HIGH-RESOLUTION DYNAMIC FUNCTIONAL CONNECTIVITY GENERATION WITH GRAPH-VARIATE FLOW MATCHING

Anonymous authors
Paper under double-blind review

ABSTRACT

High-resolution dynamic functional connectivity (DFC) is attractive for studying rapidly evolving brain-network interactions, but it is difficult to estimate and generate reliably because short temporal windows produce noisy and often low-rank covariance estimates. Graph-Variate Dynamic (GVD) connectivity addresses this problem by modulating fast instantaneous interactions with a stable trial-level support, suppressing spurious fluctuations while emphasizing persistent and informative connections. We further show that this Hadamard construction lifts low-rank instantaneous connectivity from the positive-semidefinite to the positive-definite cone, enabling high-resolution connectivity trajectories to remain on the SPD manifold without additive ridge regularisation or post-hoc projection. Building on this structure, we introduce GVD-CFM, a class-conditional generative model designed specifically for high-resolution dynamic connectivity generation. Each trial is represented as a sequence of SPD GVD matrices on a product Riemannian manifold and mapped through a global log-Euclidean diffeomorphism and an invertible temporal DCT basis. A Transformer-based conditional flow models all spectral modes jointly, allowing the complete high-resolution trajectory to be generated non-autoregressively in Euclidean coordinates while preserving exact correspondence with valid SPD connectivity sequences. Because the full DCT basis is retained, the learned representation also defines a temporal basis expansion that can be decoded on denser temporal grids without retraining. Across multiple EEG motor-imagery datasets, GVD-CFM achieves the strongest overall performance across held-out distributional fidelity, preservation of temporal dynamics, and synthetic-to-real classification, while remaining computationally efficient relative to strong raw-signal and direct GVD-space generative baselines. These results establish GVD-CFM as a framework for generating realistic, temporally coherent, high-resolution brain-network trajectories while preserving manifold structure and supporting resolution-flexible decoding from a single trained model.

1 INTRODUCTION

Dynamic functional connectivity (DFC) has been a topic of major interest in neuroscience in recent years (Hutchison et al., 2013; Preti et al., 2017). Brain-network connectivity is central to many neuroscientific studies (Fox et al., 2014; Allen et al., 2014; Roy et al., 2023; 2024), and EEG, with its very high temporal resolution (Pfurtscheller and da Silva, 1999; Smith et al., 2017), should in principle allow these networks to be followed at the time scale of cognitive events. In practice, connectivity estimated over short temporal windows is dominated by noise and spurious correlations (Leonardi and Van De Ville, 2015; Hindriks et al., 2016). Graph-variate dynamic (GVD) connectivity addresses this problem by modulating instantaneous interactions with a stable, trial-level connectivity matrix through a Hadamard product (Smith et al., 2019; Roy et al., 2025). This has been shown to give reliable time-varying dynamics at high temporal resolution, for example in the detection of connectivity changes during event-related potentials (Roy et al., 2023; 2024).

In the machine learning community, the study of connectivity has been largely limited to the static case (Sihag et al., 2022; Cavallo et al., 2024; Roy et al., 2026), and most progress in architectures

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

and generative models for EEG concerns the raw signal (Lotte et al., 2018; Lawhern et al., 2018; Schirrmeister et al., 2017; Barachant et al., 2012; Kobler et al., 2022; Hartmann et al., 2018). Among recent generative models (Goodfellow et al., 2014; Kingma and Welling, 2014; Song et al., 2021b; Ho et al., 2020), flow matching has emerged as an efficient way of training continuous normalizing flows (Chen et al., 2018; Lipman et al., 2023; Liu et al., 2023; Albergo and Vanden-Eijnden, 2023), and conditional flow matching makes its objective tractable (Tong et al., 2024; Lipman et al., 2024). Connectivity matrices are symmetric positive definite (SPD), and Riemannian geometry provides a natural setting for their study (Pennec et al., 2006; Arsigny et al., 2007; Pennec, 2006; Thanwerdas, 2022). This has led to Riemannian flow matching (Chen and Lipman, 2024) and, to avoid its computational cost, to diffeomorphic pullback flow matching, which performs Euclidean flow matching through a global diffeomorphism and is equivalent to the Riemannian process (Collas et al., 2025). These methods require every matrix to be SPD. A covariance matrix estimated from a window with fewer samples than channels is rank deficient, so the geometry that makes these methods tractable is not available for high-resolution DFC.

Here, we propose GVD-CFM, a spectral flow matching model that generates complete high-resolution EEG connectivity trajectories. We first show that the Hadamard product with a positive-definite support lifts rank-deficient window covariances to SPD matrices, so that a GVD trajectory is a point on a product Riemannian manifold. We map this manifold to Euclidean coordinates with a log-Euclidean chart followed by an orthonormal temporal DCT, and train a Transformer to predict the velocity of all DCT modes jointly. The DCT approximately decorrelates the temporal covariance of the trajectory, which simplifies the transport problem, and the generated coefficients can be decoded on finer temporal grids than the one used in training. We further show that current generative models for raw EEG may produce realistic signals but fail to reproduce the temporal dynamics of connectivity. Since GVD-CFM generates connectivity trajectories rather than raw EEG, its samples are directly useful to models that consume connectivity representations, a setting where labeled data are scarce (Lashgari et al., 2020; Roy et al., 2019; Jiang et al., 2024).

Contributions.

1. **High-resolution DFC that overcomes the covariance rank issue.** We show that GVD connectivity is SPD even when window covariances are rank deficient, provided the stable support is SPD and every channel has nonzero energy in the window. This gives a valid log-Euclidean representation when a window contains fewer samples than channels allowing high resolution temporal precision.
2. **Non-autoregressive spectral generation of complete trajectories.** GVD-CFM performs conditional flow matching on complete log-Euclidean GVD trajectories in an orthonormal DCT basis, exactly equivalent to Riemannian flow matching on the product SPD manifold. The DCT approximately diagonalizes the optimal Gaussian transport field, and the generated coefficients can be evaluated on denser temporal grids without retraining.
3. **Controlled evaluation and physiological plausibility.** We compare against four raw-EEG generators and three GVD-space controls that share the GVD targets and decoding pipeline, ablate the spectral representation and the stable support, and show that generated trajectories preserve the regional organization and time course of real motor-imagery connectivity.

2 Related Work

Dynamic functional connectivity. DFC is commonly estimated with sliding-window correlation (Hutchison et al., 2013; Allen et al., 2014; Preti et al., 2017), but short windows increase noise and can produce spurious fluctuations (Leonardi and Van De Ville, 2015; Hindriks et al., 2016). Regularized approaches improve stability through shrinkage or temporal smoothness (Ledoit and Wolf, 2004; Chen et al., 2010; Monti et al., 2014; Hallac et al., 2017), usually at the cost of temporal resolution. Graph-variate methods instead filter instantaneous interactions using a stable trial-level connectivity matrix, enabling high-resolution EEG connectivity estimates (Smith et al., 2017; 2019; Roy et al., 2023; 2024; 2025).

SPD geometry and generative modeling. EEG covariance matrices lie on the SPD manifold, motivating geometry-aware methods based on affine-invariant, log-Euclidean, and log-Cholesky

metrics (Pennec et al., 2006; Arsigny et al., 2007; Lin, 2019). These ideas have also been used in neural models for covariance data (Kobler et al., 2022; Sihag et al., 2022; Cavallo et al., 2024; Ju et al., 2025; Roy et al., 2026). Generative models on manifolds include Riemannian diffusion and flow matching (De Bortoli et al., 2022; Jo and Hwang, 2023; Chen and Lipman, 2024), while SPD-specific work has focused mainly on generating single matrices (Li et al., 2024; de Surrel et al., 2025; Marti, 2020). Diffeo-CFM (Collas et al., 2025) is closest to our approach, but considers static, full-rank connectivity matrices.

Graph, time-series and EEG generation. Graph generators mainly model static graphs (Vignac et al., 2023; Qin et al., 2025; Huang and Ruan, 2025; Williams, 2025), while time-series models operate in Euclidean spaces (Esteban et al., 2017; Rasul et al., 2021; ten Brinke et al., 2026). EEG generation has largely focused on synthesizing raw signals using GANs or denoising models (Hartmann et al., 2018; Luo and Lu, 2018; Wang et al., 2026). In contrast, GVD-CFM directly generates high-temporal-resolution connectivity trajectories rather than raw EEG.

3 Background

Notation. $\tau \in [0,1]$ denotes flow time and t EEG sample time. A trial has T samples on d channels and is divided into B temporal windows. $\mathbb{S}^d$ and $\mathbb{S}^d_{++}$ denote the symmetric and SPD $d \times d$ matrices, $\odot$ and $\oslash$ the Hadamard product and division, and $m = d(d+1)/2$. $\mathrm{svec} : \mathbb{S}^d \to \mathbb{R}^m$ is lower-triangular vectorization with off-diagonal entries scaled by $\sqrt{2}$, which preserves the Frobenius inner product.

Conditional flow matching. Flow matching (Lipman et al., 2023) learns a velocity field $v_\theta : [0,1] \times \mathbb{R}^D \to \mathbb{R}^D$ that transports a prior p_0 to a data distribution p_1. For a coupling $\pi(z_0, z_1)$, the linear path $z_\tau = (1-\tau)z_0 + \tau z_1$ has constant velocity $z_1 - z_0$, and the conditional objective

$$\mathcal{L}_{\mathrm{CFM}}(\theta) = \mathbb{E}_{\tau,\,(z_0,z_1)\sim\pi} \left\| v_\theta(\tau, z_\tau) - (z_1 - z_0) \right\|^2 \tag{1}$$

has the same gradient as the intractable marginal objective. An independent coupling gives the standard method; a minibatch optimal-transport coupling straightens the marginal paths (Villani, 2009; Tong et al., 2024).

Riemannian flow matching by diffeomorphic pullback. Riemannian flow matching (Chen and Lipman, 2024) regresses onto geodesic velocities, at the cost of computing geodesics and Riemannian norms. Collas et al. (2025) observed that this cost disappears whenever $\mathcal{M}$ admits a global diffeomorphism $\varphi : \mathcal{M} \to E$ onto a Euclidean space; under the pullback metric $\varphi^* g_E$, φ is an isometry and geodesics are pulled-back straight lines.

Proposition 1 (Pullback reduction; Collas et al., 2025). *On $(\mathcal{M}, \varphi^* g_E)$ the Riemannian CFM objective equals the Euclidean objective of Equation 1 on $z = \varphi(x)$. Integrating in E and decoding by φ^{-1} gives exactly the samples obtained by integrating on $\mathcal{M}$, and every sample lies on $\mathcal{M}$.*

4 Methods

4.1 Graph-Variate Dynamic Connectivity

Graph-variate signal analysis (GVSA) represents a multivariate time series through evolving interactions on a stable support (Smith et al., 2019). Let $u_t \in \mathbb{R}^d$ be the channel-standardized EEG sample at time t, $U = [u_1, \ldots, u_T] \in \mathbb{R}^{d \times T}$, and $W = \frac{1}{T} U U^\top$ the whole-trial correlation matrix. We use the signed outer product $J_t = u_t u_t^\top$ as the instantaneous interaction and define $\Delta_t = W \odot J_t$. Unlike the original GVSA definition, we retain the sign of both the support and the instantaneous interaction, and we retain the main diagonal, since both are needed for the positive-definite geometry of our model (Appendix B.1). When the stable support is computed from the signal itself, this is graph-variate dynamic (GVD) connectivity (Smith et al., 2019; Roy et al., 2024; 2025).

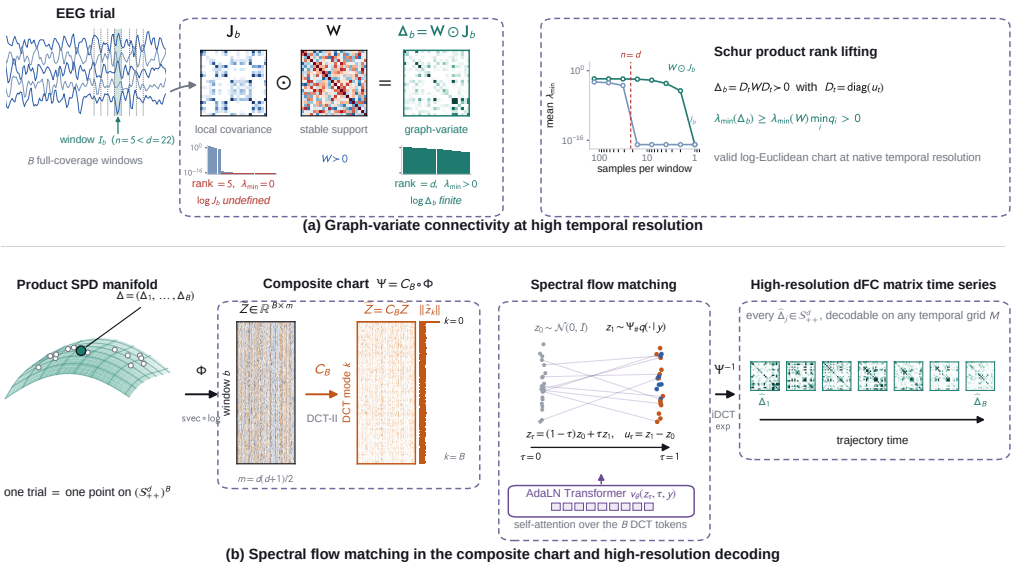


Figure 1: **Overview of GVD-CFM. (a)** When a window has fewer samples than channels ($n < d$), the window covariance J_b is rank deficient and $\log J_b$ is undefined. With a positive-definite support W, $\Delta_b = W \odot J_b$ is positive definite with $\lambda_{\min}(\Delta_b) \geq \lambda_{\min}(W) \min_i q_i > 0$, so the log-Euclidean chart remains valid at native temporal resolution. **(b)** A trial is a point on $(\mathbb{S}^d_{++})^B$. The chart Ψ applies $z_b = \mathrm{svec} \log \Delta_b$, featurewise standardization and the orthonormal temporal DCT-II; it is a global diffeomorphism and an isometry for its pullback metric (Remark 1). Flow matching along $z_\tau = (1-\tau)z_0 + \tau z_1$ is performed in these coordinates by a Transformer over DCT modes, and decoding through Ψ^{-1} returns a trajectory in which every $\widehat{\Delta}_b$ lies on $\mathbb{S}^d_{++}$.

We divide the T samples into B disjoint, full-coverage windows $\{\mathcal{I}_b\}_{b=1}^B$. Individual samples are noisy and give prohibitively long sequences, so a trial is represented by the window averages

$$\overline{\Delta}_b = \frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} \Delta_t = W \odot J_b, \qquad J_b = \frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} u_t u_t^\top. \tag{2}$$

Setting $B = T$ recovers sample-resolution connectivity, while smaller B trades temporal resolution for lower variance. The window covariance J_b has rank at most $\min(d, |\mathcal{I}_b|)$, so it is singular whenever a window contains fewer samples than channels, and $\log J_b$ does not exist. The Hadamard product with the support removes this barrier.

Proposition 2 (GVD lifts rank-deficient covariance)**.** *Let $W \in \mathbb{S}^d_{++}$. If every channel has nonzero energy in window b, $q_i = \frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} u_{i,t}^2 > 0$ for $i = 1, \ldots, d$, then $\overline{\Delta}_b = W \odot J_b$ is SPD, whatever the rank of J_b, and*

$$\lambda_{\min}(\overline{\Delta}_b) \geq \lambda_{\min}(W) \min_i q_i > 0. \tag{3}$$

In particular, a single sample with rank-one J_t gives $\Delta_t \in \mathbb{S}^d_{++}$ whenever every component of u_t is nonzero. A GVD trajectory $\overline{\boldsymbol{\Delta}} = (\overline{\Delta}_1, \ldots, \overline{\Delta}_B)$ therefore lies on the product Riemannian manifold $\mathcal{M}_B = (\mathbb{S}^d_{++})^B$.

The proof uses $W \odot u_t u_t^\top = D_t W D_t$ with $D_t = \mathrm{diag}(u_t)$ and is given in Appendix B.2. The result requires $W \succ 0$, which a sample correlation matrix need not satisfy. Here $\mathrm{rank}(W) = \mathrm{rank}(U)$, which is full whenever the $T \geq 512$ samples span $\mathbb{R}^d$ ($d \leq 30$). We verify strict positive definiteness of every support and window matrix in double precision, and no trial failed this check.

4.2 GVD-CFM

We propose *Graph-Variate Dynamic Conditional Flow Matching* (GVD-CFM). Essentially, GVD-CFM maps each trajectory to Euclidean coordinates with a global diffeomorphism, organizes these coordinates by temporal frequency, and learns a flow over the complete trajectory at once (Figure 1).

4.2.1 Composite Diffeomorphism

We apply the log-Euclidean diffeomorphism (Pennec et al., 2006; Arsigny et al., 2007) to each window, $z_b = \mathrm{svec}(\log \overline{\Delta}_b) \in \mathbb{R}^m$, and stack the results into $Z = \Phi(\overline{\mathbf{\Delta}}) = [z_1, \dots, z_B]^\top \in \mathbb{R}^{B\times m}$. We standardize each feature with training-set statistics, $\widetilde{Z} = (Z - \mathbf{1}\mu^\top) \oslash \mathbf{1}\sigma^\top$, and apply the orthonormal DCT-II matrix C_B (Ahmed et al., 1974; Strang, 1999), $C_B C_B^\top = I_B$, along the temporal axis. The complete chart is

$$\Psi : \left(\mathbb{S}^d_{++}\right)^B \longrightarrow \mathbb{R}^{B\times m}, \qquad \Psi(\overline{\mathbf{\Delta}}) = \bar{Z} = C_B \widetilde{Z}. \tag{4}$$

The matrix logarithm, svec, the standardization and the orthonormal DCT are all globally invertible, so Ψ is a diffeomorphism. Its inverse applies $C_B^\top$, reverses the standardization and exponentiates each window (Appendix A.16). GVD-CFM can therefore be trained in Euclidean coordinates while remaining in one-to-one correspondence with trajectories on the product SPD manifold. For the unmodulated window covariance this chart does not exist when $|\mathcal{I}_b| < d$.

Remark 1 (Standardization and the metric). Standardization is not an isometry of the log-Euclidean metric. The chart $C_B \circ \Phi$ is an isometry for the product log-Euclidean metric g_{LE}, while Ψ is an isometry for its pullback metric $g_\sigma = \Psi^* g_E$, the constant reweighting $\langle \xi, \eta \rangle_\sigma = \sum_{b,j} \xi_{bj}\eta_{bj}/\sigma_j^2$ of g_{LE} in log coordinates. The two metrics have the same geodesics, and the two flow-matching objectives differ only by a fixed weighting of the velocity residual and share the same population minimizer. Our equivalence statements refer to g_σ and reduce to g_{LE} when $\sigma \equiv 1$ (Corollary 1).

4.2.2 Spectral Velocity Network

The DCT organizes temporal variation by frequency (Strang, 1999; Shuman et al., 2013): persistent structure concentrates in low-order modes and faster changes in higher-order modes. All B modes are retained, so no temporal resolution is discarded. Each row $\bar{z}_k$ of $\bar{Z}$ is a Transformer token for DCT mode k (Vaswani et al., 2017), with hidden state given by a learned projection of $\bar{z}_k$ plus an embedding of the mode index. The model is class-conditional through $c_\tau = e_{\mathrm{flow}}(\tau) + e_{\mathrm{class}}(y)$, a Fourier encoding of flow time (Tancik et al., 2020) plus a learned class embedding, which modulates adaptive LayerNorm (AdaLN) Transformer blocks (Ba et al., 2016; Peebles and Xie, 2023). Self-attention acts across all DCT modes and gives hidden states $h_D \in \mathbb{R}^{B\times h}$, and an output projection gives the joint velocity $v_\theta : \mathbb{R}^{B\times m} \times [0,1] \times \mathcal{Y} \to \mathbb{R}^{B\times m}$ of the complete trajectory in a single forward pass.

Temporal branch. The amplitude of one window is spread across all B modes, while the decoder exponentiates each window separately. Errors that are unbiased in log coordinates therefore inflate the power of decoded windows (Remark 3), and a purely spectral network can produce a systematic amplitude offset (Appendix F.3). To correct this we add a small temporal branch T_θ of two AdaLN blocks, which reads the same flow state in the window basis and returns its features to DCT alignment through a learned tokenwise gate,

$$h_T = C_B\, T_\theta\big(C_B^\top z_\tau, c_\tau\big), \qquad v_\theta(z_\tau, \tau, y) = P_{\mathrm{out}}\big(h_D + g \odot h_T\big), \qquad g = \sigma\big(W_g[h_D; h_T] + b_g\big). \tag{5}$$

The flow state and the velocity remain in DCT coordinates, and since $C_B^\top$ is a fixed linear map, the objective, its minimizer and the Riemannian equivalence are unchanged. The branch's primary role is to reduce the amplitude bias (Section 6.3, Appendix A.17).

4.2.3 Training and Sampling

Let $\Psi_\# q(\cdot \mid y)$ be the Euclidean pushforward of the class-conditional distribution of GVD trajectories. We draw $z_0 \sim \mathcal{N}(0, I)$, $z_1 \sim \Psi_\# q(\cdot \mid y)$ and $\tau \sim \mathcal{U}[0,1]$, and use $z_\tau = (1-\tau)z_0 + \tau z_1$ with target $u_\tau = z_1 - z_0$. Within each minibatch, source and target samples of the same class are paired by an entropic optimal-transport plan π_y computed with Sinkhorn iterations (Cuturi, 2013;

Tong et al., 2024) on the squared Euclidean cost in the DCT chart. The orthonormal DCT preserves distances, so this cost is the squared geodesic distance under g_σ, and the coupling is computed in the geometry of the product manifold. GVD-CFM minimizes

$$\mathcal{L}_{\text{GVD-CFM}}(\theta) = \mathbb{E}_{y,\,(z_0,z_1)\sim\pi_y,\,\tau}\left[\frac{1}{Bm}\left\|v_\theta(z_\tau,\tau,y) - u_\tau\right\|_F^2\right]. \tag{6}$$

To sample, we draw $z(0) \sim \mathcal{N}(0, I)$, integrate $\mathrm{d}z/\mathrm{d}\tau = v_{\theta^\star}(z, \tau, y)$ with fourth-order Runge–Kutta (Hairer et al., 1993), and decode the result through Ψ^{-1}. Algorithms 1 and 2 in Appendix A state both procedures in full.

4.3 Why Spectral Coordinates Simplify Trajectory Flow Matching

The DCT reparametrizes the dependence structure of the complete log-Euclidean trajectory. Let $K_T \in \mathbb{R}^{B\times B}$ be the temporal covariance of the log-GVD trajectory. The Karhunen–Loève transform (KLT) diagonalizes K_T exactly, and the orthonormal DCT is a fixed, data-independent approximation to the KLT for strongly correlated, temporally smooth processes (Ahmed et al., 1974; Strang, 1999). This has a direct consequence for conditional flow matching.

Theorem 1 (Spectral decoupling of Gaussian trajectory flow)**.** *Let $x_1 \sim \mathcal{N}(0, \Sigma)$ with $\Sigma = V\Lambda V^\top$, $x_0 \sim \mathcal{N}(0, I)$, and $x_\tau = (1-\tau)x_0 + \tau x_1$. In the covariance eigenbasis $y = V^\top x$, the population-optimal squared-error conditional flow-matching field is diagonal:*

$$v^\star_{\tau,k}(y) = \frac{\tau\lambda_k - (1-\tau)}{(1-\tau)^2 + \tau^2\lambda_k}\, y_k \qquad \textit{(independent coupling)}, \tag{7}$$

$$v^\star_{\tau,k}(y) = \frac{\sqrt{\lambda_k} - 1}{(1-\tau) + \tau\sqrt{\lambda_k}}\, y_k \qquad \textit{(optimal-transport coupling)}.$$

Hence, in the exact KLT basis, no cross-mode interaction is needed to represent the optimal second-order Gaussian transport under either coupling.

The proofs are given in Appendix C (Corollaries 3 and 4). The residual off-diagonal temporal covariance measures how closely the DCT approximates the KLT; on BNCI2014_001 the DCT reduces it from 0.778 to 0.068 (Figure 5). The DCT therefore removes most of the second-order temporal coupling before the Transformer is applied, while remaining invertible and isometric.

5 Empirical Benchmarks

The main benchmark uses five two-class motor-imagery datasets from MOABB (Jayaram and Barachant, 2018; Aristimunha et al., 2023): BNCI2014_001, BNCI2014_002, BNCI2015_001, Shin2017A and Zhou2016. Every trial is band-pass filtered to 4–38 Hz, resampled to 128 Hz, standardized per channel and converted to a GVD trajectory of $B = 100$ windows. The final session, or the final run when only runs are available, is held out, and all results are averaged over three generator seeds (Appendix A).

We compare against seven generators. Three are GVD-space controls that share the GVD targets and decoding pipeline of GVD-CFM and therefore isolate the contribution of the generative model: GVD-cVAE, GVD-DDPM and Window-DIFFEO-CFM, the last of which applies diffeomorphic flow matching (Collas et al., 2025) to each window independently. Four generate raw multichannel EEG, which is converted to a GVD trajectory using its own support and window covariances: JET (Wang et al., 2026), a U-Net DDPM/DDIM (Ho et al., 2020; Song et al., 2021a), EEGGAN-2025 (Williams et al., 2025) and a conditional VAE (Sohn et al., 2015). No real support or window covariance is ever reused for a raw-EEG baseline.

We report a full-dimensional GVD Fréchet distance relative to the real-train to real-test reference (Dowson and Landau, 1982; Heusel et al., 2017); EvaGeM α-precision, β-recall and F1 (Alaa et al., 2022); the classification accuracy score (CAS) of a classifier trained only on generated trajectories and tested on held-out real ones (Ravuri and Vinyals, 2019); held-out temporal diagnostics; and novelty measures (Kynkäänniemi et al., 2019; Alaa et al., 2022). All are defined in Appendix A.

Table 1: Dataset-balanced generative performance across five EEG datasets, averaged over three generator seeds per dataset. Rel. GVD-FID and EvaGeM are evaluated jointly on trajectory positions and temporal increments. The real-data row reports the corresponding real-train/held-out-real-test reference and is excluded from generator rankings. Entries after $\pm$ are the seed standard deviation of the dataset-balanced mean, $\left(\sum_{i=1}^{5}\sigma_i^2\right)^{1/2}/5$, computed from the per-dataset seed standard deviations σ_i. Lower is better for Rel. GVD-FID; higher is better otherwise. Best generative values are **bold**; second-best generative values are *underlined italics*.

Method	Rel. GVD-FID ↓	Eva α ↑	Eva β ↑	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑
GVD-CFM	*1.022*±0.019	**0.677**±0.034	**0.610**±0.029	**0.613**±0.028	**0.790**±0.008	**0.725**±0.006
GVD-cVAE	**0.815**±0.003	0.017±0.005	0.004±0.002	0.005±0.002	*0.727*±0.015	*0.675*±0.010
GVD-DDPM	1.753±0.005	0.089±0.011	0.039±0.005	0.053±0.007	0.599±0.014	0.560±0.012
Window-DIFFEO-CFM	1.697±0.005	0.005±0.002	0.082±0.002	0.009±0.003	0.683±0.003	0.634±0.002
cVAE	1.416±0.016	*0.407*±0.036	0.006±0.002	0.012±0.003	0.631±0.006	0.572±0.017
JET	3.908±0.274	0.005±0.002	0.002±0.002	0.002±0.002	0.495±0.031	0.349±0.014
Vanilla-Diffusion	1.867±0.096	0.217±0.042	*0.366*±0.098	*0.247*±0.056	0.640±0.012	0.549±0.029
EEGGAN-2025	3.696±0.016	0.006±0.002	0.002±0.001	0.002±0.001	0.524±0.031	0.401±0.017
Real data reference	–	0.806	0.665	0.695	0.831	0.756

Table 2: Complementary data-quality diagnostics averaged across five datasets and three seeds. Dynamic and diversity ratios have ideal value 1. Temporal quantities are evaluated against held-out real trajectories after removing each trial's temporal mean, so the static support does not contribute to them. $\pm$ as in Table 1. Best values are **bold**; second-best values are *underlined italics*. For ratio metrics, ranking is by proximity to 1.

Method	Temp. corr. ↑	Lag-ACF ↑	Energy → 1	Dyn. frac. → 1	Diversity → 1	Coverage ↑
GVD-CFM	**0.910**±0.004	*0.997*±0.001	**0.953**±0.010	**0.987**±0.013	**0.987**±0.005	**0.226**±0.005
GVD-cVAE	0.774±0.014	0.995±0.001	0.260±0.007	0.289±0.009	0.396±0.008	*0.128*±0.004
GVD-DDPM	0.059±0.005	0.327±0.068	1.688±0.007	1.543±0.008	1.367±0.004	0.038±0.002
Window-DIFFEO-CFM	0.012±0.005	0.055±0.091	1.794±0.008	1.755±0.006	1.189±0.002	0.031±0.001
cVAE	0.527±0.016	0.957±0.006	0.898±0.006	*0.981*±0.014	0.817±0.008	0.034±0.001
JET	0.385±0.019	0.933±0.009	1.940±0.118	6.835±0.338	1.501±0.079	0.009±0.004
Vanilla-Diffusion	*0.783*±0.014	**0.998**±0.001	*0.941*±0.037	1.759±0.418	1.070±0.049	0.085±0.010
EEGGAN-2025	0.022±0.010	0.306±0.137	0.609±0.012	4.864±0.199	*1.033*±0.019	0.003±0.000

6 RESULTS

6.1 GENERATING HIGH-RESOLUTION CONNECTIVITY TRAJECTORIES

Table 1 reports the main results; per-dataset values are given in Appendix F.6. In downstream classification GVD-CFM performs best by a large margin, with a CAS AUC of 0.790 and CAS F1 of 0.725 against 0.727 and 0.675 for the second-best model, GVD-cVAE, and it comes within 0.041 AUC of a classifier trained on real data. GVD-cVAE attains the lowest relative GVD-FID, but its EvaGeM precision and recall are close to zero and its diversity ratio is 0.396 (Table 2), so its samples concentrate near the center of the distribution rather than covering it. GVD-CFM obtains the second-lowest relative GVD-FID (1.022) together with the highest EvaGeM α-precision, β-recall and F1 of all models, reaching 88% of the real-data EvaGeM F1. The raw-EEG generators obtain EvaGeM F1 below 0.25 once their outputs are converted to GVD trajectories; we can see that realistic raw EEG does not imply realistic connectivity dynamics.

6.2 DATA QUALITY

Table 2 shows that GVD-CFM gives the best overall balance between coverage and temporal fidelity. It achieves the highest held-out temporal-correlation agreement (0.910) and the second-highest lag-ACF agreement (0.997, against 0.998 for Vanilla-Diffusion). Its dynamic-energy, dynamic-fraction and diversity ratios are the closest to 1 of all generators, and its training-manifold coverage is the highest. The GVD-space controls show why joint spectral modeling matters. GVD-DDPM and Window-DIFFEO-CFM, which model temporal windows directly, obtain temporal-correlation

Table 3: DCT and temporal-branch ablation. *No DCT* trains the Transformer on temporal log-svec coordinates; *DCT, spectral only* is GVD-CFM without the temporal branch; the last row is the full GVD-CFM. All values are dataset-balanced means over five datasets and three generator seeds; per-dataset values are in Tables 16 and 17, and paired per-dataset tests in Table 18. Ratios have ideal value 1. $\pm$ as in Table 1. Best values are **bold**; second-best values are *underlined italics*.

Variant	Rel. GVD-FID ↓	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑	Temp. corr. ↑	Lag-ACF ↑	Energy → 1	Dyn. frac. → 1
No DCT	1.092±0.029	0.540±0.028	0.772±0.007	0.704±0.007	0.404±0.007	*0.666*±0.069	*0.973*±0.016	*1.011*±0.017
DCT, spectral only	*1.025*±0.006	*0.600*±0.019	**0.792**±0.002	*0.721*±0.004	**0.919**±0.002	**0.997**±0.001	**0.975**±0.011	**1.010**±0.010
GVD-CFM (DCT + temporal branch)	**1.022**±0.019	**0.613**±0.028	*0.790*±0.008	**0.725**±0.006	*0.910*±0.004	**0.997**±0.001	0.953±0.010	0.987±0.013

agreement below 0.06; GVD-cVAE retains temporal correlation but collapses dynamic energy to 0.260 of the real value. GVD-CFM produces no exact copies of training trajectories (Appendix F).

Static support versus generated dynamics. Every GVD window carries the trial-level support W, so class information could reside in static structure alone. The generated trajectories carry dynamics beyond it. The temporal diagnostics above remove each trial's temporal mean and are insensitive to W; a permutation test rejects exchangeability of the generated windows at $p = 0.002$, the smallest attainable value with 500 permutations, on every dataset and seed; and on real data the full GVD trajectory outperforms the static support alone in held-out classification, while a support modulated by Gaussian noise falls to near chance (Figure 10).

6.3 Ablations

The DCT gives large and consistent gains. Table 3 separates the spectral representation from the temporal branch. Moving from temporal coordinates to the DCT chart raises temporal-correlation agreement from 0.404 to 0.919 and lag-ACF agreement from 0.666 to 0.997, in line with Theorem 1. Both improvements hold on all five datasets (Welch t from 4.8 to 80.9 and from 3.4 to 13.0; Table 18). Under the same spectral-only network, a random orthogonal basis does not reproduce this gain, while the empirical KLT performs comparably to the DCT (Table 4). The benefit therefore comes from temporal decorrelation rather than from orthogonality alone.

The temporal branch corrects amplitude. Adding the temporal branch changes the dataset-balanced means by amounts comparable to their seed variation (EvaGeM F1 $0.600 \to 0.613$, CAS F1 $0.721 \to 0.725$, CAS AUC $0.792 \to 0.790$, temporal correlation $0.919 \to 0.910$), so we treat these differences as ties. Its intended effect is on amplitude: dynamic energy decreases on every dataset and moves towards 1 where the spectral-only network overshoots, as we show in Appendix F.3.

The stable support improves downstream utility. We compared GVD-CFM with the same generator trained on ridge-regularized window covariances $J_b + 10^{-6} I$, since without the support a ridge is needed to make the windows SPD (per-dataset values in Table 19; $\pm$ as in Table 1). The stable support raises CAS AUC on every dataset; the dataset-balanced CAS AUC rises from 0.736 ± 0.005 to 0.790 ± 0.008, CAS F1 from 0.679 ± 0.005 to 0.725 ± 0.006 and EvaGeM F1 from 0.436 ± 0.015 to 0.613 ± 0.028.

6.4 Regional Physiological Plausibility

Global metrics do not show whether generated connectivity is organized in space and time as in real EEG. We grouped the 22 electrodes of BNCI2014_001 into Frontal/FC, Central, Centro-parietal and Parietal/Occipital regions and computed the mean GVD edge magnitude for every pair of regions (Appendix E). Generated trajectories keep the regional organization of held-out data. Within-region connectivity is 0.835 against 0.847 for real data in Frontal/FC and 0.880 against 0.894 in Parietal/Occipital, Central–Centro-parietal coupling is 0.706 against 0.717, and the long-range Frontal/FC–Parietal/Occipital coupling stays weak (0.420 against 0.437). Figure 6 shows that the time course of these interactions is also preserved, and that its changes occur together across central, centro-parietal, frontal and posterior interactions. The generated dynamics are therefore organized around the sensorimotor regions expected to take part in motor imagery (Pfurtscheller and da Silva, 1999).

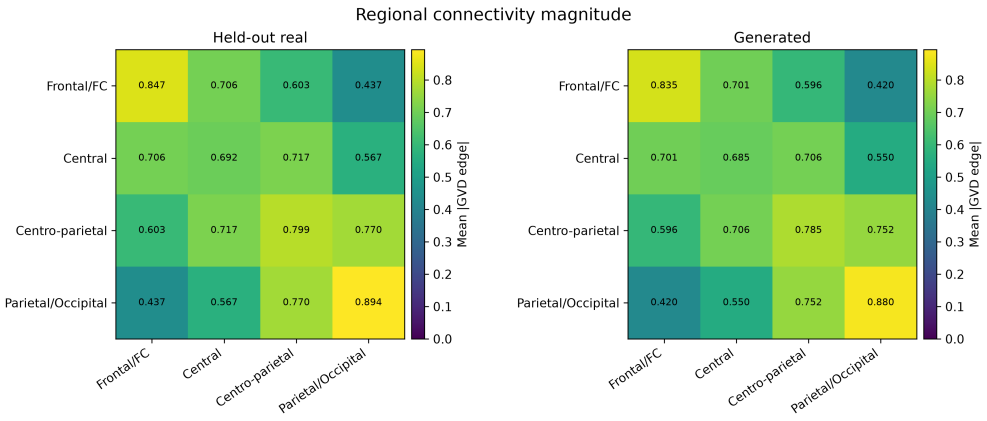


Figure 2: **Regional connectivity magnitude on BNCI2014_001.** Mean absolute GVD edge magnitude between broad scalp regions for held-out real trials and generated trials. The generated data preserve the relative regional organization of the real EEG, including strong within-region Frontal/FC and Parietal/Occipital connectivity, strong Central–Centro-parietal coupling, and weaker long-range Frontal/FC–Parietal/Occipital interactions.

6.5 Continuous-Grid Decoding and Efficiency

The generated DCT coefficients define a band-limited cosine trajectory, so they can be decoded on any temporal grid without retraining. Training the spectral-only network once at $B_{\text{train}} = 25$ windows and decoding at $M = 100$ changes generated-to-real AUC only from 0.803 to 0.802; at $M = 200$ it remains 0.794 with lag-ACF agreement 0.959, and every decoded matrix stays SPD. At $M = 400$ temporal-correlation agreement falls to 0.506, the bandwidth limit of 25 modes (Appendix D). GVD-CFM trains in about five minutes per dataset and achieves the best efficiency trade off (Appendix Figure 9).

7 Conclusions and Limitations

We introduced GVD-CFM, a conditional flow matching model for generating high-temporal-resolution DFC trajectories from EEG. When the stable support is positive definite, GVD trajectories lie on a product Riemannian manifold even when individual window covariances are rank deficient. This allows Euclidean flow matching under a global diffeomorphism to be exactly equivalent to Riemannian flow matching on the original manifold. We further showed that applying the DCT approximately decorrelates the temporal structure of these trajectories, leading to large and consistent improvements in temporal fidelity. Across five EEG datasets and seven baselines, GVD-CFM provides the best overall trade-off across the reported metrics. The generated trajectories remain on the manifold, show no evidence of memorizing the training data, and preserve both the regional organization and temporal evolution of motor-imagery connectivity. Together, these results suggest that directly modeling connectivity trajectories is a practical alternative to generating raw EEG when the downstream quantity of interest is dynamic functional connectivity.

Our work also has several limitations. First, the number of log-Euclidean coordinates per window grows quadratically with the number of channels, which may limit scalability to substantially higher-density recordings. Second, GVD-CFM ultimately depends on the quality of the underlying GVD estimator. Future work could explore using phase- or coherence-based node functions (Smith et al., 2019; Roy et al., 2025), and could explore alternative diffeomorphisms with different geometric or computational properties (Lin, 2019; David and Gu, 2019; Thanwerdas, 2022). It would also be useful to test whether the same framework generalizes beyond motor-imagery EEG and to other modalities where dynamic connectivity is of interest, particularly fMRI (Hutchison et al., 2013; Allen et al., 2014; Preti et al., 2017). More broadly, extending GVD-CFM to larger networks, frequency-resolved connectivity, and other forms of neural dynamics would help establish how well the approach scales beyond the setting studied here.

Reproducibility Statement

All five datasets are public and are obtained through MOABB (Jayaram and Barachant, 2018; Aristimunha et al., 2023). Appendix A gives the preprocessing, the GVD construction, the log-Euclidean and DCT chart, the architecture and optimizer settings of GVD-CFM, including the temporal branch and the minibatch coupling, the configuration of every baseline, and the definition of every metric. Training and sampling are stated as Algorithms 1 and 2. All reported experiments use three generator seeds and three classifier seeds, stated in the appendix, and exact train–test duplication is checked by hashing before any model is fitted. Code is included in the supplementary material.

Use of Large Language Models

Large language models were used for proofreading, notation consistency, and formatting. They were also used to check correctness of mathematical proofs. All scientific claims, mathematical arguments, experimental choices, implementations, and reported results remain the responsibility of the authors.

References

Nasir Ahmed, T. Natarajan, and Kamisetty R. Rao. Discrete cosine transform. *IEEE Transactions on Computers*, C-23(1):90–93, 1974.

Ahmed Alaa, Boris van Breugel, Evgeny S. Saveliev, and Mihaela van der Schaar. How faithful is your synthetic data? Sample-level metrics for evaluating and auditing generative models. In *39th International Conference on Machine Learning*, pages 290–306, Baltimore, MD, July 2022.

Michael S. Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *11th International Conference on Learning Representations*, Kigali, Rwanda, May 2023.

Elena A. Allen, Eswar Damaraju, Sergey M. Plis, Erik B. Erhardt, Tom Eichele, and Vince D. Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24 (3):663–676, 2014.

Bruno Aristimunha, Igor Carrara, Pierre Guetschel, Sara Sedlar, Pedro Rodrigues, Jan Sosulski, Divyesh Narayanan, Erik Bjareholt, Quentin Barthelemy, Robin T. Schirrmeister, Emmanuel Kalunga, Ludovic Darmet, Cattan Gregoire, Ali Abdul Hussain, Ramiro Gatti, Vladislav Goncharenko, Jordy Thielen, Thomas Moreau, Yannick Roy, Vinay Jayaram, Alexandre Barachant, and Sylvain Chevallier. Mother of all BCI benchmarks (MOABB), 2023.

Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *34th International Conference on Machine Learning*, pages 214–223, Sydney, Australia, August 2017.

Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2007.

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer normalization. *arXiv:1607.06450 [stat.ML]*, 2016.

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

Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass brain–computer interface classification by Riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928, 2012.

Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-Robert Müller. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine*, 25(1):41–56, 2008.

Valentin De Bortoli, Emile Mathieu, Michael Hutchinson, James Thornton, Yee Whye Teh, and Arnaud Doucet. Riemannian score-based generative modelling. In *36th Conference on Neural Information Processing Systems*, pages 2406–2422, December 2022.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. In *34th Conference on Neural Information Processing Systems*, pages 1877–1901, December 2020.

Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, and Gert Pfurtscheller. BCI Competition 2008 – Graz data set A. Technical report, Institute for Knowledge Discovery, Graz University of Technology, 2008.

Andrea Cavallo, Maosheng Sabbaqi, and Elvin Isufi. Spatiotemporal covariance neural networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 18–34, Cham, August 2024. Springer Nature Switzerland.

Ricky T. Q. Chen and Yaron Lipman. Flow matching on general geometries. In *12th International Conference on Learning Representations*, Vienna, Austria, May 2024.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. In *32nd Conference on Neural Information Processing Systems*, pages 6572–6583, Montréal, QC, December 2018.

Yilun Chen, Ami Wiesel, Yonina C. Eldar, and Alfred O. Hero. Shrinkage algorithms for MMSE covariance estimation. *IEEE Transactions on Signal Processing*, 58(10):5016–5029, 2010.

Antoine Collas, Ce Ju, Nicolas Salvy, and Bertrand Thirion. Riemannian flow matching for brain connectivity matrices via pullback geometry (DIFFEOCFM). In *39th Conference on Neural Information Processing Systems*, December 2025.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, 2013.

Paul David and Weiqing Gu. A Riemannian structure for correlation matrices. *Operators and Matrices*, 13(3):607–627, 2019.

Thibault de Surrel, Fabien Lotte, Sylvain Chevallier, and Florian Yger. Wrapped Gaussian on the manifold of symmetric positive definite matrices. In *42nd International Conference on Machine Learning*, July 2025.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat GANs on image synthesis. In *35th Conference on Neural Information Processing Systems*, pages 8780–8794, December 2021.

Haoran Ding, Noémie Jaquier, Jan Peters, and Leonel Rozo. Fast and robust visuomotor Riemannian flow matching policy. *IEEE Transactions on Robotics*, 2025.

Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using Real NVP. In *5th International Conference on Learning Representations*, Toulon, France, April 2017.

D. C. Dowson and B. V. Landau. The Fréchet distance between multivariate normal distributions. *Journal of Multivariate Analysis*, 12(3):450–455, 1982.

Cristóbal Esteban, Stephanie L. Hyland, and Gunnar Rätsch. Real-valued (medical) time series generation with recurrent conditional GANs. *arXiv:1706.02633 [stat.ML]*, 2017.

Josef Faller, Carmen Vidaurre, Teodoro Solis-Escalante, Christa Neuper, and Reinhold Scherer. Autocalibration and recurrent adaptation: towards a plug and play online ERD-BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(3):313–319, 2012.

Luca Falorsi, Pim de Haan, Tim R. Davidson, and Patrick Forré. Reparameterizing distributions on Lie groups. In *22nd International Conference on Artificial Intelligence and Statistics*, pages 3244–3253, Naha, Japan, April 2019.

P. Thomas Fletcher and Sarang Joshi. Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In *Sonka, M., Kakadiaris, I.A., Kybic, J. (eds) Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis. MMBIA CVAMIA 2004 2004.*, Lecture Notes in Computer Science, vol 3117. Springer, Berlin, Heidelberg.

Michael D. Fox, Randy L. Buckner, Hesheng Liu, M. Mallar Chakravarty, Andres M. Lozano, and Alvaro Pascual-Leone. Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. *Proceedings of the National Academy of Sciences*, 111(41):E4367–E4375, 2014.

Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. JHU Press, Baltimore, MD, 4th edition, 2013.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. In *28th Conference on Neural Information Processing Systems*, pages 2672–2680, Montréal, QC, December 2014.

Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and Matti S. Hämäläinen. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7:267, 2013.

Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved training of Wasserstein GANs. In *31st Conference on Neural Information Processing Systems*, pages 5769–5779, Long Beach, CA, December 2017.

Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, Berlin, Heidelberg, 2nd edition, 1993.

David Hallac, Youngsuk Park, Stephen Boyd, and Jure Leskovec. Network inference via the time-varying graphical lasso. In *23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 205–213, Halifax, NS, August 2017.

Kay Gregor Hartmann, Robin Tibor Schirrmeister, and Tonio Ball. EEG-GAN: generative adversarial networks for electroencephalographic (EEG) brain signals. *arXiv:1806.01875 [eess.SP]*, 2018.

Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (GELUs). *arXiv:1606.08415 [cs.LG]*, 2016.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *31st Conference on Neural Information Processing Systems*, pages 6629–6640, Long Beach, CA, December 2017.

Rikkert Hindriks, Mohit H. Adhikari, Yusuke Murayama, Marco Ganzetti, Dante Mantini, Nikos K. Logothetis, and Gustavo Deco. Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *NeuroImage*, 127:242–256, 2016.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv:2207.12598 [cs.LG]*, 2022.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *34th Conference on Neural Information Processing Systems*, pages 6840–6851, December 2020.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 2nd edition, 2012.

Xikun Huang,Tianyu Ruan, Chihao Zhang and Shihua Zhang. Graph generation with spectral geodesic flow matching. *arXiv:2510.02520 [cs.LG]*, 2025.

Guillaume Huguet, James Vuckovic, Kilian Fatras, Eric Thibodeau-Laufer, Pablo Lemos, Riashat Islam, Cheng-Hao Liu, Jarrid Rector-Brooks, Tara Akhound-Sadegh, Michael M. Bronstein, Alexander Tong, and Avishek Joey Bose. Sequence-augmented SE(3)-flow matching for conditional protein generation. In *38th Conference on Neural Information Processing Systems*, December 2024.

R. Matthew Hutchison, Thilo Womelsdorf, Elena A. Allen, Peter A. Bandettini, Vince D. Calhoun, Maurizio Corbetta, Stefania Della Penna, Jeff H. Duyn, Gary H. Glover, Javier Gonzalez-Castillo, Daniel A. Handwerker, Shella Keilholz, Vesa Kiviniemi, David A. Leopold, Francesco de Pasquale, Olaf Sporns, Martin Walter, and Catie Chang. Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage*, 80:360–378, 2013.

Vinay Jayaram and Alexandre Barachant. MOABB: trustworthy algorithm benchmarking for BCIs. *Journal of Neural Engineering*, 15(6):066011, 2018.

Weibang Jiang, Liming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous EEG data in BCI. In *12th International Conference on Learning Representations*, Vienna, Austria, May 2024.

Jaehyeong Jo and Sung Ju Hwang. Generative modeling on manifolds through mixture of Riemannian diffusion processes. *arXiv:2310.07216 [cs.LG]*, 2023.

Ce Ju, Reinmar J. Kobler, Antoine Collas, Motoaki Kawanabe, Cuntai Guan, and Bertrand Thirion. SPD learning for covariance-based neuroimaging analysis: perspectives, methods, and challenges. *arXiv:2504.18882 [cs.LG]*, 2025.

Kacper Kapusniak, Peter Potaptchik, Teodora Reu, Leo Zhang, Alexander Tong, Michael Bronstein, Avishek Joey Bose, and Francesco Di Giovanni. Metric flow matching for smooth interpolations on the data manifold. In *38th Conference on Neural Information Processing Systems*, December 2024.

Diederik P. Kingma and Jimmy Lei Ba. ADAM: A method for stochastic optimization. In *3rd International Conference on Learning Representations*, pages 1–15, San Diego, CA, May 2015.

Diederik P. Kingma and Max Welling. Auto-encoding variational Bayes. In *2nd International Conference on Learning Representations*, Banff, AB, April 2014.

Reinmar J. Kobler, Jun ichiro Hirayama, Qibin Zhao, and Motoaki Kawanabe. SPD domain-specific batch normalization to crack interpretable unsupervised domain adaptation in EEG. In *36th Conference on Neural Information Processing Systems*, pages 6219–6235, December 2022.

Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Improved precision and recall metric for assessing generative models. In *33rd Conference on Neural Information Processing Systems*, Vancouver, BC, December 2019.

Elnaz Lashgari, Dehua Liang, and Uri Maoz. Data augmentation for deep-learning-based electroencephalography. *Journal of Neuroscience Methods*, 346:108885, 2020.

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

Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.

Robert Leeb, Felix Lee, Claudia Keinrath, Reinhold Scherer, Horst Bischof, and Gert Pfurtscheller. Brain–computer communication: motivation, aim, and impact of exploring a virtual apartment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(4):473–482, 2007.

Nora Leonardi and Dimitri Van De Ville. On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage*, 104:430–436, 2015.

Yunchen Li, Zhou Yu, Gaoqi He, Yunhang Shen, Ke Li, Xing Sun, and Shaohui Lin. SPD-DDPM: denoising diffusion probabilistic models in the symmetric positive definite space. In *38th AAAI Conference on Artificial Intelligence*, pages 13709–13717, Vancouver, BC, February 2024.

Zhenhua Lin. Riemannian geometry of symmetric positive definite matrices via Cholesky decomposition. *SIAM Journal on Matrix Analysis and Applications*, 40(4):1353–1370, 2019.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *11th International Conference on Learning Representations*, Kigali, Rwanda, May 2023.

Yaron Lipman, Marton Havasi, Peter Holderrieth, Neta Shaul, Matt Le, Brian Karrer, Ricky T. Q. Chen, David Lopez-Paz, Heli Ben-Hamu, and Itai Gat. Flow matching guide and code. *arXiv:2412.06264 [cs.LG]*, 2024.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: learning to generate and transfer data with rectified flow. In *11th International Conference on Learning Representations*, Kigali, Rwanda, May 2023.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *7th International Conference on Learning Representations*, New Orleans, LA, May 2019.

Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3), 2018.

Yun Luo and Bao-Liang Lu. EEG data augmentation for emotion recognition using a conditional Wasserstein GAN. In *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2535–2538, Honolulu, HI, July 2018.

Gautier Marti. CorrGAN: sampling realistic financial correlation matrices using generative adversarial networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8459–8463, Barcelona, Spain, May 2020.

Robert J. McCann. A convexity principle for interacting gases. *Advances in Mathematics*, 128(1): 153–179, 1997.

Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M. Wood. FlowMM: generating materials with Riemannian flow matching. In *41st International Conference on Machine Learning*, Vienna, Austria, July 2024.

Ricardo Pio Monti, Peter Hellyer, David Sharp, Robert Leech, Christoforos Anagnostopoulos, and Giovanni Montana. Estimating time-varying brain connectivity networks from functional MRI time series. *NeuroImage*, 103:427–443, 2014.

Muhammad Ferjad Naeem, Seong Joon Oh, Youngjung Uh, Yunjey Choi, and Jaejun Yoo. Reliable fidelity and diversity metrics for generative models. In *37th International Conference on Machine Learning*, pages 7176–7185, July 2020.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. In *33rd Conference on Neural Information Processing Systems*, pages 8024–8035, Vancouver, BC, December 2019.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *IEEE/CVF International Conference on Computer Vision*, pages 4195–4205, Paris, France, October 2023.

Xavier Pennec. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006.

Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.

Gert Pfurtscheller and Fernando H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.

Maria Giulia Preti, Thomas A. W. Bolton, and Dimitri Van De Ville. The dynamic functional connectome: state-of-the-art and perspectives. *NeuroImage*, 160:41–54, 2017.

Yiming Qin, Manuel Madeira, Dorina Thanou, and Pascal Frossard. DeFoG: discrete flow matching for graph generation. In *42nd International Conference on Machine Learning*, July 2025.

Kashif Rasul, Calvin Seward, Ingmar Schuster, and Roland Vollgraf. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. In *38th International Conference on Machine Learning*, pages 8857–8868, July 2021.

Suman Ravuri and Oriol Vinyals. Classification accuracy score for conditional generative models. In *33rd Conference on Neural Information Processing Systems*, pages 12268–12279, Vancouver, BC, December 2019.

Om Roy, Yashar Moshfeghi, Agustin Ibanez, Francisco Lopera, Mario A. Parra, and Keith M. Smith. Robust, high temporal-resolution EEG functional connectivity detects increased connectivity coinciding with P300 in visual short-term memory binding in both familial and sporadic prodromal Alzheimer's disease. In *Complex Networks 2023: The 12th International Conference on Complex Networks and Their Applications*, pages 679–682, 2023.

Om Roy, Yashar Moshfeghi, Agustin Ibanez, Francisco Lopera, Mario A. Parra, and Keith M. Smith. FAST functional connectivity implicates P300 connectivity in working memory deficits in Alzheimer's disease. *Network Neuroscience*, 8(4):1467–1490, 2024.

Om Roy, Yashar Moshfeghi, J. Smith, Agustin Ibanez, Mario A. Parra, and Keith M. Smith. A Hodge-FAST framework for high-resolution dynamic functional connectivity analysis of higher-order interactions in EEG signals. In *47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1–6, 2025. doi: 10.1109/EMBC58623.2025.11253015.

Om Roy, Yashar Moshfeghi, and Keith M. Smith. Covariance density neural networks. *Transactions on Machine Learning Research*, 2026.

Yannick Roy, Hubert Banville, Isabela Albuquerque, Alexandre Gramfort, Tiago H. Falk, and Jocelyn Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*, 16(5):051001, 2019.

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.

Issai Schur. Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. *Journal für die reine und angewandte Mathematik*, 140:1–28, 1911.

David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98, 2013.

Saurabh Sihag, Gonzalo Mateos, Corey McMillan, and Alejandro Ribeiro. CoVariance neural networks. In *36th Conference on Neural Information Processing Systems*, pages 17003–17016, Red Hook, NY, 2022. Curran Associates Inc.

Lene Theil Skovgaard. A Riemannian geometry of the multivariate normal model. *Scandinavian Journal of Statistics*, 11(4):211–223, 1984.

Keith Smith, Javier Escudero, Mario A. Parra, Agustin Ibanez, John M. Starr, and Sergio Della Sala. Locating temporal functional dynamics of visual short-term memory binding using graph modular Dirichlet energy. *Scientific Reports*, 7:42013, 2017.

Keith Smith, Loukas Spyrou, and Javier Escudero. Graph-variate signal analysis. *IEEE Transactions on Signal Processing*, 67(2):293–305, 2019. doi: 10.1109/TSP.2018.2881658.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *32nd International Conference on Machine Learning*, pages 2256–2265, Lille, France, July 2015.

Kihyuk Sohn, Honglak Lee, and Xinchen Yan. Learning structured output representation using deep conditional generative models. In *29th Conference on Neural Information Processing Systems*, pages 3483–3491, Montréal, QC, December 2015.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *9th International Conference on Learning Representations*, May 2021a.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *33rd Conference on Neural Information Processing Systems*, Vancouver, BC, December 2019.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *9th International Conference on Learning Representations*, May 2021b.

David Steyrl, Reinhold Scherer, Josef Faller, and Gernot R. Müller-Putz. Random forests in non-invasive sensorimotor rhythm brain-computer interfaces: a practical and convenient non-linear classifier. *Biomedical Engineering / Biomedizinische Technik*, 61(1):77–86, 2016.

Gilbert Strang. The discrete cosine transform. *SIAM Review*, 41(1):135–147, 1999.

Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high-frequency functions in low-dimensional domains. In *34th Conference on Neural Information Processing Systems*, pages 7537–7547, December 2020.

Michael Tangermann, Klaus-Robert Müller, Ad Aertsen, Niels Birbaumer, Christoph Braun, Clemens Brunner, Robert Leeb, Carsten Mehring, Kai J. Miller, Gernot Müller-Putz, Guido Nolte, Gert Pfurtscheller, Hubert Preissl, Gerwin Schalk, Alois Schlögl, Carmen Vidaurre, Stephan Waldert, and Benjamin Blankertz. Review of the BCI competition IV. *Frontiers in Neuroscience*, 6:55, 2012.

Kiet Bennema ten Brinke, Koen Minartz, and Vlado Menkovski. STFlow: data-coupled flow matching for geometric trajectory simulation. In *43rd International Conference on Machine Learning*, 2026.

Yann Thanwerdas. *Riemannian and stratified geometries on covariance and correlation matrices*. PhD thesis, Université Côte d'Azur, 2022.

Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *31st Conference on Neural Information Processing Systems*, pages 5998–6008, Long Beach, CA, December 2017.

Clement Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, and Pascal Frossard. DiGress: discrete denoising diffusion for graph generation. In *11th International Conference on Learning Representations*, Kigali, Rwanda, May 2023.

Cédric Villani. *Optimal Transport: Old and New*. Springer, Berlin, Heidelberg, 2009.

Yifan Wang, Yijia Ma, Wen Li, and Chenyu You. Let EEG models learn EEG. *arXiv:2605.21280 [cs.LG]*, 2026.

Chad C. Williams, Daniel Weinhardt, Maria Wirzberger, and Sebastian Musslick. Augmenting EEG with generative adversarial networks enhances brain decoding across classifiers and sample sizes. In *45th Annual Meeting of the Cognitive Science Society* Sydney, Australia, July 2023.

Chad C. Williams, Daniel Weinhardt, Joshua Hewson, Martyna Beata Płomecka, Nicolas Langer, and Sebastian Musslick. EEG-GAN: a generative EEG augmentation toolkit for enhancing neural classification. *bioRxiv*, 2025. doi: 10.1101/2025.06.23.661164.

Robert Williams. Scalable generative modeling of weighted graphs. *arXiv:2507.23111 [cs.LG]*, 2025.

Xu F, Dong G, Li J, Yang Q, Wang L, Zhao Y, Yan Y, Zhao J, Pang S, Guo D, Zhang Y and Leng J. Deep convolution generative adversarial network-based electroencephalogram data augmentation for post-stroke rehabilitation with motor imagery. *International Journal of Neural Systems*, 32 (9):2250039, 2022. doi: 10.1142/S0129065722500393.

Qiqi Zhang and Ying Liu. Improving brain computer interface performance by data augmentation with conditional deep convolutional generative adversarial networks. *arXiv:1806.07108 [cs.HC]*, 2018.

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.

A Experimental Details

A.1 Datasets and evaluation protocol

Experiments were conducted on five motor-imagery EEG datasets distributed through MOABB (Jayaram and Barachant, 2018; Aristimunha et al., 2023). The main benchmark uses `BNCI2014_001` (Tangermann et al., 2012; Brunner et al., 2008), `BNCI2014_002` (Steyrl et al., 2016), `BNCI2015_001` (Faller et al., 2012), `Shin2017A`, and `Zhou2016` (Zhou et al., 2016), with all configured subjects for each dataset. The corresponding subject counts were 9, 14, 12, 29, and 4, respectively.

All reported experiments used a fixed cross-session or cross-run evaluation protocol. For each subject, when multiple sessions were available, the final session was reserved for testing and all preceding sessions were used for training. When multiple sessions were unavailable but multiple runs were present, the final run was held out instead. If neither structure was available, a stratified 50/50 split was used as a fallback. The data split was fixed across generator seeds.

All representation statistics, normalization parameters, model parameters, source-distribution statistics, and downstream classifiers were estimated using training data only. Exact duplicate trials between the training and held-out sets were explicitly checked, and execution was terminated if train–test leakage was detected.

All generative experiments were repeated using three generator seeds,

$$s_{\mathrm{gen}} \in \{1, 2, 3\}, \tag{8}$$

and CAS evaluation used three independent classifier seeds,

$$s_{\mathrm{CAS}} \in \{9001, 9002, 9003\}. \tag{9}$$

A.2 EEG preprocessing

Only EEG channels were retained; non-EEG channels were discarded. Signals were converted to microvolts, band-pass filtered from 4 to 38 Hz, and resampled to 128 Hz (Gramfort et al., 2013). Event-aligned trials were then extracted using the complete event interval provided by each dataset.

Before GVD construction, every EEG trial was standardized independently for each channel over the complete temporal duration of the trial. For channel c,

$$\widetilde{u}_{c,t} = \frac{u_{c,t} - \mu_c}{\max(\sigma_c, 10^{-6})}, \tag{10}$$

where

$$\mu_c = \frac{1}{T} \sum_{t=1}^{T} u_{c,t}, \tag{11}$$

and

$$\sigma_c = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (u_{c,t} - \mu_c)^2}. \tag{12}$$

Standardization was performed over time within each channel and trial.

Two motor-imagery classes were retained for each binary experiment.

A.3 GRAPH-VARIATE DYNAMIC CONNECTIVITY CONSTRUCTION

Each standardized EEG trial was transformed into a high-resolution graph-variate dynamic (GVD) connectivity trajectory. Let

$$U = \begin{bmatrix} u_1 & \cdots & u_T \end{bmatrix} \in \mathbb{R}^{d \times T}, \tag{13}$$

where

$$u_t \in \mathbb{R}^d \tag{14}$$

is the vector of standardized channel amplitudes at EEG sample t.

Stable trial-level support. The long-term support matrix was the signed whole-trial Pearson correlation matrix,

$$W = \frac{1}{T} U U^\top. \tag{15}$$

Because every channel has already been centered and normalized over the complete trial, Equation 15 is the signed channel correlation matrix under the population-standard-deviation convention used by the implementation.

No absolute-value operation, additive ridge, or nearest-SPD projection was applied to the support matrix in the reported configuration. Because every channel is standardized over the trial, W has unit diagonal and $\operatorname{rank}(W) = \operatorname{rank}(U)$, which is full whenever the $T \gg d$ samples span $\mathbb{R}^d$. The strict positive definiteness of every W is verified in double precision before construction of the trajectory, and execution terminates if the check fails. No trial in any dataset failed.

Sample-resolution instantaneous interaction. At every original EEG sample, the instantaneous interaction matrix was defined as the rank-one outer product

$$J_t = u_t u_t^\top. \tag{16}$$

The sample-resolution graph-variate matrix was then

$$\Delta_t = W \odot J_t, \tag{17}$$

where $\odot$ denotes the Hadamard product.

Using

$$D_t = \operatorname{diag}(u_t), \tag{18}$$

Equation 17 can equivalently be written as

$$\Delta_t = D_t W D_t. \tag{19}$$

Temporal aggregation. Each trial was partitioned into

$$B = 100 \tag{20}$$

disjoint full-coverage temporal bins. Their boundaries were

$$e_b = \left\lfloor \frac{bT}{B} \right\rfloor, \qquad b = 0, \ldots, B, \tag{21}$$

so that every original EEG sample belongs to exactly one bin and the complete trial is covered.

Let

$$\mathcal{I}_b = \{e_{b-1}, \ldots, e_b - 1\} \tag{22}$$

denote the sample indices in bin b. The reported GVD trajectory was obtained by averaging the already Hadamard-modulated sample-resolution matrices (this is $\overline{\Delta}_b$ in the main text; we drop the bar in the appendix):

$$\Delta_b = \frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} \left(W \odot u_t u_t^\top \right). \tag{23}$$

Since W is constant within a trial, this is equivalently

$$\Delta_b = W \odot \left(\frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} u_t u_t^\top \right), \tag{24}$$

which is the form used in the implementation.

No local re-centering was performed inside a temporal bin, because centering and scaling had already been carried out over the complete trial. Likewise, no covariance ridge or additive GVD ridge was used.

The resulting matrices were symmetrized numerically and their minimum eigenvalues were evaluated in double precision. If a trajectory failed the strict positive-definiteness check, execution terminated rather than applying an additive ridge or post-hoc nearest-SPD correction.

The canonical GVD representation used throughout the benchmark is therefore

$$\boxed{\Delta_b = \frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} W \odot u_t u_t^\top}. \tag{25}$$

A.4 Log-Euclidean trajectory representation

Each SPD GVD matrix was mapped to the log-Euclidean chart (Arsigny et al., 2007):

$$z_b = \operatorname{svec}(\log \Delta_b) \in \mathbb{R}^m, \tag{26}$$

where

$$m = \frac{d(d+1)}{2}. \tag{27}$$

The svec operator contains the lower-triangular entries of a symmetric matrix, with off-diagonal elements multiplied by $\sqrt{2}$. This preserves the Frobenius inner product under vectorization.

The complete trajectory was stacked as

$$Z = \begin{bmatrix} z_1^\top \\ \vdots \\ z_B^\top \end{bmatrix} \in \mathbb{R}^{B \times m}. \tag{28}$$

For every log-svec feature j, the normalization statistics were estimated using the real training trajectories only and pooled across training trials and temporal bins:

$$\mu_j = \frac{1}{N_{\mathrm{tr}} B} \sum_{n=1}^{N_{\mathrm{tr}}} \sum_{b=1}^{B} Z_{n,b,j}, \tag{29}$$

and

$$\sigma_j = \mathrm{Std}_{n,b}\left[Z_{n,b,j}\right]. \tag{30}$$

A minimum scale of 10^{-6} was used,

$$\sigma_j \leftarrow \max(\sigma_j, 10^{-6}), \tag{31}$$

and the standardized coordinates were

$$\widetilde{Z}_{n,b,j} = \frac{Z_{n,b,j} - \mu_j}{\sigma_j}. \tag{32}$$

The same train-estimated statistics were used for all direct GVD-space generators.

Full spectral representation. For GVD-CFM, a full orthonormal DCT-II was subsequently applied along the temporal axis:

$$\bar{Z} = C_B \widetilde{Z}, \qquad C_B^\top C_B = C_B C_B^\top = I_B. \tag{33}$$

All $B = 100$ DCT modes were retained. The DCT therefore performs no dimensionality reduction; it is an invertible orthogonal reparameterization of the complete temporal trajectory.

The inverse transformation is

$$\widetilde{Z} = C_B^\top \bar{Z}, \tag{34}$$

$$Z = \widetilde{Z} \odot \sigma + \mu, \tag{35}$$

$$\Delta_b = \exp\left(\operatorname{svec}^{-1}(z_b)\right). \tag{36}$$

For generated trajectories, extreme log-eigenvalues were stabilized before matrix exponentiation. The lower and upper bounds were estimated exclusively from the real training log-spectrum using the 0.001 and 0.999 quantiles, respectively, and each bound was expanded by a margin of 0.5. This stabilization was applied only in the log domain and did not modify the real training or held-out trajectories.

A.5 GENERATIVE MODELS

The benchmark contained eight generators:

1. GVD-CFM;
2. GVD-cVAE;
3. GVD-DDPM;
4. Window-DIFFEO-CFM (Collas et al., 2025);
5. JET (Wang et al., 2026);
6. Vanilla Diffusion / DDIM (Ho et al., 2020; Song et al., 2021a);
7. EEGGAN-2025 (Williams et al., 2025; 2023);
8. conditional VAE (Kingma and Welling, 2014; Sohn et al., 2015);

The first four models generate GVD trajectories directly. The remaining four generate raw multichannel EEG, after which each synthetic EEG trial is independently transformed into a GVD trajectory using its own support matrix W and its own window covariances J_b. The quantitative tables report GVD-CFM and seven comparators.

All trainable generators were trained for 1000 epochs. The training batch size was scaled with the number of available training trials:

$$N_{\text{batch}} = \max\left(1, \min\left(N, \text{round}\left[64\frac{N}{1000}\right]\right)\right). \tag{37}$$

This gives a batch size of 64 for 1000 training trials and keeps the number of optimizer batches per epoch approximately constant across datasets.

A.6 GVD-CFM

GVD-CFM generates complete dynamic GVD trajectories in the composite log-Euclidean/DCT coordinate system. The network operates on

$$\bar{Z} \in \mathbb{R}^{B\times m}, \tag{38}$$

where each of the B tokens corresponds to one DCT mode and each token contains the m log-`svec` connectivity coordinates.

The velocity network consists of a spectral Transformer and a small temporal branch that share the conditioning vector. The spectral branch is an AdaLN Transformer (Vaswani et al., 2017; Peebles and Xie, 2023) with model dimension 256, 8 attention heads, and 6 Transformer blocks acting on the B DCT tokens. Each block contains multi-head self-attention followed by a feed-forward network with expansion factor 4 and GELU activation (Hendrycks and Gimpel, 2016). The input and output projections map between the m-dimensional GVD coordinate and the 256-dimensional Transformer state.

The temporal branch receives $C_B^\top z_\tau$, the inverse-DCT view of the current state, as B window tokens. It projects each window to the same 256-dimensional width, adds a window-position embedding, and applies 2 AdaLN Transformer blocks of the same width and number of heads, conditioned on the same flow-time and class vector. Its output is multiplied by C_B to return to DCT-token alignment and fused with the spectral hidden states through the tokenwise sigmoid gate of Equation 5 before the shared output projection. No auxiliary loss is applied to the temporal branch. Including both branches, the network has between 9.89 and 10.18 million parameters, depending on the number of EEG channels.

DCT-mode identity is represented using a continuous learned embedding. Flow time is encoded using 16 random Fourier frequencies (Tancik et al., 2020) followed by two fully connected SiLU layers. The class condition is represented by a learned embedding and added to the flow-time representation before adaptive layer-normalization modulation.

Source distribution. The default GVD-CFM source is an isotropic standard normal distribution in the complete DCT trajectory space. Specifically, for each generated trajectory,

$$z_0 \sim \mathcal{N}(0, I), \tag{39}$$

where z_0 has the same dimensionality as the vectorized DCT representation of the target GVD trajectory. Equivalently,

$$z_0 = \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I). \tag{40}$$

The source distribution is independent of the class label and is not estimated from the training data. Class information is instead supplied to the conditional flow model through the class-conditioning mechanism. Thus, all classes share the same standard Gaussian source, while the learned conditional vector field transports samples toward the corresponding class-conditional distribution of GVD trajectories.

Minibatch coupling. Source and target samples are paired classwise. For every class present in a training minibatch, a fresh standard-normal draw of the same size is matched to that class's target trajectories by an entropic optimal-transport plan computed with Sinkhorn iterations (Cuturi, 2013) on the squared Euclidean distance between complete DCT-coordinate trajectories, and the minibatch is re-paired according to this plan (Tong et al., 2024). Since the orthonormal DCT is an isometry, this is the squared distance between standardized log-Euclidean trajectories. Coupling within class ensures that every source sample is paired with a target of the label on which the velocity is conditioned.

Flow-matching objective. For each training example,

$$\tau \sim \mathcal{U}(0, 1), \tag{41}$$

(z_0, z_1) is drawn from the classwise minibatch coupling, and a straight conditional probability path is used:

$$z_\tau = (1 - \tau)z_0 + \tau z_1, \tag{42}$$

with target velocity

$$u_\tau = z_1 - z_0. \tag{43}$$

The model is optimized using an ℓ_2 conditional flow-matching objective,

$$\mathcal{L}_{\mathrm{CFM}} = \|v_\theta(z_\tau, \tau, y) - u_\tau\|_2^2 . \tag{44}$$

AdamW (Loshchilov and Hutter, 2019) is used with learning rate 5×10^{-4} and weight decay 10^{-4}. Gradient norms are clipped at 1.0. Training minibatches are sampled with inverse class-frequency weighting.

Sampling. Sampling starts from the standard-normal source and integrates the learned velocity field from flow time 0 to 1. The default sampler is fourth-order Runge–Kutta with 50 uniform integration steps, so each trajectory requires 200 evaluations of the velocity network. Sampling is performed in batches of at most 2048 trajectories.

After integration, DCT coefficients are transformed back to temporal log-`svec` coordinates using the inverse DCT, reversed through the training-set affine standardization, stabilized in the log-spectrum, and exponentiated to obtain SPD GVD trajectories.

A.7 Direct GVD-space control models

Three additional generators operate directly on the same GVD targets as GVD-CFM: GVD-cVAE, GVD-DDPM, and Window-DIFFEO-CFM.

These controls never generate raw EEG. They therefore isolate the contribution of the generative model from that of the GVD representation itself. All three use the same training GVD matrices, log-Euclidean $\mathrm{svec}(\log(\cdot))$ representation, training-set standardization, generated log-spectrum stabilization, and SPD decoding procedure as GVD-CFM. However, unlike GVD-CFM, they operate directly on the temporal sequence of standardized log-Euclidean GVD coordinates and do not transform the trajectories into DCT modes.

GVD-cVAE and GVD-DDPM receive exactly the same training-set standardized coordinates as GVD-CFM. Thus, each trial is represented directly as a sequence of standardized log-`svec` GVD vectors, and no temporal DCT or inverse-DCT operation is used by either baseline.

A.7.1 GVD-cVAE

GVD-cVAE is a conditional VAE (Kingma and Welling, 2014; Sohn et al., 2015) defined on the complete temporal GVD trajectory in standardized log-Euclidean coordinates.

The sequence of temporal log-`svec` vectors is flattened into a single trial-level representation and concatenated with a one-hot class vector. The encoder contains two fully connected layers of width 512 with GELU activations. The latent representation has dimension 64 and is parameterized by separate mean and log-variance heads.

The decoder concatenates the sampled latent representation with the one-hot class vector and applies two width-512 GELU layers followed by a linear output layer spanning the complete temporal GVD trajectory.

The objective is

$$\mathcal{L}_{\mathrm{GVD-cVAE}} = \mathcal{L}_{\mathrm{MSE}} + 10^{-3}\mathcal{L}_{\mathrm{KL}}. \tag{45}$$

The model uses AdamW with learning rate 10^{-3}, weight decay 10^{-4}, and gradient-norm clipping at 5.0.

At generation time,

$$z \sim \mathcal{N}(0, I_{64}), \tag{46}$$

is sampled and passed to the class-conditional decoder. The generated temporal log-Euclidean trajectory is inverse-standardized and mapped directly back to a sequence of SPD GVD matrices through the same matrix-exponential decoding path used by GVD-CFM.

A.7.2 GVD-DDPM

GVD-DDPM is an ϵ-prediction diffusion model (Ho et al., 2020) operating directly on the complete temporal sequence of standardized log-Euclidean GVD coordinates.

Its noise-prediction network is a pre-norm Transformer encoder with width 256, 8 attention heads, and 6 layers, matching the width, number of heads, and depth of the GVD-CFM spectral Transformer. Each layer contains multi-head self-attention followed by a feed-forward network with expansion factor 4 and GELU activation, without dropout. The Transformer tokens correspond to temporal GVD windows. Each token is a learned projection of one standardized log-`svec` window, to which a Fourier embedding of the diffusion step, a window-position embedding, and a learned class embedding are added. Consequently, the model operates directly on the temporal log-`svec` trajectory.

The diffusion process contains 200 steps with a linear variance schedule

$$\beta_1 = 10^{-4}, \qquad \beta_{200} = 2 \times 10^{-2}. \tag{47}$$

At a randomly selected diffusion step ℓ, noise

$$\epsilon \sim \mathcal{N}(0, I) \tag{48}$$

is added according to the standard forward diffusion process. The network is trained with

$$\mathcal{L}_{\mathrm{DDPM}} = \|\epsilon_\theta(x_\ell, \ell, y) - \epsilon\|_2^2 . \tag{49}$$

AdamW uses learning rate 2×10^{-4}, weight decay 10^{-4}, and gradient clipping at 5.0. Training and sampling are performed in single precision.

Generation uses full ancestral DDPM sampling over all 200 diffusion steps. The resulting temporal coordinates are inverse-standardized and decoded directly through the log-Euclidean inverse map to obtain a sequence of SPD GVD matrices.

A.7.3 WINDOW-DIFFEO-CFM

Window-DIFFEO-CFM (Collas et al., 2025) is a per-window flow-matching control defined in the same standardized log-Euclidean GVD space. Unlike GVD-CFM and the other whole-trajectory controls, it does not model the complete temporal trajectory jointly and has no communication between different temporal windows.

Each standardized log-`svec` GVD window is treated as an independent training sample. The network is a conditional MLP with one hidden layer of width 128 and SELU activation. Its input contains:

1. the current noisy GVD window;
2. a one-hot class vector;
3. the normalized temporal window location $\xi_b \in [0, 1]$;
4. the flow time.

For each window,

$$x_0 \sim \mathcal{N}(0, I), \tag{50}$$

and the conditional interpolation path is

$$x_\tau = (1 - \tau)x_0 + \tau x_1. \tag{51}$$

The model is trained with the MSE velocity objective

$$\mathcal{L}_{\mathrm{Window-CFM}} = \|v_\theta(x_\tau, \tau, y, \xi_b) - (x_1 - x_0)\|_2^2 . \tag{52}$$

AdamW uses learning rate 10^{-3} with zero weight decay and gradient clipping at 5.0. Sampling uses RK4 with 50 steps independently for each temporal window.

This baseline therefore tests whether matching the marginal distribution of each temporal GVD window independently is sufficient, in contrast to jointly modeling the full temporal trajectory with cross-window context.

A.8 RAW-EEG GENERATORS

The remaining four methods generate raw multichannel EEG. Their generated signals are not compared directly with GVD-CFM in raw-signal space. Instead, every synthetic EEG trial is passed through the same GVD construction used for real data.

For every generated trial $U^{(g)}$, its own support matrix is computed:

$$W^{(g)} = \operatorname{corr}\left(U^{(g)}\right), \tag{53}$$

and its own window covariance sequence

$$J_1^{(g)}, \ldots, J_B^{(g)} \tag{54}$$

is constructed. The final generated GVD trajectory is

$$\Delta_b^{(g)} = W^{(g)} \odot J_b^{(g)}. \tag{55}$$

Thus, no real-data W or J_b is reused for a raw-EEG baseline.

A.8.1 JET

JET (Wang et al., 2026) is evaluated using the official Y-Research-SBU implementation and the JiT-B/16 configuration. Raw EEG trials are padded at the input boundary to a length divisible by a patch size of 200 samples. The number of input tokens is therefore determined by the number of EEG channels multiplied by the number of temporal patches.

The benchmark retains the official JET denoiser, objective, EMA updates, and sampling equations. The configuration uses a class-conditional model, label dropout probability 0.1, $P_{\mathrm{mean}} = -0.8$, $P_{\mathrm{std}} = 0.8$, Gaussian noise, and the mixed loss configuration from the release. The enabled auxiliary loss terms are statistical loss with weight 1.0, total-variation loss with weight 0.1, and correlation loss with weight 0.1; the STFT loss weight is 0.

Two EMA decay factors, 0.9999 and 0.9996, are retained from the implementation. Sampling uses the Heun solver with 50 steps.

The JET base learning rate is 5×10^{-5} and is scaled by

$$\mathrm{lr} = 5 \times 10^{-5} \frac{N_{\mathrm{batch}}}{256}. \tag{56}$$

AdamW uses $\beta = (0.9, 0.95)$ and zero weight decay.

A.8.2 Vanilla Diffusion / DDIM

The diffusion baseline follows the released Song, Meng, and Ermon DDPM/DDIM implementation (Ho et al., 2020; Song et al., 2021a). The original model uses two-dimensional convolution for images; only the convolutional operators are ported to one-dimensional convolution so that the model operates directly on multichannel EEG.

No class embedding is introduced because the released architecture is unconditional. Instead, one diffusion model is trained separately for each class.

The forward diffusion process uses 1000 steps and a linear beta schedule from

$$10^{-4} \quad \text{to} \quad 2 \times 10^{-2}. \tag{57}$$

The U-Net uses base width 128, channel multipliers

$$(1, 2, 2, 2), \tag{58}$$

two residual blocks per resolution, and dropout 0.1.

Training uses standard DDPM noise prediction with learning rate 2×10^{-4}. Sampling uses the generalized DDIM sampler with 50 sampling steps and

$$\eta = 0, \tag{59}$$

corresponding to deterministic DDIM sampling.

Input EEG is scaled to the training-derived $[-1, 1]$ range before diffusion and transformed back afterward.

A.8.3 EEGGAN-2025

EEGGAN-2025 (Williams et al., 2025; 2023) uses the official AutoResearch EEG-GAN generator, discriminator, and `GANTrainer.batch_train` update procedure.

The benchmark reproduces the source preprocessing in memory. EEG is arranged as trial × time × channel and normalized using global training-set min–max normalization. A one-step class-condition prefix is prepended and repeated over channels.

The source configuration uses temporal patch size 20, hidden dimension 16, four model layers, latent dimension 128, five critic iterations, and gradient-penalty coefficient 10. Generator and discriminator learning rates are both 10^{-4}.

The released Adam optimizer settings

$$\beta = (0, 0.9) \tag{60}$$

are retained. A compatibility wrapper is used only to express both beta values as floating-point numbers in recent PyTorch versions.

Generated normalized EEG is mapped back to the original training-data amplitude range before GVD construction.

A.8.4 Conditional raw-EEG VAE

The raw-EEG conditional VAE (Kingma and Welling, 2014; Sohn et al., 2015) operates on a flattened complete EEG trial.

The encoder concatenates the flattened signal with a learned class embedding and uses two fully connected layers of width 512 with SiLU activations. The latent dimension is 64. The decoder uses two width-512 SiLU layers before mapping back to the complete multichannel EEG trial.

The objective is

$$\mathcal{L}_{\mathrm{cVAE}} = \mathcal{L}_{\mathrm{MSE}} + 10^{-3}\mathcal{L}_{\mathrm{KL}}. \tag{61}$$

AdamW is used with learning rate 10^{-3} and weight decay 10^{-4}.

Synthetic EEG is generated from

$$z \sim \mathcal{N}(0, I_{64}), \tag{62}$$

conditioned on the desired class.

A.9 Full-dimensional GVD Fréchet distance

All generative models are compared in the same GVD feature space (Dowson and Landau, 1982; Heusel et al., 2017). No PCA or other dimensionality-reduction transform is applied.

For a trajectory

$$Z = [z_1, \ldots, z_B], \tag{63}$$

the Fréchet feature contains both all log-`svec` positions and all first temporal increments:

$$\phi(Z) = [z_1, \ldots, z_B, z_2 - z_1, \ldots, z_B - z_{B-1}]\,. \tag{64}$$

The feature dimension is therefore

$$(2B - 1)m. \tag{65}$$

For two sets of features with Gaussian moments (μ_1, Σ_1) and (μ_2, Σ_2), the Fréchet distance is

$$d_F = \|\mu_1 - \mu_2\|_2^2 + \mathrm{tr}\left(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})^{1/2}\right). \tag{66}$$

The full feature covariance is not explicitly constructed. The Bures cross term is evaluated through the algebraically equivalent sample-space nuclear norm, allowing the exact full-coordinate distance to be computed without reducing feature dimension.

Distances are computed separately by class and macro-averaged. Real-training to held-out-real distance is retained as the real–real reference. Relative Fréchet is computed by dividing the generated-to-held-out-real distance by the corresponding real-training-to-held-out-real reference within each class before averaging.

A.10 Classification accuracy score

Class-conditional utility is assessed using a train-synthetic-test-real protocol (Ravuri and Vinyals, 2019; Esteban et al., 2017).

Each GVD trajectory is encoded as the complete flattened log-`svec` sequence. A ridge of $10^{-6}I$ is added before the logarithmic chart for the classifier representation. No dimensionality reduction is applied.

The shared CAS classifier is a two-layer MLP. Each hidden layer has width 256 and consists of

$$\text{Linear} \to \text{BatchNorm} \to \text{GELU} \to \text{Dropout}(0.1). \tag{67}$$

The classifier uses AdamW with learning rate 10^{-3} and weight decay 10^{-4}. Training examples are class-balanced by weighted sampling and weighted cross-entropy. The maximum batch size is 2048.

Training runs for at most 120 epochs. The state with the lowest training loss is retained and optimization stops after 15 epochs without improvement.

For each generator, the classifier is trained using synthetic trajectories and tested only on the held-out real partition. ROC-AUC and weighted F1 are reported. A classifier trained on real training trajectories and evaluated on the same held-out test set provides the real-data reference.

A.11 NOVELTY AND MEMORIZATION ANALYSIS

The novelty analysis is performed in the complete standardized flattened log-`svec` trajectory space. Let

$$x_i \in \mathbb{R}^{Bm}$$

denote the standardized flattened feature vector of a real training trajectory and let

$$g_j \in \mathbb{R}^{Bm}$$

denote the corresponding representation of a generated trajectory, where $m = d(d+1)/2$ is the number of log-`svec` coordinates per window. All nearest-neighbor calculations described below are performed class-conditionally.

Exact copying is tested by rounding complete unstandardized log-`svec` feature vectors to six decimal places and comparing their hashes with those of the real training set.

Nearest-neighbor diagnostics include real leave-one-out distances, generated-to-training distances, generated leave-one-out distances, and held-out-real-to-training distances. A generated sample is marked as a near copy when its distance to the closest same-class real training sample is below the 5th percentile of the corresponding real leave-one-out nearest-neighbor distribution.

Manifold precision. Following the k-nearest-neighbor support construction used in the precision–recall family of generative-model metrics, we use $k = 5$. For each real training sample x_i, let

$$\rho_i^{(r)} = d_k(x_i,\, \{x_\ell : \ell \neq i,\ y_\ell = y_i\})$$

be the Euclidean distance to its kth nearest same-class real training neighbor.

A generated sample g is counted as lying on the estimated real-data manifold when it falls inside at least one same-class real support ball:

$$\mathbb{I}_{\mathrm{prec}}(g) = \mathbb{I}\left[\exists\, i : y_i = y_g,\ \|g - x_i\|_2 \leq \rho_i^{(r)}\right].$$

The reported manifold precision is

$$\mathrm{Precision} = \frac{1}{N_g}\sum_{j=1}^{N_g} \mathbb{I}_{\mathrm{prec}}(g_j).$$

Higher values indicate that a larger fraction of generated trajectories lies inside the empirical support of the real training distribution:

$$\mathrm{Precision} \uparrow .$$

Diversity ratio. Diversity is measured from leave-one-out nearest-neighbor spacing. For every real training sample,

$$d_i^{(r)} = \min_{\ell \neq i,\ y_\ell = y_i} \|x_i - x_\ell\|_2,$$

and for every generated trajectory,

$$d_j^{(g)} = \min_{\ell \neq j,\ y_\ell = y_j} \|g_j - g_\ell\|_2.$$

The diversity ratio is

$$R_{\mathrm{div}} = \frac{\mathrm{median}_j\, d_j^{(g)}}{\mathrm{median}_i\, d_i^{(r)}}.$$

The ideal value is therefore one:

$$\mathrm{Diversity} \to 1.$$

Values below one indicate that generated trajectories are more tightly clustered than the real training trajectories, whereas values above one indicate that generated samples are more dispersed.

Training coverage. For each generated trajectory, let

$$n(g_j) = \arg \min_{i:\, y_i = y_j} \|g_j - x_i\|_2$$

denote its nearest same-class real training trajectory. Training coverage is defined as the fraction of real training trajectories that are selected as the nearest neighbor of at least one generated trajectory:

$$\text{Coverage} = \frac{|\{n(g_j) : j = 1, \ldots, N_g\}|}{N_{\text{train}}}.$$

Coverage therefore measures how broadly the generator distributes samples across the empirical training set rather than repeatedly concentrating around a small subset:

$$\text{Coverage} \uparrow .$$

A.12 Temporal-coherence analysis

Temporal coherence is measured in log-`svec` coordinates after removing the temporal mean of every feature within each trajectory. For trajectory i, let

$$z_{i,b} \in \mathbb{R}^m, \qquad b = 1, \ldots, B,$$

denote its log-`svec` coordinates and define

$$\bar{z}_i = \frac{1}{B} \sum_{b=1}^{B} z_{i,b}, \qquad \widetilde{z}_{i,b} = z_{i,b} - \bar{z}_i.$$

Removing the within-trial temporal mean prevents the stable connectivity level of a trial from dominating the dynamic comparison.

For real and generated trajectories, population time-by-time correlation matrices are computed from the centred trajectories. Their agreement is reported as the Pearson correlation between corresponding upper-triangular entries and as their mean absolute error.

Temporal autocorrelation is additionally evaluated at positive lags up to 16 windows. The real and generated lag-autocorrelation curves are compared using Pearson correlation and mean absolute error.

A.13 Dynamic-signal measures

Two permutation-based diagnostics test whether generated trajectories contain nontrivial temporal dynamics.

For a trajectory

$$Z_i = [z_{i,1}, \ldots, z_{i,B}], \qquad z_{i,b} \in \mathbb{R}^m,$$

define its temporal mean

$$\bar{z}_i = \frac{1}{B} \sum_{b=1}^{B} z_{i,b}.$$

Dynamic energy. The dynamic energy of trajectory i is

$$E_{\text{dyn}}(Z_i) = \frac{1}{Bm} \sum_{b=1}^{B} \|z_{i,b} - \bar{z}_i\|_2^2 .$$

This measures the absolute amount of within-trajectory temporal variation after removing the stable temporal mean.

The table reports the generated-to-real median ratio

$$R_E = \frac{\text{median}_{i \in \text{gen}}\, E_{\text{dyn}}(Z_i)}{\text{median}_{i \in \text{real}}\, E_{\text{dyn}}(Z_i)}.$$

Consequently, the ideal value is

$$\text{Energy} \to 1.$$

Values below one indicate insufficient temporal variation, whereas values above one indicate excessive temporal variation relative to the real trajectories.

Dynamic fraction. The total log-svec energy of trajectory i is

$$E_{\mathrm{tot}}(Z_i) = \frac{1}{Bm} \sum_{b=1}^{B} \|z_{i,b}\|_2^2.$$

The fraction of trajectory energy attributable to temporal variation is

$$F_{\mathrm{dyn}}(Z_i) = \frac{E_{\mathrm{dyn}}(Z_i)}{\max(E_{\mathrm{tot}}(Z_i), \varepsilon)},$$

where ε is a numerical safeguard.

The reported dynamic-fraction ratio is

$$R_F = \frac{\mathrm{median}_{i \in \mathrm{gen}} F_{\mathrm{dyn}}(Z_i)}{\mathrm{median}_{i \in \mathrm{real}} F_{\mathrm{dyn}}(Z_i)}.$$

Again, the ideal value is

$$\text{Dyn. fraction} \to 1.$$

This quantity differs from dynamic energy because it normalizes temporal variation by the overall magnitude of the trajectory.

Adjacent-window step energy. Temporal smoothness is quantified by the mean squared displacement between successive windows:

$$E_{\mathrm{adj}}(Z_i) = \frac{1}{(B-1)m} \sum_{b=1}^{B-1} \|z_{i,b+1} - z_{i,b}\|_2^2 .$$

The complementary diagnostic table reports the generated-to-real ratio

$$R_{\mathrm{adj}} = \frac{\mathrm{median}_{i \in \mathrm{gen}} E_{\mathrm{adj}}(Z_i)}{\mathrm{median}_{i \in \mathrm{real}} E_{\mathrm{adj}}(Z_i)},$$

so the target value is

$$\text{Adjacent} \to 1.$$

Values below one correspond to trajectories that are smoother than the real data, while values above one indicate excessive frame-to-frame variation.

A.14 Efficiency measurements

Training time and generation time are measured separately using synchronized CUDA wall-clock timing.

Generation efficiency is measured using a batch of exactly 32 model-native samples. One warm-up call is performed before timing, and the reported latency is based on repeated synchronized measurements.

For direct GVD-space models, the measured generation latency already produces a GVD trajectory.

For raw-EEG generators, raw EEG generation and the subsequent EEG-to-GVD transformation are timed separately. The latter contains:

1. per-trial EEG standardization;
2. computation of the generated trial support W;
3. computation of all window covariances J_b;
4. construction of $\Delta_b = W \odot J_b$.

The benchmark therefore reports both native raw-EEG generation latency and end-to-end latency required to obtain a GVD trajectory.

Peak allocated and reserved CUDA memory are recorded during training and sampling.

For the full GVD-CFM, 1000 training epochs take 304 s on average across the five main datasets and three seeds (from 128 s on Zhou2016 to 470 s on Shin2017A), generating a batch of 32 trajectories with 50 RK4 steps takes 3.92 s, and peak allocated training memory is 1.77 GB on average.

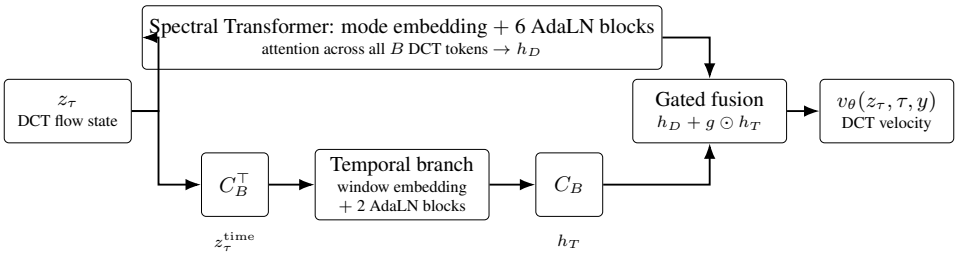


Figure 3: Velocity network of GVD-CFM. Only the DCT state z_τ is integrated. The spectral Transformer attends over all DCT modes. The temporal branch reads $z_\tau^{\text{time}} = C_B^\top z_\tau$, and its output is returned to DCT alignment by C_B before gated fusion. Both are conditioned on c_τ through adaptive LayerNorm.

A.15 Implementation

The benchmark is implemented in PyTorch (Paszke et al., 2019) and requires a CUDA-capable GPU. The reference runner was optimized for an NVIDIA A100.

Automatic mixed precision is enabled. BFloat16 is used when supported by the GPU and FP16 is used otherwise. TF32 matrix multiplication and cuDNN benchmarking are enabled. GVD log-Euclidean matrix logarithms and exponentials use GPU eigendecomposition (Golub and Loan, 2013) where available, while the SPD calculations themselves are retained in double precision where required for numerical stability.

Fused AdamW (Loshchilov and Hutter, 2019) is used when supported. The GVD-CFM and compatible baseline networks are optionally compiled using `torch.compile` with the `reduce-overhead` mode. Training arrays are kept GPU-resident where practical.

External baselines are loaded from their public source repositories. The benchmark uses the official JET repository, the official Song–Meng–Ermon DDIM release, and the AutoResearch EEG-GAN repository. Compatibility changes required by modern Python or PyTorch are restricted to syntax, datatype, and execution issues; model architectures and reported training objectives are not intentionally altered.

A.16 Inverse of the composite chart

With Ψ as defined in Section 4.2.1, its inverse is

$$\widetilde{Z} = C_B^\top \bar{Z}, \tag{68}$$

$$Z = \widetilde{Z} \odot \mathbf{1}\sigma^\top + \mathbf{1}\mu^\top, \tag{69}$$

$$\overline{\Delta}_b = \exp\left(\mathrm{svec}^{-1}(z_b)\right). \tag{70}$$

A.17 Velocity network and training path

Each DCT token of Section 4.2.2 enters the spectral Transformer with initial hidden representation

$$h_k^{(0)} = P_{\text{in}} \bar{z}_k + e_k^{\text{mode}}, \tag{71}$$

where P_{in} is a learned projection and e_k^{mode} identifies the DCT mode. Each window of the time view $z^{\text{time}} = C_B^\top \bar{Z}$ enters the temporal branch as

$$r_b^{(0)} = P_{\text{in}}^{\text{time}} z_b^{\text{time}} + e_b^{\text{win}}, \tag{72}$$

where e_b^{win} identifies the window position. The model is class-conditional, with conditioning vector

$$c_\tau = e_{\text{flow}}(\tau) + e_{\text{class}}(y), \tag{73}$$

where $\tau \in [0, 1]$ is flow time. Flow time is encoded using Fourier features followed by an MLP (Tancik et al., 2020).

Writing S_θ for the spectral Transformer and T_θ for the temporal branch, both conditioned on c_τ, a single forward pass computes (Figure 3)

$$h_D = S_\theta(\bar{Z}, c_\tau), \qquad h_T = C_B\, T_\theta\Big(C_B^\top \bar{Z}, c_\tau\Big), \tag{74}$$

$$v_\theta(\bar{Z}, \tau, y) = P_{\text{out}}\left(h_D + \sigma(W_g[h_D; h_T] + b_g) \odot h_T\right), \tag{75}$$

which is the joint velocity

$$v_\theta : \mathbb{R}^{B\times m} \times [0,1] \times \mathcal{Y} \longrightarrow \mathbb{R}^{B\times m} \tag{76}$$

of the complete trajectory.

During training of Section 4.2.3, we sample

$$z_0 \sim \mathcal{N}(0, I), \qquad z_1 \sim \Psi_{\#} q(\cdot \mid y), \qquad \tau \sim \mathcal{U}[0,1]. \tag{77}$$

The pair (z_0, z_1) is then re-drawn from the classwise minibatch Sinkhorn coupling π_y. We use the linear conditional path

$$z_\tau = (1-\tau)z_0 + \tau z_1, \qquad u_\tau = z_1 - z_0. \tag{78}$$

To generate a trajectory from class y we sample $z(0) \sim \mathcal{N}(0, I)$ and solve

$$\frac{\mathrm{d}z(\tau)}{\mathrm{d}\tau} = v_{\theta^\star}(z(\tau), \tau, y), \qquad \tau \in [0,1]. \tag{79}$$

A.18 Training and sampling algorithms

Algorithm 1 GVD-CFM training

Require: Label distribution $\pi_{\mathcal{Y}}$, training distribution q, and diffeomorphism Ψ
Ensure: Trained parameters $\theta^\star$

Initialize θ
while not converged **do**
  Sample a class-weighted minibatch $\{(\overline{\mathbf{\Delta}}^{(i)}, y^{(i)})\}_{i=1}^N$ from q
  Set $z_1^{(i)} \leftarrow \Psi(\overline{\mathbf{\Delta}}^{(i)})$
  Sample $z_0^{(i)} \sim \mathcal{N}(0, I)$
  for each class y in the minibatch **do**
    Re-pair $\{z_0^{(i)}\}$ with $\{z_1^{(i)} : y^{(i)} = y\}$ using the Sinkhorn plan π_y
  end for
  Sample $\tau^{(i)} \sim \mathcal{U}[0,1]$
  Set $z_\tau^{(i)} \leftarrow (1-\tau^{(i)})z_0^{(i)} + \tau^{(i)} z_1^{(i)}$
  Set $u_\tau^{(i)} \leftarrow z_1^{(i)} - z_0^{(i)}$
  Evaluate $v_\theta(z_\tau^{(i)}, \tau^{(i)}, y^{(i)})$
  Compute

$$\mathcal{L} \leftarrow \frac{1}{NBm} \sum_{i=1}^{N} \left\| v_\theta(z_\tau^{(i)}, \tau^{(i)}, y^{(i)}) - u_\tau^{(i)} \right\|_F^2$$

  Update $\theta \leftarrow \mathrm{OptimizerStep}(\theta, \nabla_\theta \mathcal{L})$
end while
return $\theta^\star$

Algorithm 2 GVD-CFM sampling

Require: Class label y, parameters $\theta^\star$, integration steps L, and diffeomorphism Ψ
Ensure: Generated trajectory $\widehat{\mathbf{\Delta}}$

Set $h \leftarrow 1/L$
Sample $z_0 \sim \mathcal{N}(0, I)$
for $\ell = 0, \ldots, L-1$ **do**
  Set $\tau_\ell \leftarrow \ell h$
  Set $z_{\ell+1} \leftarrow \mathrm{RK4Step}(v_{\theta^\star}, z_\ell, \tau_\ell, y, h)$
end for
Set $\widehat{\mathbf{\Delta}} \leftarrow \Psi^{-1}(z_L)$
return $\widehat{\mathbf{\Delta}}$

B Extended Proofs

Geometric overview. The log-Euclidean map $\phi : \mathbb{S}^d_{++} \to \mathbb{R}^m$, $\phi(X) = \mathrm{svec}(\log X)$, is a global coordinate system on the SPD manifold. The matrix logarithm is a diffeomorphism from $\mathbb{S}^d_{++}$ onto $\mathbb{S}^d$ and svec is a linear isometry, so ϕ is a global diffeomorphism. The log-Euclidean metric is the pullback of the Euclidean inner product through ϕ,

$$g_X(\xi, \eta) = \big\langle D\phi_X[\xi], D\phi_X[\eta] \big\rangle_{\mathbb{R}^m}, \qquad \xi, \eta \in T_X\mathbb{S}^d_{++},$$

so ϕ is an isometry by construction, and the straight line $z_\tau = (1-\tau)\phi(X_0) + \tau\phi(X_1)$ maps through ϕ^{-1} to the log-Euclidean geodesic between X_0 and X_1. The construction requires $X \succ 0$: a positive-semidefinite matrix with a zero eigenvalue has no finite logarithm and lies outside the domain of ϕ.

For a trajectory of B matrices the state space is the product manifold $\mathcal{M}_B = (\mathbb{S}^d_{++})^B$, with tangent space $\prod_{b=1}^B T_{X_b}\mathbb{S}^d_{++}$ and product metric $g^{\mathrm{prod}}(\xi, \eta) = \sum_{b=1}^B g_{X_b}(\xi_b, \eta_b)$. Applying ϕ to every factor gives the trajectory chart $\Phi(X_{1:B}) = (\phi(X_1), \ldots, \phi(X_B)) \in \mathbb{R}^{B\times m}$, which is a global diffeomorphism with $g^{\mathrm{prod}}(\xi, \eta) = \langle D\Phi[\xi], D\Phi[\eta]\rangle_F$. If a single window X_b is singular, the corresponding factor of Φ is undefined and the trajectory lies outside the domain of the chart. Strict positive definiteness (Proposition 2) must therefore be established before this geometry can be used.

The chart, including the DCT and the standardization, is a global diffeomorphism and an isometry for its pullback metric. In addition, the Euclidean conditional flow-matching loss in these coordinates equals the Riemannian loss on $\mathcal{M}_B$, because the manifold norm of a velocity residual equals the Frobenius norm of its pushforward. Finally, integrating the flow in Euclidean coordinates and decoding by Φ^{-1} gives the manifold flow, and every decoded window is SPD because the matrix exponential of a symmetric matrix is SPD; the same holds at every step of an explicit Runge–Kutta integrator.

B.1 Graph-variate signal analysis

GVSA (Smith et al., 2019) defines the modulated connectivity of Section 4.1 entrywise as

$$\theta_{ij}(t) = \begin{cases} W_{ij} F_{\mathcal{V}}(x_i(t), x_j(t)), & i \neq j, \\ 0, & i = j, \end{cases} \tag{80}$$

where W_{ij} describes the stable, long-term relationship between nodes i and j, while $F_{\mathcal{V}}$ measures their instantaneous connectivity. For correlation-based GVSA,

$$F_{\mathcal{V}}(x_i(t), x_j(t)) = \left|(x_i(t) - \bar{x}_i)\,(x_j(t) - \bar{x}_j)\right|, \tag{81}$$

where $\bar{x}_i$ is the temporal mean of node i. Substituting this into the entrywise definition yields the sample-resolution connectivity matrix of Section 4.1.

Our implementation makes two deliberate modifications to this classical formulation in order to preserve the geometry required by the generative model. First, although the original correlation-based definition uses the absolute value of the instantaneous product, we retain the sign of both the long-term support and the instantaneous interaction. Let

$$u_i(t) = x_i(t) - \bar{x}_i, \tag{82}$$

and write

$$J_t = u_t u_t^\top, \qquad \Delta_t = W \odot J_t. \tag{83}$$

Thus, rather than replacing the instantaneous interaction by $|u_i(t)u_j(t)|$, we use the signed outer product. When the support is written as a correlation matrix W and $D_t = \mathrm{diag}(u_t)$, this gives

$$\Delta_t = W \odot u_t u_t^\top = D_t W D_t. \tag{84}$$

This signed construction preserves the congruence structure of the support matrix and is therefore compatible with positive-semidefinite, and in the experimentally used binned case positive-definite, GVD trajectories. In contrast, taking entrywise absolute values destroys this exact congruence relationship and is not required for the downstream geometric construction.

Second, the original GVSA definition sets the diagonal entries to zero. We retain the diagonal during GVD construction because it contains the instantaneous node-energy terms and is necessary for treating each connectivity state as a full symmetric positive-definite matrix. After temporal binning, we therefore use

$$\Delta_b = \frac{1}{|\mathcal{I}_b|} \sum_{t \in \mathcal{I}_b} W \odot u_t u_t^\top, \tag{85}$$

with the diagonal left intact. All GVD matrices used in our experiments are explicitly verified to be positive definite before entering the log-Euclidean representation; no ridge or nearest-SPD projection is applied to the canonical GVD construction. If a zero-diagonal graph representation is desired for visualization or conventional network analysis, the diagonal can be removed post hoc without altering the matrices used for geometric learning.

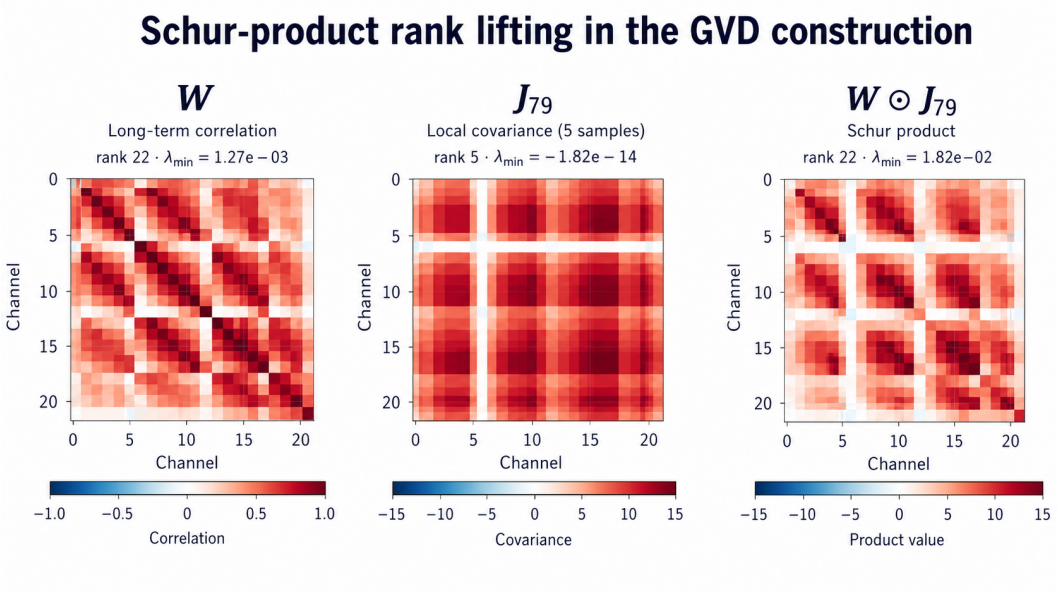


Figure 4: Schur-product rank lifting induced by the long-term support. The instantaneous covariance J_b is positive semidefinite but rank-deficient because each temporal window contains fewer samples than EEG channels. The long-term support W is positive definite, and the Hadamard composition $W \odot J_b$ preserves positive semidefiniteness while lifting the example to full rank. No ridge regularization is added to J_b or $W \odot J_b$.

B.2 Proof of Proposition 2 for single samples

Proof. Since $J_t = u_t u_t^\top \succeq 0$ and $W \succ 0$, the Schur product theorem (Horn and Johnson, 2012; Schur, 1911) guarantees that

$$\Delta_t = W \odot J_t \succeq 0. \tag{86}$$

To establish strict positive definiteness, let $D_t = \mathrm{diag}(u_t)$. Entrywise,

$$[W \odot (u_t u_t^\top)]_{ij} = W_{ij} u_{i,t} u_{j,t}, \tag{87}$$

and therefore

$$\Delta_t = D_t W D_t. \tag{88}$$

If every component of u_t is nonzero, then D_t is invertible. Hence, for every nonzero $v \in \mathbb{R}^d$,

$$v^\top \Delta_t v = (D_t v)^\top W (D_t v) > 0, \tag{89}$$

because $D_t v \neq 0$ and $W \succ 0$. Thus, $\Delta_t \in \mathbb{S}^d_{++}$. Moreover, congruence by an invertible matrix preserves rank, giving

$$\mathrm{rank}(\Delta_t) = \mathrm{rank}(D_t W D_t) = \mathrm{rank}(W) = d. \tag{90}$$

Therefore every Δ_t lies on the SPD manifold. □

If some component of u_t is exactly zero, D_t is singular and Δ_t is positive semidefinite rather than positive definite. We do not encounter this in practice however.

B.3 Further results

Proposition 3 (Equivalence of DCT-coordinate flow matching and product log-Euclidean Riemannian flow matching)**.** *Let*

$$\mathcal{M}_B = \left(\mathbb{S}^d_{++}\right)^B, \qquad m = \frac{d(d+1)}{2},$$

and define the product log-Euclidean coordinate map

$$\Phi(\Delta_{1:B}) = \begin{bmatrix} \mathrm{svec}(\log \Delta_1)^\top \\ \vdots \\ \mathrm{svec}(\log \Delta_B)^\top \end{bmatrix} \in \mathbb{R}^{B \times m},$$

where svec *is chosen to preserve the Frobenius inner product on symmetric matrices. Let*

$$\mathcal{C}(Z) = C_B Z, \qquad C_B^\top C_B = I_B,$$

be the orthonormal temporal DCT, and define the composite map

$$\Psi = \mathcal{C} \circ \Phi.$$

Equip $\mathcal{M}_B$ *with the product log-Euclidean metric*

$$g_{\mathrm{LE}}^{\mathrm{prod}}.$$

Then Ψ *is a global Riemannian isometry from*

$$(\mathcal{M}_B, g_{\mathrm{LE}}^{\mathrm{prod}})$$

to Euclidean space

$$(\mathbb{R}^{B\times m}, \langle\cdot,\cdot\rangle_F).$$

Consequently, for any endpoints $\Delta^0, \Delta^1 \in \mathcal{M}_B$*, with*

$$Z_i = \Psi(\Delta^i), \qquad i \in \{0,1\},$$

the Euclidean conditional path

$$Z_\tau = (1-\tau)Z_0 + \tau Z_1$$

is the Ψ*-image of the product log-Euclidean geodesic*

$$\Gamma_\tau = \Psi^{-1}(Z_\tau).$$

Moreover, if

$$U^E = Z_1 - Z_0$$

denotes the Euclidean conditional velocity and

$$U_\tau^{\mathcal{M}} = D\Psi_{Z_\tau}^{-1}[U^E]$$

the corresponding manifold tangent velocity, then for any manifold vector field V_θ *whose coordinate representation is*

$$v_\theta(Z,\tau) = D\Psi_{\Psi^{-1}(Z)}\left[V_\theta(\Psi^{-1}(Z),\tau)\right],$$

we have the pointwise identity

$$\boxed{\left\|V_\theta(\Gamma_\tau,\tau) - U_\tau^{\mathcal{M}}\right\|_{g_{\mathrm{LE}}^{\mathrm{prod}}}^2 = \|v_\theta(Z_\tau,\tau) - (Z_1 - Z_0)\|_F^2.}$$

Hence the Euclidean L_2 *conditional flow-matching objective in full-mode DCT coordinates is exactly equal to the corresponding Riemannian flow-matching objective on the product SPD manifold:*

$$\boxed{\mathcal{L}_{\mathrm{CFM}} = \mathcal{L}_{\mathrm{RFM}}.}$$

Proof. For each SPD factor, the matrix logarithm

$$\log : \mathbb{S}_{++}^d \to \mathbb{S}^d$$

is a global smooth diffeomorphism. Since svec is a linear isomorphism from $\mathbb{S}^d$ to $\mathbb{R}^m$, the product map

$$\Phi : \mathcal{M}_B \to \mathbb{R}^{B\times m}$$

is also a global smooth diffeomorphism.

By definition of the log-Euclidean metric, the single-factor map

$$\Delta \mapsto \mathrm{svec}(\log \Delta)$$

is an isometry from $\mathbb{S}_{++}^d$ equipped with the log-Euclidean metric to $\mathbb{R}^m$ equipped with its Euclidean metric. Therefore, for tangent vectors

$$\xi = (\xi_1,\ldots,\xi_B), \qquad \eta = (\eta_1,\ldots,\eta_B)$$

at $\Delta = (\Delta_1,\ldots,\Delta_B)$,

$$g_{\mathrm{LE},\Delta}^{\mathrm{prod}}(\xi,\eta) = \langle D\Phi_\Delta[\xi], D\Phi_\Delta[\eta]\rangle_F.$$

Now consider the temporal DCT map

$$\mathcal{C}(Z) = C_B Z.$$

Because C_B is orthogonal,

$$C_B^\top C_B = I_B,$$

and therefore, for arbitrary $X, Y \in \mathbb{R}^{B\times m}$,

$$\langle C_B X, C_B Y\rangle_F = \mathrm{tr}(X^\top C_B^\top C_B Y) = \mathrm{tr}(X^\top Y) = \langle X, Y\rangle_F.$$

Thus $\mathcal{C}$ is a Euclidean isometry.

Since

$$\Psi = \mathcal{C} \circ \Phi,$$

its differential satisfies

$$D\Psi_\Delta = D\mathcal{C}_{\Phi(\Delta)} \circ D\Phi_\Delta.$$

Hence

$$\begin{aligned}\langle D\Psi_\Delta[\xi], D\Psi_\Delta[\eta]\rangle_F &= \langle C_B D\Phi_\Delta[\xi], C_B D\Phi_\Delta[\eta]\rangle_F &(91)\\ &= \langle D\Phi_\Delta[\xi], D\Phi_\Delta[\eta]\rangle_F &(92)\\ &= g^{\mathrm{prod}}_{\mathrm{LE},\Delta}(\xi,\eta). &(93)\end{aligned}$$

Therefore Ψ is a global Riemannian isometry.

Now let

$$Z_\tau = (1-\tau)Z_0 + \tau Z_1.$$

Euclidean straight lines are geodesics, and an isometry maps geodesics to geodesics. Thus

$$\Gamma_\tau = \Psi^{-1}(Z_\tau)$$

is the corresponding product log-Euclidean geodesic on $\mathcal{M}_B$.

Differentiating

$$\Gamma_\tau = \Psi^{-1}(Z_\tau)$$

gives

$$\dot\Gamma_\tau = D\Psi^{-1}_{Z_\tau}[\dot Z_\tau].$$

Since

$$\dot Z_\tau = Z_1 - Z_0,$$

we obtain

$$U_\tau^{\mathcal{M}} = D\Psi^{-1}_{Z_\tau}[Z_1 - Z_0].$$

Because Ψ is a diffeomorphism,

$$D\Psi_{\Gamma_\tau} \circ D\Psi^{-1}_{Z_\tau} = \mathrm{Id},$$

and hence

$$D\Psi_{\Gamma_\tau}[U_\tau^{\mathcal{M}}] = Z_1 - Z_0.$$

By definition of the coordinate representation of the model vector field,

$$D\Psi_{\Gamma_\tau}[V_\theta(\Gamma_\tau,\tau)] = v_\theta(Z_\tau,\tau).$$

Therefore, using linearity of the differential,

$$\begin{aligned}&D\Psi_{\Gamma_\tau}\left[V_\theta(\Gamma_\tau,\tau) - U_\tau^{\mathcal{M}}\right] &(94)\\ &\quad = v_\theta(Z_\tau,\tau) - (Z_1 - Z_0). &(95)\end{aligned}$$

Since Ψ is a Riemannian isometry,

$$\|\xi\|^2_{g^{\mathrm{prod}}_{\mathrm{LE}}} = \|D\Psi[\xi]\|_F^2.$$

Applying this to the velocity residual yields

$$\left\|V_\theta(\Gamma_\tau,\tau) - U_\tau^{\mathcal{M}}\right\|^2_{g^{\mathrm{prod}}_{\mathrm{LE}}} = \left\|v_\theta(Z_\tau,\tau) - (Z_1 - Z_0)\right\|_F^2.$$

The equality holds pointwise for every endpoint pair and every τ. Taking expectations therefore gives

$$\mathcal{L}_{\mathrm{RFM}}(\theta) = \mathcal{L}_{\mathrm{CFM}}(\theta).$$

□

Corollary 1 (Standardized chart)**.** *Let $S(Z) = (Z - \mathbf{1}\mu^\top)D_\sigma^{-1}$ with $D_\sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_m)$, $\sigma_j > 0$, and let $\Psi_\sigma = \mathcal{C} \circ S \circ \Phi$ be the chart used in practice. Define on $\mathcal{M}_B$ the metric*

$$g_\sigma(\xi, \eta) = \sum_{b=1}^{B} \big\langle D_\sigma^{-1}\, \mathrm{D}\phi_{\Delta_b}[\xi_b],\, D_\sigma^{-1}\, \mathrm{D}\phi_{\Delta_b}[\eta_b] \big\rangle, \qquad \phi = \mathrm{svec} \circ \log .$$

Then (i) Ψ_σ is a global diffeomorphism and a Riemannian isometry from $(\mathcal{M}_B, g_\sigma)$ to $(\mathbb{R}^{B\times m}, \langle\cdot,\cdot\rangle_F)$, so Proposition 3 holds verbatim with Ψ_σ and g_σ in place of Ψ and $g_{\mathrm{LE}}^{\mathrm{prod}}$; (ii) the geodesics of g_σ and $g_{\mathrm{LE}}^{\mathrm{prod}}$ coincide as parametrized curves; (iii) for a manifold field V and target U, $\|V-U\|_{g_\sigma}^2 = \sum_b \|D_\sigma^{-1}\mathrm{D}\phi_{\Delta_b}[V_b - U_b]\|^2$, a fixed diagonal reweighting of the log-Euclidean residual, so the population minimizer over measurable fields is the same conditional expectation under both metrics.

Proof. (i) S is an invertible affine map, so Ψ_σ is a diffeomorphism, and its differential is $\mathrm{D}\Psi_\sigma[\xi] = C_B\, \mathrm{D}\Phi[\xi]\, D_\sigma^{-1}$. Since C_B is orthogonal, $\langle \mathrm{D}\Psi_\sigma[\xi], \mathrm{D}\Psi_\sigma[\eta]\rangle_F = g_\sigma(\xi, \eta)$, which is the definition of an isometry; the proof of Proposition 3 uses only this property. (ii) In log coordinates both metrics are constant, so their geodesics are affinely parametrized straight lines in these coordinates, and S maps straight lines to straight lines. (iii) The expression follows from the definition of g_σ. The conditional expectation minimizes the expected squared residual under any fixed positive-definite quadratic form, so both objectives share it as population minimizer. □

Remark 2. The proposition places no restriction on how v_θ is computed from Z beyond measurability. It therefore holds for the network with the temporal branch, whose temporal input $C_B^\top Z$ is a fixed linear function of Z, and it holds for any coupling of the endpoints, including the classwise minibatch Sinkhorn coupling used in training, because the identity is pointwise in the endpoint pair.

Proposition 4 (Spectral bounds for graph-variate connectivity)**.** *Let $u_1, \ldots, u_n \in \mathbb{R}^d$ denote the observations in a temporal bin, and define*

$$J = \frac{1}{n}\sum_{t=1}^{n} u_t u_t^\top, \qquad \Delta = W \odot J, \tag{96}$$

where $W \in \mathbb{S}_{++}^d$. Define the channelwise energy within the bin by

$$q_i := \frac{1}{n}\sum_{t=1}^{n} u_{i,t}^2, \qquad i = 1, \ldots, d, \tag{97}$$

and let

$$Q := \mathrm{diag}(q_1, \ldots, q_d). \tag{98}$$

Then Δ satisfies the Loewner-order bounds

$$\lambda_{\min}(W)\, Q \preceq \Delta \preceq \lambda_{\max}(W)\, Q. \tag{99}$$

Consequently,

$$\lambda_{\min}(\Delta) \geq \lambda_{\min}(W) \min_i q_i, \tag{100}$$

and

$$\lambda_{\max}(\Delta) \leq \lambda_{\max}(W) \max_i q_i. \tag{101}$$

If $q_i > 0$ for every channel, then in particular

$$\lambda_{\min}(\Delta) > 0, \tag{102}$$

and its spectral condition number obeys

$$\kappa_2(\Delta) \leq \kappa_2(W) \frac{\max_i q_i}{\min_i q_i}. \tag{103}$$

Thus the distance of a GVD matrix from the boundary of the SPD cone is controlled jointly by the smallest eigenvalue of the long-term support and the least energetic channel in the local bin.

Proof. For each sample u_t, let

$$D_t := \mathrm{diag}(u_t). \tag{104}$$

Using the identity

$$W \odot (u_t u_t^\top) = D_t W D_t, \tag{105}$$

and linearity of the Hadamard product, we can write

$$\Delta = W \odot \left(\frac{1}{n}\sum_{t=1}^{n} u_t u_t^\top\right) = \frac{1}{n}\sum_{t=1}^{n} D_t W D_t. \tag{106}$$

Since $W \in \mathbb{S}^d_{++}$,

$$\lambda_{\min}(W) I \preceq W \preceq \lambda_{\max}(W) I. \tag{107}$$

Congruence preserves the Loewner order, so for every t,

$$\lambda_{\min}(W) D_t^2 \preceq D_t W D_t \preceq \lambda_{\max}(W) D_t^2. \tag{108}$$

Averaging over the n samples gives

$$\lambda_{\min}(W) \left(\frac{1}{n} \sum_{t=1}^{n} D_t^2 \right) \preceq \Delta \preceq \lambda_{\max}(W) \left(\frac{1}{n} \sum_{t=1}^{n} D_t^2 \right). \tag{109}$$

Because

$$\frac{1}{n} \sum_{t=1}^{n} D_t^2 = \operatorname{diag} \left(\frac{1}{n} \sum_{t=1}^{n} u_{1,t}^2, \ldots, \frac{1}{n} \sum_{t=1}^{n} u_{d,t}^2 \right) = Q, \tag{110}$$

we obtain

$$\lambda_{\min}(W) Q \preceq \Delta \preceq \lambda_{\max}(W) Q, \tag{111}$$

which proves equation 99.

Applying the Rayleigh–Ritz characterization to the lower bound, for every unit vector x,

$$\begin{aligned} x^\top \Delta x &\geq \lambda_{\min}(W) x^\top Q x && (112) \\ &= \lambda_{\min}(W) \sum_{i=1}^{d} q_i x_i^2 && (113) \\ &\geq \lambda_{\min}(W) \min_i q_i. && (114) \end{aligned}$$

Taking the minimum over all $\|x\|_2 = 1$ therefore yields

$$\lambda_{\min}(\Delta) \geq \lambda_{\min}(W) \min_i q_i. \tag{115}$$

Likewise, the upper Loewner bound gives, for every unit vector x,

$$\begin{aligned} x^\top \Delta x &\leq \lambda_{\max}(W) x^\top Q x && (116) \\ &\leq \lambda_{\max}(W) \max_i q_i, && (117) \end{aligned}$$

and hence

$$\lambda_{\max}(\Delta) \leq \lambda_{\max}(W) \max_i q_i. \tag{118}$$

If every $q_i > 0$, the lower bound is strictly positive because $W \succ 0$, so $\lambda_{\min}(\Delta) > 0$. Finally,

$$\begin{aligned} \kappa_2(\Delta) &= \frac{\lambda_{\max}(\Delta)}{\lambda_{\min}(\Delta)} && (119) \\ &\leq \frac{\lambda_{\max}(W) \max_i q_i}{\lambda_{\min}(W) \min_i q_i} && (120) \\ &= \kappa_2(W) \frac{\max_i q_i}{\min_i q_i}, && (121) \end{aligned}$$

which proves equation 103. □

Proposition 5 (DCT-II diagonalizes discrete temporal variation)**.** *Let $D_B \in \mathbb{R}^{(B-1)\times B}$ denote the first-order temporal difference operator,*

$$D_B = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix},$$

and let

$$L_B = D_B^\top D_B$$

be the corresponding path-graph Laplacian. Let $C_B \in \mathbb{R}^{B\times B}$ be the orthonormal DCT-II matrix with entries

$$[C_B]_{k,b} = \alpha_k \cos \left(\frac{\pi k}{B} \left(b + \frac{1}{2} \right) \right), \qquad k, b = 0, \ldots, B-1,$$

where

$$\alpha_0 = \frac{1}{\sqrt{B}}, \qquad \alpha_k = \sqrt{\frac{2}{B}}, \quad k \geq 1.$$

Then the DCT-II basis diagonalizes L_B:

$$C_B L_B C_B^\top = \Lambda_B,$$

where

$$\Lambda_B = \mathrm{diag}(\lambda_0, \ldots, \lambda_{B-1}), \qquad \lambda_k = 4\sin^2\left(\frac{\pi k}{2B}\right).$$

In particular,

$$0 = \lambda_0 < \lambda_1 < \cdots < \lambda_{B-1} < 4.$$

Proof. Let $q_k \in \mathbb{R}^B$ denote the k-th DCT-II basis vector,

$$q_k(b) = \alpha_k \cos\left(\frac{\pi k}{B}\left(b + \frac{1}{2}\right)\right), \qquad b = 0, \ldots, B-1.$$

For an interior index $b = 1, \ldots, B-2$,

$$(L_B q_k)_b = 2q_k(b) - q_k(b-1) - q_k(b+1).$$

Let $\theta_k = \pi k/B$. Using

$$\cos(a - \theta_k) + \cos(a + \theta_k) = 2\cos(a)\cos(\theta_k),$$

we obtain

$$(L_B q_k)_b = 2(1 - \cos\theta_k) q_k(b).$$

Since

$$2(1 - \cos\theta_k) = 4\sin^2\left(\frac{\theta_k}{2}\right),$$

it follows that

$$(L_B q_k)_b = 4\sin^2\left(\frac{\pi k}{2B}\right) q_k(b).$$

The two boundary rows satisfy the same identity, and hence

$$L_B q_k = \lambda_k q_k, \qquad \lambda_k = 4\sin^2\left(\frac{\pi k}{2B}\right).$$

Because the DCT-II basis is orthonormal, its basis vectors form a complete orthonormal eigenbasis of L_B, which gives

$$C_B L_B C_B^\top = \Lambda_B.$$

The ordering of the eigenvalues follows from the strict monotonicity of $\sin(x)$ on $[0, \pi/2)$. □

Theorem 2 (Exact spectral decomposition of log-Euclidean temporal variation)**.** *Let*

$$\Delta_{1:B} = (\Delta_1, \ldots, \Delta_B) \in \left(\mathbb{S}^d_{++}\right)^B$$

be a GVD trajectory. Define its log-Euclidean coordinates by

$$z_b = \mathrm{svec}(\log \Delta_b) \in \mathbb{R}^m, \qquad m = \frac{d(d+1)}{2},$$

and stack them as

$$Z = \begin{bmatrix} z_1^\top \\ \vdots \\ z_B^\top \end{bmatrix} \in \mathbb{R}^{B \times m}.$$

Let

$$\bar{Z} = C_B Z$$

be the full orthonormal DCT-II representation, and denote the k-th DCT coefficient vector by

$$\bar{z}_k^\top = [\bar{Z}]_{k,:}, \qquad k = 0, \ldots, B-1.$$

Define the discrete temporal variation of the GVD trajectory under the log-Euclidean metric as

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \sum_{b=1}^{B-1} d^2_{\mathrm{LE}}(\Delta_{b+1}, \Delta_b),$$

where

$$d_{\mathrm{LE}}(A, B) = \|\log A - \log B\|_F.$$

Then

$$\boxed{\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \sum_{k=0}^{B-1} \lambda_k \|\bar{z}_k\|_2^2}$$

with

$$\lambda_k = 4\sin^2\left(\frac{\pi k}{2B}\right).$$

Equivalently,

$$\boxed{\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \sum_{k=1}^{B-1} 4\sin^2\left(\frac{\pi k}{2B}\right) \|\bar{z}_k\|_2^2}$$

since $\lambda_0 = 0$.

Hence, the DCT-II provides an exact orthogonal decomposition of the log-Euclidean temporal variation of a GVD trajectory. The zero-frequency mode contributes no temporal variation, while the weighting λ_k *increases monotonically with DCT mode index* k.

Proof. Since svec preserves the Frobenius inner product on symmetric matrices,

$$d_{\mathrm{LE}}^2(\Delta_{b+1}, \Delta_b) = \|\log \Delta_{b+1} - \log \Delta_b\|_F^2 = \|z_{b+1} - z_b\|_2^2.$$

Therefore,

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \sum_{b=1}^{B-1} \|z_{b+1} - z_b\|_2^2.$$

Using the first-difference matrix D_B from Proposition 5,

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \|D_B Z\|_F^2.$$

Hence,

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \mathrm{tr}\left(Z^\top D_B^\top D_B Z\right) = \mathrm{tr}\left(Z^\top L_B Z\right).$$

Since the DCT-II matrix is orthonormal,

$$Z = C_B^\top \bar{Z}.$$

Substituting this expression gives

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \mathrm{tr}\left(\bar{Z}^\top C_B L_B C_B^\top \bar{Z}\right).$$

By Proposition 5,

$$C_B L_B C_B^\top = \Lambda_B.$$

Thus

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \mathrm{tr}\left(\bar{Z}^\top \Lambda_B \bar{Z}\right).$$

Because Λ_B is diagonal,

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) = \sum_{k=0}^{B-1} \lambda_k \|\bar{z}_k\|_2^2.$$

Finally, substituting

$$\lambda_k = 4\sin^2\left(\frac{\pi k}{2B}\right)$$

gives the stated result. □

Interpretation. Theorem 2 gives the DCT-token representation a direct geometric interpretation. The quantity

$$E_k = 4\sin^2\left(\frac{\pi k}{2B}\right)\|\bar{z}_k\|_2^2$$

is exactly the contribution of DCT mode k to the discrete log-Euclidean temporal variation of the GVD trajectory. The DC mode $k = 0$ describes time-invariant connectivity structure and contributes zero temporal variation, whereas higher-order modes receive progressively larger temporal-variation weights. Thus, the DCT does not merely reorganize the trajectory into frequency coordinates; it diagonalizes its intrinsic temporal variation under the log-Euclidean geometry.

Corollary 2 (Exact log-Euclidean error of spectral truncation)**.** *Let $\Delta_{1:B}^{(K)}$ be obtained by retaining only DCT modes $k = 0, \dots, K-1$ and setting all remaining coefficients to zero before applying the inverse DCT and log-Euclidean decoder. Then*

$$\sum_{b=1}^{B} d_{\mathrm{LE}}^2\left(\Delta_b, \Delta_b^{(K)}\right) = \sum_{k=K}^{B-1}\|\bar{z}_k\|_2^2.$$

Moreover,

$$\mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}) - \mathcal{V}_{\mathrm{LE}}(\Delta_{1:B}^{(K)}) = \sum_{k=K}^{B-1}\lambda_k\|\bar{z}_k\|_2^2.$$

Proof. The first identity follows from the isometry of the log map, svec, and the orthonormal DCT together with Parseval's identity. The second follows directly from Theorem 2. □

Proposition 6 (Stable support attenuates local connectivity perturbations)**.** *Let*

$$W \in \mathbb{S}_{++}^d$$

be a correlation matrix and define

$$\rho_W = \max_{i \neq j}|W_{ij}|.$$

Since W is positive definite with unit diagonal,

$$0 \le \rho_W < 1.$$

Let a local covariance estimate satisfy

$$\widehat{J} = J + E,$$

where $E = E^\top$. Decompose the perturbation into diagonal and off-diagonal components,

$$E = E_{\mathrm{diag}} + E_{\mathrm{off}},$$

where

$$E_{\mathrm{diag}} = \mathrm{diag}(E), \qquad \mathrm{diag}(E_{\mathrm{off}}) = 0.$$

Define

$$\Delta = W \odot J, \qquad \widehat{\Delta} = W \odot \widehat{J}.$$

Then

$$\boxed{\|\widehat{\Delta} - \Delta\|_F^2 \le \|E_{\mathrm{diag}}\|_F^2 + \rho_W^2\|E_{\mathrm{off}}\|_F^2.}$$

In particular,

$$\|\widehat{\Delta} - \Delta\|_F \le \|\widehat{J} - J\|_F,$$

so the stable support never amplifies a covariance perturbation in Frobenius norm. Moreover, if the perturbation is purely off-diagonal,

$$E_{\mathrm{diag}} = 0,$$

then

$$\boxed{\|\widehat{\Delta} - \Delta\|_F \le \rho_W\|\widehat{J} - J\|_F,}$$

which is a strict contraction whenever $E \neq 0$.

Proof. By linearity of the Hadamard product,

$$\widehat{\Delta} - \Delta = W \odot E.$$

Because W is a correlation matrix,

$$W_{ii} = 1.$$

Therefore the diagonal perturbation is unchanged,

$$W \odot E_{\text{diag}} = E_{\text{diag}}.$$

For the off-diagonal component,

$$\begin{aligned}\|W \odot E_{\text{off}}\|_F^2 &= \sum_{i \neq j} W_{ij}^2 E_{ij}^2 \\ &\leq \rho_W^2 \sum_{i \neq j} E_{ij}^2 \\ &= \rho_W^2 \|E_{\text{off}}\|_F^2.\end{aligned}$$

Since the diagonal and off-diagonal components have disjoint support, they are orthogonal under the Frobenius inner product. Hence

$$\begin{aligned}\|\widehat{\Delta} - \Delta\|_F^2 &= \|W \odot E_{\text{diag}}\|_F^2 + \|W \odot E_{\text{off}}\|_F^2 \\ &\leq \|E_{\text{diag}}\|_F^2 + \rho_W^2 \|E_{\text{off}}\|_F^2.\end{aligned}$$

Since $\rho_W < 1$, the stated consequences follow. □

C SPECTRAL STRUCTURE OF GVD-CFM

This appendix provides the complete derivation underlying Theorem 1.

C.1 LOG-EUCLIDEAN TRAJECTORY COORDINATES

Let $\mathbf{\Delta} = (\Delta_1, \ldots, \Delta_B)$ with $\Delta_b \in \mathbb{S}_{++}^d$, and define $z_b = \text{svec}(\log \Delta_b) \in \mathbb{R}^m$. The complete trajectory is $Z = [z_1, \ldots, z_B]^\top \in \mathbb{R}^{B \times m}$, and the orthonormal DCT-II acts along the temporal dimension, $\bar{Z} = C_B Z$ with $C_B^\top C_B = I_B$. Since svec is linear, mode k is

$$\bar{z}_k = \sum_{b=1}^{B} (C_B)_{kb} z_b = \text{svec}\left(\sum_{b=1}^{B} (C_B)_{kb} \log \Delta_b \right). \tag{122}$$

The DCT does not act within an individual connectivity matrix. It reorganizes the temporal evolution of the complete log-connectivity trajectory.

C.2 TEMPORAL COVARIANCE AND THE KLT

For one log-connectivity coordinate j, define the temporal vector

$$x^{(j)} = \begin{bmatrix} Z_{1,j} & \cdots & Z_{B,j} \end{bmatrix}^\top. \tag{123}$$

Let

$$K_j = \text{Cov}(x^{(j)}) \tag{124}$$

and define the feature-averaged temporal covariance

$$K_T = \frac{1}{m} \sum_{j=1}^{m} K_j. \tag{125}$$

Because K_T is symmetric positive semidefinite,

$$K_T = U \Lambda U^\top, \tag{126}$$

with orthogonal U and diagonal

$$\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_B). \tag{127}$$

Proposition 7 (Exact KLT decorrelation)**.** *Let x be zero mean with covariance $K_T = U \Lambda U^\top$. Then for*

$$y = U^\top x, \tag{128}$$

we have

$$\text{Cov}(y) = \Lambda. \tag{129}$$

Therefore

$$\text{Cov}(y_k, y_\ell) = 0, \qquad k \neq \ell. \tag{130}$$

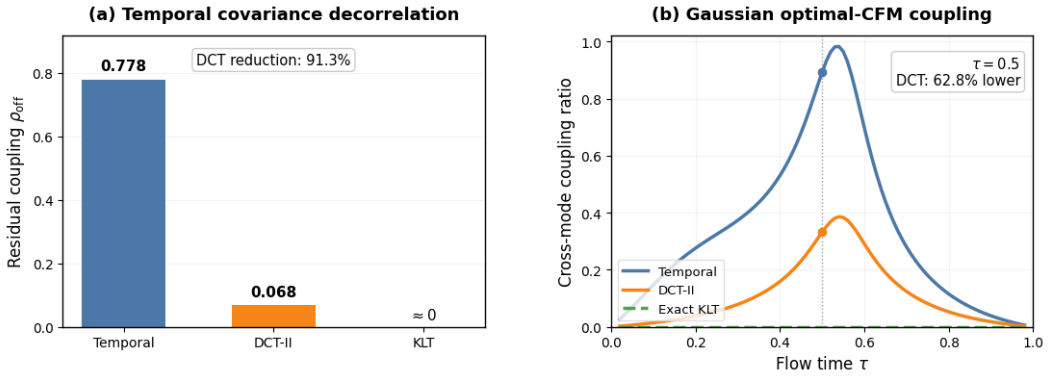


Figure 5: Empirical and theoretical motivation for the temporal DCT representation. **(a)** Residual off-diagonal temporal covariance, $\rho_{\text{off}}(K) = \|K - \text{diag}(K)\|_F / \|K\|_F$, for standardized log-GVD trajectories from BNCI2014_001. The full orthonormal DCT-II reduces residual temporal covariance coupling from 0.778 to 0.068, corresponding to a 91.3% reduction, while the empirical Karhunen–Loève transform (KLT) yields numerical zero. **(b)** Cross-mode coupling ratio of the population-optimal Gaussian conditional flow-matching field as a function of flow time τ. The exact KLT removes second-order cross-mode coupling, while DCT-II substantially reduces it relative to the original temporal coordinates; at $\tau = 0.5$, the DCT coupling ratio is 62.8% lower. Together, these results show that DCT-II acts as an invertible, data-independent approximation to the temporal KLT that reduces the cross-mode dependency structure presented to the trajectory velocity model.

Table 4: Temporal-basis ablation. Results are dataset-balanced means over five datasets and three generator seeds. All variants retain all $B = 100$ temporal coordinates. Random orthogonal uses a fixed random orthogonal basis and KLT the training-set empirical temporal Karhunen–Loève basis. The first four rows use the spectral-only velocity network and therefore isolate the basis; the last row is the full GVD-CFM. Ratios have ideal value 1. Best values are **bold**; second-best values are *underlined italics*.

Basis	Rel. GVD-FID ↓	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑	Temp. corr. ↑	Lag-ACF ↑	Energy → 1	Dyn. frac. → 1
No DCT	1.092	0.540	0.772	0.704	0.404	0.666	0.973	*1.011*
Random orthogonal	1.093	0.480	0.788	0.713	0.356	0.678	*1.017*	1.047
KLT	1.033	*0.600*	**0.795**	**0.729**	**0.951**	**0.998**	**1.004**	1.014
DCT, spectral only	*1.025*	*0.600*	*0.792*	0.721	*0.919*	*0.997*	0.975	**1.010**
GVD-CFM (DCT + temporal branch)	**1.022**	**0.613**	0.790	*0.725*	0.910	*0.997*	0.953	0.987

Proof.

$$\text{Cov}(y) = U^\top \text{Cov}(x) U \tag{131}$$
$$= U^\top K_T U \tag{132}$$
$$= U^\top U \Lambda U^\top U \tag{133}$$
$$= \Lambda. \tag{134}$$

□

C.3 The DCT as an approximate KLT

The DCT-transformed temporal covariance is

$$K_{\text{DCT}} = C_B K_T C_B^\top. \tag{135}$$

For temporally smooth or approximately stationary processes whose covariance eigenvectors are close to cosine modes,

$$C_B K_T C_B^\top \approx \Lambda. \tag{136}$$

We quantify residual temporal coupling using

$$\rho_{\text{off}}(K) = \frac{\|K - \text{diag}(K)\|_F}{\|K\|_F}. \tag{137}$$

The residual coupling in the temporal and DCT bases is then

$$\rho_{\text{time}} = \rho_{\text{off}}(K_T) \tag{138}$$

and

$$\rho_{\text{DCT}} = \rho_{\text{off}}\left(C_B K_T C_B^\top\right). \tag{139}$$

A reduction

$$\rho_{\text{DCT}} < \rho_{\text{time}} \tag{140}$$

indicates that the DCT has reduced second-order temporal dependence.

C.4 Separable trajectory covariance

Let

$$x = \text{vec}(Z^\top) \in \mathbb{R}^{Bm}, \tag{141}$$

which stacks the windows $z_1, \ldots, z_B$, and suppose

$$\Sigma = K_T \otimes K_S, \tag{142}$$

where K_S represents covariance between log-connectivity coordinates.

Proposition 8 (Temporal KLT decouples temporal covariance)**.** *If*

$$K_T = U \Lambda_T U^\top, \tag{143}$$

then with

$$Q = U^\top \otimes I_m, \tag{144}$$

we have

$$Q \Sigma Q^\top = \Lambda_T \otimes K_S. \tag{145}$$

Proof. Using the mixed-product property of the Kronecker product,

$$\begin{aligned} Q\Sigma Q^\top &= (U^\top \otimes I_m)(K_T \otimes K_S)(U \otimes I_m) && (146)\\ &= (U^\top K_T U) \otimes K_S && (147)\\ &= \Lambda_T \otimes K_S. && (148) \end{aligned}$$

□

Therefore the KLT removes cross-temporal second-order coupling while leaving the within-mode connectivity covariance K_S unchanged.

Replacing $U^\top$ with C_B yields the approximate DCT analogue

$$\Sigma_{\text{DCT}} = (C_B K_T C_B^\top) \otimes K_S. \tag{149}$$

C.5 Gaussian conditional flow matching

Assume

$$x_0 \sim \mathcal{N}(0, I), \qquad x_1 \sim \mathcal{N}(0, \Sigma), \tag{150}$$

with x_0 and x_1 independent, and define

$$x_\tau = (1-\tau)x_0 + \tau x_1. \tag{151}$$

The conditional target velocity is

$$u_\tau = x_1 - x_0. \tag{152}$$

For squared-error conditional flow matching, the population-optimal field is

$$v_\tau^\star(x) = \mathbb{E}[u_\tau \mid x_\tau = x]. \tag{153}$$

Proposition 9 (Optimal Gaussian CFM field)**.** *The optimal field is linear:*

$$v_\tau^\star(x) = A_\tau x, \tag{154}$$

where

$$A_\tau = [\tau\Sigma - (1-\tau)I]\left[(1-\tau)^2 I + \tau^2 \Sigma\right]^{-1}. \tag{155}$$

Proof. Since (u_τ, x_τ) is jointly Gaussian,

$$\mathbb{E}[u_\tau \mid x_\tau = x] = \mathrm{Cov}(u_\tau, x_\tau)\,\mathrm{Cov}(x_\tau)^{-1}x. \tag{156}$$

Using independence of x_0 and x_1,

$$\mathrm{Cov}(x_\tau) = (1-\tau)^2 I + \tau^2\Sigma, \tag{157}$$

and

$$\mathrm{Cov}(u_\tau, x_\tau) = \tau\Sigma - (1-\tau)I. \tag{158}$$

Substitution proves the result. □

Now let

$$\Sigma = V\Lambda V^\top. \tag{159}$$

Corollary 3 (Mode-wise Gaussian flow)**.** *In coordinates*

$$y = V^\top x, \tag{160}$$

the optimal field is diagonal:

$$v^\star_{\tau,k}(y) = a_{\tau,k}y_k, \tag{161}$$

with

$$a_{\tau,k} = \frac{\tau\lambda_k - (1-\tau)}{(1-\tau)^2 + \tau^2\lambda_k}. \tag{162}$$

Proof. Substituting

$$\Sigma = V\Lambda V^\top \tag{163}$$

into Proposition 9 gives

$$A_\tau = V\left[\tau\Lambda - (1-\tau)I\right]\left[(1-\tau)^2 I + \tau^2\Lambda\right]^{-1}V^\top. \tag{164}$$

Since the middle term is diagonal, the result follows. □

This proves Theorem 1. In an exact covariance eigenbasis, the optimal Gaussian conditional flow field requires no cross-mode coupling.

Corollary 4 (Mode-wise Gaussian flow under optimal-transport coupling)**.** *Let $x_0 \sim \mathcal{N}(0, I)$ and $x_1 \sim \mathcal{N}(0, \Sigma)$ with $\Sigma = V\Lambda V^\top \succ 0$, and couple them by the quadratic-cost optimal-transport map $x_1 = \Sigma^{1/2}x_0$ (Dowson and Landau, 1982; Villani, 2009). Along $x_\tau = (1-\tau)x_0 + \tau x_1$, the conditional target $u_\tau = x_1 - x_0$ is a deterministic function of x_τ, and in coordinates $y = V^\top x$*

$$v^\star_{\tau,k}(y) = \frac{\sqrt{\lambda_k} - 1}{(1-\tau) + \tau\sqrt{\lambda_k}}\,y_k. \tag{165}$$

Proof. Let $M_\tau = (1-\tau)I + \tau\Sigma^{1/2}$. Since $\Sigma^{1/2} \succ 0$, M_τ is positive definite for every $\tau \in [0,1]$, and $x_\tau = M_\tau x_0$. Hence $x_0 = M_\tau^{-1}x_\tau$ and

$$u_\tau = (\Sigma^{1/2} - I)x_0 = (\Sigma^{1/2} - I)M_\tau^{-1}x_\tau, \tag{166}$$

so $\mathbb{E}[u_\tau \mid x_\tau = x] = (\Sigma^{1/2} - I)M_\tau^{-1}x$. Both factors are diagonal in the basis V, with entries $\sqrt{\lambda_k} - 1$ and $(1-\tau) + \tau\sqrt{\lambda_k}$. □

Minibatch Sinkhorn coupling approximates this population coupling within each class, with Σ replaced by the class-conditional trajectory covariance. Under both couplings the optimal field is diagonal in the KLT basis, so the motivation for spectral coordinates does not depend on the coupling. The DCT is not generally the exact KLT, but whenever

$$C_B K_T C_B^\top \tag{167}$$

is more diagonal than K_T, the DCT reduces the amount of second-order temporal interaction that must be represented by the velocity network.

C.6 Residual coupling

Write

$$C_B K_T C_B^\top = D_T + E_T, \tag{168}$$

where

$$D_T = \operatorname{diag}\left(C_B K_T C_B^\top\right) \tag{169}$$

and

$$E_T = C_B K_T C_B^\top - D_T. \tag{170}$$

Define

$$\epsilon_{\mathrm{DCT}} = \frac{\|E_T\|_F}{\|C_B K_T C_B^\top\|_F}. \tag{171}$$

For the exact KLT,

$$\epsilon_{\mathrm{KLT}} = 0. \tag{172}$$

Hence ϵ_{DCT} measures the residual second-order coupling remaining after the fixed spectral transform.

C.7 Orthogonal invariance of the GVD-CFM representation

Proposition 10 (DCT isometry)**.** *For any trajectory Z,*

$$\|C_B Z\|_F = \|Z\|_F. \tag{173}$$

Likewise, for any two velocity fields V and U,

$$\|C_B(V-U)\|_F^2 = \|V-U\|_F^2. \tag{174}$$

Proof. Since

$$C_B^\top C_B = I_B, \tag{175}$$

$$\|C_B Z\|_F^2 = \operatorname{tr}\left(Z^\top C_B^\top C_B Z\right) \tag{176}$$

$$= \operatorname{tr}(Z^\top Z) \tag{177}$$

$$= \|Z\|_F^2. \tag{178}$$

□

Therefore the DCT does not alter the Euclidean flow-matching metric or reduce the objective by rescaling the trajectory. Its effect is purely a change of coordinates.

C.8 Energy organization

The KLT additionally orders directions by variance. If

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_B, \tag{179}$$

then the first r KLT modes retain the largest possible amount of expected second-order energy among all orthonormal r-dimensional projections.

When the DCT approximates the KLT, temporally persistent structure is therefore expected to concentrate toward lower-order cosine modes. GVD-CFM does not truncate this representation: all B modes are retained. Thus this is an organization of temporal variation rather than a dimensionality-reduction argument.

D Continuous-Grid Decoding as Band-Limited Interpolation

D.1 Cosine-basis decoding of GVD trajectories

A distinctive consequence of the DCT-coordinate representation is that a generated trajectory is not restricted to the temporal grid used during training. The generated DCT coefficients define a finite cosine trajectory over normalized time, and may therefore be evaluated at an arbitrary number of temporal locations without retraining the generative model. We test this with an intentionally low native resolution: the spectral-only GVD-CFM network is trained once at $B_{\mathrm{train}} = 25$ windows, and the same generated coefficients are decoded

at $M \in \{25, 50, 100, 200, 400\}$, with the real reference trajectories recomputed natively at each M. The protocol is given in Section D.2.

Moderate upsampling preserves both discriminative and temporal structure. A $4\times$ denser grid ($M = 100$) changes generated-to-real AUC only from 0.803 to 0.802, and at $M = 200$ it remains 0.794 with lag-ACF agreement 0.959. Every generated matrix stays SPD at every tested resolution. Evaluating at $M > B$ is interpolation of a finite-bandwidth function: it does not increase the generated temporal bandwidth. Table 5 reports the full comparison and the finite-bandwidth limit that appears at $M = 400$.

The composite diffeomorphism introduced in Section 4.2.1 maps a GVD trajectory to the DCT-coordinate representation

$$\bar{Z} = C_B \widetilde{Z} \in \mathbb{R}^{B\times m},$$

where B is the temporal resolution of the GVD trajectory, $m = d(d+1)/2$, $\widetilde{Z}$ denotes the standardized log-Euclidean trajectory, and C_B is the orthonormal DCT-II matrix. In the standard GVD-CFM decoder, a generated coefficient tensor $\widehat{\bar{Z}}$ is mapped back to the original B temporal locations using the inverse DCT,

$$\widehat{\widetilde{Z}} = C_B^\top \widehat{\bar{Z}}.$$

We additionally exploit the fact that the inverse DCT is an evaluation of a finite cosine basis. Rather than restricting decoding to the original B DCT sampling locations, the generated coefficients can therefore be interpreted as defining a continuous trajectory and evaluated on an arbitrary temporal grid.

Continuous cosine representation. Let

$$\widehat{\bar{Z}} = \begin{bmatrix} \widehat{\mathbf{c}}_0^\top \\ \vdots \\ \widehat{\mathbf{c}}_{B-1}^\top \end{bmatrix}, \qquad \widehat{\mathbf{c}}_k \in \mathbb{R}^m,$$

denote the DCT-coordinate trajectory generated by GVD-CFM. We associate these coefficients with the continuous standardized tangent trajectory

$$\widehat{\widetilde{\mathbf{z}}}(\xi) = \sum_{k=0}^{B-1} \alpha_k \, \widehat{\mathbf{c}}_k \cos(\pi k \xi), \qquad \xi \in [0,1], \tag{180}$$

where ξ denotes normalized EEG trajectory time and is distinct from the flow-time variable τ. The orthonormal DCT-II normalization is

$$\alpha_k = \begin{cases} B^{-1/2}, & k = 0, \\ \sqrt{2/B}, & k > 0. \end{cases} \tag{181}$$

Hence, GVD-CFM generates a finite set of frequency coefficients, but those coefficients define a function over continuous trajectory time. The output temporal resolution is consequently determined during decoding rather than being restricted to the resolution at which the coefficient representation was generated.

Evaluation at an arbitrary temporal resolution. Suppose that the generated trajectory is to be evaluated at M temporal locations. We use the centered grid

$$\xi_j^{(M)} = \frac{j + \frac{1}{2}}{M}, \qquad j = 0, \ldots, M-1. \tag{182}$$

Evaluating Equation 180 on this grid gives

$$\widehat{\widetilde{\mathbf{z}}}_j^{(M)} = \sum_{k=0}^{B-1} \alpha_k \, \widehat{\mathbf{c}}_k \cos\left[\pi k \frac{j + \frac{1}{2}}{M}\right]. \tag{183}$$

Equivalently, define the continuous DCT synthesis matrix

$$A_{M\leftarrow B} \in \mathbb{R}^{M\times B}, \qquad [A_{M\leftarrow B}]_{j,k} = \alpha_k \cos\left[\pi k \frac{j + \frac{1}{2}}{M}\right]. \tag{184}$$

The complete M-point standardized trajectory is then obtained by the single matrix operation

$$\boxed{\widehat{\widetilde{Z}}^{(M)} = A_{M\leftarrow B} \widehat{\widetilde{Z}}}. \tag{185}$$

Thus,

$$\widehat{\widetilde{Z}} \in \mathbb{R}^{B\times m}, \qquad \widehat{\widetilde{Z}}^{(M)} \in \mathbb{R}^{M\times m},$$

and M does not need to equal B. The same generated GVD trajectory can therefore be sampled at its native resolution, at a denser temporal resolution, or at a coarser temporal resolution without retraining GVD-CFM.

Exact agreement with the original GVD-CFM decoder. The continuous formulation is a strict extension of the inverse-DCT decoder already used by GVD-CFM. When $M = B$,

$$\xi_j^{(B)} = \frac{j + \frac{1}{2}}{B},$$

and therefore

$$[A_{B\leftarrow B}]_{j,k} = \alpha_k \cos\left[\frac{\pi}{B}\left(j + \frac{1}{2}\right)k\right]. \tag{186}$$

This is exactly the orthonormal inverse DCT-II basis, giving

$$A_{B\leftarrow B} = C_B^\top. \tag{187}$$

Consequently,

$$\widehat{\widetilde{Z}}^{(B)} = C_B^\top \widehat{\widetilde{Z}}, \tag{188}$$

which recovers the original GVD-CFM decoding operation exactly. Continuous basis decoding therefore does not alter the learned representation or introduce an additional generative model; it generalizes the temporal evaluation of the existing DCT representation.

Mapping the continuous trajectory back to the SPD manifold. After continuous basis evaluation, the same inverse log-Euclidean mapping used by GVD-CFM is applied independently at each temporal location. Using the training-set coordinate statistics $\mu, \sigma \in \mathbb{R}^m$,

$$\widehat{\mathbf{z}}_j^{(M)} = \widehat{\widetilde{\mathbf{z}}}_j^{(M)} \odot \sigma + \mu. \tag{189}$$

The corresponding symmetric tangent matrix is

$$\widehat{X}_j^{(M)} = \mathrm{svec}^{-1}\left(\widehat{\mathbf{z}}_j^{(M)}\right), \tag{190}$$

and the GVD matrix is reconstructed as

$$\widehat{\Delta}_j^{(M)} = \exp\left(\widehat{X}_j^{(M)}\right), \qquad j = 0, \ldots, M-1. \tag{191}$$

Since $\widehat{X}_j^{(M)}$ is symmetric, its matrix exponential is strictly positive definite. Hence,

$$\widehat{\Delta}_j^{(M)} \in \mathbb{S}_{++}^d \qquad \forall j, \tag{192}$$

so changing the temporal evaluation resolution does not compromise the manifold constraint.

Interpretation. The DCT-token representation used by GVD-CFM can therefore be viewed not only as a convenient Euclidean coordinate system for a discrete product-SPD trajectory, but also as the coefficient representation of a continuous cosine trajectory. All B generated modes are retained; no spectral truncation is introduced. Continuous decoding changes only the set of temporal locations at which this same generated trajectory is evaluated.

This separates *generative resolution* from *sampling resolution*: GVD-CFM learns the joint distribution of B cosine modes, while the resulting trajectory may be evaluated at any desired number M of temporal locations. In particular, choosing $M > B$ produces a denser realization of the generated dynamic connectivity trajectory while preserving the same underlying DCT coefficients and the SPD geometry of every reconstructed GVD matrix.

D.2 Method

The DCT representation used by GVD-CFM permits a generated coefficient trajectory to be evaluated on a temporal grid different from the one used during training. We evaluate this property directly by deliberately training the generative model at a low temporal resolution and decoding its outputs on progressively denser grids.

All experiments in this section use

$$B_{\mathrm{train}} = 25.$$

For a generated DCT-coordinate trajectory

$$\widehat{\widetilde{Z}} = \begin{bmatrix} \widehat{c}_0^\top \\ \vdots \\ \widehat{c}_{B-1}^\top \end{bmatrix} \in \mathbb{R}^{B\times m}, \qquad B = 25,$$

the continuous standardized tangent trajectory is

$$\widehat{z}(\xi) = \sum_{k=0}^{B-1} \alpha_k \widehat{c}_k \cos(\pi k \xi), \qquad \xi \in [0,1],$$

where

$$\alpha_k = \begin{cases} B^{-1/2}, & k = 0, \\ \sqrt{2/B}, & k > 0. \end{cases}$$

For a target resolution M, we evaluate this function on the centered grid

$$\xi_j^{(M)} = \frac{j + \frac{1}{2}}{M}, \qquad j = 0, \dots, M-1.$$

Equivalently,

$$\widehat{Z}^{(M)} = A_{M \leftarrow B} \widehat{\widetilde{Z}},$$

with

$$[A_{M \leftarrow B}]_{j,k} = \alpha_k \cos\left[\pi k \frac{j + \frac{1}{2}}{M}\right].$$

The resulting standardized tangent coordinates are inverse-standardized,

$$\widehat{z}_j^{(M)} = \widehat{\widetilde{z}}_j^{(M)} \odot \sigma + \mu,$$

mapped back to symmetric matrices using svec^{-1},

$$\widehat{X}_j^{(M)} = \mathrm{svec}^{-1}\left(\widehat{z}_j^{(M)}\right),$$

and finally reconstructed on the SPD manifold through

$$\widehat{\Delta}_j^{(M)} = \exp\left(\widehat{X}_j^{(M)}\right).$$

Because $\widehat{X}_j^{(M)}$ is symmetric,

$$\widehat{\Delta}_j^{(M)} \in \mathcal{S}_{++}^d \qquad \forall j, M.$$

We evaluate

$$M \in \{25, 50, 100, 200, 400\},$$

corresponding to $1\times$, $2\times$, $4\times$, $8\times$, and $16\times$ the training-grid density.

A single generated coefficient tensor $\widehat{\widetilde{Z}}$ is reused across all values of M. Differences across resolutions therefore arise only from evaluating the same learned continuous cosine trajectory on different temporal grids. No GVD-CFM model is retrained for any target resolution.

For each M, the corresponding real reference trajectories are recomputed directly from raw EEG using M GVD windows. The experiment therefore compares

$$\text{GVD-CFM trained at } B = 25 \quad \longrightarrow \quad \text{continuous decode at } M$$

against

$$\text{real EEG} \quad \longrightarrow \quad \text{native GVD construction at } M.$$

The same pooled cross-session/cross-run train–test protocol used in the main benchmark is retained. Results are reported for BNCI2014-001, BNCI2014-002, BNCI2015-001, Shin2017A, and Zhou2016.

D.3 Dataset-balanced results

The principal observation is that substantial temporal densification remains possible without retraining. Relative to native-resolution decoding,

$$\mathrm{AUC}_{25\to25} = 0.803, \qquad \mathrm{F1}_{25\to25} = 0.739,$$

while $4\times$ denser decoding gives

$$\mathrm{AUC}_{25\to100} = 0.802, \qquad \mathrm{F1}_{25\to100} = 0.734.$$

Thus,

$$\Delta\mathrm{AUC} = -0.001, \qquad \Delta\mathrm{F1} = -0.005.$$

At the same resolution, temporal-correlation agreement remains 0.860 and lag-ACF agreement remains 0.960.

Table 5: Dataset-balanced continuous-resolution results on BNCI2014_001, BNCI2014_002, BNCI2015_001, Shin2017A, and Zhou2016. A single spectral-only GVD-CFM is trained at $B_{\text{train}} = 25$ for each dataset and decoded at the indicated target resolution M without retraining.

M	Rel. Fréchet	Temp. corr.	Lag-ACF	Dyn. energy	SPD	Gen→Real AUC	Gen→Real F1	Real→Real AUC	Real→Real F1
25	1.098	0.944	0.995	0.978	1.000	0.803	0.739	0.838	0.764
50	0.983	0.853	0.960	0.595	1.000	0.804	0.739	0.838	0.761
100	0.919	0.860	0.960	0.350	1.000	0.802	0.734	0.831	0.756
200	0.880	0.780	0.959	0.199	1.000	0.794	0.715	0.832	0.760
400	0.872	0.506	0.822	0.105	1.000	0.785	0.647	0.810	0.732

Table 6: Per-dataset continuous-resolution results for a single spectral-only GVD-CFM trained at $B_{\text{train}} = 25$. The same generated DCT coefficient tensor is evaluated at $M \in \{25, 50, 100, 200, 400\}$ without retraining, while the corresponding real GVD trajectories are recomputed natively at each target resolution.

Dataset	M	Rel. Fréchet	Temp. corr.	Lag-ACF	Dyn. energy	Gen→Real AUC	Gen→Real F1	Real→Real AUC	Real→Real F1
BNCI2014-001	25	1.163	0.977	0.995	1.041	0.814	0.724	0.863	0.775
	50	1.021	0.875	0.977	0.622	0.824	0.748	0.866	0.784
	100	0.979	0.890	0.958	0.360	0.816	0.727	0.857	0.783
	200	0.958	0.778	0.959	0.195	0.806	0.704	0.821	0.739
	400	0.943	0.440	0.558	0.103	0.778	0.497	0.797	0.710
BNCI2014-002	25	1.008	0.911	0.999	0.964	0.791	0.746	0.797	0.721
	50	0.841	0.822	0.984	0.560	0.792	0.742	0.802	0.714
	100	0.790	0.778	0.972	0.319	0.787	0.730	0.799	0.721
	200	0.768	0.643	0.961	0.184	0.772	0.714	0.841	0.771
	400	0.789	0.377	0.884	0.097	0.769	0.736	0.792	0.717
BNCI2015-001	25	1.070	0.963	0.998	0.882	0.766	0.697	0.793	0.714
	50	1.269	0.889	0.990	0.526	0.764	0.694	0.794	0.723
	100	1.246	0.894	0.988	0.291	0.764	0.700	0.779	0.701
	200	1.183	0.806	0.988	0.160	0.756	0.682	0.763	0.693
	400	1.145	0.439	0.923	0.080	0.744	0.681	0.753	0.682
Shin2017A	25	1.199	0.945	0.984	1.219	0.673	0.621	0.759	0.691
	50	0.898	0.840	0.896	0.812	0.675	0.625	0.752	0.676
	100	0.771	0.889	0.936	0.532	0.673	0.631	0.747	0.672
	200	0.733	0.886	0.942	0.325	0.671	0.628	0.767	0.690
	400	0.723	0.843	0.829	0.187	0.677	0.553	0.742	0.660
Zhou2016	25	1.052	0.926	1.000	0.783	0.968	0.907	0.977	0.917
	50	0.888	0.838	0.953	0.451	0.965	0.889	0.973	0.910
	100	0.806	0.850	0.946	0.245	0.965	0.882	0.971	0.907
	200	0.756	0.787	0.945	0.130	0.966	0.850	0.969	0.902
	400	0.761	0.430	0.918	0.062	0.955	0.772	0.963	0.887

Even at $M = 200$, corresponding to $8\times$ denser evaluation, generated-to-real AUC remains 0.794 and lag-ACF agreement remains 0.959. The corresponding temporal-correlation agreement is 0.780.

The real-to-real classifier provides useful context for these changes. At $M = 100$, its dataset-balanced AUC is 0.831, compared with 0.802 for generated-to-real classification, a gap of

$$0.029.$$

At $M = 200$, the corresponding values are 0.832 and 0.794, respectively.

All generated matrices remain SPD at every tested resolution:

$$\text{SPD validity} = 1.000$$

for every dataset and every M.

At the highest tested resolution, $M = 400$, generated-to-real AUC remains 0.785, whereas temporal-correlation agreement decreases to 0.506. This is consistent with the finite-bandwidth limitation of evaluating a trajectory represented by only 25 learned cosine modes on a substantially denser temporal grid.

D.4 Per-dataset results

The per-dataset results show that resolution transfer is not driven by a single dataset. BNCI2014-001 preserves generated-to-real AUC from 0.814 at $M = 25$ to 0.816 at $M = 100$ and 0.806 at $M = 200$. BNCI2014-002 similarly remains near 0.79 through $M = 100$, while BNCI2015-001 changes only from 0.766 at $M = 25$ to 0.764 at $M = 100$.

Shin2017A also shows stable discriminative performance under substantial densification. Its generated-to-real AUC is

$$0.673,\ 0.675,\ 0.673,\ 0.671,\ 0.677$$

at $M = 25, 50, 100, 200, 400$, respectively. Thus, the classifier-level utility of the generated trajectories is essentially unchanged across the entire range of target resolutions. Temporal-correlation agreement is 0.945 at native resolution, 0.889 at $M = 100$, and 0.886 at $M = 200$. Lag-ACF agreement remains 0.936 at $M = 100$ and 0.942 at $M = 200$.

Zhou2016 shows particularly strong preservation of discriminative structure:

$$0.968,\ 0.965,\ 0.965,\ 0.966$$

generated-to-real AUC at $M = 25, 50, 100, 200$, respectively, and remains at 0.955 even at $M = 400$.

Across the five datasets, the effect of increasing M is therefore more apparent in the temporal-dynamics diagnostics than in generated-to-real classification. Dataset-balanced generated-to-real AUC changes only from 0.803 at $M = 25$ to 0.802 at $M = 100$, 0.794 at $M = 200$, and 0.785 at $M = 400$. In contrast, the dataset-balanced dynamic-energy ratio decreases from 0.978 at $M = 25$ to 0.350 at $M = 100$, 0.199 at $M = 200$, and 0.105 at $M = 400$.

This behavior is expected from finite-bandwidth cosine decoding. Increasing M evaluates the same $B_{\text{train}} = 25$ learned cosine modes on a denser temporal grid; it does not introduce additional high-frequency modes. Consequently, continuous-grid decoding can preserve class-discriminative and broad temporal structure under substantial densification, while increasingly fine-scale dynamic amplitudes cannot match native high-resolution GVD trajectories indefinitely.

At $M = 400$, corresponding to $16\times$ the training-grid density, the limitation is visible in the dataset-balanced temporal-correlation agreement of 0.506 and dynamic-energy ratio of 0.105. Nevertheless, generated-to-real AUC remains 0.785, lag-ACF agreement remains 0.822, and every generated matrix remains SPD. This supports interpreting the procedure as continuous-resolution evaluation of a finite-bandwidth trajectory rather than recovery of temporal frequencies absent from the original 25-mode representation.

D.5 Interpretation and limitation

The experiment demonstrates that output temporal resolution is not fixed by the grid used during GVD-CFM training. In particular, a model trained on only 25 temporal positions can be evaluated at 100 or 200 positions while retaining much of its discriminative and temporal structure.

However, increasing M does not create additional temporal bandwidth. The learned trajectory contains only the $B = 25$ cosine modes generated by the model:

$$\widehat{z}(\xi) = \sum_{k=0}^{24} \alpha_k \widehat{c}_k \cos(\pi k \xi).$$

Hence, evaluating this function on a denser grid provides a finer sampling of the same finite-dimensional trajectory rather than synthesizing additional high-frequency modes.

This distinction is visible at $M = 400$. Although generated-to-real AUC remains 0.785 and SPD validity remains perfect, temporal-correlation agreement falls to 0.506. This regime corresponds to $16\times$ denser temporal evaluation than training and exposes the finite-bandwidth limitation of the 25-mode representation.

We therefore use the term *continuous-resolution decoding* to mean that a generated GVD trajectory can be evaluated on arbitrary temporal grids without retraining, while explicitly not claiming that arbitrarily dense evaluation recovers temporal frequencies absent from the learned coefficient representation.

E Regional and Physiological Plausibility of Generated GVD Dynamics

Here, we ask whether GVD-CFM reproduces physiologically structured spatiotemporal connectivity rather than only aggregate statistics. We performed the analysis on BNCI2014_001 by grouping its 22 scalp electrodes into four broad regions: Frontal/FC, Central, Centro-parietal and Parietal/Occipital. For each pair of regions we computed the mean magnitude of the corresponding GVD edges for held-out real and generated trials, and followed representative regional interactions across all $B = 100$ windows (Figure 6 in the main text). Finally, we compared a generated single-trial sequence with its nearest held-out real trial.

Regional connectivity organization. Figure 2 shows that the generated data preserve the large-scale regional organization of held-out EEG. Mean Frontal/FC connectivity is 0.847 in held-out real data and 0.835 in generated data, and Parietal/Occipital connectivity is 0.894 and 0.880. Central–Centro-parietal connectivity is similarly preserved (0.717 real against 0.706 generated), while the weaker Frontal/FC–Parietal/Occipital interaction remains weak (0.437 against 0.420). GVD-CFM therefore preserves the relative spatial organization

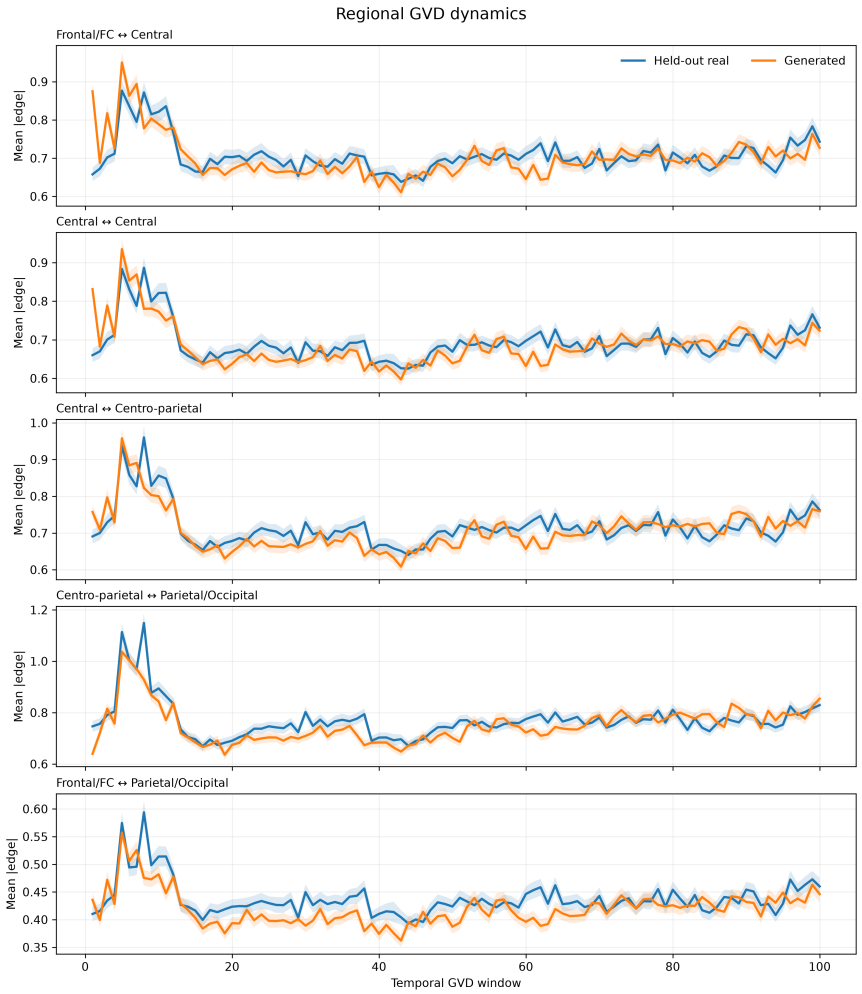


Figure 6: **Regional GVD dynamics on BNCI2014_001.** Mean GVD edge magnitude of five representative regional interactions over the $B = 100$ windows for held-out real and GVD-CFM trajectories. The generated trajectories reproduce the early transient, the fall to a lower connectivity regime, the mid-trial reduction and the gradual recovery of the real data, and these changes are aligned across regions. Shaded regions show the variability around each mean trajectory.

of interactions across the scalp, with strong local and adjacent-region coupling and weaker long-range frontal-to-posterior coupling. Motor-imagery EEG is classically associated with strong modulation of sensorimotor activity over central and neighboring centro-parietal regions, with distributed involvement of frontal and posterior areas (Pfurtscheller and da Silva, 1999). The preserved central and centro-parietal structure is consistent with this.

Temporal regional dynamics. Across the Frontal/FC–Central, Central–Central, Central–Centro-parietal, Centro-parietal–Parietal/Occipital and Frontal/FC–Parietal/Occipital interactions of Figure 6, real and generated data both show a pronounced early transient followed by a fall to a lower connectivity regime. Around the middle of the trial several interactions fall further, after which connectivity gradually recovers and continues to fluctuate. These changes occur together across related regions; the generated interactions do not fluctuate independently around a static mean. Such non-stationarity is compatible with the sequence of preparation, imagery and recovery within a motor-imagery trial.

Single-trial connectivity reconfiguration. Population-level agreement could in principle arise from a generator that learned only an average connectivity template. Figure 7 therefore compares one generated trial with its nearest held-out real trial over ten consecutive intervals. The real sequence reconfigures substantially over time, with periods of strong, spatially distributed connectivity alternating with weaker or more concen-

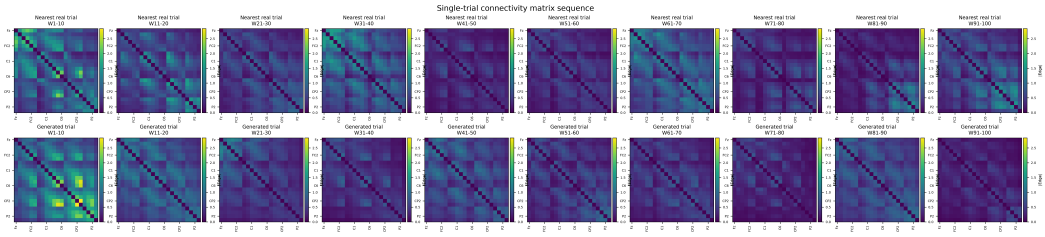


Figure 7: **Single-trial GVD connectivity-matrix sequence on BNCI2014_001.** One generated trial and its nearest held-out real trial are shown over ten consecutive temporal intervals. Both sequences exhibit substantial time-dependent reconfiguration of channel-level connectivity rather than a static graph template.

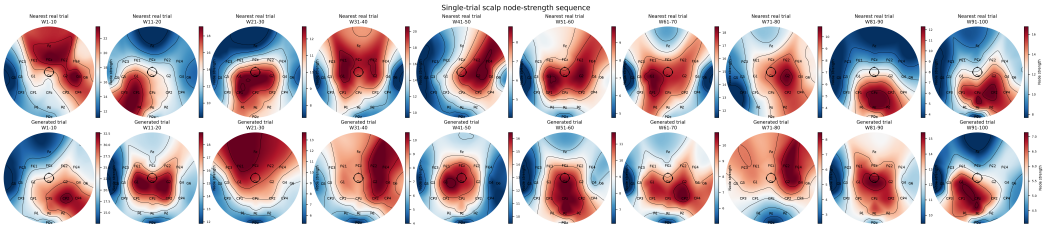


Figure 8: **Single-trial scalp node-strength sequence on BNCI2014_001.** Scalp node-strength maps for the same generated trial and nearest held-out real trial across ten temporal intervals. Both sequences display spatially heterogeneous and time-varying patterns involving frontal, central, centro-parietal, and posterior electrodes. The generated sequence preserves broad sensorimotor-centered spatial dynamics without reproducing the real trial exactly.

trated interactions. The generated sequence shows a comparable degree of restructuring and passes through a series of distinct channel-level patterns rather than holding a fixed matrix. The two sequences are not identical, which is consistent with generation of a new trajectory from the learned distribution rather than reproduction of a training trial.

Node-strength dynamics. The scalp node-strength maps of Figure 8 give a more interpretable view of the same trials. In the held-out trial, different intervals emphasize central, centro-parietal, posterior or more broadly distributed patterns. The generated trial shows similarly heterogeneous organization: high-strength regions shift across windows rather than remaining locked to one set of electrodes, prominent modulation recurs around central and centro-parietal locations, and frontal and posterior contributions vary over time. The model is therefore not scaling a fixed graph uniformly; the relative contribution of electrode groups changes over time.

Interpretation. These analyses show agreement at different levels: the time-averaged regional connectivity, its evolution over the trial, the reconfiguration of single-trial connectivity matrices, and the spatial distribution of node strength over time, which repeatedly involves the central and centro-parietal scalp regions relevant to motor imagery. We interpret this as evidence of regional physiological plausibility.

F Additional Ablations

F.1 Efficiency

Figure 9 summarizes training and generation cost. For GVD-CFM, 1000 training epochs take 304 s on average across the five main datasets and three seeds, and generating a batch of 32 trajectories with 50 RK4 steps takes 3.92 s (Appendix A).

F.2 Temporal-Resolution Audit of Stable-Support Information

To determine whether the discriminative information in GVD comes from the stable support W, from the instantaneous term J_t, or from their interaction, we evaluated four representations of real data over $B \in$

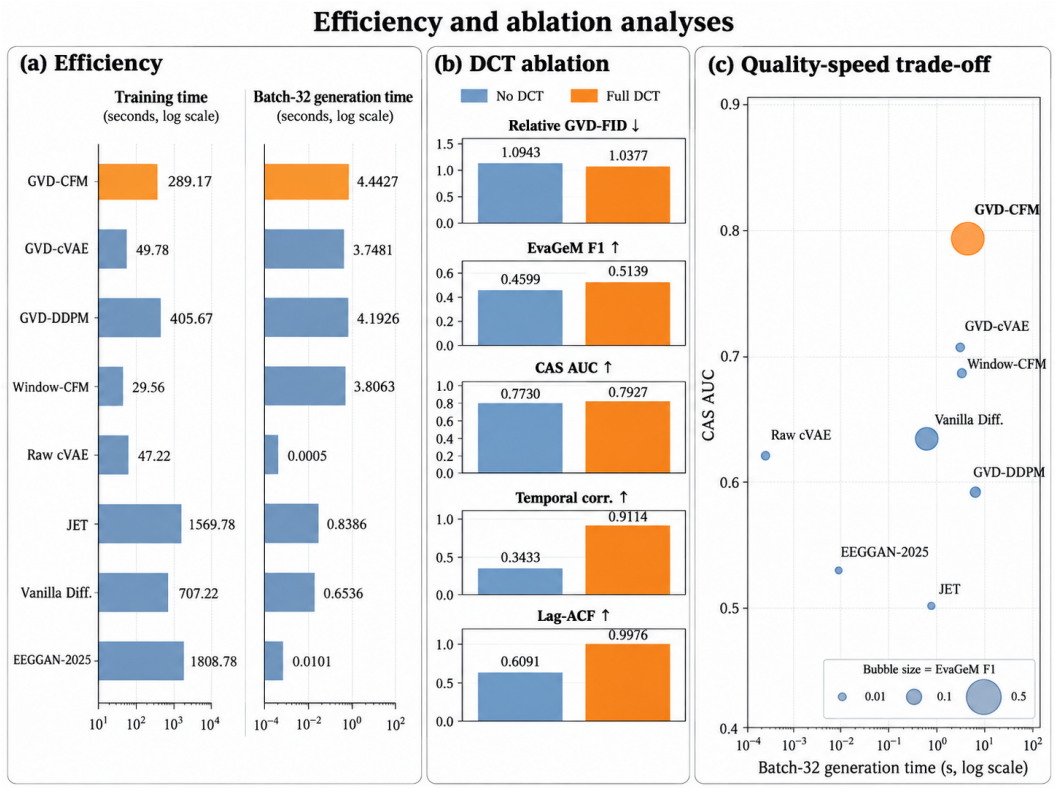


Figure 9: Efficiency and ablation analysis of GVD-CFM. **(a)** Average training time and batch-32 generation latency across five datasets and three generator seeds, shown on logarithmic axes. GVD-CFM is substantially faster to train than the large raw-signal generative baselines while retaining practical sampling cost. **(b)** DCT ablation with the spectral-only velocity network, comparing the full orthonormal DCT-II representation against direct temporal log-svec coordinates. The DCT improves relative GVD-FID, EvaGeM F1, CAS AUC, temporal-correlation agreement, and lag-autocorrelation agreement, with the largest gains observed in temporal structure. **(c)** Quality–speed trade-off across generators. The horizontal axis shows batch-32 generation time, the vertical axis shows CAS AUC, and marker size is proportional to EvaGeM F1. GVD-CFM combines strong downstream utility with substantially stronger trajectory-level distributional overlap than competing methods at practical generation cost.

$\{8, 16, 32, 64, 128, 256, 512\}$ windows with the CAS classifier and the cross-session protocol of the main benchmark: the stable-support trajectory $W \odot J_t$; the no-support trajectory $J_t + \epsilon I$; the static support W alone; and $W \odot J_t^{\text{Gaussian}}$, in which the node activity is replaced by independent Gaussian noise.

Table 7 shows that the stable-support representation keeps CAS AUC between 0.83 and 0.84 and weighted F1 between 0.76 and 0.77 from $B = 8$ to $B = 128$, and only falls to 0.807 and 0.738 at sample resolution. The no-support representation is informative at coarse resolution (AUC 0.816 at $B = 8$) but falls steadily as B increases, to 0.745 at $B = 64$ and 0.690 at $B = 512$. This is the expected behavior of a covariance estimated from fewer and fewer samples. The static support alone reaches AUC 0.822 and weighted F1 0.747, so W itself carries substantial class information. It does not explain the full representation, however: $W \odot J_t$ exceeds W alone at every resolution up to $B = 256$, and replacing the real node activity with Gaussian noise drops performance to near chance (AUC 0.51–0.57). The temporal modulation must therefore carry real structure from the EEG. At $B = 512$ the full representation falls slightly below the support alone, so at sample resolution the instantaneous noise outweighs the added dynamic information; the loss is still much smaller than for the no-support representation. Essentially, the stable support acts as a variance-reducing structural prior that keeps informative dynamic modulation while avoiding the instability of an independent covariance estimate in every short window.

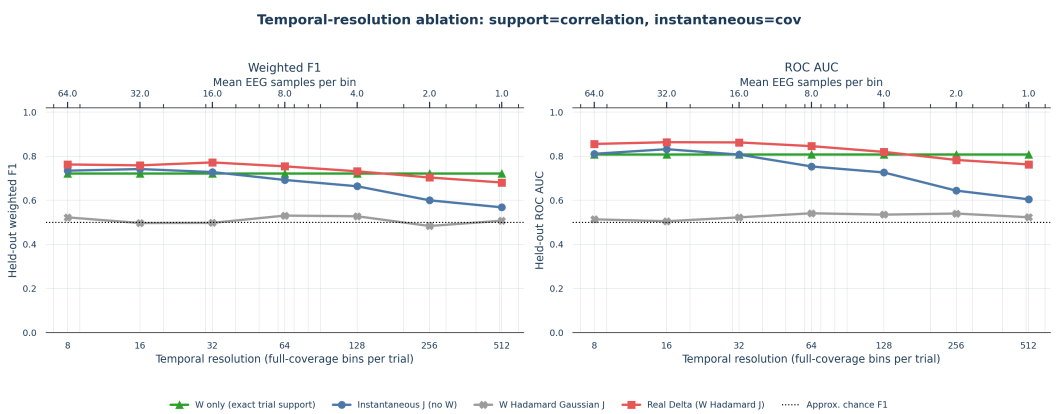


Figure 10: **Temporal-resolution ablation of graph-variate connectivity.** Weighted F1 (left) and ROC AUC (right) on the held-out evaluation split as the number of full-coverage bins increases from $B = 8$ to sample resolution ($B = 512$). The upper axis gives the mean number of EEG samples per bin. We compare the trial-level support W, the instantaneous covariance J without support, the control $W \odot J^{\text{Gaussian}}$, and the real graph-variate trajectory $\Delta = W \odot J$. The Gaussian control remains close to chance, so arbitrary temporal modulation of W does not reproduce the discriminative information of the real trajectory. At fine resolutions J degrades as its estimates become noisier, whereas modulation by the stable support is far more robust.

Table 7: Temporal-resolution support audit (Figure 10). Held-out classification of real trajectories with the CAS classifier, dataset-balanced over five datasets and three seeds, as a function of the number of windows B. *Stable support* is $W \odot J_t$, *no support* is $J_t + \epsilon I$ with $\epsilon = 10^{-6}$, *support only* is the static W and *Gaussian dynamics* is $W \odot J_t^{\text{Gaussian}}$. Best value in each column is **bold**.

Representation	$B = 8$	16	32	64	128	256	512
CAS AUC ↑							
Stable support	**0.841**	**0.838**	**0.840**	**0.833**	**0.837**	**0.827**	0.807
No support	0.816	0.797	0.765	0.745	0.756	0.738	0.690
Support only	0.822	0.822	0.822	0.822	0.822	0.822	**0.822**
Gaussian dynamics	0.512	0.518	0.527	0.545	0.574	0.558	0.528
CAS weighted F1 ↑							
Stable support	**0.770**	**0.772**	**0.766**	**0.764**	**0.759**	**0.747**	0.738
No support	0.747	0.735	0.698	0.684	0.693	0.673	0.640
Support only	0.747	0.747	0.747	0.747	0.747	**0.747**	**0.747**
Gaussian dynamics	0.505	0.512	0.513	0.526	0.550	0.535	0.506

F.3 Amplitude Bias and the Temporal Branch

GVD-CFM transports the trajectory in DCT coordinates, but the decoder exponentiates each window separately. Remark 3 shows that errors which are unbiased in log coordinates inflate the expected power of the decoded windows. Figure 11 shows this failure mode. The generated trajectory reproduces the broad shape of the real mean GVD edge trajectory, including the early transient and the subsequent recovery, but after the initial period it stays above the real trajectory for much of the trial. The global temporal shape is preserved while the time-local amplitude is biased.

The temporal branch of Section 4.2.2 addresses this bias without introducing a second dynamical state. It reads the current DCT state z_τ in the window basis, $C_B^\top z_\tau$, processes it with two AdaLN blocks, and returns the result to DCT alignment before the gated fusion (Figure 3). The velocity is still predicted in DCT coordinates and only the DCT state is integrated, but the network now sees directly the quantity that the decoder exponentiates. Figure 12 shows a generated mean GVD edge trajectory from the full model that follows the real one through the initial transient, the subsequent decline and the middle and later portions of the trial. Quantitatively, the branch reduces dynamic energy on every dataset and moves it towards 1 where the spectral-only network overshoots, while leaving the aggregate fidelity and utility metrics essentially unchanged (Table 3).

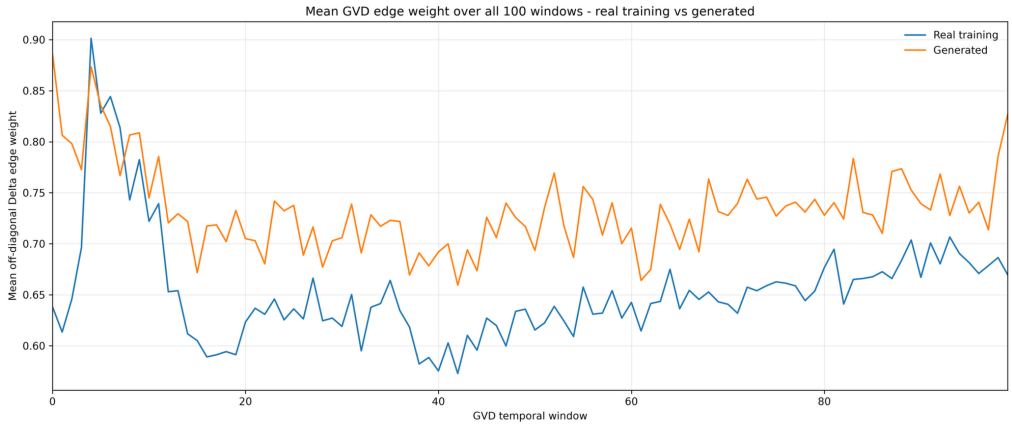


Figure 11: **Temporal amplitude bias in generated GVD dynamics.** Mean off-diagonal GVD edge weight over the $B = 100$ windows for real training trials and generated trials. The generated sequence captures the broad temporal organization, including the early transient, but shows a persistent positive amplitude offset over much of the later trajectory.

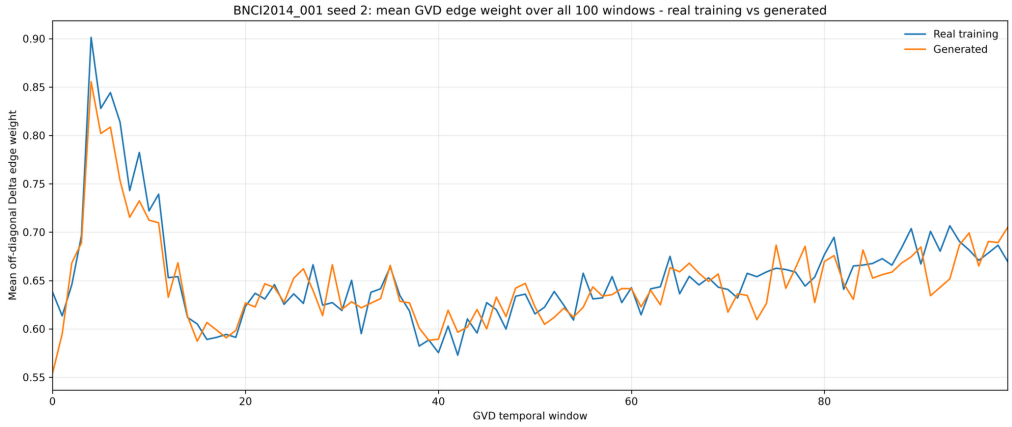


Figure 12: **Temporal alignment of the full GVD-CFM model.** Mean off-diagonal GVD edge weight over all $B = 100$ windows for real training trials and generated trials on BNCI2014_001. The generated trajectory follows the real sequence through the initial transient, the subsequent reduction, the intermediate fluctuations and the later recovery.

Remark 3 (Why spectral errors inflate decoded amplitudes). Let $\widehat{\bar{Z}} = \bar{Z}^\star + E$ be a generated DCT-coordinate trajectory with error E. Window b is decoded as $\widehat{\Delta}_b = \exp\big(\mathrm{svec}^{-1}(\widehat{z}_b)\big)$ with $\widehat{z}_b = z_b^\star + \sigma \odot (C_B^\top E)_b$, so the log-coordinate error of every window is a superposition of the errors of all modes. The map $X \mapsto \mathrm{tr}\exp(X)$ is convex on symmetric matrices. If the window error has zero mean, Jensen's inequality gives

$$\mathbb{E}\big[\mathrm{tr}\,\widehat{\Delta}_b\big] \geq \mathrm{tr}\exp\big(\mathrm{svec}^{-1}(z_b^\star)\big), \tag{193}$$

so unbiased errors in the log chart inflate the expected total power of each decoded window. The spectral Transformer sees E only mode by mode, whereas the temporal branch sees $C_B^\top \widehat{\bar{Z}}$, the quantity that is exponentiated window by window.

F.4 Nearest-Neighbor Memorization Diagnostic

Table 8 compares the ratio $M = d_1/d_2$ of the first and second nearest same-class training neighbors of generated samples with the same ratio for held-out real samples. A generator that copies training data would have M well below the held-out value. For GVD-CFM the two agree to within 0.001, and no generator produces exact copies. GVD-CFM has a near-copy rate of 0.16, against 0.98 for GVD-cVAE, whose samples collapse towards the training data.

Table 8: Nearest-neighbor memorization analysis. $M = d_1/d_2$ compares the first and second nearest same-class training neighbors. The held-out-real reference provides the target local-neighbor geometry; for GVD-CFM it is computed in the same run as its samples.

Method	M_{gen}	M_{held}	$\lvert\Delta M\rvert \downarrow$	M^2_{gen}	M^2_{held}	Near-copy ↓	Exact-copy ↓
GVD-CFM	0.99400	0.99501	0.00101	0.98803	0.99004	0.16155	0.00000
GVD-cVAE	0.99316	0.99516	0.00200	0.98636	0.99034	0.98171	0.00000
GVD-DDPM	0.99671	0.99516	0.00156	0.99344	0.99034	0.18854	0.00000
Window-DIFFEO-CFM	0.99678	0.99516	0.00162	0.99357	0.99034	0.00000	0.00000
cVAE	0.99555	0.99516	**0.00040**	0.99113	0.99034	0.01807	0.00000
JET	0.99298	0.99516	0.00218	0.98605	0.99034	0.00000	0.00000
Vanilla-Diffusion	0.99397	0.99516	0.00119	0.98799	0.99034	0.00005	0.00000
EEGGAN-2025	0.99169	0.99516	0.00347	0.98351	0.99034	0.00000	0.00000

Table 9: Dataset-balanced results over five EEG datasets and three generator seeds per dataset with generated log-eigenvalue clipping disabled. GVD-CFM is the full model with the temporal branch. Values are mean ± standard deviation across datasets after averaging generator seeds within each dataset.

Metric	**GVD-CFM**	**GVD-DDPM**	**GVD-cVAE**
Relative GVD-FID + increments	1.0158 ± 0.0692	1.7955 ± 1.5328	0.8131 ± 0.0789
Relative GVD-FID, positions only	1.0040 ± 0.0453	1.7354 ± 1.2118	0.9014 ± 0.0708
EvaGeM F1 + increments	0.6086 ± 0.2440	0.0480 ± 0.0372	0.0028 ± 0.0034
EvaGeM F1, positions only	0.6442 ± 0.1778	0.0046 ± 0.0058	0.0035 ± 0.0048
CAS AUC	0.7920 ± 0.0973	0.6071 ± 0.1317	0.7334 ± 0.0983
CAS weighted F1	0.7240 ± 0.0963	0.5759 ± 0.1020	0.6766 ± 0.0840
Temporal correlation agreement	0.9419 ± 0.0198	0.7442 ± 0.3233	0.8255 ± 0.0690
Lag-ACF agreement	0.9980 ± 0.0016	0.9462 ± 0.0885	0.9964 ± 0.0011
Dynamic-energy ratio	0.9750 ± 0.0774	1.8396 ± 2.0071	0.2301 ± 0.0596
Dynamic-fraction ratio	0.9818 ± 0.0968	1.6042 ± 1.4582	0.2587 ± 0.0787
M^2	0.9881 ± 0.0018	0.9931 ± 0.0027	0.9855 ± 0.0024
Held-out M^2	0.9900 ± 0.0025	0.9900 ± 0.0025	0.9900 ± 0.0025
Exact-copy rate	0.0000	0.0000	0.0000
Training coverage	0.2273 ± 0.0455	0.0385 ± 0.0189	0.1026 ± 0.0305

F.5 Ablation: Removing Generated Eigenvalue Clipping

To assess whether the performance of GVD-CFM depends on post-generation eigenvalue stabilization, we repeated the full generative benchmark with generated log-eigenvalue clipping disabled. All other components were unchanged: the full GVD-CFM with its temporal branch, the standard normal source, classwise Sinkhorn coupling, full orthonormal DCT coordinates, RK4 integration with 50 steps and the canonical stable-support GVD representation with $B = 100$ windows. We evaluated GVD-CFM, GVD-DDPM and GVD-cVAE on all five datasets with three generator seeds per dataset.

Table 9 reports the dataset-balanced results. Disabling clipping has a negligible effect on GVD-CFM. Its relative GVD Fréchet distance is 1.016, EvaGeM F1 0.609, CAS AUC 0.792 and CAS weighted F1 0.724, against 1.022, 0.613, 0.790 and 0.725 with clipping (Table 1). Temporal agreement remains high (temporal correlation 0.942, lag-ACF 0.998), and the dynamic-energy and dynamic-fraction ratios remain close to 1 (0.975 and 0.982). The reported performance of GVD-CFM is therefore not driven by clipping, which acts as a numerical safeguard.

The ablation also shows the different failure modes of the GVD-space controls. Without clipping, GVD-DDPM becomes unstable on some datasets, with a relative GVD Fréchet distance of 1.796 ± 1.533 across datasets and a dynamic-energy ratio of 1.840 ± 2.007. GVD-cVAE again attains a low Fréchet distance with near-zero EvaGeM F1 (0.003) and strongly attenuated dynamics (dynamic-energy ratio 0.230), so a favorable second-order distance does not imply faithful recovery of the dynamic distribution. Novelty metrics are unaffected: GVD-CFM has $M^2 = 0.988$ against a held-out value of 0.990, no exact copies, and training coverage 0.227, against 0.039 for GVD-DDPM and 0.103 for GVD-cVAE.

F.6 Extended Tables

Table 10: Full per-dataset generative performance corresponding to Table 1. Entries are mean±standard deviation over three generator seeds. The real-data reference uses real training trajectories against held-out real test trajectories for EvaGeM, and the corresponding TRTR classifier reference for CAS. Real-data references are excluded from generator rankings. Best generator values within each dataset are **bold**; second-best generator values are *underlined italics*.

Method	Dataset	Rel. GVD-FID ↓	Eva α ↑	Eva β ↑	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑
GVD-CFM	BNCI2014_001	1.006±0.014	**0.897±0.057**	**0.836±0.025**	**0.864±0.027**	**0.789±0.015**	**0.707±0.007**
	BNCI2014_002	*0.977±0.016*	**0.837±0.046**	**0.666±0.055**	**0.740±0.022**	**0.780±0.015**	**0.709±0.005**
	BNCI2015_001	*0.978±0.005*	**0.842±0.050**	*0.719±0.090*	**0.775±0.073**	**0.746±0.011**	**0.681±0.002**
	Shin2017A	*1.190±0.084*	*0.369±0.032*	**0.696±0.049**	**0.482±0.039**	**0.672±0.029**	**0.622±0.025**
	Zhou2016	*0.959±0.033*	**0.440±0.143**	0.133±0.081	*0.203±0.110*	**0.965±0.003**	**0.905±0.014**
	5-dataset avg.	*1.022*	**0.677**	**0.610**	**0.613**	**0.790**	**0.725**
GVD-cVAE	BNCI2014_001	**0.813±0.005**	0.018±0.001	0.007±0.006	0.009±0.006	*0.742±0.033*	*0.683±0.020*
	BNCI2014_002	**0.734±0.008**	0.027±0.022	0.001±0.002	0.002±0.003	*0.727±0.020*	*0.669±0.024*
	BNCI2015_001	**0.961±0.005**	0.020±0.001	0.010±0.004	0.013±0.004	*0.701±0.013*	*0.644±0.007*
	Shin2017A	**0.796±0.008**	0.008±0.013	0.002±0.002	0.000±0.000	0.575±0.026	0.552±0.016
	Zhou2016	**0.770±0.002**	0.012±0.004	0.002±0.001	0.003±0.002	0.888±0.059	*0.828±0.036*
	5-dataset avg.	**0.815**	0.017	0.004	0.005	*0.727*	*0.675*
GVD-DDPM	BNCI2014_001	*0.967±0.013*	0.064±0.025	0.023±0.012	0.034±0.016	0.530±0.049	0.513±0.031
	BNCI2014_002	1.019±0.001	0.110±0.011	0.051±0.006	0.070±0.008	0.583±0.029	0.541±0.005
	BNCI2015_001	1.000±0.013	0.186±0.044	0.107±0.022	0.135±0.029	0.536±0.031	0.516±0.026
	Shin2017A	4.778±0.014	0.006±0.008	0.000±0.000	0.001±0.001	0.529±0.004	0.507±0.013
	Zhou2016	1.004±0.007	0.077±0.018	0.014±0.004	0.024±0.006	0.815±0.032	0.723±0.039
	5-dataset avg.	1.753	0.089	0.039	0.053	0.599	0.560
Window-DIFFEO-CFM	BNCI2014_001	1.412±0.003	0.004±0.003	0.071±0.003	0.008±0.005	0.638±0.001	0.616±0.003
	BNCI2014_002	1.577±0.005	0.002±0.002	0.026±0.005	0.004±0.004	0.637±0.004	0.587±0.001
	BNCI2015_001	1.311±0.009	0.012±0.004	0.050±0.001	0.019±0.005	0.662±0.010	0.594±0.003
	Shin2017A	2.864±0.021	0.005±0.009	*0.021±0.004*	*0.006±0.010*	*0.589±0.008*	*0.563±0.010*
	Zhou2016	1.323±0.007	0.003±0.001	*0.242±0.009*	0.006±0.002	0.889±0.003	0.810±0.006
	5-dataset avg.	1.697	0.005	0.082	0.009	0.683	0.634
Real data reference	BNCI2014_001	–	0.911±0.001	0.894±0.000	0.902±0.001	0.856±0.003	0.763±0.008
	BNCI2014_002	–	0.673±0.001	0.328±0.001	0.441±0.001	0.801±0.005	0.733±0.007
	BNCI2015_001	–	0.926±0.001	0.882±0.003	0.904±0.001	0.780±0.002	0.703±0.004
	Shin2017A	–	0.639±0.003	0.883±0.002	0.741±0.002	0.745±0.001	0.676±0.009
	Zhou2016	–	0.880±0.004	0.337±0.004	0.488±0.004	0.972±0.000	0.907±0.002
	5-dataset avg.	–	0.806	0.665	0.695	0.831	0.756

Table 11: Full per-dataset generative performance (continued).

Method	Dataset	Rel. GVD-FID ↓	Eva α ↑	Eva β ↑	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑
cVAE	BNCI2014_001	1.374±0.020	*0.621±0.090*	0.002±0.002	0.004±0.003	0.588±0.013	0.555±0.015
	BNCI2014_002	1.154±0.008	*0.486±0.045*	0.006±0.004	0.012±0.008	0.550±0.022	0.491±0.066
	BNCI2015_001	1.885±0.052	0.023±0.005	0.005±0.001	0.007±0.001	0.584±0.002	0.546±0.011
	Shin2017A	1.617±0.055	**0.668±0.140**	0.002±0.002	0.003±0.003	0.514±0.013	0.470±0.043
	Zhou2016	1.049±0.003	0.240±0.043	0.018±0.006	0.034±0.011	*0.920±0.007*	0.799±0.026
	5-dataset avg.	1.416	*0.407*	0.006	0.012	0.631	0.572
JET	BNCI2014_001	4.794±0.294	0.004±0.002	0.002±0.001	0.002±0.001	0.467±0.011	0.333±0.000
	BNCI2014_002	3.402±0.053	0.000±0.000	0.000±0.000	0.000±0.000	0.494±0.033	0.333±0.000
	BNCI2015_001	3.864±0.035	0.014±0.007	0.007±0.009	0.008±0.009	0.501±0.016	0.342±0.015
	Shin2017A	3.753±0.097	0.005±0.005	0.000±0.000	0.000±0.000	0.498±0.009	0.333±0.000
	Zhou2016	3.726±1.331	0.003±0.003	0.000±0.000	0.000±0.000	0.516±0.151	0.405±0.067
	5-dataset avg.	3.908	0.005	0.002	0.002	0.495	0.349
Vanilla-Diffusion	BNCI2014_001	1.566±0.370	0.314±0.194	*0.448±0.347*	*0.366±0.250*	0.646±0.050	0.514±0.094
	BNCI2014_002	1.444±0.084	0.242±0.031	*0.370±0.273*	*0.264±0.072*	0.598±0.013	0.576±0.016
	BNCI2015_001	1.667±0.184	*0.240±0.053*	**0.746±0.096**	*0.363±0.072*	0.578±0.012	0.526±0.050
	Shin2017A	3.230±0.109	0.007±0.008	0.004±0.003	0.004±0.005	0.509±0.016	0.419±0.037
	Zhou2016	1.427±0.205	*0.284±0.059*	**0.264±0.188**	**0.240±0.074**	0.870±0.027	0.712±0.089
	5-dataset avg.	1.867	0.217	*0.366*	*0.247*	0.640	0.549
EEGGAN-2025	BNCI2014_001	4.943±0.023	0.008±0.009	0.001±0.001	0.001±0.001	0.516±0.044	0.336±0.004
	BNCI2014_002	3.424±0.033	0.000±0.000	0.000±0.000	0.000±0.000	0.513±0.043	0.465±0.019
	BNCI2015_001	4.005±0.027	0.014±0.007	0.007±0.006	0.009±0.006	0.518±0.069	0.391±0.042
	Shin2017A	3.856±0.032	0.005±0.005	0.000±0.000	0.000±0.000	0.501±0.018	0.405±0.063
	Zhou2016	2.249±0.054	0.002±0.001	0.000±0.000	0.000±0.000	0.573±0.123	0.408±0.028
	5-dataset avg.	3.696	0.006	0.002	0.002	0.524	0.401

Table 12: Full per-dataset complementary diagnostics corresponding to Table 2: temporal and dynamic fidelity. Entries are mean±standard deviation over three generator seeds. For ratio metrics, ranking is by proximity to 1.

Method	Dataset	Temp. corr. ↑	Lag-ACF ↑	Energy → 1	Dyn. frac. → 1	Adjacent → 1
GVD-CFM	BNCI2014_001	*0.920±0.006*	**0.999±0.000**	**0.995±0.023**	**1.024±0.052**	1.070±0.013
	BNCI2014_002	**0.838±0.003**	**0.999±0.000**	0.948±0.018	*0.938±0.036*	*0.977±0.017*
	BNCI2015_001	**0.958±0.001**	**0.999±0.000**	0.938±0.016	**0.976±0.007**	**0.976±0.015**
	Shin2017A	**0.958±0.002**	0.994±0.001	*1.068±0.006*	**1.161±0.011**	1.317±0.005
	Zhou2016	**0.877±0.017**	*0.992±0.005*	*0.815±0.039*	0.835±0.011	0.871±0.027
	5-dataset avg.	**0.910**	*0.997*	**0.953**	**0.987**	1.042
GVD-cVAE	BNCI2014_001	0.816±0.024	*0.998±0.000*	0.219±0.005	0.239±0.007	0.193±0.002
	BNCI2014_002	0.704±0.038	0.997±0.002	0.371±0.012	0.408±0.009	0.327±0.008
	BNCI2015_001	0.852±0.012	0.995±0.000	0.382±0.026	0.446±0.034	0.306±0.023
	Shin2017A	*0.802±0.039*	*0.995±0.002*	0.134±0.020	0.159±0.026	0.117±0.016
	Zhou2016	0.699±0.034	0.992±0.003	0.193±0.011	0.192±0.014	0.182±0.009
	5-dataset avg.	0.774	0.995	0.260	0.289	0.225
GVD-DDPM	BNCI2014_001	0.038±0.004	0.183±0.077	0.694±0.020	0.714±0.014	0.863±0.025
	BNCI2014_002	0.029±0.004	0.215±0.284	0.759±0.006	0.871±0.005	0.930±0.009
	BNCI2015_001	0.138±0.009	0.791±0.048	0.788±0.013	*0.877±0.015*	*0.972±0.014*
	Shin2017A	-0.011±0.002	-0.274±0.150	5.466±0.020	4.467±0.025	8.566±0.028
	Zhou2016	0.103±0.023	0.718±0.066	0.732±0.012	0.788±0.026	*0.916±0.016*
	5-dataset avg.	0.059	0.327	1.688	1.543	2.450
Window-DIFFEO-CFM	BNCI2014_001	0.011±0.006	0.201±0.167	1.487±0.005	1.464±0.009	1.846±0.010
	BNCI2014_002	0.003±0.009	0.157±0.268	1.746±0.009	1.794±0.010	2.142±0.012
	BNCI2015_001	0.018±0.014	0.053±0.244	1.256±0.007	1.299±0.014	1.553±0.010
	Shin2017A	-0.003±0.017	-0.214±0.070	3.174±0.035	2.980±0.018	4.981±0.055
	Zhou2016	0.029±0.001	0.080±0.210	1.308±0.019	1.238±0.011	1.641±0.024
	5-dataset avg.	0.012	0.055	1.794	1.755	2.433

Table 13: Full per-dataset temporal and dynamic diagnostics (continued).

Method	Dataset	Temp. corr. ↑	Lag-ACF ↑	Energy → 1	Dyn. frac. → 1	Adjacent → 1
cVAE	BNCI2014_001	0.399±0.031	0.948±0.017	0.986±0.014	1.381±0.031	*1.039±0.023*
	BNCI2014_002	0.512±0.060	0.976±0.006	**0.956±0.011**	0.850±0.014	**1.013±0.008**
	BNCI2015_001	0.798±0.019	0.990±0.003	**1.023±0.014**	0.520±0.001	1.038±0.013
	Shin2017A	0.216±0.033	0.913±0.026	0.651±0.018	*1.290±0.060*	**0.886±0.023**
	Zhou2016	0.710±0.018	0.958±0.004	**0.876±0.005**	*0.862±0.006*	**0.955±0.006**
	5-dataset avg.	0.527	0.957	0.898	*0.981*	**0.986**
JET	BNCI2014_001	0.357±0.037	0.937±0.012	1.897±0.235	9.197±1.470	1.268±0.096
	BNCI2014_002	0.399±0.051	0.983±0.003	1.474±0.087	9.324±0.386	1.177±0.025
	BNCI2015_001	0.490±0.040	0.985±0.006	1.542±0.109	6.338±0.175	1.146±0.019
	Shin2017A	0.375±0.047	0.785±0.037	2.080±0.138	8.183±0.626	1.456±0.056
	Zhou2016	0.306±0.035	0.977±0.017	2.709±0.502	**1.135±0.361**	1.912±0.242
	5-dataset avg.	0.385	0.933	1.940	6.835	1.392
Vanilla-Diffusion	BNCI2014_001	**0.930±0.013**	0.998±0.002	*0.992±0.024*	*1.185±0.403*	1.041±0.013
	BNCI2014_002	*0.832±0.006*	*0.998±0.000*	*0.954±0.075*	**0.943±0.256**	0.973±0.051
	BNCI2015_001	*0.955±0.003*	*0.999±0.000*	*0.949±0.011*	0.688±0.104	0.948±0.014
	Shin2017A	0.323±0.063	**0.996±0.002**	**1.023±0.096**	3.645±0.568	*1.282±0.108*
	Zhou2016	*0.875±0.027*	**0.998±0.000**	0.788±0.139	2.337±1.953	0.831±0.097
	5-dataset avg.	*0.783*	**0.998**	*0.941*	1.759	*1.015*
EEGGAN-2025	BNCI2014_001	-0.079±0.004	-0.228±0.011	0.830±0.044	3.490±0.378	**1.038±0.055**
	BNCI2014_002	0.078±0.008	0.820±0.089	0.622±0.018	8.267±0.550	0.761±0.024
	BNCI2015_001	0.098±0.023	0.759±0.142	0.574±0.020	6.277±0.214	0.706±0.024
	Shin2017A	-0.020±0.011	-0.206±0.141	0.408±0.006	2.072±0.231	0.640±0.010
	Zhou2016	0.035±0.045	0.385±0.650	0.610±0.027	4.216±0.671	0.765±0.031
	5-dataset avg.	0.022	0.306	0.609	4.864	0.782

Table 14: Full per-dataset novelty, diversity, and coverage diagnostics corresponding to Table 2. Entries are mean±standard deviation over three generator seeds. Diversity has ideal value 1.

Method	Dataset	Precision ↑	Diversity → 1	Coverage ↑
GVD-CFM	BNCI2014_001	0.967±0.012	*0.945±0.002*	**0.175±0.008**
	BNCI2014_002	0.899±0.026	*0.952±0.012*	**0.211±0.003**
	BNCI2015_001	0.859±0.037	0.956±0.007	**0.203±0.006**
	Shin2017A	0.032±0.036	**1.168±0.005**	**0.222±0.017**
	Zhou2016	**1.000±0.000**	0.914±0.022	**0.317±0.012**
	5-dataset avg.	0.751	**0.987**	**0.226**
GVD-cVAE	BNCI2014_001	**1.000±0.000**	0.363±0.013	*0.131±0.009*
	BNCI2014_002	**1.000±0.000**	0.505±0.003	*0.138±0.004*
	BNCI2015_001	**1.000±0.000**	0.497±0.018	*0.129±0.011*
	Shin2017A	**0.999±0.002**	0.276±0.027	*0.058±0.014*
	Zhou2016	*1.000±0.000*	0.338±0.021	*0.185±0.006*
	5-dataset avg.	**1.000**	0.396	*0.128*
GVD-DDPM	BNCI2014_001	*1.000±0.000*	0.827±0.014	0.030±0.006
	BNCI2014_002	*1.000±0.000*	0.871±0.000	0.025±0.002
	BNCI2015_001	*1.000±0.000*	0.884±0.011	0.034±0.001
	Shin2017A	0.000±0.000	3.391±0.000	0.039±0.004
	Zhou2016	1.000±0.001	0.863±0.002	0.063±0.006
	5-dataset avg.	*0.800*	1.367	0.038
Window-DIFFEO-CFM	BNCI2014_001	0.000±0.000	1.156±0.003	0.030±0.003
	BNCI2014_002	0.000±0.000	1.133±0.006	0.021±0.001
	BNCI2015_001	0.000±0.000	1.077±0.004	0.031±0.001
	Shin2017A	0.000±0.000	1.488±0.007	0.014±0.002
	Zhou2016	0.000±0.000	1.092±0.001	0.057±0.002
	5-dataset avg.	0.000	1.189	0.031

Table 15: Full per-dataset novelty, diversity, and coverage diagnostics (continued).

Method	Dataset	Precision ↑	Diversity → 1	Coverage ↑
cVAE	BNCI2014_001	0.799±0.090	0.709±0.002	0.027±0.001
	BNCI2014_002	0.615±0.016	0.890±0.015	0.029±0.002
	BNCI2015_001	0.101±0.031	*1.037±0.012*	0.022±0.003
	Shin2017A	*0.875±0.122*	0.503±0.034	0.008±0.000
	Zhou2016	0.572±0.024	**0.946±0.009**	0.085±0.005
	5-dataset avg.	0.593	0.817	0.034
JET	BNCI2014_001	0.000±0.000	1.340±0.169	0.011±0.005
	BNCI2014_002	0.000±0.000	0.870±0.063	0.004±0.001
	BNCI2015_001	0.000±0.000	0.833±0.032	0.004±0.001
	Shin2017A	0.000±0.000	1.528±0.112	0.006±0.000
	Zhou2016	0.000±0.000	2.936±0.331	0.023±0.020
	5-dataset avg.	0.000	1.501	0.009
Vanilla-Diffusion	BNCI2014_001	0.022±0.023	**1.046±0.006**	0.089±0.034
	BNCI2014_002	0.098±0.008	**1.032±0.038**	0.111±0.016
	BNCI2015_001	0.049±0.007	**1.027±0.006**	0.094±0.005
	Shin2017A	0.000±0.000	1.303±0.134	0.011±0.004
	Zhou2016	0.110±0.053	*0.943±0.203*	0.119±0.034
	5-dataset avg.	0.056	1.070	0.085
EEGGAN-2025	BNCI2014_001	0.000±0.000	1.392±0.060	0.004±0.000
	BNCI2014_002	0.000±0.000	0.778±0.046	0.002±0.001
	BNCI2015_001	0.000±0.000	0.707±0.033	0.001±0.000
	Shin2017A	0.000±0.000	*1.182±0.046*	0.003±0.001
	Zhou2016	0.000±0.000	1.106±0.016	0.004±0.001
	5-dataset avg.	0.000	*1.033*	0.003

Table 16: Full per-dataset DCT and temporal-branch ablation corresponding to Table 3: generative fidelity and utility. No DCT and DCT, spectral only are trained without the temporal branch; the last row of each block is the full GVD-CFM. Entries are mean±standard deviation over three generator seeds. Best values are **bold**; second-best values are *underlined italics*.

Dataset	Variant	Rel. GVD-FID ↓	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑
BNCI2014_001	No DCT	1.030±0.017	0.641±0.090	0.773±0.027	0.691±0.023
	DCT, spectral only	*1.016±0.015*	*0.825±0.021*	*0.784±0.006*	*0.695±0.005*
	GVD-CFM (DCT + temporal branch)	**1.006±0.014**	**0.864±0.027**	**0.789±0.015**	**0.707±0.007**
BNCI2014_002	No DCT	1.036±0.009	0.628±0.090	0.757±0.009	0.685±0.009
	DCT, spectral only	*0.988±0.012*	*0.723±0.042*	**0.789±0.006**	**0.716±0.006**
	GVD-CFM (DCT + temporal branch)	**0.977±0.016**	**0.740±0.022**	*0.780±0.015*	*0.709±0.005*
BNCI2015_001	No DCT	1.009±0.006	**0.889±0.026**	0.704±0.002	0.640±0.011
	DCT, spectral only	*0.992±0.002*	*0.789±0.059*	**0.761±0.003**	**0.692±0.004**
	GVD-CFM (DCT + temporal branch)	**0.978±0.005**	0.775±0.073	*0.746±0.011*	*0.681±0.002*
Shin2017A	No DCT	1.400±0.143	0.178±0.045	0.661±0.015	0.601±0.018
	DCT, spectral only	**1.153±0.022**	*0.351±0.028*	*0.664±0.008*	*0.605±0.017*
	GVD-CFM (DCT + temporal branch)	*1.190±0.084*	**0.482±0.039**	**0.672±0.029**	**0.622±0.025**
Zhou2016	No DCT	0.985±0.002	**0.364±0.021**	**0.967±0.009**	*0.903±0.007*
	DCT, spectral only	*0.978±0.003*	*0.311±0.047*	0.964±0.002	0.896±0.003
	GVD-CFM (DCT + temporal branch)	**0.959±0.033**	0.203±0.110	*0.965±0.003*	**0.905±0.014**
5-dataset avg.	No DCT	1.092	0.540	0.772	0.704
5-dataset avg.	DCT, spectral only	*1.025*	*0.600*	**0.792**	*0.721*
5-dataset avg.	GVD-CFM (DCT + temporal branch)	**1.022**	**0.613**	*0.790*	**0.725**

Table 17: Full per-dataset DCT and temporal-branch ablation: temporal and dynamic fidelity. Entries are mean±standard deviation over three generator seeds. For ratio metrics, ranking is by proximity to 1. Best values are **bold**; second-best values are *underlined italics*.

Dataset	Variant	Temp. corr. ↑	Lag-ACF ↑	Energy → 1	Dyn. frac. → 1	Adjacent → 1
BNCI2014_001	No DCT	0.065±0.022	*0.465±0.158*	*0.992±0.027*	**1.005±0.059**	1.231±0.034
	DCT, spectral only	**0.939±0.003**	**0.999±0.000**	1.013±0.026	1.025±0.031	**1.052±0.024**
	GVD-CFM (DCT + temporal branch)	*0.920±0.006*	**0.999±0.000**	**0.995±0.023**	*1.024±0.052*	*1.070±0.013*
BNCI2014_002	No DCT	0.140±0.018	*0.649±0.177*	*0.974±0.019*	**0.980±0.013**	1.193±0.024
	DCT, spectral only	**0.842±0.003**	**0.999±0.001**	**0.984±0.032**	*0.963±0.002*	**0.990±0.024**
	GVD-CFM (DCT + temporal branch)	*0.838±0.003*	**0.999±0.000**	0.948±0.018	0.938±0.036	*0.977±0.017*
BNCI2015_001	No DCT	0.934±0.003	*0.984±0.002*	**0.971±0.005**	*1.010±0.010*	1.051±0.007
	DCT, spectral only	*0.956±0.000*	**0.999±0.000**	*0.953±0.015*	**0.995±0.004**	*0.971±0.015*
	GVD-CFM (DCT + temporal branch)	**0.958±0.001**	**0.999±0.000**	0.938±0.016	0.976±0.007	**0.976±0.015**
Shin2017A	No DCT	0.016±0.020	0.252±0.251	1.084±0.074	1.199±0.051	1.695±0.116
	DCT, spectral only	**0.961±0.003**	**0.995±0.000**	*1.080±0.033*	*1.170±0.022*	*1.318±0.020*
	GVD-CFM (DCT + temporal branch)	*0.958±0.002*	*0.994±0.001*	**1.068±0.006**	**1.161±0.011**	**1.317±0.005**
Zhou2016	No DCT	0.865±0.006	0.981±0.003	*0.842±0.002*	*0.861±0.030*	**0.942±0.004**
	DCT, spectral only	**0.895±0.009**	**0.994±0.003**	**0.847±0.001**	**0.898±0.035**	*0.898±0.005*
	GVD-CFM (DCT + temporal branch)	*0.877±0.017*	*0.992±0.005*	0.815±0.039	0.835±0.011	0.871±0.027
5-dataset avg.	No DCT	0.404	*0.666*	*0.973*	*1.011*	1.222
5-dataset avg.	DCT, spectral only	**0.919**	**0.997**	**0.975**	**1.010**	*1.046*
5-dataset avg.	GVD-CFM (DCT + temporal branch)	*0.910*	**0.997**	0.953	0.987	**1.042**

Table 18: Paired per-dataset comparison for the DCT and temporal-branch ablation. Each entry is the change in the dataset mean (second variant minus first) followed by the Welch t-statistic computed from three generator seeds per variant, $t = \Delta/\sqrt{(\sigma_1^2 + \sigma_2^2)/3}$. Entries with $|t| \geq 2$ are **bold**; a dash marks zero seed variance. For Rel. GVD-FID a negative change is an improvement; for Energy the ideal value is 1.

Comparison	Metric	BNCI2014_001	BNCI2014_002	BNCI2015_001	Shin2017A	Zhou2016
No DCT → DCT, spectral only	Rel. GVD-FID	-0.014 (-1.1)	-0.048 (**-5.5**)	-0.017 (**-4.7**)	-0.247 (**-3.0**)	-0.007 (**-3.4**)
	Eva F1	+0.184 (**3.4**)	+0.095 (1.7)	-0.100 (**-2.7**)	+0.173 (**5.7**)	-0.053 (-1.8)
	CAS AUC	+0.011 (0.7)	+0.032 (**5.1**)	+0.057 (**27.4**)	+0.003 (0.3)	-0.003 (-0.6)
	CAS F1	+0.004 (0.3)	+0.031 (**5.0**)	+0.052 (**7.7**)	+0.004 (0.3)	-0.007 (-1.6)
	Temp. corr.	+0.874 (**68.2**)	+0.702 (**66.6**)	+0.022 (**12.7**)	+0.945 (**80.9**)	+0.030 (**4.8**)
	Lag-ACF	+0.534 (**5.9**)	+0.350 (**3.4**)	+0.015 (**13.0**)	+0.743 (**5.1**)	+0.013 (**5.3**)
	Energy	+0.021 (1.0)	+0.010 (0.5)	-0.018 (-2.0)	-0.004 (-0.1)	+0.005 (**3.9**)
DCT, spectral only → GVD-CFM	Rel. GVD-FID	-0.010 (-0.8)	-0.011 (-1.0)	-0.014 (**-4.5**)	+0.037 (0.7)	-0.019 (-1.0)
	Eva F1	+0.039 (2.0)	+0.017 (0.6)	-0.014 (-0.3)	+0.131 (**4.7**)	-0.108 (-1.6)
	CAS AUC	+0.005 (0.5)	-0.009 (-1.0)	-0.015 (**-2.3**)	+0.008 (0.5)	+0.001 (0.5)
	CAS F1	+0.012 (**2.4**)	-0.007 (-1.6)	-0.011 (**-4.3**)	+0.017 (1.0)	+0.009 (1.1)
	Temp. corr.	-0.019 (**-4.9**)	-0.004 (-1.6)	+0.002 (**3.5**)	-0.003 (-1.4)	-0.018 (-1.6)
	Lag-ACF	+0.000 (–)	+0.000 (0.0)	+0.000 (–)	-0.001 (-1.7)	-0.002 (-0.6)
	Energy	-0.018 (-0.9)	-0.036 (-1.7)	-0.015 (-1.2)	-0.012 (-0.6)	-0.032 (-1.4)

Table 19: Full per-dataset stable-support ablation summarized in Section 6.3. GVD-CFM uses the canonical ridge-free stable support; the no-support control uses the ridge required to obtain SPD matrices and the spectral-only velocity network. Entries are mean±standard deviation over three generator seeds. Best values are **bold**; second-best values are *underlined italics*.

Dataset	Variant	Rel. GVD-FID ↓	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑	Temp. corr. ↑	Lag-ACF ↑	Energy → 1
BNCI2014_001	GVD-CFM (stable support)	*1.006±0.014*	**0.864±0.027**	**0.789±0.015**	**0.707±0.007**	*0.920±0.006*	*0.999±0.000*	**0.995±0.023**
	No support + ridge	**0.917±0.006**	*0.292±0.032*	*0.667±0.019*	*0.614±0.017*	**0.977±0.001**	**1.000±0.000**	*0.826±0.007*
BNCI2014_002	GVD-CFM (stable support)	*0.977±0.016*	**0.740±0.022**	**0.780±0.015**	**0.709±0.005**	*0.838±0.003*	**0.999±0.000**	**0.948±0.018**
	No support + ridge	**0.909±0.001**	*0.481±0.024*	*0.736±0.009*	*0.674±0.010*	**0.929±0.001**	**0.999±0.000**	*0.850±0.010*
BNCI2015_001	GVD-CFM (stable support)	*0.978±0.005*	**0.775±0.073**	**0.746±0.011**	**0.681±0.002**	*0.958±0.001*	**0.999±0.000**	**0.938±0.016**
	No support + ridge	**0.921±0.005**	*0.372±0.050*	*0.698±0.012*	*0.638±0.006*	**0.980±0.000**	**0.999±0.000**	*0.874±0.006*
Shin2017A	GVD-CFM (stable support)	*1.190±0.084*	*0.482±0.039*	**0.672±0.029**	**0.622±0.025**	*0.958±0.002*	*0.994±0.001*	*1.068±0.006*
	No support + ridge	**0.999±0.010**	**0.800±0.035**	*0.622±0.008*	*0.584±0.005*	**0.971±0.002**	**0.996±0.000**	**0.990±0.021**
Zhou2016	GVD-CFM (stable support)	*0.959±0.033*	*0.203±0.110*	**0.965±0.003**	**0.905±0.014**	*0.877±0.017*	**0.992±0.005**	**0.815±0.039**
	No support + ridge	**0.922±0.004**	**0.237±0.021**	*0.955±0.007*	*0.883±0.010*	**0.905±0.007**	*0.980±0.007*	*0.794±0.007*
5-dataset avg.	GVD-CFM (stable support)	*1.022*	**0.613**	**0.790**	**0.725**	*0.910*	**0.997**	**0.953**
5-dataset avg.	No support + ridge	**0.934**	*0.436*	*0.736*	*0.679*	**0.952**	*0.995*	*0.867*

Table 20: Full temporal-basis ablation across the five main EEG datasets: generative fidelity and downstream utility. Each per-dataset entry is mean±standard deviation over three generator seeds. All four variants retain all $B = 100$ temporal coordinates. Random orthogonal uses a fixed random orthogonal basis, KLT the training-set empirical temporal Karhunen–Loève basis, and GVD-CFM (DCT) the fixed orthonormal DCT-II basis with the full network including the temporal branch; the other variants use the spectral-only velocity network. Best values are **bold**; second-best values are *underlined italics*.

Dataset	Basis	Rel. GVD-FID ↓	Eva F1 ↑	CAS AUC ↑	CAS F1 ↑
BNCI2014_001	No DCT	1.030±0.017	0.641±0.090	0.773±0.027	0.691±0.023
	Random orthogonal	1.071±0.003	0.461±0.045	0.780±0.021	0.695±0.011
	GVD-CFM (DCT)	**1.006±0.014**	**0.864±0.027**	**0.789±0.015**	**0.707±0.007**
	KLT	*1.023±0.006*	*0.802±0.008*	*0.787±0.012*	*0.697±0.010*
BNCI2014_002	No DCT	1.036±0.009	0.628±0.090	0.757±0.009	0.685±0.009
	Random orthogonal	1.048±0.015	0.578±0.050	*0.783±0.006*	0.700±0.015
	GVD-CFM (DCT)	**0.977±0.016**	*0.740±0.022*	0.780±0.015	*0.709±0.005*
	KLT	*0.984±0.002*	**0.758±0.025**	**0.794±0.035**	**0.737±0.039**
BNCI2015_001	No DCT	1.009±0.006	**0.889±0.026**	0.704±0.002	0.640±0.011
	Random orthogonal	1.031±0.019	0.652±0.121	*0.754±0.014*	*0.683±0.013*
	GVD-CFM (DCT)	**0.978±0.005**	0.775±0.073	0.746±0.011	0.681±0.002
	KLT	*0.994±0.003*	*0.846±0.022*	**0.755±0.013**	**0.690±0.010**
Shin2017A	No DCT	1.400±0.143	0.178±0.045	0.661±0.015	0.601±0.018
	Random orthogonal	1.304±0.041	0.075±0.009	0.659±0.014	0.596±0.015
	GVD-CFM (DCT)	*1.190±0.084*	**0.482±0.039**	*0.672±0.029*	**0.622±0.025**
	KLT	**1.183±0.036**	*0.316±0.036*	**0.675±0.006**	*0.620±0.008*
Zhou2016	No DCT	0.985±0.002	*0.364±0.021*	**0.967±0.009**	*0.903±0.007*
	Random orthogonal	1.013±0.018	**0.636±0.172**	0.964±0.002	0.891±0.006
	GVD-CFM (DCT)	**0.959±0.033**	0.203±0.110	0.965±0.003	**0.905±0.014**
	KLT	*0.979±0.003*	0.277±0.035	*0.966±0.002*	0.901±0.005
5-dataset avg.	No DCT	1.092	0.540	0.772	0.704
5-dataset avg.	Random orthogonal	1.093	0.480	0.788	0.713
5-dataset avg.	GVD-CFM (DCT)	**1.022**	**0.613**	*0.790*	*0.725*
5-dataset avg.	KLT	*1.033*	*0.600*	**0.795**	**0.729**