On Wavelet based modeling of Neural networks using Graph theoretic approach

Bharat Bhosale[1], Bouthina Ahmed[2], Anjan Biswas[3, 4*]

[1] S H Kelkar College of Arts, Commerce and Science, University of Mumbai, Devgad (M.S.), India
[2] Department of Mathematics, Girls' College, Ain Shams University, Cairo, Egypt
[3] Department of Mathematical Sciences, Delaware State University, Dover,19901-2277, USA
[4] Department of Mathematics, Faculty of Science King Abdulaziz University, Jeddah, Saudi Arabia
biswas.anjan@gmail.com

Abstract: A graph is an abstract representation of complex network. Many types of relations and process dynamics in physical, biological, social and information systems can be modeled with graphs. Graph analysis has been used in the study of models of neural networks, anatomical connectivity, and functional connectivity. These developments in the theory of complex networks have inspired new applications in the upcoming field of neuroscience. In our work, we propose a novel wavelet based neural network stochastic model that extends existing methods for processing the data represented in graph domain. Our approach is based on defining random walks on arbitrary infinite graphs representing neural networks, which itself is a stochastic process characterized by some probability distribution. More so, random walks exhibit fractal-like patterns that, in turn, attribute the use of wavelet methods. The robustness of the proposed model as against the existing ones has been justified by highlighting the potential applications in neuroscience.

[Bharat Bhosale, Bouthina Ahmed, Anjan Biswas**, On Wavelet based modeling of Neural networks using Graph theoretic approach.** *Life Sci. J.* 2013; 10(2):1509-1515]. (ISSN: 1097-8135). http://www.lifesciencesite.com. 204

Keywords: wavelet transform, graph neural networks, random process, neuroscience

1. INTRODUCTION

Graph theory pertains to the study of *graphs* which are mathematical structures used to model pairwise relations between objects. A graph structure consists of vertices/nodes and edges. Graphs can be further extended by assigning weights to each edge of the graph called weighted graphs. A directed graph or digraph with weighted edges is called a network. Network analysis have many practical applications, for example, to model and analyze traffic networks. These applications can be classified in to three categories: First, analysis to determine structural properties of a network, such as the distribution of vertex degrees and the diameter of the graph. Second, analysis to find a measurable quantity within the network, for example, for a transportation network, the level of vehicular flow within any portion of it, and the third is analysis of dynamical properties of networks.

Many practical problems can be represented by graphs that can be used to model different types of relations and process dynamics in physical, social and information systems. In mathematics, graphs are useful in geometry and certain parts of topology, e.g. Knot Theory. Algebraic graph theory has close links with group theory. In computer science, graphs are used to represent networks of communication, data organization, computational devices, the flow of computation, etc. Graph theory is also used to study molecules in chemistry and physics. In condensed matter physics, the three dimensional structure of complicated simulated atomic structures can be studied quantitatively by gathering statistics on graph-theoretic properties related to the topology of the atoms. For example, Franzblau's shortest-path (SP) rings. In chemistry, a graph makes a natural model for a molecule, where vertices represent atoms and edges bonds. This approach is especially used in computer processing of molecular structures, ranging from chemical editors to database searching. In statistical physics, graphs can represent local connections between interacting parts of a system, as well as the dynamics of a physical process on such systems. Likewise, graph theory is useful in biology and conservation efforts where a vertex can represent regions where certain species exist (or habitats) and the edges represent migration paths, or movement between the regions. This information is important when looking at breeding patterns or tracking the spread of disease, parasites or how changes to the movement can affect other species. In sum, many underlying relationships among data in several areas of science and engineering, e.g., computer vision, molecular chemistry, molecular biology, pattern recognition, and data mining, can be represented in terms of graphs and the graph theoretic approach can be employed to analyse such complex data.

In recent years many important properties of complex networks have been delineated and studied

the relationship between the structural properties, nature of dynamics taking place on these networks. For instance, the 'synchronizability' of complex networks of coupled oscillators can be determined by graph spectral analysis. These developments in the theory of complex networks have inspired new applications in the field of neuroscience such as study of models of neural networks, anatomical connectivity, and functional connectivity based upon functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and magnetoencephalography (MEG). The recent applications of network theory to neuroscience such as modelling of neural dynamics on complex networks; graph theoretical analysis of neuroanatomical networks; and applications of graph analysis to studies of functional connectivity with fMRI, EEG and MEG are discussed at length [18]. The survey of the recent literature reports several models used for analyzing complex networks. The structure of the yeast protein complex network in which weighted edges between complexes represent the number of shared proteins is studied by Masagh et al.[13]. It is reported that the network of protein complexes is a small world network with scale free behavior for many of its distributions, that the human brain can be modeled as a complex network, and may have a small-world structure both at the level of anatomical as well as functional connectivity.

The complex networks like neural networks [20] are also represented with the help of graphs by showing the computational elements, neurons of the network. Each node corresponds to one neuron and the arrows usually denote weighted sums of the values from other neurons.

Recently various approaches have been unified in neural network model called graph neural networks (GNNs) [16], which is used for processing the data represented in graph domains. It is demonstrated that the GNN can be used to process structured data inputs, e.g. acyclic graph, cyclic graph, directed or un-directed graphs. The GNN model is particularly suited for problems whose domain can be represented by a set of patterns and relationships between them. In those problems, a prediction about a given pattern can be carried out exploiting all the related information, which includes the pattern features, the pattern relationships and, in general, the whole graph that represents the domain. GNN peculiarity consists in its capability of making the prediction taking directly in input the domain graph, without any preprocessing. GNNs have been proved to be sort of universal approximator for functions on graphs and have been applied to several problems, including spam detection, object localization in images, molecule classification.

Moreover, the GNN are of two kinds viz. Biological neural networks (BNN) and Artificial neural networks (ANN) [15]. BNN are objective existence, in which the neurons are linked as a network in a certain order, e.g. human neural network is the most intelligent network system. The ANN are aimed at modeling the organization principles of central neural system, with the hope that the biologically inspired computing capabilities of ANN will allow the cognitive and sensory tasks to be performed more easily and satisfactorily. ANN models some structure, characters and functions of BNN by electronical and optical technologies [9].

Wavelets have been used in analyzing graphs and the complex networks represented by graphs. Hammond [7] proposed a method for constructing wavelet transforms of functions defined on the vertices of an arbitrary finite weighted graph. The approach is based on defining scaling using the graph analogue of the Fourier domain, namely the spectral decomposition of the discrete graph Laplacian and forming the spectral graph wavelets. Wavelet transform has been employed to analyze neural networks [19]. Useful wavelet applications in turbulence onset, spectrum analyses, fractal aggregates, and bubble-chamber particle-track pattern-recognition problems are indicated.

Since the introduction of wavelet theory for square integrable functions defined on the real line, numerous authors have introduced extensions and related transforms for signals on the plane and higher dimensional spaces. By taking separable products of one dimensional wavelets, one can construct orthogonal families of wavelets in any dimension [8], [12], [17]. A comprehensive review of the potential of state of the art wavelets, and in particular wavelet statistical methodology, in different areas of molecular biology such as genome sequence, protein structure and microarray data analysis is given in [10]. In our earlier works also, we studied the wavelet interaction with solitons arising as the solutions of non-linear partial differential equations viz. Non-linear Schrodinger Equation, Sine-Gordon equation, Korteweg-de Vries equation [1], [3], [4], [6]. Also, we studied extensively the strong relationships existing between wavelets, solitons and probability distributions [2]. Moreover, we studied the wavelet interaction to random processes, wavelet based analysis of genomic sequences [5].

We confine our discussion to neural networks, its representation in terms of graphs, and its mathematical modeling. We can define random walks on the digraphs representing neural networks. Random walks themselves are the stochastic processes that can be analysed by employing wavelet methods. This approach has not seen exploited so far

in the reported investigations on the various studies on graph neural networks. This has motivated us to undertake the present study.

In this work, we combine wavelet transform and graph theory for feature extraction and visualization of structures of neural networks especially in stochastic framework. This is an extension of wavelet methods proposed by us for genomic sequence analysis [5].

2. Notations and Terminologies

Graphs: In the most common sense of the term, a **graph** is an ordered pair $G = (E, V)$ comprising a set V of vertices or nodes together with a set E of edges or lines, which are 2-element subsets of V (i.e., an edge is related with two vertices, and the relation is represented as unordered pair of the vertices with respect to the particular edge). A vertex may exist in a graph and not belong to an edge. The presence of an edge between two vertices indicates the presence of some kind of interaction or connection between the vertices (the interpretation depends upon what is being modeled with the graph). The order of a graph $|V|$, is the number of vertices. A graph's size is $|E|$ the number of edges. The degree of a vertex is the number of edges that connect to it, where an edge that connects to the vertex at both ends (a loop) is counted twice.

As an extension of the simple graph is a weighted graph $G = (E, V, w)$ that consists of a set of vertices, a set of edges E, and a weight function $w : E \to R+$ which assigns a positive weight to each edge. The adjacency matrix A for a weighted graph G is the $N \times N$ matrix that contains the information about the connectivity structure of the graph. When an edge exists between two vertices i and j, the corresponding entry of the adjacency matrix is $A_{ij} = 1$; otherwise $A_{ij} = 0$. The likelihood $P(k)$ that a randomly chosen vertex will have degree k is given by the degree distribution: it is a plot of $P(k)$ as a function of k. The degree distribution can have different forms: Gaussian, binomial, Poisson, exponential or power law.

For a weighted graph, the degree of each vertex m, written as $d(m)$, is defined as the sum of the weights of all the edges incident to it. This implies $d(m) = \sum_n A_{m,n}$.

Every real valued function $f : V \to R$ on the vertices of the graph G can be viewed as a vector in R^N, where the value of f on each vertex defines each coordinate. This implies an implicit numbering of the vertices.

Random Walk: A random walk is a mathematical formalization of a path that consists of a succession of random steps. Various different types of random walks are of interest. Often, random walks are assumed to be Markov chains or Markov processes, but other, more complicated walks are also of interest. Some random walks are on graphs, others on the line, in the plane, or in higher dimensions, while some random walks are on groups. Random walks also vary with regard to the time parameter. A random walk of length k on a possibly infinite graph G with a root 0 is a stochastic process with random variables $X_1, X_2, X_3, \ldots, X_k$ such that $X_1 = 0$ and $X_{i=1}$ is a vertex chosen uniformly at random from the neighbors of X_i. Then the number $P_{v,w,k}(G)$ is the probability that a random walk of length k starting at v ends at w. Lovasz [11] surveyed various aspects of the theory of random walks on graphs. In particular, estimates on the important parameters of access time, commute time, cover time and mixing time are discussed. Connections with the eigenvalues of graphs and with electrical networks, and the use of these connections in the study of random walks are elaborated.

Wavelet Transform: The wavelet transform (WT) is a decomposition of a function, $f(x)$, with respect to a basic wavelet $\psi(x)$, given by the convolution of a function with a scaled and translated version of $\psi(x)$

$$W_\psi[f(x)](a,b) = |a|^{-1/2} \int f(x)\psi^*\left(\frac{x-b}{a}\right)dx$$
$$= \langle f, \frac{1}{\sqrt{|a|}}\psi\left(\frac{x-b}{a}\right)\rangle \quad (1)$$
$$= \langle f, \psi_{a,b}\rangle = \langle f, U(a,b)\psi\rangle = W_\psi f(a,b), \langle .,. \rangle$$

is the inner product.

where, $f, \psi \in L^2$, the square integrable functions. and ψ, the mother wavelet or analyzing wavelet, satisfies the admissibility condition $C_\psi = \int \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty$. Subscript '*' denotes complex conjugation, 'a' is the scale parameter, $a > 0$, 'b' is the translation parameter. The term $1/\sqrt{|a|}$ is the energy conservative term that keeps energy of the scaled mother wavelet equal to the energy of the original wavelet.

3. Modeling Neural Networks

The neural networks can be represented by graphs showing the computational elements, neurons of the network; each node corresponds to one neuron and the arrows usually denote weighted sums of the values from other neurons. This representation is called as Graph neural network (GNN), which can be used to process structured data inputs, e.g. acyclic graph, cyclic graph, directed or un-directed graphs. This class of neural networks implements a function $\tau(G, n) \in R^m$ that maps a graph G and one of its

nodes n onto an m-dimensional Euclidean space [16].

Let us consider the triplet, $G = (V;E;w)$ to be the connected weighted graph (or a tree) with n nodes and m edges; wherein analogously, nodes being the neurons, arrows as the weighted sums of the values from other neurons, and $w = A_{ij}$, the weights showing a links between the nodes in the given neural network.

A linear factor analysis model that represents the neural network expressed in the form [20],

$$x_i(t) = \sum_j A_{ij} s_j(t) + a_i + n_i(t) \quad (2)$$

where i indexes different components of the observation vector x_i, representing weighted sums of underlying latent variables, j indexes different factors and A_{ij} are the weightings of the factors s, also known as factor loadings. The factors s and noise n are assumed to have zero mean. The bias in x is assumed to be caused by a. In this model, the effect of the inaccuracies and other causes is summarised by Gaussian noise n. In anticipation of the dynamic model, the observations are indexed by t referring to time.

The probability density function for the observations $x_i(t)$ is a normal probability density function given by

$$P(x_i(t) \mid s(t), A, a_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{[x_i(t) - \sum_j A_{ij} s_j(t) - a_i]^2}{2\sigma_i^2}\right) \quad (3)$$

The expression (3) can be simply written as $x(t) \sim N(As + a, \sigma^2)$, where the vector σ^2 contains the variances σ_i^2, the variances of the Gaussian noise terms $n_i(t)$ as σ_i^2,

The normal probability distribution itself is Gaussian distribution. The Gaussian distribution emerges if a large number of independent variables are summed linearly.

With this ground work, we can define the random walk on the graph, $G = (V;E;w)$ representing the neural network as follows.

Our approach is that nodes in a graph represent objects or concepts, and edges represent their relationships. Each object/concept is naturally defined by its features and the related concepts. Thus, we can attach a *state* $x_n \in R$ to each node n, that is based on the information contained in the neighborhood of n. The variable x_n contains a representation of the concept denoted by n and can be used toproduce an *output*, i.e. a decision about the concept/object. The GNN model captures also the random walks on graphs when choosing f_w as a linear function, where f_w is a parametric function, called *local transition function*, that expresses the dependence of a node n on its neighborhood [16]. Then, x_n is defined as follows

$$x_n = f_w(l_n, l_{co}[n], x_{ne}[n], l_{ne}[n]) \quad (4)$$

where $l_n, l_{co}[n], x_{ne}[n], l_{ne}[n]$ are respectively the label of n, the labels of its edges, the states and the labels of the nodes in the neighborhood of n.

In random walks on graphs, the state x_n associated with a node is a real value and is described by

$$x_n = \sum_{u \in Pa[n]} a_{n,i} x_u \quad (5)$$

where $Pa[n]$ is the set of parents of n, and $a_{n,i} \in R$, $a_{n,i} \geq 0$ holds for each n, i. The $a_{n,i}$ are normalized so that

$\sum_{i \in Pa[n]} a_{i,n} = 1$.

The Eq. (5) represents a random walker who is traveling on the graph. The value $a_{n,i}$ represents the probability that the walker, when visiting node n, decides to go to node i. The state x_n stands for the probability that the walker is on node n in the steady state. When all x_n are stacked into a vector x, Eq. (5) becomes $x = Ax$ where $A = \{a_{n,i}\}$ and $a_{n,i}$ is defined as in Eq. (5) if $i \in Pa[n]$ and $a_{n,i} = 0$ otherwise. The corresponding random walk of length k can be defined appropriately as

$\mathcal{S} = \{s[j]; j = 1,2, \ldots, k\}$, where for any position l, we have a cumulative sum of x_n for $1 \leq j \leq k$ described by $s[l] = \sum_{n=1}^{l} x_n$.

This *random walk* of length k on graph G with a root 0 forms a stochastic process with random variables $X = \{x_n\}$. For a completely random sequence $\{x_n\}$, the path mapping gives a Brownian motion type signal [5], say, $\mathcal{S}(t) = \sum_{n=1}^{t} x_n$, where x_n is as defined at (5).

For carrying out the analysis of the neural network characterized by random walk/process, possibly we can resort to random walk with generalized step sizes. But, it is more convenient, especially to employ wavelet analysis, to map elements of random walk consisting of say n characters into points of either a real N-dimensional space with $2^N \geq n$ or to a complex space of dimension $N/2$. This now becomes a complex valued random walk defined on the graph $G = (V;E;w)$ representing the neural network (GNN).

4. Wavelet Based Analysis

The wavelet transform is well suited for characterizing the scaling properties of fractal

objects. The random processes exhibit self similar fractal like properties. The random walk on graph neural networks itself is a random/stochastic process. Hence, wavelet decomposition, which is also based on self-similarity, turns to be effective to tool to extract fractal information from the structure of graph neural network. In our earlier work [5], we have carried out the wavelet analysis of the random processes. For a random process, $X = X(t,w)$ defined on the probability space (Ω, B, P), for any function $\psi: R \to C$ satisfying the admissibility condition, the wavelet transform is given as

$W_\psi[X(u,w)](a,t) = \int_R \psi_{a,t}{}^*(u)\, X(u,w)du, t \in R, a > 0, w \in \Omega.$ (6)

where $\psi_{a,t}(u) = \frac{1}{\sqrt{a}}\psi\left(\frac{u-t}{a}\right)$, the scaled and translated version of analyzing wavelet ψ, subject to the stipulation

$$E \mid X(t) \mid^2 = \int_\Omega \mid X(t,w) \mid^2 dP(w) < \infty, \quad \forall t \in R, w \in \Omega \quad , \tag{7}$$

where E denotes the mathematical expectation.

Applying this formulation to the complex valued random walk defined on GNN, $\mathcal{S}(t) = \sum_{n=1}^{t} x_n$, where x_n is as given in Eq (5), we obtain the wavelet transform of $\mathcal{S}(t)$ with respect to the analyzing/generic wavelet ψ as

$$W_\psi[\mathcal{S}(t)](a,b) = \int_R \frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)\mathcal{S}(t)dt, \quad b \in R, a > 0 \quad , \tag{8}$$

subject to the stipulation

$$E \mid \mathcal{S}(t) \mid^2 = \int \mid \mathcal{S}(t) \mid^2 dP(x) < \infty, \forall x \in R, \tag{9}$$

where E denotes the mathematical expectation. To have the WT well defined, the signal $\mathcal{S}(t)$ should belong to the $L^2(R)$, the class of square integrable functions.

Indeed, $\mathcal{S}(t)$ belongs to the class of square integrable functions since $x_n's$ have Gaussian distribution and the Gaussian functions belong to this class of square integrable functions.

The wavelet transform (8), is a convolution of function $\mathcal{S}(t) \in L^2(R)$ with certain locally supported function $\psi(t)$ shifted and dilated, called analyzing wavelet. It should be noted that the choice of the analyzing wavelet is of vital importance as it dictates the representation and properties of the wavelet transform. In particular, the choice depends on the factors like the kind of signal to be analyzed.

As regards to the analyzing wavelet, $\psi(t)$, it is chosen from the battery of wavelets for which its first n momenta vanish. Usually, the analyzing wavelet ψ is chosen from the family of wavelets based on the Gaussian functions, which can be successively differentiated to any arbitrary order n, $\varphi^n(t) = (-1)^{n+1}\frac{d^n}{ds^n}exp\,(-t^2/2)$ and whose first moments vanish, that is, $\int t^k \varphi(t)dt = 0$, for $0 \le k \le n$.

For appropriate choice of analyzing wavelet, we can make use of the fact that the random walk formed by the x_n given in Eq. (5); of which each x_i, s as the weighted sums of the latent variables given by Eq. (3) in the neural network having induced normal probability density function,

$$f(x_i; \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}} \tag{10}$$

with $\mu_i = \sum_j A_{ij} s_j(t)$, $-\infty < x < \infty$, $-\infty \le \mu_i \le \infty$ and $\sigma_i > 0$.

By the transformation, $X = \frac{x-\mu}{\sigma}$, we can have the standard normal probability density function, $\psi(X) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$, which is a Gaussian function that can be differentiated successively to any arbitrary order, say, n, as $\psi^{(n)}(t) = (-1)^{n+1}\frac{d^n}{dx^n}exp\,(-x^2/2)$, and whose first momenta vanish.

In particular, we can obtain second order moment $(n = 2)$,

$$\psi_2(X) = \psi''(X) = \frac{(x^2-1)}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \tag{11}$$

which is a "Mexican hat wavelet". Mexican hat wavelet has been analytically proved wavelet that satisfies the admissibility condition, $C_{\psi_2} = 2\int_0^\infty \frac{|\hat{\psi}_2(\omega)|^2}{|\omega|} d\omega = 2\,\pi\Gamma(n) < \infty$. Hence, we can safely employ $\psi_2(X)$, as the analyzing wavelet in the transform equation.

Thus, the wavelet transform of the random walk $\mathcal{S}(t)$, with the analyzing wavelet, ψ_2 will be

$$W_{\psi_2}[\mathcal{S}(t)](a,b) = \int_R \frac{1}{\sqrt{a}}\psi_2\left(\frac{t-b}{a}\right)\mathcal{S}(t)dt \tag{12}$$

For numerical implementation, to compute the wavelet transform, we need to discretize the expression (12) as

$$w_{j,k} = W_{\psi_2}\mathcal{S}(j,k) = \int_R \mathcal{S}(t)\psi_{2\,j,k}(t)dt \tag{13}$$

In this discretized form, the signal $\mathcal{S}(t)$ is represented as a linear decomposition

$$\mathcal{S}(t) = \sum_j \sum_k w_{j,k}\psi_{2\,j,k}(t)\ , j,k \in Z, \tag{14}$$

where $w_{j,k}$ are the wavelet coefficients of the expansion that constitute the discrete set, and $\psi_{2\,j,k}$ is a set of wavelet functions of t given by

$$\psi_{2\,j,k}(t) = 2^{j/2}\psi_2(2^j t - k): j,k \in Z.$$

Note that the scaling and translations are discrete, and the indexes j and k are respectively related to these processes. The basis functions $\psi_{2\,j,k}$ are dilated in a dydatic form (in powers of two), when varying the value of index j, and in analogous way translated when varying the index k. In this process, translation is associated with time resolution, and dilation provides scaling.

Now, let us turn to inverting the transform so as to reconstruct the signal corresponding to the given set of transform coefficients so obtained.

Intuitively, the wavelet coefficient $w_{j,k} = W_{\psi_2}\mathcal{S}(j,k)$ provides a measure of "how much of" the wavelet $\psi_{2\,j,k}(t)$ is present in the signal $\mathcal{S}$. This suggests that the original signal may be recovered by summing the wavelets $\psi_{2\,j,k}(t)$ multiplied by each wavelet coefficient $w_{j,k}$. That is,

$$\frac{1}{C_{\psi_2}}\sum_{n=1}^{N} \int_R w_{j,k}\psi_{2\,j,k}(t)\frac{dt}{t} = \mathcal{S}^{\#} \qquad (15)$$

Finally, we calculate correlations between wavelet coefficients $W_{\psi_2}[\mathcal{S}(t)](a,b)$ at different scales. For this, we define wavelet covariance of the covariance function for all t,

$R_S(u,v) = E\psi_2{}^*(u)\varphi(v), u,v \in R$, as

$$R_W(a,b,c,d) = EW_a(b)W_b(d) = \int_R \frac{1}{\sqrt{ac}}\psi_2\left(\frac{u-b}{a}\right)\psi_2{}^*\left(\frac{v-d}{c}\right)R_S(u,v)dudv \quad , \qquad (16)$$

provided that the condition, $E\left\{\int_R \left|\frac{1}{\sqrt{a}}\psi_2{}^*\left(\frac{t-b}{a}\right)\mathcal{S}(t)dt\right|\right\}^2 < \infty$, holds good.

That is, $R_W(a,c,b-d) = \langle W_{\psi_2}(a,b), W_{\psi_2}(c,d\rangle$, where the curly brackets $\langle .,. \rangle$. mean the covariance $cov(W_1,W_2) = E\frac{(W_1 - E(W_1)(W_2 - E(W_2)}{\sqrt{DW_1.DW_2}}$ where D is the dispersion and E is the mathematical expectation. We should infer that for a random walk, wavelet coefficients correlation function will coincide with that of random signal; if no, the structure of wavelet coefficients correlation function will be different.

5. Discussion

Wavelet analysis is well suited for scale-invariant systems, that is, the systems that exhibit self similar properties at different scales. The Brownian motion is a classical example of scale-invariant physical systems. Random walks are the random processes or the stochastic processes, which in turn, constitute a Brownian motion signal. In our work, we adopted a novel approach to model the neural network in graph theoretic framework, defined a random walk on the graph and suitably employed wavelet transform to analyse the complex network in this representation, characterized by Brownian motion signal. One of the goals of random walk is to determine long-range power law correlation based on measurements of fluctuations [14]. The benefit of using this random walk is that the sequence always evolves along the positive direction on z-axis as a function of sequence index k and this facilitates transform analysis. The wavelet transform technique locates the periodicities in the random sequence defined on GNN and also gives fine time index resolution so that these patterns can be located accurately in the walk representing the sequence. Transforming the signal actually yields the discrete set of wavelet coefficients, at a particular scale and translation which tells us how well the signal and the scaled and translated analyzing wavelet match. If the signal is similar to the scaled and translated analyzing wavelet, then the wavelet coefficient will have big magnitude. The wavelet coefficient also represents the degree of correlation between two functions at a particular scale and translation.

The prime objective of wavelet analysis is to extract structural information from the complex network structure in the transform domain. We, in the mathematical analysis, obtain the wavelet transform of the random walk $\mathcal{S}$; which in turn can be plotted in space-time plane. This will demonstrate the structural pattern visually and help locating periodicities of these patterns in the sequence that represents the network. The actual values can then obtained by plotting $\left|W_{\psi_2}\mathcal{S}(a,b)\right|$ over the space-time plane. The peculiar feature of this wavelet model is that the choice of Gaussian function obtained from the normal probability density function of latent variables of the network as analyzing wavelet is robustly made owing to the fact the random walk is formed by the sequence comprised of weighted sums of the latent variables in the neural network that have probabilities distributed normally.

6. Conclusion

Complex neural networks arising in small world of neuroscience can be modeled with graphs using analogy as a natural extension. We can define random process on graph representing neural networks making it possible to apply wavelet techniques for visualization of patterns and feature

extraction from the complex networks. The present work robustly models the complex neural network in graph theoretic settings and conducts wavelet analysis in stochastic framework. The technique is novel and relatively easy to implement. This technique will provide a new insight in handling the applications of network theory to neuroscience such as modelling of neural dynamics on complex networks; graph theoretical analysis of neuroanatomical networks; and applications of graph analysis to studies of functional connectivity with fMRI, EEG and MEG and others from the small world of neuroscience.

References

1. Bhosale B. and A. Biswas, "Wavelet analysis of optical solitons and its energy aspects", *Mathematics in Engineering, Science and Aerospace,* **2012**, 3(1): 15-27
2. Bhosale B. and A. Biswas, "Wavelet analysis of soliton interaction and its relation to probability distributions", *Nonlinear Studies*, **2012**, 19(4): 563-572
3. Bhosale B. and A. Biswas, "Wavelet transform of soliton solution of Sine-Gordon equation as Random process". (submitted)
4. Bhosale B., B. S. Ahmed, A. Biswas, "Multi-resolution Analysis of Wavelet like Soliton Solution of KdV Equation". (Submitted)
5. Bhosale B., B. S. Ahmed, A. Biswas, "Wavelet Based Analysis in Bio-informatics", *Life Sci.,* **2013,** 10(2):853-859] (ISSN: 1097-8135)
6. Bhosale B., A, Yildirim and A. Biswas, "Modelling space-time-varying systems for analyzing solitons", *Advances in Science, Engineering and Medicine,* **2012**, 4: 164-170
7. Hammond D, Pierre Vandergheynstb,R´emi Gribonval, "Wavelets on graphs via spectral graph theory", *Applied and Computational Harmonic Analysis,* **2011**, 30(2): 129-150
8. Jansen M, G. P. Nason and B. W. Silverman, "Multiscale methods for data on graphs and irregular multidimensional situations", *Journal of the Royal Statistical Society*: *Series B (Statistical Methodology),* **2009,** 71(1): 97–125
9. Jim XU and BAO Zheng,"Neural networks and Graph theory", *Science in China (Series F)*, **2002**, 45(1): 1-24
10. Lio P., "Histidine biosynthetic pathway and genes", *Bioinformatics,* **2003**, 19(1): 2-9.
11. Lovasz L., "Random Walks on Graphs: A Survey", *Bolyai society, Mathematical studies, Keszthely (Hungary),* **1993**, l(2): 1-46
12. Mallat S., "A Wavelet Tour of Signal Processing", *Academic Press*, **1998**
13. Mashaghi A. R., A. Ramezanpour, V. Karimipour, "Investigation of a protein complex network", *The European Physical Journal B - Condensed Matter and Complex Systems* **2004**, 41(1): 113-121
14. Peng C., S. Byldyrev and H. Stanley,"Long-range correlations in nucleotide sequences" *Nature,* **1992,** 356: 168-171
15. Podolak IT."Functional graph model of a neural network", *IEEE Trans Syst Man Cybern B Cybern,* **1998**, 28(6):876-81
16. Scarselli F. M. Gori Ah Chung Tsoi, M. Hagenbuchner, G. Monfardini, "The Graph Neural Network Model", *Neural Networks, IEEE Transcations,* **2009**, 20(1): 61-80.
17. Smalter A., J. Huan, G. Lushington, "Graph wavelet alignment kernels for drug virtual screening", *Journal of Bioinformatics and Computational Biology,* **2009,** 7: 473–497
18. Cornelis J Stam and Jaap C Reijneveld, "Graph theoretical analysis of complex networks in the brain", *Nonlinear Biomedical Physics,* **2007**, 1:3 doi:10.1186/1753-4631-1-3
19. Szu , Xiang-Yang Yang ,Brian A. Telfer, Yunlong Sheng, "Neural network and wavelet transform for scale-invariant data classification", *Phys. Rev.,* **1993,** 48(E): 1497–1501
20. Valpola H.,"Bayesian Ensembel learning for Non-linear factor analysis", *Acta Ploytechnica, Scandinavica, Mathematics and Computing Series,* **2000,** 108:54-64