



Probabilistic Encoding Models for Multivariate Neural Data

Marcus A. Triplett and Geoffrey J. Goodhill*

Queensland Brain Institute and School of Mathematics and Physics, The University of Queensland, Brisbane, QLD, Australia

A key problem in systems neuroscience is to characterize how populations of neurons encode information in their patterns of activity. An understanding of the encoding process is essential both for gaining insight into the origins of perception and for the development of brain-computer interfaces. However, this characterization is complicated by the highly variable nature of neural responses, and thus usually requires probabilistic methods for analysis. Drawing on techniques from statistical modeling and machine learning, we review recent methods for extracting important variables that quantitatively describe how sensory information is encoded in neural activity. In particular, we discuss methods for estimating receptive fields, modeling neural population dynamics, and inferring low dimensional latent structure from a population of neurons, in the context of both electrophysiology and calcium imaging data.

Keywords: neural coding, calcium imaging, population code, brain-computer interfaces, generalized linear model, Gaussian process, factor analysis

OPEN ACCESS

Edited by:

Greg Stuart,
Australian National University, Australia

Reviewed by:

Tara Julia Hamilton,
Macquarie University, Australia
Adam Morris,
Monash University, Australia

*Correspondence:

Geoffrey J. Goodhill
g.goodhill@uq.edu.au

Received: 27 July 2018

Accepted: 07 January 2019

Published: 28 January 2019

Citation:

Triplett MA and Goodhill GJ (2019)
Probabilistic Encoding Models for
Multivariate Neural Data.
Front. Neural Circuits 13:1.
doi: 10.3389/fncir.2019.00001

1. INTRODUCTION

An animal's perceptual capabilities critically depend on the ability of its brain to form appropriate representations of sensory stimuli. However, the neural activity induced by a specific stimulus is highly variable, suggesting that neural encoding is a fundamentally probabilistic process. Characterizing the neural code thus requires statistical methods for relating stimuli to distributions of evoked patterns of activity. Modern techniques for recording such neural activity include multi-electrode arrays, which provide access to the behavior of populations of neurons at millisecond resolution, and optical imaging with genetically encoded calcium (Chen et al., 2013) and voltage indicators (Abdelfattah et al., 2018), which allow thousands of neurons to be recorded simultaneously (Ahrens et al., 2013; Chen et al., 2018). However, while improvements in multi-neuron recording allow us to probe neural circuits in great detail, they are accompanied by a need for computational techniques that scale to entire neural populations.

A statistical model for neural coding describes how a stimulus is mathematically related to a pattern of neural activity. By fitting the model one can extract important variables that quantitatively describe the encoding procedure taking place. For instance, such models enable the estimation of receptive fields and/or interneuronal coupling strengths. In contrast to other methods for inferring these variables, an approach based on statistical models situates the task of estimating salient parameters in a coherent mathematical framework, often with proof of asymptotic optimality or computational efficiency. By making explicit assumptions about how the data was generated, statistically principled approaches are often capable of identifying patterns in neural data which are challenging to find with simpler methods.

Linear and generalized linear models are among the most straightforward classes of statistical models for spike trains and assume that a neuron's activity is a noisy linear combination of the stimulus features. These models are highly effective at explaining the structure of sensory receptive fields and are computationally tractable, but do not explicitly model the temporal structure of the recorded signal and have difficulty accounting for correlations between neurons in short time windows. An important aspect of these correlations is their tendency to be modular, with distinct groups of neurons showing co-fluctuating activity. Latent factor models attempt to uncover the low dimensional structure that gives rise to this correlated variability, and recent efforts have focused on extracting low dimensional structure that evolves smoothly through time using a latent linear dynamical system or Gaussian process (Cunningham and Byron, 2014).

A further challenge is presented by calcium imaging, which provides only indirect access to neural activity through recorded fluorescence levels that reflect the concentration of calcium within a neuron. Often this data can be more difficult to interpret than electrophysiological recordings as there are a number of biophysical stages between stimulus presentation and fluorescence imaging where noise can enter and information can be lost. Using a generative model for calcium imaging data, however, one can explicitly account for the process through which action potentials are transformed into fluorescence levels. Fitting the generative model amounts to deconvolving the fluorescence signal to estimate the underlying spike train timeseries, and conventional encoding models can then be applied to deconvolved data. However, the ability to obtain spike counts from fluorescence data is highly constrained by experimental conditions, which motivates the development of encoding models specific to calcium imaging that do not necessarily involve spike train deconvolution.

While previous reviews have focused on estimating stimulus-response functions (Paninski et al., 2007; Pillow, 2007; Meyer et al., 2017), neural decoding (Paninski et al., 2007; Quiroga and Panzeri, 2009), and conceptual overviews of models and data analysis techniques (Cunningham and Byron, 2014; Paninski and Cunningham, 2018), this review instead discusses a range of recent exemplary models and their successful application to experimental data. Our goal is to provide sufficient mathematical detail to appreciate the respective strengths and weaknesses of each model, while leaving formal treatment of their associated fitting algorithms to their original sources.

2. LINEAR AND GENERALIZED LINEAR MODELS

We first briefly review now-standard material on models for single-neuron spike trains, primarily to develop the theory, terminology, and notation necessary for more recent work focused on multivariate models.

2.1. The Linear-Gaussian Model

Among the simplest probabilistic models for a neuron's response r to a stimulus vector \mathbf{s} is the linear-Gaussian model (Figure 1A), which assumes that a neuron linearly filters the features of \mathbf{s} as

$$r = \mathbf{w}^T \mathbf{s} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2) \quad (1)$$

where the vector \mathbf{w} is the stimulus filter, ϵ is an additive noise variable, and $\mathcal{N}(0, \sigma^2)$ is a Gaussian distribution with mean 0 and variance σ^2 (see Table 1 for a table of notation). In the case of visual processing the stimulus \mathbf{s} is a vector of pixel intensities for each point in the visual field, the stimulus filter \mathbf{w} corresponds to the classical visual receptive field, and the response r is either the spike count or firing rate within some time window following the stimulus. Assuming stimuli $\mathbf{s}_1, \dots, \mathbf{s}_K$ are presented over K trials yielding responses r_1, \dots, r_K with independent and identically distributed noise as in Equation 1, the maximum likelihood estimate (MLE, see Table 2 for a table of abbreviations) for the filter \mathbf{w} is given by

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\operatorname{argmax}} \prod_{k=1}^K p(r_k | \mathbf{s}_k, \mathbf{w}). \quad (2)$$

Since the noise model is Gaussian, the solution to Equation (2) is simply the ordinary least squares solution (Bishop, 2006)

$$\hat{\mathbf{w}} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{r} \quad (3)$$

where $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_K)^T$ is the stimulus design matrix and $\mathbf{r} = (r_1, \dots, r_K)^T$ is the vector of neuron responses.

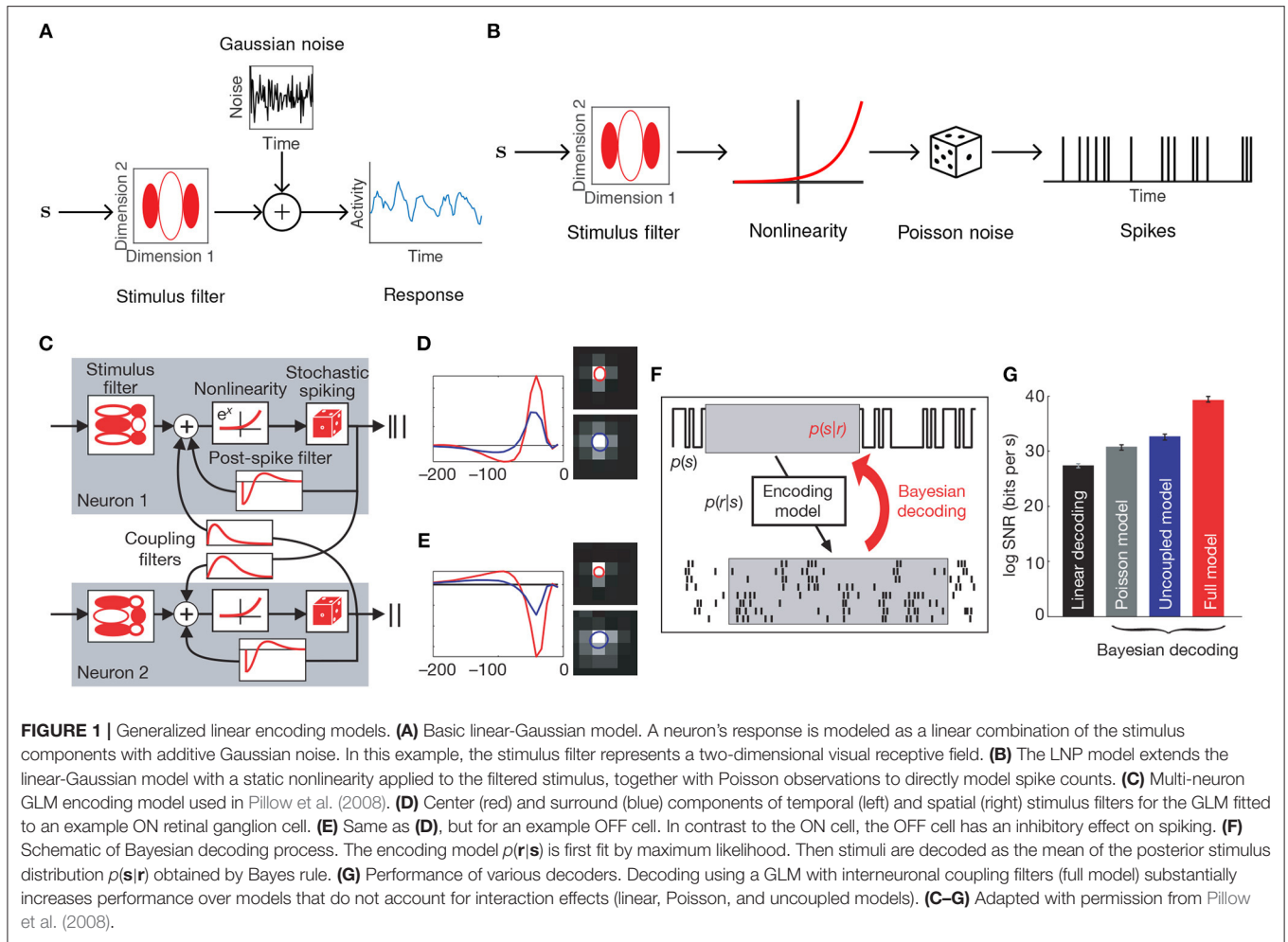
A common interpretation of the estimator in Equation (3) is in terms of the spike-triggered average (STA) of the stimulus, which is the filter obtained by averaging over the stimuli that elicited a response,

$$\hat{\mathbf{w}}_{\text{STA}} = \frac{1}{N} \mathbf{S}^T \mathbf{r} \quad (4)$$

where N is the total number of spikes. When the stimulus ensemble follows a multivariate Gaussian with independent dimensions (and is therefore not biased toward any particular region of the feature space) the STA is the optimal filter (Chichilnisky, 2001; Dayan and Abbott, 2001; Simoncelli et al., 2004) and is proportional to the MLE. In general, the MLE pre-multiplies the STA by the inverse of the autocorrelation matrix $\mathbf{S}^T \mathbf{S}$ of the stimulus ensemble to correct for bias in the presented stimuli, and thus corresponds to a whitened STA. Further discussion of the STA and its connection to the MLE can be found in Simoncelli et al. (2004) and Meyer et al. (2017).

2.2. The Linear-Nonlinear-Poisson Model

While a linear model can recover basic receptive field structure, it fails to capture the nonlinear changes in firing rate observed in electrophysiological recordings in cortex. In addition, the assumption of Gaussian noise leads to continuous (and possibly negative) estimates of spike counts. The linear-nonlinear-Poisson (LNP) model addresses these shortcomings by equipping the



generative model with a static nonlinearity following the linear filtering, and a Poisson noise model to directly model the number of spikes generated within a fixed time-window (**Figure 1B**) (Chichilnisky, 2001). Let $t = 1, \dots, T$ index over time bins. The LNP model assumes spikes follow an inhomogeneous Poisson process with time-varying firing rate $\lambda(t)$,

$$\lambda(t) = g(\mathbf{w}^\top \mathbf{s}(t)), \quad r(t) \sim \text{Pois}(\lambda(t)) \quad (5)$$

where g is a nonlinear activation function. While this nonlinearity can be estimated nonparametrically for each neuron (Simoncelli et al., 2004), it is often chosen to be $g(x) = \exp(x)$ as this ensures a non-negative intensity λ and tractable model fitting. Note that the specified firing rate $\lambda(t)$ will depend on the width Δ of the time bins or imaging rate, but for clarity here and for the remainder of the paper we omit explicit dependence of $\lambda(t)$ on Δ .

Assuming $g(x) = \exp(x)$ and that the responses $r(t)$ are count data, the MLE for the LNP model is the solution

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\operatorname{argmax}} \prod_{t=1}^T p(r(t)|\mathbf{s}(t), \mathbf{w}) = \underset{\mathbf{w}}{\operatorname{argmax}} \sum_{t=1}^T (r(t) \ln \lambda(t) - \lambda(t)) \quad (6)$$

where the second equality follows by substituting the Poisson mass function and taking logarithms. The LNP model can be fit by standard gradient-based optimization methods since the intensity function $\lambda(t)$ is differentiable with respect to the filter parameters \mathbf{w} and the log-likelihood function is concave (Paninski, 2004).

Regularization is a commonly used technique in machine learning for preventing a model from overfitting the training data. When maximizing the log-likelihood function for the LNP model with regularization, one penalizes the filter components whenever they deviate from zero

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\operatorname{argmax}} \sum_{t=1}^T (r(t) \ln \lambda(t) - \lambda(t)) - \eta \|\mathbf{w}\|_p \quad (7)$$

where $\|\cdot\|_p$ denotes the L_p norm and $\eta > 0$ is a penalty coefficient. Setting $p = 1$ or $p = 2$ corresponds, respectively, to LASSO and ridge regression (Friedman et al., 2001), encouraging a sparse filter \mathbf{w} . Maximizing the penalized log-likelihood is equivalent to performing posterior inference in a Bayesian regression model where \mathbf{w} has a Laplacian (for $p = 1$) or Gaussian (for $p = 2$) prior (Wu et al., 2006). In many circumstances,

TABLE 1 | Table of notation.

Symbol	Parameter
r	Response
\mathbf{s}	Stimulus
\mathbf{w}	Stimulus filter
\mathbf{S}	Stimulus design matrix
λ	Intensity, mean
$\epsilon, \xi, \xi_{h_\ell}$	Gaussian noise variable
σ, σ_s, ν	Noise standard deviation
μ, μ_s	Baseline activity, mean
\mathbf{h}	Spike history vector
\mathbf{J}_{ij}	Spike history coupling filter
\mathbf{x}	Latent factor
Ψ_s	Diagonal variance matrix
Λ, Λ_s	Factor loading matrix
Σ_s	Covariance matrix
τ	Gaussian process timescale
n	Spike count
c	Calcium concentration
γ	Autoregressive coefficient
α	Fluorescence scale
β	Baseline fluorescence level
f	Fluorescence level
p	Spike probability
h_ℓ	Refractory term
k_d	Dissociation constant
\mathcal{N}	Gaussian distribution
Pois	Poisson distribution
Bern	Bernoulli distribution

A subscripted s indicates stimulus-specific parameters.

such as when the data exhibits high noise levels, the ordinary (unpenalized) MLE cannot recover realistic receptive fields and needs to be constrained by regularization or priors (Sahani and Linden, 2003). Such Bayesian methods become highly effective in regimes of high noise, and a number of Bayesian extensions of receptive field inference invoke more subtle machine learning methods. For example, automatic relevance determination (Sahani and Linden, 2003) places a Gaussian prior on each element \mathbf{w}_i of the filter and iteratively updates the prior variance until the filter components corresponding to irrelevant stimulus features effectively vanish from the model. Automatic locality determination, on the other hand, involves constructing receptive field priors encoding the information that receptive fields tend to be localized in space, time relative to the stimulus, and spatiotemporal frequency (Park and Pillow, 2011).

2.3. Extensions of the LNP Model

The LNP model is a special case of a generalized linear model (GLM): a class of encoding models that generalize the simple linear-Gaussian model to models that follow linear filtering with a static nonlinearity and any noise model from the exponential family. While there is in general no probability mass function for a multivariate extension of the Poisson distribution, the GLM framework allows one to incorporate interaction effects between different neurons, thereby allowing statistical models for single

TABLE 2 | Table of abbreviations.

Abbreviation	Meaning
MLE	Maximum likelihood estimate
STA	Spike-triggered average
LNP	Linear-nonlinear-Poisson
GLM	Generalized linear model
FA	Factor analysis
EM	Expectation maximization
GP	Gaussian process
GPFA	Gaussian process factor analysis
PLDS	Poisson linear dynamical system

neurons to be used for entire populations. The LNP model is extended by the addition of spike-history filters \mathbf{J}_{ij} for all pairs of neurons i and j , intended to capture refractory effects for individual neurons (i.e., when $i = j$) and interaction effects between neurons ($i \neq j$), giving

$$\lambda_i(t) = \exp \left(\mathbf{w}_i^\top \mathbf{s}(t) + \sum_{j=1}^N \mathbf{J}_{ij}^\top \mathbf{h}_j(t) \right), \quad r_i(t) \sim \text{Pois}(\lambda_i(t)) \quad (8)$$

where \mathbf{w}_i is the stimulus filter for neuron i , $\mathbf{h}_j(t) = (r_j(t - 1), \dots, r_j(t - \tau))^\top$ is a vector of neuron j 's spike history, and τ determines the length of the spike history window. The addition of the coupling filters allows the GLM to model the correlation structure within a population of neurons, as opposed to a model consisting of independent LNP neurons. Note, however, that the GLM is only well defined for coupling filters that act on the recent spike history of other neurons within the population, and cannot model correlations that arise from coactivity with zero time-lag (Macke et al., 2011). This motivates the use of latent variable models (see below), where simultaneous correlations arise among neurons whose activity is concurrently modulated by a shared factor.

Nonetheless, the GLM has been successfully applied to many data sets (Pillow et al., 2005, 2008; Park et al., 2014). Notably, Pillow et al. (2008) applied the GLM to a population of retinal ganglion cells from the fly (**Figures 1C–E**), obtained a complete characterization of the network's spatiotemporal correlation structure, and showed how incorporating these correlations yields a $\sim 20\%$ increase in estimated information about the presented visual scene (**Figures 1F,G**).

3. LATENT FACTOR MODELS

3.1. Encoding With Factor Analysers

A frequent observation when recording population responses to the repeated presentation of identical stimuli is that variability tends to be correlated among groups of neurons. Such correlated variability (also known as shared variability or noise correlations) can substantially impact the efficacy of a neural code depending on the particular correlation structure (Abbott and Dayan, 1999; Schneidman et al., 2006; Lin et al., 2015), and suggests that there may be factors present that comodule the responses of groups

of neurons. Factor analysis (FA), a probabilistic generalization of principal components analysis, is a classical model for inferring the latent group structure that can give rise to correlated variability.

In a Gaussian coding scheme with independent neurons, a population response \mathbf{r} to a fixed stimulus s has a probability density given by

$$p(\mathbf{r}|s) = \mathcal{N}(\mathbf{r}|\boldsymbol{\mu}_s, \sigma_s^2 \mathbf{I}_N) \quad (9)$$

where the vector $\boldsymbol{\mu}_s$ is the mean population response, σ_s^2 is a noise variance common to each neuron, and \mathbf{I}_N is the $N \times N$ identity matrix. While this model is analytically tractable with closed-form expressions for $\boldsymbol{\mu}_s$ and σ_s , the diagonal covariance matrix means it fails to account for the correlation structure that may be present in the data. As shown in e.g., Pillow et al. (2008), this additional information can considerably influence decoding accuracy.

On the other hand, a Gaussian model with an unconstrained covariance matrix $\boldsymbol{\Sigma}_s$ yields a density of the form

$$p(\mathbf{r}|s) = \mathcal{N}(\mathbf{r}|\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \quad (10)$$

which, in principle, could outperform the Gaussian version that uses an unrealistic assumption of independently acting neurons (Santhanam et al., 2009). However, the covariance matrix $\boldsymbol{\Sigma}_s$ has $(N^2 + N)/2$ parameters to be learned per stimulus, requiring an amount of data that is impractically large to obtain experimentally for large N .

FA is a more moderate approach that attempts to capture shared variability in population activity by specifying a tractable parameterization of the covariance matrix. For FA the covariance matrix is defined as $\boldsymbol{\Sigma}_s = \boldsymbol{\Lambda}_s \boldsymbol{\Lambda}_s^\top + \boldsymbol{\Psi}_s$, where $\boldsymbol{\Psi}_s \in \mathbb{R}^{N \times N}$ is a diagonal matrix, $\boldsymbol{\Lambda}_s \in \mathbb{R}^{N \times q}$ is a factor loading matrix (analogous to the component loading matrix in principal components analysis), and $q < N$ determines the rank of $\boldsymbol{\Lambda}_s \boldsymbol{\Lambda}_s^\top$. Hence the population response \mathbf{r} is distributed as

$$p(\mathbf{r}|s) = \mathcal{N}(\mathbf{r}|\boldsymbol{\mu}_s, \boldsymbol{\Lambda}_s \boldsymbol{\Lambda}_s^\top + \boldsymbol{\Psi}_s). \quad (11)$$

This decomposes $\boldsymbol{\Sigma}_s$ into two matrices that capture separate aspects of the response variability: $\boldsymbol{\Lambda}_s \boldsymbol{\Lambda}_s^\top$ is a low-rank matrix that captures the variability that is shared across neurons, whereas the diagonal matrix $\boldsymbol{\Psi}_s$ captures variability private to each neuron (Churchland et al., 2010). A critical observation is that the FA covariance matrix only requires $(q + 1)N$ parameters, which is less than $(N^2 + N)/2$ whenever $q < (N - 1)/2$. Since q is usually chosen to be small, the FA covariance matrix requires much fewer parameters to be learned from the data.

An equivalent formulation of FA models the population response to a stimulus s as the projection from a low dimensional space of latent factors into the N -dimensional population space. This low dimensionality constraint forces any variability that the latent factors account for to be shared across groups of neurons, which leads to a modular correlation structure in the population recording. The generative model for the population response \mathbf{r}_s given a stimulus s is

$$\mathbf{r}_s = \boldsymbol{\Lambda}_s \mathbf{x}_s + \boldsymbol{\mu}_s + \boldsymbol{\epsilon}_s \quad (12)$$

$$\mathbf{x}_s \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_q) \quad (13)$$

$$\boldsymbol{\epsilon}_s \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_s), \quad (14)$$

where $\mathbf{x}_s \in \mathbb{R}^q$ denotes the vector of latent factors, which are assumed to be independent with a Gaussian prior. These factors are intended to reflect unobserved brain states and could be physiologically realized as, e.g., shared gain modulation by downstream circuits. Note that the formulation of FA in Equation (11) can be recovered from Equations (12–14) by marginalizing over the latent factors.

Maximum likelihood estimation of the FA parameters $\theta_s = (\boldsymbol{\mu}_s, \boldsymbol{\Psi}_s, \boldsymbol{\Lambda}_s, \sigma_s)$ is complicated by the presence of latent variables \mathbf{x} , as the MLE $\hat{\theta}_s$ depends on an estimated $\hat{\mathbf{x}}$, and vice versa. FA thus uses the Expectation Maximization (EM) algorithm, an iterative procedure for fitting latent variable models (Dempster et al., 1977; Ghahramani et al., 1996). One must also choose the dimensionality q of the latent space, typically with a standard model selection procedure such as a comparison of the cross-validated log-likelihood or with an information criterion (Schwarz, 1978).

The FA method was applied to rhesus monkeys with brain-computer interfaces implanted in area PMd (Santhanam et al., 2009). Monkeys were trained on reaching tasks and the authors attempted to infer the intended target from electrophysiological data using a decoder based on the FA encoding model. By fitting the factor analyser, the decoder inferred the latent factors that modulated neurons' responses. Incorporating this information led to substantial improvements in decoding accuracy over decoders based on independent Gaussian and Poisson encoding models.

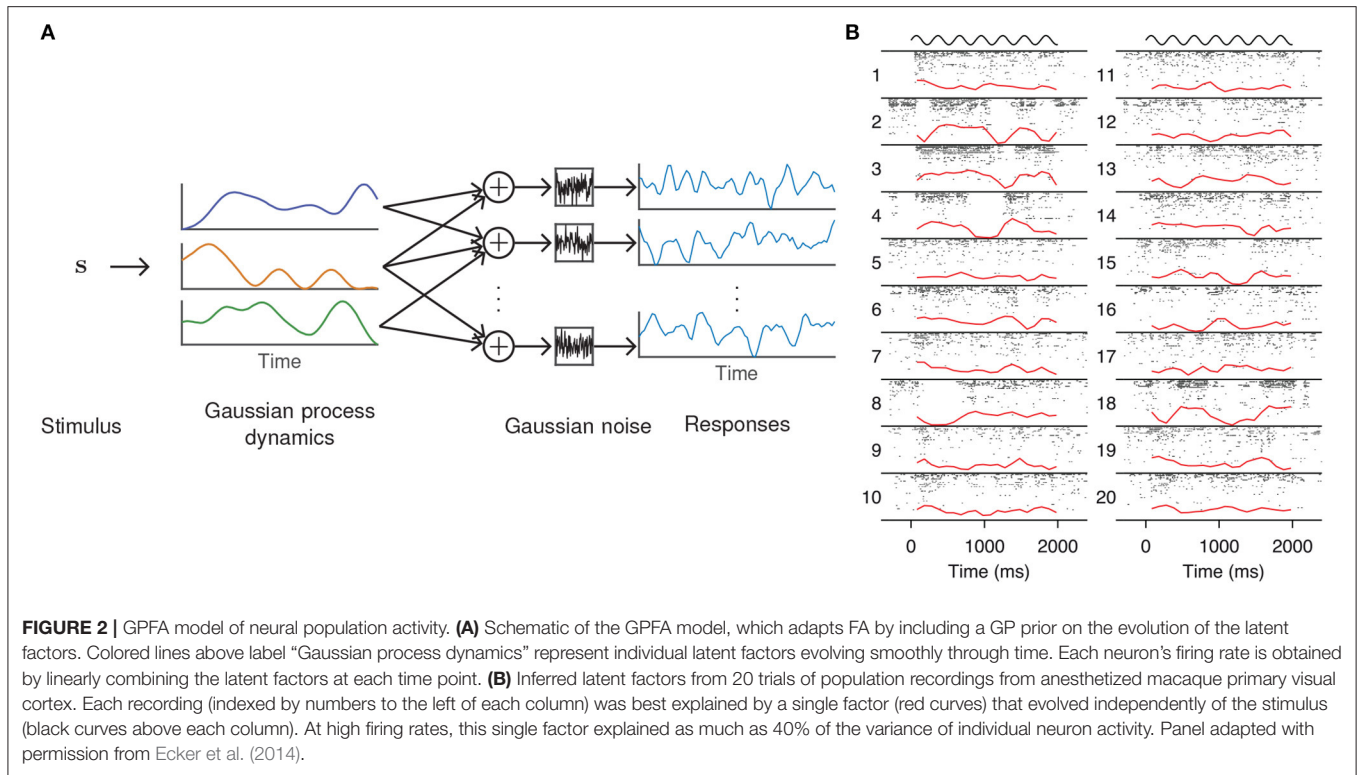
3.2. Gaussian Process Factor Analysis

The peristimulus time histogram averages spike trains over many trials to robustly estimate the aggregate effect of presenting a stimulus. Similarly, the FA encoding model is fit by pooling responses across trials to estimate the parameters θ_s . While this across-trial synthesis is necessary for fitting model parameters accurately, it will fail to reveal possibly important subtleties in neural activity within individual trials (Churchland et al., 2007, 2010; Afshar et al., 2011).

One way to adapt FA to single-trial analysis is to model the temporal evolution of the latent factors. A common technique in machine learning for enforcing temporal structure (or smoothness more generally) is Gaussian process (GP) regression, a Bayesian technique for nonparametric statistical modeling that places a GP prior on the latent variables (Williams and Rasmussen, 2006). The Gaussian process factor analysis (GPFA, Figure 2A) model (Yu et al., 2009) defines a GP for each dimension of the latent state $\ell = 1, \dots, q$, which, in the case of discretely indexed time, reduces to a collection of multivariate Gaussians

$$\mathbf{x}^{(\ell)} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}). \quad (15)$$

Here each $\mathbf{x}^{(\ell)} = (\mathbf{x}^{(\ell)}(1), \dots, \mathbf{x}^{(\ell)}(T))^\top$. Elements of the covariance matrix \mathbf{K} are typically determined by the squared



exponential kernel for encouraging smoothness

$$\mathbf{K}_{t_1, t_2} = \sigma_f^2 \exp\left(-\frac{(t_1 - t_2)^2}{2\tau^2}\right) + \sigma_n^2 \delta(t_1, t_2) \quad (16)$$

where δ is the Kronecker delta function and σ_f and σ_n are parameters controlling the variance of the GP. The observed responses are then modeled as in FA,

$$\mathbf{r}(t) = \mathbf{\Lambda} \mathbf{x}(t) + \boldsymbol{\mu} + \boldsymbol{\epsilon}(t) \quad (17)$$

$$\boldsymbol{\epsilon}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}) \quad (18)$$

where $\mathbf{x}(t)$ is the latent state at time t , $\mathbf{\Lambda}$ is the factor loading matrix, and $\boldsymbol{\mu}$ is a baseline activity level. GPFA can be viewed as a sequence of factor analysers (one for each time point) whose dimensions are linked together by smooth GPs. Note that while we have specified a single GP timescale τ , one can also assign distinct timescales τ_i to each dimension at the cost of an increase in computational overhead.

An advantage of GPFA is that the posterior over latent states $\mathbf{x}^{(t)}$ can be written down analytically because both the prior and likelihood are Gaussian, which form a conjugate pair (Bishop, 2006). This naturally leads to model fitting with the EM algorithm, where the updates for the parameter estimates are analogous to EM for FA (Ghahramani et al., 1996; Yu et al., 2009). Other examples of GP-based latent factor models are given in Nam (2015), Zhao and Park (2017), and Wu et al. (2017).

In a study of opioid anesthesia in macaque primary visual cortex, Ecker et al. (2014) used GPFA to investigate stimulus-driven patterns of population activity. The fitted model possessed

a single latent dimension that unmasked spontaneous transitions between periods of inactivity and highly elevated activity (Figure 2B). This single factor explained the observed increase in noise correlations and accounted for 40% of the variance of individual neuron firing rates. The extracted latent factors spanned a range of timescales, with some data best described by a latent factor whose strength changed slowly, on the order of several minutes. Similar up and down states had previously been seen only with non-opioid anesthetics.

3.3. The Poisson Linear Dynamical System

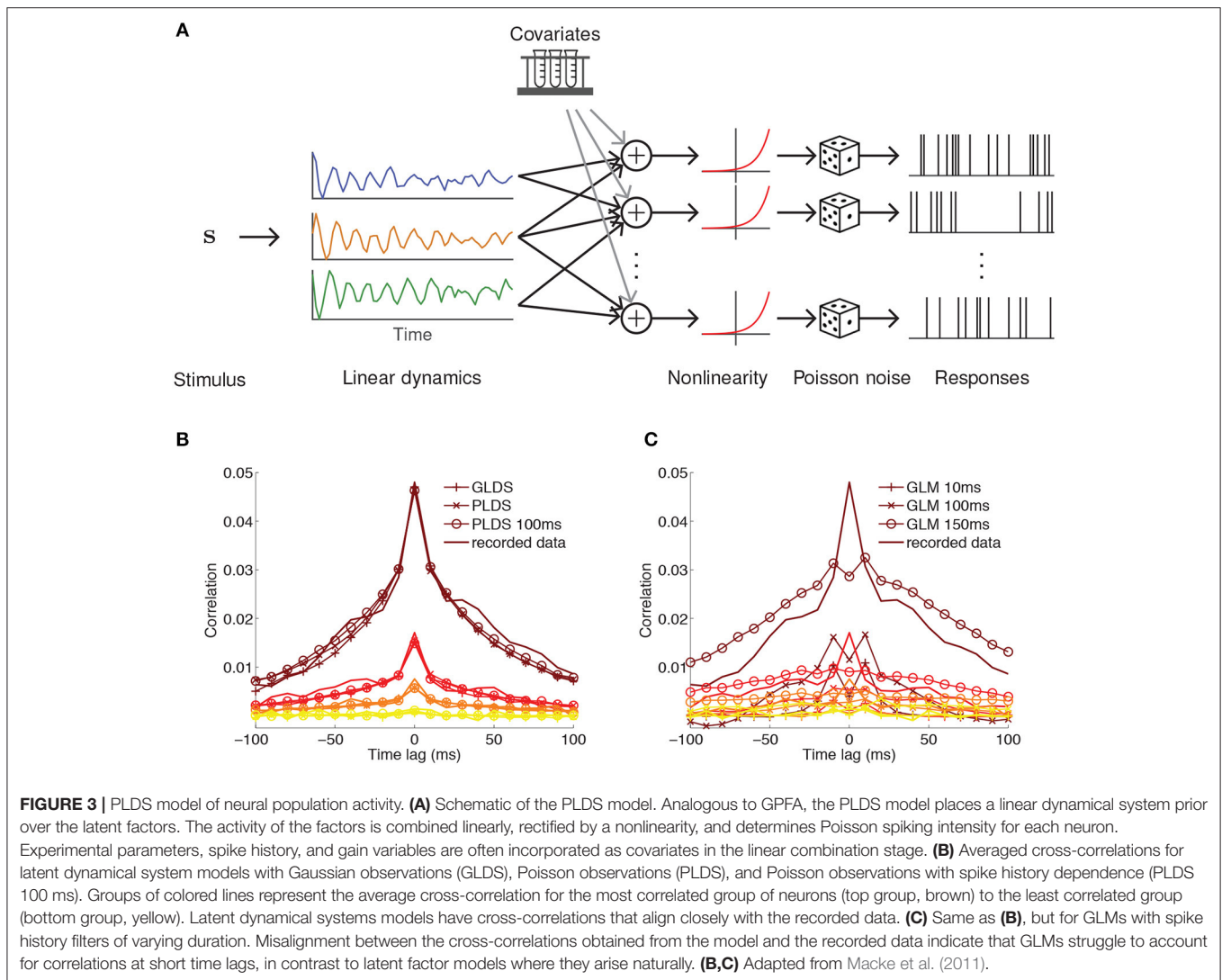
An alternative approach for latent trajectory modeling is to estimate the underlying linear dynamics of the latent state (Macke et al., 2011; Churchland et al., 2012; Pandarinath et al., 2018a). While the classical Kalman filter is the most thoroughly developed method for estimating the transition matrix in a linear dynamical system, a more appropriate generative model for neurons is the Poisson linear dynamical system (PLDS, Figure 3A) (Macke et al., 2011), which substitutes Poisson observations for the Gaussian emissions in the Kalman filter to directly model observed spike counts. The latent state $\mathbf{x}_k(t) \in \mathbb{R}^d$ on trial k at time bin t follows linear Markovian dynamics

$$\mathbf{x}_k(t+1) = \mathbf{A} \mathbf{x}_k(t) + \mathbf{b}(t) + \boldsymbol{\epsilon}_k(t+1) \quad (19)$$

$$\mathbf{x}_k(1) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_1) \quad (20)$$

$$\boldsymbol{\epsilon}_k(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}) \quad (21)$$

where \mathbf{A} is the dynamics matrix, \mathbf{Q} is the noise covariance for the latent linear dynamics, and \mathbf{Q}_1 is the covariance of the initial



state. The latent dynamics are driven by a variable $\mathbf{b}(t)$ that captures stimulus-specific effects. Note that the PLDS model is formulated with explicit dependence on the trial index k , so that $\mathbf{b}(t)$ accounts for stimulus effects that are trial-independent. Similar to the LNP model, the observed spike responses on trial k then follow a Poisson distribution with mean $\lambda_{i,k}(t)$ derived from the latent state. For neuron i this takes the form

$$\lambda_{i,k}(t) = g(\mathbf{\Lambda}_{(i)} \mathbf{x}_k(t) + \boldsymbol{\mu}_i), \quad r_{i,k}(t) \sim \text{Pois}(\lambda_{i,k}(t)). \quad (22)$$

Here the latent state influences an individual neuron i according to a row $\mathbf{\Lambda}_{(i)}$ of the factor loading matrix $\mathbf{\Lambda}$, and the low dimensionality of the latent state leads to the correlated variability as in the discussion of FA. Common choices for the nonlinearity include $g(x) = \exp(x)$ (Macke et al., 2011) and $g(x) = \ln(1 + \exp(x))$ (Buesing et al., 2017).

This model can be modified in various ways to suit the data. For example, the stimulus drive term \mathbf{b} in Equation (19) can be moved within the nonlinearity in Equation (22), so that the latent dynamics are decoupled from the stimulus and only

reflect changes internal to the brain. The intensity can be further extended by adding terms for, e.g., multiplicative gain (Buesing et al., 2017) and spike history (Macke et al., 2011) to capture refractory effects. A major advantage of latent factor models is their ability to account for correlations within short time intervals (**Figure 3B**), which GLMs struggle to match (**Figure 3C**).

The PLDS model is fit using a modified EM algorithm, which requires computing the posterior over the latent variables. Due to the Poisson observation model an analytic form of this posterior is unavailable. Typically one replaces the exact posterior by its Laplace approximation, which accelerates model fitting but violates some assumptions of the EM algorithm, resulting in an approximate inference framework (Macke et al., 2011).

An application of PLDS to multi-electrode recordings from songbird auditory cortex by Buesing et al. (2017) revealed that responses are modulated by shared variability with a single latent state, a similar result to Ecker et al. (2014). Buesing et al. histologically traced the locations of the recording sites and found a spatial gradient in the strength of the latent states. Shared variability was stronger (i.e., neurons were more strongly

coupled to the latent state) in deeper regions of auditory cortex. Interestingly, this strength was much weaker for certain stimulus classes than others, suggesting that deeper neurons selectively decouple from the latent state according to their stimulus preference. Other examples of dynamical systems-based latent factor models are given in Paninski et al. (2010), Buesing et al. (2012), Pfau et al. (2013), Semedo et al. (2014), Buesing et al. (2014), Archer et al. (2014), Kao et al. (2015), Gao et al. (2016), and Pandarinath et al. (2018b).

4. GENERATIVE MODELS FOR CALCIUM IMAGING DATA

4.1. Autoregressive Calcium Dynamics and Spike Deconvolution

The potential utility of large scale simultaneous neural recordings is constrained by our ability to make use of sophisticated techniques (such as latent factor methods) to analyse the data. While calcium imaging provides access to such large scale data, the models discussed so far assume that the data being analyzed is electrophysiological; i.e., that the neurons' responses are spike counts (for Poisson noise models) or firing rates (e.g., for Gaussian noise models). Their application to calcium imaging thus requires knowledge of how the optically recorded fluorescence signals are related to the underlying spiking activity. One approach to solving this problem involves constructing a generative statistical model where the spike counts are latent variables that are subsequently inferred from the fluorescence levels.

The presentation of a stimulus elicits a sequence of spikes across a population of neurons. For an individual neuron, we have assumed that the number of spikes within a time bin is sampled from a Poisson distribution with mean λ according to its particular receptive field. Each action potential is associated with a stereotypical rise and decay of the intracellular calcium concentration $c(t)$, usually modeled by an autoregressive process of order p (suppressing initial conditions for clarity) (Vogelstein et al., 2009),

$$c(t) = \sum_{i=1}^p \gamma_i c(t-i) + n(t), \quad n(t) \sim \text{Pois}(\lambda) \quad (23)$$

where the Poisson-distributed random variable $n(t)$ models the generation of spikes within a time bin and $\gamma_1, \dots, \gamma_p$ are the autoregressive coefficients that govern the rise and decay of the fluorescence levels. The observed fluorescence signal $f(t)$ is then obtained by a linear transformation of the calcium levels with additive noise,

$$f(t) = \alpha c(t) + \beta + \epsilon(t), \quad \epsilon(t) \sim \mathcal{N}(0, \sigma^2) \quad (24)$$

where α sets the scale of the fluorescence signal and β accounts for a baseline fluorescence that may be unique to the imaging set-up or due to specific biophysical properties of individual neurons. The Gaussian noise model is intended to encompass variability due to, e.g., light scattering and shot noise (Delaney et al., 2018).

Note that this model does not set parameters for the scale or baseline of the calcium transient in Equation (23), as they are absorbed by α and β when the calcium is transformed to obtain the fluorescence (Vogelstein et al., 2009). An illustration of the generative model is given in **Figures 4A,B**.

For imaging systems where the rise time of the indicator is fast relative to the imaging rate a first-order autoregressive process is typically used, corresponding to an instantaneous rise and exponential decay of the calcium concentration. An autoregressive process of order 2 is used in situations where the rise time is slow relative to the imaging rate, in which case the calcium transient appears to approach its maximum amplitude gradually (Pnevmatikakis et al., 2016).

Models based on Equations (23, 24) have been used for spike train deconvolution (Vogelstein et al., 2009, 2010; Friedrich and Paninski, 2016; Pnevmatikakis et al., 2016). Let the vector $\theta = (\alpha, \beta, \lambda, \sigma, \{\gamma_i\}_{i=1}^p)$ denote the model parameters, and let $\mathbf{f} = (f(1), \dots, f(T))^T$ and $\mathbf{n} = (n(1), \dots, n(T))^T$. Following Bayes' rule, the maximum *a posteriori* estimate for the spike train is

$$\hat{\mathbf{n}} = \underset{n(t) \in \mathbb{N}_0 \forall t}{\operatorname{argmax}} p(\mathbf{n}|\mathbf{f}, \theta) = \underset{n(t) \in \mathbb{N}_0 \forall t}{\operatorname{argmax}} p(\mathbf{f}|\mathbf{n}, \theta)p(\mathbf{n}|\theta) \quad (25)$$

where \mathbb{N}_0 is the set of non-negative integers. Given the spike sequence \mathbf{n} , the fluorescence levels $f(t)$ are independent and depend only on the calcium concentration $c(t)$, hence the likelihood factorizes as

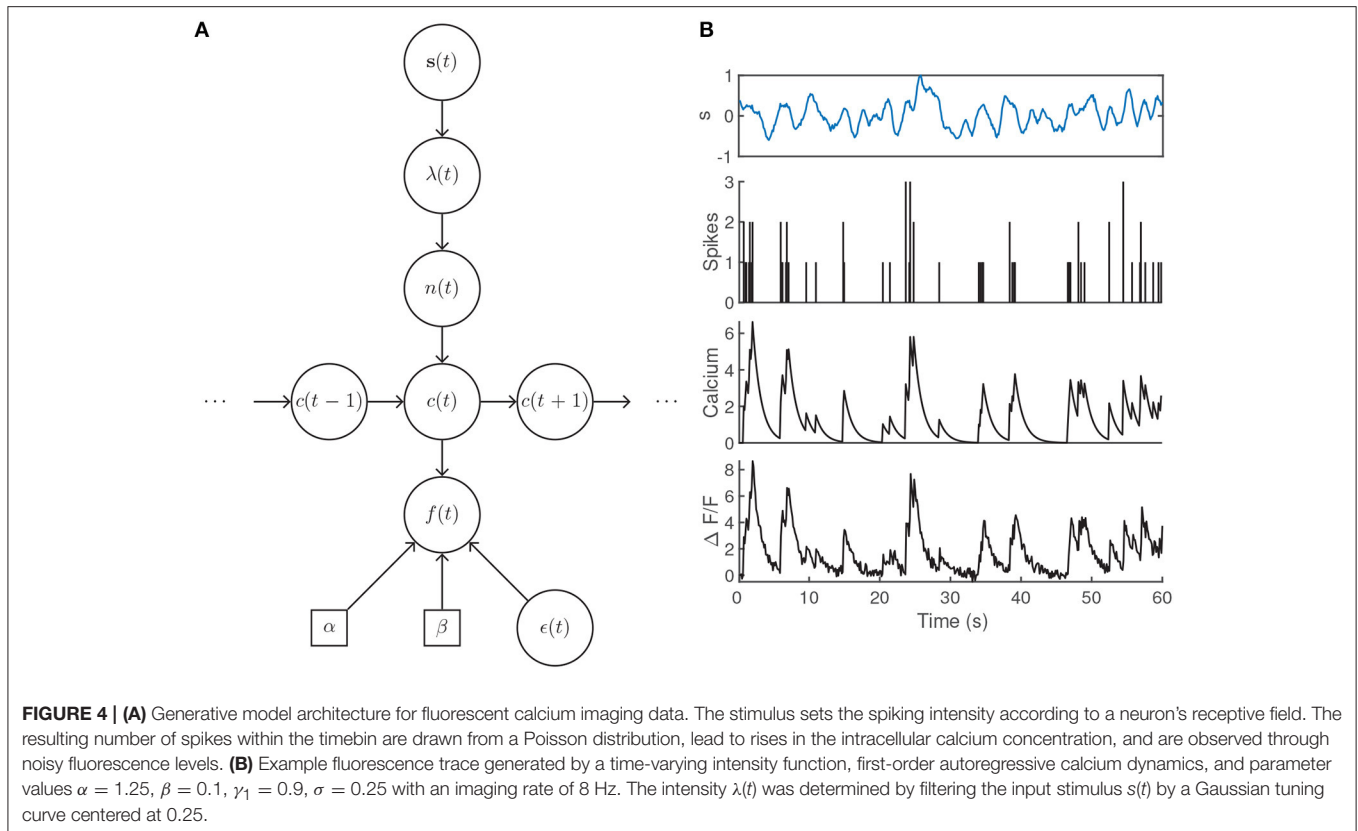
$$p(\mathbf{f}|\mathbf{n}, \theta) = \prod_{t=1}^T p(f(t)|c(t), \theta) = \prod_{t=1}^T \mathcal{N}(f(t)|\alpha c(t) + \beta, \sigma^2). \quad (26)$$

Substituting Equation (26) into (25) and taking logarithms, the optimal sequence of spikes is then

$$\hat{\mathbf{n}} = \underset{n(t) \in \mathbb{N}_0 \forall t}{\operatorname{argmax}} \sum_{t=1}^T \left\{ -\frac{1}{2\sigma^2} (f(t) - \alpha c(t) - \beta)^2 + n(t) \ln \lambda - \ln(n(t)!) \right\}. \quad (27)$$

This is a difficult optimization problem because it requires searching through an infinite discrete space of spike trains. As noted in Vogelstein et al. (2010), even imposing an upper bound on the number of spikes within a frame yields an optimization problem with exponential computational complexity. One approach for overcoming this intractability involves approximating the Poisson distribution in Equation (25) by an exponential distribution, which leads to a concave objective function but with continuous estimates of $\hat{\mathbf{n}}$ (Vogelstein et al., 2010). This approximation also allows for a time-varying intensity function $\lambda(t)$, but does not explicitly model the transformation from stimulus to spiking intensity.

Runyan et al. (2017) applied a combination of the methods described in this review to study the timescales of population codes in cortex. 2-photon calcium imaging of auditory and posterior parietal cortices was performed while mice completed a sound localization task. The resulting fluorescence data



was deconvolved according to the exponential-approximation approach described above to estimate firing rates (Vogelstein et al., 2010). They then fitted a GLM encoding model to populations from each cortical area that included coupling filters and various experimental and behavioral covariates. The fitted model was used in a decoding analysis that quantified the contribution of interneuronal coupling in the two cortical areas, and showed that stronger coupling was associated with population codes that had longer timescales. This provided evidence for a coding mechanism where tightly coupled populations of neurons prolonged the representation of stimuli through their sequential activation.

4.2. A Generalized Model for Calcium Dynamics

The calcium kinetics in Equation (23) are deterministic given the spike counts. In reality the concentration of calcium may be subject to many sources of variability, and analyses of some data sets may benefit from explicitly accounting for this noise. Vogelstein et al. (2009) modeled this by driving the calcium levels by both Bernoulli-distributed spikes and additive Gaussian noise,

$$c(t) = \gamma c(t-1) + n(t) + \xi(t) \quad (28)$$

$$n(t) \sim \text{Bern}(p(t)) \quad (29)$$

$$\xi(t) \sim \mathcal{N}(0, v^2) \quad (30)$$

where $\text{Bern}(p(t))$ is the Bernoulli distribution with time-dependent trial-success probability $p(t)$, $\gamma < 1$ is an autoregressive coefficient, and v^2 is the calcium noise variance. A simplifying assumption in models based on Equations (23) and (24) is that spikes are generated independently of their spike history. However, the spike probability can be more generally modeled with a GLM (Vogelstein et al., 2009)

$$p(t) = 1 - \exp\left(-g\left(\mathbf{w}^\top \mathbf{s}(t) + \mathbf{J}^\top \mathbf{h}(t)\right)\right) \quad (31)$$

where g is a selected nonlinearity. Unlike the standard GLM structure of Equation 8, the spike history term here takes the form $\mathbf{h}(t) = (h_1(t), \dots, h_L(t))^\top$, where each h_ℓ is an exponentially decaying refractory term that jumps following each spike

$$h_\ell(t) = \gamma_{h_\ell} h_\ell(t-1) + n(t) + \xi_{h_\ell}(t), \quad \xi_{h_\ell}(t) \sim \mathcal{N}(0, v_{h_\ell}^2). \quad (32)$$

Finally, rather than a simple linear relationship between $f(t)$ and $c(t)$, Vogelstein et al. (2009) and Vogelstein et al. (2010) also consider saturating fluorescence levels using a nonlinear Hill function with dissociation constant k_d

$$f(t) = \alpha \frac{c(t)}{c(t) + k_d} + \beta + \epsilon(t), \quad \epsilon(t) \sim \mathcal{N}(0, \sigma^2). \quad (33)$$

Importantly, saturation of the fluorescence signal causes the spike-triggered fluorescence transients to become progressively

smaller during a train of action potentials, and failure to account for this detail may limit the accuracy of spike deconvolution algorithms. The model defined by Equations (28–33) is fit using a sequential Monte Carlo method (Vogelstein et al., 2009). By including explicit stimulus and spike history filters, Vogelstein et al. (2009) could accurately infer spike times from fluorescence data with temporal superresolution; i.e., could identify when within an imaging frame each spike occurs. Some other example methods for spike deconvolution are based on compressed sensing (Pnevmatikakis and Paninski, 2013), fully Bayesian inference (Pnevmatikakis et al., 2013), and variational autoencoders (Speiser et al., 2017).

5. DISCUSSION

Probabilistic modeling provides a practical, interpretable, and theoretically grounded framework for probing how networks of neurons process information. Many of the statistical models discussed in this review are abstract mathematical descriptions of how stimuli are related to patterns of neural activity. Often the mathematical operations that define the models do not necessarily attempt to align with real biological functions or behavior. Rather, such models are intended to serve as tools to uncover interpretable patterns and relationships that may not be detectable by other approaches. On the other hand, there are cases where the goal is to infer biophysical variables, as in e.g., models for calcium imaging data or for the anatomical architecture of a neural circuit, and then greater care must be taken to constrain the model by relevant physiological data (Paninski et al., 2007; Real et al., 2017; Latimer et al., 2018).

Recent advances in statistical models of spike train data have focused on incorporating more general nonlinear transformations of the latent state, including the use of neural networks (Gao et al., 2016; Pandarinath et al., 2018b) and GPs (Wu et al., 2017). This is in contrast to e.g., the FA and GPFA encoding models, where the mean spiking intensity of a neuron is obtained by a simple linear transformation of the latent state. Bayesian methods, such as latent factor modeling, are a powerful way to incorporate prior knowledge when making inferences about the behavior of a system. While GPFA places a smoothness prior on the evolution of latent factors to encourage some degree of temporal structure, other methods place priors on, e.g., network structure for connectivity inference (Linderman S. et al., 2016) and the latent states of a hidden Markov model with Poisson observations (Linderman S. W. et al., 2016).

Although there has been a rapid expansion in the number of models for extracting receptive fields, interneuronal coupling strengths, and latent structure from multivariate electrophysiological recordings, similar models for calcium imaging data are only beginning to emerge (Aitchison et al., 2017; Khan et al., 2018). A common approach for analysing calcium imaging data involves first deconvolving fluorescence traces and then fitting conventional models, but deconvolution methods only provide coarse estimates of firing rates. Spike trains obtained by highly optimized algorithms typically only agree with ground truth recordings with a correlation coefficient less than ~ 0.75 , even with substantial training data, suggesting that there is an unavoidable loss of information associated with

spike deconvolution (Pnevmatikakis et al., 2016; Berens et al., 2018). An advantage of GPFA over earlier methods for estimating trajectories of population activity is that it condenses the two stages of dimensionality reduction and smoothing into a single stage of posterior inference. Similarly, probabilistic analysis of calcium imaging data can have the two stages of deconvolution and model fitting merged into a single step by marginalizing over possible spike trains (Ganmor et al., 2016), mitigating some of the information loss accompanied by deconvolution. Neural encoding models for calcium imaging data that avoid an explicit intermediate step of spike inference are likely to be an important future development in this area (Aitchison et al., 2017).

Many studies consider the amplitude of an evoked calcium transient as a measure of a neuron's response. This has been widely used in zebrafish larvae, for which there has been significant interest in recent years. For example, 2-photon calcium imaging of the zebrafish optic tectum has led to new insights into the circuit architecture determining selectivity to size, location, and direction of motion (Del Bene et al., 2010; Gabriel et al., 2012; Grama and Engert, 2012; Nikolaou et al., 2012; Lowe et al., 2013; Preuss et al., 2014; Avitan et al., 2016; Abbas et al., 2017), and light-sheet microscopy has allowed for the creation of brain-wide functional circuit models for motor behavior driven by vision (Naumann et al., 2016) and thermosensation (Haesemeyer et al., 2018). Similar studies in the future provide further opportunities for model-based analyses.

The techniques described in this review were developed for spike train or calcium imaging data, but some approaches are broadly applicable across systems neuroscience. For instance, suitably adapted latent factor models have been successfully applied to recordings of the local field potential, where it was found that the activity of particular latent factors could discriminate vulnerability to stress-induced behavioral dysfunction in mouse models of major depressive disorder (Gallagher et al., 2017; Hultman et al., 2018).

As the scale of multi-neuron data continues to grow, the creation of new models and their associated fitting algorithms may be spurred more by efficiency and scalability considerations than the level of statistical detail they are able to extract from experimental data (Zoltowski and Pillow, 2018). In some cases the computational issues associated with neural data analysis are more profound than simply needing a larger computer cluster. Neuropixel electrode arrays (Jun et al., 2017), for example, are capable of recording from hundreds of channels simultaneously, and may put inference algorithms under strain if computational efficiency is not sufficiently addressed. When combined with fluorescent sensors of neural activity, optogenetic photostimulation grants the ability to manipulate neural circuits in real time, and models are now beginning to explicitly integrate the effect of photostimulation on calcium transients (Aitchison et al., 2017). Moreover, genetically encoded voltage indicators operate on a timescale of tens of milliseconds (Knöpfel et al., 2015), overcoming one of the principal drawbacks of calcium imaging; namely, the slow binding kinetics of the indicator relative to the timescale of action potential generation. Combining these emerging technologies with models designed to capture their associated generative processes thus promises

to greatly improve our capacity to uncover how patterns of neural activity represent and process features of the external world.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Abbas, F., Triplett, M. A., Goodhill, G. J., and Meyer, M. P. (2017). A three-layer network model of direction selective circuits in the optic tectum. *Front. Neural Circuits* 11:88. doi: 10.3389/fncir.2017.00088
- Abbott, L. F., and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Comput.* 11, 91–101. doi: 10.1162/089976699300016827
- Abdelfattah, A. S., Kawashima, T., Singh, A., Novak, O., Liu, H., Shuai, Y., et al. (2018). Bright and photostable chemigenetic indicators for extended *in vivo* voltage imaging. *bioRxiv:436840*. doi: 10.1101/436840
- Afshar, A., Santhanam, G., Yu, B. M., Ryu, S. I., Sahani, M., and Shenoy, K. V. (2011). Single-trial neural correlates of arm movement preparation. *Neuron* 71, 555–564. doi: 10.1016/j.neuron.2011.05.047
- Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., and Keller, P. J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* 10, 413–420. doi: 10.1038/nmeth.2434
- Aitchison, L., Russell, L., Packer, A. M., Yan, J., Castonguay, P., Häusser, M., et al. (2017). “Model-based Bayesian inference of neural activity and connectivity from all-optical interrogation of a neural circuit,” in *Advances in Neural Information Processing Systems*, eds I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.), 3489–3498.
- Archer, E. W., Koster, U., Pillow, J. W., and Macke, J. H. (2014). “Low dimensional models of neural population activity in sensory cortical circuits,” in *Advances in Neural Information Processing Systems*, eds Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Curran Associates, Inc.), 343–351.
- Avitan, L., Pujic, Z., Hughes, N. J., Scott, E. K., and Goodhill, G. J. (2016). Limitations of neural map topography for decoding spatial information. *J. Neurosci.* 36, 5385–5396. doi: 10.1523/JNEUROSCI.0385-16.2016
- Berens, P., Freeman, J., Deneux, T., Chenkov, N., McColgan, T., Speiser, A., et al. (2018). Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. *PLoS Comput. Biol.* 14:e1006157. doi: 10.1371/journal.pcbi.1006157
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin; Heidelberg: Springer-Verlag.
- Buesing, L., Calabrese, A., Cunningham, J., Woolley, S., and Paninski, L. (2017). A statistical model of shared variability in the songbird auditory system. *bioRxiv:113670*. doi: 10.1101/113670
- Buesing, L., Machado, T. A., Cunningham, J. P., and Paninski, L. (2014). “Clustered factor analysis of multineuronal spike data,” in *Advances in Neural Information Processing Systems*, eds Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Curran Associates, Inc.), 3500–3508.
- Buesing, L., Macke, J. H., and Sahani, M. (2012). Learning stable, regularised latent models of neural population dynamics. *Network* 23, 24–47. doi: 10.1099/0954898X.2012.677095
- Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295–300. doi: 10.1038/nature12354
- Chen, X., Mu, Y., Hu, Y., Kuan, A. T., Nikitchenko, M., Randlett, O., et al. (2018). Brain-wide organization of neuronal activity and convergent sensorimotor transformations in larval zebrafish. *Neuron* 100, 876–890.e5. doi: 10.1016/j.neuron.2018.09.042

ACKNOWLEDGMENTS

MT is supported by an Australian Government Research Training Program Scholarship. GG is grateful for financial support from the Australian Research Council Discovery Projects DP170102263 and DP180100636. We thank Lilach Avitan, Jan Mölter, and the Reviewers for very helpful feedback on the manuscript.

- Chichilnisky, E. (2001). A simple white noise analysis of neuronal light responses. *Network* 12, 199–213. doi: 10.1080/713663221
- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., et al. (2012). Neural population dynamics during reaching. *Nature* 487, 51–56. doi: 10.1038/nature11129
- Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* 13, 369–378. doi: 10.1038/nn.2501
- Churchland, M. M., Yu, B. M., Sahani, M., and Shenoy, K. V. (2007). Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Curr. Opin. Neurobiol.* 17, 609–618. doi: 10.1016/j.conb.2007.11.001
- Cunningham, J. P., and Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* 17, 1500–1509. doi: 10.1038/nn.3776
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience*, vol. 806. Cambridge, MA: MIT Press.
- Del Bene, F., Wyart, C., Robles, E., Tran, A., Looger, L., Scott, E. K., et al. (2010). Filtering of visual information in the tectum by an identified neural circuit. *Science* 330, 669–673. doi: 10.1126/science.1192949
- Delaney, T., O'Donnell, C., and Ashby, M. (2018). *How Do Calcium Indicator Properties Affect Spike Inference?* Denver, CO: Cosyne Abstracts.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* 39, 1–38.
- Ecker, A. S., Berens, P., Cotton, R. J., Subramanian, M., Denfield, G. H., Cadwell, C. R., et al. (2014). State dependence of noise correlations in macaque primary visual cortex. *Neuron* 82, 235–248. doi: 10.1016/j.neuron.2014.02.006
- Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The Elements of Statistical Learning*. New York, NY: Springer-Verlag.
- Friedrich, J., and Paninski, L. (2016). “Fast active set methods for online spike inference from calcium imaging,” in *Advances in Neural Information Processing Systems*, eds D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc.), 1984–1992.
- Gabriel, J. P., Trivedi, C. A., Maurer, C. M., Ryu, S., and Bollmann, J. H. (2012). Layer-specific targeting of direction-selective neurons in the zebrafish optic tectum. *Neuron* 76, 1147–1160. doi: 10.1016/j.neuron.2012.12.003
- Gallagher, N., Ulrich, K. R., Talbot, A., Dzirasa, K., Carin, L., and Carlson, D. E. (2017). “Cross-spectral factor analysis,” in *Advances in Neural Information Processing Systems*, eds I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.), 6842–6852.
- Ganmor, E., Krumin, M., Rossi, L. F., Carandini, M., and Simoncelli, E. P. (2016). Direct estimation of firing rates from calcium imaging data. arXiv: 1601.00364.
- Gao, Y., Archer, E. W., Paninski, L., and Cunningham, J. P. (2016). “Linear dynamical neural population models through nonlinear embeddings,” in *Advances in Neural Information Processing Systems*, eds D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc.), 163–171.
- Ghahramani, Z., and Hinton, G. E. (1996). *The EM Algorithm for Mixtures of Factor Analyzers*. Technical report, Technical Report CRG-TR-96-1, University of Toronto.
- Gramma, A., and Engert, F. (2012). Direction selectivity in the larval zebrafish tectum is mediated by asymmetric inhibition. *Front. Neural Circuits* 6:59. doi: 10.3389/fncir.2012.00059
- Haesemeyer, M., Robson, D. N., Li, J. M., Schier, A. F., and Engert, F. (2018). A brain-wide circuit model of heat-evoked swimming behavior in larval zebrafish. *Neuron* 98, 817–831.e6. doi: 10.1016/j.neuron.2018.04.013

- Hultman, R., Ulrich, K., Sachs, B. D., Blount, C., Carlson, D. E., Ndubuizu, N., et al. (2018). Brain-wide electrical spatiotemporal dynamics encode depression vulnerability. *Cell* 173, 166–180. doi: 10.1016/j.cell.2018.02.012
- Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature* 551, 232–236. doi: 10.1038/nature24636
- Kao, J. C., Nuyujukian, P., Ryu, S. I., Churchland, M. M., Cunningham, J. P., and Shenoy, K. V. (2015). Single-trial dynamics of motor cortex and their applications to brain-machine interfaces. *Nat. Commun.* 6:7759. doi: 10.1038/ncomms8759
- Khan, A. G., Poort, J., Chadwick, A., Blot, A., Sahani, M., Mrcsic-Flogel, T. D., et al. (2018). Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nat. Neurosci.* 21, 851–859. doi: 10.1038/s41593-018-0143-z
- Knöpfel, T., Gallero-Salas, Y., and Song, C. (2015). Genetically encoded voltage indicators for large scale cortical imaging come of age. *Curr. Opin. Chem. Biol.* 27, 75–83. doi: 10.1016/j.cbpa.2015.06.006
- Latimer, K., Rieke, F., and Pillow, J. W. (2018). Inferring synaptic inputs from spikes with a conductance-based neural encoding model. bioRxiv 281089. doi: 10.1101/281089
- Lin, I.-C., Okun, M., Carandini, M., and Harris, K. D. (2015). The nature of shared cortical variability. *Neuron* 87, 644–656. doi: 10.1016/j.neuron.2015.06.035
- Linderman, S., Adams, R. P., and Pillow, J. W. (2016). “Bayesian latent structure discovery from multi-neuron recordings,” in *Advances in Neural Information Processing Systems*, eds D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc.), 2002–2010.
- Linderman, S. W., Johnson, M. J., Wilson, M. A., and Chen, Z. (2016). A Bayesian nonparametric approach for uncovering rat hippocampal population codes during spatial navigation. *J. Neurosci. Methods* 263, 36–47. doi: 10.1016/j.jneumeth.2016.01.022
- Lowe, A. S., Nikolaou, N., Hunter, P. R., Thompson, I. D., and Meyer, M. P. (2013). A systems-based dissection of retinal inputs to the zebrafish tectum reveals different rules for different functional classes during development. *J. Neurosci.* 33, 13946–13956. doi: 10.1523/JNEUROSCI.1866-13.2013
- Macke, J. H., Buesing, L., Cunningham, J. P., Byron, M. Y., Shenoy, K. V., and Sahani, M. (2011). “Empirical models of spiking in neural populations,” in *Advances in Neural Information Processing Systems*, eds J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger (Curran Associates, Inc.), 1350–1358.
- Meyer, A. F., Williamson, R. S., Linden, J. F., and Sahani, M. (2017). Models of neuronal stimulus-response functions: elaboration, estimation, and evaluation. *Front. Syst. Neurosci.* 10:109. doi: 10.3389/fnsys.2016.00109
- Nam, H. (2015). *Poisson Extension of Gaussian Process Factor Analysis for Modeling Spiking Neural Populations*. Master’s thesis, Department of Neural Computation and Behaviour, Max Planck Institute for Biological Cybernetics, Tübingen.
- Naumann, E. A., Fitzgerald, J. E., Dunn, T. W., Rihel, J., Sompolinsky, H., and Engert, F. (2016). From whole-brain data to functional circuit models: the zebrafish optomotor response. *Cell* 167, 947–960. doi: 10.1016/j.cell.2016.10.019
- Nikolaou, N., Lowe, A. S., Walker, A. S., Abbas, F., Hunter, P. R., Thompson, I. D., et al. (2012). Parametric functional maps of visual inputs to the tectum. *Neuron* 76, 317–324. doi: 10.1016/j.neuron.2012.08.040
- Pandarinath, C., Ames, K. C., Russo, A. A., Farshchian, A., Miller, L. E., Dyer, E. L., et al. (2018a). Latent factors and dynamics in motor cortex and their application to brain-machine interfaces. *J. Neurosci.* 38, 9390–9401. doi: 10.1523/JNEUROSCI.1669-18.2018
- Pandarinath, C., O’Shea, D. J., Collins, J., Jozefowicz, R., Stavisky, S. D., Kao, J. C., et al. (2018b). Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat. Methods* 15, 805–815. doi: 10.1038/s41592-018-0109-9
- Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network* 15, 243–262. doi: 10.1088/0954-898X_15_4_002
- Paninski, L., Ahmadian, Y., Ferreira, D. G., Koyama, S., Rahnama Rad, K., Vidne, M., et al. (2010). A new look at state-space models for neural data. *J. Comput. Neurosci.* 29, 107–126. doi: 10.1007/s10827-009-0179-x
- Paninski, L., and Cunningham, J. P. (2018). Neural data science: accelerating the experiment-analysis-theory cycle in large-scale neuroscience. *Curr. Opin. Neurobiol.* 50, 232–241. doi: 10.1016/j.conb.2018.04.007
- Paninski, L., Pillow, J., and Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. *Prog. Brain Res.* 165, 493–507. doi: 10.1016/S0079-6123(06)65031-0
- Park, I. M., Meister, M. L., Huk, A. C., and Pillow, J. W. (2014). Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nat. Neurosci.* 17, 1395–1403. doi: 10.1038/nn.3800
- Park, M., and Pillow, J. W. (2011). Receptive field inference with localized priors. *PLoS Comput. Biol.* 7:e1002219. doi: 10.1371/journal.pcbi.1002219
- Pfau, D., Pnevmatikakis, E. A., and Paninski, L. (2013). “Robust learning of lowdimensional dynamics from large neural ensembles,” in *Advances in Neural Information Processing Systems*, eds C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Curran Associates, Inc.), 2391–2399.
- Pillow, J. (2007). “Likelihood-based approaches to modeling the neural code,” in *Bayesian Brain: Probabilistic Approaches to Neural Coding*, eds K. Doya, S. Ishii, A. Pouget, and R. P. N. Rao. (Cambridge, MA: MIT Press), 53–70.
- Pillow, J. W., Paninski, L., Uzzell, V. J., Simoncelli, E. P., and Chichilnisky, E. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J. Neurosci.* 25, 11003–11013. doi: 10.1523/JNEUROSCI.3305-05.2005
- Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E., et al. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454, 995–999. doi: 10.1038/nature07140
- Pnevmatikakis, E. A., Merel, J., Pakman, A., and Paninski, L. (2013). “Bayesian spike inference from calcium imaging data,” in *Signals, Systems and Computers, 2013 Asilomar Conference on* (Pacific Grove, CA: IEEE), 349–353.
- Pnevmatikakis, E. A., and Paninski, L. (2013). “Sparse nonnegative deconvolution for compressive calcium imaging: algorithms and phase transitions,” in *Advances in Neural Information Processing Systems*, eds C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Curran Associates, Inc.), 1250–1258.
- Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., et al. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* 89, 285–299. doi: 10.1016/j.neuron.2015.11.037
- Preuss, S. J., Trivedi, C. A., vom Berg-Maurer, C. M., Ryu, S., and Bollmann, J. H. (2014). Classification of object size in retinotectal microcircuits. *Curr. Biol.* 24, 2376–2385. doi: 10.1016/j.cub.2014.09.012
- Quian Quiroga R., and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* 10, 173–185. doi: 10.1038/nrn2578
- Real, E., Asari, H., Gollisch, T., and Meister, M. (2017). Neural circuit inference from function to structure. *Curr. Biol.* 27, 189–198. doi: 10.1016/j.cub.2016.11.040
- Runyan, C. A., Piasini, E., Panzeri, S., and Harvey, C. D. (2017). Distinct timescales of population coding across cortex. *Nature* 548, 92–96. doi: 10.1038/nature23020
- Sahani, M., and Linden, J. F. (2003). “Evidence optimization techniques for estimating stimulus-response functions,” in *Advances in Neural Information Processing Systems*, eds S. Becker, S. Thrun, and K. Obermayer (MIT Press), 317–324.
- Santhanam, G., Yu, B. M., Gilja, V., Ryu, S. I., Afshar, A., Sahani, M., et al. (2009). Factor-analysis methods for higher-performance neural prostheses. *J. Neurophysiol.* 102, 1315–1330. doi: 10.1152/jn.00097.2009
- Schneidman, E., Berry II, M. J., Segev, R., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 1007–1012. doi: 10.1038/nature04701
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136
- Semedo, J., Zandvakili, A., Kohn, A., Machens, C. K., and Byron, M. Y. (2014). “Extracting latent structure from multiple interacting neural populations,” in *Advances in Neural Information Processing Systems*, eds Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Curran Associates, Inc.), 2942–2950.

- Simoncelli, E. P., Paninski, L., Pillow, J., and Schwartz, O. (2004). *Characterization of Neural Responses With Stochastic Stimuli*, Vol. 3. Cambridge, MA: MIT Press.
- Speiser, A., Yan, J., Archer, E. W., Buesing, L., Turaga, S. C., and Macke, J. H. (2017). “Fast amortized inference of neural activity from calcium imaging data with variational autoencoders,” in *Advances in Neural Information Processing Systems*, eds I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.), 4024–4034.
- Vogelstein, J. T., Packer, A. M., Machado, T. A., Sippy, T., Babadi, B., Yuste, R., et al. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J. Neurophysiol.* 104, 3691–3704. doi: 10.1152/jn.01073.2009
- Vogelstein, J. T., Watson, B. O., Packer, A. M., Yuste, R., Jedynak, B., and Paninski, L. (2009). Spike inference from calcium imaging using sequential Monte Carlo methods. *Biophys. J.* 97, 636–655. doi: 10.1016/j.bpj.2008.08.005
- Williams, C. K., and Rasmussen, C. E. (2006). *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.
- Wu, A., Roy, N. G., Keeley, S., and Pillow, J. W. (2017). “Gaussian process based nonlinear latent structure discovery in multivariate spike train data,” in *Advances in Neural Information Processing Systems*, eds I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.), 3499–3508.
- Wu, M. C., David, S. V., and Gallant, J. L. (2006). Complete functional characterization of sensory neurons by system identification. *Annu. Rev. Neurosci.* 29, 477–505. doi: 10.1146/annurev.neuro.29.051605.113024
- Yu, B. M., Cunningham, J. P., Santhanam, G., Ryu, S. I., Shenoy, K. V., and Sahani, M. (2009). “Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity,” in *Advances in Neural Information Processing Systems*, eds D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou (Curran Associates, Inc.), 1881–1888.
- Zhao, Y., and Park, I. M. (2017). Variational latent Gaussian process for recovering single-trial dynamics from population spike trains. *Neural Comput.* 29, 1293–1316. doi: 10.1162/NECO_a_00953
- Zoltowski, D. M., and Pillow, J. W. (2018). “Scaling the Poisson GLM to massive neural datasets through polynomial approximations,” in *Advances in Neural Information Processing Systems 31*, eds S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc.), 3521–3531.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Triplett and Goodhill. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.