Empirical Bayes interpretations of random point events

2005 J. Phys. A: Math. Gen. 38 L531

(http://iopscience.iop.org/0305-4470/38/29/L04)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **38** (2005) L531–L537 doi:10.1088/0305-4470/38/29/L04

LETTER TO THE EDITOR

Empirical Bayes interpretations of random point events

Shinsuke Koyama and Shigeru Shinomoto

Department of Physics, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

Received 8 April 2005, in final form 9 June 2005
Published 6 July 2005
Online at stacks.iop.org/JPhysA/38/L531

Abstract

Given a sequence of apparently random point events, such as neuronal spikes, one may interpret them as being derived either irregularly in time from a constant rate or regularly from a fluctuating rate. To determine which interpretation is more plausible in any given case, we employ the empirical Bayes method. For a sequence of point events derived from a rate-fluctuating gamma process, the marginal likelihood function can possess two local maxima and the system exhibits a first-order phase transition representing the switch of the most plausible interpretation from one to the other.

PACS numbers: 02.70.Rr, 02.50.Tt, 02.50.−r, 05.70.Fh

Accurate statistical descriptions of neuronal spike sequences are essential for extracting underlying information about the brain [1–4]. It had been believed that *in vivo* cortical neurons are Poisson spike generators [5–7]. This belief was due to the apparent randomness of neuronal spike events, the exponential distribution of inter-spike intervals and the fact that the values of the coefficient of variation are close to unity. However, a recent analysis using a newly introduced measure of the local variation of inter-spike intervals revealed that *in vivo* spike sequences are not uniformly random but, rather, possess statistical characteristics specific to individual neurons [8]. The neocortex consists of heterogeneous neurons that differ not only from one cortical area to another, but also from one layer to another in their spiking patterns [9].

The main reason that different spiking characteristics had been incorrectly interpreted as a simple Poisson process is that most analyses were based on the assumption of the stationarity, or constancy, of the rate. The new measure of the local variation of inter-spike intervals is efficient in classifying individual neurons, robustly against rate fluctuation [10]. But it is nevertheless desirable to formulate a systematic method of simultaneously characterizing not only the intrinsic irregularity but also the fluctuating rate of an individual spike train. Here, we examine the empirical Bayes method [11–15] on its applicability to non-stationary point processes.

In particular, we first generate sequences of point events using the rate-fluctuating gamma process, and then apply the empirical Bayes interpretation on each sequence. It is found that the 'marginal likelihood function' or the 'evidence', whose logarithm corresponds to the negative free energy in equilibrium thermodynamics, can possess two maxima in the space of the hyperparameters that represent the degree of spiking irregularity and the magnitude of the rate fluctuation. As the rate fluctuation of the underlying gamma process is increased, the marginal likelihood function obtained from the empirical Bayes method exhibits a first-order phase transition corresponding to the switch of the most plausible interpretation from (I) *irregularly derived from a nearly constant rate* to (II) *rather regularly derived from a significantly fluctuating rate*.

Rate-fluctuating gamma process

First, we consider point events (or spikes) occurring along the time axis according to the renewal process with a given inter-event interval distribution. In the present letter, we employ the gamma distribution function,

$$f_\kappa(x) = \kappa(\kappa x)^{\kappa-1}\,\mathrm{e}^{-\kappa x}/\,\Gamma(\kappa), \tag{1}$$

where $\Gamma(\kappa) \equiv \int_0^\infty x^{\kappa-1}\,\mathrm{e}^{-x}\,\mathrm{d}x$ is the gamma function. This $f_\kappa(x)$ is defined as a function of a dimensionless variable x, which makes the mean of x unity, independent of the parameter κ.

A rate-fluctuating gamma process could be constructed by rescaling the time [16, 17] of the renewal gamma process with a given time-dependent rate $\lambda(t)$ as

$$\Lambda(t) \equiv \int_0^t \lambda(u)\,\mathrm{d}u. \tag{2}$$

Then, the conditional probability for a spike to occur at t_i, given that the preceding spike occurred at t_{i-1}, is given by

$$r_\kappa(t_i|t_{i-1};\{\lambda(t)\}) = \lambda(t_i)g_\kappa(\Lambda(t_i)|\Lambda(t_{i-1})), \tag{3}$$

where

$$g_\kappa(z|y) \equiv \frac{f_\kappa(z-y)}{1-\int_y^z f_\kappa(x-y)\,\mathrm{d}x} \tag{4}$$

is the conditional intensity function or the 'hazard function' [18].

From this, we find that the probability density for spikes to occur at $\{t_i\}_{i=0}^n = \{t_0, t_1, \ldots, t_n\}$ for a given time-dependent rate $\lambda(t)$ is

$$p_\kappa\left(\{t_i\}_{i=0}^n\middle|\{\lambda(t)\}\right) = \left[\prod_{i=1}^n r_\kappa(t_i|t_{i-1};\{\lambda(t)\})\right]\exp\left(-\int_0^T r_\kappa(u|t_{N(u)};\{\lambda(t)\})\,\mathrm{d}u\right), \tag{5}$$

where the exponential of the integral on the rhs is the survivor function representing the product of the probabilities with which no spikes occur in the inter-spike intervals and $N(t)$ is the total number of spikes in the interval $(0, t]$ for $t \in (0, T]$ [19]. This rate-fluctuating gamma process is a natural extension of both the time-dependent Poisson process ($\kappa = 1$) and the renewal gamma process (for which $\lambda(t)$ is constant).

Empirical Bayes method

For a sequence of point events $\{t_i\}_{i=0}^n$ derived from the rate-fluctuating gamma process defined by equation (5), we apply the Bayes method for the inference of the time-dependent rate $\lambda(t)$. As the prior distribution of $\lambda(t)$, we introduce here a tendency to flatness in the form

$$p_\beta(\{\lambda(t)\}) = \frac{1}{Z(\beta)} \exp\left[-\beta \int_0^T \left(\frac{\mathrm{d}\lambda}{\mathrm{d}t}\right)^2 \mathrm{d}t\right], \tag{6}$$

where the hyperparameter β represents the stiffness of the rate fluctuation and $Z(\beta)$ is a normalization constant.

The posterior distribution of $\lambda(t)$ for a given set of data $\{t_i\}_{i=0}^n$ is obtained using the Bayes formula as

$$p_{\kappa,\beta}\left(\{\lambda(t)\}|\{t_i\}_{i=0}^n\right) = \frac{p_\kappa\left(\{t_i\}_{i=0}^n\middle|\{\lambda(t)\}\right) p_\beta(\{\lambda(t)\})}{p_{\kappa,\beta}\left(\{t_i\}_{i=0}^n\right)}, \tag{7}$$

where $p_{\kappa,\beta}\left(\{t_i\}_{i=0}^n\right)$ is the 'marginal likelihood function' or the 'evidence' for the hyperparameters κ and β, with the given data $\{t_i\}_{i=0}^n$:

$$p_{\kappa,\beta}\left(\{t_i\}_{i=0}^n\right) = \int p_\kappa\left(\{t_i\}_{i=0}^n\middle|\{\lambda(t)\}\right) p_\beta(\{\lambda(t)\})\, \mathrm{d}\{\lambda(t)\}. \tag{8}$$

The integration here is a functional integration over $\lambda(t)$. According to the empirical Bayes theory, the hyperparameters $\hat{\kappa}$ and $\hat{\beta}$ can be determined by maximizing the marginal likelihood function [11–15]. The log marginal likelihood function corresponds to the negative free energy in equilibrium thermodynamics.

By applying the variational method to the log posterior distribution

$$\log p_{\hat{\kappa},\hat{\beta}}\left(\{\lambda(t)\}|\{t_i\}_{i=0}^n\right) \propto \log p_{\hat{\kappa}}\left(\{t_i\}_{i=0}^n\middle|\{\lambda(t)\}\right) p_{\hat{\beta}}(\{\lambda(t)\})$$
$$= \sum_{i=1}^n \log r_{\hat{\kappa}}(t_i|t_{i-1}; \{\lambda(t)\}) - \int_0^T r_{\hat{\kappa}}(u|t_{N(u)}; \{\lambda(t)\})\, \mathrm{d}u - \hat{\beta} \int_0^T (\mathrm{d}\lambda/\mathrm{d}t)^2\, \mathrm{d}t, \tag{9}$$

the maximum *a posteriori* (MAP) estimate $\hat{\lambda}(t)$ is found to satisfy the integro-differential equation

$$2\hat{\beta}\frac{\mathrm{d}^2\hat{\lambda}}{\mathrm{d}t^2} = \frac{r_{\hat{\kappa}}(t|t_{N(t)}; \{\hat{\lambda}(t)\})}{\hat{\lambda}(t)} - \sum_{i=1}^n \frac{\delta(t-t_i)}{\hat{\lambda}(t_i)}, \tag{10}$$

where $\delta(t)$ denotes the Dirac delta function. The term $r_{\hat{\kappa}}(t|t_{N(t)}; \{\hat{\lambda}(t)\})$, defined by equations (3) and (2), contains the integration over $\hat{\lambda}(t)$. We recently realized that in [20, 21], a differential equation similar to the present integro-differential equation is derived using the saddle point approximation of quantum field theory.

Our method of analysis can be summarized as follows. First, for a given spike train, the optimal hyperparameters $\hat{\kappa}$ and $\hat{\beta}$ are determined by maximizing the marginal likelihood function, equation (8). This maximization can be carried out with the expectation maximization (EM) algorithm under the assumption that the distribution of $\lambda(t)$ is Gaussian [22, 23]. Second, the integro-differential equation (10) is solved numerically with the hyperparameters $\hat{\kappa}$ and $\hat{\beta}$ to obtain the MAP estimate of the rate of occurrence $\hat{\lambda}(t)$.

Data analysis

Here, we apply our method of analysis to sequences of point events derived from rate-fluctuating gamma processes to observe how such data are interpreted according to the

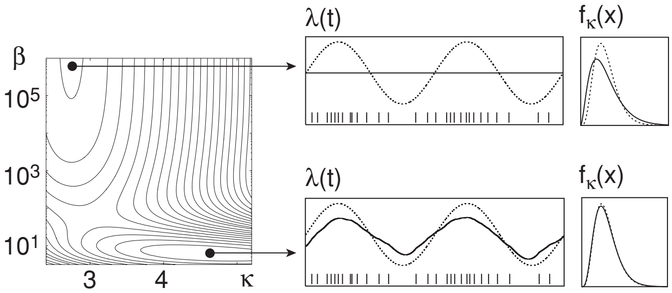

Figure 1. Left: a contour plot of the log marginal likelihood function in the space of the hyperparameters κ and β for a sequence of 10^5 point events derived from the rate-fluctuating gamma process with $\kappa = 5, \sigma/\mu = 0.6$ and $\tau\mu = 2.2$. Right: two interpretations obtained for the same sequence of point events. The dotted lines represent the original underlying rate $\lambda(t)$ and the original gamma distribution with $\kappa = 5$. The solid lines represent the inferred rate $\hat{\lambda}(t)$ and the inferred gamma distribution of $\hat{\kappa}$.

empirical Bayes method. In the present study, we consider a rate of occurrence $\lambda(t)$ that is regulated sinusoidally in time as

$$\lambda(t) = \mu + \sigma \sin \frac{t}{\tau}. \tag{11}$$

In the case that $\lambda(t) \leqslant 0$, which occurs for $|\sigma| \geqslant \mu$, we stipulate that point events are not generated.

This process is characterized by two independent dimensionless parameters, σ/μ and $\tau\mu$, respectively representing the amplitude and the time scale of the rate fluctuation. For each set of values of these dimensionless parameters, we derived a sequence of 10^5 point events, from which we numerically computed the marginal likelihood function, equation (8).

Figure 1 presents a contour plot of the log marginal likelihood function obtained for a sequence of point events generated from the rate-fluctuating gamma process. It is seen that the (log) marginal likelihood function possesses two local maxima, which correspond to the following two interpretations:

(I) The point events were derived irregularly in time ($\hat{\kappa}$ small) from a nearly constant rate ($\hat{\beta}$ large).

(II) The point events were derived rather regularly ($\hat{\kappa}$ large) from a significantly fluctuating rate ($\hat{\beta}$ small).

The log marginal likelihood function corresponds to a negative free energy. In the case that there exist multiple local minima of the free energy, the lowest minimum is chosen, as in equilibrium thermodynamics. When the amplitude, σ/μ, or the time scale, $\tau\mu$, of the underlying rate fluctuation is increased, the system exhibits a first-order phase transition corresponding to an abrupt switch of the most plausible interpretation from (I) to (II).

Figure 2(a) presents phase diagrams in the space of the original parameters, σ/μ and $\tau\mu$, of three kinds of event generating processes characterized by different intrinsic spiking (ir)regularities: $\kappa = 5$ (quasi-regular), $\kappa = 1$ (random) and $\kappa = 0.5$ (clumpy-bursty). It is seen that the parameter space is divided into four characteristic regions, in which the marginal likelihood function has (A) a single maximum, corresponding to a nearly constant rate (stiffness $\hat{\beta}$ large), (B) two maxima, corresponding to a nearly constant rate ($\hat{\beta}$ large) and a significantly fluctuating rate ($\hat{\beta}$ small), with the former representing a higher marginal likelihood, (C) two such maxima, with the latter representing a higher marginal likelihood, and (D) a single maximum, corresponding to a significantly fluctuating rate ($\hat{\beta}$ small).

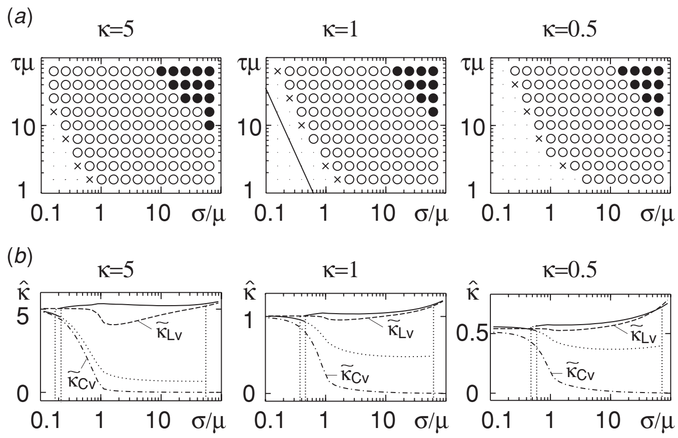

Figure 2. (*a*) Phase diagrams for sequences generated by rate-fluctuating gamma processes with three values of the intrinsic spiking regularity $\kappa = 5, \kappa = 1$ and $\kappa = 0.5$. Four characteristic phases A, B, C and D explained in the text are represented by dots, crosses, open circles and filled circles, respectively. The solid line in the figure for $\kappa = 1$ represents the phase boundary for which the optimal bin width in the time histogram method diverges, $\sigma^2\tau/\mu = \pi/8$. (*b*) The empirical Bayes estimates of the coefficient of intrinsic spiking regularity, $\hat{\kappa}$, for various values of σ/μ, with fixed $\tau\mu = 10$. The coefficient that gives a higher/lower marginal likelihood is represented by a solid/dotted line. The coefficient estimated from the values of the local variation, $\tilde{\kappa}_{L_\mathrm{V}} = 3/(2L_\mathrm{V}) - 1/2$, is represented by a dashed line. The coefficient estimated from the values of the coefficient of variation, $\tilde{\kappa}_{C_\mathrm{V}} = 1/(C_\mathrm{V})^2$, is represented by a dot-dashed line.

Figure 2(*b*) depicts how the estimated intrinsic regularity, $\hat{\kappa}$, is affected by the underlying rate fluctuation. It is interesting to compare this with the value of the local variation,

$$L_\mathrm{V} \equiv \frac{3}{n-1}\sum_{i=1}^{n}\frac{(T_i - T_{i+1})^2}{(T_i + T_{i+1})^2}, \tag{12}$$

computed for the sequence of inter-event intervals $\{T_i \equiv t_i - t_{i-1}\}_{i=1}^{n}$. It is shown in [8] that L_V corresponds to the value $3/(2\kappa + 1)$ for a renewal gamma process of κ. Therefore, the intrinsic regularity κ may be inferred from the L_V estimate as

$$\tilde{\kappa}_{L_\mathrm{V}} = 3/(2L_\mathrm{V}) - 1/2. \tag{13}$$

The intrinsic regularity κ can also be inferred from the coefficient of variation [18], which is defined as the ratio of the standard deviation of inter-spike intervals ΔT to the mean $\bar{T}$, as

$$\tilde{\kappa}_{C_\mathrm{V}} = 1/(C_\mathrm{V})^2 = \bar{T}^2/(\Delta T)^2. \tag{14}$$

It is seen from figure 2(*b*) that the empirical Bayes estimate $\hat{\kappa}$ and the L_V estimate $\tilde{\kappa}_{L_\mathrm{V}}$ indicate values comparably close to the original coefficient $\kappa = 5, 1$ or 0.5. In contrast, the C_V estimate $\tilde{\kappa}_{C_\mathrm{V}}$ does not effectively extract the original intrinsic regularity from rate-fluctuating spike trains.

In addition to allowing an estimation of the intrinsic regularity, the Bayes method has the further advantage over that based on the coefficient of local variation L_V that it simultaneously yields an estimation of the fluctuating rate. It should be noted, however, that the empirical Bayes method does not necessarily infer a time-dependent rate, even for data derived from rate-fluctuating gamma processes, as in the parameter regions (A) and (B) in figure 2(*a*).

Similar phenomena are observed with the time histogram method. In that case, the optimal bin width, which is determined by finding the best fit of the time histogram to the underlying rate of occurrence, can diverge, in what constitutes a second-order phase transition [24]. For the sinusoidally regulated Poisson processes, it was proven that the optimal bin width diverges in the region satisfying $\sigma^2\tau/\mu < \pi/8$. We plot this second-order transition line in the Bayes phase diagram for sinusoidally regulated Poisson processes, $\kappa = 1$, in figure 2(a).

In the present letter, we have applied the empirical Bayes method to sequences of point events derived from rate-fluctuating gamma processes, in particular to the case in which the rate $\lambda(t)$ is regulated sinusoidally. We have also examined the case in which the rate is driven by the Ornstein–Uhlenbeck process and found that the results are qualitatively the same. Furthermore, we have examined the robustness of the present method of analysis against noises. Namely, we added a Gaussian white noise to the original time-dependent rate and derived an event sequence from the noisy rate. We have observed that the estimated irregularity and rate are not significantly altered by the noise. A slight change due to the noise is that the local maximum of the marginal likelihood corresponding to the interpretation 'irregularly derived from a constant rate' becomes a little larger.

In all the cases, there is a first-order phase transition which corresponds to the abrupt switch of the most plausible interpretation derived from the empirical Bayes method from (I) *irregularly derived from a constant rate* to (II) *regularly derived from a fluctuating rate*.

The present analysis is based on the assumption that the intrinsic regularity characterized by the inter-event interval distribution, $f_\kappa(x)$, does not change in time. It would be interesting to construct a Bayesian framework applicable to point processes in which both the rate and the intrinsic regularity fluctuate in time. We are currently working on the application of the method to the analysis of real biological data to determine whether the assumption of fixed intrinsic regularity is plausible.

Acknowledgments

We are grateful to Yukito Iba for his advice concerning the empirical Bayes method. This study is supported in part by Grants-in-Aid for Scientific Research to SS from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by the 21st century COE 'Center for Diversity and Universality in Physics'.

References

[1] Tuckwell H C 1988 *Introduction to Theoretical Neurobiology* vol 2 (Cambridge: Cambridge University Press)
[2] Rieke F *et al* 1997 *Spikes: Exploring the Neural Code* (Cambridge, MA: MIT Press)
[3] Gabbiani F and Koch C 1998 *Methods in Neuronal Modeling: From Ions to Networks* 2nd edn (Cambridge, MA: MIT Press)
[4] Shinomoto S *et al* 1999 *Neural Comput.* **11** 935
[5] Softky W R and Koch C 1993 *J. Neurosci.* **13** 334
[6] Shadlen M N and Newsome W T 1994 *Curr. Opin. Neurobiol.* **4** 569
[7] Dayan P and Abbott L F 2001 *Theoretical Neuroscience* (Cambridge, MA: MIT Press)
[8] Shinomoto S *et al* 2003 *Neural Comput.* **15** 2823
[9] Shinomoto S *et al* 2005 *J. Neurophysiol.* at press
[10] Shinomoto S *et al* 2005 *BioSystems* **79** 67
[11] MacKay D J C 1992 *Neural Comput.* **4** 415
[12] Neal R M 1992 *Technical Report* CRG-TR-91-1 (Department of Computer Science, University of Toronto)
[13] Bruce A D and Saad D 1994 *J. Phys. A: Math. Gen.* **27** 3355
[14] Pryce J M and Bruce A D 1995 *J. Phys. A: Math. Gen.* **28** 511
[15] Carlin B P and Louis T A 2000 *Bayes and Empirical Bayes Methods for Data Analysis* 2nd edn (New York: Chapman and Hall)
[16] Berman M 1981 *Biometrica* **68** 143
[17] Barbieri R *et al* 2001 *J. Neurosci. Methods* **105** 25

[18] Cox D R and Lewis P A W 1966 *The Statistical Analysis of Series of Events* (New York: Wiley)
[19] Daley D J and Vere-Jones D 2002 *An Introduction to the Theory of Point Processes* vol 1 2nd edn (New York: Springer)
[20] Bialek W *et al* 1996 *Phys. Rev. Lett.* **77** 4693
[21] Nemenman I and Bialek W 2002 *Phys. Rev.* E **65** 026137
[22] Dempster A P *et al* 1977 *J. R. Stat. Soc.* B **39** 1
[23] Smith A C and Brown E N 2003 *Neural Comput.* **15** 965
[24] Koyama S and Shinomoto S 2004 *J. Phys. A: Math. Gen.* **37** 7255