**RESEARCH ARTICLE**

# Formal perspectives on shared interbrain activity in social communication

## Insights from information and control theories

**Rodrick Wallace[1]**

### Abstract

The mechanisms underlying a reorientation of neuroscience from a single-brain to a multi-brain frame of reference have long been with us. These revolve around the evolutionary exaptation of the inevitable second-law 'leakage' of crosstalk between co-resident cognitive phenomena. Crosstalk characterizes such processes as immune response, wound-healing, gene expression, as so on, up through and including far more rapid neural processes. It is not a great leap-of-faith to infer that similar phenomena affect/afflict social interactions between individuals within and across populations.

**Keywords** Cognition · Crosstalk · Information theory · Neuroscience · Groupoid symmetry-breaking · Phase transition

People are embedded in social interaction that shapes their brains throughout lifetime. Instead of emerging from lower-level cognitive functions, social interaction could be the default mode via which humans communicate with their environment.

> — Hari et al. (2015)

The challenge for the study of brain-to-brain coupling is to develop detailed models of the dynamical interaction that can be applied at the behavioural levels and at the neural levels.

> — Hasson and Frith (2016)

...[A] deeper understanding of inter-brain dynamics may provide unique insight into the neural basis of collective behavior that gives rise to a broad range of economic, political, and sociocultural activities that shape society.

> — Kingsbury and Hong (2020)

✉ Rodrick Wallace
  rodrick.wallace@gmail.com;
  rodrick.wallace@nyspi.columbia.edu

[1] The New York State Psychiatric Institute, New York, USA

## Introduction

A recent elegant study of a bat population by Rose et al. (2021) finds that bidirectional interbrain activity patterns are a feature of their socially interactive behaviors, and that such shared interbrain activity patterns likely play an important role in social communication between group members. Sliwa (2021) summarizes that work, and parallel material by Baez-Mendoza et al. (2021) on macaques. Kingsbury et al. (2019), in a particularly deep analysis, studied correlations in brain activity between socially interacting mice, finding strong structuring by dominance relations.

Rose et al. are careful to cite the large and growing human literature on brain-to-brain coupling in social interaction, including Hasson et al. (2012); Barraza et al. (2020); Kuhlen et al. (2017); Perez et al. (2017); Stolk (2014), and Dikker (2017). As Rose et al. put it, a wide range of species naturally interact in groups and exhibit a diversity of social structures and forms of communication involving similarities and differences in neural repertoires for social communication.

On far longer time scales, Abraham et al. (2020) found that concordance in parent and offspring cortico-basal ganglia white matter connectivity varies by parental history of major depressive disorder and early parental care. As they put it,

Social behavior is transmitted cross-generationally through coordinated behavior within attachment bonds. Parental depression and poor parental care are major risks for disruptions of such coordination and are associated with offspring's psychopathology and interpersonal dysfunction... [Study] showed diminished neural concordance among dyads with a depressed parent and that better parental care predicted greater concordance, which also provided a protective buffer against attenuated concordance among dyads with a depressed parent... [Such] disruption may be a risk factor for intergenerational transmission of psychopathology. Findings emphasize the long-term role of early caregiving in shaping neural concordance among at-risk and affected dyads.

Indeed, a broad spectrum of Holocaust studies (Dashorst et al. 2019) has followed intergenerational transmission of psycho- and other pathologies, including, but not limited to, patterns of brain function.

Here, following Wallace (2022a, b), we examine 'shared interbrain activity patterns' from the perspectives of recent developments in control and information theories, using the asymptotic limit theorems of those disciplines to develop probability models that might be converted to statistical tools of value in future observational and empirical studies of the phenomena at various time scales. There is, after all, a very long tradition of using control theory ideas in psychological research. See the review by Henry et al. (2021) for a deep and cogent summary.

Shared interbrain activity patterns are concrete representations – indeed, instantiations – of information transmission within a group, and, as Dretske (1994) indicates, the properties of any transfer of information are strongly constrained by the asymptotic limit theorems of information theory, in the same sense that the Central Limit Theorem imposes constraints leading to useful statistical models of supposedly 'random' phenomena.

We begin with 'simple' correlation of brain activity between individuals in social interaction, and then move on to more complex models of joint cognition across individuals and/or 'workgroups', in a large sense.

## Correlation

The elegant paper by Kingsbury et al. (2019) explores correlated neural activity and the encoding of behavior across the brains of socially-interacting mice. Two central findings of that work are shown in Fig. 1, (as adapted from their Figs. 2 and 8). The top row of Fig. 1 shows time series of brain activity in two mice, first with, and then without, direct contact. Correlation between the signals is
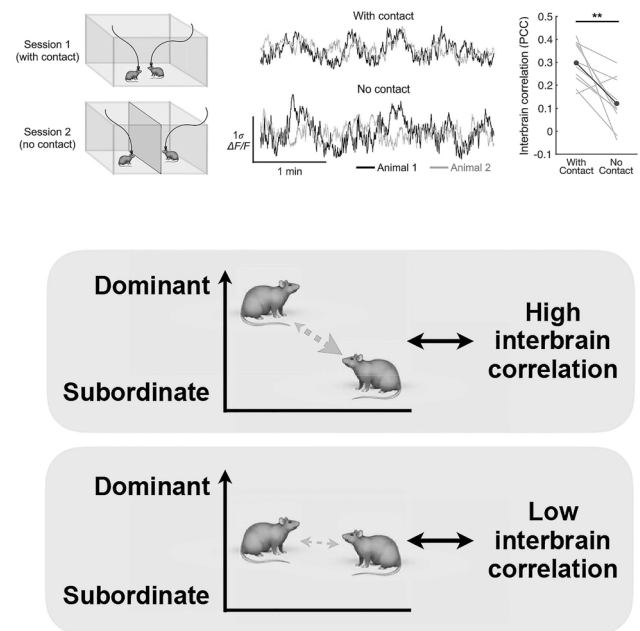


**Fig. 1** Adapted from Kingsbury et al. (2019), Figs. 1 and 8. Top: Correlation between brain activity is much higher during social interaction. Bottom: Correlation is much higher between discordantly dominant animals
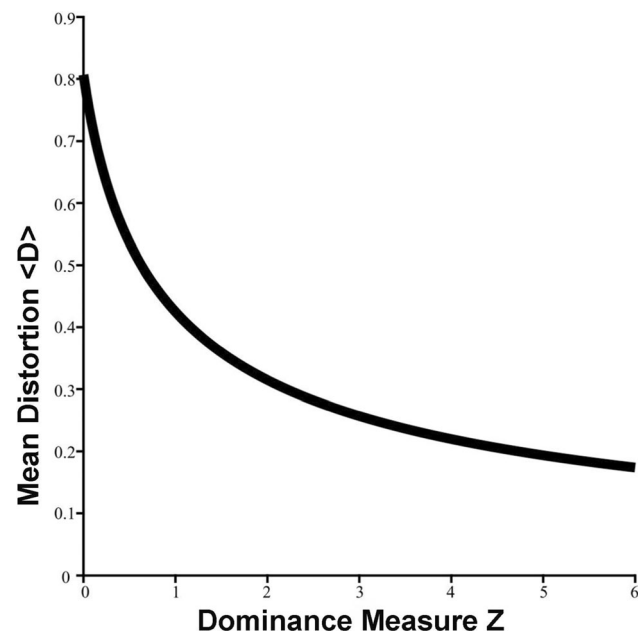


**Fig. 2** Mean distortion $<D>$ for the Gaussian channel vs. the dominance index $Z$. Higher status disjunction implies closer coupling during social interaction, according to this model

much higher during social interaction. The lower part indicates that correlations rise with difference in status between animals.

Here, we derive these results using perhaps the simplest possible dynamic model, one based on a principal

asymptotic limit theorem of information theory, the Rate Distortion Theorem.

We adapt the introductory model of Wallace (2020b, Sec. 1.3), focused on interacting institutions under conditions of conflict analogous to status disjunction.

Suppose we have developed – or been given – a robust scalar measure of social dominance between pairs of individuals, $Z$. How does $Z$ affect 'correlation', in a large sense, during interactions?

A 'dominant' partner transmits signals to a 'subordinate' in the presence of noise, sending a sequence of signals $U_i = \{u_1^i, u_2^i ...\}$ that, – again, in the presence of noise – is received as $\hat{U}^i = \{\hat{u}_1^i, \hat{u}_2^i, ...\}$. We take the $U^i$ as sent with probabilities $P(U^i)$, and define a scalar distortion measure between $U^i$ and $\hat{U}^i$ as $d(U^i, \hat{U}^i)$, defining an average distortion $D$ as

$$D \equiv \sum_i P(U^i) d(U^i, \hat{U}^i) \tag{1}$$

Following Wallace (2020b, Sec. 1.3) closely, it is then possible to apply a rate distortion argument. The Rate Distortion Theorem – for stationary, ergodic systems – states that there is a convex Rate Distortion Function that determines the minimum channel capacity, written $R(D)$, that is needed to keep the average distortion below the limit $D$. See Cover and Thomas (2006) for details. The theory can be extended, with some difficulty, to nonergodic sources via an infimum argument applied to the ergodic decomposition (e.g., Shields et al. 1978).

The 'central trick' is to construct a Boltzmann pseudo-probability in the dominance measure $Z$ as

$$dP(R, Z) = \frac{\exp[-R/g(Z)]dR}{\int_0^\infty \exp[-R/g(Z)]dR} \tag{2}$$

where the function $g(Z)$ is unknown and must be determined from first principles.

We have implicitly adopted the 'worst case coding' scenario of an analog 'Gaussian' channel (Cover and Thomas 2006). In consequence, the 'partition function' integral in the denominator of Eq. (2) has the simple value $g(Z)$.

For the Gaussian channel using the squared distortion measure, the Rate Distortion Function $R(D)$, and the distortion measure, are given as (Cover and Thomas 2006)

$$R(D) = \frac{1}{2}\log_2[\sigma^2/D]$$
$$D = \sigma^2 2^{-2R} \tag{3}$$

If $D \geq \sigma^2$, then $R = 0$.

From these relations, using Eq. (2), and after some manipulation, it is possible to calculate the 'average average distortion' $<D>$ as

$$<D> = \frac{\sigma^2}{\log(4)g(Z) + 1} \tag{4}$$

For the 'natural channel'

$$D = \frac{\sigma^2}{1 + R}$$
$$<D> = \frac{\sigma^2 e^{\frac{1}{g(Z)}} \mathrm{Ei}_1\left(\frac{1}{g(Z)}\right)}{g(Z)} \tag{5}$$

where $\mathrm{Ei}_1$ is the exponential integral of order 1.

What, then, is $g(Z)$?

Here we abduct more formalism from statistical mechanics, using the 'partition function' in Eq. (2) to define an 'iterated free energy' as

$$\exp[-F/g(Z)] = \int_0^\infty \exp[-R/g(Z)]dR = g(Z)$$
$$F = -\log[g(Z)]g(Z) \tag{6}$$
$$g(Z) = \frac{-F(Z)}{W(n, -F(Z))}$$

where $W(n, x)$ is the Lambert W-function of order $n$ solving the relation $W(n, x)\exp[W(n, x)] = x$ and is real-valued only for $n = 0, -1$ over limited ranges. These are, respectively, for $n = 0, x > -\exp[-1]$, and for $n = -1, -\exp[-1] < x < 0$. These conditions are important and impose themselves on the expressions for $<D>$.

The next step is to define an 'iterated entropy' in the standard manner as the Legendre transform of the 'free energy' analog $F$, and from it, impose a simple first-order version of the Onsager treatment of nonequilibrium thermodynamics (de Groot and Mazur 1984), in the context of a nonequilibrium steady state, so that $dZ/dt = 0$. The relations are then

$$S(Z) \equiv -F(Z) + ZdF(Z)/dZ$$
$$dZ/dt \propto dS/dZ = Zd^2F/dZ^2 = 0 \tag{7}$$
$$F(Z) = C_1 Z + C_2$$

where the $C_i$ are appropriate boundary conditions and $g(Z)$ is then given by the last expression in Eq. (6).

For the Gaussian channel, taking the Lambert W-function of order zero, and setting $C_1 = -1$, $C_2 = 1/10$, gives Fig. 2.

For the 'natural' channel, again with the W-function of order zero, taking the same values for the $C_i$ gives Fig. 3.

Higher status disjunction, in this model, implies closer coupling between individuals.

Similar results will follow for all possible Rate Distortion Functions that must be both convex in $D$ and zero-valued for $D \geq \sigma^2$ (Cover and Thomas 2006).

A fairly elementary model of social interaction under dominance relations, based on a somewhat counter-
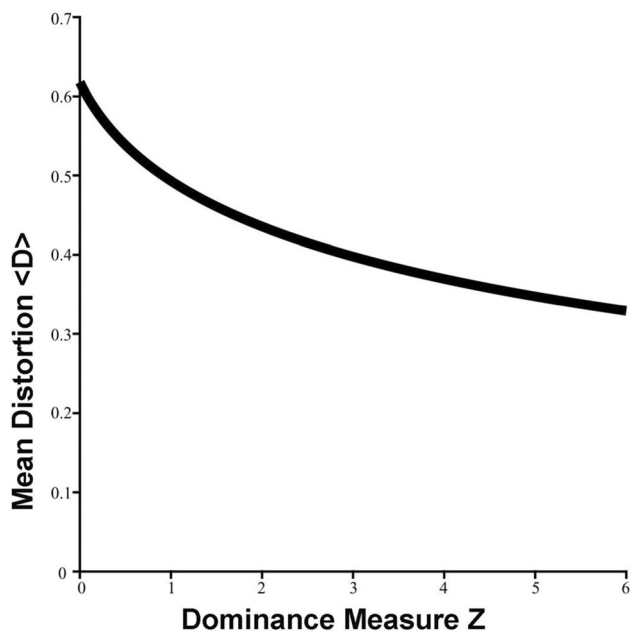
**Fig. 3** Mean distortion $<D>$ for the 'natural' channel vs. the dominance index Z. Again, higher status disjunction implies closer coupling during social interaction

intuitive dynamic adaptation of the Rate Distortion Theorem, produces results consistent with the empirical observations of Kingsbury et al. (2019). Extension of the model to both nonergodic and nonstationary conditions more likely to mirror real-world conditions, or at least going beyond the nonequilibrium steady state assumption, will requires further work.

The model can, in theory, be extended using the ergodic decomposition of a nonergodic process using the methods of Shields et al. (1978).

Otherwise, we are constrained to adiabatically, piecewise, stationary ergodic (APSE) systems that remain as close to ergodic and stationary as needed for the Rate Distortion Theorem to work, much like the Born-Oppenheimer approximation of molecular physics in which rapid electron dynamics are assumed to quasi-equilibrate about much slower nuclear oscillations, allowing calculations using 'simple' quantum mechanics models.

## Cognition

Cognition is not correlation, and requires more general address. Indeed, cognition has become a kind of shibboleth in theoretical biology, seen by some as the fundamental characterization of the living state at and across its essential scales and levels of organization (Maturana and Varela 1980). In this regard, a central inference by Atlan and Cohen (1998), in their study of the immune system, is that

cognition, via mechanisms of choice, demands reduction in uncertainty, implying the existence of information sources 'dual' to any cognitive process. The argument is unambiguous and direct, and serves as the foundation to our general approach.

A first step is to view 'information' as a biological and social resource matching the importance of metabolic free energy and other overtly material agents and agencies.

## Information and other resources

Here, we must move beyond 'simple' measures of dominance between individuals in social interaction.

At least three resource streams are required by any cognitive entity facing real-time, real-world challenges. The first is measured by the rate at which information can be transmitted between elements within the entity, determined as an information channel capacity, say $\mathscr{C}$ (Cover and Thomas 2006). The second resource stream is sensory information regarding the embedding environment – here, primarily social interaction – available at a rate $\mathscr{D}$. It is along this channel that 'neural representations' will be shared.

The third regards material resources, including metabolic free energy – in a large sense – available at a rate $\mathscr{M}$.

These three rates may well – but not necessarily – interact, a matter characterized as a 3 by 3 matrix analogous to, but not the same as, a simple correlation matrix. Let us write this as **Z**.

An $n$-dimensional square matrix has $n$ scalar invariants $r_i$ defined by the relation

$$p(\gamma) = \det[\mathbf{Z} - \gamma \mathbf{I}] =$$
$$\gamma^n - r_1 \gamma^{n-1} + r_2 \gamma^{n-2} - ... + (-1)^n r_n \tag{8}$$

**I** is the $n$-dimensional identity matrix, det the determinant, and $\gamma$ a real-valued parameter. The first invariant is usually taken as the matrix trace, and the last as $\pm$ the determinant.

These scalar invariants make it possible to project the full matrix down onto a single scalar index $Z = Z(r_1, ..., r_n)$ retaining much of the basic structure, analogous to conducting a principal component analysis. The simplest index might be $Z = \mathscr{C} \times \mathscr{D} \times \mathscr{M}$ – the matrix determinant for a system without crossinteraction. However, scalarization must be appropriate to the individual circumstance studied, and there will almost always be important cross-interactions between resource streams.

Clever scalarization, however, enables approximate reduction to a one dimensional system.

Taking $\mathscr{M}$ out of the equation – equalizing it across the social structure – might generate two independent 'orthogonal' indices, for example the determinant and the

trace of the 'interaction matrix' separately, so that $Z$ becomes a two dimensional vector. General expansion of $Z$ into vector form leads to difficult multidimensional dynamic equations (e.g., Wallace 2021c). See the Mathematical Appendix for details.

## Cognition and information

Only in the case that cross-sectional and longitudinal means are the same can information source uncertainty be expressed as a conventional Shannon 'entropy' (Khinchin 1957; Cover and Thomas 2006). Here, we only require that source uncertainties converge for sufficiently long paths, not that they fit some functional form. It is the values of those uncertainties that will be of concern so that we study 'Adiabatically Piecewise Stationary' (APS) systems, in the sense of the Born-Oppenheimer approximation for molecular systems that assume nuclear motions are so slow in comparison with electron dynamics that they can be effectively separated, at least on appropriately chosen trajectory 'pieces' that may characterize the various phase transitions available to such systems. Extension of this work to nonstationary circumstances remains to be done.

This approximation can be carried out via a fairly standard Morse Function iteration (Pettini 2007).

The systems of interest here are composed of cognitive submodules that engage in crosstalk. At every scale and level of organization all such submodules are constrained by both their own internals and developmental paths and by the persistent regularities of the embedding environment, including the cognitive intent of colleagues, in a broad sense, and the regularities of 'grammar' and 'syntax' imposed by the embedding social structure.

Further, there are structured uncertainties imposed by the large deviations possible within that environment, again including the behaviors of adversaries who may be constrained by quite different developmental trajectories and 'punctuated equilibrium' evolutionary transitions.

Recapitulating somewhat the arguments of Wallace (2018, 2020a), the Morse Function construction assumes a number of interacting factors:

- As Atlan and Cohen (1998) argue, cognition requires choice that reduces uncertainty. Such reduction in uncertainty directly implies the existence of an information source 'dual' to that cognition at each scale and level of organization. The argument is unambiguous and sufficiently compelling.
- Cognitive physiological processes, like the immune and gene expression systems, are highly regulated, in the same sense that 'the stream of consciousness' flows between cultural and social 'riverbanks'. That is, a

cognitive information source $X_i$ is generally paired with a regulatory information source $X^i$.

- Environments (in a large sense), also have sequences of very high and very low probability: night follows day, hot seasons follow cold, and so on.
- 'Large deviations', following Champagnat et al. (2006) and Dembo and Zeitouni (1998), also involve sets of high probability developmental pathways, often governed by 'entropy' analog laws that imply the existence of yet one more information source.

Full system dynamics must then be characterized by a joint, nonergodic information source uncertainty

$$H(\{X_i, X^i\}, V, L_D) \tag{9}$$

that is defined path-by-path and not represented as an 'entropy' function (Khinchin 1957). Consequently, each path will have it's own $H$-value, but the functional form of that value is not specified in terms of underlying probability distributions.

The set $\{X_i, X^i\}$ includes the internal interactive cognitive dual information sources of the system of interest and their associated regulators, $V$ is taken as the information source of the embedding environment. This may include the actions and intents of adversaries/symbionts/colleagues, as well as 'weather'. $L_D$ is the information source of the associated large deviations possible to the system, possibly including 'punctuated equilibrium' evolutionary transitions.

Again, we are projecting the spectrum of essential resources onto a scalar rate $Z$.

The underlying equivalence classes of developmental or dynamic system paths used to define groupoid symmetries can be defined fully in terms of the magnitude of individual path source uncertainties, $H(x^j)$, $x^j \equiv \{x_0^j, x_1^j, ... x_n^j, ...\}$) alone.

Recall the central conundrum of the ergodic decomposition of nonergodic information sources. It is formally possible to express a nonergodic source as the composition of a sufficient number of ergodic sources, much as it is possible to reduce planetary orbits to a Fourier sum of circular epicycles, obscuring the basic dynamics. Hoyrup (2013) discusses the problem further, finding that ergodic decompositions are not necessarily computable. Here, we need focus only on the values of the source uncertainties associated with dynamic paths.

## Dynamics

The next step is to build an iterated 'free energy' Morse Function (Pettini 2007) from a Boltzmann pseudoprobability, based on enumeration of high probability developmental pathways available to the system $j = 1, 2, ...$ so that

$$P_j = \frac{\exp[-H_j/g(Z)]}{\sum_k \exp[-H_k/g(Z)]} \qquad (10)$$

where $H_j$ is the source uncertainty of the path $j$, which – again – we do not assume to be given as a 'Shannon entropy' since we are no longer restricted to ergodic sources.

The essential point is the ability to divide individual paths into two equivalence classes, a small set of high probability paths consonant with an underlying 'grammar' and 'syntax', and a much larger set of vanishingly low probability, essentially a set of measure zero.

The temperature-analog characterizing the system, written as $g(Z)$ in Eq. (10), can be calculated – or at least approximated – via a first-order Onsager nonequilibrium thermodynamic approximation built from the partition function, i.e., the denominator of Eq. (10) (de Groot and Mazur 1984).

We define the 'iterated free energy' Morse Function as

$$\exp[-F/g(Z)] \equiv \sum_k \exp[-H_k/g(Z)] \equiv h(g(Z))$$
$$F(Z) = -\log[h(g(Z))]g(Z) \qquad (11)$$

where the sum is over all possible high probability developmental paths of the system, again, those consistent with an underlying grammar and syntax. Again, system paths not consonant with grammar and syntax constitute a set of measure zero that is very much larger than the set of high probability paths.

Feynman (2000) makes the direct argument that information itself is to be viewed as a form of free energy, using Bennett's 'ideal machine' that turns a message into work. Here, we invoke an iterated – rather than a direct – free energy construction.

$F$, taken as a free energy, then becomes subject to symmetry-breaking transitions as $g(Z)$ varies (Pettini 2007). These symmetry changes, however, are not as associated with physical phase transitions as represented by standard group algebras. Cognitive phase changes involve shifts between equivalence classes of high probability developmental pathways to be represented as *groupoids*, a generalization of the group concept where a product is not necessarily defined for every possible element pair (Brown 1992; Cayron 2006; Weinstein 1996). See the Mathematical Appendix for an outline of the theory.

The disjunction described above – into high and low probability equivalence classes representing paths consonant with, or discordant from, underlying grammar and syntax – should be seen as the primary 'groupoid phase transition' affecting cognitive systems. It is essentially the biological 'big bang' of Maturana and Varela (1980). Think about this carefully.

Dynamic equations follow from invoking a first-order Onsager approximation akin to that of nonequilibrium thermodynamics (de Groot and Mazur 1984) in the gradient of an entropy measure constructed from the 'iterated free energy' $F$ of Eq. (11):

$$S(Z) \equiv -F(Z) + ZdF(Z)/dZ$$
$$\partial Z/\partial t \approx dS/dZ = f(Z)$$
$$f(Z) = Zd^2F/dZ^2$$
$$g(Z) = \qquad (12)$$

$$\frac{-C_1 Z - \left(\int \frac{f(Z)}{Z} dZ\right)Z + C_2 + \int f(Z)dZ}{RootOf\left(e^Q - h\left(-\frac{C_1 Z + \left(\int \frac{f(Z)}{Z} dZ\right)Z - C_2 - \left(\int f(Z)dZ\right)}{Q}\right)\right)}$$

where the last relation follows from an expansion of the third part of Eq. (12) using the second expression of Eq. (11). $C_1$ and $C_2$ are two constants in the indefinite integral of the second derivative of $F(Z)$, and $Q$ is the independent variable of the function being taken roots.

Three important – and somewhat subtle – points:

(1)   The 'RootOf' construction actually generalizes the not-so-well-known Lambert W-function (e.g., Yi et al. 2010; Mezo and Keady 2015). This leads to deep waters: since 'RootOf' may have complex number solutions, the temperature analog $g(Z)$ enters the realm of the 'Fisher Zeros' characterizing phase transition in physical systems (e.g., Dolan et al. 2001; Fisher 1965; Ruelle 1964 Sec. 5).

(2)   Information sources are not microreversible, that is, palindromes are highly improbable, e.g., ' eht ' has far lower probability than ' the ' in English. In consequence, there are no 'Onsager Reciprocal Relations' in higher dimensional systems. The necessity of groupoid symmetries appears to be driven by this directed homotopy.

(3)   Typically, it is necessary to impose a delay in provision of Z, so that, for example, $dZ/dt = f(Z) = \beta - \alpha Z(t)$ and $Z \to \beta/\alpha$ at a rate determined by $\alpha$.

Suppose, in the first of Eq. (11), it is possible to approximate the sum with an integral, so that

$$\exp[-F/g(Z)] \approx \int_0^\infty \exp[-H/g(Z)]dH = g(Z) \qquad (13)$$

$g(Z)$ must be real-valued and positive. Then

$$F(Z) = -\log[g(Z)]g(Z)$$
$$g(Z) = -F(Z)/W_L[n, -F(Z)] \qquad (14)$$

Again, $W_L$ is the 'simplest' Lambert W-function that satisfies $W_L[n,x]\exp[W_L[n,x]] = x$. It is real-valued only for $n = 0, -1$ and only over limited ranges of $x$ in each case.

In theory, specification of any two of the functions $f, g, h$ permits calculation of the third. $h$, however, is determined – fixed – by the internal structure of the larger system. Similarly, 'boundary conditions' $C_1, C_2$ are externally-imposed, further sculpting dynamic properties of the 'temperature' $g(Z)$, and $f$ determines the rate at which the composite essential resource $Z$ can be delivered. Both information and metabolic free energy resources are rate-limited.

## Cognition rate

For phase transitions in physical systems, there is generally a minimum temperature for punctuated activation of the dynamics associated with given group structure – underlying symmetry changes associated with the transitions of ice to water water to steam, and so on. For cognitive processes, following the arguments of Eq. (5), there will be a minimum necessary value of $g(Z)$ for onset of the next in a series of transitions. That is, at some $T_0 \equiv g(Z_0)$, having a corresponding information source uncertainty $H_0$, a second groupoid phase transition becomes manifest.

Taking a reaction rate perspective from chemical kinetics (Laidler 1987), we can write an expression for the rate of cognition as

$$L(Z) = \frac{\sum_{H_j > H_0} \exp[-H_j/g(Z)]}{\sum_k \exp[-H_k/g(Z)]} \quad (15)$$

If the sums can be approximated as integrals, then the system's rate of cognition at resource rate index $Z$ can be written as

$$L(Z) \approx \frac{\int_{H_0}^{\infty} \exp[-H/g(Z)]dH}{\int_0^{\infty} \exp[-H/g(Z)]dH} = \exp[-H_0/g(Z)] \\ = \exp[H_0 W_L(n, -F)/F] \quad (16)$$

where $W_L(n, -F)$ is the Lambert W-function of order $n$ in the free energy index $F = -\log[g(Z)]g(Z)$.

Figure 4 shows $L(F) vs. F$, using Lambert W-functions of orders 0 and $-1$, respectively real-valued only on the intervals $x > -\exp[-1]$ and $-\exp[-1] < x < 0$.

The Lambert W-function is only real-valued for orders 0 and $-1$, and only for $F < \exp[-1]$. However, if $0 < F < \exp[-1]$, then a bifurcation instability emerges, with a transition to complex-valued oscillations in cognition rate at higher values.

This development recovers what is essentially an analog to the Data Rate Theorem from control theory (Nair et al. 2007 and the Mathematical Appendix), in the sense that the requirement $H > H_0$ in Eqs.(15) and (16) imposes stability constraints on $F$, the free energy analog, and by inference, on the resource rate index $Z$ driving it.
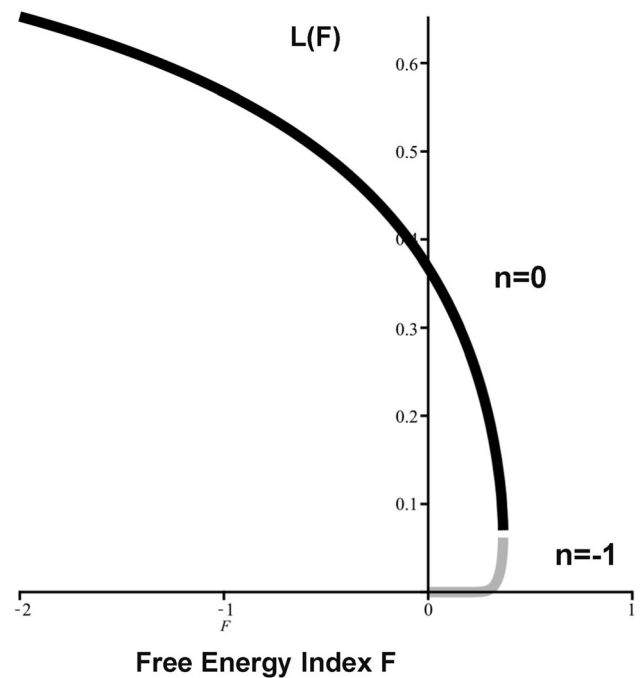


**Fig. 4** Rate of cognition from Eq. (16) as a function of the iterated free energy measure $F$, taking $H_0 = 1$. The Lambert W-function is only real-valued for orders 0 and $-1$, and only if $F < \exp[-1]$. However, if $0 < F < \exp[-1]$, then a bifurcation instability emerges, with a transition to complex-valued oscillations in cognition rate at higher values. Since $F$ is driven by $Z$, there is a minimum resource rate for stability

## An example

We again approximate the sum in Eq. (11) by an integral – so that $h(g(Z)) = g(Z)$ – and make a simple assumption on the form of $dZ/dt = f(Z)$, say $f(Z) = \beta - \alpha Z(t)$. Then

$$g(Z) = -\frac{2\ln(Z)Z\beta - Z^2\alpha + 2C_1 Z - 2Z\beta + 2C_2}{2W_L\left(n, -\ln(Z)Z\beta + \frac{Z^2\alpha}{2} - C_1 Z + Z\beta - C_2\right)} \quad (17)$$

with, again, $L = \exp[-H_0/g(Z)]$, depending on the rate parameters $\alpha$ and $\beta$, the boundary conditions $C_i$, and the degree of the Lambert W-function. Proper choice of boundary conditions generates a classic 'inverted-U' signal transduction Yerkes-Dodson Law analog (e.g., Wallace 2020a, 2021c; Diamond et al. 2007). That is, since $Z \to \beta/\alpha$, we look at the cognition rate for fixed $\alpha$ and boundary conditions $C_j$ as $\beta$ increases. The result is shown in Fig. 5, for appropriate boundary conditions.

Similar results follow if

$$\exp[-F/g(Z)] = h(g(Z)) \propto A_m g(Z)^m$$

that is, if the function $h(g(Z))$ has a strongly dominant term of order $m > 0$.

It is complicated – but not difficult – to incorporate stochastic effects into cognition rate dynamics based on Eq. (17), via standard methods from the theory of stochastic differential equations (e.g., Wallace 2021a, b, c). See the Mathematical Appendix for an example based on Fig. 5, fixing $\beta = 3$.

## Cooperation: multiple workspaces

Individual brains – and indeed, even individual cells – are composed of interacting (often spatially distributed) cognitive submodules. Social groups are constituted by interacting individuals, separated by space, time, and/or social distance. Institutions are made up of dispersed but interacting 'workgroups', in a large sense. A critical phenomenon in all such examples is that the joint uncertainty of the dual information source associated with the particular level of cognition is less than or equal to the sum of the uncertainties of independent components – the information theory chain rule (Cover and Thomas 2006). To invert the argument, preventing crosstalk between cognitive submodules requires more investment of free energy or other resources than allowing interaction, as the electrical engineers often lament. From this lemon, evolution has made lemonaide (e.g., Wallace 2022).

More specifically for the work here, the emergence of a generalized Lambert W-function in Eq. (12), reducing to a 'simple' W-function if $h(g(Z)) = g(Z)$, is particularly suggestive. Recall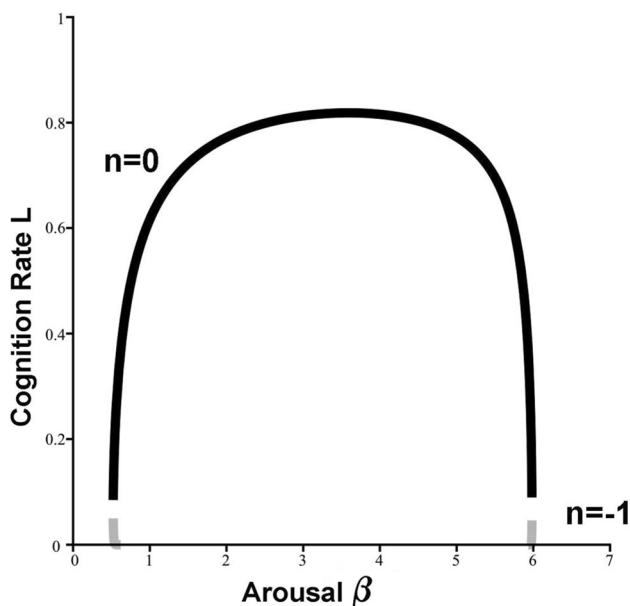 that the fraction of network nodes in a giant component of a random network of $N$ nodes with probability $P$ of linkage between them can be given as (Newman 2010)

$$\frac{W_L(0, -NP \exp[-NP]) + NP}{NP} \tag{18}$$

where, again, the Lambert W-function emerges.

This expression has punctuated onset of a giant component of linked nodes only for $NP > 1$. See Fig. 6. In general, we might expect $P$ to be a monotonic increasing function of $g(Z)$.

Within broadly 'social' groupings, interacting cognitive submodules – individuals – can become linked into shifting, tunable, temporary, workgroup equivalence classes to address similarly rapidly shifting patterns of threat and opportunity. These might range from complicated but relatively slow multiple global workspace processes of gene expression and immune function to the rapid – hence necessarily stripped-down – single-workspace neural phenomena of higher animal consciousness (Wallace 2012).

More complicated approaches to such a phase transition – involving Kadanoff renormalizations of the Morse Function free energy measure $F$ and related measures – can be found in Wallace (2005; 2012; 2022).

By contrast here, while multiple workspaces are most simply invoked in terms of a simultaneous set of the $g_j(Z_j)$, individual workspace tunability emerges from exploring equivalence classes of network topologies associated with
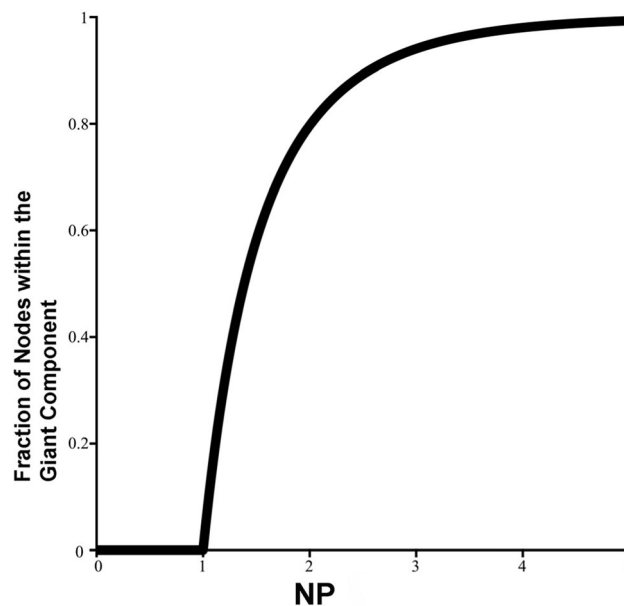


**Fig. 5** Classic 'inverted-U' signal transduction for the cognition rate based on Eq. (17), setting $\alpha = 1$, $C_1 = -2$, $C_2 = -2$, $H_0 = 1$. Increase in $\beta$ is taken as the 'arousal' measure



**Fig. 6** Fraction of an $N$-node random network within the giant component as determined by the probability of contact between nodes $P$. The essential point is the punctuated accession to 'global broadcast' if and only if $NP > 1$ (e.g., Baars 1989; Dehaene and Changeux 2011). We might expect $P$ to be a monotonic increasing function of $g(Z)$

a particular value of some designated $g_j(Z_j)$. Central matters then revolve around the equivalence class decompositions implied by the existence of the resulting workgroups, leading again to dynamic groupoid symmetry-breaking as they shift form and function in response to changing patterns of threat and opportunity. See Wallace (2021a, Sec. 12) for a parallel argument from an ergodic system perspective. Here, the $g(Z)$ substitute for the 'renormalization constant' $\omega$ in that development. For brevity, we omit a full discussion here.

## Network topology is important

We can, indeed, calculate the cognition rate of a single 'global workspace' as follows.

Suppose we have $N$ linked cognitive submodules – individuals – in a social 'giant component', i.e., we operate in the upper portion of Fig. 6. The free energy Morse Function $F$ can be expressed in terms of a full-bore partition function as

$$\exp[-F/g(Z)] = \sum_{k=1}^{N} \sum_{j=1}^{M} \exp[-H_{k,j}/g(Z)]$$

$$\approx \sum_{k=1}^{N} \int_0^\infty \exp[-H_k/g(Z)]dH_k = \sum_{k=1}^{N} g(Z) = Ng(Z)$$

$$F = -\log[Ng(Z)]g(Z), \quad g(Z) = \frac{-F}{W_L(n, -NF)}$$

(19)

The sum over $j$ represents available states within individual submodules, and the sum over $k$ is across submodules, and $W_L$ is again the Lambert W-function. Note, however, the appearance of the factor $NF$ in the last expression.

The rate of cognition of the linked-node giant component can be expressed as

$$L = \frac{\sum_{k=1}^{N} \int_{H_0^k}^\infty \exp[-H_k/g(Z)]dH_k}{\sum_{k=1}^{N} \int_0^\infty \exp[-H_k/g(Z)]dH_k}$$

$$= \frac{g(Z) \sum_k \exp[-H_0^k/g(Z)]}{Ng(Z)}$$

(20)

$$= \left(\sum_k \exp[-H_0^k/g(Z)]\right)/N \equiv <L_k>$$

where, not entirely unexpectedly, $<L_k>$ represents an averaging operation. More sophisticated averages might well be applied – at the expense of more formalism.

If we impose the approximation of Onsager nonequilibrium thermodynamics, i.e., defining $S(Z) = -F(Z) + ZdF/dZ$, and assume $dS/dZ = f(Z)$, we again obtain $f(Z) = Zd^2F/dZ^2$, and can calculate $g(Z)$ from the last expression in Eq. (19).

The appearance of $N$ in expressions for $g(Z)$ and $L$ is of some note. In particular, groupthink – high values of $H_0^k$ – may result in failure to detect important signals.

Such considerations lead to a fundamentally different picture that is not just possible, but often observed, i.e., a transmission model. Consider something like a colony of prairie dogs under predation by hawks. A 'sentinel' pattern emerges, rather than the average system of Eq. (20), through the 'bottleneck' of a single, very highly-optimized, 'subcomponent' of the larger social structure. This involves a hypervigilant individual – or a small number of such appropriately dispersed – whose special danger signal or signals can be imposed rapidly across the entire population – postures, vocalizations, or both. Then

$$L = \max_k \{L_k = \exp[-H_0^k/g(Z)]\}$$

(21)

where max is the maximization across the set $\{L_k\}$. However, Dore et al. (2019) show that, while brain activity in humans can track information sharing, there are likely to be important individual differences. That is, not everyone may respond the same to an 'alert' message.

Other network topologies are clearly possible. For example, in a rigidly hierarchical linear-chain business, military, or political setting, $L$ will often be given by the *minimization* of cognition rates across the $L_k$, that is, by a bottleneck model.

The critical dependence of a system's cognitive dynamics on its underlying 'social topology' has profound implications for theories of the 'extended conscious mind' (Clark 2009; Lucia Valencia and Froese 2020). A particular expression of such matters involves failures of institutional cognition on wickedly hard problems (Wallace 2021b).

## Time and resource constraints are important

Multiple workspaces, however, can also present a singular – and independent – problem of resource delivery, taking the scalar $Z_j$ and *time itself* as essential resources. That is, not only are the $Z_j$ both limited and delayed, in most cases, there will be a limit on possible response times, for both 'predator' and 'prey', so to speak.

If we assume an overall limit to available resources across a multiple workspace system $j = 1, 2, \dots$ as $Z = \sum_j Z_j$, and available time as $T = \sum_j T_j$, then it is possible to carry out a simple Lagrangian optimization on the rate of system cognition $\sim \sum_j \exp[-H_j^0/g_j(Z_j)]$ as

$$\mathscr{L} \equiv \sum_j \exp[-H_j^0/g_j(Z_j)]+$$

$$\lambda \left( Z - \sum_j Z_j \right) + \mu \left( T - \sum_j T_j \right) \tag{22}$$

$$\partial \mathscr{L}/\partial Z_j = 0, \ \ \partial \mathscr{L}/\partial T_j = 0$$

where we assume $dZ_j/dt = f_j(Z_j(t)) = \beta_j - \alpha_j Z_j(t)$.

This leads, after some development, to a necessary expression for individual subsystem resource rates as

$$Z_j = f_j^{-1}(\mu/\lambda) = \frac{\beta_j - \mu/\lambda}{\alpha_j} > 0 \tag{23}$$

$\mu$ and $\lambda$ are to be viewed in economic terms as *shadow prices* imposed by 'environmental constraints', in a large sense (e.g., Jin et al. 2007; Robinson 1993).

The essential point is that cognitive dynamics – the rates of cognition driven by the rate of available resources $Z_j$ and time – in this model, are strongly determined by the shadow price ratio $\mu/\lambda$. The shadow price ratio is to be interpreted as an environmental signal. A sufficiently large shadow price ratio, according to Eq. (23), can starve essential components, driving their cognition rates to failure.

## Further theoretical development

Following the arguments of Wallace (2021a), the cognition models can be extended in a number of possible directions, much as is true for 'ordinary' regression theory.

Perhaps the simplest next step is to replace the relation $dZ/dt = f(Z(t))$ with a stochastic differential equation having the form $dZ_t = f(Z_t)dt + \sigma g(Z_t)dB_t$, where $dB_t$ represents ordinary white noise.

A next 'simple' generalization might be replacing the scalar index $Z$ with a multidimensional vector quantity, $\mathbf{Z}$, leading to an intricate set of simultaneous partial differential equations requiring, at best, Lie symmetry address. See the Mathematical Appendix for an outline.

Further development could involve expanding the 'Onsager approximation' in terms of a 'generalized entropy' $S = \sum_k \epsilon_k Z^{k-1} F^{k-1}$, where $F^j \equiv d^j F/dZ^j$. Then the dynamics might also be generalized, at least for a scalar $Z$, as $\partial Z/\partial t \approx \sum_j \mu_j d^j S/dZ^j = f(Z)$, and so on toward multidimensional models.

As with regression equations much beyond $Y = mX + b$, matters can rapidly become complicated indeed.

## Discussion

Hasson et al. (2012), in a widely-cited foundational study, call for a reorientation of neuroscience from a single-brain to a multi-brain frame of reference:

> Cognition materializes in an interpersonal space. The emergence of complex behaviors requires the coordination of actions among individuals according to a shared set of rules. Despite the central role of other individuals in shaping our minds, most cognitive studies focus on processes that occur within a single individual. We call for a shift from a single-brain to a multi-brain frame of reference. We argue that in many cases the neural processes in one brain are coupled to the neural processes in another brain via the transmission of a signal through the environment. Brain-to-brain coupling constrains and simplifies the actions of each individual in a social network, leading to complex joint behaviors that could not have emerged in isolation.

Kingsbury et al. (2019), Rose et al. (2021) and Baez-Mendoza et al. (2021) extend this perspective to interacting non-human populations, while Abraham et al. (2020), for humans, extend the time scale across generations.

Here, we outline something of the formal developments needed to implement such reorientation, based on groupoid symmetry-breaking within longstanding paradigms of information and control theories, as affected and afflicted by network topologies and their dynamics. This, it can be argued, is very much a 'rocket science' problem, since the difficulty lies not in the individual components of a possible comprehensive approach, which are all well-studied, but in using the building blocks to construct a theoretical enterprise that accounts well for observational and experimental data. Very similar conundrums confront contemporary theories of consciousness (e.g., Wallace 2022), albeit, in this case, without the hindrance of quite so many longstanding philosophical and de-facto theological presuppositions.

Application of the approach developed here to stochastic systems is straightforward if somewhat complicated, as is extension to higher approximation 'Onsager-type' entropy gradient models, the analog of moving from $y = mx + b$ to $y = mx^2 + b$, and so on (Wallace 2021a, b).

In a deep-time sense, the underlying mechanisms have long been with us, i.e., the evolutionary exaptation of the inevitable second-law 'leakage' of crosstalk between co-resident cognitive processes (e.g., Wallace 2012). Crosstalk characterizes the immune system, wound-healing, tumor control, gene expression, and so on, up through and including far more rapid neural processes. It is not a great

leap-of-faith to infer that similar dynamics instantiate social interactions between individuals within and across populations.

## Mathematical appendix

### Groupoids

We follow Brown (1992) closely. Consider a directed line segment in one component, written as the source on the left and the target on the right.

$\bullet \longrightarrow \bullet$

Two such arrows can be composed to give a product **ab** if and only if the target of **a** is the same as the source of **b**

$\bullet \xrightarrow{\mathbf{a}} \bullet \xrightarrow{\mathbf{b}} \bullet$

Brown puts it this way,

> One imposes the geometrically obvious notions of associativity, left and right identities, and inverses. Thus a groupoid is often thought of as a group with many identities, and the reason why this is possible is that the product **ab** is not always defined.
>
> We now know that this apparently anodyne relaxation of the rules has profound consequences... [since] the algebraic structure of product is here linked to a geometric structure, namely that of arrows with source and target, which mathematicians call a directed graph.

Cayron (2006) elaborates this as follows,

> A group defines a structure of actions without explicitly presenting the objects on which these actions are applied. Indeed, the actions of the group G applied to the identity element e implicitly define the objects of the set G by ge = g; in other terms, in a group, actions and objects are two isomorphic entities. A groupoid enlarges the notion of group by explicitly introducing, in addition to the actions, the objects on which the actions are applied. By this approach, many identities may exist (they correspond to the actions that leave an object invariant).

It is of particular importance that equivalence class decompositions permit construction of groupoids in a highly natural manner.

Weinstein (1996) and Golubitsky and Stewart (2006) provide more details on groupoids and on the relation between groupoids and bifurcations.

An essential point is that, since there are no necessary products between groupoid elements, 'orbits', in the usual sense, disjointly partition groupoids into 'transitive' subcomponents.

## The data rate theorem

Real-world environments are inherently unstable. Organisms, to survive, must exert a considerable measure of control over them. These control efforts range from immediate responses to changing patterns of threat and affordance, through niche construction, and, in higher animals, elaborate, highly persistent, social and sociocultural structures. Such necessity of control can, in some measure, be represented by a powerful asymptotic limit theorem of probability theory different from, but as fundamental as, the Central Limit Theorem: the Data Rate Theorem, first derived as an extension of the Bode Integral Theorem of signal theory.

Consider a reduced model of a control system as follows:

For the inherently unstable system of Fig. 7, assume an initial $n$-dimensional vector of system parameters at time $t$, as $x_t$. The system state at time $t + 1$ is then – near a presumed nonequilibrium steady state – determined by the first-order relation

$$x_{t+1} = \mathbf{A}x_t + \mathbf{B}u_t + W_t \tag{24}$$

In this approximation, **A** and **B** are taken as fixed $n$-dimensional square matrices. $u_t$ is a vector of control information, and $W_t$ is an $n$-dimensional vector of Brownian white noise.
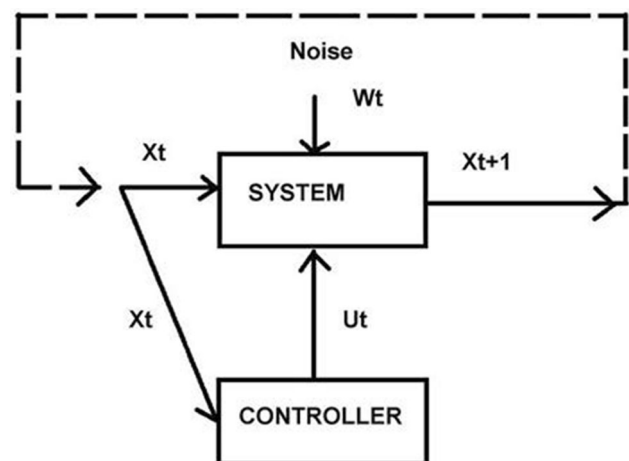


**Fig. 7** The reduced model of an inherently unstable system stabilized by a control signal $U_t$

According to the DRT, if $H$ is a rate of control information sufficient to stabilize an inherently unstable control system, then it must be greater than a minimum, $H_0$,

$$H > H_0 \equiv \log[\| \det[\mathbf{A}^m]\|] \quad (25)$$

where det is the determinant of the subcomponent $\mathbf{A}^m$ – with $m \le n$ – of the matrix $\mathbf{A}$ having eigenvalues $\ge 1$. $H_0$ is defined as the rate at which the unstable system generates 'topological information' on its own.

If this inequality is violated, stability fails.

## Stochastic analysis for Fig. 5

Here, we apply the Ito Chain Rule (Protter 2005) to the expression $L(Z) = \exp[-1/g(Z)]$, as based on Eq. (17) for $g(Z)$. We set $\alpha = 1$, $\beta = 3$, $C_1 = C_2 = -2$ and numerically calculate the solution set for the relation $<dL_t> = 0$ based on the underlying stochastic differential equation

$$dZ_t = (\beta - \alpha Z_t)dt + \sigma Z_t dB_t \quad (26)$$

where $dB_t$ is assumed to be ordinary white noise, as associated with Brownian motion.

Figure 8 shows the result, the solution equivalance class $\{\sigma, Z\}$ for $\beta = 3$, just to the left of the peak in Fig. 5. Note that, at $\sigma \approx 0.278$, the system becomes susceptible to a bifurcation instability, well before the 'standard' instability expected from a second-order Ito Chain Rule analysis
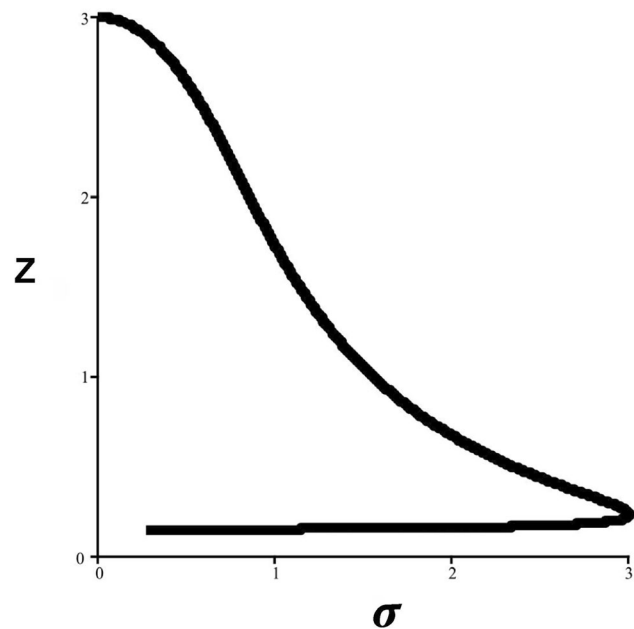


Fig. 8 Numerical solution equivalence class $\{\sigma, Z\}$ for the relation $<dL_t> = 0$ from Fig. 5, taking $\beta = 3$, just to the left of the peak. Again, $\alpha = 1$, $C_1 = C_2 = -2$. While, in this model, $Z_t$ becomes unstable in variance at $\sigma > \sqrt{2}$, the cognition rate can suffer a bifurcation instability for $\sigma > \approx 0.278$

based on Eq. (26). More specifically, the nonequilibrium steady state (nss) conditions associated with Eq. (26) are the relations $<Z_t> = \beta/\alpha$ and $\text{Var}[Z_t] = (\beta/(\alpha - \sigma^2/2))^2 - (\beta/\alpha)^2$, so that variance in $Z$ explodes as $\sigma \to \sqrt{2\alpha}$, here, $\sqrt{2}$.

Similar analyses to Fig. 8 across increasing values of $\beta$ produce increasingly complicated equivalence classes $\{\sigma, Z\}$, as constrained by the nss conditions on $Z_t$.

## Higher dimensional systems

Above, we have viewed systems as sufficiently well characterized by the single scalar parameter $Z$, mixing material resource/energy supply with internal and external flows of information. The real world, however, may often be far more complicated. That is, invoking techniques akin to Principal Component Analysis, there may be more than one independent composite entity irreducibly driving system dynamics. It may then be necessary to replace the scalar $Z$ with an n-dimensional vector $\mathbf{Z}$ having orthogonal components that, together, account for a good portion of the total variance in the rate of supply of essential resources. The dynamic equations are then in vector form:

$$\begin{aligned} F(\mathbf{Z}) &= -\log(h(g(\mathbf{Z})))g(\mathbf{Z}) \\ S &= -F + \mathbf{Z} \cdot \nabla_{\mathbf{Z}} F \\ \partial \mathbf{Z}/\partial t &\approx \hat{\mu} \cdot \nabla_{\mathbf{Z}} S = f(\mathbf{Z}) \\ -\nabla_{\mathbf{Z}} F + \nabla_{\mathbf{Z}}(\mathbf{Z} \cdot \nabla_{\mathbf{Z}} F) &= \\ \hat{\mu}^{-1} \cdot f(\mathbf{Z}) &\equiv f^*(\mathbf{Z}) \\ ((\partial^2 F/\partial z_i \partial z_j)) \cdot \mathbf{Z} &= f^*(\mathbf{Z}) \\ ((\partial^2 F/\partial z_i \partial z_j))|_{\mathbf{Z}_{nss}} \cdot \mathbf{Z}_{\mathbf{nss}} &= \mathbf{0} \end{aligned} \quad (27)$$

Here, $F$, $g$, $h$, and $S$ are scalar functions, and $\hat{\mu}$ is an n-dimensional square matrix of diffusion coefficients. The matrix $((\partial F/\partial z_i \partial z_j))$ is the obvious n-dimensional square matrix of second partial derivatives, and $f(\mathbf{Z})$ is a vector function. The last relation imposes a nonequilibrium steady state condition, i.e. $f^*(\mathbf{Z}_{nss}) = \mathbf{0}$.

For the 'simple' Rate Distortion approach, $h(g(\mathbf{Z})) \to g(\mathbf{Z})$, while, again, we assume $\mathbf{Z}(t) \to \mathbf{Z}_{nss}$.

For $n \ge 2$, this is an overdetermined system of partial differential equations (Spencer 1969). Indeed, for a general $f^*(\mathbf{Z})$ the system is inconsistent, resulting in as many as $n$ different expressions for $F(\mathbf{Z})$, and hence the same number of 'temperature' measures as determined by the relation $F = -\log(g)g$.

## Declarations

**Conflicts of interest** The author declares that there is no conflict of interest.

## References

Abraham E, Posner J, Wickramaratne P, Aw N, van Dijk M, Cha J, Weissman M, Talati A (2020) Concordance in parent and offspring cortico-basal ganglia white matter connectivity varies by parental history of major depressive disorder and early parental care. Soc Cognit Affect Neurosci. https://doi.org/10.1093/scan/nsaa118

Atlan H, Cohen I (1998) Immune information, self-organization, and meaning. Int Immunol 10:711–717

Baars B (1989) A cognitive theory of consciousness. Cambridge University Press, New York

Baez-Mendoza R, Mastrobattista E, Wang A, Williams Z (2021) Social agent identity cells in the prefrontal cortex of interacting groups of primates. Science 374:421. https://doi.org/10.1126/science.abb4149

Barraza P, Perez A, Rodriguez E (2020) Brain-to-Brain Coupling in the Gamma-Band as a Marker of Shared Intentionality. Front Hum Neurosci 14:295. https://doi.org/10.3389/fnhum.2020.00295

Brown R (1992) Out of line. R Inst Procee 64:207–243

Cayron C (2006) Groupoid of orientational variants. Acta Crystalogr Sect A A62:21040

Champagnat N, Ferriere R, Meleard S (2006) Unifying evolutionary dynamics: from individual stochastic process to macroscopic models. Theor Popul Biol 69:297–321

Clark A (2009) Spreading the joy? Why the machinery of consciousness is (probably) still in the head. Mind 118:963–993

Cover T, Thomas J (2006) Elements of information theory, 2nd edn. Wiley, New York

Dashorst P, Mooren T, Kleber R, de Jong P, Huntjens R (2019) Intergenerational consequences of the Holocaust on offspring mental health: a systematic review of associated factors and mechanisms. Eur J Psychotraumatol 10:1654065. https://doi.org/10.1080/20008198.2019.1654065

de Groot S, Mazur P (1984) Nonequilibrium thermodynamics. Dover, New York

Dehaene S, Changeux J (2011) Experimental and theoretical approaches to conscious processing. Neuron 70:200–227

Dembo A, Zeitouni O (1998) Large deviations and applications, 2nd edn. Springer, New York

Diamond D, Campbell A, Park C, Halonen J, Zoladz P (2007) The temporal dynamics model of emotional memory processing. Neural Plast. https://doi.org/10.1155/2007/60803

Dikker S et al (2017) Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. Curr Biol 27:1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Dolan B, Janke W, Johnston D, Stathakopoulos M (2001) Thin fisher zeros. J Phys A 34:6211–6223

Dore B, Scholz C, Baek E, Garcia J, O'Donnell M, Bassett D, Vettel J, Falk E (2019) Brain activity tracks population information sharing by capturing consensus judgements of value. Cereb Cortex. https://doi.org/10.1093/cercor/bhy176

Dretske F (1994) The explanatory role of information. Phil Trans R Soc A 349:59–70

Fisher M (1965) Lectures in theoretical physics, vol 7. University of Colorado Press, Boulder

Feynman R (2000) Lectures in computation. Westview Press, Boulder CO

Golubitsky M, Stewart I (2006) Nonlinear dynamics and networks: the groupoid formalism. Bull Am Math Soc 43:305–364

Hari R, Henriksson L, Malinen S, Parkkonen L (2015) Centrality of social interaction in human brain function. Neuron 88:181–193

Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C (2012) Brain-to-brain coupling: a mechanism for creating and sharing a social world. Trends Cognit Sci 16:114–121. https://doi.org/10.1016/j.tics.2011.12.007

Hasson U, Frith C (2016) Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. Philos Trans B 371:20150366. https://doi.org/10.1098/rstb.2015.0366

Henry T, Robinaugh D, Fried E (2021) On the control of psychological networks. Psychometrika online. https://doi.org/10.1007/s11336-021-09796-9

Hoyrup M (2013) Computability of the ergodic decomposition. Ann Pure Appl Logic 164:542–549

Jin H, Hu Z, Zhou X (2008) A convex stochastic optimization problem arising from portfolio selection. Math Financ 18:171–183

Khinchin A (1957) Mathematical foundations of information theory. Dover, New York

Kingsbury L, Huang S, Wang J, Gu K, Golshani P, Wu Y, Hong W (2019) Correlated neural activity and encoding of behavior across brains of socially interacting animals. Cell 178:429–446. https://doi.org/10.1016/j.cell.2019.05.022

Kingsbury L, Hong W (2020) A multi-brain framework for social interaction. Trends Neurosci 43:651–665. https://doi.org/10.1016/j.tins.2020.06.008

Kuhlen A, Bogler C, Brennan SE, Haynes JD (2017) Brains in dialogue: decoding neural preparation of speaking to a conversational partner. Soc Cogn Affect Neurosci 12:871–880. https://doi.org/10.1093/scan/nsx018

Laidler K (1987) Chemical kinetics, 3rd edn. Harper and Row, New York

Lucia Valencia A, Froese T (2020) What binds us? Inter-brain neural synchronization and its implications for theories of human consciousness. Neurosci Conscious. https://doi.org/10.1093/nc/niaa010

Maturana H, Varela F (1980) Autopoiesis and cognition: the realization of the living. Reidel, Boston

Mezo I, Keady G (2015) Some physical applications of generalized Lambert functions, arXiv:1505.01555v2 [math.CA] 22 Jun 2015

Nair G, Fagnani F, Zampieri S, Evans R (2007) Feedback control under data rate constraints: an overview. Proc IEEE 95:108137

Newman M (2010) Networks: an introduction. Oxford University Press, New York

Perez A, Carreiras M, Duñabeitia JA (2017) Brain-to-brain entrainment: EEG interbrain synchronization while speaking and listening. Sci Rep 7:4190. https://doi.org/10.1038/s41598-017-04464-4

Pettini M (2007) Geometry and topology in hamiltonian dynamics and statistical mechanics. Springer, New York

Protter P (2005) Stochastic integration and differential equations, 2nd edn. Springer, New York

Robinson S (1993) Shadow prices for measures of effectiveness II: general model. Oper Res 41:536–548

Rose M, Styr B, Schmid T, Elie J, Yartsev M (2021) Cortical representation of group social communication in bats. Science 374:422. https://doi.org/10.1126/science.aba9584

Ruelle D (1964) Cluster property of the correlation functions of classical gases, Reviews of Modern Physics, April, 580-584

Shields P, Neuhoff D, Davisson L, Ledrappier F (1978) The distortion-rate function for nonergodic sources. Ann Probab 6:138–143

Sliwa J (2021) Toward collective animal neuroscience. Science 374:397–398

Spencer D (1969) Overdetermined systems of linear partial differential equations. Bull Am Mathth Soc 75:179–239

Stolk A et al (2014) Cerebral coherence between communicators marks the emergence of meaning. Proc Natl Acad Sci USA 111:18183–18188. https://doi.org/10.1073/pnas.1414886111

Wallace R (2005) Consciousness: a mathematical treatment of the global neuronal workspace model. Springer, New York

Wallace R (2012) Consciousness, crosstalk, and the mereological fallacy: an evolutionary perspective. Phys Life Rev 9:426–453

Wallace R (2018) New statistical models of nonergodic cognitive systems and their pathologies. J Theor Biol 436:72–78

Wallace R (2020) On the variety of cognitive temperatures and their symmetry-breaking dynamics. Acta Biotheor. https://doi.org/10.1007/s10441-019-09375-7

Wallace R (2020) Cognitive dynamics on clausewitz landscapes: the control and directed evolution of organized conflict. Springer, New York

Wallace R (2021) Toward a formal theory of embodied cognition. Biosystems 202:104356

Wallace R (2021) Enbidued cognition and its pathologies: the dynamics of institutional failure on wickedly hard problems. Commun Nonlinear Sci Numer Simul 95:105616

Wallace R (2021) How AI founders on adversarial landscapes of fog and friction. J Def Model Simul. https://doi.org/10.1177/1548512920962227

Wallace R (2022) Consciousness, cognition and crosstalk: the evolutionary exaptation of nonergodic groupoid symmetry-breaking. Springer, New York

Wallace R (2022b) Major transitions as groupoid symmetry-breaking in nonergodic prebiotic, biological and social information systems. Submitted

Weinstein A (1996) Groupoids: unifying internal and external symmetry. Not Am Math Assoc 43:744–752

Yeshurun Y, Nguyen M, Hasson U (2021) The default mode network: where the idiosyncratic self meets the shared social world, Nature Reviews Neuroscience. https://www.nature.com/articles/s41583-020-00420-w

Yi S, Nelson PW, Ulsoy AG (2010) Time-delay systems: analysis and control using the lambert W function. World Scientific, New Jersey