

On the spike train variability characterized by variance-to-mean power relationship

Shinsuke Koyama

Department of Statistical Modeling, The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

We propose a statistical method for modeling the non-Poisson variability of spike trains observed in a wide range of brain regions. Central to our approach is the assumption that the variance and the mean of interspike intervals are related by a power function characterized by two parameters: the scale factor and exponent. It is shown that this single assumption allows the variability of spike trains to have an arbitrary scale and various dependencies on the firing rate in the spike count statistics, as well as in the interval statistics, depending on the two parameters of the power function. We also propose a statistical model for spike trains that exhibits the variance-to-mean power relationship, and based on this a maximum likelihood method is developed for inferring the parameters from rate-modulated spike trains. The proposed method is illustrated on simulated and experimental spike trains.

1 Introduction

The variability of neural firing is of central importance in the study of signal processing that is carried out by the nervous system. The reliable transmission of sensory signals, integration of neural information, and precise control of neural-motor systems are

significantly dependent on the variability of the neural responses to identical sensory or behavioral variables, as well as on the average responses (Mainen and Sejnowski, 1995; de Ruyter van Steveninck et al., 1997; Harris and Wolpert, 1998; Shadlen and Newsome, 1998; Ma et al., 2006; Lu et al., 2013).

Two types of measurement, inter-spike interval (ISI) and spike count, are commonly used to quantify the variability of spike trains. The variability of ISI, expressed in the variance, quantifies how irregular the firing time is on a short timescale, characterized by the typical ISI. Since the variance of ISI is computed within single spike trains, it signifies intra-trial variability. The variance of the spike count across repeated observations, by contrast, quantifies the trial-to-trial variability in relatively long time intervals. These two quantities are by no means independent variables, but are closely related (Nawrot et al., 2008). In general, the variances of both ISI and spike count are scaled by the mean, the degree of which may vary across different brain regions (Kara et al., 2000; Maimon and Assad, 2009).

In this article, we formulate a statistical framework for modeling the variability of spike trains in terms of both the ISI and counting statistics. Our approach is motivated by an observation made by Troy and Robson (1992). They reported that for steady discharges of X retinal ganglion cells of cats, in response to stationary visual patterns, the standard deviation of ISI increases as approximately the $3/2$ power of the mean ISI. Motivated by their observation, we make a single assumption about the ISI statistics:

$$\text{Var}(\text{ISI}) = \phi E(\text{ISI})^\alpha, \quad (1)$$

where ϕ is the scale factor controlling the overall variance of ISIs, and α is the exponent controlling how the variance is scaled by the mean. Presently, it should be emphasized that this statistical assumption is a generalization of the finding of Troy and Robson (1992), in the sense that $\phi(> 0)$ and α can take arbitrary values in theory. On the basis of the power law (1), we show that this allows the spike trains to have a wide range of variability in the counting statistics, as well as in the ISI statistics observed across the brain areas, depending on ϕ and α . By combining Eq. (1) with the time-rescaling transformation (Barbieri et al., 2001), we propose a “generalized” rate-modulated renewal process to model spike trains, and develop a maximum likelihood method to infer ϕ

and α from rate-modulated spike trains.

The rest of this article is organized as follows. In section 2, we develop a statistical method. In section 3, we illustrate our method on simulated and experimental data. Section 4 contains discussions on the possible implications of the results.

2 Theory

2.1 Statistical assumption

Consider spike trains whose ISIs are independent and identically distributed, with mean μ and variance σ^2 . The central assumption in our approach is that the variance of ISI has a power function of the mean, in the form

$$\sigma^2 = \phi\mu^\alpha, \quad (2)$$

where $\phi > 0$ is the scale factor controlling the overall amplitude of the power law, and α is the exponent controlling how the variance is scaled by changing the mean. For $\alpha = 2$, the scale factor ϕ corresponds to the squared coefficient of variation, whose value is unity for a Poisson process. By contrast, values of $\alpha > 2$ (< 2) imply that the timing of spike tends to be over (under) dispersed for large means, and under (over) dispersed for small means.

Next, consider the spike count. Let N_Δ be the number of spikes in the counting window of duration Δ . The variability of spike count is often measured by the Fano factor, defined by the ratio of the variance to the mean:

$$F_\Delta := \frac{\text{Var}(N_\Delta)}{\text{E}(N_\Delta)}, \quad (3)$$

where the expectation is computed over repeated observations. For a large counting window $\Delta \gg \mu$, the mean and variance of N_Δ are asymptotically evaluated as $\text{E}(N_\Delta) \sim \Delta/\mu$ and $\text{Var}(N_\Delta) \sim \sigma^2\Delta/\mu^3$, respectively (Cox, 1962). Suppose that the variance of ISIs obeys Eq. (2). Then, for large Δ the Fano factor exhibits the power law

$$F_\Delta \sim \phi\lambda^\gamma, \quad (4)$$

where

$$\lambda := \frac{E(N_\Delta)}{\Delta} \quad (5)$$

is the mean firing rate, and the exponent γ is related to that of the ISI statistics via the scaling relation:

$$\gamma = 2 - \alpha. \quad (6)$$

Eq. (4) describes the dependency of the Fano factor on the ISI parameters and the mean firing rate λ . For $\gamma = 0$ (i.e., $\alpha = 2$), the Fano factor does not depend on λ ; in other words, the variance of the spike count is proportional to the mean. If $\gamma > 0$ ($\alpha < 2$), the Fano factor increases as λ increases, while the Fano factor is inversely related to λ if $\gamma < 0$ ($\alpha > 2$).

The Fano factor depends on the length of the counting window Δ . When $\Delta \ll \mu$, the probability of two and more spikes is negligible, and the spike count can be approximated by a Bernoulli random variable with probabilities $P(N_\Delta = 1) = \lambda\Delta$ and $P(N_\Delta = 0) = 1 - \lambda\Delta$, respectively. The variance of the Bernoulli distribution is $\lambda\Delta(1 - \lambda\Delta)$, so that for any values of α and ϕ the Fano factor approaches unity (Teich et al., 1997):

$$\lim_{\Delta \rightarrow 0} F_\Delta = \lim_{\Delta \rightarrow 0} \frac{\lambda\Delta(1 - \lambda\Delta)}{\lambda\Delta} = 1, \quad (7)$$

which is different from Eq. (4). In the numerical studies presented in section 3, we choose Δ so that an average of five spikes fall in the window, which is enough for Eq. (4) to apply.

2.2 Statistical model

2.1 Generalized rate-modulated renewal process

We construct a statistical model for spike trains whose variability is characterized by the variance-to-mean power law. Consider first the stationary renewal process, a class of point processes in which ISIs are independent and identically distributed. Let X be a random variable describing ISI. It follows from Eq. (2) that by rescaling ISI as $X \rightarrow \lambda X$, $\lambda = 1/\mu$ being the mean firing rate, the parameters are rescaled as $\mu \rightarrow 1$ and $\phi \rightarrow \lambda^{2-\alpha}\phi$. Thus, a parametric probability density function $f(x; \mu, \phi)$ that has

mean μ and variance $\phi\mu^\alpha$, and is invariant under the rescaling, satisfies

$$f(x; \mu, \phi) = \lambda f(\lambda x; \lambda^{2-\alpha} \phi), \quad (8)$$

where $f(x; \phi) := f(x; 1, \phi)$. Eq. (8) suggests that one can always reparametrize an arbitrary probability density function with unit mean and variance ϕ , so that the variance has the power function of the mean (2).

We extend the stationary renewal process defined by Eq. (8), to a rate-modulated process. Let $N(t)$ be the number of spikes that have already occurred at time t , and t_i denote the i th spike time. A point process is fully defined by a conditional intensity function (Daley and Vere-Jones, 2003; Kass and Ventura, 2001),

$$r(t; H(t)) = \lim_{dt \rightarrow 0} \frac{P\{N(t+dt) - N(t) = 1; H(t)\}}{dt}, \quad (9)$$

where $H(t) = \{t_1, t_2, \dots, t_{N(t)}\}$ denotes the history of spikes up to the time t . For a renewal process whose ISI density function is given by $f(x; \phi)$, the conditional intensity function, also called the hazard function, is given by

$$r(t; t_{N(t)}, \phi) = \frac{f(t - t_{N(t)}; \phi)}{1 - \int_{t_{N(t)}}^t f(u - t_{N(t)}; \phi) du}. \quad (10)$$

Let $\lambda(t) > 0$ be an instantaneous firing rate, and define

$$\Lambda(t) = \int_0^t \lambda(u) du, \quad (11)$$

which is monotone and invertible. By rescaling the time $t \rightarrow \Lambda(t)$, we can obtain the “conventional” rate-modulated renewal process (Barbieri et al., 2001; Berman, 1981; Koyama and Kass, 2008; Koyama and Kostal, 2014; Nawrot et al., 2008; Pillow, 2008; Reich et al., 1998), whose conditional intensity function is given by

$$r(t; t_{N(t)}, \{\lambda(t)\}, \phi) = \frac{\lambda(t) f(\Lambda(t) - \Lambda(t_{N(t)}); \phi)}{1 - \int_{t_{N(t)}}^t \lambda(v) f(\Lambda(v) - \Lambda(t_{N(t)}); \phi) dv}. \quad (12)$$

Note that the expectation of Eq. (12) is equal to the following:

$$\lambda(t) = \mathbb{E}[r(t; t_{N(t)}, \{\lambda(t)\}, \phi)], \quad (13)$$

where $\lambda(t)$ is also called the “marginal” intensity function, which does not depend on the past spikes. However, the Fano factor of the process (12) does not have the power law with the exponent (6).¹

We generalize Eq. (12) such that the Fano factor has a power function of the firing rate. Analogously with Eq. (8), by rescaling the parameter $\phi \rightarrow \lambda(t)^{2-\alpha}\phi$, as well as the time $t \rightarrow \Lambda(t)$, the conditional intensity function of a “generalized” rate-modulated renewal process is obtained as

$$r(t; t_{N(t)}, \{\lambda(t)\}, \phi, \alpha) = \frac{\lambda(t)f(\Lambda(t) - \Lambda(t_{N(t)}); \lambda(t)^{2-\alpha}\phi)}{1 - \int_{t_{N(t)}}^t \lambda(v)f(\Lambda(v) - \Lambda(t_{N(t)}); \lambda(v)^{2-\alpha}\phi)dv}. \quad (14)$$

Eq. (14) is reduced to the conditional intensity function associated with Eq. (8) if $\lambda(t) = \lambda$, and corresponds to Eq. (12) if $\alpha = 2$.

2.2 Likelihood function

Using the conditional intensity function (14), the probability density of the spike trains $\{t_i\} := \{t_1, t_2, \dots, t_n\}$ in the interval $(0, T]$ can be expressed as

$$\begin{aligned} p(\{t_i\}; \{\lambda(t)\}, \phi, \alpha) &= P_1(t_1) \prod_{i=2}^n r(t_i; t_{i-1}, \{\lambda(t)\}, \phi, \alpha) \\ &\quad \times \exp \left[- \int_{t_1}^{t_n} r(u; t_{N(u)}, \{\lambda(t)\}, \phi, \alpha) du \right] P_0((t_n, T]), \end{aligned} \quad (15)$$

where $P_1(t_1)$ is the probability of the first spike occurring at time t_1 , $P_0((t_n, T])$ is the probability of no spikes occurring in the interval $(t_n, T]$, and the exponential factor represents the probability of there being no spikes in each interspike interval (Daley and Vere-Jones, 2003; Kass and Ventura, 2001). Substituting Eq. (14) into Eq. (15) yields the more

¹ In fact, this transformation results in the Fano factor being a constant, whose value is determined by ϕ .

tractable form (see Appendix A):

$$\begin{aligned}
p(\{t_i\}; \{\lambda(t)\}, \phi, \alpha) \\
= P_1(t_1) \prod_{i=2}^n \lambda(t_i) f(\Lambda(t_i) - \Lambda(t_{i-1}); \lambda(t_i)^{2-\alpha} \phi) P_0((t_n, T]). \quad (16)
\end{aligned}$$

For spike trains consisting of M repeated trials, $\{t_i^j\}_{j=1}^M := \{t_1^j, \dots, t_{n_j}^j\}_{j=1}^M$, the log likelihood function of (ϕ, α) , given $\{\lambda(t)\}$, is obtained as

$$\begin{aligned}
L(\phi, \alpha; \{\lambda(t)\}, \{t_i^j\}_{j=1}^M) &= \sum_{j=1}^M \sum_{i=2}^{n_j} \{ \log \lambda(t_i^j) + \log f(u_i^j; \xi_i^j) \} \\
&\quad + \sum_{j=1}^M \{ \log P_1(t_1^j) + \log P_0((t_{n_j}^j, T]) \}, \quad (17)
\end{aligned}$$

where $\xi_i^j = \lambda(t_i^j)^{2-\alpha} \phi$ and $u_i^j = \Lambda(t_i^j) - \Lambda(t_{i-1}^j)$. In the following analysis, we assume that there are many spikes in each trial ($n_j \gg 1$), so that the last two terms in Eq. (17) are negligible.

If the firing rate $\lambda(t)$ is not known, an estimated firing rate $\hat{\lambda}(t)$ may be used, and the maximum likelihood estimator (MLE) $(\hat{\phi}, \hat{\alpha})$ is obtained by maximizing Eq. (17) with respect to (ϕ, α) . The MLE does not generally admit closed form solutions, and is obtained by maximizing the log likelihood function numerically. In the numerical studies, we use a rectangular sliding window (27) to compute $\hat{\lambda}(t)$, and use a MATLAB function “fminsearch” to maximize Eq. (17). We will discuss alternative methods for estimating $\lambda(t)$, and for estimating (ϕ, α) together with $\lambda(t)$ rather than separately, in section 4.

The numerical studies in the following section show that the distribution of $\hat{\phi}$ is right-skewed because $\phi > 0$, but that $\log \hat{\phi}$ is approximately normally distributed (figure 2b). Therefore, we consider the variance of $\hat{\eta} = \log \hat{\phi}$. By differentiating the log likelihood (17) with respect to $\eta (= \log \phi)$ and α , the observed information matrix is

obtained as

$$\begin{aligned}
J(\eta, \alpha) &= - \begin{pmatrix} \frac{\partial^2 L}{\partial \eta^2} & \frac{\partial^2 L}{\partial \eta \partial \alpha} \\ \frac{\partial^2 L}{\partial \eta \partial \alpha} & \frac{\partial^2 L}{\partial \alpha^2} \end{pmatrix} \\
&= - \sum_{j=1}^M \sum_{i=2}^{n_j} \left\{ \xi_i^j \frac{\partial^2}{\partial \xi_i^{j2}} \log f(u_i^j; \xi_i^j) + \xi_i^j \frac{\partial}{\partial \xi_i^j} \log f(u_i^j; \xi_i^j) \right\} A_i^j, \quad (18)
\end{aligned}$$

where

$$A_i^j = \begin{pmatrix} 1 & -\log \lambda(t_i^j) \\ -\log \lambda(t_i^j) & \{\log \lambda(t_i^j)\}^2 \end{pmatrix}. \quad (19)$$

Note that the rank of matrix A_i^j is 1, but the rank of $J(\eta, \alpha)$ generally becomes 2.² Further, if the probability density function f satisfies the regularity conditions that ensure asymptotic normality of parameter estimators (Casella and Berger, 2002), the asymptotic variance matrix of the MLE is given by $J(\hat{\eta}, \hat{\alpha})^{-1}$, with which the confidence intervals for $\hat{\eta}$ and $\hat{\alpha}$ are constructed as

$$\hat{\eta} \pm z \sqrt{(J(\hat{\eta}, \hat{\alpha})^{-1})_{11}} \quad (20)$$

and

$$\hat{\alpha} \pm z \sqrt{(J(\hat{\eta}, \hat{\alpha})^{-1})_{22}}, \quad (21)$$

where z is the critical value.

2.3 Choice of ISI density function

The ISI density function $f(x; \phi)$ is one of the building blocks of the proposed statistical model. Any ISI density function with a finite variance represents a generalized rate-modulated renewal process (14). Presently, we use a Tweedie distribution, a special case of an exponential dispersion model (Jorgensen, 1987, 1997). This is a two-parameter distribution, consisting of a linear exponential family with an additional dispersion parameter. These distributions play an important role in statistics, because they are the response distributions for generalized linear models (McCullagh and Nelder, 1989). A Tweedie distribution is an exponential dispersion model that has scale invariance (8),

² The rank of $J(\eta, \alpha)$ becomes 1 when the firing rate is constant.

and includes probability distributions commonly used for describing the ISI variability, such as the gamma (for $\alpha = 2$) and inverse Gaussian (for $\alpha = 3$) distributions as special cases. These properties make a Tweedie distribution an obvious choice for $f(x; \phi)$.

Exponential dispersion models have a probability density function of the form,

$$f(x; \mu, \phi) = c(x, \phi) \exp \left[\frac{1}{\phi} \{x\theta - \kappa(\theta)\} \right], \quad (22)$$

where θ is the canonical parameter, and $\kappa(\theta)$ is the cumulant function, with derivatives being the cumulants of the distribution. In particular, its mean and variance are given by $\mu = \dot{\kappa}(\theta)$ and $\sigma^2 = \phi \ddot{\kappa}(\theta)$, respectively. The mapping from θ to the ISI mean μ is invertible, and is written $\ddot{\kappa}(\theta) = V(\mu)$ for a suitable function $V(\mu)$, called the variance function. A Tweedie distribution is identified by a particular choice of the variance function, as $V(\mu) = \mu^\alpha$. By equating $\ddot{\kappa}(\theta) = d\mu/d\theta = \mu^\alpha$ and solving for μ and κ , θ and κ are obtained as

$$\theta = \begin{cases} \frac{\mu^{1-\alpha}-1}{1-\alpha} & \alpha \neq 1 \\ \log \mu & \alpha = 1 \end{cases}, \quad (23)$$

and

$$\kappa(\theta) = \begin{cases} \frac{\mu^{2-\alpha}-1}{2-\alpha} & \alpha \neq 2 \\ \log \mu & \alpha = 2 \end{cases}, \quad (24)$$

where we chose $\kappa(\theta) = 0$ and $\mu = 1$ at $\theta = 0$, without loss of generality. The factor $c(x, \phi)$ in Eq. (22), which is determined by the normalization condition, does not have a closed form, except for in special cases. We compute it numerically, using series expansion and the Fourier inversion formula (Dunn and Smyth, 2005, 2008).

3 Results

In this section, we demonstrate with simulations that the proposed statistical model (14) describes spike trains that have a wide range of variability, characterized by α and ϕ . We illustrate on simulated and experimental data that our inference method is capable of estimating α and ϕ from rate-modulated spike trains.

3.1 Simulation study

First, we simulate spike trains. The probability of a spike occurring in a short interval $(t, t + dt]$ is given by the conditional intensity function (14):

$$P\{N(t+dt) - N(t) = 1; t_{N(t)}, \{\lambda(t)\}, \phi, \alpha\} = r(t; t_{N(t)}, \{\lambda(t)\}, \phi, \alpha)dt + o(dt). \quad (25)$$

Spike trains are simulated by discretizing the time into small bins ($dt = 10^{-5}$), and evaluating Eq. (25) in each bin. We use the firing rate function as

$$\lambda(t) = 40 + 20 \sin \frac{2\pi}{0.5}t, \quad (26)$$

and generate M spike trains in the time interval $t \in (0, 1]$. In order to compute the firing rate and the Fano factor, we use a sliding window of duration $\Delta = 0.125$, in which an average of five spikes are expected to fall. Let $N_{\Delta}^j(t)$ denote the number of spikes of the j th spike train in the counting window centered at t . The firing rate $\hat{\lambda}(t)$ and the Fano factor $\hat{F}_{\Delta}(t)$ in this window are computed, by averaging across trials, as

$$\hat{\lambda}(t) = \frac{1}{M} \sum_{j=1}^M N_{\Delta}^j(t) / \Delta, \quad (27)$$

and

$$\hat{F}_{\Delta}(t) = \frac{1}{M-1} \sum_{j=1}^M \{N_{\Delta}^j(t) - \hat{\lambda}(t)\Delta\}^2 / \hat{\lambda}(t)\Delta. \quad (28)$$

Figure 1 displays the raster plots of 20 spike trains, and $\hat{F}_{\Delta}(t)$ computed with $M = 10^4$ for different α and ϕ . We see that α and ϕ differentiate the variability of spike trains in different manners, as described in Eq. (4): ϕ scales the overall variability of spike trains, while α controls the dependency on the firing rate. The Fano factor is inversely related to the firing rate for $\alpha = 3$ (Figure 1b). For $\alpha = 2$, in which case the proposed model (14) corresponds to the “conventional” rate-modulated renewal process (12), the Fano factor is almost constant, irrespective of the firing rate (Figure 1a). For $\alpha = 2$ and $\phi = 1$, the spike trains become the inhomogeneous Poisson process (Figure 1a2).

We simulated M spike trains, from which the MLE $(\hat{\alpha}, \hat{\phi})$ was computed. We repeated the simulation 10^3 times. Figure 2 shows that $\hat{\alpha}$ and $\log \hat{\phi}$ are approximately

normally distributed, and that they are correlated. Figure 3a plots $\hat{\alpha}$ and $\log \hat{\phi}$ against the number of spike trains M (open circles). It is observed that $\hat{\alpha}$ and $\log \hat{\phi}$ converge to the true values as M increases. The errors in $\hat{\alpha}$ and $\log \hat{\phi}$ are decomposed into the bias and variance, which are shown in figure 3bc. Figure 4 displays the MLE $\hat{\alpha}$ of the exponent against the true value, ranging from $\alpha = 2$ to 3, where $\hat{\alpha}$ was computed with $M = 50$ (open circles). The MLE approximately matches the true value in this range.

For comparison, we computed an empirical estimate of $(\alpha, \log \phi)$, using linear regression of $\{\log \hat{F}_{\Delta}(t)\}$ on $\{\log \hat{\lambda}(t)\}$ from Eqs. (4) and (6). The results are plotted in figures 3 and 4 (crosses). It is confirmed that the bias and variance of $(\hat{\alpha}, \log \hat{\phi})$ are smaller than those predicted by the empirical estimates.

3.2 Experimental data

We apply our method to two experimental datasets. One dataset, labeled “nsa2004.1”, is publicly available from the Neural Signal Archive (Britten et al., 2004). The spike data was recorded from 216 neurons in the visual cortical area MT of adult rhesus macaques. The recordings were obtained while a visual stimulus, consisting of a dynamic random dot pattern, was presented. Further experimental details can be found in Britten et al. (1992). The other dataset, labeled “ia-1”, is available from the CRCNS data sharing website (Rokem et al., 2009). Spike trains were recorded from 43 auditory receptor cells of grasshoppers, while an auditory stimulus consisting of random amplitude modulations of wave was presented. See Rokem et al. (2006) for more details.

Both datasets were divided into sub-datasets, consisting of multiple spike trains recorded from one cell under identical stimulus conditions. We selected sub-datasets containing ≥ 50 trials, and with the mean firing rate ≥ 10 spikes/s, due to the sufficiency of spikes for the analysis. Consequently, 193 sub-datasets for nsa2004.1 and 138 sub-datasets for ia-1 were used. Representative sub-datasets for nsa2004.1 and ia-1 are shown in Figure 5, together with the estimated firing rate $\hat{\lambda}(t)$ and Fano factor $\hat{F}_{\Delta}(t)$, computed with the sliding window whose length Δ was taken so that an average of five spikes are encompassed.

For each sub-dataset, we obtained the MLE $(\hat{\alpha}, \hat{\phi})$. Figure 6a shows a scatter plot of $(\hat{\alpha}, \log \hat{\phi})$ (open circles stand for nsa2004.1 and crosses stand for ia-1). The mean

and standard deviations of the MLE are $\hat{\alpha} = 2.43 \pm 0.38$ and $\log \hat{\phi} = 1.52 \pm 1.69$ for nsa2004.1, and $\hat{\alpha} = 2.96 \pm 0.58$ and $\log \hat{\phi} = 3.37 \pm 2.46$ for ia-1. It is observed that a large portion of the $\hat{\alpha}$ are greater than two, and that on average $\hat{\alpha}$ of ia-1 is greater than that of nsa2004.1. This indicates that the firing variability tends to decrease as the firing rate increases, and that this tendency is stronger in ia-1 than in nsa2004.1. In order to confirm this result, we estimated the exponent γ of the Fano factor empirically for each sub-dataset, by performing linear regression of $\{\log \hat{F}_{\Delta}(t)\}$ on $\{\log \hat{\lambda}(t)\}$ (see Figure 5a2,b2). The estimated exponents, γ , are $\hat{\gamma} = -0.17 \pm 0.51$ for nsa2004.1 and $\hat{\gamma} = -0.94 \pm 0.33$ for ia-1. Figure 6b plots $\hat{\gamma}$ against $\hat{\alpha}$ in each of the sub-datasets, showing that the individual sub-datasets scatter around the line (6).

4 Discussion

This article was concerned the variability of spike trains, described by the power mean-variance relationship (1). It was shown that this single assumption allows the spike trains to have a wide range of variability, characterized by ϕ and α . By combining the power law with the time-rescaling transformation, we proposed generalized rate-modulated renewal processes, based on which a statistical method was developed for inferring (ϕ, α) from rate-modulated spike trains.

In our method, the firing rate $\lambda(t)$ was estimated separately from (ϕ, α) , using a rectangular counting window. We could use other methods, such as kernel density estimators or spline methods, which produce more precise rate estimates (Kass et al., 2003; Shimazaki and Shinomoto, 2010). Alternatively, one may estimate $\lambda(t)$ together with (ϕ, α) , rather than separately. A principled method is to adopt a Bayesian framework, introducing a prior process of $\lambda(t)$ for regularization and computing the posterior process. Parameters of the prior process and (ϕ, α) can be simultaneously optimized, by maximizing the marginal likelihood or the evidence (Cunningham et al., 2008; Koyama and Shinomoto, 2005; Koyama et al., 2013), which may improve the statistical efficiency.

It is often assumed that the variance of spike counts is proportional to their mean (Averbeck, 2009), where the coefficient of proportionality (which corresponds to the Fano factor) may differ from unity due to a deviation from Poisson spiking. In our

formulation, this assumption is relaxed, and we adopt one that the ratio of the count variance to the mean changes with the firing rate (Eq. (4)), which is observed in a wide range of brain regions (Kara et al., 2000).

The degree of irregularity of neural firing, which is measured by ISI statistics such as the *local variation* L_V (Shinomoto et al., 2003), is generally maintained *in vivo* cortical areas, while the firing rate varies in time (Maimon and Assad, 2009; Shinomoto et al., 2009). This implies that the exponent of the power law (2) in the ISI statistics is $\alpha \approx 2$, from which a linear relationship between the mean and variance of spike counts ($\gamma \approx 0$) is expected. On the other hand, steady discharges of X retinal ganglion cells, in response to stationary visual patterns, approximately obey the power law with $\alpha \approx 3$ (Troy and Robson, 1992), implying that a fixed ratio of the variance to the mean spike count no longer holds, but that the spike counts are less variable at higher rates (Berry and Meister, 1998; Reich et al., 1998).

In the nervous system, neurons produce an action potential by integrating presynaptic inputs within tens of milliseconds, in which typically a few spikes come from each presynaptic neuron. This implies that the variance of spike counts in the integration time exhibits the power law, so that the presynaptic inputs have *signal-dependent noise*³ that may be relevant to the computation carried out by the nervous system. Ma et al. (2006) hypothesized that the Poisson-like statistics in the responses of populations of cortical neurons may represent probability distributions over the stimulus, and implemented Bayesian inferences using linear combinations of the responses. A necessary condition in their hypothesis, which makes the Bayesian inferences possible, is that the variance of spike counts is proportional to the mean spike count ($\gamma = 0$). Lu et al. (2013) showed that in controlling dynamical systems with noisy signals, precise control is achievable if the control signal has sub-Poisson noise ($\gamma < 0$), while it is not achievable if the control signal has Poisson or supra-Poisson noise ($\gamma \geq 0$).

By analyzing a stochastic leaky integrate-and-fire model, we provided a possible mechanistic explanation for the origin of the power law (1) with various exponents

³ With a temporal resolution of this integration time, spike trains may be described as

$$\frac{dN(t)}{dt} \approx \lambda(t) + \xi(t),$$

where $\xi(t)$ is a white noise with $E[\xi(t)] = 0$ and $E[\xi(t)\xi(s)] = \phi\lambda(t)^{\gamma+1}\delta(t-s)$.

(Koyama, 2014): $\alpha = 3$ may imply a supra-threshold firing regime, in which firing is driven by excitatory input; $\alpha = 2$ may be interpreted as a sub-threshold firing regime, in which the membrane potential fluctuates below the threshold; and $\alpha = 1$ may emerge when firing is strongly caused by large fluctuations of the membrane potential. Therefore, it is speculated that the “intrinsic” exponent may reflect electrophysiological properties of individual cells or dynamical states of networks, and may vary across different brain areas. The proposed statistical framework offers a systematic way to explore the diversity of the variability of neural responses.

A Derivation of the likelihood function

In this appendix, we derive Eq. (16) from Eq. (15). Using Eq. (14), the second factor in the rhs of Eq. (15) is rewritten as

$$\begin{aligned} & \prod_{i=2}^n r(t_i; t_{i-1}, \{\lambda(t)\}, \phi, \alpha) \\ &= \prod_{i=2}^n \lambda(t_i) f(\Lambda(t_i) - \Lambda(t_{i-1}); \lambda(t_i)^{2-\alpha} \phi) \\ & \quad \times \prod_{i=2}^n \left[1 - \int_{t_{i-1}}^{t_i} \lambda(v) f(\Lambda(v) - \Lambda(t_{i-1}); \lambda(v)^{2-\alpha} \phi) dv \right]^{-1}. \end{aligned} \quad (29)$$

Taking the derivative of the logarithm of the last factor in Eq. (29) leads to

$$\begin{aligned} & \frac{d}{dt_i} \log \left[1 - \int_{t_{i-1}}^{t_i} \lambda(v) f(\Lambda(v) - \Lambda(t_{i-1}); \lambda(v)^{2-\alpha} \phi) dv \right] \\ &= - \frac{\lambda(t_i) f(\Lambda(t_i) - \Lambda(t_{i-1}); \lambda(t_i)^{2-\alpha} \phi)}{1 - \int_{t_{i-1}}^{t_i} \lambda(v) f(\Lambda(v) - \Lambda(t_{i-1}); \lambda(v)^{2-\alpha} \phi) dv} \\ &= -r(t_i; t_{i-1}, \{\lambda(t)\}, \phi, \alpha), \end{aligned} \quad (30)$$

where the last equality comes from Eq. (14). Thus, we obtain

$$\begin{aligned} & \prod_{i=2}^n \left[1 - \int_{t_{i-1}}^{t_i} \lambda(v) f(\Lambda(v) - \Lambda(t_{i-1}); \lambda(v)^{2-\alpha} \phi) dv \right] \\ &= \exp \left(- \int_{t_1}^{t_n} r(u; t_{N(u)}, \{\lambda(t)\}, \phi, \alpha) du \right). \end{aligned} \quad (31)$$

Substituting Eqs. (29) and (31) into Eq. (15) leads to Eq. (16).

Acknowledgments

I would like to thank K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon for uploading their experimental data to the Neural Signal Archive. I am also indebted to A. Rokem for collecting the data at the lab of A. Herz and providing it through the CRCNS program. This research was supported by JSPS KAKENHI Grant Number 24700287.

References

- Averbeck, B. B. (2009). Poisson or not Poisson: Differences in spike train statistics between parietal cortical areas. *Neuron*, 62:310–311.
- Barbieri, R., Quirk, M. C., Frank, L. M., Wilson, M. A., and Brown, E. N. (2001). Construction and analysis of non-Poisson stimulus-response models of neural spiking activity. *J Neuroscience Methods*, 105:25–37.
- Berman, M. (1981). Inhomogeneous and modulated gamma processes. *Biometrika*, 68:143–152.
- Berry, M. J. and Meister, M. (1998). Refractoriness and neural precision. *Journal of Neuroscience*, 18:2200–2211.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, 12:4745–4765.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (2004). Responses of single neurons in macaque MT/V5 as a function of motion coherence in stochastic dot stimuli. Neural Signal Archive, nsa2004.1. Available online at <http://www.neuralsignal.org>.
- Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Thomson Learning, 2nd edition.

- Cox, D. R. (1962). *Renewal Theory*. Chapman and Hal.
- Cunningham, J. P., Yu, B. M., Shenoy, K. V., and Sahani, M. (2008). Inferring neural firing rates from spike trains using Gaussian processes. In *Neural Information Processing Systems*, volume 20, pages 329–336.
- Daley, D. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*, volume 1. Springer-Verlag.
- de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R., and Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science*, 275:1805–1808.
- Dunn, P. K. and Smyth, G. K. (2005). Series evaluation of Tweedie exponential dispersion model densities. *Statistics and Computing*, 15:267–280.
- Dunn, P. K. and Smyth, G. K. (2008). Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. *Statistics and Computing*, 18:73–86.
- Harris, C. M. and Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394:780–784.
- Jorgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society. Series B*, 49:127–162.
- Jorgensen, B. (1997). *The Theory of Dispersion Models*. Chapman and Hall.
- Kara, P., Reinagel, P., and Reid, R. C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, 27:635–646.
- Kass, R. E. and Ventura, V. (2001). A spike-train probability model. *Neural Computation*, 13:1713–1720.
- Kass, R. E., Ventura, V., and Cai, C. (2003). Statistical smoothing of neuronal data. *Network: Computation in Neural Systems*, 14:5–15.
- Koyama, S. (2014). Fluctuation scaling in point processes. [arXiv.org:1409.6800](https://arxiv.org/abs/1409.6800) [physics.data-an].

- Koyama, S. and Kass, R. E. (2008). Spike train probability models for stimulus-driven leaky integrate-and-fire neurons. *Neural Computation*, 20:1776–1795.
- Koyama, S. and Kostal, L. (2014). The effect of interspike interval statistics on the information gain under the rate coding hypothesis. *Mathematical Biosciences and Engineering*, 11:63–80.
- Koyama, S., Omi, T., Kass, R. E., and Shinomoto, S. (2013). Information transmission using non-poisson regular firing. *Neural Computation*, 25:854–876.
- Koyama, S. and Shinomoto, S. (2005). Empirical bayes interpretations of random point events. *Journal of Physics A: Mathematical and General*, 38:L531–L537.
- Lu, W., Feng, J., Amari, S., and Waxman, D. (2013). Achieving precise mechanical control in intrinsically noisy systems. *New Journal of Physics*, 15:063012.
- Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9:1432–1438.
- Maimon, G. and Assad, J. A. (2009). Beyond Poisson: Increased spike-time regularity across primate parietal cortex. *Neuron*, 62:426–440.
- Mainen, Z. F. and Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268:1503–1506.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman and Hall, 2 edition.
- Nawrot, M. P., Boucsein, C., Rodriguez-Molina, V., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, 169:374–390.
- Pillow, J. W. (2008). Time-rescaling methods for the estimation and assessment of non-poisson neural encoding models. In *Neural Information Processing Systems*, volume 22, pages 1473–1481.
- Reich, D. S., Victor, J. D., and Knight, B. W. (1998). The power ratio and the interval map: spiking models and extracellular recordings. *Journal of Neuroscience*, 18:10090–10104.

- Rokem, A., Watzl, S., Gollisch, T., Stemmler, M., Herz, A. V. M., and Samengo, I. (2006). Spike-timing precision underlies the coding efficiency of auditory receptor neurons. *Journal of Neurophysiology*, 95:2541–2552.
- Rokem, A., Watzl, S., Gollisch, T., Stemmler, M., Herz, A. V. M., and Samengo, I. (2009). Recording from grasshopper (*Locusta Migratoria*) auditory receptor cells. CRCNS.org. <http://dx.doi.org/10.6080/K0BG2KWB>.
- Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18:3870–3896.
- Shimazaki, H. and Shinomoto, S. (2010). Kernel bandwidth optimization in spike rate estimation. *Journal of Computational Neuroscience*, 29:171–182.
- Shinomoto, S., Kim, H., Shimokawa, T., Matsuno, N., Funahashi, S., Shima, K., Fujita, I., Tamura, H., Doi, T., Kawano, K., Inaba, N., Fukushima, K., Kurkin, S., Kurata, K., Taira, M., Tsutsui, K., Komatsu, H., Ogawa, T., Koida, K., Tanji, J., and Toyama, K. (2009). Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Computational Biology*, 5:e1000433.
- Shinomoto, S., Shima, K., and Tanji, J. (2003). Differences in spiking patterns among cortical neurons. *Neural Computation*, 15:2823–2842.
- Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T., and Kaplan, E. (1997). Fractal character of the neural spike train in the visual system of the cat. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 14:529–546.
- Troy, J. B. and Robson, J. G. (1992). Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance. *Visual Neuroscience*, 9:535–553.

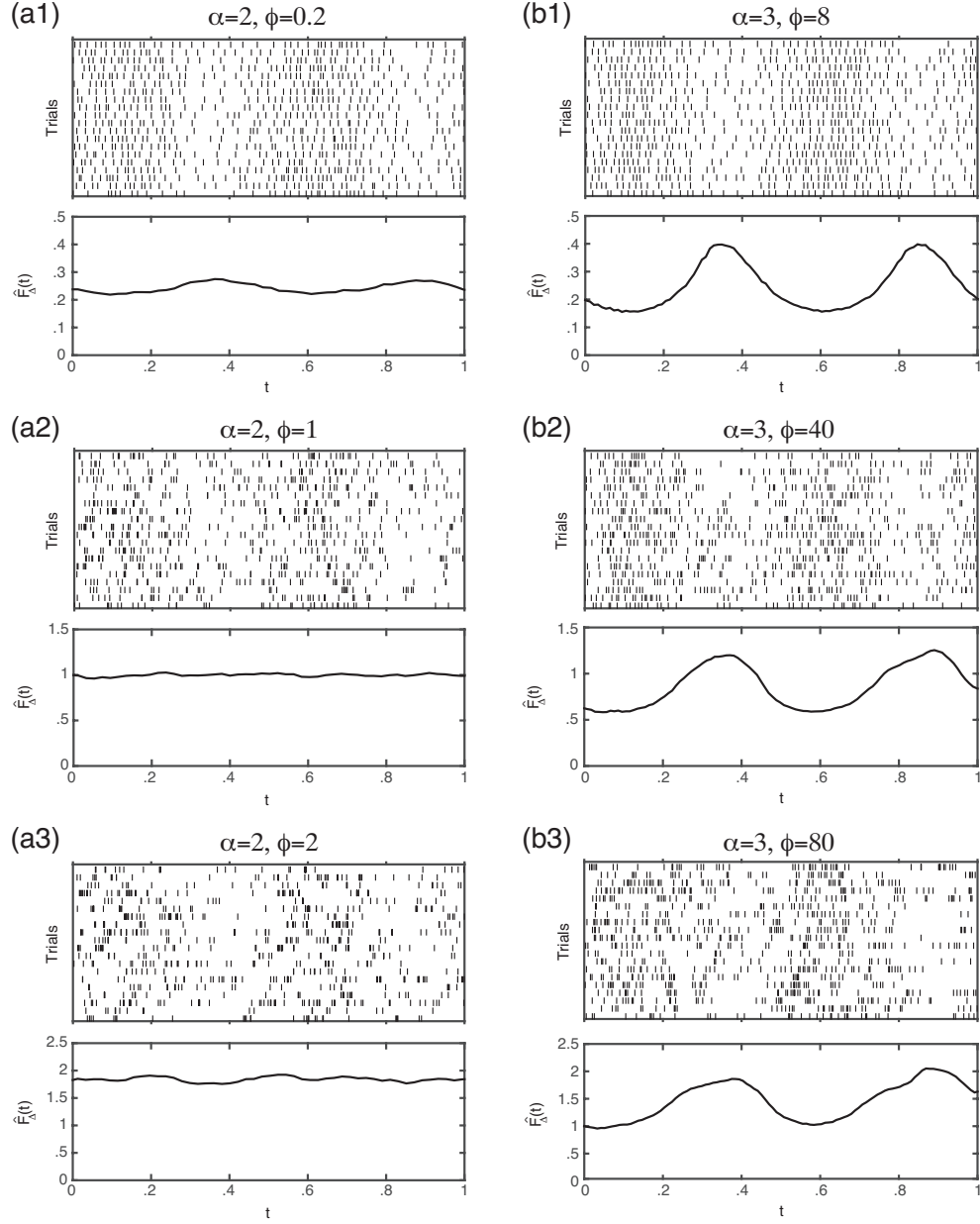


Figure 1: Raster plots of 20 spike trains simulated using Eqs. (25) and (26) and estimated Fano factor $\hat{F}_\Delta(t)$, for different α and ϕ . The Fano factor is almost constant for $\alpha = 2$ (a1,2,3), while it is inversely related to the firing rate for $\alpha = 3$ (b1,2,3). The overall Fano factor increases as ϕ increases, for fixed α .

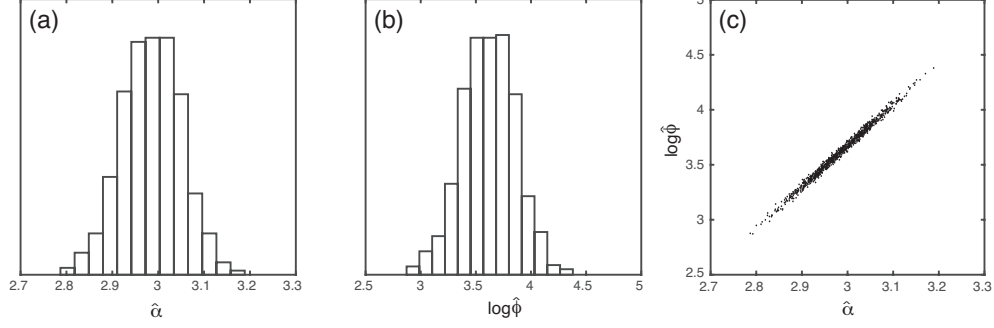


Figure 2: Histograms of $\hat{\alpha}$ (a) and $\log \hat{\phi}$ (b), obtained by 10^3 repeated simulations, with $M = 100$ spike trains. Both $\hat{\alpha}$ and $\log \hat{\phi}$ are approximately normally distributed. (c) presents a scatter plot of $(\hat{\alpha}, \log \hat{\phi})$, showing that $\hat{\alpha}$ and $\log \hat{\phi}$ are linearly related.

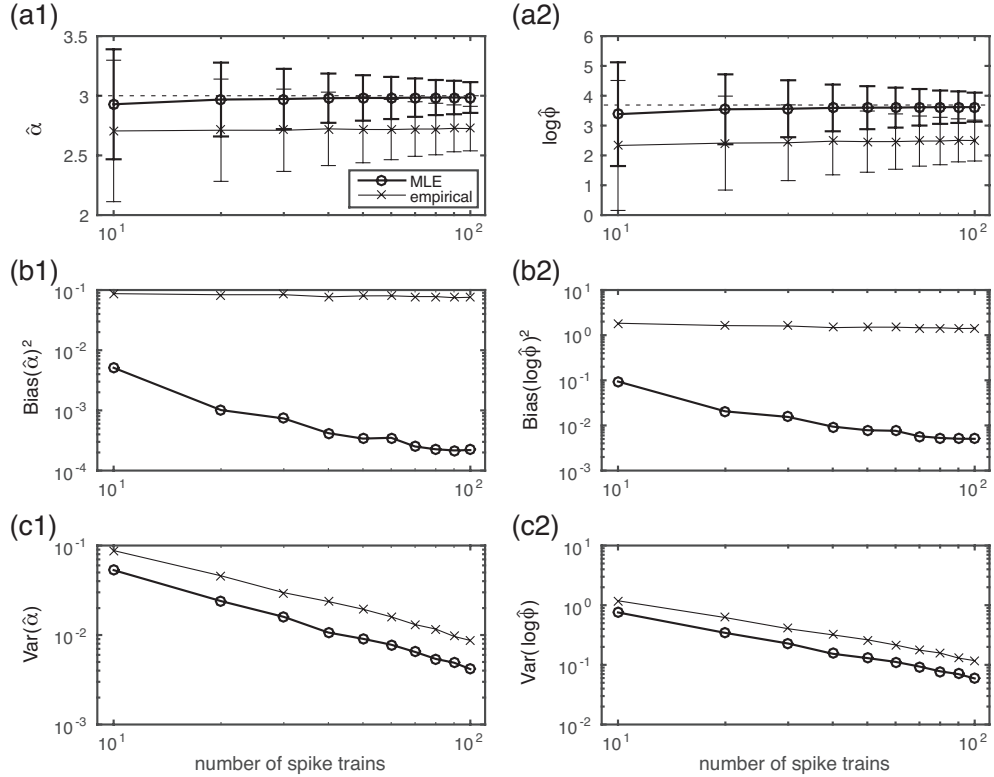


Figure 3: Estimates of α and $\log \phi$ as functions of the number of spike trains M . Results in this figure were computed by averaging across 10^3 repeated simulations for each M . Open circles and crosses represent the MLE and the empirical estimate, respectively. The true parameters are $\alpha = 3$ and $\log \phi = \log 40 (\approx 3.69)$, represented by dashed lines in (a). (a) presents the mean and 2SD of $\hat{\alpha}$ and $\log \hat{\phi}$. The errors in $\hat{\alpha}$ and $\log \hat{\phi}$ are decomposed into the squared bias (b) and variance (c). Both the bias and variance decrease as the number of spike trains M increases. The bias and variance of the MLEs are smaller than those of the empirical estimates.

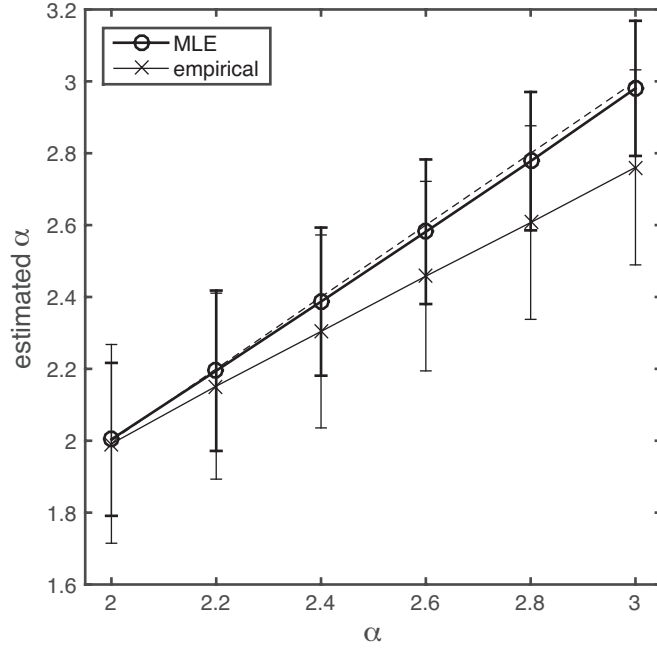


Figure 4: Plot of the MLE (open circles) and empirical estimate (crosses) of α against the true value, for $M = 50$ spike trains. The mean and 2SD error bar were computed by averaging across 10^3 repeated simulations. The dashed line represents the true value. The bias of the empirical estimate increases as α increases, while the mean of $\hat{\alpha}$ approximately matches the true value.

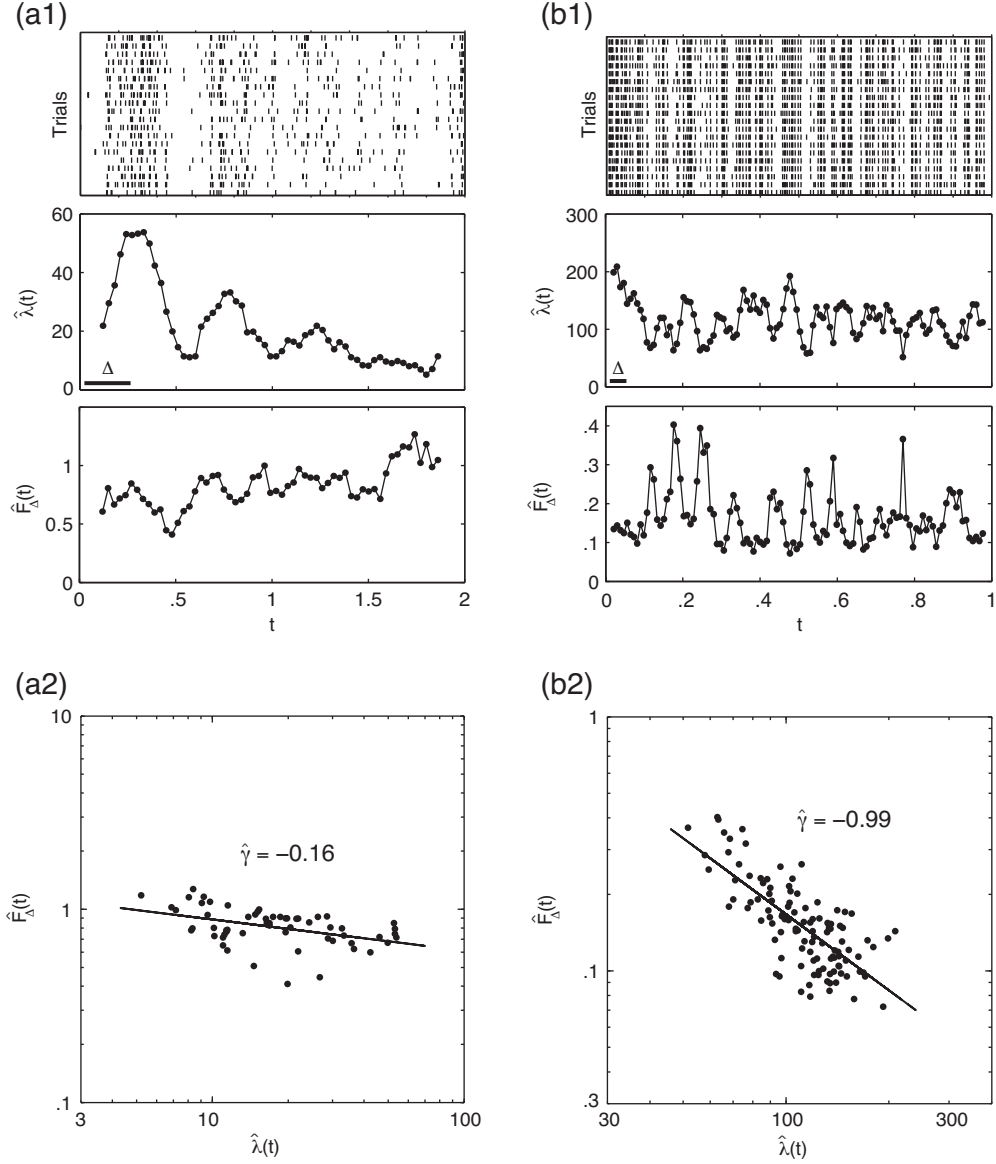


Figure 5: Representative sub-datasets for nsa2004.1 (a1) and for ia-1 (b1). (Top) raster plot of 20 spike trains; (middle) estimated firing rate (the horizontal bar indicates the length of the counting window); (bottom) the Fano factor. The Fano factor is plotted against the firing rate on a log-log scale (a2 for nsa2004.1 and b2 for ia-1), on which linear regression was performed to obtain the exponent $\hat{\gamma}$.

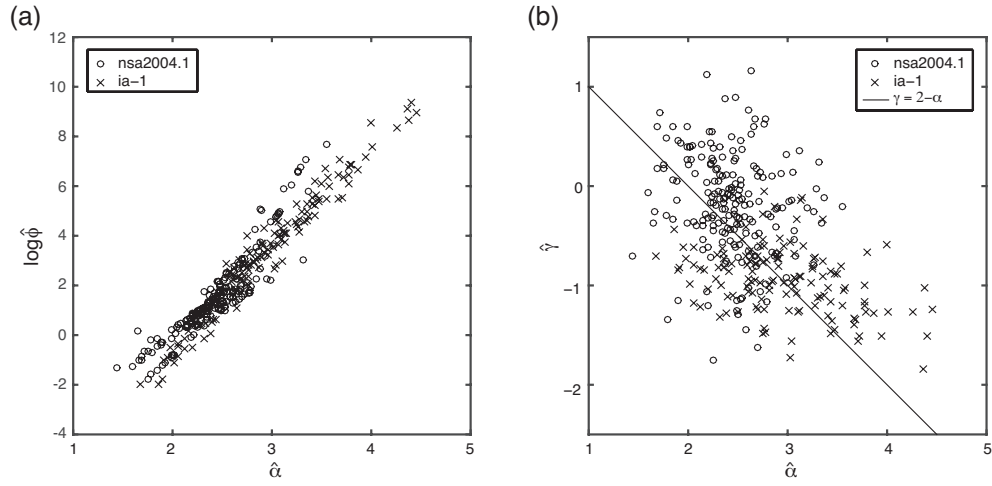


Figure 6: Scatter plot of $(\hat{\alpha}, \log \hat{\phi})$ (a) and $(\hat{\alpha}, \hat{\gamma})$ (b). Open circles represent nsa2004.1 and crosses represent ia-1. The solid line in (b) represents the scaling relation (6), around which the individual sub-datasets scatter.