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

A Causal covariance functions

The causal covariance function used in this paper induces three main features: temporal localization, causality and smoothness. In the frequency domain, temporal localization can be implemented by inducing correlations between the Fourier coefficients of neighboring frequencies. We induce these spectral correlations with a squared exponential covariance function:

$$\mathfrak{K}_{SE}(\omega_1,\omega_2) = e^{-\vartheta\frac{(\omega_2-\omega_1)^2}{2}+it_s(\omega_2-\omega_1)} = e^{-\vartheta\frac{\zeta^2}{2}+it_s\zeta} \ , \tag{1}$$

where $\zeta = \omega_2 - \omega_1$. In order to enforce causality, we take the Hilbert transform of this squared exponential covariance function. The resulting causal covariance function is given by

$$\mathfrak{K}_C(\zeta) = \mathfrak{K}_{SE}(\zeta) + i\mathcal{H}\mathfrak{K}_{SE}(\zeta) \ . \tag{2}$$

Finally, smoothness is obtained by discounting high frequency components. We use the following discounting function:

$$f(\omega_1,\omega_2) = e^{-\nu\frac{\omega_1^2+\omega_2^2}{2}} \ . \tag{3}$$

Our final covariance function is

$$\mathfrak{K}(\omega_1,\omega_2) = f(\omega_1,\omega_2)\ (\mathfrak{K}_{SE}(\zeta) + i\mathcal{H}\mathfrak{K}_{SE}(\zeta)) \ . \tag{4}$$

B Causal connectivity using Gaussian processes

Consider a set of observable frequencies $\{u_u\}_u$. The GP weights in the frequency domain are given by:

$$\hat{W} = \hat{\Gamma}_j\Big(\sum_i \hat{\Gamma}_i\hat{K}\hat{\Gamma}_i^* + \sigma^2\mathrm{I})^{-1} \ , \tag{5}$$

where K is has entries $[\hat{K}_f]_{uv} = \mathfrak{K}(\omega_u,\omega_v)$ [Rasmussen, 2006]. The matrices $\hat{\Gamma}i$ are diagonal with the entries given by $\gamma_j(\omega_u)$. The weights can be transformed back to the time domain by multiplying Eq. 5 with a matrix of Fourier basis functions from the left.

C Simulation procedure

Each coupling structure is denoted by an adjacency matrix $\mathbf{A} = \{a_{ij}\}$ with elements a_{ij} either 0 indicating the absence of a connection, or $1, -1$ for excitatory and inhibitory connections, respectively. The adjacency matrix is multiplied with a connection strength w, which is varied to investigate the effect on the reconstruction of causal influence for increasingly stronger coupling. For the true causal response function we use a canonical EPSP (IPSP) [Koch, 2004]. With the causal response structure given, action potentials are simulated according to the generative model described in the main text, using the following biologically plausible parameter settings: threshold $\phi = -53$mV, gain $a = 0.5\mathrm{mV}^{-1}$, maximum firing rate $b = 200$Hz and membrane time constant $\tau = 15$ms. With this procedure, we generated 200 random trials for each of the five coupling structures. Effect sizes for estimated connections are obtained as the mean estimated causal response function divided by its standard deviation, at the point of the peak of the EPSP (IPSP). For absent connections, this timing is indicated with a dashed line in the figures.

Estimated connectivity for one connection $\hat{a}_{ij}$ is determined by

$$\hat{a}_{ij} = (p\text{-value} < p\text{-threshold}) \times \mathrm{sign}(z\text{-score}) \ , \tag{6}$$

with $\hat{\mathbf{A}} = \{\hat{a}_{ij}\}$, with z-scores and p-values determined over (bootstrapped) trials at time point of the peak of the true EPSP in the case of present directed connectivity, and at the halfway point for absent connections. The p-values were obtained using a simple z-test, with p-threshold $= 0.005/m$, with $m = N(N-1)$ the number of potential edges in the network, to correct for multiple comparisons. Subsequently, the root-mean-squared-error is computed between the actual adjacency matrix $\mathbf{A}$ and the estimated matrix $\hat{\mathbf{A}}$.

D Data acquisition

The CRCNS.org hc-3 data set Mizuseki et al. [2014, 2013] consists of multi-unit recordings measured in multiple hippocampal regions of rats performing different behavioral tasks. Here we analyze four recordings of animal `ec012`, acquired for two behavioral tasks; in the first condition the rat walks freely in an open square of 180 by 180cm (sessions `ec012ec.228 and 229`), in the second condition the rat walks in a linear maze (sessions `ec012ec.239 and 240`). The provided recordings are readily preprocessed into spike

trains via KlustaKwik Kadir et al. [2014] and Klusters (http://neurosuite.sourceforge.net/). The pipeline detects and groups spikes into neuronal sources, with spikes looking similar expected to stem from the same neuron. To apply our inference method on the provided spike trains, the spike-sorted data was downsampled from 20kHz to 1kHz. Subsequently, we selected the three most-active neuronal clusters within this recording session for further analysis: Two of the selected neuronal clusters are located in region entorhinal cortex layer three (EC3) and one in region entorhinal cortex layer five (EC5). The clusters are labeled as ECa.b.c, where a denotes the entorhinal cortex layer number, b is the electrode number and c the identified cluster. Finally, for each considered session the first 50 000 time points of the selected neuronal clusters are partitioned into 200 samples, each of 250ms length, which results in data dimensions similar to those used in the simulations.

E Deep networks architecture

In our experiments, the conditional variational distributions $q(\boldsymbol{m}_j, |\boldsymbol{s}_k)$ were Gaussian distributions with diagonal covariance matrix. The means and log variances of these distributions were determined by the output of a network comprised 8 of dilated convolution layers with 60 one-dimensional kernels of length three and rectified linear units. The dilation factors of the first two layers were one and doubled after every subsequent layer. The number of convolution layers was chosen to be the largest possible value such that the receptive field length of the last convolution layer was less than the signal length. The output of the network was given by two fully connected layers, one determining the means of the Gaussian $\boldsymbol{\mu}$ and the other the log variances $\boldsymbol{l}$. We initialized the bias terms to zero and the weights to samples drawn from a scaled Gaussian distribution. We used Adam with initial $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$ and a mini-batch size of 100 to train the network for 20,000 epochs by iteratively minimizing a the stochastic variational loss.

F: Estimated causal response functions

Figures 1 and 2 show the estimated causal response functions for each of the five considered network structures, for both variants of SGP CaKe as well as both variants of the Hawkes process.

References

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

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

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.

S. N. Kadir, D. F. M. Goodman, and K. D. Harris. High-dimensional cluster analysis with the masked EM algorithm. *Neural Computation*, 26(11):2379–2394, 2014.

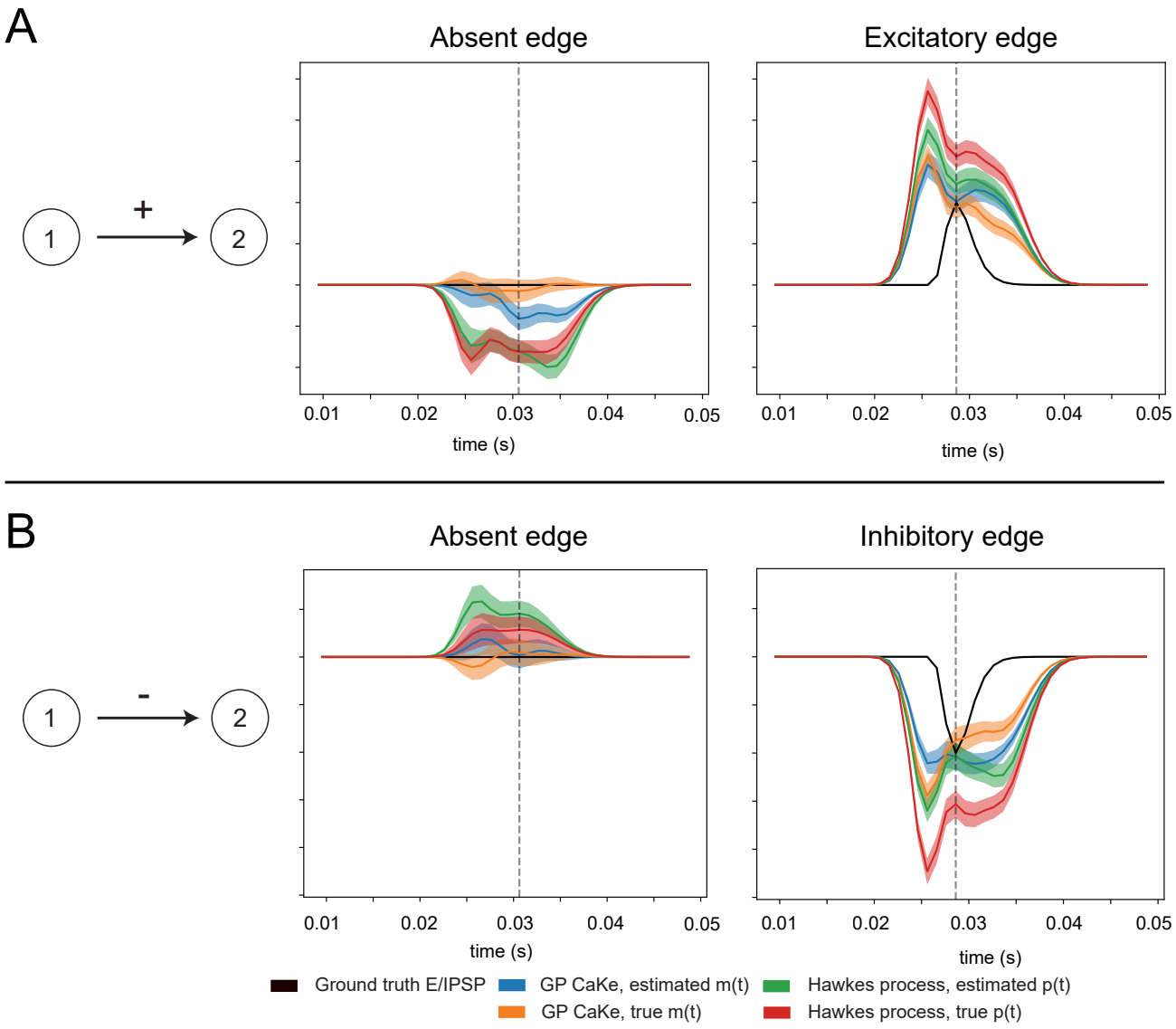
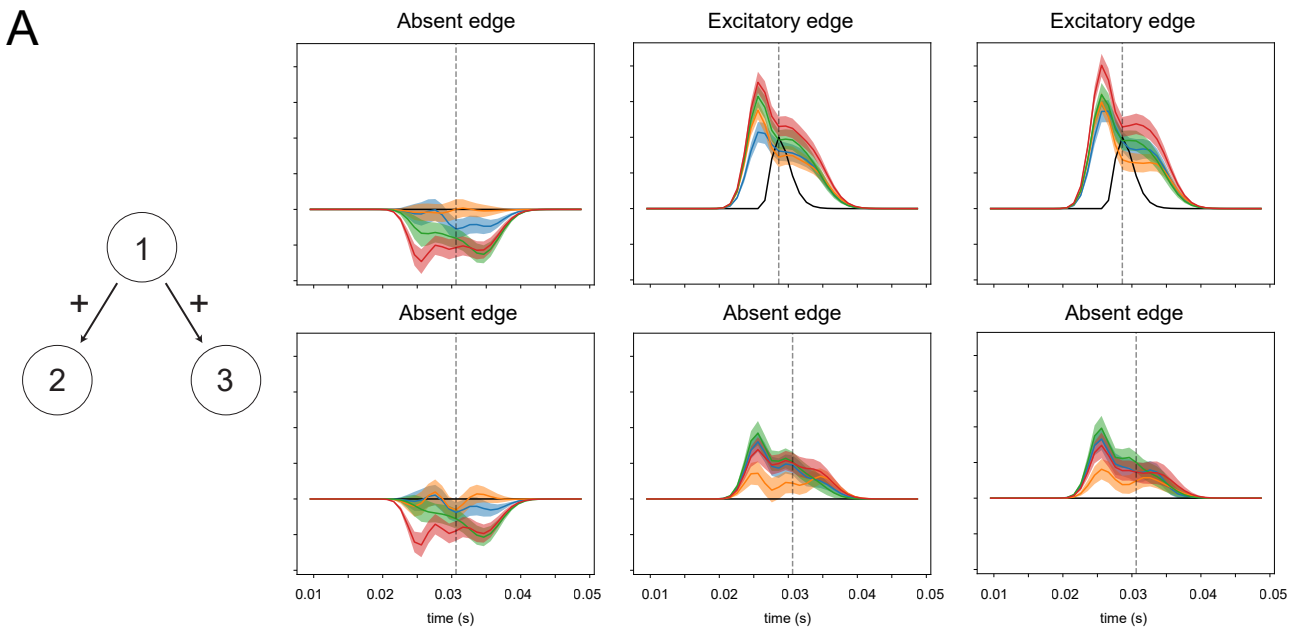
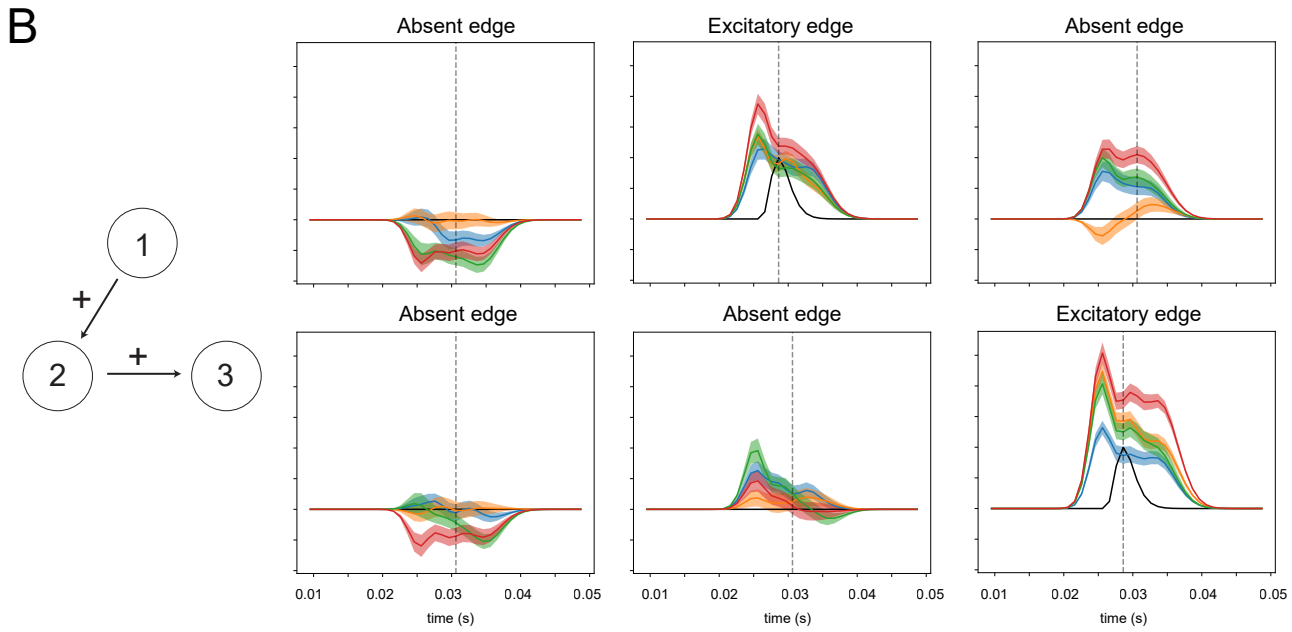
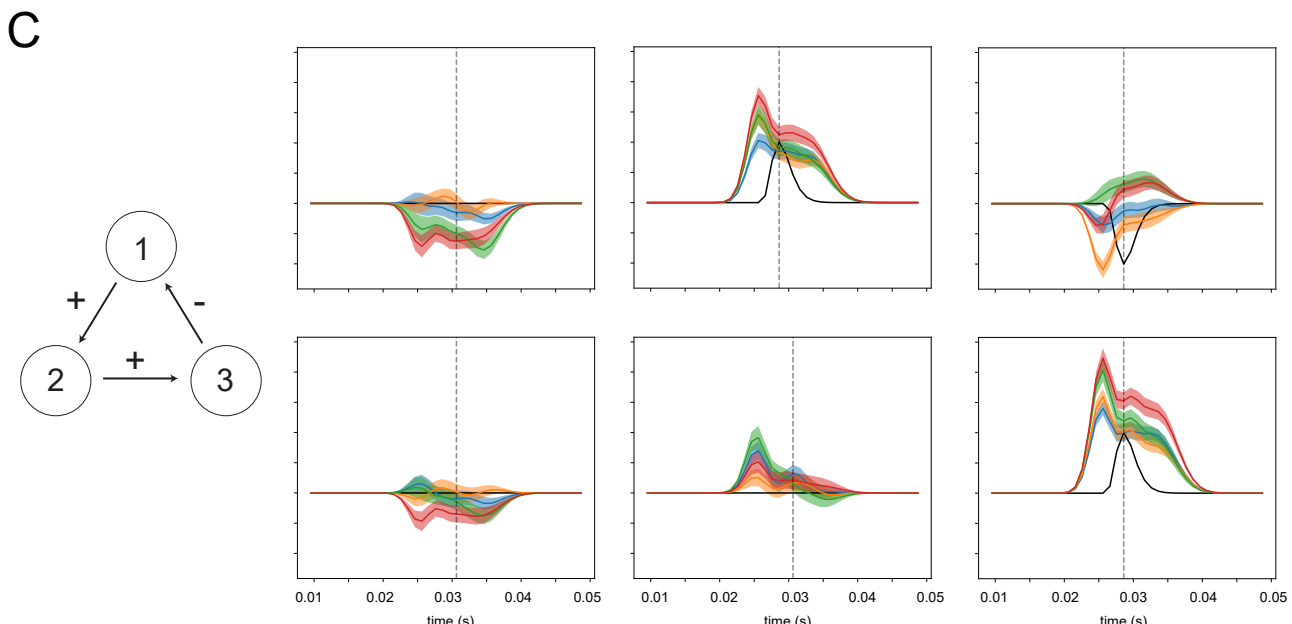


Figure 1: Estimated connectivity for networks 1 and 2.


A
Absent edge
Excitatory edge
Excitatory edge
Absent edge
Absent edge
Absent edge
1
2
3
+
+
0.01
0.02
0.03
0.04
0.05
time (s)


B
Absent edge
Excitatory edge
Absent edge
Absent edge
Absent edge
Excitatory edge
1
2
3
+
+
0.01
0.02
0.03
0.04
0.05
time (s)


C
1
2
3
+
+
-
0.01
0.02
0.03
0.04
0.05
time (s)