Asymptotic behaviors of population codes
Si Wu\textsuperscript{a,b,*}, Shun-ichi Amari\textsuperscript{b}, Hiroyuki Nakahara\textsuperscript{b}

\textsuperscript{a}Department of Computer Science, Sheffield University, UK
\textsuperscript{b}RIKEN Brain Science Institute, Saitama, Japan

Abstract

The present study investigates asymptotic behaviors of population codes, i.e., their performances when the number of neurons goes to infinity. Since the number of neurons involved in population coding is often large, asymptotic behaviors are the properties of biologic relevance. This study is based on a neural field model for the encoding process, which has a simple form, whereas, covers a broad range of important correlation cases. The asymptotic behaviors of both Fisher information and maximum likelihood type of decoding methods are clarified. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Population coding; Asymptotic behavior; Fisher information; Maximum likelihood; Center of mass

1. Introduction

Population coding is a paradigm used in the brain to represent stimuli by using the joint activities of a number of neurons. Since the number of neurons involved in coding often becomes large (can be approximated as infinity in the calculation), here we are interested in the asymptotic behaviors of population codes, i.e, their performances when the number of neurons goes to infinity.

It is well known that if neural activities are independent of each other or very weakly correlated, the asymptotic behavior of population codes is trivial, the decoding accuracy (measured by the variance of decoding error) is inversely proportional to the number of neurons. However, if neural activities are strongly correlated, this may no longer be the case.

\textsuperscript{*} Corresponding author.

E-mail addresses: s.wu@dcs.shef.ac.uk (S. Wu), amari@brain.riken.go.jp (S.-i. Amari), hiro@brain.riken.go.jp (H. Nakahara).

0925-2312/02/$-see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0925-2312(02)00460-5
The asymptotic behaviors of population codes were studied in [1,5] through analyzing the Fisher information. (According to the Cramér–Rao bound, the inverse of Fisher information is the optimal accuracy for any unbiased estimator to achieve.) An interesting phenomenon found by them is that if the correlation is strong enough, the Fisher information saturates without increasing with the number of neurons. This finding leads to the suspicion on the general belief that the brain uses more neurons to improve the accuracy of information processing. The asymptotic behavior of a decoding method can also become non-trivial in strong correlation. This was first pointed out in our work [4], in which by simulation we show that the usual picture of Gaussian convergence of decoding error does not hold for some strong correlation structures.

In this paper, we systematically study the asymptotic behaviors of population codes for correlated signals, including both Fisher information and the maximum likelihood inference (MLI) type of decoding methods. To carry out the theoretic analysis, a general model for the encoding process in a continuous neural field is proposed, in which neural activities are assumed to be pair-wise correlated with a strength distribution given by a Gaussian function of their difference in preferred stimuli. Compared with other prototypes in the literature [1,5], this model shows advantage of simplifying the calculation and provides a clear picture of results.

2. Encoding model

The encoding model we consider is a one-dimensional neural field, in which neurons are located with uniform density $\rho$. The activity of neuron at position $c$ (i.e., the preferred stimulus) is denoted by $r(c)$, whose mean value is given by a Gaussian tuning function

$$f(c - x) = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{(c-x)^2}{2a^2}\right),$$

where $a$ is the tuning width.

The population activity, $r = \{r(c)\}$, is assumed to be Gaussian pair-wise correlated, that is, the probability of observing $r$ given the stimulus $x$ is

$$Q(r|x) = \frac{1}{Z} \exp\left\{-\frac{\rho^2}{2\sigma^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [r(c) - f(c-x)]h^{-1}(c,c')ight\}$$

$$\times [r(c') - f(c'-x)] dc dc',$$

where $\rho$ is the neural density, $\sigma$ the noise strength and $Z$ the normalization factor.

The correlation structure is determined by the covariance function $h(c,c')$, which is also assumed to be of a Gaussian form

$$h(c,c') = \rho(1 - \beta)\delta(c - c') + \rho^2 \beta e^{-(c-c')^2/2b^2},$$

where $\delta$ is the Kronecker delta and $b$ are the parameters of the Gaussian.
where \( b \) is the correlation length. This correlation model contains a number of important cases.

- **No correlation:** When \( b = 0 \), neurons are uncorrelated.
- **Local correlation:** When \( b \) is of order \( 1/\rho \), that is, \( b = m/\rho \) for a fixed \( m \), neurons are correlated only within \( m \) neighboring neurons.
- **Short-range correlation:** When the correlation length is much longer than \( 1/\rho \) but shorter than \( \sqrt{2} \) times the width of the tuning function, i.e., \( 1/\rho \ll b < \sqrt{2}a \), neurons are correlated over a short range.
- **Wide-range correlation:** When \( b \gtrsim \sqrt{2}a \), neurons are correlated over a wide range.
- **Uniform correlation:** When the correlation length \( b \rightarrow \infty \), neurons are uniformly correlated.

### 3. Asymptotic behaviors of Fisher information

The Fisher information for the encoding model \( Q(r|x) \) is defined as

\[
I_F(x) = - \int Q(r|x) \frac{d^2 \ln Q(r|x)}{dx^2} \, dr. \tag{4}
\]

By using Eq. (2) and the Fourier transformations of \( f(c - x) \) and \( h^{-1}(c, c') \), we get

\[
I_F(x) = \frac{\rho^2}{2\pi\sigma^2} \int_{-\infty}^{\infty} \frac{\omega^2 e^{-\omega^2}}{\rho(1 - \beta) + \rho^2 \sqrt{2\pi}b e^{-b^2/2}} \, d\omega. \tag{5}
\]

Fig. 1 shows the asymptotic behaviors of Fisher information (the value of Eq. (5)) in different correlation cases. For no correlation and uniform correlation (Fig. 1(a)), the Fisher information increases in proportion to the neural density \( \rho \). For local correlation (Fig. 1(b)), the behavior is the same as that in Fig. 1(a), since the correlation decreases quickly with the increment of \( \rho \). For the short-range correlation (Fig. 1(c)), the Fisher information saturates. This result agrees with that in [1,5]. For the wide-range correlation, the Fisher information increases without limitation. This is a new interesting finding.

### 4. Asymptotic behaviors of decoding methods

We investigate three decoding methods. All of them are formulated as the MLI type, whereas, they differ in the probability models being used for decoding.

A MLI type estimator \( \hat{x} \) is obtained through maximization of the presumed log likelihood \( \ln P(r|x) \), i.e., by solving

\[
\nabla \ln P(r|\hat{x}) = 0, \tag{6}
\]
where $P(r|x)$ is called the decoding model. It can be different from the real encoding model $Q(r|x)$ in two scenarios. One is that the decoding system does not know the exact information of encoding one. The other is that a simple and robust decoding model is desirable from the computational point of view.

- The first method is the conventional MLI, referred to as FMLI, which utilizes all of the encoding information, i.e., the decoding model is the true encoding one,

$$P(r|x) = Q(r|x).$$  \hspace{1cm} (7)

- The second method, referred to as UMLI, utilizes the information of the tuning function, but neglects the neural correlation, that is,

$$P(r|x) = \frac{1}{Z_U} \exp \left\{ -\frac{\rho}{2\sigma^2} \int_{-\infty}^{\infty} [r(c) - f(c - x)]^2 \, dc \right\}. \hspace{1cm} (8)$$

- The third method, referred to as COM, does not utilize any information of the encoding process, but instead it assumes an incorrect but simple tuning function.
Table 1
Asymptotic behaviors of three decoding methods

<table>
<thead>
<tr>
<th></th>
<th>FMLI</th>
<th>UMLI</th>
<th>COM</th>
</tr>
</thead>
<tbody>
<tr>
<td>No correlation</td>
<td>FE</td>
<td>QFE</td>
<td>QFE</td>
</tr>
<tr>
<td>Local correlation</td>
<td>FE</td>
<td>QFE</td>
<td>QFE</td>
</tr>
<tr>
<td>Short-range correlation</td>
<td>Non-F</td>
<td>Non-F</td>
<td>Non-F</td>
</tr>
<tr>
<td>Wide-range correlation</td>
<td>Non-F</td>
<td>Non-F</td>
<td>Non-F</td>
</tr>
<tr>
<td>Uniform correlation</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
</tr>
</tbody>
</table>

It also disregards correlations by using

\[
P(r|x) = \frac{1}{Z_C} \exp \left\{ -\frac{\rho}{2\sigma^2} \int_{-\infty}^{\infty} [r(c) - \tilde{f}(c-x)]^2 dc \right\},
\]

where, \( \tilde{f}(c-x) = -(x-c)^2 + \text{const} \) is the presumed tuning function.

Note that the third method is equivalent to the conventional center-of-mass strategy [2].

The asymptotic behaviors of three methods are listed in Table 1 (for more details, please refer to [3]), where FE denotes Fisher efficient, QFE quasi-Fisher efficient, and non-F non-Fisherian. FMLI is called FE (or asymptotically efficient), if its decoding error asymptotically (in the order of \( 1/N \)) satisfies a Gaussian distribution with the variance being the inverse of Fisher information, that is, FMLI achieves the Cramér–Rao bound. UMLI or COM is called QFE, if its error asymptotically satisfies a Gaussian distribution, though the variance is normally larger than the Cramér–Rao bound since less information is used.

When a method is non-F, its decoding error satisfies the Cauchy type of distribution, and the normal picture of Gaussian convergence fails. For the Cauchy distribution, its variance does not exist. In this case, we should be careful of using variance to measure the decoding accuracy. Specifically, when FMLI is non-F, the Cramér–Rao bound is not achievable. One should be careful for carrying analysis based on the Fisher information. From Table 1, we see that for the short- and wide-ranges of correlation, all three methods are non-F.

5. Conclusion

The present study investigates the asymptotic behaviors of population codes for correlated signals.

We first study the Fisher information, and find that it increases with the number of neurons when the correlation covers the wide-range of population. This is an important complement to the result in [1,5], which states that the Fisher information saturates in the short-range correlation.

We then clarify the asymptotic efficiencies of three MLI type of decoding methods, and find that when the correlation covers a non-local range of population (excluding
the uniform correlation and weak noise), the MLI type of method is non-Fisherian, that is, the decoding error does not satisfy the Gaussian distribution but instead the Cauchy type. This finding has important implication on experimental data analysis.

References