THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Transmission of population-coded information

Citation for published version:
Renart, A & van Rossum, MCW 2012, 'Transmission of population-coded information', *Neural Computation*, vol. 24, no. 2, pp. 391-407. https://doi.org/10.1162/NECO_a_00227

Digital Object Identifier (DOI):
10.1162/NECO_a_00227

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Neural Computation

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 26. May. 2020

LETTER Communicated by Si Wu

Transmission of Population-Coded Information

Alfonso Renart
arenart@neuro.fchampalimaud.org
Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102, U.S.A., and Champalimaud Centre for the Unknown, Doca de Pedrouços 1400-038 Lisboa, Portugal

Mark C. W. van Rossum
mvanross@inf.ed.ac.uk
Institute of Adaptive and Neural Computation, School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, U.K.

As neural activity is transmitted through the nervous system, neuronal noise degrades the encoded information and limits performance. It is therefore important to know how information loss can be prevented. We study this question in the context of neural population codes. Using Fisher information, we show how information loss in a layered network depends on the connectivity between the layers. We introduce an algorithm, reminiscent of the water filling algorithm for Shannon information that minimizes the loss. The optimal connection profile has a center-surround structure with a spatial extent closely matching the neurons' tuning curves. In addition, we show how the optimal connectivity depends on the correlation structure of the trial-to-trial variability in the neuronal responses. Our results explain how optimal communication of population codes requires the center-surround architectures found in the nervous system and provide explicit predictions on the connectivity parameters.

1 Introduction

Neural information often passes through many different brain areas. Consider, for instance, a task where a subject has to reach for a visual target. Information about the target location will have to be transmitted across many stages: retina, lateral geniculate nucleus, visual, premotor, and motor cortices. Due to a variety of noise sources, single neurons are highly variable and unreliable at every stage (Dean, 1981; Shadlen & Newsome, 1998; Faisal, Selen, & Wolpert, 2008). It is believed that robustness against single-neuron variability is engendered by population codes, in which many neurons are activated by a given stimulus. Population codes have been studied extensively. In particular, the role of the tuning curve shape and the neural

 © 2011 Massachusetts Institute of Technology

noise model (Seung & Sompolinsky, 1993; Zhang & Sejnowski, 1999; Bethge, Rotermund, & Pawelzik, 2002), as well as the effect of correlations in the trial-to-trial variability of different neurons, have been studied in detail (Abbott & Dayan, 1999; Sompolinsky, Yoon, Kang, & Shamir, 2002; Wu, Amari, & Nakahara, 2002; Shamir & Sompolinsky, 2004). Moreover, several algorithms have been proposed to read out population codes (Lee, Rohrer, & Sparks, 1988; Oram, Foldiak, Perrett, & Sengpiel, 1998), including neurally plausible networks (Deneve, Latham, & Pouget, 1999).

Yet although it is well known how much information is available in a given population and how it can be read out, it is not known how information is preserved when it is transmitted between processing stages. Only a few studies have addressed the transmission of population codes (Hamaguchi, Okada, Yamana, & Aihara, 2005; Hamaguchi, Okada, & Aihara, 2007), and none have quantified transmission quality. The transmission quality is of functional importance: in the reaching task, each processing stage presumably uses a population code. Suboptimal transmission would lead to decreased performance and, hence, errors in the perceived target location or increased reaction times.

In this study, we use Fisher information (FI) to quantify how information about a stimulus feature degrades as activity is transmitted between layers. Next, we use these results to find the synaptic weight profile that minimizes information loss, and we find it to be center-surround (Mexican hat), with a width approximately matched to the neurons' tuning width. The generality of our results is confirmed using numerical simulations of networks of integrate-and-fire neurons.

2 Population Coding in Layered Networks

As a concrete example of information propagation across a hierarchy of processing areas, we analyze the transmission of population-coded information between two layers. We consider the network shown in Figure 1A. Each of the two layers consists of a homogeneous population of N neurons encoding an angular variable labeled θ, which could represent, for instance, orientation. The response of each neuron i in the input layer depends on the difference between its preferred orientation ϕ_i and θ, resulting in a population activity bump on each trial (see Figure 1A, top). Neurons in the output layer receive activity of the input layer through a synaptic weight matrix. The neurons in the output layer are noisy as well.

The purpose of this study is to determine which synaptic weight matrix minimizes information loss. Should the activity in the output layer be narrower (sharpening) or wider than the input? Narrow tuning curves in principle contribute to accuracy (Zhang & Sejnowski, 1999), but sharpening also induces correlations (Seriès, Latham, & Pouget, 2004) and leads to lower overall activity. Wide tuning curves, however, smear the input, reducing selectivity but also increase overall activity. In the presence of noise

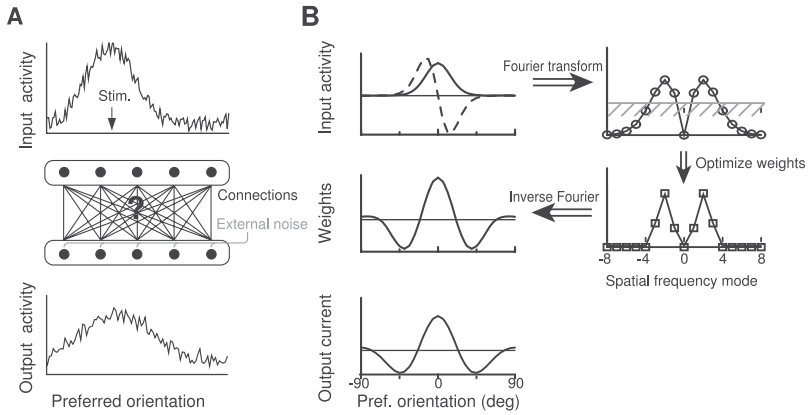

Figure 1: The network architecture and a schematic of the algorithm to find the weights that optimally preserve information. (A) Architecture of the two-layer network. The stimulus (vertical arrow) is encoded in the population activity of the input layer (top). Due to trial-to-trial variability, the activity of the population on any given trial represents the stimulus with only limited accuracy. Neurons in the output layer (bottom) pool the activity of the input layer through a set of synaptic weights (middle, black), and also receive noise (middle, gray). The algorithm optimizes the synaptic weights to maximize the coding accuracy in the output layer. (B) Schematic of the weight optimization procedure. (Top left) From the population activity (solid line), the spatial derivative is calculated (dashed line) and Fourier-transformed. In the Fourier domain, only the strong Fourier modes that exceed a certain threshold are selected and are transmitted by the weights (right). In the illustrated case, the threshold is constant; in general, it is spatial frequency dependent. Transformed back into the spatial domain, the optimal weight profile typically resembles a center-surround structure. The bottom panel shows the net current received by neurons in the output layer, which is the convolution of the input activity with the weights.

in the output layer, it is a priori not obvious what the optimal resolution of this trade-off will be.

For mathematical tractability, we start analyzing this problem with a rate description of the neural activity, but then we extend our results to simulations with spiking neurons. We denote the firing rate of neuron i in the input layer as r_i^{in}. Its average across trials is given by $f(\theta - \phi_i)$, where f specifies the tuning curve shape, which is the same for all neurons. Our theory is valid for any tuning curve shape; for the figures, we use

$$f(\phi_i - \theta) = r_p \exp\left[\frac{\cos 2(\phi_i - \theta) - 1}{a^2}\right] \equiv f_i(\theta), \tag{2.1}$$

where $\phi_i = (-1 + \frac{2i-1}{N})90^\circ$, $i = 1, \ldots, N$ is the preferred orientation of cell i, and r_p and a are the peak firing rate of the cell and tuning curve width, respectively.

Trial-to-trial response variability in the input neurons in principle can be correlated (the so-called noise correlation), which affects the information content. Under the assumption of gaussian noise, the noise correlation is fully quantified by a covariance matrix C_0. The structure of the noise correlation can be dependent on the similarity between the tuning of different cells, that is, $[C_0]_{ij} = C_0(\phi_i - \phi_j)$, but is assumed independent of the stimulus. Our theory allows for any functional form of C_0 that meets this condition. For the figures, we use correlations that decay exponentially with the difference in preferred angle,

$$[C_0]_{ij} = C_0(\phi_i - \phi_j) = (\sigma_0^2 - c_0)\delta_{ij} + c_0 \exp\left[-\frac{2|\phi_i - \phi_j|}{\rho_0}\right], \tag{2.2}$$

where σ_0^2, c_0, and ρ_0 are the firing rate variance of the neurons, peak correlation, and correlation length, respectively.

The firing rate of neuron j in the output layer on a given trial is

$$\begin{aligned} r_j^{\text{out}} &= g\left(I_j\right) \\ &= g\left(\frac{1}{N}\sum_{i=1}^{N} W_{ji} r_i^{\text{in}} + \eta_j\right), \end{aligned} \tag{2.3}$$

where I_j is the current to neuron j and $g(\ldots)$ represents an instantaneous nonlinearity that maps the input current of a neuron into its firing rate. The magnitude of the synaptic weight between an input and an output neuron is assumed to depend on the difference between their preferred angles only, $W_{ij} = W(|\phi_i - \phi_j|)$, so that activity in the input and output layers is aligned (Salinas & Abbott, 1994). The term η_j represents external noise in the current, such as coming from other inputs, synapses, or voltage-gate channels. It is assumed to be unrelated to the stimulus and the activity of the previous layer. Again allowing for correlations in noise, it is modeled as zero-mean gaussian noise current with covariance matrix C_1, which in the figures is parameterized as in equation 2.2, by σ_1^2, c_1, and ρ_1.

2.1 Information in the Input Layer. We use the Fisher information to quantify how accurately the stimulus is represented in both layers (Cover & Thomas, 1991; Kay, 1993). Under the above noise and homogeneity assumptions, the Fisher information in the input layer is

$$J_{\text{in}} = \sum_{i,j=1}^{N} f_i'(\theta)[C_0^{-1}]_{ij} f_j'(\theta), \tag{2.4}$$

where $f_i'(\theta)$ is the derivative of the tuning curve of neuron i with respect to the stimulus angle θ. But the FI is more elegantly formulated in spatial frequency space (Sompolinsky et al., 2002; Wu et al., 2002). We use the following convention for the Fourier transform: $\widetilde{h}_n = (1/N)\sum_{j=1}^{N} e^{-2\pi ijn/N} h_j$ and the inverse transform $h_j = \sum_{n=0}^{N-1} e^{2\pi ijn/N}\widetilde{h}_n$. The Fisher information becomes

$$J_{\text{in}} = \sum_{n=0}^{N-1} \frac{|\widetilde{f}'(n)|^2}{\widetilde{C}_0(n)}, \tag{2.5}$$

where $\widetilde{C}_0(n)$ is the nth Fourier component of the noise correlation matrix and $\widetilde{f}'(n)$ is the Fourier components of the derivative of the tuning curve (Sompolinsky et al., 2002). Defining the signal-to-noise ratios at spatial frequency n as $J(n) = |\widetilde{f}'(n)|^2/\widetilde{C}_0(n)$,

$$J_{\text{in}} = \sum_{n=0}^{N-1} J(n). \tag{2.6}$$

The FI is thus a sum of the signal-to-noise ratios at every spatial frequency. For a given spatial frequency, the signal is the squared magnitude of the Fourier mode of the derivative of the tuning curve, while the noise is the corresponding Fourier mode of the covariance matrix. Because typical neuronal tuning curves are smooth and broad, only low spatial frequencies carry power and, hence, information. Note, however, that the zero mode of the signal carries no information. The zero mode corresponds to the change in average network activity with respect to the stimulus, as $\widetilde{f}_0' = (1/N)\sum_{j=1}^{N} f_j'$. This quantity is effectively 0 for networks with broad tuning curves and many neurons ($N \geq 100$).

2.2 Information in the Output Layer. Next, we calculate the information in the output layer. Because the currents are the linear sum of gaussian distributed firing rates and gaussian noise sources (see equation 2.3), the currents have again a gaussian distribution with mean

$$I_i(\theta) = \frac{1}{N}\sum_j W_{ij} f_j(\theta)$$

and covariance matrix

$$[C_I]_{ij} = \frac{1}{N^2}\sum_{kl} W_{ik}[C_0]_{kl} W_{lj} + [C_1]_{ij}.$$

The frequency components are transmitted independently, and the FI in the currents is calculated as above:

$$J_{\text{out}} = \sum_{n=0}^{N-1} J(n) \left[\frac{|\widetilde{W}(n)|^2}{|\widetilde{W}(n)|^2 + T(n)} \right], \tag{2.7}$$

where $\widetilde{W}(n)$ and $\widetilde{C_1}(n)$ are the Fourier components of the weight profile and of the covariance matrix of the external noise, and $T(n)$ is defined as $T(n) = \widetilde{C_1}(n)/\widetilde{C_0}(n)$.

This expression has a straightforward interpretation. Compared to the FI in the input layer, equation 2.6, the output information at every frequency is attenuated by a factor between 0 and 1 (the term in the square brackets). The output information is therefore always less than or equal to the input information. If there is no external noise in the output layer, $C_1 = 0$, then $T(n) = 0$ and the attenuation factors are exactly 1. In that case, no information is lost between the two layers, and save for the completely flat profile, any profile will suffice to transmit the population code perfectly.

From the currents, the firing rate follows from the nonlinearity $g()$ (see equation 2.3). In general, the output firing rate distribution will not be gaussian, but as long as the nonlinearity is invertible (e.g., a sigmoid or a power law with positive argument), the FI in the output rates equals the FI in the current.

2.3 Optimal Connectivity. Using the above results, we optimize the connectivity between the input and output layers so that it minimizes the information loss between the input and the output layers. The loss is minimal if the attenuation factor in equation 2.7 approaches 1 for all frequencies. This can be trivially achieved if the weights strongly amplify all spatial frequencies, that is, if $|\widetilde{W}(n)|^2$ is sufficiently large. This is, for instance, the case if the weight matrix is a large multiple of the identity matrix. In that case, the noise in the output layer becomes negligible compared to the input, and the FI in the output layer approaches the FI of the input layer. However, this solution yields unphysiologically large currents and high firing rates in the output layer. Clearly a constraint in the magnitude of the weights has to be imposed.

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, that is, $\frac{1}{N}\sum_i |W(\phi_i)|^2 = q$, with q a constant of order unity. The square prevents the constraint from being satisfied through a cancellation between positive (excitatory) and negative (inhibitory) weights. Other types of constraints gave comparable results (see simulations below). The constrained optimization effectively introduces competition between the different frequency modes of the weights. Rather than amplifying all modes, the optimal weights transmit only the most informative spatial frequencies. Assigning power in the

weakly informative spatial frequencies would not significantly increase the transmitted information, but, due to the constraint, it would take power away from the informative modes.

In spatial frequency space, the constraint is

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

Full details of the resulting multidimensional constrained optimization process, schematically illustrated in Figure 1B, can be found in the appendix. There, it is shown that the optimal weights are given by

$$|\widetilde{W}(n)|^2 = \sqrt{T(n)}\left[\sqrt{J(n)/\lambda} - \sqrt{T(n)}\right]_+ \tag{2.8}$$

where $[x]_+ = \max(0,x)$, and the constant λ is given by

$$\frac{1}{\sqrt{\lambda}} = \frac{\sum_{n'} |\widetilde{W}(n')|^2 + T(n')}{\sum_{n'} \sqrt{J(n')T(n')}}, \tag{2.9}$$

and the sum over n' runs over all frequency modes where the weights are non-0. Equations 2.8 and 2.9 have to be solved self-consistently due to the thresholding nonlinearity in equation 2.8, which appears because power can only be positive. In equation 2.8, the ratio of noises, $T(n)$, acts as an threshold on the weights.

Note that because both synaptic weights and correlation functions are translationally invariant, the optimal solution for the synaptic weights is very general. In particular, because our calculation is nonparametric, equation 2.8 describes the optimal connectivity profile for any tuning curve shape and any correlation profile.

2.4 Optimal Connectivity Profiles. We examine the optimal weight profiles that follow from the above calculation for a few illustrative cases. Consider first the simple case in which both noises are each uncorrelated, that is, $[C_0]_{ij} = \sigma_0^2\delta_{ij}$ and $[C_1]_{ij} = \sigma_1^2\delta_{ij}$, which yields $T(n) = T = \sigma_1^2/\sigma_0^2$. The optimal weights are $|\widetilde{W}(n)|^2 = [\alpha|\widetilde{f}'(n)| - T]_+$, where α is a constant given by $\alpha = \frac{qT\sum'_n 1}{\sum'_n |\widetilde{f}'(n')|}$. In other words, the weights in spatial frequency space are a thresholded version of the tuning curve derivative. Optimal weight profiles in this case are shown in Figure 2A for three different widths of the input tuning curve: When the input tuning curves are very narrow, the intermediate frequencies are most informative, giving the optimal weight profile an oscillatory character. For medium and wide tuning curves, the low-frequency modes are most informative, yielding a smoother weight profile.

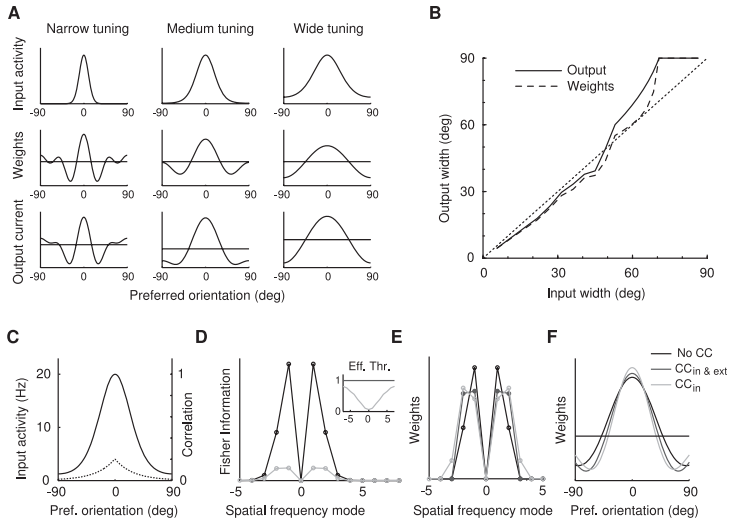

Figure 2: Exact optimization of synaptic weights in the firing rate model. (A) Input layer activity (top), optimal synaptic weights (middle), and output synaptic current (bottom) for three different tuning curve widths: narrow ($a = 1/6$, left), medium ($a = 1/3$, center), and wide tuning ($a = 1$, right). Other parameters: $q = 2, c_0 = c_1 = 0, \text{sigma}_0 = \sigma_1 = 2\,\text{Hz}^2, r_p = 20\,\text{Hz}, N = 501$. (B) Output width of optimal synaptic weights (solid) and output currents (dashed) as a function of tuning width in the input layer. Output and input widths are approximately matched. (C–F) Noise correlations sharpen the optimal weights. Three cases were considered: No noise correlations (black), noise correlation in the input layer only, and uncorrelated external noise (light gray empty circles; $c_0 = 0.2$ and $\rho_0 = 1$), and identical correlation structure in input layer and external noise (dark gray filled circles; $c_0 = c_1 = 0.2$ and $\rho_0 = \rho_1 = 1$). C, Tuning curve in the input layer (solid) and correlation structure (dotted lines). (D) Correlations strongly reduce the Fisher information in the input layer (light and dark gray overlapping), compared to the uncorrelated case (black). Inset, Effective threshold for the synaptic weights $T(n)$ (black and dark gray overlapping). (E) The optimal weights in the spatial frequency domain. In the absence of correlations, the weights amplify only the most informative, low-frequency modes (black). The effective threshold across frequencies is the same when the input and external correlations are equal, as in the case where correlations are absent. But since correlations decrease the information at the lowest-frequency modes (see panel D), more weight is devoted to higher frequencies (dark gray). Finally, when the external noise is uncorrelated, the low-frequency thresholds decrease (D, inset). The information in the lowest frequencies can now be transmitted with smaller weights, allowing further emphasis of higher frequencies (light gray). (F) After inverse Fourier transformation, the optimal weights are narrower when correlations are present. (C–F). Parameters as in A, width $a = 0.85$. We displayed the highest frequencies as negative ($n = N - m \to n = -m$) to highlight the symmetry of the Fourier transforms.

In Figure 2B, the width of the connectivity profile and the output width are plotted as a function of the tuning curve width. The width was calculated as the full width at half maximum. In general, the width of the optimal connectivity approximately matches that of the input tuning curve. The width has a complicated dependence on the Fourier components. As a result, the width can be either narrower or wider than the input, even in this highly idealized scenario. The optimal weights depend on the noise as well. If, for instance, the postsynaptic noise is small, the optimal weights will transmit intermediate spatial frequencies. If, in contrast, the postsynaptic noise is large, the threshold is high and the weights are concentrated around the large, low-frequency modes, yielding broader weight profiles.

Next, we address the optimal transmission strategy in the presence of noise correlations. Although in some cases, noise correlations can increase the information of neural codes (Oram et al., 1998; Averbeck, Latham, & Pouget, 2006), they typically reduce it, as they limit the potential for averaging out noise by pooling over neurons (Britten, Shadlen, Newsome, & Movshon, 1992; Zohary, Shadlen, & Newsome, 1994), which is reflected in a saturation of the FI in the presence of limited-range correlations (Sompolinsky et al., 2002; Wu et al., 2002). Noise correlations affect the optimal weights in two ways. First, correlations change the amount of information present at each spatial frequency in the input layer (see Figure 2D). In particular, limited-range correlations reduce the signal-to-noise ratio at low spatial frequencies, which are the most informative. Second, noise correlations in the input lower the effective thresholds $T(n)$, so that less power is required in the low-frequency weights (see equation 2.8). Thus, in the presence of limited range correlations, the optimal weights emphasize the intermediate frequencies more than for uncorrelated noise, leading to a narrower profile (see Figure 2E and 2F).

Although the precise shape of the optimal weight profile depends on detailed properties of the population firing statistics (see equation 2.8), it has several general properties. Because there is no information about the stimulus in the average activity of the population, the optimal weights are balanced: their zero-frequency mode (the area under the weight profile) is 0. Second, if the tuning curves are not too narrow, most information is in the lowest spatial frequencies, which are amplified the most by the optimal profile. In summary, the optimal weights have the most power in the low-frequency modes and none at 0, which implies that they have a center-surround or Mexican hat structure.

3 Application to Spiking Networks

The calculation in the previous section made a number of assumptions that might not hold for spiking neuron models. In particular, the noise will not be gaussian nor independent of the stimulus. Moreover, it is more natural to constrain the peak firing rate in the output layer rather than the sum of

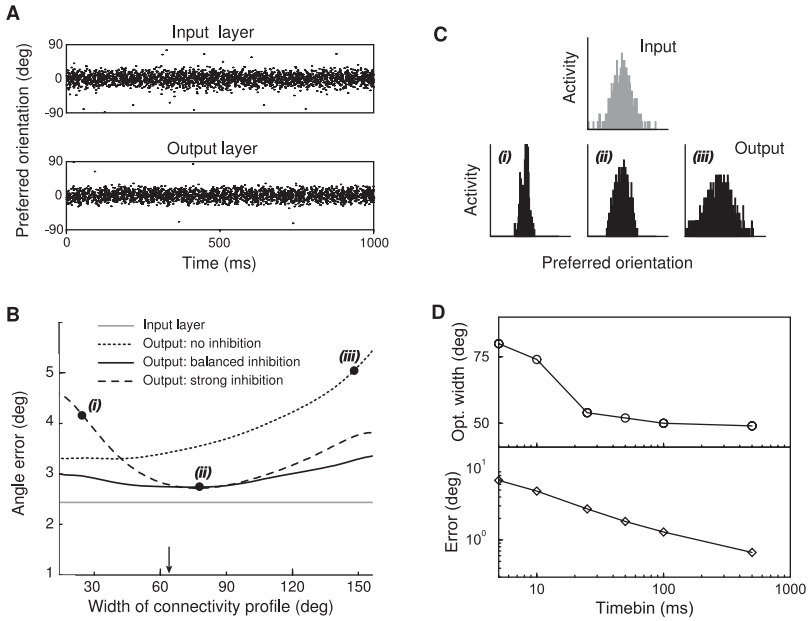

Figure 3: Optimal connectivity in a network of spiking integrate-and-fire neurons. (A) Snapshot of 1 second of the activity in the input (top) and output (bottom) layers (optimal connectivity). (B) Error of the optimal linear estimator of the stimulus as a function of the width of the connectivity profile for various levels of surround inhibition. The error in the estimator in the input layer is shown by the solid gray line. Without inhibition, estimation is suboptimal for all widths (dotted curve). With balanced inhibition (solid curve), the error reduces considerably and approaches the input error. The width of the optimal profile is similar to but slightly larger than the tuning curve width (arrow). Stronger inhibition (dashed curve) works well if the width is matched to the tuning curve width, but the error is large when the profile is mistuned. (C) Three examples of the input and output bumps of activity. (i) Narrow connectivity with strong inhibition. In this case, the strong inhibition leads to narrow and highly variable activity in the output. (ii) Balanced inhibition at the optimal width. Input and output activity profiles are very similar in the optimal case. (iii) Wide connectivity with no inhibition. The output layer blurs in activity profile. (D) The optimal width (top) as a function of the duration of the decoding time window. For shorter windows, a broader connectivity is optimal. The minimal error (bottom) decreases as the square root of the decoding time window.

the weights squared. To test the generality of the results, we simulated a two-layer network of spiking integrate-and-fire neurons (see Figure 3A).

We use a network consisting of 128 current-based integrate-and-fire neurons per layer with periodic boundary conditions. The biophysical

parameters of the neurons are 100 MΩ input resistance R_m, 20 ms time constant τ, −60 mV reset and resting potential V_r, −50 mV firing threshold and 1 ms refractory period (van Rossum, Turrigiano, & Nelson, 2002). The dynamics are $\tau_m \frac{dV}{dt} = -(V(t) - V_r) + R_m I_{in}$, where I_{in} is the input current consisting of both synaptic current and noise currents. On top of that, the input layer also receives a static profile of current injections with a gaussian shape according to equation 2.1, with a full width at half maximum of 60 degrees. After a spike, the membrane potential resets to the resting potential, where it remains clamped for 1 ms (absolute refractory period).

All neurons are injected with a gaussian-distributed, low-pass-filtered noisy background current with positive mean (mean 55 pA, standard deviation 70 pA, filter time constant 2 ms). This noise current mimics the intrinsic noise in the neurons and unspecific input from other neurons. The noise prevents synchronization of the neural activity and ensures that firing statistics, response latencies, and membrane potential distributions mimic those found in vivo, yielding so-called rate-mode propagation (Knight, 1972; Gerstner, 2000; van Rossum et al., 2002). We note that in more synchronous modes of activity, such as the synfire chain, information quantification is much more complicated because information can be coded in spike timing and packet size (Hamaguchi et al., 2005).

The synapses are modeled as current sources, which are positive (negative) for excitatory AMPA (inhibitory GABA) synapses; both have simple exponential decay with a 5 ms time constant. Conductance-based synapses lead to very similar results but do not allow us to make as precise claims about balance as the current-based synapses. The synapses have a release probability of 50%. For simplicity, neurons can make both excitatory and inhibitory synapses.

3.1 Parameterization of the Connectivity Profile. The output layer is driven by the input layer through excitatory and inhibitory synapses via a weight matrix. A full optimization of the weight matrix is not feasible in the simulations. Instead, the synaptic connections between layers are parameterized by a rectified cosine with a uniform inhibitory surround. The connection strength between a neuron at position i and a neuron at position j in the next layer equals $W_{ij} = A\cos(\frac{\phi_{ij}}{\phi_w}90^\circ)$ wherever $|\phi_{ij}| = |\phi_i - \phi_j| \leq \phi_w$, while $W_{ij} = -c\sum_i [W_{ij}]_+$ otherwise. 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 give very similar results; the parameterization used allows a straightforward adjustment of width ϕ_w, inhibition strength c, and amplitude A. The amplitude A is calibrated for every choice of connectivity profile, so that the peak firing rates in the input and output layer are identical (50 Hz). This is important as the gain

influences the transmission quality: the more output spikes are produced, the more accurate the estimate can be. The calibration prevents this effect.

3.2 Measuring Information Transmission in Spiking Networks. Although these simulations have the advantage of being realistic, measuring the true FI in simulations of spiking neurons is computationally prohibitive for nongaussian, correlated noise. 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 25 ms time bins. We estimated the position of the stimulus using an optimal linear discriminator, which provides a lower bound on the Fisher information (Shamir & Sompolinsky, 2004; Seriès et al., 2004). On each trial, spike counts were used as the input to the discriminator. On half the trials, the stimulus had an angle $+\Delta\phi$ (7°), while on the other trial, the angle was $-\Delta\phi$. A linear discriminator had to distinguish these two stimuli by calculating the dot product between the weights of the discriminator $\mathbf{v}$ and the firing rates of the population for either stimulus, $z_{\pm\Delta\phi} = \mathbf{v}.\mathbf{r}(\pm\Delta\phi)$. 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 output of the discriminator was, on average, different between the two stimuli, but fluctuated on each trial, thereby limiting the discrimination. The quality of the discriminator was characterized by its signal-to-noise ratio (SNR). The discriminator weights were optimized by numerically minimizing the SNR with respect to the Fourier coefficients a_k (eight Fourier coefficients proved sufficient; using more did not improve the SNR). This optimization was done independently for every setting of the connectivity. The SNR of the discriminator is converted to the standard deviation in the angle estimate, which equals the inverse square root of the Fisher information estimate (Seriès et al., 2004; Shamir & Sompolinsky, 2004).

3.3 Effect of Connectivity on Information Transmission. First, we research the effect of the connectivity profile on transmission quality. Previously the connectivity profile in the simulations was parameterized by the width of the excitatory center and the strength of the inhibition. In Figure 3B the trial-to-trial standard deviation in the stimulus angle (a lower bound on the inverse square root of the FI) is plotted as a function of the width of the connectivity profile and various levels of inhibition. Compare the error to the error in the input layer (the gray line).

First, we used an excitatory-only weight profile. The error in the output layer (the dotted curve) is substantially higher than in the error in the input. The error in the output is largest for wide weight profiles, as in this case, the output activity is smeared, (see Figure 3C, example iii). Next, we tested profiles with a balanced inhibitory surround. Now, neurons away from the center of the stimulus will experience a net inhibition, preventing them from firing. Adding surround inhibition substantially improves the

accuracy in the output layer. For a balanced, center-surround profile (solid curve in Figure 3B) with a half-width approximately matching the width of the activity profile (arrow), the error is minimal. The information loss between the layers is remarkably small (compare the gray to solid black curve). In line with the analytical results on the width of input and output (see Figure 2B), the output activity profile is in the optimal case very similar to the input profile (example ii in Figure 3C).

The exact amount of inhibition required turned out not to be critical as long as the width of the profile is matched to that of the tuning curve. A three-fold increase in the inhibition yields similar performance for a profile with an optimally tuned width (see the dashed curve in Figure 3B). However, strong inhibition decreases performance if the weight and tuning widths are mismatched (e.g., Figure 3C, example 1). These simulation results are in line with the theoretical analysis above: in order to preserve information, the connectivity profile needs to amplify spatial frequencies that convey the most information, while filtering out the high and zero spatial frequencies.

To compare this to the theory, we use the derived algorithm to calculate the optimal profile for the tuning curve used in the simulation. We adjusted the constraint parameter q so that the peak output firing rate equaled the peak input firing rate. The theoretically optimal profile had a width of 64 degrees, close to the numerical optimum. Using this profile, the standard deviation in angle estimate in the output was 2.73 degrees, which equals the minimal error found numerically.

Next, we examined the role of the readout window on the accuracy and the shape of the optimal profile. We used readout windows (Δt) of 5, 10, 25, 50, 100, and 500 ms. A longer readout time gives a more accurate estimate because noise is averaged out more. The error in the readout decreased approximately as $err \propto 1/\sqrt{\Delta t}$, as expected from averaging uncorrelated noise (see Figure 3D, bottom). The width of the optimal profile is weakly dependent on the time bin (see Figure 3D, top). For short time windows, a slightly wider profile is optimal. This is similar to findings that optimal tuning curves are broader for short integration times to ensure that it still captures the few spikes in the network (Brunel & Nadal, 1998; Bethge et al., 2002). For longer time bins, the noise is more gaussian, and the optimal width converges to a smaller value.

4 Discussion

We have considered a scenario where information is processed in stages and thus needs to be represented in multiple brain regions. Based on the Fisher information, our results show that the information encoded by a population of broadly tuned neurons is best transmitted through center-surround connectivity profiles. The width of the center-surround profile roughly matches the width of the tuning curves in the input population,

and as a result, the tuning width in the output layer approximately matches that of the input layer. Thus, the tuning width in deeper layers cannot be sharpened or widened arbitrarily without a cost; instead, it is dictated by the input.

Our findings have implications for the interpretation of neural connectivity. In sensory systems, center-surround architecture has been interpreted as a contrast enhancement mechanism and as a way to minimize redundancy at early states of the visual processing stream by removing stimulus correlations (Atick & Redlich, 1990). However, it has been unclear why contrast enhancement would be needed at every stage. Our results suggests that the abundance of center-surround connectivity in many parts of the nervous system (Hartline, Wagner, & Ratcliff, 1956; Hubel & Wiesel, 1962; Enroth-Cugell & Robson, 1966; Trappenberg, Dorris, Munoz, & Klein, 2001) is important for accurate population code processing throughout the nervous system and is dictated by the need to filter out the uninformative high and zero spatial-frequency modes.

We note that our result on the optimal weight profile is reminiscent of the water-filling algorithm used to maximize the transmission of Shannon information in a temporal signal with a power-limited transmitter (Cover & Thomas, 1991). In both cases, signals are transmitted independently through different (spatial resp. temporal) frequency channels, each channel is corrupted with channel-specific noise, and there is a limit to the total power. Although the exact results are mathematically distinct, the constraints and the Shannon and Fisher information are different; in both cases, the most informative frequencies are transmitted, while other frequencies are thresholded out.

We believe that quantifying information transmission between populations presents an important concept that should be contrasted to the problem of how to read out a given population. A read-out algorithm can be thought of as a map from high-dimensional population activity to a low-dimensional encoded quantity, such as stimulus orientation. But in the nervous system, although the encoded quantity is processed as activity flows through, say, the sensory-motor stream, subsequent stages do not read out the previous one. Instead, the high-dimensional nature of the encoding process is preserved until the very last stages, where a small set of muscles is driven by the corresponding motor neurons.

While this study considers only the transmission of activity, the ultimate goal is to understand information processing. The findings here are also applicable to computing networks. In an earlier study using recurrent networks to implement an optimal read-out algorithm, the width of the connectivity profile had to be set approximately equal to the tuning curve width (Deneve et al., 1999). Another study found that spiking feedforward networks employing a width-matched center-surround connectivity accurately computed with population codes (van Rossum & Renart, 2004). Finally, in a man-made readout of population codes, matching

the connectivity to the activity profile was found to give good performance (Georgopoulos, Taira, & Lukashin, 1993). Our results help to explain those findings and might also be of use for recurrent networks.

Appendix: Optimal Connectivity Profile

The optimal connectivity profile is found by maximizing equation 2.7 subject to the constraint $\sum_n |\widetilde{W}(n)|^2 = q$. This can be performed by using Lagrange multipliers. Define the Lagrangian

$$L = \sum_n \frac{J(n)|\widetilde{W}(n)|^2}{|\widetilde{W}(n)|^2 + T(n)} - \lambda \left(\sum_n |\widetilde{W}(n)|^2 - q \right),$$

where λ is the Lagrange multiplier. Imposing $\partial L / \partial |\widetilde{W}(n)| = 0$ results in

$$|\widetilde{W}(n)|^2 = \sqrt{T(n)} \left[\sqrt{J(n)/\lambda} - \sqrt{T(n)} \right]_+ . \tag{A.1}$$

The rectification appears because $|\widetilde{W}(n)|^2 > 0$, and λ has to be chosen so that the constraint is satisfied. We solve these equations self-consistently. Given an initial arbitrary value of λ, one calculates $|\widetilde{W}(n)|^2$ using equation A.1. Let us denote by $\{n'\}$ those frequency modes for which $|\widetilde{W}(n)|^2 > 0$. Imposing the constraint, we obtain

$$\frac{1}{\sqrt{\lambda}} = \frac{q + \sum_{n'} T(n')}{\sum_{n'} \sqrt{J(n')T(n')}}.$$

This new value is used to calculate the updated $|\widetilde{W}(n)|^2$, and this procedure is repeated until λ no longer changes, at which point $\sum_{n'} |\widetilde{W}(n)|^2 = q$. Substituting this final value for λ into equation A.1 results in the solution for the optimal weights in equation 2.8. The final weight matrix is retrieved by transforming back into real space using

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

Note that formally, our solution specifies only the squared magnitude of the weights, that is, their power spectrum. Thus, any solution obtained by multiplying the $|\widetilde{W}(n)|$ by any phase $e^{i\beta_n}$ is also a solution. However, because we require the weight profile to be an even function, the Fourier coefficients can only be real.

Acknowledgments

We thank Larry Abbott, David Barber, Mike Oram, Peggy Series, Gina Turrigiano, and Si Wu for discussions. M.vR. was supported by the EPSRC and HFSP. A. R. was supported by a start-up Ramon y Cajal grant from the Spanish Ministry of Education. A. R. thanks X-J. Wang and K. D. Harris for support during the initial and final stages of this project.

References

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

Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Comput.*, *2*(3), 308–320.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nat. Rev. Neurosci.*, *7*(5), 358–366.

Bethge, M., Rotermund, D., & Pawelzik, K. (2002). Optimal short-term population coding: When Fisher information fails. *Neural Comput.*, *14*, 2317–2351.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J. Neurosci.*, *12*, 4745–4765.

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

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. Hoboken, NJ: Wiley.

Dean, A. F. (1981). The variability of discharge of simple cells in cat striate cortex. *Exp. Brain. Res.*, *44*, 437–440.

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

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol.*, *187*(3), 517–552.

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci. 9*(4), 292–303.

Georgopoulos, A. P., Taira, M., & Lukashin, A. (1993). Cognitive neurophysiology of the motor cortex. *Science*, *260*, 47–52.

Gerstner, W. (2000). Population dynamics of spiking neurons: Fast transients, asynchronous state, and locking. *Neural Comput.*, *12*, 43–89.

Hamaguchi, K., Okada, M., & Aihara, K. (2007). Variable timescales of repeated spike patterns in synfire chain with Mexican-hat connectivity. *Neural Comput.*, *19*(9), 2468–2491.

Hamaguchi, K., Okada, M., Yamana, M., & Aihara, K. (2005). Correlated firing in a feedforward network with Mexican-hat-type connectivity. *Neural Comput.*, *17*(9), 2034–2059.

Hartline, H. K., Wagner, H. G., & Ratcliff, F. (1956). Inhibition in the eye of limulus. *J. Gen. Physiol.*, *39*, 651–673.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, *160*, 106–154.

Kay, S. M. (1993). *Fundamentals of statistical signal processing: Estimation theory*. Upper Saddle River, NJ: Prentice Hall.

Knight, B. W. (1972). Dynamics of encoding in a population of neurons. *J. Gen. Physiol.*, *59*, 734–766.

Lee, C., Rohrer, W. H., & Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, *332*, 357–360.

Oram, M. W., Foldiak, P., Perrett, D. I., & Sengpiel, F. (1998). The "ideal homunculus": Decoding neural population signals. *Trends Neurosci.*, *21*, 259–265.

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

Seriès, P., Latham, P. E., & Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nat. Neurosci.*, *7*(10), 1129–1135.

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

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci.*, *18*, 3870–3896.

Shamir, M., & Sompolinsky, H. (2004). Nonlinear population codes. *Neural Comput.*, *16*(6), 1105–1136.

Sompolinsky, H., Yoon, H., Kang, K., & Shamir, M. (2002). Population coding in neuronal systems with correlated noise. *Phys. Rev. E*, *64*, 51904.

Trappenberg, T. P., Dorris, M. C., Munoz, D. P., & Klein, R. M. (2001). A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *J. Cogn. Neurosci.*, *13*(2), 256–271.

van Rossum, M.C.W., & Renart, A. (2004). Computation with population codes in layered networks of integrate-and-fire neurons. *Neurocomputing*, *58*, 265–270.

van Rossum, M.C.W., Turrigiano, G. G., & Nelson, S. B. (2002). Fast propagation of firing rates through layered networks of noisy neurons. *J. Neurosci.*, *22*, 1956–1966.

Wu, S., Amari, S., & Nakahara, H. (2002). Population coding and decoding in a neural field: A computational study. *Neural Comput.*, *14*, 999–1026.

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

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, *370*, 140–144.

Received April 5, 2011; accepted August 7, 2011.