Estimating Neural Firing Rates: An Empirical Bayes Approach

Shinsuke Koyama

Department of Statistical Modeling, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
skoyama@ism.ac.jp

Abstract. A lot of neurophysiological findings rely on accurate estimates of firing rates. In order to estimate an underlying rate function from sparse observations, i.e., spike trains, it is necessary to perform temporal smoothing over a short time window at each time point. In the empirical Bayes method, in which the assumption for the smoothness is incorporated in the Bayesian prior probability of underlying rate, the time scale of the temporal average, or the degree of smoothness, can be optimized by maximizing the marginal likelihood. Here, the marginal likelihood is obtained by marginalizing the complete-data likelihood over all possible latent rate processes. We carry out this marginalization using a path integral method. We show that there exists a lower bound of rate fluctuations below which the optimal smoothness parameter diverges. We also show that the optimal smoothness parameter obeys asymptotic scaling laws, the exponent of which depends on the smoothness of underlying rate processes.

Keywords: Neural firing rate, empirical Bayes method, Path integrals.

1 Introduction

Neural activity, particularly in the cortex, is known to be highly variable [1]. The spike sequences generated in response to identical behavioral stimuli applied on separate occasions are not identical. In such a situation that the spiking response contains uncertainties, it is often assumed that the spike trains are generated from a smooth underlying function of time (i.e., the firing rate) and that this function carries a significant portion of the neural information. Thus, estimating the firing rate is important for understanding how the brain performs neural computations.

The difficulty in estimating the firing rate lies in the fact that spike data gives only a sparse observation of its underlying rate. In such a case, one usually repeat the same trial a number of times to find a smooth estimate. However, averaging across many trials can obscure important temporal features. Estimating the underlying rate from a few spike trains, or even from single spike trains, is therefore an important problem [2].

In order to obtain a smooth estimate of the firing rate from sparse data, it is necessary to perform temporal averaging over a short time interval at each time

T. Huang et al. (Eds.): ICONIP 2012, Part II, LNCS 7664, pp. 51–59, 2012.

point. The time scale of the averaging window (e.g., the bin size of a *peristimulus time histogram*, PSTH) must be optimized to produce a plausible estimate of the underlying rate. A method for selecting the bin size of PSTH has been proposed, according to minimizing the (estimated) mean integrated squared error (MISE) [3].

Here, we use the empirical Bayes method for estimating the firing rate, in which the assumption for the smoothness is incorporated in the Bayesian prior probability of the underlying rate, and the optimal smoothness parameter is determined by maximizing the marginal likelihood [4–7].

We utilize the path integral method [8, 9], a technique developed in the fields of quantum mechanics and statistical mechanics, in order to carry out the marginalization over the latent path space. We show that there exists a lower bound of the degree of rate fluctuations, below which the smoothness parameter diverges. We also show that the optimal smoothness parameter obeys asymptotic scaling laws, the exponent of which depends on the smoothness of the underlying rate processes.

2 Inhomogeneous Gamma Process

We assume the underlying firing rate $\lambda(t)(\geq 0)$, and then consider that the model of spike trains is given by a conditionally inhomogeneous renewal process, given $\lambda(t)$, constructed in the following manner. First, we consider spikes $\{s_0, \ldots, s_N\}$ occurring along the time axis according to the renewal process. Here, we employ the gamma distribution for the interspike interval distribution:

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

where $\Gamma(\kappa) = \int_0^\infty x^{\kappa-1} e^{-x} dx$ 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 shape parameter κ. An inhomogeneous gamma process can be constructed by rescaling the time of the renewal gamma process with $\lambda(t)$ as $t_i = \Lambda^{-1}(s_i)$, where $\Lambda(t) := \int_0^t \lambda(u)du$. Accordingly, the probability density of a sequence of spikes $\{t_i\} := \{t_0, \ldots, t_N\}$ on the interval $[0, T]$ is given by

$$p_\kappa(\{t_i\}|\{\lambda(t)\}) = p_0(t_0|\{\lambda(t)\}) \cdot \prod_{i=1}^{N} \lambda(t_i) f_\kappa(\Lambda(t_i) - \Lambda(t_{i-1})) \cdot p_T(t_N|\{\lambda(t)\}), \tag{2}$$

where $p_0(t_0|\{\lambda(t)\})$ is the density of the first spike occurring at t_0, and $p_T(t_N|\{\lambda(t)\})$ is the probability of no spikes being observed on $(t_N, T]$. This inhomogeneous gamma process is a natural extension of both the inhomogeneous Poisson process ($\kappa = 1$) and the renewal gamma process (for which $\lambda(t)$ is constant) [10].

3 Empirical Bayes Method

We employ the empirical Bayes method to decode the underlying firing rate $\lambda(t)$ from n spike trains $\{t_i^j\} := \{t_0^j, \ldots, t_{N_j}^j\}_{j=1}^n$, ($N_j$ being the number of spikes in the jth trial,) independent and identically derived from the inhomogeneous gamma process (2). The likelihood of n spike trains is then given by the product of the likelihood of single spike train. Let $x(t) \in \mathbb{R}$ a latent process that is transformed from $\lambda(t)$ via the log-link function $x(t) = \log \lambda(t)$. For the inference of $\lambda(t)$ from $\{t_i^j\}$, we use a prior distribution of $x(t)$, such that the large gradient of $x(t)$ is penalized with

$$p_\gamma(\{x(t)\}) = \frac{1}{Z(\gamma)} \exp\left[-\frac{1}{2\gamma^2} \int_0^T \left(\frac{dx(t)}{dt} \right)^2 dt \right], \tag{3}$$

where the hyperparameter γ controls the smoothness of the latent process $x(t)$. $Z(\gamma)$ is the normalization constant given by

$$\begin{aligned} Z(\gamma) &= \int \exp\left[-\frac{1}{2\gamma^2} \int_0^T \left(\frac{dx(t)}{dt} \right)^2 dt \right] \mathcal{D}\{x(t)\} \\ &= \frac{1}{\sqrt{2\pi\gamma^2 T}} \exp\left[-\frac{\{x(T)-x(0)\}^2}{2\gamma^2 T} \right], \end{aligned} \tag{4}$$

where $\int \mathcal{D}\{x(t)\}$ represents integration over all latent processes, or the Wiener integral over all paths of $x(t)$ [11]. Eq. (3) is the same as the smoothing-spline penalization with the first-order derivative [12]. By inverting the conditional probability distribution with the Bayes rule, the posterior distribution of $\{x(t)\}$ is obtained as

$$p_{\nu,\gamma}(\{x(t)\}|\{t_i^j\}) = \frac{p_\nu(\{t_i^j\}|\{x(t)\}) p_\gamma(\{x(t)\})}{p_{\nu,\gamma}(\{t_i^j\})}. \tag{5}$$

The hyperparameter, γ and ν, which represent the inverse smoothness of the latent process and the shape of the gamma distribution, can be determined by maximizing the marginal likelihood defined by

$$p_{\nu,\gamma}(\{t_i^j\}) = \int p_\nu(\{t_i^j\}|\{x(t)\}) p_\gamma(\{x(t)\}) \mathcal{D}\{x(t)\}. \tag{6}$$

Under a set of hyperparameters $(\hat{\gamma}, \hat{\kappa}) = \arg\max_{\gamma,\nu} p_{\nu,\gamma}(\{t_i^j\})$, that optimizes the smoothness of the rate process and the shape of the gamma distribution, we can determine the maximum a posteriori (MAP) estimate of the latent process $\hat{x}(t)$. The estimate of the firing rate is then obtained as $\hat{\lambda}(t) = e^{\hat{x}(t)}$.

4 Path Integral Analysis

We suppose that the underlying firing rate is given by

$$\lambda(t) = \mu + \sigma f(t), \tag{7}$$

where μ is the mean firing rate, and $f(t)$ represent the rate fluctuation such that $\langle f(t)\rangle = 0$ and $\langle f(t)f(t')\rangle = \phi(t-t')$, $\langle\cdot\rangle$ denoting the ensemble average.

We also assume the following conditions for the asymptotic analysis: (A) the time scale of the temporal modulation of $f(t)$ is longer than the mean interspike interval; (B) σ/μ is small; (C) a large observation interval $T \gg 1$, or equivalently, a large number of spikes $N_j \gg 1$ for $i = 1, \dots, n$.

4.1 Evaluation of the Marginal Likelihood

From the condition (A), the firing rate in each interspike interval could be approximated to be constant, from which we can separate the rate fluctuation from interspike interval as $\log(\Lambda(t_i^j) - \Lambda(t_{i-1}^j)) \approx x(t_i^j) + \log(t_i^j - t_{i-1}^j)$. We further decompose the state $x(t)$ into the mean $\log\mu$ and fluctuation $y(t)$ as $x(t) = \log\mu + y(t)$. Accordingly, log of the likelihood function of the n spike trains is decomposed into two parts as $\log p_\nu(\{t_i^j\}|\{x(t)\}) = \mathcal{H} + \mathcal{I}$, where $\mathcal{H}$ represents the log likelihood function of the gamma distribution,

$$\mathcal{H} = \sum_{j=1}^{n}\Bigg[-T\nu\mu + N_j\nu\log\mu + N_j\nu\log\nu - N_j\log\Gamma(\nu) + (\nu-1)\sum_{i=1}^{N_j}\log(t_i^j - t_{i-1}^j)\Bigg], \tag{8}$$

whereas $\mathcal{I}$ represents the contribution of rate fluctuation. Note that we have omitted $p_0(t_0^j|\{\lambda(t)\})$ and $p_T(t_{N_j}^j|\{\lambda(t)\})$ due to $T \gg 1$. Substituting $\mathcal{H}$ and $\mathcal{I}$ into Eq. (6), the marginal likelihood function is expressed as $p_{\nu,\gamma}(\{t_i^j\}) = e^{\mathcal{H}}\mathcal{F}/Z(\gamma)$. Here, the contribution of the rate fluctuation can be represented in the form of a path integral [9]:

$$\mathcal{F} = \int \exp\left[-\int_0^T L(y,\dot{y})dt\right]\mathcal{D}\{y(t)\}, \tag{9}$$

where $L(y,\dot{y})$ is a "Lagrangian" of the form:

$$L(y,\dot{y}) = \frac{1}{2\gamma^2}\dot{y}^2 + n\nu\mu(e^y - 1) - \nu\sum_{j=1}^{n}\sum_{i=1}^{N_j}\delta(t - t_i^j)y. \tag{10}$$

The fluctuation in the apparent spike count is given by the variance to mean ratio as represented by the Fano factor. For the renewal process in which interspike intervals are drawn from a given distribution function, it is proven that the Fano factor is related to the interspike interval variability with $F \approx C_V^2$, where C_V is the coefficient of variation defined as the standard deviation of the ISIs divided by the mean [13]. The interspike interval variability of the gamma distribution (1) is given by $C_V = 1/\sqrt{\kappa}$. Thus, in each realization of a spike generation, the occurrence of events fluctuates around the underlying rate, and the superimposed spike train in Eq. (10) can be represented as a stochastic process, $\sum_{j=1}^{n}\sum_{i=1}^{N_j}\delta(t-t_i^j) = n\lambda(t) + \sqrt{n\lambda(t)/\kappa}\xi(t)$, where $\xi(t)$ is a fluctuating process such that $\langle\xi(t)\rangle = 0$ and $\langle\xi(t)\xi(t')\rangle = \delta(t-t')$. Note that this approximation holds under the condition (A). Using this, and taking up to the second-order terms with respect to y, the Lagrangian (10) becomes

$$L(y, \dot{y}) = \frac{1}{2\gamma^2}\dot{y}^2 + \frac{n\nu\mu}{2}y^2 - n\nu\left[\sigma f(t) + \sqrt{\frac{\lambda(t)}{n\kappa}}\xi(t)\right]y, \tag{11}$$

which holds in $O((\sigma/\mu)^{3/2})$ due to $y \sim \sigma/\mu$.

The MAP estimate $\hat{y}$ is obtained by taking the extremum of the action integral $S[y(t)] := \int_0^T L(y, \dot{y})dt$ in Eq. (9). The extremum condition for S is expressed by the variational equation $\delta S = 0$, and an integration by part in δS with fixed boundary values leads to the Euler-Lagrange equation for $\hat{y}(t)$:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\hat{y}}}\right) - \frac{\partial L}{\partial \hat{y}} = 0. \tag{12}$$

In the following analysis, we consider a long interval $T \gg 1$, in which the boundary effect is negligible.

By considering the deviation from the MAP path as $y(t) = \hat{y}(t) + \eta(t)$, ($\eta(0) = \eta(T) = 0$), and approximating the action integral to a range quadratic with respect to $\eta(t)$, the path integral (9) can be expressed as $\mathcal{F} = Re^{-S[\hat{y}(t)]}$, where $e^{-S[\hat{y}(t)]}$ represents the contribution of the mode to the path integral, whereas R represents that of quadratic derivation:

$$R = \int \exp\left[-\frac{1}{2}\int_0^T \left(\frac{1}{\gamma^2}\dot{\eta}^2 + n\nu\mu\eta^2\right)dt\right]\mathcal{D}\{\eta(t)\}. \tag{13}$$

Using this, the marginal likelihood function is computed analytically as

$$p_{\nu,\gamma}(\{t_i\}) = \frac{e^{\mathcal{H}}}{Z(\gamma)}Re^{-S[\hat{y}(t)]}. \tag{14}$$

Note that the path integral method presented here can be regarded as a functional version of the Laplace approximation. In the following, we evaluate the three factors in Eq. (14): $e^{-S[\hat{y}(t)]}$, R and $e^{\mathcal{H}}$.

First, consider the contribution of the mode to the path integral $e^{-S[\hat{y}(t)]}$. The MAP path $\hat{y}(t)$ is obtained by solving the Euler-Lagrange equation (12) associated with the Lagrangian (11) as

$$\hat{y}(t) = \frac{\gamma}{2}\sqrt{\frac{n\nu}{\mu}}\int_0^T e^{-\gamma\sqrt{n\nu\mu}|t-s|}\left[\sigma f(s) + \sqrt{\frac{\lambda(s)}{n\kappa}}\xi(s)\right]ds. \tag{15}$$

By using Eqs. (11) and (12), we obtain

$$S[\hat{y}(t)] = \int_0^T \left[\frac{1}{2\gamma^2}\frac{d}{dt}(\hat{y}\dot{\hat{y}}) - \frac{n\nu}{2}\left\{\sigma f(t) + \sqrt{\frac{\lambda(t)}{n\kappa}}\xi(t)\right\}\hat{y}\right]dt. \tag{16}$$

For $T \gg 1$, the boundary effect is negligible so that the first-term in the rhs of the above equation vanishes. Substituting the MAP path (15) into Eq. (16) leads to

$$\frac{1}{T}S[\hat{y}(t)] = -\frac{\gamma\sqrt{n\nu\mu}}{4}\left\{\frac{\nu}{\kappa} + \frac{2n\nu\sigma^2}{\mu}\int_0^\infty \phi(u)e^{-\gamma\sqrt{n\nu\mu}u}du\right\}. \tag{17}$$

Consider next R given by Eq. (13). An integration by part leads to

$$\int_0^T \left(\frac{1}{\gamma^2}\dot{\eta}^2 + n\nu\mu\eta^2\right)dt = \int_0^T \eta\left(-\frac{1}{\gamma^2}\partial_t^2 + n\nu\mu\right)\eta dt, \tag{18}$$

where we have used the boundary condition $\eta(0) = \eta(T) = 0$. Let $\{\vartheta_i(t)\}$ be a complete set of orthogonal eigenfunctions of $\left(-\frac{1}{\gamma^2}\partial_t^2 + n\nu\mu\right)$ vanishing at the boundaries, and $\{\theta_i\}$ be its eigenvalues. Then, $\eta(t)$ can be expressed as $\eta(t) = \sum_i a_i\vartheta_i(t)$. Accordingly, the measure is transformed as $\mathcal{D}\{\eta(t)\} = C\prod_i (2\pi)^{-\frac{1}{2}}da_i$, where C is a constant chosen so that the integral over this measure corresponds to the Wiener integral (4), and we find

$$\begin{aligned} R &= C\prod_i \int_{-\infty}^{\infty} \frac{da_i}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\theta_i a_i^2\right) \\ &= C\prod_i \theta_i^{-\frac{1}{2}} := C\det\left(-\frac{1}{\gamma^2}\partial_t^2 + n\nu\mu\right)^{-\frac{1}{2}}. \end{aligned} \tag{19}$$

From Eqs. (4) and (19), we also obtain $C\det\left(-\frac{1}{\gamma^2}\partial_t^2\right)^{-\frac{1}{2}} = 1/\sqrt{2\pi\gamma^2 T}$. Thus, R is obtained as

$$R = \frac{1}{\sqrt{2\pi\gamma^2 T}}\left[\frac{\det\left(-\frac{1}{\gamma^2}\partial_t^2 + n\nu\mu\right)}{\det\left(-\frac{1}{\gamma^2}\partial_t^2\right)}\right]^{-\frac{1}{2}} = \frac{1}{\sqrt{2\pi\gamma^2 T}}\left[\frac{\varphi_1(T)}{\varphi_2(T)}\right]^{-\frac{1}{2}}. \tag{20}$$

It has been proved that the determinants can be computed by solving the associated differential equations [8, 9]:

$$\begin{aligned} \left(-\frac{1}{\gamma^2}\partial_t^2 + \frac{\kappa}{\mu}\right)\varphi_1(t) = 0, \quad \varphi_1(0) = 0, \quad \dot{\varphi}_1(0) = 1, \\ -\frac{1}{\gamma^2}\partial_t^2\varphi_2(t) = 0, \quad \varphi_2(0) = 0, \quad \dot{\varphi}_2(0) = 1. \end{aligned}$$

These differential equations are solved as $\varphi_1(t) = \frac{1}{\gamma\sqrt{n\nu\mu}}\sinh\gamma\sqrt{n\nu\mu}t$ and $\varphi_2(t) = t$, from which R is obtained as

$$\lim_{T\to\infty}\frac{1}{T}\log R = -\frac{\gamma\sqrt{n\nu\mu}}{2}. \tag{21}$$

In order to evaluate $\mathcal{H}$ given by Eq. (8), we need to compute $\frac{1}{N_j}\sum_i \log(t_i^j - t_{i-1}^j)$. Let $\{t_i^{j(\lambda)} - t_{i-1}^{j(\lambda)}\}$ be a set of interspike intervals derived from the gamma distribution with the rate λ in the jth spike train, and $N_j^{(\lambda)}$ be the number of the interspike intervals in this set. Then, we obtain

$$\frac{1}{N_j^{(\lambda)}}\sum_{i=1}^{N_j^{(\lambda)}}\log(t_i^{j(\lambda)}-t_{i-1}^{j(\lambda)})\to\psi(\kappa)-\log\kappa-\log\lambda,\quad \text{as } N_j^{(\lambda)}\to\infty,$$

where $\psi(\kappa)$ is the digamma function. On the other hand, $N_j^{(\lambda)}/N_j\to\lambda p(\lambda)d\lambda/\mu$ as $N_j\to\infty$, from the law of large number. Using these, we obtain

$$\lim_{N_j\to\infty}\frac{1}{N_j}\sum_{i=1}^{N_j}\log(t_i^j-t_{i-1}^j)=\int_0^\infty[\psi(\kappa)-\log\kappa-\log\lambda]\frac{\lambda p(\lambda)}{\mu}d\lambda.$$

Expanding up to the second-order with respect to σ/μ, the above equation can be evaluated as $\psi(\kappa)-\log\kappa-\log\mu-\frac{\sigma^2\phi(0)}{2\mu^2}$. Substituting this into Eq. (8), we obtain the factor $e^{\mathcal{H}}$.

Substituting the three factors into Eq. (14), the log marginal likelihood function is obtained as

$$\begin{aligned}\mathcal{L}(\gamma,\nu):=\frac{1}{T}\log p_{\nu,\gamma}(\{t_i\})&=\frac{1}{T}\left(\log R-\log Z(\gamma)-\int_0^T L(\hat{y},\dot{\hat{y}})dt+\mathcal{H}\right)\\&=-\frac{\gamma\sqrt{n\nu\mu}}{4}\left\{2-\frac{\nu}{\kappa}\left(1+\frac{2n\kappa\sigma^2}{\mu}\int_0^\infty\phi(u)e^{-\gamma\sqrt{n\nu\mu}u}du\right)\right\}\\&\quad+n\sqrt{\mu\sigma}\left\{\left(\log\mu-\nu+\nu\log\nu-\log\Gamma(\nu)\right.\right.\\&\quad\left.\left.+(\nu-1)[\psi(\kappa)-\log\kappa]\right)\left(\frac{\sigma}{\mu}\right)^{-\frac{1}{2}}-\frac{(\nu-1)\phi(0)}{2}\left(\frac{\sigma}{\mu}\right)^{\frac{3}{2}}\right\},\end{aligned}\qquad(22)$$

in the limit of $T\to\infty$. Note that Eq. (22) holds in $O((\sigma/\mu)^{3/2})$.

4.2 Results

Divergence of the Optimal Smoothness Parameter. In the range of the parameter space (γ,ν) that is valid for the asymptotic analysis (in which $o((\sigma/\mu)^{3/2})$ is negligible), the log marginal likelihood function (22) can have a maximum at $(\gamma,\nu)=(0,\hat{\kappa}_0)$ or at $(\hat{\gamma},\hat{\kappa})$, $\hat{\gamma}>0$, which correspond to constant and fluctuating rate estimations, respectively.

For the case of $\gamma=0$, the fluctuation in the rate estimation (15) vanishes, and thus the rate estimation becomes constant $\hat{\lambda}(t)=\mu$. Taking $\partial\mathcal{L}/\partial\nu=0$ leads to $\psi(\hat{\kappa}_0)-\log\hat{\kappa}_0-[\psi(\kappa)-\log\kappa-\frac{\sigma^2\phi(0)}{2\mu^2}]=0$, the solution of which is obtained as

$$\hat{\kappa}_0=\kappa-\frac{\sigma^2\phi(0)}{2\mu^2 I(\kappa)}+o((\sigma/\mu)^2),\qquad(23)$$

where $I(\kappa) = \dot{\psi}(\kappa) - 1/\kappa$ is the Fisher information of the gamma distribution.

We next evaluate the fluctuating rate estimation ($\hat{\gamma} > 0$) if it exists. From (23), it must be that $\hat{\kappa} = \kappa + O((\sigma/\mu)^2)$. Then, the log marginal likelihood function becomes

$$\mathcal{L}(\gamma, \hat{\kappa}) = \mathcal{L}(\gamma, \kappa) = -\frac{1}{4\Delta}\left[1 - 2\beta \int_0^\infty \phi(u) e^{-u/\Delta} du\right] + \mathcal{L}(0, \hat{\kappa}_0), \tag{24}$$

in $O((\sigma/\mu)^{3/2})$, where we defined $\beta := n\kappa\sigma^2/\mu$, and $\Delta := 1/(\gamma\sqrt{n\kappa\mu})$ represents the time scale of smoothness. $\mathcal{L}(\gamma, \kappa)$ satisfies $\mathcal{L}(0, \hat{\kappa}) = 0$ and $\mathcal{L}(\infty, \hat{\kappa}) = -\infty$, and has the global maximum either at $\gamma = \hat{\gamma} > 0$ or $\gamma = 0$, depending on the value of β. $\mathcal{L}(\gamma, \kappa)$ has the global maximum at $\gamma = \hat{\gamma} > 0$ if β exceeds the critical value:

$$\beta_c = \frac{1}{2 \max_\Delta \int_0^\infty \phi(u) e^{-u/\Delta} du}. \tag{25}$$

In other words, the optimal time scale of smoothness $\hat{\Delta} = 1/(\hat{\gamma}\sqrt{n\kappa\mu})$ diverges if $\beta < \beta_c$ (or the degree of rate fluctuations σ^2/μ is smaller than $\beta_c/(n\kappa)$), in which the empirical Bayes method cannot detect the underlying rate fluctuations.

Asymptotic Scaling Laws. For $\beta \gg \beta_c$, which can be achieved by a large number of trials n for instance, the optimal γ becomes $\hat{\gamma} \gg 1$, or $\hat{\Delta} \ll 1$. Then, expanding Eq. (24) and taking the derivative with respect to $\hat{\Delta}$ leads to

$$1 + 2\beta\{\dot{\phi}(0)\hat{\Delta}^2 + 2\ddot{\phi}(0)\hat{\Delta}^3 + \cdots\} = 0. \tag{26}$$

If the underlying rate fluctuates smoothly ($\dot{\phi}(0) = 0$ and $\ddot{\phi}(0) < 0$), the optimal time scale obeys the scaling law:

$$\hat{\Delta} \sim \{-4\ddot{\phi}(0)\beta\}^{-\frac{1}{3}}. \tag{27}$$

If, on the other hand, the underlying rate fluctuation is nowhere differentiable ($\dot{\phi}(0) < 0$), such as a path of Brownian motions, $\hat{\Delta}$ obeys

$$\hat{\Delta} \sim \{-2\dot{\phi}(0)\beta\}^{-\frac{1}{2}}. \tag{28}$$

The asymptotic MISE between the true and estimated rates is also obtained as

$$MISE := \lim_{T\to\infty} \frac{1}{T} \int_0^T \{\hat{\lambda}(t) - \lambda(t)\}^2 dt \sim \frac{\sigma^2}{4\beta\hat{\Delta}}.$$

The MISE for smooth rate fluctuations is then given by

$$MISE \sim 4^{-\frac{2}{3}}\{-\ddot{\phi}(0)\}^{\frac{1}{3}}\sigma^2\beta^{-\frac{2}{3}}, \tag{29}$$

while the MISE for rate fluctuations that are nowhere differentiable is scaled as

$$MISE \sim 2^{-\frac{3}{2}}\{-\dot{\phi}(0)\}^{\frac{1}{2}}\sigma^2\beta^{-\frac{1}{2}}. \tag{30}$$

5 Discussion

We carried out the marginalization (6) using the path integral method, and obtained the optimal smoothness parameter by maximizing the likelihood function. This analysis enabled us to derive the lower bound of the degree of rate fluctuations, below which the optimal time scale of smoothness diverges: in other words, the empirical Bayes estimator cannot detect the underlying rate fluctuations. We also derived the asymptotic characteristics of the rate estimator for a larger number of trials.

Note that there commonly exists a lower bound below which the underlying rate fluctuations are undetectable, not only in the empirical Bayes method with the above prior distribution (3), but also with other prior distributions, and in other rate estimators such as PSTH [14]. The condition for the lower bound in these methods is similar to Eq. (25). The asymptotic characteristics (27) and (28) are also respectively the same as those for the optimal bin size of PSTH determined with the MISE criteria [15]. It would be interesting to investigate how general these results are to other inference frameworks.

References

1. Softky, W.R., Koch, C.: The Highly Irregular Firing of Cortical Cells Is Inconsistent with Temporal Integration of Random EPSPs. J. Neurosci. 13, 334–350 (1993)
2. Kass, R.E., Ventura, V., Brown, E.N.: Statistical Issues in the Analysis of Neuronal Data. J. Neurophysiol. 94, 8–25 (2005)
3. Shimazaki, H., Shinomoto, S.: A Method for Selecting the Bin Size of a Time Histogram. Neural Comp. 19, 1503–1700 (2007)
4. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)
5. Cunningham, J.P., Yu, B.M., Shenoy, K.V., Sahani, M.: Inferring Neural Firing Rates from Spike Trains Using Gaussian Processes. In: Neural Information Processing Systems, vol. 20, pp. 329–336 (2008)
6. Koyama, S., Shinomoto, S.: Empirical Bayes Interpretations of Random Point Events. J. Phys. A: Math. Gen. 38, L531–L537 (2005)
7. Koyama, S., Shinomoto, S.: Phase Transitions in the Estimation of Event Rate: a Path Integral Analysis. J. Phys. A. Math. Theor. 40, F383–F390 (2007)
8. Coleman, S.: Aspects of Symmetry. Cambridge University Press (1988)
9. Kleinert, H.: Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets, 5th edn. World Scientific Publishing Company (2009)
10. Berman, M.: Inhomogeneous and Modulated Gamma Processes. Biometrica 68, 143–152 (1981)
11. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus, 2nd edn. Springer (1997)
12. Ramsay, J., Silverman, B.W.: Functional Data Analysis, 2nd edn. Springer (2010)
13. Cox, D.R.: Renewal Theory. Chapman and Hall (1962)
14. Shintani, T., Shinomoto, S.: Detection Limit for Rate Fluctuations in Inhomogeneous Poisson Processes. Phys. Rev. E 85, 041139 (2012)
15. Koyama, S., Shinomoto, S.: Histogram Bin Width Selection for Time-Dependent Poisson Processes. J. Phys. A: Math. Gen. 37, 7255–7265 (2004)