Published in final edited form as:

Proc Conf Inf Sci Syst. 2018 March ; 2018: . doi:10.1109/ciss.2018.8362310.

Optimal Information Transmission by Overlapping Retinal Cell Mosaics

Yilun Zhang^{*,†}, David B. Kastner[‡], Stephen A. Baccus[§], Tatyana O. Sharpee^{*,†}

^{*}Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, California, USA

[†]Department of Physics, University of California San Diego, La Jolla, California, USA

[‡]Department of Psychiatry, University of California San Francisco, San Francisco, California, USA

[§]Department of Neurobiology, Stanford University, Palo Alto, California, USA

Abstract

The retina provides an excellent system for understanding the trade-offs that influence distributed information processing across multiple neuron types. We focus here on the problem faced by the visual system of allocating a limited number neurons to encode different visual features at different spatial locations. The retina needs to solve three competing goals: 1) encode different visual features, 2) maximize spatial resolution for each feature, and 3) maximize accuracy with which each feature is encoded at each location. There is no current understanding of how these goals are optimized together. While information theory provides a platform for theoretically solving these problems, evaluating information provided by the responses of large neuronal arrays is in general challenging. Here we present a solution to this problem in the case where multi-dimensional stimuli can be decomposed into approximately independent components that are subsequently coupled by neural responses. Using this approach we quantify information transmission by multiple overlapping retinal ganglion cell mosaics. In the retina, translation invariance of input signals makes it possible to use Fourier basis as a set of independent components. The results reveal a transition where one high-density mosaic becomes less informative than two or more overlapping lower-density mosaics. The results explain differences in the fractions of multiple cell types, predict the existence of new retinal ganglion cell subtypes, relative distribution of neurons among cell types and differences in their nonlinear and dynamical response properties.

Keywords

Mutual Information; Efficient Coding; Retina Ganglion Cell; Photoreceptor Noise; Functional Neuronal Types

I. Introduction

The efficient coding hypothesis plays an important role in understanding the design principles of neural systems, especially sensory systems [1]. Efficient coding hypothesis can be commonly formulated in the language of information theory [2] by stating that a neural

system should maximize the mutual information between its input stimuli and neural response, subject to certain metabolic constraints [3]. However, in practice the application of efficient coding hypothesis is largely restricted to linear systems with Gaussian noise [4], systems using local Gaussian approximation [5], or small neural systems consisting of only few neurons [6]-[9]. The reasons are that, on the one hand, mutual information can be computed analytically for large systems only with strong simplifying assumptions, such as assuming linear input-output functions for individual neurons. On the other hand, numeric computation of mutual information is challenging for high dimensional variables, since the computational complexity is exponential in the number of dimensions. One approach to overcome this problem is to use Fisher information as a bound [10] or approximation [11] to mutual information, but this approach does not work with high dimensional stimuli. Another approach is to use variational approximation with certain decoder [12], but such methods only provide a bound on mutual information, and the result depends on the choice of a specific decoder. Thus, a major open problem remains for how to characterize information transmitted by large neuronal populations taking into account both the nonlinear aspects of neuronal responses and the fact that these neurons might be sensitive to different components of multi-dimensional input signals.

To test the efficient coding hypothesis in large neural populations, we developed a method that could efficiently compute the mutual information provided about the high dimensional inputs by the responses of a large neuronal array. The method works for a correlated Gaussian distribution or any other stimulus distribution that can be factorized. Applying this approach to multiple overlapping retinal ganglion cell mosaics, consisting of ~ 1000 neurons, we identify a transition that determines when it is optimal for a system to shift from one high-density mosaic to two or more overlapping lower-density mosaics. We further show how the relative densities in the two mosaics depend on different types of input noise and their spatiotemporal filtering properties. The results account for the observed properties of retinal ganglion cell types in the salamander cells and predict new cell types in the primate and guinea pig retinas.

II. Results

A. Mutual Information for Large Neuronal Arrays

For a large group of linear-nonlinear (LN) neurons encoding a high dimensional stimulus $\vec{s}$ using population neural response $\vec{r}$, the mutual information between stimulus and neural response $I(\vec{s}, \vec{r})$ is computationally intractable due to its high dimensionality. We show that the problem can be converted to a sum of one-dimensional problems provided one can identify a basis where the input distribution factorizes. Importantly, the neural responses depend simultaneously on multiple stimulus components, so this problem is more complex than that of independent response channels encoding independent inputs.

To model neural responses, we assume that their responses can be approximately described by a saturating logistic nonlinearity as a function of the stimulus component along the neuron's receptive field (RF). Specifically, the probability to observe a spike from the i th neuron (in a small time window) is given by:

$$p(r_i = 1 \mid \vec{s}) = \frac{1}{1 + e^{-\frac{u_i - \mu_i}{\nu_i^{\text{eff}}}}}, \quad (1)$$

where $u_i = \vec{f}^i \cdot \vec{s}$ represents the component of the stimulus $\vec{s}$ along the i th neuron's RF $\vec{f}^i$. The parameter μ_i represents the mid-point (threshold) of the spiking nonlinearity where the spiking probability is 0.5. The parameter ν_i^{eff} is the inverse slope of the nonlinearity which represents the cumulative effect of multiple noise sources in the input neural circuitry [7], [13]-[15]. The response across the population is described by a response vector $\vec{r}$ whose components r_i take values 1 or 0 corresponding to the presence or absence of a spike from the i th neuron. The probability to observe the response vector $\vec{r}$ across the population is then given by:

$$p(\vec{r} \mid \vec{s}) = \frac{1}{Z(\vec{s})} \prod_i \exp\left(r_i \frac{u_i - \mu_i}{\nu_i^{\text{eff}}}\right), \quad (2)$$

with the normalization factor $Z(\vec{s}) = \prod_i \left[1 + \exp\left(\frac{u_i - \mu_i}{\nu_i^{\text{eff}}}\right)\right]$. We note that although this

expression is obtained assuming the neural responses are conditionally independent given the stimulus, the final result will also hold in the presence of stimulus-independent noise correlations as long as the population response function still belongs to an exponential family.

Now if there is a basis within which the input distribution factorizes, $P(\vec{s}) = \prod_k P(s_k)$, then we can use this basis to represent both the stimuli and neural RFs. For concreteness, we use the Fourier components, because this is the basis within which translation invariant signals factorize, and this basis would be the most appropriate for analyzing the retinal arrays. Both stimuli and neural RFs can be described using their Fourier components $S_{\vec{k}}$ and $F_{\vec{k}}^j$ as:

$$s(\vec{x}) = \sum_{\vec{k}} S_{\vec{k}} e^{i\vec{k} \cdot \vec{x}}, \quad f^j(\vec{x}) = \sum_{\vec{k}} F_{\vec{k}}^j e^{i\vec{k} \cdot (\vec{x} - \vec{x}_j)}, \quad (3)$$

where $\vec{x}_j$ marks the coordinates of the peak of j th neuron RF and $F_{\vec{k}}^j$ is a Fourier component of the j th neuron RF computed relative to $\vec{x}_j$. In these expressions, the Fourier components should satisfy the following constraints $S_{\vec{k}} = S_{-\vec{k}}^*$ and $F_{\vec{k}}^j = F_{-\vec{k}}^{j*}$ that derive from the fact that stimuli and RFs have to be real-valued functions. The stimulus component along j th neuron RF is given by $u_j = \sum_{\vec{k}} S_{\vec{k}} F_{\vec{k}}^{j*} e^{i\vec{k} \cdot \vec{x}_j}$.

Now, the full response probability from Eq. (2) can be rewritten as

$$p(\vec{r} | \vec{s}) = \frac{1}{Z(\vec{s})} \exp \left[\sum_k \vec{S}_k T_k(\vec{r}) \right] h(\vec{r}), \quad (4)$$

where the quantities T_k are obtained as weighted averages of neural responses as

$T_k = \sum_j r_j \frac{F_k^{j*}}{v_j} e^{i \vec{k} \cdot \vec{x}_j}$, and $h(\vec{r}) = \exp(-\sum_i r_i \frac{\mu_i}{v_i})$. Because of the exponential structure of the full response probability, these quantities comprise the sufficient statistics [16] of the neural responses. In other words, the mutual information between stimuli and the set of neural responses $\vec{r}$ is equal to the mutual information between the stimuli and the set of T_k values: $I(\{S_k\} | \vec{r}) = (I(\{S_k\} | \{T_k\}))$. Importantly, we show in Materials and Methods that S_k is also the sufficient statistics of T_k so that $p(T_k | \vec{s}) = p(T_k | S_k)$, and that $\vec{T}$ also factorizes in the same basis $p(\vec{T}) = \prod_k p(T_k)$. Therefore, in this basis we obtain:

$$I(\vec{s}, \vec{r}) = \sum_k I(S_k | T_k). \quad (5)$$

This approach makes it possible to convert a multi-dimensional information calculation into a sum of one dimensional calculation for the mutual information between independent stimulus components S_k and the corresponding sufficient statistics variables $T_k(\vec{r})$. The variables $T_k(\vec{r})$ are computed as linear combinations of the neural responses across the population. Thus the approach does take different patterns of population responses into account. It also indicates which patterns of neural responses carry information about which stimulus components.

B. Transitions in the Coordinated Encoding by Retinal Ganglion Cell Mosaics

We now use this approach to compute and analyze information transmission in large arrays of retinal ganglion cells composed of multiple sub-types. The first question that we would like to address is whether it is better to encode the same type of the visual feature with higher spatial resolution or to encode different types of visual features each with lower spatial resolution. “Better” in this context is quantified as conveying greater amounts of mutual information about the full spatial stimulus pattern. We model each retinal array using a hexagonal grid (Fig. 1, insets). The spatial profile of each neuron is modeled as a difference of Gaussian to capture the center-surround structure [18]

$$f(\vec{x}) = C(\sigma, w, \gamma) \left(\frac{1}{2\pi\sigma^2} e^{-\frac{(\vec{x} - \vec{x}_c)^2}{2\sigma^2}} - \frac{w}{2\pi\gamma^2\sigma^2} e^{-\frac{(\vec{x} - \vec{x}_c)^2}{2\gamma^2\sigma^2}} \right), \quad (6)$$

where $C(\sigma, w, \gamma)$ is chosen so that $\|f\|_2 \equiv \int f(\vec{x})^2 d\vec{x} = 1$, and $\vec{x}_c$ is the receptive field center.

The lattice spacing a , defined as the difference between nearest neighbor RF peaks, relates to the RF center size σ as $a = 2\sigma$, in agreement with experimental observations [18]-[22]. This relationship also constrains the number of neurons N in a given lattice as $Na^2 = \text{const}$. Both configurations – either a single array or two overlapping arrays – are constrained to have the same total number of neurons and to produce the same number of spiking responses, on average across the stimulus ensemble.

We find that a transition occurs when noise level v_{eff} decreases. For a given average spiking rate across the population, when noise level is high, a single higher density array conveys more information. When noise level decreases below a certain value, two overlapping mosaics convey more information than one higher-density mosaic. In Fig. 1A we plot results as a function of noise component v associated with spike generation. Here one observes a transition boundary (white line) that separates a range of values for the noise level and neural thresholds where two overlapping mosaics convey more information from those values for which a single higher density mosaic convey more information.

The transition can also be induced by noise in the afferent circuitry. We model this type of noise by a Gaussian white noise added to stimulus, so that the input signal u according to which spikes are produced is described by:

$$u = \vec{f} \cdot (\vec{s} + \vec{n}), \quad (7)$$

where $\langle n(\vec{x})n(\vec{x}') \rangle = \delta(\vec{x} - \vec{x}')\sigma_n^2$. We will refer to this noise as photoreceptor noise because studies find it to be the dominant component prior to ganglion cell nonlinearity [13]-[15]. The presence of an independent additive noise does not change the form of sufficient statistics, so our previous information calculation still holds. Yet, because the photoreceptor noise is added before ganglion cell integrate (spatial-temporal) inputs, the magnitude of this noise can be reduced by integration summarized by the spatio-temporal filters. In particular, when neurons form multiple mosaics, they have broader spatial filters compared to the case where there is only one type of neurons. Thus, forming multiple mosaics provides a powerful way to reduce input noise. Correspondingly, increasing input noise can by itself force the transition from encoding using one array to encoding using multiple arrays (Fig. 1B).

C. Interaction Between Multiple Noise Sources Impacts Optimal Coordination Between Retinal Cell Types

In the presence of both the photoreceptor noise and the spiking noise, the effective noise in Eq. 1 becomes (see Materials and Methods for derivation):

$$\nu^{\text{eff}} = \nu \sigma_{\vec{f} \cdot \vec{s}} \sqrt{1 + \sigma_{\text{eff}}^2 + \frac{\pi \sigma_{\text{eff}}^2}{8 \nu^2}} \quad (8)$$

σ_{eff}^2 is the effective photoreceptor noise that depends on both the power spectrum of stimuli and neural filters:

$$\sigma_{\text{eff}}^2 = \frac{\sigma_n^2}{\sigma_{\vec{f} \cdot \vec{s}}^2} = \frac{\sigma_n^2}{\int \hat{f}(\vec{k})^2 k^{-2} d\vec{k} \int \hat{g}(\omega)^2 \omega^{-2} d\omega}. \quad (9)$$

Here f and g are the spatial and temporal Fourier components of the filter describing neuron's spatiotemporal RF. Eq. 9 takes into account statistics of natural visual scenes showing $1/f^2$ dependence on both spatial and temporal frequencies [23], [24].

Using Eq. 8 and 9 we can now analyze how the interaction between these noise sources impacts properties of the optimal mosaics. There are two main effects. The first effect is mainly quantitative, showing that with increasing photoreceptor noise, the transition boundary between one-array and two-array encoding shifts towards larger ν (Fig. 1C).

The second effect induced by photoreceptor noise can lead to large changes in the predicted fraction of different neuronal types. For example, for low values of the photoreceptor noise, near the transition from one-array encoding to two-array encoding, when two overlapping mosaics first start to provide more information than a single dense mosaics, the number of neurons is very unevenly distributed between the arrays. Most of the neurons form one (denser) array and only a small fraction of neurons form the other array (Fig. 2A, D). Neurons in the lower density array have higher thresholds and spike less often. In contrast, for high photoreceptor noise, neurons in the lower density array have lower thresholds and spike more often (Fig. 2C, F). These theoretical observations can help explain the observed differences in the number of adapting and sensitizing neurons for fast Off, medium Off, and slow Off ganglion cells in salamander [17]. Because fast Off neurons have faster temporal filters, according to Eq. 9, the effective photoreceptor noise for these neurons will be higher than that for the medium/slow Off neurons which have comparable temporal kernels. As a result, the expected ratio between adapting and sensitizing cells for these three classes will be different, with largest differences between fast and medium/slow subtypes (Fig. 3).

The theoretical analyses also predict the existence of new cell types. In Fig. 4A we show data from salamander cell types [17] relative to the transition boundary of the region in the parameter space where one array encoding provides more information than the two array encoding. We find that Off neurons are located on one side of the boundary where two

arrays provide more information whereas most On neurons are located on the other side of the boundary. In the case of the salamander neurons with fast time courses these predictions agree with experimental measurements because it is known that the fast-Off neurons split into subtypes whereas fast-On neurons do not. At the same time, theory predicts that slow On neurons should split into subtypes. This is a new prediction to be tested experimentally.

In Fig. 4B and C we compare theoretical predictions for the number of cell types with vertebrate data from primate [25] and guinea pig [26], respectively. Here the predictions are that Off neurons from the guinea pig as well as primate parvo cells should split into subtypes whereas the corresponding On neurons in those species should not. The theoretical predictions are made for larger photoreceptor noise than in the salamander because photoreceptor noise increases with temperature. We estimate the photoreceptor noise for guinea pig to be approximately 2 times the value of salamander using linear interpolation [27]. Meanwhile, primate is shown to have smaller noise in single photon response than guinea pig [28], thus we use a smaller photoreceptor noise for primate than for guinea pig.

D. Optimal Coordinated Encoding Explains Transient Adaptation

Our results predict that, for larger effective photoreceptor noise where the fraction of adapting cell is larger (Fig. 2C), the optimal threshold difference between the adapting and sensitizing cells increase when the intrinsic noise level ν decreases. Such prediction explains the observation that the threshold difference between adaptive and sensitizing cells increases transiently after signal variance suddenly drops [17]. Theoretical results for lower effective photoreceptor noise, such as for medium and slow Off neurons that have slower temporal filters, indicate the opposite behavior. For example, in Fig. 2A the difference between thresholds of the two subtypes would decrease following a decrease in intrinsic noise level ν .

E. Optimal Redundant Coding

Finally, it is worth commenting on the possibility that two overlapping arrays can be formed by neurons with identical response characteristics. We find such possibility, but only for relative large noise levels associated with both spiking and in photoreceptors (Fig. 5). So far, none of the experimental data we have fell into this region. The observation of optimal redundant coding creates the prediction that the usual association between mosaics and cell types in the retina can be violated in the presence of high noise.

III. Summary

This paper described a theoretical approach for computing information in large neural populations in the case where inputs can be decomposed into a set of approximately independent components. For each of the stimulus components we derived the corresponding sufficient statistics as a combination of neural responses across the population that captures all of the information available in the responses of the neural population about this component. Applying this approach to populations of retinal ganglion cells, we characterized conditions when it becomes optimal to shift from encoding of a single stimulus feature at high spatial resolution to encoding of two or more features at lower

resolutions. The results accounts for properties of known retinal ganglion cell types, predict the existence of new cell types as well as novel dynamical properties of these cell types following adaptation to stimulus variance.

IV. Materials and Methods

A. Sufficient Statistics of Binary Neurons

We showed that $T_{\rightarrow k}$ is the sufficient statistics of stimulus Fourier component $S_{\rightarrow k}$. We now prove that $S_{\rightarrow k}$ is also the sufficient statistics of $T_{\rightarrow k}$. To show this, we write down the conditional probability

$$\begin{aligned} p(\vec{S} | \vec{T}) &= \frac{p(\vec{T} | \vec{S})p(\vec{S})}{p(\vec{T})} = \frac{p(\vec{r} | \vec{S})p(\vec{S})}{p(\vec{T})} \left| \frac{\partial \vec{r}}{\partial \vec{T}} \right| \\ &= \frac{1}{Z[\vec{s}(\vec{S})]} \exp \left[\sum_k S_{\rightarrow k} T_{\rightarrow k} \right] h[\vec{r}(\vec{T})] \frac{p(\vec{S})}{p(\vec{T})} \left| \frac{\partial \vec{r}(\vec{T})}{\partial \vec{T}} \right|, \end{aligned}$$

where $\left| \frac{\partial \vec{r}}{\partial \vec{T}} \right|$ is the determinant of Jacobian, and in the last line we used Eq. 4. Therefore $S_{\rightarrow k}$ is also the sufficient statistics for $T_{\rightarrow k}$, and that $p(T_{\rightarrow k} | \vec{S}) = p(T_{\rightarrow k} | S_{\rightarrow k})$. We can further show that $p(\vec{T})$ factorizes

$$\begin{aligned} p(\vec{T}) &= \int p(\vec{T} | \vec{S}) p(\vec{S}) d\vec{S} = \int \left[\prod_k p(T_{\rightarrow k} | \vec{S}) \right] \left[\prod_k p(S_{\rightarrow k}) \right] d\vec{S} \\ &= \int \prod_k p(T_{\rightarrow k} | S_{\rightarrow k}) p(S_{\rightarrow k}) dS_{\rightarrow k} = \prod_k p(T_{\rightarrow k}). \end{aligned}$$

Thus, the mutual information between $\vec{s}$ and $\vec{r}$ can be decomposed into

$$\begin{aligned} I(\vec{s}, \vec{r}) &= H(\vec{T}) - H(\vec{T} | \vec{S}) = \sum_k H(T_{\rightarrow k}) - \sum_k H(T_{\rightarrow k} | \vec{S}) \\ &= \sum_k H(T_{\rightarrow k}) - \sum_k H(T_{\rightarrow k} | S_{\rightarrow k}) = \sum_k I(S_{\rightarrow k}, T_{\rightarrow k}). \end{aligned}$$

B. Spiking Nonlinearity in Presence of Photoreceptor Noise

In presence of photoreceptor noise defined by Eq. 7, the spiking probability given filtered stimulus $p(r|u)$ is a sigmoid function $p(r = 1 | u) = \frac{1}{1 + e^{-\frac{u - \mu}{\nu'}}$. The spiking probability given

stimulus is a convolution of sigmoid and gaussian $p(r|s) = \int p(r|u)p(ul s)du$, which is

approximately another sigmoid function with different slope $p(r = 1 | s) = \frac{1}{1 + e^{-\frac{u - \mu}{\nu^{\text{eff}}}}}$, where

the new slope satisfies $(\nu^{\text{eff}})^2 = \nu'^2 + \frac{\pi}{8}\sigma_n^2$. Thus, given sigmoid spiking probability $p(r = 1 | u)$ and gaussian noise, $p(r = 1 | s)$ is a sigmoid function with higher effective ν .

To find out the exact expression for ν^{eff} , we note that ν' scales with the standard deviation of u such that $\nu' = \sigma_{\nu}\nu = \nu\sigma_{\vec{f} \cdot \vec{s}}\sqrt{1 + \sigma_{\text{eff}}^2}$, where ν is the intrinsic inverse slope in unit of σ_u as we reported in our figures, σ_{eff} is defined in Eq. 9, and $\sigma_{\vec{f} \cdot \vec{s}}$ is the standard deviation of $\vec{f} \cdot \vec{s}$. Thus we have $(\nu^{\text{eff}})^2 = (1 + \sigma_{\text{eff}}^2)\nu^2\sigma_{\vec{f} \cdot \vec{s}}^2 + \frac{\pi}{8}\sigma_n^2$.

C. Transforming Experimental Data

To compare our predictions with experiments, we transform the measured values to the intrinsic ν and μ that are independent from photoreceptor noise or neuron's filters. The measured inverse slope ν_{exp} is simply ν^{eff} in unit of the standard deviation of $\vec{f} \cdot \vec{s}$ such that $\nu_{\text{exp}} = \nu^{\text{eff}} / \sigma_{\vec{f} \cdot \vec{s}}$, so we have $\nu_{\text{exp}}^2 = (1 + \sigma_{\text{eff}}^2)\nu^2 + \frac{\pi}{8}\sigma_{\text{eff}}^2$. Similarly we have $\mu_{\text{exp}}^2 = (1 + \sigma_{\text{eff}}^2)\mu^2$. Using these relations, we can transform measured values ν_{exp} and μ_{exp} (averaged across all neuron data of the same type) to ν and μ and compare with our calculation.

Acknowledgement

Authors thank John Berkowitz for discussions. This research was supported by the National Science Foundation (NSF) award numbers IIS-1254123, IOS-1556388, and IIS-1724421.

References

- [1]. Barlow HB, "Possible principles underlying the transformation of sensory messages," Sensory communication, pp. 217–234, 1961.
- [2]. Shannon CE and Weaver W, "The mathematical theory of information," 1949.
- [3]. Linsker R, "Self-organization in a perceptual network," Computer, vol. 21, no. 3, pp. 105–117, 1988.
- [4]. Atick JJ and Redlich AN, "What does the retina know about natural scenes?" Neural computation, vol. 4, no. 2, pp. 196–210, 1992.
- [5]. Karklin Y and Simoncelli EP, "Efficient coding of natural images with a population of noisy linear-nonlinear neurons," Advances in neural information processing systems, vol. 24, p. 999, 2011. [PubMed: 26273180]
- [6]. Laughlin S, "A simple coding procedure enhances a neuron's information capacity," Zeitschrift für Naturforschung c, vol. 36, no. 9-10, pp. 910–912, 1981.
- [7]. Kastner DB, Baccus SA, and Sharpee TO, "Critical and maximally informative encoding between neural populations in the retina," Proceedings of the National Academy of Sciences, vol. 112, no. 8, pp. 2533–2538, 2015.
- [8]. Gjorgjieva J, Sompolinsky H, and Meister M, "Benefits of pathway splitting in sensory coding," The Journal of Neuroscience, vol. 34, no. 36, pp. 12 127–12 144, 2014.

- [9]. Nikitin AP, Stocks NG, Morse RP, and McDonnell MD, "Neural population coding is optimized by discrete tuning curves," *Physical review letters*, vol. 103, no. 13, p. 138101, 2009. [PubMed: 19905542]
- [10]. Brunel N and Nadal J-P, "Mutual information, fisher information, and population coding," *Neural Computation*, vol. 10, no. 7, pp. 1731–1757, 1998. [PubMed: 9744895]
- [11]. Wei X-X and Stocker AA, "Mutual information, fisher information, and efficient coding," *Neural computation*, vol. 28, no. 2, pp. 305–326, 2016. [PubMed: 26654209]
- [12]. Brinkman BA, Weber AI, Rieke F, and Shea-Brown E, "How do efficient coding strategies depend on origins of noise in neural circuits?" *PLOS Comput Biol*, vol. 12, no. 10, p. e1005150, 2016. [PubMed: 27741248]
- [13]. Ala-Laurila P, Greschner M, Chichilnisky EJ, and Rieke F, "Cone photoreceptor contributions to noise and correlations in the retinal output," *Nat Neurosci*, vol. 14, no. 10, pp. 1309–1316, 2011. [PubMed: 21926983]
- [14]. Rieke F and Baylor DA, "Origin and functional impact of dark noise in retinal cones," *Neuron*, vol. 26, pp. 181–186, 2000. [PubMed: 10798402]
- [15]. Borghuis BG, Sterling P, and Smith RG, "Loss of sensitivity in an analog neural circuit," *J Neurosci*, vol. 29, pp. 3045–3058, 2009. [PubMed: 19279241]
- [16]. Cover TM and Thomas JA, *Information theory*. New York: John Wiley & Sons, INC., 1991.
- [17]. Kastner DB and Baccus SA, "Coordinated dynamic encoding in the retina using opposing forms of plasticity," *Nature neuroscience*, vol. 14, no. 10, pp. 1317–1322, 2011. [PubMed: 21909086]
- [18]. Borghuis BG, Ratliff CP, Smith RG, Sterling P, and Balasubramanian V, "Design of a neuronal array," *J Neurosci*, vol. 28, pp. 3178–3189, 2008. [PubMed: 18354021]
- [19]. Devries SH and Baylor DA, "Mosaic arrangement of ganglion cell receptive fields in rabbit retina," *J Neurophysiol*, vol. 78, pp. 2048–2060, 1997. [PubMed: 9325372]
- [20]. Liu YS, Stevens CF, and Sharpee TO, "Predictable irregularities in retinal receptive fields," *PNAS*, vol. 106, pp. 16 499–16 504, 2009.
- [21]. Field GD and Chichilnisky EJ, "Information processing in the primate retina: Circuitry and coding," *Ann. Rev. Neurosci*, vol. 30, pp. 1–30, 2007. [PubMed: 17335403]
- [22]. Field GD, Sher A, Gauthier JL, Greschner M, Shlens J, Litke AM, and Chichilnisky EJ, "Spatial properties and functional organization of small bistratified ganglion cells in primate retina," *J Neurosci*, vol. 27, pp. 11 261–13 272, 2007.
- [23]. Field DJ, "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA A*, vol. 4, no. 12, pp. 2379–2394, 1987.
- [24]. Dong DW and Atick JJ, "Statistics of natural time-varying images," *Network: Comput. Neural Syst*, vol. 6, pp. 345–358, 1995.
- [25]. Sincich LC, Horton JC, and Sharpee TO, "Preserving information in neural transmission," *J. Neurosci*, vol. 29, pp. 6207–6216, 2009. [PubMed: 19439598]
- [26]. Zaghloul KA, Boahen K, and Demb JB, "Different circuits for on and off retinal ganglion cells cause different contrast sensitivities," *The Journal of Neuroscience*, vol. 23, no. 7, pp. 2645–2654, 2003. [PubMed: 12684450]
- [27]. Aho AC, Donner K, Hyden C, Larsen LO, and Reuter T, "Low retinal noise in animals with low body temperature allows high visual sensitivity," *Nature*, vol. 334, pp. 348–350, 1988. [PubMed: 3134619]
- [28]. Field GD and Rieke F, "Mechanisms regulating variability of the single photon responses of mammalian rod photoreceptors," *Neuron*, vol. 35, no. 4, pp. 733–747, 2002. [PubMed: 12194872]

Fig. 1.

(A) Mutual information difference between optimal two-array encoding and one-array encoding as a function of neuron intrinsic noise level and mean spiking probability. We map mean spiking probability to mean spiking threshold to better compare with experimental data. White line show zero mutual information difference marking the transition of optimal encoding strategy from one array to two arrays. We use photoreceptor noise $\sigma_{eff} = 0.3$. (B) Mutual information difference between optimal two-array encoding and one-array encoding as a function of photoreceptor noise level. Black dashed line shows the transition from one array to two arrays. The intrinsic noise level and mean spiking threshold for each line are chosen to match the data from salamander [17]. Experimentally measured temporal kernels of each cell type (as shown in Fig. 3A) are used to calculate the effective photoreceptor noise. (C) Transition boundary for different levels of photoreceptor noise.

Fig. 2.

Optimal threshold and neuron fraction of high threshold cell (red) and low threshold cell (blue) under low (A, $\sigma_{\text{eff}} = 0.3$), intermediate (B, $\sigma_{\text{eff}} = 0.6$) and high (C, $\sigma_{\text{eff}} = 1.0$) photoreceptor noise. Solid/dashed line shows the regime where two/one array is optimal. For low photoreceptor noise, only the regime where two arrays are optimal is shown, since only the transition from two arrays to one array take places, and redundant coding (two neuron types have the same threshold and neuron fraction) is not optimal. For high photoreceptor noise, two arrays are always preferable than one array, and a transition to redundant coding occurs. For intermediate photoreceptor noise, both transitions could coexist, with the transition from two arrays to one array occurs at low intrinsic noise level, while the transition to redundant coding occurs at high intrinsic noise level. Another noticeable difference between different photoreceptor noise is the ratio of cells of high and low spiking thresholds. High threshold cells have few neurons than low threshold cells when photoreceptor noise is low, while the opposite is true when photoreceptor noise is high. For intermediate photoreceptor noise, both ratio could be optimal, with a potential transition take places as intrinsic noise changes.

Fig. 3.

(**A**) Average temporal receptive fields measured in experiments. Faster temporal kernel results in higher effective photoreceptor noise, thus have more adapting cells. Similarly, slower temporal kernel results in more sensitizing cells. (**B**) (Adapted from [17]) Fractions of adapting (red) and sensitizing (blue) cells of different cell types measured in Salamander.

Fig. 4.

Mutual information difference and predicted transition boundary for salamander, primate, and guinea pig, compared with experimental data. Markers with white/black face color indicate the data of ON/OFF cells, respectively. (A) Same as Fig. 1A but for photoreceptor noise $\sigma_{\text{eff}} = 0.25$, and data points from salamander [17]. Circles represent fast cell, up triangle represents medium cell, down triangles represent slow cell. (B) Same as (A), but for photoreceptor noise $\sigma_{\text{eff}} = 0.4$ and data from primate parvo cells [25]. (C) Same as (A), but for photoreceptor noise $\sigma_{\text{eff}} = 0.5$ and data from guinea pig [26].

Fig. 5.

Optimal spiking threshold difference (**A**) and neuron fraction difference (**B**) of two neuron sub-types in two arrays. White lines shows the transition boundary from different threshold (**A**) and receptive field size (**B**) to the same threshold/RF size. Redundant coding (two cell types in two arrays have the same spiking threshold/receptive field size) is optimal on the right side of the boundary. Photoreceptor noise is $\sigma_{\text{eff}} = 1.0$.