Hierarchical network structure as the source of hierarchical dynamics (power-law frequency spectra) in living and non-living systems: How state-trait continua (body plans, personalities) emerge from first principles in biophysics

R. Goekoop [a,*], R. Kleijn de [b]

[a] Free University Amsterdam, Department of Behavioral and Movement Sciences, Parnassia Academy, Parnassia Group, PsyQ, Department of Anxiety Disorders, Early Detection and Intervention Team (EDIT), Lijnbaan 4, 2512VA The Hague, the Netherlands
[b] Faculty of Social and Behavioral Sciences, Department of Cognitive Psychology, Pieter de la Courtgebouw, Postbus 9555, 2300 RB Leiden, the Netherlands

ARTICLE INFO

ABSTRACT

Living systems are hierarchical control systems that display a *small world* network structure. In such structures, many smaller clusters are nested within fewer larger ones, producing a fractal-like structure with a 'power-law' cluster size distribution (a mereology). Just like their structure, the dynamics of living systems shows fractal-like qualities: the timeseries of inner message passing and overt behavior contain high frequencies or 'states' (treble) that are nested within lower frequencies or 'traits' (bass), producing a power-law frequency spectrum that is known as a 'state-trait continuum' in the behavioral sciences. Here, we argue that the power-law dynamics of living systems results from their power-law network structure: organisms 'vertically encode' the deep spatio-temporal structure of their (anticipated) environments, to the effect that many small clusters near the base of the hierarchy produce high frequency signal changes and fewer larger clusters at its top produce ultra-low frequencies. Such ultra-low frequencies exert a tonic regulatory pressure that produces morphological as well as behavioral traits (i.e., body plans and personalities). Nested-modular structure causes higher frequencies to be embedded within lower frequencies, producing a power-law state-trait continuum. At the heart of such dynamics lies the need for efficient energy dissipation through networks of coupled oscillators, which also governs the dynamics of non-living systems (e.q., earthquakes, stock market fluctuations). Since hierarchical structure produces hierarchical dynamics, the development and collapse of hierarchical structure (e.g., during maturation and disease) should leave specific traces in the system dynamics (shifts in lower frequencies, i.e. morphological and behavioral traits) that may serve as early warning signs to system failure. The applications of this idea range from (bio)physics and phylogenesis to ontogenesis and clinical medicine.

* Corresponding author.
E-mail address: R.Goekoop@vu.nl (R. Goekoop).

1. On the structure (spatial characteristics) of living systems

What causes organisms to have different body plans and personalities? In this paper, we address this question by looking at universal principles that govern the structure and dynamics of living systems. Living systems such as cells, organs, organisms and social networks are known to share a generic network structure that is called a *small world* topology (Watts and Strogatz, 1998). *Small world* networks are a class of networks in which the number of connections per system component (node 'degree') is unevenly distributed across system components, i.e., most nodes have few connections but some have many. For example, most genes, cells, neurons, neural circuits and social individuals have few connections, but some have many (e.g., hub genes, hub neurons, alpha males, community workers). The degree distribution of small world networks follows a characteristic inverse relationship called a 'power-law' degree distribution (Clauset et al., 2009) (Fig. 1). This pattern deviates markedly from the Gaussian or 'normal' distributions of attributes such as body weight or height. A well-known hallmark of power-law distributions is that the natural logarithm of the distribution produces a straight line (Fig. 1). Another feature is that it is possible to zoom in or out on a power-law curve and still observe the same shape. For this reason, power-law degree distributions are said to have 'scale invariant' or 'scale free' features, which is alternatively referred to as 'self-similarity' or 'fracticality'[1] (Barabasi, 2009; Song et al., 2005).

In *small world* networks, highly connected nodes ('hubs') converge onto other hubs to form a hierarchy of hub nodes (a 'rich club' (Colizza et al., 2006)) (Fig. 2). This can be compared to teams of horses that are kept in check by a number of horse cart drivers (hubs), which in turn serve as horses that are kept in check by yet higher order drivers (hubs), etcetera, to form a pyramidal structure with a broad base and a narrow apex. As a result of this network topology, messages can travel along hub structures across highly efficient routes, causing any two nodes in the network to be connected via only a small number of intermediate steps (hence the term '*small world*'). Apart from producing efficient pathways, hubs contract parts of the network into so called clusters (modules), which are communities of nodes that share more connections amongst themselves than with their environments (Newman, 2012). This combination of high clustering and low average pathlength is called a *small world* network topology.

In *small world* networks, clusters may themselves serve as hub nodes at a higher spatial scale level of observation that contract collections of other clusters into superclusters and so on, to produce a nested modular, hierarchical structure (Fig. 2A). For instance, a set of hub genes in one cluster may connect to hub genes within other clusters to produce a clustering of clusters. Such nested clustering continues until only a few large modules form the top of a hierarchy of part-whole relationships. A *small world* network topology can be identified in such structures regardless of the spatial scale level of observation, which is why they are called scale invariant or scale free systems (Barabasi, 2009). When examining the distribution of cluster sizes in such fractal-like systems, a power-law is again obtained (Fig. 2B).

2. On the dynamics (temporal characteristics) of living systems

Apart from their network structure, the dynamics of living systems is known to show signs of scale invariance. The timeseries of inner message
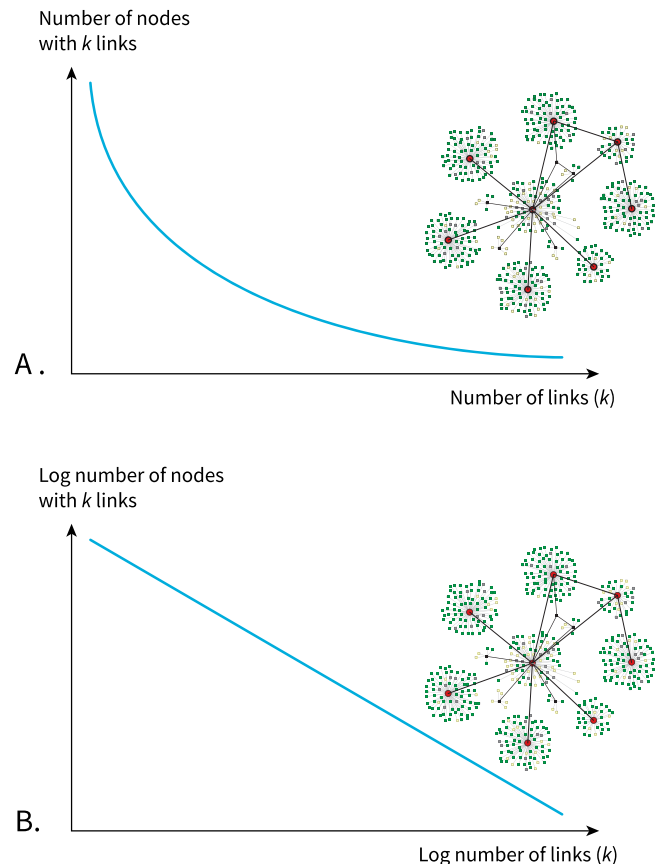
Fig. 1. Living systems: small world networks of which the distribution of links across nodes follows a power-law. *A. The distribution of links across nodes (the degree distribution) in small world networks follows a power-law: most nodes have few connections but some (hubs) have many. B. The natural logarithm (ln) of this power-law degree distribution produces a straight line $Y = b* ln (s)$, where b is the 'power-law exponent', which indicates the steepness of the slope of the line.*

---

[1] *Technically, there is a distinction between scale invariant and scale free that is usually unpacked in relation to the renormalisation group. In essence, scale invariance requires the conservation of a system's dynamics (or Lagrangian) at successive scales of coarse-graining. Scale free is a slightly stronger notion of scaling variance: scale free systems are scale invariant systems with no characteristic scale. Similarly, there is a technical distinction between self-similarity and fractality. Fractal refers to a fractional dimension that usually presents with self-similarity.*
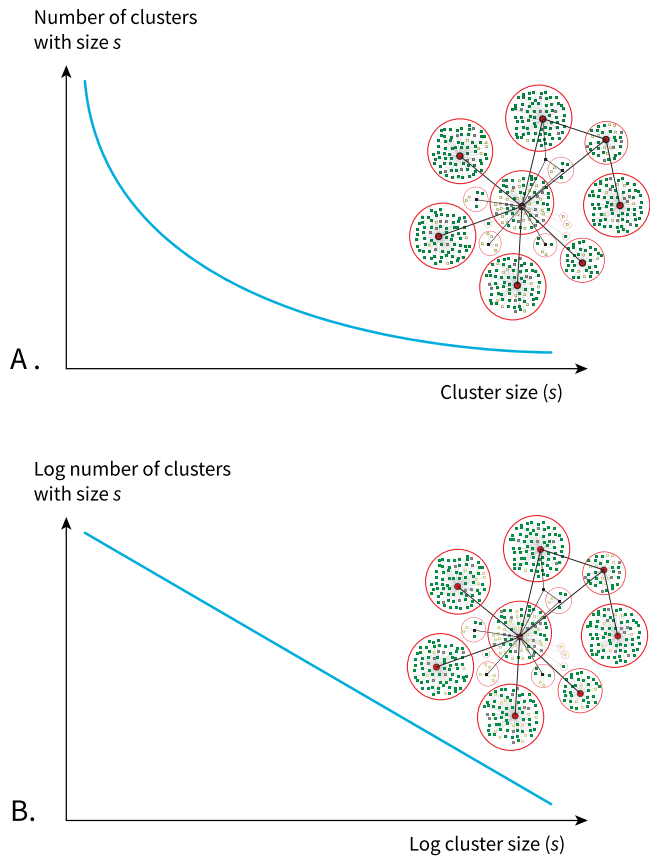
Fig. 2. Living systems: nested modular (hierarchical) network structures of which the cluster size distribution follows a power-law. A. In small world network structures, network clusters show conditional dependencies in space, i.e., a given cluster can only exist conditionally upon the existence of its constituent subclusters, producing a spatial hierarchy of part-whole relationships (a 'mereology'). The distribution of cluster size in such scale free networks (e.g., living systems) follows a power-law. B. The natural logarithm of this power-law cluster size distribution produces a straight line.

passing or overt behavior of such systems shows fast fluctuations that are 'nested' within slower fluctuations, producing a hierarchy of part-whole relationships (He, 2014; He et al., 2010). This can be observed by decomposing the timeseries into their constituent frequencies that are represented by sine waves (this is called Fourier transformation: Fig. 3A). Each frequency within this frequency spectrum can be assigned a value that indicates the average amplitude at which that frequency is present within the timeseries. It turns out that the lower (base) frequencies are expressed at the highest amplitudes, after which amplitude smoothly falls off as a power of frequency, producing a typical 'power-law' frequency distribution (Fig. 3B, C). The various frequency components show conditional dependencies, often such that low-frequency phase changes produce high frequency 'bursts': a phenomenon called phase-amplitude coupling (Fig. 3D) (He et al., 2010; Velarde et al., 2019; Canolty et al., 2006). During task performance, certain intermediate frequencies temporally gain in prominence, which shows up as 'bumps' on the power-law curve (Fig. 3E) (He, 2014).

Power-law frequency distributions dominate the timeseries of living systems across a wide range of spatiotemporal scales, varying from sub-millisecond responses of individual photoreceptors and membrane potentials to the distribution of clades across evolutionary timescales (He, 2014; He et al., 2010). They occur anywhere from gene expression and protein synthesis to photosynthesis, respiratory cycles, neural firing rates, neural field potentials and the dynamics of neural systems at large (Clauset et al., 2009). Power-law frequency distributions govern the dynamics of externally observable behavior such as locomotion, posture
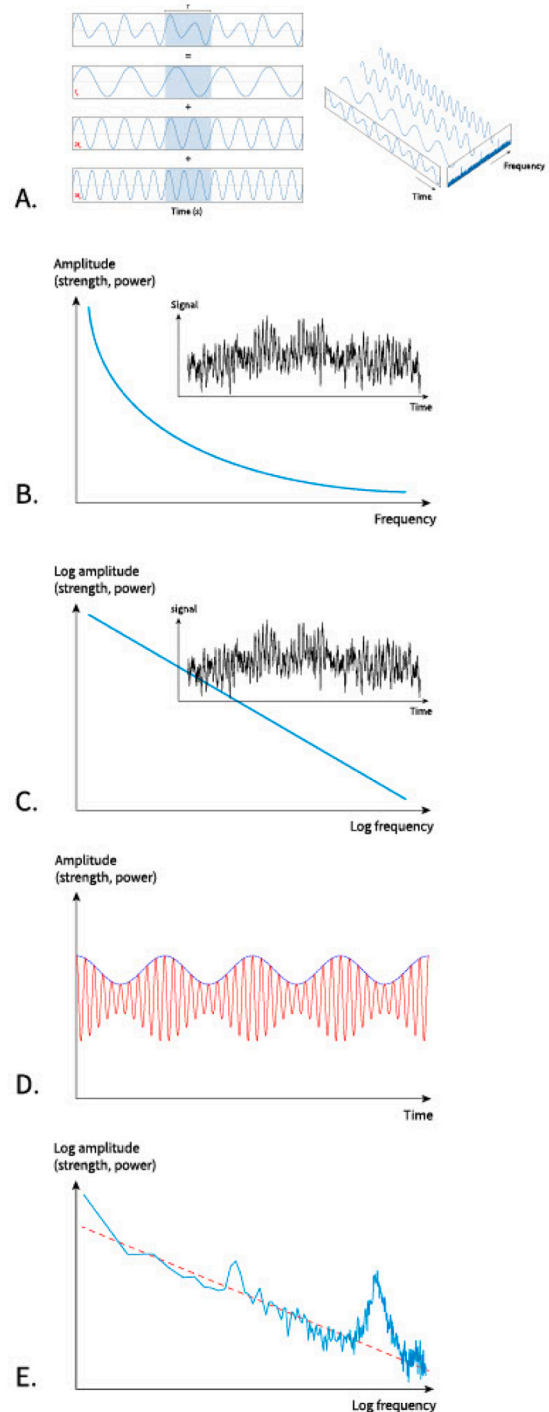
Fig. 3. The inner message passing and overt behavior of living systems show nested modular (hierarchical) dynamics of which the distribution of frequency components follows a power-law. A. Fourier transform is a method by which any signal can be decomposed into its constituent frequencies. B. The distribution of frequency components of a signal (e.g., an EEG scalp recording) follows a power-law $P \sim 1/f^{\beta}$, where P is power or amplitude, f is frequency, and β is the power law exponent (typically in the range of 0–3). C. The natural logarithm of this power-law frequency distribution follows a straight line (see Fig. 1). D. The various frequencies show conditional dependencies. Often, the phase of lower frequencies determines the amplitudes of higher frequencies, so that lower frequencies act as carrier 'waves' for higher frequencies (phase-amplitude coupling or amplitude modulation). E. Task performance produces 'bumps' on the power-law curve (figure shows an EEG power spectrum).

control, finger tapping, key pressure, reaction time, hit rate, applied force and speech (He, 2014; He et al., 2010). This is also true for the dynamics of subjective inner experience such as mood fluctuations (emotions), motivational changes (mania) and cognitive functions (memory and attention) (Nakamura et al., 2008; Perdikis et al., 2011). Power-law dynamics are found in social networks and the internet, making it a truly scale invariant phenomenon (Barabasi, 2005; Clegg et al., 2010; de Menezes and Barabasi, 2004). In essence, a power-law frequency distribution describes a state-trait continuum, in which the rapid fluctuations or 'states' of a system (behavioral 'weather') are superposed onto lower frequencies or 'traits' (behavioral 'climate') (Steyer et al., 1999). A power-law state-trait continuum therefore appears to be a universal feature that governs the dynamics of living systems.

3. On the functional significance of power-law dynamics in living systems

The ubiquity of the power-law frequency distribution has sparked questions regarding its functional significance. At first, power-laws were considered to reflect measurement artifacts or mere by-products of complex systems, especially since they are equally present in living systems as in abiotic systems such as sand piles (hence the mildly dismissive term $1/f$ or pink 'noise'). However, recent studies have found evidence to the contrary (He, 2014; He et al., 2010). Changes in scale-free brain activity have been found during development (Fransson et al., 2013; Smit et al., 2011), sleep (Tagliazucchi et al., 2013), task performance (He, 2011) and various physical, neurological and psychiatric disorders (Maxim et al., 2005; Tolkunov et al., 2010; Wei et al., 2013; Montez et al., 2009). Such changes involve a flattening or steepening of the power-law curve, indicating a shift in the degree to which different frequencies contribute to the signal at large. Such findings suggest that distinct generative mechanisms underwrite the presence of power-law frequency distributions as well as changes in power-law dynamics during (extreme) task performance, development and disease. Nevertheless, it remains unclear what mechanisms are involved or what degree of universality they might display.

In this paper, we propose that the hierarchical dynamics of living systems (power-law frequency spectra) result from their hierarchical network structure (*small world* network structure). To fully appreciate this relationship, we will first show that living systems can be abstracted as nested modular (hierarchical) networks of coupled oscillators, with an information bottleneck or 'bowtie' motif that allows them to function as hierarchical (Bayesian) control systems. We then argue that oscillations in such systems result from circularly causal relationships between excitatory and inhibitory nodes that underwrite a process of hierarchical message passing and predictive coding. Next, we discuss evidence that organisms 'vertically encode' the deep spatiotemporal structure of their environments to the effect that lower hierarchical regions encode rapidly changing events and higher hierarchical regions encode the slow dynamics of their surroundings. In other words, hub structures at the top of a (regulatory) hierarchy produce slow oscillations (traits) whereas nodes at the base of a hierarchy display fast fluctuations (states), with each intermediate hierarchical level producing its own characteristic frequency (a state-trait continuum). Cross-frequency coupling should therefore reflect the coordination between different levels of a nested modular hierarchy, providing a definition for bottom-up and top-down control. We then argue that the typical runoff of amplitude with frequency in power-law spectra results from the fact that living systems are open dissipative systems (i.e., systems that are open to the exchange of energy or matter with the outside world). Such systems are forced to dissipate energy efficiently across multiple hierarchical levels and corresponding frequency bands, with equal amounts of energy being distributed across lower and higher frequencies, producing higher amplitudes for lower frequencies and vice versa. The same mechanism has previously been proposed as an explanation for the power-law dynamics of non-living systems, such as earthquake dynamics or stock market fluctuations. Crucially, this means that a single biophysical principle may explain the dynamics of living as well as non-living systems.

Our theory has several remarkable consequences, which center around the fact that the top of a regulatory hierarchy is responsible for producing the stable behavior or 'traits' of a system (i.e., the offsets in timeseries). In living systems, these include the stable aspects of inner experience and overt behavior (personalities) as well as morphological traits (body plans). Small alterations in key regulatory systems at the top of the control hierarchy may produce large changes at its base (behavior), leading to a spectrum of individual differences in personalities and physical traits. Such differences allow organisms to specialize in different econiches. A subtle 'tweaking' of regulatory areas (e.g., point mutations) may therefore suffice to distribute organisms across widely different (social) econiches and optimize survival rates. This provides a principled account of the specialization and speciation of living systems and puts bowties center stage as 'hotspots of evolution'. On a practical note, alterations in power-law frequency spectra may serve as early warning signs to the collapse of hierarchical control, which is a key aspect of many (medical) disorders and failing physical systems.

4. Living systems as nested modular, hierarchical control systems

In a previous paper (Goekoop and de Kleijn, 2021a), we proposed that the nested modular network structure of living systems allows them to function as hierarchical control systems. In this view, the global structure of organisms resembles that of central heating systems: the input to the system (heat) is encoded by the input part of the system (a heat sensor) and compared to a reference state (the setpoint of a thermostat), after which the difference (the error) is conveyed to the output parts of the system (the radiator) to affect the environment (an increase in environmental temperature). This cycle is repeated until the error is reduced to a minimum. This is nicely illustrated by woodlice, which keep on running around erratically until their surrounding humidity levels reach near 100%, which is why we find these creatures in all sorts of nooks and crannies (Fortier and Friedman, 2018). Rather than directly smelling or seeing a damp corner a few yards away and moving directly towards it in a controlled fashion, woodlice keep their motor systems active and vary their output pseudorandomly until they bump into a set of environmental conditions (an 'econiche') that fits their preset reference values (e.g. ~100% humidity), which is when they finally come to rest. Such behavior protects these creatures from desiccation and predation and keeps them stable and intact (homeostasis, survival). Thus, behavior is in the service of producing novel percepts by altering the environment, the aim of which is to bring the organism closer to its reference state: a phenomenon known as 'active sensing' or perceptual control (Schroeder et al., 2010). This leaves organisms free to generate any kind of behavior that contributes to achieving a specified reference state (e.g. rolling up, digging in or hiding in tight spaces all help to prevent desiccation and evade predators). Such behavioral flexibility allows them to solve many different and unexpected challenges, which greatly adds to their stability (Powers, 1973).

The nested modular (hierarchical) network structure of organisms turns out to be ideally suited to solve an important puzzle that every organism faces, which is known as the binding problem (Rohe et al., 2019). In a capricious world, organisms have to decide whether some set of observations is caused by a set of independent causal factors (e.g., three different rivals) or by a single causal factor (e.g., a single mate). In other words, the co-occurrence of several events may have a different meaning and require different actions depending on whether such events have a common underlying (hidden) cause, or rather several independent (hidden) causes. In case of three distinct rivals, these should be encoded separately (segregated, unbinded), whereas a single mate should be encoded as a single factor (integrated, binded). In hierarchical networks, low-level nodes or clusters are used to encode statistically separable contextual cues (e.g., juicy, sweet, green and round), whereas

higher-level hub structures encode the instantaneous co-occurrence of these factors, which corresponds to a more contextually integrated yet parsimonious (abstract) model of the environment ('apple'). This principle can be repeated across several hierarchical levels, which allows organisms to model their econiches with increasing amounts of contextual integration yet parsimony, i.e., abstraction (Tenenbaum et al., 2011). This allows (hierarchically) higher organisms to encode categories such as 'fruit' instead of apples, pears and grapes (Rigoli et al., 2017). Thus, the top of a nested modular hierarchy serves as a natural funnel or 'information bottleneck' (Tishby and Zaslavsky, 2015) that causes its input at its base to be encoded by a shrinking number of network nodes (compressed). This allows organisms to capture the key properties of the econiches into which they are embedded.

Organisms do not merely encode the current states of their environments at increasing levels of abstraction, which would amount to hierarchical perception. Rather, their input layers smoothly transition into throughput layers that encode states (i.e., have values) that deviate increasingly from current reality (Fig. 4). Such high level throughput layers (Ha and Schmidhuber, 2018; Friston et al., 2021a) encode econiches that are not yet realized but are to be approached (or rather avoided) by generating an output sequence (Goekoop and de Kleijn, 2021b). Such (possible) models of possible worlds are called 'goal states' or 'world models'. Interestingly, hierarchical structure can also be used to encode the *sequential* co-occurrence of events (e.g. first blossom, then apple). This has the effect of modeling causal relationships (Friston et al., 2018; Cui et al., 2016; Kiebel et al., 2008) and anticipating future events (Pezzulo et al., 2022). Organisms use their hierarchical structure to make increasingly long-term predictions: when one moves further up the throughput hierarchy, (hub) nodes encode increasingly abstract events that are projected ever more distantly into the future (e.g., 'future fruit', e.g., the next harvest). Thus, a 'goal-hierarchy' is produced in which long-term and abstract goals are encoded at the top of the hierarchy. Such global goals are increasingly unpacked (decoded or decompressed) into their constituent subgoals when following a goal hierarchy from the top down towards its base. Goal hierarchies smoothly transition into the top of the output hierarchy, where they are further unpacked into executive functions and subfunctions, to eventually form detailed 'output commands' that activate various output organs in a sequence (e.g. vesicles, flagella, muscles, endocrine glands) (Pezzulo et al., 2022, 2015; Friedlander et al., 2015; Maisto et al., 2015; Adams et al., 2013). This process is called 'action control' (Pezzulo et al., 2015; Maisto et al., 2015; Pezzulo et al., 2018; de Kleijn et al., 2014). The subsequent output is a hierarchically controlled sequence of basic action-perception cycles (behavioral 'primitives' or 'reflexes') that constitutes the overt behavior of an organism. Such behavior serves to alter the environment and, hence, perception, to bring the organism closer to its goals.

In short, signaling pathways in living systems naturally form a dual information bottleneck motif, in which multiple input streams converge onto fewer throughput systems (the top of the hierarchy), which in turn diverge onto multiple output systems, producing dedicated compartments for hierarchical perception (input), goal setting (throughput) and action control (output) (Kirchhoff et al., 2018; Kitano, 2004). In systems biology, information bottleneck motifs are called 'bowtie' (2D) or 'hourglass' (3D) motifs because of their physical resemblance to such objects (Fig. 4) (Kitano, 2004). The 'wings' of the bowtie (or the bulbs of the hourglass) produce action-perception patterns (behavior), whereas the 'knot' of the bowtie (or the waist of the hourglass) exerts top-down (goal-directed) control over such patterns. The hub nodes that reside at the knot of the bowtie can be compared to generals at the top of a hierarchical chain of command, who acquire a global overview of the battlefield by combining multiple sources of information, after which they need to snap their fingers only occasionally to start a cascade of orders down the executive hierarchy that will eventually make their troops move in different directions (Fig. 4). Bowtie motifs have been observed in network systems at all spatial scale levels of observation,

including molecular signaling (Citri and Yarden, 2006), gene regulatory networks (Li et al., 2012; Yu and Gerstein, 2006), neurons, nerves and neural systems (Kitano, 2004), whole brains (Markov et al., 2013), large-scale social networks and the Internet (Fujita et al., 2019). In statistical physics, information bottlenecks are called Markov Blankets. These are limited sets of nodes (hubs) that separate two larger sets of nodes (clusters) into statistically separable compartments (Kirchhoff et al., 2018). Bowtie motifs engage in a dynamic loop with their environments, allowing organisms to function as hierarchical control systems (Fig. 4).

Folded bowties can be stacked next to and on top of each other to produce nested modular and hierarchically organized control systems (Fig. 4 and next section). The level of behavioral complexity that such structures produce depends on their hierarchical breadth as well as their hierarchical depth (Pezzulo et al., 2022). Broad hierarchies allow for a detailed articulation (factorization, orthogonalization) of context factors. Deep hierarchies allow for high levels of integration across such context factors, producing highly contextualized yet parsimonious ('abstract') models of the world. Hierarchies that are both broad and deep allow for complex behavioral repertoires that are known as goal-directed behavior (e.g. 'harvesting' requires ploughing the field, seeding the grain, watering the shoots, fertilizing the soil, fending off birds, etcetera, in a logical order). The term 'sophistication' has been reserved for such behavior (Pezzulo et al., 2022; Friston et al., 2021a), which includes abilities such as 'agency' (self-functioning, self-directedness, autonomy), social functioning (communion, cooperativeness) and normative functioning (e.g., following moral guidelines or rules) (Goekoop and de Kleijn, 2021b; Constant et al., 2019). This relates to the cybernetic literature, which speaks of 'homeostatic control' (or system 1) when referring to lower hierarchical levels that control relatively simple processes such as blood pressure or ventilation that are aimed at short-term stability of 'allostatic control' (or system 2) when involving higher hierarchical levels that inspire more complex, effortful and future-oriented behavior that is aimed at securing long-term stability by cycling continuously through different strategies and activities (e.g. collective hunting, foraging, or farming) (Friston et al., 2018; Pezzulo et al., 2022; Evans, 2008; Seth, 2014).

To summarize, organisms have a mereological structure that allows them to capture the key properties of the larger mereology into which they are embedded and to act upon such models. The macroscopic environment of the organism acts as a latent state that contextualizes the dynamics of lower levels in such hierarchical models (bottom-up control). Similarly, the dynamics at higher levels serves as a latent state that contextualizes dynamics at lower levels (top-down control). By definition, content will change more quickly than context, so the higher levels of a control system should encode slower dynamics than lower hierarchical levels. This speaks to the separation of temporal scales across hierarchical levels that underlies the current thesis. Below, we will further explore the nature of hierarchical message passing in living systems and how such message passing may give rise to hierarchical dynamics (i.e. power-law frequency spectra).

5. Living systems as hierarchical *Bayesian* control systems that are engaged in active inference

The goal states of organisms can be compared to the setpoint of a thermostat, which encodes a room temperature that deviates from current reality, making it a model of a possible environment (Ha and Schmidhuber, 2018; Friston et al., 2021a). Mathematically, the encoding of possible worlds is equivalent to making predictive models of such worlds (Friston et al., 2021a). In this view, the setpoint of a thermostat encodes a prediction of what actual room temperature will be at some point within the future, provided the system will keep on running indefinitely (Kruglanski et al., 2020). The realization that living systems make predictive models of their environments has opened up a whole new field in biophysics known as 'active inference', which aims to
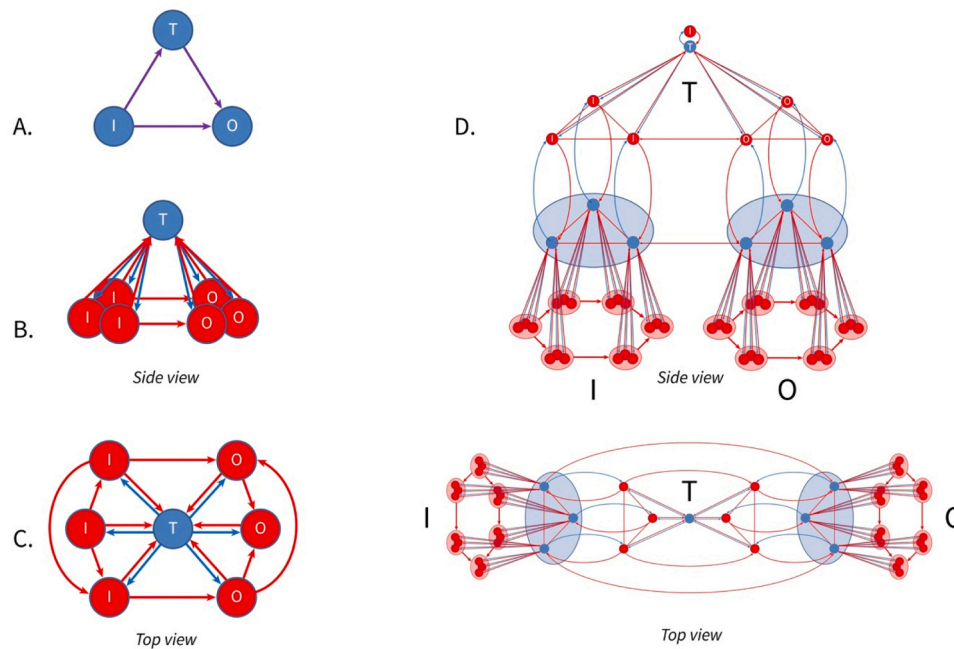
Fig. 4. Living systems as hierarchical control systems with a scale free information bottleneck structure. A. In systems biology, canonical microcircuits or 'network motifs' have been discovered within the signalling pathways of living systems that repeat across scales to produce a macroscopic network structure with a similar shape (Alon, 2007). This figure shows the 'feed forward loop' (FFL) motif: one of the most ubiquitous network motifs in living systems. It constitutes three nodes that are connected by arrows such that messages are passed 'forward' without any intrinsic feedback (reverse arrows) taking place (such feedback is provided by the environment of the motif). I: input, T: throughput, O: output. B. Collections of feed-forward loops may form three-dimensional sets of input nodes that converge onto smaller sets of throughput nodes (hubs), which then diverge onto larger sets of output nodes. Such structures have a basic hierarchical structure, where throughput nodes modulate the traffic that takes place between input and output parts. In many cases, 'top-down' (T = >I and T = >O) and inhibitory connections are thought to convey predictions (blue), whereas 'bottom-up' (I=>T and O=>T) and excitatory connections convey prediction errors (red, see next section). C. When viewed from the top, such structures display a so-called 'bowtie' or hourglass motif (Kitano, 2004; Kawakami et al., 2016). The input- and output parts form the 'wings' of the bowtie whereas the central throughput node (a hub structure) is called the 'knot' of the bowtie. Horizontal connections between input and output parts cause the bowtie to fold back onto itself. D. Folded bowties may be stacked next to and on top of each other to form an extensive nested modular hierarchy (see next section). The ensuing macrostructure again conforms to a feed-forward loop, or folded bowtie network motif. Such aggregation may repeat across several scales to form a hierarchy of part-whole relationships (a mereology). Only 1 scale-step is shown. Bowtie motifs control the flow of information between themselves and their environments, hence the term (hierarchical) 'control systems'.

produce a first-principle account of the structure and dynamics of living systems (Friston et al., 2006). According to this theory, the difference between the state of the world as predicted by a reference state and its current value is called a 'prediction error'. As in any control system, prediction error is used to initiate an action sequence that is aimed at altering the environment and (hence) the input to the system, which may reduce prediction error and bring the system closer its goals. In active inference, however, the same prediction error is used to update the predictive model itself (i.e., alter the reference state), allowing the organism to meet environmental conditions halfway. Organisms can therefore reduce prediction error in two fundamental ways: either by changing their environments through action (a process called niche construction (Constant et al., 2018)) or by changing themselves through model revision (a process called 'belief updating' or 'learning') (Friston, 2010). The combined use of action and model revision allows organisms to iteratively build better models of their worlds: a process called inference (Hohwy, 2016). This speaks to the active sensing or perceptual control literature (see above), although active inference puts more (cognitivist) emphasis on inference rather than perception per se.

According to information theory, reducing prediction error (improving model fitness) is to reduce a quantity called 'variational free energy' (see Box 1). This means that while optimizing their predictive models of the world, organisms are actually trying to descend upon a low-energy stable state (Friston et al., 2017). By doing so, they follow the second law of thermodynamics, which states that any system that is open to the in- and efflux of matter or energy must seek its lowest possible energy state (i.e. maximum stability) despite a continuous

influx of energy or matter. In this respect, organisms are not much different from rivers flowing downstream, ping-pong balls rolling into pits, or coins that start rolling on their sides to reduce friction when dropped to the ground. These systems are all compelled to seek their lowest possible (potential) energy states, exploring a variety of intermediate states and configurations in the process. Such configurations can be seen as 'models' that encode the state of the system's environment (Friston et al., 2006). Thus, information processing (hierarchical message passing and predictive coding) ultimately involves the dissipation of free energy across time (Friston, 2010, 2012; Ramstead et al., 2018), which is as much a physical imperative for living systems as it is for rivers to flow downhill (Box 1). Interestingly, the low-energy stable (predictable) state that organisms seek in this way is called 'homeostasis' or 'survival' in biology and the process of seeking stability through change is called 'allostasis'. Optimal information processing (building accurate models of the world) is therefore a prerequisite for survival in living systems (Hesp et al., 2019).

The equations that describe the process of active inference rely heavily on variational Bayesian statistics, which describe the constant updating of predictive models in the face of novel evidence (e.g. sensory input). At the heart of these equations is the free energy principle, which tells us that the best way to get rid of an excess of (variational) free energy is to make accurate yet parsimonious models of the world and vice versa (Box 1). To achieve this, organisms are thought to use a physical instantiation of hierarchical Bayesian inference (Ramírez and Marshall, 2017), which is important to the central argument of this paper i.e. that the (power-law) dynamics of living systems is produced

---

Box 1
Active Inference and the Free Energy Principle.

According to the second law of thermodynamics, energy or matter must flow in the direction where there is minimal constraint to its flow, i.e. *energy gradients must be overcome* (Jeffery et al., 2019)*. As long as its surroundings are cooler and/or less dense, all systems that are open to the in- and efflux of energy or matter will look as if they are ridding themselves of ('dissipate') an excess of energy or matter (e.g. rivers flowing downhill). The second law applies to any type of energy gradient that must be overcome, including 'free energy'. Free energy is a type of 'unbound' energy that is free to do work (i.e, to displace things across a distance). A system that tries to minimize its free energy therefore appears to seek a state of minimal flux or change (come to rest, e.g. a calm flow of water). In the process of dissipating (free) energy, the system adopts a spatiotemporal configuration that allows for optimal energy dissipation (e.g. the meandering paths of rivers through a mountain range). Such 'streamlining' makes it more predictable in space (self-similarity, e.g. river deltas) and/or in time (oscillations or permanence). The predictable configurations in space and time ('patterns') that thus emerge encode a 'model' of their environments with which the system is in an open exchange (its context; e.g. rivers modelling mountains, mountains modelling rivers)* (Friston et al., 2006; Roweis and Ghahramani, 1999; Friston, 2019a)*. Note that this local accumulation of predictability (order, negative entropy) is only possible as long as it contributes to a global increase in unpredictability (disorder, entropy) of the universe at large (i.e. the second law). The spontaneous emergence of (local) order is called 'self-organization'* (Kauffman, 1996)*.*

Under certain assumptions, the free energy of a (living) system is proportional to the (squared) prediction error of that system (i.e. the difference between perceived and expected states). Living systems abide by the second law by iteratively changing their environments (through action) or themselves (through model revision, 'learning') to adopt a spatiotemporal configuration (a predictive model) that is optimal in reducing prediction error (produce a model with optimal predictive accuracy). This process is called 'active inference' (Friston et al., 2006)*. The low-error, stable state that organisms thus achieve is called 'homeostasis' or 'survival' in biology and the process of seeking stability through change is known as 'allostasis'* (McEwen and Wingfield, 2003)*. Mathematically, active inference can be described as a gradient descent on free energy* (Friston et al., 2017)*. In information theory, the dissipation of free energy aligns with information processing (where information is negative entropy, entropy is the long-term average of surprise (as negative log probability) and variational free energy is used as a proxy for calculating surprise (it provides an upper bound) since the exact calculation of surprise is an intractable problem). In engineering, such gradient descents are used in Bayesian filtering schemes whereas in neuroscience and machine learning, they correspond to predictive coding* (Bastos et al., 2012)*. Interestingly, optimal dissipation of (variational) free energy is achieved by balancing model accuracy with model complexity (i.e. the number of variables used to explain the data), which yields a model with sufficient predictive accuracy yet parsimony (Ockham's Razor)* (Ramstead et al., 2018)*. This 'free energy principle' (FEP) applies to living as well as non-living systems* (Friston, 2012; Ramstead et al., 2018; Friston, 2019a)*. It predicts that hierarchical (sparse) structure (information bottlenecks, bowties, Markov Blankets) emerges spontaneously because of a need to optimize (free) energy dissipation, which has indeed been confirmed experimentally* (Jarman et al., 2017; Hellrigel et al., 2019; Annila and Kuismanen, 2009; Mäkelä and Annila, 2010; Rentzeperis et al., 2022)*. In this paper, we argue that the same principle also explains the emergence of hierarchical (sparse) dynamics,* i.e.*, power-law frequency spectra in living as well as non-living systems.*

The FEP describes a process of filtering out signals that accurately and parsimoniously predict events from signals that carry less reliable predictions (i.e., noise)*, as a precondition to remaining stable. In doing so, organisms are thought to use a physical implementation of variational Bayesian statistics that allows them to 'reason back' from a sequence of observed events (e.g., some sensory input) to the unobserved (hidden) causes of such events (i.e., the 'latent causal structure' of the observed effects). This process, called 'model inversion', allows organisms to peak through the veil of their imperfect sensory samples at the underlying events that probably caused them. The whole process of estimating the hidden causal structure behind a given input is called Bayesian inference (from Latin 'in-ferre', meaning 'to bring or carry (meaning) into'). As such, active inference can be read as a formal theory of enactivism* (Gallagher and Allen, 2018))*: a philosophical stance that emphasizes the embeddedness of living systems into a physical (body) and ecological environment (econiche) with which they intimately exchange signals. Such signals are intrinsically meaningless and organisms must actively project meaning into such signals in order to survive,* i.e. *signals must be identified as reliable predictors of events that will change their prediction error, i.e. signals that may ultimately affect their stability and survival.*

---

by their (power-law) network structure. To substantiate this claim, we will now briefly discuss a network structure that has previously been put forward as a consensus architecture that universally underwrites the process of active inference in living systems.

## 6. On the structure and dynamics of hierarchical Bayesian control systems

Fig. 5 shows a proposed consensus network architecture that explains the dynamics of living systems in terms of optimized variational free energy dissipation, or active inference. This model combines key findings from graph theory (e.g (Barabasi, 2009; Broido and Clauset, 2019)), systems biology (Friedlander et al., 2015; Kitano, 2004), machine learning (e.g (Gershman et al., 2015)) and hierarchical predictive coding as proposed by Karl Friston (Constant et al., 2021; Friston, 2019b). For a more detailed discussion of this structure and its dynamics under stress, see (Goekoop and de Kleijn, 2021b,a). For a mathematical model that approaches this topology, see (Friedman and Landsberg, 2013). For the equations of motion that describe the dynamics of such systems, see (Bastos et al., 2012). This model is still under development and may undergo adaptations in the future.

Fig. 5 shows a 'dual' hierarchical (nested-modular) bowtie structure with an input hierarchy (left wing of the bowtie, allowing for

perception) that smoothly transitions into a throughput hierarchy (information bottleneck or 'knot' of the bowtie, allowing for high-level goal setting) and an output hierarchy (right wing of the bowtie, allowing for action control). Prediction errors (variational free energy[2]) grossly flow from input via throughput to output hierarchy to eventually affect output organs, whereas predictions 'flow' in the opposite direction to bias perception. When following the base of the input-hierarchy to its top, network connections converge onto higher-level hub structures that integrate across multiple input streams (functional integration). This forces higher level structures to more parsimoniously encode multiple low-level events (e.g., round, green, sweet and juicy) as a single abstract

---

[2] *Technically speaking, under Gaussian assumptions about random fluctuations, the free energy of any node in a hierarchical generative model corresponds to the sum of squared prediction errors, weighted by their precision (summation is possible because free energy is an extensive property). Practically speaking this is a useful observation in machine learning schemes that use predictive coding. It means that free energy can be minimised in a local fashion; as in biotic self-organisation* (Millidge et al., 2022; Salvatori et al., 2022; Tschantz et al., 2022)*. Note that variational free energy in machine learning is also known as an evidence (lower) bound (ELBO)* (Winn et al., 2005)*. Interestingly, this evidence lower bound is the objective function used in variational autoencoders that have exactly the bowtie architecture illustrated in* Fig. 5 (Mescheder et al., 2017; Marino, 2022)*.*
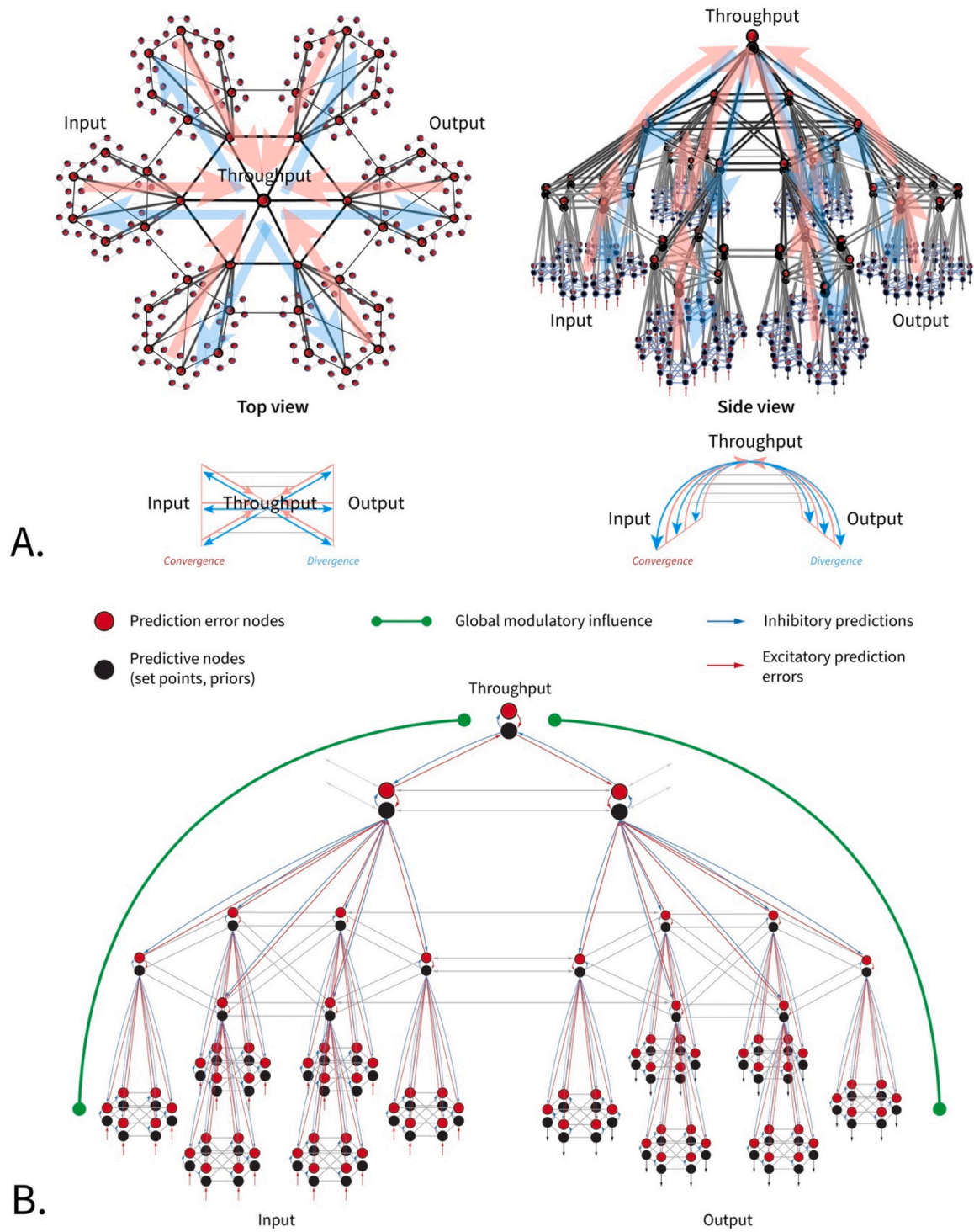
(caption on next page)

Fig. 5. A network structure that proposes a universal explanation of the dynamics of living systems in terms of optimizing variational free energy dissipation ('active inference'). *Overview showing a proposed consensus model of a generic network structure that underwrites the process of hierarchical message passing and predictive coding in living systems ('active inference'). A. The global shape of the object is that of a nested modular (dual hierarchical) and folded bowtie network structure. Note the global flow of prediction errors from the bottom of the input- (perception) and output- (executive functioning, action control) hierarchies to the top of the structure (throughput hierarchy; goal setting). Also, note the global flow of predictions in the reverse direction, causing perception to be biased by top-down predictions of what will be perceived (selective attention) and output to constitute predictions of movements rather than output commands. These predictions are eventually 'fulfilled' by prediction error units that connect to output organs to correct any deviations from the predicted path (these units are not shown in this figure, but see text). Horizontal connections: skip-connections that bypass higher-level processing and cause the bowtie structure to fold back onto itself. These short action-perception loops underwrite automatic responses, the level of complexity of which is a function of hierarchical depth: short cycles that bypass higher level processing correspond to simple reflex arcs, whereas progressively longer loops that visit deeper levels correspond to instinct patterns, habitual behavior, and goal-directed behavior, respectively. B. Excerpt of Figure A, showing nodes and connections in more detail. Black spheres: prediction units (priors, setpoints, predictive models). Red spheres: prediction error units, encoding deviations of some prior value relative to available evidence. Black arrows: inhibitive connections from priors to prediction errors (explaining away evidence). Red arrows: excitatory connections from prediction errors to priors (belief updating). Note the hierarchy of black nodes that encodes a deep (i.e., multilayered) hierarchical predictive model of the inner and outer econiche of the organism (i.e. a world model). Lower levels converge onto higher levels that encode the environment at increasing levels of contextual integration yet parsimony (abstraction). Deeper levels also encode progressively slower (stable) aspects of events in the outside world (the vertical encoding of timescale, see text). A hierarchy of red nodes encodes deviations of such hierarchical world models wrt. available evidence (i.e. unexplained evidence). Any unexplained evidence is simultaneously projected upward in the input hierarchy as a residual prediction error that is to be suppressed (inhibited, explained away) by a more sophisticated model of the world (black nodes, see text), as well as relayed to the output (action control) hierarchy to produce an output sequence at a matching level of sophistication ('adaptive behavior'). Thus, 'gradients' of prediction error are actively overcome, which amounts to variational free energy dissipation (active inference). Green arcs: global modulatory influences (e.g. neuromodulatory neurotransmitters) that control the overall precision (signal-to-noise ratio) of bottom-up versus top-down signalling by modulating the gain of self-loops (not shown, but see next figure). See text for further details .*

event (e.g., 'apple'). As observed in the previous section, successive hierarchical levels encode aspects of the environment that are progressively distant from current reality, with the top of the regulatory hierarchy (the knot of the bowtie) encoding the most sophisticated (contextualized, long-term) goal states of the organism. The reverse happens when moving from the knot of the bowtie towards the base of the output hierarchy. Here, connections diverge from high-level hub nodes across subordinate nodes, allowing for a 'decompression' of long-term and abstract goal states into a hierarchical succession of goals and corresponding subgoals to eventually produce detailed 'output commands' that connect to output organs to produce behavior (Pezzulo et al., 2015; Adams et al., 2013). Note that Fig. 5 shows prediction errors flowing upward and predictions downward in both input and output hierarchies (which appears to be the current consensus), causing output signals to be 'predictions rather than commands' (Adams et al., 2013). Such predictions are only corrected at the lowest possible levels by sensory (proprioceptive) input from the output organs, i.e. predictions are compared to actual movements to produce a prediction error that forms the actual output command. Thus, instead of predictions (black) directly connecting to output organs (e.g. muscles), the actual output signal is given by prediction error units that correct any deviations with respect to predicted output, making it look as if they *fulfill* predictions or 'see them done' (Adams et al., 2013).

The bowtie is folded because of 'skip connections' that run horizontally between same-level nodes within the input and output hierarchy, creating shortcuts. These ensure that input signals can skip higher level processing to produce automatic and faster, more energy efficient (but less well-informed) responses. The level of complexity of such action-perception cycles is a function of hierarchical depth: low-level shortcuts correspond to simple reflex arcs whereas progressively deeper loops correspond to instinct patterns (e.g., fight-flight responses), habits and goal-directed behavior, respectively (Pezzulo et al., 2015). The horizontal exchange of signals allows input hierarchies to bias output hierarchies in the direction of actions that are preferred (predicted) under the current states. Conversely, output hierarchies may bias input hierarchies in the direction of percepts that are predicted under the current intended action (e.g., corollary discharge, selective attention). This ensures that organisms pay attention to perceptual cues that are relevant to the planned behavior (Kanai et al., 2015; Novák and Tyson, 2008). An increase in the precision (inverse variance, newsworthiness) of signals that predict a positive perceptual outcome (a drop in perceptual prediction error) under some current action has been associated with increased motivation (Pezzulo et al., 2018). The precision of top-down versus bottom-up signalling can be tuned by global modulatory influences (see below) (Gilbert and Li, 2013; Kanai et al., 2015). This may shift the balance between top-down generative

modeling (i.e., generating events from presumed causes) and bottom-up predictive coding (i.e. inferring causes from perceived events) in either favor, depending on the circumstances. Task-negative circumstances favor top-down generative processing (a resting state), whereas task-positive circumstances favor bottom-up predictive coding and belief updating (an active state) (Bastos et al., 2012).

In Fig. 5, a hierarchy of priors (black nodes, setpoints) can be observed that encodes a hierarchical predictive (generative) model of the current world that smoothly transitions into a goal hierarchy. A hierarchy of prediction error units (red nodes) can be observed next to it that encodes the deviation of predictive models with respect to the available evidence. Prediction errors keep ascending in the input hierarchy until they are sufficiently suppressed (explained away) by a set of higher-level priors. Thus, perceptive models automatically obtain a hierarchical depth and corresponding breadth (a level of sophistication) that optimally matches the available evidence from the outside world. Any residual prediction error then crosses over to the output hierarchy via horizontal skip-connections to activate a cascade of events further down the output hierarchy ('action control') that results in an action sequence (behavior) that may bring the organisms closer to its goals. When sensory input becomes more complex, more prediction error is produced that engages ever higher levels of the goal hierarchy to initiate ever more complex (and long-term) executive plans. As a result, the level of sophistication at which action plans are initiated in the output hierarchy automatically matches that of the perceptual model and corresponding goal state (Pezzulo et al., 2015). This correspondence of sophistication between perception, goal setting and action control ensures that organism show behavior that matches actual environmental demands, i.e. that organisms produce *'adaptive behavior'* (Goekoop and De Kleijn, 2021). The selected behavioral policy then changes the environment, which changes the input to the system, after which the cycle repeats. Thus, much like the Baron of Münchhausen who pulled himself out of a swamp by his own hair, living systems save themselves from drowning in a flood of prediction error through a circular process of action and model formation (active inference). Macroscopically, this corresponds to organisms exploring and exploiting their optimal ecological niches (Constant et al., 2019, 2018).

7. Explaining power-law frequency distributions in the timeseries of living systems

To explain power-law frequency distributions (i.e. system dynamics) from the workings of hierarchical Bayesian control systems as shown in Fig. 5 (i.e., system structure), at least four things should be addressed: first, we must explain the process that produces oscillations at different frequencies. Next, we must explain the conditional dependencies that

exist between these frequencies as observed in nature, i.e., nested-modular frequency distributions and cross-frequency coupling. Thirdly, we have to explain the inverse relationship between frequency and amplitude in power-law frequency distributions (i.e., why no specific intermediate frequencies dominate the power-law during the resting state). Finally, we must explain the flattening of the slope of the power-law curve ('bumps') during task performance and high levels of stress. These points will be addressed below.

7.1. Explaining the origin of oscillations at different frequencies in the timeseries of living systems

In the active inference literature, oscillations arise as a result of circularly causal relationships between priors (setpoints, models) and prediction error units (evidence), where priors inhibit prediction errors ('explain away evidence') and prediction errors excite priors ('perform belief updating') as shown in Fig. 5 (Constant et al., 2021; Friston, 2019b). Although knowledge on oscillatory dynamics in living systems in many cases derives from neurodynamics, similar oscillations have been observed in non-neural network systems at various scale levels of observation (Friston, 2012), such as genomes, proteomes and metabolomes (Novák and Tyson, 2008), calcium waves in networks of pancreatic cells and islets of Langerhans (Zmazek et al., 2021; Stožer et al., 2022), social networks and the internet (Nekovee et al., 2007). Throughout the years, many explanations have been proposed for the emergence of such oscillations and the various frequencies they display (Novák and Tyson, 2008). More recently, evidence converges on a key role for hierarchical structure. This will be explained below.

As observed above, hierarchical (nested modular) structure itself may develop spontaneously as a result of the optimization of energy dissipation (Kirchhoff et al., 2018; Jarman et al., 2017; Hellrigel et al., 2019; Annila and Kuismanen, 2009; Mäkelä and Annila, 2010; Rentzeperis et al., 2022), Interestingly, simulation studies show that hierarchical dynamics may emerge from the same principle. Non-randomly connected (hierarchical) network systems spontaneously engage in a 'vertical encoding of timescale' from the sole requirement of optimized energy dissipation (Paine and Tani, 2005). This means that timescale of oscillations is encoded vertically as gradients across the various hierarchical levels of the system, with high frequencies being encoded at lower levels and low frequencies at higher levels of the hierarchy. By now, a large body of empirical studies has confirmed that timescale is distributed vertically as gradients across the brains of many different species, including nematodes, rodents and primates, revealing it as a general organizing principle of brain function (Mahjoory et al., 2020; Atasoy et al., 2018, 2016; Raut et al., 2020; Piasini et al., 2021; Henin et al., 2021; Murray et al., 2014). In humans, mesocortical structures form an information bottleneck that encodes slower dynamics than its wings (the sensorimotor cortices) (Huang et al., 2017). A more fine-grained mapping of such gradients at successive scales using functional magnetic resonance imaging shows a frequency distribution with a power-law scaling exponent of 1.14 (Friston et al., 2021b), i.e., when ascending hierarchical (and mereological) scale levels, dynamics becomes consistently slower. More recent studies show asymmetries in the frequency between ascending and descending neuronal messages (Bosman et al., 2012; Bastos et al., 2015; Giraud and Poeppel, 2012; Hovsepyan et al., 2018), suggesting that neuronal coupling from low to high levels or scales may use faster (e.g., gamma) frequencies than the corresponding reciprocal messages from higher to lower levels (sometimes expressed at slower beta frequencies). This kind of separation of temporal scales at a microscopic scale fits comfortably with the vertical encoding of timescale observed macroscopically in brain imaging experiments and associated models as described e.g., in Kiebel et al. (2008); Murray et al. (2014); Wang and Kennedy (2016); Hasson et al. (2008). A hierarchical encoding of timescale has further been observed in molecular networks (Henzler-Wildman et al., 2007) as well as decision making processes in human organizations (Purcell and Kiani,

2016), suggesting it is a scale free phenomenon. To explain it, we have to take a closer look at the 'equations of motion' that describe hierarchical message passing in nested modular systems such as shown in Fig. 5 (Kiebel et al., 2008; Bastos et al., 2012; Friston and Kiebel, 2009).

The dynamics of hierarchical message passing (predictive coding) in structures like Fig. 5 can best be described as a hierarchy of attractor systems, where the state-space trajectories of higher level attractors shape the trajectories of attractor system at lower levels (Fig. 6). From Figs. 5 and 6, it can be observed that bottom-up hierarchical message passing in Bayesian control systems involves prediction errors rather than raw input signals themselves. Each subsequent hierarchical level therefore encodes deviations relative to expected deviations (i.e., prediction errors relative to predicted prediction errors). Mathematically, deviations that change in time can be expressed as derivatives, which indicate the amount of change of a certain variable across a small time-interval. When derivatives are hierarchically organized with respect to each other, each derivative encodes a different aspect of the environment. For instance: a change in location indicates speed and a change in speed indicates acceleration. Thus, higher hierarchical levels predict 'changes in changes', which tend to involve increasingly slow (rare) events in a spatiotemporally structured environment. Intuitively, this may produce slower dynamics at higher levels. However, this explanation focuses on bottom-up signalling pathways only and disregards top-down influences. Mathematical models of hierarchical predictive coding show that ascending prediction errors drive the dynamics at higher levels in a linear fashion and contribute little (if at all) to the emergence of oscillations at different frequencies. Rather, the separation of timescales inherits from descending, predictive signals (Friston and Kiebel, 2009; Kiebel et al., 2009). One reason for this is that higher-level hub units integrate across many prediction errors, causing higher frequencies to cancel each other out (a form of temporal smoothing) (Bastos et al., 2012). Additionally, higher-level predictive hub units modulate the effects that multiple lower level (prediction error) units have onto each other. Thus, higher hierarchical levels provide the 'context' or 'control parameters' for 'content' (dynamics) that unfolds at the level below. Predictive coding formulations show that such 'modulatory' effects are essentially non-linear in nature and produce higher frequencies at lower levels (Smit et al., 2011;Tagliazucchi et al., 2013). This can be intuitively understood as a form of 'period doubling' (i.e., squaring a sinewave doubles its frequency), causing a separation of time scales across hierarchical levels. Thus, a gradient is formed where lower hierarchical levels encode higher frequencies (treble) and higher hierarchical levels encode lower frequencies (bass). This cumulative suppression of higher frequencies as a function of hierarchical level can be regarded as a (nonlinear) form of low-pass (Bayesian) filtering, to produces 'smooth' estimates of hidden causes at increasingly higher levels (Kiebel et al., 2008; Knill and Pouget, 2004). Empirical studies show that this mechanism drives things like theta-gamma coupling in the hippocampus and is in line with the functional asymmetries in ascending and descending messages in neuronal hierarchies (Shipp, 2005; Sherman and Guillery, 2011). The same principle is used in computational neuroscience and machine learning, e.g., for temporal memory implementations and deep temporal models for active inference (Friston et al., 2018; George and Hawkins, 2009).

In engineering and cybernetics, the vertical encoding of timescale relates to the good regulator theorem (or the law of requisite variety), which states that to control certain aspects of the environment, a system must encode models that are at least as sophisticated as the aspects of the environment it tries to control (Conant and Ross Ashby, 1970). In other words, in order to be an effective controller, a control system must encode the deep spatiotemporal structure of its environment and act upon such models. The state-trait continua that emerge from hierarchical Bayesian controls systems such as Fig. 5 essentially mirror the complex and nonlinear hidden causal structure of world in which the system is embedded. Thus, nested modular and hierarchical structure allows for more sophisticated world models and a progressive finetuning
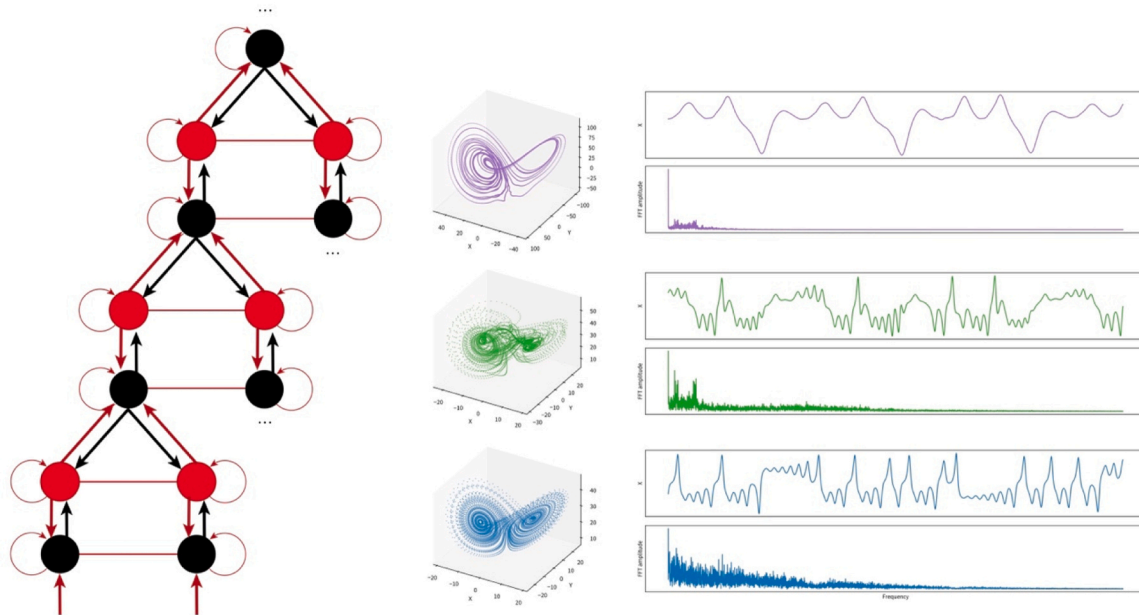
Fig. 6. The vertical encoding of timescale: why each level within a hierarchical (Bayesian) control system encodes its own characteristic frequency. *This figure shows the dynamics that emerge at different hierarchical levels within a hierarchical Bayesian control system such as shown in* Fig. 5 *(after* (Friston and Kiebel, 2009)*). Left images show a closeup of the input hierarchy in* Fig. 5 *that consists of three levels. Each level consists of a subnetwork of coupled nodes that generate oscillatory dynamics. Black circle: prior node, Red circle: prediction error node. Red arrows: excitatory connections, black arrows: inhibitory connections. Middle images show the network dynamics at each level as represented by a state space with coordinate axes that represent the degrees of freedom of the system (e.g., number of nodes or clusters). State space trajectories (purple, green, blue) represent the timeseries of the system in three dimensions (2-dimensional representations of these timeseries are shown on the right, with their frequency distributions below each timeseries). As can be seen in the figure, each state-space trajectory defines two volumes of state space around which the system orbits, which are called 'attractors'. Like the subnetworks that produce them, the dynamics at each level are hierarchically coupled: the dynamics of an attractor network at some level acts as a control parameter that is passed as a prediction to the level below it. This influences its dynamics in a nonlinear fashion, producing faster (higher frequency) dynamics at lower hierarchical levels (see text). Thus, each hierarchical level produces dynamics a characteristic dominant frequency (a carrier wave) that contextualizes dynamics at lower levels in a non-linear way. As a result, a vertically organized gradient of frequencies emerges, with traits being produced at the top and states at the base of the hierarchy (see frequency spectra at each level). The lowest level incorporates all the frequencies (control parameters) of the levels above it (see the Fourier spectrum). Another way to refer to such a frequency spectrum is a 'state-trait continuum'. In input hierarchies, such dynamics is used to bias sensory processing in the direction of expected causes (i.e. attentional biasing). In output hierarchies, such dynamics is used to generate complex sequences of reflexes (i.e. behavior). Note that the position of each attractor shows relatively little changes in time, making it an ultra-low frequency 'trait' relative to the faster dynamics of the trajectory itself. At a larger timescale, however, the positions of attractors do vary, since they are influenced by the dynamics that unfolds at other levels. This causes traits to show ultra-slow dynamics in time (e.g. personality development). The attractor states in this figure have been produced using code for coupled Lorenz attractors provided in* Quinn et al. (2020).

of a system's behavior, which involves anticipating increasingly abstract events and seeking stability (survival) through constant pre-emptive action (i.e., allostasis). Overall, this means that hierarchical structure conveys an important reproductive advantage.

Due to the hierarchical encoding of spatiotemporal structure, each level in a nested modular network structure has its own characteristic frequency. In most natural systems, hierarchical levels are not as neatly stratified as shown in Fig. 5, in which each scale level can be assigned a discrete number (scalar dimensions). Rather, hierarchical scale levels smoothly dissolve into each other, causing a continuous or 'broken' dimensional scale (hence the term fractal). This may explain why Fourier analyses of the timeseries of the inner message passing and overt behavior of living systems tend to produce a smear of frequencies (i.e., frequency *spectrum*) rather than a set of discrete spikes. In the next section, we will examine the nested modular relationships between the various frequencies that are produced in hierarchical (Bayesian) control systems.

7.2. Explaining nested modular frequency distributions (with phase-amplitude coupling) in the timeseries of living systems

The frequency components of the timeseries of living systems are far from independent. A burgeoning neuroimaging literature highlights the importance of cross-frequency coupling in coordinating brain activity across spatial and temporal scales (Canolty and Knight, 2010). In many cases, the phase of the lower frequencies controls the amplitude of the

higher frequency components in the human brain (Hyafil et al., 2015; Tort et al., 2010). The exact mechanism by which such phase-amplitude coupling occurs remains unclear, however, although our theory agrees most with a previously posited explanation stating that lower frequencies modulate the gain (precision) of circularly causal loops between excitatory and inhibitory units (Fig. 6) (Canolty and Knight, 2010; Hyafil et al., 2015; Onslow et al., 2014) in ways comparable to selective attention (Sadaghiani et al., 2010). Simulations show that this causes lower frequencies to shape the trajectories of higher frequencies, to automatically produce a phase-locking between slower and higher frequencies (Canolty and Knight, 2010; Hyafil et al., 2015; Onslow et al., 2014). Thus, low-frequency oscillations may initiate high-frequency 'bursts' and vice versa. Another feature of cross-frequency dependencies is that higher frequencies are iteratively embedded within lower frequencies. In our view, such nested modulatory in time can be explained by the nested modularity in space of the systems that produce the various frequencies: since larger clusters can only exist conditionally upon the presence of their constituent smaller subclusters, their activity must be conditionally dependent as well, producing nested frequencies (Fig. 7). Simulations studies show that such mereological dependencies indeed produce nested modular (power-law) dynamics (Wang et al., 2011), even in the absence of criticality (Friedman and Landsberg, 2013). Thus, both hierarchical depth (of vertically stacked systems such as shown in Fig. 6) and hierarchical breadth (mereological relationships) contribute to cross-frequency dependencies.

To summarize, each level within a nested modular hierarchy has its
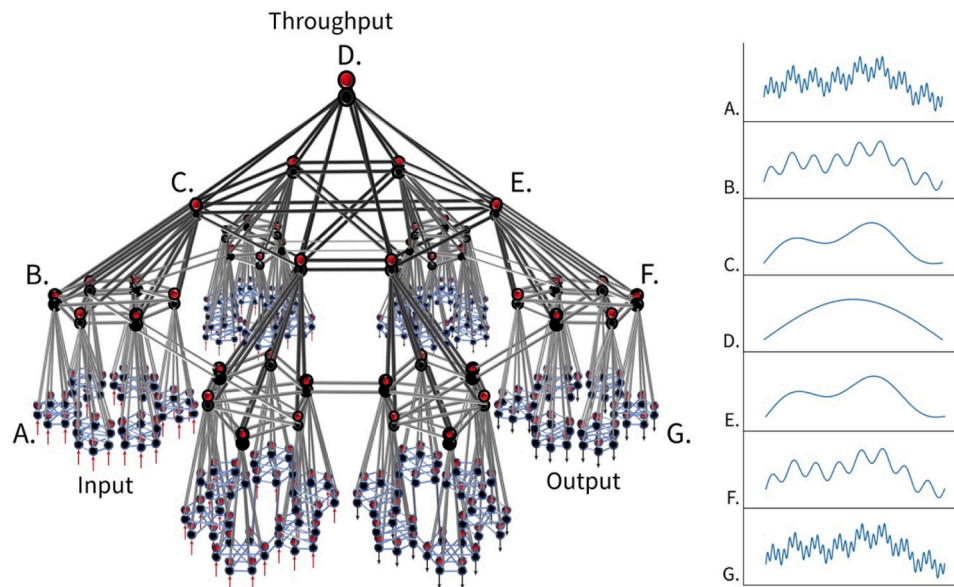
Fig. 7. **How hierarchical structure produces hierarchical dynamics: power-law frequency spectra reflect hierarchical message passing in nested modular systems.** *In nested modular systems, each hierarchical level produces its own characteristic frequency (see* Fig. 6*). Cross-frequency coupling therefore represents the communication between hierarchical levels (top-down versus bottom-up control). Because of their mereological dependencies (part-whole relationships), amplitude changes at level A (the signal) depend on amplitude and phase changes at level B (the carrier wave) and so on. Thus, higher frequencies become nested within lower frequencies, explaining the nested modularity of the frequency components of power-law frequency spectra. Only cross-frequency dependencies of amplitudes are shown. Cross-frequency phase-amplitude modulation is omitted in this figure for visualization purposes. See text for further details.*

own characteristic frequency, with higher levels producing lower frequencies and lower levels producing higher frequencies (Fig. 7). Cross-frequency coupling should therefore reflect the coupling between different levels in a hierarchical control system. This can be read as a definition of top-down and bottom-up control (Liu et al., 2012): like generals forming a grand overview of the battlefield based on which they issue out highly global orders, a small number of hub nodes and clusters (at the top of a bowtie) encode global goal states that manifest as low-frequency signals. Such signals subsequently control vast numbers of lower-level signals through a (nonlinear) cascade of interlocked frequencies until they shape the trajectories of action-perception cycles at the base of the hierarchy (the wings of the bowtie), defining the process of top-down control (e.g., action control or perception control/selective attention). Higher hierarchical levels essentially impose their tonicity onto the output of their subordinate systems, so that the timeseries of low-level action-perception cycles (overt behavior) incorporate the frequencies of all superordinate levels that were involved in controlling the output of the organism. This would explain the universal presence of power-law dynamics in the behavioral timeseries of living systems, as discussed above (Kaplan et al., 2020). A similar (linear) process may define bottom-up control, in which high-frequency signals at lower levels shape the trajectories of low-frequency signals within higher regions, reflecting the process of belief updating in the face of novel evidence (i.e., novel percepts). Cross-frequency coupling should therefore have different (causal) directions depending on whether one examines the input/perception, output/action or throughout/goal-setting parts of a control system. This is all but confirmed by studies that show that bottom-up signalling involves high-frequency signals relative to top-down signals, which involve lower frequencies (Bosman et al., 2012; Bastos et al., 2015; Giraud and Poeppel, 2012; Hovsepyan et al., 2018). This is also in line with results from numerical as well as simulation studies based on empirical data, showing that high-degree (high-level) hub nodes in biologically plausible networks may both phase-lag and phase-lead signal changes in lower degree nodes, which may indicate a balance between top-down and bottom-up control (Woo et al., 2020). A long history of psychometrics further shows that power-laws-govern human inner experience just as well as overt behavior: state-trait

continua have been observed in the timeseries of nearly every aspect of the human mental phenotype, varying from attention levels global experiences such as self-image or theories of mind, e.g., (Marsh and Shavelson, 1985; Kotov et al., 2021). We expect that the hierarchical network structure of the human brain explains the state-trait continua of the human mental phenotype in the same way as they do for overt behavior (Figs. 6 and 7 and see below).

In a seminal paper on the subject, He et al. consider the possibility that the power-law (broadband) signal results from the summation of many narrowband oscillations, precisely as we propose, but dismiss it as 'near magic', stating that 'overwhelming evidence' now suggests that there are at least two distinct phenomena in the brain, which are the 1/f power-law activity and the rhythmic oscillations often studied using EEG or MEG recordings which show up as bumps on the power-law curve (He, 2014). We believe that this view rests upon a notion of the brain prior to the full recognition of its nested modular topology and precise knowledge of the predictive coding and active inference framework as discussed above, which ultimately concerns the optimized dissipation of variational free energy across the full spectrum of frequency bands. Also, the bumps seem to have another point of origin. These issues will be discussed below.

7.3. Explaining the inverse amplitude-to-frequency relationship in the timeseries of living systems

The smooth run-off of amplitude with increasing frequency that can be observed in power law frequency spectra can be explained from the perspective of energy dissipation in networks of coupled oscillators (see Box 1). As shown above, living systems can be abstracted as networks of coupled oscillators that are in the act of dissipating (metabolic or free) energy, which is what defines information processing. The energy that is required to produce oscillations is supplied by metabolic pathways that ultimately require glucose and ATP (Lane, 2015). During hierarchical message passing, sensory prediction errors (variational free energy) drive nodes that use metabolic energy to update internal states and drive the output of the system (active inference; see above). Information processing in living systems can thus be understood as a process by

which metabolic energy dissipation (oscillations) is put in the service of variational free energy dissipation (prediction error reduction). In non-living systems, the metabolic component is lacking and nodes simply pass on the energy from their surroundings to other nodes within their network structure until it leaves the system.

Empirical studies as well as computer simulations show that highly connected (hub) nodes oscillate at higher amplitudes (Moon et al., 2015), although some studies find more variability (Woo et al., 2020). This observation fits well with our prediction that hierarchical (nested modular) network structure underwrites the power-law dynamics of living systems. In *small-world* networks, the most strongly connected nodes (hubs) reside at the top of the hierarchy (the knot of the bowtie). Coupling strength (weighted node degree) then diminishes in a power-law fashion when moving down the hierarchy. In such networks, strongly coupled (constrained) oscillators are 'rigid' oscillators, i.e. they have a faster relaxation rates than loosely coupled nodes, which means they show a quick recovery after perturbation (Bordyugov et al., 2011; Marković et al., 2014). Relaxation rate is directly related to the dissipation of (metabolic and / or free) energy (Gosak et al., 2015). Thus, hub units dissipate (and require) large amounts of energy, but are relatively few in number, whereas peripheral nodes each emit (and require) less amounts of energy but come in larger numbers. Assuming that a comparable amount of energy is required and emitted by a small number of hub nodes with low frequencies as by a large number of (peripheral) nodes with fixed high frequencies, this should cause hubs to oscillate at higher amplitudes than peripheral nodes. In other words, 1/f noise can be explained by the necessity of open dissipative systems to distribute energy dissipation equally across its constituent coupled oscillators, hierarchical levels and corresponding frequency bands. As a result, amplitudes should fall in a power-law manner when moving from the top (low frequencies) to the bottom (high frequencies) of the hierarchy, producing an inverse amplitude-to-frequency relationship. In other words, the non-egalitarian coupling of nodes and modules in *small world* networks imposes a non-randomness of amplitudes onto their frequency spectra. We believe this to be the first time that the universal runoff of amplitude with frequency in power-law frequency spectra is explained in terms of the non-randomness of coupling strength and energy dissipation of oscillators in nested modular (scale free) networks and, hence, hierarchical network structure (see Fig. 1).

This proposal is in line with seminal work done by Per Bak, who showed that the size of cascading events in simulated sandpiles displays a 1/f frequency distribution (Bak et al., 1988). Whereas Per Bak concluded that a self-organized critical state is responsible for producing power-law scaling and even doubted the generality of his models' applicability, it seems that this model can nonetheless be generalized by explaining such cascades (e.g., avalanches in grains of sand, neural avalanches or snow avalanches) as special cases of granular systems involved in the optimization of (free) energy dissipation. This idea receives support from a recently published universal model of earthquake statistics, which uses a simple mean field model to define the coupling between granula (Uhl et al., 2015). According to this model, the earth's crust can be modelled as a network of coupled oscillators of different sizes, varying from (a few) large chunks of earth to a large number of smaller rocks and grains of sand that use vibration and friction to dissipate kinetic energy in the form of heat. The heavier chunks of rock act as (mass-constrained) 'rigid oscillators' that vibrate at lower frequencies and higher amplitudes than the lighter granules and grains of sand, producing 1/f noise. This model is consistent with neurobiological findings that show that hierarchical structure promotes self-sustained 1/f noise by interconnecting modules that display self-organized criticality (Wang et al., 2011). In this view, the brain is a nested modular system in which each node or module is like a (metabolically activated) granule of a particular size that can be excited by internal or external energy to produce scale-free (avalanche) dynamics at a global level. Thus, a single (mean field) model can be used to explain the dynamics of living as well as non-living systems. Interestingly, mean field models

connect directly (through the Bogoliubov inequality) to the equations of motion that describe the process of active inference (Kuzemsky, 2015). It therefore seems that efficient energy dissipation ultimately drives the dynamics of living as well as non-living systems. We believe this is the first time that a single physical explanation is given for power-law (1/f) dynamics of living as well as non-living systems.

Previously, it has been proposed that (changes in) power-law frequency spectra represent (shifts in) the excitation/inhibition balance in the human brain (Gao et al., 2017). This is in line with another proposal that power-law frequency distributions in EEG signals result from fluctuations in the precision of prediction errors and predictive signals (Sadaghiani et al., 2010). In this view, precision is encoded by the amplitude-to-noise ratio of such signals, which is modulated at the synaptic level by neuromodulatory neurotransmitter systems that affect the gain of self-loops of prediction error (excitatory) and predictive (inhibitory) units (Figs. 5, 6). In our view, shifts in the global excitation-inhibition balance may certainly affect power-law dynamics, but only to the extent that such shifts alter the global balance between top-down (inhibitory, predictive) and bottom up (excitatory, corrective) processing and, therefore, the overall hierarchical depth or level of sophistication of information processing. In other words, we consider fluctuations in the excitation/inhibition balance in themselves to be insufficient to explain the full phenomenology of power-law frequency spectra, which in our view requires the idea of hierarchical structure.

In living systems, most metabolic energy is consumed by a small number of hub structures at the top of a regulatory hierarchy (so called rich clubs (Colizza et al., 2006; Griffa and Van den Heuvel, 2022), which encodes highly compressed world models (i.e. long-term predictive models of abstract events). This means that producing such models is allowed at the expense of a metabolic penalty, putting an upper bound on the level of compression or sophistication an organism can achieve when encoding its environment. The 'costs of compression' can be read from the total amount of spectral energy involved in encoding some aspect of the world, which is given by the area under the power-law curve E. Theoretically, the area right of any vertical line $f = x_{(m-n)}$ (where f and x are frequencies, f equals an index frequency, x_m is the lowest frequency and x_n the highest frequency of the signal, and $x_m < f < x_n$) is a measure of the total amount of spectral energy that is required to encode some aspect of the world up to the hierarchical level that is given by the index frequency. The area left of this line represents the amount of energy required to construct even more abstract or compressed models. The ratio between these two areas is a measure of the metabolic costs per time unit that an organism is prepared to spend on making more sophisticated models of the world than those given by the index level. The well-known deviations of naturally obtained frequency spectra from perfect power-law curves that occur near their asymptotes (Broido and Clauset, 2019) can be explained by the fact living systems are not perfectly scale free hierarchies: their scale levels have a lower bound at the level of molecules and an upper bound at the level of organisms or (eco)systems. This prevents a scenario where infinite amounts of metabolic energy are required by the apex of a hierarchy to compress infinite amounts of uncompressed information encoded at its base.

In summary, we propose that the steepness of the power-law curve conveys information about the hierarchical depth of the generative models that a system encodes: steep curves signal deeper hierarchies whereas flatter curves indicate flatter hierarchies. This means that power-law frequency spectra should change as a function of the hierarchical depth of a generative model as observed e.g. during development, task performance, stress, and physical or mental illness. This will be discussed below.

7.4. Explaining changes in power-law frequency spectra during development, task performance, stress and disease (bumps and tipping points)

In task-negative conditions, intrinsically generated stochastic noise produces prediction errors that are escalated upwards in the hierarchy to be suppressed by predictive models. This occurs along the full depth of the regulatory hierarchy, producing a (day)dreamlike state that reflects the free exploration of the network's state space in the absence of external stimuli (Wang et al., 2011; Deco et al., 2011). In our view, this 'resting state' explains the emergence of the full breadth of the power-law frequency spectrum, as well as the characteristic non-dominance of any intermediate component. During task performance, however, prediction error is projected upwards in the hierarchy to a hierarchical depth and breadth that is optimal in solving that task (see above and Fig. 5). Thus, task performance usually recruits an intermediate hierarchical level that lies somewhere in between the base of the hierarchy (minimal effort) and its top (peak performance). Since each hierarchical level has its own characteristic frequency, a switch from a resting state to task performance should therefore increase the amplitudes of the frequency components that correspond to the hierarchical levels that are involved in solving that particular task (this is because external stimuli (task performance) produce stronger

oscillations than intrinsic noise (resting state)). This may explain why power-law frequency distributions develop 'bumps' at intermediate frequencies during task performance (e.g., alpha, beta, or gamma bands in EEG (He et al., 2010)) (Fig. 3). The amplified frequencies that make up these bumps should reflect the hierarchical depth of the chosen model and, therefore, task difficulty.

In this view, power-law frequency spectra should covary with task difficulty: simple tasks that require tax only the base of a regulatory hierarchy (i.e. reflexes or habits) should produce bumps at higher frequencies, whereas increasingly demanding tasks should increasingly recruit lower frequencies, leading to a more global flattening of the power-law curve (Fig. 8). This has been confirmed experimentally by examining the effect of varying task loads on the steepness of power-law curves (Tinker and Velazquez, 2014; Ciuciu et al., 2014). fMRI studies show that more difficult tasks recruit lower frequencies in brain areas that are involved in interregional communication (i.e., high-level hub structures) (Ciuciu et al., 2014). This also relates to the band-filtered EEG literature on the theta-beta ratio, which shows a shift towards the slower theta component when tasks become more demanding (Puma et al., 2018; Clarke et al., 2019). Even bacteria respond to varying growth conditions primarily by shifts in low-frequency components within the timeseries of gene expression, which is indicative of high-level regulatory changes (Almaas et al., 2004). An increase in bass
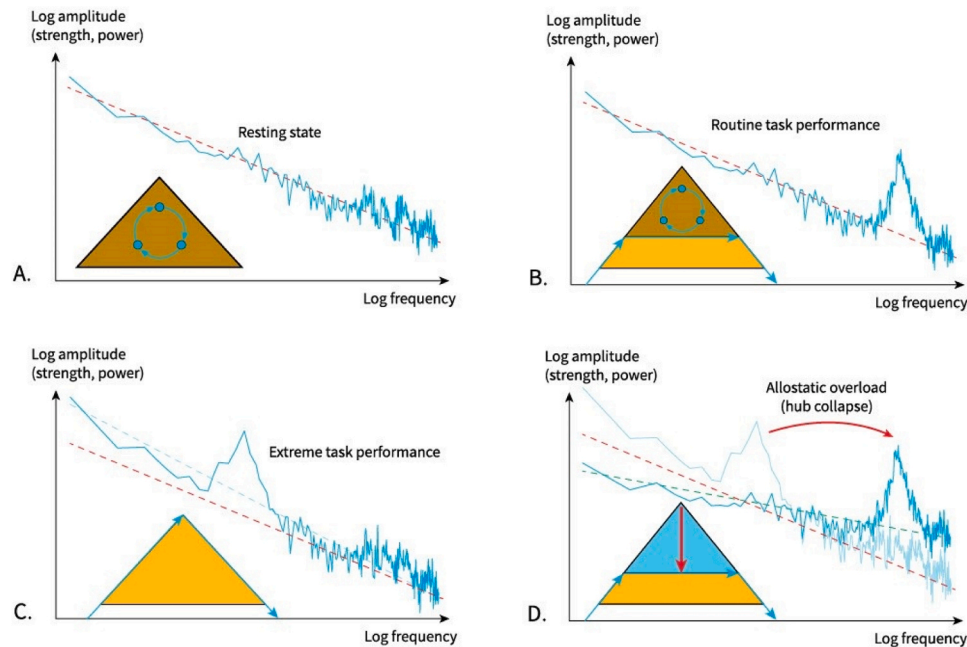
Fig. 8. Changes in power-law frequency distributions as a function of task performance (levels of stress). *This figure shows EEG scalp recordings of individuals engaged in increasingly demanding tasks (A-D). Triangles symbolize a hierarchical control system such as Fig. 5. Arrows indicate the presence of action-perception cycles, i.e. task performance. Dark orange color signifies resting state activity, bright orange color signifies active state (executive) processing and blue indicates a shutdown of activity. Horizontal lines depict the highest hierarchical level that is recruited during task performance. We propose that increasingly difficult tasks tax increasingly deeper layers of a hierarchical control system until peak performance, after which the system collapses in a top-down fashion due to allostatic overload (hub collapse and cascading failure, see text). We predict that such changes show up as specific alternations in power-law frequency spectra. A. Resting state: a state of relatively little action-perception activity at lower levels that shifts the balance between top-down generative modelling (internal processing as symbolized by the loop) and bottom-up predictive coding in favor of generative modelling (musing, daydreaming). No bumps are observed. The power-law frequency distribution provides information on the full depth of the structural hierarchy. B. Simple tasks produce high-frequency bumps (flat tail flattening), indicating increased engagement of lower hierarchical levels involved in simple action-perception cycles (automatic or homeostatic control), i.e., reflexes or habits. The balance between resting state (generative modelling) and active state (predictive coding, belief updating) now shifts more in favor of the active state. C. More demanding tasks recruit intermediate to higher levels of the hierarchy, causing bumps at intermediate to lower frequencies (a mid-tail flattening). In our view, such changes reflect the recruitment of higher hierarchical levels that exert more sophisticated forms of control (effortful, allostatic, slow goal-directed behavior). The increased recruitment of higher levels manifests as an increase in the power of low frequencies until D. A top-down (functional) collapse of the hierarchical model occurs as a result of hub overload and cascading failure. We predict that such a collapse of hierarchical control manifests as a discrete event (a tipping point) that coincides with a sudden flattening of the power-law curve, i.e. a drop in low-frequency power, corresponding to a reduced involvement of higher levels and associated sophisticated traits (e.g. changes in personality traits during acute episodes of mental illness). The system falls back to less sophisticated means of coping (habitual or routine reflexive behavior), as indicated by an increase in the amplitude of the flat tail (the high frequency runoff). The loss of top-down central integrative control by higher level hub regions decreases synchronization between their subordinate areas, which reduces the predictability of activity changes at lower levels. The ensuing unpredictability of the timeseries can quantified in terms of permutation entropy, which is a measure of 'disorder' (see text).*

frequencies as a function of task difficulty therefore appears to be quite universal. From the perspective of active inference (see above), bumps on a power-law frequency distribution signal barriers to the efficient dissipation of variational free energy that need to be overcome by action (task performance) and belief updating (learning). Once the problem is solved, the bumps are 'smoothed out' and organisms can return to a low-free energy resting state (homeostasis, i.e., a power-law curve without bumps, Fig. 8).

In a previous study, we proposed that extremely challenging tasks produce prediction errors that reach the top of a regulatory hierarchy, reflecting the peak performance of the system. In this case, the organism must deploy its most sophisticated world models and corresponding action strategies to escape a difficult situation. Continuous peak performance may then cause the regulatory capacity of a system to overload. This is due to the region's disproportionately high density of hub units, which have the highest rates of energy dissipation and, thus, the highest (metabolic) energy demand (see above). The knot of a regulatory hierarchy is therefore most vulnerable to metabolic energy depletion. When energy demand exceeds energy supply, hub units shut down and fail in a cascading manner as a function of node degree, causing a top-down collapse of hierarchical control (Goekoop and de Kleijn, 2021b, a). This coincides with a shift in behavior from so called 'slow' to 'fast' survival strategies, i.e. from high-level (allostatic, goal-directed, integrated, contextualized, abstract, socially inclusive and long-term) to low-level (homeostatic, habitual or reflexive, segregated, decontextualized, concrete, self-centered and short-term) strategies (Wingfield, 2006). Together, such changes are referred to as 'allostatic overload' in the biological literature (McEwen and Wingfield, 2003). Allostatic overload (due to hub collapse) may continue until a tipping point is reached where large proportions of the network become functionally segregated, marking a discrete transition from a centrally coordinated and ordered state towards an uncoordinated, disordered state that is associated with malfunction, disease, or death (i.e., a tipping point). The ensuing disorder (a loss of predictability) can be quantified by a single term called permutation entropy and turns out to be a hallmark of many physical and mental disorders (Goekoop and De Kleijn, 2021a), see below.

Since power-law frequency spectra may convey information about the hierarchical status of a system, the collapse of hierarchical structure or function (allostatic overload) should coincide with specific changes in curve characteristics (Fig. 8). This is confirmed by a growing number of studies reporting temporary or permanent changes in power-law frequency distributions in various physical (Huikuri et al., 1998; Weissman and Binah, 2014) and mental disorders (Veerakumar et al., 2019; Lee et al., 2021; Robertson et al., 2019; Ramsay et al., 2021) as well as during social isolation (Weber et al., 2020), various sleep stages (Horváth et al., 2022), development (Ostlund et al., 2022; Hill et al., 2022) and aging (Voytek et al., 2015; Tran et al., 2020). The combined results of these studies appear to support the conclusion that increased task difficulty (such as REM sleep versus deep sleep or social engagement versus social isolation) involves a steepening of power-law curves, i.e., a gain in the amplitude of lower frequencies. In our view, this indicates the (increased) recruitment of deeper hierarchical layers that allow for more sophisticated (goal directed, allostatic) forms of control. Conversely, less difficult tasks appear to involve a flattening of the power-law curve (a relative decrease in lower frequency amplitudes), which is consistent with a greater reliance on lower hierarchical levels of control (e.g., habits, instinct patterns or reflexes). For instance, a loss of high-level regulatory hubs (allostatic control) has been observed in diabetes mellitus (Johnston et al., 2016), neurological disorders (Stam, 2014), major depression or schizophrenia (Bassett et al., 2008). This coincides with a flattening of power-law curves in such disorders (Veerakumar et al., 2019; Lee et al., 2021; Robertson et al., 2019; Ramsay et al., 2021). Also, flatter power-law curves have been found in disorders in which regulatory hierarchies fail to reach sufficient depth during development, producing a permanently underregulated state

(such as autism spectrum disorders, personality disorders or ADHD) (Robertson et al., 2019; Bruining et al., 2020). In some cases, however, pathology is related to a steepening of the power-law curve, e.g., during traumatic brain injury or aging. This may reflect situations in which higher (prefrontal) level are recruited (effortful processing) to compensate for the loss of intermediate-level functions. Steeper slopes may occur also occur in cases where aging or disease interfere with the necessity to keep certain frequencies tightly within limits, e.g. heart rate variability (Huikuri et al., 1998). The increased involvement of lower frequencies may signal a regulatory failure of systems that keep such variability under control. Overall, four logical types of regulatory deficits may occur in any type of organism across its lifespan, each of which may have both internal and external causes (i.e. control may fail in the presence or absence of extreme environmental circumstances). Such deficits should translate into specific changes in power-law frequency spectra that can be systematically parameterized using existing techniques (Donoghue et al., 2020) (Fig. 9).

In summary, we propose that the mereological (nested modular, hierarchical) structure of living system imposes a power-law imprint onto its dynamics (1/f 'pink noise'). This theory may have several consequences. Rather than a mere by-product, pink noise appears to be a cardinal feature of hierarchical message passing (information processing) in living systems. The current theory has implications for studies of the ontogeny (developmental aspects), phylogeny (evolutionary aspects) and the practical management of living systems in clinical medicine and ecology. These implications will be discussed below.

## 8. Implications of the theory: Ontological and phylogenetic aspects

In our view, power-law frequency spectra reflect the hierarchical structure of a system: higher (mereological) levels exert a tonic pressure onto the dynamics of their (constituent) lower-level systems through cross-frequency coupling ('top-down control'). The top of a regulatory hierarchy is therefore responsible for producing the zero-frequency 'offsets' (DC-components) in the timeseries of a system's overt behavior (Fig. 10). Such offsets reflect the default or mean expression levels of a system's inner message passing and overt behavior (Fig. 10). In other words, the top of a regulatory hierarchy controls the behavioral 'climate' of an organism (its personality) as opposed to its behavioral 'weather' (fast changing action-perception sequences), which is produced at its base. This conclusion is supported by studies in different fields of science. For instance, translational studies show that nearly all species express inter-individual differences in behavioral traits, i.e., 'personalities' (Carere and Maestripieri, 2013; Cabrera et al., 2021). The cause of such differences has been disproportionately localized within higher regulatory areas. For instance, high-level hub neurons and brain regions have been implicated in controlling stable behavioral traits in nematodes, smaller animals and humans (Zhang et al., 2021; van Oers and Mueller, 2010; Tompson et al., 2018). In bacteria, individual specimens from the same strain may differ with respect to exploratory or social behavior, which has been linked to differential expression levels of higher-level regulatory genes in genetic information bottlenecks (Keegstra et al., 2017; Lyon, 2015; Shapiro, 2007; Waite et al., 2018; Davidson and Surette, 2008; Ni et al., 2017). In plants, regulatory genes have been implicated in controlling individual differences in plant communication strategies, causing scientists to consider the possibility of 'plant personalities' (i.e. more exploratory or avoidant growth patterns) (Karban et al., 2022). Finally, human studies show that hub brain regions control the development of adult personalities and that developmental deficits or pathological changes in such areas produce personality-disorders (van den Heuvel and Sporns, 2019; DeYoung et al., 2022).

Apart from controlling overt behavior such as locomotion or fight- or flight responses, high-level regulatory systems control the morphological aspects of organisms (Kuchling et al., 2020; Pio-Lopez et al., 2022).
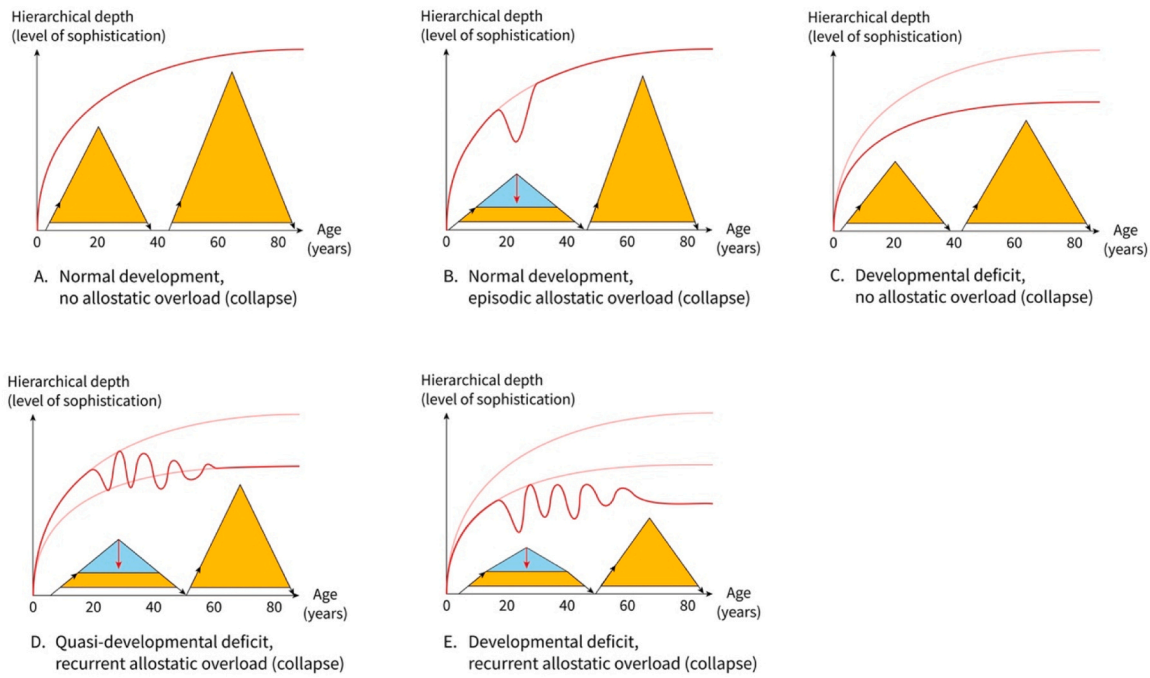
Fig. 9. Optimal development and four logical deficits of hierarchical control in living systems across the lifespan. *Triangles symbolize a hierarchical control system such as* Fig. 5. *Curve shows hierarchical depth against time (the organism's life span). See caption of* Fig. 8 *for further details. a. Optimal development, no episodic or chronic deficit. b. An (episodic) loss of hierarchical control relative to a previous attained level. c. A developmental (chronic) deficit of hierarchical control without episodic collapses. d. A prolonged sequence of episodic collapses, interfering with development. e. A developmental deficit leading to a prolonged sequence of episodic collapses (vicious cycle).*
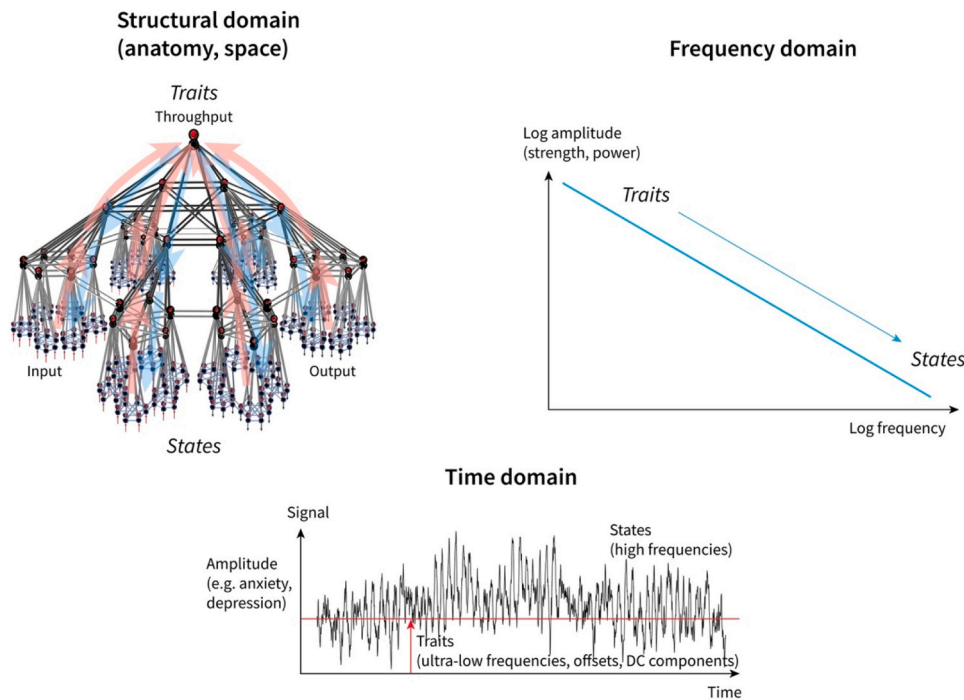
Fig. 10. How the top of a regulatory hierarchy controls the expression of stable aspects of inner experience and overt behavior (personality traits, morphological traits). *The top of a regulatory hierarchy contains high-level priors (setpoints, 'thermostats') that anticipate certain stable aspects of the environment (e.g. how warm, safe or violent an econiche will be). High-level priors are rarely updated by low-level prediction errors, causing their values to be relatively stable. Thus, high-level priors impose ultra-low frequency signal components onto the timeseries of their subordinate systems, down to the level of simple action-perception cycles (reflexes) or more complex reflexes (instinct patterns), which are involved e.g. in producing fight, flight, or freezing responses in living systems. Such stable signal components are called 'offsets' in statistics, or 'DC components' in engineering. Individual differences in the values of high-level priors produce more anxious (shy) or aggressive (dominant) individuals. A trait can therefore be grossly defined as a coordinate in state space that is given by the values of the offsets of the timeseries of the independent components or 'dimensions' that together define the state space (an attractor state). This is true for both behavioral traits and morphological traits (see text). Thus, high-level regulatory systems control the emergence of personalities and body plans in living systems.*

For example, homeotic genes are regulatory genes that control the timing and duration of gene expression, protein synthesis, cell division, and cell migration during embryogenesis and morphogenesis. This eventually governs the outgrowth of e.g., arms, legs, spines, tails or wings, producing different physical phenotypes (Pearson et al., 2005). Other regulatory genes control the transition from the body plans of youngsters into the adult phenotype during ontogenesis. Individual differences in the expression of such genes explain within-specifies differences in physical appearance (Mallo and Alonso, 2013). Problems in regulatory genes involved in physical development (e.g., growth hormone receptor) may lead to physical malformations, depending on the type of gene or tissue that is involved (Mallo and Alonso, 2013; Maurano et al., 2012). Thus, bowtie-structures in many cases control within-species inter-individual differences in morphological traits (body plans) as well as behavioral traits (personalities), which is collectively referred to as 'phenotypic variance'.

As observed, high-level regulatory systems contain prior units that encode the abstract and temporally stable aspects of an organism's econiche (such as meteorological or social climates). In other words, phenotypic variance reflects an individual's tendency to anticipate physically or behaviorally to certain stable aspects of the environment, i. e. a degree of *specialization* (Bergmüller and Taborsky, 2010). Such specialization allows for a 'division of labor' that prevents organisms from having to compete for similar (social) econiches (Cooper and West, 2018). In most organisms, morphological traits covary with personality traits to produce an overall phenotype and corresponding division of labor (e.g. in ants, individual differences in the body size, strength and aggression ants causes them to take on different social roles such as workers, nurses and soldiers). Such phenotypes are controlled by a limited number of regulatory hub genes that form the knot of a bowtie motif (Trible and Kronauer, 2021; Mikheyev and Linksvayer, 2015). Likewise, many species display clear sex differences in morphology and behavior (sexual dimorphism), which allows for a division of labor with respect to the production and rearing of offspring and fending off threats (Barr et al., 2018; Korb, 2016). Such differences are driven by a differential expression of high-level regulatory hub genes during embryogenesis and ontogenesis (Nijhout, 2003; Erwin and Davidson, 2009). In short, phenotypic variance of organisms reflects a degree of specialization that in many cases is controlled by high-level regulatory areas (information bottlenecks, the knots of bowties), rather than the base of the hierarchy (the wings of a bowtie).

As discussed, high-level regulatory areas are information bottlenecks that consist of a small number of hub nodes (a rich club) that control large areas of the network (Colizza et al., 2006). Because of their central positions, only minor changes in such key areas suffice to produce widely different body plans and personalities. When such variants are actively selected upon, novel species may emerge. This idea is supported by a growing number of observations that link bowtie motifs to the robustness and evolvability of living systems (Kitano, 2004). For instance, studies show that evolutionary changes in the past often involved the introduction of small mutations in so called 'hotspot genes' that are overrepresented in regulatory areas (Erwin and Davidson, 2009; Davidson and Erwin, 2006; Stern and Orgogozo, 2009). Studies in cichlids and voles show that diverse species may quickly emerge from a common ancestor as a result of only minor modifications in high-level regulatory systems, leading to a wide variety of body plans, diet specializations, social preferences and corresponding geographical distributions: a phenomenon known as adaptive radiation (Carruthers et al., 2022; Mehta et al., 2021; Insel and Shapiro, 1992). Similar findings involve the evolution of social behavior in different ant species (Mikheyev and Linksvayer, 2015; Shubin et al., 2009; Warner et al., 2019) and plant species (Subrahmaniam et al., 2018). In mammals, most between-species variance in brain architecture is found within mesocortical, prefrontal and anterior temporal areas, which involve regulatory (hub) regions. In contrast, sensorimotor cortices show relatively little cross-species differences (Striedter and Northcutt, 2019). Rather

than the base of the hierarchy (the wings of a bowtie), therefore, it's the *regulation* of such basic machinery (the knot of the bowtie) that makes the difference. A subtle tweaking of regulatory systems may be sufficient to distribute organisms across widely different econiches and social roles to optimize survival rates. Information bottlenecks therefore seem to play a key role in the specialization and speciation of living systems, which puts them center stage as 'hotspots of evolution' (Fig. 11).

To summarize, we propose that power-law frequency distributions result from the hierarchical (mereological) architecture of living systems. The top of the hierarchy controls the expression of behavioral as well as physical traits (i.e., the stable phenotype). Subtle modifications to information bottlenecks may cause strong phenotypic variance and drive the specialization and speciation of living systems.

9. Implications of the theory: Practical aspects

The idea that power-law frequency spectra convey information about the hierarchical dynamics of living systems has several practical consequences. For instance, such spectra could be used to monitor the quality of hierarchical information processing in living systems. A nonpathological example is the universal response of a brain to an unpredicted stimulus. This is known as event related desynchronisation and is characterized by an exuberance of high frequencies, relative to low frequencies. This indicates a relative absence of higher level (low frequency) regions that would otherwise predict and explain away lower-level prediction errors, causing the subject to be less 'surprised'. Conversely, low frequency alpha oscillations in the occipital lobe are usually characterized by states of mind that involve introspection and withdrawal from the sensorium (Pfurtscheller et al., 1996; Breakspear, 2002; Moruzzi and Magoun, 1949; Stam and De Bruin, 2004). This indicates a relatively strong involvement of high-level predictive regions involved in top-down generative modelling (imagination, mind wandering) relative to low-level action-perceptive loops that characterize the active state (see also Fig. 8).

In more pathological cases, the overburdening of regulatory capacity of living systems leads to a metabolic failure of hub nodes at the top of the regulatory hierarchy, producing a top-down collapse of hierarchical control ('allostatic overload' (Goekoop and de Kleijn, 2021b,a). This flattening of the 'command structure' produces a form of desynchronization, leaving subordinate levels in an underregulated state of disorder. Such disorder can be quantified by a single measure called 'permutation entropy', which expresses the amount of randomness or unpredictability in the timeseries of a system. The permutation entropy measure
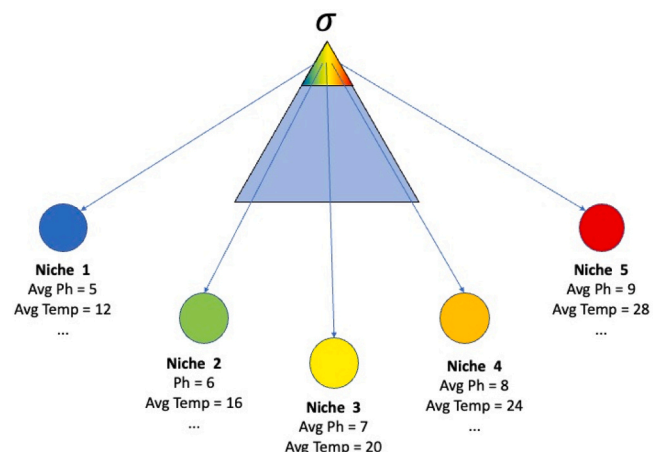
Fig. 11. Information bottlenecks as 'hotspots of evolution'. *Subtle alterations at the top of a regulatory hierarchy (the knot of a bowtie) produce a wide spectrum of different phenotypes (physical traits and personality traits) that may occupy a range of different ecological niches. Such phenotypical spectra are subjected to natural selection to quickly produce new species (adaptive radiation).*

(especially the more sophisticated versions such as defined in Li et al. (2021) takes Shannon's information entropy as a starting point and adapts this specifically for the purpose of quantifying randomness in timeseries (i.e., sequences of amplitude changes). This includes adaptations for taking amplitude changes of varying magnitudes into account and calculating such changes across multiple temporal scale levels (by systematically coarse graining / merging contiguous datapoints). Also, the obtained results are compared with those of white noise (pure randomness) as a reference value. Altogether, this yields a measure that captures more aspects of timeseries than the original Shannon's entropy. As such, is also more informative than simple descriptive statistics that track (changes in) e.g., the numbers of variables in a system, their correlations, autocorrelations, or signal variance. Such statistics are often used to describe the phenomenon of critical slowing down (CSD): a situation where regulatory systems slow to recover from perturbation, which is frequently found in undercontrolled living systems that struggle under severe levels of stress (Goekoop and De Kleijn, 2021a). A rise in CSD can presage the onset of tipping points (i.e. system failure). In a previous paper on the topic, we argued that permutation entropy can be used to more parsimoniously summarize CSD (Goekoop and De Kleijn, 2021a). This has practical benefits, such as a reduced need for calculations. However, it also has a conceptual advantage, since permutation entropy helps us to link the phenomenology of struggling systems (as witnessed by CSD) to the concept of disorder (higher levels of entropy). This makes it more consistent with existing notions of pathophysiological phenomena ('disorders') in clinical medicine, which often manifest by CSD and tipping points towards system failure. Together, these arguments speak in favor of using (novel versions of) permutation entropy rather than Shannon's entropy, or simple descriptive statistics, to detect regulatory deficits and corresponding pathological situations in living systems.

Indeed, increases in permutation entropy scores have been used as an early warning sign to discrete points that mark the transition from healthy and stable behavior to various physical and mental 'disorders', allowing for precautionary measures. Such tipping points (bifurcations or catastrophes) occur in any system under a significant amount of stress and signal the collapse of information bottleneck structures (Markov blankets) such as cell membranes during lysis or high-level neural systems during mental disorders. Permutation entropy has been used successfully to predict heart attacks, arrhythmias, epileptic seizures or mental disorders, relapses in autoinflammatory diseases, patients with mental disorders, social systems in disarray, polluted ecosystems and to seek out weak spots in a system's ability to control certain environmental challenges (e.g. selecting suitable antibiotics to treat bacterial infections) (Zhu et al., 2020). Rising levels of permutation entropy have been used to predict the collapse of single organisms as well as social systems, ecosystems, earthquakes, landslides, and stock market crashes (Goekoop and De Kleijn, 2021a; van de Leemput et al., 2014). Although living and non-living systems differ with respect to their higher order statistics of power-law frequency distributions (He, 2014), we expect these phenomena to nonetheless connect at the level of energy dissipation, which powers the dynamics of all open systems, whether living or non-living (Friston, 2019a; Uhl et al., 2015). Likewise, permutation entropy may be used as a universal quantifier of 'disorder' in any open dissipative system (Goekoop and De Kleijn, 2021a).

To permutation entropy, we now add a second measure from which to judge the quality of hierarchical information processing, which is the (steepness of the) power-law frequency distribution. Whereas permutation entropy is a more distal measures that consequences of failing hierarchical control (i.e,. disorder), power-law frequency spectra may be considered a more proximal measure that convey information on the depth of hierarchical information processing itself. In our view, a collapse of hierarchical control leads to a proportional loss of lower frequencies (and possibly an increase in higher frequencies), which is signalled by a sudden flattening of the power-law curve (Fig. 8D). Like permutation entropy, flattening power-law curves may serve as an early

warning sign for tipping points that mark the transition from ordered (healthy) states to disordered (unhealthy) states. These predictions can be tested in any hierarchical (Bayesian) control system that is taxed beyond its abilities to correct for environmental disturbances. Monitoring power-law exponents in relation to entropy (disorder) scores may thus serve preventive purposes and can be used to study the weaknesses of other systems.

Power-law dynamics may also serve to monitor the effects of interventions into dynamic systems. The degree to which power-law curves recover, disorder levels diminish and symptoms abate after certain interventions may serve as a quantifier for the restoration of hierarchical functioning and therapeutic success in clinical medicine. In psychiatry, for instance, the balance between top down and bottom-up information processing (hierarchical depth) may be tuned by neuromodulatory neurotransmitter systems, using pharmacotherapeutic agents such as antidepressants or other psychoactive drugs that modulate the efficiency of neural processing. The effect of such interventions can be tested by examining changes in power-law curves of brain function before and after treatment.

10. Discussion

We have proposed that hierarchical network structure produces hierarchical dynamics (state-trait continua) in (open dissipative) systems of coupled oscillators. Such systems engage in a vertical encoding of the deep spatiotemporal structure of their environments, which explains the emergence of different frequencies: the top of a hierarchy encodes slow events and produces low frequencies whereas lower reaches of the hierarchy encode fast events and produce fast frequencies. The typical run-off of amplitude with frequency is explained by an inverse relationship between coupling strength and energy dissipation rate: strongly coupled (hub) nodes at the top produce strong oscillations whereas loosely coupled nodes at the base of the hierarchy produce weaker oscillations. Since each hierarchical level is characterized by its own frequency, cross-frequency coupling should reflect top-down and bottom up 'control'. Thus, we propose that power-law frequency spectra quantify the process of hierarchical control in nested modular network structures. Crucially, this proposal applies both to living and non-living systems.

In the resting (not actively exploring) state, living systems display smooth power-law curves that represent the equal distribution of energy dissipation across all hierarchical levels and corresponding frequency bands. When actively sampling their environments (during task performance), living systems develop bumps on their power-law curves. In our view, such bumps involve the amplification of frequencies that are produced at a self-organized optimal hierarchical depth of information processing. Extreme task performance recruits the highest hierarchical levels of a control system to solve the most complex tasks. This selectively overburdens (metabolically demanding) hub nodes that are overrepresented in such areas. When such nodes overload and fail, this causes a top-down cascading failure of central-integrative hierarchical control ('allostatic overload'). Since hub nodes keep a network together and synchronized, their failure causes disconnection and desynchronization (disorder) of signal changes at lower levels. This can be quantified in terms of the permutation entropy score of the timeseries of a system. Since power-law frequency spectra may quantify hierarchical control, the top-down collapse of hierarchical control during allostatic overload should coincide with a flattening of (log-normalized) power-law curves, reflecting a decreased involvement of higher levels of control (lower frequencies). In this view, increased permutation entropy describes the more distal effects of collapsing hierarchical control (i.e., disorder), whereas flattening slopes may convey more proximal information (i.e. dysregulation: collapse of hierarchical control). Together, such changes may serve as early warning signs to system failure.

Testing this theory requires a transdisciplinary approach, which should avoid subjecting sentient creatures to unnecessarily stressful circumstances. One can consider studies in microbes under antibiotic

challenges, plants under conditions of severe drought, or situations in which 'natural experiments' have brought living systems to the limits of their regulatory capacities, such as patients suffering from various physical or mental conditions. We suspect that our theory will apply to the various sleep stages and undulating levels of consciousness that can be observed in various physiological and pathological situations (e.g., intoxication) and general anesthesia, although these have not been reviewed here. An open question is the degree to which the proposed theory and measures of hierarchical control apply to large-scale social systems such as ant colonies, beehives, human communities, organizations, governments or biotopes. All such systems may show shifts in powerlaw frequency distributions as a function of stress or the collapse of hierarchical structure. Also, such systems may display low-frequency traits or 'personalities', i.e., a default expression-level of certain policies that reflect certain stable aspects of the environment. Studies that address such questions can make use of *in silico* models to simulate the effects of changes in hierarchical structures on the dynamics of a system. Such studies require models of deep oscillatory neural networks with a folded information bottleneck motif that are involved in Bayesian inference. To date, many of such models are yet to be developed (for an early exception, see (Soman et al., 2018)). Because of its scale free nature, the free energy principle and its corollary, active inference, have the exciting prospect of marrying the physical sciences and the humanities, with obvious practical, philosophical and ethical implications (Parr et al., 2022). Overall, the 21st century promises to be an extraordinary time for anyone who wishes to optimize their predictive models of the world.

Acknowledgments

References

Adams, R.A., Shipp, S., Friston, K.J., 2013. Predictions not commands: active inference in the motor system. Brain Struct. Funct. 218 (3), 611–643.

Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z.N., Barabasi, A.L., 2004. Global organization of metabolic fluxes in the bacterium Escherichia coli. Nature 427 (6977), 839–843.

Alon, U., 2007. Network motifs: theory and experimental approaches. Nat. Rev. Genet. 8 (6), 450–461.

Annila, A., Kuismanen, E., 2009. Natural hierarchy emerges from energy dispersal. Biosystems 95 (3), 227–233.

Atasoy, S., Donnelly, I., Pearson, J., 2016. Human brain networks function in connectome-specific harmonic waves. Nat. Commun. 7 (1), 10340.

Atasoy, S., Deco, G., Kringelbach, M.L., Pearson, J., 2018. Harmonic brain modes: a unifying framework for linking space and time in brain dynamics. Neuroscientist 24 (3), 277–293.

Bak, P., Tang, C., Wiesenfeld, K., 1988. Self-organized criticality. Phys. Rev. A Gen. Phys. 38 (1), 364–374.

Barabasi, A.L., 2005. The origin of bursts and heavy tails in human dynamics. Nature 435 (7039), 207–211.

Barabasi, A.L., 2009. Scale-free networks: a decade and beyond. Science 325 (5939), 412–413.

Barr, M.M., Garcia, L.R., Portman, D.S., 2018. Sexual dimorphism and sex differences in caenorhabditis elegans neuronal development and behavior. Genetics 208 (3), 909–935.

Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia. J. Neurosci. 28 (37), 9239–9248.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J., 2012. Canonical microcircuits for predictive coding. Neuron 76 (4), 695–711.

Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J.R., et al., 2015. Visual areas exert feedforward and feedback influences through distinct frequency channels. Neuron 85 (2), 390–401.

Bergmüller, R., Taborsky, M., 2010. Animal personality due to social niche specialisation. Trends Ecol. Evol. 25 (9), 504–511.

Bordyugov, G., Granada, A., Herzel, H., 2011. How coupling determines the entrainment of circadian clocks. Eur. Phys. J. B 82 (3), 227–234.

Bosman, C.A., Schoffelen, J.-M., Brunet, N., Oostenveld, R., Bastos, A.M., Womelsdorf, T., et al., 2012. Attentional stimulus selection through selective synchronization between monkey visual areas. Neuron 75 (5), 875–888.

Breakspear, M., 2002. Nonlinear phase desynchronization in human electroencephalographic data. Hum. Brain Mapp. 15 (3), 175–198.

Broido, A.D., Clauset, A., 2019. Scale-free networks are rare. Nat. Commun. 10 (1), 1017.

Bruining, H., Hardstone, R., Juarez-Martinez, E.L., Sprengers, J., Avramiea, A.-E., Simpraga, S., et al., 2020. Measurement of excitation-inhibition ratio in autism spectrum disorder using critical brain dynamics. Sci. Rep. 10 (1), 9195.

Cabrera, D., Nilsson, J.R., Griffen, B.D., 2021. The development of animal personality across ontogeny: a cross-species review. Anim. Behav. 173, 137–144.

Canolty, R.T., Knight, R.T., 2010. The functional role of cross-frequency coupling. Trends Cogn. Sci. 14 (11), 506–515.

Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., et al., 2006. High gamma power is phase-locked to theta oscillations in human neocortex. Science 313 (5793), 1626–1628.

Carere, C., Maestripieri, D., 2013. Animal Personalities: Behavior, Physiology, and Evolution. University of Chicago Press,.

Carruthers, M., Edgley, D.E., Saxon, A.D., Gabagambi, N.P., Shechonge, A., Miska, E.A., et al., 2022. Ecological speciation promoted by divergent regulation of functional genes within African cichlid fishes. bioRxiv.

Citri, A., Yarden, Y., 2006. EGF-ERBB signalling: towards the systems level. Nat. Rev. Mol. Cell Biol. 7 (7), 505–516.

Ciuciu, P., Abry, P., He, B.J., 2014. Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks. Neuroimage 95, 248–263.

Clarke, A.R., Barry, R.J., Karamacoska, D., Johnstone, S.J., 2019. The EEG theta/beta ratio: a marker of arousal or cognitive processing capacity? Appl. Psychophysiol. Biofeedback 44, 123–129.

Clauset, A., Shalizi, C.R., Newman, M.E., 2009. Power-law distributions in empirical data. SIAM Rev. 51 (4), 661–703.

Clegg, R.G., Di Cairano-Gilfedder, C., Zhou, S., 2010. A critical look at power law modelling of the Internet. Comput. Commun. 33 (3), 259–268.

Colizza, V., Flammini, A., Serrano, M.A., Vespignani, A., 2006. Detecting rich-club ordering in complex networks. Nat. Phys. 2 (2), 110–115.

Conant, R.C., Ross Ashby, W., 1970. Every good regulator of a system must be a model of that system. Int. J. Syst. Sci. 1 (2), 89–97.

Constant, A., Ramstead, M.J.D., Veissiere, S.P.L., Campbell, J.O., Friston, K.J., 2018. A variational approach to niche construction. J. R. Soc. Interface 15 (141), 20170685.

Constant, A., Ramstead, M.J.D., Veissiere, S.P.L., Friston, K., 2019. Regimes of expectations: an active inference model of social conformity and human decision making. Front Psychol. 10, 679.

Constant, A., Clark, A., Friston, K.J., 2021. Representation wars: enacting an armistice through active inference. Front. Psychol. 11, 598733.

Cooper, G.A., West, S.A., 2018. Division of labour and the evolution of extreme specialization. Nat. Ecol. Evol. 2 (7), 1161–1167.

Cui, Y., Ahmad, S., Hawkins, J., 2016. Continuous online sequence learning with an unsupervised neural network model. Neural Comput. 28 (11), 2474–2504.

Davidson, C.J., Surette, M.G., 2008. Individuality in bacteria. Annu Rev. Genet 42, 253–268.

Davidson, E.H., Erwin, D.H., 2006. Gene regulatory networks and the evolution of animal body plans. Science 311 (5762), 796–800.

Deco, G., Jirsa, V.K., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat. Rev. Neurosci. 12 (1), 43–56.

DeYoung, C.G., Beaty, R.E., Genç, E., Latzman, R.D., Passamonti, L., Servaas, M.N., et al., 2022. Personality neuroscience: an emerging field with bright prospects. Personal. Sci. 3, 1–21.

Donoghue, T., Haller, M., Peterson, E.J., Varma, P., Sebastian, P., Gao, R., et al., 2020. Parameterizing neural power spectra into periodic and aperiodic components. Nat. Neurosci. 23 (12), 1655–1665.

Erwin, D.H., Davidson, E.H., 2009. The evolution of hierarchical gene regulatory networks. Nat. Rev. Genet. 10 (2), 141–148.

Evans, J.S., 2008. Dual-processing accounts of reasoning, judgment, and social cognition. Annu. Rev. Psychol. 59, 255–278.

Fortier M., Friedman D.A. Of woodlice and men. 2018:17.

Fransson, P., Metsaranta, M., Blennow, M., Aden, U., Lagercrantz, H., Vanhatalo, S., 2013. Early development of spatial patterns of power-law frequency scaling in FMRI resting-state and EEG data in the newborn brain. Cereb. Cortex 23 (3), 638–646.

Friedlander, T., Mayo, A.E., Tlusty, T., Alon, U., 2015. Evolution of bow-tie architectures in biology. PLOS Comput. Biol. 11 (3), e1004055.

Friedman, E.J., Landsberg, A.S., 2013. Hierarchical networks, power laws, and neuronal avalanches. Chaos 23 (1), 013135.

Friston K. 2019a. A free energy principle for a particular physics. arXiv preprint arXiv: 190610184.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11 (2), 127–138.

Friston, K., 2012. A free energy principle for biological systems. Entropy 14 (11), 2100–2121.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B: Biol. Sci. 364 (1521), 1211–1221.

Friston, K., Kilner, J., Harrison, L., 2006. A free energy principle for the brain. J. Physiol. 100 (1–3), 70–87.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G., 2017. Active inference: a process theory. Neural Comput. 29 (1), 1–49.

Friston, K., Moran, R.J., Nagai, Y., Taniguchi, T., Gomi, H., Tenenbaum, J., 2021a. World model learning and inference. Neural Netw. 144, 573–590.

Friston, K., Da Costa, L., Hafner, D., Hesp, C., Parr, T., 2021a. Sophisticated inference. Neural Comput. 33 (3), 713–763.

Friston, K.J., 2019b. Waves of prediction. PLOS Biol. 17 (10), e3000426.

Friston, K.J., Rosch, R., Parr, T., Price, C., Bowman, H., 2018. Deep temporal models and active inference. Neurosci. Biobehav. Rev. 90, 486–501.

Friston, K.J., Fagerholm, E.D., Zarghami, T.S., Parr, T., Hipolito, I., Magrou, L., et al., 2021b. Parcels and particles: Markov blankets in the brain. Netw. Neurosci. 5 (1), 211–251.

Fujita, Y., Kichikawa, Y., Fujiwara, Y., Souma, W., Iyetomi, H., 2019. Local bow-tie structure of the web. Appl. Netw. Sci. 4 (1), 1–15.

Gallagher, S., Allen, M., 2018. Active inference, enactivism and the hermeneutics of social cognition. Synthese 195 (6), 2627–2648.

Gao, R., Peterson, E.J., Voytek, B., 2017. Inferring synaptic excitation/inhibition balance from field potentials. Neuroimage 158, 70–78.

George, D., Hawkins, J., 2009. Towards a mathematical theory of cortical micro-circuits. PLOS Comput. Biol. 5 (10), e1000532.

Gershman, S.J., Horvitz, E.J., Tenenbaum, J.B., 2015. Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. Science 349 (6245), 273–278.

Gilbert, C.D., Li, W., 2013. Top-down influences on visual processing. Nat. Rev. Neurosci. 14 (5), 350–363.

Giraud, A.-L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 15 (4), 511–517.

Goekoop, R., de Kleijn, R., 2021. Permutation entropy as a universal disorder criterion: how disorders at different scale levels are manifestations of the same underlying principle. Entropy 23 (12), 1701.

Goekoop, R., de Kleijn, R., 2021. How higher goals are constructed and collapse under stress: a hierarchical Bayesian control systems perspective. Neurosci. Biobehav. Rev. 123, 257–285.

Gosak, M., Stožer, A., Markovič, R., Dolenšek, J., Marhl, M., Slak Rupnik, M., et al., 2015. The relationship between node degree and dissipation rate in networks of diffusively coupled oscillators and its significance for pancreatic beta cells. Chaos: Interdiscip. J. Nonlinear Sci. 25 (7), 073113.

Griffa, A., Van den Heuvel, M.P., 2022. Rich-club neurocircuitry: function, evolution, and vulnerability. Dialog-. Clin. Neurosci.

Ha D. , Schmidhuber J, 2018. World models. arXiv preprint arXiv:180310122.

Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N., 2008. A hierarchy of temporal receptive windows in human cortex. J. Neurosci. 28 (10), 2539–2550.

He, B.J., 2011. Scale-free properties of the functional magnetic resonance imaging signal during rest and task. J. Neurosci. 31 (39), 13786–13795.

He, B.J., 2014. Scale-free brain activity: past, present, and future. Trends Cogn. Sci. 18 (9), 480–487.

He, B.J., Zempel, J.M., Snyder, A.Z., Raichle, M.E., 2010. The temporal structures and functional significance of scale-free brain activity. Neuron 66 (3), 353–369.

Hellrigel, S., Jarman, N., van Leeuwen, C., 2019. Adaptive rewiring in weighted networks. Cogn. Syst. Res. 55, 205–218.

Henin, S., Turk-Browne, N.B., Friedman, D., Liu, A., Dugan, P., Flinker, A., et al., 2021. Learning hierarchical sequence representations across human cortex and hippocampus. Sci. Adv. 7 (8), eabc4530.

Henzler-Wildman, K.A., Lei, M., Thai, V., Kerns, S.J., Karplus, M., Kern, D., 2007. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature 450 (7171), 913–916.

Hesp, C., Ramstead, M., Constant, A., Badcock, P., Kirchhoff, M., Friston, K., 2019. A multi-scale view of the emergent complexity of life: a free-energy proposal. Evolution, Development and Complexity. Springer,, pp. 195–227.

van den Heuvel, M.P., Sporns, O., 2019. A cross-disorder connectome landscape of brain dysconnectivity. Nat. Rev. Neurosci. 20 (7), 435–446.

Hill, A.T., Clark, G.M., Bigelow, F.J., Lum, J.A., Enticott, P.G., 2022. Periodic and aperiodic neural activity displays age-dependent changes across early-to-middle childhood. Dev. Cogn. Neurosci. 54, 101076.

Hohwy, J., 2016. The self-evidencing brain. Noûs 50 (2), 259–285.

Horváth CG , Szalárdy O. , Ujma P.P. , Simor P. , Gombos F. , Kovács I., et al. 2022. Overnight dynamics in scale-free and oscillatory spectral parameters of NREM sleep EEG: can polysomnography provide an insight into sleep regulation?.

Hovsepyan, S., Olasagasti, I., Giraud, A.-L., 2018. Combining predictive coding with neural oscillations optimizes on-line speech processing. BioRxiv, 477588.

Huang, X., Xu, K., Chu, C., Jiang, T., Yu, S., 2017. Weak higher-order interactions in macroscopic functional networks of the resting brain. J. Neurosci. 37 (43), 10481–10497.

Huikuri, H.V., Mäkikallio, T.H., Airaksinen, K.J., Seppänen, T., Puukka, P., Räihä, I.J., et al., 1998. Power-law relationship of heart rate variability as a predictor of mortality in the elderly. Circulation 97 (20), 2031–2036.

Hyafil, A., Giraud, A.-L., Fontolan, L., Gutkin, B., 2015. Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. Trends Neurosci. 38 (11), 725–740.

Insel, T.R., Shapiro, L.E., 1992. Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. Proc. Natl. Acad. Sci. USA 89 (13), 5981–5985.

Jarman, N., Steur, E., Trengove, C., Tyukin, I.Y., van Leeuwen, C., 2017. Self-organisation of small-world networks by adaptive rewiring in response to graph diffusion. Sci. Rep. 7 (1), 13158.

Jeffery, K., Pollack, R., Rovelli, C., 2019. On the statistical mechanics of life: Schrödinger revisited. Entropy 21 (12), 1211.

Johnston, N.R., Mitchell, R.K., Haythorne, E., Pessoa, M.P., Semplici, F., Ferrer, J., et al., 2016. Beta cell hubs dictate pancreatic islet responses to glucose. Cell Metab. 24 (3), 389–401.

Kanai, R., Komura, Y., Shipp, S., Friston, K., 2015. Cerebral hierarchies: predictive processing, precision and the pulvinar. Philos. Trans. R. Soc. B: Biol. Sci. 370 (1668), 20140169.

Kaplan, H.S., Thula, O.S., Khoss, N., Zimmer, M., 2020. Nested neuronal dynamics orchestrate a behavioral hierarchy across timescales. Neuron 105 (3), 562–576 e9.

Karban, R., Grof-Tisza, P., Couchoux, C., 2022. Consistent individual variation in plant communication: do plants have personalities? Oecologia 199 (1), 129–137.

Kauffman, S., 1996. At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford University Press,.

Kawakami, E., Singh, V.K., Matsubara, K., Ishii, T., Matsuoka, Y., Hase, T., et al., 2016. Network analyses based on comprehensive molecular interaction maps reveal robust control structures in yeast stress response pathways. NPJ Syst. Biol. Appl. 2, 15018.

Keegstra, J.M., Kamino, K., Anquez, F., Lazova, M.D., Emonet, T., Shimizu, T.S., 2017. Phenotypic diversity and temporal variability in a bacterial signaling network revealed by single-cell FRET. eLife 6, e27455.

Kiebel, S.J., Daunizeau, J., Friston, K.J., 2008. A hierarchy of time-scales and the brain. PLOS Comput. Biol. 4 (11), e1000209.

Kiebel, S.J., Daunizeau, J., Friston, K.J., 2009. Perception and hierarchical dynamics. Front. Neuroinform. 3, 569.

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., Kiverstein, J., 2018. The Markov blankets of life: autonomy, active inference and the free energy principle. J. R. Soc. Interface 15 (138), 20170792.

Kitano, H., 2004. Biological robustness. Nat. Rev. Genet. 5 (11), 826–837.

de Kleijn, R., Kachergis, G., Hommel, B., 2014. Everyday robotic action: lessons from human action control. Front. Neurorobot. 8, 13.

Knill, D.C., Pouget, A., 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 27 (12), 712–719.

Korb, J., 2016. Genes underlying reproductive division of labor in termites, with comparisons to social hymenoptera. Front. Ecol. Evol. 4.

Kotov, R., Krueger, R.F., Watson, D., Cicero, D.C., Conway, C.C., DeYoung, C.G., et al., 2021. The hierarchical taxonomy of psychopathology (HiTOP): a quantitative nosology based on consensus of evidence. Annu. Rev. Clin. Psychol. 17, 83–108.

Kruglanski, A.W., Jasko, K., Friston, K., 2020. All thinking is 'wishful'thinking. Trends Cogn. Sci. 24 (6), 413–424.

Kuchling, F., Friston, K., Georgiev, G., Levin, M., 2020. Morphogenesis as Bayesian inference: a variational approach to pattern formation and control in complex biological systems. Phys. Life Rev. 33, 88–108.

Kuzemsky, A.L., 2015. Variational principle of Bogoliubov and generalized mean fields in many-particle interacting systems. Int. J. Mod. Phys. B 29 (18), 1530010.

Lane, N., 2015. The Vital Question: Energy, Evolution, and the Origins of Complex Life. WW Norton & Company,.

Lee, Y.-J., Huang, S.-Y., Lin, C.-P., Tsai, S.-J., Yang, A.C., 2021. Alteration of power law scaling of spontaneous brain activity in schizophrenia. Schizophr. Res. 238, 10–19.

Li, J., Hua, X., Haubrock, M., Wang, J., Wingender, E., 2012. The architecture of the gene regulatory networks of different tissues. Bioinformatics 28 (18), i509–i514.

Li, Y., Geng, B., Jiao, S., 2021. Refined composite multi-scale reverse weighted permutation entropy and its applications in ship-radiated noise. Entropy 23 (4), 476.

Liu, Y.Y., Slotine, J.J., Barabasi, A.L., 2012. Control centrality and hierarchical structure in complex networks. PLOS One 7 (9), e44459.

Lyon, P., 2015. The cognitive cell: bacterial behavior reconsidered. Front. Microbiol. 6, 264.

Mahjoory, K., Schoffelen, J.-M., Keitel, A., Gross, J., 2020. The frequency gradient of human resting-state brain oscillations follows cortical hierarchies. eLife 9, e53715.

Maisto, D., Donnarumma, F., Pezzulo, G., 2015. Divide et impera: subgoaling reduces the complexity of probabilistic inference and problem solving. J. R. Soc. Interface 12 (104), 20141335.

Mäkelä, T., Annila, A., 2010. Natural patterns of energy dispersal. Phys. Life Rev. 7 (4), 477–498.

Mallo, M., Alonso, C.R., 2013. The regulation of Hox gene expression during animal development. Development 140 (19), 3951–3963.

Marino, J., 2022. Predictive coding, variational autoencoders, and biological connections. Neural Comput. 34 (1), 1–44.

Markov, N.T., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., Kennedy, H., 2013. Cortical high-density counterstream architectures. Science 342 (6158), 1238406.

Markovič, R., Gosak, M., Marhl, M., 2014. Broad-scale small-world network topology induces optimal synchronization of flexible oscillators. Chaos, Solitons Fractals 69, 14–21.

Marsh, H.W., Shavelson, R., 1985. Self-concept: Its multifaceted, hierarchical structure. Educ. Psychol. 20 (3), 107–123.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., et al., 2012. Systematic localization of common disease-associated variation in regulatory DNA. Science 337 (6099), 1190–1195.

Maxim, V., Şendur, L., Fadili, J., Suckling, J., Gould, R., Howard, R., et al., 2005. Fractional Gaussian noise, functional MRI and Alzheimer's disease. Neuroimage 25 (1), 141–158.

McEwen, B.S., Wingfield, J.C., 2003. The concept of allostasis in biology and biomedicine. Horm. Behav. 43 (1), 2–15.

Mehta, T.K., Koch, C., Nash, W., Knaack, S.A., Sudhakar, P., Olbei, M., et al., 2021. Evolution of regulatory networks associated with traits under selection in cichlids. Genome Biol. 22 (1), 1–28.

de Menezes, M.A., Barabasi, A.L., 2004. Fluctuations in network dynamics. Phys. Rev. Lett. 92 (2), 028701.

Mescheder, L., Nowozin, S., Geiger, A. (Eds.), 2017. Adversarial Variational Bayes: Unifying Variational Autoencoders and Generative Adversarial Networks. International Conference on Machine Learning. PMLR.

Mikheyev, A.S., Linksvayer, T.A., 2015. Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. eLife 4.

Millidge, B., Tschantz, A., Buckley, C.L., 2022. Predictive coding approximates backprop along arbitrary computation graphs. Neural Comput. 34 (6), 1329–1368.

Montez, T., Poil, S.-S., Jones, B.F., Manshanden, I., Verbunt, J.P., van Dijk, B.W., et al., 2009. Altered temporal correlations in parietal alpha and prefrontal theta oscillations in early-stage Alzheimer disease. Proc. Natl. Acad. Sci. USA 106 (5), 1614–1619.

Moon, J.-Y., Lee, U., Blain-Moraes, S., Mashour, G.A., 2015. General relationship of global topology, local dynamics, and directionality in large-scale brain networks. PLOS Comput. Biol. 11 (4), e1004225.

Moruzzi, G., Magoun, H.W., 1949. Brain stem reticular formation and activation of the EEG. Electroencephalogr. Clin. Neurophysiol. 1 (1–4), 455–473.

Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., et al., 2014. A hierarchy of intrinsic timescales across primate cortex. Nat. Neurosci. 17 (12), 1661–1663.

Nakamura, T., Takumi, T., Takano, A., Aoyagi, N., Yoshiuchi, K., Struzik, Z.R., et al., 2008. Of mice and men—universality and breakdown of behavioral organization. PLOS One 3 (4), e2050.

Nekovee, M., Moreno, Y., Bianconi, G., Marsili, M., 2007. Theory of rumour spreading in complex social networks. Phys. A: Stat. Mech. Appl. 374 (1), 457–470.

Newman, M.E., 2012. Communities, modules and large-scale structure in networks. Nat. Phys. 8 (1), 25–31.

Ni, B., Ghosh, B., Paldy, F.S., Colin, R., Heimerl, T., Sourjik, V., 2017. Evolutionary remodeling of bacterial motility checkpoint control. Cell Rep. 18 (4), 866–877.

Nijhout, H.F., 2003. Development and evolution of adaptive polyphenisms. Evol. Dev. 5 (1), 9–18.

Novák, B., Tyson, J.J., 2008. Design principles of biochemical oscillators. Nat. Rev. Mol. Cell Biol. 9 (12), 981–991.

Onslow, A.C., Jones, M.W., Bogacz, R., 2014. A canonical circuit for generating phase-amplitude coupling. PLOS One 9 (8), e102591.

Ostlund, B., Donoghue, T., Anaya, B., Gunther, K.E., Karalunas, S.L., Voytek, B., et al., 2022. Spectral parameterization for studying neurodevelopment: how and why. Dev. Cogn. Neurosci. 54, 101073.

Paine, R.W., Tani, J., 2005. How hierarchical control self-organizes in artificial adaptive systems. Adapt. Behav. 13 (3), 211–225.

Parr, T., Pezzulo, G., Friston, K.J., 2022. Active Inference: The Free Energy Principle in Mind, Brain, and Behavior. MIT Press,.

Pearson, J.C., Lemons, D., McGinnis, W., 2005. Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet. 6 (12), 893–904.

Perdikis, D., Huys, R., Jirsa, V.K., 2011. Time scale hierarchies in the functional organization of complex behaviors. PLOS Comput. Biol. 7 (9), e1002198.

Pezzulo, G., Rigoli, F., Friston, K., 2015. Active Inference, homeostatic regulation and adaptive behavioural control. Prog. Neurobiol. 134, 17–35.

Pezzulo, G., Rigoli, F., Friston, K.J., 2018. Hierarchical active inference: a theory of motivated control. Trends Cogn. Sci. 22 (4), 294–306.

Pezzulo, G., Parr, T., Friston, K., 2022. The evolution of brain architectures for predictive coding and active inference. Philos. Trans. R. Soc. B 377 (1844), 20200531.

Pfurtscheller, G., Stancak Jr, A., Neuper, C., 1996. Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: a review. Int. J. Psychophysiol. 24 (1–2), 39–46.

Piasini, E., Soltuzu, L., Muratore, P., Caramellino, R., Vinken, K., Op de Beeck, H., et al., 2021. Temporal stability of stimulus representation increases along rodent visual cortical hierarchies. Nat. Commun. 12 (1), 1–19.

Pio-Lopez, L., Kuchling, F., Tung, A., Pezzulo, G., Levin, M., 2022. Active inference, morphogenesis, and computational psychiatry. Front. Comput. Neurosci. 16, 988977.

Powers, W.T., 1973. Feedback: beyond behaviorism: stimulus-response laws are wholly predictable within a control-system model of behavioral organization. Science 179 (4071), 351–356.

Puma, S., Matton, N., Paubel, P.-V., Raufaste, É., El-Yagoubi, R., 2018. Using theta and alpha band power to assess cognitive workload in multitasking environments. Int. J. Psychophysiol. 123, 111–120.

Purcell, B.A., Kiani, R., 2016. Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. Proc. Natl. Acad. Sci. USA 113 (31), E4531–E4540.

Quinn, C., O'Kane, T.J., Kitsios, V., 2020. Application of a local attractor dimension to reduced space strongly coupled data assimilation for chaotic multiscale systems. Nonlin Process. Geophys 27 (1), 51–74.

Ramírez, J.C., Marshall, J.A., 2017. Can natural selection encode Bayesian priors? J. Theor. Biol. 426, 57–66.

Ramsay, I., Lynn, P., Lee, E., Schermitzler, B., Leipold, D., Sponheim, S., 2021. Disturbances in aperiodic neural activity during resting state in patients with schizophrenia. Biol. Psychiatry 89 (9), S254–S255.

Ramstead, M.J.D., Badcock, P.B., Friston, K.J., 2018. Answering Schrödinger's question: a free-energy formulation. Phys. Life Rev. 24, 1–16.

Raut, R.V., Snyder, A.Z., Raichle, M.E., 2020. Hierarchical dynamics as a macroscopic organizing principle of the human brain. Proc. Natl. Acad. Sci. USA 117 (34), 20890–20897.

Rentzeperis, I., Laquitaine, S., van Leeuwen, C., 2022. Adaptive rewiring of random neural networks generates convergent–divergent units. Commun. Nonlinear Sci. Numer. Simul. 107, 106135.

Rigoli, F., Pezzulo, G., Dolan, R., Friston, K., 2017. A goal-directed Bayesian framework for categorization. Front. Psychol. 8, 408.

Robertson, M.M., Furlong, S., Voytek, B., Donoghue, T., Boettiger, C.A., Sheridan, M.A., 2019. EEG power spectral slope differs by ADHD status and stimulant medication exposure in early childhood. J. Neurophysiol. 122 (6), 2427–2437.

Rohe, T., Ehlis, A.C., Noppeney, U., 2019. The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. Nat. Commun. 10 (1), 1907.

Roweis, S., Ghahramani, Z., 1999. A unifying review of linear Gaussian models. Neural Comput. 11 (2), 305–345.

Sadaghiani, S., Hesselmann, G., Friston, K.J., Kleinschmidt, A., 2010. The relation of ongoing brain activity, evoked neural responses, and cognition. Front. Syst. Neurosci. 4, 20.

Salvatori T., Song Y., Xu Z., Lukasiewicz T., Bogacz R., 2022. editors. Reverse differentiation via predictive coding. In: Proceedings of the AAAI Conference on Artificial Intelligence.

Schroeder, C.E., Wilson, D.A., Radman, T., Scharfman, H., Lakatos, P., 2010. Dynamics of active sensing and perceptual selection. Curr. Opin. Neurobiol. 20 (2), 172–176.

Seth, A.K., 2014. The cybernetic Bayesian brain. Open Mind: Open MIND. MIND Group,, Frankfurt am Main.

Shapiro, J.A., 2007. Bacteria are small but not stupid: cognition, natural genetic engineering and socio-bacteriology. Stud. Hist. Philos. Sci. Part C: Stud. Hist. Philos. Biol. Biomed. Sci. 38 (4), 807–819.

Sherman, S.M., Guillery, R., 2011. Distinct functions for direct and transthalamic corticocortical connections. J. Neurophysiol. 106 (3), 1068–1077.

Shipp, S., 2005. The importance of being agranular: a comparative account of visual and motor cortex. Philos. Trans. R. Soc. B: Biol. Sci. 360 (1456), 797–814.

Shubin, N., Tabin, C., Carroll, S., 2009. Deep homology and the origins of evolutionary novelty. Nature 457 (7231), 818–823.

Smit, D.J., de Geus, E.J., van de Nieuwenhuijzen, M.E., van Beijsterveldt, C.E., van Baal, G.C.M., Mansvelder, H.D., et al., 2011. Scale-free modulation of resting-state neuronal oscillations reflects prolonged brain maturation in humans. J. Neurosci. 31 (37), 13128–13136.

Soman, K., Muralidharan, V., Chakravarthy, V.S., 2018. An oscillatory neural autoencoder based on frequency modulation and multiplexing. Front. Comput. Neurosci. 12, 52.

Song, C., Havlin, S., Makse, H.A., 2005. Self-similarity of complex networks. Nature 433 (7024), 392–395.

Stam, C.J., 2014. Modern network science of neurological disorders. Nat. Rev. Neurosci. 15 (10), 683–695.

Stam, C.J., De Bruin, E.A., 2004. Scale-free dynamics of global functional connectivity in the human brain. Hum. Brain Mapp. 22 (2), 97–109.

Stern, D.L., Orgogozo, V., 2009. Is genetic evolution predictable? Science 323 (5915), 746–751.

Steyer, R., Schmitt, M., Eid, M., 1999. Latent state–trait theory and research in personality and individual differences. Eur. J. Personal. 13 (5), 389–408.

Stožer, A., Šterk, M., Leitgeb, E.P., Markovič, R., Klemen, M.S., Ellis, C.E., et al., 2022. From isles of Königsberg to islets of Langerhans: examining the function of the endocrine pancreas through network science. Front. Endocrinol. 13.

Striedter, G.F., Northcutt, R.G., 2019. Brains through Time: A Natural History of Vertebrates. Oxford University Press,.

Subrahmaniam, H.J., Libourel, C., Journet, E.P., Morel, J.B., Muños, S., Niebel, A., et al., 2018. The genetics underlying natural variation of plant–plant interactions, a beloved but forgotten member of the family of biotic interactions. Plant J. 93 (4), 747–770.

Tagliazucchi, E., von Wegner, F., Morzelewski, A., Brodbeck, V., Jahnke, K., Laufs, H., 2013. Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep. Proc. Natl. Acad. Sci. USA 110 (38), 15419–15424.

Tenenbaum, J.B., Kemp, C., Griffiths, T.L., Goodman, N.D., 2011. How to grow a mind: Statistics, structure, and abstraction. science 331 (6022), 1279–1285.

Tinker, J., Velazquez, J.L.P., 2014. Power law scaling in synchronization of brain signals depends on cognitive load. Front. Syst. Neurosci. 8, 73.

, 2015Tishby N., Zaslavsky N., 2015. editors. Deep learning and the information bottleneck principle. In: Proceedings of the 2015 IEEE Information Theory Workshop (itw), IEEE.

Tolkunov, D., Rubin, D., Mujica-Parodi, L.R., 2010. Power spectrum scale invariance quantifies limbic dysregulation in trait anxious adults using fMRI: adapting methods optimized for characterizing autonomic dysregulation to neural dynamic time series. Neuroimage 50 (1), 72–80.

Tompson, S.H., Falk, E.B., Vettel, J.M., Bassett, D.S., 2018. Network approaches to understand individual differences in brain connectivity: opportunities for personality neuroscience. Personal. Neurosci. 1.

Tort, A.B., Komorowski, R., Eichenbaum, H., Kopell, N., 2010. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. J. Neurophysiol. 104 (2), 1195–1210.

Tran, T.T., Rolle, C.E., Gazzaley, A., Voytek, B., 2020. Linked sources of neural noise contribute to age-related cognitive decline. J. Cogn. Neurosci. 32 (9), 1813–1822.

Trible, W., Kronauer, D.J., 2021. Hourglass model for developmental evolution of ant castes. Trends Ecol. Evol. 36 (2), 100–103.

Tschantz, A., Barca, L., Maisto, D., Buckley, C.L., Seth, A.K., Pezzulo, G., 2022. Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using active inference. Biol. Psychol. 169, 108266.

Uhl, J.T., Pathak, S., Schorlemmer, D., Liu, X., Swindeman, R., Brinkman, B.A., et al., 2015. Universal quake statistics: from compressed nanocrystals to earthquakes. Sci. Rep. 5 (1), 1–10.

van de Leemput, I.A., Wichers, M., Cramer, A.O., Borsboom, D., Tuerlinckx, F., Kuppens, P., et al., 2014. Critical slowing down as early warning for the onset and termination of depression. Proc. Natl. Acad. Sci. USA 111 (1), 87–92.

van Oers, K., Mueller, J.C., 2010. Evolutionary genomics of animal personality. Philos. Trans. R. Soc. B: Biol. Sci. 365 (1560), 3991–4000.

Veerakumar, A., Tiruvadi, V., Howell, B., Waters, A.C., Crowell, A.L., Voytek, B., et al., 2019. Field potential 1/f activity in the subcallosal cingulate region as a candidate signal for monitoring deep brain stimulation for treatment-resistant depression. J. Neurophysiol. 122 (3), 1023–1035.

Velarde, O.M., Urdapilleta, E., Mato, G., Dellavale, D., 2019. Bifurcation structure determines different phase-amplitude coupling patterns in the activity of biologically plausible neural networks. NeuroImage 202, 116031.

Voytek, B., Kramer, M.A., Case, J., Lepage, K.Q., Tempesta, Z.R., Knight, R.T., et al., 2015. Age-related changes in 1/f neural electrophysiological noise. J. Neurosci. 35 (38), 13257–13265.

Waite, A.J., Frankel, N.W., Emonet, T., 2018. Behavioral variability and phenotypic diversity in bacterial chemotaxis. Annu. Rev. Biophys. 47, 595.

Wang, S.-J., Hilgetag, C.C., Zhou, C., 2011. Sustained activity in hierarchical modular neural networks: self-organized criticality and oscillations. Front. Comput. Neurosci. 5, 30.

Wang, X.-J., Kennedy, H., 2016. Brain structure and dynamics across scales: in search of rules. Curr. Opin. Neurobiol. 37, 92–98.

Warner, M.R., Qiu, L., Holmes, M.J., Mikheyev, A.S., Linksvayer, T.A., 2019. Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. Nat. Commun. 10 (1), 1–11.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393 (6684), 440–442.

Weber, J., Klein, T., Abeln, V., 2020. Shifts in broadband power and alpha peak frequency observed during long-term isolation. Sci. Rep. 10 (1), 17987.

Wei, M., Qin, J., Yan, R., Li, H., Yao, Z., Lu, Q., 2013. Identifying major depressive disorder using Hurst exponent of resting-state brain networks. Psychiatry Res.: Neuroimaging 214 (3), 306–312.

Weissman, A., Binah, O., 2014. The fractal nature of blood glucose fluctuations. J. Diabetes Complicat. 28 (5), 646–651.

Wingfield J.C., 2006. Control of behavioural strategies for capricious environments. Essays in animal behaviour celebrating 50 years of Animal Behaviour. pp. 115–133.

Winn, J., Bishop, C.M., Jaakkola, T., 2005. Variational message passing. J. Mach. Learn. Res. 6 (4).

Woo, J.H., Honey, C.J., Moon, J.-Y., 2020. Phase and amplitude dynamics of coupled oscillator systems on complex networks. Chaos: Interdiscip. J. Nonlinear Sci. 30 (12), 121102.

Yu, H., Gerstein, M., 2006. Genomic analysis of the hierarchical structure of regulatory networks. Proc. Natl. Acad. Sci. USA 103 (40), 14724–14731.

Zhang, J., Xu, L., Cui, Z., 2021. Convergent developmental principles between Caenorhabditis elegans and human connectomes. Trends Cogn. Sci. 25 (12), 1015–1017.

Zhu, Z., Surujon, D., Ortiz-Marquez, J.C., Huo, W., Isberg, R.R., Bento, J., et al., 2020. Entropy of a bacterial stress response is a generalizable predictor for fitness and antibiotic sensitivity. Nat. Commun. 11 (1), 1–15.

Zmazek, J., Klemen, M.S., Markovič, R., Dolenšek, J., Marhl, M., Stožer, A., et al., 2021. Assessing different temporal scales of calcium dynamics in networks of beta cell populations. Front. Physiol. 12, 612233.