

Estimating Instantaneous Irregularity of Neuronal Firing

Takeaki Shimokawa

shimokawa@ton.scphys.kyoto-u.ac.jp

Shigeru Shinomoto

shinomoto@scphys.kyoto-u.ac.jp

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

Cortical neurons *in vivo* had been regarded as Poisson spike generators that convey no information other than the rate of random firing. Recently, using a metric for analyzing local variation of interspike intervals, researchers have found that individual neurons express specific patterns in generating spikes, which may symbolically be termed regular, random, or bursty, rather invariantly in time. In order to study the dynamics of firing patterns in greater detail, we propose here a Bayesian method for estimating firing irregularity and the firing rate simultaneously for a given spike sequence, and we implement an algorithm that may render the empirical Bayesian estimation practicable for data comprising a large number of spikes. Application of this method to electrophysiological data revealed a subtle correlation between the degree of firing irregularity and the firing rate for individual neurons. Irregularity of firing did not deviate greatly around the low degree of dependence on the firing rate and remained practically unchanged for individual neurons in the cortical areas V1 and MT, whereas it fluctuated greatly in the lateral geniculate nucleus of the thalamus. This indicates the presence and absence of autocontrolling mechanisms for maintaining patterns of firing in the cortex and thalamus, respectively.

1 Introduction

Spike trains generated by cortical neurons appear irregular, changing from trial to trial for identical behavioral stimuli. Such irregular firing may be regarded as Poisson random events that convey no information other than the rate of occurrence (Shadlen & Newsome, 1994, 1998; Dayan & Abbott, 2001). This idea was supported by various observations, such as the exponential distribution of interspike intervals (ISIs) and the coefficient of variation being close to unity (Softky & Koch, 1993). However, these statistics are easily affected by fluctuations in the firing rate, which often occur *in vivo* (Gabbiani & Koch, 1998).

Recent sophisticated analyses have shown that neuronal firing *in vivo* is not exactly a Poisson random phenomenon (Richmond & Optican, 1990;

Pillow, Paninski, Uzzell, Simoncelli, & Chichilnisky, 2005; Kostal & Lansky, 2006). Furthermore, firing irregularity has been shown to be rather specific to individual neurons and, for some cortical areas, invariant with respect to time and the modulation of firing rate (Shinomoto, Shima, & Tanji, 2003; Shinomoto, Miyazaki, Tamura, & Fujita, 2005; Nawrot et al., 2008). There is also a contrary report that in some cortical areas, firing irregularity varies in time and according to the behavioral context (Davies, Gerstein, & Baker, 2006). Therefore, it is desirable to examine the variability in firing irregularity by a method that captures it along with the firing rate simultaneously in a manner that is more systematic than the ad hoc metrics used thus far.

The Bayesian method proposed by Koyama and Shinomoto (2005) is robust against rate fluctuations in its estimation of firing irregularity, but it assumes that the irregularity is constant throughout a sequence. In this study, we extended this framework so that variations in both the irregularity and the rate can be captured simultaneously and implemented using a practical algorithm that is applicable to data comprising a large number of spikes. We analyzed biological data by employing this new Bayesian method and revealed that the firing irregularity of individual neurons exhibits a certain systematic tendency to vary with the firing rate. In addition, we found that the degree of irregularity fluctuates less in cortical areas V1 and MT than in the lateral geniculate nucleus (LGN) of the thalamus. This indicates that the mechanism responsible for the firing events is autocontrolled against the rate modulation in the cortex (Shu, Hasenstaub, & McCormick, 2003; Haider, Duque, Hasenstaub, & McCormick, 2006; Berg, Alburda, & Hounsgaard, 2007; Miura, Tsubo, Okada, & Fukai, 2007; Roudi & Latham, 2008), but uncontrolled in the thalamus.

2 Bayesian Method for Estimating the Firing Irregularity

Throughout this study, we consider a spike train as a realization of a stochastic process characterized by two time-dependent parameters, the rate $\lambda(t)$ and the irregularity or regularity $\kappa(t)$. Given a single spike train, we estimate the two parameters $\lambda(t)$ and $\kappa(t)$ by inverting the conditional probability distribution with the Bayes theorem.

2.1 Spike Generation Model. First, we consider the renewal process in which interspike intervals (ISIs) derived independently from an identical distribution function are accumulated as a series of spike times $\{t_i\}$. We adopt here the gamma distribution function (Cox & Lewis, 1966), whose shape is parameterized by κ ,

$$f_{\kappa}(x) = \kappa(\kappa x)^{\kappa-1} e^{-\kappa x} / \Gamma(\kappa). \quad (2.1)$$

This $f_{\kappa}(x)$ is defined as a function of a dimensionless variable x , which makes the mean of x unity. The gamma distribution is frequently assumed

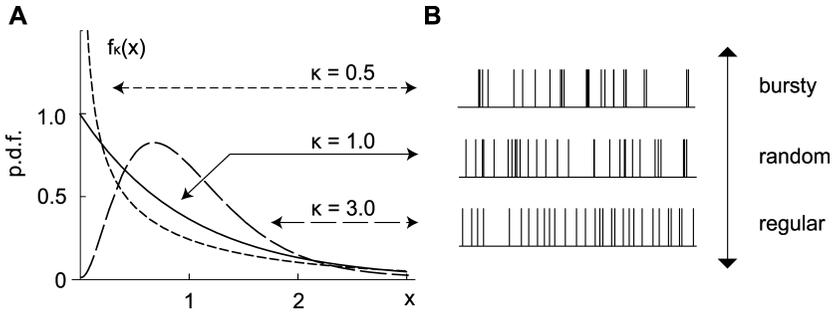


Figure 1: (A) Gamma distributions with $\kappa = 0.5, 1,$ and $3.$ (B) Sample sequences of spikes or ISIs derived from the distributions with identical rates, which may be called bursty, random, or regular.

as a model for neuronal ISI distributions, notably in the cortex and LGN (Kuffler, Fitzhugh, & Barlow, 1957; Stein, 1965; Troy & Robson, 1992; Teich, Heneghan, Lowen, Ozaki, & Kaplan, 1997; Baker & Lemon, 2000; Badoual, Rudolph, Piwkowska, Destexhe, & Bal, 2005). The gamma processes with the shape parameter κ larger than, equal to, and smaller than unity generate firing patterns termed regular, random (Poisson), and bursty, respectively (Softky & Koch, 1993; Holt, Softky, Koch, & Douglas, 1996; Shadlen & Newsome, 1998) (see Figure 1). Note that the following framework is generally applicable to any parameterized family of distributions, such as the inverse gaussian distribution family (Tuckwell, 1988) or the log-normal distribution family.

Next, we consider the case in which the rate and the regularity are modulated independently as explicit functions of time, $\lambda(t)$ and $\kappa(t)$. A rate-fluctuating gamma process could be constructed by rescaling the time (Berman, 1981; Reich, Victor, & Knight, 1998) of the renewal gamma process with a given time-dependent rate $\lambda(t)$. The conditional probability for a spike to occur at t_i , given that the preceding spike occurred at t_{i-1} , is obtained from

$$h(t_i | t_{i-1}; \{\lambda(t)\}, \{\kappa(t)\}) = \lambda(t_i) f_{\kappa(t_{i-1})}(\Lambda(t_i) - \Lambda(t_{i-1})), \tag{2.2}$$

where $\Lambda(t) \equiv \int_0^t \lambda(u) du$, and we have assumed that the regularity parameter κ does not vary during an ISI, in which there is no spike. The probability of the occurrence of spikes at times $\{t_i\}_{i=0}^n$ given the two time-dependent parameters, $\{\lambda(t)\}$ and $\{\kappa(t)\}$, is given by the product of the conditional probabilities (see equation 2.2),

$$p(\{t_i\}_{i=0}^n | \{\lambda(t)\}, \{\kappa(t)\}) = \prod_{i=1}^n h(t_i | t_{i-1}; \{\lambda(t)\}, \{\kappa(t)\}). \tag{2.3}$$

2.2 Empirical Bayesian Estimation. Second, we consider inverting the conditional distribution given in equation 2.3 to estimate the time-dependent parameters, $\lambda(t)$ and $\kappa(t)$, from a given sequence of spikes, $\{t_i\}_{i=0}^n$ using the Bayes formula,

$$\begin{aligned} p(\{\lambda(t)\}, \{\kappa(t)\} \mid \{t_i\}_{i=0}^n; \gamma_\lambda, \gamma_\kappa) \\ = \frac{p(\{t_i\}_{i=0}^n \mid \{\lambda(t)\}, \{\kappa(t)\}) p(\{\lambda(t)\}; \gamma_\lambda) p(\{\kappa(t)\}; \gamma_\kappa)}{p(\{t_i\}_{i=0}^n; \gamma_\lambda, \gamma_\kappa)}. \end{aligned} \quad (2.4)$$

We introduce the following gaussian process prior distributions of $\lambda(t)$ and $\kappa(t)$ so that both parameters are independently modulated and their large gradients are penalized:

$$p(\{\lambda(t)\}; \gamma_\lambda) = \frac{1}{Z(\gamma_\lambda)} \exp \left[-\frac{1}{2\gamma_\lambda^2} \int_0^T \left(\frac{d\lambda(t)}{dt} \right)^2 dt \right], \quad (2.5)$$

$$p(\{\kappa(t)\}; \gamma_\kappa) = \frac{1}{Z(\gamma_\kappa)} \exp \left[-\frac{1}{2\gamma_\kappa^2} \int_0^T \left(\frac{d\kappa(t)}{dt} \right)^2 dt \right], \quad (2.6)$$

where γ_λ and γ_κ are hyperparameters that represent the temporal gradients of rate and regularity, respectively (Rasmussen & Williams, 2006; Cunningham, Yu, Shenoy, & Sahani, 2008).

The hyperparameters γ_λ and γ_κ can be determined by maximizing the (marginal) likelihood (MacKay, 1992) for having a sequence of spikes $\{t_i\}_{i=0}^n$,

$$\begin{aligned} p(\{t_i\}_{i=0}^n; \gamma_\lambda, \gamma_\kappa) \equiv \int \int p(\{t_i\}_{i=0}^n \mid \{\lambda(t)\}, \{\kappa(t)\}) p(\{\lambda(t)\}; \gamma_\lambda) \\ \times p(\{\kappa(t)\}; \gamma_\kappa) d\{\lambda(t)\} d\{\kappa(t)\}. \end{aligned} \quad (2.7)$$

With the hyperparameters determined, we can obtain the maximum a posteriori (MAP) estimates of the rate $\lambda(t)$ and regularity $\kappa(t)$, so that their posterior distribution,

$$\begin{aligned} p(\{\lambda(t)\}, \{\kappa(t)\} \mid \{t_i\}_{i=0}^n; \gamma_\lambda, \gamma_\kappa) \\ \propto p(\{t_i\}_{i=0}^n \mid \{\lambda(t)\}, \{\kappa(t)\}) p(\{\lambda(t)\}; \gamma_\lambda) p(\{\kappa(t)\}; \gamma_\kappa), \end{aligned} \quad (2.8)$$

is maximized.

3 Numerical Method for Performing the Bayesian Estimation

In this section, we construct an algorithm that may render the empirical Bayes estimation practicable. Firstly, we employ the expectation

maximization (EM) method to maximize the marginalized likelihood function (see equation 2.7). Under the hyperparameters determined by the EM method, we next use the relaxation method to obtain the MAP estimates of the time-dependent rate and regularity.

3.1 Maximization of Marginal Likelihood. The hyperparameters $\gamma \equiv (\gamma_\lambda, \gamma_\kappa)^T$ for the degree of fluctuation in rate and regularity are determined by maximizing the marginal likelihood (see equation 2.7). We have implemented an algorithm for selecting the hyperparameters γ by employing the EM method and the Kalman filter under the assumption that the distributions of $\{\lambda(t)\}$ and $\{\kappa(t)\}$ are gaussian (Smith & Brown, 2003).

3.1.1 State-Space Model. Under the assumption that the parameters representing the rate and the regularity $\theta \equiv (\lambda, \kappa)^T$ do not vary greatly during each ISI in which there is no spike, we performed the marginalization path integral of equation 2.7 by discretizing the time coordinate for every occasion on which an event occurred. Thus, the occurrence of an event could be treated with the state-space model or the hidden Markov model, in which $\{\theta_i \equiv \theta(t_i)\}_{i=0}^{n-1}$ are the states or the hidden variables, and $\{T_i \equiv t_{i+1} - t_i\}_{i=0}^{n-1}$ are the observations. With the given prior distributions (see equations 2.5 and 2.6), a state transition can be represented as

$$p(\lambda_{j+1} | \lambda_j; \gamma_\lambda) = \frac{1}{\sqrt{2\pi \gamma_\lambda^2 T_j}} \exp \left[-\frac{(\lambda_{j+1} - \lambda_j)^2}{2\gamma_\lambda^2 T_j} \right], \tag{3.1}$$

$$p(\kappa_{j+1} | \kappa_j; \gamma_\kappa) = \frac{1}{\sqrt{2\pi \gamma_\kappa^2 T_j}} \exp \left[-\frac{(\kappa_{j+1} - \kappa_j)^2}{2\gamma_\kappa^2 T_j} \right], \tag{3.2}$$

or

$$p(\theta_{j+1} | \theta_j; \gamma) = \frac{1}{\sqrt{2\pi |Q_j|}} \exp \left[-\frac{1}{2} (\theta_{j+1} - \theta_j)^T Q_j^{-1} (\theta_{j+1} - \theta_j) \right], \tag{3.3}$$

$$Q_j = \begin{pmatrix} \gamma_\lambda^2 T_j & 0 \\ 0 & \gamma_\kappa^2 T_j \end{pmatrix}. \tag{3.4}$$

The corresponding state equation can be written as

$$\theta_{j+1} = \theta_j + w_j, \tag{3.5}$$

where w_j is the gaussian noise with the mean 0 and the variance-covariance matrix Q_j .

In this framework, the gamma process (see equation 2.2) corresponds to the observation equation, in which each interval T_j is observed under a

state $\theta_j = (\lambda_j, \kappa_j)^T$, as

$$p(T_j | \theta_j) = h(t_{j+1} | t_j; \lambda_j, \kappa_j) = \frac{\lambda_j^{\kappa_j} \kappa_j^{\kappa_j}}{\Gamma(\kappa_j)} T_j^{\kappa_j-1} \exp(-\lambda_j \kappa_j T_j). \tag{3.6}$$

3.1.2 EM Algorithm. The EM algorithm is used for finding the maximum likelihood estimate of parameters in a model that depends on latent variables (Dempster, Laird, & Rubin, 1977). In the model presented here, observable variables are the sequence of ISIs $\{T_i\}_{i=0}^{n-1}$, while latent variables are the rate and regularity parameters $\{\theta_i\}_{i=0}^{n-1}$. In the EM algorithm, the set of hyperparameters $\gamma \equiv (\gamma_\lambda, \gamma_\kappa)^T$ that determines the smoothness of the rate and regularity can be determined by iteratively maximizing the expected value of the log likelihood, the Q function,

$$Q(\gamma | \gamma^{(p)}) = E[\log p(\{T_j\}_{j=0}^{n-1}, \{\theta_j\}_{j=0}^{n-1}; \gamma) | \{T_i\}_{i=0}^{n-1}; \gamma^{(p)}], \tag{3.7}$$

$$= E \left[\sum_{j=0}^{n-2} \log p(\theta_{j+1} | \theta_j; \gamma) + \sum_{j=0}^{n-1} \log p(T_j; \theta_j) | \{T_i\}_{i=0}^{n-1}; \gamma^{(p)} \right], \tag{3.8}$$

where $\gamma^{(p)} = (\gamma_\lambda^{(p)}, \gamma_\kappa^{(p)})^T$ is the set of hyperparameters of the p th iteration. The $p + 1$ st γ is obtained by maximizing this Q function or, equivalently, the conditions for $dQ/d\gamma = 0$,

$$\gamma_\lambda^{(p+1)} = \frac{1}{n-1} \left\{ \sum_{j=0}^{n-2} \frac{1}{T_j} E[(\lambda_{j+1} - \lambda_j)^2 | \{T_i\}_{i=0}^{n-1}; \gamma^{(p)}] \right\}^{-1}, \tag{3.9}$$

$$\gamma_\kappa^{(p+1)} = \frac{1}{n-1} \left\{ \sum_{j=0}^{n-2} \frac{1}{T_j} E[(\kappa_{j+1} - \kappa_j)^2 | \{T_i\}_{i=0}^{n-1}; \gamma^{(p)}] \right\}^{-1}. \tag{3.10}$$

The expected values in equations 3.9 and 3.10 can be obtained using the state-space model (see equations 3.5 and 3.6) with the Kalman filter and the smoothing algorithm under the assumption that the distributions of $\{\lambda(t)\}$ and $\{\kappa(t)\}$ are gaussian. The details of the numerical algorithms are given in appendix A. The determination of hyperparameters is feasible with Kalman filtering, which requires computing the linear order of the number of data $O(n)$.

3.2 Derivation of the Exact MAP Paths. After determining the hyperparameters, we evaluate the MAP estimate of the time-dependent rate and regularity by the relaxation method (Press, Teukolsky, Vetterling, & Flannery, 1992; Nemenman & Bialek, 2002).

In order to maximize the posterior distribution (see equation 2.8), the rate and regularity should satisfy the following differential equations:

$$\frac{1}{\gamma_\lambda^2} \frac{d^2 \hat{\lambda}(t)}{dt^2} = \sum_{i=1}^n \left\{ \hat{\kappa}(t_{i-1}) - \frac{\hat{\kappa}(t_{i-1}) - 1}{\int_{t_{i-1}}^{t_i} \hat{\lambda}(u) du} \right\} I_{(t_{i-1}, t_i]}(t) - \sum_{i=1}^n \frac{\delta(t - t_i)}{\hat{\lambda}(t_i)}, \quad (3.11)$$

$$\begin{aligned} \frac{1}{\gamma_\kappa^2} \frac{d^2 \hat{\kappa}(t)}{dt^2} = & - \sum_{i=1}^n \left\{ 1 - \int_{t_{i-1}}^{t_i} \hat{\lambda}(u) du + \log \int_{t_{i-1}}^{t_i} \hat{\lambda}(u) du \right. \\ & \left. + \log \hat{\kappa}(t_{i-1}) - \psi(\hat{\kappa}(t_{i-1})) \right\} \delta(t - t_i), \end{aligned} \quad (3.12)$$

where $\psi(x)$ is the digamma function, and

$$I_{(t_{i-1}, t_i]}(t) = \begin{cases} 1 & \text{if } t \in (t_{i-1}, t_i], \\ 0 & \text{otherwise.} \end{cases} \quad (3.13)$$

Equations 3.11 and 3.12 can be converted into recurrence equations of state variables at the instance of the spike occurrence, and therefore the complexity of the entire computation is also of linear order of data size, $O(n)$. The practical numerical algorithm constructed on the basis of the relaxation method is given in appendix B.

4 Analysis of Spike Trains

4.1 Examination of the Bayesian Method with Simulation Data. In this section, we first examine the Bayesian method for its suitability in estimating the parameters of the simulation data. Figure 2A shows a process in which the firing activity was increased two times, while the mode of event occurrence was switched from irregular (low κ) to regular (high κ). Figure 2B is a raster diagram of the sequence of events (simulated spikes) generated under a time-dependent gamma process.

We used the Bayesian method for estimating the regularity and the rate for the single train of spikes shown in Figure 2B. The solid lines in Figure 2C represent the maximum a posteriori (MAP) estimates of the regularity $\hat{\kappa}(t)$ and the rate $\hat{\lambda}(t)$. The dashed lines indicate the 95% confidence intervals estimated using gaussian approximation (see appendix A). The Bayesian method successfully captured the transition of firing mode as well as the variation in firing activity.

Furthermore, the accuracy of inference was examined by the Kolmogorov-Smirnov (K-S) test (Brown, Barbieri, Ventura, Kass, & Frank,

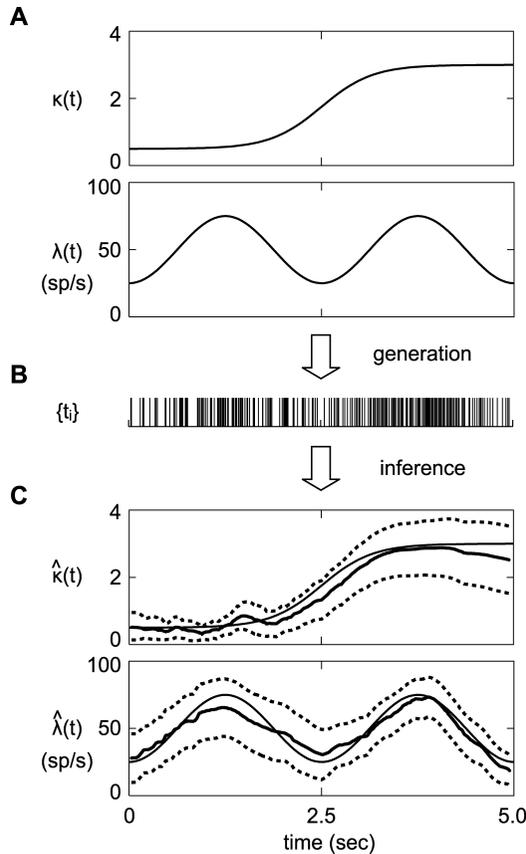


Figure 2: Inference of the regularity and the rate. (A) The regularity $\kappa(t)$ and the rate $\lambda(t)$ of a time-dependent gamma process ($\kappa(t) = 0.5 + 2.5/(1 + e^{-3(t-2.5)})$, $\lambda(t) = 50 + 25 \sin(4\pi t/5 - \pi/2)$). (B) Sample sequence of events derived from the gamma process of A. (C) The regularity $\hat{\kappa}(t)$ and the rate $\hat{\lambda}(t)$ estimated from the sequence of B. The solid and dashed lines represent the maximum a posteriori (MAP) estimates and the 95% confidence intervals, respectively.

2002; Kass, Ventura, & Brown, 2005). We estimated the time-dependent parameters and rescaled the sequences with the rate and regularity for 100 ISIs generated from the regularity and the rate, obeying the Ornstein-Uhlenbeck processes introduced to mimic the biological circumstances in which the rate and regularity fluctuate in time (the means, $\mu_\lambda = 50$ sp/s, $\mu_\kappa = 1.0$; the standard deviations, $\sigma_\lambda = 25$ sp/s, $\sigma_\kappa = 1.0$; fluctuation timescales $\tau_\lambda = \tau_\kappa = 600$ ms). Among the 10,000 time-dependent processes, 9,982 sequences satisfied the 95% confidence of the K-S test.

4.2 Analysis of Biological Data. Having confirmed the validity of the method with simulations in the preceding section, we applied the Bayesian estimation method to the biological spike data publicly available from the Neural Signal Archive (Bair & Movshon, 2004a, 2004b; Kohn & Movshon, 2003, 2004a, 2004b; Bair, Cavanaugh, Smith, & Movshon, 2002). The spike data were recorded from the visual cortical areas V1 and MT and the LGN of *Macaca fascicularis* anesthetized with sufentanil and paralyzed with vecuronium. The recordings from V1 and LGN were obtained while a drifting sinusoidal grating was presented (duration, 6000 or 3000 ms for V1 and 5138 ms for LGN). The recordings from MT were obtained while a drifting sinusoidal grating or a sparse dot pattern was presented (duration, 1280 or 1000 ms). (Further experimental details are provided online at <http://www.neuralsignal.org>.) In the analysis, we excluded spike sequences having a mean firing rate of less than 10 spikes per second due to the insufficiency of the data for analysis. Consequently, 44, 43, and 52 neurons were selected, each represented by approximately 15, 25, and 15 trials, containing approximately 250, 50, and 150 spikes for V1, MT, and LGN, respectively.

4.2.1 Analysis of a Single Neuron. The firing characteristics of a single neuron were analyzed in the following manner. Each spike sequence for a single trial was analyzed by the Bayesian method in order to obtain the MAP estimate for the regularity $\kappa(t)$ and the rate $\lambda(t)$ (see Figure 3A). Then they were mapped at every spike as a scatter diagram in the log-log $\{\lambda, \kappa\}$ plane (see Figure 3B). We repeated this mapping for a single neuron for a number of trials in order to obtain the distribution of the regularity and rate to characterize the variation in spiking characteristics within and across trials (see Figure 3C). The distribution of the scatter points was summarized into one ellipse representing a two-dimensional gaussian distribution fitted to the data.

4.2.2 Analysis of a Group of Neurons. The firing characteristics of a group of neurons can be summarized in one figure by plotting the ellipses representing the individual neurons for V1, MT, and LGN, respectively (see Figure 4A). It was observed from the inclined ellipsoids that there was a systematic dependency of regularity on the rate for individual neurons. The deviation from the regression line was generally small for neurons in cortical areas V1 and MT, as indicated by the flat ellipsoids, while the degree of regularity fluctuated greatly in the LGN, as indicated by the fat ellipsoids.

Note that the covariances of the logarithms of (λ, κ) in Figure 4 can be arbitrarily exaggerated by scaling the axes. Therefore, these distributions should be compared in a standardized metric multiplied with Fisher information (see appendix C), as

$$d_\lambda(t) = (\lambda(t) - \bar{\lambda})\sqrt{I_\lambda}, \quad (4.1)$$

$$d_\kappa(t) = (\kappa(t) - \bar{\kappa})\sqrt{I_\kappa}, \quad (4.2)$$

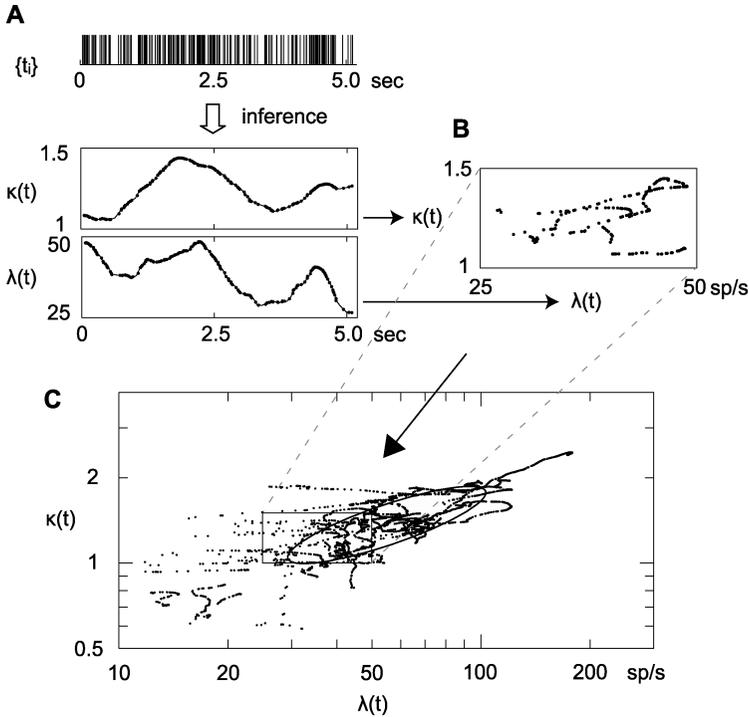


Figure 3: Summary of the spiking characteristics for a single V1 neuron. (A) MAP estimates of the regularity $\hat{\kappa}(t)$ and the rate $\hat{\lambda}(t)$ for a spike sequence of a single trial. (B) Scatter plot of the estimated $\hat{\lambda}$ and $\hat{\kappa}$ at every spike in the log-log plane of $\{\lambda, \kappa\}$. (C) Scatter plot for a single neuron within and across trials. An ellipse represents the $\frac{1}{2}$ quantile of a two-dimensional gaussian distribution fitted to the data.

where I_λ and I_κ denote the Fisher information of λ and κ , respectively. Figure 4B represents ellipses under these standardized metrics with the centers set together. The average slopes of the regression lines were, 0.25, 0.26, and 0.19 for V1, MT, and LGN, respectively. The ratios of the standard deviations of $d_\kappa(t)$ from the regression lines to the standard deviations of $d_\lambda(t)$ were 0.28, 0.37, and 0.99 for V1, MT, and LGN, respectively, thus affirming the simple impression given by Figure 4A, with the standardized information-geometric measures.

4.2.3 Covariation of the Regularity and the Rate. Figure 4 shows that regularity covaried positively with the rate for almost all neurons. It is noteworthy that these systematic correlations could be mimicked by adding a refractory period to the gamma distribution. The dashed lines in Figure 5

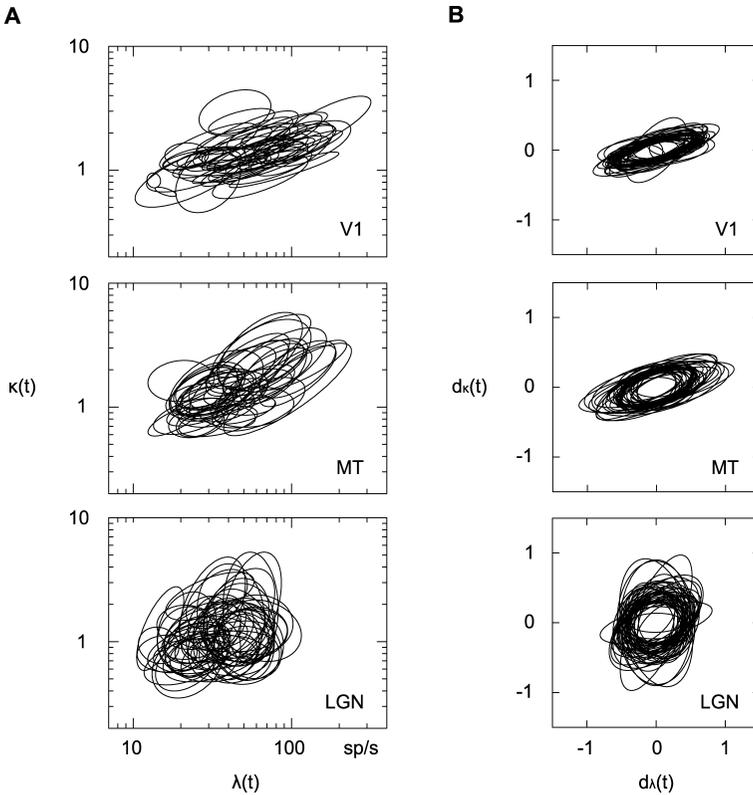


Figure 4: (A) Distributions of regularity κ and rate λ of all neurons in V1, MT, and LGN represented in a log-log plot. Each ellipse represents a single neuron, as in Figure 3C. (B) Covariations of the (λ, κ) represented on a standardized metric (see equations 4.1 and 4.2) for all neurons with the centers set together.

represent the dependence of regularity κ on rate λ , which was induced by introducing a refractory period of 2 ms; we numerically generated long spike trains using gamma processes adding the fixed refractory period and interpolated the resulting parameters generated under the same κ . These lines agreed well with the covariations of neurons in V1, MT, and LGN.

4.2.4 Goodness-of-Fit of the Spike Generation Model. Next, we examined the suitability of the time-dependent gamma process in representing the biological spike sequences by means of the K-S test applied to the data time rescaled with the estimated rate and regularity (Brown et al., 2002). In order to standardize the examination, we considered the first 100 ISIs of

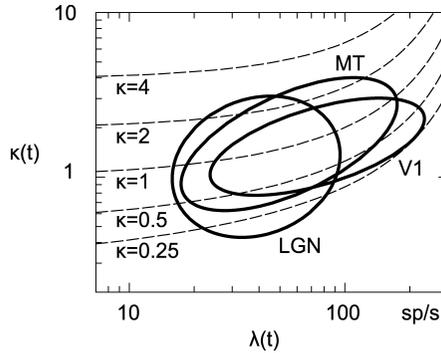


Figure 5: Three ellipses represent the $\frac{3}{4}$ quantiles of gaussian distributions for all neurons of V1, MT, and LGN. The dashed lines represent the gamma process of various shape parameters κ , accompanied by a refractory period of 2 ms.

each of the trials (509, 111, and 689 trials for V1, MT, and LGN, respectively). Trials that contained fewer than 100 spikes were excluded from the analysis. The time-dependent Poisson process may account for the small fractions of spike sequences (41%, 10%, and 13% were within the 95% confidence range for V1, MT, and LGN, respectively); however, the gamma process improves the fraction of acceptable data drastically (72%, 86%, and 35% for V1, MT, and LGN, respectively). Figure 6 represents the K-S plots for five sample spike sequences from V1, MT, and LGN.

5 Discussion

In this study, we developed a Bayesian method for simultaneously estimating the firing regularity or irregularity, and the firing rate for a sequence of spikes, and applied the method to biological spike data recorded from V1, MT, and LGN.

First, we found that the degree of firing irregularity does not fluctuate widely for individual neurons in cortical areas V1 and MT compared with those in the LGN of the thalamus. This result may be consistent with the assertion that excitatory and inhibitory inputs are balanced out in cortical networks (Shu et al., 2003; Haider et al., 2006; Berg et al., 2007; Miura et al., 2007; Roudi & Latham, 2008). It would be interesting to see whether the thalamus possesses mechanisms for providing the balanced inhibition.

Second, there was a subtle correlation between firing irregularity and the firing rate for individual neurons; firing tended to be more regular as the rate increased, which is consistent with several reports providing neurophysiological data (Troy & Robson, 1992; Kara, Reinagel, & Reid, 2000; Shinomoto et al., 2005; Mitchell, Sundberg, & Reynolds, 2007; Meier, Egert, Aertsen, & Nawrot, 2008). The facilitation of regularity may be mimicked

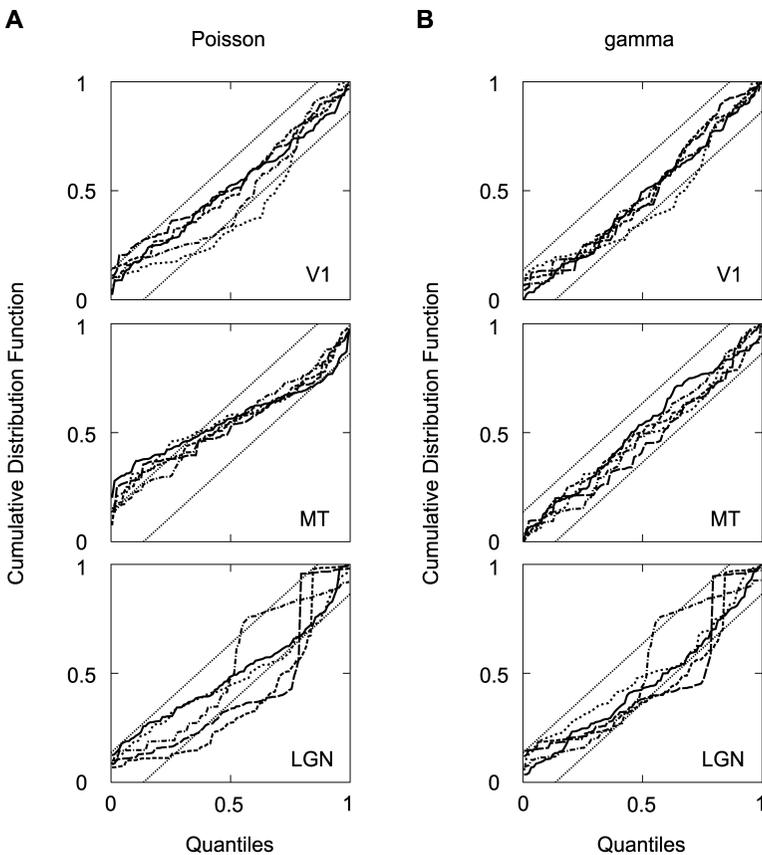


Figure 6: Kolmogorov-Smirnov (K-S) plot for five sample spike sequences of V1, MT, and LGN. Dotted lines represent the 95% confidence bounds. (A) K-S plot for the data fitted by inhomogeneous Poisson processes. (B) K-S plot for the same data fitted by the time-dependent gamma processes.

by introducing the refractory period into the point process (Johnson, 1996). We suggest that this dependence can be reproduced quantitatively by introducing a refractory period of 2 ms into the gamma distribution. The gamma distribution was originally adopted in the field of neuroscience as a parametric ISI distribution, which can suitably represent the relative refractory period that captures transient afterhyperpolarizing effects (Kuffler et al., 1957; Stein, 1965; Baker & Lemon, 2000). The fact that another dead time of 2 ms is required for reproducing the covariation might imply that this dead time represents an absolute refractory period, which captures sodium channel inactivation (Levitan & Kaczmarek, 2002) and does not covary with the firing rate (Nawrot et al., 2008).

Third, the Kolmogorov-Smirnov test revealed that the characterization of the biological spike trains can be improved by including non-Poisson characteristics specific to individual neurons. However, it should be noted that the gamma process was still unsuccessful in fitting the spike sequences of the LGN (35%), whereas the process fitted the spike sequences of the cortical areas V1 and MT, relatively well (72% and 86%, respectively). There is a scope for improvements to overcome the remaining percentage, such as the replacement of the gamma distribution with another family of ISI distributions. It would also be interesting to consider replacing the prior probabilities of rate and regularity given in equations 2.5 and 2.6 with ones that allow abrupt changes.

The Bayesian method of analysis developed in this letter is a general framework applicable to any type of point event, including earthquakes and economic events. The analysis of the instantaneous irregularity of the occurrence of the event may reveal hidden mechanisms underlying the event generation, such as a transition of active faults leading to earthquakes or a change in agents leading to economic events (Goh & Barabási, 2008). Similarly, a change in neuronal firing characteristics, if it were observed by the analysis, would indicate the occurrence of a drastic transition in the neuronal environmental conditions.

Appendix A: Kalman Filter and Smoothing Algorithm _____

In order to use the EM algorithm (equations 3.9 and 3.10), we require the conditional probability distribution given that the event intervals are $\{T_i\}_{i=0}^{n-1}$ and the hyperparameters are $\gamma^{(p)}$. Under the gaussian assumption, the conditional distribution can be obtained by Kalman filtering and smoothing applied to the mean, variance, and covariance of the distribution, defined as follows:

$$\theta_{j|l} \equiv E[\theta_j | \{T_i\}_{i=0}^l; \gamma^{(p)}], \quad (\text{A.1})$$

$$V_{j|l} \equiv E[\{\theta_j - \theta_{j|l}\}\{\theta_j - \theta_{j|l}\}^T | \{T_i\}_{i=0}^l; \gamma^{(p)}], \quad (\text{A.2})$$

$$V_{j,k|l} \equiv E[\{\theta_j - \theta_{j|l}\}\{\theta_k - \theta_{k|l}\}^T | \{T_i\}_{i=0}^l; \gamma^{(p)}]. \quad (\text{A.3})$$

- Prediction:

$$\theta_{j|j-1} = \theta_{j-1|j-1}, \quad (\text{A.4})$$

$$V_{j|j-1} = V_{j-1|j-1} + Q_{j-1}. \quad (\text{A.5})$$

- **Filtering:**

The algorithm recursively computes the posterior probability density (Mendel, 1995),

$$\begin{aligned}
 p(\theta_j | \{T_i\}_{i=0}^j) &= \frac{p(\theta_j | \{T_i\}_{i=0}^{j-1}) p(T_j | \theta_j, \{T_i\}_{i=0}^{j-1})}{p(T_j | \{T_i\}_{i=0}^{j-1})} \\
 &\propto p(\theta_j | \{T_i\}_{i=0}^{j-1}) p(T_j | \theta_j). \tag{A.6}
 \end{aligned}$$

If we assume $p(\theta_j | \{T_i\}_{i=0}^{j-1})$ to be gaussian, the log posterior probability density can be expressed by

$$\begin{aligned}
 \log p(\theta_j | \{T_i\}_{i=0}^j) &= \left[-\frac{1}{2}(\theta_j - \theta_{j|j-1})^T V_{j|j-1}^{-1}(\theta_j - \theta_{j|j-1}) \right] \\
 &\quad + \log p(T_j | \theta_j) + \text{const.} \tag{A.7}
 \end{aligned}$$

By further assuming the posterior probability density $p(\theta_j | \{T_i\}_{i=0}^j)$ to be gaussian, the mean $\theta_{j|j}$ and variance $V_{j|j}$ are given as the mode and negative inverse of the second derivative of the log posterior probability density (Tanner, 1996),

$$\frac{d}{d\theta_j} \log p(\theta_j | \{T_i\}_{i=0}^j) \Big|_{\theta_j = \theta_{j|j}} = 0, \tag{A.8}$$

$$V_{j|j} = \left[\frac{d^2}{d\theta_j^2} \log p(\theta_j | \{T_i\}_{i=0}^j) \Big|_{\theta_j = \theta_{j|j}} \right]^{-1}. \tag{A.9}$$

- **Fixed-interval smoothing algorithm (Ansley & Kohn, 1982):**

$$\theta_{i|n} = \theta_{i|i} + A_i(\theta_{i+1|n} - \theta_{i+1|i}), \tag{A.10}$$

$$V_{i|n} = V_{i|i} + A_i(V_{i+1|n} - V_{i+1|i})A_i^T, \tag{A.11}$$

where

$$A_i = V_{i|i} V_{i+1|i}^{-1}. \tag{A.12}$$

- **Covariance algorithm (de Jong & Mackinnon, 1988):**

$$V_{i+1,i|n} = A_i V_{i+1|n}. \tag{A.13}$$

From these, we may obtain the variances of equations 3.9 and 3.10:

$$\begin{aligned}
 E[(\lambda_{j+1} - \lambda_j)^2 | \{T_i\}_{i=0}^{n-1}; \mathcal{Y}^{(p)}] &= V_{j+1|n-1}^{(1,1)} - 2V_{j+1,j|n-1}^{(1,1)} + V_{j|n-1}^{(1,1)} \\
 &\quad - (\theta_{j+1|n-1}^{(1)} - \theta_{j|n-1}^{(1)})^2, \tag{A.14}
 \end{aligned}$$

$$\begin{aligned}
 E[(\kappa_{j+1} - \kappa_j)^2 | \{T_i\}_{i=0}^{n-1}; \mathcal{Y}^{(p)}] &= V_{j+1|n-1}^{(2,2)} - 2V_{j+1,j|n-1}^{(2,2)} + V_{j|n-1}^{(2,2)} \\
 &\quad - (\theta_{j+1|n-1}^{(2)} - \theta_{j|n-1}^{(2)})^2. \tag{A.15}
 \end{aligned}$$

Appendix B: Relaxation Method

The MAP path equations 3.11 and 3.12 for rate and regularity can be converted into the following recurrence equations of the state variables defined at the instance of the spike occurrence $\Theta_i \equiv (\lambda(t_i), \dot{\lambda}(t_i), \kappa(t_i), \dot{\kappa}(t_i))^T$,

$$\Theta_{i+1} = F_i(\Theta_i, \Theta_{i+1}), \quad (\text{B.1})$$

or, more explicitly,

$$\lambda(t_{i+1}) = \lambda(t_i) + \dot{\lambda}(t_i)T_i + \frac{1}{2}R(\Theta_i, \Theta_{i+1})T_i^2, \quad (\text{B.2})$$

$$\dot{\lambda}(t_{i+1}) = \dot{\lambda}(t_i) + R(\Theta_i, \Theta_{i+1})T_i - \frac{\gamma_\lambda^2}{\lambda(t_{i+1})}, \quad (\text{B.3})$$

$$\kappa(t_{i+1}) = \kappa(t_i) + \dot{\kappa}(t_i)T_i, \quad (\text{B.4})$$

$$\dot{\kappa}(t_{i+1}) = \dot{\kappa}(t_i) - U(\Theta_i, \Theta_{i+1}), \quad (\text{B.5})$$

where

$$R(\Theta_i, \Theta_{i+1}) \equiv \gamma_\lambda^2 \left\{ \kappa(t_i) - \frac{\kappa(t_i) - 1}{\Lambda(\Theta_i, \Theta_{i+1})} \right\}, \quad (\text{B.6})$$

$$U(\Theta_i, \Theta_{i+1}) \equiv \gamma_\kappa^2 \left\{ 1 - \Lambda(\Theta_i, \Theta_{i+1}) + \log \Lambda(\Theta_i, \Theta_{i+1}) + \log \kappa(t_i) - \psi(\kappa(t_i)) \right\}, \quad (\text{B.7})$$

$$\Lambda(\Theta_i, \Theta_{i+1}) \equiv \lambda(t_i)T_i + \frac{1}{6} \left\{ \dot{\lambda}(t_{i+1}) + 2\dot{\lambda}(t_i) + \frac{\gamma_\lambda^2}{\lambda(t_{i+1})} \right\} T_i^2. \quad (\text{B.8})$$

In principle, it is possible to solve the recurrence equations by the shooting method. For a train of a large number of events, however, it is difficult to explore the initial condition that does not lead to divergence.

Here we suggest a solution of the recurrence equation with the following relaxation method (Press et al., 1992) starting from the provisional solution obtained under the gaussian approximation with the Kalman filter and smoothing. We assume the provisional states $\{\tilde{\Theta}_i\}$ to be close to the true states $\{\hat{\Theta}_i\}$,

$$\hat{\Theta}_i = \tilde{\Theta}_i + \Delta\Theta_i, \quad (\text{B.9})$$

and derive the algorithm to diminish a deviation of the provisional state from the true solution. The deviation $\Delta\Theta_i$ satisfies the relation,

$$\nabla_i F_i \Delta\Theta_i + (\nabla_{i+1} F_i - 1) \Delta\Theta_{i+1} = E_i, \quad (\text{B.10})$$

where $E_i \equiv \tilde{\Theta}_{i+1} - F_i(\tilde{\Theta}_i, \tilde{\Theta}_{i+1})$ and $\nabla_j F_i \equiv \nabla_{\Theta_j} F_i(\Theta_i, \Theta_{i+1})|_{\Theta_i=\tilde{\Theta}_i, \Theta_{i+1}=\tilde{\Theta}_{i+1}}$.

This relation, equation B.10, can be summarized as

$$\begin{aligned}
 & \begin{bmatrix} \nabla_0 F_0 & \nabla_1 F_0 - 1 & 0 & \dots & 0 \\ 0 & \nabla_1 F_1 & \nabla_2 F_1 - 1 & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & \nabla_{n-1} F_{n-1} & \nabla_n F_{n-1} - 1 \end{bmatrix} \begin{bmatrix} \Delta\Theta_0 \\ \Delta\Theta_1 \\ \vdots \\ \Delta\Theta_n \end{bmatrix} \\
 & = \begin{bmatrix} E_0 \\ E_1 \\ \vdots \\ E_{n-1} \end{bmatrix}. \tag{B.11}
 \end{aligned}$$

As the matrix is a block diagonal matrix, the set of equations can be solved with the $O(n)$ computational complexity. We add the $\Delta\Theta_i$ obtained from this equation to revise the provisional solutions $\tilde{\Theta}_i$. The solution of the recurrence equations is obtained by repeating this relaxation algorithm of the $O(n)$ computational complexity.

If $\{\hat{\Theta}_i\}$ is obtained, the values for each interval $t \in (t_i, t_{i+1}]$ can be obtained by

$$\hat{\lambda}(t) = \hat{\lambda}(t_i) + \hat{\lambda}'(t_i)(t - t_i) + \frac{1}{2}R(\hat{\Theta}_i, \hat{\Theta}_{i+1})(t - t_i)^2, \tag{B.12}$$

$$\hat{\kappa}(t) = \hat{\kappa}(t_i) + \hat{\kappa}'(t_i)(t - t_i). \tag{B.13}$$

Appendix C: The Fisher Information Metric _____

We consider comparing the degree of variability between different quantities (λ, κ) by standardizing the metrics. The Fisher information may be used for standardizing metrics of parameters (Amari & Nagaoka, 2000) on the basis of the Cramér-Rao bound,

$$E[(\theta - \bar{\theta})^2] \geq \frac{1}{nI(\theta)}, \tag{C.1}$$

where n is the number of data. Metrics of different parameters such as λ or κ may be standardized by multiplying them by the square root of the Fisher information, given as $\Delta\lambda\sqrt{I_\lambda}$ or $\Delta\kappa\sqrt{I_\kappa}$.

Parameters λ and κ of the gamma distribution can be treated as independent parameters, as they span orthogonally in the manifold of information geometry, or diagonalize the Fisher information matrix (Ikeda, 2005; Miura,

Okada, & Amari, 2006) as follows:

$$I(\theta) = - \int \frac{\partial^2 \log p(T; \theta)}{\partial \theta^2} p(T; \theta) dT = \begin{pmatrix} I_\lambda & 0 \\ 0 & I_\kappa \end{pmatrix}, \quad (\text{C.2})$$

where

$$I_\lambda = \frac{\kappa}{\lambda^2}, \quad I_\kappa = \psi'(\kappa) - \frac{1}{\kappa}, \quad (\text{C.3})$$

where $\psi'(\kappa)$ is a trigamma function.

Acknowledgments

We thank Shinsuke Koyama and Rob Kass for stimulating discussion. We are also indebted to W. Bair, A. Kohn, and J. A. Movshon for uploading their experimental data to the Neural Signal Archive. This study was supported in part by Grants-in-Aid for Scientific Research to S.S. from the MEXT Japan (20300083, 20020012). T.S. is supported by the Research Fellowship of the JSPS for Young Scientists.

References

- Amari, S., & Nagaoka, H. (2000). *Methods of information geometry*. New York: Oxford University Press.
- Ansley, C. F., & Kohn, R. (1982). A geometrical derivation of the fixed interval smoothing algorithm. *Biometrika*, *69*(2), 486–487.
- Badoual, M., Rudolph, M., Piwkowska, Z., Destexhe, A., & Bal, T. (2005). High discharge variability in neurons driven by current noise. *Neurocomputing*, *65*–*66*, 493–498.
- Bair, W., Cavanaugh, J. R., Smith, M. A., & Movshon, J. A. (2002). The timing of response onset and offset in macaque visual neurons. *Journal of Neuroscience*, *22*(8), 3189–3205.
- Bair, W., & Movshon, J. A. (2004a). Response vs. temporal frequency of drifting sinusoidal gratings for V1 complex direction selective cells. *Neural Signal Archive*, nsa2004.4. Available online at <http://www.neuralsignal.org>.
- Bair, W., & Movshon, J. A. (2004b). Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *Journal of Neuroscience*, *24*(33), 7305–7323.
- Baker, S. N., & Lemon, R. N. (2000). Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance levels. *Journal of Neurophysiology*, *84*(4), 1770–1780.
- Berg, R. W., Alaburda, A., & Hounsgaard, J. (2007). Balanced inhibition and excitation drive spike activity in spinal half-centers. *Science*, *315*(5810), 390–393.

- Berman, M. (1981). Inhomogeneous and modulated gamma processes. *Biometrika*, 68(1), 143–152.
- Brown, E. N., Barbieri, R., Ventura, V., Kass, R. E., & Frank, L. M. (2002). The time-rescaling theorem and its application to neural spike train data analysis. *Neural Comput.*, 14(2), 325–346.
- Cox, D. R., & Lewis, P. A. W. (1966). *The statistical analysis of series of events*. New York: Wiley.
- Cunningham, J. P., Yu, B. M., Shenoy, K. V., & Sahani, M. (2008). Inferring neural firing rates from spike trains using gaussian processes. In J. C. Platt, D. Koller, Y. Singer, & S. Rowers (Eds.), *Advances in neural information processing systems*, 20 (pp. 329–336). Cambridge, MA: MIT Press.
- Davies, R. M., Gerstein, G. L., & Baker, S. N. (2006). Measurement of time-dependent changes in the irregularity of neural spiking. *Journal of Neurophysiology*, 96(2), 906–918.
- Dayan, P., & Abbott, L. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.
- de Jong, P., & Mackinnon, M. J. (1988). Covariances for smoothed estimates in state space models. *Biometrika*, 75(3), 601–602.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1), 1–38.
- Gabbiani, F., & Koch, C. (1998). *Methods in neuronal modeling: From ions to networks* (2nd ed.). Cambridge, MA: MIT Press.
- Goh, K. I., & Barabási, A. L. (2008). Burstiness and memory in complex systems. *Europhysics Letters*, 81, 48002 (1–5).
- Haider, B., Duque, A., Hasenstaub, A. R., & McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *Journal of Neuroscience*, 26(17), 4535–4545.
- Holt, G. R., Softky, W. R., Koch, C., & Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of Neurophysiology*, 75(5), 1806–1814.
- Ikeda, K. (2005). Information geometry of interspike intervals in spiking neurons. *Neural Comput.*, 17(12), 2719–2735.
- Johnson, D. H. (1996). Point process models of single-neuron discharges. *Journal of Computational Neuroscience*, 3(4), 275–299.
- Kara, P., Reinagel, P., & Reid, R. C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, 27(3), 635–646.
- Kass, R. E., Ventura, V., & Brown, E. N. (2005). Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology*, 94(1), 8–25.
- Kohn, A., & Movshon, J. A. (2003). Neuronal adaptation to visual motion in area MT of the macaque. *Neuron*, 39(4), 681–691.
- Kohn, A., & Movshon, J. A. (2004a). Contrast response and direction tuning in macaque MT/V5. *Neural Signal Archive*, nsa2004.6. Available online at <http://www.neuralsignal.org>.
- Kohn, A., & Movshon, J. A. (2004b). LGN responses to drifting sinusoidal gratings at various temporal frequencies. *Neural Signal Archive*, nsa2004.3. Available online at <http://www.neuralsignal.org>.

- Kostal, L., & Lansky, P. (2006). Classification of stationary neuronal activity according to its information rate. *Network: Computation in Neural Systems*, 17(2), 193–210.
- Koyama, S., & Shinomoto, S. (2005). Empirical Bayes interpretations of random point events. *Journal of Physics A: Mathematical and Theoretical*, 38(29), L531–L537.
- Kuffler, S. W., Fitzhugh, R., & Barlow, H. B. (1957). Maintained activity in the cat's retina in light and darkness. *Journal of General Physiology*, 40(5), 683–702.
- Levitin, I. B., & Kaczmarek, L. K. (2002). *The neuron cell and molecular biology* (3rd ed.). Oxford: Oxford University Press.
- MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation*, 4(3), 415–447.
- Meier, R., Egert, U., Aertsen, A., & Nawrot, M. P. (2008). FIND—A unified framework for neural data analysis. *Neural Networks*, 21(8), 1085–1093.
- Mendel, J. M. (1995). *Lessons in estimation theory for signal processing, communications, and control*. Upper Saddle River, NJ: Prentice Hall.
- Mitchell, J. F., Sundberg, K. A., & Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*, 55(1), 131–141.
- Miura, K., Okada, M., & Amari, S. (2006). Estimating spiking irregularities under changing environments. *Neural Comput.*, 18(10), 2359–2386.
- Miura, K., Tsubo, Y., Okada, M., & Fukai, T. (2007). Balanced excitatory and inhibitory inputs to cortical neurons decouple firing irregularity from rate modulations. *Journal of Neuroscience*, 27(50), 13802–13812.
- Nawrot, M. P., Boucsein, C., Rodriguez-Molina, V., Riehle, A., Aertsen, A., & Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, 169(2), 374–390.
- Nemenman, I., & Bialek, W. (2002). Occam factors and model independent Bayesian learning of continuous distributions. *Physical Review E*, 65(2), 026137.
- Pillow, J. W., Paninski, L., Uzzell, V. J., Simoncelli, E. P., & Chichilnisky, E. J. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *Journal of Neuroscience*, 25(47), 11003–11013.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in C: The art of scientific computing* (2nd ed.). Cambridge: Cambridge University Press.
- Rasmussen, C., & Williams, C. (2006). *Gaussian processes for machine learning*. Cambridge, MA: MIT Press.
- Reich, D. S., Victor, J. D., & Knight, B. W. (1998). The power ratio and the interval map: Spiking models and extracellular recordings. *Journal of Neuroscience*, 18(23), 10090–10104.
- Richmond, B. J., & Optican, L. M. (1990). Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. II. Information transmission. *Journal of Neurophysiology*, 64(2), 3021–3033.
- Roudi, Y., & Latham, P. E. (2008). A balanced memory network. *PLoS Computational Biology*, 3, e141.
- Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4(4), 569–579.
- Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18(10), 3870–3896.

- Shinomoto, S., Miyazaki, Y., Tamura, H., & Fujita, I. (2005). Regional and laminar differences in in vivo firing patterns of primate cortical neurons. *Journal of Neurophysiology*, *94*(1), 567–575.
- Shinomoto, S., Shima, K., & Tanji, J. (2003). Differences in spiking patterns among cortical neurons. *Neural Comput.*, *15*(12), 2823–2842.
- Shu, Y., Hasenstaub, A., & McCormick, D. A. (2003). Turning on and off recurrent balanced cortical activity. *Nature*, *423*(6937), 288–293.
- Smith, A. C., & Brown, E. N. (2003). Estimating a state-space model from point process observations. *Neural Comput.*, *15*(5), 965–991.
- Softky, W. R., & Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience*, *13*(1), 334–350.
- Stein, R. B. (1965). A theoretical analysis of neuronal variability. *Biophysical Journal*, *5*(2), 173–194.
- Tanner, M. A. (1996). *Tools for statistical inference*. New York: Springer-Verlag.
- Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T., & Kaplan, E. (1997). Fractal character of the neural spike train in the visual system of the cat. *Journal of the Optical Society of America*, *14*(3), 529–546.
- Troy, J. B., & Robson, J. G. (1992). Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance. *Visual Neuroscience*, *9*(6), 535–553.
- Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology* (Vol. 2). Cambridge: Cambridge University Press.