Optimal Tuning Widths in Population Coding of Periodic Variables

Journal Item

LETTER Communicated by Kechen Zhang

Optimal Tuning Widths in Population Coding of Periodic Variables

Marcelo A. Montemurro
m.montemurro@manchester.ac.uk
Stefano Panzeri
s.panzeri@manchester.ac.uk
Faculty of Life Sciences, University of Manchester, Manchester M60 1QD, U.K.

We study the relationship between the accuracy of a large neuronal population in encoding periodic sensory stimuli and the width of the tuning curves of individual neurons in the population. By using general simple models of population activity, we show that when considering one or two periodic stimulus features, a narrow tuning width provides better population encoding accuracy. When encoding more than two periodic stimulus features, the information conveyed by the population is instead maximal for finite values of the tuning width. These optimal values are only weakly dependent on model parameters and are similar to the width of tuning to orientation or motion direction of real visual cortical neurons. A very large tuning width leads to poor encoding accuracy, whatever the number of stimulus features encoded. Thus, optimal coding of periodic stimuli is different from that of nonperiodic stimuli, which, as shown in previous studies, would require infinitely large tuning widths when coding more than two stimulus features.

1 Introduction

The width of the tuning curves of individual neurons to sensory stimuli plays an important role in determining the nature of a neuronal population code (Pouget, Deneve, Ducom, & Latham, 1999; Zhang & Sejnowski, 1999). On the one hand, narrow tuning makes single neurons highly informative about a specific range of stimulus values. On the other hand, coarse tuning increases the size of the population activated by the stimulus, but at the price of making individual neurons less precisely tuned.

An important question is whether there is an optimal value of tuning width that allows a most effective trade-off between these two partly conflicting requirements of high encoding accuracy by single neurons and engagement of a large population. When considering the encoding of nonperiodic stimulus variables, such as the Cartesian coordinates of position in space, Zhang and Sejnowski (1999) demonstrated that under very general conditions, the information about the stimulus conveyed by the population

scales as σ^{D-2}, where σ stands for the width of the tuning curve and D is equal to the number of encoded stimulus features. Thus, extremely narrow tuning curves are better for encoding one nonperiodic stimulus feature, whereas extremely coarse tuning curves are better for encoding more than two nonperiodic features (Zhang & Sejnowski, 1999).

However, many important stimulus variables are described by periodic quantities. Example of such variables are the direction of motion and the orientation of a visual stimulus. Thus, it is crucial to investigate how the population accuracy in encoding periodic stimuli depends on the tuning width. Here, we address this problem, and we find that under general conditions and in marked contrast with the case of nonperiodic stimuli, the population accuracy in encoding periodic stimuli decreases quickly for large tuning widths, whatever the stimulus dimensionality. For stimulus dimensions $D > 2$, there is a finite optimal value of the tuning curve width for which the population conveys maximal information. If D is in the range 2 to 6, the optimal tuning widths predicted by the model are similar in magnitude to those observed in visual cortex. This suggests that the tuning widths of visual cortical neurons are efficient at transmitting information about a small number of periodic stimulus features.

2 Model of Encoding Periodic Variables

We consider a population made up of a large number N of neurons. The response of each individual neuron is quantified by the number of spikes fired in a certain poststimulus time window. Thus, the overall neuronal population response is represented as an N-dimensional spike count vector. We assume that the neurons are tuned to a small number D of periodic stimulus variables, such as the orientation or the direction of motion of a visual object. The stimulus variable will be described as an angular vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_D)$ of dimension D. A real cortical neuron may also encode sensory variables that are not periodic, such as retinal position or speed of motion. However, for simplicity here, we will focus entirely on periodic variables.

The neuronal tuning curves, which quantify the mean spike count of each neuron to the presented D-dimensional stimulus, are all taken to be identical in shape but having their maxima at different angles. Thus, each neuron is uniquely identified by its preferred angle ϕ. For concreteness, we choose the following multidimensional circular normal form for the tuning function,

$$f(\boldsymbol{\theta} - \phi) \equiv b + f_0(\boldsymbol{\theta} - \phi) = b + m \prod_{i=1}^{D} \exp\left(\frac{\cos(\nu(\theta_i - \phi_i)) - 1}{(\nu\sigma)^2}\right), \tag{2.1}$$

where b is the baseline (nonstimulus-induced) firing. The stimulus-modulated part of the tuning curve $f_0(\boldsymbol{\theta} - \boldsymbol{\phi})$ depends on σ, the width of tuning, and on m, the response modulation. As in Zhang and Sejnowski (1999), we assume that the preferred angles of each neurons are distributed across the population uniformly over the D-dimensional cuboid region $[-\pi/\nu, \pi/\nu]^D$. The parameter ν sets the period of the tuning function, which is equal to $(2\pi)/\nu$. For example, $\nu = 1$ corresponds to a period equal to 2π and would describe a motion direction angle, whereas $\nu = 2$ corresponds to a period equal to π and would describe an orientation angle.

For simplicity, we assume that different stimulus dimensions are mapped in a separable way. Thus, the stimulus-dependent part of the multidimensional tuning curve in equation 2.1 is written as a product of a one-dimensional circular normal function over the different stimulus dimensions:

$$g(\theta_i - \phi_i) = \exp\left(\frac{\cos(\nu(\theta_i - \phi_i)) - 1}{(\nu\sigma)^2}\right). \tag{2.2}$$

This tuning function has been used in neural coding models (see, e.g., Pouget, Zhang, Deneve, & Latham, 1998 and Sompolinsky, Yoon, Kang, & Shamir, 2001). It was chosen here because, unlike the most commonly used gaussian models, it is a genuinely periodic function of the stimulus, and it fits accurately experimental tuning curves for both orientation-sensitive (Swindale, 1998) and direction-sensitive (Kohn & Movshon, 2004) neurons in primary visual cortex. In Figure 1 we plot two examples of the one-dimensional circular normal distribution in equation 2.2, compared with their respective gaussian approximations. If σ is smaller than $\approx 20^0$, the circular normal function closely resembles a gaussian tuning curve (see equation). For tuning widths much larger than $\approx 20^0$, the shape of the circular tuning functions differs substantially from the gaussian.

The above assumption that the multidimensional tuning curve is just a product of one-dimensional tuning curves of individual features has been used extensively in population coding models. In addition to being, mathematically convenient for its simplicity, the multiplicative form of the multidimensional tuning function describes well the tuning of neurons in higher visual cortical areas to, for example, complex multidimensional rotations (Logothetis, Pauls, & Poggio, 1995) or to local features describing complex boundary configurations (Pasupathy & Connor, 2001).

We assume that the neurons in the population are uncorrelated at fixed stimulus. Thus, the stimulus-conditional probability of population response $P(\mathbf{r}|\boldsymbol{\theta})$ is a product of the spike count distribution of each individual neuron. Although it is a simplification, this assumption is useful because it makes our model mathematically tractable, and it is sufficient to account in most cases for the majority of information transmitted by real neuronal populations (Nirenberg, Carcieri, Jacobs, & Latham, 2001; Petersen,

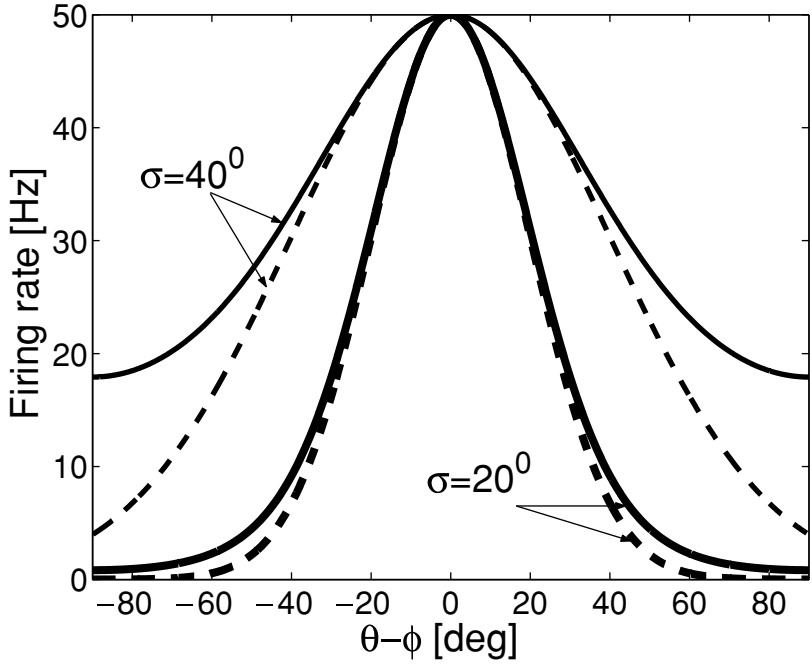

Figure 1: Comparison of periodic and nonperiodic tuning functions for an orientation-selective neuron coding one stimulus variable. The curves show mean firing rates as a function of the difference between the presented and the preferred stimulus of the neuron. Solid lines: periodic tuning function given by equation 2.2 for $\nu = 2$ (orientation selectivity) and for two values of the tuning width σ. Dashed lines: nonperiodic (gaussian) tuning function for the same tuning widths. The difference between the periodic and nonperiodic tuning functions becomes apparent for large tuning widths and for angles away from the preferred one.

Panzeri, & Diamond, 2001; Oram, Hatsopoulos, Richmond, & Donoghue, 2001). However, in section 8 we introduce a specific model that takes into account cross-neuronal correlations and demonstrates that the conclusions obtained with the uncorrelated assumption are still valid in that correlated case.

Following Zhang and Sejnowski (1999), we choose a general model of the activity of the single neuron in the population by requiring that the probability that the neuron with preferred angle $\boldsymbol{\phi}$ emits r spikes in response to stimulus $\boldsymbol{\theta}$ is an arbitrary function of the mean spike count only:

$$P(r|\boldsymbol{\theta}) = S(r, f(\boldsymbol{\theta} - \boldsymbol{\phi})). \tag{2.3}$$

In this article, some specific cases of single-neuron models that satisfy this assumption are studied in detail. We shall also derive scaling rules of the encoding efficiency as a function of σ and D that are valid for any model of single-neuron firing that satisfies equation 2.1.

3 Fisher Information

The ability of the population to encode accurately a particular stimulus value can be quantified by means of Fisher information (Cover & Thomas, 1991). When the stimulus is a D-dimensional periodic variable, Fisher information is a $D \times D$ matrix, $\mathbf{J}$, whose elements $i,\ j$ are measured in units of deg^{-2} and are defined as follows:

$$J_{i,j}(\boldsymbol{\theta}) = -\int d\mathbf{r} P(\mathbf{r}|\boldsymbol{\theta}) \left(\frac{\partial^2}{\partial\theta_i\, \partial\theta_j} \log P(\mathbf{r}|\boldsymbol{\theta}) \right) . \tag{3.1}$$

Fisher information provides a good measure of stimulus encoding because it sets a limit on the accuracy with which a particular stimulus value can be reconstructed from a single observation of the neuronal population activity. In fact, it satisfies the following generalized Cramér-Rao matrix inequality (Cover & Thomas, 1991),

$$\Sigma \geq \mathbf{J}^{-1} , \tag{3.2}$$

where Σ is the covariance matrix of the D-dimensional error made by any unbiased estimation method reconstructing the stimulus from the neuronal population activity.

Since the neurons fire independently, the population Fisher information is simply given by the sum over all neurons of the single-neuron Fisher information (Cover & Thomas, 1991). The latter, denoted as $J_{i,j}^{(neuron)}(\boldsymbol{\theta} - \boldsymbol{\phi})$, has the following expression:

$$J_{i,j}^{(neuron)}(\boldsymbol{\theta} - \boldsymbol{\phi}) = -\int dr\, S(r,\, f(\boldsymbol{\theta} - \boldsymbol{\phi})) \left(\frac{\partial^2}{\partial\theta_i\, \partial\theta_j} \log S(r,\, f(\boldsymbol{\theta} - \boldsymbol{\phi})) \right) . \tag{3.3}$$

By computing explicitly the derivatives in the above expression, one can rewrite the single-neuron Fisher information in equation 3.3 as follows:

$$J_{i,j}^{(neuron)}(\boldsymbol{\theta} - \boldsymbol{\phi}) = \left[f_0^2(\boldsymbol{\theta} - \boldsymbol{\phi}) \int dr\, \frac{S'(r,\, f(\boldsymbol{\theta} - \boldsymbol{\phi}))^2}{S(r,\, f(\boldsymbol{\theta} - \boldsymbol{\phi}))} \right] \times \frac{\sin((\nu(\theta_i - \phi_i))\sin(\nu(\theta_j - \phi_j))}{\nu^2\sigma^4} , \tag{3.4}$$

where $S' = \frac{\partial}{\partial z} S(r, z)$, and we require the integral over responses in the above equation to be convergent, so that the single-neuron Fisher information is finite.

Since the number of neurons in the population is assumed to be large and since a neuron is uniquely identified by the center $\boldsymbol{\phi}$ of its tuning curve, we can compute the population Fisher information by approximating the sum over the neurons by an integral over the preferred angles $\boldsymbol{\phi}$,

$$J_{i,j}(\boldsymbol{\theta}) = \frac{N\nu^D}{(2\pi)^D} \int_{-\pi/\nu}^{\pi/\nu} d^D\boldsymbol{\phi}\, J_{i,j}^{(neuron)}(\boldsymbol{\theta} - \boldsymbol{\phi}), \tag{3.5}$$

where $\int_{-\pi/\nu}^{\pi/\nu} d^D\boldsymbol{\phi}$ denotes the angular integration over the D-dimensional cuboid region $[-\pi/\nu, \pi/\nu]^D$. Since the term in square brackets in equation 3.4 is an even function of each of the angular variables, the integral in equation 3.5 will be nonzero only when $i = j$. It is also clear that because of symmetry over index permutations, the population Fisher information $J_{i,j}(\boldsymbol{\theta})$ is proportional to the identity matrix. Moreover, since the integrand in equation 3.5 is a periodic function integrated over its whole period, the Fisher information in the equation does not depend on the value of angular stimulus variable. Thus, dropping index notation and the $\boldsymbol{\theta}$-dependency in the argument, we will denote by J the diagonal element of the Fisher information matrix, and call it simply Fisher information.

4 Poisson Model

We begin our analysis of population coding of periodic stimuli by considering a neuronal firing model satisfying equation 2.3: we assume that single-neuron statistics at fixed stimulus is described by a Poisson process with mean given by the neuronal tuning function, equation 2.1:

$$P(r|\boldsymbol{\theta}) = \frac{(f(\boldsymbol{\theta} - \boldsymbol{\phi}))^r}{r!} \exp\left(-f(\boldsymbol{\theta} - \boldsymbol{\phi})\right). \tag{4.1}$$

The Fisher information conveyed by the Poisson neuronal population is as follows,

$$J = \frac{N\nu^D}{(2\pi)^D} \int_{-\frac{\pi}{\nu}}^{\frac{\pi}{\nu}} d^D\boldsymbol{\phi} \left[\frac{f_0^2(\boldsymbol{\theta} - \boldsymbol{\phi})}{f(\boldsymbol{\theta} - \boldsymbol{\phi})} \frac{\sin^2(\nu(\theta_1 - \phi_1))}{\nu^2\sigma^4} \right], \tag{4.2}$$

where the term in square brackets is the single-neuron Fisher information, and θ_1 and ϕ_1 denote the projections of $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ along the first stimulus dimension. In the following, we will study how the Poisson population Fisher information depends on the model parameters.

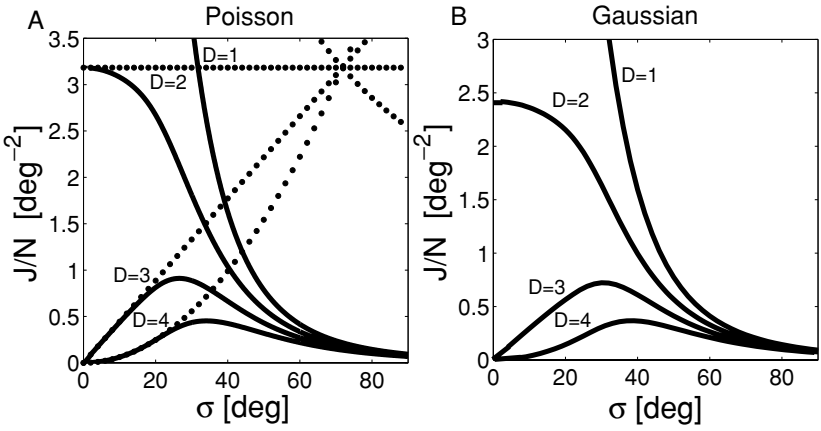

Figure 2: (A) Fisher information per neuron J/N as a function of the tuning curve width σ for a population of orientation-selective "Poisson" neurons, for stimulus dimensions $D = 1$, 2, 3, and 4. Solid lines correspond to the periodic stimulus Fisher information J and were calculated with equation 4.3. For plotting, the parameter m (which has only a trivial multiplicative effect) was fixed to 5. Dotted lines correspond to the nonperiodic Fisher information J_{np} and were calculated using gaussian tuning curves of width σ (Zhang & Sejnowski, 1999). (B) Fisher information per neuron J/N for a population of gaussian independent neurons, for different dimensions of the periodic stimulus variable. Parameters were as follows: $m = 5$; $b = 0.5$, $\alpha = 1$, and $\beta = 1$.

4.1 Analytical Solution with No Baseline Firing. First, we consider the case in which there is no baseline firing: $b = 0$. In this case, it is possible to integrate exactly equation 4.2 and obtain the following analytical solution for J,

$$J = \frac{Nm}{\sigma^2} K_1(\nu^2\sigma^2) K_0^{D-1}(\nu^2\sigma^2), \tag{4.3}$$

where

$$K_n(x) = e^{-1/x} I_n(1/x), \tag{4.4}$$

and $I_n(x)$ stands for the nth order modified Bessel function of the first kind. As in the nonperiodic case (Zhang & Sejnowski, 1999), N and m affect the Fisher information in equation 4.3 only as trivial multiplicative factors. Thus, we can focus on the dependence of Fisher information on σ and D. Figure 2A compares the periodic-stimulus Fisher information J (normalized in the plot as Fisher information per neuron J/N) to that obtained in the nonperiodic case (which we will denote as J_{np}, and has

a very simple σ-dependence: $J_{np} \propto \sigma^{D-2}$; see Zhang & Sejnowski, 1999). It is apparent that the periodic-stimulus Fisher information J depends on σ in a more complex way than the nonperiodic one J_{np}. While for $D = 1$ there is no qualitative difference between the periodic and the nonperiodic case (with both J and J_{np} being divergent at $\sigma = 0$), significant differences appear for $D \geq 2$. If $D = 2$, J is not constant with σ, but it has a maximum at $\sigma = 0$ and then decays rapidly. If $D > 2$, in sharp contrast with J_{np}, J exhibits a maximum at finite σ. The optimal values of σ that maximized J were 26.6, 34.1, 39.9, 44.9 degrees for $D = 3, 4, 5$, and 6, respectively.

The dependence of J on σ and D and its relation with J_{np} can be understood by comparing their respective expressions and taking into account the properties of the functions $K_n(\nu^2\sigma^2)$, as follows.

For small σ, $K_n(\nu^2\sigma^2)$ scales to zero as σ. Thus, the small-σ scaling of the periodic-stimulus model is identical to that of Zhang & Sejnowski (1999). This is because as $\sigma \to 0$, the periodic tuning function tends to a gaussian and σ can be rescaled away from the angular integrals as in the nonperiodic case.

For large σ, $K_0(\nu^2\sigma^2)$ increases monotonically toward 1, whereas $K_1(\nu^2\sigma^2)$ (which has a maximum at $\nu\sigma \approx 0.8$) decreases toward zero as $1/\sigma^2$. Thus, J decreases rapidly to zero as $1/\sigma^4$ for any D. This is very different from the nonperiodic case, in which J_{np} grows to infinity for large σ if $D > 2$ (Zhang & Sejnowski, 1999).

The occurrence of a finite-σ maximum of J can also be understood in terms of the K_n functions. If D is 1 or 2, then the dominant factor is $1/\sigma^2$, which leads to a maximum of J at $\sigma = 0$. For larger (but finite) D, the term K_0^{D-1} becomes more and more important, and thus the maximum of J is shifted toward larger σ. However, unlike in the nonperiodic case, since K_0 saturates at 1 and K_1 goes to zero for $\sigma \to \infty$, this maximum must be reached at a finite σ value. An infinite optimal σ value can only be reached in the $D \to \infty$ limit, where K_0^{D-1} is dominant and the other terms can be neglected.

It is interesting to compare the optimal values of tuning width obtained with our model with the widths of orientation tuning curves observed in visual cortex (summarized in Table 1). The width of orientation tuning of V1 neurons is typically in the range of 17 to 21 degrees. In higher cortical visual areas, the tuning curves get progressively broader: σ is approximately 26 degrees in MT and 38 degrees in V4. Thus, tuning widths of cortical neurons are neither too narrow nor too wide and are similar in magnitude to the optimal ones obtained from information-theoretic principles when considering multidimensional periodic stimuli.

What is the advantage of using tuning widths in this intermediate range of 15 to 40 degrees observed in cortex? Our model results suggest that unlike very narrow or very wide tuning widths, tuning widths in this intermediate range are rather efficient at conveying information over a range of stimulus dimensions. Intermediate tuning widths

Table 1: Typical Values of Tuning Widths to Either Orientation or Motion Direction of Neurons in Different Visual Cortex Areas.

Species	Area	Stimulus	σ[deg]
Ferret	V1	O	15–17*
Cat	V1	O	14.6†
Macaque	V1	O	19–22*
Macaque	MT	O	24–27*
Macaque	V4	O	38†
Macaque	V1	D	25–29*
Macaque	MT	D	35–40*

Notes: These values were taken or derived from published reports. † = the original data were reported as the standard deviation σ of the experimental tuning curve. * = the original published values were given as full widths at half height, and we then converted into σ using equation 2.1 with $D = 1$. Since the conversion depends on the value of the baseline firing b, we report the converted σ assuming a baseline firing ranging from 10% of the response modulation m (lower σ value) to zero baseline firing (upper σ value). Sources: Usrey, Sceniak, and Chapman (2003); Henry, Dreher, and Bishop (1974); Albright (1984); McAdams and Maunsell (1999).

(e.g., $\sigma = 25$ degrees) would not be efficient for $D = 1$. However, they would be highly efficient at encoding a handful of periodic stimulus features (e.g., $D = 2, \cdots, 5$). On the other hand, using a small width of tuning, say, $\sigma = 5$ degrees, would be more efficient for D = 1, only marginally more efficient for $D = 2$, but very inefficient for $D > 3$. Using a large σ of 90 degrees would lead to poor information transmission for any value of D below $\approx$15. Therefore, the advantage of intermediate tuning widths is that they offer a highly efficient information transmission across a range of stimulus dimensions, provided that the population encodes more than one stimulus feature.

The results obtained for orientation tuning are easily extended to coding of direction stimulus variables (i.e., $\nu = 1$). It is easy to see that apart from an overall multiplicative factor ν^2, the Fisher information in equation 4.3 depends on σ only through the product $\nu\sigma$ (Notice that this is true not only for the Poisson model solution in equation 4.3, but also for the general Fisher information in equation 3.5). Thus, the dependence of the information on σ for direction-selective populations will be the same as that obtained for orientation-selective neurons, with an overall rescaling of σ by 2. Therefore, for any D, optimal tuning widths for direction-selective populations are exactly twice those corresponding to orientation-selective populations. Table 1 shows that in both V1 and MT, the motion direction tuning widths are larger than the orientation tuning widths, by a factor of $\approx$1.5. Cortical direction tuning widths are hence also efficient for $D > 1$.

4.2 Effect of Baseline Firing. If there is baseline firing, that is, $b > 0$, it is not possible to express J by a simple analytical formula such as equation 4.3. However, we can gain insight into the effect of baseline firing by obtaining approximate solutions for the limiting cases of very small and very large b, as follows. Here we will focus on how robust the optimal σ values are that were found above in the $D > 2$ case.

When b is very small, we can expand the integrand in equation 4.2 in powers of $b/f_0(\boldsymbol{\theta} - \boldsymbol{\phi})$.[1] Keeping up to first order in $b/f_0(\boldsymbol{\theta} - \boldsymbol{\phi})$, we obtain:

$$J = \frac{Nm}{\sigma^2}\left[K_1(\nu^2\sigma^2)K_0^{D-1}(\nu^2\sigma^2) - \frac{b}{2m}\right]. \tag{4.5}$$

The first term corresponds to the Fisher information, equation 4.3, for the case $b = 0$, which has a maximum at a certain value of tuning curve width, which we indicate by $\sigma*$. The second term is a perturbative correction that decreases monotonously when increasing σ. Consequently, for $D > 2$, the effect of a small baseline firing is to increase slightly the optimal σ with respect to the $\sigma*$ values obtained for $b = 0$.

How much can the optimal σ values increase when increasing the baseline firing? This can be established by considering the opposite limit, $b \to \infty$. In this case, $f_0(\boldsymbol{\theta} - \boldsymbol{\phi})/b \ll 1$ for all angles, and we can expand the integrand in equation 4.2 in powers of $f_0(\boldsymbol{\theta} - \boldsymbol{\phi})/b$. Up to first order in b^{-1}, we find:

$$J = \frac{Nm^2}{b\,\sigma^2}\left[K_1\left(\frac{\nu^2\sigma^2}{2}\right)K_0^{D-1}\left(\frac{\nu^2\sigma^2}{2}\right) - \frac{m}{b}K_1\left(\frac{\nu^2\sigma^2}{3}\right)K_0^{D-1}\left(\frac{\nu^2\sigma^2}{3}\right)\right]. \tag{4.6}$$

The first contribution is the asymptotic behavior for $b \to \infty$ and has a maximum at $\sigma = \sqrt{2}\sigma^*$. It can be shown that the correction term tends to decrease the optimal value of σ from its $\sqrt{2}\sigma^*$ asymptotic large-b value. This suggests that the maximal effect of baseline firing on optimal σ values consists in an increase of a factor $\sqrt{2}$ with respect to the optimal value for $b = 0$, and this maximal effect is reached when baseline firing dominates. We have verified this prediction by integrating numerically equation 4.2 for different values of b. We found that (data not shown) for $D = 3$ and 4, the optimal σ values were located, for any b, between $\sigma*$ and $\sqrt{2}\sigma*$. Thus, optimal tuning width values are robust to the introduction of baseline firing.

[1] This expansion is valid if $b/f_0(\boldsymbol{\theta} - \phi) \ll 1$ for all angles. This condition can be met if σ is nonzero and b is sufficiently small.

5 Gaussian Model

We consider next another model of single-cell firing, the gaussian model, which describes well the statistics of spike counts of visual neurons when spikes are counted over sufficiently long windows (Abbott, Rolls, & Tovée, 1996; Gershon, Wiener, Latham, & Richmond, 1998). This model, unlike the Poisson one, permits considering the effect of autocorrelations between the spikes emitted by the same neuron. The statistics of single-neuron spike counts for the gaussian model are defined as follows:

$$P(r|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\psi(\zeta)} \exp\left(-\frac{(r - f(\boldsymbol{\theta} - \boldsymbol{\phi}))^2}{2\psi^2(\zeta)}\right), \tag{5.1}$$

where the standard deviation of spike counts, ψ, is an arbitrary function of the mean: $\zeta = f(\boldsymbol{\theta} - \boldsymbol{\phi})$. Under these assumptions, it is easy to show that the population Fisher information J has the following expression,

$$J = \frac{N\nu^D}{(2\pi)^D} \int_{-\frac{\pi}{\nu}}^{\frac{\pi}{\nu}} d^D\boldsymbol{\phi} \left[f_0^2(\boldsymbol{\theta} - \boldsymbol{\phi}) \left(\frac{1}{\psi^2(\zeta)} + \frac{2\psi'^2(\zeta)}{\psi^2(\zeta)} \right) \frac{\sin^2(\nu(\theta_1 - \phi_1))}{\nu^2\sigma^4} \right], \tag{5.2}$$

where the term in square brackets is the single-neuron Fisher information, and $\psi'(\zeta) = \frac{\partial}{\partial\zeta}\psi(\zeta)$.

Since the variance of spike counts of cortical neurons is well described by a power law function of the mean spike count (Tolhurst, Movshon, & Thompson, 1981; Gershon et al., 1998), from now on we will assume that ψ is a power law function of the tuning curve:

$$\psi = \alpha^{1/2} f^{\beta/2}(\boldsymbol{\theta} - \boldsymbol{\phi}). \tag{5.3}$$

In this case, equation 5.2 becomes

$$J = \frac{N\nu^D}{(2\pi)^D} \int_{-\frac{\pi}{\nu}}^{\frac{\pi}{\nu}} d^D\boldsymbol{\phi} \times \left[f_0^2(\boldsymbol{\theta} - \boldsymbol{\phi}) \left(\frac{1}{\alpha f^\beta(\boldsymbol{\theta} - \boldsymbol{\phi})} + \frac{\beta^2}{2f^2(\boldsymbol{\theta} - \boldsymbol{\phi})} \right) \frac{\sin^2(\nu(\theta_1 - \phi_1))}{\nu^2\sigma^4} \right]. \tag{5.4}$$

For cortical neurons, the parameters α and β are typically close to 1, both distributed within the range 0.8 to 1.4 (Gershon et al., 1998; Dayan & Abbott, 2001). If all spike times are independent, then the spiking process is Poisson, and α and β would both be 1. Deviations of α and β from 1 require the presence of autocorrelations. Therefore, the study of how J depends on α and

β allows an understanding of how autocorrelations influence population coding.

The dependence of the gaussian model Fisher information on σ and D is plotted in Figure 2B. For this plot, we chose $m = 5$. The parameters α and β were both set to 1, so that, as for the Poisson process, the variance of spike counts equals the mean. In this case, the scaling of Fisher information is essentially identical to that of the Poisson model in Figure 2A.

We investigated numerically the dependence of J on α and β. We found that the shape of Fisher information plotted in Figures 2A and 2B is conserved across the entire range analyzed. In particular, J scaled to zero for large σ for any D and scaled as σ^{D-2} for small σ. For $D = 1$, J was always divergent at $\sigma = 0$. For $D = 2$, the maximum was always at $\sigma = 0$. For $D > 2$, J always had a maximum at finite values of σ. The position of the maximum varied slightly as a function of α and β. Results for the position of the maximum for $D = 3$ and $D = 4$ are reported in Figure 3. The position of the maxima was almost unchanged when varying α. They varied within less than 3 degrees for $D = 3$ and 5 degrees for $D = 4$ when β varied within the typical cortical range 0.8 to 1.4.

The similarity between the gaussian model Fisher information, equation 5.4, and the Poisson model Fisher information, equation 4.3, can be explained by noting that if m is large and $\beta < 2$, then the second additive term within parentheses in equation 5.4 can be neglected and the gaussian model Fisher information has the following approximated solution:

$$J \approx \frac{Nm^{2-\beta}}{\alpha(2-\beta)\sigma^2} K_1\left(\frac{\nu^2\sigma^2}{2-\beta}\right) K_0^{D-1}\left(\frac{\nu^2\sigma^2}{2-\beta}\right). \tag{5.5}$$

This expression is (apart from an overall multiplicative factor) identical to the population Fisher information for the Poisson model, equation 4.3, with an overall rescaling of the arguments of the K functions by a factor $2 - \beta$. Therefore, if the exponent β of the power law variance-mean relationship, equation 5.3, is approximately 1 (as for real cortical neurons), the optimal values of the gaussian model in the large-m case are almost identical to the ones obtained for the Poisson population. It is worth noting that the α- and β-dependence of J arising from the large-m approximation in equation 5.5 are compatible with the intermediate-m numerical results of Figure 3, which showed that the optimal σ values depend very mildly on α and decrease monotonically with β at fixed D.

If m is not very large, then the second additive term within parentheses in equation 5.4 contributes to the gaussian Fisher information. However, we have verified that under a wide range of parameters, this contribution is less dependent on σ than the other one and does not shift the maxima or alter the dependence of J on σ and D in a prominent way.

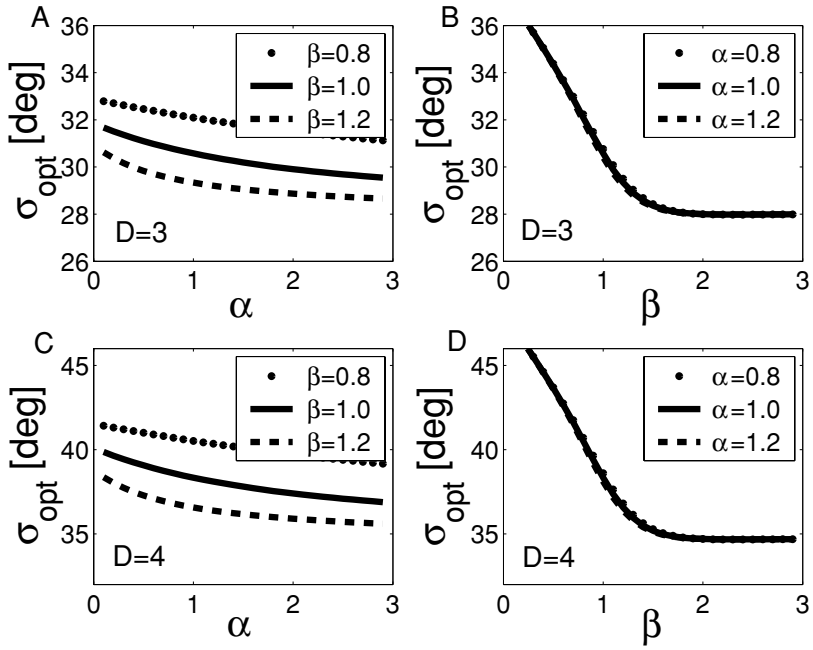

Figure 3: Optimal tuning width, σ_{opt} (corresponding to a maximum of Fisher information J), for a population of orientation-selective "gaussian" neurons, as a function of the parameters α and β defining the power law spike count mean variance relationship, equation 5.3. (A) The number of encoded stimulus variables is $D = 3$. Here β is kept fixed while α is varied. (B) Here, α is kept fixed, β is varied, and $D = 3$. (C) Now D is 4; β is kept fixed, while α is varied. (D) Again D is 4. Here, α is kept fixed, and β is varied.

Thus, we conclude that the values of optimal tuning widths obtained with the Poisson model are robust to changes in model details such as the introduction of autocorrelations parameterized by α and β, as long as these parameters remain within the realistic cortical range.

6 General Multiplicative Noise Model

To further check the robustness of the above conclusions, we introduced a more general model of single-neuron firing: the multiplicative noise model. This model, unlike the Poisson and the gaussian models, has the advantage of not assuming a particular functional form for the variability of neuronal responses at fixed stimulus. It assumes that the variability of spike counts in response to any stimulus is generated by an arbitrary stochastic process

modulated by an arbitrary function ψ of the mean spike rate. In this case, the spike rate of each neuron is given by the following equation,

$$r = f(\boldsymbol{\theta} - \boldsymbol{\phi}) + \epsilon\, \psi(\zeta) z\,, \tag{6.1}$$

where z is an arbitrary stochastic process of zero mean and unit variance with distribution $Q(z)$, $\psi(\zeta) \equiv \psi\,(f(\boldsymbol{\theta} - \boldsymbol{\phi}))$, and ϵ is a parameter that modulates the overall strength of the response variability. Under these assumptions, the single-neuron's spike count probability is

$$P(r|\boldsymbol{\theta}) = \frac{1}{\epsilon\psi(\zeta)} Q\left(\frac{(r - f(\boldsymbol{\theta} - \boldsymbol{\phi}))}{\epsilon\psi(\zeta)}\right), \tag{6.2}$$

and the single-neuron Fisher information, equation 3.3, has the following form:

$$\begin{aligned} J_{i,i}^{(neuron)}(\boldsymbol{\theta}) = & \frac{f_0^2(\boldsymbol{\theta} - \boldsymbol{\phi}) \sin^2(\nu(\theta_i - \phi_i))}{\epsilon^2 \psi^2(\zeta) \nu^2 \sigma^4} \\ & \times \left[T_0(Q) + 2\epsilon\psi'(\zeta) T_1(Q) + \epsilon^2 \psi'^2(\zeta) T_2(Q)\right]. \end{aligned} \tag{6.3}$$

The coefficients $T_i(Q)$ are a function of the noise distribution only and are defined as follows:

$$\begin{aligned} T_0(Q) &= \int \frac{Q'^2(z)}{Q(z)} dz\,; \qquad T_1(Q) = \int Q'(z) dz + \int \frac{Q'^2(z) z}{Q(z)} dz \\ T_2(Q) &= 1 + \int \frac{Q'^2(z)}{Q(z)} z^2 dz + 2 \int Q'(z) z dz. \end{aligned} \tag{6.4}$$

Although equation 6.3 permits the computation of the population Fisher information for any multiplicative noise model, in the following we will concentrate on examining two interesting limiting cases: very low and very high noise strengths. In examining these two cases, we will assume that (as for the gaussian model) the variance of the noise ψ^2 is in a power law relationship with the mean, equation 5.3.

6.1 Low Noise Limit. We first consider the low noise limit $\epsilon \ll 1$. In this case, responses are almost deterministic, and single neurons convey information by stimulus-induced changes in the mean spike rate (Brunel & Nadal, 1998). In this limit, the population Fisher information J can be calculated by keeping the leading order in ϵ only in the single-neuron

Fisher information, equation 6.3, and then integrating it over the preferred stimuli, obtaining:

$$J = \frac{Nm^{2-\beta}T_0(Q)}{\alpha(2-\beta)\sigma^2} K_1\left(\frac{\nu^2\sigma^2}{2-\beta}\right) K_0^{D-1}\left(\frac{\nu^2\sigma^2}{2-\beta}\right). \tag{6.5}$$

The dependence of the low-noise Fisher information on σ and D is thus affected only by β, with all other model parameters entering only as an overall multiplicative factor. Equation 6.5 is identical (apart from an overall multiplicative factor) to the large-m approximation of the gaussian model Fisher information, equation 5.5. It is also identical (apart from a rescaling of the argument of K_n) to the Poisson model exact solution in equation 4.3.

6.2 High Noise Limit. We considered next the case of very noisy neurons: $\epsilon \gg 1$. In this limit, information is transmitted entirely by stimulus-modulated changes of the variance ψ^2. (If the variance was not stimulus dependent, then information would be zero in the high noise limit; Samengo & Treves, 2001). Taking the $\epsilon \to \infty$ of the single-neuron Fisher information, integrating it over the preferred angles, and taking into account equation 5.3, we obtain the following expression for the population Fisher information:

$$J = \frac{N\nu^D}{(2\pi)^D} \int_{-\frac{\pi}{\nu}}^{\frac{\pi}{\nu}} d^D\boldsymbol{\phi} \left[f_0^2(\boldsymbol{\theta}-\boldsymbol{\phi}) \left(\frac{T_2(Q)}{\alpha f^\beta(\boldsymbol{\theta}-\boldsymbol{\phi})}\right) \frac{\sin^2(\nu(\theta_1-\phi_1))}{\nu^2\sigma^4}\right]. \tag{6.6}$$

In this limit, J is thus independent of the noise strength ϵ and depends on the details of the single-neuron model only through a multiplicative factor $T_2(Q)$. It can be seen that its expression is similar to the first (and dominant) additive term of the gaussian solution, equation 5.4. Because of this similarity, when integrating numerically equation 6.6, we found that the dependence of the high-noise population Fisher information on σ and D is remarkably consistent with that obtained for the Poisson and gaussian models (see Figure 2), and that the changes in the position of the maxima when varying α and β were again similar to those reported in Figure 3 for the gaussian model (data not shown).

In summary, the dependence on σ and D of the information transmitted by a population of neurons described by the general multiplicative model behaves in a way consistent with the results obtained above in both the high-noise and the low-noise limit.

7 General Scaling Limits for Large and Small σ

After having analyzed three different classes of single-neuron firing models, in this section we switch back to the most general firing model in equation

2.3, in which the single-neuron statistics is an arbitrary function of the mean spike rate; we consider its small- and large-σ scaling. We will derive that for any such firing model, Fisher information scales in an universal way as σ^{D-2} for small σ and goes to zero as $1/\sigma^4$ for large σ. Thus, for $D > 2$, Fisher information reaches a maximum for a finite value of the tuning width, whatever the firing model considered.

7.1 Small σ Scaling. When $\sigma \ll 1$, the exponential in equation 2.1 gives a nonzero contribution to the tuning function only when $\boldsymbol{\theta} - \boldsymbol{\phi} \sim 0$. In this regime, we can take a Taylor expansion of the cosines in the exponent of equation 2.1, obtaining the following,

$$f(\boldsymbol{\theta} - \boldsymbol{\phi}) \simeq b + m \exp\left(\frac{-|\boldsymbol{\theta} - \boldsymbol{\phi}|^2}{2\sigma^2}\right) \equiv G\left(\frac{|\boldsymbol{\theta} - \boldsymbol{\phi}|^2}{\sigma^2}\right), \tag{7.1}$$

where in the above $G(|\boldsymbol{\theta} - \boldsymbol{\phi}|^2/\sigma^2)$ is the standard gaussian tuning function. The population Fisher information becomes

$$J = \frac{N\nu^D}{(2\pi)^D} \int_{-\pi/\nu}^{\pi/\nu} d^D\boldsymbol{\phi}\, \tilde{A}\left(\frac{|\boldsymbol{\theta} - \boldsymbol{\phi}|^2}{\sigma^2}\right) \frac{(\theta_1 - \phi_1)^2}{\sigma^2}, \tag{7.2}$$

where, following Zhang and Sejnowski (1999), the function $\tilde{A}$ is defined as follows:

$$\tilde{A}\left(\frac{|\boldsymbol{\theta} - \boldsymbol{\phi}|^2}{\sigma^2}\right) = \exp\left(-\frac{|\boldsymbol{\theta} - \boldsymbol{\phi}|^2}{\sigma^2}\right) \int dr \frac{S'\left(r, G\left(\frac{|\boldsymbol{\theta}-\boldsymbol{\phi}|^2}{\sigma^2}\right)\right)^2}{S\left(r, G\left(\frac{|\boldsymbol{\theta}-\boldsymbol{\phi}|^2}{\sigma^2}\right)\right)}. \tag{7.3}$$

By introducing new integration variables $\xi_i = (\theta_i - \phi_i)/\sigma$, one can rewrite equation 7.2 as follows:

$$J = \sigma^{D-2} \frac{N\nu^D}{(2\pi)^D} \int_{-\pi/(\nu\sigma)}^{\pi/(\nu\sigma)} d^D\xi\, \tilde{A}\left(|\xi|^2\right) \xi_1^2. \tag{7.4}$$

By taking the small-σ limit of the above expression, one gets

$$J = \sigma^{D-2} \frac{N\nu^D}{(2\pi)^D} \int_{-\infty}^{+\infty} d^D\xi\, \tilde{A}\left(|\xi|^2\right) \xi_1^2. \tag{7.5}$$

If the improper integral above converges, then the periodic Fisher information scales as σ^{D-2} for small σ, coinciding with the universal scaling rule for nonperiodic stimuli found by Zhang and Sejnowski (1999). It is important

to note that for a given neuronal model defined by a probability function S, equation 7.5 is the Fisher information obtained if the tuning curve was gaussian with variance σ and the stimulus variable ξ was nonperiodic (Zhang & Sejnowski, 1999). Thus, the small-σ scaling of the periodic-stimulus Fisher information J is $\propto \sigma^{D-2}$ whenever the Fisher information of the corresponding gaussian nonperiodic tuning model is well defined (see Wilke & Eurich, 2001, for cases and parameters in which the gaussian model nonperiodic Fisher information is not well defined).

7.2 Large σ Scaling. When $\sigma \gg 1$, the argument of the exponentials in equation 2.1 is very small. Thus, the following expansion will be valid:

$$f(\boldsymbol{\theta} - \boldsymbol{\phi}) \simeq b + m + m \sum_{i=1}^{D} \left(\frac{\cos(\nu(\theta_i - \phi_i)) - 1}{(\nu\sigma)^2} \right) \approx f(\mathbf{0}) + O\left(\frac{1}{\sigma^2}\right). \tag{7.6}$$

Consequently, the population Fisher information becomes

$$J \approx \frac{1}{\sigma^4} \left[m^2 \int dr \frac{S'(r, m+b)^2}{S(r, m+b)} \right] \int d\phi \frac{\sin^2(\theta_1 - \phi_1)}{\nu^2}. \tag{7.7}$$

Thus, for large σ, Fisher information goes to zero as σ^4 for any stimulus dimensionality.

8 Correlations Between Neurons

The analysis above considered a population of independent neurons. In this section, we address the effect of correlated variability among the populations on the position of the optimal values of the tuning curve width. For simplicity, we shall consider that the firing statistics are governed by a multivariate gaussian distribution as follows,

$$P(\mathbf{r}|\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}|}} e^{-\frac{1}{2}(\mathbf{r}-\mathbf{f})^T \mathbf{C}^{-1} (\mathbf{r}-\mathbf{f})} \tag{8.1}$$

where $\mathbf{C}$ is the population correlation matrix and $\mathbf{f}$ stands for a column vector whose components are the neuron tuning functions, that is, $\mathbf{f} \equiv [f(\boldsymbol{\theta} - \boldsymbol{\phi}^{(1)}), \ldots, f(\boldsymbol{\theta} - \boldsymbol{\phi}^{(N)})]^T$, $\phi^{(k)}$ being the preferred stimulus of the kth neuron. The correlation matrix is defined as follows,

$$C^{(kl)} = [\delta_{kl} + (1 - \delta_{kl})q]\, \psi(\zeta^{(k)}) \psi(\zeta^{(l)}), \tag{8.2}$$

where $\zeta^{(p)} \equiv f(\boldsymbol{\theta} - \boldsymbol{\phi}^{(p)})$, and (to ensure that the correlation matrix $\mathbf{C}$ is positive definite) the cross-correlation strength parameter q is allowed to vary in the range $0 \leq q < 1$. The Fisher information for this gaussian-correlated model is given by the following expression (Abbott & Dayan, 1999):

$$J_{ij}(\boldsymbol{\theta}) = \frac{\partial \mathbf{f}^T}{\partial \theta_i} \mathbf{C}^{-1} \frac{\partial \mathbf{f}}{\partial \theta_j} + \frac{1}{2} \mathbf{Tr} \left[\frac{d\mathbf{C}}{d\theta_i} \mathbf{C}^{-1} \frac{d\mathbf{C}}{d\theta_j} \mathbf{C}^{-1} \right]. \tag{8.3}$$

By inserting the correlation matrix definition given by equation 8.2, into equation 8.3, and taking the continuous limit for $N \gg 1$ (Abbott & Dayan, 1999; Wilke & Eurich, 2001), one arrives at the following expression:

$$J_{ij}(\boldsymbol{\theta}) = \frac{N\nu^{D-2}}{(2\pi)^D(1-q)\sigma^4} \int_{-\frac{\pi}{\nu}}^{\frac{\pi}{\nu}} d^D\boldsymbol{\phi}\, f_0^2(\boldsymbol{\theta} - \boldsymbol{\phi}) \left(\frac{1}{\psi^2(\zeta)} + \frac{(2-q)\psi'^2(\zeta)}{\psi^2(\zeta)} \right) \times \sin(\nu(\theta_i - \phi_i)) \sin(\nu(\theta_j - \phi_j)). \tag{8.4}$$

Note that as for the uncorrelated model discussed in section 2, the only nonzero elements of the Fisher information matrix are the diagonal ones; these elements are all identical and do not depend on the value of angular stimulus variable. Thus, dropping index and $\boldsymbol{\theta}$-dependency notation, we will again simply denote by J the diagonal element of the Fisher information matrix in equation 8.4.

The expression of J for the correlated model in equation 8.4 is almost identical to the one for the uncorrelated gaussian model in equation 5.2, the only difference being a q-dependent relative weight of the two additive terms within parentheses in equation 8.4. Since, as explained in section 5, the first additive term in parentheses is the prominent one in shaping the σ- and D-dependence of J, the correlated model J behaves very much like the uncorrelated-gaussian-model J in equation 5.2. In particular, the cross-correlation parameter q affects the optimal σ values in a very marginal way. Thus, we expect that the only appreciable effect of the cross-correlation strength q is to shift slightly the position of the maximum for $D > 2$. The variations of the optimal σ values of the correlated model as a function of q (obtained integrating numerically equation 8.4) are reported in Figure 4 for $D = 3$ and $D = 4$. $\psi(\zeta)$ was again chosen according to equation 5.3 with $\alpha = 1$ and $\beta = 1$. It is apparent that the values of the optimal tuning widths found for the uncorrelated model are virtually unchanged throughout the entire allowed range of cross-correlation strength q.

9 Discussion

Determining how the encoding accuracy of a population depends on the tuning width of individual neurons is crucial for understanding the

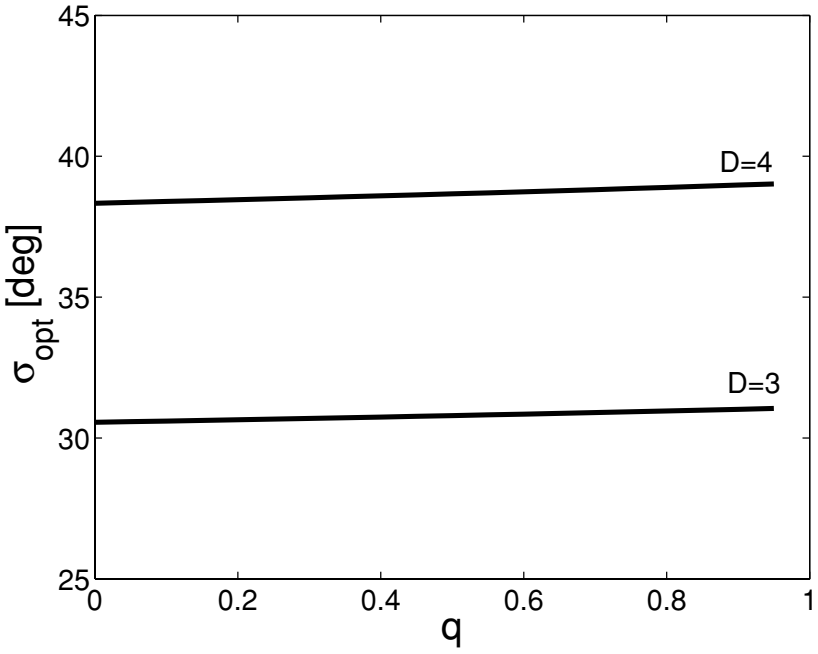

Figure 4: Optimal tuning widths for a population of orientation-selective neurons with gaussian firing statistics in the presence of uniform cross correlations, as a function of the cross-correlation strength parameter q. The two cases $D = 3$ (lower curve) and $D = 4$ (upper curve) are separately reported.

transformation of the sensory representation across different levels of the cortical hierarchy (Hinton, McClelland, & Rumelhart, 1986; Zohary, 1992).

Generalizing previous results obtained for nonperiodic stimuli (Zhang & Sejnowski, 1999), here we have determined how encoding accuracy of periodic stimuli depends on the tuning width. Although we found no universal scaling rule, the dependence of the encoding accuracy of periodic stimuli on the width of tuning was remarkably consistent across neural models and model parameters. This indicates that the key properties of encoding of periodic variables are general.

The encoding accuracy of angular variables differs significantly from that of nonperiodic stimuli. The two major differences are that (1) whatever the number of stimulus features D, very large tuning widths are inefficient for encoding a finite number of periodic variables, and (2) for $D > 2$, intermediate values of tuning widths (within the range observed in cortex) provide the population with the largest representational capacity. These differences suggest that population coding of periodic stimuli may be influenced by

computational constraints that differ from those influencing the coding of nonperiodic stimuli.

As for the nonperiodic case (Zhang & Sejnowski, 1999), the neural population information about periodic stimuli depends crucially on D, the number of stimulus features being encoded. This number is not known exactly; therefore, it is difficult to derive precisely the optimal tuning widths in each cortical area. However, some evidence indicates that neurons may encode a small number of stimulus features, and in many cases the number of encoded features is more than one. For example, visual neurons extract a few features out of a rich dynamic stimulus (Brenner, Bialek, & de Ruyter van Steveninck, 2000; Touryan, Lau, & Dan, 2002), and a small number of different stimulus maps are often found to coexist over the same area of cortical tissue (Swindale, 2004). Our results show that in this regime in which more than one (but no more than a few) periodic features are encoded, tuning widths within the range observed in cortex are efficient at transmitting information.

We showed that the optimal tuning width for population coding increases with the number of periodic features being encoded. Neurophysiological data in Table 1 show a progressive increase of tuning widths across the cortical hierarchy, consistent with the idea that higher visual areas encode complex objects described by a combination of several stimulus parameters (Pasupathy & Connor, 2001), thus requiring larger tuning widths for efficient coding.

In summary, our results demonstrate that tuning curves of intermediate widths offer computational advantages when considering the encoding of periodic stimuli.

Acknowledgments

We thank I. Samengo and R. Petersen for interesting discussions. This research was supported by Human Frontier Science Program, Royal Society and Wellcome Trust 066372/Z/01/Z.

References

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

Abbott, L. F., Rolls, E. T., & Tovée, M. J. (1996). Representational capacity of face coding in monkeys. *Cerebral Cortex, 6*, 498–505.

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol., 52*, 1106–1130.

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron, 26*, 695–702.

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

Cover, T., & Thomas, J. (1991). *Elements of information theory*. New York: Wiley.
Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.
Gershon, E. D., Wiener, M. C., Latham, P. E., & Richmond, B. J. (1998). Coding strategies in monkey V1 and inferior temporal cortices. *J. Neurophysiol., 79*, 1135–1144.
Henry, G. H., Dreher, B., & Bishop, P. O. (1974). Orientation specificity of cells in cat striate cortex. *Journal of Neurophysiology, 37*, 1394–1409.
Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing* (Vol. 1, pp. 77–109). Cambridge, MA: MIT Press.
Kohn, A., & Movshon, A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience, 7*, 764–772.
Logothetis, N., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology, 5*, 552–563.
McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci., 19*(1), 431–441.
Nirenberg, S., Carcieri, S. M., Jacobs, A., & Latham, P. E. (2001). Retinal ganglion cells act largely as independent encoders. *Nature, 411*, 698–701.
Oram, M. W., Hatsopoulos, N., Richmond, B., & Donoghue, J. (2001). Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures, *J. Neurophysiol., 86*, 1700–1716.
Pasupathy, A., & Connor, C. E. (2001). Shape representation in area v4: Position-specific tuning for boundary conformation. *J. Neurophysiol., 86*, 2505–2519.
Petersen, R. S., Panzeri, S., & Diamond, M. (2001). Population coding of stimulus location in rat somatosensory cortex. *Neuron, 32*, 503–514.
Pouget, A., Deneve, S., Ducom, J.-C., & Latham, P. (1999). Narrow versus wide tuning curves: What's best for a population code? *Neural Computation, 11*, 85–90.
Pouget, A., Zhang, K., Deneve, S., & Latham, P. E. (1998). Statistically efficient estimation using population coding. *Neural Comp., 10*, 373–401.
Samengo, I., & Treves, A. (2001). Representational capacity of a set of independent neurons. *Physical Review E 63*, 0119101–01191014.
Sompolinsky, H., Yoon, H., Kang, K., & Shamir, M. (2001). Population coding in neuronal systems with correlated noise. *Phys. Rev. E, 64*, 051904.
Swindale, N. V. (1998). Orientation tuning curves: Empirical description and estimation parameters. *Biological Cybernetics, 78*, 45–56.
Swindale, N. V. (2004). How different feature spaces may be represented in cortical maps. *Network, 15*, 217–242.
Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Experimental Brain Research, 41*, 414–419.
Touryan, J., Lau, B., & Dan, Y. (2002). Isolation of relevant visual features from random stimuli for cortical complex cells. *J. Neurosci., 22*, 10811–10818.
Usrey, W. M., Sceniak, M. P., & Chapman, B. (2003). Receptive fields and response properties of neurons in layer 4 of ferret visual cortex. *J. Neurophysiol., 89*, 1003–1015.

Wilke, S. D., & Eurich, C. W. (2001). Representational accuracy of stochastic neural populations. *Neural Comp., 14*, 155–189.

Zhang, K., & Sejnowski, T. (1999). Neuronal tuning: To sharpen or to 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.

Received February 7, 2005; accepted November 3, 2005.