A new measure of heterogeneity for complex networks

Rinku Jacob and K. P. Harikrishnan
Department of Physics, The Cochin College, Cochin-682 002, India

R. Misra
Inter University Centre for Astronomy and Astrophysics, Pune-411 007, India

G. Ambika
Indian Institute of Science Education and Research, Pune-411 008, India

We propose a novel measure of heterogeneity for unweighted and undirected complex networks that can be derived from the degree distribution of the network instead of the degree sequences, as is done at present. We show that the proposed measure can be applied to all types of topology with ease and shows direct correlation with the diversity of node degrees in the network. The measure is mathematically well behaved and is normalised in the interval $[0,1]$. The measure is applied to compute the heterogeneity of synthetic (both random and scale free) and real world networks. We specifically show that the heterogeneity of an evolving scale free network decreases as a power law with the size of the network $N$, implying a scale free character for the proposed measure. Finally, as a specific application, we show that the proposed measure can be used to compare the heterogeneity of recurrence networks constructed from the time series of several low dimensional chaotic attractors, thereby providing a single index to compare the structural complexity of chaotic attractors.

PACS numbers: 89.75.Hc, 89.75.Da, 05.45.Ac

I. INTRODUCTION

A network is an abstract entity consisting of a certain number of nodes connected by links or edges. The number of nodes that can be reached from a reference node $i$ in one step is called its degree denoted by $k_i$. If equal number of nodes can be reached in one step from all the nodes, the network is said to be regular or homogeneous. A regular lattice where nodes are associated with fixed locations in space and each node connected to equal number of nearest neighbours, is an example of a regular network. However, in the general context of complex networks, a metric is not associated with the network as it is defined in an abstract space with a set of nodes $N=1,2,3,...,N$ and a set of links denoted by $K=k_1,k_2,k_3,...,k_{N-1}$. As the spectrum of $k$ values of the nodes increases, the network becomes more and more irregular and complex. Over the last two decades, the study of such complex networks has developed into a major field of inter-disciplinary research spanning across mathematics, physics, biology and social sciences.

Many real world structures and interactions can be modeled using the underlying principles of complex networks and analysed using the associated network measures. In such contexts, the corresponding complex network can be weighted or unweighted and directed or undirected depending on the system or interaction it represents. In this paper, we restrict ourselves to unweighted and undirected networks and the possible extensions for weighted and directed networks are discussed in the end. The topology or structure of a complex network is determined by the manner in which the nodes are connected in the network. For example, in the case of the classical random graphs (RG) of Erdős and Rényi (E-R) who started the mathematical study of complex networks, two nodes are connected with a constant and random probability $p$. In contrast, many real world networks are found to have a tree structure with the network being a combination of small number of hubs on to which large number of individual nodes are connected. An important measure that distinguishes between different topologies of complex networks is the degree distribution $P(k)$ that determines how many nodes in the network have a given degree $k$. For the RGs, $P(k)$ is a Poisson distribution around the average degree $\langle k \rangle$ while most real world networks follow a fat-tailed power law distribution given by $P(k) \propto k^{-\gamma}$, with the value of $\gamma$ typically between 1 and 3. Such networks are called scale free (SF) due to the inherent scale invariance of the distribution.

*Electronic address: rinku.jacob.vallanat@gmail.com
†Electronic address: kp.hk2002@yahoo.co.in
‡Electronic address: rmisra@iucaa.in
§Electronic address: g.ambika@iiserpune.ac.in
Though topology is an important measure of a complex network, that alone is not sufficient to characterize and compare the interactions that are so vast and diverse. A number of other statistical measures have been developed for this purpose, each of them being useful in different contexts. Two such commonly used quantifiers are the clustering coefficient (CC) and the characteristic path length (CPL). There are also measures of local structure used to compare the complexity of networks in particular cases, such as, the hierarchy or community structure \[5\] in social networks and motifs \[14\] and super family profiles \[5\] in genetic and neuronal networks. However, a single index that can quantify the diversity of connections between nodes in networks even with different topologies, is the heterogeneity measure \[15\]. It is also indicative, in many cases, of how stable and robust \[16\] a network is with respect to perturbations from various external parameters. An important example is the technological network of North American power grid \[1\]. Recent studies have also revealed the significance of the heterogeneity measure in various other contexts, such as, epidemic spreading \[17\], traffic dynamics in networks \[18\] and network synchronization \[19\].

The network heterogeneity has been defined in various ways in the literature which we will discuss in detail in the next section where, we will also present the motivations and need for a new measure. While the existing measures are based on the degree correlations \(k_i\) and \(k_j\) of nodes \(i\) and \(j\) in the network, the measure proposed in this paper uses only the degree distribution \(P(k)\) to compute the heterogeneity of the network. We show that this new measure has a direct correlation with the \(k\) spectrum, or the spectrum of \(k\) values in the network, and hence gives a true representation of the diversity of node degrees present in the network and also serves as a single index to quantify the network diversity.

In this work, we also include a class of networks not considered so far in the context of heterogeneity measure in any of the previous works. These are complex networks constructed from the time series of chaotic dynamical systems, called recurrence networks (RNs) \[20\]. They have a wide range of practical applications \[21, 22\] and the measures from these networks are used to characterize strange attractors in state space typical of chaotic dynamical systems, as discussed in §V. The diversity of node degrees in the RNs was actually one of the motivations for us to search for a heterogeneity measure that could be used to compare the structural complexities of different chaotic attractors through the construction of RNs.

Our paper is organized as follows: In the next section, we discuss briefly all the previous measures of heterogeneity and give reasons why we have to look for a new measure. The measure that we propose is based on the idea of what we consider as a completely heterogeneous network of \(N\) nodes, that is illustrated in §III. The proposed measure of heterogeneity is presented in §IV while §V and §VI are devoted to computation of this new measure for various synthetic as well as real world networks. Our conclusions are summarised in §VII.

## II. EXISTING MEASURES OF HETEROGENEITY

A detailed discussion on all the earlier heterogeneity measures is given in the recent paper by Estrada \[15\]. Earlier attempts to measure the heterogeneity of complex networks found in the literature were mainly motivated by the random graph theory. For example, the first person to propose a measure of heterogeneity was Snijders \[23\] in the context of social networks and it was modified by Bell \[24\] as the variance of node degrees:

\[
VAR = \frac{1}{N} \sum_{i}^{N} (k_i - \langle k \rangle)^2
\]

where \(\langle k \rangle\) represents the average degree in the network. Though this is still one of the popular measures of heterogeneity, its applicability is mainly limited to RGs where one can effectively define an average \(k\).

Another measure was proposed by Albertson \[25\] again based on graph theory as:

\[
A = \sum_{i,j} |k_i - k_j|
\]

which is a sum of the local measures of irregularity. This index also is not useful in quantifying correctly different varieties of complex networks. There was also an attempt to define heterogeneity in terms of the structural entropy of a network \[26\] in the context of SF networks, which is more of an entropy measure than a heterogeneity measure. A unified index to quantify the heterogeneity of networks with any distribution has been proposed \[27\] based on the Gini coefficient \[28\] which is a measure originally used in microeconomics to study income equality. In short, none of these measures, though useful in particular contexts, truly reflects the heterogeneity in a network.

Of late, a much better index has been proposed by Estrada \[15\] based on the irregularity of two connecting nodes in the network as:

\[
\rho = \sum_{i,j} \left( \frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}} \right)^2
\]
FIG. 1: A comparison of the completely heterogeneous networks (see text) with $N = 4, 5, 6$ and 7. In each case, all the possible $k$-values from 1 to $(N - 1)$ are present in the network as shown. One degree (one $k$ value) has to be shared by two nodes since the $N^{th}$ node will have the degree of any one of other nodes. It is empirically shown that this degree of $N^{th}$ node, denoted by $k^*$, is automatically fixed (if the network has all possible degrees from 1 to $(N - 1)$) and is $N/2$ if $N$ is even and $(N - 1)/2$ if $N$ is odd. For example, for $N = 4$ and 5, $k^* = 2$ and for $N = 6$ and 7, $k^* = 3$ and so on.

which can also be normalised to get a measure $\rho_n$ within the unit interval $[0, 1]$ as:

$$\rho_n = \frac{\rho}{N - 2\sqrt{(N - 1)}}$$

This measure has some interesting mathematical properties as it can be expressed as a quadratic form of the Laplacian matrix of the network and also allows a spectral representation of the heterogeneity. Moreover, it can be used to quantify a wide range of networks of different topologies as shown by the author [15].

However, this measure is so defined, as indicated by the author, to give maximum heterogeneity for a star network of given number of nodes $N$ so that even the SF networks, which are normally considered to be quite heterogeneous,
have a maximum heterogeneity of only about 10% of the corresponding star network [15]. The basic issue here is the subjectivity to choose which network of a given size of \( N \) nodes is most heterogeneous. If the definition of the most heterogeneous network changes, as we show below, the above measure has to be modified. In the star network, \((N - 1)\) nodes have degree 1 and the \( N \)th node has degree \((N - 1)\). One is generally inclined to believe that the star network is the most heterogeneous because it has a degree that is maximum possible in a network of \( N \) nodes, namely, \((N - 1)\). What we argue here is that it is the diversity in the node degrees that really matters in determining the heterogeneity measure of a complex network, rather than the presence of a single node with large degree. Accordingly, our definition of a completely heterogeneous network of \( N \) nodes is presented in the next section. Note that, even from other aspects, a star network fails to be promoted as most heterogeneous. For example, if one correlates the robustness or stability of a network with heterogeneity, the star network is the most vulnerable since disruption of just one node can destroy the entire network. Since any SF network is a combination of several star like networks, it is difficult to conceive that the heterogeneity of a SF network is much less than that of a star network. So our initial effort is to define properly the completely heterogeneous complex network of \( N \) nodes which then can naturally lead to a heterogeneity measure applicable to all the different types of networks.

### III. COMPLETELY HETEROGENEOUS COMPLEX NETWORK

Here we present what we consider as the logical limit of a completely heterogeneous network of \( N \) nodes. The reader may find that this is an ideal case. Nevertheless, it helps to put the concept of heterogeneity of a complex network in a proper perspective. Consider an unweighted and undirected complex network of \( N \) nodes, with all the nodes connected to the network having a degree of at least one. If all the nodes have the same degree \( k \), the network is completely homogeneous with the degree distribution \( P(k) \) being a \( \delta \)function peaked at \( k \).

Let us now consider the other extreme where no two nodes have the same degree. The maximum possible degree for a node is \((N - 1)\). Let the nodes be arranged in the ascending order of their degree. It is obvious that the \( N \)th node will have to take a degree equal to that of any one of the other nodes having degree from 1 to \((N - 1)\). To find out what degree is possible for the \( N \)th node under the given condition, we start with taking small number of nodes as shown in Fig. 1 where we show 4 different cases of \( N \) ranging from 4 to 7. In each case, the \( N \)th node is represented as a pentagon shape with its degree denoted as \( k^* \). It is clear that if all the node degrees are to be different, there is only one possible value of \( k^* \) for the \( N \)th node, which is \( \frac{N}{2} \) if \( N \) is even and \( \frac{(N-1)}{2} \) if \( N \) is odd.

We now give a simple argument that this result is true in general for any \( N \). The degree of node 1 is 1 which means that it is connected only to the node with degree \((N - 1)\). That is, node 1 is not connected to \( N \)th node. Node 2 is connected only to two nodes with degree \((N - 1)\) and \((N - 2)\) and hence it is also not connected to node \( N \). By induction, one can easily show that the \( r \)th node is connected only to nodes with degree \((N - 1), (N - 2), \ldots, (N - r)\). Suppose \( N \) is even. When \( r = \frac{N}{2} \), this node is connected to nodes with degree from \((N - 1)\) to \( \frac{N}{2} \). To avoid self loop, this node should be connected to node \( N \). Thus all nodes with higher degree from \( \frac{N}{2} \) to \((N - 1)\) are connected to node \( N \) whose degree becomes \( \frac{N}{2} \). By a similar argument, one can show that the degree of \( N \)th node is \( \frac{(N-1)}{2} \) if \( N \) is odd.

![Fig. 2: Change in the degree distribution for a typical complex network as it is transformed from complete homogeneity to complete heterogeneity, for \( N = 10 \).](image)

The diagram illustrates the change in the degree distribution of a complex network from complete homogeneity to complete heterogeneity, with \( N = 10 \). The degree distribution \( P(k) \) is shown for different nodes, highlighting the transformation from a homogeneous to a heterogeneous state.
FIG. 3: A snapshot of different types of complex networks in the increasing order of their heterogeneity ($H_m$), taking $N = 50$ in all cases. From top to bottom, the nature of the network varies from completely homogeneous, star, RG, SF and finally to completely heterogeneous network. The degree distribution and $k$ spectrum are also shown for each case to indicate that $H_m$ is a measure of the diversity in the node degrees.

Let us now consider the degree distribution $P(k)$ of this completely heterogeneous network. All the nodes have different $k$ values and only two nodes share the same $k$ value, $k^*$. One can easily show that:

\[
P(k) = P_0 = \frac{1}{N}, (k \neq k^*)
\]

\[
P(k) = \frac{2}{N}, (k = k^*)
\]
Our definition of heterogeneity is derived in such a way that this network has maximum heterogeneity, which is done in the next section.

IV. A NEW MEASURE OF HETEROGENEITY

It is very well accepted that a network of $N$ nodes with all nodes having equal degree $k$ is a completely homogeneous network with $P(k)$ being a δfunction centered at $k$. The value of $k$ can be anything in the range $2 \leq k \leq (N - 1)$ and all these networks have heterogeneity measure zero, for any $N$. In principle, the heterogeneity of a network should measure the diversity in the node degrees with respect to a completely homogeneous network of same number of nodes. All the measures defined so far in the literature directly use the $k$ values present in the network for computing the heterogeneity measure. Here we argue that a much better candidate to define such a measure is $P(k)$ rather than $k$. Since $P(k)$ is a probability distribution, as the spectrum of $k$ values increase, the value of $P(k)$ gets shared between more and more nodes with the condition $\sum_k P(k) = 1$. In other words, this variation in $P(k)$ reflects the diversity of node degrees and hence the heterogeneity of the network. A typical variation of $P(k)$ as the network changes from complete homogeneity to complete heterogeneity is shown in Fig.2 Note that for RGs, this variation in $P(k)$ is with respect to $P(<k>)$, with $<k>$ being the average degree, while for SF networks, it is with respect to $P(k_{\text{min}})$.

To get the heterogeneity measure, we first define a heterogeneity index $h$ for a network of $N$ nodes as the variance of $P(k)$ with respect to the peak value corresponding to the completely homogeneous case:

$$h = \frac{1}{N} \sum_{k_{\text{min}}}^{k_{\text{max}}} (1 - P(k))^2, P(k) \neq 0 \quad (5)$$

The condition implies that the summation is only over $k$ values for which $P(k) \neq 0$. For a completely homogeneous network, $P(k)$ is non zero only for one value of $k$, say $k^*$, and $P(k^*) = 1$, making $h = 0$, for all $N$.

We now consider the other extreme of completely heterogeneous case. From the results in the previous section for the completely heterogeneous case, we have

$$h_{\text{het}} = \frac{1}{N} \sum_{k=1}^{(N-1)} (1 - P(k))^2 \quad (6)$$

Putting the values of $P(k)$ and simplifying, we get

$$h_{\text{het}} = 1 - \frac{3}{N} + \frac{N + 2}{N^3} \quad (7)$$

This is the maximum possible heterogeneity measure for a network of $N$ nodes. For large $N$, as a first approximation, we have

$$h_{\text{het}} \approx 1 - \frac{3}{N} \quad (8)$$

For finite $N$, its value is $<1$ and as $N \to \infty$, $h_{\text{het}} \to 1$. To define the heterogeneity measure ($H_m$) for a network, we normalise the heterogeneity index of the network with respect to the completely heterogeneous network of same number of nodes to get the value in the unit interval $[0,1]$:

$$H_m = \frac{h}{h_{\text{het}}} \quad (9)$$

If $N$ is sufficiently large, say $N > 1000$ as is the case for most practical networks, $h_{\text{het}} \sim 1$ and $H_m \approx h$.

We note the following features regarding $H_m$:

i) It is defined here for unweighted and undirected complex networks and represents a unique measure applicable to networks of different topologies and degree distributions.

ii) $H_m$ for a star network is very close to zero and hence the star network is nearly homogeneous in our definition. This is because, the degree of only one node is different from the rest of the nodes.

iii) For two networks of the same size $N$ independent of the topology, the measure we propose is in direct correlation to the degree diversity in the network. To show this explicitly, we present the $k$ spectrum, the spectrum of $k$ values in the network in the form of a discrete line spectrum. In Fig.3 we compare some standard networks in the increasing order of their $H_m$, taking $N = 50$. For each network, we show the degree distribution (as histogram), the $k$ spectrum
FIG. 4: Degree distribution of E-R networks (RGs) for four different $p$ values with $N$ fixed at 2000. The $k$ spectrum is shown below the degree distribution. The value of $H_m$ and $<k>$ are also indicated in each case. Note that $H_m$ is directly correlated with the degree diversity or the spectrum of $k$ values in the network.

and the value of $H_m$. Note that of different topologies, the SF network is the most heterogeneous. Here the star network has a reasonably high value of $H_m$ since $N$ is only 50. We also show the completely heterogeneous network with $H_m = 1$, for comparison.

iv) The heterogeneity index $h$ is defined as a measure normalised with respect to the size of the network $N$. For large $N$, since $h \sim H_m$, the measure $H_m$ can also be used to compare the heterogeneities of two networks even if $N$ is different. This is especially important for real world networks where $N$ varies from one network to another, as discussed in §6. However, a network with larger $N$ generally tends to have lower $H_m$ since, to keep the same heterogeneity, the range of non zero $k$ values should also increase correspondingly. In other words, a network with 100 nodes attains complete heterogeneity if the $k$ values range from 1 to 99 whereas, to attain complete heterogeneity for a network of 1000 nodes, the $k$ values should range from 1 to 999.

The above result also implies that for any network that is evolving or growing, for example the SF network where the nodes are added with preferential attachment [39], the value of $H_m$ generally keeps on decreasing with increasing $N$. In the next section, we numerically study the variation of $H_m$ with different network parameters for various synthetic networks.

V. HETEROGENEITY OF SYNTHETIC NETWORKS

In this section, we analyze 3 different classes of complex networks, namely, the RGs of Erdos-Renyi, the SF networks and the networks derived from the time series of chaotic dynamical systems, called recurrence networks (RNs) whose details are discussed in §V.C.

A. Classical random graphs

For RGs, the degree distribution is Poissonian centered around an average degree $<k> = pN$ where $p$ is the probability that two nodes in the network is connected. In Fig. 4, we show the degree distribution and the $k$ spectrum
FIG. 5: Variation of $H_m$ with $p$ for RGs for a fixed value of $N$, as shown. We expect the profile for a higher $N$ value to be within that of a lower $N$ as $H_m$ decreases with $N$ for any fixed $p$.

for RGs of 4 different $p$ values with $N$ fixed at 2000. The values of $H_m$ for all these networks are also shown. The main result here is that the value of $H_m$ has a direct correlation with $k$ spectrum for a fixed $N$.

We next consider how $H_m$ depends on $p$ and $N$, the two basic parameters of the RG. The effect of changing $p$ for a fixed $N$ as well as changing $N$ for a fixed $p$ are to shift the average $k$ value of the nodes in the RG. Since the degree distribution is approximately Gaussian, the spectrum of $k$ values depends directly on the variance of the Gaussian profile. As $p$ increases from zero for a fixed $N$, the spectrum of $k$ values and hence $H_m$ increase correspondingly. Due to the obvious symmetry, as $p$ increases beyond 0.5, $H_m$ starts decreasing. Thus the maximum value of $H_m$ is obtained for $p = 0.5$ for any fixed $N$. On the other hand, by increasing $N$ for any fixed $p$, one expects the Gaussian profile of the degree distribution to become sharper, thus decreasing $H_m$. These results are compiled in Fig. 5 for three values of $N$. Note that the minimum $p$ value that can be used for $N = 500$ is 0.004 and this decreases as $N$ increases. In the figure, we show the results starting from $p = 0.1$. Higher values of $N$ would involve very large computer memory requirements for large $p$. However, we have checked the variation of $H_m$ with $N$ for smaller $p$ values, say 0.005 and 0.01, for $N$ up to 5000 and have found that the decrease is approximately exponential.
FIG. 6: Degree distributions (inset) and the distributions in log scale along with the \( k \) spectrum for synthetic SF networks with four different values of \( \gamma \) and \( N \) fixed at 2000. In all cases, the values of \( H_m \) and the minimum degree \( k_{\text{min}} \) of the network are also shown. As the spectrum of \( k \) values increases, \( H_m \) increases correspondingly for a fixed \( N \).

FIG. 7: Variation of \( H_m \) with \( N \) for synthetic SF networks with two different values of \( k_{\text{min}} \). The same variation is shown in the inset in log scale indicating a clear power law in both cases.
B. Scale free networks

For SF networks, the degree distribution obeys a power law $P(k) \propto k^{-\gamma}$. To construct the SF network synthetically, we use the basic scheme provided in the website: www.barabasilab.com. In this scheme, we start with a small number of initial nodes denoted as $m_0$. As a new node is connected, a fixed number of edges, say $m$, is added to the network. This number represents the minimum number of node degree, $k_{\min}$, in the network. The two parameters, $m_0$ and $k_{\min}$, determines the value of $\gamma$ as the network evolves. We have constructed SF networks of different $\gamma$ by changing both $m_0$ and $k_{\min}$.

In Fig. 6, we show the degree distribution and the corresponding $k$ spectrum for SF networks with four different $\gamma$ and $k_{\min}$, with $N$ fixed at 2000. We find that the $k$ spectrum and hence the value of $H_m$ depend directly on $k_{\min}$ as can be seen from the figure. In other words, for a SF network of fixed $N$, $H_m$ increases as the value of $k_{\min}$ increases. The variation is approximately linear for $k_{\min}$ in the range 1 to 10. More interesting is the variation of $H_m$ with $N$ for a fixed $k_{\min}$. In Fig. 7 we show the variation of $H_m$ as $N$ increases from 1000 to 10000 for two different SF networks with $k_{\min} = 5$ and 10. This variation is also shown in the inset in a log scale in the same figure indicating that $H_m$ obeys a power law, $H_m \propto N^{-\rho}$, where the value of $\rho$ is found to be 0.6288 for $k_{\min} = 5$ and 0.6581 for $k_{\min} = 10$. Thus the scale free character in the evolution of the network is also reflected in the variation of $H_m$ with $N$.

C. Recurrence networks

Recently, a new class of complex networks has been proposed for the characterization of the structural properties of chaotic attractors, called the recurrence networks [20, 29]. They are constructed from the time series of any one variable of a chaotic attractor. From the single scalar time series, the underlying attractor is first constructed in an embedding space of dimension $M$ using the time delay embedding [30] method. Any value of $M$ equal to or greater than the dimension of the attractor can be used for the construction of the attractor. The topological and the structural properties of this attractor can be studied by mapping the information inherent in the attractor to a complex network and analyzing the network using various network measures.

To construct the network, an important property of the trajectory of any dynamical system is made use of, namely, the recurrence [31]. By this property, the trajectory tends to revisit any infinitesimal region of the state space of a dynamical system over a certain interval of time. To convert the attractor to a complex network, one considers all
the points on the embedded attractor as nodes and two nodes \( i \) and \( j \) are considered to be connected if the distance \( d_{ij} \) between the corresponding points on the attractor in the embedded space is less than or equal to a recurrence threshold \( \epsilon \). Selection of this parameter is crucial in getting the optimum network that represents the characteristic properties of the attractor. The resulting complex network is the RN which, by construction, is an unweighted and undirected network. The adjacency matrix \( A \) of the RN is a binary symmetric matrix with elements \( A_{ij} = 0 \) or 1. In Fig. 9 we show the construction of the RN from the time series of the celebrated Lorenz attractor. The error bar in the degree distribution represents statistical error due to finiteness in the number of nodes. Its value is given by \( \frac{\sqrt{\pi(k)}}{N} \), where \( N \) is the number of nodes and \( n(k) \) the number of nodes with a given degree \( k \). The error gets normalised to \( \frac{n(k)}{N} \) as \( n(k) \to 0 \).

To select the value of \( \epsilon \), the generally accepted criterion is to look for the emergence of a giant component as suggested by Donges et al. \cite{32} and Eroglu et al. \cite{33}. We have recently proposed \cite{34} a general framework for the selection of \( \epsilon \) and have shown that the value of \( \epsilon \) is closely linked to the embedding dimension \( M \). Here we follow this criterion to construct the RN from chaotic time series using the natural dimension of the attractor and the corresponding \( \epsilon \). To get a comparison between different chaotic attractors, we have rescaled the time series into the unit interval taking the uniform deviate transformation before constructing the network as discussed in \cite{34}. A RN from a random time series is also constructed for comparison taking \( M = 3 \).

We first study how the value of \( H_m \) varies with \( N \) for RNs. In Fig. 10 (top panel), we show the results for the Lorenz attractor and the Henon attractor. It is evident that for large value of \( N \), \( H_m \) converges to a finite value. We have checked and verified that this is true for other low dimensional chaotic attractors as well, considered in this paper. It is found that once the basic structure of the attractor is formed, the value of \( H_m \) remains independent for further increase in \( N \). In other words, the range of \( k \) values increases with \( N \) to keep the value of \( H_m \) constant. This result also follows from the statistical invariance of the degree distribution of the RN as has already been shown \cite{34}.

Next, we consider the variation of \( H_m \) with embedding dimension \( M \). This is also shown in Fig. 10 (bottom panel) for two standard chaotic attractors. It is clear that the value of \( H_m \) converges for \( M \geq 3 \) in both cases. To understand this, we compute the degree distribution of the two attractors as a function of \( M \). The results, shown in Fig. 11, indicate that the degree distribution itself shows convergence for any dimension \( M \) equal to or greater than the actual dimension of the system. Thus, \( H_m \) can be considered to be a unique measure for any chaotic attractor independent of both \( M \) and \( N \).

From the construction of RNs, the range of interaction between two nodes is limited by the recurrence threshold \( \epsilon \). Hence the degree of a node in the RN and the probability density around the corresponding point over the attractor are directly correlated. For example, for the RN from a random time series, every node has degree close to the average value \( < k > \) since the probability density over the attractor is approximately the same. One can show that the degree distribution of the RN from a random time series is Gaussian for large \( N \). Thus, the \( k \) spectrum of the RN is indicative of the range in the probability density variations over the attractor, which in turn, is characteristic of the structural complexity of the attractor.

We have already shown that the measure \( H_m \) proposed here is indicative of the diversity in the \( k \) spectrum. Moreover, it is found to have a specific value for a given attractor independent of \( M \) and \( N \). These two results imply that \( H_m \) can serve as a single index to quantify the structural complexity of a chaotic attractor through RN construction. In Fig. 12 we show the \( k \) spectrum of the RN from the Lorenz and Henon attractors. For comparison, we also show the \( k \) spectrum for the RN from a random time series, a RG of same \( < k > \) as the random time series and that for a typical SF network. In all cases, we have used \( N = 2000 \). The \( k \) spectrum of the RN from standard chaotic attractors is almost continuous having a wide range of \( k \) values characteristic of the probability density variations over the attractor. The values of \( H_m \) for all these networks as well as the RNs from other standard chaotic attractors are shown in Table I. For each attractor, the value of \( H_m \) converged upto \( M = 5 \) is shown. In each case, five different RNs are constructed changing the initial conditions and the average is shown with standard deviation as the error bar. For RG and SF networks, average of five different simulations is taken.

Note that, even though the largest degree \( (k_{max}) \) is more for the SF network compared to the RN of Lorenz, its \( H_m \) is less since the number of non zero \( k \) values is comparatively less. Also, from our results it can be concluded that among the continuous systems compared, the Lorenz attractor is structurally the most complex while in the case of 2D discrete systems, Lozi attractor is found to be the most diverse in terms of the probability density variations. Finally, the \( k \)-spectrum and \( H_m \) for the RN from a random time series is exactly same as that of a RG with identical degree distribution.

VI. HETEROGENEITY OF REAL WORLD NETWORKS

So far, we have been discussing the heterogeneity measure of synthetic networks of different topologies. In this section, we consider some unweighted and undirected complex networks from the real world and see what information
FIG. 9: Top panel shows the variation of $H_m$ with $N$ for RNs constructed from Lorenz (filled triangle) and Henon (filled circle) attractor time series. For Lorenz, $M = 3$ and $\epsilon = 0.1$ while for Henon, $M = 2$ and $\epsilon = 0.06$. Bottom panel shows the variation of $H_m$ with $M$ for fixed $N$ for Lorenz (filled triangle) and Duffing (filled circle) attractors. In both graphs, the error bar comes from the standard deviation of values for $H_m$ computed from time series with five different initial conditions.

regarding the heterogeneity of such networks can be deduced using the proposed measure. We use data on networks from a cross section of fields, such as, biological, technological and social networks. In Fig. 12, we show the degree distribution and $k$ spectrum of two such networks. In Table II, we compile the details of these networks and the values of $H_m$ computed by us for each.

Since we have to restrict to the case of unweighted and undirected networks, we could use only a small subset from the very large variety of real world networks that are mostly weighted or directed. To extend the measure to directed networks, one has to consider the in-degree and out-degree distributions and find the heterogeneity separately. In order to generalise the measure to weighted networks, the distribution of the weight or strength of the nodes in the network rather than the simple degree distribution is to be considered and define the measure accordingly. For example, for unweighted and undirected networks, all the links are equivalent and hence the degree of $i^{th}$ node $k_i$ is just the sum of the links connected to node $i$. On the other hand, for weighted networks, each link is associated with a weight factor $w_{ij}$ and hence the degree $k_i$ should be generalised to the sum of the weights of all the links attached
FIG. 10: Variation of the degree distribution of the RN with $M$ for the Lorenz and Duffing attractors. In both cases, red thin solid line is for $M = 2$, green thick solid line is for $M = 3$, dashed line in magenta is for $M = 4$ and blue dotted line is for $M = 5$.

Thus the degree distribution needs to be generalised to the strength distribution $P(s)$, which is the probability that a given node has a strength equal to $s$ [40, 41]. The equation for heterogeneity for weighted networks can be modified accordingly. However, it should be noted that the weight factors are assigned based on different criteria depending on the specific system or interaction the network tries to model. Modifying the measure by incorporating the specific aspects of interaction, the measure itself becomes network specific. The measure that we propose here is independent of such details and is representative of only the diversity of node degrees in a network, determined completely by the simple degree distribution.

to node $i$:

$$s_i = \sum_j w_{ij}$$  \hspace{1cm} (10)
FIG. 11: Comparison of the $k$ spectrum for the RNs of random time series and two chaotic attractors along with that for a RG and a SF synthetic network with $N = 2000$ in all cases. The natural dimension of the attractor is chosen as the embedding dimension for constructing the RN as indicated.

VII. CONCLUSION

Complex networks and the network based quantifiers have become useful tools for the analysis of many real world phenomena. An important measure for the characterization of any complex network is its heterogeneity. Though the degree distributions can give some insight on how heterogeneous a network is, they are unable to give a unique characterization particularly when different distributions approximately fit the same network and when two networks with completely different degree distributions are compared. Several attempts have been made previously to define a unique measure applicable to networks with different topologies, but none of them have so far addressed the question of what logical limit can be set for getting complete heterogeneity in a network of size $N$. This idea leads us to the proposed measure of heterogeneity that is normalised between 0 and 1.

Also, the measure that we propose uniquely quantifies the diversity in the node degrees in the network which is characteristic of the type and range of interactions the network represents. The diversity also depends on the topology of the resulting network. For example, for RGs, the diversity is limited since most degrees are centered around the average value $< k >$ while the SF networks are comparatively more diverse due to the presence of hubs. The proposed measure can quantify this diversity in the node degrees irrespective of the topology of the network.

By applying the proposed measure, we compute the heterogeneity of various unweighted and undirected networks, synthetic as well as real world. We study numerically how the heterogeneity varies for RG with respect to the two parameters $p$ and $N$, while for SF networks we specifically show that the heterogeneity decreases as a power law as a function of the network size $N$. We also show how the measure can be used for the analysis of RNs constructed from the time series of chaotic dynamical systems and its utility as a quantifier to compare the structural complexities of different chaotic attractors.

Acknowledgments

RJ and KPH acknowledge the financial support from the Science and Engineering Research Board (SERB), Govt. of India in the form of a Research Project No. SR/S2/HEP-27/2012. KPH acknowledges the computing facilities in
FIG. 12: Degree distribution and the $k$ spectrum of the protein interaction network are shown in the top panel. In the bottom panel, the same for the network of Western Power Grid.
### TABLE I: Comparison of $H_m$ for several chaotic attractors along with that for a RG ($p = 0.0035$) and a SF network ($\gamma = 2.124$).

| System                      | $H_m$       |
|-----------------------------|------------|
| Random Graph                | 0.0079 ± 0.0011 |
| Scale Free Network          | 0.0128 ± 0.0032 |
| Random Time Series          | 0.0079 ± 0.0037 |
| Lorenz Attractor            | 0.0368 ± 0.0028 |
| Rössler Attractor           | 0.0312 ± 0.0014 |
| Duffing Attractor           | 0.0294 ± 0.0034 |
| Ueda Attractor              | 0.0285 ± 0.0018 |
| Henon Attractor             | 0.0724 ± 0.0024 |
| Lozi Attractor              | 0.0813 ± 0.0046 |
| Cat Map                     | 0.0154 ± 0.0022 |

### TABLE II: Comparison of the heterogeneity measure of five real world networks

| System          | Reference                              | $N$   | $H_m$ |
|-----------------|----------------------------------------|-------|-------|
| US Power Grid   | [http://cdg.columbia.edu/cdg/datasets](http://cdg.columbia.edu/cdg/datasets) | 4941  | 0.0029 |
| Protein Interaction | [www3.nd.edu/ networks/resources.html](http://www3.nd.edu/ networks/resources.html) | 1846  | 0.0132 |
| Budding Yeast   | [math.nist.gov/](http://math.nist.gov/) | 2353  | 0.0233 |
| US Patent Citation | [https://snap.stanford.edu/data](https://snap.stanford.edu/data) | 7253  | 0.0081 |
| Dolphin Interaction | [https://snap.stanford.edu/data](https://snap.stanford.edu/data) | 62    | 0.1486 |

IUCAA, Pune. We use the open source software GEPHI [www.gephi.org/](http://www.gephi.org/) for the construction of all networks.

---

1. Newman, M. E. J. (2010) *Networks: An Introduction*, (Oxford Univ. Press, Oxford).
2. Barzel, B. & Barabasi, A. L. (2013) Universality in network dynamics, *Nature Phys.*, 9, 673-681.
3. Newman, M. E. J. (2012) Communities, modules and large scale structure in networks, *Nature Phys.*, 8, 25-31.
4. Albert, R., Albert, I. & Nakarado, G. L. (2004) Structural vulnerability of the North American power grid, *Phys. Rev. E*, **69**, 025103(R).
5. Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Orr, S. S., Ayzenshtat, I., Sheffer, M. & Alon, U. (2004) Super families of evolved and designed networks, *Science*, **303**, 1538-42.
6. Boccaletti, S., Latora, V., Moreno, Y., Chavez M. & Hwang, D. U. (2006) Complex Networks: Structure and Dynamics, *Phys. Reports*, **424**, 175-308.
[7] Yook, S. H., Jeong, H., Barabasi, A. L. & Tu, Y. (2001) Weighted evolving networks, Phys. Rev. Lett., 86, 5835-38.
[8] Barrat, A., Barthelemy, M., Pastor-Satorras, R. & Vespignani, A. (2004) The architecture of complex weighted networks, Proc. Nat. Acad. Sci. USA, 101, 3747-52.
[9] Erdős, P. & Rényi, A. (1960) On the evolution of random graphs, Publ.Math.Inst.Hung.Acad. Sci., 5, 17-61.
[10] Caldarelli, G. (2007) Scale-free networks, (Oxford Univ. Press, Oxford).
[11] Barabasi, A. L. & Albert, R. (1999) Emergence of scaling in complex networks, Science, 286, 509-512.
[12] Albert, R. & Barabasi, A. L. (2002) Statistical mechanics of complex networks, Rev. Mod. Phys., 74, 47-97.
[13] Estrada, E. (2011) The structure of complex networks: Theory and applications, (Oxford Univ. Press, Oxford).
[14] Xu, X., Zhang, J. & Small, M. (2008) Super family phenomena and motifs of networks induced from time series, Proc. Nat. Acad. Sci. USA, 105, 19601-05.
[15] Estrada, E. (2010) Quantifying network heterogeneity, Phys. Rev. E, 82, 066102.
[16] Crucitti, P., Latora, V. & Marchioni, M. (2003), Efficiency of scale-free networks: error and attack tolerance, Physica A, 320, 622.
[17] Pastor-Satorras, R. & Vespignani, A. (2001), Epidemic dynamics and endemic states in complex networks, Phys. Rev. E, 63, 06617.
[18] Zhao, L., Lai, Y-C., Park, K. & Ye, N. (2007), Heterogeneity of scale free networks, Phys. Rev. E, 71, 026125.
[19] Motter, A. E., Zhou, C.-S. & Kurths, J. (2005), Network synchronization, diffusion and the paradox of heterogeneity, Phys. Rev. E, 71, 016116.
[20] Donner, R. V., Zou, Y., Donges, J. F., Marwan, N. & Kurths, J. (2010) Recurrence networks: A novel paradigm for nonlinear time series analysis, New J. Phys., 12, 033025.
[21] Marwan, N. & Kurths, J. (2015) Complex networks based techniques to identify extreme events and transitions in spatio-temporal systems, CHAOS, 25, 097609.
[22] Avila, R., Gapelyuk, A., Marwan, N., Walther, T., Stepan, H., Kurths, J. & Wessel, N. (2013) Classification of cardio-vascular time series based on different coupling structures using recurrence network analysis, Philos. T. Roy. Soc. A, 371, 20110623.
[23] Snijders, T. A. B. (1981) Social Networks, 3, 163-68.
[24] Bell, F. K. (1992) Linear Algebra Appl., 161, 45-54.
[25] Albertson, M. O. (1997) Ars. Comb., 46, 219-26.
[26] Wu, J., Tan, Y. J., Deng H. Z. & Zhu, D. Z. (2008) Heterogeneity of scale free networks, Systems Engg.-Theor. and Practice, 27, 101-05.
[27] Hu, H.-B. & Wang, X.-F. (2008), Unified index to quantifying heterogeneity measure of complex networks, Physica A, 387, 3769-80.
[28] Colander, D. C. (2001) Microeconomics (McGraw-Hill, Boston).
[29] Donner, R. V., Small, M., Donges, J. F., Marwan, N., Zou, Y., Xiang, R. & Kurths, J. (2011) Recurrence based time series analysis by means of complex network methods, Int. J. Bif. Chaos, 21, 1019-46.
[30] Sprott, J. C. (2003) Chaos and Time Series Analysis (Oxford Univ. Press, Oxford).
[31] Eckmann, J. P., Kamphorst, S. O. & Ruelle, D. (1987) Recurrence plot of dynamical systems, Europhys Letters, 5, 973-977.
[32] Donges, J. F., Heitzig, J., Donner R. V. & Kurths, J. (2012) Analytical framework for recurrence network analysis of time series, Phys. Rev. E, 85, 046105.
[33] Eroglu, D., Marwan, N., Prasad, S. & Kurths, J. (2014) Finding recurrence network threshold adaptively for a specific time series, Nonlin. Processes Geophys., 21, 1085.
[34] Rinku Jacob., Harkrishnan, K. P., Misra, R. & Ambika, G. (2016) Uniform framework for the recurrence-network analysis of chaotic time series, Phys. Rev. E, 93, 012202.
[35] Watts, D. J. & Strogatz, S. H. (1998) Collective dynamics of small world networks, Nature, 393, 440-42.
[36] Jeong, H., Mason, S., Barabasi, A. L. & Oltvai, Z. N. (2001) Contrality and lethality of protein networks, Nature, 411, 41-44.
[37] Sun, S., Ling, L., Zhang, N., Li, G. & Chen, R. (2003) Topological structure analysis of protein-protein interaction network in budding yeast, Nucleic Acids Research, 31, 2443-50.
[38] Lusseau, D. (2003) The emergent properties of a dolphin social network, Proc. R Soc. Lond. B (Suppl.), 270, S186-188.
[39] Albert, R. & Barabasi, A. L. (2000) Topology of evolving networks: local events and universality, Phys. Rev. Lett., 85, 5234-37.
[40] Barthelemy, M., Barrat, A., Pastor-Satorras, R. & Vespignani, A. (2005) Characterisation and modelling of weighted networks, Physica A, 346, 34-43.
[41] Latora, V. & Marchiori, M. (2003) Economic small world behavior in weighted networks, European Phys. J. B, 32, 249-63.