A considerable number of recent experimental and computational studies suggest that subtle impairments of excitatory to inhibitory balance or regulation are involved in many neurological and psychiatric conditions. The current paper aims to relate, specifically and quantitatively, excitatory to inhibitory imbalance with psychotic symptoms in schizophrenia.

Considering that the brain constructs hierarchical causal models of the external world, we show that the failure to maintain the excitatory to inhibitory balance results in hallucinations as well as in the formation and subsequent consolidation of delusional beliefs. Indeed, the consequence of excitatory to inhibitory imbalance in a hierarchical neural network is equated to a pathological form of causal inference called ‘circular belief propagation’. In circular belief propagation, bottom-up sensory information and top-down predictions are reverberated, i.e. prior beliefs are misinterpreted as sensory observations and vice versa. As a result, these predictions are counted multiple times. Circular inference explains the emergence of erroneous percepts, the patient’s overconfidence when facing probabilistic choices, the learning of ‘unshakable’ causal relationships between unrelated events and a paradoxical immunity to perceptual illusions, which are all known to be associated with schizophrenia.

Keywords: Bayesian; network; belief; predictive coding; hallucinations; delusions
Abbreviations: GABA = gamma-aminobutyric acid; LOR = log-odd ratio; NMDA = N-methyl-D-aspartate

Introduction

Schizophrenia is a devastating mental disorder that afflicts ~1% of the world’s population during their lifetime (McGrath et al., 2008). The disorder can be defined as a complex array of non-specific symptoms that can be clustered into at least two to three different clinical dimensions (Crow, 1981; Liddle, 1987): the ‘positive symptoms’, which include reality distortions such as hallucinations, delusions and other bizarre beliefs; the ‘negative symptoms’, comprising psychomotor poverty, blunt affects, social withdrawal and lack of motivation; and ‘cognitive disorganization’, encompassing dissociation, illogical speech and behaviours.

Although the exact underlying pathophysiological mechanisms for this disorder remain enigmatic, a substantial amount of physiological (Uhlhaas and Singer, 2010; Mulert et al., 2011) and post-mortem (Lewis et al., 2005) evidence converges to suggest an impairment of gamma-aminobutyric acid (GABA) transmission or of N-methyl-D-aspartate (NMDA) receptor plasticity in schizophrenia (Stephan et al., 2009). These findings have contributed to a re-conceptualization of this disorder as a possible disruption in the neural excitatory to inhibitory balance (O’Donnell, 2011).

Evidence of an excitatory to inhibitory imbalance in schizophrenia includes recent support for an alteration in the inhibitory GABAergic transmission of cortical microcircuits, comprising
dopaminergic, glutamatergic and other GABAergic pathways (Lisman et al., 2008). Using rodent models, the experimental blockage of parvalbumin interneurons (Bartos et al., 2002), or the suppression of their activity using optogenetic methods (Sohal et al., 2009) was shown to induce significant reductions in θ-oscillations, an impairment replicated several times in schizophrenia (for review see Uhlfahs and Singer, 2010). Using induced pluripotent stem cell technology, another group evidenced general defects in neuronal connectivity, also compatible with the excitatory to inhibitory imbalance hypothesis (Brennand et al., 2011). By creating differentiated neurons from patients with schizophrenia, these authors effectively demonstrated impairments in neuronal glutamate receptor expression.

This excitatory to inhibitory imbalance was shown to specifically occur when mesocortical dopaminergic (dopamine) pathways were activated, resulting in schizophrenia-related phenomena in rodents (Gruber et al., 2010). This phenomenon of GABAergic interneurons recruitment by dopamine remains inefficient during the juvenile period (Tseng and O’Donnell, 2007), suggesting that insults of this system would only manifest at later developmental stages (Insel, 2010). This mechanism supports the classical emergence period of first episodes of schizophrenia during adolescence or early adulthood. Interestingly, the reduced GABAergic inhibition observed in patients with schizophrenia was found to be related to perceptual deficits (Yoon et al., 2010), such as reduced vulnerability to contrast-contrast illusions (Dakin et al., 2005).

Similarly to reduced GABAergic transmission, the downregulation of NMDA receptors may also affect excitatory to inhibitory balance. In animal models, global NMDA receptor hypofunction causes an increase in intrinsic pyramidal cell excitability and a selective disruption of parvalbumin-expressing interneurons (Gandal et al., 2012). Psychotomimetic models of psychosis, especially those based on ketamine, also support the excitatory to inhibitory imbalance hypothesis. Schizophrenia-like symptoms have been described for ketamine use (Krystal et al., 2005) and in autoimmune anti-NMDA receptor encephalitis (Dalmau et al., 2008). Specifically, the drug was shown to bind D2R and induce striatal dopamine release (Smith et al., 1998). Corlett et al. (2009a) suggested that under ketamine, the subject may experience both perceptual aberrations (due to AMPA upregulation) and a reduced capacity to accommodate and ignore these aberrations (due to NMDA blockade).

However, the mechanism by which excitatory to inhibitory dysfunctions at the cellular level relates to the abovementioned heterogeneous clinical symptoms of schizophrenia remains unclear. We consider here that the specific dimensions of schizophrenia are more reliable constructs than the whole disorder and specifically focus on the potential underlying mechanisms for positive symptoms, which have been shown to cause severe distress in >70% of affected individuals (Andreasen and Flaum, 1991). The question may thus be simplified as to how a failure of the excitatory to inhibitory balance regulation may result in erroneous perceptions (i.e. hallucinations) or inflexible beliefs that are not directly corroborated by the senses (i.e. delusions). Intuitively, less inhibition may result in too much excitation or propagation within cortical circuits, but this hypothesis on its own seems too reductionistic. Here, we propose moving beyond this relatively vague concept by relating impaired inhibition to aberrant belief formation, considering hierarchical Bayesian inference as a basic principle of brain function (Friston, 2008).

The Bayesian formalism interprets perception as the inference on what may have caused the sensory observations, optimally combining information from sensory receptors with predictions based on prior knowledge (von Helmholtz, 1866; Knill and Richards, 1996; Doya et al., 2007). This framework accounts for erroneous percepts in a non-pathological context, such as perceptual illusions (Geisler and Kersten, 2002; Weiss et al., 2002). Imbalances between the gain (or confidence) attributed to bottom-up sensory information compared with top-down prior knowledge provide a simple account not only for the positive symptoms of schizophrenia (Friston, 2005a), but also for sensory integration deficits in autistic spectrum disorders (Pellicano and Burr, 2012) and functional motor and sensory symptoms in hysteria (Edwards et al., 2012).

Interestingly, the role of inhibitory loops in the brain mechanisms that address ambiguity was recently experimentally strengthened by van Loon et al. (2013). These authors showed that the modulation of GABA_A receptors by the administration of a benzodiazepine to healthy subjects, induced changes in both the excitatory to inhibitory balance, which were measurable through magnetic resonance spectroscopy of the visual cortex, and the dynamics of bistable perception.

In the present paper, we show that the maintenance of a precise excitatory to inhibitory balance is crucial to the implementation of Bayesian inference in a recurrent biological neural network, and, in particular, in the proper adjustment of the effective gain of (or confidence in) feed-forward sensory information versus top-down priors. A dominance of excitation results in a pathological form of inference called ‘circular belief propagation’, in which ‘bottom-up’ and ‘top-down’ messages are reverberated and taken into account multiple times. We show how this concept can parsimoniously account for hallucinations and delusions while remaining compatible with recent theoretical and empirical findings. Overall, the presented theory is quantitative and thus testable through behavioural or neurophysiological paradigms.

Materials and methods

Hierarchical inference and excitatory to inhibitory balance

The brain can be seen as constructing a representation of the world as a hierarchy of progressively more abstract causes. As a simplified toy example, let us imagine the processing of evidence for the colour ‘green’ in a visual scene. Green could be caused by the presence of a leaf. Leaves are themselves predicted by the presence of a tree. Meanwhile, the presence of a tree could be due to the observer being located in a forest (Fig. 1A). Hierarchical Bayesian inference interprets sensory inputs by finding the set of causes that best accounts for sensory observations. This type of inference is a result of optimally combining two types of information: ‘bottom-up’ (magenta arrows in Fig. 1) and ‘top-down’ (violet arrows in Fig. 1).
Bottom-up processing sends information from low-level sensory observations to higher levels. For example, detecting the colour green increases the probability of a leaf and, in turn, of a tree. Top-down processing sends prior expectations from high-level causes to their lower-level sensory consequences. For example, prior knowledge of being located in a forest increases the expectations of observing a tree, a leaf and the colour green. In the Bayesian formalism, top-down messages that correspond to sensory likelihoods (magenta) and descending messages that correspond to prior expectations (violet) are propagated only once in each direction, appropriately sharing prior and sensory information among all of the neurons.

Figure 1

**Figure 1** Belief propagation in a hierarchical neural network. (A) Toy example of a hierarchical causal model with six hidden variables represented as ‘nodes’. Arrows represent causal relationships. Sensory evidence is provided at the bottom and prior evidence at the top of the hierarchy. (B) Neural network that implements causal inference in this toy example. $M_X$ represents the message sent by node $x$ to node $y$. The overall flow of messages in the network is under the control of inhibitory loops (yellow lines: feed-forward connections; blue lines: feedback connections; green circles: upward inhibitory loops; black circles: downward inhibitory loops). (C) Belief propagation in the intact and impaired networks. **Left:** normal network that performs belief propagation. Because of intact inhibitory loops, ascending messages correspond to prior likelihoods and descending messages correspond to prior expectations. **Right:** impaired network without inhibitory loops. Top-down and bottom-up messages are uncontrollably propagated multiple times in each direction, which results in the multiple overcounting of the same redundant sensory and prior information. (D) Consequences of circular belief propagation. The belief obtained through iterating the network dynamics is out of proportion with the available sensory and prior evidence. This construct results in overconfidence (i.e. probabilities only slightly above 0.5 are perceived as being close to 1, and probabilities just below 0.5 are perceived as being close to zero).

Bottom-up processing sends information from low-level sensory observations to higher levels. For example, detecting the colour green increases the probability of a leaf and, in turn, of a tree. Top-down processing sends prior expectations from high-level causes to their lower-level sensory consequences. For example, prior knowledge of being located in a forest increases the expectations of observing a tree, a leaf and the colour green. In the Bayesian formalism, top-down information corresponds to the prior, and bottom-up information corresponds to the likelihood of sensory inputs. Because this hierarchy of progressively more abstract representations is loosely reflected in the hierarchy of sensory-to-associative brain areas, bottom-up and top-down processing can be interpreted as ‘feed-forward’ and ‘feedback’ neural processing (Fig 1A and B; Lee and Mumford, 2003; Lochmann and Deneve, 2011).

The necessity of combining bottom-up and top-down information raises a crucial (albeit usually disregarded) issue: that of disentangling ‘new’ and ‘old’ information in the constant flow of neural inputs. The simultaneous presence of feed-forward and feedback connections creates multiple potential loops in the corresponding neural network. Without efficient control mechanisms, top-down information, after descending the hierarchy, could be reverberated back up, resulting in top-down expectations being misunderstood as ‘real’ sensory observations. The opposite is also true: bottom-up sensory information could be reverberated back down and misinterpreted as top-down expectations. As a result, sensory evidence and priors would be propagated many times in the network and accounted for many times rather than only once. This problem is illustrated in Fig. 1C. Considering our toy example, walking in the forest can generate expectations for leaves and the colour green. In the presence of loops, the system could misinterpret the resulting internally generated expectations as external sensory information, causing the aberrant perception of green leaves, even during winter.

The cortex is a highly recurrent structure in which most long-range connections are excitatory, and a large proportion of these connections are top-down (Douglas et al., 1995). In these conditions, circular inference through a reverberation of activity appears to be unavoidable (Fig. 1C). Fortunately, in spite of the apparent prevalence of excitatory connections, the balance between excitation and inhibition is tightly maintained at all levels of neural processing, from single dendrites and neurons to entire brain areas (Wehr and Zador, 2003; Hensch and Fagiolini, 2004; Shu et al., 2007; Okun and Lampl, 2008). Moreover, stimulus-driven or spontaneous temporal modulations of excitatory and inhibitory drive are temporally correlated (Shu et al., 2003; Wehr and Zador, 2003). In a recurrent network with large numbers of feed-forward and feedback excitatory connections, excitatory to inhibitory balance only exists if each excitatory loops is compensated by the presence of an equally strong inhibitory loop (Fig. 1B). As shown in the next section and illustrated in Fig. 1B and C, such
inhibitory loops can predict and cancel out redundant information that reverberates through both top-down and bottom-up excitatory loops, avoiding circular inferences. This mechanism ensures that sensory information and prior expectations are accounted for only once, and not multiple times, in spite of the strong level of recurrence in the network.

We can now start to see how perturbations of the excitatory to inhibitory balance result in aberrant belief formation. Indeed, scaling down these inhibitory loops (or scaling up excitation without an equivalent increase in inhibition) results in circular inferences and the generation of erroneous percepts. Moreover, these predictions are quantifiable (regardless of their specific neural implementation) because lowering the strength of inhibition relative to excitation results in circular propagation of beliefs in the corresponding normative causal model, thus predicting specific perceptual and behavioural impairments. In the following section, we present, in more technical terms, the belief propagation algorithm, its neural interpretation and the proposed role of inhibitory loops.

**Belief propagation and its neural analogy**

For the sake of simplicity, we present here a belief propagation algorithm specific to binary causes. ‘Binary’ implies that the causes can take only two states, for example, a leaf can be either present or absent. Moreover, we assume that each of the binary variables has at most one cause in the layer above. We chose binary variables for illustration purposes and because of their direct interpretability in terms of decisions, choices and confidence. We assumed only one parent to keep equations in the main text as simple as possible. However, our approach readily generalizes to other hierarchical causal models, as shown in the Supplementary material.

The belief propagation algorithm works by propagating ‘beliefs’, which can be understood as the current estimate for the probability of each cause. To facilitate an interpretation in terms of neural processing, we express these beliefs as a log-odd ratio (LOR). As an example, the LOR associated with ‘tree’ is the log of the ratio between the probability that the tree is present and the probability that the tree is absent. If the tree is as likely to be present as absent (and thus, the state of ‘tree’ is completely uncertain), its LORlog(1) = 0. On the other hand, if the tree is three times more likely to be present than absent, its LOR is log(3) ≈ 1.1. More generally, LORs are large and positive if there is a high confidence in state ‘1’ (tree present), they are large and negative if there is a high confidence in state ‘0’ (tree absent), and they are close to zero when the state is uncertain. We call the LORs of cause i at step nBnij.

To illustrate the consequences of circular inference, we used a binary causal graph with four causal layers and six variables (Fig. 1A and B). Except when otherwise stated, we set the conditional probabilities to p0 = 0.9 (the probability that j is present given that i is present in 0.9) and p0 = 0.1 (the probability that j is absent given that i is absent in 0.1) for all of the connected causal pairs. For example, the presence of a tree causes the presence of a leaf with a probability of 0.9, while the probability of observing a leaf in the absence of a tree is 0.1. Note that the qualitative effects reported here are not dependent on the exact setting of p0 and p0.

Sensory information is provided as messages clamped in the sensory layer. For example, M0 = 3 implies that there is strong sensory evidence for the colour green (i.e. the probability of the colour green based on sensory information alone is e/e ≈ 0.95). M0 = 0 implies that there is no sensory information available about the presence or absence of a stem (i.e. the probability of the presence of a stem based on sensory information alone is 0.5). Prior expectations are implemented as a constant message clamped in the top layer. For example, M0 = −1 implies that the prior probability of being in the forest (before taking into account sensory information) is 0.5. We assumed constant, fixed priors and likelihoods for the sake of simplicity. Readers should note that the results are similar when ‘noisy’ sensory evidence is provided as random samples from the corresponding likelihood distributions.

Beliefs, Bnij, can be computed iteratively by propagating messages between causal nodes. These messages convey predictions of one node by another. For example, the message sent from ‘leaf’ to ‘tree’ can be understood as the strength of the evidence for ‘tree’ given the current belief for ‘leaf’. We will call the message sent from node i to node j at step nMnij. Initial messages and beliefs are set to zero. Sensory evidence (i.e. the sensory evidence for green) is provided by clamping messages in the bottom layer (see toy examples in the next section). Similarly, prior expectations (i.e. the prior probability of being in a forest) are implemented by messages in the top layer. Other messages and beliefs are initialized at zero. Belief propagation implements the following recursive equations until convergence:

\[
M_{n+1}^{ij} = W_{ij}(B_{n}^{j} - M_{n}^{i})
\]

\[
B_{n+1}^{i} = \sum_{j} M_{n+1}^{ij}
\]

The belief propagation algorithm is warranted to converge to the a posteriori probability distribution for each cause. For example, at the end of convergence, the belief associated with ‘tree’ is the posterior probability of a tree given sensory information and prior knowledge. It is clear from these equations that belief propagation is analogous to the dynamics of a recurrent neural network, with the activity of single units interpreted as beliefs that are associated with the variable that they represent (Fig. 1B). Wij can be understood as representing the strength of the connection between neuron i and neuron j. Note that Wij is not a single synaptic weight; instead, it is a sigmoid function whose shape depends on the conditional probability of cause j, given cause i. More exactly, it depends on the conditional probabilities that link node j and node i, p0 = p(xj = 1|xj = 0) and p0 = p(xj = 0|xj = 0), as follows (see Supplementary material for details).

\[
W_{ij}(B) = \log \left( \frac{p_{ij}e^{B} + p_{0j}}{1 - p_{ij}e^{B} + p_{0j}} \right)
\]

Most importantly, the message M_{n+1}^{ij} is a function of the belief of the presynaptic node, B_{n}^{j}, corrected by the message previously sent in the opposite direction. This structure prevents the reverberation of beliefs between node i and node j by predicting out the redundant messages. We propose specifically that this correction is implemented by inhibitory loops, as shown in Fig. 1B.

**Circular belief propagation**

We parameterize the strength of the inhibitory loops by a scale factor α, between 0 (no inhibition at all) and 1 (normal level of inhibition). This scaling factor is applied to the message correction, i.e. M_{n+1}^{ij} = W_{ij}(B_{n}^{j} - αM_{n}^{i}). We also investigate the contributions of different types of inhibitory loops by distinguishing between ‘downward loops’ for redundant messages descending the hierarchy and ‘upward
loops’ for redundant messages climbing the hierarchy. Thus, we distinguish two scale factors for the inhibition of top-down and bottom-up messages:

If $i$ is above $j$:

$$M_{ij}^{n+1} = W_{ij} \left( B_{i}^{n} - \alpha_{i} M_{ij}^{n} \right)$$

If $j$ is above $i$:

$$M_{ij}^{n+1} = W_{ji} \left( B_{j}^{n} - \alpha_{j} M_{ij}^{n} \right)$$

$\alpha_{i}$ and $\alpha_{j}$ parameterize the scaling of inhibition for climbing (upward) and descending (downward) loops, respectively. Thus, downward loops can be seen as removing from bottom-up ‘sensory’ messages the redundant ‘top-down’ predictions by higher level areas. The opposite is also true: upward loops can be seen as removing from ‘top-down’ messages the redundant ‘bottom-up’ information from lower areas that has been already accounted for.

Learning the precision of the causal links

As in all hierarchical inference problems solved by the brain, we distinguish between the inference of the unknown states of the world (binary hidden causes, such as ‘trees’ and ‘leaves’) and the parameters mediating causal dependencies (e.g. what is the probability of a leaf given the presence of a tree). Belief propagation computes the probability of the hidden causes on a fast time scale after exposure to a single stimulus. In contrast, the learning of the causal dependencies occurs over a slower timescale, during exposure to multiple stimuli. We will demonstrate how abnormalities of inference lead to severe abnormalities in learning.

Uncertainty not only is at the heart of decision strategies, but also controls the learning of internal predictive models (Dayan and Yu, 2003). For example, if a high level of confidence for ‘green’ often co-occurs with a high level of confidence for ‘stem’, then there is support for the existence of a higher order common cause for these two observations, e.g. a ‘leaf’ (Fig. 1A). Here, however, we will not consider the learning of latent variables or of the structure of the graph; instead, we will concentrate on the much simpler problem of learning the strength of causal relationships, or precision with which one variable predicts another, i.e. the parameters $p_{ij}^1$ and $p_{ij}^0$, from experience.

To illustrate the impact of circular inferences on learning causal relationships from experience, we compared the result of learning based on the beliefs that are computed by the intact and circular belief propagation networks (i.e. belief propagation and circular belief propagation, respectively). For illustration purposes, we only learned the connection between ‘tree’ and ‘leaves’, setting other connections to their true values. We also used a simplified toy example, without the variables ‘root’ and ‘stem’. Thus, in this simplified network, ‘forest’ caused the presence of a tree with probability $p_{21}^{true}$ and $p_{21}^{false}$, ‘tree’ caused the presence of a leaf with probability $p_{23}^{true}$ and $p_{23}^{false}$, and ‘tree’ caused the presence of green with probabilities $p_{15}^{true}$ and $p_{15}^{false}$. We initialized the network connections with parameters $p_{12}^1 = p_{12}^{true}$, $p_{12}^0 = p_{12}^{false}$, $p_{15}^1 = p_{15}^{true}$, $p_{15}^0 = p_{15}^{false}$, and $p_{23}^1 = p_{23}^{true}$, whereas the unknown (i.e. to be learned) connections between ‘tree’ and ‘leaf’ $p_{23}^1$ and $p_{23}^0$ were set to a random value that was close to 0.5 (Supplementary material).

For each training example, we sampled the presence or absence of a tree randomly from a probability of 0.5 and then sampled tree, leaves and green from their true causal probabilities. ‘Tree’ and ‘green’ were assumed to be completely observable (i.e. their beliefs were set to (0,1) if present and (1,0) if absent). We relied on belief propagation or circular belief propagation to compute the beliefs of the hidden variables (i.e. ‘tree’ and ‘leaves’) and then applied a learning algorithm called ‘expectation maximization’ (Supplementary material) to update the estimates for $p_{23}^1$ and $p_{23}^0$. Note that the results are not crucially dependent of the learning algorithm. In the Supplementary material and Supplementary Fig. 2, we show that a less standard (but more biologically plausible) Hebbian learning rule leads to the same qualitative effects.

Results

We describe here the main impairments in causal inference and perceptual decisions as predicted by the circular belief propagation algorithm. Interpretations of these impairments, both in terms of the positive symptoms of schizophrenia as well as the other dimensions of this disorder, are provided in the ‘Discussion’ section.

Normal and circular inference

We first consider the case in which upward and downward loops are equally impaired, i.e. when $\alpha_{i} = \alpha_{j} < 1$. The impairment of inhibitory loops causes redundant top-down and bottom-up information to be propagated multiple times in the network (Fig. 1C and D). Quite surprisingly, however, this circular belief propagation does not generate completely ‘false’ percepts, i.e. the signs of LORs are unaffected (Fig. 1D). For example, if the presence of a tree is more likely than its absence according to a normal inference, then it is also found to be more likely by the impaired network. This pattern is illustrated with the binary toy example in Fig. 2A and B. In Fig. 2A, strong sensory evidence is provided for the presence of ‘green’ and ‘stem’, which leads the network to correctly conclude the presence of trees and a forest. In Fig. 2B, no sensory evidence is provided (e.g. the eyes are closed), but there is a strong prior expectation against the forest, which leads both belief propagation and circular belief propagation to correctly predict the absence of trees, leaves, green and stems.

The fact that circular belief propagations converge to posterior probabilities on the correct side of 0.5 has been reported previously in the context of circular binary graphs (Frey and McKay, 1998; Murphy et al., 1999). These results predict specifically that networks with equally impaired upward and downward loops (the ‘equally’ is important, as we will see later) can still infer the most probable interpretations of their sensory input and priors. Even a total absence of inhibitory controls does not predict any systematic bias in perceptual or behavioural decisions as long as these decisions are based on reliable sensory evidence and/or on general agreement with prior expectations. The main explanation for this phenomenon is as follows: while top-down and bottom-up messages are indeed counted too many times, the number of over-counted messages is equal in both directions, leaving the relative weighting of the sensory evidence and priors intact.
Figure 2  Results for normal and circular inference when inhibitory loops are equally impaired. In each panel, the computed beliefs for all of the nodes $x_1$ to $x_6$ are plotted for different iterations of the network equations (blue circles: physiological network; red lines: impaired network with $\alpha_c = 0.1$, $\alpha_d = 0.1$). (A) If sensory evidence is entered as an Log-Odd Ratio (LOR) of 1 for « green » and an LOR of 2 for « stem », then the network converges toward the correct percepts (i.e. the beliefs are on the same side of 0.5), but the confidence is overestimated. (B) The same as in A, but with no sensory evidence and a prior belief that corresponds to an LOR for a forest of −3 (i.e. a strong prior belief against the forest). (C) If the network receives very weak sensory evidence for green and stem, then the final belief corresponds to technically correct percepts (on the correct side of 0.5) but with an excessively high confidence for such weak sensory evidence. (D) Finally, when the network receives contradictory priors and sensory evidence, i.e. the prior is against a forest (LOR of −2), but when sensory evidence suggests the presence of green and stems (LORs of 1), intermediate representations become frustrated and oscillate between prior-based and sensory-based beliefs, specifically for « tree » and « root ».
Over-confidence

Even if the LOR’s sign is correct, the levels of confidence that are reached by the impaired and ‘normal’ networks, i.e. the amplitude of the LOR, are widely different in the circular belief propagation and belief propagation networks. Because circular belief propagation accounts for sensory and prior evidence not once, but multiple times, it generates levels of confidence that are out of proportion when compared with the available sensory and prior information (Fig. 1D). In the example shown in Fig. 2C, the sensory evidence provided to the network is extremely weak. Such small likelihoods could be the result of: (i) processing extremely unreliable sensory inputs; (ii) completely ambiguous sensory inputs; or even (iii) a pure background sensory noise.

Weak sensory evidence can be partly processed and analysed (e.g. we can sometimes see faces in the random shapes of clouds), but the confidence that is associated with these interpretations should be very low. Indeed, the unimpaired network correctly maintains LORs close to zero in these situations, which indicates the absence of any reliable conclusions. If perceptual decisions (and any other forms of decisions) are based on the confidence in these interpretations, then the system should not perceive trees or forest when the evidence is too weak. In contrast, the circular belief propagation network ‘jumps to conclusions’ and reaches a high level of confidence even when the sensory evidence and/or the prior are extremely weak. In fact, the circular belief propagation network cannot stably maintain the LORs of any variable near zero but will amplify any small biases until reaching the maximum level of confidence that is allowed by the strength of the causal links $W_{ij}$. In other words, it is forced to derive a highly confident interpretation from meaningless or non-significant sensory information. The same phenomena are observed in the presence of a weak prior (not illustrated here). For example, even a small expectation of being in the forest could lead the system to perceive trees, leaves and the colour green.

Dissociation between ‘low-level’ and ‘high-level’ representations

In addition to being overconfident, the circular belief propagation network exhibits a striking dissociative effect between low-level sensory areas and high-level representational areas in cases in which sensory evidence contradicts prior beliefs (Fig. 2D). For example, this type of situation could occur when we are inside a city building and see stems and the colour green (e.g. from a potted plant). This phenomenon goes against prior expectations because we are not in a forest, but rather a building, where plants are rare. Normal belief propagation in the unimpaired network correctly concludes that while the presence of ‘leaves’ is likely, the presence of a tree is unlikely. Indeed, while ‘tree’ is supported by ‘leaves’, it is contradicted by the absence of ‘forest’. Top-down and bottom-up evidence cancel each other at an intermediate level in the hierarchy, preventing the propagation of false beliefs up and down the hierarchy. Only a few iterations for the belief propagation algorithm are sufficient to reach this stable belief state.

Instead of quickly converging to low confidence in intermediate layers, however, the circular belief propagation algorithm alternates (over successive iterations of the circular belief propagation algorithm) between strong beliefs that are based on sensory evidence alone (a consequence of over-counting bottom-up messages) and strong beliefs based on prior expectations alone (a consequence of over-counting top-down messages) (Fig. 2D). Because both belief states are mutually incompatible, this pattern results in incoherence between ‘high level’ and ‘low level’ sensory representations. These oscillations are strongest in the intermediate layers (where prior and sensory evidence should cancel to give an LOR that is close to zero). At these intermediate stages, the network alternates between one interpretation of the sensory observations and its exact opposite. When the loops are completely impaired ($\alpha_c = \alpha_d = 0$) and the causal links are strong ($p_{ij}^b = 1$, $p_{ij}^d = 0$), such oscillations can even be sustained indefinitely, which prevents the network from converging to any stable conclusions. When loops are only partially impaired ($1 > \alpha_c = \alpha_d > 0$) and causal links are less extreme, the amplitude of these oscillations is progressively dampened until the network reaches stable beliefs, but not before many iterations of the circular belief propagation algorithm have occurred.

Impact of a circular inference on learning causal relationships

Uncontrolled reverberation of messages is as detrimental to learning as it is to inference. For example, if the node ‘tree’ is frequently co-activated with ‘leaf’, the strength of their causal link should be reinforced, but only if the activation of ‘tree’ is not solely due to processing information from ‘leaf’ and comes from other (non-redundant) sources of information, such as the presence of ‘forest’ or the colour ‘green’ or ‘stem’ (Fig. 1). Otherwise, the causal link would be reinforced indefinitely, even in the absence of any true causality, simply because any pre-existing connection between ‘leaf’ and ‘tree’ will introduce spurious correlations between their associated beliefs, leading to a reinforcement of this link and thus leading to more correlations.

Even when the causal model is accurate, as in the examples above, circular inference causes over-confidence and dissociative beliefs on a trial-to-trial basis. However, on a more permanent (and most likely disruptive) basis, it also results in the learning of delusional causal relationships, as illustrated in Fig. 3. Figure 3A represents the learned conditional probability estimate $p_{23}^{\text{true}}$ as a function of the number of training examples, in a situation of high sensory and prior uncertainty (i.e. $p_{12}^{\text{true}} = p_{34}^{\text{true}} = 0.7$, $p_{12}^{\text{true}} = p_{34}^{\text{true}} = 0.3$). Let us assume that ‘green’ and ‘leaf’ are in fact completely independent events in the real world, i.e. $p_{23}^{\text{true}} = p_{23}^{\text{true}} = 0.5$. This construct could correspond to the imaginary case of an observer living in a world where leaves are not carried by trees. The intact belief propagation network accurately learns that the two variables are not causally linked, i.e. $p_{23}^{\text{true}}$ converges to a value close to 0.5. In contrast, the estimate of the circular belief propagation network ($\alpha_c = \alpha_d = 0.1$) does not stay close to 0.5; instead, it diverges to a value that is close to 0 or 1, which implies that the model gradually learns a strong but completely ‘delusional’ (because it does not exist in the real world) causal relationship between ‘green’ and ‘leaf’.
With respect to overconfidence, this problem arises only in situations that have a high level of uncertainty, i.e. when sensory and prior evidence are relatively unreliable and between variables that are completely uncorrelated. If priors and likelihoods are strong and unambiguous, or if the causal relationship is real (i.e. most of the time), both the intact and circular networks learn accurate causal links. For example, by using more accurate evidence in Fig. 3c ($p^{true}_{12} = p^{true}_{34} = 0.9$, $p^{true}_{12} = p^{true}_{34} = 0.1$, while $p^{true}_{13} = p^{true}_{23} = 0.5$), we found that the circular belief propagation network estimates converged to the true estimate of 0.5, inferring that ‘tree’ is unrelated to ‘leaf’. Similarly, even when sensory and prior evidence are weak ($p^{true}_{12} = p^{true}_{34} = 0.7$, $p^{true}_{12} = p^{true}_{34} = 0.3$), both intact and circular networks can accurately learn a non-random causal relationship $p^{true}_{13} = 0.8$, $p^{true}_{23} = 0.2$ (Fig. 3B). These examples illustrate how the acquisition of delusional causal relationships is observable only between ‘highly uncertain’ variables that are truly unrelated, and at the same time, are not strongly corroborated by direct sensory evidence. For example, a tingling sensation in the arm is only weak evidence for the presence of a micro-chip and, in principle, is not related to the existence of ‘invisible’ aliens; nevertheless, a patient suffering from schizophrenia can strongly believe that aliens (that he never directly observed) implanted a micro-chip (that he did not really feel) in his arm.

**Figure 3** Learning in circular and intact belief propagation networks using the expectation-maximization (E-M) algorithm. (A) Representation of the learned conditional probability estimates as a function of the number of trials in a situation of high sensory and prior uncertainty. Although the belief propagation network quickly converges to a value that is close to 0.5 (green), the circular belief propagation network gradually diverges to 0 or 1 (red), which models how such an impaired system could aberrantly detect coincidences that do not rely on observable causal relationships. Note that the different curves were obtained by running the E-M algorithm on a different random sequence of training trials, with different initial values for the parameters. (B) Learned conditional probabilities between uncorrelated causes as a function of the number of trials, when sensory and prior evidence is more accurate. (C) Learned conditional probabilities between truly correlated causes as a function of the number of trials in a situation of high sensory and prior uncertainty.

**Asymmetric impairments of upward and downward loops**

Because upward and downward inhibitory loops most likely rely on distinct sub-populations of interneurons and/or cortico-basal or cortico-thalamic loops (see ‘Discussion’ section), both types of loops could be differentially affected in schizophrenia. Here, we consider the consequence (for inference) when upward loops are altered while downward loops remain relatively intact, or vice versa. Figure 4 illustrates the differential effects of disturbing mainly upward loops (Fig. 4A, black) or mainly downward loops (Fig. 4B, green). As described previously, both types of dysfunctions lead to overconfidence (Fig. 4C). Similarly, contradictions between sensory evidence and priors lead to oscillations in intermediate layers and dissociations between high-level and low-level representations, albeit to a lesser extent (Fig. 4D and 4E). In addition, however, the two types of impairments affect the relative weighting of the sensory evidence and priors in strikingly different fashions.

When upward loops are affected, but downward loops are not, sensory evidence that is sent up the hierarchy is reverberated back down (Fig. 4A). As a result, high-level interpretations of the current sensory information are mistaken for prior knowledge. However, prior knowledge sent down the hierarchy is not reverberated back up, i.e. the system does not mistake its own expectations for sensory evidence (Fig. 4A). As a result, sensory evidence accumulates over successive layers of the hierarchy and becomes severely over-counted. However, priors (descending messages) are processed normally because the downward loops are intact. Because they are not reverberated, prior evidence is not over-counted. This scenario introduces an effective imbalance in the relative weighting of sensory evidence and priors. The network with impaired upward loops relies more on its current sensory observation than its prior knowledge or its past experience (Fig. 4D and E).

In contrast, when downward loops are affected, but upward loops are not, prior expectations that were sent down the hierarchy are reverberated back as if they were additional sensory evidence (Fig. 4B). Thus, the system mistakes its own prior expectations for new (supporting) sensory observations. In contrast, sensory evidence that is sent up the hierarchy is not reverberated back down, i.e. the system does not mistake sensory evidence for prior expectations (Fig. 4B). As a result, predictions based on priors are over-counted and accumulate over successive layers. In contrast, sensory evidence is not reverberated and is counted only once. As a consequence, the network with impaired downward
Figure 4  Results of circular inference with an imbalance between upward and downward loops. (A) When upward loops are impaired but downward loops are not, sensory evidence that is sent up in the hierarchy is reverberated back as if it were additional prior evidence (magenta arrows). However, prior evidence sent down the hierarchy is not reverberated back up, i.e. the system does not mistake its own expectations for sensory evidence (violet arrow). As a result, sensory evidence is overcounted, but not priors. (B) When downward loops are impaired, but not upward loops, the prior expectation sent down the hierarchy is reverberated back as if it were additional sensory evidence (violet arrows). However, sensory evidence sent up the hierarchy is not reverberated back down, i.e. the system does not mistake its sensory evidence as prior expectations (magenta arrow). As a result, prior expectations are overcounted, but not sensory evidence. (C) Impaired networks in the presence of weak sensory evidence (in this example, against the presence of green and stem). Blue: intact network. Green: networks with impaired downward loops, i.e. $C_{11}^d = 0.9, C_{11}^c = 0.1$ (as illustrated in (B)). Black: networks with impaired upward loops, i.e. $C_{11}^c = 0.1, C_{11}^d = 0.9$ (as illustrated in (A)). Both dysfunctions result in overconfidence, i.e. weak sensory evidence leads to highly confident beliefs. (D) Impaired networks in the presence of contradicting prior information (for the presence of forest) and sensory evidence (against the presence of green and stem). The network with impaired upward loops bases its highly over-confident conclusions entirely on the prior information. The two networks reach such opposite conclusions for the low-level variables, even when a normal network typically concludes that the states of these variables are highly uncertain. (E) The same as in D, but for a different combination of likelihoods and priors.
loops relies more on its priors that on its new sensory evidence when making perceptual decisions (Fig. 4C and D).

We thus can say that, quite paradoxically, an imbalance between upward and downward loops (Fig. 4) has an even more drastic and disruptive influence on perception and belief formation than when both loops are equally impaired (Fig. 2). In contrast to the case of equally impaired loops, a selective impairment creates an imbalance and results in ‘false percepts’, i.e. in subjects believing in interpretations that are opposite to the results of normal inference (LORs are of the wrong sign).

Let us finally consider the case of perceptual illusions. In a normal inference (belief propagation), a strong prior (i.e. slow motions are more likely than fast motions) combined with weak sensory likelihood (i.e. a visual motion stimulus at a low contrast) results in a systematic bias in the direction of the most a priori probable interpretations (i.e. the motion of a low contrast visual stimulus is perceived as slower than it actually is because slower motions are more likely). This effect is gradual, i.e. the weaker the sensory likelihood, the stronger the effect of the prior. This phenomenon accounts specifically for many perceptual illusions (Fullle and Büttthoff, 1996; Weiss et al., 2002; Bogadhi et al., 2011). Figure 4D and E show two examples in which weak sensory evidence is combined with a strong prior. The circular belief propagation framework predicts a strong disruption of the normal perceptual biases when the upward and downward loops are differentially impaired. Whereas a selective impairment of the downward loops predicts dominance of the prior and thus a stronger effect of perceptual illusions, a selective impairment of the upward loops results paradoxically in fewer perceptual illusions than are normally expected. Thus, assessing the relative weighting of sensory observations and priors is an indirect way to assess whether upward or downward loops are primarily impaired (or vice versa).

Discussion

The aim of the current paper was to introduce the concept of circular belief propagation in causal networks as a result of an excitatory to inhibitory imbalance in hierarchical cortical networks. We now discuss more conceptually how it could account not only for inappropriate causal attributions, bizarre beliefs and hallucinations in the schizophrenia spectrum but also for emerging aberrant coincidence detections during the transition phase to psychosis.

The ‘jumping to conclusions’ phenomena

The most salient effect of circular belief propagation is to reach levels of confidence that are out of proportion with the true levels of uncertainty that are associated with the sensory data and/or prior knowledge. Such over-interpretation of weak sensory observations could account for the ‘jumping to conclusions’ bias that is observed in schizophrenia patients (Huq et al., 1988). The jumping to conclusions phenomenon is traditionally described using probabilistic reasoning tasks such as the ‘beads task’. Two jars are presented to the participant with different proportions of ‘red’ and ‘black’ beads, and the task is to successively sample beads from one of the two jars. The subject must decide from which jar the beads are taken (based on the proportion of beads in each jar). Using such a paradigm, a number of behavioural studies have showed that patients with paranoid schizophrenia, when faced with probabilistic choices, base their decision on far less evidence while reporting much higher confidence than healthy controls. Crucially, the jumping to conclusions phenomenon seems to be specifically associated with delusional tendencies (Huq et al., 1988; Garety et al., 1991; Moritz and Woodward, 2005; Speechley et al., 2010; Averbeck et al., 2011).

Importantly, the circular belief propagation framework can be distinguished from Bayesian models assuming degraded sensory information or the use of suboptimal decision criteria (Averbeck et al., 2011; Moutoussis et al., 2011), but considering that the inference itself is not affected (Fear and Healy, 1997). In line with an impaired inference process, a recent experiment confirmed that when explicitly required to report their confidence level in a variant of the ‘beads task’, patients with schizophrenia systematically overestimated the weight of the sensory evidence (Speechley et al., 2010).

Uncertainty, hallucinatory and delusional experiences

Interestingly, the dysfunctions observed in the circular belief propagation models were not found to be perpetual in the sense that the system was able to interpret appropriately the stimuli most of the time. Erroneous beliefs (e.g. jumping to conclusions (phenomenon)) were shown to emerge only when the system was facing ambiguous or contradictory information, i.e. when the probabilities were close to 0.5. Such situations are presumably relatively rare in everyday life, where our senses and priors provide us (most of the time) with strong and unambiguous information. The low frequency of ambiguous situations could account for the very intermittent nature of hallucinatory experiences (Jardri et al., 2013) but could also explain why hallucinations are more frequent when there is sensory deprivation (Zubek, 1964). A lack of sensory input could, in this situation, put causal networks into a state of high uncertainty, leading to an increased occurrence of erroneous percepts. Furthermore, this scenario could account for the frequent persecutory nature of paranoid delusions because social inferences are tainted with much more uncertainty than simple perceptual inferences in which the relationship between the cause and the evidence is more straightforward (Chambon et al., 2011).

Learning of delusional causal models

We showed that circular belief propagation disrupts the control of messages and results in the reverberation of top-down and bottom-up evidence, which occurs multiple times. Thus, even in the absence of any true causal relationships in the outside world, top-down and bottom-up messages can be strongly correlated (e.g. the messages are mixtures of violet and magenta arrows, in Fig. 4). As a result, inaccurate causal links can be progressively and recursively strengthened, which leads to highly trusted causal
models that are not corroborated by sensory experience. This dynamic phenomenon fits especially well with the abnormal experiences that are described in the initial phase of psychotic experiences (Woods et al., 2009). During this prodromal period, patients describe the appearance of mild or brief non-specific symptoms, such as a noticeable feeling of strangeness or aberrant coincidence detections. Interestingly, a low dose administration in healthy volunteers of ketamine, an NMDA receptor antagonist drug, also appears to mimic prodromal states rather than the full picture of schizophrenia (Pomarol-Clotet et al., 2006). Such a drug-induced prepsychotic state supports the progressive learning of aberrant causal associations that has been demonstrated by circular belief propagation models.

Beyond the emergence of delusional ideas, circular belief propagation models also account for their remarkable persistence. Indeed, an insufficient compensation for internally generated correlations by inhibitory loops could lead to the consolidation of extremely strong delusional causal models based on weak or insignificant sensory evidence or prior beliefs, thus forcing paranoid extremely strong delusional causal models based on weak or insignificant sensory evidence or prior beliefs, thus forcing paranoid beliefs (Teufel et al., 2010). Knowledge that the biases that were introduced by prior beliefs are considered to be at the root of perceptual illusions (Geisler and Kersten, 2002; Weiss et al., 2002), these findings support an over-estimation of the strength of sensory evidence as top-down rather than bottom-up. This, however, does not rule out the possibility that some patients are also impaired in their downward loops, and thus, over-count their priors (Friston, 2008; Chambon et al., 2011). Importantly, these two hypotheses could be tested experimentally by measuring how patients weight their likelihood and priors during decisions. Finally, an impairment of downward loops (i.e. the reverberation of top-down expectations as if they were sensory evidence) could cause vivid mental imagery, as observed, for example, during normal human development before complete frontal GABAergic maturation (Uhlhaas et al., 2010).

We would like to highlight that, in the circular belief propagation framework, positive symptoms are understood in terms of false inferences in a causal hierarchy and not as the consequence of focal brain damage (the previously described diffuse inhibitory loops impairment is clearly different from what could be expected from a neural lesion). As shown in Supplementary Fig. 3, low-level damage (e.g. at the level of the retina or of the pedunculus) can generate erroneous sensory inputs (Supplementary Fig. 3A), but because inferences are correctly processed during belief propagation, the subject continues to be able to challenge and disprove these perceptions. Such a preserved insight is effectively a prominent feature of hallucinosis (Ey, 1973; Blom, 2010), compared with the psychotic hallucinations that are observed in schizophrenia. In contrast, when lesions are located in high-level brain structures (e.g. at the level of the pre-frontal cortex), incorrect priors can be generated, which results in common symptoms of the dysexecutive syndrome, such as perseverations; however, these patients do not suffer from abnormal sensory judgements (Supplementary Fig. 3B).

Selective impairment of upward loops in schizophrenia?

The circular belief propagation framework distinguishes between different classes of circular inference by considering bottom-up and top-down streams separately. We have seen that impaired upward loops result in sensory evidence that is reverberated as if it were a prior expectation, which causes sensory evidence to be over-counted. The opposite also occurs: impaired downward loops result in prior expectations being reverberated as if they were sensory evidence, which causes priors to be over-counted. These two forms of impairment could both be implicated in the schizophrenia spectrum, which would account in part for the large heterogeneities that are observed in patients’ behaviours (Tandon et al., 2009).

It has been previously proposed in the literature that positive symptoms could result from an underestimation or impairment of prediction errors, which would lead patients with hallucinations and delusions to rely more on their priors (Grossberg, 2000; Friston, 2005a; Fletcher and Frith, 2009). Such a hypothesis is in line with an impairment of downward loops in the current framework. However, an over-counting of the prior does not appear to be fully compatible with the spectrum of behaviour that is observed in schizophrenia. For example, let us return to the jumping to conclusions phenomena that were previously described. In the beads task, patients with delusional tendencies base their decision on far fewer samples (often only one sample suffices) while reporting much higher confidence compared with healthy control subjects. If these patients under-estimate the reliability of new sensory observations (i.e. a newly sampled ‘black’ bead) compared to what they already know (i.e. the probability of each urn based on beads that were observed previously), they should, in contrast, give a smaller weight to each new sample, thus accumulating evidence and confidence more slowly than healthy controls. Jumping to conclusions suggests, in contrast, that patients over-interpret their sensory evidence compared to their prior knowledge, which could be the result of impaired upward loops.

In additional support of this hypothesis, patients with schizophrenia are paradoxically less vulnerable to illusory percepts than healthy control subjects (Dakin et al., 2005; Tschacher et al., 2006; Crawford et al., 2010; Williams et al., 2010). This attenuation specifically appears in patients with schizophrenia prone to positive symptoms (Shergill et al., 2005), but may also be observed in healthy individuals that score highly on delusional beliefs (Teufel et al., 2010). Knowing that the biases that were introduced by prior beliefs are considered to be at the root of perceptual illusions (Geisler and Kersten, 2002; Weiss et al., 2002), these findings support an over-estimation of the strength of sensory evidence and an under-weighting of the prior, which is compatible with an impairment of the upward loops.
The relationship between circular belief propagation and gain control

False inference has now become a dominant theme in the study of the positive symptoms (reality distortion) of schizophrenia and related syndromes. Generally, this is cast in terms of abnormal representations of uncertainty or—more specifically—estimates of the precision of sensory evidence relative to top-down prior beliefs (Friston, 2005a; Fletcher and Frith, 2009). In the Bayesian formalism, the precision of sensory evidence and prior knowledge determines their gain, i.e. their relative contributions. Interestingly, one of the main consequences of circular belief propagation is the resulting change in the confidence assigned to sensory evidence and top-down priors. This is mediated by multiple reverberations of the same redundant message, and will worsen with the number of layers and the number of iterations of the algorithm. As a result, the ‘normal’ gain of the sensory evidence and/or prior drastically increase as $a_s$ and $a_d$ decrease from 1 (normal inhibition) to zero (no inhibition).

The relationship with predictive coding

In this paper, we focus on simple models with discrete states that do not change with time. In the case of continuous time-varying variables, belief propagation can be efficiently approximated (or in some cases exactly implemented) by a form of generalized predictive coding. As in belief propagation, the generalized predictive coding framework has been used to interpret hallucinations and delusions as an impairment in the neural representation of uncertainty (Friston, 2005a; Adams et al., 2013). In addition, belief propagation and generalized predictive coding are algorithmically similar. Both belief propagation and generalized predictive coding are minimizing an energy function (Friston, 2005b; Friston et al., 2010; Balaji and Friston, 2011). As for belief propagation, generalized predictive coding estimates the precision (gain) of bottom-up and top-down evidence online in order to combine them properly. Finally, generalized predictive coding updates internal representations according to the error between the bottom-up sensory evidence and its top-down predictions (i.e. the prediction error). This mechanism is analogous to the downward loops in belief propagation, in which ‘bottom-up’ messages are corrected by the previous redundant ‘top-down’ message (and vice versa). Although a systematic and rigorous comparison of two theoretical approaches extends beyond the scope of this paper, these similarities clearly imply that the two frameworks are related and mutually compatible.

Possible neural substrates

Different models have been proposed for the specific neural implementation of belief propagation in cortical circuits (Lee and Mumford, 2003; Deneve, 2004; Friston, 2005b; Bishop, 2006; George and Hawkins, 2009; Steimer et al., 2009; Markov and Kennedy, 2013). Importantly, regardless of its specific implementation, belief propagation relies on local corrective loops to avoid circular inference, a property shared with ‘predictive coding’ (see above). The presence of these corrective loops has been shown to be highly compatible with the architecture, connectivity and dynamics of the cortical column (Bastos et al., 2012).

In addition, the notion of separate feed-forward and feedback inhibitory pathways (implementing on upward and downward loops, respectively) builds on a series of anatomical and physiological arguments. It can—at least roughly—be motivated by the distinction between superficial and deeper pyramidal cells that are the source of forward and backward connections, respectively (Coogan and Burkhalter, 1993; Markov et al., 2011). Furthermore, recent electrophysiological studies suggest that feed-forward connections rely on higher synchronization frequencies than feedback connections (Bosman et al., 2012). Another potential implementation for these inhibitory pathways relies on the mediation or control of cortical communications by subcortical structures. For example, the ventral striatum was proposed as a key structure in the pathophysiology of positive symptoms. Abnormally salient experiences in schizophrenia could relate to an increased tone within the mesolimbic dopamine system (Kapur, 2003), whereas $D_2$ receptor blockage by antipsychotic medication was shown to be critical to symptom resolution (Creese et al., 1976; Johnstone et al., 1978). Moreover, increased activity within the mesolimbic system correlates with the severity of hallucinatory experiences (Howes et al., 2007).

Accounting for other clinical dimensions of schizophrenia?

Circular belief propagation provides a new framework for interpreting the generation of unusual sensory experiences or strange beliefs, which progressively become unshakable. From a macro-anatomical view, we can conclude that circular belief propagation models with impaired inhibitory loops are fully compatible with dysconnectivity hypotheses that are supported by several imaging findings in schizophrenia and, more specifically, in hallucinations (Stephan et al., 2009; de Weijer et al., 2011; Amad et al., in press). Indeed, circular belief propagation acts by disrupting communications between brain areas. However, does circular belief propagation account for the more rarely modelled dimensions of schizophrenia, such as cognitive disorganization or negative symptoms?

When upward inhibitory loops are selectively impaired, the system could be forced to take into account ‘overweighed’ sensory information even if such information is unreliable. This possibility perfectly matches with experimental findings from the classical latent inhibition paradigm, which measures the ability to filter irrelevant stimuli, a capability that is found to be decreased in patients with schizophrenia (Lubow and Weiner, 2010). This proposal also fits with the high distractedness and attentional impairments that are observed in schizophrenia (Tandon et al., 2009).

Finally, the emergence of frustrated states in circular belief propagation networks provides a potential mechanism for the occurrence of concomitant ambivalent ideas in schizophrenia patients who suffer from dissociation (Fig. 2D). We effectively showed that circular belief propagation predicts incoherence and instability in the beliefs that were generated by intermediate levels of representations whenever sensory evidence and prior contradicted each
Circular inferences in schizophrenia

Conclusions

Within the broader field of psychiatry, a background argument has been to determine whether psychosis was essentially characterized by anomalous perception leading to incorrect (but essentially rational) inferences or, on the contrary, may rely on normal perceptions but irrational inferences. Crucially, as suggested by Fletcher and Frith (2009), the Bayesian framework renders the argument itself unimportant because a single deficit acting at multiple levels may account for both perceptual and belief/inferential anomalies in psychosis. In line with this approach, we used an analogy between hierarchical neural processing and belief propagation to illustrate how excitatory to inhibitory imbalances might result in circular inference. Our framework is normative, highly testable, and nicely relates to the clinical concepts of ‘double orientation to reality’ by Jaspers (1963) or ‘double entry bookkeeping’ by Bleuler (1950), which both refer to the fact that patients with schizophrenia might live with two separate and conflicting sets of beliefs.

Acknowledgements

We would like to thank Philippe Domenech, Izzet Burak Yildiz and Delphine Pins for reading and discussing a preliminary version of this work. We also thank the reviewers for their help in improving the manuscript.

Funding

RJ was partially supported by the Pierre Houriex Foundation (hosted by the Fondation de France), grant no. FdF-200462/2009-02.

Supplementary material

Supplementary material is available at Brain online.

References


Corlett PR, Frith CD, Fletcher PC. From drugs to deprivation: a Bayesian framework for understanding models of psychosis. Psychopharmacology (Berl) 2009a; 206: 515–30.


de Weijer AD, Mandl RC, Diederren KM, Negers SF, Kahn RS, Hulshoff Pol HE, et al. Microstructural alterations of the arcuate fasciculus in...
schizophrenia patients with frequent auditory verbal hallucinations. Schizophr Res 2011; 130: 68–77.


Lewis DA, Hashimoto T, Volk DW. Cortical inhibitory neurons and schizophrenia. Nat Rev Neurosci 2005; 6: 312–24.

Liddle PF. Schizophrenic syndromes, cognitive performances and neurological dysfunction. Psychol Med 1987; 17: 49–57.


