
A Latent Neural ODE-VAE for Modeling Hippocampal Population Activity on Low-Dimensional Manifolds

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 Neural population activity traces trajectories in a high-dimensional state space, yet
2 accumulating evidence suggests these trajectories are confined to low-dimensional
3 manifolds that encode both task variables and internal state. Existing manifold
4 inference pipelines can recover geometry and explain variability, but often rely on
5 multi-stage local models and do not impose globally smooth continuous-time dy-
6 namics. We develop a latent Neural ODE variational autoencoder (ODE-VAE) that
7 jointly learns (i) a low-dimensional stochastic initial condition, (ii) continuous-time
8 latent dynamics parameterized by a mixture-of-experts ODE, and (iii) a decoder
9 back to neural activity. To better align reconstruction with temporally structured
10 variability, our v5 implementation adds transition-consistency regularization in
11 observation space and a soft locally linear embedding (LLE) constraint in latent
12 space. On synthetic random-foraging sequences, the model achieves high recon-
13 struction accuracy ($R^2 = 0.9789$) while exhibiting seed sensitivity. On the E65
14 hippocampal calcium dataset, we observe moderate reconstruction in archived
15 runs ($R^2 = 0.4368$) and substantial dependence on preprocessing, split strategy,
16 and evaluation space. Together, these results highlight both the promise and cur-
17 rent fragility of end-to-end continuous-time manifold models for noisy biological
18 recordings.

19 1 Introduction

20 Neural activity can be described as a point in a high-dimensional coordinate system, where each
21 coordinate axis represents a single neuron's activity [Cunningham and Yu, 2014]. Underlying
22 properties of the network and its inputs can confine neural trajectories to a subregion of this space,
23 often referred to as a neural manifold [Cunningham and Yu, 2014, Gallego et al., 2017]. The
24 neural manifold has been proposed to underlie motor movements [Gallego et al., 2017, Russo et al.,
25 2018], head direction cells [Chaudhuri et al., 2019], and hippocampal maps of physical variables
26 [O'Keefe and Dostrovsky, 1971, Frank et al., 2000, Wood et al., 2000, O'Keefe and Nadel, 1978].
27 The conceptual ideas in these studies suggest a general principle of hippocampal computation: the
28 construction of organized maps of learned knowledge instantiated by neural manifolds [Tolman, 1948,
29 O'Keefe and Nadel, 1978, Stachenfeld et al., 2017, Bellmund et al., 2018, Nieh et al., 2021].

30 Nonlinear dimensionality reduction has demonstrated that neural population activity can often be
31 described by 4–6 latent variables, suggesting that activity is constrained to a low-dimensional neural
32 manifold that displays a geometric representation of both physical and abstract variables [Low
33 et al., 2018, Chaudhuri et al., 2019, Nieh et al., 2021]. Existing approaches are limited to multi-
34 stage machine-learning pipelines, using forest-based transition models (with probabilistic principal
35 component analysis in decision-tree leaves) to define distances between population states, which are
36 then embedded into a low-dimensional manifold and mapped back to neural activity for reconstruction

[Low et al., 2018, Tipping and Bishop, 1999, Breiman, 2001, Tenenbaum et al., 2000, Yu et al., 2009]. This piecewise approach partitions state space and models dynamics locally, hence lacking explicit enforcement of globally smooth latent dynamics and can exhibit saturation of reconstruction decoding performance with low-dimensional embeddings [Low et al., 2018]. Thus, we hypothesize that generative deep learning models offer a complementary framework: neural network architectures can be trained directly on biological neural population recordings to jointly learn low-dimensional latent coordinates, their temporal evolution, and the mapping back to neural activity [Kingma and Welling, 2014, Chen et al., 2018, Rubanova et al., 2019].

In this paper, we propose a novel approach to modeling the neural manifold by constructing a Neural Ordinary Differential Equation variational autoencoder (ODE-VAE): a deep generative model that (i) encodes high-dimensional population activity into a low-dimensional latent state, (ii) models the evolution of that latent state as a continuous-time dynamical system parameterized by a neural ODE, and (iii) decodes the resulting latent trajectory back into neural activity [Kingma and Welling, 2014, Chen et al., 2018, Rubanova et al., 2019]. By training the encoder, dynamics, and decoder end-to-end under a variational objective, this approach aims to capture nonlinear manifold structure while imposing smooth temporal dynamics. Our v5 implementation introduces mixture-of-experts latent dynamics and adds two regularizers inspired by manifold inference—transition-consistency in observation space and a soft locally linear embedding (LLE) constraint in latent space [Low et al., 2018, Saul and Roweis, 2003]. We evaluate this family on synthetic and hippocampal calcium datasets and analyze the sensitivity of performance to preprocessing and evaluation choices.

Contributions.

- We formalize an ODE-VAE for trialized population sequences with mixture-of-experts latent dynamics and explicit geometric regularizers.
- We instantiate this formulation in a versioned codebase (v1–v6) and present v5 as the primary model with transition and soft-LLE regularization.
- We provide a reproducible evaluation on synthetic and hippocampal calcium datasets and identify protocol factors that strongly affect reconstruction metrics.

Related Work

Our approach lies at the intersection of manifold-based neuroscience and latent dynamical systems. In hippocampus, the cognitive map framework and subsequent experimental work motivate geometric organization of population codes [Tolman, 1948, O’Keefe and Dostrovsky, 1971, O’Keefe and Nadel, 1978, Eichenbaum and Cohen, 2014], including abstract and non-spatial representations [Aronov et al., 2017, Tavares et al., 2015, Constantinescu et al., 2016, Schuck and Niv, 2019, Park et al., 2020, Nieh et al., 2021]. Beyond classical place coding, hippocampal population activity reflects trajectory and sequential organization [Frank et al., 2000, Pastalkova et al., 2008, MacDonald et al., 2011, Taxidis et al., 2020], episodic variables at shared locations [Wood et al., 2000, Gill et al., 2011, McKenzie et al., 2014], and multimodal/task variables such as odor and taste [Eichenbaum et al., 1987, Herzog et al., 2019]. Manifold inference methods can recover low-dimensional structure and explain structured variability beyond measured task variables [Low et al., 2018, Chaudhuri et al., 2019, Rubin et al., 2019].

In machine learning, variational autoencoders [Kingma and Welling, 2014] and neural ODEs [Chen et al., 2018] provide a principled framework for continuous-time latent-variable modeling. Latent ODEs extend this idea to irregularly sampled sequences [Rubanova et al., 2019]. We adopt this framework but tailor the encoder, evaluation protocol, and regularization to the neuroscience setting, emphasizing trialized sequences, explicit geometric constraints, and comparisons to MIND-style evaluation pipelines [Low et al., 2018]. For calcium imaging recordings, related methodological work has emphasized motion correction and demixing/denoising [Pnevmatikakis et al., 2016, Pnevmatikakis and Giovannucci, 2017], highlighting the importance of preprocessing choices when evaluating reconstruction metrics.

86 **3 Problem Setup and Data**

87 We study trialized population activity sequences. Let $y_b(t_\ell) \in \mathbb{R}^N$ denote the raw activity of
 88 N simultaneously recorded units/ROIs on trial $b \in \{1, \dots, B\}$ at resampled time t_ℓ , where $\ell \in$
 89 $\{1, \dots, L\}$ indexes a fixed-length grid. We write $Y_b \in \mathbb{R}^{L \times N}$ for the stacked sequence.

90 **Observation space.** In v5, the model is trained on a PCA-projected representation of activity.
 91 Let $x_b(t_\ell) \in \mathbb{R}^K$ be the K -dimensional PCA coordinate at time t_ℓ , and let $X_b \in \mathbb{R}^{L \times K}$ be the
 92 corresponding trial sequence. Unless otherwise stated, all losses and reported R^2 values for v5 are
 93 computed in this PCA space. For some evaluations (e.g., the optional sweep path), reconstructions
 94 are mapped back to raw ROI space via inverse PCA and de-normalization.

95 **Time grid.** Trials are resampled to a common duration and the time vector is normalized to $[0, 1]$;
 96 we denote the resulting grid by $0 = t_1 < \dots < t_L = 1$. The latent dimension is denoted by D .

97 **E65 dataset.** We use the Schottorf Lab E65 dataset (E65_data.npz), containing calcium activity
 98 ($\Delta F/F$) from $N = 375$ ROIs over $T = 7434$ frames, along with trial IDs, timestamps, and aligned
 99 behavioral covariates. In the v5 preprocessing path: (i) PCA is fit to the full recording and retains
 100 95% variance, producing $K = 129$ components; (ii) frames are grouped by trial, the first 10 trials
 101 are dropped, and each trial is linearly interpolated to a fixed length $L = 120$ (trial_len_s=12,
 102 fps=10); (iii) the time vector is normalized to $[0, 1]$; (iv) each PCA component is standardized over
 103 time (session-level z-score) and an optional per-trial baseline is removed by subtracting the mean of
 104 the first 5 resampled bins. After filtering, 180 trials are available; default validation holds out the
 105 last 3 trials (train 177 / val 3). For efficiency, the default configuration further subsamples to 100
 106 sequences via greedy landmark selection.

107 **Synthetic benchmark.** We additionally evaluate on synthetic_rat_data.npz (4000 frames,
 108 300 neurons, 20 trials), which provides a controlled benchmark for recoverability of smooth low-
 109 dimensional dynamics.

110 **4 Model: Latent Neural ODE-VAE**

111 **4.1 Stochastic encoder**

112 For each trial, the encoder uses only $x_b(t_1)$ and outputs a diagonal Gaussian posterior on the latent
 113 initial state:

$$q_\phi(z_{0,b} | x_b(t_1)) = \mathcal{N}(\mu_b, \text{diag}(\sigma_b^2)), \quad (1)$$

114 with reparameterization

$$z_{0,b} = \mu_b + \sigma_b \odot \epsilon, \quad \epsilon \sim \mathcal{N}(0, I). \quad (2)$$

115 Here $z_{0,b} \in \mathbb{R}^D$, $\mu_b \in \mathbb{R}^D$, $\sigma_b \in \mathbb{R}_{>0}^D$, and \odot denotes elementwise multiplication.

116 **4.2 Continuous-time latent dynamics**

117 Latent trajectories are generated by a neural ODE:

$$\frac{dz_b(t)}{dt} = f_\theta(z_b(t), t), \quad z_b(t_1) = z_{0,b}. \quad (3)$$

118 In v5, f_θ is a mixture of experts:

$$f_\theta(z, t) = \sum_{e=1}^E \pi_e(z) f_e(z), \quad \pi(z) = \text{softmax}(g(z)), \quad (4)$$

119 with $E = 4$ latent experts by default and Dormand–Prince integration (dopri5). Each expert
 120 $f_e : \mathbb{R}^D \rightarrow \mathbb{R}^D$ is an MLP and $\pi_e(z) \in [0, 1]$ are gating weights satisfying $\sum_e \pi_e(z) = 1$.

121 **4.3 Decoder family**

122 A decoder maps latent states back to observations:

$$\hat{x}_b(t_\ell) = g_\psi(z_b(t_\ell)). \quad (5)$$

123 The codebase supports MLP, neuron-aware, local-attention, and MoE decoders; v5 default is MoE
 124 decoder with 8 decoder experts. In all cases, $g_\psi : \mathbb{R}^D \rightarrow \mathbb{R}^K$ outputs the mean of a factorized
 125 Gaussian observation model in PCA space.

126 **5 Training Objective and Regularization**

127 We optimize a variational objective with auxiliary regularizers. Under a Gaussian observation model
 128 $p_\psi(x_b(t_\ell) | z_b(t_\ell)) = \mathcal{N}(g_\psi(z_b(t_\ell)), \sigma^2 I)$ with fixed σ^2 , maximizing the ELBO corresponds (up to
 129 constants and a scale factor) to minimizing mean-squared reconstruction error plus a KL penalty.

130 The base objective combines reconstruction and KL terms:

$$\mathcal{L}_{\text{base}} = \mathcal{L}_{\text{rec}} + \beta \mathcal{L}_{\text{KL}}, \quad (6)$$

131 where

$$\mathcal{L}_{\text{rec}} = \frac{1}{B L K} \sum_{b=1}^B \sum_{\ell=1}^L \|\hat{x}_b(t_\ell) - x_b(t_\ell)\|_2^2, \quad (7)$$

$$\mathcal{L}_{\text{KL}} = \frac{1}{B} \sum_b D_{\text{KL}}(q_\phi(z_{0,b} | x_b(t_1)) \| \mathcal{N}(0, I)). \quad (8)$$

133 Equivalently, the (negative) ELBO per trial is

$$\mathcal{L}_{\text{ELBO}} = -\mathbb{E}_{q_\phi(z_{0,b} | x_b(t_1))} \left[\sum_{\ell=1}^L \log p_\psi(x_b(t_\ell) | z_b(t_\ell)) \right] + \beta D_{\text{KL}}(q_\phi(z_{0,b} | x_b(t_1)) \| p(z_{0,b})), \quad (9)$$

134 with prior $p(z_{0,b}) = \mathcal{N}(0, I)$. In practice, the code uses a single Monte Carlo sample of $z_{0,b}$ per trial
 135 and minibatch.

Smoothness regularization.

$$\mathcal{L}_{\text{smooth}} = \frac{1}{B(L-1)D} \sum_{b,\ell} \left\| \frac{z_b(t_{\ell+1}) - z_b(t_\ell)}{t_{\ell+1} - t_\ell} \right\|_2^2. \quad (10)$$

Transition-aware regularization (v5).

$$\mathcal{L}_{\text{trans}} = \frac{1}{B(L-1)K} \sum_{b,\ell} \left\| (\hat{x}_b(t_{\ell+1}) - \hat{x}_b(t_\ell)) - (x_b(t_{\ell+1}) - x_b(t_\ell)) \right\|_2^2. \quad (11)$$

136 This term is linearly warmed up for the first 30 epochs.

137 **Soft LLE latent regularization (v5).** For flattened latent points $\{z_i\}_{i=1}^M \subset \mathbb{R}^D$, with k -NN set
 138 $\mathcal{N}_k(i)$, we add a soft locally linear embedding penalty [Saul and Roweis, 2003]:

$$\mathcal{L}_{\text{LLE}} = \frac{1}{M} \sum_{i=1}^M \left\| z_i - \sum_{j \in \mathcal{N}_k(i)} w_{ij} z_j \right\|_2^2, \quad w_{ij} \propto \exp\left(-\frac{\|z_i - z_j\|_2}{\tau}\right). \quad (12)$$

139 Default parameters: $k = 8$, $M \leq 256$, $\tau = 0.1$.

Total loss.

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \beta_t \mathcal{L}_{\text{KL}} + \lambda_{\text{smooth}} \mathcal{L}_{\text{smooth}} + \lambda_{\text{trans},t} \mathcal{L}_{\text{trans}} + \lambda_{\text{LLE}} \mathcal{L}_{\text{LLE}}. \quad (13)$$

140 The KL coefficient β_t is warmed up over 30 epochs to a final value $\beta = 0.02$.

Table 1: Synthetic random-foraging benchmark from `seed_sweep_results.txt`.

Seed	Final R^2	Best validation loss
1	0.9789	0.06642
42	0.6757	0.40521
1337	0.9116	0.16298
2025	0.7467	0.31145
777	0.0000 (NaN collapse)	∞
Mean (all seeds)	0.6626	—
Mean (non-collapsed seeds)	0.8282	—

141 6 Experimental Protocol

142 6.1 Configurations

143 Main v5 settings from `config.txt`: latent dimension $D = 5$ (with a sweep to $D = 8$), batch size 8,
 144 150 epochs, Adam optimizer (learning rate 0.002, weight decay 10^{-5}), $\lambda_{\text{smooth}} \in \{5 \times 10^{-4}, 2 \times 10^{-4}\}$,
 145 $\lambda_{\text{trans}} = 0.01$ with a 30-epoch warmup, $\lambda_{\text{LLE}} = 0.01$, landmark count 100, and baseline correction
 146 enabled.

147 **Implementation details.** The encoder is an MLP with hidden widths 512–256–128. The latent
 148 vector field uses $E = 4$ experts with hidden width 128 and a learned gating network; derivatives
 149 are layer-normalized for stability. For reconstruction, the default v5 decoder is a mixture-of-experts
 150 network with 8 decoder experts and hidden width 256. Latent dynamics are integrated with Dormand–
 151 Prince (`dopri5`) using tolerances `rtol=1e-3` and `atol=1e-4`. Gradients are clipped to max norm
 152 1.0.

153 6.2 Metrics

154 The primary training metric is coefficient of determination,

$$R^2 = 1 - \frac{\sum_{b,\ell} \|x_b(t_\ell) - \hat{x}_b(t_\ell)\|_2^2}{\sum_{b,\ell} \|x_b(t_\ell) - \bar{x}\|_2^2}, \quad (14)$$

155 where $\bar{x} = \frac{1}{BL} \sum_{b,\ell} x_b(t_\ell)$ denotes the mean activity vector across all validation entries in the
 156 evaluation space. Our implementation supports two evaluation styles: (i) PCA-space R^2 and (ii) strict
 157 raw-neuron-space R^2 via inverse PCA and de-normalization. We report values as saved in the run
 158 metadata for each experiment.

159 7 Results

160 7.1 Synthetic benchmark: high ceiling with seed sensitivity

161 Table 1 summarizes a five-seed sweep on synthetic data. The best seed reaches $R^2 = 0.9789$.
 162 Excluding the divergent run, the mean performance is $R^2 = 0.828 \pm 0.141$ (std. dev.), indicating
 163 sensitivity to initialization even in a controlled setting.

164 7.2 E65 hippocampal data: mixed performance across variants

165 Table 2 reports available E65 run metrics for the ODE-VAE family. An archived ODE-VAE run
 166 reports $R^2 = 0.4368$. In contrast, v5 sweep artifacts are substantially lower (best-epoch $R^2 \leq 0.0569$,
 167 negative final R^2 in three runs), while a no-PCA v6 run records $R^2 = 0.0956$. These discrepancies
 168 are consistent with the codebase supporting multiple evaluation spaces and split protocols; therefore,
 169 comparisons across runs should be interpreted cautiously unless the evaluation definition is matched.

Table 2: E65 reconstruction metrics extracted from saved run artifacts. Sweep runs report both best-epoch and final R^2 as logged during training.

Variant	Space	D	λ_{smooth}	Best R^2	Final R^2
archived ODE-VAE	PCA	—	—	—	0.4368
v5 sweep	PCA	5	5×10^{-4}	0.0354	-0.2391
v5 sweep	PCA	5	2×10^{-4}	0.0353	-0.1382
v5 sweep	PCA	8	5×10^{-4}	0.0569	-0.2022
v6 (no PCA)	raw	—	—	—	0.0956

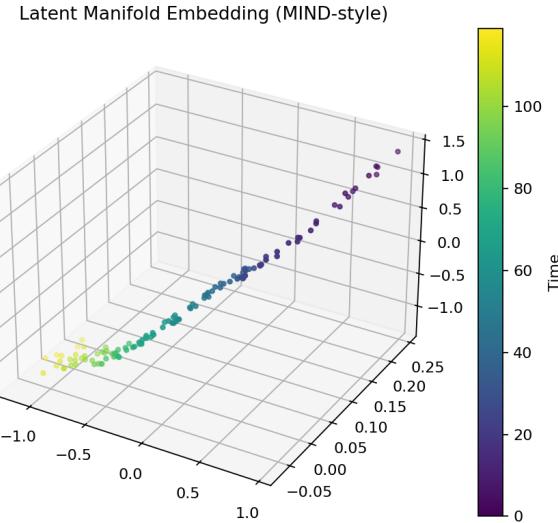


Figure 1: Latent manifold embedding produced by the ODE-VAE analysis pipeline.

170 7.3 Manifold interpretability

171 The codebase saves latent manifold projections (MDS) and reconstruction diagnostics for each run.
 172 Figure 1 shows an example latent trajectory embedding from the trained model artifacts.

173 8 Discussion

174 The model captures the intended inductive bias: low-dimensional continuous latent trajectories
 175 with explicit geometric regularization. On synthetic data, this bias is highly effective. On real E65
 176 recordings, however, results are sensitive to implementation and evaluation choices.

177 Three factors emerge from the saved run artifacts:

- 178 **Metric-space mismatch.** PCA-space training can look favorable while strict raw-space R^2
 179 may degrade. This is particularly salient when comparing to MIND-style evaluations, which
 180 reconstruct back to neuron space (via inverse PCA) before scoring [Low et al., 2018].
- 181 **Data-efficiency tradeoff.** Landmark subsampling (100 selected sequences from 180 usable
 182 trials) accelerates training but may reduce generalization. In MIND, landmarks primarily
 183 support graph construction and embedding; the learned mapping is then applied to all
 184 eligible time points [Low et al., 2018].
- 185 **Optimization stability.** Strong regularization with small validation sets (3 trials) and stiff
 186 latent dynamics can produce unstable or negative final R^2 , despite early high points.

187 These observations suggest that future gains likely require protocol-level changes in addition to
 188 architectural changes: larger and randomized holdout splits, early stopping on a stable cross-validated
 189 objective, trial-level (not frame-level) landmark selection, and direct raw-space reconstruction losses.

190 **9 Future Work**

191 A central motivation of this project is to connect end-to-end continuous-time latent dynamical
192 modeling with the multi-stage manifold inference pipeline used in MIND [Low et al., 2018]. Our
193 current codebase already adopts several MIND-inspired components (global PCA preprocessing,
194 greedy landmark selection for visualization, and MDS-based manifold plots), but the modeling
195 philosophy differs: MIND estimates a graph of transition structure via a PPCA regression forest and
196 learns explicit local mappings between ambient activity and manifold coordinates [Low et al., 2018,
197 Tipping and Bishop, 1999, Breiman, 2001], whereas the ODE-VAE learns a single global generative
198 model (encoder + latent dynamics + decoder) by optimizing a reconstruction objective [Kingma and
199 Welling, 2014, Chen et al., 2018, Rubanova et al., 2019]. Below we outline concrete directions to
200 tighten this connection and improve robustness on calcium recordings.

201 **9.1 Match MIND-style evaluation protocols and metrics**

202 Many apparent discrepancies across saved E65 runs are consistent with evaluation-definition mis-
203 match. In the MIND Matlab cross-validation script, trials are split randomly (e.g., 90/10), reconstruc-
204 tion is scored in the original neuron space after mapping back through inverse PCA, and performance
205 is visualized both as an overall score and as per-trial dots [Low et al., 2018]. Aligning our training and
206 reporting with this protocol would make comparisons substantially more interpretable. Concretely,
207 we plan to (i) report both Pearson correlation on vectorized activity blocks,

$$r = \text{corr}\left(\text{vec}(Y_{\text{test}}), \text{vec}(\hat{Y}_{\text{test}})\right), \quad (15)$$

208 and variance-explained R^2 under repeated random trial splits, and (ii) include held-out neuron
209 evaluation where latents are inferred from a subset of neurons and used to predict excluded neurons,
210 mirroring the “cell prediction” analyses in MIND [Low et al., 2018]. This will also require revisiting
211 the current practice of validating on the final 3 trials, which can conflate generalization with drift.

212 **9.2 Use landmarks for geometry, not for shrinking the training set**

213 In MIND, landmarks are an efficiency device for graph construction and embedding; the learned
214 mapping is then applied to all eligible time points [Low et al., 2018]. In contrast, the default v5
215 configuration further subsamples the dataset down to 100 landmarked sequences (from 180 trials),
216 which likely increases estimator variance and can bias which trials are emphasized during training.
217 A straightforward next step is to train the ODE-VAE on all trials/time points and reserve landmark
218 selection for: (i) visualization, (ii) neighbor graph construction for local regularizers, and (iii)
219 lightweight geometric diagnostics (e.g., random-walk distance embeddings). This change should
220 directly improve stability without changing the model class.

221 **9.3 Hybrid decoders: combine global reconstruction with MIND-like local mappings**

222 The MIND pipeline learns mappings between ambient PCA space and manifold coordinates using
223 locally weighted methods (e.g., LLE regression) [Saul and Roweis, 2003, Low et al., 2018]. This
224 provides a natural mechanism to capture sharp, local irregularities that global regressors may smooth
225 out. Our current v5 decoders are global function approximators (MLP/MoE), which can yield good
226 coarse reconstructions but may miss neuron-specific transients. An appealing hybrid is a global
227 decoder plus a local residual term defined over nearby latent states,

$$\hat{x}(t) = g_\psi(z(t)) + \sum_{j \in \mathcal{N}_k(z(t))} \alpha_j(z(t)) r_j, \quad (16)$$

228 where $\mathcal{N}_k(\cdot)$ are neighbors in latent space (or in a MIND-style random-walk metric), $r_j \in \mathbb{R}^K$ are
229 learned prototype residuals, and α_j are normalized weights (e.g., softmax over distances). This would
230 preserve the interpretability and global smoothness of the ODE while injecting the kind of local
231 adaptivity that MIND’s mapping stage provides.

232 **9.4 Optimize and score in raw neuron space (with PCA as an internal linear layer)**

233 Several E65 runs in this repository train and score in different spaces (PCA vs raw ROI), making R^2
234 values hard to compare. MIND keeps PCA primarily as a compression step but reconstructs back

235 to the original activity space before computing reconstruction scores [Low et al., 2018]. A direct
 236 analogue for the ODE-VAE is to keep a fixed (or lightly fine-tuned) PCA projection for computational
 237 efficiency, but decode back to raw ROI space and compute the main reconstruction loss on $y_b(t) \in \mathbb{R}^N$.
 238 One implementation is to parameterize a raw-space decoder as $\hat{y}(t) = W_{\text{PCA}}^\top \hat{x}(t) + \mu$, using the
 239 PCA loading matrix W_{PCA} and mean μ from preprocessing, and to define \mathcal{L}_{rec} in raw space. This
 240 would more closely match the scientific question—reconstructing neural activity—and reduce the
 241 chance that good PCA-space fits hide biologically relevant errors.

242 9.5 Make latent dynamics probabilistic to better match MIND transition structure

243 MIND estimates transition structure via a probabilistic model of next-step activity (a PPCA regression
 244 forest) and then derives a random-walk geometry from transition probabilities [Low et al., 2018,
 245 Tipping and Bishop, 1999, Breiman, 2001]. Our latent ODE is deterministic given $z_{0,b}$, which
 246 can be brittle when real data exhibit unmodeled inputs, nonstationarities, or observation noise. A
 247 natural extension is to introduce process noise (Neural SDEs) or discrete-time stochastic residuals,
 248 $z(t_{\ell+1}) = z(t_\ell) + \int_{t_\ell}^{t_{\ell+1}} f_\theta(z(t), t) dt + \eta_\ell$, which can absorb variability not explained by the initial
 249 condition while retaining smooth latent structure. This direction also creates a clearer conceptual
 250 bridge between ODE-based dynamics and MIND’s transition-probability graph.

251 9.6 Geometry-aware objectives beyond soft-LLE

252 Our current soft-LLE penalty encourages local linearity in the learned latent point cloud, but it
 253 does not directly use transition structure. The MIND code constructs local distances from transition
 254 probabilities (e.g., $d_{ij} \propto \sqrt{-\log p_{ij}}$) and then computes geodesic distances on the resulting graph
 255 before embedding [Low et al., 2018]. This is conceptually related to geodesic-distance embeddings in
 256 nonlinear dimensionality reduction [Tenenbaum et al., 2000]. A promising direction is to import this
 257 idea as a regularizer: estimate a transition graph among landmarked latent points, compute a random-
 258 walk geodesic distance matrix, and penalize distortions between these distances and Euclidean
 259 distances in the latent embedding. Such a constraint could encourage the latent representation to
 260 respect the sequential structure that MIND leverages, while still permitting an end-to-end generative
 261 model.

262 10 Limitations and Reproducibility

263 This study is bounded by the available run artifacts and inherits version-specific logging differences.
 264 In particular, some run files report “best” and “final” R^2 under different conditions, and not all check-
 265 points include identical metadata fields. We therefore report values exactly as saved in each artifact
 266 path. The implementation also exhibits training fragility (including occasional NaN divergence),
 267 which should be addressed before drawing definitive biological conclusions.

268 11 Conclusion

269 We presented a mathematically grounded latent Neural ODE-VAE framework for neural manifold
 270 modeling and analyzed a sequence of model variants (v1–v6), with v5 as the primary model. The
 271 method can recover smooth low-dimensional dynamics and high synthetic reconstruction quality,
 272 but real-data performance remains sensitive to preprocessing and evaluation protocol. This work
 273 provides a formal foundation and concrete directions for improving robustness of ODE-VAE manifold
 274 modeling for neuroscience.

275 References

- 276 Dmitriy Aronov, Rachel Nevers, and David W. Tank. Mapping of a non-spatial dimension by the
 277 hippocampal–entorhinal circuit. *Nature*, 543(7647):719–722, 2017.
- 278 Jacob L. S. Bellmund, Peter Gärdenfors, Edvard I. Moser, and Christian F. Doeller. Navigating
 279 cognition: Spatial codes for human thinking. *Science*, 362(6415):eaat6766, 2018.
- 280 Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

- 281 Rishidev Chaudhuri, Burak Gerçek, Bikash Pandey, Adrien Peyrache, and Ila Fiete. The intrinsic
282 attractor manifold and population dynamics of a canonical cognitive circuit across waking and
283 sleep. *Nature Neuroscience*, 22:1512–1520, 2019.
- 284 Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary
285 differential equations. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.
- 286 Alexandra O. Constantinescu, Jill X. O'Reilly, and Timothy E. J. Behrens. Organizing conceptual
287 knowledge in humans with a grid-like code. *Science*, 352(6292):1464–1468, 2016.
- 288 John P. Cunningham and Byron M. Yu. Dimensionality reduction for large-scale neural recordings.
289 *Nature Neuroscience*, 17(11):1500–1509, 2014.
- 290 Howard Eichenbaum and Neal J. Cohen. Can we reconcile the declarative memory and spatial
291 navigation views on hippocampal function? *Neuron*, 83(4):764–770, 2014.
- 292 Howard Eichenbaum, Menachem Kuperstein, Andrew Fagan, and Janet Nagode. Cue-sampling and
293 goal-approach correlates of hippocampal unit activity in rats performing an odor-discrimination
294 task. *Journal of Neuroscience*, 7:716–732, 1987.
- 295 Loren M. Frank, Emery N. Brown, and Matthew Wilson. Trajectory encoding in the hippocampus
296 and entorhinal cortex. *Neuron*, 27:169–178, 2000.
- 297 Juan A. Gallego, Matthew G. Perich, Lee E. Miller, and Sara A. Solla. Neural manifolds for the
298 control of movement. *Neuron*, 94(5):978–984, 2017.
- 299 P. R. Gill, Sheri J. Y. Mizumori, and David M. Smith. Hippocampal episode fields develop with
300 learning. *Hippocampus*, 21:1240–1249, 2011.
- 301 Lauren E. Herzog et al. Interaction of taste and place coding in the hippocampus. *Journal of
302 Neuroscience*, 39:3057–3069, 2019.
- 303 Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *International Conference
304 on Learning Representations (ICLR)*, 2014.
- 305 Ryan J. Low, Sean Lewallen, Dmitriy Aronov, Rachel Nevers, and David W. Tank. Probing variability
306 in a cognitive map using manifold inference from neural dynamics. *bioRxiv*, 2018. doi: 10.1101/
307 418939.
- 308 Chris J. MacDonald, Kyle Q. Lepage, Uri T. Eden, and Howard Eichenbaum. Hippocampal “time
309 cells” bridge the gap in memory for discontiguous events. *Neuron*, 71:737–749, 2011.
- 310 Sam McKenzie et al. Hippocampal representation of related and opposing memories develop within
311 distinct, hierarchically organized neural schemas. *Neuron*, 83:202–215, 2014.
- 312 Edward H. Nieh et al. Geometry of abstract learned knowledge in the hippocampus. *Nature*, 595
313 (7865):80–84, 2021.
- 314 John O'Keefe and Jonathan Dostrovsky. The hippocampus as a spatial map. preliminary evidence
315 from unit activity in the freely-moving rat. *Brain Research*, 34(1):171–175, 1971.
- 316 John O'Keefe and Lynn Nadel. *The Hippocampus as a Cognitive Map*. Clarendon Press, 1978.
- 317 Sang Ah Park, David S. Miller, Hamed Nili, Charan Ranganath, and Erie D. Boorman. Map making:
318 Constructing, combining, and inferring on abstract cognitive maps. *Neuron*, 107(6):1226–1238.e8,
319 2020.
- 320 Eva Pastalkova, Vladimir Itskov, A. Amarasingham, and Gyorgy Buzsaki. Internally generated cell
321 assembly sequences in the rat hippocampus. *Science*, 321:1322–1327, 2008.
- 322 Eftychios A. Pnevmatikakis and Andrea Giovannucci. Normcorre: An online algorithm for piecewise
323 rigid motion correction of calcium imaging data. *Journal of Neuroscience Methods*, 291:83–94,
324 2017.

- 325 Eftychios A. Pnevmatikakis et al. Simultaneous denoising, deconvolution, and demixing of calcium
326 imaging data. *Neuron*, 89:285–299, 2016.
- 327 Yulia Rubanova, Ricky T. Q. Chen, and David Duvenaud. Latent ordinary differential equations for
328 irregularly-sampled time series. In *Advances in Neural Information Processing Systems (NeurIPS)*,
329 2019.
- 330 Amir Rubin et al. Revealing neural correlates of behavior without behavioral measurements. *Nature
331 Communications*, 10:1–14, 2019.
- 332 Abigail A. Russo et al. Motor cortex embeds muscle-like commands in an untangled population
333 response. *Neuron*, 97(4):953–966.e8, 2018.
- 334 Lawrence K. Saul and Sam T. Roweis. Think globally, fit locally: Unsupervised learning of low
335 dimensional manifolds. *Journal of Machine Learning Research*, 4:119–155, 2003.
- 336 Nicolas W. Schuck and Yael Niv. Sequential replay of nonspatial task states in the human hippocam-
337 pus. *Science*, 364(6447):eaaw5181, 2019.
- 338 Kimberly L. Stachenfeld, Matthew M. Botvinick, and Samuel J. Gershman. The hippocampus as a
339 predictive map. *Nature Neuroscience*, 20(11):1643–1653, 2017.
- 340 Rita M. Tavares et al. A map for social navigation in the human brain. *Neuron*, 87:231–243, 2015.
- 341 Jiannis Taxidis et al. Differential emergence and stability of sensory and temporal representations in
342 context-specific hippocampal sequences. *Neuron*, 108:984–998.e9, 2020.
- 343 Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for
344 nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.
- 345 Michael E. Tipping and Christopher M. Bishop. Probabilistic principal component analysis. *Journal
346 of the Royal Statistical Society: Series B*, 61(3):611–622, 1999.
- 347 Edward C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948.
- 348 Eric R. Wood, Paul A. Dudchenko, Rebekka J. Robitsek, and Howard Eichenbaum. Hippocampal
349 neurons encode information about different types of memory episodes occurring in the same
350 location. *Neuron*, 27:623–633, 2000.
- 351 Byron M. Yu, John P. Cunningham, Gopal Santhanam, Stephen I. Ryu, Krishna V. Shenoy, and
352 Maneesh Sahani. Gaussian-process factor analysis for low-dimensional single-trial analysis of
353 neural population activity. *Journal of Neurophysiology*, 102(1):614–635, 2009.