ABSTRACT

Neural computation based on principles of quantum mechanics can provide improved models of memory processes and brain functioning and is of primary importance for the realization of quantum computing machines. To this end, this chapter studies neural structures with weights that follow the model of the quantum harmonic oscillator. The proposed neural networks have stochastic weights which are calculated from the solution of Schrödinger’s equation under the assumption of a parabolic (harmonic) potential. These weights correspond to diffusing particles, which interact with each other as the theory of Brownian motion (Wiener process) predicts. The learning of the stochastic weights (convergence of the diffusing particles to an equilibrium) is analyzed. In the case of associative memories the proposed neural model results in an exponential increase of patterns storage capacity (number of attractors). It is also shown that conventional neural networks and learning algorithms based on error gradient can be conceived as a subset of the proposed quantum neural structures. Thus, the complementarity between classical and quantum physics is also validated in the field of neural computation.
INTRODUCTION

In this chapter, neural structures with weights that follow the model of the quantum harmonic oscillator will be studied. Connectionist structures which are compatible with the theory of quantum mechanics and demonstrate the particle-wave nature of information, have been analyzed in (Hagan et. al., 2002), (Perus et. al., 2004), (Rigatos & Tzafestas, 2007a). This chapter extends results on the compatibility of neural structures with quantum mechanics principles, presented in (Rigatos &. Tzafestas, 2002), (Rigatos & Tzafestas, 2006a).

It is assumed that the neural weights are stochastic variables which correspond to diffusing particles, and interact to each other as the theory of Brownian motion predicts. Brownian motion is the analogous of the quantum harmonic oscillator (Q.H.O.), i.e. of Schrödinger’s equation under harmonic (parabolic) potential. It will be shown that the update of the stochastic weights is described by Langevin’s stochastic differential equation which is a generalization of conventional gradient algorithms. It will also be shown that weights following the Q.H.O. model give to associate memories significant properties: (i) the learning of the weights is a Wiener process, and (ii) the number of attractors increases exponentially comparing to conventional associative memories.

The structure of the chapter is as follows: In Section “Equivalence between Schrödinger’s equation and a diffusion process”, an analysis of the diffusive motion of particles (stochastic weights) is given and the background that relates Schrödinger’s equation with diffusive motion is analyzed. It is shown that Schrödinger’s equation with harmonic potential is equivalent to a stationary diffusion, and that the motion of the diffusing particles is described by Langevin’s equation. In section “Interacting diffusing particles as a model of neural networks”, a neural model based on interacting diffusing particles, is proposed. In Section “Compatibility with principles of quantum mechanics”, it is shown that using directly Schrödinger’s equation in place of the previously analyzed stochastic processes, the stochastic weights of the neural network can be described by a probability density function which stems again from the model of the quantum harmonic oscillator. In section “Attractors in Associative Memories based on the Q.H.O. model”, it is shown that the Q.H.O. model in associative memories results in exponential increase of the number of attractors while the the update of the weights stands for a Wiener process. In section “Spectral analysis of associative memories that follow the QHO model”, spectral analysis shows that the stochastic weights satisfy a relation analogous to the principle of uncertainty. In section “Simulation tests”, simulation results are presented about the convergence of Brownian weights to attractors and about the exponential increase of the number of attractors in associative memories with Brownian weights. Finally, in Section “Conclusions”, concluding remarks are stated.

BACKGROUND

A detailed analysis of the current status of research on neural networks and cognitive models based on the principles of quantum mechanics will be given. As it can be observed from the relevant bibliography the field is wide and with enormous potential. Without excluding other approaches on neural structures with quantum mechanical features, in this chapter three main research directions are distinguished: (1) Neural structures which use as activation functions the eigenstates of the quantum harmonic oscillator, (2) Neural structures with stochastic weights. Sub-topics in this area are (a) neural structures with stochastic weights which stem from the solution of Schrödinger’s linear equation for constant or zero potential (b) stochastic neural networks which can be modelled as gene networks (3). Neural structures with stochastic weights which stem from the model of the quantum harmonic oscillator

1. Results on neural structures which use as activation functions the eigenstates of the quantum harmonic oscillator: Feed-forward neural networks (FNN) are the most popular neural architectures due to their structural flexibility, good representational capabilities, and availability of a large number of training algorithms. The hidden units in a FNN usually have the same activation functions and are usually selected as sigmoidal functions or gaussians. Feed-forward neural networks that use the eigenstates of the quantum harmonic oscillator (QHO) as basis functions have some interesting properties: (i) the basis functions are invariant under the Fourier transform,
subject only to a change of scale (ii) the basis functions are the eigenstates of the QHO, and are derived from the solution of Schrödinger’s harmonic equation. The proposed neural networks belong to the general category of nonparametric estimators and are suitable for function approximation, system modelling and image processing. Two-dimensional QHO-based neural networks can be also constructed by taking products of the one-dimensional basis functions.

Feed-forward neural networks that use the eigenstates of the quantum harmonic oscillator demonstrate the particle-wave nature of information as described by Schrödinger’s diffusion equation (Cohen-Tannoudji et al., 1998), (Strauss, 1992). Attempts to enhance the connectionist neural models with quantum mechanics properties can be also found in (Kosko, 1992), (Ventura & Martinez, 2000), (Resconi et al., 2002), (Perus, 2001). An analysis of NNs which use as activation functions the QHO eigenstates is given in (Rigatos & Tzafestas, 2006b), (Rigatos, 2006c). It is considered that the input variable \( x \) of the neural network can be described not only by crisp values (particle equivalent) but also by the normal modes of a wave function (wave equivalent). Since the basis functions of the proposed FNN are the eigenstates of the quantum harmonic oscillator, the FNN’s output will be the weighted sum of these eigenfunctions, while the the square of the output will be the probability that the input of the neural network (quantum particle equivalent) is found in the uncertainty interval \([x, x + \Delta x]\). Thus, in these NNs the weights provide a measure of the probability to find the input on the neural network in the region associated with the eigenfunction \( \psi_k(x) \).

Furthermore, issues related to the uncertainty principle have been examined in case of the QHO-based neural network. The uncertainty principle can be expressed first through the Balian-Low theorem for the case of Gabor frames. Next an expression of the uncertainty principle for basis functions that follow the model of the quantum harmonic oscillator can be given. The uncertainty principle is a measure of the time-frequency localization of the activation functions in the QHO-based neural network and evaluates the degradation of localization when successive elements of these orthonormal basis functions are considered. It is shown that the Hermite basis functions as well as their Fourier transforms cannot be uniformly concentrated in the time-frequency plane (Jamming et al., 2007), (Powell, 2005).

2. Results on neural structures with stochastic weights

2a. Neural structures with stochastic weights which stem from the solution of Schrödinger’s linear equation for zero or constant potential. Modelling of brain processes has emerged as a promising research topic, which can help in the understanding of the human cognition (Hopfield, 1982), (Pribram, 1991), (Kosko, 1992), (Petritis, 1996), (Resconi, 2002). Investigations in this field have delivered successful implementations of learning and memory along with lines inspired by neural architectures. These findings have promoted optimism that a sufficiently complex artificial neural network would reproduce the full spectrum and extent of the relevant brain processes involved in human cognition and consciousness. However, physical effects in the functioning of the nervous system, which lie outside classical physics suggest that conventional neural networks may ultimately prove insufficient for this task. One finds ample support for this in an analysis of the sensory organs, the operation of which is quantized at levels varying from the reception of individual photons by the retina, to thousands of phonon quanta in the auditory system. Of further interest is the argument that synaptic signal transmission has a quantum character, although the debate on this issue has not been conclusive. In any case it is known that quantum modes exist in peptides, DNA and proteins. For instance, in (Hagan et al., 2002) it is argued that superposition of quantum states for a significant time interval takes place in the microtubules of the brain and that memory recall is the result of quantum measurement. Moreover, it has been argued that human cognition must involve an element inaccessible to simulation on classical neural networks and this might be realized through a biological instantiation of quantum computation. In the same direction, in (Petritis, 2008), the statistics of the genomic sequences are studied (which are usually confronted with stochastic or quantum models of computing) and it is established that certain aspects of these statistics are not in contradiction with quantum mechanical Turing computation, corroborating thus the thesis of quantum nature of cognitive processes (at least not invalidating this thesis).

Quantum computation gives a new perspective in computer science and can help in the modelling of cognitive processes (Perus, 2001), (Rigatos & Tzafestas, 2006a). The association of neural structures with quantum infor-
Attractors and Energy Spectrum of Neural Structures

mation processing results in an exponential increase of patterns storage capacity and can explain the extensive memorization and inferencing capabilities of humans (Kak, 1999), (Ventura, 1999). Several models of quantum neural networks containing stochastic parameters have been proposed: in (Ventura, 1997) the weight vector which is associated with a quantum neuron is a linear superposition (wave function) of crisp weight vectors, while in (Purosathaman et al., 1997) the quantum neuron has a multilevel transfer function, which is the result of the linear superposition of sigmoid functions. In (Ventura et al., 2000) quantum associative memories with exponential storage capacity are presented. The stored patterns are considered to be basis states of a quantum wave function. Rotation operators increase the probability to recall the basis state associated with the input pattern. This work is extended in (Ezhov et al., 2000) where fuzzy queries to a quantum memory end up at the retrieval of valid stored patterns. In (Resconi et al., 2002) it is stated that the property of linear superposition indicates an equivalence between neural and quantum computation. The proposed quantum neural structures are non-diagonal recurrent neural networks where the output of each neuron corresponds to a quantum state. The evolution of the states in time is realized through unitary operators. In (Behrmann et al., 2000) quantum dot molecules which are spatially distributed on a suitable substrate and which affect each other through phononic and Coulombic interactions are viewed as a quantum Hopfield net. Several other papers have been published outlining similar ideas (De Garis, 2003). Finally, the use of quantum diodes in the realization of nanocircuits with neural computing attributes is shown in (Banyopadhyay et al., 2002).

In (Rigatos & Tzafestas, 2006a), (Rigatos, 2006d) neural connectionist structures with stochastic weights that stem from Schrödinger’s linear equation for zero or constant potential have been studied and a model of quantum associative memories has been introduced. The main features of this memory model are summarized in the following:

a. The proposed quantum associative memories result from neural associative memories if the elements of the weight matrix are taken to be stochastic (fuzzy) variables. The probability density function of each weight is given by the solution of Schrödinger’s diffusion equation. The correspondence of quantum mechanics to fuzzy logic has been demonstrated.

b. The weights of the proposed associative memories are updated with the use of a learning algorithm that satisfies quantum mechanics postulates. Instead of Hebbian learning, it is assumed that the incremental changes of the weights are performed with the use of fuzzy inference. This fuzzy learning rule is proved to satisfy two basic postulates of quantum mechanics: (i) existence in superimposing states, (ii) evolution between the superimposing states with the use of unitary operators. Therefore it can be considered as a quantum learning algorithm.

c. Taking the elements of the weight matrix of the associative memory to be stochastic (fuzzy) variables means that the initial weight matrix can be decomposed into a superposition of associative memories. This is equivalent to mapping the fundamental memories (attractors) of the associative memory into the vector spaces which are spanned by the eigenvectors of the superimposing matrices and which are related to each other via unitary rotations. In this way, it can be shown that the storage capacity of the associative memories with stochastic weights increases exponentially with respect to the storage capacity of conventional associative memories.

2b. Stochastic neural networks which can be modelled as gene networks. Although the term quantum is not used explicitly in their description there are stochastic neural networks with features that are reminiscent of quantum mechanical principles and which are worth to be presented in this background overview. The pattern formation ability of some genetic circuits has been studied and the behaviour of such networks under stochastic disturbances has been examined (Vakulenko & Grigoriev, 2006). Fundamental learning theory for pattern recognition problems was proposed by L. Valiant and developed by many authors. After, evolutionary algorithms for machine learning (for example, for boolean formula recognition) are proposed by L. Valiant (Valiant, 2007). In the boolean case, many algorithms need an exponentially large running time. The following ideas have been proposed: To recognize real patterns one can use results (Grigoriev et al., 1991a), (Grigoriev et al., 1990), (Grigoriev et al. 1994), (Grigoriev et al. 1991b) that allow to construct fast algorithms of adaptive learning. These ideas are applicable to patterns which are sparse sums of polynomial, rational functions, or, in general, functions
which are eigenfunctions of some linear operators (Grigoriev et al., 1991a), (Grigoriev et al., 1990), (Grigoriev et al. 1994), (Grigoriev et al. 1991b). Such algorithms can be realized physically by networks consisting of quantum or optical devices.

To extend these algorithms to a large class of patterns, a new idea has been proposed (Valiant, 2007), (Grigoriev et al, 1991a), (Grigoriev et al., 1990), (Grigoriev et al. 1994), (Grigoriev et al. 1991b): a technique based on pfaffian functions (Khovanskii, 1991) together with evolutionary algorithms from (Valiant 2007). Earlier this pfaffian approach has been applied to some important genetic and neural network problems (Vakulenko et al 2003a), (Vakulenko et al 2005a), (Vakulenko et al 2003b), (Vakulenko et al. 2003c), (Vakulenko et al 2005b), (Vakulenko et al 2006), (Vakulenko et al, 2002). The main idea is that one can recognize a complicated pattern step by step, using previous patterns and trying to present a given pattern as a sparse sum of functions of previous patterns and space variables.

The goal of the above mentioned stochastic neural networks is to construct a physical model of networks which can realize new algorithms. Such a network can be based on quantum or optical physical mechanisms. Quantum systems to resolve complicated problems are proposed in (Grigoriev et al., 2007). Non-quantum physical realizations of neural networks have been proposed in (Vakulenko 2000). In (Vakulenko 2000), (Vakulenko 2002) it was also shown that Hopfield networks generate any structurally stable dynamics and control algorithms have been proposed.

3. Results on neural structures with stochastic weights which stem from the model of the quantum harmonic oscillator. Conventional neural networks may prove insufficient for modelling memory and cognition, as indicated by physical effects in the functioning of the nervous system which can be better interpreted with the use of quantum mechanics instead of classical physics (Hagan,2002), (Behrman et. al., 2000), (Deutch, 1989), (Feynman, 1986) and (Perus, 2001). To evaluate the validity of the aforementioned arguments, in this chapter elements of stochastic calculus will be used (Klebaner, 2005). It will be assumed that the neural weights are stochastic variables that follow the model of the quantum harmonic oscillator. Stochastic variables that are associated with quantum mechanics equations and which demonstrate the particle-wave nature of information, have been analyzed in (Mahler & Weberuss, 1998), (Muller, 1998), (Cohen-Tannoudji et al., 1998), and (Nielsen et al., 2000). The use of neural networks compatible with quantum mechanics principles in image processing, function approximation and system modelling can be found in (Perus,2000), (Perus et al., 2004) and (Resconi et al, 2002), while the relation between random oscillations and diffusion equations has been studied in (Soong et al, 1992) and (Gitterman, 2005). Other studies on neural models with quantum mechanical properties can be found in (Refregier, 2003), (Rigatos et al., 2006b), (Rigatos, 2007b) and (Ventura et al, 2000). In (Rigatos, 2008b) it has been assumed that the weights of neural networks are stochastic variables. These weights correspond to diffusing particles which interact with each other as the theory of Brownian motion (Wiener process) predicts. This assumption has been also used in (Iwasaki et. al., 1998).

Brownian motion is the analogous of the quantum harmonic oscillator (Q.H.O.), i.e. of Schrödinger’s equation under harmonic (parabolic) potential (Comets et al., 2006). Instead of trying to solve Schrödinger’s equation for various types of the potential $V(x)$ one can study the time evolution of the particle through an equivalent diffusion equation, assuming probability density function depending on Q.H.O’s ground state, (Faris, 2006). This is the model of the Ornstein-Uhlenbeck diffusion which has previously been analyzed in (Basseville et al, 1993). It is shown that the diffusive motion of the stochastic particles (weights’ update) can be described by Langevin’s equation which is a stochastic linear differential equation (Gitterman, 2005) and (Faris, 2006). Using the theoretical analysis of gradient algorithms given in (Dufllo,1996) and (Benveniste et al., 1990), it is proved that Langevin’s equation is a generalization of conventional gradient algorithms, (Rigatos, 2007c), (Rigatos, 2008c). Therefore neural structures with crisp numerical weights can be considered as a subset of NN with stochastic weights based on the Q.H.O model. Additionally, more practical issues have been taken into account and a neural associative memory with stochastic weights following the Q.H.O. model is studied. From the results given in (Levine et al, 2000) it is known that the convergence of multi-particle systems to an attractor can be studied with the use of Lyapunov stability theory. Lyapunov stability analysis has been performed in (Rigatos 2008a) and following this methodology one can provide sufficient conditions for the weights (Brownian particles) to converge to stability points (attractors). It has been shown that: (i) the learning of the weights stands for a Wiener process, (ii) in
Attractors and Energy Spectrum of Neural Structures

Q.H.O-based associative memories the number of attractors increases exponentially, comparing to conventional associative memories. This in turn gives an explanation for the advanced memorization and inference capabilities of the human brain. Furthermore, the energy spectrum of the stochastic weights that follow the Q.H.O model is studied. To this end, previous results on wavelets’ energy spectrum are used (Addison, 2002), (Debauchies, 1990), (Mallat, 1998), and (Torresani, 1995). Spectral analysis of the stochastic weights has shown that: (i) the Gaussian membership functions of the weights express the distribution of energy with respect to the weights’ value. The smaller the spread of the basis functions is, the larger becomes the spectral (energy) content that can be captured therein, (ii) the stochastic weights satisfy a relation analogous to the principle of uncertainty.

4. Modelling of diffusion processes with the use of density estimators. Another important issue for research on quantum neural networks is the modelling of diffusion processes. It has been assumed that the diffusing particles interact to each other as the theory of Brownian motion predicts (Rigatos, 2007c), (Rigatos, 2008c). Brownian motion is the analogous of the quantum harmonic oscillator (Q.H.O.), i.e. of Schrödinger’s equation under harmonic (parabolic) potential. However, the analytical or numerical solution of Schrödinger’s equation, is computationally intensive, since for different values of the potential \( V(x) \) it is required to calculate the modes \( \psi_j(x) \) in which the particle’s wave-function \( \psi(x) \) is decomposed (Chung 2002), (Chuong et al., 2007). Moreover, the solution of Schrödinger’s equation contains non-easily interpretable terms such as the complex number probability amplitudes which are associated with the modes \( \psi(x) \), or the path integrals that constitute the particle’s trajectory. On the other hand, instead of trying to solve Schrödinger’s equation for various types of \( V(x) \) one can study the time evolution of the particles through an equivalent diffusion equation, assuming probability density function depending on Q.H.O’s ground state, i.e. (Faris, 2006), (Rigatos, 2008c).

An outline of the research work contained in this chapter is given in Table 1.

### EQUIVALENCE OF SCHRÖDINGER’S EQUATION TO A DIFFUSION PROCESS

1. Wiener Walk And Wiener Process
First, the Wiener walk will be analyzed and the Wiener process will be derived as a limit case of the walk. The Wiener walk describes a simple symmetric random walk. Assume \( \xi_1, \xi_2, \ldots, \xi_n \) a finite sequence of independent random variables, each one of which takes the values \( \pm 1 \) with the same probability. The random walk is the sequence

\[
s_k = \xi_1 + \xi_2 + \ldots + \xi_k, \quad 0 \leq k \leq n
\]

(1)

### Table 1. Associative memories based on the QHO model

| 1. The weights of the quantum associative memory correspond to interacting Brownian particles. |
| 2. The dynamic behaviour of the weights can be described by a stochastic process (Ornstein-Uhlenbeck diffusion) which is shown to be analogous to the model of the quantum harmonic oscillator. |
| 3. The dynamic behaviour of the weights is also described by the probability density function, which is calculated from the solution of Schrödinger’s harmonic equation. |
| 4. For a large number of weights (high-dimensional neural network) the dynamics of the associative memory can be studied with the use of statistical methods, such as the Central Limit Theorem (C.L.T.) |
| 5. The probability distribution of the stochastic weights can be substituted by a fuzzy possibility function. This enables to show the decomposition of the weight matrix into a superposition of individual matrices and the existence of stochastic attractors |
| 6. Comparing to conventional associative memories, the number of attractors of the proposed quantum associative memories increases exponentially. |
| 7. Spectral analysis shows that the stochastic weights satisfy a relation which is equivalent to the principle of uncertainty. |
1. The $n$-step Wiener walk is considered in the time interval $[0, T]$, where the time step $\Delta t$ is associated with particle’s displacement $\Delta x$. Thus the following random variable $w(t_n)$ can be defined:

$$w(t_n) = \xi_0 \Delta x + \ldots + \xi_n \Delta x \Rightarrow w(t_n) = w(t_{n-1}) + \xi_n \Delta x$$

The random function that is described by Eq. (2) is the Wiener walk (Faris, 2006). A sample of the Wiener walk is depicted in Fig. 1(a). The Wiener walk (also known as Brownian motion) is an important topic in the theory of stochastic processes since it provides a model for the motion of a particle under the effect of a potential (Comets & Meyre, 2006). The Wiener process is the limit of the Wiener walk for $n \to \infty$, and using the central limit theorem (C.L.T.) it can be shown that the distribution of the Wiener process is Gaussian (Rigatos & Tzafestas, 2006a), (Faris, 2006). Indeed, since the random variable $w(t_n)$ of Eq. (2) is the sum of an infinitely large number of increments, then according to the C.L.T. it must follow a Gaussian distribution. Thus one obtains (Faris, 2006):

$$E[w(t)] = 0, \quad \text{while } E[w(t) - E[w(t)]]^2 = \sigma^2 \cdot t$$

2. The Wiener Process Corresponds To A Diffusion Partial Differential Equation

It has been shown that the limit of the Wiener walk for $n \to \infty$ is the Wiener process, which corresponds to the partial differential equation (p.d.e) of a diffusion (Faris, 2006). For $t > 0$, function $\rho(x,t)$ is defined as the probability density function (p.d.f) of the Wiener process $f(w(t))$, i.e.

$$E[f(w(t))] = \int_{-\infty}^{\infty} f(x)\rho(x,t)dx$$

The p.d.f. $\rho(x,t)$ is a Gaussian variable with mean value equal to 0 and variance equal to $\sigma^2$ and satisfies a diffusion p.d.e of the form

$$\frac{\partial \rho}{\partial t} = \frac{1}{\sigma^2} \frac{\partial^2 \rho}{\partial x^2}$$

which is the simplest diffusion equation (heat equation). The generalization of the Wiener process in an infinite dimensional space is the Ornstein-Uhlenbeck process, where the joint probability density function is also Gaussian.
Attractors and Energy Spectrum of Neural Structures

Faris, 2006), (Basseville & Nikiforov, 1993). In that case there are \( n \) Brownian particles, and each one performs a Wiener walk, given by \( w(t_i), \ i=1,\ldots, n \) of Eq. (2).

3. Schrödinger’s Equation With Nonzero Potential Stands For A Diffusion with Drift

The equivalence between Wiener process and the diffusion process was demonstrated. Next, the direct relation between diffusion and quantum mechanics will be shown (Faris, 2006). The basic equation of quantum mechanics is Schrödinger’s equation, i.e.

\[
\imath \frac{\partial \psi}{\partial t} = H \psi (x,t)
\]  

(6)

where \( \psi (x,t) \) is the probability density function of finding the particle at position \( x \) at time instant \( t \), and \( H \) is the system’s Hamiltonian, i.e. the sum of its kinetic and potential energy, which is given by \( H = \frac{p^2}{2m} + V \), with \( p \) being the momentum of the particle, \( m \) the mass and \( V \) an external potential. It holds that \( \frac{p^2}{2m} = - \frac{\hbar}{2 \sigma^2} \) thus the Hamiltonian can be also written as \( H = - \frac{\hbar}{2 \sigma^2} \). The solution of Eq. (6) is given by (Cohen-Tannoudji et. al, 1998)

\[
\psi (x,t) = e^{-\imath H t} \psi (x,0)
\]  

(7)

A simple way to transform Schrödinger’s equation into a diffusion equation is to substitute variable \( \imath t \) with \( t \). This passage from imaginary time to real time is convenient but artificial. However, in the domain of non-relativistic quantum mechanics there is a closer connection between diffusion theory and quantum theory. In stochastic mechanics, the real time of quantum mechanics is also the real time of diffusion and in fact quantum mechanics is formulated as conservative diffusion (Faris, 2006). This change of variable results in the Fokker-Planck partial differential equation (Gitterman, 2005), (Faris, 2006).

\[
\frac{\partial \rho}{\partial t} = \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} - V(x) \rho
\]  

(8)

Eq. (8) can be also written as

\[
\frac{\partial \rho}{\partial t} = -H \rho
\]  

(9)

where \( H \) is the associated Hamiltonian and the solution is of the form \( \rho(x,t) = e^{-\imath H t} \rho(x) \), and variable \( \sigma^2 \) is a diffusion constant.

The Fokker-Planck equation can be used for the calculation of the mean position of the diffused particle, as well as for the calculation of its variance (Gitterman, 2005). An equivalent model of the motion of the diffused particle is based on Langevin’s equation, and will be analyzed in the sequel.

Now, as known from quantum mechanics, particle’s probability density function \( \rho(x) \) is a wave-function for which holds \( \rho(x) = | \psi(x) |^2 \), with \( \psi (x) = \sum_{k=0}^{\infty} c_k \psi_k (x) \), where \( \psi_k (x) \) are the associated eigenfunctions (Müller, 1998)), (Cohen-Tannoudji, 1998). It can be assumed that \( | \rho_0 (x) = |\psi_0 (x) | \), i.e. the p.d.f includes only the basic
mode, while higher order modes are truncated. Thus, it is considered that the initial probability density function is \( \rho(x) = \rho_0(x) \), which is independent of time. Consequently, from Eq. (9) one has \( H\rho_0 = 0 \), which means that the p.d.f. remains independent of time and the examined diffusion process is a stationary one, i.e. \( \rho(x,t) = \rho_0(x) \ \forall t \).

A form of the probability density function for the stationary diffusion is that of shifted, partially overlapping Gaussians, which is depicted in Fig. 1(b). In place of Gaussian p.d.f, symmetric triangular possibility distributions have been also proposed (Rigatos & Tzafestas, 2002), (Dubois et. al., 2004). The equation that describes the shifted Gaussians is (Faris, 2006)

\[
\rho_0(x) = \frac{1}{2} C e^{-\frac{x^2}{a^2}} + \frac{1}{2} C e^{-\frac{x^2}{b^2}}
\]

(10)

4. Study Of The Q.H.O Model Through The Ornstein-Uhlenbeck Diffusion

The Ornstein-Uhlenbeck diffusion is a model of the Brownian motion (Basseville & Nikiforov, 1993). The particle tries to return to the equilibrium \( x = 0 \) under the influence of a linear force, i.e. there is a spring force applied to the particle as a result of the potential \( V(x) \). The corresponding phenomenon in quantum mechanics is that of the quantum harmonic oscillator (Q.H.O.) (Cohen-Tannoudji, 1998) and (Gitterman, 2005). In the Q.H.O. model the motion of the particle is affected by the parabolic (harmonic) potential

\[
V(x) = \frac{1}{2} \omega^2 x^2
\]

(11)

It is known that the ground mode of the quantum harmonic oscillator of Eq. (6) is a Gaussian function (Rigatos & Tzafestas, 2006b), (Faris, 2006) i.e.

\[
\psi_0(x) = Ce^{-\frac{x^2}{2a^2}}
\]

(12)

while it can be proved easily that the associated eigenvalue is \( \lambda_0 = \frac{1}{2} \omega \). A diagram of \( V(x) - \lambda_0 \) is given in Fig. 2(a).

For the diffusion constant \( \sigma \) holds \( \sigma^2 = \hbar / m \) where \( \hbar \) is Planck’s constant and this finally gives \( V(x) = \frac{1}{2} m\omega^2 x^2 \).

Figure 2.
Attractors and Energy Spectrum of Neural Structures

Assuming the stationary p.d.f. of Eq. (10), i.e. \( p(x) = \psi(x)^2 = C^2 e^{-\frac{x^2}{\sigma^2}} \), the force applied to the particle due to the harmonic potential \( V(x) \) is given by Eq. (12), and is of the form

\[
u(x) = -kx
\]

which means that the drift is a spring force applied to the particle and which aims at leading it to an equilibrium position. The drift force is depicted in Fig. 2(b).

5. The Motion Of The Q.H.O. Particle Is A Generalization Of Gradient Algorithms

As mentioned above the Q.H.O. model describes the motion of a particle. Here, a kinematic model for particle’s motion will be derived, using Langevin’s equation. The stochastic differential equation for the position of the particle is (Faris, 2006):

\[
dx(t) = \nu(x(t))dt + dw(t)
\]

where \( \nu(x) = -kx \) is the drift function of Eq. (13), i.e. a spring force generated by the harmonic potential \( V(x) \), which tries to bring the particle to the equilibrium \( x = 0 \). The term \( w(t) \) denotes a random force (due to interaction with other particles) and results in a Wiener walk. For each continuous random path \( w(t) \), a continuous random path \( x(t) \) is also generated, which can be written in the form

\[
x(t) = x(0) + \int_0^t \nu(x(s))ds + w(t)
\]

The integration of Langevin’s equation and certain assumptions about the noise \( w(t) \), for instance white noise, dichotomic noise (also known as Ornstein-Uhlenbeck noise) etc., enable the calculation of the mean position of the particle \( E(x) \) and of its variance \( E(x-E(x))^2 \) (Gitterman, 2005).

Langevin’s equation gives a model of an harmonic oscillator, driven by noise. Apart from the spring force, a friction force that depends on the friction coefficient \( \gamma \) and on the velocity of the particle can be considered. Thus, the model of motion of the particle can then be also written as (Gitterman, 2005), (Astrom, 2006):

\[
\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + kx = \xi(t)
\]

Knowing that the Q.H.O. model imposes to the particle a spring force Eq. (16), can be formulated as

\[
dx(t) = -kdx + w(t).
\]

Equation (16) is close to robotics and control models (Tzafestas & Rigatos, 2000), (Rigatos, 2002). Eq. (17) can be also written as

\[
dx(t) = h(x(t))dt + w(t)
\]

where \( h(x(t)) = \alpha \frac{\partial V(x)}{\partial t} \), with \( \alpha \) being a learning gain, \( V(x) \) being the harmonic potential and \( w(t) \) being a noise function. Eq. (18) is a generalization of gradient algorithms based on the ordinary differential equation (O.D.E) concept, where the gradient algorithms are described as trajectories towards the equilibrium of an ordinary differential equation (Duflo, 1996), (Benveniste et. al., 1990). Indeed, conventional gradient algorithms with diminishing step are written as
\[ dx(t) = h(x(t))dt \]  \hspace{1cm} (19)

The comparison of Eq. (18) and Eq. (19) verifies the previous argument. The update of the neural weights that follow the model of the quantum harmonic oscillator is given by Eq. (18), which is a description of Schrödinger’s equation for parabolic potential. The force that drives \( x(t) \) to the equilibrium is the derivative of the harmonic potential, and there is also an external noisy force \( w(t) \) which is the result of collisions or repulsive forces due to interaction with neighbouring particles. Similarly, in the update of neural weights with a conventional gradient algorithm, the weight is driven towards an equilibrium under the effect of a potential’s gradient, and the potential is usually taken to be a quadratic error cost function. Feedback from neighbouring neurons can affect the weight’s update.

**INTERACTING DIFFUSING PARTICLES AS A MODEL OF NEURAL NETWORKS**

The concept of a neural network with weights defined by the position \( x^i \) of interacting Brownian particles can be also found in (Iwasaki et al. 1998). The interaction \( g \) between the stochastic weights (Brownian particles) contains attractive forces (spring-mass forces) and repulsive forces (defined by a Gaussian term). In this neural network the update of the weights is finally the kinematic model of the Q.H.O. particles and can be approximated by:

\[ x^i(t+1) = x^i(t) + \gamma^i(t)[-\nabla x^i V^i(x^i) + \epsilon^i(t+1)] + \sum_{j=1, j \neq i}^M g(x^i - x^j) \]  \hspace{1cm} (20)

with \( i = 1, 2, ..., M \) to be the weight’s (particle’s) index, \( \gamma^i(t) = 1 \) and \( \epsilon^i(t+1) \) to be a noise term.

The term \( g(x^i - x^j) \) is defined as

\[ g(x^i - x^j) = -(x^i - x^j)(\alpha - be^{\frac{-|x^i - x^j|^2}{\sigma^2}}) \]  \hspace{1cm} (21)

and denotes the interaction between particle \( i \) and particle \( j \). Coefficient \( \alpha \) defines an attractive (spring-type) force between the particles, while coefficient \( b \) defines a repulsive force between the particles.

Eq. (20) is in accordance with Eq. (18) and shows that weights’ update is finally described by interacting gradient algorithms. The stability of the learning algorithm, i.e. the convergence of the Brownian weights to the
desirable values is equivalent to the convergence of the interacting particles to an attractor and can be studied using Lyapunov stability analysis and LaSalle’s theorem (Gazi & Passino 2004), (Rigatos, 2008a).

**COMPATIBILITY WITH PRINCIPLES OF QUANTUM MECHANICS**

The neural network which is based on the motion interacting diffusing particles was shown to stem from a stochastic process (Ornstein-Uhlenbeck diffusion) which is analogous to the model of the quantum harmonic oscillator. It was shown that the particle’s diffusive motion is characterised by a probability density function (p.d.f.) which is represented by a mixture of Gaussians (see Eq. (8) and Fig. 1(b)). The same result can be reached if theory of quantum mechanics is used.

1. Free Schrödinger’s Equation and Wave-Packets

It is assumed that each weight of the neural network is a stochastic variable with probability density function which is calculated from the solution of Schrödinger’s equation. The case in which the external potential \( V \) is zero or constant at every point of the space is considered. Then, a solution of Schrödinger’s equation is the plane wave \( \psi(x, t) = Ae^{i(kx-\omega t)} \), where \( k, \omega \) satisfy the L. de Broglie relations \( k = p/h \) and \( \omega = Eh \). According to the principle of superposition any linear combination of these plane waves (wavepacket) is a solution of Schrödinger’s equation, \( \psi(x, t) = \int \psi(k) e^{i(kx-\omega t)} dk \). The wave-packet can be interpreted from the point of view of spectral decomposition. Assume that a particle at time instant \( t = 0 \) is described by the wave function \( \psi(x, 0) = Ae^{ikx} \) and has a well defined momentum, which means that a measurement of the momentum at this time instant will certainly give \( p = hk \). Thus one can deduce that \( e^{ikx} \) characterizes the eigenstate which corresponds to \( p = hk \). Since, on the other hand, there is a plane wave for each real value of \( k \), the eigenvalues that one can expect in a certain measurement of the momentum are all real values (in this case there is no quantification of the possible results: as in classical mechanics all values of the momentum are permitted).

At \( t = 0 \) variables \( \psi(x, 0) \) and \( g(k) \) are connected through a Fourier transform pair (Cohen-Tannoudji et al., 1998), which is \( \psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int g(k) e^{ikx} dx \) and \( g(k) = \frac{1}{\sqrt{2\pi}} \int \psi(x, 0) e^{-ikx} dx \). Setting \( g(k) = g(k) e^{a_0x} \) and taking the Taylor expansion of \( a(k) \) round \( k_0 \), permits to rewrite the wave-packet at \( t = 0 \) as

\[
\psi(x, 0) \approx \frac{e^{a_0x}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{i(k-k_0)x} dk
\]

with \( x_0 = -\frac{da(k)}{dk} |_{k=k_0} \). When \( x - x_0 \) is large Eq.(22) represents a function of \( k \) which performs very fast oscillations in the interval \( \Delta k \). In case that \( g(k) \) is a Gaussian, and \( |x - x_0| > \frac{1}{\Delta k} \), the real part of function \( g(k) e^{i(k-k_0)x} \) is depicted in Fig. 4(a), while for \( |x - x_0| < \frac{1}{\Delta k} \), the real part of function \( g(k) e^{i(k-k_0)x} \) is depicted in Fig. 4(b).

When \( x \) moves far from \( x_0 \), \( \psi(x, 0) \) decreases. This decrease becomes significant when \( e^{i(k-k_0)x} \) completes one oscillation while \( k \) covers the domain \( \Delta k \), i.e. when \( \Delta k(x-x_0) \approx 1 \). If \( \Delta x \) is the approximative width of the wave packet one obtains \( \Delta k \Delta x \approx 1 \), which is the classical relation between the width of two functions, related to each other through a Fourier transform. One can write the above relation as

\[
\Delta x \Delta p \geq \hbar \tag{23}
\]

where \( \Delta p = \hbar \Delta k \) is the spread of the curve representing \( g(k) \). The interpretation of Eq. (23) is as follows: it is impossible to define at a certain time instant both the position and the momentum of the particle with arbitrary precision. When the limit imposed by Eq. (23) is approached, the increase of the accuracy of the position measurement (decrease of \( \Delta x \)) implies a decrease in the accuracy of the momentum measurement (increase of \( \Delta p \)) and vice-versa. This relation is Heisenberg’s principle of uncertainty.
2. The Gaussian Wave-Packet

The wave packet associated with the free quantum particle \( V(x) = 0 \) is considered and function \( g(k) \) is taken to be Gaussian. In that case \( g(x) \) is found to be
\[
g(k) = \frac{1}{\sqrt{2\pi \Delta k}} e^{-k^2 \Delta k^2},
\]
where \( C \) is a normalization coefficient.

The probability to find the particle (stochastic weight), at time instant \( t = 0 \), in the interval \( [x, x + \Delta x] \) is given by
\[
|g(x)|^2 = C e^{-k^2 \Delta k^2}.
\]

Therefore, the probability density function of the particle is given by the well-known bell-shaped curve, centered at \( x = \) (see Fig. 5(a)). Similarly, the probability to find the quantum particle in the intervals \( x_i \) \( [x, x + \Delta x] \) is given by Gaussian functions, centered at \( x_i \) (Fig. 5(b)). The constraint that the total probability equals 1 should be satisfied.

3. Eigenstates of the Quantum Harmonic Oscillator

In case that a potential \( V(x) \) exists, Schrödinger’s equation (Eq. (6)) becomes
\[
i \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(x,t) + V(x) \psi(x,t) \Rightarrow i \frac{\partial \psi(x,t)}{\partial t} = H \psi(x,t)
\]
(24)

where \( H \) is the Hamiltonian, i.e. the sum of the potential \( V(x) = \frac{p^2}{2m} + E \) and of the Laplacian
\[
\frac{\hbar^2}{2m} \nabla^2 \psi(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x,t).
\]
The solution of Eq. (24) can be found using the separation of variables method, i.e. \( \psi(x,t) = X(x)T(t) \). It has been proved that \( T(t) \) is given by (Strauss, 1992)
\[
\frac{\partial T}{\partial t} = \lambda T(t) \Rightarrow T(t) = e^{\alpha t}
\]
(25)

while \( \lambda(x) \) corresponds to an undamped oscillation, with boundary condition \( \lim_{x \to \pm \infty} X(x) = 0 \). In case of Eq. (24), i.e. when Schrödinger’s equation contains a potential term, the following solution of the harmonic quantum oscillator is considered
\[
X_k(x) = H_k(x)e^{-x^2/2} \quad k = 0, 1, 2, ...
\]
(26)
where $H_n(x)$ are the Hermite orthogonal functions. The Hermite functions $H_n(x)$ are the eigenstates of the quantum harmonic oscillator (Fig. 6(a) and Fig. 6(b)). The general relation for the Hermite polynomials is

$$H_n(x) = (-1)^k e^{-x^2} \frac{d^k}{dx^k} e^{x^2}$$  \hspace{1cm} (27)

According to Eq. (27) the first five Hermite polynomials are $H_0(x) = 1$, $H_1(x) = 2x$, $H_2(x) = 4x^2 - 2$, $H_3(x) = 8x^3 - 12x$, $H_4(x) = 16x^4 - 48x^2 + 12$, and are depicted in Fig. 6(b). The general solution of the Schrödinger’ equation both in time and space is written as:

$$\psi(x,t) = H_k(x)e^{-x^2/2} e^{-i(\omega x - k \omega t)} \hspace{1cm} k = 0,1,2,...$$  \hspace{1cm} (28)

Eq. (28) satisfies the boundary condition $\lim_{r \to \infty} X(x) = 0$. It can be observed that if only the basic mode of the Q.H.O. is maintained then again the position of the particle (stochastic weight) is described by the p.d.f. depicted in Fig. 1(b).
Weights Learning In Associative Memories is a Wiener Process

The patterns to be stored in an associative memory are denoted by $x_k^j$, $k=1,2,...,p$ and their elements are given by $x_k^j$, $j = 1, 2, ..., N$ (see Fig. 3(a)). The weights of the associative memory, are selected using Hebb's postulate of learning, i.e. $w_{ji} = \frac{1}{N} \sum_{k=1}^{p} x_k^j x_k^i$, $j \neq i$ and $w_{ji} = 0$, $j = i$, where $i, j = 1, 2, ..., N$. The above can be summarized in the network’s correlation weight matrix which is given by $W = \frac{1}{N} \sum_{k=1}^{p} x_k^T x_k$. In the case of binary fundamental memories the weights learning is given by Hebb’s rule

$$w_{ji}(n+1) = w_{ji}(n) + \text{sgn}(x_{k+1}^j x_{k+1}^i)$$

(29)

where $k$ denotes the $k$-th fundamental memory. The increment $\text{sgn}(x_{k+1}^j x_{k+1}^i)$ may have the values +1 or -1. Thus, the weight $w_{ji}$ is increased or decreased by 1 each time a new fundamental memory is presented to the network. The maximum number of fundamental memories that can be retrieved successfully from an associative memory of $N$ neurons (error probability <1%) is given by $P_{\text{max}} = \frac{N}{2 \ln N}$ (Ayier et al., 1990), (Haykin, 1994). Thus, for $N = 16$ the number of fundamental memories should not be much larger than 2.

The equivalence between Eq. (2) and Eq. (29) is obvious and becomes more clear if in place of $\text{sgn}(x_{k+1}^j x_{k+1}^i)$, the variable $\xi_{k+1} = \pm 1$ is used, and if $\xi_{k+1}$ is also multiplied with the increment $\Delta x$. Thus the learning of the weights of an associative memory is a Wiener process. Consequently, the weights can be considered as Brownian particles and for large $N$ their mean value and variance can be calculated using the Central Limit Theorem.

### 2. Attractors in High Dimensional Spaces Coincide with Eigenvectors Of Weight Matrix $W$

It can be shown that the memory vectors of an associative memory become collinear to the eigenvectors of matrix $W$, thus constituting an orthogonal basis if the following two conditions are satisfied: (i) the number of neurons $N$, is large (high dimensional spaces), and (ii) the memory vectors are chosen randomly. The following lemmas give sufficient conditions for the fundamental memory vectors (attractors) to coincide with the eigenvectors of the weight matrix $W$.

**Lemma 1:** If the fundamental memory vectors (attractors) of the associative memory are chosen to be orthogonal, then they are collinear to the eigenvectors of matrix $W$ (Rigatos & Tzafestas, 2006a).

**Proof:** The fundamental memory vectors $\hat{x}$ are taken to be orthogonal to each other, i.e. $\hat{x}_i \cdot \hat{x}_j = \delta (i-j)$, where $i, j = 1, 2, ..., N$. The weight matrix $W$ is given by Eq. $W = \frac{1}{N} \sum_{k=1}^{p} \hat{x}_k^T \hat{x}_k$. Thus, the following holds

$$W \hat{x}_k = \frac{1}{N} \left\{ \sum_{i=1}^{p} \hat{x}_i \hat{x}_i^T \right\} \hat{x}_k = \frac{1}{N} \left\{ \sum_{i=1}^{p} \hat{x}_i (\hat{x}_i \cdot \hat{x}_k) \right\} \Rightarrow W \hat{x}_k = \frac{1}{N} \hat{x}_k$$

(30)

From Eq. (30) it can be deduced that if the memory vectors are orthogonal then they are collinear to the eigenvectors of the matrix $W$ ($Q.E.D.$).

**Lemma 2:** If the memory vectors of an associative memory are chosen randomly and the number of neurons $N$ is large, then there is high probability for them to be orthogonal (Rigatos & Tzafestas, 2006a).

**Proof:** The normalized internal product of the memory vectors $x_i$ and $\hat{x}_i$ is considered
\[ \frac{1}{N} x_i x_i^T = \frac{1}{N} \sum_{j=1}^{N} x_j x_j^T = \frac{1}{N} \sum_{j=1}^{N} Y_j \]  

(31)

For large \( N \) and \( x_j, x_i \) randomly chosen from the discrete set \{-1, 1\} it holds that the mathematical expectation of \( Y_j \), denoted by \( E(Y_j) \), is \( E(Y_j) = 0 \) and

\[ E(Y_j - \bar{Y}_j)^2 = \frac{1}{N} \sum_{j=1}^{N} (Y_j - \bar{Y}_j)^2 = \frac{1}{N} \sum_{j=1}^{N} y_j^2 = \frac{1}{N} \sum_{j=1}^{N} (x_j x_j^T) \]

Assuming patterns \( x_j \in \{-1, 1\} \) then \( (x_j x_j^T)^2 = 1 \), i.e.

\[ \sum_{j=1}^{N} (x_j x_j^T)^2 = N \]

Thus, taking into account the orthogonality of the memory vectors and Lemma 1, it can be deduced that memory patterns in high dimensional spaces practically coincide with the eigenvectors of the weight matrix \( W \).

It can be also shown that the learning of the Brownian weights, can be written in the form of unitary operators, as the theory of quantum mechanics predicts (Rigatos & Tzafestas, 2002).

3. Decomposition Of The Weight Matrix \( W \) into a Superposition of Matrices \( \bar{W}_i \)

In the NN models that consists of interacting Brownian particles the weights are stochastic variables. This has important consequences in the case of associative memories. Taking the weights \( w_{ij} \) of the weight matrix \( W \) to be stochastic variables with p.d.f. (or possibility distribution) as the one depicted in Fig. 1(b) means that \( W \) can be decomposed into a superposition of associative memories (see Fig. 7(a)). In that case, the overall associative memory \( W \) equals a weighted averaging of the individual weight matrices \( \bar{W}_i \), i.e. \( W = \sum_{i=1}^{N} \mu_i \bar{W}_i \), where the nonnegative weights \( \mu_i \) are possibility values (fuzzy memberships) that indicate the contribution of each local associative memory \( \bar{W}_i \) to the aggregate outcome (Rigatos & Tzafestas 2006a). To make clear the above, a 3× 3 weight matrix of a neural associative memory is considered:

\[ W = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \]  

(34)
It is also assumed that the weights $w_{ij}$ are stochastic (fuzzy) variables, as described in Fig. 1(b). The two adjacent fuzzy sets to which the weight $w_{ij}$ belongs are denoted as $A_i$ and $A_{j-1}$.

The weights satisfy the condition $\sum_{i=1}^{N} \mu_{A_i} w_{ij} = 1$ (strong fuzzy partition). The membership of the weight $w_{ij}$ to the fuzzy set $A_i$ is denoted as $\mu_{A_i}$, while the membership to the fuzzy set $A_{j-1}$ is denoted as $1-\mu_{A_i}$. Then, the following combinations of membership values of the elements of the matrices $W$ are possible:

$$
\begin{align*}
\tilde{W}_1 & : \mu_{12}, \mu_{13}, \mu_{23} \\
\tilde{W}_2 & : \mu_{12}, \mu_{13}, 1-\mu_{23} \\
\tilde{W}_3 & : \mu_{12}, 1-\mu_{13}, \mu_{23} \\
\tilde{W}_4 & : \mu_{12}, 1-\mu_{13}, 1-\mu_{23} \\
\tilde{W}_5 & : 1-\mu_{12}, \mu_{13}, \mu_{23} \\
\tilde{W}_6 & : 1-\mu_{12}, \mu_{13}, 1-\mu_{23} \\
\tilde{W}_7 & : 1-\mu_{12}, 1-\mu_{13}, \mu_{23} \\
\tilde{W}_8 & : 1-\mu_{12}, 1-\mu_{13}, 1-\mu_{23}
\end{align*}
$$

The centers of these fuzzy sets are are shown in Fig. 3(b), and are denoted as $\alpha_{Ai}$ and $\alpha_{Aj-1}$ respectively.

The diagonal elements of the matrices $W_i$ are taken to be 0 (no self-feedback in neurons is considered), while the membership value of the element $w_{ij}$, $i=1,...,3$ is indifferent and is denoted by *. Thus the weight matrix $W$ can be decomposed into a set of superimposing matrices $\tilde{W}_i$ as shown in Fig. 7(a). The submatrices $\tilde{W}_i$ of this decomposition are as follows:

$$
\begin{align*}
\tilde{W}_1 & = \begin{bmatrix}
* & \mu_{12} & \mu_{13} \\
\mu_{12} & * & \mu_{23} \\
\mu_{13} & \mu_{23} & *
\end{bmatrix}
\begin{bmatrix}
0 & \alpha_{12} & \alpha_{13} \\
\alpha_{12} & 0 & \alpha_{23} \\
\alpha_{13} & \alpha_{23} & 0
\end{bmatrix} \\
\tilde{W}_2 & = \begin{bmatrix}
* & \mu_{12} & \mu_{13} \\
\mu_{12} & * & 1-\mu_{23} \\
\mu_{13} & 1-\mu_{23} & *
\end{bmatrix}
\begin{bmatrix}
0 & \alpha_{12} & \alpha_{13} \\
\alpha_{12} & 0 & \alpha_{23} \\
\alpha_{13} & \alpha_{23} & 0
\end{bmatrix} \\
\tilde{W}_3 & = \begin{bmatrix}
* & \mu_{12} & 1-\mu_{13} \\
\mu_{12} & * & \mu_{23} \\
1-\mu_{13} & \mu_{23} & *
\end{bmatrix}
\begin{bmatrix}
0 & \alpha_{12} & \alpha_{13} \\
\alpha_{12} & 0 & \alpha_{23} \\
\alpha_{13} & \alpha_{23} & 0
\end{bmatrix} \\
\tilde{W}_4 & = \begin{bmatrix}
* & \mu_{12} & 1-\mu_{13} \\
\mu_{12} & * & 1-\mu_{23} \\
1-\mu_{13} & 1-\mu_{23} & *
\end{bmatrix}
\begin{bmatrix}
0 & \alpha_{12} & \alpha_{13} \\
\alpha_{12} & 0 & \alpha_{23} \\
\alpha_{13} & \alpha_{23} & 0
\end{bmatrix} \\
\tilde{W}_5 & = \begin{bmatrix}
* & 1-\mu_{12} & \mu_{13} \\
1-\mu_{12} & * & \mu_{23} \\
\mu_{13} & \mu_{23} & *
\end{bmatrix}
\begin{bmatrix}
0 & \alpha_{12} & \alpha_{13} \\
\alpha_{12} & 0 & \alpha_{23} \\
\alpha_{13} & \alpha_{23} & 0
\end{bmatrix}
\end{align*}
$$
The associated \( ||L_i|| \) norms are calculated using the sum \( \sum_{i=1}^{N} \sum_{j=1}^{N} |\mu_{ij}| \), which in turn is divided by the number of the non-diagonal elements, i.e. \( N(N-1) \). The following lemma holds (Rigatos & Tzafestas, 2006a):

**Lemma 3:** The \( ||L_i|| \) of the matrices \( M_i \) (i.e. \( \sum_{i,j=1}^{N} |\mu_{ij}| \)) divided by the number of the non-diagonal elements, i.e. \( N(N-1) \) and by \( 2^{N-1} \), where the number of neurons \( N \), equals unity.

\[
\frac{1}{N(N-1)2^{N-1}} \sum_{i,j=1}^{N} |\mu_{ij}| = 1 \tag{35}
\]

**Proof:** There are \( 2^{N-1} \) couples of matrices \( M_i \). Due to the strong fuzzy partition there are always two matrices \( M_i \) and \( M_j \) with complementary elements, i.e., \( \mu(w_{ij}) \) and \( 1-\mu(w_{ij}) \). Therefore the sum of the corresponding \( L_1 \) norms \( ||M_i|| + ||M_j|| \) normalized by the number of the nonzero elements (i.e., \( N(N-1) \)) equals unity. Since there are \( 2^{N-1} \) couples of \( L_1 \) norm sums \( ||M_i|| + ||M_j|| \) it holds \( \frac{1}{2^{N-1}} \sum_{i} ||M_i|| = 1 \). This normalization procedure can be used to derive the membership values of the weight matrices \( W_i \) (Q.E.D.).

Lemma 3 enables to rewrite the decomposition of matrix \( W \) into the weight matrices \( W_i \) as:

**Figure 7.**

Fig. 7(a) Decomposition of the weight matrix \( W \) into a superposition of weight matrices \( W_i \).

Fig. 7(b) Fourier transform of Gaussians of different variances, ranging between small \( \sigma \) in diagram (a) to large \( \sigma \) in diagram (d).
The decomposition of the matrix $W$ into a group of matrices $\tilde{W}_i$ reveals the existence of non-observable attractors. According to Lemmas 1 and 2 these attractors coincide with the eigenvectors $v_i$ of the matrices $\tilde{W}_i$. Thus the patterns that can be recalled from an associative memory, are more than the ones associated with the initial matrix $W$. For an associative memory of $N$ neurons, the possible patterns become $N \times 2^N$.

It has been also shown that the transition between the vector spaces which are associated with matrices $\tilde{W}_i$ is described by unitary rotations (Rigatos & Tzafestas, 2006a). Therefore the transition from the reference system $\tilde{W}_i$ to the reference system $\tilde{W}_j$ is described by unitary operators, which is a result compatible with quantum mechanics postulates (Cohen-Tanoudji et al. 1998). This is stated in the following theorem (Rigatos & Tzafestas, 2006a):

**Theorem 1:** The rotations between the spaces which are spanned by the eigenvectors of the weight matrices $\tilde{W}_i$ are unitary operators.

**Proof:** Let $x_i, y_i, z_i$ and $x_j, y_j, z_j$ be the unit vectors of the bases which span the spaces associated with the matrices $\tilde{W}_i$ and $\tilde{W}_j$, respectively. Then a memory vector $p$ can be described in both spaces as $p = (p_s, p_{s_j}, p_{s_j})^T$ and $p = (p_{s_i}, p_{s_i}, p_{s_i})^T$. Transition from the reference system $\tilde{W}_i \rightarrow \{x_i, y_i, z_i\}$ to the reference system $\tilde{W}_j \rightarrow \{x_j, y_j, z_j\}$ is expressed by the rotation matrix $R$, i.e. $p_{\tilde{W}_j} = R \cdot p_{\tilde{W}_i}$. The inverse transition is expressed by the rotation matrix $Q$, i.e. $p_{\tilde{W}_i} = Q \cdot p_{\tilde{W}_j}$. Furthermore it is true that

$$
\begin{pmatrix}
 p_s \\
 p_{s_i} \\
 p_{s_j}
\end{pmatrix} =
\begin{pmatrix}
 x_i x_j & x_i y_j & x_i z_j \\
 y_i x_j & y_i y_j & y_i z_j \\
 z_i x_j & z_i y_j & z_i z_j
\end{pmatrix}
\begin{pmatrix}
 p_{s_j} \\
 p_{s_j} \\
 p_{s_j}
\end{pmatrix}
$$

Thus, using Eq. (36) the rotation matrices $R$ and $Q$ are given by

$$
W = \frac{\mu_{12} + \mu_{13} + \mu_{23}}{3} \begin{pmatrix}
 0 & a_{12}^A & a_{13}^A \\
 a_{12}^A & 0 & a_{23}^A \\
 a_{13}^A & a_{23}^A & 0
\end{pmatrix} + \frac{\mu_{12} + \mu_{13} - \mu_{23} + 1}{3} \begin{pmatrix}
 0 & a_{12}^A & a_{13}^A \\
 a_{12}^A & 0 & a_{23}^A \\
 a_{13}^A & a_{23}^A & 0
\end{pmatrix} + \frac{\mu_{12} - \mu_{13} + \mu_{23} + 1}{3} \begin{pmatrix}
 0 & a_{12}^A & a_{13}^A \\
 a_{12}^A & 0 & a_{23}^A \\
 a_{13}^A & a_{23}^A & 0
\end{pmatrix}
$$

$$
\begin{pmatrix}
 \mu_{12} - \mu_{13} + \mu_{23} + 1 \\
 -\mu_{12} + \mu_{13} + \mu_{23} + 2 \\
 -\mu_{12} - \mu_{13} + \mu_{23} + 2
\end{pmatrix}
$$
Attractors and Energy Spectrum of Neural Structures

\[ R = Q = \begin{pmatrix} x_i x_i & x_i y_i & x_i z_i \\ y_i x_i & y_i y_i & y_i z_i \\ z_i x_i & z_i y_i & z_i z_i \end{pmatrix} \]  \( (37) \)

Since “dot products” are commutative, from Eq. (37) one obtains \( Q = R^{-1} = R^T \). Therefore the transition from the reference system \( \bar{W}_i \) to the reference system \( \bar{W}_j \) is described by unitary operators, i.e., \( QR = R^T R = R^{-1} R = I \) (Q.E.D).

SPECTRAL ANALYSIS OF ASSOCIATIVE MEMORIES THAT FOLLOW THE Q.H.O. MODEL

Spectral analysis of associative memories with weights described by interacting Brownian particles (quantum associative memories) will be carried out following previous studies on wavelets power spectra (Addison, 2002). To find a signal’s spectrum, Fourier transform has to be carried out. For instance, the Fourier transform of Gaussian functions of different variances is given in Fig. 7(b).

Spectral analysis in quantum associative memories shows that: (i) the Gaussian basis functions in which the weight \( w_{ij} \) is decomposed express the distribution of energy with respect to the weight’s values (ii) the weights \( w_{ij} \) satisfy the principle of uncertainty.

1. Spectral Analysis of Wavelets

The Morlet wavelet is the most commonly used complex wavelet and is given by

\[ \psi(x) = \pi^{-\frac{1}{2}} (e^{i2\pi f_0 x} - e^{-\frac{(2\pi f_0)^2}{2}})e^{-\frac{x^2}{2\sigma^2}} \]  \( (38) \)

This wavelet is simply a complex wave within a Gaussian envelope. The complex sinusoidal waveform is contained in the term \( e^{2\pi f_0 x} = \cos(2\pi f_0 x) + i \sin(2\pi f_0 x) \). The real and the imaginary part of the Morlet wavelet for various central frequencies are depicted in Fig. 7(a). It can be seen that the real and the imaginary part of the wavelet differ in phase by a quarter period. The \( \frac{1}{2\pi} \) term is a normalization factor which ensures that the wavelet has unit energy.

2. Spectral Decomposition and Heisenberg Boxes

The Fourier transform of the Morlet wavelet is given by

\[ \hat{\psi}(f) = \pi^{-\frac{1}{2}} \sqrt{2\pi} e^{\frac{1}{2}(2\pi f - 2\pi f_0)^2} \]  \( (39) \)

which has the form of a Gaussian function displaced along the frequency axis by \( f_0 \). The energy spectrum (the squared magnitude of the Fourier transform) is given by (Addison, 2002)

\[ |\hat{\psi}(f)|^2 = 2\pi^2 e^{-2(2\pi f - 2\pi f_0)^2} \]  \( (40) \)

which is a Gaussian centered at \( f_0 \). The integral of Eq. (40) gives the energy of the Morlet wavelet. The energy spectrum of the Morlet wavelet depicted in Fig. 8(a)-diagram (c), for different values of the variance \( \sigma^2 \) is given in Fig. 8(b).

The central frequency \( f_0 \) is the frequency of the complex sinusoid and its value determines the number of significant sinusoidal waveforms contained within the envelope. The dilated and translated Morlet wavelet
Attractors and Energy Spectrum of Neural Structures

Figure 8.

\[ \psi \left( \frac{x - b}{a} \right) \] is given by

\[ \psi \left( \frac{x - b}{a} \right) = \pi^{-\frac{1}{4}} e^{-\frac{\left( 2\pi f_0 \right)^2}{a^2}} e^{-\frac{1}{2} \frac{x^2}{a^2}} \]  

(41)

A tool that enables to visualize the space-frequency characteristics of signals is the so-called Heisenberg box.

The Heisenberg boxes in the x-frequency plane for a wavelet at different scales, are shown in Fig. 9(a). To evaluate frequency composition a sample of a long region of the signal is required. If instead, a small region of the signal is measured with accuracy, then it becomes very difficult to determine the frequency content of the signal in that region. That is, the more accurate the temporal measurement (smaller \( \sigma_x \)) is, the less accurate the spectral measurement (larger \( \sigma_f \)) becomes, and vice-versa (Addison, 1998).

The central frequency \( f_0 \) sets the location of the Heisenberg box in the x-frequency plane. If the x-length of the wavelets remains the same, then no matter the change of the central frequency \( f_0 \) the associated Heisenberg boxes will have the same dimensions. This is depicted in Fig. 9(b).

Finally, in Fig. 10(a) are shown the Heisenberg boxes in the x-frequency plane for a number of wavelets with three different spectral frequencies (low, medium and high). The confining Gaussian windows have the same dimensions along the x axis. Therefore, altering the central frequency of the wavelet shifts the Heisenberg box up and down the x-frequency plane without altering its dimensions.

3. Energy Spectrum of Stochastic Weights that Follow the Q.H.O. Model

It is assumed that the stochastic weight \( w_{ij} \) is described by the probability (possibility) distributions shown in Fig. 1(b). Then the following theorem holds (Rigatos, 2006).

**Theorem 2:** The Gaussian basis functions of the weights \( w_{ij} \) of a quantum associative memory express the distribution of energy with respect to the value of \( w_{ij} \). The smaller the spread \( \sigma \) of the basis functions is, the larger becomes the spectral (energy) content that can be captured therein.

**Proof:** The Fourier transform of \( g(x) = e^{-\frac{x^2}{2}} \) is \( G(f) = \sqrt{\frac{\pi}{2}} e^{-\frac{f^2}{4}} \). Furthermore the mean value of the stochastic

396
**Attractors and Energy Spectrum of Neural Structures**

**Figure 9.**

Fig. 9(a) Heisenberg boxes in the $x$-frequency plane for a wavelet at various scales $f_1, f_2, f_3$ and of different length $\sigma_x$ along the $x$-axis. $\sigma_x$ is the standard deviation of the squared wavelet function $\psi(x)^2$; $\sigma_f$ is the standard deviation of the spectrum around the mean spectral components $f_1, f_2$ and $f_3$

Fig. 9(b) Heisenberg boxes in the $x$-frequency plane for wavelet at various scales $f_1, f_2, f_3$ and of the same length $\sigma_x$ along the $x$-axis. Changes of the central frequency $f_0$ do not affect the size of the Heisenberg boxes.

**Figure 10.**

Fig. 10(a) Heisenberg boxes in the $x$-frequency plane for a number of superimposing wavelets of the same length $\sigma_x$ along the $x$-axis and at three different spectral frequencies set to low, medium and high value

Fig. 10(b) Heisenberg boxes in the $x$-frequency plane for a number of Gaussian basis functions which have the same variance and which are shifted on the $x$-axis
weight $< w_j >$ is described by shifted overlapping Gaussians, i.e.,

\[ < w_j > = \sum_{k=1}^{n} e^{-\frac{(w_j - k\sigma)^2}{2\sigma^2}} a_k. \]

Consequently, the associated Fourier transform will be:

\[
G(\sum_{k=1}^{n} e^{-\frac{(w_j - k\sigma)^2}{2\sigma^2}} a_k) = \sum_{k=1}^{n} k \cdot c \cdot e^{jfk} \sqrt{2\pi} e^{-\frac{j^2 s^2}{2\sigma^2}}
\]

where $a_k = k \cdot c$ with $c$ being the distance between the centers of two adjacent fuzzy basis functions. Keeping the real part of $G(f)$ is kept one obtains

\[
G(\sum_{k=1}^{n} e^{-\frac{(w_j - k\sigma)^2}{2\sigma^2}} a_k) ≈ \sum_{k=1}^{n} k \cdot c \cdot \cos(fk) \sqrt{2\pi} e^{-\frac{j^2 s^2}{2\sigma^2}}
\]

From Eq. (43) it can be seen that if the weight $w_j$ is decomposed in the $x$ domain into shifted Gaussians of the same variance $\sigma^2$ then, in the frequency domain, $w_j$ is analyzed in chirps, like the ones given in Fig. 10(b). The energy of the stochastic weights can now be found using Rayleigh’s theorem. This states that the energy of a signal $f(x)$, $x \in [-\infty, +\infty]$ is given by $\int f^2(x) dx$. Equivalently, using the Fourier transform $F(s)$ of $f(x)$, the energy is given by $\int F^2(s) ds$. The energy distribution of the quantum particle is proportional to the probability $|\psi(w_j)|^2$ of finding the particle between $w_j$ and $w_j + \Delta w_j$. Therefore, the energy of the particle will be given by the integral of the squared Fourier transform $|\psi(f)|^2$. Thus, basis functions of small spread $\sigma$ in the $x$ domain, become ample functions in the frequency domain (see for instance Fig. 7(b)) which defines the associated energy content (Q.E.D).

The Fourier transform of the weights of the quantum associative memories is depicted in the diagram given in the bottom of Fig. 10(b), and the associated spectrum consists again of Gaussian functions.

4. Weights that Follow the Q.H.O. Model and the Principle of Uncertainty

The significance of Eq. (43) is that the product between the information in the space domain $g(w_j)$ and the information in the frequency domain $G(s)$ cannot be smaller than a constant. In the case of a Gaussian function $g(w_j) = e^{-\frac{w_j^2}{2\sigma^2}}$ with Fourier transform $G(s) = e^{-\frac{1}{2\sigma^2}s^2}$ it can be observed that: (i) if $\sigma$ is small then $g(w_j)$ has a peak at $w_j = 0$ while $G(s)$ tends to become flat, (ii) if $\sigma$ is large then $g(w_j)$ is flat at $w_j = 0$ while $G(s)$ makes a peak at $s = 0$. These can be observed in Fig. 9(a).

This becomes more clear if the dispersion of function $g(w_j) = e^{-\frac{w_j^2}{2\sigma^2}}$ round $w_j = 0$ is used (Pinsky, 1991). The dispersion of $g(w_j)$ and of its Fourier transform $G(s)$ becomes

\[
D(g) = \frac{\int_{-\infty}^{\infty} w_j^2 e^{-\frac{w_j^2}{2\sigma^2}} dw_j}{\int_{-\infty}^{\infty} e^{-\frac{w_j^2}{2\sigma^2}} dw_j} = \frac{1}{2\sigma^2}, \quad D(G) = \frac{\int_{-\infty}^{\infty} s^2 e^{-\frac{s^2}{2\sigma^2}} ds}{\int_{-\infty}^{\infty} e^{-\frac{s^2}{2\sigma^2}} ds} = \frac{1}{2\sigma^2}
\]

which results into the uncertainty principle for the weights of quantum associative memories

\[
D(g)D(G) = 1/4
\]
Eq. (45) means that the accuracy in the calculation of the weight \( w_{ij} \) is associated with the accuracy in the calculation of its spectral content. When the spread of the Gaussians of the stochastic weights is large (small) then their spectral content is poor (rich). Eq. (45) is an analogous of the quantum mechanics uncertainty principle, i.e., \( \Delta x \Delta p \geq \hbar \), where \( \Delta x \) is the uncertainty in the measurement of particle’s position, \( \Delta p \) is the uncertainty in the measurement of the particle’s momentum and \( \hbar \) is Planck’s constant.

It should be noted that Eq. (45) expresses a general property of the Fourier transform and that similar relations can be found in classical physics. For instance in electromagnetism it is known that it is not possible to measure with arbitrary precision, at the same time instant, the variation of a wave function both in the time and frequency domain. What is really quantum in the previous analysis, is the association of a wave function with a particle (stochastic weight) and the assumption that the wave length and the momentum of the particle satisfy a relation equivalent to \( p = \hbar k \) with \( |k| = \frac{2\pi}{\lambda} \).

SIMULATION TESTS

1. **Convergence Of The Stochastic Weights To An Equilibrium**

The update of the stochastic weights is given by the gradient algorithm, of Eq. (20). The objective is to make the stochastic weights (particles) \( \tilde{w}_{ij} \) converge simultaneously to the desirable value \( \tilde{w}_{ij}^* \), or equivalently make the errors of the weight values \( e_{ij} = w_{ij} - \tilde{w}_{ij}^* \) converge simultaneously to the attractor \([e^*, e^*] = [0, 0] \).

In the conducted simulation experiments the multi-particle system consisted of \( N \) particles (weights) which were randomly initialized in the 2-D field \([x, y] = [e, e] \). Two cases were distinguished: (i) update without constraints (Fig.10) and (ii) update under constraints (Fig. 12). In the latter case there were areas in the 2-D plane which could not be accessed by the trajectory of the particles towards the equilibrium.

As expected, the relative values of the parameters \( a \) and \( b \) that appear in the term \( g(x' - x') \) of Eq. (21), affect the trajectories of the individual particles. For \( a > b \) the cohesion of the particles was maintained and abrupt displacements of the particles were avoided.

For weights update without constraints the Lyapunov function of each stochastic weight (individual particle), is shown in Fig. 13(a). The aggregate Lyapunov function is in Fig. 13(b).

Finally, in case of weights update under constraints the Lyapunov function of the individual particles, is depicted in Fig. 14(a). The aggregate Lyapunov function is shown in Fig. 14(b).
Figure 12.

Fig. 12(a) Convergence of the individual neural weights that follow the Q.H.O. model to an attractor assuming constraints.

Fig. 12(b) Convergence of the mean of the weights position to the attractor, assuming constraints.

Figure 13.

Fig.13(a) Lyapunov function of the individual stochastic weights (Brownian particles) in a 2D-attractors plane without constraints.

Fig. 13(b) Lyapunov function of the mean of the stochastic weights (multi-particle system) in a 2D-attractors plane without constraints.

Figure 14.

Fig.14(a) Lyapunov function of the individual stochastic weights (Brownian particles) in a 2D-attractors plane with constraints (there are prohibited regions exist).

Fig. 14(b) Lyapunov function of the mean of the stochastic weights (multi-particle system) in a 2D-attractors plane with constraints (there are prohibited regions).
Attractors and Energy Spectrum of Neural Structures

It should be noted that the difference between the neural structures that follow the Q.H.O model and neural structures that follow Schrödinger’s equation with zero or constant potential, is that convergence to an attractor is controlled by the drift force imposed by the harmonic potential. A particle can be steered to an attractor, through the drift force, which in turn is tuned by the parameters \(a\) and \(b\) of Eq. (21), or through an external potential.

2. Attractors in Associative Memories That Follow The Q.H.O. Model

The theoretical results about the increased number of attractors in associative memories that follow the Q.H.O. model will be verified through a numerical example.

a). Superposition of weight matrices:

Assume that the fundamental memory patterns are the following binary vectors \(s_1 = [1,1,1], s_2 = [1,1,-1], s_3 = [1,-1,1]\) which are linearly independent but not orthogonal. Orthogonality should be expected in high dimensional vector spaces, if the elements of the memory vectors are chosen randomly. In this example, to obtain orthogonality of the memory vectors, Gramm-Schmidt orthogonalization is used. This gives the orthogonal vectors \(u_1 = [1,1,1], u_2 = [2/3,2/3,-4/3], u_3 = [1,-1,0]\). The weight \(W\) which derived from the above memory patterns is \(W = (1/3)[u_1^T u_1 + u_2^T u_2 + u_3^T u_3]\), i.e.

\[
W = \begin{bmatrix}
0.8141 & 0.1481 & 0.0369 \\
0.1481 & 0.8141 & 0.0369 \\
0.0369 & 0.0369 & 0.9247 \\
\end{bmatrix}
\]

(46)

It can be easily shown that \(W_{u_1} = u_1, W_{u_2} = u_2, W_{u_3} = u_3\) i.e. \(u_1, u_2, u_3\) are stable states (attractors) of the network. The eigenvalues of matrix \(W\) are \(\lambda_1 = 0.667, \lambda_2 = 0.888\) and \(\lambda_3 = 1.0\). The associated eigenvectors of \(W\) are \(v_1 = [0.7071, -0.7071, 0]^T, v_2 = [-0.4066, -0.4066, 0.8181]^T\) and \(v_3 = [0.5785, 0.5785, 0.5750]^T\). It can be observed that \(v_1\) is collinear to \(u_2\), \(v_2\) is collinear to \(u_3\), and \(v_3\) is collinear to \(u_1\).

Next, the elements of the weight matrix \(W\) are considered to be stochastic variables, with p.d.f. (possibility distribution) as the one depicted in Fig. 1(b). The universe of discourse of these fuzzy variables is shown in Fig. 3(b). Thus matrix \(W\), given by Eq. (44), can be decomposed into a superposition of weight matrices \(W_i\). Assume that only the non-diagonal elements of \(W\) are considered and that the possibility distribution of the stochastic variables \(w_{ij}\) is depicted in Fig. 1(b). Then, the weight matrix \(W\) is decomposed into a superposition of weight matrices \(W_i, i = 1, ..., 8:\)

\[
W = \begin{bmatrix}
* & 0.405 & 0.155 \\
0.405 & * & 0.155 \\
0.155 & 0.155 & * \\
\end{bmatrix}
\begin{bmatrix}
0 & 0.14 & 0.02 \\
0.14 & 0 & 0.02 \\
0.02 & 0.02 & 0 \\
\end{bmatrix}
+ \begin{bmatrix}
* & 0.405 & 0.155 \\
0.405 & * & 0.845 \\
0.155 & 0.845 & * \\
\end{bmatrix}
\begin{bmatrix}
0 & 0.14 & 0.04 \\
0.14 & 0 & 0.04 \\
0.02 & 0.04 & 0 \\
\end{bmatrix}
+ \begin{bmatrix}
* & 0.405 & 0.155 \\
0.405 & * & 0.845 \\
0.155 & 0.845 & * \\
\end{bmatrix}
\begin{bmatrix}
0 & 0.16 & 0.02 \\
0.16 & 0 & 0.02 \\
0.02 & 0.02 & 0 \\
\end{bmatrix}
+ \begin{bmatrix}
* & 0.595 & 0.155 \\
0.595 & * & 0.845 \\
0.155 & 0.845 & * \\
\end{bmatrix}
\begin{bmatrix}
0 & 0.16 & 0.04 \\
0.16 & 0 & 0.04 \\
0.02 & 0.04 & 0 \\
\end{bmatrix}
+ \begin{bmatrix}
* & 0.595 & 0.155 \\
0.595 & * & 0.845 \\
0.155 & 0.845 & * \\
\end{bmatrix}
\begin{bmatrix}
0 & 0.16 & 0.02 \\
0.16 & 0 & 0.02 \\
0.04 & 0.04 & 0 \\
\end{bmatrix}
\]
The membership $\mu_i$ of each matrix $W_i$ is taken to be the normalized $\| L_i \|$ of the matrix with elements the membership values of the weights $w_{ij}$, i.e., $\frac{1}{N \cdot (N - 1)} \sum_{j \neq i} \mu(w_{ij})$. This gives $W = \mu_1 W_1 + \mu_2 W_2 + \ldots + \mu_s W_s$, where the membership $\mu_i$ are: $\mu_1 = 0.0596$, $\mu_2 = 0.1171$, $\mu_3 = 0.1171$, $\mu_4 = 0.1746$, $\mu_5 = 0.0754$, $\mu_6 = 0.1329$, $\mu_7 = 0.1329$ and $\mu_8 = 0.1904$. By calculating the eigenvectors of matrices $W_i$, the associated memory patterns can be found. These are non-observable attractors different from the attractors $u_i$, $u_2$ and $u_3$ of the initial weight matrix $W$. Thus, the number of memory patterns is increased by a factor $2^8 = 8$.

Indeed:

The eigenstructure analysis of matrix $\bar{W}_1$ gives: $\lambda_1 = -0.14$, $\lambda_2 = 0.1455$, $\lambda_3 = -0.0055$, with associated eigenvectors $v_1^{W_1} = \begin{bmatrix} 0.7071 \\ -0.7071 \\ 0 \end{bmatrix}$, $v_2^{W_1} = \begin{bmatrix} 0.6941 \\ 0.6942 \\ 0.1908 \end{bmatrix}$ and $v_3^{W_1} = \begin{bmatrix} 0.1349 \\ 0.1349 \\ -0.9816 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_2$ gives: $\lambda_1 = -0.1415$, $\lambda_2 = -0.0104$, $\lambda_3 = 0.1519$, with associated eigenvectors $v_1^{W_2} = \begin{bmatrix} 0.6921 \\ -0.7143 \\ 0.1041 \end{bmatrix}$, $v_2^{W_2} = \begin{bmatrix} -0.2648 \\ -0.1176 \\ 0.9572 \end{bmatrix}$ and $v_3^{W_2} = \begin{bmatrix} 0.6715 \\ 0.6900 \\ 0.2701 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_3$ gives: $\lambda_1 = -0.1415$, $\lambda_2 = 0.00104$, $\lambda_3 = 0.1519$, with associated eigenvectors are $v_1^{W_3} = \begin{bmatrix} 0.1415 \\ 0.1415 \\ -0.00104 \end{bmatrix}$, $v_2^{W_3} = \begin{bmatrix} -0.1170 \\ -0.2648 \\ 0.9572 \end{bmatrix}$ and $v_3^{W_3} = \begin{bmatrix} 0.6715 \\ 0.6900 \\ 0.2701 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_4$ gives: $\lambda_1 = -0.14$, $\lambda_2 = 0.16$, $\lambda_3 = -0.02$, with associated eigenvectors $v_1^{W_4} = \begin{bmatrix} 0.7071 \\ -0.7071 \\ 0 \end{bmatrix}$, $v_2^{W_4} = \begin{bmatrix} 0.6667 \\ 0.6667 \\ 0.3333 \end{bmatrix}$ and $v_3^{W_4} = \begin{bmatrix} 0.2357 \\ 0.2357 \\ -0.9428 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_5$ gives: $\lambda_1 = 0.16$, $\lambda_2 = 0.1649$, $\lambda_3 = 0.0049$, with associated eigenvectors $v_1^{W_5} = \begin{bmatrix} 0.6921 \\ 0.6969 \\ 0.1691 \end{bmatrix}$ and $v_3^{W_5} = \begin{bmatrix} 0.1196 \\ 0.1196 \\ 0.9856 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_6$ gives: $\lambda_1 = -0.1613$, $\lambda_2 = -0.0093$, $\lambda_3 = 0.1706$, with associated eigenvectors $v_1^{W_6} = \begin{bmatrix} 0.6957 \\ -0.7126 \\ 0.0905 \end{bmatrix}$, $v_2^{W_6} = \begin{bmatrix} -0.2353 \\ -0.1071 \\ 0.9660 \end{bmatrix}$ and $v_3^{W_6} = \begin{bmatrix} 0.6787 \\ 0.6933 \\ 0.2421 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_7$ gives: $\lambda_1 = -0.1613$, $\lambda_2 = 0.0093$, $\lambda_3 = 0.1706$, with associated eigenvectors $v_1^{W_7} = \begin{bmatrix} -0.7126 \\ 0.6957 \\ 0.0905 \end{bmatrix}$, $v_2^{W_7} = \begin{bmatrix} -0.1071 \\ -0.2353 \\ 0.9660 \end{bmatrix}$ and $v_3^{W_7} = \begin{bmatrix} -0.6787 \\ -0.6933 \\ 0.2421 \end{bmatrix}$.

The eigenstructure analysis of matrix $\bar{W}_8$ gives: $\lambda_1 = 0.1600$, $\lambda_2 = 0.1780$, $\lambda_3 = -0.0180$, with associated eigenvectors $v_1^{W_8} = \begin{bmatrix} 0.7071 \\ -0.7071 \\ 0 \end{bmatrix}$, $v_2^{W_8} = \begin{bmatrix} 0.6739 \\ 0.6739 \\ 0.3029 \end{bmatrix}$, and $v_3^{W_8} = \begin{bmatrix} 0.2142 \\ 0.2142 \\ -0.9530 \end{bmatrix}$.

b) Unitarity of the rotation operators:

Here it will be verified that the transition between matrices $\bar{W}_1$ and $\bar{W}_2$ in which the weight matrix $W$ of the associative memory is decomposed, is described by unitary operators. Take for instance matrices $\bar{W}_1$ and $\bar{W}_2$ of the previous example. The matrix $R$ which performs a rotation from the basis defined by the eigenvectors $v_1^{W_1}$, $v_2^{W_1}$, $v_3^{W_1}$, to the basis defined by the vectors $v_1^{W_2}$, $v_2^{W_2}$, $v_3^{W_2}$ is calculated as follows: $v_1^{W_2} = \begin{bmatrix} 0.7071 \\ -0.7071 \\ 0 \end{bmatrix}$, $v_2^{W_2} = \begin{bmatrix} 0.6941 \\ 0.6941 \\ 0.1908 \end{bmatrix}$ and $v_3^{W_2} = \begin{bmatrix} 0.1349 \\ 0.1349 \\ -0.9816 \end{bmatrix}$, while $v_1^{W_2} = \begin{bmatrix} 0.6921 \\ -0.7143 \\ 0.1041 \end{bmatrix}$, $v_2^{W_2} = \begin{bmatrix} -0.2648 \\ -0.1176 \\ 0.9572 \end{bmatrix}$ and $v_3^{W_2} = \begin{bmatrix} 0.6715 \\ 0.6900 \\ 0.2701 \end{bmatrix}$.

The rotation matrix $R$ is given by

$$ R = \begin{bmatrix} v_1^{W_1} & v_2^{W_1} & v_3^{W_1} \\ v_1^{W_2} & v_2^{W_2} & v_3^{W_2} \\ v_1^{W_3} & v_2^{W_3} & v_3^{W_3} \end{bmatrix} = \begin{bmatrix} 0.9945 & -0.1041 & -0.0131 \\ 0.0045 & -0.0752 & 0.9966 \\ -0.1052 & -1.0304 & -0.0815 \end{bmatrix} $$
and it holds that

\[
RR^T = \begin{bmatrix}
1.000 & -0.0008 & 0.0037 \\
-0.0008 & 0.9989 & -0.0042 \\
0.0037 & -0.0042 & 1.0794
\end{bmatrix} \Rightarrow RR^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

Thus, matrix \( R \) represents a unitary rotation from the vector space \( v_1^{\tilde{w}}, v_2^{\tilde{w}}, v_3^{\tilde{w}} \) to the vector space \( v_1^{\tilde{w}_i}, v_2^{\tilde{w}_i}, v_3^{\tilde{w}_i} \). On the other hand matrix \( R^T \) represents a unitary rotation from the vector space \( v_1^{\tilde{w}_i}, v_2^{\tilde{w}_i}, v_3^{\tilde{w}_i} \) to the vector space \( v_1^{\tilde{w}}, v_2^{\tilde{w}}, v_3^{\tilde{w}} \).

**CONCLUSION**

The objective of this chapter was to study neural structures with weights which are stochastic variables and which follow the model of the quantum harmonic oscillator (Q.H.O.). The neural weights were taken to correspond to diffusing particles, which interact to each other as the theory of Wiener process (Brownian motion) predicts. The values of the weights are the positions of the particles and the probability density function \( |\psi(x,t)|^2 \) that describes their position is derived by Schrödinger’s equation. Therefore the dynamics of the neural network is given by the solution of Schrödinger’s equation under a parabolic (harmonic) potential. Assuming a probability density function that depends only on the ground state of the Q.H.O., i.e. \( \rho_0(x) = |\Psi_0(x)|^2 \), the solution of Schrödinger’s equation was approximated by a stationary diffusion.

Next, it was shown that in neural structures with weights that follow the Q.H.O. model, the weights update is described by Langevin’s stochastic differential equation. It was proved that conventional gradient algorithms are a subcase of Langevin’s equation. Moreover, it was demonstrated that the weights which follow the Q.H.O. model give to associative memories significant properties: (i) the learning of the stochastic weights is a Wiener process, (ii) the number of attractors increases exponentially comparing to conventional associative memories.

Spectral analysis of the weights of that follow the Q.H.O. model was also carried out, based on previous studies of wavelets’ energy spectrum. It was shown that: (i) the Gaussian basis functions of the weights express the distribution of the energy with respect to the weights’ value. The smaller the spread of the basis functions is, the larger becomes the spectral (energy) content that can be captured therein. (ii) The stochastic weights satisfy an equation which is analogous to the principle of uncertainty.

Furthermore, simulation tests were performed to demonstrate that (i) the weights which follow the Q.H.O. model converge to attractors, (ii) the storage capacity of the Q.H.O-based associative memories increases in an exponential way. Finally, it was pointed out that the basic difference of neural weights that follow the Q.H.O. model from neural weights which are based on Schrödinger’s equation with zero or constant potential, is that convergence to one of the possible attractors is controlled by the drift force that is imposed by the harmonic potential.

**FUTURE RESEARCH DIRECTIONS**

Quantum information processing in neural structures is a promising research topic. Relevant areas are wavelets and wavelet networks, neural structures using the eigenstates of the quantum harmonic oscillator, harmonic analysis, while future research directions could be: (i) multi-scale information processing, (ii) control of multi-particle systems and perspectives for emerging fields, such as nanotechnology, biotechnology and quantum computing, (ii) theoretical issues, such as computation models, limits of multi-scale analysis and uncertainty principle, (iii) applications in image processing (iv) applications in communication systems, (v) application in biomedical systems.
To implement the proposed quantum neural structures, learning issues have to be examined. The learning in neural networks compatible with principles of quantum mechanics is a diffusion process. Thus learning is associated to the control of the diffusion stochastic differential equations. The modelling of the probability density functions of quantum diffusions can be carried out with the use of wavelet functions, Hermite functions of fuzzy basis functions. To succeed control of the quantum diffusions, open-loop control approaches can be examined. On the other hand closed-loop control methods for particles at micro and nano scale can be based on the theory of optimal control or robust control approaches. Feedback-based control methods require knowledge of the complete state vector that describes the particle’s motion and in case that the state vector is not completely measurable this has to be reconstructed with the use of a filter or state estimator.

REFERENCES


Attractors and Energy Spectrum of Neural Structures


Attractors and Energy Spectrum of Neural Structures


Attractors and Energy Spectrum of Neural Structures


ADDITIONAL READING


