Journal of Statistical Physics, Vol. 76, Nos. 5/6, 1994

Multiplicity of Metastable Retrieval Phases in Networks of Multistate Neurons

S. Bös[1] **and R. Kühn**[2]

Received February 1, 1994

It is demonstrated that networks of multistate neurons storing an ensemble of multistate patterns exhibit a multiplicity of metastable retrieval phases. These phases are described by solutions of fixed-point equations with characteristic retrieval errors. They emerge if the gain of the neural input–output relation is varied. The number of these phases increases as Q^2 with the number Q of gray levels available to each neuron. Implications for the optimal gain function and basins of attraction are briefly discussed. Again, networks endowed with pseudoinverse couplings are found to perform better than networks with Hebbian couplings: at moderate loading levels phases with retrieval errors are destabilized, whereas the error-free phase remains stable up to the theoretically possible maximum, if the gain parameter is properly chosen.

KEY WORDS: Multistate neurons; metastable retrieval phases.

The collective behavior of networks of two-state neurons being well understood (see, e.g., ref. 1), interest has in the last few years shifted to multistate(2–6) and analog(7–15) neuron systems. As for the two-state systems, storage capacities of networks with prescribed forms for their synaptic organization(2,3,7–13) as well as optimal capacities of multistate(4–6) or analog(16) perceptrons and their generalization abilities(14,15) have been studied.

In the present short communication, we wish to report a finding about attractor networks of multistate neurons that appears to have escaped the attention of previous investigators, namely that such systems, if designed to store an ensemble of multistate (gray-toned) patterns, exhibit a very

---

[1] Instituut voor Theoretische Fysica, K. U. Leuven, B-3001 Leuven, Belgium. Present address: Institut für Theoretische Physik, Universität Heidelberg, D-69120 Heidelberg, Germany.

[2] Institut für Theoretische Physik, Universität Heidelberg, D-69120 Heidelberg, Germany.

righ structure of coexisting metastable retrieval phases.(15) We analyze the nature of these phases and their respective domains of existence for both networks with Hebbian couplings and for networks with pseudoinverse-type couplings. Moreover, we discuss implications for optimal parameter settings.

The phenomenon can be studied in networks with continuous-time dynamics(11,12) or with asynchronous dynamics.(2) In the deterministic ($T=0$) limit, to which we restrict our attention here, both setups have, in fact, the same stationary states. To fix our conventions, we will therefore introduce only one of them, continuous-time dynamics.(7) It is described by

$$\frac{dU_i}{dt}=\sum_j J_{ij}V_j-U_i, \qquad V_j=g(\gamma U_j) \tag{1}$$

with i, j enumerating the neurons, J_{ij} designating synaptic couplings, V_j neural outputs, and U_j membrane potentials. The latter two quantities are related by the neutral input-output relation g, with a gain parameter γ fixing relative scales. For the sake of definiteness, we consider the problem of storing *multistate* patterns $\{\xi_i^\mu\}$, with

$$\xi_i^\mu \in \{\xi_k := -1+2(k-1)/(Q-1);\, k=1,\dots, Q\} \tag{2}$$

where $\mu=1,\dots,p$ enumerates the patterns. That is, the ξ_i^μ can assume Q equidistant values between -1 and $+1$ with, say, equal probability. A stable *representation* of such patterns in recursive neural nets can be accomplished by neurons with input-output relations $g(x)$ that approach a $(Q-1)$-step structure, with a step height $\Delta g=2/(Q-1)$ determined by the number of admissible steps and a step width Δx that can, in principle, be left open as an adjustable parameter determining the "average slope" $\Delta g/\Delta x$ of the gain function near $x=0$. Formally,

$$g(x)=\frac{1}{Q-1}\sum_{k=1}^{Q-1}\operatorname{sgn}(x-x_k) \tag{3}$$

with $x_k \in \{0, \pm\Delta x,\dots, \pm\frac{1}{2}(Q-2)\Delta x\}$ if Q is even, and $x_k \in \{\pm\frac{1}{2}\Delta x, \pm\frac{3}{2}\Delta x,\dots, \pm\frac{1}{2}(Q-2)\,\Delta x\}$ if Q is odd. Here, we adopt the convention that $g(x)$ has $\Delta g/\Delta x=1$ and that the neural response is given by $V_i=g(\gamma U_i)$, with an adjustable gain parameter γ. Using these conventions, we have opted for homogeneous pattern distributions and symmetric input-output relations with *uniform* step widths and step heights. It should, however, be remarked at the outset that most of the results presented below are *not* restricted to this maximally regular setup. Since a complete systematic exploration of the set of possible multistep input-output relations and pattern distributions is out of the question anyway, we have restricted our study to the homogeneous and symmetric situation, both for convenience and to keep the analytic structure of our equations as simple and transparent as possible.

Storage of multistate patterns as defined in (1) can be achieved, for instance, in networks with synapses of the Hebb–Hopfield form

$$J_{ij}=\frac{1}{N\Delta_0}\sum_{\mu=1}^{p}\xi_i^{\mu}\xi_j^{\mu}, \qquad i\neq j \tag{4}$$

or in networks endowed with so called pseudoinverse synapses,

$$J_{ij}=\frac{1}{N\Delta_0}\sum_{\mu,\nu=1}^{p}\xi_i^{\mu}(C^{-1})_{\mu\nu}\xi_j^{\nu}, \qquad i\neq j \tag{5}$$

with C denoting the correlation matrix of the pattern set, with elements

$$C_{\mu\nu}=\frac{1}{N\Delta_0}\sum_{i=1}^{N}\xi_i^{\mu}\xi_i^{\nu} \tag{6}$$

In (4)–(6), the ξ_i^{μ} take a wider range of values, as described by (2), and the normalization constant

$$\Delta_0=\langle(\xi_i^{\mu})^2\rangle=\frac{1}{3}\frac{Q+1}{Q-1}$$

is chosen to fix the J_{ij} scale in a Q-independent manner. Note that we have also chosen to normalize the correlation matrix (4) such that $C_{\mu\mu}=1$, independently of Q.

The fixed-point structure of the neutral dynamics—be it the asynchronous dynamics studied in ref. 2 or the continunous-time dynamics investigated in refs. 11 and 12—is given by the local minima of the energy function$^{(7,11,12)}$

$$\mathscr{H}_N(\mathbf{V})=-\frac{1}{2}\sum_{i,j=1}^{N}J_{ij}V_iV_j+\frac{1}{\gamma}\sum_{i=1}^{N}G(V_i) \tag{7}$$

where G is the integrated inverse input–output relation. The "inverse" of a multistep input–output relation can be defined by a limiting process, and has a multistep structure itself, with discontinuities at the plateau values $y_k=-1+2(k-1)/(Q-1)$, $k=1,\ldots,Q-1$ of g. This leads to $G(y_k)=\frac{1}{2}y_k^2+\text{const}$, where const $=0$ for odd Q and const $=-(Q-1)^{-2}$ if Q is even, with *linear* interpolation *between* these values. If restricted to the discrete configurations allowed by the limiting multistep input–output relation, the energy function (7) reduces to the Gathak–Sherrington Hamiltonian,(17) which was also used by Rieger,(2) albeit here with variable effective gain. Indeed, as we shall see below, by varying the gain, the phase structure of the model can be changed considerably.

As advocated before,(11,12) the fixed-point structure of the deterministic neural dynamics (1) is revealed by studying the zero-temperature phases of the free energy

$$f_N(\beta) = -(\beta N)^{-1} \log \mathrm{Tr}_{\mathrm{V}} \exp[-\beta \mathscr{H}_N(\mathrm{V})] \tag{8}$$

where Tr_{V} stands for a classical sum over states. The evaluation of (8) utilizes mean-field technique and the replica method to deal with the quenched randomness in the J_{ij} due to the randomness of the patterns. It follows the general scheme of Amit *et al.*(18) in case of the Hebbian synapses (4) and that of Kanter and Sompolinsky(19) in case of the network with pseudoinverse synapses (5), (6), with modifications in both cases to account for the nonbinary nature of the V_i. The solution is obtained by adapting the methods outlined in refs. 11 and 12 to the more complicated pattern statistics, and it provides a *macroscopic characterization* of the stable attractors of the deterministic dynamics (1) in terms of order parameters.

Before discussing the ensuing results, let us obtain an estimate of the capacity for *perfect* storage through a signal-to-noise approximation. For fixed γ, the step width of the input–output relation scales as $\Delta U \propto 2/(Q-1)$ with the number Q of output levels. Safe storage requires that the noise term generated by a pattern, found to be $\mathcal{O}((\alpha\Delta_0)^{1/2})$, with $\alpha = p/N$ denoting the storage level, must be small compared to the step width ΔU. For $\gamma = 1$, this gives $p_{\max}(Q) = 3p_{\max}(2)/(Q^2-1)$ for the number of patterns that can be stored without errors in a network of Q-state neurons, with a nonextensive $p_{\max}(2) = N/4 \ln N$. It exhibits the same Q^{-2} scaling of the large-Q storage capacity as obtained by Rieger(2) in the limit of extensive loading, $\alpha > 0$, and confirmed numerically in ref. 20.

Turning to questions other than the storage capacity, we note that in $Q > 2$ systems there exists the possibility of having a multiplicity of retrieval phases which do have an $\mathcal{O}(1)$ overlap

$$m = \frac{1}{N\Delta_0} \sum_i \xi_i^1 V_i \tag{9}$$

with, say, pattern 1 only, but which exhibit a variety of retrieval errors, characteristic for the different phases in question. These errors occur because at some nodes the V_i do *not* assume the value ξ_i^1 they should, but are instead systematically shifted to either larger or smaller values in the sense that—depending on the value of γ—either $|V_i| \geqslant |\xi_i^1|$ for *all* i, or $|V_i| \leqslant |\xi_i^1|$ for *all* i. The structure of these phases is most easily explained in the $\alpha = 0$ limit, where mean-field theory predicts[11,12] $V_i = \hat{V}(\xi_i^1) = g(\gamma m \xi_i^1)$ for the stable stationary states.

The simplest nontrivial multistate system to consider is the $Q = 3$ system, with $\xi_i^1 \in \{0, \pm 1\}$ and $g(x) = \frac{1}{2}[\operatorname{sgn}(x + 1/2) + \operatorname{sgn}(x - 1/2)]$. For this system, the only phase exhibiting retrieval errors (aside from the spin-glass phase; see Fig. 2 below) is a trivial null phase. It occurs when nodes with $\xi_i^1 = \pm 1$ have $V_i = 0$, because $\hat{V}(\pm 1) = g(\pm \gamma m) = 0$. Since $g(0) = 0$, we have $m = 0$ for this phase, and the picture is self-consistent.

Of greater complexity is the case $Q = 4$, for which $\xi_i^1 \in \{\pm 1/3, \pm 1\}$, so that $\Delta_0 = 5/9$, and $g(x) = \frac{1}{3}[\operatorname{sgn}(x + 2/3) + \operatorname{sgn}(x) + \operatorname{sgn}(x - 2/3)]$. In this system, two nontrivial phases with different types of retrieval error may be observed. The first type occurs when nodes with $\xi_i^1 = \pm 1$ are mapped onto $V_i = \pm 1/3$. This gives rise to a phase with $m = 2/5$. The second type of error arises when nodes with $\xi_i^1 = \pm 1/3$ are mapped onto $V_i = \pm 1$, entailing $m = 6/5$. In addition, there is of course the error-free phase with $V_i = \xi_i^1$, hence $m = 1$.

The γ ranges where these phases exist (at $\alpha = 0$) are also computed. For instance, as long as $\gamma m > 2$, one has $\hat{V}(\pm 1) = g(\pm \gamma m) = \pm 1$ *and* $\hat{V}(\pm 1/3) = g(\pm \gamma m/3) = \pm 1$, so that $m = 6/5$. This solution is thus self-consistent for $\gamma^{-1} < 3/5$. In a similar vein, for $2 > \gamma m > 2/3$ one has $\hat{V}(\xi_i^1) = g(\gamma m_i^1) = \xi_i^1$, hence $m = 1$. Thus $m = 1$ is observed in the range $1/2 < \gamma^{-1} < 3/2$. For $\gamma m < 2/3$, on the other hand, one has $\hat{V}(\pm 1) = \hat{V}(\pm 1/3) = \pm 1/3$, entailing $m = 2/5$. Such a reduced overlap will therefore be observed for all $\gamma^{-1} > 3/5$.

Three things may be learnt from these simple examples. First, there is a multiplicity of retrieval phases with retrieval m *not* close to 1. It can be shown that the energies of these phases are in general nondegenerate. In this (restricted) sense, all but one of these phases are *metastable* for given values of the system parameters. Second, transitions between these phases are first order, and there are strong hysteresis effects. The limiting γ values computed above denote points in parameter space where phases become unstable, and a kind of spinodal decomposition occurs. Since we are dealing with a deterministic dynamics, these instabilities are the points of interest, rather than degeneracies at which equilibrium first-order transitions could occur at $T > 0$. Third, null phases can occur only in systems with *odd* Q.

The same type of analysis can be carried out for other Q values. Our results for $Q = 3$, 4, 5, and 6 are summarized in Tables Ia–Id, in which the retrieval overlap m and the Edwards–Anderson order parameter $q_1 = (1/N) \sum_i V_i^2$ of various phases are shown along with the range of inverse-gain values where these phases exist. The complexity of the phase structure as well as that of the hysteresis pattern clearly increases quite dramatically with Q (see Fig. 1).

A relatively straightforward argument based on the graphical solution of the mean-field equations(11,12) in the $\alpha \to 0$ limit shows that the total number $\mathcal{N}(Q)$ of different retrieval phases at $\alpha = 0$ (excluding the null phase in the odd-Q case) is

$$\mathcal{N}(Q) = \tfrac{1}{4}Q^2 - \tfrac{1}{2}Q + \text{const}$$

Table I. Metastable Retrieval Phases at $\alpha = 0$ for Networks with $Q = 3, 4, 5, 6$[a]

m	q_1	$\gamma^{-1}_{\min}$	$\gamma^{-1}_{\max}$
	(a) Phases for $Q = 3$		
0	0	0	∞
1	2/3	0	2
	(b) Phases for $Q = 4$		
2/5	1/9	3/5	∞
1	5/9	1/2	3/2
6/5	1	0	3/5
	(c) Phases for $Q = 5$		
0	0	0	∞
2/5	1/10	4/5	8/5
3/5	1/5	4/5	6/5
1	1/2	2/3	4/3
6/5	4/5	0	4/5
	(d) Phases for $Q = 6$		
9/35	1/25	9/14	∞
19/35	11/75	57/70	19/14
5/7	19/75	25/28	15/14
1	7/15	3/4	5/4
41/35	17/25	41/70	123/140
43/35	59/75	43/140	53/70
9/7	1	0	9/28

[a] Liested are retrieval overlap m and Edwards–Anderson order parameter q_1 of the various phases, along with the ranges of inverse gain, $\gamma^{-1}_{\min}$ and $\gamma^{-1}_{\max}$, where these phases exist. Overlapping inverse-gain ranges of different phases indicate hysteresis.

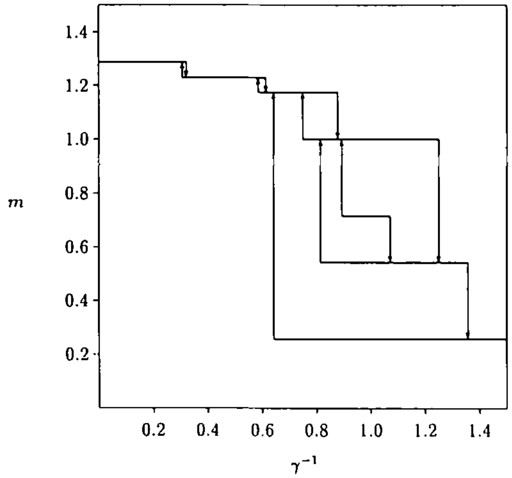

Fig. 1. Hysteresis pattern for the metastable retrieval phases of a $Q=6$ system at $\alpha=0$. Shown are overlaps of the various phases over the ranges of inverse gain γ^{-1} where they exist. Arrows indicate decay channels of the various phases. Note that the $m=5/7$ phase is not accessible through hysteresis effects simply by changing γ.

where const = 1 if Q is even and const = 1/4 if Q is odd. For some of these phases, the γ range where they exist is strictly zero, however. In the $Q=9$ system, for example, this happens for exactly one out of 16 phases, in the $Q=10$ system for two out of 21 phases. The total number of different retrieval phases just given is thus only an upper bound, but rather a tight one, for the number of phases that are actually observable—a number that scales as Q^2 for both odd and even Q. The γ ranges in which these phases *are* actually observed (at $\alpha=0$) are decreasing functions of Q for all but the lowest-m and the largest-m phase. To mention numbers just for the $m=1$ phase, we find that it exists for $\gamma^{-1}\in[1-1/(Q-2),\ 1+1/(Q-2)]$—an interval, which shrinks to zero as $Q\to\infty$. At extensive levels of loading, $\alpha>0$, this interval becomes smaller with increasing α, remaining symmetric about 1, so $\gamma=1$ appears to be the optimal gain for the system with Hebbian synapses. Note, however, that even at $\gamma=1$, phases with retrieval errors exist.

We find that all these metastable retrieval phases persist at sufficiently small values of extensive loading, i.e., for $\alpha>0$, as shown in Fig. 2. Moreover, other such phases, which do not exist at $\alpha=0$, may appear at extensive levels of loading, as exemplified for the $Q=4$ system in Fig. 2. Let us remark again that—qualitatively—our results carry over to input–output relations less symmetric than those explicitly considered above.

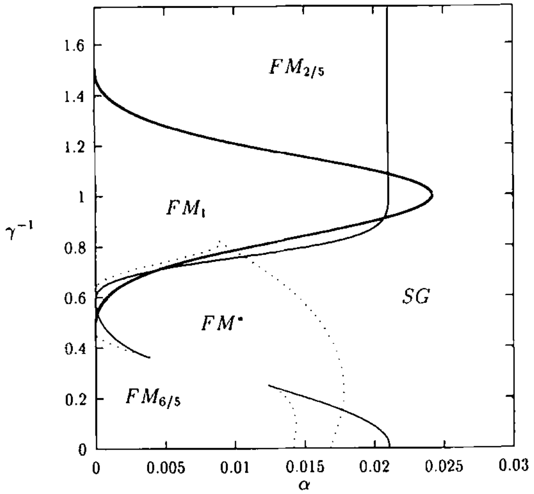

Fig. 2. Phase diagram of the $Q=4$ phase in the γ–α plane. Beyond the phases FM_m with $\alpha=0$ magnetizations $m=2/5$, $m=1$, and $m=6/5$ there is a fourth metastable retrieval phase—denoted FM* in the figure, with phase boundaries given as dotted lines—which exists only for $\alpha>0$. For intermediate α, this phase can be deformed continuously into the $FM_{6/5}$ phase; for small *and* for larger α, though, discontinuous transitions between these two phases occur, and two pairs of spinodals coalescing in critical endpoints at $(\alpha, \gamma^{-1}) \simeq (0.0038, 0.3615)$ and $(\alpha, \gamma^{-1}) \simeq (0.0124, 0.2510)$, respectively, are shown in the figure. A spin-glass phase (SG) with $m=0$ and $q_1 \neq 0$ coexists with the FM_m phases at all $\alpha>0$. For sufficiently large α, it is the only phase, though.

The metastable retrieval phases act as attractors for the neural dynamics, and states with the indicated values for the overlaps have been observed in simulations that tried to measure the basins of attraction for pure pattern states(20) when initial configurations were sufficiently distorted version of the stored patterns.

Very much the same phenomena are observed in a network coupled via the pseudoinverse synapses (5), (6). For these, we find that the mean-field equations(11,12) describing the phase structure of the network have a stable retrieval solution characterized by stationary membrane potential *without* a noise contribution in a signal-to-noise decomposition, *even* in the case of extensive levels of loading, $\alpha > 0$. It leads to a self-consistency equation for the corresponding retrievel overlap given by

$$m = \langle (\xi^1/\Delta_0)\, g((1-\alpha)\gamma m \xi^1) \rangle \tag{10}$$

where the angular brackets denote an average over the ξ^1 distribution. However, this "noiseless" solution exists *only* for the *true* retrieval phase, having $m=1$. For example, for $Q=4$, this phase can be observed in the range $2>\gamma(1-\alpha)>2/3$. A similar replacement $\gamma\rightarrow\gamma(1-\alpha)$ applies to the range of γ values for which the $m=1$ phase is observed in systems with orther Q values (see Table I).

The other metastable retrieval phases exist in the system with pseudoinverse synapses, too. But for these, the stationary membrane potentials do have a noise contribution which destabilizes the $m\neq 1$ retrieval states at loading levels comparable to those in the system with Hebbian synapses. Hence, for moderate loading levels they no longer compete with the true $m=1$ retrieval phase, and we are left with this error-free phase alone. The fact that noise contributions to the membrane potential are absent *only* in the $m=1$ phase follows from the observation that the influence of these contributions is proportional to $q_1-\Delta_0 m^2$, and that (9) entails $q_1\geqslant\Delta_0 m^2$ by a Schwarz inequality, which becomes an equality *if and only if* $V_i\equiv\xi_i^1$, thus $m=1$.

In summary, we have demonstrated that networks of multistate neurons exhibit many metastable retreival phases if one uses simple gain functions with variable gain parameters. The phenomenon is observed in networks with both Hebbian synapses and pseudoinverse couplings. The number of such phases was shown to scale with the number Q of gray levels as Q^2. Transitions between these phases are usually first-order, and there are strong hysteresis effects. The γ range in which the $m=1$ phase is observed was shown to be decreasing to zero as $Q\rightarrow\infty$. Optimal gain parameters were found to be $\gamma=1$ in the system with Hebbian couplings and $\gamma=(1-\alpha)^{-1}$ in the system with pseudoinverse-type couplings. One should be aware that even at optimal gain, basins of attraction are strongly modified due to coexisting metastable retrieval phases. Qualitatively the same phenomena will be observed for systems less regular than those explicitly considered in the present paper.

Let us finally remark that very recently the thermodynamics of these systems has been the subject of an independent investigation by Bollé *et al.*(21)

ACKNOWLEDGMENTS

Part of this work was done while one of us (S.B.) was visiting the Instituut voor Theoretische Fysica at the K. U. Leuven. It is a pleasure to thank D. Bollé for the invitation and the neural networks group at Leuven for their hospitality and for illuminating discussions. Thanks are also due to K. Müller for communicating his simulation results prior to publication, and for useful comments.

REFERENCES

1. D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, 1989).
2. H. Rieger, *J. Phys. A* **23**:L1273 (1990).
3. G. A. Kohring, *J. Stat. Phys.* **62**:563 (1991).
4. S. Mertens, H. M. Köhler, and S. Bös, *J. Phys. A* **24**:4941 (1991).
5. D. Bollé, P. Dupont, and J. van Mourik, *Europhys. Lett.* **15**:893 (1991).
6. F. Gerl, K. Bauer, and U. Krey, *Z. Phys. B* **88**:339 (1992).
7. J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **81**:3088 (1984).
8. C. M. Marcus and R. M. Westervelt, *Phys. Rev. A* **40**:501 (1989); C. M. Marcus, F. M. Waugh, and R. M. Westervelt, *Phys. Rev. A* **41**:3355 (1990).
9. A. Treves, *J. Phys. A* **23**:2631 (1990); *Phys. Rev. A* **42**:2418 (1990).
10. M. Shiino and T. Fukai, *J. Phys. A* **23**:L1009 (1990); *J. Phys. A* **25**:L375 (1992).
11. R. Kühn, in *Statistical Mechanics of Neural Networks*, L. Garrido, ed. (Springer-Verlag, Berlin, 1990), p. 19; R. Kühn, S. Bös, and J. L. van Hemmen, *Phys. Rev. A* **43**:2084 (1991).
12. R. Kühn and S. Bös, *J. Phys. A* **26**:831 (1993).
13. D. Bollé, B. Vinck, and V. A. Zagrebnov, *J. Stat. Phys.* **70**:1099 (1993); D. Bollé, G. M. Shim, B. Vinck, and V. A. Zagrebnov, K. U. Leuven preprint (1993).
14. S. Bös, W. Kinzel, and M. Opper, *Phys. Rev. E* **47**:1384 (1993).
15. S. Bös, Ph.D. thesis, Universität Giessen (1992), unpublished.
16. D. Bollé, R. Kühn, and J. van Mourik, *J. Phys. A* **26**:3149 (1993).
17. S. K. Ghatak and D. Sherrington, *J. Phys. C* **10**:3149 (1977).
18. D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Ann. Phys.* (N.Y.) **173**:30 (1987).
19. I. Kanter and H. Sompolinsky, *Phys. Rev. A* **35**:380 (1987).
20. T. Stiefvater and K. R. Müller, *J. Phys. A* **25**:5919 (1992).
21. D. Bollé, H. Rieger, and G. M. Shim, K. U. Leuven preprint (1993).