LETTER Communicated by Peter Dayan

Population Coding and Decoding in a Neural Field: A Computational Study

Si Wu
s.wu@dcs.ac.uk
RIKEN Brain Science Institute, Wako-shi, Saitama, Japan, and Department of Computer Science, Sheffield University, U.K.

Shun-ichi Amari
amari@brain.riken.go.jp
RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

Hiroyuki Nakahara
hiro@brain.riken.go.jp
RIKEN Brain Science Institute, Wako-shi, Saitama, Japan, and Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, Japan

This study uses a neural field model to investigate computational aspects of population coding and decoding when the stimulus is a single variable. A general prototype model for the encoding process is proposed, in which neural responses are correlated, with strength specified by a gaussian function of their difference in preferred stimuli. Based on the model, we study the effect of correlation on the Fisher information, compare the performances of three decoding methods that differ in the amount of encoding information being used, and investigate the implementation of the three methods by using a recurrent network. This study not only rediscovers main results in existing literatures in a unified way, but also reveals important new features, especially when the neural correlation is strong. As the neural correlation of firing becomes larger, the Fisher information decreases drastically. We confirm that as the width of correlation increases, the Fisher information saturates and no longer increases in proportion to the number of neurons. However, we prove that as the width increases further—wider than $\sqrt{2}$ times the effective width of the turning function—the Fisher information increases again, and it increases without limit in proportion to the number of neurons. Furthermore, we clarify the asymptotic efficiency of the maximum likelihood inference (MLI) type of decoding methods for correlated neural signals. It shows that when the correlation covers a nonlocal range of population (excepting the uniform correlation and when the noise is extremely small), the MLI type of method, whose decoding error satisfies the Cauchy-type distribution, is not asymptotically efficient. This implies that the variance is no longer adequate to measure decoding accuracy.

Neural Computation **14, 999–1026** (2002)

1 Introduction

Population coding is a method to represent stimuli by using the joint activities of a number of neurons. Experimental studies have revealed that this coding paradigm is widely used in the sensor and motor areas of the brain. For example, in the visual area MT, neurons are tuned to the moving direction (Maunsell & Van Essen, 1983). In response to an object moving in a particular direction, many neurons in MT fire, with a noise-corrupted and bell-shaped activity pattern across the population. The moving direction of the object is retrieved from the population activity, to be immune from the fluctuation existing in a single neuron's signal.

From the theoretical point of view, population coding, which concerns how information is represented in the brain, is prerequisite for more complex issues of the brain functions. It is also one of a few mathematically well-formulated problems in neuroscience, in the sense that it grasps the essential features of neural coding and yet, is simple enough for theoretic analysis. There has been extensive work to understand population coding theoretically (Abbott & Dayan, 1999; Brunel & Nadal, 1998; Georgopoulos, Kalaska, Caminiti, & Massey, 1982; Nakahara, Wu, & Amari, 2001; Paradiso, 1988; Pouget, Zhang, Deneve, & Latham, 1998; Salinas & Abbott, 1994; Seung & Sompolinsky, 1993; Wu, Nakahara, Murata, & Amari, 2000; Wu, Nakahara, & Amari, 2001; Yoon & Sompolinsky, 1999)—for example, how much information is included in population coding, what the optimal decoding accuracy is given an encoding model, how to construct a simple and yet accurate enough decoding strategy, how to implement a decoding method in a biological network, and what the effect of correlation is on decoding accuracy. These issues have been studied by using various models for the encoding process.

In this article, based on a unified encoding model in a neural field, we systematically study all the above computational aspects. This article is more than a review of the literature; it elucidates the published results in a more transparent way. Furthermore, we explore important new features concerning how the behavior of the Fisher information changes as the width of correlation increases. The efficiencies of various decoding methods are also evaluated.

To study population coding, a prototype model for the encoding process needs to be constructed. We consider a general case that neural activities are correlated, as seen in experimental data (Fetz, Yoyama, & Smith, 1991; Gawne & Richmond, 1993; Lee, Port, Kruse, & Georgopoulos, 1998; Zohary, Shadlen, & Newsome, 1994), under the assumption of gaussian additive noise. The uncorrelated case is handled as a special one when the correlation length is zero. The correlation model we consider is a continuous neural field (Amari, 1977; Giese, 1999) in which neurons are pair-wise correlated with a strength distribution given by a gaussian function of their difference in preferred stimuli. Depending on the width of gaussian function, the correlation

form varies to include noncorrelation, local correlation, short-range correlation, wide-range correlation, and uniform correlation. Compared with other prototype models in the literature (Abbott & Dayan, 1999; Johnson, 1980; Snippe & Koenderink, 1992; Yoon & Sompolinsky, 1999), this one shows the advantage of simplifying the calculation and provides a clear picture of the results. As we show, due to the property of the gaussian function and the continuous extension of the model, many calculations can be done more easily in the Fourier domain. It is also possible to apply the proposed method of neural fields to a more general case of firing-rate-dependent variances, including the multiplicative and Poisson-type noises, as treated in Wilke and Eurich (in press).

Based on the proposed encoding model, we first calculate the Fisher information. The inverse of Fisher information, the Cramér-Rao bound, defines the optimal accuracy for an unbiased estimator to achieve. When no correlation exists, that is, the correlation width is zero, the Fisher information increases in proportion to the number of neurons or the neural density in the field. As the width of correlation increases, the Fisher information decreases rapidly. Abbott and Dayan (1999) and Yoon and Sompolinsky (1999) found that if the neural correlation covers a nonlocal range of population, the Fisher information does not increase but saturates even when the number of neurons increases. The same behavior is observed in a simpler way in this article. We also show that as the range of correlation becomes wider, that is, when the width of correlation becomes larger than $\sqrt{2}$ times the effective length of the tuning function, the Fisher information increases again in proportion to the number of neurons. This is an interesting new finding. Such a phenomenon is observed more generally in the case of multiplicative noise (Wilke & Eurich, in press), which we also confirmed by our method, although it is not described in this article.

We then study three decoding methods and compare their performances. All of them are formulated in the maximum likelihood inference (MLI) type (including the conventional center of mass method), whereas they differ in the knowledge on the true encoding scheme. It turns out that a decoding method that keeps the knowledge of the tuning function but neglects the neural correlation is a good compromise between computational complexity and decoding accuracy, supporting the findings in (Wu, Nakahara, et al., 2000; Wu et al., 2001).

In previous work, we have pointed out that the MLI type of decoding method may not be asymptotically efficient for some strong correlation structures (Wu, Nakahara, et al., 2000; Wu et al., 2001). In this article, we prove that this is indeed the case when the correlation covers a nonlocal range of population (excepting the uniform correlation and when the noise is extremely small). Here, the estimated or decoded position of stimulus is no longer subject to the gaussian distribution, but is subject to the Cauchy-type distribution in which the mean and variance diverge. In other words, the standard paradigm of statistical estimation does not hold. This is also a

new finding in population coding, and we discuss the consequence of this property.

We also investigate the network implementation of the three decoding methods, following the idea of Pouget et al. (Deneve, Latham, & Pouget, 1999; Pouget & Zhang, 1997; Pouget et al., 1998). A recurrent neural field is constructed in such a way that its steady state has a shape similar to the tuning function and is noise free. The peak position of the steady state gives the estimator of the stimulus (Amari, 1977; Giese, 1999). When there is no external input to the network, the system is neutrally stable on a line attractor. A small input will cause the state to drift to the position corresponding to the estimator in the three methods. Throughout the article, the effect of correlation on decoding accuracies is discussed as a by-product of calculations.

In this study, we consider only that the stimulus is a single variable. Population coding also works in more complex cases, as studied by Eurich, Wilke, and Schwegler (2000), Treue, Hol, and Rauber (2000), Zemel, Dayan, and Pouget (1998), Zhang and Sejnowski (1999), and Zohary (1992). We need to study a high-dimensional neural field. These issues are not in the scope of the work presented here.

The article is organized as follows. In section 2, we introduce a discrete encoding model, which is extended to be the continuous version in section 3. In section 4, the Fisher information for the encoding model is calculated. In section 5, we compare three decoding methods. Their network implementations are discussed in section 6. Conclusions and a discussion are given in section 7.

2 Encoding Model

We begin with a discrete encoding model. Consider an ensemble of N neurons coding a variable x, which represents the position of the stimulus. Let us denote by c_i the preferred stimulus position of the ith neuron, and let r_i denote the response of the neuron, so that $\mathbf{r} = \{r_i\}$, for $i = 1, \ldots, N$, denote the population activity.

The neural responses are correlated, and the ith neuron's activity is given by

$$r_i = f_i(x) + \sigma\epsilon_i, \qquad i = 1, \ldots, N, \tag{2.1}$$

where $f_i(x)$ is the tuning function of the ith neuron representing the mean value of the response when stimulus x is applied, and $\sigma\epsilon_i$ is noise.

In this study, we consider only the gaussian tuning function, that is,

$$f_i(x) = \frac{1}{\sqrt{2\pi}a} e^{-(c_i - x)^2/2a^2}, \tag{2.2}$$

where the parameter a is the tuning width.

The parameter σ represents the noise intensity, and ϵ_i is a gaussian random variable with mean 0 and variance 1. We decompose ϵ_i as

$$\epsilon_i = \epsilon_i' + \epsilon_i'', \tag{2.3}$$

where $\{\epsilon_i'\}$ are independent to all the others with zero mean and variance $1 - \beta$, while ϵ_i'' and ϵ_j'' are correlated. We assume the gaussian correlation,

$$\langle \epsilon_i'' \cdot \epsilon_j'' \rangle = \beta e^{-(c_i - c_j)^2/2b^2}, \tag{2.4}$$

where $<\quad>$ denotes expectation over many trials. Then, the noise satisfies

$$\langle \epsilon_i \rangle = 0, \tag{2.5}$$

$$\langle \epsilon_i \epsilon_j \rangle = A_{ij}, \tag{2.6}$$

where the covariance matrix is given by

$$A_{ij} = (1 - \beta)\delta_{ij} + \beta e^{-(c_i - c_j)^2/2b^2}. \tag{2.7}$$

The parameter satisfies $0 \leq \beta \leq 1$, and the width b is called the effective correlation length. The model captures the fact that the correlation strength between neurons decreases with the dissimilarity in their preferred stimuli $|c_i - c_j|$.

The encoding process of population coding is fully specified by the conditional probability density of $\boldsymbol{r}$ when stimulus x is given as

$$Q(\mathbf{r}|x) = \frac{1}{\sqrt{(2\pi\sigma^2)^N \det(\mathbf{A})}} \times \exp\left[-\frac{1}{2\sigma^2}\sum_{ij} A_{ij}^{-1}(r_i - f_i(x))(r_j - f_j(x))\right]. \tag{2.8}$$

3 From Discrete to Continuous

Because we are interested only in the case when the number N of neurons is large (more accurately, the neuron density is large), it is useful to extend the discrete model to the continuous case. Mathematically, there are a lot of benefits from coping with a continuous neural field model (Amari, 1977; Giese, 1999). For example, as shown later, many operations can be done much more easily in the Fourier domain due to the continuous extension.

Let us consider a one-dimensional neural field in which neurons are located with uniform density ρ. The activity of neuron at position c is denoted by $r(c)$. The neural response function $r(c)$ is given by

$$r(c) = f(c - x) + \sigma\epsilon(c), \tag{3.1}$$

when stimulus x is applied, where quantities $r(c)$, $f(c - x)$, and $\epsilon(c)$ are the counterparts of r_i, f_i, and ϵ_i in the discrete version. The tuning function $f(c - x)$ has the same form as that of $f_i(x)$ except that c_i is replaced by c.

The noise term $\epsilon(c)$ satisfies

$$\langle\epsilon(c)\rangle = 0, \tag{3.2}$$

$$\langle\epsilon(c)\epsilon(c')\rangle = h(c, c')/\rho^2, \tag{3.3}$$

where $h(c, c')$ is the covariance function divided by the neuron density ρ^2.

We assume that the covariance function has the same form as the discrete case,

$$h(c, c') = D_1(1 - \beta)\delta(c - c') + D_2\beta e^{-(c-c')^2/2b^2}, \tag{3.4}$$

where $\delta(c - c')$ is the delta function. In order to determine the coefficients D_1 and D_2, we use the correspondence principle: The covariance matrix A_{ij} and the correlation function $h(c, c')$ correspond to each other to give the quadratic form for an arbitrary vector $\mathbf{k} = (k_i)$ and its continuous version $k(c)$,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(c)h(c, c')k(c')\,dc\,dc' = \sum_{ij} k_i A_{ij} k_j, \tag{3.5}$$

Substituting equations 2.7 and 3.4 into 3.5, we get (see appendix A)

$$h(c, c') = \rho(1 - \beta)\delta(c - c') + \rho^2\beta e^{-(c-c')^2/2b^2}. \tag{3.6}$$

Therefore, the continuous form of the encoding process is

$$Q(\mathbf{r}|x) = \frac{1}{Z}\exp\left\{-\frac{\rho^2}{2\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[r(c) - f(c - x)]h^{-1}(c, c') \times[r(c') - f(c' - x)]\,dc\,dc'\right\}, \tag{3.7}$$

where $\mathbf{r} = \{r(c)\}$, and the parameter Z is the normalization factor. The function $h^{-1}(c, c')$ is the inverse kernel of $h(c, c')$, satisfying

$$\int_{-\infty}^{\infty} h^{-1}(c, c')h(c', c'')\,dc' = \delta(c - c''). \tag{3.8}$$

The above correlation model contains a number of important forms depending on the correlation length:

- *No correlation*: When the correlation length $b = 0$, neurons are uncorrelated, with the covariance function $h(c, c') = \rho(1 - \beta)\delta(c - c')$ or the correlation matrix $A_{ij} = (1 - \beta)\delta_{ij}$.
- *Local correlation*: When the correlation length is of order $1/\rho$, that is, $b = m/\rho$ for a fixed m, neurons are correlated only within m neighboring neurons. When the density ρ is large, neurons are correlated only extremely locally.
- *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, that is, $1/\rho \ll b < \sqrt{2}a$, neurons are correlated over a short range.[1]
- *Wide-range correlation*: When $b \geq \sqrt{2}a$, neurons are correlated over a wide range.
- *Uniform correlation*: When the correlation length $b \to \infty$, neurons are uniformly correlated with strength β, that is, $h(c, c') = \rho(1 - \beta)\delta(c - c') + \rho^2\beta$ or $A_{ij} = (1 - \beta)\delta_{ij} + \beta$.

4 The Fisher Information

The Fisher information is a useful measure in the study of population coding. By knowing the Fisher information, we have an idea of the minimum error one may achieve.

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

$$I_F(x) = -\int Q(\mathbf{r}|x)\frac{d^2 \ln Q(\mathbf{r}|x)}{dx^2}\, d\mathbf{r}. \tag{4.1}$$

From equation 3.7, we get

$$I_F(x) = \frac{\rho^2}{\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f'(c - x)h^{-1}(c, c')f'(c' - x)\, dc\, dc', \tag{4.2}$$

where $f'(c-x) = df(c-x)/dx$. Note that $I_F(x)$ does not depend on x because of the homogeneity of the field.

The Fourier transform of a function $g(t)$ is defined as

$$\mathcal{F}[g(t)] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega t}g(t)\, dt. \tag{4.3}$$

[1] We consider only the case that $a \gg 1/\rho$, as suggested by experiments, that is, neurons are widely tuned by stimulus.

By using the Fourier transformation, equation 4.2 becomes

$$I_F(x) = \frac{\rho^2}{2\pi\sigma^2}\int_{-\infty}^{\infty}\frac{\omega^2 F(\omega)^2}{H(\omega)}\,dw, \tag{4.4}$$

where $F(\omega) = \mathcal{F}[f(c-x)]$ and $H(\omega) = \mathcal{F}[h(c-c')]$. Here, we use the relations $\mathcal{F}[f'(c-x)] = i\omega F(\omega)$ and $\mathcal{F}[h^{-1}(c,c')] = 1/H(\omega)$.

From equations 2.2 and 3.6,

$$F(\omega) = e^{-a^2\omega^2/2}, \tag{4.5}$$

$$H(\omega) = \rho(1-\beta) + \rho^2\sqrt{2\pi}\beta b e^{-b^2\omega^2/2}, \tag{4.6}$$

where we put $x = 0$ without loss of generality. Therefore,

$$I_F(x) = \frac{\rho^2}{2\pi\sigma^2}\int_{-\infty}^{\infty}\frac{\omega^2 e^{-a^2\omega^2}}{\rho(1-\beta) + \rho^2\sqrt{2\pi}\beta b e^{-b^2\omega^2/2}}\,d\omega. \tag{4.7}$$

Let us analyze the property of Fisher information in various forms:

- *No correlation*: The Fourier transform of the covariance function is $H(\omega) = \rho(1-\beta)$, and the Fisher information $I_F(x) = \rho/[4\sqrt{\pi}a^3\sigma^2(1-\beta)]$ increases in proportion to the neuron density ρ. Since each neuron carries independent information for x and the total Fisher information is their sum, I_F increases in proportion to ρ.
- *Local correlation*: $H(\omega) = \rho(1-\beta+\sqrt{2\pi}m\beta e^{-m^2\omega^2/2\rho^2})$ is of order ρ. In this case, the total Fisher information I_F is not a simple sum of component neurons, but correlations disappear so quickly that the total information is still in proportion to ρ.
- *Short-range correlation*: For large ρ, $H(\omega)$ is the order of ρ^2, so that we have

  $$H(\omega) \approx \rho^2\sqrt{2\pi}\beta b e^{-b^2\omega^2/2}. \tag{4.8}$$

  Hence, I_F is written as

  $$I_F \approx \frac{1}{(2\pi)^{3/2}\beta b\sigma^2}\int \omega^2 e^{(-a^2+b^2/2)\omega^2}\,d\omega. \tag{4.9}$$

  Since $b < \sqrt{2}a$, the integral converges to a constant. Hence, I_F is of finite value even when ρ goes to infinity. This result agrees with Abbott and Dayan (1999) and Yoon and Sompolinsky (1999) under the additive noise distribution.

- *Wide-range correlation*: When $b \geq \sqrt{2}a$, the integral, equation 4.9, diverges, and we cannot use this approximation. If we evaluate the integral more accurately by taking the term of $1/\rho$ into account, we have

$$I_F = \frac{1}{2\pi\sigma^2} \int \frac{\omega^2 e^{-a^2\omega^2}}{(1/\rho)(1-\beta) + \sqrt{2\pi}\beta b e^{-b^2\omega^2/2}} \, d\omega$$
$$= \frac{\rho}{2\pi\sigma^2} \int \frac{\omega^2}{(1-\beta)e^{a^2\omega^2} + \sqrt{2\pi}\rho\beta b e^{(a^2 - b^2/2)\omega^2}} \, d\omega. \tag{4.10}$$

  Hence, I_F increases with ρ for large ρ (see Figure 2d).

- *Uniform correlation*: $H(\omega) = \rho(1-\beta)$, and $I_F(x) = \rho/[4\sqrt{\pi}a^3\sigma^2(1-\beta)]$ is proportional to ρ, which has the same value as in the uncorrelated case.[2] In this case, the noise is decomposed as (Wu, Nakahara, et al., 2000; Wu et al., 2001)

$$\epsilon(c) = \epsilon'(c) + \epsilon'', \tag{4.11}$$

  where $\epsilon'(c)$ is independent noise, $\langle \epsilon'(c)\epsilon'(c') \rangle = (1-\beta)\delta(c-c')/\rho^2$, and ϵ'' is common to all neurons, $\langle (\epsilon'')^2 \rangle = \beta/\rho^2$. The term ϵ'' increases all $r(c)$ by a common factor $\sigma\epsilon''$, which does not have any effect for decoding x.

- *Special case with $\beta = 1$*: In order to make the situation clear, we consider this special case. In this case, when $b < \sqrt{2}a$, the Fisher information is constant not depending on ρ, as is similar to the case of short-range correlation. However, when $b \geq \sqrt{2}a$, the Fisher information diverges to ∞ as is seen from equation 4.7. This implies that x can be decoded accurately. In this case, the noise $\epsilon(c)$ is so strongly correlated among neurons in an interval of b that $\epsilon(c)$ is regarded as constant in this range. Since the range of width b covers the effective range a of the tuning function, the noise shifts the tuning function to $r(c) = f(x-c) + \epsilon$, so that we can decode x from $r(c)$ as if there were no noise.

Figure 1 shows how the Fisher information behaves as the correlation width changes (with fixed neural density). It first decreases drastically when the correlation width is small and then increases again when the width is large. Figure 2 shows the asymptotic behavior of the Fisher information in different correlation cases, as the density ρ increases.

When a multiplicative noise model is used, the Fisher information always increases with the neural density ρ (Wilke & Eurich, in press). This is im-

[2] This conclusion is different from that in Abbott and Dayan (1999), where the uncorrelated case is defined as $A_{ij} = \delta_{ij}$ (by setting $\beta = 0$) instead of $A_{ij} = (1-\beta)\delta_{ij}$ used here (by setting $b = 0$).

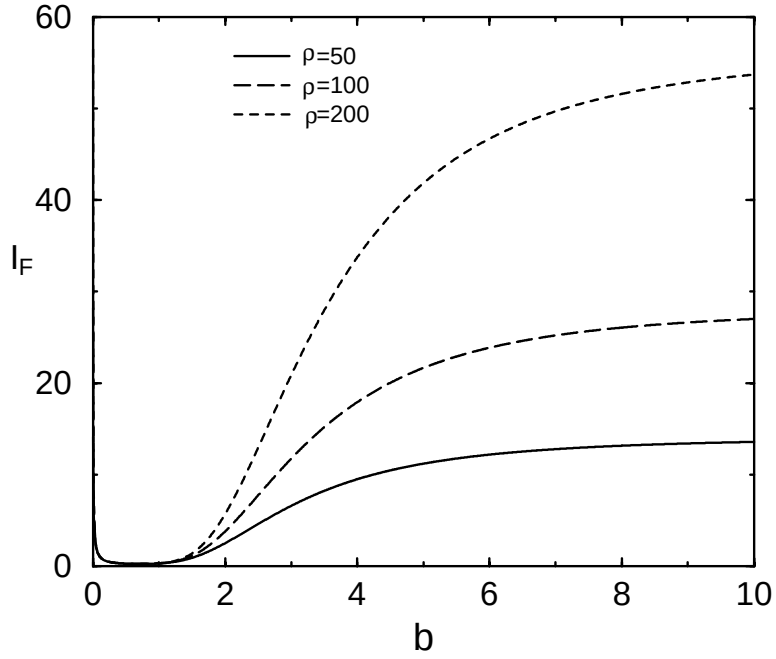


Figure 1: The Fisher information changes with the correlation width. The parameters are $a = 1$, $\sigma = 1$, and $\beta = 0.5$.

portant because the multiplicative noise is believed to be more biologically plausible. Our method can be extended in such a case.

5 Population Decoding

The Fisher information tells us only the optimal decoding accuracy for an unbiased estimator to achieve. When a practical decoding method is concerned, its performance needs to be evaluated individually depending on the decoding model. We compare three decoding methods in this study. All are formulated as the MLI type, that is, the maximizer of a likelihood function, whereas they differ in the probability models used for decoding.

5.1 Three Decoding Methods. An MLI type estimator $\hat{x}$ is obtained through maximization of the presumed log likelihood $\ln P(\mathbf{r}|x)$, that is, by solving

$$\nabla \ln P(\mathbf{r}|\hat{x}) = 0, \tag{5.1}$$

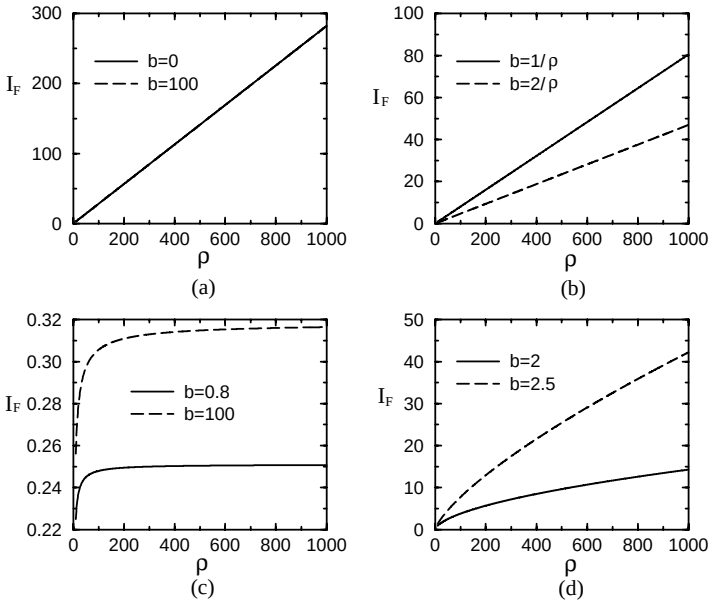


Figure 2: The asymptotic behavior of Fisher information in different correlation cases. The parameters are $a = 1$, $\sigma = 1$, and $\beta = 0.5$. (a) No correlation ($b = 0$) and uniform correlation ($b = 100$, approximated as infinity). The two curves coincide. (b) Limited-range correlations. Parameters are $b = 1/\rho$ and $b = 2/\rho$, respectively. (c) Short-range correlations. $b = 0.8$ and 1. (d) Wide-range correlations. $b = 2$ and 2.5.

where $\nabla k(x)$ denotes $dk(x)/dx$. $P(\mathbf{r}|x)$ is called the decoding model, which can be different from the real encoding model $Q(\mathbf{r}|x)$. This is because the decoding system usually does not know the exact information of the encoding system. Moreover, a simple and robust decoding model is computationally desirable. We consider three decoding models, defined as follows:

1. The conventional MLI, referred to as FMLI (MLI based on the faithful model), uses all of the encoding information, that is, the decoding model is the true encoding model,

$$P_F(\mathbf{r}|x) = Q(\mathbf{r}|x). \tag{5.2}$$

2. The UMLI method (MLI based on an unfaithful model) (Wu, Nakahara et al., 2000) uses the information on the shape of the tuning function

but neglects the neural correlation, so that the probability density

$$P_U(\mathbf{r}|x) = \frac{1}{Z_U} \exp\left\{-\frac{\rho}{2\sigma^2}\int_{-\infty}^{\infty}[r(c) - f(c-x)]^2\,dc\right\}, \tag{5.3}$$

is used for decoding.

3. The center of mass (COM) method does not use any information of the encoding process; instead, it assumes an incorrect but simple tuning function. It also disregards correlations by using

$$P_C(\mathbf{r}|x) = \frac{1}{Z_C} \exp\left\{-\frac{\rho}{2\sigma^2}\int_{-\infty}^{\infty}[r(c) - \tilde{f}(c-x)]^2\,dc\right\}, \tag{5.4}$$

where $\tilde{f}(c-x) = -(x-c)^2 + \text{const}$ is used as a presumed tuning function.[3]

It is easy to check that the third method is equivalent to the conventional COM decoding strategy (Baldi & Heiligenberg, 1988; Georgopoulos et al., 1982; Wu, Nakahara et al., 2000),[4] with the solution given by

$$\hat{x} = \frac{\int_{-\infty}^{\infty} cr(c)\,dc}{\int_{-\infty}^{\infty} r(c)\,dc}. \tag{5.5}$$

We should point out that the use of an unfaithful model (e.g., UMLI or COM) has important meaning. When experimental scientists reconstruct the stimulus from the recorded data, they in fact use an unfaithful model, since the real encoding process is never known. Furthermore, the real neural correlation is often complex and may constantly change over time, so that it is hard, even if possible, for the brain to store and use all this information. MLI based on an unfaithful model (neglecting some part of information) is a key to solving this information curse.

5.2 Performance of Decoding and Asymptotic Efficiency. Since the three methods are of the same type, their decoding errors can be calculated in similar ways. We show only the derivation of the decoding error for UMLI. For FMLI, see appendix B.

[3] In the strict sense, $\tilde{f}(c-x)$ cannot be a tuning function, since its value diverges when $(x-c)$ goes to infinity. However, this does not matter in practice, since $(x-c)$ is always restricted to a finite region when COM is practically used. See Snippe (1996), Wu, Nakahara, et al. (2000), and Wu et al. (2001).

[4] This explains why COM performs comparably well to MLI when neurons are uncorrelated and the cosine tuning function is used (Salinas & Abbott, 1994), since the cosine function, for example, $\cos[(x-c)/T]$, has an approximated quadratic form when $(x-c)$ is restricted in a small region.

For convenience, two notations are introduced: $E_Q[k(\mathbf{r}, x)]$ and $V_Q[k(\mathbf{r}, x)]$ denote, respectively, the mean and variance of $k(\mathbf{r}, x)$ with respect to the distribution $Q(\mathbf{r}|x)$.

Suppose $\hat{x}$ is close enough to x. We expand $\nabla \ln P_U(\mathbf{r}|\hat{x})$ at x,

$$\nabla \ln P_U(\mathbf{r}|\hat{x}) \simeq \nabla \ln P_U(\mathbf{r}|x) + \nabla\nabla \ln P_U(\mathbf{r}|x)(\hat{x} - x). \tag{5.6}$$

Since the estimator $\hat{x}$ satisfies $\nabla \ln P_U(\mathbf{r}|\hat{x}) = 0$,

$$\frac{1}{\rho}\nabla\nabla \ln P_U(\mathbf{r}|x)(\hat{x} - x) \simeq -\frac{1}{\rho}\nabla \ln P_U(\mathbf{r}|x). \tag{5.7}$$

We put

$$\begin{aligned} R &= \frac{1}{\rho}\nabla \ln P_U(\mathbf{r}|x) = \frac{1}{\sigma^2}\int_{-\infty}^{\infty}[r(c) - f(c - x)]f'(c - x)\,dc \\ &= \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\epsilon(c)f'(c - x)\,dc, \end{aligned} \tag{5.8}$$

$$\begin{aligned} S &= \frac{1}{\rho}\nabla\nabla \ln P_U(\mathbf{r}|x) = \frac{1}{\sigma^2}\int_{-\infty}^{\infty}[r(c) - f(c - x)]f''(c - x)\,dc \\ &\quad + \frac{1}{\sigma^2}\int_{-\infty}^{\infty}[f'(c - x)]^2\,dc \\ &= \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\epsilon(c)f''(c - x)\,dc + D, \end{aligned} \tag{5.9}$$

where

$$D = \frac{1}{4\sqrt{\pi}a^3\sigma^2}. \tag{5.10}$$

Then the estimating equation is

$$S(\hat{x} - x) = -R \tag{5.11}$$

or

$$\hat{x} - x = -\frac{R}{S}. \tag{5.12}$$

Here, both R and S are random variables depending on $\epsilon(c)$. It is easy to show that

$$E_Q[R] = 0, \tag{5.13}$$

$$E_Q[S] = \frac{1}{4\sqrt{\pi}a^3\sigma^2}. \tag{5.14}$$

Their variances are given, by using the Fourier transforms, as

$$V_Q[R] = \frac{1}{2\pi\rho^2\sigma^2}\int_{-\infty}^{\infty}\omega^2 F(\omega)^2 H(\omega)\,d\omega, \tag{5.15}$$

$$V_Q[S] = \frac{1}{2\pi\rho^2\sigma^2}\int_{-\infty}^{\infty}\omega^4 F(\omega)^2 H(\omega)\,d\omega. \tag{5.16}$$

These show that R is a zero-mean gaussian random variable. The random variable S is composed of two terms, the first one being random and the second term D being a fixed constant (see the right-hand side of equation 5.9).

Remark. The above procedure is the standard way to analyze the asymptotic error of estimation. In the standard statistical model with repeated independent observations that is, the independent and identically distributed (i.i.d.) case, R is gaussian, while the constant term D dominates over the random one because of the law of large numbers. In particular, when the faithful model is used (FMLI), D is the Fisher information, and so is $V_Q[R]$. Therefore, the asymptotic error is given by $1/N$ times of the inverse of Fisher information, where N is the number of observations. This shows that the Cramér-Rao bound is asymptotically attained. Such an estimator is said to be asymptotically efficient or Fisher efficient. When an unfaithful model is applied (still to the i.i.d. case), $V_Q[R]$ and D are different. Hence, the Fisher efficiency is not attained in general (Akahira & Takeuchi, 1981; Murata, Yoshizawa, & Amari, 1994). However, the asymptotic gaussianity of the estimator and the $1/N$ convergence property of the error are guaranteed. In this article, we say that such an estimator is quasi-asymptotic efficient or quasi-Fisher efficient.

Apart from the above two standard cases, population coding includes the third one (because of correlation), where the estimator is not subject to the gaussian distribution, although the asymptotic error may be small (Wu et al., 2001). Such an unusual case occurs when the random term S is not negligible, as compared to the constant D. In such a situation, the estimator is represented by a ratio of two gaussian random variables, so it is subject to the Cauchy-type distribution. The $1/N$ convergence of error does not hold either. Such an estimator is said to be non-Fisherian, since the Cramér-Rao paradigm using the Fisher information does not hold. This is a new fact in the population coding literature.

Let us return to calculating the UMLI decoding error. From equations 5.10 and 5.16, we see that the constant term in S dominates over the random one

in the two cases:

- $H(\omega)$ is of order ρ. In this case, the random term is of $O(\frac{1}{\rho})$, and D is of $O(1)$.
- $H(\omega)$ is of order ρ^2, but the noise variance σ^2 is sufficiently small. In this case, the random one is $O(\frac{1}{\sigma})$, and the constant term D is $O(\frac{1}{\sigma^2})$.

The first case corresponds to uncorrelated, local range, and uniformly correlated cases. The second case holds for the short- and wide-range correlations with small noise.[5] In these cases, we may neglect the random term in R, so that we have asymptotically

$$S \approx D, \tag{5.17}$$

$$\hat{x} - x \approx \frac{R}{D}, \tag{5.18}$$

which is normally distributed with zero mean and variance,

$$E_Q[(\hat{x} - x)^2] = \frac{8a^6\sigma^2}{\rho^2} \int_{-\infty}^{\infty} \omega^2 F(\omega)^2 H(\omega)\, d\omega. \tag{5.19}$$

The above equation holds asymptotically when $\rho \to \infty$ and $H(\omega)$ is of order ρ, or $\sigma \to 0$.

The cases of short- and wide-range correlations and the noise are strong. In these cases, since $H(\omega)$ is of order ρ^2, the random and constant terms in the variable S are of the same order. It is rather difficult to analyze the behavior of $\hat{x}$. Equation 5.12 shows that $\hat{x}-x$ is a ratio of two gaussian random variables, so that its distribution is of the Cauchy type, which means that UMLI is not Fisherian. We may note that the neural field is finite practically; the variance of decoding error does not diverge even in the Cauchy case. But the variance is no longer adequate to measure the decoding accuracy.

A similar analysis is applicable to FMLI (see appendix B) and COM. It turns out that they have the same asymptotic behaviors. Table 1 summarizes the asymptotic behaviors of the three decoding methods (the special case of weak noise, in which the MLI type of methods is always approximately asymptotically or quasi-asymptotically efficient, is not included) and the Fisher information in different correlation cases.

[5] Note that the weak noise assumption is used in Deneve et al. (1999) and Pouget et al. (1998) to get their results.

Table 1: The Asymptotic Behaviors of Fisher Information and Three MLI Type of Decoding Methods.

	I_F	FMLI	UMLI	COM
No correlation	$\propto \rho$	FE	QFE	QFE
Local correlation	$\propto \rho$	FE	QFE	QFE
Short-range correlation	Saturating	Non-F	Non-F	Non-F
Wide-range correlation	$\propto \rho$	Non-F	Non-F	Non-F
Uniform correlation	$\propto \rho$	FE	QFE	QFE

Notes: The special case of weak noise is exluded. FE = Fisher efficiency, QFE = quasi-Fisher efficiency; Non-F = non-Fisherian.

Table 2: Comparing the Decoding Errors of FMLI, UMLI, and COM When $H(\omega)$ is Order of ρ.

	$b = 0$	$b \to \infty$	$b = m/\rho$
FMLI	$4\sqrt{\pi}a^3\sigma^2(1-\beta)/\rho$	$4\sqrt{\pi}a^3\sigma^2(1-\beta)/\rho$	$4\sqrt{\pi}a^3\sigma^2[1+(\sqrt{2\pi}m-1)\beta]/\rho$
UMLI	$4\sqrt{\pi}a^3\sigma^2(1-\beta)/\rho$	$4\sqrt{\pi}a^3\sigma^2(1-\beta)/\rho$	$4\sqrt{\pi}a^3\sigma^2[1+(\sqrt{2\pi}m-1)\beta]/\rho$
COM	$18a^3\sigma^2(1-\beta)/\rho$	$18a^3\sigma^2(1-\beta)/\rho$	$18a^3\sigma^2[1+(\sqrt{2\pi}m-1)\beta]/\rho$

When FMLI and COM are asymptotically efficient, their decoding errors are calculated to be

$$E_Q[(\hat{x}-x)^2_{\text{FMLI}}] \sim \frac{2\pi\sigma^2}{\rho^2\int_{-\infty}^{\infty}\omega^2 F(\omega)^2/H(\omega)\,d\omega} = \frac{1}{I_F}, \tag{5.20}$$

$$E_Q[(\hat{x}-x)^2_{\text{COM}}] \sim \frac{\sigma^2}{\rho^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ch(c,c')c'\,dc\,dc', \tag{5.21}$$

respectively.

5.2.1 Performance Comparison. We compare the performances of three decoding methods. Table 2 lists the decoding errors of three decoding methods when $H(\omega)$ is the order of ρ.[6] We see that UMLI and FMLI have comparable performances. Both are much better than COM.

Figure 3 compares the decoding errors of FMLI, UMLI and COM in the case of nonlocal range correlation and weak noise (note that FMLI, UMLI, and COM are asymptotically or quasi-asymptotically efficient in this case

[6] Two conditions are used to obtain the results in Table 2. (1) To calculate the decoding error of COM, we need to restrict the preferred stimulus c within a finite range $[-L, L]$; otherwise, the error diverges. This restriction means sampling only those neurons that are active enough. This approximation benefits only COM and does not affect the results of FMLI and UMLI much (see Wu, Nakahara et al., 2000; Wu et al., 2001). In this article, we choose $L = 3a$. (2) To get the results for the case of $b = m/\rho$, we use the condition $a \gg 1/\rho$.

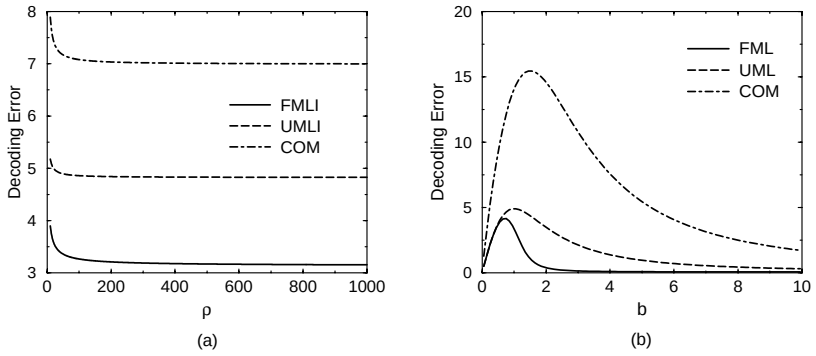


Figure 3: Comparing the decoding errors of FMLI, UMLI, and COM in the case of strong correlation and weak noise. The unit of decoding errors in the figure is σ^2. Parameters $a = 1$ and $\beta = 0.5$. (a) Decoding errors change with the neuron density ρ. Parameter $b = 1$. (b) Decoding errors change with the correlation length b. Neural density $\rho = 50$.

because of weak noise.). It shows that UMLI has a larger error than FMLI and a lower error than COM. When the correlation covers the short range of the population, the decoding errors of three methods saturate as the neural density increases (see Figure 3a). This is similar to the behavior of Fisher information. For a fixed neural density ρ, the decoding errors of the three methods first increase with the correlation length and then decrease subsequently (see Figure 3b). This is understandable, since the extremes on both ends correspond to the cases when neurons are either uncorrelated or uniformly correlated.

The computational complexity of the three methods can be roughly compared as follows. Consider maximization of the log likelihood of UMLI and FMLI by using the standard gradient-descent method. The amounts of computation for obtaining the derivative of the log likelihood of UMLI and FMLI are proportional to N and N^2, respectively. UMLI is significantly simpler than FMLI when N is large. For COM, due to the quadratic form of the tuning function, the estimation can be done in one shot by equation 5.5. Therefore, UMLI is a good compromise between decoding accuracy and computational complexity.

6 Network Implementation

So far, we have been concerned only with the accuracy and simplicity of decoding methods. To be realistic, it is essential for the strategy to be bi-

ologically achievable. We investigate implementing the three methods by using a recurrent network, following the ideas in Deneve et al. (1999), Pouget and Zhang (1997), and Pouget et al. (1998). UMLI is studied first.

Let us consider a fully connected one-dimensional homogeneous neural field, in which c denotes the position coordinate. Let U_c denote the (average) internal state of neurons at c and $W_{c,c'}$ the recurrent connection weights from neurons at c to those at c'. We propose that the dynamics of neural excitation is governed by

$$\frac{dU_c}{dt} = -U_c + \int W_{c,c'} O_{c'}\, dc' + I_c, \tag{6.1}$$

where

$$O_c = \frac{U_c^2}{1 + \mu \int U_c^2\, dc}. \tag{6.2}$$

This O_c is the activity of the neurons at c, and I_c is the external input arriving at c. The recurrent interaction is assumed to be

$$W_{c,c'} = e^{-(c-c')^2/2a^2}. \tag{6.3}$$

We first look at the network dynamics, when there is no external input—$I_c = 0$. It is not difficult to check that the network has a one-parameter family of nontrivial steady states,

$$\tilde{O}_c = Ae^{-(c-z)^2/2a^2}, \tag{6.4}$$

with the corresponding internal state,

$$\tilde{U}_c = Be^{-(c-z)^2/4a^2}, \tag{6.5}$$

including z as a free parameter, where the coefficients A and B are easily determined. The parameter z denotes the peak position of the population activity comprised by $\{\tilde{O}_c\}$. Note that the stable state has a similar shape to the tuning function (see equation 2.2) in this case. Equation 6.4 or 6.5, parameterized by z, defines a line attractor for the network, on which the system is neutrally stable. This is due to the translation invariance of the network interactions (Amari, 1977; Deneve et al., 1999; Giese, 1999; Seung, 1996; Zhang, 1996).

The decoded estimator $\hat{x}$ is given by the peak position of the final population activity O_c, that is, the value of z. In order to achieve this goal, one considers that the external input is transient, that is, $I_c \sim r_c^{1/2}\delta(t)$. This stimulates the network whose initial state is set equal to the original noisy version, $O_c(0) = r_c$ (Pouget et al., 1998; Deneve et al., 1999). After relaxation, the system reaches the desired state, in which the noises are smoothed out and the peak position gives $\hat{x}$. However, if I_c disappears, the attractor of the field is only neutrally stable, so that the peak position may fluctuates. Hence, we consider that a small input $I_c = \epsilon r_c$ persists after the initial state is set.[7] We further assume that the input is sufficiently small (ϵ is small enough), such that the change it brings to the form of the stable state (see equation 6.4) is negligible.

Instead of being trapped into complicated mathematical calculation, we adopt an approximate but simple way to understand the above network dynamics, being backed up by simulation results. Since the steady state of the network is assumed always to be on the line attractor and its position is determined by the input, independent of the initial value, we can, without risk of losing any information, see that the network is, from the beginning on the line attractor, at a random position, evolving along it, until it reaches a stable position. This modified dynamic picture simplifies the relationship between O_c and U_c as

$$O_c = DU_c^2, \tag{6.6}$$

where equations 6.4 and 6.5 are used, and $D = B^2/A$.

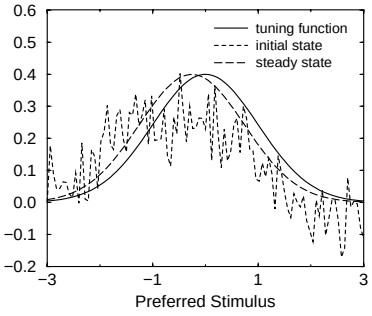
For the dynamics specified by equations 6.1 and 6.6, there exists a Lyapunov function (Cohen & Grossberg, 1983),

$$\begin{aligned} L = &-\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{c,c'}O_cO_{c'}\,dc\,dc' \\ &+ \int_{-\infty}^{\infty}\int_{0}^{U_c} zg'(z)\,dz\,dc - \epsilon\int_{-\infty}^{\infty} r_cO_c\,dc, \end{aligned} \tag{6.7}$$

with the function $g(z) = Dz^2$.

We consider a two-step perturbation procedure to minimize the Lyapunov function. Minimizing the first two terms of equation 6.7, which are

[7] This mechanism was proposed in Amari (1977) and Pouget and Zhang (1997).



(a)

	$V_Q[\hat{z}]$	$V_Q[\hat{x}]$	t
$b = 0.01$	0.00031	0.00030	0.013
$b = 0.1$	0.00112	0.00109	0.019
$b = 1$	0.00503	0.00507	0.032
$b = 2$	0.00371	0.00362	0.016

(b)

Figure 4: Performance of the recurrent network. The parameters are $a = 1$, $\mu = 0.5$, $\sigma^2 = 0.01$, and $\beta = 0.5$. (a) The typical states of network before and after relaxation. The steady state is scaled to match the tuning function. (b) Comparing the decoding errors of the network estimation and UMLI. The results are obtained after averaging over 100 trials.

of order 1, determines that the stable state is on the line attractor. Minimizing the third term of order of ϵ determines the peak positions of the stable state. A justification of this two-step perturbation procedure is given in appendix C. This two-step optimization is equivalent to

$$\begin{aligned} &\text{Min}_{\{z\}} \quad -\int_{-\infty}^{\infty} r_c O_c \, dc, \\ &\text{subject to} \quad O_c = Ae^{-(c-z)^2/2a^2}. \end{aligned} \tag{6.8}$$

Recall that the solution of UMLI is given by

$$\begin{aligned} &\text{Max}_{\{x\}} \quad \ln P_U(\mathbf{r}|x), \\ &\qquad = \frac{\rho}{\sigma^2}\int_{-\infty}^{\infty} r_c f(c-x)\, dc + \text{terms not on } x. \end{aligned} \tag{6.9}$$

Comparing equations 6.8 with 6.9 and 2.2, we see that the final state of the network gives the same estimator as that of UMLI.

The above result is confirmed by the simulation experiment (see Figure 4), which was done with 101 neurons uniformly distributed in the region $[-3, 3]$, with the true stimulus having the value of zero. Figure 4a shows the

typical behaviors of the recurrent network before and after relaxation. The steady state of the network becomes smooth after the noise in the initial state is cleaned out. To see the coincidence between the two methods, we measure a quantity t, defined as $t =< (\hat{x} - \hat{z})(\hat{x} - \hat{z}) > /\sqrt{V_Q[\hat{x}]V_Q[\hat{z}]}$, where $\hat{x}$ and $\hat{z}$ denote the estimations of UMLI and the network, respectively. $V_Q[\hat{x}]$ and $V_Q[\hat{z}]$ are their variances, and $<>$ represents averaging over many trials. This quantity measures the statistic difference between the two estimators. The smaller the value of t is, the more the two methods agree with each other. In the extreme case of $\hat{x} = \hat{z}$ for each trial, the value of t is zero. If the two estimators are completely independent of each other and have zero means, as in this example, the value of t is 2. Figure 4b shows that t is quite small in all correlation cases (the largest value is 0.032), which means that the network estimation can be regarded as the same as that of UMLI.

In a similar way, we can show that FMLI can be implemented by the same recurrent neural field (since both methods use the same tuning function to fit the data) but by using a different external input: $I_c = \epsilon \int h^{-1}_{c,c'} r_{c'}\, dc'$ (see appendix D). This result is different from that in Deneve et al. (1999), where FMLI and UMLI are implemented by using different recurrent interactions.

To implement COM, the external input is the same as that in UMLI (since both methods discard the correlation). However, a different form of network interaction is needed to ensure that the line attractor has the same shape of the corresponding tuning function.

It is interesting to compare the complexity of three methods in terms of network implementation. Obviously, FMLI is more complicated than UMLI, since it uses neural correlation. However, COM and UMLI are around the same level.

7 Conclusion and Discussion

We have proposed a new unified encoding field model for population coding, in which neural responses are correlated with a gaussian correlation function. This model serves as a good prototype for theoretical study. It has the advantages of simplifying the calculation and providing a clear picture of the results, which are often vague when other models are used.

Based on the proposed model, we calculate the Fisher information and elucidate its asymptotic behaviors for various correlation lengths. We confirm that as the correlation covers the short range of population, the Fisher information saturates and does not increase any more in proportion to the number of neurons. Moreover, we prove that as the correlation covers the wide range of the population, the Fisher information increases again without limit in proportion to the number of neurons. This finding, together with others, needs experimental confirmation on the range of neural correlation in the brain.

Three decoding methods are compared in this study. All are formulated as the MLI type, including the conventional COM method, whereas they differ in the knowledge of encoding process being used. It turns out that UMLI, which uses the shape of the tuning function but neglects the neural correlation, stands out due to the good balance between computational complexity and decoding accuracy. Furthermore, we investigate the network implementation of the above three methods. It shows that UMLI and FMLI can be achieved by the same recurrent network with different external inputs.

We also clarify the asymptotic efficiency (Fisher efficiency) of the MLI type of decoding methods for correlated signals. It is proved that when the neural correlation covers a nonlocal range of population (excluding the cases of uniform correlation and weak noise), the MLI type of method is non-Fisherian. The Cramér-Rao bound is not achievable in this case, and hence one should be careful carrying out analysis based on the Fisher information.

It is also important to be aware of the quasi-asymptotic efficiency of MLI based on unfaithful models. Only when the quasi-asymptotic efficiency is ensured can one calculate the decoding errors of UMLI and COM by equations 5.19 and 5.21, respectively. Otherwise, the decoding errors are subject to the Cauchy type of distribution and are difficult to quantify.

An interesting case that is not studied in this article is the multiplicative correlation, in which the correlation strength depends on firing rates (Abbott & Dayan, 1999; Nakahara & Amari, in press; Wilke & Eurich, in press). To cope with this situation, we can, for example, extend the correlation matrix A_{ij} in equation 2.7 to be a new one $A'_{ij} = f_i^{\alpha}(x) A_{ij} f_j^{\alpha}(x)$. In the case of $\alpha = 1$ and $b \to \infty$, it returns to $A'_{ij} = f_i(x)[(1-\beta)\delta_{ij} + \beta] f_j(x)$, which is the case studied by Abbott and Dayan (1999) and Wu et al. (2000). The continuous version of A'_{ij} in a neural field is $h'(c, c', x) = f^{\alpha}(x - c) h(c, c') f^{\alpha}(x - c')$, where $h(c, c')$ is given by equation 3.6, and its inverse is $(h')^{-1}(c, c', x) = f^{-\alpha}(x - c) h^{-1}(c, c') f^{-\alpha}(x - c')$. We can calculate the Fisher information and the performance of the MLI type of decoding methods much as we have done in this article. The results will be reported in a future publication.

Finally, we should point out that a nonlocal range correlation does not imply the failure of MLI-type methods. In addition to the counter-example of uniform correlation, another one is the multiplicative correlation, for example, $A_{ij} = [\delta_{ij} + c(1 - \delta_{ij})] f_i(x) f_j(x)$. It has been proved that the MLI-type method is asymptotically efficient in this case, although the correlation covers the whole range of the population (Wu, Chen, & Amari, 2000). There is an intuitive way to understand the reason (similarly for the uniform correlation). The idea is to decompose the fluctuations of neural responses into two parts, $r_i - f_i(x) = \sigma f_i(x)(\gamma + \epsilon_i)$, where γ and $\{\epsilon_i\}$, for $i = 1, \ldots, N$, are independent random variables having zero mean and variance c and $1 - c$, respectively. Neurons are correlated through the common factor γ,

which can be calculated by $\gamma = \sum_i [r_i - f_i(x)]/[\sigma \sum_i f_i(x)]$. This information can be used when MLI is performed.

Appendix A: The Covariance Function of the Continuous Encoding Model

Without loss of generality, we consider that the preferred stimulus c_i is uniformly distributed in a range $[-\frac{L}{2}, \frac{L}{2}]$, that is,

$$c_i = -\frac{L}{2} + i\frac{L}{N}, \qquad \text{for} \quad i = 1, \ldots, N. \tag{A.1}$$

By choosing a particular form of $\{k_i\}$, e.g., $k_i = 1$, for $i = 1, \ldots, N$, equation 3.3 becomes

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} \int_{-\frac{L}{2}}^{\frac{L}{2}} h(c, c')\, dc\, dc' = \sum_{ij} A_{ij}, \tag{A.2}$$

which has an intuitive meaning; the total correlation is reserved after the continuous extension.

From equation 3.2,

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} \int_{-\frac{L}{2}}^{\frac{L}{2}} h(c, c')\, dc\, dc' = D_1(1-\beta)L + D_2\beta\sqrt{2\pi}bL. \tag{A.3}$$

From equation 2.6,

$$\sum_{ij} A_{ij} = N(1-\beta) + N^2\beta\sqrt{2\pi}b/L, \tag{A.4}$$

in the large N limit.

Combining equations A.3 and A.4, we get

$$h(c, c') = \rho(1-\beta)\delta(c - c') + \rho^2\beta e^{-(c-c')^2/2b^2}, \tag{A.5}$$

where $\rho = \frac{N}{L}$ is the neuron density.

Appendix B: The Performance of FMLI

The performance of FMLI can be similarly analyzed as done for UMLI. Expanding $\nabla \ln Q(\mathbf{r}|\hat{x})$ at x and using the condition $\nabla \ln Q(\mathbf{r}|x) = 0$, we obtain an estimating equation,

$$\hat{x} - x = -\frac{R_F}{S_F}, \tag{B.1}$$

where the variables R_F and S_F are defined as

$$R_F = \frac{1}{\rho^2}\nabla \ln Q(\mathbf{r}|x)$$
$$= \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\epsilon(c)h^{-1}(c,c')f'(c'-x)\,dc\,dc', \tag{B.2}$$

$$S_F = \frac{1}{\rho^2}\nabla\nabla \ln Q(\mathbf{r}|x)$$
$$= \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\epsilon(c)h^{-1}(c,c')f''(c'-x)\,dc\,dc' + D_F, \tag{B.3}$$

where

$$D_F = \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f'(c-x)h^{-1}(c,c')f'(c'-x)\,dc\,dc'. \tag{B.4}$$

It is easy to check that the mean values of R_F and S_F are zero and D_F, respectively. Their variances and the value of D_F are given by using the Fourier transforms as

$$V_Q[R_F] = \frac{1}{2\pi\rho^2\sigma^2}\int_{-\infty}^{\infty}\omega^2 F(\omega)^2/H(\omega)\,d\omega, \tag{B.5}$$

$$V_Q[S_F] = \frac{1}{2\pi\rho^2\sigma^2}\int_{-\infty}^{\infty}\omega^4 F(\omega)^2/H(\omega)\,d\omega, \tag{B.6}$$

$$D_F = \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty}\omega^2 F(\omega)^2/H(\omega)\,d\omega. \tag{B.7}$$

These show that R_F is a zero-mean gaussian variable. The random variable S_F is composed of two terms, the first one being random and the second one, D_F, being a constant. The constant term dominates over the random one in the two cases:

- $H(\omega)$ is of order ρ. In this case, the random term is $O(\frac{1}{\rho^{3/2}})$, and the constant one is $O(\frac{1}{\rho})$.
- $H(\omega)$ is of order ρ^2, but the noise variance σ^2 is sufficiently small. In this case, the random one is $O(\frac{1}{\sigma})$, and the constant term is $O(\frac{1}{\sigma^2})$.

The first case corresponds to uncorrelated, local range, and uniformly correlated cases. The second case holds for the short- and wide-range correlations with small noise. In these cases, we may neglect the random term in S_F, so

that we have asymptotically

$$\begin{aligned} S_F &\approx D_F, \\ \hat{x} - x &\approx \frac{R_F}{D_F}, \end{aligned} \tag{B.8}$$

which is normally distributed with zero mean and variance,

$$E_Q[(\hat{x} - x)^2_{\mathrm{FMLI}}] \sim \frac{2\pi\sigma^2}{\rho^2 \int_{-\infty}^{\infty} \omega^2 F(\omega)^2 / H(\omega)\, d\omega}. \tag{B.9}$$

In the cases of short- and wide-range correlations and when the noise is strong, the random and constant terms in S_F are of the same order. Equation B.1 shows that the distribution of $\hat{x} - x$ is of the Cauchy type. FMLI is not asymptotically efficient in this case.

Appendix C: Minimizing the Lyapunov Function

When $\epsilon = 0$, the solution of minimizing the Lyapunov function of the form 6.7 (only the first terms are concerned) is given by

$$\begin{aligned} U_c &= \tilde{U}_c = Be^{-(c-z)^2/4a^2}, \\ O_c &= \tilde{O}_c = Ae^{-(c-z)^2/2\sigma^2}, \end{aligned} \tag{C.1}$$

for any value of z.

When a small input ϵr_c ($\epsilon \to 0$) is added, the solution will be uniquely determined and has a form slightly deviated from equation C.1, which can be approximated as (in the first order of ϵ)

$$\begin{aligned} U_c &\approx \tilde{U}_c + \epsilon E_c, \\ O_c &\approx \tilde{O}_c + \epsilon(2D\tilde{U}_c E_c), \end{aligned} \tag{C.2}$$

where ϵE_c and $\epsilon(2D\tilde{U}_c E_c)$ denote deviations from the linear attractor.

Substituting equation C.2 into 6.7, we get

$$\begin{aligned} L = &-\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{c,c'}\tilde{O}_c\tilde{O}_{c'}\, dc\, dc' + \int_{-\infty}^{\infty}\int_{0}^{\tilde{U}_c} zg'(z)\, dz\, dc \\ &- \epsilon\int_{-\infty}^{\infty} r_c\tilde{O}_c\, dc + \text{terms of order } \epsilon^2 \text{ and higher.} \end{aligned} \tag{C.3}$$

Note that the deviation of the waveform of the solution from the linear attractor contributes to the Lyapunov function only in the order of ϵ^2 or

higher. The contribution in the first order of ϵ, however, is determined by the overlapping between r_c and $\tilde{O}_c$ (the third term in equation C.3). Minimizing this term determines the position of stable state on the linear attractor and gives the estimation of stimulus.

The above procedure can be formally done in a two-step perturbation procedure described in the text.

Appendix D: The Network Implementation of FMLI

For FMLI, we consider that it is implemented by using the same recurrent neural field as for UMLI, but with a different external input, which is

$$I_c = \epsilon \int h^{-1}_{c,c'} r_{c'}\, dc'. \tag{D.1}$$

Following the same line as for UMLI, we obtain a Lyapunov function,

$$\begin{aligned} L = &-\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{c,c'} O_c O_{c'}\, dc\, dc' + \int_{-\infty}^{\infty}\int_{0}^{U_c} z g'(z)\, dz\, dc \\ &- \epsilon \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} O_c h^{-1}_{c,c'} r_{c'}\, dc\, dc', \end{aligned} \tag{D.2}$$

which can be approximately minimized in a two-step perturbation procedure, with the solution determined by

$$\begin{aligned} &\mathrm{Min}_{\{z\}} \quad -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} O_c h^{-1}_{c,c'} r_{c'}\, dc\, dc', \\ &\text{subject to} \quad O_c = A e^{-(c-z)^2/2\sigma^2}. \end{aligned} \tag{D.3}$$

Recall that the solution of FMLI is given by

$$\begin{aligned} &\mathrm{Max}_{\{x\}} \quad \ln P_F(\mathbf{r}|x), \\ &\qquad = \frac{\rho^2}{\sigma^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(c-x) h^{-1}_{c,c'} r_{c'}\, dc\, dc' + \text{terms not on } x. \end{aligned} \tag{D.4}$$

Comparing equations D.3 and D.4, we see that the network estimation is the same as that of FMLI.

Acknowledgments

We thank the two anonymous reviewers for their valuable comments. S. W. acknowledges helpful discussions with Peter Dayan. H. N. is supported by Grants-in-Aid 11780589 and 13210154 from the Ministry of Education, Japan.

References

Abbott, L. F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation, 11*, 91–101.

Akahira, M., & Takeuchi, K. (1981). *Asymptotic efficiency of statistical estimation: Concepts and high order asymptotic efficiency*. Berlin: Springer-Verlag.

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological Cybernetics, 27*, 77–87.

Baldi, P., & Heiligenberg, W. (1988). How sensory maps could enhance resolution through ordered arrangements of broadly tuned receivers. *Biological Cybernetics, 59*, 313–318.

Brunel, N., & Nadal, J.-P. (1998). Mutual information, Fisher information, and population coding. *Neural Computation, 10*, 1731–1757.

Cohen, M., & Grossberg, S. (1983). Absolute stability of global pattern formulation and parallel memory storage by competitive neural networks. *IEEE Trans. SMC, 13*, 815–826.

Deneve, S., Latham, P. E., & Pouget, A. (1999). Reading population codes: A neural implementation of ideal observers. *Nature Neuroscience, 2*, 740–745.

Eurich, C. W., Wilke, S. D., & Schwegler, H. (2000). Neural representation of multi-dimensional stimuli. In S. A. Solla, T. K. Leen, & K.-R. Müller (Eds.), *Advances in neural information processing systems, 12* (pp. 115–121). Cambridge, MA: MIT Press.

Fetz, E., Yoyama, K., & Smith, W. (1991). Synaptic interactions between cortical neurons. In A. Peters & E. G. Jones (Ed.), *Cerebral cortex, 9*. New York: Plenum Press.

Gawne, T. J., & Richmond, B. J. (1993). How independent are messages carried by adjacent inferior temporal cortical neurons? *J. Neuroscience, 13*, 2758–2771.

Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci., 2*, 1527–1537.

Giese, M. A. (1999). *Dynamic neural field theory for motion perception*. Norwell, MA: Kluwer Academic.

Johnson, K. O. (1980). Sensory discrimination: Neural processes preceding discrimination decision. *J. Neurophysiology, 43*, 1793–1815.

Lee, D., Port, N. L., Kruse, W., & Georgopoulos, A. P. (1998). Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J. Neuroscience, 18*, 1161–1170.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the Macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiology, 49*, 1127–1147.

Murata, M., Yoshizawa, S., & Amari, S. (1994). Network information criterion—determining the number of hidden units for artificial neural network model. *IEEE Trans. Neural Networks, 5*, 865–872.

Nakahara, H., & Amari, S. (in press). Attention modulation of neural tuning through peak and base rate in correlated firing. *Neural Networks*.

Nakahara, H., Wu, S., & Amari, S. (2001). Attention modulation of neural tuning through peak and base rate. *Neural Computation, 13*, 2031–2047.

Paradiso, M. A. (1988). A theory for use of visual orientation information which exploits the columnar structure for striate cortex. *Biological Cybernetics, 58*, 35–49.

Pouget, A., & Zhang, K. (1997). Statistically efficient estimation using cortical lateral connections. In M. Mozer, M. Jordan, & T. Petsche (Eds.), *Advances in neural processing systems, 9* (pp. 97–103). Cambridge, MA: MIT Press.

Pouget, A., Zhang, K., Deneve, S., & Latham, P. E. (1998). Statistically efficient estimation using population coding. *Neural Computation, 10*, 373–401.

Salinas, E., & Abbott, L. F. (1994). Vector reconstruction from firing rates. *J. Comp. Neurosci., 1*, 89–107.

Seung, H. S. (1996). How the brain keeps the eyes still. *Proc. Acad. Sci. USA, 93*, 13339–13344.

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA, 90*, 10749–10753.

Snippe, H. P. (1996). Parameter extraction from population codes: A critical assessment. *Neural Computation, 8*, 511–529.

Snippe, H. P., & Koenderink, J. J. (1992). Information in channel-coded systems: Correlated receivers. *Biological Cybernetics, 67*, 183–190.

Treue, S., Hol, K., & Rauber, H.-J. (2000). Seeing multiple directions of motion-physiology and psychophysics. *Nature Neuroscience, 3*, 270–276.

Wilke, S. D., & Eurich, C. W. (in press). Representational accuracy of stochastic neural population. *Neural Computation*.

Wu, S., Chen, D., & Amari, S. (2000). Unfaithful population decoding. In *Proceedings of the International Joint Conference on Neural Networks*.

Wu, S., Nakahara, H., & Amari, S. (2001). Population coding with correlation an unfaithful model. *Neural Computation, 13*, 775–797.

Wu, S., Nakahara, H., Murata, N., & Amari, S. (2000). Population decoding based on an unfaithful model. In S. A. Solla, T. K. Leen, & K.-R. Müller (Eds.), *Advances in neural information systems, 12* (pp. 192–198). Cambridge, MA: MIT Press.

Yoon, H., & Sompolinsky, H. (1999). The effect of correlations on the Fisher information of population codes. In M. S. Kearns, S. Solla, & D. Cohn (Eds.), *Advances in neural information processing systems, 11* (pp. 167–173). Cambridge, MA: MIT Press.

Zemel, R. S., Dayan, P., & Pouget, A. (1998). Probabilistic interpolation of population codes. *Neural Computation, 10*, 403–430.

Zhang, K-C. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *J. Neuroscience, 16*, 2112–2126.

Zhang, K., & Sejnowski, T. J. (1999). Neural tuning: To sharpen or broaden. *Neural Computation, 11*, 75–84.

Zohary, E. (1992). Population coding of visual stimuli by cortical neurons tuned to more than one dimension. *Biological Cybernetics, 66*, 265–272.

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neural discharge rate and its implications for psychophysical performance. *Nature, 370*, 140–143.

Received February 2, 2001; accepted September 18, 2001.