Published in final edited form as:
Neuron. 2016 January 20; 89(2): 409–422. doi:10.1016/j.neuron.2015.12.037.

Structures of Neural Correlation and How They Favor Coding

Felix Franke[1,2,3], **Michele Fiscella**[1], **Maksim Sevelev**[2,3], **Botond Roska**[4], **Andreas Hierlemann**[1], and **Rava Azeredo da Silveira**[2,3,5,*]

[1]Department of Biosystems Science and Engineering, ETH Zürich, 4058 Basel, Switzerland [2]Department of Physics, Ecole Normale Supérieure, 75005 Paris, France [3]Laboratoire de Physique Statistique, Centre National de la Recherche Scientifique, Université Pierre et Marie Curie, Université Denis Diderot, 75005 Paris, France [4]Friedrich Miescher Institute, 4058 Basel, Switzerland [5]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA

Summary

The neural representation of information suffers from “noise”—the trial-to-trial variability in the response of neurons. The impact of correlated noise upon population coding has been debated, but a direct connection between theory and experiment remains tenuous. Here, we substantiate this connection and propose a refined theoretical picture. Using simultaneous recordings from a population of direction-selective retinal ganglion cells, we demonstrate that coding benefits from noise correlations. The effect is appreciable already in small populations, yet it is a collective phenomenon. Furthermore, the stimulus-dependent structure of correlation is key. We develop simple functional models that capture the stimulus-dependent statistics. We then use them to quantify the performance of population coding, which depends upon interplays of feature sensitivities and noise correlations in the population. Because favorable structures of correlation emerge robustly in circuits with noisy, nonlinear elements, they will arise and benefit coding beyond the confines of retina.

Introduction

There is a long history of studies of the representation of information by single neurons, but the investigation of coding by populations of neurons is far less advanced. High dimensionality makes the problem difficult: on the experimental side, because access to the simultaneous activity of an entire population is necessary; on the theoretical side, because the way in which information is coded in this simultaneous activity has to be elucidated. The problem is high dimensional, in particular, since different neurons can process their inputs in different ways and since their outputs can be correlated. Coding in the face of these two aspects—the diversity of cell types and neural response properties, and the correlation in neural activity—is the topic of the present paper.

*Correspondence: rava@ens.fr.

Author Contributions

Conceptualization: F.F. and R.A.d.S.; Analytical work: F.F., M.S., and R.A.d.S.; Numerical work: F.F.; Neural recordings: M.F.; Writing—Original Draft: F.F. and R.A.d.S.; Writing—Reviewer & Editing: F.F., M.F., B.R., A.H., and R.A.d.S.; Supervision: B.R., A.H., and R.A.d.S.; Project Administration: R.A.d.S.; Funding Acquisition: B.R., A.H., and R.A.d.S.

Europe PMC Funders Author Manuscripts

Neural noise is a theme that runs through any discussion of neural coding. In the simplest case, the noise in the output of a neuron is independent from that in other neurons. Noise then makes the output of each neuron less precise and thereby harms the coding performance. In contrast to this idealized case, many brain areas display “noise correlation,” i.e., correlation in the variability of the output of a population of neurons in response to a given stimulus (Hatsopoulos et al., 1998; Mastronarde, 1989; Ozden et al., 2008; Perkel et al., 1967; Sasaki et al., 1989; Zohary et al., 1994; Shlens et al., 2008; Usrey and Reid, 1999; Vaadia et al., 1995; Bair et al., 2001; Fiser et al., 2004; Kohn and Smith, 2005; Smith and Kohn, 2008; Lee et al., 1998; Ecker et al., 2010; Graf et al., 2011; Goris et al., 2014; Lin et al., 2015). The problem now becomes higher dimensional: beyond the degree of noise in each neuron, one has to be concerned with all the other quantities that describe the statistical relations among neurons. In other words, when examining the impact of noise on coding, one has to take into account not only the magnitude of the noise, but also its structure.

In past years, a number of theoretical proposals on the way in which noise correlation may impact coding have been put forth (Johnson, 1980; Vogels, 1990; Oram et al., 1998; Abbott and Dayan, 1999; Panzeri et al., 1999; Sompolinsky et al., 2001; Wilke and Eurich, 2002; Romo et al., 2003; Golledge et al., 2003; Pola et al., 2003; Averbeck and Lee, 2003, 2006; Shamir and Sompolinsky, 2004, 2006; Averbeck et al., 2006; Josić et al., 2009; Ecker et al., 2011; Hu et al., 2014; da Silveira and Berry, 2014). Many of these have argued that noise correlation is detrimental to the coding performance, and this statement was, for some time, taken as a rule of thumb. More recently, some studies have demonstrated that correlation can be beneficial to neural coding, depending upon its structure (Ecker et al., 2011; Hu et al., 2014; da Silveira and Berry, 2014; Lin et al., 2015; see also Wilke and Eurich, 2002; Moreno-Bote et al., 2014). In most cases, however, a model of the structure of correlation was assumed, and its consequences were derived: there are, as yet, very few direct demonstrations of the role of correlation in coding, as these require large volumes of simultaneous recordings of populations of neurons over a range of stimuli or tasks. Analyses of cortical data (Averbeck and Lee, 2004, 2006; Montani et al., 2007; Graf et al., 2011; Lin et al., 2015) indicate that correlation can benefit coding, and a number of earlier studies (Maynard et al., 1999; Romo et al., 2003; Cohen and Newsome, 2008; Poort and Roelfsema, 2009; Gutnisky and Dragoi, 2008) similarly suggest, though in a less direct manner, that correlation can enhance the coding performance in cortex. There exist few studies along these lines in the context of retina; one such study (Nirenberg et al., 2001) argued that correlation affects coding only weakly, while others highlighted a potential greater relevance of correlation (Schnitzer and Meister, 2003; Schneidman et al., 2003).

Determining whether, in a given population of neurons, noise correlation is beneficial or detrimental to coding, and quantitatively how important the effect is, will be essential for understanding neural coding beyond the single cell. Here, we address these issues in the context of populations of retinal neurons of a given type, direction-selective cells, and we then extend our considerations to more general populations of broadly tuned neurons. Since we recorded the simultaneous responses of populations of direction-selective cells to a battery of stimuli, it was possible to extract from our data the structure of the noise in these populations. With this knowledge at hand, we show not only that correlation is beneficial to coding, but also that it is its specific structure that grants it its favorable role. We then

introduce simple models of correlated activity that capture the beneficial structure of correlation; these models are not only simple but also robust, as they rely on common circuit properties like diverging synapses and non-linear neural transfer functions. Given their generality, these models extend our results beyond the confines of the retina.

Several broad conceptual points emerge from our analyses:

- Noise correlations can benefit sensory coding appreciably even in small populations of four or eight cells, yet this follows from a collective phenomenon that operates beyond individual pairs of neurons.
- The coding enhancement results from the form of stimulus-dependent correlations observed in data. Simple models that rely on common circuit properties show how these structured correlations emerge in a robust manner.
- In large populations of correlated neurons, an unexpected physiological organization can be advantageous, namely using a set of discrete tuning preferences in the population to code for a continuous stimulus, rather than covering the dynamic range of the stimulus uniformly.

A few earlier studies (Panzeri et al., 1999; Pola et al., 2003; Shamir and Sompolinsky, 2004) have considered the effect of stimulus-dependent noise correlation upon coding. Our approach is complementary to theirs, as it differs in the overall information theoretic framework, the models of noise correlation, and especially the mechanism by which coding benefits from correlation. Furthermore, here we present a direct demonstration that noise correlation benefits coding in an identified neural population.

Results

Measurements of Noise Correlation in Retinal Direction-Selective Neurons

We address the role of noise correlation in sensory coding using simultaneous recordings from populations of identified cell types, namely retinal direction-selective ganglion cells in rabbit retina. These recordings were obtained with a high-density microelectrode array (Frey et al., 2009), which featured an electrode density comparable to the density of retinal ganglion cells. After localizing and selecting the cells of interest, we recorded from each of the identified direction-selective ganglion cells with at least five electrodes, with high signal-to-noise ratio (Fiscella et al., 2012). From these population recordings, we derived response and correlation properties in the population.

We presented the rabbit retina with moving bar stimuli. A direction-selective cell responds most vigorously to a bar that moves in its so-called preferred direction (Figure 1A) and with a broad tuning curve about this preferred direction (Figures 1B and 1C) (see Fiscella et al., 2015 for details on experimental procedures and a thorough study of response properties of direction-selective cells). We focused upon the four types of ON-OFF retinal direction-selective cells whose preferred directions point along the four cardinal directions (Figure 1B). While the peak amplitude of the response could vary from cell to cell (Figure 1C), in particular as a function of the location of the stimulus with respect to the cell's receptive field, the shapes of the tuning curves were comparable for different cells, as seen when

normalizing their peak amplitude (Figure 1B). Typically, a tuning curve covered a little more than 180°, i.e., each direction-selective neuron "monitored" the stimulus direction in a range somewhat over 90° on each side of its preferred direction (Figure 1C).

Neural responses to repetitions of the same stimulus varied appreciably from trial to trial (clouds of points in Figure 1D). Thus, for any given stimulus, noise in the response was substantial. This noise, furthermore, came with a specific structure across stimuli. Because direction-selective neurons can have one of the four cardinal directions as preferred direction, there are three types of pairwise correlations to consider: between two neurons with similar preferred directions, with preferred directions that differ by 90°, and with preferred directions that differ by 180° (Figure 1E). When plotted as functions of the stimulus direction, these three correlation values displayed a consistent modulation about their mean value (Figure 1E). Roughly, in all three cases the pairwise correlation was large when the product of the tuning curves was large. In other words, the correlation between two neurons behaved similarly to the geometric mean of their two tuning curves (Figure 1E).

We note that the correlation values we observed in pairs of neurons sometimes exceeded by an appreciable amount values often quoted in the literature on cortex (Hatsopoulos et al., 1998; Mastronarde, 1989; Ozden et al., 2008; Perkel et al., 1967; Sasaki et al., 1989; Zohary et al., 1994; Shlens et al., 2008; Usrey and Reid, 1999; Vaadia et al., 1995; Bair et al., 2001; Fiser et al., 2004; Kohn and Smith, 2005; Smith and Kohn, 2008; Lee et al., 1998; Ecker et al., 2010; Graf et al., 2011; Goris et al., 2014). Most of the values quoted in the literature, however, correspond to averages over cell pairs and over stimuli. If we also average the correlation values that we find over pairs and stimuli, the resulting mean correlations are comparable to those previously reported. In a recent cortical study (Lin et al., 2015) in which correlation values were not averaged over either tuning preferences or stimuli, these took much larger values, ranging from 0.2 to 0.8, with a large fraction of pairs above 0.5—values not only comparable, but in fact larger than the ones we observed in retina.

From Spiking Data to Coding Performance: The Impact of Noise Correlation

Noise in the response of direction-selective cells is appreciable, and correlation shapes it in a particular manner. Because of its non-negligible magnitude, one expects that noise has an effect on the coding performance in a population and that the correlations it carries can modulate this effect. To begin, we quantified the impact of noise correlation upon the coding performance through direct data analysis.

For a quantitative assessment, we compared the experimental, correlated populations of neurons to "equivalent" independent populations, in which all mean responses and variances of individual neurons were preserved as in the data, but the correlations were removed (see Supplemental Experimental Procedures, Sec. 1.3). As a measure of the coding performance, we used the Cramér-Rao bound (the inverse square root of the Fisher information). Given the level of noise in the circuit and its structure, this provides a lower bound, $\delta\theta^2_{\text{correlated}}(\theta)$, on the performance of any estimator of motion direction: $\text{var}\,(\hat{\theta}) \geq \delta\theta^2_{\text{correlated}}(\theta)$, where $\text{var}(\hat{\theta})$ is the variance of the decoded direction, $\hat{\theta}$, given a stimulus direction, θ (see Supplemental Experimental Procedures, Sec. 1.2) (Abbott and Dayan, 1999, Sompolinsky et

al., 2001, Wilke and Eurich, 2002, Josić et al., 2009). We compared this quantity to its counterpart, $\delta\theta^2_{\text{independent}}(\theta)$, the Cramér-Rao bound obtained from a version of the data in which we removed correlations. Finally, we quantified the impact of noise correlation through the "mean percent improvement":

$$\Delta R = \left(1 - \left\langle \frac{\delta\theta^2_{\text{correlated}}}{\delta\theta^2_{\text{independent}}} \right\rangle_\theta \right) \times 100, \qquad \text{(Equation 1)}$$

where $\langle\cdot\rangle_\theta$ denotes an average over all stimulus directions. When $\Delta R = 0\%$, correlations do not affect coding on average; but when ΔR is larger, say $\Delta R = 50\%$, this corresponds to a boost in coding precision by a factor of two in a correlated population as compared with an independent population.

We analyzed data obtained in several experiments (see Experimental Procedures and Fiscella et al., 2015). For each experiment, we randomly selected populations of direction-selective neurons among all simultaneously recorded cells; we repeated this sampling many times, with the same fixed number of cells in the population. As expected, the accuracy of the coding of direction, as described by $\delta\theta_{\text{correlated}}$, improved with the size of the population (Figure 2A). The surprise was that this enhancement was stronger for the correlated neurons in the real data than for independent populations (Figure 2B). Furthermore, this relative enhancement grew with the size of the population, i.e., correlation was increasingly beneficial in larger populations (Figure 2B). A study based on paired intracellular recordings supports a similar benefit of correlation for coding in retinal direction-selective cells (Zylberberg et al., 2016, this issue).

An appreciable fraction of the effect of correlation was present already in small populations of cells (Figure 2B). Owing to this observation, and because our data afforded more reliable statistics for small populations, we turned to examining populations of four direction-selective neurons with preferred directions aligned with the four cardinal directions ("quadruplets," Figure 3, insets) and of eight direction-selective neurons with four pairs of neurons with preferred directions aligned with the four cardinal directions ("octuplets," Figure 3, insets). For quadruplets and octuplets, bounds to the error in direction coding averaged at 10° and 5°, respectively. More interesting was the behavior of the relative improvement due to correlation. We observed a large effect: the average percent improvement over all quadruplets was 19%, and the corresponding value for octuplets was 23% (Figure 3A). In a number of examples, the percent improvement reached 50%, which corresponds to a 2-fold enhancement of precision due to noise correlation (Figure 3A). Also noteworthy is the observation that correlation was beneficial for all quadruplets and octuplets: we did not see any for which coding accuracy was degraded by the presence of correlation.

Given the consistent stimulus dependence in the correlation (Figure 1E), we wondered to what extent this structure contributed to the favorable impact of correlation on coding. In order to investigate this point, we replaced the correlations in pairs of neurons by their average over all stimulus presentations; the tuning curves of the cells and the variances in

their output were kept unchanged and depended upon the stimulus, but now correlation took the same value for different stimuli: its structure was “flattened” (dashed lines in Figure 1E). To be explicit, we considered three cases: populations of (1) correlated neurons with stimulus-dependent correlation, as given by the data; (2) correlated neurons with stimulus-independent correlation; and (3) independent neurons (see also Supplemental Experimental Procedures, Sec. 1.3). We then calculated the percent improvement (Equation 1) for 1 versus 3 and for 1 versus 2. (We refer to “favorable correlation” when the coding performance is greater in case 1 or 2 than in case 3. To avoid any possible confusion, we emphasize that we do not mean here that noise itself is favorable, but rather, given that noise is present, that correlation can sculpt it in a manner beneficial to coding.)

In the case of quadruplets, a generic improvement of the coding accuracy survived when the stimulus dependence in correlation was removed, but it became modest in magnitude (6% on average, Figure 3B). The average improvement in the case of octuplets was weaker, because in an appreciable fraction of examples, correlation became detrimental (negative percent improvement, Figure 3B). In a study in the context of mouse retina, Zylberberg et al. (2016) observed an improvement of the coding performance accompanied by a similar shaping of the noise; this indicates that the findings apply more broadly to retinal processing. Furthermore, Lin et al. (2015) reported a fine structure of the noise correlations as a function of tuning preference and stimulus very similar to the one we observed in retina, as well as its incidence upon coding, in the context of cortex. The convergence of these observations indicates that the phenomenon under scrutiny extends beyond retina and applies more generally to information processing in neural circuits. It also suggests that the observed structure in noise correlation, which we examine in detail next, is key to the improvement of the coding performance.

Functional Models of Correlated Neurons Coding for Direction

These observations—namely, that noise correlations can impact the neural code appreciably and that their stimulus-dependence is critical—motivated us to formulate functional models of neural response that capture the tuning properties of direction-selective cells as well as the correlated fluctuations in the population activity. These models can then be used (and we use them; see below) to analyze in greater detail the effect of various response properties, including the stimulus-dependence of noise correlations, upon the coding performance.

Likely, several biological mechanisms are at play in generating the observed correlations (Zylberberg et al., 2016). Rather than pursuing a mechanistic model, however, we looked for simple, phenomenological models that may capture the measured population responses. Recently, such models have proven successful in describing the correlated variability of cortical cells (Ecker et al., 2014; Goris et al., 2014; Lin et al., 2015); these models assume a doubly stochastic process, in which the intrinsic variability of single-cell response is combined with a secondary variability in the output gain. We borrowed this framework and extended it to allow for diverse sources of secondary variability.

In order to construct functional models, we assumed that the spiking output of a neuron labeled μ, in a given stimulus trial, was governed by a rate, λ_μ, so that its mean response was

equal to λ_μ and the variance about this mean was $\lambda_\mu^{1-\beta}$, where $\beta < 0$, $\beta = 0$, and $\beta > 0$ corresponded to supra-Poisson variability, Poisson variability, and sub-Poisson variability, respectively. This intrinsic variability was interpreted as internal to the neuron, rather than coming from circuit effects. In addition, we assumed that the rate, λ_μ, is itself a random variable, which experiences secondary sources of variability according to

$$\lambda_\mu = \left(1+\sigma_3\eta_\mu^{(3)}\right) g\left(\left(1+\sigma_2\eta_\mu^{(2)}\right)\varphi_\mu(\theta)+\sigma_1\eta_\mu^{(1)}\right), \quad \text{(Equation 2)}$$

where $g(\cdot)$ is the (non-linear) transfer function of the neuron, $\varphi_\mu(\theta)$ is its stimulus-dependent input, the three positive constants (σ_1, σ_2, and σ_3) define the magnitude of each source of noise, and $\eta_\mu^{(1)}, \eta_\mu^{(2)}, \text{ and } \eta_\mu^{(3)}$ are secondary random variables with vanishing means and unit variances (Figure 4A, see also Supplemental Experimental Procedures, Sec. 2.1). Here, the term labeled by (3) (Model III) amounts to the output gain modulation considered previously in the context of cortex (Ecker et al., 2014; Goris et al., 2014), which can be interpreted as gain modulation in the spiking mechanism; the term labeled by (1) (Model I) can be interpreted as stimulus-independent input noise, and the term labeled by (2) (Model II) can be interpreted as stimulus-independent synaptic gain modulation. We assumed that the secondary source of noise, $\eta_\mu^{(i)}$, did not depend upon stimulus but was correlated among neurons (see Supplemental Experimental Procedures, Sec. 2, for a detailed discussion of the models). This phenomenological formulation of the generation of noise correlation encompasses a specific proposal of a mechanistic model (Zylberberg et al., 2016) (see also below).

Even though the noise terms are stimulus independent, stimulus-dependent variances and covariances—and, hence, correlations—emerge in the neural population because of the doubly stochastic nature of the model. The shape of these variances and covariances depend upon the structure of the model. We examined the predictions of each of the three models, with $\sigma_1 \neq 0$ and $\sigma_2 = \sigma_3 = 0$ (Model I), $\sigma_2 \neq 0$ and $\sigma_1 = \sigma_3 = 0$ (Model II), and $\sigma_3 \neq 0$ and $\sigma_1 = \sigma_2 = 0$ (Model III), separately. We found that, in a large fraction of our fits, all three models reproduced closely the shape of the observed tuning functions (Figure 4B) and that of the observed variances and covariances (Figure 4C). Statistically, the variance explained was 60% ± 17% in Models II and III and somewhat lower, 56% ± 16%, in Model I (see also Supplemental Experimental Procedures, Sec. 2.2).

The convergence of the various phenomenological models suggests that, if a doubly stochastic process is at play—and, given the many possible sources of noise in retinal pathways, one indeed expects multi-stochastic processes—then the form of the stimulus dependence of the correlation is robust in that it is not expected to depend strongly upon details of the biological mechanisms involved. Indeed, a similar, doubly stochastic, phenomenological model (Lin et al., 2015) yielded comparable fits to cortical data and generated noise correlations with a fine structure similar to the one we observed in retina. In the case of retina, this fine structure predicted by phenomenological models can be explained also by a particular mechanistic instantiation of correlated noise in the circuit,

based upon intracellular recordings (Zylberberg et al., 2016). The convergence of phenomenological models in retina and cortex, and their mechanistic support found in retina, suggest that they may serve as useful tools for analyzing neural circuits.

Exploiting the Functional Models: Beneficial Effect of Correlation and Harmful Effect of Noise

With simple phenomenological models at hand, we investigated the ways in which the structure of the noise and other parameters affect the coding performance. We first asked how the coding performance in a population of model neurons compared to its counterpart in a population either of neurons with stimulus-independent correlation or of independent neurons. All three models (Models I, II, and III) led to qualitatively similar outcomes, and these outcomes agreed with our data analyses. Specifically, stimulus-independent correlations yielded a modest improvement of the coding accuracy for a wide range of the magnitudes of the noise, but it could also lead to a weak suppression of the coding accuracy in a small-noise regime. By contrast, the effect was strong and always favorable in the case of stimulus-dependent correlations (Figure 5A). That is, input noise (Model I), input gain modulation (Model II), and output gain modulation (Model III) all impose a stimulus-dependent shape to the correlation that is appreciably more beneficial to coding than a mean-matched stimulus-independent correlation.

In general, noise in neural responses is expected to be detrimental to coding accuracy. It is instructive to examine the harmful effect of noise in the context of our models of noise correlation: we compare the coding performance in the presence of one source of secondary noise ($\sigma_i \neq 0$, with $i = 1$, 2, or 3) and that in the absence of secondary noise ($\sigma_1 = \sigma_2 = \sigma_3 = 0$). We find that, as expected, coding accuracy is harmed by any one of the secondary sources of variability in the stimulus-independent case (Figure 5B). However, in the stimulus-dependent case, the secondary source of variability has only a marginal impact upon the coding accuracy, which remains nearly as good as if there were no secondary source of noise whatsoever (Figure 5B). The response of the population is still fluctuating, given the presence of the intrinsic variability, but an additional source of noise added on top of this barely affects coding. Thus, another way to view our results can be summarized as follows: stimulus dependence shapes noise in such a way as to prevent it from harming the coding performance.

The form of Model II suggests an additional, intriguing possibility: if the function $g(\cdot)$ (see Equation 2) is sufficiently non-linear, the coding accuracy can in fact be improved by the presence of secondary noise (Figure 5B). This effect occurs because the magnitude of the noise, σ_2, affects not only the correlation but also the tuning curve; this increase in “signal” can overcompensate the increase in “noise,” itself shaped advantageously as a function of stimulus, to enhance coding.

Why Are Stimulus-Dependent Correlations Beneficial to Coding?

A basic mechanism by which correlation can benefit coding has been discussed by a number of authors (Abbott and Dayan, 1999; Wilke and Eurich, 2002; Averbeck and Lee, 2006; Averbeck et al., 2006; Hu et al., 2014; da Silveira and Berry, 2014; Moreno-Bote et al.,

2014, Lin et al., 2015): in brief, neurons with different preferences ought to be positively correlated. While this mechanism is often illustrated in the case of pairs of neurons, it has also been examined in larger populations (Hu et al., 2014; da Silveira and Berry, 2014; Moreno-Bote et al., 2014; Lin et al., 2015) where it can lead to a quantitatively strong effect. A more general formulation of this mechanism stipulates that the harmful effect of noise is minimized if it lies in a subspace of the response space orthogonal to the "informative subspace," namely, that in which the mean response of the population varies. In the case of direction-selective neurons, the tuning curves define an informative curve in the space of population responses, parametrized by the stimulus value (Figure 6A). Since correlations benefit coding, we expect that these sculpt the noise in such a way that its variations mostly lie perpendicularly to this informative curve. In the case of the pair of cells illustrated in Figure 6B, for example, positive correlations enhance the coding performance while negative correlations suppress it, over a range of stimulus directions. It is tempting to terminate the argument at this stage and to claim that this intuition on coding with pairs of neurons suffices to understand coding in populations of (direction-selective) cells. But the argument would be incomplete: the stimulus dependence of the correlation is a key ingredient, and a complete intuition requires more than considering pairs of neurons separately.

Typically, a tuning curve spans somewhat more than 180°. Consider for a minute the hypothetical, simpler case in which tuning curves would span at most 180°. For a quadruplet of direction-selective cells, each stimulus value, θ, would then be monitored by the two neurons whose preferences lie within 90° of θ and at roughly 90° of each other. Since these two neurons have distinct tuning preferences, constant (stimulus-independent) positive correlation affects coding favorably (compare the left and middle panels in Figure 6C). Now, stimulus dependence enhances this effect because it accumulates the positive correlation in the region that lies in between their preferred directions, while it depletes it near their preferred directions (see Figure 1E). That is, the stimulus-dependent structure boosts the positive correlation where the two tuning curves overlap and, hence, where it enhances the coding performance (compare the middle and right panels in Figure 6C).

This argument breaks down when the stimulus direction is close to the preferred direction of one of the two cells under scrutiny, because there the stimulus-dependent structure depletes the correlation and, hence, suppresses the coding performance of this pair of neurons. Here, it is necessary to consider the simultaneous impact of other pairs of neurons. Indeed, when we consider the actual situation, in which a tuning curve spans somewhat more than 180°, then each stimulus direction is monitored by more than two neurons in the quadruplet. When the stimulus direction is close to the tuning preference of a given neuron, it is also monitored by the two other neurons flanking it (in tuning preference) and whose preferred directions lie at 180° of each other. The stimulus-dependent structure boosts the correlation between those two neurons (Figure 1E), so that correlation now enhances the information in that pair of neurons.

In other words, the impact of correlation upon coding in a given pair of neurons depends upon the stimulus direction; when correlation has the appropriate stimulus-dependent structure, the combined effect of all pairs of neurons is favorable. The geometrical

interpretation is that the stimulus dependence in the correlation reorients the variability so that it extends in a direction orthogonal to the informative curve in the high-dimensional (four-dimensional in the case of a quadruplet) space of the population response (in analogy with the mechanism depicted in Figure 6C). Here, correlated coding is a collective phenomenon, which cannot be fully understood by considering a single pair of neurons. This is true even if only pairwise correlations are considered and higher-order correlations are neglected. We note in passing that, in general, noise correlations can modify, and indeed suppress, the noise entropy, as compared to the independent case; what matters for coding, however, is the magnitude of the noise along the informative curve.

The situation is similar in the case of octuplets of direction-selective neurons (Figure 6D): each stimulus direction is monitored by many cells, and contributions from all pairs have to be taken into account. This insight also sheds light on the different behaviors noted for quadruplets and for octuplets (Figure 3): when the stimulus dependence in the correlation is taken into account (Figure 3A), all quadruplets and octuplets benefit from the presence of noise correlation; when correlation is stimulus independent (Figure 3B), it still benefits nearly all quadruplets (even though the effect is weaker), but it is harmful to an appreciable fraction of octuplets. Because positive correlations between two neurons with similar preference—there are eight such pairs in an octuplet—are detrimental, when correlations are averaged over stimuli, these can overrun the beneficial effect of positive correlation between neurons with distinct preferences.

In order to examine how the interplay between structured correlation and tuning curves affects coding, we focused on a given, arbitrary value of the stimulus, and we considered general variances and covariances in a population of direction-selective neurons. We then investigated theoretically the coding performance in the space of these parameters, by evaluating bounds on the coding error for each choice of correlation values. We simplified the problem by assuming that the form of the correlation can be summarized by three parameters, namely, the correlation coefficients among pairs of cells with preferred directions that differ by 0°, 90°, and 180° (Figure 7A). We computed the percent improvement in coding accuracy as a function of these parameters (see Supplemental Experimental Procedures, Secs. 3.1, 3.2, and 3.3, for derivations); we then used the outcome as a benchmark for the experimentally measured values (Figure 7B). The comparison with data is simplest in the case of quadruplets, as there only two parameters survive, namely the correlation between cells whose preferred directions differ by 90° and 180°. While the measured correlation values are far from the limiting values allowed by mathematical constraints on covariances, and for which very large percent improvements of the coding accuracy occur, a fraction of measured correlation values correspond to an appreciable improvement. An extrapolation of this picture to the case of 12 and 20 neurons (i.e., populations with groups of $m = 3$ or 5 neurons with preferences aligned with each of the four cardinal directions) suggests that the effect of correlation values observed in data can become appreciably stronger already in these small populations (Figure 7B).

Beyond the Retina: Coding with Favorable Correlations in Populations of Broadly Tuned Neurons

The modeling approach is relevant not only to retinal direction-selective neurons, but more generally to populations of correlated neurons that respond to stimuli with broad tuning curves. Because there are many examples of broadly tuned neurons responsive to continuous stimuli, in several brain areas, we decided to conclude our study by examining a more general problem, which extends beyond the constraints imposed by retinal data. We considered a model population with N_0 neurons (Figure 8A). Whereas retinal direction-selective neurons align their preferences with the four cardinal directions, in the more general model we allowed maxima of the tuning curves to take any value of the continuous stimulus; this occurs, for example, for cortical direction-selective cells. It is known that noise correlations among broadly tuned neurons harm coding if the magnitude of the correlation in pairs of neurons with similar preferences exceeds that in pairs of neurons with different preferences (Abbott and Dayan, 1999; Sompolinsky et al., 2001; Wilke and Eurich, 2002). By contrast, in populations of retinal direction-selective cells, where correlation improved the coding performance, stimulus dependence sculpted the correlation such that, in a large fraction of the examples studied and for a range of stimuli, it was greater in pairs of neurons with different preferences than in pairs of neurons with similar preferences (Figure 8B). We wanted to encompass this possibility in our general model, so we allowed for forms of the correlation in which neurons with distinct preferences could be slightly more correlated than neurons with similar preferences (Figure 8C) (Wilke and Eurich, 2002); we called $\theta_{\max}$ the difference in preference for which pairs of neurons reached the maximum value of the correlation. We then asked how the coding performance varied with the population size and how the arrangement of the neural preferences in the population influenced it. To explore different arrangements, we allowed the population to break up in N pools, with any $N < N_0$, while keeping the total number of neurons in the population fixed (Figure 8A).

We found that the coding accuracy can be enhanced appreciably as a function of the number of neurons in the population (Figure 8D; see Supplemental Experimental Procedures, Sec. 3.4, for derivations). The optimal arrangement of the preferred directions, which led to this coding enhancement, was interesting (Figures 8E and 8F): initially, as we increased the number of neurons in the population, their preferred directions were distributed uniformly over all directions; once the distance (in stimulus space) between neighboring preferred directions became comparable to $\theta_{\max}$, then neurons clumped into discrete pools as N_0 was increased. In this regime, as N_0 increased, the total number of pools settled to a constant, while each pool grew in size (Figures 8E and 8F). The presence of correlation leads to a percent improvement of the coding accuracy, ΔR, which can be expressed as

$$\Delta R = \bar{c}_{\max}\left[(N_0/N)\, G(N, \theta_{max}) - (N_0/N - 1)\,\bar{c}_0\right] \times 100, \quad \text{(Equation 3)}$$

where $\bar{c}_{\max}$ is proportional to the maximum value taken by the correlation and $\bar{c}_0$ is the correlation of two neurons with identical preference. The function $G(N; \theta_{\max})$ is controlled by the "correlation length," $\theta_{\max}$, and becomes large when the number of pools, N, is such that the distance (in stimulus space) between successive pools becomes comparable to $\theta_{\max}$,

i.e., when $2\pi/N \sim \theta_{max}$ (see Supplemental Experimental Procedures, Sec. 3.4, for details). This arrangement ensures that positive correlations are concentrated in such a way as to maximize their favorable effect upon the coding accuracy. These observations imply that, even for coding a continuous stimulus, it may be advantageous for a neural population to break up in discrete pools with distinct preferences (as illustrated in Figure 8A), depending upon the structure of correlation.

Discussion

In this study, we have characterized the structure of noise correlation in populations of direction-selective retinal ganglion cells and shown that noise correlation benefits the accuracy of coding appreciably, as compared with the case of independent neurons. The stimulus dependence of the noise correlation was key, and coding improvement reflected a collective phenomenon operating beyond individual pairs. In order to examine this phenomenon, we introduced functional models of direction-selective cells, which captured the statistics of responses to moving bars. We then used these models to investigate the performance of population coding. The generality of the functional models suggests that similar phenomena may occur in other brain areas, and we extended theoretical considerations to large populations of broadly tuned neurons.

The effect of stimulus-dependent noise correlations has been examined in a number of earlier studies (Panzeri et al., 1999; Pola et al., 2003; Shamir and Sompolinsky, 2004). These, however, differ from our work in their information theoretic framework, the models of noise correlation they assume, or, more importantly, the conclusions they reach (but see Lin et al., 2015, which we discuss below). Panzeri and colleagues (Panzeri et al., 1999; Pola et al., 2003) break up the mutual information between stimulus and neural output in a sum of terms that elucidate the various contributions of noise correlations to the mutual information. In analyses of the activity of pairs of neurons in rat S1 and of multi-unit recordings in monkey area MT in the context of discrimination tasks, they show that stimulus-dependent correlation can (marginally) enhance the mutual information and, hence, benefit neural coding. However, because their approach consists of a direct data analysis without the use of an encoding or noise model, it is not possible to address the question of which aspect of the structure of noise correlations is beneficial. A more drastic disconnect between their conclusions and ours pertains to the mechanism by which correlation improves benefits coding. There, coding improves because information about the stimulus is stored in the magnitudes of the pairwise correlations themselves, independently of the tuning properties of the cells. Similar remarks apply to a study by Shamir and Sompolinsky (Shamir and Sompolinsky, 2004). In their purely theoretical investigation, they consider two alternative noise models, one of which features stimulus-dependent correlations. However, the form of the noise correlation is different from the one we measured and modeled. On a conceptual level, again the mechanism put forth by the authors for the beneficial effect of correlation amounts to storing information about stimuli in the values of the correlations, as in Panzeri et al., 1999; Pola et al., 2003.

Our work comes to complement these earlier studies. In our case, the coding improvement due to correlations follows from a basic and often discussed (Abbott and Dayan, 1999;

Wilke and Eurich, 2002; Romo et al., 2003; Averbeck and Lee, 2006; Averbeck et al., 2006; Josić et al., 2009; Hu et al., 2014; da Silveira and Berry, 2014) mechanism: for a given stimulus, positive correlation among neurons with different preferences is advantageous. But when a possible stimulus dependence of correlations is taken into account, as in our case, the problem becomes richer: this stimulus dependence can act in concert with the basic mechanism by enhancing correlations where they are useful and suppressing them where they are not. We demonstrated, quantitatively, that the corresponding boost can be substantial even in small populations with less than ten neurons, and grows with population size. The increase of the performance improvement as a function of the number of neurons in small populations has also been noted, in a different context (in Jeanne et al., 2013). Our conclusions can also be stated differently, namely, that stimulus-dependent correlation shapes the noise so as to prevent its deleterious effect on coding. This requires a specific shape of the stimulus-dependent correlation, which was present in the data.

We also developed simple functional encoding models of direction-selective neurons. These models are “functional” because they are not based upon detailed mechanistic knowledge of the circuitry and physiology, but rather summarize these in a few mathematical features. Furthermore, a full-fledged model would convert a spatio-temporal input into a neural output, and one would aim for it to be valid for the broadest possible ensemble of spatio-temporal inputs. We were concerned with a more modest formulation of models, which converted stimulus features into neural output. (We did not aim at modeling the first stage of processing, from the general spatio-temporal input into stimulus features.) Constructing a more general model would require recordings of direction-selective cells in response to a broader ensemble of visual stimuli. In our models, the favorable correlation structure, matching that in data, came about as a consequence of the presence of non-linear transfer functions and the doubly stochastic nature of the neural response. As such, favorable correlation structures are expected to be insensitive to biophysical and connectivity details: when noise trickles down a circuit, favorable stimulus-dependent correlations emerge in a robust manner.

We examined various possible sources of noise in our modeling approach, including additive input noise as well as input and output gain fluctuations. Similar phenomenological models, with both additive and multiplicative noise, were recently fitted also to cortical data (Ecker et al., 2014; Goris et al., 2014; Lin et al., 2015). These models can then be used to investigate the relation between noise and its correlation structure on the one hand, and coding of information on the other hand. As it turns out, both the structure of noise correlation and its impact upon coding shared a large degree of similarity between the case of retina, discussed here (see also Zylberberg et al., 2016), and that of cortex (Ecker et al., 2014; Goris et al., 2014; Lin et al., 2015). Our theoretical work complements and extends the analysis of cortical coding (Lin et al., 2015); in particular, we point out that not only is the structure of correlation relevant, but also the arrangement of tuning curves in the population, and that the two interact. In models of cortex, it is customary to assume a uniform coverage of tuning preferences in the population; even a small modulation away from uniformity can impact coding appreciably in the presence of correlation. Beyond such quantitative considerations, the convergence between findings in retina and cortex suggests

that the phenomenology reported is a general one and that functional models can serve as a broadly useful tool for its investigation.

Both in retina and in cortex (Ecker et al., 2014; Goris et al., 2014; Lin et al., 2015), it was necessary to assume the presence of multiplicative noise for improved fits to the spiking data. The requirement of multiplicative noise is interesting, because gain fluctuation is a widespread phenomenon in neural systems and is likely involved in many different computations. In particular, a number of adaptation phenomena result from the variation of some gain in the processing stream. Thus, one can expect gain fluctuations to occur naturally in a circuit where elements adapt dynamically as a function of stimulus history. A similar point has been made about gain fluctuations induced by top-down attentional signals (Lin et al., 2015). It is thus advantageous that those correlations that emerge from common gain fluctuation come with a form of stimulus dependence that prevents the noise from impairing the coding accuracy.

The presence of different tuning curves and of stimulus-dependent correlations in a population invests it with a higher degree of heterogeneity. Other forms of heterogeneity than the ones considered here can also enhance the coding performance (Wilke and Eurich, 2002; Shamir and Sompolinsky, 2006; Ecker et al., 2011; da Silveira and Berry, 2014). In particular, heterogeneity in the magnitudes of the tuning curves has been shown to lead to an enhancement of the coding performance (Ecker et al., 2011), and this mechanism may also play a role in retina. When viewed in the context of linear algebra, the beneficial role of heterogeneity is expected to be generic: greater variability in the correlations generically results in smaller eigenvalues of the covariance matrix, so that neural activity in response to a given stimulus is confined to a narrower region in the space of population responses.

Finally, when we turned the problem on its head and asked, “Given a correlation structure, how should the tuning curves be arranged in a population?”, we found that it could be advantageous for a neural population to clump into pools with discrete preferences, even when coding a continuous stimulus. Similar forms of “discrete coding” were put forth by a number of authors (Harper and McAlpine, 2004; Nikitin et al., 2009; Sharpee, 2014; Kastner et al., 2015). In their cases, however, discreteness followed from the maximization of an information-theoretic quantity calculated for a population of independent neurons, whereas in our case the discreteness is related to the correlation structure. Our result was derived using a relatively general model of correlated neurons, which extends beyond the confines of retina. If beneficial structures of stimulus-dependent correlation emerge generically from noise trickling down neural circuits, then correlation may be the quantity informing the way in which response properties should be arranged for high-performance coding.

Experimental Procedures

All animal experiments and procedures were approved by the Swiss Veterinary Office. New Zealand female albino rabbits (females, 2.5–3 kg) were obtained from Charles River Laboratories. The eyes were dissected under dim red light conditions in Ames solution (Sigma, A1420) continuously equilibrated with 5% CO_2, 95% O_2. The vitreous was removed, and a patch of retina (2×2 mm^2) was isolated between 1 mm and 4 mm

eccentricity, along the ventral direction. The retina patch was placed, ganglion cell side down, on the high-density multielectrode array (HDMEA) (Frey et al., 2009; Fiscella et al., 2012) featuring 11,011 platinum electrodes with diameters of 7 μm and electrode center-to-center distances of 18 μm over an area of 2 × 1.75 mm^2 (3,161 electrodes/mm^2) and perfused by Ames solution (pH 7.4, 36° C), equilibrated with 5% CO_2, 95% O_2. The retina was optically stimulated (projector Acer K10 with 60 Hz refreshing rate using blue [460 ± 15 nm] and green [523 ± 23 nm] projector LEDs) with white noise and flashing square stimuli to map receptive fields and ON-OFF responses. Direction selectivity was measured with moving bar stimuli (1 mm length, background irradiance 0.2 μW/cm^2, bar stimulus irradiance 2.2. μW/cm^2, constant velocity of 1.6 mm/s along 36 equidistant angular directions radially spaced at 10°). Two different bar widths were used, 0.5 mm and 1 mm, respectively.

Of the 11,000 electrodes, 126 electrodes can be recorded with a sampling frequency of 20 kHz at the same time and arbitrarily selected during the experiment. To place the electrodes preferentially below direction-selective cells, we scanned the electrode array with several non-overlapping electrode configurations while stimulating the retina with moving bar stimuli. The data were spikesorted manually using the software UltraMegaSort2000 (Hill et al., 2011) on 5–7 electrodes at a time, online during the experiment. After this initialization period, 5–7 electrodes were placed close to each identified direction-selective cell, and the electrode configuration was fixed for the rest of the experiment. One hundred repetitions of each of the 36 moving bar stimuli were subsequently projected onto the retina in a randomized order. The resulting data were spikesorted in the same fashion as before, but offline, after the experiment. This yielded 6–18 (mean 12.8 ± 4.5) well-isolated (fraction of inter-spike intervals shorter than 1.5 ms below 2%) ON-OFF direction-selective ganglion cells per experiment (total of 90 cells in seven experiments). Spike counts were defined as the total number of spikes of each neuron during the 1.6-s-long stimulus presentations. The range of the recorded outputs observed from experiment to experiment (Figure 2) may be explained, at least in part, by the variations in the retinal eccentricity of the patch of retina we recorded from.

Supplemental Information

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the CNRS through UMR 8559, SNSF Sinergia Project CRSII3_141801, the ERC Advanced Grant “NeuroCMOS” under contract No. AdG 267351, the French government through a *Bourse de mobilité* No. 806711B, and the Swiss SystemsX interdisciplinary PhD grant No. 2009_031.

References

Abbott LF, Dayan P. The effect of correlated variability on the accuracy of a population code. Neural Comput. 1999; 11:91–101. [PubMed: 9950724]

Averbeck BB, Lee D. Neural noise and movement-related codes in the macaque supplementary motor area. J Neurosci. 2003; 23:7630–7641. [PubMed: 12930802]

Averbeck BB, Lee D. Coding and transmission of information by neural ensembles. Trends Neurosci. 2004; 27:225–230. [PubMed: 15046882]

Averbeck BB, Lee D. Effects of noise correlations on information encoding and decoding. J Neurophysiol. 2006; 95:3633–3644. [PubMed: 16554512]

Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. Nat Rev Neurosci. 2006; 7:358–366. [PubMed: 16760916]

Bair W, Zohary E, Newsome WT. Correlated firing in macaque visual area MT: time scales and relationship to behavior. J Neurosci. 2001; 21:1676–1697. [PubMed: 11222658]

Cohen MR, Newsome WT. Context-dependent changes in functional circuitry in visual area MT. Neuron. 2008; 60:162–173. [PubMed: 18940596]

da Silveira RA, Berry MJ 2nd. High-fidelity coding with correlated neurons. PLoS Comput Biol. 2014; 10:e1003970. [PubMed: 25412463]

Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, Tolias AS. Decorrelated neuronal firing in cortical microcircuits. Science. 2010; 327:584–587. [PubMed: 20110506]

Ecker AS, Berens P, Tolias AS, Bethge M. The effect of noise correlations in populations of diversely tuned neurons. J Neurosci. 2011; 31:14272–14283. [PubMed: 21976512]

Ecker AS, Berens P, Cotton RJ, Subramaniyan M, Denfield GH, Cadwell CR, Smirnakis SM, Bethge M, Tolias AS. State dependence of noise correlations in macaque primary visual cortex. Neuron. 2014; 82:235–248. [PubMed: 24698278]

Fiscella M, Farrow K, Jones IL, Jäckel D, Müller J, Frey U, Bakkum DJ, Hantz P, Roska B, Hierlemann A. Recording from defined populations of retinal ganglion cells using a high-density CMOS-integrated microelectrode array with real-time switchable electrode selection. J Neurosci Methods. 2012; 211:103–113. [PubMed: 22939921]

Fiscella M, Franke F, Farrow K, Müller J, Roska B, da Silveira RA, Hierlemann A. Visual coding with a population of direction-selective neurons. J Neurophysiol. 2015; 114:2485–2499. [PubMed: 26289471]

Fiser J, Chiu C, Weliky M. Small modulation of ongoing cortical dynamics by sensory input during natural vision. Nature. 2004; 431:573–578. [PubMed: 15457262]

Frey U, Egert U, Heer F, Hafizovic S, Hierlemann A. Microelectronic system for high-resolution mapping of extracellular electric fields applied to brain slices. Biosens Bioelectron. 2009; 24:2191–2198. [PubMed: 19157842]

Golledge HD, Panzeri S, Zheng F, Pola G, Scannell JW, Giannikopoulos DV, Mason RJ, Tovée MJ, Young MP. Correlations, feature-binding and population coding in primary visual cortex. Neuroreport. 2003; 14:1045–1050. [PubMed: 12802200]

Goris RL, Movshon JA, Simoncelli EP. Partitioning neuronal variability. Nat Neurosci. 2014; 17:858–865. [PubMed: 24777419]

Graf AB, Kohn A, Jazayeri M, Movshon JA. Decoding the activity of neuronal populations in macaque primary visual cortex. Nat Neurosci. 2011; 14:239–245. [PubMed: 21217762]

Gutnisky DA, Dragoi V. Adaptive coding of visual information in neural populations. Nature. 2008; 452:220–224. [PubMed: 18337822]

Harper NS, McAlpine D. Optimal neural population coding of an auditory spatial cue. Nature. 2004; 430:682–686. [PubMed: 15295602]

Hatsopoulos NG, Ojakangas CL, Paninski L, Donoghue JP. Information about movement direction obtained from synchronous activity of motor cortical neurons. Proc Natl Acad Sci USA. 1998; 95:15706–15711. [PubMed: 9861034]

Hill DN, Mehta SB, Kleinfeld D. Quality metrics to accompany spike sorting of extracellular signals. J Neurosci. 2011; 31:8699–8705. [PubMed: 21677152]

Hu Y, Zylberberg J, Shea-Brown E. The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. PLoS Comput Biol. 2014; 10:e1003469. [PubMed: 24586128]

Jeanne JM, Sharpee TO, Gentner TQ. Associative learning enhances population coding by inverting interneuronal correlation patterns. Neuron. 2013; 78:352–363. [PubMed: 23622067]

Johnson KO. Sensory discrimination: decision process. J Neurophysiol. 1980; 43:1771–1792. [PubMed: 6251182]

Josić K, Shea-Brown E, Doiron B, de la Rocha J. Stimulus-dependent correlations and population codes. Neural Comput. 2009; 21:2774–2804. [PubMed: 19635014]

Kastner DB, Baccus SA, Sharpee TO. Critical and maximally informative encoding between neural populations in the retina. Proc Natl Acad Sci USA. 2015; 112:2533–2538. [PubMed: 25675497]

Kohn A, Smith MA. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J Neurosci. 2005; 25:3661–3673. [PubMed: 15814797]

Lee D, Port NL, Kruse W, Georgopoulos AP. Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. J Neurosci. 1998; 18:1161–1170. [PubMed: 9437036]

Lin I-C, Okun M, Carandini M, Harris KD. The nature of shared cortical variability. Neuron. 2015; 87:644–656. [PubMed: 26212710]

Mastronarde DN. Correlated firing of retinal ganglion cells. Trends Neurosci. 1989; 12:75–80. [PubMed: 2469215]

Maynard EM, Hatsopoulos NG, Ojakangas CL, Acuna BD, Sanes JN, Normann RA, Donoghue JP. Neuronal interactions improve cortical population coding of movement direction. J Neurosci. 1999; 19:8083–8093. [PubMed: 10479708]

Montani F, Kohn A, Smith MA, Schultz SR. The role of correlations in direction and contrast coding in the primary visual cortex. J Neurosci. 2007; 27:2338–2348. [PubMed: 17329431]

Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. Nat Neurosci. 2014; 17:1410–1417. [PubMed: 25195105]

Nikitin AP, Stocks NG, Morse RP, McDonnell MD. Neural population coding is optimized by discrete tuning curves. Phys Rev Lett. 2009; 103:138101. [PubMed: 19905542]

Nirenberg S, Carcieri SM, Jacobs AL, Latham PE. Retinal ganglion cells act largely as independent encoders. Nature. 2001; 411:698–701. [PubMed: 11395773]

Oram MW, Földiák P, Perrett DI, Sengpiel F. The 'Ideal Homunculus': decoding neural population signals. Trends Neurosci. 1998; 21:259–265. [PubMed: 9641539]

Ozden I, Lee HM, Sullivan MR, Wang SS. Identification and clustering of event patterns from in vivo multiphoton optical recordings of neuronal ensembles. J Neurophysiol. 2008; 100:495–503. [PubMed: 18497355]

Panzeri S, Treves A, Schultz S, Rolls ET. On decoding the responses of a population of neurons from short time windows. Neural Comput. 1999; 11:1553–1577. [PubMed: 10490938]

Perkel DH, Gerstein GL, Moore GP. Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. Biophys J. 1967; 7:419–440. [PubMed: 4292792]

Pola G, Thiele A, Hoffmann KP, Panzeri S. An exact method to quantify the information transmitted by different mechanisms of correlational coding. Network. 2003; 14:35–60. [PubMed: 12613551]

Poort J, Roelfsema PR. Noise correlations have little influence on the coding of selective attention in area V1. Cereb Cortex. 2009; 19:543–553. [PubMed: 18552357]

Romo R, Hernández A, Zainos A, Salinas E. Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. Neuron. 2003; 38:649–657. [PubMed: 12765615]

Sasaki K, Bower JM, Llinás R. Multiple purkinje cell recording in rodent cerebellar cortex. Eur J Neurosci. 1989; 1:572–586. [PubMed: 12106116]

Schneidman E, Bialek W, Berry MJ 2nd. Synergy, redundancy, and independence in population codes. J Neurosci. 2003; 23:11539–11553. [PubMed: 14684857]

Schnitzer MJ, Meister M. Multineuronal firing patterns in the signal from eye to brain. Neuron. 2003; 37:499–511. [PubMed: 12575956]

Shamir M, Sompolinsky H. Nonlinear population codes. Neural Comput. 2004; 16:1105–1136. [PubMed: 15130244]

Shamir M, Sompolinsky H. Implications of neuronal diversity on population coding. Neural Comput. 2006; 18:1951–1986. [PubMed: 16771659]

Sharpee TO. Toward functional classification of neuronal types. Neuron. 2014; 83:1329–1334. [PubMed: 25233315]

Shlens J, Rieke F, Chichilnisky E. Synchronized firing in the retina. Curr Opin Neurobiol. 2008; 18:396–402. [PubMed: 18832034]

Smith MA, Kohn A. Spatial and temporal scales of neuronal correlation in primary visual cortex. J Neurosci. 2008; 28:12591–12603. [PubMed: 19036953]

Sompolinsky H, Yoon H, Kang K, Shamir M. Population coding in neuronal systems with correlated noise. Phys Rev E Stat Nonlin Soft Matter Phys. 2001; 64:051904. [PubMed: 11735965]

Usrey WM, Reid RC. Synchronous activity in the visual system. Annu Rev Physiol. 1999; 61:435–456. [PubMed: 10099696]

Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A. Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature. 1995; 373:515–518. [PubMed: 7845462]

Vogels R. Population coding of stimulus orientation by striate cortical cells. Biol Cybern. 1990; 64:25–31. [PubMed: 2285759]

Wilke SD, Eurich CW. Representational accuracy of stochastic neural populations. Neural Comput. 2002; 14:155–189. [PubMed: 11747537]

Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature. 1994; 370:140–143. [PubMed: 8022482]

Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F. Direction-selective circuits shape noise to ensure a precise population code. Neuron. 2016; 89(this issue):369–383. [PubMed: 26796691]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Highlights

- Neural noise correlations have stimulus-dependent structure in retina and cortex
- Analysis of neural data demonstrates that this structure improves visual coding
- Response models capture the emergence of this structure from circuit properties
- Neural coding models indicate that it can affect coding strategies qualitatively

Europe PMC Funders Author Manuscripts

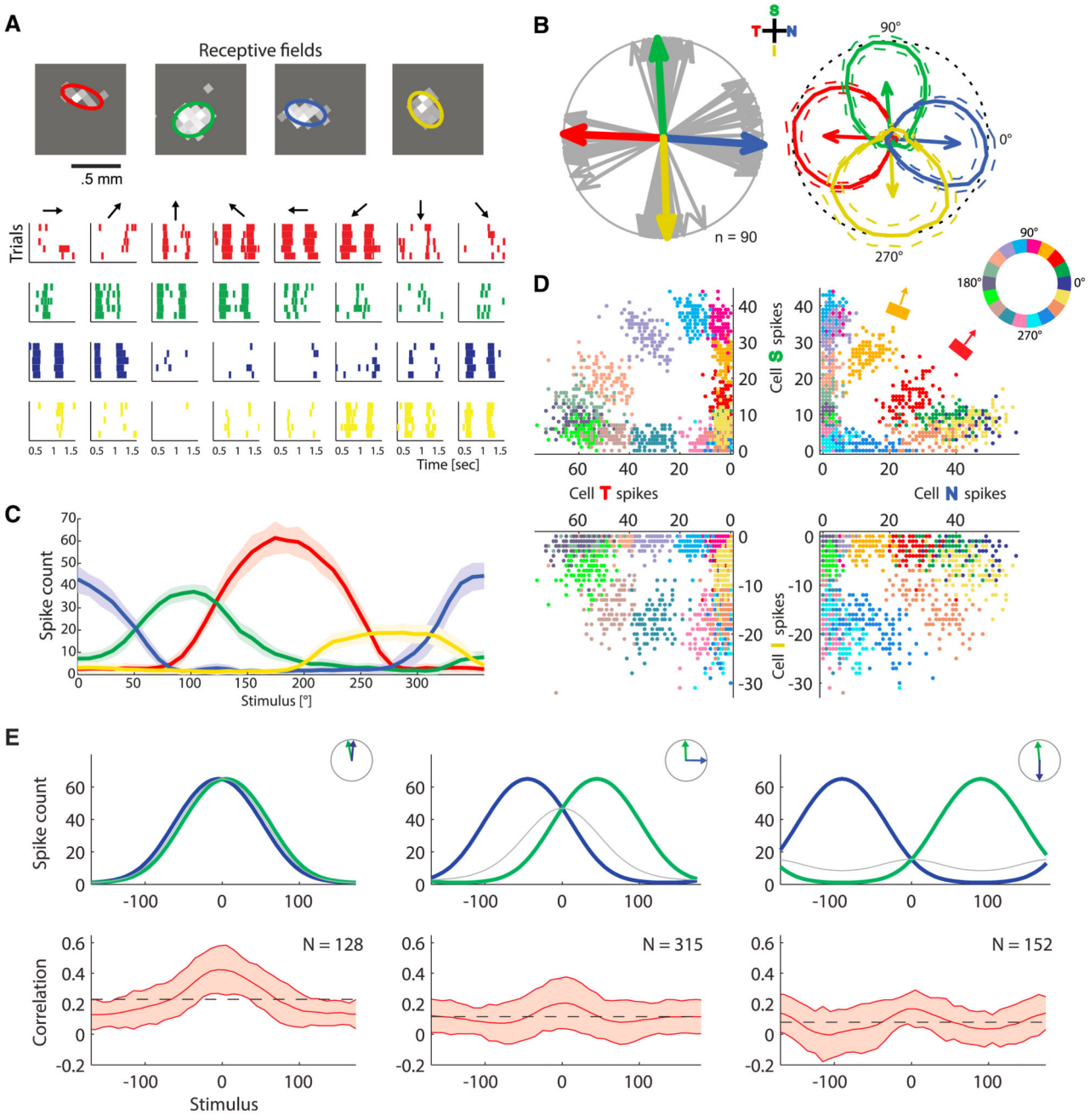

Figure 1. Spiking Responses, Tuning Curves, and Noise Statistics in Retinal Direction-Selective Neurons

(A) Spike raster of a “quadruplet” of neurons, four direction-selective retinal ganglion cells with preferred directions aligned along the four cardinal directions (temporal, red; superior, green; nasal, blue; inferior, yellow). For each stimulus direction (black arrows), five trials of spiking responses are displayed. The top panel exhibits the receptive fields of this example quadruplet, as obtained from white noise stimulation and reverse correlation; the ellipses illustrate the shapes of the receptive fields.

(B) The preferred directions of all 90 recorded direction-selective cells (gray arrows), from seven experiments, point along the four cardinal directions (S, superior; I, inferior; T, temporal; N, nasal; same color coding as in A). One example quadruplet of neurons is indicated by arrows in colors, and their tuning curves are represented in polar plots (solid lines in the top right panel, peaks normalized to 1, dotted lines show the mean ± SD).
(C) Same tuning curves as in (B), in a Cartesian plot and unnormalized.
(D) Spike counts from the example quadruplet in (A), (B), and (C) in the presence of moving bar stimuli: each quadrant exhibits the responses of two cells whose preferred directions differ by 90°; one point per trial. Although stimulus directions were separated by 10° in experiments, here, for the sake of clarity, we display every other instance (stimulus directions separated by 20°). Colors label the direction of motion of the bar in the stimulus, as illustrated in the top right quadrant: the orange bar and red bar illustrate the stimulus and the corresponding arrow its direction; the color wheel indicates the directions of motion corresponding to each choice of color.
(E) Schematic tuning curves and measured noise correlations in pairs of simultaneously recorded cells as a function of the direction of motion in the stimulus, for two neurons with preferred directions that differ by 0° (left panel), 90° (middle panel), and 180° (right panel). The top panels present schematic tuning curves (in Cartesian plots, blue and green curves) for these three configurations (illustrated in the insets), as well as the geometric means of these two quantities (thin, gray curves). The bottom panels display the measured correlations, where the mean correlations are plotted as solid lines and the standard deviations appears as shaded areas.

Europe PMC Funders Author Manuscripts

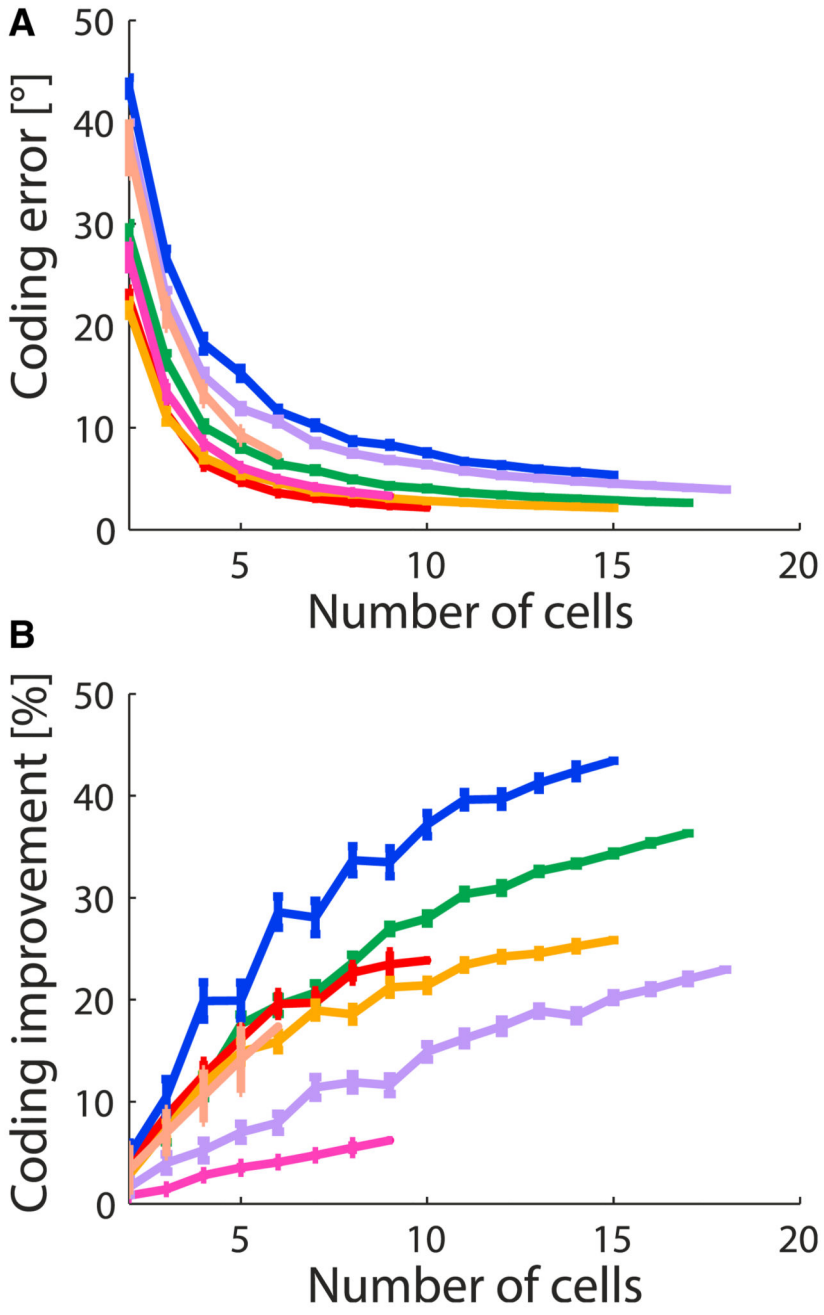

Figure 2. Improvement of the Coding Performance by Correlations in Populations of Retinal Direction-Selective Neurons

(A and B) The coding error (in angular degrees, A) and the percent improvement in resolution due to noise correlations (B), as functions of the number of simultaneously recorded neurons. Different colors correspond to different experiments; each point represents an average over up to 50 distinct groups of cells selected randomly within each experiment; error bars indicate standard errors of the mean.

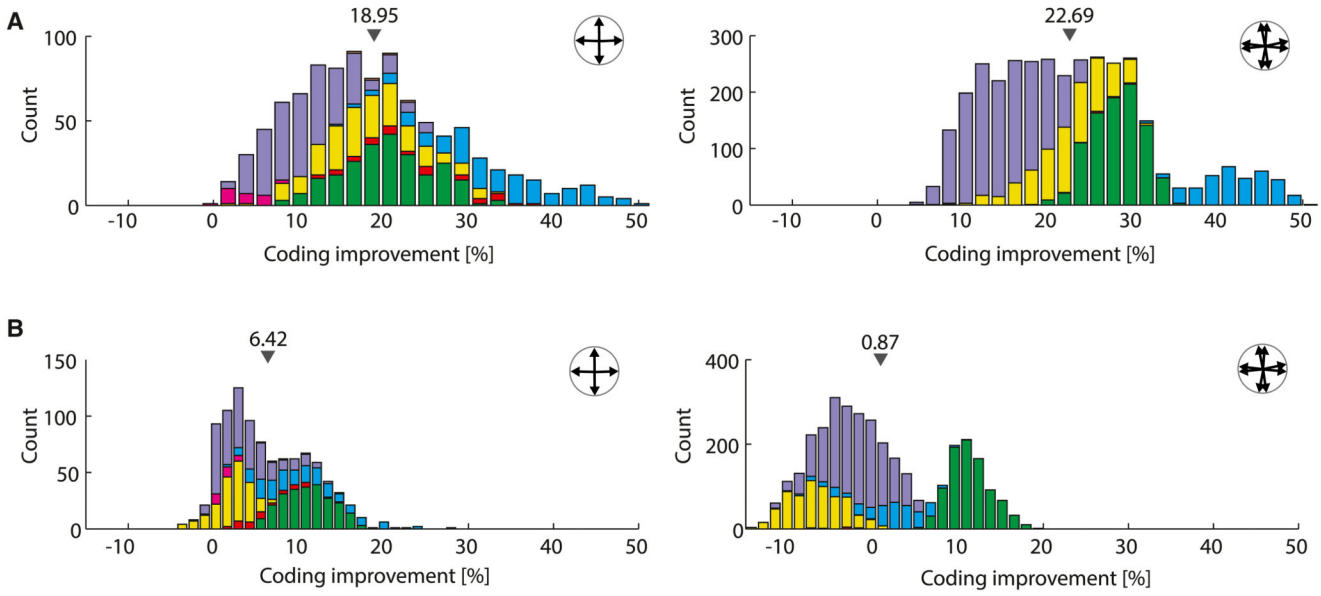

Figure 3. Improvement of the Coding Performance in Quadruplets and Octuplets, with Stimulus-Dependent versus Stimulus-Independent Correlations

(A) Histograms of percent improvement of the coding resolution from the full data, in which correlations depend upon the stimulus, for quadruplets (left panel) and octuplets (right panel) recorded simultaneously. Small gray triangles indicate the means of these empirical distributions, quoted above them. Different colors correspond to different experiments.

(B) Same as (A) but with histograms derived in the case in which noise correlations in the data were replaced by their average over all stimuli and hence became stimulus independent. The effect of the stimulus dependence of the correlations on coding can be seen by comparing (A) and (B).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

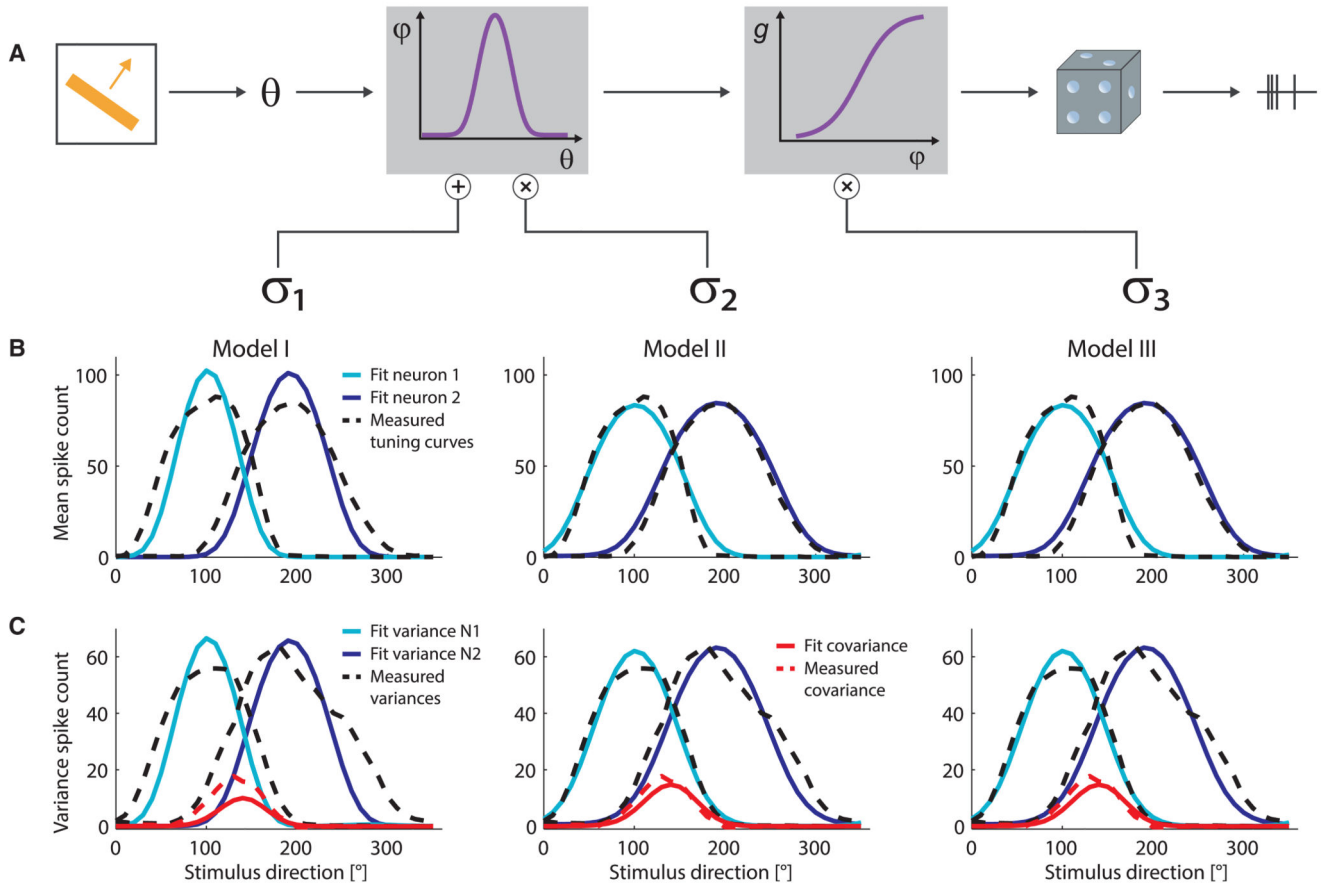

Figure 4. Functional Models of Response and Noise Statistics of Direction-Selective Neurons

(A) Illustration of the models defined in Equation 2: the variability of neural response and the correlation among the responses of neurons can originate from additive input noise (Model I), input gain modulation (Model II), or output gain modulation (Model III).

(B) Examples of model fits of the tuning curves of two neurons (model, solid curves; data, dashed curves).

(C) Examples of model fits of the variances and covariance in the activity of the same two neurons as in (B) (model, solid curves; data, dashed curves).

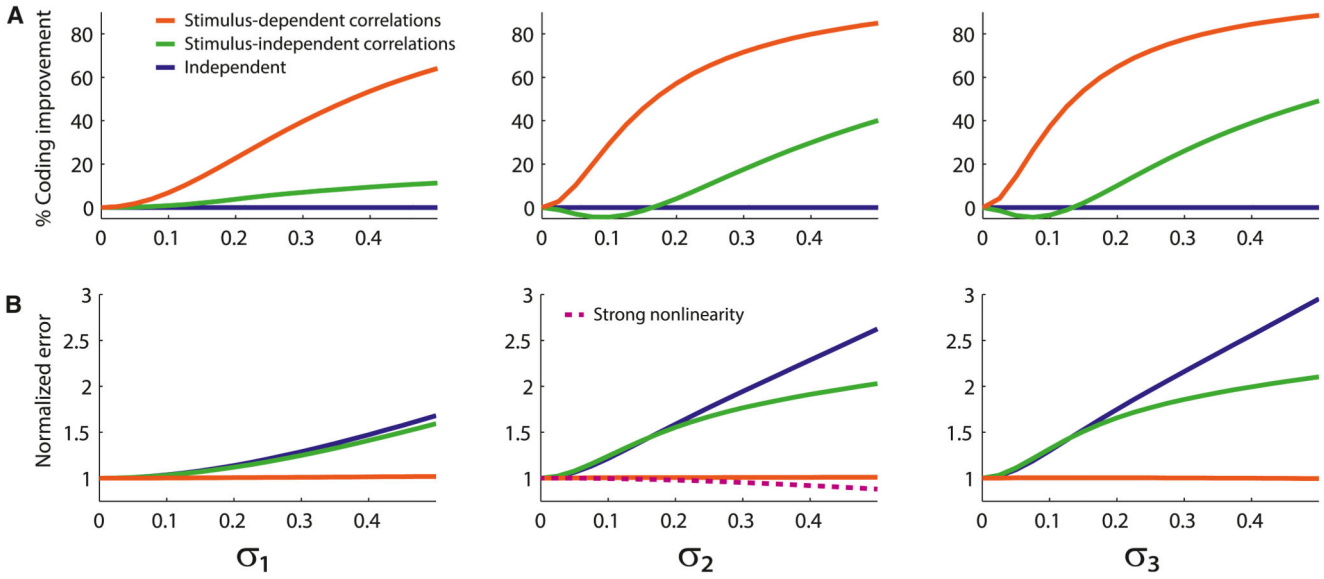

Figure 5. Behavior of the Coding Error and the Coding Improvement as a Function of Correlated Noise

(A) Percent improvement in the accuracy of direction coding as a function of the magnitude of the secondary noise, in Model I (left panel), Model II (middle panel), and Model III (right panel), for the stimulus-dependent correlations generated by the models (red curves), as compared to the mean-matched stimulus-independent case (green curves). For independent neurons (blue curves), the percent improvement is constant and vanishing.

(B) Deterioration of the coding accuracy by the secondary sources of noise, as described by the error normalized by the error in the noiseless case with $\sigma_1 = \sigma_2 = \sigma_3 = 0$. In the presence of the stimulus-dependent correlations generated by Model I (left panel), Model II (middle panel), and Model III (right panel), the coding accuracy is largely insensitive to the secondary noise (red curves), whereas in the cases of stimulus-independent correlation (green curves) and independent neurons (blue curves), noise is detrimental to the coding accuracy. For a sufficiently strong non-linearity, g, input-gain modulation can improve the coding accuracy (dashed line, middle panel).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

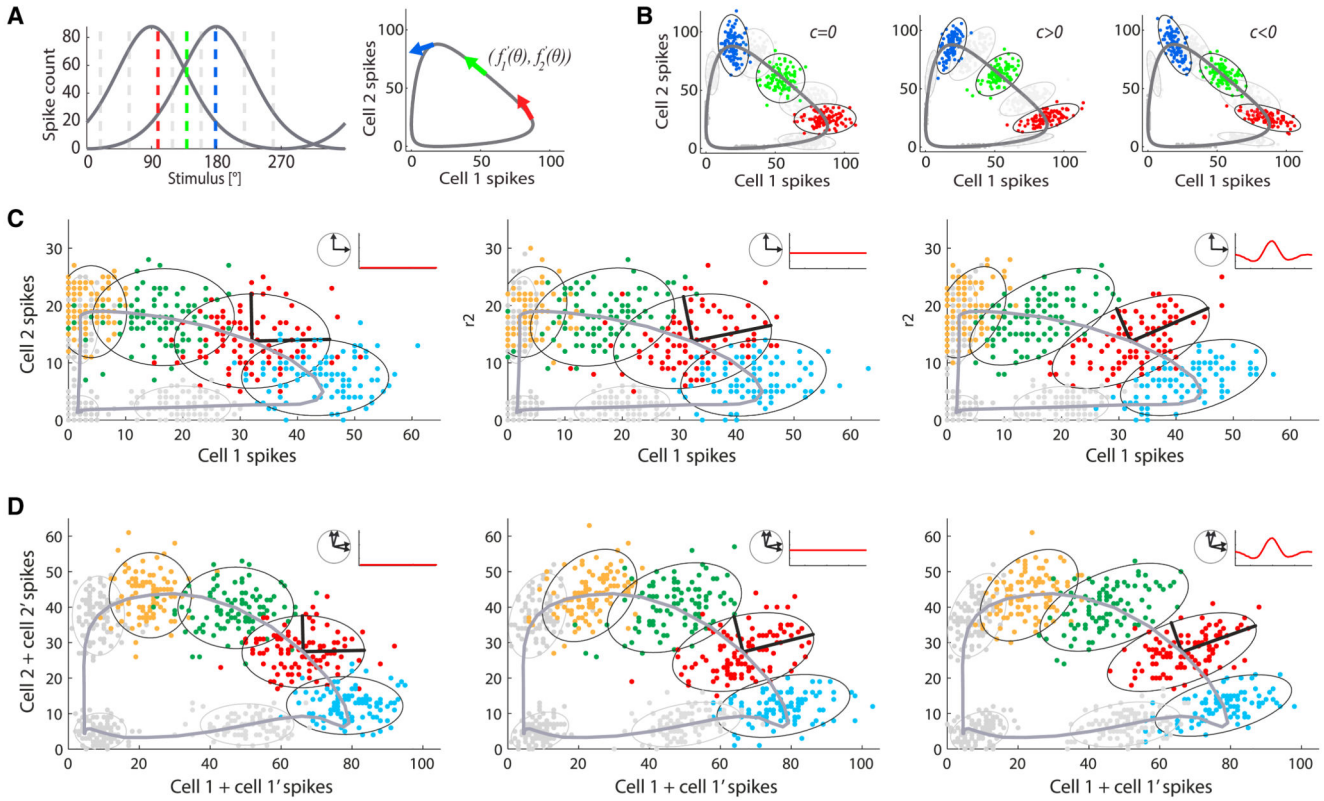

Figure 6. Impact of Different Structures of Noise Correlation upon Population Coding

(A) Left panel: two model direction-selective neurons respond to different stimuli (dashed lines) according to tuning curves (solid gray curves), $f_1(\theta)$ and $f_2(\theta)$, with two direction preferences that differ by 90°. Right panel: the two tuning curves are represented as a solid gray line parametrized by the stimulus direction, θ. In the space of the two-cell output, this gray line forms an informative subspace: the location of the pair response along the gray line yields information about the stimulus presented. More precisely, for each stimulus, θ, the tangent vector, $(f_1'(\theta), f_2'(\theta))$, defines the informative direction (arrows in colors corresponding to the stimulus values in the left panel).

(B) For each stimulus presented, noise correlation distorts the cloud of two-cell responses about the mean over trials; depending upon the geometry of this distortion with respect to the informative direction, it can either benefit or harm the coding accuracy. Positive correlation in the pair ($c > 0$) favors the reliability of coding with respect to the independent case ($c = 0$), while negative correlation ($c < 0$) is detrimental. Specifically, when $c > 0$, the responses for nearby stimulus directions overlap less, and, hence, coding is more reliable. (Conversely, if the two tuning curves have similar preference, $c < 0$ is favorable whereas $c > 0$ is detrimental; illustration not shown.) More precisely, coding is favored if the eigenvector of the covariance matrix parallel to the tangent vector, $(f_1'(\theta), f_2'(\theta))$, comes with a small eigenvalue: correlation then relegates the noise in the orthogonal, uninformative direction. Ellipses are contours of equal probability, drawn at 2.5 standard deviations.

(C) Responses of a pair of neurons with 90° difference in preferred direction; each point is a trial, colors represent different stimuli (direction of moving bar). Data were simulated using

the measured tuning functions and covariance matrices. Three different scenarios are visualized: two independent neurons (left panel), two neurons with stimulus-independent correlations (middle panel), two neurons with the full stimulus-dependent correlation present in data (right panel); ellipses are contours of equal probability for each stimulus, at 2 standard deviations; thick black lines illustrate eigenvectors of the covariance matrix (for one example stimulus, in red). The ellipses are reoriented by noise correlations so that their long axis moves away from the informative direction along which stimuli are coded (thick gray line, given by the tuning functions). Insets: arrangements of the preferred directions and sketches of the dependence of the correlation upon stimulus.

(D) Same as (C) but for two pairs of neurons, corresponding to half of an octuplet. For the sake of illustration, spike counts of neurons with similar preferred directions were summed. The effects are similar to (C), and stronger.

Europe PMC Funders Author Manuscripts

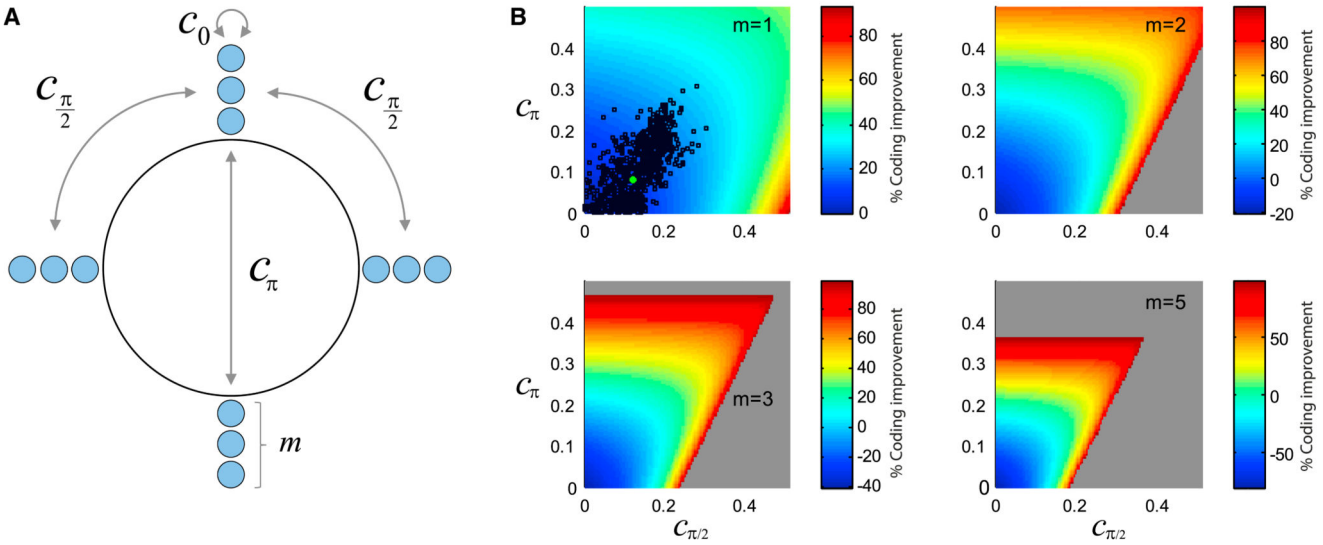

Figure 7. Effect of Correlation in Small Populations of Direction-Selective Neurons

(A) Illustration of the model of direction-selective neurons, for a given stimulus, in which there are m neurons whose preferred directions point along each of the four cardinal directions. The correlation structure is parametrized by c_0, $c_{\pi/2}$, and c_π, the values of correlations in pairs of neurons whose preferred directions differ by 0°, 90°, and 180°. (B) Profiles of the percent improvement in coding accuracy as functions of $c_{\pi/2}$ and c_π, for different values of m; in the cases m = 2, 3, and 5, c_0 was set to 0.2, comparable to the observed values. Black dots represent the values observed in the data; the green dot represents the average over all data. The gray area represents forbidden values of the correlations (with negative eigenvalues of the covariance matrix).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

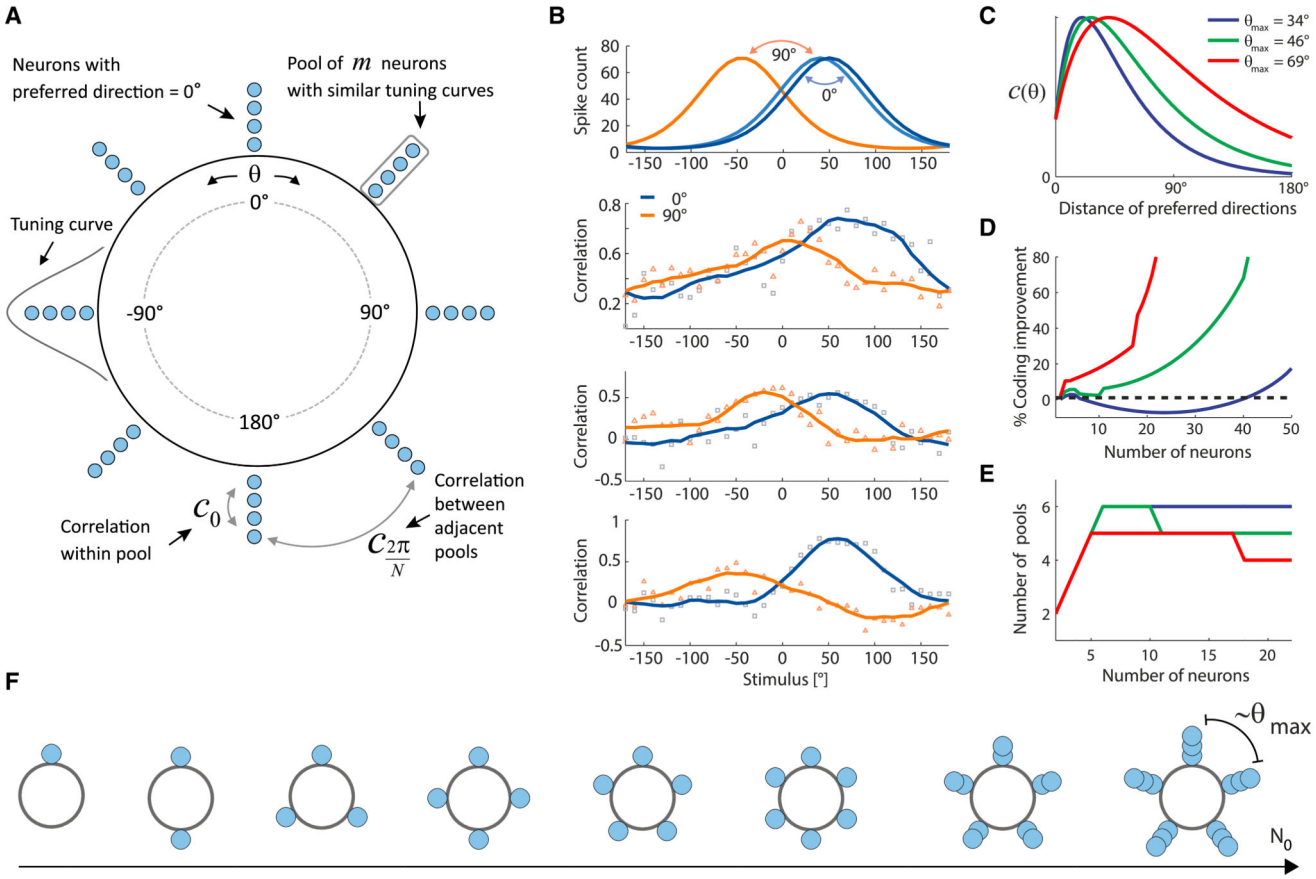

Figure 8. Beyond the Retina: Model of Coding with a Population of Broadly Tuned, Correlated Neurons

(A) Illustration of the model of a population of N_0 broadly tuned, correlated neurons. The population is divided into N pools, each of which contains m neurons with similar stimulus preference.

(B) Measured stimulus-dependent correlations in pairs of neurons. Top panel: illustration of the tuning curves. Bottom panels: three examples of measured correlations in pairs of direction-selective cells with 0° difference in preferred direction (squares) and with 90° difference (triangles), from given groups of three cells. Solid lines are smoothed versions of the raw correlation values. For a range of stimulus directions, the correlation between neurons whose preferred directions differ by 90° exceeds the correlation between neurons with similar preferred directions; the examples were chosen from three different experiments.

(C) Shapes of the correlation function used in the model; $c(\theta)$ is the correlation between two neurons with preferred stimuli that differ by θ. Different colors correspond to different parameter choices, as labeled in the legend.

(D) Percent improvement of the coding accuracy as a function of the number of neurons in the population, N_0. Color coding as in (C).

(E) Optimal number of pools, N, as a function of the number of neurons in the population, N_0. Color coding as in (C).

(F) Illustration of the optimal configurations in a correlated population of different sizes (total number of neurons, N_0). For small populations, neurons cover the stimulus range uniformly, i.e., there are as many pools as there are neurons ($N = N_0$) and their separation decreases as N_0 increases. When the separation becomes comparable to the "correlation length" set by the shape of the correlation function, the number of pools, N, stops growing with increasing N_0; instead, each pool becomes more populated.

Europe PMC Funders Author Manuscripts