*LETTER* Communicated by Andreas Tolias

# Population Coding with Motion Energy Filters: The Impact of Correlations

**F. Klam**
*fklam@salk.edu*
*Vision Center Laboratory, Salk Institute, La Jolla, CA 92037, U.S.A.*

**R. S. Zemel**
*zemel@cs.toronto.edu*
*Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 3H5*

**A. Pouget**
*alex@bcs.rochester.edu*
*Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, U.S.A.*

**The codes obtained from the responses of large populations of neurons are known as population codes. Several studies have shown that the amount of information conveyed by such codes, and the format of this information, is highly dependent on the pattern of correlations. However, very little is known about the impact of response correlations (as found in actual cortical circuits) on neural coding. To address this problem, we investigated the properties of population codes obtained from motion energy filters, which provide one of the best models for motion selectivity in early visual areas. It is therefore likely that the correlations that arise among energy filters also arise among motion-selective neurons. We adopted an ideal observer approach to analyze filter responses to three sets of images: noisy sine gratings, random dots kinematograms, and images of natural scenes. We report that in our model, the structure of the population code varies with the type of image. We also show that for all sets of images, correlations convey a large fraction of the information: 40% to 90% of the total information. Moreover, ignoring those correlations when decoding leads to considerable information loss—from 50% to 93%, depending on the image type. Finally we show that it is important to consider a large population of motion energy filters in order to see the impact of correlations. Study of pairs of neurons, as is often done experimentally, can underestimate the effect of correlations.**

 

## 1 Introduction

One of the main goals of systems neuroscience is to understand how information about the external world is encoded and subsequently processed by neural circuits. Two important results have emerged from electrophysiological recordings. First, cortical cells have variable responses to repeated presentations of a stimulus (Schreiner, Essick, & Whitsel, 1978; Tolhurst, Movshon, & Thompson, 1981). Second, cortical codes involve large numbers of neurons, the resulting codes being known as population codes (O'Keefe & Dostrovsky, 1971; Georgopoulos, Schwartz, & Kettner, 1986; Paradiso, 1988).

Several schemes have been proposed for decoding population codes (Pouget, Dayan, & Zemel, 2003). Those algorithms can recover the value of the encoded variable with high accuracy provided the neuronal noise is independent. Cortical neurons however, are not independent; in particular, their spike counts are often correlated (e.g., Zohary, Shadlen, & Newsome, 1994; Lee, Port, Kruse, & Georgopoulos, 1998).

A number of theoretical studies have investigated the impact of these correlations (von der Malsburg, 1981; Vogels, 1990; Oram, Foldiak, Perrett, & Sengpiel, 1998; Abbott & Dayan, 1999; Yoon & Sompolinsky, 1999; Wu, Nakahara, & Amari, 2001; Wu, Amari, & Nakahara, 2002; Schneidman, Bialek, & Berry, 2003). Most of those studies addressed this issue from the encoding perspective, asking whether correlations can increase the coding capacity of population codes. Surprisingly, this question does not have a simple answer: correlations can increase information (synergistic codes) or decrease information (redundant), depending on the pattern of correlations (Abbott & Dayan, 1999; Yoon & Sompolinsky, 1999). The implications of those studies for the neural codes in cortex remain unclear because they are not based on the pattern of correlations in vivo but rather on arbitrary correlations patterns that are set by hand (often chosen for their analytical tractability). The authors had no choice: the pattern of correlations in vivo is not known yet. A few measurements are available for small cell sets (2–10 cells; Zohary et al., 1994; Lee et al., 1998; Maynard et al., 1999; Averbeck & Lee, 2003; Kohn & Smith, 2005), but those are insufficient to extrapolate the entire correlation pattern.

Here we adopt a different approach. We construct a model of how population responses arise from stimuli and then study the resulting correlation patterns and their consequences. More specifically, we investigate the synergy and redundancy of a population code composed of motion energy filters (MEFs) responding to various types of images. We chose MEFs because they are one of the best available models of motion-selective V1 neurons (Basole, White, & Fitzpatrick, 2003), and they have been utilized as building blocks for modeling motion-sensitive MT cells (Simoncelli & Heeger, 1998). In this data set, the variability comes from the multiplicity of images that can instantiate the same motion, and the correlations between units arise

from stimulus-driven factors, such as the collection of motions present in a random dot kinematogram. Therefore, unlike previous studies, we focus on correlations due to the stimulus, and we ignore any contributions from internal noise; hence, we use the term *external* noise. The advantage of this approach is that we are studying correlations that are guaranteed to emerge in cortical circuits (to the extent that MEFs provide a good model of cortical responses). Any estimate of motion direction will be affected by the response variability due to the fact that the same motion can be instantiated by a wide variety of images. In the rest of this letter, we call this variability *noise*, since it carries no information about the direction of motion. It is important to keep in mind that this "noise" is in fact carrying information about other image features such as luminance or texture.

In addition to considering the impact of correlations in model units on encoding, we consider their impact from the point of view of decoding. The decoding perspective evaluates if a downstream population needs to know correlations to recover the coded information (Wu et al., 2001). Depending on the covariance matrix considered, strategies ignoring correlations can be close to optimal or not (Wu et al., 2001). Hence our second goal is to evaluate the importance of taking correlations into account in recovering information from population responses.

## 2 Methods

**2.1 Motion Energy Filters.** Motion energy filters (MEF; Watson & Ahumada, 1985; Adelson & Bergen, 1985) are built from subunits with separable spatial and temporal receptive fields, as is the case for many V1 simple cells (Robson, 1966; Tolhurst & Movshon, 1975; Field & Tolhurst, 1986; DeAngelis, Ohzawa, & Freeman, 1995; De Valois, Cottaris, Mahon, Elfar, & Wilson, 2000).

*2.1.1 Spatial Subfield.* Spatial subfields are composed by 2D Gabor functions:

$$G(\omega_s, \theta, \phi_s, x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \times \cos[\omega_s(x\cos\theta + y\sin\theta) + \phi_s], \tag{2.1}$$

where $\omega_s = 2\pi f_s$, is the spatial pulsation ($\omega_s = 2\pi f_s$ where $f_s$ is the spatial frequency), and $\theta$ and $\phi_s$ are the orientation and phase of the spatial receptive field. The parameter $\sigma$ is set by (cf. Watson & Ahumada, 1985)

$$\sigma = \frac{3\sqrt{\ln 2}}{\pi f_s}. \tag{2.2}$$

Note that the Gabor functions are centered at the same spatial location (0,0) degrees.

*2.1.2 Temporal Subfield.* The causal temporal subfield is composed of an alpha function multiplied by a cosine:

$$T(\omega_t, \phi_t, t) = \begin{cases} \dfrac{(-t)^x e^{t/\tau}}{(x\tau)^x e^{-x}} \cos(\omega_t t + \phi_t) & \text{if } t < 0 \\ 0 & \text{if } t \geq 0 \end{cases}, \tag{2.3}$$

where $\omega_t$ is the temporal pulsation ($\omega_t = 2\pi f_t$ where $f_t$ is the temporal frequency), $\phi_t$ is the temporal phase, and $x$ and $\tau$ are constants set to fit the experimentally determined response profile of V1 cells ($x = 5$, $\tau = 0.1/f_t$) (DeAngelis et al., 1995).

*2.1.3 Model of V1 Simple Cells.* The response of the simple cells, $R_{s,}$ is obtained by convolving the image with the spatiotemporal receptive field of the cells:

$$R_s(\omega_s, \theta, \phi_s, \omega_t, \phi_t, t) = \int_{-\infty}^{t} T(\omega_t, \phi_t, \tau - t) d\tau \iint_{\Re} \times G(\omega_s, \theta, \phi_s, x, y) I(x, y, \tau) dx dy, \tag{2.4}$$

where $I(x, y, t)$ is the spatiotemporal image, and $G(.)$ and $T(.)$ are the spatial and temporal subfields of the simple cell as defined in equations 2.1 and 2.3.

*2.1.4 Model of V1 Complex Cell: Motion Energy.* The responses of the simple cells $R_s$ depend on the phase of the stimulus. In order to obtain pure motion energy filters, equivalent to V1 complex cells, we have to combine units in phase quadrature in space and time. Slight numerical adjustments are necessary to obtain quadrature pairs with unbiased responses (i.e., units with no response when the stimulus contrast is 0). Specifically, the parameters $\sigma$ and $\phi_t$ are adjusted such that the following relationship is verified when $I(x, y, t)$ is constant:

$$\begin{cases} R_s(\omega_s, \theta, \phi_s, \omega_t, \phi_t, t) = 0 \\ R_s(\omega_s, \theta, \phi_s, \omega_t, \phi_t + \pi/2, t) = 0, \end{cases}$$

with $\phi_s \in \{0, \pi/2\}$.

Pure motion energy (i.e., space-time quadrature) is obtained through the following additional step (cf. Adelson & Bergen, 1985):

$$R_c(\omega_s, \theta, \omega_t, t) = \left[ R_s(\omega_s, \theta, 0, \omega_t, 0, t) - R_s\left(\omega_s, \theta, \frac{\pi}{2}, \omega_t, \frac{\pi}{2}, t\right) \right]^2 + \left[ R_s\left(\omega_s, \theta, 0, \omega_t, \frac{\pi}{2}, t\right) + R_s\left(\omega_s, \theta, \frac{\pi}{2}, \omega_t, 0, t\right) \right]^2. \tag{2.5}$$

**2.2 Model Units' Distribution and Spatiotemporal Arrangement.** MEFs are distributed in polar coordinates on logarithmic scales, such that the whole range of spatial [0.5,3 cycles/deg] and temporal [2,6 cycles/s] frequencies is covered uniformly by five filters for each of the five chosen preferred motion speed sensitivities. In other words, units' receptive fields are aligned in Fourier coordinates along straight lines in ($\omega_s$, $\omega_t$) space corresponding to constant speeds (see Figure 1A; Simoncelli & Heeger, 1998).

Five preferred speeds are chosen on a logarithmic scale to cover uniformly the Fourier space {2, 2.82, 4, 5.66, 8 deg/s} (see Figure 1A, 5 lines). Then finally, for all speeds, five preferred motion direction are chosen evenly $\{-32, -16, 0, 16, 32 \text{ deg}\}$ (see Figure 1B), leading to a total of $5 \times 5 \times 5 = 125$ units covering uniformly a portion of the ($\omega_x$, $\omega_y$, $\omega_t$) space close to the one spanned by V1 cells (Foster, Gaska, Nagler, & Pollen, 1985). This results in 25 preferred motions (5 directions $\times$ 5 speeds) evenly distributed in the ($V_x$,$V_y$) velocity space (see Figure 1C).

**2.3 Image Types.** All images are circular, with a diameter of 9.56 degrees of visual angle, the pixel size is 1/25 degree, and the contrast interval is [1 2], with 1 corresponding to the background level.

*2.3.1 Noisy Sinusoidal Grating.* This type of image is composed by a sine wave of fixed spatial frequency (0.97 cycles/s), whose orientation is perpendicular to the direction of motion, with a constant speed of 4 degrees per sec (see Figure 2A). The image is then corrupted by space-time independent noise. At each time step, the value of each pixel is perturbed by the addition of a noise term, drawn independently for each pixel from a gaussian distribution with zero mean and standard deviation, $\sigma \in [0.05 \;\; 100]$:

$$I(x, y, t) = \cos(\omega[x \cos\alpha + y \sin\alpha - vt]) + n(0, \sigma), \tag{2.6}$$

where $\omega$ is the spatial pulsation, $\alpha$ the motion direction, $v$ the speed, and $n$ the noise term. After the addition of noise, $I$ is renormalized so that contrast values stay within the [1 2] interval.

*2.3.2 Random Dots Kinematograms.* Random dots kinematograms (see Figure 2B) consist of 11 to 2963 dots (density range: 0.156–41.6 dot/deg$^2$) whose initial position is drawn from a 2D uniform distribution. Individual dots are small gaussians ($\sigma = 1/20^\circ$) chosen so that their Fourier gain above 10 cycles per degree is less than 1% of their maximal gain. In simulations, all dots move in the same direction with the same speed ($4^\circ$/s: 100% coherence). In this type of image, the variability from trial to trial is due to the variations in dot density and the relative dot positions (Barlow & Tripathy, 1997).

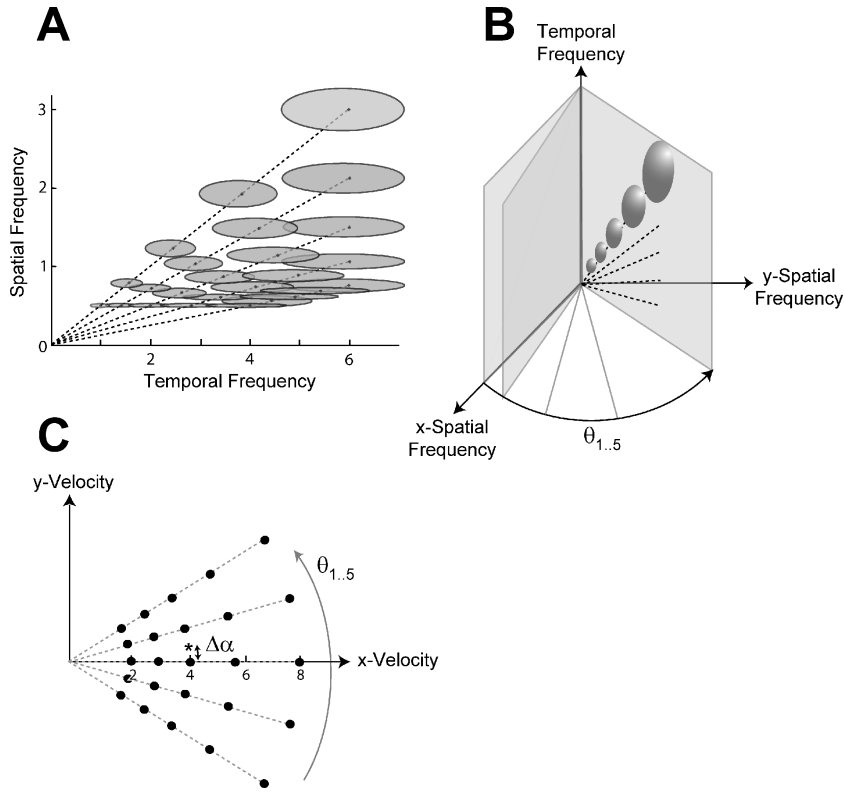


Figure 1: MEF distribution. (A) Distribution of the units in spatial-temporal frequency space. For a given preferred speed, a straight line in spatiotemporal Fourier space, five units were exponentially spaced along the range of spatial [0.5, 3 c/°] and temporal [1, 6 c/s] frequencies. Moreover, five preferred speeds were chosen between 2°/s and 8°/s. The ellipses show the spatiotemporal region where each unit responds within 95% of its maximal power. (B) Distributions of units in three-dimensional frequency space. For each of the five preferred orientations, that is, a direction in the $(x,y)$-spatial frequency plane, 25 units are distributed according to the arrangement described in *A*. This results in coverage of the 3D frequency space with five preferred directions, five preferred speeds, and five preferred spatial-temporal frequencies. (C) Preferred velocities of the units represented in velocity space. Also shown (asterisks) are the two velocities used to train the local optimal estimator, which we used to compute Fisher information.

*2.3.3 Natural Images.* We used 500 pictures of natural outdoor scenes (see Figure 2C) from the van Hateren databank (van Hateren & van der Schaaf, 1998). These large pictures consist of 1536 × 1024 pixels, each pixel covering approximately 1 minute of arc. For each trial, a picture and the initial

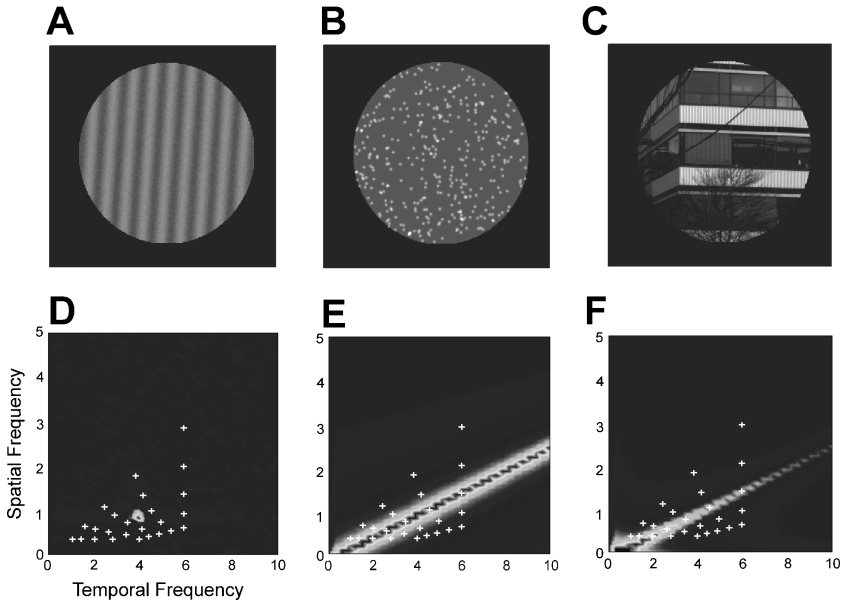


Figure 2: Image types. (A,B,C) Examples of the three different types of images used in this study. (A) Single sinusoidal grating plus spatiotemporal independent gaussian random noise. (B) Random dot pattern with all dots moving in the same direction with the same speed; the source of variability comes from the random distribution of the dots. (C) Natural scenes: the variability in this case is due to the fact that different images are used to instantiate the same motion. (D,E,F) Spatiotemporal Fourier power spectra of the images in *A*,*B*,*C*, respectively. The crosses indicate the preferred spatial and temporal frequencies of the MEFs.

position of our image patch were chosen randomly. Each patch was then convolved with a 2D gaussian to obtain a pixel resolution of 1/25 degree and motion was simulated by translating the patch at the appropriate velocity.

**2.4 Data Analysis and Fisher Information.** Each data set comprises 1000 trials of responses of the 125 MEF to two directions of motion [0 8.1°] and one speed (4°/s), for one of the three types of images. To obtain data during the steady state, the total duration of each trial is set to three times the longest temporal period (lowest temporal frequency unit). Responses are averaged over a 300 ms exposure to the steady state. The sampling rate is 78 frames per sec.

*2.4.1 Estimating Fisher Information.* Our approach to estimating Fisher information is similar to that of Series, Latham, and Pouget (2004). First, we divide our data in three sets for the two motion directions: a training set

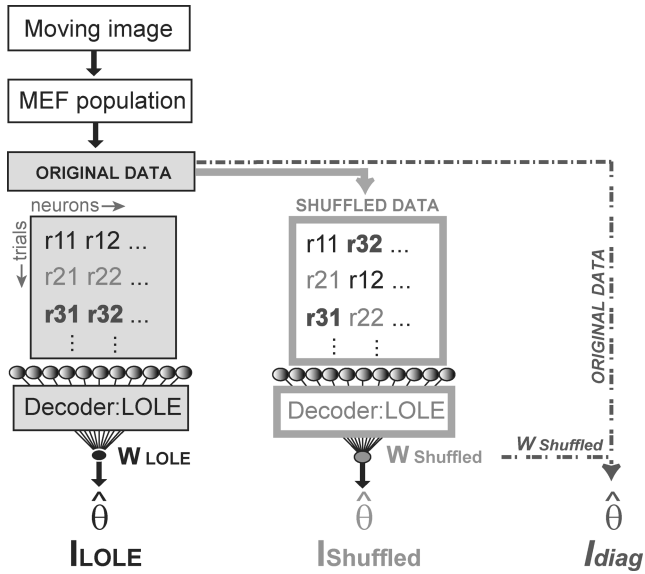


Figure 3: Information measures. We quantified information in terms of the smallest change in motion direction that can be detected based on the MEF population response. We decoded the activities of the MEFs using a local optimal linear estimator parameterized by a set of weights $\mathbf{W}_{LOLE}$ (see equation 2.7; black weights). This linear estimate was then used to obtain a lower bound on Fisher information called $I_{LOLE}$. To evaluate the synergy or redundancy of the population code, artificially decorrelated data sets were generated by shuffling the responses of individual units across trials, and a new set of weights $\mathbf{W}_{shuffled}$ was computed (light gray), leading to a second measure of Fisher information, $I_{shuffled}$. To assess the cost of ignoring correlations, we computed Fisher information using the weights optimized for the decorrelated data ($\mathbf{W}_{shuffled}$) and applying them to the original data set, yielding $I_{diag}$ (dash-dotted line). This is equivalent to decoding the population assuming that the MEFs are independent.

(400 trials), a validation set (200 trials), and a test set (400 trials) used for optimization, early stopping, and generalization, respectively.

Second, the total information ($I_{LOLE}$, see Figure 3: black inset) contained in the MEF population is estimated by training a locally optimal linear estimator (LOLE) of the direction of motion on the training data set. The LOLE is equivalent to a linear perceptron of the form

$$\hat{\theta} = \mathbf{WX}, \tag{2.7}$$

where $\mathbf{W}$ is a 125 weight vector, and $\mathbf{X}$ the matrix responses of filter responses to which the mean responses have been subtracted (125 rows $\times$ [500 $\times$ 2]columns). The weights $\mathbf{W}$ are optimized by conjugate gradient descent to estimate the two neighboring directions of motion [0°, 8.1°]. The

validation set is used to monitor the generalization performance and stop the optimization when the performance starts decreasing on this set. This early-stopping procedure avoids overfitting.

Third, after training, the mean $\{E(\hat{\theta}_i)\}_{i=1,2}$ and the variance $\{V(\hat{\theta}_i)\}_{i=1,2}$ of the estimates for the two directions of motion are computed on the test set. The lower-bound estimate of Fisher information is obtained with

$$I_{LOLE} = \frac{1}{\Delta\theta^2} \frac{(E(\hat{\theta}_2) - E(\hat{\theta}_1))^2}{[V(\hat{\theta}_2) + V(\hat{\theta}_1)]\backslash 2}, \tag{2.8}$$

where $\Delta\theta = 8.1^\circ$ is the angular difference between the two directions of motion.

This particular method for estimating Fisher information provides only a lower bound, which ignores any information that could be obtained through nonlinear decoders. However, as mentioned in section 4 ("Lower Bound on Fisher Information"), this bound appears to be quite tight because various nonlinear methods failed to obtain information beyond what we observed with the LOLE.

*2.4.2 $I_{shuffled}$ and $I_{diag}$.* In order to evaluate the information conveyed by correlations within the population code, we generate artificially decorrelated data sets by shuffling the responses of individual units across different trials in the same motion direction. Shuffling is performed by circular permutations on the rows of matrix $\mathbf{X}$ chosen in random order. The first chosen row is left intact, the second is shifted by 2 values, the next by 4, and so on (see Figure 3, light gray inset). Analyses confirmed that on each data set, this shuffling indeed reduced the correlation to near zero. Once the correlations are removed, we train a LOLE on the shuffled data set with the same cross-validation technique as described above and compute another estimate of Fisher information ($I_{shuffled}$; see Figure 3, light gray inset) using equation 2.8.

To assess the impact of correlations from a decoding point of view, we compute $I_{diag}$, the Fisher information recovered when ignoring correlations ("diag" refers to the fact that the decoder assumes a diagonal covariance matrix). This quantity is obtained by using the weights optimized for the decorrelated (shuffled) data and applying them to the original data set (see equation 2.7), but replacing $\mathbf{W}$ with $\mathbf{W}_{shuffled}$ (see Figure 3, dash-dotted line). The motion direction estimates are then used to compute $I_{diag}$ (see equation 2.8). Note that because $\mathbf{W}_{shuffled}$ is not optimized for the original data set, $I_{diag} \leq I_{LOLE}$.

**2.5 Normalizations.** In the case of random dot kinematograms and natural images, the contrast can greatly vary from one image to the next. To determine whether these contrast variations contribute to our results, we retrained the LOLE on population responses that have been normalized. We tried several different types of normalization, with two basic forms:

Norm 1: Normalization by maximum activity:

$$a_{i,p} \mapsto \frac{a_{i,p}}{\max_k(a_{k,p})}, \tag{2.9}$$

where $a_{i,p}$ is the response of unit $i$ in trial $p$. This normalization ensures that all units have their activities between 0 and 1.
Norm 2: Divisive normalization:

$$a_{i,p} \mapsto \frac{a_{i,p}^2}{c_1 + c_2 \sum_{k=1}^{n} a_{k,p}^2}. \tag{2.10}$$

We used two different sets of values for the parameters $c_1$ and $c_2$: $c_1 = 0.01$ and $c_2 = 0.25$, as used by Simoncelli and Heeger (1998) to model complex cells and $c_1 = 0.1$ and $c_2 = 0.002$, as in Deneve, Latham, and Pouget (1999).

## 3 Results

We used a population of 125 motion energy filters (MEF; Watson & Ahumada, 1985; Adelson & Bergen, 1985) spanning uniformly a large portion of the 3D Fourier space ($\omega_x$, $\omega_y$, $\omega_t$) (see section 2 and Figure 1), their receptive fields spatially overlapping.

These deterministic MEFs were presented with three sets of moving images, which displayed diverse types of image noise, or intrinsic variability (see Figure 2). The first set of images consisted of a single sinusoidal grating, corrupted with random gaussian space-time independent pixel noise (see Figure 2A). The level of noise was controlled through the standard deviation of the random gaussian contrast perturbations. The second set of images contained random dots patterns that varied in their spatial configuration across trials (a type of variability known as correspondence noise; Barlow & Tripathy, 1997). The level of variability was varied by changing the number (or density) of dots (see Figure 2B) as is done in psychophysical experiments (Watamaniuk, 1993; Barlow & Tripathy, 1997). The last set of images was natural images (see Figure 2C). In this case, the image variability was caused by choosing different images moving in the same direction. Contrary to the other sets, the level of noise could not be parametrically varied in the last set.

The responses of the MEF population to the constant velocity moving images were used to discriminate between two directions of motion differing by a small angle $\delta\theta$, where $\delta\theta$ was set to 4 degrees. Unless specified otherwise, the estimate of the direction of motion was obtained with a local optimal linear estimator (LOLE), and a lower bound on the Fisher information of this estimator was computed as described by equation 2.8.

The main goal of this work was to evaluate the role of correlations of motion energy filters in response to natural images and images that are

commonly used in psychophysics experiments. Note that all correlations are stimulus induced in these model units, since we did not add any noise to the responses. To the extent that MEFs provide a good model of motion-selective neurons, the same correlations must arise in the cortex.

**3.1 Distribution of the Pairwise Correlations Coefficients Between MEF Filters.** Figure 4 shows the distribution of the pairwise correlations coefficients between MEF filters. Interestingly, the pairwise correlations are very similar to what has been reported in vivo for the same type of stimuli (Zohary et al., 1994; Bair, Zohary, & Newsome, 2001). For random dots, the correlations tend to be mostly positive and small (see Figure 4A). Moreover, the amplitude of the correlations decreases as a function of the difference in preferred directions (see Figure 4B). Cell pairs with small differences in preferred directions (less than 30 degrees) are positively correlated, while cells with large differences in preferred directions (more than 30 degrees) show very little correlation. The same trend has been reported in MT by Zohary and colleagues.

Sinusoidal gratings induce a similar pattern of correlations as for random dots (see Figure 4C) but not natural images. With natural images, correlations tend to be stronger, with an average value of 0.38 compared to 0.12 for random dots, and 0.05 for sinusoidal gratings. We are not aware of any measurement of correlations in MT in response to natural images. It would therefore be interesting to test experimentally whether correlations are indeed stronger when using natural images.

**3.2 Total Information as a Function of Image Variability Level.** First, we checked that the performance of the MEF population parallels that of human observers in similar conditions. The variation of the total information in the system as a function of the magnitude of image variability is presented in Figure 5, for the sinusoidal grating plus pixel noise (see Figures 5A and 5C) and random dots pattern images (see Figures 5B and 5D) for which the level of noise was under parametric control. As expected, and similar to psychophysical data (Eckstein, Whiting, & Thomas, 1996a, 1996b), Fisher information (see Figure 5A) as well as the performance of classification (see Figure 5C) dropped with increasing (pixel) noise level. Note that a 75% correct classification was obtained when the standard deviation of the pixel noise was above 60, thus more than 60 times the peak-to-peak contrast of the grating. Hence, this system is very robust to pixel noise. Fisher information as well as correct classification increased as a function of the number of dots in the image, as has been observed psychophysically (Barlow & Tripathy, 1997; Watamaniuk, 1993). Note that with only 11dots in the image (dot density $= 0.156$ dots/deg$^2$) the percentage of correct classification is above 80% (see Figure 5D) and that Fisher information saturates at about 520 when the number of dots is above 700 (density$\sim$10 dots/deg$^2$; see Figure 5C) but correct classification never reaches 100%.

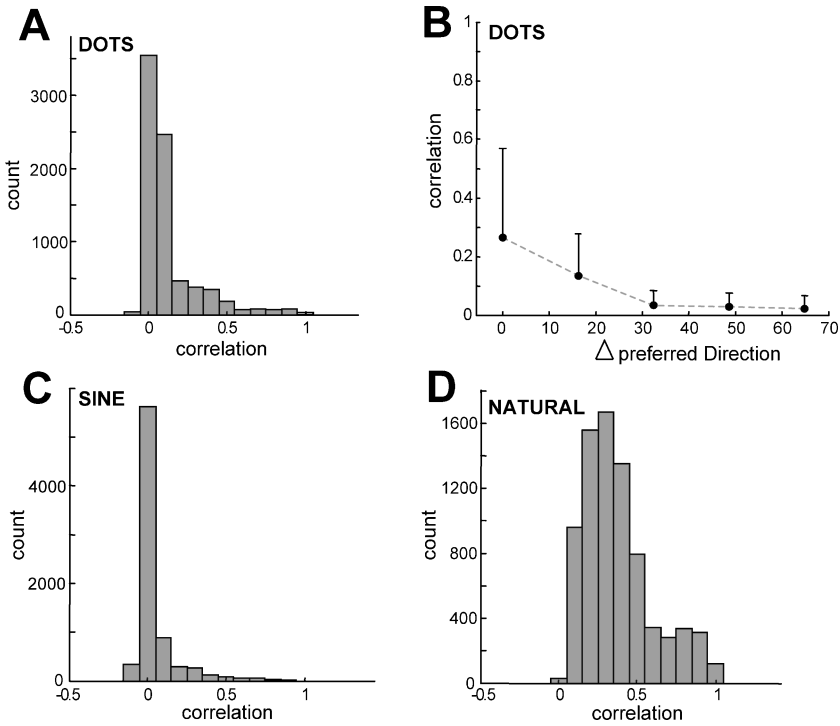


Figure 4: Pairwise correlations for the different types of images. (A) Distribution of pairwise correlations obtained with random dots (DOTS; %Correct = 97%, nDots = 371). m = 0.12 ± 0.19, med = 0.06 (mean, standard deviation, median). (B) Mean correlation (and standard deviation) as a function of difference in preferred direction between two MEFs. (C,D) Distribution of pairwise correlations for noisy sinusoidal grating images (SINE; %Correct = 98%, $\sigma_{\text{noise}} = 30$) m = 0.05 ± 0.14, med = 0.01, and for natural images (NATURAL) m = 0.38 ± 0.22, med = 0.33.

**3.3 $I_{shuffled}$: The Impact of Correlations on Encoding.** To determine whether MEF population codes are redundant (correlations encode information about the motion direction beyond individual responses) or synergistic (correlations decrease the amount of information), we compared the Fisher information computed on the original data set ($I_{LOLE}$) to the information computed on an artificially decorrelated data set in which individual responses have been shuffled across trials ($I_{shuffled}$; Nirenberg & Latham, 1998, 2003; Panzeri, Golledge, Zheng, Tovée, & Young, 2001; Series et al., 2004; see Figure 6). In a synergistic code, $I_{shuffled}$ should be less than $I_{LOLE}$ ($I_{LOLE} - I_{shuffled} > 0$), and conversely for a redundant code.

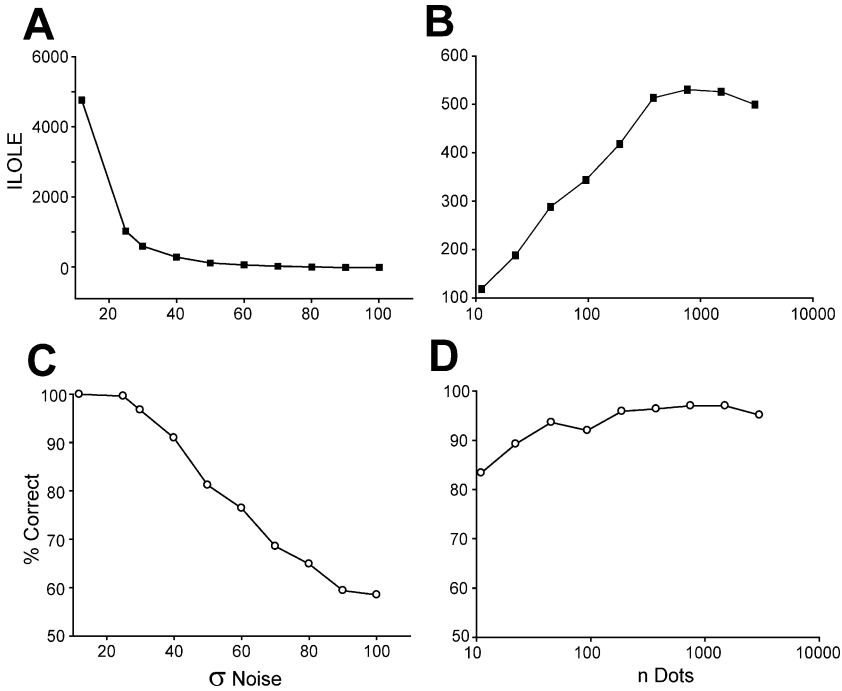


Figure 5: Fisher information and discrimination performance as a function of noise level. (A) Fisher information ($I_{LOLE}$) as a function of the standard deviation of the pixel noise for sinusoidal gratings (see Figure 2A). (B) Fisher information ($I_{LOLE}$) as a function of the number of dots in the random dots kinematograms (see Figure 2B). (C,D) Percentage of correct classification as a function of the standard deviation of the pixel noise (C) and the number of dots (D). Note that for the random dot patterns (B,D), the variability in the image decreases as the number of dots increases.

Note that the terms *synergy* and *redundancy* have been defined in different ways over the years (Barlow, 2001; Averbeck, Latham, & Pouget, 2006). According to a recent definition proposed by Schneidman et al. (2003), a population code with $N$ neurons is said to be synergistic (resp. redundant) if $I - \sum_{i=1}^{N} I_i$, where $I$ is the total Shannon information and $I_i$ is the Shannon information for neuron $i$, is positive (resp. negative). We generalized this notion to Fisher information by simply noting that $\sum_i I_i = I_{shuffled}$ in the case of Fisher information, in which case the relevant quantity for synergy and redundancy is indeed $I_{LOLE} - I_{shuffled}$.

For sinusoidal gratings with a noise level corresponding to a performance of 98% correct, we found that $I_{shuffled}$ is 26% greater than $I_{LOLE}$ (see Figure 6A, SINE), indicating that the code is redundant in this case. By

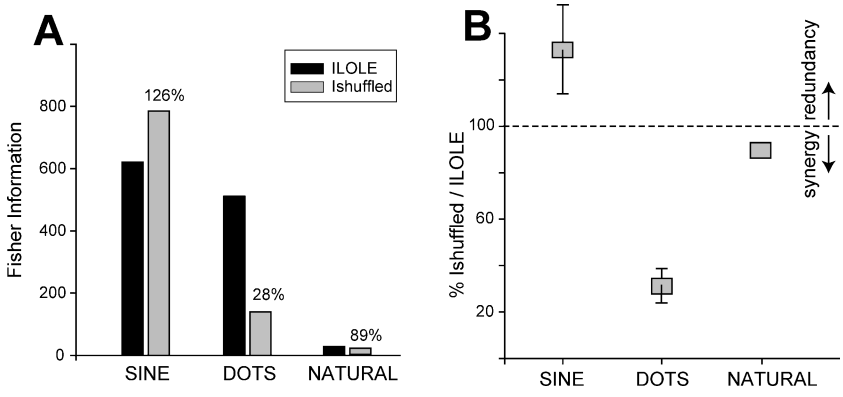


Figure 6: Evaluation of synergy or redundancy of the MEF population code. (A) $I_{LOLE}$ and $I_{shuffled}$ for population responses obtained with noisy sinusoidal gratings (SINE) for an intermediate level of noise (%Correct = 98%, $\sigma_{\text{noise}} = 30$), with random dot kinematograms (DOTS; %Correct = 97%, nDots = 371 or DotDensity = 5.2 dots/deg$^2$), and with natural images (NATURAL). (B) Percentage of $I_{shuffled}$ relative to $I_{LOLE}$ as a function of noise level for the three different image types. Average percentage (squares) and standard deviations (bars) are given. Note that the standard deviation is 0 for natural images: only one level of noise was available since the variability of the stimuli was determined only by the image data set. The code was found to be redundant for sinusoidal gratings ($I_{shuffled} > I_{LOLE}$) and synergistic for random dot kinematograms and natural images ($I_{shuffled} < I_{LOLE}$).

contrast, for random dots with a dot density allowing 97% correct classification, $I_{shuffled}$ was found to be 72% below $I_{LOLE}$ (see Figure 6A, DOTS). In other words, the MEF codes are synergistic for random dots. A similar situation was observed for natural images, although the decrease in $I_{shuffled}$ relative to $I_{LOLE}$ was only 11% (see Figure 6A, NATURAL).

The general trend, a redundant code for sinusoidal gratings and a synergistic code for random dots and natural images, is conserved for all levels of variability (see Figure 6B, $\%I_{shuffled}/I_{LOLE} = 133 \pm 19\%$ and $31 \pm 7\%$ for grating and random dots, respectively). Therefore, the same population of MEFs presented with different types of noisy images can have encoding characteristics exhibiting either synergy or redundancy.

The synergy ($I_{shuffled} < I_{LOLE}$) observed for random dots and natural images could simply be due to large contrast variability from one image to the next (see Figure 7). Indeed, the natural images varied greatly in contrast, and with random dots, the actual number of dots appearing in the receptive field of the MEFs can vary significantly from trial to trial (see Figure 7A). This shared variability can be removed easily from the population response by taking differences between neurons, in the same way noise is reduced in

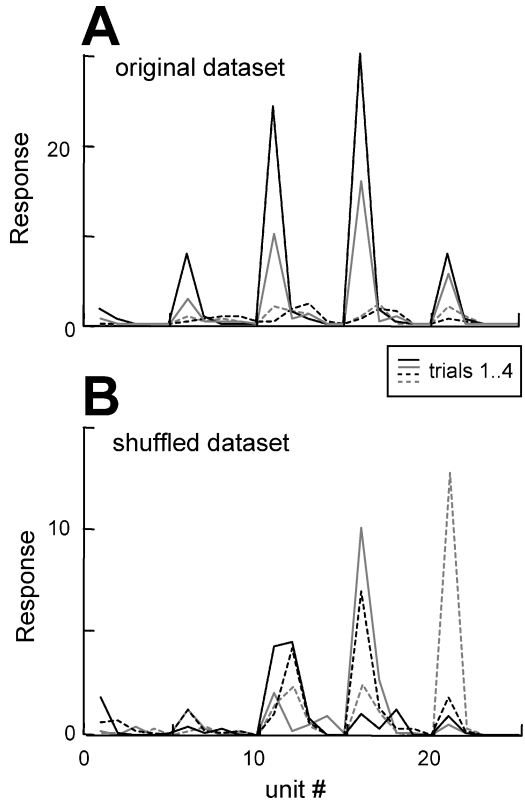


Figure 7: Cross-trial variability and shuffling. Activity of 25 units during four trials of random dots image motion, each trial being represented by a single line, for the original data (A) and shuffled data (B). On average, units tend to be either strongly or weakly active during a given trial (A). This common variability across trials disappears after shuffling (B).

a differential amplifier. Once the data are shuffled, however (see Figure 7B), this shared variability can no longer be removed, which could explain the decrease in information after shuffling.

If shared variability due to overall image contrast underlies synergy, it should be possible to normalize the activity of the MEFs to increase $I_{shuffled}$ (hence, reducing the synergy), while keeping $I_{LOLE}$ almost identical. This is indeed what we found for natural images. When we used divisive normalization (normalization 2 in section 2), the code went from synergistic to redundant ($I_{shuffled}$ equal to 95% of $I_{LOLE,}$ before normalization, 120% after) with only a minor change in $I_{LOLE,}$($I_{LOLE,NORM2} = 27$, compared to $I_{LOLE} = 29$ without normalization).

For random dot kinematograms, we also found a reduction in synergy ($I_{shuffled}$ went from 33% of $I_{LOLE,}$ to 85% after divisive normalization). However, this decrease in synergy was accompanied by a substantial decrease of $I_{LOLE}$ ($I_{LOLE,NORM2} = 93 \pm 53$ compared to $I_{LOLE} = 406 \pm 172$). We also tried other forms of normalization (e.g., normalization 1 in section 2), but they failed to reduce synergy. This indicates that the synergy for random dot kinematograms is not simply the result of variability in contrast across images but could be related to the variability in other stimulus parameters, such as the local dot density.

**3.4 $I_{diag}$: The Impact of Correlations on Decoding.** $I_{shuffled}$ measures the impact of correlation from a encoding perspective. To measure the impact of correlations from the point of view of decoding, one can measure the information recovered when decoding the patterns of activity under the assumption that the MEF units are independent (Wu et al., 2001; Pola, Thiele, Hoffmann, & Panzeri, 2003; Nirenberg & Latham, 2003). $I_{diag}$, the information recovered when the covariance matrix is assumed to be diagonal (see section 2), is guaranteed to be less than or equal to the true information (due to the independence assumption). In the case of sinusoidal gratings, with a noise level corresponding to a performance of 98% correct, the ratio $I_{diag}/I_{LOLE}$ was found to be equal to 50% (see Figure 8A, SINE). For random dot patterns, at a performance level of 97% correct (see Figure 8A, DOTS), this ratio is down to 12%. For natural images, the ratio is even lower, down to 7% (see Figure 8A, NATURAL). Accordingly, when ignoring correlations, classification performance drops from 98% to 92% for sinusoidal grating, from 97% to 72% for random dots and from 83% to 54% for natural images.

These small ratios of $I_{diag}/I_{LOLE}$ are observed for all levels of image variability (see Figure 8B: $\%I_{diag}/I_{LOLE} = 54 \pm 13\%$ and $14 \pm 5\%$ for noisy grating and random dots, respectively).

Hence, it is clear that correlations in MEF population codes have a significant impact on decoding performance for sine grating stimuli and a very large impact on the other two stimulus sets.

### 3.5 Population Code Properties and Image Variability

*3.5.1 Fourier Spectra of Images.* One simple observation that may explain the different behavior of the MEF population across the three image types is that the Fourier spectra of the different types of images differ substantially. The spectrum of a noisy sinusoidal grating is very localized (see Figure 2D); the spectrum of a random dot stimulus is roughly constant in the spatiotemporal frequency range spanned by the model neurons (see Figure 2E); and the spectrum of natural images follows the classical $1/f$ power function (see Figure 2F; van Hateren & van der Schaaf, 1998), hence, displaying a maximal power for low spatial and temporal frequencies ($\omega_s$, $\omega_t$).

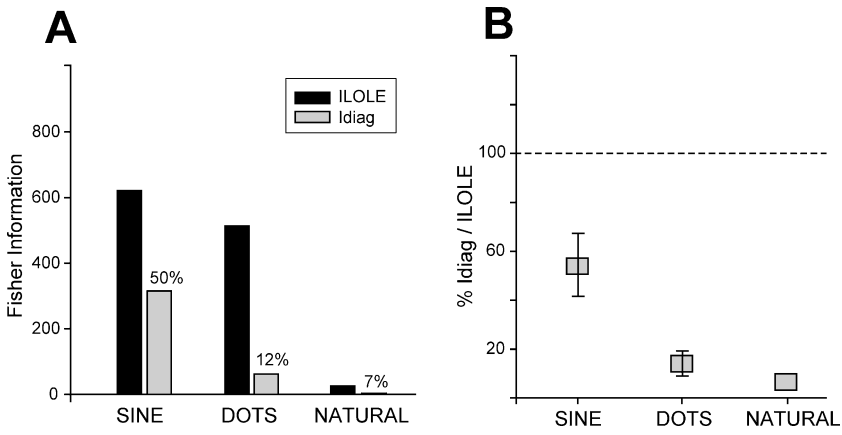


Figure 8: Impact of correlations from a decoding perspective. (A) $I_{LOLE}$ and $I_{diag}$ for population responses obtained with sinusoidal grating plus noise (SINE) for an intermediate level of noise (%Correct = 98%, $\sigma_{\text{noise}} = 30.0$), with random dots (DOTS; %Correct = 97%, number of dots = 371), and with natural images (NATURAL). (B) Percentage of $I_{diag}$ relative to $I_{LOLE}$ as a function of noise level for the three different image types. Average percentage (squares) and standard deviations (bars) are given. Note that the ratio $I_{diag}/I_{LOLE}$ is guaranteed to be less than 100% because $I_{diag}$ is obtained with weights that are not optimized for the normal data set (see Figure 3). Moreover, the standard deviation is 0 for natural images: only one level of noise was available since the variability of the stimuli was determined only by the image data set. For all images, ignoring correlations led to large information losses.

In order to test this hypothesis, we generated two new sets of images. The first consisted of a sum of noisy drifting gratings, with the spatiotemporal frequencies selected such that the resulting Fourier spectrum of the image was indistinguishable from that of random dots. The second new type of image consisted of a single natural image corrupted with pixel noise. For both images, the standard deviation of the gaussian noise was chosen so that the Fisher information ($I_{LOLE}$) was within 98.0 $\pm$ 1.2% correct for all images ($\sigma = 3$ and $\sigma = 1.6$ for the sum of noisy sine gratings and the single noisy natural image, respectively). The comparison of $I_{LOLE}$, on one hand, and $I_{shuffled}$ and $I_{diag}$, on the other hand, for the single grating plus noise and the two new image types is shown in Figure 9 (SINE, sum of SINES and NATURAL+noise, respectively). Overall, despite the fact that the Fourier spectra of these images are similar to the three depicted in Figures 2D to 2F, the relations between the different information measures are qualitatively similar to that of the single noisy grating (see Figure 9). Hence, the observed differences in statistical properties cannot be explained solely by differences

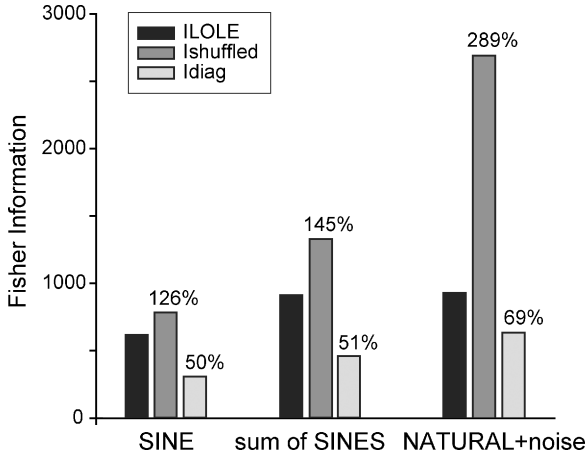


Figure 9: Influence of image Fourier spectra on information. Comparison of $I_{LOLE}$ versus $I_{shuffled}$ and $I_{diag}$ for sinusoidal gratings plus pixel noise (SINE); a combination of sinusoidal gratings plus pixel noise, adjusted so that the image Fourier spectrum is similar to the random dots images (sum of SINES); and a single natural image plus pixel noise (NATURAL+noise; in this case, the image spectrum is identical to natural image spectra). Note that noise levels used to obtain similar levels of information ($I_{LOLE}$) differ substantially (SINE: $\sigma_{\text{noise}} = 30$, sum of SINES: $\sigma_{\text{noise}} = 3$, NATURAL+noise: $\sigma_{\text{noise}} = 1.6$).

in Fourier spectra. More relevant to information is the structure of the image variability, which is not captured by the spectra.

*3.5.2 Responses of Motion Energy Filters.* The response distributions of motion energy filters across trials for a given motion direction are not gaussian, thus violating the standard assumption of multivariate gaussian noise found in most papers dealing with correlations (Abbott & Dayan, 1999; Yoon & Sompolinsky, 1999; Wu, Amari, & Nakahara, 2004). To quantify these deviations, we measured the skewness (indication of asymmetry) and kurtosis (indication of peakiness) of the distributions. When the motion direction is 0 degree (movement to the right), units stimulated with noisy sinusoidal grating have an average skewness of 1.56 ($SD = 0.33$) and a mean kurtosis of 3.82 ($SD = 1.81$) when both of these values should be 0 for gaussian distributions. This indicates that response distributions, which look similar to gamma distributions, are not symmetric and are more peaky than gaussian distributions. Both tendencies are increased when the motion energy filters are stimulated by random dots (skewness (m $\pm$ SD): 2.12 $\pm$ 0.32 and kurtosis (m $\pm$ SD): 6.90 $\pm$ 3.63) and by natural images (skewness = 10.94 $\pm$ 4.12 and kurtosis = 182.02 $\pm$ 125.81). These data reflect the observation that for random dots, and even more so for natural images, response distributions

tend to peak near 0 because for a sizable proportion of trials, a particular unit is often poorly stimulated.

To further analyze the statistics of MEF responses, we computed the variance-mean relations. In log-log plot, the variance of response was indeed a linear function of the mean, with slopes of 2.16, 2.04, and 2.03 for responses obtained with noisy gratings, random dots, and natural images, respectively. These relations do not significantly differ between conditions and are in the range of those reported in the literature (Softy & Koch, 1993: 1.25 in areas V1,MT; Lee et al., 1998: 1.1–1.17 in parietal cortex, M1; Snowden, Treue, & Andersen, 1992: 0.7–1.8 in V1).

A final step to characterize the responses across the different stimulation conditions, given that we used a motion direction discrimination task, is to examine the units' responses to the direction of motion. The tuning curves are plotted in Figure 10, for the noisy gratings (for %correct = 98%, see Figure 10A), the random dots (for %correct = 97%, see Figure 10B), and the natural images (see Figure 10C), with the units arranged according to their preferred velocity (rows) and preferred spatial (-temporal) frequency (column). Two main observations emerge from the tuning curves. First, the shapes of the directional tuning curves change with image types. As expected from the motion energy characteristics, a narrow gaussian tuning arises in response to sinusoidal grating plus pixel noise images, and the peak response is aligned with the unit's preferred direction of motion (see Figure 10A). With random dots, the tuning curves become bimodal, both peaks being at exactly 90 degrees away from the unit's preferred direction (see Figure 10B). Such bimodal responses have been observed in V1 cells (Snowden et al., 1992; Skottun, Zhang, & Grosof, 1994; Basole et al., 2003), confirming that motion energy filters provide a good model of motion-selective V1 neurons. The shapes of tuning curves to natural images are similar to those obtained with random dots images, although somewhat noisier because motion direction responses also depend on the contrast power in each orientation in the natural images. The second important observation is that the maximally active units are not the same for all types of images, although the speed and direction of motion are identical. For noisy sine gratings, the most active unit has a preferred speed of 4 degrees per sec and a spatial frequency of 0.87 c per degree, those parameters corresponding closely to the grating stimulus (see Figure 10A). In contrast, the most active unit for both random dots and natural images is the lowest speed ($2^\circ$/s) and lowest spatial frequency (0.5 c/$^\circ$; see Figures 10B and 10C for random dots and natural images, respectively). More generally, active units have preferred speeds below those of the stimulus and low spatial frequencies. Hence, depending on the type of images, the same image motion activates different units in the population. Moreover, the shape of the tuning curves change radically.

This result makes a specific prediction: that natural images and random dots moving at the same physical speed as a sinusoidal grating

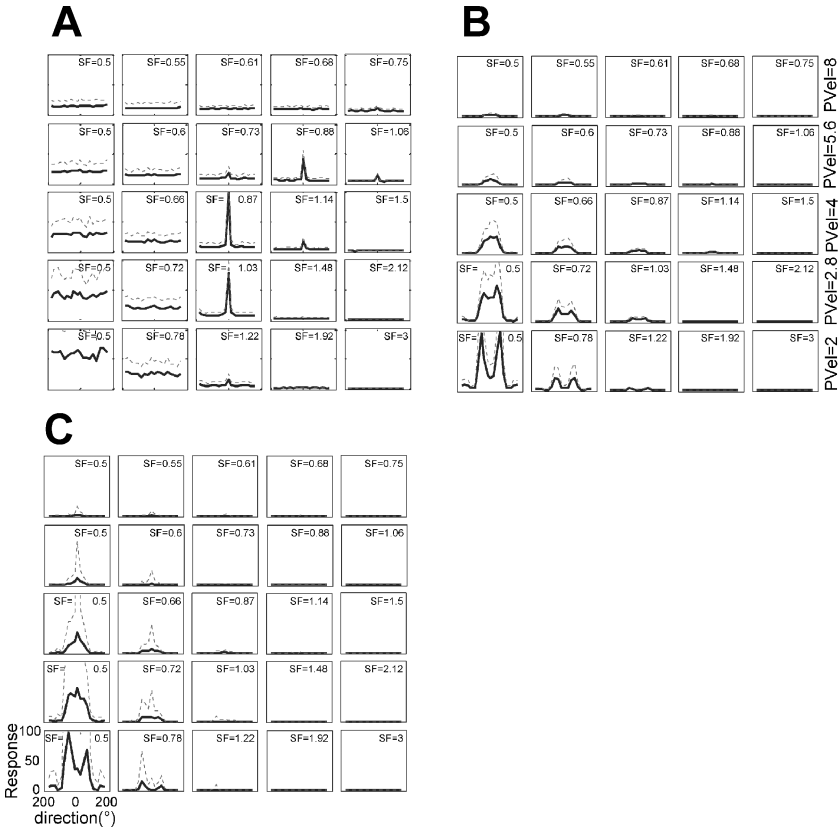


Figure 10: Motion direction tuning curves of the MEF population. (A,B,C) Tuning of MEFs to sinusoidal grating plus noise, to random dots, and to natural images, respectively. Each inset shows the response of a given unit to the direction of motion. Units are arranged as a function of their preferred spatial frequency ($x$-axis) and their preferred velocity ($y$-axis; the bottom row corresponds to the units depicted along the steepest line in Figure 1A, and the top row those lying on the flattest line). Mean responses (thick black lines) and standard deviations (dotted gray lines) for $n = 100$ trials in $A$ and $B$ and $n = 200$ trials in $C$.

should be perceived as being slower. To our knowledge, this has never been tested.

*3.5.3 Discrimination with Subsets of Motion Energy Filters.* The tuning curves to motion direction seem to indicate that different subsets of units in the population might support the motion discrimination task for different

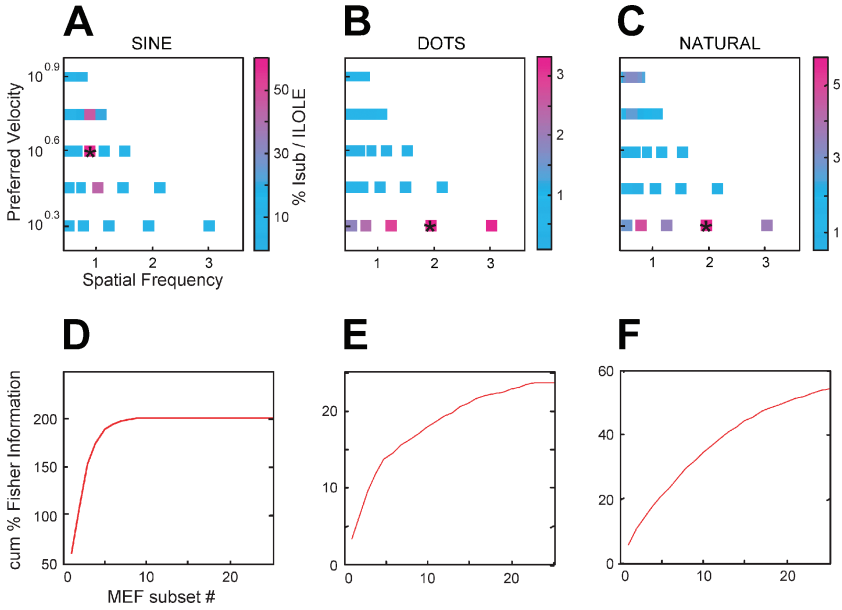


Figure 11: Subset of the most efficient units for discrimination. (A,B,C) Percentage of total information (grayscale) recovered by each subsets of model neurons ($I_{sub}/I_{LOLE}$) with identical preferred speed ($y$-axis) and spatial frequency ($x$-axis), for sinusoidal grating plus noise (A), random dots kinematograms (B), and natural images (C). The star symbol indicates which subset yields the most information. (D,E,F) Cumulative percentage of the total information ($I_{LOLE}$) yielded by summing the information obtained individually on each subsets, when including 1 to 25 subsets, for sinusoidal grating plus noise (D), random dots kinematograms (E), and natural images (F).

images. To test that hypothesis, we grouped the units according to their preferred speed and spatial frequency, across all spatial orientations, and then evaluated the discrimination performance by computing Fisher information while including only members of a given subset at the time ($I_{sub}$). Figure 11 shows the information of each subpopulation compared to the total information $I_{LOLE}$. For noisy gratings (see Figure 11A), the subset yielding the most information includes units whose preferred speed (4°/s) and spatial frequency (0.87 c/°) are closest to the grating stimulus. Using only those 25 units recovers about 60% of $I_{LOLE}$. The situation is significantly different when using random dots or natural images (see Figures 11B and 11C). The subsets yielding most information now include low (2°/s) preferred speed and intermediate (1.9 c/°) spatial frequency units, but the information based on those subgroups is only a few percent of $I_{LOLE}$ ($I_{sub\text{-}max} = 3.3\%$, and $I_{sub\text{-}max} = 5.8\%$ for random dots and natural images, respectively),

suggesting that the size of the subgroups we considered is too small in those cases for accurate motion discrimination. Those tendencies are confirmed by the cumulative percentages of $I_{sub}/I_{LOLE}$, including 1 to 25 of the subgroups (see Figures 11D to 11F). For noisy gratings (see Figure 11D), the sum of the information obtained with each subset is about twice as large as the information obtained on the whole population, whereas it is only 24% and 54% for, respectively, random dots and natural images (see Figures 11E and 11F). Note that taking subgroups separately (i.e., selecting a reduced number of covariances) is somewhat related to shuffling data (i.e., making all off-diagonal covariances zero), and similarly as the cumulative $I_{sub}$, $I_{shuffled}$ is greater than $I_{LOLE}$ for noisy gratings and lower otherwise (see Figure 6).

In summary, for a given motion direction, not only do different subsets of units support most efficiently the discrimination task depending on the type of images, but for image variabilities exhibited by random dots or natural images, discrimination is relatively poor when considering only subsets of the whole population.

**3.6 Information in Cells Pairs Versus Whole Population.** Most experimental studies report that ignoring correlations when decoding does not lead to a substantial loss of information (Nirenberg, Carcieri, Jacobs, & Latham, 2001; Pola et al., 2003; Averbeck & Lee, 2003). In contrast with those reports, our results show that decoding the population activity pattern under the assumption that neurons are independent leads to dramatic loss of information. At first sight, it would appear that our results are incompatible with those experimental findings since we report a large information loss with MEFs when ignoring correlations ($I_{diag}$ versus $I_{LOLE}$; see Figure 8). However, experiments typically consider pairs of neurons, while we studied a large population of motion energy filters. It is possible that correlations have little impact on the information in pairs of neurons, while having a large impact of the information conveyed by the entire population. To determine whether this is the case, we computed $I_{LOLE}$ and $I_{diag}$ in all pairs of motion energy filters for sine grating stimuli with $\sigma_{\text{noise}} = 30$. Following Nirenberg et al. (2001), we plot in Figure 12 the percentage of information recovered when the pairs are assumed to be independent $(1 - [I_{LOLE} - I_{diag}]/I_{LOLE})$.

Despite the fact that $I_{diag}$ is only 12% of $I_{LOLE}$ for the entire population, we found that for most cell pairs, $I_{diag}$ is 90% or more of $I_{LOLE}$. In other words, most cell pairs fail to reflect the large impact of correlations. There are a few pairs for which the $I_{diag}$ is less than 90% of $I_{LOLE}$, but such pairs represent less than 3% of the 7750 pairs in our data. It is therefore possible that the failure to find a large impact of correlations on population codes is due to the small size of the population being tested. We suspect that this is not an issue for the study of Nirenberg et al. (2001) because they used a very large number of cells pairs, but other studies could be subject to this problem.

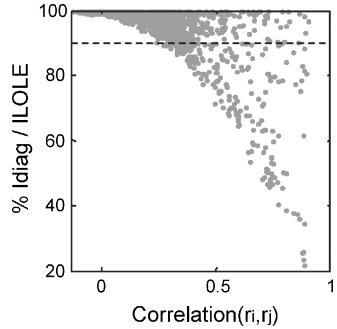


Figure 12: Underestimation of information in cell pairs. Ratio $I_{diag}/I_{LOLE}$ for pairs of MEFs expressed in percentage as a function of the correlation between the two members of the pair. Each point corresponds to one pair. A total of 7750 pairs is shown; 97% of the cells pairs show a ratio of $I_{diag}/I_{LOLE}$ above 90%, even though this ratio is 12% for the entire population. Therefore, the impact of correlations on cell pairs can greatly underestimate the impact at the population level.

## 4 Discussion

This work investigated the role of statistical dependencies among neurons in a population code, offering a novel perspective by studying the responses of a population of motion energy filters, in which the response variability is due to the image itself rather than a form of variability chosen for its suitability for subsequent analytical study. We evaluated the role of correlations in population responses on a range of image types, including natural images and images commonly used in psychophysics experiments.

**4.1 Encoding with Correlations: Redundancy or Synergy?** We assessed the impact of correlations from an encoding point of view by comparing the discrimination ability of true population responses to those obtained by shuffling the individual unit responses across trials. We found that depending on the type of image, the same population of motion energy filters can exhibit either redundancy ($I_{LOLE} < I_{shuffled}$ when using images made of a sine grating with pixel noise) or strong synergy ($I_{LOLE} > I_{shuffled}$ with random dots displays or natural images).

Our finding of strong synergy with random dot kinematograms stands in contrast to the report of Zohary et al. (1994) that neural responses in MT are primarily redundant. It is important to note, however, that Zohary et al. did not measure the correlation patterns for a large number of cells. They instead extrapolated the correlations for the population from pairwise measurements. As we saw in section 3.6, trying to infer the impact of correlations

on a population of neurons from pairwise correlations can be highly inaccurate, particularly when the values of only a few pairwise correlations are known. Therefore, it is too early to tell whether recordings in area MT are consistent or inconsistent with our results; only multielectrode recordings will answer this question. It is conceivable that in the Zohary et al. experiment, the internal noise swamps the image-induced variability that we considered here. Therefore, to test our predictions, future experiments will have to use large amounts of external noise—as is done, for instance, in perceptual learning experiments (Dosher & Lu, 1998)—to ensure that the internal variability is dominated by image-induced variability.

**4.2 Decoding Assuming Independence.** We studied the impact of correlations from a decoding perspective by comparing the discrimination ability of a system trained on the actual, correlated population responses to one trained on uncorrelated responses (the shuffled data set). This is a more biologically relevant comparison, as both systems are evaluated based on true population responses, which always contain correlations, to a common stimulus. Here the results were consistent across image types: in all cases, decoding MEF population codes assuming that the responses are independent leads to considerable loss of information ($I_{diag} \ll I_{LOLE}$). We controlled that our results did not depend on the fact that all visual receptive fields are superimposed: using only partially overlapping units, all the above results remained unchanged (not shown). Thus, for MEF population codes, decoding assuming independence is a bad strategy in terms of recovered information.

This result seems inconsistent with a number of experimental studies on neuron pairs that report that decoding while ignoring correlations does not lead to significant loss of information (Nirenberg et al., 2001; Pola et al., 2003; Averbeck & Lee, 2003). As we have pointed out, however, large information loss at the population level can be very difficult to detect in cell pairs. In our study, we found the majority of cell pairs failed to reflect the large impact of correlations.

One difference between our study and previous studies is that we used Fisher information, while those studies typically relied on Shannon information (Nirenberg et al., 2001; Pola et al., 2003; Schneidman et al., 2003). Although we have no reason to believe at this stage that correlations have a different impact on these two information measures, we cannot definitively rule out this possibility.

**4.3 Limitations of Our Study.** Lateral interactions in neural circuits could possibly modify correlations so as to lessen the impact of ignoring correlations. Since we did not model lateral interactions, we could not include this factor. Additionally, we considered only noise induced by the images, while the variability in neural populations is likely due to a combination of image-induced variability and variability generated within the

nervous system. If the latter source of variability is substantial and if the correlations for the internal noise introduce redundancies, the resulting code might be more redundant than we predict.

**4.4 Lower Bound on Fisher Information.** Our results are based on a lower bound of Fisher information obtained with a local optimal linear estimator (LOLE). This is an appealingly simple estimator that nonetheless takes into account the full set of population responses. It also has the advantage of being biologically plausible (Deneve et al., 1999; Latham, Deneve, & Pouget, 2003). However, if the derived lower bound is not tight, it is conceivable that our conclusions apply to this bound but not to the actual Fisher information. To test for this possibility, we attempted to get a better bound on Fisher information by applying a nonlinear classification algorithm, a support vector machine (SVM; Suykens, van Gestel, De Bradanter, De Moor, & Vandewalle, 2002). We found that the SVM did not significantly outperform the LOLE on any of the image types. This result provides evidence that our bound is reasonably tight since the SVM is considered to be one of the best classification techniques currently available. It also shows that a locally linear estimator can go a long way toward recovering all the information in a population with correlations. This is an important finding because Shamir and Sompolinsky (2004) have argued that for biologically realistic correlations, a nonlinear decoder might be required to recover Fisher information. This does seem to be the case for our data sets.

Still, despite this encouraging result, we cannot rule out the possibility that our bound is not tight since there might still be other nonlinear estimators that could extract all the Fisher information. The only way to ensure that we have a tight bound would be to estimate Fisher information directly. Unfortunately, this is intractable because we would have to estimate the probability distribution over the MEF responses, a distribution defined over a 125-dimensional space (corresponding to the number of motion energy filters). Directly estimating a probability distribution in such a large space would require an astronomical number of data.

This is a general problem faced by the nervous system, as an unrealistic number of data would be needed to estimate the statistics of 125 neural responses. Note that information-theoretic quantities such as multiinformation have been used recently to estimate (marginal, not conditional) correlations beyond pairs, for groups of as many as 10 neurons (Schneidman, Berry, Segev, & Bialek, 2006), but estimating these quantities for the large population studied here is quite difficult. Therefore, even if it is the case that Fisher information does not behave exactly in the same way as what we reported for our lower bound, one would have to wonder whether this is truly relevant for the nervous system. In that respect, the bound we have obtained with the LOLE is a sensible choice because it has been shown to be recoverable with a biologically plausible architecture (Deneve et al., 1999; Pouget, Zhang, Deneve, & Latham, 1998), and the

LOLE can be trained with a biologically plausible learning rule (i.e., the delta rule, Bishop, 1995).

**4.5 Signal or Noise Correlations?** Studies dealing with correlations often distinguish between noise and signal correlations. Noise correlations refer to correlations between neurons in response to the same stimulus over many trials. By contrast, signal correlations are the correlations due to changes in the stimulus from trial to trial. Recent experimental work has demonstrated the important effect that stimulus changes have on neuronal correlations (Kohn & Smith, 2005). Most theoretical studies of the impact of correlations in population codes have considered noise correlations.

One might argue that the correlations we have studied are not consistent with these earlier studies. Indeed, what we call the "stimulus" is not the image itself but the direction of motion in the image. Therefore, our claim that the stimulus was kept constant across a set of images meant that all images in this set contain the same motion, but the images themselves differed from trial to trial. This does not conform to the strict definition of noise correlations, which typically refer to correlations observed in response to repeated presentations of the same image with the same motion. In our case, this would actually result in no variability, and therefore no correlations, since we used a deterministic model.

Therefore, we are not studying "noise" correlations, as defined in previous work. Instead, our approach focuses on one component of noise correlations: those due to image-driven factors that are not directly relevant to the "stimulus" (i.e., motion), such as the dots moving in random directions, the texture in the image, luminance, or other visual features unrelated to motion. We focus on these correlations because any system trying to infer the direction of motion in sets of images has to deal with the fact that the same motion can be instantiated with completely different images. As we have seen, this variability induces complex correlations, which have a strong impact on the neural codes.

A few experimental studies have measured correlations with a stimulus set in which the random detail of the dot patterns varied within trials while the direction of motion is kept constant (Zohary et al., 1994; Bair et al., 2001). Both of these studies have concluded that image-induced variability has little impact on the correlation patterns in the brain, suggesting that internal variability might dominate in the nervous system. There are, however, several problems with this conclusion. First, as we pointed out earlier, inferring the impact of correlations on the whole population from a few pairwise measurements can lead to very large errors. Hence, what would appear as minor changes in pairwise correlations could in fact have a large impact on the overall population. Moreover, the best way to study the impact of image-induced correlations is to select a stimulus set in which the performance of the subject is known to be limited by image noise, as is done in perceptual learning studies. It is not know whether this was the case

for the stimuli used by Zohary et al. (1994) and Bair et al. (2001). Therefore, neurophysiology experiments are needed to explore this issue further.

Finally, it is important to stress that all of our conclusions apply to locally linear decoders of direction of motion. Whether they generalize to other variables is unclear. We chose to focus on direction alone because there are situations, such as catching a ball, in which we need to know the motion of objects but not other factors such as contrast. However, it is possible that our conclusions would have changed had we considered decoders that extract multiple features in parallel, such as direction of motion, spatial frequency, and contrast. Moreover, even when catching a ball, our nervous system does not need to estimate motion directly. Instead, it needs to compute a motor command, that is, a complex nonlinear function of motion. Correlations may have a different impact depending on the function being computed. This question would be worth investigating further since recent work has shown that the optimality of a code depends on the computation performed (Salinas, 2006).

## Acknowledgments

This work was supported by the Fyssen Foundation, Paris and the Canadian Institute for Advanced Research.

## References

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

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A, 2*, 284–299.

Averbeck, B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Review Neuroscience, 7*, 358–366.

Averbeck, B. B., & Lee, D. (2003). Neural noise and movement-related codes in the macaque supplementary motor area. *J. Neurosci., 23*, 7630–7641.

Bair, W., Zohary, E., & Newsome, W. T. (2001). Correlated firing in macaque visual area MT: Time scales and relationship to behavior. *J. Neurosci., 21*, 1676–1697.

Barlow, H. (2001). Redundancy reduction revisited. *Network, 12*, 241–253.

Barlow, H., & Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *J. Neurosci., 17*, 7954–7966.

Basole, A., White, L. E., & Fitzpatrick, D. (2003). Mapping multiple features in the population response of visual cortex. *Nature, 423*, 986–990.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. New York: Oxford University Press.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1995). Receptive-field dynamics in the central visual pathways. *Trends Neurosci., 18*, 451–458.

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

De Valols, R. L., Cottaris, M. P., Mahon, L. E., Elfar, S. D., & Wilson, J. A. (2000). Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision Res., 40*(27), 3685–3702.

Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. USA, 95*, 13988–13993.

Eckstein, M. P., Whiting, J. S., & Thomas, J. P. (1996a). Role of knowledge in human visual temporal integration in spatiotemporal noise. *J. Opt. Soc. Am. A, 13*, 1960–1968.

Eckstein, M. P., Whiting, J. S., & Thomas, J. P. (1996b). Detection and contrast discrimination of moving signals in uncorrelated gaussian noise. In H. Kundel (Ed.), *Medical imaging, image perception: Proc. SPIE* 2712 (pp. 9–25). Bellingham, WA: SPIE.

Field, D. J., & Tolhurst, D. J. (1986). The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc. R. Soc. Lond. B. Biol. Sci., 228*, 379–400.

Foster, K. H., Gaska, J. P., Nagler, M., & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J. Physiol., 365*, 331–363.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science, 233*, 1416–1419.

Kohn, A., & Smith, M. A. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci., 25*, 3661–3673.

Latham, P. E., Deneve, S. & Pouget, A. (2003). Optimal computation with attractor networks. *Journal of Physiology (Paris), 97*, 683–694.

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

Maynard, E., Hatsopoulos, N., Ojakangas, C., Acuna, B., Sanes, J., Normann, R., & Donoghue, J. (1999). Neuronal interactions improve cortical population coding of movement direction. *J. Neurosci., 19*, 8083–8093.

Nirenberg, S., Carcieri, S. M., Jacobs, A. L., & Latham, P. E. (2001). Retinal ganglion cells act largely as independent encoders. *Nature, 411*, 698–701.

Nirenberg, S., & Latham, P. E. (1998). Population coding in the retina. *Curr. Opin. Neurobiol., 8*, 488–493.

Nirenberg, S., & Latham, P. E. (2003). Decoding neuronal spike trains: How important are correlations? *Proc. Natl. Acad. Sci. USA, 100*, 7348–7353.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Res., 34*, 171–175.

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

Panzeri, S, Golledge, H. D. R., Zheng, F., Tovée, M. J., & Young, M. P. (2001). Objective assessment of the functional role of spike train correlations using information measures. *Vis. Cognition, 8*, 531–547.

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

Pola, G., Thiele, A., Hoffmann, K. P., & Panzeri, S. (2003). An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network, 14*, 35–60.

Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. *Annu. Rev. Neurosci., 26*, 381–410.

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

Robson, J. C. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *J. Opt. Soc. Am. A, 56*, 1141–1142.

Salinas, E. (2006). How behavioral constraints may determine optimal sensory representations. *PLoS Biol., 4*, e387.

Schneidman, E., Berry, M. J. II, Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature, 440*, 1007–1012.

Schneidman, E., Bialek, W., & Berry, M. J. II. (2003). Synergy, redundancy, and independence in population codes. *J. Neurosci., 23*, 11539–11553.

Schreiner, R. C., Essick, G. K., & Whitsel, B. L. (1978). Variability in somatosensory cortical neuron discharge: Effects on capacity to signal different stimulus conditions using a mean rate code. *J. Neurophysiol., 41*, 338–349.

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

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

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Res., 38*, 743–761.

Skottun, B. C., Zhang, J., & Grosof, D. H. (1994). On the directional selectivity of cells in the visual cortex to drifting dot patterns. *Vis. Neurosci., 11*, 885–897.

Snowden, R. J., Treue, S., & Andersen, R. A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Exp. Brain Res., 88*, 389–400.

Softky, W. R., & Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci., 13*(1), 334–350.

Suykens, J. A. K., van Gestel, T., De Bradanter, J., De Moor, B., & Vandewalle, J. (2002). *Least squares support vector machines*. Singapore: World Scientific.

Tolhurst, D. J., & Movshon, J. A. (1975). Spatial and temporal contrast sensitivity of striate cortical neurones. *Nature, 257*, 674–675.

Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp. Brain Res., 41*, 414–419.

van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. B: Biol. Sci., 265*, 359–366.

Vogels, R. (1990). Population coding of stimulus orientation by striate cortical cells. *Biol. Cybern., 64*, 25–31.

von der Malsburg, C. (1981). The correlation theory of brain function. In E. Domany, J. L. van Hemmen, & K. Schulten (Eds.), *Models of neural networks II* (pp. 95–119). Berlin: Springer-Verlag.

Watamaniuk, S. N. (1993). Ideal observer for discrimination of the global direction of dynamic random-dot stimuli. *J. Opt. Soc. Am. A, 10*, 16–28.

Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *J. Opt. Soc. Am. A, 2*, 322–341.

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

Wu, S., Amari, S., & Nakahara, H. (2004). Information processing in a neuron ensemble with the multiplicative correlation structure. *Neural Netw., 17*, 205–214.

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

Yoon, H., & Sompolinsky, H. (1999). The effect of correlations on the Fisher information of population codes. In M. J. Kearns, S. A. Solla, & D. A. Cohn (Eds.), *Advances in neural information processing systems* (pp. 167–174). Cambridge, MA: MIT Press.

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

---

Received June 19, 2006; accepted January 28, 2007.