SpikeCaKe: Semi-Analytic Nonparametric Bayesian Inference for Spike-Spike Neuronal Connectivity

L. Ambrogioni
Radboud University

P. Ebel
University of Amsterdam

M. Hinne
Radboud University

U. Güçlü
Radboud University

M. van Gerven
Radboud University

E. Maris
Radboud University

Abstract

In this paper we introduce a semi-analytic variational framework for approximating the posterior of a Gaussian processes coupled through non-linear emission models. While the semi-analytic method can be applied to a large class of models, the present paper is devoted to the analysis of effective connectivity between biological spiking neurons. Estimating effective connectivity between spiking neurons from measured spike sequences is one of the main challenges of systems neuroscience. This semi-analytic method exploits the tractability of GP regression when the membrane potential is observed. The resulting posterior is then marginalized analytically in order to obtain the posterior of the response functions given the spike sequences alone. We validate our methods on both simulated data and real neuronal recordings.

1 Introduction

Gaussian process (GP) regression is an elegant technique for performing inference on infinite-dimensional functional spaces from a finite number of data points. When the emission model is Gaussian and linear, the posterior distribution of a GP regression can be obtained in closed-form [Rasmussen, 2006]. This is an enormous advantage since approximate inference schemes are often unreliable when the dimensionality of the latent space is large. In this paper we introduce a semi-analytic variational method that decouples the emission model from the latent Gaussian process and, consequently, leverages the analytic solution in the latent space. Using this general framework, we derive a method for the estimation of the response function between biological spiking neurons. The method relies on an important neurophysiological fact concerning neuronal communication: the membrane potential responds approximately linearly to weak synaptic inputs while spike initiation is a highly non-linear function of the membrane potential [Jagadeesh et al., 1993, Spruston, 2008].

Action potentials (spikes) are the fundamental units of neuronal communication [Rieke, 1999]. Spikes originate from the axon hillock and propagate through the axon towards the synaptic terminal, where the release of neurotransmitters affects the membrane potential of the downstream neurons. While there is a great deal of computation in the dynamics of a single neuron [Koch, 2004], most of the computational capabilities of biological neuronal networks depend on their pattern of interconnections [Sporns, 2010]. Mapping coupling between spiking neurons is therefore a major goal in system neuroscience. However, inferring effective connectivity from spike sequences is a challenging data analysis problem as networks of spiking neurons are highly non-linear dynamical systems [Izhikevich, 2007].

We begin with a simpler but practically relevant inference where both spikes and membrane potentials are observed variables. In this case, the posterior distribution of the resulting connectivity model can be obtained analytically and is related to the GP-CaKe method for field-field effective connectivity [Ambrogioni et al., 2017]. The main methodological contribution of the paper is to extend the framework to the case when only the spike sequences are measured. In this situation, the Bayesian model cannot be

Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS) 2019, Naha, Okinawa, Japan. PMLR: Volume 89.

solved in closed form since the spike initiation model is non-Gaussian. Furthermore, approximate inference is complicated by the intractability of the inhomogeneous Poisson likelihood [Kingman, 1964, Adams et al., 2009]. To resolve these difficulties, we use our new semi-analytic variational approximation that combines the analytic solution of the response function given the membrane potential with a likelihood-free stochastic estimation of the membrane potential.

2 Semi-analytic variational Gaussian process regression

We start by introducing the semi-analytic method with a simpler class of problems where a latent function $f(x)$ generates observations y through a non-Gaussian likelihood. The most common example of this class of problems is GP classification, where the observations are binary variables [Rasmussen, 2006]. Consider the following observation model:

$$y_j \sim p(y \mid f(t_j)) \ , \tag{1}$$

where p is an arbitrary distribution conditional on $f(x_j)$. We assume that the latent function f follows a GP distribution with zero mean and covariance function $k(t, t')$. Given a set $\{t_1, ..., t_N\}$ of observation time points, we collect the observations and the latent values into the vectors $\boldsymbol{y} = (y_{t_1}, ..., y_{t_n})$ and $\boldsymbol{f} = (f(t_1), ..., f(t_n))$. We denote the covariance matrix of f at the observation time points as K. Our aim is to obtain the posterior distribution of the values of the latent functions on an arbitrary set of target time points $\{t_1^*, ..., t_m^*\}$. We collect these values in the vector $\boldsymbol{f}^*$ and we denote the cross-covariance matrix between observation and target time points as K_*. Furthermore, we denote the covariance matrix at the target points as K_{**}. To find an approximate posterior distribution, we adopt a joint-contrastive variational scheme [Ambrogioni et al., 2018, Huszár, 2017]. We begin by defining the factorized variational joint distribution:

$$q(\boldsymbol{f}, \boldsymbol{f}^*, \boldsymbol{y}) = q(\boldsymbol{f}^* \mid \boldsymbol{f})\, q(\boldsymbol{f} \mid \boldsymbol{y})\, p(\boldsymbol{y}) \tag{2}$$

We obtain the variational distributions $q(\boldsymbol{f}^* \mid \boldsymbol{f})$ and $q(\boldsymbol{f} \mid \boldsymbol{y})$ by minimizing the forward KL divergence between the variational joint and the model joint:

$$\begin{aligned} D_{KL}(p\|q) &= \mathbb{E}_{\boldsymbol{f},\boldsymbol{f}^*,\boldsymbol{y}\sim p}\left[\log \frac{p(\boldsymbol{f}, \boldsymbol{f}^*, \boldsymbol{y})}{q(\boldsymbol{f}, \boldsymbol{f}^*, \boldsymbol{y})}\right] \\ &\doteq_q \mathbb{E}_{\boldsymbol{f},\boldsymbol{f}^*,\boldsymbol{y}\sim p}\left[\log \frac{p(\boldsymbol{f}^* \mid \boldsymbol{f})}{q(\boldsymbol{f}^* \mid \boldsymbol{f})}\right] - \mathbb{E}_{\boldsymbol{f},\boldsymbol{y}\sim p}[\log q(\boldsymbol{f} \mid \boldsymbol{y})] \ , \end{aligned} \tag{3}$$

where $\doteq_q$ denotes equality up to additive terms that are constant in the variational distributions $q(\boldsymbol{f}^* \mid \boldsymbol{f})$ and $q(\boldsymbol{f} \mid \boldsymbol{y})$. The first term of this expression is the KL divergence between $p(\boldsymbol{f}^* \mid \boldsymbol{f})$ and $q(\boldsymbol{f}^* \mid \boldsymbol{f})$. Since $p(\boldsymbol{f}^* \mid \boldsymbol{f})$ is the posterior distribution of a noiseless GP regression, the variational distribution $q(\boldsymbol{f}^* \mid \boldsymbol{f})$ is a multivariate Gaussian distribution and can be optimized analytically:

$$q(\boldsymbol{f}^* \mid \boldsymbol{f}) = \mathcal{N}\big(\boldsymbol{f}^*; K_* K^{-1} \boldsymbol{f}, K_{\text{post}}\big) \ ,$$

where $K_{\text{post}} = K_{**} - K_* K^{-1} K_*^T$ is the posterior covariance matrix. The remaining term involves the variational distribution $q(\boldsymbol{f} \mid \boldsymbol{y})$ and cannot be optimized in closed-form. We parameterize this term as a mixture of Gaussians:

$$q(\boldsymbol{f} \mid \boldsymbol{y}) = \sum_k^K w(\boldsymbol{y}) \mathcal{N}(\boldsymbol{f}; \mu_k(\boldsymbol{y}), \Sigma_k(\boldsymbol{y})) \ , \tag{4}$$

where the weights, means and covariances are given by parameterized nonlinear functions of $\boldsymbol{y}$. These transformations can be trained by minimizing the remaining part of the variational loss using stochastic gradient descent:

$$\mathcal{L}[q] = -\mathbb{E}_{\boldsymbol{f},\boldsymbol{y}\sim p}[\log q(\boldsymbol{f} \mid \boldsymbol{y})] \ . \tag{5}$$

This loss has a simple interpretation. We sample both $\boldsymbol{f}$ and synthetic observations $\boldsymbol{y}$ from the generative model and we train a model to invert this operation by maximizing the probability of the observations given the latent. This is an example of amortized inference since we train a whole family of posterior distributions at once. Finally, we can obtain the marginalized variational posterior $q(\boldsymbol{f}^* \mid \boldsymbol{y})$ by marginalizing out $\boldsymbol{f}$ analytically:

$$\begin{aligned} q(\boldsymbol{f}^* \mid \boldsymbol{y}) &= \int p(\boldsymbol{f}^* \mid \boldsymbol{f})\, q(\boldsymbol{f} \mid \boldsymbol{y})\, \mathrm{d}\boldsymbol{f} \\ &= \sum_k^K w(\boldsymbol{y}) \int p(\boldsymbol{f}^* \mid \boldsymbol{f})\, \mathcal{N}(\boldsymbol{f}; \mu_k(\boldsymbol{y}), \Sigma_k(\boldsymbol{y})) \\ &= \sum_k^K w(\boldsymbol{y}) \mathcal{N}\big(\boldsymbol{f}^*; K_* K^{-1} \mu_k(\boldsymbol{y}), C_k\big) \ , \end{aligned} \tag{6}$$

where $C_k = K_{\text{post}} + K_* K^{-1} \Sigma_k(\boldsymbol{y}) K^{-1} K_*^T$.

3 Spike-membrane effective connectivity analysis

We can now move to neuronal connectivity analysis. We begin by introducing a nonparametric Bayesian method for estimating the response function when both spike sequences and membrane potentials are observed variables. Besides its intrinsic relevance in several experimental settings, this method is also an important analytically tractable component of our

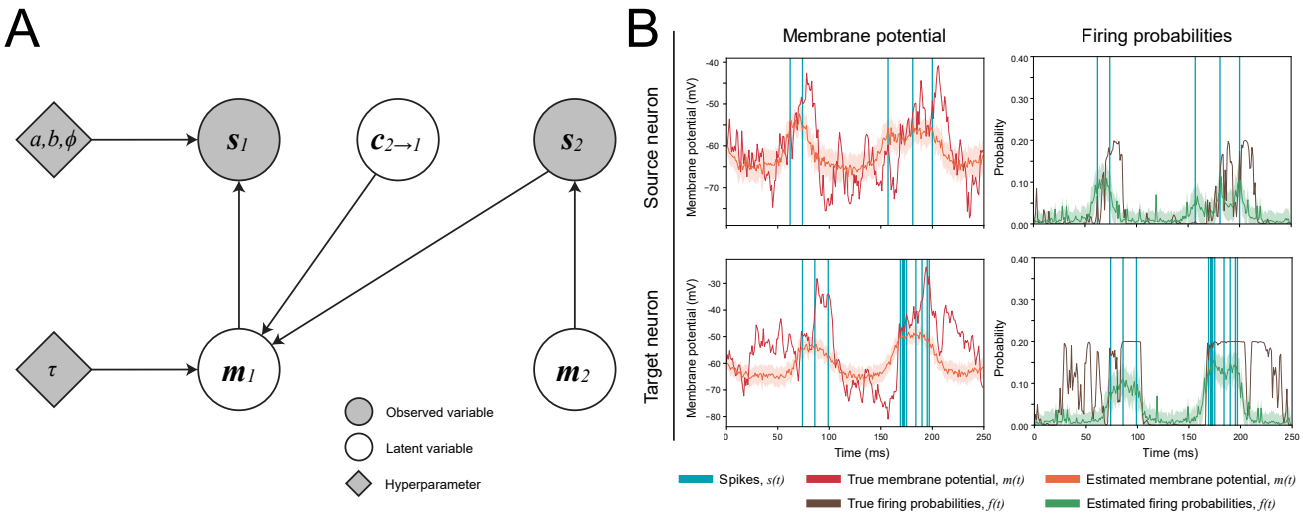

Figure 1: **A.** The generative model as explained in the text. For simplicity of the notation, the relevant variables are only shown for two neurons. Note that the membrane potentials m can be either observed or latent. In the latter case we use the variational approach. **B.** A draw from the generative model for two neurons connected through a single unidirectional excitatory connection as well as the variational recovery of the membrane and action potentials. The shaded regions indicate one standard deviation.

method for spike-spike connectivity. We begin by defining a linear dynamical model of the membrane potential that captures the linear response of the membrane potentials to weak synaptic inputs.

3.1 A linear dynamical model of the membrane potential

Consider a network of N interconnected neurons. In the following, we will denote the membrane potential of the j-th neuron as $m_j(t)$ and its spike sequence as the sum of delta functions $s_k(t) = \sum_k \delta(t - t_{j,k})$ where $t_{j,k}$ is the timestamp of the k-th spike of the j-th neuron. The linear response of a neuronal membrane to a synaptic input can be described using a differential equation [Dayan and Abbott, 2001]:

$$\tau \frac{\mathrm{d}}{\mathrm{d}t} m_j(t) = -m_j(t) + \mathcal{I}_j(t) \ , \tag{7}$$

where the time constant τ determines the time that the membrane needs to return to baseline after a perturbation. The synaptic input from the other $N-1$ neurons in the network is given by the following function:

$$\mathcal{I}_j(t) = \sum_{k=1, k\neq j}^{N} c_{k\to j}(t) \star s_k(t) + w_j(t) \ , \tag{8}$$

where the operator $\star$ denotes convolution. The additional stochastic term $w_j(t)$ is Gaussian white noise with variance σ^2 and accounts for unmeasured perturbations. The causality of the neuronal network is guaranteed as the response function $c_{k\to j}(t)$ vanishes for negative values of t.

3.2 Analytic GP regression for spike-membrane effective connectivity

We use the dynamical model specified by Eq. 7 and Eq. 8 as an implicit likelihood of a nonparametric Bayesian model. The model is defined by assigning a GP prior over the space of response functions $c_{k\to j}$. The posterior distribution of $c_{k\to j}$ is a GP and can be obtained in closed-form because both the derivative and the convolution in Eq. 7 and Eq. 8 are linear operators. In the frequency domain, Eq. 7 and Eq. 8 can be jointly written as

$$m_j(\omega) = \sum_{k=1}^{N} c_{k\to j}(\omega)\gamma_k(\omega) + \tilde{w}_j(\omega) \ , \tag{9}$$

where

$$\gamma_k(\omega) = (-i\omega\tau + 1)^{-1} s_k(\omega) = (-i\omega\tau + 1)^{-1} \sum_j e^{-i\omega t_{k,j}}$$

and

$$\tilde{w}_j(\omega) = (-i\omega\tau + 1)^{-1} w_j(\omega) \ .$$

Eq. 7 defines a nonparametric regression problem where $m_j(\omega)$ is the observed data, $\gamma_k(\omega)$ are known mixing functions and $c_{k\to j}(\omega)$ are the unknowns. Problems of this form have an analytic solution when the prior distributions over $c_{k\to j}(\omega)$ are GPs [Rasmussen, 2006]. To assure the causality of the response

functions we adopt the causal covariance function that was introduced in [Ambrogioni et al., 2017]. In the frequency domain this covariance function can be expressed as

$$\begin{aligned}&\mathfrak{K}(\omega_1,\omega_2) = \\ &f(\omega_1,\omega_2)\ \left(\mathfrak{s}_{SE}(\omega_2-\omega_1)+i\mathcal{H}\mathfrak{s}_{SE}(\omega_2-\omega_1)\right)\ ,\end{aligned} \tag{10}$$

where $f(\omega_1,\omega_2)$ is a function that induces smoothness by discounting the high frequency components, $\mathfrak{s}_{SE}(\omega)$ is the spectral density of a squared exponential covariance function and $\mathcal{H}$ denotes the Hilbert transform which enforces causality. The resulting GP prior induces causality, smoothness and temporal localization of the response function. See Appendix A for more details on the construction of this covariance function.

Consider a set of M time points $\{t_1,...,t_M\}$ and a vector of measured membrane potentials $m_u = m(t_u)$. The posterior expected value of $c_{k\to j}$ is given by

$$\bar{c}_{k\to j}(t) = \sum_{u,v} W_{uv} m_v \mathfrak{K}(t,t_u)\ , \tag{11}$$

where the GP weights W_{uv} depend on the covariance function and can be obtained using standard GP regression techniques in the frequency domain. The matrix formula for the weights is given in Appendix B. The time domain covariance function in Eq. 11 is the inverse Fourier transform of Eq. 10 with respect to both of its arguments.

4 Spike-spike effective connectivity analysis

We can now use the results of the previous section in order to derive a semi-analytic solution to the more challenging problem of spike-spike connectivity.

4.1 A non-linear model of spike initiation

In biological neurons, spike initiation depends on the non-linear dynamics of the membrane potential and of several ionic channels [Izhikevich, 2007]. We approximate these dynamics using a stochastic model. Specifically, the firing rate $f(t)$ is obtained by passing the rescaled membrane potential through a compressive non-linearity:

$$f(t) = a\ \sigma(b\ m_j(t)+\phi)\ , \tag{12}$$

where a is the maximum firing rate and $\sigma(\cdot)$ is the logistic sigmoid with b and ϕ its gain and threshold parameters respectively. The resulting spike sequence follows a nonhomogeneous Poisson process with density function $f(t)$ [Kingman, 1964]. This model is admittedly a simplification. For example, it does not take into account the refractory period [Kandel et al., 2000]. However, we adopt this model only for the sake of simplicity and the variational Bayesian model that is introduced in the next section can be used with any other spike initiation model without substantial modifications.

4.2 Semi-analytic variational GP regression for spike-spike effective connectivity

To simplify the notation we will explain the analysis for the case of two neurons. All results generalize straightforwardly to arbitrary network structures. Given a set of M sample time points $\{t_1,...,t_M\}$, we organize the sampled time-series in the arrays $\boldsymbol{s}_j = (s_j(t_1),..,s_j(t_M))$, $\boldsymbol{m}_j = (m_j(t_1),..,m_j(t_M))$ and $\boldsymbol{c}_{2\to1} = (c_{2\to1}(-t_M/2),..,c_{2\to1}(t_M/2))$. The graphical model is shown in Fig. 1. This model is summarized by the following factorized joint distribution:

$$\begin{aligned}&p(\boldsymbol{s}_1,\boldsymbol{m}_1,\boldsymbol{c}_{2\to1}\mid\boldsymbol{s}_2) = \\ &p(\boldsymbol{s}_1\mid\boldsymbol{m}_1)\,p(\boldsymbol{m}_1\mid\boldsymbol{c}_{2\to1},\boldsymbol{s}_2)\,p(\boldsymbol{c}_{2\to1})\ ,\end{aligned} \tag{13}$$

where we conditioned on the spike sequence $\boldsymbol{s}_2$. Our aim is to obtain $p(\boldsymbol{c}_{2\to1}\mid\boldsymbol{s}_1,\boldsymbol{s}_2)$, i.e., the posterior distribution of the response function given the two spike sequences. Most existing variational methods do not directly leverage the analytic solution of $p(\boldsymbol{c}_{2\to1}\mid\boldsymbol{m}_1,\boldsymbol{s}_2)$ and require the evaluation of the intractable likelihood $p(\boldsymbol{s}_1\mid\boldsymbol{m}_1)$ [Hoffman et al., 2013, Ranganath et al., 2014, Rezende et al., 2014]. Therefore we developed a new semi-analytic variational approximation that fully exploits the analytic tractability of the latent GP analysis. We begin by defining the following structured joint variational distribution:

$$\begin{aligned}&q(\boldsymbol{s}_1,\boldsymbol{m}_1,\boldsymbol{c}_{2\to1}\mid\boldsymbol{s}_2) = \\ &q(\boldsymbol{c}_{2\to1}\mid\boldsymbol{m}_1,\boldsymbol{s}_2)\,q(\boldsymbol{m}_1\mid\boldsymbol{s}_1)\,p(\boldsymbol{s}_1)\ ,\end{aligned} \tag{14}$$

where $p(\boldsymbol{s}_1)$ is the real marginal distribution of $\boldsymbol{s}_1$. In this variational factorization we assumed that the distribution of the membrane potential $\boldsymbol{m}_1$ solely depends on the spike sequence $\boldsymbol{s}_1$. We can find the distributions $q(\boldsymbol{c}_{2\to1}\mid\boldsymbol{m}_1,\boldsymbol{s}_2)$ and $q(\boldsymbol{m}_1\mid\boldsymbol{s}_1)$ by minimizing the following functional:

$$\mathcal{L}[q] = \mathbb{E}_{p(\boldsymbol{s}_2)}[D_{KL}(p\|q)]\ . \tag{15}$$

Note that this functional is a proper (joint-contrastive) variational loss since it is always non-negative and vanishes if and only if $p(\boldsymbol{s}_1,\boldsymbol{m}_1,\boldsymbol{c}_{2\to1},\boldsymbol{s}_2) =$

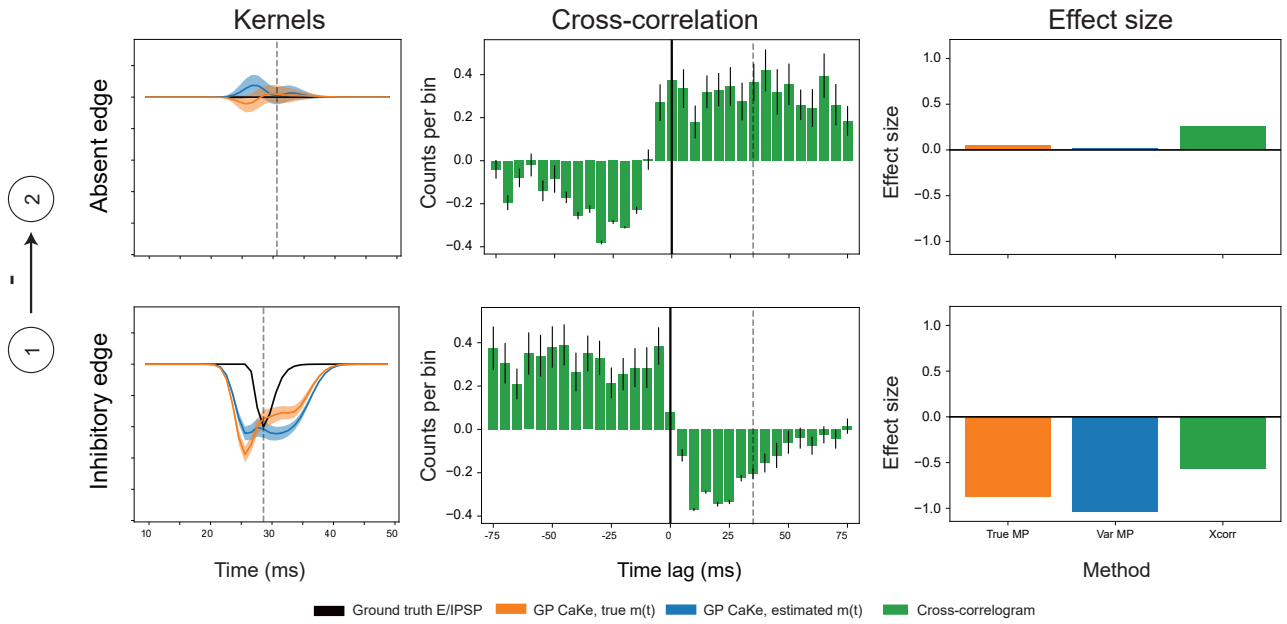

Figure 2: Examples of reconstructed causal kernels for both variants of SpikeCaKe together with the corresponding cross-correlograms and effect sizes for each of the three methods, per connection. Shaded intervals and error bars indicate 95% confidence intervals.

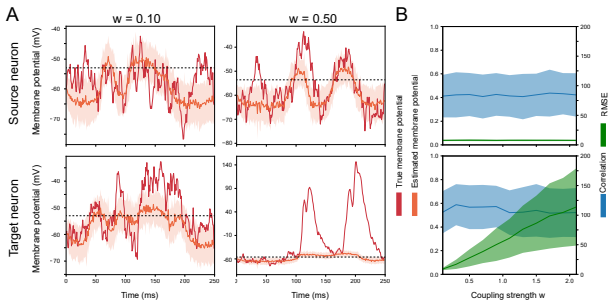

Figure 3: **A.** Variational inference of the membrane potentials for the source and target neurons of network 1 (see Fig. 4A, for $w \in \{0.1, 0.5\}$). The dashed line indicates the membrane threshold. **B.** The correlation and root-mean-squared error between the true and estimated membrane potentials. Shaded intervals indicate one standard deviation.

$q(\boldsymbol{s}_1, \boldsymbol{m}_1, \boldsymbol{c}_{2\to1}, \boldsymbol{s}_2)$. We can rearrange the loss as:

$$
\begin{aligned}
&\mathbb{E}_{p(\boldsymbol{s}_2)}[D_{KL}(p\|q)] = \\
&\mathbb{E}_{p(\boldsymbol{s}_1,\boldsymbol{m}_1,\boldsymbol{c}_{2\to1},\boldsymbol{s}_2)}\left[\log \frac{p(\boldsymbol{s}_1,\boldsymbol{m}_1,\boldsymbol{c}_{2\to1} \mid \boldsymbol{s}_2)}{q(\boldsymbol{s}_1,\boldsymbol{m}_1,\boldsymbol{c}_{2\to1} \mid \boldsymbol{s}_2)}\right] \\
&\doteq_q \mathbb{E}_{p(\boldsymbol{m}_1,\boldsymbol{c}_{2\to1},\boldsymbol{s}_2)}\left[\log \frac{p(\boldsymbol{c}_{2\to1} \mid \boldsymbol{m}_1,\boldsymbol{s}_2)}{q(\boldsymbol{c}_{2\to1} \mid \boldsymbol{m}_1,\boldsymbol{s}_2)}\right] \\
&\quad - \mathbb{E}_{p(\boldsymbol{m}_1,\boldsymbol{s}_1)}[\log q(\boldsymbol{m}_1|\boldsymbol{s}_1)] \\
&= \mathbb{E}_{p(\boldsymbol{s}_2,\boldsymbol{m}_1)}[D_{KL}(p(\boldsymbol{c}_{2\to1}|\boldsymbol{s}_2,\boldsymbol{m}_1)\|q(\boldsymbol{c}_{2\to1}|\boldsymbol{s}_2,\boldsymbol{m}_1))] \\
&\quad - \mathbb{E}_{p(\boldsymbol{m}_1,\boldsymbol{s}_1)}[\log q(\boldsymbol{m}_1|\boldsymbol{s}_1)] \ , \qquad (16)
\end{aligned}
$$

where $\doteq_q$ denotes that the expressions are equal up to terms that are constant in q. The first term of this expression is an expectation of a KL divergence and therefore vanishes when $q(\boldsymbol{c}_{2\to1}|\boldsymbol{s}_2,\boldsymbol{m}_1)$ is equal to the real posterior $p(\boldsymbol{c}_{2\to1}|\boldsymbol{s}_2,\boldsymbol{m}_1)$, which can be expressed analytically (see Eq. 11). We can parameterize the remaining term as a mixture of Gaussian distributions:

$$
q(\boldsymbol{m}_1|\boldsymbol{s}_1) = \sum_h \alpha_h(\boldsymbol{s}_1)\mathcal{N}(\boldsymbol{m}_1; \boldsymbol{\mu}_h(\boldsymbol{s}_1), Q_h(\boldsymbol{s}_1)) \ , \quad (17)
$$

where the scalar-valued functions $\alpha_h(\boldsymbol{s}_1)$, the vector-valued functions $\boldsymbol{\mu}_h(\boldsymbol{s}_1)$ and the matrix-valued functions $Q_h(\boldsymbol{s}_1)$ are determined by expressive regression models such as deep convolutional networks [Goodfellow et al., 2016]. The parameters of these networks can be trained by minimizing the remaining term of the variational loss

$$
\begin{aligned}
&\mathcal{L}[p(\boldsymbol{c}_{2\to1} \mid \boldsymbol{m}_1, \boldsymbol{s}_2) \doteq_q \\
&q(\boldsymbol{m}_1 \mid \boldsymbol{s}_1)\, p(\boldsymbol{s}_1)] - \mathbb{E}_{p(\boldsymbol{m}_1,\boldsymbol{s}_1)}[\log q(\boldsymbol{m}_1|\boldsymbol{s}_1)] \ , \quad (18)
\end{aligned}
$$

whose gradient can be easily sampled without bias by sampling from the model marginal $p(\boldsymbol{m}_1, \boldsymbol{s}_1)$. Optimizing Eq.18 requires to train the regression models $\alpha_h(\boldsymbol{s}_1)$, $\boldsymbol{\mu}_h(\boldsymbol{s}_1)$ and $Q_h(\boldsymbol{s}_1)$ separately every time we want to analyze a new network structure since the distribution $p(\boldsymbol{m}_1, \boldsymbol{s}_1)$ includes the (marginalized) effects of all neurons. In order to increase the efficiency of the method we approximate $p(\boldsymbol{m}_1, \boldsymbol{s}_1)$ with the joint distribution of a single uncoupled neuron. This is a weak coupling approximation since we are assuming

the (cumulative) coupling strength between neurons to be small compared to the stochastic input. We analyze the consequences of this approximation in our experiments below. We can now obtain the variational posterior $q(\boldsymbol{c}_{2\to1} \mid \boldsymbol{s}_1, \boldsymbol{s}_2)$ by marginalizing the variational distribution analytically:

$$q(\boldsymbol{c}_{2\to1} \mid \boldsymbol{s}_1, \boldsymbol{s}_2) = \int p(\boldsymbol{c}_{2\to1} \mid \boldsymbol{m}_1, \boldsymbol{s}_2)\, q(\boldsymbol{m}_1 \mid \boldsymbol{s}_1)\, \mathrm{d}\boldsymbol{m}_1$$
$$= \sum_h \alpha_h(\boldsymbol{s}_1) \mathcal{N}\big(\boldsymbol{c}_{2\to1}; W\boldsymbol{\mu}_h(\boldsymbol{s}_1), K_p + WQ_h(\boldsymbol{s}_1)W^T\big) \ , \tag{19}$$

where W are the GP weights (see Eq. 11) and K_p is the covariance matrix of the posterior $p(\boldsymbol{c}_{2\to1} \mid \boldsymbol{m}_1, \boldsymbol{s}_2)$. We refer to our method of spike-spike connectivity estimation as SpikeCaKe (Spike GP Causal Kernels).

5 Related work

Several techniques have been used to identify spike-spike connectivity. Simple nonparametric methods such as histograms have a long history and are still widely applied [Perkel et al., 1967]. Parametric methods based on the generalized linear model (GLM) often offer a better signal-to-noise ratio [Brown et al., 2004]. The models introduced in this paper are strictly related to GP classification and can therefore be considered as the nonparametric generalization of GLM based methods [Rasmussen, 2006]. Other modern approaches are based on dynamic Bayesian networks [Eldawlatly et al., 2010] and Cox processes [Berry et al., 2012]. We will now devote special attention to methods based on Hawkes processes, given their theoretical similarity to our approach.

5.1 Spike-spike connectivity with Hawkes Processes

The multivariate stochastic process defined in this paper has some similarity with a Hawkes process [Liniger, 2009]. While most of the existing literature based on Hawkes processes assumes a simple parametrization for the response functions, several new studies introduced the use of nonparametric methods [Zhou et al., 2013, Rousseau et al., 2018, Yang et al., 2017]. Hawkes processes have been successfully used in neuroscience settings in order to infer spike-spike effective connectivity [Reynaud-Bouret et al., 2013, Lambert et al., 2017]. In a Hawkes process the spike density of the j-th unit is a linear functional of the spike sequences of the other units:

$$f_j(t) - \mu_j = \sum_k c_{k\to j}(t) \star s_k(t) \ , \tag{20}$$

where μ_j is the baseline spike density. Note that Eq. 20 is strikingly similar to our Eq. 7. The difference is that in a Hawkes Process the spike density is a linear functional of the input spike sequences while in our model the linear response is defined at the level of the latent membrane potential. From a biophysical point of view, linearity of the spike density response is not a realistic assumption since spike initiations in biological neurons are determined by highly non-linear 'threshold' events [Izhikevich, 2007]. Another obvious problem of Eq. 20 is that the spike density could become negative in the presence of inhibitory responses. Conversely, in our model a highly negative membrane potential simply corresponds to a very low but positive spike density. The similarity between Eq. 7 and Eq. 20 implies that both the analytic and the semi-analytic methods introduced in this paper can be applied to Hawkes processes as well. The analytic method cannot be applied on real data since the spike density is not directly measurable. Nevertheless we will use it as an idealized baseline comparison in our simulation studies where we know the ground truth.

6 Simulated effective connectivity

Here we validate the reconstructions by SpikeCaKe. The details of the deep neural networks used for the estimation of the membrane are given in appendix E. The performance of spike-spike connectivity methods and non-linear regression in general is strongly affected by the form of regularization used. In order to have a balanced comparison we compare the performance of our method with its equivalent Hawkes process model where the prior covariance function and the approximative inference methods are exactly the same. We also include a comparison with a simpler nonparametric method based on spike-spike histograms [Perkel et al., 1967].

First we define five different network structures, as shown in Fig. 4A. For each of these structures, which may contain both excitatory and inhibitory interactions, we generate 200 trials of observable membrane potentials and spikes according to the generative model of Section 4.2. The true connection strength w is varied to investigate its effect on the recovery of the response function. More details of the simulation procedure can be found in Appendix C. The first two networks simply demonstrate the recovery of either excitatory, inhibitory or absent coupling. An example of a single trial of simulated data is shown in Fig. 1B. The leftmost subfigures show the recovery of the membrane potential using the variational procedure. Note that as expected the spike density of neuron 2 is temporally concentrated near the spikes of the input neuron. Importantly, the transformation from membrane potentials to firing probabilities is non-linear, which is one of the main differences between SpikeCaKe and

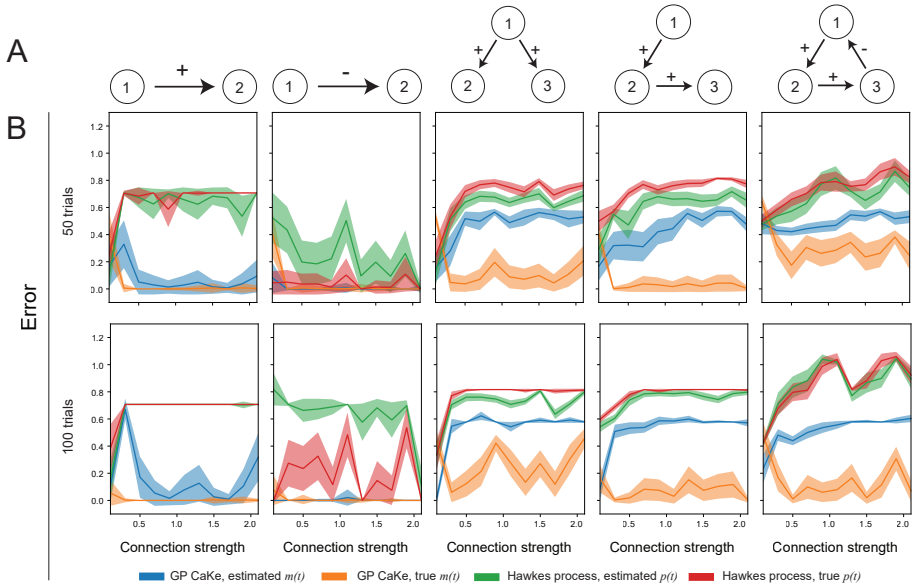

Figure 4: **A.** The considered network structures. **B.** Root-mean-squared-error between the actual connectivity matrix and the different recovery approaches shown for different numbers of bootstrapped subsamples. Interval widths indicate one standard deviation over 1 000 runs of the indicated number of bootstrapped samples.

the Hawkes process. Similarly, the rightmost figures show how the firing rates may be reconstructed using the variational method.

Figure 3 shows the variational approximation of the membrane potential for the first network, this time for different coupling strengths w. As the figure shows, the membrane potentials recovered well for the neurons that received no input (i.e. the top row of the figure). For the neurons that did receive input the membrane potential approximation deteriorates in its estimation of the magnitude when the coupling strength is increased. This is due to the violation of the assumption that neurons are only weakly coupled and have their activity predominantly driven by internal dynamics. Despite this, the correlation between the true and the estimated membrane potential remains high and, as we will show below, sufficient to recover the coupling structure.

6.1 Recovery of effective connectivity

As an example, Fig. 2 shows the recovered response functions for the two-neuron network with a single inhibitory connection. Both variants of SpikeCaKe successfully distinguish present and absent coupling and correctly identify that the present connection is inhibitory. In addition, we show the cross-correlation estimation of this connection. While this more traditional approach also identifies the inhibitory coupling, it fails to classify the other connection as absent, as can be seen from the estimated effect sizes (see Appendix C) for the two connections.

To further quantify these results we use the estimated response functions to recover the coupling structures from Fig. 4A. The presence of a connection is estimated via a z-test at the peak of the true response function while its directionality is given by the sign of the corresponding z-score (more details are provided in Appendix C). The performance is scored using the root-mean-squared-error between the true adjacency matrix describing the coupling structure and the estimated structure. The results of this analysis are shown in Fig. 4B. From these results it is apparent that both SpikeCaKe variants consistently provide the best estimates of the coupling structure. In many cases the recovery is (near) perfect.

When the coupling strength w is increased the assumption of weak coupling is again violated. We see that this is particularly detrimental for the networks with common causes and transitive effects. However, when the true membrane potential is observed, SpikeCaKe still estimates the coupling structure nearly perfectly. Also, even for these more complex cases, the variational SpikeCaKe approach outperforms both variants of the Hawkes process, even the idealized case where the true firing rates are known. Interestingly, for some networks the Hawkes process in fact performs better with the estimated firing rates than with the true ones. Presumably this is due to the smoothing induced by the variational approach, which causes the estimated firing rates to be more similar to membrane potentials.

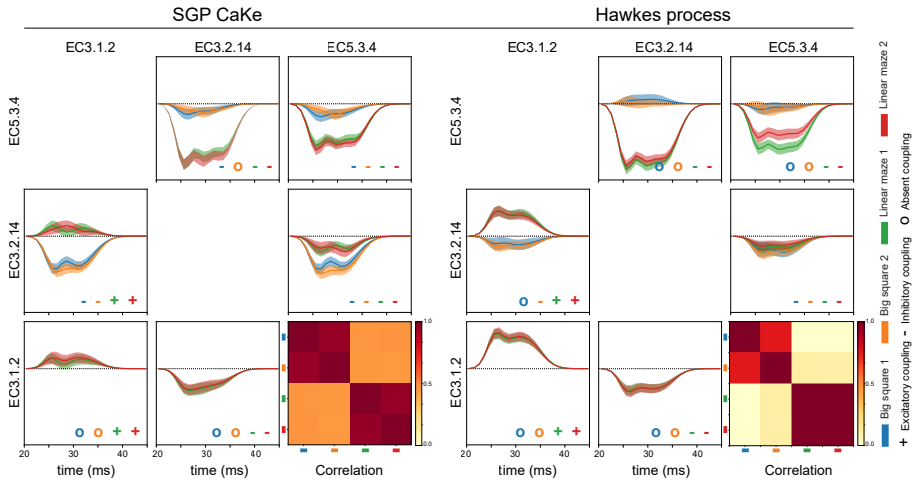

Figure 5: Estimated effective connectivity between three neuronal clusters in entorhinal cortex for four conditions (see Appendix D. The insets show the (average) correlation between the response functions for the four different conditions.

7 Analysis of real spike trains: connectivity in rat entorhinal cortex

To illustrate a more realistic application of our proposed methods we applied both the SpikeCaKe and the Hawkes process to multi-unit recordings of rat entorhinal cortex [Mizuseki et al., 2014, 2013]. Details of the data acquisition and preprocessing can be found in Appendix D. For these data sets only the spikes were observed so we estimated the membrane potentials and firing rates for SpikeCaKe and the Hawkes process respectively using the semi-analytic variational approach. As the ground truth is obviously unavailable, we estimated coupling in two different conditions (condition one consists of the rat moving freely in an open square; condition two consists of the rat navigating through a linear maze) and looked at the between-session reproducibility for validation of the procedures. The estimated response functions between three electrodes are shown in Fig. 5. Overall, SpikeCaKe and the Hawkes process resulted in similar response functions, although slight differences may be observed in the estimated coupling structure. Clearly there is strong correspondence in the response functions within conditions, while at the same time the response functions between the conditions are fairly different, showing the sensitivity of the methods. The reproducibility is further quantified in the correlations between the response functions for each pair of conditions (see inset in Fig. 5).

8 Conclusion

We introduced two new nonparametric Bayesian models for spike-membrane and spike-spike connectivity analysis. We obtain an approximate semi-analytic posterior for the spike-spike problem by minimizing a new likelihood-free variational loss. This semi-analytic method has wide applicability outside our current model since it can be used every time a latent GP regression is coupled to a non-linear emission model. For example, our semi-analytic variational method can be directly used in a calcium imaging setting where the spikes are observed through a non-linear calcium response [Packer et al., 2015].

References

C. E. Rasmussen. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

B. Jagadeesh, H. S Wheat, and D. Ferster. Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science*, 262(5141):1901–1904, 1993.

N. Spruston. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9(3):206, 2008.

F. Rieke. *Spikes: Exploring the Neural Code*. The MIT press, 1999.

C. Koch. *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press, 2004.

O. Sporns. *Networks of the Brain*. The MIT press, 2010.

E. M. Izhikevich. *Dynamical Systems in Neuroscience*. MIT Press, 2007.

L. Ambrogioni, M. Hinne, M.A.J. van Gerven, and E. Maris. GP CaKe: Effective brain connectivity

with causal kernels. *Advances in Neural Information Processing Systems*, pages 951–960, 2017.

J. F. C. Kingman. On doubly stochastic Poisson processes. *Mathematical Proceedings of the Cambridge Philosophical Society*, 60(4):923–930, 1964.

R. P. Adams, I. Murray, and D. J. C. MacKay. Tractable nonparametric Bayesian inference in Poisson processes with Gaussian process intensities. *International Conference on Machine Learning*, pages 9–16, 2009.

Luca Ambrogioni, U. Güçlü, J. Berezutskaya, E. W.P. van den Borne, Y. Güçlütürk, M. Hinne, E. Maris, and M. .J. van Gerven. Forward amortized inference for likelihood-free variational marginalization. *arXiv preprint arXiv:1805.11542*, 2018.

F. Huszár. Variational inference using implicit distributions. *arXiv preprint arXiv:1702.08235*, 2017.

P. Dayan and L. F. Abbott. *Theoretical Neuroscience*. MIT Press, 2001.

E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth. *Principles of Neural Science*. McGraw-Hill, 2000.

M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley. Stochastic variational inference. *The Journal of Machine Learning Research*, 14(1):1303–1347, 2013.

R. Ranganath, S. Gerrish, and D. Blei. Black box variational inference. *International Conference on Artificial Intelligence and Statistic*, 2014.

D. J. Rezende, S. Mohamed, and D. Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *International Conference on Machine Learning*, 2014.

I. Goodfellow, A. Courville, and Y. Bengio. *Deep Learning*, volume 1. The MIT Press, 2016.

D. H. Perkel, G. L. Gerstein, and G. P. Moore. Neuronal spike trains and stochastic point processes: II. simultaneous spike trains. *Biophysical Journal*, 7(4): 419–440, 1967.

E. N. Brown, R. E. Kass, and P. P. Mitra. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, 7(5):456, 2004.

S. Eldawlatly, Y. Zhou, R. Jin, and K. G. Oweiss. On the use of dynamic Bayesian networks in reconstructing functional neuronal networks from spike train ensembles. *Neural Computation*, 22(1):158–189, 2010.

T. Berry, F. Hamilton, N. Peixoto, and T. Sauer. Detecting connectivity changes in neuronal networks. *Journal of Neuroscience Methods*, 209(2):388–397, 2012.

T. J. Liniger. *Multivariate Hawkes processes*. ETH Zurich, 2009.

K. Zhou, H. Zha, and L. Song. Learning triggering kernels for multi-dimensional hawkes processes. *International Conference on Machine Learning*, pages 1301–1309, 2013.

J. Rousseau, S. Donnet, and V. Rivoirard. Nonparametric Bayesian estimation of multivariate Hawkes processes. *arXiv preprint arXiv:1802.05975*, 2018.

Y. Yang, J. Etesami, N. He, and N. Kiyavash. Online learning for multivariate Hawkes processes. *Advances in Neural Information Processing Systems*, pages 4944–4953, 2017.

P. Reynaud-Bouret, V. Rivoirard, and C. Tuleau-Malot. Inference of functional connectivity in neurosciences via Hawkes processes. *Global Conference on Signal and Information Processing*, pages 317–320, 2013.

R. C. Lambert, C. Tuleau-Malot, T. Bessaih, V. Rivoirard, Y. Bouret, N. Leresche, and P. Reynaud-Bouret. Reconstructing the functional connectivity of multiple spike trains using Hawkes models. *Journal of Neuroscience Methods*, 2017.

K. Mizuseki, K. Diba, E. Pastalkova, J. Teeters, A. Sirota, and G. Buzsáki. Neurosharing: large-scale data sets (spike, LFP) recorded from the hippocampal-entorhinal system in behaving rats. *F1000Research*, 2014.

K. Mizuseki, A. Sirota, E. Pastalkova, K. Diba, and G. Buzsáki. Multiple single unit recordings from different rat hippocampal and entorhinal regions while the animals were performing multiple behavioral tasks, 2013.

A. M. Packer, L. E Russell, H. W. P. Dalgleish, and M. Häusser. Simultaneous all-optical manipulation and recording of neural circuit activity with cellular resolution in vivo. *Nature Methods*, 12(2):140, 2015.