Physica D 41 (1990) 137–172
North-Holland

CLUSTERING, CODING, SWITCHING, HIERARCHICAL ORDERING, AND CONTROL IN A NETWORK OF CHAOTIC ELEMENTS

Kunihiko KANEKO[1]
Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Received 16 May 1989
Revised manuscript received 28 September 1989
Accepted 2 October 1989
Communicated by A.V. Holden

A network of chaotic elements is investigated with the use of globally coupled maps. A simple coding of many attractors with clustering is shown. Through the coding, the attractors are organized so that their change exhibits bifurcation-like phenomena. A precision-dependent tree is constructed which leads to the similarity of our attractor with those of spin-glasses. Hierarchical dynamics is constructed on the tree, which leads to the dynamical change of trees and the temporal change of effective degrees of freedom. By a simple input on a site, we can switch among attractors and tune the strength of chaos. A threshold on a cluster size is found, beyond which a peculiar "posi–nega" switch occurs. Possible application to biological information processing is discussed with the emphasis on the fuzzy switch (chaotic search) and hierarchical code (categorization).

Contents

1. Introduction

The study of networks of chaotic elements is important not only as a model for nonlinear systems with many degrees of freedom, but also from the viewpoint of biological information processing and possible engineering applications. Here we introduce the "globally coupled map" (GCM) as the extreme limit of long-range couplings and a simple network model of chaotic elements. The simplest example is given by

$$x_{n+1}(i) = (1-\epsilon) f(x_n(i)) + \frac{\epsilon}{N} \sum_{j=1}^{N} f(x_n(j)), \tag{1}$$

where n is a discrete time step and i is the index of an element ($i = 1, 2, \ldots, N$ = system size)[#1]. We choose here the logistic map $f(x) = 1 - ax^2$ as a prototype for a system of globally coupled chaotic systems.

The model is a mean-field theory-type extension of coupled map lattices (CML), which have been proposed as simple models for spatiotemporal chaos and extensively investigated [1–15][#2].

A CML is a dynamical system with a discrete time, discrete space, and continuous state [1, 4–7]. It is a nonlinear model which consists of successive parallel procedures on a lattice. A typical

[1] On leave from: Institute of Physics, College of Arts and Sciences, University of Tokyo, Komaba, Meguro-ku, Tokyo 153, Japan.

[#1] This model is an extremely simple case of globally coupled maps and might be termed as "completely coupled".

[#2] See for the intermittency in CML refs. [1, 3, 4, 10].

example is the following diffusively coupled model:

$$x_{n+1}(i) = (1-\epsilon) f(x_n(i)) + \tfrac{1}{2}\epsilon\left[f(x_n(i+1)) + f(x_n(i-1))\right], \tag{2}$$

where n is a discrete time step and i is a lattice point ($i = 1, 2, \ldots, N$ = system size). The function $f(x)$ is a nonlinear mapping. The model has been investigated as a prototype for chaos in spatially extended systems, including extensions to a higher-dimensional lattice, a different choice of nonlinear function $f(x)$, and different types of couplings [1–12]. In CML, a parallel local nonlinear transformation and the parallel diffusion process are successively carried out.

The model (2) is the limit of short-ranged coupling model ("nearest neighbor"). Our model (1) is the other limit, of long-ranged coupling [15], and can be regarded as the mean-field version of (2).

Our dynamics (1) consists of a parallel nonlinear transformation and a feedback from the "mean field". Here the mean field obeys the relation $\Sigma_i x_{n+1}(i) = \Sigma_i f(x_n(i))$.

We note that the equivalent dynamics to (1) is obtained by the transformation $y_n(i) = f(x_n(i))$,

$$y_{n+1}(i) = f\left((1-\epsilon) y_n(i) + \frac{\epsilon}{N}\sum_{j=1}^{N} y_n(j)\right). \tag{3}$$

This form may be more familiar with researchers in neural networks, if one chooses a sigmoid function (e.g., $\tanh(\beta x)$) as $f(x)$ and the coupling term ϵ depending on elements.

The motivation of the present paper is as follows:

(I) In a short-ranged CML, we have seen a rich variety of pattern dynamics and phase transitions (see ref. [4]). Does our GCM give a kind of "mean-field theory" for the rich variety of phases in the pattern dynamics in CML [4] and give any better understanding of the transition?

(II) If the frozen random state in CML is related with a "glassy" phase [4], can the model (1) play a similar role for the state as that the "Sherrington–Kirkpatrick (SK) model" has played for spin glass (SG) [16]? Is there any significant difference between our frozen state and SG?

(III) Is there a way to "code" many attractors (some of which are chaotic and others of which are periodic)? Are there any bifurcation-like phenomena associated with the change of attractors?

(IV) Is it possible to construct a tree-like structure and attach any hierarchical code corresponding to the tree? Can we write down the hierarchical dynamics which describes the dynamics of the motion in different levels?

(V) Can we switch among attractors as we like through a simple input? Is it possible to "tune" the strength of chaos by this switch? If the switch works, is there any peculiar switching phenomenon with possible applications to information processing?

(VI) Is it possible to construct dynamics whose effective degrees of freedom are slowly changing, as are typically seen in the problem of the turbulence with a coherent structure and also in adaptive systems?

All of these questions are answered in the affirmative.

The importance of globally coupled chaotic systems is not restricted to dynamical systems. They are relevant to biological information processing, physical systems with nonlinear elements and a global coupling (e.g., fluid turbulence, Josephson junction circuit), ecological models, evolutionary models, and economics.

Recent extensive investigations on neural networks have given a simple example of applications of statistical mechanics to a simple model of information processing. The real neural dynamics in the brain consists of the dynamics of an ensemble of complex elements with complex coupling. Most of current neural network studies adopt oversimplified elements (0–1 or a sigmoid function) with a moderately simplified coupling. On the other hand, it is known that even a single neuron or a small

ensemble of neurons can exhibit complex dynamical behavior like chaos or frequency locking [17]. Then, it is natural and important to ask the following question: What happens if we use simplified elements (with chaotic response) with oversimplified couplings instead, as a different limit of simplification from the neural dynamics? Our model gives the first endeavor towards this direction. As we will see, the results are rather promising. We can have a huge number of coded attractors. Some attractors have tree-like structures with a hierarchical code and have similar complexity with the SK model in spin glass [16], which is the basis of one of the most popular models in neural networks [18–20]. Upon the tree structure one can construct hierarchical dynamics in which the motion is governed by the "slaving" from the higher-level cluster, strong interactions among clusters of the same level, and a slight feedback from lower-level clusters.

We also have to note that the tree structure in our model is attached to each attractor. Our model can have many attractors with this tree structure. Thus we may say that our model forms a "forest", rather than a tree.

Possible applications of our model include a regular posi–nega switch, a fuzzy hierarchical switch, and categorization as hierarchical clustering.

Other examples of globally coupled nonlinear elements include the dynamics of vortices in fluid dynamics. The motion of a vortex is governed by global coupling to all other vortices [21]. Coupled Josephson junction circuits [15] are another example from physics.

In population dynamics, global constraint from the "environment" is important. The dynamics of each species is described by the inherent dynamics of the species and the interaction with the environment. The dynamics of the environment, on the other hand, depends on all species.

In a similar context, Eigen and Schuster [22] have proposed an equation for evolutionary dynamics of the population of RNAs. It consists of a growth equation for each RNA and a constraint from the food source which is globally coupled to all the numbers of RNA species.

In economics (e.g., stock market), again, local complex dynamics and global feedback are important.

Of course, our model may be too simple to discuss these systems. Especially, the choice of the identical function $f(x)$ and the identical coupling for every element must be too simple to give a realistic model in the above fields. However, we may hope that some of the novel features presented here are useful in these systems, too. Up to now, we have very few examples of simulation with complex dynamics and global feedback. Thus it is important to investigate the simplest model with these features and to find novel notions and phenomena which our model can provide. We can then proceed to consider their applications and extensions to more realistic models for the above examples.

The organization of the present paper is as follows: In section 2, we classify attractors by the structure of clusterings and introduce a simple coding of attractors. Section 3 is devoted to the phase diagram of our model and the correspondence of our phases with those in a short-ranged CML [4]. Suppression of chaos in a few-cluster attractor is shown. In section 4, the change of dynamical nature with the clustering structure is discussed. In section 5, "precision-dependent clustering" is introduced to investigate the hierarchical structure of clusterings. A tree-like structure is constructed. Construction of hierarchical dynamics in our system is presented in section 6, where the "chaotic revolt" against hierarchy is found. We discuss the switch among coded attractors by simple successive inputs in section 7. We can jump among coded attractors as we choose (deterministic switch). The posi–nega switch with intermittency, stochastic switch, and hierarchical switch are also found. In section 8, we discuss the transient process before the system falls into attractors. Section 9 is devoted to discussions on the relevance of our results to problems of biological information processing.

2. Clustering and coding of attractors

Let us discuss the possible types of attractors in our model (1). The simplest attractor is a coherent one with $x(i) = x(j)$ for all i, j, in which case the motion is governed just by the single logistic map $x_{n+1} = f(x_n)$. The stability of this attractor is calculated by the Jacobi matrix J_n for the dynamics (1), i.e., the product of the matrices

$$J_n = f'(x_n)[(1-\epsilon)I + (\epsilon/N)D],$$

where the matrix I is a diagonal matrix with elements 1 ($\delta_{i,j}$), and D is a constant matrix whose elements are all 1. The stability is calculated by

$$\Gamma_j = \lim_{n\to\infty} \left(j\text{th eigenvalue of } \prod_{k=1}^{n} J_k \right)^{1/n} \quad (j = 1, \ldots, N). \tag{4}$$

After simple algebra, we obtain the eigenvalues $\Gamma_1 = \gamma$ and $\Gamma_j = \gamma(1-\epsilon)$ $(j = 2, \ldots, N)$ $((N-1)$-fold degeneracy). Here $\gamma = \exp(\lambda_0)$, where λ_0 is the Lyapunov exponent of the single logistic map. The eigenvector corresponding to the eigenvalue γ is given by $(1/\sqrt{N})(1, 1, \ldots, 1)^{\mathrm{T}}$. Thus, the amplification of a disturbance along this eigenvector *does not* destroy the coherence. Eigenvectors for the other $N-1$ eigenvalues are not uniform of course, and the amplification (which occurs if $|\Gamma_j| > 1$) along these vectors destroys the coherence. Thus the stability condition of the coherent attractor is given by $|\gamma(1-\epsilon)| < 1$, or in other words, by

$$\lambda_0 + \log(1-\epsilon) < 0. \tag{5}$$

Besides the above single-clustered coherent attractor, we have attractors with clusterings. A cluster is defined as the set of elements in which $x(i) = x(j)$ for i, j, that is, $x(i) = x(j)$ for i, $j \in$ the same cluster.

We can classify the state of our GCM by the number of clusters k and the number of elements for each cluster N_k: (see fig. 1 schematically). Unless otherwise mentioned, we use such labelling of clusters that $N_1 \geq N_2 \geq \cdots \geq N_k$.

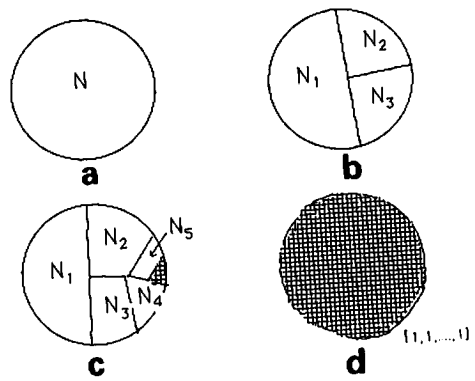


Fig. 1. Schematic figure for clusterings: (a) Coherent attractor. (b) Few clusters ($k = 3$). (c) Many-cluster attractor with unequal partition. (d) Many-cluster attractor with $k = N$.

If our system is attracted exactly to a k-cluster solution with $(N_1, N_2, \ldots, N_k)$, our system never goes out of the state with this clustering, since the motion for $x_n(i)$ and $x_n(j)$ at time $n > m$ are governed by exactly the same dynamics if $x_m(i) = x_m(j)$ holds ($x_n(i) = x_n(j)$ if $x_m(i) = x_m(j)$). Our attractor is therefore characterized by the clustering condition $[k, (N_1, N_2, \ldots, N_k)]$[#3].

Besides the coherent attractor (1), we have the following types of attractors (see fig. 1).

(2) Attractors with a small (much smaller than N) number of clusters.

(3) Attractors with a large number of clusters (of the order of N), with large N_1. An example of clustering is $[k = N/2 + 1, (N/2, 1, 1, \ldots, 1)]$.

(4) Attractors with a large number of clusters (of the order of or equal to N), and all N_j's are small (1 or 2). A typically observed case is $k = N$ and $N_j = 1$ for all j.

The distinction between (2) and (3) looks somewhat arbitrary, but we can make a clear distinction by calling attractors type (2) if the number of clusters does not increase with N. The distinction

[#3] Approximate (not exact) clusterings can change temporally, as will be discussed in section 6.

between (3) and (4) can be judged by the condition whether the maximum of N_i increases with N or not.

Some examples of the dynamical behavior of these attractors are shown in fig. 2.

If our system is fallen into a k-cluster attractor with $(N_1, N_2, \ldots, N_k)$, the dynamics is written by the following k-dimensional map:

$$X^{\nu}_{n+1} = (1-\epsilon) f(X^{\nu}_n) + \sum_{\mu=1}^{k} \epsilon_{\mu} f(X^{\mu}_n), \tag{6}$$

where X^{ν}_n denotes the value of x_n in the νth cluster, and the "effective coupling" ϵ_{μ} is given by $\epsilon_{\mu} = \epsilon(N_{\mu}/N)$.

The stability of the k-cluster state, however, is determined not only by the above k-dimensional map. To destroy the coherence within a cluster, amplification of a small disturbance $x(j) - x(i)$ must occur for the elements i, j belonging to the same cluster. This condition is again calculated by the products of Jacobi matrices for the N-dimensional map (1), as in eq. (4). If all the absolute values of eigenvalues of (4) are less than 1, the state is obviously stable (the attractor is periodic). If all the elements of an eigenvector corresponding to the eigenvalue whose absolute value is larger than unity take an identical value in each cluster (e.g., $v(i) = 1$ for $i \in$ cluster A, and $v(j) = 0$ for $j \notin$ A), the amplification of a disturbance along this eigenvector does not destroy the coherence within each cluster. Since the number of independent vectors with the above condition is less or equal to k, we get the following proposition:

Proposition. The number of positive Lyapunov exponents for a k-cluster attractor in our system (1) cannot exceed k.

2.1. *Coding versus partition*

Since a fixed set of numbers $[k, (N_1, N_2, \ldots, N_k)]$ is attached to each attractor, we can code each attractor by this set of numbers. If the k-dimensional map (6) has a unique attractor[#4] which is stable in the criterion of the N-dimensional Jacobi matrix (4), this coding is complete. A unique attractor is assigned by the above code.

In principle, the k-dimensional map (3) can have more than one attractor. It can happen that more than one of these attractors are stable in the original N-dimensional dynamical systems (1). If this is the case, the coding by $(N_1, N_2, \ldots, N_k)$ is imperfect and we can have more than one attractor corresponding to this clustering. Although this case is rather rare, we have encountered a few examples for $k = 3$ numerically.

In a 2-cluster case, the coding is very simple. Attractors are coded just by a single number N_1, since $N_2 = N - N_1$. Similarly a 3-cluster attractor is coded just by N_1 and N_2.

If we distinguish each element i, there are $N!/(N_1!N_2!\ldots N_k!)$ ways of partitions with the clustering condition $(N_1, N_2, \ldots, N_k)$. Thus we have exponentially many attractors for each clustering condition. Even in the simplest case of 2-clusters, we have $N!/(N_1!(N-N_1)!)$ attractors for each N_1 condition.

If we use our system for the storage of information, and try to store the information in each attractor, we can store $N!/(N_1!N_2!\ldots N_k!)$ patterns for the clustering condition $(N_1, N_2, \ldots, N_k)$. We can have large capacity of information storage. To quantify the above capacity we introduce the clustering entropy defined by

$$S = -\sum_{j=1}^{k} p(j) \log p(j) \tag{7}$$

with

$$p(k) = N_k/N. \tag{8}$$

[#4] We exclude the attractor with $X^{\mu} = X^{\nu}$ for some pair of μ and ν, since the equality leads to clustering with a smaller number. Our system is described by a lower-dimensional map.

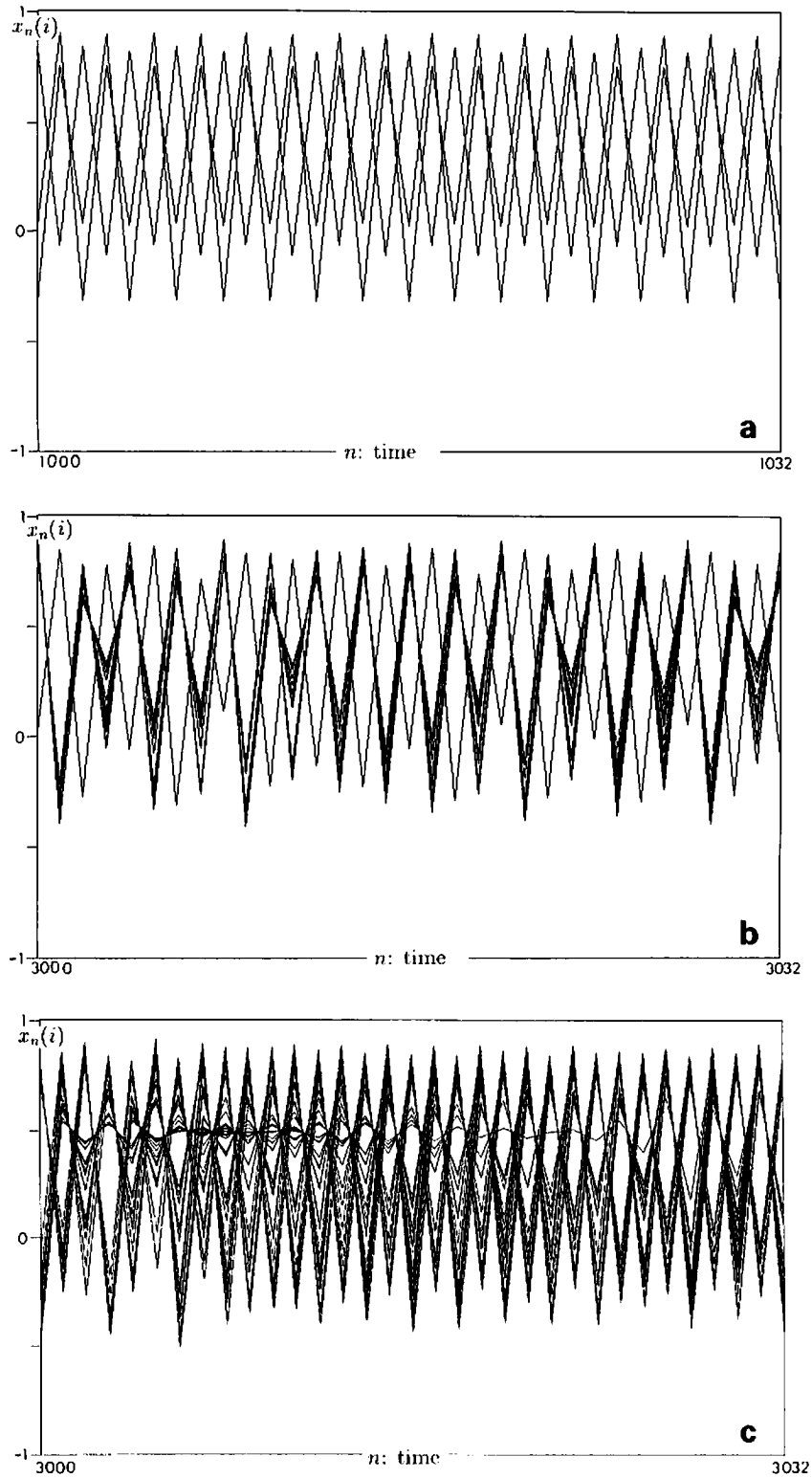


Fig. 2. Examples of time series of attractors: $x_n(i)$ for all i plotted as a function of time n. If there are only k lines ($k < N$), the system has fallen to k-clusters. $N = 50$, (a) 3-cluster attractor: $a = 1.8$, $\epsilon = 0.2$; $N_1 = 18$, $N_2 = 18$, $N_3 = 14$ ($1000 \le n \le 1032$). (b) Attractor with $k = 27$: $a = 1.95$, $\epsilon = 0.2$; $N_1 = 24$ and $N_j = 1$ ($27 \ge j \ge 2$ ($1000 \le n \le 1032$). (c) N-cluster attractor ($N_j = 1$, for all j): $a = 2.0$, $\epsilon = 0.2$ ($1000 \le n \le 1032$).

As the distribution is closer to equal partition (i.e., $N_i \approx N_j$), the possible combinatorial number $N!/(N_1!, N_2!, \ldots, N_k!)$ gets larger and the clustering entropy takes a larger value. If many elements are concentrated on one cluster (e.g., $N_1 \gg N_j$), the combinatorics and the entropy take a small value even if k is large.

3. Phases and basin distribution

We take an ensemble of initial conditions, to see a variety of attractors and the basin volume for each attractor. By sampling many initial conditions, we can calculate the cluster distribution function $Q(k)$, which is defined as the basin volume ratio for a k-cluster attractor. Numerical calculation is carried out in the following way: Take randomly chosen, many (M) initial conditions. By counting the number of initial conditions which lead to a k-cluster attractor, and dividing it by M, we obtain $Q(k)$. $Q(k)$ is an analogue of the pattern distribution function studied in CMLs [4].

According to the numerical results we can classify the phases as follows:

(i) *Coherent phase*: coherent attractors have occupied (almost) all basin volumes.

(ii) *Ordered phase*: few-cluster attractors have occupied (almost) all basin volumes.

(iii) *Partially ordered phase*: coexistence of many-cluster and few-cluster attractors.

(iv) *Turbulent phase*: all attractors have many ($\approx N$) clusters.

In the coherent phase, almost all attractors are coherent, i.e., $x(i) = x(j)$ for all i, j. $Q(1) = 1$ in this phase.

In the ordered phase, only a few-cluster attractor has a basin volume. More precisely, $Q(k) = 0$ for $k > k_c$ with k_c independent of N. For numerical simplicity, we can roughly check this by the condition $Q_L \equiv \Sigma_{k > N/2} Q(k) = 0$[#5] and $Q(1) \neq 1$. In most cases numerically observed so far, only few $Q(k)$'s take nonvanishing values.

In the partially ordered phase, both few-cluster attractors and many-cluster attractors coexist. Although $Q(k)$ for small k takes a nonzero value, there is no upper bound for the number of clusters. Numerically it is distinguished by the condition $\Sigma_{k > N/2} Q(k) > 0$ and $\Sigma_{k < N/2} Q(k) > 0$. We call the partially ordered phase as "intermittent phase" when it is adjacent to the turbulent phase, while it is called as "glassy phase" if the phase is between two ordered (or coherent) phases in the phase diagram. Reasons for this nomenclature are that (a) the dynamics in the "intermittent phase" gives typical intermittent behavior in spacetime [1, 4], and (b) in the glassy phase the competition of some attractors with different cluster size leads to the frustration (note that the randomness and frustration are two basic ingredients of glassy behavior).

In the turbulent phase, most attractors have many clusters, i.e., $\mathcal{O}(N)$. Here it is distinguished numerically by the condition $\Sigma_{k > N/2} Q(k) = 1$.

A simple quantification of the attractors is the use of average cluster number $R \equiv \Sigma_{k=1}^{N} k Q(k)$. In the coherent phase $R = 1$, while $R = b \ll N$ for the ordered phase, and $R = rN$ for partially ordered ($r < 1$) and turbulent phases ($r = 1$).

A rough phase diagram is given in fig. 3. It is obtained through the simulation of our model for the parameters $a = 1.4, 1.41, \ldots, 2.0$ and $\epsilon = 0.02, 0.04, 0.06, \ldots, 0.4$. The phases are determined from $Q(k)$ which is calculated through $M = 500$ samples, after the transients of 2000 steps, and $N = 200$. In the diagram, "(1, 2)" means the ordered phase of dominant cluster size 1 or 2, and "(1, 2, 3)" for the dominant cluster sizes 1, 2, 3, etc. Here we call the cluster size dominant if $Q(k) > 0.1$. Again this criterion is not important, since the dominant cluster occupies almost all basin volumes, except in the vicinity of phase boundaries. Actually in the region marked as "(2)", $Q(2) \approx 1$, and in "(2, 3)", $Q(2) + Q(3) \approx 1$, etc.

$Q(k)$ for $k < 5$ and Q_L as a function of a are plotted in figs. 4, 7, 8, and 10 for $\epsilon = 0.1$, 0.2, 0.3, and 0.4. Let us survey the change in each ϵ.

[#5] This criterion for many-cluster is flexible. One can choose a different criterion, like $\Sigma_{k > N/4} Q(k)$. Our phase diagram essentially is independent of the criterion, although minor changes of boundary between phases are possible.

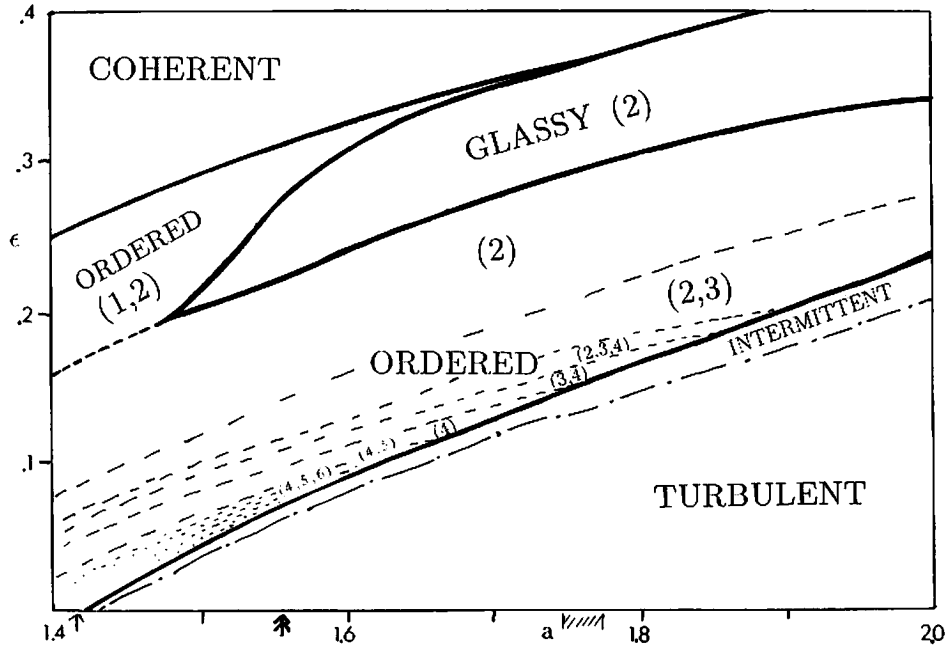


Fig. 3. Rough phase diagram: Phases are determined by $Q(k)$, calculated from 500 randomly chosen initial conditions and $N = 200$. The parameters are changed from $a = 1.4$ to 2.0 by 0.01 and $\epsilon = 0.02$ to 0.4 by 0.02. Numbers such as (1, 2, 3) represent dominant cluster numbers (with basin volume ratio more than 10%). The single arrow at the bottom line shows accumulation of period-doubling bifurcations, while the double arrow denotes the band merging point for the single logistic map. The region with oblique lines correspond to the period-3 window in the logistic map.

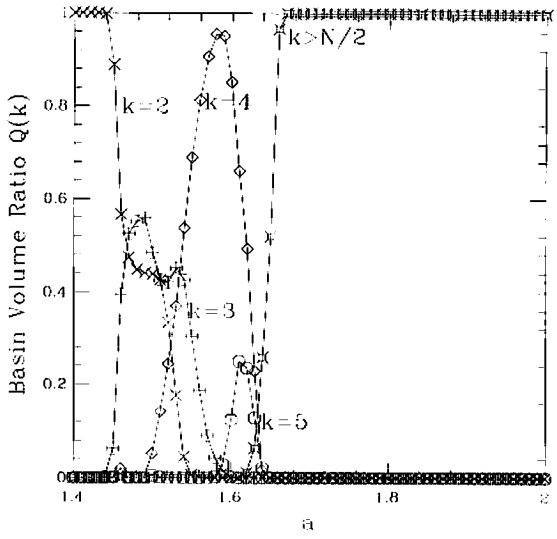


Fig. 4. Cluster distribution $Q(k)$ for $\epsilon = 0.1$ $(1 < k \leq 5)$ and Q_t plotted as a function of a for $1.4 < a < 2.0$. $\epsilon = 0.1$, $N = 200$. $Q(1) = 0$ for these parameters. Calculated from 1000 randomly chosen initial conditions, after 3000 transients.

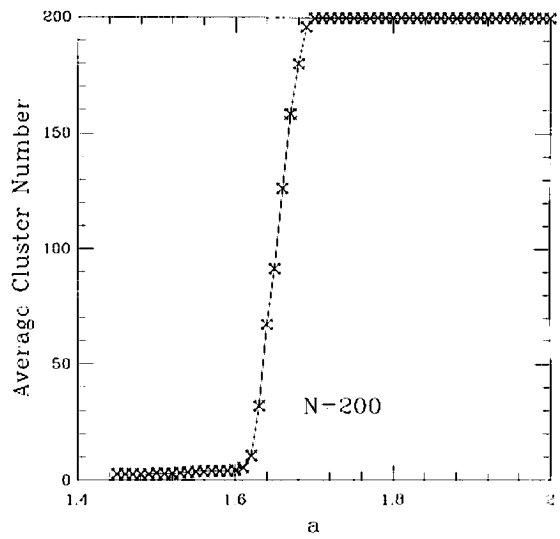


Fig. 5. Average cluster number as a function of a. Calculated as in fig. 4. $\epsilon = 0.1$.

(i) $\epsilon = 0.1$

The coherent attractor has a too small basin volume to be observed even in the parameter region where $\lambda_0 + \log(1-\epsilon) < 0$. For $a < 1.62$, the ordered phase appears. We note that the dominant cluster size increases from $k = 2$ to $k = 3$ and then to $k = 4$. The second-order type transition is clearly seen in fig. 5: The average cluster number increases for $a > 1.61$, and reaches N at $a \approx 1.69$. Examples of the dynamics of 4-cluster and 6-cluster attractors are shown in fig. 6. Note the band splitting and periodic motion of these attractors.

(ii) $\epsilon = 0.2$

Around $1.48 < a < 1.52$, a glassy state appears. Except this small parameter regime, we have the ordered phase with a dominant 2-cluster attractor for $a < 1.76$. Especially for $1.56 \leq a \leq 1.68$, the 2-cluster attractor has occupied all the basin volumes within our calculation. The motion there is quasiperiodic if $N_1 \approx N_2$. As $N_1 - N_2$ is increased

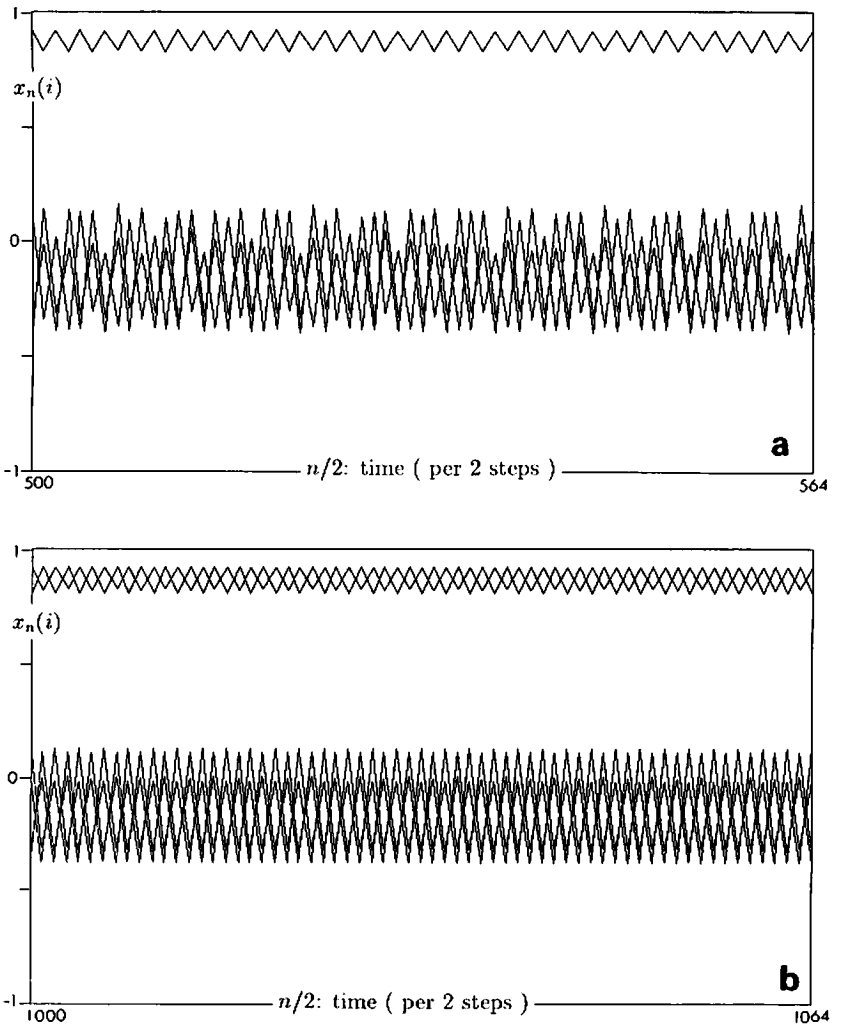


Fig. 6. Some examples of attractors for $\epsilon = 0.1$, $N = 50$. Time series of $x_{2n}(i)$ $(n = 500, 501, \ldots, 564)$ plotted for (a) $a = 1.58$, $(k = 4, [19, 15, 15, 1])$; (b) $a = 1.60$, $(k = 6, [12, 9, 8, 8, 7, 6])$.

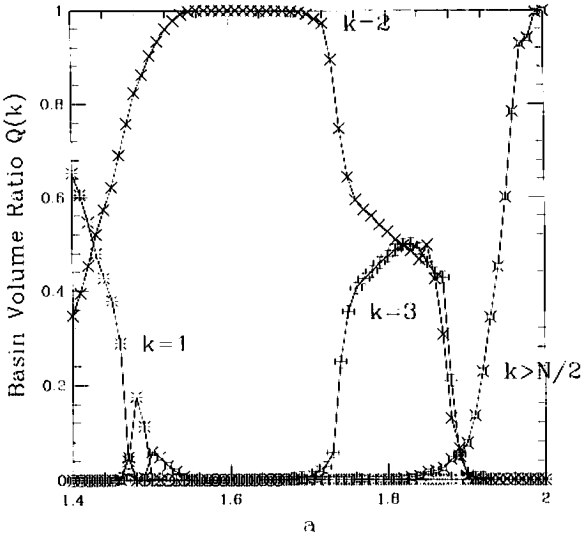


Fig. 7. $Q(k)$ for $\epsilon = 0.2$ plotted as a function of a, calculated in the same manner as in fig. 4.

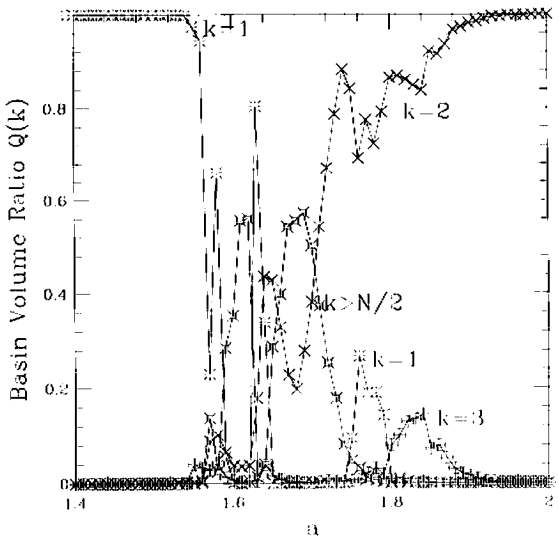


Fig. 8. $Q(k)$ for $\epsilon = 0.3$ plotted as a function of a, calculated in the same manner as in fig. 4.

the motion changes to period-2, and then exhibits the period-doubling to chaos (see section 4). For $1.68 < a < 1.86$, the basin volume is divided into 2-cluster and 3-cluster attractors. Chaos is strongly suppressed in these parameter regimes. For larger a, the basin for many-($\mathcal{O}(N)$-) cluster attractors increases till it occupies almost all volumes at $a \geq 1.99$.

(iii) $\epsilon = 0.3$

For $a \leq 1.54$, we have found only a coherent attractor. For $1.56 \leq a \leq 1.80$, our system is in the glassy phase, in which few-cluster and many-cluster attractors coexist. The maximal number of clusters can be N, while the minimum is 1 or 2. For $a > 1.7$, the basin volume for 2-cluster attractors starts to occupy a large ratio. The ratio increases with a, and for $a > 1.88$, the system has fallen into the ordered phase in which the 2-cluster attractor occupies almost all basin volumes. The motion of 2-cluster attractor is much more regular than the single logistic map.

(iv) $\epsilon = 0.4$

The coherent attractor occupies all the basin volumes for $a < 1.84$. The basins both for many-cluster and 2-cluster attractors increase with a, for $a > 1.84$. We have not seen the ordered phase with $k \geq 2$.

For $\epsilon < 0.25$, the coherent attractor has a very small basin volume even in the periodic window regime of the logistic map, where the existence of such attractor is assured by eq. (5). For example, $Q(1) \approx 0$ even in the parameter regime for the period-3 window in the logistic map ($a \approx 1.75$)[#6].

In all these examples, the motion of a 2-cluster attractor is period-2 band (chaotic/periodic) (see fig. 11). The two clusters oscillate out of phase with each other: In one cluster $x(i)$ changes as $+-+-\ldots$, while the other as $-+-+\ldots$, where $+, -$ is distinguished by whether $x(i) > x^*$ or not, with x^* as the unstable fixed point of logistic map $(\sqrt{1+4a}-1)/2a$ (this distinction is also used in a short-range CML [4]). If $N_1 = N_2$, the motion of two clusters is symmetric. As $N_1 - N_2$ increases, the asymmetry between the motion of two clusters grows (see fig. 11), accompanied by the bifurcation to chaos to be discussed in section 4.

[#6] At $\epsilon = 0.3$, we can see the increase of $Q(1)$ in the period-3 window of the logistic map (see fig. 8).

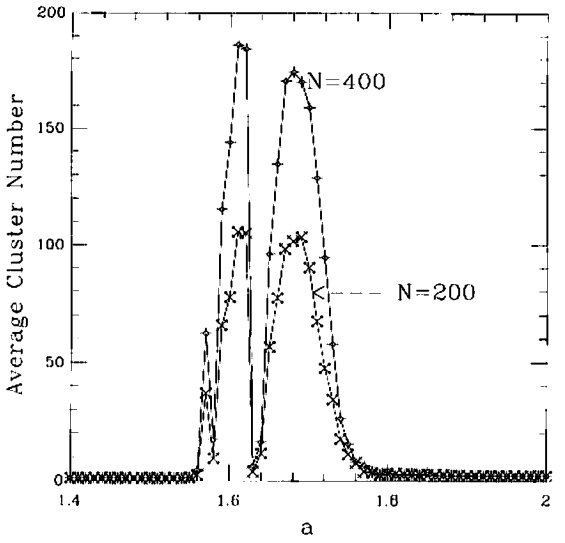


Fig. 9. Average cluster number as a function of a. $\epsilon = 0.3$, for $N = 200$ and $N = 400$; calculated in the same way as in fig. 4.

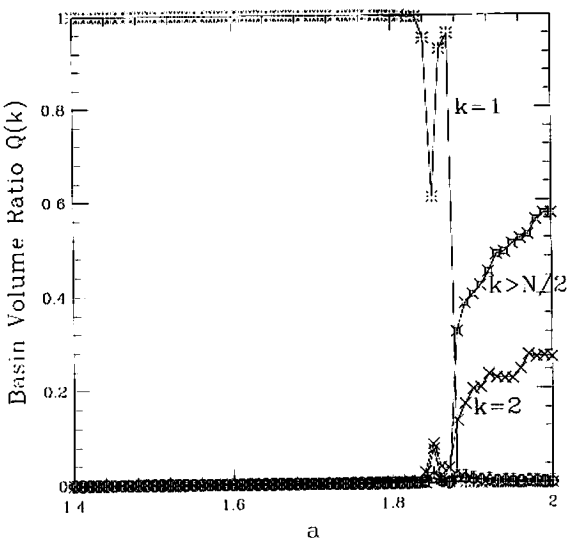


Fig. 10. $Q(k)$ for $\epsilon = 0.4$ plotted as a function of a, calculated in the same manner as in fig. 4.

In a 3-cluster attractor, the system splits into 2 "mega-clusters" which oscillate out of phase just like 2-cluster attractors. One of the two mega-clusters splits into two subclusters which oscillate in-phase in the period-2 band, but out of phase in the period-4 (see fig. 12). 4-cluster attractors and many-cluster attractors are formed in a similar way (see fig. 6). This leads to the hierarchical clustering to be investigated in section 5.

3.1. *Correspondence of our phases with those in short-ranged CML*

Since our GCM is a mean-field version of a short-ranged CML, it is important to compare our

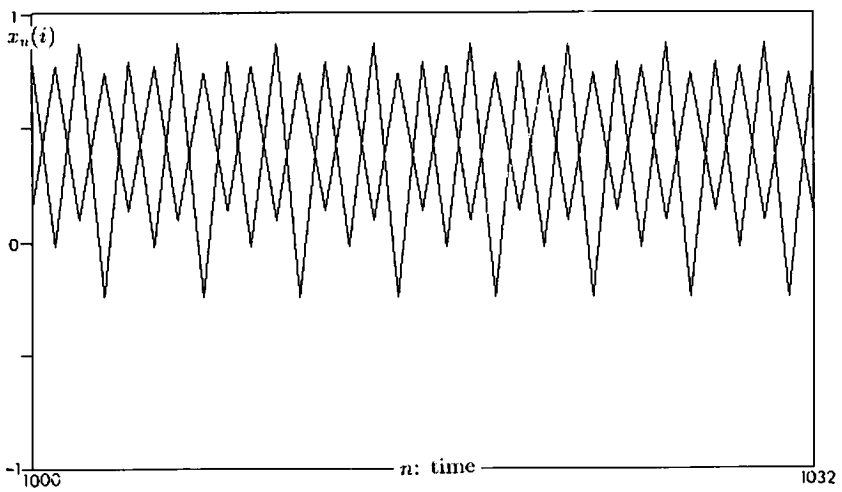


Fig. 11. Time series of 2-cluster attractors for $a = 1.88$, $\epsilon = 0.3$, and $N = 100$. $x_n(i)$ is plotted as a function of time n ($n = 1000, 1001, \ldots, 1032$). ($k = 2, [59, 41]$).

phases with those in the short-ranged CML, presented in ref. [4].

(A) *Ordered phase (GCM) versus pattern selection (CML)*

First, few domain sizes (in CML) or few cluster numbers (in GCM) are selected.

Second, chaos is suppressed in both cases. Actually, the motion of a few-cluster attractor is more regular than that of the single logistic map for the same parameter. The 2-cluster attractors, for example, are period-2 or in a period-2 band with quasiperiodic or weakly chaotic modulation. On the other hand, suppression of chaos in the pattern selection in a short-ranged CML has been investigated in ref. [4].

The reason for the suppression of chaos in a few-cluster attractor is seen in the proposition about positive Lyapunov exponents in section 2. If the chaos were too strong in a k-cluster state, there would be more than k positive exponents. Then the state should be unstable against a higher-dimensional perturbation. The state could not be confined within the k-dimensional dynamics (6). Thus strong chaos is inhibited in a few-cluster attractor. (For numerical results on the Lyapunov spectra see section 4.)

(B) *Glassy phase (GCM) versus frozen random pattern (CML)*

First, both phases have a huge number of possible attractors; second, the size of clusters (or domains) can take almost arbitrary values, depending on the initial conditions; third, the motion there splits into some bands and the temporal power spectrum of $x_n(i)$ has δ peaks together with the continuous parts. Fourth, the strength of chaos is enhanced as the domain size (CML) or cluster size (GCM) is increased.

(C) *Intermittent transition*

The transition to the turbulent state is common between GCM and CML. The motion at the transition regime is temporally intermittent. In GCM, there is a critical behavior as

$$Q_L = (a - a_c)^{\beta}. \tag{9}$$

Corresponding critical behavior of disordered parameter is seen in CML as $1 - \Sigma_p Q(p) \propto (a - a_c)^{\beta'}$, where the sum runs over all possible selected domain sizes p [4]. Detailed study on the exponents will be left for future.

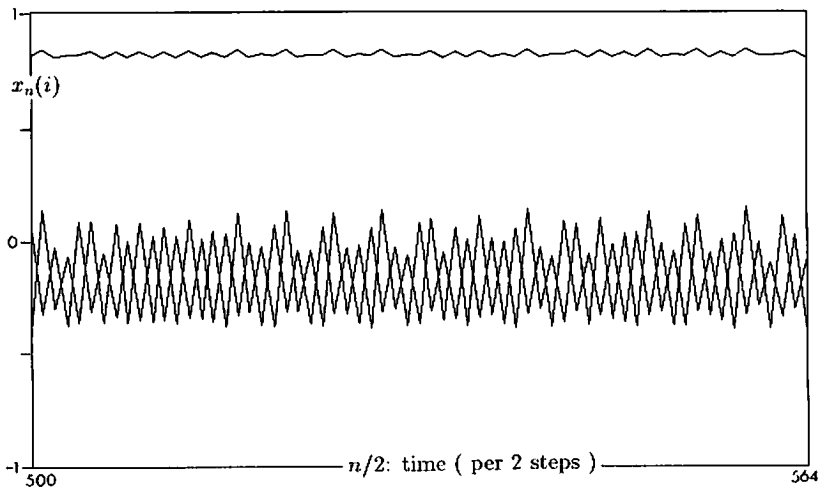


Fig. 12. Time series of 3-cluster attractors for $a = 1.75$, $\epsilon = 0.2$, and $N = 50$. $x_{2n}(i)$ is plotted as a function of time n ($n = 500, 501, \ldots, 564$). ($k = 3$,[23, 14, 13]).

(D) *Turbulent phase (GCM) versus fully developed spatiotemporal chaos (CML)*

In both phases, there is a small correlation among elements. In CML, the spatial correlation decays exponentially, while there is almost no correlation among elements in GCM. The motion in both phases gives fully developed chaos without any δ peaks in the temporal power spectrum. Both motions are well described by the random process.

(E) *Coherent phase versus absence of frozen patterns*

In GCM, a coherent phase is dominant for large ϵ. No ordered phase is observed. Corresponding to it, there is neither frozen random pattern nor pattern selection in a higher dimensional CML with a large coupling [4].

4. Cluster bifurcation

How does the dynamical nature of each attractor change with $[k,(N_1, N_2,\ldots, N_k)]$? This problem is directly related with the bifurcation, since the change of N_j leads to the change of effective coupling parameters ϵ_j in the k-dimensional map (6).

As the simplest case, we start with 2-cluster attractors. If we confine ourselves to 2-cluster solutions, the dynamics (1) is written as the two-dimensional coupled map [23]

$$X^1_{n+1} = (1-\epsilon_2) f(X^1_n) + \epsilon_2 f(X^2_n),$$
$$X^2_{n+1} = (1-\epsilon_1) f(X^2_n) + \epsilon_1 f(X^1_n) \qquad (10)$$

with X^1_n, X^2_n as $x_n(i)$ for each cluster, and $\epsilon_{1,2} = \epsilon(N_{1,2}/N)$[#7].

Thus the change of N_1 corresponds to the change of bifurcation parameter in the two-dimensional map. We have to note, however, that the above reduction is possible only after the system has fallen onto 2-clusters.

[#7]Here we adopt a slight different labelling of clusters (N_1, N_2). We call the cluster "1" if the site 1 belongs to it. Thus $N_1 \geq N_2$ is not necessarily satisfied.

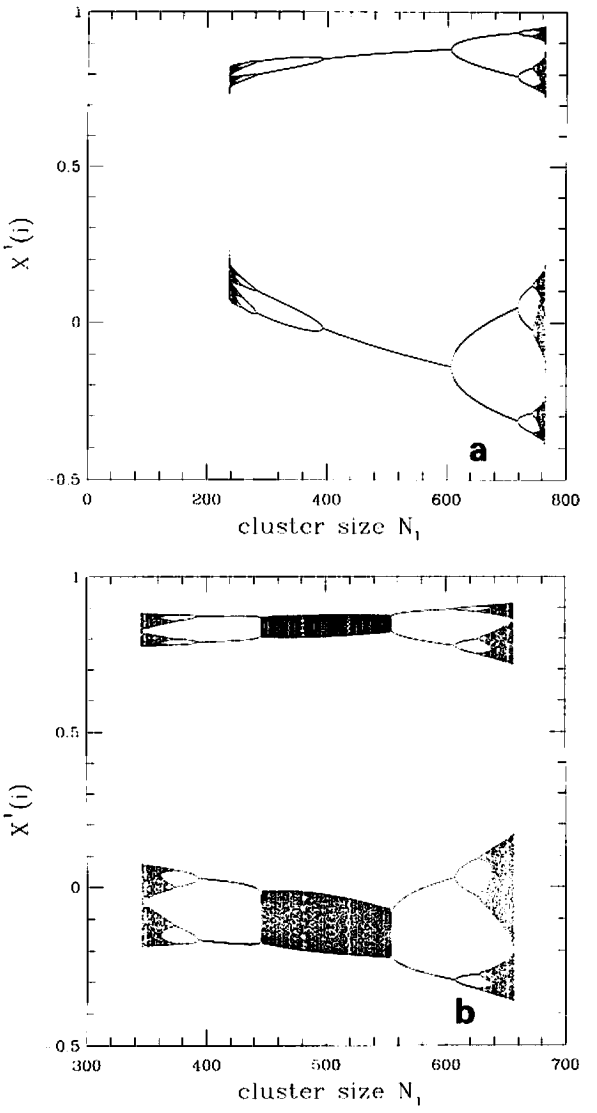


Fig. 13. Cluster bifurcation for 2-cluster attractors: $X^1_n = x_n(1)$ $(n = 2000, 2001, \ldots, 2260)$ is plotted as a function of N_1. $N = 1000$. $\epsilon = 0.2$. To change the attractors, the switching method in section 7 is used.

In fig. 13, two examples of cluster bifurcation of 2-cluster attractors are shown, where $X^1_n = x_n(1)$ $(n = 2000, 2001, \ldots, 2260)$ are plotted as a function of N_1.

The above 2-cluster attractors exist for $N - N_{\text{thr}} \leq N_1 \leq N_{\text{thr}}$ $(N_{\text{thr}} \geq N/2)$. The threshold N_{thr} is numerically obtained. It is proportional to N

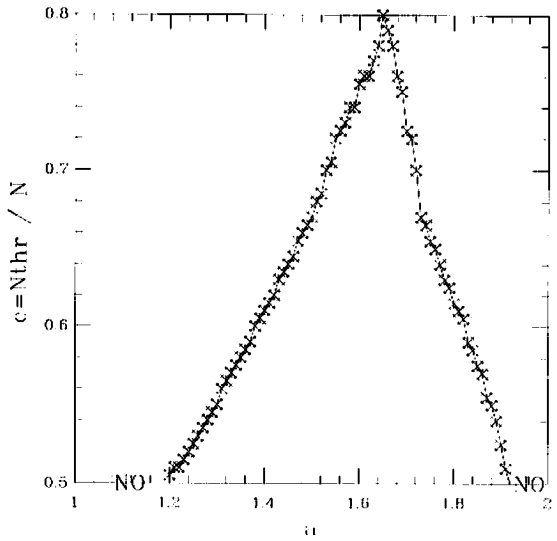


Fig. 14. $c = N_{thr}/N$ as a function of a (for $\epsilon = 0.2$). Calculated from a system with $N = 200$, and with the use of switching in section 7. c goes to 0.5 at $a = 1.2$ and $a = 1.92$. Below 1.2 and beyond 1.92, any 2-cluster state is no longer an attractor.

($N_{thr} = cN$). The coefficient c depends on a as in fig. 14. It increases with a and then decreases down to 0.5, where the 2-cluster solution loses its stability.

So far, we have observed two types of bifurcations in the cluster-size change.

One is the period-doubling bifurcation, which is typical at $\epsilon = 0.3$. In fig. 13a, X_n^1 clearly exhibits the period-doubling "bifurcation" cascade to chaos as N_1 is changed. As is stated, this is not a bifurcation in the usual sense. The parameter is fixed here and all what we have done is arrange attractors in the order of N_1. In other words, we have found a simple way to organize many attractors, through which the change of attractors can be seen just as in a bifurcation diagram.

The other is a quasiperiodicity–chaos transition with period-doubling bifurcation, which is observed for smaller ϵ (e.g., 0.2). For an example, see fig. 13b. The motion is quasiperiodic[#8] with some lockings if $N_1 \approx N_2$. As $N_1 - N_2$ is increased, the attractors exhibit period-doubling to chaos from period-2. After some band mergings, the 2-cluster attractors become unstable (N_1 hits N_{thr}).

In 3-cluster attractors, we can see a co-dimension-2 bifurcation with the change of N_1 and N_2. Again, we have observed quasiperiodicity, lockings, and period-doubling to chaos. There is a threshold on N_1 and N_2, beyond which 3-cluster attractors are no longer stable. Near the edge of threshold, the motion is chaotic. As N_1 and N_2 approach the threshold, the number of positive Lyapunov exponents increases from 0 to 2, and their magnitude grows. We can tune the dynamical state of our system, through changes of cluster sizes.

In a similar manner, we can see the bifurcation in a k-dimensional space for k-cluster attractors. If we take all the possible attractors of different cluster numbers, we can have bifurcations *not only in the parameter space with a fixed number of parameters, but with a variable number of parameters and dimension*.

In all the examples of few-cluster attractors, the motion is periodic or quasiperiodic if the clustering is close to the equal partition ($N_i \approx N_j$). The motion becomes chaotic, as the differences among N_j's are increased and the system approaches the edge of the stability of this cluster solution.

To see the change of strength of chaos by attractors quantitatively, Lyapunov spectra are calculated from the products of Jacobi matrices.

In fig. 15, we have plotted the Lyapunov spectra for attractors with different numbers of clusters. As a simple way to see the change of strength of chaos, (i) Kolmogorov–Sinai (KS) entropy calculated by the sum of positive Lyapunov exponents, and (ii) maximal Lyapunov exponents are plotted with the change of number of clusters (see fig. 16). In the example, the number of positive Lyapunov exponents is about $k - 3 \sim k$ for attractors with large k (>30). It roughly increases linearly with k. The maximal Lyapunov exponent increases slowly with k, and KS entropy increases faster than it with k.

[#8]The quasiperiodicity is checked by the calculation of Lyapunov exponents. In fact, the maximal exponent is zero within the numerical error for $0.45 \lesssim N_1/N \lesssim 0.55$ in the example of fig. 13b.

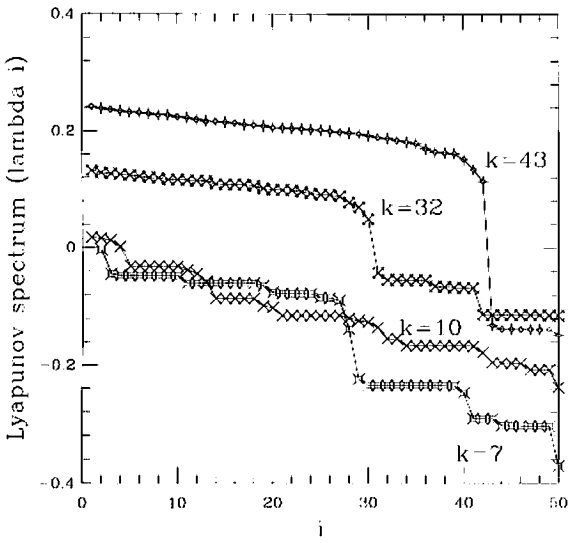


Fig. 15. Examples of Lyapunov spectra for $a = 1.65$, $\epsilon = 0.1$, and $N = 50$. Calculated from 2000 steps after discarding 5000 transients. Four examples with different initial conditions are shown which lead to different clusterings: $(k = 43, [8, 1, \ldots, 1])$, $(k = 32, [11, 9, 1, \ldots, 1])$, $(k = 10, [11, 10, 8, 7, 7, 3, 1, 1, 1, 1])$, and $(k = 7, [13, 10, 9, 8, 6, 3, 1])$.

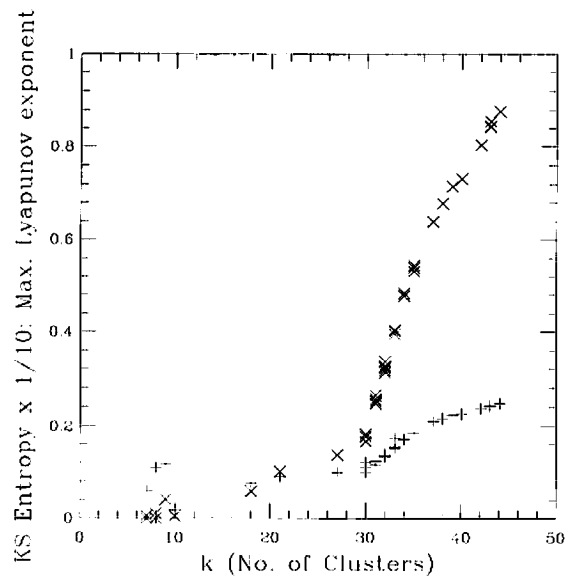


Fig. 16. KS entropy (×) and maximal Lyapunov exponent (+) are plotted as a function of the number of clusters k. Calculated in the same way and with the same parameters as in fig. 15. Twenty-five randomly chosen initial conditions.

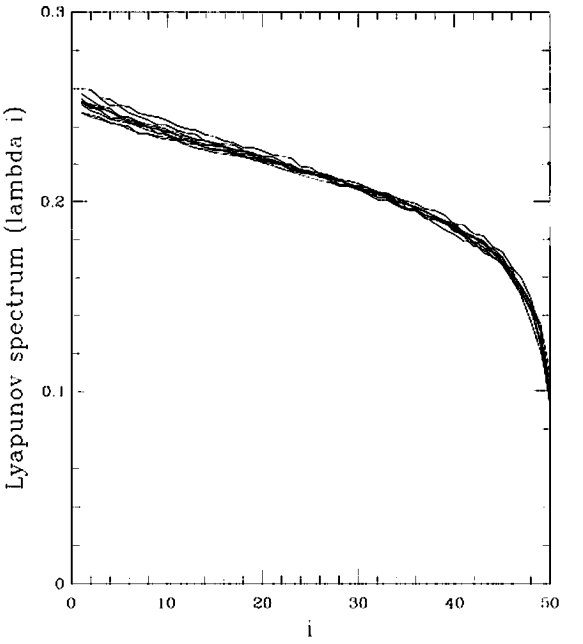


Fig. 17. Lyapunov spectra for the turbulent attractor for $a = 1.75$, $\epsilon = 0.1$, and $N = 50$. Calculated in the same way as in fig. 15, taking randomly chosen 10 initial conditions, all of which lead to the attractor with $k = N$.

These results clearly show the suppression of chaos in a smaller number of clusters. We note that not only the number of chaotic degrees of freedom, but also the strength of instability in each cluster (e.g., maximal Lyapunov exponent) increases with the number of clusters.

In the turbulent regime, the spectra have smooth forms. We have sampled many initial conditions, but all of them give the same spectra within our precision. This suggests that the attractor is unique in our turbulent regime.

5. Precision-dependent clustering and hierarchical code

In the examples of figs. 6, the distance between values of some clusters are much closer than the distance between those of other clusters. We have also observed an attractor with one large cluster and many small clusters (e.g., $N_j = (N/2, 1, 1, \ldots, 1)$ in the partially ordered phase. If we

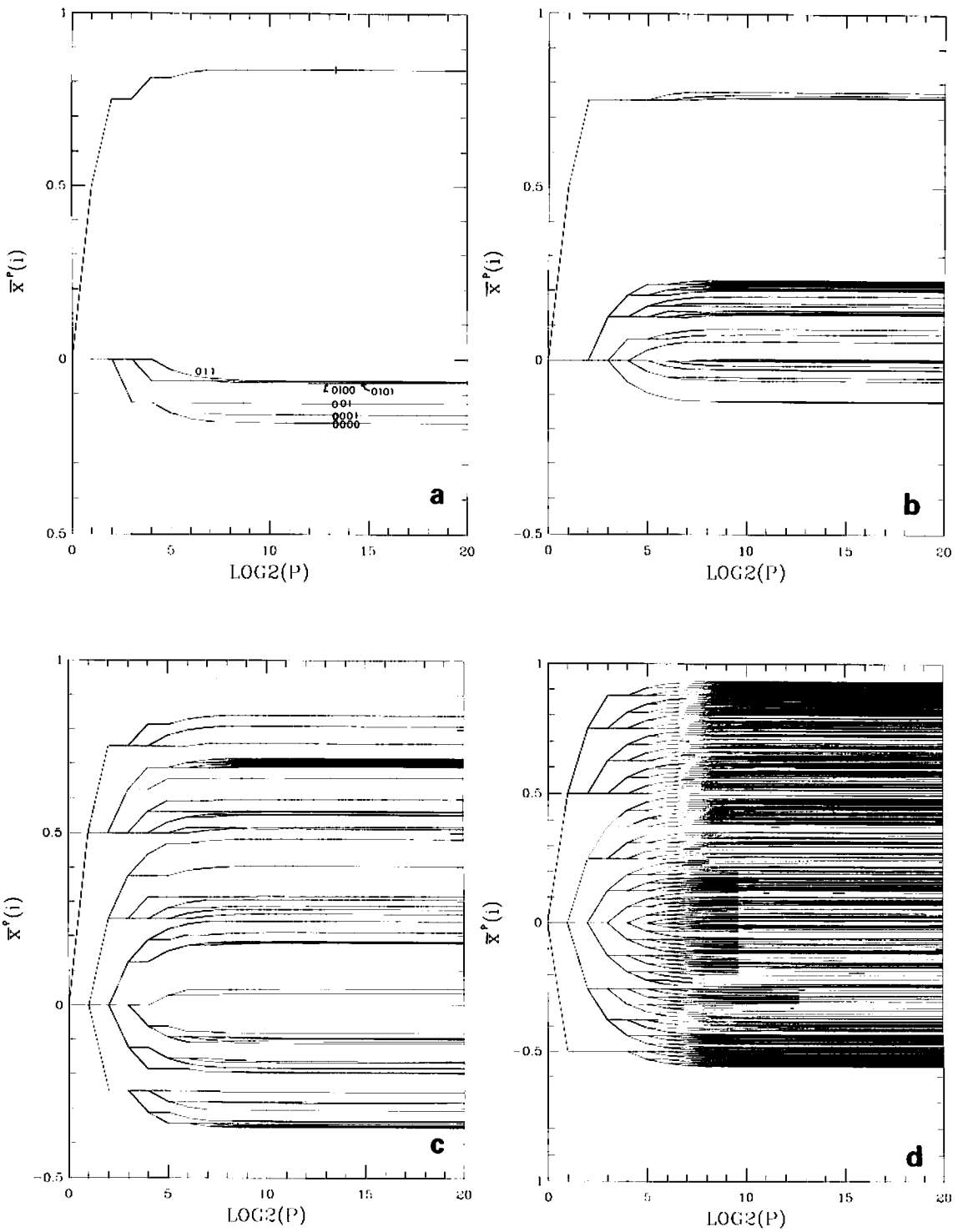


Fig. 18. Precision-dependent tree: $\bar{x}_n^P(i)$ plotted for all i, at $n = 50000$, with the change of the precision P as $P = 2^m$ ($m = 1, 2, \ldots, 30$). $N = 1000$. (a)–(c): $a = 1.92$, $\epsilon = 0.2$ from three randomly chosen initial conditions. (d) $a = 1.90$, $\epsilon = 0.1$. (We have checked other initial conditions, which lead to essentially the same tree structure.)

look at these clusters closely, we notice that the distance between two small clusters $|X^{\mu} - X^{\nu}|$ $(\mu, \nu > 1)$ is much smaller than the distance between the large cluster and a small one $|X^{1} - X^{\nu}|$. To discuss this kind of structure, we need the notion of "distance" between clusters.

For this purpose, we introduce the coarse-grained measurement of $x_n(i)$ and then the notion of *precision-dependent clustering*.

Here a coarse-grained measurement is defined as

$$\bar{x}_n^P(i) \equiv [Px_n(i)]/P, \tag{11}$$

where $[\ldots]$ is the integer part of... and P is an integer which gives the precision. The above coarse-graining gives digital values with m/P $(m = 0, \pm 1, \pm 2, \ldots)$. The precision-dependent clustering is defined by the clustering for $\bar{x}_n^P(i)$. If $\bar{x}_n^P(i) = \bar{x}_n^P(j)$ holds, the elements i and j are regarded as belonging to the same cluster within that precision P. The precision-dependent cluster number k_n^P is defined as the number of clusters at time n with precision P.

As the precision is increased, the clusters split and their number k^P increases until it goes to the exact cluster number k. To see this process, we introduce the *precision-dependent tree*. In fig. 18, $\bar{x}_n^P(i)$ is plotted with the change of precision P, for $i = 1, \ldots, N$, for fixed n. The number of lines at a given P gives the number of clusters at the precision. We can see the tree-like organization.

Four tree structures of different attractors are shown in figs. 18a–18c for $a = 1.92$, and $\epsilon = 0.2$. Note the variety of tree structures.

The tree structure is close to be symmetric in the turbulent region (see fig. 18d), while it is strongly asymmetric in a glassy or intermittent phase. A typical asymmetric tree consists of one big branch without any more smaller branchings and many branches with small bifurcations (see fig. 18b). In most attractors of this type, the number of elements N_1 in one big branch is comparable to the sum of all numbers of elements in other branches.

Based on this tree structure, we can have a binary hierarchical code for the clustering, as has been used in various fields. Each cluster is coded by a sequence of binary numbers $[0, 1], [0, 1], \ldots$ as in fig. 18a[#9]. In some clusters this code has a short length, while some other clusters have many branchings. The length also depends on attractors. We often use this binary code instead of the code in section 2, to discuss the hierarchical nature of clustering.

In the partially ordered phase, we can have a lot of attractors of different tree structures. This kind of tree has already been discussed in the spin glass system (SK model) [16]. Comparison of our glassy state with the SK model is summarized in table 1.

To understand the whole structure of our system with an ensemble of trees, we have to study the ensemble of trees ("forest"). To understand the forest, it may be relevant to extend some notions developed in the study of SK-model, like the overlapping among trees and replica methods [16]. This study will be left for future.

5.1. *Storage versus precision*

The clustering entropy in section 2 is easily extended to the precision-dependent clustering entropy as

$$S^P \equiv -\sum_{j=1}^{k^P} \frac{N_j^P}{N} \log\left(\frac{N_j^P}{N}\right), \tag{12}$$

where N_j^P is the number of elements in the jth cluster of the precision P. As P is increased, the entropy S^P is increased till it takes the value of clustering entropy S. S^P gives the capacity of information storage in our system with precision P.

[#9]The binary code here is not compressed. In fig. 18a, $k = 6 < 2^3$, but we require 4 bit lengths, since some of binary codes are degenerated (e.g., the codes 1, 001, 011 in the figure have no more branching). We may use the notations $1***$, $001*$, and $011*$ instead, where $*$ gives the mark for "do not care", as is used in the classifier systems [24].

Table 1
Comparison of our model (GCM) with SK model for spin glasses.

System	spin glass	network of chaos
Model	SK model	GCM
Randomness	given	created by chaos
	static	dynamic
Tree	for a metastable state	for an attractor
	static	dynamic
Forest	for an ensemble of metastable states	for an ensemble of attractors

In fig. 19, we have plotted k_n^P and S_n^P as a function of precision P, for ten randomly chosen initial conditions. For attractors with many clusterings, k_n^P and S_n^P increase almost proportionally to $\log(P)$ up to some precision and approach k and S. In the turbulent regime, k^P and S^P smoothly increase till they approach N and $\log(N)$. Results from different initial conditions give the same curve, which again supports the conjecture that there is a unique attractor in this regime.

6. Hierarchical dynamics

The dynamics of our system may be better represented by a hierarchical code, which leads to the description of hierarchical dynamics.

Here the meaning of *hierarchical dynamics* is as follows: In a hierarchical dynamical system, there are many units of different levels, organized in a tree structure. A unit interacts strongly with the other units of the same level. A unit at the lower level is slaved by the unit of its upper level, while there is a small feedback from a lower-level to the higher level. A metaphorical example of hierarchical dynamics can be seen in the dynamics of society. In a society, there are hierarchies like nation, states, city, town and so on. The importance of this kind of system on our mind is typically seen in our ability of "categorization", and is stressed by Minsky as "Society of Mind" [25]. Also we can easily expect the importance of the hierarchical dynamics in ecology and economics.

Haken has proposed the slaving principle [26][#10], which focuses on the constraint from a higher level to lower level. The dynamics in the tree structure is also discussed as ultra-diffusion in ref. [28], and in the context of the dynamics of spin glass [16].

Here we show explicitly that our system belongs to the hierarchical dynamical system in the above sense, if it has hierarchical precision-dependent clusters in section 5.

As the simplest case we consider a 3-cluster-attractor with the tree structure in fig. 21. The three clusters can be coded by 1, 01, and 00. From eq. (6) with $k = 3$, the dynamics of each cluster X^1, X^{01}, X^{00} is written as

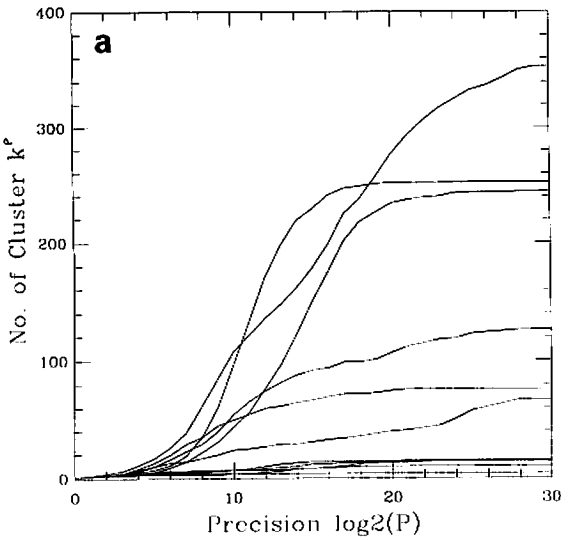
$$X_{m+1}^1 = (1 - n^0\epsilon) f(X_m^1) + n^{01}\epsilon f(X_m^{01}) + n^{00}\epsilon f(X_m^{00}), \tag{13}$$

$$X_{m+1}^{01} = [1 - (n^1 + n^{00})\epsilon] f(X_m^{01}) + n^1\epsilon f(X_m^1) + n^{00}\epsilon f(X_m^{00}), \tag{14}$$

$$X_{m+1}^{00} = [1 - (n^1 + n^{01})\epsilon] f(X_m^{00}) + n^1\epsilon f(X_m^1) + n^{01}\epsilon f(X_m^{01}), \tag{15}$$

where $n^1 = N^1/N$, $n^{01} = N^{01}/N$, $n^{00} = N^{00}/N$ and $n^0 = n^{01} + n^{00}$. Let us introduce the dynamics of a mega-cluster which is a higher-level than 11 and 10. The simplest way for this is the introduction of the weighted average of the two clusters X^{01} and X^{00}, as given by $X^0 = (n^{01}X^{01} + n^{00}X^{00})/n^0$. This gives the dynamics of the node

[#10] For the slaving principal in a stochastic system, see ref. [27].



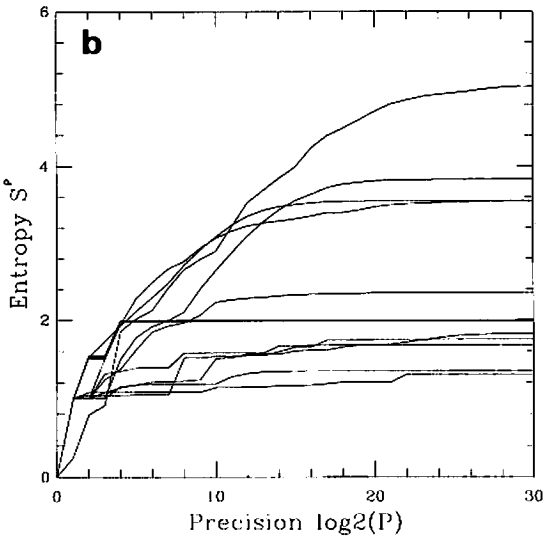
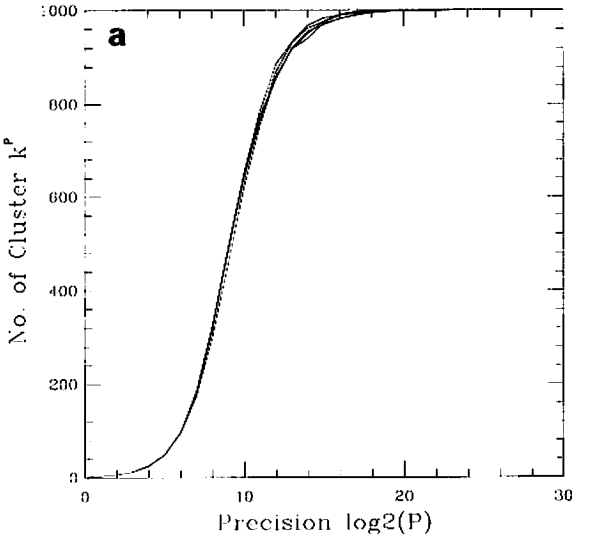


Fig. 19. Number of precision-dependent clusters k_n^P (a) and precision-dependent cluster entropy S_n^P (b), at $n = 5000$. Calculated in the same way as in fig. 18. Results from 12 randomly chosen initial conditions are overlayed. $a = 1.92$, $\epsilon = 0.2$, and $N = 1000$.



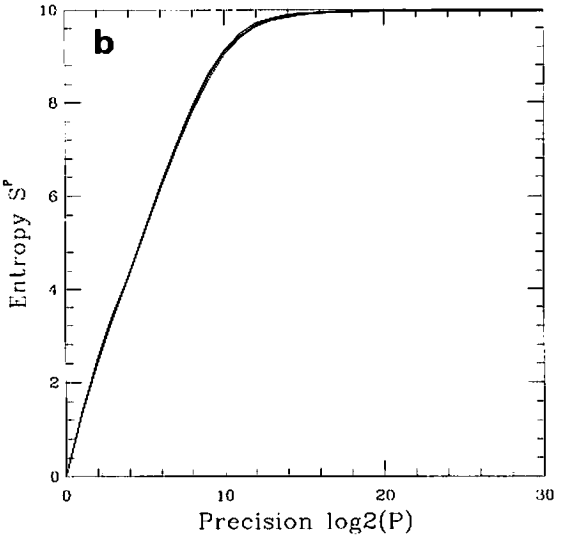


Fig. 20. Number of precision-dependent clusters k_n^P (a) and precision-dependent cluster entropy S_n^P (b). Calculated in the same way as in fig. 19. Results from 10 randomly chosen initial conditions are overlayed. $a = 1.95$, $\epsilon = 0.1$, and $N = 1000$.

at the higher level[#11]. The motion of lower-level clusters X^{01} and X^{00} is described by X^0 and a small discrepancy from it[#12]. Introducing the notation of $X^{01} = X^0 + \delta^{01}$ and $X^{00} = X^0 + \delta^{00}$ (note that $n^{01}\delta^{01} + n^{00}\delta^{00} = 0$), and taking the order up to $(\delta)^2$, we obtain

$$X^1_{m+1} = (1 - n^0\epsilon) f(X^1_m) + n^0\epsilon f(X^0_m) + \frac{\epsilon n^0 n^{00}}{2n^{01}} (\delta^{00}_m)^2 f''(X^0_m), \tag{16}$$

$$X^0_{m+1} = (1 - n^1\epsilon) f(X^0_m) + n^0\epsilon f(X^0_m) + \frac{(1 - n^0\epsilon) n^{00}}{2n^{01}} (\delta^{00}_m)^2 f''(X^0_m), \tag{17}$$

$$\delta^{00}_{m+1} = (1 - \epsilon) f'(X^0_m)\delta^{00}_m + \tfrac{1}{2}(1 - \epsilon)(1 - n^{00}/n^{01})(\delta^{00}_m)^2 f''(X^0_m). \tag{18}$$

(For the logistic map $f(y) = 1 - ay^2$, the above equations are rigorous even in any higher order, since $f''(y) = -2a$ (constant) and higher-order derivatives of $f(y)$ vanish.)

First we note that the dynamics of X^0 and X^1 is the same as eq. (10) for the 2-cluster, up to the order of δ. Our system behaves as if 2-clusters were interacting. In second order of δ there appears corrections to this dynamics. Thus we can view our 3-cluster dynamics as that of 2-clusters with some additional small corrections by the motion of sub-clusters (X^{00} and X^{01}). The motion of sub-clusters is governed by the higher-level cluster ("slaved"). There is a first-order correction which separates the sub-clusters from the higher-level cluster. This correction dynamics (the first term in (18)) again depends on the higher-level cluster. We note that this dynamics is a linear map whose slope is $(1 - \epsilon) f'(X^0_m)$ in the first order: The separation of the two sub-clusters is proportional to the absolute value of the derivative of the motion of the higher-level cluster.

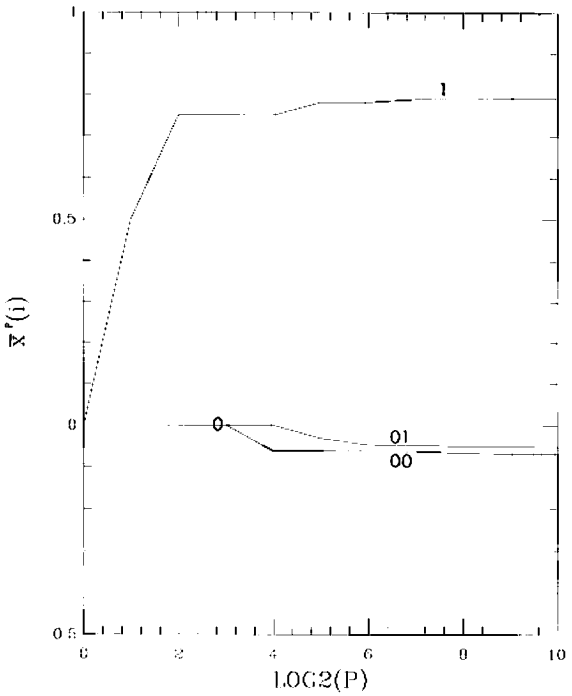


Fig. 21. Precision-dependent clustering of a 3-cluster-attractor, with a hierarchical code.

If the condition

$$\left| \prod_m (1 - \epsilon) f'(X^0_m) \right| > 1 \tag{19}$$

holds, the above δ is amplified with time. Then a small difference between the two sub-clusters is amplified till it becomes the order of the higher-level cluster. The above tree structure is no longer fixed, and changes with time. In this case, our attractor is represented as a dynamically changing tree as is shown below (see for speculation on a

[#11] This introduction of x^0 is not unique. We can take a different way of averaging. Another useful way is the definition by

$$f(X'^0) \equiv [n^{01} f(X^{01}) + n^{00} f(X^{00})]/n^0.$$

If we use this average, we can remove the dependence of the dynamics of x^1 on δ (eq. (16)), and the equation for X^1 is the same as that for the 2-cluster case (eq. (10)). (Instead the equations of X^1 and δ are more complicated.) Anyway, other choices of average give the same result up to the first order of δ.

[#12] By the notion of "precision-dependent clustering" we can assume that the discrepancy is a smaller order than X^0.

dynamical tree, in the content of the generation of language, ref. [29]).

Note that the condition (19) is the same as the stability condition of a coherent attractor (5). The condition means that the separation of sub-clusters occurs if the chaos is strong enough to overcome the stabilizing effect of global coupling. Thus we may call the above chaotic collapse of the tree structure as "chaotic revolt against slaving principle".

The above perturbative derivation of higher-level dynamics can be straightforwardly extended to a case of more-than-2 clusters of the same level or to a branching of more than 2. By applying the above procedure at every level of branching, we can get the equations of hierarchical dynamics. In other words, the motion of a sub-cluster at a given level is slaved by the dynamics of the higher-level cluster and is also governed by the interaction among the sub-clusters at the same level. The interaction from lower level to higher level appears as the second-order perturbation.

6.1. *Dynamical trees*

Once the system has fallen into an attractor, the exact clustering is fixed. However, the precision-dependent clustering can be time dependent. Since our dynamics involves chaos, a small difference $|x(i) - x(j)|$ can be amplified with the temporal evolution. Thus the condition whether $\bar{x}_n^P(i) = \bar{x}_n^P(j)$ or not can change in time. The mechanism of the amplification is clearly seen in the above chaotic revolt. If the chaos is strong enough, small differences among sub-clusters are amplified till they destroy the tree structure. The tree structure can change with time. In fig. 22, we have shown temporal changes of trees.

In figs. 22a1–22a3 we have three "mega"-clusters. One of them has no branching. These fea-

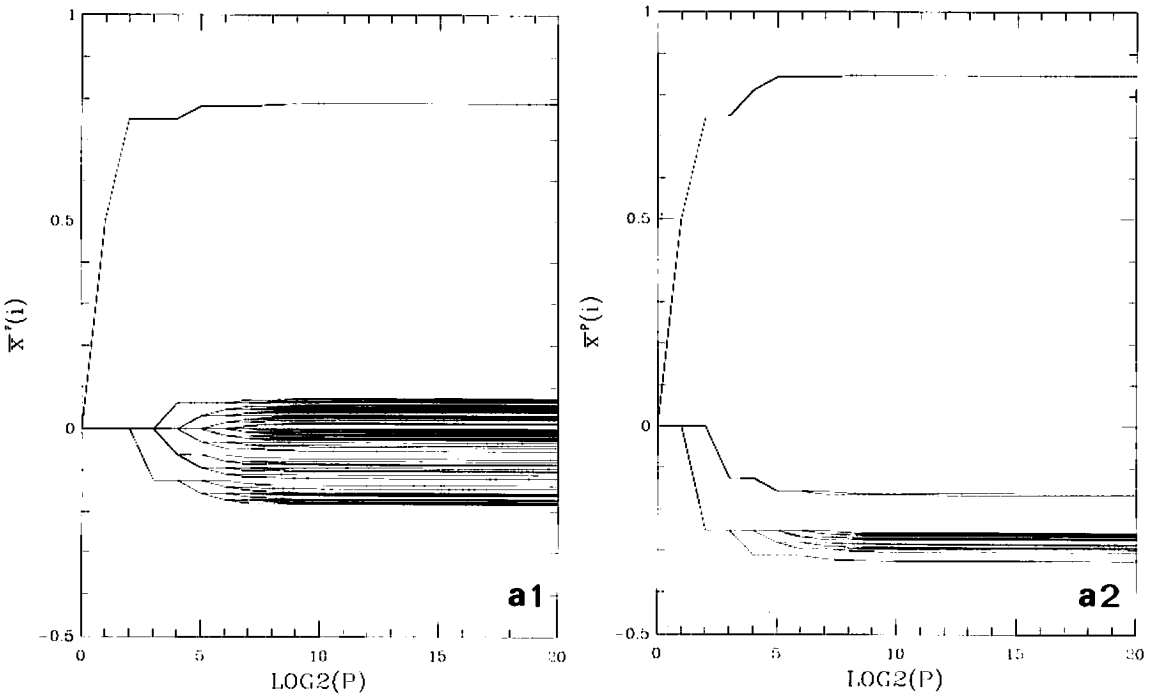


Fig. 22. Temporal evolution of precision-dependent clustering. $\bar{x}_n^P(i)$ is plotted for all i, at $n = 50032$ (a1), (b1); 50064 (a2), (b2), and 50096 (a3), (b3) steps. $a = 1.91$, $\epsilon = 0.2$, and $N = 1000$. From two randomly chosen initial conditions (respectively (a1)–(a3) and (b1)–(b3)).

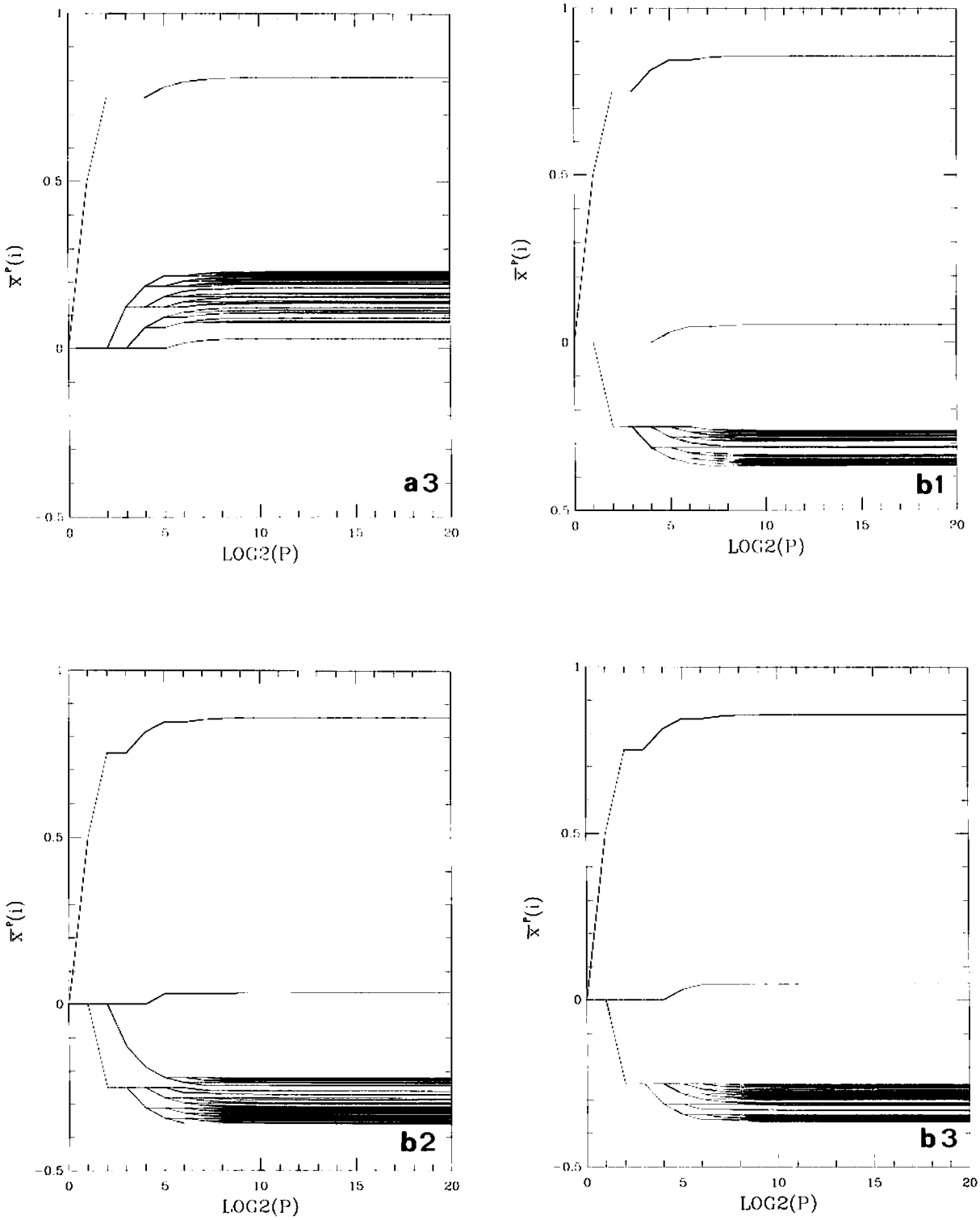


Fig. 22. Continued.

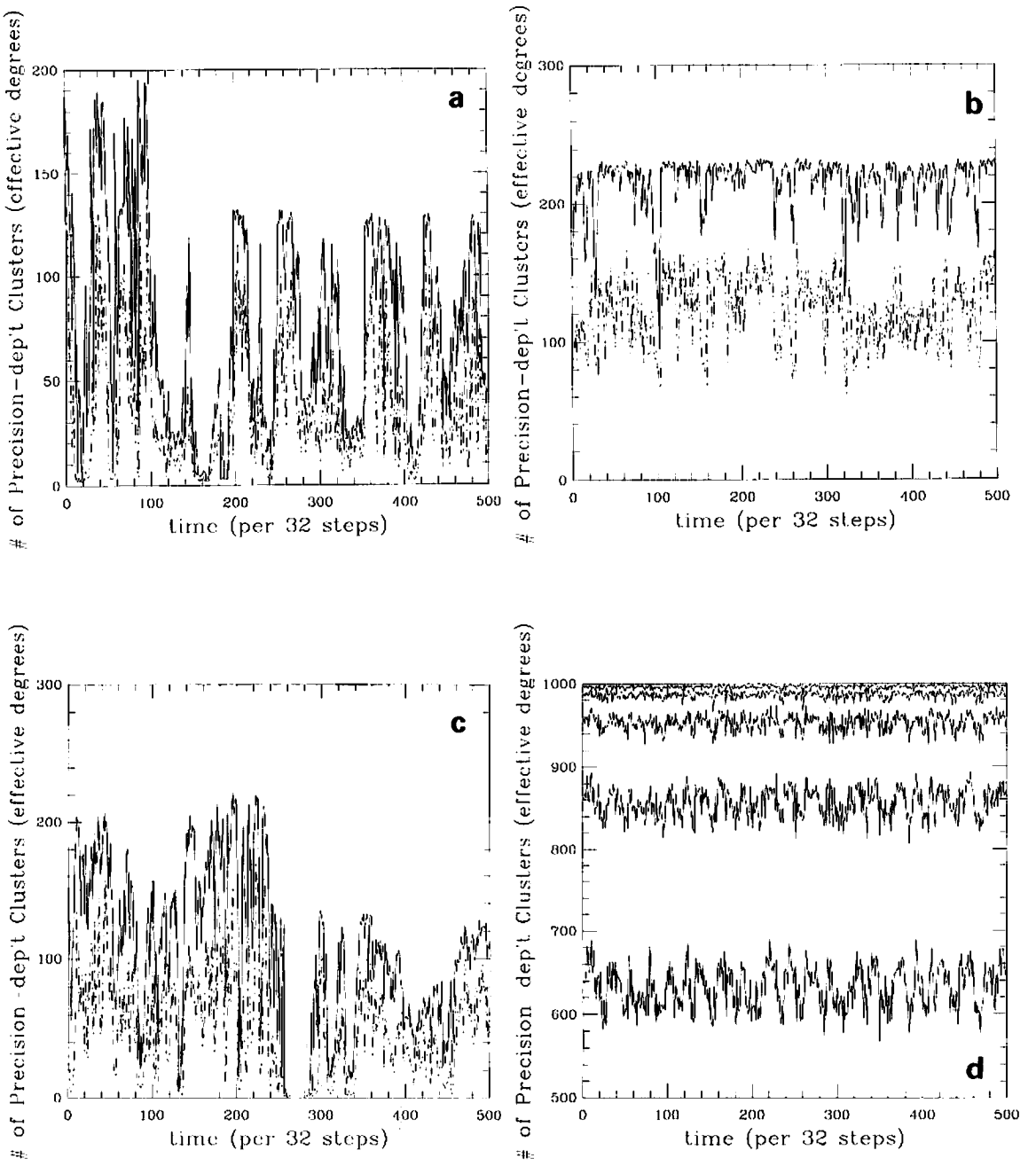


Fig. 23. Temporal evolution of number of precision-dependent clusters. k^{P}_{32n} is plotted as a function of time step n (per 32 steps). Calculated after 50000 transients. $N = 1000$, and starting with a random initial condition. $P = 2^{12}$ (dashed line) and $P = 2^{18}$ (solid line). (a) $a = 1.92$, $\epsilon = 0.2$. (b) Same parameter values as (a) but starting from a different initial condition. (c) $a = 1.62$, $\epsilon = 0.3$. (d) $a = 1.9$, $\epsilon = 0.1$. $P = 2^{10}, 2^{12}, 2^{14}, 2^{16}, 2^{18}$ from bottom to top.

tures are time independent. The clustering within the other two "mega"-clusters, however, are time independent. For example, one of them has two sub-clusters up to fine precision in fig. 22a2. They start to split again in fig. 22a3. In figs. 22b1–22b3, we have three "mega"-clusters, two of which have no branchings and no dynamical change. The tree structure of the other "mega"-cluster has many branchings and is changing slightly with time.

As a simple way to see the temporal change of precision-dependent clusterings, we have plotted k_n^P and S_n^P in figs. 23 and 24. In the partially ordered phase, they exhibit a slow and large variation and give a typical intermittent time series even in a large P (fine precision)[#13]. In fig. 23a k_n^P varies from 2 to 200, even for $P = 2^{18}$, while k_n^P stays at 1 for ~ 600 steps and grows to 100 in fig. 23c (see fig. 25 for the time series of this example). These examples give the intermittent switch from a state with large k^P to an ordered state with small k^P. For more disordered attractors, k_n^P is close to k for most of the time, but it intermittently goes down to smaller value as in fig. 23c.

[#13]Note that k_n^P and S_n^P for $P \to \infty$ should be time independent in an attractor.

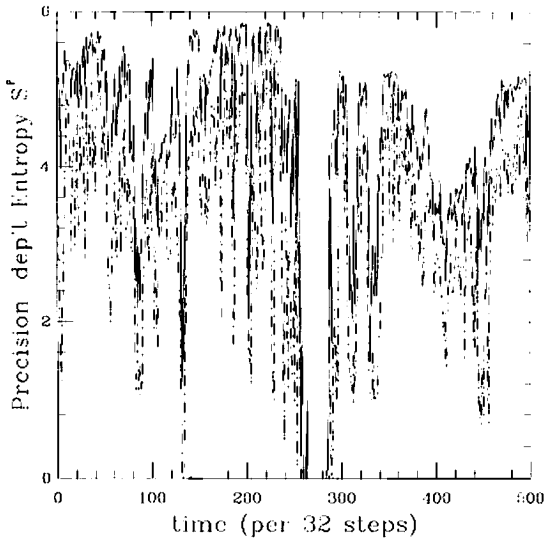


Fig. 24. Temporal evolution of precision-dependent entropy S_{32n}^P. Calculated in the same way as in fig. 23. $P = 2^{12}$ (dashed line) and $P = 2^{18}$ (solid line). $a = 1.62$, $\epsilon = 0.3$ (corresponding to fig. 23c).

The entropy S_n^P exhibits the similar temporal change as k_n^P, as is seen in fig. 24.

In the turbulent regime, the temporal variation k_n^P is very small, and takes almost constant value for large P (see fig. 23d).

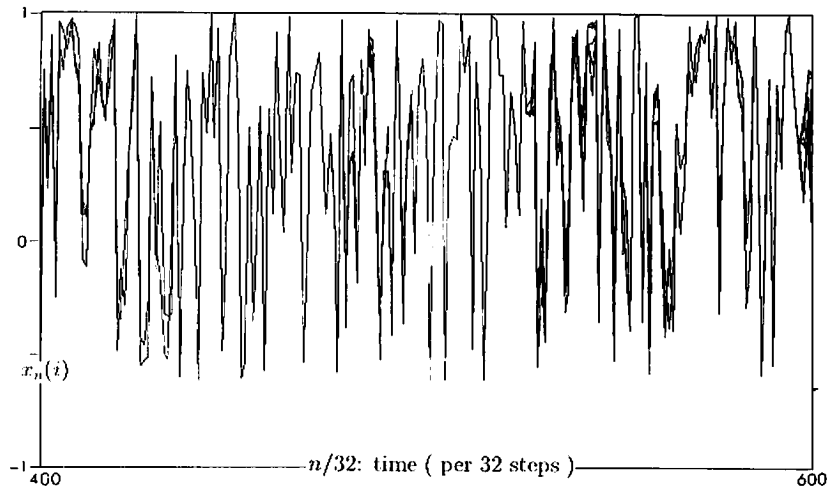


Fig. 25. Time series of an intermittent attractor whose effective degrees of freedom varies from 1 to $\mathcal{O}(N)$. $a = 1.62$, $\epsilon = 0.3$, and $N = 100$. $x_{32n}(i)$ plotted as a function of time n ($n = 400, 401, \ldots, 600$). Note that our system stays very close to a $k = 1$ state for a long time.

6.2. *Coherent structure as dynamics whose effective degrees of freedom changes*

The dynamics of our system $x_n(i) \to x_{n+1}(i)$ at time step n is described by the k_n^P-dimensional map (6), up to precision P. Thus k_n^P can be regarded as the *effective degrees of freedom* at time n up to precision P, to determine the next step. In our examples in fig. 23a–23c, the effective degrees k_n^P changes slowly with time. In fig. 25 (or in fig. 23c), for example, the dynamics is essentially governed just by a single logistic map or a two-coupled logistic map for a time interval of more than 3000 steps, and then the state moves to a higher-dimensional dynamical state.

At the time step where the effective degrees of freedom is very small, the motion is ordered. In each cluster the motion is approximately coherent. In this sense, we may call this motion "coherent structure" [30], because it is low-dimensional ordered motion emerging from a very high-dimensional phase space. We have to note that our system exhibits the *intermittent change between the self-organization towards the coherent structure and collapse of the structure to the high-dimensional disordered motion*. This process is quite analogous with the turbulent process in fluids (see also the discussion in section 9).

The relevance of dynamics whose effective degrees of freedom change with time has been discussed in models of adaptive systems, evolution, and neural dynamics [31]. For most of these studies, we have to put some additional dynamics externally to change the degrees of freedom. A remarkable point in our simple model is that it does not require any additional external change of parameters, but can exhibit this class of dynamics spontaneously as far as we observe our system in a finite precision.

7. Switching among attractors

Let us consider a switch among attractors by a simple input. Since our attractors are coded by a set of numbers, the switch among attractors is represented by the transition between two sets of numbers. By the switch we can change the distribution of clusterings. If we can find a rule of switches, we can tune the dynamical nature of our system, because there is a bifurcation-like phenomenon associated with the clusterings as is discussed in section 4.

Here we consider the simplest case of an input, i.e., $\delta_n(j)$ onto a single site j, at a single time step n. By this input we change only the value of element j from $x_n(i)$ to $x_n(i) + \delta_n(j)$. After the input we iterate the dynamics without any input and wait for our system settling down to another (or the same) attractor[#14]. If $|\delta|$ is small, the system goes back to the original attractor after few steps. If $|\delta|$ is large enough, we can make a switch from one attractor to another.

The switch to change the cluster distribution as $N_m \to N_m - 1$ and $N_l \to N_l + 1$ is simple: Choose i belonging to the mth cluster, and change its value from X_n^m to X_n^l. The switch for this is just to put a pulse $\delta_n(i) \approx X_n^l - X_n^m$ after the nth step's iteration[#15].

First we consider the switching among 2-cluster attractors. Here we call the two clusters + and − depending on the condition $x_{2n}(i) > x^*$ (+) or $x_{2n}(i) < x^*$ (−), where x^* is the unstable fixed point of logistic map $(\sqrt{1+4a} - 1)/2a$ [12].

By an input on site j belonging to the + cluster this site is switched from the + cluster to the −, or vice versa. Thus we can change N_+, by successive inputs of $\delta_n(j)$ (see figs. 26 and 27). In this manner, we can "control" the attractors through an input.

Then, what happens if we try to increase N_+ beyond N_{thr} or decrease below $N - N_{\text{thr}}$? After intermittent-chaotic [10] transients, all + sites change to − and vice versa (see fig. 26), unless the system comes back to the original attractor.

[#14] For a similar approach to cellular automata, see ref. [32].

[#15] Here we choose an input close to, but not equal to $X_n^l - X_n^m$, to avoid an accidental unstable state, and to check the robustness of our switch. We take $|\delta - (X_n^l - X_n^m)| < 0.1$ typically, which makes a switch successful.

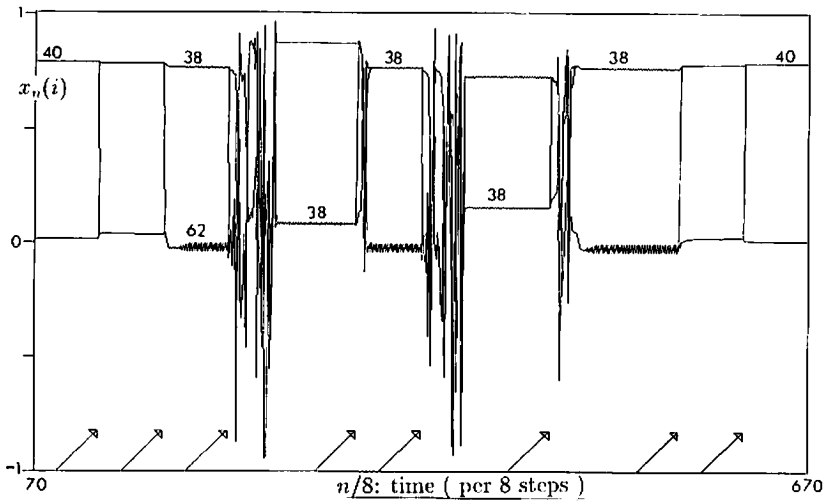


Fig. 26. The time series with switches among attractors. $x_{32n}(i)$ for all i plotted as a function of time. If there are only two lines, the system is fallen to 2-clusters at the time step. Arrows indicate inputs described in the text. $a = 1.95$, $\epsilon = 0.3$, and $N = 100$. $N_{thr} = 62$. Numbers on the time series of $x(i)$ denote the size of the cluster.

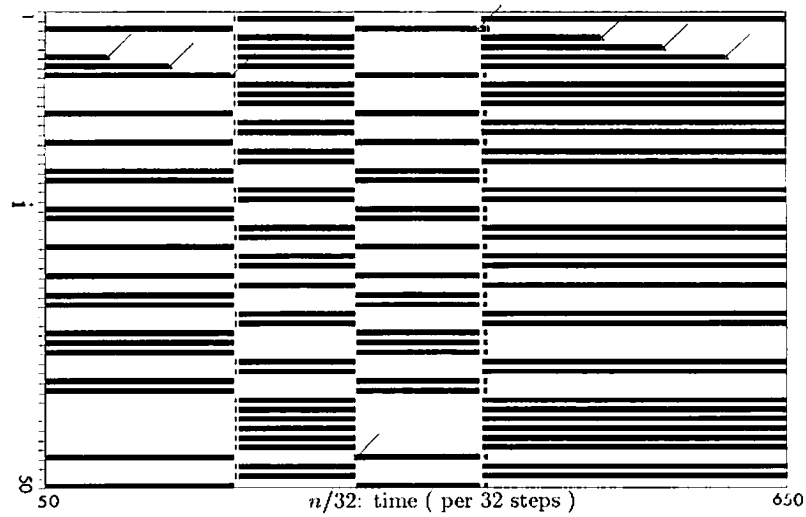


Fig. 27. Site–time diagram with switching: If $x_{32n}(i) > x^*$, the corresponding pixel (i, n) is painted black, otherwise left blank. Arrows indicate the inputs on the corresponding sites and time steps. $a = 1.96$, $\epsilon = 0.3$, and $N = 50$. By successive inputs on site in the + cluster, N_- is increased from 29 to $N_{thr} = 31$, and then a posi–nega switch occurs. Next, by an input to the − cluster, the switch again occurs.

By the switch from $(2,[N_1, N_2])$, we have a transient 3-cluster state $(3,[N_1-1, N_2, 1])$. In the above example, there is a path from this state to the state $(2,[N_1, N_2])$. The above regular posi–nega switch means that paths to other attractors are too narrow to be observed. Schematically this is written as:

$$\text{“}n\text{”} \rightleftharpoons \text{“}n+1\text{”} \rightleftharpoons \cdots \rightleftharpoons \text{“}N_{\text{thr}}\text{”}$$
$$\rightleftharpoons \text{(intermittent transient)}$$
$$\rightleftharpoons \text{“}N-N_{\text{thr}}\text{”} \rightleftharpoons \text{“}N-N_{\text{thr}}+1\text{”} \rightleftharpoons \ldots,$$

where “n” denotes the 2-cluster attractor with $N_+ = n$.

This posi–nega switch belongs to “crisis” [33] in dynamical system theory. The crisis here has an orbit connecting from a two-dimensional phase space to a three-dimensional space, and then again coming back to the original two-dimensional phase space.

For some parameter values and sizes (e.g., $a = 1.93$, $\epsilon = 0.3$, $N = 50$), there is another 2-cluster attractor corresponding to the above intermittent state [11]. The above switching mechanism still works, and can be schematically written as

$$\text{“}n\text{”} \rightleftharpoons \text{“}n+1\text{”} \rightleftharpoons \cdots \rightleftharpoons \text{“}N_{\text{thr}}-1\text{”} \rightleftharpoons$$
$$\text{“}N_{\text{thr}} \rightleftharpoons \text{chaotic attractor with}$$
$$\text{spontaneous intermittent switches”}$$
$$\rightleftharpoons \text{“}N-N_{\text{thr}}+1\text{”} \rightleftharpoons \text{“}N-N_{\text{thr}}+2\text{”} \rightleftharpoons \ldots.$$

In the intermittent attractor, we can make a switch only to the direction from $N_{\text{thr}} \rightarrow N_{\text{thr}}-1$. If we try to increase N_+ beyond N_{thr}, our state comes back to the original attractor with $N_+ = N_{\text{thr}}$. The control of N_+ by simple inputs works well; we can switch from the intermittently switching state to the ordered two-cluster and vice versa as we like.

Next we discuss the switching among clusters more than two. There are two types of switches: (i) inter-cluster switch (a transition within the attractors with the same number of clusters) and (ii) intra-cluster switch (the number of cluster changes by the transition). By the latter switch, we can change even the relevant degrees of freedom, i.e., (even the integral part of) the dimension of chaotic attractors.

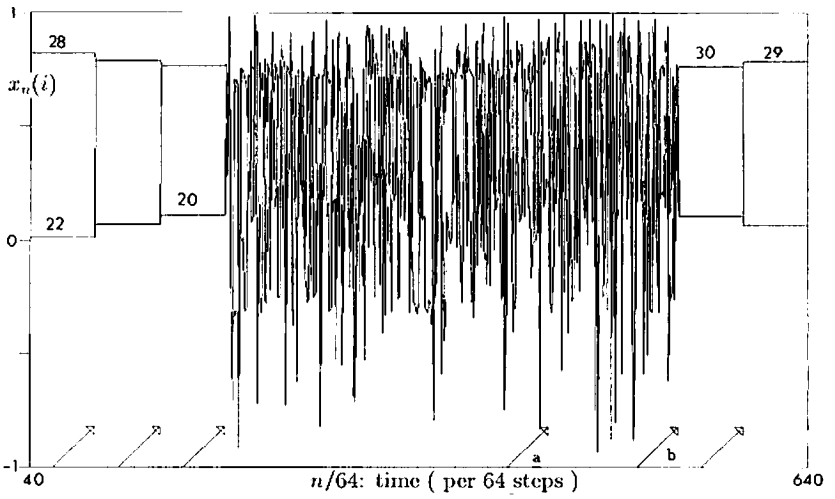


Fig. 28. Switching: Time series with switches among attractors. Calculated and plotted in the same manner as in fig. 26. $a = 1.93$, $\epsilon = 0.3$, and $N = 50$. The intermittent state remains even if we iterate longer time steps. In the input at (a), we have tried to increase N_1 from 31 to 32, but our state has come back to the original attractor with $N_1 = 31$. By the input to decrease N_1 from 31 to 30, we can eliminate the intermittency (the arrow (b)).

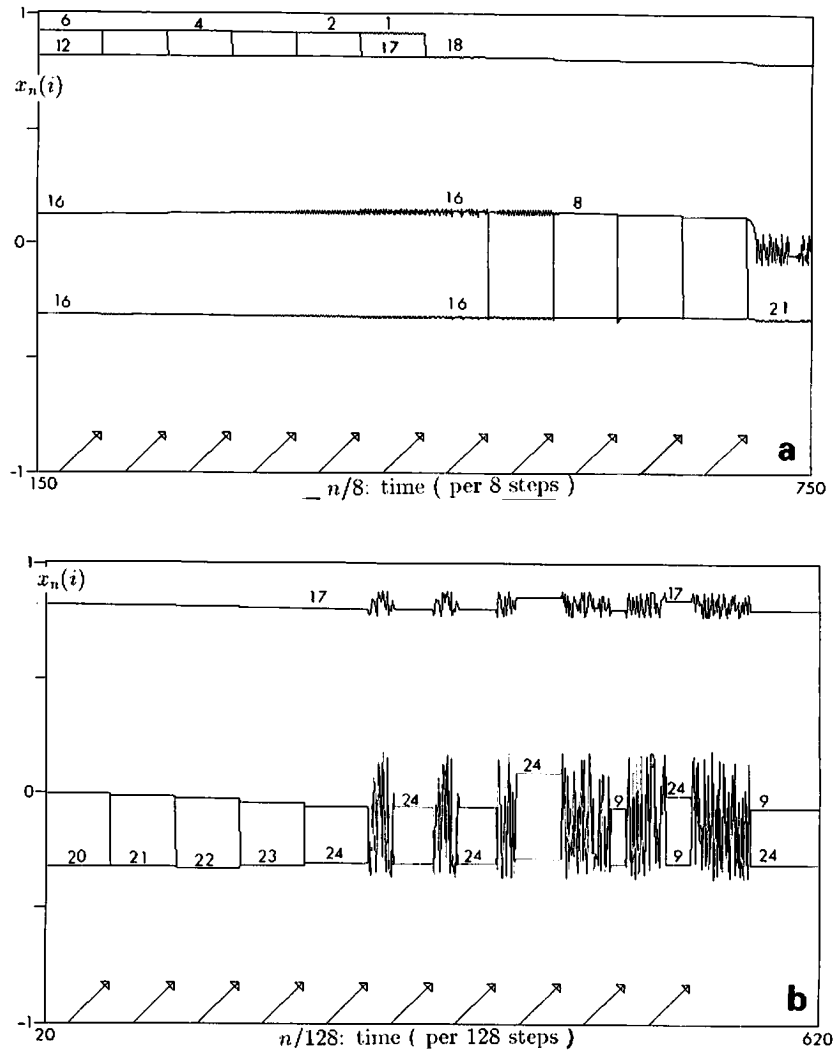


Fig. 29. Time series with switches among attractors. Calculated and plotted in the same manner as in fig. 26. Numbers on the time series of $x(i)$ denote the size of the cluster. (a) $a = 1.59$, $\epsilon = 0.1$, $N = 50$ (removal of one cluster); (b) $a = 1.76$, $\epsilon = 0.2$, $N = 50$ (internal posi–nega switch); (c) $a = 1.83$, $\epsilon = 0.2$, $N = 50$ (hierarchical switch with chaotic revolt); (d) continued from (c); (e) another example of $a = 1.83$, $\epsilon = 0.2$, $N = 50$; (f) $a = 1.58$, $\epsilon = 0.1$, $N = 50$ (remote switch).

Here we explain a variety of functions of switches with clusters more than 2, by taking simple cases (the number of clusters is 3 or 4).

(i) *Hierarchical switch*

(a) *Ordered switching in sub-clusters*

If clusters are hierarchically organized as in sections 5 and 6, the switches are also hierarchically organized. We have switches between lower-level clusters or between higher-level clusters. Of course the switches at the higher-level affect the lower-level clusters. On the other hand, switches at the lower-level do or do not affect the higher-level clusters. In the example of fig. 29a, we can remove one cluster of a lower level by successive inputs on an element belonging to it. This removal of cluster does not affect other cluster structures.

In other examples (fig. 29b–29f), there is a threshold on N_j, such that the state with $N_j =$

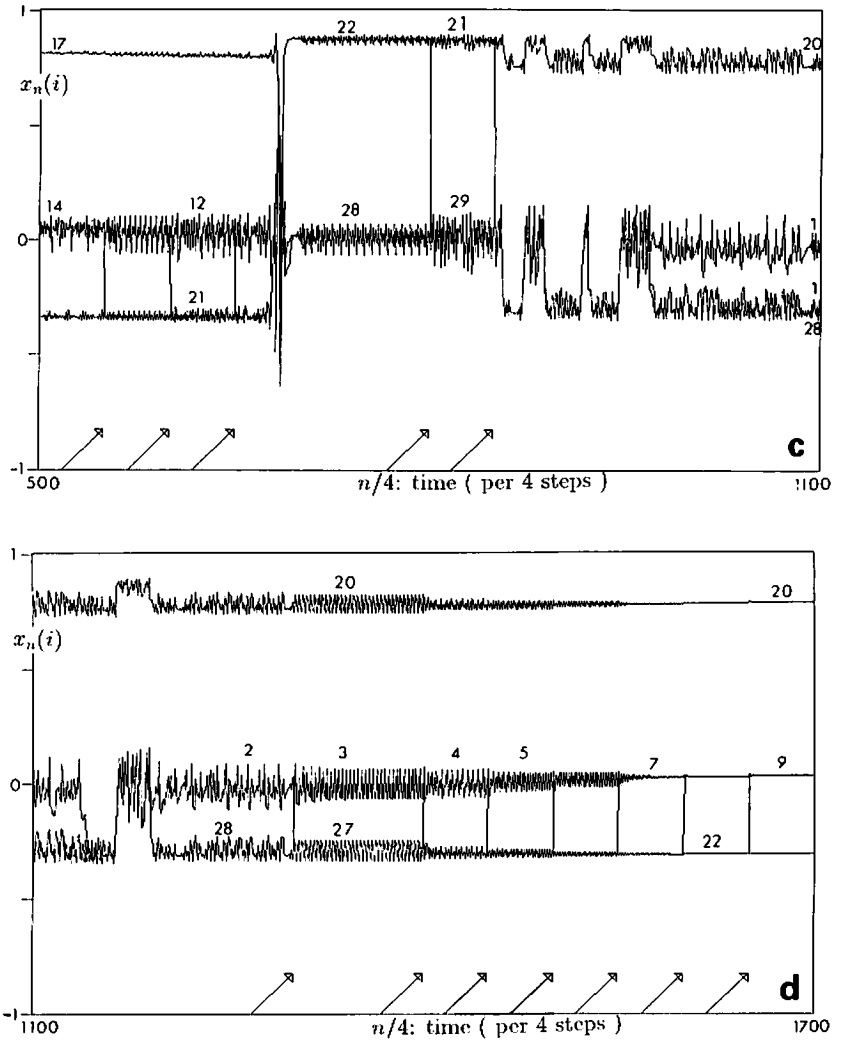


Fig. 29. Continued.

$N_{thr} - 1$ cannot exist as an attractor. Down to the threshold, we can change the number of sub-clusters. The switch here is well organized and is deterministic as in the case for the switch in 2-clusters. In the 3-cluster attractor of fig. 29, we can change N_{01} and N_{00} as we like with fixing N_1.

If we try to increase the difference of the numbers in subclusters beyond a threshold, two types of switches ((b) and (c) as follows) are possible depending on the parameters.

(b) *Internal posi–nega switch*

A posi–nega switch between two sub-clusters occurs, without any change of other tree structures (inter-cluster switch). For example, we can have posi–nega switches only in the two subclusters of 01 and 00, without any change of the number of other clusters (N_1). See fig. 29b, where the threshold is 9 with the code (3, [24, 17, 9]). By this switch the cluster with size = 17 (the one at the top) is not affected.

(c) *Switch with chaotic revolt*

A switch in a lower level propagates to the higher level. This is understood as the "chaotic revolt" in section 6. As the difference between the sizes of the two subclusters is increased, the instability by chaos gets larger, till the condition in (19) for the revolt is satisfied. The cluster structures are no longer stable, and δ in (18) is

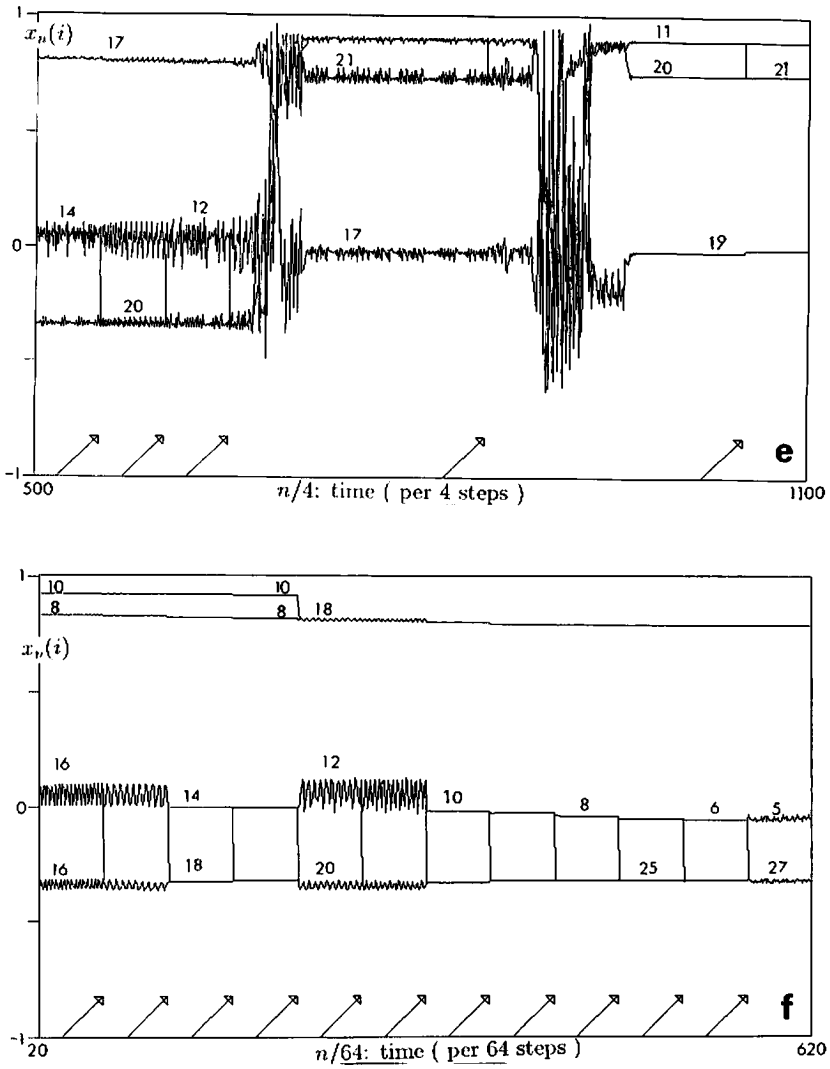


Fig. 29. Continued.

amplified till it destroys the higher-order structure. This is clearly seen in fig. 29c (after the third input). What occurs after the revolt is beyond the perturbative analysis (18). In the above example, a switch to a 2-cluster attractor follows.

Generally speaking, our system can jump to attractors with different clusterings, after the chaotic revolt. There are two possibilities in the switch; one is the intra-cluster switch as is seen in the above example ($k = 3 \to 2$) and the other is the inter-cluster switch with different weights on a tree, like a switch from $(3,[N_{11}, N_{10}, N_0])$ to $(3,[N'_{11}, N'_{10}, N'_0])$.

We have intra-cluster switches from 3-cluster to 2-cluster attractors, and then from 2 to 3 in figs. 29c, 29d. The former switch in the figure occurs from an attractor $(3,[21,17,12])$ to $(2,[28,22])$ via the transient state $(4,[21,17,11,1])$. Note that the above decrease of cluster numbers from 4 to 2 comes from the merging of clusters in the transient region (i.e., $28 = 17 + 11$, $22 = 21 + 1$). The latter switch involves the splitting of clusters, and

merging. In the figure, we have a switch from (2,[29,21]) to (3,[28,20,2]) via the transient states (3,[29,20,1]) and (4,[28,20,1,1]) successively.

An example of inter-switch is given in fig. 29e, where a switch from (3,[21, 17, 12]) to (3,[20,19,11]) occurs via the successive transient states (4,[21,17,11,1]) and (5,[20,17,11,1,1]).

The above hierarchical switch may be useful in the search in a categorized structure, as will be discussed in section 9.

(d) *Remote switch*

Another class of switches in a hierarchical switch is seen in the effect on the change of subclusters on which the input is not applied. In fig. 29f, we have 4-clusters with the tree structure of 11, 10, 00, and 01. Here we have applied inputs successively on elements belonging to a cluster 01 so that a switch from 01 to 00 occurs. Down to a threshold, this switch affects only the subclusters on which the inputs are applied (i.e., $N_{01} \to N_{01} - 1$, $N_{00} \to N_{00} + 1$, without any change of N_{11} and N_{10}). At the threshold, by the input, besides $N_{01} \to N_{01} - 1$ and $N_{00} \to N_{00} + 1$, the cluster structure for 10 and 11 (N_{10} and N_{11}) changes simultaneously. In the figure, cluster 11 is suddenly absorbed into cluster 10. The intra-switch from $k = 4$ to $k = 3$ thus occurs. A remarkable point here is that the switch occurs in clusters with which the input is not related.

(ii) *Deterministic versus stochastic switch*

Another classification of switching is if it is deterministic or stochastic. Of course, our switch should be deterministic if we prescribe all the values of $x(i)$ and the input exactly. The first question is whether it is deterministic if we only prescribe the code of attractors (e.g., $(k,[N_1, N_2, \ldots, N_k])$) and the input. The answer to this question depends on the attractor and parameter. If the attractor is periodic, the switch is deterministic in so far as we have observed. If the attractor is chaotic, we have encountered some examples of stochastic switch, especially in more-than-2-clusters. The next question is how our switch depends on the input. If the boundary between basins of attractors is fractal [34], the switch must be stochastic, unless we prescribe the input exactly. As an example, compare fig. 29c with 29e where we have obtained two different attractors after the third inputs (which are slightly different), even if we start from the same attractor.

To sum up, a switch in our system consists of the following two transient processes. One is "*merging* of clusters", i.e., $(k,[\ldots N_j, \ldots, N_m, \ldots]) \to (k-1,[\ldots, N_j + N_m, \ldots, \ldots])$, and the other is "*splitting* of a cluster", i.e., $(k,[\ldots, N_j, \ldots]) \to (k+1,[\ldots, N_j - 1, \ldots, 1])$. The latter process comes from the chaotic instability.

The above two processes appear only in the transient regime. In the examples we have seen so far, only one or two merging and splitting processes intervene within a switch. This empirical fact leads to strong restriction on a possible variety of attractors after the switch. This kind of restricted transition is also seen in a class of cellular automata[#16].

Detailed study on the switching process with the use of dynamical entropy [32] and basin structures will be discussed in the future.

8. Formation process of clusters and transients

In the ordered phase, if we start from a random initial condition, we can see the temporal ordering process. The number of clusters decreases successively (see fig. 30a). In the late stage, the number

[#16] The regularity of the transition by a switch is characterized by the transition matrix and the dynamical entropy defined by the mutual information of the two successive attractors: This definition and calculation for CA are seen in ref. [32].

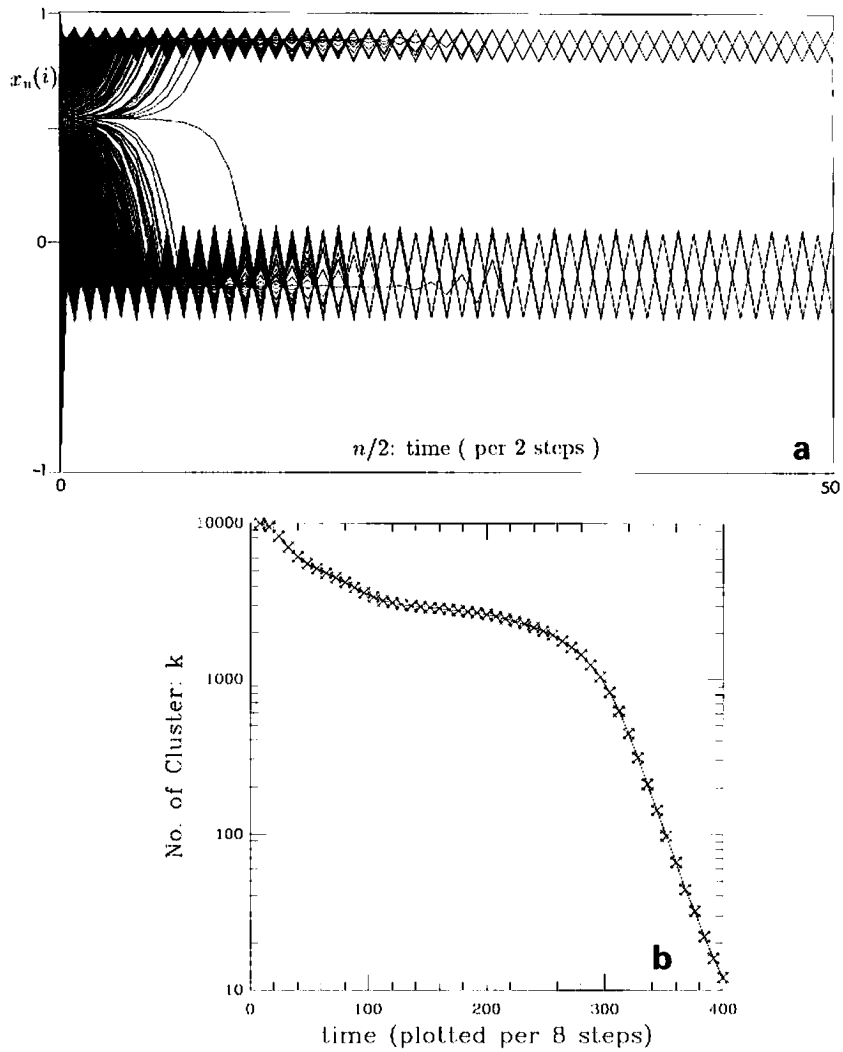


Fig. 30. Ordering process: (a) Time series $x_{2n}(i)$ for all i plotted as a function of time. $a = 1.6$, $\epsilon = 0.1$, and $N = 5000$. Time steps $2n = 2, 4, \ldots, 50$ are plotted starting from a random initial condition. (b) Number of clusters as a function of time: k_n^P, calculated with $P = 10^7$ ($\approx \infty$) is plotted as a funciton of time n. $a = 1.6$, $\epsilon = 0.1$, and $N = 10000$, starting from a random initial condition. Plotted per 8 steps.

of clusters decreases almost exponentially (see fig. 30b). In the partially ordered phase, the decrease, even if it occurs, is very slow. Our system stays at a transient intermittent state for a long time. For this parameter regime, the supertransients discussed by Crutchfield and the author must be important [11].

Another interesting transient phenomenon is seen in the temporal clustering process towards a many-cluster attractor, starting from an almost coherent initial condition. The number of clusters increases successively here[#17]. This class of transient process may be common with the spin glass problem, since the process also passes through successive bifurcations.

Detailed studies on the above two transient processes will be reported elsewhere.

[#17] This may lead to a metaphorical model for the "differentiation" in the developmental biology, since the initially same cells differentiate as time. The formation process of hierarchy in our model may be relevant to the hierarchy we encounter in the differentiation.

9. Discussion: Towards biological information processing

Since our model may capture a novel aspect in neural dynamics as is discussed in section 1, it is important to discuss the capability of information processing of our model. In fact, some novel and important features have been found in our model.

The first one is the hierarchical clustering. Our attractors can have many clusters. These clusters have turned out to be coded by a tree. This kind of tree structure can be frequently seen in the vortices in the turbulence, evolutionary history, categorization in the brain, and the organization of society. We note that our model gives a simple dynamical system which leads to the hierarchical organization.

The most familiar model with hierarchical order is the SK model for spin glasses, which is used as the basis of some neural networks. The hierarchical organization of our model gives us the hope that our system may lead to a novel class of information processing. Since our model has complex dynamics in each element, it may be possible to go beyond SG-type neural networks (in other words, "from tree to forest"), as has been discussed in section 5. Up to now we do not have a theory to characterize the forest, and remain at the stage of the proverb "*don't see the forest for the trees*".

We are not yet sure if our mechanism of the creation of hierarchy is related with examples of hierarchy like the brain, evolution, and society. We have to note, however, that these examples include nonlinear self-reproducing units and a global interaction among units. These two features are included in our model, since the logistic map comes from a nonlinear population dynamics model, and the coupling among units is the simplest global one. If our results are relevant to these systems, we may say that *chaos is the source of hierarchical complexity in nature and society*, since chaos in the logistic map leads to infinitely many possible states and successive splittings of clusters. We have to wait for the future study with more realistic modellings, to confirm the above picture.

We have also frequently encountered with the hierarchical organization in economics. Economics again is governed by the dynamics of nonlinear elements and global feedback.

The second aspect which may have some relevance to biological information processing is the switch among coded attractors. Chaotic switches between attractors are first studied by Davis [35], with the use of adaptive change of a parameter. In our case, all the attractors are organized so that a switch is possible only by a simple input. We have found the intermittent posi–nega switch, hierarchical switch, and stochastic switch.

The intermittent transient in the posi–nega switch reminds us of our mental state when we look at some of Escher's figures[#18]. By looking at them, we wonder which is the "figure" and the "ground". Before we decide from a higher level of mental state, our state changes intermittently. If we fix which is which (corresponding to a switch to a different attractor), we can understand the figure easily.

An important application may be the use of the hierarchical switch. As is discussed in section 7, we can make a switch in a sub-cluster only up to a threshold. If we try to switch beyond the threshold, chaotic revolt occurs, and our system chooses an attractor with a different tree structure.

Let us take a situation of a recall in categorized data. For example, assume that you see a charming lady, and you recall that she was a classmate at some school. You start to recall all the classmates in high-school like Yoko, Kyoko, Akina, and so on, but it turns out that you cannot find the lady in your storage. You hit the capacity of storage ("threshold" in our model). You try to go up in the hierarchy and to search in a different category, like the classmates in junior-high, or in college. In our model, this corresponds to the switch to attractors of different cluster structures (as in section 7) by chaotic revolt. In our model the switch as this level seems to be stochastic. You

[#18] Figures by M.C. Escher (e.g., Circle Limit IV). A metaphorical relation of Escher's figures with dynamical systems is suggested in ref. [5].

may start to search among the junior-high or college. Then you search the memory in that class (in that clustering structure), and may finally find "Aha, it's Ikuko".

This example suggests that our model may be useful in hierarchical search. When we cannot find at the lower-level (i.e., we reach the threshold of the cluster size or memory size), we can automatically switch to a different category and start searching there.

Our remote switch may also be useful to control a system from a different area of storage. A stochastic switch may be relevant in fuzzy control. As is discussed in section 7, the switch is not random, but is strongly restricted. With a high probability the switch leads to few specific attractors. This kind of restricted transition, which is also seen in a class of CA, may be important in the information processing.

To close this subsection, we note that nontrivial switches in our system use the edge of cluster distribution, where the attractors are chaotic. This is consistent with the recent observations in CA [36, 37] and CML [4] that the edge of chaotic phase is relevant to the information processing, as is beautifully phrased by the title of ref. [36].

9.1. *Relevance to neural dynamics*

Although our model is too simple to have a direct connection with neural dynamics, some of our novel notions may be relevant to the understanding of it.

(i) *Search with chaos*: Assume that we store a categorized pattern in each attractor. Searching is carried out by the switching in section 7. What is remarkable here is that our attractor is chaotic at the edge of a condition of clusterings. If we try to switch to a different tree structure by going beyond the threshold, our dynamics exhibits a strongly chaotic behavior, and then switches to an attractor with different clustering conditions. In other words, we have to pass through a strongly chaotic state to switch to a novel class of memory in our model.

An important experiment has recently been performed by Freeman [38]. He has discovered that the activity of neurons of the olfactory bulb exhibits chaos when a rabbit encounters with novel odors, while for the well-known odor the activity gives regular time series. This result is quite similar to our searching process with chaos.

(ii) *Why are there many results on the dimension of EEG?*: There are increasing interests in dynamical states of EEG [39, 40]. The measurement of the dimension of the time series of EEG has been carried out, which suggests that the dimension strongly depends on our state of brain. This possibly means that our brain dynamics has many attractors with different dimensions. In the partially ordered phase in our model, we have many attractors whose dimension varies from $\mathcal{O}(1)$ to $\mathcal{O}(N)$, depending on the clustering condition. Furthermore, even in a single attractor, the effective dimension (measured by the effective degrees of freedom) can vary intermittently as in fig. 23. The dynamics whose effective degrees of freedom varies must be relevant to the neural dynamics.

(iii) *Epilepsy*[#19]: In epilepsy, an ensemble of neurons exhibit a large spike due to the coherent oscillation of the neurons [38]. When one falls in this state, one's capability of information processing goes down. In our system, there are some parameter ranges in which the coherent attractor and attractors of some clusters coexist. Once we choose the coherent attractor, we can no longer have the ability of information processing.

9.2. *Other models*

We have discussed the simplest system of nonlinear elements with global feedback. It is important to extend our model and see how relevant our observations are to other systems.

Examples include:

(i) Coupled Josephson junction: This can be modelled by a globally coupled circle map [15];

[#19] This is pointed out by Walter Freeman and Bernardo Huberman independently.

that is, the use of a local map $f(x) = x + (K/2\pi)\sin(2\pi x) + \omega$ with global coupling. Preliminary results of globally coupled circle maps show the existence of a similar transition among coherent, ordered, glassy, and turbulent phases. A novel discovery therein is a quasiperiodic state with partial attraction.

(ii) Use of a map with an excitable state: Activity of a single neuron may be represented by a one-dimensional map with an excitable state and refractory periods. With the choice of the map and a global coupling, it may be possible to have a closer connection with the neural dynamics.

(iii) Population dynamics model among species. With a coupling with the environment, we can write down a set of differential equations with global couplings. Is it possible to see the hierarchical organization in ecology as in our toy model?

(iv) Another relevance of our results lies in turbulence as the interaction of vortices. Let us recall the intermittent change of precision-dependent clusters in fig. 23. In the figure the effective degree of freedom varies from 2 to 200. At the time step where the effective degree of freedom is 2, our system forms a "coherent structure", as has been discussed in section 6. In turbulence, we have infinitely many vortices, but at some time steps, the motion is ordered ("coherent structure") and may be described by few degrees of vortices. Then this structure again splits into many vortices. If we replace the term "vortex" by "cluster", this is exactly what occurs in our system in section 6. Since the dynamics of vortices is described by a globally coupled dynamical system [21], we may expect that the similar mechanism to our hierarchical dynamics underlies in the motion. A study towards this direction will be reported in future [41].

Acknowledgements

The author would like to thank P. Davis, I. Tsuda, M. Sano, Y. Kimura, G. Mayer-Kress, B. Huberman, and W. Freeman for stimulating discussions. He would also like to thank D. Campbell, M. Casdagli, S. Rasmussen, S. Forrest, C. Langton, and J.D. Farmer for useful discussions.

Note added in proof

Hierarchical dynamics of our intermittent phase (sections 3, 6) gives "chaotic itenerancy over ordered states with low effective degrees of freedom. Chaotic itenerancy is independently discovered by K. Ikeda (in optical systems) and by I. Tsuda (in neural dynamics) [42].

References

[1] K. Kaneko, Prog. Theor. Phys. 72 (1984) 480; 74 (1985) 1033.
[2] K. Kaneko, Physica D 23 (1986) 436.
[3] K. Kaneko, Phys. Lett. A 125 (1987) 25; Europhys. Lett. 6 (1988) 193.
[4] K. Kaneko, Physica D 34 (1989) 1; D 37 (1989) 60.
[5] K. Kaneko, Ph.D. Thesis, Collapse of Tori and Genesis of Chaos in Dissipative Systems (1983) (enlarged version is published by World Scientific, Singapore, 1986).
[6] J.P. Crutchfield and K. Kaneko, Phenomenology of spatiotemporal chaos, in: Directions in Chaos (World Scientific, Singapore, 1987) pp. 272–353.
[7] I. Waller and R. Kapral, Phys. Rev. A 30 (1984) 2047; R. Kapral, Phys. Rev. A 31 (1985) 3868; G.L. Oppo and R. Kapral, Phys. Rev. A 33 (1986) 4219.
[8] K. Kaneko, in: Dynamical Problems in Soliton Systems, S. Taneko, ed. (Springer, Berlin, 1985) pp. 272–277.
[9] K. Kaneko, Phenomenology and characterization of spatio-temporal chaos, in: Dynamical Systems and Singular Phenomena, G. Ikegami, ed. (World Scientific, Singapore, 1987).
[10] J.D. Keeler and J.D. Farmer, Physica D 23 (1986) 413; H. Chate and P. Manneville, C.R. Acad. Sci. 304 (1987) 609; Phys. Rev. A 38 (1988) 4351; Europhys. Lett. 6 (1988) 591; Physica D 32 (1988) 409.
[11] J.P. Crutchfield and K. Kaneko, Phys. Rev. Lett. 60 (1988) 2715; and in preparation.
[12] K. Kaneko, Phys. Rev. Lett. 63 (1989) 219.
[13] K. Kaneko, Phys. Lett. A 139 (1989) 47.
[14] R.J. Deissler, Phys. Lett. A 120 (1984) 334; Y. Aizawa, Prog. Theor. Phys. 72 (1984) 662; T. Yamada and H. Fujisaka, Prog. Theor. Phys. 72 (1984) 885; 74 (1985) 918; K. Kaneko, Phys. Lett. A 111 (1985) 321;

R.J. Deissler and K. Kaneko, Phys. Lett. A 119 (1987) 397;
T. Bohr, G. Grinstein, Y. He and C. Jayaprakash, Phys. Rev. Lett. 58 (1987) 2155;
F. Kaspar and H.G. Schuster, Phys. Lett. A 113 (1986) 451; Phys. Rev. A 36 (1987) 842;
R. Shaw, Boiling chaos, movie presented at the Conference on Spatiotemporal Chaos, Los Alamos, 1986;
N.H. Packard, private communications (1986);
P. Alstrom and R.K. Ritala, Phys. Rev. A 35 (1987) 300;
H. Daido, Prog. Theor. Phys. 75 (1986) 1460;
S. Coppersmith, Phys. Rev. A 38 (1988) 375;
K. Aoki and N. Mugibayashi, Phys. Lett. A 128 (1988) 349;
P. Grassberger, Lumped and distributed dynamical systems, preprint (1987);
K. Kaneko and T. Konishi, J. Phys. Soc. Japan 56 (1987) 2993; preprint (1988);
H. Kanz and P. Grassberger, J. Phys. A 21 (1988) L127;
G. Mayer-Kress and K. Kaneko, J. Stat. Phys. 54 (1989) 1489;
I. Tsuda and H. Shimizu, in: Complex Systems—Operational Approaches, H. Haken, ed. (Springer, Berlin, 1985);
M. Rotenberg, Physica D 30 (1988) 192;
J. Brindley and R.M. Everson, Phys. Lett. A 134 (1989) 229;
L.A. Bunimovich and Ya.G. Sinai, Nonlinearity 1 (1989) 491;
I.S. Aronson, A.V. Gaponov-Grekhov and M.I. Rabinovich, Physica D 33 (1988) 1;
Y. Oono and S. Puri, Phys. Rev. Lett. 58 (1986) 836; Phys. Rev. A 38 (1988) 1542;
G. Grinstein, preprint (1988);
D. Rand and T. Bohr, preprint (1988);
M.H. Jensen, Phys. Rev. Lett 62 (1989) 1361;
L.A. Bunimovich, A. Lambert and R. Lima, preprint (1989).

[15] P. Hadley and K. Wiesenfeld, Phys. Rev. Lett. 62 (1989) 1335.

[16] M. Mezard, G. Parisi, and M.A. Virasoro, eds., Spin Glass Theory and Beyond (World Scientific, Singapore, 1987).

[17] K. Aihara and G. Matsumoto, in: Chaos, A.V. Holden, ed. (Princeton Univ. Press, Princeton, 1986);
H. Hayashi, S. Ishizuka and K. Hirakawa, Phys. Lett. A 88 (1982) 265.

[18] L.N. Cooper, in: Proceedings of Nobel Symposium on Collective Properties of Physical Systems, B. Lundruist and S. Lundruist, eds. (Academic Press, New York, 1973).

[19] W.A. Little and G.I. Shaw, Behavioral Biology 14 (1975) 115.

[20] J.J. Hopfield, Proc. Natl. Acad. Sci. US 79 (1982) 2554.

[21] L. Onsager, Nuovo Cimento 6 (1949) 279;
H. Aref, Ann. Rev. Fluid Mech. 15 (1983) 345.

[22] M. Eigen and P. Schuster, The Hypercycle (Springer, Berlin, 1979).

[23] K. Kaneko, Prog. Theor. Phys. 69 (1983) 1477.

[24] J.H. Holland, Escaping brittleness, in: Machine Learning II, R.S. Mishalski, J.G. Carbonell and T.M. Mitchell, eds. (Kaufman, Los Altos, CA, 1986).

[25] M. Minsky, Society of Mind (Simon and Schuster, New York, 1986).

[26] H. Haken, Syngergetics (Springer, Berlin, 1979).

[27] K. Kaneko, Prog. Theor. Phys. 66 (1981) 129.

[28] B. Huberman and M. Kerszberg, J. Phys. A 18 (1985) 2338.

[29] I. Tsuda and H. Nomura, Shizengengo-shori 71-5 (1989) (in Japanese).

[30] A.K.M.F. Hussain, J. Fluid Mech. 173 (1986) 303.

[31] J.D. Farmer, A. Lapedes, N.H. Packard and B. Wendroff, eds., Evolution, Games and Learning (North-Holland, Amsterdam, 1986), special issue, Physica D 22 (1986).

[32] K. Kaneko, in: Theory and Applications of Cellular Automata, S. Wolfram, ed. (World Scientific, Singapore, 1986) pp. 367–399.

[33] C. Grebogi, E. Ott and J.A. Yorke, Phys. Rev. Lett. 48 (1982) 1507; 57 (1986) 1284.

[34] C. Grebogi, E. Ott and J.A. Yorke, Phys. Rev. Lett. 50 (1983) 935.

[35] P. Davis, Chaotic switch in delayed-feedback model, preprint ATR (1988).

[36] N.H. Packard, Adaptation towards the edge of chaos, preprint (1988).

[37] C.G. Langton, to be published;
W. Li and N.H. Packard, to be published.

[38] W. Freeman and C.A. Skarda, Brain Res. Rev. 10 (1985) 147;
W. Freeman, Brain Res. Rev. 11 (1986) 259; Biol. Cybernetics 55 (1987) 1.

[39] A. Babloyanz, C. Nicolis and M. Salazar, Phys. Lett. A 111 (1985) 152.

[40] S.P. Layne, G. Mayer-Kress and J. Holzfuss, in: Dimensions and Entropies in Chaotic Systems, G. Mayer-Kress, ed. (Springer, Berlin, 1986).

[41] Y. Kimura and K. Kaneko, to be published.

[42] K. Ikeda, Prog. Theor. Suppl. 100 (1990), in press;
I. Tsuda, preprint.