



ELSEVIER

Neural processing as causal inference

Timm Lochmann¹ and Sophie Deneve^{2,3}

Perception is about making sense, that is, understanding what events in the outside world caused the sensory observations. Consistent with this intuition, many aspects of human behavior confronting noise and ambiguity are well explained by principles of causal inference. Extending these insights, recent studies have applied the same powerful set of tools to perceptual processing at the neural level. According to these approaches, microscopic neural structures solve elementary probabilistic tasks and can be combined to construct hierarchical predictive models of the sensory input. This framework suggests that variability in neural responses reflects the inherent uncertainty associated with sensory interpretations and that sensory neurons are active predictors rather than passive filters of their inputs. Causal inference can account parsimoniously quantitatively for non-linear dynamical properties in single synapses, single neurons and sensory receptive fields.

Addresses

¹ College of Computer, Mathematical, and Natural Sciences, 1207 Bioscience Research Building, University of Maryland, College Park, MD 20742, United States

² Laboratoire de Neurosciences Cognitives, U960, INSERM, France

³ Département d'études Cognitives, Ecole Normale Supérieure, 29, rue d'Ulm, 75005 Paris, France

Corresponding author: Deneve, Sophie (sophie.deneve@ens.fr)

Current Opinion in Neurobiology 2011, 21:774–781

This review comes from a themed issue on
Networks, Circuits and Computation
Edited by Peter Dayan, Dan Feldman, Marla Feller

Available online 11th July 2011

0959-4388/\$ – see front matter
Published by Elsevier Ltd.

DOI 10.1016/j.conb.2011.05.018

Introduction

Computers can beat the best chess players but cannot yet manipulate objects, walk or segment visual scenes like a 3-year-old child. What makes these apparently simple tasks so hard is the presence of multiple sources of uncertainty. The sensory feedback we receive from the environment is by nature incomplete, noisy and ambiguous. In addition, the consequences of our actions are often hard to predict. Recent years have seen the growing use of normative Bayesian models to describe behavior, perception and reasoning [1–5]. These models formalize sensory perception, motor control or behavioral strategies as inferring the probability of underlying causes for the sensory observations and rewards.

In contrast, the neural basis of probabilistic inference remains largely unknown. Neural systems processing noisy and ambiguous inputs have to solve the same elementary statistical problems as those faced, at a much higher level, by behaving subjects. This motivated recent works interpreting neural processing as probabilistic inference in underlying causal models.

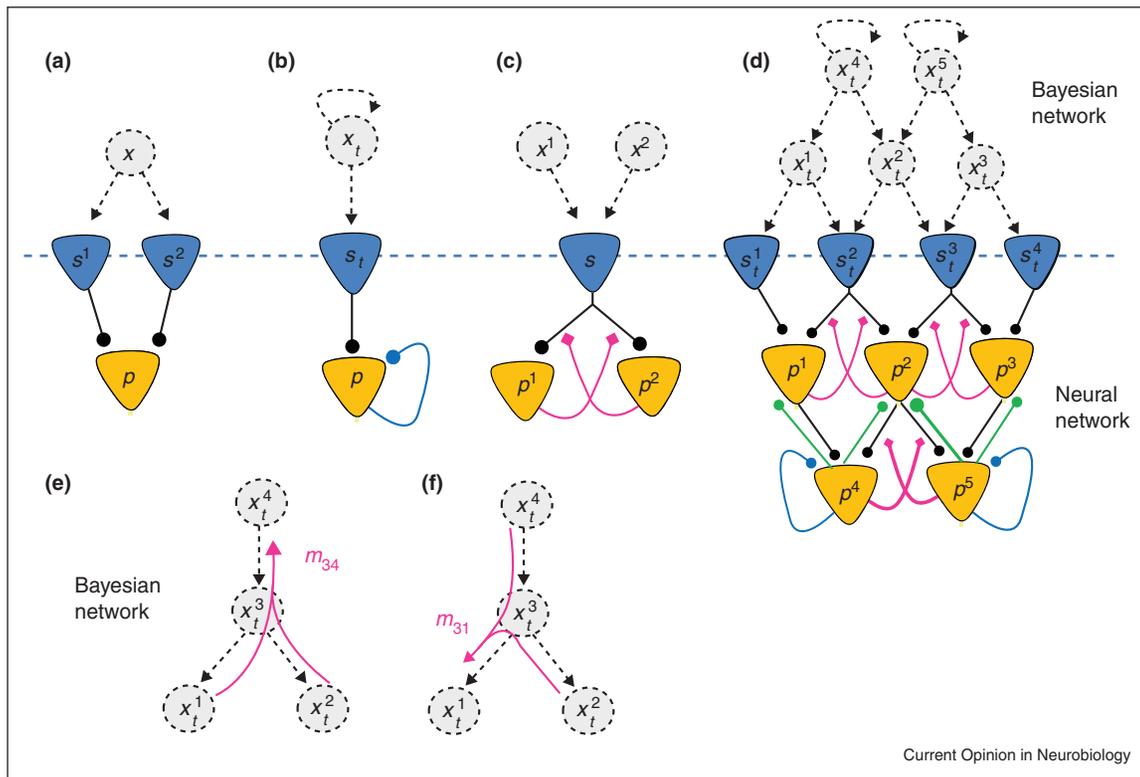
Bayesian networks and neural networks

Most statistical problems can be formalized by a *Bayesian network* describing the causal relationship between ‘observed variables’ and ‘hidden variables’. For example, localizing an object given visual and auditory cues can be done by inferring the object location x that most probably caused the noisy visual and auditory sensory inputs. This can be represented by the Bayesian network shown in [Figure 1a](#) (dashed nodes and arrows), with causal arrows going from the hidden variable x to the observed variables s^1 and s^2 . Humans are able to solve this cue combination task nearly optimally, taking into account the reliability of each modality to infer the most probable multimodal estimate [2,6]. Similarly, sensory neurons combine multiple noisy synaptic inputs (s^1 and s^2) in order to respond to the stimulus x that caused these inputs.

It was shown recently that such a cue combination could be performed by single neurons [7,8] or populations of neurons [9] through a linear combination of their inputs. The corresponding neural network is represented in the lower half of [Figure 1a](#) (plain connections and triangles). A neuron or neural population (triangular orange unit) sums feed-forwards inputs from s^1 and s^2 in order to infer the probability of stimulus, $p(x|s^1, s^2)$. Note that the neural network resembles a mirror image of the Bayesian network it implements. Causal arrows go ‘down’ from the hidden cause x to the sensory observations s while connections in the neural network go ‘up’ from the sensory inputs to the neuron coding for x .

Sensory and motor variables are not static as they typically change over time: they not only cause the sensory input but also predict (cause) their own future states (circular dashed arrows in [Figure 1b](#)), corresponding to a ‘Hidden Markov model’ (HMM, upper half). For example, such models have been used to understand how humans optimally track the position of their arm in space [10]. The same problem is faced by a neuron or neural population responding to stimuli in a dynamic environment (orange triangle). Recent approaches used HMMs to understand how biochemical mechanisms of phototransduction [11], single synapses [12,13], single neurons [8,14] or recurrent neural networks [15–17,18,19] could optimally

Figure 1



Bayesian networks and neural networks. (a) Cue combination. Bayesian network (upper half): a stimulus (hidden variable x) causes several sensory inputs (observed variables s^1 and s^2) represented by blue triangles. Dashed circle represents hidden variables; dashed arrows represent causal relationships. Its neural implementation is represented in the lower half: a neuron or neural population (orange triangle) combines the two sensory inputs through feed-forward connections (thick black lines) and infers the probability of the stimulus (p). (b) Temporal integration. Sensory stimulus x_t can change state (circular arrow) over time and causes the sensory input. Circular dashed arrow represents the past state causing the future state of x_t . A neuron or recurrent neural network integrates the sensory input over time (circular blue connection), in effect predicting the temporal dynamics of the stimulus. (c) Explaining away. Two possible stimuli, x^1 and x^2 , can account for the same sensory spike train s . Two competing neurons (or neural populations) infer the probability of the two stimuli. In addition to feed-forward connections, lateral inhibition connections (magenta lines) shunt inputs to the competing neuron when they can be predicted [22**]. (d) Elementary causal models in A, B and C and their neural implementation can be used as building blocks to implement hierarchical predictive models of the sensory input. In addition to feedforward (black), lateral (magenta) and recurrent (blue) connections, the network must also possess feedback connections (green) transmitting beliefs from high level neural representations to sensory representations. (e) Control of message direction in belief propagation. Messages sent between 'pre' and 'post' nodes combine all messages converging to the 'pre' node, except the message coming from 'post' node. Left and right panels represent messages sent from node x^3 at different times during the belief propagation algorithm. A message is sent from x^3 to x^4 in the left panel, while a message is sent from x^3 to x^1 in the right panel. Excluding the messages sent in the reverse direction (from x^4 to x^3 and from x^1 to x^3) prevents redundant messages from looping into the Bayesian network.

integrate their noisy inputs over time. Once again, the neural implementation is a mirror image of the underlying Bayesian network. Recurrent or self connections in the neural network (blue) predict the future state of x and thus integrate sensory evidence over time.

Finally, sensory observations are often ambiguous. For example, at a cocktail party, speech sounds could come from multiple speakers. This can be represented by two hidden variables, x^1 and x^2 , causing the same sensory input s (dashed arrows and nodes in Figure 1c). Experimental evidence suggests that humans use probabilistic inference to infer common sources of visual and auditory inputs [20]. A similar problem is faced by sensory neurons

representing overlapping stimuli (orange triangles). Inputs from the overlap could provide evidence for either stimulus, requiring a competition between the neurons detecting them [21,22**,23]. This competition is called 'explaining away': a neuron that can explain the sensory input s should prevent another neuron from responding to this same input, since it is already accounted for (magenta connections). In effect, neurons remove their own prediction of s from the input to other neurons through lateral inhibition [22**,24]. The way neuron p^1 responds to the sensory input s depends on the activity of neuron p^2 . As a result, neuron p^2 reshapes the receptive field of neuron p^1 (Lochmann *et al.*: **Perceptual Inference predicts contextual modulations of sensory responses**. *J Neurosci*, 2011,

unpublished data). This was recently shown to account for many contextual and adaptive changes in sensory receptive fields [22^{••},25,26]. Note that because of explaining away, the neural implementation of a Bayesian network is not a 'perfect' mirror image of the causal model but contains additional connections and complexity.

Elementary Bayesian networks (upper halves of Figure 1a–d) can be used to formalize basic problems solved by neural micro-circuits (lower halves). In turn, these Bayesian networks and their neural mirror image can be combined to implement hierarchical causal models ('belief networks') of the external world (Figure 1d). The success of 'Infomax' and 'sparse coding' approaches suggests that hierarchical causal inference is essential to understand how the brain processes sensory information. These approaches used unsupervised learning on large sets of natural stimuli to learn a set of independent features generating natural sensory statistics, which is equivalent to maximizing information transfer, or learning a two layer Bayesian network [27]. While these Infomax approaches reproduced the receptive fields in retinal ganglion cells [28], V1 simple cells [25,29] and auditory nerve fibers [30], these theories based purely of information theory (as opposed to Bayesian models) may sometime obscure the true nature of sensory representations. Elementary features could be very different from sensory receptive fields. For example, inhibitory surrounds and 'repulsive' effects of adaptation (i.e. tilt illusion) often reflect effects of explaining away and *not* properties of the elementary features [25].

More importantly, the Bayesian framework can re-interpret top-down and bottom-up processing in the hierarchical brain as propagation of messages (i.e. belief propagation [31]) from sensory areas to abstract representations and back. Recent models proposed that 'beliefs' could be computed by the firing rates of neural populations [32,33], or transmitted as 'particles', that is, multiple samples from probability distributions [34]. Alternatively, recurrent networks of integrate and fire neurons could perform belief propagation with spikes coding for sudden increases in probability [18^{••},35].

However, special care must be taken to control the direction of message propagation in probabilistic inference and their neural counterpart (Figure 1e). A message sent from x^3 to x^4 should contain all the information received by x^3 , *except* the information sent from x^4 to x^3 (Figure 1e, left panel). Vice versa, a message sent from x^3 to x^1 should contain all information received by x^3 , *except* information sent from x^1 to x^3 (Figure 1e, right panel). Otherwise, inference (belief propagation) is loopy and generates false beliefs. This can be avoided in a neural structure by transferring only *prediction errors*, that is, the difference between the inputs and a prediction of these inputs by the next layer [36] or by local inhibitory loops cancelling reverberated spikes [35].

Several recent models postulated specific neural implementations of hierarchical inference in canonical cortical microcircuits [32,36,37]. These models hypothesized how different types of messages (i.e. top down versus bottom up, prediction errors) are encoded by different neuron subtypes and cortical layers. However, this remains highly speculative because of our incomplete knowledge of the anatomy and dynamics of cortical microcircuits.

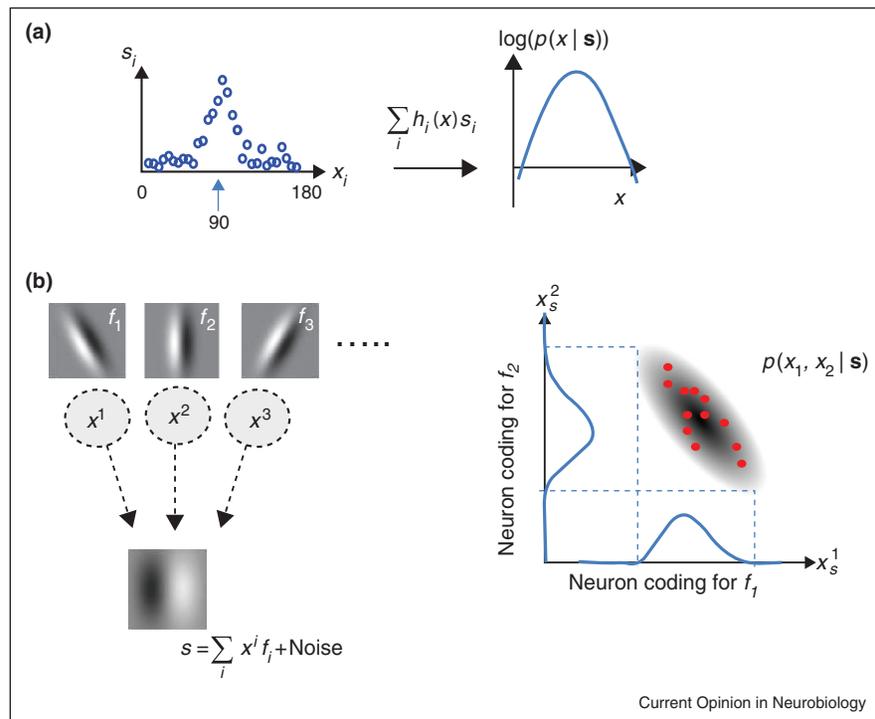
Neural coding of probability

Other recent approaches focused on how neurons compute and represent uncertainty. In particular, plausible models should account for the large variability of cortical responses. To a first approximation, cortical spike trains can be described as Poisson processes whose only reproducible feature is their instantaneous firing rate [38]. Spikes look like random samples from this rate occurring at unpredictable times. How can such variability be reconciled with (or exploited for) probabilistic computations?

The framework of 'Probabilistic population codes' (PPC) proposes that this variability implicitly represents probabilities [9]. The log of the probability distribution of a variable encoded by a population of neurons can be computed as a weighted sum of their spike counts (Figure 2a). In particular, weights correspond to the log of the tuning curves when neurons are uncorrelated [39]. This implies that cue combinations (Figure 1a) can be performed by a weighted sum of spikes, and thus by simple convergence of multiple inputs onto an output layer. Note, however, that other forms of probabilistic inference such as marginalization or temporal prediction require non-linear computations and/or approximations. PPC predicts, in particular, that rates are gain modulated by certainty. This was proposed to account for ramp-like growth of neural responses in the parietal cortex during a motion discrimination task [40[•]].

An alternative theory interprets noisy neural responses as samples from underlying probability distributions [41^{••}]. Rather than propagating beliefs, the neural activity pattern performs 'Gibbs sampling': it repeatedly draws the state of each unit from one possible interpretation of its input given the states of other units in the network. Since sensory inputs are inherently ambiguous, the network activity will constantly explore a large range of possible states, resulting in variable neural responses. For example, the same image s could be caused by many different linear mixtures of gabors (feature f_1 and f_2) with different 'contrasts' (x^1 and x^2), for example, $s = x^1 f^1 + x^2 f^2 + \text{Noise}$ (Figure 2b). Rather than computing probability distribution of continuous variables, as in PPC, the neurons repeatedly sample from likely combinations of 'discrete' feature contrasts. For example, PPC and sampling codes differ drastically in their interpretations of orientation tuning curves in visual cortices. PPC

Figure 2



Neural code for probability. (a) Probabilistic population code (PPC). Response of a population of neurons (left panel) tuned to 'orientation' x . Neurons have bell-shaped tuning curves peaking at their 'preferred' orientation x_i , and their spike counts in response to one presentation of orientation $x = 90^\circ$ are plotted as a function of x_i . Because of neural variability (here, Poisson independent noise) the population response is a noisy hill of activity. Neurons with preferred orientation close to 90 are maximally activated. The log probability distribution of x can be computed by summing the spike counts s_i , weighted by 'kernels' $h_i(x)$. For independent Poisson noise, $h_i(x)$ is the log of the tuning curve of neuron i . (b) Sampling code. Left panel: a Bayesian network for a local image patch: given a set of gabor features f_i , many different linear combinations of gabors contrast x^i (hidden variables) could give rise to same sensory input s (equation in left schema). Right panel: result of sampling from likely combinations of contrast in neurons selective to f_1 and f_2 . Neural activity sample from possible interpretations of s (i.e. generates multiple successive 'guesses', red dots in the right panel). Summing the network responses over time recovers the posterior probability $p(x_1, x_2 | s)$ (grey hill in right panel). Because the sensory input is highly ambiguous, neural responses are correlated (tilt in grey hill) and highly variable (as seen on neural activity projected on each axis). This figure was adapted from [41**].

assumes that 'orientation' is a single 'hidden' sensory variable. Sampling codes assume that images are combinations of many discrete features (i.e. gabors) and orientation tuning curves are a by-product of inherent uncertainty on the correct combination.

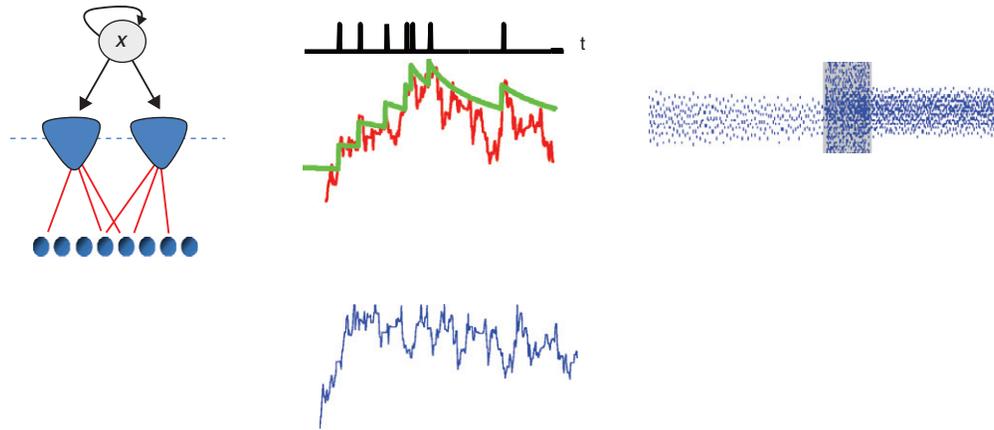
The sampling model predicts that neural responses should become *less variable* when certainty increases (in contrast, PPC predicts higher response gains). Indeed, response variability decreases in many areas during sensory presentation or movement [21]. It also predicts that spontaneous neural responses should be sampled from the prior, that is, the distribution of responses experienced in the past. This was recently shown to be true during development of the ferret visual cortex, where spontaneous activity becomes gradually more similar to responses evoked by natural scenes [41**].

While postulating a neural code for probability, these theories do not specify how such a representation may arise from neural sensory processing. A recent set of

studies [7,8*,18**] suggested that integrate and fire dynamics in recurrent neural networks allow them to naturally represent uncertainty with a spike-based code (Figure 3). This approach postulates that a neuron emits a spike whenever the log probability of the stimulus inferred from the sensory inputs differs significantly from the log probability encoded by output spikes (Figure 3a and b). The membrane potential of each neuron simply integrates the difference between its inputs and a prediction of this input by lateral connections and its own reset (Figure 3b). Using this 'predictive coding' mechanism, noisy sensory signals can be translated into a PPC, with sensory uncertainty reflected in neural variability [18**]. However, apparently unpredictable spike times and Poisson statistics (Figure 3c) stem from sensory uncertainty and do not reflect noisy neural processing [8*,18**].

Note that there are interesting similarities between the neural dynamics predicted by the sampling hypothesis (Figure 2b) and those caused by spike-based predictive

Figure 3



coding (Figure 3a). In both cases, a neuron responds by taking into account its input and the state of other neurons. Neural activity measured in successive time slices can be treated as conditionally independent samples, and the posterior probability can be decoded by summing these neural activities. If we replace ‘log probability’ of orientation x (L and G in Figure 3) by ‘probability of presence’ for gabors at a different orientation, the spiking network generates a sampling-based code (as in Figure 2b) despite its deterministic dynamics.

Neural signatures of inference

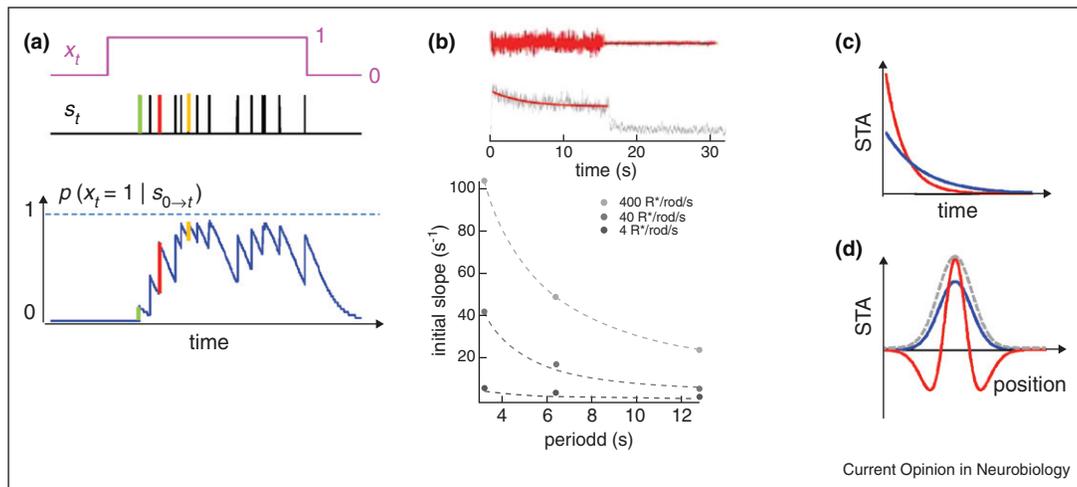
Even without postulating whether and how neurons represent probabilities, Bayesian models can be applied to neural processing as examples of feature detection from noisy and ambiguous inputs. As in behavioral studies, simple causal models can capture highly non-linear and adaptive aspects of neural processing.

‘Classical’ models of neural processing are based on applying fixed spatiotemporal filters to the input [42].

In contrast, probabilistic inference treats inputs differently depending on the context in which they occur (Figure 4a). Because probabilities are normalized, sensory inputs are more informative and more strongly weighted when they occur in situations of uncertainty, for example right after the sudden appearance or disappearance of a stimulus.

For this reason, probabilistic inference can account for many non-linear and adaptive properties of sensory and neural integration. For example, Lengyel *et al.* [12**] proposed that synapses with short-term synaptic facilitation and depression optimally infer the membrane potential of a pre-synaptic neuron (Figure 4a). In another example, Fairhall *et al.* [43**] modeled the dynamics of contrast and light adaptation of retinal ganglion cells (RGCs). After a change in luminance or contrast, RGCs adapt their sensitivity more slowly when changes are smaller or after more time was spent at the previous level of luminance/contrast. This is quantitatively predicted by probabilistic inference (Figure 4b). Another causal

Figure 4



Neural signature of causal inference. A. Simulation with causal model presented in Figure 1b. A binary stimulus x_t (magenta) randomly switches on (from 0 to 1) and off (from 1 to 0), modulating the firing rate of a sensory neuron (black line). This generates a sequence of sensory input spikes S_t (vertical black lines) with Poisson statistics. Bottom panel: Posterior probability of the stimulus (blue line). Each input spike increases the probability, followed by a temporal decay, resembling synaptic integration of input spikes. However, 'EPSPs' do not all have the same size. When certainty is high (green and yellow highlights) the effective sensory weight is low. When certainty is low, the effective sensory weight is high (red highlight). (b) Results from Wark *et al.* [43**]. When responding to full-field stimuli with randomly varying luminance (top red curve) retinal ganglion cells adapt their response gain to the stimulus variance, or contrast. Middle panel: mean synaptic current to a RGC cell in response to a contrast shift. An exponential fit (red line) was performed to measure the speed of adaptation. Bottom panel: initial slope (speed of adaptation) as a function of the period between each switch in contrast and the mean luminance level. Dashed line: fit by a Bayesian model. (c) Temporal RF (spike triggered average STA) of a model neuron inferring the presence of a binary stimulus (Figures 1b and 4a). Simulations are performed with model from [8*,22**]. Inputs used to map the STA were Poisson spike trains with rate sampled from a Gaussian distribution at each time step. The variance of this Gaussian determines the stimulus contrast. STA is longer at low contrast (blue line) than at high contrast (red line). (d) Spatial RF of model neurons inferring the presence of a bell-shaped visual pattern (dashed grey line) in the visual scene. This neuron competes with other neurons representing overlapping stimuli (overlapping dashed grey hills at different locations). Explaining away leads to surround suppression at high contrast (red line) with a small excitatory center. Competition is weaker at low contrast (blue line) resulting in larger center and no suppressive surrounds. The low contrast RF is more similar to stimulus represented by the model neuron (dashed grey 'causal field'). Results adapted from [22**].

inference model was used to account for the dynamics of sensory adaptation and synaptic short-term depression [13*].

Similar mechanisms account for changes in visual temporal receptive fields (tRFs) as a function of contrast (Figure 4c): Noisy inputs are integrated more slowly, leading to longer and more integrative tRFs, as observed in the retina [44], LGN [45] and visual cortices [46]. 'explaining away' (Figure 1c) has also been used to capture spatial aspects [22**] and temporal aspects [13*] of sensory adaptation. For example, competition between neurons representing overlapping stimuli (i.e., with overlapping 'causal fields') creates an 'automatic' coarse-to-fine change in sensory receptive fields (without any associated changes in model parameters) as a function of time and contrast (Figure 4d). Such contextual changes in RFs have been reported in the retina [44] and visual cortex [47,48].

Conclusion

Little experimental evidence has been collected about an explicit neural representation of probability, but recent theories and experimental results suggest that the

responses of spiking neural networks can be interpreted as representing not only sensory and motor variables but also their posterior probability distributions.

Moreover, using causal models could challenge our 'classical' view of neural processing in terms of 'feed-forward' linear filtering of inputs. Examples presented above suggest that neural selectivity could be more accurately and parsimoniously described as a causal (predictive) model of the input rather than by fixed spatiotemporal RFs. Highly non-linear, complex and context dependent neural processes may in fact correspond to detection of much simpler features.

Finally, the inherently modular property of the Bayesian framework brings hope that we will eventually be able to link different levels of analysis, from single synapses to canonical microcircuits and hierarchies of brain areas, in a single coherent framework. In turn, this might allow us to link deficits or pharmacological effects on cortical microcircuits to their effects on behavior [49].

Acknowledgement

This work was supported by an EU grant BIND MECT-CT-20095-024831.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Körding KP: **Decision theory: what 'should' the nervous system do?** *Science* 2007, **318**:606-610.
 2. Ernst MO, Banks MS: **Humans integrate visual and haptic information in a statistically optimal fashion.** *Nature* 2002, **415**:429-433.
 3. Daw ND *et al.*: **Cortical substrates for exploratory decisions in humans.** *Nature* 2006, **441**:876-879.
 4. Kemp C *et al.*: **A probabilistic model of theory formation.** *Cognition* 2009.
 5. Knill DC, Richards W: *Perception as Bayesian inference.* Cambridge, MA: Cambridge University Press; 1996.
 6. Clemens IA *et al.*: **Multisensory processing in spatial orientation: an inverse probabilistic approach.** *J Neurosci* 2011, **31**:5365-5377.
 7. Deneve S: **Bayesian spiking neurons II: learning.** *Neural Comput* 2008, **20**:118-145.
 8. Deneve S: **Bayesian spiking neurons I: inference.** *Neural Comput* 2008, **20**:91-117.
- The problem faced by a neuron detecting a stimulus from its noisy synaptic input is modelled as a probabilistic inference task. On the basis of principle of self consistency, the author shows that this predicts an integrate and fire dynamics where spikes represent sudden increase in probability.
9. Ma WJ *et al.*: **Bayesian inference with probabilistic population codes.** *Nat Neurosci* 2006, **9**:1432-1438.
 10. Wolpert D, Ghahramani Z, Jordan M: **An internal model for sensorimotor integration.** *Science* 1995, **269**:1880-1882.
 11. Houillon A, Bessiere P, Droulez J: **The probabilistic cell: implementation of a probabilistic inference by the biochemical mechanisms of phototransduction.** *Acta Biotheor* 2010, **58**:103-120.
 12. Pfister JP, Dayan P, Lengyel M: **Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials.** *Nat Neurosci* 2010, **13**:1271-1275.
- Synaptic short-term facilitation and depression could allow the post-synaptic neuron to estimate accurately the membrane potential of a pre-synaptic neuron. STP is due the fact that inputs are not equally informative depending on the current level of confidence.
13. Stevenson IH *et al.*: **Sensory adaptation and short term plasticity as Bayesian correction for a changing brain.** *PLoS One* 2010, **5**:e12436.
- Sensory networks face constant 'global' changes in state and excitability that are irrelevant for sensory processing. The authors propose that adaptive mechanisms such as synaptic short-term depression infer those excitability states, exploiting the fact that they vary more slowly than their sensory input, in order to reduce input variability.
14. Lochmann T, Deneve S: **Information transmission with spiking Bayesian neurons.** *New J Phys* 2008, **10**:055019 19 pp.
 15. Rao RP: **Bayesian computation in recurrent neural circuits.** *Neural Comput* 2004, **16**:1-38.
 16. Huys QJ *et al.*: **Fast population coding.** *Neural Comput* 2007, **19**:404-441.
 17. Beck JM, Pouget A: **Exact inferences in a neural implementation of a hidden Markov model.** *Neural Comput* 2007, **19**:1344-1361.
 18. Boerlin M, Deneve S: **Spike-based population coding and working memory.** *PLoS Comput Biol* 2011, **7**:e1001080.
- Recurrent networks of integrate and fire neurons can infer the probability distribution of a dynamic variable from their noisy inputs. In particular, they can perform optimal sensory integration and maintain information in working memory. The model generates poisson-distributed output spike train reflecting uncertainty.
19. Bobrowski O, Meir R, Eldar YC: **Bayesian filtering in spiking neural networks: noise, adaptation, and multisensory integration.** *Neural Comput* 2009, **21**:1277-1320.
 20. Kording KP *et al.*: **Causal inference in multisensory perception.** *PLoS One* 2007, **2**:e943.
 21. Churchland MM *et al.*: **Stimulus onset quenches neural variability: a widespread cortical phenomenon.** *Nat Neurosci* 2010, **13**:369-378.
 22. Lochmann T, Deneve S: In *Contextual modulation of visual receptive fields: a Bayesian perspective, in sensory cue integration.* Edited by Trommershauser J, Kording K, Landy MS. Oxford: Oxford University Press; 2011:448-456.
- In order to analyse sensory scenes, sensory neurons must identify what combination of features best predict their sensory input. 'Explaining away' implies a dynamic reshaping of sensory receptive fields by the spatial and temporal context. Inference in a simple causal model of visual scenes accounts for complex effects of adaptation and surround stimuli on visual responses.
23. Lucke J, Sahani M: **Maximal causes for non-linear component extraction.** *J Mach Learn Res* 2008, **9**:1227-1267.
 24. Srinivasan MV, Laughlin SB, Dubs A: **Predictive coding: a fresh view of inhibition in the retina.** *Proc R Soc Lond B Biol Sci* 1982, **216**:427-459.
 25. Schwartz O, Sejnowski TJ, Dayan P: **Perceptual organization in the tilt illusion.** *J Vis* 2009, **9**:191-220.
 26. Sprattling MW: **Predictive coding as a model of response properties in cortical area V1.** *J Neurosci* 2010, **30**:3531-3543.
 27. Hinton GE, Ghahramani Z: **Generative models for discovering sparse distributed representations.** *Phil Trans R Soc Lond B* 1997, **352**:1177-1190.
 28. Atick JJ, Redlich AN: **What does the retina know about natural scenes?** *Neural Comput* 1992, **4**:196-210.
 29. Olshausen BA, Field DJ: **Sparse coding with an overcomplete basis set: a strategy employed by V1?** *Vis Res* 1997, **37**:3311-3325.
 30. Smith EC, Lewicki MS: **Efficient auditory coding.** *Nature* 2006, **439**:978-982.
 31. Frey B: *Graphical models for machine learning and digital communication.* Cambridge, MA: MIT Press; 1998.
 32. Litvak S, Ullman S: **Cortical circuitry implementing graphical models.** *Neural Comput* 2009, **21**:3010-3056.
 33. Steimer A, Maass W, Douglas R: **Belief propagation in networks of spiking neurons.** *Neural Comput* 2009, **21**:2502-2523.
 34. Lee TS, Mumford D: **Hierarchical Bayesian inference in the visual cortex.** *J Opt Soc Am A Opt Image Sci Vis* 2003, **20**:1434-1448.
 35. Deneve S: **Bayesian inference with recurrent spiking networks.** In *Advances in neural information processing system.* Edited by Lawrence KS, Weiss Y, Bottou L. Cambridge: MIT Press; 2004: 353-360.
 36. Friston K, Kiebel S: **Predictive coding under the free-energy principle.** *Phil Trans R Soc Lond B Biol Sci* 2009, **364**:1211-1221.
 37. George D, Hawkins J: **Towards a mathematical theory of cortical micro-circuits.** *PLoS Comput Biol* 2009, **5**:e1000532.
 38. Tolhurst D, Movshon J, Dean A: **The statistical reliability of signals in single neurons in cat and monkey visual cortex.** *Vis Res* 1982, **23**:775-785.
 39. Graf AB *et al.*: **Decoding the activity of neuronal populations in macaque primary visual cortex.** *Nat Neurosci* 2011, **14**:239-245.
 40. Beck JM *et al.*: **Probabilistic population codes for Bayesian decision making.** *Neuron* 2008, **60**:1142-1152.
- The authors apply the 'probabilistic population code' model to data recorded in medio-temporal cortex (MT) and lateral intraparietal cortex

(LIP), while a monkey performed a motion integration task. Noisy inputs from the population of MT neurons are integrated by a population of LIP cells. This predicts the temporal dynamics of their firing rates and the behavioral data.

41. Berkes P *et al.*: **Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment.** *Science* 2011, **331**:83-87.

Cortical networks could build 'optimal' internal models of the environment by representing the statistical structure of the natural visual input. This predicts that 'spontaneous' activity samples from 'probable' neural responses to natural movies. Indeed, statistics of spontaneous activity in the visual cortex of developing ferrets become progressively more similar to responses to natural movies.

42. Carandini M *et al.*: **Do we know what the early visual system does?** *J Neurosci* 2005, **25**:10577-10597.

43. Wark B, Fairhall A, Rieke F: **Timescales of inference in visual adaptation.** *Neuron* 2009, **61**:750-761.

RGCs adapt to changes in stimulus variance (contrast) and luminance by decreasing their response gains. This adaptation is not instantaneous but requires an accumulation of evidence because of sensory noise. As predicted by Bayesian inference, adaptation is faster when the change

in luminance/contrast is larger, or when the previous level of luminance/contrast lasted longer.

44. Solomon SG, Lee BB, Sun H: **Suppressive surrounds and contrast gain in magnocellular-pathway retinal ganglion cells of macaque.** *J Neurosci* 2006, **26**:8715-8726.
45. Butts DA *et al.*: **Temporal precision in the neural code and the timescales of natural vision.** *Nature* 2007, **449**:92-95.
46. Bair W, Movshon JA: **Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex.** *J Neurosci* 2004, **24**:7305-7323.
47. Sceniak MP *et al.*: **Contrast's effect on spatial summation by macaque V1 neurons.** *Nat Neurosci* 1999, **2**:733-739.
48. Sceniak MP, Hawken MJ, Shapley R: **Contrast-dependent changes in spatial frequency tuning of macaque V1 neurons: effects of a changing receptive field size.** *J Neurophysiol* 2002, **88**:1363-1373.
49. Corlett PR *et al.*: **Glutamatergic model psychoses: prediction error, learning, and inference.** *Neuropsychopharmacology* 2011, **36**:294-315.