Uncovering Neural Encoding Variability with Infinite Gaussian Process Factor Analysis

Anonymous Author(s)
Affiliation
Address
email

Abstract

Gaussian Process Factor Analysis (GPFA) is a powerful factor analysis model for extracting low-dimensional latent processes underlying population neural activities. However, one limitation of standard GPFA models is that the number of latent factors needs to be pre-specified or selected through heuristic-based processes. We propose the infinite GPFA model, a fully Bayesian non-parametric extension of the classical GPFA model by incorporating an Indian Buffet Process (IBP) prior over the factor loading process, such that it is possible to infer the potentially infinite set of likely latent factors active at each time points, in a probabilistically principled manner. Learning and inference in the infinite GPFA model is performed through variational expectation-maximisation, and we additionally propose a scalable extension based on sparse variational Gaussian Process methods. We empirically demonstrate that the infinite GPFA model correctly infers dynamically changing activations of latent factors on synthetic dataset. Through fitting the infinite GPFA model to population activities of hippocampal pyramidal cells during spatial navigation, we identify non-trivial and behavioural meaningful variability in neural encoding process, and interpret neural variability from a novel perspective.

1 Introduction

Latent variable modelling is a popular class of unsupervised approaches for discovering low-dimensional manifolds underlying high-dimensional neural population activities [Churchland et al., 2007, Cunningham and Yu, 2014, Pei et al., 2021]. Accurate inference over the dynamical latent processes allows us to perform exploratory analysis for identifying relevant behavioural correlates of target neuron ensembles. However, a key limitation for such modelling is the necessity for pre-specifying latent dimensions. This is usually performed based on model-selection apporaches, such as cross-validation and various information measures [Doya, 2007]. In the absence of prior knowledge of encoded behavioural covariates underlying target neurons, heuristically selecting the latent manifold dimensions lacks interpretability, and the selection is often sensitive with respect to model hyperparameters and sampling process, hence leading to inconsistent inference outcomes. Alternative approaches based on regularisation methods, such as automatic relevance determination [ARD; Wipf and Nagarajan, 2007, Jensen et al., 2021], requires maximum likelihood (ML) learning based on marginalisation over all training samples. Therefore, selection of the set of latent factors that are mostly likely accounting for *all* observations, but not for *each* observation.

Here we propose a novel, probabilistically principled model that enables simultaneous posterior inference of the number of latent factors and the set of activated latent factors pertinent to each observation. Specifically, we develop a fully Bayesian nonparametric extension of the Gaussian Process Factor Analysis (GPFA) model [Yu et al., 2008], a popular latent variable model for extracting latent Gaussian process factors underlying population activities over single trials. The resulting model,

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

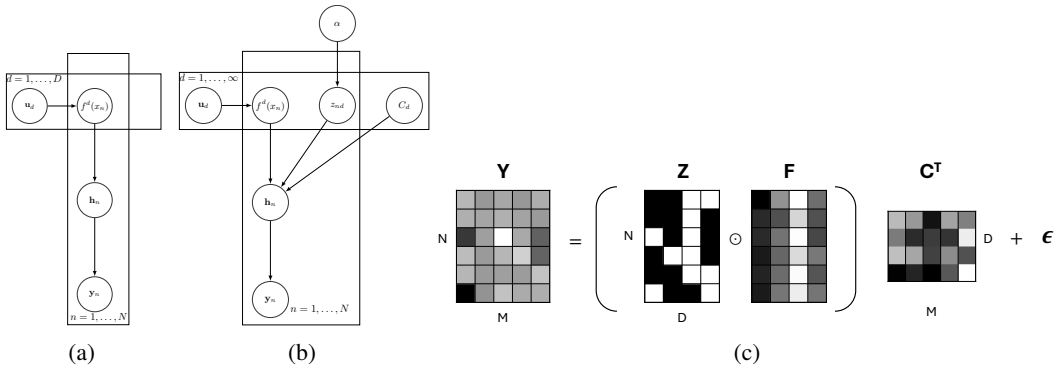


Figure 1: **Graphical demonstration of generative processes of GPFA and IBP models.** Generative models for standard **(a)** and infinite **(b)** GPFA models with sparse variational approximation. **(c)** Graphical illustration of IBP prior, in the form of weighted factor analysis model, with binary activations. Each observation $\mathbf{y}_n$ is generated as a weighted sum of different set of latent factors with some additive noise. By taking the limit $D \to \infty$, we essentially place an IBP prior on the binary latent activation, $\mathbf{Z}$.

infinite GPFA, incorporates stochastic activation of latent factors in the loading process, which is modelled by the Indian Buffet Process (IBP) prior [Ghahramani and Griffiths, 2005]. The IBP defines a distribution over binary matrices with finite number of rows and infinite number of columns, hence enabling inference over the potentially infinite number of features, as well as tracking uncertainty associated with factor activations for each observation. Importantly, the latter feature allows us to investigate the nature of neural variability from a novel perspective: the variability in the expression of latent factors, potentially due to changes in internal states of the animal [Kelemen and Fenton, 2010, Flavell et al., 2022]. Through empirical evaluations on synthetic datasets, we show that the infinite GPFA model yields similar performance as standard GPFA model on dataset with constant generative process, but significantly outperforms GPFA when variability is introduced to the generative process. We further apply our model to population activities of hippocampal place cells recorded during spatial navigation tasks, and identify non-trivial variability in the neural encoding process, which is additionally contingent on the engaged task context.

2 Background

2.1 Gaussian Process Factor Analysis

GPFA extends standard factor analysis models, by replacing Gaussian factors with Gaussian Process factors for capturing non-trivial temporal dependencies over the latent space [Yu et al., 2008]. Similar to standard factor analysis model, GPFA assumes conditional independence between observation dimensions, given the latents. The generative model of GPFA is defined as following (Figure 1a).

$$\begin{aligned} \mathbf{f}_d(\cdot) &\sim \mathcal{GP}\left(m^d(\cdot), k^d(\cdot,\cdot)\right), && \text{for } d = 1, \ldots .D, \\ \mathbf{h}(x_n) &= \mathbf{C} \cdot \mathbf{F}(x_n) + \mathbf{d}, && \text{for } n = 1, \ldots, N, \\ \mathbf{y}(x_n) &\sim p\left(\mathbf{y}(x_n) | \phi(\mathbf{h}(x_n)), \theta\right), && \text{for } n = 1, \ldots, N, \end{aligned} \tag{1}$$

where $m^d(\cdot)$ and $k^d(\cdot,\cdot)$ are the mean and kernel functions for the d-th latent factors, respectively[1], $\mathbf{C} \in \mathbb{R}^{M \times D}$ is the loading matrix that projects the latent factors to the neural space, with M being the number of neurons, $\mathbf{d}$ is the offset for the linear transformation, $\mathbf{F}(x_n) = [\mathbf{f}_1(x_n) \quad \cdots \quad \mathbf{f}_D(x_n)]$ is the column-stack of all latent factors at input location x_n, $\phi(\cdot)$ is some (non-linear) link function, and θ is some auxiliary generative parameters. Learning and inference with GPFA model can be performed using variational expectation-maximisation (EM), which we briefly review in the appendix.

[1]We assume $m^d(\cdot) = 0$ unless stated otherwise.

2.2 Indian Buffet Process

The IBP defines a distribution over binary matrices with finite number of rows (observations) and infinite number of columns (latent factors) [Ghahramani and Griffiths, 2005]. Hence, by factorising the loading process in a factor analysis model into independent binary activation matrix, $\mathbf{Z}$, and activation weight matrix, $\mathbf{C}$, we can place an IBP prior over $\mathbf{Z}$ for allowing stochastic loading of latent factors into each observation. Moreover, posterior inference over $\mathbf{Z}$ allows for determination of the optimal set of latent factors pertinent to each observation, in a probabilistically principled manner. For interpretation, we consider the following Gaussian factor analysis model with stochastic latent activations (Figure 1c).

$$
\begin{aligned}
& f_d \sim \mathcal{N}(0, \sigma_d^2), \mathbf{C}_d \sim \mathcal{N}(\mathbf{0}, \nu_d^2 \mathbf{I}), && \text{for } d = 1, \ldots, \infty\,, \\
& \pi_d \sim \text{Beta}(\frac{\alpha}{D}, 1), && \text{for } d = 1, \ldots, D\,, \\
& p(\mathbf{Z}|\boldsymbol{\pi}) = \prod_{d=1}^{D} \pi_d^{m_d} (1 - \pi_d)^{N - m_d}\,, && \\
& \mathbf{y}_n = \mathbf{C}(\mathbf{Z}_n \odot \mathbf{F}) + \boldsymbol{\epsilon}_n\,, && \text{for } n = 1, \ldots, N\,.
\end{aligned}
\tag{2}
$$

We observe that $p(Z)$ models the probability for the n-th observation possessing the k-th factor, for all n and k. Taking the limit $D \to \infty$, it can be shown that the marginal distribution over $\mathbf{Z}$ following the IBP distribution, and α controls the expected total number of latent factors (see details in the appendix). Posterior inference over the IBP-distributed $\mathbf{Z}$ is intractable, but it is possible to perform approximate inference leveraging either MCMC or variational methods [Ghahramani and Griffiths, 2005, Doshi et al., 2009]. Here we use the mean-field variational inference approach, which we briefly review in the appendix Doshi et al. [2009].

3 Infinite GPFA

Under the similar motivation behind the original proposal of IBP, we now propose infinite GPFA, the fully Bayesian nonparametric extension of standard GPFA that allows simultaneous inference over the optimal number of latent features and the set of most likely active latent factors underlying *each* observation. Specifically, the generative process of infinite GPFA is as following (Figure 1b).

$$
\begin{aligned}
& \mathbf{f}_d(\cdot) \sim \mathcal{GP}\left(0, k^d(\cdot, \cdot)\right)\,, \quad \pi_d \sim \text{Beta}\left(\frac{\alpha}{D}, 1\right)\,, \quad z_{nd}|\pi_d \sim \text{Bernoulli}(\pi_d)\,, \\
& \mathbf{C}_d \sim \mathcal{N}(\mathbf{0}, \nu_d^2 \mathbf{I})\,, \quad \mathbf{h}(x_n) = \mathbf{C} \cdot (\mathbf{Z} \odot \mathbf{F}(x_n)) + \mathbf{d}\,, \quad \mathbf{y}(x_n) \sim p(\mathbf{y}(x_n)|\phi(\mathbf{h}(x_n)))\,, \quad \forall n, d\,,
\end{aligned}
\tag{3}
$$

Here we use the finite Beta-Bernoulli approximation of the IBP distribution (Equation 2; [Ghahramani and Griffiths, 2005]). For the simplicity of demonstration, we assume both the loading weight matrix, $\mathbf{C}$, and concentration parameter, α, to be deterministic (but setting priors over $\mathbf{C}$ and α is also possible, see Section 4.1 and appendix for further details). Note that the major difference between the generative processes of standard and infinite GPFA lies in their implementation of factor loading process, where standard GPFA assumes each latent GP factor is deterministically loaded into the observations, and infinite GPFA allows stochastic binary expression of latent factors that varies across each timestep and observation. Both learning of generative parameters and approximate inference over latent variables ($\mathbf{f}$, $\boldsymbol{\pi}$ and $\mathbf{Z}$) in the infinite GPFA model is achieved through variational learning, leveraging mean-field variational approximations. We further develop a sparse-variational extension of the infinite GPFA (infinite svGPFA), which greatly improves scalability. Further mathematical details of model learning and inference can be found in the appendix due to space constraints.

4 Results

4.1 Empirical Evaluation on Synthetic Data

We first consider synthetic population spikings generated from two sinusoidal latent processes, following the generic GPFA generative process with exponential link function and Poisson conditional likelihood (Equation 1). To demonstrate variations in neural encoding within a single trial, we

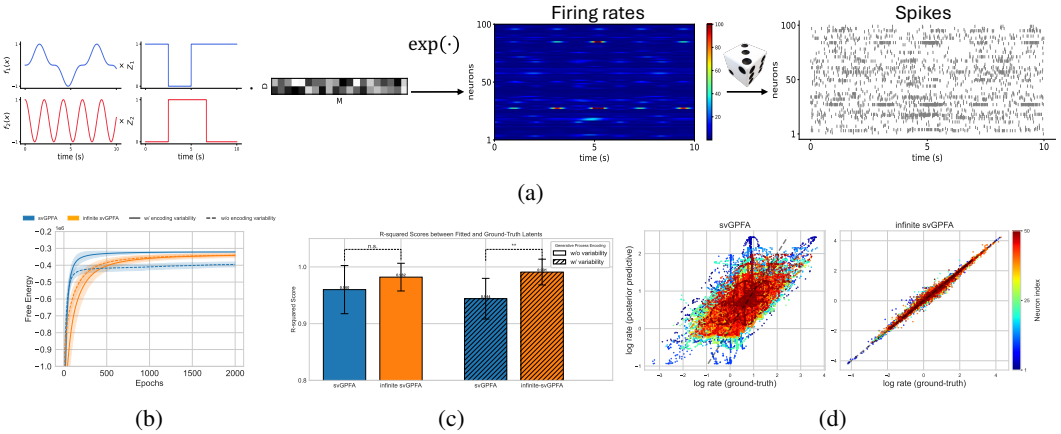


Figure 2: **Empirical evaluation of infinite svGPFA on synthetic dataset.** (**a**) Generative process for synthetic data, following the standard GPFA generative model with sinusoidal latent processes ($f_1(x) = \cos^3(x)$ and $f_2(x) = \sin(3x)$) and (optional) binary masking. The latents are linearly projected to the neural space, and passed through an exponential link function to generate firing rates, which are then used to generate spikes following time-inhomogeneous Poisson process. The binary masking, Z, represents within-trial variability in expression of latent factors in the neural activities. (**b**) Variational free energy objective during training for different models. (**c**) R-squared score between posterior means over latent processes and ground-truth latents, for svGPFA (blue) and infinite svGPFA (orange), on data given both generative processes (with and without encoding variability). (**d**) Log-log plot between logarithm of predicted and ground-truth firing rates for svGPFA (left) and infinite svGPFA (right). Different color represents different neurons. All evaluations are performed based on averaging over 10 random seeds where applicable.

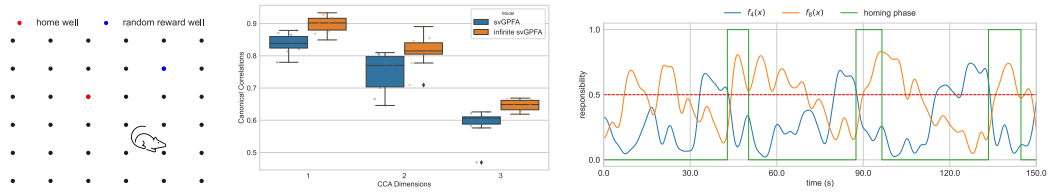
optionally apply a multiplicative binary mask to the latent processes before projecting them to the neural space (Equation 3). We generate synthetic data for 100 neurons over 10 trials, each lasting 10 seconds in duration, for both cases with and without encoding variability (Figure 2a).

Under both generative processes, we fit standard and infinite svGPFA to corresponding population activities. We place Gaussian and Gamma priors over $\mathbf{C}$ and α, respectively, hence additional marginalisation over them is required to compute the posterior distribution over latent processes (corresponding prior parameters are identical between svGPFA and infinite svGPFA where applicable). Both methods converge quickly under either data generative process (Figure 2b). Upon training completion[2], to validate the fidelity of fitted latents, we compare the R-squared score between the posterior means over the latent processes and the ground-truth latents for both models. For the baseline case with trivial binary masks, we observe that both models perform comparably well, reaching almost perfect discovery of latent processes driving the generation of neural activities. When encoding variability is introduced to binary masking, we observe that inferred latents of infinite svGPFA explains the ground-truth latents significantly better than those of standard svGPFA (one-sided student-t test, $p = 0.0028$). Such performance difference in model fitting is exacerbated through examining the accuracy of predicted firing rate: the svGPFA prediction is significantly noisier than the infinite svGPFA prediction, and the mean squared error of predicted log-rates is 0.40 ± 0.87, which is again significantly higher than infinite svGPFA (0.0043 ± 0.025). Absence of explicit mechanisms accounting for variability in factor loading process in standard GPFA leads to greater deficits in learning the correct generative process. This is due to increased prediction errors in spiking observations induced by periods when at least one of the factors is not activated, which leads to learning of the wrong generative parameters to account for the gap.

4.2 Variability in Neural Encoding in Multi-Phase Spatial Navigation Tasks

We now probe the existence of variability in neural encoding and potential behavioural implications in real neural recordings. We apply our model to simultaneously recorded population activities of 204 pyramidal cells recorded from rat dorsal hippocampal CA1, whilst the rat is performing a spatial memory task [Pfeiffer and Foster, 2013]. Within each trial, given 36 uniformly arranged feeding

[2]All hyperparameters used in training can be found in Appendix



(a) (b) (c)

Figure 3: **Probing within-trial encoding variability in place cell population activities during spatial navigation with alternating behavioural phases.** (**a**) Illustration of behavioural task [Pfeiffer and Foster, 2013]. Rats navigate in a $2m \times 2m$ box, with 36 feeding wells uniformly arranged in the box. Animals alternate between searching for reward in a random well (*foraging* phase), and navigate back to a home well (*homing* phase). (**b**) We perform CCA between posterior mean over latent processes and selected behavioural variables for both infinite svGPFA and standard svGPFA. We show comparison of first three canonical correlations for the two models (dots represent different random seeds). (**c**) Temporal trace of posterior responsibilities associated with selected latent processes, and binary behavioural phase (green line, 0 and 1 indicate foraging and homing phases, respectively).

wells within a $2m \times 2m$ open-field arena (Figure 3a), rats learn to alternate between foraging for food in an unknown and random location (*foraging* phase), and returning to a fixed home location (*homing* phase). The transition to the next phase or trial is automatic upon consumption of the reward. We fit both standard and infinite svGPFA, with 10-dimensional latents to one recording session lasting 2187 seconds, binning spike trains into spike counts within each 30 ms time window.

We perform canonical correlation analysis (CCA) between posterior mean of latent factors and relevant behavioural variables, including the 2-dimensional allocentric location, speed, and head direction of the animal [Hardoon et al., 2004, O'keefe and Nadel, 1978, Geisler et al., 2007]. Conforming with our findings from the synthetic experiment, we found that inferred latents from the infinite svGPFA model comprise more faithful representations of behvaioural covariates than those from the standard svGPFA, indicated by the higher canonical correlations over all three principal directions.

We examine the activation of each latent factors across all timesteps. We observe high variability over time in posterior responsibilities for each latent (Figure 3c). Hence, despite stationarity in the marginal distribution of behavioural variables, the infinite svGPFA model predicts that the expression of these variables in population neural activities is not deterministic over time. By separating the continuous recording into alternating homing and foraging phases, we identify latent processes exhibiting selective activations in accordance with different behavioural phases. Specifically, we observe one latent process, $f_4(x)$, is usually increasingly activated during foraging phases and deactivated during homing phases (Figure 3c). From standard correlational analysis, we found that $f_4(x)$ is most strongly correlated with the speed of the animal. We identify another latent process, $f_8(x)$, being most strongly correlated with spatial location of the animal, which is activated at the beginning of homing phases, and decreasingly activated over the foraging phase. These comprise a coherent interpretation: speed is more actively represented during random foraging, potentially due to the importance of speed information in path integration (especially given extended trajectories), whereas during homing phases, the rat is usually running in straight trajectories back to the home location, potentially leveraging a pre-fixed strategy, hence leading to decreased representation of speed information, but increased representation of allocentric spatial location, in population activities. Collectively, we identify non-trivial temporal variability in encoding of behavioural correlates in population neural activities, and show that such variability is mediated by behavioural state of the animal through empirical verification.

5 Discussion

We introduce the infinite GPFA, a fully Bayesian nonparametric generalisation of standard GPFA models. The incorporation of the IBP prior over latent activations enables simultaneous inference over both the number of latent factors, as well as the most likely active set of latent factors underlying each observation. Through extensive evaluations on both synthetic and real neural datasets, we demonstrate improved empirical performance comparing to standard GPFA models. More importantly, we show that the infinite GPFA model is suited for exploring a gap in interpreting neural variability: the variability in neural encoding arising from changes in internal states of the animal.

References

Mark M Churchland, M Yu Byron, Maneesh Sahani, and Krishna V Shenoy. Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Current opinion in neurobiology*, 17(5):609–618, 2007.

John P Cunningham and Byron M Yu. Dimensionality reduction for large-scale neural recordings. *Nature neuroscience*, 17(11):1500–1509, 2014.

Felix Pei, Joel Ye, David Zoltowski, Anqi Wu, Raeed H Chowdhury, Hansem Sohn, Joseph E O'Doherty, Krishna V Shenoy, Matthew T Kaufman, Mark Churchland, et al. Neural latents benchmark'21: evaluating latent variable models of neural population activity. *arXiv preprint arXiv:2109.04463*, 2021.

Kenji Doya. *Bayesian brain: Probabilistic approaches to neural coding*. MIT press, 2007.

David Wipf and Srikantan Nagarajan. A new view of automatic relevance determination. *Advances in neural information processing systems*, 20, 2007.

Kristopher Jensen, Ta-Chu Kao, Jasmine Stone, and Guillaume Hennequin. Scalable bayesian gpfa with automatic relevance determination and discrete noise models. *Advances in Neural Information Processing Systems*, 34:10613–10626, 2021.

Byron M Yu, John P Cunningham, Gopal Santhanam, Stephen Ryu, Krishna V Shenoy, and Maneesh Sahani. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *Advances in neural information processing systems*, 21, 2008.

Zoubin Ghahramani and Thomas Griffiths. Infinite latent feature models and the indian buffet process. *Advances in neural information processing systems*, 18, 2005.

Eduard Kelemen and André A Fenton. Dynamic grouping of hippocampal neural activity during cognitive control of two spatial frames. *PLoS biology*, 8(6):e1000403, 2010.

Steven W Flavell, Nadine Gogolla, Matthew Lovett-Barron, and Moriel Zelikowsky. The emergence and influence of internal states. *Neuron*, 110(16):2545–2570, 2022.

Finale Doshi, Kurt Miller, Jurgen Van Gael, and Yee Whye Teh. Variational inference for the indian buffet process. In *Artificial Intelligence and Statistics*, pages 137–144. PMLR, 2009.

Brad E Pfeiffer and David J Foster. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79, 2013.

David R Hardoon, Sandor Szedmak, and John Shawe-Taylor. Canonical correlation analysis: An overview with application to learning methods. *Neural computation*, 16(12):2639–2664, 2004.

John O'keefe and Lynn Nadel. *The hippocampus as a cognitive map*. Oxford university press, 1978.

Caroline Geisler, David Robbe, Michaël Zugaro, Anton Sirota, and György Buzsáki. Hippocampal place cell assemblies are speed-controlled oscillators. *Proceedings of the National Academy of Sciences*, 104(19):8149–8154, 2007.

Michalis Titsias. Variational learning of inducing variables in sparse gaussian processes. In *Artificial intelligence and statistics*, pages 567–574. PMLR, 2009.

Lea Duncker and Maneesh Sahani. Temporal alignment and latent gaussian process factor inference in population spike trains. *Advances in neural information processing systems*, 31, 2018.

Stephen Keeley, David Zoltowski, Yiyi Yu, Spencer Smith, and Jonathan Pillow. Efficient non-conjugate gaussian process factor models for spike count data using polynomial approximations. In *International conference on machine learning*, pages 5177–5186. PMLR, 2020.

Yee Whye Teh, Dilan Grür, and Zoubin Ghahramani. Stick-breaking construction for the indian buffet process. In *Artificial intelligence and statistics*, pages 556–563. PMLR, 2007.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

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

A Further Model Details

A.1 Variational Learning for GPFA

We consider the standard GPFA generative process (Equation 1). We use variational expectation-maximisation (EM) methods for approximate inference over latent processes in GPFA. Specifically, we leverage the mean-field sparse-variational approximation based on inducing points for scalability purposes [Titsias, 2009], which renders desirable conditional independence in the variational free energy objective.

$$q(\mathbf{F}, \mathbf{U}) = \prod_{d=1}^{D} p(\mathbf{f}_d|\mathbf{u}_d)q(\mathbf{u}_d)\,, \quad \mathcal{F}[q] = \sum_{x} \langle \log p(\mathbf{y}|\phi(\mathbf{h}))\rangle_{q(\mathbf{h})} - \sum_{d=1}^{D} \mathrm{KL}\left[q(\mathbf{u}_d)||p(\mathbf{u}_d)\right]\,, \quad (4)$$

where $\mathbf{u}_d$ are the inducing points for the d-th latent factor.

From standard Gaussian identity, we know that the conditional likelihood $p(\mathbf{f}_d(\mathbf{x})|\mathbf{u}_d)$ is also Gaussian, with mean $\mathbf{K}^d_{\mathbf{xw}}(\mathbf{K}^d_{\mathbf{wz}})^{-1}\mathbf{u}_d$ and covariance $\mathbf{K}^d_{\mathbf{xx}} - \mathbf{K}^d_{\mathbf{x,w}}(\mathbf{K}^d_{\mathbf{ww}})^{-1}(\mathbf{K}^d_{\mathbf{xw}})^T$, where $\mathbf{K}^d_{\mathbf{x,w}} \in \mathbb{R}^{N\times S}$ such that $\left(\mathbf{K}^d_{\mathbf{x,w}}\right)_{nd} = k^d(x_n w_d)$. Hence, given $q(\mathbf{u}_d) = \mathcal{N}(\boldsymbol{\mu}^u_d, \mathbf{S}^u_d)$, we could easily compute the marginal variational approximation for $\mathbf{f}$, $q(\mathbf{f}_d) = \mathcal{N}(\boldsymbol{\mu}^f_d, \mathbf{S}^f_d)$.

$$\mu^f_{nd} = k^d(x_n, \mathbf{w})\left(\mathbf{K}^d_{\mathbf{ww}}\right)^{-1}\boldsymbol{\mu}^f_d\,, \quad (s^f_{nd})^2 = k^d_{nn} + k^d_{n\mathbf{w}}\left((\mathbf{K}^d_{\mathbf{ww}})^{-1}\mathbf{S}^u_d(\mathbf{K}^d_{\mathbf{ww}})^{-1} - (\mathbf{K}^d_{\mathbf{ww}})^{-1}\right)k^d_{\mathbf{w}n}\,, \quad (5)$$

Note that $q(\mathbf{h})$ is additively GP-distributed (Equation 1). In general, the expected log conditional likelihood can only be evaluated approximately [Duncker and Sahani, 2018, Keeley et al., 2020]. However, it is possible to compute the expected log conditional-likelihood under certain assumptions of conditional likelihood and link function (e.g., Gaussian observation and identity link function). The KL divergence between the variational approximation and GP prior over the inducing points can be evaluated analytically.

A.2 Variational Inference for IBP

We re-iterate the weighted factor analysis generative process below.

$$\begin{aligned}
&f_d \sim \mathcal{N}(0, \sigma^2_d), \mathbf{C}_d \sim \mathcal{N}(\mathbf{0}, \nu^2_d\mathbf{I}), && \text{for } d = 1, \ldots, \infty\,,\\
&\pi_d \sim \text{Beta}(\frac{\alpha}{D}, 1), && \text{for } d = 1, \ldots, D\,,\\
&p(\mathbf{Z}|\boldsymbol{\pi}) = \prod_{d=1}^{D} \pi_d^{m_d}(1-\pi_d)^{N-m_d}\,, &&\\
&\mathbf{y}_n = \mathbf{C}(\mathbf{Z}_n \odot \mathbf{F}) + \boldsymbol{\epsilon}_n\,, && \text{for } n = 1, \ldots, N\,.
\end{aligned} \quad (6)$$

Given the conjugacy between Beta and binomial distributions, we can analytically marginalise π out.

$$p(\mathbf{Z}) = \prod_{d=1}^{D} \frac{\frac{\alpha}{D}\Gamma(m_d + \frac{\alpha}{D})\Gamma(N - m_d + 1)}{\Gamma(N + 1 + \frac{\alpha}{D})}\,, \quad (7)$$

Taking the limit $D \to \infty$, the IBP places a prior on $[\mathbf{Z}]$, the canonical form of $\mathbf{Z}$ that is permutation-invariant [Ghahramani and Griffiths, 2005].

$$p([\mathbf{Z}]) = \frac{\alpha^{\mathfrak{D}}\exp(-\alpha H_N)}{\prod_{h\in\{0,1\}^N\setminus\mathbf{0}}\mathfrak{D}_h!}\prod_{d=1}^{\mathfrak{D}}\frac{(N-m_d)!(m_d-1)!}{N!} \quad (8)$$

where $\mathfrak{D}$ is the number of non-zero columns in $\mathbf{Z}$, $H_N = \sum_{n=1}^{N}\frac{1}{n}$ is the N-th harmonic number, m_d is the number of one-entries in the d-th column of $\mathbf{Z}$, $\mathfrak{D}_h$ is the number of occurrences of non-zero binary column vector h in $\mathbf{Z}$, α is the prior parameter that controls the expected number of features present in each observation.

A useful interpretation of IBP is based on the stick-breaking formulation [Teh et al., 2007], which interprets π_d being constructed by stick-breaking weights, $\pi_d = \prod_{i=1}^{d} v_d$, where $v_d \sim \text{Beta}(\alpha, 1)$. We hence see that the probability of employing the $d-$th latent factor decreases exponentially with d, and α controls the expected number of latents.

Inference given the IBP prior can be performed with either MCMC or variational methods [Ghahramani and Griffiths, 2005, Doshi et al., 2009]. Here we briefly review the finite mean-field variational inference approach outlined in Doshi et al. [2009].

$$\begin{aligned} q(\pi_d|a_d, b_d) &= \text{Beta}(a_d, b_d), \quad \forall d\,, \\ q(\mathbf{C}_d|\boldsymbol{\mu}_d, \mathbf{S}_d) &= \mathcal{N}(\boldsymbol{\mu}_d, \mathbf{S}_d), \quad \forall d\,, \\ q(z_{nd}|\tau_{nd}) &= \text{Bernoulli}(\tau_{nd}), \quad \forall n, d\,, \end{aligned} \tag{9}$$

Given the conditional independence within the generative model, the variational free energy objective takes the following expression[3].

$$\begin{aligned} \mathcal{F}[q] &= \langle \log p(\boldsymbol{\pi}, \mathbf{C}, \mathbf{Z}, \mathbf{Y}) - \log q(\boldsymbol{\pi}) q(\mathbf{C}) q(\mathbf{Z}) \rangle \\ &= \sum_{d=1}^{D} \langle \log p(\pi_d) \rangle + \sum_{d=1}^{D} \langle \log p(\mathbf{C}_d) \rangle + \sum_{n=1}^{N} \sum_{d=1}^{D} \langle \log p(z_{nd}|\pi_d) \rangle + \sum_{n=1}^{N} \langle \log p(\mathbf{y}_n|\mathbf{Z}_n, \mathbf{C}) \rangle + \mathbb{H}[q] \end{aligned} \tag{10}$$

A.3 Variational Learning for Infinite GPFA

We perform variational learning using the finite mean-field variational approximations, $q(\mathbf{U}, \boldsymbol{\pi}, \mathbf{Z}) = \prod_{d=1}^{D} [q(\mathbf{u}_d) q(\pi_d) \prod_n p(z_{nd})]$.

$$\begin{aligned} q(\mathbf{u}_d|\boldsymbol{\mu}_d^u, \mathbf{S}_d^u) &= \mathcal{N}(\mathbf{u}_d|\boldsymbol{\mu}_d^u, \mathbf{S}_d^u)\,, \quad \forall d\,, \\ q(\pi_d|a_d, b_d) &= \text{Beta}(\pi_d|a_d, b_d)\,, \quad \forall d\,, \\ q(z_{nd}|\tau_{nd}) &= \text{Bernoulli}(\tau_{nd})\,, \quad \forall n, d\,, \end{aligned} \tag{11}$$

Note that in the above formulation, by default we have assumed sparse variational approximation treatment for scalability purposes.

Given the conditional independence in the generative process, we could express the variational free energy as following.

$$\begin{aligned} \mathcal{F}[q] &= \langle \log p(\mathbf{Y}, \mathbf{F}, \boldsymbol{\pi}, \mathbf{Z}) - \log q(\mathbf{F}, \boldsymbol{\pi}, \mathbf{Z}) \rangle \\ &= \sum_{n=1}^{N} \langle \log p((\mathbf{y}_n|\mathbf{F}_n, \mathbf{Z}_n)) \rangle - \sum_{d=1}^{D} \text{KL}\left[q(\mathbf{u}_d)||p(\mathbf{u}_d)\right] - \sum_{d=1}^{D} \text{KL}\left[q(\pi_d)||p(\pi_d)\right] - \sum_{n,d} \langle \text{KL}\left[q(z_{nd})||p(z_{nd})\right] \rangle_{q(\pi_d)} \end{aligned} \tag{12}$$

Given the variational distributions, all terms always admit analytical expression apart from expected conditional log-likelihoods. Due to the non-Gaussian nature of $q(\mathbf{h})$, previous approximation approaches based on Gaussian quadrature no longer applies [Duncker and Sahani, 2018]. Instead, we leverage second-order Taylor expansion for approximating the expected conditional log-likelihood, which offers an effective tradeoff between computational efficiency and approximation accuracy (see Supplemental Section 2 for details). However, we note that under the special case of Gaussian conditional likelihood with identity link function, it is possible to evaluate such expectation analytically.

$$\langle \log p((\mathbf{y}_n|\mathbf{F}_n, \mathbf{Z}_n)) \rangle = -\frac{1}{2\sigma^2} \sum_{m=1}^{M} \langle (y_{nm} - h_{nm})^2 \rangle = -\frac{1}{2\sigma^2} \sum_{m=1}^{M} \left((y_{nm} - \langle h_{nm} \rangle)^2 + \text{Var}[h_{nm}] \right)\,, \tag{13}$$

[3] Unless necessary, we do not explicitly show the variational distributions the expectation is taken with respect to for notational simplicity.

Given the mean-field assumption, we could analytically evaluate the expectation and variance of h_{nm} with respect to the variational distributions.

$$\begin{aligned}\langle h_{nm} \rangle &= \mathbf{C}_m \cdot (\boldsymbol{\tau}_n \odot \boldsymbol{\mu}_n^f) + d_m \,, \\ \text{Var}[h_{nm}] &= \mathbf{C}_m^{\odot 2} \cdot \left(\boldsymbol{\tau}_n^2 \odot (\mathbf{s}_n^f)^2 + ((\boldsymbol{\mu}_n^f)^{\odot 2} + (\mathbf{s}_n^f)^{\odot 2}) \odot \boldsymbol{\tau}_n \odot (1 - \boldsymbol{\tau}_n) \right)\end{aligned} \tag{14}$$

where $\boldsymbol{\mu}_n^f$ and $(\mathbf{s}_n^f)^2$ are the mean and diagonal-variance of $\mathbf{F}(x_n)$, respectively, and $^{\odot 2}$ represents the elementwise square operation. We have leveraged the law of total variance, $\text{Var}[XY] = \mathbb{E}\left[\text{Var}[XY]\right] + \text{Var}\left[\mathbb{E}[XY]\right]$. The complete derivation of variational free energy for the finite variational approach can be found in Supplemental Section 2.

The model is learned via variational EM, iteratively updating the variational parameters (Equation 11), and the generative model parameters (i.e., $\mathbf{C}$, $\mathbf{d}$ and α), via gradient-based updates that maximises the free energy objective. In practice, we employ standard automatic differentiation framework for such gradient-based learning [Paszke et al., 2019].

B Experiment Details

All models are trained with Adam optimiser [Kingma and Ba, 2014], with learning rate 0.01. For the main experimental evaluations, we train all models over 2000 epochs. All evaluations are based on averaging over 10 random seeds where applicable.

Synthetic Data. We instantiate both the standard GPFA and infinite GPFA models with stochastic $\mathbf{C}$, where $\nu_d^2 = 0.1$. We set the number of inducing points to be 30 for the main evaluations, and the corresponding inducing locations are randomly initialised and trained. For all models, we use the squared exponential (SE) kernels, with trainable scale and lengthscale parameters.

$$k^d(x, x') = s_d^2 \exp\left(-\frac{||x - x'||}{\tau_d^2} \right) \,, \tag{15}$$

Neural Data. We preprocess the spiking train data into spike counts, with $30ms$ time window. The instantaneous firing rates for each neuron are computed via dividing the spike counts by the time window size, followed by Gaussian smoothing. The loading matrix, $\mathbf{C}$, is assumed to be deterministic, hence is learned through the variational M-step. The concentration parameters, α, is again assumed to be stochastic, with Gamma prior and parameters $s_1 = 1.0$, $s_2 = 1.0$. For all models, the number of inducing points are 100, and corresponding inducing locations are fixed as equally spaced location along the input (time) domain. We again use the SE kernels for the latent GPs with trainable scale and lengthscale parameters.

For segmenting the continuous recordings into separate foraging and homing phases, we note that animals often lowers their speed upon consuming the food, and we can use this feature as a marker for the segmentation. Hence, we identify all periods with low speed ($< 1cm/s$) and within proximity of the reward location ($< 5cm$) as the end of the homing phase, and all other periods with speed ($< 1cm/s$) over an extended time span ($> 10s$) as the end of the foraging phase.