



Precision and False Perceptual Inference

Thomas Parr*, David A. Benrimoh, Peter Vincent and Karl J. Friston

Institute of Neurology, Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom

Accurate perceptual inference fundamentally depends upon accurate beliefs about the reliability of sensory data. In this paper, we describe a Bayes optimal and biologically plausible scheme that refines these beliefs through a gradient descent on variational free energy. To illustrate this, we simulate belief updating during visual foraging and show that changes in estimated sensory precision (i.e., confidence in visual data) are highly sensitive to prior beliefs about the contents of a visual scene. In brief, confident prior beliefs induce an increase in estimated precision when consistent with sensory evidence, but a decrease when they conflict. Prior beliefs held with low confidence are rapidly updated to posterior beliefs, determined by sensory data. These induce much smaller changes in beliefs about sensory precision. We argue that pathologies of scene construction may be due to abnormal priors, and show that these can induce a reduction in estimated sensory precision. Having previously associated this precision with cholinergic signaling, we note that several neurodegenerative conditions are associated with visual disturbances and cholinergic deficits; notably, the synucleinopathies. On relating the message passing in our model to the functional anatomy of the ventral visual stream, we find that simulated neuronal loss in temporal lobe regions induces confident, inaccurate, empirical prior beliefs at lower levels in the visual hierarchy. This provides a plausible, if speculative, computational mechanism for the loss of cholinergic signaling and the visual disturbances associated with temporal lobe Lewy body pathology. This may be seen as an illustration of the sorts of hypotheses that may be expressed within this computational framework.

Keywords: active inference, saccades, visual system, synucleinopathy, precision

OPEN ACCESS

Edited by:

Ehud Ahissar,
Weizmann Institute of Science, Israel

Reviewed by:

Claire O'Callaghan,
University of Cambridge,
United Kingdom
Daniel Collerton,
Newcastle University, United Kingdom
Ishan Walpola,
University of Sydney, Australia, in
collaboration with reviewer CO

*Correspondence:

Thomas Parr
thomas.parr.12@ucl.ac.uk

Received: 15 March 2018

Accepted: 30 August 2018

Published: 20 September 2018

Citation:

Parr T, Benrimoh DA, Vincent P and
Friston KJ (2018) Precision and False
Perceptual Inference.
Front. Integr. Neurosci. 12:39.
doi: 10.3389/fnint.2018.00039

INTRODUCTION

The brain's visual system must overcome formidable inferential challenges. Despite receiving sequentially sampled, spatially limited sensory information from a two-dimensional array of photoreceptors, we perceive spatially and temporally continuous three-dimensional visual scenes, populated with complex objects. However, despite this remarkable capacity for scene construction (Mirza et al., 2016; Parr and Friston, 2017a), the visual system is not infallible. It depends upon a delicate balance between prior beliefs about perceptual hypotheses (Gregory, 1980), and the sensory evidence that supports or refutes them (Geisler and Kersten, 2002; Brown and Friston, 2012). This paper is about the computational mechanisms that could maintain this balance, and the consequences of their failure (e.g., Collerton et al., 2005).

We begin by outlining active inference (Friston et al., 2015), a theory of optimal behavior that accommodates both perceptual inference and foraging for new sensory data (Mirza et al., 2016). For our purposes, this foraging takes place through saccades to locations in the visual scene. We

simulate visual foraging and the concurrent optimisation of beliefs about the precision of (i.e., confidence in) sensations, using a simple visual scene. This serves to illustrate the sensitivity of perceptual inference to the prior beliefs chosen, and the potential for (pathological) false inference (Friston, 2017).

These considerations serve as a foundation for the next section, which explores the source of abnormal prior beliefs. To do so, we extend the model and relate inferential message passing to the interactions between the ventral visual stream and subcortical structures. This draws upon previous suggestions that acetylcholine may act as a sensory precision signal (Yu and Dayan, 2005; Parr and Friston, 2017c). This rests upon several empirical observations. First, nicotinic acetylcholine receptors are found on the presynaptic terminals of cells in layers 3 and 4 of the cortex (Sahin et al., 1992; Lavine et al., 1997). These laminae are the targets of sensory relays from the thalamus (Shipp, 2007). Secondly, cholinergic manipulations modulate the gain of visually evoked responses (Gil et al., 1997; Disney et al., 2007). This renders it almost tautologically true that one of the roles of acetylcholine is to modulate the precision of (some types of) sensory input, as precision and gain are mathematically identical. Finally, in both behavioral (Marshall et al., 2016) and neuroimaging (Moran et al., 2013) studies in humans that explicitly test the association between precision and acetylcholine, differences following cholinergic manipulations are best accounted for by altered precision. If we lesion our model to simulate the Lewy body pathology associated with visual disturbances in synuclein disorders (Harding et al., 2002), we find that the downstream consequences include decreased sensory precision, neuronal gain, and false perceptual inferences. These are consistent with the cholinergic deficits, occipital hypometabolism, and visual hallucinations that occur with synucleinopathies (McKeith et al., 2004). While this is speculative, we draw upon this notion throughout to illustrate the potential relevance of these ideas to pathology.

A prominent, but controversial, hypothesis in research on neurodegenerative conditions is that deficits in acetylcholine underwrite the mnemonic deficits characteristic of Alzheimer's disease (Perry, 1986; Contestabile, 2011). This has sometimes been extended to other neurodegenerative conditions (Terry and Buccafusco, 2003), and to normal aging (Bartus et al., 1982). It is important to emphasize that the theory and simulations in this paper are not intended to address the cholinergic hypothesis in neurodegeneration. While we appeal to computational theories concerning the role of acetylcholine in health and hallucinatory disorders, this work should not be seen as supporting (or refuting) the cholinergic hypothesis of cognitive decline in dementia: this work just offers a computational (i.e., teleological) formulation of cholinergic neurotransmission that may help understand psychopathology.

ACTIVE INFERENCE

Under active inference, creatures minimize their free energy through action and perception (Friston et al., 2010). Free energy is a functional of beliefs about the processes causing

sensations, and a function of sensory data. It is an upper bound on the negative evidence for these beliefs (Beal, 2003), meaning minimisation of free energy is formally equivalent to maximization of evidence for a model of the world (Hohwy, 2016). This (generative) model expresses beliefs about how sensory data are generated, and can take subtly different forms, depending on the type of inference being performed (Friston et al., 2017b). In this paper, we describe inference using a Markov Decision Process (MDP) model:

$$P(\bar{o}, \bar{s}, \pi, \zeta) = \underbrace{P(\pi)P(s_1)P(\zeta)}_{\text{priors}} \prod_{\tau} \underbrace{P(s_{\tau+1}|s_{\tau}, \pi)}_{\text{transitions}} \underbrace{P(o_{\tau}|s_{\tau}, \zeta)}_{\text{likelihood}}$$

Here, $\bar{o} = [o_1, o_2, \dots, o_T]^T$ is a sequence of sensory observations over time. \bar{s} is the sequence of latent (hidden) states that are not directly observable. π is the policy (or sequence of actions) the subject pursues. The equation above says that observations depend only upon the latent states at the current time step and a precision term, ζ (Parr and Friston, 2017c). Intuitively, if $\zeta = \infty$ this probability (i.e., likelihood) becomes deterministic. Conversely, if $\zeta = 0$ the likelihood of any observation is completely random; irrespective of the state of the world.

For people who are not familiar with the notion of precision, precision quantifies the inverse variability or uncertainty associated with a probability distribution. For example, if the probability of an observed consequence, given a particular cause is very precise, then one can be confident that the observed outcome can be attributed to a particular cause. Conversely, if contingencies are imprecise, there is no definitive relationship between causes and consequences and observations do little to resolve uncertainty about causes. In this sense, precision corresponds to the confidence with which one can infer a cause from observations or data. If one believes contingencies are very precise, one will afford greater weight to sensory evidence in terms of updating beliefs about their causes. This is why optimizing the precision *per se* has been associated with attention. Clearly, getting precision wrong can have profound effects on inference; particularly, in hierarchical inference where the relative precision at different levels of an inference hierarchy becomes especially important—and sometimes counterintuitive (as we will see below). Heuristically, getting precision wrong can lead to both false positives and false negatives, which—in the context of psychopathology—may be the formal homologue of hallucinosis and (attentional) neglect.

States depend upon the states at the previous time step, and the policy currently being pursued. Technically, these probabilistic dependencies are called empirical priors. In addition to the likelihood and empirical priors, three (full) prior distributions are needed: these are the prior over policies, the initial state, and the precision. Active inference mandates that the first of these has a very specific form. However, before we specify this form, we have to consider the form of the (approximate) posterior beliefs, Q , an MDP agent holds about the latent variables. To make inference tractable, we employ a mean-field approximation (Friston and Buzsáki, 2016) that assumes the

beliefs may be expressed as a product of (marginal) factors:

$$Q(\tilde{s}, \pi, \zeta) = Q(\pi)Q(\zeta) \prod_{\tau} Q(s_{\tau}|\pi)$$

These are the beliefs that are optimized to minimize free energy, where the approximation allows us to optimize each factor independently. We can now write the free energy under a given policy as

$$F(\pi) = E_Q[\ln Q(\tilde{s}, \zeta|\pi) - \ln P(\tilde{o}, \tilde{s}, \zeta|\pi)]$$

The notation $E_Q[\]$ means the expectation (or average) under the approximate posterior beliefs. For actions to minimize this quantity, policies should be selected that lead to lower free energies. However, the free energy is a function of observations, and cannot be defined in relation to observations that have yet to occur. To solve this problem, we can define an expected free energy, and use this to define prior beliefs about policies that minimize this quantity.

$$G(\pi) = E_{\tilde{Q}}[\ln Q(\tilde{s}, \zeta|\pi) - \ln P(\tilde{o}, \tilde{s}, \zeta|\pi)]$$

$$\tilde{Q}(\tilde{o}, \tilde{s}|\pi) = Q(\tilde{s}|\pi)P(\tilde{o}|\tilde{s})$$

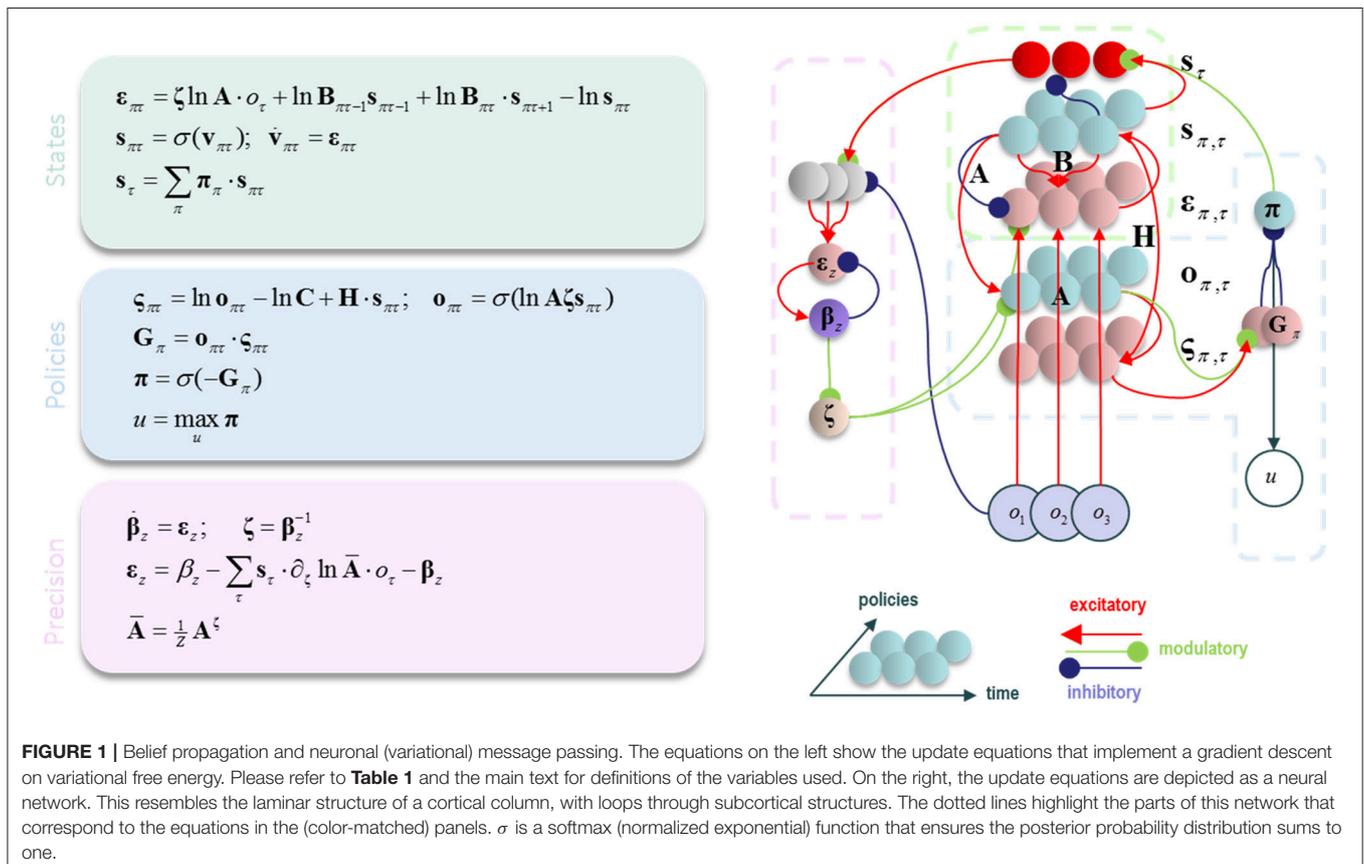
$$P(\pi) = \sigma(-G(\pi))$$

The key move here is to augment the distribution used for the expectation so that it includes beliefs about future observations.

This allows us to move from a free energy to an expected free energy. The third line says that those policies considered *a priori* more probable are those that lead to a low expected free energy. Minimizing expected free energy is equivalent to minimizing uncertainty about states of the world. In the absence of any explicit rewards (i.e., prior preferences over outcomes), this uncertainty resolving behavior dominates policy selection—to select policies with the greatest epistemic value or affordance.

This sort of evidence accumulation depends upon the precision afforded sensory evidence, relative to prior beliefs. Crucially, the likelihood or sensory precision ζ itself has to be inferred (*via* minimization of free energy). This means that while policies are selected to minimize uncertainty, this selection is based upon posterior expectations about the precision of sensory evidence; namely, expected uncertainty that is determined by the ability of predictions to explain the outcomes encountered.

Having specified the generative model, we can now compute the free energy gradients for each posterior belief (please see the **appendix**), and use these to derive the variational message passing equations needed for belief updating over time (Friston et al., 2017a,b; Parr and Friston, 2017c). These are shown in **Figure 1** both as belief update equations and graphically as a neural network. There are two important points to draw from this sort of belief update or propagation scheme. First, the structure of the neural network closely resembles that of a cortical column, with loops through subcortical structures (Friston et al., 2017c).



Second, ζ appears as a multiplicative factor, controlling the gain of messages derived from sensations (Feldman and Friston, 2010). We have previously associated this term with cholinergic projections to the cortex (Parr and Friston, 2017c). This is consistent with empirical data suggesting that acetylcholine

mediates the gain of visual cortical responses (Disney et al., 2007), and with the cortical hypometabolism that occurs following lesions to cholinergic forebrain nuclei (Motohiro et al., 1987).

THE GENERATIVE MODEL

We have outlined above the form of a generic Markov Decision Process, and of the message passing (Dauwels, 2007) it entails. We now focus on a more concrete example. This is the model we will use to illustrate inferences about precision (Figure 2). It comprises a very simple visual scene, containing only four features. Each of these is a circle that may be absent (white), green, or blue. Only one of these features may be foveated at any one time (Mirza et al., 2016). This means that there are four hidden state variables representing visual features, and a fifth that represents fixation location (Parr and Friston, 2017c). The transition probabilities associated with the visual features are all identity matrices, ensuring that the precision of prior beliefs about hidden states is determined only by the prior probability over initial states, $P(s_1)$. The transitions between fixation locations depend upon the action (i.e., saccade) selected, and it is this that facilitates the active interrogation of the visual scene (Gibson, 1966).

Two types of sensory data are generated by this model; visual and proprioceptive. The latter is generated by an identity mapping from the fixation location hidden state. Visual data are caused by both the fixation location and the visual feature found at that location. Each location is equipped with its own sensory

TABLE 1 | Glossary of variables (in the text and figures).

Variable	Definition
G_π	Expected free energy
π	Policy posterior
$s_{\pi\tau}^{(i)}$	State posterior belief (for a given policy and time) at hierarchical level i
$o_{\pi\tau}$	Outcome belief (for a given policy and time)
o_τ	Observed outcome
$A_{ij} = P(o_\tau = i s_\tau = j, \zeta = 1)$	Likelihood matrix (mapping states to outcomes)
$\bar{A}_{ij} = P(o_\tau = i s_\tau = j, \zeta)$	Likelihood matrix (mapping states to outcomes)
$B_{ij}(u) = P(s_{\tau+1} = i s_\tau = j, \pi(\tau) = u)$	Transition matrix (mapping states to states)
$C_{\tau i} = P(o_\tau = i)$	Outcome prior
$H_j = \sum_i P(o_\tau = i s_\tau = j, \zeta) \ln P(o_\tau = i s_\tau = j, \zeta)$	Entropy of the likelihood mapping
$\beta_Z P(\zeta) = \text{Gamma}(1, \beta_Z)$	Parameter of prior and posterior beliefs about sensory precision
$\beta_Z Q(\zeta) = \text{Gamma}(1, \beta_Z)$	Parameter of prior and posterior beliefs about sensory precision
$\zeta = E_{Q(\zeta)}[\zeta] = \beta_Z^{-1}$	Posterior expectation of sensory precision

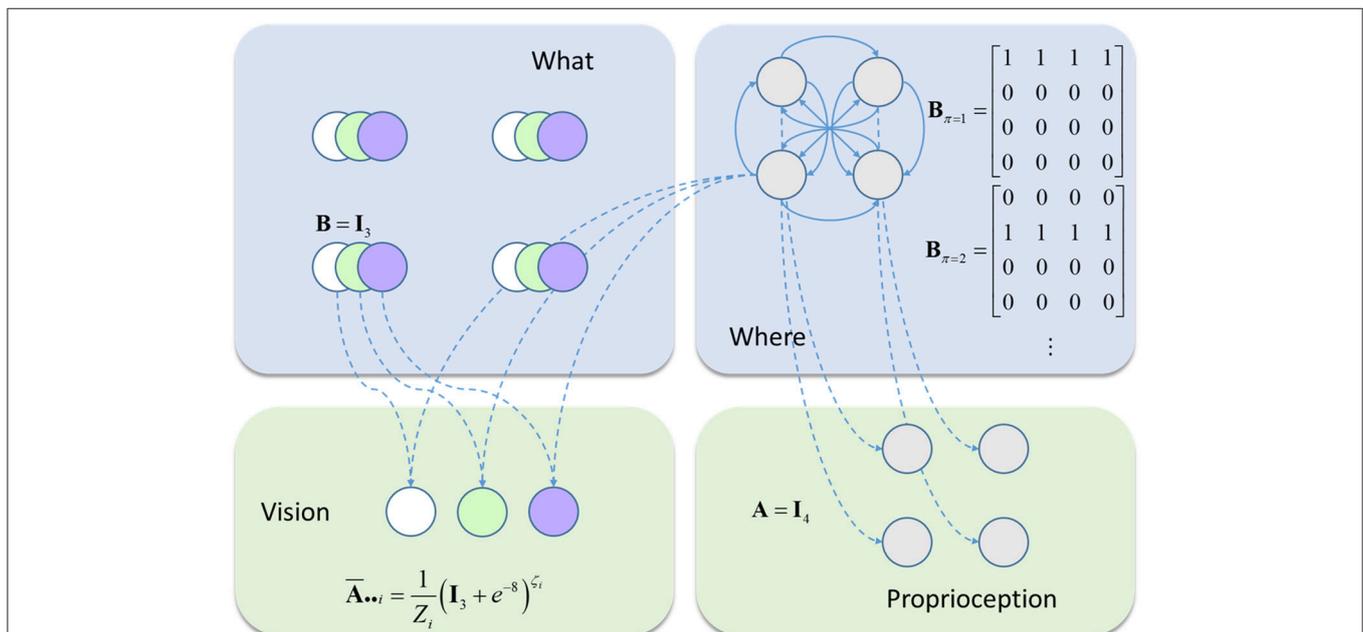


FIGURE 2 | The generative model. This schematic illustrates the form of the generative model used in our simulations. Blue panels show the hidden states: four stimuli at different locations (upper left) and the fixation location (upper right). The latter is a control state (entailed by the policy), with transition probabilities that depend upon policy choices. The green panels show the observable outcomes that are caused by the hidden states. There is one visual and one proprioceptive outcome. The visual outcome depends on the fixation location and what is at that location. The proprioceptive outcome depends only upon the fixation location. The precision of the visual outcome depends upon the current fixation location.

precision, and beliefs about these precisions can be optimized independently. This generative model should be seen as a toy example that serves to illustrate a subset of pathological changes in visual processing. It is not intended as a complete account of perceptual changes following neurodegenerative insults to the visual system. To demonstrate the behavior of this model, we simulated several different combinations of stimuli and prior belief.

SIMULATIONS

Given this generative model, we can now simulate scene construction via active vision by solving the equations in **Figure 1**—and examining the simulated behavior and posterior beliefs of a synthetic subject. This approach to simulating visual search behavior has been described in detail in (Mirza et al., 2016; Parr and Friston, 2017c). We will present two sets of simulations. The first uses a simple (single hierarchical level) generative model to illustrate the basics of perceptual inference—and how this depends upon the precision afforded sensory evidence, relative to (empirical) prior beliefs about state transitions. In the second simulation, we equip the model with a second (hierarchical) level that embodies the belief that outcomes are generated by a scene (i.e., a combination of visual objects at four spatial locations) that remains constant over successive (five saccade) visual searches. While this, deliberately simple, form of visual scene does not capture the rich phenomenology associated with real scene construction (Hassabis et al., 2007), it enables us to simulate visual processing under lesions that are hierarchically remote from the (neuromodulatory) effects of expected precision. We offer this as a formal explanation for the sort of (functional) diaschisis that characterizes synucleinopathies; particularly those associated with visual hallucinosis.

Figure 3 shows the results of simulating a visual search for 10 saccades under different prior beliefs and stimuli. First, we chose a set of prior beliefs that matched the true states of the world (**Figure 3A**). There is little change in the estimated sensory precision over time, and the posterior belief matches both the prior and the true states. We then tested the case for which the prior and the true states are different. **Figure 3B** shows a prior belief with the same content as 3a, but held with a lower degree of confidence (i.e., the prior belief is less precise). Again, there is little change in sensory precision, but now the posterior reflects the true states and not the prior. In other words, the sensory likelihood dominates perceptual inference. In **Figure 3C**, we simulate another mismatch between the prior and sensory evidence. This time, the prior belief is held with a high degree of confidence (i.e., a very precise prior), and this dominates inference. The posterior belief matches the prior, and is inconsistent with the sensory data sampled. The conflict between the prior and the sensory evidence is resolved in this case by a decrease in the precision associated with the contradictory locations. Heuristically, if a prior belief is held very confidently, evidence to the contrary is disregarded or ignored.

We have demonstrated that excessively precise prior beliefs lead to a compensatory decrease in the precision of the likelihood

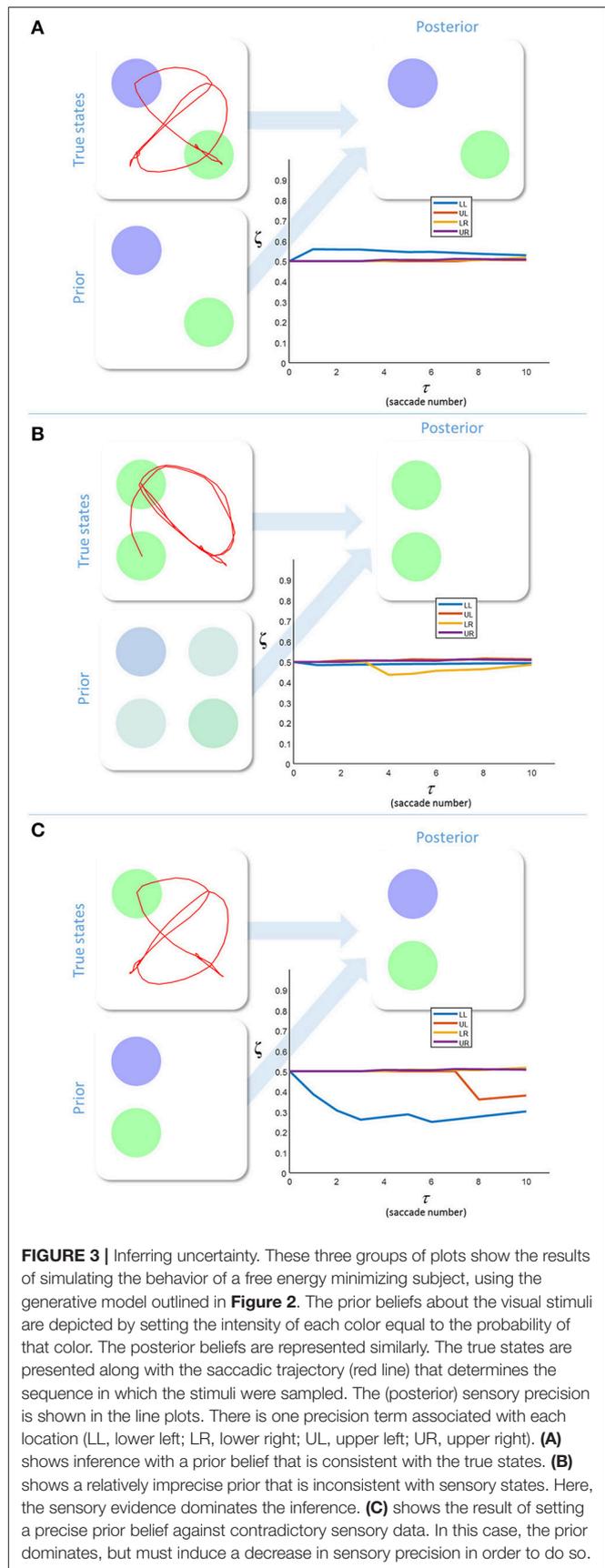


FIGURE 3 | Inferring uncertainty. These three groups of plots show the results of simulating the behavior of a free energy minimizing subject, using the generative model outlined in **Figure 2**. The prior beliefs about the visual stimuli are depicted by setting the intensity of each color equal to the probability of that color. The posterior beliefs are represented similarly. The true states are presented along with the saccadic trajectory (red line) that determines the sequence in which the stimuli were sampled. The (posterior) sensory precision is shown in the line plots. There is one precision term associated with each location (LL, lower left; LR, lower right; UL, upper left; UR, upper right). **(A)** shows inference with a prior belief that is consistent with the true states. **(B)** shows a relatively imprecise prior that is inconsistent with sensory states. Here, the sensory evidence dominates the inference. **(C)** shows the result of setting a precise prior belief against contradictory sensory data. In this case, the prior dominates, but must induce a decrease in sensory precision in order to do so.

distribution. Given the association between sensory precision and cholinergic signals (Dayan and Yu, 2001; Yu and Dayan, 2002; Vossel et al., 2014; Marshall et al., 2016), this provides a possible mechanism for the decrease in activity in the nucleus basalis of Meynert in several neurodegenerative disorders (Candy et al., 1983). Of these, the synucleinopathies, including Lewy body dementia, show an especially dramatic decrease in cholinergic signaling (Perry et al., 1994), and are associated with false visual inferences (i.e., hallucinations). These are, by definition, the imposition of prior beliefs on perception in the absence of supportive sensory evidence. Furthermore, decreased sensory gain means a smaller response to visual stimulation (**Figure 1**, first equation), consistent with the combination of reduced occipital cholinergic activity (Kuhl et al., 1996) and occipital hypometabolism (Lobotesis et al., 2001; Heitz et al., 2015) in synuclein disorders.

The above raises an important question: What is the source of the abnormal prior beliefs in conditions such as Lewy body dementia? We have previously argued that pathological prior beliefs might arise through anatomically defined vascular lesions (Parr and Friston, 2017b). Here, too, we can appeal to the anatomical distribution of the lesions to try to understand the relationship between tissue pathology and computational (network level) dysfunction. Lewy body pathology occurs in many brain regions, but it is their presence in parts of the temporal lobe that is associated with visual hallucinations (Harding et al., 2002). This leads us to consider the ventral visual hierarchies and their computational homologs.

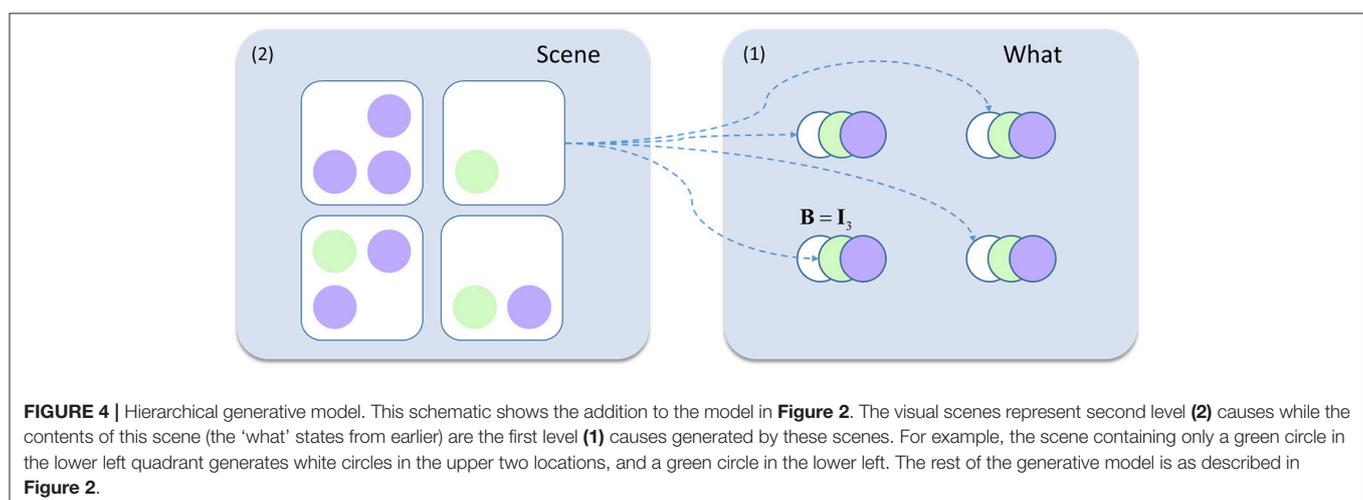
VISUAL HIERARCHIES

The visual system, like other sensory systems, is known to be hierarchically organized (Desimone et al., 1985; Zeki and Shipp, 1988; Felleman and Van Essen, 1991; Markov et al., 2013). We have previously appealed to this hierarchical structure to model reading (Friston et al., 2017c) and visual working memory tasks (Parr and Friston, 2017d). We now draw upon the same idea

to account for the source of the prior beliefs above, and to show how inaccurate but highly precise beliefs can develop. The visual system is organized into two broad hierarchical streams. These are the ventral (what) and the dorsal (where) pathways (Ungerleider and Haxby, 1994). It is the former that is of relevance here, as it leads from the occipital cortex to the temporal cortex, and represents stimulus identity at increasing levels of abstraction. While regions earlier in this pathway tend to respond to simple visual features (Hubel and Wiesel, 1959), later regions are selective for more complex visual objects (Valdez et al., 2015) or scenes (Epstein et al., 1999), constructed from lower level features. This is very important in accounting for the phenomenology of visual hallucinations in neurodegenerative conditions, as hallucinatory components of the percept appear in a consistent and plausible way in the context of the scene. This implies there is no impairment in scene construction *per se*. Instead, it is the wrong scene that is constructed. Crucially, this suggests hallucinated scenes are constructed based upon hierarchical principles, leading to the integration of a false percept in a way that is contextualized by the rest of the scene. This does not imply any impairment in the posterior precision of the overall percept.

To account for this hierarchical structure, we can augment our generative model so that visual stimuli (the “what” panel in **Figure 2**) are themselves generated by “scene” states. Both the “what” and the “scene” variables are types of hidden state. We refer to the former as a “first level” and the latter as a “second level” state. The second level is much simpler in this case (**Figure 4**), as there is only one type of state with no policies. Furthermore, all transitions at the second level are taken to be identity matrices, expressing the belief that the scene remains constant over time. This type of generative model allows the first level (empirical) priors to be generated by the second level. While this generative model is too abstract to map directly to the real visual system, this type of hierarchy does express cardinal features of the organization of the ventral visual stream.

Importantly, although inference about a visual feature can be performed within a given fixation, it takes a multiple saccades to



make inferences about the scenes at the second level. This implies that beliefs at this level should be updated more slowly (Friston et al., 2017c), consistent with the slower response properties higher in sensory hierarchies (Hasson et al., 2008; Kiebel et al., 2008; Murray et al., 2014).

Figure 5 shows the inferences made when we simulate responses with this hierarchical model. The upper part of the figure shows the beliefs about each of the second level states through time. At the start, the second level beliefs are combined (weighted by their probabilities) to generate an empirical prior at the first level. In both **Figures 5A,B**, this prior is relatively imprecise. A sequence of 5 saccades is performed, and the

observations made are used to refine the first level posteriors in exactly the same way as in **Figure 3**. These posterior beliefs are used to update the second level beliefs. These then generate a new empirical prior and this sequence repeats. The sensory precision is reset to its prior value whenever a new empirical prior is set (at the start of each sequence of 5 saccades).

The reason the precision re-sets to its prior value periodically is due to the separation of temporal scales inherent in the generative model. This is analogous to processes like reading, for which a sentence provides a high level context linking sequential words. Letters in one word only inform inferences about the next word via sentence-level representations. This means all lower

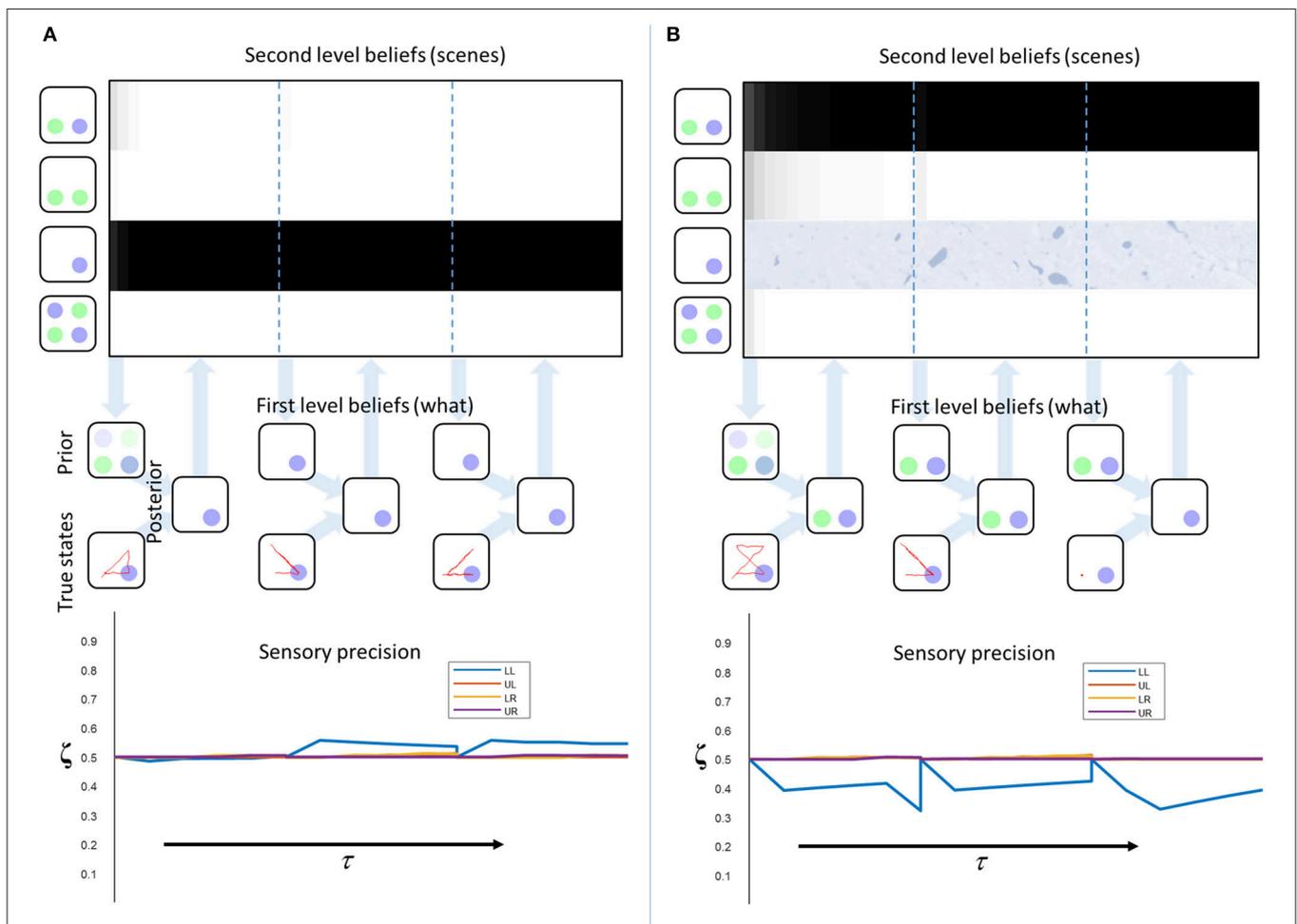


FIGURE 5 | Empirical priors and pathology. These plots illustrate the evolution of beliefs about second level states, first level states, and sensory precision. The upper plots show the beliefs about scenes over time. Each row of these represents a given scene (indicated by the images on the left). The shading indicates the belief that this is the scene responsible for the sensory input. Black indicates a belief that the probability is 1, white indicates 0. The descending arrows represent the computation of a first level empirical prior from the second level beliefs. A new empirical prior is generated after every 5 saccades (demarcated by dashed blue lines). The empirical prior and the sensory consequences of saccadic exploration combine to form first level posterior beliefs (exactly as in **Figure 3**). The beliefs from each set of 5 fixations are used to update the second level beliefs (ascending arrows). The lower plots show the beliefs about the sensory precision, aligned to the beliefs at the higher level. The precision is reset at the vertical dotted lines. **(A)** shows “healthy” second level priors that associate an equal probability to each scene at $\tau = 0$. Under these priors, the correct scene is inferred and the consistency between priors and sensory data leads to an increase in sensory precision. **(B)** shows the same model but with the prior probability of the third scene set to zero to simulate the loss of this neuron (or neuronal population). Here, the conflict between priors and sensory evidence leads to a decrease in precision. This also demonstrates the importance of action in perception as, at the final time-step, consistently fixating on the lower left location leads to a correct percept. This illustrates the point that collecting more data can compensate for the diminished precision of those data.

level representations are set to their (empirical) prior values every time we move from one word to the next. In our setting, the same is true of all lower level representations, including the precision. In a more complex model, it would be possible to condition the prior belief about the precision upon slowly changing variables at the higher level. While beyond the material presented in this paper, this would allow inferences about the precision to transcend the time-scale of the lower level.

Figure 5A illustrates this process in a model with “healthy” second level priors. There is a very rapid inference that the third scene is the most likely cause at the second level, and the first level beliefs, following saccadic interrogation, invariably match the true states. A moderate increase in estimated precision occurs under the second empirical prior, because confident prior beliefs match the sensory inputs. Note that the start location is in the lower left, so there is a larger effect for the precision at this location. This illustrates the fact that the prior has a greater influence at the start of the trial, where fewer observations have been made. **Figure 5B** shows a simulated synucleinopathy. Neuronal loss (or disconnection) high in the ventral stream has been simulated by setting the second level prior belief for one of the scenes (the correct one) to zero. This is as if we had removed the neuron that represents this second level hypothesis. Interestingly, this does not impede the formation of confident (false) first level empirical priors. As we saw earlier, these induce a decrease in the estimated sensory precision, and false perceptual inference. This demonstrates that pathology high in ventral visual hierarchies can, in principle, induce changes in distant brain areas—something that has been characterized in terms of a functional or dynamic diaschisis (Price et al., 2001; Carrera and Tononi, 2014). The idea that damage to a neuronal population preserves the confidence in the beliefs they represent may seem counterintuitive. However, even with the loss of neurons representing the correct inference, there is still a clear “best” explanation at the level of scenes. This leads to confident posterior beliefs about the scene, giving rise to confident (but incorrect) empirical prior beliefs about the contents of that scene. This is further facilitated by the permissive decrease in sensory precision.

This result recapitulates the idea that, for hallucinations to occur, prior beliefs must be held with a high degree of confidence (precision) relative to that associated with contradictory sensory evidence. This has previously been demonstrated in the context of auditory hallucinations in schizophrenia (Adams et al., 2013). Our account does, however, provide a different perspective on the initial computational insult. While this has previously been formulated as a false prior belief that something is present, we have demonstrated that hallucinations may be induced by a false prior that a given scene is not a good explanation for sensory data. This forces the brain to resort to an alternative explanation, associated with other, spurious, perceptual content.

COMPUTATIONAL NEUROPATHOLOGY

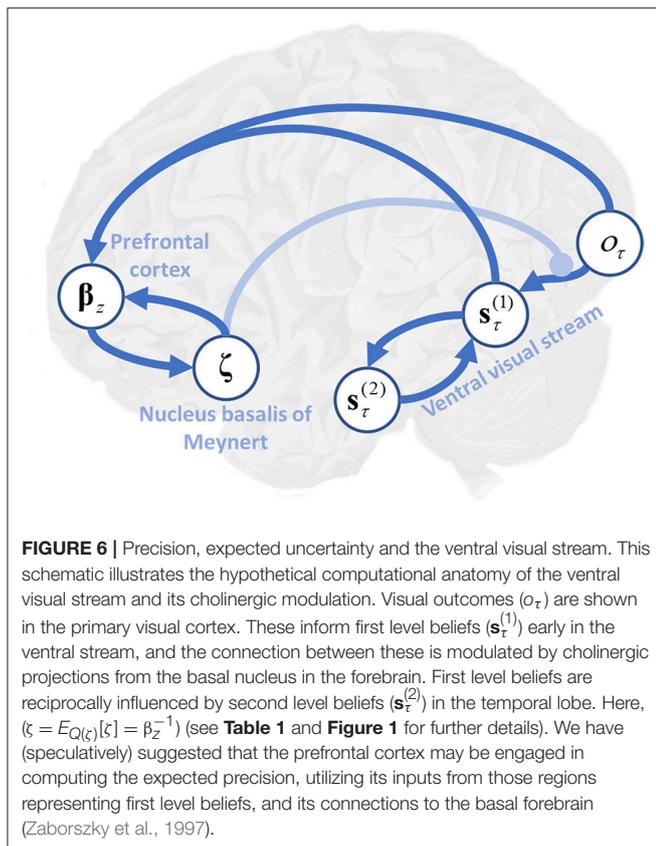
In the above, we have presented a model that relates temporal lobe pathology to the development of complex

visual hallucinations and reduced cholinergic signaling to the occipital cortex. Crucially, although the primary pathology only affects temporal components of the simulated network, its computational consequences are felt throughout the brain via a dynamic diaschisis (from Greek *διασχισις* meaning “shocked throughout”). This type of account is necessary in explaining the patterns of diaschisis observed in neuropathological processes.

The synucleinopathies [including Lewy body dementia, Parkinson’s disease, and Multiple system atrophy (Tsuboi and Dickson, 2005; McCann et al., 2014)] provide important examples that illustrate the need to connect tissue pathology to computational dysfunction. Despite the presence of physiological changes in the occipital cortices (Kuhl et al., 1996), and visual symptoms (McKeith et al., 2004; Weil et al., 2016), the histopathological processes in these disorders tend not to affect occipital cortex directly (Khundakar et al., 2016). While impaired dopamine signaling to the cortex in these disorders might contribute, occipital regions tend to receive relatively few dopaminergic projections (Javoy-Agid et al., 1989). The absence of these processes in primary visual areas, and the association between visual symptoms and temporal lobe Lewy bodies (Harding et al., 2002), calls for an explanation of physiological changes in the former in terms of their computational relationship to the latter.

We note that, for prior beliefs to dominate inference, the sensory precision must be low relative to the precision of prior beliefs. This means that hallucinations could occur with intact prior beliefs and a primary lesion to systems encoding sensory precision, or an increase in prior confidence without any change in sensory precision. Computationally, these are equivalent as they each change the balance of precisions in the same way. However, they are not necessarily biologically equivalent. The former implies a primary lesion to neuromodulatory systems that modulate synaptic gain in sensory cortices, while the latter implies damage to higher regions of cortex that provide empirical priors to sensory areas. In the context of visual hallucinations in Lewy body disease, both of these are present. While these may be two independent primary lesions, a simpler explanation would be that one is a downstream effect of the other. In this paper, we have suggested a mechanism by which damage to higher cortical areas could lead to disruption of synaptic gain in early visual cortex.

In short, the formal account of active inference or vision on offer here also provides an explanation in terms of a functional diaschisis—a dysfunction of one region as a consequence of a distant lesion (Price et al., 2001; Carrera and Tononi, 2014). **Figure 6** illustrates a plausible computational anatomy that could underwrite this account. While this anatomy is speculative, it serves to illustrate the importance of the functional interactions between brain regions to the understanding of neurological disease. Damage to temporal regions, representing second level beliefs, induces changes in the first level beliefs. This leads to inconsistencies between perceptual beliefs and sensory data, which down regulates cholinergic projections to the occipital cortex. Decreased cholinergic signaling uncouples beliefs about states from sensations they cause, facilitating hallucinations (Perry et al., 1991; O’Callaghan et al., 2017). As this model



would predict, treatment with cholinesterase inhibitors increases occipital blood flow, while attenuating visual hallucinations (Mori et al., 2006). A complementary, and more biophysically detailed, perspective on this is formulated in terms of impaired conductance (Tsukada et al., 2015) in the synapses between visual cortical neurons, and those in the ventral visual stream.

An influential model of recurrent complex visual hallucinations (Collerton et al., 2005) implicates these same regions, but makes the point that many other disorders involve similar changes. For example, cholinergic deficits are also associated with Alzheimer's disease (Minoshima et al., 2004), although to a lesser extent (Perry et al., 1994; Tiraboschi et al., 2000). Like those for Lewy body disease, pharmacological therapeutics have focused on correcting this neurochemical deficit (Lam et al., 2009). There is evidence to implicate changes in the temporal lobes in this disorder, as it tends to impact these structures early. However, this is typically more medial than in Lewy body dementia (Minoshima et al., 2002). Furthermore, it is unlikely that the cholinergic deficits in Alzheimer's disease are consequences of temporal lobe changes, as there is good evidence for a primary pathological insult to the nucleus basalis (Etienne et al., 1986; Samuel et al., 1994; Liu et al., 2015). This renders it improbable that this condition exhibits a similar set of computational deficits to those described above. The lower prevalence of visual hallucinations in Alzheimer's disease, despite overlapping pathological features with Lewy body disease, illustrates an important point. It is not sufficient to have

temporal lobe damage and cholinergic dysfunction to give rise to hallucinations. The interplay between the two is crucial in characterizing this type of diaschisis.

In this paper, we have focused upon false positive inferences (i.e., hallucinations). However, brain damage often leads to false negative inferences (i.e., agnosia) (Warrington and James, 1967). These manifest as a failure to perceive a stimulus, despite it being present. The approach we have described could be used to account for these phenomena in several ways. We outline these here, but emphasize that determining which of these best accounts for agnosia remains an open question that requires further investigation. The first way in which we could account for this is by setting a prior belief that a given object is present to zero. If the most probable alternative explanation is the absence of any object, this inference will result. It is important to distinguish this inference of absence from uncertain inferences, in which the presence or absence of an object cannot be inferred with any certainty. These could result from disconnections that render this object conditionally independent from sensory data in the generative model. This would ensure beliefs about the presence or absence of a given object would not depend upon these data. A third way in which certain stimuli may fail to enter into perceptual awareness is the failure to attend to certain kinds of stimuli, as in visual neglect (Halligan and Marshall, 1998). We have previously argued that this syndrome, in which stimuli on the left of space are ignored, depends upon a failure to actively engage with stimuli on the left (Parr and Friston, 2017b).

A number of outstanding questions are raised by the approach we have taken, which require empirical resolution. The first concerns our use of the term "visual features." We have illustrated a feature as the color of a circle in a given location, but this is not mandated by the mathematics used in our generative model. In principle, relevant features could be shape, luminance, contrast, or any other experienced attribute. We would need to present patients with a task like that illustrated above, but with different sorts of stimuli, to elucidate which of these afford the right level of description—and whether the ensuing responses are conserved over patients. The second question concerns the fixed parameters of the generative model—such as the prior belief about sensory precision. These are likely to be subject specific, but could be estimated from eye-tracking data collected during the above task (Mirza et al., 2018). Finally, we would hope to use these data to fit the prior beliefs (at the second level) of both patients and healthy participants (i.e., a quantitative and belief based computational phenotyping). We predict that patient data would offer greatest evidence for a reduced (Friston et al., 2016) version of the model that best fits healthy participants. If patient data afforded greater evidence for any other model, this would provide evidence against the hypotheses advanced in this paper. Given the associations between the parameters of our model and their biological substrates, this makes additional, falsifiable, predictions. First, if we were to perform this behavioral experiment combined with neuroimaging, we would predict a reduced effective connectivity between early visual areas and regions in the ventral visual stream that correlates with behaviorally derived precision parameters. Second, chemical neuroimaging to assess cholinergic

dysfunction should correlate with both behaviorally derived precision parameters, and decreases in effective connectivity in the visual system. Evidence against these associations would represent evidence against our hypothesis.

CONCLUSION

In the first part of this paper, we illustrated the computational mechanisms that could act to maintain the perceptual balance between prior beliefs and sensory evidence. We simulated inferences about the precision associated with the likelihood, and demonstrated that confident, but incorrect, prior beliefs cause a decrease in the expected sensory precision, and false perceptual inferences. In the second part, we asked what the computational mechanisms might be that give rise to pathological empirical priors, and motivated this through an appeal to the neurobiology of synuclein disorders. We described a plausible mechanism by which tissue pathology in higher visual areas could cause in occipital hypometabolism, cholinergic deficits, and visual hallucinations. Crucially, this calls upon the computational (network level) pathologies induced by regional synucleinopathies. This accounts for several empirical findings, including the association of temporal lobe changes with hallucinations in Lewy body disease and the improvement in hallucinations and occipital metabolism when these patients are treated with cholinesterase inhibitors. The ideas and

simulations presented here emphasize the importance of relating neuropathological processes to computational dysfunction to understand neurological disease.

SOFTWARE NOTE

Although the generative model changes from application to application, the belief updates described in this paper are generic and can be implemented using standard routines (here a customized version of `spm_MDP_VB_X.m`). These routines are available as Matlab code in the SPM academic software: <http://www.fil.ion.ucl.ac.uk/spm/>. Simulations of the sort reported above can be reproduced (and customized) via a graphical user interface by typing in `>> DEM` and selecting the “visual foraging” demo.

AUTHOR CONTRIBUTIONS

TP wrote and performed the simulations using an inversion scheme written by KF. TP, DB, PV, and KF contributed to conception and writing of the manuscript.

FUNDING

Rosetrees Trust (Award Number 173346) to TP. KF is a Wellcome Principal Research Fellow (Ref: 088130/Z/09/Z).

REFERENCES

- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., and Friston, K. J. (2013). The computational anatomy of psychosis. *Front. Psychiatry* 4:47. doi: 10.3389/fpsy.2013.00047
- Bartus, R. T., Dean, R. L., Beer, B., and Lippa, A. (1982). The cholinergic hypothesis of geriatric memory dysfunction. *Science* 217, 408–414. doi: 10.1126/science.7046051
- Beal, M. J. (2003). *Variational Algorithms for Approximate Bayesian Inference*. London, UK: University of London United Kingdom.
- Botvinick, M., and Toussaint, M. (2012). Planning as inference. *Trends Cogn. Sci.* 16, 485–488. doi: 10.1016/j.tics.2012.08.006
- Brown, H., and Friston, K. J. (2012). Free-energy and illusions: the cornsweet effect. *Front. Psychol.* 3:43. doi: 10.3389/fpsyg.2012.00043
- Candy, J. M., Perry, R. H., Perry, E. K., Irving, D., Blessed, G., Fairbairn, A. F., et al. (1983). Pathological changes in the nucleus of meynert in Alzheimer's and Parkinson's diseases. *J. Neurol. Sci.* 59, 277–289. doi: 10.1016/0022-510X(83)90045-X
- Carrera, E., and Tononi, G. (2014). Diaschisis: past, present, future. *Brain* 137, 2408–2422. doi: 10.1093/brain/awu101
- Collerton, D., Perry, E., and McKeith, I. (2005). Why people see things that are not there: a novel perception and attention deficit model for recurrent complex visual hallucinations. *Behav. Brain Sci.* 28, 737–757. doi: 10.1017/S0140525X05000130
- Contestabile, A. (2011). The history of the cholinergic hypothesis. *Behav. Brain Res.* 221, 334–340. doi: 10.1016/j.bbr.2009.12.044
- Dauwels, J. (2007). “On variational message passing on factor graphs. Information Theory ISIT 2007,” in *IEEE International Symposium on, IEEE (Nice)*.
- Dayan, P., and Yu, A. J. (2001). “ACh, uncertainty, and cortical inference” in *NIPS (Vancouver, BC)*.
- Desimone, R., Schein, S. J., Moran, J., and Ungerleider, L. G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Res.* 25, 441–452. doi: 10.1016/0042-6989(85)90069-0
- Disney, A. A., Aoki, C., and Hawken, M. J. (2007). Gain modulation by nicotine in macaque V1. *Neuron* 56, 701–713. doi: 10.1016/j.neuron.2007.09.034
- Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23, 115–125. doi: 10.1016/S0896-6273(00)80758-8
- Etienne, P., Robitaille, Y., Wood, P., Gauthier, S., Nair, N. P., and Quirion, R. (1986). Nucleus basalis neuronal loss, neuritic plaques and choline acetyltransferase activity in advanced Alzheimer's disease. *Neuroscience* 19, 1279–1291. doi: 10.1016/0306-4522(86)90142-9
- Feldman, H., and Friston, K. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215
- Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1
- Friston, K., and Buzsáki, G. (2016). The functional anatomy of time: what and when in the brain. *Trends Cogn. Sci.* 20, 500–511. doi: 10.1016/j.tics.2016.05.001
- Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., and Pezzulo, G. (2015). Active inference and epistemic value. *Cogn. Neurosci.* 6, 187–214. doi: 10.1080/17588928.2015.1020053
- Friston, K. J. (2017). Precision psychiatry. *Biol. Psychiatry Cogn Neurosci Neuroimaging* 2, 640–643. doi: 10.1016/j.bpsc.2017.08.007
- Friston, K. J., Daunizeau, J., Kilner, J., and Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biol. Cybern.* 102, 227–260. doi: 10.1007/s00422-010-0364-z
- Friston, K. J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., van Wijk, B. C. M., et al. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. *NeuroImage* 128(Suppl. C), 413–431. doi: 10.1016/j.neuroimage.2015.11.015
- Friston, K. J., Parr, T., and de Vries, B. (2017b). The graphical brain: belief propagation and active inference. *Netw. Neurosci.* 1, 381–414. doi: 10.1162/NETN_a_00018
- Friston, K. J., Rosch, R., Parr, T., Price, C., and Bowman, H. (2017c). Deep temporal models and active inference. *Neurosci. Biobehav. Rev.* 77, 388–402. doi: 10.1016/j.neubiorev.2017.04.009

- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2017a). Active inference: a process theory. *Neural Comput.* 29, 1–49. doi: 10.1162/NECO_a_00912
- Geisler, W. S., and Kersten, D. (2002). Illusions, perception and Bayes. *Nat. Neurosci.* 5, 508–510. doi: 10.1038/nn0602-508
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Oxford, UK: Houghton Mifflin.
- Gil, Z., Connors, B. W., and Amitai, Y. (1997). Differential regulation of neocortical synapses by neuromodulators and activity. *Neuron* 19, 679–686. doi: 10.1016/S0896-6273(00)80380-3
- Gregory, R. L. (1980). Perceptions as hypotheses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 290, 181–197. doi: 10.1098/rstb.1980.0090
- Halligan, P. W., and Marshall, J. C. (1998). Neglect of awareness. *Conscious. Cogn.* 7, 356–380. doi: 10.1006/ccog.1998.0362
- Harding, A. J., Broe, G. A., and Halliday, G. M. (2002). Visual hallucinations in Lewy body disease relate to Lewy bodies in the temporal lobe. *Brain* 125, 391–403. doi: 10.1093/brain/awf033
- Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1726–1731. doi: 10.1073/pnas.0610561104
- Hasson, U., Yang, E., Vallines, I., Heeger, D. J., and Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* 28, 2539–2550. doi: 10.1523/JNEUROSCI.5487-07.2008
- Heitz, C., Noblet, V., Cretin, B., Philippi, N., Kremer, L., Stackfleth, M., Hubele, F., et al. (2015). Neural correlates of visual hallucinations in dementia with Lewy bodies. *Alzheimers Res. Ther.* 7:6. doi: 10.1186/s13195-014-0091-0
- Hohwy, J. (2016). The self-evidencing brain. *Nous* 50, 259–285. doi: 10.1111/nous.12062
- Hubel, D. H., and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148, 574–591. doi: 10.1113/jphysiol.1959.sp006308
- Javoy-Agid, F., Scatton, B., Ruberg, M., L'Heureux, R., Cervera, P., Raisman, R., Maloteaux, J. M., et al. (1989). Distribution of monoaminergic, cholinergic, and GABAergic markers in the human cerebral cortex. *Neuroscience* 29, 251–259. doi: 10.1016/0306-4522(89)90055-9
- Khundakar, A. A., Hanson, P. S., Erskine, D., Lax, N. Z., Roscamp, J., Karyka, E., et al. (2016). Analysis of primary visual cortex in dementia with Lewy bodies indicates GABAergic involvement associated with recurrent complex visual hallucinations. *Acta Neuropathol. Commun.* 4:66. doi: 10.1186/s40478-016-0334-3
- Kiebel, S. J., Daunizeau, J., and Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Comput. Biol.* 4:e1000209. doi: 10.1371/journal.pcbi.100209
- Kuhl, D. E., Minoshima, S., Fessler, J. A., Ficaro, E., Wieland, D., Koeppe, R. A., et al. (1996). *In vivo* mapping of cholinergic terminals in normal aging, Alzheimer's disease, and Parkinson's disease. *Ann. Neurol.* 40, 399–410. doi: 10.1002/ana.410400309
- Lam, B., Hollingdrake, E., Kennedy, J. L., Black, S. E., and Masellis, M. (2009). Cholinesterase inhibitors in Alzheimer's disease and Lewy body spectrum disorders: the emerging pharmacogenetic story. *Hum. Genomics* 4:91. doi: 10.1186/1479-7364-4-2-91
- Lavine, N., Reuben, M., and Clarke, P. (1997). A population of nicotinic receptors is associated with thalamocortical afferents in the adult rat: laminar and areal analysis. *J. Compar. Neurol.* 380:175–190. doi: 10.1002/(SICI)1096-9861(19970407)380:2<175::AID-CNE3>3.0.CO;2-0
- Liu, A. K., Chang, R. C., Pearce, R., K. B., and Gentleman, S. M. (2015). Nucleus basalis of Meynert revisited: anatomy, history and differential involvement in Alzheimer's and Parkinson's disease. *Acta Neuropathol.* 129, 527–540. doi: 10.1007/s00401-015-1392-5
- Lobotesis, K., Fenwick, J. D., Phipps, A., Ryman, A., Swann, A., Ballard, C., et al. (2001). Occipital hypoperfusion on SPECT in dementia with Lewy bodies but not AD. *Neurology* 56, 643–649. doi: 10.1212/WNL.56.5.643
- Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., et al. (2013). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J. Comp. Neurol.* 522, 225–259. doi: 10.1002/cne.23458
- Marshall, L., Mathys, C., Ruge, D., de Berker, A. O., Dayan, P., Stephan, K. E., et al. (2016). Pharmacological fingerprints of contextual uncertainty. *PLoS Biol.* 14:e1002575. doi: 10.1371/journal.pbio.1002575
- McCann, H., Stevens, C. H., Cartwright, H., and Halliday, G. M. (2014). α -Synucleinopathy phenotypes. *Parkinson. Relat. Disord.* 20(Suppl. 1), S62–S67. doi: 10.1016/S1353-8020(13)70017-8
- McKeith, I., Mintzer, J., Aarsland, D., Burn, D., Chiu, H., Cohen-Mansfield, J., Dickson, D., et al. (2004). Dementia with Lewy bodies. *Lancet Neurol.* 3, 19–28. doi: 10.1016/S1474-4422(03)00619-7
- Minoshima, S., Foster, N. L., Petrie, E. C., Albin, R. L., Frey, K. A., and Kuhl, D. E. (2002). Neuroimaging in dementia with Lewy bodies: metabolism, neurochemistry, and morphology. *J. Geriatr. Psychiatry Neurol.* 15, 200–209. doi: 10.1177/089198870201500405
- Minoshima, S., Frey, K. A., Cross, D. J., and Kuhl, D. E. (2004). Neurochemical imaging of dementias. *Semin. Nuclear Med.* 34, 70–82. doi: 10.1053/j.semnuclmed.2003.09.008
- Mirza, M. B., Adams, R. A., Mathys, C., and Friston, K. J. (2018). Human visual exploration reduces uncertainty about the sensed world. *PLoS ONE* 13:e0190429. doi: 10.1371/journal.pone.0190429
- Mirza, M. B., Adams, R. A., Mathys, C. D., and Friston, K. J. (2016). Scene construction, visual foraging, and active inference. *Front. Comput. Neurosci.* 10:56. doi: 10.3389/fncom.2016.00056
- Moran, R. J., Campo, P., Symmonds, M., Stephan, K. E., Dolan, R. J., and Friston, K. J. (2013). Free energy, precision and learning: the role of cholinergic neuromodulation. *J. Neurosci.* 33, 8227–8236. doi: 10.1523/JNEUROSCI.4255-12.2013
- Mori, T., Ikeda, M., Fukuhara, R., Nestor, P. J., and Tanabe, H. (2006). Correlation of visual hallucinations with occipital rCBF changes by donepezil in DLB. *Neurology* 66, 935–937. doi: 10.1212/01.wnl.0000203114.03976.b0
- Motohiro, K., Sabina, P., Danielle, D., Danielle, R., Henri, C., Bernard, M., Yves, S., et al. (1987). Cortical Hypometabolism and its recovery following nucleus Basalis Lesions in Baboons: a PET study. *J. Cereb. Blood Flow Metab.* 7, 812–817. doi: 10.1038/jcbfm.1987.139
- Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., et al. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* 17, 1661–1663. doi: 10.1038/nn.3862
- O'Callaghan, C., Hall, J. M., Tomassini, A., Muller, A. J., Walpola, I. C., Moustafa, A. A., et al. (2017). Visual Hallucinations are characterized by impaired sensory evidence accumulation: insights from hierarchical drift diffusion modeling in Parkinson's Disease. *Biol. Psychiatry* 2, 680–688. doi: 10.1016/j.bpsc.2017.04.007
- Parr, T., and Friston, K. J. (2017a). The active construction of the visual world. *Neuropsychologia* 104, 92–101. doi: 10.1016/j.neuropsychologia.2017.08.003
- Parr, T., and Friston, K. J. (2017b). The computational anatomy of visual neglect. *Cereb. Cortex* 28, 777–790. doi: 10.1093/cercor/bhx316
- Parr, T., and Friston, K. J. (2017c). Uncertainty, epistemics and active inference. *J. R. Soc. Interf.* 14:136. doi: 10.1098/rsif.2017.0376
- Parr, T., and Friston, K. J. (2017d). Working memory, attention, and salience in active inference. *Sci. Rep.* 7:14678. doi: 10.1038/s41598-017-15249-0
- Perry, E. K. (1986). The cholinergic hypothesis—ten years on. *Br. Med. Bull.* 42, 63–69. doi: 10.1093/oxfordjournals.bmb.a072100
- Perry, E. K., Haroutunian, V., Davis, K. L., Levy, R., Lantos, P., Eagger, S., et al. (1994). Neocortical cholinergic activities differentiate Lewy body dementia from classical Alzheimer's disease. *Neuroreport* 5, 747–749. doi: 10.1097/00001756-199403000-00002
- Perry, E. K., McKeith, I., Thompson, P., Marshall, E., Kerwin, J., Jabeen, S., Edwardson, J. A., et al. (1991). Topography, extent, and clinical relevance of neurochemical deficits in Dementia of Lewy body type, Parkinson's disease, and Alzheimer's disease. *Ann. N. Y. Acad. Sci.* 640, 197–202. doi: 10.1111/j.1749-6632.1991.tb00217.x
- Price, C. J., Warburton, E., Moore, C., Frackowiak, R. S., and Friston, K. J. (2001). Dynamic diaschisis: anatomically remote and context-sensitive human brain lesions. *J. Cogn. Neurosci.* 13, 419–429. doi: 10.1162/08989290152001853
- Sahin, M., Bowen, W. D., and Donoghue, J. P. (1992). Location of nicotinic and muscarinic cholinergic and μ -opioid receptors in rat cerebral neocortex: evidence from thalamic and cortical lesions. *Brain Res.* 579, 135–147. doi: 10.1016/0006-8993(92)90752-U

- Samuel, W., Terry, R. D., DeTeresa, R., Butters, N., and Masliah, E. (1994). Clinical correlates of cortical and nucleus basalis pathology in Alzheimer dementia. *Arch. Neurol.* 51, 772–778. doi: 10.1001/archneur.1994.00540200048015
- Shipp, S. (2007). Structure and function of the cerebral cortex. *Curr. Biol.* 17, R443–R449. doi: 10.1016/j.cub.2007.03.044
- Terry, A. V., and Buccafusco, J. J. (2003). The cholinergic hypothesis of age and Alzheimer's disease-related cognitive deficits: recent challenges and their implications for novel drug development. *J. Pharmacol. Exp. Ther.* 306, 821–827. doi: 10.1124/jpet.102.041616
- Tiraboschi, P., Hansen, L. A., Alford, M., Sabbagh, M. N., Schoos, B., Masliah, E., et al. (2000). Cholinergic dysfunction in diseases with Lewy bodies. *Neurology* 54, 407–407. doi: 10.1212/WNL.54.2.407
- Tsuboi, Y., and Dickson, D. W. (2005). Dementia with Lewy bodies and Parkinson's disease with dementia: are they different? *Parkinson. Relat. Disord.* 11(Suppl. 1), S47–S51. doi: 10.1016/j.parkreldis.2004.10.014
- Tsukada, H., Fujii, H., Aihara, K., and Tsuda, I. (2015). Computational model of visual hallucination in dementia with Lewy bodies. *Neural Netw.* 62, 73–82. doi: 10.1016/j.neunet.2014.09.001
- Ungerleider, L. G., and Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Curr. Opin. Neurobiol.* 4, 157–165.
- Valdez, A. B., Papesh, M. H., Treiman, D. M., Smith, K. A., Goldinger, S. D., and Steinmetz, P. N. (2015). Distributed representation of visual objects by single neurons in the human brain. *J. Neurosci.* 35, 5180–5186. doi: 10.1523/JNEUROSCI.1958-14.2015
- Vossel, S., Bauer, M., Mathys, C., Adams, R. A., Dolan, R. J., Stephan, K. E., et al. (2014). Cholinergic stimulation enhances bayesian belief updating in the deployment of spatial attention. *J. Neurosci.* 34, 15735–15742. doi: 10.1523/JNEUROSCI.0091-14.2014
- Warrington, E. K., and James, M. (1967). Disorders of visual perception in patients with localised cerebral lesions. *Neuropsychologia* 5, 253–266. doi: 10.1016/0028-3932(67)90040-1
- Weil, R. S., Schrag, A. E., Warren, J. D., Crutch, S. J., Lees, A. J., and Morris, H. R. (2016). Visual dysfunction in Parkinson's disease. *Brain* 139, 2827–2843. doi: 10.1093/brain/aww175
- Yu, A. J., and Dayan, P. (2002). Acetylcholine in cortical inference. *Neural Netw.* 15, 719–730. doi: 10.1016/S0893-6080(02)0058-8
- Yu, A. J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692. doi: 10.1016/j.neuron.2005.04.026
- Zaborszky, L., Gaykema, R. P., Swanson, D. J., and Cullinan, W. E. (1997). Cortical input to the basal forebrain. *Neuroscience* 79, 1051–1078. doi: 10.1016/S0306-4522(97)00049-3
- Zeki, S., and Shipp, S. (1988). The functional logic of cortical connections. *Nature* 335, 311–317. doi: 10.1038/335311a0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Parr, Benrimoh, Vincent and Friston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

In this appendix, we outline the derivations for the update equations we have used in our simulation. For more technical accounts, please see (Friston et al., 2017a,b; Parr and Friston, 2017c).

Inference

The free energy for a given policy is given as

$$F(\pi) = E_Q[\ln Q(\tilde{s}|\pi) - \ln P(\tilde{o}, \tilde{s}|\pi, \zeta)]$$

If we are interested in finding $Q(s_\tau|\pi)$, we can omit all terms that are constant with respect to this

$$\begin{aligned} F(\pi) = & E_{Q(s_\tau|\pi)}[\ln Q(s_\tau|\pi)] - E_{Q(s_{\tau-1}|\pi)Q(s_\tau|\pi)}[\ln P(s_\tau|s_{\tau-1}, \pi)] \\ & - E_{Q(s_\tau|\pi)Q(s_{\tau+1}|\pi)}[\ln P(s_{\tau+1}|s_\tau, \pi)] \\ & - E_{Q(s_\tau|\pi)}[\ln P(o_\tau|s_\tau, \zeta)] \end{aligned}$$

In terms of sufficient statistics, this is

$$\begin{aligned} \mathbf{F}_\pi = & \mathbf{s}_{\pi\tau} \cdot (\ln \mathbf{s}_{\pi\tau} - \ln \mathbf{B}_{\pi\tau-1} \mathbf{s}_{\pi\tau-1} - \ln \mathbf{B}_{\pi\tau} \cdot \mathbf{s}_{\pi\tau+1} \\ & - \zeta \ln \mathbf{A} \cdot \mathbf{o}_\tau) \end{aligned}$$

We can then set up a gradient descent on this using an auxiliary variable, $\mathbf{v}_{\pi\tau}$:

$$\begin{aligned} \mathbf{s}_{\pi\tau} &= \sigma(\mathbf{v}_{\pi\tau}) \\ \dot{\mathbf{v}}_{\pi\tau} &= -\nabla_{\mathbf{s}} \mathbf{F}_\pi = \mathbf{e}_{\pi\tau} \end{aligned}$$

$$\mathbf{e}_{\pi\tau} = \ln \mathbf{B}_{\pi\tau-1} \mathbf{s}_{\pi\tau-1} + \ln \mathbf{B}_{\pi\tau} \cdot \mathbf{s}_{\pi\tau+1} + \zeta \ln \mathbf{A} \cdot \mathbf{o}_\tau - \ln \mathbf{s}_{\pi\tau}$$

Here, σ is a softmax (normalized exponential) function.

Planning

In order to treat planning as inference (Botvinick and Toussaint, 2012), we must define the free energy to include beliefs about policies. This is:

$$\mathbf{F} = \boldsymbol{\pi} \cdot (\mathbf{F}_\pi + \mathbf{G}_\pi + \ln \boldsymbol{\pi})$$

If we set the gradient of the free energy to zero, we find:

$$\nabla_{\boldsymbol{\pi}} \mathbf{F} = 0 \Leftrightarrow \boldsymbol{\pi} = \sigma(-\mathbf{F}_\pi - \mathbf{G}_\pi)$$

Precision

We can include beliefs about the precision of the likelihood by defining $\bar{\mathbf{A}} = \frac{1}{\zeta} \mathbf{A}^\zeta$. We then express beliefs about ζ as gamma distributions (for which $\zeta = E_Q[\zeta] = \boldsymbol{\beta}_z^{-1}$), and write down the free energy as:

$$\begin{aligned} \mathbf{F} = & -\boldsymbol{\pi} \cdot \sum_{\tau} \mathbf{s}_{\pi\tau} \cdot (\ln \bar{\mathbf{A}} \cdot \mathbf{o}_\tau + \ln \mathbf{B}_{\pi\tau} \cdot \mathbf{s}_{\pi\tau+1}) + \ln \boldsymbol{\beta}_z \\ & - \ln \boldsymbol{\beta}_z + \zeta \boldsymbol{\beta}_z \\ \nabla_{\zeta} \mathbf{F} = 0 \Leftrightarrow & \boldsymbol{\beta}_z = \boldsymbol{\beta}_z - \sum_{\tau} \mathbf{s}_{\pi\tau} \cdot \nabla_{\zeta} \ln \bar{\mathbf{A}} \cdot \mathbf{o}_\tau \\ \dot{\boldsymbol{\beta}}_z = & \boldsymbol{\beta}_z - \sum_{\tau} \mathbf{s}_{\pi\tau} \cdot \nabla_{\zeta} \ln \bar{\mathbf{A}} \cdot \mathbf{o}_\tau - \boldsymbol{\beta}_z \end{aligned}$$