Review article

Metastability demystified – the foundational past, the pragmatic present and the promising future

Fran Hancock [1,24] ✉, Fernando E. Rosas[2,3,4,5,6,7,24] ✉, Andrea I. Luppi [5,8,9], Mengsen Zhang[10], Pedro A. M. Mediano [11,12], Joana Cabral [5,13], Gustavo Deco [14,15,16,17], Morten L. Kringelbach [5,18], Michael Breakspear[19], J. A. Scott Kelso [20,21,22] & Federico E. Turkheimer[1,23]

Abstract

Sections

Healthy brain function depends on balancing stable integration between brain areas for effective coordinated functioning, with coexisting segregation that allows subsystems to express their functional specialization. Metastability, a concept from the dynamical systems literature, has been proposed as a key signature that characterizes this balance. Building on this principle, the neuroscience literature has leveraged the phenomenon of metastability to investigate various aspects of brain function in health and disease. However, this body of work often uses the notion of metastability heuristically, and sometimes inaccurately, making it difficult to navigate the vast literature, interpret findings and foster further development of theoretical and experimental methodologies. Here, we provide a comprehensive review of metastability and its applications in neuroscience, covering its scientific and historical foundations and the practical measures used to assess it in empirical data. We also provide a critical analysis of recent theoretical developments, clarifying common misconceptions and paving the road for future developments.

A full list of affiliations appears at the end of the paper. ✉e-mail: fran.hancock@kcl.ac.uk; f.rosas@sussex.ac.uk

Review article

Introduction

The notion of metastability is used in neuroscience to characterize two coexisting tendencies in neuronal populations: at any point in time, there is a tendency for some brain regions to work collectively by coordinating their activity while other brain regions remain autonomous to allow the performance of specialized functions. Thus, metastability is thought to reflect a delicate mixture of cooperation and relative independence between brain areas or neuronal populations in response to inputs from the environment. As such, metastability is considered a fundamental feature of brain function[1,2].

Although metastability has been increasingly featured in the recent neuroscience literature, its empirical and theoretical origins are sometimes lost in the process, thus depriving the wider neuroscience community of an understanding of its origins and meaning[3,4]. Indeed, the plethora of studies of metastability have involved heterogeneous methodologies, and it can be challenging for neuroscientists to build a fundamental understanding about what metastability is and how different applications relate to each other. Moreover, researchers unfamiliar with this area of research could easily misinterpret applications of different signatures of metastability as indicators of conflicting definitions[5], thereby hindering the use of metastability as a lens into brain function.

The aim of this Review is to provide a comprehensive account of metastability in neuroscience while illustrating common misunderstandings and offering guidance for future developments. We first provide an introductory account of the origin of metastability in diverse literatures before exploring several signatures of metastability that have been measured in both empirical and computational modelling studies. Next, we examine a number of 'dynamical routes' to metastability (that is, different types of dynamical scenario that can give rise to metastability) that have originated from various theoretical models[6–10], and we review empirical studies that propose neural correlates of metastability or describe its relevance for healthy and pathological brain function. Finally, we suggest several future avenues for theoretical and empirical studies of metastability. In summary, this Review aims to provide a unified account of metastability in neuroscience covering both its technical and practical dimensions, complementing recent efforts focusing on metastability from the micro-level circuits viewpoint[11,12] or from a purely dynamical systems perspective[13].

The past: origins of metastability
What is metastability?

When studying the brain, it is helpful to distinguish between the following: (a) the physical system of interest, be it the entire brain or a circuit of neurons, (b) the system's attractor landscape, a theoretical construct that describes the rules governing the dynamics of that system (Box 1), and (c) observable spatiotemporal activity, such as brain signals captured with electroencephalography (EEG), functional MRI (fMRI) or electrophysiology (including local field potentials and neuronal spikes), reflecting features of the system at different spatial and temporal scales. These three aspects of the brain are closely related: the shape of the attractor landscape depends on the organization of the physical system, such as its connectivity and the properties of individual elements (that is, regions and neurons), and the observed activity is generated by the physical system as it moves across the landscape (Fig. 1). Metastability, essentially, is a way of describing attractor

---

Box 1 | A dynamical systems theory primer

Dynamical systems theory is a branch of mathematics that studies how the state of a system evolves over time, using either discrete-time difference equations or continuous-time differential equations. These equations can be solved either analytically (pencil and paper), geometrically (shapes) or numerically (approximations using a computer) to find out how a system transits between its states. Solutions of the equations specify the stability of the system (Fig. 2a). A state is a configuration of the variables of the system at a particular point in time. The phase space of a dynamical system is the set of all possible states and, hence, contains all the allowed combinations of values of the variables of the system. A trajectory of the system is a path of the dynamical system through a succession of states.

Attractors, repellers and saddles

When trajectories from different initial conditions converge to a set of states, that set is called an attractor (Fig. 2b). An attractor is surrounded by its basin of attraction, which is all the points in phase space that flow onto the attractor. When many trajectories migrate away from a set of states, that set is called a repeller (Fig. 2b). Some dynamical systems can be illustrated as an attractor landscape, wherein balls illustrate the system transitioning between the peaks (repellers) and the valleys (attractors) (Fig. 2c). In some situations, noise can kick the system between stable attractors (Fig. 2d). When trajectories are attracted to a set of states in one direction but migrate away from the set in another direction, that

set is called a saddle (Fig. 2e). Saddles can be joined together into cycles of saddles (Fig. 2f). Attractors can be a single point, a line or a cycle, or they can have more complex geometries (usually called a strange or chaotic attractor). A chaotic attractor exhibits dynamics that are highly sensitive to their initial conditions. The trajectories within a chaotic attractor have rich and complex geometries (for example, Fig. 2g shows two synchronized chaotic attractors). Finally, a Milnor attractor is often observed in chaotic systems and can possess repelling trajectories, allowing the system to escape to a non-contiguous attractor, like a 'wormhole' (Fig. 2h and Supplementary information S2). Thus, the Milnor attractors of chaotic systems exhibit saddle-like behaviour.

The shape of the attractor landscape is a function of control parameters, which modify the attractor landscape and can cause attractors (or repellers) to appear, disappear or be altered in their stability (Fig. 2i). This happens either at a bifurcation or at a crisis (when an attractor hits its boundary following parameter tuning). In contrast to control parameters, which have influence on the system, an order parameter is a collective observable property of the system, such as synchrony between the elements. An order parameter is a feature of the system (or a subsystem) and not of an individual component, and it acts as a 'readout' of the system as a whole.

Because states in neuronal systems are not easy to define[165], dynamic instabilities, which refer to behavioural changes of the system in the vicinity of a bifurcation, are studied to identify relevant order parameters[166].

---

a The physical system   **b** The attractor landscape   **c** Data patterns

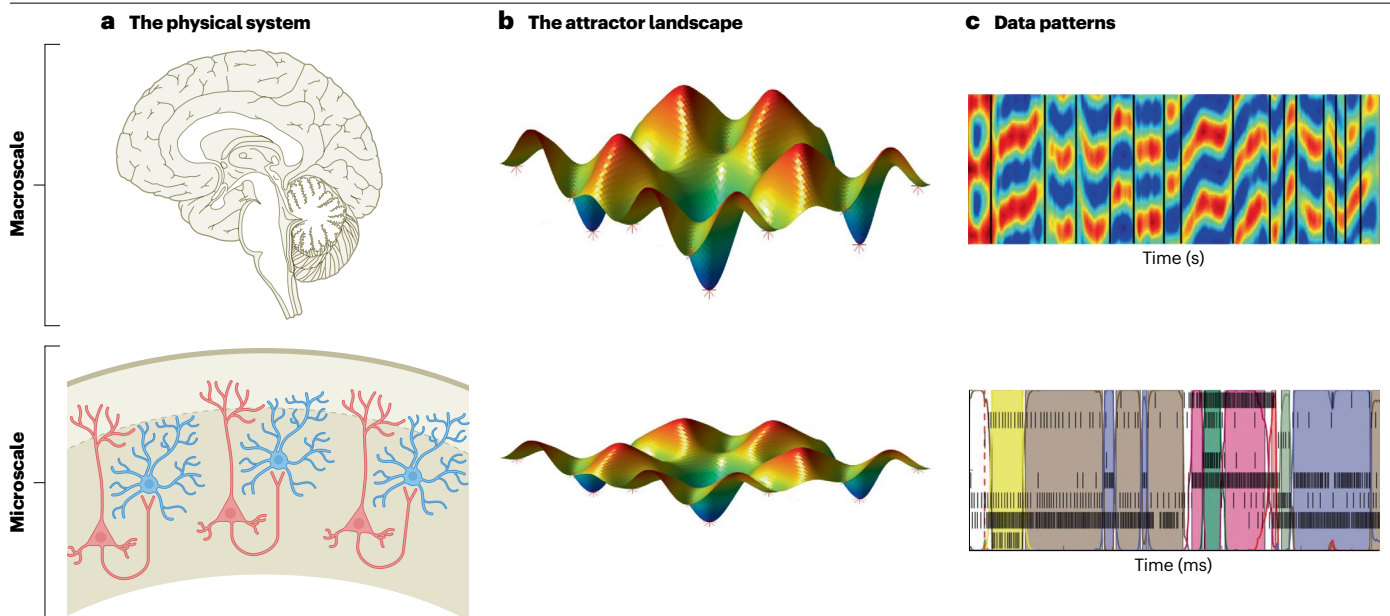Macroscale

Microscale

Time (s)

Time (ms)

Fig. 1 | From physical substrate to attractor landscapes and spatiotemporal patterns in data. Macroscale (top row) and microscale (bottom row) representations of the three aspects that need to be distinguished when discussing metastability in neuroscience. **a**, The first aspect is the physical system of interest, be it the entire brain (top) or a circuit of neurons (bottom). **b**, The second aspect is the organization of this system that determines the attractor landscape, a theoretical construct that describes the rules governing the dynamics of that system. **c**, The third aspect is the observable spatiotemporal activity (for example, the regional blood oxygen level-dependent signals from functional MRI at the macroscale or neuronal spike trains at the microscale) generated by the physical system according to the attractor landscape, as then viewed through a noisy observation. At both the microscale and macroscale, it is possible to identify features of metastable dynamics such as the presence of transient states. Part **b** is reprinted from ref. 160, CC BY 3.0. Top image in part **c** is adapted from ref. 105, Springer Nature Limited. Bottom image in part **c** is adapted from ref. 127, CC BY 4.0.

landscapes that exhibit specific features, that is, it refers to how the dynamics are governed.

The characterizing features of metastable systems can be described by comparing them to other types of stability described by dynamical systems theory (Box 1). The simplest form of stability is mon-ostability (Fig. 2), which corresponds to when the system has a single attractor (that is, when there is one stable solution to the differential equation describing its dynamics). A slightly richer type of dynamics can be observed in systems with two attractors that exhibit bistability (Fig. 2): the system will settle into one of the attractors depending on the initial conditions, but sufficiently large noise fluctuations can force the system to switch from one attractor to another via a sequence of states within the phase space (known as a trajectory) (Fig. 2). More generally, a system is said to exhibit multi-stability when multiple attractors in the dynamics exist[14] (for more technical discussions on multi-stability, see refs. 15–17). Finally, a system is said to exhibit metastability when it exhibits 'unstable attraction', that is, its attractor landscape includes regions that attract in some directions but repel in others (called 'saddles'; Box 1), pulling the system towards the region then pushing it away. Hence, a metastable system approaches successive saddle-like regions, dwells near each for some time, and spontaneously escapes to visit another. In contrast to multi-stable scenarios with attractors that pull and trap trajectories, metastability works more like a children's slide: when the slide is nearly flat, the dynamics slow down, giving the impression of nearly stopping ('attraction'), then set the system free to continue its trajectory. Therefore, while in multi-stable systems, a trajectory can only escape from an attractor owing to noise, and trajectories in metastable systems transition between repelling and attracting regions 'by design' (Box 2). As such, metastability is more flexible than multi-stability but more structured than randomness[1].

A consequence of the previous argument is that metastability refers to a specific type of dynamics that may take place in systems that exhibit patterns that recur either in repeatable sequences[9,18] or flexible alternation[10,19]. Thus, metastability corresponds to a specific type of dynamic behaviour, that is, something certain systems do, and not to a single specific mechanism, that is, how these systems do it[20]. In fact, metastability can be realized by several different mechanisms (as discussed in the section 'The present: dynamical mechanisms and drivers of metastability').
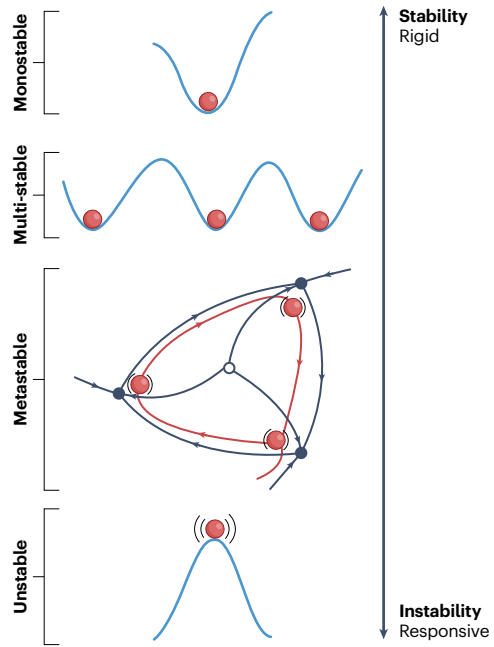
A brief history of metastability

At this point, it is helpful to understand how different fields independently came to develop analogous notions of metastability, in turn influencing different perspectives of neuroscientists on metastability and its application to study brain function. Below, we provide an outline of how metastability has been studied in thermodynamics[21], human motor coordination[22–24], mathematics[25,26], and complex systems and chaos[27,28].
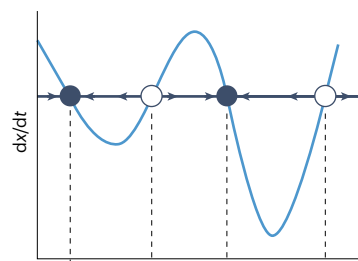
Arguably, the first report of a metastable phenomenon can be found at the turn of the nineteenth century, when a chemical system (a supersaturated solution of sodium nitrate) was reported to be in a 'metastable condition' such that its transition time to a stable state exceeded the relaxation time[21], that is, the system took longer
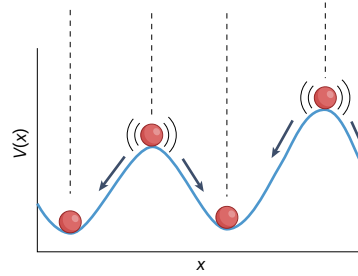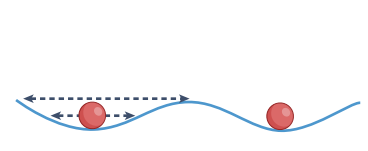
a Types of stability

Monostable

Multi-stable
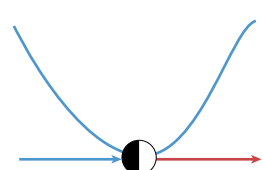
Metastable

Unstable

Stability
Rigid

Instability
Responsive

b Trajectory map

dx/dt

c Attractor landscape

V(x)

x

d Multi-stable fluctuations

e Saddle node

f Cycle of saddles
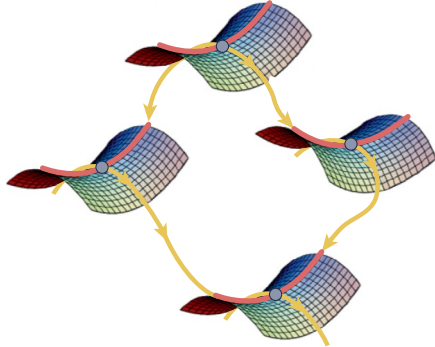
g Synchronized chaotic attractors

Z

W

V
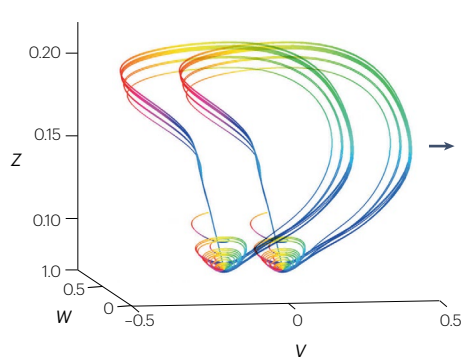
h Milnor attractors ('wormholes')
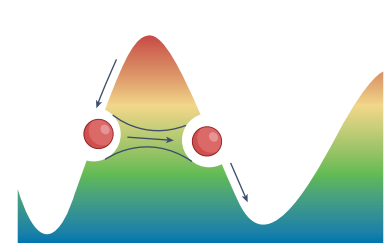
i Fixed-point memory
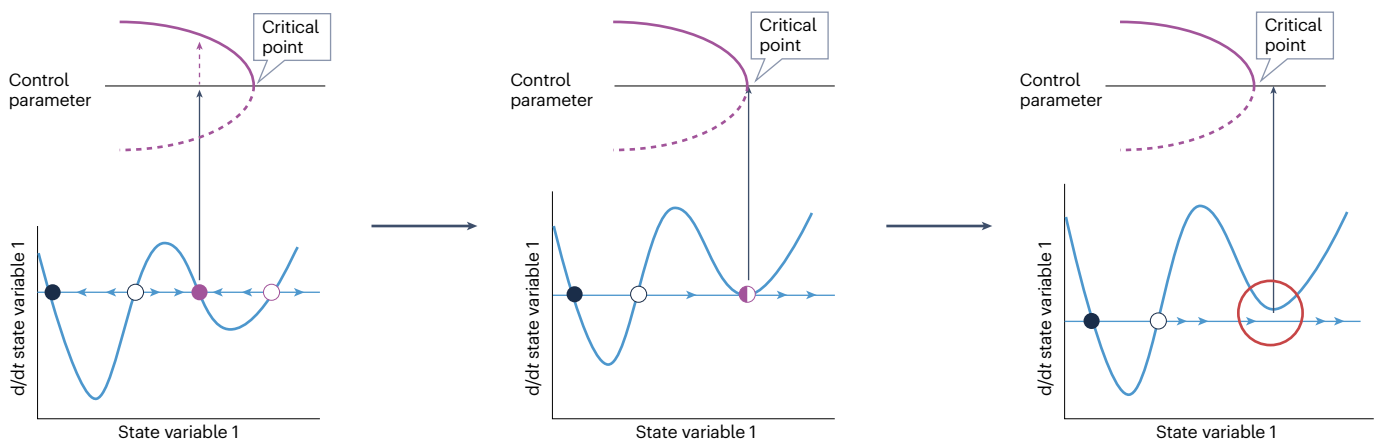
Control
parameter

Critical
point

d/dt state variable 1

State variable 1

Control
parameter

Critical
point

d/dt state variable 1

State variable 1

Control
parameter

Critical
point

d/dt state variable 1

State variable 1

Review article

b, This part shows a trajectory map of a one-dimensional dynamical system. The horizontal axis presents a one-dimensional phase space (x). The horizontal dark blue arrows indicate the direction of the flow in the phase space determined by the differential equation (dx/dt), which is plotted in light blue as a function of x. Where dx/dt is above or below the horizontal line, the trajectories in the state space flow to the right or left corresponding to a positive or negative change of the state variable x. Where the function intersects with the horizontal axis, there is no change to the state, making that value of x a fixed point. The slope of the function at the fixed point indicates its stability: a negative slope implies an attractor, whereas a positive slope indicates a repeller. By convention, attractors are represented by a filled circle, whereas repellers are shown as empty circles. **c**, This part illustrates the attractor landscape of the same system. The balls illustrate the system transitioning between valleys (attractors) and peaks (repellers). **d**, In systems with two or more attractors, noise can allow the system to spontaneously switch between stable attractors. **e**, In a one-dimensional system, the saddle is represented as a dual-coloured circle. The blue line shows the flow towards the saddle fixed point and the red line shows the flow away from the saddle. **f**, Trajectories sometimes link saddles when they are repelled from one saddle and attracted towards another. This part illustrates a cycle of four saddles. **g**, This part shows synchronized chaotic attractors[161]. The dynamic variables are the mean pyramidal cell membrane potential V, the

mean inhibitory cell membrane potential Z and the mean proportion of open potassium channels W. **h**, Synchronized chaotic attractors can give rise to Milnor attractors with riddled basins (see Supplementary information S2). A small number of trajectories are attracted to this special attractor. However, because its basin of attraction is full of holes that connect to other regions of the phase space, some trajectories flow out to a non-contiguous attractor basin. **i**, This part shows how multi-stable dynamics change to metastable dynamics as a control parameter varies owing to fixed-point memory. The top row is the bifurcation diagram and the bottom row is the trajectory map. The left panel shows two stable attracting fixed points (black and purple filled circles) and two unstable repellers (open circles) in the multi-stable regime when the control parameter of the model is below the critical point. The middle panel shows that as the control parameter varies and reaches the critical point (top), the trajectory curve shifts vertically upwards (bottom), which brings a stable and an unstable fixed point together, creating a saddle node. The right panels shows that if the control parameter goes just beyond the critical point (top), the saddle is annihilated, leaving the memory of the fixed point in which the saddle once existed. Although the attractor no longer exists, trajectories are still attracted to this memory wherein they slow down before eventually escaping (a so-called ghost attractor). Part **f** is adapted with permission from ref. [162], SPIE. Part **g** is adapted with permission from ref. [161], Elsevier.

than expected to reach its stable state, so that its period of instability was paradoxically more persistent than expected (hence the name). A more recent example of such long transient behaviour is found in supercooled water, which remains liquid after being cooled below 0 °C before abruptly freezing[29]. This phenomenon inspired neuroscientific investigations using neural network models in the 1990s[30–32]. However, what is now studied as metastability in neuroscience does not follow this form (and, hence, these approaches will not be discussed further).

The origins of metastability as currently studied in neuroscience can be traced to a behaviour first observed and studied in *Labrus* fish: these fish coordinate their swimming with their fins 'almost-in-sync' but intermittently lose this synchronization[33]. This type of relative coordination was later systematically investigated and formalized in the study of the collective dynamics of human coordination[23,34]. For example, when performing experiments in which human participants flexed their index finger against a metronome, it was observed that as the metronome frequency (a control parameter set by the experimenter) was increased, participants progressively switched from off-beat to on-beat flexing, became almost-in-sync with the metronome, and finally lost synchronization between the finger movements and the metronome[34,35]. The relative coordination of the almost-in-sync condition was described as metastability in dynamical systems terms[1] and formalized into a mathematical model[22,34,36] (Supplementary information S1). In addition, the cyclic behaviour of being almost-in-sync, to escaping through desynchronization and to returning to being almost-in-sync, was described as metastable, as no stable states existed, just a succession of tendencies for integration (between movement and metronome to maintain synchronization) and segregation (loss of rhythm between movement and metronome)[1,3,34,35]. This cycling phenomenon was quickly recognized as potentially relevant for studying the interplay of integration and segregation in brain dynamics[7,37,38].

Around the same time, studies in computational brain modelling using chaotic systems have used the concept of itinerant dynamics to explain patterns of transiently stable brain activity[7,10,39] (see Supplementary information S2 for an account of chaotic systems and chaotic itinerancy[40]). This literature developed the characterization of

metastable behaviour to include cluster formation[41,42] and sequential switching between multiple metastable states[9,10,43]. Complementing this line of work, studies on the dynamics of percolation[26], in which nodes are added to a network, have also led to useful applications to study brain dynamics[44–48].

In summary, different scientific fields arrived at similar notions of 'metastable behaviour', which set the stage for how metastability is investigated today in neuroscience. We next discuss how metastability has been used to investigate brain function over the past 20 years.

The present: signatures of metastability and empirical applications
Signatures of metastability

Studying metastability in brain data — whether from electrophysiological (for example, individual spiking or bursting neurons) or neuroimaging (for example, fMRI) recordings — presents a formidable methodological challenge. In effect, a thorough evaluation of metastability would require a full reconstruction of the attractor landscape of the neural dynamics, which would often include a large number of attractors (Supplementary information S1). This difficulty has motivated the search for heuristic signatures of metastability that are suitable for efficient estimation from current electrophysiological and neuroimaging data. This section reviews several heuristic markers of metastability that have been introduced in the neuroscience literature that allow tractable computation and have led to neuroscientific insights. Their key conceptual features are illustrated in Fig. 3, and we have made a MATLAB[49] code library available to calculate these signatures from empirical data (see Code availability statement). Discussion about additional markers of metastability and associated equations can be found in Supplementary information S3.

It is important to note that these signatures do not disambiguate metastable or multi-stable dynamics, as will become clear when describing each signature. Thus, these signatures are not meant to be used as tests for deciding whether a system is metastable or not, but only to assess the degree to which a given system displays specific necessary, but not sufficient signatures for metastability. In practice, it

---

Box 2 | Common misconceptions about metastability

There are a number of misconceptions relating to metastability. The first misconception is that observing switching between distinct states with long periods of stability is enough to infer metastability. However, this is not strictly correct: within dynamical systems theory, this is a necessary but not sufficient condition for metastability. Indeed, this feature is shared by both metastable and multi-stable systems. For metastability, there must be coexisting tendencies for attraction and repulsion, that is, there must be unstable (saddle-like) features in the attractor landscape[40,50].

The second misconception is that signatures of metastability are alternative definitions of this concept. In fact, the different signatures of metastability (as discussed in the section 'The present: signatures of metastability and empirical applications') are not alternative definitions but represent different ways to assess the dynamics associated with the phenomenon of transient states, irrespective of their dynamical origins. The signatures are heuristic, being particularly useful when a full analysis of the dynamical system at hand (that is, reconstruction of its attractor landscape) is not

feasible, often because the required volume or granularity of data is not available.

The third misconception is that metastability is in the data. Actually, metastability is a property of a dynamical system (that is, a mathematical or theoretical model aimed at capturing empirically observed phenomena) and not of data. More specifically, it is a way of characterizing certain attractor landscapes. Although no attractors (only saddles) exist in a metastable system, and so more correctly, we could refer to a dynamical landscape, we retain the standard nomenclature from dynamical systems theory of attractor landscapes.

Another misconception is that noise is necessary for metastability. Although noise is necessary for switching within multi-stable systems, metastable cycling can occur in the absence of noise. In dynamical systems, metastability can arise from time delays and heterogeneities in network connectivity that create asymmetries. The addition of noise to these models can affect the stability of features of the dynamical system, such as duration of[143], or switching between, transient states[101], but does not negate their metastability.

---

is common to contrast the values of these markers obtained in groups of individuals (for example, patients and control individuals) or experimental conditions or manipulations, and assess whether a specific signature of metastability has increased or decreased.

Temporal variability of the Kuramoto order parameter. Metastability in the brain has been claimed to reflect coexisting tendencies for integration and segregation[40,50]. When conceptualizing oscillatory dynamics as supporting cortical processing, integration can be interpreted as synchronization (as discussed, for example, in 'communication through coherence'[51,52] and other similar frameworks[53], but note also recent critiques of such approaches[54,55]), and segregation would involve a loss of synchrony with regions pursuing specialist functionality. Building on these intuitions, Shanahan proposed the temporal variance of the global synchronization, technically known as the Kuramoto order parameter[56] (KOP), as a signature of metastability[57]. A small variance, as measured by the standard deviation of the KOP (std-KOP), implies that the degree of synchronization is stable over time, whereas if the variance is high, then the degree of synchronization

between the nodes must be constantly changing (see Fig. 3a and Supplementary information S4 for more technical details and Video 1 for an animation of the dynamics). Note, however, that a high variance of the KOP could reflect either metastable or multi-stable dynamics, or it could even reflect mere random fluctuations within the system (see the section 'The future: open questions and opportunities'). This signature has proven to be extremely versatile, leading to numerous applications spanning both empirical studies and computational models[58–69], which are reviewed in later sections.

Temporal variability of dynamic functional connectivity. Time-varying functional connectivity among brain regions can be captured with statistical relationships such as Pearson correlation or phase difference[70] extracted from whole-brain images recorded over time. The resulting patterns of correlation or phase-difference relationships between the regions in the time-varying functional connectivity matrices represent the repertoire of configurations spontaneously visited by a system over time. The first eigenvalue of the instantaneous functional connectivity matrix, known as the spectral radius, captures the

---

Fig. 3 | Practical signatures of metastability. Metastability has been associated with simultaneous tendencies for coupling and decoupling of synchronized regions and for integration and segregation of disparate regions. Signatures of metastability in empirical data are grounded in these fundamental associations. The signatures are estimated from instantaneous phase, amplitude correlations or relative phase differences. **a**, The most common signature of metastability is the variability over time of the Kuramoto order parameter (KOP), which tracks the overall synchrony in a system of oscillators, wherein the variability is typically measured as the standard deviation of the KOP (std-KOP). **b**, A signature that can be estimated directly from the functional MRI (fMRI) data or from Hilbert-transformed fMRI data is the variability measured as the standard deviation over time of the spectral radius (largest eigenvalue obtained with eigenvector decomposition) of time-resolved functional connectivity matrix (std-SPECT) (note that many ways of obtaining time-resolved functional connectivity exist). **c**, Using the Hilbert transform to extract instantaneous phase from fMRI data, the variability of the intrinsic ignition over time, which is based on relative

phase differences, is another signature. Intrinsic-driven ignition is obtained by identifying 'driver events' (unusually high spontaneous activity in fMRI[72,163]) and measuring the magnitude of the concomitant activity occurring in the rest of the brain within a short time window (in fMRI studies, it is common to use the duration of the haemodynamic response function). The variability is measured as the standard deviation of intrinsic ignition (std-IGNITE). **d**, Again, using Hilbert-transformed fMRI data, estimates of relative phase differences between regions segregate the regions into those that are aligned with the overall phase alignment of the brain and those that are aligned in antiphase. Over time, the contribution of each region to this antiphase alignment varies, and the mean temporal variance of this contribution (mean-VAR) is another signature of metastability. The contribution variance is calculated from the time-varying leading eigenvector (the eigenvector with the largest eigenvalue) of the phase-aligned connectivity matrix. TR, repetition time. Part **a** is adapted from ref. 61, CC BY 4.0. Part **c** is adapted from ref. 73, CC BY 4.0.

effective dimensionality of the matrix. A large value indicates that the configuration at that time has only a few degrees of freedom (that is, elements behave similarly), whereas a small value suggests many degrees of freedom (that is, elements are free to behave relatively independently). Hence, variability (for example, the standard deviation) of the spectral radius over time (std-SPECT) captures variability between integration (large spectral radius, reflecting interdependence) and segregation (small spectral radius, reflecting independence) and, thus, has been proposed as a signature of metastability[71] (Fig. 3b). This signature has exhibited empirical usability, for example, in identifying an association between reduced metastability and ageing in fMRI recordings from a longitudinal study in rats[71].

Temporal variability of intrinsic ignition. The spontaneous activation of one brain region or neuronal ensemble (as captured by a moment of unusually high activity) often propagates to other regions or ensemble, respectively, 'igniting' similar high-activity events elsewhere within a short time window. The extent of propagation of such ignition events may be seen as reflecting the level of neural integration. Accordingly, the temporal variability (for example, the standard deviation) of the breadth of intrinsic ignition (std-IGNITE) (that is, the number of other regions or ensembles that are ignited) reflects the variability in integration, which can be seen as part of the integration–segregation tendency that is intuitively associated with metastability[72] (Fig. 3c). This signature has been successfully applied to fMRI data to characterize

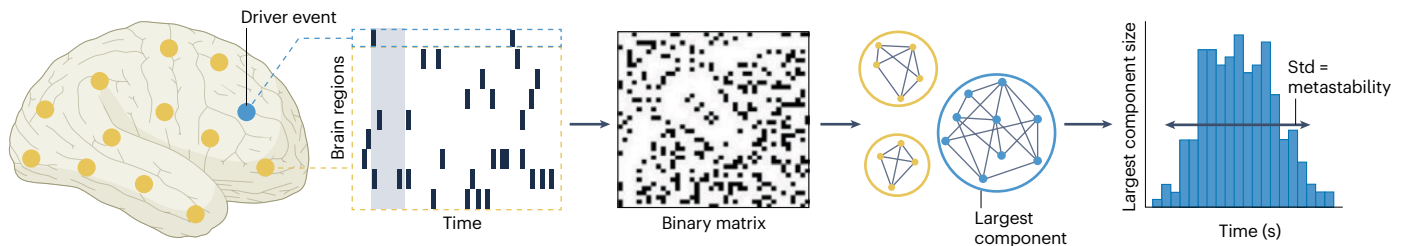**a** **Standard deviation of the Kuramoto order parameter (std-KOP)**
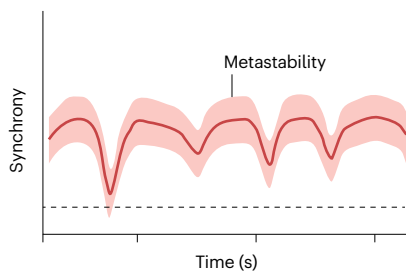
b **Standard deviation of the spectral radius (std-SPECT)**

Time-varying functional connectivity matrices

Metastability = variance $\lambda_1(t)$
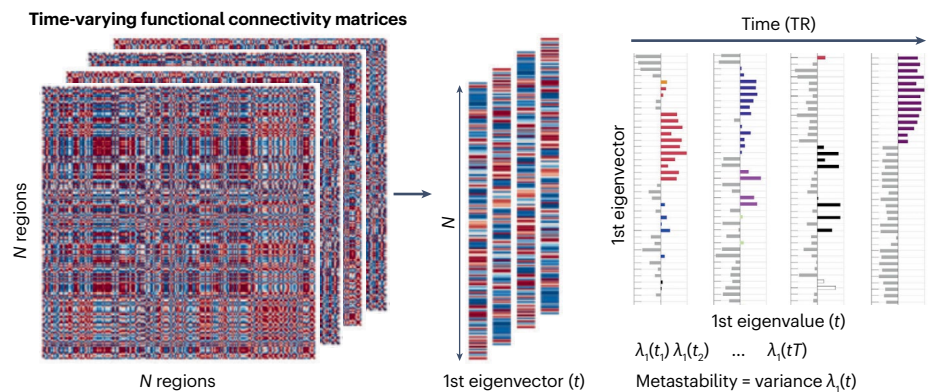
c **Standard deviation of intrinsic ignition (std-IGNITE)**

d **Mean variance of the learning eigenvector (mean-VAR)**

Metastability = mean $<V_n .. V_m>$
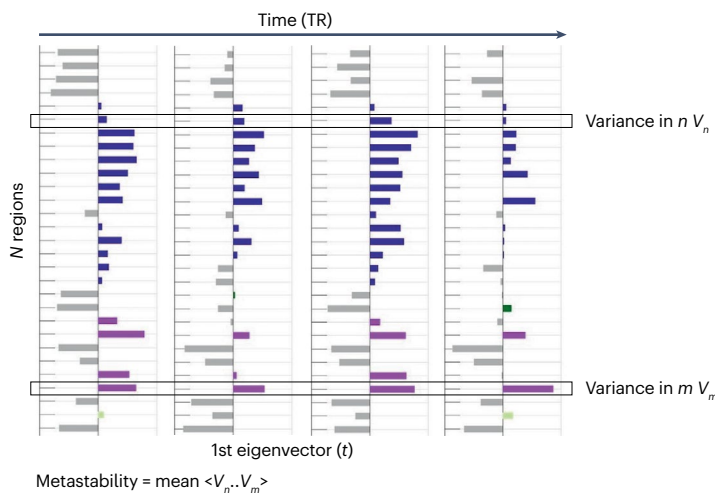
the dynamical complexity underlying disorders of consciousness[73], ageing[74] and meditation[75].

Temporal variability of phase alignment. Several studies have investigated the relationship between time-varying functional connectivity and signatures of metastability[61,66,76], wherein the former is assessed in terms of phase alignment and the latter is based on overall synchrony. However, it is not straightforward to conceptually align these two approaches, as they use different types of phase relationship. To overcome this methodological limitation, a new signature based on relative phase was proposed[77] that replaces overall synchrony (which averages phase overall regions[57]) with overall phase alignment (which considers whether each individual region is in phase or in antiphase with respect to the overall phase direction) (Fig. 3d). The contribution of individual regions to phase alignment is captured in the leading eigenvector (or first principal component) of the functional connectivity matrix[78], and thus the temporal variance of its entries (and not the eigenvalues as in std-SPECT) captures the flexibility of each brain region to modify its phase alignment and, hence, their participation in transient coalitions with other regions. The mean of this temporal variance (mean-VAR) overall brain regions (or a subset of brain regions) has been proposed as a signature of metastability, reflecting the balance between global integration and functional segregation (see Supplementary information S5). This signature of metastability was higher in people with early psychosis and individuals with chronic schizophrenia than in healthy control individuals, based on fMRI recordings[77].

Empirical studies of metastability
The signatures of metastability at the macroscopic scale described above focus on variability in the dynamics associated with transient states. The interpretation of the transient states themselves – and their relationship with anatomy, cognition, individual differences, pathology and even consciousness – has been the subject of numerous investigations through a variety of approaches for studying time-varying functional connectivity in fMRI and EEG in humans and other species[79–90] (see ref. 91 for an authoritative review). Below, we review how the signatures of metastability above have been used to identify differences in behaviour and cognition across conditions and groups on (macroscopic) neuroimaging data. For an overview of the complementary insights provided by electrophysiological recordings at the level of neurons and local field potentials, see Supplementary information S8.

Currently, normative values are not available for the different signatures, and so the most common approach is to compare different groups or conditions against a reference, such as healthy controls. Moreover, these signatures alone do not confirm metastability in the underlying dynamics but rather suggest links between the dynamics underlying the transient states and brain function across a spectrum of experimental and neurodevelopmental conditions.

Studies in humans[59] and rats[71] indicate that reductions in metastability signatures generally occur with ageing, when older groups are compared with middle-aged people (std-IGNITE) and when time-point signatures are compared across the lifespan of rats (std-SPECT). Signatures of metastability seem to be reliable across multiple scans in healthy young adults[61] (std-KOP). Interestingly, estimates of metastability in babies born at full-term were higher than in premature-born babies[92] (std-IGNITE). These results suggest that metastability may increase during in-utero development and then decrease in healthy ageing.

Signatures of metastability have been applied to study neurodegenerative and neuropsychiatric illnesses at the macroscopic scale.

For example, in a study of neurodegeneration in patients with Alzheimer disease (AD), the magnitude of a signature of metastability (std-KOP) was lower in these individuals than in healthy control individuals[59]. Additionally, a significant association was found between global cognitive performance and the signature of metastability both in individuals with AD and in people with mild cognitive impairment (MCI). Congruent with these findings, another empirical study has found gradual decreases in std-KOP from people with (healthier) subjective cognitive impairment (SCI) to people with partially affected (MCI) and more severely affected cognitive states (AD)[93]. The signature of metastability that unveiled these graded differences reflects the variability in the global synchronization across the entire brain. These results suggest that decreases in the magnitude of signatures of metastability across ageing may be associated with cognitive impairment or vice versa. However, there is evidence that reductions in signatures of metastability can also occur owing to perturbations, whether pharmacologically by anaesthesia[94] or physically by traumatic brain injury[62].

The previous results should not lead to the conclusion that any increase in metastability at the macroscale is associated with beneficial effects. Indeed, a study on people with a diagnosis of early psychosis or chronic schizophrenia[77] has revealed that mean-VAR was higher in these individuals than in healthy control individuals. Moreover, another study has reported a modular increase in metastability (std-KOP) in people with schizophrenia compared with healthy individuals[64].

Together, these results (although coming from different signatures) may be thought to suggest that both increases and decreases in signatures of metastability can be associated with different cognitive and neuropsychiatric disorders as outlined in Table 1. In other words, there may be a 'sweet spot' for metastability in healthy brain functioning[2,95].

The present: dynamical mechanisms and drivers of metastability
The many routes to metastability and transient states
In contrast to empirical studies, computational models are fully available for inspection and manipulation by the researcher, allowing relationships between system behaviour and model parameterization to be probed to address questions about causality[96]. This is particularly relevant for identifying the potential neural underpinnings of metastability, which could be amenable to treatment in disorders in which metastability is disrupted. Below, we review a selection of computational studies highlighting the different dynamical mechanisms that can give rise to metastability. Additionally, to address ambiguity in the literature, we also appraise studies in which transient states and complex spatiotemporal patterns arise from noise-driven multi-stability. Note that Fig. 4 provides a summary of the reviewed dynamical routes to metastability, with more technical information about these routes provided in Supplementary information S1, S2, S4 and S6. The seminal models of metastable coordination dynamics in behavioural studies are reviewed in Supplementary information S1.

A key outcome from the modelling work is the identification of symmetry breaking as a key enabler of metastability. Symmetry breaking is a process in which a model parameter is varied so that asymmetry suddenly arises in the system. In dynamical systems terms, symmetry refers to the extent to which the dynamic of the system is invariant of permutation of its components, and hence asymmetry happens when this invariance is broken[97]. Symmetry breaking can arise from systematic variations of different model parameters, including the connectivity between regions, the local properties of

Review article

Table 1 | Investigations of metastability in functional MRI empirical studies

Study group[a]	Signature of metastability	Findings	Ref.
Healthy young adults	std-KOP	Global metastability was reliable across multiple scanning sessions	61
Individuals with AD or MCI	std-KOP	Global metastability was reduced in people with AD or MCI compared with controls, but it was higher in people with AD than in those with MCI with concomitant reductions in cognitive performance	59
Individuals with AD, MCI or SCI	std-KOP	Global metastability was highest in SCI and progressively reduced in MCI and AD	93
Individuals with schizophrenia	std-KOP	Increased metastability in the salience network was associated with the negative symptoms of schizophrenia	64
People with early psychosis or schizophrenia	mean-VAR	Global and local metastability was higher in people with schizophrenia or early psychosis than in healthy controls	77
Healthily ageing rats	std-SPECT	Metastability decreased progressively with ageing	71
Healthily ageing individuals	std-IGNITE	Global metastability was lower in the older group than in the middle-aged group	74
Premature babies	std-IGNITE	Preterm neonates exhibited reduced metastability in the default mode, dorsal attention and the salience networks compared with neonates born at full-term	92
Anaesthetized macaques	std-KOP	The awake condition had higher levels of global metastability than the anaesthetized state, with different anaesthetics reducing global metastability by different degrees	94
Individuals with TBI	std-KOP	Global metastability was reduced in people with TBI compared with healthy control individuals, with the reductions in TBI being found in the salience network, a left-fronto-parietal network and a dorsal attention network	62
People practising meditation	std-IGNITE	Experienced meditators showed higher metastability during resting state than inexperience meditators	75

AD, Alzheimer disease; MCI, mild cognitive impairment; mean-VAR, mean variance of the leading eigenvector; SCI, subjective cognitive impairment; std-KOP, standard deviation of the Kuramoto order parameter; std-SPECT, standard deviation of the spectral radius; std-IGNITE, standard deviation of intrinsic ignition; TBI, traumatic brain injury. [a]Unless stated otherwise, the studies involved humans.

the individual regions, or the addition of noise (to name a few)[98,99]. We point out the specific source of symmetry breaking in the models reviewed below.

Metastability via fixed-point memory. In 1990, Kelso et al. published a mathematical model[34] that aimed to explain the metastable behaviour observed in bimanual coordination studies of finger flexing with a pacing metronome[34]. This first model of metastability was based on the broken symmetry between the intrinsic frequencies of the metronome and the flexing finger. The results showed that phase and frequency synchronization between the metronome and the flexing finger were replaced by tendencies for in-phase and antiphase coordination, interspersed with desynchronization and phase wandering. Furthermore, when the control parameter (the metronome frequency) was set to a particular frequency, a bifurcation occurred wherein a fixed point (on-beat attractor) and a repeller (off-beat repeller) collided, resulting in their annihilation. Crucially, although no attractor was now present in the system, trajectories in the system were still attracted to where it once was, that is, to the memory of the fixed point (Fig. 2). Approaching this region, the system slowed down, dwelling for some time, before escaping to other regions of the attractor landscape through desynchronization and phase wandering. The spontaneous dwell and escape behaviour of the system indicated metastable dynamics (see Supplementary information S1 for the model equations and extensions and Video 2 for the behaviour of the generalized model).

Metastability via cycles of saddles with time delays. In 1993, Hansel et al. published a model of weakly coupled Hodgkin–Huxley neurons[100] without added noise but with broken symmetry parameterized through time delays[101]. This model exhibited clustering of oscillators into two groups in antiphase with individual oscillators

switching between the clusters, with the frequencies of the individual oscillators progressively slowing down before switching. These dynamics are owing to connected saddles, thereby demonstrating that the system was metastable (see Video 3 for the behaviour and Supplementary information S6 for the model equations and parameter settings for the video). Kori and Kuramoto[102] showed that this behaviour was also present in a model of neuronal bursting[103] and demonstrated that the slow switching between the clusters of oscillators could be invoked when the symmetry breaking was parameterized through randomly distributed time delays or heterogeneous coupling[102]. More recently, metastable behaviour was observed without noise, in both a whole-brain model of coupled oscillators using a conduction delay matrix derived from structural connectivity[69], and a structural connectome derived from diffusion MRI modelled together with time delays[104,105]. Together, these studies illustrate how time delays or heterogeneous connectivity can lead – even in the absence of noise – to metastability.

Metastability via cycles of saddles with asymmetric stimuli. Repeated sequences of metastable states were observed in a generalized high-dimensional Lotka–Volterra model that described the dynamics of firing rates in an inhibitory network subject to an external stimulus[106]. This dynamical model is the canonical model for winnerless competition in the evolution of species[9,106,107]. In the computational model, trajectories in the system moved towards a saddle, slowed down and eventually moved on towards the next saddle in the sequence. Symmetry breaking in the model was achieved through an asymmetric stimulus matrix, which governed the stimulation strength to any neuron depending on a connectivity rule[106]. This model can be used to explain how such dynamics arise, why they can be robust to noise, and how the nature of the saddles ensures flexibility.
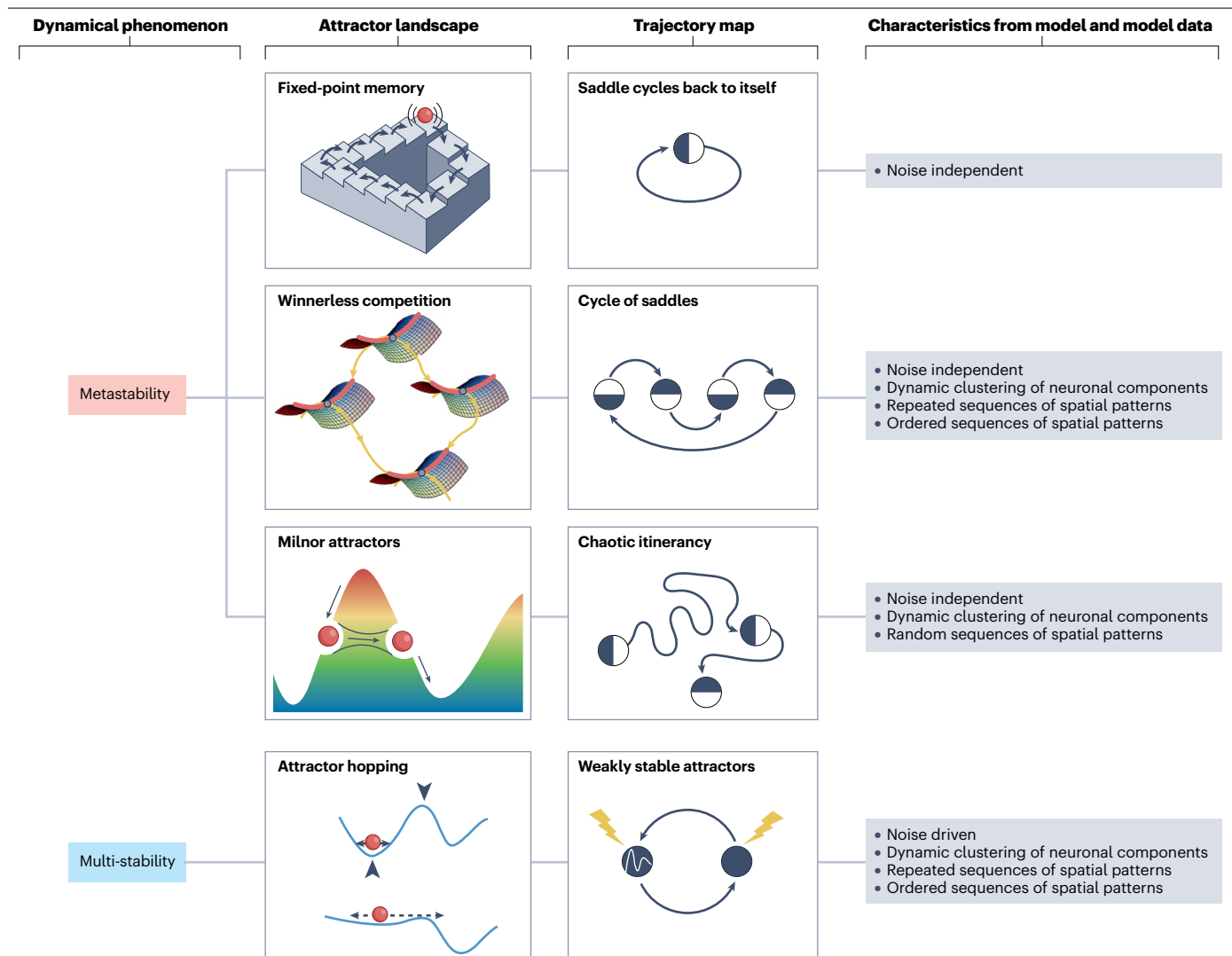
Dynamical phenomenon	Attractor landscape	Trajectory map	Characteristics from model and model data

Fig. 4 | A graphical overview of routes to transient state switching.
The presence of switching between multiple transient states (or between coordination and periods of lack thereof) can arise from metastability and multi-stability[164]. Multiple dynamical mechanisms can account for each family of dynamical stability, yielding sometimes different metastable characteristics. In the dynamics wherein the system cycles back to itself, there are actually no attractors, only the memory of the fixed point that was annihilated at a bifurcation. This fixed-point memory still attracts trajectories, but they slow down as they approach the metastable regime, giving the impression of stability, before escaping the region to eventually being attracted again to the memory. When the dynamics are explained as a cycle of saddles, the system moves to one saddle and remains there for a while before moving on again to the next saddle in the cycle. The cycle of saddles gives rise to either a repeated sequence of spatial patterns or a network containing repeated and unique sequences of spatial patterns. When the dynamics are explained through chaotic itinerancy, the system stays close to a destabilized Milnor attractor for a while before bursting out through the riddled basin (that is, full of holes) of the chaotic attractor to arrive close to another Milnor attractor − like travelling through a 'wormhole'. In multi-stable switching, the attractors are weakly stable, that is, they are close to a critical point, and the landscape has consequently flattened. Noise can now drive the system away from the weakly stable attractor towards another attractor in the system. The winnerless competition figure image is adapted with permission from ref. 162, SPIE.

Metastability via time delays. Metastable 'chimaera' states[108–110], wherein partial synchronization and desynchronization coexist, were found in a community-structured Kuramoto model with symmetrically coupled identical phase-lagged oscillators without added noise[57]. Symmetry breaking in the model was induced through the phase lags, that is, via time delays. The system transitioned from disorder to partial order, wherein spontaneously formed coalitions of phase-locked oscillators coexisted with desynchronized individual oscillators. The existence of saddles in this specific generalization of a Kuramoto model (that is, oscillators identical in size and natural frequency, community structure, and no noise) has, to our knowledge, not been formally investigated. This model of metastability was later used, in a subsequent work, to investigate the mechanisms underlying a proposed account of consciousness[111]. Building on the global workspace theory[112,113],

Review article

conscious processing was posited to entail competition between coalitions of specialist processes to access the global workspace and broadcast their influence throughout the brain[114-116] (see Supplementary information S4).

Metastability via chaotic itinerancy. Complex dynamics suggestive of metastability were found with synchronized chaotic attractors in a local discrete-time model with noise[42] and a whole-brain continuous-time network of conductance-based[117] neural mass models that were parameterized to exhibit chaotic behaviour[118] without added noise. These studies are focused on a phenomenon known as chaotic itinerancy which corresponds to relatively long periods of high synchronization that are punctuated by brief bursts of desynchronization[10,40,42] (Box 1 and Supplementary information S2). Moreover, a reduction in coupling between the chaotic oscillators was found to induce symmetry breaking in these systems. The models exhibited chaotic synchronization (see Supplementary information S2 for how synchronization is computed in chaotic systems that are not periodic)[119,120] and desynchronization, yielding complex dynamics including the appearance of phase-shifted synchronized clusters and oscillators switching among these clusters[42,118]. These studies illustrate metastability in chaotic systems, in which switching can occur between 'Milnor attractors' (which repel trajectories in some directions similar to a saddle) that are not directly contiguous in the dynamic landscape (see Supplementary information S2 for a more detailed and technical account of chaotic itinerancy and Milnor attractors).

We have reviewed several models of metastable dynamics. However, transient states and complex spatiotemporal patterns also arise from noise-driven multi-stability. We, therefore, review some models of multi-stability to gain an understanding of the differences between these dynamical phenomena.

Multi-stability in a whole-brain model. Resting-state functional connectivity, as observed, for example, in human fMRI data, has been proposed to arise as the result of noise-driven switching between bistable attractors in a corticothalamic model[121,122] or noise-driven exploration of multi-stable states in a whole-brain model of neuronal activity[123]. To investigate these proposals, a whole-brain global spiking attractor network model incorporating human structural connectivity was fitted to empirical fMRI data[124]. The results revealed a bifurcation determined by the global (inter-areal) coupling strength, at which the number of attractors grow from one to five. These results suggested that resting-state functional connectivity was influenced by latent or ghost attractors, that is, 'memories' of fixed points beyond the critical point (Box 1) that induced structure in the noise-driven explorations at the brink of the bifurcation[124]. The transient recurrent patterns of functional connectivity — or 'resting-state networks' — observed consistently across subjects in resting-state fMRI, have been proposed to be signatures of ghost attractors[125].

Multi-stability in spontaneous and evoked cortical activity. Recurrent sequences of spatiotemporal patterns, extracted and decoded with hidden Markov models, have been found in studies of individual spiking neurons in the gustatory cortex of rats[126-128]. These patterns were investigated in a recurrent network of spiking neurons organized into excitatory, inhibitory and unstructured populations[127]. Bifurcations were extracted analytically to identify fixed points (that is, attractors) as synaptic weights within excitatory clusters varied. It was proposed that internally generated noise — as a consequence

of excitation–inhibition balance — drove the system among these attractors, yielding repeated sequences of spatiotemporal patterns of neuronal activity. Similar patterns were found in simulations of evoked and resting-state conditions pointing to the potential role of synaptic plasticity for learning external stimuli.

Multi-stability in models with turbulent behaviour. Turbulence, a property of fluids when a smooth flow breaks up into whorls and eddies, has been found to provide a fundamental principle governing optimal mixing and efficient transfer of energy over space and time[129]. Recent research has applied ideas from turbulent dynamics to brain activity, proposing that the scale-free nature of turbulence-like behaviour provides a dynamical regime in which hierarchical information cascades allow the brain to function optimally despite its relative slowness[130-132]. Turbulence-like properties have been found in fast local field potentials in local brain regions[133] and in whole-brain dynamics measured with magnetoencephalography[134]. Based on Kuramoto's pioneering insight[56], signatures of turbulence can be quantified from the variability across space-time of the local KOP[135]. It was proposed that this local signature of turbulence is a signature of local vorticity, perhaps analogous to brain spiralling[136], and is in essence complementary with the concept of global metastability[135] but focusing on the local dynamics rather than the whole system described here[50,56,57,137-139]. However, to show signatures of turbulence in fast neuronal dynamics of magnetoencephalography, which has poorer spatial resolution than fMRI, a new signature of local turbulence was introduced[134]. This metric captures pairwise phase differences between regions rather than their mean-field phase and is defined as the standard deviation across space and time of this local signature (Supplementary information S7).

Neurobiological drivers of metastability

The computational neuroscience studies reviewed above reveal multiple dynamical mechanisms that can serve as alternative paths towards exhibiting transient states. These paths — and the additional paths to multi-stability — arise owing to specific features of the physical system. However, it is important to note that these models (and other models such as neuropercolation[44]) are not mutually exclusive. Indeed, all these models exhibit some features of metastability, including dynamic clustering, partial synchronization followed by rapid desynchronization, and repeated or random sequences of transient states. In fact, it is highly non-trivial to disambiguate these dynamical mechanisms solely using data (see the section 'The future: open questions and opportunities').

For this reason, instead of focusing on how to perform inferences from these models, a more fruitful avenue towards revealing the neurobiological origins of metastability is to investigate the drivers of the features that these models are capturing at microscopic and macroscopic levels. Here, we provide a brief overview of existent work in this direction.

Drivers at a microscopic level. Work on realistic models of recurrent spiking networks here implicate key roles for clustering, excitatory synaptic weights, recurrent inhibition and intrinsic noise in the dynamic behaviour of these networks. Specifically, grouping excitatory neurons into clusters yields slow firing rate fluctuations in the presence of intrinsic noise and in the absence of stimuli[140]. Furthermore, having heterogeneous cluster sizes has been observed to increase the probability of the system visiting large clusters, thus initiating recurrent sequences of transient states[127]. Additionally, recurrent inhibition has been found to reduce the number of simultaneously active clusters[127],

Review article

Glossary

Asymmetric
A systematic imbalance in some property of the system.

Attractor
A set of states on which many trajectories converge.

Basin of attraction
All the points in phase space that flow onto the attractor.

Bifurcation
A qualitative change in dynamics produced when a control parameter reaches a critical point.

Bistability
A form of dynamic stability wherein two attractors exist in the dynamics, that is, when there are two stable solutions of the differential equation describing the dynamics.

Chaos
A form of dynamical behaviour that can arise in a time-invariant nonlinear system. Chaos is characterized by sustained aperiodic (nonrepeating) oscillations, leading to extreme sensitivity of future states to small changes in present values of the system.

Chaotic attractor
An attractor that holds dynamics that are highly sensitive to their initial conditions.

Chaotic itinerancy
The behaviour of complicated systems with weakly attracting sets, in which destabilized attractors allow the system to leave its basin of attraction for another through a trajectory of connected saddles.

Control parameters
Parameters that modify a system of differential or difference equations, hence deforming the corresponding flows through phase space.

Crisis
The collision of an unstable periodic orbit and a coexisting chaotic attractor.

Critical fluctuations
Stochastic fluctuations that are orders of magnitude larger than normal, which occur when a system is close to a critical point. They may be sufficient to kick the system out of its basin and into the region of another attractor. A switch will occur, even though the original fixed point may still be classified as stable.

Critical point
The value of a control parameter at which a bifurcation occurs.

Dynamical systems theory
A branch of mathematics that studies how the state of systems evolves over time based on either an analytical (pencil and paper), a geometric (shapes) or a numeric (approximations using a computer) study of deterministic evolution equations.

Dynamic instabilities
Behavioural changes of the system in the vicinity of a bifurcation.

Fixed point
A point in the state space wherein the rate of change of the system with respect to time is equal to zero, corresponding to states at which the system remains unchanged unless perturbed.

Ghost attractors
Regions of phase space wherein the memory of a fixed point is attractive for the system. The memory is created by the annihilation of a fixed point and a repeller when a control parameter is changed.

Hidden Markov models
Statistical models that are used to describe the probabilistic relationship between sequences of observations and sequences of hidden states. They are used to classify sequences or predict future observations based on the underlying hidden processes that generate the data.

Metastability
A specific type of dynamics that may take place in a system with coexisting tendencies of attraction and repulsion, and is characterized by patterns that recur either in repeatable sequences (pattern) or flexible alternation (no pattern).

Milnor attractor
An attractor that no longer attracts all trajectories in its basin of attraction following an arbitrary small perturbation.

Monostability
A form of dynamic stability wherein a single attractor exists in the dynamics, that is, when there is one stable solution to the differential equation describing the dynamics.

Multi-stability
A form of dynamic stability wherein multiple stable attractors exist in the dynamics, that is, when there are two or more stable solutions to the differential equation describing the dynamics.

Order parameter
A single variable that captures the collective or macro-behaviour of a system composed of microscopic elements.

Phase space
The set of all possible states and, hence, contains all the allowed combinations of values of the variables of a system (also known as state variables).

Repeller
A set of states from which many trajectories migrate.

Saddles
Fixed points that are stable in one direction but unstable in another. Trajectories approach a saddle and are repelled away from it at the fixed point.

States
A state is a configuration of the variables of a system that is a solution to the equations.

Trajectory
A sequence of states within the phase space that satisfies the dynamics of the system as defined by its differential equation.

---

and intrinsic noise owing to excitatory–inhibitory balance has been observed to drive the system around multi-stable states[127].

Drivers at a macroscopic level. Drivers of metastability from macroscale computational models include influences from structural connectivity, network topology, spectral heterogeneities and intrinsic noise.

The relationship between topology and the std-KOP signature of metastability was established in models with community structure[57], based on healthy human connectomes[69], the connectomes of people with traumatic brain injuries[68], and the connectome of individuals with AD[59]. Specifically, compromised macroscopic organization of the human structural connectome, derived from subject-level diffusion tractography data, was associated with a reduction in a signature of metastability (std-KOP) in a cohort with AD (compared with controls)[59] together with reductions in higher-order network metrics including clustering coefficient and eigenvector centrality (influence of the node on a network) of the structural connectome. For traumatic brain injury, the model was informed by subject-level connectomes, and the same

signature of metastability was applied to the model data[62]. It was found that high-order network metrics – such as clustering coefficient and small worldness of the structural connectome – positively correlated with the signature of metastability and were also reduced.

In another study, a model of human functional connectivity constrained by the white matter connectome was used in combination with a focal lesioning approach to investigate the relationship between structural connectivity of macroscopic brain regions and large-scale neural dynamics[68]. Using the std-KOP signature of metastability, it was found that removing nodes with high eigenvector centrality (influence of the node on a network), or hub nodes joining topologically segregated network modules, led to increases in global metastability.

Pertaining to the role of heterogeneity, the first model of metastability[34] incorporated symmetry breaking through the heterogeneous intrinsic frequencies of the metronome and the flexing finger. A model structured with a human connectome also generated metastable behaviour when the nodes in the model were endowed with heterogeneous frequencies[141]. A similar model, but with stochastic dynamics, found that when the model was set at maximum metastability (std-KOP), the heterogeneous frequencies led to clustering of node phases and node frequencies[142].

Finally, in systems with metastable dynamics, studies have shown that noise can stabilize the system[68], reorganize frequency clusters and influence characteristic state duration times dependent on the magnitude of the noise[143].

The future: open questions and opportunities

In the previous sections, we reviewed the past and present of metastability research in neuroscience. In the following section, we discuss what may lay ahead for this area of research.

Metastability as a driver in behaviour, cognition and associated disorders

Using diverse signatures of metastability, we have seen convergent evidence that metastability in the brain may peak at some point during healthy development and then decline – a decline that seems to accelerate in cases of injury or disease. Indeed, both increases and decreases of metastability have been associated with altered cognitive function in psychiatric illnesses[59,64,77,93], healthy ageing[71,74], pharmacological interventions[94], extreme preterm birth[92] and traumatic brain injury[68]. However, abnormally high brain metastability may occur in psychosis[77]. This body of evidence (though obtained from diverse signatures) may point to an 'optimal range' for metastability, associated with healthy cognitive function. This conjecture may be reinforced by the notion – reflected in many of the proposed signatures of metastability – that the brain needs to simultaneously balance the degree of integration required for global coordination with segregation of neuronal ensembles to achieve functional specialization[4,144,145]. Metastability, as operationalized by these signatures, can be seen as a natural way of accomplishing this balancing act. Additionally, as one may argue, switching between transient states may allow the brain to visit available configurations of neuronal interdependency in anticipation of moment-to-moment changes in the environment[146].

Although these are compelling reasons for having metastability in the brain, throughout this article, we have taken care to conceptually disambiguate metastability and multi-stability. Crucially, we have also highlighted that commonly used signatures of metastability do not disambiguate metastability and multi-stability. The same is true for the above-mentioned reasons that supposedly make metastability

valuable for the brain: balancing integration and segregation and switching in preparation to a changing environment – both of which could probably be accomplished by metastability and multi-stability. Furthermore, it is worth noting that there may be scenarios wherein systems that have metastable dynamics at a microscopic level could appear as multi-stable at a macroscopic level owing to simplifications (for example, mean-field approximations) of the underlying model, and hence analyses and interpretations must be handled with care.

Overall, although the literature provides conclusive evidence for the value of transient dynamics, neither the theoretical reasons nor the empirical studies using signatures of metastability have conclusively demonstrated yet why (or whether) the brain should exhibit metastability over multi-stability. This does not undermine the empirical value that some of these signatures have proven to have as informative biomarkers – it only states that their interpretation requires nuance. Arguably, metastability presents the advantage that a metastable system does not need noise or other perturbations to escape from being 'trapped' in an attractor, hence allowing transient dynamics 'by design'. However, because noise is inherent in biological systems, and the state in which the brain resides may be in the vicinity of a critical point wherein fluctuations may be amplified (Box 3), it remains to be

---

Box 3 | Metastability and criticality

The concept of criticality — and its multiple signatures — has been widely applied to study brain activity across multiple scales[167] either by investigating phase transitions[8,46,168–170] or following the paradigm of self-organized criticality, in which a link between the control and order parameters spontaneously sets the system near a critical regime[171–176]. Criticality is a special phenomenon in physical systems that characterizes conditions wherein a small change on a control parameter (Box 1) can elicit a huge change on an order parameter — characterizing, for example, what happens when ice turns into water (with temperature being the control parameter)[177]. Systems that undergo such abrupt changes are said to undergo a phase transition, and the value of the control parameter at which this happens is known as the critical point. Systems at their critical state (that is, at their critical point) exhibit distinctive properties, including power-law scaling and long-range spatial and temporal correlations[177]. Although criticality was originally studied in systems in thermodynamic equilibrium, its study has also included systems far from equilibrium (for example, biological systems) undergoing non-trivial dynamics. Such systems have been shown to exhibit characteristic dynamical properties, including fluctuation enhancement (critical fluctuations) and slowing down of global variables[1,166].

Criticality and metastability are different concepts, although both refer to behaviours that may take place simultaneously. Multi-stable switching could be a consequence of critical fluctuations at a weakly stable state, that is, a state with a shallow valley in an attractor landscape. By contrast[164], metastable switching in the non-chaotic systems described elsewhere[9,57,101,143] does not require critical fluctuations for the behaviour to take place.

In summary, criticality can have a role for the emergence of metastable behaviours, but metastable dynamics do not require the brain to be operating exclusively 'at criticality', or to refer to criticality at all.

Review article

determined how essential this advantage of metastability would be for the brain.

This situation, however, does not call for conflating metastability and multi-stability – as has often been done in the literature. On the contrary, it calls for greater scrutiny into their differences, and greater care to distinguish them in computational and empirical work. Below, we outline promising directions for accomplishing this, which are now beginning to emerge: generative models, null hypotheses and the distribution of dwell times.

Generative models and null hypotheses for inferring metastability

Arguably, inferring metastability is best achieved by inverting a generative model of the system and finding that it best fits the data when tuned into a metastable regime (that is, saddles are present). Inferring 'metastability' from 'variability of data' is a weak claim because that variability may be present in both metastable and multi-stable systems. Other signatures without this weakness, but which are less tractable in large systems, have been proposed, including the k number[4], fluctuations of relative phase[147,148] (Supplementary information S3), and phase difference derivative[149]. In the setting in which suggestions of metastability are made based directly on data metrics, additional signatures to disambiguate the various candidates outlined in this Review should be assessed. The key differentiator is the manner (and, hence, the statistics) of how the system dwells (multi-stability), cycles or sequences (metastability), or bursts (itinerancy) or exhibits other behaviour not covered in this Review, such as a random walk with no dwelling (criticality)[5,150].

With the myriad of signatures that can be found in the literature, choosing the signature most appropriate to address a given scientific question at hand is not always an easy task. Whereas some signatures conceptually correspond to properties related to integration and segregation (they themselves are not clearly defined concepts, which have been operationalized in a variety of ways), other signatures are associated with the repertoire of recurring metastable spatial patterns of brain activity. In the absence of a strong rationale to pick a specific signature, it would be advisable to compare and contrast several of the reported signatures against one's chosen null to obtain a more complete overview. Additionally, it is worth keeping in mind that the claim that a particular model yields metastable[151] or multi-stable[124] dynamics often hinges on the assumptions made and the aspects of the system that were chosen as the relevant control parameters and/or state variables[152]. One interesting avenue to further investigate metastability and its multiple signatures is by building approaches based on alternative models to contrast hypotheses. Along this direction, two types of approach are particularly promising: (1) null models that allow assessment of the relevance of a specific effect and (2) generative models to describe how a system behaves under certain conditions.

Null models represent a powerful approach to derive inferences, not only against an empirical control group but also against controls generated according to desired specifications[153,154]. For example, some studies[75,93,155] have compared observed dynamics against dynamics obtained from a null model (such as phase randomization), using this approach to determine whether the data could have arisen from purely random switching (as would be the case in noise-driven multi-stability). The more stringently one's null model can reflect the difference between metastability and multi-stability, the stronger is the inference that can be drawn. For example, rather than randomizing phases, one could build different biophysically plausible models of how the data

may have been generated: one multi-stable and another metastable. Owing to the flexibility of computational models, the type of stability can be determined a priori. Therefore, one can then use a suitable measure of goodness-of-fit to determine whether the empirically observed data are more consistent with the metastability-generating or multi-stability-generating mechanism.

More broadly, computational models can also be used for further clarifying the computational advantages of metastability, and simulated dynamics can be obtained from metastable and multi-stable systems that are otherwise identical. These dynamics can then be compared across a battery of measures, such as different indicators of integration, segregation and information transmission[156]. Does the metastable system systematically outperform the multi-stable one on any such dimension? If so, it would be a plausible candidate for further theoretical investigation.

The same approach can also be used to evaluate the neural mechanisms (that is, the properties of the physical system) underlying metastability, building on the studies reviewed above. For example, regional neurobiological correlates of a given signature of metastability may be identified in empirical data (for example, through databases such as NeuroMaps[157], which provides access to brain maps of receptor density, transcriptomic patterns, and other features of structural and functional brain architecture). Maps showing significant associations with metastability may then be incorporated into generative computational models of brain dynamics that allow for regional heterogeneities, comparing the resulting dynamics against models with random or absent heterogeneity[154]. Greater markers of metastability under the true map would provide evidence for a mechanistic role, beyond mere correlation. Establishing whether any cellular, molecular and metabolic disease markers are associated with metastability could lead to improved understanding of the pathophysiology of neurodevelopmental and neurodegenerative disorders and potentially lead to a more sophisticated understanding of treatment targets and efficacy.

Distribution of dwell times as a potential differentiator between metastability and multi-stability

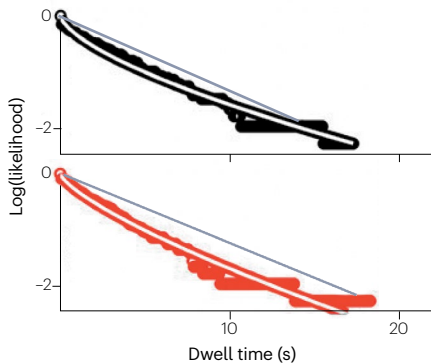A promising – yet still only partially explored[5,158] – way to determine whether observed dynamics come from a metastable or multi-stable system is to investigate the distribution of state dwell times. In effect, switching between two or more stable states in multi-stable systems is driven by noise, and hence the dwell time will (under standard assumptions on the noise distribution) display an exponential distribution[15,121,122] (Fig. 5a). By contrast, switching within a cycle of saddles in metastable dynamics does not involve trapping, and hence dwell times can exhibit distributions that are different from exponential ones. In a computational whole-brain model of coupled neural masses, the distribution of dwell times for different patterns of large-scale spontaneous activity appears to follow a gamma distribution[105] (Fig. 5b). Additionally, in a neural mass model informed by the macaque connectome, chaotic phase synchronization and desynchronization between dynamic clusters of neuronal ensembles as a consequence of chaotic itinerancy also yielded a duration of synchronized episodes that followed a gamma distribution[158] (Fig. 5c). However, it is worth mentioning that a biophysical model has successfully reproduced the 'waxing and waning' of alpha-band (8–13 Hz) EEG activity in humans at rest using noise-driven switching between a low-amplitude and a high-amplitude oscillatory state, resulting in a (stretched-)exponential (that is, not a gamma) distribution (Fig. 5a).
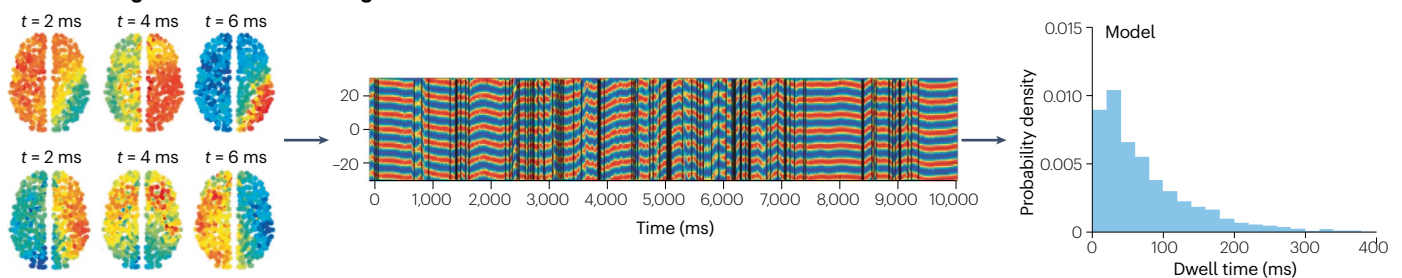
Overall, although computational observations tend to support the relationship between metastability and the gamma distribution of dwell times, more work is needed to build a more principled understanding of this phenomenon. If such observations continue to be corroborated and a robust underlying rationale can be found, goodness-of-fit of the gamma distribution via, for example, the Kullback–Leibler divergence[159], could be a way to obtain an all-or-none test for metastability (versus multi-stability), unlike signatures that are graded and conflate the two.
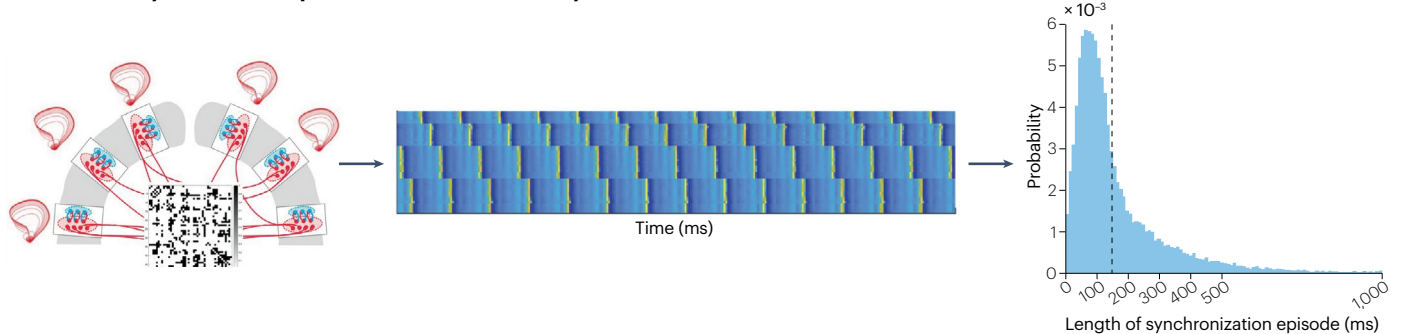

Fig. 5 | Duration statistics in models of multi-stable and metastable dynamics. a, Dwell-time distributions for a modelled state-dependent noise-driven multi-stable system. Cumulative distributions of dwell times of oscillations in the alpha frequency band (8–13 Hz) during resting state. The instantaneous power in the alpha band switches spontaneously and erratically between distinct low-amplitude (black curve) and high-amplitude (red curve) modes, and their dwell times follow long-tailed stretched-exponential forms (white curves). The grey curves indicate a simple exponential form. Cumulative distributions were separately rescaled to have a mean value of 1. **b**, Dwell time gamma distributions arising from metastable waves in a network of neural mass models. The left image shows large-scale wave patterns in the model, with six snapshots of the dynamics of pyramidal mean membrane potential at various latencies. The middle panel highlights that each pattern has a signature wherein vertical lines depict low values of inter-hemispheric cross-correlation corresponding to wave transitions. The right chart shows dwell time distributions for the model at particular parameter settings. **c**, Duration of synchronization episodes in a network of neural mass models. The left panel shows a macaque connectome-based network of chaotic oscillators. Inhibitory cells are shown in blue and pyramidal cells are shown in red. In the middle panel, neuronal clusters synchronize and desynchronize in phase and in antiphase with other clusters. The right panel shows that the duration of synchronization episodes resembles a gamma distribution. The mean duration is shown with a dashed vertical line. Part **a** is adapted with permission from ref. 122, Society for Neuroscience. Part **b** is adapted from ref. 105, Springer Nature Limited. Left image in part **c** is adapted from ref. 6, Springer Nature Limited. Middle image in part **c** is adapted from ref. 118, CC BY 4.0. Right image in part **c** is adapted with permission from ref. 158, copyright (2007) National Academy of Sciences, U.S.A.

Review article

Concluding remarks

Metastability is a fascinating phenomenon that offers valuable ways to frame brain function. In this Review, we have provided a broad overview of the historical foundations of metastability, its dynamical mechanisms, its multiple signatures and their applications in empirical studies, and the most promising avenues of progress for the field. Our core message is that metastability can be rigorously characterized in terms of the rules that govern the dynamics of a system, and this rigour is essential for understanding that (1) metastability can arise through different paths but is still one unifying construct and (2) it should not be conflated with any of its practical signatures, nor with related but distinct concepts of complexity, criticality or multi-stability.

The future of metastability as a fundamental construct for neuroscience rests on demonstrating its role (over and above multi-stability) in behaviour, cognition and associated disorders and on understanding its neurobiological underpinnings. Such a feat would not only deepen our understanding of the dynamics that govern brain activity but also open the door towards devising interventions aimed at restoring a healthy regime of brain dynamics through appropriate pharmacological, psychotherapeutic or non-invasive brain stimulation control strategies.

Overall, metastability has an increasingly prominent role in contemporary neuroscience, informing computational models and offering promising biomarkers — a role we expect to keep growing in the future. It is our hope that the synthesis provided in this Review will foster rigour and clarity in future studies, further extending these promising lines of investigation.

Code availability

All codes used to estimate the signatures of metastability from fMRI data and to generate Videos 1, 2 and 3 are publicly available. See Supplementary information for details.

References

1. Kelso, J. A. S. *Dynamic Patterns*: *The Self-Organization of Brain and Behavior* 334 (MIT Press, 1995).
   This is an excellent book that promotes viewing the brain as a complex system and understanding behaviour as a result of cooperation between neuronal ensembles.
2. Kelso, J. A. S. & Tognoli, E. in *Neurodynamics of Cognition and Consciousness* (eds Perlovsky, L. I. & Kozma, R.) 39–59 (Springer, 2007).
3. Kelso, J. A. S. in *Evolution of Dynamical Structures in Complex Systems* (eds Friedrich, R. & Wunderlin, A.) 223–234 (Springer, 1992).
4. Kelso, J. A. S. An essay on understanding the mind. *Ecol. Psychol.* **20**, 180–208 (2008).
5. Cocchi, L., Gollo, L. L., Zalesky, A. & Breakspear, M. Criticality in the brain: a synthesis of neurobiology, models and cognition. *Prog. Neurobiol.* **158**, 132–152 (2017).
6. Breakspear, M. Dynamic models of large-scale brain activity. *Nat. Neurosci.* **20**, 340–352 (2017).
7. Friston, K. J. Transients, metastability, and neuronal dynamics. *NeuroImage* **5**, 164–171 (1997).
   This paper highlights and attempts to reconcile metastability in coordination dynamics and chaotic itinerancy.
8. Fuchs, A., Kelso, J. A. S. & Haken, H. Phase transitions in the human brain: spatial mode dynamics. *Int. J. Bifurc. Chaos* **02**, 917–939 (1992).
9. Rabinovich, M., Huerta, R., Varona, P. & Afraimovich, V. S. Transient cognitive dynamics, metastability, and decision making. *PLoS Comput. Biol.* **4**, e1000072 (2008).
   This paper introduces stable heteroclinic channels as a mathematical object to explain metastability.
10. Tsuda, I. Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behav. Brain Sci.* **24**, 793–810 (2001).
11. La Camera, G., Fontanini, A. & Mazzucato, L. Cortical computations via metastable activity. *Curr. Opin. Neurobiol.* **58**, 37–45 (2019).
12. Brinkman, B. A. W. et al. Metastable dynamics of neural circuits and networks. *Appl. Phys. Rev.* **9**, 011313 (2022).
13. Rossi, K. L. et al. Dynamical properties and mechanisms of metastability: a perspective in neuroscience. Preprint at https://doi.org/10.48550/arXiv.2305.05328 (2024).
14. Attneave, F. Multistability in perception. *Sci. Am.* **225**, 62–71 (1971).
15. Feudel, U. & Grebogi, C. Multistability and the control of complexity. *Chaos* **7**, 597–604 (1997).
16. Kraut, S., Feudel, U. & Grebogi, C. Preference of attractors in noisy multistable systems. *Phys. Rev. E* **59**, 5253–5260 (1999).
17. Pisarchik, A. N. & Feudel, U. Control of multistability. *Phys. Rep.* **540**, 167–218 (2014).
18. Ashwin, P. & Postlethwaite, C. On designing heteroclinic networks from graphs. *Phys. D Nonlinear Phenom.* **265**, 26–39 (2013).
19. Tsuda, I. Chaotic itinerancy and its roles in cognitive neurodynamics. *Curr. Opin. Neurobiol.* **31**, 67–71 (2015).
20. Rosas, F. E. et al. Disentangling high-order mechanisms and high-order behaviours in complex systems. *Nat. Phys.* **18**, 476–477 (2022).
21. Miets, H. A. & Chevalier, J. On the crystallization of sodium nitrate. *Mineral. Mag.* **14**, 123–133 (1906).
22. Haken, H., Kelso, J. A. S. & Bunz, H. A theoretical model of phase transitions in human hand movements. *Biol. Cybern.* **51**, 347–356 (1985).
23. Kelso, J. A. S. Phase transitions and critical behavior in human bimanual coordination. *Am. J. Physiol.* **246**, R1000–R1004 (1984).
24. Schöner, G. & Kelso, J. A. S. Dynamic pattern generation in behavioral and neural systems. *Science* **239**, 1513–1520 (1988).
25. Ashwin, P., Buescu, J. & Stewart, I. Bubbling of attractors and synchronisation of chaotic oscillators. *Phys. Lett. A* **193**, 126–139 (1994).
26. Aizenman, M. & Lebowitz, J. L. Metastability effects in bootstrap percolation. *J. Phys. A Math. Gen.* **21**, 3801–3813 (1988).
27. Afraimovich, V., Verichev, N. N. & Rabinovich, M. Stochastic synchronization of oscillation in dissipative systems. *Radiophys. Quantum Electron.* **29**, 795–803 (1986).
28. Kaneko, K. Clustering, coding, switching, hierarchical ordering, and control in a network of chaotic elements. *Phys. D Nonlinear Phenom.* **41**, 137–172 (1990).
29. Shlosman, S. in *Encyclopedia of Mathematical Physics* (eds Françoise, J.-P., Naber, G. L. & Tsun, T. S.) 417–420 (Academic, 2006).
30. Kryukov, V. The metastable and unstable states in the brain. *Neural Netw.* **1**, 264 (1988).
31. Niebur, E., Schuster, H. G. & Kammen, D. M. Collective frequencies and metastability in networks of limit-cycle oscillators with time delay. *Phys. Rev. Lett.* **67**, 2753–2756 (1991).
32. Niebur, E., Schuster, H. G. & Kammen, D. M. in *Neural Network Dynamics* (eds Taylor, J. G. et al.) 226–233 (Springer, 1992).
33. Holst, E. R. M. von. *The Behavioural Physiology of Animals and Man*: *The Collected Papers of Erich von Holst* (Univ. of Miami Press, 1973).
34. Kelso, J. A. S., Del Colle, J. D. & Schöner, G. in *Attention and Performance 13*: *Motor Representation and Control* 139–169 (Erlbaum, 1990).
   This paper highlights how metastability arises from broken symmetry.
35. DeGuzman & Kelso, J. A. S. in *Principles Of Organization In Organisms* (ed. Mittenthal, J. E.) (Addison-Wesley, 1992).
36. Schöner, G., Haken, H. & Kelso, J. A. S. A stochastic theory of phase transitions in human hand movement. *Biol. Cybern.* **53**, 247–257 (1986).
37. Fingelkurts, A. & Fingelkurts, A. A. Making complexity simpler: multivariability and metastability in the brain. *Int. J. Neurosci.* **114**, 843–862 (2004).
38. Rabinovich, M. et al. Dynamical encoding by networks of competing neuron groups: winnerless competition. *Phys. Rev. Lett.* **87**, 068102 (2001).
39. Kaneko, K. On the strength of attractors in a high-dimensional system: Milnor attractor network, robust global attraction, and noise-induced selection. *Phys. D Nonlinear Phenom.* **124**, 322–344 (1998).
40. Kaneko, K. & Tsuda, I. Chaotic itinerancy. *Chaos* **13**, 926–936 (2003).
   This paper provides a summary of chaotic itinerancy and its applications at the start of this century, wherein switching occurs between fully developed chaos and ordered behaviour characterized by low-dimensional dynamics.
41. Breakspear, M. Perception of odors by a nonlinear model of the olfactory bulb. *Int. J. Neural Syst.* **11**, 101–124 (2001).
42. Breakspear, M., Terry, J. R. & Friston, K. J. Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a biophysical model of neuronal dynamics. *Netw. Comput. Neural Syst.* **14**, 703–732 (2003).
   This paper provides a thorough theoretical explanation and computational demonstration of metastability and chaotic itinerancy in chaotic systems.
43. Afraimovich, V., Rabinovich, M. & Varona, P. Heteroclinic contours in neural ensembles and the winnerless competition principle. *Int. J. Bifurc. Chaos* https://doi.org/10.1142/S0218127404009806 (2003).
44. Kozma, R. & Puljic, M. Random graph theory and neuropercolation for modeling brain oscillations at criticality. *Curr. Opin. Neurobiol.* **31**, 181–188 (2015).
45. Kozma, R. & Puljic, M. Hierarchical random cellular neural networks for system-level brain-like signal processing. *Neural Netw.* **45**, 101–110 (2013).
46. Kozma, R., Puljic, M., Balister, P., Bollobás, B. & Freeman, W. J. Phase transitions in the neuropercolation model of neural populations with mixed local and non-local interactions. *Biol. Cybern.* **92**, 367–379 (2005).
47. Breakspear, M. Nonlinear phase desynchronization in human electroencephalographic data. *Hum. Brain Mapp.* **15**, 175–198 (2002).
48. Seliger, P., Tsimring, L. S. & Rabinovich, M. I. Dynamics-based sequential memory: winnerless competition of patterns. *Phys. Rev. E* **67**, 011905 (2003).
49. MATLAB. *Version 9.11.0.1769968 (R2021b)* (The MathWorks Inc., 2021).
50. Tognoli, E. & Kelso, J. A. S. The metastable brain. *Neuron* **81**, 35–48 (2014).
   This seminal paper introduces metastability to the wider neuroscience community.

Review article

51. Fries, P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* **9**, 474–480 (2005).

52. Fries, P. Rhythms for cognition: communication through coherence. *Neuron* **88**, 220–235 (2015).

53. Varela, F., Lachaux, J.-P., Rodriguez, E. & Martinerie, J. The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* **2**, 229–239 (2001).

54. Dowdall, J. R. & Vinck, M. Coherence fails to reliably capture inter-areal interactions in bidirectional neural systems with transmission delays. *NeuroImage* **271**, 119998 (2023).

55. Sancristóbal, B., Vicente, R. & Garcia-Ojalvo, J. Role of frequency mismatch in neuronal communication through coherence. *J. Comput. Neurosci.* **37**, 193–208 (2014).

56. Kuramoto, Y. in *Chemical Oscillations, Waves, and Turbulence,* Vol. 19, 60–88 (Springer, 1984).

57. Shanahan, M. Metastable chimera states in community-structured oscillator networks. *Chaos* **20**, 013108 (2010).
   This seminal paper provides the first tractable signature of metastability that becomes commonly used in empirical and computational studies.

58. Abeysuriya, R. G. et al. A biophysical model of dynamic balancing of excitation and inhibition in fast oscillatory large-scale networks. *PLoS Comput. Biol.* **14**, e1006007 (2018).

59. Alderson, T. H., Bokde, A. L. W., Kelso, J. A. S., Maguire, L. & Coyle, D. Metastable neural dynamics in Alzheimer's disease are disrupted by lesions to the structural connectome. *NeuroImage* **183**, 438–455 (2018).

60. Deco, G., Kringelbach, M. L., Jirsa, V. K. & Ritter, P. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Sci. Rep.* **7**, 3095 (2017).

61. Hancock, F. et al. Metastability, fractal scaling, and synergistic information processing: what phase relationships reveal about intrinsic brain activity. *NeuroImage* **259**, 119433 (2022).

62. Hellyer, P. J., Scott, G., Shanahan, M., Sharp, D. J. & Leech, R. Cognitive flexibility through metastable neural dynamics is disrupted by damage to the structural connectome. *J. Neurosci.* **35**, 9050–9063 (2015).

63. Jobst, B. M. et al. Increased stability and breakdown of brain effective connectivity during slow-wave sleep: mechanistic insights from whole-brain computational modelling. *Sci. Rep.* **7**, 4634 (2017).

64. Lee, W. H., Doucet, G. E., Leibu, E. & Frangou, S. Resting-state network connectivity and metastability predict clinical symptoms in schizophrenia. *Schizophr. Res.* **201**, 208–216 (2018).

65. Lee, W. H. & Frangou, S. Emergence of metastable dynamics in functional brain organization via spontaneous fMRI signal and whole-brain computational modeling. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (*EMBC*) 4471–4474 (IEEE, 2017).

66. Lord, L.-D. et al. Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. *NeuroImage* **199**, 127–142 (2019).

67. Mediano, P. A. M. et al. Integrated information as a common signature of dynamical and information-processing complexity. *Chaos* **32**, 013115 (2022).

68. Váša, F. et al. Effects of lesions on synchrony and metastability in cortical networks. *NeuroImage* **118**, 456–467 (2015).

69. Cabral, J., Hugues, E., Sporns, O. & Deco, G. Role of local network oscillations in resting-state functional connectivity. *NeuroImage* **57**, 130–139 (2011).
   To our knowledge, this is the first paper to use a signature of metastability in whole-brain modelling.

70. Breakspear, M. & Terry, J. R. Nonlinear interdependence in neural systems: motivation, theory, and relevance. *Int. J. Neurosci.* **112**, 1263–1284 (2002).

71. De Alteriis, G. et al. EiDA: a lossless approach for dynamic functional connectivity; application to fMRI data of a model of ageing. *Imaging Neurosci.* **2**, 1–22 (2024).

72. Deco, G., Tagliazucchi, E., Laufs, H., Sanjuán, A. & Kringelbach, M. L. Novel intrinsic ignition method measuring local-global integration characterizes wakefulness and deep sleep. *eNeuro* https://doi.org/10.1523/ENEURO.0106-17.2017 (2017).

73. Luppi, A. I. et al. Reduced emergent character of neural dynamics in patients with a disrupted connectome. *NeuroImage* **269**, 119926 (2023).

74. Escrichs, A. et al. Whole-brain dynamics in aging: disruptions in functional connectivity and the role of the rich club. *Cereb. Cortex* **31**, 2466–2481 (2021).

75. Escrichs, A. et al. Characterizing the dynamical complexity underlying meditation. *Front. Syst. Neurosci.* **13**, 27 (2019).

76. Alonso Martínez, S., Deco, G., Ter Horst, G. J. & Cabral, J. The dynamics of functional brain networks associated with depressive symptoms in a nonclinical sample. *Front. Neural Circuits* **14**, 570583 (2020).

77. Hancock, F. et al. Metastability as a candidate neuromechanistic biomarker of schizophrenia pathology. *PLoS ONE* **18**, e0282707 (2023).

78. Cabral, J. et al. Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Sci. Rep.* **7**, 1–13 (2017).

79. Allen, E. A. et al. Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* **24**, 663–676 (2014).

80. Zamani Esfahlani, F. et al. High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proc. Natl Acad. Sci. USA* **117**, 28393–28401 (2020).

81. Vidaurre, D., Smith, S. M. & Woolrich, M. W. Brain network dynamics are hierarchically organized in time. *Proc. Natl Acad. Sci. USA* **114**, 12827–12832 (2017).

82. Stevner, A. B. A. et al. Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep. *Nat. Commun.* **10**, 1035 (2019).

83. Demertzi, A. et al. Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci. Adv.* **5**, eaat7603 (2019).

84. Luppi, A. I. et al. Consciousness-specific dynamic interactions of brain integration and functional diversity. *Nat. Commun.* **10**, 4616 (2019).

85. Shine, J. M. et al. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* **92**, 544–554 (2016).

86. Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O. & Betzel, R. F. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nat. Neurosci.* **23**, 1644–1654 (2020).

87. Michel, C. M., Brechet, L., Schiller, B. & Koenig, T. Current state of EEG/ERP microstate research. *Brain Topogr.* **37**, 169–180 (2024).

88. Mortaheb, S. et al. Mind blanking is a distinct mental state linked to a recurrent brain profile of globally positive connectivity during ongoing mentation. *Proc. Natl Acad. Sci. USA* **119**, e2200511119 (2022).

89. Barttfeld, P. et al. Signature of consciousness in the dynamics of resting-state brain activity. *Proc. Natl Acad. Sci. USA* **112**, 887–892 (2015).

90. Gutierrez-Barragan, D. et al. Unique spatiotemporal fMRI dynamics in the awake mouse brain. *Curr. Biol.* **32**, 631–644.e6 (2022).

91. Lurie, D. J. et al. Questions and controversies in the study of time-varying functional connectivity in resting fMRI. *Netw. Neurosci.* **4**, 30–69 (2020).

92. Padilla, N. et al. Disrupted resting-sate brain network dynamics in children born extremely preterm. *Cereb. Cortex* **33**, 8101–8109 (2023).

93. Córdova-Palomera, A. et al. Disrupted global metastability and static and dynamic brain connectivity across individuals in the Alzheimer's disease continuum. *Sci. Rep.* **7**, 1–14 (2017).

94. Signorelli, C. M., Uhrig, L., Kringelbach, M., Jarraya, B. & Deco, G. Hierarchical disruption in the cortex of anesthetized monkeys as a new signature of consciousness loss. *NeuroImage* **227**, 117618 (2021).

95. Kelso, J. A. S. & Engstrom, D. A. *The Complementary Nature* (MIT Press, 2008).

96. Luppi, A. I. et al. Computational modelling in disorders of consciousness: closing the gap towards personalised models for restoring consciousness. *NeuroImage* **275**, 120162 (2023).

97. Golubitsky, M. & Stewart, I. *The Symmetry Perspective* (Birkhäuser, 2002).

98. Jirsa, V. & Sheheitli, H. Entropy, free energy, symmetry and dynamics in the brain. *J. Phys. Complex.* **3**, 015007 (2022).

99. Pillai, A. S. & Jirsa, V. K. Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior. *Neuron* **94**, 1010–1026 (2017).

100. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544 (1952).

101. Hansel, D., Mato, G. & Meunier, C. Clustering and slow switching in globally coupled phase oscillators. *Phys. Rev. E* **48**, 3470–3477 (1993).
   This paper uses a Kuramoto model of weakly coupled oscillators that included the first harmonic demonstrating metastability between antiphase clusters of synchronized oscillators.

102. Kori, H. & Kuramoto, Y. Slow switching in globally coupled oscillators: robustness and occurrence through delayed coupling. *Phys. Rev. E* https://doi.org/10.1103/PhysRevE.63.046214 (2000).

103. Hindmarsh, J. L., Rose, R. M. & Huxley, A. F. A model of neuronal bursting using three coupled first order differential equations. *Proc. R. Soc. Lond. B Biol. Sci.* **221**, 87–102 (1984).

104. Cabral, J. et al. Exploring mechanisms of spontaneous functional connectivity in MEG: how delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations. *NeuroImage* **90**, 423–435 (2014).

105. Roberts, J. A. et al. Metastable brain waves. *Nat. Commun.* **10**, 1–17 (2019).
   This paper provides an expansion of metastable dynamics to brain waves in a computational model.

106. Afraimovich, V., Zhigulin, V. P. & Rabinovich, M. On the origin of reproducible sequential activity in neural circuits. *Chaos* **14**, 1123–1129 (2004).

107. Afraimovich, V., Gong, X. & Rabinovich, M. Sequential memory: binding dynamics. *Chaos* **25**, 103118 (2015).

108. Kuramoto, Y. & Battogtokh, D. in *Nonlinear Phenomena in Complex Systems,* Vol. 5 (ed. Haken, H.) 380–385 (2002).

109. Abrams, D. M. & Strogatz, S. H. Chimera states for coupled oscillators. *Phys. Rev. Lett.* **93**, 174102 (2004).

110. Abrams, D. M., Mirollo, R. E., Strogatz, S. H. & Wiley, D. A. Solvable model for chimera states of coupled oscillators. *Phys. Rev. Lett.* **101**, 084103 (2008).

111. Shanahan, M. *Embodiment and the Inner Life*: *Cognition and Consciousness in the Space of Possible Minds* (Oxford Univ. Press, 2010).

112. Baars, B. J. *A Cognitive Theory of Consciousness* (Cambridge Univ. Press, 1988).

113. Baars, B. J. in *Progress in Brain Research* 150 (ed. Laureys, S.) 45–53 (Elsevier, 2005).

114. Dehaene, S. & Changeux, J.-P. Experimental and theoretical approaches to conscious processing. *Neuron* **70**, 200–227 (2011).

115. Dehaene, S., Kerszberg, M. & Changeux, J.-P. A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl Acad. Sci. USA* **95**, 14529–14534 (1998).

116. Mashour, G. A., Roelfsema, P., Changeux, J.-P. & Dehaene, S. Conscious processing and the global neuronal workspace hypothesis. *Neuron* **105**, 776–798 (2020).

117. Morris, C. & Lecar, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **35**, 193–213 (1981).

Review article

118. Heitmann, S. & Breakspear, M. Putting the "dynamic" back into dynamic functional connectivity. *Netw. Neurosci.* **2**, 150–174 (2018).

119. Fujisaka, H. & Yamada, T. A new intermittency in coupled dynamical systems. *Prog. Theor. Phys.* **74**, 918–921 (1985).

120. Pecora, L. M. & Carroll, T. L. Synchronization in chaotic systems. *Phys. Rev. Lett.* **64**, 821–824 (1990).

121. Freyer, F., Aquino, K., Robinson, P. A., Ritter, P. & Breakspear, M. Bistability and non-Gaussian fluctuations in spontaneous cortical activity. *J. Neurosci.* **29**, 8512–8524 (2009).

122. Freyer, F. et al. Biophysical mechanisms of multistability in resting-state cortical rhythms. *J. Neurosci.* **31**, 6353–6361 (2011).

123. Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R. & Jirsa, V. K. Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Comput. Biol.* **4**, e1000196 (2008).

124. Deco, G. & Jirsa, V. K. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J. Neurosci.* **32**, 3366–3375 (2012).

125. Vohryzek, J., Deco, G., Cessac, B., Kringelbach, M. L. & Cabral, J. Ghost attractors in spontaneous brain activity: recurrent excursions into functionally-relevant BOLD phase-locking states. *Front. Syst. Neurosci.* **14**, 20 (2020).

126. Jones, L. M., Fontanini, A., Sadacca, B. F., Miller, P. & Katz, D. B. Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proc. Natl Acad. Sci. USA* **104**, 18772–18777 (2007).

127. Mazzucato, L., Fontanini, A. & Camera, G. L. Dynamics of multistable states during ongoing and evoked cortical activity. *J. Neurosci.* **35**, 8214–8231 (2015).

128. Mazzucato, L., La Camera, G. & Fontanini, A. Expectation-induced modulation of metastable activity underlies faster coding of sensory stimuli. *Nat. Neurosci.* **22**, 787–796 (2019).

129. Frisch, U. *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge Univ. Press, 1995).

130. Deco, G., Kemp, M. & Kringelbach, M. L. Leonardo da Vinci and the search for order in neuroscience. *Curr. Biol.* **31**, R704–R709 (2021).

131. Deco, G. & Kringelbach, M. L. Turbulent-like dynamics in the human brain. *Cell Rep.* **33**, 108471 (2020).

132. Deco, G. et al. Rare long-range cortical connections enhance human information processing. *Curr. Biol.* **31**, 4436–4448.e5 (2021).

133. Sheremet, A., Qin, Y., Kennedy, J. P., Zhou, Y. & Maurer, A. P. Wave turbulence and energy cascade in the hippocampus. *Front. Syst. Neurosci.* https://doi.org/10.3389/fnsys.2018.00062 (2019).

134. Deco, G., Liebana Garcia, S., Sanz Perl, Y., Sporns, O. & Kringelbach, M. L. The effect of turbulence in brain dynamics information transfer measured with magnetoencephalography. *Commun. Phys.* **6**, 1–8 (2023).

135. Kawamura, Y., Nakao, H. & Kuramoto, Y. Noise-induced turbulence in nonlocally coupled oscillators. *Phys. Rev. E* **75**, 036209 (2007).

136. Xu, Y., Long, X., Feng, J. & Gong, P. Interacting spiral wave patterns underlie complex brain dynamics and are related to cognitive processing. *Nat. Hum. Behav.* **7**, 1196–1215 (2023).

137. Cabral, J., Kringelbach, M. L. & Deco, G. Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* **114**, 102–131 (2014).

138. Kitzbichler, M. G., Smith, M. L., Christensen, S. R. & Bullmore, E. Broadband criticality of human brain network synchronization. *PLoS Comput. Biol.* **5**, e1000314 (2009).

139. Wildie, M. & Shanahan, M. Metastability and chimera states in modular delay and pulse-coupled oscillator networks. *Chaos* **22**, 043131 (2012).

140. Litwin-Kumar, A. & Doiron, B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat. Neurosci.* **15**, 1498–1505 (2012).

141. Ponce-Alvarez, A. et al. Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. *PLoS Comput. Biol.* **11**, e1004100 (2015).

142. Pang, J. C., Gollo, L. L. & Roberts, J. A. Stochastic synchronization of dynamics on the human connectome. *NeuroImage* **229**, 117738 (2021).

143. Ashwin, P., Orosz, G., Wordsworth, J. & Townley, S. Dynamics on networks of cluster states for globally coupled phase oscillators. *SIAM J. Appl. Dyn. Syst.* **6**, 728–758 (2007).

144. Tononi, G., Sporns, O. & Edelman, G. M. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc. Natl Acad. Sci. USA* **91**, 5033–5037 (1994).

145. Sporns, O. *Networks of the Brain* 412 (MIT Press, 2011).

146. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**, 43–56 (2011).

147. Zhang, M., Beetle, C., Kelso, J. A. S. & Tognoli, E. Connecting empirical phenomena and theoretical models of biological coordination across scales. *J. R. Soc. Interface* **16**, 20190360 (2019).
    This paper provides a generalization on the Haken–Kelso–Bunz model for self-organization of behaviour from dyadic to multiple agents in social interactions.

148. Zhang, M., Kelso, J. A. S. & Tognoli, E. Critical diversity: divided or united states of social coordination. *PLoS ONE* **13**, e0193843 (2018).

149. Breakspear, M., Williams, L. M. & Stam, C. J. A novel method for the topographic analysis of neural activity reveals formation and dissolution of 'dynamic cell assemblies'. *J. Comput. Neurosci.* **16**, 49–68 (2004).

150. Roberts, J. A., Boonstra, T. W. & Breakspear, M. The heavy tail of the human brain. *Curr. Opin. Neurobiol.* **31**, 164–172 (2015).

151. Stratton, P. & Wiles, J. Global segregation of cortical activity and metastable dynamics. *Front. Syst. Neurosci.* **9**, 119 (2015).

152. Schirner, M., Kong, X., Yeo, B. T. T., Deco, G. & Ritter, P. Dynamic primitives of brain network interaction. *NeuroImage* **250**, 118928 (2022).

153. Váša, F. & Mišić, B. Null models in network neuroscience. *Nat. Rev. Neurosci.* **23**, 493–504 (2022).

154. Luppi, A. I. et al. A synergistic core for human brain evolution and cognition. *Nat. Neurosci.* **25**, 771–782 (2022).

155. López-González, A. et al. Loss of consciousness reduces the stability of brain hubs and the heterogeneity of brain dynamics. *Commun. Biol.* **4**, 1–15 (2021).

156. Shine, J. M., Aburn, M. J., Breakspear, M. & Poldrack, R. A. The modulation of neural gain facilitates a transition between functional segregation and integration in the brain. *eLife* **7**, e31130 (2018).

157. Markello, R. D. et al. neuromaps: structural and functional interpretation of brain maps. *Nat. Methods* **19**, 1472–1479 (2022).

158. Honey, C. J., Kötter, R., Breakspear, M. & Sporns, O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl Acad. Sci. USA* **104**, 10240–10245 (2007).

159. Andriulli, M., Starling, J. K. & Schwartz, B. Distributional discrimination using Kolmogorov-Smirnov statistics and Kullback-Leibler divergence for gamma, log-normal, and Weibull distributions. In *2022 Winter Simulation Conference* (WSC) 2341–2352 (IEEE, 2022).

160. McKinley, J. et al. Third party stabilization of unstable coordination in systems of coupled oscillators. *J. Phys. Conf. Ser.* **2090**, 012167 (2021).

161. Roberts, J. A., Friston, K. J. & Breakspear, M. Clinical applications of stochastic dynamic models of the brain, part I: a primer. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2**, 216–224 (2017).

162. Bociort, F. & van Grol, P. Systematics of the design shapes in the optical merit function landscape. *Proc. SPIE* https://doi.org/10.1117/12.853924 (2010).

163. Deco, G. & Kringelbach, M. L. Hierarchy of information processing in the brain: a novel 'intrinsic ignition' framework. *Neuron* **94**, 961–968 (2017).

164. Kelso, J. A. S. Multistability and metastability: understanding dynamic coordination in the brain. *Philos. Trans. Biol. Sci.* **367**, 906–918 (2012).

165. Dezhina, Z. et al. Establishing brain states in neuroimaging data. *PLoS Comput. Biol.* **19**, e1011571 (2023).

166. Haken, H. *Synergetics: An Introduction* (Springer, 1983).

167. O'Byrne, J. & Jerbi, K. How critical is brain criticality? *Trends Neurosci.* **45**, 820–837 (2022).

168. Kelso, J. A. S. et al. A phase transition in human brain and behavior. *Phys. Lett. A* **169**, 134–144 (1992).

169. Jirsa, V. K., Friedrich, R., Haken, H. & Kelso, J. A. S. A theoretical model of phase transitions in the human brain. *Biol. Cybern.* **71**, 27–35 (1994).

170. Kelso, J. A. S. in *Criticality in Neural Systems* 67–104 (Wiley, 2014).

171. Beggs, J. M. & Plenz, D. Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**, 11167–11177 (2003).

172. Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M. & Ilmoniemi, R. J. Long-range temporal correlations and scaling behavior in human brain oscillations. *J. Neurosci.* **21**, 1370–1377 (2001).

173. Lombardi, F., Shriki, O., Herrmann, H. J. & de Arcangelis, L. Long-range temporal correlations in the broadband resting state activity of the human brain revealed by neuronal avalanches. *Neurocomputing* **461**, 657–666 (2021).

174. Petermann, T. et al. Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. *Proc. Natl Acad. Sci. USA* **106**, 15921–15926 (2009).

175. Shriki, O. et al. Neuronal avalanches in the resting MEG of the human brain. *J. Neurosci.* **33**, 7079–7090 (2013).

176. Tagliazucchi, E., Balenzuela, P., Fraiman, D. & Chialvo, D. R. Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Front. Physiol.* **3**, 15 (2012).

177. Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, 1971).

Review article

[1]Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [2]Department of Informatics, University of Sussex, Brighton, UK. [3]Sussex Centre for Consciousness Science, University of Sussex, Brighton, UK. [4]Centre for Psychedelic Research, Department of Brain Science, Imperial College London, London, UK. [5]Centre for Eudaimonia and Human Flourishing, University of Oxford, Oxford, UK. [6] Sussex AI, University of Sussex, Brighton, UK. [7]Centre for Complexity Science, Department of Brain Science, Imperial College London, London, UK. [8]St John's College, University of Cambridge, Cambridge, UK. [9]Department of Psychiatry, University of Oxford, Oxford, UK. [10]Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI, USA. [11]Department of Computing, Imperial College London, London, UK. [12]Division of Psychology and Language Sciences, University College London, London, UK. [13]Life and Health Sciences Research Institute School of Medicine, University of Minho, Braga, Portugal. [14]Computational Neuroscience Group, Center for Brain and Cognition, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain. [15]Institución Catalana de la Recerca i Estudis Avancats (ICREA), Barcelona, Spain. [16]Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. [17]School of Psychological Sciences, Monash University Clayton, Melbourne, Victoria, Australia. [18]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark. [19]School of Psychological Sciences, College of Engineering, Science and the Environment, University of Newcastle, Newcastle, New South Wales, Australia. [20]Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA. [21]Intelligent Systems Research Centre, Ulster University, Derry~Londonderry, Northern Ireland. [22]The Bath Institute for the Augmented Human, University of Bath, Bath, UK. [23]The Institute for Human and Synthetic Minds, King's College London, London, UK. [24]These authors contributed equally: Fran Hancock, Fernando E. Rosas.