nature reviews neuroscience

https://doi.org/10.1038/s41583-024-00795-0

Perspective

Check for updates

Response sub-additivity and variability quenching in visual cortex

Robbe L. T. Goris [1] ✉, Ruben Coen-Cagli [2,3,4], Kenneth D. Miller [5,6,7,8,9], Nicholas J. Priebe [10] & Máté Lengyel [11,12]

Abstract

Sub-additivity and variability are ubiquitous response motifs in the primary visual cortex (V1). Response sub-additivity enables the construction of useful interpretations of the visual environment, whereas response variability indicates the factors that limit the precision with which the brain can do this. There is increasing evidence that experimental manipulations that elicit response sub-additivity often also quench response variability. Here, we provide an overview of these phenomena and suggest that they may have common origins. We discuss empirical findings and recent model-based insights into the functional operations, computational objectives and circuit mechanisms underlying V1 activity. These different modelling approaches all predict that response sub-additivity and variability quenching often co-occur. The phenomenology of these two response motifs, as well as many of the insights obtained about them in V1, generalize to other cortical areas. Thus, the connection between response sub-additivity and variability quenching may be a canonical motif across the cortex.

Sections

[1]Center for Perceptual Systems, University of Texas at Austin, Austin, TX, USA. [2]Department of Systems and Computational Biology, Albert Einstein College of Medicine, Bronx, NY, USA. [3]Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY, USA. [4]Department of Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, Bronx, NY, USA. [5]Center for Theoretical Neuroscience, Columbia University, New York, NY, USA. [6]Kavli Institute for Brain Science, Columbia University, New York, NY, USA. [7]Dept. of Neuroscience, College of Physicians and Surgeons, Columbia University, New York, NY, USA. [8]Morton B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA. [9]Swartz Program in Theoretical Neuroscience, Columbia University, New York, NY, USA. [10]Center for Learning and Memory, University of Texas at Austin, Austin, TX, USA. [11]Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge, UK. [12]Center for Cognitive Computation, Department of Cognitive Science, Central European University, Budapest, Hungary. ✉e-mail: Robbe.Goris@utexas.edu

Introduction

The primary visual cortex (V1) has long been a model system for studying cortical circuitry and computations. In recent decades, two major foci of study in V1 have been response sub-additivity and response variability. Response sub-additivity involves phenomena in which the neuronal response to two simultaneously presented stimuli is less than the sum of the responses to the two stimuli presented independently (also referred to as sublinear response summation). For example, V1 cells have distinct spatial receptive fields (defined as the locations in visual space in which the presentation of a stimulus elicits an increase in activity), beyond which lies the receptive field surround (Fig. 1a). Although ineffective by itself, the presentation of a stimulus within the surround often suppresses the response to a stimulus presented within the receptive field[1] (Fig. 1a). Similarly, for many neurons, doubling the contrast of a medium contrast stimulus presented within the classical receptive field does not double the neural response. Thus, sub-additive effects occur for stimuli presented outside[1–4] and within[5–9] the classical receptive field, and for preferred and non-preferred stimulus orientations[3,6,7,9]. The sub-additivity of responses to a compound stimulus is better described as a divisive than as a subtractive interaction between the responses to its two constituent stimuli[10,11]. Response variability involves phenomena in which repeated presentations of the same stimulus elicit variable responses in cortical cells (Fig. 1a). This variability is evident both in the cells' membrane potential[12–14] and in their spiking activity[15–17]. Response variability in the visual cortex appears largely random[18,19], exhibits strong dependence on features of the stimulus (such as its contrast)[20,21], has a non-trivial spatio-temporal structure[22–27] and is often well described by a doubly stochastic process model of spike generation[28–33].

Response sub-additivity and variability feature prominently in the literature because they provide directly observable indications of the brain processes that enable us to perform natural perceptual tasks on the one hand (nonlinear neural transformations[34,35]) and those that are traditionally believed to limit our ability to do so on the other hand (neural information loss[36,37]). These response motifs have largely been studied independently from each other. However, it has become apparent that experimental manipulations that elicit sub-additivity often also change variability. In particular, response sub-additivity often co-occurs with variability quenching (a decrease in neural variability at stimulus onset, or when two stimuli are presented simultaneously compared with either being presented alone). This has been observed for experimental manipulations of stimuli within[21,33] and beyond the classical receptive field[38–41] (Fig. 1b,c). We propose that this is not mere coincidence. Recent theoretical studies of the functional operations, representational objectives and circuit mechanisms underlying V1 activity all suggest that response sub-additivity and variability quenching may be intimately connected. The aim of this Perspective is to describe what is known about this connection. We first discuss the connection between sub-additivity and quenching through the lens of models that seek to offer an economic description of the transformations that govern stimulus–response relations in V1. Next, we consider this connection from the viewpoint of models that seek to identify the computational and representational goals that shape V1 activity. Finally, we discuss this connection from a mechanistic perspective on cortical circuitry and computation. Note that we primarily discuss data collected in cat and monkey. Recent work in rodent V1 has revealed similar sub-additive phenomena (and is rapidly advancing our understanding of the underlying circuit mechanisms[42–48]), but studies of response variability in rodent V1 have, thus far, been less extensive than in these other species (but see refs. 49–51).

Experimental observations

Response sub-additivity

In cat and monkey V1, layer 4 neurons exhibit numerous response properties that are fundamentally different from those of the thalamic neurons from which they receive feedforward inputs. These properties include selectivity for the orientation, direction of motion, and distance of a stimulus[52,53]. A longstanding view, pioneered by Hubel and Wiesel, holds that this selectivity arises from the alignment of the receptive fields of the presynaptic thalamic relay cells[52]. In its simplest form, this framework predicts that V1 receptive fields perform a spatio-temporal linear filtering operation on visual input, followed by a thresholding operation to transform intracellular signals into spikes. This model for V1 responses is at once simple, elegant and powerful. It enabled neuroscientists to approach a fundamental biological question — how do cortical sensory circuits transform their input into a novel representation of the visual environment? — through the principled abstraction of a linear system at a time when little was known about the cortical representation of visual information. The chief benefit of this approach is that it readily generates quantitative predictions of the responses that will be elicited by arbitrary visual stimuli, and, to a first approximation, these predictions are quite good. For example, a model that includes a linear receptive field and a static threshold nonlinearity can explain V1 selectivity for elementary stimulus attributes such as position, scale, orientation, and speed and direction of motion[54]. However, cortical cells also exhibit clear violations of linearity that go beyond a static threshold nonlinearity: these often manifest through the phenomenology of sub-additivity.

One prominent example of sub-additivity arises when a masking stimulus is superimposed on a stimulus whose orientation, spatial frequency and size match the preference of the cell under study (Fig. 1c). Masking stimuli can suppress responses to the preferred stimulus across a broad range of spatial frequencies[6,8], orientations[6,7] and temporal frequencies[7,55,56]. As discussed in the 'Introduction', responses of V1 neurons to stimuli presented within the receptive field can also be substantially diminished by stimuli presented outside the receptive field[2,3,11,57]. The strength of this 'surround suppression' depends on the exact position of the surround stimulus (the larger the distance from the receptive field centre, the weaker the suppression[58]) and its similarity to the stimulus presented within the classical receptive field (the larger the resemblance, the stronger the suppression[4,11,59,60]).

A different example of sub-additivity comes from contrast summation experiments (Fig. 1c), in which the same stimulus is presented at various contrasts. For a linear system, scaling the contrast of an effective stimulus will scale the response by the same factor, a property known as 'response homogeneity' in linear systems analysis. However, this is not what happens in the visual cortex. Instead, with increasing contrast, the responses of V1 cells typically grow at a faster than linear rate for stimuli with low contrasts. Above a threshold (typically relatively low) level of contrast, the responses grow at a slower than linear rate and, at higher contrasts, they may approach saturation[5,9]. This sublinear or saturated response is immediately present at response onset and occurs for preferred and non-preferred stimuli[61]. Thus, as soon as stimulus contrast exceeds a low level, sub-additivity is a general property of cortical contrast summation.

Sub-additivity induced by stimuli in the receptive field centre and that induced by stimuli in the surround share important properties,

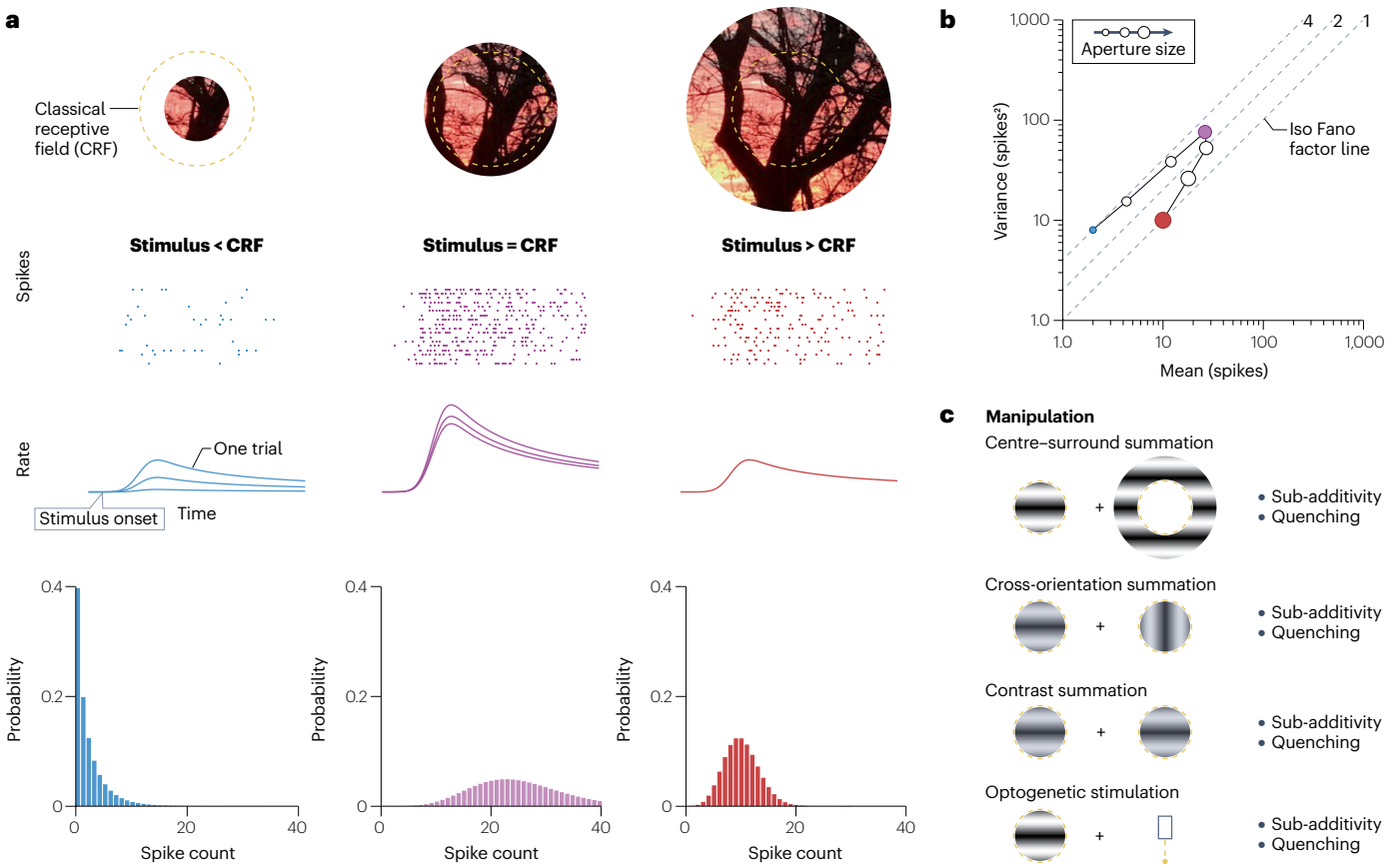

Fig. 1 | Response sub-additivity and variability quenching co-occur under various experimental manipulations. a, Simulated responses of a V1 neuron to three patches cropped from the same image using differently sized apertures. Details of the model simulation are described in Supplementary Information. In brief, spikes were generated by simulating a Poisson point process with an underlying stimulus-dependent rate that is multiplied by a stimulus-independent response gain that varies across trials[28]. Spike times are plotted as a raster (second row), with each tick representing one spike. Each row depicts a repeated presentation of the same stimulus. Different aperture sizes are associated with different levels of responsiveness, as also shown in neurons recorded in vivo[3,4]. Within each raster, there is considerable variability in spiking activity across trials, again replicating the findings of experiments in vivo[12,17]. This variability in part arises from variability in firing rate (third row). The rate variability decreases with stimulus size, as is the case for real V1 cells both at the level of spike counts[39] and, relatedly[131,184], membrane potentials[13,21]. The associated spike count distributions (fourth row) were obtained by using a counting window whose length matches the total stimulus duration. In experimental studies, response mean and response variance are computed from this distribution, not from the generative process itself. **b**, Spike count variance plotted as a function of spike count mean for the simulated size tuning experiment shown in part **a**. Larger symbol sizes indicate larger stimulus apertures. Responses initially increase with stimulus size, but begin to decrease as the stimulus engages the suppressive surround. This suppressive effect co-occurs with a change in the relative amount of response variability, measured as the variance-to-mean ratio (the Fano factor), as also seen in experiments in vivo[39]. On these logarithmic axes, lines with a slope equal to one trace out a constant Fano factor (dashed lines show Fano factors of 1, 2 and 4). The blue, purple and red circles indicate the three conditions shown in part **a**. **c**, Summary of some other classical experimental manipulations that elicit both response sub-additivity[3,5,6,41] and variability quenching[21,33,39,41]. In these experiments, a stimulus presented in the classical receptive field is combined with either an annulus presented in the surround (top row), an orthogonally oriented mask (second row), an identical mask (third row) or optogenetic stimulation of the cortex (bottom row, icon on the right represents a laser).

but also differ in critical regards. In both cases, the masking stimulus tends to act as if it divides the contrast of the stimulus presented in the receptive field by a constant fraction[3,9,62]. This observation has motivated attempts to build stimulus–response models of V1 activity that capture both types of sub-additive effects with a single model component. However, surround suppression is partially delayed relative to response onset[63–65], exhibits interocular transfer[64] and is modified by contrast adaptation[64]. None of this is true for within-receptive-field violations of linearity[61,64,66] (though see ref. 67), suggesting that the two types of sub-additive effects have distinct mechanistic origins.

V1 response sub-additivity might arise at any stage of the visual pathway: the retina, lateral geniculate nucleus of the thalamus, or V1. Recent studies have explored the mechanistic origin of V1 response sub-additivity by combining visual stimuli with direct optogenetic stimulation of the visual cortex (Fig. 1c). In macaque and marmoset V1, responses to optogenetic and visual stimulation combine sub-additively[41,68], suggesting that cortical circuits contribute

to at least some forms of sub-additivity. Note, however, that in these studies, the illuminated patch of cortex was large enough to engage lateral connections thought to be involved in surround suppression[57]. Results may differ if the optogenetic stimulation is confined to the receptive field centre.

Response variability

Sensory neurons transmit information about the external world via sequences of action potentials that are inherently variable. This is also true of area V1: repeated presentations of identical visual stimuli elicit different patterns of spiking activity[16,17,20]. The origins of this variability are still unknown, but its consequences may be profound. If this variability represents irreducible noise, then it will limit the reliability with which neural populations can represent sensory events and, ultimately, the capacity of the organism to perform perceptual tasks. Motivated by this insight, physiologists have made considerable efforts to directly compare neuronal and psychophysical sensitivity — an enterprise pioneered by Werner and Mountcastle[69]. This comparison is most meaningful when it involves neurons that are suitably tuned for the task under consideration and when sensitivity estimates are based on physiological and behavioural data obtained in the same animals from the same set of trials[70]. Studies that meet these criteria have consistently reported that the sensitivity of some individual sensory neurons rivals the behavioural capacity of well-trained macaque monkeys. For example, the ability of V1 cells to signal changes in stimulus orientation closely approximates perceptual orientation acuity[71–73], whereas cells in visual cortical area MT exhibit sensitivity for visual motion that is very similar to perceptual motion sensitivity[74]. This is puzzling if neural variability is considered simple noise, as one would expect that behavioural decisions should rely on averaging the noisy activities of many neurons and they should, thus, be more reliable than the responses of individual neurons. By violating this expectation, these findings sparked an enormous interest in the origins and role of neural response variability. Here, too, many found it fruitful to approach these fundamental biological questions by considering a principled abstraction that generates quantitative predictions for arbitrary stimuli: the Poisson point process.

Which aspects of a spike train are signal and which are noise? One extreme possibility is that only the number of spikes realized during a temporal interval matters, and that there is no information in the exact timing of each spike[18]. This concept is formalized by the Poisson point process. The Poisson point process is the simplest stochastic point process and is fully characterized by a single firing rate parameter that represents a reproducible response to a sensory stimulus. If this rate is fixed over time, the process is said to be homogeneous; if it varies over time, it is inhomogeneous[75]. Both variants give rise to Poisson-distributed spike counts. A hallmark of this distribution is a spike count variance across repeated measurements that matches the spike count mean, regardless of the length or placement of the time interval over which spikes are counted. In other words, the ratio of the variance to the mean, a statistic known as the Fano factor, is always one. In visual cortex, this prediction enjoys some support. Spike count variance often approximately equals the mean[17,76]. However, some cortical spiking statistics exhibit clear deviations from a Poisson distribution. This most commonly manifests in the form of excess variance. When the mean count is high, either owing to a high firing rate or owing to the use of a long counting window, super-Poisson variability (more variability than expected from a Poisson process) becomes apparent[17,28,31,76]. Statistically, both cases of super-Poisson variability can be explained by extending the Poisson process with a slowly fluctuating gain signal that modulates the rate and varies from trial to trial[28,29] (thus, creating a doubly stochastic process known as the 'modulated Poisson model', Fig. 1a). Empirically, however, the dependency of response variance on response mean makes it difficult to identify changes in response variability that are not simply a consequence of changes in response mean. One way to overcome this challenge is to estimate the Fano factor using an analysis procedure that corrects for differences in mean response level[22] (but see ref. 77). This statistic is called the mean-matched Fano factor. It generally exceeds 1 in cortex and, under the Poisson assumption, represents a measure of cross-trial variability in firing rate.

Firing-rate variability is stimulus dependent in a manner that resembles phenomena of sub-additivity. Across the cortex, it is maximal in the absence of stimulation and decreases rapidly following stimulus onset[22]. The magnitude of the decrease depends on the amount of stimulus energy. For example, in area V1, low-contrast stimuli placed within the receptive field of a neuron are associated with stronger rate fluctuations than high-contrast stimuli[21,33]. Such variability quenching occurs for preferred and non-preferred stimuli that drive a neuron[21,22,33], thus resembling response saturation in contrast summation experiments. It also occurs for stimuli that do not drive a neuron[22], thus resembling the broad tuning of suppressive effects in masking experiments. Finally, stimuli presented outside of the receptive field can quench neural response variability beyond the reduction of variability caused by increasing stimulation inside the receptive field[39] (Fig. 1), thus resembling surround suppression. The strength of this effect weakly depends on the similarity between the centre and surround stimulus[39]. A recent preprint has reported that the effect also depends on the cortical layer[78] and it has also been suggested that it depends on the size and location of the surround stimulus[40,78].

The co-occurrence of response sub-additivity and variability quenching can be illustrated by plotting the relationship between the variance and the mean for stimuli of different sizes[33] (see simulation in Fig. 1b). This has shown that increases in stimulus size initially increase both the variance and the mean of the spike count response, but decrease their ratio (the Fano factor)[39]. Further increases in stimulus size cause surround suppression, reducing both the variance and the mean spike count, and monotonically decrease the Fano factor[39]. This relationship is illustrated in Fig. 1b, in which the simulated activity eventually reaches the line corresponding to a Fano factor equal to 1.

Is it a coincidence that experimental manipulations that elicit response sub-additivity often also quench variability (Fig. 1c)? Are these phenomena partly related or are they distinct manifestations of shared underlying mechanisms? These questions are difficult to answer because we cannot directly observe the signals of interest. There is no empirical measurement that directly reveals the strength of the 'sub-additive signal' of a cell, whereas firing rate variability is a statistical construct that cannot be mapped onto an observable biophysical quantity. Answering these questions, thus, requires a theoretical exploration of the issues at stake.

Models of V1 activity

The trail-blazing work of Hubel and Wiesel inspired many to build, test and refine models of V1 activity. There is a great deal of diversity among these models, reflecting differences in their underlying aspirations. For example, some models seek to explain V1 responses in a manner that remains faithful to known physiological mechanisms, thus revealing how the structure of neural circuits gives rise to their function[54,79–82].

We will refer to such models as 'mechanistic' accounts of V1 activity. Other models seek to explain V1 responses on the basis of theoretical coding principles[21,83–88], thereby revealing the computational objectives that shape neural function ('normative' accounts). Conversely, other models aspire to describe quantitatively the transformation of visual stimuli into neural responses using a limited set of operations and parameters that can be fit to neural data[9,89–93] ('descriptive' accounts).

Descriptive accounts of response sub-additivity and variability quenching in V1

Descriptive models are useful to simulate V1 activity and, hence, can provide insight into V1's representation of visual information that goes beyond experimentally feasible measurements[94]. They are also an essential point of comparison for studies that aim to connect V1 representations to downstream transformations[95], perceptual capabilities[96,97], other sensory modalities[98] and artificial visual systems[99]. We will focus here on one prominent descriptive framework, the 'normalization' model[90,98,100]. This model proposes that the firing rate of V1 neurons is determined by the ratio of the output of a narrowly tuned excitatory channel to that of a broadly tuned inhibitory channel. The excitatory channel usually consists of a linear spatio-temporal stimulus filtering operation followed by a nonlinear response pooling operation and determines the stimulus selectivity of the neuron (Fig. 2a). The inhibitory channel is built from the same operations but has much weaker tuning. Specifically, it processes visual input over a larger area of visual space than the excitatory channel and is weakly tuned for orientation and spatial phase (Fig. 2a). Because this model incorporates a divisive operation, its responses saturate with increasing stimulus contrast – that is, once stimulus contrast exceeds a low level, the responses to higher stimulus contrast are sub-additive[90,100]. This occurs for preferred and non-preferred stimuli, as is the case for real V1 cells[5,9]. Moreover, because the inhibitory channel is broadly tuned, the normalization model also exhibits cross-orientation suppression and surround suppression[3,86,90]. Importantly, in the normalization model, all sub-additive effects arise from a single operation. This is an extreme proposition, yet it provides a remarkably accurate description of classic sub-additive phenomena[3,9].

As discussed in 'Response variability' section, V1 neurons often exhibit super-Poisson variability in a manner that resembles the effects of a noisy response gain. This behaviour naturally arises in the normalization model when we assume that spikes arise from a Poisson process[93] and allow for noise in the normalization signal[33] (Box 1). This model variant is known as the stochastic normalization model. As shown in ref. 33, including noise in the normalization signal has almost no effect on the mean responses of the model but alters response variability in a number of ways. First, because the firing rate now varies across repeated presentations of the same stimulus, spike generation results from a doubly stochastic process, yielding super-Poisson variability. Second, because the normalization signal provided by the inhibitory channel rescales the output of the excitatory channel, the additive noise has a multiplicative effect on firing rate, that is, it introduces gain fluctuations. Last, the strength of these gain fluctuations depends on the output of the inhibitory channel (Fig. 2b). As is evident from the expressions that govern the behaviour of the model[33], excitatory drive and normalization noise increase excess response variance, whereas inhibitory drive has the opposite effect (Box 1, equation (5)). For this reason, the stochastic normalization model predicts that response sub-additivity and variability quenching will often co-occur, as illustrated for some classical experimental manipulations in Fig. 2c. It is important to note that these predictions have not yet been tested in great quantitative detail. Indeed, it is probable that additional model complexity will be required to capture the exact relationship between response sub-additivity and variability quenching. Note for example that, in this model (Fig. 2c), Fano factor initially increases with stimulus contrast. The available evidence suggests that this occurs for some cortical cells but that, in most cases, Fano factor decreases monotonically with stimulus contrast[32]. The model also predicts that Fano factor initially increases with increasing stimulus size (Fig. 2c): based on the limited evidence that is available, this appears true for most, but not all V1 cells[78] (see also ref. 39). It is possible that simple variation in parameter values can account for this cross-neural diversity[32]. However, it is also possible that additional model components will be required to fully capture these diverse empirical behaviours.

An alternative version of the stochastic normalization model replaces the Poisson point process with a Gaussian noise source in the excitatory channel[32]. Stimulus–response relationships are governed by different quantitative expressions, but they qualitatively behave in a very similar manner to those in the stochastic normalization model described in ref. 33. Most importantly, this variant also predicts a general quenching effect of normalization on neural response variability[32]. With the additional flexibility afforded by the stochastic excitatory channel, this model can capture empirical deviations from the stochastic normalization model described above. Another important feature of this model version is that it can be inverted to estimate the single-trial strength of the normalization signal (which is a statistical construct that cannot be measured empirically), from the measured neural activity. Using this method, it has been shown that even when the stimulus is constant, normalization strength fluctuates substantially across trials, and that the variability of V1 responses is more strongly quenched during trials in which normalization is stronger[32].

In summary, from the vantage point of this descriptive model of V1 activity, many of the classical phenomena of sub-additivity appear to result from the same operation (divisive normalization) and neural responses appear to contain two layers of variability: variability of spiking and variability of rate. The latter (variability in firing rate) is quenched by the suppressive signal.

Computational and representational objectives that shape V1 activity

The operations implemented by sensory systems are shaped by evolution, development and learning and are adapted to the tasks that the organism must perform in its natural environment. This suggests that key features of sensory systems might be understood by studying artificial systems designed either to optimally perform such tasks[101,102] or to realize computational goals essential to these tasks[103,104]. Several theoretical coding principles provide such normative insight into response sub-additivity and variability quenching in V1.

The most successful theoretical proposal concerning the goal of early sensory processing is the efficient coding hypothesis[103,104]. Applied to V1, this hypothesis states that the goal of V1 activity is to represent natural inputs with less statistical redundancy than is present in those inputs. This notion enjoys strong empirical support. When a set of linear filters is optimized such that filter responses to a generic ensemble of natural stimuli are both as informative and as statistically independent as possible, the resulting filters resemble the visual receptive field structure of V1 simple cells[83,84], particularly if the stimuli are natural movies rather than static natural images[105].

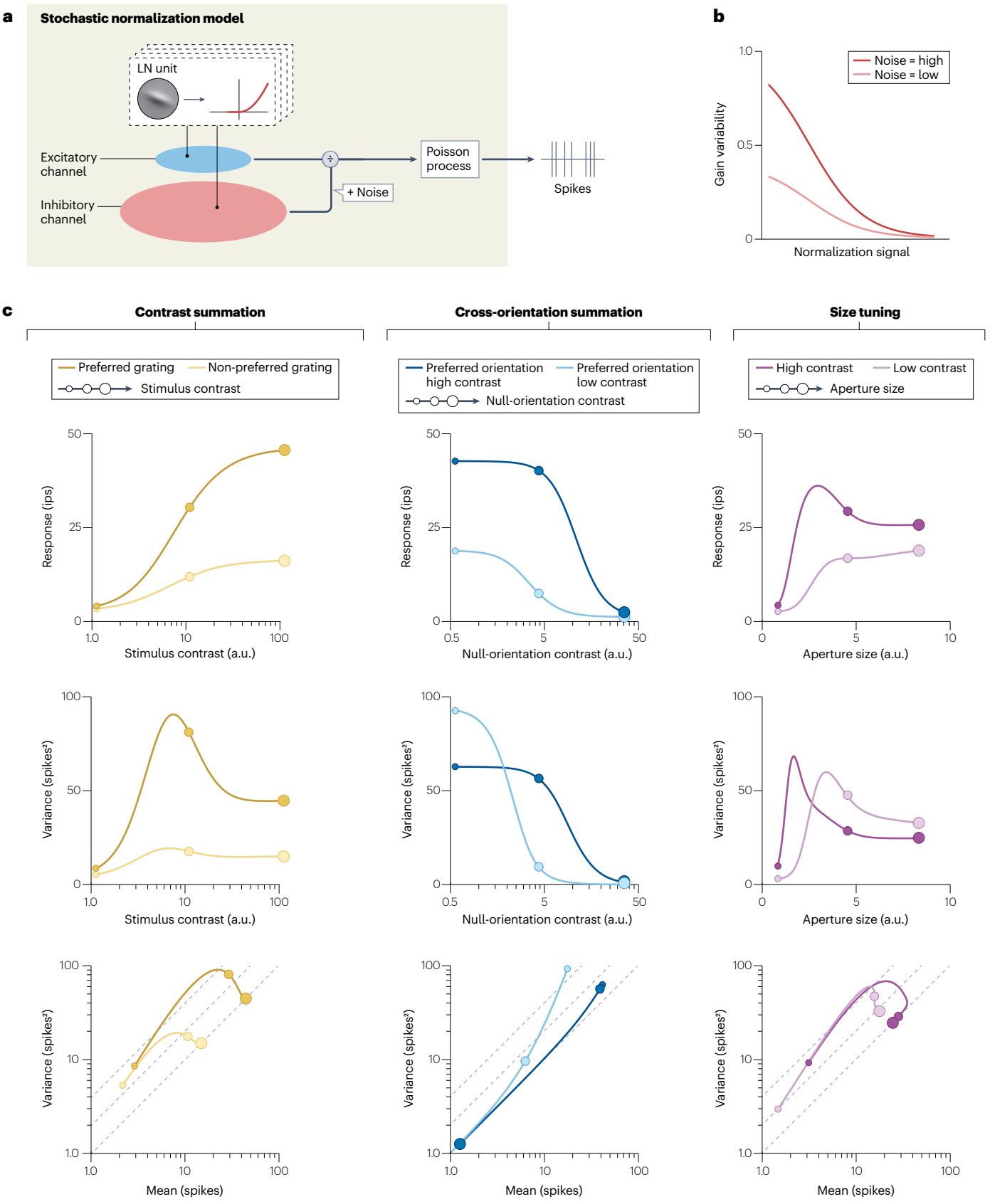
a
Stochastic normalization model
LN unit
Excitatory channel
Inhibitory channel
÷
+ Noise
Poisson process
Spikes
b
Noise = high
Noise = low
Gain variability
1.0
0.5
0
Normalization signal
c
Contrast summation
Cross-orientation summation
Size tuning
Preferred grating
Non-preferred grating
Stimulus contrast
Preferred orientation high contrast
Preferred orientation low contrast
Null-orientation contrast
High contrast
Low contrast
Aperture size
Response (ips)
Stimulus contrast (a.u.)
Null-orientation contrast (a.u.)
Aperture size (a.u.)
Variance (spikes²)
Mean (spikes)

Fig. 2 | Response sub-additivity and variability quenching under a stochastic normalization model. a, Model schematic. Stimuli are first processed by a bank of linear–nonlinear (LN) units, whose responses are pooled to form a narrowly tuned excitatory channel and a broadly tuned inhibitory channel that modulates response gain through a divisive operation[3,9,86,90,93,185]. The normalization signal provided by the inhibitory channel is subject to stimulus-independent, additive noise and spikes are generated by a Poisson process[33]. **b**, Under the stochastic normalization model, gain variability depends on the normalization signal and on the level of normalization noise. This plot is a schematic representation of this relationship, as demonstrated in ref. 33 and shown in equation (5) of Box 1. **c**, In this model, response sub-additivity and variability quenching often co-occur. The top row shows simulated mean responses of the stochastic normalization model to some classical experimental manipulations. Details of the simulation are described in Supplementary Information. The predictions of the model provide a good quantitative account for the behaviour of V1 cells[3,5,90]. The middle row shows the associated predicted response variance, obtained from equation (5) in Box 1, taken from ref. 33. The bottom row shows the resulting variance-to-mean relationships, which directly illustrate the joint occurrence of response sub-additivity and variability quenching in the stochastic normalization model. a.u., arbitrary unit; ips, impulses per second.

However, a simple linear transformation is insufficient to produce fully independent responses to natural images. Removing the remaining statistical dependencies requires additional nonlinear response transformations. In particular, dividing the response of each filter by the weighted sum of the rectified responses of neighbouring filters increases response independence[86]. When optimized for natural image statistics, the resulting divisive normalization model exhibits response sub-additivity that is reminiscent of that seen in cortical cells. For example, simulations of classical contrast summation, cross-orientation summation and size tuning experiments all yield model responses that qualitatively match the behaviour of V1 cells[86]. This framework can also account for the intricate cortical suppression phenomena elicited by natural images, namely, the observation that the strength of surround suppression varies widely for different images and is predicted by the statistical similarity between the image features presented within the receptive field and those presented outside the receptive field[4]. Phenomena of sub-additivity can, thus, be understood as a direct consequence of the attempt of the visual system to efficiently encode the statistical structure of natural images. In summary, this line of work demonstrates that the nonlinear response sub-additivity of sensory neurons captured by divisive normalization-based descriptive models (described in 'Descriptive accounts of response sub-additivity and variability quenching in V1' section) is well predicted by the normative principle of efficient coding. In other words, divisive normalization is "not an accident of biological implementation, but has an important functional role"[86].

Although redundancy reduction leads to normalization, which can account for response sub-additivity, it does not provide a normative justification for variability, or its quenching, per se. In fact, neural variability seems at odds with the normative goal of efficient coding because it limits coding capacity[106]. However, sensory systems seek to do more than just represent sensory input. Ultimately, they must construct perceptual interpretations of the environment that facilitate behavioural performance. The most relevant aspects of the environment (such as the presence of potential prey or a potential predator) typically have a complex and ambiguous relationship with raw sensory input and, thus, need to be inferred. Inevitably, these inferences have varying degrees of certainty. To achieve optimal behavioural outcomes, the uncertainty of perceptual inferences needs to be taken into account[107–109]. How neural circuits do so is debated and an important topic of modern research[21,27,33,39,110–117].

A prominent hypothesis suggests that the structure of neural response variability in the sensory cortex may facilitate the assessment of perceptual uncertainty by downstream circuits. Theorists have proposed several variants of this idea[114,118]. In particular, the neural sampling hypothesis proposes that neural responses in the sensory cortex represent samples from a probabilistic model of the environment[21,118,119]. It follows that neural response variability reflects uncertainty about the inferred stimulus feature. Consistent with this idea, factors that improve the quality of perceptual orientation estimates, such as image contrast and aperture size[120], often also quench the response variability of V1 neurons[21,33,39].

From a computational perspective, if sensory systems must take uncertainty into account, the optimal way to do so is to learn the causes of sensory inputs by analysing their statistical regularities and forming a so-called generative model that reproduces those statistics. Probabilistic inference involves inverting this generative model to correctly map an observed input onto a probability distribution of the causes of that input (the so-called posterior distribution)[75].

This foundational idea has recently been adapted to provide a unified account for the phenomenologies of response sub-additivity and response variability in V1 (refs. 21,39). The theory, which combines critical elements of efficient coding and neural sampling, proposes that V1 activity represents approximate probabilistic inferences based on a generative model of local image structure[21] (Fig. 3a,b). The theory further postulates that images are generated by combining local features with a global modulator that represents luminance or contrast (Fig. 3a). The theory assumes that the computational goal of V1 is to represent local image features by 'undoing' the effect of nuisance variables (such as the global modulator), a computation termed marginalization. Marginalization has the effect of removing redundancies that are present in the raw visual inputs (Fig. 3c). The theory additionally assumes that V1 activity represents samples from the inferred posterior probability distribution of the feature coefficients, that is, it is a neural sampling-based representation (Fig. 3d). In this way, the average neural response (the sample mean) represents the mean of the posterior distribution (the estimate of the feature coefficients that is expected to have minimal squared-error), whereas neural variability (sample variance) represents posterior variance (the uncertainty about this estimate).

When this model is optimized for natural image statistics, response sub-additivity and variability quenching often co-occur[21,39,121]. The intuitive explanation for this connection is that marginalization results in more certain inferences about local image features and is achieved via divisive normalization (Fig. 3b). Thus, the average neural response (representing the posterior mean) in this model inherits — at least qualitatively — the sub-additive effects predicted by the divisive normalization model of efficient coding discussed above[21,86]. For example, the responses of the linear filters representing the receptive field of a neuron and those representing its surround are typically informative about the global modulator. Homogeneous images that extend beyond the receptive field elicit similar responses in all of these filters, suggesting that the value of the global modulator is large and, therefore, evoking strong normalization from the surround[4,122]. Importantly, variability quenching arises as a consequence of the same computation:

Box 1

The stochastic normalization model

In the stochastic normalization model (Fig. 2), inhibitory drive suppresses response strength and quenches response variability[33]. Consider the simplest instantiation of the normalization framework[98]:

$$\lambda(S) = \frac{E(S)}{\beta + I(S)} \quad (1)$$

where λ is firing rate, S is an image, E is the excitatory drive obtained by measuring the contrast energy of a stimulus at a specific location in visual space and within a narrow range of orientations and spatial frequencies, I is the inhibitory drive obtained by measuring energy across a more extended region and across a broad range of orientations and spatial frequencies, and β is a stimulus-independent constant[90]. Equation (1) specifies a deterministic relationship between stimulus and firing rate, a statistic that is not directly observable but can be inferred from trial-averaged measurements. The simplest way to obtain a full generative model of spiking activity from this model is to include a Poisson point process[93,95]. Together, these model components suffice to express the probability of every possible spike count for arbitrary visual stimuli:

$$p(N|S, \Delta t) = \frac{(\lambda(S)\Delta t)^N}{N!} e^{-\lambda(S)\Delta t} \quad (2)$$

where N is the spike count and Δt is the duration of the counting window. Under this model, response variance equals the response mean:

$$\mathrm{Var}[N|S, \Delta t] = \mathrm{Mean}[N|S, \Delta t] = \lambda(S)\,\Delta t \quad (3)$$

The stochastic normalization model[33] assumes that the normalization signal is not deterministic but is subject to additive Gaussian noise with zero mean and variance σ_N^2. On a single trial, spikes are generated from a Poisson process with a firing rate given by the following equation:

$$\lambda_i(S) = \frac{E(S)}{\beta + I(S) + \epsilon_i} \quad (4)$$

where subscript i is a trial index and ϵ_i is the Gaussian noise (with zero mean and σ_N standard deviation). Because the denominator in equation (4) rescales the excitatory drive, this additive noise has a multiplicative effect on firing rate, that is, it introduces gain fluctuations[33]. The strength of these gain fluctuations depends on the inhibitory drive and is well approximated by the following equation:

$$\sigma_G = \frac{\sigma_N}{\beta + I(S)} \quad (5)$$

where σ_G expresses the standard deviation of the gain[33]. The inhibitory drive reduces gain fluctuations and, hence, reduces response variability. The stochastic normalization model describes a doubly stochastic process. It follows from the law of total variance that spike count variance is composed of the sum of the expected Poisson variance (equation (3)) and a term that represents the contribution of rate variability[187]. This term is the product of the variance of the gain signal and the squared mean response[28] and is given by the following equation:

$$\mathrm{Var}[\lambda(S)\,\Delta t] = \frac{(\sigma_N E(S))^2}{(\beta + I(S))^4}\,\Delta t^2 \quad (6)$$

observing the image content in the surround lowers the estimation uncertainty about local image structure inside the receptive field and, thus, results in less response variability[39]. Conversely, when the image is confined to the receptive field (Fig. 3d), or when the surround image is part of a different object (termed heterogeneous centre–surround configuration[4]), the responses of the linear filters representing the receptive field and surround are less redundant (resulting in weaker normalization) and uncertainty about local image features is higher (resulting in higher response variability).

In summary, studies of the computational and representational objectives underlying V1 activity offer a parsimonious explanation for the co-occurrence of response sub-additivity and variability quenching: divisive normalization in V1 serves to compute probabilistic inferences about visual inputs, relating sub-additive phenomena that maximize coding efficiency and quenching phenomena that express uncertainty about inferred image features.

Circuit mechanisms that govern V1 activity

V1 activity is shaped by retinal, thalamic and cortical circuit mechanisms. How do cortical response sub-additivity and variability quenching arise mechanistically from the interplay of these distinct anatomical components? The descriptive and normative modelling approaches discussed thus far offer little insight into this. These models are formulated in terms of normative principles and phenomenological operations — including linear filtering, divisive normalization and noisy spike generation — whose biophysical and anatomical substrates are not specified. To address this question, we must turn to neural circuit models.

Response sub-additivity and variability quenching need not be produced by circuits within V1; they could instead arise from the inputs to V1. There is evidence that this is the case for the sub-additivity that is induced by stimuli presented in the receptive field centre[54–56,123–126] (but see ref. 67). Furthermore, variability in the firing rate of lateral geniculate nucleus cells decreases with stimulus contrast[13,127]. Thus, some forms of response sub-additivity and variability quenching in the cortex may have, at least in part, a feedforward origin.

In other cases, cortical circuitry might be the source of response sub-additivity and variability quenching. For example, there is evidence that cortical circuitry has a major role in the sub-additivity of surround suppression[64,128] and the quenching of variability elicited by stimuli presented in the surround[40,78]. How cortical circuits produce response sub-additivity and variability quenching is an active topic

of contemporary research. The proposed models that have received the most attention (for example, see refs. 129–131) share two common features. First, excitatory cortical circuitry is structured: excitatory neurons with shared selectivity are more strongly coupled whereas those with distinct preferences exhibit weak coupling. Second, inhibitory connections are strong enough to stabilize the network despite the excitatory connectivity being strong enough to potentially cause instability. These features are apparent in the cortex[45,132–136], although there is as yet limited direct evidence for their role in response sub-additivity or variability quenching (but see ref. 132).

Cortical neurons receive inputs from many cells. A neuron receiving a large number of excitatory inputs, without compensating inhibition, would have a large mean input and exhibit regular clock-like spiking, inconsistent with the variable firing observed in cortical cells[18,137]. Therefore, a number of additional mechanisms have been proposed to account for spiking variability. Two possible sources of variability are stochasticity in the cellular and synaptic mechanisms that drive firing[138–140] and input correlations that prevent the variability of individual inputs from being averaged out[141,142]. Spiking variability can also arise from network dynamics: if inhibition balances excitation sufficiently that the mean input to a cell is below or close to the threshold required to drive firing, the neuron will be driven to fire by input fluctuations – brief imbalances in excitation and inhibition – that occur at random times[143–146]. The 'balanced network' model[144] demonstrated that network dynamics can automatically yield such balancing in a broad parameter regime, without requiring the fine tuning of parameters. This model produced 'tight balance', which means that the excitatory and inhibitory inputs are much larger than the net input remaining after they cancel one another[147]. However, tightly balanced network models do not naturally produce the nonlinear input–output transformations that could give rise to response sub-additivity. They also do not generate super-Poisson variability or variability quenching. These limitations can be addressed by considering more loosely balanced models[147] and/or structured[129] or heterogeneous[148] connectivity, as we now discuss.

What additional sources of variability yield the super-Poisson variability that is characteristic of cortical neurons, and why is this additional variability quenched by the presentation of a masking or higher contrast stimulus? In one family of models, additional variability arises during spontaneous activity because the network is wandering among many states, each corresponding to one of the possible responses to many different stimuli, and neurons fire at different rates in different states[129,149]. A stimulus 'pins' the network to one state, quenching the variability. Some of these models depend on specific connectivity. For example, given stronger connections within and weaker connections between distinct clusters of excitatory units in an otherwise balanced network, in which the activity of one cluster inhibits the others, network activation can be largely restricted to one cluster at a time and wander between clusters over time[129]. A similar mechanism involves neurons being most strongly connected to neurons with similar response properties, thus forming a continuum of clusters rather than discrete clusters. Consistent with this idea, spontaneous activity in V1 has been

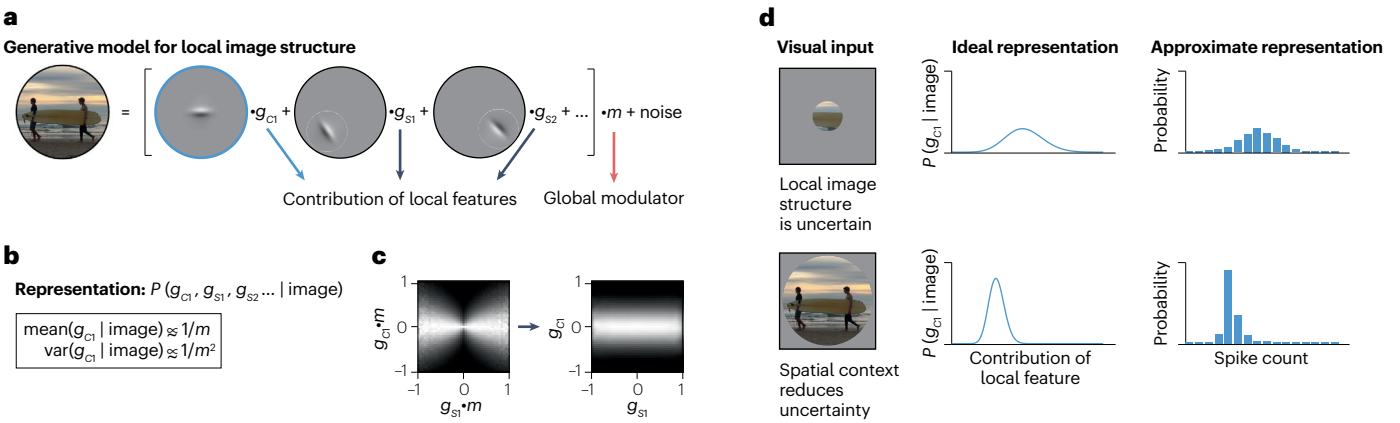

Fig. 3 | A normative account for the relationship between response sub-additivity and variability quenching in V1. a, The local structure of natural images can be described as a linear combination of a set of spatially localized image features (g_{C1}, g_{S1}, g_{S2} and so on) that is subject to global modulation (m) and noise (that is, as a Gaussian scale mixture model)[186]. **b**, Encoding image information by inferring the contribution of each local feature (that is, by computing the posterior distribution, top equation) naturally results in divisive response suppression and response variability quenching for higher values of the global modulator[39] (for example, for images with higher contrast levels). This is shown mathematically in the two equations inside the box: both the mean and the variance of the posterior distribution approximately scale with the inverse of the global modulator. **c**, Joint histograms of the simulated responses of a nearby pair of local image filters before (left) and after (right) normalizing by the global modulator. Responses were random samples from a simulation performed using the Gaussian scale mixture model described in ref. 186. Details of the simulation are described in Supplementary Information. The greyscale value in each bin is proportional to the number of cases in that bin, re-scaled for each column separately so that black corresponds to zero and white corresponds to the maximum number of cases in that column. The 'bow tie' shape of the joint histogram on the left indicates statistical redundancy between the simulated responses of the two filters because the variance of the responses of one filter (spread in the vertical direction) depends on the magnitude of the response of the other filter (horizontal position). Normalization reduces response redundancy among these filters, as indicated by the absence of the bow tie shape in the right panel, wherein the variance of the responses of one filter does not depend on the magnitude of the response of the other filter. This is also shown in ref. 86. **d**, Schematic representation of sampling-based inference in the Gaussian scale mixture model[39]. If the inferred contribution of local features is represented probabilistically, informative image content in the surround will lower the peak and narrow the width of the posterior belief in the contribution of a local feature positioned at the centre of the image (centre panels). If neural responses represent samples from the posterior distribution, this will manifest as response suppression and variability quenching (right panels).

shown to wander through states resembling stimulus responses more often than expected by chance, in both experiments involving laboratory stimuli[26,150,151] and those involving natural stimuli[21,27]. In these models, a stimulus that pins the wandering reduces the variability of all neurons, including those not driven by the stimulus[152,153], consistent with the finding that presentation of a stimulus quenches variability both in neurons that are driven by the stimulus and those that are not[22]. A related proposal is that the variability of a network is generated by chaotic dynamics of spontaneous activity, as occurs in tightly balanced networks with sufficient variability in their weights[148]. According to this model, a stimulus can then suppress variability by suppressing the chaos[154].

It is important to note that, in these models, there is no connection between the mechanisms that alter variability and sub-additivity of responses. If such a model shows response sub-additivity, this must arise owing to mechanisms distinct from the stimulus-induced pinning of network state that quenches variability. However, in these models, changes in firing rates (which need not involve sub-additivity) are naturally coupled to changes in variability: for example, a decrease in stimulus strength decreases firing rates and increases variability at the same time[153]. A similar increase in variability accompanying a decrease in firing rates has been reported in a recent preprint in some cases of surround suppression[78], and it has been suggested[78,153] that, in these cases, surround suppression arises primarily from the suppression of feedforward inputs (consistent with experimental evidence for feedforward contributions to surround suppression in macaque V1 (refs. 64,128) and, as reported in two recent preprints, in mouse V1 (refs. 155,156)). However, in most cases, a decrease in variability accompanies surround suppression[39,78] (consistent with experimental evidence for cortical contributions to surround suppression[64,128]). To explain this phenomenon, other mechanisms relying on the intrinsic dynamics of V1 are necessary.

An alternative model of V1 dynamics that may account for both variability quenching and response sub-additivity posits that the network randomly fluctuates about a single steady state for a given fixed external input (including the input driving spontaneous activity), but the amplitude of the fluctuations decreases as the external input strength increases, quenching variability. In this model, the fluctuations arise owing to external input noise and are amplified by an excitatory network that is stabilized by inhibitory cells[131]. It turns out that the strength of this inhibitory stabilization increases with increasing external input, driving both response sub-additivity and variability quenching. A key feature of this stabilized supralinear network (SSN) model[81,130,131] is that neuronal input–output functions are supralinear (Fig. 4a). This means that neuronal gain – the change in output per change in input (that is, the slope of the input–output function) – increases with increasing neuronal activation. The result is that the effective connection strengths – the change in postsynaptic firing rate per change in presynaptic firing rate – increase with increasing strength of the external input of the network[130] (Box 2). This increase has a central role in both response sub-additivity and variability quenching in the SSN.

For very weak external input – below or around the level that drives spontaneous activity – effective synaptic strengths are weak. As a result, inputs from monosynaptic pathways (the external input) are much stronger than the inputs from disynaptic and polysynaptic pathways that they evoke (via the recurrent connections between cortical cells). Thus, responses largely follow the supralinear input–output function of decoupled cells and, hence, sum supralinearly. With increasing input strength, the relative contribution of recurrently driven input increases, eventually exceeding a point at which the excitatory subnetwork alone would become unstable. At this point, the network enters an inhibition-stabilized regime[45,130,132,157]. This transition probably occurs for weaker external input than that received during spontaneous activity because, at least in mouse, the cortex is already in an inhibition-stabilized state during spontaneous activity[45]. Stabilization occurs through 'loose balancing'[130,147] (Fig. 4b), which means that the recurrent input largely cancels the external input, so that the net input grows sublinearly as a function of the external input and that the net input is comparable in size to the factors that cancel (that is, the balance is 'loose'). This loose balance is sufficient to yield irregular spiking[158] yet allows nonlinear behaviours such as response sub-additivity that are absent when balance is tight. In particular, when a second stimulus is added to a first, most of the extra feedforward input is cancelled. This loosely balanced regime in turn divides into two parameter sub-regimes, in only one of which contrast saturation occurs[130]. However, for most parameters in either sub-regime, when two different stimuli are added, response summation is sublinear, regardless of the location of the second stimulus (receptive field or surround)[81,130].

This mechanism also creates and quenches super-Poisson variability, as illustrated for a simple model in which there is one excitatory and one inhibitory population[131] (Fig. 4). With both strong amplifying excitatory and strong stabilizing inhibitory connections, the network shows 'balanced amplification'[159]: small input imbalances favouring excitation (or inhibition) strongly increase (or decrease) the drive to both excitatory and inhibitory cells. This can be mathematically summarized by formulating the dynamics in terms of the strengths of two patterns of activity: a difference pattern and a sum pattern, in which the excitatory and inhibitory population activities have opposite signs or the same signs, respectively (Fig. 4a). Any actual pattern of excitatory and inhibitory population activities can be expressed as a linear combination of these two patterns. Each pattern effectively inhibits (or damps) its own activity, with weights λ_d and λ_s. The difference pattern excites the sum pattern with a weight w_{FF}, but there is no connection in the opposite direction (that is, there is a feedforward connection between the two patterns[159]).

In this model, as external input h increases from 0, the variability, as measured by voltage standard deviation, first increases to a peak before thereafter being suppressed (Fig. 4c). The peak occurs around the transition between the external input-dominated and recurrent input-dominated regimes, in which the recurrent input starts to balance the external input (Fig. 4b). As h increases from zero, effective connection weights rapidly increase and w_{FF}, the feedforward drive from difference patterns to sum patterns, rapidly grows (Fig. 4d). Thus, variability is increased by increasingly strong balanced amplification: small excitatory or inhibitory differences in the external noise will drive large joint fluctuations of excitatory and inhibitory activity. The decrease in the self-inhibition λ_s of the sum pattern also contributes, decreasing the damping of fluctuations of the sum pattern. Beyond the regime transition, the growth of w_{FF} greatly slows, whereas λ_s and, later, λ_d grow. This represents increasingly strong inhibitory stabilization, which damps fluctuations and so quenches variability. The net result (Fig. 4e) is that the fluctuations, initially driven by the input noise ($h = 0$), are greatly amplified in the sum direction, producing the peak in voltage variability ($h = 2$), before being quenched by the increasingly strong inhibitory damping ($h = 15$).

A signature of the SSN is a non-monotonic dependence of variability on stimulus strength (Fig. 4c), in agreement with the stochastic

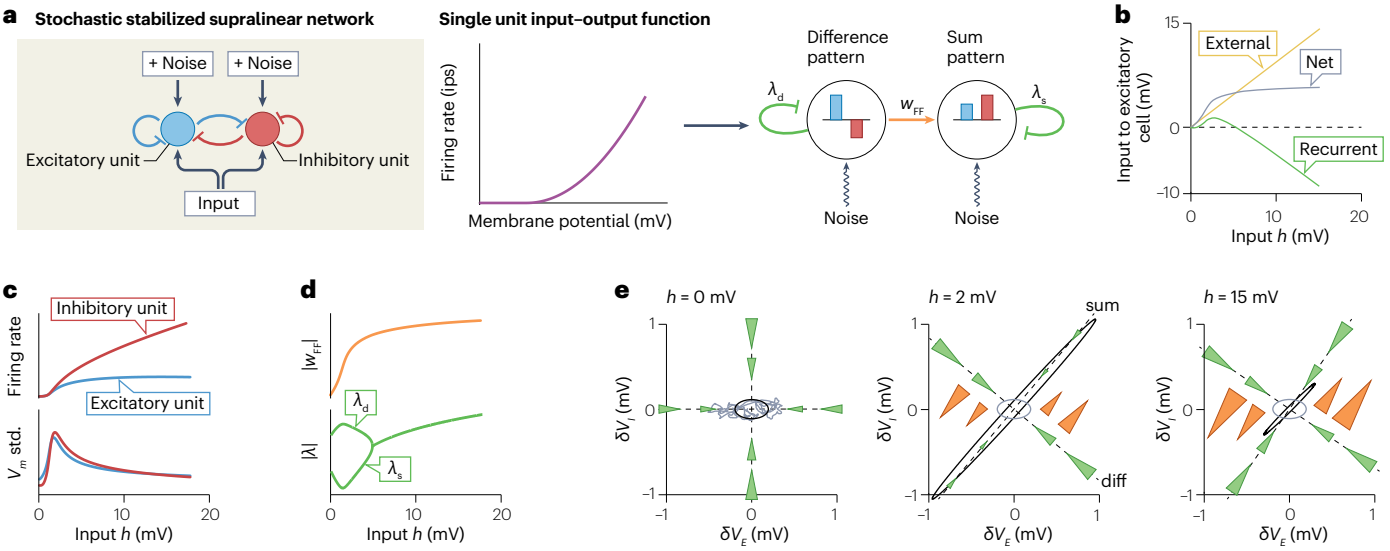

Fig. 4 | Response sub-additivity and variability quenching in the stochastic stabilized supralinear network model. a, The left panel shows the model network studied in ref. 131: two recurrently connected units, representing the activity of populations of excitatory and inhibitory neurons, each receive private input noise and a common mean input. As shown in the middle panel, a supralinear (threshold–quadratic) neural input–output function determines the instantaneous firing rate of these units as a function of membrane potential. The dynamics of this network can be expressed in terms of two activity patterns: a difference pattern, in which excitatory and inhibitory activities (blue and red bars, respectively) have opposite signs and a sum pattern, in which they have the same sign. The patterns inhibit themselves with weights λ_d (diff) and λ_s (sum), respectively, whereas the difference pattern excites the sum pattern with weight w_{FF}. **b**, Change in external input, recurrent input and net input (external input plus recurrent input) to the excitatory unit with increasing external input h. When the external input h is weak, it dominates recurrent input (green). However, when the external input is stronger, the SSN dynamics leads the recurrent input to largely cancel the external input. Thus, net input grows sub-linearly as a function of external input. Over the dynamic range of this model, balance is 'loose' (that is, the net input is similar in size to the other two, cancelling, inputs) but would become 'tight' (with a net input much smaller than the others) for very strong (probably non-physiological) input. Loose, but not tight, balance can yield response sub-additivity. **c**, Mean firing rate (top) and voltage standard deviation (bottom) for the excitatory and inhibitory units, as a function of mean input strength h. **d**, Dependence of the weights λ_d, λ_s and w_{FF} on input strength. **e**, How the dependence of these weights on input strength causes variability to be first amplified and then quenched with increasing input strength. We study the voltage fluctuations δV_E and δV_I of the excitatory and inhibitory units about their mean voltages. The black ellipses represent 1-standard-deviation contours of the fluctuations. When $h = 0$, the excitatory and inhibitory units are effectively uncoupled, so the ellipse represents only the external input noise filtered by the time constant of each isolated unit (repeated in other panels in grey, for comparison). The grey lines in the $h = 0$ panel are representative trajectories of δV_E and δV_I in simulations. The green triangular arrows and dashed lines between them illustrate the effects of λ_d and λ_s on the trajectories of the system. These drive fluctuations toward the origin, along the direction indicated by the arrows, and with a strength indicated by the area of the arrow. These directions, indicated by the dashed lines, correspond to the activities of the inhibitory unit and the excitatory unit when $h = 0$ (when the units are effectively uncoupled), and to the activities of the difference pattern and the sum pattern when $h > 0$. The orange triangles similarly represent the effects of w_{FF}, by which positive (or negative) amplitudes of the activity of the difference pattern drive fluctuations in the positive (or negative) direction of the activity of the sum pattern. The rapid increase in w_{FF} with the initial increase of h from 0 (part **d**) causes variability to be strongly amplified in the direction of the sum pattern (seen for $h = 2$), creating super-Poisson variability. The subsequent growths of λ_d and λ_s (part **d**) quench this variability (seen for $h = 15$). Parts **a**, **c**, **d** and **e** are adapted from ref. 131.

normalization model (Fig. 2c). This has been demonstrated in the SSN for changes in contrast[131], but data are not yet available from in vivo experiments in which sufficiently fine manipulations of contrast have been used to test this prediction comprehensively. In those in vivo experiments that have been performed, a decrease in variability with increasing contrast was robustly seen for larger contrasts[21,22,32,160]. However, the very low contrasts, for which an increase in variability would be expected, have not been carefully studied (although an increase in variability with increasing stimulus contrast at moderately lower contrasts was seen in a minority of cells[32]). Other experiments have, however, seen a clear non-monotonic change in variability with increasing spatial size of a stimulus, at least in some layers of V1: here, variability increased for the smallest sizes then decreased for larger sizes[39,78]. The SSN mechanism reproduces surround suppression of firing rates[81], but whether it can account for the dependence of variability on stimulus size has not been explicitly studied. Nevertheless, we expect to see the same non-monotonic dependence of variability on stimulus size as for stimulus contrast, as the same mechanisms should apply in both cases.

Experimental data suggest that multiple mechanisms of surround suppression are engaged in different cortical layers and by different spatio-temporal stimulus configurations[57,64,78,128]. Modulation of variability might also vary accordingly[39,40,78]. For example, it is possible that (as described above) surround suppression of feedforward inputs to the cortex might replicate the effects of a decrease in contrast and, thus, increase variability, whereas surround suppression that derives from recurrent cortical mechanisms may represent increasingly strong inhibitory stabilization that decreases variability. Similar considerations apply to masking suppression, which has been shown to be composed of a weaker cortical component[67] that can be described by the

Box 2

Effective connection strengths in the SSN model

Effective connectivity scales with input strength in the stabilized supralinear network (SSN)[130] (see also refs. 81,131). In the SSN (Fig. 4), at steady state, the contribution of a particular presynaptic cell to the input of a postsynaptic cell, I, is simply given by the firing rate of the presynaptic cell, r_{pre}, scaled by the strength of the connection between the two cells, W:

$$I = W\ r_{pre} + \ldots \quad (7)$$

where ... denotes terms that are independent of the presynaptic neuron, such as recurrent inputs from other neurons in the network and external, feedforward inputs from upstream areas. The firing rate of the postsynaptic neuron is then determined by the neuronal input–output function:

$$r_{post} = f(I) \quad (8)$$

Combining equations (7) and (8) yields a self-consistent equation that expresses the (steady-state) relationship between the firing rates of the presynaptic and postsynaptic neuron:

$$r_{post} = f(W\ r_{pre} + \ldots) \quad (9)$$

This means that for a sufficiently small deviation in the activity of a presynaptic neuron (for example, owing to a change in its external drive or random fluctuations), δr_{pre}, the change in the response of the postsynaptic neuron, δr_{post}, is given by the following equation:

$$\delta r_{post} = f'(I)\ W\ \delta r_{pre} \quad (10)$$

where $f'(I)$ is the 'neural gain' — the slope of the input–output function at the steady-state input of the postsynaptic neuron. Thus, the effective connection strength between the two neurons is the actual connection strength scaled by the neuronal gain:

$$W_{eff} = \frac{\delta r_{post}}{\delta r_{pre}} = f'(I)\ W \quad (11)$$

The supralinearity of f means that the rate $f(I)$ and the gain $f'(I)$ both grow with the input I, and thus so does the effective connection strength W_{eff}.

SSN (ref. 81) and a stronger component that arises owing to masking effects on the feedforward inputs to cortex[54–56,123–126].

In summary, in mechanistic models of V1 activity, response sub-additivity and variability quenching can both arise via network effects that yield increasingly strong inhibitory stabilization as stimulus strength increases. Variability quenching can also arise through stimulus pinning of wandering network activity, without any necessary connection to response sub-additivity. In all of these models, suppression of feedforward input will suppress responses and is expected to increase variability. Particular forms of sub-additivity or suppression may involve different sets of these mechanisms in different locations. Thus, mechanistic models suggest that sub-additivity will often, but not always, co-occur with variability suppression.

Conclusions

We have seen that response sub-additivity in V1 often co-occurs with variability quenching. Response sub-additivity arises from nonlinear input transformations whereas response variability results from the accumulation and amplification of small amounts of noise as signals flow through neural circuits. It is, therefore, not obvious that both phenomena should have common origins. Yet, that is exactly what we propose. This proposal is motivated by recent model-based insights into the functional operations, computational objectives and circuit mechanisms that govern V1 activity. Although these modelling approaches address different aspects of cortical activity and rely on very different model architectures, they all predict that response sub-additivity and variability quenching will often co-occur. We do not wish to suggest that a single circuit mechanism underlies this relationship: indeed, different forms of response sub-additivity and variability quenching probably arise from distinct circuit mechanisms. Moreover, more work is needed to establish whether the discussed models are rich enough to account for the diversity of neural behaviours seen within the same experimental paradigm. That said, the converging insights naturally raise new questions. We end this Perspective by considering three questions that seem particularly important to us: can the modelling insights be unified, is the connection between response sub-additivity and variability quenching a canonical motif across cortex, and do specific model components map onto specific subtypes of neurons?

Descriptive, normative and mechanistic modelling approaches offer different levels of explanation, but they are not mutually exclusive enterprises. Progress at one level can spark progress at another level. For example, refining descriptive models to better capture the diverse effects of surround stimulation on response suppression[3,11] has provided critical guidance for normative models of V1 activity[39,86]. Likewise, descriptive accounts of variability quenching across the cortex[22] inspired progress in mechanistic models of spiking activity[129,131]. More direct examples of cross-level interactions are offered by recent attempts to combine different levels of explanation in a single model[82,121]. One study[121] has bridged normative and mechanistic levels by optimizing the connectivity of the SSN architecture for probabilistic inference, so that SSN response variability closely matched the variability produced by a sampling-based normative model for stimuli with a cross-orientation mask. The network optimized for this variability structure was precisely in the SSN loosely balanced regime described above that shows response sub-additivity and variability quenching. The study also showed that this SSN regime reproduced other cortical phenomena for which it was not directly optimized, including contrast-controlled oscillations (see also the findings of a recent preprint[161]) and stimulus-onset transients, each of which had a functionally well-defined role in network computations.

Another recent study has developed a model that bridges descriptive and a mechanistic levels. Specifically, the family of dynamic circuit models called oscillatory recurrent gated neural integrator circuits (ORGaNICs (ref. 82)) was explicitly designed to produce a steady state

exactly described by the equations of divisive normalization. Preliminary findings suggest that, similar to the SSN framework, recurrent inhibition in stochastic variants of ORGaNICs stabilizes the network when recurrent excitation would otherwise make it unstable, producing both sub-additivity and variability quenching (S. Martiniani and D. Heeger, personal communication). ORGaNICs relies on recurrent amplification through a multiplicative interaction between recurrent drive and recurrent gain that can be regarded as a phenomenological description of actual circuit mechanisms. One advantage of this model family is that its steady state and its variability and covariability can all be computed analytically, simplifying the study of large-scale and multi-area networks.

Looking forward, training deep neural networks whose connectivity resembles visual cortical circuitry to either perform visual tasks[162–165] or to predict responses of visual neurons[166–168] holds promise as a powerful approach to build bridges between descriptive, normative and mechanistic approaches. However, thus far, this approach has not yielded any insight into neural response variability or its quenching – this is an important open challenge for future research.

Response sub-additivity and variability quenching are not limited to V1. Suppression of neural responses to a preferred stimulus by the simultaneous presentation of a non-preferred stimulus has been documented for many sensory[169–172] and non-sensory brain areas[173,174]. Likewise, the quenching of response variability by stimulus onset is thought to be a general property of cortical neurons[22]. This raises the question of whether the connection between response sub-additivity and variability quenching is a canonical motif across the cortex. The insights provided by the V1 models that we have discussed suggest that this may be the case. Specifically, these models suggest that both phenomena result from neural mechanisms that implement divisive normalization. This operation is considered a canonical neural computation that is repeated in a modular fashion in many distinct brain systems through a variety of circuits and mechanisms[98]. Determining the generality of the co-occurrence of response sub-additivity and variability quenching may reveal a canonical aspect of neural activity and as such represents a crucial step for developing a principled understanding of cortical computation.

The models that we have discussed offer abstracted descriptions of neural stimulus–response transformations. As we have highlighted, such abstractions can provide valuable insight into brain function even if the model components cannot be mapped onto biophysical substrates. Nevertheless, establishing such mapping is a quintessential goal of systems neuroscience. The recent advent of circuit-dissection tools capable of distinguishing the functional role of specific sub-types of cortical neurons[41,68,175] brings this goal within experimental reach.

In this Perspective, we have focused on sub-additivity of firing rate and on response variability, both single-neuron response statistics. We have reviewed modelling frameworks that suggest unified descriptions and explanations for those phenomena, but that also help us to distinguish separate mechanisms underlying similar phenomena. A natural and important extension of this work is to additionally consider pairwise and population-level response statistics (such as pairwise noise correlations and the geometry of population activity). These statistics have been studied extensively in cortical areas[176–178], are influenced by similar factors as those that elicit response sub-additivity and variability quenching[38,51,179–183], and further constrain models of neural activity.

Published online: 19 February 2024

References

1. Born, R. T. & Tootell, R. B. Single-unit and 2-deoxyglucose studies of side inhibition in macaque striate cortex. *Proc. Natl Acad. Sci. USA* **88**, 7071–7075 (1991).
2. Sceniak, M. P., Hawken, M. J. & Shapley, R. Visual spatial characterization of macaque v1 neurons. *J. Neurophysiol.* **85**, 1873–1887 (2001).
3. Cavanaugh, J. R., Bair, W. & Movshon, J. A. Nature and interaction of signals from the receptive field center and surround in macaque v1 neurons. *J. Neurophysiol.* **88**, 2530–2546 (2002).
4. Coen-Cagli, R., Kohn, A. & Schwartz, O. Flexible gating of contextual influences in natural vision. *Nat. Neurosci.* **18**, 1648 (2015).
5. Albrecht, D. G. & Hamilton, D. B. Striate cortex of monkey and cat: contrast response function. *J. Neurophysiol.* **48**, 217–237 (1982).
6. Morrone, M. C., Burr, D. C. & Maffei, L. Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proc. R. Soc. Lond. B Biol. Sci.* **216**, 335–354 (1982).
7. Bonds, A. B. Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis. Neurosci.* **2**, 41–55 (1989).
8. DeAngelis, G. C., Robson, J. G., Ohzawa, I. & Freeman, R. D. Organization of suppression in receptive fields of neurons in cat visual cortex. *J. Neurophysiol.* **68**, 144–163 (1992).
9. Carandini, M., Heeger, D. J. & Movshon, J. A. Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* **17**, 8621–8644 (1997).
10. Tolhurst, D. & Heeger, D. Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. *Vis. Neurosci.* **14**, 293–309 (1997).
11. Cavanaugh, J. R., Bair, W. & Movshon, J. A. Selectivity and spatial distribution of signals from the receptive field surround in macaque v1 neurons. *J. Neurophysiol.* **88**, 2547–2556 (2002).
12. Azouz, R. & Gray, C. M. Cellular mechanisms contributing to response variability of cortical neurons in vivo. *J. Neurosci.* **19**, 2209–2223 (1999).
13. Sadagopan, S. & Ferster, D. Feedforward origins of response variability underlying contrast invariant orientation tuning in cat visual cortex. *Neuron* **74**, 911–923 (2012).
14. Andoni, S., Tan, A. & Priebe, N. J. in *The New Visual Neurosciences* 367–380 (MIT Press, 2013).
15. Tomko, G. J. & Crapper, D. R. Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Res.* **79**, 405–418 (1974).
16. Heggelund, P. & Albus, K. Response variability and orientation discrimination of single cells in striate cortex of cat. *Exp. Brain Res.* **32**, 197–211 (1978).
17. Vogels, R., Spileers, W. & Orban, G. A. The response variability of striate cortical neurons in the behaving monkey. *Exp. Brain Res.* **77**, 432–436 (1989).
18. Shadlen, M. & Newsome, W. T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).
19. Talluri, B. C. et al. Activity in primate visual cortex is minimally driven by spontaneous movements. *Nat. Neurosci.* **26**, 1953–1959 (2023).
20. Tolhurst, D. J., Movshon, J. A. & Thompson, I. D. The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp. Brain Res.* **41**, 414–419 (1981).
21. Orbán, G., Berkes, P., Fiser, J. & Lengyel, M. Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron* **92**, 530–543 (2016).
22. Churchland, M. M. et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369–378 (2010).
23. Arandia-Romero, I., Tanabe, S., Drugowitsch, J., Kohn, A. & Moreno-Bote, R. Multiplicative and additive modulation of neuronal tuning with population activity affects encoded information. *Neuron* **89**, 1305–1316 (2016).
24. Rosenbaum, R., Smith, M., Kohn, A., Rubin, J. & Doiron, B. The spatial structure of correlated neuronal variability. *Nat. Neurosci.* **20**, 107–114 (2017).
25. Davis, Z. W., Muller, L., Martinez-Trujillo, J., Sejnowski, T. & Reynolds, J. H. Spontaneous travelling cortical waves gate perception in behaving primates. *Nature* **587**, 432–436 (2020).
26. Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A. & Arieli, A. Spontaneously emerging cortical representations of visual attributes. *Nature* **425**, 954–956 (2003).
27. Berkes, P., Orbán, G., Lengyel, M. & Fiser, J. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* **331**, 83–87 (2011).
28. Goris, R. L. T., Movshon, J. A. & Simoncelli, E. P. Partitioning neuronal variability. *Nat. Neurosci.* **17**, 858–865 (2014).
29. Ecker, A. S. et al. State dependence of noise correlations in macaque primary visual cortex. *Neuron* **82**, 235–248 (2014).
30. Rabinowitz, N. C., Goris, R. L., Cohen, M. & Simoncelli, E. P. Attention stabilizes the shared gain of V4 populations. *eLife* **4**, e08998 (2015).
31. Goris, R. L., Ziemba, C. M., Movshon, J. A. & Simoncelli, E. P. Slow gain fluctuations limit benefits of temporal integration in visual cortex. *J. Vis.* **18**, 8 (2018).
32. Coen-Cagli, R. & Solomon, S. S. Relating divisive normalization to neuronal response variability. *J. Neurosci.* **39**, 7344–7356 (2019).
33. Hénaff, O. J., Boundy-Singer, Z. M., Meding, K., Ziemba, C. M. & Goris, R. L. T. Representation of visual uncertainty through neural gain variability. *Nat. Commun.* **11**, 2513 (2020).
34. DiCarlo, J. J., Zoccolan, D. & Rust, N. C. How does the brain solve visual object recognition? *Neuron* **73**, 415–434 (2012).
35. Pitkow, X. & Angelaki, D. E. Inference in the brain: statistics flowing in redundant population codes. *Neuron* **94**, 943–953 (2017).
36. Moreno-Bote, R. et al. Information-limiting correlations. *Nat. Neurosci.* **17**, 1410 (2014).

37. Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E. & Pouget, A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* **74**, 30–39 (2012).
38. Snyder, A. C., Morais, M. J., Kohn, A. & Smith, M. A. Correlations in v1 are reduced by stimulation outside the receptive field. *J. Neurosci.* **34**, 11222–11227 (2014).
39. Festa, D., Aschner, A., Davila, A., Kohn, A. & Coen-Cagli, R. Neuronal variability reflects probabilistic inference tuned to natural image statistics. *Nat. Commun.* **12**, 3635 (2021).
40. Henry, C. A. & Kohn, A. Feature representation under crowding in macaque v1 and v4 neuronal populations. *Curr. Biol.* **32**, 5126–5137 (2022).
41. Nassi, J. J., Avery, M. C., Cetin, A. H., Roe, A. W. & Reynolds, J. H. Optogenetic activation of normalization in alert macaque visual cortex. *Neuron* **86**, 1504–1517 (2015).
42. Isaacson, J. & Scanziani, M. How inhibition shapes cortical activity. *Neuron* **72**, 231–243 (2011).
43. Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J. & Scanziani, M. A neural circuit for spatial summation in visual cortex. *Nature* **490**, 226–231 (2012).
44. Atallah, B. V., Bruns, W., Carandini, M. & Scanziani, M. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73**, 159–170 (2012).
45. Sanzeni, A. et al. Inhibition stabilization is a widespread property of cortical networks. *eLife* **9**, e54875 (2020).
46. Keller, A. J. et al. A disinhibitory circuit for contextual modulation in primary visual cortex. *Neuron* **108**, 1181–1193 (2020).
47. Millman, D. J. et al. VIP interneurons in mouse primary visual cortex selectively enhance responses to weak but specific stimuli. *eLife* **9**, e55130 (2020).
48. Veit, J., Handy, G., Mossing, D. P., Doiron, B. & Adesnik, H. Cortical VIP neurons locally control the gain but globally control the coherence of gamma band rhythms. *Neuron* **111**, 405–417 (2023).
49. Stringer, C. et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, eaav7893 (2019).
50. Stringer, C., Michaelos, M., Tsyboulski, D., Lindo, S. E. & Pachitariu, M. High-precision coding in visual cortex. *Cell* **184**, 2767–2778 (2021).
51. Weiss, O., Bounds, H. A., Adesnik, H. & Coen-Cagli, R. Modeling the diverse effects of divisive normalization on noise correlations. *PloS Comput. Biol.* **19**, e1011667 (2023).
52. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106-54 (1962).
53. Barlow, H. B., Blakemore, C. & Pettigrew, J. D. The neural mechanism of binocular depth discrimination. *J. Physiol.* **193**, 327–342 (1967).
54. Priebe, N. J. & Ferster, D. A. Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron* **57**, 482–497 (2008).
55. Priebe, N. J. & Ferster, D. Mechanisms underlying cross-orientation suppression in cat visual cortex. *Nat. Neurosci.* **9**, 552–561 (2006).
56. Freeman, T. C., Durand, S., Kiper, D. C. & Carandini, M. Suppression without inhibition in visual cortex. *Neuron* **35**, 759–771 (2002).
57. Angelucci, A. et al. Circuits and mechanisms for surround modulation in visual cortex. *Annu. Rev. Neurosci.* **40**, 425–451 (2017).
58. Nurminen, L. & Angelucci, A. Multiple components of surround modulation in primary visual cortex: multiple neural circuits with multiple functions? *Vis. Res.* **104**, 47–56 (2014).
59. Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J. & Davis, J. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* **378**, 492–496 (1995).
60. Walker, G. A., Ohzawa, I. & Freeman, R. D. Asymmetric suppression outside the classical receptive field of the visual cortex. *J. Neurosci.* **19**, 10536–10553 (1999).
61. Albrecht, D. G., Geisler, W. S., Frazor, R. A. & Crane, A. M. Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function. *J. Neurophysiol.* **88**, 888–913 (2002).
62. Geisler, W. S. & Albrecht, D. G. Cortical neurons: isolation of contrast gain control. *Vis. Res.* **32**, 1409–1410 (1992).
63. Bair, W., Cavanaugh, J. R. & Movshon, J. A. Time course and time-distance relationships for surround suppression in macaque v1 neurons. *J. Neurosci.* **23**, 7690–7701 (2003).
64. Webb, B. S., Dhruv, N. T., Solomon, S. G., Tailby, C. & Lennie, P. Early and late mechanisms of surround suppression in striate cortex of macaque. *J. Neurosci.* **25**, 11666–11675 (2005).
65. Henry, C. A., Joshi, S., Xing, D., Shapley, R. M. & Hawken, M. J. Functional characterization of the extraclassical receptive field in macaque v1: contrast, orientation, and temporal dynamics. *J. Neurosci.* **33**, 6230–6242 (2013).
66. Walker, G. A., Ohzawa, I. & Freeman, R. D. Binocular cross-orientation suppression in the cat's striate cortex. *J. Neurophysiol.* **79**, 227–239 (1998).
67. Sengpiel, F. & Vorobyov, V. Intracortical origins of interocular suppression in the visual cortex. *J. Neurosci.* **25**, 6394–6400 (2005).
68. Chen, S. C.-Y. et al. Similar neural and perceptual masking effects of low-power optogenetic stimulation in primate v1. *eLife* **11**, e68393 (2022).
69. Werner, G. & Mountcastle, V. B. The variability of central neural activity in a sensory system, and its implications for the central reflection of sensory events. *J. Neurophysiol.* **26**, 958–977 (1963).
70. Parker, A. J. & Newsome, W. T. Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* **21**, 227–277 (1998).
71. Vogels, R. & Orban, G. A. How well do response changes of striate neurons signal differences in orientation: a study in the discriminating monkey. *J. Neurosci.* **10**, 3543–3558 (1990).
72. Goris, R. L., Ziemba, C. M., Stine, G. M., Simoncelli, E. P. & Movshon, J. A. Dissociation of choice formation and choice-correlated activity in macaque visual cortex. *J. Neurosci.* **37**, 5195–5203 (2017).
73. Jasper, A. I., Tanabe, S. & Kohn, A. Predicting perceptual decisions using visual cortical population responses and choice history. *J. Neurosci.* **39**, 6714–6727 (2019).
74. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
75. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, 2005).
76. Tolhurst, D. J., Movshon, J. A. & Dean, A. F. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vis. Res.* **23**, 775–785 (1983).
77. Charles, A. S., Park, M., Weller, J. P., Horwitz, G. D. & Pillow, J. W. Dethroning the Fano factor: a flexible, model-based approach to partitioning neural variability. *Neural Comput.* **30**, 1012–1045 (2018).
78. Nurminen, L., Bijanzadeh, M. & Angelucci, A. Size tuning of neural response variability in laminar circuits of macaque primary visual cortex. Preprint at *bioRxiv* https://doi.org/10.1101/2023.01.17.524397 (2023).
79. Troyer, T. W., Krukowski, A. E., Priebe, N. J. & Miller, K. D. Contrast-invariant orientation tuning in cat visual cortex: feedforward tuning and correlation-based intracortical connectivity. *J. Neurosci.* **18**, 5908–5927 (1998).
80. McLaughlin, D., Shapley, R., Shelley, M. & Wielaard, D. J. A neuronal network model of macaque primary visual cortex (v1): orientation selectivity and dynamics in the input layer 4cα. *Proc. Natl Acad. Sci. USA* **97**, 8087–8092 (2000).
81. Rubin, D. B., Van Hooser, S. D. & Miller, K. D. The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron* **85**, 402–417 (2015).
82. Heeger, D. J. & Zemlianova, K. O. A recurrent circuit implements normalization, simulating the dynamics of v1 activity. *Proc. Natl Acad. Sci. USA* **117**, 22494–22505 (2020).
83. Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).
84. Bell, A. J. & Sejnowski, T. J. The 'independent components' of natural scenes are edge filters. *Vis. Res.* **37**, 3327–3338 (1997).
85. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
86. Schwartz, O. & Simoncelli, E. P. Natural signal statistics and sensory gain control. *Nat. Neurosci.* **4**, 819–825 (2001).
87. Wiskott, L. & Sejnowski, T. J. Slow feature analysis: unsupervised learning of invariances. *Neural Comput.* **14**, 715–770 (2002).
88. Coen-Cagli, R., Dayan, P. & Schwartz, O. Statistical models of linear and nonlinear contextual interactions in early visual processing. In *Advances in Neural Information Processing Systems 22* (eds Bengio, Y. et al.) https://proceedings.neurips.cc/paper_files/paper/2009/file/be3159ad04564bfb90db9e32851ebf9c-Paper.pdf (Curran Associates, 2009).
89. Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol.* **283**, 79–99 (1978).
90. Heeger, D. J. Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* **9**, 181–197 (1992).
91. Rust, N. C., Schwartz, O., Movshon, J. A. & Simoncelli, E. P. Spatiotemporal elements of macaque V1 receptive fields. *Neuron* **46**, 945–956 (2005).
92. Vintch, B., Movshon, J. A. & Simoncelli, E. P. A convolutional subunit model for neuronal responses in macaque v1. *J. Neurosci.* **35**, 14829–14841 (2015).
93. Goris, R. L., Simoncelli, E. P. & Movshon, J. A. Origin and function of tuning diversity in macaque visual cortex. *Neuron* **88**, 819–831 (2015).
94. Freeman, J. & Simoncelli, E. P. Metamers of the ventral stream. *Nat. Neurosci.* **14**, 1195–1201 (2011).
95. Rust, N. C., Mante, V., Simoncelli, E. P. & Movshon, J. A. How MT cells analyze the motion of visual patterns. *Nat. Neurosci.* **9**, 1421–1431 (2006).
96. Goris, R. L., Putzeys, T., Wagemans, J. & Wichmann, F. A. A neural population model for visual pattern detection. *Psychol. Rev.* **120**, 472 (2013).
97. Hénaff, O. J., Goris, R. L. & Simoncelli, E. P. Perceptual straightening of natural videos. *Nat. Neurosci.* **22**, 984–991 (2019).
98. Carandini, M. & Heeger, D. J. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13**, 51–62 (2012).
99. Cadena, S. A. et al. Deep convolutional models improve predictions of macaque v1 responses to natural images. *PloS Comput. Biol.* **15**, e1006897 (2019).
100. Albrecht, D. G. & Geisler, W. S. Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis. Neurosci.* **7**, 531–546 (1991).
101. Geisler, W. S. Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.* **59**, 167–192 (2008).
102. Yamins, D. L. K. & DiCarlo, J. J. Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* **19**, 356–365 (2016).
103. Barlow, H. Possible principles underlying the transformations of sensory messages. *Sens. Commun.* **1**, 217–234 (1961).
104. Simoncelli, E. P. & Olshausen, B. A. Natural image statistics and neural representation. *Annu. Rev. Psychol.* **24**, 1193–1216 (2001).
105. van Hateren, J. H. & Ruderman, D. L. 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.* **265**, 2315–2320 (1998).
106. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27** 379–423 (1948).

107. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
108. Knill, D. C. & Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
109. Boundy-Singer, Z. M., Ziemba, C. M. & Goris, R. L. T. Confidence reflects a noisy decision reliability estimate. *Nat. Hum. Behav.* **7**, 142–154 (2023).
110. Ma, W. J., Beck, J. M., Latham, P. E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438 (2006).
111. Jazayeri, M. & Movshon, J. A. Optimal representation of sensory information by neural populations. *Nat. Neurosci.* **9**, 690–696 (2006).
112. Buesing, L., Bill, J., Nessler, B. & Maass, W. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PloS Comput. Biol.* **7**, e1002211 (2011).
113. Lochmann, T. & Deneve, S. Neural processing as causal inference. *Curr. Opin. Neurobiol.* **21**, 774–781 (2011).
114. Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E. Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* **16**, 1170–1178 (2013).
115. Meyniel, F., Sigman, M. & Mainen, Z. F. Confidence as Bayesian probability: from neural origins to behavior. *Neuron* **88**, 78–92 (2015).
116. Pouget, A., Drugowitsch, J. & Kepecs, A. Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* **19**, 366–374 (2016).
117. Lange, R. D. & Haefner, R. M. Task-induced neural covariability as a signature of approximate Bayesian learning and inference. *PloS Comput. Biol.* **18**, e21009557 (2022).
118. Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* **14**, 119–130 (2010).
119. Hoyer, P. & Hyvarinen, A. Interpreting neural response variability as Monte Carlo sampling of the posterior. *Advances in Neural Information Processing Systems 15* (2002).
120. Mareschal, I. & Shapley, R. M. Effects of contrast and size on orientation discrimination. *Vis. Res.* **44**, 57–67 (2004).
121. Echeveste, R., Aitchison, L., Hennequin, G. & Lengyel, M. Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. *Nat. Neurosci.* **23**, 1138–1149 (2020).
122. Coen-Cagli, R., Dayan, P. & Schwartz, O. Cortical surround interactions and perceptual salience via natural scene statistics. *PloS Comput. Biol.* **8**, e1002405 (2012).
123. Barbera, D., Priebe, N. J. & Glickfeld, L. L. Feedforward mechanisms of cross-orientation interactions in mouse v1. *Neuron* **110**, 297–311 (2022).
124. Lauritzen, T. Z., Krukowski, A. E. & Miller, K. D. Local correlation-based circuitry can account for responses to multi-grating stimuli in a model of cat V1. *J. Neurophysiol.* **86**, 1803–1815 (2001).
125. Kayser, A. S., Priebe, N. J. & Miller, K. D. Contrast-dependent nonlinearities arise locally in a model of contrast-invariant orientation tuning. *J. Neurophysiol.* **85**, 2130–2149 (2001).
126. Carandini, M., Heeger, D. J. & Senn, W. A synaptic explanation of suppression in visual cortex. *J. Neurosci.* **22**, 10053–10065 (2002).
127. Kara, P., Reinagel, P. & Reid, R. C. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron* **27**, 635–646 (2000).
128. Henry, C. A., Jazayeri, M., Shapley, R. M. & Hawken, M. J. Distinct spatiotemporal mechanisms underlie extra-classical receptive field modulation in macaque V1 microcircuits. *eLife* **9**, e54264 (2020).
129. Litwin-Kumar, A. & Doiron, B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat. Neurosci.* **15**, 1498–1505 (2012).
130. Ahmadian, Y., Rubin, D. B. & Miller, K. D. Analysis of the stabilized supralinear network. *Neural Comput.* **25**, 1994–2037 (2013).
131. Hennequin, G., Ahmadian, Y., Rubin, D. B., Lengyel, M. & Miller, K. D. The dynamical regime of sensory cortex: stable dynamics around a single stimulus-tuned attractor account for patterns of noise variability. *Neuron* **98**, 846–860 (2018).
132. Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D. & Ferster, D. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron* **62**, 578–592 (2009).
133. Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. & Suarez, H. H. Recurrent excitation in neocortical circuits. *Science* **269**, 981–985 (1995).
134. Bosking, W. H., Zhang, Y., Schofield, B. & Fitzpatrick, D. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* **17**, 2112–2127 (1997).
135. Borg-Graham, L. J., Monier, C. & Frégnac, Y. Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature* **393**, 369–373 (1998).
136. Gilbert, C. D. & Wiesel, T. N. Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. *Nature* **280**, 120–125 (1979).
137. Softky, W. R. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* **13**, 334–350 (1993).
138. Mainen, Z. F. & Sejnowski, T. J. Reliability of spike timing in neocortical neurons. *Science* **268**, 1503–1506 (1995).
139. Schneidman, E., Freedman, B. & Segev, I. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput.* **10**, 1679–1703 (1998).
140. O'Donnell, C. & van Rossum, M. C. Systematic analysis of the contributions of stochastic voltage gated channels to neuronal noise. *Front. Comput. Neurosci.* **8**, 105 (2014).
141. Stevens, C. F. & Zador, A. M. Input synchrony and the irregular firing of cortical neurons. *Nat. Neurosci.* **1**, 210–217 (1998).
142. DeWeese, M. R. & Zador, A. M. Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *J. Neurosci.* **26**, 12206–12218 (2006).
143. Tsodyks, M. V. & Sejnowski, T. Rapid state switching in balanced cortical network models. *Network* **6**, 111–124 (1995).
144. Van Vreeswijk, C. & Sompolinsky, H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274**, 1724–1726 (1996).
145. Troyer, T. W. & Miller, K. D. Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. *Neural Comput.* **9**, 971–983 (1997).
146. Amit, D. & Brunel, N. Dynamics of a recurrent network of spiking neurons before and following learning. *Netw. Comput. Neural Syst.* **8**, 373–404 (1997).
147. Ahmadian, Y. & Miller, K. D. What is the dynamical regime of cerebral cortex? *Neuron* **109**, 3373–3391 (2021).
148. Kadmon, J. & Sompolinsky, H. Transition to chaos in random neuronal networks. *Phys. Rev. X* **5**, 041030 (2015).
149. Deco, G. & Hugues, E. Neural network mechanisms underlying stimulus driven variability reduction. *PloS Comput. Biol.* **8**, e1002395 (2012).
150. Smith, G. B., Hein, B., Whitney, D. E., Fitzpatrick, D. & Kaschube, M. Distributed network interactions and their emergence in developing neocortex. *Nat. Neurosci.* **21**, 1600–1608 (2018).
151. Trägenap, S., Whitney, D. E., Fitzpatrick, D. & Kaschube, M. Visual experience drives the development of novel and reliable visual representations from endogenously structured networks. *J. Vis.* **23**, 5225–5225 (2023).
152. Ponce-Alvarez, A., Thiele, A., Albright, T. D., Stoner, G. R. & Deco, G. Stimulus-dependent variability and noise correlations in cortical MT neurons. *Proc. Natl Acad. Sci. USA* **110**, 13162–13167 (2013).
153. Bressloff, P. C. Stochastic neural field model of stimulus-dependent variability in cortical neurons. *PloS Comput. Biol.* **15**, e1006755 (2019).
154. Rajan, K., Abbott, L. & Sompolinsky, H. Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E* **82**, 011903 (2010).
155. Mossing, D. P., Veit, J., Palmigiano, A., Miller, K. D. & Adesnik, H. Antagonistic inhibitory subnetworks control cooperation and competition across cortical space. Preprint at *bioRxiv* https://doi.org/10.1101/2021.03.31.437953 (2021).
156. Di Santo, S. et al. Unifying model for three forms of contextual modulation including feedback input from higher visual areas. Preprint at *bioRxiv* https://doi.org/10.1101/2022.05.27.493753 (2022).
157. Tsodyks, M. V., Skaggs, W. E., Sejnowski, T. J. & McNaughton, B. L. Paradoxical effects of external modulation of inhibitory interneurons. *J. Neurosci.* **17**, 4382–4388 (1997).
158. Ekelmans, P., Kraynyukova, N. & Tchumatchenko, T. Targeting operational regimes of interest in recurrent neural networks. *PloS Comput. Biol.* **19**, e1011097 (2023).
159. Murphy, B. K. & Miller, K. D. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).
160. Finn, I. M., Priebe, N. J. & Ferster, D. The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron* **54**, 137–152 (2007).
161. Holt, C. J., Miller, K. D. & Ahmadian, Y. The stabilized supralinear network accounts for the contrast dependence of visual cortical gamma oscillations. Preprint at *bioRxiv* https://doi.org/10.1101/2023.05.11.540442 (2023).
162. Spoerer, C. J., Kietzmann, T. C., Mehrer, J., Charest, I. & Kriegeskorte, N. Recurrent neural networks can explain flexible trading of speed and accuracy in biological vision. *PloS Comput. Biol.* **16**, e1008215 (2020).
163. Kar, K., Kubilius, J., Schmidt, K., Issa, E. B. & DiCarlo, J. J. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat. Neurosci.* **22**, 974–983 (2019).
164. Nayebi, A. et al. Task-driven convolutional recurrent models of the visual system. *Advances in Neural Information Processing Systems 31* (2018).
165. Miller, M., Chung, S. & Miller, K. D. Divisive feature normalization improves image recognition performance in AlexNet. In *International Conference on Learning Representations* (Vienna, 2021).
166. Maheswaranathan, N. et al. Interpreting the retinal neural code for natural scenes: from computations to neurons. *Neuron* **111**, 2742–2755 (2023).
167. Burg, M. F. et al. Learning divisive normalization in primary visual cortex. *PloS Comput. Biol.* **17**, e1009028 (2021).
168. Pan, X., DeForge, A. & Schwartz, O. Generalizing biological surround suppression based on center surround similarity via deep neural network models. *PloS Comput. Biol.* **19**, e1011486 (2023).
169. Zoccolan, D., Cox, D. D. & DiCarlo, J. J. Multiple object response normalization in monkey inferotemporal cortex. *J. Neurosci.* **25**, 8150–8164 (2005).
170. Olsen, S. R., Bhandawat, V. & Wilson, R. I. Divisive normalization in olfactory population codes. *Neuron* **66**, 287–299 (2010).
171. Rabinowitz, N. C., Willmore, B. D., Schnupp, J. W. & King, A. J. Contrast gain control in auditory cortex. *Neuron* **70**, 1178–1191 (2011).
172. Ohshiro, T., Angelaki, D. E. & DeAngelis, G. C. A neural signature of divisive normalization at the level of multisensory integration in primate cortex. *Neuron* **95**, 399–411 (2017).
173. Louie, K., Grattan, L. E. & Glimcher, P. W. Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* **31**, 10627–10639 (2011).
174. Churchland, A. K., Kiani, R. & Shadlen, M. N. Decision-making with multiple alternatives. *Nat. Neurosci.* **11**, 693–702 (2008).
175. Fenno, L., Yizhar, O. & Deisseroth, K. The development and application of optogenetics. *Annu. Rev. Neurosci.* **34**, 389–412 (2011).

176. Cohen, M. R. & Kohn, A. Measuring and interpreting neuronal correlations. *Nat. Neurosci.* **14**, 811 (2011).
177. Chung, S. & Abbott, L. Neural population geometry: an approach for understanding biological and artificial neural networks. *Curr. Opin. Neurobiol.* **70**, 137–144 (2021).
178. Kriegeskorte, N. & Wei, X.-X. Neural tuning and representational geometry. *Nat. Rev. Neurosci.* **22**, 703–718 (2021).
179. Kohn, A. & Smith, M. A. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* **25**, 3661–3673 (2005).
180. Cohen, M. R. & Maunsell, J. H. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12**, 1594–1600 (2009).
181. Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. *Neuron* **63**, 879–888 (2009).
182. Ruff, D. A. & Cohen, M. R. Global cognitive factors modulate correlated response variability between v4 neurons. *J. Neurosci.* **34**, 16408–16416 (2014).
183. Verhoef, B.-E. & Maunsell, J. H. Attention-related changes in correlated neuronal activity arise from normalization mechanisms. *Nat. Neurosci.* **20**, 969–977 (2017).
184. Hennequin, G. & Lengyel, M. Characterizing variability in nonlinear recurrent neuronal networks. Preprint at https://doi.org/10.48550/arXiv.1610.03110 (2016).
185. Geisler, W. S. & Albrecht, D. Bayesian analysis of identification performance in monkey visual cortex: nonlinear mechanisms and stimulus certainty. *Vis. Res.* **35**, 2723–2730 (1995).
186. Wainwright, M. J. & Simoncelli, E. P. in *Adv. Neural Information Processing Systems* (*NIPS*99*) Vol. 12 (eds Solla, S. A. et al.) 855–861 (MIT Press, 2000).
187. Churchland, A. K. et al. Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).

Acknowledgements

We thank D. Heeger, S. Martiniani and Y. Ahmadian for the helpful discussions. This work was supported by the US National Institutes of Health grants EY032999 (R.L.T.G.), EY030578 and DA056400 (R.C.-C.), EY025102, EY024071 and NS120562 (N.J.P), and U01NS108683 and U19NS107613 (K.D.M.), by CAREER award #2146369 (R.L.T.G.), by award DBI-1707398 (K.D.M.) from the National Science Foundation, by a Wellcome Trust Investigator Award in Science 212262/Z/18/Z (M.L.), and by Simons Foundation award 543017 and the Gatsby Charitable Foundation (K.D.M.).

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41583-024-00795-0.

Peer review information *Nature Reviews Neuroscience* thanks Ralf Haefner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.