Probabilistic Inference and the Brain
SECTION 4: CRITICAL GAPS IN MODELS

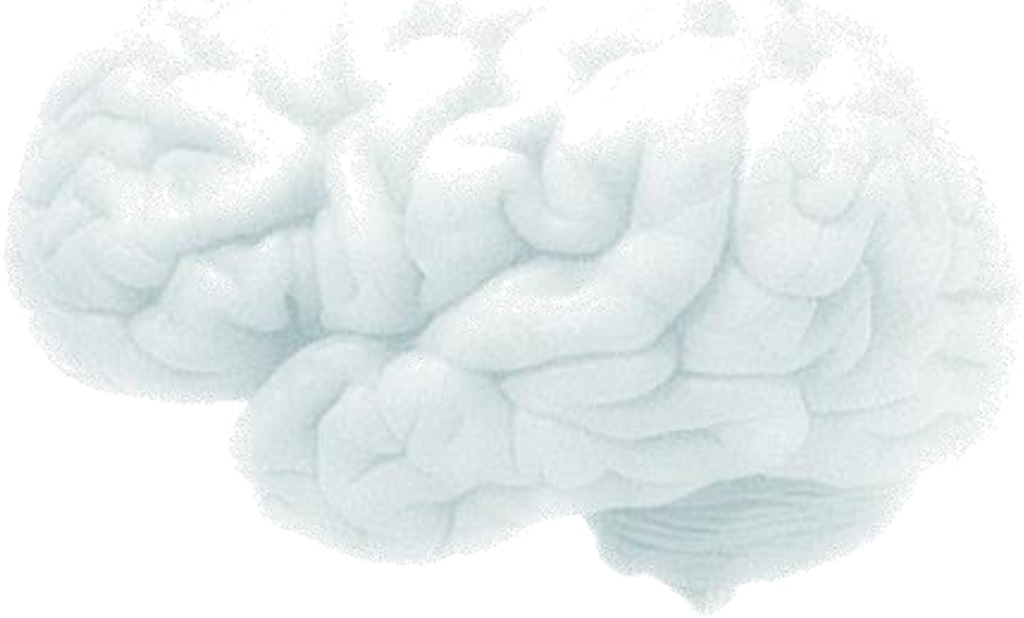
JOINING THE DOTS IN THEORETICAL NEUROBIOLOGY

Karl Friston, University College London

ABSTRACT: My treatment of critical gaps in models of probabilistic inference will focus on the potential of unified theories to "close the gaps" between probabilistic models of perception and evidence accumulation - and how these models can be understood in terms of embodied inference and action. Formally speaking, models of motor control and choice behaviour can be cast in terms of (active) probabilistic inference; however, there are two key outstanding issues. Both pertain to the implementation of active inference in the brain. The first speaks to the distinction between discrete and continuous state-space models – and the requisite message passing schemes (e.g., variational Bayes versus predictive coding). The second key distinction is between mean field descriptions in terms of population dynamics (i.e., sufficient statistics) and their microscopic implementation in terms of spiking neurons (i.e., sampling approaches to probabilistic inference). These are potentially important issues that constrain the interpretation of empirical data – and how these data can be used to adjudicate among different models of the Bayesian brain.
.

Does the brain use continuous or discrete state space models?

Does the brain encode beliefs with ensemble densities or sufficient statistics?



Does the brain use continuous or discrete state-space models?

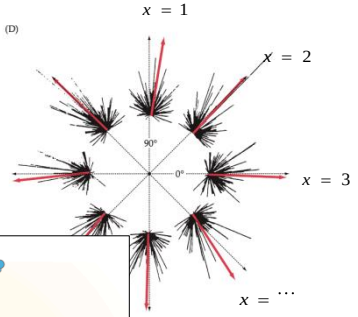
$x \in \mathbb{R}$

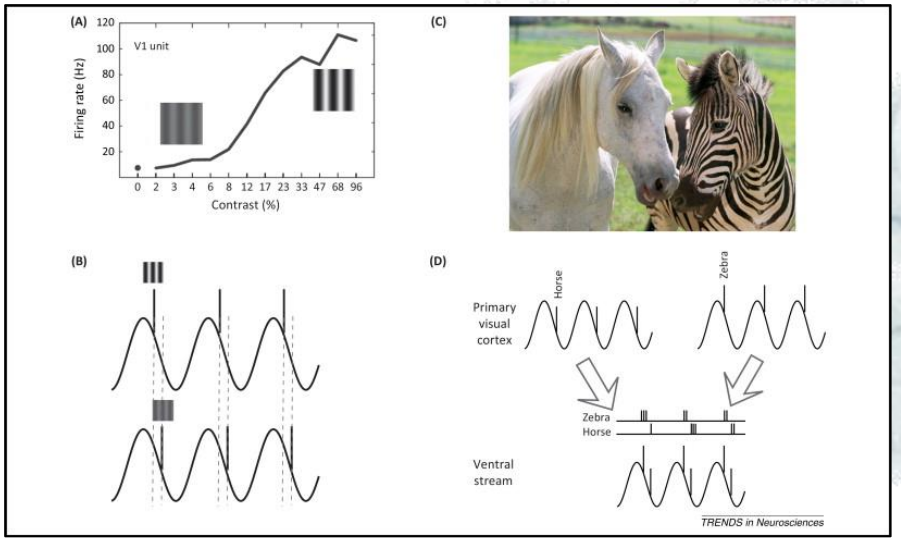
$\mu \in \mathbb{R}$

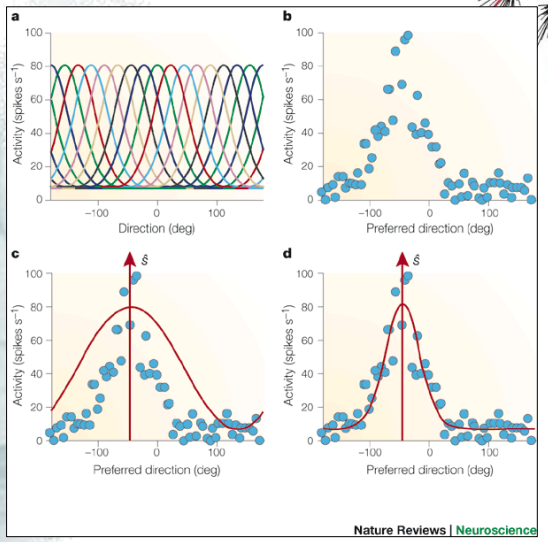
States and their Sufficient statistics

$x \in \{1, \cdots, N\}$

$\mu \in [0, 1]$







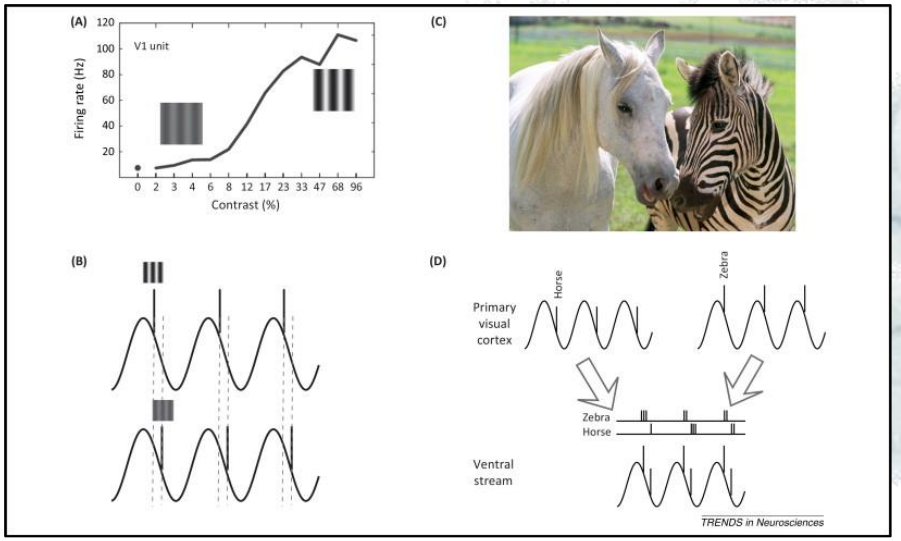
$$q(x \mid \mu) \approx p(x \mid s, m)$$

Approximate Bayesian inference for continuous states: Bayesian filtering

$$x \in \mathbb{R}$$

$$\mu \in \mathbb{R}$$

States and their
Sufficient statistics



$$\mu^{*} = \arg\min F(s, \mu)$$

$$\underbrace{F(s, \mu)}_{\text{Free energy}} \geq -\ln \underbrace{p(s \mid m)}_{\text{Model evidence}}$$

$$q(x \mid \mu) \approx p(x \mid s, m)$$

Hermann von Helmholtz

"Objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism" - von Helmholtz

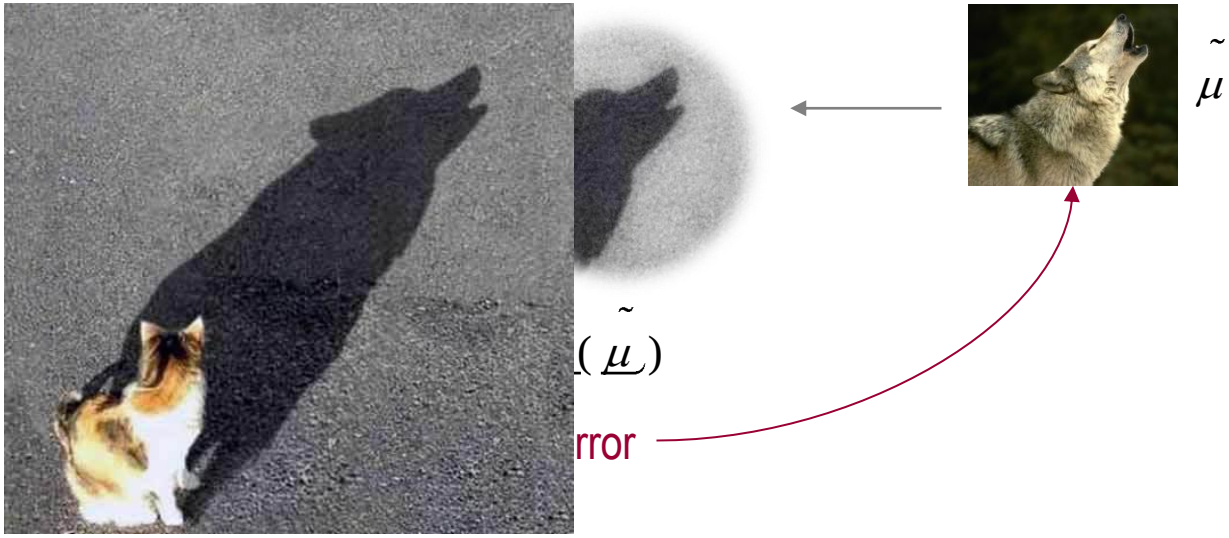
Richard Gregory

Impressions on the Markov blanket…

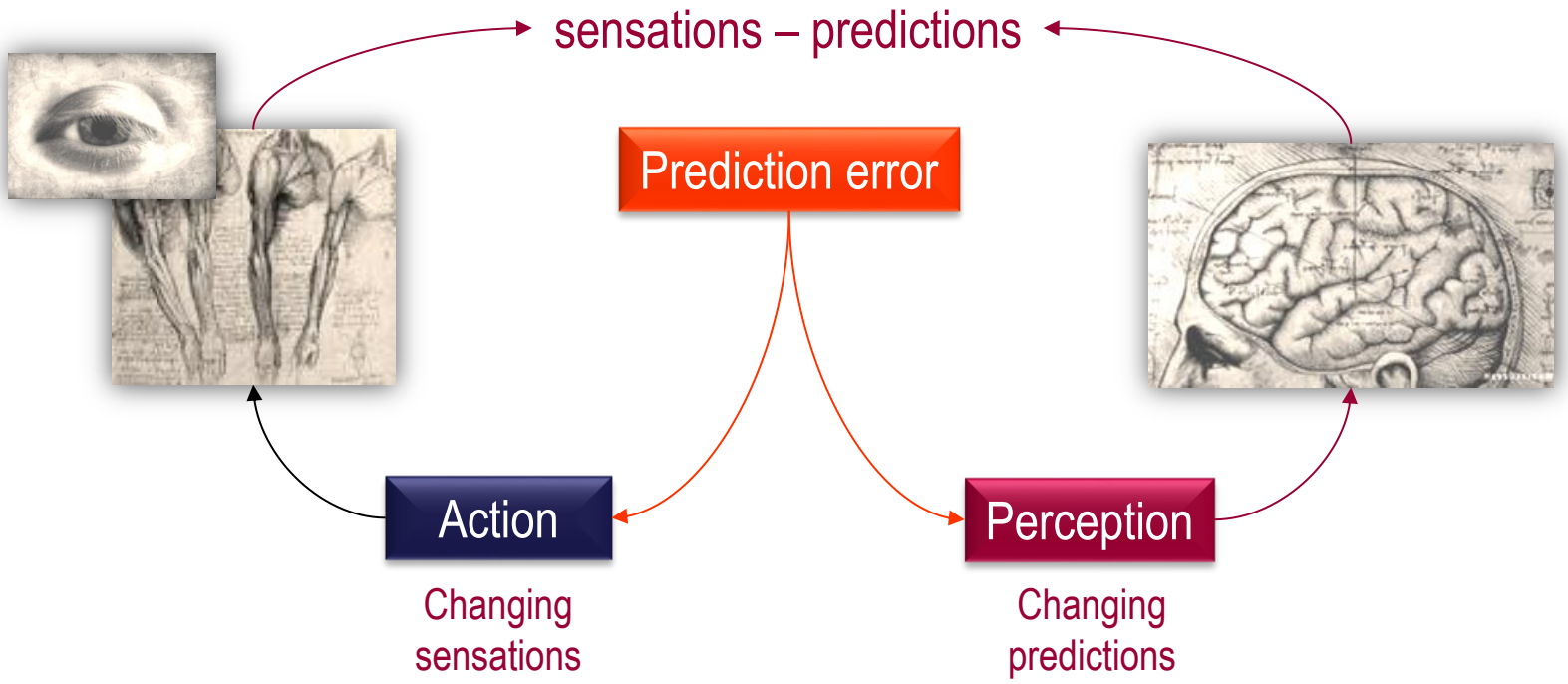
Bayesian filtering and predictive coding

$$\dot{\tilde{\mu}} = (Q - \Gamma)\nabla F(\tilde{s}, \tilde{\mu})$$
$$= D\tilde{\mu} - \Gamma\nabla\varepsilon_{\cdot} \Pi\cdot\varepsilon$$

prediction update

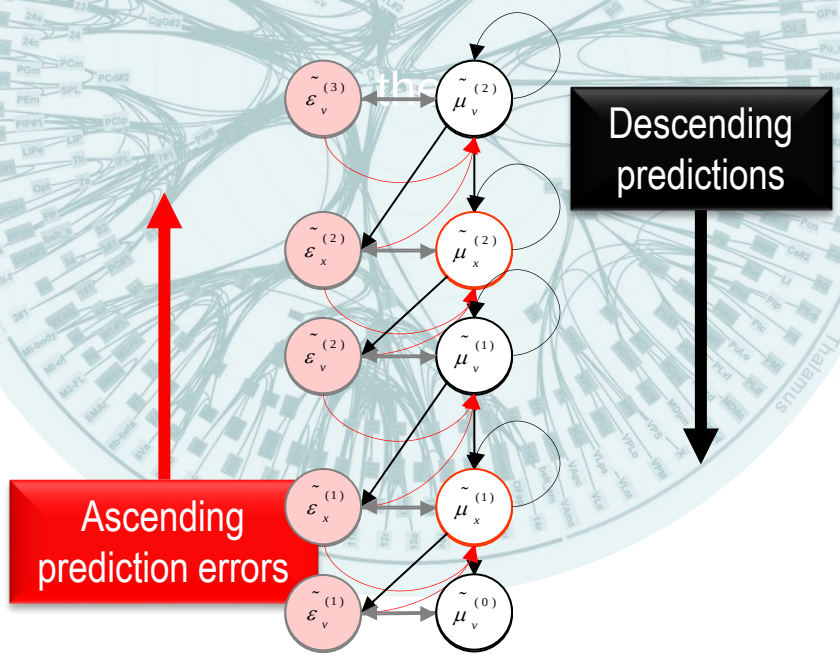


Making our own sensations

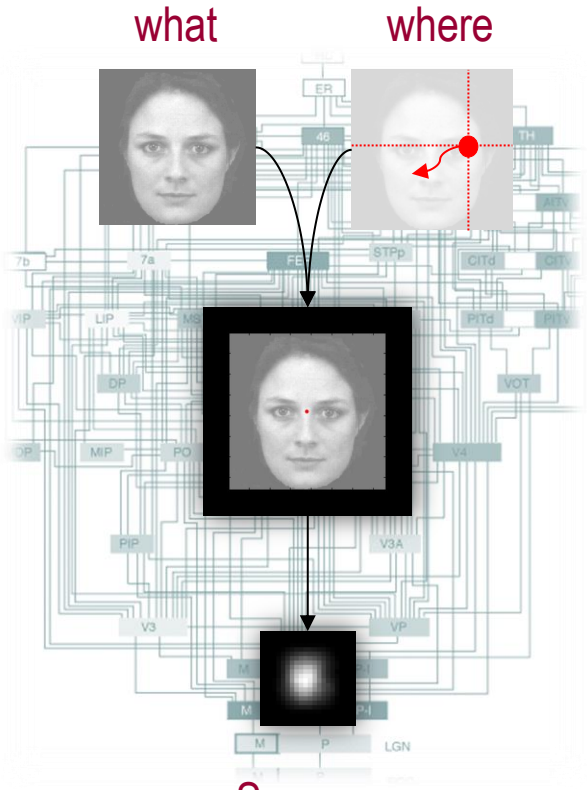


Hierarchical generative models

A simple hierarchy



$$\dot{\tilde{\mu}} = D\,\tilde{\mu} - \Gamma\,\nabla\,\tilde{\varepsilon}\cdot\Pi\cdot\tilde{\varepsilon}$$



What does Bayesian filtering explain?

OPEN ACCESS Freely available online PLOS COMPUTATIONAL BIOLOGY

Hierarchical Models in the Brain

- The hierarchical organisation of cortical areas (c.f., [39])
- Each area comprises distinct neuronal subpopulations, encoding expected states of the world and prediction error (c.f., [72]).
- Extrinsic forward connections convey prediction error (from superficial pyramidal cells) and backward connections mediate predictions, based on hidden and causal states (from deep pyramidal cells) [49].
- Recurrent dynamics are intrinsically stable because they are trying to suppress prediction error [54,64].
- Functional asymmetries in forwards (linear) and backwards (nonlinear) connections may reflect their distinct roles in recognition (c.f., [44]).
- Principal cells elaborating predictions (e.g., deep pyramidal cells) may show distinct (low-pass) dynamics, relative to those encoding error (e.g., superficial pyramidal cells)
- Lateral interactions may encode the relative precision of prediction errors and change in a way that is consistent with classical neuromodulation (c.f., [63,71]).
- The rescaling of prediction errors by recurrent connections, in proportion to their precision, affords a form of cortical bias or gain control [73,74].
- The dynamics of plasticity and modulation of lateral interactions encoding precision or uncertainty (which optimise a path-integral of variational energy) must be slower than the dynamics of neuronal activity (which optimise variational energy *per se*)
- Neuronal activity, synaptic efficacy and neuromodulation must all affect each other; activity-dependent plasticity and neuromodulation shape neuronal responses and:
- Neuromodulatory factors play a dual role in modulating postsynaptic responsiveness (e.g., through modulating in after-hyperpolarising currents) and synaptic plasticity [66,67].

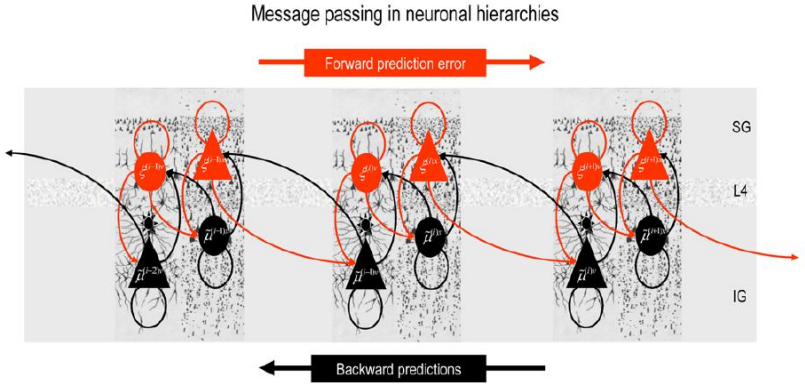


Figure 9. Schematic detailing the neuronal architectures that encode an ensemble density on the states and parameters of hierarchical models. This schematic shows how the neuronal populations of the previous figure may be deployed hierarchically within three cortical areas (or macro-columns). Within each area the cells are shown in relation to the laminar structure of the cortex that includes supra-granular (**SG**) granular (**L4**) and infra-granular (**IG**) layers.
doi:10.1371/journal.pcbi.1000211.g009

What does Bayesian filtering explain?

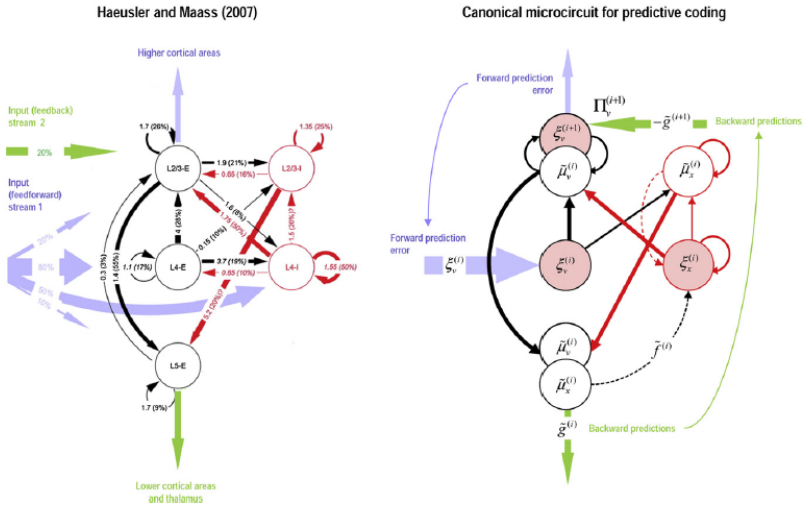
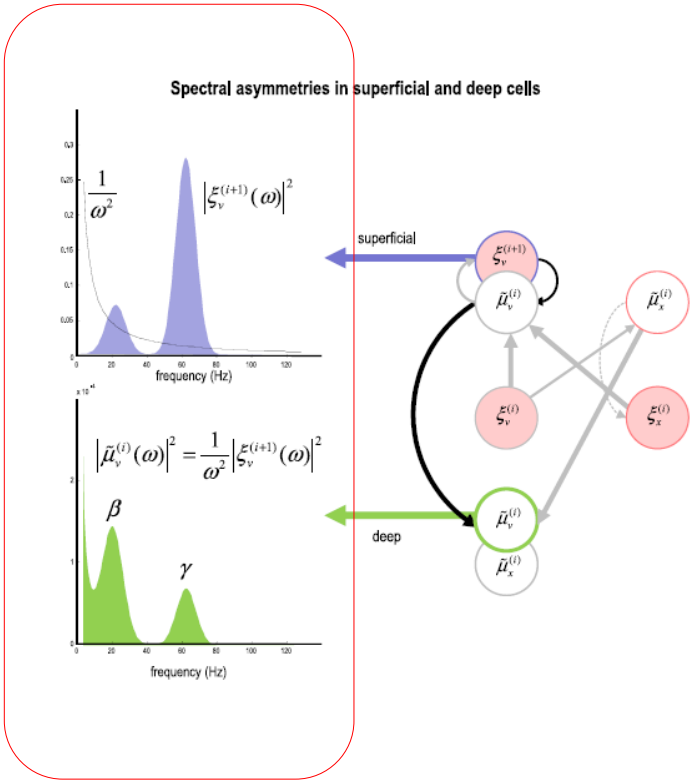
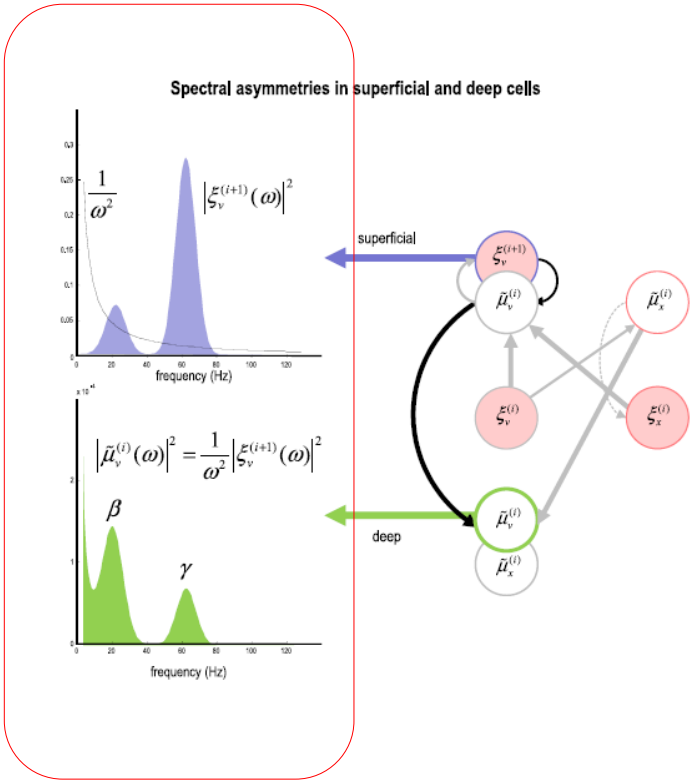


Figure 5. A Canonical Microcircuit for Predictive Coding
Left: the canonical microcircuit based on Haeusler and Maass (2007), in which we have removed inhibitory cells from the deep layers because they have very little interlaminar connectivity. The numbers denote connection strengths (mean amplitude of PSPs measured at soma in mV) and connection probabilities (in parentheses) according to Thomson et al. (2002). Right: the proposed cortical microcircuit for predictive coding, in which the quantities of the previous figure have been associated with various cell types. Here, prediction error populations are highlighted in pink. Inhibitory connections are shown in red, while excitatory connections are in black. The dotted lines refer to connections that are not present in the microcircuit on the left (but see Figure 2). In this scheme, expectations (about causes and states) are assigned to (excitatory and inhibitory) interneurons in the supragranular layers, which are passed to infragranular layers. The corresponding prediction errors occupy granular layers, while superficial pyramidal cells encode prediction errors that are sent forward to the next hierarchical level. Conditional expectations and prediction errors on hidden causes are associated with excitatory cell types, while the corresponding quantities for hidden states are assigned to inhibitory cells. Dark circles indicate pyramidal cells. Finally, we have placed the precision of the feedforward prediction errors against the superficial pyramidal cells. This quantity controls the postsynaptic sensitivity or gain to (intrinsic and top-down) presynaptic inputs. We have previously discussed this in terms of attentional modulation, which may be intimately linked to the synchronization of presynaptic inputs and ensuing postsynaptic responses (Feldman and Friston, 2010; Fries et al., 2001).



What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics



What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

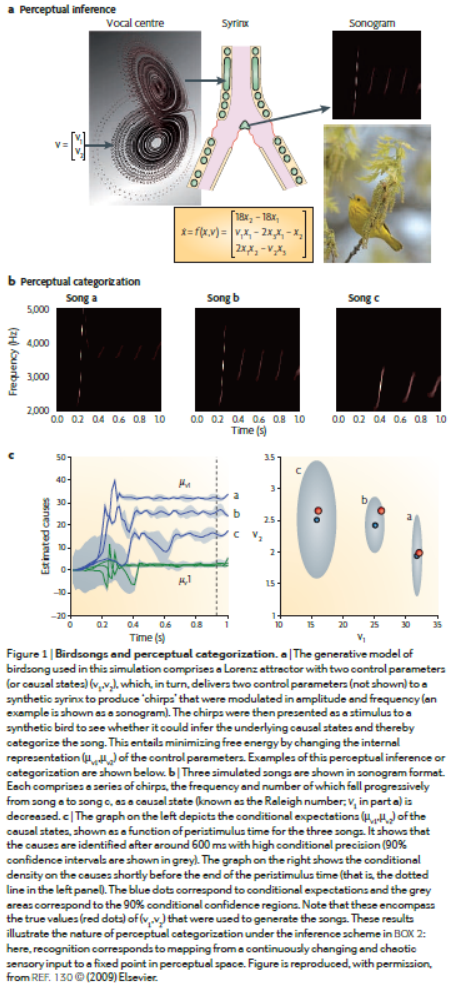


Figure 1 | **Birdsongs and perceptual categorization. a** | The generative model of birdsong used in this simulation comprises a Lorenz attractor with two control parameters (or causal states) (v_1,v_2), which, in turn, delivers two control parameters (not shown) to a synthetic syrinx to produce 'chirps' that were modulated in amplitude and frequency (an example is shown as a sonogram). The chirps were then presented as a stimulus to a synthetic bird to see whether it could infer the underlying causal states and thereby categorize the song. This entails minimizing free energy by changing the internal representation (μ_{v1},μ_{v2}) of the control parameters. Examples of this perceptual inference or categorization are shown below. **b** | Three simulated songs are shown in sonogram format. Each comprises a series of chirps, the frequency and number of which fall progressively from song a to song c, as a causal state (known as the Raleigh number; v_1 in part **a**) is decreased. **c** | The graph on the left depicts the conditional expectations (μ_{v1},μ_{v2}) of the causal states, shown as a function of peristimulus time for the three songs. It shows that the causes are identified after around 600 ms with high conditional precision (90% confidence intervals are shown in grey). The graph on the right shows the conditional density on the causes shortly before the end of the peristimulus time (that is, the dotted line in the left panel). The blue dots correspond to conditional expectations and the grey areas correspond to the 90% conditional confidence regions. Note that these encompass the true values (red dots) of (v_1,v_2) that were used to generate the songs. These results illustrate the nature of perceptual categorization under the inference scheme in BOX 2: here, recognition corresponds to mapping from a continuously changing and chaotic sensory input to a fixed point in perceptual space. Figure is reproduced, with permission, from REF. 130 © (2009) Elsevier.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

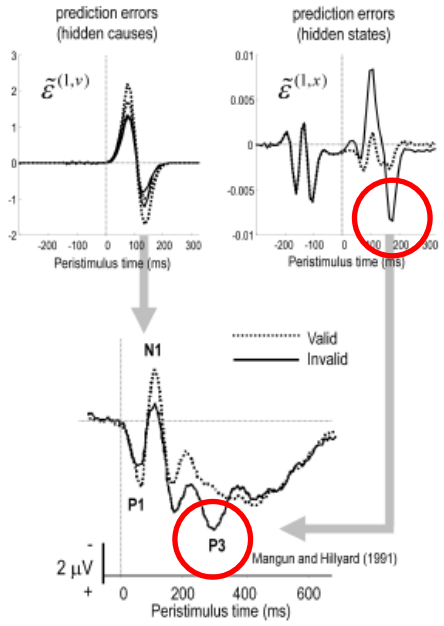


FIGURE 6 | Simulated EEG data from our simulations (upper panels) and empirical EEG data (lower panel) from Mangun and Hillyard (1991). The EEG traces were created from the prediction errors on the hidden causes (left) and states (right). The empirical data were recorded via EEG from the occipital cortex contralateral to the target (i.e., the cortex processing the target). The simulated data exhibits two important features of empirical studies: early in peristimulus time, stimulus-driven responses are greater for valid cues (upper left panel) relative to invalid cues. This is often attributed to a validity enhancement of early (e.g., N1) components. Conversely, later in peristimulus time, invalid responses are greater in amplitude. This can be related to novelty (and salience) responses usually associated with late waveform components (e.g., P3). In the simulations, this invalidity effect is explained simply by greater prediction errors on inferred hidden states encoding precision (upper right panel). It is these prediction errors that report a surprising or novel context, following the failure to predict invalidly cued stimuli in an optimal fashion.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

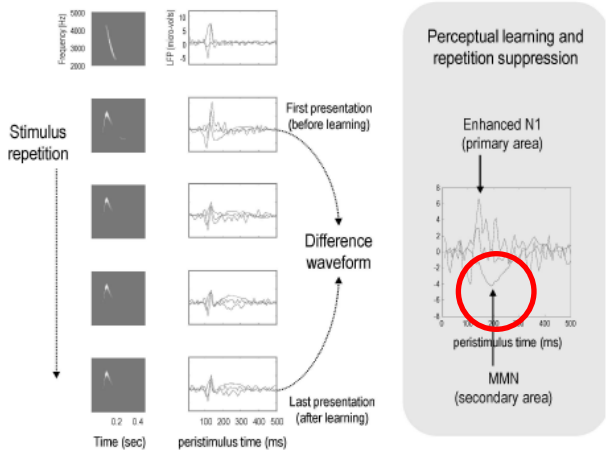


Fig. 6: A demonstration of perceptual learning. This figure shows the results of a simulated roving oddball paradigm, in which a stimulus is changed sporadically to elicit an oddball (i.e., deviant) response. The stimuli used here are chirps of the same sort as those used in Fig. 4. Left panels: The left column shows the percepts elicited in sonogram format. These are simply the predictions of sensory input, based on their inferred causes (i.e., the expectations about hidden states). The right column shows the evolution of prediction error at the first (dotted lines) and second (solid line) levels of a simple linear convolution model (in which a causal state produces time-dependent amplitude and frequency modulations). The results are shown for one learned chirp (top graph) and the first four responses to a new chirp (lower graphs). The new chirp was generated by changing the parameters of the underlying equations of motion. It can be seen that following the first oddball stimulus, the prediction errors show repetition suppression (i.e., the amplitudes of the traces get smaller). This is due to learning the model parameters over trials (see synaptic plasticity and gain in Fig. 3). Of particular interest is the difference in responses to the first and last presentations of the new stimulus: these correspond to the deviant and standard responses, respectively. Right panel: This shows the difference between standard and oddball responses, with an enhanced negativity at the first level early in peristimulus time (dotted lines for inferred amplitude and frequency), and a later negativity at the higher or second level (solid line for the causal state). These differences could correspond to phenomena like enhanced N1 effects and the mismatch negativity (MMN) found in empirical difference waveforms. Note that superficial pyramidal cells (see Fig. 3) dominate event related potentials and that these cells may encode prediction error[47,146].

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

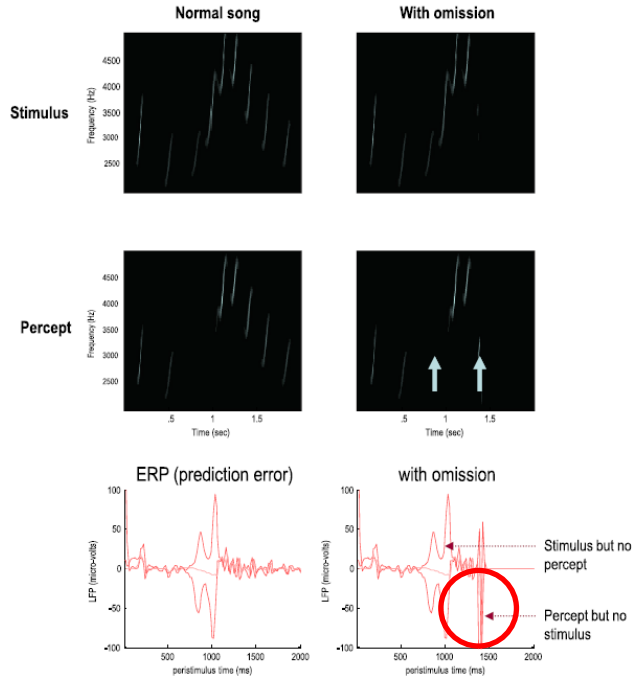


Fig. 5 – Omission-related responses: The left panels show the original song and responses evoked. The right panels show the equivalent responses to omission of the last chirps. The top panels show the stimulus and the middle panels the corresponding percept in sonogram format. The interesting thing to note here is the occurrence of an anomalous percept after termination of the song on the lower right. This corresponds roughly to the chirp that would have been perceived in the absence of omission. The lower panels show the corresponding (precision weighted) prediction error under the two stimuli at both levels. These show a burst of prediction error when a stimulus is missed and at the point that the stimulus is omitted (at times indicated by the arrows on the sonogram). The solid lines correspond to sensory prediction error and the broken lines correspond to extrasensory prediction error at the second level of the generative model.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

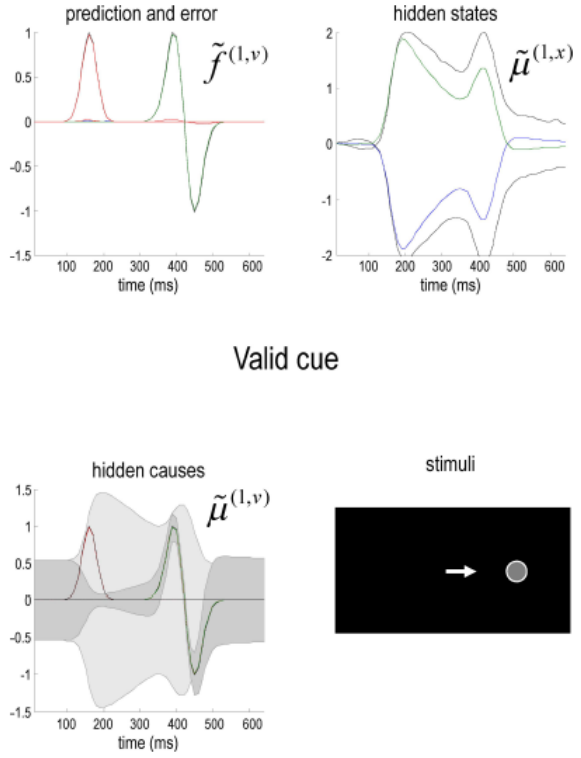
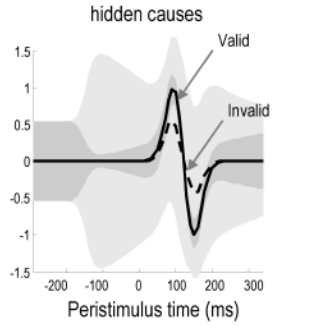


FIGURE 2 | Simulation of the Posner task (validly cued target). Upper left panel: the time-dependent expression of the cue and target stimuli are shown as broken gray lines, while the respective predictions are in red s_c and green s_t respectively. The dotted red lines show the prediction error and reflect the small amount of noise we used in these simulations. Lower left panel: the ensuing conditional expectations of the underlying hidden causes $v_c^{(1)}, v_t^{(1)}, v_t^{(1)}$ are shown below. The gray areas correspond to 90% conditional confidence tubes; this confidence reflects the estimated precision of the sensory data, which is encoded by the expectations of the hidden states in the upper right panel. The green line corresponds to a precision or attentional bias to the right $x_r^{(1)}$ and the blue line to the left $x_l^{(1)}$. They gray lines are the true precisions. Lower right panel: this insert indicates the sort of stimuli that would be generated by these hidden causes.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics



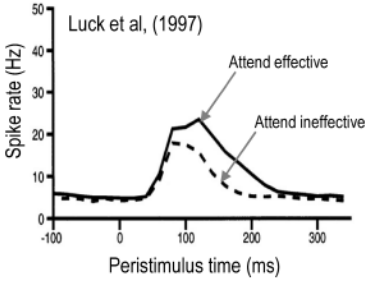


FIGURE 8 |This figure demonstrates how generalized predictive coding reproduces some quantitative aspects of biased competition. The simulation (upper panel) reproduces the conditional expectations in the previous figure about valid (solid line) and invalid (dashed line) targets, when presented simultaneously. These two responses resemble those reported in Luck et al. (1997). Lower panel: peristimulus histograms (over 20 ms bins) redrawn from Luck et al. (1997), following simultaneous presentation of two (effective and ineffective) stimuli averaged over 29 V4 neurons that showed a significant attention effect. The solid line reports trials in which attention was directed to the effective stimulus (cf, responses to a valid target) and the dashed line when attention was directed to the ineffective stimulus (cf, responses to an invalid target). Note that the empirical data are non-negative spike counts, whereas the simulated activity represent firing rate deviations around baseline levels.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

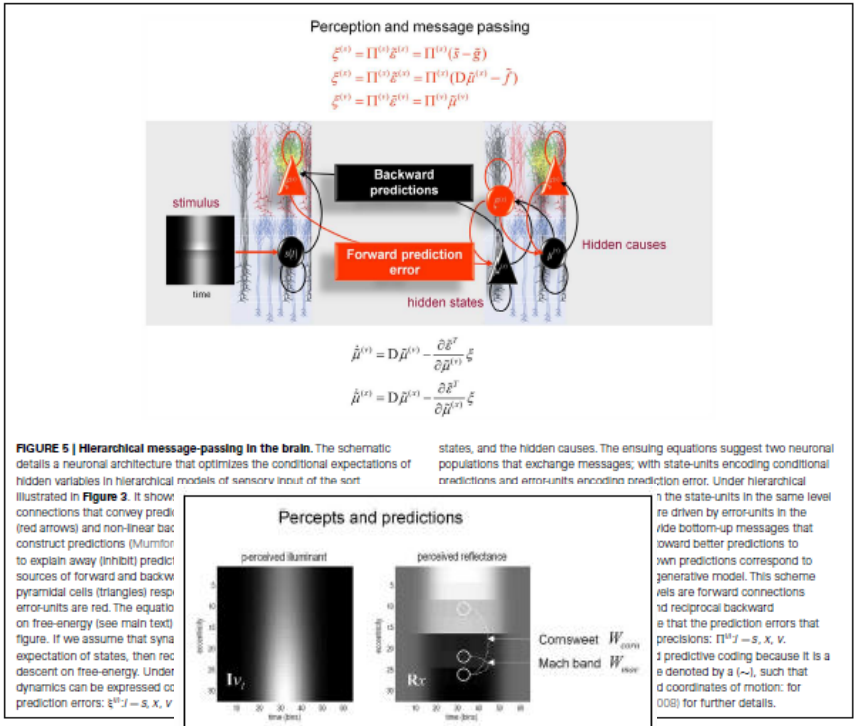


FIGURE 5 | Hierarchical message-passing in the brain. The schematic details a neuronal architecture that optimizes the conditional expectations of hidden variables in hierarchical models of sensory input of the sort illustrated in **Figure 3**. It shows connections that convey predi (red arrows) and non-linear bac construct predictions (Mumfor to explain away (inhibit) predict sources of forward and backw pyramidal cells (triangles) resp error-units are red. The equatio on free-energy (see main text) figure. If we assume that syna expectation of states, then rec descent on free-energy. Under dynamics can be expressed co prediction errors: $\xi^{(v)}: I = s, x, v$ states, and the hidden causes. The ensuing equations suggest two neuronal populations that exchange messages; with state-units encoding conditional predictions and error-units encoding prediction error. Under hierarchical n the state-units in the same level re driven by error-units in the ide bottom-up messages that oward better predictions to wn predictions correspond to generative model. This scheme els are forward connections d reciprocal backward e that the prediction errors that precisions: $\Pi^{(v)}: I = s, x, v$. predictive coding because it is a e denoted by a (~), such that d coordinates of motion: for 008) for further details.

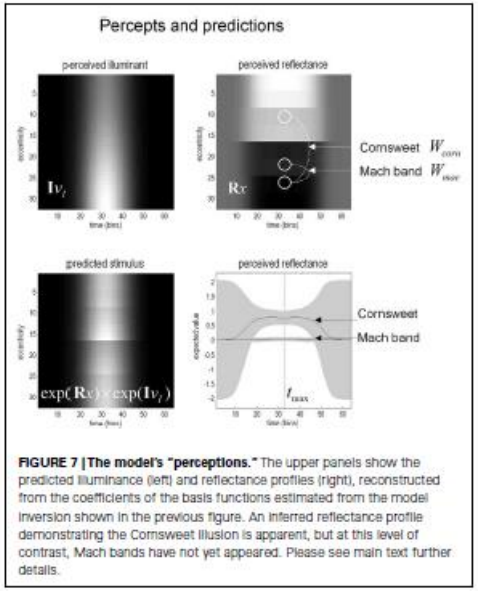


FIGURE 7 | The model's "perceptions." The upper panels show the predicted illuminance (left) and reflectance profiles (right), reconstructed from the coefficients of the basis functions estimated from the model inversion shown in the previous figure. An inferred reflectance profile demonstrating the Cornsweet illusion is apparent, but at this level of contrast, Mach bands have not yet appeared. Please see main text further details.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

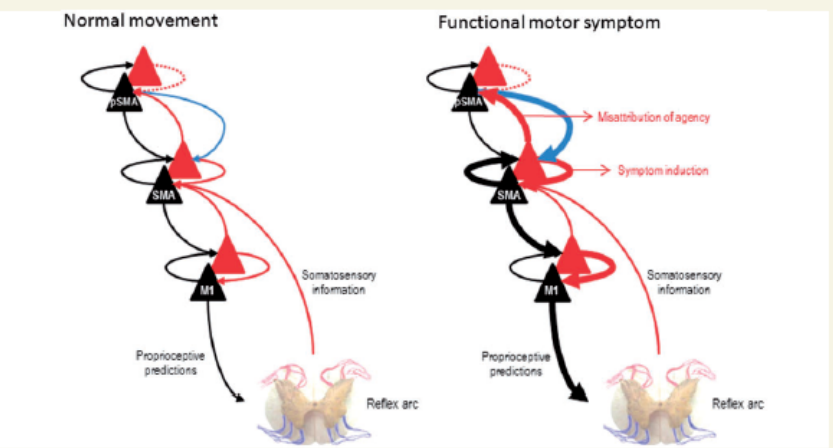


Figure 4 This schematic illustrates the hierarchical anatomy we presume underlies false inference in patients with functional motor symptoms (both weakness and 'positive' phenomena such as tremor). In normal movement, we propose that predictions regarding the sensory consequences of intended movement arise at a high hierarchical level (here pre-supplementary motor area) and are propagated down the motor hierarchy, producing a proprioceptive prediction error (peripherally) that is fulfilled by movement. In functional motor symptoms we propose that an abnormal prior expectation related to the dynamics/scaling of movement is formed within an intermediate motor area (here the supplementary motor area). This prior is afforded abnormal precision by attentional processes (thick blue arrow) that cause intermediate level motor predictions (thick black arrow) to elicit movement and prediction errors (thick red arrow) to report the unpredicted content of that movement to higher cortical areas (here, pre-supplementary motor area). The secondary consequence of these prediction errors is that prefrontal regions will try to explain them away in terms of a symptomatic interpretation or misattribution of agency to external causes; in short, a failure to realize the movement was intended. Forward connections convey prediction error (red), backward connections convey predictions (black) and descending attentional modulatory connections (blue). pSMA = pre-supplementary motor area; M1 = primary motor cortex; SMA = supplementary motor area.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

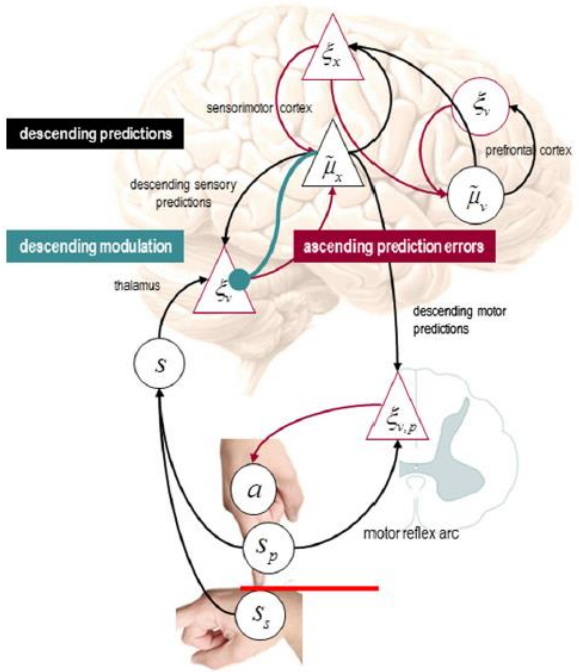


Fig. 2 Functional anatomy: Speculative mapping of Eq. (3) onto neuroanatomy. Somatosensory and proprioceptive prediction errors are generated by the thalamus, while conditional expectations and prediction errors about hidden states (*circles*) (the forces) are placed in sensorimotor cortex. The expectations and prediction errors about the hidden causes of forces (*triangles*) have been placed in the prefrontal cortex. In active inference, proprioceptive predictions descend to the spinal cord and elicit output from alpha motor neurons (playing the role of proprioceptive prediction error units) via a classical reflex arc. Red connections originate from prediction error units (ξ cells) and can be regarded as intrinsic connections or ascending (forward) extrinsic connections from superficial principal cells. Conversely, the *black connections* represent intrinsic connections and descending (backward) efferents from (deep) principal cells encoding conditional expectations ($\tilde{\mu}$ cells). The *cyan connections* denote descending neuromodulatory effects that mediate sensory attenuation. The crucial point to take from this schematic is that conditional expectations of sensory states (encoded in the pyramidal cell $\tilde{\mu}_v$) can either be fulfilled by descending proprioceptive predictions (that recruit classical reflex arcs), or they can be corrected by ascending sensory prediction errors. In order for descending motor efferents to prevail, the precision of the sensory prediction errors must be attenuated

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

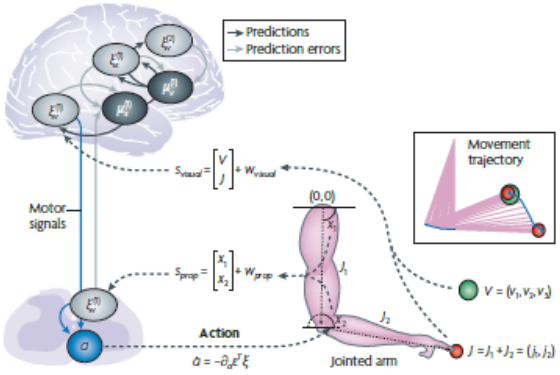


Figure 2 | **A demonstration of cued reaching movements.** The lower right part of the figure shows a motor plant, comprising a two-jointed arm with two hidden states, each of which corresponds to a particular angular position of the two joints; the current position of the finger (red circle) is the sum of the vectors describing the location of each joint. Here, causal states in the world are the position and brightness of the target (green circle). The arm obeys Newtonian mechanics, specified in terms of angular inertia and friction. The left part of the figure illustrates that the brain senses hidden states directly in terms of proprioceptive input (S_{prop}) that signals the angular positions (x_1,x_2) of the joints and indirectly through seeing the location of the finger in space (J_1,J_2). In addition, through visual input (S_{visual}) the agent senses the target location (v_1,v_2) and brightness (v_3). Sensory prediction errors are passed to higher brain levels to optimize the conditional expectations of hidden states (that is, the angular position of the joints) and causal (that is, target) states. The ensuing predictions are sent back to suppress sensory prediction errors. At the same time, sensory prediction errors are also trying to suppress themselves by changing sensory input through action. The grey and black lines denote reciprocal message passing among neuronal populations that encode prediction error and conditional expectations; this architecture is the same as that depicted in BOX 2. The blue lines represent descending motor control signals from sensory prediction-error units. The agent's generative model included priors on the motion of hidden states that effectively engage an invisible elastic band between the finger and target (when the target is illuminated). This induces a prior expectation that the finger will be drawn to the target, when cued appropriately. The insert shows the ensuing movement trajectory caused by action. The red circles indicate the initial and final positions of the finger, which reaches the target (green circle) quickly and smoothly; the blue line is the simulated trajectory.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

Cell PRESS

Neuron
Perspective

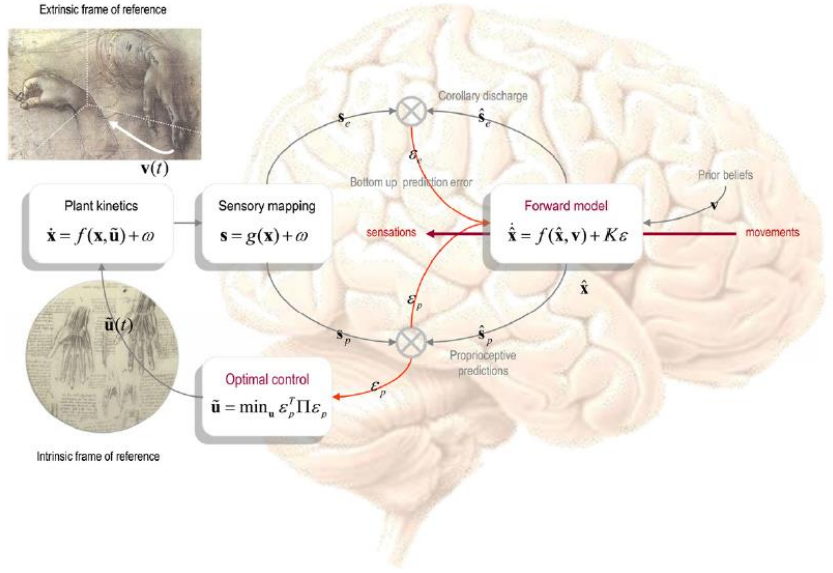


Figure 3. Active Inference
This figure represents the final simplification of the predictive coding scheme of the previous figure. Here, cost functions have been replaced by prior beliefs about (desired) trajectories in an extrinsic frame of reference. These beliefs enter the Bayesian filter to guide predictions of sensory inputs. Proprioceptive predictions are fulfilled in the periphery through classical motor reflex arcs, while predictions of exteroceptive inputs correspond to corollary discharge and are an integral part of perceptual inference. Note that optimal control now reduces to simply suppressing proprioceptive prediction errors. This is active inference.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

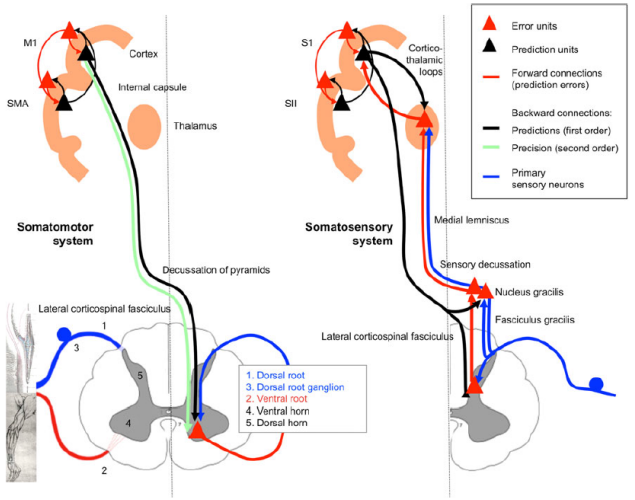


Fig. 9 Somatomotor and somatosensory connections in active inference: In this figure, we have focused on monosynaptic reflex arcs and have therefore treated alpha motor neurons as prediction error units. In this scheme, descending (corticospinal) proprioceptive predictions (from upper motor neurons in M1) and (primary sensory) proprioceptive afferents from muscle spindles converge on alpha motor neurones on the ventral horn of the spinal cord. The comparison of these signals generates a prediction error. The gain of this prediction error is in part dependent upon descending predictions of its precision (for further explanation see 'CM neurons and predictions of precision' in the "Discussion"). The associated alpha motor neuron discharges elicit (extrafusal) muscle fibre contractions until prediction error is suppressed. Ascending proprioceptive and somatosensory information does not become a prediction error until it encounters descending predictions, whether in the (ventral posterior nucleus of the) thalamus, the dorsal column nuclei, or much earlier in the dorsal horn. In the cortex, error units at a given level receive predictions from that level and the level above, and project to prediction units at that level and the level above (only two levels are shown). In this way, discrepancies between actual and predicted inputs—resulting in prediction errors—can either be resolved at that level or passed further up the hierarchy (Friston et al. 2006). Prediction units project to error units at their level and the level below, attempting to explain away their activity. Crucially, active inference suggests that both proprioceptive (motor) and somatosensory systems use a similar architecture. It is generally thought that prediction units correspond to principal cells in infragranular layers (deep pyramidal cells) that are the origin of backward connections; while prediction error units are principal cells in supragranular layers (superficial pyramidal cells) that elaborate forward projections (Mumford 1992; Friston and Kiebel 2009). Note that we have implicitly duplicated proprioceptive prediction errors at the spinal (somatomotor) and thalamic (somatosensory) levels. This is because the gain of central (somatosensory) principal units encoding prediction error is set by neuromodulation (e.g. synchronous gain or dopamine), while the gain of peripheral (somatomotor) prediction error units is set by NMDA-Rs and gamma motor neuron activity. In predictive coding, this gain encodes the precision (inverse variance) of prediction errors (see Feldman and Friston 2010). Algorithmically, the duplication of prediction errors reflects the fact that somatomotor prediction errors drive action, while somatosensory prediction errors drive (Bayes-optimal) predictions. For reasons of clarity we have omitted connections ascending the cord in the somatomotor system, e.g. spinal projections to M1 and the transcortical reflex pathway from S1 (in particular the proprioceptive area 3a) to M1: these are described in the "Discussion"

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

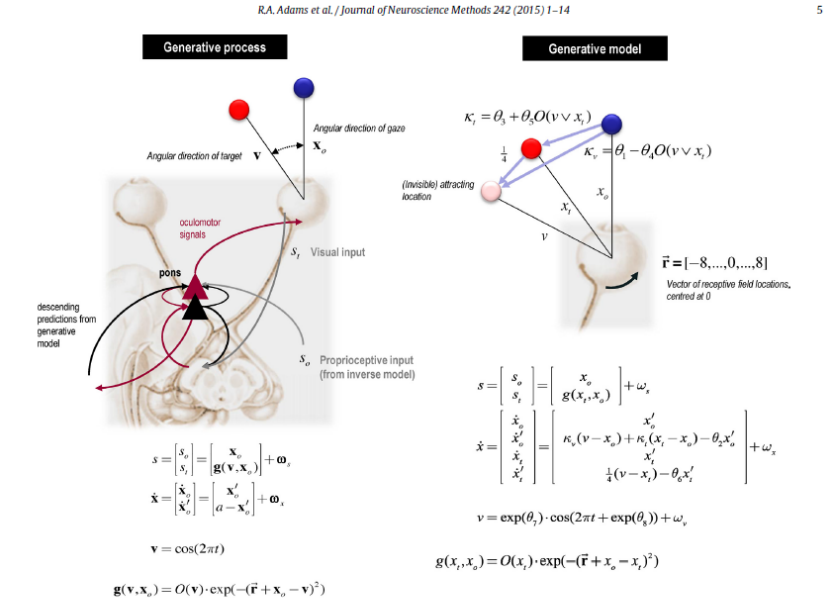


Fig. 2. Generative process and model of oculomotor pursuit movements. This schematic illustrates the process (left panel) and generative model of that process (right panel) used to simulate Bayes optimal pursuit. The graphics on the left show a putative predictive coding scheme (with superficial pyramidal cells in red and deep pyramidal cells in black in the pontine nuclei) processing proprioceptive information during oculomotor pursuit. These cells receive proprioceptive information from an inverse model in the subcortical oculomotor system and respond reflexively to minimise proprioceptive prediction error through action. This prediction error rests on descending predictions from the generative model on the right. The actual movement of the target is determined by a hidden cause (target location), which determines the visual input for any given direction of gaze. The generative model entails beliefs about how the target and eyes move. In brief, this model includes an invisible location that attracts the target, causing it to move. Crucially, the agent believes that its centre of gaze is attracted to this location (and the target), where the forces of attraction may (or may not) depend upon occlusion of the target and its attracting location. These forces of attraction are illustrated with lilac arrows in the top right; the arrows are labelled with their respective multipliers from the equations directly below. Please see main text for a description of the variables in the equations describing the motion of hidden states and how they depend upon hidden causes. Note that real states that are hidden from observation in the real world are in bold, whereas the hidden states assumed by the generative model are in italics. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

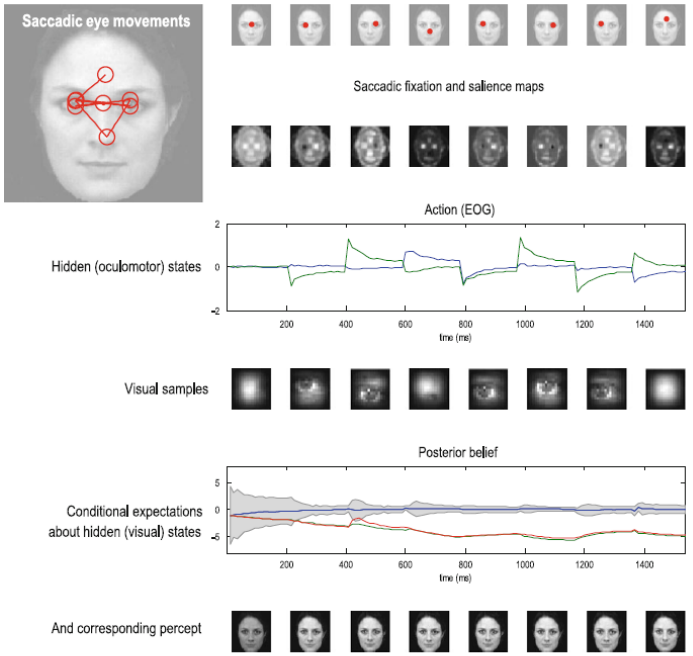


Fig. 3 This figure shows the results of simulations in which a face was presented to an agent, whose responses were simulated using the active inference scheme described in the main text. In this simulation, the agent had three internal images or hypotheses about the stimuli it might sample (an upright face, an inverted face and a rotated face). The agent was presented with an upright face and its posterior expectations were evaluated over 16 (12 ms) time bins, until the next saccade was emitted. This was repeated for eight saccades. The ensuing eye movements are shown as *red dots* at the location (in extrinsic coordinates) at the end of each saccade in the upper row. The corresponding sequence of eye movements is shown in the insert on the *upper left*, where the *red circles* correspond roughly to the proportion of the image sampled. These saccades are driven by prior beliefs about the direction of gaze—based upon the saliency maps in the second row. Note that these maps change with successive saccades as posterior beliefs about the hidden states, including the stimulus, become progressively more confident. Note also that salience is depleted in locations that were foveated in the previous saccade. This reflects an inhibition of return that was built into the prior beliefs. The resulting posterior beliefs provide both visual and proprioceptive predictions that suppress visual prediction errors and drive eye movements, respectively. Oculomotor responses are shown in the third row in terms of the two hidden oculomotor states corresponding to vertical and horizontal displacements. The associated portions of the image sampled (at the end of each saccade) are shown in the fourth row. The final two rows show the posterior beliefs and inferred stimulus categories, respectively. The posterior beliefs are plotted in terms of posterior expectations and the 90 % confidence interval about the true stimulus. The key thing to note here is that the expectation about the true stimulus supervenes over its competing expectations and—as a result—posterior confidence about the stimulus category increases (the confidence intervals shrink to the expectation). This illustrates the nature of evidence accumulation when selecting a hypothesis or percept the best explains sensory data. Within-saccade accumulation is evident even during the initial fixation with further stepwise decreases in uncertainty as salient information is sampled by successive saccades

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

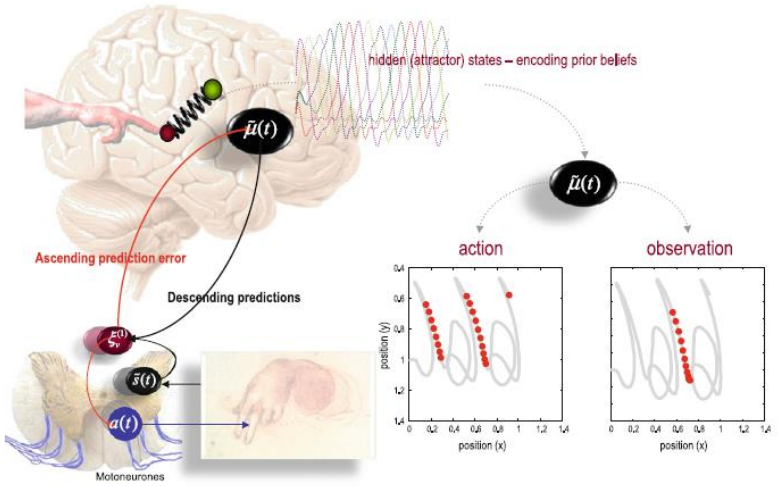


Fig. 2 This schematic summarizes the results of the simulations of action observation reported in Friston et al. (2011). The *left panel* pictures the brain as a forward or generative model of itinerant movement trajectories (based on a Lotka-Volterra attractor, whose states are shown as a function of time in *coloured lines*). This model furnishes predictions about visual and proprioceptive inputs, which prescribe movement through reflex arcs at the level of the spinal cord (insert on the *lower left*). The variables have the same meaning as in the previous figure. The mapping between attractor dynamics and proprioceptive consequences is modelled with Newtonian mechanics on a two jointed arm, whose extremity (*red ball*) is drawn to a target location (*green ball*) by an imaginary spring. The location of the target is prescribed (in an extrinsic frame of reference) by the currently active state in the attractor. These attractor dynamics and the mapping to an extrinsic (movement) frame of reference constitute the agent's prior beliefs. The ensuing posterior beliefs are entrained by visual and proprioceptive sensations by prediction errors during the process of inference, as summarized in the previous figure. The resulting sequence of movements was configured to resemble handwriting and is shown as a function of location over time on the *lower right* (as *thick grey lines*). The *red dots* on these trajectories signify when a particular neuron or neuronal population encoding one of the hidden attractor states was active during *action* (*left panel*) and *observation* of the same action (*right panel*): More precisely, the dots indicate when responses exceeded half the maximum activity and are shown as a function of limb position. The *left panel* shows the responses during action and illustrates both a place-cell-like selectivity and directional selectivity for movement in an extrinsic frame of reference. The equivalent results on the right were obtained by presenting the same visual information to the agent but removing proprioceptive sensations. This can be considered as a simulation of action observation and mirror neuron-like activity

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

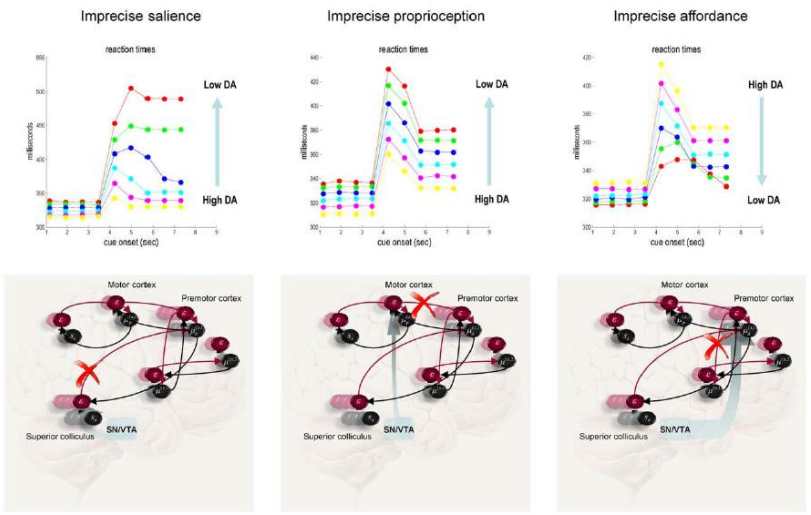


Figure 10. This figure represents behavioral results in terms of reaction times for depleting dopamine in three regions: the superior colliculus encoding sensory salience (as in previous figure), the motor cortex encoding proprioception (middle column) and the premotor cortex encoding affordance (right column). These results are shown using the same format as in previous figure and illustrate the qualitatively different effects of dopamine depletion in different parts of the brain (or model). The lower panels indicate the implicit projections, from the substantia nigra or ventral tegmental area, have been selectively depleted (where a red cross highlights the forward prediction errors affected). The key thing to take from these simulations is that reducing the precision of prediction errors on sensory salience induces bradykinesia and perseveration; whereas the equivalent reduction in proprioceptive affordance causes bradykinesia without perseveration. Finally, compromising the precision of changes in affordance increases perseveration and decreases bradykinesia.
doi:10.1371/journal.pcbi.1002327.g010

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

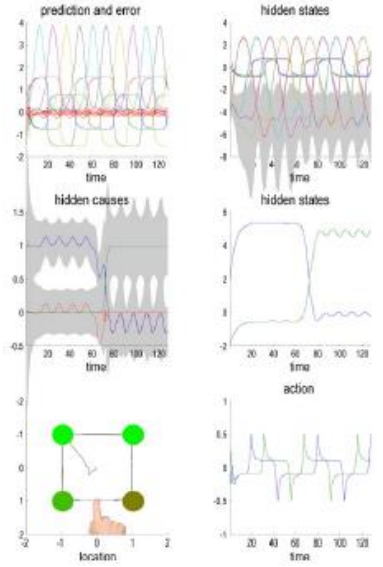


Figure 5. This figure summarizes the results of simulations under normal levels of dopamine (using a log precision of four for all prediction errors). The conditional predictions and expectations are shown as functions of time over 128 time bins, each modeling 64 ms of time. The upper left panel shows the conditional predictions (colored lines) and prediction errors (red lines) based upon the expected in states on the upper right. In this panel and throughout, the grey areas denote 90% Bayesian confidence intervals. The inferred speed of itinerant cycling among affordance states corresponds to the first of the hidden causes at the second level (left middle panel). These hidden causes are a softmax function of their associated hidden states (right middle panel). The blue lines encode a sequential context, while the green lines encode the converse (random) context. The switching in these conditional expectations occurs after sufficient sensory evidence has accumulated following a reversal of the presentation order. The lower left panel shows the trajectory (dotted lines) in an extrinsic frame of reference, in relation to the cue locations (green circles), while the lower right panel shows action in terms of horizontal and vertical angular forces causing these movements.
doi:10.1371/journal.pcbi.1002327.g005

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

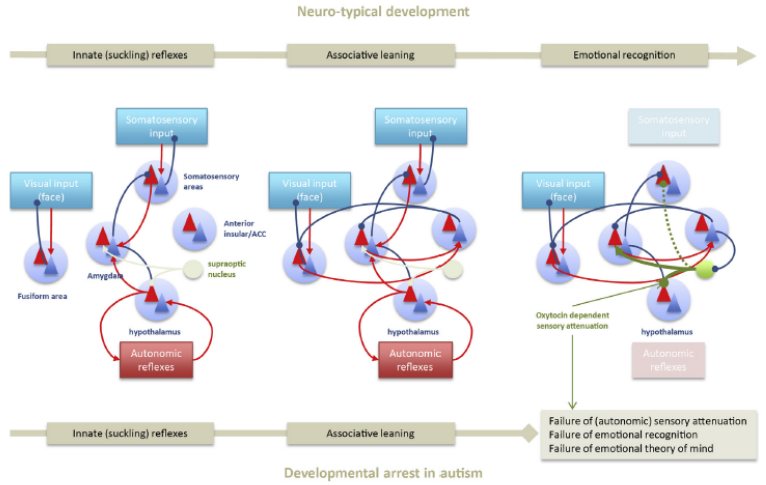


Fig. 3. *Oxytocin and the development of emotional affordance.* This schematic describes normal (and autistic) neurodevelopmental trajectories, in terms of (simplified) neural architectures underlying predictive coding of autonomic (emotional) signals. The three panels illustrate the development of associative connections we imagine underlie the acquisition of emotional responses during three stages of development. The anatomical designations should not be taken too seriously—they are just used to illustrate how predictive coding can be mapped onto neuronal systems. In all of these schematics, red triangles correspond to neuronal populations (superficial pyramidal cells) encoding prediction error, while blue triangles represent populations (deep pyramidal cells) encoding expectations. These populations provide descending predictions to prediction error populations in lower hierarchical levels (blue lines). The prediction error populations then reciprocate ascending prediction errors to adjust the expectations (red lines). Arrows denote excitatory connections, while circles denote inhibitory effects (mediated by inhibitory interneurons). **Left panel**: in the first panel, connections are in place to mediate innate (epigenetically specified) reflexes – such as the suckling reflex – that elicit autonomic (e.g., vasovagal) reflexes in response to appropriate somatosensory input. These reflexes depend upon high-level representations predicting both the somatosensory input and interoceptive consequences. The representations are activated by somatosensory prediction errors and send interoceptive predictions to the hypothalamic area--to elicit interoceptive prediction errors that are resolved in the periphery by autonomic reflexes. Oxytocin is shown to project to the high-level representations (the amygdala) and the hypothalamic area, to modulate the gain or precision of prediction error units. In this schematic, its effects are twofold: oxytocin attenuates the gain of hypothalamic prediction error units and augments the gain of higher level units. **Middle panel**: this shows the architecture after associative learning, during which high-level representations in the anterior cingulate or insular cortex have learned the coactivation of amygdala representations and exteroceptive cues (e.g., the mother's face during suckling). These high-level representations now predict the exteroceptive visual input and (through the amygdala) somatosensory and autonomic consequences. **Right panel**: in this schematic, visual input (e.g., the mother's face) is recognized using the high-level representation in the anterior insular or cingulate cortex. However, in this case, interoceptive prediction error is attenuated so that it does not elicit an autonomic response. In other words, although the high-level emotional representation is used to recognize exteroceptive cues, lower-level transcortical reflexes are inhibited. In autism, we presume that oxytocin is deficient, such that sensory attenuation is impaired – leading to disinhibition of autonomic responses and the failure to recognize a mother's face in any other context – other than during suckling. This failure of sensory attenuation may underlie autonomic hypersensitivity, failure of emotional recognition, attention to emotional cues, theory of mind and central coherence. The dotted green line in this figure acknowledges that there may not be any direct projections from the origin of oxytocin cells (in the supraoptic and paraventricular nuclei of the hypothalamus) to secondary or primary somatosensory cortex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

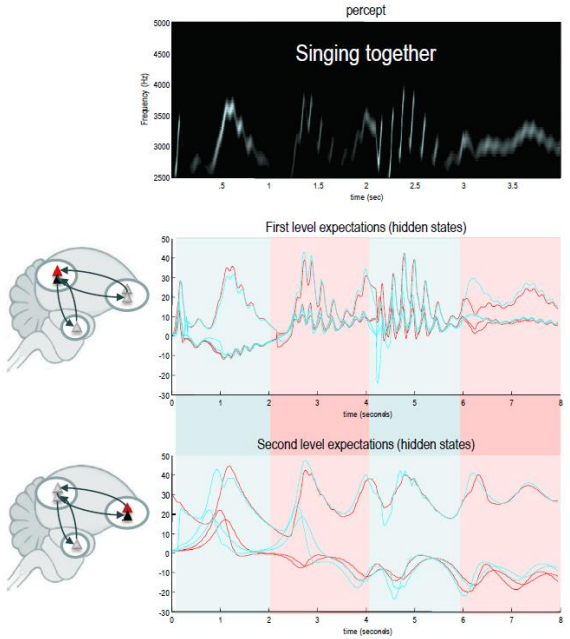


Fig. 7 – Communication and generalised synchrony. This figure uses the same format as Fig. 6; however, here, we have juxtaposed the two birds so that they can hear each other. In this instance, the posterior expectations show identical synchrony at both the sensory and extrasensory hierarchical levels – as shown in the middle and lower panels respectively. Note that the sonogram is continuous over successive 2 sec epochs – being generated alternately by the first and second bird.

What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

❖Specify a deep (hierarchical) generative model

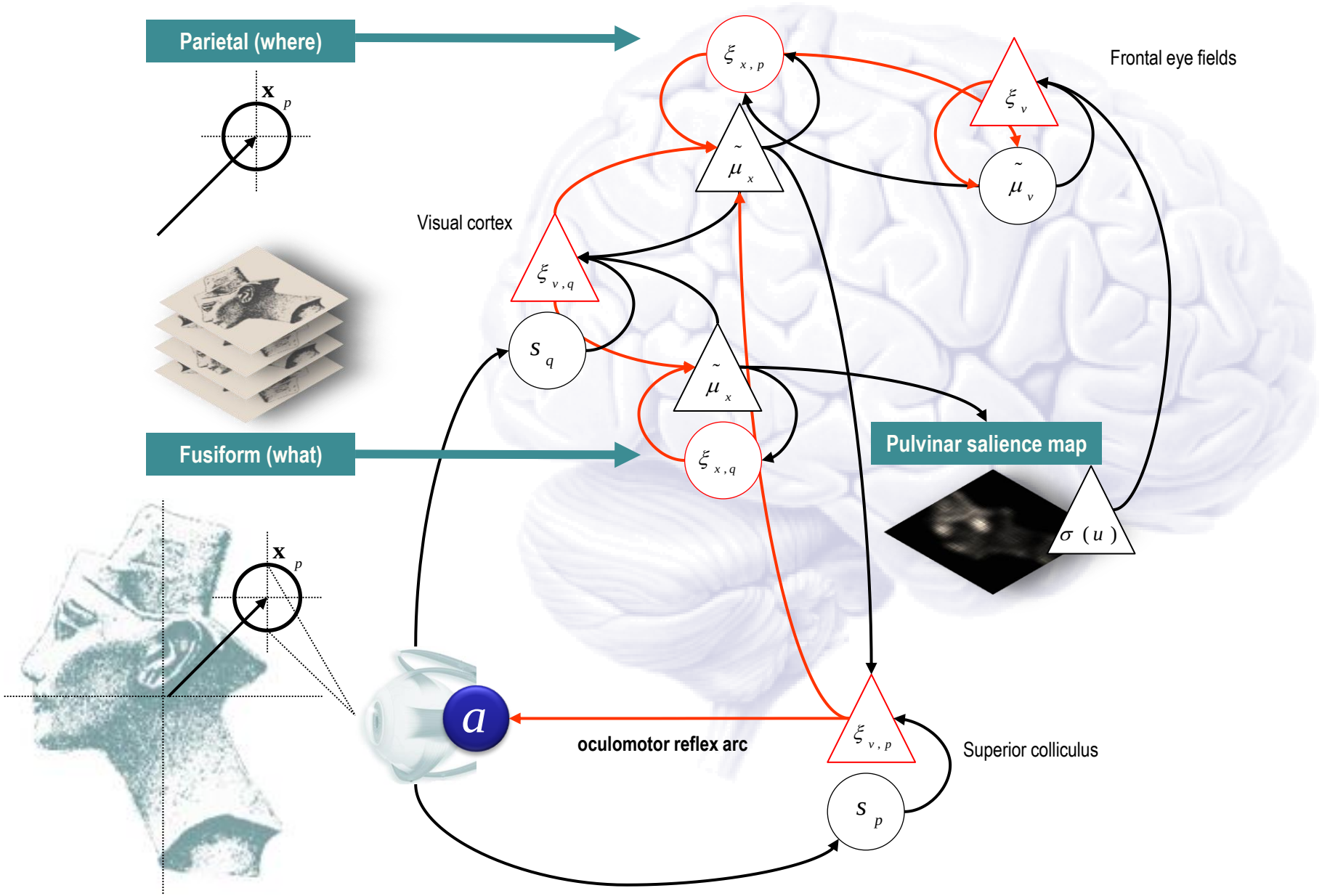
❖Simulate approximate (active) Bayesian inference by solving:

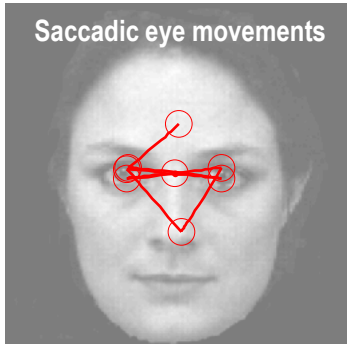
$$\dot{\tilde{\mu}} = D\tilde{\mu} - \Gamma\nabla F(\tilde{s}, \tilde{\mu})$$

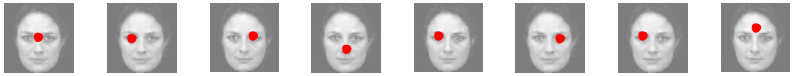
prediction update

An example: Visual searches and saccades

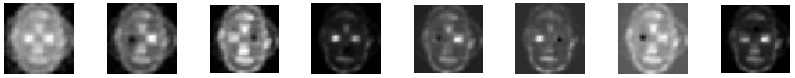
Deep (hierarchical) generative model

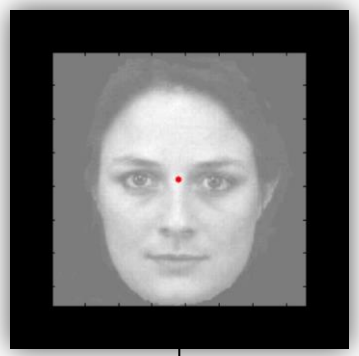




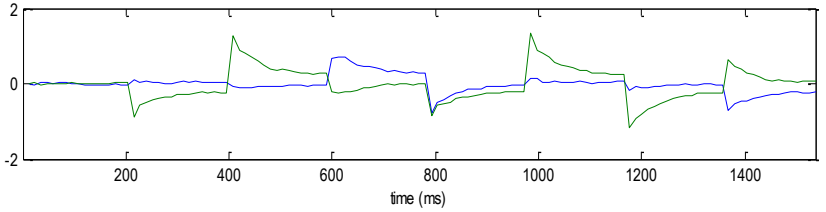


Saccadic fixation and salience maps

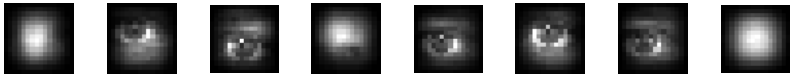




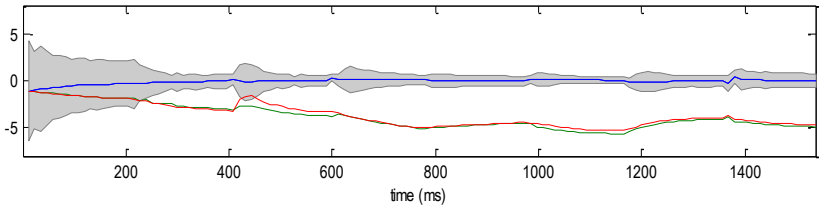
Hidden (oculomotor) states



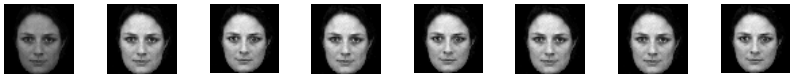
Visual samples



Conditional expectations about hidden (visual) states



And corresponding percept

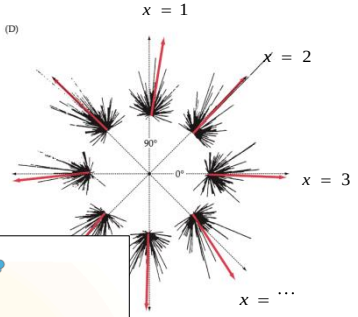


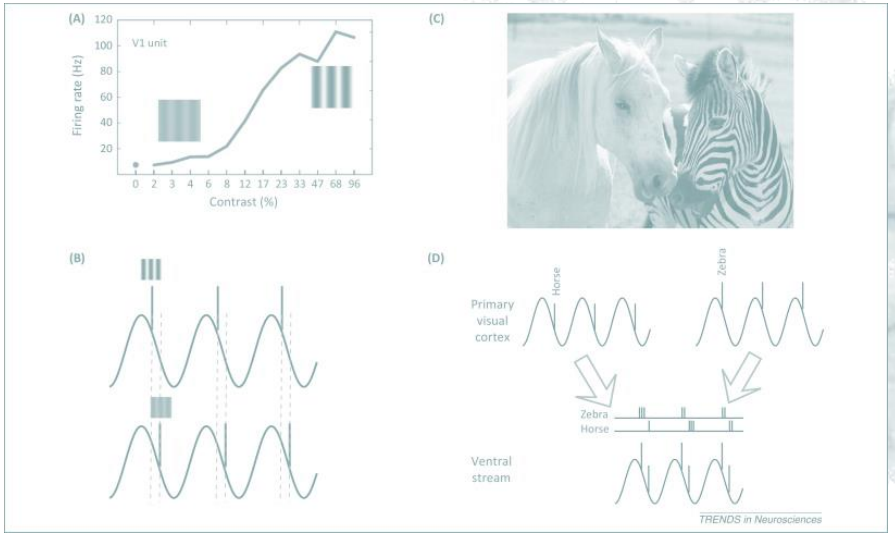
Is Bayesian filtering the only process theory for approximate Bayesian inference in the brain?

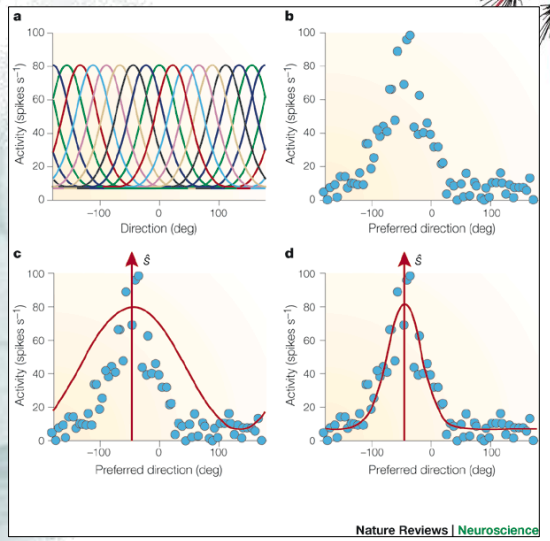
What about state space models?

$$x \in \{1, \cdots, N\}$$

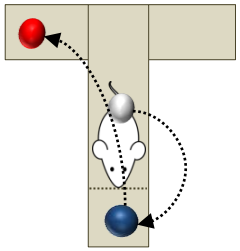
$$\mu \in [0,1]$$







$$q(x \mid \mu) \approx p(x \mid s, m)$$



A (Markovian) generative model

$$P(\tilde{o}, \tilde{s}, \tilde{u}, \gamma \mid \tilde{a}, m) = P(\tilde{o} \mid \tilde{s})\, P(\tilde{s} \mid \tilde{a})\, P(\tilde{u} \mid \gamma)\, P(\gamma \mid m)$$

$$P(\tilde{o} \mid \tilde{s}) = P(o_0 \mid s_0) P(o_1 \mid s_1) \cdots P(o_t \mid s_t)$$

$$P(o_t \mid s_t) = \mathbf{A}$$

Likelihood

$$P(\tilde{s} \mid \tilde{a}) = P(s_t \mid s_{t-1}, a_t) \cdots P(s_1 \mid s_0, a_1) P(s_0 \mid m)$$

$$P(s_{t+1} \mid s_t, u_t) = \mathbf{B}(u_t)$$

Empirical priors – hidden states

$$P(\tilde{u} = \pi \mid \gamma) = \sigma(-\gamma \cdot \hat{\mathbf{F}})$$

– control states

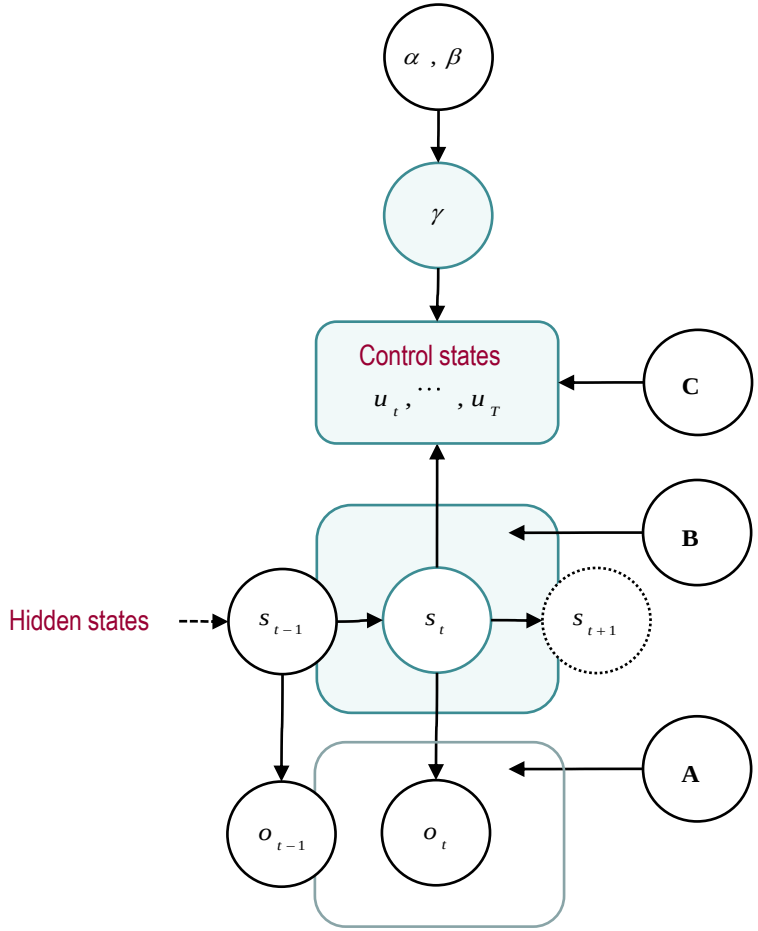
$$\hat{\mathbf{F}} = \sum_{\tau} E_{Q(o_\tau, s_\tau \mid \pi)}[\ln P(o_\tau, s_\tau \mid \pi) - \ln Q(s_\tau \mid \pi)]$$

$$P(o_\tau \mid m) = \mathbf{C}$$

$$P(s_0 \mid m) = \mathbf{D}$$

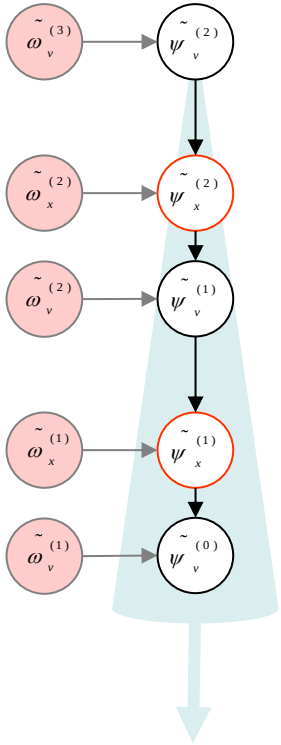
$$P(\gamma \mid m) = \Gamma(\alpha, \beta)$$

Full priors

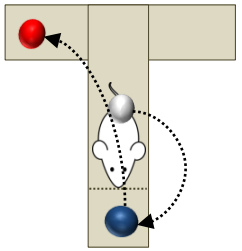


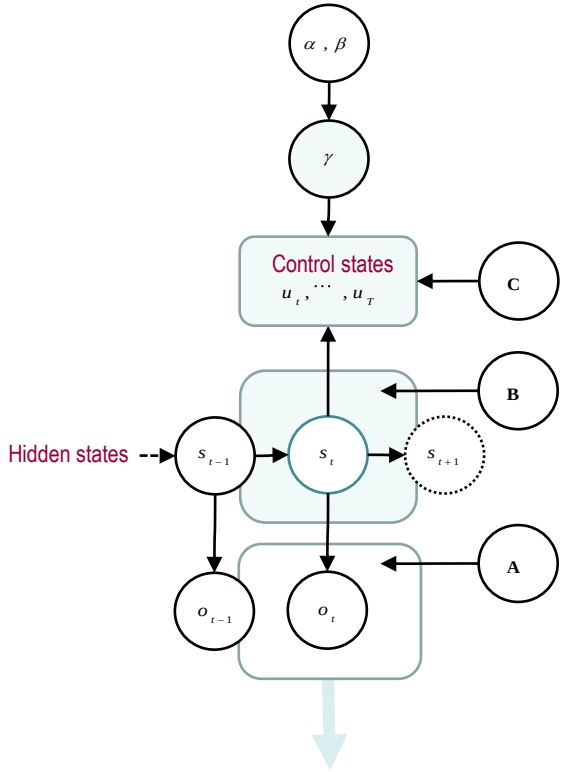
Generative models

Continuous states



Discrete states





Bayesian filtering (predictive coding)

Variational Bayes (belief updating)

$$\dot{\tilde{\mu}}^{(i)} = D\,\tilde{\mu}^{(i)} - \nabla\,\tilde{\varepsilon}^{(i)} \cdot \Pi^{(i)} \cdot \tilde{\varepsilon}^{(i)} \xrightarrow{\dot{\tilde{\mu}}^{(i)} = 0} \mathbf{s}_t^{\pi} = \sigma\left(\mathbf{A} \cdot o_t + \mathbf{B}_{t-1}^{\pi}\mathbf{s}_{t-1}^{\pi} + \mathbf{B}_t^{\pi} \cdot \mathbf{s}_{t+1}^{\pi}\right)$$

Variational updates

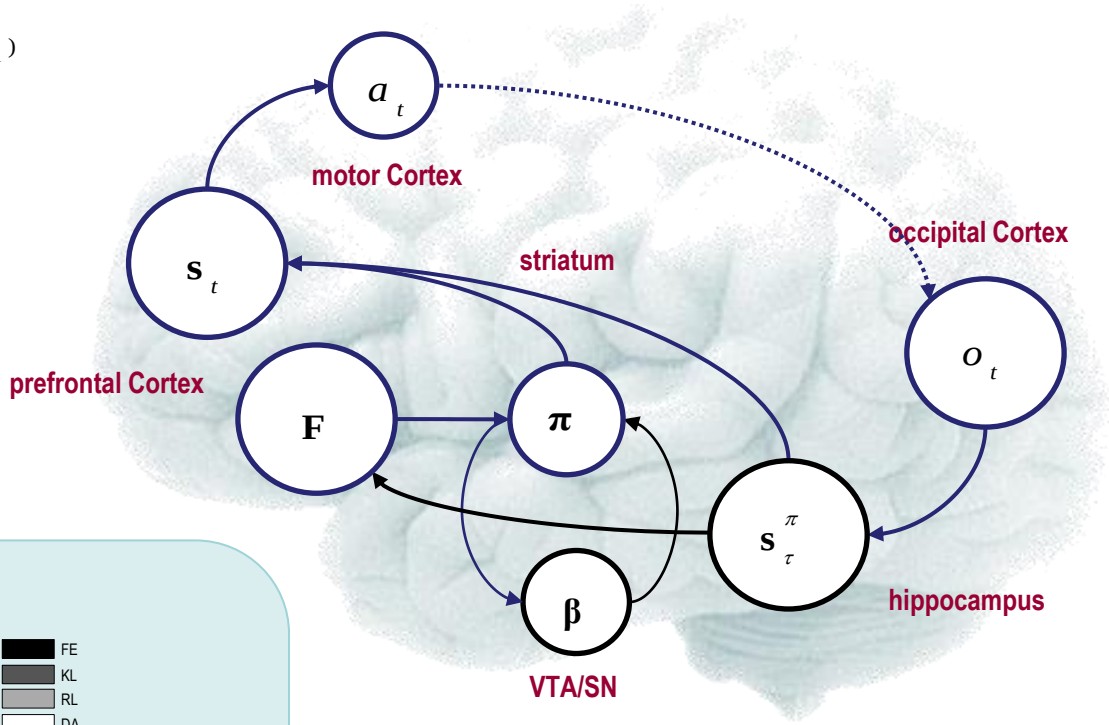
Perception $\mathbf{s}_t^{\pi} = \sigma(\mathbf{A} \cdot o_t + \mathbf{B}_{t-1}^{\pi}\mathbf{s}_{t-1}^{\pi} + \mathbf{B}_t^{\pi} \cdot \mathbf{s}_{t+1}^{\pi})$

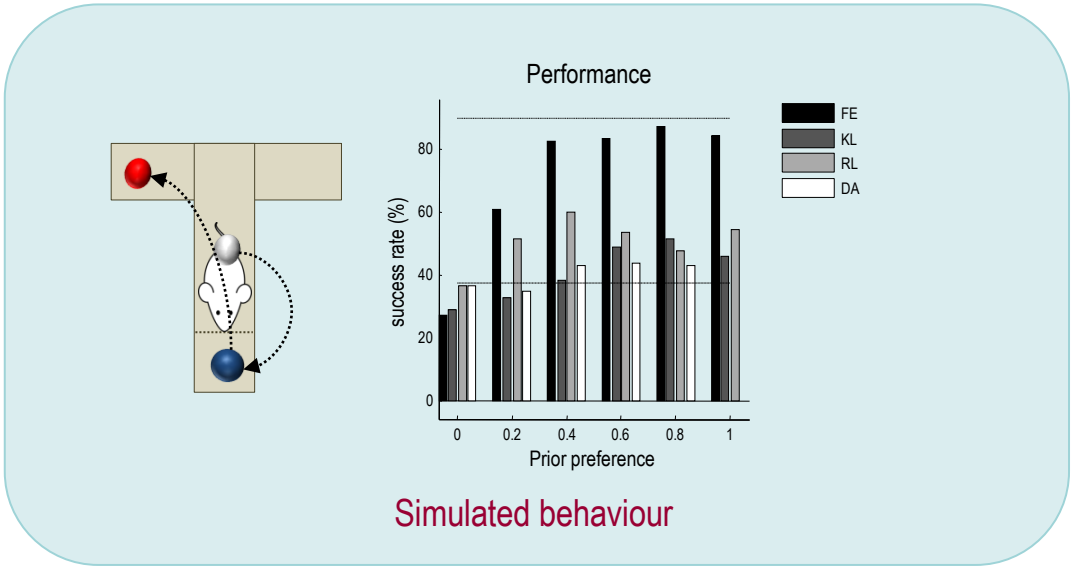
Action selection $\boldsymbol{\pi} = \sigma(-\gamma \cdot \hat{\mathbf{F}})$

Incentive salience $\boldsymbol{\beta} = \beta - \boldsymbol{\pi} \cdot \ln \boldsymbol{\pi}$

$$\tilde{\mu}^* = \arg\min F(o_1, \cdots, o_t, \tilde{\mu})$$

Functional anatomy





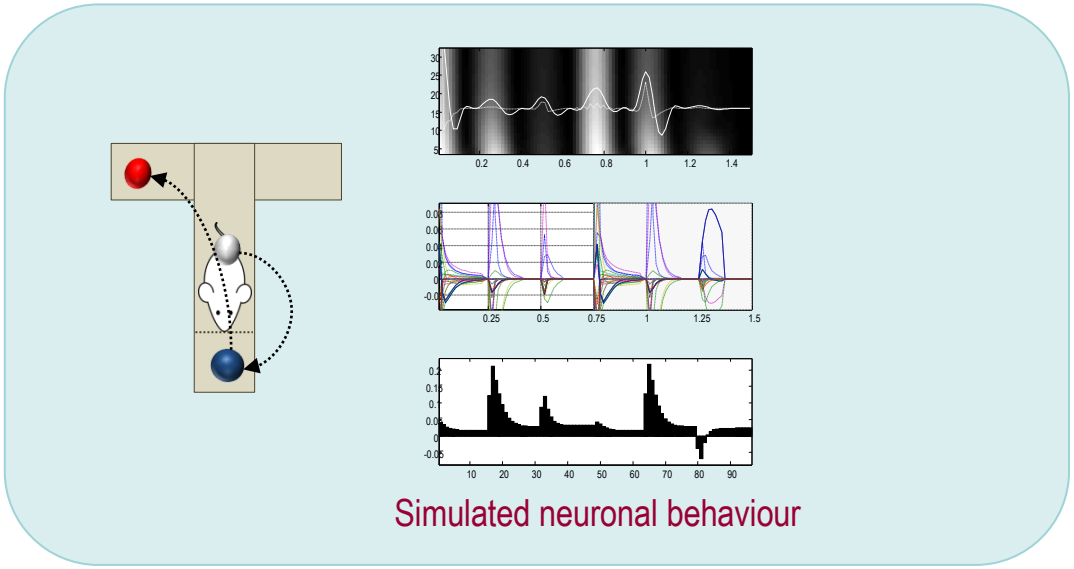
Variational updating

Perception $\dot{\mathbf{s}}_t^{\pi} = \sigma'(\mathbf{A} \cdot o_t + \mathbf{B}_{t-1}^{\pi}\mathbf{s}_{t-1}^{\pi} + \mathbf{B}_t^{\pi} \cdot \mathbf{s}_{t+1}^{\pi} - \ln \mathbf{s}_t^{\pi})$

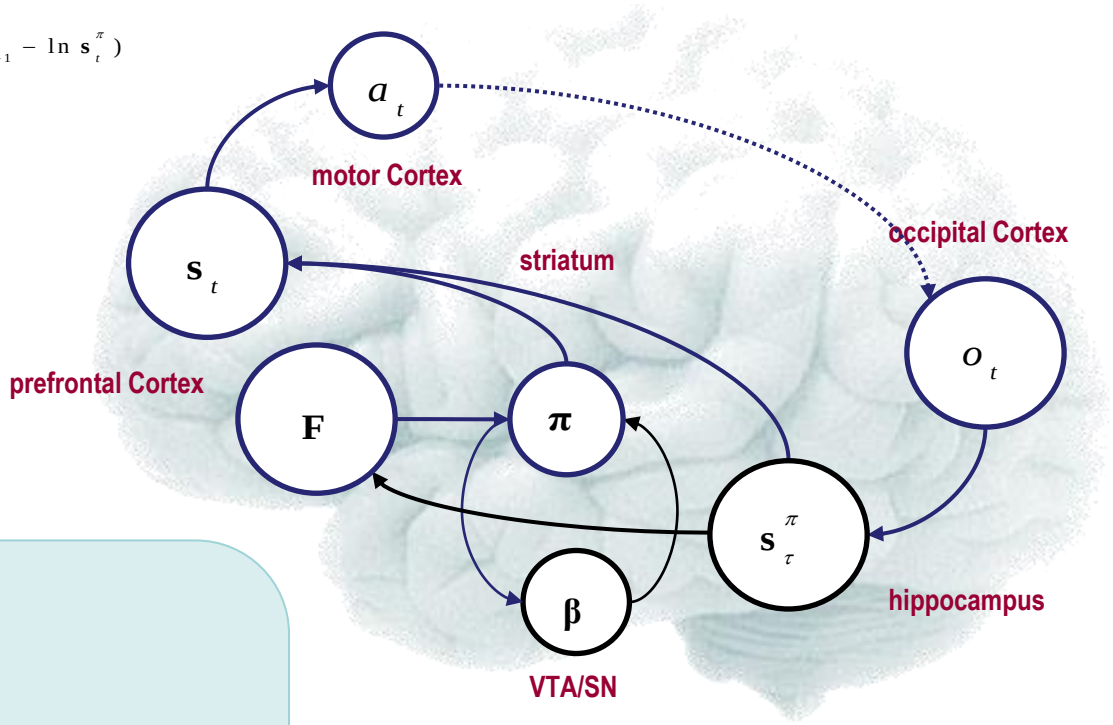
Action selection $\boldsymbol{\pi} = \sigma(-\gamma \cdot \hat{\mathbf{F}})$

Incentive salience $\dot{\boldsymbol{\beta}} = \beta - \boldsymbol{\pi} \cdot \ln \boldsymbol{\pi} - \boldsymbol{\beta}$

$$\dot{\tilde{\mu}} = -\nabla F(o_1, \cdots, o_t, \tilde{\mu})$$



Functional anatomy



What does believe updating explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

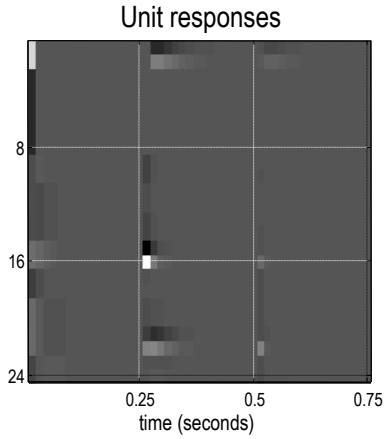
- Specify a deep (hierarchical) generative model
- Simulate approximate (active) Bayesian inference by solving:

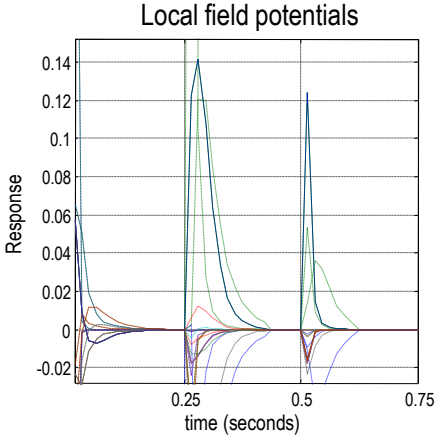
$$\dot{\tilde{\mu}} = -\nabla F\,(o_1, \cdots, o_t, \tilde{\mu})$$

What does Bayesian filtering explain?

Cross frequency coupling (phase precession)
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

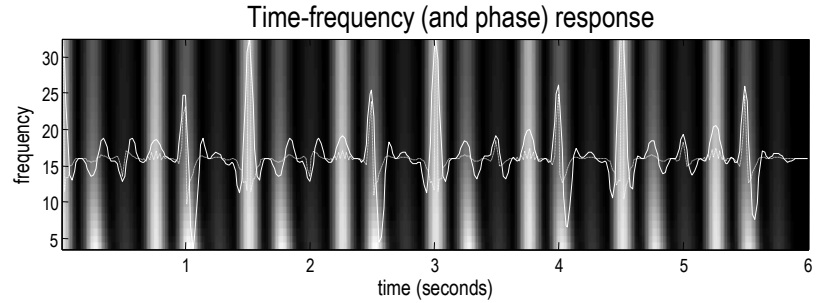
$$\tilde{\mu} = \begin{Bmatrix} \mathbf{s}_{t=1} \\ \mathbf{s}_{t=2} \\ \mathbf{s}_{t=3} \end{Bmatrix}$$

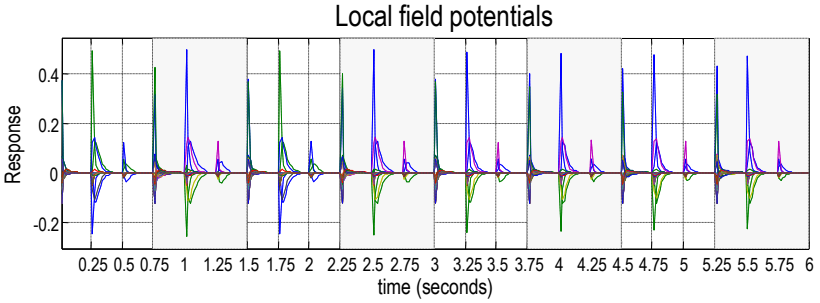


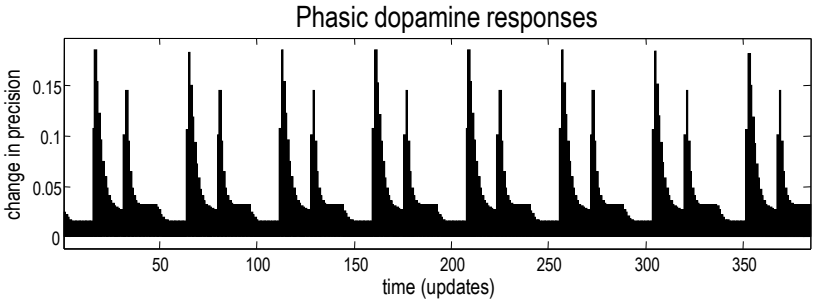


What does Bayesian filtering explain?

Cross frequency coupling (phase synchronisation)
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

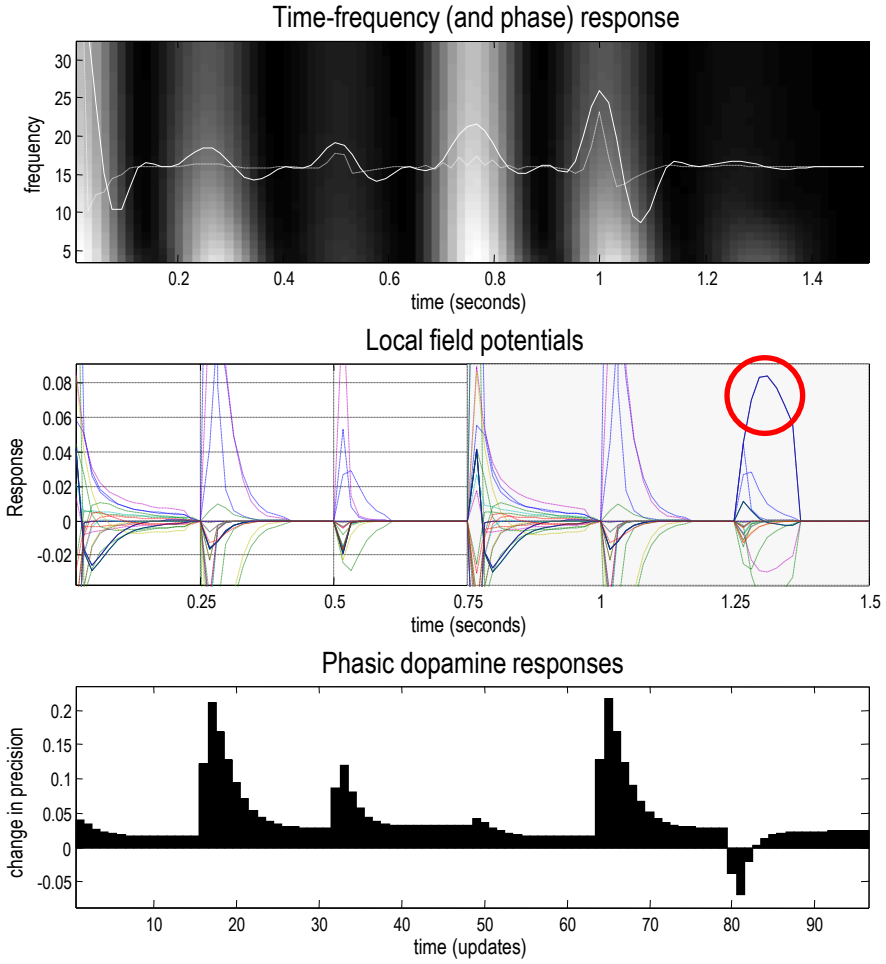






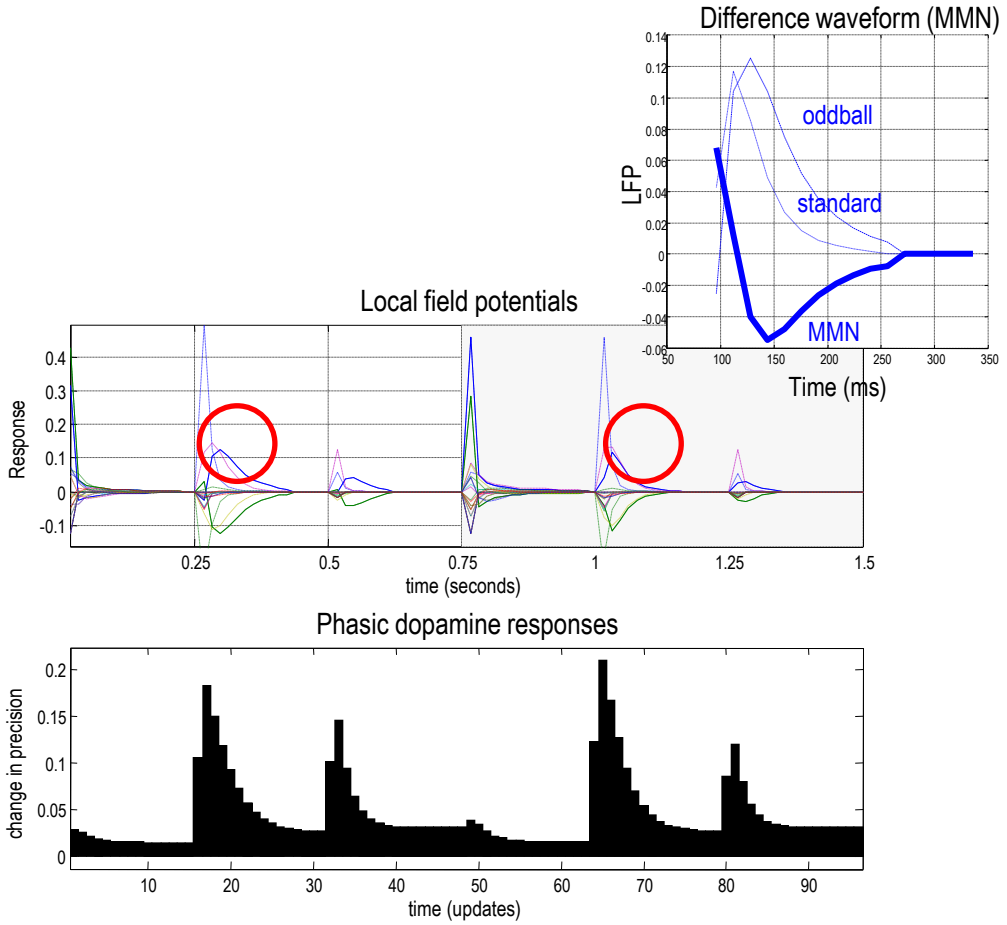
What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics



What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics



What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance (transfer of responses)
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics

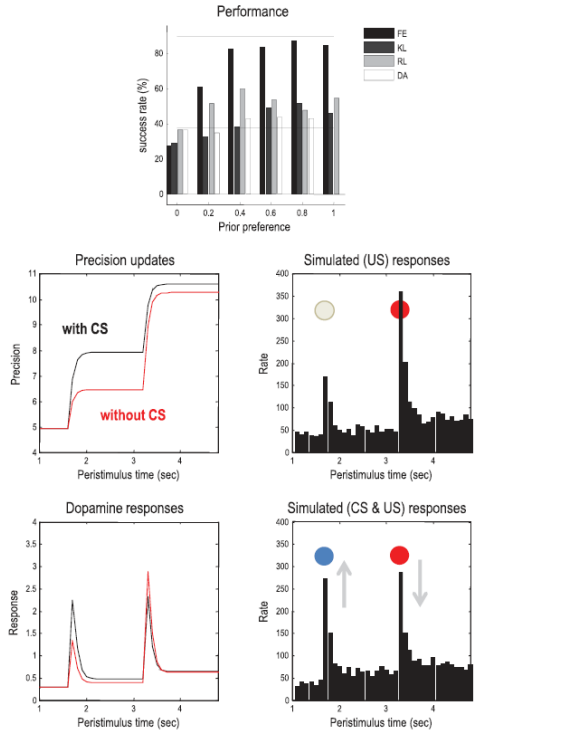
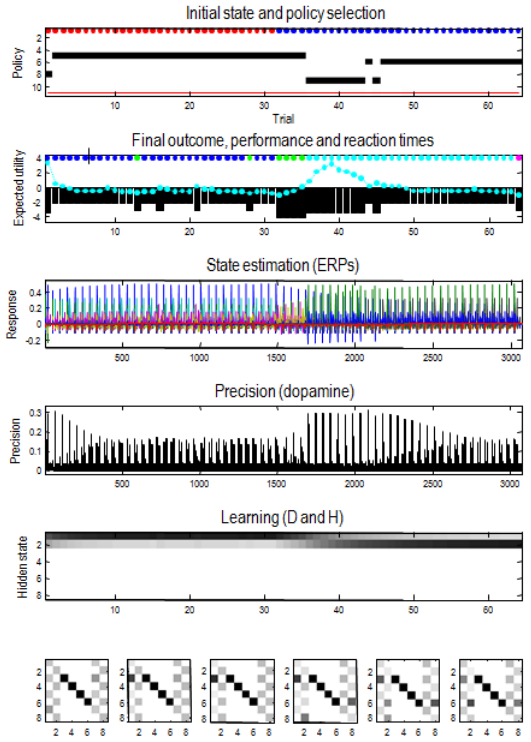


Figure 3. Upper panel: The results of 128 simulated trials assessed in terms of the probability of obtaining a reward. This performance is shown as a function of prior preference over six equally spaced levels. The four profiles correspond to active inference (FE), risk-sensitive control (KL), expected utility (RL), and active inference under fixed levels of precision (DA). See main text for a description of these schemes and how they relate to each other. The two horizontal lines show chance (bottom line) and optimal (top line) performance, respectively. Lower left panels: These report expected precision as a function of time within a trial (comprising three movements). The black lines correspond to a trial in which the cue (CS) was first accessed in the lower arm of the maze in the previous figure, after which the reward (US) was secured. The equivalent results, when staying at the center location and accessing the reward directly, are shown as red lines. The upper panel shows the expected precision and the lower panel shows simulated dopamine responses (that produce an increase in precision, which subsequently decays). Lower right panels: These show the equivalent results in terms of simulated dopamine discharges. The key thing to note here is that the responses to the cue (CS) are increased when it is informative (i.e., accessed in the lower arm), while subsequent responses to the reward (US) are decreased. See main text for details of these simulated responses.

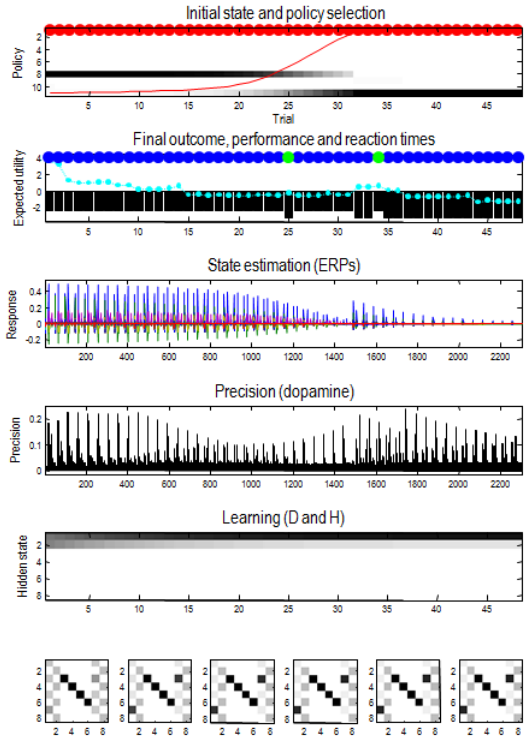
What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (reversal learning)
Interoceptive inference
Communication and hermeneutics



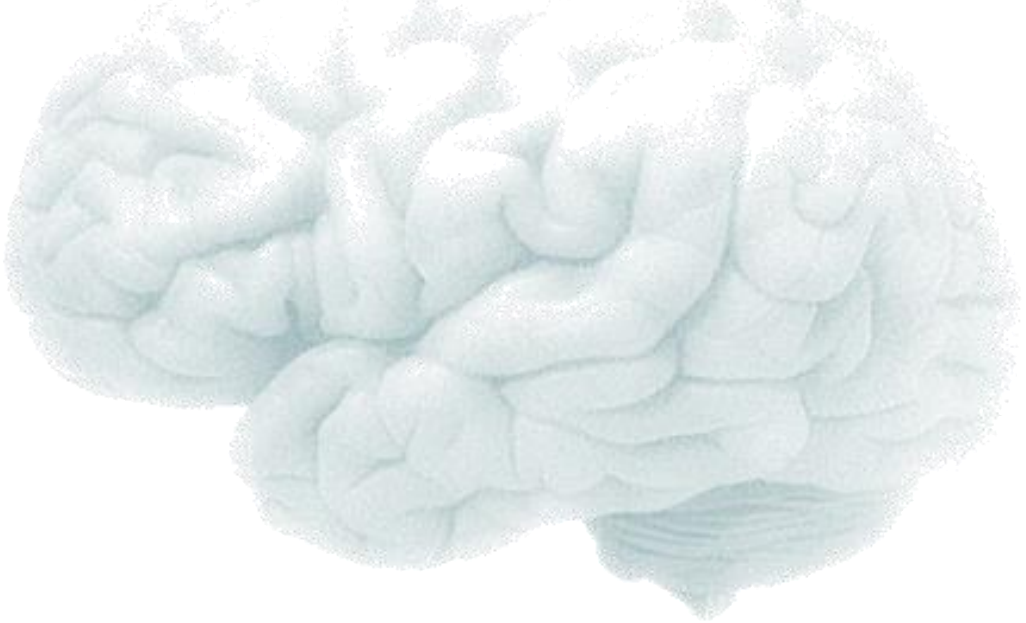
What does Bayesian filtering explain?

Cross frequency coupling
Perceptual categorisation
Violation (P300) responses
Oddball (MMN) responses
Omission responses
Attentional cueing (Posner paradigm)
Biased competition
Visual illusions
Dissociative symptoms
Sensory attenuation
Sensorimotor integration
Optimal motor control
Motor gain control
Oculomotor control and smooth pursuit
Visual searches and saccades
Action observation and mirror neuron responses
Dopamine and affordance
Action sequences (devaluation)
Interoceptive inference
Communication and hermeneutics

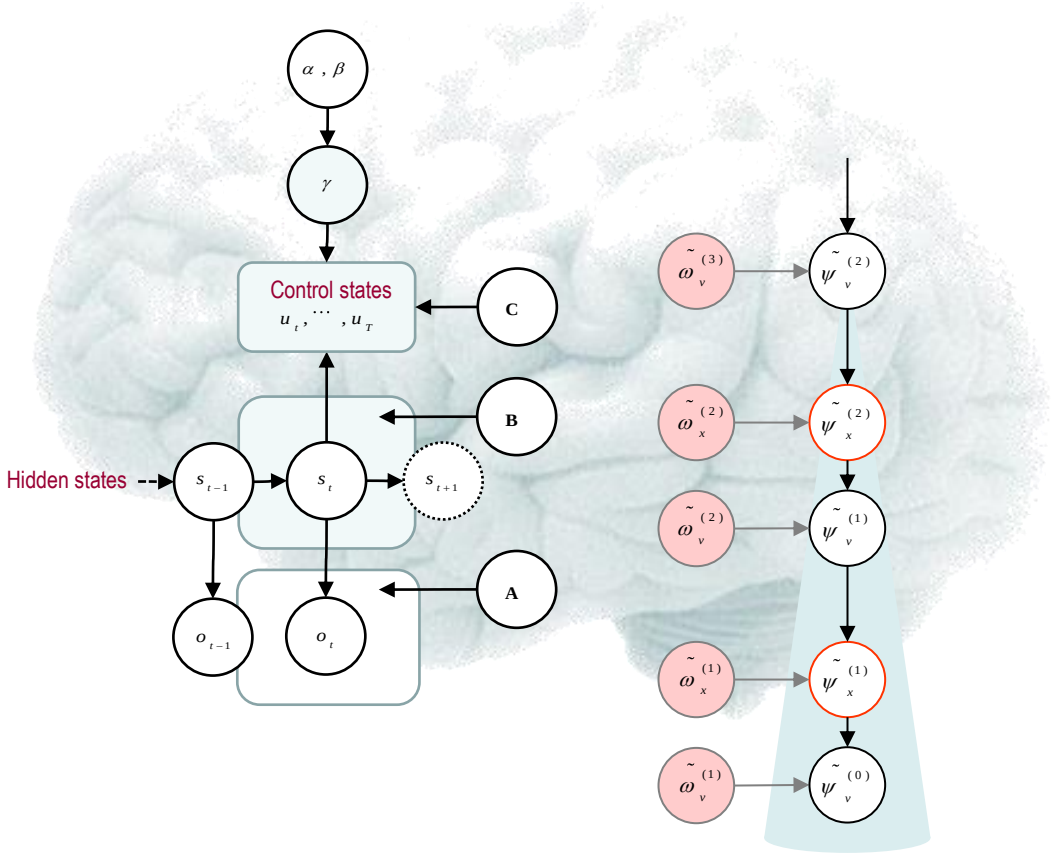


Does the brain use continuous or discrete state space models or both?

Does the brain encode beliefs with ensemble densities or sufficient statistics?



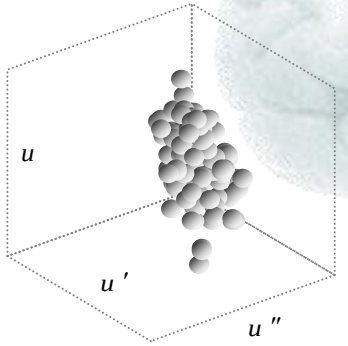
Does the brain use continuous or discrete state space models or both?



Does the brain use continuous or discrete state space models or both?

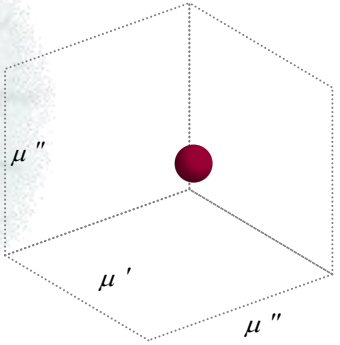
Does the brain encode beliefs with ensemble densities or sufficient statistics?

Ensemble code



or

Laplace code



$$q(x \mid \tilde{u}) \approx N\left(\tilde{\mu} = \frac{1}{n}\sum_i \tilde{u}_i, \Sigma = \frac{1}{n}\sum_i (\tilde{u}_i - \tilde{\mu})(\tilde{u}_i - \tilde{\mu})^T\right)$$

Ensemble density

$$q(x \mid \tilde{\mu}) = N\left(\tilde{\mu}, \Sigma(\tilde{\mu})\right)$$

Parametric density

Variational filtering with ensembles

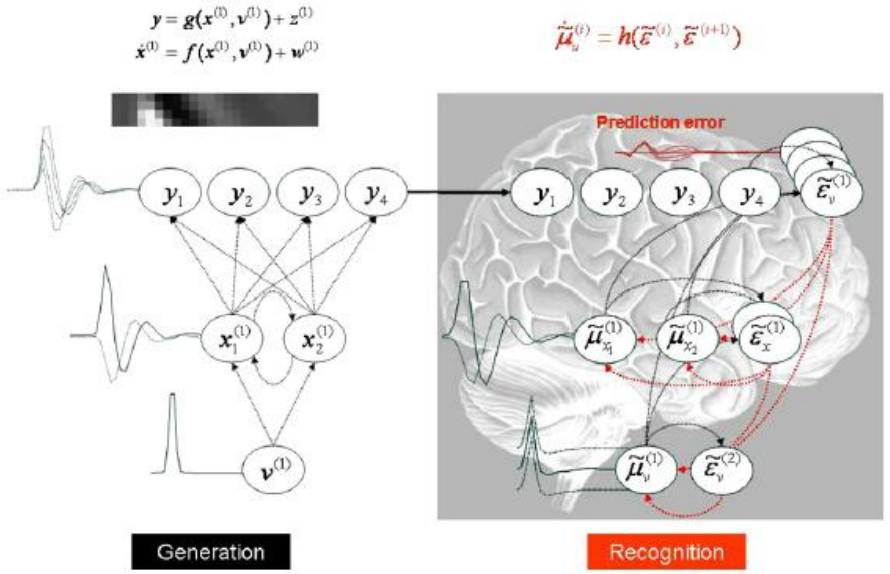


Fig. 5. Diagram showing the generative model (left) and corresponding recognition; i.e., neuronal model (right) used in the simulations. Left panel: this is the generative model using a single cause $v^{(1)}$, two dynamic states $x_1^{(1)}, x_2^{(1)}$ and four outputs $y_1, \ldots, y_4$. The lines denote the dependencies of the variables on each other, summarised by the equation on top (in this example both the equations were simple linear mappings). This is effectively a linear convolution model, mapping one cause to four outputs, which form the inputs to the recognition model (solid arrow). The architecture of the corresponding recognition model is shown on the right. This has a similar architecture, apart from the inclusion of prediction error units; $\tilde{\varepsilon}_u^{(i)}$. The combination of forward (red lines) and backward influences (black lines) enables recurrent dynamics that self-organise (according to the recognition equation; $\dot{\tilde{\mu}}_u^{(i)} = h(\tilde{\varepsilon}^{(i)}, \tilde{\varepsilon}^{(i+1)})$) to suppress and hopefully eliminate prediction error, at which point the inferred causes and real causes should correspond. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

Variational filtering

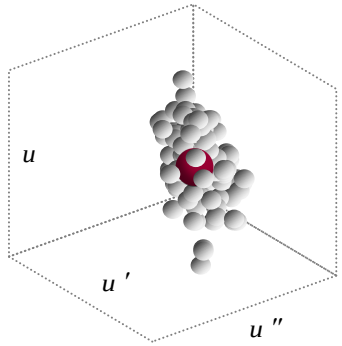
$$F = G - H$$

$$G = \left\langle \ln p(\tilde{s}, \tilde{u}) \right\rangle_q = \left\langle U(\tilde{u}) \right\rangle_q$$

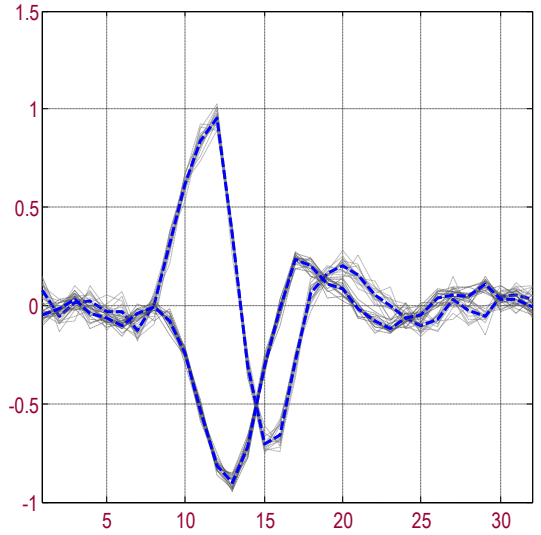
$$H = \left\langle \ln q(\tilde{u}) \right\rangle_q$$

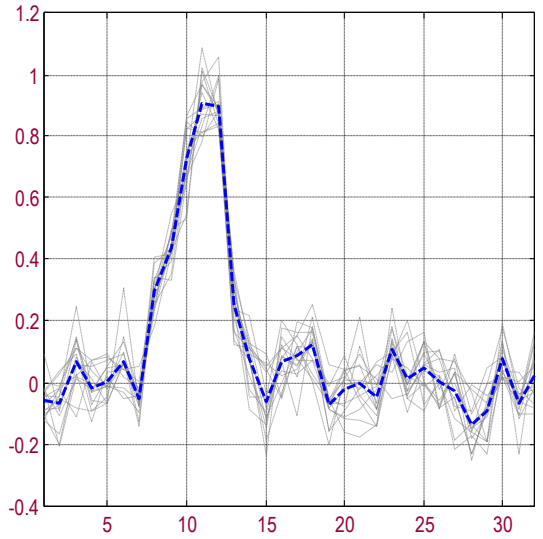
$$\dot{\tilde{u}} = D\tilde{\mu} - \Gamma \nabla U(\tilde{s}, \tilde{u}) + \Gamma$$

$$\tilde{u} = (v, v', v'', \cdots x, x', x'', \cdots)$$

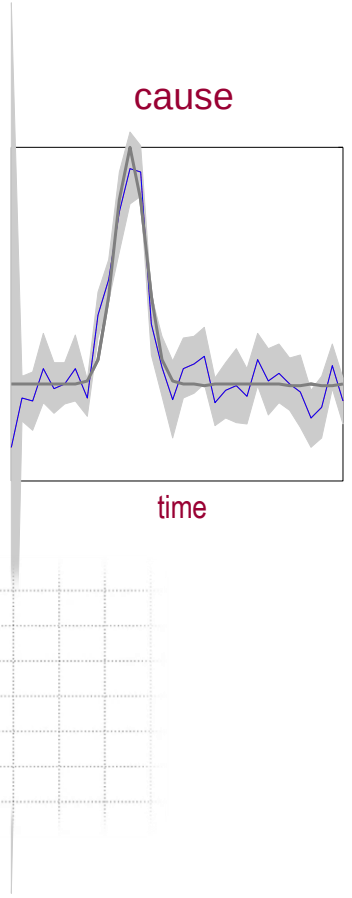


$v(t)$





$x(t)$



From ensemble coding to predictive coding (Bayesian filtering)

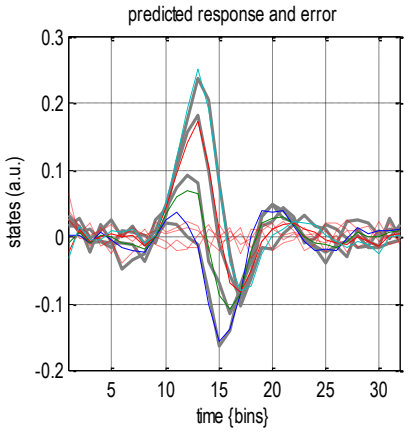
Taking the expectation of the ensemble dynamics, under the Laplace assumption,we get:

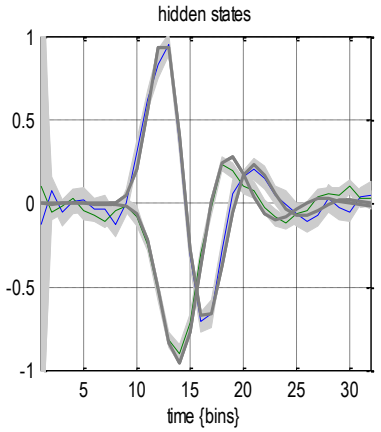
$$\dot{\tilde{u}} = D\tilde{\mu} - \nabla U + \Gamma \Rightarrow$$
$$\dot{\tilde{\mu}} = D\tilde{\mu} - \nabla F \Leftrightarrow \dot{\tilde{\mu}} - D\tilde{\mu} = \nabla F$$

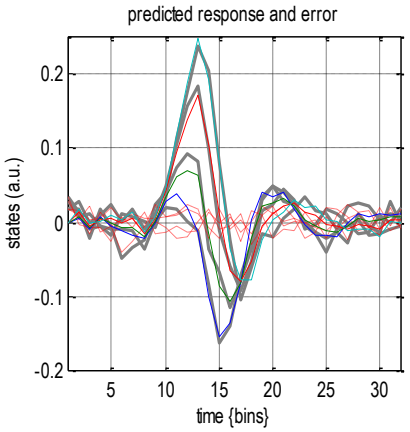
This can be regarded as a gradient ascent in a frame of reference that moves along the trajectory encoded in generalised coordinates. The stationary solution, in this moving frame of reference, maximises variational action by the Fundamental lemma.

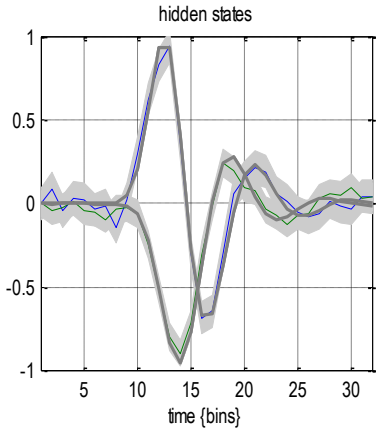
$$\dot{\tilde{\mu}} - D\tilde{\mu} = 0 \Rightarrow \nabla F = 0 \Leftrightarrow \delta_u F = 0$$

c.f., Hamilton's principle of stationary action.



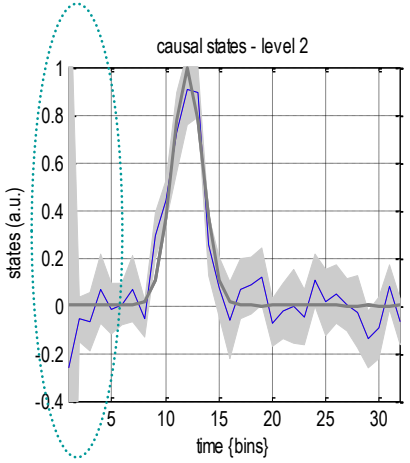




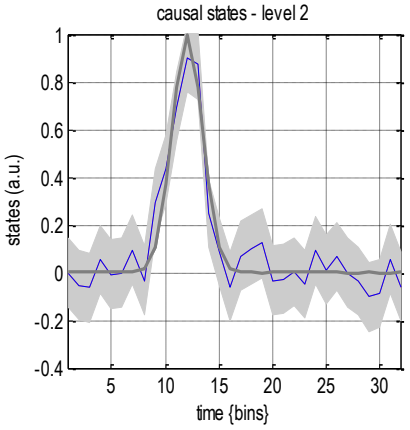


Deconvolution with variational filtering (SDE) – free-form

Deconvolution with Bayesian filtering (ODE) – fixed-form



$$\dot{\tilde{u}} = D\tilde{\mu} - \nabla U + \Gamma$$



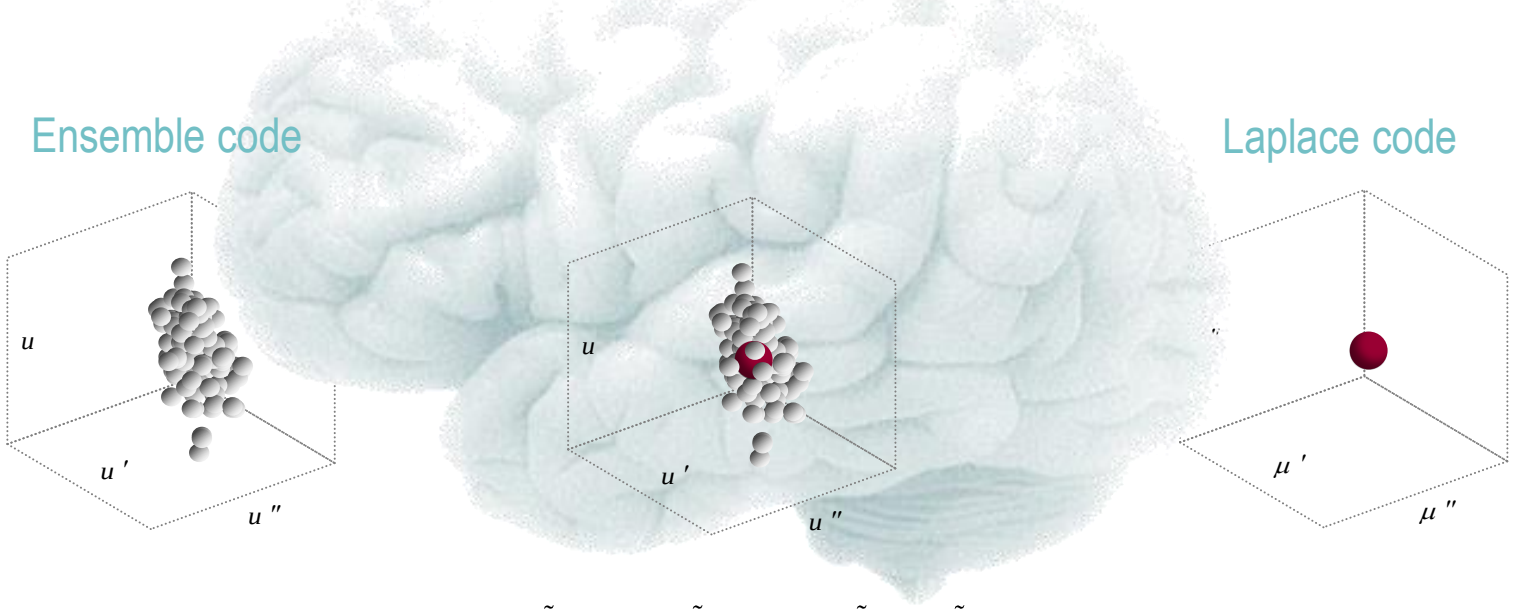
$$\dot{\tilde{\mu}} = D\tilde{\mu} - \nabla F$$

Does the brain use continuous or discrete state space models or both?

Does the brain encode beliefs with ensemble densities or sufficient statistics or both?

Does the brain use continuous or discrete state space models or both?

Does the brain encode beliefs with ensemble densities or sufficient statistics or both?



$$q(x \mid \tilde{u}) \approx N\left(\tilde{\mu} = \frac{1}{n}\sum_i \tilde{u}_i, \Sigma(\tilde{\mu})\right)$$

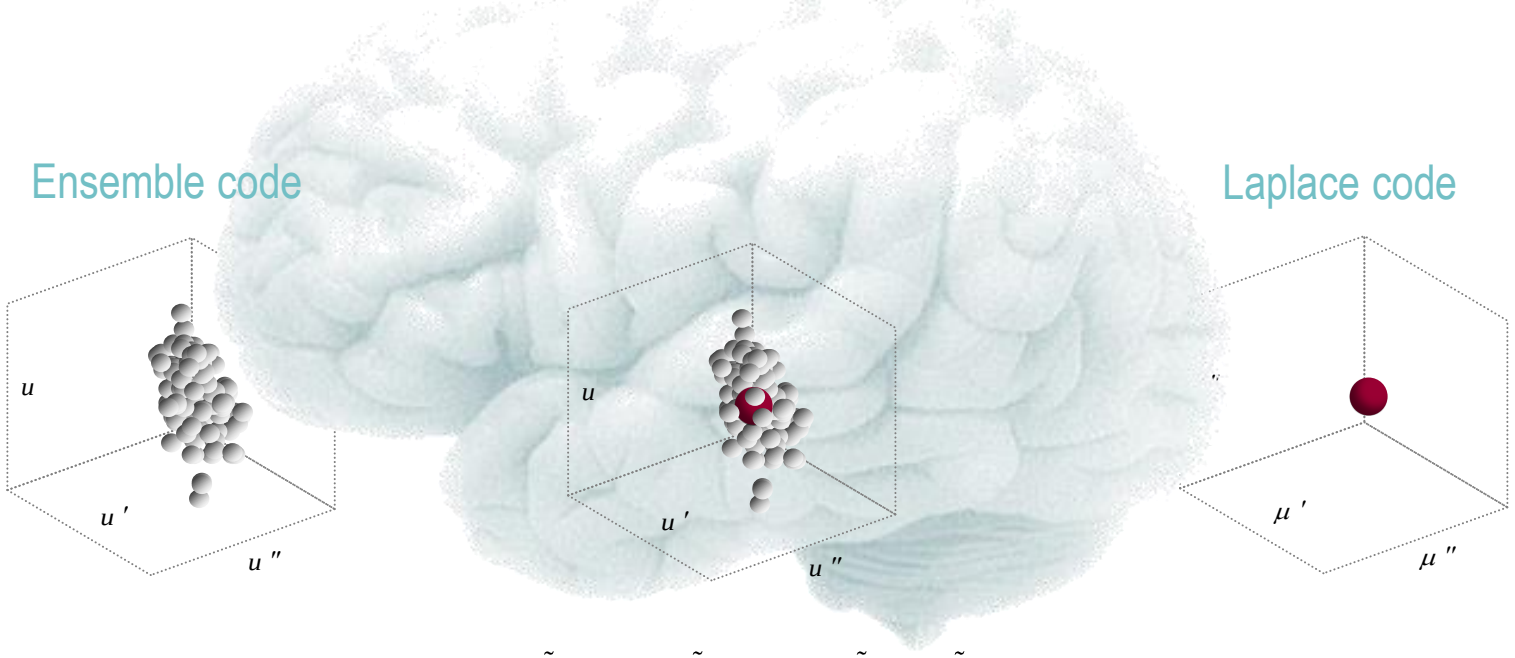
Parametric description of ensemble density

In other words, do we have:

An approximate description of (nearly) exact Bayesian inference

or

A (nearly) exact description of approximate Bayesian inference



$$q(x \mid \tilde{u}) \approx N\left(\tilde{\mu} = \frac{1}{n}\sum_i \tilde{u}_i, \Sigma(\tilde{\mu})\right)$$

Parametric description of ensemble density

Thank you

And thanks to collaborators:

Rick Adams
Ryszard Auksztulewicz
Andre Bastos
Sven Bestmann
Harriet Brown
Jean Daunizeau
Mark Edwards
Chris Frith
Thomas FitzGerald
Xiaosi Gu
Stefan Kiebel
James Kilner
Christoph Mathys
Jérémie Mattout
Rosalyn Moran
Dimitri Ognibene
Sasha Ondobaka
Will Penny
Giovanni Pezzulo
Lisa Quattrocki Knight
Francesco Rigoli
Klaas Stephan
Philipp Schwartenbeck

And colleagues:

Micah Allen
Felix Blankenburg
Andy Clark
Peter Dayan
Ray Dolan
Allan Hobson
Paul Fletcher
Pascal Fries
Geoffrey Hinton
James Hopkins
Jakob Hohwy
Mateus Joffily
Henry Kennedy
Simon McGregor
Read Montague
Tobias Nolte
Anil Seth
Mark Solms
Paul Verschure

And many others