

Chapter 13

Neuronal Spike Train Analysis Using Gaussian Process Models

Babak Shahbaba, Sam Behseta, and Alexander Vandenberg-Rodes

Abstract Statistical analysis of simultaneously recorded neurons plays an important role in understanding complex behaviors, decision making process, and neurophysiological disorders. Here, we briefly review several statistical methods specifically developed for analysis of neuronal spike trains. We then focus on application of Gaussian process models for estimating time-varying firing rates of neurons and show how this approach can be extended for modeling synchrony among multiple neurons. We finish this chapter by discussing some possible future directions where more advanced nonparametric Bayesian methods can be utilized to improve existing models.

13.1 Introduction

A common approach in neuroscience involves recording spiking activities or action potentials of neurons using microelectrodes. Subsequently, neuronal data may be represented as the times at which spikes occur. The main objective of a considerably large number of statistical methods then is to model the temporal evolution of the firing patterns of a group of neurons (Brillinger 1988; Brown et al. 2004; Kass et al. 2005; Gerstner and Kistler 2002; Tuckwell 1988; H.C. 1989; Riccardi 1977; Holden 1976; West 2007; Rigat et al. 2006). For a comprehensive review of the topic see Kass et al. (2005).

B. Shahbaba (✉) • A. Vandenberg-Rodes
UC Irvine, CA, USA
e-mail: babaks@uci.edu; vandenbe@uci.edu

S. Behseta
CSUF, Fullerton, CA, USA
e-mail: sbehseta@exchange.fullerton.edu

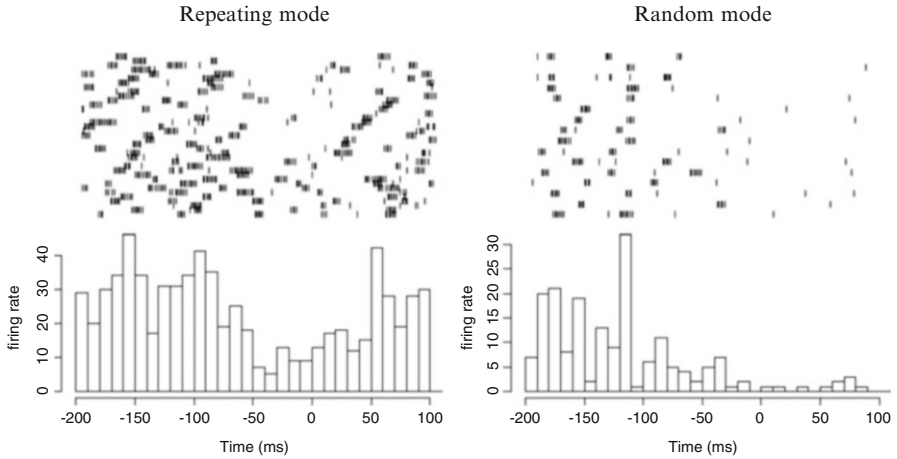


Fig. 13.1 Raster and PSTH plots for a neuron under repeating (*left panel*) and random (*right panel*) modes

As an example, consider a study of neurons recorded from the primary motor cortex (M1) area of a Macaque monkey, performing a sequential task of reaching five targets arranged horizontally on a touch sensitive screen (Matsuzaka et al. 2007). The targets were numbered 1 to 5 from left to right and could be illuminated upon reaching them. The animal was trained to respond to the visual stimuli under two experimental conditions or modes. In the “repeating mode,” a sequence of targets appeared on the screen in a repeating order. In the “random mode,” targets appear in a pseudo-random order. An experimental window of 300 milliseconds (ms) was used. This time window began at 200 ms prior to the target reach and continued for 100 ms after that. The upper segment of Fig. 13.1 shows the corresponding raster plots for a neuron recorded under both modes of this task. Rows in the raster plot represent trials, and the tick-marks are spike time occurrences.

Typically, neuronal data are summarized through peri-stimulus time histograms (PSTH). For the above example, by dividing the window of 300 ms into bins of 10 ms, and pooling the spike occurrences within each bin, one can create the PSTH plots shown in the lower segment of Fig. 13.1.

Let Y_1, \dots, Y_n denote the number of spike occurrences within the bins centered at times t_1, \dots, t_n . A common approach to modeling the neuronal firing rates is by discretizing an inhomogeneous Poisson point process, resulting in a hierarchical model of the form

$$\begin{aligned} Y_j &\sim p(y_j | \theta_j, \zeta) \\ \theta_j &= f(t_j), \end{aligned} \tag{13.1}$$

where the data model $p(y_j | \theta_j, \zeta)$ is usually a $\text{Poisson}(\theta_j)$ density. If the bins are narrow enough they can safely be assumed (or thresholded) to contain at most one spike, so that Y_j can be modeled as a Bernoulli random variable. The model includes a vector of nuisance parameters ζ to allow for generality.

The (latent) firing rates $f(t_j)$ are often the key quantity of interest. In particular—ignoring the details of binning or the data model—one should think of $f(t)$ as a latent *function* over the whole time interval of interest. Thus in the Bayesian approach a key challenge is to produce an appropriately realistic and flexible prior distribution over latent functions, and to then provide computationally efficient procedures for approximating the posterior distribution of f given the observed spike data Y_1, \dots, Y_n .

One highly flexible approach is known as Bayesian adaptive regression splines (BARS) (Dimatteo et al. 2001). In this model the latent function f is assumed to be a spline having knots at unknown locations ξ_1, \dots, ξ_k . Writing $f(t)$ in terms of basis functions $b_{\xi,h}(t)$ as $f(t) = \sum_h b_{\xi,h}(t) \beta_{\xi,h}$, the function evaluations $f(t_1), \dots, f(t_n)$ may be collected into a vector $(f(t_1), \dots, f(t_n))^T = \mathbf{X}_\xi \beta_\xi$, where \mathbf{X}_ξ is the design matrix and β_ξ the coefficient vector. BARS then employs a reversible jump MCMC algorithm (Green 1995) to sample from a suitable approximate posterior distribution on the knot set ξ . Eventually, curves are fitted via model averaging.

One advantage of using BARS for modeling PSTH is the ability to develop inferential methods, suitable for comparing the patterns of spiking activities for comparative problems similar to the one depicted in Fig. 13.1 (Behseta and S. 2011; Behseta et al. 2005). Kottas et al. (2012) and Kottas and Behseta (2010) also treated the problem of comparing the spike trains resulting in the experiments similar to the ones shown in Fig. 13.1, and subsequently developed a fully-Bayesian inferential methodology for such comparative studies; however, they used a Dirichlet process mixture of Beta densities as the prior for f .

Although single-neuron analysis of this type has led to many interesting discoveries, it is widely perceived that complex behaviors are driven by networks of neurons instead of a single neuron (Buzsáki 2010). Therefore, investigators have been recording neuronal activity from multi-probe electrodes. From the statistical point of view, multiple channel recordings greatly facilitate assessing the temporal properties of networks of neurons in real time.

Early analysis of simultaneously recorded neurons focused on correlation of activity across pairs of neurons using cross-correlation analyses (Narayanan and Laubach 2009) and analyses of changes in correlation over time, i.e., by using a Joint Peri-Stimulus Time Histogram or PSTH (Gerstein and Perkel 1969). Similar analyses were performed in the frequency domain by using coherence analysis of neuron pairs using Fourier-transformed neural activity (Brown et al. 2004). For the Bayesian correction for attenuation of correlation in multi-trial spike see Behseta et al. (2009). There are also a number of multivariate analysis techniques for the investigation of simultaneously recorded populations of neurons (Chapin 1999; Nicolelis 1999; Grün et al. 2002; Pillow et al. 2008; Harrison et al. 2013; Brillinger 1988; Brown et al. 2004; Kass et al. 2005; West 2007; Rigat et al. 2006; Patnaik et al. 2008; Diekmann et al. 2009; Sastry and Unnikrishnan 2010; Kottas et al. 2012).

Recently, Kelly and Kass (2012) proposed a new method to quantify synchrony among multiple neurons. The authors argued that separating stimulus effects from history effects would allow for a more precise estimation of the instantaneous

conditional firing rate. Specifically, given the firing history H_t , define $\lambda^A(t|H_t^A)$, $\lambda^B(t|H_t^B)$, and $\lambda^{AB}(t|H_t^{AB})$ to be the conditional firing intensities of neuron A, neuron B, and their synchronous spikes respectively. Independence between the two point processes may be examined by testing the null hypothesis $H_0: \zeta(t) = 1$, where $\zeta(t) = \frac{\lambda^{AB}(t|H_t^{AB})}{\lambda^A(t|H_t^A)\lambda^B(t|H_t^B)}$. where ζ represents the excess firing rate ($\zeta > 1$) or the suppression of firing rate ($\zeta < 1$) due to dependence between two neurons (Ventura et al. 2005; Kelly and Kass 2012). That is, ζ accounts for the excess joint spiking beyond what is explained by independence.

In this chapter we discuss alternative approaches that place a Gaussian process (GP) prior over the latent function in order to model the time-varying and history-dependent firing rate for each neuron. The joint distribution of spikes for multiple neurons is connected to their marginals using a parametric copula model. We first provide a brief overview of univariate GP models in Sect. 13.2. Then, in Sect. 13.3 we discuss the application of GP for single neuron analysis. The copula model for simultaneously recorded neurons is presented in Sect. 13.4. In Sect. 13.5, we discuss some future directions.

13.2 Gaussian Process Models

A Gaussian process (GP) on the real line is a random real-valued function $x(t)$, with statistics determined by its mean function $\mathbb{E}x(s)$ and *kernel* $\kappa(s, t) = \text{Cov}(x(s), x(t))$. More precisely, all finite-dimensional distributions $(x(t_1), \dots, x(t_n))$ are multivariate Gaussian with mean $(\mathbb{E}x(t_1), \dots, \mathbb{E}x(t_n))$, and with covariance matrix $(\kappa(t_k, t_\ell))_{k, \ell=1}^n$. Since the latter must be positive semi-definite for every finite collection of inputs t_1, \dots, t_n , only certain kernels κ are *valid*. Thus when using Gaussian processes, a practitioner often chooses from among the few popular classes of kernels, such as the Squared Exponential (SE), Ornstein–Uhlenbeck (OU), Matérn, Polynomial, and linear combinations of these. For example, we can use the following covariance form, which combines a random constant with the SE kernel and iid observation noise (Rasmussen and Williams 2006; Neal 1998):

$$\begin{aligned} C_{ij} &= \text{Cov}[x(t_i), x(t_j)] \\ &= \lambda^2 + \eta^2 \exp[-\rho^2(t_i - t_j)^2] + \delta_{ij}\sigma_\varepsilon^2. \end{aligned} \quad (13.2)$$

Here, λ , η , ρ , and σ_ε are hyperparameters with their own hyperpriors. In general, the choice of kernel encodes our *qualitative* beliefs about the underlying signal. For instance, samples from a GP with OU kernel are always non-differentiable functions $x(t)$, and the SE kernel generates only infinitely differentiable functions. Despite such differences, both kernels have the inverse *length-scale* ρ as a hyperparameter: smaller values of ρ result in more slowly varying functions. In practice we only observe GPs at a finite number of points, hence local properties of GPs such as

differentiability are irrelevant—in Cunningham et al. (2007), for example, it was observed that using the Matérn instead of SE kernel resulted in negligible differences when modeling spike trains.

It should be remarked that many dynamical models such as autoregressive processes with Gaussian noise are also multivariate Gaussian and hence can be situated within the GP framework, albeit with a usually less interpretable kernel.

13.3 Gaussian Process Model of Firing Rates

With the model (13.1), note that the latent firing rates $f(t_i)$ need to be non-negative, hence a Gaussian process cannot be directly used as a prior distribution for f . In the case of Poisson observations one can use an exponential link function, letting $f(t) = \exp(x(t))$, where $x(t)$ is a GP. In Cunningham et al. (2007) it was instead proposed to set the constant mean function $\mu(t) = \mu > 0$ as an additional hyperparameter for a GP, and then to let the latent rate f be this GP conditioned to be non-negative.¹ In a recent work, Shahbaba et al. (2014) also use the model (13.1) to estimate the underlying firing rate of neurons, but after discretizing time so that there is at most one spike within each time interval, resulting in a binary time series Y_1, \dots, Y_n comprised of 1 s (spike) and 0 s (silence). To model the latent firing probabilities $f(t_i) = P(Y_i = 1)$, they apply the sigmoidal transformation

$$f(t_i) = \frac{1}{1 + \exp[-u(t_i)]},$$

where $u(t)$ has a GP prior. Note that as $u(t)$ increases, so does $f(t_i)$. The prior autocorrelation imposed by this model allows the firing rate to change smoothly over time. When there are R trials (i.e., R spike trains) for each neuron, we can model the corresponding spike trains as conditionally independent given the latent variable $u(t)$. Figure 13.2 shows the posterior expectation of firing rate (blue curve) overlaid on the PSTH plot of a single neuron with 5 ms bin intervals.

13.4 Detecting Synchrony Among Multiple Spike Trains

For multiple neurons, Shahbaba et al. (2014) propose to use a generalization of the method by Kelly and Kass (2012) (see Sect. 13.1) to model the joint distribution as a function of marginals. In general, models that couple the joint distribution of two (or more) variables to their individual marginal distributions are called *copula* models. See Nelsen (1998) for detailed discussion of copula models. Onken et al. (2009) and Berkes et al. (2009) also use copula models for capturing neural dependencies.

¹ Their data model is somewhat different from (13.1), as the spike times are assumed to follow a conditionally inhomogeneous gamma-interval process instead of a Poisson process.

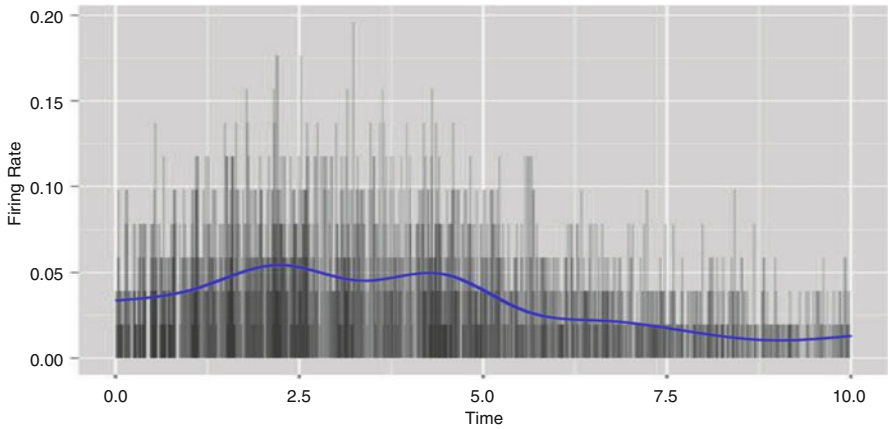


Fig. 13.2 Using the Gaussian process model of Shahbaba et al. (2014) to capture the underlying firing rate of a single neuron from prefrontal cortical areas in rat’s brain. There are 51 spike trains recorded over 10 s. The PSTH plot is generated by creating 5 ms intervals. The curve shows the estimated firing rate (posterior expectation)

Let H be n -dimensional distribution functions with marginals F_1, \dots, F_n . Then, an n -dimensional copula is a function of the following form:

$$H(y_1, \dots, y_n) = \mathcal{H}(F_1(y_1), \dots, F_n(y_n)), \text{ for all } y_1, \dots, y_n.$$

Here, \mathcal{H} defines the dependence structure between the marginals. For example, the Farlie–Gumbel–Morgenstern (FGM) copula family (Farlie 1960; Gumbel 1960; Morgenstern 1956; Nelsen 1998) is defined as follows:

$$\mathcal{H} = \left[1 + \sum_{k=2}^n \sum_{1 \leq j_1 < \dots < j_k \leq n} \beta_{j_1 j_2 \dots j_k} \prod_{l=1}^k (1 - F_{j_l}) \right] \prod_{i=1}^n F_i, \tag{13.3}$$

where $F_i = F_i(y_i)$. As shown by Wilson and Ghahramani (2012), this idea can be generalized to multivariate processes. Restricting the above model to second-order interactions, we have

$$H(y_1, \dots, y_n) = \left[1 + \sum_{1 \leq j_1 < j_2 \leq n} \beta_{j_1 j_2} \prod_{l=1}^2 (1 - F_{j_l}) \right] \prod_{i=1}^n F_i, \tag{13.4}$$

where $F_i = P(Y_i \leq y_i)$. Here, we use y_1, \dots, y_n to denote the firing status of n neurons at time t ; $\beta_{j_1 j_2}$ captures the relationship between the j_1^{th} and j_2^{th} neurons.

For a pair of neurons with firing probabilities p and q respectively, we can show that $\beta = \frac{\zeta - 1}{(1-p)(1-q)}$. As discussed in Sect. 13.1, ζ represents the excess firing rate ($\zeta > 1$) or the suppression of firing rate ($\zeta < 1$) due to dependence between two neurons (Ventura et al. 2005; Kelly and Kass 2012). In our model, $\beta = 0$ indicates that the two neurons are independent; the excess firing rate and the suppression of firing rate between two dependent neurons are represented by $\beta > 0$ and $\beta < 0$ respectively.

To ensure that probability distribution functions remain within $[0, 1]$, the following constraints on all $\binom{n}{2}$ parameters $\beta_{j_1 j_2}$ are imposed:

$$1 + \sum_{1 \leq j_1 < j_2 \leq n} \beta_{j_1 j_2} \prod_{l=1}^2 \varepsilon_{j_l} \geq 0, \quad \varepsilon_1, \dots, \varepsilon_n \in \{-1, 1\}.$$

Considering all possible combinations of ε_{j_1} and ε_{j_2} in the above condition, there are $n(n-1)$ linear inequalities, which can be combined into the following inequality:

$$\sum_{1 \leq j_1 < j_2 \leq n} |\beta_{j_1 j_2}| \leq 1.$$

13.4.1 Computation

Sampling from the posterior distribution of β 's in the above copula model is quite challenging because of the imposed constraints. Lan et al. (2014) developed a novel Markov Chain Monte Carlo algorithm for constrained target distributions of this type based on Hamiltonian Monte Carlo (HMC) (Duane et al. 1987; Neal 2011). They show that in many cases, bounded connected constrained D -dimensional parameter spaces can be bijectively mapped on to the D -dimensional unit ball. Their method then augments the original D -dimensional parameter θ with an extra auxiliary variable θ_{D+1} to form an extended $(D+1)$ -dimensional parameter $\tilde{\theta} = (\theta, \theta_{D+1})$ such that $\|\tilde{\theta}\|_2 = 1$ so $\theta_{D+1} = \pm\sqrt{1 - \|\theta\|_2^2}$. This way, the domain of the target distribution is changed from the unit ball to the D -dimensional sphere. Using the above transformation, they define the Hamiltonian dynamics on the sphere. This way, the resulting HMC sampler can move freely on the sphere, \mathbf{S}^D , while implicitly handling the constraints imposed on the original parameters. As illustrated in Fig. 13.3, the boundary of the constraint, i.e., $\|\theta\|_2 = 1$, corresponds to the equator on the sphere \mathbf{S}^D . Therefore, as the sampler moves on the sphere, passing across the equator from one hemisphere to the other translates to “bouncing back” off the boundary in the original parameter space.

Lan et al. (2014) show that by defining HMC on the sphere, besides handling the constraints implicitly, the computational efficiency of the sampling algorithm could be improved since the resulting dynamics has a partial analytical solution (geodesic flow on the sphere). They used this approach, called Spherical HMC, for sampling from the posterior distribution of β 's in the above copula model and showed that the resulting sampler is substantially more efficient than alternative methods.

13.4.2 Results for Experimental Data

We now consider an experiment designed to investigate the role of the prefrontal cortex in rats in conjunction with reward-seeking behaviors and inhibition of

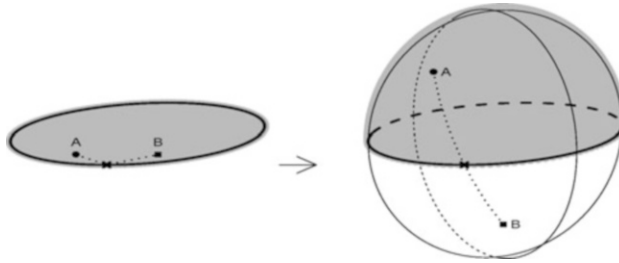


Fig. 13.3 Transforming unit ball $\mathbf{B}_0^D(1)$ to sphere S^D

reward-seeking in the absence of a rewarded outcome. The neural activity (spike trains) of several prefrontal neurons were recorded simultaneously. There are two conditions during the experiment: rewarded and non-rewarded. During the recording/test sessions, two different stimuli were presented: tone 1 (10 KHz) or tone 2 (5 KHz) individually and in pseudorandom order. At the same time, one of two levers was presented: an active-lever, paired with tone 1 (Rewarded-Stimulus—RS) and an inactive-lever paired with tone 2 (Non-rewarded Stimulus—NS). Pressing the active lever resulted in the offset of tone 1, retraction of the lever, and illumination of the reward receptacle. If the rat then went to the reward receptacle, 0.1 ml of 15 % sucrose solution was delivered as a reward. Pressing the inactive lever produced no effect. See Moorman and Aston-Jones (2014) for more details.

Here, we focus on five simultaneously recorded neurons. There are 51 trials per neuron under each scenario. We set the time intervals to 5 ms. Tables 13.1 and 13.2 show the estimates of $\beta_{i,j}$, which capture the association between the i th and j th neurons, under the two scenarios. Figure 13.4 shows the schematic representation of these results under the two experimental conditions. The solid line indicates significant association.

These results show that neurons recorded simultaneously in the same brain area are correlated in some conditions and not others. This strongly supports the hypothesis that population coding among neurons (here though correlated activity) is a meaningful way of signaling differences in the environment (rewarded or non-rewarded stimulus) or behavior (going to press the rewarded lever or not pressing) (Buzsáki 2010). It also shows that neurons in the same brain region are differentially involved in different tasks, an intuitive perspective but one that is neglected by much of behavioral neuroscience. Finally, these results indicate that network correlation is dynamic and that functional pairs—again, even within the same brain area—can appear and disappear depending on the environment or behavior. This suggests (but does not confirm) that correlated activity across separate populations within a single brain region can encode multiple aspects of the task. For example, the pairs that are correlated in reward and not in non-reward could be related to reward-seeking whereas pairs that are correlated in non-reward could be related to response inhibition. Characterizing neural populations within a single brain region

Table 13.1 Estimates of β 's along with their 95 % probability intervals for the first scenario (Rewarded) based on the copula model. Statistically significant values are shown in bold

| β | 2 | 3 | 4 | 5 |
|---------|------------------------|------------------|------------------|------------------------|
| 1 | 0.22(0.07,0.39) | 0.00(-0.07,0.04) | 0.03(-0.02,0.15) | 0.01(-0.04,0.08) |
| 2 | | 0.03(-0.02,0.18) | 0.06(-0.02,0.22) | 0.07(0.00,0.25) |
| 3 | | | 0.08(-0.01,0.26) | 0.21(0.04,0.38) |
| 4 | | | | 0.23(0.09,0.40) |

Table 13.2 Estimates of β 's along with their 95 % probability intervals for the second scenario (Non-rewarded) based on the copula model. Statistically significant values are shown in bold

| β | 2 | 3 | 4 | 5 |
|---------|------------------|------------------------|------------------------|------------------|
| 1 | 0.05(-0.02,0.25) | -0.01(-0.09,0.04) | 0.15(-0.01,0.37) | 0.05(-0.03,0.22) |
| 2 | | 0.21(0.03,0.41) | 0.18(0.00,0.37) | 0.03(-0.02,0.19) |
| 3 | | | 0.17(0.00,0.34) | 0.03(-0.02,0.19) |
| 4 | | | | 0.07(-0.01,0.24) |

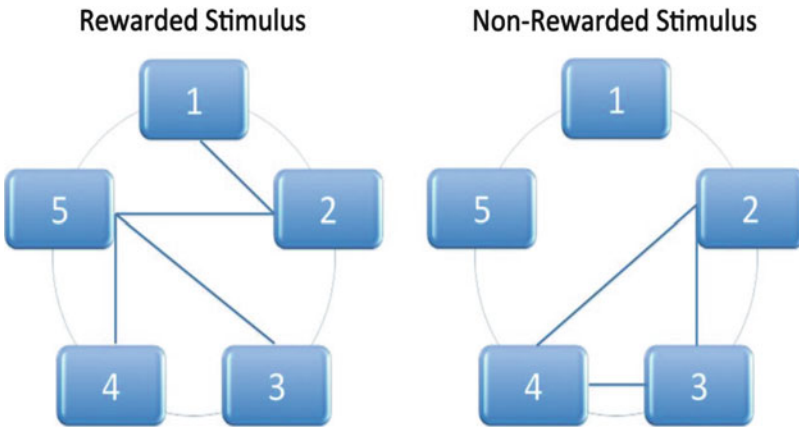


Fig. 13.4 A schematic representation of connections between five neurons under two experimental conditions. The *solid line* indicates significant association

based on task-dependent differences in correlated firing is a less-frequently studied phenomenon compared to the frequently pursued goal of identifying the overall function of the brain region based on individual neural firing (Stokes et al. 2013).

13.5 Future Directions

The methods discussed here can be generalized in several ways as discussed below.

13.5.1 Multivariate GPs

The multivariate model presented in the previous section uses univariate Gaussian processes for the marginal distributions and a copula model for the joint distribution of multiple neurons in terms of these marginals. Alternatively, we can use a *multivariate* GP for modeling the joint distribution of multiple neurons directly. A multivariate Gaussian process can be defined in a similar way as a univariate GP, but this time the kernel function depends on two pairs of inputs. For simplicity we can assume that the mean of each process is the zero function. The kernel κ is now defined for $i, j = 1, \dots, p$ and $s, t \in \mathbb{R}$ as

$$\kappa([i, s], [j, t]) = \mathbb{E}x_i(s)x_j(t). \quad (13.5)$$

The initial challenge within the Gaussian process context is to produce a valid and interpretable kernel. A common technique for generating multivariate GP kernels is known as *co-kriging*, borrowed from the geostatistical literature (Cressie 1993).

One variant of co-kriging describes $(x_1(t), \dots, x_p(t))$ as linear combinations of *latent* factors. We suppose $u_1(t), \dots, u_q(t)$ are independent mean zero Gaussian processes, and let

$$x_i(t) = \sum_{k=1}^q a_{i,k} u_k(t), \quad \text{for } i = 1, 2, \dots, p. \quad (13.6)$$

Let $\kappa_i(s, t) = \mathbb{E}u_i(s)u_i(t)$ be the kernel for the i th latent process. Then the observed processes $\mathbf{x}(t) = (x_1(t), \dots, x_p(t))$ are jointly mean-zero Gaussian with covariances

$$\mathbb{E}x_i(s)x_j(t) = \sum_{k=1}^q a_{i,k} a_{j,k} \kappa_k(s, t). \quad (13.7)$$

This is the *semi-parametric latent factor model* of Teh et al. (2005), so-called because the linear combination of latent GPs is parameterized by the matrix of coefficients $A = (a_{i,k})$, while each Gaussian process is of course a non-parametric model. See Alvarez et al. (2011) for a survey of co-kriging and other multivariate GPs seen in the literature.

Recently, Vandenberg-Rodes and Shahbaba (2015) proposed a multivariate Gaussian processes model for multiple time series $X(t) = (x_1(t), \dots, x_p(t))$ such that each marginal process $x_j(t)$ is a stationary mean-zero Gaussian process with Matérn kernel. Crucially, the marginal processes are not required to share the same hyperparameter values. This approach can be used to model the joint distribution of the firing rates of multiple neurons directly, and allows for significant heterogeneity among neurons while also providing a high degree of interpretability.

13.5.2 *Dynamic Networks*

The static (stationary) model discussed here aggregates cross-neuronal spike-train interactions over time. This can lead to misleading results. Although there exist many dynamic methods developed for modeling brain functional and effective connectivity (Friston et al. 1997; Cribben et al. 2013; Ombao et al. 2005; Ombao and Van Belleghem 2008; Motta and Ombao 2012; Park et al. 2014; Lindquist et al. 2014), these approaches are primarily designed for continuous-valued signals such as functional magnetic resonance imaging (fMRI) and electroencephalogram (EEG) data. The GP-based method discussed here can be extended to model neuronal connections dynamically.

13.5.3 *Community Detection*

Besides allowing for time-varying firing rates and interactions among neurons, the GP-based method can also be extended to cluster neurons based on their cross-dependencies in order to detect subnetworks (communities). To this end, stochastic block models could be used to identify network partitions (Holland et al. 1983). For example, Rodriguez (2012) recently proposed a stochastic block model for network analysis where interactions among factors are observed at multiple time points. This method uses a Bayesian hierarchical stochastic block model to detect possible structural changes in a network. Alternatively, one can use a method similar to the product partition model (PPM) of Müller and Quintana (2010). In general, these methods assume a prior probability on all possible partitions. The assumed prior probability could be influenced by some covariates. This approach can be used to partition neurons into subnetworks.

References

- Alvarez, M. A., Rosasco, L., and Lawrence, N. D. (2011). Kernels for Vector-Valued functions: a review.
- Behseta, S. and Chenouri, S. (2011). Comparison of two population of curves with an application in neuronal data analysis. *Statistics in Medicine*, **30**, 1441–1454.
- Behseta, S., Kass, R. E., and Wallstrom, G. L. (2005). Hierarchical models for assessing variability among functions. *Biometrika*, **92**(2), 419–434.
- Behseta, S., Berdyeva, T., Olson, C. R., and Kass, R. E. (2009). Bayesian correction for attenuation of correlation in Multi-Trial spike count data. *Journal of Neurophysiology*, pages 90727.2008+.

- Berkes, P., Wood, F., and Pillow, J. W. (2009). Characterizing neural dependencies with copula models. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 129–136. Curran Associates, Inc.
- Brillinger, D. R. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cybernetics*, **59**, 189–200.
- Brown, E. N., Kass, R. E., and Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, **7**(5), 456–461.
- Buzsáki, G. (2010). Neural syntax: Cell assemblies, synapse ensembles, and readers. *Neuron*, **68**(3), 362–385.
- Chapin, J. (1999). *Population-level analysis of multi-single neuron recording data: multivariate statistical methods*. In *Methods for Neural Ensemble Recordings*, M. Nicolelis Editor, CRC Press: Boca Raton (FL).
- Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.
- Cribben, I., Wager, T., and Lindquist, M. (2013). Detecting functional connectivity change points for single-subject fmri data. *Frontiers in Computational Neuroscience*, **7**(143).
- Cunningham, J. P., Yu, B. M., Shenoy, K. V., and Sahani, M. (2007). Inferring neural firing rates from spike trains using gaussian processes. In J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis, editors, *NIPS*.
- Diekman, C. O., Sastry, P. S., and Unnikrishnan, K. P. (2009). Statistical significance of sequential firing patterns in multi-neuronal spike trains. *Journal of neuroscience methods*, **182**(2), 279–284.
- Dimatteo, I., Genovese, C. R., and Kass, R. E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika*, **88**(4), 1055–1071.
- Duane, S., Kennedy, A., Pendleton, B. J., and Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, **195**(2), 216–222.
- Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika*, **47**(3/4).
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, **6**(3), 218–229.
- Gerstein, G. L. and Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science*, **164**(3881), 828–830.
- Gerstner, W. and Kistler, W. M. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, 1 edition.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**(4), 711–732.
- Grün, S., Diesmann, M., and Aertsen, A. (2002). Unitary events in multiple single-neuron spiking activity: I. detection and significance. *Neural Computation*, **14**(1), 43–80.
- Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, **55**, 698–707.

- Harrison, M. T., Amarasingham, A., and Kass, R. E. (2013). Statistical identification of synchronous spiking. In P. D. Lorenzo and J. Victor., editors, *Spike Timing: Mechanisms and Function*. Taylor and Francis.
- Tuckwell, H. C. (1989). Philadelphia, Pa. Society for Industrial and Applied Mathematics (SIAM).
- Holden, A. (1976). *Models of the Stochastic Activity of Neurons*. Springer Verlag.
- Holland, P. W., Laskey, K., and Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, **5**(2), 109–137.
- Kass, R. E., Ventura, V., and Brown, E. N. (2005). Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology*, **94**, 8–25.
- Kelly, R. C. and Kass, R. E. (2012). A framework for evaluating pairwise and multiway synchrony among Stimulus-Driven neurons. *Neural Computation*, pages 1–26.
- Kottas, A. and Behseta, S. (2010). Bayesian nonparametric modeling for comparison of single-neuron firing intensities. *Biometrics*, pages 277–286.
- Kottas, A., Behseta, S., Moorman, D. E., Poynor, V., and Olson, C. R. (2012). Bayesian nonparametric analysis of neuronal intensity rates. *Journal of Neuroscience Methods*, **203**(1).
- Lan, S., Zhou, B., and Shahbaba, B. ((2014)). Spherical Hamiltonian Monte Carlo for constrained target distributions. In *Proceedings of the 31th International Conference on Machine Learning (ICML)*.
- Lindquist, M., Xu, Y., Nebel, M., and Caffo, B. (2014). Evaluating dynamic bivariate correlations in resting-state fMRI: A comparison study and a new approach. *Neuroimage*, **101**, 531–546.
- Matsuzaka, Y., Picard, N., and Strick, P. L. (2007). Skill representation in the primary motor cortex after long-term practice. *Journal of neurophysiology*, **97**(2), 1819–1832.
- Moorman, D. and Aston-Jones, G. (2014). Orbitofrontal cortical neurons encode expectation-driven initiation of reward-seeking. *Journal of Neuroscience*, **34**(31), 10234–46.
- Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. *Mitteilungsblatt für Mathematische Statistik*, **8**, 234–235.
- Motta, G. and Ombao, H. (2012). Evolutionary factor analysis of replicated time series. *Biometrics*, **68**, 825–836.
- Müller, P. and Quintana, F. (2010). Random partition models with regression on covariates. *J Stat Plan Inference*, **140**(10), 2801–2808.
- Narayanan, N. S. and Laubach, M. (2009). Methods for studying functional interactions among neuronal populations. *Methods in molecular biology*, **489**, 135–165.
- Neal, R. M. (1998). Regression and classification using Gaussian process priors. *Bayesian Statistics*, **6**, 471–501.
- Neal, R. M. (2011). MCMC using Hamiltonian dynamics. In S. Brooks, A. Gelman, G. Jones, and X. L. Meng, editors, *Handbook of Markov Chain Monte Carlo*, pages 113–162. Chapman and Hall/CRC.
- Nelsen, R. B. (1998). *An Introduction to Copulas (Lecture Notes in Statistics)*. Springer, 1 edition.

- Nicolelis, M. (1999). *Methods for Neural Ensemble Recordings*. CRC Press: Boca Raton (FL).
- Ombao, H. and Van Bellegem, S. (2008). Coherence analysis: A linear filtering point of view. *IEEE Transactions on Signal Processing*, **56**, 2259–2266.
- Ombao, H., von Sachs, R., and Guo, W. (2005). Slex analysis of multivariate non-stationary time series. *Journal of the American Statistical Association*, **100**, 519–531.
- Onken, A., Grünewälder, S., Munk, M., and Obermayer, K. (2009). Modeling short-term noise dependence of spike counts in macaque prefrontal cortex. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1233–1240. Curran Associates, Inc.
- Park, T., Eckley, I., and Ombao, H. (2014). Estimating the time-evolving partial coherence between signals via multivariate locally stationary wavelet processes. *IEEE Transactions on Signal Processing*, **accepted**.
- Patnaik, D., Sastry, P., and Unnikrishnan, K. (2008). Inferring neuronal network connectivity from spike data: A temporal data mining approach. *Scientific Programming*, **16**, 49–77.
- Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., and Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, **454**(7207), 995–9.
- Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press, 2nd edition.
- Riccardi, L. (1977). *Diffusion Processes and Related Topics in Biology*. Springer Verlag.
- Rigat, F., de Gunst, M., and van Pelt, J. (2006). Bayesian modeling and analysis of spatio-temporal neuronal networks. *Bayesian Analysis*, pages 733–764.
- Rodriguez, A. (2012). Modeling the dynamics of social networks using Bayesian hierarchical blockmodels. *Statistical Analysis and Data Mining*, **5**(3), 218–234.
- Sastry, P. S. and Unnikrishnan, K. P. (2010). Conditional probability-based significance tests for sequential patterns in multineuronal spike trains. *Neural Comput.*, **22**(4), 1025–1059.
- Shahbaba, B., Zhou, B., Ombao, H., Moorman, D., and Behseta, S. (2014). A Semi-parametric Bayesian Model for Detecting Synchrony Among Multiple Neurons. *Neural Computation*, **26**(9), 2025–51.
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, **78**(2), 364–375.
- Teh, Y. W., Seeger, M., and Jordan, M. I. (2005). Semiparametric latent factor models. In *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, volume 10.
- Tuckwell, H. (1988). *Introduction to Theoretical Neurobiology: Volume 1, Linear Cable Theory and Dendritic Structure*. Cambridge University Press.
- Vandenberg-Rodes, A. and Shahbaba, B. (2015). Dependent Matérn Processes for Multivariate Time Series.

- Ventura, V., Cai, C., and Kass, R. E. (2005). Statistical assessment of time-varying dependency between two neurons. *J Neurophysiol*, **94**(4), 2940–7.
- West, M. (2007). Hierarchical mixture models in neurological transmission analysis. *Journal of the American Statistical Association*, **92**, 587–606.
- Wilson, A. G. and Ghahramani, Z. (2012). Modelling input dependent correlations between multiple responses. In N. C. P. Flach, T. De Bie, editor, *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD)*, Bristol, UK. Springer.