Stability Criteria for the Contextual Emergence of Macrostates in Neural Networks

Peter beim Graben*
School of Psychology and Clinical Language Sciences,
University of Reading, UK

Adam Barrett
Department of Informatics,
University of Sussex, UK

Harald Atmanspacher
Institute for Frontier Areas of Psychology and Mental Health,
Freiburg, Germany

July 2, 2009

Accepted Manuscript to appear in *Network: Computation in Neural Systems.*

Abstract

More than thirty years ago, Amari and colleagues proposed a statistical framework for identifying structurally stable macrostates of neural networks from observations of their microstates. We compare their stochastic stability criterion with a deterministic stability criterion based on the ergodic theory of dynamical systems, recently proposed for the scheme of contextual emergence and applied to particular inter-level relations in neuroscience. Stochastic and deterministic

*Email: p.r.beimgraben@reading.ac.uk

stability criteria for macrostates rely on macro-level contexts, which make them sensitive to differences between different macro-levels.

1 Introduction

One of the most important issues in neuroscience are relations between different levels of description. In *cognitive neuroscience*, this refers to the relationship between the brain states and their dynamics and mental states relevant for phenomena such as cognition and even consciousness at a higher level. Corresponding ideas have been put forward, e.g., by Smolensky (1988, 2006), beim Graben (2004) and Atmanspacher and beim Graben (2007) for the relationship between neurodynamics on the one hand and cognitive computation or mental states in general on the other.

In *computational neuroscience*, relations between microscopic states (of ion channels, individual neurons, synapses), mesoscopic states (of neural assemblies, cortical columns) and macroscopic states (of functional networks as observable by techniques such as EEG or fMRI) and their associated levels of description are concerned. A recent review by Atmanspacher and Rotter (2008) outlines numerous examples, achievements and problems for specific inter-level relations between descriptions of the brain and its components.

Inter-level relations in general have been a topic of discussion for decades, and key questions have not been ultimately resolved even today. Are higher-level descriptions strongly reducible to lower-level descriptions? Do higher-level descriptions supervene upon lower-level descriptions? Or do higher-level descriptions emerge from lower-level descriptions?

Recently, Bishop and Atmanspacher (2006) suggested a classification of inter-level relations in terms of necessary and sufficient conditions. If a lower-level description bears both necessary and sufficient conditions for a higher-level description, the higher level can be *strongly reduced* to the lower level. If a lower-level description possesses sufficient but not necessary conditions for a higher-level description, the latter *supervenes* upon the former. For situations in which a lower-level description is necessary but not sufficient for a higher-level description, Bishop and Atmanspacher (2006) propose the term *contextual emergence.* The remaining, rather unattractive, case in which the lower-level description provides neither necessary nor sufficient conditions for the higher level description has been called *radical emergence*, resembling a patchwork scenario with basically unrelated domains.

The idea of contextual emergence has been successfully used to clarify inter-level relations between statistical mechanics and thermodynamics and between quantum mechanics and physical chemistry (Primas 1998, Bishop and Atmanspacher 2006). It has also been shown to be a viable tool to formally address neural correlates of consciousness (Chalmers 2000) in terms of partitioned neural state spaces (Atmanspacher and beim Graben 2007). This methodology, which is based on the ergodic theory of *deterministic* dynamical systems, can also be applied to study relations between the (lower-level) dynamics of neural networks and the (higher-level) behavior of local field potentials or the EEG (Allefeld et al. 2009).

It is the aim of this paper to demonstrate that elements of the same kind of contextual emergence are applicable, and in fact have been applied earlier, to inter-level relations between *statistical* descriptions of neural systems. Remarkably, the basic ingredients for such an investigation have been worked out more than thirty years ago by Amari (1974) and Amari et al. (1977) in two influential papers on macrostates in random neural networks.

Referring to higher-level states as macrostates and lower-level states as microstates, Amari (1974, p. 203) introduced a theory of statistical neurodynamics in the following way:

> "Statistical neurodynamics investigates such properties of random nets that are possessed in common by almost all random nets in an ensemble rather than those that are possessed on the average. When random nets are composed of a sufficiently large number of neurons, it is anticipated from the law of large numbers that such properties surely exist. These properties, if they exist, do not depend on the precise values of net parameters but only on their statistics. They are structurally stable in the sense that a minor change of parameters do not destroy the properties. These properties are analyzed in the following by introducing the concept of macrostates."

The paper is structured as follows: In Sect. 2 we review the contextual emergence of (higher-level) macrostates and their associated properties from (lower-level) microstates and their associated properties. Particular emphasis is placed on the issue of structural stability, referred to in the quotation above. In Sect. 3 we demonstrate contextual emergence in neural networks in three steps. In a first step (Sect. 3.1) we recapitulate Amari's approach by translating his original formalism in the light of algebraic statistical mechanics (Sewell 2002) and dynamical system theory (Guckenheimer and Holmes 1983). In a second step (Sect. 3.2) we introduce the notion of contextuality by epistemic observables in the sense of beim Graben and Atmanspacher (2006). In a third step (Sect. 3.3) we show how Amari's macrostate criteria implement appropriate stability conditions for the contextual emergence of macroscopic descriptions for a random neural network.

The paper concludes with a discussion of similarities and differences of the approaches. In addition, a particular example is addressed, contextual emergence of macrostates in liquid state machines (Maass et al. 2002) that might be relevant for the discussion of neural correlates of consciousness (Chalmers 2000, Atmanspacher and beim Graben 2007).

2 Contextual Emergence: The Basic Idea

For the idea of contextual emergence it is assumed that the description of features of a system at a particular level offers *necessary but not sufficient* conditions to derive features at a higher level of description. In logical terms, the necessity of conditions at the lower level of description means that higher-level features *imply* those of the lower level of description. The converse — that lower-level features also *imply* the features at the higher level of description — does not hold in contextual emergence. This is the meaning of the absence of sufficient conditions at the lower level of description. Additional, contingent contexts for the transition from the lower to the higher level of description are required in order to provide such sufficient conditions.

For the contextual emergence of temperature, the notion of thermal equilibrium represents such a context. Thermal equilibrium is not available at the lower-level description of Newtonian or statistical mechanics. Implementing thermal equilibrium in terms of a particular *stability condition* (the Kubo-Martin-Schwinger (KMS) condition) and considering the thermodynamic limit of infinitely many particles ($N \to \infty$) at the level of statistical mechanics, temperature can be obtained as an emergent property at a higher-level thermodynamical description.[1] It is of paramount importance for this

[1]A non-technical presentation of the detailed argumentation can be found in At-

procedure that KMS states satisfy a *stability condition* that derives from a context at the level of thermodynamics and implemented at the level of statistical mechanics.

In addition to the contextual emergence of temperature as a new observable that is not contained in the algebra of observables of statistical mechanics, the concept of a thermal state differs substantially from the concept of a statistical state. The introduction of KMS states in the phase space of statistical mechanics entails a coarse graining (a change of topology) that leads to equivalence classes of microstates with properties implying the same temperature. These equivalent microstates are multiple realizations of one and the same thermal state. In this sense, thermal states supervene on microstates, although thermal properties emerge from properties of microstates.

The significance of contextual emergence in combination with supervenience as opposed to strict reduction in this example is clear. Of course, it would be interesting to extend the general construction scheme outlined above to other cases. More physical examples are indicated and discussed, for example, in Primas (1998) and Batterman (2002). However, the concept of stability, in the sense of stability against perturbations or fluctuations, should serve as a key principle for the construction of a contextual topology and an associated algebra of contextual observables in examples even beyond physics.

As mentioned in the introduction, possible, and ambitious, cases refer to emergent features in the framework of cognitive science and neuroscience (Atmanspacher 2007). As the brain definitely operates far from equilibrium, the general approach must be able to incorporate a non-equilibrium stability criterion. Based on empirical material that suggests to consider neurodynamics in terms of deterministic nonlinear dynamics, Atmanspacher and beim Graben (2007) suggested a suitable implementation, based on ergodic theory, of appropriate higher-level contexts as lower-level stability conditions.

Depending on the precise nature of the dynamics, basins of attraction (e.g., for fixed points) or invariant hyperbolic sets (e.g., for chaotic attractors) provide partitions of the phase space. These coarse-grainings are directly prescribed by the deterministic dynamics of the system considered and can be investigated in terms of ergodic Markov chains (see appendix). At the

manspacher (2007). The KMS condition induces a partition into equivalence classes of mechanical states defining statistical states whose mean energy can be assigned a particular temperature.

higher level of description, the coarse grains or partition cells represent new (macro-) states with new associated observables, respectively.

3 Emergence of Macrostates in Neural Networks

In this section we demonstrate how Amari's macrostate criteria for statistical neurodynamics (Amari 1974, Amari et al. 1977) compares to macrostate criteria for the contextual emergence of macroscopic descriptions for neural networks.

3.1 Amari's Macrostate Conditions

Amari (1974) and Amari et al. (1977) discussed ensembles of random (or stochastic) networks of McCulloch-Pitts units.[2] This can be formalized by a phase space $X_n \subset \mathbb{R}^n$ of n randomly connected model neurons, obeying a nonlinear difference equation:

$$\boldsymbol{x}(t+1) = \Phi_\omega(\boldsymbol{x}(t)) \,. \tag{1}$$

Here $\boldsymbol{x}(t) \in X_n$ is the activation vector (the *microstate*) of the network at time t and Φ_ω is a nonlinear map, parameterized by ω. The network as a dynamical system is then given by (X_n, Φ_ω).

The map Φ_ω is often assumed to be of the form

$$\Phi_\omega(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{W} \cdot \boldsymbol{x} - \boldsymbol{\theta}) \,, \tag{2}$$

with *synaptic weight matrix* $\boldsymbol{W} \in \mathbb{R}^{n^2}$, *activation threshold vector* $\boldsymbol{\theta} \in \mathbb{R}^n$, and a nonlinear squashing function $\boldsymbol{f} = (f_i)_{1 \le i \le n} : X_n \to X_n$, the *activation function* of the network. For $f_i = \Theta$ (where Θ denotes the Heaviside jump function), equations (1, 2) describe a network of McCulloch-Pitts neurons (McCulloch and Pitts 1943).

Another popular choice for the activation function is the logistic function

$$f_i(x) = \frac{1}{1 + \mathrm{e}^{-x}} \,,$$

[2] Another important example are Hebbian auto-associator networks (Hopfield 1982) trained with random patterns that have been investigated by Amari and Maginu (1988).

describing firing rate models (cf., e.g., beim Graben (2008)). Replacing Eq. (2) by the map

$$\Phi_\omega(\boldsymbol{x}) = (1 - \Delta t)\boldsymbol{x} + \Delta t \boldsymbol{f}(\boldsymbol{W} \cdot \boldsymbol{x} - \boldsymbol{\theta}) \tag{3}$$

yields a time-discrete leaky integrator network (Wilson and Cowan 1972, beim Graben and Kurths 2008). For numerical simulations, $\Delta t < 1$ will be the choice for the time step of the Euler method.

According to (2) and (3), the network parameters are given as

$$\omega = (\boldsymbol{W}, \boldsymbol{\theta}) \in \mathbb{R}^{n^2} \times \mathbb{R}^n \, . \tag{4}$$

In a random neural network, the parameters ω are regarded as stochastic variables drawn from a probability space $\Omega_n = \mathbb{R}^{n^2} \times \mathbb{R}^n$ with measure

$$\mu_n : \mathcal{B}(\Omega_n) \to [0, 1] \tag{5}$$

for measurable sets from $\mathcal{B}(\Omega_n)$. We refer to a particular network realization from this probability space as $N(\omega)$.

In order to obtain limit theorems for random neural networks, one has to assure that network realizations of different size n behave similarly. Thus, we restrict ourselves to particular network topologies, such as directed Erdős-Rényi graphs with fixed connectivity p (Bollobás 2001, beim Graben and Kurths 2008, Maass et al. 2002) or networks with Gaussian synaptic weights $\boldsymbol{W} \sim \mathcal{N}(\boldsymbol{m}, n\sigma^2)$. For an Erdős-Rényi network of size n the expected number of connections is then np.

Amari (1974) and Amari et al. (1977) treat the network parameters ω as strictly stochastic variables, such that the evolution equations (2) and (3) describe stochastic processes where ω assumes another value after each temporal iteration. Amari (1974, p. 203) concedes that such an "ensemble of random nets is quite different to the Gibbs ensemble in statistical mechanics in this respect, because the latter consists of dynamical systems of the same structure but only in different states".

Simplifying Amari's treatment of networks, we follow Amari and Maginu (1988) and Touboul et al. (2008) who describe a random neural network as one particular realization of the stochastic variable ω that is considered frozen during the temporal evolution of the network states. This assumption does not only facilitate the theory, it is also more plausible for certain scenarios. The changes of network parameters during development usually take place at a larger time scale than the intrinsic network dynamics. Note, however, that time scales related to synaptic plasticity cannot always be separated so easily.

Touboul et al. (2008) also consider noise-driven networks, which would again require a fully stochastic treatment. We will refrain from this complication as well and study random neural networks in the original sense of Gibbs ensembles, namely as ensembles of identical systems (characterized by the same realization of the stochastic parameter ω for a given network size n). Such systems differ only in their initial conditions. This approach is consistent with the treatment of deterministic dynamical systems where initial conditions can be drawn according to probability distributions over phase space. Particular distributions, so-called Sinai-Ruelle-Bowen (SRB) measures (Guckenheimer and Holmes 1983), obey stability conditions that are crucial for the contextual emergence of macrostates in dynamical systems.

A macro-observable of a neural network (X_n, Φ_ω) is a real-valued function $s : X_n \to \mathbb{R}$ that is measurable by suitable techniques. Such observables are, e.g., local field potentials or the EEG (Freeman 2007). Let $s_{i;n} : X_n \to \mathbb{R}$ be a family of m observables (for $1 \leq i \leq m$) for a network of size n, such that $\boldsymbol{y} = \boldsymbol{s}_n(\boldsymbol{x}) = (s_{i;n}(\boldsymbol{x}))_{1 \leq i \leq m}$ is a vector in m-dimensional space $Y \subset \mathbb{R}^m$. We call $Y = \boldsymbol{s}_n(X_n)$ the *macrostate space* generated by the family $(s_{i;n})_{1 \leq i \leq m}$. Note that we are interested in a macrostate space Y that is the same for a sequence of network realizations $(X_n, \Phi_\omega)_{n \in \mathbb{N}}$.

Amari (1974) and Amari et al. (1977) asked for conditions under which the images $\boldsymbol{y} = \boldsymbol{s}_n(\boldsymbol{x})$ of a microstate $\boldsymbol{x}$ under the observables $\boldsymbol{s}_n$ can be regarded as *macrostates* obeying a macroscopic evolution law. In order to formulate appropriate constraints, Amari (1974, p. 203) discussed two possible conditions which should be fulfilled if the number n of network components is sufficiently large:[3]

1. In order for $\boldsymbol{s}_n(\boldsymbol{x})$ to represent a macrostate whose state-transition law is identical for almost all nets, it is required that the values of $\boldsymbol{y}(1) = \boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}(0)))$ are identical for almost all nets in the ensemble, even though the values of $\Phi_\omega(\boldsymbol{x}(0))$ differ for different $N(\omega)$'s.

2. The value of the macrostate $\boldsymbol{y}(1) = \boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}(0)))$ is required to depend on $\boldsymbol{x}(0)$ not directly but only through the value of the initial macrostate $\boldsymbol{y}(0) = \boldsymbol{s}_n(\boldsymbol{x}(0))$.

[3] We present Amari's proposal in our own notation here and in the following.

Amari (1974, pp. 203f) proposed that these verbal criteria give rise to the following formal *macrostate condition* (see also Amari et al. (1977, p. 98)): The images $\boldsymbol{y} = \boldsymbol{s}_n(\boldsymbol{x})$ of microstates $\boldsymbol{x}$ under the sequence of observables $\boldsymbol{s}_n$ is called a macrostate if it fulfils, for every $\boldsymbol{x} \in X_n$, the following criteria:

1. There exists a function $\varphi : Y \to Y$ for which

$$\lim_{n\to\infty} \mathrm{E}_{\mu_n}(\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}))) = \varphi(\boldsymbol{s}_n(\boldsymbol{x})) \,. \tag{6}$$

2. And, furthermore

$$\lim_{n\to\infty} \mathrm{V}_{\mu_n}(||\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}))||) = 0 \,. \tag{7}$$

Here, E_{μ_n} and V_{μ_n} denote, respectively, the ensemble mean and ensemble variance (i.e., the expectation and variance of stochastic variables with respect to the probability measure μ_n), and the Euclidian norm

$$||\boldsymbol{y}||^2 = \sum_{i=1}^{m} y_i^2 \,.$$

The macrostate condition formulated by Eqs. (6,7) can be graphically depicted as a diagram that is asymptotically ($n \to \infty$) commutative (Fig. 1). Note that a similar *quasi-commutativity* has been observed by Gaveau and Schulman (2005) for the coarse-graining of Markov chains.

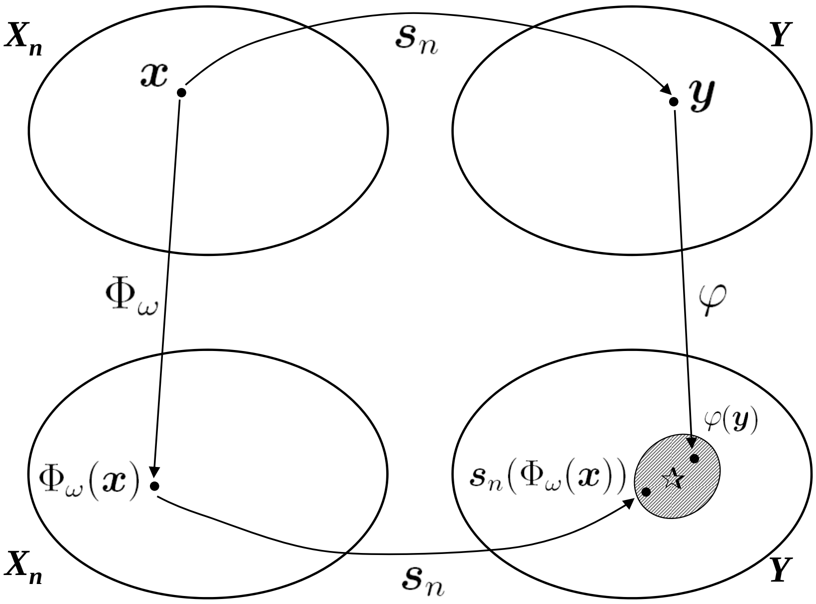


Figure 1: Illustration of Amari's macrostate condition (Amari 1974, p. 204). The observable $\boldsymbol{s}_n$ maps a microstate $\boldsymbol{x} \in X_n$ onto a macrostate $\boldsymbol{y} = \boldsymbol{s}_n(\boldsymbol{x})$ (upper row), and the time iterate $\Phi_\omega(\boldsymbol{x})$ of $\boldsymbol{x}$ onto $\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}))$ (bottom row). The asterisk indicates the expectation value due to Amari's first criterion, whereas the shaded area illustrates the dispersion due to the second.

If a sequence of observables $\boldsymbol{s}_n$ gives rise to a macrostate in the sense of Eqs. (6) and (7), one can construct another sequence of maps $\varphi_n : Y \to Y$, converging towards φ for $n \to \infty$, such that

$$\boldsymbol{y}(t+1) = \varphi_n(\boldsymbol{y}(t)) + \boldsymbol{e}_\omega(\boldsymbol{x}(t)) \, ,$$

where $\boldsymbol{e}_\omega(\boldsymbol{x}(t))$ is an error depending on the microstate $\boldsymbol{x}(t)$ (Amari 1974, p. 204). In the limiting case of $\boldsymbol{e}_\omega(\boldsymbol{x}(t)) \to 0$ for $n \to \infty$, which is not guaranteed by the macrostate condition yet, one obtains the desired macrostate evolution equation

$$\boldsymbol{y}(t+1) = \varphi(\boldsymbol{y}(t)) \, , \tag{8}$$

for all times t (Amari 1974, Amari et al. 1977). Equation (8) is non-deterministic if network parameters fluctuate stochastically, or when the network is exposed to stochastic forcing.

However, even for the simplified deterministic micro-dynamics discussed here, $\boldsymbol{e}_\omega(\boldsymbol{x}(t))$ does in general not converge towards zero for $n \to \infty$. The reason are temporal correlations in $\boldsymbol{x}(t)$. An illustrative example is a Hopfield network trained on random patterns, discussed by Amari and Maginu (1988). This makes Amari's macrostate criteria (6) and (7) necessary but not sufficient conditions for the macrostate dynamics (8). As a sufficient condition Amari (1974, p. 204) *postulates*

$$\lim_{n \to \infty} \mathrm{Prob} \left\{ \sup_t ||\boldsymbol{s}_n(\Phi_\omega^t(\boldsymbol{x}(0))) - \varphi^t(\boldsymbol{x}(0))|| > \epsilon \right\} = 0 \tag{9}$$

for arbitrary $\epsilon > 0$.

Since Eq. (9) refers to the images of microstates in macrostate space under the observables $\boldsymbol{s}_n$ it can be regarded as a decorrelation condition. Amari (1974) compares Eq. (9) with Boltzmann's *Stoßzahlansatz* in statistical mechanics. Amari et al. (1977) proved Eq. (9) under several assumptions about the microscopic dynamics.

3.2 Contextual Observables

Any real-valued function $s : X_n \to \mathbb{R}$ of a neural network (X_n, Φ_ω) that can be measured is an observable. Since observables are usually defined with respect to a particular scientific or pragmatic *context* (Freeman 2007), we refer to them as *contextual observables.* A context provides a reference frame for the meaningful usage of observables.[4]

In general, the mapping $\boldsymbol{s}_n : X_n \to Y, \boldsymbol{x} \mapsto \boldsymbol{y}$ obtained from a family $s_{i;n} : X_n \to \mathbb{R}$ of m observables (for $1 \le i \le m$) is not injective, such that different neural microstates $\boldsymbol{x}, \boldsymbol{x}' \in X_n$ are mapped onto the same state $\boldsymbol{y} \in Y$ in macrostate space. Following beim Graben and Atmanspacher (2006), we call the states $\boldsymbol{x}, \boldsymbol{x}' \in X$ *epistemically equivalent* with respect to the observables $s_{i;n}$. Epistemic equivalence induces a partition of the neural phase spaces X_n into disjoint classes such that all members of one class are mapped onto the same point $\boldsymbol{y}$ in the macrostate space Y. Call $A_{\boldsymbol{y}} = \boldsymbol{s}_n^{-1}(\boldsymbol{y}) \subset X_n$ the equivalence class of all pre-images of $\boldsymbol{y}$.

Arbitrarily defined observables are unlikely to obey stability conditions such as Amari's macrostate condition. However, we can generally construct

[4]This situation resembles quantum mechanical complementarity where observables such as the position and momentum of an electron refer to different, mutually excluding measurement contexts. For a treatment of complementary observables in classical systems see beim Graben and Atmanspacher (2006).

a sequence of mappings $\tilde{\varphi}_n : Y \to \wp(Y)$, where $\wp(Y)$ denotes the power set of the macrostate space Y, by setting

$$\tilde{\varphi}_n(\boldsymbol{y}) = (\boldsymbol{s}_n \circ \Phi_\omega \circ \boldsymbol{s}_n^{-1})(\boldsymbol{y}) \,. \tag{10}$$

These functions bring $\boldsymbol{y} \in Y$ to a set $\tilde{\varphi}_n(\boldsymbol{y}) = \boldsymbol{s}_n(\Phi_\omega(A_{\boldsymbol{y}})) \subset Y$. Hence, the deterministic microscopic dynamics given by Φ_ω at X_n is mapped onto a non-deterministic dynamics in macrostate space. Figure 2 illustrates the mapping (10). Amari's macrostate conditions together with his decorrelation postulate guarantee that this non-deterministic macroscopic dynamics converges to a deterministic one in the thermodynamic limit $n \to \infty$.

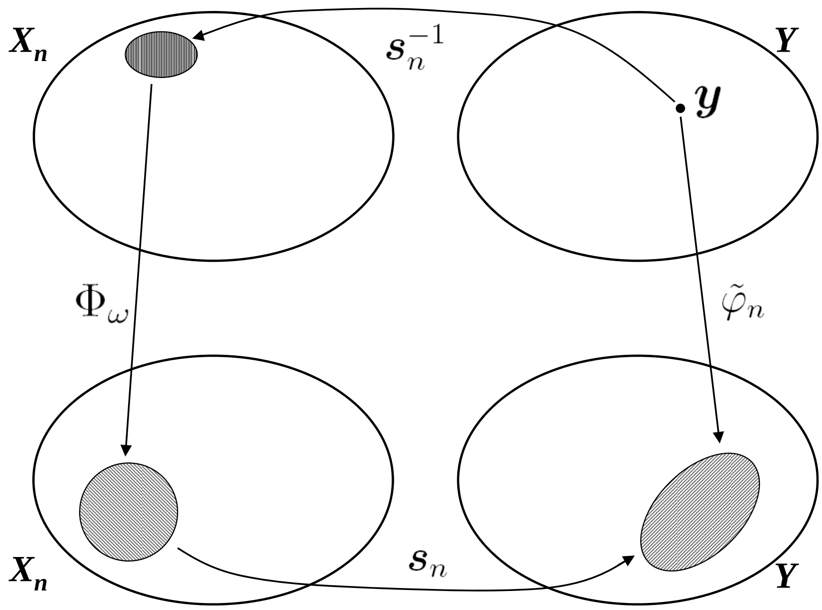


Figure 2: Construction of a non-deterministic dynamics $\tilde{\varphi}$ in macrostate space. Upper row: the pre-image of a macrostate $\boldsymbol{y}$ under the observable $\boldsymbol{s}_n$ is an equivalence class of microstates, $A_{\boldsymbol{y}} = \boldsymbol{s}_n^{-1}(\boldsymbol{y}) \subset X_n$. Bottom row: The temporal iterates of epistemically equivalent microstates, $\Phi_\omega(A_{\boldsymbol{y}})$, belong to a set $\tilde{\varphi}_n(\boldsymbol{y}) = \boldsymbol{s}_n(\Phi_\omega(A_{\boldsymbol{y}})) \subset Y$ in macrostate space.

The observable macro-dynamics (10) inherits an important property from its microscopic counterpart Eq. (1). Since the latter is described by a first-order difference equation, the former becomes a first-order Markov process (Shalizi and Moore 2003). Thus, we can define transfer operators

$$T_n(\boldsymbol{y}', \boldsymbol{y}) = \begin{cases} 0 & \text{for} \quad \boldsymbol{y}' \notin \tilde{\varphi}_n(\boldsymbol{y}) \\ 1 & \text{for} \quad \boldsymbol{y}' \in \tilde{\varphi}_n(\boldsymbol{y}) \,. \end{cases} \tag{11}$$

The transfer operators do not vanish if the state $\boldsymbol{y}' \in Y$ belongs to the image of $\tilde{\varphi}_n(\boldsymbol{y})$, for $\boldsymbol{y} \in Y$.

3.3 Structural Stability

3.3.1 Stochastic Structural Stability

In a first step, we show that Amari's macrostate criterion entails a stochastic condition for structural stability. To this aim, we formalize Amari's first verbal criterion "that the values of $\boldsymbol{y}(1) = \boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}(0)))$ are identical for almost all nets in the ensemble, even though the values of $\Phi_\omega(\boldsymbol{x}(0))$ differ for different $N(\omega)$'s" as follows:

For all $\epsilon, \delta > 0$ there exists a k such that for all $n > k$, $\boldsymbol{x} \in X_n$,

$$\text{Prob}\{||\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x})) - \boldsymbol{s}_n(\Phi_{\omega'}(\boldsymbol{x}))|| > \epsilon\} < \delta \tag{12}$$

if ω and ω' are drawn from $\Omega_n = \mathbb{R}^{n^2} \times \mathbb{R}^n$ using the probability measure μ_n. Equation (12) expresses that, for two random realizations of the microstate dynamics, the macrostate dynamics will, with high probability, be very similar if the number of neurons is large. (Note that the exact interpretation of (12) is slightly different from Amari's formulation in that the values of $\boldsymbol{y}(1) = \boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}(0)))$ are *almost identical for almost all nets in the ensemble.*)

Amari's second macrostate criterion is equivalent with the probabilistic structural stability condition (12), provided that for all $\epsilon > 0$, there exists a k such that for all $n > k$, $\boldsymbol{x} \in X_n$,

$$\text{V}_{\mu_n}(||\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}))||) < \epsilon \,, \tag{13}$$

This assertion is proved as follows: Consider a one-dimensional macrostate obeying a Gaussian distribution. (Higher-dimensional macrostates are treated similarly.) Suppose S, S' are two one-dimensional stochastic variables, obeying an $\mathcal{N}(m, \sigma^2)$ distribution. Then, for any ϵ,

$$\text{Prob}\{|S - S'| > \epsilon\} = 2\left(1 - F\left(\frac{\epsilon}{\sqrt{2}\sigma}\right)\right) \,, \tag{14}$$

where F is the cumulative distribution function for a $\mathcal{N}(0,1)$ random variable. Using

$$\frac{zf(z)}{z^2+1} < 1 - F(z) < \frac{f(z)}{z}\,, \tag{15}$$

where f is the probability density function of a $\mathcal{N}(0,1)$ random variable, two inequalities follow:

$$\text{Prob}\{|S - S'| > \epsilon\} \;<\; \frac{2\sigma}{\sqrt{\pi}\epsilon}\mathrm{e}^{-\epsilon^2/4\sigma^2}\,, \tag{16}$$

$$\text{Prob}\{|S - S'| > \epsilon\} \;>\; \frac{2\sigma\epsilon}{\sqrt{\pi}(\epsilon^2+2\sigma^2)}\mathrm{e}^{-\epsilon^2/4\sigma^2}\,. \tag{17}$$

Assume now that Amari's macrostate criterion, as given by (13), holds. Choose any $\epsilon, \delta > 0$. Then, by Amari's condition (13), and the fact that the RHS of (16) tends to zero as $\sigma \to 0$, we can choose a k such that for any $n > k$ and any $\boldsymbol{x} \in X_n$, stochastic structural stability (12) holds.

If, conversely, Amari's macrostate criterion does not hold, then there is a $\delta > 0$ such that, for any n, we can find an $\boldsymbol{x} \in X_n$ such that

$$\mathrm{V}_{\mu_n}(\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x}))) > \delta\,. \tag{18}$$

Then, since the RHS of (17) increases with σ, for any n, we can find an $\boldsymbol{x} \in X_n$ such that

$$\text{Prob}\{|\boldsymbol{s}_n(\Phi_\omega(\boldsymbol{x})) - \boldsymbol{s}_n(\Phi_{\omega'}(\boldsymbol{x}))| > \epsilon\} > \frac{2\sqrt{\delta}\epsilon}{\sqrt{\pi}(\epsilon^2+2\delta)}\mathrm{e}^{-\epsilon^2/4\delta}\,. \tag{19}$$

Hence, stochastic structural stability (12) fails in this case.

3.3.2 Deterministic Structural Stability

Let us now assume that the contextual observables $\boldsymbol{s}_n$ partition the phase spaces X_n into a finite number $\ell \in \mathbb{N}$ of epistemic equivalence classes. Then, the Markov process in macrostate space Y, described by (11), becomes a finite-state Markov chain, or, in other words, a shift of finite type (Lind and Marcus 1995). Such a non-deterministic dynamical system can be obtained from a finite partition $\mathcal{P}_n = \{A_1, \ldots, A_\ell\}$ of X_n into pairwise disjoint subsets A_i that cover the whole phase space X_n, by choosing the characteristic functions $s_{i;n} = \chi_{A_i}$ as contextual observables. Then, $\boldsymbol{y} = \boldsymbol{s}_n(\boldsymbol{x})$ is the ℓ-dimensional canonical basis vector $\boldsymbol{y} = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ with 1 in the ith position if $\boldsymbol{x} \in A_i$.

Structural stability ensures that, in the limit $n \to \infty$, there is a well-defined transfer operator (11). For shifts of finite type this is a transition matrix

$$T_{ik} = \begin{cases} 0 & \text{for} \quad \Phi_\omega(A_k) \cap A_i = \emptyset \\ 1 & \text{for} \quad \Phi_\omega(A_k) \cap A_i \neq \emptyset \,, \end{cases} \tag{20}$$

that can be regarded as an adjacency matrix of a transition graph. As indicated in Sect. 4, Atmanspacher and beim Graben (2007) used stability conditions for shifts of finite type to construct contextually emergent observables from neurodynamics. If the transition matrix $\boldsymbol{T} = (T_{ik})$ is diagonal, there are only transitions from every state to itself, indicating fixed points in macrostate space. If a power $\boldsymbol{T}^q$ for $q \in \mathbb{N}$ is diagonal, the corresponding Markov chain is periodic. In both cases, the observable dynamics possesses invariant and ergodic states that are structurally stable.

If the transition matrix $\boldsymbol{T}$ is irreducible, i.e., if a power $\boldsymbol{T}^q > 0$ for $q \in \mathbb{N}$, the Markov chain is irreducible and aperiodic. In this case the observable dynamics possesses invariant, ergodic and mixing states (Ruelle 1968, 1989). They generalize the KMS states of algebraic statistical mechanics (Olesen and Pedersen 1978, Pinzari et al. 2000, Exel 2004) to structurally stable non-equilibrium SRB measures in the microscopic phase spaces X_n.

Now we can identify the kind of Markov chain that is obtained for an observable dynamics obeying both of Amari's macrostate conditions. Since this condition assures the existence of the map $\varphi : Y \to Y$, we obtain the deterministic macrostate dynamics (8)

$$\boldsymbol{y}(t+1) = \varphi(\boldsymbol{y}(t))$$

under our simplifications and in the limit $n \to \infty$. Thus, the pre-image $B = \boldsymbol{s}_n^{-1}(\varphi(\boldsymbol{y}(t)))$ is again an equivalence class in phase space. This means that cells of the partition $\mathcal{P} = \{A_1, \ldots, A_\ell\}$ are faithfully mapped onto cells of the partition. The resulting shift of finite type is therefore a deterministic Markov chain where every vertex in the transition graph is the source of one link at most. Hence, there must be a positive integer number q making $\boldsymbol{T}^q$ diagonal. Therefore, the macroscopic dynamics is either multistable or periodic.

We see that the restriction to deterministic microscopic dynamics yields, under Amari's macrostate condition, a deterministic macroscopic dynamics which is a shift of finite type for a finite number of distinct macrostates. As a consequence we obtain structurally stable fixed points or limit tori in the macrostate space as contextually emergent macro-features.

However, it is straightforward to relax Amari's macrostate condition to Markovian macrostate dynamics by demanding *Markov partitions* in the microscopic neural phase spaces (Sinai 1968, Ruelle 1989), but keeping the structural stability condition. For the simplified case of expanding maps, Markov partitions have the property that cells are mapped onto joins of cells; i.e. cell bounderies are mapped onto cell bounderies, without necessarily mapping cells onto cells — which was actually the result of this section for deterministic Markov chains — (for hyperbolic maps things are slightly more involved), such that $\Phi_\omega(A_k) \cap A_i \neq \emptyset$ implies $A_i \subset \Phi_\omega(A_k)$. Markov partitions entail aperiodic, irreducible Markov chains, which in turn possess invariant, ergodic and mixing SRB measures, i.e., structurally stable chaotic attractors, in the macrostate space.

3.3.3 Macrostates in a Random Neural Network

Consider a random network of n randomly connected McCulloch-Pitts units, described by Equations (1) and (2) with Heaviside activation function,

$$x_i(t+1) = \Theta\left(\sum_{j=1}^{n} w_{ij} x_j(t) - hn\right) , \tag{21}$$

where $0 < h < 1$. Let the synaptic weights w_{ij} be independently identically distributed random normal variables with mean m and variance $n\sigma^2$, that is

$$w_{ij} \sim \mathcal{N}(m, n\sigma^2) . \tag{22}$$

Following Amari (1974), we can define the *activity level*

$$r(t) = \frac{1}{n} \sum_{i=1}^{n} x_i(t) \tag{23}$$

of the network that serves as a kind of “model EEG”. Given an appropriately chosen constant r_0 satisfying $0 < r_0 < 1$, a contextual observable may be defined by

$$s(t) = \begin{cases} 0 , & r(t) \leq r_0 , \\ \dfrac{r(t) - r_0}{1 - r_0} , & r(t) > r_0 \end{cases} . \tag{24}$$

In order to prove that s yields a macrostate satisfying the previously defined stability condition (13), we first determine the probability that unit

x_i is active at time $t+1$ as

$$\begin{aligned}\text{Prob}\{x_i(t+1)=1\} &= \text{Prob}\left\{\Theta\left(\sum_{j=1}^{n} w_{ij}x_j(t)-hn\right)=1\right\}\\ &= \text{Prob}\left\{\sum_{j=1}^{n} w_{ij}x_j(t)-hn>0\right\}\\ &= \text{Prob}\left\{\sum_{j=1}^{n} w_{ij}x_j(t)>hn\right\}.\end{aligned}$$

Since $x_j(t)\in\{0,1\}$ for all times t, the weighted sum

$$\sum_{j=1}^{n} w_{ij}x_j(t)=\sum_{j_k} w_{ij_k}\,,$$

when $x_{j_k}(t)=1$. Now, the number of active units j_k at time t is given through the activity level as $nr(t)$. Because the synaptic weights w_{ij} are normally distributed according to $\mathcal{N}(m,n\sigma^2)$, their sum is normally distributed according to $\mathcal{N}(mnr(t),n^2\sigma^2 r(t))$. Thus,

$$\text{Prob}\{x_i(t+1)=1\}=\text{Prob}\left\{\mathcal{N}(0,1)>\frac{h-mr(t)}{\sigma\sqrt{r(t)}}\right\}.$$

On the other hand, the activity level (23) is described by a Bernoulli process obeying a binomial distribution

$$nr(t+1)\sim\mathcal{B}(n,q(t))\,,$$

where $\mathcal{B}(n,q(t))$ denotes the binomial distribution with n trials and probability q for a positive outcome in each trial, with

$$q(t)=\pi(r(t))=1-F\left(\frac{h-mr(t)}{\sigma\sqrt{r((t)}}\right)$$

with F being the cumulative standard normal distribution function as above. Note that $\pi:[0,1]\to[0,1]$ satisfies $\pi(0)=0$, $\pi(r)>0$ for $r>0$, and $\pi'(r)\to+\infty$ as $r\to 0$. See Figure 3 for typical plots of π against r for positive and negative m.

In the case $m>0$, let r^* be the larger solution of

$$\pi(r)=r\,,$$

while we define for $m < 0$

$$r^* = \max_{r\in[0,1]} (\pi(r)) \, .$$

Choose now $\epsilon > 0$ and $\delta > 0$, and let us consider the uncertainty in $s(t+1)$ given $s(t)$. Suppose that $s(t) = 0$ initially, which is consistent with $r(t) \leq r_0$. Then, for n sufficiently large,

$$\begin{aligned}
\text{Prob}\{s(t+1) > \epsilon\} &= \text{Prob}\{nr(t+1) > n(r_0 + \epsilon(1-r_0))\} \\
&= \text{Prob}\{\mathcal{B}(n, \pi(r(t))) > n(r_0 + \epsilon(1-r_0))\} \\
&\approx \text{Prob}\left\{\mathcal{N}(0,1) > \sqrt{n}\frac{r_0 - \pi(r(t)) + \epsilon(1-r_0)}{\sqrt{\pi(r(t))[1-\pi(r(t))]}}\right\} \\
&< \delta \, , \qquad (25)
\end{aligned}$$

under the condition

$$r_0 \geq r^* \, . \qquad (26)$$

Moreover, if $r(t) > r_0$, then for n sufficiently large,

$$\begin{aligned}
\text{Prob}\{|s(t+1) - \text{E}_{\mu_n}[s(t+1)]| > \epsilon\} &< \text{Prob}\{|s(t+1) - \text{E}_{\mu_n}[s(t+1)]| > \epsilon(1-r_0)\} \\
&= 2\,\text{Prob}\left\{\mathcal{N}(0,1) > \frac{\sqrt{n}\epsilon}{\sqrt{\pi(r(t))[1-\pi(r(t))]}}\right\} \\
&< \delta \, . \qquad (27)
\end{aligned}$$

Combining these inequalities, we deduce that there exists a k such that for all $n > k$, and all possible initial microstates $\boldsymbol{x}(t) \in X_n$,

$$\text{Prob}\{|s(t+1) - \text{E}_{\mu_n}[s(t+1)]| > \epsilon\} < \delta \, . \qquad (28)$$

This is sufficient for the observable s to yield well-defined macrostates according to Amari's stability criteria.

For another coarse-grained contextual observable

$$\tilde{s}(\boldsymbol{x}) = \begin{cases} 0\,, & r(\boldsymbol{x}) \leq r_0\,, \\ 1\,, & r(\boldsymbol{x}) > r_0\,, \end{cases} \qquad (29)$$

the same argument holds for all possible initial microstates $\boldsymbol{x}(t) \in X_n$. The observable $\tilde{s}$ provides a Markov chain that is periodic (either monoor bistable) if Amari's stability criteria hold. This confirms a result of Amari (1974) regarding the dynamics of the activity level of a random neural network.

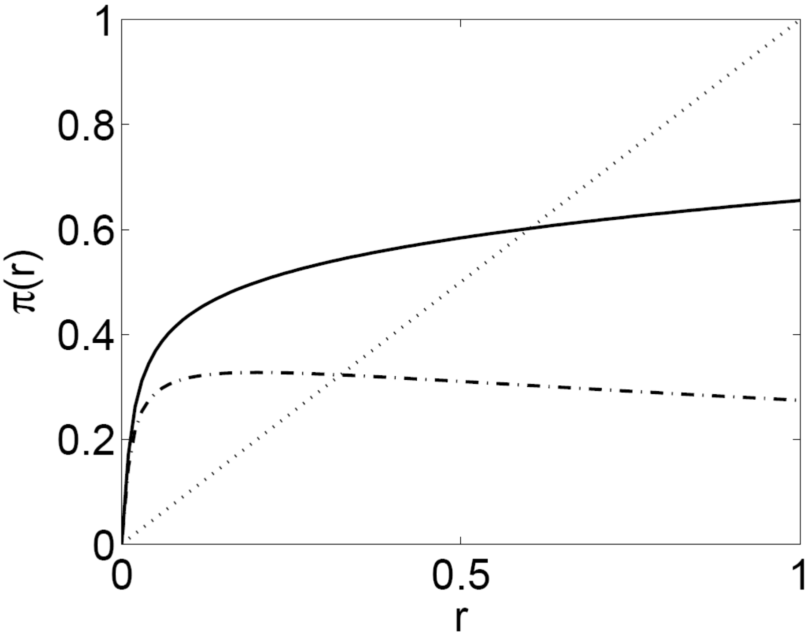


Figure 3: $\pi(r)$ against r, the activity level, for $m = +0.5$ (solid) and $m = -0.5$ (dashed-dotted). Since π gives the next time-step iterate of an initial r value, there are two stable points for r (indicated by the intersections with the identity (dotted)).

4 Discussion

In this paper we have conceptually and formally analyzed a "method of statistical neurodynamics", suggested by Amari and colleagues more than thirty years ago (Amari 1974, Amari et al. 1977). Rephrasing their ideas in terms of algebraic statistical mechanics (Sewell 2002) and dynamical system theory (Guckenheimer and Holmes 1983), we are able to demonstrate that Amari's criteria for identifying macrostates in random neural networks resemble stability criteria for the *contextual emergence* of a higher-level macroscopic description, implemented at the lower level of network dynamics.

Contextual emergence is a non-reductive way to address inter-level relations. It was suggested by Bishop and Atmanspacher (2006) and successfully applied to problems in computational and cognitive neuroscience (Atmanspacher and beim Graben 2007, Atmanspacher and Rotter 2008, Allefeld et al. 2009). The scheme of contextual emergence expresses that a lower-level description comprises necessary but not sufficient conditions for a higher-level description of a system. The lacking sufficient conditions can be provided by a contingent higher-level context that imposes stability conditions on the system's dynamics. Such contexts usually distinguish between different epistemic frameworks[5] and give rise to *contextual observables* (beim Graben and Atmanspacher 2006).

Contextual observables in Amari's statistical neurodynamics are mappings from a microscopic state space of a random neural network to a macroscopic state space, providing a suitable *coarse-graining* of the dynamics. In a first step, a sequence of observables $\boldsymbol{s}_n$ for a random neural network (X_n, Φ_ω) with parameters $\omega \in \Omega_n$ is chosen with respect to a context of the desired macroscopic description. These observables span the macrostate space $Y = \boldsymbol{s}_n(X_n)$. Secondly, the criteria given by equations (6) and (7), or given by Eq. (12), *implements* a condition for *structural stability* of the contextual macroscopic level at the level of microscopic dynamics. However, this is only a *necessary condition* for the existence of a deterministic macroscopic evolution law (8). The *decorrelation postulate* (9) serves as a *sufficient condition* which cannot be derived from (lower-level) properties of microstates and needs to be selected (postulated) with respect to a chosen context.

As Amari's criteria for identifying proper macrostates implement stability conditions upon the microstates, macrostates are *contextually emergent.* His macrostate criteria, representing necessary conditions at the lower-level description, are supplemented by a sufficient (higher-level) condition implemented as a decorrelation criterion yielding the macro-dynamics according to Eq. (8). For a finite number of macrostates, the coarse-graining partitions the microscopic state space into a finite number of classes of epistemically equivalent states. The resulting macroscopic dynamics is a Markov chain that can be studied by means of ergodic theory. We have shown that Amari's criteria directly lead to periodic Markov chains, possessing invariant and ergodic — but not mixing — Sinai-Ruelle-Bowen (SRB) equilibrium measures (Guckenheimer and Holmes 1983).

In order to obtain equilibrium states at the macroscopic level, the macro-

[5]An illustrative example is the particle-wave dualism in quantum mechanics: An electron behaves as a particle in one particular measurement context and as a wave in another.

scopic Markov chain must be aperiodic and irreducible. A sufficient condition for this is that the emergent observables arise from a *Markov partition* of the microscopic state space. In this case, the approach by Atmanspacher and beim Graben (2007) suggests to replace Amari's criteria by demanding a Markov partition that leads to aperiodic, irreducible Markov chains. The relevant equivalence classes of microstates are derived from a spectral analysis of the matrix of transition probabilities between microstates (cf. Allefeld et al. (2009) for details). The obtained partition is stable under the dynamics, i.e. the resulting macrostates are dynamically robust in this sense.

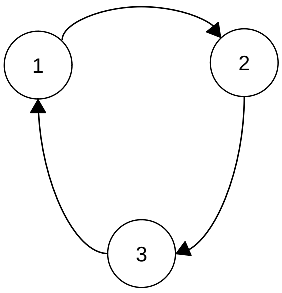
Our results could be of particular significance for research at the interface between cognitive and computational neuroscience. High-dimensional (and in the limit $n \to \infty$ infinite-dimensional) random neural networks have been suggested by Maass et al. (2002) as a new paradigm for neural computation, called *liquid computing*. Liquid state machines are large-scale neural networks with random connectivity that are perturbed by input signals. Their high-dimensional, transient trajectories are measured by so-called *read-out neurons*, that can be trained to perform particular computational tasks, especially for signal classification.

These read-out neurons implement particular observables upon the microscopic state space. By training different assemblies of read-out neurons on different tasks, one can impose different contexts for how to interpret the high-dimensional dynamics of the liquid state machine. Convergence of the macroscopic read-out states toward a classification task finally supplies a stability criterion in terms of deterministic Markov chains. Thus, liquid computing is a way to implement contextual emergence.

To conclude with a rather speculative idea, one could think of read-out neurons for neural correlates of consciousness (Chalmers 2000), interpreting the high-dimensional, transient dynamics of "liquid" cortical networks by low-dimensional, contextually given mental states. A similar idea has been discussed by Atmanspacher and beim Graben (2007).

Appendix: Stability of Markov Chains

Let us consider two simple examples of Markov chains, resulting from coarse-grainings of the state spaces of dynamical systems. Figure 4 displays the transition graphs of two simple 3-state Markov chains. The numbered nodes denote the distinct states of the processes, referring to cells of a state space partition. The lines connecting two nodes represent permitted transitions. (For a more detailed treatment, the lines are labeled with transition probabilities.)



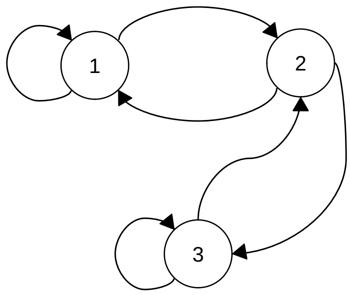


Figure 4: Transition graphs of two Markov chains with (a) only one periodic orbit and (b) an invariant, ergodic and mixing equilibrium state.

Each finite state Markov chain can be characterized by its transition matrix

$$T_{ik} = \begin{cases} 1 & \text{if there is an arc connecting node } k \text{ with node } i, \\ 0 & \text{otherwise} \end{cases}$$

The matrix $\boldsymbol{T} = (T_{ik})$ is thus the adjacency matrix of the corresponding transition graph. For the Markov chain from Fig. 4(a) we obtain

$$\boldsymbol{T}_a = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} ,$$

while the Markov chain in Fig. 4(b) is characterized by

$$\boldsymbol{T}_b = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} .$$

Paths through a transition graph correspond to integer powers of the transition matrix. Two cases are important: (1) there is an integer q such that $\boldsymbol{T}^q$ is a diagonal matrix, and (2) there is an integer q such that $\boldsymbol{T}^q$ is positive. Consider the matrix $\boldsymbol{T}_a^q$ with $q = 3$,

$$\boldsymbol{T}_a^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

which is the identity matrix in three-dimensional space. It indicates a closed path of length three through the graph shown in Fig. 4(a), corresponding to three period-three orbits (distinguished only by their initial conditions) of the Markov chain. Periodic orbits for Markov chains can be related to stationary and ergodic — but not mixing — probability distributions over state space.

On the other hand, we find that all elements of

$$\boldsymbol{T}_b^2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

are positive, such that $\boldsymbol{T}_b^2$ is positive, i.e. contains no zeros. This means that every node in the transition graph shown in Fig. 4(b) can be reached from every other node through a path of length two, which is not a periodic orbit. The corresponding Markov chain is thus irreducible and aperiodic.

For the normalized transition matrix

$$\boldsymbol{N} = \frac{1}{\lambda_1}\boldsymbol{T} ,$$

where λ_1 is the largest eigenvalue of $\boldsymbol{T}$, one gets by virtue of the Frobenius Perron theorem a unique eigenstate of $\boldsymbol{N}$ for eigenvalue one, i.e. an *invariant equilibrium state* (Norris 1998).

Further properties of this equilibrium state are due to the features of the transition graph. The fact that every state is accessible from every other state through a path of sufficient length is known as *ergodicity.* Moreover, as paths of sufficient length intersect with each other, the system exhibits decaying temporal correlations, which is known as the *mixing property.*

To conclude, two interesting cases for finite state Markov chains can be distinguished. If an integer power of the transition matrix is diagonal, the Markov chain possesses invariant and ergodic, but not mixing, equilibrium states. If, on the other hand, an integer power of the transition matrix is strictly positive, the Markov chain has invariant, ergodic and mixing equilibrium states corresponding to thermal equilibrium states according to the KMS criterion. For more details see the relevant literature, for instance Norris (1998).

Acknowledgements

This work has been supported by an EPSCR Bridging the Gaps grant on *Cognitive Systems Sciences*. Adam Barrett was partly supported by a postdoctoral fellowship at the Institute for Adaptive and Neural Computation in the School of Informatics at the University of Edinburgh, funded by the Human Frontier Science Program. We thank Slawomir Nasuto and two anonymous referees for substantial help improving the paper.

References

Allefeld, C., Atmanspacher, H., and Wackermann, J. (2009). Mental states as macrostates emerging from EEG dynamics. *Chaos*, 19:015102.

Amari, S.-I. (1974). A method of statistical neurodynamics. *Kybernetik*, 14:201 – 215.

Amari, S.-I. and Maginu, K. (1988). Statistical neurodynamics of associative memory. *Neural Networks*, 1(1):63 – 73.

Amari, S.-I., Yoshida, K., and Kanatani, K.-I. (1977). A mathematical foundation for statistical neurodynamics. *SIAM Journal of Applied Mathematics*, 33(1):95 – 126.

Anderson, J. A. and Rosenfeld, E., editors (1988). *Neurocomputing. Foundations of Research*, volume 1. MIT Press, Cambridge (MA).

Atmanspacher, H. (2007). Contextual emergence from physics to cognitive neuroscience. *Journal of Consciousness Studies*, 14(1/2):18 – 36.

Atmanspacher, H. and beim Graben, P. (2007). Contextual emergence of mental states from neurodynamics. *Chaos and Complexity Letters*, 2(2/3):151 – 168.

Atmanspacher, H. and Rotter, S. (2008). Interpreting neurodynamics: concepts and facts. *Cognitive Neurodynamics*, 2(4):297 – 318.

Batterman, R. (2002). *The Devil in the Details*. Oxford University Press, Oxford.

Bishop, R. C. and Atmanspacher, H. (2006). Contextual emergence in the description of properties. *Foundations of Physics*, 36(12):1753 – 1777.

Bollobás, B. (2001). *Random Graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (UK).

Chalmers, D. J. (2000). What is a neural correlate of consciousness? In Metzinger (2000), pages 17 – 39.

Exel, R. (2004). KMS states for generalized gauge actions on Cuntz-Krieger algebras (An application of the Ruelle-Perron-Frobenius theorem). *Bulletin of the Brazilian Mathematical Society*, 35(1):1 – 12.

Freeman, W. J. (2007). Definitions of state variables and state space for brain-computer interface. part 1. multiple hierarchical levels of brain function. *Cognitive Neurodynamics*, 1:3 – 14.

Gaveau, B. and Schulman, L. S. (2005). Dynamical distance: coarse grains, pattern recognition, and network analysis. *Bulletin des Sciences Mathematiques*, 129:631 – 642.

beim Graben, P. (2004). Incompatible implementations of physical symbol systems. *Mind and Matter*, 2(2):29 – 51.

beim Graben, P. (2008). Foundations of neurophysics. In beim Graben, P., Zhou, C., Thiel, M., and Kurths, J., editors, *Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain Networks*, pages 3 – 48. Springer, Berlin.

beim Graben, P. and Atmanspacher, H. (2006). Complementarity in classical dynamical systems. *Foundations of Physics*, 36(2):291 – 306.

beim Graben, P. and Kurths, J. (2008). Simulating global properties of electroencephalograms with minimal random neural networks. *Neurocomputing*, 71(4):999 – 1007.

Guckenheimer, J. and Holmes, P. (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer, New York.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the U.S.A.*, 79(8):2554 – 2558.

Lind, D. and Marcus, B. (1995). *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge (UK).

Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531 – 2560.

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115 – 133. Reprinted in J. A. Anderson and E. Rosenfeld (1988), pp. 18ff.

Metzinger, T., editor (2000). *Neural Correlates of Consciousness*. MIT Press, Cambridge (MA).

Norris, J.R. (1998). *Markov Chains*. Cambridge University Press, Cambridge.

Olesen, D. and Pedersen, G. K. (1978). Some C*-dynamical systems with a single KMS state. *Mathematica Scandinavia*, 42:111 – 118.

Pinzari, C., Watatani, Y., and Yonetani, K. (2000). KMS states, entropy and the variational principle in full C*-dynamical systems. *Communications of Mathematical Physics*, 213:331 – 379.

Primas, H. (1998). Emergence in exact natural sciences. *Acta Polytechnica Scandinavica*, Ma-91:83 – 98.

Ruelle, D. (1968). Statistical mechanics of a one-dimensional lattice gas. *Communications of Mathematical Physics*, 9:267 – 278.

Ruelle, D. (1989). The thermodynamic formalism for expanding maps. *Communications of Mathematical Physics*, 125:239 – 262.

Sewell, G. L. (2002). *Quantum Mechanics and its Emergent Macrophysics*. Princeton University Press.

Shalizi, C. R. and Moore, C. (2003). What is a macrostate? subjective observations and objective dynamics. arXiv: cond-mat/0303625.

Sinai, Y. G. (1968). Construction of Markov partitions. *Functional Analysis and its Applications*, 2(3):245 – 253.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11(1):1 – 74.

Smolensky, P. (2006). Harmony in linguistic cognition. *Cognitive Science*, 30:779 – 801.

Touboul, J., Faugeras, O., and Cessac, B. (2008). Mean field analyis of multi-population neural networks with random synaptic weights and stochastic input. Technical Report 6454, INRIA, Sophia Antipolis.

Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12:1 – 24.