Parametrical neural network based on the four-wave mixing process

B.V. Kryzhanovsky\textsuperscript{a}, L.B. Litinskii\textsuperscript{a,\dagger,1,2}, A. Fonarev\textsuperscript{b}

\textsuperscript{a} Institute of Optical Neural Technologies RAS, 44/2 Vavilov Str., Moscow 119333, Russia
\textsuperscript{b} Department of Eng. Sci. and Phys., The College of Staten Island, CUNY, SI, New York 10314, USA

Abstract

We develop a formalism allowing us to describe operating of a network based on the parametrical four-wave mixing process that is well-known in nonlinear optics. It is shown that the storage capacity of such a network is higher compared with the Potts-glass neural networks.

© 2003 Elsevier Science B.V. All rights reserved.

PACS: 87.10.+e

Keywords: Associative memory; Multi-state neuron; Nonlinear optics

0. Introduction

In Refs. [1,2] a network based on the well-known in nonlinear optics the parametrical four-wave mixing process (FWM) [3] was examined. Such a network is able to hold and handle information that is encoded to be of the form of the phase-frequency modulation. In the network, the signals propagate along interconnections in the form of quasi-monochromatic pulses at \( q \) different frequencies \( \{\omega_r\}_q \equiv \{\omega_1, \omega_2, \ldots, \omega_q\} \). The model is based on a parametrical neuron that is a cubic nonlinear element capable of transforming and generating frequencies in the parametrical FWM processes \( \omega_i - \omega_j + \omega_k \rightarrow \{\omega_r\}_q \). Schematically, this model of a neuron can be assumed to be a device composed of a summator of input signals, a set of \( q \) ideal frequency filters \( \{\omega_r\}_q \), a block comparing the amplitudes of the signals and \( q \) generators of quasi-monochromatic signals \( \{\omega_r\}_q \). The network operates as follows: the input signals are summarized; the summarized signal propagates through \( q \) parallel frequency filters; the output signals from the filters are compared as regards their amplitudes; the signal with the maximal amplitude activates generation of an output signal, whose frequency and phase are the same as the frequency and phase of the initiating signal. For this scheme the condition of the frequencies noncommensurability is of key importance:

\[
\omega_i - \omega_j + \omega_k \notin \{\omega_r\}_q \quad \text{when} \quad i \neq j \neq k \neq i. \tag{1}
\]

This condition having been satisfied, the internal noises get suppressed to a great extent.

\dagger Corresponding author.
E-mail address: litin@hppi.troitsk.ru (L.B. Litinskii).
\( ^1 \) Grant 01-01-00090 RBRF.
\( ^2 \) Participation in ACAT’2002 is supported by HP-grant.
We call this network the parametrical neural network (PNN). The neurons of PNN can be in \( q \) different states. Then, \[4\] the states of neurons have to be described in terms of vectors. In the next section we present the respective vector formalism.

1. Vector formalism for PNN

Let us examine a network consisting of \( N \) neurons connected with each other. In order to describe the states of neurons we use the set of basic vectors \( \mathbf{e}_k \) in the space \( \mathbb{R}^q \), where \( q \geq 1 \). The state of the \( i \)th neuron is described by a vector \( \mathbf{x}_i = x_i \mathbf{e}_i \), where \( x_i = \pm 1 \), \( \mathbf{e}_i \in \mathbb{R}^q \), \( k = 1, \ldots, q \), \( i = 1, \ldots, N \). Then, the state of the network as a whole is determined by a set of \( N \) \( q \)-dimensional vectors \( \mathbf{x}_i: \mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \). Since neurons are vectors in this model, the local field \( \mathbf{h}_i \) affecting the \( i \)th neuron is a vector too. By analogy with the Hopfield model, we write

\[
\mathbf{h}_i = \sum_{j=1}^{N} T_{ij} \mathbf{x}_j = \sum_{k=1}^{q} A_{k}^{(i)} \mathbf{e}_k
\]

where the \((q \times q)\)-matrices \( T_{ij} \) describe the interconnections between the \( i \)th and the \( j \)th neurons. Its elements can be calculated using the generalized Hebb rule:

\[
T_{ij}^{(kl)} = (1 - \delta_{ij}) \sum_{\mu=1}^{p} (\mathbf{e}_k \mathbf{x}_\mu)(\mathbf{x}_\mu \mathbf{e}_l)
\]

where \( k, l = 1, \ldots, q \) and \( p \) stored patterns are \( \mathbf{X}_\mu = (\mathbf{x}_\mu, \ldots, \mathbf{x}_\mu^N) \). The matrices (3) are consistent with the principle of noncommensurability of frequencies (1).

Let \( \mathbf{X}(t) \) be the state of the system at the time \( t \). Then, the amplitudes \( A_{k}^{(i)} \) in the right-hand side of Eq. (2) are equal

\[
A_{k}^{(i)} = \sum_{j(\neq i)}^{N} \sum_{\mu=1}^{p} (\mathbf{e}_k \mathbf{x}_\mu)(\mathbf{x}_\mu \mathbf{x}_j(t)).
\]

We use the subscript max to denote an amplitude \( A_{k}^{(i)} \) that is maximal in modulus. Then the neuron dynamics is as follows: due to the action of the local field the \( i \)th neuron at the time \( t + 1 \) is oriented along the basic vector \( \mathbf{e}_{\text{max}} \), if \( A_{k}^{(i)} > 0 \), and it is oriented in the opposite direction if \( A_{k}^{(i)} < 0 \), i.e., \( \mathbf{x}_i(t + 1) = \text{sgn}(A_{k}^{(i)}) \mathbf{e}_{\text{max}} \).

This network is equivalent of the PNN described above. When \( q = 1 \), the network transforms into the Hopfield model.

2. Storage capacity of PNN when \( N \gg 1 \)

Let the network start from the distorted pattern \( \mathbf{X}_m \)

\[
\mathbf{X}_m = (a_1 \mathbf{b}_1 \mathbf{x}_m, \ldots, a_N \mathbf{b}_N \mathbf{x}_m^N).
\]

Here \( \{a_i\}_N \) and \( \{b_i\}_N \) define a multiplicative noise: \( a_i \) is a random value that is equal to \(-1\) or \(+1\) with the probabilities \( a \) and \( 1 - a \) respectively; \( b \) is the probability that the operator \( b_i \) changes the state of the vector \( \mathbf{x}_m \), and \( 1 - b \) is the probability that vector \( \mathbf{x}_m \) remains unchanged. Then, substituting Eq. (5) in Eq. (4), we obtain

\[
A_{k}^{(i)} = (\mathbf{e}_k \mathbf{x}_m) \sum_{j=1}^{N} \sum_{\mu=1}^{L} \xi_j + \sum_{r=1}^{L} \eta_r, \quad \xi_j = a_j(\mathbf{x}_m \mathbf{b}_j \mathbf{x}_m), \quad \eta_{j} = a_j(\mathbf{e}_k \mathbf{x}_m)(\mathbf{x}_m \mathbf{b}_j \mathbf{x}_m), \quad j \neq i, \mu \neq m, \quad L = (N - 1)(p - 1).
\]

Since the patterns \( \{X_\mu\}^p \) are uncorrelated, the quantities \( \xi_j \) and \( \eta_r \) can be considered as independent random variables with the probability distributions

\[
\xi_j = \begin{cases} +1, & (1 - b)(1 - a) \\ 0, & b \\ -1, & (1 - b)a \end{cases}, \quad \eta_r = \begin{cases} +1, & 1/2q^2 \\ 0, & \frac{q^2 - 1}{q^2} \\ -1, & 1/2q^2 \end{cases}.
\]

We use the well-known Chebyshev-Chernov method [5] to estimate the probability of the error of the pattern \( \mathbf{X}_m \) recognition, \( \text{Pr}_{\text{err}} \):

\[
\text{Pr}_{\text{err}} = N \exp\left(-\frac{N(1 - 2a)^2}{2p} q^2 (1 - b)^2 \right). \tag{6}
\]

When \( N \) increases, this probability tends to zero, if \( p \) increases as a function of \( N \) slower than

\[
p_c = \frac{N(1 - 2a)^2}{2 \ln N} q^2 (1 - b)^2. \tag{7}
\]

\( p_c \) is an asymptotically possible value of the PNN storage capacity.

When \( q = 1 \), Eqs. (6) and (7) transform into well-known results for the Hopfield model. When \( q \) increases, the noise immunity of PNN increases noticeably. At the same time, the storage capacity
of the network increases as $q^2$. In contrast to the Hopfield model, the number of the patterns $p$ can by many times in excess of the number of neurons. We simulated PNN with $N = 100$ and $q = 25$. When $p = 200$, PNN reconstructed the pattern that was 80–90% noisy after one cycle. When we have $p = 1000(1)$, the network reconstructed a 65% noisy pattern after 4 or 5 cycles.

The same analysis of the Potts-glass neural network (PGNN) [4] has shown that its storage capacity is twice as low as the storage capacity of PNN (7). As a result, the noise immunity of PNN is much higher than the noise immunity of PGNN.

Acknowledgements

The work was supported by RBRF grant 02-01-00457 and the program “Intellectual Computer Systems” (the project 4.5).

References