Bayesian Nonparametric Analysis of Neuronal Intensity Rates

Athanasios Kottas[1], **Sam Behseta**[2], **David E. Moorman**[3], **Valerie Poynor**[4], and **Carl R. Olson**[5]

KEYWORDS: Bayesian nonparametrics, Dependent Dirichlet process prior, Dirichlet process mixture models, Multiple experimental conditions, PSTH, Sensory motor neurons, Supplementary eye field.

[1]Department of Applied Mathematics and Statistics, University of California, Santa Cruz, California 95064, U.S.A., Email: thanos@ams.ucsc.edu

[2]**Corresponding Author**, Department of Mathematics, California State University Fullerton, Fullerton, California 92834, U.S.A. Email: sbehseta@fullerton.edu

[3]Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina 29425, USA., Email: moorman@musc.edu

[4]Department of Applied Mathematics and Statistics, University of California, Santa Cruz, California 95064, U.S.A., Email: vpoynor@ams.ucsc.edu

[5]Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, Email: colson@cnbc.cmu.edu

Abstract

We propose a flexible hierarchical Bayesian nonparametric modeling approach to compare the spiking patterns of neurons recorded under multiple experimental conditions. In particular, we showcase the application of our statistical methodology using neurons recorded from the supplementary eye field region of the brains of two macaque monkeys trained to make delayed eye movements to three different types of targets. The proposed Bayesian methodology can be used to perform either a global analysis, allowing for the construction of posterior comparative intervals over the entire experimental time window, or a pointwise analysis for comparing the spiking patterns locally, in a predetermined portion of the experimental time window. By developing our nonparametric Bayesian model we are able to analyze neuronal data from three or more conditions while avoiding the computational expenses typically associated with more traditional analysis of physiological data.

1 Introduction

The fundamental statistical question in this work, and in a wide range of similar studies involving distinct experimental conditions, revolves around the calibration of the similarities of multiple firing patterns along with the identification of sharp differences between them. This idea is motivated from the observation that the presence or absence of firing activity is considered as the main marker of the degree of involvement of the neuron in the studied behavior. Our interest is in developing a method whereby the overall pattern of activity across multiple conditions can be described which will allow us to better define exactly what role these neurons play in variable sensorimotor contexts. Consequently, it would be useful to devise a suitable statistical methodology to address such comparative inquiries, mainly on two fronts: first, a global analysis over the entire experimental time window, enabling neurophysiologists to decide whether the neuron should be considered for further study; second, a pointwise analysis, to pinpoint differential patterns at specific time points in the experimental time interval. The method proposed in this work is well suited to address these scientific goals.

The need for a comparative study of spiking patterns in multiple conditions may be justified by investigating the neuronal activities presented in Figure 1, where a peri-stimulus time histogram (PSTH) for a single neuron recorded from an awake, behaving monkey is plotted under three experimental conditions described below. A 4000 millisecond window is considered. The time is aligned on a saccadic eye movement. Condition 1 ("Space" condition on the top panel) reflects a trimodal firing pattern: a peak in firing activity in about 1000 millisecond prior to the saccade, followed by a significantly less-pronounced peak at the saccade time, yielding to yet another strong peak at about 1500 milliseconds after the saccade time. The response in condition 2 ("Dot" condition on the middle panel) is inherently different: A series of comparably weaker bursts of activity, more like a random noise, nonetheless with a seemingly noticeable decline in firing activity at the saccade time. Finally, in condition 3 ("Ring" condition on the lower panel) a multimodal pattern is suggested with the most notice-

able peak at the saccade time. We now can restate the main objectives of this work in the context of the data presented in this figure: 1- Are the differences and similarities in the spiking patterns of the three conditions in Figure 1 statistically significant or should they be interpreted as perturbations due to chance and hence be ignored? 2- When studied in predetermined slices of the entire time segment, could such differences shed light on the sensorimotor properties? These are among the questions that motivate the statistical strategies we adopt in order to compare the firing patterns of each neuron in multiple conditions. Here, we address the above two questions by developing a Bayesian nonparametric model that allows us to compare neuronal intensity rates under a number of distinct experimental conditions within a coherent probabilistic framework for inference.

Insert Figure 1 Here

The outline of the paper is as follows. Section 2 develops the methodology with technical details on implementation included in an Appendix. In Section 3, we provide more details for the experiment used to illustrate the proposed methodology, and in Section 4, we present the results from the analysis of the corresponding data. Finally, Section 5 concludes with an overview and discussion.

2 Models

In Section 2.1, we discuss the stochastic model underlying our approach and provide a brief review of the class of models from the field of Bayesian nonparametrics that provides the foundation for the proposed methodology. Section 2.2 develops the modeling approach for comparison of neuronal spiking patterns under multiple experimental conditions.

2.1 Motivation and Background

2.1.1 Poisson process modeling for neuronal firing intensities

Stochastic modeling and statistical estimation techniques for the analysis of data from single-recording neurophysiological experiments have received considerable attention in the neuroscience, as well as the statistics literature (see, e.g., Brillinger, 1992; Ventura et al., 2002; Kass et al., 2005). Predominantly, the focus of statistical modeling approaches is on the temporal evolution of the neuronal firing activity. Historically, the stochastic modeling of spike trains may be traced back to the original forms of the so-called Integrate and Fire (IF) models (Gerstein and Mandelbrot, 1964; Stein, 1965). In these stochastic models, the output is taken as a one dimensional voltage while the inputs consist of current and membrane conductance. The error is typically captured by a Brownian motion, representing the stochastic feature of the model. As noted in Paninski et al. (2010), from the statistical point of view, the IF model may also be studied via hidden Markov (or state space) models in which the unobserved (hidden) voltage is modeled through a Markovian process evaluated at the observed spiking times (Volgestein and Paninski, 2007; Brown et al., 1998).

An alternative approach, leading eventually to the methodology proposed in this paper, is to view the spike train as a realization of a point process, a random sequence of times associated with spike occurrences, and subsequently model the spike counts with a time-varying intensity function formulated through a Non-Homogeneous Poisson Process (NHPP) as described below. Reviews of the analysis of neuronal data using point processes, from either a single neuron or from multiple neurons, can be found in, e.g., Brillinger (1992), Brown et al. (2004), Kass et al. (2005).

Let $N_{(t_a,t_b)}$ denote the number of spike occurrences in the time interval (t_a, t_b). By definition, a NHPP point process model is constructed over two conditions:

(a) For any interval (t_a, t_b), $N_{(t_a,t_b)}$ follows a Poisson distribution with mean $\int_{t_a}^{t_b} \lambda(u)du$. Here, $\lambda(\cdot)$ is the NHPP intensity function, a non-negative and locally integrable

function (i.e., $\int_D \lambda(u)du < \infty$ for any bounded subset D of the positive real line).

(b) For any non-overlapping intervals, (t_a, t_b), and (t_c, t_d), the random variables $N_{(t_a,t_b)}$, and $N_{(t_c,t_d)}$ are independent.

Subsequently, if a generic set of n spike times, $\{s_1, \cdots, s_n\}$, observed in time window (A, B), is assumed to arise from a NHPP, the corresponding likelihood for the intensity function is given by $e^{-\int_A^B \lambda(u)du} \prod_{i=1}^n \lambda(s_i)$.

Statistically, the problem of modeling the spike trains will then revolve around estimating the NHPP intensity function $\lambda(\cdot)$ from which inference on several features of the neuronal spiking pattern can be obtained. In this paper, we adopt a Bayesian nonparametric point of view for such an estimation problem developing a practically important methodological extension of the approach proposed in Kottas and Behseta (2010). This approach and the relevant class of nonparametric Bayesian prior models are reviewed in the following section.

2.1.2 Background on Bayesian nonparametric mixture models

A nonparametric modeling approach for the NHPP intensity treats the entire function $\lambda(\cdot)$ as the unknown parameter, which under the Bayesian paradigm, necessitates placing a prior over a space of functions (i.e., over an infinite dimensional parameter). The field of Bayesian nonparametrics deals with the problem of placing prior probability models on spaces of distributions (or, in general, functions), including theory that studies the definition and properties of such priors (e.g., Ghosh and Ramamoorthi, 2003), and methods for modeling and posterior inference with nonparametric priors (see, e.g., Walker et al., 1999; Müller and Quintana, 2004; Hanson et al., 2005, for related reviews). Nonparametric Bayesian methods combine the flexibility of data-driven modeling, which is not restricted by specific parametric forms, with the inferential power of the Bayesian statistical framework. In particular, in the context of modeling neuronal firing intensities, a flexible Bayesian nonparametric model can uncover non-standard

neuronal spiking patterns and at the same time provide full uncertainty quantification for the intensity estimates, which does not rely on large-sample asymptotic results that may be suspect for spike trains of small to moderate dimension.

The modeling approach of Section 2.2 utilizes Dirichlet process (DP) mixtures, a widely applicable class of Bayesian nonparametric mixture models. The DP prior (Ferguson, 1973) was the first nonparametric probability model developed for spaces of random distributions (equivalently, distribution functions); informally, it can be viewed as a stochastic process with realizations that are distributions. The DP prior model, denoted by $\mathrm{DP}(\alpha, G_0)$, is characterized by two parameters: α a positive scalar parameter, and $G_0 \equiv G_0(\varphi)$, a distribution specified up to a number of parameters φ. In particular, G_0 is the center (or base distribution) of the process whereas α can be interpreted as a precision parameter; the larger α is, the closer a DP realization is to G_0. The most useful definition of the DP is its constructive definition (Sethuraman, 1994), according to which a random distribution G generated from a $\mathrm{DP}(\alpha, G_0)$ prior is (almost surely) given by

$$G(\cdot) = \sum_{j=1}^{\infty} \omega_j \delta_{\theta_j}(\cdot). \tag{1}$$

Here, $\delta_a(\cdot)$ denotes a point mass at a, the θ_j are i.i.d. from G_0, and the ω_j are defined through a *stick-breaking* procedure. Specifically, $\omega_1 = \zeta_1$, $\omega_j = \zeta_j \prod_{s=1}^{j-1}(1-\zeta_s)$, for $j \geq 2$, where the ζ_s are i.i.d. from a $\mathrm{Beta}(1, \alpha)$ distribution; moreover, $\{\zeta_s,\ s = 1,2,...\}$ and $\{\theta_j,\ j = 1,2,...\}$ are independent sequences of random variables. Hence, the DP yields discrete distributions with a countable number of possible values drawn from G_0 and associated probabilities ω_j generated as described above. Note that the DP parameter α controls the effective number of probabilities ω_j, in particular, small α values favor DP realizations that are discrete distributions with most of their probability mass placed on a small number of values. Moreover, DP realizations can be readily approximated to any desired level of accuracy by truncation: $G(\cdot) \approx \sum_{j=1}^{J} q_j \delta_{\theta_j}(\cdot)$, where the finite number of q_j are normalized to sum up to 1; e.g., the stick-breaking construction can be used for the first $J-1$ of the q_j with the last one set equal to $1 - \sum_{j=1}^{J-1} q_j$.

The discreteness of DP realizations is an asset when the DP prior is used for later stages of a hierarchical model, as in DP mixture models,

$$f(\cdot; G) = \int k(\cdot; \theta) dG(\theta), \quad G \sim \mathrm{DP}(\alpha, G_0),$$

where $k(\cdot; \theta)$ is the density for a parametric family of distributions indexed by θ. DP mixtures increase considerably the scope of mixture modeling, since they replace parametric specifications for the mixing distribution G with a general nonparametric prior, which supports the entire space of possible mixing distributions. Note that $f(\cdot; G)$ is a random density function since G is random. Moreover, using the constructive definition of the DP, we obtain $f(\cdot; G) = \sum_{j=1}^{\infty} \omega_j k(\cdot; \theta_j)$, i.e., a representation for $f(\cdot; G)$ as a countable mixture with the ω_j and θ_j arising from the DP prior as discussed above. Of course, this is the prior probability model for the random density; in practice, given a sample of size n, the number of distinct components $n^* < n$ is driven by the data, making the class of DP mixture models appealing for applications where clustering is anticipated as in, e.g., density estimation, classification, and regression.

Kottas and Behseta (2010) proposed a DP mixture modeling approach for neuronal firing intensities under the NHPP stochastic model setting discussed in Section 2.1.1. Key to the approach is an equivalent representation for the NHPP intensity in terms of a density function over a bounded interval (taken without loss of generality to be $(0, 1)$), and a scale parameter (details are provided in Section 2.2). In Kottas and Behseta (2010), a DP mixture of Beta densities is utilized for the NHPP density to obtain flexible inference for neuronal firing intensities, and the methodology is applied to neurons recorded from the primary motor cortex area of a monkey's brain while performing a sequence of reaching tasks under two distinct experimental conditions. A limitation of the approach is that the mixture model was applied separately to the neuronal data from each condition, that is, no *borrowing of strength* across different conditions was allowed in the prior model. This aspect of the approach becomes more restrictive as the number of experimental conditions increases, which will typically result in unbalanced sample sizes for the condition-specific firing times. In Section

2.2, we develop a methodologically novel and practically important extension of the approach in Kottas and Behseta (2010), using a dependent DP mixture model, which can be applied to neuronal firing times from any number of experimental conditions. The hierarchical model structure retains the flexibility of the DP mixture setting for each condition-specific neuronal firing intensity, and at the same time, enables borrowing of strength in inference for the collection of firing intensities across the different conditions. The key practical implication is with regard to uncertainty quantification in the estimation of the neuronal firing intensities and of pairwise differences among neuronal firing intensities for all conditions of interest.

2.2 The Nonparametric Bayesian Modeling Approach

In Section 2.2.1, we present the general framework for modeling collections of intensity functions of neuronal firing times corresponding to a number of distinct experimental conditions. In Section 2.2.2, this framework is utilized to build inference for contrasts over pairs of condition-specific neuronal intensity functions.

2.2.1 A dependent nonparametric mixture model for collections of NHPP intensity functions

Focusing on the data from a particular condition (which will be indexed by $i = 1, ..., I$), the observed neuron spiking times comprise the data vector $\{t_{i,km} : k = 1, ..., K;\ m = 1, ..., M_{ki}\}$, where $t_{i,km}$ is the m-th firing time in the k-th trial under condition i. Note that for the experiment considered here the number of trials is the same for all conditions. To develop the statistical tools needed for comparing the condition-specific firing intensities, it suffices to consider the firing times aggregated over all trials from each condition. Consequently, within each condition, $i = 1, ..., I$, there will be $n_i = \sum_{k=1}^{K} M_{ki}$ firing times from all trials.

The proposed dependent model for the collection of condition-specific intensity functions is applied to vectors of responses (firing times) across conditions. These vectors

of response replicates arise naturally under the experiment through $(t_{i,km}, ..., t_{I,km})$, that is, by grouping firing times from the same trial across conditions. We denote the response vectors by $\boldsymbol{t}_j = \{t_{ij} : i = 1, ..., I\}$, for $j = 1, ..., n$, where $n = \sum_{k=1}^{K} \max_i M_{ki}$. Each vector $\boldsymbol{t}_j$ can range in dimension from 1 to I. To account for this in the notation, we use missing value indicators s_{ij}, where $s_{ij} = 1$, if there is a j-th response for condition i, and $s_{ij} = 0$, otherwise. Note that the s_{ij} are fully determined by the data, and $\sum_{j=1}^{n} s_{ij} = n_i$. Finally, we assume, without loss of generality, that the point patterns of firing times are observed in the unit time interval; inference over the original time interval can be readily obtained through transformation.

Our approach is based on a NHPP stochastic model for the underlying point process generating the firing times under each condition. As has been documented in the relevant literature (see, e.g., the discussion in Ventura et al., 2002), the NHPP provides a plausible model for the aggregated firing times based on both empirical evidence as well as theoretical results, which yield that pooled point patterns across a large number of replicated trials follow approximately a NHPP model.

Under the NHPP assumption, and using conditional independence given the intensity functions, $\lambda_i(t)$, for $i = 1, ..., I$, the first-stage specification (likelihood) of the Bayesian model for the neuronal data from all conditions is given by

$$\prod_{i=1}^{I} \exp\{-\textstyle\int_0^1 \lambda_i(u)du\} \prod_{j=1}^{n} \{\lambda_i(t_{ij})\}^{s_{ij}}. \tag{2}$$

The full Bayesian model builds dependence across the different experimental conditions through a dependent nonparametric prior for $\{\lambda_i(t) : i = 1, ..., I\}$, the collection of neuronal intensities.

The modeling framework builds on a flexible nonparametric prior for each NHPP intensity $\lambda_i(t)$, which can provide full inference for the neuronal firing intensity under the corresponding condition without relying on specific parametric forms or asymptotic arguments. Moreover, the prior model allows for dependence between the different intensities, which enables borrowing of strength in the posterior inferences given the data from all conditions. In particular, we use a DP mixture formulation for a density

function directly connected with the intensity function, and relate the DP mixture models for the condition-specific densities through a dependent DP prior developed by extending the DP constructive definition in (1). By casting the modeling in a density estimation framework, we can use a flexible class of nonparametric mixture models that allows relatively easy prior specification and Markov chain Monte Carlo (MCMC) simulation for posterior inference. The approach was originally developed by Kottas (2006) and Kottas and Sansó (2007) in the context of modeling a single NHPP intensity defined over a time interval or a spatial region, respectively, and was extended in Taddy and Kottas (2009) to model marked NHPPs. As discussed in Section 2.1.2, Kottas and Behseta (2010) adapted the methodology in the context of neuronal data analysis focusing on inference for global and local comparison of neuronal intensities under two conditions. The modeling approach developed here offers a methodological advance on modeling and inference for collections of related NHPP intensity functions.

The starting point for the modeling approach is an equivalent representation for $\lambda_i(t)$, $t \in (0,1)$, through the density function $f_i(t) = \lambda_i(t)/\gamma_i$, $t \in (0,1)$, where $\gamma_i = \int_0^1 \lambda_i(u)du$. The density function $f_i(\cdot)$ (which may be referred to as the NHPP density) fully controls the shape of the intensity function, since the parameter γ_i provides only the scale for $\lambda_i(\cdot)$. Hence a flexible nonparametric prior model for $f_i(\cdot)$ can capture non-standard intensity shapes. Note that the NHPP density has bounded support given by the unit interval. The model used in Kottas (2006) and Kottas and Behseta (2010) was based on a DP mixture of Beta densities for the NHPP density. Here, we work instead with a logit-normal density for the kernel of the DP mixture, that is,

$$k(t;\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2} t^{-1}(1-t)^{-1} \exp\{-[\log(t/(1-t)) - \mu]^2/2\sigma^2\}, \quad t \in (0,1).$$

Note that this density arises through the logistic transformation, $t = \exp(y)/(1+\exp(y))$, of a $\mathrm{N}(\mu,\sigma^2)$ density for y, and this provides a more convenient platform for MCMC posterior inference than the Beta kernel, which is particularly important for the model extension we are considering here. Hence, the model for the NHPP density

$f_i(\cdot)$, that defines the neuronal intensity under condition i, is given by

$$f_i(t) \equiv f(t; G_i, \sigma_i^2) = \int k(t; \mu, \sigma_i^2) dG_i(\mu), \quad t \in (0,1), \tag{3}$$

that is, a location mixture of logit-normal densities with nonparametric random mixing distribution G_i. The mixture specification in (3) strikes a good balance between model flexibility and computational feasibility. Mixing over the location parameter of the logit-normal kernel can result in skewness or multimodality for $f(t; G_i, \sigma_i^2)$ when such features are suggested by the data; this may be at the expense of a larger number of mixture components than what would be needed under the more general model that includes mixing also with respect to the scale parameter of the logit-normal kernel. However, this more general location-scale mixture requires a more complex prior model for the G_i and more complicated methods for posterior simulation.

We build dependence in the prior model for the neuronal densities $\{f_i(\cdot) : i = 1, ..., I\}$ through a dependent nonparametric prior for the collection of corresponding mixing distributions $\boldsymbol{G} = \{G_i : i = 1, ..., I\}$. Specifically, we use a dependent DP prior (MacEachern, 2000) motivated by the DP constructive definition in (1), which is extended from a nonparametric prior model for a single distribution to one for the collection $\boldsymbol{G}$. A powerful, and at the same time practical, way to accomplish this is to retain the probabilities ω_j exactly as defined in (1) through stick-breaking, but replace the locations θ_j with vectors $\boldsymbol{\theta}_j = (\theta_{j1}, ..., \theta_{jI})$. Therefore, the univariate centering distribution G_0 needs to be replaced by a multivariate distribution $\boldsymbol{G}_0$ of dimension I. In our context, a multivariate normal distribution provides a natural choice for $\boldsymbol{G}_0$; in particular, we take $\boldsymbol{G}_0 = \mathrm{N}_I(\lambda \mathbf{1}_I, \Lambda)$, a normal distribution of dimension I with common mean λ ($\mathbf{1}_I$ denotes an I-dimensional vector with all its elements equal to 1) and covariance matrix Λ. Hence, the dependent DP prior, denoted by $\mathrm{DDP}(\alpha, \boldsymbol{G}_0)$, yields (almost sure) realizations of the form

$$\boldsymbol{G} = \sum_{j=1}^{\infty} \omega_j \delta_{\boldsymbol{\theta}_j}, \tag{4}$$

where the $\boldsymbol{\theta}_j$ are i.i.d. from $\boldsymbol{G}_0$. Dependence across $i = 1, ..., I$ in the locations $\boldsymbol{\theta}_j = (\theta_{j1}, ..., \theta_{jI})$ implies dependent G_i, $i = 1, ..., I$, in the prior model. At the same time, for any $i = 1, ..., I$, the mixing distribution G_i admits the representation $G_i(\cdot) = \sum_{j=1}^{\infty} \omega_j \delta_{\theta_{ji}}(\cdot)$, with the θ_{ji} i.i.d. from $G_{0i} = \mathrm{N}(\lambda, \Lambda_{ii})$, where the variance Λ_{ii} is defined by the i-th diagonal element of Λ. Hence, marginally, each G_i has a $\mathrm{DP}(\alpha, G_{0i})$ prior, thus retaining the flexibility and useful interpretation for the nonparametric mixture in (3). This dependent DP prior construction – where the dependence is built through the locations with the corresponding probabilities defined as in the regular DP prior – is referred to as the "single-p" DDP prior. For further details on general properties of single-p DDP models and a range of applications we refer to MacEachern (2000), De Iorio et al. (2004), Gelfand et al. (2005), Kottas et al. (2008), Rodriguez and ter Horst (2008), Kottas and Krnjajić (2009), and Fronczyk and Kottas (2010).

To complete the dependent DP prior formulation and obtain the full Bayesian model, we place priors on the scale parameters γ_i for the NHPP intensities, on the scale parameters σ_i^2 of the mixture kernels for the NHPP densities, and on the DDP prior hyperparameters. In particular, key are hyperpriors for λ and Λ, since these parameters define $\boldsymbol{G}_0$, where the covariance matrix in particular specifies the dependence of the centering distribution for the DDP prior. We assign a normal prior to λ and an inverse-Wishart prior to Λ. Moreover, we place a gamma prior distribution on α. Regarding parameters γ_i, a convenient form for their priors is given by a gamma distribution with parameters that can be specified based on the role γ_i plays as the mean of the NHPP over the observation time interval. Alternatively, a more automatic choice arises by using the reference prior for each γ_i, which can be shown to be given by $p(\gamma_i) \propto \gamma_i^{-1}$, for $\gamma_i > 0$ (Kottas and Behseta, 2010). Finally, we place a hierarchical inverse-gamma prior distribution on the σ_i^2. More details on the form of the priors above as well as on specification of their parameters are provided in the Appendix.

2.2.2 Inference for comparison of multiple experimental conditions

Given neuronal firing times arising under I distinct experimental conditions, we employ the Bayesian nonparametric approach of Section 2.2.1 to model the collection of corresponding neuronal intensity functions through $\lambda_i(t) \equiv \gamma_i f(t; G_i, \sigma_i^2)$, $t \in (0,1)$, for $i = 1, ..., I$, where $f(t; G_i, \sigma_i^2)$ is the logit-normal mixture in (3), and the collection of mixing distributions $\boldsymbol{G} = \{G_i : i = 1, ..., I\}$ is assigned the dependent DP prior developed in the previous section.

The full Bayesian model involves the collection of mixing distributions $\boldsymbol{G}$, the DDP hyperparameters $(\alpha, \lambda, \Lambda)$, the vector of NHPP mean parameters, $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_I)$, and the vector of mixture kernel scale parameters, $\boldsymbol{\sigma}^2 = (\sigma_1^2, ..., \sigma_I^2)$. Samples from the joint posterior distribution, $p(\boldsymbol{\gamma}, \boldsymbol{\sigma}^2, \boldsymbol{G}, \alpha, \lambda, \Lambda \mid \text{data})$, where data $= \{\boldsymbol{t}_j : j = 1, ..., n\}$, yield full inference for each neuronal firing intensity, $\gamma_i f(t_0; G_i, \sigma_i^2)$, at any collection of points t_0 in the observation time window. Based on the form of the NHPP likelihood, it can be shown that the joint posterior distribution factorizes into the marginal posterior for $\boldsymbol{\gamma}$ and the posterior distribution for all other model parameters, $p(\boldsymbol{\sigma}^2, \boldsymbol{G}, \alpha, \lambda, \Lambda \mid \text{data})$. This is evident from the expression for the likelihood in (2) which can be rewritten as $\prod_{i=1}^{I} \exp(-\gamma_i)\gamma_i^{n_i} \prod_{j=1}^{n} \prod_{i=1}^{I} \{f(t_{ij}; G_i, \sigma_i^2)\}^{s_{ij}}$. In particular, under a gamma prior or the reference prior for each γ_i, the posterior distribution for $\boldsymbol{\gamma}$ is available analytically as $p(\boldsymbol{\gamma} \mid \text{data}) = \prod_{i=1}^{I} p(\gamma_i \mid \text{data})$, where $p(\gamma_i \mid \text{data})$ is given by a gamma distribution, for instance, a gamma$(n_i, 1)$ distribution under the reference prior $p(\gamma_i) \propto \gamma_i^{-1}$. The technical details on MCMC posterior simulation from $p(\boldsymbol{\sigma}^2, \boldsymbol{G}, \alpha, \lambda, \Lambda \mid \text{data})$ can be found in the Appendix.

In addition to inference for each condition-specific neuronal firing intensity, our interest is in comparison of all pairs of firing intensities $\lambda_i(\cdot)$ and $\lambda_{i'}(\cdot)$, for $i, i' = 1, ..., I$ with $i \neq i'$. Note the key advantage of working with the NHPP density functions as opposed to the intensity functions. Because firing activities on each pair of conditions are expressed in terms of their densities, they are scale-free and hence can be contrasted directly either locally at a small number of time points of interest, or globally over the

observation time interval. Therefore, for any pair (i, i') of distinct conditions, posterior realizations for the neuronal firing densities $f_i(\cdot)$ and $f_{i'}(\cdot)$ can be used to launch both point as well as interval estimation for comparing the respective firing intensities $\lambda_i(\cdot)$ and $\lambda_{i'}(\cdot)$. For example, we may construct point estimates and associated posterior uncertainty bands for the function $f_i(\cdot) - f_{i'}(\cdot)$ over the entire experimental window, as well as obtain the posterior distribution for $f_i(t_0) - f_{i'}(t_0)$ for specific points t_0 in the experimental time interval.

3 Experimental design and methods

In primates, the supplementary eye field (SEF) is a brain region involved in sensorimotor and cognitive processing. Neurons in this region exhibit changes in firing rate in relation to a number of sensory (e.g., upon the presentation of visual stimuli), cognitive (e.g., during the maintenance of spatial memories), and behavioral (e.g., during the generation of eye movements) factors. Intracortical microstimulation at moderate levels of current (40-50 A) elicits eye movements (Fujii et al. 1995; Lee and Tehovnik 1995; Mann et al. 1988; Martinez-Trujillo et al. 2003a; Martinez-Trujillo et al. 2004; Martinez-Trujillo et al. 2003b; Missal and Heinen 2001, 2004; Mitz and Godschalk 1989; Russo and Bruce 1993; Schall 1991a; Schlag and Schlag-Rey 1985, 1987a, b; Tehovnik and Lee 1993; Tehovnik et al. 1994; Tehovnik and Slocum 2000; Tehovnik et al. 1999; Tehovnik et al. 1998; Tehovnik and Sommer 1996, 1997; Tian and Lynch 1995). Single neurons in the area are active during the planning and execution of eye movements toward targets in restricted response fields (Bon and Lucchetti 1991, 1992; Chen and Wise 1995a, b, 1996, 1997; Coe et al. 2002; Fujii et al. 2002; Hanes et al. 1995; Lee and Tehovnik 1995; Moorman and Olson 2007; Mushiake et al. 1996; Olson and Gettner 1995, 1999, 2002; Olson et al. 2000; Olson and Tremblay 2000; Russo and Bruce 1996, 2000; Schall 1991a, b; Schlag-Rey et al. 1997; Schlag and Schlag-Rey 1985, 1987b; Schlag et al. 1992; Tremblay et al. 2002). Targets, however, may be

represented in multiple ways. For example, saccade targets may be physically present as in the case of a dot on a screen, or they may be internally-represented, as in the case of a remembered location in empty space. The question of how neurons in the SEF respond during eye movements planned to internally – vs. externally – represented targets has not been investigated until now. In the results described here, we recorded from neurons of the SEF of two monkeys during performance of memory-guided saccades to one of three target-types (see below). These results are a subset from a larger study (Moorman, Medler, and Olson, in preparation). In that study there were four conditions in each target type: targets were located at one of four cardinal locations on each condition resulting in 12 conditions to analyze. The task and targets are shown in Figure 2. Here, however, we focus on three conditions per recording: one direction per condition, matched across condition (so, for example, dots, ring, and space targets located at one position to the left of the screen center). We chose to focus specifically on variations in target type so as to directly address the issue of whether changes in firing pattern related to the targets could be statistically described. This type of classification would be particularly interesting given that, in the conditions studied, manipulations are categorical, not parametric (as might be seen in changes of target size or brightness). In the latter case, direct relationships between neuronal activity and target dimensions might be more easily observed. The goal of the analysis of the present data, therefore, was to develop a means for characterizing the activity of single neurons under categorically different conditions. Two adult male rhesus monkeys were used. Experimental procedures were approved by the Carnegie Mellon University Animal Care and Use Committee and were in compliance with the guidelines set forth in the United States Public Health Service Guide for the Care and Use of Laboratory Animals. The aim of this task was to require monkeys to make memory guided saccades under conditions in which the possible target locations were either marked with visible detail or unmarked. There were four possible target locations surrounding the central fixation point at 90 intervals although, for the purposes of the present analysis,

we focus on one direction per recording (direction was determined post-hoc based on the directional selectivity of the recorded neuron). Marking could take any of three forms. Space: the targets were unmarked. Dot: the targets were marked by individual dots. Ring: an annulus centered on fixation intersected all four target locations. The timing of events is summarized in Figure 2. Each trial began with presentation of a central fixation spot and attainment of fixation. After 400 ms, white marks appeared in the dot and ring conditions whereas the peripheral screen remained blank in the space condition. After an additional 500 ms, a green cue spot was presented at one of the possible target locations for 100 ms. Following a random delay of 800 to 900 ms, the fixation spot was extinguished and the monkey was required to make a saccade to the previously cued location. Reward was delivered only if the monkey made a saccade directly to the cued location within 150 ms of breaking fixation and then fixated on the cued location for 200-300 ms. The cue stimulus reappeared as a visual feedback target during reward presentation. The twelve conditions were presented in pseudo-random interleaved sequence until, typically, 16 trials had been completed successfully under each condition. Recordings from single neurons in the SEF were conducted as described previously (Moorman and Olson, 2007a,b). The data for our analysis consist of the counts, accumulated across trials, beginning 2000 ms before the onset of the spatial cue and extending 2000 ms after.

Insert Figure 2 Here

4 Results

Results are presented from an application of the analysis to three neurons recorded from the SEF of one monkey during different recording sessions and serve as examples of how firing-rate data analysis might benefit from its use. Figure 1 shows the firing rate histograms of the activity of a single SEF neuron (sp259a.1) recorded in three conditions: 1) saccades directed to an unmarked location (Space, top), 2) saccades

directed to one of four marked targets (Dot, middle), and 3) saccades directed to a location on a peripheral ring (Ring, bottom) (see Section 3 for details). In all cases, the saccade target was positioned at the same location. A qualitative inspection of the data reveals a number of differences in firing patterns across the three conditions. For example, in the Space condition only, activity is moderate until the period of time just at and after the saccade (around 1000 ms) at which point the neuron fires a large burst of activity. In the Ring condition, on the other hand, neural activity builds up prior to the cue onset (time 0) and the neuron fires a robust, brief response just after cue onset. Intriguingly, neural activity in the Dot condition is moderate across the whole trial, and the neuron fires no salient bursts. The firing patterns of this neuron across multiple conditions underscore a number of difficulties in interpretation of firing rate data. First, neurons fire differentially across conditions and a different time points across conditions. So, for example, this neuron exhibits more activity in the Space than Dot conditions, but only at select periods in time (e.g., around the saccade). Thus characterizing this neuron as firing more in the Space condition than the Dot condition is correct only if we focus on the peri-/post-saccade epoch. In comparing the Ring condition to the Dot condition, on the other hand, it is clear that the differentiation between the two conditions primarily occurs around the time of the visual stimulus/cue onset.

A question arises, consequently, as to what portion of the firing patterns should be compared across conditions in analysis. One common solution to this question, which has been employed in multiple studies (e.g., Moorman and Olson, 2007a,b), is to choose epochs on the basis of either periods of behavioral interest (e.g., +/- 500 ms around the cue) or based on regions in the histogram that display clear differences across conditions (e.g., comparing the peri-saccadic epoch between Space and Dot conditions). Ultimately, these solutions are subjective in nature and may ignore subtle differences that exist across conditions. This issue is both exemplified and potentially solved by considering the difference of density function plots shown in Figure 3. These plots

display the posterior mean (solid red lines) and 95% interval bands (dashed lines) for the difference of neuronal firing density functions between any pair of conditions. The x-axis shows time across the trial, and the y-axis shows difference of posterior probability densities. Functionally speaking the figures characterize the difference in firing between the two histograms compared, with the black horizontal line demarcating no difference, curves above that line demarcating stronger firing in the first condition (e.g., Space in Figure 3, top panel) and curves below the black line demarcating stronger firing in the second condition (Dot in Figure 3, top panel). The probability interval bands produced by the Bayesian model are key for proper uncertainty quantification, as they represent statistically significant differences in firing patterns between the two conditions. In particular, the analysis presented in this figure advances our understanding of the differences between two firing rates in a number of major ways. First, the graph displays the periods of time at which there are differences in firing between the two conditions and, in particular, it describes the periods that firing patterns are statistically different from one another. Second, it does so in a continuous fashion over the entire experimental time interval. This allows transitional time points, such as onset or offset of statistical differences, to be calculated. It also precludes the necessity to compare average data culled from arbitrarily-defined epochs.

Insert Figure 3 Here

Perhaps most importantly, the analysis reveals subtle effects not necessarily observable by qualitatively comparing firing patterns. This is well-exemplified in Figure 3. Roughly, between the time-points -1000 and 1000 ms, there is a prominent, and statistically significant, difference in firing between Space and Dot conditions, with stronger firing occurring in the Dot, as compared to the Space, condition. Although this difference is somewhat apparent in the histograms, the onset, duration, and consistency of the difference is hard to identify without aid of the model-based formal inference reported in Figure 3. Of interest from a functional perspective, this significant difference appears at approximately at or after fixation onset (approximately -1000 ms)

and persists throughout the duration of fixation until saccade onset (approximately +1000 ms). After saccade onset, firing is significantly stronger in the Space condition, and this predominance of Space persists until reward/reward-cue presentation. From a comparison of these two conditions one might come to a number of conclusions regarding the sensorimotor properties of this neuron: that it is, in some way, visually driven, exemplified by the Dot-condition-bias preceding the saccade (with the array of target dots potentially driving neural activity) as well as by the onset of Space selectivity in the period of time around or following the saccade (with the presentation of the reward-cue stimulus potentially driving activity). The hypothesis that this neuron is driven by visual stimuli is further advanced by comparing the Space and Ring conditions (histograms in the top and the bottom panels of Figure 1, and difference of densities plot in Figure 3). Focusing on the difference of densities plot, it is clear that activity in the Ring condition is stronger at the onset of the ring stimulus approximately -500 ms, decays following stimulus onset, and increases again during cue presentation. A similar pattern is seen when comparing Ring and Dot conditions (histograms in middle and lower panels of Figure 1, and difference of densities plot in Figure 3), and it may be assumed that the ring stimulus produces a stronger visual response based on something as simple as the fact that it is a larger, brighter stimulus. Interpreting this neuron as solely visually-responsive, however, is precluded by the fact that there is a strong burst of activity following the spatial cue in the Ring condition compared to the Dot and Space conditions, despite the fact that the spatial cue is the same in all three conditions.

Clearly SEF neurons fire with complex patterns under different conditions, and the activity of SEF neurons have been correlated with numerous sensorimotor and cognitive behaviors (see Experimental Design and Methods). This fact is further exemplified in observing the firing rate histograms and inference for the difference in densities for the other tested neuron (sp220b.1) shown in Figures 4 and 5. Figure 4 shows an example of a neuron that responds both to the presentation of visual stimuli as well as to saccades

directed to visually-marked targets. This can be seen as increased firing following the onset of stimuli and during saccades in Dot and Ring histograms (middle and lower panels of Figures 4) as well as in the difference of densities plots, which reveal statistically significant differences between these conditions and the Space condition (top and middle panels of Figure 5).

Insert Figures 4 to 5 Here

For each of the two neurons, Figure 6 plots posterior mean estimates and 95% interval bands for the firing intensities associated with each of the three conditions. The estimates of the firing intensity functions arising from the Bayesian nonparametric model smooth the PSTH histograms, but at the same time, uncover multimodal patterns, most notably, for the Ring condition. It is also noteworthy that the uncertainty bands have more uniform width across conditions compared to the interval estimates (not shown) obtained from the model in Kottas and Behseta (2010), which is applied separately to the data from each condition. This result highlights the borrowing of strength across conditions facilitated by the dependent nonparametric mixture model developed in Section 2.2.

Insert Figure 6 Here

The results from these two neurons, both the traditional histogram-based data representation as well as difference of densities plots representing significance characterized by the analysis described here, demonstrate the complexity, both temporal (across trials) and categorical (across conditions), of firing patterns of (here SEF) neurons even in relatively simple eye-movement tasks. A thorough characterization of the firing properties of the population of SEF neurons from which these were selected is beyond the scope of this paper and will be presented in a separate manuscript (Moorman, Medler, and Olson, in preparation). These data demonstrate, however, how the proposed statistical methodology can be used to identify differences in firing patterns across multiple conditions, with appropriate uncertainty quantification.

5 Discussion

The problem of comparing neuronal intensity rates of a neuron recorded under two or more experimental conditions has been examined in Behseta and Chenouri (2011), Behseta et.al (2007), Behseta et al. (2005), and Behseta and Kass (2005). To model PSTH, the authors employ a technique called Bayesian Adaptive Regression Splines or BARS (DiMateo et.al, 2001). BARS utilizes Bayesian generalized linear modeling over a natural cubic spline basis via the reversible-jump Markov Chain Monte Carlo (MCMC) technique of Green (1995). Subsequently, a test statistic is developed to compare the firing intensities of two conditions. A modified form of Hotelling's T^2 test-statistic, as well as Bayes factors were employed to test the hypothesis that the two curves behave the same (Behseta and Kass, 2005). To test the hypothesis of equality of more than two curves – specifically more than two firing intensity functions – Behseta et al. (2007) developed a functional Multivariate Analysis of Variance (fMANOVA) paradigm that may be applied over a population of functional neuronal data estimated with BARS.

In more recent work Behseta and Chenouri (2011) showed that a test based on a nonparametric spatial signed-rank statistic of Möttönen and Oja (1995) attains a very high power. Also, in the case of more than two conditions, tests constructed based on likelihood ratio statistics are shown to be quite powerful (Behseta et.al, 2007). The works cited above rely on asymptotic properties of the proposed testing procedures. Also, as discussed in Behseta et al. (2007), in multi-trial neurophysiological data, there are scenarios in which trial-averaged firing rates may be misleading. Alternatively, this paper provides a more general framework through developing fully inferential nonparametric Bayesian models for temporal point patterns assumed to arise from non-homogeneous Poisson processes over time.

Understanding how neurons fire under different behavioral contexts is often challenging, particularly when studying neurons in complicated sensorimotor cortex such as the SEF. While such neurons may have selectivity for a particular movement di-

rection or a particular visual stimulus, responses are often complex combinations of the two (Bon and Lucchetti 1991, 1992; Chen and Wise 1995a, b, 1996, 1997; Coe et al. 2002; Fujii et al. 2002; Hanes et al. 1995; Lee and Tehovnik 1995; Moorman and Olson 2007a,b; Mushiake et al. 1996; Olson and Gettner 1995, 1999, 2002; Olson et al. 2000; Olson and Tremblay 2000; Russo and Bruce 1996, 2000; Schall 1991a, b; Schlag-Rey et al. 1997; Schlag and Schlag-Rey 1985, 1987b; Schlag et al. 1992; Tremblay et al. 2002). Furthermore, as exemplified by the data presented in this paper, neuronal responses often change over time with selectivity for one condition (manifested by selectively enhanced firing rate) predominating during one timepoint in a behavioral task and selectivity for another condition predominating at another. Identifying periods of time during a trial for analysis is a somewhat uncontrolled process. Although in some cases an investigator can be guided by behavioral epochs (e.g., period of time when stimulus is on), using predetermined analysis epochs could bias investigators to see effects only where they are expected. Previous experimental electrophysiology studies (e.g., Moorman and Olson 2007a,b) have used multivariate ANOVA and its associated post hoc analyses for comparing the differences in firing rate across conditions, the use of which is computationally expensive simply due to the fact that in the presence of k experimental conditions, $\binom{k}{2}$ pairwise comparisons need to be considered. This is typically coupled by the familiar problems associated with multiple testing requiring some sort of a type-I error adjustment. In sharp contrast to the above approach, under the methodology presented in this work, statistically significant differences across multiple conditions are evaluated at multiple time points in a continuous fashion. One can easily imagine applying this technique to multiple neurons recorded from the same area, perhaps in combination with an awareness of periods of behavioral significance, to better understand how neurons in a region differentially characterize behavioral conditions.

In addition to defining specific time periods for further analysis, the methodology presented here provides additional benefits by virtue of being a continuous data analysis technique. For example, precise onset and offsets of specific differences across

conditions can be quantified based on significant differences in firing rate. Future developments will consist of utilizing the differences between selected pairs of conditions to develop a clustering paradigm to quantitatively identify subsets of neurons whose firing intensity rates share common patterns. These types of clustering can be used not only to characterize differences across neurons but to identify the precise features in differential firing maximally distinguishing neurons in a population using, for example, principal components analysis to identify primary distinguishing characteristics. Characterization based on clustering will also be invaluable in understanding results from model-based simultaneous analysis of more than three conditions, noted above. These extensions of the current model are currently under investigation in order to further develop the sophistication and general utility of the methodology.

With regard to the stochastic model for the point process of firing times, a more realistic approach would consider modeling $\lambda(t|H_t)$, the instantaneous firing rate at time t given H_t, the spiking history preceding time t. The main reasoning behind this strategy is to avoid pooling spike trains over multiple trials. Conditioning on the history allows one to take into account the effects of refractory period and bursting delays, which in turn would contribute to the departure from the NHPP assumption. A natural starting point for development of Bayesian nonparametric methodology is structured non-Poisson point process models, for instance, in the spirit of the inhomogeneous Markov interval (IMI) and the additive IMI models of Kass and Ventura (2001), or the time-rescaled renewal process model of Brown et al. (2002). Future work will explore this methodologically interesting and scientifically relevant research direction.

Finally, in early sensory or motor areas, neurons have more direct response properties: a neuron in primary visual cortex, for example, will parametrically fire based on variations in visual stimulus properties (see, for example, Chen et al., 2009 and references therein); though certain secondary factors such as attention have recently been shown to produce more complex responses (Roelfsema et al., 2007; Smith et al., 2007). Neurons in associative cortex, such as in the SEF, respond conjunctively, based

on interactions of sensory and motor variables, as well as dynamically, with selectivity changing over the course of a single trial. Analytical techniques used to disentangle the relative contributions of behavioral variables are critical in understanding what role these complex cortical areas play in shaping behavior. The approach to statistical modeling and inference presented here is a first step in a new line of analyses designed to address this complexity and to identify key components in neuronal activation driving behavior.

Acknowledgments

The authors wish to thank two referees for helpful comments. The work of A. Kottas and V. Poynor was supported in part by the National Science Foundation under award SES 1024484.

Appendix: Implementation details

Here, we provide technical details on prior specification and the approach to posterior inference for the Bayesian nonparametric model developed in Section 2.2.

Based on the definition of the logit-normal distribution, note that fitting the location mixture of logit-normals model, $\int t^{-1}(1-t)^{-1}\mathrm{N}(\log(t/(1-t)); \mu, \sigma_i^2)dG_i(\mu)$ (where $t \in (0,1)$), to the data t_{ij} on the unit interval is equivalent to fitting the location mixture of normals model, $\int \mathrm{N}(y; \mu, \sigma_i^2)dG_i(\mu)$ (for which $y \in \mathbb{R}$), to the logit-transformed data $y_{ij} = \log(t_{ij}/(1-t_{ij}))$. The latter is computationally more efficient, and we thus work with the logit-transformed data. Of course, all the inferences can be reported on the original scale by applying the inverse logistic transformation to evaluate the neuronal firing densities $f(t_0; G_i, \sigma_i^2)$, for $i = 1, ..., I$, at any grid of time points t_0.

Prior specification

Regarding the hyperparameters of the DDP prior, we place a gamma(a_α, b_α) prior distribution on α (with mean a_α/b_α), a N(a_λ, b_λ^2) prior distribution on λ, and an inverse Wishart prior distribution on Λ with scale and matrix parameters a_Λ and B_Λ, respectively (thus, $p(\Lambda) \propto |\Lambda|^{-(a_\Lambda+I+1)/2} \exp(-0.5\text{tr}(B_\Lambda \Lambda^{-1}))$). For each of the NHPP mean parameters γ_i, we used the reference prior given in Section 2.2.1. Finally, the mixture kernel scale parameters σ_i^2 are assigned a hierarchical inverse gamma prior distribution, inv-gamma(c, β), with fixed shape parameter c and random scale parameter β (such that the mean is $\beta/(c-1)$, provided $c > 1$); an exponential prior with mean m_β is placed on β.

To specify the parameters of the prior distributions discussed above, we work with a single component of the mixture model (i.e., the limiting case of the DDP prior with $\alpha \to 0^+$) and use a relatively noninformative approach based on a guess, say R, at the range of values for the logit-transformed data. In fact, we take the same range for the data from all conditions, thus resulting in the same prior for all condition-specific parameters. Then, based on the prior structure, we can use the following approximation to the variance of the population:

$$\begin{aligned}
(R/4)^2 &\approx \text{Var}(Y) = \text{E}(\text{Var}(Y|\theta_{ji}, \sigma_i^2)) + \text{Var}(\text{E}(Y|\theta_{ji}, \sigma_i^2)) \\
&= \text{E}(\sigma_i^2) + \text{Var}(\theta_{ji}) = \text{E}(\text{E}(\sigma_i^2|\beta)) + \text{Var}(\text{E}(\theta_{ji}|\lambda, \Lambda_{ii})) + \text{E}(\text{Var}(\theta_{ji}|\lambda, \Lambda_{ii})) \\
&= \text{E}((c-1)^{-1}\beta) + \text{Var}(\lambda) + \text{E}(\Lambda_{ii}) = (c-1)^{-1} m_\beta + b_\lambda^2 + (a_\Lambda - I - 1)^{-1} B_\Lambda^{ii}
\end{aligned}$$

where B_Λ^{ii} is the i-th diagonal element of matrix B_Λ. We set $a_\lambda = 0$ and $c = 2$, the latter providing an infinite prior variance for each σ_i^2. We set $a_\Lambda = 5$, which is the smallest value for the degrees of freedom of an inverse Wishart distribution having a well-defined mean. From the range, we subtract a value of 0.5, and allocate it to the first term, which represents the prior expectation of σ_i^2. The remaining amount of the range is divided equally between the second and third terms. Solving the three terms provides values for m_β, b_λ^2, and B_Λ^{ii}. Matrix B_Λ is then specified by setting

its off-diagonal elements to 0. Prior sensitivity analysis indicated that this default specification does not effect the model's ability to capture the required correlations implied by the posterior distribution for Λ. Finally, we use a relatively dispersed gamma prior for α, setting $a_\alpha = 3$ and $b_\alpha = 0.5$.

MCMC posterior simulation

As discussed in Section 2.2.2, the marginal posterior distribution for the NHPP mean parameters γ_i, which provide the scale for the firing intensity functions, is available analytically. Here, we develop the approach to MCMC simulation from the posterior distribution of all other model parameters.

To sample the posterior distribution of the dependent DP mixture model, we utilize an extension of blocked Gibbs sampling for standard DP mixtures (e.g., Ishwaran and James, 2001). The approach builds from a finite truncation approximation to the DDP($\alpha, \boldsymbol{G}_0$) prior, using its stick-breaking constructive definition. Specifically, we can approximate $\boldsymbol{G}$ in (4) with $\boldsymbol{G}^L = \sum_{\ell=1}^{L} p_\ell \delta_{\boldsymbol{\theta}_\ell}$, where the $\boldsymbol{\theta}_\ell = (\theta_{\ell,1}, ..., \theta_{\ell,I})$ are again i.i.d. from $\boldsymbol{G}_0 = \mathrm{N}_I(\lambda \mathbf{1}_I, \Lambda)$, and the p_ℓ arise from a truncated version of the stick-breaking construction. In particular, based on variables V_ℓ, $\ell = 1, ..., L-1$, which are i.i.d. from a Beta$(1, \alpha)$ distribution, we define: $p_1 = V_1$; $p_\ell = V_\ell \prod_{r=1}^{\ell-1}(1 - V_r)$, for $\ell = 2, ..., L-1$; and $p_L = \prod_{r=1}^{L-1}(1 - V_r) = 1 - \sum_{\ell=1}^{L-1} p_\ell$. Under this particular truncation definition for the vector of probabilities $\boldsymbol{p} = (p_1, ..., p_L)$, the implied prior for $\boldsymbol{p}$ is given by the generalized Dirichlet distribution with density

$$h(\boldsymbol{p}; \alpha) = \alpha^{L-1} p_L^{\alpha-1} (1 - p_1)^{-1} (1 - (p_1 + p_2))^{-1} \times ... \times (1 - \sum_{\ell=1}^{L-2} p_\ell)^{-1}.$$

The approximation can be made accurate to any desired tolerance, with the choice of the truncation level L facilitated through use of relevant DP properties. For instance, for the probabilities in the countable representation for DDP($\alpha, \boldsymbol{G}_0$) realizations in (4), we have $\mathrm{E}(\sum_{j=L}^{\infty} \omega_j \mid \alpha) = \{\alpha/(\alpha+1)\}^{L-1}$. Given a suitable tolerance level and averaging the expectation over the prior distribution for α, this expression pro-

vides the corresponding truncation value L. For the neuronal data analysis results reported in Section 4, we used $L = 100$, which implies $\mathrm{E}(\sum_{j=1}^{L} \omega_j) \approx 0.99986$ under the gamma$(3, 0.5)$ prior for α (posterior inference results were robust to values of L in the range 50 - 100).

As also done with finite mixture models, MCMC sampling proceeds by breaking the mixture through introduction of latent configuration indicators that identify the mixture components. Denote by $\boldsymbol{w} = \{w_j : j = 1, ..., n\}$ the configuration variables taking values in $\{1, ..., L\}$ such that $w_j = \ell$, for $\ell = 1, ..., L$, signifies that the j-th response vector $\boldsymbol{y}_j = \{y_{ij} : i = 1, ..., I\}$ is assigned to the ℓ-th mixture component. Let $\boldsymbol{\theta} = \{\boldsymbol{\theta}_\ell : \ell = 1, ..., L\}$. Then, the hierarchical model for the logit-transformed data can be written as follows:

$$\begin{aligned}
\boldsymbol{y}_j \mid \boldsymbol{w}, \boldsymbol{\theta} &\overset{ind}{\sim} \prod_{i=1}^{I} \{\mathrm{N}(y_{ij}; \theta_{w_j,i}, \sigma_i^2)\}^{s_{ij}}, \;\; j = 1, ..., n \\
w_j \mid \boldsymbol{p} &\overset{iid}{\sim} \sum_{l=1}^{L} p_l \delta_l(w_j), \;\; j = 1, ..., n \\
(\boldsymbol{p}, \boldsymbol{\theta}) \mid \alpha, \lambda, \Lambda &\sim h(\boldsymbol{p}; \alpha) \times \prod_{\ell=1}^{L} \mathrm{N}_I(\boldsymbol{\theta}_\ell; \lambda \mathbf{1}_I, \Lambda) \\
\sigma_i^2 \mid \beta &\overset{iid}{\sim} \text{inv-gamma}(c, \beta), \;\; i = 1, ..., I,
\end{aligned}$$

with the hyperpriors for $(\alpha, \lambda, \Lambda)$ and β discussed above.

Now, Gibbs sampling can be applied to draw from the posterior distribution $p(\boldsymbol{\sigma}^2, \boldsymbol{G}, \alpha, \lambda, \Lambda, \boldsymbol{w}, \beta \mid \text{data})$, where $\boldsymbol{G}$ is defined through $(\boldsymbol{p}, \boldsymbol{\theta})$ based on its truncation approximation. Details on the required posterior full conditional distributions are provided below.

Regarding the Gibbs sampling updates for the $\boldsymbol{\theta}_\ell$, $\ell = 1, ..., L$, let n^* be the number of distinct values in vector $\boldsymbol{w}$, denote the distinct elements by w_k^*, for $k = 1, ..., n^*$, and let $M_\ell = |\{w_j : w_j = \ell\}|$ be the number of configuration variables corresponding to mixture component $\ell = 1, ..., L$. Then, all the $\boldsymbol{\theta}_\ell$ for which $\ell \notin \{w_k^* : k = 1, ..., n^*\}$ are generated from the prior, that is, they are drawn independently from the $\mathrm{N}_I(\lambda \mathbf{1}_I, \Lambda)$

distribution. The posterior full conditional for $\boldsymbol{\theta}_\ell$ corresponding to an active component, that is, $\ell = w_k^*$ for some $k = 1, ..., n^*$, is given by

$$\mathrm{N}_I(\boldsymbol{\theta}_{w_k^*}; \lambda \mathbf{1}_I, \Lambda) \prod_{\{j: w_j = w_k^*\}} \prod_{i=1}^{I} \{\mathrm{N}(y_{ij}; \theta_{w_k^*, i}, \sigma_i^2)\}^{s_{ij}}.$$

We extend the Gibbs sampler to each component of vector $\boldsymbol{\theta}_{w_k^*} = (\theta_{w_k^*,1}, ..., \theta_{w_k^*,I})$, where the posterior full conditional for $\theta_{w_k^*,i}$, $i = 1, ..., I$, is normal with mean $(m_i\sigma_i^2 + S_i^2 \sum_{\{j:w_j=w_k^*\}} s_{ij} y_{ij})/(\sigma_i^2 + S_i^2 \sum_{\{j:w_j=w_k^*\}} s_{ij})$ and variance $\sigma_i^2 S_i^2/(\sigma_i^2 + S_i^2 \sum_{\{j:w_j=w_k^*\}} s_{ij})$. Here, m_i and S_i^2 denote the mean and variance, respectively, of the conditional distribution for $\theta_{w_k^*,i}$, given all the other $\theta_{w_k^*,i'}$, $i' \neq i$, arising from the joint $\mathrm{N}_I(\boldsymbol{\theta}_{w_k^*}; \lambda \mathbf{1}_I, \Lambda)$ distribution.

Each configuration variable w_j, $j = 1, ..., n$, has a discrete posterior conditional distribution taking values in $\{1, ..., L\}$ with probabilities which are proportional to $p_\ell \prod_{i=1}^{I} \{\mathrm{N}(y_{ij}; \theta_{\ell,i}, \sigma_i^2)\}^{s_{ij}}$, for $\ell = 1, ..., L$. The Gibbs sampling updates for vector $\boldsymbol{p}$ are the same with a generic DP mixture model (see, e.g., Ishwaran and James, 2001).

The priors assigned to the DDP prior hyperparameters are conditionally conjugate. Specifically, λ has a normal posterior full conditional distribution with mean $(a_\lambda + b_\lambda^2 \mathbf{1}_I^T \Lambda^{-1} (\sum_{k=1}^{n^*} \boldsymbol{\theta}_{w_k^*}))/(1 + n^* b_\lambda^2 \mathbf{1}_I^T \Lambda^{-1} \mathbf{1}_I)$ and variance $b_\lambda^2/(1 + n^* b_\lambda^2 \mathbf{1}_I^T \Lambda^{-1} \mathbf{1}_I)$; the posterior full conditional distribution for Λ is inverse Wishart with updated scale parameter $a_\Lambda + n^*$ and matrix parameter $B_\Lambda + \sum_{k=1}^{n^*} (\boldsymbol{\theta}_{w_k^*} - \lambda \mathbf{1}_I)(\boldsymbol{\theta}_{w_k^*} - \lambda \mathbf{1}_I)^T$; and α has a gamma posterior full conditional distribution with shape parameter $L + a_\alpha - 1$ and rate parameter $b_\alpha - \log p_L$.

Finally, the posterior full conditional distribution for each σ_i^2 is inverse gamma with shape parameter $c + 0.5 \sum_{j=1}^{n} s_{ij}$ and scale parameter $\beta + 0.5 \sum_{j=1}^{n} s_{ij} (y_{ij} - \theta_{w_j,i})^2$, and β has a gamma posterior full conditional distribution with shape parameter $1 + Ic$ and rate parameter $m_\beta^{-1} + \sum_{i=1}^{I} \sigma_i^{-2}$.

REFERENCES

Behseta, S. and Chenouri, S. (2011). Comparison of two populations of curves with an application in the neuronal data analysis. *Statistics in Medicine* **30**, 1441-1454.

Behseta, S. and Kass, R.E. (2005). Testing equality of two functions using BARS. *Statistics in Medicine* **24**, 3523-3534.

Behseta, S., Kass, R.E., Moorman D.E., Olson, C.R. (2007). Testing equality of several functions: Analysis of single-unit firing-rate curves across multiple experimental conditions. *Statistics in Medicine* **26**, 3958-3975.

Behseta, S., Kass, R.E., and Wallstrom, G.L. (2005). Hierarchical models for assessing variability among functions. *Biometrika* **92**, 419-434.

Bon L, Lucchetti C. Behavioral and motor mechanisms of dorsomedial frontal cortex of macaca monkey. Int J Neurosci, 1991; 60: 187-93.

Bon L, Lucchetti C. The dorsomedial frontal cortex of the macaca monkey: fixation and saccade-related activity. Exp Brain Res, 1992; 89: 571-80.

Brillinger, DR. Nerve cell spike train data analysis: A progression of technique. Journal of the American Statistical Association 1992; 87:260-271.

Brown E, Barbieri R, Ventura V, Kass R, and Frank L. The time-rescaling theorem and its application to neural spike train data analysis. Neural Computation 2002, 14:325-346.

Brown E, Frank L, Tang D, Quirk M, Wilson M. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. Journal of Neuroscience 1998; 18:7411-7425.

Brown E, Kass RE, and Mitra PP. Multiple neural spike train data analysis: state-of-the-art and future challenges. Nature Neuroscience 2004; 7:456-461.

Chen LL, Wise SP. Conditional oculomotor learning: population vectors in the supplementary eye field. J Neurophysiol, 1997; 78: 1166-9.

Chen LL, Wise SP. Evolution of directional preferences in the supplementary eye field during acquisition of conditional oculomotor associations. J Neurosci, 1996; 16: 3067-81.

Chen LL, Wise SP. Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. J Neurophysiol, 1995a; 73: 1101-21.

Chen LL, Wise SP. Supplementary eye field contrasted with the frontal eye field during acquisition of conditional oculomotor associations. J Neurophysiol, 1995b; 73: 1122-34.

Coe B, Tomihara K, Matsuzawa M, Hikosaka O. Visual and anticipatory bias in three cortical eye fields of the monkey during an adaptive decision-making task. J Neurosci, 2002; 22: 5081-90.

De Iorio M, Müller P, Rosner GL, MacEachern SN. An ANOVA Model for Dependent Random Measures. Journal of the American Statistical Association, 2004; 99: 205-215.

DiMatteo, I., Genovese, C.R., and Kass, R.E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika* **88**, 1055-1071.

Dunson DB. Nonparametric Bayes applications to biostatistics. In Bayesian Nonparametrics, Hjort NL, Holmes C, Muller P, and Walker SG eds. Cambridge: Cambridge University Press 2010; 223-268.

Escobar M, West M. Bayesian density estimation and inference using mixtures. Journal of the American Statistical Association 1995; 90: 577-588.

Ferguson TS. A Bayesian analysis of some nonparametric problems. Annals of Statistics 1973; 1: 209-230.

Ferguson, TS. Prior distributions on spaces of probability measures. Annals of Statistics 1974; 2: 615-629.

Fronczyk, K., Kottas A. A Bayesian Nonparametric Modeling Framework for Developmental Toxicity Studies. Technical Report UCSC-SOE-10-11, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, 2010.

Fujii N, Mushiake H, Tamai M, Tanji J. Microstimulation of the supplementary eye field during saccade preparation. Neuroreport, 1995; 6: 2565-8.

Fujii N, Mushiake H, Tanji J. Distribution of eye- and arm-movement-related neuronal activity in the SEF and in the SMA and Pre-SMA of monkeys. J Neurophysiol, 2002; 87: 2158-66.

Gelfand AE, Kottas A, MacEachern SN. Bayesian Nonparametric Spatial Modeling With Dirichlet Process Mixing. Journal of the American Statistical Association, 2005; 100: 1021-35.

Gerstein G, Mandelbrot B. Random walk models for spike activity of a single neuron. Biophysical Journal 1967; 4:41-68.

Ghosh JK, RV Ramamoorthi. 2003. Bayesian Nonparametrics. New York: Springer.

Green PG.,Reversible jump Markov chain Monte Carlo computation and Bayesian model determination , Biometrika (1995) 82 (4): 711-732.

Hanes DP, Thompson KG, Schall JD. Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. Exp Brain Res, 1995; 103: 85-96.

Hanson, T., A. Branscum, and W. Johnson. 2005. Bayesian nonparametric modeling and data analysis: An introduction. In Bayesian Thinking: Modeling and Computation (Handbook of Statistics, vol. 25), Dey DK and Rao CR (eds). Elsevier: Amsterdam, pp. 245-278.

Ishwaran H, James LF. Gibbs Sampling Methods for Stick-Breaking Priors. Journal of the American Statistical Association 2001; 96: 161-173.

Ishwaran H, James LF. Computational methods for multiplicative intensity models using weighted gamma processes: Proportional hazards, marked point processes, and panel count data. Journal of the American Statistical Association 2004; 99: 175-190.

Kass RE, Ventura V. A spike-train probability model. Neural Computation 2001; 13: 1713-1720.

Kass RE, Ventura V, Brown, EN. Statistical issues in the analysis of neuronal data. Journal of Neurophysiology 2005; 94: 8-25.

Kottas A. Dirichlet process mixtures of Beta distributions, with applications to density and intensity estimation. Proceedings of the Workshop on Learning with Nonparametric Bayesian Methods. In 23rd International Conference on Machine Learning Pittsburgh, PA, 2006.

Kottas A, Behseta S. Bayesian nonparametric modeling for comparison of single-neuron firing intensities. Biometrics 2010; 66: 277-286.

Kottas A, Krnjajić M. Bayesian Semiparametric Modelling in Quantile Regression. Scandinavian Journal of Statistics, 2009; 36: 297-319.

Kottas A, Sansó B. Bayesian Mixture Modeling for Spatial Poisson Process Intensities, with Applications to Extreme Value Analysis. Journal of Statistical Planning and Inference 2007; 137:3151-3163.

Kottas A, Duan JA, Gelfand AE. Modeling Disease Incidence Data with Spatial and Spatio-temporal Dirichlet Process Mixtures. Biometrical Journal, 2008; 50: 29-42.

MacEachern SN. Dependent Dirichlet Processes. Technical Report, Department of Statistics, The Ohio State University, 2000.

Mann SE, Thau R, Schiller PH. Conditional task-related responses in monkey dorsomedial frontal cortex. Exp Brain Res, 1988; 69: 460-8.

Lee K, Tehovnik EJ. Topographic distribution of fixation-related units in the dorsomedial frontal cortex of the rhesus monkey. Eur J Neurosci, 1995; 7: 1005-11.

Lo AY, Weng, CS. On a class of Bayesian nonparametric estimates: II. Hazard rate estimates. Annals of the Institute of Statistical Mathematics 1989; 41: 227-245.

Martinez-Trujillo JC, Klier EM, Wang H, Crawford JD. Contribution of head movement to gaze command coding in monkey frontal cortex and superior colliculus. J Neurophysiol, 2003a; 90: 2770-6.

Martinez-Trujillo JC, Medendorp WP, Wang H, Crawford JD. Frames of Reference for Eye-Head Gaze Commands in Primate Supplementary Eye Fields. Neuron, 2004; 44: 1057-66.

Martinez-Trujillo JC, Wang H, Crawford JD. Electrical stimulation of the supplementary eye fields in the head-free macaque evokes kinematically normal gaze shifts. J Neurophysiol, 2003b; 89: 2961-74.

Missal M, Heinen SJ. Facilitation of smooth pursuit initiation by electrical stimulation in the supplementary eye fields. J Neurophysiol, 2001; 86: 2413-25.

Missal M, Heinen SJ. Supplementary eye fields stimulation facilitates anticipatory pursuit. J Neurophysiol, 2004; 92: 1257-62.

Mitz AR, Godschalk M. Eye-movement representation in the frontal lobe of rhesus monkeys. Neurosci Lett, 1989; 106: 157-62.

J. Mottonen and H. Oja, Multivariate spatial sign and rank methods, Nonparametric Statist. 1995; 5: 201-213.

Moorman DE, Olson CR. Combination of neuronal signals representing object-centered location and saccade direction in macaque supplementary eye field. J Neurophysiol, 2007a; 97: 3554-66.

Moorman DE, Olson CR. Impact of experience on the representation of object-centered space in the macaque supplementary eye field. J Neurophysiol, 2007b; 97: 2159-73.

Müller, P, Quintana, FA. Nonparametric Bayesian Data Analysis. Statistical Science 2004; 19:95-110.

Mushiake H, Fujii N, Tanji J. Visually guided saccade versus eye-hand reach: contrasting neuronal activity in the cortical supplementary and frontal eye fields. J Neurophysiol, 1996; 75: 2187-91.

Neal, RM. Markov chain sampling methods for Dirichlet process mixture models. Journal of Computational and Graphical Statistics 2000; 9:249-265.

Olson CR, Gettner SN. Macaque SEF neurons encode object-centered directions of eye movements regardless of the visual attributes of instructional cues. J Neurophysiol, 1999; 81: 2340-6.

Olson CR, Gettner SN. Neuronal activity related to rule and conflict in macaque supplementary eye field. Physiol Behav, 2002; 77: 663-70.

Olson CR, Gettner SN. Object-centered direction selectivity in the macaque supplementary eye field. Science, 1995; 269: 985-8.

Olson CR, Gettner SN, Ventura V, Carta R, Kass RE. Neuronal activity in macaque supplementary eye field during planning of saccades in response to pattern and spatial cues. J Neurophysiol, 2000; 84: 1369-84.

Olson CR, Tremblay L. Macaque supplementary eye field neurons encode object-centered locations relative to both continuous and discontinuous objects. J Neurophysiol, 2000; 83: 2392-411.

Paninski L, Brown EN, Iyengar S, Kass RE. Statistical models of spike trains. In Stochastic Methods in Neuroscience, Laing C, Lord G, eds. Newyork: Oxford, 2010.

Rodriguez A, ter Horst E. Bayesian dynamic density estimation. Bayesian Analysis, 2008; 3: 339-366.

Russo GS, Bruce CJ. Neurons in the supplementary eye field of rhesus monkeys code visual targets and saccadic eye movements in an oculocentric coordinate system. J Neurophysiol, 1996; 76: 825-48.

Russo GS, Bruce CJ. Effect of eye position within the orbit on electrically elicited saccadic eye movements: a comparison of the macaque monkey's frontal and supplementary eye fields. J Neurophysiol, 1993; 69: 800-18.

Russo GS, Bruce CJ. Supplementary eye field: representation of saccades and relationship between neural response fields and elicited eye movements. J Neurophysiol, 2000; 84: 2605-21.

Schall JD. Neuronal activity related to visually guided saccades in the frontal eye fields of rhesus monkeys: comparison with supplementary eye fields. J Neurophysiol, 1991a; 66: 559-79.

Schall JD. Neuronal activity related to visually guided saccadic eye movements in the supplementary motor area of rhesus monkeys. J Neurophysiol, 1991b; 66: 530-58.

Schlag-Rey M, Amador N, Sanchez H, Schlag J. Antisaccade performance predicted by neuronal activity in the supplementary eye field. Nature, 1997; 390: 398-401.

Schlag J, Schlag-Rey M. Evidence for a supplementary eye field. J Neurophysiol, 1987; 57: 179-200.

Schlag J, Schlag-Rey M. Unit activity related to spontaneous saccades in frontal dorsomedial cortex of monkey. Exp Brain Res, 1985; 58: 208-11.

Schlag J, Schlag-Rey M, Pigarev I. Supplementary eye field: influence of eye position on neural signals of fixation. Exp Brain Res, 1992; 90: 302-6.

Sethuraman J. A constructive definition of Dirichlet priors. Statistica Sinica 1994; 4:639-650.

Taddy M, Kottas A. Dirichlet Process Mixture Modeling for Marked Poisson Processes. Technical Report UCSC-SOE-09-31, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, 2009.

Tehovnik EJ, Lee K. The dorsomedial frontal cortex of the rhesus monkey: topographic representation of saccades evoked by electrical stimulation. Exp Brain Res, 1993; 96: 430-42.

Tehovnik EJ, Lee K, Schiller PH. Stimulation-evoked saccades from the dorsomedial frontal cortex of the rhesus monkey following lesions of the frontal eye fields and superior colliculus. Exp Brain Res, 1994; 98: 179-90.

Tehovnik EJ, Slocum WM. Effects of training on saccadic eye movements elicited electrically from the frontal cortex of monkeys. Brain Res, 2000; 877: 101-6.

Tehovnik EJ, Slocum WM, Schiller PH. Behavioural conditions affecting saccadic eye movements elicited electrically from the frontal lobes of primates. Eur J Neurosci, 1999; 11: 2431-43.

Tehovnik EJ, Slocum WM, Tolias AS, Schiller PH. Saccades induced electrically from the dorsomedial frontal cortex: evidence for a head-centered representation. Brain Res, 1998; 795: 287-91.

Tehovnik EJ, Sommer MA. Compensatory saccades made to remembered targets following orbital displacement by electrically stimulating the dorsomedial frontal cortex or frontal eye fields of primates. Brain Res, 1996; 727: 221-4.

Tehovnik EJ, Sommer MA. Electrically evoked saccades from the dorsomedial frontal cortex and frontal eye fields: a parametric evaluation reveals differences between areas. Exp Brain Res, 1997; 117: 369-78.

Tian JR, Lynch JC. Slow and saccadic eye movements evoked by microstimulation in the supplementary eye field of the cebus monkey. J Neurophysiol, 1995; 74: 2204-10.

Tremblay L, Gettner SN, Olson CR. Neurons with object-centered spatial selectivity in macaque SEF: do they represent locations or rules? J Neurophysiol, 2002; 87: 333-50.

Ventura V, Carta R, Kass RE, Gettner SN, Olson CR. Statistical analysis of temporal evolution in single-neuron firing rates. Biostatistics 2002; 3: 1-20.

Walker, S.G., P. Damien, P.W. Laud, and A.F.M. Smith. 1999. Bayesian nonparametric inference for random distributions and related functions (with discussion). Journal of the Royal Statistical Society, Series B 61:485-527.

West M. Hierarchical mixture models in neurological transmission analysis. Journal of the American Statistical Association 1997; 92:587-606.

Wolpert RL, Ickstadt K. Poisson/Gamma random field models for spatial statistics. Biometrika 1998; 85:251-267.

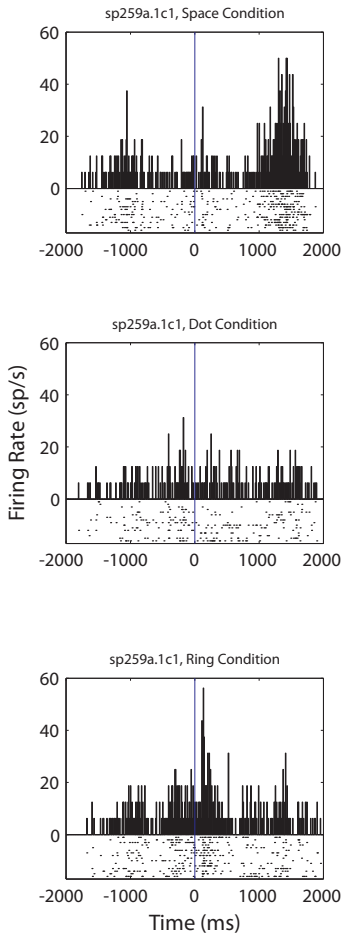

Figure 1: Neuron sp259a.1. PSTH is shown for three experimental conditions (conditions 1, 2, and 3). A 4000 millisecond window is considered. The time is aligned on the spatial cue onset. Conditions 1 (Space, top panel), 2 (Dot, middle panel), and 3 (Ring, lower panel) demonstrate different responses.

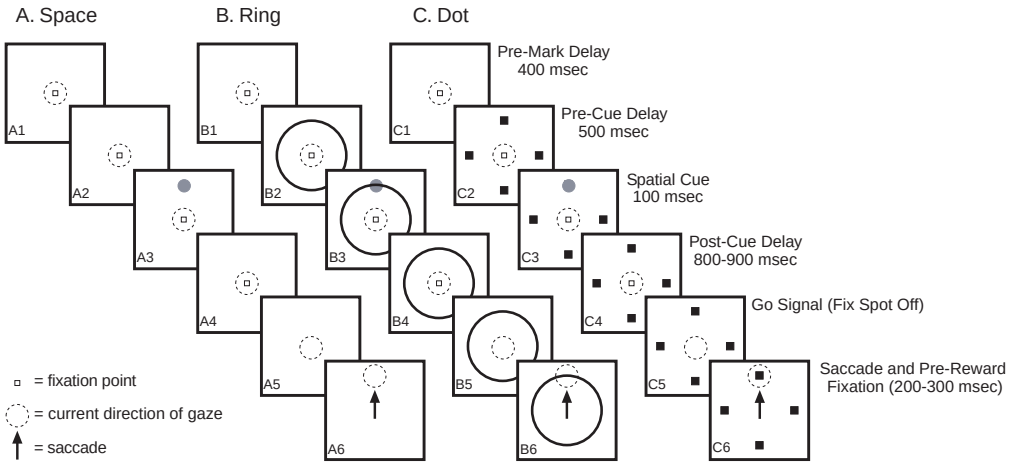

Figure 2: The experiment. The three different memory-guided saccade tasks performed in the variable target experiment. Saccades were made to one of three different target types at one of four locations separated by 90 degree angles. Note for the analyses described here, only one target location per target type is described. Monkeys fixated a central spot to initiate the trial (figures A1-2, B1-2, C1-2). Small central square represents the fixation spot, and the dotted ring indicates the location of the monkey's gaze. After 400 ms following attainment of fixation, one of three events occurred: an array of four dots representing the four potential targets appeared (Dot condition, C2), a circular ring appeared at equal eccentricity to and passing through the positions of the array of dots (Ring condition, B2), or no targets appeared (Space condition, A2). After an additional 500 ms a green cue spot was flashed in one of the four target locations for 100 ms (A3-C3). Following a random delay of 800 to 900 ms (A4-C4), the fixation spot was extinguished and the monkey was required to make a saccade to the previously cued location (A5-6, B5-6, C5-6). Arrow represents the executed saccade. Reward was delivered after the monkey made a saccade directly to the cued location within 150 ms of breaking fixation and then fixated on the cued location for 200-300 ms.

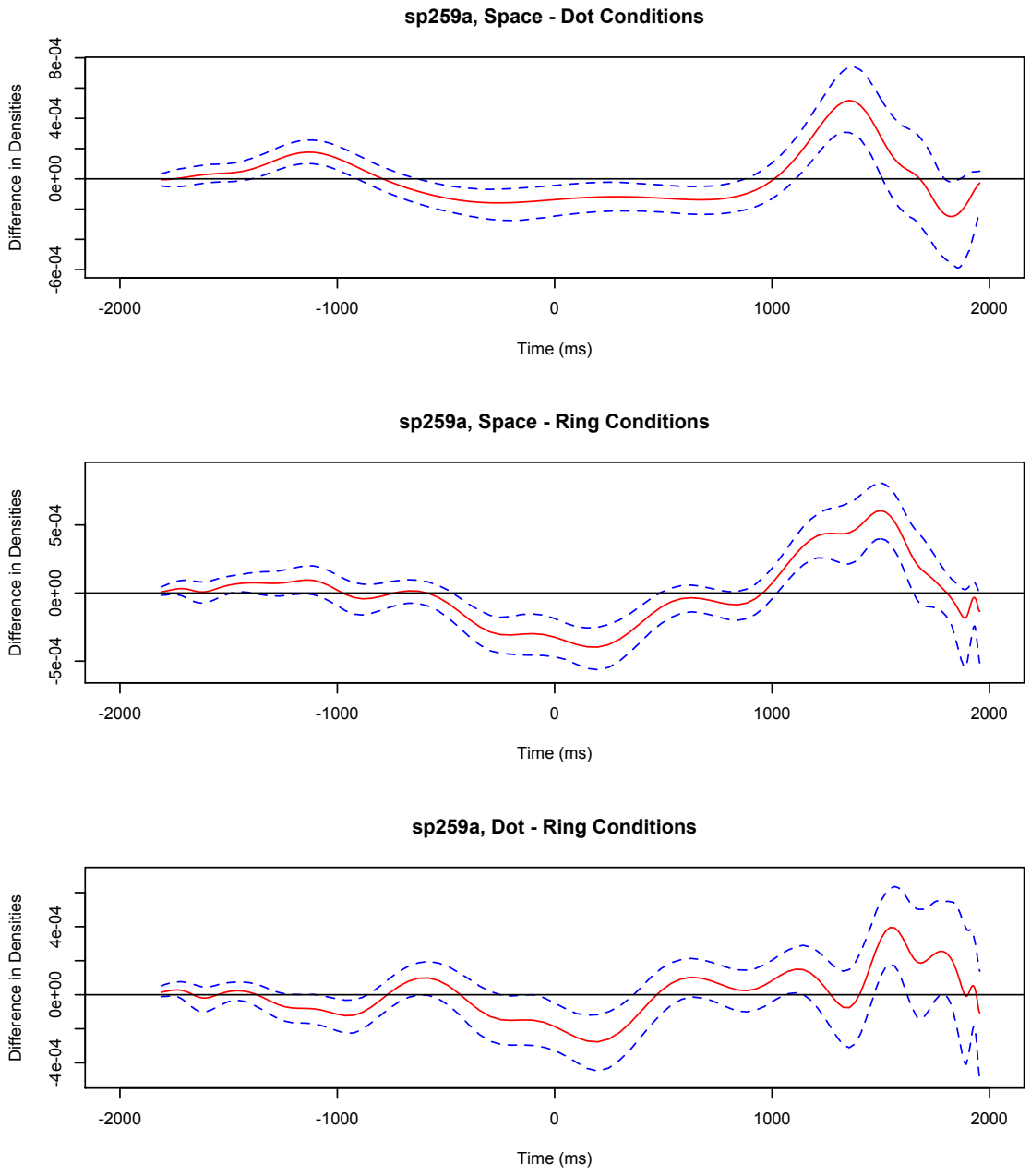

Figure 3: Pairwise differences in firing densities for neuron sp259a.1. The top panel contains the posterior mean difference (solid red line), along with 95% posterior bands (dashed lines) for conditions 1 and 2. Conditions 1 and 3 are compared in the middle panel, and conditions 2 and 3 in the lower panel. The solid black line is placed at the 0 difference mark.

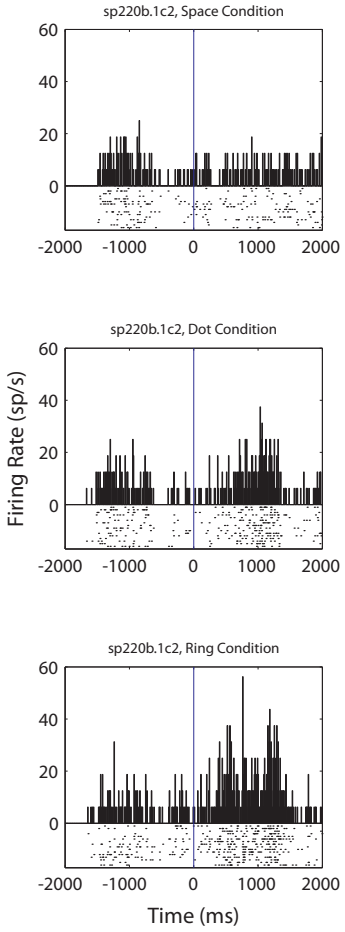

Figure 4: Neuron sp220b.1. PSTH is shown for three experimental conditions (conditions 1, 2, and 3). The time is aligned on the spatial cue onset. Conditions 1 (Space, top panel), 2 (Dot, middle pattern), and 3 (Ring, lower panel) demonstrate different responses.

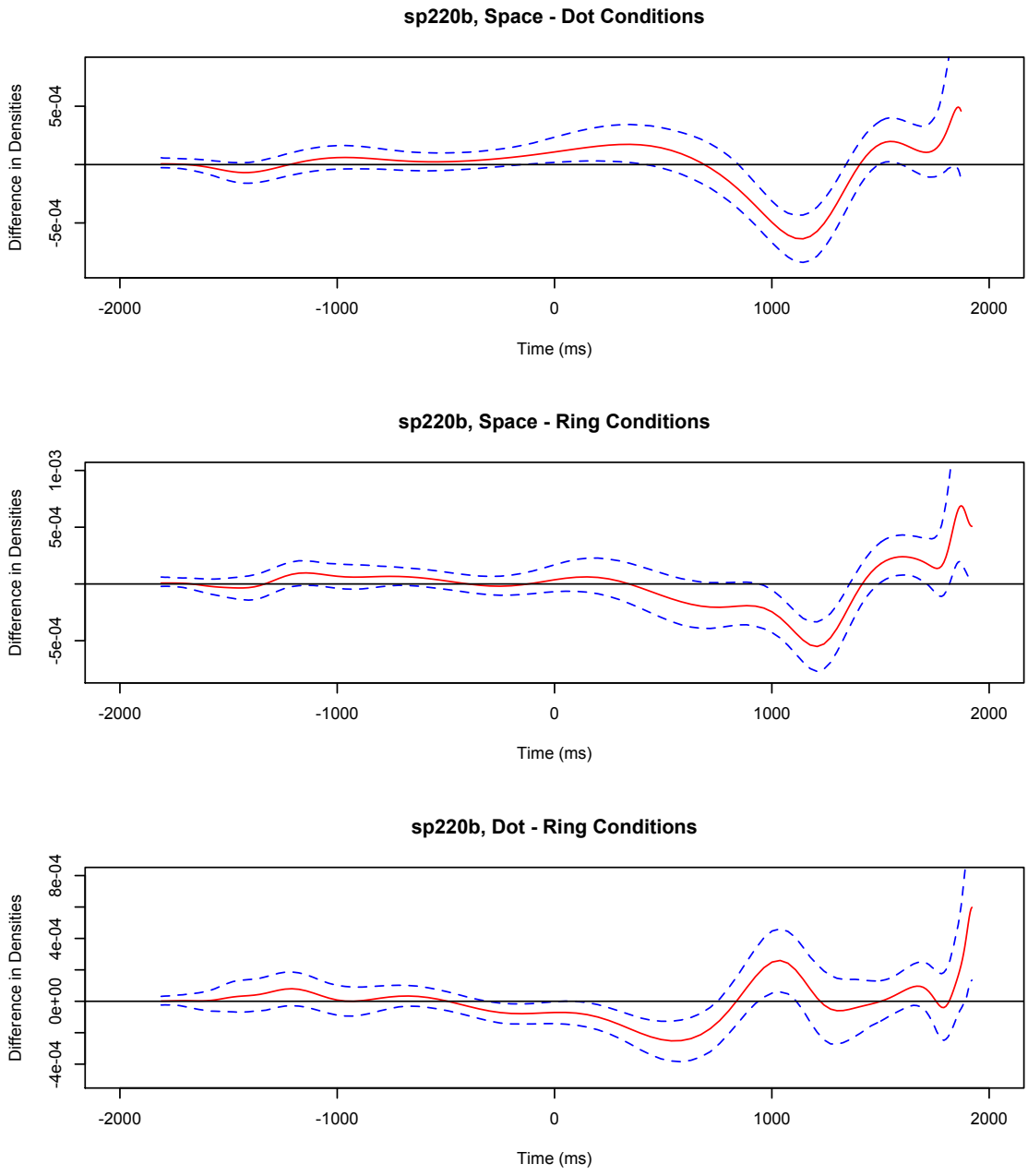

Figure 5: Pairwise differences in firing densities for neuron sp220b.1. The top panel contains the posterior mean difference (solid red line), along with 95% posterior bands (dashed lines) for conditions 1 and 2. Conditions 1 and 3 are compared in the middle panel, and conditions 2 and 3 in the lower panel. The solid black line is placed at the 0 difference mark.

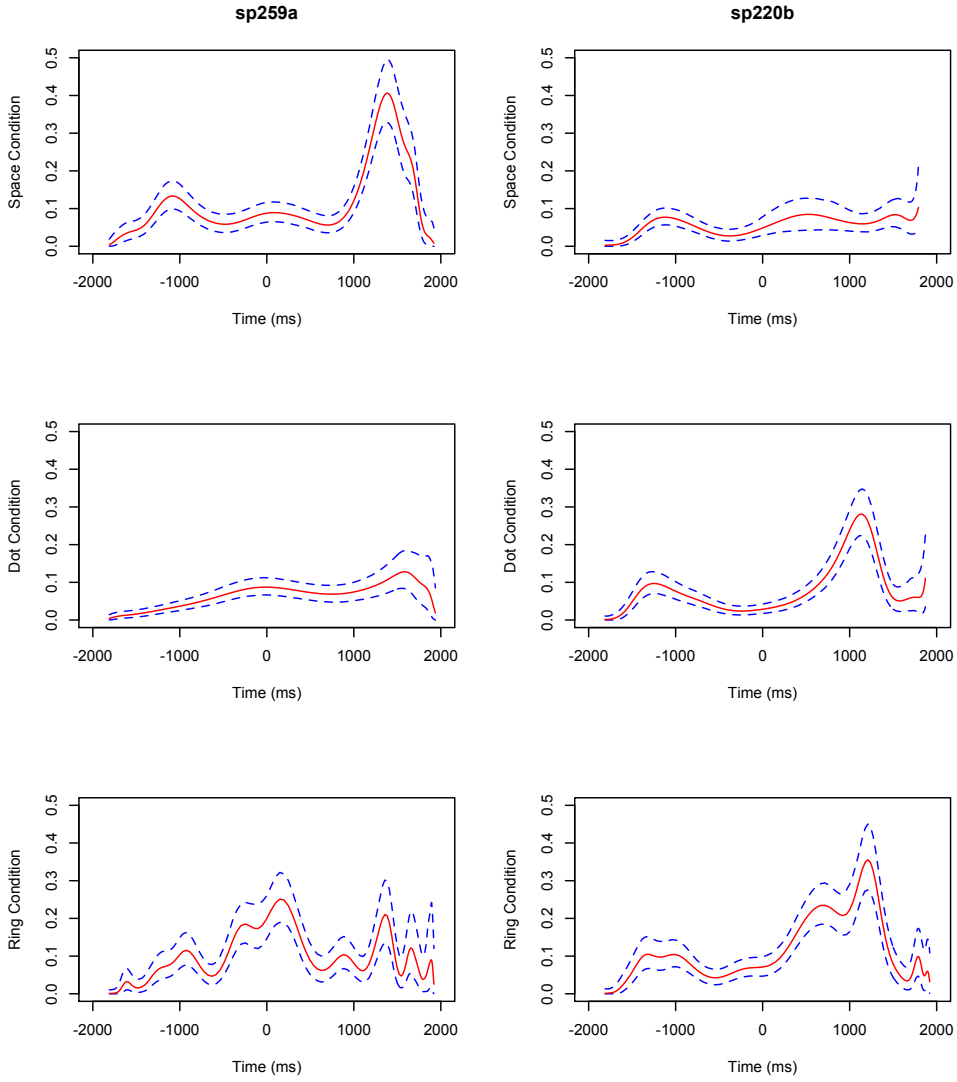

Figure 6: Posterior mean (solid line) and 95% interval estimates (dashed lines) for the neuronal intensity function of each condition for neuron sp259a.1 (left column) and neuron sp220b.1 (right column).