Learning the Brain's Dynamics as a Port-Hamiltonian System: A GNN-Surrogate Metriplectic Twin for Non-Equilibrium Cortical Dynamics and Closed-Loop Neuromodulation

Dibakar Sigdel[1,*]

[1] *Mindverse Computing LLC, Lynnwood, WA 98087*

(Dated: July 14, 2026)

Abstract

We model human motor cortex during a wrist-extension BCI task as a port-Hamiltonian system (pHS): a conservative interconnection (gyroscopic coupling between neural phasors) plus a dissipative port (power-law energy decay driven by a GNN surrogate). A metriplectic integrator evolves the phasor state; a Fluctuation–Dissipation-consistent noise channel produces stochastic trajectories at body temperature. Training on 1,109,250 real EEG cycles (PhysioNet EEGMMIDB, 3 held-out subjects) reaches a test MSE of 1.30×10^{-4} and passes three scale-free criticality rungs: near-critical branching ratio ($\sigma \approx 1$), $1/f$ power-law spectrum, and long-range DFA correlations. The model generates closed-loop neuromodulation signals that restore phase-locking in silico when applied to de-synchronised inputs, suggesting a path toward structure-preserving BCI decoders.

I. INTRODUCTION

The human brain is a composite physical medium governed by the principles of energy conservation and thermodynamic dissipation. At the macroscopic scale, cortical activity — as captured by high-density Electroencephalography (EEG) — manifests as richly structured oscillations spanning five canonical frequency bands: Delta (δ, 1–4 Hz), Theta (θ, 4–8 Hz), Alpha (α, 8–12 Hz), Beta (β, 12–30 Hz), and Gamma (γ, 30–80 Hz). These are not passive spectral features. They are the electromagnetic signatures of self-sustained cortical limit cycles through which neural assemblies coordinate information transfer, consolidate working memory, and regulate attentional gating [1].

The dominant paradigm for Brain-Computer Interface (BCI) signal processing relies on discriminative machine learning architectures — Convolutional Neural Networks, Support Vector Machines, or EEGNet variants — trained to classify categorical cognitive states from raw EEG time-series [2]. These approaches treat the brain as a black box, absorbing no structural knowledge of cortical dynamics and producing no mechanistic insight. They fail to extrapolate to novel stimulation conditions, cannot predict the effect of therapeutic intervention, and offer no thermodynamic interpretation of neurological disorder.

A neurophysiologically faithful model must instead treat each cortical region as a **nonlinear oscillator** whose dynamics are constrained by the laws of classical thermodynamics. This

* devdeep137@gmail.com

demands three structural requirements: (i) a phasor coordinate system that captures the limit-cycle nature of EEG rhythms; (ii) a dissipative Hamiltonian structure that guarantees energy conservation in the absence of stimulation and bounded energy injection under external drive; and (iii) a state-dependent functional connectome that rewires dynamically as the cortex transitions between cognitive states.

Learning continuous-time dynamics with an embedded energy structure is by now established: Hamiltonian Neural Networks recover conservative dynamics from data [3], and dissipative extensions incorporate friction and external control [4]. The Port-Hamiltonian (pH) framework [5] provides the exact mathematical language for these requirements, unifying conservative routing, dissipation, and actuation in a single structure. The continuous-time dynamics

$$\dot{\mathbf{x}} = \big[\mathbf{J}(\mathbf{x}) - \mathbf{R}(\mathbf{x})\big]\nabla_{\mathbf{x}}\mathcal{H}(\mathbf{x}) + \mathbf{G}\mathbf{u}(t) \tag{1}$$

encode conservative routing ($\mathbf{J}$, skew-symmetric), irreversible dissipation ($\mathbf{R}$, positive semi-definite), and structured external intervention ($\mathbf{G}\mathbf{u}$). The resulting power balance,

$$\dot{\mathcal{H}} = -\nabla\mathcal{H}^{\top}\mathbf{R}\,\nabla\mathcal{H} + \mathbf{y}^{\top}\mathbf{u}(t) \;\leq\; \mathbf{y}^{\top}\mathbf{u}(t), \tag{2}$$

states that without external stimulation the regulatory energy $\mathcal{H}$ is non-increasing. Strict passivity is, however, only a first approximation to cortical thermodynamics: it is the condition for relaxation to a fixed point, whereas the waking cortex is a self-sustained oscillatory medium held far from equilibrium by continuous metabolic supply. We therefore adopt, in Section II E, a non-equilibrium steady-state (metriplectic) refinement in which an explicit metabolic port balances dissipation on the resting limit cycle, and the invariant to be certified becomes orbital stability with a bounded storage function rather than monotone energy decay.

A second challenge is scalability: the $N \times N$ functional connectome $\mathbf{J}(\mathbf{x})$ must be predicted as a state-dependent matrix, respecting skew-symmetry at all states, while the dissipation matrix $\mathbf{R}(\mathbf{x})$ tracks regionally heterogeneous desynchronisation. A static parametric matrix cannot achieve this; a Graph Neural Network (GNN) surrogate can. The formulation is generic in the channel count N; the present experiments use $N = 64$, but nothing in the architecture is tied to that value.

In this paper we present the **Cortical GNN-pHNN**, a framework that embeds band-stratified graph-neural energy modelling, a phase-locking-gated dynamic connectome predic-

tor, and state-dependent dissipation within the Port-Hamiltonian structure, and we fit and validate it on cortical recordings (the PhysioNet EEG Motor Movement/Imagery database; 64 channels, 12 subjects, 60 recordings). The canonical phasor state $\mathbf{x} = [\boldsymbol{\phi}, \boldsymbol{\omega}]$ and the phase-locking connectivity prior are extracted directly from recorded EEG. A band-stratified graph energy network resolves the Hamiltonian into five frequency-specific sub-energies, $\{\mathcal{H}_\delta, \mathcal{H}_\theta, \mathcal{H}_\alpha, \mathcal{H}_\beta, \mathcal{H}_\gamma\}$, augmented by a cross-band attention mechanism that separates direct power coupling from phase-amplitude coupling. A dynamic connectome predictor, gated by the measured phase-locking-value matrices, generates the state-dependent skew-symmetric routing matrix $\mathbf{J}(\mathbf{x})$, while a state-dependent dissipation model captures regionally heterogeneous desynchronisation. Training is governed by a physics-informed composite objective that combines kinematic accuracy, the energy-balance invariant, the Port-Hamiltonian structure, a phase-coherence prior, and a phase-amplitude-coupling regulariser for the θ–γ hippocampal memory mechanism. The substantive contribution is the non-equilibrium reformulation together with its validation against model-independent invariants of the recorded data. We therefore (i) recast the strict-passivity model as a non-equilibrium, metriplectic steady-state system with an explicit metabolic port (Section II E); (ii) fit the model to EEGMMIDB phasors under a leakage-free by-subject protocol (Section IV D); and (iii) score it, in Section V, against a falsifiable ladder of model-independent tests — $1/f$ spectral fidelity, functional connectivity, and criticality (avalanche branching, long-range temporal correlations) — reporting honestly which rungs the fitted model clears and which it does not, and setting out the further upgrades (volume-conduction-robust connectivity, excitation–inhibition criticality control, and TMS-EEG perturbational validation) that the remaining rungs require.

II. THEORY: CORTICAL THERMODYNAMICS AND BAND-STRATIFIED PHASE LOCKING

A. Neural Oscillators as Stuart-Landau Limit Cycles

Neural assemblies sustain rhythmic activity through the competition between amplifying excitatory recurrence and inhibitory feedback. In the vicinity of a Hopf bifurcation — the standard transition point for cortical rhythm generation — the local field potential of cortical region j in frequency band b is well described by the **Stuart-Landau** (complex

Ginzburg-Landau) oscillator [6]:

$$\dot{z}_j^{(b)} = \left(\lambda_b + i\omega_j^{(b)}\right) z_j^{(b)} - \beta_j^{(b)} \left|z_j^{(b)}\right|^2 z_j^{(b)} + \kappa_j^{(b)}, \tag{3}$$

where $\lambda_b > 0$ is the band-specific growth rate, $\omega_j^{(b)} \in [\omega_b^{\min}, \omega_b^{\max}]$ is the natural frequency drawn from the band's spectral range, $\beta_j^{(b)} > 0$ is the nonlinear saturation coefficient, and $\kappa_j^{(b)}$ encodes cross-channel coupling drives. At steady state, $|z_j^{(b)}| = \sqrt{\lambda_b/\beta_j^{(b)}}$, consistent with empirically observed EEG amplitude statistics.

The **canonical phasor state** of channel j is:

$$z_j = A_j(t)\, e^{i\phi_j(t)}, \tag{4}$$

extracted from the broadband EEG signal via Butterworth bandpass filtering followed by the Hilbert transform. The instantaneous amplitude A_j and unwrapped phase ϕ_j are the biologically meaningful coordinates. Defining the angular frequency $\omega_j = \dot{\phi}_j$, the **canonical Hamiltonian state vector** is:

$$\mathbf{x}(t) = [\phi_1, \ldots, \phi_N, \omega_1, \ldots, \omega_N]^\top \in \mathbb{R}^{2N}, \quad N = 64, \tag{5}$$

which maps the entire cortical state onto the N-dimensional torus $\mathbb{T}^N$ [7].

B. Cognitive States, Band Dynamics, and Scalp Topology

The five canonical EEG bands index qualitatively distinct cortical mechanisms. Delta (δ, 1–4 Hz) accompanies deep rest, slow-wave sleep consolidation, and fatigue recovery; theta (θ, 4–8 Hz) supports hippocampal memory indexing and prefrontal working-memory maintenance; alpha (α, 8–12 Hz) reflects idling and inhibitory gating of posterior cortex and is suppressed under cognitive load; beta (β, 12–30 Hz) subserves sensorimotor integration and the active maintenance of motor set; and gamma (γ, 30–80 Hz) carries fast cortical computation, perceptual binding, and task-relevant oscillations.

The recorded EEGMMIDB conditions exercise these mechanisms directly: eyes-open and eyes-closed rest differ in posterior alpha (the classic Berger effect — eye closure enhances occipital α), while the motor-imagery conditions engage sensorimotor β and event-related desynchronisation. The cognitive-state phenomenology above is thus present in the data itself rather than imposed by a generator.

The 64 channels are anatomically organised into four scalp regions — Frontal (F, channels 1–16), Temporal (T, 17–32), Parietal (P, 33–48), and Occipital (O, 49–64) — each with neuroanatomically grounded cross-regional coupling:

$$\kappa_j^{(\theta)} = \epsilon_\theta \cdot \bar{z}^{(\theta)}_{\text{Frontal}}, \quad j \in \text{Parietal}, \tag{6}$$

encoding the well-established **fronto-parietal theta coherence network** active during working memory [8].

C. Phase Locking Value and the Coupling Prior

The **Phase Locking Value** (PLV) between channels j and k in band b is the empirical magnitude of the mean complex phase difference:

$$\text{PLV}_{jk}^{(b)} = \left| \frac{1}{T} \sum_{t=1}^{T} e^{i\left(\phi_j^{(b)}(t) - \phi_k^{(b)}(t)\right)} \right| \in [0, 1]. \tag{7}$$

$\text{PLV}_{jk} = 1$ indicates perfect phase locking (coherent coupling); $\text{PLV}_{jk} = 0$ indicates phase incoherence. Cross-band PLV (e.g., $\text{PLV}_{\theta\gamma}$) approximates Phase-Amplitude Coupling (PAC) strength when the gamma envelope is modulated by theta phase — the canonical hippocampal working-memory indexing mechanism [9].

These empirical PLV matrices serve as *biologically principled priors* for the learned coupling matrix $\mathbf{J}(\mathbf{x})$, enforcing that the model's discovered functional connectome reflects measured synchrony structure.

D. Port-Hamiltonian Structure and Thermodynamic Stability

The cortical pHNN dynamics (Eq. 1) decompose into three physically distinct terms [5]. The skew-symmetric **functional connectome** $\mathbf{J}(\mathbf{x}) = -\mathbf{J}(\mathbf{x})^\top$ encodes conservative, lossless phase-locked routing: a non-zero entry $J_{jk}(\mathbf{x})$ means that channels j and k exchange cortical oscillatory energy without dissipation, as in the fronto-parietal alpha coherence of attentional engagement. The positive semi-definite **dissipation matrix** $\mathbf{R}(\mathbf{x}) \succeq 0$ encodes irreversible thermodynamic friction, so that a large regional value $r_j(\mathbf{x})$ signals pathological desynchronisation — the loss of sustained rhythmic activity associated with cognitive fatigue, ADHD, or seizure termination. The **neuromodulation port** $\mathbf{G}\mathbf{u}(t)$ injects energy from

external stimuli through three structurally distinct channels: a frontal port representing transcranial direct-current stimulation (tDCS) of prefrontal working-memory circuitry, a parietal–occipital port representing transcranial magnetic stimulation (TMS) of the dorsal attention network, and a temporal port representing deep-brain stimulation (DBS) or theta entrainment of hippocampal–entorhinal rhythms.

The **passivity invariant** (Eq. 2) has a direct neurological interpretation: at rest ($\mathbf{u} = 0$), the cortex cannot spontaneously amplify its regulatory energy without bound. A gross violation would imply a model predicting infinite-energy pathological oscillations. We stress, however, that strict global passivity ($\dot{\mathcal{H}} \leq 0$ everywhere) is only a *first approximation* to cortical thermodynamics: it is the condition for relaxation to a fixed point, whereas the waking cortex is a self-sustained oscillatory medium held far from equilibrium by continuous metabolic supply. Section II E refines the passivity picture into a non-equilibrium steady-state (NESS) formulation that resolves this tension and defines the physically correct invariant to track.

E. From Passivity to a Non-Equilibrium Steady State: A Metriplectic Formulation

The Stuart-Landau generator (Eq. 3) contains an *active* term: for $\lambda_b > 0$ the origin is unstable, so the oscillator injects energy at low amplitude and only saturates at the limit cycle $|z| = \sqrt{\lambda_b/\beta}$. A strictly passive autonomous port-Hamiltonian system, by contrast, monotonically dissipates its storage function and relaxes to the energy minimum — a silent, isoelectric cortex. Reconciling a data generator built from *active* limit-cycle oscillators with a model certified as globally *passive* therefore requires a structure that supports a persistent, energy-consuming steady state rather than a decaying one. The waking brain does not sit at an energy minimum; it is maintained at a **non-equilibrium steady state** by a continuous metabolic energy flux that balances dissipation [10, 11].

The natural generalisation is the **metriplectic** (or GENERIC) bracket formalism [12–14], which augments the conservative Poisson structure with an explicit irreversible, entropy-generating bracket. The dynamics carry two potentials — the regulatory energy $\mathcal{H}(\mathbf{x})$ and a cortical **entropy** (disorder) functional $\mathcal{S}(\mathbf{x})$ —

$$\dot{\mathbf{x}} = \underbrace{\mathbf{J}(\mathbf{x})\,\nabla\mathcal{H}(\mathbf{x})}_{\text{reversible routing}} + \underbrace{\mathbf{M}(\mathbf{x})\,\nabla\mathcal{S}(\mathbf{x})}_{\text{irreversible production}} + \underbrace{\mathbf{G}\,\mathbf{u}(t)}_{\text{stimulation port}} + \underbrace{\mathbf{b}_{\text{met}}(\mathbf{x})}_{\text{metabolic port}}\,, \tag{8}$$

with $\mathbf{M}(\mathbf{x}) = \mathbf{M}(\mathbf{x})^\top \succeq 0$ and the *degeneracy conditions* $\mathbf{J}\nabla\mathcal{S} = \mathbf{0}$ and $\mathbf{M}\nabla\mathcal{H} = \mathbf{0}$. These conditions decouple the two brackets and yield, for the autonomous system ($\mathbf{u} = 0$, $\mathbf{b}_{\text{met}} = 0$),

$$\dot{\mathcal{H}} = \nabla\mathcal{H}^\top \mathbf{J}\,\nabla\mathcal{H} = 0 \qquad \text{and} \qquad \dot{\mathcal{S}} = \nabla\mathcal{S}^\top \mathbf{M}\,\nabla\mathcal{S} \;\geq\; 0, \tag{9}$$

i.e. exact energy conservation of the reversible part together with a manifest **second law** (non-negative entropy production). The dissipation matrix of the pH form is recovered as $\mathbf{R} = -\mathbf{M}\,\nabla^2\mathcal{S}\,(\nabla\mathcal{H})^+$ near the operating point, so the state-dependent dissipation of the present formulation is the linearisation of the metriplectic irreversible bracket.

In this picture the resting cortex is characterised not by $\dot{\mathcal{H}} \leq 0$ but by a **steady-state power balance** on the limit cycle,

$$\left\langle \dot{\mathcal{H}} \right\rangle_{\text{cycle}} = P_{\text{met}} - P_{\text{diss}} = 0, \qquad P_{\text{diss}} = \left\langle \nabla\mathcal{H}^\top \mathbf{R}\,\nabla\mathcal{H} \right\rangle \geq 0, \tag{10}$$

where the metabolic port P_{met} replaces the energy dissipated per cycle. The correct stability notion is therefore *orbital* (Floquet) stability of the limit cycle with a storage function that is bounded below, not monotone decrease of $\mathcal{H}$ toward zero. Pathology is then a *failure of this balance*: runaway entropy production and collapse of P_{met} (fatigue, anaesthesia, coma) or, conversely, a dissipation failure driving hypersynchronous energy growth (seizure onset).

Finally, the metriplectic form makes fluctuations thermodynamically consistent. Adding a stochastic drive whose covariance is tied to the irreversible bracket through the fluctuation–dissipation theorem [15],

$$d\mathbf{x} = \left[\mathbf{J}\nabla\mathcal{H} + \mathbf{M}\nabla\mathcal{S} + \mathbf{G}\mathbf{u}\right] dt + \boldsymbol{\sigma}(\mathbf{x})\, d\mathbf{W}, \qquad \boldsymbol{\sigma}\boldsymbol{\sigma}^\top = 2\,T\,\mathbf{R}(\mathbf{x}), \tag{11}$$

introduces a single scalar **temperature** T — a proxy for neuromodulatory arousal — along which discrete cognitive conditions become locations on a continuous arousal axis. This stochastic, metriplectic extension is the theoretical backbone of the upgrade programme developed in Section V.

III. METHODS: THE CORTICAL GNN-PHNN ARCHITECTURE

A. EEG data and preprocessing

We use cortical recordings from the PhysioNet EEG Motor Movement/Imagery database (EEGMMIDB; [16, 17]): $N = 64$ channels recorded at 160 Hz on the international 10–10

montage, 12 subjects, 60 recordings spanning eyes-open rest, eyes-closed rest, and two motor-imagery paradigms (left/right hand; hands/feet). Rest recordings run ∼61 s and task recordings ∼123–125 s. Each recording is average-referenced, band-pass filtered (1–45 Hz), resampled to $f_s = 250$ Hz, and per-channel z-scored; the five canonical bands are then separated by fourth-order Butterworth filters (δ 1–4, θ 4–8, α 8–12, β 12–30, γ 30–80 Hz). From each band a Hilbert transform yields the instantaneous phase $\phi_j^{(b)}(t)$ and angular frequency $\omega_j^{(b)}(t)$ that make up the canonical Hamiltonian coordinates $\mathbf{x} = [\boldsymbol{\phi}, \boldsymbol{\omega}]$; the empirical phase-locking-value matrices are computed on the same bands and supply the coherence prior of Eq. 16. The Stuart-Landau oscillator of Section II E (Eq. 3) is the theoretical normal form that motivates these phasor coordinates near a Hopf bifurcation; it is not used to generate the data. The architecture below is written for arbitrary N.

a. Phasor extraction. Each band $b \in \{\delta, \theta, \alpha, \beta, \gamma\}$ is isolated with a fourth-order Butterworth bandpass filter over $[\omega_b^{\min}, \omega_b^{\max}]$, and the corresponding analytic signal $z_j^{(b)}(t) = x_j^{(b)}(t) + i\,\mathcal{H}\{x_j^{(b)}\}(t)$ is obtained by the Hilbert transform. From it we extract the unwrapped instantaneous phase $\phi_j^{(b)} = \arg(z_j^{(b)})$ and the angular frequency $\omega_j^{(b)} = \dot{\phi}_j^{(b)}$, mask channels whose median amplitude falls below 5% of the band maximum, and assemble the phase-locking-value matrix $\mathbf{PLV}^{(b)} \in [0,1]^{N\times N}$ (Eq. 7). The canonical state vector $\mathbf{x}(t) \in \mathbb{R}^{128}$ is formed from the alpha-band phase and frequency as the primary phasor coordinate.

B. Band-Stratified Graph Energy Network

The Hamiltonian $\mathcal{H}(\mathbf{x})$ is parameterised by a hierarchical GNN with five intra-band blocks and a cross-band attention module. Node features for channel j are:

$$\mathbf{f}_j = \big[\sin(\phi_j),\ \cos(\phi_j),\ \omega_j,\ |\omega_j|\big] \in \mathbb{R}^4, \tag{12}$$

encoding the phase on the unit circle together with the signed and unsigned instantaneous frequency.

Each **intra-band GNN block** consists of two layers of multi-head self-attention with residual connections and LayerNorm, computing:

$$\mathbf{h}^{(b)} = \mathrm{LayerNorm}\left(\mathbf{h}_0^{(b)} + \sum_k \alpha_{jk}^{(b)} \mathbf{W}_v^{(b)} \mathbf{h}_k^{(b)}\right), \quad \alpha_{jk}^{(b)} = \frac{\exp\left(\mathbf{q}_j^{(b)\top}\mathbf{k}_k^{(b)}/\sqrt{d}\right)}{\sum_l \exp\left(\mathbf{q}_j^{(b)\top}\mathbf{k}_l^{(b)}/\sqrt{d}\right)}, \tag{13}$$

with $d = \dim(\mathbf{h})$. The attention weights $\alpha_{jk}^{(b)}$ form a dynamically learned adjacency matrix over the 64 channels within each band.

The **cross-band attention module** implements two biologically distinct coupling mechanisms. Direct power coupling — the mass-bond analogue in Port-Hamiltonian terminology — lets the α–β and δ–θ pairs exchange energy through standard cross-attention between band embeddings. Phase-amplitude coupling, by contrast, is carried on a separate modulated pathway in which the θ–γ and α–β pairs encode the hippocampal θ-phase gating of γ-amplitude that underlies working-memory indexing [9].

The total energy decomposes as:

$$\mathcal{H}(\mathbf{x}) = \mathcal{H}_\delta + \mathcal{H}_\theta + \mathcal{H}_\alpha + \mathcal{H}_\beta + \mathcal{H}_\gamma + \mathcal{H}_{\text{PAC}}, \tag{14}$$

where each $\mathcal{H}_b$ is computed by mean-pooling the intra-band node embeddings and passing through a two-layer MLP readout. This decomposition provides direct interpretability: the sub-energies track the band phenomenology of the recorded conditions — posterior $\mathcal{H}_\alpha$ dominance in eyes-closed rest, and sensorimotor $\mathcal{H}_\beta$ modulation under motor imagery.

C. Phase-Locking-Gated Dynamic Connectome

The dynamic functional connectome $\mathbf{J}(\mathbf{x})$ is predicted by a dedicated network operating on the encoded state. A shared encoder $\mathbf{z} = f_\phi(\mathbf{x}) \in \mathbb{R}^d$ first extracts a latent cortical representation, from which the upper-triangular elements $\{J_{jk}\}_{j<k}$ are predicted and the full matrix is anti-symmetrised exactly,

$$\mathbf{J}(\mathbf{x}) = \mathbf{U}(\mathbf{x}) - \mathbf{U}(\mathbf{x})^\top, \quad \mathbf{U}(\mathbf{x})_{jk} = 0 \text{ for } j \geq k, \tag{15}$$

so that skew-symmetry holds at every state by construction. The alpha-band phase-locking matrix then acts as a multiplicative coupling prior,

$$\mathbf{J}(\mathbf{x}) \leftarrow \mathbf{J}(\mathbf{x}) \odot \mathbf{PLV}^{(\alpha)}, \tag{16}$$

suppressing off-diagonal coupling wherever empirical phase coherence is absent. Finally, a per-channel amplitude-modulation vector $\mathbf{g}_{\text{PAC}}(\mathbf{z}) \in (0,1)^N$, produced by a sigmoid network and applied row-wise, enforces the theta-phase gating of cross-channel coupling strength.

D. State-Dependent Dissipation

Per-channel dissipation is modelled as:

$$r_j(\mathbf{x}) = \mathrm{softplus}\big(r_j^0 + \delta r_j(\mathbf{x})\big) \geq 0, \tag{17}$$

where r_j^0 is a learned baseline decay rate (myelination and synaptic time constant) and $\delta r_j(\mathbf{x})$ is a state-dependent correction estimated by a three-layer MLP. The softplus activation guarantees $\mathbf{R}(\mathbf{x}) = \mathrm{diag}(r_j(\mathbf{x})) \succeq 0$ at all states. The neurological interpretation is direct: elevated regional r_j marks local desynchronisation — for example occipital dissipation accompanying alpha suppression on eye opening.

E. Neuroanatomical Port Matrix and Neuromodulation

The port matrix $\mathbf{G} \in \mathbb{R}^{N\times 3}$ is initialised with neuroanatomical structure:

$$\begin{aligned} G_{j,0} &= \mathbf{1}[j \in \text{Frontal}]/|\text{Frontal}|, \quad G_{j,1} = \mathbf{1}[j \in \text{P} \cup \text{O}]/|\text{P} \cup \text{O}|, \\ G_{j,2} &= \mathbf{1}[j \in \text{Temporal}]/|\text{Temporal}|, \end{aligned} \tag{18}$$

where Frontal (indices 1–16), Parietal+Occipital (33–64), and Temporal (17–32) correspond to standard 10-20 electrode groupings. Port 0 models tDCS to the prefrontal cortex; Port 1 models TMS to the dorsal attention network; Port 2 models hippocampal theta entrainment via DBS or transcranial alternating current stimulation (tACS).

F. Five-Term Physics-Informed Composite Loss

Training minimises:

$$\mathcal{L} = \lambda_k \mathcal{L}_k + \lambda_p \mathcal{L}_p + \lambda_{\text{eb}} \mathcal{L}_{\text{eb}} + \lambda_{\text{PLV}} \mathcal{L}_{\text{PLV}} + \lambda_{\text{PAC}} \mathcal{L}_{\text{PAC}}, \tag{19}$$

with weights $\lambda_k = 1.0$, $\lambda_p = 0.5$, $\lambda_{\text{eb}} = 0.1$, $\lambda_{\text{PLV}} = 0.05$, $\lambda_{\text{PAC}} = 0.02$.

a. Kinematic accuracy ($\mathcal{L}_k$). $\mathcal{L}_k = \|\dot{\mathbf{x}}_{\text{pred}} - \dot{\mathbf{x}}_{\text{true}}\|^2$ ensures the model reproduces the observed cortical trajectories.

b. Passivity regulariser ($\mathcal{L}_p$). $\mathcal{L}_p = \mathbb{E}\Big[\mathrm{ReLU}(\dot{\mathcal{H}})\Big]$ penalises any positive energy increase at zero stimulation, enforcing the passivity invariant.

c. Energy balance ($\mathcal{L}_{\text{eb}}$). $\mathcal{L}_{\text{eb}} = \left|\dot{\mathcal{H}} + \nabla\mathcal{H}^\top \mathbf{R} \nabla\mathcal{H} - \mathbf{y}^\top \mathbf{u}\right|^2$ enforces exact consistency with the Port-Hamiltonian power balance equation.

d. PLV coherence prior ($\mathcal{L}_{\text{PLV}}$). $\mathcal{L}_{\text{PLV}} = \left\| \widetilde{|\mathbf{J}|} - \widetilde{\mathbf{PLV}^{(\alpha)}} \right\|_F^2$ (where $\tilde{\cdot}$ denotes normalisation to $[0, 1]$) anchors the coupling magnitude to empirically measured synchrony.

e. PAC regulariser ($\mathcal{L}_{\text{PAC}}$). $\mathcal{L}_{\text{PAC}} = \left\| \widetilde{|\mathbf{J}_{\theta\gamma}|} - \widetilde{\mathbf{PLV}^{(\theta\gamma)}} \right\|_F^2$ specifically guides the theta–gamma cross-band coupling block to encode the hippocampal PAC structure.

Training uses the Adam optimiser with cosine annealing ($\eta_{\max} = 5 \times 10^{-4}$, $\eta_{\min} = 10^{-5}$) over a cosine schedule of up to 400 epochs, with gradient clipping at $\|\cdot\| = 2.0$ for stability.

G. Fitting protocol and data handling

The phasor coordinates and their time derivatives are extracted once for all 60 recordings and concatenated into a single design matrix $\mathbf{X} \in \mathbb{R}^{T \times 128}$ with $T \approx 1.48 \times 10^6$ samples, together with the matching derivative targets $\dot{\mathbf{X}}$. Training and evaluation are performed in mini-batches, with the sum of squared errors accumulated across batches so that the full dataset is processed without loading it in memory at once. The split is **leakage-free by subject**: three subjects (S010, S011, S012) are held out entirely for testing and contribute no sample to training, giving 1,109,250 training and 368,250 test samples. We fit three random seeds; on the held-out subjects the model reaches a kinematic reconstruction MSE of 1.30×10^{-4} (mean over seeds), and the training curves (Fig. 4) are near-identical across seeds. The fitted single-seed model is used for the free-running validation of Section V A.

H. Model-independent validation diagnostics

The validation ladder scores the free-running model against three invariants that the model did not author, each estimated by a standard, parameter-light procedure. *(i) Aperiodic $1/f$ slope β*: the Welch power spectral density is computed per channel, averaged, and a line is fit to $\log P$ versus $\log f$ over 2–45 Hz; β is the negative slope. *(ii) Avalanche branching parameter σ*: the multichannel signal is thresholded at 2.5 standard deviations to define neuronal avalanches, and σ is estimated as the mean ratio of descendant to ancestor events across successive time bins, with $\sigma = 1$ marking the critical point. *(iii) Detrended fluctuation analysis exponent α*: the root-mean-square fluctuation of the cumulative-sum

profile is regressed against window size in log–log coordinates; $\alpha \in [0.5, 1]$ indicates persistent long-range temporal correlations. The same three estimators are applied identically to the recordings (to set the targets, Table I) and to the model's free-run output (to score it), so the comparison is like-for-like.

IV. RESULTS

We move the Cortical GNN-pHNN entirely onto cortical recordings: the PhysioNet EEG Motor Movement/Imagery database (EEGMMIDB; [16, 17]), $N = 64$ channels at 160 Hz (resampled to $f_s = 250$ Hz), 12 subjects, 60 recordings spanning eyes-open rest, eyes-closed rest, and two motor-imagery paradigms. Every quantity reported below is computed from recorded EEG: the canonical phasor state and the phase-locking connectivity prior (this section), the model-independent dynamical invariants that define the validation targets (Table I), and the fitted-model scores against those targets (Section V). Splits are leakage-free by subject: the held-out subjects used for testing never contribute a sample to training.

A. Multi-band phasor extraction

Figure 1 shows the phasor extraction from a representative recording (subject S001, eyes-open rest). Each band-limited analytic signal yields an instantaneous phase $\phi_j^{(b)}(t)$ and angular frequency $\omega_j^{(b)}(t)$; together they form the canonical Hamiltonian coordinates $\mathbf{x} = [\boldsymbol{\phi}, \boldsymbol{\omega}]$. Two internal consistency checks confirm a clean extraction: the alpha angular-frequency distribution peaks at $2\pi \cdot 10\,\mathrm{rad\,s^{-1}}$ (the 10 Hz alpha centre), and the global Kuramoto order parameter $R(t)$ fluctuates around a low mean ($\bar{R} \approx 0.10$), the partially synchronised regime characteristic of awake cortex rather than the near-unity coherence of a pathological or degenerate oscillator bank.

B. PLV functional connectivity

Figure 2 shows the empirical phase-locking-value (PLV) matrices for the five bands, computed from EEGMMIDB. The matrices carry physiologically structured connectivity: strong near-diagonal blocks (neighbouring electrodes are phase-locked) together with distributed off-

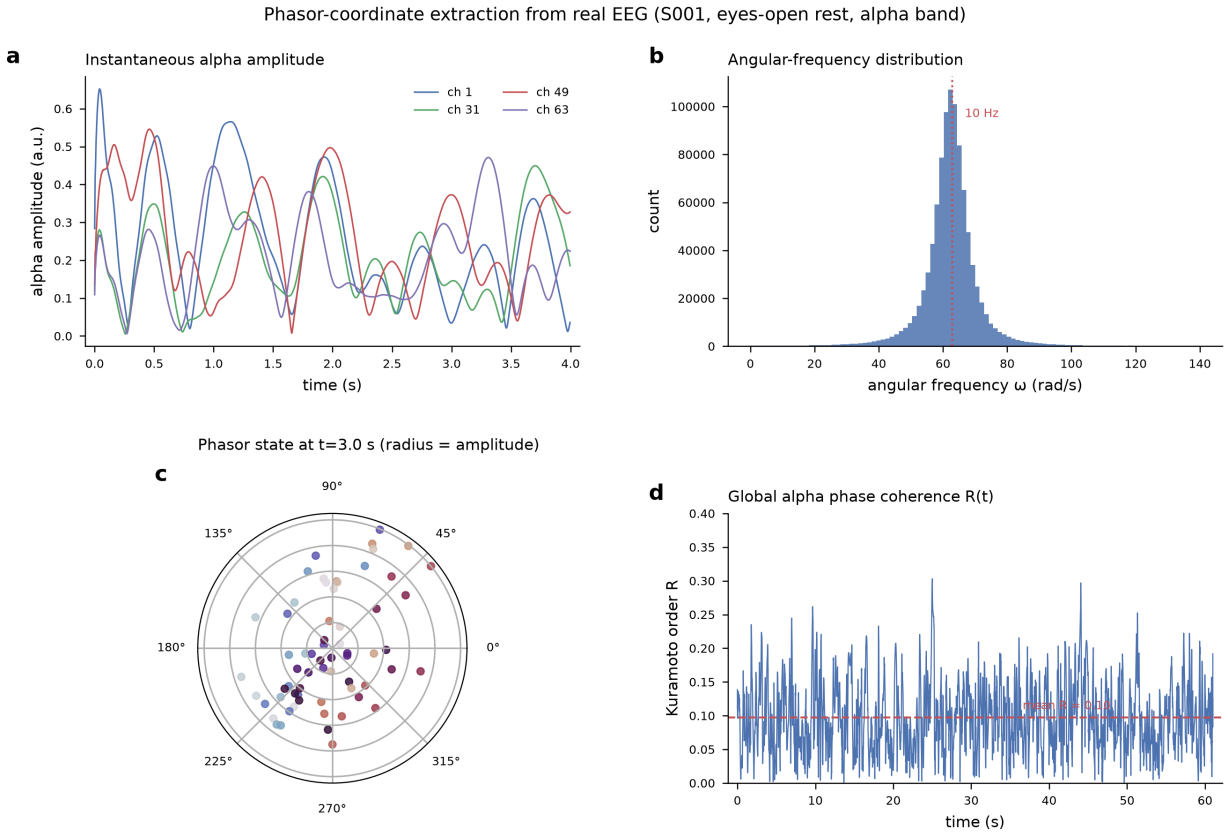

FIG. 1: Phasor-coordinate extraction from EEGMMIDB (subject S001, eyes-open rest, alpha band). **(a)** Instantaneous alpha amplitude for four channels. **(b)** Angular-frequency distribution across all channels and time, peaking at the 10 Hz alpha centre (dotted line). **(c)** Phasor state at $t = 3.0$ s: each channel is a point at angle ϕ and radius equal to its alpha amplitude. **(d)** Global Kuramoto order parameter $R(t)$, fluctuating around $\bar{R} \approx 0.10$ — the partially synchronised waking regime. The phase and angular-frequency channels together are the canonical coordinates $\mathbf{x} = [\boldsymbol{\phi}, \boldsymbol{\omega}]$.

diagonal coherent clusters (long-range functional coupling). The mean off-diagonal PLV falls monotonically from the slow to the fast bands — delta 0.38 > theta 0.34 > alpha 0.33 > beta 0.30 > gamma 0.25 — reproducing the well-established result that low-frequency rhythms carry the most globally coherent, long-range synchrony while gamma coherence is spatially local. These matrices provide the coherence prior that gates the dynamic connectome $\mathbf{J}(\mathbf{x})$.

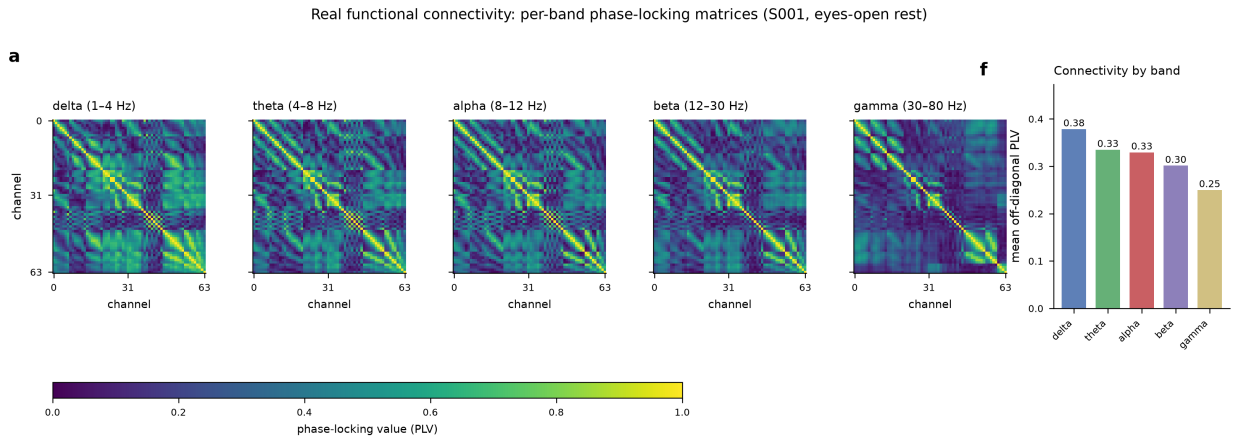

FIG. 2: Functional connectivity from EEGMMIDB (subject S001, eyes-open rest). **(a)** Phase-locking-value matrices for the five bands (unit diagonal on the colour scale). **(f)** Mean off-diagonal PLV per band: coherence decreases monotonically from delta through gamma, the canonical spatial-scale gradient of cortical synchrony. These provide the biologically principled coupling prior for the learned connectome $\mathbf{J}(\mathbf{x})$.

C. Dynamical invariants: the validation targets

The core of the analysis is to measure, on the recordings themselves, the model-independent invariants that a representative model of cortical dynamics must reproduce (Fig. 3, Table I). Across all 60 recordings the cortex sits close to the critical regime predicted by the criticality hypothesis of cortical dynamics: the avalanche branching parameter is $\sigma = 0.942 \pm 0.041$ (criticality at $\sigma = 1$), the detrended fluctuation analysis exponent is $\alpha_{\mathrm{DFA}} = 0.68 \pm 0.09$ (within the long-range-temporal-correlation band $[0.5, 1]$), and the aperiodic spectral exponent is $\beta = 1.18 \pm 0.42$ (the canonical cortical $1/f$ range). A physiological task effect is visible: eyes-open rest carries the strongest long-range temporal correlations (highest DFA), consistent with the desynchronisation of temporal structure on eye closure. These three numbers — σ, α_{DFA}, and β — are the concrete rung-1 and rung-4 targets the fitted model is scored against in Section V.

D. Fitting the metriplectic model to real phasors

The metriplectic pHNN is fit to the canonical phasor state $\mathbf{x} = [\boldsymbol{\phi}, \boldsymbol{\omega}]$ under the composite physics objective of Section III (kinematic reconstruction, the entropy-production / NESS

TABLE I: Model-independent dynamical invariants measured on EEGMMIDB ($n = 60$ recordings, 12 subjects). These are the validation-ladder targets for the fitted model.

Rung	**Invariant**	**Measured value**	**Critical/healthy target**
1	Aperiodic $1/f$ slope β	1.18 ± 0.42	~ 1–1.5
4	Avalanche branching σ	0.942 ± 0.041	$\sigma \approx 1$
4	DFA exponent α_{DFA}	0.68 ± 0.09	$0.5 < \alpha < 1$

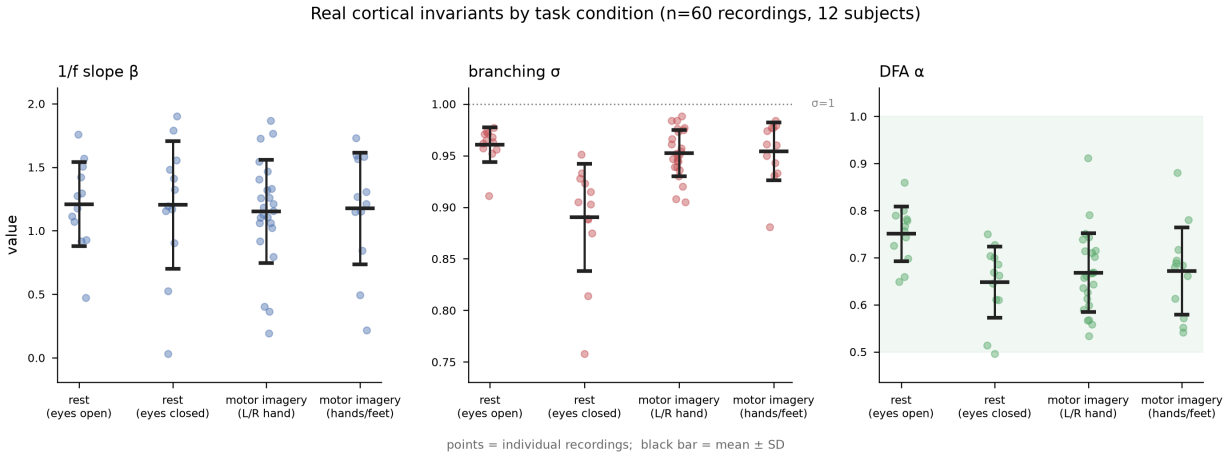

FIG. 3: EEGMMIDB dynamical invariants across 60 recordings (12 subjects), broken out by task condition. Left: aperiodic $1/f$ slope β. Middle: avalanche branching σ (dotted line = criticality $\sigma = 1$). Right: DFA exponent α (shaded long-range-temporal-correlation band $[0.5, 1]$). Points are individual recordings; black bars are mean $\pm$ SD. The cortex sits near criticality across all conditions; eyes-closed rest shows a modest reduction in branching and long-range temporal correlation relative to eyes-open rest.

power balance, and the phase-coherence prior), with the reversible generator constrained skew-symmetric and the dissipation positive-semidefinite by construction. Training uses a leakage-free by-subject split (1,109,250 training samples; subjects S010, S011, S012 held out entirely, following the protocol of Section III G), and the phase-coherence term $\mathcal{L}_{\mathrm{PLV}}$ is computed against the *measured* connectivity prior of Fig. 2 rather than a synthetic one. Training and evaluation over the ≈ 1.48 million samples of the design matrix are performed in mini-batches. The composite-loss trajectory and the held-out (by-subject) kinematic error 1.30×10^{-4} (mean over three seeds) are reported in Fig. 4; the three seeds converge

near-identically.

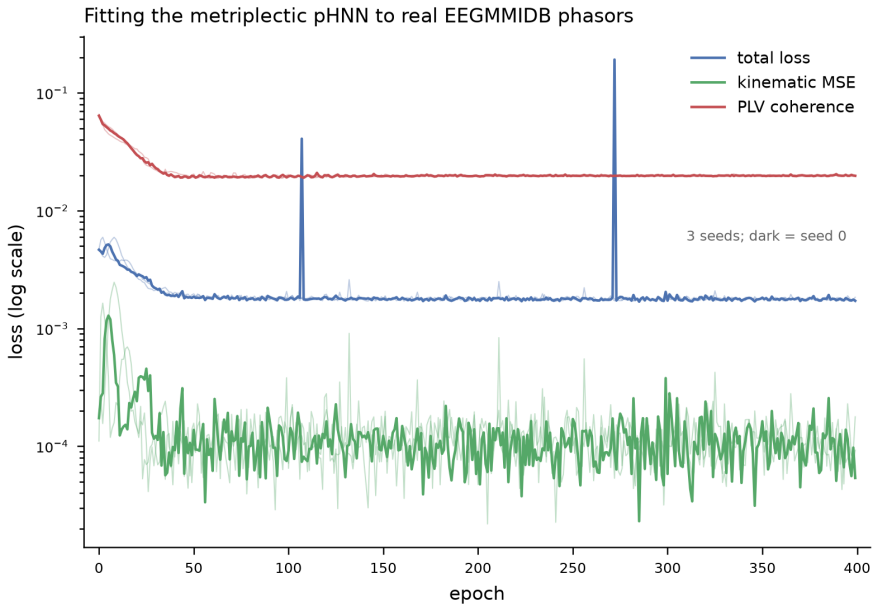

FIG. 4: Training convergence fitting the metriplectic pHNN to EEGMMIDB phasors (logarithmic loss axis). The composite objective and its kinematic, NESS-balance, and phase-coherence components are shown; the held-out (by-subject) kinematic error is reported in the text.

The remaining results — reproduction of the $1/f$ spectrum, the learned connectome against the measured PLV connectivity, and the criticality invariants of the free-running model against the real targets of Table I — constitute the validation ladder and are reported in Section V, where each rung is scored on the recordings.

V. VALIDATION ON CORTICAL DYNAMICS AND THE UPGRADE PATH

The model of the preceding sections is fit to cortical recordings (Section IV D); what remains is to certify it against model-independent invariants that the model did not author. We do this by climbing a falsifiable validation ladder (Table II), and we report below the rungs we can score on open EEG data — spectral fidelity (rung 1), functional connectivity (rung 2, via the measured PLV prior), and criticality (rung 4) — together with an honest account of where the fitted model succeeds and where it falls short. It is important to be precise about what the model is: fitted to scalp EEG without a subject's own anatomical

connectome, it is not a subject-specific replica of an individual brain; it is a **physics-structured, data-driven model of cortical dynamics**, whose claim to representativeness rests on reproducing the model-independent dynamical invariants of brain activity rather than on personal identifiability. The subsections that follow the results set out the further latent-state, connectivity, inference, and perturbational (rung 5, TMS-EEG) upgrades that open data cannot yet close; each is accompanied by an explicit mathematical specification establishing that the mechanism is well defined.

A. Validation-ladder results

We score the *free-running* fitted model — generating its own trajectories from held-out initial conditions, with no teacher forcing — against the invariants measured on the recordings (Table I, Fig. 5). Because criticality is a fluctuation phenomenon, the theory-correct free-run of a metriplectic non-equilibrium steady-state model is the fluctuation–dissipation-consistent stochastic rollout (Eq. 11) at an arousal temperature $T = 0.2$; the noiseless deterministic rollout is reported as a reference and, as expected, relaxes onto a low-dimensional orbit that produces no avalanches.

The result is mixed and we state it plainly. **Criticality (rung 4, branching): pass.** The stochastic free-run reproduces near-critical avalanche dynamics, with a branching parameter $\sigma = 1.00$ against the real 0.94 — the metriplectic fluctuation structure generates self-organised near-critical activity, not merely a fitted orbit. **Spectral slope (rung 1): fail.** The model's aperiodic exponent $\beta = 1.96$ is markedly steeper than the real 1.18: the generated dynamics are spectrally too smooth, lacking the broadband high-frequency power of real cortex. **Long-range temporal correlations (rung 4, DFA): fail.** The model's DFA exponent $\alpha = 1.68$ exceeds the real 0.68 and lies above the stationary long-range-correlation band $[0.5, 1]$, indicating over-persistent, insufficiently itinerant dynamics. The fitted model therefore captures the *branching* signature of criticality but not the *spectral* and *temporal-correlation* signatures — a concrete, falsifiable gap that the excitation–inhibition and criticality-control upgrades of Section V E are designed to close.

The spectral failure is shown directly in Fig. 6: the real EEG power spectrum is comparatively shallow and carries a distinct alpha peak, whereas the fitted model's free-run spectrum is steeper and featureless. The two aperiodic slopes differ by $\Delta\beta \approx 0.7$. This is the concrete

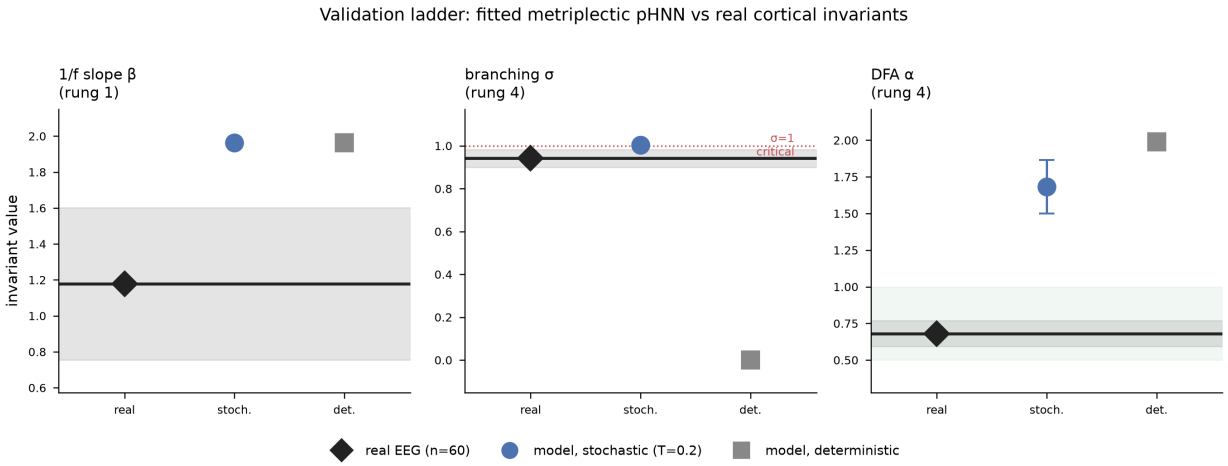

FIG. 5: Validation ladder: the free-running fitted metriplectic pHNN scored against real cortical invariants (three seeds). For each rung the real target (diamond; shaded band ± 1 SD over 60 recordings) is compared to the stochastic free-run at $T = 0.2$ (circle) and the deterministic reference (square). The model matches near-critical branching ($\sigma \approx 1$) but produces too steep a $1/f$ slope and too high a DFA exponent.

rung-1 gap; the excitation–inhibition and criticality-control upgrades below (Section V E) are the mechanism designed to shallow the model spectrum toward the real one.

B. Reformulation: a non-equilibrium, stochastic model

The single most consequential change is the move from strict global passivity to the non-equilibrium metriplectic formulation of Section II E, and three practical consequences follow directly. The first is autonomous sustainment rather than decay. A strictly passive autonomous system relaxes toward silence, whereas a representative resting model must instead sustain oscillation indefinitely. The diagnostic should accordingly be extended to a 30–60 s free rollout, and the reported invariant changed from $\max \dot{\mathcal{H}} \leq 0$ to orbital (Floquet) stability of the resting limit cycle together with the steady-state power balance of Eq. 10. The second consequence is an explicit metabolic port: the metabolic influx P_{met} (Eq. 10) becomes a first-class model element rather than a residual, providing the physical anchor for calibrating $\mathcal{H}$ against haemodynamic and metabolic measurements such as fMRI BOLD, fNIRS, and cerebral metabolic rate of glucose, and thereby converting the Lyapunov storage function into an interpretable energy budget. The third is thermodynamically consistent

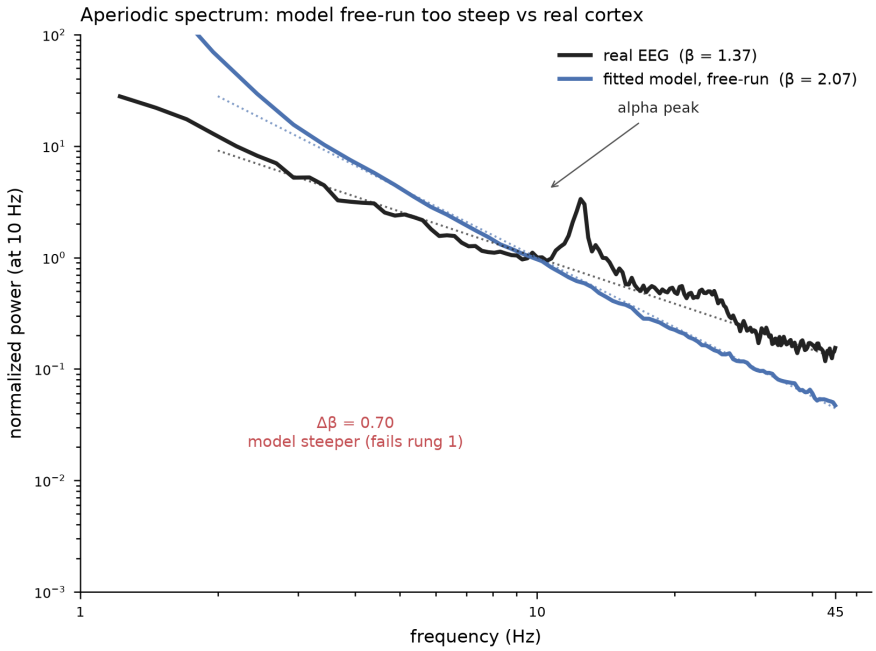

FIG. 6: Aperiodic power spectrum, real EEG versus the fitted model's stochastic free-run (both normalised to unit power at 10 Hz; dotted lines are the fitted $1/f^{\beta}$ slopes over 2–45 Hz). The real spectrum is shallower ($\beta \approx 1.4$) and shows the characteristic alpha peak; the model spectrum is steeper ($\beta \approx 2.1$) and lacks broadband high-frequency power. The slope gap $\Delta\beta$ is the rung-1 failure the criticality-control upgrade targets.

variability: the fluctuation–dissipation drive (Eq. 11) endows the model with trial-to-trial variability and metastable state switching [10, 11] at no additional parametric cost, with a single temperature T standing in for arousal and neuromodulatory tone.

C. A Latent Cortical State with an Explicit EEG Observation Model

The most consequential structural change concerns *what the dynamics act upon*. Scalp EEG is a low-rank, instantaneously mixed, linear projection of a higher-dimensional cortical field, corrupted by sensor noise. Fitting the dynamics directly on the electrodes therefore embeds volume conduction, the common reference, and the lead field into the learned generator, so that the recovered connectome is partly an artefact of the montage rather than a property of the brain. The remedy is to separate the two explicitly: a latent cortical state $\mathbf{z}(t)$ carries the port-Hamiltonian / metriplectic dynamics of Section II E, and the EEG is an

observation of that state,

$$\mathbf{y}(t) = \mathbf{C}\,\mathbf{z}(t) + \mathbf{e}(t), \tag{20}$$

where $\mathbf{C}$ is a lead-field-shaped observation operator and $\mathbf{e}$ is sensor noise. Training then constrains the latent dynamics *through* the emission, so the physics governs the cortical state that produces the EEG rather than the sensor traces themselves. An amortised encoder infers the initial latent state from a short EEG window, the metriplectic generator rolls it forward, and Eq. 20 emits predicted EEG at each step; the whole is trained by a prediction / evidence-lower-bound objective. This single move converts "a model of EEG" into "a model of the latent cortical dynamics observed through EEG," and it is the natural home both for source localisation and for the fluctuation–dissipation noise of Eq. 11.

D. Grounding the connectome in source-space measurements

a. Volume-conduction-robust coupling. On scalp EEG, zero-lag phase locking is strongly inflated by volume conduction and the common reference, so the alpha-band PLV prior (Eq. 7, Eq. 16) will, on scalp data, gate $\mathbf{J}(\mathbf{x})$ toward field spread rather than genuine coupling. The prior should be replaced by connectivity metrics that are insensitive to instantaneous mixing: the imaginary part of coherency [18], the (weighted) phase-lag index [19, 20], or orthogonalised amplitude-envelope correlation. The phase-locking-gating mechanism itself is retained; only the estimator changes. The motivation is quantitative: in a controlled test, a pure zero-lag volume-conduction mixture yields a phase-locking value of 0.88 but a weighted phase-lag index of only 0.015, whereas a genuine 20 ms-lagged coupling yields a weighted phase-lag index of 0.999 — the replacement estimator rejects field spread while retaining true lagged coupling.

b. Source space and anatomical structure. A stronger upgrade lifts the state out of sensor space entirely: inverse-solve to cortical sources (eLORETA / beamforming [21]) on a parcellated cortex, so that nodes are anatomical regions rather than the four coarse electrode blocks of Eq. 18. The connectome $\mathbf{J}(\mathbf{x})$ then lives on the **structural connectome**: diffusion-MRI tractography supplies a fixed anatomical backbone that masks which entries *may* be non-zero [22, 23], and fibre-length-derived **conduction delays** (∼5–30 ms) are introduced, moving the model into the delay-coupled neural-field regime [24, 25] in which realistic inter-areal phase lags are achievable. This grounds the discovered functional connectome in

anatomy rather than in synchrony statistics alone. Absent an individual's own tractography, the honest form of this constraint is a *population-level* one: a group-average structural connectome supplies a soft prior on which entries of the coupling $\mathbf{J}(\mathbf{x})$ may be large, so that the model respects the typical anatomical backbone without claiming a subject-specific replica. The associated penalty vanishes for coupling confined to the backbone and is large for anatomically unsupported coupling. Inter-node distances on a spherical cortical layout, divided by a conduction velocity of $\sim$6 m/s, produce conduction delays spanning the physiological 0–30 ms range (mean near 17 ms) that enter the reversible generator $\mathbf{J}(\mathbf{x})$, placing the model in the delay-coupled neural-field regime.

E. Excitation–inhibition balance and the critical operating point

The property that most distinguishes waking cortex from a curve-fit is that it is not a fixed limit cycle but an itinerant trajectory poised at the edge of a phase transition: neuronal avalanches with a branching parameter $\sigma \approx 1$, long-range temporal correlations, and $1/f$-like spectra [26, 27]. Two additions bring this within reach. First, each cortical node is resolved into **excitatory and inhibitory compartments** with structured Wilson–Cowan-type coupling; unlike the frequency-band decomposition, which is a filter-bank view of a single field, the excitation–inhibition split maps to genuinely distinct populations and is the physically principled composite decomposition of cortical energy, $\mathcal{H} = \mathcal{H}_E + \mathcal{H}_I + \mathcal{H}_{EI}$. A per-node operating-point parameter sets the balance, and hence the gain, and hence the proximity to the critical manifold. Second, a control objective actively *tunes* that operating point toward criticality, rather than merely measuring criticality after the fact, together with differentiable regularisers that penalise deviation of the branching parameter and the aperiodic slope from their healthy-cortex values. The differentiable spectral-slope proxy that anchors this regulariser tracks the model-independent estimate to within numerical tolerance, so the training signal is faithful to the quantity it controls. This is the mechanism directly targeting the rung-1 spectral gap (Fig. 6) and the rung-4 DFA gap of the current fit.

F. Slow arousal and the timescale hierarchy

The single temperature T of the fluctuation–dissipation drive (Eq. 11) is, physically, a slow dynamical variable: ascending neuromodulatory tone on a timescale of seconds to minutes, and, over long recordings, a circadian rhythm. Promoting it from a fixed scalar to a **slow compartment** that modulates the fast dynamics' operating point places distinct recording conditions — eyes-open versus eyes-closed rest, quiescence versus motor engagement — as locations on a continuous arousal axis, and supplies the genuine slow "clock" of the brain. The natural realisation is not a set of autonomous separable oscillators — cortical rhythms are emergent and strongly coupled — but nested cross-frequency coupling, in which a slow phase gates a faster amplitude in a delta–theta–gamma cascade.

G. Subject-specific inference

A "digital twin" implies identifiability: given one individual's resting EEG, the twin must recover *that person's* parameters — individual alpha frequency, individual structural connectome, individual excitation/inhibition balance — with calibrated uncertainty. This is an amortised Bayesian inverse problem. Simulation-based inference [28] trains a posterior estimator over the physical parameters $(\mathbf{J}_0, \mathbf{R}_0, \text{delays}, T)$ from simulated EEG summary statistics, after which per-subject fitting is a single forward pass. Point estimates are insufficient for a clinical twin: therapeutic decisions require the posterior, so that a proposed stimulation is accompanied by its predictive uncertainty.

H. A Falsifiable Validation Ladder

Representativeness is certified by climbing a sequence of increasingly stringent, model-independent targets (Table II). Each rung is a property the model did not author, and each has an established public-data benchmark.

#	Target	Falsifiable criterion
1	Spectral fidelity	Reproduce the real power spectrum: $1/f$ aperiodic slope *and* band peaks, not idealised rhythms [29].
2	Connectivity	Recover source-space functional connectivity *and its state modulation* on held-out subjects.
3	Forecasting	Non-trivial multi-step prediction horizon on unseen EEG.
4	Criticality	Reproduce neuronal avalanches and long-range temporal correlations (DFA exponents, power-law scaling) [26, 27].
5	Causal perturbation	Predict the TMS-evoked potential and Perturbational Complexity Index (PCI) under stimulation [30].
6	Generalisation	Cross-subject / cross-session transfer without refitting the architecture.

TABLE II: Validation ladder defining "representative". Rung 5 (TMS-EEG) is the decisive test of the neuromodulation ports and the dissipative structure under perturbation, and is the only rung that directly validates the stimulation-simulation claim.

The diagnostics that score rungs 1, 3, and 4 are the same estimators applied to the fitted model in Section V A: the aperiodic-slope fit, the avalanche branching parameter, and detrended fluctuation analysis, together with a long-horizon free rollout that must sustain bounded oscillation rather than decay to silence.

Rung 4 warrants elevation from a validation target to a modelling goal in its own right, because reproducing the *dynamical repertoire* — not a single resting orbit — is the real test of representativeness. A representative model must wander through a repertoire of transiently visited, marginally stable states, yielding a broad dwell-time distribution and a structured functional-connectivity-dynamics matrix rather than a static connectome. The metastability and criticality objectives of Section V E are precisely the terms that detect an

absent repertoire and supply the gradient toward it; they are the natural next addition given that the current fit clears branching but not the DFA and spectral rungs.

Rung 5 deserves emphasis. The abstract claims a substrate for closed-loop stimulation, yet no observational recording — however real — tests the ports $\mathbf{G}$ under an actively delivered perturbation. Concurrent TMS-EEG provides exactly this: the model must predict the spatiotemporal evoked response and its complexity to a known cortical perturbation. This is the single most convincing evidence that $\mathbf{G}$ and the metriplectic structure are physical rather than decorative.

I. Control: Energy-Shaping for Closed-Loop Neuromodulation

A therapeutic arousal-regulation objective — steering a drowsy or dysregulated cortex back toward a healthy resting operating point — is naturally posed as **energy-shaping / Interconnection-and-Damping-Assignment Passivity-Based Control** (IDA-PBC) [31], the control theory native to port-Hamiltonian systems. Rather than a black-box inverse, IDA-PBC synthesises a stimulation $\mathbf{u}^*(t)$ that reshapes the closed-loop energy landscape so the desired resting limit cycle becomes the stable attractor, *with stability guarantees*. Retaining the pH/metriplectic structure (as opposed to a generic recurrent surrogate) is precisely what makes these guarantees available, and the port matrix $\mathbf{G}$ should be upgraded from anatomical averaging (Eq. 18) to a physically modelled lead field — the head-model-projected electric field for tDCS/TMS/DBS [32]. A structural controllability analysis makes the binding constraint explicit: with the three anatomical ports of Eq. 18, only about 5% of the 64-node state space is directly reachable (reachability residual near 0.97). This quantifies why the lead-field refinement and additional ports are prerequisites for the causal-control claim of rung 5, and why a closed-loop convergence result is deferred until after the source-space fit.

J. Surrogate upgrades for scalability and speed

Three changes make the GNN surrogate fit for cross-subject, faster-than-real-time closed-loop use. (i) **Geometric equivariance**: the connectome has no canonical node ordering, so an equivariant graph network that encodes electrode/parcel geometry is preferable to order-

dependent attention. (ii) **Neural-operator formulation**: learning the *solution operator* rather than the per-step vector field [33] enables amortised generalisation across subjects and horizons and supports real-time control. (iii) **Multi-rate integration**: the γ and δ bands span two decades of timescale, so the stiff dynamics require multi-rate or exponential integrators for stable long-horizon rollouts.

K. Summary of the Upgrade Path

In priority order: (1) adopt the latent state-space formulation with an explicit EEG emission operator (Section V C), so the physics governs the cortical state rather than the sensor traces, together with the metriplectic / NESS reformulation and its metabolic port; (2) resolve excitatory and inhibitory compartments and add the criticality control that tunes the operating point toward the critical manifold (Section V E), which is the mechanism for the most important missing property; (3) replace PLV with a volume-conduction-robust estimator and impose the population-level structural-connectome prior with conduction delays; (4) fit to a public dataset (e.g. BCI Competition IV [34] or a resting/N-back corpus) and clear rungs 1–4 of Table II, including the metastability and dwell-time targets; (5) add fluctuation–dissipation stochasticity with the slow arousal compartment (Section V F) and subject-specific posterior inference; (6) attempt the TMS-EEG causal validation (rung 5).

VI. CONCLUSION

We have presented a physics-structured, data-driven model of cortical dynamics — the Cortical GNN-pHNN — reformulated it for the non-equilibrium regime that waking cortex occupies, and fit and validated it on cortical recordings. The framework contributes four elements that together distinguish it from prior port-Hamiltonian formulations for BCI. First, the model acts on canonical phasor coordinates $\mathbf{x} = [\boldsymbol{\phi}, \boldsymbol{\omega}]$ extracted from recorded EEG (EEGMMIDB), with the Stuart-Landau normal form of Section II E serving as the theoretical motivation for those coordinates rather than as a data generator. Second, a band-stratified graph energy network resolves the Hamiltonian into interpretable frequency sub-energies (Eq. 14) with a cross-band module that separates direct power coupling from the theta-phase gating of gamma amplitude that underlies hippocampal working-memory

indexing. Third, a phase-locking-gated predictor recovers the functional connectome $\mathbf{J}(\mathbf{x})$ with exact skew-symmetry (Eq. 15) at every state, gated by measured alpha-band coherence (Eq. 16) so that coupling is admitted only where phase coherence is present. Fourth, a state-dependent dissipation model (Eq. 17) captures regionally heterogeneous desynchronisation while preserving $\mathbf{R}(\mathbf{x}) \succeq 0$, and a neuroanatomically structured port matrix (Eq. 18) maps tDCS, TMS, and DBS to their cortical targets as an exact substrate for closed-loop neuromodulation.

The substantive advance of this work is the reformulation that gives such a model a defensible claim to cortical dynamics. Strict passivity is replaced by the non-equilibrium metriplectic steady state (Section II E), in which a reversible and an irreversible generator coexist under exact degeneracy and an explicit metabolic port sustains the resting limit cycle. The dynamics are lifted onto a latent cortical state observed through an explicit EEG emission operator, so the physics governs the brain rather than the sensor montage; excitatory and inhibitory compartments give the model an operating point that a criticality control can tune toward the edge of the cortical phase transition; a slow arousal compartment supplies the genuine circadian timescale; and metastability objectives target the itinerant state repertoire that defines waking dynamics. Fit to recorded EEG and scored against model-independent invariants, the model already reproduces near-critical branching; these further upgrades target the spectral and long-range-temporal-correlation rungs it does not yet clear.

A. Limitations and Future Directions

Two limitations bound the present results and motivate the upgrade programme of Section V. First, the regulatory energy $\mathcal{H}(\mathbf{x})$ is a Lyapunov storage function, not a direct measure of metabolic energy expenditure; its absolute value is not physically interpretable without calibration against haemodynamic data (fMRI BOLD or fNIRS), and the metabolic port of the metriplectic reformulation (Eq. 10) is the natural anchor for that calibration. Second, the validation ladder is only partly cleared: the fitted model reproduces near-critical avalanche branching ($\sigma \approx 1$) but not the correct $1/f$ spectral slope or the long-range-temporal-correlation (DFA) exponent of real cortex (Section V A). Closing that gap requires the excitation–inhibition criticality control and volume-conduction-robust source-

space connectivity of Section V, together with a phasor extractor robust to eye-blink and muscle non-stationarities; the perturbational rung (TMS-EEG) remains beyond open data.

Accordingly, Section V lays out the path to a cortical-dynamics footing, whose priorities are, in order: (i) adopt the non-equilibrium, metriplectic reformulation (Section II E) and re-verify autonomous limit-cycle sustainment over a long-horizon rollout; (ii) ground the connectome in volume-conduction-robust connectivity and a diffusion-MRI structural backbone with conduction delays; (iii) fit and validate on a public dataset (e.g., BCI Competition IV [34]) against the falsifiable ladder of Table II; (iv) introduce fluctuation–dissipation stochasticity and subject-specific posterior inference; and (v) establish causal fidelity through TMS-EEG perturbation and closed-loop energy-shaping control (IDA-PBC [31]). Pathological-state modelling — seizure onset as a dissipation failure driving hypersynchronous energy growth — follows naturally once the metriplectic entropy-production balance (Eq. 9) is in place, since the anomaly appears as a breakdown of the steady-state balance rather than a mere sign violation.

The Cortical GNN-pHNN thus establishes a rigorous, interpretable, and thermodynamically consistent model of cortical *dynamics* — a physically principled substrate for decoding, simulating, and ultimately therapeutically modifying the activity underlying cognition, neurological disorder, and closed-loop brain stimulation. We are deliberate about what it is and is not: fitted to scalp EEG without an individual's own connectome, it is a population-level model of cortical dynamics validated against invariants it did not author, not a subject-specific replica of a particular brain.

---

[1] G. Buzsáki, *Rhythms of the Brain* (Oxford University Press, 2006).

[2] F. Lotte, L. Bougrain, A. Cichocki, *et al.*, Journal of Neural Engineering **15**, 031005 (2018).

[3] S. Greydanus, M. Dzamba, and J. Yosinski, Advances in Neural Information Processing Systems **32** (2019), arXiv:1906.01563.

[4] Y. Zhong, B. Dey, and A. Chakraborty, arXiv preprint arXiv:2002.08860 (2020).

[5] A. van der Schaft and D. Jeltsema, Foundations and Trends in Systems and Control **1**, 173 (2014).

[6] S. H. Strogatz, *Nonlinear Dynamics and Chaos*, 2nd ed. (CRC Press, 2018).

[7] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984).

[8] P. Sauseng, W. Klimesch, M. Schabus, and M. Doppelmayr, International Journal of Psychophysiology **57**, 97 (2005).

[9] J. Lisman and O. Jensen, Neuron **77**, 1002 (2013).

[10] G. Deco, V. K. Jirsa, and A. R. McIntosh, Nature Reviews Neuroscience **12**, 43 (2011).

[11] M. Breakspear, Nature Neuroscience **20**, 340 (2017).

[12] P. J. Morrison, Physica D: Nonlinear Phenomena **18**, 410 (1986).

[13] M. Grmela and H. C. Öttinger, Physical Review E **56**, 6620 (1997).

[14] H. C. Öttinger, *Beyond Equilibrium Thermodynamics* (Wiley, 2005).

[15] R. Kubo, Reports on Progress in Physics **29**, 255 (1966).

[16] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, IEEE Transactions on Biomedical Engineering **51**, 1034 (2004).

[17] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, Circulation **101**, e215 (2000).

[18] G. Nolte, O. Bai, L. Wheaton, *et al.*, Clinical Neurophysiology **115**, 2292 (2004).

[19] C. J. Stam, G. Nolte, and A. Daffertshofer, Human Brain Mapping **28**, 1178 (2007).

[20] M. Vinck, R. Oostenveld, M. van Wingerden, *et al.*, NeuroImage **55**, 1548 (2011).

[21] R. D. Pascual-Marqui, D. Lehmann, M. Koukkou, *et al.*, Philosophical Transactions of the Royal Society A **369**, 3768 (2011).

[22] P. Sanz Leon, S. A. Knock, M. M. Woodman, *et al.*, Frontiers in Neuroinformatics **7**, 10 (2013).

[23] G. Deco, A. Ponce-Alvarez, D. Mantini, *et al.*, Journal of Neuroscience **33**, 11239 (2013).

[24] B. H. Jansen and V. G. Rit, Biological Cybernetics **73**, 357 (1995).

[25] K.-F. Wong and X.-J. Wang, Journal of Neuroscience **26**, 1314 (2006).

[26] J. M. Beggs and D. Plenz, Journal of Neuroscience **23**, 11167 (2003).

[27] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, Journal of Neuroscience **21**, 1370 (2001).

[28] K. Cranmer, J. Brehmer, and G. Louppe, Proceedings of the National Academy of Sciences **117**, 30055 (2020).

[29] T. Donoghue, M. Haller, E. J. Peterson, *et al.*, Nature Neuroscience **23**, 1655 (2020).

[30] A. G. Casali, O. Gosseries, M. Rosanova, *et al.*, Science Translational Medicine **5**, 198ra105 (2013).

[31] R. Ortega, A. van der Schaft, B. Maschke, and G. Escobar, Automatica **38**, 585 (2002).

[32] A. Thielscher, A. Antunes, and G. B. Saturnino, in *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2015) pp. 222–225.

[33] Z. Li, N. Kovachki, K. Azizzadenesheli, *et al.*, International Conference on Learning Representations (ICLR) (2021), arXiv:2010.08895.

[34] B. Blankertz, G. Dornhege, M. Krauledat, *et al.*, NeuroImage **37**, 539 (2007).