On the Variability of Cortical Neural Responses: A Statistical Interpretation

Sheng Li and Si Wu

Department of Informatics, Sussex University, UK

Abstract

Neural responses in the cerebral cortex exhibit tremendous variability. Understanding the origin and the functional meaning of this variability is of critical importance for our understanding of neural coding. The present study investigates the neural response variability from the view of statistical inference. We show that high variability can also arise due to the inferential sensitivity in neural coding. This view is support by the simulation on the encoding of nature images.

Key words: neural coding; neural response variability; statistical inference

1 Introduction

External stimuli are encoded in neural responses in the brain. Neural responses have a unique feature, that is, all spikes are of the same stereotyped shape, with the variable quantity being the time interval between successive spikes (ISI). The pattern of ISIs is then the medium that conveys the stimulus information. It is expected that if the relationship between stimuli and the patterns of ISI is revealed, the neural cipher is broken. Motivated by this idea, experiments have recorded a large quantity of data on neuronal responses in different areas of the brain when stimuli are presented. The finding on the responses of a single neuron in the cortex, however, reveals a quite puzzling result: the pattern of ISI seems to be generated by a random process; its correlation with the stimulus is quite low. Indeed, the variability of neural response is so high, when measured by the variance-mean ratio of ISI, the value is typically in the range of $(1, 1.5)$, being comparable or even larger than the corresponding value of the random Poisson process (see e.g., (4; 3)).

Email addresses: `sheng@sussex.ac.uk` (Sheng Li), `siwu@sussex.ac.uk` (Si Wu).

Giving a reasonable justification of the origin and the functional meaning of this high variability is of critical importance for our understanding of neural coding. A large volume of work have been done on this issue. In terms of the origin, there are two possibilities being suggested. One is that the irregularity is generated at the single neuron level, where the stochasticity comes from, e.g., the unreliable release of neuro-transmitter at synapses, the randomness in the channel gating, and the fluctuations in membrane potential (see e.g., (2)). Alternatively, the irregularity may arise at the network level. For instance, as shown by (3; 5), irregular firing can be generated due to the balanced contribution between excitatory and inhibitory inputs. In terms of the function meaning that justifies why the brain works in this highly variable dynamic regime, there are also a few discussions. For instances, it is suggested that the irregularity enables neuron response fast to external stimulus, provides a mechanism similar to stochastic resonance and serves as a framework for encoding the uncertainty of stimulus. Although these studies answer our concern in many aspects, it is still not fully convincing as why neural activity has to be so irregular in the first place. Anyhow, the above mentioned functional advantages can in principle be also implemented in deterministic systems. What are missed are perhaps strong reasonings that directly link the irregularity of neural response with its goal of representing information.

In the present study we will investigate a new source of the irregularity in neural response. Our view is being motivated by a general property in statistical inference, that is, under certain circumstance the inferred result can be very sensitive to noise. An intuitive illustration of this idea is present in Fig.1. Fig.1a exhibits a natural image. Two pictures that are independently down-sampled from this image are shown in Fig.1c and Fig.1d. For a large sampling rate, their discrepancy is small, and is almost visually indistinguishable. We may expect that their neural representations will have similar values. The result turns out to be a big surprise, as shown in Fig.1b, their neural representations are significantly different in the activities of individual neurons (the details in Sec.3). This implies that neural representation (i.e., neural responses) can be very sensitive to the input noise (Here the notion of noise summarizes the minor discrepancy between two pictures). It is this inferential sensitivity we will attribute as a source of variability in neural response. In the below we will analyze and quantify this sensitivity in neural coding.

2 Statistical Inferential Sensitivity

An ultimate goal of neural coding is to represent and retrieve external stimuli in a noisy environment, which, mathematically, is described as statistical inference. Since we are only interested in the general property of statistical inference (the inferential sensitivity), we will only consider a simple model for
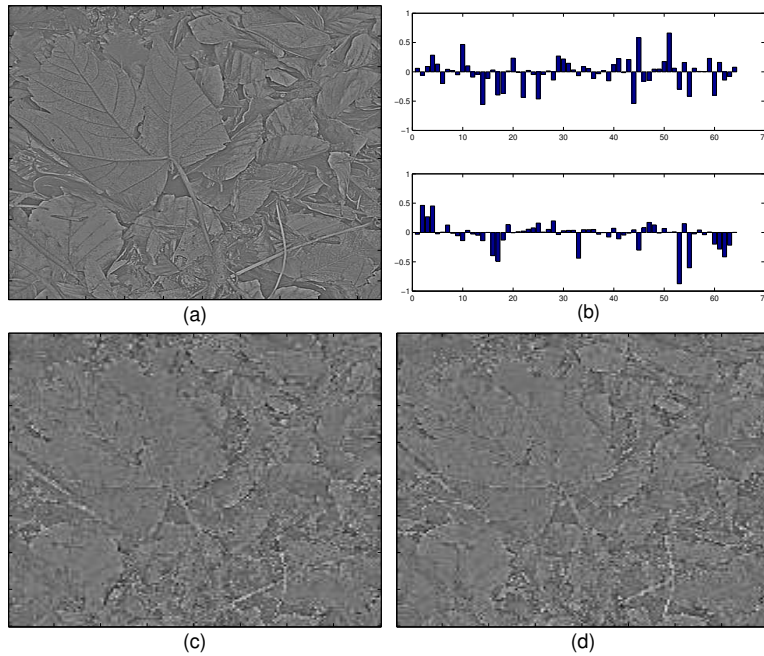
2

Fig. 1. An illustration of inferential sensitivity. (a) A natural image; (b) The neural representations for the two sampled pictures in (c) and (d); (c) and (d) Two copies of independent sampling the image in (a) with the sampling rate 1/9.

.

neural coding. The conclusions to be derived below, however, is applicable to general cases. In particular, we consider that external stimulus is encoded by a linear superposition of basis functions (e.g. (1)),

$$I(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{\phi}(\boldsymbol{x}) = \sum_i a_i \phi_i(\boldsymbol{x}), \tag{1}$$

where $I(\boldsymbol{x})$ denotes the stimulus, and $\boldsymbol{\phi}(\boldsymbol{x}) = \{\phi_i(\boldsymbol{x})\}$ the basis functions which are related to the neuronal receptive fields. The coefficients $\boldsymbol{a} = \{a_i\}$ represents the neural responses, which are the representation of the stimulus in the basis set $\boldsymbol{\phi}$. In practice, the data observed by the estimator is noisy, which can be written by the true value plus a random variable,

$$y_i = I(x_i) + \varepsilon_i, \quad \text{for} \quad i = 1, \ldots, N, \tag{2}$$

where N is the number of data points being observed. The task of neural estimator is to infer the neural representation $\boldsymbol{a}$ from the observed input value $y(\boldsymbol{x})$.

To illustrate the idea, we further simplify the model (just for the purpose of being mathematically trackable) by considering there are only two neurons and the basis functions are Gaussian (Fig.2a), and the noise form is also independent Gaussian. Because of the noise, the finite number of data points,

3

the intrinsic structures of the stimulus, and the decoding method used, the inferred result will generally deviate from the true representation. We can formally prove that in the above model setting, the minimum variability of neural response is given by,

$$P(a_1, a_2) = \frac{1}{Z} \exp\{-\frac{1}{2} \boldsymbol{a}^T \boldsymbol{\Sigma} \boldsymbol{a}\}, \tag{3}$$

where the vector $\boldsymbol{a} = (a_1 - a_1^0, a_2 - a_2^0)$, with a_1^0, a_2^0 being the true representation, and the matrix $\boldsymbol{\Sigma}$ is the Fisher information

$$\boldsymbol{\Sigma} = \frac{1}{\sigma^2} \begin{pmatrix} \sum_i \phi_1(x_i)^2 & \sum_i \phi_1(x_i)\phi_2(x_i) \\ \sum_i \phi_1(x_i)\phi_2(x_i) & \sum_i \phi_2(x_i)^2 \end{pmatrix} \tag{4}$$

The minimum variability means that for any un-biased estimator with the co-variance matrix $\boldsymbol{\Omega}$ of the estimation, $\boldsymbol{\Omega} \geq \boldsymbol{\Sigma}^{-1}$, according to the Cramér-Rao bound. The above distribution quantifies the range and the corresponding probability where the inferred result will fall when a noisy stimulus is presented. Intuitively, the broader the distribution, the larger the fluctuation in the estimation can be expected. High variability also implies the inferred results for two independent samplings of the same stimulus can be dramatically different.

The marginal distributions of (4) quantify the response variability of individual components. They are calculated to be

$$P(a_i) = \frac{1}{Z_i} \exp\{-\frac{1}{2\tau_i^2}(a_i - a_i^0)^2\}, \qquad \text{for} \quad i = 1, 2, \tag{5}$$

with τ_1^2 (similarly for τ_2^2) given by

$$\tau_1^2 = \frac{\sigma^2 \sum_i \phi_2(x_i)^2}{\sum_i \phi_1(x_i)^2 \sum_i \phi_2(x_i)^2 - (\sum_i \phi_1(x_i)\phi_2(x_i))^2}. \tag{6}$$

From eq.(4) and (6), we see that the overlap between the basis functions, measured by $OP = \sum_x \phi_1(x)\phi_2(x)/\sqrt{\sum_x \phi_1(x)^2 \sum_x \phi_2(x)^2}$, plays a crucial role in determining the magnitude of the inferential sensitivity. For a fixed noise level, the larger the overlap, the higher the variability is. The value of variability can even be arbitrarily large depending on the amount of overlap (Fig.2b). This overlapping effect is also intuitively understandable. In the overlapping region, the contributions from different basis functions on the generation of data are obscured. As a result data points in this region are less informative in 'indistinguishing' the activities of neurons. More overlap implies more ambiquity.
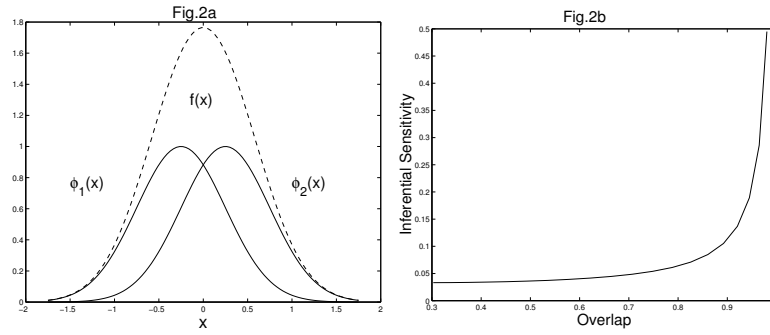
4

Fig. 2. An illustration of a simple model study. (a) The stimulus and the basis functions; (b) Variability vs. Overlap. The noise level is fixed.

3   The Neural Response Variability

The above observation on the inferential sensitivity turns out to have an important implication for neural coding. Typically, stimuli, such as natural images, to be represented/extracted by cortical neuron have a complicated structure that covers a large area of the input space. In order to infer such a coherent structure from noisy data, information from different sites need to be integrated, which computationally requires the receptive fields of neurons to be largely overlapping as is evident in experiment. From the function approximation point of view, this overlapping is essential for accurately interpolating function values between basis centers. Thus, neural responses are inevitably sensitive to input noise. In biological systems noise is ubiquitous. For instance, it may be due to the fluctuation of external stimulus, or be induced during the transfer of neural signals. These noises, even if small, can lead to large variability in neural responses because of inferential sensitivity.

Simulation Experiment. To confirm the above analysis, we do simulation on the encoding of natural image. The noisy input is generated by repeatly and independently down-sampling a fixed natural scene. This down-sampling process mimics the time-varying input stream in real-time computation in neural systems. The inference strategy we use is the sparse coding algorithm (1). In the simulation, five different sampling rates are considered. Here the sampling rate is defined as the percentage of data points being used in the inference. At each sampling rate, 100 copies are sampled.

Fig.1 illustrate the effect of inferential sensitivity, which confirms that minor discrepancy in the input can lead to dramatic fluctuations in the output. We may point out that we did not calculate the variance-mean ratio to compare with experimental data. This is because in the sparse coding model (1), neural activity can has negative value, which makes the mean extremely small. It is unfair to do comparison in this case. But consider the mean of the absolute value of neural activity is in the order of $(0, 0.2)$ (not shown here), the variance values in the Table.1 indeed reveal the irregular responses.

Table 1
Variance vs. Sampling Rate

Sampling rate	1/4	1/9	1/16	1/25	1/36
Variance	0.0306	0.0496	0.0576	0.0615	0.0627
Mean	0.0004	0.0010	0.0013	0.0006	0.0014

4 Conclusions

In the present study we investigate a novel source of irregularity in neural response from the perspective of statistical inference. It shows that irregularity can also arise due to the inferential sensitivity of neural coding. More precisely, it is the overlap between neuronal receptive fields that account for this sensitivity. This overlapping, however, is unavoidable due to that the neural estimator needs to extract the smooth structure of stimulus. This interpretation is different from the others in the literature, which more concentrate on the dynamical process of spike generation at either a single neuron or network level. These views, however, are not necessarily contradictory, instead they supplement with each other to give us a more complete picture of how neural response variability is generated.

References

[1] B. Olshausen and D. Field (1996) Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Nature Images. *Nature*, p.607, v.381.

[2] H. Robinson (2003) The Biophysical Basis of Firing Variability in Cortical Neurons. In: *Computational Neuroscience: A Comprehensive Approach*, ed. by J. Feng. CRC Press.

[3] M. N. Shadlen and W. Newsome (1998) The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *The Journal of Neuroscience*, p.3870-3896, v.15.

[4] W. Softky and C. Koch (1993) The Highly Irregular Firing of Cortical Cells is Inconsistent with Temporal Integration of Random EPSPs. *Journal of Neuroscience*, p.334-350, v.13.

[5] C. van Vreeswijk and H. Sompolinsky (1996) Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science*, p.1724-1726, v.274.