: 10.1016/j.clinph.2014.05.037.

Yinsong Yan, Qu Yang, Yujie Wu, Hanwen Liu, Malu Zhang, Haizhou Li, Kay Chen Tan, and Jibin Wu. Efficient and robust temporal processing with neural oscillations modulated spiking neural networks. *Nature communications*, 16(1):8651, 2025.

Chaoqi Yang, M Brandon Westover, and Jimeng Sun. Biot: Biosignal transformer for cross-data learning in the wild. In *Thirty-seventh Conference on 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, 2023b. ISSN 2817-1705. doi: https://doi.org/10.2196/46769. URL https://www.sciencedirect.com/science/article/pii/S2817170523000273.

Weibo Yi, Shuang Qiu, Kun Wang, Hongzhi Qi, Lixin Zhang, Peng Zhou, Feng He, and Dong Ming. Evaluation of eeg oscillatory patterns and cognitive process during simple and compound limb motor imagery. *PloS one*, 9(12):e114853, 2014.

Yulin Wang, Wei Duan, Lihong Ding, Debo Dong, and Xu Lei. A test-retest resting and cognitive state eeg dataset, 2022. URL https://openneuro.org/datasets/ds004148/versions/1.0.1.

Wei-Long Zheng and Bao-Liang Lu. A multimodal approach to estimating vigilance using eeg and forehead eog. *Journal of Neural Engineering*, 14(2):026017, 2017. URL http://stacks.iop.org/1741-2552/14/i=2/a=026017.

Bangyan Zhou, Xiaopei Wu, Zhao Lv, Lei Zhang, and Xiaojin Guo. A fully automated trial selection method for optimization of motor imagery based brain-computer interface. *PloS one*, 11(9): e0162657, 2016.

Natalia Zhozhikashvili, Maria Protopova, Tatiana Shkurenko, Marie Arsalidou, Ilya Zakharov, Boris Kotchoubey, Sergey Malykh, and Yuri Pavlov. Sternberg, 2024. URL https://openneuro.org/datasets/ds005095/versions/1.0.1.

Zoltan Kekecs, Kyra Girán, Vanda Vizkievicz, Anna Lutoskin, and Yeganeh Farahzadi. The effects of sham hypnosis techniques, 2026. URL https://openneuro.org/datasets/ds004572/versions/1.3.2.

Igor Zyma, Sergii Tukaev, Ivan Seleznov, Ken Kiyono, Anton Popov, Mariia Chernykh, and Oleksii Shpenkov. Electroencephalograms during mental arithmetic task performance. *Data*, 4(1):14, 2019.

A Further Explanation of Neural Coherence Encoding

This section provides a more detailed explanation of the Neural Coherence Encoding (NCE) module. Table 4 summarizes the main quantities in NCE, including their type and functional role.

Table 4: Summary of parameters and variables in the NCE module.

Symbol	Type	Role	Description
Φ_{stem}	Learnable	Latent projection	Projects H_{in} into compact latent oscillator space Z.
Z	Dynamic representation	Latent sequence	Input-dependent latent representation used for graph construction and oscillator driving.
W	Dynamically computed	Structural prior	Functional connectivity graph computed from temporal correlations among latent dimensions.
ρ_W	Hyperparameter	Graph diagonal control	Controls the residual self-connection strength in $W = (1-\rho_W)\widetilde{W} + \rho_W I_d$.
d	Hyperparameter	Latent oscillator dimension	Number of latent oscillators / latent dimensions in NCE.
D	Hyperparameter	Internal oscillator dimension	Number of internal phase dimensions for each latent oscillator, yielding phase states in $\mathbb{R}^{d\times D}$.
r_i	Learnable	Intrinsic rhythm	Controls the natural phase evolution of the i-th oscillator.
λ	Hyperparameter	Coupling strength	Controls the global strength of graph-aware oscillator synchronization.
ψ_{ij}	Dynamic / model-internal	Phase lag	Introduces partial synchronization between oscillators and avoids trivial phase collapse.
η_i	Learnable	Sensory-drive strength	Controls how strongly the input-derived drive attracts the oscillator phase.
$\zeta_{i,t}$	Dynamic representation	Input drive	Input-derived phase target for the sensory-drive term.
τ	Hyperparameter	Phase-delay step	Determines how far the phase state is delayed when generating the rhythm signal.
g_t	Dynamically computed	Macroscopic rhythm signal	Phase-derived rhythm signal computed from delayed oscillator phases, used as the top-down rhythmic modulation source.
$\widetilde{g}_t$	Dynamically computed / learnable projection	Spiking rhythm gate	Projected gate derived from g_t that directly modulates membrane update and spike emission in the rhythm-gated LIF dynamics.
u_t	Dynamic representation	Synaptic drive	Converts phase and rhythm information into input current for the spiking dynamics.
κ	Learnable or fixed by design	Membrane decay	Controls temporal leakage of the membrane potential.
ϑ	Hyperparameter	Firing threshold	Determines when the membrane potential produces a spike.
v_t	State variable	Membrane potential	Accumulates synaptic drive over time under rhythm-gated LIF dynamics.
o_t	State variable	Spike output	Sparse binary spike train generated by the rhythm-gated spiking neuron.
$\mathcal{H}(\cdot)$	Non-learnable operator	Spike generation	Heaviside function used for binary spike emission in the forward pass.

A.1 Latent Projection

Given the input representation H_{in}, the latent sequence Z is mapped by:

$$Z = \Phi_{\text{stem}}(H_{\text{in}}) = \text{PWConv}\left(\underset{k\in\{3,5,9\}}{\text{Concat}}\ \text{DWConv}_k\left(\text{PWConv}(H_{\text{in}})\right)\right) \in \mathbb{R}^{d\times T'} \tag{8}$$

Here, PWConv denotes a pointwise 1×1 convolution, and DWConv_k denotes a depthwise temporal convolution with kernel size k. The stem reduces the feature dimension to the latent oscillator dimension d while preserving the temporal length T'.

A.2 Coupled Oscillator Dynamics

Given the latent sequence Z and the structural prior W (calculated by Eq. (4)), the phase dynamics consist of three terms in Eq. 5:

$$\underbrace{r_i}_{\text{intrinsic rhythm}} \underbrace{+\lambda \sum_{j=1}^{d} W_{ij} \sin(\varphi_{j,t} - \varphi_{i,t} - \psi_{ij})}_{\text{graph-coupled synchronization}} + \underbrace{\eta_i \odot \sin(\zeta_{i,t} - \varphi_{i,t})}_{\text{input-driven phase attraction}}.$$

$\dot{\varphi}_{i,t}$: Phase velocity of the i-th oscillator; dynamic quantity determined by the oscillator equation.

$\varphi_{i,t}$: Phase state of the i-th latent oscillator at time t; dynamic state variable. The initial phase is initialized as a fixed state (zero default).

r_i: Intrinsic rhythm of the i-th oscillator; learnable parameter.

λ: Global coupling strength among oscillators; hyperparameter.

W_{ij}: Structural connection strength from oscillator j to oscillator i.

d: Number of latent oscillators, corresponding to the latent dimension of Z; hyperparameter.

D: The dimension of the phase states of the oscillators; hyperparameter.

ψ_{ij}: Phase lag between oscillator i and oscillator j; model-internal variable used to induce partial synchronization.

η_i: Sensory-drive strength of the i-th oscillator; learnable parameter.

$\zeta_{i,t}$: Input-derived sensory drive for the i-th oscillator at time t. The sensory drive is computed as

$$\zeta_t = f_{\mathrm{enc}}\left(f_{\mathrm{proj}}(Z_{:,t})\right) \in \mathbb{R}^{d \times D}, \qquad \zeta_{i,t} = [\zeta_t]_i. \tag{9}$$

f_{proj}: Learnable projection block. In the implementation, it consists of a linear layer, layer normalization, LeakyReLU activation, and dropout, mapping the instantaneous latent feature into a hidden drive representation.

f_{enc}: Learnable encoding block. In the implementation, it first uses a linear layer to map the hidden drive representation back to the d-dimensional oscillator space, and then broadcasts the oscillator-wise drive along the internal oscillator dimension D to match the vector phase state.

Equation Explanation: The first term describes the intrinsic phase evolution of each oscillator. The second term promotes synchronization among connected latent oscillators according to the structural prior W, while allowing phase offsets through ψ_{ij}. The third term anchors the oscillator phase to the current EEG-derived latent input, preventing the phase dynamics from evolving independently of the observed signal.

A.3 Rhythm Signal

Given the synchronized phase state φ_t, the macroscopic rhythm signal is computed as

$$g_t = \frac{1}{2}\left(1 + \sin(\bar{\varphi}_{t-\tau})\right), \qquad g_t \in [0,1]^d. \tag{10}$$

g_t: Macroscopic rhythm signal computed from the delayed oscillator phase; dynamically computed variable. It serves as the top-down rhythmic modulation source for the subsequent spiking dynamics.

$\bar{\varphi}_{t-\tau}$: Vector of averaged delayed phases for all latent oscillators, calculated by averaging the internal D-dimensional phase state of each oscillator:

$$\bar{\varphi}_{i,t-\tau} = \frac{1}{D} \sum_{q=1}^{D} \varphi_{i,t-\tau,q}, \qquad \bar{\varphi}_{t-\tau} \in \mathbb{R}^d. \tag{11}$$

τ: Phase-delay step used to generate the rhythm signal; hyperparameter.

A.4 Rhythm-Gated Spiking Dynamics

Given the phase state φ_t and the macroscopic rhythm signal g_t, the rhythm-gated spiking dynamics are defined as (Eq. 6):

$$\begin{cases} h_t = \kappa v_{t-1} + (1-\kappa)u_t - \vartheta o_{t-1}, \\ v_t = \widetilde{g}_t \odot h_t + (1-\widetilde{g}_t) \odot v_{t-1}, \\ o_t = \widetilde{g}_t \odot \mathcal{H}(v_t - \vartheta). \end{cases}$$

u_t: Synaptic drive at time t; dynamically computed from the oscillator phase and the macroscopic rhythm signal:

$$u_t = f_u\left(\sin(\varphi_t) \odot g_t\right). \tag{12}$$

f_u: Learnable projection layer. In the implementation, it is a linear layer, i.e., $\mathrm{Linear}(D, 1)$.

$\widetilde{g}_t$: Spiking rhythm gate at time t; dynamically derived from the macroscopic rhythm signal g_t:

$$\widetilde{g}_t = \sigma\left(f_{\mathrm{gate}}(g_t)\right). \tag{13}$$

f_{gate}: Learnable gate projection layer. In the implementation, it is a linear layer, i.e., $\mathrm{Linear}(d, d)$.

$\sigma(\cdot)$: Sigmoid function used to map the projected rhythm representation into a soft gate in $[0, 1]$.

h_t: Candidate membrane state at time t; dynamic quantity before rhythm-gated membrane update.

v_t: Membrane potential at time t; dynamic state variable of the spiking neuron.

o_t: Spike output at time t; sparse binary state generated by thresholding the membrane potential.

κ: Membrane decay factor; learnable parameter controlling the leakage of the membrane potential from the previous time step.

ϑ: Firing threshold; hyperparameter determining when the membrane potential produces a spike.

$\mathcal{H}(\cdot)$: Heaviside step function; non-learnable spike-generation operator. During training, *Vibe-Spike* approximates its backward pass using the MG surrogate-gradient function.

Equation Explanation: The candidate membrane state h_t contains three operations. The first term preserves part of the previous membrane potential. The second term integrates the current synaptic drive derived from the rhythm-modulated oscillatory phase. The third term resets the membrane after a previous spike. The projected rhythm gate $\widetilde{g}_t$ then selectively updates the membrane potential and modulates spike generation. A spike is emitted when the updated membrane potential exceeds the firing threshold, i.e., $v_t > \vartheta$.

A.5 Spiking readout

Given $O = [o_1, o_2, \ldots, o_{T'}]$, the spike vector is first projected by a linear readout transformation: $a_t = W_{\mathrm{read}} o_t + b_{\mathrm{read}}$. The projected signal is then accumulated by a non-rhythmic spiking readout neuron (the same dynamics in Eq. 6 but without gating):

$$y_t = \Phi_{\mathrm{read}}(a_t, y_{t-1}), \tag{14}$$

where y_t denotes the readout state at time t.

B Pre-training data sources

Table 5 lists the public data sources used in the REVE open-access pre-training subset.

Table 5: Public datasets included in the REVE open-access pre-training subset.

Dataset / source	Reference
Spatial memory and non-invasive closed-loop stimulus timing	[Rudoler et al., 2023]
FakeFaceEmo_data	[Makowski et al., 2023]
Music and Speech Elicit Similar Subcortical Responses in Human Listeners	[Shan et al., 2024]
EEG-attention-rsvp-exp2	[Grootswagers et al., 2023b]
Search Superiority Recollection Familiarity	[Helbing et al., 2024]
Capacity for movement is an organisational principle in object representations: EEG data from Experiment 2	[Shatek et al., 2022]
The time-course of feature-based attention effects dissociated from temporal expectation and target-related processes	[Moerel et al., 2022]
Capacity for movement is an organisational principle in object representations: EEG data from Experiment 1	[Shatek et al., 2023]
Features-EEG	[Grootswagers et al., 2024]
Human electroencephalography recordings from 50 subjects for 22,248 images from 1,854 object concepts	[Grootswagers et al., 2022]
EEG-attention-rsvp-exp1	[Grootswagers et al., 2023a]
Dataset of electrophysiological signals (EEG, ECG, EMG) during Music therapy with adult burn patients in the Intensive Care Unit.	[Cordoba-Silva et al., 2023]
ArEEG: Arabic Inner Speech EEG dataset	[Metwalli et al., 2025]
PES - Pandemic Emergency Scenario	[Tasos Papastylianou et al., 2023]
Neural representation of consciously seen and unseen information	[Esteban et al., 2024]
Illusion of Agency over Electrically-Actuated Movements	[Veillette et al., 2023]
Braille letters - EEG	[Haupt et al., 2024]
ToonFaces	[Chacón and Wriessnegger, 2023]
STERNBERG	[Zhozhikashvili et al., 2024]
EEG meditation study	[Delorme and Brandmeyer, 2024]
EEG, ECG and pupil data from young and older adults: rest and auditory cued reaction time tasks	[Ribeiro and Castelo-Branco, 2021]
Visual Attribute-Specific Contextual Trajectory Paradigm	[Benjamin Lowe (Ben.Lowe@Mq.Edu.Au) et al., 2023]
Meditation vs thinking task	[Delorme and Braboszcz, 2021]
Average Task Value	[Hassall et al., 2024]
Imagined Emotion Study	[Onton and Makeig, 2022]
An EEG dataset recorded during affective music listening	[Daly et al., 2024]
Drum Trainer	[Hassall et al., 2022a]
Competition_EEG	[Aguado-Lopez et al., 2024]
Continuous Feedback Processing	[Hassall et al., 2022b]
Steer the Ship	[Hassall et al., 2022c]
Mood Manipulation and PST, Experiment 1	[Cavanagh and Jackson, 2022]

Continued on next page

Dataset / source	Reference
Invariant encoding of phonemes in neural responses to continuous speech	[Bialas et al., 2023]
Siefert2024	[Siefert et al., 2024]
SRM Resting-state EEG	[Hatlestad-Hall et al., 2022]
The effects of sham hypnosis techniques	[Zoltan Kekecs et al., 2026]
UC San Diego Resting State EEG Data from Patients with Parkinson's Disease	[Rockhill et al., 2022]
Prediction Error	[Gehrke et al., 2024]
Large Spanish EEG	[Araya et al., 2024]
A test-retest resting and cognitive state EEG dataset	[Yulin Wang et al., 2022]
A Resting-state EEG Dataset for Sleep Deprivation	[Chuqin Xiang et al., 2024]
Go-nogo categorization and detection task	[Delorme, 2021]
eeg-neuroforecasting	[Veillette et al., 2022]
Penn Electrophysiology of Encoding and Retrieval Study (PEERS)	[Kahana and Rudoler, 2023]
Healthy Brain Network (HBN) EEG - Release 1	[Shirazi et al., 2024a]
Healthy Brain Network (HBN) EEG - Release 2	[Shirazi et al., 2025a]
Healthy Brain Network (HBN) EEG - Release 3	[Shirazi et al., 2025b]
Healthy Brain Network (HBN) EEG - Release 5	[Shirazi et al., 2025c]
Healthy Brain Network (HBN) EEG - Release 6	[Shirazi et al., 2024c]
Healthy Brain Network (HBN) EEG - Release 7	[Shirazi et al., 2025d]
Healthy Brain Network (HBN) EEG - Release 8	[Shirazi et al., 2025e]
Healthy Brain Network (HBN) EEG - Release 9	[Shirazi et al., 2025f]
Healthy Brain Network (HBN) EEG - Release 4	[Shirazi et al., 2024b]
PerceiveImagine	[Weilong Li and Jiaxin Zhao, 2024]
A repeated awakening study exploring the capacity of complexity measures to capture dreaming during propofol sedation	[Bajwa1 et al., 2025]
Alphabetic Decision Task	[Taylor et al., 2024]
EEG responses to the number of objects in partially occluded and uncovered scenes	[Baykan and Schütz, 2024]
Commande robuste d'un effecteur par une interface cerveau machine EEG asynchrone	[Barachant, 2012]
EEG datasets for motor imagery brain–computer interface	[Cho et al., 2017]
EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy	[Lee et al., 2019a]
An EEG motor imagery dataset for brain computer interface in acute stroke patients	[Liu et al., 2024]
Upper limb movements can be decoded from the time-domain of low-frequency EEG	[Ofner et al., 2017]
BCI2000: a general-purpose brain-computer interface (BCI) system	[Schalk et al., 2004]
PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals	[Goldberger et al., 2000]
Evaluation of EEG oscillatory patterns and cognitive process during simple and compound limb motor imagery	[Yi et al., 2014]
A fully automated trial selection method for optimization of motor imagery based brain-computer interface	[Zhou et al., 2016]
Deep learning with convolutional neural networks for EEG decoding and visualization	[Schirrmeister et al., 2017]
Using Riemannian geometry for SSVEP-based brain computer interface	[Kalunga et al., 2015]

Continued on next page

Dataset / source	Reference
EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy	[Lee et al., 2019c]
Brain Invaders calibration-less P300-based BCI using dry EEG electrodes Dataset (bi2014a)	[Korczowski et al., 2019c]
Brain invaders solo versus collaboration: Multi-user P300-based brain-computer interface dataset (bi2014b)	[Korczowski et al., 2019d]
Brain Invaders calibration-less P300-based BCI with modulation of flash duration Dataset (bi2015a)	[Korczowski et al., 2019a]
Brain invaders cooperative versus competitive: multi-user P300-based brain-computer interface dataset (BI2015b)	[Korczowski et al., 2019b]
Spatial filters for auditory evoked potentials transfer between different experimental conditions.	[Sosulski and Tangermann, 2019]
EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy	[Lee et al., 2019b]
I-CARE: International Cardiac Arrest REsearch consortium Database	[Amorim et al., 2023]

C Additional pre-training settings

Table 6 7 provides the detailed configuration used for self-supervised pre-training.

Table 7: Default implementation settings of *Vibe-Spike*.

Category	Setting
Input encoder	Enabled
Encoder hidden dimension	32
Coupling strength λ	1.0
Euler integration step Δt	1.0
Phase-delay step τ	1
Phase-lag construction	Structural-prior-derived by default
SNN hidden dimension	128
Rhythm-gated SNN layers	1
SNN firing threshold	0.5
Readout firing threshold	0.5
Initial membrane state	Zero initialization
Initial spike state	Zero initialization
Surrogate gradient	MG surrogate gradient

D Downstream dataset preprocessing

This section summarizes the preprocessing procedures used for all downstream datasets. For datasets that have been included in prior EEG foundation model benchmarks, we follow the reported preprocessing protocols and train/validation/test splits whenever applicable [Wang et al., 2025, Jiang et al., 2024, Yang et al., 2023a, Ouahidi et al., 2026]. For newly added datasets, we adopt the same subject-wise or recording-wise splitting principle to avoid data leakage, and we use each baseline model with its recommended task-specific configuration. All inputs to our model are resampled to 100 Hz, whereas baseline models are evaluated using their prescribed sampling rates and preprocessing settings. For bipolar derivations, we keep the original bipolar signal. For a channel denoted as $A-B$, we use the coordinate of the first electrode A as the channel position for montage alignment.

We follow REVE's pretraining data curation and exclusion protocol, as described in Section 3.1. For the additional downstream datasets, ADFTD-RS, SADT, and SEED-VIG, we also cross-checked their source cohorts against the constituent pretraining datasets listed in Appendix B. These checks confirmed that no subjects or patients are shared between our pretraining corpus and any downstream dataset.

Table 6: Detailed pre-training configuration.

Category	Setting
Dataset	REVE open-access subset
Input window length	2000 samples
Sampling rate	100 Hz
Target montage	10-20 standard 19 montage
Batch size	100 per process
Training epochs	80
STFT n_{fft}	128
STFT window length	128
STFT hop length	16
Frequency range	0–45 Hz
Masking strategy	Structured spatio-temporal block masking
Masking window	200 samples
Masking overlap	20 samples
Masking ratio	0.25
Spatial masking radius	0.03
Temporal masking radius	3
Masking dropout ratio	0.10
Masking dropout radius	0.04
Target channels	19
Projection dimension	128
Temporal stem kernels	$\{3, 5, 9\}$
Vibe-Spike output dimension	128
Number of *Vibe-Spike* layers	1
Oscillator dimension	2
Graph construction	Internal correlation graph from projected features
Graph diagonal value	0.5
Decoder hidden dimension	256
Decoder dropout	0.00
Decoder block token	Enabled
Optimizer	AdamW
Learning rate	3.0×10^{-4}
Weight decay	1.0×10^{-3}
Adam betas	$(0.9, 0.99)$
Adam epsilon	1.0×10^{-8}
Scheduler	Warmup-stable-decay
Warmup steps	5000
Stable steps	0
Minimum learning rate	1.0×10^{-5}
Gradient clipping	1.0
Mixed precision	Disabled
Gradient accumulation	1
Hardware	8 NVIDIA RTX 5090 GPUs
Training time	Approximately one day

D.1 ISRUC

ISRUC-Sleep Khalighi et al. [2016] is used for automatic sleep staging. We use ISRUC group 1, which contains 100 overnight polysomnography recordings with expert sleep-stage annotations. Following the preprocessing protocol of REVE and CBraMod Ouahidi et al. [2026], Wang et al. [2025], raw signals are band-pass filtered at 0.3–35 Hz, notch filtered at 50 Hz, and segmented into 30-second epochs. Consecutive epochs are further grouped into 20-epoch sequences, and the original annotations are mapped into five sleep-stage classes: Wake, N1, N2, N3, and REM. For subject-wise evaluation, subjects 1–80 are used for training, subjects 81–90 for validation, and subjects 91–100 for testing.

D.2 HMC

HMC Alvarez-Estevez and Rijsman [2022] is used for automatic sleep staging. It contains 151 overnight polysomnography recordings collected at the Haaglanden Medisch Centrum sleep center, with expert sleep-stage annotations. Following REVE and CBraMod Ouahidi et al. [2026], Wang et al. [2025], we retain four EEG channels (F4-M1, C4-M1, O2-M1, and C3-M2), remove non-EEG channels, apply band-pass and notch filtering, and segment the recordings into 30-second epochs. The labels are mapped into five classes: Wake, N1, N2, N3, and REM. Following the REVE split protocol Ouahidi et al. [2026], the first 100 recordings are used for training, the next 25 recordings for validation, and the remaining recordings for testing.

D.3 Mumtaz2016

Mumtaz2016 Mumtaz [2016] is used for mental disorder diagnosis. It contains EEG recordings from 34 subjects with major depressive disorder (MDD) and 30 healthy controls, acquired from 19 electrodes following the 10–20 system. Following REVE Ouahidi et al. [2026], we use only the eyes-open and eyes-closed sessions, discard task recordings, apply band-pass and notch filtering, and segment the signals into 5-second windows. The task is binary MDD-vs-control classification. The split includes 24 MDD and 19 control subjects for training, 5 MDD and 4 controls for validation, and 5 MDD and 5 controls for testing.

D.4 CHB-MIT

CHB-MIT Guttag [2010] is used for seizure detection. It contains long-term scalp EEG recordings from pediatric subjects with epileptic seizures. Following CBraMod Wang et al. [2025], we use the common 16 bipolar montage channels, segment the continuous recordings into 10-second windows, and formulate the task as binary seizure-vs-non-seizure classification. A segment is labeled as seizure if the annotated seizure onset or offset falls within the window. Additional positive windows are extracted around seizure intervals with a 5-second stride to alleviate class imbalance. For evaluation, `chb21` and `chb22` are used for validation, `chb23` and `chb24` for testing, and the remaining processed subjects for training.

D.5 DFTD-RS

DFTD-RS Miltiadous et al. [2023] is used for Alzheimer's disease detection. It contains resting-state eyes-closed EEG recordings from subjects with Alzheimer's disease (AD), frontotemporal dementia (FTD), and cognitively normal controls (CN). We use only AD and CN subjects and formulate the task as binary AD-vs-CN classification. We retain the 19 standard 10–20 scalp channels and segment the continuous recordings into non-overlapping 30-second windows. CN and AD samples are labeled as 0 and 1, respectively. For evaluation, AD subjects `sub-034`–`sub-036` and CN subjects `sub-063`–`sub-065` are used for validation, while the remaining AD subjects `sub-001`–`sub-033` and CN subjects `sub-037`–`sub-062` are used for training.

D.6 SADT

SADT Cao et al. [2019] is used for vigilance state classification. It contains multi-channel EEG recordings collected during a sustained-attention driving task. We preprocess the dataset independently and retain 30 scalp EEG channels after removing the two reference channels. The signals are band-pass filtered and segmented into 5-second windows before each deviation onset. Trial labels are derived from response time: response-time z-scores are computed relative to the first 10 minutes of each session, where trials with $z \leq 1$ are labeled as optimal and trials with $z > 2$ are labeled as poor. We only use optimal and poor trials and formulate the task as binary vigilance classification, while suboptimal trials are excluded. For evaluation, we use a subject-wise split with 70%, 15%, and 15% of subjects for training, validation, and testing, respectively.

D.7 SEED-VIG

SEED-VIG Zheng and Lu [2017], Cao et al. [2019] is used for vigilance state classification. It contains multi-channel EEG recordings collected during a sustained-attention simulated driving task,

with vigilance levels measured by the PERCLOS indicator. We preprocess the dataset independently and retain the 17 scalp EEG channels used in the standard SEED-VIG montage. The signals are downsampled and segmented into non-overlapping 8-second windows. Different from CBraMod, which formulates SEED-VIG as a regression task, we convert PERCLOS into binary vigilance labels to better match practical alert/drowsy detection scenarios and to facilitate unified classification-based evaluation across datasets. Windows with PERCLOS ≥ 0.35 are labeled as low vigilance (drowsy), while windows with PERCLOS < 0.35 are labeled as high vigilance (alert). Following BIOT and CBraMod Yang et al. [2023a], Wang et al. [2025], we use subjects 1–15 for training, subjects 16–19 for validation, and subjects 20–23 for testing.

D.8 MentalArithmetic

MentalArithmetic Zyma et al. [2019] is used for stress detection. It contains EEG recordings from subjects under resting baseline and mental arithmetic conditions. Following REVE Ouahidi et al. [2026], we formulate the task as binary baseline-vs-stress classification, where files with suffix `_1` are labeled as baseline and files with suffix `_2` are labeled as stress. We retain the 20 channels used in the REVE version, apply band-pass filtering, and segment the continuous recordings into 5-second windows. For evaluation, `Subject00`–`Subject27` are used for training, `Subject28`–`Subject31` for validation, and `Subject32`–`Subject35` for testing.

D.9 TUEV

TUEV Obeid and Picone [2016] is used for event type classification. It contains EEG event annotations from the Temple University Hospital EEG corpus. Following BIOT Yang et al. [2023a], we use the common 16 TCP bipolar montage channels and formulate the task as 6-class event classification, including SPSW, GPED, PLED, EYEM, ARTF, and BCKG. For each annotated event, we extract a 5-second EEG segment centered around the event interval. For evaluation, the official training split is used for training, and the official evaluation split is used for validation.

D.10 TUAB

TUAB Obeid and Picone [2016] is used for abnormal EEG detection. It contains EEG recordings from the Temple University Hospital EEG corpus annotated as normal or abnormal. Following BIOT Yang et al. [2023a], we use the common 16 TCP bipolar montage channels, segment the continuous recordings into non-overlapping 10-second windows, and formulate the task as binary normal-vs-abnormal classification. For evaluation, the official training split is further divided into training and validation sets with an 80:20 subject-wise ratio for each class, and the official evaluation split is used for testing.

E Additional Downstream Task Performance

In this section, we provide task-wise supplementary downstream results of *Vibe-Spike*, including evaluation metrics beyond the balanced accuracy reported in the main text.

E.1 Result Provenance and Statistical Testing

Table 8 summarizes the provenance of the downstream results. N denotes the number of independent runs per method.

We compare the methods with the highest and second-highest mean balanced accuracy on each dataset using two-sided Welch's t-tests. The tests are calculated from the reported means, standard deviations, and run counts. They quantify variability across runs under fixed data splits.

Table 9 reports the test results. Rank 1 and Rank 2 denote the two compared methods, and N_1/N_2 gives their respective run counts. Δ is the balanced accuracy of Rank 1 minus Rank 2 in percentage points (pp). All p-values are unadjusted for multiple comparisons. Significance is denoted by $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$; n.s. denotes $p \geq 0.05$.

Table 8: Sources, run counts, and evaluation protocols for downstream results.

Benchmark setting	**Datasets**	**Published results**	**Evaluated in this study**	**Evaluation protocol**
Protocol aligned	TUAB, TUEV, ISRUC, HMC, Mumtaz2016, MentalArithmetic	All baselines reported in REVE [Ouahidi et al., 2026] $N = 10$	*Vibe-Spike* $N = 10$	REVE-released preprocessed files and original benchmark splits
Partially rerun	CHB-MIT	Baseline results from BIOT [Yang et al., 2023a] and CBraMod [Wang et al., 2025] $N = 5$	REVE and *Vibe-Spike* $N = 10$	BIOT-released processed files and original benchmark split
Fully rerun	ADFTD-RS, SADT, SEED-VIG	—	All methods $N = 10$	Shared preprocessing and deterministic subject-level splits (Appendix D)

Table 9: Two-sided Welch's t-tests for the top two methods on each dataset.

Dataset	**Rank 1**	**Rank 2**	N_1/N_2	Δ (pp)	t	**df**	p	**Sig.**
TUAB	REVE	CBraMod	10/10	$+0.26$	3.153	15.26	0.0065	$**$
TUEV	REVE	CBraMod	10/10	$+0.88$	1.101	12.75	0.2913	n.s.
ISRUC	*Vibe-Spike*	CBraMod	10/10	$+1.12$	2.537	17.01	0.0213	$*$
HMC	*Vibe-Spike*	REVE	10/10	$+4.47$	10.110	15.25	3.71×10^{-8}	$***$
Mumtaz2016	*Vibe-Spike*	REVE	10/10	$+0.50$	1.335	16.13	0.2005	n.s.
CHB-MIT	*Vibe-Spike*	CBraMod	10/5	$+9.25$	5.988	7.94	3.37×10^{-4}	$***$
ADFTD-RS	*Vibe-Spike*	REVE	10/10	$+9.14$	6.146	16.13	1.36×10^{-5}	$***$
SADT	*Vibe-Spike*	CNN-Transformer	10/10	$+0.07$	0.587	15.10	0.5661	n.s.
SEED-VIG	*Vibe-Spike*	SPaRCNet	10/10	$+0.83$	0.648	10.32	0.5312	n.s.
MentalArithmetic	REVE	*Vibe-Spike*	10/10	$+3.83$	2.552	17.75	0.0202	$*$

E.2 Sleep Staging

We report additional results on ISRUC and HMC for multi-class sleep-stage classification.

Table 10: Additional results on ISRUC sleep staging (5-class).

Method	Params	Balanced Accuracy	Cohen's Kappa	Weighted F1
EEGNet	0.003M	0.7154 ± 0.0121	0.7040 ± 0.0173	0.7513 ± 0.0124
EEGConformer	0.55M	0.7400 ± 0.0133	0.7143 ± 0.0162	0.7634 ± 0.0151
SPaRCNet	0.79M	0.7487 ± 0.0075	0.7097 ± 0.0132	0.7624 ± 0.0092
ContraWR	1.6M	0.7402 ± 0.0126	0.7178 ± 0.0156	0.7610 ± 0.0137
CNN-Transformer	3.2M	0.7363 ± 0.0087	0.7129 ± 0.0121	0.7719 ± 0.0105
FFCL	2.4M	0.7277 ± 0.0182	0.7016 ± 0.0291	0.7614 ± 0.0197
ST-Transformer	3.5M	0.7381 ± 0.0205	0.7013 ± 0.0352	0.7681 ± 0.0175
DeepSleepNet	–	0.7419 ± 0.0144	0.7036 ± 0.0241	0.7643 ± 0.0122
USleep	–	0.7586 ± 0.0116	0.7209 ± 0.0143	0.7805 ± 0.0105
BIOT	3.2M	0.7527 ± 0.0121	0.7192 ± 0.0231	0.7790 ± 0.0146
LaBraM	5.8M	0.7633 ± 0.0102	0.7231 ± 0.0182	0.7810 ± 0.0133
CBraMod	4.0M	0.7865 ± 0.0110	0.7442 ± 0.0152	0.8011 ± 0.0099
REVE	69M	0.7819 ± 0.0078	$\mathbf{0.7500} \pm 0.0156$	0.8005 ± 0.0135
Vibe-Spike	0.41M	$\mathbf{0.7977} \pm 0.0086$	0.7474 ± 0.0155	$\mathbf{0.8022} \pm 0.0114$

Table 11: Additional results on HMC sleep staging (5-class).

Method	Params	Balanced Accuracy	Cohen's Kappa	Weighted F1
EEGNet	0.003M	0.6856 ± 0.0064	0.5050 ± 0.0092	0.5493 ± 0.0127
EEGConformer	0.55M	0.7210 ± 0.0102	0.5742 ± 0.0153	0.6384 ± 0.0188
SPaRCNet	0.79M	0.4756 ± 0.1109	0.3147 ± 0.1315	0.4108 ± 0.1310
ContraWR	1.6M	0.4242 ± 0.0541	0.2340 ± 0.0554	0.2987 ± 0.0288
CNN-Transformer	3.2M	0.6573 ± 0.0141	0.5961 ± 0.0105	0.6896 ± 0.0065
FFCL	2.4M	0.4427 ± 0.0702	0.2542 ± 0.0654	0.2902 ± 0.0485
ST-Transformer	3.5M	0.2559 ± 0.0141	0.0503 ± 0.0183	0.1428 ± 0.0122
BIOT	3.2M	0.6862 ± 0.0041	0.6295 ± 0.0113	0.7091 ± 0.0147
LaBraM	5.8M	0.7286 ± 0.0101	0.6812 ± 0.0073	0.7554 ± 0.0024
CBraMod	4.0M	0.7335 ± 0.0047	0.6094 ± 0.0134	0.6829 ± 0.0162
REVE	69M	0.7401 ± 0.0075	0.6982 ± 0.0078	0.7638 ± 0.0074
Vibe-Spike	0.41M	$\mathbf{0.7848} \pm 0.0125$	$\mathbf{0.7713} \pm 0.0164$	$\mathbf{0.8237} \pm 0.0122$

E.3 Seizure Detection

We report additional results on CHB-MIT for binary seizure-vs-non-seizure classification.

Table 12: Additional results on CHB-MIT seizure detection (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.5658 ± 0.0106	0.1914 ± 0.0182	0.8048 ± 0.0136
EEGConformer	0.55M	0.5976 ± 0.0141	0.2209 ± 0.0215	0.8226 ± 0.0170
SPaRCNet	0.79M	0.5876 ± 0.0191	0.1247 ± 0.0119	0.8143 ± 0.0148
ContraWR	1.6M	0.6344 ± 0.0002	0.2264 ± 0.0174	0.8097 ± 0.0114
CNN-Transformer	3.2M	0.6389 ± 0.0067	0.2479 ± 0.0227	0.8662 ± 0.0082
FFCL	2.4M	0.6262 ± 0.0104	0.2049 ± 0.0346	0.8271 ± 0.0051
ST-Transformer	3.5M	0.5915 ± 0.0195	0.1422 ± 0.0094	0.8237 ± 0.0491
BIOT	3.2M	0.7068 ± 0.0457	0.3277 ± 0.0460	0.8761 ± 0.0284
LaBraM	5.8M	0.7075 ± 0.0358	0.3287 ± 0.0402	0.8679 ± 0.0199
CBraMod	4.0M	0.7398 ± 0.0284	0.3689 ± 0.0382	0.8892 ± 0.0154
REVE	69M	0.6919 ± 0.0193	0.2573 ± 0.0630	0.8919 ± 0.0140
Vibe-Spike	0.41M	**0.8323** ± 0.0278	**0.3773** ± 0.0320	**0.9114** ± 0.0346

E.4 Mental Disorder Diagnosis

We report additional results on Mumtaz2016 for binary MDD-vs-control classification.

Table 13: Additional results on Mumtaz2016 mental disorder diagnosis (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.9232 ± 0.0104	0.9626 ± 0.0095	0.9639 ± 0.0093
EEGConformer	0.55M	0.9308 ± 0.0117	0.9684 ± 0.0105	0.9702 ± 0.0101
SPaRCNet	0.79M	0.9316 ± 0.0095	0.9754 ± 0.0065	0.9781 ± 0.0083
ContraWR	1.6M	0.9195 ± 0.0115	0.9589 ± 0.0102	0.9621 ± 0.0092
CNN-Transformer	3.2M	0.9305 ± 0.0068	0.9757 ± 0.0074	0.9742 ± 0.0059
FFCL	2.4M	0.9314 ± 0.0038	0.9717 ± 0.0021	0.9753 ± 0.0033
ST-Transformer	3.5M	0.9135 ± 0.0103	0.9578 ± 0.0086	0.9594 ± 0.0059
BIOT	3.2M	0.9358 ± 0.0052	0.9736 ± 0.0034	0.9758 ± 0.0042
LaBraM	5.8M	0.9409 ± 0.0079	0.9798 ± 0.0093	0.9782 ± 0.0057
CBraMod	4.0M	0.9560 ± 0.0056	0.9923 ± 0.0032	0.9921 ± 0.0025
REVE	69M	0.9644 ± 0.0097	0.9961 ± 0.0013	**0.9957** ± 0.0015
Vibe-Spike	0.41M	**0.9694** ± 0.0068	**0.9964** ± 0.0009	0.9942 ± 0.0014

E.5 AD Diagnosis

We report additional results on ADFTD-RS for binary AD-vs-CN classification.

Table 14: Additional results on ADFTD-RS AD diagnosis (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.6885 ± 0.0374	0.7935 ± 0.0362	0.7414 ± 0.0388
EEGConformer	0.55M	0.7250 ± 0.0186	0.8446 ± 0.0105	0.7785 ± 0.0154
SPaRCNet	0.79M	0.7280 ± 0.0167	0.8557 ± 0.0111	0.8061 ± 0.0139
ContraWR	1.6M	0.6885 ± 0.0155	0.8360 ± 0.0137	0.7679 ± 0.0252
CNN-Transformer	3.2M	0.7335 ± 0.0238	0.8482 ± 0.0148	0.7996 ± 0.0172
FFCL	2.4M	0.7127 ± 0.0208	0.8220 ± 0.0181	0.7648 ± 0.0228
ST-Transformer	3.5M	0.7144 ± 0.0145	0.8272 ± 0.0205	0.7864 ± 0.0174
BIOT	3.2M	0.7330 ± 0.0141	0.8606 ± 0.0112	0.7879 ± 0.0313
LaBraM	5.8M	0.6962 ± 0.0184	0.8356 ± 0.0238	0.7585 ± 0.0378
CBraMod	4.0M	0.7564 ± 0.0193	0.8548 ± 0.0194	0.7817 ± 0.0300
REVE	69M	0.7750 ± 0.0385	0.8985 ± 0.0291	0.8583 ± 0.0434
Vibe-Spike	0.41M	$\mathbf{0.8664} \pm 0.0270$	$\mathbf{0.9384} \pm 0.0241$	$\mathbf{0.9423} \pm 0.0191$

E.6 Vigilance Classification

We report additional results on SADT and SEED-VIG for vigilance-state classification.

Table 15: Additional results on SADT vigilance classification (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.6890 ± 0.0036	0.3748 ± 0.0066	0.6993 ± 0.0048
EEGConformer	0.55M	0.7301 ± 0.0106	0.4401 ± 0.0242	0.7910 ± 0.0102
SPaRCNet	0.79M	0.7558 ± 0.0121	0.5519 ± 0.0102	$\mathbf{0.8244} \pm 0.0091$
ContraWR	1.6M	0.7013 ± 0.0069	0.5469 ± 0.0144	0.7827 ± 0.0033
CNN-Transformer	3.2M	0.7561 ± 0.0032	0.5402 ± 0.0068	0.8151 ± 0.0102
FFCL	2.4M	0.7289 ± 0.0043	0.5692 ± 0.0117	0.8147 ± 0.0061
ST-Transformer	3.5M	0.7098 ± 0.0067	0.4827 ± 0.0127	0.7866 ± 0.0054
BIOT	3.2M	0.7322 ± 0.0106	0.5130 ± 0.0171	0.8091 ± 0.0069
LaBraM	5.8M	0.7036 ± 0.0057	0.4673 ± 0.0050	0.7657 ± 0.0045
CBraMod	4.0M	0.7494 ± 0.0055	0.5128 ± 0.0044	0.8106 ± 0.0035
REVE	69M	0.6869 ± 0.0196	0.4890 ± 0.0228	0.7519 ± 0.0097
Vibe-Spike	0.41M	$\mathbf{0.7568} \pm 0.0026$	$\mathbf{0.5731} \pm 0.0179$	0.8105 ± 0.0072

Table 16: Additional results on SEED-VIG vigilance classification (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.6843 ± 0.0480	0.7701 ± 0.1916	0.7319 ± 0.0542
EEGConformer	0.55M	0.7262 ± 0.0476	0.8276 ± 0.1710	0.7999 ± 0.0718
SPaRCNet	0.79M	0.7265 ± 0.0391	0.8183 ± 0.1630	0.7968 ± 0.0524
ContraWR	1.6M	0.7223 ± 0.0265	0.8336 ± 0.1538	$\mathbf{0.8060} \pm 0.0457$
CNN-Transformer	3.2M	0.6956 ± 0.0308	0.7916 ± 0.1767	0.7509 ± 0.0531
FFCL	2.4M	0.7177 ± 0.0434	0.8139 ± 0.1710	0.7858 ± 0.0613
ST-Transformer	3.5M	0.6721 ± 0.0348	0.7746 ± 0.1713	0.7325 ± 0.0419
BIOT	3.2M	0.7242 ± 0.0364	0.8382 ± 0.1442	0.8039 ± 0.0506
LaBraM	5.8M	0.6630 ± 0.0593	0.7685 ± 0.1916	0.7257 ± 0.0699
CBraMod	4.0M	0.6488 ± 0.0331	0.7370 ± 0.1949	0.7038 ± 0.0484
REVE	69M	0.6569 ± 0.1096	0.7651 ± 0.1368	0.7041 ± 0.1287
Vibe-Spike	0.41M	$\mathbf{0.7348} \pm 0.0106$	$\mathbf{0.9033} \pm 0.0137$	0.8029 ± 0.0142

E.7 Stress Classification

We report additional results on MentalArithmetic for binary baseline-vs-stress classification.

Table 17: Additional results on MentalArithmetic stress detection (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.6770 ± 0.0116	0.5763 ± 0.0102	0.7321 ± 0.0108
EEGConformer	0.55M	0.6805 ± 0.0123	0.5829 ± 0.0134	0.7424 ± 0.0128
SPaRCNet	0.79M	0.6879 ± 0.0107	0.5825 ± 0.0193	0.7418 ± 0.0132
ContraWR	1.6M	0.6631 ± 0.0097	0.5787 ± 0.0164	0.7332 ± 0.0082
CNN-Transformer	3.2M	0.6779 ± 0.0268	0.5777 ± 0.0285	0.7258 ± 0.0336
FFCL	2.4M	0.6798 ± 0.0142	0.5786 ± 0.0266	0.7330 ± 0.0198
ST-Transformer	3.5M	0.6631 ± 0.0173	0.5672 ± 0.0259	0.7132 ± 0.0174
BIOT	3.2M	0.6875 ± 0.0186	0.6004 ± 0.0195	0.7536 ± 0.0144
LaBraM	5.8M	0.6909 ± 0.0125	0.5999 ± 0.0155	0.7721 ± 0.0093
CBraMod	4.0M	0.7256 ± 0.0132	0.6267 ± 0.0099	0.7905 ± 0.0073
REVE	69M	$\mathbf{0.7660} \pm 0.0355$	$\mathbf{0.7470} \pm 0.0807$	$\mathbf{0.8450} \pm 0.0514$
Vibe-Spike	0.41M	0.7277 ± 0.0315	0.6120 ± 0.0421	0.7697 ± 0.0338

E.8 EEG Event Type Classification

We report additional results on TUEV for six-class EEG event type classification.

Table 18: Additional results on TUEV event type classification (6-class).

Method	Params	Balanced Accuracy	Cohen's Kappa	Weighted F1
EEGNet	0.003M	0.3876 ± 0.0143	0.3577 ± 0.0155	0.6539 ± 0.0120
EEGConformer	0.55M	0.4074 ± 0.0164	0.3967 ± 0.0195	0.6983 ± 0.0152
SPaRCNet	0.79M	0.4161 ± 0.0262	0.4233 ± 0.0181	0.7024 ± 0.0104
ContraWR	1.6M	0.4384 ± 0.0349	0.3912 ± 0.0237	0.6893 ± 0.0136
CNN-Transformer	3.2M	0.4087 ± 0.0161	0.3815 ± 0.0134	0.6854 ± 0.0293
FFCL	2.4M	0.3979 ± 0.0104	0.3732 ± 0.0188	0.6783 ± 0.0120
ST-Transformer	3.5M	0.3984 ± 0.0228	0.3765 ± 0.0306	0.6823 ± 0.0190
BIOT	3.2M	0.5281 ± 0.0225	0.5273 ± 0.0249	0.7492 ± 0.0082
LaBraM-Base	5.8M	0.6409 ± 0.0065	0.6637 ± 0.0093	0.8312 ± 0.0052
LaBraM-Large	46M	0.6581 ± 0.0156	0.6622 ± 0.0136	0.8315 ± 0.0040
LaBraM-Huge	369M	0.6616 ± 0.0170	0.6745 ± 0.0195	0.8329 ± 0.0086
CBraMod	4.0M	0.6671 ± 0.0107	0.6772 ± 0.0096	0.8342 ± 0.0064
REVE	69M	$\mathbf{0.6759} \pm 0.0229$	$\mathbf{0.6783} \pm 0.0199$	$\mathbf{0.8451} \pm 0.0129$
Vibe-Spike	0.41M	0.5385 ± 0.0148	0.5325 ± 0.0047	0.7553 ± 0.0019

E.9 Abnormal EEG Detection

We report additional results on TUAB for binary normal-vs-abnormal EEG classification.

Table 19: Additional results on TUAB abnormal EEG detection (2-class).

Method	Params	Balanced Accuracy	AUC-PR	AUROC
EEGNet	0.003M	0.7642 ± 0.0036	0.8299 ± 0.0043	0.8412 ± 0.0031
EEGConformer	0.55M	0.7758 ± 0.0049	0.8427 ± 0.0054	0.8445 ± 0.0038
SPaRCNet	0.79M	0.7896 ± 0.0018	0.8414 ± 0.0018	0.8676 ± 0.0012
ContraWR	1.6M	0.7746 ± 0.0041	0.8421 ± 0.0104	0.8456 ± 0.0074
CNN-Transformer	3.2M	0.7777 ± 0.0022	0.8433 ± 0.0039	0.8461 ± 0.0013
FFCL	2.4M	0.7848 ± 0.0038	0.8448 ± 0.0065	0.8569 ± 0.0051
ST-Transformer	3.5M	0.7966 ± 0.0023	0.8521 ± 0.0026	0.8707 ± 0.0019
BIOT	3.2M	0.7959 ± 0.0057	0.8792 ± 0.0023	0.8815 ± 0.0043
LaBraM-Base	5.8M	0.8140 ± 0.0019	0.8965 ± 0.0016	0.9022 ± 0.0009
LaBraM-Large	46M	0.8226 ± 0.0015	0.9130 ± 0.0005	0.9127 ± 0.0005
LaBraM-Huge	369M	0.8258 ± 0.0011	0.9204 ± 0.0011	0.9162 ± 0.0016
CBraMod	4.0M	0.8289 ± 0.0022	0.9258 ± 0.0008	0.9227 ± 0.0011
REVE	69M	$\mathbf{0.8315} \pm 0.0014$	$\mathbf{0.9281} \pm 0.0009$	$\mathbf{0.9245} \pm 0.0013$
Vibe-Spike	0.41M	0.8048 ± 0.0014	0.8894 ± 0.0011	0.8806 ± 0.0023

F Additional Visualization Results

In this section, we provide additional visualization results for downstream datasets not shown in the main text. The main text presents visualization results on CHB-MIT and TUAB, while this appendix includes the remaining downstream datasets.

F.1 Sleep Staging: ISRUC

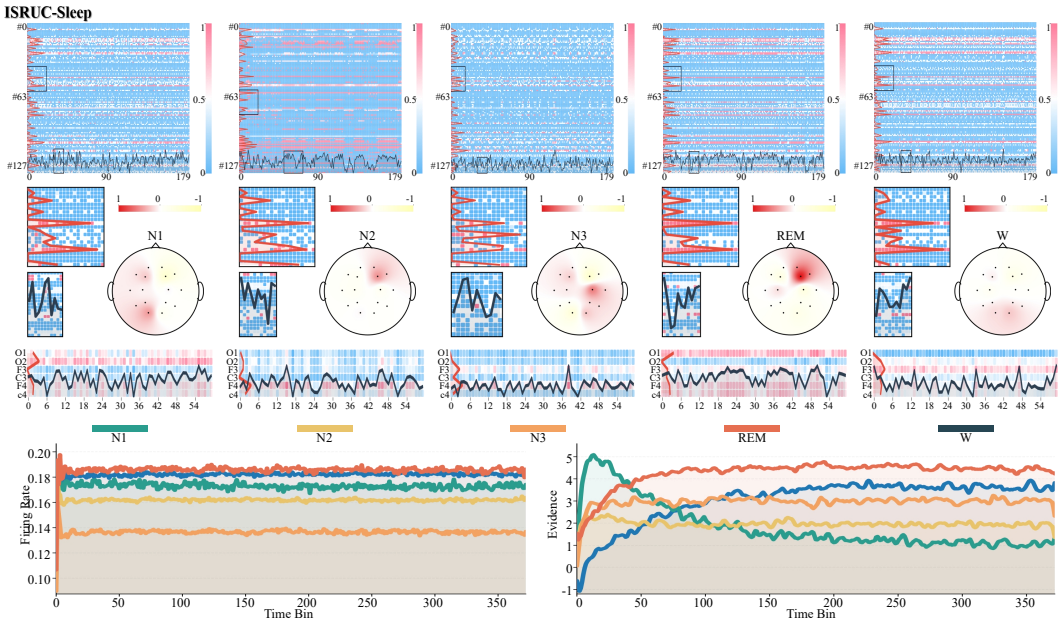

Figure 6: Visualization results of *Vibe-Spike* on ISRUC

F.2 Sleep Staging: HMC

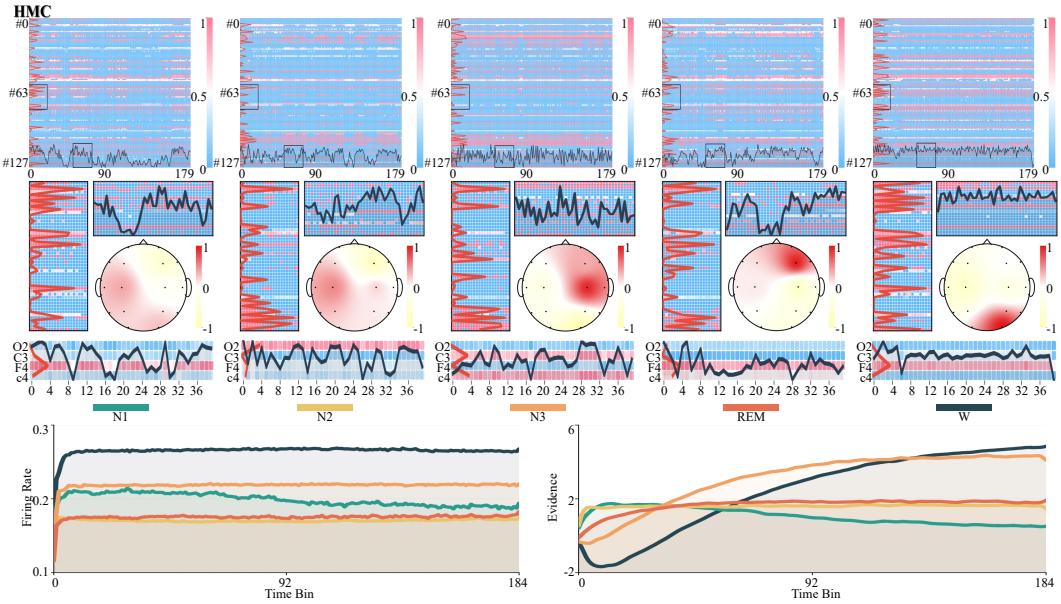

Figure 7: Visualization results of *Vibe-Spike* on HMC

Interpretability for sleep staging. Figs. 6 and 7 provide task-specific interpretability evidence for sleep staging. From the spike-rate perspective, different sleep stages rapidly enter distinct and relatively stable firing-rate regimes after the initial transient, indicating that *Vibe-Spike* encodes sleep stages as sustained rhythmic states rather than isolated transient responses. Notably, N1 and N2 exhibit partially overlapping rate patterns, consistent with the transitional nature of N1 and its physiological proximity to light sleep, whereas N3, REM, and wakefulness show more separated rate trajectories.

From the scalp-manifold perspective, the learned activations show stage-dependent spatial organization that agrees with the visible topographic patterns. N1 and N2 exhibit relatively diffuse and partially similar scalp activations, while N2 shows a more structured central/fronto-central pattern. N3 shows a more concentrated central/centro-parietal activation pattern. REM displays a distinct anterior/fronto-central activation pattern compared with N1–N3. Wakefulness shows a more posterior/occipital-dominant pattern. These observations are compatible with sleep EEG physiology: N2 is associated with organized light-sleep graphoelements such as spindles and K-complexes, N3 reflects more synchronized slow-wave activity, REM is characterized by low-amplitude mixed-frequency activity with ocular/frontal components, and wakefulness commonly involves posterior alpha rhythms [American Academy of Sleep Medicine, 2023]. Importantly, these stage-dependent activation and suppression patterns are observed consistently across ISRUC and HMC, suggesting that *Vibe-Spike* captures sleep-stage-related spatiotemporal rhythms rather than dataset-specific artifacts.

F.3 Mental Disorder Diagnosis: Mumtaz2016

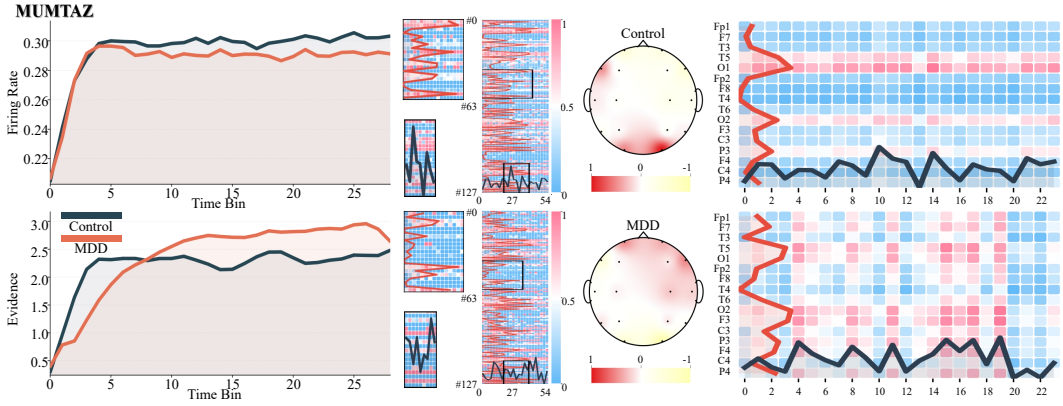

Figure 8: Visualization results of *Vibe-Spike* on Mumtaz

Interpretability for depression recognition. Fig. 8 provides task-specific interpretability evidence for MDD recognition on the Mumtaz dataset. From the spike-rate perspective, both control and MDD samples rapidly enter stable firing regimes, indicating that the model treats depression recognition as a sustained brain-state discrimination problem rather than an event-detection task. Meanwhile, the evidence curves progressively separate the two classes, suggesting that *Vibe-Spike* accumulates persistent rhythmic differences over the EEG segment instead of relying on isolated transient responses.

From the scalp-manifold perspective, the learned activations show class-dependent spatial organization. Compared with the control group, MDD samples exhibit more distributed frontal/frontotemporal-related activation patterns in both the scalp maps and channel-time activation maps. This observation is compatible with depression-related EEG findings, where frontal asymmetry and altered frontal/prefrontal rhythmic activity have been widely investigated as neurophysiological markers of MDD [Henriques and Davidson, 1991, Allen et al., 2004, Luo et al., 2025]. These consistent rate-level and spatial activation differences suggest that *Vibe-Spike* encodes MDD-related EEG abnormalities as sparse, sustained, and spatially organized rhythmic spike manifolds.

F.4 AD Diagnosis: ADFTD-RS

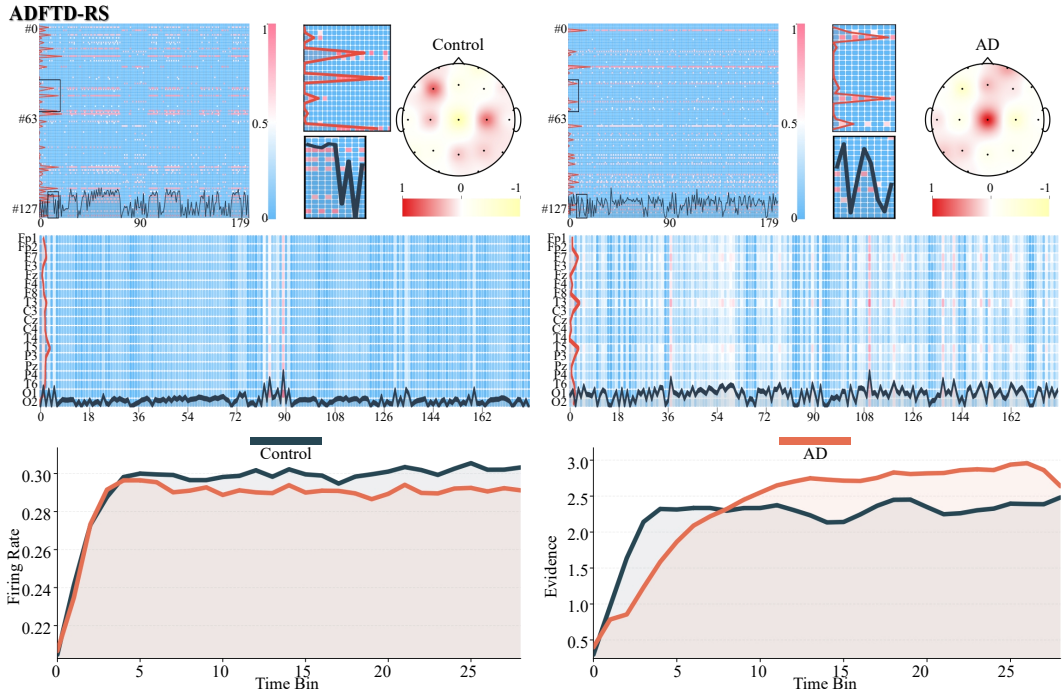

Figure 9: Visualization results of *Vibe-Spike* on ADFTD-RS

Interpretability for AD recognition. Fig. 9 provides task-specific interpretability evidence for AD recognition. From the spike-rate and evidence perspectives, both groups rapidly enter stable firing-rate regimes after the initial transient, whereas the evidence trajectory progressively favors the AD class over time. This suggests that *Vibe-Spike* treats AD recognition as a sustained brain-state discrimination problem, and accumulates class-relevant rhythmic evidence rather than relying on isolated transient responses.

From the scalp-manifold perspective, the learned topographic activations show clear class-dependent differences: compared with the control group, the AD sample exhibits a more pronounced midline frontal–central activation pattern, whereas the control sample shows relatively weaker and more spatially distributed responses. This tendency is further supported by the channel-time activation maps. The control heatmap remains comparatively sparse, with lower-amplitude and more localized activations, while the AD heatmap shows stronger and more temporally persistent activations distributed across multiple channels, especially over frontal-to-central leads. Such a more diffuse and sustained channel-level pattern is physiologically plausible, as AD is widely associated with EEG slowing, increased low-frequency synchronization, and disruption of normal posterior alpha organization, often accompanied by more widespread functional disorganization rather than focal abnormalities [Jeong, 2004, Moretti et al., 2012]. These results suggest that *Vibe-Spike* encodes AD-related EEG abnormalities as sparse but spatially organized rhythmic spike manifolds, with the channel-time maps further revealing their temporally persistent and spatially distributed nature.

F.5 Vigilance Classification: SADT

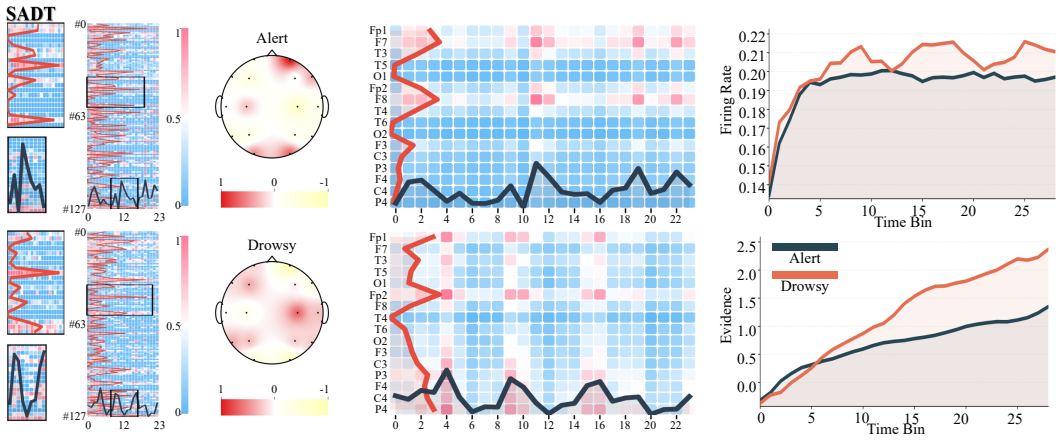

Figure 10: Visualization results of *Vibe-Spike* on SADT

F.6 Vigilance Classification: SEED-VIG

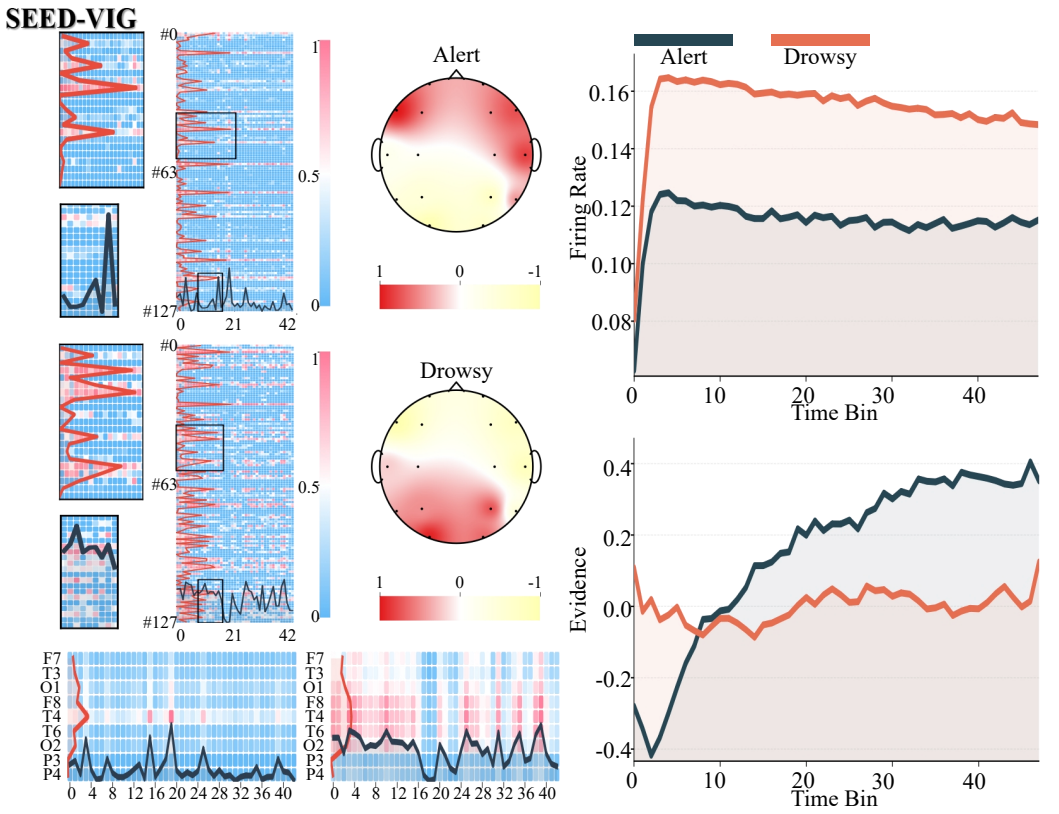

Figure 11: Visualization results of *Vibe-Spike* on SEED

Interpretability for vigilance classification. Figs. 10 and 11 provide task-specific interpretability evidence for vigilance classification. From the spike-rate perspective, the drowsy class consistently

exhibits higher firing rates than the alert class in both datasets, and the evidence trajectories progressively separate the two states over time. This indicates that *Vibe-Spike* represents vigilance as a sustained neural-state difference rather than a transient event, with drowsiness corresponding to a persistently elevated latent spiking regime.

From the latent-spike and channel-time perspectives, the two vigilance states also show qualitatively different activation styles. Drowsy samples exhibit denser and more continuously distributed spike activity, but with a comparatively blunted and less sharply structured pattern, whereas alert samples show clearer, more contrastive, and more phasic spike motifs. This difference is visible both in the latent spike maps and in the channel heatmaps, where drowsy states display broader and more persistent activation across time, while alert states retain more discrete and structured temporal responses. From the scalp-manifold perspective, both SADT and SEED-VIG show a consistent anterior–posterior contrast: alert states are associated with stronger frontal/anterior activation, whereas drowsy states show relatively stronger posterior/posterior-lateral involvement. This pattern is physiologically plausible, as alertness is more strongly supported by frontal attentional control processes, whereas drowsiness is commonly accompanied by reduced frontal vigilance control and a shift toward more diffuse low-frequency, posterior-dominant activity.[Krause et al., 2017, Thomas et al., 2000] Importantly, these rate-level, latent-spike, channel-time, and topographic patterns are highly consistent across SADT and SEED-VIG, suggesting that *Vibe-Spike* captures task-relevant vigilance dynamics rather than dataset-specific artifacts.

F.7 Stress Detection: MentalArithmetic

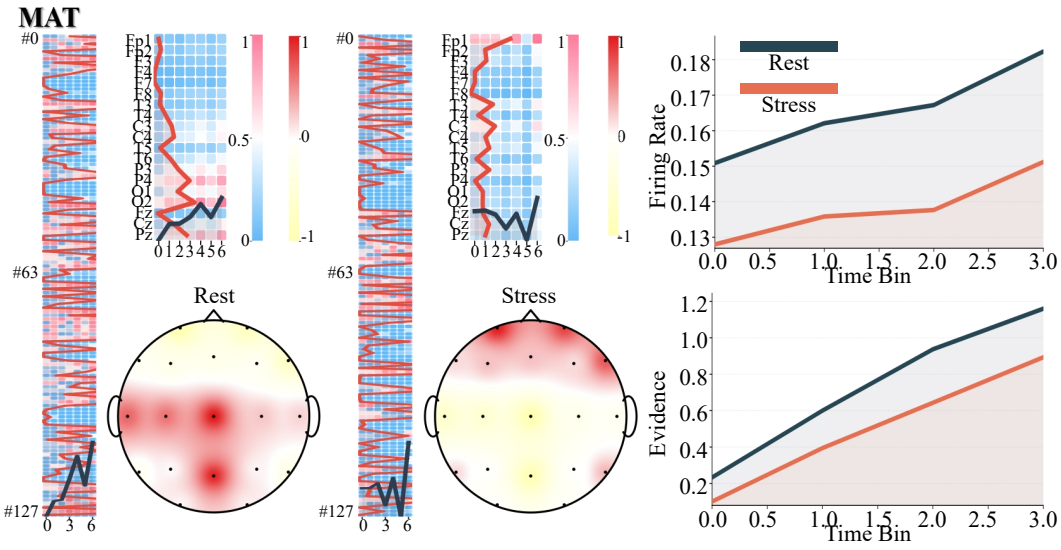

Figure 12: Visualization results of *Vibe-Spike* on MAT

Interpretability for stress recognition. Fig. 12 provides task-specific interpretability evidence for stress recognition on the MAT dataset. From the spike-rate perspective, both rest and stress samples show monotonic evidence accumulation over time, while the rest class maintains a higher firing-rate trajectory than the stress class. This suggests that *Vibe-Spike* separates the two mental states through sustained rhythmic differences rather than isolated transient responses.

From the latent-spike and channel-time perspectives, the two states exhibit clearly different activation patterns. The rest sample shows broader and stronger latent spike activity, together with more distributed channel-time activation over central and posterior channels. In contrast, the stress sample shows a more selective and anterior-dominant activation pattern, with the scalp manifold highlighting stronger frontal responses. This spatial shift is physiologically meaningful, as mental stress is commonly associated with altered frontal EEG rhythms and increased cognitive-control/arousal demands, whereas resting states often preserve stronger posterior rhythmic organization [Sugimoto et al., 2026, Halkiopoulos et al., 2026]. These spike-rate, channel-time, and scalp-topographic differences suggest

that *Vibe-Spike* encodes stress-related EEG changes as sparse, temporally accumulated, and spatially organized rhythmic spike manifolds.

F.8 EEG Event Type Classification: TUEV

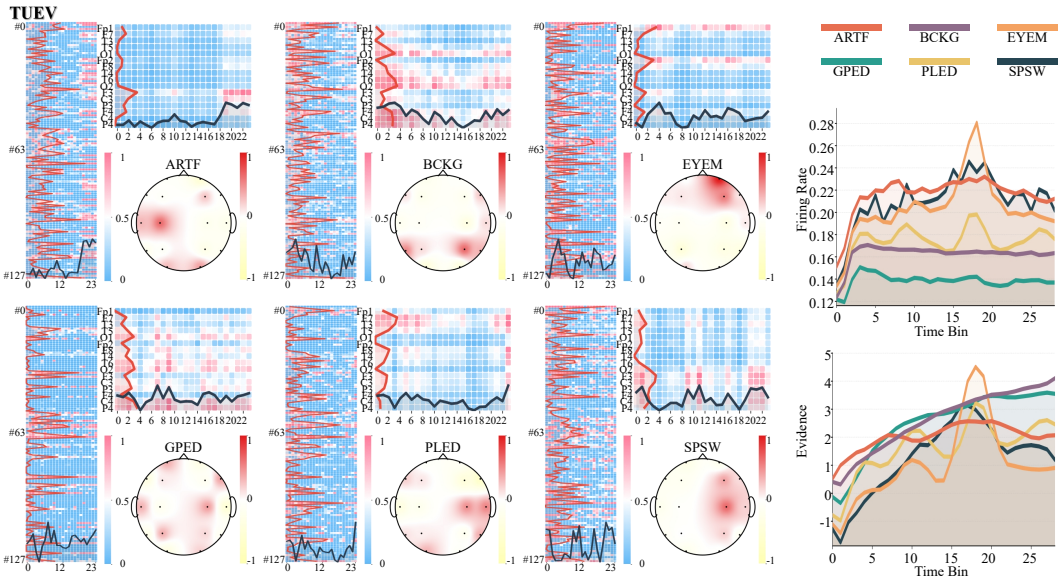

Figure 13: Visualization results of *Vibe-Spike* on TUEV

Interpretability for EEG event classification. Fig. 13 provides task-specific interpretability evidence for multi-class EEG event classification on TUEV. From the spike-rate perspective, different event categories show clearly separated temporal firing profiles. Background (BCKG) maintains the lowest and most stable firing rate, consistent with its non-event nature, whereas artifact- and epileptiform-related classes exhibit stronger and more dynamic firing responses. In particular, EYEM shows a sharp transient firing-rate peak, matching the burst-like nature of eye-movement artifacts, while SPSW, PLED, and GPED accumulate discriminative evidence over time, suggesting that *Vibe-Spike* captures temporally structured epileptiform patterns rather than isolated noisy fluctuations.

From the latent-spike, channel-time, and scalp-manifold perspectives, the learned representations also align with the semantic structure of TUEV events. EYEM shows prominent anterior/frontal activation, consistent with ocular activity being most visible over frontal leads. BCKG exhibits weak and spatially sparse activation, reflecting the absence of salient pathological or artifactual events[Mormann et al., 2007]. ARTF displays more diffuse and irregular activation patterns, compatible with non-cerebral contamination[Urigüen and Garcia-Zapirain, 2015]. For epileptiform events, GPED shows more broadly distributed activation, whereas PLED and SPSW present more focal or lateralized scalp patterns, consistent with the distinction between generalized periodic discharges, lateralized periodic discharges, and spike/sharp-wave events [Hirsch et al., 2013, Westover et al., 2015]. These class-specific firing-rate, channel-time, and topographic patterns suggest that *Vibe-Spike* learns sparse rhythmic spike manifolds that preserve the physiological and annotation-level structure of EEG events, rather than merely separating classes through dataset-specific artifacts.

G Computational Cost and Energy Consumption

We evaluate computational cost, inference latency, and device energy on HMC using a batch size of 1. Full inference uses 60 s EEG segments. Streaming updates are scheduled every 2 s.

G.1 Arithmetic Cost Estimates

We estimate arithmetic cost for the software implementation and an event-driven implementation. For *Vibe-Spike*, software FLOPs include profiler-supported operations, explicit arithmetic counts, and an analytical FFT estimate. The rhythm-gate projection and continuous gating multiplications are included. The event-driven estimate replaces eligible dense propagation operations with sparse MACs counted from the nonzero support of continuous-amplitude events.

Arithmetic energy is estimated using 4.6 pJ per FP32 MAC and 0.9 pJ per addition-like operation for a normalized 45-nm implementation [Cao et al., 2024, Horowitz, 2014]. Sine and sigmoid evaluations are included in the runtime measurements. The sparse-MAC energy is included in the event-driven total.

Table 20: Estimated arithmetic cost per full inference on 60 s HMC segments.

Model	Software FLOPs (G/op)	MAC-eq. (G/op) Software / Event-driven	Sparse event MACs (M/op)	Energy (mJ/op) Software / Event-driven	Sparse-MAC energy (μJ/op)
Vibe-Spike	**0.394**	**0.175 / 0.171**	2.360	**0.838 / 0.821**	10.854
REVE	75.838	37.919 / –	–	174.428 / –	–
BIOT	3.085	1.542 / –	–	7.095 / –	–
LaBraM	2.772	1.386 / –	–	6.376 / –	–
CBraMod	3.808	1.904 / –	–	8.759 / –	–

Table 20 shows that *Vibe-Spike* has the lowest software FLOPs and estimated arithmetic energy among the evaluated models. Its software implementation requires 0.394 GFLOPs and an estimated 0.838 mJ per inference. The event-driven estimate is 0.821 mJ per inference.

G.2 Inference Latency

CPU inference is evaluated on an Intel Xeon Gold 6459C with one PyTorch thread. GPU inference is evaluated on an NVIDIA GeForce RTX 5090. Each benchmark uses 100 warm-up iterations and 100 timed iterations, repeated over ten independent rounds. We report median and 95th-percentile (P95) latency.

The real-time factor is $\mathrm{RTF} = t_{\mathrm{median}}/T_{\mathrm{ref}}$, where latency is expressed in seconds. The reference duration is 60 s for full inference and 2 s for streaming updates.

Table 21: CPU and GPU latency for full inference on 60 s HMC segments.

	CPU latency (ms)		GPU latency (ms)		RTF	
Model	**Median**	**P95**	**Median**	**P95**	**CPU**	**GPU**
Vibe-Spike	168.850	172.542	181.624	184.305	0.002814	0.003027
REVE	619.910	637.244	6.842	6.910	0.010332	0.000114
BIOT	**23.224**	**24.634**	**1.248**	**1.280**	**0.000387**	**0.000021**
LaBraM	33.760	34.113	9.028	9.495	0.000563	0.000150
CBraMod	41.018	42.424	6.695	6.786	0.000684	0.000112

We further evaluate deployment on a Raspberry Pi 5 with 16 GB RAM. The post-hoc setting performs full inference on each 60 s segment. The streaming setting performs updates at a target interval of 2 s. Table 22 reports latency, miss rate, and RTF for both settings.

BIOT achieves the lowest full-inference latency on CPU, GPU, and Raspberry Pi. In the streaming setting, *Vibe-Spike* achieves the lowest Raspberry Pi update latency, with a median of 29.779 ms and a P95 of 44.941 ms. It completes updates within the 2 s interval with no recorded misses.

G.3 Measured Power and Energy

We measure whole-system power on the Raspberry Pi 5 using an external power meter as the board's sole power source. Each setting is evaluated over three 15 min runs. The measured idle power is $P_{\mathrm{idle}} = 2.94\,\mathrm{W}$.

Table 22: Latency of full inference and streaming updates on a Raspberry Pi 5.

Model	Setting	Median (ms/op)	P95 (ms/op)	Miss rate (%)	RTF
Vibe-Spike	Post-hoc	730.214	764.233	0.0	0.0122
REVE	Post-hoc	2681.913	2724.687	0.0	0.0447
BIOT	Post-hoc	**107.909**	**138.208**	0.0	**0.0018**
LaBraM	Post-hoc	158.847	190.737	0.0	0.0026
CBraMod	Post-hoc	199.580	233.524	0.0	0.0033
Vibe-Spike	Streaming	**29.779**	**44.941**	0.0	**0.0149**
REVE	Streaming	2872.689	3090.000	100.0	1.4363
BIOT	Streaming	107.174	137.630	0.0	0.0536
LaBraM	Streaming	163.333	194.150	0.0	0.0817
CBraMod	Streaming	208.256	242.210	0.0	0.1041

Let Q_{start} and Q_{end} denote the meter readings in Wh. For each run, energy and average system power are calculated as

$$\begin{aligned} E_{\mathrm{run}} &= 3600\,(Q_{\mathrm{end}} - Q_{\mathrm{start}})\,, \\ P_{\mathrm{system}} &= \frac{E_{\mathrm{run}}}{T_{\mathrm{run}}}, \end{aligned} \tag{15}$$

where T_{run} is the run duration in seconds. Power above idle and energy per completed operation are

$$\begin{aligned} P_{\mathrm{above\ idle}} &= P_{\mathrm{system}} - P_{\mathrm{idle}}, \\ E_{\mathrm{above\ idle/op}} &= \frac{E_{\mathrm{run}} - P_{\mathrm{idle}} T_{\mathrm{run}}}{N_{\mathrm{completed}}}. \end{aligned} \tag{16}$$

Here, $N_{\mathrm{completed}}$ counts completed full inferences in the post-hoc setting or completed updates in the streaming setting. We calculate these quantities for each run and report their means over the three runs.

Table 23: Measured system power and energy above idle on a Raspberry Pi 5.

Model	Setting	System power (W)	Power above idle (W)	Energy above idle (J/op)
Vibe-Spike	Post-hoc	**3.558**	**0.618**	**1.236**
REVE	Post-hoc	7.392	4.452	17.808
BIOT	Post-hoc	4.260	1.320	2.640
LaBraM	Post-hoc	5.862	2.922	5.844
CBraMod	Post-hoc	6.084	3.144	6.288
Vibe-Spike	Streaming	**3.420**	**0.480**	**0.960**
REVE	Streaming	6.431	3.491	10.022
BIOT	Streaming	4.350	1.410	2.820
LaBraM	Streaming	6.312	3.372	6.744
CBraMod	Streaming	4.170	1.230	2.460

Table 23 shows that *Vibe-Spike* has the lowest measured system power and energy above idle in both settings. Its energy above idle is 1.236 J per full inference and 0.960 J per streaming update. The corresponding average system powers are 3.558 and 3.420 W.

H Parameter Sensitivity Analysis

This section reports the detailed numerical results of the parameter sensitivity analysis, complementing the visualization in Fig. 4.

Table 24: Parameter sensitivity analysis on the SADT dataset.

Group	λ	τ	d	ρ_W	Bal. Acc.	AUROC	AUPRC
Default	1	1	2	0.5	0.7049 ± 0.0086	0.7734 ± 0.0082	0.5420 ± 0.0143
λ	0.1	1	2	0.5	0.7091 ± 0.0133	0.7709 ± 0.0027	0.5383 ± 0.0045
λ	10	1	2	0.5	0.7078 ± 0.0105	0.7722 ± 0.0056	0.5419 ± 0.0139
τ	1	0	2	0.5	0.7051 ± 0.0138	$\mathbf{0.7764} \pm 0.0122$	0.5493 ± 0.0176
τ	1	2	2	0.5	0.7019 ± 0.0118	0.7647 ± 0.0163	0.5411 ± 0.0127
τ	1	5	2	0.5	0.7033 ± 0.0171	0.7688 ± 0.0155	0.5472 ± 0.0306
d	1	1	1	0.5	0.7103 ± 0.0064	0.7689 ± 0.0125	0.5474 ± 0.0195
d	1	1	4	0.5	0.7076 ± 0.0127	0.7707 ± 0.0105	0.5443 ± 0.0123
d	1	1	8	0.5	$\mathbf{0.7114} \pm 0.0038$	0.7753 ± 0.0070	$\mathbf{0.5548} \pm 0.0172$
ρ_W	1	1	2	0.3	0.7085 ± 0.0136	0.7692 ± 0.0031	0.5407 ± 0.0162
ρ_W	1	1	2	0.7	0.7098 ± 0.0108	0.7678 ± 0.0043	0.5383 ± 0.0160
ρ_W	1	1	2	1.0	0.7047 ± 0.0115	0.7712 ± 0.0058	0.5442 ± 0.0149

I Data Scale and Sampling Rate Analysis

I.1 Pretraining Data Scale

The compact architecture of *Vibe-Spike* encodes EEG-specific inductive biases. Pretraining learns transferable oscillator and spiking dynamics, together with their coupling, across recording conditions. We use the full public REVE corpus to cover diverse recording protocols, channel configurations, tasks, and populations.

We evaluate three progressively larger pretraining subsets and the full corpus on CHB-MIT, SADT, and HMC. Table 25 compares their downstream performance with scratch training. The smallest subset reduces average balanced accuracy by 1.18 percentage points (pp), consistent with negative transfer from limited source coverage. The 101.42 GB and 301.90 GB subsets improve average balanced accuracy by 0.93 and 1.77 pp, respectively. The full 4.85 TB corpus achieves the highest average balanced accuracy of 0.7913, a gain of 4.43 pp over scratch training. Performance improves across the evaluated pretraining scales, supporting the use of broad data coverage for this compact model.

Table 25: Downstream balanced accuracy (mean $\pm$ std) at different pretraining scales. Avg. is the mean across datasets; Δ is its change from scratch training (pp). Bold indicates the highest means.

Pretraining data	Volume	CHB-MIT	SADT	HMC	Avg. BA (Δ, pp)
None (scratch)	0	0.7816 ± 0.0334	0.7049 ± 0.0086	0.7546 ± 0.0157	0.7470 (ref.)
REVE groups 100–104	9.64 GB	0.7538 ± 0.0357	0.6993 ± 0.0089	0.7526 ± 0.0190	0.7352 (-1.18)
REVE groups 100–110	101.42 GB	0.7927 ± 0.0290	0.7151 ± 0.0081	0.7611 ± 0.0171	0.7563 ($+0.93$)
REVE groups 100–150	301.90 GB	0.8013 ± 0.0272	0.7316 ± 0.0065	0.7614 ± 0.0185	0.7648 ($+1.77$)
All public REVE groups	4.85 TB	$\mathbf{0.8323} \pm 0.0278$	$\mathbf{0.7568} \pm 0.0020$	$\mathbf{0.7848} \pm 0.0118$	$\mathbf{0.7913}$ ($+4.43$)

I.2 Sampling Rate and Oscillator Dynamics

The released REVE signals are already sampled at 200 Hz. Our pipeline downsamples these signals once to 100 Hz. We analyze this choice from three perspectives: physiological frequency coverage and downstream performance, learned oscillator representations, and computational cost.

Physiological frequency coverage and performance. We select 100 Hz to retain the principal EEG rhythms within approximately 0.5–30 Hz. Its Nyquist frequency is 50 Hz, covering this range and part of the gamma band. Higher-frequency gamma activity provides task-dependent information, while the 50/60 Hz range is also susceptible to power-line interference. Reducing the sampling rate to

50 or 20 Hz lowers the Nyquist frequency to 25 or 10 Hz, respectively, removing progressively more of the frequency range containing these rhythms.

We evaluate 20, 50, 100, and 200 Hz inputs on five downstream datasets using ten paired random seeds. Table 26 shows that 100 and 200 Hz achieve similar average balanced accuracies of 0.7879 and 0.7880. At 50 and 20 Hz, average balanced accuracy decreases by 0.74 and 4.53 percentage points, respectively. These results support 100 Hz as a suitable sampling rate for retaining the spectral information used by the evaluated tasks.

Table 26: Balanced accuracy (mean ± std) at different input sampling rates. Avg. is the mean across datasets; Δ is its change from 100 Hz (pp). Bold indicates the highest means.

Sampling rate	CHB-MIT	ISRUC	HMC	SADT	Mumtaz2016	Avg. BA (Δ, pp)
100 Hz	$\mathbf{0.7816} \pm 0.0334$	0.7734 ± 0.0061	0.7546 ± 0.0157	$\mathbf{0.7049} \pm 0.0086$	$\mathbf{0.9248} \pm 0.0121$	0.7879 (ref.)
200 Hz	0.7808 ± 0.0335	$\mathbf{0.7744} \pm 0.0122$	$\mathbf{0.7559} \pm 0.0160$	0.7045 ± 0.0086	0.9244 ± 0.0122	**0.7880** (+0.01)
50 Hz	0.7771 ± 0.0362	0.7584 ± 0.0065	0.7476 ± 0.0220	0.6964 ± 0.0088	0.9228 ± 0.0128	0.7805 (−0.74)
20 Hz	0.6971 ± 0.0384	0.7315 ± 0.0102	0.6896 ± 0.0170	0.6938 ± 0.0115	0.9008 ± 0.0150	0.7426 (−4.53)

Learned oscillator representations. The Kuramoto module receives STFT-derived latent features. The learned time-frequency projection compresses and filters the additional high-frequency content available at 200 Hz. We assess the resulting representations on CHB-MIT using centered kernel alignment (CKA). Between 100 and 200 Hz, CKA reaches 0.9792 for the time-frequency projection, 0.9168 for oscillator phases, and 0.9017 for spike-rate representations (Table 27). These similarities indicate that the learned oscillator and spiking representations are largely preserved at 100 Hz.

At 50 and 20 Hz, oscillator-phase CKA decreases to 0.7774 and 0.6644, while spike-rate CKA decreases to 0.7436 and 0.6207. The narrower frequency range removes more of the EEG content that drives the oscillators. The resulting changes in phase and spike-rate representations are consistent with the accompanying decline in downstream performance.

Table 27: Representation similarity to the 100 Hz model on CHB-MIT, averaged over ten paired seeds. Higher CKA indicates greater similarity.

Comparison	TF-projection CKA	Oscillator-phase CKA	Spike-rate CKA
100 vs. 200 Hz	0.9792	0.9168	0.9017
100 vs. 50 Hz	0.9732	0.7774	0.7436
100 vs. 20 Hz	0.9256	0.6644	0.6207

Computational cost. Table 28 compares 100 and 200 Hz under matched configurations using 30 s EEG epochs. Increasing the sampling rate to 200 Hz doubles the raw input size and increases the number of STFT frequency bins from 65 to 129. Peak inference activation memory increases from 27.88 to 49.68 MB, a factor of 1.78. Peak training activation memory and training-step time increase by factors of 1.05 and 1.03, respectively. Inference latency remains nearly unchanged at 24.93 and 24.94 ms per sample. The 100 Hz setting reduces input and activation memory while maintaining similar performance and oscillator representations to 200 Hz.

J Limitations Results

J.1 Task-dependent limitation on motor imagery.

Tables 29 and 30 further illustrate a boundary condition of the proposed inductive bias. While *Vibe-Spike* performs well on tasks where discriminative information is expressed through sustained rhythmic states, pathological temporal dynamics, or spatially coherent brain-state variations, it is less effective on motor imagery. This should not be interpreted as a general failure of the model, but rather as a task-dependent mismatch. Motor imagery classification relies heavily on fine-grained, subject-specific sensorimotor spatial modulation, where class differences are often encoded in subtle local spatial-frequency changes rather than in global rhythmic state transitions. Therefore, a strongly

Table 28: Input size and computational cost at 100 and 200 Hz for 30 s EEG epochs.

Quantity	**100 Hz**	**200 Hz**	**Ratio (200/100)**
Raw samples per epoch	3,000	6,000	2.00
Raw input per batch (MB)	0.87	1.74	2.00
Nyquist frequency (Hz)	50	100	2.00
STFT frequency bins	65	129	1.99
STFT temporal frames	188	188	1.00
Inference latency (ms/sample)	24.93	24.94	1.00
Inference throughput (samples/s)	40.11	40.10	1.00
Inference activation peak (MB)	27.88	49.68	1.78
Training-step time (ms/batch)	451.20	464.74	1.03
Training activation peak (MB)	478.45	502.51	1.05

constrained rhythmic-dynamical backbone may reduce the flexibility needed to learn task-specific spatial projections for this type of task.

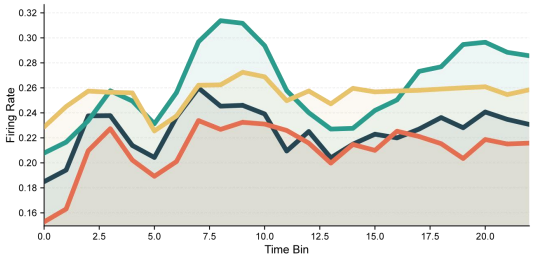

Figure 14: Spiking rate results of *Vibe-Spike* on BCIC-IV-2a

This limitation is also reflected in the visualization results. Unlike sleep staging, vigilance classification, or clinical abnormality detection, the firing-rate curves on motor imagery do not form clearly separated and stable class-specific regimes. Instead, different classes show partially overlapping and fluctuating firing trajectories, suggesting that the latent rhythmic spikes are less able to organize motor imagery classes into distinct neural manifolds. Together with the weaker quantitative performance in Table 30, these results indicate that *Vibe-Spike* is most suitable for EEG tasks dominated by rhythmically or spatially coherent neural dynamics, but may be less optimal for tasks primarily requiring highly adaptive, localized sensorimotor discrimination.

Table 29: Task-property analysis of the proposed inductive bias.

Task type	**Dominant discriminative cue**	**Expected fit to NCE bias**
Sleep staging / vigilance / seizure / AD	Sustained rhythmic or pathological temporal dynamics	High
Motor imagery	Fine-grained sensorimotor spatial modulation	Lower

Table 30: Additional results on BCIC-IV-2a motor imagery classification (4-class).

Method	Params	Balanced Accuracy	Cohen's Kappa	Weighted F1
EEGNet	0.003M	0.4482 ± 0.0094	0.2693 ± 0.0121	0.4226 ± 0.0108
EEGConformer	0.55M	0.4696 ± 0.0106	0.2924 ± 0.0141	0.4533 ± 0.0128
SPaRCNet	0.79M	0.4635 ± 0.0117	0.2847 ± 0.0147	0.4432 ± 0.0126
ContraWR	1.6M	0.4678 ± 0.0125	0.2905 ± 0.0160	0.4413 ± 0.0142
CNN-Transformer	3.2M	0.4600 ± 0.0108	0.2800 ± 0.0148	0.4460 ± 0.0114
FFCL	2.4M	0.4470 ± 0.0143	0.2627 ± 0.0176	0.4238 ± 0.0139
ST-Transformer	3.5M	0.4575 ± 0.0145	0.2733 ± 0.0198	0.4471 ± 0.0142
BIOT	3.2M	0.4748 ± 0.0093	0.2997 ± 0.0139	0.4607 ± 0.0125
LaBraM-Base	5.8M	0.4869 ± 0.0085	0.3159 ± 0.0154	0.4758 ± 0.0103
CBraMod	4.0M	0.5138 ± 0.0066	0.3518 ± 0.0094	0.4984 ± 0.0085
REVE	69M	$\mathbf{0.6396} \pm 0.0095$	$\mathbf{0.5194} \pm 0.0126$	$\mathbf{0.6339} \pm 0.0110$
Vibe-Spike	0.41M	0.4058 ± 0.0534	0.2078 ± 0.0712	0.3913 ± 0.0642

J.2 Failure Analysis

Vibe-Spike achieves a balanced accuracy of 0.5385 ± 0.0148 on TUEV (Table 2). It outperforms all task-specific baselines and BIOT but falls below LaBraM, CBraMod, and REVE. We examine class-level predictions to characterize its failure modes. Table 31 reports the corresponding performance and error destinations.

Table 31: Class-level performance and the two most frequent error destinations of *Vibe-Spike* on TUEV.

Class	Support	Precision	Recall	Main errors
SPSW	567	0.0251	0.0600	GPED, BCKG
GPED	4,677	0.6925	0.6665	BCKG, ARTF
PLED	1,998	0.4677	0.6011	GPED, BCKG
EYEM	329	0.5572	0.5775	BCKG, ARTF
ARTF	2,204	0.2869	0.6497	BCKG, GPED
BCKG	19,646	0.8810	0.7025	ARTF, PLED
Macro average	**29,421**	**0.4851**	**0.5429**	—

The clearest class-specific failure occurs in spike/sharp-wave (SPSW) recognition. SPSW recall is 0.0600, compared with 0.5775–0.7025 for the other five classes. Of 567 SPSW samples, 34 are correctly classified. The main errors are predictions of GPED (312), background activity (105), and artifacts (89). The precision of 0.0251 also indicates frequent false-positive SPSW predictions. SPSW accounts for 1.93% of the 29,421 evaluation samples.

GPED and PLED exhibit periodic temporal patterns and structured spatial distributions. SPSW events can be brief, focal, isolated, and nonperiodic. Such events offer limited temporal recurrence and cross-regional information for the model's synchronization prior. During recurrent processing, an isolated response may decay or be interpreted as background or artifact. Coincident epileptiform transients may also produce apparent phase alignment and be confused with periodic discharges. These mechanisms offer a possible explanation for the observed error destinations.

This interpretation relates to the physiological origin of EEG. Scalp EEG primarily reflects synchronized postsynaptic dendritic currents across neuronal populations. Individual axonal spikes cannot be directly resolved from these recordings. *Vibe-Spike* maps these macroscopic observations to latent neural dynamics. The Kuramoto module encodes phase synchronization, and the SNN maps this representation to membrane integration and spiking activity.

A focal SPSW event can arise from a brief hypersynchronous discharge within a spatially restricted neuronal population. Local hypersynchrony need not involve sustained propagation or coordination across distributed regions. Scalp recordings of such events provide limited observations of the underlying source dynamics. Inferring latent dynamics from these brief observations is therefore difficult. The higher recall for GPED and PLED is consistent with the model's emphasis on recurrent,

spatially organized activity. SPSW errors expose the corresponding trade-off when discriminative evidence is brief and localized.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Please check Abstract and Introduction.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: please check Discussion.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [N/A]

Justification: The paper does not include formal theoretical results, theorems, or proofs.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: We provide the dataset descriptions, preprocessing procedures, model architecture, training settings, evaluation protocols, and implementation details in Sections 3 and Appendix C D. These details are intended to support reproduction of the main experimental results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer:[Yes]

Justification: We will provide the source code, configuration files, preprocessing scripts, and training scripts needed to reproduce the main experimental results. The datasets used in this work are publicly available, and the corresponding access links and preprocessing details are described in the paper and appendix.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification:We specify the training and testing settings, including dataset splits, preprocessing procedures, model configurations, optimizer, learning rate, batch size, training epochs, evaluation metrics, and model-selection criteria in Section 3 and Appendix C D.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: We report mean $\pm$ standard deviation over repeated runs and conduct statistical significance analysis for the main comparisons. For datasets with established fixed subject-wise splits, we follow the standard protocols used in prior studies; for datasets where random subject-wise splits are commonly adopted, we repeat experiments with multiple random seeds. Thus, the reported variability reflects random initialization and data-split variability where applicable, while preserving dataset-specific evaluation conventions.

   Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer:[Yes]

   Justification: We report the computational resources used for the main experiments, including the GPU type, number of GPUs, training time, and relevant implementation settings. Please check Experiments section.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research conforms to the NeurIPS Code of Ethics. We use publicly available EEG datasets or datasets released under established research protocols, and no new human-subject data are collected in this work. The study is intended for methodological research and does not make claims of clinical deployment or replace professional medical decision-making.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer:[Yes]

Justification: The paper discusses the potential positive impacts of improving EEG representation learning, including more accessible, efficient, and interpretable analysis tools for neurological, cognitive, and clinical EEG applications. It also acknowledges potential negative impacts, including the risk of over-reliance on automated predictions and possible performance disparities across populations or recording protocols. We emphasize that the proposed method is intended for research use and should not replace professional medical decision-making without further clinical validation.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper does not release data or models with high misuse risk, such as generative models or scraped datasets. The experiments are conducted on publicly available and anonymized EEG datasets, and the proposed model is intended for methodological research rather than direct clinical deployment.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: We properly cite the original sources of the public EEG datasets, benchmark datasets, and baseline models used in this work. Dataset access links, original references, and license or usage terms are summarized in the dataset descriptions and appendix. All existing assets are used for research purposes and in accordance with their corresponding terms of use.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: We introduce a new EEG foundation model and will provide documentation and usage instructions.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This work does not conduct new crowdsourcing experiments or newly recruit human subjects. All experiments are performed on existing public EEG datasets released under their original research protocols and usage terms.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This work does not conduct new crowdsourcing experiments or newly recruit human subjects. All experiments are performed on existing public EEG datasets that were released under their original research protocols, ethical approvals, and usage terms.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs are not used as an important, original, or non-standard component of the core methodology. Any use of LLMs, if applicable, is limited to writing, editing, or formatting assistance and does not affect the scientific claims, experiments, or originality of the proposed method.

    Guidelines:

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.