

---

# Identifying signal and noise structure in neural population activity with Gaussian process factor models

---

**Stephen L. Keeley**  
Princeton Neuroscience Institute  
Princeton University,  
Princeton, NJ 08544  
skeeley@princeton.edu

**Mikio C. Aoi**  
Princeton Neuroscience Institute  
Princeton University,  
Princeton, NJ 08544

**Yiyi Yu**  
Dept. of Electrical  
Computer Engineering  
University of California Santa Barbara  
Santa Barbara, CA

**Spencer L. Smith**  
Dept. of Electrical  
Computer Engineering  
University of California Santa Barbara  
Santa Barbara, CA

**Jonathan W. Pillow**  
Princeton Neuroscience Institute  
Princeton University,  
Princeton, NJ 08544

## Abstract

Neural datasets often contain measurements of neural activity across multiple trials of a repeated stimulus or behavior. An important problem in the analysis of such datasets is to characterize systematic aspects of neural activity that carry information about the repeated stimulus or behavior of interest, which can be considered “signal”, and to separate them from the trial-to-trial fluctuations in activity that are not time-locked to the stimulus, which for purposes of such analyses can be considered “noise”. Gaussian Process factor models provide a powerful tool for identifying shared structure in high-dimensional neural data. However, they have not yet been adapted to the problem of characterizing signal and noise in multi-trial datasets. Here we address this shortcoming by proposing “signal-noise” Poisson-spiking Gaussian Process Factor Analysis (SNP-GPFA), a flexible latent variable model that resolves signal and noise latent structure in neural population spiking activity. To learn the parameters of our model, we introduce a Fourier-domain black box variational inference method that quickly identifies smooth latent structure. The resulting model reliably uncovers latent signal and trial-to-trial noise-related fluctuations in large-scale recordings. We use this model to show that predominantly, noise fluctuations perturb neural activity within a subspace orthogonal to signal activity, suggesting that trial-by-trial noise does not interfere with signal representations. Finally, we extend the model to capture statistical dependencies across brain regions in multi-region data. We show that in mouse visual cortex, models with shared noise across brain regions out-perform models with independent per-region noise.

## 1 Introduction

Recent advances in electrophysiological and calcium fluorescence imaging technologies have enabled the collection of increasingly high-dimensional neural datasets. Making sense of such datasets will rely on the development of flexible statistical methods for extracting relevant structure. Gaussian process factor models provide one powerful tool for identifying low-dimensional latent structure from high-dimensional neural response data. These models seek to characterize neural time-series data in terms of a small number of smoothly evolving latent variables, and have been successfully used to characterize neural representations in a variety of contexts [1, 2, 3, 4, 5, 6].

Standard Gaussian Process factor analysis (GPFA) uses a Gaussian process prior to impose smoothness on inferred latent variables, but do not explicitly consider stimulus or task conditions. However, neural data often exist in the form of repeated trials, whereby the same condition is presented to an animal multiple times. These repeated presentations give rise to neural activity that varies across trials around some time-varying “signal” component that is typically estimated using the peri-stimulus time histogram (PSTH). Understanding this signal, and its relationship to trial-to-trial variability, is of central importance to the models of coding in the nervous system [7, 8], yet latent factor models have not been developed to explicitly study this question. Here we address this shortcoming by developing an extension to Gaussian process factor analysis with Poisson spiking (P-GPFA) which we call signal and noise P-GPFA (SNP-GPFA). This model incorporates both signal and independent per-trial components that vary across trials. We refer to these latter components as “noise”, in the sense that they are not time-locked to the repeated stimulus, though they may well reflect other signals unrelated to the experimental stimulus of interest.

In both P-GPFA and SNP-GPFA models, because the Gaussian process is not a conjugate prior for a Poisson observation model, posterior inference is intractable in closed form. Variational inference methods have become increasingly common for applications of Gaussian processes [9, 10, 5]. They achieve tractability by approximating the posterior distribution  $p_{\theta}(\mathbf{x}|\mathbf{y})$  with a well-behaved variational distribution  $q_{\phi}(\mathbf{x}|\mathbf{y})$  [11]. For P-GPFA and SNP-GPFA, because the calculation of the expectation under  $q_{\phi}(\mathbf{x}|\mathbf{y})$  of the joint distribution  $p_{\theta}(\mathbf{x}, \mathbf{y})$  is intractable, we use a ‘black-box’ approach, which works via sampling of the joint distribution [12].

However, black-box variational inference approaches for Gaussian Process Factor models with long time-series can be computationally cumbersome. Therefore, we introduce a variant of black-box variational inference which uses a Fourier-transformed latent representation that factorizes across Fourier modes. This procedure diagonalizes the Gaussian Process (GP) covariance, avoiding a large matrix inversion during inference, thereby providing speed and computational improvements. We demonstrate the inference technique is fast and flexible in a simpler P-GPFA framework, and then use it to learn the SNP-GPFA model quickly and efficiently.

The SNP-GPFA model recovers separate signal and noise subspaces, which allows us to answer a number of scientific questions regarding these facets of neural activity. Here, we address two scientific questions with SNP-GPFA. **1)** We characterize the overlap between signal and noise subspaces, and **2)** We characterize the extent to which noise is shared across cortical region using multi-region neural recordings.

For the first, the alignment of subspaces that reflect different aspects of neural activity has been explored in other contexts [6, 13], as well as the characterization of the subspace of neural noise [14]. Previous work suggests that signal and noise subspaces may be orthogonal [14], and such orthogonal representations may not limit neural information [15]. Our model directly addresses this question. Using SNP-GPFA on primate data we find that there is indeed more noise activity orthogonal to the signal subspace than in the signal subspace, particularly when a visual stimulus is present. This suggests that trial-by-trial variability does not interfere with stimulus encoding.

To address the second scientific question, we include SNP-GPFA analyses on simultaneously recorded visual regions in rodent cortex to ask if trial-varying activity is shared or independent across cortical regions. We compare performance of SNP-GPFA models that varying in their number of shared and independent noise latents across cortical regions. We find that the model that has shared noise latents performs best on cross-validation measures, suggesting trial-by-trial variability has shared structure across cortical regions.

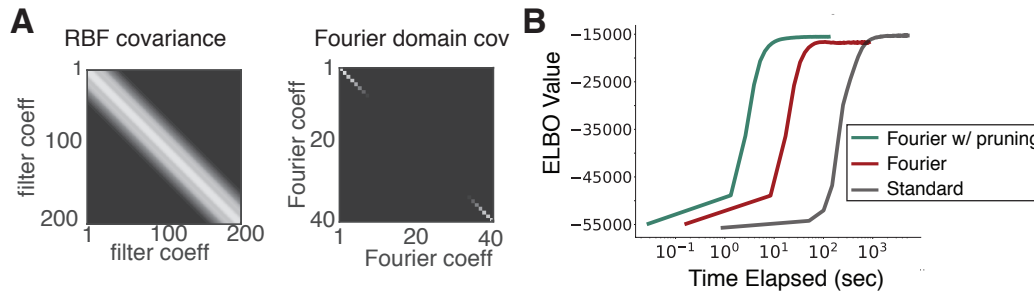


Figure 1: Efficient representation of GP covariance. (A) Standard GP covariance matrix for 1D vectorization of 200 timepoints, with length scale  $\ell = 15$ , and its Fourier representation, pruned. (B) Time to maximization of the ELBO in the time-domain inference and in the Fourier domain with and without a minimum frequency.

## 2 Poisson Gaussian Process Factor Analysis (P-GPFA)

We begin by introducing the Poisson-GPFA model, which has been used previously to identify continuous latent states from population spike train recordings [10, 16, 6]. The observations of our model are spike-train data, represented by the neurons-by-time matrix  $\mathbf{Y} \in \mathbb{N}^{N \times T}$ .

We seek to learn a  $P$ -dimensional latent variable  $\mathbf{x}(t) \in \mathbb{R}^P$  that linearly maps to the data via a loadings matrix  $\mathbf{W} \in \mathbb{R}^{N \times P}$ , followed by some nonlinear function  $f$  and Poisson observations.

$$\mathbf{Y} = \text{Poiss}(f(\mathbf{W}^\top \mathbf{X})) \quad (1)$$

Our choice of non-linear function  $f$  is the softplus  $f(x) = \log(1 + \exp(x))$ .

Each latent  $x_j(t)$  ( $j \in \{1 \dots P\}$ ,  $t \in \{1, 2 \dots T\}$ ) evolves according to a Gaussian process,  $x_j(t) \sim \mathcal{GP}(0, \mathbf{K}(\theta))$ , with covariance matrix  $\mathbf{K}(\theta)$  defined by a squared exponential kernel  $[\mathbf{K}(\theta)]_{tt'} = \rho \exp(-|\mathbf{z}(t) - \mathbf{z}(t')|^2 / 2\ell^2)$ , where hyperparameters  $\theta = \{\ell, \rho\}$  include a length scale  $\ell$  controlling smoothness and a marginal variance  $\rho$  controlling magnitude.

Given that the marginal likelihood of this model,  $p(\mathbf{Y}|\mathbf{W}) = \int p(\mathbf{Y}|\mathbf{W}, \mathbf{X})p(\mathbf{X}|\theta)d\mathbf{X}$  is not available in closed-form, it is common to use a variational inference approach to learn the parameters of such models [16, 10]. Recall that variational inference seeks to maximize an evidence lower bound (ELBO) using a variational distribution [11]. Here, because the expectation term in the ELBO,  $\mathbb{E}_{q_\phi}[\log(p(y|x, w))]$ , cannot be calculated analytically, we employ a 'black box approach' which uses Monte-Carlo samples to estimate the expectation term [17]. This inference method is called black-box variational inference (BBVI).

### 2.1 Fourier-domain black-box variational inference

BBVI can be computationally cumbersome. Therefore, to learn the P-GPFA and SNP-GPFA models, we introduce a novel inference method which performs BBVI over a Fourier-represented latent space, which increases both inference tractability and speed. Factorizing and learning the time series in the Fourier domain, rather than the time domain, allows us to take advantage of computational savings conferred by a diagonal covariance matrix while overcoming problems of uncorrelated timepoints which is typical when the variational distribution is factorized over time [18].

Our motivation for this approach is that the GP prior over  $x_j(t)$  describes a stationary process, as its covariance only depends on pairwise distances. This allows us to diagonalize the covariance  $\mathbf{K}$  by the Fourier transform (Fig 1A). Here, the covariance matrix  $\mathbf{K}$  is diagonalized by  $\tilde{\mathbf{K}} = \mathbf{B}\mathbf{K}\mathbf{B}^\top$  where  $\mathbf{B}$  is the orthonormal discrete Fourier transform matrix with  $[\mathbf{B}]_{\omega, t} = \frac{1}{\sqrt{P}}e^{-i2\pi\omega t/m}$ ,  $i \equiv \sqrt{-1}$ . The diagonalized kernel is represented as  $\tilde{c}(\tilde{\omega}) = \tilde{\rho}e^{-\frac{1}{2}\tilde{\omega}^2\ell^2}$  where  $\tilde{\rho} = \sqrt{2\pi}\rho\ell$  is the frequency-domain variance and  $\tilde{\omega} = \frac{2\pi}{m}\omega$  represents an adjusted frequency of the GP kernel.

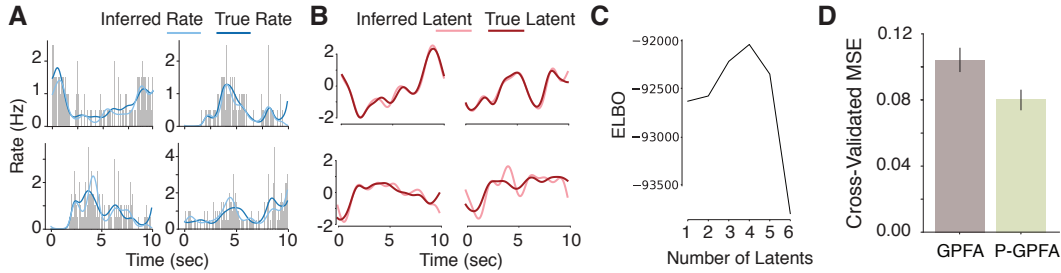


Figure 2: Learning a P-GPFA latent model using Fourier-domain black box variational inference. **(A)** The learned and true rates of four simulated neurons. Grey bars indicate spike PSTHs **(B)** The four generative latents and learned latents rotated to their optimal mapping via regression. **(C)** The maximum ELBO value after inference with a varying number of latent dimensions. The ELBO achieves its maximal value at the true number of latent variables. **(D)** Reconstruction mean-squared error (MSE) on real rodent data using P-GPFA and GPFA.

Inference can be conducted completely in the Fourier domain, precluding the need to invert the prior covariance  $\mathbf{K}$ . The joint likelihood is expressed as

$$\begin{aligned} \ell(\mathbf{Y}, \mathbf{X} | \mathbf{W}, \theta) &= \ell(\mathbf{Y}, \tilde{\mathbf{X}} | \mathbf{W}, \theta) & (2) \\ &\approx \sum_i \log(f(\mathbf{w}_i^\top \tilde{\mathbf{X}} \mathbf{B})) y_i + \left( f(\mathbf{w}_i^\top \tilde{\mathbf{X}} \mathbf{B}) - \log(y_i!) \right) \mathbb{1}_T \\ &\quad - \frac{1}{2} \left( P d \log 2\pi + P \sum_{\tilde{\omega}} \log \tilde{c}_\theta(\tilde{\omega}) + \sum_p \tilde{\mathbf{x}}_p^\top \text{diag}(\tilde{\mathbf{W}}_{n\theta})^{-1} \tilde{\mathbf{x}}_p \right), & (3) \end{aligned}$$

where  $\tilde{\mathbf{X}}$  represents the Fourier-transformed latents and  $\mathbb{1}_T$  is a length- $T$  vector of ones, and  $i \in \{1, 2 \dots N\}$  denotes neuron index. The diagonalized representation demonstrably speeds up computational time (Fig 1 B). Moreover, the inversion of the time-domain  $\mathbf{K}$  can present tractability challenges due to computer precision [19], however, the inversion of  $\tilde{\mathbf{K}}$  is trivial so long as the vector along the diagonal,  $\tilde{\mathbf{W}}_\theta$ , does not contain values that are too small. When small values are present, we regularize  $\tilde{\mathbf{W}}_\theta$  by adding a small constant value ( $10^{-7}$ ).

This Fourier-represented GP has additional computational advantages, including methods to prune unnecessary Fourier coefficients that do not substantially contribute to explaining variability in  $\mathbf{Y}$ . Pruning frequencies constrains the number of coefficients in the Fourier representation to a much smaller number than would be necessary in the time-domain. For a more detailed treatment, see [20]. For our purposes, pruning of the Fourier representation has the consequence of pruning the variational distribution, increasing inference efficiency. Fourier BBVI also preserves time-correlations while permitting a diagonal variational covariance. Ultimately, this Fourier-domain BBVI method can be viewed as an alternative to many other methods that work to make Gaussian processes computationally efficient, such as inducing points and sparse GP approximations [21, 22, 23].

We use our Fourier-domain BBVI to learn the Fourier latents  $\tilde{\mathbf{X}}$  via direct optimization of the variational distribution  $q_\phi(\tilde{\mathbf{X}})$ , factor loading parameters  $\mathbf{W}$ , and hyperparameters  $\ell$ . (Note there is an invariance between hyperparameter  $\rho$ , and the loadings matrix  $\mathbf{W}$ , so we need not directly learn  $\rho$  in this model.) The speed up from optimization with Fourier domain BBVI can be realized most starkly in the domain where the time-series is very long. Figure 1B compares Fourier-domain inference to time-domain inference for a Poisson observation GPFA model with a single latent (i.e.  $P = 1, T = 1500$ ) and  $N = 10$  neurons. Inference is sped up by conversion to the Fourier domain as the bottleneck in time-domain inference is the inversion of a  $1500 \times 1500$  covariance matrix  $\mathbf{K}$ . By additionally specifying a minimum frequency, sufficiently small frequencies are pruned and the variational distribution and prior covariance can be cut from 1500 values to 62. This provides an additional substantial speed advantage of approximately an order of magnitude. It is important to note that the speed-up of our method depends on the specifics of the number of neurons, latents, latent length, and pruning. For subsequent analyses in this paper, the speed-up of BBVI due to the Fourier-domain implementation is anywhere from 20-70%.

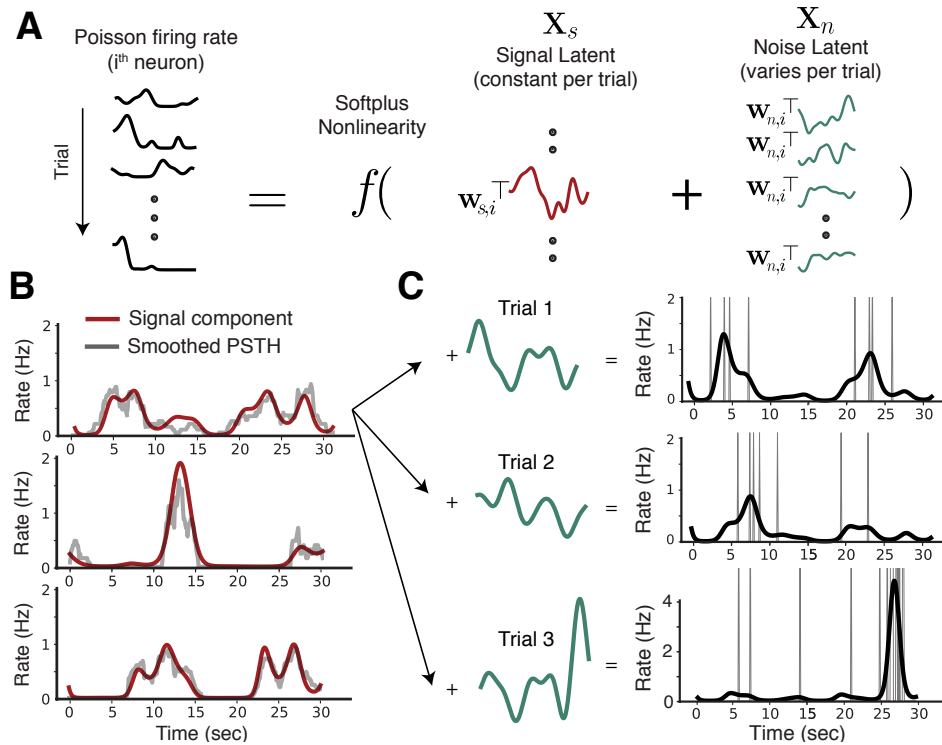


Figure 3: SNP-GPFA model. **(A)** The model describes trial-by-trial variability with signal latents common to all trials and noise latents independent across trials (one dimension for each shown here). Resultant rates are predictive for a single neuron across trials. **(B)** Reconstructed PSTHs for three example rodent V1 neurons given by signal model component. **(C)** Demonstration of trial-by-trial variability explained by noise component for three trials for the top example neuron in **B**.

We verify our Fourier-BBVI inference method using simulated neural spikes drawn from a Poisson distribution with four GP latent dimensions. Our inference approach is able to reconstruct rates of individual neurons as well as latent structure (Fig 2). The inference procedure also identifies correct latent dimensionality, as the ELBO is maximal at the true number of latent dimensions (Fig 2C).

The non-conjugacy of P-GPFA (and SNP-GPFA), and thus the reason we need to use the sophisticated inference of Fourier-BBVI, is due to the fact that the observations are Poisson, as opposed to Gaussian. This is an important choice as Poisson observations better describe neural data. We show, using data from rodent visual cortex (Fig 2D), the cross-validated mean squared error of the inferred spike rate to smoothed spike rate from held-out trials. The model with Poisson observations performs significantly better than the GPFA model with Gaussian observations. Others have noted similar advantages to Poisson observation factor models for neural data in other settings [10]. For this reason we wish to use a Poisson observation factor model for our SNP-GPFA model.

### 3 SNP-GPFA

To isolate noise and signal subspaces in the P-GPFA framework, we introduce a model that includes separate noise and signal latent structure (SNP-GPFA). We assess the model on two neural data sets. One contains multi-neuron spiking activity from 65 neurons recorded in primate V1 during passive viewing of a drifting sinusoidal grating stimulus, with 72 different orientations for  $D = 35$  repeated trials. The second consists of spiking activity from 67 neurons from two regions of rodent visual cortex, recorded during passive viewing of  $D = 20$  repeats of a 32-second sinusoidal grating stimulus. Gratings had 8 different orientations which persisted for 4 seconds each. For more information, see [24, 25] and the supplemental materials.

The SNP-GPFA model describes neural activity on trial  $j$  as

$$\mathbf{y}_j = \text{Pois}(f(\mathbf{W}_s^\top \mathbf{X}^s + \mathbf{W}_n^\top \mathbf{X}_j^n)) \quad (4)$$

where  $P$  signal latents are drawn from a “signal” Gaussian process,  $\mathbf{x}_p^s \sim \mathcal{GP}(0, \mathbf{K}_s)$  with covariance  $\mathbf{K}_s$  and concatenated to form  $\mathbf{X}^{s\top} = (\mathbf{x}_1^s, \mathbf{x}_2^s, \dots, \mathbf{x}_P^s)$ , which are shared across trials. On each trial,  $Q$  independent noise latents are drawn from a “noise” Gaussian process,  $\mathbf{x}_q^n \sim \mathcal{GP}(0, \mathbf{K}_n)$  with covariance  $\mathbf{K}_n$ , forming  $\mathbf{X}^{n\top} = (\mathbf{x}_1^n, \mathbf{x}_2^n, \dots, \mathbf{x}_Q^n)$ . Loading weights  $\mathbf{W}_s^\top$  and  $\mathbf{W}_n^\top$  parametrize a mapping from the dimensionality of the signal space  $P$  or the noise space  $Q$  to the full  $N$ -dimensional neural response space. Thus,  $\mathbf{W}_s$  is of size  $P \times N$  and  $\mathbf{W}_n$  is  $Q \times N$ . Covariance matrices  $\mathbf{K}_s$  and  $\mathbf{K}_n$  are constructed by evaluating a radial basis covariance at all pairs of time points in a trial. The SNP-GPFA model is outlined schematically in Fig 3A. For clarity, we visualize the model with only one signal and noise latent dimension, but the dimensions of the two spaces can be arbitrary.

To perform inference for the SNP-GPFA model, we develop a variational approach similar to that for P-GPFA. We use a variational distribution  $q_\phi$  for the latents, parametrized as a fully-independent multivariate normal distribution of dimension  $\tilde{T}(P+QD)$  where  $\tilde{T}$ , which corresponds to the number of Fourier coefficients needed to represent the signal. We determine  $\tilde{T}$  by assuming a minimum length scale of 10 ( $\ell \geq 10$ ), which substantially shrinks the number of Fourier coefficients required to represent the latent signal and noise processes (from 321 to 44 dimensions for rodent data, and from 511 to 108 dimensions for primate data). This choice is appropriate, as we typically we do not learn length scales smaller than this value.

We fit the SNP-GPFA model to data, and found that the the signal component successfully captures the PSTH, or mean response across trials (Fig. 3B). Importantly, the model also identifies a noise component that accurately predicts trial-by-trial spiking variability. Figure 3C shows 3 example trials for the top neuron in 3B. The per-trial rate deviations, given by a  $\mathbf{w}_{n,i}^\top \mathbf{X}_n$ , where  $\mathbf{w}_{n,i}$  is an isolated row of  $\mathbf{W}_n$ , accurately capture per-trial spiking deviations. This can be easily seen in trial three, where a sharp burst at the end of the trial is captured by the noise component of the model.

### 3.1 Learning dimensionality

To identify the dimensionality of signal and noise latents, we used a cross-validation known as co-smoothing training [26]. We first trained on a subset of randomly selected trials (10 for rodent data, 20 for primate data) and used Fourier-BBVI to learn  $\mathbf{W}_s^\top$ ,  $\mathbf{W}_n^\top$  and  $\mathbf{X}^s$ . To test the accuracy of the learned parameters ( $\mathbf{W}_s^\top$ ,  $\mathbf{W}_n^\top$ ,  $\ell$ ) and signal latents, we withhold a small random selection of neurons and then learn the noise latents  $\mathbf{X}^n$  over the held-out trials. We evaluate the cross-validated log-likelihood of the held out neurons using these new noise latents and the inferred structure from the initial trials. We performed five-fold cross validation and averaged over folds. For additional information about the data preprocessing and cross-validation, see supplemental materials.

We find the signal dimensionality by first leaving out the noise component, and increasing the number of signal dimensions until there is a decrease in CV performance. We then incrementally increase the number of noise dimensions until CV performance decreases. Interestingly, noise components increased CV performance suggesting population-level structure in trial-to-trial variability. For the primate data, this procedure identified 5 signal dimensions and 6 noise dimensions. (See section 3.3 for details of dimensionality for the multi-region rodent data.)

### 3.2 Visualizing signal and noise subspaces

We visualize the resulting signal and noise subspaces for neural population data recorded from primates. For this experiment, a drifting gratings stimulus was present for the first half of the of a 2.5 second trial (see [24] and the supplemental materials for more details). For a particular stimulus orientation ( $0^\circ$ ), we show the first 3 PCs of the five dimensional signal subspace. Here, we only show three PC dimensions for clarity. Note that during the stimulus presentation, there is a strong sinusoidal component to the latent neural activity. However, after the stimulus presentation period, this structure is no longer present. This latent signal structure in our SNP-GPFA model is nearly identical to the results of P-GPFA with no noise component, and to the plot in [10], which also uses traditional P-GPFA run on the same data. However, unlike P-GPFA, SNP-GPFA additionally extracts

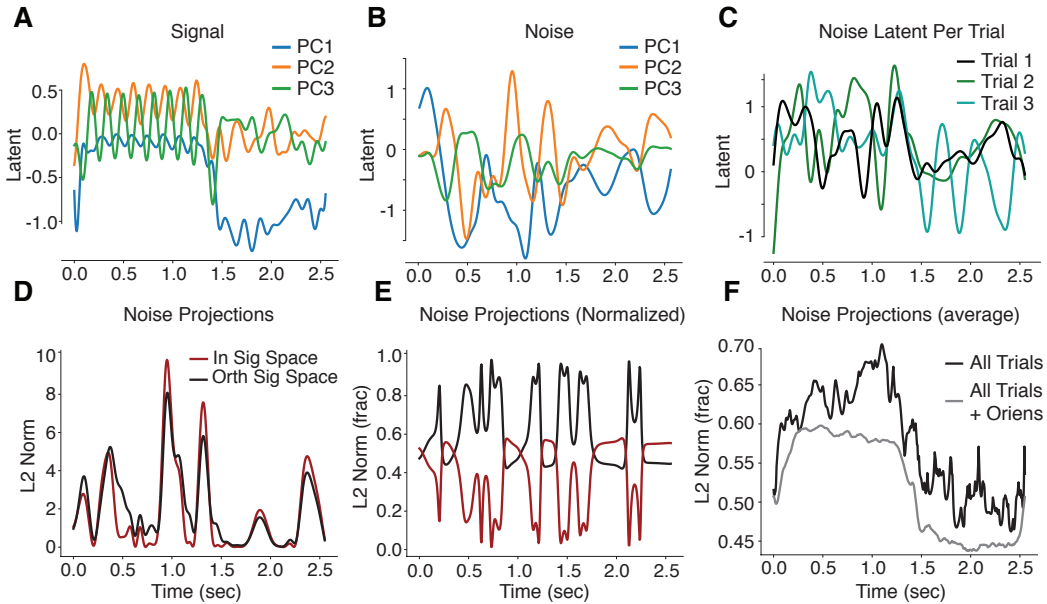


Figure 4: Visualizing noise and signal subspaces. First three PCs of signal (A) and noise (B) subspaces for 20 trials of a single stimulus presentation. (C) The first PC of the noise latents for three random example trials. (D) Total L2 norm of noise activity projected into signal subspace (red) and orthogonal to signal subspace (black) for an example trial. (E) Same as D but normalized to show percentage of noise activity variance. (F) Fraction of noise variance orthogonal to signal subspace averaged over all trials for a single orientation (black) and all trials and orientations (grey).

a noise subspace (Fig 4 B). This subspace has no obvious structure and does not include the same sinusoidal component in the first half of the trial. As expected, plotting the first noise latent PC across three example trials, there is no obvious pattern to the noise deviations, reflecting idiosyncratic variations in population firing rates across trials (Fig 4 C).

An important question that arises with this model is whether or not the noise subspace overlaps with the signal subspace. Overlap of these subspaces implies that trial-to-trial variability in the noise components can corrupt the population response along the signal dimensions, thereby interfering with representation of the signal. Previous work [15] suggests that noise only interferes with the signal representation if it lies in a direction defined by the derivatives of neural tuning curves, and recent work suggests that noise and signal subspaces may indeed be nearly orthogonal [14, 27, 28]. Because our model contains separate latent components for signal and noise, we can explicitly compare the relative angle between these subspaces. More specifically, we look to assess how strongly the pure-noise component of neural activity projects into the signal subspace. Under the SNP-GPFA model, the noise-subspace component of neural activity is  $\mathbf{Y}_n = \mathbf{W}_n^T \mathbf{X}_n$ . To assess overlap with the signal subspace, we compute the singular value decomposition of the signal-component loading weights,  $\mathbf{W}_s = \mathbf{U}\mathbf{S}\mathbf{V}^T$ , which provides a basis for the signal subspace via the columns of  $\mathbf{V}$ . The portion of variance of the noise within the signal subspace for each time point is then given by  $\|\mathbf{V}\mathbf{V}^T \mathbf{Y}_{n,t}\|_2^2$ , where  $\mathbf{Y}_{n,t}$  is the  $t^{\text{th}}$  column of  $\mathbf{Y}_n$ , or the L2 norm along the six noise dimensions. The portion of variance orthogonal to the signal subspace is thus simply  $\|\mathbf{Y}_{n,t} - \mathbf{V}\mathbf{V}^T \mathbf{Y}_{n,t}\|_2^2$ .

Figure 4D shows the resulting L2-norm time-series both within and orthogonal to the signal subspace. This is for a single example trial for the same single stimulus as in A-C (orientation of  $0^\circ$ ). To visualize the fractional variance into and out of the signal subspace, we normalize each trace by the total variance at each time point,  $\|\mathbf{Y}_{n,t}\|_2^2$  (Fig 4 E). For this trial, the noise activity tends to mostly lie in the subspace orthogonal to the stimulus activity. However, this is not true when the overall noise variance is high. At these moments, the noise exists mostly in the signal subspace. Figure 4F shows the fractional noise variance orthogonal to the signal subspace averaged over all trials for the orientation of  $0^\circ$  (black line) and all trials and orientations (grey line). It is primarily during the stimulus presentation time that the noise activity is preferentially orthogonal to the stimulus subspace.

When there is no stimulus, after the halfway point of the trial, there is a slight preference for the noise activity to lie in the signal subspace.

### 3.3 Shared and independent noise in multi-region data

The rodent dataset we examined contained data from two simultaneously-recorded visual cortical regions, an upstream area “V1” and a downstream area “AL”. The SNP-GPFA model can therefore be extended to allow for a characterization of shared variability across these regions. For simplicity, let’s consider two versions of the model: (1) a “shared-noise” model, which is the SNP-GPFA model applied to both regions simultaneously (see eq 4); or (2) an independent noise model, which includes a block-diagonalization of  $\mathbf{W}_n$  into a V1 component  $\mathbf{W}_n^{V1}$  and AL component  $\mathbf{W}_n^{AL}$ . The independent noise model describing neural activity for trial  $j$  is thus:

$$\begin{bmatrix} \mathbf{y}_j^{V1} \\ \mathbf{y}_j^{AL} \end{bmatrix} = \text{Pois} \left( f(\mathbf{W}_s^\top \mathbf{X}^s + \begin{bmatrix} \mathbf{W}_n^{V1} & 0 \\ 0 & \mathbf{W}_n^{AL} \end{bmatrix} \begin{bmatrix} \mathbf{X}_j^{n,V1} \\ \mathbf{X}_j^{n,AL} \end{bmatrix}) \right) \quad (5)$$

where  $\mathbf{X}_j^{n,V1}$  are noise latents that map exclusively to V1 activity, and  $\mathbf{X}_j^{n,AL}$  are noise latents that map exclusively to AL activity. By contrast, departures from a block-diagonal structure in the loadings matrix reflects the degree to which latent variability is shared across brain regions.

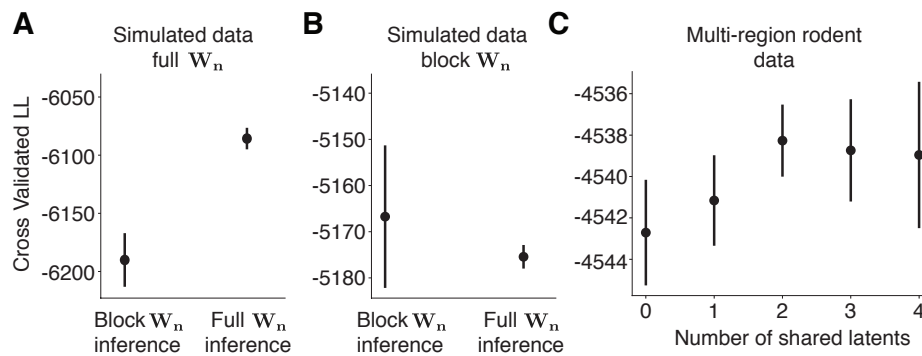


Figure 5: Two-region SNP-GPFA model. (A and B) Validation of our cross-validation (CV) approach on simulated data. (C) CV results on real neural data with varied block-diagonal structure for  $\mathbf{W}_n$ .

We first validate the approach on simulated data. We generate two datasets, one with a full  $\mathbf{W}_n$  matrix, and this other with block-diagonal structure. We perform inference on each of these datasets, one using the original model outlined in eq 4, and the other using the block diagonal structure in eq 5. We show that the model with its corresponding generated data exhibits higher CV performance (Fig 5 A,B). For additional details see supplement.

On the multi-region rodent data, we compare the SNP-GPFA model with a block-diagonal  $\mathbf{W}_n$  to models where  $\mathbf{W}_n$  has an increasing number of shared noise latents. We determine a six dimensional signal component. For the noise dimensionality, we start with a complete block representation (eq 5) of 5 V1 dimensions and 4 AL dimensions, which contain no shared components. We then compare CV performance of this model to ones where increasing numbers of noise latents are shared between the regions (Fig 5C). For information regarding how we select the proper number of noise and signal dimensions in this framework, see supplemental materials. We find that models with at least two shared noise dimensions perform better than models with one or fewer shared noise dimensions. Means and standard error shown over five-fold CV (Fig 5). This suggests that there is trial-varying structure in neural population activity that is shared across cortical regions.

## 4 Conclusion

We have introduced a Gaussian process factor analytic model for spike train data that extracts separate signal and noise latent structure from trial-structured data. To learn this model we employ a novel inference method based on black-box variational inference in the Fourier domain, which allows for faster and more stable inference by diagonalizing the posterior covariance and pruning unnecessary



frequencies. The resulting SNP-GPFA model is able to extract signal latents that characterize population PSTHs in real neural data, and noise latents that capture trial-to-trial variability that is shared across neurons. We then use the results of the model fit to answer scientific questions about trial-based neural data. We find that noise activity tends to project primarily in a subspace orthogonal to signal activity, especially when a stimulus is present, suggesting an optimal type of neural encoding [15]. We additionally use our model on multi-region rodent data and compare performance where noise is shared between cortical regions as opposed to being independent to each region. We find that noise models with shared structure better describes neural data suggesting cortical variability is shared across cortical regions. Overall, the model is a promising method for understanding the relationship between stimulus-locked and trial-varying neural activity at the population level. We believe there is a great number of additional scientific questions that the SNP-GPFA model can help answer, including determining how signal and noise representations relate to behavior, and further exploring how block-diagonal loadings matrices may partition signal and noise latent representations in multi-region data. We provide downloadable code for the community to use the SNP-GPFA model on their own trial-based neural data.

## Broader Impact

Here, we propose a new model for neuroscientists to uncover latent structure in trial-based neural population data. Trial-based neural recordings with identical stimuli are ubiquitous in neuroscience research. However, trial-by-trial variability in neural activity is not well understood. More broadly, it is unclear in general what the function of neural noise is in the brain. Our model works on neural population data to separate out neural noise latent representations from stimulus-locked representations. It additionally uses a novel inference technique that is rapid and stable. Here, we provide a general, easy-to-use tool for neuroscientists, and we hope others are encouraged to employ it to understand trial-based neural information in their own experimental set-up. We provide code for download. We do not foresee any negative consequences to society resulting from this work.

## Acknowledgements

SLK was supported by NIH grant F32MH115445-03, MCA and JWP were supported by grants from the Simons Collaboration on the Global Brain (SCGB AWD543027), and a U19 NIH-NINDS BRAIN Initiative Award (5U19NS104648). JWP was also supported by the NIH BRAIN initiative (NS104899 and R01EB026946). YY and SLS were supported by the NIH (R01EY024294 and R01NS091335), the NSF (1450824 and 1707287) the Simons Foundation (SCGB 325407) and the McKnight Foundation.

## References

- [1] BM Yu, JP Cunningham, G Santhanam, SI Ryu, KV Shenoy, and M Sahani. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. In *Adv neur inf proc sys*, pages 1881–1888, 2009.
- [2] John P Cunningham and B M Yu. Dimensionality reduction for large-scale neural recordings. *Nature neuroscience*, 17(11):1500–1509, 2014.
- [3] Anqi Wu, Nicholas G Roy, Stephen Keeley, and Jonathan W Pillow. Gaussian process based nonlinear latent structure discovery in multivariate spike train data. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 3499–3508. Curran Associates, Inc., 2017.
- [4] KC Lakshmanan, PT Sadtler, EC Tyler-Kabara, AP Batista, and BM Yu. Extracting low-dimensional latent structure from time series in the presence of delays. *Neural computation*, 2015.
- [5] Evan Archer, Il Memming Park, Lars Buesing, John Cunningham, and Liam Paninski. Black box variational inference for state space models. *arXiv preprint arXiv:1511.07367*, 2015.
- [6] Yuan Zhao, Jacob L Yates, Aaron J Levi, Alexander C Huk, and Il Memming Park. Stimulus-choice (mis) alignment in primate area mt. *PLOS Computational Biology*, 16(5):e1007614, 2020.

- [7] 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.
- [8] Jacob L Yates, Il Memming Park, Leor N Katz, Jonathan W Pillow, and Alexander C Huk. Functional dissection of signal and noise in mt and lip during decision-making. *Nature neuroscience*, 20(9):1285, 2017.
- [9] AC Damianou, MK Titsias, and ND Lawrence. Variational inference for uncertainty on the inputs of gaussian process models. *arXiv preprint arXiv:1409.2287*, 2014.
- [10] Yuan Zhao and Il Memming Park. Variational latent gaussian process for recovering single-trial dynamics from population spike trains. *Neural computation*, 29(5):1293–1316, 2017.
- [11] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- [12] Rajesh Ranganath, Sean Gerrish, and David Blei. Black box variational inference. In *Artificial Intelligence and Statistics*, pages 814–822, 2014.
- [13] João D Semedo, Amin Zandvakili, Christian K Machens, M Yu Byron, and Adam Kohn. Cortical areas interact through a communication subspace. *Neuron*, 2019.
- [14] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893, 2019.
- [15] Rubén Moreno-Bote, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. Information-limiting correlations. *Nature neuroscience*, 17(10):1410, 2014.
- [16] Lea Duncker and Maneesh Sahani. Temporal alignment and latent gaussian process factor inference in population spike trains. In *Advances in Neural Information Processing Systems*, pages 10445–10455, 2018.
- [17] Diederik P Kingma, Tim Salimans, and Max Welling. Variational dropout and the local reparameterization trick. In *Advances in Neural Information Processing Systems*, pages 2575–2583, 2015.
- [18] Richard E Turner and Maneesh Sahani. Two problems with variational expectation maximisation for time-series models. *Bayesian Time series models*, 1(3.1):3–1, 2011.
- [19] Magda Peligrad and Wei Biao Wu. Central limit theorem for fourier transforms of stationary processes. *The Annals of Probability*, pages 2009–2022, 2010.
- [20] Mikio Aoi and Jonathan W Pillow. Scalable bayesian inference for high-dimensional neural receptive fields. *bioRxiv*, page 212217, 2017.
- [21] MK Titsias and ND Lawrence. Bayesian gaussian process latent variable model. In *AISTATS*, volume 9, pages 844–851, 2010.
- [22] James Hensman, Nicolo Fusi, and Neil D Lawrence. Gaussian processes for big data. In *Uncertainty in Artificial Intelligence*, page 282, 2013.
- [23] Edward Snelson and Zoubin Ghahramani. Local and global sparse gaussian process approximations. In *Artificial Intelligence and Statistics*, pages 524–531, 2007.
- [24] Arnulf BA Graf, Adam Kohn, Mehrdad Jazayeri, and J Anthony Movshon. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature neuroscience*, 14(2):239, 2011.
- [25] Yiyi Yu, Jeffery N Stirman, Christopher R Dorsett, and Spencer LaVere Smith. Mesoscale correlation structure with single cell resolution during visual coding. *bioRxiv*, page 469114, 2018.
- [26] Jakob H Macke, Lars Buesing, John P Cunningham, M Yu Byron, Krishna V Shenoy, and Maneesh Sahani. Empirical models of spiking in neural populations. In *Advances in neural information processing systems*, pages 1350–1358, 2011.
- [27] Oleg I Rumyantsev, Jérôme A Lecoq, Oscar Hernandez, Yanping Zhang, Joan Savall, Radosław Chrapkiewicz, Jane Li, Hongkui Zeng, Surya Ganguli, and Mark J Schnitzer. Fundamental bounds on the fidelity of sensory cortical coding. *Nature*, 580(7801):100–105, 2020.
- [28] Ramon Bartolo, Richard C Saunders, Andrew R Mitz, and Bruno B Averbeck. Information-limiting correlations in large neural populations. *Journal of Neuroscience*, 40(8):1668–1678, 2020.