Transmission of population coded information

Alfonso Renart and Mark C. W. van Rossum

September 10, 2009

Supplementary Information

Contents

1 Analytical Procedures 2
  1.1 Input and Output Fisher Information ......................... 3
  1.2 Optimization of weight profile ................................ 5

2 Numerical Procedures 8
1 Analytical Procedures

We studied the loss of accuracy in a two layered population code (main text Fig. 1A). The input layer encodes a one-dimensional circular quantity (e.g., orientation) $\theta$, taking values between $-90^\circ$ and $+90^\circ$. The activity of neuron $i = 1, \ldots, N$ in the input layer is given by its firing rate $r_{i}^\text{in}$. Its average across trials is $\langle r_{i}^\text{in} \rangle = f(\phi_i - \theta)$, where $\phi_i = [(2i-1)/N-1]90^\circ$ is the preferred angle of the neuron. The neuron’s tuning curve $f$ is a symmetric, bell-shaped curve, in particular we use

$$f(\phi_i - \theta) = r_p \exp \left[ \frac{\cos(2(\phi_i - \theta)) - 1}{a^2} \right]$$

where $r_p$ denotes the peak rate and $a$ the width of the tuning curve. The activity of the input layer population is written as the vector $\mathbf{r}_{\text{in}}$ with elements $r_{i}^\text{in}$. Correlated trial-to-trial variability in the population is described by an $N$-dimensional Gaussian distribution, so that the probability density of $\mathbf{r}_{\text{in}}$ given the stimulus $\theta$ is

$$\rho(\mathbf{r}_{\text{in}} | \theta) = \frac{1}{Z} \exp \left[ -\frac{1}{2} (\mathbf{r}_{\text{in}} - \mathbf{f}(\theta))^T C_0^{-1} (\mathbf{r}_{\text{in}} - \mathbf{f}(\theta)) \right]$$

where $Z$ is a normalization constant, $\mathbf{f} = \langle \mathbf{r}_{\text{in}} \rangle$ is a vector with elements $f(\phi_i - \theta)$, and $C_0(\phi_i, \phi_j) = \langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle$ is the covariance matrix of the noise, where the angular brackets represent averages over the probability density $\rho(\mathbf{r}_{\text{in}} | \theta)$.

The correlation between neurons with preferred angles $\phi_i$ and $\phi_j$ is assumed to be independent of $\theta$ and a function of $|\phi_i - \phi_j|$ only. The results below hold for any such correlation structure, but for the figures we use the following parametrization for the noise covariance matrix

$$C_0(|\phi_i - \phi_j|) = (\sigma_0^2 - c_0)\delta_{ij} + c_0 \exp \left[ -\frac{2|\phi_i - \phi_j|}{\rho_0} \right] = [C_0]_{ij}$$

where $\sigma_0^2$ is the trial-to-trial variance in the firing rate, $c_0$ the peak covariance between neurons with very close preferred orientations, and $\rho_0$ the spatial extend of the correlations.
The neurons in the input layer project to the output layer. The connection strength between neuron \(i\) in the input layer and neuron \(j\) in the output layer is given by \(W(\phi_i, \phi_j)/N\). The scaling factor \(1/N\) guarantees that the mean current from the input layer will be independent of the network size \(N\) for large networks. We assume that the weight profile obeys \(W(\phi_i, \phi_j) = W(|\phi_i - \phi_j|)\), which guarantees that the activity profile in the input and output layers is aligned (Salinas and Abbott, 1994). The firing rate of neuron \(j\) in the output layer on a given trial is given by

\[
r_j^{\text{out}} = g[I_j] = g\left[\frac{1}{N} \sum_{i=1}^{N} W(|\phi_i - \phi_j|) r_i^{\text{in}} + \eta_j\right]
\]

(2)

where \(g\) is a invertible non-linearity that relates the current \(I_j\) into a neuron to its firing rate, incorporating synaptic integration and the spiking mechanism. The term \(\eta_j\) represents stimulus-independent noise to neuron \(j\). We assume these inputs to be zero on average, i.e., \(\langle \eta \rangle = 0\). The stimulus-independent noise in the output layer is Gaussian distributed and characterized by a covariance matrix \(C_1(|\phi_i - \phi_j|)\), with parameters \(\sigma_1^2\), \(c_1\) and \(\rho_1\).

1.1 Input and Output Fisher Information

To quantify the accuracy of the population code we use the Fisher Information (Cover and Thomas, 1991; Kay, 1993). The Fisher Information (FI) gives a lower bound on the variance of the best possible unbiased estimator of \(\theta\), and has been extensively used in the context of population coding (Seung and Sompolinsky, 1993; Abbott and Dayan, 1999; Zhang and Sejnowski, 1999; Deneve, Latham, and Pouget, 1999; Sompolinsky et al., 2002; Wu, Amari, and Nakahara, 2002). Let \(\hat{\theta}(r)\) be an unbiased estimator of the stimulus \(\theta\) given the activity \(r\) of a neural population in a given trial. According to the Cramer-Rao theorem (Cover and Thomas, 1991; Kay, 1993), the variance of \(\hat{\theta}(r)\) cannot be smaller than the inverse of the FI. That is, \(\langle \delta \hat{\theta}^2 \rangle \geq 1/J\), where the FI is given by
\[ J = -\left( \frac{\partial^2}{\partial \theta^2} \log \rho(r|\theta) \right) \]. Because the population is dense and homogeneous, the FI does not depend on \( \theta \). When \( \rho(r|\theta) \) is Gaussian and the covariance \( C_0 \) does not depend on \( \theta \), the FI is given by the simple expression

\[ J = f'(\theta)C_0^{-1}f'(\theta) \]  

(3)

As the correlation between a pair of neurons depends only on the difference between their preferred directions, \( C_0 \) is diagonal in Fourier space. Thus, it is natural to work in the Fourier domain. We use the following convention for the discrete Fourier transform of an array \( f_j \) and its inverse

\[ \tilde{f}_n = \frac{1}{N} \sum_{j=1}^{N} e^{-2\pi i j n/N} f_j \]

\[ f_j = \sum_{n=0}^{N-1} e^{2\pi i j n/N} \tilde{f}_n \]

so that \( \sum_n |\tilde{f}_n|^2 = \sum_j |f_j|^2 / N \). With this convention, if a matrix \( A_{ij} = A(\phi_i - \phi_j) \) has Fourier transform \( \tilde{A}_n \), then the Fourier transform of its inverse \( B_{ij} \) (defined such that \( \sum_k A_{ik}B_{kj} = \delta_{ij} \)) is given by \( \tilde{B}_n = 1/(N^2 \tilde{A}_n) \). We will in general write the spatial frequency modes of different quantities as going from \( n = 0, \ldots, N - 1 \). However, in the figures of the main text we displayed the highest spatial frequencies as negative, to highlight the symmetry of the Fourier transforms, e.g., \( n = N - m \rightarrow n = -m \).

Each spatial frequency contributes independently to the FI in Equation 3 (Sompolinsky et al., 2002; Wu, Amari, and Nakahara, 2002). This follows from the fact that for any matrix for which \( X_{ij} = X_{i-j} \)

\[
\sum_{jj'=1}^{N} a_j X_{jj'} a_{j'} = \sum_{jkl'}^{N} a_j \left[ \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i (j-k)n/N} \right] \left[ \frac{1}{N} \sum_{m=0}^{N-1} e^{2\pi i (l'-j')m/N} \right] X_{kl} a_{j'} = \sum_{m,n=0}^{N-1} \tilde{a}_n^* \tilde{a}_m \left[ \sum_{l=1}^{N} e^{2\pi i (m-n)l} \right] \left[ \sum_{k=1}^{N} e^{-2\pi i (k-l)l} \right] X_{k-l} = N^2 \sum_{n=0}^{N-1} |\tilde{a}_n|^2 \tilde{X}_n
\]
which, recalling that $\tilde{X}_n\tilde{X}_n^{-1} = 1/N^2$, implies that the FI in the input layer is given by

$$J_{\text{in}} = \sum_{n=0}^{N-1} \frac{|\tilde{f}'(n)|^2}{C_0(n)}$$

(Equation 1 in the main text).

Although the distribution of activity on a given trial in the output layer is not Gaussian due to the non-linear transformation $g(\ldots)$ between currents and rates, as long as this transformation is invertible, it can be shown that no information is lost in the mapping from currents to rates. Thus the FI in the output activity is the same as that in the distribution of afferent currents to the neurons in the output layer. The current $I_j$ to neuron $j$ in the output layer is a weighted sum of Gaussian random variables (Equation 2) and thus is Gaussian distributed itself as well, with mean

$$\langle I(\phi_j) \rangle = \frac{1}{N} \sum_{i=1}^{N} W(|\phi_j - \phi_i|) f(\phi_i - \theta)$$

and covariance

$$C_1(|\phi_i - \phi_j|) = \frac{1}{N^2} \sum_{kl=1}^{N} W(|\phi_i - \phi_k|)C_0(|\phi_k - \phi_l|)W(|\phi_l - \phi_j|) + C_1(|\phi_i - \phi_j|) \equiv [C_1]_{ij}$$

Using the definition of the Fourier transform, if $h_i = \sum_{j=1}^{N} f_{i-j}g_j$, then $\tilde{h}_n = N\tilde{f}_n\tilde{g}_n$. Thus, the FI in the output layer $J_{\text{out}} \equiv \sum_{ij=1}^{N} \partial_\theta \langle I(\phi_i) \rangle [C_1^{-1}]_{ij} \partial_\theta \langle I(\phi_j) \rangle$ is given in the Fourier domain by

$$J_{\text{out}} = \sum_{n=0}^{N-1} \frac{|\tilde{f}'(n)|^2|\tilde{W}(n)|^2}{C_0(n)|\tilde{W}(n)|^2 + C_1(n)}$$

### 1.2 Optimization of weight profile

Let us denote by $J_{\text{in}}(n)$ the FI available in the $n^{th}$ Fourier mode of the activity of the input layer, so that $J_{\text{in}} = \sum_n J_{\text{in}}(n)$. In our network, the activity in each Fourier mode is transmitted independently to the output layer,

$$J_{\text{out}}(n) = J_{\text{in}}(n) \left[ \frac{|\tilde{W}(n)|^2}{|\tilde{W}(n)|^2 + C_1(n)/C_0(n)} \right]$$
Each factor in square brackets has a value between zero and one, and measures how much information is lost on that frequency mode between the input and output layers. As explained in the main text, unless $\tilde{C}_1(n) = 0$, the output layer information is always less than the input layer information. The amount of information loss could be minimized if $|\tilde{W}(n)| \gg \tilde{C}_1(n)/\tilde{C}_0(n)$ for all $n$, e.g., by having $W_{ij} = A \delta_{ij}$ with $A$ sufficiently large. However, this trivial solution is unphysiological. Clearly, there has to be some constraint in the magnitude of the weights. Since the scaling of the weights has already been factored out, it is reasonable to constrain the magnitude of the weights by imposing that the spatial average of $|W(\phi_i - \phi_j)|^2$ be of order unity, with the square preventing the constraint to be satisfied through a cancellation between positive (excitatory) and negative (inhibitory) weights. With our convention for the Fourier transform this is equivalent to the constraint

$$\sum_{n=0}^{N-1} |\tilde{W}(n)|^2 = q$$

with $q$ of order unity (in the figures in the main text we used $q = 2$).

The problem is thus to maximize $J_{out}$ subject to the constraint $\sum_n w_n^2 = q$. The constraint induces a ‘competition’ between the different Fourier modes of the weights, which is resolved by a solution in which power is only assigned to the optimal weights in the frequencies that are more informative. To show this, let us rewrite $J_{out}$ as

$$J_{out} = \sum_{n=0}^{N-1} \frac{j_n w_n^2}{w_n^2 + t_n}$$

with $j_n \equiv J_{in}(n)$, $w_n \equiv |\tilde{W}(n)|$ and $t_n \equiv \tilde{C}_1(n)/\tilde{C}_0(n)$. The $t_n$ represent an effective threshold on the $n^{th}$ Fourier mode of the weights, so that if $w_n^2 \gg t_n$, most information in that frequency mode is transmitted, while if $w_n^2 \ll t_n$, most information in that mode is lost.

We perform the optimization using Lagrange-multipliers. The Lagrangian is

$$L = \sum_n \frac{j_n w_n^2}{w_n^2 + t_n} - \lambda \left( \sum_n w_n^2 - q \right)$$
where we introduced a Lagrange multiplier $\lambda$ for the $q$ constraint. To find the maximum output information, we take the gradients with respect to $\{w_n\}$ and $\lambda$, and set them to zero

$$\frac{\partial L}{\partial w_n} = \frac{j_n t_n}{(w_n^2 + t_n)^2} - \lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = q - \sum_n w_n^2 = 0$$

The first equation yields

$$w_n^2 = \sqrt{t_n} \left[ \frac{\sqrt{j_n/\lambda} - \sqrt{t_n}}{\sum_{n'} \sqrt{j_{n'}/\lambda} - \sqrt{t_{n'}}} \right]$$

(4)

where the parameter $\lambda$ has to be chosen so that the second equation $\sum_n w_n^2 = q$, implementing the constraint, is satisfied. The $w_n^2$ for which Equation 4 is negative, are set to zero. Due to this, these equations have to be solved self-consistently. Given an initial arbitrary value of $\lambda$, we calculate $w_n^2$. Let’s call $\{n'\}$ those $\{n\}$ for which $w_{n'}^2$ is positive. From the constraint $\sum w_n^2 = q$, we obtain a new value for $\lambda$

$$\frac{1}{\sqrt{\lambda}} = \frac{q + \sum_{n'} t_{n'}}{\sum_{n'} \sqrt{j_{n'}/\lambda}}$$

This new value is used to calculate the updated $w_n^2$, and this procedure is repeated until $\lambda$ no longer changes at which point $\sum_{n'} w_{n'}^2 = q$. Substituting this final value for $\lambda$ into Equation 4 results in

$$w_n^2 = \sqrt{t_n} \left[ \left( \frac{\sum_{n'} (w_{n'}^2 + t_{n'})}{\sum_{n'} \sqrt{j_{n'}/\lambda}} \right) \sqrt{j_n} - \sqrt{t_n} \right]_+$$

where $[x]_+ = \max(0, x)$. The final weight matrix is retrieved by transforming back into real space using

$$W(\phi_i - \phi_j) = \sum_{n=0}^{N-1} e^{i n (\phi_i - \phi_j)} w_n$$

Note that, formally, our solution only specifies the squared magnitude of the weights, i.e., their ‘power spectrum’. Thus, any solution obtained by multiplying the $w_n$ by any phase
$e^{i\beta_n}$ is also a solution. We consider the most parsimonious solution for which $\beta_n = 0$ for all $n$.

As an example consider the simple case where neurons in the input layer are independent and the external noise is also uncorrelated, i.e., $c_{0,1} = 0$, as in Figure 2A of the main text. In this case $t_n = (\sigma_1/\sigma_0)^2 \equiv t$ and $j_n = N|\tilde{f}(n)|^2/\sigma_0^2$, so that the optimal weights become

$$w_n^2 = \left[\alpha|\tilde{f}(n)| - t\right]_+$$

with $\alpha = (q + N\prime t)/\sum_{n'}|\tilde{f}(n')|$, where $N'$ is the number of non-zero weights. In this case the squared Fourier modes of the optimal weights are a thresholded version of the absolute value the Fourier transform of the derivative of the tuning curve. The threshold comes from the constraint on the power in the weights and is reminiscent of the ‘water-filling’ analogue used to maximize Shannon information in noisy parallel channels (Cover and Thomas, 1991). In both cases informative modes are enhanced, whereas modes without enough information are not transmitted. Note however, that in contrast to ‘water-filling’ the optimal weights in Equation 1.2 also depend on $t_n$, which for non-white noise depends on $n$.

### 2 Numerical Procedures

For the simulation study we use a layered network consisting of leaky integrate-and-fire neurons with 100 M$\Omega$ input resistance, 20 ms time constant, -60 mV resting potential, and -50 mV firing threshold. After a spike the membrane potential resets to the resting potential at which it remains clamped for 1 ms (absolute refractory period). The network has 128 neurons per layer and periodic boundary conditions are imposed to prevent edge effects. The layers are connected in a feed-forward fashion, with stochastic synapses that have a release probability of 50%.
As we want to be able to exactly balance excitatory against inhibitory input current, we use a current-based simulation of the synapses (Vogels and Abbott, 2005). Thus the synapses are modeled as current sources which are positive (negative) for excitatory AMPA (inhibitory GABA) synapses. Both synapses have a 5ms time-constant. For simplicity neurons can make both excitatory and inhibitory synapses. Conductance-based simulations lead to very similar results, but do not allow us to make as precise claims about balance. All neurons are injected with a Gaussian-distributed, low-pass filtered noisy background current with positive mean (mean 55 pA, standard deviation 70 pA, filtering time constant 2 ms). This noise current mimics the intrinsic noise in the neurons and unspecific input of other neurons. The noise prevents synchronization of the neural activity ensures that firing statistics, response latencies, and membrane potential distributions mimic those found in vivo (Knight, 1972; Gerstner, 2000; van Rossum, Turrigiano, and Nelson, 2002).

The input layer is stimulated by a static profile of current injections with a Gaussian shape according to Equation 1. The activity in the input layer has a width of 60° degrees (full width at half maximum) and a peak response of 50Hz. Input to the subsequent layer is provided by excitatory and inhibitory synapses via a weight matrix. A full optimization of the weight matrix is not feasible in the simulations. Instead, the strength of the connections is parametrized by a rectified cosine with a uniform inhibitory surround. The connection strength between a neuron at a position \(i\), and a neuron at position \(j\) in the next layer, equals \(W_{ij} = A \cos(\frac{\phi_{ij}}{\phi_w} 90^\circ)\) whenever \(\phi_{ij} \leq \phi_w\), while \(W_{ij} = -c \sum_i [W_{ij}]^+\) otherwise. The angle \(\phi_w\) defines the width of the excitatory center. The strength surround inhibition is given by \(c\). For a purely excitatory profile \(c = 0\), while for a balanced profile \(c\) is adjusted such that \(\sum_i W_{ij} = 0\). Other weight profiles such as a difference of Gaussians gives very similar results; the parametrization used allows for a straightforward adjustment of width, inhibition and gain. For every choice of connectivity profile the
amplitude $A$ is calibrated so that the peak firing rate in the input and output layer of the network are the same. This is important as the gain influences the transmission quality, because the more output spikes are produced, the more accurate the estimate can be.

The quality of the transmission is calculated as follows: A long lasting stimulus is applied and after removal of the onset transient, the population activity of a layer of interest is sliced into time bins of 25 ms. Based on the spikes in these time bins the stimulus orientation is estimated. In trials where there are no spikes in any bin, a random position estimate is assigned. We estimated the position of the stimulus using an optimal linear discriminator, which provides a lower bound on the Fisher Information. On half the trails the stimulus had an angle $+\Delta \phi$ (7°), while on the other trial the angle was $-\Delta \phi$. A linear discriminator has to distinguish these two stimuli by calculating $o_{\pm \Delta \phi} = v \cdot r(\pm \Delta \phi)$, where $v$ are the weights of the discriminator. Because of symmetry, the discriminator weights can be written as a discrete Fourier Series $v_i = \sum_{k=1}^{n} a_k \sin(2\pi ik/N)$. The discriminator was optimized by numerically minimizing the SNR with respect to the Fourier coefficients $a_k$ (8 Fourier coefficients proved sufficient). This optimization was done independently for every setting of the connectivity. The signal to noise ratio of the discriminator is used to calculate the standard deviation in the angle estimate, which equals the inverse square root of $J$ (Seriès, Latham, and Pouget, 2004; Shamir and Sompolinsky, 2004).

References


