

**Stochastic resonance in pattern recognition by a holographic neuron model**R. Stoop,<sup>1,2,\*</sup> J. Buchli,<sup>1</sup> G. Keller,<sup>1</sup> and W.-H. Steeb<sup>2,†</sup><sup>1</sup>*Institut für Neuroinformatik, Swiss Federal Institute of Technology ETHZ, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*<sup>2</sup>*University of Applied Technical Sciences of Northwestern Switzerland, Olten, Switzerland*

(Received 6 November 2002; published 30 June 2003)

The recognition rate of holographic neural synapses, performing a pattern recognition task, is significantly higher when applied to natural, rather than artificial, images. This shortcoming of artificial images can be largely compensated for, if noise is added to the input pattern. The effect is the result of a trade-off between optimal representation of the stimulus (for which noise is favorable) and keeping as much as possible of the stimulus-specific information (for which noise is detrimental). The observed mechanism may play a prominent role for simple biological sensors.

DOI: 10.1103/PhysRevE.67.061918

PACS number(s): 87.16.Ac, 02.50.Ey, 42.30.Sy, 07.05.Mh

**I. INTRODUCTION**

The traditional view on noise in information processing is that the more noise occurs in a process, the worse the process performance is. Based on the observation that noise is ubiquitous in natural systems, whose processing capabilities are still unchallenged by artificial systems, this view has recently changed. Gradually, the idea has emerged that noise actually could be used to *improve* the efficiency of computations. As the paradigm of such a phenomenon, the principle of stochastic resonance has been identified. In its early days, the phenomenon of stochastic resonance was strongly tied to the existence of a periodic weak subthreshold oscillation. When, to this signal, relatively large-scale noise was added, the system was able to cross the threshold, and an improvement of the signal was obtained. This concept was first discussed in the context of climate dynamics [1], and then found in electronic circuits [2]. Later, the phenomenon was proposed [3] and verified in lasers [4,5], and finally found in magnetic systems [6], in neurons [7], and in chemical reactions [8]. More recently, numerous examples were found in the analysis of biological sensors. As an impressive example, the crayfish has been shown to use stochastic resonance to catch its prey (e.g., Refs. [9,10]). In its most general form, stochastic resonance can be defined as a nonlinear cooperative effect, whereby the addition of a random process, or noise, to a weak signal, results in an enhanced response of the system (thus dropping the condition of periodicity). Investigations on the use of noise in the context of signal processing are of great technological importance. Miniaturization of computer chips naturally generates conditions, where the (thermal) noise is of the order of the signal. For signal processing, the cortex in many respects still is the most efficient device, and it operates at conditions where the noise level is comparable to the level of the signal. This motivates the expectation that the scientific focus will shift from noiseless computation at high signal power to computation at high levels of noise and low signal power.

In our paper, we report on the observation of a stochastic resonance effect that occurs in holographic neural synapses (for short: holographs) during pattern recognition, providing an example of stochastic resonance without underlying periodicity. Holographs are part of the family of analog, correlation-based, associative, stimulus-response memories, where information is mapped onto the phase orientation of complex numbers (operating, however, differently from standard connectionist models). The holographic method [11–14] is of interest in itself as it exhibits some remarkable efficiency characteristics. Unfortunately, and in spite of the long tradition of work on closely related approaches in optical holography [15–17], the method seems to have lacked widespread scientific interest. Holographs have been shown to be effective for associative memory tasks, generalization, and pattern recognition with changeable attention [11–14]. More specifically, investigations have shown [13] that efficient learning of arbitrary relationships between input and output with no constraints on topology or separability, high encoding densities, robustness with respect to low numerical resolution, good saturation, generalization and classification properties, fast learning rates and low steady-state error rates are characteristics of the method. From extended studies, it has been observed that the performance of the holograph depends in a surprising way on the statistical properties of the input data [12,13], which can be condensed in an asymmetry index. The lower this index, the better the performance of the holograph. In particular, it has been found [11,13] that there is an important distinction between artificial and natural images, since the latter tend to have lower asymmetry indices, implying that learning of artificial images is more difficult. We will show that for this class, a dramatic performance improvement can be obtained if noise is added to the input signal. This observation of stochastic resonance is the main result of the paper. Understanding this observation sheds some light on how biological signal processing could successfully operate in noisy environments.

**II. HOLOGRAPH SETUP**

Given a family of input patterns  $\mathcal{S}$ , for each pattern  $s$  an input vector  $\mathbf{S}$  is derived, and related to a desired response vector  $\mathbf{R}$ , forming in this way an association pair. The input

\*Email address: ruedi@ini.phys.ethz.ch

†Present address: Rand Afrikaans University, RSA-2000 Johannesburg, Republic of South Africa.

vector and the response vector entries are composed of complex numbers, i.e.,  $\mathbf{S}=(S_1, \dots, S_n)$ , with  $S_j=\lambda_j e^{I\theta_j}$ ,  $j \in \{1, \dots, n\}$ , where  $n$ , e.g., denotes the number of pixels in the input picture, and  $I$  denotes the imaginary unit. Similarly, the output vector has the form  $\mathbf{R}=(R_1, \dots, R_m)$ , with  $R_j = \gamma_j e^{I\phi_j}$ ,  $j \in \{1, \dots, m\}$ . In a more general formulation,  $S_j$  and  $R_j$  can be generalized as multidimensional complex numbers [12,13]. The exact form of the coding may depend on the typical input pattern. For any coding, the essential information is captured in the phase, whereas the associated modulus may be used as an attention parameter in the input and to express the confidence level of the output. In our investigation, this feature will remain unexplored: the moduli will be set to unity.

The heart of the holograph is an  $m \times n$  matrix  $\mathbf{X}$ , with arbitrarily chosen complex initial entries. During learning, presented patterns  $\mathbf{s} \in \mathcal{S}$  update the matrix iteratively as

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \bar{\mathbf{S}}_{a(i)}^T \cdot \left( \mathbf{R}_{a(i)} - \frac{1}{c} \mathbf{S}_{a(i)} \cdot \mathbf{X}_i \right), \quad (1)$$

where  $c = \sum_{i=1}^n \lambda_i$  is an input pattern specific normalization constant, and where index  $a(i) \in \{1, \dots, s = |\mathcal{S}|\}$  indicates the pattern presented at step  $i$ . If this iterative scheme converges, the ‘‘relation’’ between  $\mathbf{S}$  and  $\mathbf{R}$  is stored in  $\mathbf{X}_\infty$ . From Eq. (1), with the help of the fixed point, the correspondence between input and response patterns is given by

$$\{\mathbf{R}_l\} = \left\{ \frac{1}{c} \mathbf{S}_l \cdot \mathbf{X}_\infty \right\}, \quad l \in \{1, \dots, s\}. \quad (2)$$

The algorithmic complexity of the process is  $O(nm)$ . To implement  $s$  associations  $\{\mathbf{s}_l, \mathbf{r}_l\}_{l=1, \dots, s}$  between stimulus and classification (in their complex vector representations, this relationship is expressed as  $\{\mathbf{S}_l, \mathbf{R}_l\}_{l=1, \dots, s}$ ), thus  $O(nms)$  operations are needed, and the storage requirement is of the order of  $mn$ . It is of importance to note that the storage space does not grow, if more patterns are to be learned, and the time needed for the learning process only grows linearly with  $s$ . With growing  $s$ , the holograph may enter a saturation region, leading to a decreased performance. To avoid this, Khan [14] proposes  $L = s/n < 0.08$ . In our investigations, we operate at  $L = 0.0056$ , which is way below the proposed saturation threshold.

### III. HOLOGRAPHIC PATTERN RECOGNITION

When an object, e.g., a picture  $\mathbf{s} \in \mathcal{S}$ , is to be processed by the holograph, it first is encoded into complex numbers by the coding function  $m_s: \mathbf{s} \rightarrow \mathbf{S}$ . At the end of the process, a decoding function  $m_r$  will convert the complex-valued response  $\mathbf{R}$  to a real-valued response (or feature) vector,  $m_r: \mathbf{R} \rightarrow \tilde{\mathbf{r}} \in \mathbb{R}^m$ . The particular response  $\tilde{\mathbf{r}}$  will then be compared to a set of desired responses  $\mathcal{R} = \{\mathbf{r}_1, \dots, \mathbf{r}_k\}$ . During the pattern recognition process, association pairs are determined as  $\{\mathbf{s}_i, \mathbf{r}_j\}$ , where  $\mathbf{r}_j$  is the element of  $\mathcal{R}$  with the minimal distance to  $\tilde{\mathbf{r}}$ . To evaluate the distances in the space of output vectors, we use the Euclidean distance  $\|\cdot\|$ , normalized by the dimension of the output-vector space,  $m$ .

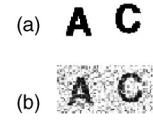


FIG. 1. (a) Artificial gray-scale image. (b) Artificial image with noise (mean  $\mu=0$  and  $\sigma^2=0.1$ ). (c) Natural gray-scale image.

To measure the performance, we determine the ratio between the number of correct to the number of total associations, called the (statistical, nontemporal) recognition rate  $r$ . The errors  $d(\mathbf{r}_j, \tilde{\mathbf{r}}_j)$  between generated and desired responses yield another measure of performance. However, it is well-known that the correlation between the recognition rate  $r$  and  $\sum_{j=1}^s d(\mathbf{r}_j, \tilde{\mathbf{r}}_j)$  is generally not too strong.

For our experiments, we used gray-level pictures (see Fig. 1), represented as matrices  $G=(g_{ij})$ , where  $g_{ij}$  denotes the gray-scale intensity value in the range  $\{0, \frac{1}{256}, \dots, \frac{255}{256}\}$ , at location  $i, j$ . From the matrices, via concatenation of the rows, the input vectors are obtained. When mapping their entries to complex numbers, it is desirable to keep the latter away from the target phase-space boundaries  $\{0, 2\pi\}$ . Otherwise, by small recall errors or noise, points may be pushed over the boundary [12], which may have devastating effects (e.g., changing very small brightness differences into maximal brightness differences). To prevent this, all values were shifted by  $\frac{1}{512}$ . To the shifted values  $\hat{g}_l$ ,  $l \in \{1, \dots, 256\}$ , optionally noise  $z$  from a clipped Gaussian random variable was added, before they were mapped onto the complex domain according to  $m_s(g_j): \theta_j = 2\pi(\hat{g}_j + z_j)$ . The desired output vectors  $\mathbf{r}$  were composed of uniformly distributed random numbers from the unit interval. Similarly to the input vectors, the output vectors are mapped into the complex domain by the map  $m_r(\phi_j): w_j = \phi_j/2\pi$ . Of course, other, more intrinsically application-related output vectors could be chosen. They should, however, be well separated from one another. The higher the dimension  $m$  of  $\mathbf{r}$ , the easier it is to meet this condition. The linear dependence of the computation time on  $m$ , however, makes a good selection of  $m$  worthwhile.

### IV. NATURAL VERSUS ARTIFICIAL IMAGES

We performed four sets of experiments, in which the recognition rates  $r$  were computed after the presentation of  $s$

TABLE I. Asymmetry index  $A = |\sum_i^n S_i / \sum_i^n \lambda_i|$ , from natural, noise-free, and noisy artificial images.

Pattern No.	A		
	Natural	$\sigma_l^2=0$	$\sigma_l^2=0.1$
1	0.94	0.92	0.56
2	0.81	0.91	0.57
3	0.39	0.91	0.60
4	0.60	0.91	0.55
5	0.28	0.91	0.56
6	0.58	0.91	0.56
7	0.54	0.91	0.56
8	0.59	0.91	0.57
9	0.19	0.91	0.57
10	0.73	0.91	0.56

input vectors. It will be indicated whether each input vector was presented exactly once, or, when the pattern was selected by random, whether repeated occurrence was tolerated. In the latter case, we chose the association index  $a(i)$  from a uniformly distributed discrete random variable in  $\{1, \dots, s\}$ . In this way, over a long learning history, every association pair appeared with equal probability. During all experiments, the boundary distance was kept at  $\epsilon=0.05$ , the noise had mean  $\mu=0$  and the elements of the matrix  $X$  were initially set to 0.

In the first experiment, associations between windows from natural photographic pictures [see Fig. 1(c)] and a set of random vectors of dimension  $m=40$  were learned ( $s=10$  association pairs). From each input picture, an arbitrary  $30 \times 60$  pixel-sized window was selected. This leads to an input-vector length of  $n=1800$ , which is equal to the size of the artificial images used. In Ref. [14] it was proposed to choose an asymmetry  $A < A_0 = 0.6$ . The asymmetries of our pictures massively violate this condition, see Table I. Nevertheless, the results shown in Figs. 2 and 3 (curves labeled by  $n$ ) confirm the earlier reported excellent holograph performance. Even for sets where more than half of the images failed to satisfy the condition massively, we found fast convergence to the optimal recognition rate of 1 within 4 training epochs of random input pattern selection (where an epoch denotes one pass through the whole set of input vectors).

In the second experiment, the photographic windows were replaced by pictures of letters of the same size, see Figs. 1(a) and 1(b). Quite astonishingly, on this set of pictures, the holograph failed to achieve a comparable performance (see Figs. 2 and 3, curves labeled by  $a$ ). The recognition rate  $r$  was found to fluctuate heavily, with a median now lying at 0.6.

**V. NOISE ENHANCES ARTIFICIAL IMAGES PERFORMANCE**

A natural assumption is that if the asymmetry index of artificial pictures can be reduced, this will improve the holograph performance. To achieve this, we added noise at sev-



FIG. 2. Transient behavior of the recognition rate  $r$  when images are associated with random output vectors of length  $m=40$ . Curves  $a-d$ : artificial images, gliding averages for  $\lambda=0.2$ . Noise combinations  $(\sigma_l^2, \sigma_r^2)$  as follows:  $a$ ,  $(0, 0)$ ;  $b$ ,  $(0, 0.05)$ ;  $c$ ,  $(0.1, 0.05)$ ;  $d$ ,  $(0.1, 0)$ . Curve  $n$ : natural images, where noise is added only during recall ( $\sigma_r^2=0.05$ ). Dashed curve: gliding average for  $\lambda=0.2$ .

eral stages of the process. When noise is added *before* processing the input patterns [Fig. 1(b)], this drastically improves the holograph performance, see Figs. 2 and 3 (curves  $c$  and  $d$ ). In these cases, the median of  $r$  increased from initially 0.5 to 0.75, where in some samples even a perfect recognition rate of 1 was observed. When also noise is added during recall, this has little effect. The immunity towards small additive noise in the recall step is found to increase with the length  $m$  of the response vector. When noise is added exclusively during recall, the average recognition rate increases only slightly. To understand these effects in detail, the recognition rates were explored for noise added during the learning (noise variances  $\sigma_l^2$ ) and during recall (noise variances  $\sigma_r^2$ ). From the numerical evaluations we generated a plot of the recognition rate over the parameter space  $\sigma_l^2 \times \sigma_r^2 = [0, 0.5] \times [0, 0.5]$ , using a resolution of  $\delta=0.05$  and  $s=9$  association pairs. At the upper bound of the interval, letters are hardly recognizable by the eye. To keep the computation time on an affordable level, the output-vector length was held fixed at  $m=5$ . For every combination of  $\sigma_l^2$  and  $\sigma_r^2$ , 32 experiments comprising  $n=30$  training epochs were performed, and the averaged recognition rate  $r$  was calculated as a function of  $\sigma_l^2, \sigma_r^2$  and training epoch  $\in \{1, \dots, n\}$ . The results indicate that the best asymptotic performance is obtained from a combination of  $\sigma_l^2=0.1$  with  $\sigma_r^2=0.05$  [Fig. 4(a)]. This is the amount of noise used for

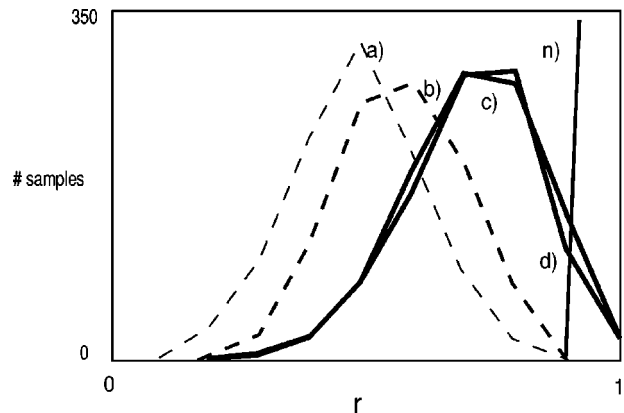


FIG. 3. Recognition rate histograms corresponding to Fig. 2, based on 1000 training epochs.

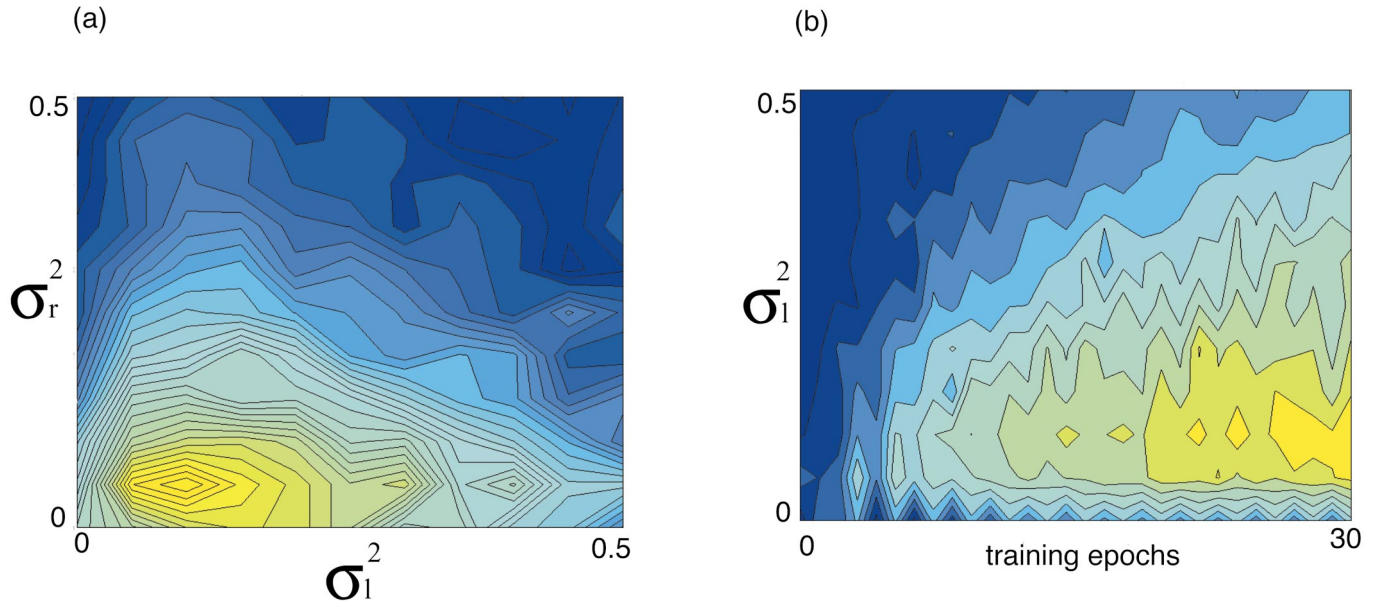


FIG. 4. (Color) Gray-level letter association to random output vectors of length  $m=5$ , based on 30 training epochs: (a) Average recognition rate as a function of  $\sigma_l^2$  and  $\sigma_r^2$ . Optimal combination:  $(\sigma_l^2, \sigma_r^2) = (0.1, 0.05)$ . (b) Average recognition rate evolution, as a function of  $\sigma_l^2$  (keeping  $\sigma_r^2 = 0.05$  fixed).

Figs. 2 and 3. Figure 4(b) shows how this performance enhancement is acquired during learning. In Fig. 4(c), sections through Fig. 4(a) along the axes  $\sigma_l^2 = 0.1$  and  $\sigma_r^2 = 0.05$ , respectively, are reported. The peaks at a nonzero level of noise are clearly visible. By performing analogous exhaustive experiments, the dependence of  $r$  on  $m$  was investigated. In Fig. 5, the sections using values of  $m = 1, 10, 20$  evidence that the higher the choice of  $m$ , the better the performance. In our experiments, random-generated desired response patterns were used. Further investigations have shown that a bad response vector choice can introduce effects that are of the order of the influence of the noise.

VI. DISCUSSION

How are these findings related with the known principles of stochastic resonance? As was mentioned above, the classical examples of stochastic resonance are connected with (mostly periodic) subthreshold oscillations. Our investigations will show that the reported effect does neither belong to this class nor to the class of noise-enhanced pattern recognition methods that are based on quantization improvement [19].

To investigate its nature, the origins of the problem with artificial pictures need to be analyzed. Holographic processing is based on a summation of column vectors in the complex domain. It can be observed that during the iterative formation of the holograph, single elements can display (1) convergent, (2) limit cycle (only possible for repeated non-random sequential learning), or (3) chaotic behavior (for the verification of this property, time series methods were used [18]). An efficient holograph is characterized by a convergent correlation matrix (which will provide the stability of the procedure) and a decent representation of the information

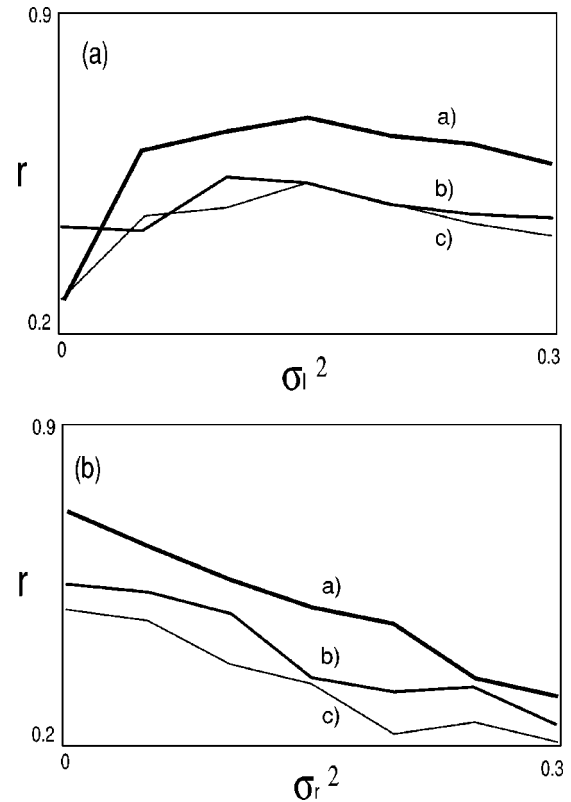


FIG. 5. Dependence of the recognition rate  $r$  on  $m$ , for (curve a)  $m=20$ , (curve b)  $m=10$ , (curve c)  $m=1$ , (a) as a function of  $\sigma_l^2$ , keeping  $\sigma_r^2 = 0.05$  fixed, (b) as a function of  $\sigma_r^2$ , keeping  $\sigma_l^2 = 0.1$  fixed. With growing  $m$ , the performance increases.

to be processed. The latter is the key issue in understanding the mechanism of stochastic resonance that we deal with.

For the following, we will maintain for generality the moduli in the complex representations. For simplicity, we will assume that the response vectors be one dimensional [12]. After the encoding using different stimuli at times  $t$ , the correlation matrix has the form

$$\mathbf{X} = \left\{ \sum_{t=1, \dots, T} l_{1,t} \gamma_t e^{I(\phi_t - \theta_{1,t})}, \dots, \sum_{t=1, \dots, T} l_{n,t} \gamma_t e^{I(\phi_t - \theta_{n,t})} \right\}, \quad (3)$$

where  $n$  is the number of elements of the stimulus field encoded as  $\{l_{i,t} e^{I\theta_{i,t}}\}_{i=1, \dots, n}$ , and  $\{\gamma_t e^{I\phi_t}\}$  is the response vector. From a new stimulus  $S^*$ , the response

$$\begin{aligned} \mathbf{R} &= 1/c \sum_{t=1, \dots, T} \gamma_t e^{I\phi_t} \sum_{k=1, \dots, n} l_k^* l_{k,t} e^{I(\theta_k^* - \theta_{k,t})} \\ &=: 1/c (L_1 e^{I\phi_1^*} + L_2 e^{I\phi_2^*} + \dots) \end{aligned} \quad (4)$$

is generated, where  $L_i, i=1, \dots, T$ , are confidence levels proportional to the degree to which the new stimulus falls close to a stimulus previously encoded at time  $t$ :

$$\begin{aligned} L_t e^{I\phi_t^*} &= \gamma_t e^{I\phi_t} \sum_{k=1, \dots, n} l_k^* l_{k,t} e^{I(\theta_k^* - \theta_{k,t})} \\ &= \gamma_t e^{I\phi_t^*} \left[ \left( \sum_{k=1, \dots, n} l_k^* l_{k,t} \cos(\theta_k^* - \theta_{k,t}) \right)^2 \right. \\ &\quad \left. + \left( \sum_{k=1, \dots, n} l_k^* l_{k,t} \sin(\theta_k^* - \theta_{k,t}) \right)^2 \right]^{1/2}, \end{aligned} \quad (5)$$

where  $l_k^*, \theta_k^*$  are the input data characteristics,  $l_{k,t}, \theta_{k,t}$  are the previously recorded input data, and

$$\begin{aligned} \phi_t^* &= \arctan \left[ \sum_{k=1, \dots, n} l_k^* l_{k,t} \sin(\theta_k^* - \theta_{k,t} + \phi_t) / \right. \\ &\quad \left. \sum_{k=1, \dots, n} l_k^* l_{k,t} \cos(\theta_k^* - \theta_{k,t} + \phi_t) \right]. \end{aligned}$$

The above expression shows that the largest contribution comes from the closest stimulus in the past. The more equidistributed each of the input vector sets are, the higher the discrimination [12].

Artificial pictures tend to generate stimulus vectors that lack equidistribution, as they contain large areas of identical elements. Moreover, these regions may coincide throughout the set of stimuli (for example, all letter backgrounds are white). Let us take two stimuli  $s_1, s_2$  and call a learning trial when  $s_1$  is first encoded and  $s_2$  is decoded, and later the same procedure with interchanged roles of  $s_1, s_2$  is applied. If the two stimuli are different, by the learning trial a net contribution

$$2 \sin\left(\frac{\phi_1 - \phi_2}{2}\right) 2 \sin\left(\frac{\theta_{k,1} - \theta_{k,2}}{2}\right) \quad (7)$$

adds to the correlation matrix element [12], which drives the two responses mutually away from their response average. For stimuli that are identical over a stimulus subfield, the contributions cancel, so that the generated response converges to the average of the responses. Although the individual magnitudes of the elements over these fields are attenuated in this way, a large number of such elements, nevertheless, may mask the salient stimulus features and confound the recognition process. If noise is added, formerly identical pieces become different. This, however, only works if the information content of the different stimuli is not completely destroyed.

The degree of equidistribution of a field can be expressed by the asymmetry index, defined as the average complex vector length over the field. As in our numerical calculations, the moduli are set equal to 1, the maximal asymmetry index is 1, whereas a fully symmetric, i.e., optimal, stimulus field yields zero asymmetry. Starting from large asymmetries, upon addition of noise, we were able to arrive below the critical asymmetry  $A_0$  (see Table I). This enhanced holographic pattern recognition substantially, however, without attaining the performance of natural images. The explanation of this fact is that artificial images contain less information than natural images, especially, if the spatial distribution also is taken into account. As a consequence, in the source(image)-channel(encoding)-receiver(holograph) picture of the holograph, the entropy of the source is smallest for artificial images. Adding noise to the source makes the received information unreliable, decreasing the mutual information on which the holographs' learning is based. In order to achieve optimal performance, an encoding of high symmetry (which can be achieved by the addition of noise) and relatively intact image structures are required. These requirements, however, are contradictory. This, ultimately, is the origin of the reported stochastic resonance effect.

As a consequence, the nature of the effect that we deal with is distinct from the previously found stochastic resonance principles in the field (e.g., noise-induced learning enhancement, as encountered in simulated annealing [20], or recently found stochastic resonance in associative memory approaches [22]). Rather, its appearance is strongly tied to the representation of the neuron by means of complex numbers, involving intrinsically notions of modulus and phases [24]. This, notably, is the case in the field of hearing and speech recognition, where it was observed that Gaussian noise added on the peripheral level enhances discrimination in hearing [21,23], forming a class of stochastic resonance effects of their own.

As symmetric encoding can be established by the inclusion of higher orders of correlations [12], we speculate that in biology, stochastic resonance will be beneficial in conditions where the sensors are too simple to provide higher-order correlations. Presently, we are investigating related questions.

- [1] R. Benzi, S. Suter, and A. Vulpiani, *J. Phys. A* **14**, L453 (1981).
- [2] S. Fauve and F. Heslot, *Phys. Lett.* **97A**, 5 (1983).
- [3] K. Wiesenfeld and B. McNamara, *Phys. Rev. Lett.* **55**, 13 (1985).
- [4] B. Derighetti, M. Ravani, R. Stoop, P.F. Meier, E. Brun, and R. Badii, *Phys. Rev. Lett.* **55**, 1746 (1985).
- [5] B. McNamara, K. Wiesenfeld, and R. Roy, *Phys. Rev. Lett.* **60**, 2626 (1988).
- [6] L. Gammaitoni, M. Martinelli, L. Pardi, and S. Santucci, *Phys. Rev. Lett.* **67**, 1799 (1991).
- [7] A. Bulsara, E. Jacobs, T. Zhou, F. Moss, and L. Kiss, *J. Theor. Biol.* **152**, 531 (1991).
- [8] M. Dykman, T. Horita, and J. Ross, *J. Chem. Phys.* **103**, 966 (1995).
- [9] K. Wiesenfeld and F. Moss, *Nature (London)* **373**, 33 (1995).
- [10] S. Bahar, A. Neiman, L.A. Wilkens, and F. Moss, *Phys. Rev. E* **65**, 050901 (2002).
- [11] J.G. Sutherland, *Int. J. Neural Syst.* **1**, 259 (1990).
- [12] J.G. Sutherland, in *Fuzzy, Holographic and Parallel Intelligence*, edited by B. Sucek (Wiley, New York, 1992), pp. 7–92.
- [13] J.I. Khan, Ph.D. thesis, University of Hawaii, 1995.
- [14] J.I. Khan, *IEEE Trans. Neural Netw.* **9**, 389 (1998).
- [15] D. Psaltis, D. Brady, X.G. Gu, and S. Lin, *Nature (London)* **343**, 325 (1990).
- [16] H.-Y.S. Li, Y. Qiao, and D. Psaltis, *Appl. Opt.* **26**, 5026 (1993).
- [17] A. Pu, R. Denkewalter, and D. Psaltis, *Opt. Eng.* **10**, 2737 (1997).
- [18] J. Peinke, J. Parisi, O.E. Roessler, and R. Stoop, *Encounter with Chaos* (Springer, Berlin, 1992).
- [19] E. Simonotto, M. Riani, C. Seife, M. Roberts, J. Twitty, and F. Moss, *Phys. Rev. Lett.* **78**, 1186 (1997).
- [20] B. Müller, J. Reinhardt, and M.T. Strickland, *Neural Networks* (Springer, Berlin, 1995).
- [21] R.P. Morse and E.F. Evans, *Nat. Med.* **2**, 928 (1996).
- [22] Z. Tan and M. Ali, *Int. J. Mod. Phys. C* **11**, 1585 (2001).
- [23] M.T. Moskowicz and B.W. Dickinson, *Proceedings of the International Symposium on Circuits and Systems, 2002* (unpublished), pp. 855–858.
- [24] Originally [12], holographic neurons emerged as an alternative model of biological neurons, where phases play an essential role (in contrast to the usual approach where phase relationships are neglected). Recent findings from biological neurons indicate that phase relationships may indeed be more important than previously expected. See, e.g., R. Stoop, K. Schindler, and L.A. Bunimovich, *Neurosci. Res.* **36**, 81 (2000).