Strong, but not Weak, Noise Correlations are Beneficial for Population Coding

Gabriel Mahuas,[1,2] Thomas Buffet,[1] Olivier Marre,[1] Ulisse Ferrari,[1,*] and Thierry Mora[2,*]
[1]*Institut de la Vision, Sorbonne Université, CNRS, INSERM, 17 Rue Moreau, 75012 Paris, France*
[2]*Laboratoire de Physique de École Normale Supérieure, CNRS, PSL University, Sorbonne University, Université Paris-Cité, 24 Rue Lhomond, 75005 Paris, France*

(Received 12 March 2025; accepted 30 July 2025; published 14 August 2025)

Neural correlations play a critical role in sensory information coding. They are of two kinds: signal correlations, when neurons have overlapping sensitivities, and noise correlations from network effects and shared noise. In experiments from early sensory systems and cortexes, many pairs of neurons typically show both types of correlations to be positive and large, especially between nearby neurons with similar stimulus sensitivity. However, theoretical arguments have suggested that the stimulus and noise correlations should have opposite signs to improve coding, at odds with experimental observations. We analyze retinal recording in response to a large variety of stimuli and show that, contrary to common belief, large noise correlations are beneficial for coding, even if aligned with signal correlations. To understand this result, we develop a theory of visual information coding by correlated neurons, which resolves that paradox. We show that noise correlations are always beneficial if they are strong enough, unless neurons are perfectly correlated by the stimulus. Finally, using neuronal recordings and modeling, we show that for high-dimensional stimuli noise correlation benefits the encoding of fine-grained details of visual stimuli, at the expense of large-scale features, which are already well encoded.

DOI: 10.1103/nhvz-6grt

I. INTRODUCTION

Neurons from sensory systems encode information about incoming stimuli in their collective spiking activity. This activity is noisy: Repetitions of the very same stimulus can drive different responses [1–6]. It has been shown that the noise is shared among neurons and synchronizes them, an effect called noise correlations, as opposed to signal correlations induced by the stimulus [6–10]. Noise correlations have been observed since the first synchronous recordings of multiple neurons [11,12] and at all levels of sensory processing, from the retina [2,13–21] to the visual cortex [1,4,5,22–25] and other brain areas [6,7,9,26–29].

Strong noise correlations have been measured mostly between nearby neurons with similar stimulus sensitivity [1,5,12,23,27,28,30–32]. This behavior is particularly evident in the retina between nearby ganglion cells of the same type [2,14,16,17,21]. This observation is however surprising, since previously it was thought that these correlations are detrimental to information coding: A theoretical argument [1,33–36] suggests that noise correlations are detrimental to information transmission if they have the same sign as signal correlations [7,9,10]. This rule is sometimes called the sign rule [37] and is related to the notion of information-limiting correlations [38]. Since nearby neurons with similar tuning are positively correlated by the signal, the theory would predict that their positive noise correlations should be detrimental, making the code less efficient. However, a large body of literature has reported the beneficial effects of noise correlations on coding accuracy, with various approaches and reasonings: by computing the Fisher information [39,40] or mutual information [41,42] in models of large populations of correlated neurons, by studying specific cases where the sign rule applies [18,43], by showing explicitly that decorrelating noise reduces decoding accuracy [44], or by showing that ignoring noise correlations in the decoder is detrimental [15,20,45]. Because of these contradictions, the effect of shared variability on information transmission is still unclear and remains a largely debated topic in neuroscience [7–9].

Here we aim to resolve these tensions by developing a general framework that is grounded in the analysis of multielectrode array recordings of rat and mouse retinas and builds on previous theoretical work [46]. While previous studies have considered the impact of noise correlations either for particular stimuli [1,15,18,29] or for particular models [36,40,41], our approach is general and covers both low- and high-dimensional stimuli. We show that the sign rule can be broken in a specific regime that we observed in retinal responses: When noise correlations are strong enough compared to signal correlations, they have a beneficial effect on information transmission. Our results unravel the complex interplay between signal and noise correlations and predict when and how noise correlations are beneficial or detrimental. In the case of high-dimensional stimuli, such as images or videos, our theory predicts different effects of noise correlations depending on stimulus features. In particular,

*These authors contributed equally to this work.

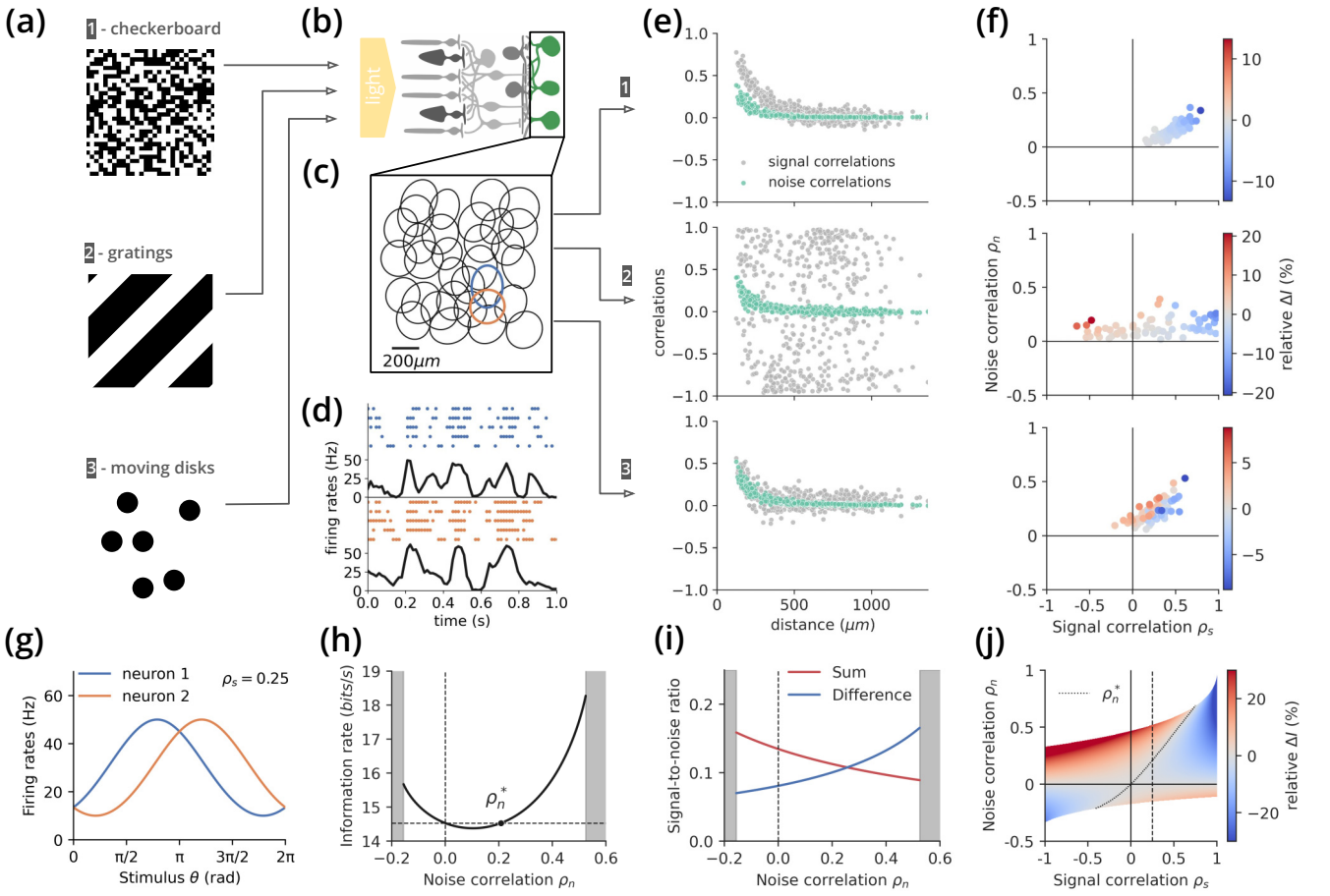

FIG. 1. The effect of noise correlations on information coding depends on the stimulus. (a) Three stimuli with different spatiotemporal statistics were presented to a rat retina. (b) Retinal ganglion cells were recorded using a multielectrode array. (c) We isolated a nearly complete population of OFF-α cells, with receptive fields (RFs) that tile the visual field following approximately a triangular lattice. (d) Example raster plots and firing rates for two cells with neighboring RFs. (e) Signal and noise correlations for each pair of neurons in the population versus their distance. Each plot corresponds to one of the three stimuli of (c). (f) Noise synergy induced by noise correlations for all pairs of nearby neurons (greater than or equal to 300 μm) for each stimulus of (e). (g) Example pair of von Mises tuning curves with moderate signal correlation level ($\rho_s = 0.25$). (h) Mutual information between stimulus and response for the example pair of (a) vs the strength of noise correlations. Gray areas correspond to forbidden correlations zones. (i) The nonmonotonicity of (b) may be explained by examining how well the stimulus is represented by the sum and difference of the two neurons' activities, as measured by their signal-to-noise ratios. Noise correlations enhance noise in the sum but reduce it in the difference. (j) Heatmap representing the noise synergy, defined as the relative gain of mutual information induced by noise correlations compared to the uncorrelated case. The dotted vertical line corresponds to the example pair of (a) and (b).

it explains how large noise correlations between neurons with similar stimulus sensitivity help encode fine details of the stimulus.

We first analyze the impact of correlated activity of rat retinal ganglion cells for a variety of stimuli and conditions. We observe violations of the sign rule, with large noise correlations being beneficial even for positive signal correlations. We then build on Ref. [46] to develop a theory of information coding by correlated neurons which extends previous theoretical efforts and accounts for the observed regimes beyond the sign rule. We extend our analysis to large populations of sensory neurons and propose a spectral analysis suggesting that local noise correlations enhance information by favoring the accurate encoding of fine-grained details. Finally, we validate that prediction by combining data from the mouse retina with accurate convolutional neural network models.

II. RESULTS

A. Benefit of noise correlations in pairs of retinal ganglion cells

We first asked under what conditions noise correlations could be beneficial or detrimental in a population of sensory neurons, taking the retina as a test case. We recorded *ex vivo* the joint spiking activity of rat retinal ganglion cells (RGCs) (see Sec. IV) [47,48]. We subjected the same retinal preparation to three stimuli with distinct spatiotemporal patterns: a random flickering checkerboard, drifting gratings, and randomly moving disks [Fig. 1(a)]. The activity of RGCs was recorded using a multielectrode array [Fig. 1(b)] and data were processed to assign spikes to each neuron [49]. The time-dependent response was then defined as the number of spikes in each 20-ms time windows [chosen to capture the peak of noise correlations in pairwise cross correlograms (see

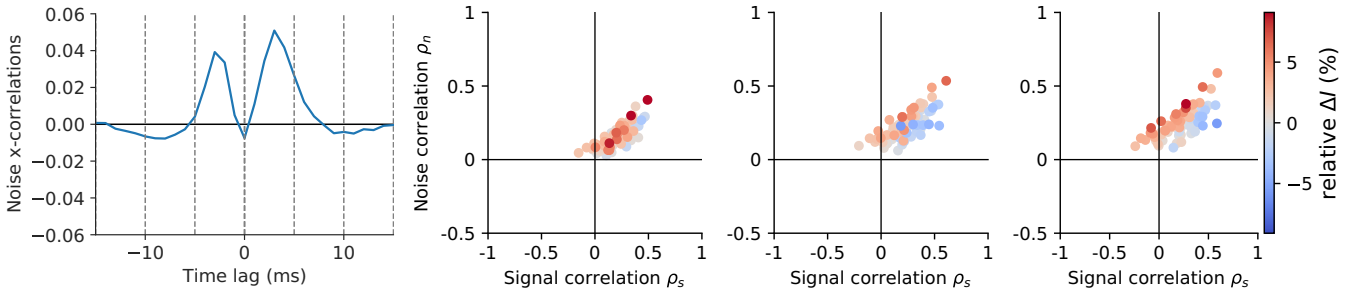

FIG. 2. Choice of bin size. Shown on the left is an example cross correlogram between two neighboring retinal ganglion cells. The two symmetric peaks, which are attributed to the effect of gap junctions, cover a short timescale of 10 ms. Shown on the right is an analysis of noise synergy ΔI [same as Fig. 1(f), bottom] for a time bin of 10, 20, and 30 ms, from left to right.

Fig. 2)]. We identified cells belonging to a nearly complete OFF-α population forming a regular mosaic pattern of their receptive fields [Fig. 1(c)].

Each of the three stimulus movies was repeated multiple times [Fig. 1(d)], which allowed us to compute the noise and signal correlation functions ρ_n and ρ_s [Fig. 1(e)] (see Sec. IV). All three stimuli produced similar structures of noise correlations across the network, with positive correlations between cells with nearby receptive fields. This is consistent with the fact that noise correlations are a property of the network, independent of the stimulus [21,50], and likely come here from gap junctions coupling neighboring RGCs [16,51]. In contrast, signal correlations strongly depend on the statistical structure of the presented stimulus and may be positive or negative, with varying strengths.

We define noise synergy as the gain in information afforded by noise correlations $\Delta I = I(\rho_n) - I(\rho_n = 0)$. To estimate it, we compute exactly the mutual information between stimulus and response for all pairs of cells whose receptive fields were closer than 300 μm and then subtract the same quantity computed on the data after shuffling across repetitions [Fig. 1(f)]. The case of the drifting gratings with fixed orientation offers an illustration of the sign rule. That stimulus induces strong positive or negative signal correlations between many cells, depending on their positions relative to the gratings direction. Since noise correlations are positive, they may be either of the same sign as signal correlations, and therefore detrimental, or of opposite sign, hence beneficial. In the case of the checkerboard stimulus, noise correlations were found to be generally detrimental. This again agrees with the sign rule since they have the same sign as signal correlations. Finally, the case of the moving disks provides an example of a third regime, which violates the sign rule: Noise correlations are of the same sign as the signal correlations, but also of comparable magnitude; yet they are beneficial for information. We checked that those results were robust to the choice of bin (Fig. 2).

Overall, the three stimuli illustrate three possible regimes when noise correlations are positive: a beneficial effect when signal correlations are negative, a detrimental effect if signal correlations are positive and large, and a beneficial effect when noise and signal correlations are both positive and of the same magnitude.

B. Strong pairwise noise correlations enhance information transmission

To make sense of our experimental findings, we consider a simple model of a pair of spiking neurons encoding an angle θ, for instance, the direction of motion of a visual stimulus, in their responses r_1 and r_2. These responses are correlated through two sources: signal correlations ρ_s due to an overlap of the tuning curves [Fig. 1(g)] and constant noise correlations ρ_n due to shared noise (see Sec. IV for mathematical definitions). We asked how this shared noise affects the encoded information, for a fixed level of noise in neurons.

To quantify the joint coding capacity of the two neurons, we computed exactly the mutual information $I(\theta; r_1, r_2)$ between their activities and the stimulus θ. For fixed tuning curves, we find that the mutual information depends nonmonotonically on the noise correlation ρ_n [Fig. 1(h)]. For small absolute values of ρ_n, the sign rule is satisfied, meaning that negative noise correlations are beneficial and weak positive ones are detrimental [7,35,37,39]. However, the mutual information increases again and noise correlations become beneficial if they are larger than a certain threshold ρ_n^*, violating the sign rule. This nonmonotonic dependence may be intuitively explained as the interplay between two opposite effects [Fig. 1(i)]. Negative noise correlations are beneficial because they reduce noise in the total activity of the neurons. By contrast, positive noise correlations reduce noise in their differential activity, but this effect only dominates when they are strong enough.

Mirroring the experimental results of Fig. 1(f), Fig. 1(j) shows theoretically how noise synergy depends on both the noise and signal correlation, where the latter is varied in the model by changing the overlap between the tuning curves. Very generally and beyond the cases predicted by the sign rule, noise correlations are beneficial also when they are stronger than the signal correlations, as we had observed in the retinal data [bottom plot of Fig. 1(f)]. We can gain insight into this behavior by computing an approximation of the mutual information that is valid for small correlations, following Ref. [46] (see Sec. IV). The noise synergy can be expressed as

$$\Delta I \approx \frac{\alpha}{2}\rho_n(\rho_n - \rho_n^*), \tag{1}$$

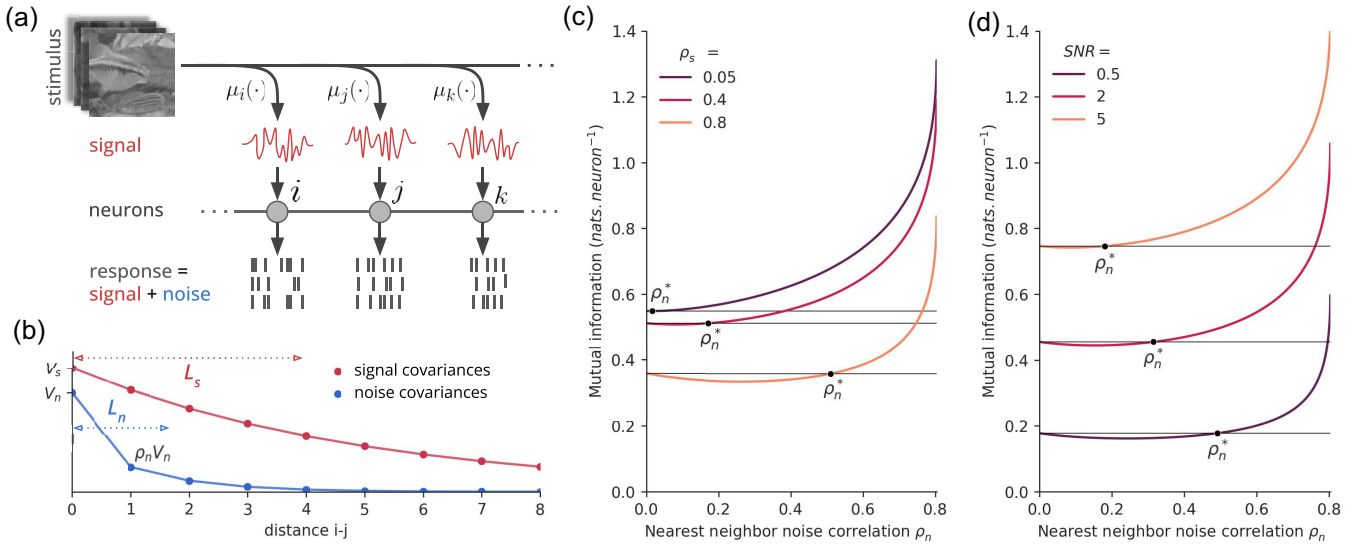

FIG. 3. Population analysis. (a) Neurons are assumed to be spatially arranged along sensory space. They combine features of the stimulus through a response function μ_i. Noise is added to the neural responses. (b) Signal and noise covariances versus distance between neurons. Signal and noise covariances decay exponentially with distance with spatial scales L_s and L_n. Mutual information is plotted as a function of the noise correlation between neighbors for (c) varying levels of signal correlations, with fixed $V_s = 2$, $V_n = 1$, and $L_n = 2$, and (d) varying levels of signal-to-noise ratio (SNR $= V_s/V_n$), with $L_s = L_n = 2$ ($\rho_s \approx 0.6$).

where $\alpha \leqslant 1$ is a prefactor that grows with the signal-to-noise ratio (SNR) of the neurons. Equation (1) captures the behavior of Figs. 1(f) and 1(j), in particular the observation that noise correlations are beneficial if $\rho_n \rho_s < 0$, as the sign rule predicts, or if they are strong enough, $|\rho_n| > |\rho_n^*|$. We can show (see Sec. IV) that the threshold ρ_n^* scales with the signal correlation strength ρ_s:

$$\rho_n^* = \beta \rho_s. \tag{2}$$

This result holds for either discretely spiking or Gaussian neurons (see Sec. IV) and the prefactor $\beta \leqslant 1$ gets smaller and even approaches 0 as the SNR increases. It is also smaller when these SNRs are dissimilar between cells, consistent with previous reports [40]. When the SNRs are weak and similar, we have $\beta \approx 1$. This analysis indicates that noise correlations are beneficial when they are of the same strength as signal correlations, but also that this benefit is enhanced when neurons are reliable.

Our definition of the noise synergy relies on comparing the noise-correlated and uncorrelated cases at fixed noise level or SNR. However, increasing noise correlations at constant SNR decreases the effective variability of the response, as measured by the noise entropy of the joint response of the pair (see Sec. IV). This means that high noise correlations imply a more precise response, which could explain the gain in information. To study this possible confounding factor, we also computed ΔI at equal noise entropy, instead of equal SNR, and found that strong noise correlations are still beneficial, with modified $\rho^* = 2\rho_s/(1+\rho_s^2) \leqslant 1$ (see Sec. IV).

Our theoretical predictions rely on the assumption that noise (Pearson) correlations do not vary with the stimulus. We wondered whether the theory was robust to stimulus-dependent noise correlations. Previous work based on Fisher, rather than Shannon, information suggested that noise correlations are detrimental when aligned to the signal direction in each point of response space [38,52]. This structure, called differential or information-limiting correlations, implies that noise correlations vary as a function of the stimulus itself. Their detrimental effect can be intuited from the definition of the Fisher information [38] and in fact provide a sort of worst-case scenario for the effect of noise correlations, as the alignment of noise and signal is satisfied locally for all values of the stimulus and thus maximally detrimental according to the sign rule. To investigate this scenario, we performed a numerical analysis of the noise synergy in a two-neuron system with information-limiting correlations [see Appendix A and Fig. 8(a)]. We find that information-limiting correlations become increasingly beneficial to the mutual information as their strength increases [Fig. 8(b)], while they are always detrimental to the Fisher information [Fig. 8(c)], consistent with previous results. While perhaps counterintuitive, this result can be understood by considering the effect of the curvature of the response curve, which induces nonlocal, nonlinear effects (ignored by the Fisher information) that become stronger with increasing noise correlations [Fig. 8(a)]. Again, it is worth stressing that information-limiting correlations were designed to be maximally detrimental. Our finding that even those are beneficial as long as noise correlation is strong enough provides strong evidence that our conclusion is very general, the only exception being the trivial case of perfectly correlated neurons ($\rho_s = 1$).

C. Large sensory populations in high dimension

We then asked how these results extend from pairs to large populations, by considering a large number of neurons tiling sensory space [Fig. 3(a)]. To go beyond neurons tuned to a single stimulus dimension and account for the ability of neurons to respond to different stimuli in a variety of natural contexts, we assume that each neuron responds to high-dimensional

stimulus, such as a whole image, a temporal sequence, or a movie. As different stimuli are shown, the spike rate of each neuron will vary. For computational ease, we take these fluctuations to be Gaussian of variance V_s.

To account for the empirical observation that nearby neurons tend to have close receptive fields, we correlate the responses of any two neurons with a strength that decreases as a function of their distance in sensory space, with characteristic decay length L_s [Fig. 3(b)]. The value of the correlation between nearest neighbors quantifies the signal correlation ρ_s. For simplicity, the response noise is also assumed to be Gaussian of variance V_n. To model positive noise correlations between nearby neurons observed in both the retina [2,14,16,17,21] and cortex [1,5,12,23,27,28,30–32], we assume that they also decay with distance, but with a different length L_n [Fig. 3(b)]. The noise correlation between nearest neighbors, defined as ρ_n, quantifies their strength.

In this setting, both signal and noise correlations are positive and the sign rule alone would predict a detrimental effect of noise correlations. The mutual information can be computed analytically in terms of simple linear algebra operations over the neurons' covariance matrices (see Sec. IV) [53]. Using these exact formulas, we examined how the mutual information changes as a function of the noise correlation ρ_n for different values of the signal correlation ρ_s [Fig. 3(c)] and of the SNR V_s/V_n [Fig. 3(d)].

The results qualitatively agree with the case of pairs of neurons considered previously. Weak noise correlations impede information transmission, in accordance with the sign rule. However, they become beneficial as they increase past a critical threshold ρ_n^*, and this threshold grows with the signal correlation strength. It also decreases and even vanishes as the SNR is increased [Fig. 3(d)] (see Sec. IV for a discussion of the large-SNR limit). This means that more reliable neurons imply an enhanced benefit of noise correlations. We further proved that, even at low SNR, there always exists a range of noise correlation strengths where noise correlations are beneficial (see Sec. IV). The general dependence of ρ_n^* on the correlation ranges L_s and L_n is shown in Fig. 4.

Based on the analysis of pairs of neurons, we expect inhomogeneities in the SNR V_s/V_n of neurons to enhance the benefit of noise correlations. To study this effect, we let the power of the signal V_s vary between cells, while the noise level V_n is kept constant. Assuming that each cell is assigned a random value of V_s, we can compute the correction to the critical noise correlation ρ_n^*. We find that ρ_n^* decreases at leading order with the magnitude of the inhomogeneity (see Sec. IV). This result confirms that, in large populations of neurons as well, variability among neurons makes it more likely for noise correlations to have a beneficial effect.

D. Spectral decomposition

Mutual information is a single number that provides a global quantification of coding efficiency, but says nothing about what is being transmitted. Likewise, a positive noise synergy indicates that noise correlations are beneficial overall, but it does not tell us what feature of the stimulus is better encoded, nor which specific interactions between signal and noise allow for that benefit. We wondered what features of the signal were enhanced by strong positive noise correlations in our population encoding model.

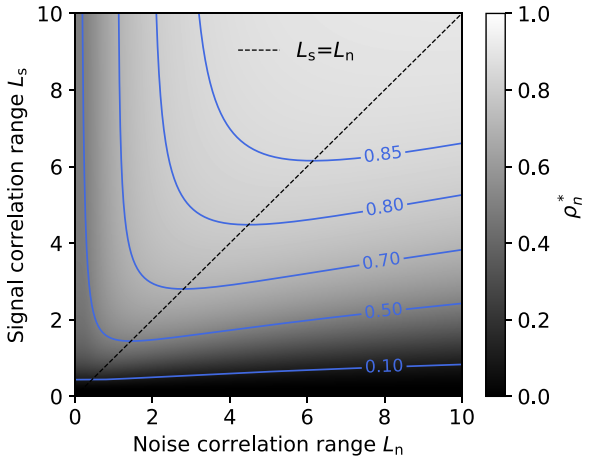

FIG. 4. Behavior of ρ_n^*. Here ρ_n^* changes nonmonotonically with the signal L_s and noise L_n correlation ranges and is concave with respect to these parameters. The maximum value of ρ_n^* at a given L_s is achieved when $L_n = L_s$.

Due to the translation-invariant structure of the model, the mutual information and noise synergies may be decomposed spectrally as a sum over spatial frequencies k (expressed in units of inverse distance between nearest neighbors),

$$\Delta I = \frac{n}{2}\int_{-1/2}^{1/2} \mathrm{d}k \log\left(\frac{1+S(k)/N(k)}{1+S(k)/V_n}\right), \tag{3}$$

where $S(k)$ is the power spectrum of the stimulus, $N(k)$ is that of the noise (see Sec. IV), and $n \to \infty$ is the total number of neurons. In this decomposition, low frequencies correspond to long-range collective modes, while high frequencies correspond to fine-grain features.

Natural stimuli involve spatially extended features impacting many neurons. This causes neural responses to exhibit strong long-range signal correlations between neurons, corresponding in our model to large L_s [Fig. 3(b)]. Most information is then carried by low-frequency modes of the response [Fig. 5(a)].

Noise correlations concentrate noise power at low frequencies and decrease noise power at high frequencies for a fixed noise level V_n [inset of Fig. 5(b)]. As a result, noise correlations enhance information in the high-frequency modes of the signal ($k \geqslant k^*$), at the expense of the low-frequency features [Fig. 5(b)], which are already well represented. Figure 5(c) shows the spectral decomposition of the noise synergy as a function of the noise correlation range L_n. The critical frequency $k^* = (1/2\pi)\arccos(e^{-1/L_n})$ above which noise correlations are beneficial only depends on L_n [Fig. 5(c)]. However, the relative information gains in each frequency domain depends on the strengths of the signal and noise correlations.

In summary, noise correlations enhance fine details of the stimulus to the detriment of its broad features, which are already sufficiently well encoded. This redistribution of the noise across the spectrum drives the gain in information. This effect is generic to any choice of the correlation lengths, and

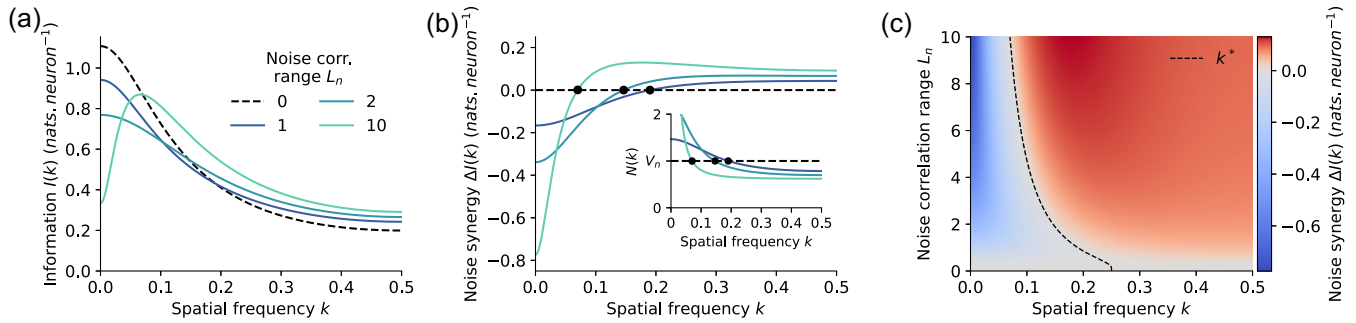

FIG. 5. Spectral analysis of stimulus information encoding. (a) Spatial spectral decomposition $I(k)$ of the mutual information between stimulus and response for a system with SNR $= 2$, $L_s = 2$, and $\rho_n = 0.4$ for various ranges of the noise correlations ($L_n = 0$ corresponds to the absence of noise correlations). (b) Spectral decomposition of the noise synergy $\Delta I(k) = \log\{[1 + S(k)/N(k)]/[1 + S(k)/V_n]\}$. The inset shows the power spectrum of the noise. (c) Heatmap showing the noise synergy spectral decomposition as a function of the noise correlation range L_n. The critical spatial frequency k^* above which noise correlations are beneficial is shown as a black dotted line.

we expect it to hold for other forms of the power spectra and receptive field geometries.

E. Noise correlations in the retina favor the encoding of fine stimulus details

To test our predictions, we studied experimentally the impact of noise correlations on the encoding of features at different spatial scales in the retina. We recorded *ex vivo* the spiking activity of seven OFF-α retinal ganglion cells from a mouse retina using the same experimental technique as described before. We presented the retina with a multiscale checkerboard stimulus composed of frames made of random black and white checkers, flashed at 4 Hz. Each frame was made of a checkerboard with a given spatial resolution (checks of sizes 12, 24, 36, 72 and 108 μm). From the recorded activity, we inferred an artificial convolutional neural network (CNN) with interactions [50] and used the inferred model to build a large synthetic population of 49 cells organized on a triangular lattice [Fig. 6(a)]. Previous studies have demonstrated the ability of CNNs to capture the stimulus response of RGCs accurately [54–58]. We checked that the resulting correlation structure was consistent with previous reports, following an exponential decay as a function of distance with a characteristic length of 110 μm (Fig. 7), comparable to observations of 170–290 μm in the rat [20]. We then generated a large dataset of repeated responses to regular black and white checker flashes. Each checker was composed of checks of a given size (sizes ranging from 140 to 420 μm, with 28-μm increments) and for each check size, 50 spatially offset versions of the checker were showed. We trained a linear decoder of each pixel value (black or white) on this synthetic dataset and a second decoder on the synthetic data in which the activity of each cell was shuffled across repetitions to destroy noise correlations (see Sec. IV).

The two decoders were then applied to the testing datasets, synthetically generated in the same way as the training sets, to decode each pixel from the response. For a fair comparison, the second decoder was applied to data in which noise correlations were removed by shuffling, as in the training.

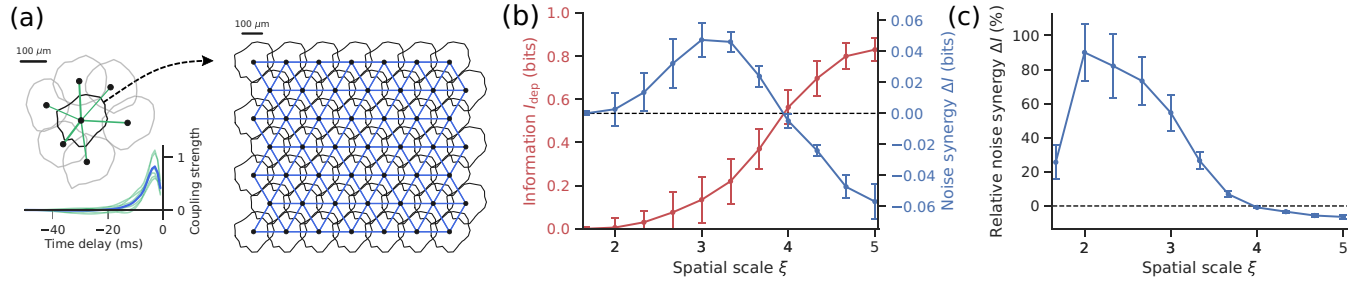

FIG. 6. Noise correlations benefit small-scale features at the detriment of large-scale ones. (a) We built a large population of 49 RGCs based on seven neurons recorded from the mouse retina. A deep GLM [50] was fit to the experimental population and its central neuron model was tiled on a triangular lattice to create a large RGC network. Couplings between the central experimental cell and its neighbors were symmetrized (green links in the population plot and green lines in the inset plot) and averaged to obtain the coupling filters between nearest neighbors in the synthetic population (blue links in the synthetic mosaic and blue lines in the inset). (b) Information I_{dep} and noise synergy ΔI per pixel for stimulus features of increasing scales ($\xi = 2$ check sizes in units of interneuron distance). These quantities were computed via a decoding approach applied to a binary flashed checkerboard stimulus with various check sizes. Error bars are the standard error obtained by repeating the analysis on bootstrapped data. In the absence of noise correlations, little information is transmitted about small stimulus features. By contrast, large-scale features are well encoded and information per pixel saturates towards one bit as check size grows. The noise synergy is positive for small and intermediate check sizes and negative for larger checks, in line with the theoretical results highlighted in Fig. 5. (c) Noise correlations nearly double the amount of information encoded about stimulus features of small and intermediate sizes, while only decreasing information for the largest checks by less than 10%.

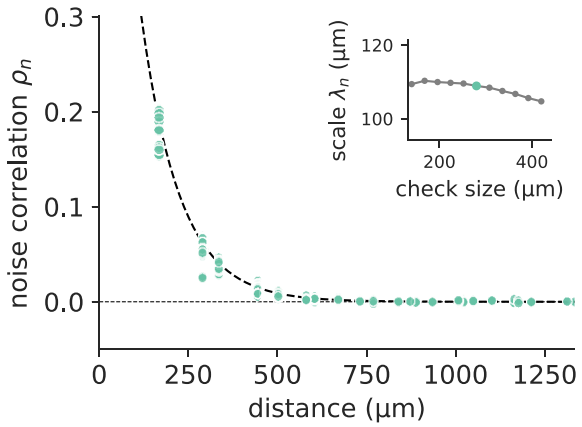

FIG. 7. Correlation structure in augmented data. Noise correlations in data were generated from the convolutional neural network model (obtained with a time bin of 50 ms). An exponential fit as a function of distance between neurons returns a robust correlation length of 110 μm.

The mutual information carried by the decoders was then estimated separately for each checker size. To limit border effects, the mutual information was estimated for each pixel within a small hexagon centered on the central cell of the synthetic population, of size (distance between opposite sides of the hexagon) equal to the distance between cells.

We found that the gain in mutual information afforded by noise correlations is large and positive for small and intermediate check sizes, while moderately negative for large checks [Figs. 6(b) and 6(c)]. These results suggest that noise correlations benefit the encoding of small-scale features of the stimulus, at the expense of the large-scale ones, which are easier to encode. Noise correlations can therefore trade the encoding power of large-scale features to improve sensitivity to the small-scale ones.

III. DISCUSSION

Many experimental works have shown that neurons with the strongest positive noise correlations are similarly tuned to the stimulus [1,5,12,23,27,28,30–32]. Here the sign rule [7,9,37] would predict a detrimental effect of shared variability, at odds with the efficient coding hypothesis [59], which is supported by a large body of work showing that noise correlations are indeed beneficial [15,18,20,39,40,43]. Our work resolves this inconsistency by showing that beyond a critical value ρ_n^*, noise correlations become beneficial to information encoding regardless of their sign. We experimentally demonstrated this effect in recordings of retinal neurons subject to stimuli with different statistics and showed that it generalizes to large populations of sensory neurons. The effect depends on the spatial scale: Large-scale (low-dimensional) modes give rise to strong signal correlations, making positive noise correlations detrimental, while small-scale (high-dimensional) modes benefit from positive noise correlations since their signal correlations are small. These small-scale features correspond to high frequencies, and thus to local differential features, consistent with the analysis on pairs of neurons showing that accuracy in the difference of neural activities is enhanced by noise correlations [Fig. 1(i)].

Since much theoretical work has been done on the impact of noise correlations, it is useful to emphasize the different ways in which our approach differs from previous studies. First, we consider nonlinear effects in the noise parameter [Eq. (1)]. Crucially, the second-order term is responsible for the violation of the sign rule. Second, we study mutual information rather than Fisher information. However, most of our results also hold for the Fisher information (see Appendix B). Third, we used constant stimulus-independent noise correlations, in the sense of Pearson. However, we also showed that, counterintuitively, strong noise correlations could be beneficial even in the case of information-limiting correlations [38,52] (see Appendix A and Fig. 8), which is the most unfavorable case of stimulus-dependent noise correlation. More generally, we considered the impact of such stimulus-dependent noise correlations within our framework (see Appendix A) and showed that these fluctuations can improve the noise synergy in two ways: by being large and by being correlated with the noise level $V_n(\theta)$, also assumed to be stimulus dependent. Our results thus extend and clarify previous theoretical work [35,60] under a common information-theoretic framework. Finally, relative to Ref. [36], we take the large population limit by considering large networks with finite correlation length, in units of interneuron separation, rather than by increasing the density of neurons; that scaling makes it possible to uncover the benefit of noise correlations at short length scales.

Our spectral analysis of scales and dimensionality helps understand apparently contradictory claims in the literature. On the one hand, noise correlation should be detrimental for coding, because it impedes denoising by pooling the signal of multiple neurons [1,36]. On the other hand, studies focusing on the stimulus response of large sensory populations have observed a positive gain [15,18,20,40,45]. Our study suggests the following interpretation: When the neural population encodes a low-dimensional stimulus, as the angle of a drifting gratings, similarly tuned nearby neurons become strongly signal correlated and their noise correlations are detrimental [36]. In the case of high-dimensional stimuli, like naturalistic images or videos, signal correlations between them are positive but weak, so that noise correlations become larger than the threshold ρ_n^* and therefore beneficial.

Our work also shares some similarity with Ref. [41], where the authors predicted the optimal patterns of noise correlations maximizing information transmission by a population of neurons. They showed that at high SNR, optimal noise correlations follow the sign rule. This result does not rule out that high levels of noise correlations violating the sign rule could be beneficial (albeit not optimal), in agreement with our theory. However, a direct comparison with our results is difficult because in Ref. [41] noise correlations were tuned through interneuron couplings that affect the mean response of each neuron to the stimulus, which is kept constant in our analysis. In fact, this effect leads to optimal noise correlations of the same sign as signal correlations in the low-SNR regime [41]. It was also shown to improve positional coding in the hippocampus through the sharpening of stimulus tuning [42]. This apparent violation of the sign rule is however indirect

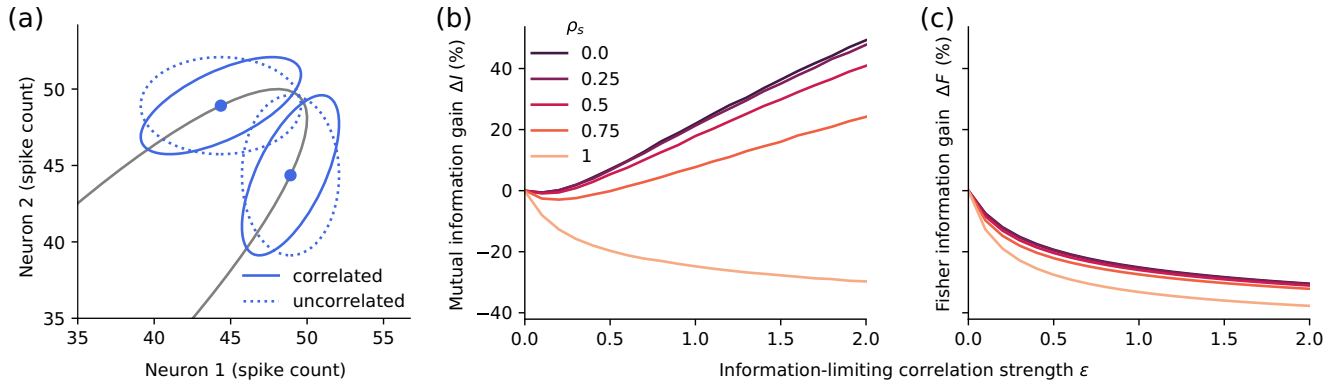

FIG. 8. Impact of information-limiting correlations on stimulus information. An angular stimulus θ is encoded by a pair of neurons characterized by von Mises tuning curves (with parameters $a = 40$, $b = 10$, and $w = 5$). Their response is Gaussian of means $\mu_1(\theta)$ and $\mu_2(\theta)$. Information-limiting correlations are defined by a covariance of the form: $\Sigma_n(\theta) = V_0\mathbb{I} + \epsilon\mu'(\theta)\mu'(\theta)^\top$, were ϵ controls their strength, and where we set $V_0 = V_s/2$. Note that Σ_n now depends on θ. (a) Noise covariance ellipses for the correlated (plain line) and uncorrelated (doted line) cases, for two example stimuli (blue points) on the manifold (grey), for an example of the system described above (with $\rho_s \approx 0.9$). Here, information limiting correlations will decrease the overlap between the two conditional response distributions. The mutual information accounts not only for the detrimental local effects of information-limiting correlations, but also for such potentially beneficial non-linear and global effects. (b) Relative mutual information gain (noise synergy) $\Delta I = I_{dep}/I_{indep} - 1$ as a function of ϵ, where I_{dep} and I_{indep} quantify the mutual information with and without (off diagonal terms of Σ_n set to 0) noise correlations, for different levels of signal correlation ρ_s. Mutual information was computed via Monte Carlo integration. Information-limiting noise correlations become beneficial to the mutual information if they are strong enough, except when cells are perfectly signal correlated. (c) By contrast, information-limiting correlations are always detrimental to the Fisher information $F(\theta) = \mu'(\theta)^\top \Sigma_n^{-1}(\theta)\mu'(\theta)$. The relative Fisher information gain $\Delta F = \langle F_{dep}(\theta)/F_{indep}(\theta) - 1\rangle_\theta$, where $F_{dep}(\theta)$ and $F_{indep}(\theta)$ denote the Fisher information with and without noise correlations, is always negative and decreases with ϵ and ρ_s.

and distinct from the direct beneficial effect of strong noise correlations that we discuss in this work.

Several studies have focused on the effect of noise correlations on the Fisher information [4,18,36,39,40]. While our main results are based on the mutual information, they equivalently apply to the Fisher information in the Gaussian case [33] (see Appendix B). To further test the robustness of our conclusions, we demonstrated that our results are model independent and hold for both binary and Gaussian neurons. In addition, empirical results from the retinal recordings [Fig. 1(j)] were obtained without any approximation or model choice and agree with the theory.

We validated our theoretical predictions experimentally on recordings of neurons from the retina. Our theory could be further tested in large-scale retinal recordings in which the population size is large enough to observe network effects, by presenting stimuli in which signal correlations are weak or tunable. An interesting question is whether the magnitude of signal correlations in natural scenes is consistent with a beneficial effect of noise correlations, in line with arguments of efficient coding in the retina [61,62]. Applying our approach to data in sensory cortical areas where similar noise correlation structures have been observed [4,5] could lead to new understanding of the role of noise correlations in sensory information processing. Another key open question is what stimulus ensembles most benefit from noise correlations, as well as where naturalistic stimuli stand in that regard. We have further shown that noise correlations benefit the encoding of high-frequency features of the stimulus, which correspond to fine-grained neural activity patterns. Extending these results to higher cortical areas would require understanding which features from the stimulus drive such activity patterns.

IV. METHODS

A. Covariance and correlation measures

The average responses of two neurons 1 and 2 are given as function of the stimulus θ by the tuning curves $\mu_1(\theta) = \langle r_1\rangle_\theta$ and $\mu_2(\theta) = \langle r_2\rangle_\theta$. Signal correlations are defined as $\rho_s = \mathrm{Corr}_\theta(\mu_1, \mu_2)$ and noise correlations as $\rho_n(\theta) = \mathrm{Corr}(r_1, r_2|\theta)$. The sum of these two coefficients does not have a simple interpretation in terms of total correlation or covariance, but we can also decompose the total correlation coefficient between r_1 and r_2 as $\mathrm{Corr}(r_1, r_2) = r_s + r_n$, with $r_s = \mathrm{Cov}_\theta(\mu_1, \mu_2)/\sqrt{\mathrm{Var}(r_1)\mathrm{Var}(r_2)}$ and $r_n = \langle\mathrm{Cov}(r_1, r_2|\theta)\rangle_\theta/\sqrt{\mathrm{Var}(r_1)\mathrm{Var}(r_2)}$.

B. Pairwise analysis

1. Tuning curves

We consider a pair of neurons encoding an angle θ. The responses of the two neurons, r_1 and r_2, are assumed to be binary (spike or no spike in a 10-ms time window) and correlated. Their average responses $\mu_1(\theta)$ and $\mu_2(\theta)$ are given by von Mises functions [Fig. 1(a)]

$$\mu_i(\theta) = a\frac{\exp\left[\cos\left(\theta - \theta_c^i\right)/w\right] - \exp(-1/w)}{\exp(1/w) - \exp(-1/w)} + b. \quad (4)$$

Signal correlations between the two neurons can be tuned by varying the distance between the center of the two tuning curves θ_c^1 and θ_c^2. The tuning curve width w was set arbitrarily to 5, the amplitude a to 0.4, and the baseline b to 0.1. The strength of noise correlations is set to a constant of θ, $\rho_n(\theta) = \rho_n$.

2. Small correlation expansion

When noise correlations ρ_n are constant and small, the noise synergy may be expanded as [46]

$$\Delta I \approx -r_s r_s + \frac{1}{2}\left(\rho_n^2 - r_n^2\right) = \frac{\alpha}{2}\rho_n(\rho_n - \rho_n^*), \tag{5}$$

where the second equality highlights the dependence on ρ_n. The critical ρ_n^* may be written as

$$\rho_n^* = \beta\rho_s, \tag{6}$$

with $\beta = \frac{2V_sV_n}{V_{tot}^2 - V_n^2}$, and the prefactor $\alpha = 1 - V_n^2/V_{tot}^2$, with the shorthand $V_{tot} = \sqrt{\mathrm{Var}(r_1)\mathrm{Var}(r_2)}$, $V_n = \langle\sqrt{\mathrm{Var}(r_1|\theta)\mathrm{Var}(r_2|\theta)}\rangle_\theta$, and $V_s = \sqrt{\mathrm{Var}[\mu_1(\theta)]\mathrm{Var}[\mu_2(\theta)]}$ corresponding to measures of total, noise, and signal variances in the two cells.

By Cauchy-Schwartz inequality we have

$$V_n^2 \leqslant \langle\mathrm{Var}(r_1|\theta)\rangle_\theta\langle\mathrm{Var}(r_2|\theta)\rangle_\theta, \tag{7}$$

which entails

$$\beta \lesssim \frac{1}{\cosh\frac{\Delta\ln R}{2} + R/2} \leqslant 1, \tag{8}$$

where $R = \sqrt{R_1R_2}$ and $\Delta\ln R = \ln(R_1/R_2)$, with $R_i = \mathrm{Var}(\mu_i)/\langle\mathrm{Var}(r_i|\theta)\rangle_\theta$ the signal-to-noise ratio of the cells; R measures the overall strength of signal-to-noise ratios, while $\Delta\ln R$ measures their dissimilarity. The last inequality implies that noise correlations are always beneficial for $\rho_n > \rho_s$.

3. Gaussian case

To test the theory's robustness to modeling choices, we also considered a continuous rather than binary neural response, $r_i = \mu_i(\theta) + \delta r_i$, where both μ_i and δr_i are Gaussian variables defined by their covariance matrices $\Sigma_{s,ij} = \mathrm{Cov}_\theta(\mu_i, \mu_j)$ and $\Sigma_{n,ij} = \langle\mathrm{Cov}(r_i, r_j|\theta)\rangle_\theta$. The noise synergy can be calculated through classic formulas for the entropy for Gaussian variables, yielding

$$\Delta I = \frac{1}{2}\log\left(\frac{|\Sigma_s + \Sigma_n||V_n|}{|\Sigma_s + V_n||\Sigma_n|}\right), \tag{9}$$

where $|X|$ denotes the determinant of matrix X and V_n is the diagonal matrix containing the noise variances of the cells $V_{n,ii} = \Sigma_{n,ii}$. Note that this formula is general for an arbitrary number of correlated neurons. In the pairwise case considered here, matrices are of size 2×2. The condition for beneficial noise correlations $\Delta I \geqslant 0$ is satisfied for $\rho_n \geqslant \rho_n^*$, with

$$\rho_n^* = \beta\rho_s, \tag{10}$$

with

$$\beta = \frac{1}{\cosh\frac{\Delta\ln R}{2} + \left(1 - \rho_s^2\right)R/2} \leqslant 1,$$

which has a similar form to Eq. (8).

4. Noise synergy at constant noise entropy

Increasing noise correlations at constant V_n decreases the effective variability of the response, as measured by the noise entropy $H(\{r_1, r_2\}|\theta) = \ln(2\pi e|\Sigma_n|^{1/2})$, with $|\Sigma_n| = V_n^2(1 - \rho_n^2)$ in the case two neurons with the same noise level. To correct for this effect, we also computed ΔI at constant noise entropy, by rescaling the noise variances in the correlated and uncorrelated cases $V_{n,c}$ and $V_{n,u}$, so that their resulting noise entropies are equal $|\Sigma_n| = V_{n,c}^2(1 - \rho_n^2) = V_{n,u}^2$. The critical noise correlation at which $\Delta I \geqslant 0$ is then given by

$$\rho_n^* = 2\frac{\rho_s}{1 + \rho_s^2} \leqslant 1, \tag{11}$$

where the last inequality implies that strong enough noise correlations are always beneficial.

5. Retinal data

Retinal data were recorded *ex vivo* from a rat retina using a microelectrode array [47] and sorted using SPYKING CIRCUS [49] to isolate single neuron spike trains. From the ensemble of single cells, we could isolate a population of 32 OFF-α ganglion cells. Three stimulus movies with different spatiotemporal statistics were presented to the retina: a checkerboard movie consisting of black and white checks changing color randomly at 40 Hz and repeated 79 times; a drifting grating movie consisting of black and white stripes of width 333 μm moving in a fixed direction relatively to the retina, at speed 1 mm/s, and repeated 120 times; and finally a movie composed of ten black disks jittering according to a Brownian motion on a white background, repeated 54 times.

C. Gaussian population and spectral analysis

We consider a population of n neurons organized along a one-dimensional lattice with constant interneuron spacing. Their mean response and noise are assumed to be Gaussian, with their noise and signal covariances given by an exponentially decaying function of their pairwise distances:

$$\Sigma_{s,ij} = V_s e^{-|i-j|/L_s}, \tag{12}$$

$$\Sigma_{n,ij} = V_n\left(\delta_{ij} + \rho_n^0 e^{-|i-j|/L_n}\right). \tag{13}$$

Here V_s and V_n are the signal and noise variance of the single cells. The parameter ρ_n^0 sets the strength of noise correlations such that nearest neighbors have noise correlation $\rho_n \equiv \rho_n^0\exp(-1/L_n)$. When n is large and boundary effects can be ignored, the system is invariant by translation and we can diagonalize Σ_s and Σ_n in the Fourier basis $v_{k,l} = \frac{1}{\sqrt{n}}\exp(-i2\pi kl/n)$. Denoting the spectra of Σ_s and Σ_n by $S(l/n)$ and $N(l/n)$, the expression of the noise synergy [Eq. (9)] can then be written as a sum over modes

$$\Delta I = \frac{1}{2}\sum_{l=-(n-1)/2}^{(n-1)/2}\log\left(\frac{1 + S(l/n)/N(l/n)}{1 + S(l/n)/V_n}\right), \tag{14}$$

which simplifies in the $n \to \infty$ limit to

$$\frac{\Delta I}{n} = \frac{1}{2}\int_{-1/2}^{1/2}\log\left(\frac{1 + S(k)/N(k)}{1 + S(k)/V_n}\right)\mathrm{d}k, \tag{15}$$

with

$$S(k) = V_s\frac{1 - \rho_s^2}{1 - 2\rho_s\cos(2\pi k) + \rho_s^2}, \tag{16}$$

$$N(k) = V_n\left(1 - \rho_n^0 + \rho_n^0\frac{1 - \lambda_n^2}{1 - 2\lambda_n\cos(2\pi k) + \lambda_n^2}\right), \tag{17}$$

where $\rho_s = \exp(-1/L_s)$ is the nearest-neighbor signal correlation, $\lambda_n = \exp(-1/L_n)$, and k is a wave vector interpretable as a spatial frequency in units of the system's size, up to a 2π factor. Examining Eq. (15), we see that noise correlations are beneficial for frequencies for which $N(k) \leqslant V_n$, which happens for $k \geqslant k^*$, where $k^* = (1/2\pi)\arccos(e^{-1/L_n})$.

In the low-noise regime $R = V_s/V_n \gg 1$, the noise synergy reduces to

$$\frac{\Delta I}{n} \approx -\frac{1}{2}\int_{-1/2}^{1/2} \log[N(k)/V_n]dk \geqslant 0, \tag{18}$$

where the inequality stems from Jensen's inequality, because $-\log$ is a convex function, and $-\log[\int_{-1/2}^{1/2} dkN(k)/V_n] = 0$. Therefore, in that regime, noise correlations are always beneficial.

In the high-noise limit $R \ll 1$, the noise synergy becomes

$$\frac{\Delta I}{n} \approx \frac{1}{2}\int_{-1/2}^{1/2}\left(\frac{S(k)}{N(k)} - \frac{S(k)}{V_n}\right)dk. \tag{19}$$

Computing this integral gives the critical noise correlation

$$\rho_n^* = \rho_s \frac{1-\lambda_n^2}{1-2\lambda_n\rho_s+\rho_s^2} \leqslant \rho_n^{\max}, \tag{20}$$

where $\rho_n^{\max} = (1+\lambda_n)/2$ is the maximum possible value of ρ_n [ensuring that the noise spectrum $N(k)$ is non-negative for all k]. The last inequality in Eq. (20) implies that there always exists a regime in which strong noise correlations are beneficial.

Nonidentical neurons

To study the effect of inhomogeneities among neurons, we considered the case where the signal variance of each cell is different and drawn at random as $\sqrt{V_s^i} = \mu + \eta_i$, where η_i is normally distributed with zero mean and variance ν^2. The noise synergy can be rewritten in the high-noise regime ($R \ll 1$) as

$$\Delta I \approx \tfrac{1}{2}\mathrm{Tr}\left(\Sigma_s\Sigma_n^{-1} - \Sigma_s V_n^{-1}\right). \tag{21}$$

Averaging this expression over η_i yields

$$\Delta I \approx \Delta I_u + \frac{1}{2}\frac{\nu^2}{\mu^2+\nu^2}\mathrm{Tr}\left[\Sigma_n^{-1}(\bar{R}\mathbb{I} - \Sigma_s)\right], \tag{22}$$

where ΔI_u is the noise synergy in a uniform population (with $V_s = \mu^2 + \nu^2$) and the second term is always positive, with $\bar{R} = \langle V_s^i\rangle/V_n = (\mu^2+\nu^2)/V_n$.

Taking the continuous limit ($n \to \infty$) in Eq. (22), similarly to the integral limit of Eq. (15), allows us to write the critical noise correlation ρ_n^* as

$$\rho_n^* = \frac{\rho_n^{*,u}}{1+\gamma} + \frac{1-\rho_n^{*,u}}{2}\left(1 - \sqrt{1 + \frac{4\gamma\rho_s^2(1-\lambda_n^2)}{(1+\gamma)^2(1-\lambda_n\rho_s)^2}}\right), \tag{23}$$

where $\gamma = \nu^2/\mu^2$ quantifies the relative magnitude of inhomogeneities and $\rho_n^{*,u}$ is the critical noise correlation value in a uniform population [Eq. (20)]. This modified critical noise correlation value is always smaller than in the uniform case and scales linearly at leading order with the inhomogeneity parameter γ:

$$\rho_n^* = \rho_n^{*,u}\left(1 - \gamma\frac{1-\rho_s^2}{1-\rho_s\lambda_n}\right) + o(\gamma). \tag{24}$$

D. Decoding analysis

1. Experimental and synthetic data

We presented a mouse retina with a stimulus consisting of a black and white random checkerboard flashed at 4 Hz, each frame with a given spatial resolution (checks of sizes 12, 24, 36, 72, and 108 μm). Retinal ganglion cell activity was recorded *ex vivo* using a microelectrode array and single-neuron activity isolated via spike sorting using SPYKING CIRCUS [49]. We isolated a population of $N_{\text{cells}} = 7$ OFF-α retinal ganglion cells, which presents strong noise correlations in their response [17]. The original recording contained a 15-s checkerboard movie repeated 90 times as well as 90 different 22.5-s-long unrepeated movies.

We inferred a deep generalized linear model (GLM) of the central cell among seven from the experimental population [Fig. 6(a)], consisting of a stimulus-processing filter, and filters for the spiking history of the cell as well as its six neighbors (couplings). The stimulus-processing part of the model consisted of a deep neural network composed of two spatiotemporal convolutional layers followed by a readout layer. The whole model was fit to the data using the two-step inference approach [50].

A synthetic population of 49 OFF-α ganglion cells was then constructed by arranging them on a triangular lattice of 7×7 points. Each cell responds according to the inferred GLM with translated receptive fields. Nearest neighbors were coupled with the average of the GLM couplings inferred between the central cell from the experiment and its neighbors.

To stimulate this synthetic population, we generated a synthetic stimulus ensemble from 550 regular black and white checker frames, each with a given check size ranging from 140 to 420 μm (with increments of 28 μm). Every checker of a given size was presented for five different regularly spaced offsets ranging from 0 to 224 μm in both the horizontal and vertical directions, resulting in 25 different frames per size. To further ensure that the color of each pixel in the stimulus ensemble is black or white with equal probability, each checker frame also had its color-reversed version in the set, resulting in 50 different frames for a given check size. A single snippet from the synthetic stimulus ensemble consisted of a 250-ms white frame followed by one of the 550 aforementioned checker frames.

We built a training, a validation, and testing set for the dependent and independent decoders by simulating the synthetic population for sets of 3750, 1250, and 5000 repetitions (respectively) of each synthetic stimulus snippets.

2. Decoders

The binary decoders are logistic regressors taking in the integrated response of the population over the $N_\tau = 5$ past bins of 50 ms to predict the ongoing stimulus frame. The

predicted stimulus at time t and repeat k is given by

$$\hat{X}(x, y, t, k) = f(A_{x,y} r(t, k) + \beta_{x,y}), \tag{25}$$

where x and y are the pixel indices along the two dimensions of the stimulus, $f(x) = (1 + e^{-x})^{-1}$ is the sigmoidal function, $A_{x,y}$ is a matrix of size $(N_\tau, N_{\text{cells}})$, $r(t, k)$ is a matrix of size $(N_{\text{cells}}, N_\tau)$ containing the spike history of the population at time t and repeat k, and $\beta_{x,y}$ is a pixelwise bias. Each decoder was trained by minimizing the average binary cross entropy (BCE) between predicted the stimulus $\hat{X}(x, y, t, k)$ and the true stimulus $X(x, y, t)$, $\langle \text{BCE}(x, y, t, k)\rangle_{x,y,t,k}$, where

$$\begin{aligned}\text{BCE}(x, y, t, k) = &- X(x, y, t) \ln [\hat{X}(x, y, t, k)] \\ &- [1 - X(x, y, t)] \ln [1 - \hat{X}(x, y, t, k)].\end{aligned} \tag{26}$$

Training was done by stochastic gradient descent on the synthetic datasets using the training (3750 repetitions) and validation (1250 repetitions) sets. Optimization was done using stochastic gradient descent with momentum, with early stopping when the validation loss did not improve over six consecutive epochs. During that procedure the learning rate was divided by 4 whenever the validation loss did not improve for three consecutive epochs.

We probed the decoders' abilities to decode features of different spatial scales by decoding the simulated responses of the synthetic population to the checker stimuli with varying check size from the testing set. Performances of the decoders were assessed by computing the mutual information between each pixel's color X and its decoded value $\hat{X}$, separately for the different sizes of checks. The noise synergy was then computed as the difference between the mutual information averaged over pixels for the dependent and independent decoders.

Error bars were computed as follows. We inferred ten deep GLMs on bootstraps of the original training set, obtained by resampling with replacement the stimulus-response pairs used for training. These ten models were used to generate ten surrogate training sets, from which ten separate decoders were inferred with noise correlations and another ten without noise correlations. Then synthetic test sets for the checker decoding task were generated from each of the ten models and the performance of each decoder was computed separately with and without noise correlations, yielding ten values of the mutual information and ten values of the noise synergy (both averaged over pixels). The error bars are the standard deviations of the resulting information, noise synergy, and synergy-to-information ratios (i.e., relative noise synergy) over the ten bootstraps.

ACKNOWLEDGMENTS

We thank Stéphane Deny for help with the experimental data used in the paper and Matthew Chalk and Simone Azeglio for useful discussions. This work was supported by ANR Grants No. ANR-22-CE37-0023 LOCONNECT and No. ANR-21-CE37-0024 NatNetNoise, by IHU FOReSIGHT (Grant No. ANR-18-IAHU-01), and by Sorbonne Center for Artificial Intelligence, Sorbonne University, IDEX SUPER Grant No. 11-IDEX-0004. This work was also supported by the Bettencourt Schueller Foundation.

DATA AVAILABILITY

Some of the data that support the findings of this article are openly available [63]. Data from [47] are not publicly available because they are owned by a third party and the terms of use prevent public distribution. The data are available from the authors upon reasonable request.

APPENDIX A: STIMULUS-DEPENDENT CORRELATIONS

Theoretical developments in the present work focused on the effect of constant noise correlation (in the sense of Pearson) on the mutual information between stimulus and response. In reality, noise correlations, i.e., correlations of the response conditioned on the stimulus, may also depend on the stimulus value. This dependence may arise via a dependence on either the firing rates of the cells [64] or the stimulus itself [18,65]. Although a complete discussion of the effect of nonconstant noise correlations on mutual information is out of this paper's scope, we succinctly extend our framework to give intuition about this phenomenon.

Starting again from the framework described in Sec. IV B and relaxing the assumption of constant noise correlations, we can show [Eq. (A7) in Ref. [46]] that the noise synergy expanded and truncated at second order in the correlations for a pair of cells encoding for stimulus θ is given, with the definitions from Sec. IV, by

$$\Delta I = -r_s r_n + \tfrac{1}{2}\left[\left\langle \rho_n^2(\theta)\right\rangle_\theta - r_n^2\right]. \tag{A1}$$

We can decompose this information gain as

$$\Delta I = \Delta I_c + \Delta I_{f,1} + \Delta I_{f,2}, \tag{A2}$$

where ΔI_c is given by Eq. (5) with constant noise correlations $\bar{\rho}_n = \langle \rho_n(\theta)\rangle_\theta$ and the two other terms contain the effects of the variations of $\rho_n(\theta)$. The first term $\Delta I_{f,1} = \frac{1}{2}\text{Var}_\theta[\rho_n(\theta)] \geqslant 0$ directly accounts for the effect of fluctuations of $\rho_n(\theta)$, while $\Delta I_{f,2}$ is linked to its correlation with the noise variance:

$$\begin{aligned}\Delta I_{f,2} = &- \left\langle [\rho_n(\theta) - \bar{\rho}_n] \frac{V_n(\theta)}{V_{\text{tot}}} \right\rangle_\theta \\ &\times \left(\frac{1}{2} \left\langle [\rho_n(\theta) + \bar{\rho}_n] \frac{V_n(\theta)}{V_{\text{tot}}} \right\rangle_\theta + r_s \right).\end{aligned} \tag{A3}$$

This contribution is zero when noise correlations $\rho_n(\theta)$ are not correlated to the noise variance of the pair $V_n(\theta)$, but it can in general be positive or negative depending on how these two quantities covary. We saw in the analysis in Fig. 1 that the effect of strong positive noise correlations between OFF-α cells from the rat retina can be qualitatively explained by the second-order approximation with constant noise correlations. This suggests that, at least for this type of neuron, the contribution of noise correlation fluctuations to the mutual information does not constitute the bulk of the effect.

Previous work investigated the effect of nonconstant noise correlations on sensory coding. When focusing on the mutual information [35,60], studies essentially relied on decompositions of the mutual information derived from correlation

measures that stray from classical Pearson correlations and their intuitive meaning (see Appendix D in Ref. [46] for a discussion on that point). As a consequence, interpretation of their results in terms of pairwise noise correlations is not straightforward.

Another line of work derived the notion of information-limiting correlations [38,52], which are correlations that always align noise with the signal's direction, at any point of the manifold. Such correlations, however, are only truly information limiting in the case of the Fisher information (they were in fact derived from the linear approximation of the Fisher information [38]). The Fisher information only accounts for how an ideal decoder can distinguish two infinitely close stimuli θ and $\theta + \mathrm{d}\theta$ with vanishing $\mathrm{d}\theta$. It thus relies on a linear analysis (first order in a small error parameter) and ignores higher-order effects. In contrast, the mutual information is a global measure of information that accounts for how the conditional response distributions are distinguishable from the marginal response distribution and includes all orders of the error parameter. While information-limiting correlation are always detrimental to the Fisher information, they can benefit the mutual information by decreasing the overlap between conditionals. Figure 8(a) illustrates an example of such an effect, in which the high curvature of the response curve causes information-limiting correlations to actually increase the discriminability of nearby stimuli. Since it is a function of the curvature, it is a nonlinear effect that cannot be captured by the Fisher information, but it does impact the mutual information.

Figure 8(b) quantifies the effect on the mutual information in a concrete case with von Mises tuning curves (see the caption for details). It shows that strong information-limiting correlations are always beneficial to the mutual information, while always being detrimental to the Fisher information [Fig. 8(c)]. One must note, however, that in the particular case of perfectly stimulus correlated neurons ($\rho_s = 1$), information-limiting correlations will always be detrimental, as they will increase the overlap between conditional response distributions and the marginal response distribution.

APPENDIX B: STRONG NOISE CORRELATIONS AND FISHER INFORMATION

We consider a pair of neurons encoding an arbitrary variable θ. The responses r of these neurons are assumed to be Gaussian with mean $\mu(\theta)$ and constant covariance matrix Σ_n, where $\mu(\theta)$ are the tuning curves. In this context, the Fisher information is defined as [38]

$$F_{\mathrm{dep}}(\theta) = \mu'(\theta)^\top \Sigma_n^{-1} \mu'(\theta). \tag{B1}$$

Expanding this expression for a pair of neurons with equal noise variance $\Sigma_n = V_n \begin{pmatrix} 1 & \rho_n \\ \rho_n & 1 \end{pmatrix}$ yields

$$F_{\mathrm{dep}}(\theta) = \frac{\mu_1'(\theta)^2 + \mu_2'(\theta)^2}{V_n(1-\rho_n^2)} \left(1 - \rho_n \frac{2\mu_1'(\theta)\mu_2'(\theta)}{\mu_1'(\theta)^2 + \mu_2'(\theta)^2}\right). \tag{B2}$$

In the absence of noise correlations ($\rho_n = 0$), the Fisher information simplifies to

$$F_{\mathrm{indep}}(\theta) = \frac{\mu_1'(\theta)^2 + \mu_2'(\theta)^2}{V_n}. \tag{B3}$$

To quantify the overall Fisher improvement, we introduce the quantity $\Delta F = \langle F_{\mathrm{dep}}(\theta)/F_{\mathrm{indep}}(\theta) - 1\rangle_\theta$. Defining $\xi(\theta) = \frac{2\mu_1'(\theta)\mu_2'(\theta)}{\mu_1'(\theta)^2+\mu_2'(\theta)^2}$, the Fisher improvement becomes

$$\Delta F = \frac{\rho_n[\rho_n - \langle \xi(\theta)\rangle_\theta]}{1-\rho_n^2}. \tag{B4}$$

Therefore, strong positive noise correlations will benefit the Fisher information whenever they exceed the critical value $\rho_n^{*,F} = \langle \xi(\theta)\rangle_\theta$.

[1] E. Zohary, M. N. Shadlen, and W. T. Newsome, Correlated neuronal discharge rate and its implications for psychophysical performance, Nature (London) **370**, 140 (1994).

[2] M. Meister, L. Lagnado, and D. A. Baylor, Concerted signaling by retinal ganglion cells, Science **270**, 1207 (1995).

[3] T. Gawne and B. Richmond, How independent are the messages carried by adjacent inferior temporal cortical neurons? J. Neurosci. **13**, 2758 (1993).

[4] M. A. Smith and A. Kohn, Spatial and temporal scales of neuronal correlation in primary visual cortex, J. Neurosci. **28**, 12591 (2008).

[5] S. B. Hofer, H. Ko, B. Pichler, J. Vogelstein, H. Ros, H. Zeng, E. Lein, N. A. Lesica, and T. D. Mrsic-Flogel, Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex, Nat. Neurosci. **14**, 1045 (2011).

[6] M. R. Cohen and A. Kohn, Measuring and interpreting neuronal correlations, Nat. Neurosci. **14**, 811 (2011).

[7] B. Averbeck, P. Latham, and A. Pouget, Neural correlations, population coding and computation, Nat. Rev. Neurosci. **7**, 358 (2006).

[8] A. Kohn, R. Coen-Cagli, I. Kanitscheider, and A. Pouget, Correlations and neuronal population information, Annu. Rev. Neurosci. **39**, 237 (2016).

[9] R. A. da Silveira and F. Rieke, The geometry of information coding in correlated neural populations, Annu. Rev. Neurosci. **44**, 403 (2021).

[10] S. Panzeri, M. Moroni, H. Safaai, and C. D. Harvey, The structures and functions of correlations in neural population codes, Nat. Rev. Neurosci. **23**, 551 (2022).

[11] D. H. Perkel, G. L. Gerstein, and G. P. Moore, Neuronal spike trains and stochastic point processes: II. Simultaneous spike trains, Biophys. J. **7**, 419 (1967).

[12] D. N. Mastronarde, Correlated firing of retinal ganglion cells, Trends Neurosci. **12**, 75 (1989).

[13] S. Nirenberg, S. M. Carcieri, A. L. Jacobs, and P. E. Latham, Retinal ganglion cells act largely as independent encoders, Nature (London) **411**, 698 (2001).

[14] J. Shlens, F. Rieke, and E. Chichilnisky, Synchronized firing in the retina, Curr. Opin. Neurobiol. **18**, 396 (2008).

[15] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli, Spatio-temporal correlations and visual signalling in a complete neuronal population, Nature (London) **454**, 995 (2008).
[16] B. Völgyi, S. Chheda, and S. Bloomfield, Tracer coupling patterns of the ganglion cell subtypes in the mouse retina, J. Comp. Neurol. **512**, 664 (2009).
[17] B. Völgyi, F. Pan, D. L. Paul, J. T. Wang, A. D. Huberman, and S. A. Bloomfield, Gap junctions are essential for generating the correlated spike activity of neighboring retinal ganglion cells, PLoS One **8**, e69426 (2013).
[18] F. Franke, M. Fiscella, M. Sevelev, B. Roska, A. Hierlemann, and R. Azeredo da Silveira, Structures of neural correlation and how they favor coding, Neuron **89**, 409 (2016).
[19] J. Zylberberg, J. Cafaro, M. Turner, E. Shea-Brown, and F. Rieke, Direction-selective circuits shape noise to ensure a precise population code, Neuron **89**, 369 (2016).
[20] K. Ruda, J. Zylberberg, and G. D. Field, Ignoring correlated activity causes a failure of retinal population codes, Nat. Commun. **11**, 4605 (2020).
[21] O. Sorochynskyi, S. Deny, O. Marre, and U. Ferrari, Predicting synchronous firing of large neural populations from sequential recordings, PLoS Comput. Biol. **17**, e1008501 (2021).
[22] J. Fiser, C. Chiu, and M. Weliky, Small modulation of ongoing cortical dynamics by sensory input during natural vision, Nature (London) **431**, 573 (2004).
[23] A. Kohn and M. A. Smith, Stimulus dependence of neuronal correlation in primary visual cortex of the macaque, J. Neurosci. **25**, 3661 (2005).
[24] A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias, Decorrelated neuronal firing in cortical microcircuits, Science **327**, 584 (2010).
[25] I. C. Lin, M. Okun, M. Carandini, and K. Harris, The nature of shared cortical variability, Neuron **87**, 644 (2015).
[26] W. M. Usrey and R. C. Reid, Synchronous activity in the visual system, Annu. Rev. Physiol. **61**, 435 (1999).
[27] D. Lee, N. L. Port, W. Kruse, and A. P. Georgopoulos, Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex, J. Neurosci. **18**, 1161 (1998).
[28] W. Bair, E. Zohary, and W. T. Newsome, Correlated firing in macaque visual area MT: Time scales and relationship to behavior, J. Neurosci. **21**, 1676 (2001).
[29] O. Hazon, V. H. Minces, D. P. Tomàs, S. Ganguli, M. J. Schnitzer, and P. E. Jercog, Noise correlations in neural ensemble activity limit the accuracy of hippocampal spatial representations, Nat. Commun. **13**, 4276 (2022).
[30] D. N. Mastronarde, Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X- and Y-cells, J. Neurophysiol. **49**, 303 (1983).
[31] D. A. Gutnisky and V. Dragoi, Adaptive coding of visual information in neural populations, Nature (London) **452**, 220 (2008).
[32] B. B. Averbeck and D. Lee, Effects of noise correlations on information encoding and decoding, J. Neurophysiol. **95**, 3633 (2006).
[33] N. Brunel and J. P. Nadal, Mutual information, Fisher information, and population coding, Neural Comput. **10**, 1731 (1998).
[34] S. Panzeri, A. Treves, S. Schultz, and E. T. Rolls, On decoding the responses of a population of neurons from short time windows, Neural Comput. **11**, 1553 (1999).
[35] G. Pola, A. Thiele, K. Hoffmann, and S. Panzeri, An exact method to quantify the information transmitted by different mechanisms of correlational coding, Network **14**, 35 (2003).
[36] H. Sompolinsky, H. Yoon, K. Kang, and M. Shamir, Population coding in neuronal systems with correlated noise, Phys. Rev. E **64**, 051904 (2001).
[37] Y. Hu, J. Zylberberg, and E. Shea-Brown, The sign rule and beyond: Boundary effects, flexibility, and noise correlations in neural population codes, PLoS Comput. Biol. **10**, e1003469 (2014).
[38] R. Moreno-Bote, J. Beck, I. Kanitscheider, X. Pitkow, P. Latham, and A. Pouget, Information-limiting correlations, Nat. Neurosci. **17**, 1410 (2014).
[39] L. F. Abbott and P. Dayan, The effect of correlated variability on the accuracy of a population code, Neural Comput. **11**, 91 (1999).
[40] A. S. Ecker, P. Berens, A. S. Tolias, and M. Bethge, The effect of noise correlations in populations of diversely tuned neurons, J. Neurosci. **31**, 14272 (2011).
[41] G. Tkačik, J. S. Prentice, V. Balasubramanian, and E. Schneidman, Optimal population coding by noisy spiking neurons, Proc. Natl. Acad. Sci. USA **107**, 14419 (2010).
[42] M. Nardin, J. Csicsvari, G. Tkačik, and C. Savin, The structure of hippocampal CA1 interactions optimizes spatial coding across experience, J. Neurosci. **43**, 8140 (2023).
[43] R. A. da Silveira and M. J. Berry, High-fidelity coding with correlated neurons, PLoS Comput. Biol. **10**, e1003970 (2014).
[44] J. C. Boffi, B. Bathellier, H. Asari, and R. Prevedel, Noisy neuronal populations effectively encode sound localization in the dorsal inferior colliculus of awake mice, eLife **13**, RP97598 (2024).
[45] A. B. Graf, A. Kohn, M. Jazayeri, and J. A. Movshon, Decoding the activity of neuronal populations in macaque primary visual cortex, Nat. Neurosci. **14**, 239 (2011).
[46] G. Mahuas, O. Marre, T. Mora, and U. Ferrari, Small-correlation expansion to quantify information in noisy sensory systems, Phys. Rev. E **108**, 024406 (2023).
[47] S. Deny, U. Ferrari, E. Macé, P. Yger, R. Caplette, S. Picaud, G. Tkačik, and O. Marre, Multiplexed computations in retinal ganglion cells of a single type, Nat. Commun. **8**, 1964 (2017).
[48] V. Botella-Soler, S. Deny, G. Martius, O. Marre, and G. Tkačik, Nonlinear decoding of a complex movie from the mammalian retina, PLoS Comput. Biol. **14**, e1006057 (2018).
[49] P. Yger, G. L. B. Spampinato, E. Esposito, B. Lefebvre, S. Deny, C. Gardella, M. Stimberg, F. Jetter, G. Zeck, S. Picaud *et al.*, A spike sorting toolbox for up to thousands of electrodes validated with ground truth recordings in vitro and in vivo, eLife **7**, e34518 (2018).
[50] G. Mahuas, G. Isacchini, O. Marre, U. Ferrari, and T. Mora, in *Advances in Neural Information Processing Systems 33*, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Curran, Red Hook, 2020), pp. 5070–5080.
[51] I. Brivanlou, D. Warland, and M. Meister, Mechanisms of concerted firing among retinal ganglion cells, Neuron **20**, 527 (1998).
[52] I. Kanitscheider, R. Coen-Cagli, and A. Pouget, Origin of information-limiting noise correlations, Proc. Natl. Acad. Sci. USA **112**, E6973 (2015).
[53] J. J. Atick and A. N. Redlich, Towards a theory of early visual processing, Neural Comput. **2**, 308 (1990).

[54] L. McIntosh, N. Maheswaranathan, A. Nayebi, S. Ganguli, and S. Baccus, in *Advances in Neural Information Processing Systems 29*, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Curran, Red Hook, 2016), pp. 1369–1377.

[55] N. Maheswaranathan *et al.*, Interpreting the retinal neural code for natural scenes: From computations to neurons, Neuron **111**, 2742 (2023).

[56] X. Ding, D. Lee, J. Melander, G. Sivulka, S. Ganguli, and S. Baccus, in *Advances in Neural Information Processing Systems 36*, edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Curran, Red Hook, 2023), pp. 44310–44322.

[57] S. Azeglio, V. C. Garcia, G. Glaziou, P. Neri, O. Marre, and U. Ferrri, Higher-order convolution improves neural predictivity in the retina, arXiv:2505.07620.

[58] M. A. Goldin, B. Lefebvre, S. Virgili, M. K. P. V. Cang, A. Ecker, T. Mora, U. Ferrri, and O. Marre, Context-dependent selectivity to natural images in the retina, Nat. Commun. **13**, 5556 (2022).

[59] H. B. Barlow, Possible principles underlying the transformations of sensory messages, *Sensory Communication*, edited by W. A. Rosenblith (MIT Press, Cambridge, MA, 1961).

[60] S. Panzeri, S. R. Schultz, A. Treves, and E. T. Rolls, Correlations and the encoding of information in the nervous system, Proc. R. Soc. London B **266**, 1001 (1999).

[61] J. J. Atick, Could information theory provide an ecological theory of sensory processing? Network **3**, 213 (1992).

[62] X. Pitkow and M. Meister, Decorrelation and efficient coding by retinal ganglion cells, Nat. Neurosci. **15**, 628 (2012).

[63] https://github.com/gmahuas/noisecorr.

[64] J. de la Rocha, B. Doiron, E. Shea-Brown, K. Josić, and A. Reyes, Correlation between neural spike trains increases with firing rate, Nature (London) **448**, 802 (2007).

[65] S. Trenholm, A. J. McLaughlin, D. J. Schwab, M. H. Turner, R. G. Smith, F. Rieke, and G. B. Awatramani, Nonlinear dendritic integration of electrical and chemical synaptic inputs drives fine-scale correlations, Nat. Neurosci. **17**, 1759 (2014).