Beyond the Euclidean brain: inferring non-Euclidean latent trajectories from spike trains

Neuroscience faces a growing need for scalable data analysis methods that reduce the dimensionality of population recordings yet retain key aspects of the computation or behaviour. To extract interpretable latent trajectories from neural data, it is critical to embrace the inherent topology of the features of interest: head direction evolves on a ring or torus, 3D body rotations on the special orthogonal group, and navigation is best described in the intrinsic coordinates of the environment. Accordingly, the manifold Gaussian process latent variable model (mGPLVM) was recently proposed to simultaneously infer latent representations on non-Euclidean manifolds and how neurons are tuned to these representations. This probabilistic method generalizes previous Euclidean models and allows principled selection between candidate latent topologies. While powerful, mGPLVM makes two unjustified approximations that limit its practical applicability to neural datasets. First, consecutive latent states are assumed independent *a priori*, whereas behaviour is continuous in time. Second, its Gaussian noise model is inappropriate for positive integer spike counts. Previous work in Euclidean LVMs such as GPFA has shown significant improvements in performance when modeling such features appropriately. Here, we extend mGPLVM by incorporating temporally continuous priors over latent states and flexible count-based noise models. This improves inference on synthetic data, avoiding negative spike count predictions and discontinuous jumps in latent trajectories. On real data, we also mitigate these pathologies while improving model fit compared to the original mGPLVM formulation. In summary, our extended mGPLVM provides a widely applicable tool for the study of (non-)Euclidean neural representations with flexible priors and noise models. We provide an efficient GPU implementation in python, relying on recent advances in approximate inference to e.g. fit 10,000 time bins of recording for 100 neurons in five minutes.

---

mGPLVM The details of mGPLVM are described in [1]. Briefly, neural activity is assumed to arise from a set of latent states $\{g\}$ on some potentially non-Euclidean manifold $\mathcal{M}$ through a Gaussian process observation model defined on the manifold:

$$\{g\} \sim p^{\mathcal{M}}(\{g\}) \qquad \text{(prior over latents)} \qquad (1)$$

$$f_i \sim \mathcal{GP}(0, k_i^{\mathcal{M}}(\cdot,\cdot)) \qquad \text{(prior over tuning)} \qquad (2)$$

$$y_{it}|g_t \sim p(y_{it}|f_i(g_t)) \qquad \text{(noise model)} \qquad (3)$$

This model succesfully infers non-Euclidean latent states and tuning curves from synthetic and experimental data, and correctly identifies the underlying latent topology on manifolds ranging from rings and tori to the group of 3D rotations. In [1], the prior over latents (Equation 1) was assumed to factorize over time, $p^{\mathcal{M}}(\{g_t\}) = \prod_t p^{\mathcal{M}}(g_t)$ with $p^{\mathcal{M}}(g)$ being uniform on $\mathcal{M}$. This is an inappropriate model for processes with known temporal dependencies such as navigation or recurrent network dynamics. However, extending the model to non-factorized priors is challenging due to the non-Euclidean nature of the latent space. Further, the noise model (Equation 3) was assumed to be Gaussian which is inappropriate for e.g. count-based data such as that arising from electrophysiological recordings. Here, we extend mGPLVM to include continuous temporal priors and fit count-based observations.

Continuous temporal priors We start from the mGPLVM Evidence Lower Bound (ELBO) [1]:

$$\mathcal{L} = \underbrace{H(Q)}_{\text{entropy}} + \underbrace{\mathbb{E}_Q[\log p(\boldsymbol{Y}|\{g\})]}_{\text{likelihood}} + \underbrace{\mathbb{E}_Q[\log p(\{g\})]}_{\text{prior}}, \qquad (4)$$

where Q is a variational distribution with parameters θ, and $\boldsymbol{Y} \in \mathbb{R}^{N \times T}$ is the recorded activity of N neurons at T time steps. We encourage continuous latent variable trajectories by introducing a Markovian prior:

$$\log p^{\mathcal{M}}(\{g_t\}) = \log p_0^{\mathcal{M}} + \sum_{t=1} \log p^{\mathcal{M}}(g_t|g_{t-1}), \qquad (5)$$

where the latent state at time t depends on that at $t-1$. Here, we propose a tractable density $p^{\mathcal{M}}(g_t|g_{t-1})$ defined on Lie groups which allows for Monte Carlo estimation of the prior. In Euclidean space, we can define a random walk prior as $p^{\mathbb{R}^n}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{x}_t; \boldsymbol{x}_{t-1} + \boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are learnable parameters. Generalizing this to non-Euclidean Lie groups, we define a reference distribution on the tangent space, $r(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, and project it onto the manifold:

$$p^{\mathcal{M}}(g_t|g_{t-1}) = \sum_{\substack{\boldsymbol{x} \in \mathbb{R}^n: \\ \mathrm{Exp}_G(\boldsymbol{x}) = g_{t-1}^{-1} \circ g_t}} r(\boldsymbol{x}) |\boldsymbol{J}(\boldsymbol{x})|^{-1}, \qquad (6)$$

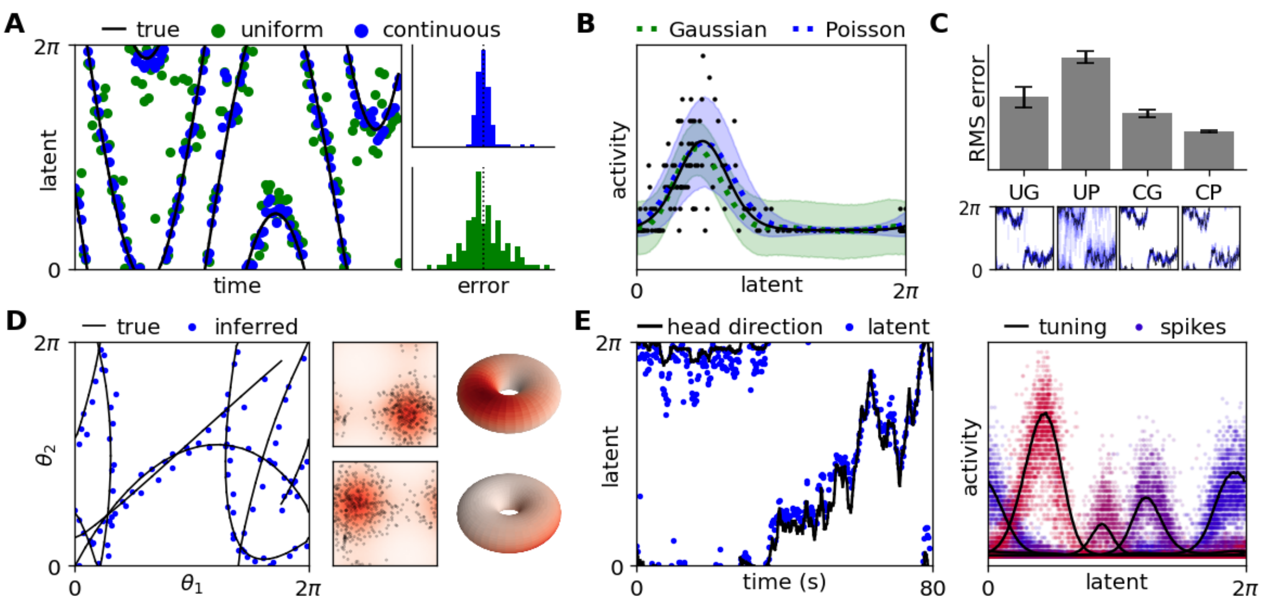
where $\boldsymbol{J}(\boldsymbol{x})$ is the Jacobian of the exponential map Exp_G at $\boldsymbol{x}$ [1,2], $\boldsymbol{\mu}$ models any systematic drift on $\mathcal{M}$, and $\boldsymbol{\Sigma}$ models the degree of continuity.

Count-based observations To generalize the observation model to fit count data, we use a variational distribution $q(\boldsymbol{u}) = \mathcal{N}(\boldsymbol{u}|\boldsymbol{m}, \boldsymbol{S})$ to lower-bound the GP likelihood term in the ELBO [3]:

$$\log p(\boldsymbol{y}|\{g\}) \geq \mathbb{E}_{q(\boldsymbol{f})}[\log p(\boldsymbol{y}|\boldsymbol{f})] - \mathrm{KL}(q(\boldsymbol{u})|p(\boldsymbol{u}|\{g\})),$$

where $\boldsymbol{y}$ is a single row of $\boldsymbol{Y}$, $\boldsymbol{f}$ are the corresponding function values of the GP, and $q(\boldsymbol{f}) = \int_{\boldsymbol{u}} p(\boldsymbol{f}|\boldsymbol{u})q(\boldsymbol{u})d\boldsymbol{u}$. Under a Poisson noise model with an exponential link function commonly used for neural count data, this lower bound can be evaluated analytically. For more general noise models, we instead approximate $\mathbb{E}_{q(\boldsymbol{f})}[\log p(\boldsymbol{y}|\boldsymbol{f})]$ using Gauss-Hermite quadrature, which is applicable to any noise distribution with a closed-form likelihood $p(\boldsymbol{y}|\boldsymbol{f})$. This allows us to fit general count data and account for e.g the overdispersion commonly found in neural recordings.

Results For synthetic data with a smooth latent trajectory on a ring, the continuous mGPLVM recovers the true latents better than with a uniform prior (Fig. A), with the learned parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ approximating the distribution of consecutive displacements. This leads to lower uncertainty in the variational distribution and improved log marginal likelihoods (ΔLL $= (0.10 \pm 0.02) \times NT$). For synthetic count data, the Poisson noise model improves the inferred tuning curves, capturing changes in uncertainty with preferred orientation and avoiding negative spike counts (Fig. B). Combining these features, we compare models with continuous or uniform priors and Gaussian or Poisson noise models on count data from a synthetic Poisson head direction circuit with random walk trajectories. We find that including both a continuous prior and Poisson noise improves the ability to infer ground truth latent states and reduces posterior uncertainty (Fig. C). Our method also generalizes to higher dimensional non-Euclidean manifolds. We fit mGPLVM with a toroidal latent and continuous prior to count data from a 2D synthetic head direction circuit and recover ground truth latent states and tuning curves in a completely unsupervised manner (Fig. D). Finally we fit electrophysiological data from the mouse anterodorsal thalamic nucleus [4] using a circular latent. Neural firing is highly overdispersed for this data, and we find that a negative binomial noise model with a continuous prior outperforms a Poisson model and uncovers both the measured head direction and appropriate tuning curves (ΔRMS error $= (0.98 \pm 0.1) \times NT$, Fig. E).



(a) True and inferred latents for mGPLVM fitted on a ring with uniform (U) or continuous (C) priors. **(b)** Tuning curves for synthetic count data fitted with a Gaussian (G) or Poisson (P) noise model. **(c)** Recovery of latent trajectories with U/C priors and G/P noise models for synthetic head direction data (top; mean RMSE ± sem) and example latent posteriors (bottom). **(d)** Latent trajectory (left) and example tuning curves for mGPLVM with a continuous prior and Poisson noise fitted to synthetic count data on a torus. **(e)** Mouse head direction and inferred latents with a negative binomial noise model (left) and four example tuning curves (right).

[1] Jensen et al. 2020 [2] Falorsi et al. 2019 [3] Hensman et al. 2015 [4] Peyrache et al. 2015