Coexistence of exponentially many chaotic spin-glass attractors

Y. Peleg,1 M. Zigzag,1 W. Kinzel,2 and I. Kanter1

1Department of Physics, Bar-Ilan University, 52900 Ramat-Gan, Israel
2Institute for Theoretical Physics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany

A chaotic network of size $N$ with delayed interactions which resembles a pseudo-inverse associative memory neural network is investigated. For a load $\alpha = P/N < 1$, where $P$ stands for the number of stored patterns, the chaotic network functions as an associative memory of $2P$ attractors with macroscopic basin of attractions which decrease with $\alpha$. At finite $\alpha$, a chaotic spin glass phase exists, where the number of distinct chaotic attractors scales exponentially with $N$. Each attractor is characterized by a coexistence of chaotic behavior and freezing of each one of the $N$ chaotic units or freezing with respect to the $P$ patterns. Results are supported by large scale simulations of networks composed of Bernoulli map units and Mackey-Glass time delay differential equations.

Over two decades ago a link between statistical physics and neural network theory was established, where associative memory models were mapped onto magnetic Hamiltonian systems with random and frustrated interactions [1, 2]. The capacity, i.e. the number of stored patterns, was calculated using techniques borrowed from random frustrated magnetic systems and in particular advanced method to solve infinite-range spin-glass models [3]. In parallel to the progress in interdisciplinary research on random magnetic systems, the theory of non-linear dynamics developed and recently there have been important steps forward in the understanding of conditions to achieve chaotic networks synchronization [4, 5].

It is reasonable to assume that the two phenomena, spin-glasses and chaos, represent two conflicting trends. The spin-glass (SG) phenomenon is characterized by the existence of exponential number of meta-stable states, in which the local degrees of freedom (spins) are frozen [3, 6, 7]. On the other hand, chaotic dynamics is characterized such that two nearby trajectories diverge from each other as a consequence of the positive Lyapunov exponent [3]. Nevertheless, a network of coupled chaotic units can synchronize [4, 3] to an identical chaotic trajectory. In this paper we show that chaotic networks can have a SG phase, characterized by an exponential number of chaotic attractors with the system size, $N$. Each one of the exponentially many attractors is characterized by a coexistence of chaotic behavior and substantial freezing of each one of the $N$ chaotic units. We start in section I with the definition of the model for a network of chaotic maps. The associative memory features of the model are analyzed at section II and the spin glass phase is examined at section III. The generalization of the chaotic map networks to time delay differential equations is discussed in section IV and concluding remarks are given at section V.

I. MODEL

One of the first examined neural network models within the statistical mechanics framework is an associative recall of memory without error known also as pseudo-inverse model [3, 4]. The model consists of a fully connected network governed by the Hamiltonian

$$H = -\frac{1}{2} \sum_{i \neq j}^{N} W_{ij} S_i S_j$$

where $S_i = \pm 1$ is an Ising spin and the sum is over all pairs of spins. The coupling matrix $W_{ij}$ is defined as

$$W_{ij} = \frac{1}{N} \sum_{\mu, \nu=1}^{P} \xi_{i}^{\mu} (C^{-1})_{\mu \nu} \xi_{j}^{\nu}, \quad C_{\mu \nu} = \frac{1}{N} \sum_{i=1}^{N} \xi_{i}^{\mu} \xi_{i}^{\nu}$$

where $N$ is the size of the network. The stored patterns $\xi_{i}^{\mu}$ ($\mu = 1 \ldots P$, $i = 1 \ldots N$) are $P$ vectors of length $N$ and $\xi_{i}^{\mu} = \pm 1$ with equal probability. All the patterns are minima of the Hamiltonian and the pseudo-inverse network operates as an associative memory up to $\alpha = P/N = 1$ [10], which is the upper bound for the capacity of such networks with symmetric interactions. In the case where the diagonal interactions, $W_{ii}$, are introduced the capacity was found to be $\alpha = 0.5$ [8]. The matrix $W_{ij}$ has $P$ eigenvalues equal to 1 where the patterns are the corresponding eigenvectors,

$$\sum_{j=1}^{N} W_{ij} \xi_{j}^{\mu} = \xi_{i}^{\mu},$$

and all other $N-P$ eigenvalues are identically zero.

The counterpart chaotic map version of the pseudo-inverse model is given for unit $i$, for instance, by

$$X_i^t = (1-\epsilon)F(X_i^{t-1}) + \epsilon \sum_{j=1}^{N} W_{ij} G(X_j^{t-\tau})$$

where $W_{ij}$ is given by [8] including the diagonal elements, $\epsilon$ is the strength of the coupling, $1-\epsilon$ is the weight of the internal dynamics and $F(\cdot)$ and $G(\cdot)$ stand for (non-linear) response functions. For simplicity of the discussion below we mainly present results for the simplest chaotic map, the generalized Bernoulli map [11]

$$F(x) = G(x) = \text{sign}(x)((a|x)| \mod 1)$$

where $a \geq 2$. In this model and for certain values of the parameter $\epsilon$, chaotic behavior and freezing of each one of the $N$ chaotic units is observed.
where $F(x) \in [-1, 1]$ and the map is known to be chaotic for $a > 1$. For all simulations exemplified below the selected parameters are $a = 1.1$, $\tau = 40$ and $\alpha = 0.5$, unless otherwise mentioned.

The state vector $X^t$ of the $N$ units at time $t$ can be written as

$$X^t = \sum_{\mu=1}^{P} b_{\mu}^i \xi_{\mu} + \delta_{\perp}^i$$

The state vector $X^t$ of the $N$ units at time $t$ can be written as

$$(1)$$

where $\delta_\perp^i$ represents the orthogonal part of $X^t$ to the subspace spanned by the $P(\leq N)$ patterns. Since $W_{ij}$ has zero eigenvalues in directions perpendicular to the subspace spanned by the patterns, $\delta_\perp^i$ decays exponentially to zero for the discussed region $a(1-\epsilon) < 1$.

II. MATTIS STATES AND ASSOCIATIVE MEMORY

An initial condition equals to one of the patterns is a fix point of the Hamiltonian system eqs. (1-2). Nevertheless, for the chaotic dynamics, eqs. (15), an initial condition which has a unity overlap with pattern $\mu$

$$X^t = b_{\mu}^i \xi_{\mu} \quad (1 \leq t \leq \tau, b_{\mu}^i > 0),$$

which is known to be chaotic for $a > 1$ and describes merely the behavior of the coefficient $b_{\mu}^i$. For binary patterns, $\xi_{\mu}^i \pm 1$, jumps of the Bernoulli map (i.e. $|X_1^t| > 1/a$) occur for all units simultaneously as exemplified by Fig. 1. The network has solely a unity projection on pattern $\mu$, and the $N$ dimensional dynamics remains in the direction of the pattern.

For a random initial condition of the system composed of projections to all patterns, $X^t = \sum_{\mu=1}^{P} b_{\mu}^i \xi_{\mu}^i$, jumps do not occur simultaneously for all units, Fig. 1b. The dynamics of the coefficients $\{b_{\mu}^i\}$ are coupled and cannot be described independently as in (3). A nontrivial effective coupling among the patterns is generated and time dependent overlaps with the $P$ patterns is observed.

When the initial condition of the system has a projection with one of the patterns which is much larger than all other projections, this overlap remains the maximal one and furthermore increases towards an overlap close to unity. Hence, the chaotic pseudo-inverse model (1) functions as an associative memory with a macroscopic basin of attraction (Fig. 2). Quantitatively, the basin of attraction is defined as the minimal initial overlap with one of the patterns, and random with the other $N - 1$ directions, where the dynamics converges asymptotically to an average overlap greater than 0.9 with this pattern. Figure 2 indicates that the size of the basin of attraction decreases with $\alpha$ and vanishes as $\alpha \to 1$. The size of the basin of attraction also increases with the size of the network, $N$, as exemplified in Fig. 2. Similar trends of the basin of attractions have been observed for the
The Hamiltonian version of the model [11]. Since the dynamics of two different initial conditions inside the basin of attraction converge to an overlap close to unity with the same pattern, it is important to verify whether the dynamics is still chaotic. The inset of Fig. 2 presents the Lyapunov exponent $\lambda$, measured from the divergence of two close initial conditions within an attractor for a given set of $(a, \epsilon)$ with $N = 400$ and $\tau = 50$. Results indicate a positive Lyapunov exponent

$$\lambda \approx \lambda_\infty = \frac{1}{\tau} \ln \left| \frac{ae}{1 - a(1 - \epsilon)} \right| > 0$$

for a large range of $a(1 - \epsilon) < 1$, as was confirmed analytically at for a single chaotic unit with infinite delayed feedback, $\tau \to \infty$ [12].

We exemplify a chaotic network with the coexistence of at least $2P$ chaotic attractors related to the unit eigenvectors, i.e. the $P$ patterns constructing $W_{ij}$. Note that the patterns can take any real values instead of the binary patterns in the Ising Hamiltonian, [11]. In the next section we examine the existence of other attractors besides the patterns and try to estimate their number.

### III. SPIN GLASS STATES

Starting from random initial conditions, which are typically out of the basin of attraction and consist of overlaps $\sim 1/\sqrt{N}$ with all patterns, result in a SG chaotic attractor. The dynamical evolution of the $P$ overlaps from a random initial condition up to $10^3\tau$ steps is depicted in Fig. 3, where each point is averaged over a window of $\tau$ steps. The average is essential since the overlaps dramatically change from one step to another as a consequence of the discrete time nature of the chaotic map. The $P$ overlaps varies substantially with time. In addition, crossings among the overlap amplitudes are observed during a transient time.

In contrast, after $10^3\tau$ steps, the average overlaps settled into a well defined order as depicted in Fig. 3a. A similar fixed order in the overlap amplitudes was observed up to $10^6\tau$ steps. The overlaps are still not macroscopic and are typically bounded (in their absolute value) by $3/\sqrt{N}$ as was confirmed in simulations of network up to $N = 1000$ units. Starting from two close initial conditions within a SG chaotic attractor, a similar Lyapunov exponent $\lambda$ was measured as for Mattis chaotic attractors, eq. [9], independent of $(a, \epsilon)$. The dynamics are still chaotic, but the trajectory stays (on average) in the neighborhood of a unique direction characterized by a given set of projections with the patterns, Fig. 3b.

#### A. Number of Attractors

Quantitative estimation of the freezing of the $P$ overlaps requires a sophisticated method. The overlap amplitudes are ranked in a decreasing order. The pattern

with the largest overlap is labeled as $P$, the pattern with the second largest overlap as $P - 1$ . . . and the pattern with the smallest overlap as 1. This ranking results in a vector $R$ which is a permutation of $\{1 \ldots P\}$. Next we introduce the Manhattan distance [13]

$$\Delta^{\alpha,\beta} = \frac{1}{P} \sum_{\mu=1}^{P} |R^\mu_\alpha - R^\mu_\beta|$$

between two rankings, $\alpha$ and $\beta$. Note that $\Delta^{\alpha,\beta} \in [0, P/2]$. For a random order of $R_\beta$ with respect to $R_\alpha$ the approximate mean value is

$$\langle \Delta^{\alpha,\beta} \rangle \approx \frac{1}{3} \left( P - \frac{1}{P} \right) ,$$

FIG. 3: (color online). Spin-glass chaotic attractors are exemplified in simulations with $\epsilon = 0.9$ and $N = 100$. (a) The dynamical evolution of the overlaps with the $P$ patterns as a function of time, where each point is averaged over a window of $\tau$ steps. Crossings among the $P$ overlaps are observed during the initial transient time. (b) The overlaps of panel a after the network settle into one of the SG chaotic attractors and the overlaps are constant. (c) Probability density of the Manhattan distance along a trajectory while starting from a random initial condition, $\Delta^{\alpha,\beta}$. The distribution is constructed from the calculation of the averaged $\Delta^{\alpha,\beta}$ over segments of $10^3\tau$ steps and time offset of $10^3\tau$ between consecutive segments. (d) Probability density of the Manhattan distance $\Delta^{\alpha,\beta}$ between the trajectories obtained from two random initial conditions and with an offset of $10^3\tau$. The distribution was constructed similar to panel c.

where correlations among $R^\mu_\alpha - R^\mu_\beta$ for different $\mu$’s are neglected.
Results of distribution of the Manhattan distance along different segments of the same trajectory $\alpha$, $\Delta^{\alpha,\alpha}$, when starting from a random initial condition and with an offset time of $10^4 \tau$ is depicted in Fig. [3]. The distribution is constructed from the Manhattan distance between each pair of 100 measurements of $R_i$ with a time offset between two consecutive measurements along the trajectory being $10^3 \tau$. Results indicate an almost complete freezing, where on the average less than a single change in the Manhattan order occurs after an offset of $10^3 \tau$ steps. In order to exclude an accumulated drift in the Manhattan order on much larger time scales, we repeated the distribution of Fig. [3] with an offset of $10^4 \tau$ and $10^5 \tau$ steps among successive measurements and obtained similar distributions. Figure [4] depicts the distribution of $\Delta^{\alpha,\beta}$ between the trajectories of two random initial conditions and with an offset of $10^3 \tau$. The most probable distance, $\Delta^{\alpha,\beta}$, is indeed close to the approximated one, $\approx P/3 = 50/3 \approx 16$, where the lower tail of the distribution ends approximately at 10.

A comparison between Fig. [3]-d for various $N$ indicates that there is a substantial gap between the tails of the distributions of $\Delta^{\alpha,\alpha}$ and $\Delta^{\alpha,\beta}$ as long as $N \gtrsim 30$. This gap is the fundamental observation which enables us to estimate the number of different SG chaotic attractors. Assume all SG attractors have similar basin of attractions, one can estimate the number of attractors for a given $N, \alpha, \epsilon$ and $\alpha$ using the following procedure. We start sequentially from $N_0$ random initial conditions for the network, eq. (1), and after a time offset of at least $10^3 \tau$ the stationary ranking $R$ is identified. In the event the Manhattan distances of this new identified vector with all previously stored ranking vectors are greater than 1, a new SG attractor is identified. Next, its ranking vector $R$ is added to the list of stored SG attractors. In order to well sample the number of SG chaotic attractors, $N_0$ is selected such that $\gamma_i N_0$ ends in previously identified chaotic SG attractors where $\gamma_i \approx 0.1$. We indeed observed that the revisited SG attractors in this procedure occur almost exclusively once, which strengthens our self-consistent assumption that all SG attractors have similar basin of attractions. At the end of this step of the procedure $(1 - \gamma_1) N_0$ distinct SG attractors are revealed. In the second step, we select $N_1$ new random initial conditions and calculate the fraction $\gamma_2$ of revisited $(1 - \gamma_1) N_0$ recorded attractors similar to the previous step. The number of attractors is then estimated to be

$$\frac{(1 - \gamma_1) N_0}{\gamma_2}.$$  \hspace{1cm} (12)

We expect this procedure to yield a lower bound for the number of SG attractors, since attractors with slightly larger basin of attractions are expected to be revealed with higher probability in the first step of the algorithm. Figure 3 depicts the estimated number of SG chaotic attractors as a function of $N$ following the above procedure for $\epsilon = 0.5$ and 0.9. The estimation was carried out in the range of $N = [30, 60]$, since for $N < 30$ the distributions of $\Delta^{\alpha,\alpha}$ and $\Delta^{\alpha,\beta}$ (Fig. [3]-d), are significantly overlapped and for $N > 60$ the number of attractors is too large to be reliably estimated in a reasonable computational time. Results indicate that the number of SG chaotic attractors scales exponentially with the size of the system, $e^{AN}$ with $A \sim 0.19$. It seems that the prefactor $A$ slightly increase with $\epsilon$ as the local dynamics term in (4) is weakened. As a conclusion, a SG phase with exponentially many chaotic attractors is found.

### B. Spin-Glass Phase

A conventional method to identify a SG phase is a freezing of each degree of freedom $X_i$, which cannot be directly deduced from the freezing estimated by the Manhattan distance, $\langle X_i \rangle$. To measure the freezing of $X_i$ we introduce a quantity

$$\phi_i = \frac{\langle X_i \rangle^2}{\langle X_i^2 \rangle}$$  \hspace{1cm} (13)

where $\langle \rangle$ stands for an average over time steps. Eq. (13) is analogous to the Edwards-Anderson order parameter, $q_{EA} = \frac{1}{N} \sum_i \phi_i$ [8]. Figure [5] indicates that along a trajectory inside a basin of attraction of a given pattern (blue curve) all units are almost completely frozen, $\phi_i \sim 0.95$. Note that for a uniform distribution of $X_i$ with a given sign, e.g., $X_i \in [0,1]$, one can verify that $\phi_i = 0.75$; hence $\phi_i \sim 0.95$ indicates a freezing beyond the preservation of the sign of each $X_i$. Indeed, a closer look at the distribution of each $X_i$ indicates that the distribution has considerably narrowed, e.g. mostly within the limited range $X_i \in [0.5,1]$. A typical freezing within a SG chaotic attractor is depicted by the red curve of
Fig. 5. Almost all the units are highly frozen; however, a slight degradation in the freezing is observed in comparison to the blue curve.

The relative freezing direction between two chaotic SG trajectories is defined by

$$q^{α,β} = \frac{⟨⟨X^α⟩⟩ · ⟨⟨X^β⟩⟩}{∥⟨⟨X^α⟩⟩∥ · ∥⟨⟨X^β⟩⟩∥}$$

(14)

similarly to the definition of the overlap between two thermodynamic states of a SG system [12] and $q^{α,β} ∈ [-1, 1]$. It is expected that the absolute overlap between two different SG chaotic trajectories, $|q^{α,β}|$, $α ≠ β$ is close to zero whereas the overlap between two vastly separated time windows along a trajectory of a given SG chaotic trajectory, $|q^{α,α}|$, is close to unity. Figure 5 shows such a behavior where a substantial gap between the distributions of $|q^{α,β}|$ and $|q^{α,α}|$ is observed, which is in agreement also with Fig. 3c-d.

A crucial question is whether the SG chaotic attractors are closely related to the metastable states of the pseudo-inverse Ising spin system, Eqs. (6), with or without the diagonal terms. This analogy is stimulating since the number of metastable states for the infinite-range SG system also scales exponentially with $N$ and with $A ≈ 0.2$ [12]. We examined this question by clipping the sign of the frozen units in a SG attractor, $S^{0}_i = \text{sign}(⟨⟨X⟩⟩)$, and then evolved the pseudo-inverse system, eq. (11), using zero temperature Monte-Carlo dynamics to the nearest metastable state with the Hamiltonian scenario, $W_{ij} = 0$, or with $W_{ij} ≠ 0$ and with long dynamics and obtained a configuration $\{S^{i}_i\}$. In both cases the overlap $\frac{1}{N} \sum_i S^{0}_i S^{i}_i$ was found to be close to zero, indicating no simple interplay between metastable states of the energy surface, eq. (11), and the chaotic SG attractors. The lack of correlation was also obtained in the reverse scenario where the chaotic dynamics, $\{X_i\}$, was initiated by a metastable state of eq. (11).

IV. CONTINUOUS TIME SYSTEMS

Finally we report on the extension of the exponentially many chaotic attractors to a network of coupled differential equations, the Mackey-Glass (MG) equation [14], originated from physiological control systems. This equation was also found to have many applications in hematology, cardiology, neurology and psychiatry [14, 16]. The counterpart network version of the MG equation is given for unit $i$, for instance, by

$$\frac{d}{dt} X_i(t) = -\gamma X_i(t) + β \sum_{j=1}^{N} W_{ij} \frac{X_j(t−τ)}{1+|X_j(t−τ)|^c}$$

(15)

where in the notations of eq. (11)

$$F(x) = -\gamma x(t)$$

(16)

and

$$G(x) = \frac{x(t−τ)}{1+|x(t−τ)|^c} .$$

As before, $N$ stands for the number of units in the network and the matrix $W_{ij}$ is given by eq. (2). The parameters used in our simulations are $β = 2$, $γ = 1$, $τ = 2$, $c = 10$.

Using same arguments as in (10), also in this case $δ^t$ decreases exponentially to zero. Hence, the stability of the Mattis states as well as the size of the basin of attractions can be estimated as in section (11). Similarly, the number of SG chaotic attractors can be estimated using the same procedure used in section (11A). Results for the number of chaotic SG attractors for $α = 0.5$ are depicted in Fig. 4 indicating again an exponential scaling with the network size, $N$.

V. CONCLUSIONS

We investigated discrete and continuous time networks of chaotic units with delayed interactions. It has been shown that the networks function as an associative memory with a macroscopic basin of attraction (Fig. 2).

Moreover, a SG phase which is characterized by an exponential number of attractors (Fig. 4) is identified, where two nearby trajectories within a given attractor diverge from each other with a positive Lyapunov exponent (3).

It is worthwhile to note that the slope of the exponential number of chaotic attractors is very similar for both the Bernoulli map and the Mackey-Glass system, $≈ 0.19$ and $≈ 0.21$, respectively, which might indicate
a universal behavior. Both the Bernoulli map and the Mackey-Glass equation are sign preserving, even in the
generalized forms eqs. \[6\] and \[15\]. An open question
is whether multiple attractors and a similar SG behavior
be can found in a chaotic dynamics with the lack of sign
preservation.

[1] A. Engel and C. Van den Broeck, Statistical Mechanics of Learning, Cambridge Univ. Press 2001.
[2] H. G. Schuster, Applications of Neural Networks, John Wiley & Sons 1992.
[3] H. Nishimori, Statistical Physics of Spin Glasses and Information Processing, Claredon Press (2001).
[4] A. Arenas, A. Daz-Guilera, J. Kurths, Y. Moreno, C. Zhou, Phys. Rep. 469, 93 (2008).
[5] I. Kanter, M. Zigzag, A. Englert, F. Geissler and W. Kinzel, EPL 93, 60003 (2011).
[6] K. Binder and A. P. Young, Rev. Mod. Phys. 58, 801 (1986).
[7] M. Mezard, G. Parisi and M. A. Virasoro, Spin Glass Theory and Beyond, World Scientific, Singapore 1987.
[8] H. G Schuster and W. Just, Deterministic Chaos, Wiley-VCH Verlag GmbH & Co. KGaA, Germany 2005.
[9] I. Personnaz, I. Guyon and G. Dreyfus, J. Phys. (Paris) Lett. 56, L359 (1985).
[10] I. Kanter and H. Sompolinsky, Phys. Rev. A 35, 380 (1987).
[11] S. Lepri, G. Giacomelli, A. Politi, F.T. Arecchi, Physica D 70, 235 (1994).
[12] A. Englert, S. Heiligenthal, W. Kinzel, and I. Kanter, Phys. Rev. E 83, 046222 (2011).
[13] E. F. Krause, Taxicab Geometry: An Adventure in Non-Euclidean Geometry. New York: Dover, 1986.
[14] M. C. Mackey and L. Glass, Science 197, 208 (1977).
[15] M. C. Mackey and L. Glass, From Clocks to Chaos: The Rythms of Life. Princeton Univ. Press, Princeton, NJ, 1988.
[16] M. C. Mackey and J. G. Milton, Ann. NY. Acad. Sci. 504, 16 (1987).