



HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2015 September 17.

Published in final edited form as:

Neuron. 2014 September 17; 83(6): 1329–1334. doi:10.1016/j.neuron.2014.08.040.

Towards functional classification of neuronal types

Tatyana O. Sharpee

Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037

Tatyana O. Sharpee: sharpee@salk.edu

Abstract

How many types of neurons are there in the brain? This basic neuroscience question remains unsettled despite many decades of research. Classification schemes have been proposed based on anatomical, electrophysiological or molecular properties. However, different schemes do not always agree with each other. This raises the question of whether one can classify neurons based on their function directly. For example, among sensory neurons, can a classification scheme be devised that is based on their role in encoding sensory stimuli? Here I outline theoretical arguments for how this can be achieved using information theory by looking at optimal numbers of cell types and paying attention to two key properties: correlations between inputs and noise in neural responses. This theoretical framework could help to map the hierarchical tree relating different neuronal classes within and across species.

As other publications in this issue of *Neuron* describe, we now stand at the threshold of being able to measure the anatomy and connections within neural circuits at an exquisite spatial resolution. The increasing availability of large-scale connectomics datasets holds the promises of answering one of the most basic and long-standing questions in neuroscience – how many different types of neurons are there in the brain (Jonas and Kording, 2014). Yet, the large size of these datasets also presents a set of challenges. To tackle these challenges, it would be advantageous to focus any analysis on those statistical properties of neural circuits that are likely to generalize across samples, brain regions, and perhaps species. In this respect, theoretical frameworks based on optimization principles can be very helpful, in particular in making predictions as to how many different neuron types one can expect to find in any given brain region, and pinpointing experiments that make differences among neural types more apparent. In this article I review a set of previous results and outline directions for future research on how we might be able to understand the relationship between neuronal types across species and brain regions, from the operation of identified neurons in invertebrate circuits to large neural populations in mammals.

Neurons can be classified based on anatomical (Boycott and Waessle, 1991; Davis and Sterling, 1979; Masland, 2004), molecular (Baraban and Tallent, 2004; Kodama et al., 2012;

© 2014 Elsevier Inc. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Nelson et al., 2006; Trimarchi et al., 2007) or electrophysiological properties (Markram et al., 2004). Yet, definitive mapping between these classification schema has proved difficult (Mott and Dingledine, 2003) because some of the diversity is due to morphological differences and some to differences in intrinsic firing or synaptic properties (Nelson et al., 2006). To complicate the matter further, sometimes vastly different molecular characteristics yield the same electrophysiological properties (Schulz et al., 2006). When such discrepancies arise, perhaps an insight based on neuronal function within the circuits could help. Functional classifications typically rely, in the case of sensory neurons, on differences in the types of input features encoded by different neurons (Balasubramanian and Sterling, 2009; Garrigan et al., 2010; Segev et al., 2006). (Analogous definitions are possible in motor circuits but will not be discussed here). The relevant input features represent the stimulus components that can modulate neural responses most strongly. (The strength of this modulation can be quantified using either mutual information [Adelman et al., 2003; Rajan and Bialek, 2013; Sharpee et al., 2004] or by the amount of change in the mean or variance between distributions of inputs that elicit and do not elicit spikes [Bialek and de Ruyter van Steveninck, 2005; Schwartz et al., 2006].) Examples of classification based on functional properties include the separation of neurons based on their chromatic preferences (Atick et al., 1992; Chichilnisky and Wandell, 1999; Derrington et al., 1984; Wiesel and Hubel, 1966), into On- and Off types of visual neurons (Chichilnisky and Kalmar, 2002; Masland, 1986; Werblin and Dowling, 1969), with further subdivisions into parvo- and magnocellular visual neurons (Sincich and Horton, 2005) that can be distinguished according to different spatial and temporal frequencies. These functional distinctions are often reflected in anatomical differences. For example, the dendrites of On and Off neurons segregate in different sublaminae of the inner plexiform layer (Masland, 1986).

Less commonly, neurons can sometimes be classified based on their nonlinear properties. Two of the most famous examples include the distinction between X and Y cells in the retina (Hochstein and Shapley, 1976) and between simple and complex cells in the primary visual cortex (Hubel and Wiesel, 1962; Hubel and Wiesel, 1968). While the existence of separate classes of simple and complex classes has been questioned (Chance et al., 1999; Mechler and Ringach, 2002), recent studies provide other examples of where neurons are primarily distinguished based on their nonlinear properties (Kastner and Baccus, 2011). For example, the so-called adapting and sensitizing cells differ primarily in terms of their thresholds but not their preferred visual features (Kastner and Baccus, 2011). These observations call attention to the need to understand the circumstances under which neuron types can be defined primarily based on features, nonlinearities, or both.

In prior work, the application of information theoretic principles has been very useful in accounting for the relative proportions of cells tasked with encoding different stimulus features (Balasubramanian and Sterling, 2009; Ratliff et al., 2010). Further, searching for an optimal solution could account for spatial and temporal filtering properties of retinal neurons (Atick and Redlich, 1990; Atick and Redlich, 1992; Bialek, 1987; Bialek and Owen, 1990; Rieke et al., 1991) as well as for changes in the filtering properties as a function of signal to noise ratio (Doi et al., 2012; Potters and Bialek, 1994; Srinivasan et al., 1982). Building on these previous successes, we will examine here the optimal solutions that maximize mutual information assuming that the responses of each neuron are binary and are described by

logistic tuning curves (Pitkow and Meister, 2012). Logistic tuning curves correspond to the least constrained model nonlinearity that can be made consistent with measured input-output correlations (Fitzgerald et al., 2011; Globerson et al., 2009). It is likely that similar conclusions will be obtained if one were to consider Fisher information (Dayan and Abbott, 2001) instead of the Shannon mutual information (Brunel and Nadal, 1998; Harper and McAlpine, 2004; Salinas, 2006) and to allow for graded neural responses. This is because using Fisher information one can construct a bound for the mutual information and at least in the case of low noise and large neural ensembles, map a population of binary neurons to single-neuron encoding using graded responses (Brunel and Nadal, 1998).

A key feature of the environment that neural circuits have to operate in is that the signals are not only multidimensional but also highly correlated. This is uniformly true across sensory modalities (Ruderman and Bialek, 1994; Simoncelli and Olshausen, 2001) and motor outputs (Todorov, 2004). For example, the inputs to the fly visual system derive from 10^6 photoreceptors each of which conveys a slightly distinct message from its neighbors, but overall correlations between the inputs are strong. For the sake of simplicity, we will approximate input signals as correlated Gaussian, ignoring for now the (fairly substantial) non-Gaussian effects in natural inputs (Ruderman and Bialek, 1994; Simoncelli and Olshausen, 2001; Singh and Theunissen, 2003). It should be noted that non-Gaussian effects can change the maximally informative dimensions within the input ensemble (Bell and Sejnowski, 1997; Karklin and Simoncelli, 2011; Olshausen and Field, 1996). Yet many insights can be obtained within the Gaussian approximation which we will focus on here. In the Gaussian approximation, the contours of equal probability are ellipses. The fact that inputs are correlated means that these ellipses are stretched in certain directions (Figure 1). The stronger the correlations, the more elongated the equal probability contours become. The fact that natural signals are strongly correlated means that the variance along certain stimulus dimensions in natural signals can be orders of magnitude larger than in other dimensions (Aljadeff et al., 2013; van Hateren and Ruderman, 1998). These observations on the properties of natural sensory signals have led Horace Barlow to propose the idea of a redundancy reduction, also known as “decorrelation” (Barlow, 2001; Barlow and Foldiak, 1989; Barlow). Indeed, in the high signal-to-noise regime, it would make sense to first encode the input value along the dimension that has the largest variance, and to allocate subsequent neurons to encode dimensions of progressively smaller variance. The high signal-to-noise regime presupposes that the ratio of noise to input variance is small even for dimensions along which input variance is small. With Gaussian inputs, one of the optimal configurations corresponds to sets of neurons that each encode inputs along orthogonal dimensions (Fitzgerald and Sharpee, 2009). For non-Gaussian inputs, optimal solutions specify that the sensitivity of single neurons to multiple inputs simultaneously (Horwitz and Hass, 2012; Sharpee and Bialek, 2007) and orthogonal dimensions should be substituted by sets of independent components (Bell and Sejnowski, 1997; Olshausen and Field, 1996; van Hateren and Ruderman, 1998) in the natural input ensemble.

The decorrelation comes with its own cost, however, in terms of its potential for noise amplification. Such problems are well known in the field of receptive field estimation where the goal is to estimate the relevant features from the neural responses (Sharpee, 2013; Theunissen et al., 2001, Sharpee et al., 2004). However, neural representations also have to

solve the same issues. To illustrate the main ideas we will consider correlations between two input dimensions and how they might optimally be encoded by two neurons. We will discuss along the way how these results might generalize to the higher dimensionality of the input space and neural populations. When two signals are correlated with strength ρ , the ratio of variances along these dimensions is given by $(1+\rho)/(1-\rho)$, which diverges as the correlation coefficient ρ increases towards 1. Assuming that input signals can be measured with a certain precision v that is the same for all input dimensions, this neural noise v will have a greater impact on encoding the dimension of smaller input variance. Specifically, the effective noise along the secondary dimension will be $(1+\rho)/(1-\rho)$ larger than the effective noise along the first dimension (Figure 1B). In the case of natural stimuli where variances along different input dimensions differ by orders of magnitude (Aljadeff et al., 2013; van Hateren and Ruderman, 1998) the effective noise along certain dimensions will be greatly amplified. In this situation, it might be better to allocate both neurons to encoding signal values along the primary dimension of variance as opposed to tasking them with encoding both input dimensions. In other words, when the amount of variance along the secondary dimension becomes smaller or comparable to neural noise, signals become effectively one-dimensional.

Neurons that have the same threshold and encode the same dimension would then functionally be classified as belonging to one class. Thus, the amount of neural noise together with correlations between inputs may determine the optimal number of neural classes. When the ratio of the neural noise to the smallest variance in input dimensions remains small, it will be beneficial to have as many neuronal classes as the input dimensions. Given that the noise in neural responses is substantial and can further enhance correlations between afferent neurons beyond correlations due to inputs (Ala-Laurila et al., 2011), multiple neurons might be assigned to encode the same input dimension, thereby reducing the number of distinct neuronal classes. Here, however, an interesting complication arises. It turns out that when multiple neurons encode the same input dimension, it might be optimal (in the sense of maximizing the Shannon mutual information about this input dimension) to assign these neurons different thresholds (Kastner et al., 2014; McDonnell et al., 2006; Nikitin et al., 2009). In this case, the neurons that encode the same dimensions would belong to separate classes, as has been observed in the retina (Kastner and Baccus, 2011). Whether or not this happens, i.e. whether mutual information is maximized when neurons have the same or different thresholds, again depends on the amount of neural noise (Figure 2). When noise in neural responses is large, information is maximal when the thresholds are the same (Figure 2A). In this regime, which might be termed redundant coding (Potters and Bialek, 1994; Tkacik et al., 2008), both neurons belong to the same functional class. As the noise in neural responses decreases, it may become beneficial to assign different thresholds (Figure 2) to the two neurons (Kastner et al., 2014).

It is also worth pointing out that in some cases correlations between inputs can be substantially increased beyond those present in the environment by noise fluctuations in afferent inputs (Ala-Laurila et al., 2011). The presence of these so-called noise correlations (Gawne and Richmond, 1993) could make encoding of stimulus values along the dimensions

of largest variance less important. In this case, multiple neurons could be assigned to encode stimulus values along dimensions or smaller variance.

In a more general case, where more than two neurons are encoding the same input dimension, one would expect to find a series of bifurcations with decreasing noise level (McDonnell et al., 2006). When the noise is large, all neurons have the same thresholds, and therefore form a single neuronal class. As the noise decreases, neurons might first split up into two groups, and if noise can be reduced further, subsequent neurons will separate from either of the two initial classes to form new classes. Thus, multiple neuronal classes may emerge even among neurons that encode the same input dimension and one could envision making a hierarchical tree of how different neuronal classes relate to each other.

Considering the effects due to input correlations and neural noise together, we observe that in some cases separate neural classes appear because neurons are encoding different input dimensions whereas in other cases neural classes appear because they have different thresholds along the same input dimension. For example, in the hypothetical case where correlations between inputs are zero, information is maximized by encoding separate input dimensions (Figure 3), regardless of neural noise level. Although this case is unrealistic, encoding along the orthogonal features will continue to be optimal provided that the correlations between inputs do not exceed a certain critical value. This critical value increases as the neural noise decreases (Figure 3). More generally, encoding along the orthogonal features corresponds to encoding along the so-called independent components of the input distribution, taking into account non-Gaussian characteristics of input statistics. However, the main observation that neurons in this regime will be distinguished based on their most relevant input feature is still likely to remain valid.

In the opposite regime of strong input correlations and small neural noise, neurons are distinguished based on their non-linear properties. As the neural noise increases and correlations remain strong, optimal solutions specify neurons to have both the same relevant features and the same thresholds. Overall, the ratio between input correlations and neural noise emerges as the main parameter controlling the separation of neurons into classes based on either their relevant features or nonlinearities.

Could neurons be segregated by both their relevant features and their nonlinearities? The answer is likely no, or at least the differences in thresholds for neurons tuned to different features are likely to be small. To see why this is the case, we recall that independent solutions are optimal only when correlations are weak, meaning along dimensions with similar variance. In this case, the effective noise along these dimensions will be comparable. Thus, the difference in noise level will not drive neural thresholds to different positions relative to the peak of the distribution. At the same time the mutual information (Cover and Thomas, 1991) is maximized when the spikes from two neurons are used equally often (this maximizes the so-called response entropy (Brenner et al., 2000)). Thus, the thresholds will be positioned at similar distances from the peak of the input distribution for two neurons. These arguments carry over to the case of multiple neurons encoding multiple input dimensions.

It is important to distinguish the effects of noise that drive the separation of neurons into two classes from the effects that the noise may have on changes in neural encoding within a given class. For example, changes in both the environmental conditions, such as the average light intensity and the metabolic state of the animal, may lead to an overall amount of noise across different inputs and neurons. Reducing the average light intensity will reduce correlations between photoreceptor outputs and change the direction of the maximal variance in the input distribution. Correspondingly, the optimal features that neurons encode will (and real features are known to) change (Atick and Redlich, 1990; Atick and Redlich, 1992; Potters and Bialek, 1994; Srinivasan et al., 1982). However, provided the correlation between inputs stays above the critical value (marked by points in Figure 3 for different noise values), the number of neuron classes will not change. Neurons will just change their relevant features (the stimulus features will “rotate” within the input space). Decreasing the signal-to-noise ratio can also make input signals more Gaussian, which would initiate the change from the relevant features that are aligned with independent components to features that are aligned with principal components (Karklin and Simoncelli, 2011). Upon a further increase in the noise level, neurons might temporarily switch their features to encode the same input dimension. Depending on the noise level, their thresholds might temporarily merge, but that does not mean that they belong to the same class, provided environmental conditions exist where the neural responses differ qualitatively. Thus, the number of neuronal classes is determined by the smallest noise level that the circuit may experience in different environmental conditions.

Overall, the information-theoretic arguments discussed above highlight the difficulties and opportunities for creating a taxonomy of neuronal classes that spans across brain regions and species. These difficulties stem from the fact that both features and nonlinearities are context-dependent, and might change depending on experimental conditions. Functional differences among neurons might be prominent in some experiments and absent in other. A theoretical framework describing how the features and thresholds should change could help pinpoint critical experiments that can reveal differences between classes of neurons that may have been missed up to now. In turn, a more detailed physiological description could help bridge differences in classification based on electrophysiological (Markram et al., 2004), molecular (Trimarchi et al., 2007), and anatomical parameters (Masland, 2004). Notable in that regard is the recent computational effort to characterize neuronal types based on statistical analysis of connectomics datasets (Jonas and Kording, 2014). At the same time, many opportunities exist for creating a catalogue of optimal neural classes that could help relate differences and commonalities that exist across species. For example, in the case of retinal processing, where the neural classes are perhaps known in the most exquisite detail (Masland, 2001), noticeable differences between neural types in the rabbit, mouse, or primate exist. While these differences have largely been left unexplained so far, perhaps a theoretical framework that takes into account differences in environmental niches could offer clues. More generally, it now seems feasible to obtain a unified framework that can bridge the divide between encoding using identified neurons in invertebrate circuits to encoding based on neural populations in mammalian and other species. Assuming that the intrinsic noise level is similar across species, one could potentially find a mapping between

single identified neurons within the smaller invertebrate neural circuit with the corresponding populations of neurons in larger mammalian circuits.

Acknowledgments

The author thanks David Kastner, Stephen Baccus, and John Berkowitz for many helpful conversations. The work was supported by the NSF Career Award 1254123 and NIH grant R01EY019493.

References

- Adelman TL, Bialek W, Olberg RM. The information content of receptive fields. *Neuron*. 2003; 40:823–833. [PubMed: 14622585]
- Ala-Laurila P, Greschner M, Chichilnisky EJ, Rieke F. Cone photoreceptor contributions to noise and correlations in the retinal output. *Nat Neurosci*. 2011; 14:1309–1316. [PubMed: 21926983]
- Aljadeff J, Segev R, Berry MJ 2nd, Sharpee TO. Spike triggered covariance in strongly correlated gaussian stimuli. *PLoS Comput Biol*. 2013; 9:e1003206. [PubMed: 24039563]
- Atick JJ, Li Z, Redlich AN. Understanding retinal color coding from first principles. *Neural Comput*. 1992; 4:559–572.
- Atick JJ, Redlich AN. Towards a theory of early visual processing. *Neural Comput*. 1990; 2:308–320.
- Atick JJ, Redlich AN. What does the retina know about natural scenes? *Neural Comput*. 1992; 4:196–210.
- Balasubramanian V, Sterling P. Receptive fields and functional architecture in the retina. *J Physiol*. 2009; 587:2753–2767. [PubMed: 19525561]
- Baraban SC, Tallent MK. Interneuron Diversity series: Interneuronal neuropeptides--endogenous regulators of neuronal excitability. *Trends Neurosci*. 2004; 27:135–142. [PubMed: 15036878]
- Barlow H. Redundancy reduction revisited. *Network*. 2001; 12:241–253. [PubMed: 11563528]
- Barlow, H.; Foldiak, P. Adaptation and decorrelation in the cortex. In: Durbin, R.; Miall, C.; Mitchinson, G., editors. *The computing neuron*. New York: Addison-Wesley; 1989.
- Barlow, HB. Possible principles underlying the transformation of sensory messages. In: Rosenblith, WA., editor. *Sensory Communication*. Cambridge, MA: MIT Press; 1961. p. 217-234.
- Bell AJ, Sejnowski TJ. The “independent components” of natural scenes are edge filters. *Vision Res*. 1997; 37:3327–3338. [PubMed: 9425547]
- Bialek W. Physical limits to sensation and perception. *Ann Rev Biophys Biophys Chem*. 1987; 16:455–478. [PubMed: 3297091]
- Bialek, W.; de Ruyter van Steveninck, RR. Features and dimensions: Motion estimation in fly vision. 2005. p. 1-18. arxiv.org/abs/q-bio.NC/0505003
- Bialek W, Owen WG. Temporal filtering in retinal bipolar cells. Elements of an optimal computation? *Biophys J*. 1990; 58:1227–1233. [PubMed: 2291942]
- Boycott BB, Waessle H. Morphological classification of bipolar cells of the primate retina. *European J Neurosci*. 1991; 3:1069–1088. [PubMed: 12106238]
- Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck RR. Synergy in a neural code. *Neural Comput*. 2000; 12:1531–1552. [PubMed: 10935917]
- Brunel N, Nadal JP. Mutual information, Fisher information, and population coding. *Neural Comput*. 1998; 10:1731–1757. [PubMed: 9744895]
- Chance FS, Nelson SB, Abbott LF. Complex cells as cortically amplified simple cells. *Nat Neurosci*. 1999; 2:277–282. [PubMed: 10195222]
- Chichilnisky EJ, Kalmar RS. Functional asymmetries in ON and OFF ganglion cells of primate retina. *J Neurosci*. 2002; 22:2737–2747. [PubMed: 11923439]
- Chichilnisky EJ, Wandell BA. Trichromatic opponent color classification. *Vision Res*. 1999; 39:3444–3458. [PubMed: 10615508]
- Cover, TM.; Thomas, JA. *Elements of Information Theory*. New York: Wiley-Interscience; 1991.

- Davis TL, Sterling P. Microcircuitry of cat visual cortex: Classification of neurons in layer {IV} of area 17, and identification of the patterns of lateral geniculate input. *J Comp Neurol.* 1979; 188:599–628. [PubMed: 521508]
- Dayan, P.; Abbott, LF. *Theoretical neuroscience: Computational and mathematical modeling of neural systems.* Cambridge, Massachusetts: The MIT Press; 2001.
- Derrington AL, Krauskopf J, Lennie P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *J Physiol.* 1984; 357:241–265. [PubMed: 6512691]
- Doi E, Gauthier JL, Field GD, Shlens J, Sher A, Greschner M, Machado TA, Jepson LH, Mathieson K, Gunning DE, et al. Efficient coding of spatial information in the primate retina. *J Neurosci.* 2012; 32:16256–16264. [PubMed: 23152609]
- Fitzgerald JD, Sharpee TO. Maximally informative pairwise interactions in networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2009; 80:031914. [PubMed: 19905153]
- Fitzgerald JD, Sincich LC, Sharpee TO. Minimal models of multidimensional computations. *PLoS Comput Biol.* 2011; 7:e1001111. [PubMed: 21455284]
- Garrigan P, Ratliff CP, Klein JM, Sterling P, Brainard DH, Balasubramanian V. Design of a trichromatic cone array. *PLoS Comput Biol.* 2010; 6:e1000677. [PubMed: 20168996]
- Gawne TJ, Richmond BJ. How independent are the messages carried by adjacent inferior temporal cortical neurons? *J Neurosci.* 1993; 13:2758–2771. [PubMed: 8331371]
- Globerson A, Stark E, Vaadia E, Tishby N. The minimum information principle and its application to neural code analysis. *Proc Natl Acad Sci U S A.* 2009; 106:3490–3495. [PubMed: 19218435]
- Harper NS, McAlpine D. Optimal neural population coding of an auditory spatial cue. *Nature.* 2004; 430:682–686. [PubMed: 15295602]
- Hochstein S, Shapley RM. Quantitative analysis of retinal ganglion cell classifications. *J Physiol.* 1976; 262:237–264. [PubMed: 994039]
- Horwitz GD, Hass CA. Nonlinear analysis of macaque V1 color tuning reveals cardinal directions for cortical color processing. *Nat Neurosci.* 2012; 15:913–919. [PubMed: 22581184]
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962; 160:106–154. [PubMed: 14449617]
- Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol.* 1968; 195:215–243. [PubMed: 4966457]
- Jonas E, Kording K. Automatic discovery of cell types and microcircuitry from neural connectomics. 2014 arXiv:1407.4137.
- Karklin, Y.; Simoncelli, EP. Efficient coding of natural images with a population of noisy Linear-Nonlinear neurons. Paper presented at: *Advances in Neural Information Processing Systems*; Denver. 1989; Curran Associates, Inc; 2011.
- Kastner DB, Baccus SA, Sharpee TO. Critical and maximally informative encoding between neural populations in the retina. 2014 arXiv: 1409.2604 [q-bio.NC].
- Kastner DB, Baccus SA. Coordinated dynamic encoding in the retina using opposing forms of plasticity. *Nat Neurosci.* 2011; 14:1317–1322. [PubMed: 21909086]
- Kodama T, Guerrero S, Shin M, Moghadam S, Faulstich M, du Lac S. Neuronal classification and marker gene identification via single-cell expression profiling of brainstem vestibular neurons subserving cerebellar learning. *J Neurosci.* 2012; 32:7819–7831. [PubMed: 22674258]
- Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci.* 2004; 5:793–807. [PubMed: 15378039]
- Masland RH. The functional architecture of the retina. *Scientific American.* 1986 Dec.:102–111. [PubMed: 2881349]
- Masland RH. The fundamental plan of the retina. *Nat Neurosci.* 2001; 4:877–886. [PubMed: 11528418]
- Masland RH. Neuronal cell types. *Curr Biol.* 2004; 14:R497–500. [PubMed: 15242626]
- McDonnell MDGSN, Pearce CEM, Abbott D. Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance. *Physics Letters.* 2006; 352:183–189.
- Mechler F, Ringach DL. On the classification of simple and complex cells. *Vis Res.* 2002; 42:1017–1033. [PubMed: 11934453]

- Mott DD, Dingledine R. Interneuron Diversity series: Interneuron research--challenges and strategies. *Trends Neurosci.* 2003; 26:484–488. [PubMed: 12948659]
- Nelson SB, Sugino K, Hempel CM. The problem of neuronal cell types: a physiological genomics approach. *Trends Neurosci.* 2006; 29:339–345. [PubMed: 16714064]
- Nikitin AP, Stocks NG, Morse RP, McDonnell MD. Neural population coding is optimized by discrete tuning curves. *Phys Rev Lett.* 2009; 103:138101. [PubMed: 19905542]
- Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature.* 1996; 381:607–609. [PubMed: 8637596]
- Pitkow X, Meister M. Decorrelation and efficient coding by retinal ganglion cells. *Nat Neurosci.* 2012; 15:628–635. [PubMed: 22406548]
- Potters M, Bialek W. Statistical mechanics and visual signal processing. *Journal de Physique I.* 1994; 4:1755.
- Rajan K, Bialek W. Maximally informative “stimulus energies” in the analysis of neural responses to natural signals. *PLoS One.* 2013; 8:e71959. [PubMed: 24250780]
- Ratliff CP, Borghuis BG, Kao YH, Sterling P, Balasubramanian V. Retina is structured to process an excess of darkness in natural scenes. *Proc Natl Acad Sci U S A.* 2010; 107:17368–17373. [PubMed: 20855627]
- Rieke, F.; Owen, WG.; Bialek, W. Optimal filtering in the salamander retina. Paper presented at: *Advances in Neural Information Processing Systems*; Denver. 1989; San Mateo: Morgan Kaufmann; 1991.
- Ruderman DL, Bialek W. Statistics of natural images: Scaling in the woods. *Phys Rev Lett.* 1994; 73:814–817. [PubMed: 10057546]
- Salinas E. How behavioral constraints may determine optimal sensory representations. *PLoS Biol.* 2006; 4:e387. [PubMed: 17132045]
- Schulz DJ, Goaillard JM, Marder E. Variable channel expression in identified single and electrically coupled neurons in different animals. *Nat Neurosci.* 2006; 9:356–362. [PubMed: 16444270]
- Schwartz O, Pillow JW, Rust NC, Simoncelli EP. Spike-triggered neural characterization. *J Vis.* 2006; 6:484–507. [PubMed: 16889482]
- Segev R, Puchalla J, Berry MJ 2nd. Functional organization of ganglion cells in the salamander retina. *J Neurophysiol.* 2006; 95:2277–2292. [PubMed: 16306176]
- Sharpee T, Rust NC, Bialek W. Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Comput.* 2004; 16:223–250. [PubMed: 15006095]
- Sharpee TO. Computational identification of receptive fields. *Annu Rev Neurosci.* 2013; 36:103–120. [PubMed: 23841838]
- Sharpee TO, Bialek W. Neural Decision Boundaries for Maximal Information Transmission. *PLOS One.* 2007
- Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annu Rev Neurosci.* 2001; 24:1193–1216. [PubMed: 11520932]
- Sincich LC, Horton JC. The circuitry of V1 and V2: integration of color, form, and motion. *Annu Rev Neurosci.* 2005; 28:303–326. [PubMed: 16022598]
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am.* 2003; 114:3394–3411. [PubMed: 14714819]
- Srinivasan MV, Laughlin SB, Dubs A. Predictive coding: A fresh view of inhibition in the retina. *Proc R Soc Lond.* 1982; B216:427–459. [PubMed: 6129637]
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network.* 2001; 12:289–316. [PubMed: 11563531]
- Tkacik G, Callan CG Jr, Bialek W. Information flow and optimization in transcriptional regulation. *Proc Natl Acad Sci U S A.* 2008; 105:12265–12270. [PubMed: 18719112]
- Todorov E. Optimality principles in sensorimotor control. *Nat Neurosci.* 2004; 7:907–915. [PubMed: 15332089]

- Trimarchi JM, Stadler MB, Roska B, Billings N, Sun B, Bartch B, Cepko CL. Molecular heterogeneity of developing retinal ganglion and amacrine cells revealed through single cell gene expression profiling. *J Comp Neurol.* 2007; 502:1047–1065. [PubMed: 17444492]
- van Hateren JH, Ruderman DL. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proc R Soc Lond B Biol Sci.* 1998; 265:2315–2320.
- Werblin FS, Dowling JE. Organization of the retina of the mudpuppy, *Necturus maculosus*. II. Intracellular recording. *J Neurophysiol.* 1969; 32:339–355. [PubMed: 4306897]
- Wiesel TN, Hubel DH. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J Neurophysiol.* 1966; 29:1115–1156. [PubMed: 4961644]

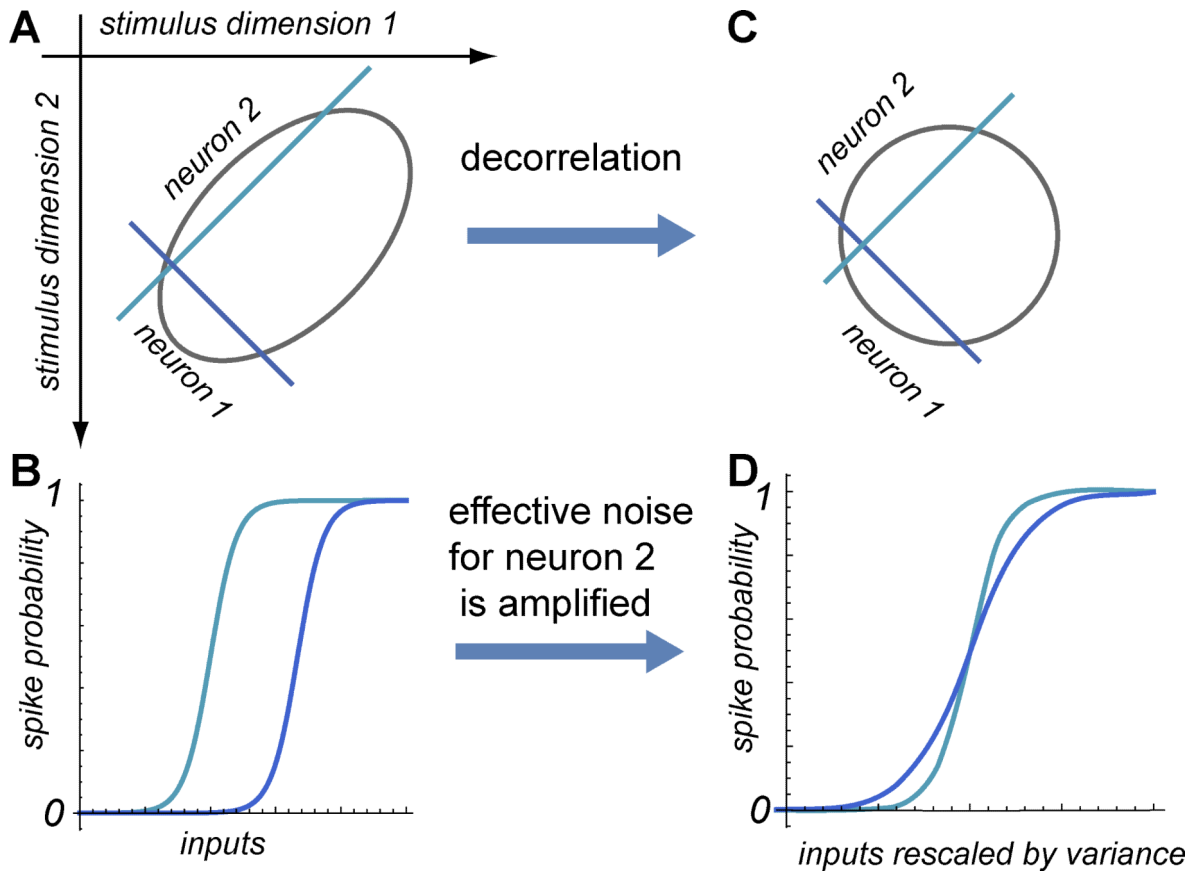


Figure 1. Decorrelation amplifies the effect of noise in neural coding

(A) One of the possible optimal configurations of neural thresholds in the presence of correlations. Neuron 1 encodes the input value along the direction of largest variance, whereas neuron 2 encodes the input value along the other dimension. (B) The threshold values of the two neurons are proportional to the input variances along their respective relevant dimensions. (C and D) are equivalent to (A) and (B) except that the input scales are rescaled by their respective variances. This representation emphasizes that decorrelation increases the effective noise for neuron 2.

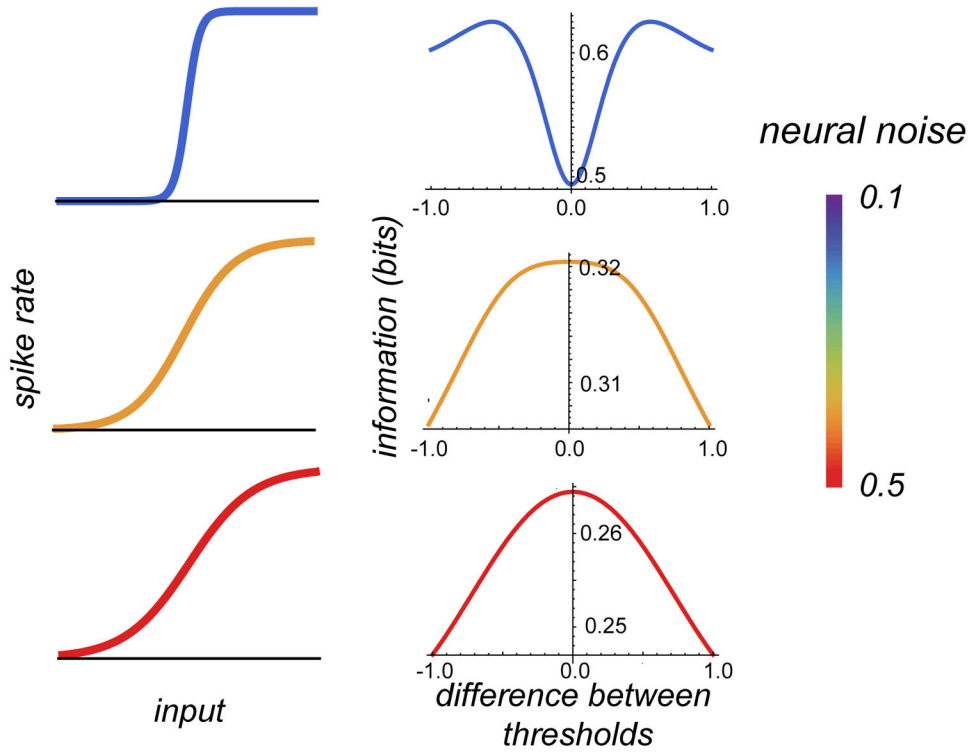


Figure 2. Separate classes of neurons encoding the same input dimension can appear with decreasing neural noise
 The noise values measured relative to the input standard deviation are $\nu=0.1, 0.3, 0.5$, top to bottom row. The probability of a spike from a neuron with a threshold μ and noise value ν is modeled using logistic function $P(\text{spike}|x)=\frac{1}{1+\exp(\frac{x-\mu}{\nu})}$, with x is a Gaussian input of unit variance. For two neurons, there are four different patterns (00, 01, 10, 11 where each neuron is assigned response 0 or 1 if it produces a spike). Neurons are assumed to be conditionally independent given x . With these assumptions, it is possible to evaluate mutual information as the difference between response entropy $R = -p_{00}\log_2 p_{00} - p_{01}\log_2 p_{01} - p_{10}\log_2 p_{10} - p_{11}\log_2 p_{11}$ and noise entropy

$$\begin{aligned}
 N = & \int dx \frac{1}{1+\exp(\frac{\mu_1-x}{\nu})} \log_2 \left[1+\exp\left(\frac{\mu_1-x}{\nu}\right) \right] \\
 & + \int dx \frac{1}{1+\exp(\frac{x-\mu_1}{\nu})} \log_2 \left[1+\exp\left(\frac{x-\mu_1}{\nu}\right) \right] \\
 & + \int dx \frac{1}{1+\exp(\frac{\mu_2-x}{\nu})} \log_2 \left[1+\exp\left(\frac{\mu_2-x}{\nu}\right) \right] \\
 & + \int dx \frac{1}{1+\exp(\frac{x-\mu_2}{\nu})} \log_2 \left[1+\exp\left(\frac{x-\mu_2}{\nu}\right) \right], \text{ where } \mu_1 \text{ and } \mu_2 \text{ are the thresholds for}
 \end{aligned}$$

the neurons 1 and 2, respectively. The mutual information $I=R-N$ is plotted here as a function of the difference between thresholds $\mu_1 - \mu_2$. For each value of $\mu_1 - \mu_2$ and ν , the average threshold between the two neurons is adjusted to ensure that the spike rate remains constant for all information values shown. In these simulation, the average spike rate across

all inputs x and summed for neurons 1 and 2 was set to 0.2. In this case, transition from redundant coding observed for large noise values to coding based on different thresholds (observed at low noise values) occurs at $\nu \sim 0.4$.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

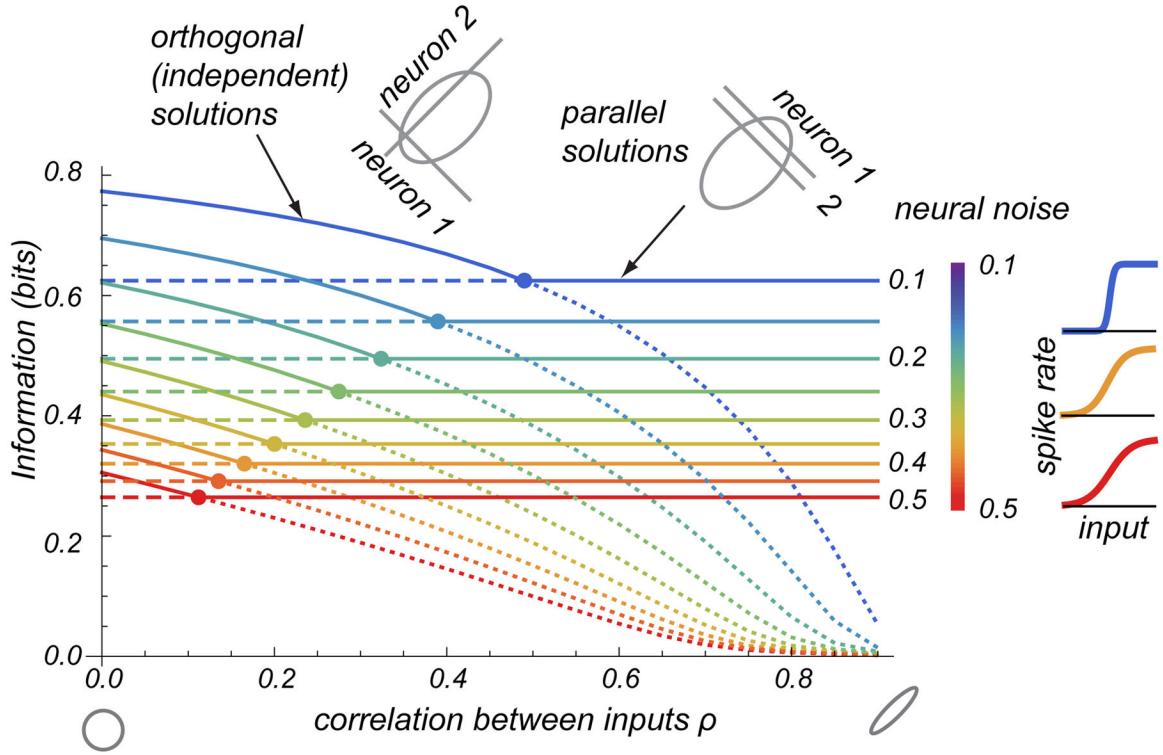


Figure 3. Cross-over between different encoding schemes

In scheme 1, neuron classes are defined based on different input. The mutual information depends on both correlation between inputs and the neural noise level. In scheme 2, neuron classes are defined based on differences in thresholds for the same input feature. The mutual information is independent of stimulus correlations and is shown here with lines parallel to the x-axis. Color denotes the noise level relative to the standard deviation along the dimension of largest variance. All simulations have been matched to have the same average spike rate across inputs (values are as in Figure 2). Suboptimal solutions (after the cross-over) are shown using dashed and dotted lines, respectively for the case of parallel and orthogonal solutions.