Emergence of Chaos in Asymmetric Networks

I. Kanter

Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel

(Received 1 April 1996)

The dynamics of a network of N nonlinear elements interacting via random asymmetric weights is studied analytically. A transition from an ordered phase to a chaotic one is obtained when the number of relevant modes used to construct the weights exceeds N^{δ} $1/2 \leq \delta \leq 1$. In the ordered phase the dynamics of each element is characterized by an embedding dimension equal to one, or is dominated by one Fourier component. The transition to chaos reflecting that N coupled elements cannot follow more than N^{δ} modes encoded in the weights was confirmed numerically. [S0031-9007(96)01776-0]

PACS numbers: 87.10.+e, 05.20.–y, 05.45.+b

One of the most promising directions of statistical mechanics during the last decade is its application to the realm of dynamical behavior of asymmetric networks. Although asymmetric coupling is not typical of physical systems, it appears on a macroscopic level in many biological and artificial systems, such as neural networks, Markov processes, and Turing machines. Hence, the ability to develop and extend the theoretical concepts and analytical methods of statistical mechanics to include asymmetric couplings is essential to uncovering the generic properties of such macroscopic systems.

Various asymmetric networks have been considered to date in the literature. Of particular interest is the case of the general asymmetric random network considered in Refs. [1,2]. The system consists of N continuous nonlinear variables connecting with random couplings, $\{J_{ij}\}$. Each J_{ij} is an independent random variable with zero mean and variance of order $1/N$. For large N the system undergoes a transition from the trivial fixed point at zero to a chaotic state as the nonlinearity of the activation function is increased. This picture of a transition to a chaotic flow or to a phase characterized by a fast decay of the autocorrelation function seems to be the generic behavior of *any* unbiased random asymmetric network, including even systems with only weak asymmetry where J_{ij} is strongly correlated with J_{ji}. (The nature of such systems in the presence of weak asymmetry but with binary variables is still in question [3,4].) The existence of a nontrivial ordered phase, at least at a finite gain, is possible only in the presence of strong enough attractive couplings.

A second class of asymmetric networks is those with *spatial* structure of the couplings, as opposed to random couplings. A prototype of such an asymmetric network consists of $N + 1$ degrees of freedom (DOF) with a very particular structure of couplings, corresponding to a Toeplitz matrix [5,6]

$$J_{ij} = W_{D(i,j)}, \tag{1}$$

where $D(i,j) = j - i$ for $j - i > 0$ and $D(i,j) = N + 1 + j - i$ for $j - i < 0$ and $J_{ii} = 0$. (The structure of the mth row of the matrix J, for instance, is $W_{N+1-m}, W_{N-m}, \ldots, W_N, 0, W_1, \ldots, W_{N-m}$.) This network under sequential updating can be considered also as a Sequence Generator (SGen), which is a perceptron with N input units and a single continuous output [6,7]. The input at each successive time step is chosen as follows: the inputs from the previous time step are shifted one unit to the right with the state of the leftmost input unit set equal to the state of the output unit in the previous time step

$$S_l = \tanh\left[\beta \sum_{j=1}^{N} W_j S_{j+l}\right], \tag{2}$$

where for concreteness we choose *tanh* as the activation function. The asymmetric network, Eq. (2), with the Toeplitz structure, as opposed to random asymmetric networks, exhibits in the generic case under sequential updating a very regular periodic or quasiperiodic behavior, independent of the complexity of the weights.

The following questions naturally arise and will be at the focus of our discussion.

First, we concern ourselves with the robustness of the regular behavior of the SGen. Do quasiperiodic flows appear only in the particular case of the Toeplitz coupling structure or they are common also to more general spatial coupling matrices? This question is fundamental since the dynamical evolution of the SGen reflects the dynamics of a single DOF whose successive state depends on its own previous N steps, Eq. (2). This compression of a system with $N + 1$ DOF to a single one with a long-term memory effect on its dynamics is possible only in the case of the Toeplitz structure. Hence, the existence of nonchaotic flows in asymmetric networks whose dynamics are of true N coupled DOF is in question.

Second, what are the general conditions on the coupling matrix which ensure chaotic flows? This issue differs from the generic chaotic flow of one (or a few) DOF whose dynamics has been exhaustively examined. In this case, chaos is induced by the special form of the activation function as for the logistic map, or is sensitive to detailed structure of the differential equation governing the dynamics. In the discussed asymmetric networks, by contrast, the dynamical evolution of each DOF is well

 0031-9007/96/77(23)/4844(4)$10.00

behaved, and complex trajectories are a result of the lack of a collective behavior. From previous studies it seems that even weak asymmetry is enough to destroy an ordered phase; hence the dynamics does not reflect the internal structure of the couplings. The interplay, if any, between the complexity of the coupling matrix and the complexity of the flow (chaotic, quasiperiodic, attractor dimension, etc.) of the $N + 1$ DOF case is in question, and will be discussed in the second part of our work.

Let us now examine in detail the first question, where a few classes of perturbations (but not small ones) around the Toeplitz structure are analytically examined. The dynamical evolution of a fully connected general asymmetric network under sequential updating (in the order $S_{N+1}, S_N, \ldots, S_1, S_{N+1} \ldots$) is given by

$$S_j^{t+1} = \tanh\left[\beta\left(\sum_{m=1}^{j-1} J_{jm} S_m^t + \sum_{m=j+1}^{N+1} J_{jm} S_m^{t+1}\right)\right], \quad (3)$$

where β is the strength of the gain and $t = 1, 2, \ldots, \infty$ is the number of times each DOF has been updated [note that Eq. (3) with $j = N + 1$ and J as in Eq. (1) reduces to Eq. (2)]. The first perturbation to be examined is a coupling matrix of the form

$$J_{ij} = R(i) \cos\left[\frac{2\pi}{N} K D(i, j)\right], \quad (4)$$

where $D(i, j)$ stands for a spatial distance defined by the Toeplitz structure [see Eq. (1)], and $R(i) > 0$. For a constant amplitude, $R(i) = R$ independent of i, the coupling matrix has a Toeplitz structure where each element is a pure cos of one of the possible wave numbers, $K > 0$, and hence the dynamics under a sequential updating is equivalent to the SGen Eq. (2). However, for the generic case where the amplitudes are independent (or even one of the amplitudes differs from the others), the dynamics, Eq. (3), is of true $N + 1$ DOF. The network in this case cannot be compressed onto a SGen, Eq. (2), consisting of only one DOF. The dynamical solution of Eqs. (3) and (4) is given to leading order in N by

$$S_j^t \sim \tanh\left\{A(j) \cos\left[\frac{2\pi}{N} K(j - t)\right]\right\}. \quad (5)$$

This solution can be verified by evaluating the sum over m in Eq. (3) and expanding S_j^t in a power series in $A(j)$. The constant $A(j)$ depends on β through the equation

$$A(j) = \frac{\beta R(j)}{2} \sum_{m(\neq j)} f_1(A(m)) \sim \frac{\beta R(j)}{2} N \overline{f}_1, \quad (6)$$

where $f_k(A)$ is defined by $\tanh[A \cos(x)] \equiv \sum_{m>0} f_{2m-1}(A)\,(e^{i(2m-1)x} + e^{-i(2m-1)x})$ and $\overline{f}_1 \equiv N^{-1} \times \sum f_1(A(m))$. Note that in leading order, $A(j)$ depends on $A(m)$ only through $\overline{f}_1$. Expanding Eq. (6) for small $\{A(m)\}$ we found that the network undergoes a Hopf bifurcation at β_c given by

$$1 = \beta_c \sum_{m=1}^{N+1} \frac{R(m)}{2 + \beta_c R(m)} \sim \frac{\beta_c}{2} \sum_{m=1}^{N+1} R(m), \quad (7)$$

where for $R(m) = O(1)$, $\beta_c = O(1/N)$ and for constant amplitudes, the Toeplitz case, $\beta_c = 2/NR$, as expected [6]. For $\beta < \beta_c$ the only solution is the trivial fixed point $S_j^t \equiv 0$. At β_c, this solution becomes unstable and the system undergoes a Hopf bifurcation to a periodic flow, Eqs. (5) and (6). The periodicity of the flow of the state of the $N + 1$ DOF is $N(N + 1)$, since the periodicity of each DOF is N, Eq. (5). Hence, both the power spectrum of the spatial structure of the network, S_i^t $i = 1, 2, \ldots$ and that of the temporal power spectrum of each DOF, S_i^t $t = 1, 2, \ldots$, exhibit a dominant component at the same wave number, K. However, one should remember that Eqs. (5) and (6) stand for the solution only in the leading order. There are small corrections at some wave numbers mK (m integer) and their weights are related to the power spectrum of $\{f_{2j-1}(A(m))\}$ [8]. These small higher harmonic corrections ensure that effectively on a long-time scale the flow of each DOF will be quasiperiodic. This effect is demonstrated in Fig. 1 where S_i^{t+1} versus S_i^t is plotted for $i = 1, \ldots, 5$. Note that since the corrections are related to the power spectrum of a function of the amplitudes, $f(A(m))$, the results, S_j^t and β_c, should be sensitive to the detailed arrangement of the amplitudes. This effect was indeed observed in networks with a small number of DOF but was found to be invisible for large N of the order of a few hundreds.

These results indicate that although the network consists of $N + 1$ nonredundant DOF, the dynamical evolution of each one of them is basically the same as for the SGen [Eq. (2)] but with the appropriate amplitudes. The correlations among the DOF are expressed only in the strength of the amplitudes and β_c [see Eqs. (6) and (7)].

In the same spirit of the above discussion one can generalize the results to the case

$$J_{ij} = R(i) \cos\left[\frac{2\pi}{N} K D(i, j) + \pi\phi\right], \quad (8)$$

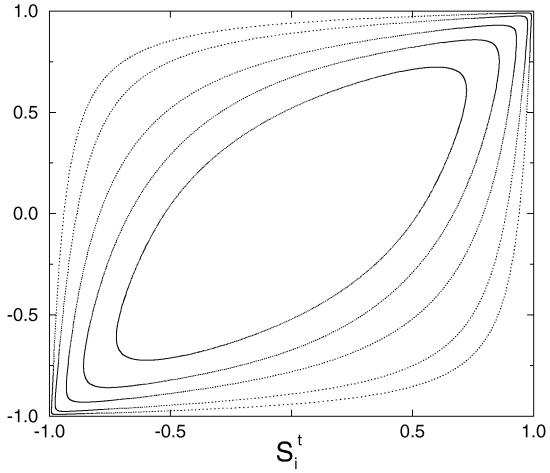

FIG. 1. Result of simulations for S_i^{t+1} vs S_i^t for Eqs. (3) and (4), with $N = 101$, $R(i) \in (0.5, 1.5)$, $\beta = 0.035$, and the five quasiperiodic flows are related to the first five DOF ($i = 1, \cdots, 5$).

where all rows have the same phase shift $\pi\phi$ and $\phi \in [-1:1]$. For $K \gg 1$ one can show that S_i^t is given in the leading order of N and K by

$$S_j^t \sim \tanh\left\{A(j)\cos\left[\frac{2\pi}{N}(K_1 j - K_N t)\right]\right\}, \qquad (9)$$

where

$$K_1 = K + \phi \qquad K_N = K + (N+1)\phi \qquad (10)$$

and the amplitudes and β_c are given by

$$A(j) = \frac{\beta N R(j)}{2}\frac{\sin(\pi\phi)}{\pi\phi}\overline{f}_1, \qquad (11)$$

$$\beta_c = \frac{2}{\sum R(m)}\frac{\pi\phi}{\sin(\pi\phi)}. \qquad (12)$$

The results of Eqs. (9) and (10) indicate that a phase shift, $\pi\phi$, in the weights results in a frequency shift, $K_1 = K + \phi$, of the spatial state of the network at a given time . However, the frequency shift in the weights has a surprisingly dramatic effect on the temporal behavior of the network. The frequency shift of the temporal successive states of each one of the DOF is proportional to the *size of the network*, $K_N = K + (N+1)\phi$. Results for the power spectrum of the spatial structure of a network and for that of the temporal sequence of one of the DOF are presented in Fig. 2. In the next more general case where the power spectrum of each row consists of the same single component

$$J_{ij} = R(i)\cos\left[\frac{2\pi}{N}KD(i,j) + \pi\phi_i\right] \qquad (13)$$

results are very similar to Eqs. (9)–(12); see [8] for details. The only remarkable differences are that ϕ is replaced now by the average phase $\overline{\phi} \equiv \sum_j \phi_j/(N+1)$ and, for instance, Eq. (9) is given now by

$$S_j^t \sim \tanh[\overline{A}(j)\cos(2\pi/N\{(K+\overline{\phi})j - [K+(N+1)\overline{\phi}]t\}) + \pi\phi_j]$$

and $\overline{A}(j) \propto A(j)\sum_i \cos[i\pi(\phi_i - \overline{\phi})]$. The power spectrum of both the spatial and the temporal evolution of one of the DOF for the case $N = 500$ and ϕ_j a random number uniform in the interval $[0.1:0.3]$ is given in Fig. 2. Note that in case that ϕ_j is distributed homogeneously in between $[-1:1]$, $\overline{A} = 0$, the result obtained for (13) is no longer valid and a broadened power spectrum was observed in simulations [8]. Hence we can conclude that *as long as rows are somewhat correlated the dynamic of the $N+1$ DOF is coherent and is reflected by averaged features like K and $\overline{\phi}$.*

Let us now turn to examine the interplay, if any, between the complexity of the weights and the complexity of the trajectories, where the above discussion hints that the complexity is related to the correlations among rows of the weights. There are *two independent ways* to enhance the complexity of the matrix **J**. In the first, rows are decorrelated but each row is still associated with only one Fourier component. In particular, one can define α prototypes of different rows, $\{K_m, \phi_m\}$ $m = 1, \ldots, \alpha$, and each row is associated at *random* with one types

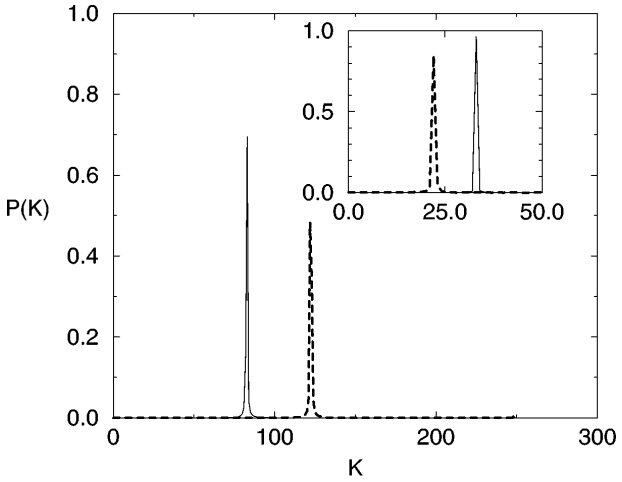

FIG. 2. Results for the temporal power spectrum S_i^t for a system defined by Eq. (8) with $N = 500$, $K = 33$, $\phi = 0.1$, $\beta = 0.005$, and $R(i) = 1$ (full line), and the same system but with $K = 22$ and random ϕ_i, uniformly distributed in the interval and (0.1, 0.3) (dashed line); see Eq. (13). Inset: The same as for the figure but for the spatial power spectrum, S_i^t, $i = 1, \cdots, N+1$.

$$J_{ij} = R(i)\cos\left[\frac{2\pi}{N}K_{m(i)}D(i,j) + \pi\phi_{m(i)}\right]. \qquad (14)$$

The second way to enhance the complexity of the weights is to increase the complexity of each row; however, *all* rows are still correlated. In this case each row is a linear combination of the same γ components

$$J_{ij} = \sum_{m=1}^{\gamma} R(i,m)\cos\left[\frac{2\pi}{N}K_m D(i,j) + \pi\phi_{m(i)}\right]. \qquad (15)$$

For $\gamma = 1$ the system reduces to the one component case, Eq. (13), where for $\gamma = N$ and $\phi_{m(i)} \in [-1:1]$, correlations among rows are absent and the system is equivalent to a random asymmetric network (the amplitudes should follow a Rayleigh distribution).

For finite α in the first direction, Eq. (14), the solution is very similar to that of Eq. (9), but each DOF follows the properties of its own row, $K_1 = K_{m(j)} + \phi_{m(j)}$, $K_N = K_{m(j)} + (N+1)\phi_{m(j)}$ and β in Eq. (12) should be rescaled by $1/\alpha$, since this is the fraction of rows associated with each row's prototype. Hence, different behavior for the spatial and the temporal power spectrum of DOF associated with different prototypes is obtained. Results of simulations with $\alpha = 2$ are presented in Fig. 3. The attractor dimension of each DOF is one (or dominated by one Fourier component), but the attractor dimension of the whole network is expected to be α, since there are α independent modes (generically ϕ_i and ϕ_j are incommensurate). As α increases, the average number of rows associated with each prototype decreases as N/α. Hence, the

signal to a solution of the type of Eq. (9) is N/α, the number of coherent DOF. The rest of the DOF, $N(1 - 1/\alpha)$, are uncorrelated with this solution (associated with other components of the power spectrum) and result in an effective noise of width $\sqrt{N(1 - 1/\alpha)} = O(\sqrt{N})$ on the embedding field of Eq. (3). Hence, the signal to noise ratio becomes $O(1)$ for

$$\alpha \sim O(\sqrt{N}) \tag{16}$$

and a transition to a disordered flow, which may be defined as a chaotic flow, is expected. This behavior was indeed observed in simulations on systems with $N \leq 1000$. For $\alpha > O(\sqrt{N})$, a solution of type (9) becomes unstable and was replaced by a broadened power spectrum characterizing the temporal evolution of each DOF (with lack of correlations among different DOF); see [8] for details.

These results indicate that one has to distinguish between the autocorrelation function of each DOF $C_i^\tau = \sum_{t=1}^{t_0} S_i^t S_i^{\tau+t}/t_0$, and the averaged autocorrelation function $C_{\rm av}^\tau = \sum_{i=1}^{N+1} C_i^\tau/(N+1)$. While C_i^τ is characterized by one dominated component in the power spectrum up to $\alpha = O(\sqrt{N})$, $C_{\rm av}^\tau$ decays exponentially to zero even for finite α, since $C_{\rm av}^\tau$ is a combination of α quasiperiodic functions.

Simulations for the second direction to enhance the complexity of the weights, Eq. (15), indicate that up to $\gamma = O(\sqrt{N})$ *all* DOF follow the same one component, among the γ possibilities, of the power spectrum and the solution is as for Eq. (13). For a greater γ, a broadened power spectrum was observed [8]. This upper bound for the complexity of the rows can be explained by the corrections for the solution of type Eq. (9). The amplitude of the corrections for each one of the other $\gamma - 1$ components is related to the Fourier spectrum of $\{f_{2j-1}(A(m))\}$ [8] which is typically of $O(1/\sqrt{N})$. Inserting these corrections into the dynamical evolution of the system, Eq. (3), one can find that the signal to noise ratio is $N^{1/2}/\gamma$ and is equal to one when $\gamma = O(\sqrt{N})$.

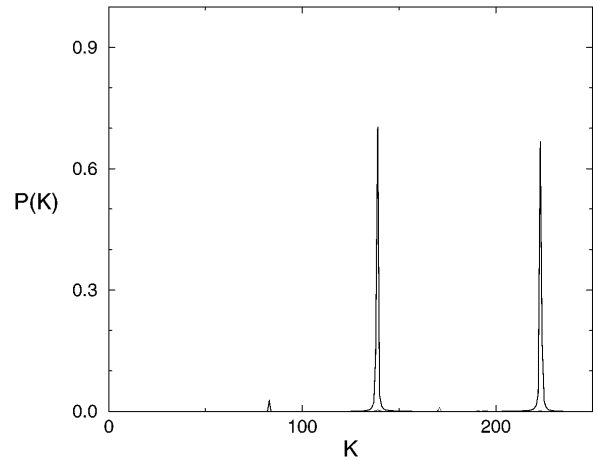

FIG. 3. The system is defined by Eq. (14) with $N = 500$, $\alpha = 2$, $K_1 = 23$, $R(1) = 0.5$, $\phi_1 = 0.4$, $K_2 = 89$, $R(2) = 1$, $\phi_2 = 0.1$, and $\beta = 0.013$. Result of simulations for the temporal power spectrum for two DOF associated with these two prototypes presenting peaks at wave numbers $223(= 23 + 500 \times 0.4)$ and $139(= 89 + 500 \times 0.1)$ following K_N in Eq. (10).

Finally, these two independent ways of increasing the complexity of the weights, Eqs. (14) and (15), can be combined. In this case there are α prototypes of rows where each prototype is a linear combination of γ components of the power spectrum. Dynamically, up to some complexity measure of the weights the dynamical evolution of each degree of freedom is expected to be quasiperiodic (or dominated by one Fourier component). All rows belonging to one prototype follow the same one of its γ components as discussed above. An upper bound for the complexity of a matrix **J** with ordered dynamics is given by $\alpha \times \gamma \leq O(N)$, where by complexity we mean $\alpha \times \gamma$—the number of independent modes constructing the matrix. This upper bound is deduced from the above-mentioned results that $\alpha, \gamma \leq N^{1/2}$. A better bound can be obtained by the extension of the above signal to noise arguments resulting in $\alpha^2 \times \gamma \leq O(N)$ for the combined case [8]. The bound or the exact value may be located by large scale simulations. Nevertheless, one can conclude that the transition to chaotic behavior indicates the following general rule: *N coupled DOF cannot follow more than N^δ $1/2 \leq \delta \leq 1$ "instructions" or different modes reflecting the structure of the weights.* This interplay between the complexity of the weights and the dynamical behavior of the system explains in a natural way the emergence of chaos in a large assembly of well-behaved coupled DOF. A similar interplay was recently found to be common also to static properties of Hamiltonian systems [9].

I especially thank D. A. Kessler for many fruitful discussions and critical comments during the process of the work. I also thank E. Domany and W. Kinzel for valuable discussions.

[1] H. Sompolinsky, A. Crisanti, and H. J. Sommer, Phys. Rev. Lett. **61**, 259 (1988).
[2] A. Crisanti and H. Sompolinsky, Phys. Rev. A **37**, 4865 (1988).
[3] P. Spitzner and W. Kinzel, Z. Phys. B **77**, 511 (1989).
[4] H. Eissfeller and M. Opper, Phys. Rev. E **50**, 709 (1994).
[5] T. H. Berlin and M. Kac, Phys. Rev. **86**, 821 (1952).
[6] I. Kanter, D. Kessler, A. Priel, and E. Eisenstein, Phys. Rev. Lett. **75**, 2614 (1995).
[7] E. Eisenstein, I. Kanter, D. Kessler, and W. Kinzel, Phys. Rev. Lett. **74**, 6 (1995).
[8] I. Kanter (unpublished).
[9] E. Eisenstein, I. Kanter, and D. Kessler, Phys. Rev. E **50**, 3526 (1994).