LEARNING NEUROVASCULAR CODES FOR EEG-TO-FMRI SYNTHESIS

Anonymous authors
Paper under double-blind review

ABSTRACT

In this paper, we introduce NEUROVASCULAR CODES as shared EEG representations for fMRI generation. We first train a separate EEG encoder to predict each fMRI signal, learning target-specific temporal and spectral features from the same EEG window. After training, we freeze the encoders and project their concatenated hidden features into a shared neurovascular code. Conditioned on this code, *stochastic interpolant* transforms Gaussian noise into a vector of fMRI signals, allowing different responses to be sampled for the same EEG window. We evaluate our approach on eight public simultaneous EEG-fMRI datasets, including four resting-state and four task-based datasets with 61,646 volume-level pairs. Under within-subject and held-out-subject evaluation, our model improves average Pearson correlation over the strongest encoder baselines by 19.1% and 19.8%, respectively. The corresponding improvements over the strongest generator baselines are 18.8% and 47.9%. Beyond reconstruction accuracy, we also evaluate the generated signals using subject-identification classifiers trained on real fMRI. Our method achieves the highest mean accuracy among the evaluated generators. We hope that our work opens up new opportunities for cross-modal brain modeling.

1 INTRODUCTION

Learning representations that transfer across tasks is a central aim of foundation models (Bommasani et al., 2021). This approach has been extended to brain signals, including electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). EEG models (Kostas et al., 2021; Jiang et al., 2024; Wang et al., 2024; Ouahidi et al., 2025) learn across subjects and recording settings, while fMRI models such as BrainLM (Caro et al., 2023) learn spatiotemporal patterns that support clinical prediction (Caro et al., 2023). These models learn from each modality separately. Here, we study how fMRI prediction can supervise EEG representations for cross-modal generation.

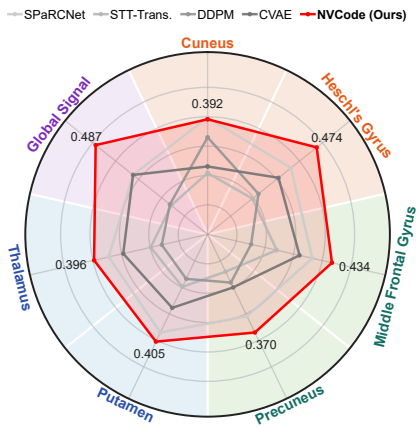


Figure 1: **EEG-to-fMRI synthesis on held-out subjects.** Pearson correlation for our method and representative encoder and generator baselines across seven fMRI targets, averaged over eight datasets.

EEG and fMRI provide complementary measurements of brain activity. EEG captures rapid electrical fluctuations and is widely used in brain-computer interfaces (McFarland & Wolpaw, 2011; Lotte et al., 2018). fMRI measures blood-oxygen-level-dependent (BOLD) signals (Ogawa et al., 1990; Logothetis, 2008) and helps characterize the spatial organization of brain regions and networks (Smith et al., 2013; Glasser et al., 2016). Simultaneous EEG-fMRI recordings (Huster et al., 2012; Murta et al., 2015; Bullock et al., 2021) provide paired observations for learning the relationship between these measurements. Acquiring these recordings requires specialized recording hardware (Allen et al., 2000) and extensive artifact suppression (Niazy et al., 2005; Huster et al., 2012; Bullock et al., 2021), making paired data difficult to collect at scale. We focus on EEG-to-fMRI synthesis because EEG can be collected outside the scanner, allowing a trained model to generate

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.

fMRI signals without an accompanying fMRI scan. Previous EEG-to-fMRI studies (Kovalev et al., 2022; Li et al., 2024a;b) show that EEG contains predictive information about individual fMRI signals. We thus use these prediction tasks to learn representations for fMRI generation.

In this paper, we introduce NEUROVASCULAR CODES, termed **NVCode**, for EEG-to-fMRI synthesis. We first train a separate EEG encoder to predict each fMRI signal from the same EEG window. Through these prediction tasks, the encoders learn target-specific spatiotemporal (Vaswani et al., 2017; Dosovitskiy et al., 2020) and spectral (Schirrmeister et al., 2017; Banville et al., 2021) features. After training, we discard the prediction heads and freeze the encoders, retaining their hidden features for generation. A learned projection maps the concatenated features into a shared neurovascular code that conditions a stochastic fMRI generator.

An EEG window may correspond to more than one plausible fMRI response (Figure 2). Scalp EEG has spatial ambiguity (Huster et al., 2012), and neurovascular coupling varies across regions and individuals (Logothetis, 2008). We therefore condition a *stochastic interpolant* on the neurovascular code, using conditional velocity matching (Lipman et al., 2022; Tong et al., 2023). The model learns a velocity field that transforms Gaussian noise into a vector of fMRI signals. Each velocity prediction depends on the current values of all target signals, allowing the field to model dependencies among them. Different noise samples yield different response vectors for the same EEG window. Averaging these samples gives a predictive mean for each fMRI target.

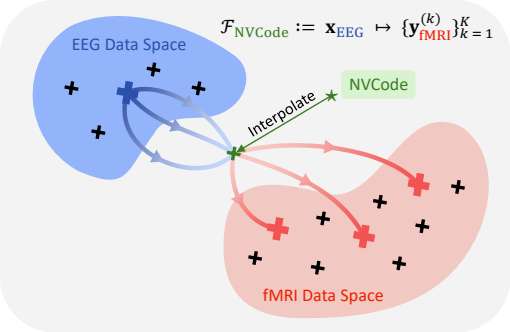


Figure 2: **NVCode** uses a shared neurovascular code to map the same EEG window to multiple plausible joint fMRI response vectors.

We evaluate our approach on eight public simultaneous EEG-fMRI datasets, including four resting-state and four task-based datasets with 61,646 volume-level pairs. The benchmark uses common target definitions, data splits, and evaluation metrics. We compare EEG encoders using a common generator architecture, and compare generators using the same frozen EEG encoders. We also ablate which target-specific representations are supplied to the generator. Under within-subject and held-out-subject evaluation, our method achieves relative gains in average Pearson correlation of 19.1% and 19.8% over the strongest encoder baselines, respectively. The corresponding gains over the strongest generator baselines are 18.8% and 47.9%.

Beyond prediction accuracy, we evaluate marginal predictive distributions and functional connectivity computed from predictive-mean time series. Our method achieves the lowest average functional connectivity error in both comparisons and evaluation settings. With classifiers trained on real fMRI and held fixed, our generated signals also achieve the highest mean subject-identification accuracy among the evaluated generators, exceeding accuracy on real fMRI test signals by 3.1 percentage points on average. Our contributions are as follows:

1. We introduce NEUROVASCULAR CODES, shared EEG representations learned through target-specific fMRI prediction and used to condition stochastic fMRI generation.
2. We establish a benchmark across eight public simultaneous EEG-fMRI datasets, covering four resting-state and four task-based datasets. It uses common target definitions, data splits, and evaluation metrics for comparisons in both within-subject and held-out-subject settings.
3. We compare EEG encoders, conditional generators, and target-specific representation sets under controlled experimental settings. Beyond reconstruction accuracy, we evaluate the marginal predictive distributions and functional connectivity computed from predictive-mean time series. We assess subject identification performance using classifiers trained on real fMRI signals.

2 RELATED WORK

EEG and fMRI representation learning. Representation learning is widely used in brain–computer interfaces (McFarland & Wolpaw, 2011; Lotte et al., 2018). Large-scale pretraining provides one way to learn features that transfer across tasks (Bommasani et al., 2021). For EEG, BENDR (Kostas

et al., 2021) uses masked contrastive learning, while BIOT (Yang et al., 2023a) enables pretraining across heterogeneous biosignal recordings. LaBraM (Jiang et al., 2024) predicts masked neural tokens, EEGPT (Wang et al., 2024) pairs masked reconstruction with representation alignment, and REVE (Ouahidi et al., 2025) uses masked autoencoding with electrode-coordinate embeddings. This line of work addresses variation in electrode layouts (Yang et al., 2023a; Ouahidi et al., 2025) and supports a range of EEG decoding tasks (Lotte et al., 2018; Wang et al., 2024). For fMRI, BrainLM (Caro et al., 2023) learns spatial and temporal structure through masked prediction and supports clinical prediction (Caro et al., 2023). During pretraining, these methods construct learning targets from the input signals themselves. Our work instead uses fMRI prediction to supervise EEG representation learning.

EEG-to-fMRI modeling. EEG–fMRI studies relate electrophysiological activity to the BOLD response (Ogawa et al., 1990; Logothetis, 2008) to investigate neurovascular coupling (Logothetis, 2008; Murta et al., 2015) and constrain EEG source localization (Huster et al., 2012). fMRI also supports the study of functional networks (Biswal et al., 1995; Van Den Heuvel & Pol, 2010) through statistical dependencies across brain regions (Biswal et al., 1995; Friston, 2011). BEIRA (Kovalev et al., 2022) learns interpretable mappings from EEG to BOLD signals, while SIREN-based prediction (Li et al., 2024a) uses sinusoidal networks to extract frequency information. NeuroBOLT (Li et al., 2024b) learns spatiotemporal and multiscale spectral features and evaluates prediction on both resting-state and task recordings. These methods produce point predictions for individual fMRI targets. NVCode retains target-specific prediction as a representation-learning task, then uses the resulting features to condition generation of the fMRI signal vector.

Conditional generative models. EEG-to-fMRI ambiguity and variable neurovascular coupling motivate probabilistic modeling (Logothetis, 2008; Murta et al., 2015). Models use stochastic latents (Kingma & Welling, 2013; Sohn et al., 2015), adversarial training (Mirza & Osindero, 2014; Isola et al., 2017), diffusion denoising (Ho et al., 2020; Song et al., 2020; Karras et al., 2022), or cross-time consistency for few-step sampling (Song et al., 2023; Yang et al., 2024). Neural ODEs and FFJORD formulate continuous normalizing flows (Chen et al., 2018; Grathwohl et al., 2018), whereas flow matching, rectified flow, and optimal-transport conditional flow matching learn path velocities (Lipman et al., 2022; Liu et al., 2022; Tong et al., 2023). Stochastic interpolants connect distributions through deterministic or stochastic dynamics (Albergo et al., 2025). NVCode conditions this framework on neurovascular codes and the full fMRI state, allowing interactions among targets.

3 Neurovascular Codes

NVCode learns EEG representations through individual fMRI prediction tasks and uses these representations to condition fMRI generation. We first train a separate EEG encoder for each fMRI target (Section 3.2). With the encoders frozen, we then train a projector and a conditional velocity field together (Sections 3.3 and 3.4). The projector maps the encoder features to a neurovascular code, which guides generation from Gaussian noise to a vector of fMRI signals. Appendix B provides architectural specifications, derivations, and training and sampling details.

3.1 Preliminaries

Each observation pairs an EEG window $\mathbf{x} \in \mathbb{R}^{C\times T}$ with an aligned fMRI target vector $\mathbf{y} \in \mathbb{R}^N$. Here, C, T, and N count EEG channels, time samples, and fMRI targets, respectively. The vector $\mathbf{y}$ contains one volume's regional and global signals, modeled conditionally on EEG.

Projecting the target-specific EEG encoders' hidden features yields the neurovascular code $\mathbf{c}_\psi(\mathbf{x}) \in \mathbb{R}^{d_c}$ (Section 3.3). It conditions $p_\theta(\mathbf{y} \mid \mathbf{c}_\psi(\mathbf{x}))$, with projector parameters ψ and generator parameters θ. Each generated vector represents one fMRI volume; applying the model to successive aligned EEG windows yields a time series for evaluation. Let $\boldsymbol{\mu}, \boldsymbol{\sigma} \in \mathbb{R}^N$ denote the per-target mean and standard deviation estimated only from the training split. The standardized generation endpoint is

$$\mathbf{z}_1 = (\mathbf{y} - \boldsymbol{\mu}) \oslash \boldsymbol{\sigma}, \tag{1}$$

where $\oslash$ denotes elementwise division. Normalization provides a common scale for velocity regression; its inverse restores the fMRI target scale before evaluation.

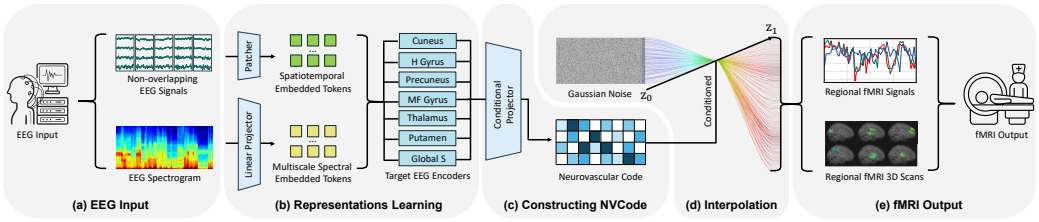


Figure 3: **NVCode pipeline. (a)** An EEG window provides non-overlapping signal patches and spectrograms. **(b)** Target-specific encoders learn spatiotemporal and multiscale spectral representations under fMRI supervision. **(c)** A shared projector maps their concatenated features to a neurovascular code. **(d)** The code conditions joint generation from Gaussian noise through a learned interpolant. **(e)** Generated regional and global fMRI signals are displayed as time series and anatomical projections.

3.2 Learning Target-Specific EEG Representations

Each EEG encoder extracts spatiotemporal and multiscale spectral features to predict one fMRI target from the same window. For target n, we attach a linear prediction head to the encoder features and minimize the squared difference between its prediction and the corresponding fMRI target:

$$\begin{aligned} \mathbf{e}_n = g_{\phi_n}(\mathbf{x}) \in \mathbb{R}^d, \quad \widehat{y}_n^{\text{aux}} = \mathbf{a}_n^\top \mathbf{e}_n + b_n, \\ \mathcal{L}_{\text{rep}}^{(n)} = \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim\mathcal{D}_{\text{tr}}}\left[(\widehat{y}_n^{\text{aux}} - y_n)^2\right], \quad n = 1, \dots, N. \end{aligned} \tag{2}$$

Here, $\mathcal{D}_{\text{tr}}$ is the training set and d the feature dimension. Encoder–head pairs are fitted independently and selected by validation mean squared error (MSE). We discard the temporary heads and freeze the selected encoders, retaining their hidden features for code construction.

3.3 Constructing Neurovascular Codes

NVCode concatenates the frozen encoder outputs and applies a learned projector Π_ψ. Its first-layer weight matrix $\mathbf{W}_1 = [\mathbf{W}_1^{(1)} \ \cdots \ \mathbf{W}_1^{(N)}]$ is partitioned by encoder output. We define

$$\begin{aligned} \mathbf{e}(\mathbf{x}) = [\mathbf{e}_1; \dots; \mathbf{e}_N] \in \mathbb{R}^{Nd}, \quad \mathbf{c}_\psi(\mathbf{x}) = \Pi_\psi(\mathbf{e}(\mathbf{x})) \in \mathbb{R}^{d_c}, \\ \mathbf{W}_1\mathbf{e}(\mathbf{x}) + \mathbf{b}_1 = \sum_{n=1}^{N} \mathbf{W}_1^{(n)}\mathbf{e}_n + \mathbf{b}_1. \end{aligned} \tag{3}$$

Here, $[\cdot;\cdot]$ denotes concatenation, and $\mathbf{W}_1^{(n)}$ linearly transforms target n's features. The projector uses two linear layers with layer normalization and sigmoid linear unit (SiLU) activations (Appendix B.3). With target-supervised encoders frozen, generation jointly trains the projector and velocity field (Section 3.4). One code per EEG window conditions all targets throughout sampling.

3.4 Learning the Conditional Velocity Field

For each training pair $(\mathbf{x}, \mathbf{y})$, we compute the neurovascular code $\mathbf{c} = \mathbf{c}_\psi(\mathbf{x})$ and the normalized target $\mathbf{z}_1$ from Eq. (1). We draw $\mathbf{z}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$ and $t \sim \mathcal{U}(0,1)$ independently of the observation and each other. A linear interpolant connects the Gaussian source and normalized fMRI target, with its path velocity given by (Lipman et al., 2022; Tong et al., 2023; Albergo et al., 2025):

$$\mathbf{z}_t = (1-t)\mathbf{z}_0 + t\mathbf{z}_1, \qquad \mathbf{u}_t = \frac{d\mathbf{z}_t}{dt} = \mathbf{z}_1 - \mathbf{z}_0. \tag{4}$$

Here, t is interpolation time, not recording time. The state $\mathbf{z}_t$ and its velocity target $\mathbf{u}_t$ are computed directly from the endpoints, without integrating an ODE during training.

We project the full state $\mathbf{z}_t$ through a learned linear map and concatenate it with the code and time embedding $\boldsymbol{\tau}_t = \boldsymbol{\tau}(t)$. A multilayer perceptron (MLP) then predicts one velocity per target:

$$\begin{aligned} \mathbf{r}_t = \mathbf{A}\mathbf{z}_t + \mathbf{b}_A, \quad \mathbf{h}_t = [\mathbf{c}; \mathbf{r}_t; \boldsymbol{\tau}_t], \\ v_\theta(\mathbf{z}_t, t, \mathbf{c}) = \text{MLP}_\theta(\mathbf{h}_t), \quad D_{\mathbf{z}_t} v_\theta = \mathbf{J}_{\text{MLP},r}\mathbf{A}. \end{aligned} \tag{5}$$

The parameters θ include the state projection $(\mathbf{A}, \mathbf{b}_A)$ and the MLP. With the code and time fixed, $\mathbf{J}_{\text{MLP},r}$ is the Jacobian of the MLP output with respect to $\mathbf{r}_t$. The resulting state Jacobian need not be diagonal, allowing the velocity of each target to depend on other entries in $\mathbf{z}_t$.

We train the projector and velocity field by minimizing

$$\mathcal{L}_{\mathrm{NVCode}}(\psi,\theta)=\mathbb{E}_{(\mathbf{x},\mathbf{y}),\mathbf{z}_0,t}\left[\|v_\theta(\mathbf{z}_t,t,\mathbf{c}_\psi(\mathbf{x}))-\mathbf{u}_t\|_2^2\right]. \tag{6}$$

The expectation averages over training pairs, Gaussian source samples, and interpolation times. The loss sums squared velocity errors across all fMRI targets and jointly updates the projector and velocity field parameters. The EEG encoder parameters remain fixed throughout velocity field training.

Training. Training proceeds in two stages. We first fit each encoder–head pair to its corresponding fMRI target with AdamW:

$$\min_{\phi_n,\mathbf{a}_n,b_n}\mathcal{L}_{\mathrm{rep}}^{(n)}(\phi_n,\mathbf{a}_n,b_n),\quad n=1,\dots,N.$$

Validation MSE selects each encoder checkpoint $\phi_n^\star$. With these encoders fixed, the second stage optimizes

$$\min_{\psi,\theta}\mathcal{L}_{\mathrm{NVCode}}(\psi,\theta;\phi_1^\star,\dots,\phi_N^\star),$$
$$\phi_{1:N}=\phi_{1:N}^\star \text{ fixed.}$$

We use AdamW with OneCycle scheduling and resample $(\mathbf{z}_0,t)$ independently for every training example. The projector and velocity field are selected together, giving $(\psi^\star,\theta^\star)$. Algorithm 1 summarizes the two stages.

Algorithm 1 NVCode fitting

Input: Training data $\mathcal{D}_{\mathrm{tr}}$; validation data $\mathcal{D}_{\mathrm{val}}$; N targets.
Output: Frozen encoders $g_{\phi_{1:N}}$; projector Π_ψ; field v_θ; statistics $(\boldsymbol{\mu},\boldsymbol{\sigma})$.
Estimate $(\boldsymbol{\mu},\boldsymbol{\sigma})$ on training targets. ▷ **Representation fitting**
for $n=1,\dots,N$ **do**
  Fit g_{ϕ_n} and its head with Eq. (2); select by validation MSE.
  Freeze g_{ϕ_n} and discard its head.
end for
Initialize Π_ψ, v_θ. ▷ **Interpolant fitting**
repeat
  Draw a training batch; compute $\mathbf{c},\mathbf{z}_1$ and draw $\mathbf{z}_0,t$ independently per example.
  Form $(\mathbf{z}_t,\mathbf{u}_t)$; update (ψ,θ) on $\mathcal{L}_{\mathrm{NVCode}}$.
until the split-specific training schedule ends
Restore the selected projector–field checkpoint.

3.5 Generating fMRI Signals

For an unseen EEG window $\mathbf{x}^*$, we compute $\mathbf{c}_{\psi^\star}(\mathbf{x}^*)$ once from the frozen encoder features. This shared condition remains fixed for all generated samples. We draw K independent Gaussian initial states and integrate each trajectory through the learned velocity field from $t=0$ to $t=1$:

$$\frac{d\mathbf{s}^{(k)}(t)}{dt}=v_{\theta^\star}\left(\mathbf{s}^{(k)}(t),t,\mathbf{c}_{\psi^\star}(\mathbf{x}^*)\right),\qquad \mathbf{s}^{(k)}(0)\sim\mathcal{N}(\mathbf{0},\mathbf{I}_N). \tag{7}$$

Although the training paths in Eq. (4) are linear, the trajectories induced by the learned field need not be. Each trajectory uses the same EEG condition and a different initial noise sample.

With S forward Euler steps of size $1/S$, the endpoint is $\mathbf{s}_S^{(k)}$. We restore the fMRI target scale and average the K samples to estimate the predictive mean:

$$\mathbf{y}^{(k)}=\boldsymbol{\mu}+\boldsymbol{\sigma}\odot\mathbf{s}_S^{(k)},\quad \widehat{\mathbf{y}}=\frac{1}{K}\sum_{k=1}^{K}\mathbf{y}^{(k)},\quad \mathrm{Cov}(\widehat{\mathbf{y}}\mid\mathbf{x}^*)=\frac{1}{K}\mathrm{Cov}(\mathbf{y}^{(1)}\mid\mathbf{x}^*). \tag{8}$$

Here, $\odot$ denotes elementwise multiplication. For a fixed trained model, the covariance identity follows from independent sampling. Increasing K reduces the Monte Carlo variability of the estimated mean, not the spread of the generated response distribution.

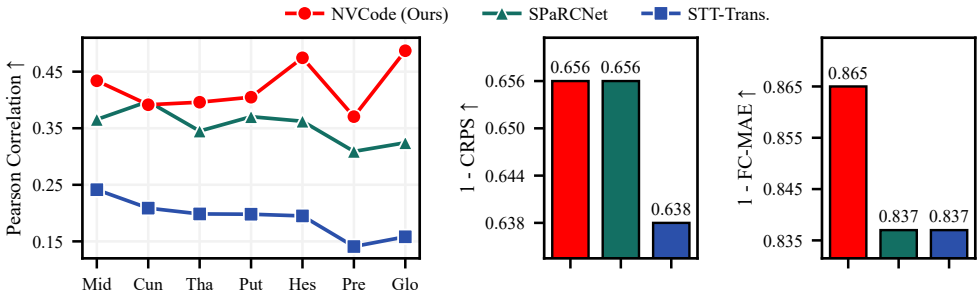
Algorithm 2 returns samples and their predictive mean. Across M aligned EEG windows, these form a $K\times M\times N$ sample tensor and an $M\times N$ predictive-mean time series. Pearson correlation and functional connectivity mean absolute error (FC-MAE) evaluate the time series; the continuous ranked probability score (CRPS) evaluates marginal sample distributions separately for each fMRI target.

Algorithm 2 NVCode sampling

Input: EEG window $\mathbf{x}^*$; frozen $g_{\phi_{1:N}},\Pi_\psi,v_\theta$; statistics $(\boldsymbol{\mu},\boldsymbol{\sigma})$; steps S; samples K.
Output: Samples $\{\mathbf{y}^{(k)}\}_{k=1}^K$ and predictive mean $\widehat{\mathbf{y}}$.
Compute $\mathbf{c}=\Pi_\psi([g_{\phi_1}(\mathbf{x}^*);\cdots;g_{\phi_N}(\mathbf{x}^*)])$ once.
for $k=1,\dots,K$ **do**
  Draw $\mathbf{s}_0^{(k)}\sim\mathcal{N}(\mathbf{0},\mathbf{I}_N)$ independently.
  for $q=0,\dots,S-1$ **do**
    $\mathbf{s}_{q+1}^{(k)}\leftarrow\mathbf{s}_q^{(k)}+S^{-1}v_\theta(\mathbf{s}_q^{(k)},q/S,\mathbf{c})$.
  end for
  $\mathbf{y}^{(k)}\leftarrow\boldsymbol{\mu}+\boldsymbol{\sigma}\odot\mathbf{s}_S^{(k)}$.
end for
Compute $\widehat{\mathbf{y}}\leftarrow K^{-1}\sum_{k=1}^K\mathbf{y}^{(k)}$.
return $\{\mathbf{y}^{(k)}\}_{k=1}^K,\widehat{\mathbf{y}}$.

Table 1: **EEG encoders with a fixed generator.** Higher Avg. R (mean Pearson correlation) and lower marginal CRPS and FC-MAE of predictive means are better. Best values are **darker purple, bold**; second best values are light purple, underlined. Appendix C.2 gives results by dataset.

Encoder	Intra-Subject			Inter-Subject		
	Avg. $R\uparrow$	CRPS↓	FC-MAE↓	Avg. $R\uparrow$	CRPS↓	FC-MAE↓
BEIRA (Kovalev et al., 2022)	0.346	0.393	0.327	0.158	0.369	0.204
BIOT (Yang et al., 2023a)	0.262	0.363	0.328	0.246	0.348	0.172
BrainOmni (Xiao et al., 2025)	0.277	0.360	0.331	0.274	0.347	0.168
CNN-Trans. (Peh et al., 2022)	0.320	0.363	0.332	0.250	0.347	0.190
ContraWR (Yang et al., 2023b)	0.329	0.390	0.348	0.238	0.358	0.221
FFCL (Li et al., 2022)	0.315	0.389	0.351	0.264	0.363	0.197
LaBraM (Jiang et al., 2024)	0.265	0.360	0.319	0.304	0.345	0.187
SIREN-based (Li et al., 2024a)	0.403	0.383	0.315	0.301	0.354	0.189
NeuroBOLT (Li et al., 2024b)	0.362	**0.351**	0.331	0.306	0.345	0.183
REVE (Ouahidi et al., 2025)	0.220	0.365	0.340	0.174	0.351	0.200
SPaRCNet (Jing et al., 2023)	0.330	0.379	0.330	0.353	0.344	0.163
STT-Trans. (Song et al., 2021)	0.258	0.381	0.323	0.192	0.362	0.163
NVCode (Ours)	**0.480**	0.352	**0.280**	**0.423**	**0.344**	**0.135**



(a) Regional Pearson Correlation

(b) Transformed CRPS And FC-MAE

Figure 4: **Inter-subject EEG encoder generalization. (a):** Regional Pearson correlation across seven fMRI targets for NVCode, SPaRCNet, and STT-Trans. **(b):** One minus the mean marginal CRPS and predictive-mean FC-MAE across eight held-out-subject datasets, so higher bars indicate lower tabulated error. Mid/Cun/Tha/Put/Hes/Pre/Glo denote middle frontal gyrus, cuneus, thalamus, putamen, Heschl's gyrus, precuneus, and global signal.

4 Experiments

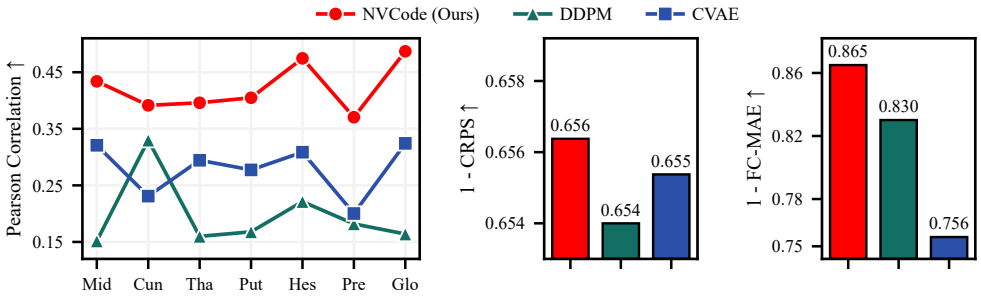
We evaluate NVCode on eight public simultaneous EEG-fMRI datasets under within-subject and held-out-subject protocols. Section 4.1 compares EEG encoders and conditional generators under matched pipelines; Section 4.2 examines temporal and spectral fidelity in representative scans; Section 4.3 tests whether generated fMRI retains identity information accessible to a classifier trained only on real data; and Section 4.4 examines the effects of the main model components. Appendix A.1 reports the eight datasets, target construction, split controls, normalization, metrics, and computing environment; the remainder of Appendix A gives baseline adaptation and evaluation details, and Appendix C provides dataset-level results.

4.1 Main Experiment

Accuracy. We first evaluate reconstruction for participants represented during training (Intra). NVCode achieves the highest average Pearson correlation, 0.480, in both matched comparisons. It exceeds the strongest encoder, SIREN-based (0.403), by 19.1% with a fixed generator (Table 1).

Table 2: **Conditional generators with a fixed EEG encoder.** Generators receive identical stored EEG conditions and use the same data splits, normalization, and evaluation metrics. Results average eight datasets equally; Appendix C.3 provides the complete results for individual datasets.

Generator	Intra-Subject			Inter-Subject		
	Avg. $R\uparrow$	CRPS↓	FC-MAE↓	Avg. $R\uparrow$	CRPS↓	FC-MAE↓
cGAN (Mirza & Osindero, 2014)	0.310	0.462	0.505	0.211	0.368	0.297
Consistency (Yang et al., 2024)	0.328	0.452	0.357	0.283	0.431	0.240
CVAE (Sohn et al., 2015)	0.404	0.378	0.336	0.280	0.345	0.244
DDPM (Ho et al., 2020)	0.302	0.358	0.309	0.197	0.346	0.170
EDM (Karras et al., 2022)	0.351	0.361	0.323	0.286	0.343	0.211
Rectified Flow (Liu et al., 2022)	0.302	0.364	0.328	0.234	0.350	0.242
Score SDE (Song et al., 2020)	0.343	0.361	0.346	0.280	**0.339**	0.202
NVCode (Ours)	**0.480**	**0.352**	**0.280**	**0.423**	0.344	**0.135**



(a) Regional Pearson Correlation

(b) Transformed CRPS And FC-MAE

Figure 5: **Inter-subject conditional generator generalization. (a):** Regional Pearson correlation across seven fMRI targets for NVCode, DDPM, and CVAE under the same EEG condition. **(b):** One minus the mean marginal CRPS and predictive-mean FC-MAE across eight held-out-subject datasets, so higher bars indicate lower tabulated error. Region abbreviations follow Figure 4.

With fixed EEG conditions, it exceeds the strongest generator, CVAE (0.404), by 18.8% (Table 2). The gains therefore hold in comparisons of both EEG representations and conditional generators.

NVCode retains its correlation advantage when evaluated on unseen participants. Under Inter evaluation, it reaches $R = 0.423$ across eight equally weighted datasets. This exceeds SPaRCNet (0.353), the strongest encoder with the same generator, by 19.8% (Table 1). With identical stored EEG conditions, the gain over EDM (0.286), the strongest generator, is 47.9% (Table 2). Figures 4(a) and 5(a) show the regional pattern, including gains at Heschl's gyrus and the global signal. NVCode exceeds DDPM and CVAE at every target. The larger generator gap under Inter than Intra evaluation highlights the value of the conditional mapping for transfer across previously unseen participants.

NVCode also improves the accuracy of connectivity estimated from predictive means. CRPS scores each target's marginal distribution; FC-MAE compares correlation matrices of predictive means and real fMRI. Under Inter evaluation, NVCode achieves the lowest FC-MAE, 0.135, in both comparisons (Tables 1 and 2). FC-MAE is 17.2% and 20.6% lower than the best encoder and generator alternatives (0.163 and 0.170). Its CRPS of 0.344 matches the best encoder at reported precision and improves on CVAE (0.345) and DDPM (0.346). Figures 4(b) and 5(b) visualize these differences. Together, these metrics support accurate regional signals with coordinated variation across their predictive means.

4.2 Analysis

Temporal fidelity. On the same EEG input, NVCode follows the Cuneus waveform more closely than EDM (Figure 6). Spectrograms, alpha scalp power and real targets are identical. NVCode recovers local reversals and the late positive excursion, yielding $R = 0.782$ versus 0.501 for EDM. At 54.6 s, the real value is 0.703; NVCode predicts 0.593 and EDM 0.224, giving absolute differences

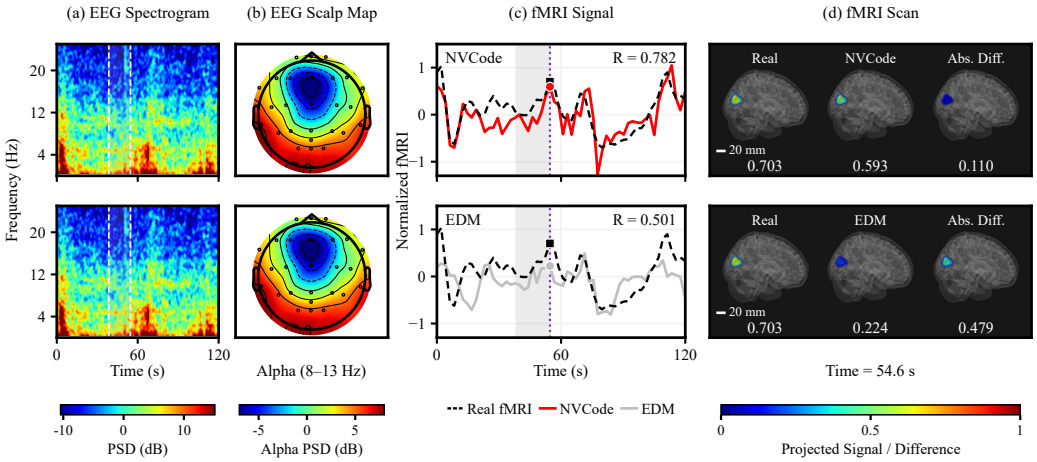


Figure 6: **Matched EEG–fMRI evidence for Cuneus. (a, b):** Identical EEG inputs in both rows. **(c):** Real fMRI (black dashed), NVCode (top, red), and EDM (bottom, grey), with correlations. **(d):** Real, predicted, and absolute-error amplitudes projected onto Cuneus at 54.6 s (purple guides) with a shared color scale. Additional scans and anatomical projection details appear in Appendix C.4.

of 0.110 and 0.479. Smaller differences indicate closer agreement with real fMRI: NVCode reduces this snapshot error by 77.0%. Panel (d) uses identical anatomical weights and a common color scale to compare the projected amplitudes. The smaller error and stronger waveform agreement link local signal dynamics to the amplitude shown in the brain projection of this same Cuneus scan.

Spectral fidelity. NVCode also recovers the dominant low frequency structure of the selected global signal (Figure 7). Resting BOLD activity concentrates power at low frequencies (Fox & Raichle, 2007; Cordes et al., 2001). We compare spectra of predictive means on the scan selected for closest alignment with real fMRI. NVCode follows the rise into the dominant band, approaches the real peak and tracks the subsequent decay. EDM peaks at a higher frequency, shifting the dominant variation to a shorter timescale. On the logarithmic axes, peak location identifies the principal timescale, while the descending branch shows attenuation of faster fluctuations. Agreement in both features supports recovery of slow variation in the preprocessed target, complementing the local signal dynamics examined in the waveform comparison.

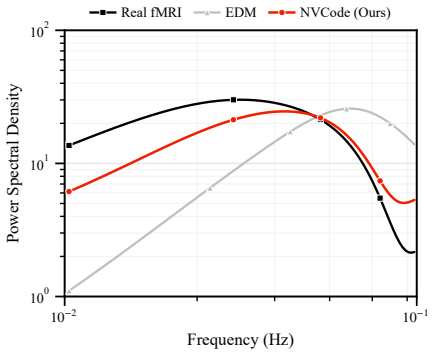


Figure 7: **Spectral fidelity.** Log-scale PSD on the best-aligned global signal.

4.3 Downstream Utility

We next test whether generated fMRI retains identity information usable by a classifier trained only on real data. For each dataset, we fit and validate an identity classifier on real fMRI segments and then freeze its parameters. At test time, we replace real inputs with predictions from matched EEG epochs. Labels, evaluation segments and classifier parameters remain unchanged across generators. The accuracy change therefore measures how well the generated inputs support the same identity decision rule. Table 3 shows that NVCode achieves the highest mean gain of 3.1 percentage points across eight datasets, compared with 2.4 for DDPM. This probe complements signal reconstruction metrics by evaluating identity patterns accessible to the same classifier trained on real fMRI.

The downstream gain is consistent across both resting and task recordings. NVCode improves accuracy on all eight datasets, ranking first on seven and second on XP1. Its mean gain exceeds the strongest alternative, DDPM, by 0.7 percentage points. NatView (4.5 percentage points), NeuroBOLT (4.3) and gradCPT (3.6) show the largest gains, covering both recording contexts. This pattern indicates that the generated inputs retain identity features usable by classifiers learned from real fMRI. The probe asks a different question from correlation and FC-MAE, which evaluate temporal

Table 3: **Synthetic fMRI transfers to a classifier trained only on real fMRI.** The first row gives real input accuracy; synthetic rows report the change Δ using predictions from matched EEG epochs. Classifier parameters stay fixed, and Mean averages the eight datasets equally. Appendix A.3 gives the protocol and score definition; Appendix A.4 analyzes regional representations.

Method	Dataset accuracy / Δ								Mean
	Sleep EEG–fMRI	MULTI-CLARID	gradCPT	NeuroBOLT	XP1	Affective Music	NatView	Experience Sampling	
Real test	0.861	0.787	0.746	0.869	0.874	0.782	0.747	0.702	0.796
cGAN (Mirza & Osindero, 2014)	+0.022	+0.018	+0.013	+0.014	+0.008	+0.019	+0.027	+0.002	+0.015
Consistency (Yang et al., 2024)	+0.017	+0.024	+0.026	+0.019	+0.010	+0.011	+0.021	+0.014	+0.018
CVAE (Sohn et al., 2015)	+0.011	+0.016	+0.034	+0.037	+0.017	+0.005	+0.033	+0.012	+0.021
DDPM (Ho et al., 2020)	+0.026	+0.019	+0.020	+0.035	**+0.024**	+0.023	+0.032	+0.014	+0.024
EDM (Karras et al., 2022)	+0.007	+0.021	+0.026	+0.022	+0.009	+0.025	+0.038	+0.008	+0.020
Rectified Flow (Liu et al., 2022)	+0.015	+0.008	+0.018	+0.032	+0.012	+0.010	+0.041	+0.011	+0.018
Score SDE (Song et al., 2020)	+0.021	+0.016	+0.027	+0.029	+0.015	+0.006	+0.026	+0.015	+0.019
NVCode (Ours)	**+0.028**	**+0.027**	**+0.036**	**+0.043**	+0.019	**+0.031**	**+0.045**	**+0.020**	**+0.031**

Table 4: **Component ablation under Inter evaluation.** Scores average the eight datasets equally under shared splits and evaluation rules. Output module capacity is matched where applicable; the shared encoder changes target supervision and encoder capacity together. Appendix A.4.2 details the variants, training and scoring; Figure 8 visualizes the three criteria across all four configurations.

Variant	Model components			Inter-subject evaluation		
	Regional Encoders	Joint Output	Stochastic Interpolant	Avg. $R\uparrow$	CRPS $\downarrow$	FC-MAE $\downarrow$
Deterministic Joint Regression	✓	✓	×	0.317	0.368	0.250
Independent Interpolants	✓	×	✓	0.405	0.377	0.174
NVCode w/ Shared Encoder	×	✓	✓	0.349	0.362	0.182
NVCode (Ours)	✓	✓	✓	**0.423**	**0.344**	**0.135**

agreement and regional organization. Together, these evaluations connect accurate reconstruction with identity information usable by an existing classifier across diverse acquisition contexts.

4.4 Ablation Study

We examine stochastic interpolation, joint state modeling and regional representations through three targeted comparisons (Table 4). All variants share eight datasets, subject disjoint splits and evaluation rules, with output capacity matched where applicable. NVCode leads all three criteria: 0.423 Avg. R, 0.344 CRPS and 0.135 FC-MAE. Relative to deterministic joint regression, stochastic interpolation raises Avg. R by 0.106 and reduces FC-MAE by 0.115 under the same regional representations. Relative to independent interpolants, joint generation reduces CRPS by 0.033 and FC-MAE by 0.039. This comparison supports giving each regional velocity access to the complete evolving fMRI state and shared EEG condition. The full encoder configuration raises correlation by 0.074 over the tested shared encoder. Together, the comparisons favor specialized EEG representations with joint stochastic generation for predictive mean accuracy, marginal prediction quality and regional connectivity.

5 Conclusion

We presented Neurovascular Codes for EEG-to-fMRI synthesis. NVCode learns target-specific EEG features through individual fMRI prediction tasks and projects them into a shared code for stochastic fMRI generation. Across eight public datasets, it achieves the highest average Pearson correlation and the lowest predictive-mean FC-MAE in both encoder and generator comparisons, under within-subject and held-out-subject evaluation. Its generated signals also achieve the highest mean subject-identification accuracy among the evaluated generators when evaluated by fixed classifiers trained on real fMRI.

Discussion. More broadly, our results suggest that predicting one brain modality can provide useful supervision for learning representations of another. In this work, the representations learned from individual prediction tasks support generation of multiple fMRI signals from the same EEG window. Whether these representations can transfer across datasets or support generation in both directions remains an open question. We hope this work opens up new opportunities for brain models that learn representations and generate signals across modalities.

6 AI USE STATEMENT

In accordance with ICLR policy, we disclose our use of generative artificial intelligence (AI) tools in manuscript preparation. These tools assisted with proofreading, grammatical correction, language refinement, LaTeX formatting, and presentation-oriented editing of figure layouts and visualization code based on author-provided data and decisions. They did not originate the research question, define the methodological contribution, execute the reported experiments, or generate the underlying data and numerical results. All AI-assisted text, code, figures, citations, and formatting were reviewed and verified by the authors. The authors take full responsibility for the final content of this work.

REFERENCES

Rodolfo Abreu, Alberto Leal, and Patrícia Figueiredo. Eeg-informed fmri: a review of data analysis methods. *Frontiers in human neuroscience*, 12:29, 2018.

Michael Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.

Philip J Allen, Oliver Josephs, and Robert Turner. A method for removing imaging artifact from continuous eeg recorded during functional mri. *Neuroimage*, 12(2):230–239, 2000.

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

Hubert Banville, Omar Chehab, Aapo Hyvärinen, Denis-Alexander Engemann, and Alexandre Gramfort. Uncovering the structure of clinical eeg signals with self-supervised learning. *Journal of Neural Engineering*, 18(4):046020, 2021.

Danielle S Bassett and Olaf Sporns. Network neuroscience. *Nature neuroscience*, 20(3):353–364, 2017.

Bharat Biswal, F Zerrin Yetkin, Victor M Haughton, and James S Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.

Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2021.

Madeleine Bullock, Graeme D Jackson, and David F Abbott. Artifact reduction in simultaneous eeg-fmri: a systematic review of methods and contemporary usage. *Frontiers in neurology*, 12: 622719, 2021.

György Buzsáki and Brendon O Watson. Brain rhythms and neural syntax: implications for efficient coding of cognitive content and neuropsychiatric disease. *Dialogues in clinical neuroscience*, 14 (4):345–367, 2012.

Vince D Calhoun and Jing Sui. Multimodal fusion of brain imaging data: a key to finding the missing link (s) in complex mental illness. *Biological psychiatry: cognitive neuroscience and neuroimaging*, 1(3):230–244, 2016.

Josue Ortega Caro, Antonio H de O Fonseca, Christopher Averill, Syed A Rizvi, Matteo Rosati, James L Cross, Prateek Mittal, Emanuele Zappala, Daniel Levine, Rahul M Dhodapkar, et al. Brainlm: A foundation model for brain activity recordings. *BioRxiv*, pp. 2023–09, 2023.

Younghwa Cha, Yeji Lee, Eunhee Ji, SoHyun Han, Sunhyun Min, Hyoungkyu Kim, Minseo Cho, Hae Seong Lee, Youngjai Park, and Joon-Young Moon. "sustained attention task (gradcpt) dataset using simultaneous eeg-fmri and dti", 2026.

Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

Michael X Cohen. Where does eeg come from and what does it mean? *Trends in neurosciences*, 40 (4):208–218, 2017.

Dietmar Cordes, Victor M Haughton, Konstantinos Arfanakis, John D Carew, Patrick A Turski, Chad H Moritz, Michelle A Quigley, and M Elizabeth Meyerand. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *American journal of neuroradiology*, 22(7):1326–1333, 2001.

Wenhui Cui, Woojae Jeong, Philipp Thölke, Takfarinas Medani, Karim Jerbi, Anand A Joshi, and Richard M Leahy. Neuro-gpt: Towards a foundation model for eeg. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–5. IEEE, 2024.

Kamalaker Dadi, Gaël Varoquaux, Antonia Machlouzarides-Shalit, Krzysztof J Gorgolewski, Demian Wassermann, Bertrand Thirion, and Arthur Mensch. Fine-grain atlases of functional modes for fmri analysis. *NeuroImage*, 221:117126, 2020.

Ian Daly, Nicoletta Nicolaou, Duncan Williams, Faustina Hwang, Alexis Kirke, Eduardo Miranda, and Slawomir J. Nasuto. "a dataset recording joint eeg-fmri during affective music listening ", 2020.

Salman UH Dar, Mahmut Yurt, Levent Karacan, Aykut Erdem, Erkut Erdem, and Tolga Cukur. Image synthesis in multi-contrast mri with conditional generative adversarial networks. *IEEE transactions on medical imaging*, 38(10):2375–2388, 2019.

Fani Deligianni, Maria Centeno, David W Carmichael, and Jonathan D Clayden. Relating resting-state fmri and eeg whole-brain connectomes across frequency bands. *Frontiers in neuroscience*, 8: 258, 2014.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, 8(9):700–711, 2007.

Karl J Friston. Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36, 2011.

Matthew F Glasser, Timothy S Coalson, Emma C Robinson, Carl D Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F Beckmann, Mark Jenkinson, et al. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, 2016.

Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.

Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

Will Grathwohl, Ricky TQ Chen, Jesse Bettencourt, Ilya Sutskever, and David Duvenaud. Ffjord: Free-form continuous dynamics for scalable reversible generative models. *arXiv preprint arXiv:1810.01367*, 2018.

Yameng Gu, Feng Han, Lucas E. Sainburg, Margeaux M. Schade, and Xiao Liu. "simultaneous eeg and fmri signals during sleep from humans", 2026.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

René J Huster, Stefan Debener, Tom Eichele, and Christoph S Herrmann. Methods for simultaneous eeg-fmri: an introductory review. *Journal of Neuroscience*, 32(18):6053–6060, 2012.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.

Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous eeg data in bci. *arXiv preprint arXiv:2405.18765*, 2024.

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

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 35:26565–26577, 2022.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Demetres Kostas, Stephane Aroca-Ouellette, and Frank Rudzicz. Bendr: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of eeg data. *Frontiers in Human Neuroscience*, 15:653659, 2021.

Alexander Kovalev, Ilia Mikheev, and Alexei Ossadtchi. fmri from eeg is only deep learning away: the use of interpretable dl to unravel eeg-fmri relationships. *arXiv preprint arXiv:2211.02024*, 2022.

Aaron Kucyi, Lotus Shareef-Trudeau, David Braun, Huiling Peng, Tiara Bounyarith, and Janet Z. Li. "a multi-session simultaneous eeg-fmri dataset with online experience sampling", 2026.

Helmut Laufs, Andreas Kleinschmidt, Astrid Beyerle, Evelyn Eger, Afraim Salek-Haddadi, Christine Preibisch, and Karsten Krakow. Eeg-correlated fmri of human alpha activity. *Neuroimage*, 19(4): 1463–1476, 2003.

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

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

Yamin Li, Ange Lou, Ziyuan Xu, Shiyu Wang, and Catie Chang. Leveraging sinusoidal representation networks to predict fmri signals from eeg. In *Medical Imaging 2024: Image Processing*, volume 12926, pp. 795–800. SPIE, 2024a.

Yamin Li, Ange Lou, Ziyuan Xu, Shengchao Zhang, Shiyu Wang, Dario J Englot, Soheil Kolouri, Daniel Moyer, Roza G Bayrak, and Catie Chang. Neurobolt: Resting-state eeg-to-fmri synthesis with multi-dimensional feature mapping. *Advances in neural information processing systems*, 37: 23378–23405, 2024b.

Giulia Lioi, Claire Cury, Lorraine Perronnet, Marsel Mano, Elise Bannier, Anatole Lecuyer, and Christian Barillot. "a multi-modal human neuroimaging dataset for data integration: simultaneous eeg and fmri acquisition during a motor imagery neurofeedback task: Xp1", 2021.

Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

Nikos K Logothetis. What we can do and what we cannot do with fmri. *Nature*, 453(7197):869–878, 2008.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

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):031005, 2018.

James E Matheson and Robert L Winkler. Scoring rules for continuous probability distributions. *Management science*, 22(10):1087–1096, 1976.

Dennis J McFarland and Jonathan R Wolpaw. Brain-computer interfaces for communication and control. *Communications of the ACM*, 54(5):60–66, 2011.

Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

Kevin Murphy, Rasmus M Birn, Daniel A Handwerker, Tyler B Jones, and Peter A Bandettini. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage*, 44(3):893–905, 2009.

Teresa Murta, Marco Leite, David W Carmichael, Patrícia Figueiredo, and Louis Lemieux. Electrophysiological correlates of the bold signal for eeg-informed fmri. *Human brain mapping*, 36(1): 391–414, 2015.

Rami K Niazy, Christian F Beckmann, Gian Domenico Iannetti, J Michael Brady, and Stephen M Smith. Removal of fmri environment artifacts from eeg data using optimal basis sets. *Neuroimage*, 28(3):720–737, 2005.

Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv preprint arXiv:2211.14730*, 2022.

Seiji Ogawa, Tso-Ming Lee, Alan R Kay, and David W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *proceedings of the National Academy of Sciences*, 87(24):9868–9872, 1990.

Yassine El Ouahidi, Jonathan Lys, Philipp Thölke, Nicolas Farrugia, Bastien Pasdeloup, Vincent Gripon, Karim Jerbi, and Giulia Lioi. Reve: A foundation model for eeg–adapting to any setup with large-scale pretraining on 25,000 subjects. *arXiv preprint arXiv:2510.21585*, 2025.

Muzaffer Özbey, Onat Dalmaz, Salman UH Dar, Hasan A Bedel, Şaban Özturk, Alper Güngör, and Tolga Cukur. Unsupervised medical image translation with adversarial diffusion models. *IEEE Transactions on Medical Imaging*, 42(12):3524–3539, 2023.

Karl Pearson. Vii. note on regression and inheritance in the case of two parents. *proceedings of the royal society of London*, 58(347-352):240–242, 1895.

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

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

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.

Chitwan Saharia, William Chan, Huiwen Chang, Chris Lee, Jonathan Ho, Tim Salimans, David Fleet, and Mohammad Norouzi. Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 conference proceedings*, pp. 1–10, 2022.

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.

Paul Scotti, Atmadeep Banerjee, Jimmie Goode, Stepan Shabalin, Alex Nguyen, Aidan Dempster, Nathalie Verlinde, Elad Yundler, David Weisberg, Kenneth Norman, et al. Reconstructing the mind's eye: fmri-to-image with contrastive learning and diffusion priors. *Advances in Neural Information Processing Systems*, 36:24705–24728, 2023.

Stephen M Smith, Christian F Beckmann, Jesper Andersson, Edward J Auerbach, Janine Bijsterbosch, Gwenaëlle Douaud, Eugene Duff, David A Feinberg, Ludovica Griffanti, Michael P Harms, et al. Resting-state fmri in the human connectome project. *Neuroimage*, 80:144–168, 2013.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. pmlr, 2015.

Kihyuk Sohn, Honglak Lee, and Xinchen Yan. Learning structured output representation using deep conditional generative models. *Advances in neural information processing systems*, 28, 2015.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in neural information processing systems*, 32, 2019.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *Proceedings of the 40th International Conference on Machine Learning*, pp. 32211–32252, 2023.

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

Jörg Stadler, Torsten Stöter, Nicole Angenstein, Andreas Fügner, Marcel Lommerzheim, Artur Mathysiak, Anke Michalsky, Gabriele Schöps, Johann van der Meer, Susann Wolff, and André Brechmann. "multi-clarid (multimodal category learning and resting-state imaging data)", 2025.

Akara Supratak, Hao Dong, Chao Wu, and Yike Guo. Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg. *IEEE transactions on neural systems and rehabilitation engineering*, 25(11):1998–2008, 2017.

Yu Takagi and Shinji Nishimoto. High-resolution image reconstruction with latent diffusion models from human brain activity. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 14453–14463, 2023.

Qawi K Telesford, Eduardo Gonzalez-Moreira, Ting Xu, Yiwen Tian, Stanley J Colcombe, Jessica Cloud, Brian E Russ, Arnaud Falchier, Maximilian Nentwich, Jens Madsen, et al. An open-access dataset of naturalistic viewing using simultaneous eeg-fmri. *Scientific Data*, 10(1):554, 2023.

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. *arXiv preprint arXiv:2302.00482*, 2023.

Martijn P Van Den Heuvel and Hilleke E Hulshoff Pol. Exploring the brain network: a review on resting-state fmri functional connectivity. *European neuropsychopharmacology*, 20(8):519–534, 2010.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Diego Vidaurre, Romesh Abeysuriya, Robert Becker, Andrew J Quinn, Fidel Alfaro-Almagro, Stephen M Smith, and Mark W Woolrich. Discovering dynamic brain networks from big data in rest and task. *NeuroImage*, 180:646–656, 2018.

Guagnyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. Eegpt: Pretrained transformer for universal and reliable representation of eeg signals. *Advances in Neural Information Processing Systems*, 37:39249–39280, 2024.

Qinfan Xiao, Ziyun Cui, Chi Zhang, Siqi Chen, Wen Wu, Andrew Thwaites, Alexandra Woolgar, Bowen Zhou, and Chao Zhang. Brainomni: A brain foundation model for unified eeg and meg signals. *arXiv preprint arXiv:2505.18185*, 2025.

Chaoqi Yang, M Westover, and Jimeng Sun. Biot: Biosignal transformer for cross-data learning in the wild. *Advances in Neural Information Processing Systems*, 36:78240–78260, 2023a.

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

Ling Yang, Zixiang Zhang, Zhilong Zhang, Xingchao Liu, Minkai Xu, Wentao Zhang, Chenlin Meng, Stefano Ermon, and Bin Cui. Consistency flow matching: Defining straight flows with velocity consistency. *arXiv preprint arXiv:2407.02398*, 2024.

BT Thomas Yeo, Fenna M Krienen, Jorge Sepulcre, Mert R Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L Roffman, Jordan W Smoller, Lilla Zöllei, Jonathan R Polimeni, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 2011.

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

APPENDICES

CONTENTS

This appendix consolidates the experimental protocols, method specifications, and extended evidence that support the main paper.

- A covers experimental setup, evaluation, downstream utility, and ablations.
- B details theory, model architecture, optimization, and generation.
- C presents aggregate, per dataset, and paired EEG–fMRI results.

A Experiment Details **17**

A.1 Experiment Setup: Data Preprocessing and Machine Requirements 17

A.2 Main Experiment Details 19

A.3 Downstream Utility Details 22

A.4 Ablation Study Details 22

B Method Details **25**

B.1 Velocity Matching Identity 25

B.2 Regional EEG Evidence Encoders 27

B.3 Shared Conditioning Module 28

B.4 Field Parameterization 29

B.5 Optimization Details 30

B.6 Generation Details 30

C Additional Results **32**

C.1 Main EEG to fMRI Synthesis Results 32

C.2 Per Dataset Encoder Results 33

C.3 Per Dataset Generator Results 36

C.4 Scan Level EEG–fMRI Evidence 37

A Experiment Details

This appendix specifies the data construction, baseline adaptation, metric formulas, downstream utility probe and ablation protocol used in Section 4.

A.1 Experiment Setup: Data Preprocessing and Machine Requirements

All experiments use identical EEG to fMRI pairs, identical target signal definitions and identical split dependent normalization.

Datasets and splits. Table 5 summarizes eight public simultaneous EEG–fMRI datasets (Huster et al., 2012; Murta et al., 2015), comprising four resting-state (Biswal et al., 1995; Fox & Raichle, 2007) and four task-evoked (Laufs et al., 2003; Murta et al., 2015) settings. Their source cohorts report 184 participants, while the 189 scans retained after scan-level exclusions yield 61,646 paired volumes. Each fMRI volume is paired with the immediately preceding 16-s EEG window, so future EEG does not enter the model input. Intra-subject evaluation uses 80% of volumes for training, 10% for validation, and the remainder for testing. A $\lceil 20/\mathrm{TR} \rceil$-volume buffer separates validation and test portions to prevent leakage from adjacent volumes and overlapping EEG contexts (Logothetis, 2008). Inter-subject evaluation (Lotte et al., 2018) uses disjoint training and held-out subjects. Held-out participants contribute neither normalization statistics nor checkpoint selection, and all splits are fixed and deterministic.

Table 5: **Eight simultaneous EEG–fMRI datasets span complementary resting-state and task-evoked settings.** The source cohorts report 184 participants; after scan-level exclusions, 189 retained scans yield 61,646 paired volumes. Labels follow official titles and citations retain repository accessions.

Type	Dataset	EEG		fMRI		Paired Acquisition			Paradigm
		Rate (Hz)	Channels	TR (s)	Pairs	Duration (s)	Scans	Cohort Size	
Rest	Sleep EEG–fMRI (Gu et al., 2026)	5000	32	2.1	5,714	600	20	33	Sleep stage
	MULTI-CLARID (Stadler et al., 2025)	5000	64	2.0	10,200	600	34	24	Eyes open rest
	gradCPT (Cha et al., 2026)	5000	64	2.0	5,040	360	28	28	Rest subset
	NeuroBOLT (Li et al., 2024b)	5000	26	2.1	16,571	1200	29	22	Multi session rest
Active	XP1 (Lioi et al., 2021)	5000	64	2.0	2,400	480	10	10	Motor imagery feedback
	Affective Music (Daly et al., 2020)	5000	31	2.0	6,300	600	21	21	Classical music listening
	Experience Sampling (Kucyi et al., 2026)	5000	64	2.0	14,850	660	45	24	Sustained attention task
	NatView (Telesford et al., 2023)	5000	61	2.1	571	600	2	22	Naturalistic film viewing

Target signals, baselines and metrics. Each volume defines a seven-dimensional target vector containing six DiFuMo ROIs and one global signal (Dadi et al., 2020; Murphy et al., 2009). The ROIs cover primary sensory targets (cuneus and Heschl's gyrus), higher-order cognitive targets (middle frontal gyrus and precuneus), and subcortical targets (putamen and thalamus). Every model predicts this same ordered vector from the same retained EEG–fMRI pairs and training-only normalization statistics. In Table 1, the generator architecture, sampler, target ordering, and split policy are held fixed while the EEG encoder configuration changes. In Table 2, the seven frozen NVCode encoders and fitted projector supply identical conditions while the downstream generation network changes. Stochastic generators return $K{=}50$ samples for each held-out EEG epoch. Avg. R and FC-MAE evaluate the predictive mean, whereas mean marginal CRPS uses the complete empirical sample distribution. Dataset scores are first summarized over their internal splits and then macro-averaged with equal dataset weight. This aggregation prevents datasets with more retained volumes from dominating the headline comparison. Appendix A.2 gives the adaptation rules, metric definitions, and aggregation procedure.

Cohort and scan-count conventions. Table 5 distinguishes three levels of data accounting. *Cohort Size* gives the participant count documented for each source cohort. *Scans* counts runs retained after scan-level inclusion, and *Pairs* counts the volume-anchored EEG–fMRI observations constructed from those runs, totaling 61,646. Scan inclusion precedes pair construction and model fitting. These columns describe source participants, retained recordings, and paired observations, respectively; subject-level train/test grouping follows the inter-subject protocol below.

Paired EEG to fMRI observations. Each paired observation is anchored at one scanner volume onset trigger. For a volume acquired at time t_v, we extract the EEG segment in the half open interval $[t_v - 16\,\mathrm{s}, t_v)$ and pair it with the fMRI signal vector of that volume. This convention uses only

preceding EEG and does not expose future EEG. The repetition time is dataset dependent: $\mathrm{TR} = 2.0\,\mathrm{s}$ for XP1, Affective Music, MULTI-CLARID, gradCPT, and Experience Sampling, and $\mathrm{TR} = 2.1\,\mathrm{s}$ for Sleep EEG–fMRI, NeuroBOLT, and NatView. We do not apply an explicit hemodynamic response function lag correction. The 16 s EEG context is fixed across datasets and covers the time scale over which EEG evidence can condition the corresponding BOLD response (Logothetis, 2008).

EEG preprocessing and channel harmonization. EEG-informed fMRI analysis combines electrophysiological preprocessing with BOLD modeling (Abreu et al., 2018) within broader multimodal brain imaging workflows (Calhoun & Sui, 2016). All EEG segments are resampled to 200 Hz with the polyphase finite impulse response resampler in `mne.io.Raw.resample(..., method="polyphase")`. The default anti aliasing low pass filter is applied at the Nyquist frequency and no time shift is introduced. Each segment is cropped to $T = 3200$ samples, and raw amplitudes are divided by 100 before being fed to the encoders. Native channel counts vary across datasets, so all recordings are harmonized to a fixed 26 channel 10–20 montage:

$$\begin{aligned}\mathcal{C}_{26} = \{&\mathrm{FP1}, \mathrm{FP2}, \mathrm{F3}, \mathrm{F4}, \mathrm{C3}, \mathrm{C4}, \mathrm{P3}, \mathrm{P4}, \mathrm{O1}, \mathrm{O2},\\ &\mathrm{F7}, \mathrm{F8}, \mathrm{T7}, \mathrm{T8}, \mathrm{P7}, \mathrm{P8}, \mathrm{FPZ}, \mathrm{FZ}, \mathrm{CZ}, \mathrm{PZ},\\ &\mathrm{POZ}, \mathrm{OZ}, \mathrm{FT9}, \mathrm{FT10}, \mathrm{TP9}, \mathrm{TP10}\}.\end{aligned}$$

Non EEG channels such as ECG, EOG and EMG are removed, along with carbon wire loop sensors. Missing target channels are reconstructed by spherical spline interpolation against the standard 10–20 montage using MNE Python with `standard_1020`. Channels present in the source recording but outside the target set are discarded. Dataset provided referencing and artifact correction decisions are preserved. No dataset specific EEG augmentation is applied.

fMRI signal construction. For each 4D fMRI scan, we extract atlas level signals with nilearn's `NiftiMapsMasker` using the 64 component DiFuMo atlas (Dadi et al., 2020) at 2 mm resolution. The masker is used only for soft mask spatial weighting; masker standardization and detrending are disabled so that all reported statistics are controlled by the split dependent preprocessing below. From the 64 DiFuMo components, we select six regional signals by label matching: Cuneus, Heschl's gyrus, Middle Frontal Gyrus anterior, Precuneus anterior, Putamen and Thalamus. Under `nilearn` 0.10.3, these correspond to component IDs $\{16, 52, 46, 3, 28, 30\}$ in the listed order. The seventh signal is the global signal (Murphy et al., 2009), computed as the unweighted mean over all 64 DiFuMo components before filtering or scaling. The seven signals are stacked into $\mathbf{r}_{s,v} \in \mathbb{R}^7$ for scan s and volume v.

Temporal filtering. We apply a fifth order Butterworth low pass filter at $f_c = 0.15\,\mathrm{Hz}$ to the seven fMRI signals, retaining the low frequency BOLD range emphasized in resting state fMRI analysis (Cordes et al., 2001; Fox & Raichle, 2007):

```
scipy.signal.butter(5, 0.15, btype='low', fs=1/TR, output='sos')
```

followed by `scipy.signal.sosfiltfilt`. In intra subject experiments, filtering is applied separately within the training, validation and test portions before the buffer is dropped. In inter subject experiments, filtering is applied per scan before subject level evaluation. This filtering is part of signal construction and is shared by all methods.

Normalization and split control. We use two normalizations. The first defines the robust-normalized fMRI label scale used for reported metrics. In the intra subject protocol, a robust scaler is fitted for each scan s and signal n using only the corresponding temporal training portion:

$$m_{s,n} = \mathrm{median}_{v \in \mathcal{I}^s_{\mathrm{train}}} r_{s,v,n}, \qquad a_{s,n} = \frac{\mathrm{IQR}_{v \in \mathcal{I}^s_{\mathrm{train}}} r_{s,v,n}}{1.349} + \epsilon, \qquad \epsilon = 10^{-6}.$$

In the inter subject protocol, the same definitions are evaluated over all volumes from training subjects in dataset d, yielding $m_{d,n}$ and $a_{d,n}$. These parameters remain fixed and are applied unchanged to every held out subject in dataset d. Let $\rho = s$ for intra subject evaluation and $\rho = d(s)$ for inter subject evaluation, where $d(s)$ denotes the dataset containing scan s. The robust-normalized label is

$$y_{s,v,n} = \frac{r_{s,v,n} - m_{\rho,n}}{a_{\rho,n}}.$$

This fixed transformation defines the fMRI supervision and evaluation scale. The second normalization is the split specific generator z score used inside the stochastic interpolant. In the intra subject

protocol, the generator mean and standard deviation are fitted on the training rows of the scan. In the inter subject protocol, they are fitted only on training subjects. Generated samples are mapped back to the robust-normalized fMRI scale before metrics are computed.

Intra subject split. For a scan with V fMRI volumes, we form four contiguous temporal portions in order: $\lfloor 0.8V \rfloor$ training volumes, $\lfloor 0.1V \rfloor$ validation volumes, a $\lceil 20/\mathrm{TR} \rceil$ volume buffer and all remaining test volumes. The buffer is removed from evaluation. It prevents leakage from adjacent fMRI volumes and from overlapping 16 s EEG windows. All trainable components are fitted on the training portion, selected on the validation portion and evaluated only on the held out test portion.

Inter subject split. Inter-subject evaluation groups recordings by subject and tests transfer to held-out people within each dataset. Both normalization layers, model fitting, checkpoint selection, and early stopping use the training-subject partition. Regional representation maps and feature caches are constructed within each training fold; evaluation applies these fixed maps to held-out EEG. The benchmark therefore measures subject transfer across eight simultaneous EEG–fMRI acquisition settings.

Reproducibility controls. The reconstruction comparisons use three runs with random seeds 42, 2026, and 114514, keeping the data splits fixed. For each seed r, we set `torch.manual_seed(r)` and `numpy.random.seed(r)` at process start. All methods share deterministic settings: we enable `torch.use_deterministic_algorithms(True)`, disable cuDNN benchmarking and disable TF32. Training uses fp32 without mixed precision or data augmentation.

Machine requirements and software. All experiments were implemented in PyTorch and trained in parallel on four NVIDIA A100 80 GB or H100 80 GB GPUs. At inference time, the EEG derived context is computed once for each held out epoch and then shared across all $S \cdot K$ velocity evaluations. With $S = 100$ Euler steps and $K = 50$ Monte Carlo samples, inference for a typical scan with 50 to 200 held out epochs uses one accelerator. The full eight dataset inter subject evaluation requires independent generator fits across subject splits and approximately 20 aggregate GPU days after caching per scan EEG features. The software environment uses PyTorch 2.1, SciPy 1.11, NumPy 1.26, MNE-Python $\geq$1.7, nilearn 0.10.3, `linear attention transformer` 0.19.1, `umap-learn` 0.5 (McInnes et al., 2018), CUDA 12.1 and cuDNN 8.9.

A.2 Main Experiment Details

Tables 1 and 2 compare EEG encoder configurations and conditional generation configurations, respectively. The EEG encoder outputs the feature vector $\mathbf{e}(\mathbf{x})$. The complete generator includes both the conditional projector Π_ψ and the generation network F_θ, with sampling randomness denoted by ϵ:

$$G_{\psi,\theta}(\mathbf{e}(\mathbf{x}), \epsilon) = F_\theta(\Pi_\psi(\mathbf{e}(\mathbf{x})), \epsilon)\,.$$

The protocols below specify which generator parameters are trained in each comparison. Table 4 separately controls stochastic interpolation, joint state access, and target-specific encoding.

EEG encoder baselines. Related EEG representation work uses convolutional feature extraction (Lawhern et al., 2018; Roy et al., 2019), recurrent sleep staging (Supratak et al., 2017), and Transformer-based sleep staging (Phan et al., 2022). All encoder baselines are adapted to the same sequential EEG to fMRI prediction task. Each baseline family uses one EEG encoder with a 1400-dimensional output representation. A temporary fMRI readout supervises this encoder with mean squared error on the seven robust-normalized target signals. Validation MSE selects the checkpoint, after which the temporary readout is discarded and the encoder is frozen. NVCode instead fits seven target-specific encoders, each with a temporary scalar head, and concatenates their 200-dimensional representations into the same 1400-dimensional interface. Each row fits a separate NVCode generator with the same architecture, jointly optimizing its projector and velocity field on the frozen encoder features. The projector maps the 1400 features to a d_c=512 context. In inter-subject evaluation, all fitting and validation selection occur within the training subjects of each split. The comparison fixes representation width, fMRI targets, split policy, normalization, generator architecture, sampler, and metrics while varying the EEG encoder configuration.

Generative model baselines. The baseline families build on adversarial learning (Goodfellow et al., 2014), diffusion (Sohl-Dickstein et al., 2015; Dhariwal & Nichol, 2021), and score estimation (Song & Ermon, 2019), with conditional diffusion also supporting image translation (Saharia et al., 2022). All

rows use the same seven frozen NVCode encoders. The projector is jointly trained with the NVCode velocity field, then frozen and reused when fitting the baseline generation networks. Thus, every complete generator has the form $G_j(\mathbf{e}, \epsilon) = F_{\theta_j}(\Pi_{\psi^\star}(\mathbf{e}), \epsilon)$, with shared projector parameters $\psi^\star$ and separately fitted network parameters θ_j. The projector belongs to the generator; this comparison holds that component fixed and evaluates generation networks on the same stored NVCode-learned conditions. Each model uses the same z-scored labels and data splits. At test time, each stochastic model returns $K{=}50$ samples per EEG epoch, evaluated using the same predictive-mean, marginal CRPS, and predictive-mean FC-MAE routines. The Rectified Flow control (Liu et al., 2022) shares the Gaussian source, linear probability path, and squared velocity objective with NVCode. Their training procedures compare direct fitting to observed endpoints with one round of reflow. Let $\mathcal{D}_{\mathrm{tr}} = \{(\mathbf{c}_i, \mathbf{z}_{1,i})\}_{i=1}^{M}$, where $\mathbf{c}_i = \mathbf{c}_\psi(\mathbf{x}_i)$ is the fixed EEG condition and $\mathbf{z}_{1,i}$ is the normalized joint fMRI endpoint. Define

$$\gamma_t(\mathbf{a}, \mathbf{b}) = (1-t)\mathbf{a} + t\mathbf{b}, \qquad u(\mathbf{a}, \mathbf{b}) = \mathbf{b} - \mathbf{a}. \tag{9}$$

With $\mathbf{z}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$ and $t \sim \mathcal{U}(0,1)$, the field optimum for a given projector is

$$\theta^\star(\psi) \in \arg\min_\theta \mathbb{E}\left[\left\| v_\theta(\gamma_t(\mathbf{z}_0, \mathbf{z}_1), t, \mathbf{c}_\psi(\mathbf{x})) - u(\mathbf{z}_0, \mathbf{z}_1)\right\|_2^2\right]. \tag{10}$$

Equation (10) expresses the field optimization within Eq. (6). NVCode learns the projector and field jointly through that objective. The initial Rectified Flow fit optimizes the field with $\psi = \psi^\star$ fixed. Both use observed EEG-paired fMRI endpoints, the same field architecture, and the same source–endpoint coupling. The common path and objective provide the starting point for the training-procedure comparison below.

The Rectified Flow row in Tables 2 and 11 denotes a specified one-reflow implementation. Its first field $v_{\theta^{(1)}}$ is fitted with Eq. (10). For each training condition $\mathbf{c}_i$, we draw and store one source $\bar{\mathbf{z}}_{0,i}$, then integrate the first field from that same source to construct its paired model-induced endpoint

$$\widehat{\mathbf{z}}_{1,i} = \Phi_{1\leftarrow 0}^{\theta^{(1)}}(\bar{\mathbf{z}}_{0,i}; \mathbf{c}_i), \qquad \frac{d\mathbf{z}(t)}{dt} = v_{\theta^{(1)}}(\mathbf{z}(t), t, \mathbf{c}_i), \qquad \mathbf{z}(0) = \bar{\mathbf{z}}_{0,i}, \tag{11}$$

where $\Phi_{1\leftarrow 0}^{\theta^{(1)}}$ is the numerical flow map. We store the paired bank $\mathcal{B}_{\mathrm{RF}} = \{(\mathbf{c}_i, \bar{\mathbf{z}}_{0,i}, \widehat{\mathbf{z}}_{1,i})\}_{i=1}^{M}$ without re-pairing sources and endpoints. A freshly initialized second field is fitted on straight segments from these stored pairs:

$$\theta_{\mathrm{RF}} \in \arg\min_\theta \mathbb{E}_{t, \mathcal{B}_{\mathrm{RF}}}\left[\left\| v_\theta(\gamma_t(\bar{\mathbf{z}}_0, \widehat{\mathbf{z}}_1), t, \mathbf{c}) - u(\bar{\mathbf{z}}_0, \widehat{\mathbf{z}}_1)\right\|_2^2\right], \tag{12}$$

Thus, NVCode directly fits observed condition–endpoint pairs once, whereas the Rectified Flow control adds a teacher-generated endpoint bank, fresh second-stage initialization, and a one-reflow objective. The label "Rectified Flow" consistently denotes this specified one-reflow implementation in our tables.

For the NVCode–Rectified Flow comparison, both rows use the same seven frozen regional encoders, the same 1400-dimensional concatenated representation and 512-dimensional context, the same seven-dimensional target ordering, training-only normalization statistics, subject or temporal split, Gaussian source law, time distribution, linear interpolation, vector-valued velocity-network parameterization, optimizer settings, and validation criterion. The same total optimization budget is used: NVCode applies it to one direct fit, while Rectified Flow divides it equally between the initial fit and the one-reflow fit. Final inference is also matched at $S{=}100$ Euler steps and $K{=}50$ sources per EEG epoch, after which both methods use the same inverse normalization and evaluation code. Rectified Flow receives no held-out fMRI value when forming a test condition or trajectory.

This control fixes the EEG evidence pathway, joint output space, total optimization budget, and inference effort. The two rows compare direct fitting to observed EEG-paired endpoints with the complete one-reflow training procedure. The evaluated intervention comprises stage count, initialization, budget allocation, and teacher-generated endpoints. Encoder quality, conditioning dimensionality, regional output factorization, normalization, and sampling effort remain matched.

Predictive mean Pearson correlation. Let $\hat{Y}, Y \in \mathbb{R}^{M\times N}$ stack M epoch level predictive mean predictions and ground truths within one evaluated scan or split. The predictive mean Pearson correlation (Pearson, 1895) averaged over fMRI signals is

$$\overline{R} \;=\; \tfrac{1}{N}\sum_{n=1}^{N} \frac{\sum_{i=1}^{M}(\hat{Y}_{i,n} - \bar{\hat{Y}}_n)(Y_{i,n} - \bar{Y}_n)}{\sqrt{\sum_{i=1}^{M}(\hat{Y}_{i,n} - \bar{\hat{Y}}_n)^2}\,\sqrt{\sum_{i=1}^{M}(Y_{i,n} - \bar{Y}_n)^2}},$$

where $\bar{\hat{Y}}_n, \bar{Y}_n$ are the column means. Higher is better.

Mean marginal CRPS. The mean marginal continuous ranked probability score is computed per epoch from the K Monte Carlo samples, averaged over fMRI signals and then averaged over the scan:

$$\overline{\mathrm{CRPS}} = \frac{1}{M}\sum_{i=1}^{M}\left[\frac{1}{KN}\sum_{k=1}^{K}\|\hat{\mathbf{y}}_i^{(k)}-\mathbf{y}_i\|_1 - \frac{1}{2K^2N}\sum_{j,k=1}^{K}\|\hat{\mathbf{y}}_i^{(j)}-\hat{\mathbf{y}}_i^{(k)}\|_1\right].$$

The bracketed term equals $\frac{1}{N}\sum_{n=1}^{N}\mathrm{CRPS}(\widehat{F}_{i,n}, y_{i,n})$, where $\widehat{F}_{i,n}$ is the empirical distribution of $P_{i,n} = \{\hat{y}_{i,n}^{(k)}\}_{k=1}^{K}$. CRPS is a strictly proper scoring rule for each marginal predictive distribution (Matheson & Winkler, 1976; Gneiting & Raftery, 2007). The reported average summarizes marginal probabilistic prediction quality across the seven targets. Lower is better. Equivalently, if $\widehat{F}_{i,n} = K^{-1}\sum_{k=1}^{K}\delta_{\hat{y}_{i,n}^{(k)}}$ is the empirical predictive distribution, then

$$\mathrm{CRPS}(\widehat{F}_{i,n}, y_{i,n}) = \frac{1}{K}\sum_{k=1}^{K}|\hat{y}_{i,n}^{(k)} - y_{i,n}| - \frac{1}{2K^2}\sum_{j,k=1}^{K}|\hat{y}_{i,n}^{(j)} - \hat{y}_{i,n}^{(k)}|. \tag{13}$$

This is the exact CRPS of the empirical predictive CDF. It follows by substituting independent draws from $\widehat{F}_{i,n}$ into the energy representation $\mathrm{CRPS}(F, y) = \mathbb{E}|X - y| - \frac{1}{2}\mathbb{E}|X - X'|$.

Functional connectivity error. Correlation-based functional networks summarize statistical associations among brain signals (Bassett & Sporns, 2017; Deligianni et al., 2014; Yeo et al., 2011). Let $Q_{:,n}$ denote the nth column of any $M \times N$ matrix and $\rho_{ab}(Q) = \mathrm{corr}(Q_{:,a}, Q_{:,b})$ be the Pearson correlation (Pearson, 1895) between columns a and b. The signal dependence diagnostic is the off diagonal mean absolute error between predicted and ground truth 7×7 signal correlation matrices:

$$\text{FC-MAE} = \frac{1}{N(N-1)}\sum_{a\neq b}\left|\rho_{ab}(\hat{Y}) - \rho_{ab}(Y)\right|.$$

Here $\hat{Y}$ contains the K-sample predictive mean for each EEG window. The metric matches off-diagonal entries of the predictive-mean and observed signal correlation matrices (Van Den Heuvel & Pol, 2010; Friston, 2011). It evaluates interregional organization across EEG windows through the resulting predictive-mean trajectories. Repeated samples for each fixed EEG condition enter this metric through their mean. Lower is better.

Lemma 1 (Affine properties of the reported metrics)**.** *Let $\mathcal{T}(\mathbf{q}) = \boldsymbol{\sigma} \odot \mathbf{q} + \boldsymbol{\mu}$ with $\sigma_n > 0$ for every signal. Applying $\mathcal{T}$ to both predictions and targets leaves every per-signal Pearson correlation and the FC-MAE unchanged. For the n-th signal,*

$$\mathrm{CRPS}(\mathcal{T}_n \# F, \mathcal{T}_n(y)) = \sigma_n\ \mathrm{CRPS}(F, y). \tag{14}$$

Proof. For any temporal vector $\mathbf{q}$, $\mathcal{T}_n(\mathbf{q}) - \overline{\mathcal{T}_n(\mathbf{q})} = \sigma_n(\mathbf{q} - \bar{\mathbf{q}})$. The positive scale cancels between the covariance numerator and standard-deviation denominator of Pearson correlation. The same cancellation holds independently for both columns of every functional-connectivity entry. Therefore, the complete correlation matrices and their elementwise differences are unchanged. For CRPS, both absolute differences in the energy representation acquire the common factor σ_n, which proves Eq. (14). □

Lemma 1 explains why inverse generator normalization preserves the temporal and cross-regional correlation diagnostics exactly. It also restores the signalwise scale on which marginal CRPS is aggregated and reported.

Aggregation. Within each seed run, reconstruction metrics are computed per held out temporal segment in the intra-subject protocol and per held out subject split in the inter-subject protocol. If $m_{d,s}$ denotes a metric for split s of dataset d in that run, its dataset and benchmark summaries are

$$\bar{m}_d = \frac{1}{|\mathcal{S}_d|}\sum_{s\in\mathcal{S}_d} m_{d,s}, \qquad \bar{m} = \frac{1}{8}\sum_{d=1}^{8}\bar{m}_d. \tag{15}$$

This two-level macro average gives each dataset equal weight after its internal splits are summarized. Reported reconstruction means average these summaries over the three seed runs. For each dataset, method, evaluation setting, and target signal, the standard deviation in Appendix C is computed across the three run-level correlation scores after split aggregation. It describes run-to-run variability under the fixed data splits. The sampled response vectors enter the predictive-mean and marginal-score calculations within each split.

A.3 Downstream Utility Details

The downstream utility experiment asks whether synthetic fMRI can serve as a useful test input for a model trained only on real fMRI. This probe evaluates whether generated fMRI preserves discriminative information used by a downstream classifier.

TrainReal TestSynth protocol. For each dataset, we train a subject identification classifier on real robust-normalized fMRI vectors from the real fMRI training temporal segments and validate it on real validation segments. The subject label space is fixed within the dataset. Classifier fitting, validation, and checkpoint selection use exclusively real fMRI data. After selection, the classifier is frozen. The real test row in Tables 3 and 6 evaluates this classifier on held out real fMRI. A synthetic row evaluates the same classifier on generated fMRI from matched held out EEG epochs. Thus, Δ measures test-input utility through the change in subject-identification accuracy. Classifier capacity and training data remain fixed when synthetic signals replace real test inputs.

Classifier architecture and optimization. The classifier is a lightweight two layer network:

$$\mathrm{Linear}(7, 64) \to \mathrm{LayerNorm} \to \mathrm{GELU} \to \mathrm{Dropout}(0.1) \to \mathrm{Linear}(64, S_{\mathrm{subj}}),$$

where S_{subj} is the number of subject labels in the dataset and LayerNorm follows Ba et al. (2016). The classifier is trained with cross entropy, AdamW (Kingma & Ba, 2014; Loshchilov & Hutter, 2017), learning rate 10^{-3}, weight decay 10^{-4}, batch size 256 and early stopping on real validation accuracy with patience 20 epochs.

Utility score. For each dataset and synthetic method, we report

$$\Delta = \mathrm{acc}_{\mathrm{TrainReal,TestSynth}} - \mathrm{acc}_{\mathrm{TrainReal,TestReal}}.$$

A higher Δ means that the frozen classifier attains higher identity accuracy on synthetic fMRI. The headline $\bar{\Delta} = \frac{1}{8}\sum_{d=1}^{8} \Delta_d$ weights each dataset equally. Tables report gains to three decimal places; multiplying an unrounded gain by 100 expresses it in percentage points. This score complements reconstruction metrics by measuring classifier-accessible identity structure under the shift from real training inputs to synthetic test inputs.

A.4 Ablation Study Details

Tables 4 and 6 detail two component studies: inter-subject encoder/generator variants and fixed-generator downstream transfer from the complete versus single-encoder representation sets.

A.4.1 Single-signal regional-encoder ablation

This study compares the complete seven-encoder representation set with each single-encoder alternative while fixing the generator, sampler, generation space, downstream classifier, and evaluation metric.

Single-signal representation variants. Each alternative retains the Cuneus, Heschl's, Middle Frontal, Precuneus, Putamen, Thalamus, or Global Signal representation map. The retained map enters the same type of learned projector, preserving condition dimensionality. Generator training and downstream evaluation remain identical to full NVCode; generator capacity is fixed while the supplied representation set changes.

Results and interpretation. Full NVCode achieves the highest mean utility lift, $\bar{\Delta} = +0.031$, compared with $+0.018$ for Precuneus only and $+0.013$ for Global Signal only. Its gains reach $+0.045$ on NatView, $+0.043$ on NeuroBOLT, and $+0.036$ on gradCPT, with positive gains on all eight datasets. Table 6 reports the same full NVCode results as Table 3. Under the shared evaluation protocol, the complete seven-encoder representation set provides the highest average classifier-accessible identity utility among the tested representation sets.

Table 6: **Full regional aggregation provides the strongest downstream transfer.** The generator, sampler, training protocol, and real-trained classifier are fixed; only the encoder set supplied to the context projector varies. Best values are **darker purple, bold**; second-best values are light purple, underlined.

Encoder	Dataset accuracy / Δ								Mean
	Sleep EEG–fMRI	MULTI-CLARID	gradCPT	NeuroBOLT	XP1	Affective Music	NatView	Experience Sampling	
Real test	0.861	0.787	0.746	0.869	0.874	0.782	0.747	0.702	0.796
Cuneus only	+0.004	+0.016	+0.018	+0.000	-0.011	-0.001	+0.007	-0.001	+0.004
Heschl's only	-0.013	-0.017	+0.021	-0.022	+0.002	-0.001	-0.081	-0.010	-0.015
Mid. Frontal only	+0.006	-0.015	-0.018	-0.002	-0.006	+0.001	-0.016	-0.004	-0.007
Precuneus only	+0.022	+0.008	+0.033	-0.003	**+0.040**	+0.021	+0.008	+0.014	+0.018
Putamen only	-0.016	-0.015	-0.024	-0.019	+0.005	-0.024	-0.082	-0.036	-0.026
Thalamus only	-0.018	-0.009	-0.025	-0.041	-0.054	-0.034	-0.074	+0.002	-0.032
Global Signal only	+0.024	+0.003	-0.006	+0.003	+0.036	+0.021	-0.001	**+0.021**	+0.013
NVCode (Ours)	**+0.028**	**+0.027**	**+0.036**	**+0.043**	+0.019	**+0.031**	**+0.045**	+0.020	**+0.031**

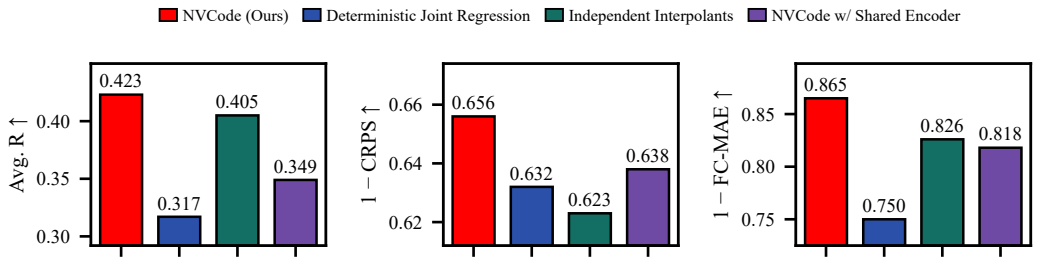


Figure 8: **The complete NVCode architecture performs best across all three reconstruction criteria.** The four matched variants from Table 4 are compared under inter-subject evaluation. From left to right, the figure reports Avg. R, $1 - \text{CRPS}$, and $1 - \text{FC-MAE}$; the two error scores are transformed only to align directionality, so taller bars are favorable throughout.

A.4.2 Controlled joint-output and stochastic-interpolant ablation

This study compares full NVCode with operational variants that remove stochastic interpolation, factorize the output field, or replace the seven signal-specific encoders with one shared encoder. Table 4 reports the three metrics used throughout the main comparison under the inter-subject protocol. The output modules are capacity matched where applicable; the shared-encoder variant changes target-specific supervision and total encoder-side capacity together.

Shared experimental base. Every variant is trained and evaluated separately on all eight datasets in Table 5, using the same retained EEG–fMRI pairs, robust-normalized targets, fixed subject-disjoint splits, and seven-dimensional output ordering. The regional-encoder variants start from the same seven frozen checkpoints selected by validation MSE. Each variant receives a new copy of the two-layer context projector in Appendix B.3, with $d_c = 512$, and matched optimization, early stopping, and checkpoint selection. Output-module widths are adjusted so that their trainable parameter counts differ by at most 1%. The shared-encoder configuration jointly varies encoder-side capacity and target-specific supervision. Reported scores are equal-weight macro averages over the eight dataset-level results.

Deterministic joint regression. The deterministic joint regressor replaces the seven scalar heads with one multi-output network,

$$\hat{\mathbf{y}} = g_\omega(\mathbf{c}_\psi(\mathbf{x})) \in \mathbb{R}^N, \qquad \mathcal{L}_{\text{joint-reg}} = \frac{1}{N} \|\hat{\mathbf{y}} - \mathbf{y}\|_2^2 .$$

Its MLP follows the hidden-width sequence of the NVCode velocity network after removing the state and time inputs, and its width is adjusted only for parameter matching. This variant retains the complete regional encoder aggregation and one shared seven-dimensional output while removing Gaussian source sampling, velocity matching, and ODE integration. The comparison evaluates conditional stochastic interpolation against an output-module-capacity-matched deterministic joint predictor.

Independent regional interpolants. The independent-interpolant variant uses seven scalar conditional velocity fields. Each signal has an independent Gaussian source and scalar interpolant,

$$z_{0,n} \sim \mathcal{N}(0,1), \qquad z_{t,n} = (1-t)z_{0,n} + tz_{1,n}, \qquad u_{t,n} = \dot{z}_{t,n} = z_{1,n} - z_{0,n}.$$

Each scalar field is trained to match $u_{t,n}$ and generates with $ds_n(t)/dt = v_{\theta_n}(s_n(t), t, \mathbf{c}_\psi(\mathbf{x}))$. Each field receives its own evolving coordinate together with the shared EEG condition. Their total parameter count is matched to the joint NVCode field, and each field is sampled with the same $S = 100$ Euler steps and $K = 50$ test sources. This comparison tests a single field with cross-coordinate state access against capacity-matched factorized fields under the same condition and sampling effort.

Shared-encoder control and full NVCode. The shared-encoder control trains one EEG encoder with a 1400-dimensional representation and a seven-output temporary regression head. After validation selection, the head is discarded and the frozen encoder supplies its representation directly to the generator's context projector. This matches the 7×200 feature interface of NVCode while replacing seven independently supervised encoders with one shared encoder. The projector architecture and velocity network remain unchanged. Full NVCode retains the seven signal-specific encoders, joint context projector, seven-dimensional Gaussian source, and joint velocity field defined in Sections 3.2–3.5.

Evaluation and aggregation. Stochastic variants return $K = 50$ samples per held-out EEG epoch, while deterministic variants return one point prediction. Avg. R and FC-MAE are computed from the predictive mean using the definitions in Appendix A.2. For a deterministic prediction, CRPS is evaluated as the point-mass score $\mathrm{CRPS}(\delta_{\hat{y}}, y) = |\hat{y} - y|$, which places all rows on the same lower-is-better scale. Scores are summarized first within held-out subjects and then by the equal-weight dataset-level macro average in Eq. (15).

Component contributions. Figure 8 places the three criteria on a common higher-is-better scale. Full NVCode leads every tested variant, reaching an Avg. R of 0.423, a CRPS of 0.344, and an FC-MAE of 0.135. Compared with deterministic joint regression, NVCode improves Avg. R by 0.106 and reduces predictive-mean FC-MAE by 0.115. Independent regional interpolants achieve an Avg. R of 0.405, a marginal CRPS of 0.377, and a predictive-mean FC-MAE of 0.174; joint NVCode generation improves all three criteria. Relative to the shared-encoder configuration, NVCode raises Avg. R by 0.074 and reduces CRPS and FC-MAE by 0.018 and 0.047, respectively. The first two controlled comparisons support stochastic interpolation and joint state interaction within the evaluated output module. The shared-encoder comparison establishes the predictive advantage of the complete seven-encoder configuration under the stated supervision and capacity settings.

Pareto separation. Full NVCode also exceeds the strongest ablated alternative for each individual criterion. Its Avg. R is 0.018 higher than independent regional interpolants, its marginal CRPS is 0.018 lower than the shared-encoder control, and its predictive-mean FC-MAE is 0.039 lower than independent regional interpolants. The complete configuration thus leads all three criteria simultaneously, combining temporal alignment, marginal probabilistic quality, and predictive-mean interregional organization among the tested variants.

B METHOD DETAILS

NVCode connects target-supervised EEG representation learning to joint fMRI generation through a shared neurovascular code. This appendix specifies that representation-to-generation pathway and derives its conditional velocity-matching foundation (Lipman et al., 2022; Albergo et al., 2025). The component comparisons in Appendix A.4 evaluate the predictive contribution of this model organization.

B.1 VELOCITY MATCHING IDENTITY

This section specializes the conditional velocity-matching identity (Lipman et al., 2022; Albergo et al., 2025) to a fixed neurovascular code map. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $(\mathbf{X}, \mathbf{Z}_1)$ be a paired EEG input and normalized fMRI endpoint, and set $\mathbf{C} = \mathbf{c}_\psi(\mathbf{X})$. Draw $\mathbf{Z}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$ independently of $(\mathbf{X}, \mathbf{Z}_1)$ and $T \sim \mathcal{U}(0,1)$ independently of $(\mathbf{X}, \mathbf{Z}_0, \mathbf{Z}_1)$. Define

$$\mathbf{Z}_T = (1-T)\mathbf{Z}_0 + T\mathbf{Z}_1, \qquad \mathbf{U} = \mathbf{Z}_1 - \mathbf{Z}_0. \tag{16}$$

Assume $\mathbb{E}(\|\mathbf{Z}_0\|_2^2 + \|\mathbf{Z}_1\|_2^2) < \infty$, so the conditional velocities below are square integrable.

Conditional L_2 projection. Let $\mathcal{H} = \sigma(\mathbf{Z}_T, T, \mathbf{C})$ and define $\mathbf{m} = \mathbb{E}[\mathbf{U} \mid \mathcal{H}]$.

Proposition 1 (Optimal conditional velocity)**.** *For a fixed conditioning map $\mathbf{C} = \mathbf{c}_\psi(\mathbf{X})$, the minimizer of the population squared velocity-matching risk over square-integrable measurable fields is*

$$v^\star(\mathbf{z}, t, \mathbf{c}) = \mathbb{E}[\mathbf{Z}_1 - \mathbf{Z}_0 \mid \mathbf{Z}_T = \mathbf{z},\, T = t,\, \mathbf{C} = \mathbf{c}]\,. \tag{17}$$

Lemma 2 (Orthogonal risk decomposition)**.** *For every square integrable $\mathcal{H}$-measurable vector $\mathbf{V}$,*

$$\mathbb{E}\|\mathbf{U} - \mathbf{V}\|_2^2 = \mathbb{E}\|\mathbf{U} - \mathbf{m}\|_2^2 + \mathbb{E}\|\mathbf{m} - \mathbf{V}\|_2^2. \tag{18}$$

Consequently, $\mathbf{m}$ is the unique minimizer up to almost-sure equality.

Proof. Expanding around $\mathbf{m}$ gives

$$\begin{aligned}\mathbb{E}\|\mathbf{U} - \mathbf{V}\|_2^2 = \mathbb{E}\|\mathbf{U} - \mathbf{m}\|_2^2 + \mathbb{E}\|\mathbf{m} - \mathbf{V}\|_2^2 \\ + 2\,\mathbb{E}\big[(\mathbf{U} - \mathbf{m})^\top(\mathbf{m} - \mathbf{V})\big]\,.\end{aligned} \tag{19}$$

Since $\mathbf{m} - \mathbf{V}$ is $\mathcal{H}$-measurable, the cross term vanishes by the tower property:

$$\mathbb{E}\big[(\mathbf{U} - \mathbf{m})^\top(\mathbf{m} - \mathbf{V})\big] = \mathbb{E}\big[\mathbb{E}[\mathbf{U} - \mathbf{m} \mid \mathcal{H}]^\top(\mathbf{m} - \mathbf{V})\big] = 0. \tag{20}$$

The remaining excess-risk term is nonnegative and vanishes precisely when $\mathbf{V} = \mathbf{m}$ almost surely. □

With $\mathbf{V} = v(\mathbf{Z}_T, T, \mathbf{C})$, Lemma 2 proves Proposition 1 and yields

$$\mathcal{L}(v) - \mathcal{L}(v^\star) = \mathbb{E}\big[\|v(\mathbf{Z}_T, T, \mathbf{C}) - v^\star(\mathbf{Z}_T, T, \mathbf{C})\|_2^2\big]\,. \tag{21}$$

The excess risk is therefore the squared field error averaged over the conditional interpolation path.

Conditional continuity equation. The optimal field induces the continuity equation for the complete conditional probability path.

Theorem 1 (Conditional probability-path identity)**.** *Fix a condition $\mathbf{c}$. Suppose the conditional law of $\mathbf{Z}_t$ has a differentiable density $p_t(\mathbf{z} \mid \mathbf{c})$, and differentiation may pass through the conditional expectation. Then $v^\star$ satisfies*

$$\partial_t p_t(\mathbf{z} \mid \mathbf{c}) + \nabla_{\mathbf{z}} \cdot [p_t(\mathbf{z} \mid \mathbf{c}) v^\star(\mathbf{z}, t, \mathbf{c})] = 0 \tag{22}$$

in the weak sense.

Proof. Let $\varphi \in C_c^\infty(\mathbb{R}^N)$ be a smooth compactly supported test function. Since $\dot{\mathbf{Z}}_t = \mathbf{U}$,

$$\begin{aligned}\frac{d}{dt}\int \varphi(\mathbf{z}) p_t(\mathbf{z} \mid \mathbf{c})\, d\mathbf{z} &= \frac{d}{dt}\mathbb{E}[\varphi(\mathbf{Z}_t) \mid \mathbf{C} = \mathbf{c}] \\ &= \mathbb{E}[\nabla\varphi(\mathbf{Z}_t)^\top \mathbf{U} \mid \mathbf{C} = \mathbf{c}] \\ &= \mathbb{E}[\nabla\varphi(\mathbf{Z}_t)^\top v^\star(\mathbf{Z}_t, t, \mathbf{c}) \mid \mathbf{C} = \mathbf{c}] \\ &= \int \nabla\varphi(\mathbf{z})^\top v^\star(\mathbf{z}, t, \mathbf{c}) p_t(\mathbf{z} \mid \mathbf{c})\, d\mathbf{z}.\end{aligned} \tag{23}$$

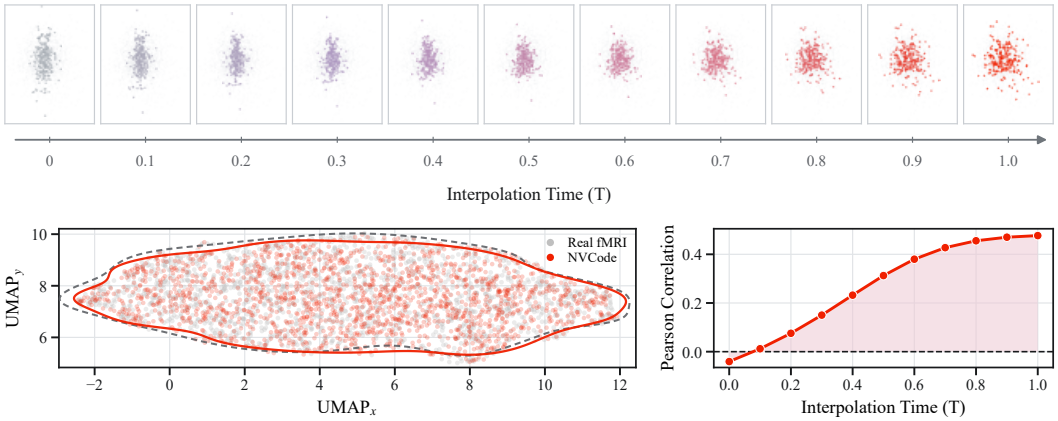


Figure 9: **Projected diagnostic of EEG-conditioned transport on a held-out scan. Top:** projected samples evolve from the Gaussian source at $t = 0$ to the generated endpoint at $t = 1$. **Bottom left:** real and generated endpoints overlap in the displayed UMAP projection. **Bottom right:** target correlation increases along the numerically integrated learned trajectory.

The tower property gives the third line. Integration by parts, with vanishing boundary term for compactly supported φ, gives Eq. (22). □

Proposition 2 (Conditional endpoint transport)**.** *Assume the ODE*

$$\frac{d\mathbf{z}(t)}{dt} = v^\star(\mathbf{z}(t), t, \mathbf{c}), \qquad \mathbf{z}(0) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N), \tag{24}$$

has a unique solution map $\Phi_t^{\mathbf{c}}$, and Eq. (22) has a unique probability-density solution. Then

$$(\Phi_t^{\mathbf{c}})_{\#}\mathcal{N}(\mathbf{0}, \mathbf{I}_N) = p_t(\cdot \mid \mathbf{c}), \qquad (\Phi_1^{\mathbf{c}})_{\#}\mathcal{N}(\mathbf{0}, \mathbf{I}_N) = \mathcal{L}(\mathbf{Z}_1 \mid \mathbf{C} = \mathbf{c}). \tag{25}$$

Proof. The ODE pushforward and the interpolant density from Theorem 1 solve the same continuity equation. Independence of $\mathbf{Z}_0$ and $\mathbf{C}$ gives both paths the initial law $\mathcal{N}(\mathbf{0}, \mathbf{I}_N)$. Uniqueness identifies the paths, and $\mathbf{Z}_t = \mathbf{Z}_1$ at $t = 1$ gives the endpoint law. □

Proposition 2 identifies the endpoint law represented by a fixed code. When that code is conditionally sufficient for the target, the same transport also recovers the EEG-conditional law:

$$\mathbf{Z}_1 \perp\!\!\!\perp \mathbf{X} \mid \mathbf{C}, \qquad \mathcal{L}(\mathbf{Z}_1 \mid \mathbf{C} = \mathbf{c}_\psi(\mathbf{x})) = \mathcal{L}(\mathbf{Z}_1 \mid \mathbf{X} = \mathbf{x}). \tag{26}$$

These identities hold for the population-optimal field under the stated regularity and uniqueness assumptions. NVCode fits a neural approximation to this field using target-supervised codes and the velocity objective in Eq. (6).

Together, Lemma 2, Theorem 1, and Proposition 2 connect conditional velocity regression to transport of the joint fMRI endpoint law.

Empirical transport diagnostic. Figure 9 visualizes the numerically integrated field on a held-out scan. Eleven projected states trace the evolution from Gaussian sources to generated endpoints, alongside endpoint overlap in UMAP and increasing target correlation along the displayed trajectory.

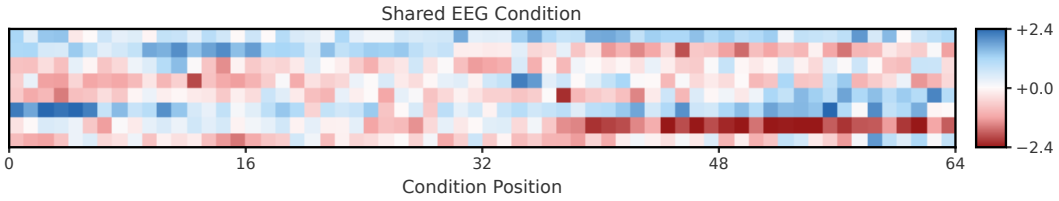


Figure 10: **Schematic layout of the shared EEG condition.** The $d_c = 512$ positions are displayed as an 8×64 matrix. Entries are deterministic centered artwork illustrating the vector layout; red, white, and blue denote negative, near-zero, and positive illustrative values.

B.2 Regional EEG Evidence Encoders

NVCode learns a target-specific EEG representation with each encoder g_{ϕ_n} in Eq. (2). Each encoder combines a spatiotemporal branch with a multi-scale spectral branch. The N encoders have independent parameters and receive supervision from their corresponding fMRI targets, producing seven EEG feature vectors for the shared conditioning module.

EEG representation learning also uses masked segment reconstruction (Cui et al., 2024) and geometry-aware pretraining across electrode layouts (Yi et al., 2023). Here, the regional prediction objectives define the supervision for the temporal and spectral features specified below.

Spatiotemporal branch. Patching provides a tokenized interface for time-series representation learning (Nie et al., 2022). The EEG window $\mathbf{x} \in \mathbb{R}^{C\times T}$ is divided into $P = T/w = 16$ non-overlapping patches of $w = 200$ samples per channel. A convolutional patcher $\mathcal{P}_{\phi_n} : \mathbb{R}^w \to \mathbb{R}^d$, channel embedding $\mathbf{a}_c^{\text{ch}}$, and temporal embedding $\mathbf{a}_p^{\text{time}}$ form each token:

$$\mathbf{q}_{c,p}^{\text{ts}} = \mathcal{P}_{\phi_n}\left(\mathbf{x}_{c,(p-1)w+1:pw}\right) + \mathbf{a}_c^{\text{ch}} + \mathbf{a}_p^{\text{time}}, \qquad c = 1,\dots,C, \quad p = 1,\dots,P. \tag{27}$$

The resulting sequence $\mathbf{Q}^{\text{ts}} \in \mathbb{R}^{CP\times d}$ retains each patch's channel–time identity.

A learnable [CLS] token (Devlin et al., 2019) is inserted to support global attention exchange during the Transformer update. The sequence is processed by a pre-norm Transformer with depth 12, hidden dimension $d = 200$, 10 attention heads, MLP expansion ratio 4, qkv bias enabled, stochastic depth rate 0.1, and layer-scale initialization 0.1. If $\mathcal{T}_{\phi_n}^{\text{ts}}$ denotes this Transformer and $[\cdot]_{c,p}$ selects a contextualized EEG token, the branch output is

$$h_{\phi_n}^{\text{ts}}(\mathbf{x}) = \frac{1}{CP}\sum_{c=1}^{C}\sum_{p=1}^{P}\left[\mathcal{T}_{\phi_n}^{\text{ts}}\left([\mathbf{q}_{\text{CLS}}; \mathbf{Q}^{\text{ts}}]\right)\right]_{c,p} \in \mathbb{R}^d. \tag{28}$$

The [CLS] token participates in attention, while the branch output averages the CP contextualized EEG tokens.

Multi-scale spectral branch. For each channel, the spectral branch computes Short-Time Fourier Transform magnitudes at four FFT lengths:

$$L_\ell \in \{100, 200, 400, 800\}.$$

Non-overlapping windows with hop length L_ℓ give $W_\ell = T/L_\ell$ time windows and $F_\ell = L_\ell/2 + 1$ frequency bins at scale ℓ:

$$\mathbf{S}_\ell(\mathbf{x}) \in \mathbb{R}^{C\times F_\ell \times W_\ell}.$$

Learned projections align every scale to the same frequency and temporal dimensions. For a channel slice $\mathbf{S}_{\ell,c}(\mathbf{x}) \in \mathbb{R}^{F_\ell\times W_\ell}$, use $\mathbf{P}_\ell^{\text{freq}} \in \mathbb{R}^{F_\ell\times d}$ and $\mathbf{P}_\ell^{\text{time}} \in \mathbb{R}^{W_\ell\times 32}$ to obtain

$$\bar{\mathbf{S}}_{\ell,c}(\mathbf{x}) = (\mathbf{P}_\ell^{\text{freq}})^\top \mathbf{S}_{\ell,c}(\mathbf{x})\mathbf{P}_\ell^{\text{time}} \in \mathbb{R}^{d\times 32}.$$

The four aligned tensors are summed before tokenization. For aligned temporal coordinate $u \in \{1,\dots,32\}$, define

$$\mathbf{q}_{c,u}^{\text{mss}} = \sum_{\ell=1}^{4} \bar{\mathbf{S}}_{\ell,c}(\mathbf{x})_{:,u} + \mathbf{b}_c^{\text{ch}} + \mathbf{b}_u^{\text{time}} \in \mathbb{R}^d. \tag{29}$$

A four-layer linear-attention Transformer $\mathcal{T}^{\text{mss}}_{\phi_n}$ processes the stacked sequence $\mathbf{Q}^{\text{mss}} \in \mathbb{R}^{32C\times d}$, followed by mean pooling:

$$h^{\text{mss}}_{\phi_n}(\mathbf{x}) = \frac{1}{32C}\sum_{c=1}^{C}\sum_{u=1}^{32}\left[\mathcal{T}^{\text{mss}}_{\phi_n}(\mathbf{Q}^{\text{mss}})\right]_{c,u} \in \mathbb{R}^d. \tag{30}$$

Equation (29) combines the four spectral resolutions in a fixed $32C$-token sequence.

Representation output. The final regional EEG representation combines the temporal and spectral branches:

$$\mathbf{e}_n = g_{\phi_n}(\mathbf{x}) = h^{\text{ts}}_{\phi_n}(\mathbf{x}) + h^{\text{mss}}_{\phi_n}(\mathbf{x}).$$

The resulting $d = 200$ feature vector carries both branches into the generator's shared conditioning module after its temporary scalar head is discarded.

Encoder-level comparison with NeuroBOLT. The adapted NeuroBOLT baseline in Table 1 uses one encoder with a $D = 1400$ output, as do the other encoder baselines. NVCode uses $N = 7$ independently parameterized regional encoders, each returning $d = 200$ features. Writing h_η for the complete baseline feature map, the two representation interfaces are

$$\mathbf{e}_{\text{base}}(\mathbf{x}) = h_\eta(\mathbf{x}) \in \mathbb{R}^D, \quad \mathbf{e}_{\text{NVCode}}(\mathbf{x}) = [g_{\phi_1}(\mathbf{x}); \ldots; g_{\phi_N}(\mathbf{x})] \in \mathbb{R}^{Nd}, \quad D = Nd = 1400.$$

Here D specifies the baseline output width, while d specifies each NVCode encoder's feature width. Each regional objective in Eq. (2) updates its own encoder, preserving the corresponding feature block:

$$\nabla_{\phi_m}\mathcal{L}^{(n)}_{\text{rep}} = \mathbf{0} \quad (m \neq n), \qquad \mathbf{e}_{\text{NVCode}}[(n-1)d + 1 : nd] = g_{\phi_n}(\mathbf{x}).$$

This organization preserves seven target-supervised feature groups at the same total representation width. Table 7 separates encoder organization, supervision, and parameter allocation from the downstream modules.

Table 7: **Encoder-level comparison at a matched representation width.** NeuroBOLT (Li et al., 2024b) denotes the single-encoder adaptation in Table 1. Both configurations output 1400 features; NVCode organizes these as seven independently supervised 200-dimensional regional representations.

Encoder aspect	NeuroBOLT (adapted encoder)	NVCode encoders
Encoder organization	One feature map h_η for each EEG window.	Seven feature maps $g_{\phi_1}, \ldots, g_{\phi_7}$ for the same EEG window.
Feature allocation	One 1400-dimensional representation.	Seven 200-dimensional regional blocks, concatenated into 1400 dimensions.
Supervision	The temporary fMRI readout supervises one common EEG representation.	A separate scalar head supervises each regional representation through $\mathcal{L}^{(n)}_{\text{rep}}$.
Parameter updates	Target losses act on the shared encoder parameters η.	Target n updates only its encoder ϕ_n and auxiliary head; the other six encoders are separate.
Encoder parameters	$P_{\text{base}} = \lvert\eta\rvert$, belonging to one encoder.	$P_{\text{NVCode}} = \sum_{n=1}^{7}\lvert\phi_n\rvert$, distributed across seven independently fitted encoders.

Here $|\cdot|$ counts scalar encoder parameters and excludes temporary readout heads, the conditional projector, and the generation network. The comparison matches the 1400-dimensional representation interface; encoder parameter allocation follows the configurations listed above.

B.3 Shared Conditioning Module

The shared conditioning module is the input component of the complete generator. It maps the concatenated $Nd = 7 \times 200 = 1400$ regional features to a $d_c = 512$ neurovascular code, implementing Eq. (3) with two projection layers:

$$\mathbf{h}_\psi(\mathbf{x}) = \text{SiLU}(\text{LN}_1(\mathbf{W}_1\mathbf{e}(\mathbf{x}) + \mathbf{b}_1)),$$
$$\mathbf{c}_\psi(\mathbf{x}) = \text{SiLU}(\text{LN}_2(\mathbf{W}_2\mathbf{h}_\psi(\mathbf{x}) + \mathbf{b}_2)).$$

Here $\mathbf{W}_1 \in \mathbb{R}^{d_c \times Nd}$, $\mathbf{W}_2 \in \mathbb{R}^{d_c \times d_c}$, and $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^{d_c}$. Both linear layers use biases and PyTorch default initialization.

The first projection makes the aggregation across regional encoders explicit. Partition $\mathbf{W}_1 = [\mathbf{W}_1^{(1)} \ \cdots \ \mathbf{W}_1^{(N)}]$, where $\mathbf{W}_1^{(n)} \in \mathbb{R}^{d_c \times d}$. Then

$$\mathbf{W}_1 \mathbf{e}(\mathbf{x}) + \mathbf{b}_1 = \sum_{n=1}^{N} \mathbf{W}_1^{(n)} \mathbf{e}_n + \mathbf{b}_1. \tag{31}$$

Thus, every first-layer coordinate can aggregate target-supervised evidence from all seven encoders.

Figure 10 schematically illustrates the 512-position layout. Each regional objective trains an EEG feature map; the generation objective trains how their hidden features are combined. For each EEG window, the resulting code is shared by all seven outputs throughout sampling.

Source-conditioned synthesis also appears in MRI contrast translation (Dar et al., 2019; Özbey et al., 2023) and visual reconstruction from measured fMRI (Takagi & Nishimoto, 2023; Scotti et al., 2023). In NVCode, the source is the EEG window: $\mathbf{c}_\psi(\mathbf{x})$ carries its target-supervised features into the velocity network within the complete generator.

B.4 Field Parameterization

The velocity field combines the current seven-signal state $\mathbf{z}_t$, interpolation time t, and shared EEG code $\mathbf{c}_\psi(\mathbf{x})$. Its parameters θ include the following state mixer and the subsequent MLP:

$$\mathbf{r}_t = \mathbf{A}\mathbf{z}_t + \mathbf{b}_A, \qquad \mathbf{A} \in \mathbb{R}^{N \times N}, \quad \mathbf{b}_A \in \mathbb{R}^N.$$

The state mixer exposes a learned combination of the current fMRI coordinates to each output velocity.

The time value is encoded by a sinusoidal embedding $\boldsymbol{\tau}(t) \in \mathbb{R}^{d_t}$ with $d_t = 32$. For $j = 0, \ldots, d_t/2 - 1$,

$$\tau_{2j}(t) = \sin(\omega_j t), \qquad \tau_{2j+1}(t) = \cos(\omega_j t), \qquad \omega_j = 10000^{-2j/d_t}.$$

The velocity field is then

$$v_\theta(\mathbf{z}_t, t, \mathbf{c}_\psi(\mathbf{x})) = \mathrm{MLP}_\theta([\mathbf{c}_\psi(\mathbf{x}); \mathbf{r}_t; \boldsymbol{\tau}(t)]) \,.$$

Each output velocity receives the complete evolving state through $\mathbf{r}_t$. The factorized-interpolant ablation in Appendix A.4 tests this joint state access: its scalar fields receive their respective current coordinates and the same shared EEG condition.

Writing $\mathbf{J}_{\mathrm{MLP},r}$ for the Jacobian of the MLP output with respect to its state-input block gives

$$\frac{\partial v_\theta}{\partial \mathbf{z}_t} = \mathbf{J}_{\mathrm{MLP},r} \mathbf{A}. \tag{32}$$

The learned factors permit cross-coordinate responses within each joint velocity evaluation. Shared conditioning supplies the EEG evidence, while access to the evolving joint state supplies the cross-target interactions tested by the independent-interpolant control. The MLP has widths $(d_c{+}N{+}d_t) \to h \to h/2 \to h/4 \to N$, with $h = 512$. Its module sequence is:

1. $\mathrm{Linear}(d_c + N + d_t, h)$, followed by LayerNorm and SiLU;
2. $\mathrm{Linear}(h, h/2)$, followed by LayerNorm and SiLU;
3. $\mathrm{Linear}(h/2, h/4)$, followed by SiLU;
4. $\mathrm{Linear}(h/4, N)$, with no activation.

All linear layers, including the output layer, use biases and PyTorch default Kaiming uniform initialization.

B.5 Optimization Details

NVCode first optimizes the regional representation objectives, then trains its complete generator by fitting the projector and joint field with the selected encoders fixed. The two training objectives are

$$\min_{\phi_n, \mathbf{a}_n, b_n} \quad \mathcal{L}_{\text{rep}}^{(n)}(\phi_n, \mathbf{a}_n, b_n), \qquad n = 1, \ldots, N, \tag{33}$$

$$\min_{\psi, \theta} \quad \mathcal{L}_{\text{NVCode}}\big(\psi, \theta; \phi_1^\star, \ldots, \phi_N^\star\big). \tag{34}$$

Each regional encoder and its temporary scalar head are fitted independently with AdamW. Minimum validation MSE selects the checkpoint $\phi_n^\star$; its head is discarded and the encoder is frozen for generation.

With the selected encoders fixed, AdamW and OneCycleLR jointly optimize Π_ψ and v_θ. Training-only statistics define each endpoint $\mathbf{z}_1$, and each example receives independently sampled $\mathbf{z}_0$ and t. Training follows the split-specific epoch and stopping settings in Table 8; validation selection retains the projector and field together as one matched checkpoint.

In the encoder comparison, each configuration supplies 1400 frozen features, either from seven regional encoders or from one shared encoder. A separate projector and velocity field are jointly fitted for each row using the same generator architecture. In the generator comparison, NVCode retains its jointly selected projector–field pair. Each baseline reuses that frozen projector and fits its generation network on the resulting stored conditions. The trained parameters are therefore (ψ, θ) for NVCode and θ_j for baseline j, with the shared $\psi^\star$ retained inside every complete generator. Appendix A.2 specifies the matched controls.

Hyperparameters. Table 8 gives the training and sampling settings.

Table 8: Key NVCode hyperparameters.

Hyperparameter	**Value**
Regional representation dimension d	200
Spatiotemporal patch length / patches	$w = 200$ / $P = 16$
Representation Transformer depth / heads / MLP ratio	12 / 10 / 4
Spectral branch FFT lengths	$\{100, 200, 400, 800\}$
Spectral branch temporal projection size	32
Shared context dimension d_c	512
Velocity MLP hidden width h / time embedding d_t	512 / 32
Representation fitting optimizer	AdamW
Representation fitting learning rate / weight decay / batch / epochs	5×10^{-4} / 0.05 / 64 / 100
Interpolant fitting optimizer	AdamW + OneCycleLR
Interpolant fitting peak learning rate / weight decay / batch	3×10^{-4} / 10^{-4} / 128
Interpolant fitting max epochs / early stopping	300 / 30 for within subject; fixed 300 for cross subject
Gradient norm clip / precision	1.0 / fp32
Random seeds	42 / 2026 / 114514
Euler steps S / step size Δt	100 / $1/S$
Monte Carlo samples K	50 test / 10 validation

B.6 Generation Details

For each evaluation EEG window, NVCode computes the shared code once and draws independent Gaussian source vectors. With $t_q = q/S$ and $\Delta t = 1/S$, left-endpoint Euler integration updates all seven signals jointly:

$$\mathbf{s}_{q+1}^{(k)} = \mathbf{s}_q^{(k)} + \frac{1}{S} v_{\theta^\star}\left(\mathbf{s}_q^{(k)}, t_q, \mathbf{c}_{\psi^\star}(\mathbf{x}^\star)\right), \quad \mathbf{s}_0^{(k)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N), \quad q = 0, \ldots, S-1.$$

Here $\mathbf{s}_q^{(k)}$ follows the learned field, while $\mathbf{z}_t$ in Eq. (4) defines its training path. We use $S = 100$ steps and $K = 50$ source draws per test window.

Proposition 3 (Sampling-error decomposition). *Fix a condition* $\mathbf{c}$ *and write* $f_\theta(\mathbf{z}, t; \mathbf{c}) = v_{\theta^\star}(\mathbf{z}, t, \mathbf{c})$. *Assume* f_θ *is* L*-Lipschitz in* $\mathbf{z}$*, uniformly over* t*, and the exact learned-field trajectory* $\mathbf{s}_\theta(t)$ *satisfies* $\|\ddot{\mathbf{s}}_\theta(t)\|_2 \le B_\theta$. *Let* $\mathbf{z}^\star(t)$ *be the population ODE trajectory from Eq. (24), initialized from the same source sample. Then the* S*-step Euler endpoint obeys*

$$\|\mathbf{s}_S - \mathbf{z}^\star(1)\|_2 \le \frac{B_\theta}{2LS}\left(e^L - 1\right) + e^L \int_0^1 \|f_\theta(\mathbf{z}^\star(t), t; \mathbf{c}) - v^\star(\mathbf{z}^\star(t), t, \mathbf{c})\|_2 \, dt, \tag{35}$$

with the first term interpreted as $B_\theta/(2S)$ *when* $L = 0$.

Proof. Let $h = 1/S$, $t_q = qh$, and $e_q = \|\mathbf{s}_\theta(t_q) - \mathbf{s}_q\|_2$. Taylor's theorem and the derivative bound give a local remainder of at most $B_\theta h^2/2$. Lipschitz continuity then yields

$$e_{q+1} \le (1 + Lh)e_q + \frac{B_\theta h^2}{2}, \qquad e_0 = 0. \tag{36}$$

Summing the geometric recursion gives

$$e_S \le \frac{B_\theta h}{2L}\left((1 + Lh)^S - 1\right) \le \frac{B_\theta}{2LS}\left(e^L - 1\right). \tag{37}$$

Next set $d(t) = \|\mathbf{s}_\theta(t) - \mathbf{z}^\star(t)\|_2$ and $\epsilon(t) = \|f_\theta(\mathbf{z}^\star(t), t; \mathbf{c}) - v^\star(\mathbf{z}^\star(t), t, \mathbf{c})\|_2$. The integral forms of the two ODEs imply

$$d(t) \le L \int_0^t d(s)\, ds + \int_0^t \epsilon(s)\, ds. \tag{38}$$

Grönwall's inequality gives $d(1) \le e^L \int_0^1 \epsilon(t)\, dt$. The triangle inequality combines this field-approximation term with Eq. (37), proving Eq. (35). □

Under these assumptions, Proposition 3 separates the $O(S^{-1})$ Euler discretization term from the trajectory-integrated field-approximation error.

Each terminal state is mapped to the robust-normalized fMRI supervision scale by reversing the generator z-score:

$$\mathbf{y}^{(k)} = \boldsymbol{\sigma} \odot \mathbf{s}_S^{(k)} + \boldsymbol{\mu}.$$

For one EEG window, the K-sample predictive mean is

$$\bar{\mathbf{y}}_K = \widehat{\mathbf{y}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{y}^{(k)}.$$

Independent sources make this average unbiased for the finite-S model-sampler mean, with Monte Carlo covariance decreasing as $1/K$:

$$\mathbb{E}[\bar{\mathbf{y}}_K \mid \mathbf{x}^\star] = \mathbb{E}[\mathbf{y}^{(1)} \mid \mathbf{x}^\star] =: \mathbf{m}_{\theta,S}(\mathbf{x}^\star), \qquad \mathrm{Cov}(\bar{\mathbf{y}}_K \mid \mathbf{x}^\star) = \frac{1}{K}\mathrm{Cov}(\mathbf{y}^{(1)} \mid \mathbf{x}^\star). \tag{39}$$

Equation (39) follows from linearity of expectation and independence of the source draws. For M EEG windows, the sample tensor is $\mathbf{Y}^{\mathrm{samp}} \in \mathbb{R}^{K \times M \times N}$ and the stacked means form $\widehat{\mathbf{Y}} \in \mathbb{R}^{M \times N}$. Each window receives independent Gaussian sources, while its seven outputs evolve jointly. Temporal correspondence and interregional organization are evaluated across windows using $\widehat{\mathbf{Y}}$ through Pearson correlation and predictive-mean FC-MAE. Mean marginal CRPS evaluates the per-target sample distributions. This distinguishes same-window vector generation from the temporal organization of the predictive-mean trajectories.

C ADDITIONAL RESULTS

This appendix expands the controlled comparison from aggregate rankings to dataset- and scan-level evidence. Table 9 summarizes all seven target signals and three evaluation metrics under matched encoder and generator controls. Tables 10 and 11 resolve those comparisons over all eight datasets, pairing intra-subject and inter-subject blocks. These per-dataset tables report signalwise correlations as mean ± standard deviation across three runs with random seeds 42, 2026, and 114514. Each run uses the same data splits and the aggregation in Appendix A.2; the standard deviations describe variability across the three runs. XP1 and NatView retain mean-only entries in both evaluation settings. Figures 11–17 provide target-specific paired EEG–fMRI examples.

Table 9 summarizes two controlled comparisons. Its encoder block compares seven regional encoders with single-encoder baselines under the same generator architecture, fitted separately for each row. Its generator block compares generation networks using the same seven frozen encoders and NVCode-trained projector. The seven signal columns span primary-sensory, high-level-cognitive, subcortical, and global activity. Avg. R measures predictive-mean temporal correspondence, CRPS evaluates the seven marginal predictive distributions, and FC-MAE compares predictive-mean interregional correlation matrices. Higher Avg. R and lower CRPS and FC-MAE are favorable.

C.1 MAIN EEG TO FMRI SYNTHESIS RESULTS

Table 9: **Combined main EEG to fMRI synthesis results with different encoders and generators.** Best values are **darker purple, bold**; second-best values are light purple, underlined.

	Encoder	Primary Sensory		High level Cognitive		Subcortical		Global Signal	Avg. R↑	CRPS↓	FC-MAE↓
		Cuneus	Heschl's Gyrus	Middle Frontal	Precuneus	Putamen	Thalamus				
Intra subject	**BEIRA**	0.302	0.372	0.257	0.372	0.384	0.344	0.392	0.346	0.393	0.327
	BIOT	0.275	0.218	0.299	0.239	0.280	0.245	0.278	0.262	0.363	0.328
	BrainOmni	0.240	0.386	0.297	0.308	0.258	0.256	0.196	0.277	0.360	0.331
	CNN-Trans.	0.247	0.415	0.264	0.309	0.362	0.314	0.332	0.320	0.363	0.332
	ContraWR	0.311	0.370	0.351	0.314	0.334	0.264	0.356	0.329	0.390	0.348
	FFCL	0.342	0.264	0.361	0.288	0.378	0.292	0.278	0.315	0.389	0.351
	LaBraM	0.280	0.201	0.280	0.282	0.303	0.278	0.231	0.265	0.360	0.319
	SIREN-based	0.443	0.406	**0.407**	0.321	0.438	0.404	0.400	0.403	0.383	0.315
	NeuroBOLT	0.327	0.380	0.330	0.387	0.349	0.412	0.352	0.362	**0.351**	0.331
	REVE	0.116	0.302	0.214	0.239	0.242	0.155	0.271	0.220	0.365	0.340
	SPaRCNet	0.334	0.341	0.243	0.257	0.359	0.367	0.410	0.330	0.379	0.330
	STT-Trans.	0.276	0.296	0.268	0.257	0.227	0.269	0.216	0.258	0.381	0.323
	NVCode (Ours)	**0.449**	**0.503**	0.400	**0.506**	**0.500**	**0.477**	**0.524**	**0.480**	0.352	**0.280**
Inter subject	**BEIRA**	0.250	0.180	0.128	0.177	0.118	0.102	0.148	0.158	0.369	0.204
	BIOT	0.232	0.204	0.215	0.324	0.231	0.281	0.236	0.246	0.348	0.172
	BrainOmni	0.216	0.304	0.257	0.308	0.316	0.293	0.227	0.274	0.347	0.168
	CNN-Trans.	0.309	0.299	0.186	0.242	0.334	0.235	0.144	0.250	0.347	0.190
	ContraWR	0.192	0.128	0.355	0.333	0.240	0.230	0.189	0.238	0.358	0.221
	FFCL	0.287	0.250	0.320	0.293	0.203	0.267	0.225	0.264	0.363	0.197
	LaBraM	0.304	0.346	0.289	0.332	0.290	0.286	0.280	0.304	0.345	0.187
	SIREN-based	0.346	0.347	0.215	0.338	0.266	0.307	0.289	0.301	0.354	0.189
	NeuroBOLT	0.233	0.350	0.252	0.240	0.292	**0.411**	0.367	0.306	0.345	0.183
	REVE	0.218	0.203	0.162	0.177	0.165	0.110	0.182	0.174	0.351	0.200
	SPaRCNet	**0.397**	0.362	0.365	0.309	0.370	0.345	0.324	0.353	0.344	0.163
	STT-Trans.	0.209	0.195	0.241	0.141	0.198	0.199	0.158	0.192	0.362	0.163
	NVCode (Ours)	0.392	**0.475**	**0.434**	**0.370**	**0.405**	0.396	**0.487**	**0.423**	**0.344**	**0.135**
	Generator	**Primary Sensory**		**High level Cognitive**		**Subcortical**		**Global Signal**	**Avg. R↑**	**CRPS↓**	**FC-MAE↓**
		Cuneus	**Heschl's Gyrus**	**Middle Frontal**	**Precuneus**	**Putamen**	**Thalamus**				
Intra subject	**cGAN**	0.232	0.312	0.297	0.308	0.389	0.301	0.335	0.310	0.462	0.505
	Consistency	0.294	0.315	0.362	0.333	0.318	0.314	0.360	0.328	0.452	0.357
	CVAE	0.351	0.409	**0.439**	0.411	0.435	0.344	0.441	0.404	0.378	0.336
	DDPM	0.285	0.243	0.345	0.286	0.302	0.280	0.375	0.302	0.358	0.309
	EDM	0.329	0.331	0.352	0.323	0.383	0.378	0.363	0.351	0.361	0.323
	Rectified Flow	0.338	0.288	0.282	0.356	0.316	0.311	0.223	0.302	0.364	0.328
	Score SDE	0.300	0.348	0.372	0.345	0.394	0.343	0.296	0.343	0.361	0.346
	NVCode (Ours)	**0.449**	**0.503**	0.400	**0.506**	**0.500**	**0.477**	**0.524**	**0.480**	**0.352**	**0.280**
Inter subject	**cGAN**	0.175	0.309	0.119	0.171	0.265	0.224	0.216	0.211	0.368	0.297
	Consistency	0.328	0.178	0.297	0.318	0.294	0.313	0.254	0.283	0.431	0.240
	CVAE	0.231	0.309	0.321	0.200	0.277	0.294	0.324	0.280	0.345	0.244
	DDPM	0.330	0.221	0.152	0.182	0.168	0.160	0.164	0.197	0.346	0.170
	EDM	0.211	0.340	0.266	0.216	0.360	0.252	0.354	0.286	0.343	0.211
	Rectified Flow	0.194	0.248	0.207	0.154	0.297	0.313	0.223	0.234	0.350	0.242
	Score SDE	0.203	0.329	0.268	0.220	0.334	0.309	0.295	0.280	**0.339**	0.202
	NVCode (Ours)	**0.392**	**0.475**	**0.434**	**0.370**	**0.405**	**0.396**	**0.487**	**0.423**	0.344	**0.135**

C.2 PER DATASET ENCODER RESULTS

Table 10: **Per-dataset EEG encoder comparison across all eight datasets.** Seven regional encoders and single-encoder baselines use the same generator architecture, fitted separately for each row. Signal columns report mean ± standard deviation across three random seeds (42, 2026, 114514); XP1 and NatView show means only in both intra-subject and inter-subject settings. Avg. R, CRPS, and FC-MAE report means across the same three runs. Best values are **darker purple, bold**; second-best values are light purple, underlined.

	Encoder	Primary Sensory		High level Cognitive		Subcortical		Global Signal	Avg. R↑	CRPS↓	FC-MAE↓
		Cuneus	Heschl's Gyrus	Middle Frontal	Precuneus	Putamen	Thalamus				
Sleep EEG–fMRI (Resting State)											
Intra subject	**BEIRA**	0.311±0.161	0.263±0.183	**0.525±0.186**	0.372±0.231	0.378±0.181	0.443±0.175	0.302±0.202	0.371	0.393	0.247
	BIOT	0.239±0.109	0.318±0.105	0.265±0.107	0.217±0.111	0.301±0.103	0.292±0.126	0.252±0.116	0.269	0.342	0.276
	BrainOmni	0.276±0.098	0.319±0.158	0.301±0.110	0.307±0.143	0.306±0.122	0.262±0.114	0.260±0.124	0.290	0.347	0.255
	CNN-Trans.	0.351±0.134	0.319±0.179	0.419±0.177	0.387±0.183	0.309±0.161	0.293±0.126	0.336±0.185	0.345	**0.335**	0.245
	ContraWR	0.351±0.195	0.308±0.175	0.442±0.187	0.297±0.239	0.396±0.195	0.315±0.207	0.392±0.203	0.357	0.361	0.281
	FFCL	0.324±0.200	0.418±0.157	0.487±0.159	0.361±0.221	0.487±0.180	0.405±0.195	0.340±0.193	0.403	0.367	0.290
	LaBraM	0.221±0.114	0.181±0.121	0.228±0.126	0.236±0.112	0.282±0.088	0.271±0.128	0.223±0.127	0.235	0.344	0.280
	SIREN-based	0.268±0.196	0.423±0.172	0.461±0.133	0.475±0.198	0.384±0.190	0.303±0.159	0.368±0.153	0.383	0.358	0.259
	NeuroBOLT	0.355±0.112	0.382±0.151	0.271±0.150	0.346±0.110	0.338±0.111	0.300±0.137	0.343±0.128	0.334	0.341	0.272
	REVE	0.275±0.102	0.380±0.119	0.233±0.133	0.299±0.146	0.342±0.155	0.334±0.108	0.304±0.123	0.309	0.354	0.309
	SPaRCNet	**0.384±0.188**	0.436±0.179	0.381±0.173	0.459±0.198	0.433±0.151	0.366±0.160	0.353±0.164	0.402	0.348	0.277
	STT-Trans.	0.253±0.121	0.414±0.160	0.387±0.120	0.335±0.111	0.296±0.130	0.239±0.171	0.363±0.145	0.327	0.348	0.262
	NVCode (Ours)	0.366±0.175	**0.522±0.186**	0.472±0.143	**0.524±0.135**	**0.608±0.118**	**0.491±0.131**	**0.534±0.145**	**0.502**	0.339	**0.239**
Inter subject	**BEIRA**	0.140±0.058	0.290±0.233	0.248±0.125	0.405±0.080	0.283±0.203	0.170±0.068	0.248±0.125	0.255	0.332	0.126
	BIOT	0.235±0.073	0.320±0.127	0.334±0.104	0.299±0.185	0.324±0.119	0.302±0.182	0.343±0.092	0.308	0.331	0.100
	BrainOmni	0.201±0.108	0.219±0.125	0.207±0.113	0.272±0.221	0.332±0.208	0.332±0.209	0.064±0.054	0.232	0.338	0.093
	CNN-Trans.	0.212±0.121	0.291±0.315	0.227±0.137	0.183±0.099	0.267±0.211	0.152±0.082	0.185±0.100	0.217	0.336	0.097
	ContraWR	0.153±0.115	0.292±0.400	0.489±0.094	**0.543±0.076**	**0.573±0.069**	0.213±0.170	**0.677±0.052**	0.420	0.332	0.208
	FFCL	0.221±0.204	0.191±0.163	**0.582±0.075**	0.475±0.113	0.395±0.179	0.266±0.327	0.625±0.066	0.394	0.336	0.098
	LaBraM	0.310±0.300	0.452±0.078	0.249±0.165	0.358±0.152	0.492±0.065	0.282±0.258	0.378±0.127	0.360	0.333	0.129
	SIREN-based	**0.328±0.366**	0.539±0.089	0.459±0.125	0.529±0.093	0.437±0.141	**0.451±0.131**	0.401±0.177	0.449	0.329	0.158
	NeuroBOLT	0.119±0.052	0.413±0.075	0.342±0.139	0.246±0.121	0.287±0.217	0.278±0.182	0.323±0.181	0.287	0.335	0.106
	REVE	0.239±0.172	0.139±0.086	0.228±0.154	0.078±0.066	0.262±0.223	0.180±0.108	0.126±0.081	0.179	0.338	0.132
	SPaRCNet	0.250±0.266	0.536±0.086	0.449±0.126	0.362±0.233	0.565±0.078	0.415±0.153	0.394±0.177	0.425	0.328	**0.092**
	STT-Trans.	0.147±0.133	0.151±0.136	0.126±0.120	0.105±0.110	0.027±0.082	0.207±0.193	0.079±0.098	0.120	0.347	0.107
	NVCode (Ours)	0.194±0.207	**0.613±0.142**	0.553±0.178	0.410±0.400	0.557±0.175	0.433±0.362	0.430±0.373	**0.456**	**0.327**	0.096
MULTI-CLARID (Resting State)											
Intra subject	**BEIRA**	0.408±0.127	0.466±0.131	0.470±0.126	**0.544±0.112**	0.523±0.131	0.608±0.113	0.464±0.114	0.498	0.442	0.316
	BIOT	0.452±0.089	0.406±0.096	0.360±0.102	0.377±0.103	0.414±0.106	0.377±0.085	0.492±0.100	0.411	0.406	0.312
	BrainOmni	0.428±0.087	0.441±0.091	0.466±0.090	0.371±0.086	0.418±0.100	0.424±0.093	0.423±0.096	0.424	0.395	0.308
	CNN-Trans.	0.415±0.126	0.472±0.117	0.533±0.138	0.351±0.082	0.552±0.114	0.494±0.119	0.439±0.115	0.465	0.421	0.346
	ContraWR	0.409±0.102	0.428±0.120	0.591±0.103	0.367±0.109	0.540±0.106	0.512±0.097	0.409±0.094	0.465	0.437	0.335
	FFCL	0.438±0.094	0.474±0.110	0.460±0.108	0.457±0.086	0.430±0.113	0.468±0.119	0.440±0.120	0.453	0.448	0.325
	LaBraM	0.492±0.072	0.465±0.108	0.432±0.090	0.426±0.090	0.507±0.107	0.478±0.113	0.410±0.084	0.458	0.400	0.295
	SIREN-based	0.506±0.099	0.478±0.141	0.578±0.131	0.404±0.113	**0.651±0.106**	0.638±0.120	0.388±0.106	0.520	0.424	0.310
	NeuroBOLT	0.397±0.104	0.470±0.118	0.431±0.123	0.408±0.094	0.569±0.109	0.469±0.126	0.364±0.091	0.444	0.402	0.332
	REVE	0.330±0.085	0.342±0.070	0.363±0.068	0.337±0.071	0.392±0.064	0.390±0.076	0.395±0.057	0.364	0.408	0.311
	SPaRCNet	0.428±0.096	0.424±0.123	0.442±0.124	0.485±0.109	0.609±0.124	0.618±0.124	**0.564±0.105**	0.510	0.426	0.344
	STT-Trans.	0.422±0.101	0.403±0.109	0.396±0.130	0.383±0.107	0.377±0.105	0.234±0.124	0.326±0.093	0.363	0.446	0.350
	NVCode (Ours)	**0.524±0.100**	**0.484±0.128**	**0.639±0.124**	0.416±0.096	0.606±0.119	**0.650±0.117**	0.482±0.124	**0.543**	**0.384**	**0.276**
Inter subject	**BEIRA**	0.163±0.053	**0.421±0.058**	0.394±0.070	0.311±0.183	0.468±0.045	0.403±0.065	0.314±0.172	0.354	0.394	0.196
	BIOT	0.282±0.113	0.364±0.063	0.381±0.054	0.355±0.069	0.269±0.092	0.302±0.157	0.308±0.137	0.323	0.381	0.098
	BrainOmni	0.256±0.079	0.170±0.039	0.341±0.083	0.297±0.155	0.290±0.132	0.257±0.080	0.290±0.134	0.271	0.385	0.106
	CNN-Trans.	0.219±0.086	0.193±0.071	0.308±0.220	0.356±0.109	0.563±0.034	0.525±0.039	0.325±0.161	0.355	0.377	0.187
	ContraWR	0.162±0.048	0.251±0.097	0.257±0.104	0.344±0.102	0.447±0.046	0.356±0.089	0.313±0.161	0.304	0.401	0.122
	FFCL	0.215±0.058	0.267±0.099	0.372±0.065	0.376±0.063	0.387±0.057	0.289±0.141	0.291±0.149	0.314	0.399	0.144
	LaBraM	0.326±0.100	0.285±0.120	0.356±0.069	0.323±0.106	0.467±0.032	0.399±0.048	0.319±0.113	0.353	0.382	0.179
	SIREN-based	0.237±0.066	0.293±0.146	0.323±0.108	0.313±0.129	0.537±0.025	0.544±0.024	0.282±0.119	0.361	0.380	0.173
	NeuroBOLT	0.374±0.062	0.273±0.106	0.299±0.171	0.353±0.076	0.615±0.020	0.582±0.022	0.317±0.124	0.402	**0.376**	0.121
	REVE	0.302±0.115	0.337±0.063	0.319±0.081	0.306±0.106	0.404±0.034	0.346±0.056	0.308±0.101	0.332	0.382	**0.094**
	SPaRCNet	0.178±0.073	0.153±0.063	0.293±0.250	0.224±0.102	0.610±0.034	0.547±0.041	0.249±0.131	0.322	0.384	0.147
	STT-Trans.	**0.398±0.047**	0.361±0.064	0.408±0.044	0.078±0.025	0.306±0.141	0.353±0.070	0.313±0.122	0.317	0.406	0.137
	NVCode (Ours)	0.324±0.134	0.361±0.197	**0.428±0.229**	**0.395±0.349**	**0.768±0.038**	**0.742±0.041**	**0.436±0.206**	**0.493**	0.380	0.105
gradCPT (Resting State)											
Intra subject	**BEIRA**	**0.476±0.173**	0.316±0.199	0.322±0.191	0.380±0.139	0.431±0.155	0.296±0.183	0.439±0.172	0.380	0.383	0.359
	BIOT	0.345±0.142	0.249±0.148	0.208±0.132	0.289±0.134	0.279±0.124	0.375±0.141	0.315±0.110	0.294	0.332	0.326
	BrainOmni	0.214±0.118	0.346±0.117	0.388±0.146	0.313±0.152	0.333±0.152	0.309±0.126	0.290±0.158	0.313	0.333	0.332
	CNN-Trans.	0.332±0.126	0.363±0.140	0.446±0.129	0.409±0.153	0.464±0.151	0.344±0.149	0.409±0.166	0.395	0.330	0.336
	ContraWR	0.347±0.163	0.447±0.164	0.399±0.144	0.478±0.181	0.545±0.157	0.400±0.168	0.509±0.161	0.447	0.352	0.353
	FFCL	0.344±0.163	0.378±0.165	0.347±0.169	0.446±0.161	0.382±0.151	0.393±0.146	0.365±0.159	0.379	0.359	0.368
	LaBraM	0.417±0.130	0.384±0.161	0.388±0.146	0.245±0.093	0.336±0.152	0.294±0.140	0.325±0.138	0.341	0.333	0.326
	SIREN-based	0.383±0.165	0.446±0.158	0.393±0.142	**0.483±0.137**	0.439±0.148	0.448±0.149	0.441±0.155	0.433	0.329	0.354
	NeuroBOLT	0.330±0.130	0.423±0.156	0.375±0.140	0.391±0.138	0.351±0.151	0.378±0.150	0.455±0.151	0.386	0.329	0.329
	REVE	0.337±0.136	0.410±0.144	0.313±0.135	0.281±0.131	0.331±0.132	0.377±0.135	0.287±0.137	0.334	0.335	0.357
	SPaRCNet	0.311±0.143	0.412±0.150	**0.493±0.168**	0.265±0.116	0.411±0.153	0.361±0.149	0.400±0.145	0.379	0.345	0.372
	STT-Trans.	0.395±0.140	0.411±0.128	0.393±0.126	0.303±0.146	0.308±0.127	0.359±0.154	0.311±0.150	0.354	0.337	0.346
	NVCode (Ours)	0.392±0.200	**0.601±0.181**	0.424±0.144	0.463±0.180	**0.577±0.188**	**0.538±0.174**	**0.609±0.166**	**0.515**	**0.329**	**0.309**
Inter subject	**BEIRA**	0.238±0.133	0.412±0.089	0.221±0.114	0.359±0.137	0.250±0.151	0.256±0.162	0.302±0.324	0.291	0.318	0.149
	BIOT	0.182±0.116	0.183±0.117	0.177±0.113	0.274±0.277	0.267±0.252	0.309±0.376	0.353±0.198	0.249	0.321	0.104
	BrainOmni	0.210±0.124	0.287±0.306	0.204±0.119	0.249±0.177	0.272±0.235	0.362±0.157	0.361±0.160	0.278	0.318	0.098
	CNN-Trans.	0.275±0.169	0.230±0.100	0.245±0.115	0.268±0.154	0.351±0.121	0.301±0.271	0.214±0.087	0.269	0.321	0.064
	ContraWR	0.175±0.106	0.222±0.148	0.431±0.102	0.217±0.142	0.290±0.349	0.174±0.105	0.176±0.106	0.241	0.325	0.210
	FFCL	0.245±0.187	0.268±0.248	0.299±0.400	0.051±0.061	0.058±0.063	0.118±0.080	0.158±0.098	0.171	0.328	0.112
	LaBraM	0.243±0.114	0.266±0.151	0.313±0.209	0.149±0.058	0.301±0.269	0.280±0.184	0.200±0.079	0.250	0.319	0.085
	SIREN-based	0.365±0.342	0.546±0.126	**0.505±0.147**	0.378±0.306	0.469±0.173	0.380±0.301	0.554±0.123	0.457	**0.311**	0.131
	NeuroBOLT	0.389±0.147	0.299±0.400	0.387±0.149	0.182±0.119	0.404±0.131	0.404±0.131	0.419±0.119	0.355	0.320	0.127
	REVE	0.205±0.122	0.227±0.145	0.195±0.113	0.116±0.073	0.189±0.109	0.234±0.155	0.305±0.367	0.210	0.323	0.145
	SPaRCNet	0.485±0.142	**0.579±0.100**	0.355±0.336	0.268±0.400	**0.516±0.125**	**0.419±0.201**	0.498±0.134	0.446	0.313	0.137
	STT-Trans.	0.248±0.158	0.257±0.176	0.222±0.123	**0.463±0.072**	0.258±0.178	0.263±0.190	0.282±0.252	0.285	0.319	0.153
	NVCode (Ours)	**0.576±0.219**	0.560±0.236	0.482±0.388	0.222±0.217	0.509±0.317	0.412±0.400	**0.587±0.208**	**0.478**	0.314	**0.044**

Continued on next page

	Encoder	Primary Sensory		High level Cognitive		Subcortical		Global Signal	Avg. R↑	CRPS↓	FC-MAE↓
		Cuneus	Heschl's Gyrus	Middle Frontal	Precuneus	Putamen	Thalamus				
	NeuroBOLT (Resting State)										
Intra subject	**BEIRA**	0.447±0.105	0.453±0.107	0.356±0.126	0.387±0.102	0.303±0.102	0.322±0.102	0.490±0.113	0.394	0.317	0.229
	BIOT	0.212±0.066	0.200±0.062	0.219±0.063	0.208±0.070	0.144±0.059	0.203±0.069	0.210±0.072	0.199	0.305	0.182
	BrainOmni	0.335±0.088	0.326±0.087	0.273±0.103	0.323±0.098	0.225±0.074	0.208±0.082	0.393±0.100	0.297	0.301	0.249
	CNN-Trans.	0.411±0.108	0.400±0.096	0.326±0.112	0.304±0.104	0.292±0.075	0.274±0.090	0.400±0.110	0.344	0.314	0.224
	ContraWR	0.386±0.106	0.338±0.112	0.256±0.092	0.364±0.103	0.237±0.094	0.218±0.084	0.402±0.109	0.314	0.347	0.209
	FFCL	0.447±0.105	0.377±0.110	0.290±0.103	0.353±0.112	0.266±0.096	0.287±0.093	0.440±0.115	0.351	0.328	0.191
	LaBraM	0.463±0.113	0.452±0.107	0.375±0.101	0.389±0.121	0.343±0.076	0.371±0.091	0.503±0.096	0.414	0.289	0.188
	SIREN-based	0.527±0.096	0.469±0.105	0.427±0.080	0.436±0.109	0.348±0.092	0.435±0.097	0.562±0.093	0.458	0.300	0.188
	NeuroBOLT	0.572±0.101	0.545±0.086	0.348±0.099	0.429±0.108	0.310±0.088	0.371±0.080	0.529±0.094	0.444	0.285	0.185
	REVE	0.174±0.063	0.149±0.062	0.103±0.075	0.161±0.067	0.085±0.065	0.058±0.059	0.149±0.066	0.125	0.310	0.259
	SPaRCNet	0.531±0.099	0.525±0.096	0.379±0.109	0.489±0.109	0.405±0.104	0.457±0.080	0.647±0.081	0.491	0.312	**0.162**
	STT-Trans.	0.104±0.066	0.203±0.070	0.180±0.066	0.136±0.068	0.128±0.070	0.106±0.082	0.178±0.062	0.148	0.326	0.231
	NVCode (Ours)	**0.639±0.113**	**0.671±0.107**	**0.489±0.153**	**0.564±0.130**	**0.439±0.175**	**0.515±0.145**	**0.653±0.110**	**0.567**	**0.219**	0.167
Inter subject	**BEIRA**	0.379±0.052	0.380±0.052	0.282±0.071	0.343±0.058	0.126±0.176	0.142±0.153	0.389±0.051	0.291	0.290	0.183
	BIOT	0.284±0.081	0.303±0.076	0.224±0.105	0.261±0.088	0.178±0.135	0.195±0.122	0.318±0.072	0.252	0.294	0.152
	BrainOmni	0.291±0.067	0.297±0.065	0.228±0.087	0.264±0.074	0.160±0.129	0.204±0.099	0.346±0.056	0.256	0.293	0.109
	CNN-Trans.	0.405±0.077	0.487±0.063	0.360±0.087	0.424±0.073	0.263±0.122	0.304±0.104	0.495±0.062	0.391	0.284	0.098
	ContraWR	0.293±0.076	0.270±0.082	0.301±0.073	0.297±0.074	0.166±0.141	0.209±0.109	0.353±0.062	0.270	0.302	0.102
	FFCL	0.404±0.065	0.394±0.067	0.301±0.089	0.314±0.085	0.231±0.118	0.252±0.108	0.415±0.063	0.330	0.290	0.101
	LaBraM	0.445±0.064	0.447±0.064	0.356±0.082	0.421±0.068	0.200±0.152	0.352±0.082	0.482±0.059	0.386	0.285	0.065
	SIREN-based	0.475±0.076	0.521±0.069	0.414±0.088	0.420±0.087	0.318±0.116	0.401±0.091	0.546±0.066	0.442	0.279	0.112
	NeuroBOLT	0.606±0.032	0.508±0.038	0.361±0.055	0.469±0.042	0.097±0.243	0.251±0.081	0.535±0.036	0.404	0.281	0.181
	REVE	0.154±0.041	0.059±0.133	0.081±0.089	0.138±0.046	0.031±0.079	0.105±0.064	0.168±0.037	0.105	0.303	0.153
	SPaRCNet	0.489±0.064	0.515±0.061	0.435±0.073	0.506±0.062	0.252±0.131	0.438±0.072	0.554±0.057	0.455	0.275	0.104
	STT-Trans.	0.090±0.082	0.099±0.073	0.080±0.096	0.076±0.103	0.093±0.079	0.088±0.085	0.131±0.052	0.094	0.303	0.133
	NVCode (Ours)	**0.616±0.049**	**0.614±0.050**	**0.436±0.070**	**0.523±0.058**	**0.341±0.090**	**0.468±0.065**	**0.585±0.052**	**0.512**	**0.268**	**0.043**
	XP1 (Active)										
Intra subject	**BEIRA**	0.467	0.233	0.216	**0.616**	0.379	0.376	0.502	0.398	0.539	0.390
	BIOT	0.009	-0.054	0.228	0.246	0.086	0.086	0.098	0.100	0.515	0.413
	BrainOmni	0.142	0.347	0.223	0.373	0.339	0.064	0.273	0.252	0.502	0.400
	CNN-Trans.	0.356	0.447	0.208	0.184	0.240	0.212	0.198	0.263	0.493	0.408
	ContraWR	0.128	0.171	0.375	0.121	0.064	0.119	-0.010	0.138	0.547	0.431
	FFCL	0.384	-0.049	0.292	0.227	0.410	0.269	0.232	0.252	0.580	0.446
	LaBraM	0.239	0.004	0.143	0.240	0.277	0.179	0.197	0.183	0.506	0.412
	SIREN-based	0.473	0.418	**0.471**	0.236	0.394	**0.408**	**0.537**	0.420	0.512	0.388
	NeuroBOLT	0.466	0.282	0.253	0.258	0.347	0.278	0.193	0.297	0.493	0.426
	REVE	-0.097	0.133	0.011	0.196	0.218	0.233	0.286	0.140	0.511	0.397
	SPaRCNet	0.364	0.250	0.311	0.067	0.372	0.329	0.188	0.269	0.531	0.419
	STT-Trans.	0.263	0.387	0.226	0.115	0.213	0.221	0.285	0.244	0.512	0.405
	NVCode (Ours)	**0.707**	**0.495**	0.291	0.461	**0.608**	0.328	0.536	**0.489**	**0.432**	**0.361**
Inter subject	**BEIRA**	0.178	-0.076	-0.129	0.024	-0.062	-0.141	0.142	-0.009	0.513	0.269
	BIOT	0.381	0.109	0.236	0.195	0.470	0.371	0.194	0.279	0.483	0.272
	BrainOmni	0.438	0.224	0.225	0.365	0.228	0.348	0.337	0.310	0.490	0.314
	CNN-Trans.	0.254	0.242	0.276	0.208	0.459	0.354	0.070	0.266	0.489	0.348
	ContraWR	0.497	0.136	0.259	0.352	0.196	0.233	0.186	0.265	0.490	0.384
	FFCL	0.485	0.143	0.211	0.353	0.307	0.313	0.125	0.277	0.493	0.356
	LaBraM	0.428	0.387	0.202	0.507	0.280	0.337	0.307	0.350	0.480	0.282
	SIREN-based	0.495	0.276	0.135	0.578	0.406	0.266	0.338	0.356	0.492	0.343
	NeuroBOLT	0.417	0.277	0.170	0.362	**0.559**	**0.464**	0.203	0.350	0.481	0.303
	REVE	0.367	0.233	0.138	0.520	0.232	0.053	0.504	0.292	0.497	0.308
	SPaRCNet	0.587	0.481	0.259	0.465	0.432	0.313	0.435	0.425	0.499	0.318
	STT-Trans.	0.353	0.225	0.198	0.099	0.310	0.100	0.178	0.209	0.489	**0.265**
	NVCode (Ours)	**0.654**	**0.660**	**0.342**	**0.588**	-0.012	0.182	**0.660**	**0.439**	**0.466**	0.310
	Affective Music (Active)										
Intra subject	**BEIRA**	0.279±0.118	0.413±0.137	0.477±0.126	0.317±0.107	0.469±0.103	0.455±0.108	0.273±0.112	0.383	0.357	0.283
	BIOT	0.194±0.083	0.235±0.065	0.252±0.086	0.218±0.050	0.198±0.055	0.251±0.081	0.258±0.074	0.230	0.357	0.261
	BrainOmni	0.276±0.104	0.320±0.085	0.393±0.099	0.293±0.086	0.297±0.067	0.354±0.063	0.261±0.093	0.313	**0.345**	0.248
	CNN-Trans.	0.324±0.083	0.396±0.092	0.419±0.147	0.330±0.064	0.393±0.133	0.449±0.103	0.290±0.088	0.371	0.369	0.239
	ContraWR	0.314±0.123	0.367±0.091	**0.553±0.116**	0.314±0.111	0.410±0.117	0.481±0.105	0.265±0.092	0.386	0.396	0.248
	FFCL	0.309±0.117	0.360±0.094	0.476±0.108	0.255±0.092	0.396±0.122	0.486±0.074	0.278±0.079	0.366	0.405	0.245
	LaBraM	0.250±0.074	0.270±0.097	0.299±0.121	0.172±0.088	0.282±0.089	0.367±0.093	0.210±0.072	0.264	0.360	0.243
	SIREN-based	0.311±0.097	0.463±0.099	0.446±0.088	0.294±0.080	0.450±0.112	0.469±0.110	0.311±0.074	0.392	0.386	0.241
	NeuroBOLT	0.239±0.085	0.378±0.075	0.403±0.073	0.258±0.060	0.336±0.087	0.346±0.117	0.236±0.086	0.314	0.365	0.266
	REVE	0.202±0.049	0.259±0.051	0.245±0.059	0.213±0.047	0.274±0.055	0.289±0.056	0.216±0.055	0.242	0.360	0.246
	SPaRCNet	0.324±0.132	0.426±0.099	0.357±0.140	0.336±0.095	0.376±0.145	0.387±0.112	0.269±0.088	0.353	0.382	0.261
	STT-Trans.	0.207±0.104	0.252±0.077	0.316±0.110	0.225±0.071	0.170±0.094	0.215±0.085	0.204±0.066	0.227	0.423	0.229
	NVCode (Ours)	**0.405±0.084**	**0.542±0.100**	0.496±0.113	**0.442±0.069**	**0.491±0.098**	**0.543±0.120**	**0.381±0.121**	**0.471**	0.395	**0.226**
Inter subject	**BEIRA**	0.328±0.056	0.341±0.052	0.434±0.034	0.337±0.053	0.285±0.075	0.412±0.037	0.219±0.158	0.337	0.344	0.158
	BIOT	0.229±0.056	0.289±0.030	0.190±0.077	0.263±0.037	0.227±0.057	0.196±0.087	0.150±0.043	0.220	0.339	0.160
	BrainOmni	0.253±0.085	0.263±0.076	0.452±0.027	0.439±0.028	0.323±0.048	0.361±0.039	0.283±0.064	0.339	**0.328**	0.157
	CNN-Trans.	0.207±0.079	0.237±0.048	0.183±0.066	0.109±0.029	0.163±0.048	0.186±0.070	0.217±0.065	0.186	0.335	0.112
	ContraWR	0.465±0.042	0.376±0.059	**0.508±0.037**	0.430±0.047	0.342±0.070	0.473±0.041	0.286±0.101	**0.411**	0.345	0.211
	FFCL	0.349±0.072	0.368±0.065	0.485±0.042	**0.474±0.043**	0.335±0.077	**0.491±0.041**	0.339±0.076	0.406	0.344	0.161
	LaBraM	0.219±0.090	0.200±0.130	0.318±0.033	0.360±0.026	0.251±0.058	0.292±0.040	0.280±0.044	0.274	0.344	0.123
	SIREN-based	0.321±0.050	0.353±0.042	0.317±0.052	0.363±0.040	0.194±0.176	0.397±0.034	0.221±0.133	0.309	0.335	0.137
	NeuroBOLT	0.261±0.088	0.358±0.045	0.260±0.089	0.240±0.111	0.283±0.072	0.299±0.064	0.279±0.074	0.283	0.333	**0.110**
	REVE	0.267±0.044	0.266±0.045	0.310±0.032	0.242±0.058	0.141±0.048	0.176±0.074	0.187±0.089	0.227	0.337	0.177
	SPaRCNet	**0.498±0.029**	**0.411±0.040**	0.365±0.048	0.353±0.051	**0.388±0.044**	0.335±0.057	0.198±0.234	0.364	0.334	0.143
	STT-Trans.	0.228±0.066	0.295±0.033	0.159±0.055	0.194±0.096	0.278±0.038	0.243±0.053	0.142±0.045	0.220	0.380	0.131
	NVCode (Ours)	0.406±0.094	0.366±0.130	0.494±0.059	0.451±0.072	0.266±0.187	0.331±0.194	**0.389±0.108**	0.386	0.351	0.167

Continued on next page

	Encoder	Primary Sensory		High level Cognitive		Subcortical		Global Signal	Avg. R↑	CRPS↓	FC-MAE↓
		Cuneus	Heschl's Gyrus	Middle Frontal	Precuneus	Putamen	Thalamus				
Experience Sampling (Active)											
Intra subject	**BEIRA**	0.446±0.110	0.426±0.111	0.360±0.108	**0.467±0.104**	0.367±0.096	0.410±0.099	0.430±0.104	0.415	0.368	0.337
	BIOT	0.409±0.089	0.314±0.077	0.341±0.075	0.433±0.089	0.385±0.082	0.403±0.070	0.371±0.094	0.379	**0.338**	0.292
	BrainOmni	0.331±0.073	0.335±0.059	0.422±0.072	0.375±0.076	0.336±0.068	0.414±0.064	0.320±0.076	0.362	0.343	0.323
	CNN-Trans.	0.436±0.075	0.415±0.085	0.382±0.095	0.314±0.079	0.443±0.089	0.395±0.089	0.393±0.094	0.397	0.339	0.337
	ContraWR	0.450±0.101	0.432±0.089	**0.460±0.098**	0.373±0.098	0.386±0.077	0.379±0.082	0.388±0.099	0.410	0.352	0.340
	FFCL	0.435±0.119	0.416±0.100	0.424±0.113	0.416±0.116	0.382±0.086	0.405±0.098	0.369±0.110	0.407	0.356	0.325
	LaBraM	0.369±0.077	0.333±0.079	0.377±0.074	0.399±0.089	0.373±0.088	0.375±0.077	0.283±0.069	0.359	0.340	0.289
	SIREN-based	**0.469±0.101**	0.490±0.079	0.373±0.096	0.413±0.090	0.430±0.091	0.426±0.092	0.444±0.068	0.435	0.352	0.331
	NeuroBOLT	0.422±0.087	0.467±0.068	0.379±0.079	0.388±0.082	0.392±0.080	0.439±0.084	0.407±0.077	0.413	0.345	0.308
	REVE	0.420±0.089	0.313±0.082	0.290±0.071	0.363±0.075	0.394±0.084	0.379±0.083	0.319±0.065	0.354	0.346	0.322
	SPaRCNet	0.459±0.128	0.322±0.095	0.399±0.100	0.431±0.104	0.419±0.095	**0.468±0.103**	0.412±0.092	0.416	0.388	0.349
	STT-Trans.	0.395±0.075	0.346±0.082	0.425±0.091	0.315±0.084	0.320±0.079	0.377±0.081	0.364±0.087	0.363	0.343	0.300
	NVCode (Ours)	0.452±0.117	**0.527±0.119**	0.306±0.089	0.363±0.108	**0.478±0.107**	0.447±0.106	**0.478±0.113**	**0.436**	0.351	**0.274**
Inter subject	**BEIRA**	0.154±0.040	0.209±0.055	0.115±0.034	0.206±0.054	0.307±0.158	0.331±0.164	0.283±0.108	0.229	0.331	0.100
	BIOT	0.147±0.039	0.303±0.146	0.281±0.106	0.269±0.092	0.262±0.085	0.231±0.065	0.183±0.047	0.239	0.335	0.130
	BrainOmni	**0.347±0.108**	0.246±0.063	0.272±0.082	0.290±0.102	0.306±0.134	0.377±0.074	0.283±0.093	0.303	0.328	0.080
	CNN-Trans.	0.175±0.051	0.345±0.146	0.150±0.045	0.302±0.163	0.203±0.060	0.290±0.136	0.247±0.084	0.245	0.334	0.089
	ContraWR	0.059±0.037	0.209±0.074	0.297±0.178	0.251±0.102	0.262±0.114	0.240±0.093	0.290±0.159	0.230	0.345	0.135
	FFCL	0.170±0.036	0.346±0.103	0.328±0.142	0.278±0.083	0.462±0.037	0.275±0.080	0.434±0.044	0.327	0.339	0.125
	LaBraM	0.298±0.135	0.309±0.164	0.250±0.076	**0.395±0.073**	0.456±0.048	0.348±0.124	0.380±0.085	0.348	**0.326**	0.092
	SIREN-based	0.258±0.098	0.278±0.123	0.290±0.143	0.261±0.101	0.249±0.091	0.246±0.088	0.251±0.092	0.262	0.328	0.087
	NeuroBOLT	0.172±0.048	0.176±0.049	0.259±0.090	0.253±0.085	0.345±0.139	0.317±0.211	0.270±0.101	0.256	0.335	0.096
	REVE	0.156±0.045	0.300±0.154	0.199±0.058	0.260±0.092	0.256±0.089	0.311±0.190	0.206±0.060	0.241	0.334	**0.049**
	SPaRCNet	0.286±0.072	**0.373±0.057**	0.339±0.090	0.321±0.128	0.295±0.081	0.336±0.094	0.277±0.064	0.318	0.329	0.070
	STT-Trans.	0.195±0.053	0.307±0.163	0.195±0.053	0.172±0.046	0.326±0.189	0.261±0.088	0.205±0.056	0.237	0.334	0.090
	NVCode (Ours)	0.235±0.126	0.348±0.232	**0.342±0.256**	0.378±0.160	**0.575±0.053**	**0.423±0.111**	**0.440±0.098**	**0.392**	0.332	0.066
`NatView` (Active)											
Intra subject	**BEIRA**	-0.418	0.406	-0.671	-0.105	0.221	-0.155	0.236	-0.070	0.347	0.456
	BIOT	0.344	0.072	**0.521**	-0.075	**0.436**	-0.026	0.228	0.214	0.307	0.563
	BrainOmni	-0.085	**0.652**	-0.085	0.107	-0.187	0.012	-0.651	-0.034	0.312	0.529
	CNN-Trans.	-0.646	0.506	-0.620	0.190	0.205	0.053	0.187	-0.018	0.301	0.521
	ContraWR	0.100	0.471	-0.269	0.198	0.095	-0.313	0.494	0.111	0.328	0.585
	FFCL	0.053	-0.260	0.111	-0.214	0.267	-0.378	-0.238	-0.094	0.271	0.618
	LaBraM	-0.207	-0.479	-0.003	0.150	0.024	-0.114	-0.301	-0.133	0.309	0.521
	SIREN-based	**0.607**	0.059	0.106	-0.174	0.405	0.106	0.148	0.180	0.405	0.448
	NeuroBOLT	-0.168	0.093	0.183	0.615	0.149	**0.713**	0.288	0.268	**0.250**	0.533
	REVE	-0.712	0.428	0.150	0.062	-0.102	-0.816	0.213	-0.111	0.295	0.517
	SPaRCNet	-0.133	-0.066	-0.815	-0.477	-0.148	-0.049	0.447	-0.177	0.300	0.453
	STT-Trans.	0.168	-0.047	-0.178	0.244	0.004	0.398	-0.308	0.040	0.314	0.462
	NVCode (Ours)	0.103	0.184	0.082	**0.816**	0.194	0.301	**0.518**	**0.314**	0.368	**0.385**
Inter subject	**BEIRA**	0.422	-0.541	-0.538	-0.570	-0.711	-0.760	-0.714	-0.488	0.426	0.453
	BIOT	0.115	-0.238	-0.102	**0.676**	-0.150	0.341	0.039	0.097	0.297	0.358
	BrainOmni	-0.267	**0.722**	0.125	0.290	**0.616**	0.103	-0.150	0.205	0.292	0.387
	CNN-Trans.	**0.726**	0.369	-0.260	0.088	0.403	-0.232	-0.602	0.070	0.300	0.525
	ContraWR	-0.265	-0.734	0.299	0.229	-0.357	-0.057	-0.773	-0.237	0.320	0.398
	FFCL	0.208	0.025	-0.019	0.027	-0.547	0.129	-0.589	-0.109	0.375	0.475
	LaBraM	0.163	0.420	0.272	0.147	-0.127	0.002	-0.107	0.110	**0.288**	0.544
	SIREN-based	0.291	-0.031	-0.727	-0.137	-0.480	-0.227	-0.278	-0.227	0.379	0.372
	NeuroBOLT	-0.473	0.493	-0.063	-0.188	-0.251	**0.692**	**0.591**	0.115	0.298	0.423
	REVE	0.051	0.067	-0.172	-0.242	-0.195	-0.529	-0.346	-0.195	0.294	0.544
	SPaRCNet	0.406	-0.149	0.429	-0.027	-0.095	-0.043	-0.010	0.073	0.289	0.296
	STT-Trans.	0.011	-0.135	**0.543**	-0.060	-0.014	0.073	-0.065	0.050	0.320	0.288
	NVCode (Ours)	0.127	0.274	0.394	-0.003	0.235	0.177	0.369	**0.225**	0.311	**0.249**

C.3 PER DATASET GENERATOR RESULTS

Table 11: **Per-dataset conditional generator comparison across all eight datasets.** Generation networks share the seven frozen NVCode encoders and the frozen NVCode-trained projector. Signal columns report mean ± standard deviation across three random seeds (42, 2026, 114514); XP1 and NatView show means only in both intra-subject and inter-subject settings. Avg. R, CRPS, and FC-MAE report means across the same three runs. Best values are **darker purple, bold**; second-best values are light purple, underlined.

	Generator	Primary Sensory		High level Cognitive		Subcortical		Global Signal	Avg. R↑	CRPS↓	FC-MAE↓
		Cuneus	Heschl's Gyrus	Middle Frontal	Precuneus	Putamen	Thalamus				
	Sleep EEG–fMRI (Resting State)										
Intra subject	**cGAN**	0.276±0.152	0.307±0.113	0.440±0.111	0.330±0.115	0.398±0.103	0.358±0.139	0.304±0.146	0.344	0.459	0.567
	Consistency	0.413±0.108	0.405±0.109	0.471±0.136	0.383±0.110	0.326±0.109	0.379±0.096	0.396±0.154	0.396	0.439	0.287
	CVAE	**0.538±0.126**	0.438±0.112	**0.573±0.149**	0.432±0.149	0.428±0.136	0.325±0.140	0.489±0.142	0.461	0.351	0.276
	DDPM	0.491±0.132	0.496±0.112	0.436±0.134	0.407±0.161	0.424±0.146	0.386±0.127	0.405±0.133	0.435	0.343	0.249
	EDM	0.436±0.154	0.510±0.173	0.550±0.116	**0.524±0.176**	0.472±0.182	0.388±0.174	0.529±0.177	0.487	0.340	0.258
	Rectified Flow	0.467±0.150	0.437±0.131	0.428±0.116	0.438±0.098	0.457±0.106	0.458±0.116	0.363±0.135	0.435	0.346	0.280
	Score SDE	0.455±0.116	**0.554±0.138**	0.422±0.145	0.472±0.138	0.463±0.143	0.408±0.149	0.499±0.120	0.468	0.343	0.299
	NVCode (Ours)	0.366±0.175	0.522±0.186	0.472±0.143	0.524±0.135	**0.608±0.118**	**0.491±0.131**	**0.534±0.145**	**0.502**	**0.339**	**0.239**
Inter subject	**cGAN**	0.400±0.113	0.412±0.110	0.330±0.134	0.219±0.192	0.310±0.142	0.306±0.143	0.384±0.117	0.337	0.338	0.180
	Consistency	**0.425±0.105**	0.450±0.099	0.399±0.111	0.381±0.116	0.383±0.115	0.316±0.137	0.365±0.120	0.388	0.433	0.169
	CVAE	0.267±0.252	**0.642±0.113**	0.470±0.151	**0.436±0.162**	0.433±0.163	**0.462±0.154**	0.437±0.162	0.449	0.334	0.272
	DDPM	0.261±0.157	0.443±0.098	0.357±0.119	0.364±0.117	0.317±0.132	0.340±0.124	0.396±0.108	0.354	0.331	0.128
	EDM	0.220±0.214	0.589±0.088	0.362±0.138	0.312±0.158	0.339±0.146	0.405±0.125	0.352±0.141	0.368	0.335	0.173
	Rectified Flow	0.381±0.139	0.446±0.121	0.446±0.120	0.254±0.199	0.276±0.185	0.349±0.151	**0.475±0.114**	0.375	0.336	0.189
	Score SDE	0.290±0.178	0.594±0.092	0.399±0.133	0.322±0.162	0.371±0.143	0.406±0.132	0.335±0.156	0.388	0.329	0.157
	NVCode (Ours)	0.194±0.207	0.613±0.142	**0.553±0.178**	0.410±0.400	**0.557±0.175**	0.433±0.362	0.430±0.373	**0.456**	**0.327**	**0.096**
	MULTI-CLARID (Resting State)										
Intra subject	**cGAN**	0.374±0.107	0.365±0.125	0.415±0.108	0.414±0.092	0.571±0.101	0.504±0.131	0.382±0.116	0.432	0.519	0.548
	Consistency	0.400±0.087	0.400±0.109	0.414±0.106	0.376±0.099	0.428±0.105	0.447±0.112	0.311±0.101	0.397	0.505	0.386
	CVAE	0.390±0.108	0.408±0.116	0.459±0.131	0.395±0.085	0.566±0.116	0.509±0.131	0.377±0.111	0.444	0.431	0.331
	DDPM	0.341±0.094	0.401±0.102	0.438±0.119	0.331±0.080	0.512±0.119	0.479±0.134	0.366±0.107	0.410	0.397	0.334
	EDM	0.420±0.094	0.405±0.104	0.458±0.107	0.389±0.084	0.506±0.118	0.477±0.122	0.343±0.099	0.428	0.413	0.322
	Rectified Flow	0.405±0.081	0.382±0.113	0.434±0.103	**0.425±0.089**	0.484±0.118	0.469±0.114	0.321±0.097	0.417	0.417	0.359
	Score SDE	0.363±0.090	0.386±0.119	0.425±0.119	0.333±0.091	0.552±0.102	0.472±0.118	0.350±0.104	0.411	0.409	0.327
	NVCode (Ours)	**0.524±0.100**	**0.484±0.128**	**0.639±0.124**	0.416±0.096	**0.606±0.119**	**0.650±0.117**	**0.482±0.124**	**0.543**	**0.384**	**0.276**
Inter subject	**cGAN**	0.316±0.106	**0.378±0.091**	0.461±0.076	0.365±0.094	0.695±0.052	0.654±0.055	0.363±0.094	0.462	0.394	0.263
	Consistency	0.354±0.092	0.270±0.117	**0.466±0.072**	0.385±0.085	0.697±0.049	0.666±0.051	0.360±0.091	0.457	0.472	0.292
	CVAE	0.389±0.092	0.318±0.111	0.447±0.081	0.328±0.108	0.694±0.054	0.627±0.059	0.376±0.095	0.454	0.381	0.191
	DDPM	**0.399±0.104**	0.316±0.128	0.418±0.099	**0.469±0.090**	0.616±0.069	0.538±0.079	0.300±0.134	0.437	0.382	0.101
	EDM	0.339±0.113	0.315±0.121	0.436±0.090	0.390±0.100	0.632±0.064	0.591±0.068	**0.459±0.086**	0.452	**0.374**	**0.089**
	Rectified Flow	0.368±0.078	0.207±0.129	0.415±0.070	0.313±0.090	0.705±0.043	0.629±0.047	0.362±0.079	0.429	0.377	0.135
	Score SDE	0.382±0.094	0.265±0.131	0.331±0.107	0.322±0.110	0.695±0.054	0.614±0.061	0.308±0.114	0.417	0.379	0.203
	NVCode (Ours)	0.324±0.134	0.361±0.197	0.428±0.229	0.395±0.349	**0.768±0.038**	**0.742±0.041**	0.436±0.206	**0.493**	0.380	0.105
	gradCPT (Resting State)										
Intra subject	**cGAN**	0.285±0.163	0.272±0.160	0.398±0.126	0.369±0.142	0.404±0.131	0.377±0.165	0.319±0.142	0.346	0.451	0.533
	Consistency	0.389±0.115	0.385±0.125	0.361±0.097	0.335±0.121	0.356±0.104	0.370±0.129	0.377±0.095	0.367	0.418	0.357
	CVAE	0.394±0.157	0.315±0.166	**0.480±0.124**	0.333±0.169	0.374±0.156	0.228±0.129	0.326±0.138	0.350	0.345	0.348
	DDPM	0.330±0.154	0.300±0.167	0.391±0.143	0.441±0.140	0.425±0.171	0.353±0.142	0.396±0.171	0.377	0.332	0.327
	EDM	0.428±0.131	0.436±0.175	0.383±0.144	0.393±0.136	0.463±0.149	0.423±0.140	0.420±0.151	0.421	0.330	0.324
	Rectified Flow	**0.507±0.118**	0.416±0.148	0.312±0.147	0.344±0.152	0.411±0.156	0.361±0.138	0.425±0.143	0.397	0.333	0.332
	Score SDE	0.255±0.136	0.444±0.138	0.379±0.151	**0.518±0.138**	0.427±0.133	0.417±0.148	0.401±0.152	0.406	0.333	0.359
	NVCode (Ours)	0.392±0.200	**0.601±0.181**	0.424±0.144	0.463±0.180	**0.577±0.188**	**0.538±0.174**	**0.609±0.166**	**0.515**	**0.329**	**0.309**
Inter subject	**cGAN**	0.292±0.228	0.427±0.162	0.284±0.233	**0.403±0.171**	0.408±0.169	0.366±0.186	0.413±0.167	0.370	0.317	0.161
	Consistency	0.325±0.152	0.244±0.195	0.298±0.164	0.342±0.145	0.399±0.126	0.341±0.145	0.425±0.119	0.339	0.410	0.195
	CVAE	0.385±0.153	0.417±0.142	0.355±0.164	0.088±0.400	0.441±0.135	0.350±0.166	0.463±0.129	0.357	0.322	0.188
	DDPM	0.315±0.195	0.437±0.145	0.353±0.176	0.324±0.190	0.458±0.139	0.357±0.174	0.459±0.139	0.386	0.317	0.082
	EDM	0.436±0.144	0.420±0.149	0.257±0.231	0.347±0.177	0.449±0.140	0.321±0.190	0.452±0.140	0.383	0.317	0.131
	Rectified Flow	0.361±0.144	0.326±0.158	0.313±0.164	0.146±0.311	0.321±0.161	0.298±0.172	0.445±0.120	0.316	0.325	0.178
	Score SDE	0.451±0.158	0.409±0.172	0.380±0.184	0.267±0.253	0.433±0.164	0.340±0.204	0.449±0.158	0.390	0.321	0.182
	NVCode (Ours)	**0.576±0.219**	**0.560±0.236**	**0.482±0.388**	0.222±0.217	**0.509±0.317**	**0.412±0.400**	**0.587±0.208**	**0.478**	**0.314**	**0.044**
	NeuroBOLT (Resting State)										
Intra subject	**cGAN**	0.459±0.106	0.418±0.095	0.271±0.083	0.367±0.081	0.213±0.099	0.266±0.089	0.444±0.090	0.348	0.326	0.198
	Consistency	0.370±0.104	0.375±0.100	0.228±0.093	0.277±0.117	0.114±0.091	0.192±0.086	0.317±0.120	0.268	0.378	0.228
	CVAE	0.487±0.113	0.458±0.084	0.312±0.110	0.384±0.109	0.272±0.088	0.264±0.097	0.480±0.098	0.380	0.300	0.185
	DDPM	0.451±0.108	0.407±0.076	0.199±0.096	0.319±0.097	0.211±0.101	0.203±0.090	0.414±0.098	0.315	0.287	0.187
	EDM	0.491±0.099	0.445±0.091	0.301±0.098	0.376±0.111	0.225±0.089	0.258±0.103	0.463±0.102	0.366	0.288	0.199
	Rectified Flow	0.449±0.113	0.433±0.100	0.274±0.093	0.363±0.103	0.215±0.098	0.271±0.082	0.436±0.108	0.349	0.290	0.180
	Score SDE	0.472±0.105	0.450±0.087	0.315±0.085	0.379±0.095	0.235±0.092	0.282±0.083	0.461±0.094	0.371	0.282	0.193
	NVCode (Ours)	**0.639±0.113**	**0.671±0.107**	**0.489±0.153**	**0.564±0.130**	**0.439±0.175**	**0.515±0.145**	**0.653±0.110**	**0.567**	**0.219**	**0.167**
Inter subject	**cGAN**	0.569±0.059	0.515±0.064	0.374±0.086	0.474±0.069	0.257±0.120	0.287±0.109	0.519±0.064	0.428	0.276	0.227
	Consistency	0.567±0.057	0.506±0.064	0.367±0.085	0.467±0.069	0.223±0.132	0.281±0.108	0.525±0.062	0.420	0.359	0.182
	CVAE	0.580±0.055	0.501±0.063	0.340±0.090	0.444±0.071	0.235±0.125	0.255±0.116	0.503±0.063	0.408	0.271	0.237
	DDPM	0.540±0.050	0.482±0.056	0.304±0.085	0.399±0.066	0.165±0.142	0.270±0.094	0.470±0.057	0.376	0.280	0.079
	EDM	0.511±0.060	0.496±0.061	0.363±0.082	0.449±0.067	0.229±0.122	0.284±0.102	0.499±0.061	0.404	0.272	0.125
	Rectified Flow	0.519±0.049	0.473±0.054	0.302±0.081	0.433±0.058	0.153±0.143	0.269±0.090	0.455±0.056	0.372	0.277	0.108
	Score SDE	0.573±0.061	0.519±0.067	0.370±0.092	0.486±0.071	0.259±0.126	0.330±0.102	0.530±0.066	0.438	**0.268**	0.189
	NVCode (Ours)	**0.616±0.049**	**0.614±0.050**	**0.436±0.070**	**0.523±0.058**	**0.341±0.090**	**0.468±0.065**	**0.585±0.052**	**0.512**	0.268	**0.043**

Continued on next page

	Generator	Primary Sensory		High level Cognitive		Subcortical		Global Signal	Avg. R↑	CRPS↓	FC-MAE↓
		Cuneus	Heschl's Gyrus	Middle Frontal	Precuneus	Putamen	Thalamus				
	XP1 (Active)										
Intra subject	**cGAN**	0.312	0.324	0.004	0.081	0.534	0.276	0.168	0.243	0.619	0.603
	Consistency	0.353	0.286	0.364	0.437	0.284	**0.348**	0.177	0.321	0.605	0.382
	CVAE	0.575	0.479	0.432	0.346	0.464	0.305	0.452	0.436	0.508	0.411
	DDPM	0.523	0.365	0.334	0.379	0.532	0.252	0.339	0.389	0.489	0.432
	EDM	0.533	0.515	0.376	0.411	0.472	0.336	0.382	0.432	0.479	0.402
	Rectified Flow	0.480	0.458	0.299	0.362	0.543	0.339	0.325	0.401	0.483	0.395
	Score SDE	0.548	**0.586**	**0.467**	0.250	0.394	0.340	0.433	0.431	0.489	0.431
	NVCode (Ours)	**0.707**	0.495	0.291	**0.461**	**0.608**	0.328	**0.536**	**0.489**	**0.432**	**0.361**
Inter subject	**cGAN**	0.170	-0.101	-0.278	-0.020	**0.456**	**0.481**	-0.124	0.083	0.495	0.436
	Consistency	0.006	-0.046	0.128	0.131	0.188	0.141	-0.122	0.061	0.577	0.422
	CVAE	0.018	-0.130	-0.006	0.070	0.185	0.099	-0.153	0.012	0.484	0.415
	DDPM	0.537	0.157	-0.181	0.189	-0.315	-0.101	0.139	0.061	0.484	**0.258**
	EDM	0.168	0.058	-0.164	0.043	0.303	0.292	-0.078	0.089	0.476	0.374
	Rectified Flow	0.129	-0.160	-0.020	-0.049	0.352	0.261	-0.197	0.045	0.487	0.443
	Score SDE	0.251	-0.088	-0.163	0.131	0.190	0.148	-0.122	0.049	0.475	0.337
	NVCode (Ours)	**0.654**	**0.660**	**0.342**	**0.588**	-0.012	0.182	**0.660**	**0.439**	**0.466**	0.310
	Affective Music (Active)										
Intra subject	**cGAN**	0.329±0.091	0.428±0.091	0.470±0.069	0.356±0.068	0.439±0.097	0.500±0.102	0.301±0.083	0.403	0.418	0.333
	Consistency	0.301±0.091	0.437±0.095	0.447±0.065	0.366±0.077	0.426±0.073	0.385±0.088	0.364±0.075	0.389	0.445	0.303
	CVAE	0.398±0.104	0.475±0.099	0.530±0.083	0.411±0.065	0.475±0.097	0.523±0.108	0.376±0.094	0.455	0.395	0.287
	DDPM	0.336±0.100	0.440±0.076	0.448±0.096	0.370±0.068	0.441±0.065	0.458±0.088	0.350±0.084	0.406	**0.363**	0.273
	EDM	0.390±0.089	0.426±0.080	0.515±0.084	0.388±0.046	0.436±0.077	0.482±0.114	0.349±0.079	0.427	0.374	0.261
	Rectified Flow	0.388±0.080	0.399±0.090	0.479±0.075	0.382±0.083	0.449±0.075	0.492±0.092	0.350±0.076	0.420	0.377	0.265
	Score SDE	0.357±0.089	0.490±0.099	**0.532±0.092**	0.395±0.062	0.470±0.090	0.497±0.104	**0.402±0.082**	0.449	0.364	0.245
	NVCode (Ours)	**0.405±0.084**	**0.542±0.100**	0.496±0.113	**0.442±0.069**	**0.491±0.098**	**0.543±0.120**	0.381±0.121	**0.471**	0.395	**0.226**
Inter subject	**cGAN**	0.365±0.065	0.413±0.058	0.345±0.069	0.322±0.073	0.359±0.066	0.321±0.073	0.293±0.079	0.346	0.371	0.204
	Consistency	0.384±0.070	**0.474±0.058**	0.429±0.064	0.412±0.066	0.331±0.081	**0.366±0.074**	0.349±0.077	0.392	0.405	0.178
	CVAE	0.323±0.065	0.451±0.048	0.304±0.069	0.294±0.071	0.288±0.072	0.365±0.058	0.325±0.065	0.336	**0.332**	**0.115**
	DDPM	0.326±0.074	0.461±0.054	0.456±0.055	0.413±0.060	0.311±0.078	0.354±0.069	0.287±0.083	0.373	0.336	0.137
	EDM	0.351±0.078	0.457±0.061	0.448±0.062	0.435±0.064	**0.371±0.074**	0.290±0.092	**0.394±0.070**	**0.392**	0.335	0.158
	Rectified Flow	0.333±0.056	0.351±0.053	0.263±0.069	0.269±0.067	0.242±0.074	0.278±0.065	0.263±0.068	0.285	0.345	0.136
	Score SDE	0.301±0.073	0.456±0.050	0.327±0.068	0.278±0.078	0.242±0.088	0.304±0.072	0.321±0.069	0.318	0.336	0.147
	NVCode (Ours)	**0.406±0.094**	0.366±0.130	**0.494±0.059**	**0.451±0.072**	0.266±0.187	0.331±0.194	0.389±0.108	0.386	0.351	0.167
	Experience Sampling (Active)										
Intra subject	**cGAN**	0.264±0.073	0.377±0.090	0.286±0.081	0.304±0.086	0.332±0.089	0.334±0.088	0.375±0.087	0.325	0.464	0.502
	Consistency	0.441±0.069	0.386±0.074	0.385±0.078	0.326±0.076	0.383±0.085	0.386±0.089	0.335±0.087	0.378	0.438	0.340
	CVAE	0.327±0.088	0.420±0.093	0.362±0.088	0.314±0.093	0.403±0.080	0.379±0.102	0.351±0.092	0.365	0.370	0.347
	DDPM	0.364±0.073	0.363±0.079	**0.411±0.062**	**0.372±0.092**	0.347±0.082	0.358±0.092	0.385±0.081	0.371	0.341	0.311
	EDM	0.396±0.084	0.359±0.082	0.370±0.078	0.302±0.080	0.318±0.074	0.432±0.088	0.390±0.081	0.367	0.346	0.330
	Rectified Flow	0.305±0.085	0.394±0.088	0.365±0.073	0.368±0.092	0.405±0.067	0.387±0.102	0.354±0.077	0.368	**0.340**	0.357
	Score SDE	0.395±0.078	0.393±0.080	0.325±0.087	0.360±0.087	0.415±0.086	0.408±0.097	0.345±0.080	0.377	0.351	0.349
	NVCode (Ours)	**0.452±0.117**	**0.527±0.119**	0.306±0.089	0.363±0.108	**0.478±0.107**	**0.447±0.106**	**0.478±0.113**	**0.436**	0.351	**0.274**
Inter subject	**cGAN**	-0.071±0.085	-0.046±0.110	0.012±0.180	-0.074±0.083	-0.021±0.153	0.022±0.151	-0.009±0.192	-0.027	0.345	0.214
	Consistency	-0.003±0.131	-0.020±0.093	-0.030±0.080	-0.083±0.045	0.023±0.089	-0.013±0.106	-0.014±0.104	-0.020	0.417	0.119
	CVAE	-0.017±0.122	0.009±0.142	0.016±0.125	-0.043±0.084	-0.007±0.148	0.046±0.081	0.015±0.126	0.003	**0.328**	0.220
	DDPM	0.014±0.122	0.002±0.157	-0.012±0.127	-0.028±0.097	-0.056±0.068	-0.044±0.078	-0.034±0.089	-0.023	0.331	**0.066**
	EDM	-0.213±0.036	0.000±0.227	0.005±0.204	-0.123±0.056	0.078±0.078	-0.066±0.086	0.033±0.125	-0.041	0.336	0.221
	Rectified Flow	-0.031±0.121	-0.059±0.086	-0.025±0.131	-0.097±0.062	-0.007±0.182	0.012±0.163	-0.064±0.082	-0.039	0.344	0.281
	Score SDE	-0.092±0.076	-0.034±0.135	-0.000±0.249	-0.136±0.057	0.038±0.128	0.065±0.095	-0.023±0.157	-0.026	0.329	0.183
	NVCode (Ours)	**0.235±0.126**	**0.348±0.232**	**0.342±0.256**	**0.378±0.160**	**0.575±0.053**	**0.423±0.111**	**0.440±0.098**	**0.392**	0.332	0.066
	`NatView` (Active)										
Intra subject	**cGAN**	-0.445	0.004	0.089	0.240	0.222	-0.207	0.387	0.041	0.441	0.760
	Consistency	-0.313	-0.155	0.226	0.167	0.228	0.008	0.605	0.109	0.388	0.571
	CVAE	-0.300	**0.279**	**0.360**	0.672	**0.502**	0.216	**0.674**	**0.343**	0.323	0.502
	DDPM	-0.555	-0.828	0.105	-0.330	-0.476	-0.250	0.349	-0.284	**0.315**	**0.362**
	EDM	-0.459	-0.449	-0.136	-0.204	0.172	0.231	0.028	-0.117	0.320	0.489
	Rectified Flow	-0.294	-0.616	-0.335	0.163	-0.437	-0.291	-0.793	-0.372	0.327	0.455
	Score SDE	-0.441	-0.523	0.109	0.051	0.199	-0.080	-0.521	-0.172	0.320	0.564
	NVCode (Ours)	**0.103**	0.184	0.082	**0.816**	0.194	**0.301**	0.518	0.314	0.368	0.385
Inter subject	**cGAN**	-0.638	0.474	-0.580	-0.325	-0.339	-0.646	-0.112	-0.309	0.409	0.687
	Consistency	**0.567**	-0.451	0.317	**0.510**	0.105	0.405	0.143	0.228	0.377	0.360
	CVAE	-0.097	0.261	**0.641**	-0.015	-0.050	0.151	0.628	0.217	0.305	0.315
	DDPM	0.248	-0.527	-0.481	-0.674	-0.153	-0.435	-0.705	-0.390	0.307	0.508
	EDM	-0.121	0.382	0.418	-0.123	**0.480**	-0.098	**0.720**	0.237	0.294	0.419
	Rectified Flow	-0.508	0.399	-0.037	-0.038	0.331	**0.409**	0.042	0.086	0.309	0.463
	Score SDE	-0.529	**0.508**	0.498	0.088	0.445	0.269	0.561	**0.263**	**0.273**	**0.214**
	NVCode (Ours)	0.127	0.274	0.394	-0.003	0.235	0.177	0.369	0.225	0.311	0.249

C.4 Scan Level EEG–fMRI Evidence

Figures 11–17 connect paired EEG inputs to fMRI predictions through selected scan examples for all seven targets. Each target is illustrated by one plate containing seven scans. These qualitative examples complement the eight-dataset comparisons above. Columns (a) and (b) describe the aligned EEG input through spectrograms and alpha-band scalp power (8–13 Hz); column (c) compares real and predicted fMRI trajectories. The shaded 16-s EEG window ends at the purple snapshot guide, whose time also defines column (d). The three-dimensional views project the real amplitude, predicted amplitude, and their absolute difference through the same anatomical weights.

The frequency-resolved input views describe EEG rhythms within the broader study of neural oscillations (Buzsáki & Watson, 2012; Cohen, 2017). Temporal profiles complement the aggregate correlation summaries, reflecting the time-varying organization of brain activity studied in resting and task recordings (Vidaurre et al., 2018).

Snapshot amplitudes and normalization. Each row links the EEG window, fMRI trajectory, and anatomical snapshot at the same marked time $t_\star$ in scan s. The numbers beneath the three views in column (d) are scalar signal amplitudes on the archived supervision scale. For these archived examples, let $q_{s,n}(t)$ denote the low-pass-filtered fMRI signal for target n. Its scanwise centering and robust scaling are

$$\bar{q}_{s,n} = \frac{1}{T_s}\sum_{j=1}^{T_s} q_{s,n}(t_j), \qquad y_{s,n}(t) = \frac{q_{s,n}(t) - \bar{q}_{s,n}}{Q_{0.95}\Big(\{|q_{s,n}(t_j) - \bar{q}_{s,n}|\}_{j=1}^{T_s}\Big) + \epsilon}, \tag{40}$$

where T_s is the number of source volumes, $Q_{0.95}$ is the empirical 95th percentile, and $\epsilon = 10^{-8}$. The archived filter is a fifth-order Butterworth low-pass filter at 0.15 Hz, applied forwards and backwards. The displayed trajectories and snapshot values share this archived target scale. Quantitative tables use the split-specific normalization and evaluation protocol in Appendix A.1; the plates provide paired visual comparisons on their stated scale.

Real, predicted, and absolute-difference values. At $t_\star$, the real value is read from the corresponding target trajectory. The NVCode value is the mean of the $K = 50$ archived generated samples, already returned to the same supervision scale:

$$a_{s,n} = y_{s,n}(t_\star), \qquad \widehat{a}_{s,n} = \frac{1}{K}\sum_{k=1}^{K} y_{s,n}^{(k)}(t_\star), \qquad e_{s,n} = \big|a_{s,n} - \widehat{a}_{s,n}\big|. \tag{41}$$

The three printed numbers are $a_{s,n}$, $\widehat{a}_{s,n}$, and $e_{s,n}$, respectively. The first two are signed, dimensionless amplitudes relative to the scan's centered baseline. The third is a nonnegative pointwise amplitude error, with smaller values indicating closer agreement at the marked time. This snapshot statistic complements trajectory-level Pearson correlation; the anatomical mask defines the displayed regional extent.

Anatomical projection and shared color scale. Let Ω_n be the displayed target mask and $W_n(v)$ its nonnegative template weights at location v. Regional weights come from the corresponding DiFuMo component; global-signal weights use anatomical intensity over the whole-brain mask. Both are normalized by their peak within the displayed mask:

$$w_n(v) = \begin{cases} W_n(v) \big/ \max\limits_{u \in \Omega_n} W_n(u), & v \in \Omega_n, \\ 0, & v \notin \Omega_n. \end{cases} \tag{42}$$

The same w_n is applied to both amplitudes and their absolute difference:

$$\begin{aligned} I_{s,n}^{\mathrm{real}}(v) = a_{s,n} w_n(v), \qquad I_{s,n}^{\mathrm{NVCode}}(v) = \widehat{a}_{s,n} w_n(v), \\ D_{s,n}(v) = \big|I_{s,n}^{\mathrm{real}}(v) - I_{s,n}^{\mathrm{NVCode}}(v)\big| = e_{s,n} w_n(v). \end{aligned} \tag{43}$$

Regional projections use nested three-dimensional surfaces at atlas-weight levels 0.1, 0.3, 0.5, 0.7, and 0.9, with transparent MNI152 anatomy indicating location. Global-signal surface colors average normalized anatomical weights sampled 2, 4, and 6 mm inward toward the brain center. All triplets share the same geometry, orthographic view, illumination, and opacity settings; the scale bars denote 20 mm in the viewing plane. One signed color scale, spanning $[-2, 2]$, applies to every plate, with absolute differences occupying its nonnegative half. Each view projects a regional scalar through fixed anatomical weights: the anatomy supplies the spatial pattern, the scalar sets its amplitude, and shading provides depth cues. Together with column (c), these matched projections relate snapshot amplitude agreement to the full predicted trajectory.

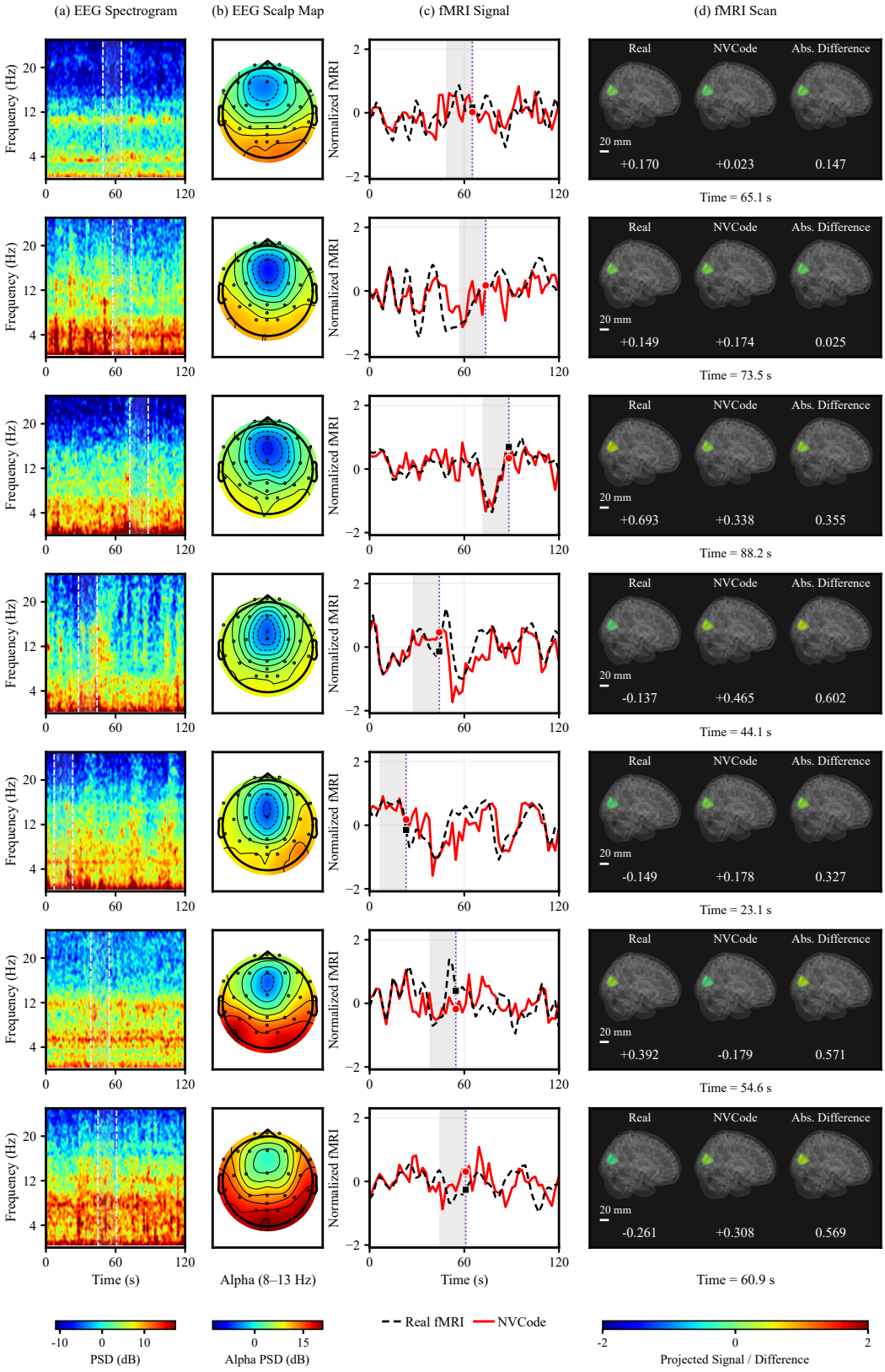


Figure 11: **Cuneus** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (red); (d) projected real, predicted and absolute-difference amplitudes.

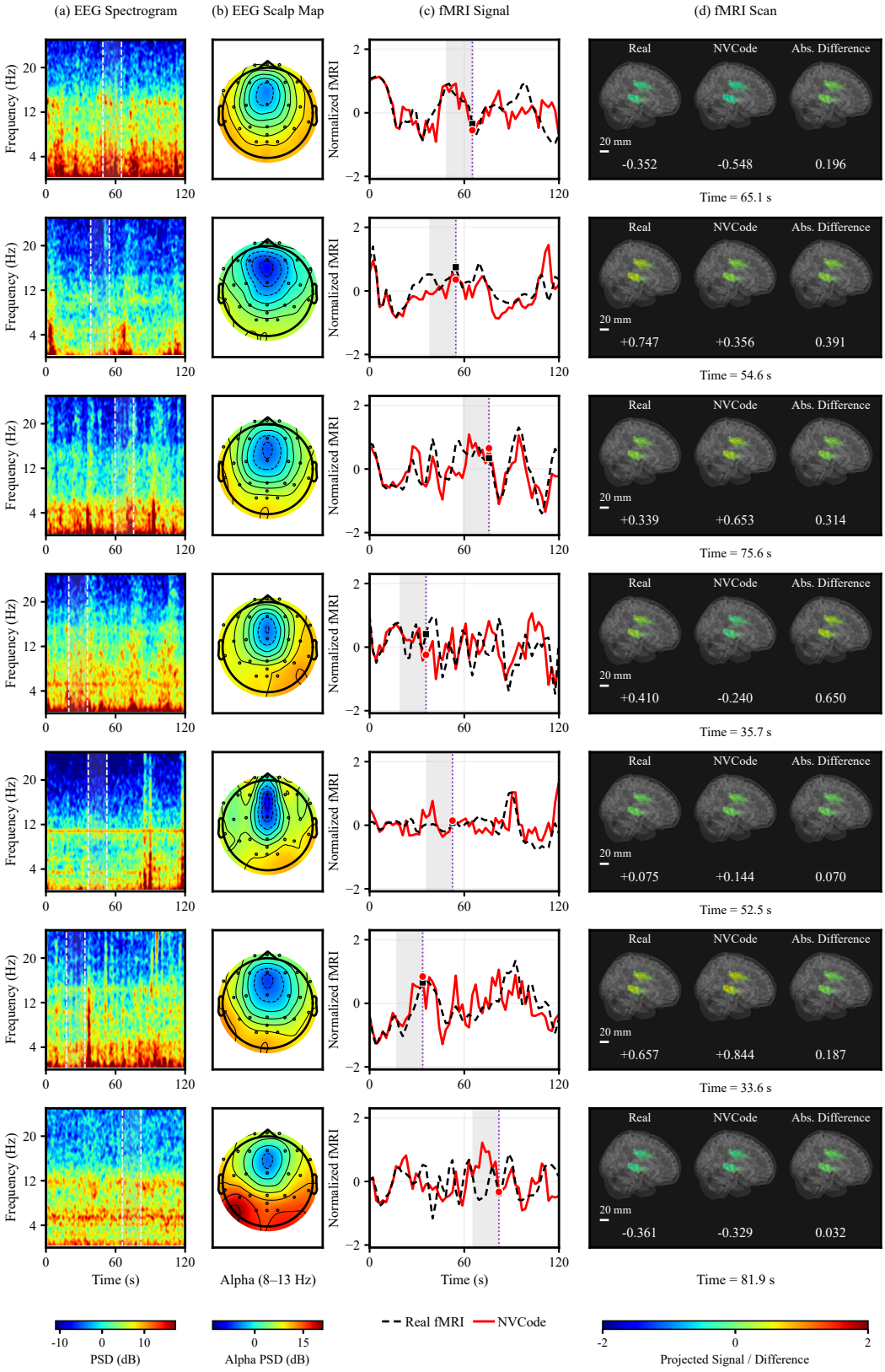


Figure 12: **Heschl's gyrus** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (red); (d) projected real, predicted and absolute-difference amplitudes.

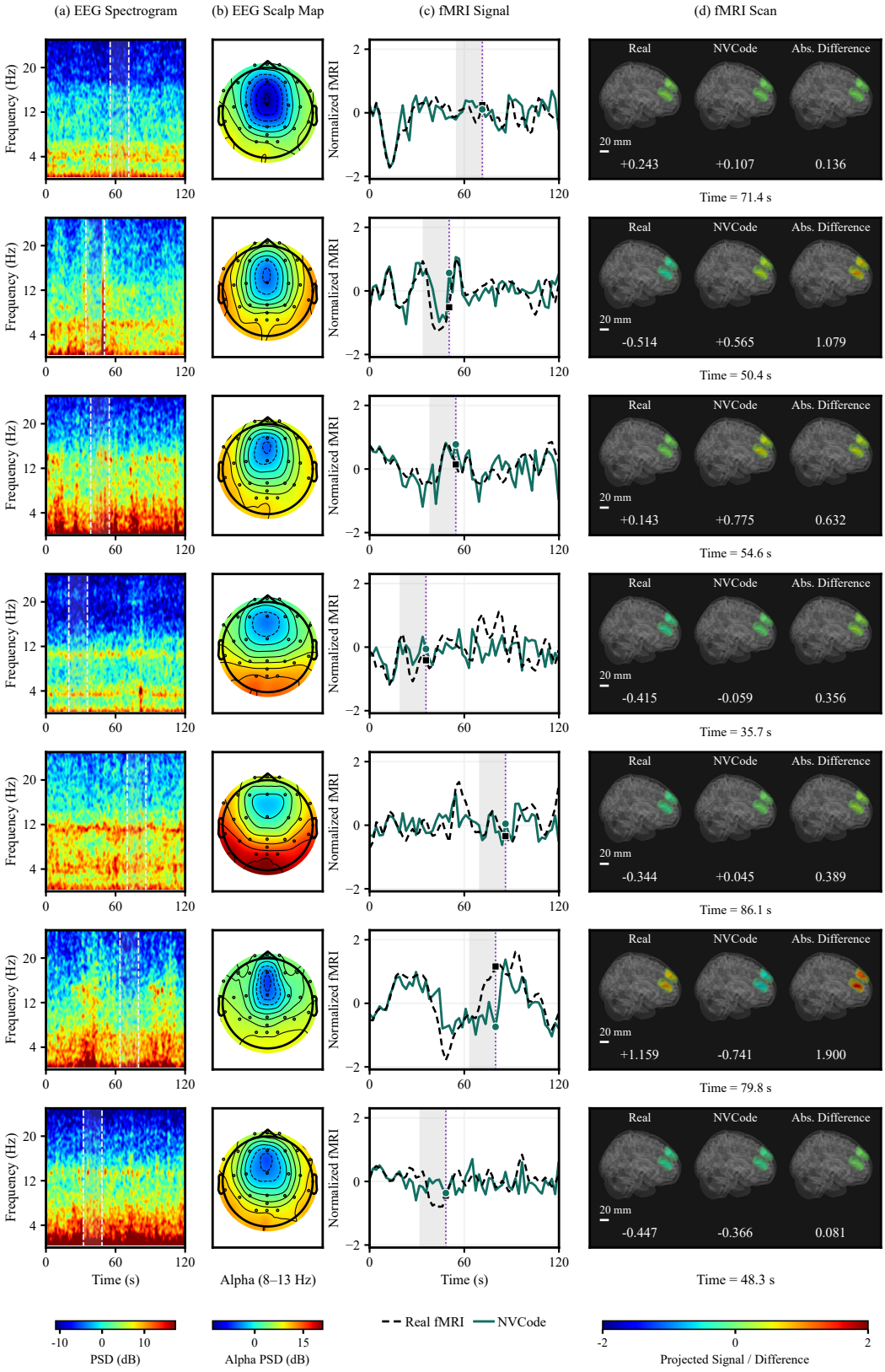


Figure 13: **Middle frontal gyrus** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (green); (d) projected real, predicted and absolute-difference amplitudes.

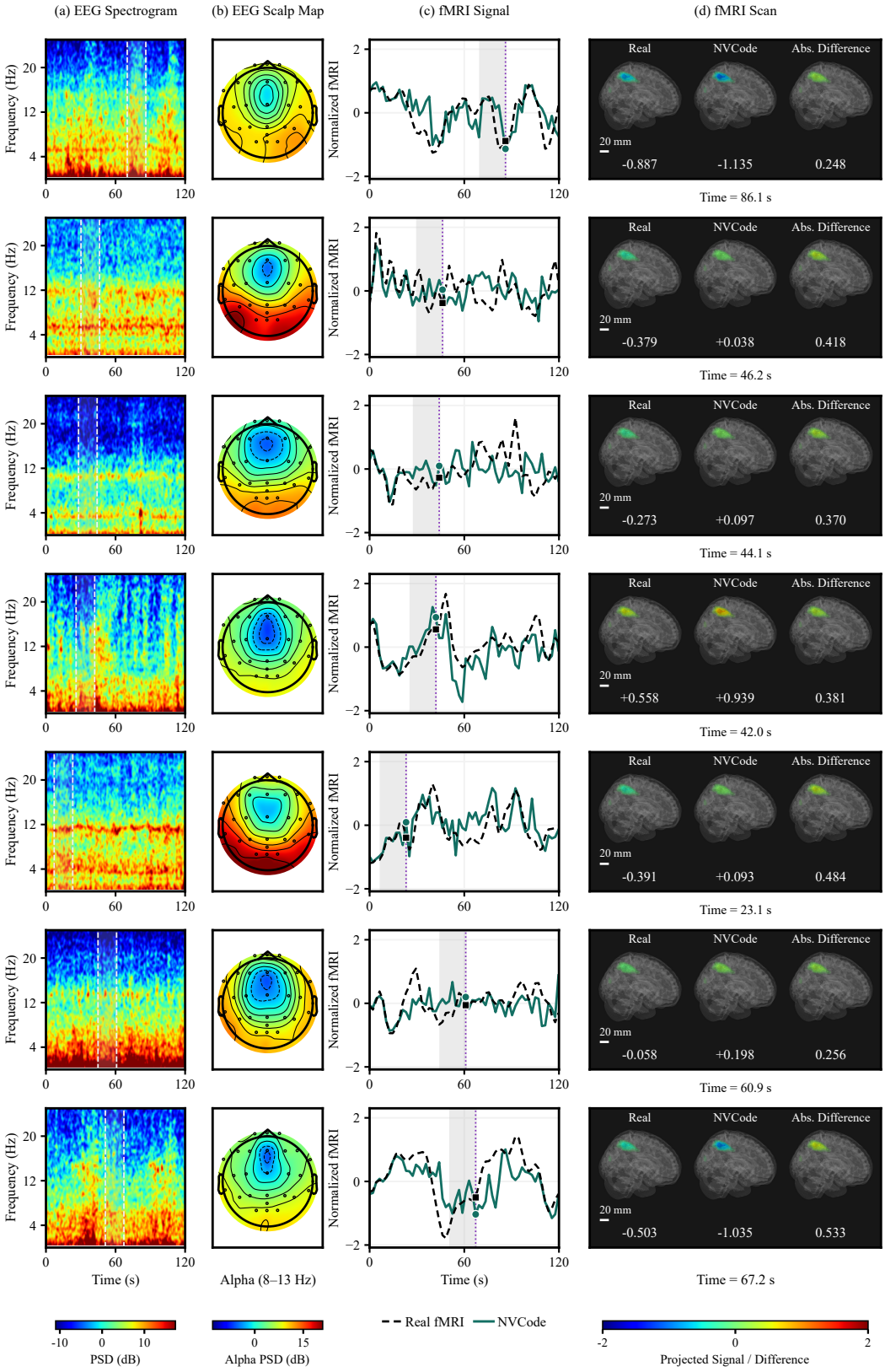


Figure 14: **Precuneus** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (green); (d) projected real, predicted and absolute-difference amplitudes.

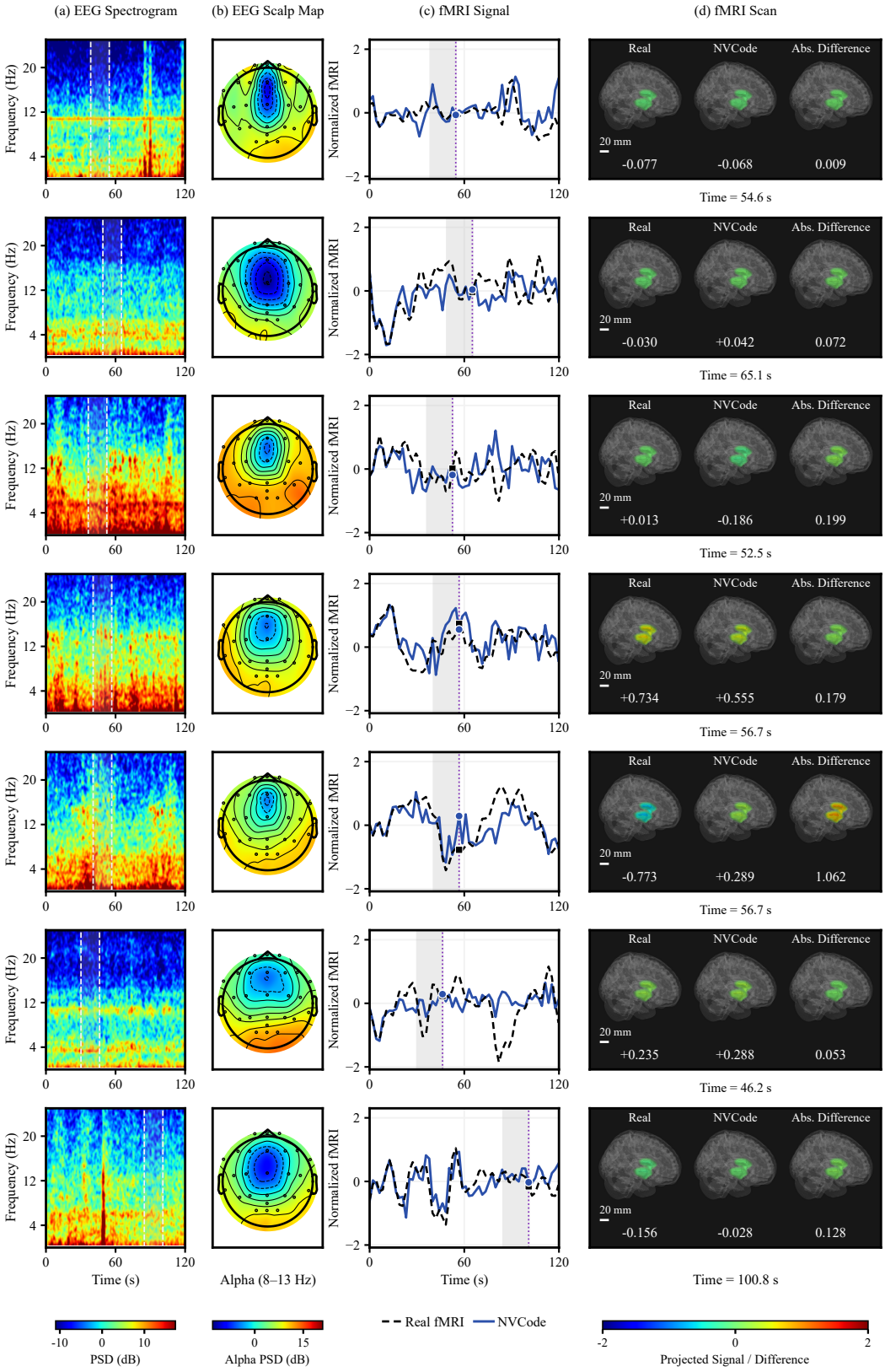


Figure 15: **Putamen** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (blue); (d) projected real, predicted and absolute-difference amplitudes.

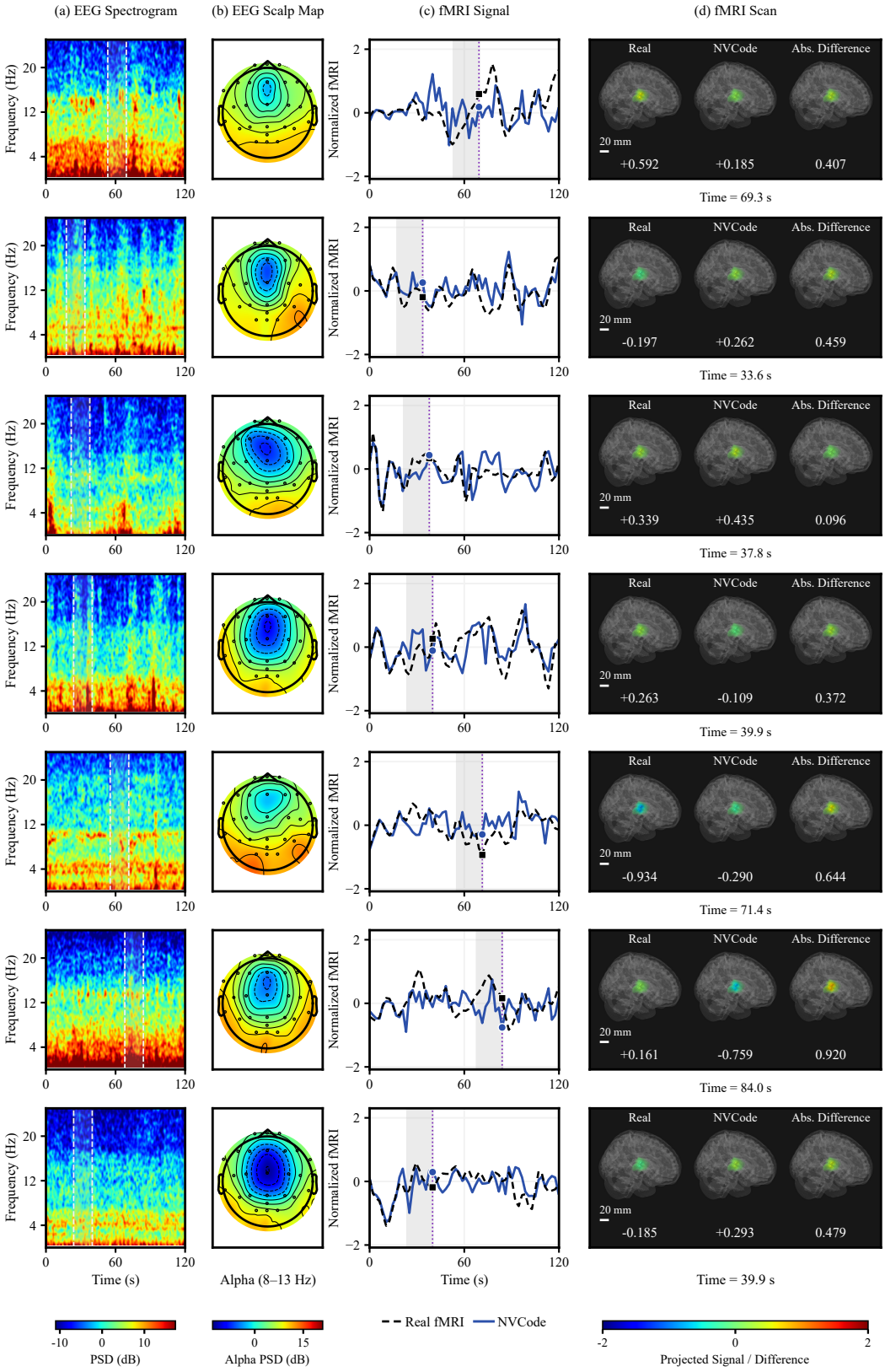


Figure 16: **Thalamus** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (blue); (d) projected real, predicted and absolute-difference amplitudes.

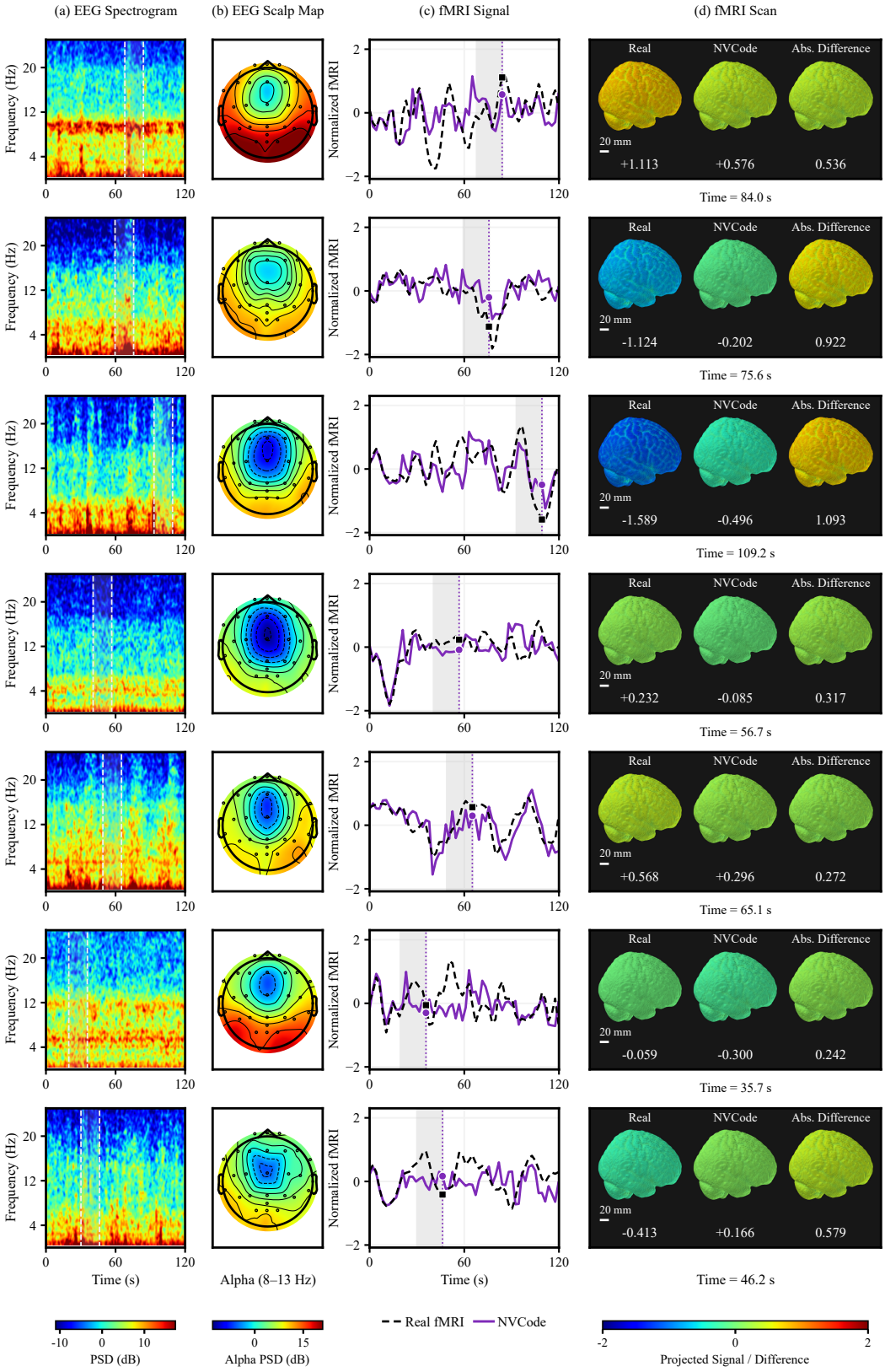


Figure 17: **Global signal** examples. (a) EEG spectrogram; (b) alpha scalp power; (c) real fMRI (black dashed) and NVCode (purple); (d) whole-brain projections of real, predicted and absolute-difference amplitudes.