



OPEN ACCESS

EDITED BY

Yan Mark Yufik,
Virtual Structures Research Inc., United States

REVIEWED BY

Joseph Monaco,
SelfMotion Labs, United States
Prakash Chandra Kavi,
University of Girona, Spain

*CORRESPONDENCE

James Joseph Wright
✉ jj.w@extra.co.nz

RECEIVED 11 August 2025

ACCEPTED 14 October 2025

PUBLISHED 21 November 2025

CITATION

Wright JJ and Bourke PD (2025) Minimization of prediction errors during cerebral embryogenesis and the emergence of agency. *Front. Syst. Neurosci.* 19:1683448. doi: 10.3389/fnsys.2025.1683448

COPYRIGHT

© 2025 Wright and Bourke. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Minimization of prediction errors during cerebral embryogenesis and the emergence of agency

James Joseph Wright* and Paul David Bourke

Centre for Brain Research, School of Medicine, University of Auckland, Auckland, New Zealand

A theory of self-organization in the central nervous system is described, proposing that additive and dissipative synaptodendritic summation leads to synchronous oscillation as the equilibrium state, thereby underpinning a primary mechanism of prediction error minimization. As a consequence, synaptic connections become arranged in mirror-symmetric paired patterns, wherein exchanges of synaptic flux within each pattern form coupled spatial eigenmodes. The mirror-reflection axis between each pair functions as a Markov blanket that maintains excitatory–inhibitory equilibrium, while multiway exchanges among mirror pairs converge toward overall error minimization and mutual organization. The primary organization of this type is evident in the spinal cord. During cortical embryogenesis, connections develop in topographies interpretable as mirror reflections with broken symmetry, aligning along the radial and circumferential axes of cortical growth, as described by the Structural Model, and subsequently manifest at the millimetric scale throughout the cortex. The proposed framework integrates a diverse range of experimental data and provides an explanatory basis for how generative models with agency can emerge through both species evolution and individual learning.

KEYWORDS

agency, free energy principle, predictive error minimization, cortical embryogenesis, structural model, synchronous oscillation

1 Introduction

This study reviews the authors' earlier work on cortical dynamics and self-organization in cerebral embryogenesis in the context of current controversy about the utility of the free energy principle and the importance of prediction error minimization in brain function.

The free energy principle, with its corollary, prediction error minimization, has become influential throughout the brain sciences over the last two decades as a significant modification of earlier top-down vs. bottom-up concepts of brain organization. It generalizes formal parallels between laws of nature, including least action, thermodynamics, and Bayesian inference (e.g., [Friston, 2002, 2005, 2008, 2010a,b](#); [Hohwy, 2016](#); [Ramstead et al., 2022](#)), and depends upon the idea that all systems separated by a boundary from a wider environment evolve until they have minimized their perturbation by that environment. This implies error minimization across the boundary, as, in the limit, equal and opposite signals are continuously exchanged, resulting in maximum mutual information. Such a boundary is termed a Markov blanket by Friston ([Kirchhoff et al., 2018](#); [Friston et al., 2017](#); [Tschacher and Haken, 2007](#); [Fields and Levin, 2019](#); [Kiebel and Friston, 2011](#); [Levin, 2019](#); [Hohwy, 2016](#)).

Applying the principle to the central nervous system (CNS), it is supposed that an internal representation of the self, acting within the world, arises in the higher centers of the brain as a generative model with the capacity to bring about novel consequences—that is, with the capacity for agency.

Due to the principle's generality and unity, and its implications for natural general intelligence, its importance for the future development of artificial general intelligence is apparent. However, it also faces strong criticism. An early criticism was that of the dark room (Friston et al., 2012; Clarke, 2013; Sun and Firestone, 2020)—that escape from stimulus alone would minimize error. Recent criticisms (Kwisthout and van Rooij, 2020; Mangalam, 2025) assert that the principle is metaphoric and not testable; that the computation of Bayesian priors and posteriors could not be achieved within hierarchical architectures; that intermediate values of these measures could not be stored; and that assumptions made about the statistical properties of pulses and electrocortical waves are unphysiological.

These criticisms conflate high-level cognitive abstractions with basic physiological mechanisms. They would be alleviated if it were shown that neurons minimize prediction error in interactions at all scales, so they are not bound solely by hierarchical interactions; that the mechanisms are robust with respect to the statistical characteristics of pulse and wave generation; and that such neural mechanisms tend to evolve toward agency. What follows is a theoretical account of neuroanatomical order that meets these requirements and could be further tested in connectomics.

2 The neural field

We begin with definitions of minimal network characteristics (Wright and Bourke, 2024a,b) necessary to justify the simulations upon which our arguments rely. The network properties overlap with those of classical artificial networks (Hopfield, 1982; Rumelhart et al., 1986; Xie and Seung, 2003), but are not identical. When stochastic approximations are assumed, it does not imply that Gaussian or linear properties are necessary.

2.1 Presynaptic flux and Hebbian plasticity

The n unidirectional flows of presynaptic flux in a neural field can be represented as n elements of a square matrix $\Phi(t)$, with elements φ_{ij} , each representing the presynaptic flux received at the neuron at position i from the neuron at position j , over all pathways.

$$\Phi(t) = GQ(t) \quad (1)$$

$G(t)$ is a square matrix operator of presynaptic gains and axo-dendritic conduction times, transforming $Q(t)$, a vector of action potential pulse rates of all neurons, so that in the steady state

$$\varphi_{ij} \left(t + \frac{|i-j|}{v} \right) = \pm \varepsilon_{ij} g_{ij} \rho_{ij} Q_j(t) \quad (2)$$

Q_j is the pulse rate of the j -th neuron, v is the speed of signal spread, and $\frac{|i-j|}{v}$ is the delay from pulse generation to arrival of peak pulse density at pre-synapses on the i -th neuron, averaged

over all routes. Synaptic gains are on three time-scales, ε_{ij} , g_{ij} , ρ_{ij} —the transient synaptic efficacy, the slow dynamic synaptic gain, and the structural synapse gains, respectively, with synaptic flux either excitatory or inhibitory.

It is assumed that gains are competitive on all three time-scales (Miller, 1996; Okamoto and Ichikawa, 2000) and consistent with the unification of fast and slow synaptic learning rules proposed by Izhikevich and Desai (2003), combining short-term plasticity (STP) and short-term depression (STD) with the slower and more permanent Bienenstock-Cooper-Monro (BCM) rule. Thus, with ongoing network activity, the rapid and transient synaptic efficacies in response to afferent pulses lead to increasing dynamic and structural gains as the time average of synaptic flux.

2.2 Dendritic summation and pulse generation

Transformations of synaptic flux leading to pulse generation can be represented as

$$V_i(t) = \delta(t - \tau) * \sum \varphi_{ij}(t) \quad (3)$$

$\delta(t - \tau)$ is a dendritic delay function convolving the aggregate presynaptic flux to generate the dendritic potential, $V_i(t)$, and may describe any form of leaky integration. The large numbers of independent non-linear summations at the dendritic membrane are assumed to be stochastically smoothed to an additive approximation, where $Q_i(t)$ is the efferent pulse rate of the i -th neuron,

$$Q_i(t) = f_\theta(V_i(t)) \quad (4)$$

and f_θ is any sigmoid-like function operating as a probabilistic threshold for action potential generation.

2.3 Anti-Hebbian plasticity and steady state

Anti-Hebbian synaptic plasticity, reflecting multiple homeostatic metabolic pathways, acts to normalize excitatory and inhibitory synaptic gains, leaving the relative strengths of Hebbian influences unchanged (Keck et al., 2017). A steady state is maintained in which excitatory and inhibitory flux remain balanced so that where $\sum \varphi_e$ is the total excitatory presynaptic flux and $\sum \varphi_i$ is the total inhibitory presynaptic flux, then

$$\sum \Phi_e \rightarrow -\sum \varphi_i \rightarrow \text{constancy} \quad (5)$$

This steady-state assumption applies to summed afferent pulses reaching individual neurons as well as the population average; however, the constancy may vary with the level of network activation.

2.4 Goals of this model

From these assumptions about network properties, we must demonstrate how the external environmental field, $E(x, y, z, t)$,

where x, y, z, t are spatial and temporal coordinates, is mapped onto the developed form of the synaptic connections. That is, we must explain how synapses that converge to a static state can represent sequences of spatiotemporal sensory and motor images. The electrodynamics of synchronous oscillation provide a starting point for understanding.

3 Synchrony

3.1 Mechanism of synchrony

Synchronous oscillatory firing (Singer, 1993) of cortical neurons is prominent in the 40 Hz (gamma) range but is apparent over all frequencies. Possible mechanisms for its generation are multiple, as it may arise from identical inputs to separate neurons or through non-linear phase-locking of neurons at high or pathological levels of excitation. However, it can be shown that synchrony emerges in any network with zero-mean summing junctions interposed between elements, even when the network is driven by temporally and spatially diffuse white noise. Principal component spatial eigenmode analysis of simulations of the neural field reveals that the first principal component of the field exhibits widespread zero-lag synchrony with relatively low spatial damping and is the response to the even components of the inputs driving the field. The second principal component is characterized by low-magnitude, asynchronous, highly spatially damped, and rapidly dissipated activity and is the response to the odd components of the inputs (Chapman et al., 2002). These properties can account for the bulk of findings on synchronous oscillation in the brain (Wright et al., 2000). Closely related simulations that also account for other EEG properties, including the background spectrum and major cerebral rhythms (Wright and Liley, 1995, 1996; Robinson et al., 1997, 2001; Rennie et al., 2000, 2002), incorporate prominent zero-lag synchrony as an inherent property.

These numerical findings can be simply explained. Oscillation arises from the to-and-fro exchange of signals between excitatory neurons and local inhibitory neurons. Synchrony is then generated by summation of synaptic afferents at dendritic membranes, as in Equation 3, with a set level as in Equation 5. Synaptic pulse trains that arrive in opposite phase with respect to the sustained mean level cancel in summation, while those in phase summate positively—thus out-of-phase signals are dissipative, whereas in-phase pulses, when exchanged through the neural population, converge to coordinated synchrony at all frequencies. This applies to the summation of both excitatory and inhibitory synaptic fluxes at the membrane, since the inhibitory flux is effectively of the opposite phase to the excitatory flux. This argument is independent of the details of pulse generation and exchange statistics, as only the interactions of odd and even signal components are relevant.

Although point, limit cycle, or chaotic attractor dynamics may appear transiently in single neurons, this model is not restricted to such dynamics. A stochastic process results in synchrony as a ground state of electrocortical activity, consistent with observations that electrocortical waves exhibit energy equipartition and minimal free energy (Wright, 1989).

3.2 Synchrony as prototypical predictive error minimization

Since the exchange of signals between any two neurons tends to converge toward their synchronous firing, this is a prototypical example of mutual predictive error minimization leading to the maximization of mutual information, I . Roughly measured in linear approximation,

$$I = -\frac{1}{2} \log_2 (1 - r^2) \quad (6)$$

where r is the correlation coefficient of their synaptic flux exchange.

From this simple form of prediction error minimization between individual cells, the formation of geometrically ordered multiway exchanges between cell populations can be inferred, as follows.

4 The generation of mirror pairs

4.1 Minimization of free energy and flux equilibrium

Hebbian learning will lead to a fall in the variational free energy of presynaptic flux, F , as cross-connections form.

$$F = A - C \rightarrow 0 \quad (7)$$

where A is the total presynaptic flux autocorrelation, and C is the total presynaptic flux cross-correlation.

A further assumption is made here—that although the asymptotic limit cannot be reached in life, it will be sufficiently and continuously approached despite perturbation, to such a degree that connections will emerge observably close to those expected at the limit.

At the limit at which $F = 0$ (or approximately, if the limit is sufficiently closely approached), in the exchange of synaptic fluxes φ_{ij} and φ_{ji} , over all synaptic routes between cells at any positions i and j , and time-lag τ , the sum of their autocorrelation terms becomes equal to the sum of their cross-correlations:

$$\begin{aligned} \varphi_{ij}(t) \varphi_{ij}(t - \tau) + \varphi_{ji}(t) \varphi_{ji}(t - \tau) &= \varphi_{ij}(t) \varphi_{ji}(t - \tau) \\ &+ \varphi_{ji}(t) \varphi_{ij}(t - \tau) \end{aligned} \quad (8)$$

According to the free energy principle, the synaptic fluxes now represent the spatial and temporal associations in the network's inputs, and their probability distributions satisfy Bayes' theorem.

Absolute equilibrium requires that energy approach equipartition: $\varphi_{ij}(t) \rightarrow \varphi_{ij}(t - \tau) \rightarrow \varphi_{ji}(t) \rightarrow \varphi_{ji}(t - \tau)$. This condition is met when pairs of excitatory cells, or pairs of inhibitory cells, exchange equal and opposite flux at zero lag ($\tau = 0$), and when pairs of cells, one inhibitory, the other excitatory, exchange flux by firing with lag $\tau = \frac{|i-j|}{v}$. This is synchronous oscillation, in which excitatory or inhibitory pairs fire in phase at zero lag, and excitatory and inhibitory pairs fire in antiphase at half the period of oscillation. The ground state of universal synchrony is approached in the face of external perturbations.

Equation 8 admits many solutions in which free energy is zero, but energy is not equipartitioned. These cases represent forces perturbing the system and external inputs operating upon the neural network and are analogous to pdV forces in equilibrium thermodynamics, where free energy is equivalent to the Gibbs free energy.

4.2 Prediction error minimization and free energy gradient

Free energy gradients must vanish as the free energy is minimized—but since the system is under continuing input perturbations, then, where $\Delta\Phi^+$ is a vector representing flux induced by the externally imposed signals, there must arise an oppositely directed vector $\Delta\Phi^-$. In effect, the neural network predicts and neutralizes its inputs with minimal error, so

$$\Delta\Phi^+(t) - \Delta\Phi^-(t) \rightarrow 0; \frac{dF}{dt} \rightarrow 0, \text{ and } \frac{d^2F}{dt^2} \rightarrow -ve \quad (9)$$

Since all fluctuations are minimized, synaptic exchanges in zero-lag synchrony are maximized to the extent possible in the face of ongoing perturbation.

4.3 The emergence of mirror pair systems and excitatory/inhibitory balance

Fields of synchrony form the spatial eigenmodes of the network. Opposite and equal compensating asymmetric exchanges cannot be within the same eigenmode system, as this would result in a frozen steady state in which a representation of input time variations would not be possible (This would be the dark room outcome.). Therefore, the minimization of prediction errors requires that spatial eigenmodes develop in paired systems with mirror reversal, each with time-varying flux exchanges between eigenmodes oppositely directed to those in the mirror partner.

Paired mirror systems must also be able to maintain overall excitatory/inhibitory balance, as required by Equation 5. This can be provided by the collision of traveling waves at the line of interaction of the mirror pair, modulated by the adaptations of anti-Hebbian plasticity, and mediated by cross-couplings of all excitatory/inhibitory combinations, as diagrammed in Wright and Bourke (2024a,b). The line of collision is acting as a Markov blanket, with prediction errors in exchanges over the blanket minimized until equilibrium is reached. The paired structure conforms to the good regulator theorem (Conant and Ashby, 1970), which states that every good regulator must be a good model of the system it is controlling. Thus, each of the paired mirrors is a good regulator of the other. When the mirror pair arrangement is generalized to the multiway interactions of many systems, each approaching approximate mirror symmetry with its neighbors, complicated patterns of mutual regulation and distributed systems of partial generative models can arise, influencing one another.

Figure 1 (left) illustrates a system composed of a pair of mirror-symmetric, coupled spatial eigenmodes, each generating oppositely directed, colliding, traveling waves. The diagram shows

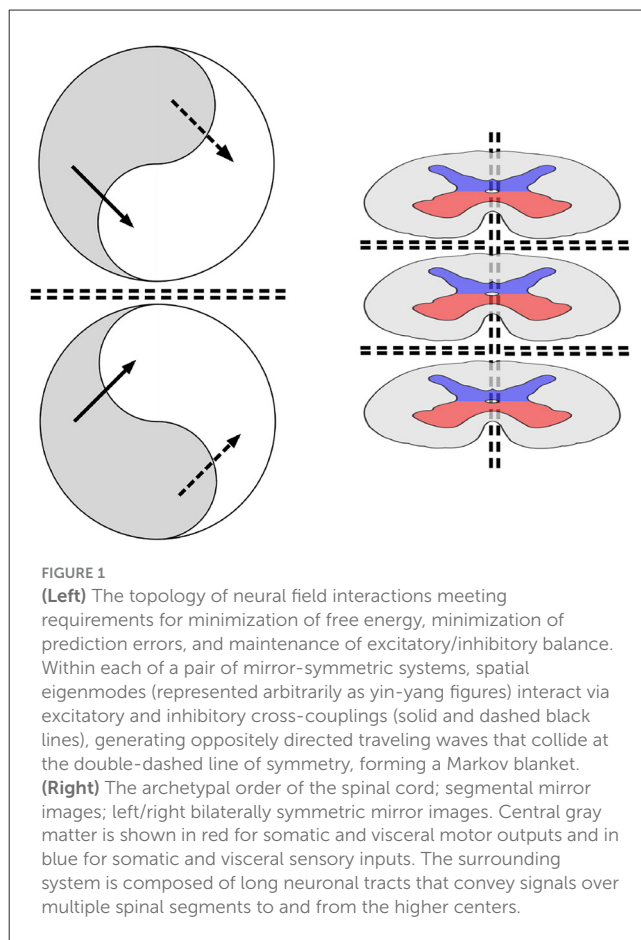


FIGURE 1

(Left) The topology of neural field interactions meeting requirements for minimization of free energy, minimization of prediction errors, and maintenance of excitatory/inhibitory balance. Within each of a pair of mirror-symmetric systems, spatial eigenmodes (represented arbitrarily as yin-yang figures) interact via excitatory and inhibitory cross-couplings (solid and dashed black lines), generating oppositely directed traveling waves that collide at the double-dashed line of symmetry, forming a Markov blanket. (Right) The archetypal order of the spinal cord; segmental mirror images; left/right bilaterally symmetric mirror images. Central gray matter is shown in red for somatic and visceral motor outputs and in blue for somatic and visceral sensory inputs. The surrounding system is composed of long neuronal tracts that convey signals over multiple spinal segments to and from the higher centers.

the topology of the connections and flux exchanges—not a specific topography. The mirror-twin eigenmode systems might be separated by some distance, or their cell soma positions might be interdigitated.

4.4 Signal exchanges across Markov blankets

It is emphasized that a Markov blanket is a mechanism for prediction error minimization, but it is not essential that free energy reach the zero limit. Perturbation of the system short of the limit will generate exchanges of signals across the blanket, enabling sequences of activity to be coordinated between mirror pairs, although exchanges across the blanket boundary maintaining excitatory/inhibitory balance result always in convergence toward signal mean and zero-lag synchrony on either side. This requires exchanges that may be mutually excitatory, mutually inhibitory, or involve excitatory/inhibitory cancellation. Exchanges may be out of phase and therefore dissipative, or in phase and conservative. Thus, perturbations of signals across the blankets can be information-rich, and multiway exchanges between assemblies of mirror pairs could organize sensorimotor sequences. This can be observed in the simplest form in the spinal cord.

4.5 A prototypical example: the spinal cord

All neural development depends upon cell differentiation in stages, from early precursors into populations of cells with characteristic differences in axonal and dendritic trees and paths of mobilization into anatomical order. Mirror symmetries are apparent from early in the evolution of vertebrates. [Figure 1](#) (right) shows cross-sections of a generic vertebrate spinal cord, introduced here as a harbinger of the development of higher centers.

In [Figure 1](#), right, the dorsal horns of gray matter, colored blue, mark the sites of sensory inputs to each spinal segment. The ventral horns, colored red, are the sites of motor commands sent out via the spinal nerves. The dorsal and ventral horns on each side in a single spinal segment can be regarded as forming a system of coupled spatial eigenmodes. Exchanges across each of the Markov blankets could then maximize the mutual information between sensory and motor signals in the left and right parts of the body, as well as between higher and lower spinal levels. Thus, elementary sensory experiences from within and without the organism, as well as motor expression, would be integrated in a generative model of self-in-the-world, in the simplest sense. Perturbations crossing these boundaries would enable the organization of activity in temporal sequences. This provides a principled basis for the well-studied mechanisms that coordinate locomotion in all creatures with a segmental nervous system, from worms to humans ([Danner et al., 2016](#)).

This symmetry could not be generated by the above hypothetical neural mechanisms alone. The structure must be genetically determined and originally selected in evolution by mechanical demands for articulation and the development of segmental reflexes, but it offers a starting point for the evolution of neural mirror symmetries in further development. It is well-established that the organization of the spinal cord extends into the brainstem and subcortex, preserving bilateral symmetry and the extension of structures related to the dorsal and ventral horns of the spinal cord. We assume without proof that the later evolved developments within the subcortical systems could themselves be decomposed into mirror pairs. Instead, we will show how the formation of mirror pairs is apparent in the formation of connections throughout the neocortex.

5 Embryogenesis

5.1 Selection of neurons and synapses early in the developing brain

Synchronous firing emerges early in neuronal development, coinciding with the establishment of small-world connectivity ([Downes et al., 2012](#); [Markov et al., 2011](#); [Vergara et al., 2019](#); [Bassett and Bullmore, 2006](#)). Concurrent apoptosis ([Hollville et al., 2019](#)) favors cells that fire with zero-lag synchrony ([Heck et al., 2008](#); [Sang et al., 2021](#)), apparently because synchrony promotes high metabolic turnover and resistance to cell suicide ([Vergara et al., 2019](#)). Initially, random action potential pulses can generate polysynaptic flows, within which bidirectional monosynaptic couplings may form between cells through symmetric exchanges. Arising first as rostral extensions of the early embryonic spinal

cord and brainstem, brain cell precursors that will give rise to the limbic system radiate outward to form the neocortex, in accordance with the Structural Model. As this expansion continues, cortical layers develop in depth, establishing pathways of signal flow and increasingly detailed local patterns of connectivity.

5.2 The structural model

The structural model provides an account of the embryogenesis of the neocortex ([Barbas, 1986](#); [Barbas and Rempel-Clower, 1997](#); [Barbas, 2015](#); [Garcia-Cabezas et al., 2019, 2020](#); [Sancha-Velasco et al., 2023](#); [Uceda-Heras et al., 2024](#)). Radial lines of cells grow outwards ([Sanides, 1962, 1964, 1970](#)) from two primary cell groups related to the hippocampus and olfactory allocortices, respectively, expanding in concentric rings. Cellular and molecular features vary systematically along the radial cortical gradient of laminar complexity ([Sancha-Velasco et al., 2023](#); [Garcia-Cabezas et al., 2017](#); [Puelles et al., 2019, 2024](#); [Rakic, 2002, 2007](#); [Barbas and Garcia-Cabezas, 2016](#); [Barbas and Garcia-Cabezas, 2015](#); [Garcia-Cabezas et al., 2019, 2020](#)).

As differentiation proceeds, cell connectivity also undergoes modification under the distance rule ([Markov et al., 2011](#); [Vezoli et al., 2021](#)), favoring cross-connections via the shortest and most locally dense pathways ([Aparicio-Rodriguez and Garcia-Cabezas, 2023](#)), thereby imposing circumferential connections among the radial lines. The radial lines and circumferential rings become folded into the neocortex, and at the most peripheral limit of development, they include the neurons in the primary sensory and motor cortices.

Along the radial lines of development, centrifugal signals from the limbic system interact with centripetal signals from the special sensory and motor cortex, bypassing each other between cortical layers in a counterflow between a hub-like center vs. the primary cortical areas that establish direct sensory and motor interactions with the environment ([Adams et al., 2013](#)). The counterflow is organized according to layers in the cortical depth, with layer complexity increasing from the limbic to the special sensory cortex. The flow outward from the limbic system passes from layers 5 and 6 centrally to layers 5 and 6 and layers 2 and 3 more peripherally. Meanwhile, the counterflow in layer 4, generated by peripheral inputs, interacts with neurons in layers 2 and 3 and is then relayed further via layer 4. Circumferential connections connect layers all-to-all.

The structural model reiterates ancestral neocortical evolution (citations above, [Steffen et al., 2022](#)), and the radial hierarchical order can be regarded as an ultimate extension of spinal segmental order. Radial development provides a hierarchical order for minimizing theoretical prediction error ([Shipp, 2007](#); [Shipp and Friston, 2022](#); [Bastos et al., 2012, 2020](#)).

5.3 Columns and patches at millimetric scale

Further order develops at the millimetric scale in the upper layers, where the superficial patch system ([Muir et al., 2011](#); [Muir](#)

and Douglas, 2011; Martin et al., 2014), made up of patches of cells that make lateral connections, skipping from patch to neighboring patches in several steps, forms gridworks believed to distribute information between all intracortical locales. Some parts of the cortex are organized into zones (Molnar, 2020; Horton and Adams, 2005), consisting of short-axon neurons surrounded by groups of superficial patch cells, which create variably resolved, often diffuse, macrocolumns. Within these macrocolumns, individual cells are organized according to their stimulus preferences (e.g., Obermayer and Blasdel, 1993). Similar neuron response preferences can be detected whether or not columns are present (Meng et al., 2012), so the presence or absence of columnar structure appears to reflect the degree to which the patch system and intervening shorter axon cells form discrete or merging systems.

Organized response columns in animals, which are clearly defined and have relatively long gestation periods, such as rhesus monkeys, appear before birth (Wiesel and Hubel, 1974). Thus, their formation does not depend upon organized visual inputs. Yet, when deprived of organized visual inputs during the post-natal period, the response architecture collapses (Blakemore and Van Sluyters, 1975; Sherk and Stryker, 1976).

This embryological order can be interpreted as the formation of mirror pairs in a number of distinct topologies, initially driven antenatally by random pulsation and subsequently responding post-natally to ordered inputs.

6 All mirrors

6.1 Whole brain and interareal approximations to mirror symmetries

In gross anatomy, approximate left/right mirror symmetry extends from the spinal cord to the cerebral hemispheres. Approximate mirror symmetries are also apparent throughout the cerebral white matter, since interareal cortico-cortical connections form U-shaped projections, each area tending to mirror others to which it projects reciprocally.

6.2 Hierarchical counterflows as mirror symmetry

The lines of counterflow in the structural model, whereby interactions between signal streams in a hierarchical order occur in cortical layers 3 and 4, can be construed as a mirror reflection of a mirror pair, with the layer 3/4 interaction interpreted as an extended Markov blanket. This aligns with the influential theory of prediction error minimization proposed by Bastos et al. (2012, 2020), in which prediction errors are minimized through excitatory/inhibitory interactions in counterflows.

We suggest that the Bastos et al. models are a special case of a more general and ubiquitous process at all scales—and particularly at millimetric scales—in which paired, mirror-symmetric cells and synaptic arrangements develop, thereby greatly increasing the degrees of freedom for signal interchange and facilitating overall minimization of prediction error.

6.3 Simulations of the generation of mirror maps at mesoscale

Detailed accounts of how the selection of developing neurons for maximum synchrony by apoptosis can shape anatomical ordering and the tuning of individual cortical neurons are given in Wright and Bourke (2013, 2016, 2022, 2023, 2024a,b); Wright et al. (2014).

Simulations (Wright and Bourke, 2013, 2016) were devised to show that continuity between columnar and non-columnar cortex could be explained when total synchrony was maximized and axonal lengths were minimized according to the rules of ultra-small world connectivity. We employed a simple force equilibrium algorithm that leveraged the unit-dimension analogy between mechanical force and synaptic flow, considering the axonal lengths of two populations—specifically, those with long vs. short axons. With an emphasis on small-world cell organization alone, the pattern of cell connections was diffuse. With emphasis on synchrony maximization alone, short axon cells formed clusters surrounded by pools of the long axon cells, the latter in systems with hexagonal, square, or irregular tiling.

With both optimization criteria considered concurrently, simulation outcomes varied depending on the absolute and relative lengths of the short vs. the long axons. Where all axonal lengths were shorter, and when the long axon lengths were closer to those of the short axons, then a more diffuse small-world order predominated. With increasing difference of relative axonal lengths, a more clearly columnar structure emerged. This accounted for the variation in columnar and non-columnar appearances, as well as the ordering of patch cells, with the implication that a single modular system is common throughout, and modular interpenetration in apparently non-columnar areas is made possible by the sparsity of connectivity throughout the cortex.

Using the patterns of cell body positions generated in the simulations, we next reconstructed the patterns of symmetric bidirectional monosynaptic connections that should emerge as synchrony is maximized in the growing network. These matched the distribution of synapses formed by superficial patch cells, with a neighborhood arrangement of cortical columns in mirror symmetry to each neighbor. The closest approximation to mirror symmetry between columns required patch cells in square patterns, with broken symmetry where patch connections were in a hexagonal array.

The requirement to maximize synchrony with bidirectionally symmetric connections in all cases implied a rather strange but necessary arrangement in the connection between patch cells and the cells within each column. The interpenetration of the sparse networks implied that maximization of synchrony among the short axon cells required their linkage in closed loops, while the connections made to the short axon cells by the patch cells would be constrained within arcs radiating from columnar centers. The upshot was that projection from the larger patch cell scale to the shorter axon cell scale would be similar to the projection of a Euclidean plane to a Möbius strip in a Riemann projection. We referred to this as a projection from a “global” to a “local map” scale,

which could be represented as follows:

$$P \leftrightarrow p \text{ where } p = \left\{ \pm \sqrt{-1}k \frac{(P-p_0)^2}{|P-p_0|} + p_0 \right\} \quad (10)$$

where P is a position at the global scale of the patch system expressed as a complex number, and p is a corresponding position within one of the local maps. Curly brackets indicate that there is a set of such maps tiling the cortical surface. $\frac{(P-p_0)^2}{|P-p_0|}$ describes the doubling of the angle in the projection from P to p , $\sqrt{-1}k$ defines rotation by 90 degrees, and the scale of the projection from global to local scales. Chirality is indicated as \pm , and p_0 is the center of a short-axon cell cluster. This implies that a mirroring relationship is created between scales, as well as mirror pairs being generated between adjacent column-like systems throughout the cortex.

6.4 Explanatory power of the mesoscale mirror pair model

The above model offers a unitary explanation of the following experimental observations:

- Columnar vs. non-columnar variation (Wright and Bourke, 2016).
- The antenatal development of retinotopic maps and their post-natal loss with deprivation of visual experience (Wright and Bourke, 2013).
- Retinotopic response maps in cortical area V1—Orientation Preference (OP) singularities, linear zones, and saddle points. A Möbius-like internal configuration within a column accounts for the organization of OP from 0 to 180 degrees over the 360 degrees around an OP singularity. Linear zones and saddle points of OP are accounted for by mirror-symmetric reversals with some broken symmetry at columnar boundaries (Wright and Bourke, 2013, 2016).
- Ocular dominance (OD) columns in a square array (Wright and Bourke, 2024a).
- “Like to like” patch cell connections—i.e., the preferential connections of patch cells to columnar cells with a common OP (Wright and Bourke, 2013, 2024b).
- The numerical relationship $TFP = \text{object velocity} \times \text{SFP}$ between preferred spatial frequency (SFP) and preferred temporal frequency (TFP) in individual cortical neurons (Wright and Bourke, 2022).
- The differential distribution of SFP and TFP in relation to OP (Wright and Bourke, 2022).
- OP variation in response to moving line stimuli of given orientation and varied speed, angle of attack, and length (Wright and Bourke, 2013)—as opposed to fixed OP selection on the input pathway alone.
- Recordings made by vertical penetrations of the somatosensory cortex showing that receptor fields for cells exhibit continuities of the fields as the recording electrode is advanced, as well as sudden breaks and reversals of continuity—consistent with the organization of local cell connections in a Möbius-like configuration that spirals within the cortical layers (Wright et al., 2014).

6.5 Orderly organization of mirror pairs

As development unfolds, initially highly asymmetric signal exchanges occur between layers 2, 3, and 4 in radial development and between all layers in circumferential development. These exchanges would proceed imperfectly, with later and local self-organization further minimizing total free energy in a manner similar to the formation of eddies in turbulent flow.

Figure 2 shows the stages of the proposed development of the mirror pair organization.

In Figure 2a, the central long dashed arrow indicates the orientation of the neuraxis, with the three sections of the spinal cord representing the cord and brain stem. The fine dashed radial lines indicate the direction of embryonic development from limbic origins toward special cortical sensory/motor areas. Lateral intra-cortical and cortico-cortical connections (gray arrows) add circumferential connectivity.

The radial development in depth of the cortical layers of the ventral and dorsal divisions, respectively, is shown on the left and right hand sides from center. Laterally projecting red arrows indicate cortico-cortical connections projecting from higher in the sensory hierarchy to lower, and diagonal blue arrows mark the counterflow of signals from lower to higher in the hierarchy. Small blue vertical arrows mark the onward projections of axons preferentially directed toward the surface from layer 4 to layers 2 and 3, and small red vertical arrows mark counterflows of traffic from layers 2 and 3. Interaction between the counterflows at layers 2, 3, and 4 is characterized as a Markov blanket, indicated by the double-dashed lines.

Figure 2b represents the development of interareal and interhemispheric projection via U-shaped cortico-cortical connections, with approximate mirror symmetry. Superficial patch cells and columnar cells are shown embedded in the cortical sheets (An area on the lower left is shown with overlapping patch connections as a reminder that apparent columnar order can be absent without loss of modular order).

Figure 2c shows the arrangement of superficial patch cells on the cortical surface, forming a communication network between column-like groups of shorter axon excitatory cells that connect in an interpenetrating meshwork, akin to systems of interlocked Möbius strips. Projections from the patch cells to the columns thus form a mirroring between scales at millimetric distances.

Figure 2d illustrates how the maximization of synchrony between adjacent columns necessitates mirrored arrays in neighborhoods, with a degree of broken symmetry imposed by the roughly hexagonal tiling of the patch cell network.

Figure 2e, shown in cortical depth as in Figure 2a, shows how the mirror arrangement would develop as lateral connectivity in layers 2/3 and 5/6, respectively.

7 Dimensions and spatiotemporal images

7.1 Dimensional consistency in synaptic storage of a generative model

Sensory and motor events can be regarded as images in four-dimensional space-time, arriving in the cortex as image projections

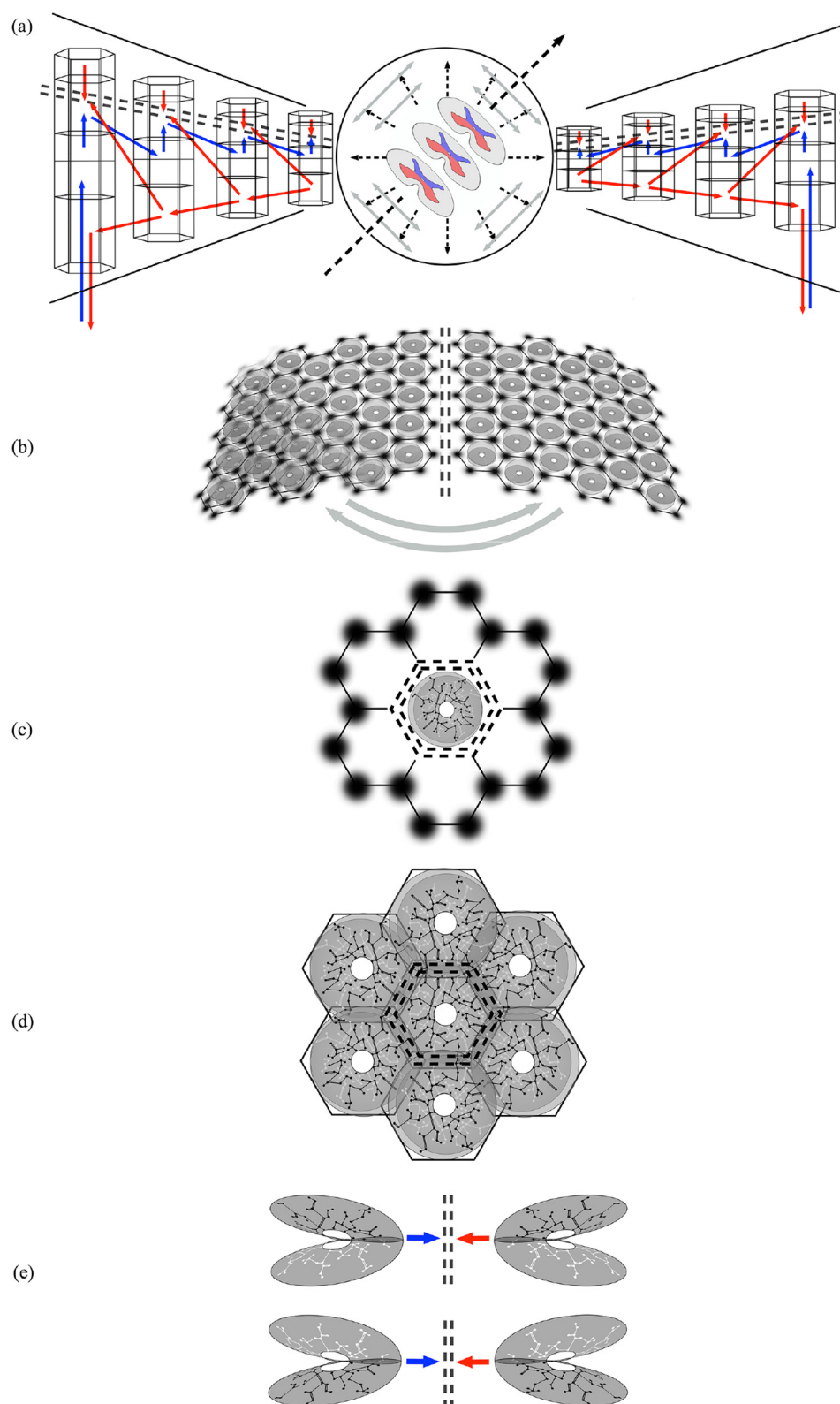


FIGURE 2

Mirrored representations in approximate developmental sequence. Double black dashed lines indicate lines of mirror symmetry and associated Markov blankets. Red, blue, and gray arrows indicate presynaptic flow along the radial and circumferential lines of development. **(a)** Development along the radial lines of the structural model (see text) indicates the interaction of counterflows at intervening Markov blankets. **(b)** Circumferential development between cortical areas. **(c)** Superficial patch cells, depicted as dark patches, form a communication network that generates local maps in a Möbius strip-like configuration, producing a mirroring between scales at the mesoscale. **(d)** Adjacent local maps, with varying degrees of overlap, interact with each other across homologous mapping positions within Markov blankets. **(e)** Along the lines of radial development, mirror pairs can form in layers 2/3 and layers 5/6.

to a two-dimensional surface. Specific images in a time-varying three-dimensional world require an object scale/rotation frame of three dimensions plus a position reference frame of three dimensions. The inclusion of information on image velocities then requires a translation within both of these frames. Therefore, when compressed to an effectively flat surface and stored as static cortical synaptic connections, image representations must be defined by sets of twelve points, and these sets must be laid out within an equivalent of a four-dimensional frame—arising as follows.

Where O is a pattern of activity at the scale of the patch network, and o is the projected image to any given column-like short-axon system, then, according to mapping (10)

$$O(P, t) \leftrightarrow \left\{ o \left(\pm(p - p_0)^2, t - \frac{|P-p|}{v} \right) \right\} \quad (11)$$

Representation of $O(P, t)$ has been dispersed across a set of synaptic connections that form a four-dimensional reference framework, with positions defined by the complex numbers, P (global map) and $p - p_0$ (within each of the set of local maps). Within this frame, associations within sets of twelve neurons can define spatiotemporal images.

Because (10) and (11) are invertible 1:1 maps, and since projections at a large scale are topographically ordered and receive and project to primary sensory and motor areas as topographic 1:1 projections, generative models developed at the mesoscale would maintain topological consistency.

7.2 Broken symmetry and sequences of spatiotemporal images

Organizing into mirror pairs would minimize, but not eliminate, broken symmetry. As the gradient of free energy diminishes, in contrast to the idealization of Equation 9, with a degree of asymmetry, an unstable, rather than a stable, fixed point would be approached. i.e.,

$$\text{as } F \text{ and } \frac{dF}{dt} \longrightarrow 0 \quad \frac{d^2F}{dt^2} \rightarrow +ve \quad (12)$$

This may be favorable rather than fatal in functionality. Spatiotemporal images must be stored in sequences much longer than the timescale of intracortical conduction. Longer sequences require cortical areas to activate each other sequentially, creating cognitive order. Therefore, asymmetry would enable the switching and manipulation of spatiotemporal images on longer time scales.

8 The evolution of agency

Given the above properties, it is possible to argue that agency can emerge progressively in the evolutionary sequence. The caudal-to-rostral evolutionary sequence can be regarded as a progression from hard-wired to soft-wired connectivity, as is apparent in the progression from lower plasticity to higher plasticity seen in enzyme expression in the radial lines of the structural model (Garcia-Cabezas et al., 2017, 2019; Gomez et al., 2019; Sancha-Velasco et al., 2023). Concurrently, the expansion of mirror pairs in circumferential order provides for parallel and feedback relations

with higher degrees of freedom than the hierarchical order alone—a systematic variant of sleep/waking network optimization (Hinton et al., 1995, 2006), in which network connectivity evolves toward optimum in the partially randomized dreaming phase.

We can compare the neocortex, which is considered soft-wired, vs. the limbic, more caudal components, and apply the Nyquist and Shannon relations to measure the channel capacities of both parts of the CNS. The separation into neocortical and limbic systems, as a division between soft- and hard-wired systems, is a convenience, but the argument is general and applicable at various stages as the number of neurons increases in a caudal-to-rostral sequence throughout evolution.

Treating each of n synaptic fluxes as binary valued for simplicity, the channel capacity of the entire CNS network, D , measured in bits, is as follows:

$$D = n \log_2 \left(1 + \frac{C}{A} \right) \quad (13)$$

where C and A are the total cross-correlation and autocorrelations of synaptic flux (Equation 7). $\frac{C}{A}$ is equivalent to a signal/noise ratio, and as the asymptotic limit is approached, $C \rightarrow A$ and therefore $D \rightarrow n$.

If the neocortical component is composed of n_{neo} fluxes and a limbic and subcortical component of n_{limb} fluxes, then there are $2^{n_{neo}}/2$ distinguishable neocortical states in mirror pairs and $2^{n_{limb}}/2$ limbic states.

Survival requires that an organism avoid conditions incompatible with any of the strongly inherited limbic states that constrain and bind its behavior. However, if $n_{neo} > n_{limb}$, there are $2^{n_{neo}-n_{limb}}$ possible neocortical states available for each limbic state. These neocortical states arise internally through couplings that minimize synaptic free energy. Consequently, as an organism interacts with its environment, various combinations of neocortical and limbic states are generated, and some of the internal neocortical states, by linking to limbic states, gain behavioral expression. The expanded range of states enables the emergence of novel responses that can be tested for their survival potential—either incurring risk or achieving unprecedented success—and, if successful, become incorporated into the organism's behavioral repertoire. This mechanism may underlie the phylogenetic evolution of agency. A parallel process can be proposed for individual development.

9 Conclusion

9.1 Summing up

We have argued that prediction error minimization is built into the signal exchanges between every neuron in the brain and every other, so that cortical neural networks tend to always maximize synchronous oscillation in the face of ongoing perturbations. Flexible evolution toward this elementary equilibrium at fast, intermediate, and slow time scales of synaptic consolidation would impose the formation of neuronal assemblies in mirror-pair topologies at micro-, meso-, and macroscales, and between scales. This, in turn, allows a four-dimensional generative model to emerge, mapping the four-dimensional interactions of a self in the world.

We have extended the argument to demonstrate that it is consistent with the evolution of the brain being shaped by apoptotic selection interacting with genetic sequences and chemical organizers and, by allowing the addition of extra and flexible information-processing capacity, to progression toward increasing agency.

This theoretical account, if valid, goes some way to countering strong criticisms that have been made of the Bayesian brain and the free energy principle. We have argued that information management is not limited to hierarchical processing, and the primary mechanism of synchrony is robust to the effects of non-linearities and statistical variations that deviate from Gaussian ideals.

9.2 Comparisons and further considerations

Our proposal builds upon and extends earlier concepts of brain function and organization.

We have earlier mentioned adherence to the “good regulator theorem,” which requires control systems to model what is controlled and allows for interacting and collaborative control subsystems. The classical Hebbian concept of phase sequences and cell assemblies (Hebb, 1949) could be expressed in terms of coupled spatial eigenmodes. The same notion of self-organization into dynamic groups of neurons with partial computational autonomy, bounded by surrounding Markov blankets, is expressed in the Neuronal Packet Hypothesis (Yufik and Friston, 2016; Yufik, 2019). The present view posits that minimization of prediction error is a universal principle in the interactions of any pair of neurons, as well as in emergent groups with mirror symmetry.

The development of this skeletal concept of neural function into a more comprehensive model of brain function is beyond the scope of this paper; however, areas of interest include, firstly, the interactions between limbic/neocortical systems. Functional correspondences between the structural model and alerting, surprise, habituation, and sleep have been analyzed by Tucker et al. (2022), Tucker and Luu (2021, 2023), Luu and Tucker (2023), Luu et al. (2024), and Tucker et al. (2025). The activation and regulatory roles of limbic components, positioned as they are at the roots of neocortical development and therefore at the crossroads of signal control, offer a guide to the way the old brain and new brain collaborate in the development toward increasing agency. By minimizing the risk of fatal outcomes, they would influence overall “policy” and, in maneuvering toward optimum adjustment of limbic and neocortical components, would be executing active inference (Ramstead et al., 2021; Kavi et al., 2025). Moreover, the role of sleep proposed in these models draws upon the resemblance to Hinton’s “dreaming” networks (earlier cited) and the idea of self-organization through the maximization of synchrony in the presence of noise, with noise and high degrees of freedom assisting movement toward stable basins of attraction. This aligns with the present model’s essentially steady-state dynamics, characterized by convergence after perturbation toward maximum synchrony.

Secondly, the interaction of cortical areas, e.g., Yeo networks (Yeo et al., 2011), requires explanation. High asymmetry in the

mirror matching of many areas would limit their demarcation via unstable Markov blankets, not only due to the conformation of their boundaries and connections, but also to the mismatching of spatial frequencies resulting from differing fiber ranges for cells in different areas. This would result in relatively weak coupling between some cortical areas vs. others. Such relative isolation between cortical areas might account for cortical area specializations.

Finally, the idea of the brain being organized into systems with mirror symmetries invites comparison with the role of mirror neurons (Bonini et al., 2022). Since mirror neurons show that the neurons of individuals watching other individuals respond as if the observer were experiencing the other’s sensations and actions, it might be thought that the term “mirror” may be merely a linguistic link. This is not necessarily the case. The internal generation of mirror neuron activity might itself arise as an imposition upon mirroring of sensory vs. motor activity within the “host.” If so, then there is scope for extending the concept of mirroring more widely than just the individual and the immediate environment, but also to the world of interpersonal and social interactions.

9.3 Testing against connectomic data

The occurrence of approximate mirror symmetries at spinal and telencephalic levels is readily apparent, as is the mirroring between cortical areas by cortico-cortical fibers. To a first approximation, the existence of mirror symmetries is uncontroversial—but a critical proposition requiring testing is that superficial patch connections project to shorter axon cells within columnar-sized areas in a manner analogous to the projection of a Euclidean plane onto a Möbius strip. The existence of such structural order is theoretically required for topographic consistency in signal exchanges, and the possible existence of this hypothetical structure is already supported by a range of indirect observations. Three-dimensional schematic diagrams reconstructing the expected connectivity are given in Wright and Bourke (2024b).

It is necessary to show that among the many sparsely connected neurons within a column, patch cells with somas diametrically opposite across any column make synaptic connections in arcs radiating outwards from the central OP singularity and do so over superimposed arcs outwards from the singularity in surface view—but, crucially, that opposite patch cells project to closely interpenetrating, but separate, systems of local cells. This should apply systematically to all patch cell groups around each singularity and should hold for each patch cell’s connections to all columns in the vicinity.

Testing hypotheses in this way faces considerable difficulty, particularly with data from small animals or in cortical data obtained from non-columnar cortex. In these cases, the analysis must deal with the interpenetration of the hypothetical columns, as well as the intertwining of local connections within each column. Data for the mouse brain obtained in the MICrONS Project (The MICrONS Consortium, 2025) have a suitable resolution; however, since the mouse is small, the connectivity to be analyzed will face the problem of interpenetration. Mosaic-like connectivity

surrounding the line of minicolumns is revealed in the MICrONS mouse data (Miyashita, 2025). This corresponds to the anticipated interweaving of cells found with vertical electrode penetrations (see Section 6.4), but explicit consideration has not been given to patch cells in relation to the interwoven local cells in the connectome data. An easier test case would require data from the columnar visual cortex of a large mammal.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

JW: Project administration, Writing – review & editing, Writing – original draft, Conceptualization, Investigation, Supervision. PB: Writing – review & editing, Software, Data curation, Visualization.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

References

- Adams, R. A., Shipp, S., and Friston, K. (2013). Predictions not commands: active inference in the motor system. *Brain Struct. Funct.* 218, 611–643. doi: 10.1007/s00429-012-0475-5
- Aparicio-Rodriguez, G., and Garcia-Cabezas, M. A. (2023). Comparison of the predictive power of two models of cortico-cortical connections in primates: the distance rule model and the structural model. *Cereb. Cortex* 33, 8131–8149. doi: 10.1093/cercor/bhad104
- Barbas, H. (1986). Pattern in the laminar origin of cortico-cortical connections *J Comp Neurol* 252, 415–422. doi: 10.1002/cne.902520310
- Barbas, H. (2015). General cortical and special prefrontal connections: principles from structure to function. *Annu. Rev. Neurosci.* 38, 269–289. doi: 10.1146/annurev-neuro-071714-033936
- Barbas, H., and Garcia-Cabezas, M. A. (2016). How the prefrontal executive got its stripes. *Curr. Opin. Neurobiol.* 40, 125–134. doi: 10.1016/j.conb.2016.07.003
- Barbas, H., and Garcia-Cabezas, M. A. (2015). Motor cortex layer 4: less is more. *Trends Neurosci.* 38, 259–261. doi: 10.1016/j.tins.2015.03.005
- Barbas, H., and Rempel-Clower, N. (1997). Cortical structure predicts the pattern of cortico-cortical connections. *Cereb. Cortex* 7, 635–646. doi: 10.1093/cercor/7.7.635
- Bassett, D. S., and Bullmore, E. (2006). Small-world brain networks. *Neuroscientist* 12, 512–513. doi: 10.1177/1073858406293182
- Bastos, A. M., and Lundkvist, M., Waite A. S., Miller E. K. (2020). Layer and rhythm specificity for predictive routing. *Proc. Natl. Acad. Sci. U. S. A.* 117, 31459–31469. doi: 10.1073/pnas.2014868117
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., Friston, K. F., et al. (2012). Canonical microcircuits for predictive coding *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038
- Blakemore, C., and Van Sluyters, R. C. (1975). Innate and environmental factors in the development of the kitten's visual cortex. *J. Physiol.* 248, 663–716. doi: 10.1113/jphysiol.1975.sp010995
- Bonini, L., Rotunno, C., Arcuri, E., and Gallese, V. (2022). Mirror neurons 30 years later: implications and applications. *Trends Cogn. Sci.* 26, 767–781. doi: 10.1016/j.tics.2022.06.003
- Chapman, C. L., Wright, J. J., and Bourke, P. D. (2002). Spatial eigenmodes and synchronous oscillation: coincidence detection in simulated cerebral cortex. *J. Math. Biol.* 45, 57–78. doi: 10.1007/s002850200141
- Clarke, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 35 181–204 doi: 10.1017/S0140525X12000477
- Conant, R. C., and Ashby, W. R. (1970). Every Good Regulator of a system must be a model of that system. *Int. J. Systems Sci.* 1, 89–97. doi: 10.1080/00207727008920220
- Danner, S. M., Wilshin, S. D., Shevtsova, N. A., and Rybak, I. A. (2016). Central control of interlimb coordination and speed-dependent gait expression in quadrupeds. *J. Physiol.* 8, 6947–6967. doi: 10.1113/JP272787
- Downes, J. H., Hammond, M. W., Xydias, D., Spencer, M. C., Becerra, V. M., Warwick, K., et al. (2012). Emergence of a small-world functional network in cultured neurons. *PLoS Comput. Biol.* 8:e1002522. doi: 10.1371/journal.pcbi.1002522
- Fields, C., and Levin, M. (2019). Somatic multicellularity as a satisficing solution to the prediction-error minimization problem. *Commun. Integr. Biol.* 12, 119–132. doi: 10.1080/19420889.2019.1643666
- Friston, K. (2002). Functional integration and inference in the brain. *Prog. Neurobiol.* 68, 113–143. doi: 10.1016/S0301-0082(02)00076-X
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:e1000211. doi: 10.1371/journal.pcbi.1000211
- Friston, K. (2010a). The free energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Friston, K. (2010b). Is the free-energy principle neurocentric? *Nat. Rev. Neurosci.* 11:605. doi: 10.1038/nrn2787-c2
- Friston, K., Thornton, C., and Clark, A. (2012). Free energy minimization and the dark room problem. *Front. Psychol.* 3:130. doi: 10.3389/fpsyg.2012.00130
- Friston, K. J., Fitzgerald, T., Rigoli, F., Swartenbeck, P., and Pezzulo, G. (2017). Active inference: a process theory. *Neural Comput.* 29, 1–49. doi: 10.1162/NECO_a_00912

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Garcia-Cabezas, M. Á., Hacker, J. L., and Zikopoulos, B. (2020). A protocol for cortical type analysis of the human neocortex applied on histological samples, the Atlas of Von Economo and Koskinas, and magnetic resonance imaging. *Front. Neuroanat.* 14:576015. doi: 10.3389/fnana.2020.576015
- Garcia-Cabezas, M. A., Joyce, M. K. P., John, Y. J., Zikopoulos, B., and Barbas, H. (2017). Mirror trends of plasticity and stability indicators in primate prefrontal cortex. *Eur. J. Neurosci.* 46, 2392–2405. doi: 10.1111/ejn.13706
- Garcia-Cabezas, M. A., Zikopoulos, B., and Barbas, H. (2019). The structural model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Struct. Funct.* 224, 985–1008. doi: 10.1007/s00429-019-01841-9
- Gomez, J., Zhen, Z., and Weiner, K. S. (2019). Human visual cortex is organized along two genetically opposed hierarchical gradients with unique developmental and evolutionary origins. *PLoS Biol.* 17:e3000362. doi: 10.1371/journal.pbio.3000362
- Hebb, D. O. (1949). *The Organization of Behaviour*. New York, NY: Wiley and Sons.
- Heck, N., Golbs, A., Riedemann, T., Sun, J.-J., Lessman, V., and Luhmann, H. J. (2008). Activity dependent regulation of neuronal apoptosis in neonatal mouse cerebral cortex. *Cereb. Cortex* 18, 1335–1349. doi: 10.1093/cercor/bhm165
- Hinton, G. E., Dayan, P., Frey, B. J., and Radford, M. N. (1995). The “wake-sleep” algorithm for unsupervised neural networks. *Science* 268 1158–1161. doi: 10.1126/science.7761831
- Hinton, G. E., Osindero, S., and The, Y.-., W. (2006). A fast learning algorithm for deep belief nets. *Neural Comput.* 18, 1527–1554. doi: 10.1162/neco.2006.18.7.1527
- Hohwy, J. (2016). The self-evidencing. *Brain* 139, 259–285. doi: 10.1111/nous.12062
- Hollville, E., Romero, S. E., and Deshmukh, M. (2019). Apoptotic cell death regulation in neurons. *FEBS J.* 286, 3276–3298. doi: 10.1111/febs.14970
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. U. S. A.* 79 2554–2558. doi: 10.1073/pnas.79.8.2554
- Horton, J. C., and Adams, D. L. (2005). The cortical column: a structure without a function. *Philos. Trans. R. Soc.* 360, 837–862. doi: 10.1098/rstb.2005.1623
- Izhikevich, E. M., and Desai, N. S. (2003). Relating STDP to BCM. *Neural Comput.* 15, 1511–1523. doi: 10.1162/089976603321891783
- Kavi, P. C., Zamora-Lopez, G., Friedman, D. A., and Patow, G. (2025). Thoughtseeds: a hierarchical and agentic framework for investigating thought dynamics in meditative states. *Entropy* 27:459. doi: 10.3390/e27050459
- Keck, T., Toyozumi, T., Chen, L., Doiron, B., Feldman, D. E., Fox, K., et al. (2017). Integrating Hebbian and homeostatic plasticity: the current state of the field and future research directions. *Philosoph. Trans. R. Soc. B.* 372:1715. doi: 10.1098/rstb.2016.0158
- Kiebel, S. J., and Friston, K. J. (2011). Free energy and dendritic self-organization. *Front. Syst. Neurosci.* 5:80. doi: 10.3389/fnsys.2011.00080
- Kirchhoff, M., Parr, T., Palacios, E., Friston, K., and Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. *J. R. Soc. Interface* 15:20170792. doi: 10.1098/rsif.2017.0792
- Kwisthout, J., and van Rooij, I. (2020). Computation resource demands of a predictive Bayesian brain. *Comput. Brain Behav.* 3, 174–188. doi: 10.1007/s42113-019-00032-3
- Levin, M. (2019). The computational boundary of a “self”: developmental bioelectricity drives multicellularity and scale-free cognition. *Front. Psychol.* 10:2688. doi: 10.3389/fpsyg.2019.02688
- Luu, P., and Tucker, D. M. (2023). Continuity and change in neural plasticity through embryonic morphogenesis, fetal activity-dependent synaptogenesis and infant memory consolidation. *Dev. Psychobiol.* 65:e22439. doi: 10.1002/dev.22439
- Luu, P., Tucker, D. M., and Friston, K. (2024). From active affordance to active inference: vertical integration of cognition in the cerebral cortex through dual subcortical control systems. *Cereb. Cortex* 34:bhad458. doi: 10.1093/cercor/bhad458
- Mangalam, M. (2025). The myth of the Bayesian brain. *Eur. J. Appl. Physiol.* 125, 2643–2677. doi: 10.1007/s00421-025-05855-6
- Markov, N. T., Misery, P., Falchier, A., Lamy, C., Vezoli, J., Quilodran, R., et al. (2011). Weight consistency specifies regularities of macaque cortical networks. *Cereb. Cortex* 21, 1254–1272. doi: 10.1093/cercor/bhq201
- Martin, K. A. C., Roth, S., and Rusch, E. S. (2014). Superficial layer pyramidal cells communicate heterogeneously between multiple functional domains of cat primary visual cortex. *Nat. Commun.* 5:55252. doi: 10.1038/ncomms6252
- Meng, Y., Tanaka, S., and Poon, C.-S. (2012). Comment on “Universality in the evolution of orientation columns in the visual cortex. *Science* 336:413. doi: 10.1126/science.1205737
- Miller, K. D. (1996). Synaptic economics : competition and cooperation in synaptic plasticity. *Neuron* 17, 371–374. doi: 10.1016/S0896-6273(00)80169-5
- Miyashita, T. (2025). Complicated architecture of cortical microcircuit: a comprehensive review. *Anat. Sci. Int.* 100, 614–634. doi: 10.1007/s12565-025-00877-8
- Molnar, Z. (2020). “Cortical columns,” in *Comprehensive Developmental Neuroscience: Neural Circuit Development and Function in the Brain*, Vol. 3, eds. J. L. R. Rubenstein and P. Rakic (Amsterdam: Elsevier), 109–129.
- Muir, D. R., Da Costa, N. M. A., Girardin, C. C., Naaman, S., Omer, D. B., Ruesch, E., et al. (2011). Embedding of cortical representations by the superficial patch system. *Cereb. Cortex* 21, 2244–2260. doi: 10.1093/cercor/bhq290
- Muir, D. R., and Douglas, R. J. (2011). From neural arbours to daisies. *Cereb. Cortex* 21, 1118–1133. doi: 10.1093/cercor/bhq184
- Obermayer, K., and Blasdel, G. G. (1993). Geometry of orientation and ocular dominance columns in monkey striate cortex. *J. Neurosci.* 13, 4114–4129. doi: 10.1523/JNEUROSCI.13-10-04114.1993
- Okamoto, H., and Ichikawa, K. (2000). A model for molecular mechanisms of synaptic competition for a finite resource. *BioSystems* 55, 65–71. doi: 10.1016/S0303-2647(99)00084-2
- Puelles, L., Alonso, A., and Garcia-Calero, E. (2024). genoarchitectural definition of the adult mouse mesocortical ring: a contribution to cortical ring theory. *J. Comp. Neurol.* 532:e25647. doi: 10.1002/cne.25647
- Puelles, L., Alonso, A., and Garcia-Calero, E., Martinez-de-la-Torre, M. (2019). Concentric ring topology of mammalian cortical sectors and relevance for patterning studies. *J. Comp. Neurol.* 527, 1731–1752. doi: 10.1002/cne.24650
- Rakic, P. (2002). Neurogenesis in adult primate neocortex: an evaluation of the evidence. *Nat. Rev. Neurosci.* 3, 65–71. doi: 10.1038/nrn700
- Rakic, P. (2007). The radial edifice of cortical architecture: from neuronal silhouettes to genetic engineering. *Brain Res. Rev.* 55, 204–219. doi: 10.1016/j.brainresrev.2007.02.010
- Ramstead, M. J., Hesp, C., Tschantz, A., Smith, R., Constant, A., Friston, K., et al. (2021). Neural and phenotypic representation under the free energy principle. *Neurosci. Biobehav. Rev.* 120, 109–122. doi: 10.1016/j.neubiorev.2020.11.024
- Ramstead, M. J. D., Sakthivadivel, D. A. R., Heins, C., Koudahl, M., Millidge, B., Da Costa, L., et al. (2022). On Bayesian mechanics: a physics of and by beliefs. *Interf. Focus* 13:20220029. doi: 10.1098/rsfs.2022.0029
- Rennie, C. J., Robinson, P. A., and Wright, J. J. (2002). Unified neurophysical model of EEG spectra and evoked potentials. *Biol. Cybern.* 86 457–471. doi: 10.1007/s00422-002-0310-9
- Rennie, C. J., Wright, J. J., and Robinson, P. A. (2000). Mechanisms of cortical electrical activity and emergence of gamma rhythm. *J. Theor. Biol.* 205, 17–35. doi: 10.1006/jtbi.2000.2040
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* 56 826doi: 10.1103/PhysRevE.56.826
- Robinson, P. A., Rennie, C. J., Wright, J. J., Bahramali, H., Gordon, E., Rowe, D. L., et al. (2001). Prediction of electroencephalographic spectra from neurophysiology. *Phys. Rev. E* 63:021930. doi: 10.1103/PhysRevE.63.021903
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by backpropagating errors. *Propag. Nat.* 323, 533–536. doi: 10.1038/323533a0
- Sancha-Velasco, A., Uceda-Heras, A., and Garcia-Cabezas, M. A. (2023). Cortical type: a conceptual tool for meaningful biological interpretation of high-throughput gene expression data in the human cerebral cortex. *Front. Neuroanat.* 17:1187280. doi: 10.3389/fnana.2023.1187280
- Sang, I. E. W. F., Schroer, J., Halbhuber, L., Warm, D., Yang, J.-Y., Luhmann, H. J., et al. (2021). Optogenetically controlled activity pattern determines survival rate of developing neocortical neurons. *Int. J. Mol. Sci.* 22:6575. doi: 10.3390/ijms22126575
- Sanides, F. (1962). Architectonics of the human frontal lobe of the brain. With a demonstration of the principles of its formation as a reflection of phylogenetic differentiation of the cerebral cortex. *Monogr. Gesamteig. Neurol. Psychiatr.* 98, 1–201.
- Sanides, F. (1964). The cyto-myeloarchitecture of the human frontal lobe and its relation to phylogenetic differentiation of the cerebral cortex. *J. Hirnforsch* 47, 269–282.
- Sanides, F. (1970). “Functional architecture of motor and sensory cortices in primates in the light of a new concept of neocortex evolution,” in *The Primate Brain: Advances in Primatology*, eds. C. R. Noback and W. Montagna (New York, NY: Appleton-Century-Crofts Educational Division/Meredith Corporation), 137–208.
- Sherk, H., and Stryker, M. P. (1976). Quantitative study of cortical orientation selectivity in visually inexperienced kitten. *J. Neurophysiol.* 39, 63–70. doi: 10.1152/jn.1976.39.1.63
- Shipp, S. (2007). Structure and function of the cerebral cortex. *Curr. Biol.* 17, R443–R449. doi: 10.1016/j.cub.2007.03.044
- Shipp, S., and Friston, K. (2022). “Predictive coding: forward and backward connectivity,” in *The Cerebral Cortex and the Thalamus*, eds. W. Martin Usrey and S. Murray Sherman (Oxford Academic (OUP)), 436–445. doi: 10.1093/med/9780197676158.003.0041
- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* 55, 349–374. doi: 10.1146/annurev.ph.55.030193.002025
- Steffen, P. R., Hedges, D., and Matheson, R. (2022). The brain is adaptive, not triune; how the brain responds to threat, challenge, and change. *Front. Psychiatry* 13:802606. doi: 10.3389/fpsyg.2022.802606

- Sun, Z., and Firestone, C. (2020). The dark room problem. *Trends Cogn. Sci.* 24, 346–348. doi: 10.1016/j.tics.2020.02.006
- The MICrONS Consortium (2025). Functional connectomics spanning multiple areas of mouse visual cortex. *Nature* 640, 435–447. doi: 10.1038/s41586-025-08790-w
- Tschacher, W., and Haken, H. (2007). Intentionality in non-equilibrium systems? The functional aspects of self-organised pattern formation. *New Ideas Psychol.* 25, 1–15. doi: 10.1016/j.newideapsych.2006.09.002
- Tucker, D. M., and Luu, P. (2021). Motive control of unconscious inference: the limbic basis of adaptive Bayes. *Neurosci. Biobehav. Rev.* 128, 328–345. doi: 10.1016/j.neubiorev.2021.05.029
- Tucker, D. M., and Luu, P. (2023). Adaptive control of functional connectivity: dorsal and ventral limbic divisions regulate the dorsal and ventral neocortical networks. *Cereb. Cortex* 33, 7870–7895. doi: 10.1093/cercor/bhad085
- Tucker, D. M., Luu, P., and Friston, K. J. (2025). Adaptive consolidation of active inference: excitatory and inhibitory mechanisms for organizing feedforward and feedback memory systems in sleep. *Cereb. Cortex* 35:bhaf122. doi: 10.1093/cercor/bhaf122
- Tucker, D. M., Luu, P., and Johnson, M. (2022). Neurophysiological mechanisms of implicit and explicit memory in the process of consciousness. *J. Neurophysiol.* 128, 872–891. doi: 10.1152/jn.00328.2022
- Uceda-Heras, A., Aparicio-Rodriguez, G., and Garcia-Cabezas, M. A. (2024). Hyperphosphorylated tau in Alzheimer's disease disseminates along pathways predicted by the Structural model for cortico-cortical connections. *J. Comp. Neurol.* 532:e25623. doi: 10.1002/cne.25623
- Vergara, R. C., Jaramillo-Riveri, S., Luarte, A., Moenne-Loccoz, C., Fuentes, R., Couve, A., et al. (2019). The energy homeostasis principle: neuronal energy regulation drives local network dynamics generating behavior. *Front. Comp. Neurosci.* 10:49. doi: 10.3389/fncom.2019.00049
- Vezoli, J., Magrou, L., Goebel, R., Wang, X.-J., Knoblauch, K., Vinck, M., et al. (2021). Cortical hierarchy, dual counterstream architecture and the importance of top-down generative networks. *Neuroimage* 225, 117479. doi: 10.1016/j.neuroimage.2020.117479
- Wiesel, T. N., and Hubel, D. H. (1974). Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.* 158, 307–318. doi: 10.1002/cne.901580306
- Wright, J. J. (1989). "Linearity and nonlinearity in electrocortical waves and their elementary statistical dynamics," in *Brain Dynamics*, eds. E. Basar and T. H. Bullock (Berlin, Heidelberg: Springer-Verlag), 202–213. doi: 10.1007/978-3-642-74557-7_16
- Wright, J. J., and Bourke, P. D. (2013). On the dynamics of cortical development: synchrony and synaptic self-organization. *Front. Comp. Neurosci.* 7:4. doi: 10.3389/fncom.2013.00004
- Wright, J. J., and Bourke, P. D. (2016). Further work on the shaping of cortical development and function by synchrony and metabolic competition. *Front. Comp. Neurosci.* 10:127. doi: 10.3389/fncom.2016.00127
- Wright, J. J., and Bourke, P. D. (2022). Unification of free energy minimization, spatio-temporal energy, and dimension reduction models of V1 organization: postnatal learning on an antenatal scaffold. *Front. Comput. Neurosci.* 16:869268. doi: 10.3389/fncom.2022.869268
- Wright, J. J., and Bourke, P. D. (2023). The mesoanatomy of the cortex, minimization of free energy, and generative cognition. *Front. Comput. Neurosci.* 12:1169772. doi: 10.3389/fncom.2023.1169772
- Wright, J. J., and Bourke, P. D. (2024a). Markov blankets and mirror symmetries. Free energy minimization and mesocortical anatomy. *Entropy* 26:287. doi: 10.3390/e26040287
- Wright, J. J., and Bourke, P. D. (2024b). Cortical development in the structural model and free energy minimization. *Cereb. Cortex* 34:bhae416. doi: 10.1093/cercor/bhae416
- Wright, J. J., Bourke, P. D., and Chapman, C. L. (2000). Synchronous oscillation in the cerebral cortex and object coherence: simulation of basic electrophysiological findings. *Biol. Cybern.* 83, 341–353. doi: 10.1007/s004220000155
- Wright, J. J., Bourke, P. D., and Favorov, O. V. (2014). Mobius-strip-like columnar functional connections are revealed in somato-sensory receptive field centroids. *Front. Neuroanat.* 8:119. doi: 10.3389/fnana.2014.00119
- Wright, J. J., and Liley, D. T. J. (1995). Simulation of electrocortical waves. *Biol. Cybern.* 72, 347–356. doi: 10.1007/BF00202790
- Wright, J. J., and Liley, D. T. J. (1996). Dynamics of the brain at global and microscopic scales: neural networks and the EEG. *Behav. Brain Sci.* 19, 285–295. doi: 10.1017/S0140525X00042679
- Xie, X., and Seung, H. S. (2003). Equivalence of backpropagation and contrastive Hebbian learning in a layered network. *Neural Comput.* 15, 441–454. doi: 10.1162/089976603762552988
- Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011
- Yufik, Y. M. (2019). The understanding capacity and information dynamics in the human brain. *Entropy* 21:308. doi: 10.3390/e21030308
- Yufik, Y. M., and Friston, K. (2016). Life and understanding: the origins of "understanding" in self-organizing nervous systems. *Front. Syst. Neurosci.* 10:98. doi: 10.3389/fnsys.2016.00098