Lyapunov exponents and dimensions of chaotic neural networks

1991 J. Phys. A: Math. Gen. 24 4557

(http://iopscience.iop.org/0305-4470/24/19/019)

J. Phys. A: Math. Gen. **24** (1991) 4557–4566. Printed in the UK

Lyapunov exponents and dimensions of chaotic neural networks

M Bauer and W Martienssen

Physikalisches Institut Universität Frankfurt, Robert-Mayer-Straße 2–4, D-6000 Frankfurt 1, Federal Republic of Germany

Received 7 November 1990, in final form 12 April 1991

Abstract. Dynamical properties of a network model composed of N continuous elements with randomly chosen asymmetrical couplings and reduced connectivity are studied. We calculate numerically Lyapunov exponents and dimensions for chaotic and non-chaotic states of the network. The dependence on the parameters can be described analytically, in good agreement with the numerical data.

1. Introduction

The physical properties of neural networks have been studied frequently by means of models that show fixed point attractors [1–4]. However, in neurobiological measurements one finds time-dependent, normally non-periodic behaviour. This behaviour has been analysed in terms of nonlinear dynamics [5–7]. It has been claimed in recent studies [8–11] that the time dependence is essential for the data processing in the brain. Therefore the dynamical properties of network models that show complex temporal behaviour have attracted more and more attention during the past few years [12–22]. There have even been examples of possible applications for the chaotic network dynamics [23, 24].

In this paper we study a network model [25] that exhibits deterministic chaotic behaviour. The global structure of the network is neural, i.e. a number of identical simple elements (neurons) that sum up weighted inputs from other elements and give an output with respect to a sigmoid function. In order to distinguish this special kind of architecture from others like cellular automata and coupled map lattices it has become common [13, 15, 22, 26] to call this kind of network with neural architecture a 'neural network' even when no specific learning algorithm is involved. We analyse the nonlinear dynamical properties of such a network numerically as well as analytically. We calculate the spectra of Lyapunov exponents [27, 28] and the Kaplan–Yorke dimension [29] of the non-chaotic and chaotic (strange) attractors. The dependence of the largest Lyapunov exponent on the parameters and the shape of the Lyapunov spectra can be described analytically with only a few assumptions. Even though we consider a deterministic model and study the properties of the attractor, the main results hold also for a similar model, where the model parameters (i.e. the connection strengths) are changed randomly after every time step and a deterministic attractor is never reached.

2. Description of the model

The model consists of N neurons which are represented by *continuous* variables s_j. These neurons are coupled through an asymmetrical matrix $\mathbf{J}$ with $J_{ij} \in [-1,1]$ chosen randomly from an ensemble with equal distribution. The parallel deterministic dynamics is given by

$$s_i(t+1) = \tanh(gh_i(t))$$
$$h_i(t) = \sum_{|i-j|<k} J_{ij} s_j \qquad (1)$$

where g is the gain parameter, and the integer number k represents the connectivity. The topology of the system can be interpreted as a one-dimensional ring of N neurons, where every neuron is connected to the k nearest neighbours in both directions. In the numerical simulations therefore only the synaptic values of $2kN$ (instead of N^2) connections have to be stored. However, for the analytical treatment it is more convenient to formulate this connection scheme in terms of a quadratic band matrix $\mathbf{J}$, where only the k off-diagonal elements in both directions are non-zero. We simulate this network using a random start vector $\boldsymbol{S}(0) = (s_1(0), s_2(0), ..., s_N(0))$ and iterate it according to the dynamics (1). In order to study the properties of the attractor the first transient iterations (approximately 1000) are ignored.

In a preceding paper [25] we have shown that, depending on the values of the control parameters g and k, one can observe stationary, periodic, quasiperiodic and chaotic behaviour (figure 1)†. The line of transition from stationary to time-dependent behaviour turns out to be $k_c \propto g^{-2}$, almost independently of N. This can be understood in terms of a linear stability analysis [20, 26].

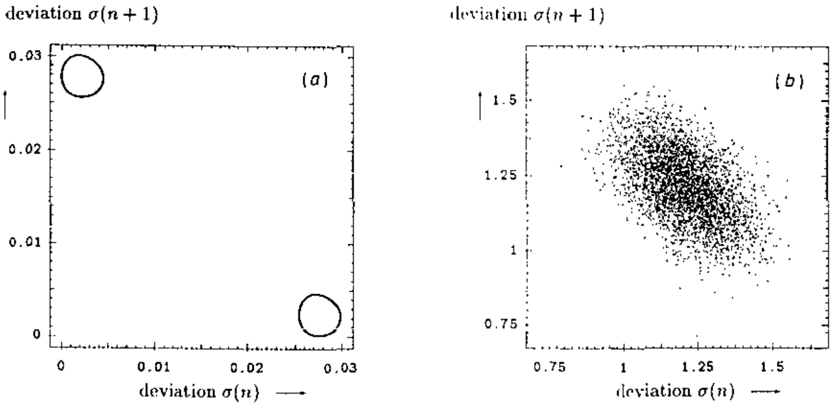

Figure 1. Phase portraits for two different gain parameters and for constant connectivity $k = 5$ ($N = 400$). As an observable the mean square deviation $\sigma(t) = \sum_{i=1}^{N}(s_i(t_0) - s_i(t))$ is plotted. The starting time $s_i(t_0)$ is chosen such that transients have disappeared (approximately 1000 iterations). (a) $g = 0.475$: the system is in a quasiperiodic state; (b) $g = 1.5$: the phase portrait exhibits chaotic behaviour.

† If the system is in this state, a trapping by a periodic orbit has not been found, even after several millions of iterations.

We wish to point out that the details of the model (e.g. topology, distribution of the random numbers) are chosen in way described for the sake of convenience of simulation and analysis only. When not mentioned explicitly, the same model is used always in this paper. It turns out, however, that similar results can also be obtained from a large number of different models with asymmetrical random couplings, when the parameters are chosen in an appropriate way.

3. Numerical results

3.1. Lyapunov spectra

The dynamical behaviour of the system can be characterized in more detail by means of the spectrum of Lyapunov exponents λ_i. The Lyapunov exponents measure the mean exponential separation between closely adjacent points under the action of the dynamics in the N different directions of the phase space. A positive Lyapunov exponent means sensitive dependence on the initial conditions, i.e. chaos. In order to calculate the λ-spectra, the neural network is considered as a N-dimensional nonlinear mapping $\boldsymbol{s}_{t+1} = \boldsymbol{f}(\boldsymbol{s}_t)$. The Lyapunov exponents of $\boldsymbol{f}$ can be calculated as described by Eckmann and Ruelle [27] using the Jacobian of the mapping, defined as

$$\mathbf{T}(\boldsymbol{s}(t)) = \mathbf{D}_s \boldsymbol{f}(\boldsymbol{s}(t)) \tag{2}$$

where $\mathbf{D}_s$ denotes the matrix of partial derivatives. The Jacobian can be calculated analytically for every given state of the system:

$$T_{ij}(\boldsymbol{s}(t)) = \frac{gJ_{ij}}{\cosh^2[g\sum J_{ij} s_j(t)]}. \tag{3}$$

In order to separate the stretching—described by the Lyapunov exponents—from the rotation, we use the method of subsequent **QR** decomposition [27] of the Jacobian, where **Q** is an orthogonal matrix and **R** is upper triangular. The characteristic exponents λ_i can now be calculated from the logarithm of the averaged diagonal elements R_{ii}. The values are indexed such that λ_1 is the largest Lyapunov exponent and the following are in decreasing order: $\lambda_1 > \lambda_2 > \ldots > \lambda_N$; such an ordered set of λ_i represents the Lyapunov spectrum of the system.

We now compare the Lyapunov spectra for different sizes N of the network considered. In order to do this, we average the Lyapunov exponents over a number of realizations (about 20) with different coupling matrices and we normalize the index i to the number N of exponents. We observe (figure 2) that the normalized Lyapunov spectra are almost identical†, i.e. the global structure of the characteristic exponents is independent of the size of the network. Furthermore, we observe that only a few (about 3) Lyapunov exponents are positive, whereas the rest (100 or more) are negative. This shows that only a few degrees of freedom are chaotic and the rest show no sensitive dependence on the initial conditions. This behaviour can be quantified by means of the Kaplan–Yorke dimension, as discussed in the next subsection.

† If not mentioned explicitly, the statistical errors in all figures are smaller than the symbols.

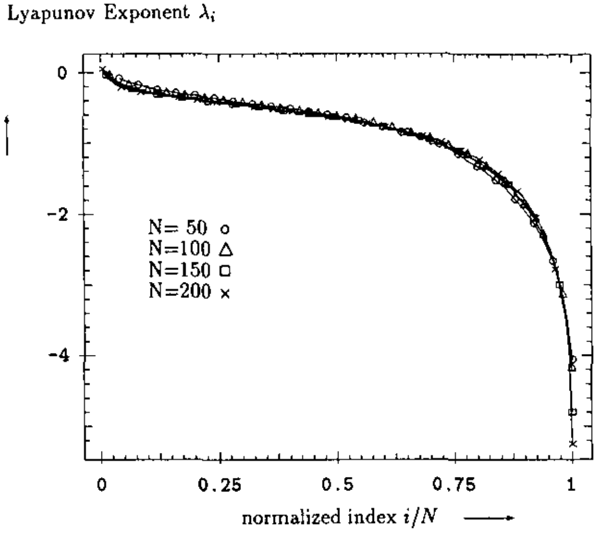

Figure 2. The Lyapunov spectrum for different sizes (N = 50, 100, 150, 200) of the network. The resulting Lyapunov exponents are plotted against the index of the exponent normalized to the size N of the network. These normalized spectra are almost identical.

3.2. Kaplan–Yorke dimension

The static properties such as the fractal dimension of the attractor are connected to the dynamical properties, such as the Lyapunov exponents, by the Kaplan–Yorke relation. If the spectrum of Lyapunov exponents is known, the dimension of an attractor can be calculated according to the Kaplan–Yorke conjecture [28, 29] via the following formalism:

$$D_{\mathrm{KY}} = j + \frac{\sum_{i=1}^{j} \lambda_i}{|\lambda_{j+1}|} \tag{4}$$

where j is the largest integer for which $\sum_{i=1}^{j} \lambda_i > 0$.

This enables us to extract for every given set of system parameters a single quantity that describes the static properties of the attractor. In figure 3(a) we show the dependence of the KY dimension D_{KY} on the gain parameter g. We find that first the dimension increases with the gain reaching a maximum at about $g = 1.5$. After this point the KY dimension decreases even though the largest Lyapunov exponent still increases after this point. It should be pointed out that even the largest attractor dimension is quite low compared to the dimension ($N = 100$) of the phase space.

The dependence of the dimension on the connectivity k for a network of $N = 300$ neurons and a gain near the maximum in figure 3(a) is depicted in figure 3(b). These results show us another feature: the KY dimension starts with a slope of one, which means that the attractors' dimension is about half the value of the number of directly connected neurons (i.e. $D_{\mathrm{KY}} = k$). For higher values of k the slope decreases due to the effects of the finite size of the network.

3.3. Stochastical variation of the coupling matrix

Up to now, we have described the properties of the attractors generated by the deterministic network dynamics. These considerations can be extended on a model where

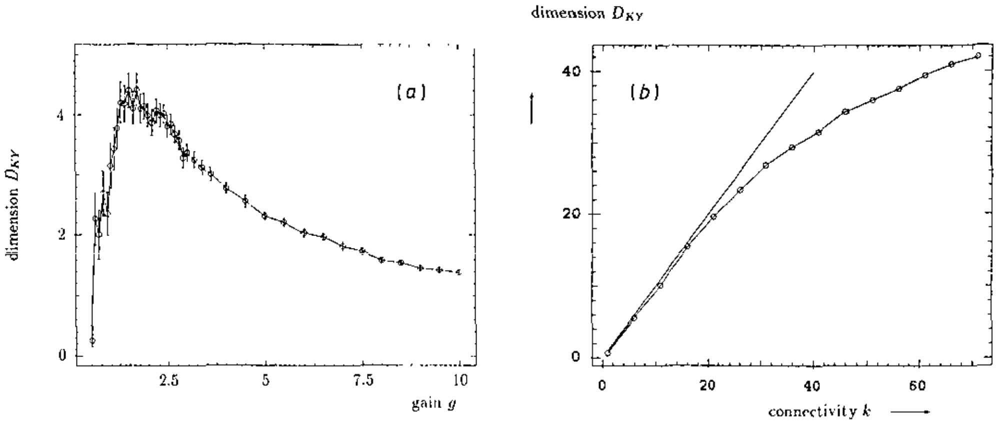

Figure 3. (a) The KY dimension plotted against the gain parameter g ($N = 100$, $k = 5$). The statistical errors are denoted by error bars (3σ). (b) The KY dimension plotted against the connectivity k ($N = 300$, $g = 1$). The line $D_{KY} = k$ is marked.

the existing connections between the neurons are changed in a random fashion after every iteration, i.e. after each timestep new elements J_{ij} of the coupling matrix are chosen from an ensemble of random numbers with an equal distribution in the intervall $[-1, 1]$. The other details of the dynamics (1) are kept unchanged. In the rest of the paper this model will be referred to as the *stochastical model*. Even though an attractor in the deterministic sense is not always reached with this dynamics†, we will apply the same numerical methods on the set of points in phase space, that are reached during iteration. This can be done using the actual (random) elements of J_{ij} and the actual state vector $s(t)$, that are generated in every step of the iteration, together with equation (3), in order to calculate the linearization of the mapping. Using these data we can apply the **QR** decomposition mentioned in §3.1 and described in detail in [27].

This calculation of the Lyapunov exponents of the stochastical model converges approximately within the same number of iterations as for the original model. In figure 4(a) we plot the Lyapunov spectra for both models with different sets of control parameters. For the stochastical model we average the spectra of 300 iterations (i.e. 300 different coupling matrices), and in the deterministic model we plot the average spectra of 20 different coupling matrices with 300 iterations of each of them. We observe that the corresponding spectra of the two models are almost identical. For higher values of the nonlinearity ($g > 1$) the spectra show small differences, but the largest Lyapunov exponents (that mainly determine the dynamics) of the two models are the same for a wide range of parameter values.

This feature can be observed only for sufficiently large networks. In figure 4(b) the dependence of the largest Lyapunov exponents on the size of the network for the deterministic and stochastical model are compared. It turns out that for small networks ($N \leq 10$) the Lyapunov exponent of the stochastical model is larger. For larger networks the Lyapunov exponents become almost identical.

† In all cases we ignore the first (few thousand) iterations (with new coupling matrices in every timestep) to allow global trends to settle down.

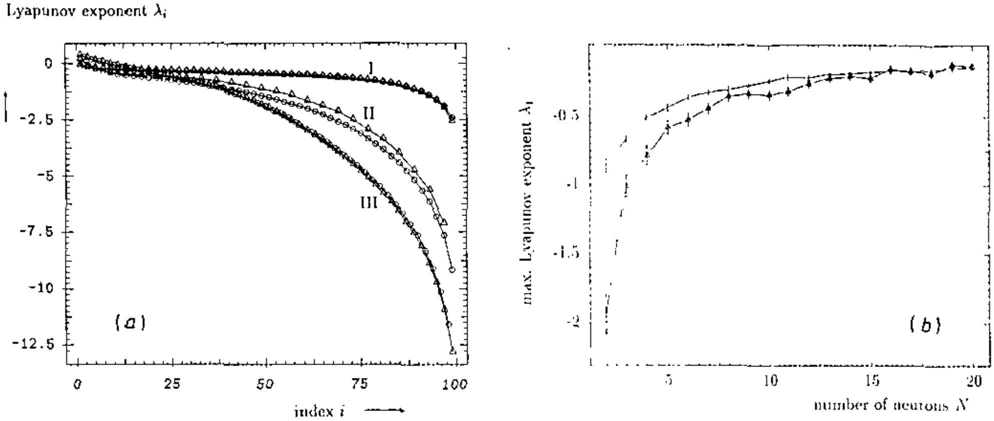

Figure 4. Comparison of the Lyapunov spectra of the deterministic (triangles) and the stochastical (circles) model. (a) For both models ($N = 100$) the spectra are plotted for three parameter combinations: I, $g = 0.5$, $k = 5$; II, $g = 1.5$, $k = 5$; III, $g = 1$, $k = 20$. (b) The largest Lyapunov exponent for the two models is plotted against the size of the network. The statistical errors are denoted by error bars (3σ).

4. Analytical description

4.1. Dependence on the parameters

The quantity that mainly determines the dynamical behaviour of the system is the largest Lyapunov exponent. In order to understand the dependence of λ_1 on the control parameters we want to find an estimate for the largest eigenvalue of the Jacobian, $\mathbf{T}$, equation (3). As a simple ansatz we interpret the Jacobian as a random matrix and approximate the spectrum of eigenvalues by the spectrum of a matrix $\mathbf{G}$ with entries chosen from a Gaussian ensemble [26]. The Gaussian ensemble should have the same mean and variance as the Jacobian $\mathbf{T}$. If the variance of the random matrix is known, the largest eigenvalue $\omega_{\max}$ can be calculated using the Wigner semicircle rule [30]:

$$\beta := \mathrm{var}[G]/N \;\Rightarrow\; \omega_{\max} = \sqrt{\beta}. \tag{5}$$

With the variance v of the non-zero *entries* in the Jacobian (3), we therefore get for the largest Lyapunov exponent

$$\lambda_{\max} = \ln(\sqrt{2kv}). \tag{6}$$

In order to calculate this variance v we replace J_{ij} and $\sum J_{ij} s_j(t)$ in the Jacobian (3) by the random variables† x and y with a Gaussian distribution ($\langle x \rangle = \langle y \rangle = 0$). The uniform distribution of the non-zero elements in the Jacobian $\mathbf{J}$ corresponds to the variance‡ $\sigma_x = \sigma_{y0} = \sqrt{1/3}$,

$$z = \frac{gx}{\cosh^2[gy]}. \tag{7}$$

† The state variable s_j is assumed to be $+/-1$ randomly.

‡ The variance σ_y of the random variable y that replaces the sum in (3) has to be calculated, taking into account that there are only $2k$ non-zero elements in every column of J: $\sigma_y = \sqrt{2k}\,\sigma_{y0}$.

This gives for the variance of z

$$v = \int_{-\infty}^{\infty} \rho(z) z^2 \, \mathrm{d}z \tag{8}$$

where $\rho(z)$ is the probability density of the distribution that is generated by inserting the Gaussian random numbers in (7). It is calculated in equation (9) below by integrating the (Gaussian) probability density of the input variables x and y multiplied with the derivative of the inverse of equation (7):

$$\rho(z) = \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} \mathrm{d}y \, \exp\left(-\frac{y^2}{2\sigma_y^2} - \frac{z^2 \cosh^4(gy)}{2\sigma_x^2 g^2}\right) \frac{\cosh^2(gy)}{g}. \tag{9}$$

The k dependence of the integral (8) is determined by the dependence of σ_y on k. This enables us to fit the k dependence of the largest Lyapunov exponent, if the result of (8) for one k value is obtained by numerical integration. Using the exponent $\lambda_{1/2}$ for $k = \frac{1}{2}$, the equation reads

$$\lambda_{\max} = \lambda_{1/2} + \ln((2k)^{1/4}). \tag{10}$$

In figure 5 the results of this equation are compared to the numerical data. The average maximal Lyapunov exponent has been calculated for 20 realizations ($N = 300$, $g = 1$) of the coupling matrix $\mathbf{J}$. We want to emphasize that in equation (10) no fit of the actual numerically calculated exponents is involved, it is derived completely from the described statistical considerations. If we fit, for example, the $\lambda_{1/2}$ to the numerical data, the agreement would be even better.

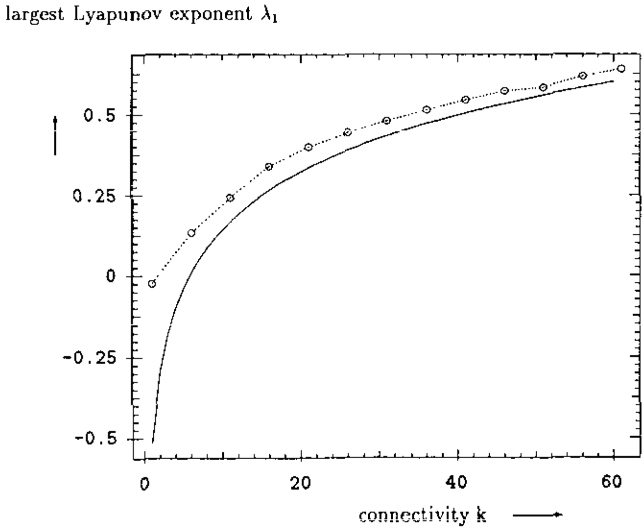

Figure 5. The dependence of the largest Lyapunov exponent on the connectivity k. The numerical results (circles) ($N = 300$, $g = 1$) are compared to the analytical results of equation (10) (full curve).

4.2. Lyapunov Spectra

In this subsection we want to generalize these results to the complete spectrum of Lyapunov exponents. For this reason we consider not only the largest eigenvalue, but the complete spectrum of a Gaussian matrix **G**, again with the same mean and variance as the Jacobian†:

$$\mathrm{var}[G] = \frac{\beta}{N}. \tag{11}$$

This gives a distribution of eigenvalues with the same probability in a circle with radius $\omega_{\max} = \sqrt{\beta}$, therefore we get for the density ρ of eigenvalues

$$\rho(\omega) = \begin{cases} (\pi\omega_{\max}^2)^{-1/2} & \text{if } |\omega| \leq \omega_{\max} \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Hence we get for the probability to find an eigenvalue with a absolute value less than ω

$$g(\omega) = \rho\pi\omega^2. \tag{13}$$

The continuous range between the largest ($g(\omega) = 1$) and the smallest eigenvalue ($g(\omega) = 0$) is now parametrized by means of the variable α. This variable corresponds to the normalized index i/N in the numerical calculations in the limit of $N \to \infty$, where the spectrum of Lyapunov exponents is continuous. Now the eigenvalue and the value of the Lyapunov exponent with the normalized index α can be calculated by setting $g(\omega) = 1 - \alpha$:

$$\begin{aligned} &\omega(\alpha) = \omega_{\max}\sqrt{1-\alpha} \qquad \text{with } 0 < \alpha < 1 \\ \Rightarrow\ &\lambda(\alpha) = \ln\left(\omega_{\max}\sqrt{1-\alpha}\right) \end{aligned} \tag{14}$$

where the largest eigenvalue $\omega_{\max}$ for a given set of control parameters can be calculated according to the results of (10) and (8). Even though the Jacobian of the considered network is not a Gaussian random matrix chosen new in every timestep, we see in figure 6 that the results of this simple approximation describe the numerical data surprisingly well (especially for small gain values $g \sim 1$). Again we want to point out that no fit is involved in this result.

5. Discussion

In this paper we have used nonlinear dynamics, in particular Lyapunov spectra, as a tool to describe a network with neural architecture. Generally, it turns out that the attractor dimensions are quite low compared to the dimensions N of the phase space‡, i.e. the system is determined by only a few of the possible degrees of freedom. This

† This ansatz allows us to calculate generally the Lyapunov spectrum of a dynamical system with a Jacobian which is a Gaussian matrix, i.e. this result would be obtained, if one applies the method described in section 3.1 to Gaussian random matrices independently chosen in every timestep.

‡ For example $D_{KY} \approx 3$ for $N = 100$ ($g = 1$, $k = 5$), growing linearly with k for $k \ll N$.

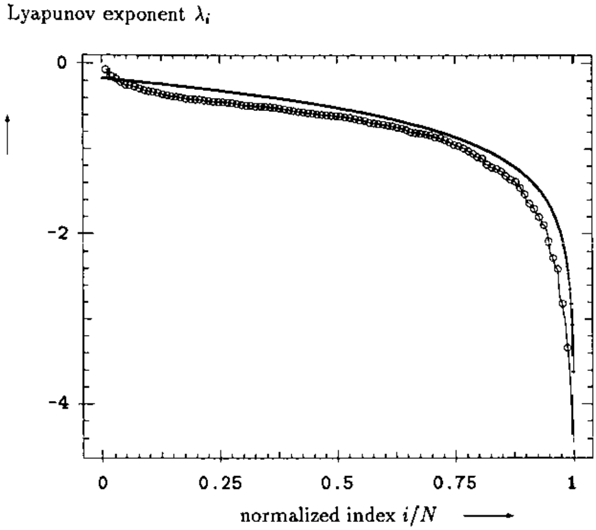

Figure 6. Comparison to the numerical results. The bold curve shows the result of equation (14), whereas the circles denote the numerical results for a network of 400 neurons ($k = 5$, $g = 0.8$). The parameter $\omega_{\max} = 0.846$ is chosen according to (8).

fact shows that it may be possible to find low-dimensional models that describe the integral features of chaotic neural networks.

The comparison between the deterministic model and a stochastic model (where a new random connection matrix is chosen in every timestep) shows that the Lyapunov spectra are almost identical for a wide range of parameter values. This is in agreement with the results for conservative dynamical systems (billiards) with many degrees of freedom [31–34], where the spectra for random evolution and deterministic dynamics have been compared. We conclude from this behaviour that the stochastical system is near enough to the deterministic chaotic attractor, so that the Lyapunov exponents are not systematically different.

Furthermore, we have been able to describe analytically the dependence of the largest Lyapunov exponent on the control parameters, as well as the shape of the Lyapunov spectrum if the largest exponent is known. In order to do this, we have only to calculate the variance of the Jacobian assuming that the entries are independent random variables. The success of this procedure implies that the dynamical behaviour of the system does not depend too much on the actual construction of the model. For example, this kind of treatment holds also for random connection matrices with other (e.g. Gaussian) distributions or different network topologies, as long as the constrains are not too strong (e.g. symmetrical couplings). As discussed before, even stochastical systems can be described with the presented analytical ansatz.

From these results we conclude that the combination of nonlinear dynamics and statistical methods is a promising method for the treatment of highly connected nonlinear systems (e.g. neural networks), that perhaps can even be extended towards a description of the dynamics of the much more complicated biological neural systems.

Acknowledgments

Helpful discussions with C M Marcus are gratefully acknowledged. This work is supported by the Deutsche Forschungsgemeinschaft via the Sonderforschungsbereich 185 'Nichtlineare Dynamik'.

References

[1] Hopfield J J 1982 *Proc. Natl Acad. Sci., USA* **79** 2554
[2] Little W A 1974 *Math. Biosci.* **19** 101
[3] Amit D J, Gutfreund H and Sompolinsky H 1986 *Ann. Phys., NY* **173** 30
[4] Haken H (ed) 1988 *Neural and Synergetic Computers* (Berlin: Springer)
[5] Babloyanz A and Salazar J M 1985 *Phys. Lett.* **111A** 152
[6] Babloyanz A and Destexhe A 1988 *From Chemical to Biological Organisation (Springer Series in Synergetics 39)* ed M Markus and G Nicolis (Berlin: Springer) p 307
[7] Skarda C A and Freeman W J 1987 *Behavioral and Brain Sci.* **10** 161
[8] van der Malsburg Ch and Schneider W 1986 *Biol. Cybern.* **54** 29
[9] Gray C M and Singer W 1987 *Neurosci. Lett. Suppl.* **22** 434
[10] van der Malsburg Ch and Singer W 1988 *The Neurobiology of Neocortex* ed P Rakic, W Singer and S Bernhard (Chichester: Wiley) p 69
[11] Gray C M, Koenig P, Engel A K and Singer W 1989 *Nature* **338** 334
[12] Sompolinsky H and Kanter I 1986 *Phys. Rev Lett.* **57** 2861
[13] Kürten K E and Clark J W 1986 *Phys. Lett.* **114A** 413
[14] Forrest B M 1988 *J. Phys. A: Math. Gen.* **21** 245
[15] Sompolinsky H, Crisanti A and Sommers H J 1988 *Phys. Rev. Lett.* **61** 259
[16] Riedel U, Kühn R and van Hemmen J L 1988 *Phys. Rev.* A **38** 1105
[17] Kürten K E 1988 *Phys. Lett.* **129A** 157
[18] Derrida B and Meir R 1988 *Phys. Rev.* A **38** 3116
[19] Spitzner P and Kinzel W 1989 *Z. Phys.* B **74** 539
[20] Marcus C M and Westervelt R M 1989 *Phys.Rev.* A **39** 347
[21] Marcus C M and Westervelt R M 1989 *Phys.Rev.* A **40** 501
[22] Azamy N and Vibert J-F 1990 *Proc. INNC 90 (Paris)* (Dordrecht: Kluwer)
Bertille J M and Perez J C 1990 *Proc. INNC 90 (Paris)* (Dordrecht: Kluwer)
Wang L and Ross J 1990 *Proc. INNC 90 (Paris)* (Dordrecht: Kluwer)
Banzhaff W, Ishii T, Nara S and Nakayama T 1990 *Proc. INNC 90 (Paris)* (Dordrecht: Kluwer)
[23] Parisi G 1986 *J. Phys. A: Math. Gen.* **19** L675
[24] Basti G, Perrone A, Cimagalli V, Giona M, Pasero E and Mmorgavi G 1990 *Proc. INNC 90 (Paris)* (Dordrecht: Kluwer)
[25] Bauer M and Martienssen W 1989 *Europhys. Lett.* **10** 427
[26] Marcus C M, private communication
[27] Eckmann J-P and Ruelle D 1985 *Rev. Mod. Phys.* **57** 617
[28] Schuster H G 1987 *Deterministic Chaos* 2nd edn (Weinheim: VCH)
[29] Kaplan J and Yorke J 1979 *Functional Differential Equations and Approximation of Fixed Points* ed H O Peitgen and H O Walther (Berlin: Springer)
[30] Wigner E P 1958 *Ann. Math* **67** 325
[31] Rechester A B, Rosenblut N M and White R B 1979 *Phys. Rev. Lett.* **42** 1247
Benettin G 1984 *Physica* **13D** 211
[32] Paladin G and Vulpiani A 1986 *J. Phys. A: Math. Gen.* **19** 1881
[33] Livi R, Politi A, Ruffo S and Vulpiani A 1987 *J. Stat. Phys.* **46** 147
[34] Newman C M 1986 *Commun. Math. Phys.* **103** 121