

# Transients, Metastability, and Neuronal Dynamics

Karl J. Friston

*The Wellcome Department of Cognitive Neurology, Institute of Neurology, Queen Square, London WC1N 3BG, United Kingdom*

Received September 3, 1996

---

**This paper is about neuronal dynamics and how their special complexity can be understood in terms of nonlinear dynamics. There are many aspects of neuronal interactions and connectivity that engender the complexity of brain dynamics. In this paper we consider (i) the nature of this complexity and (ii) how it depends on connections between neuronal systems (e.g., neuronal populations or cortical areas). The main conclusion is that *simulated neural systems show complex behaviors, reminiscent of neuronal dynamics, when these extrinsic connections are sparse. The patterns of activity that obtain, under these conditions, show a rich form of intermittency with the recurrent and self-limiting expression of stereotyped transient-like dynamics. Despite the fact that these dynamics conform to a single (complex) attractor this metastability gives the illusion of a dynamically changing attractor manifold (i.e., a changing surface upon which the dynamics unfold). This metastability is characterized using a measure that is based on the entropy of the time series' spectral density.*** © 1997 Academic Press

---

## INTRODUCTION

This paper is concerned with the nature and genesis of the complicated dynamics observed in the brain. The question we started with was "Are brain dynamics best described by a single dynamical system (i.e., a global attractor) or an ensemble of separable systems (i.e., a collection of smaller attractors)?" We concluded that both viewpoints can be reconciled by representing brain dynamics as a global attractor, where this single attractor has a special complexity that emulates a succession of transient-like dynamics, each with its own distinct and recurring spatiotemporal organization. This behavior can arise when the connectivity, among simulated neuronal populations, is sparse. The aim of this paper is to describe how these conclusions were reached.

### Complexity and Functional Integration

The brain appears to conform to two fundamental principles of organization: *functional segregation* and

*functional integration* (e.g., Zeki, 1990). Functional segregation requires the dynamics of each area to be distinct, in terms of its intrinsic activity and responses to input. Functional integration, on the other hand, requires segregated areas to influence each other in a way that facilitates coherent integration and the motor behaviors that ensue. It has been proposed that the resolution of this dialectic, between the preservation of regionally specific dynamics and global coherence, is a hallmark of complexity (Tononi *et al.*, 1994; Friston *et al.*, 1995a). A measure of this complexity, based on the theory of stochastic processes and information theory, has been described (Tononi *et al.*, 1994). The present work uses a nonlinear framework to address the dynamic or temporal complexity of systems like the brain.

There are clearly many aspects of neuronal interactions and connectivity that can render neuronal dynamics complex. In this paper we first consider what happens when the coupling between simulated neuronal populations is systematically increased. The degree of coupling or integration among neuronal populations [e.g., neuronal groups (Edelman, 1993), functionally specialized patches in extrastriate cortex or functionally segregated cortical areas like V5 (Zeki, 1990)] is particularly important when considering functional integration in the context of functional segregation. Functional integration can be emulated by increasing the extrinsic connectivity among simulated neuronal populations that, in the absence of extrinsic connections, each express their own dynamics. On the basis of simulations and electrophysiology one can make some predictions about the effects of modulating extrinsic connectivity: As the coupling increases the dynamics should come to resemble the complicated, intermittent dynamics seen in the real brain. In this regime several populations may interact in a coherent way [e.g., phase locking among units (e.g., Gray and Singer, 1989) or populations (e.g., Sporns *et al.*, 1989)], creating spatiotemporal patterns of activity that include many, if not all, of the system's components. In real brains these transient dynamics are generally short-lived, with new patterns being continually created, destroyed, and subsequently recreated. This regime could be likened to *intermittency* in simple nonlinear systems (Tsonis, 1992)

or to *dynamic instabilities* in self-organizing systems (Kelso, 1995). As extrinsic connectivity is increased further the dynamics cease to be complex, with every component locked into a single, coherent pattern of activity (see Friston *et al.*, 1995a). The intermediate regime of intermittent and dynamic instability is the subject of this work.

The term “transient” is used here to denote a distinct, self-limiting stereotyped pattern of activity, by analogy to its use in dynamical systems theory. Generally a transient describes the behavior of a system that occurs in the initial period as the system approaches an attractor. In this work, systems are considered to be perpetually in an “initial period” by virtue of continuous changes in the underlying attractor. These changes may be real, due to changes in control parameters (e.g., changes in connection strengths caused by modulatory interactions), or apparent. This paper is concerned with *apparent* changes in the attractor that arise when the system moves to a different part of the attractor, giving the impression that the attractor itself has changed. This apparent “switching” from one dynamic to another is characterized here as dynamic instability or metastability (Kelso, 1995). See Kelso (1995) for a description of dynamic instability and its relationship to other aspects of dynamical systems.

This paper is divided into three sections. The first section addresses the effect of increasing the extrinsic connectivity between simulated neuronal populations. In this section we observe that, as predicted, the dynamics move from stable incoherence, where each population preserves its own unique oscillatory dynamics, through a regime of metastability (transients and periods of stable coherence that are themselves inherently unstable) (Kelso, 1995), to, finally, a regime of stable coherence with phase locking and complete entrainment. A key feature of the metastable regime is a dynamic modulation of the frequencies expressed by the succession of transients. This characteristic changeability in the spectral density is used as a measure of metastability in the subsequent section.

The second section attempts to characterize metastability using the uncertainty (entropy) of the spectral density measured repeatedly over short periods of time. This measure is then used to demonstrate that metastability shows an inverted U dependency on the density of extrinsic couplings between neuronal populations, reaching a maximum when these connections are present but sparse. In contradistinction the “strangeness” or dimensional complexity (as measured with the correlation dimension) falls monotonically, suggesting that metastable complexity and chaos are fundamentally different things. These findings are consistent with the analysis presented in Friston *et al.* (1995a) using a linear measure of complexity.

The third section presents a further perspective on

metastability in terms of attractor manifolds. An attractor is simply the surface (i.e., manifold) that contains the trajectory traced out when time series from all the system’s components [e.g., channels in multichannel magnetoencephalography recordings] are plotted against each other. If the attractor is a strange attractor (i.e., has a fractional dimensionality greater than 2) the system is called chaotic. Strange or “chaotic” attractors are characterized by a deterministic unpredictability in their evolution, sometimes referred to as “sensitivity to initial conditions” and related to an exponential divergence of trajectories on the attractor manifold (Tsonis, 1992). The distinction between “chaos” and “complexity” is a crucial one and represents a key focus of this paper.

## THE EFFECT OF INCREASING EXTRINSIC CONNECTIVITY

This section describes the simulated neuronal system used to examine the effect of changing the sparsity of extrinsic connectivity among neuronal populations. In brief we show that as extrinsic connectivity is increased the system passes through a regime of complicated metastable behavior into a regime of global coherence. In what follows vectors are denoted by bold lowercase letters and matrices by bold uppercase letters.

### The Nonlinear Simulation

The simulations comprised three groups of 6, 8, and 16 units. Within each group every unit was vicariously connected to every other unit with one excitatory connection. All the units within a group were directly interconnected with inhibitory connections (cf., GABA inhibitory interneurons). These intrinsic or within-group connectivities were chosen to ensure chaotic dynamics. The extrinsic connections between groups were excitatory (cf., glutaminergic corticocortical projections). The within-group excitatory and inhibitory connectivity matrices ( $\mathbf{E}_w$  and  $\mathbf{I}$ ) comprised connection strengths selected from a uniform random distribution [0, 1] and scaled such that the sum of squares over all elements was 0.2. To ensure that the system was dissipative we added 0.125 to the random elements of the leading diagonal of  $\mathbf{I}$  (before normalization). This models a decay in activity or adaptation in the real brain. Between-group connections were based on the matrix  $\mathbf{E}_b$ , whose elements were selected from the unit normal distribution. These connections strengths were then transformed using a sigmoid function to lie in the range [0, 0.0008]. The sparsity or distribution of the latter connections was determined by a control parameter  $\alpha$  (see below). Dynamics were obtained by integra-

tion of

$$\begin{aligned} \partial s_i(t)/\partial t &= f_i[s_1(t) \cdots s_n(t)] \\ &= \sum E_{ij} \cdot s_j(t) - s_i(t) \cdot \sum I_{ij} \cdot s_j(t), \end{aligned}$$

or, in matrix notation

$$\partial \mathbf{s}(t)/\partial t = \mathbf{E} \cdot \mathbf{s}(t) - \text{diag}(\mathbf{s}(t)) \cdot \mathbf{I} \cdot \mathbf{s}(t), \quad (1)$$

where

$$\mathbf{E} = \mathbf{E}_w + c(\alpha, \mathbf{E}_b)$$

and

$$c(\alpha, \mathbf{E}_b) = 0.0008 \cdot (\tan h(\alpha + 2 \cdot \mathbf{E}_b) + 1)/2.$$

$E_{ij}$  are the elements of  $\mathbf{E}$  and  $I_{ij}$  are the elements of  $\mathbf{I}$ .  $\mathbf{s}(t)$  is a column vector with elements  $s_i(t)$  representing the activity of the  $i$ th unit.  $\partial s_i(t)/\partial t$  is the change in activity of unit  $i$  per unit time.  $c(\alpha, \mathbf{E}_b)$  is an element-wise matrix function of  $\mathbf{E}_b$  and returns a matrix of connection strengths that constitutes the between-group components of excitatory connectivity  $\mathbf{E}$ . This contribution increases with  $\alpha$  in a way that allowed us to manipulate the number of relatively strong connections (i.e., sparsity) in a continuous fashion.

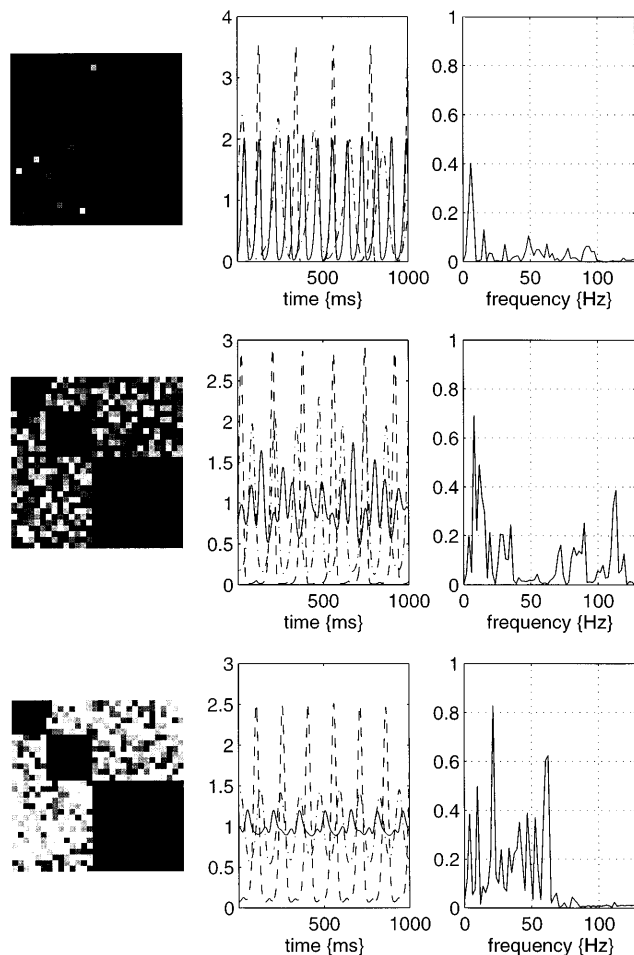
The form of these equations means that excitatory inputs from unit  $j$  increase activity in unit  $i$  according to the excitatory connection strength  $E_{ij}$ . The inhibitory inputs, mediated by the inhibitory connections  $I_{ij}$  are modulated by activity intrinsic to the unit in question. This nonlinear interaction emulates voltage-dependent inhibition, where the effect is only realized in the presence of postsynaptic depolarization. This form of state equation [Eq.(1)] also ensures positivity of the activities  $s_i(t)$  given all  $E_{ij}$  and  $I_{ij} > 0$ . The resulting activities are then interpretable as instantaneous firing rates. Although Eq.(1) may seem very simple it can lead to markedly nonlinear dynamics reminiscent of neuronal systems with spontaneous periodic bursting. We have previously used this model to estimate modulatory interactions in human visual cortex using fMRI data (Friston *et al.*, 1995b).

Each simulation comprised  $2^{14}$  iterations following a 2048 iteration “burn in.” This initial period ensured that transients due simply to the initial conditions had died away. The initial activities were selected from the uniform random distribution [0, 1]. Each iteration corresponded to a unit of time which, to relate these simulations to real neuronal dynamics, was considered to be a millisecond (i.e.,  $\partial t = 1$  ms). Each simulation can then be thought of as lasting about 16 s. The dominant oscillatory dynamics that result from these simulations then correspond to the  $\alpha$  range seen in the brain. The

intrinsic time constants (i.e., width of the autocorrelation function of unit activity) for the systems simulated was in the order of 24–64 ms.

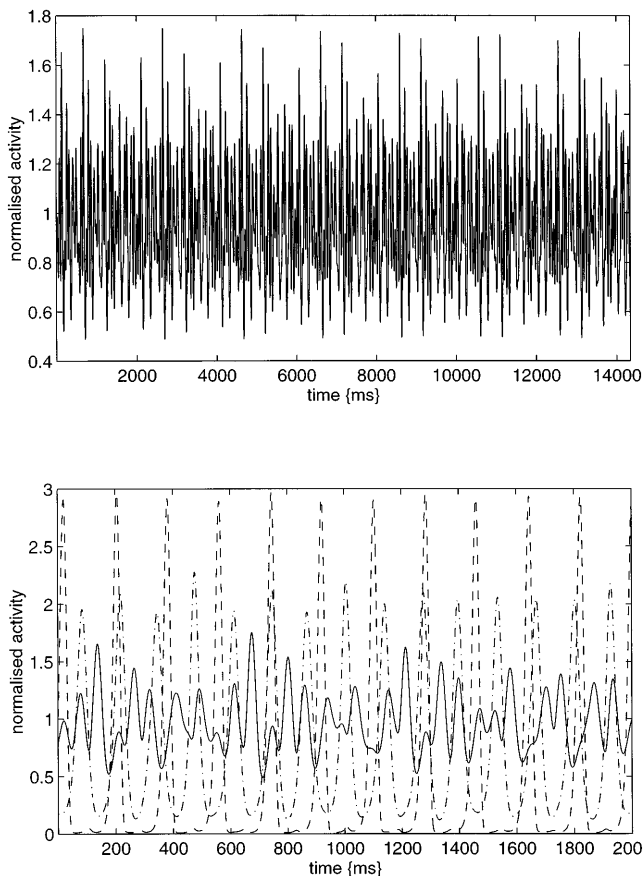
### Increasing Extrinsic Connectivity

The effect of increasing extrinsic connectivity was investigated by repeating the simulations, as described above, using low, intermediate, and high values of  $\alpha$  (i.e., scant, sparse, and dense interconnectivity). Figure 1 shows the results of a typical set of simulations for a given set of connectivity matrices. On the left are image representations of the extrinsic connectivities  $c(\alpha, \mathbf{E}_b)$ , all scaled to the same maximum. The middle graphs



**FIG. 1.** (Left) Extrinsic connectivity matrices corresponding to  $c(\alpha, \mathbf{E}_b)$  in the main text. The gray scale is (0–0.0008). (Middle) Example of the dynamics over 1000 iterations of Eq. (1). The activity of the first unit from each of the three groups is shown (first group—solid line, second group—dot-dash line, and third group—dashed line). These time series have been normalized. Top row, scant connectivity; middle row, sparse connectivity; and bottom row, dense connectivity. (Right) Coherence functions between the first unit of the first and third groups estimated using Welch’s averaged periodogram method and a window length of 512 ms.

show a segment of the resulting dynamics (over 1000 of the  $2^{14}$  iterations). The time series from the first unit in each of the groups are shown. The right graphs show the coherence between the first unit of the first and last groups estimated using Welch's averaged periodogram method. These functions reflect the degree of phase-locked coherent interactions between the two groups over the frequencies shown. The top row ( $\alpha = -5$ ) represents scant extrinsic connectivity, wherein the dynamics of each of the three groups are largely independent. This stable but uncoupled pattern of activity is evident from the distinct frequencies at which the three units show periodic bursts of activity (middle graph) and the minimal amount of coherence between the groups (right graph). When the extrinsic connectivity is increased (middle row,  $\alpha = -1$ ) the dynamics show pronounced but complicated interactions. These dynamics are neither independent or coherent. These dynamics are shown in more detail in Fig. 2 for an extended period of time (top, first unit only) and for 2000 ms (bottom, all three units). The



**FIG. 2.** Metastable dynamics. (Top) The activity of the first unit of the first group over the entire simulation (after the initial period was discarded). (Bottom) The dynamics of all three first units over 2000 ms as in the middle row of Fig. 1. These time series have been normalized.

coherence between the first and third groups has increased and is fairly “broad band” (right in Fig. 1). The first unit (solid line) reflects a degree of metastability, with short-lived doublets or triplets of bursts, interrupted by apparent suppression by the third unit (broken line). This “unstable stability” is a hallmark of metastable complexity and can be characterized as the *successive expression of a series of stereotyped transients*. On further increasing the connectivity (bottom row of Fig. 1,  $\alpha = 1.2$ ) the dynamics become coherent with clear phase locking and coherence at about 20 and 60 Hz (right). The interactions between the first and the second (dot-dash line) speak to a degree of metastability, but less pronounced than in the previous simulation.

If metastability is characterized by transient periods of stability, or the recurrent expression of different transients, then the frequency composition, or spectral density of the time series should change with time. However, if the dynamics are stable, then the corresponding spectral densities will not change, irrespective of whether that stability results from the expression of independent intrinsic dynamics (i.e., no connectivity) or from complete entrainment and coherence (i.e., dense connectivity). Consequently the changeability or stability of the spectral densities could be used to measure metastability.

## MEASURING METASTABILITY

In this section we describe a simple measure of metastability, framed in terms of the instability or entropy of the spectral density of a time series. This measure is then applied to the simulations of the previous section, to characterize the relationship between the sparsity of extrinsic connections, dimensional complexity, and metastability.

### Spectral Density

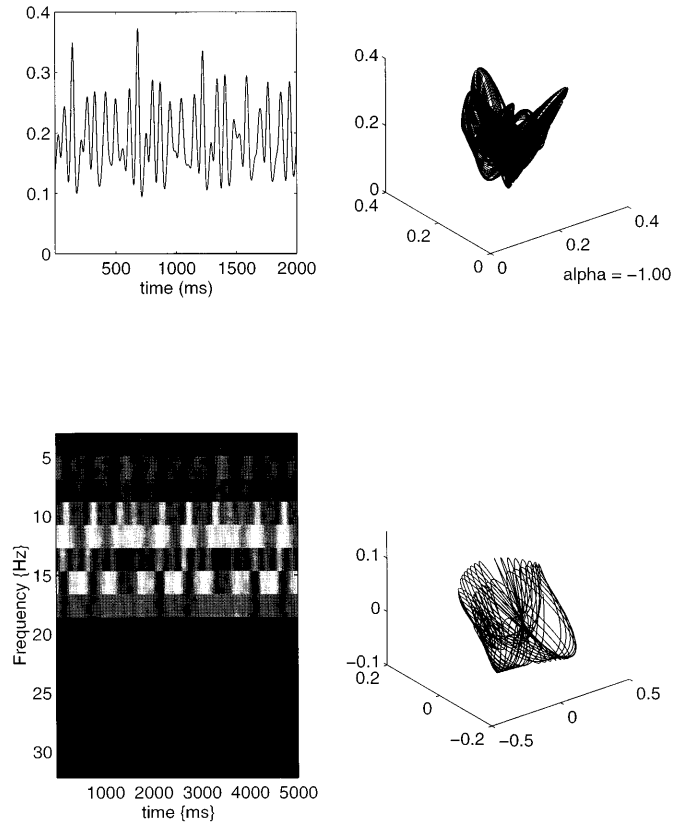
Using a continuous time formulation, for any given time series  $s(t)$  the (time-dependent) spectral density  $g(\omega, t)$  can be estimated with:

$$g(\omega, t) = |f(\omega, t)|^2,$$

where

$$\begin{aligned} f(\omega, t) &= s(t) \otimes [h(t) \cdot \exp(-j\omega t)] \\ &= \int h(u) \cdot \exp(-j\omega u) \cdot s(t - u) \cdot du. \end{aligned} \tag{2}$$

$\omega$  is  $2\pi$  times the frequency in question and  $j = \sqrt{-1}$ . Here  $\otimes$  denotes convolution and  $h(t)$  is some suitable windowing function of length  $l$ . A Hanning function [a bell-shaped function  $= (1 - \cos(2\pi t/(l + 1)))/2$ ] with  $l = 512$  iterations or milliseconds was used in this paper.



**FIG. 3.** (Top left) Dynamics of the first unit of the first group over 2000 iterations. (Top right) Phase portrait or dynamical attractor of the same time series using temporal embedding and a lag of 64 iterations. (Lower left) Spectral density as a function of time shown in image format. The data have been mean corrected. The gray scale is arbitrary. (Lower right) The spectral density attractor depicted in terms of the first three principal components or modes of the spectral density time series.

The temporal length of  $h(t)$  determines the period over which the spectral density is estimated. Figure 3 presents a spectral density analysis of the time series from the first unit of the first group (the data presented in Fig. 2). A 2000-ms segment of this time series is shown on the upper left. The same data are shown on the upper right as a phase portrait using temporal embedding. This is simply a plot of  $s(t)$ ,  $s(t - \tau)$ , and  $s(t - 2 \cdot \tau)$ , where in this instance  $\tau = 64$  iterations (roughly the decay of the autocorrelation function). This phase portrait does not constitute an analysis but is simply a way of visualizing the underlying attractor. The lower left is an image representation of  $g(\omega, t)$  over 5000 ms. It can be seen that the spectral densities themselves change with time and display chaotic behavior. The corresponding *spectral density attractor* is shown on the lower right and was constructed by plotting the activities of the first three principal components or modes, of the spectral density time series, against each other. These activities are given by the dot

product  $g(\omega, t) \cdot p_i(\omega)$ , where  $p_i(\omega)$  is the  $i$ th eigenvector of the spectral density covariance matrix  $\text{Cov} [g(\omega, t)]$ . The elements in the  $i$ th row and  $j$ th column of  $\text{Cov} [g(\omega, t)]$  comprise the covariance between the spectral densities at  $\omega_i$  and  $\omega_j$  over time. This is simply a device to plot a three-dimensional view of a  $m$ -dimensional attractor in a fashion that reveals the most structure.  $m$  is the number of frequencies that were measured, in this case 16.

The spectral density attractor has an interesting interpretation wherein each region corresponds to one or more transients in  $s(t)$ : Each time a transient is expressed (with a different spectral density) the trajectory in spectral density space moves to a new region. In other words one region in spectral density space corresponds to a “transient” in the original time series and, equivalently, to a particular submanifold of the original dynamical attractor (top right).

### A Measure of Metastability

Intuitively it can be seen that a large spectral density attractor “covers” more regions and reflects a greater number and diversity of transients (i.e., metastability). A measure of the volume of the spectral density attractor is provided by the entropy  $H$  where, under Gaussian assumptions (Jones, 1979),

$$H = \log (2\pi e^m \det [\text{Cov} \{g(\omega, t)\}]/2). \quad (3)$$

det  $\{\cdot\}$  means the determinant of a matrix. This simple expression provides the measure used below to assess metastability as a function of extrinsic connectivity.

### Dimensional Complexity and Metastability

It is important to appreciate that the complexity embodied in metastable dynamics is very different from that measured using other nonlinear characterizations. The correlation dimension ( $D_2$ ) is a commonly used nonlinear measure and reflects the degree of chaos, or strangeness, of the underlying attractor. The  $D_2$  is often referred to as “dimensional complexity” and reflects the space-filling nature of the trajectory: It is related to the average exponential divergence of nearby trajectories on the attractor manifold. In this work we estimated  $D_2$  using the Lyapunov exponents of the system’s trajectory according to the Kaplan–Yorke conjecture (Kaplan and Yorke, 1979). The Lyapunov exponents measure the degree of exponential divergence of nearby trajectories and correspond to  $\lambda$ , the eigenvalues of the Jacobian matrix (Tsonis, 1992). The elements of the Jacobian  $\mathbf{J}$  are  $\partial f_i / \partial s_j$  and were derived using Eq. (1). This method of calculating the Lyapunov exponents is known as the Jacobian method and depends on knowing the state equation governing the system’s behavior. It is simple to show that small perturbations

$\mathbf{s}^*(t)$  from the trajectory  $\mathbf{s}(t)$  evolve according to  $\partial \mathbf{s}^*(t) / \partial t \approx \mathbf{J} \cdot \mathbf{s}^*(t)$  with the solution  $\mathbf{s}^*(t) \approx \text{expm} \{ \mathbf{J} \cdot t \} \cdot \mathbf{s}^*(0)$  (Tsonis, 1992), where  $\text{expm} \{ \cdot \}$  is a matrix exponential. Consider the evolution of an initial perturbation equal to the  $k$ th eigenvector of  $\mathbf{J}$ , i.e.,  $\mathbf{s}^*(0) = \mathbf{e}_k$  which is given by  $\mathbf{s}^*(t) \approx \text{expm} \{ \mathbf{J} \cdot t \} \cdot \mathbf{e}_k = \mathbf{e}_k \cdot \exp(\lambda_k \cdot t)$ . In other words, the perturbation along the  $k$ th principal axis of divergence increases exponentially with exponent  $\lambda_k$ . Because the Lyapunov exponents are themselves time dependent (Tsonis, 1992) we used the expectation of their real components evaluated over the system's trajectory.

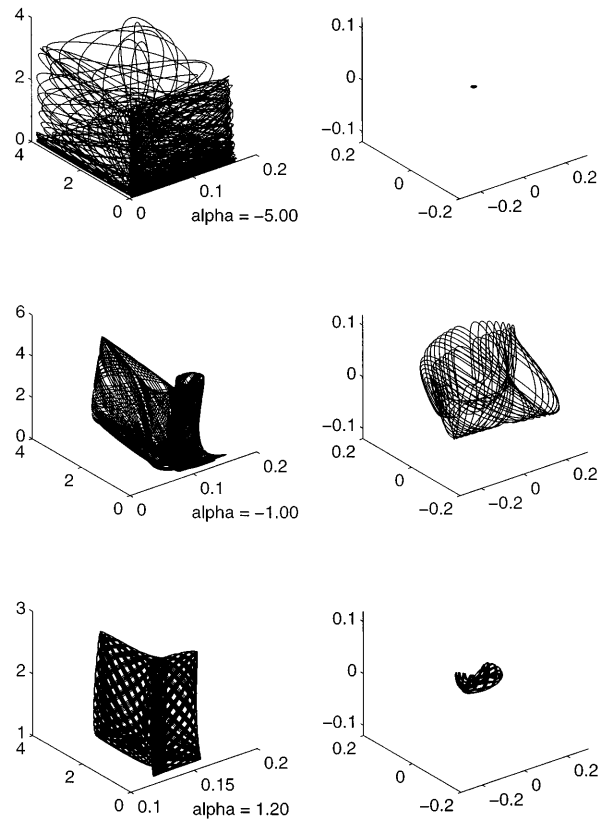
Our previous analyses (Friston *et al.*, 1995a) suggested that dimensional complexity falls monotonically as extrinsic connectivity is increased. This can be seen intuitively if one considers that, in the absence of any connectivity, each isolated group contributes its own dimensions to the overall dimension. As connectivity increases the system ceases to be a collection of separate chaotic systems and starts to behave as a single system with a relatively lower dimensionality. This is in contrast to the expression of metastability, which rises and then falls. This point is made in Fig. 4. Figure 4 presents the dynamical and spectral density attractors for each of the three simulations presented in Fig. 1 (top—scant connectivity, middle—sparse connectivity, and lower—dense connectivity). The dynamical attractors shown correspond to a trajectory traced out by the first unit of each group. As extrinsic connectivity increases, the space-filling nature of these attractors falls with a corresponding reduction in dimensional complexity (see below). Conversely the space-filling nature of the spectral density attractor increases with the expression of metastable dynamics and then falls again.

We measured the degree of metastability ( $H$ , for the first unit) and the dimensional complexity ( $D_2$ ) as a function of extrinsic connectivity  $\alpha$ . As predicted, the dimensional complexity fell monotonically with increasing  $\alpha$ . Conversely  $H$  rose and fell (see Fig. 5). The distribution of extrinsic connection strengths that gave rise to the greatest metastability corresponded to a sparsity of 0.12, using a threshold a 50% of the maximum strength to define a connection as “present.”

### METASTABILITY AND COMPLEX ATTRACTOR MANIFOLDS

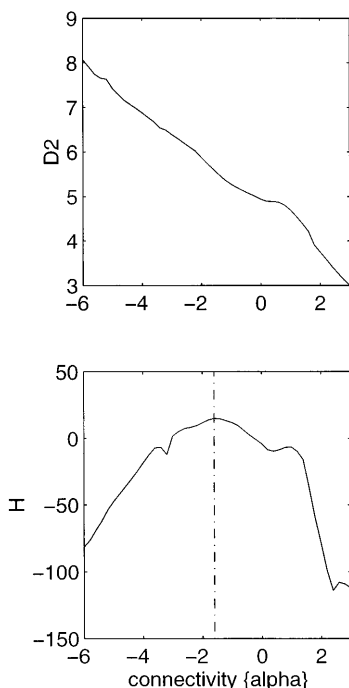
We have framed metastability in terms of the successive expression of transients that emerge when simulated neuronal populations are loosely coupled or sparsely connected. In this section we provide a heuristic reformulation of metastability in terms of the underlying attractor.

Although chaotic systems can be represented by a single attractor manifold, the recurrent creation and



**FIG. 4.** (Left) Phase portrait or dynamical attractor obtained by plotting the activities of the first unit from each group against each other. (Right) The corresponding spectral density attractors based on the unit from the first group. Top row, scant connectivity; middle row, sparse connectivity; and bottom row, dense connectivity. The axes of the three spectral density attractors have been made the same to enable comparison.

destruction of transient-like dynamics can create an impression of instability, where the attractor manifold itself appears to change with time. Of course this is not the case because there is only one attractor surface, but if the system were observed for short periods at a time, one would see one transient dynamic, then another, and then another or the first again. This succession of self-limiting, recurring patterns has been referred to as *metastability* following Kelso (1995). This phenomenon can be understood in terms of a complex attractor surface that traps the trajectory locally in some “sub-manifold” before it escapes to another locally structured part of the attractor surface. These submanifolds are exactly the same as a normal attractor manifold but for the fact that they are connected to, or embedded in, a larger surface. At some point the trajectory will find this connection and a new transient will emerge as the trajectory moves off to another submanifold. We suggest that attractors that have this property are complex. Note that the complexity of the attractor is not directly related to its dimensionality or strangeness



**FIG. 5.** (Top) Dimensional complexity ( $D_2$ ) and (bottom) metastability ( $H$ ) as functions of extrinsic connectivity ( $\alpha$ ) between the simulated neuronal groups. The key thing to note is that the two forms of complexity are dissociable and that  $H$  peaks in a regime of sparse connectivity.

(e.g., as measured by the  $D_2$ ). The latter measures pertain to the *space-filling nature of the manifold*, averaged over its entire surface. Complexity, as discussed here, relates to the *shape of the manifold*, where this shape can be characterized as a set of connected submanifolds, each capable of sequestering the trajectory for a limited period of time [see also Freeman and Barrie (1994) for a convergent discussion].

If the system were observed over a short period it may not be possible to differentiate between a true (simple) attractor manifold and a submanifold. However, with continued observation, if the observed manifold is part of a complex manifold it will, ultimately, change. It is this apparent dynamical change in the attractor that characterizes metastability and provides the basis for the measure introduced in the previous section (i.e., the uncertainty about the attractor when repeatedly observed for short periods). More exactly we have used the entropy of a multivariate “signature” (spectral density) of the extant submanifold.

## CONCLUSION

In summary complicated metastable dynamics can occur when the connectivity among simulated neuronal populations is sparse. Complexity of this sort is characterized by a succession of transient-like dynamics that

gives the illusion of a continuously changing attractor manifold.

Neuronal dynamics can be characterized as a temporal succession of transients (Friston, 1995). See Mayer-Kress *et al.* (1991) and Fuchs *et al.* (1992) for compelling examples and Pfurtscheller and Aranibar (1979) for spectral density changes in relation to self-paced movement. On the basis of this, and in the light of the simulations presented above, we infer that neuronal dynamics are modeled by neither an ensemble of separate attractors nor a simple low-dimensional attractor, but are consistent with the attractor surface that ensues when many separate attractors are loosely coupled together. This manifold has a special complexity, where the trajectories upon it show complicated metastable dynamics, with the recurrent appearance and destruction of transient-like dynamics. A complicated manifold is not necessarily associated with a high-dimensional complexity, because its main feature is one of local entrapment or “lingering” of the trajectory in submanifolds (as opposed to its space-filling nature).

The technique presented here for measuring the degree of metastability is very simple and may provide another perspective when characterizing dynamical systems. This is particularly important because nonlinear analysis procedures such as the correlation dimension are often very difficult to apply to empirical data and a more complete picture emerges when several complementary approaches are used.

It is proposed that the complex nature of nonlinear systems like the brain includes metastability. Furthermore, in keeping with much of the current thinking on complexity in self-organizing systems, this rich form of intermittency, dynamical instability, or metastability is found in regimes of parameter space near critical points or phase transitions (e.g., Kauffman, 1992; Kelso, 1995). This work suggests that, for the brain, critical regions involve *sparse* extrinsic connectivity.

## ACKNOWLEDGMENTS

K.J.F. was funded by the Wellcome Trust. I thank my colleagues for invaluable discussions and help with the presentation of this work, particularly, Chris Frith, Ray Dolan, Richard Frackowiak, and Cathy Price.

## REFERENCES

- Edelman, G. M. 1993. Neural Darwinism: Selection and reentrant signalling in higher brain function. *Neuron* **10**:115–125.
- Freeman, W., and Barrie, J. 1994. Chaotic oscillations and the genesis of meaning in cerebral cortex. In *Temporal Coding in the Brain* (Buzsaki, R. Llinas, W. Singer, A. Berthoz, and T. Christen, Eds.), pp. 13–38. Springer-Verlag, Berlin.
- Friston, K.J. 1995. Neuronal transients. *Proc. R. Soc. London Ser. B* **261**:401–405.
- Friston, K. J., Tononi, G., Sporns, O., and Edelman, G. M. 1995a. Characterising the complexity of neuronal interactions. *Hum. Brain Map.* **3**:302–314.

- Friston, K. J., Ungerleider, L. G., Jezzard, P., and Turner, R. 1995b. Characterising modulatory interactions between V1 and V2 in human cortex with fMRI. *Hum. Brain Map.* **2**:211–224.
- Fuchs, A., Kelso, J. A. S., and Haken, H. 1992. Phase transitions in the human brain: Spatial mode dynamics. *Int. J. Bifurcat. Chaos* **2**:917–939.
- Gray, C. M., and Singer, W. 1989. Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* **86**:1698–1702.
- Jones, D. S. 1979. *Elementary Information Theory*. Clarendon Press, Oxford.
- Kaplan, J., and Yorke, J. 1979. Chaotic behavior of multidimensional difference equations. In *Functional Differential Equations and Approximation of Fixed Points* H. O. Peitgen and H. O. Walther, (Eds.). Springer, Berlin/New York.
- Kauffman, S. A. 1992. The sciences of complexity and “origins of order.” In *The Principles of Organization in Organisms* (J. E. Mittenthal and A. B. Baskin, Eds.), pp. 303–320. Addison–Wesley, New York.
- Kelso, J. A. S. 1995. *Review of Dynamic Patterns: The Self-Organisation of Brain and Behavior*. MIT Press, Cambridge, MA.
- Mayer-Kress, G., Barczys, C., and Freeman, W. 1991. Attractor reconstruction from event-related multi-electrode EEG-data. In *Mathematical Approaches to Brain Functioning Diagnostics* (I. Dvorak and A. V. Holden, Eds.), pp. 315–336. Manchester Univ. Press, New York.
- Pfurtscheller, G., and Aranibar, A. 1979. Evaluation of event-related desynchronisation (ERD) preceding and following voluntary self-paced movement. *Electroencephalogr. Clin. Neurophysiol.* **46**:138–146.
- Sporns, O., Gally, J. A., Reeke, G. N., and Edelman, G.M. M. 1989. Reentrant signalling among simulated neuronal groups leads to coherence in their oscillatory activity. *Proc. Natl. Acad. Sci. USA* **86**:7265–7269.
- Tononi, G., Sporns, O., and Edelman, G. M. 1994. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA* **91**:5033–5037.
- Tsonis, A. A. 1992. *Chaos: From Theory to Applications*. Plenum, New York.
- Zeki, S. 1990. The motion pathways of the visual cortex. In *Vision: Coding and Efficiency* (C. Blakemore, Ed.), pp. 321–345. Cambridge Univ. Press, Cambridge, UK.