

OPEN

# Controlling the Multifractal Generating Measures of Complex Networks

Ruochen Yang &amp; Paul Bogdan\*

Mathematical modelling of real complex networks aims to characterize their architecture and decipher their underlying principles. Self-repeating patterns and multifractality exist in many real-world complex systems such as brain, genetic, geoscience, and social networks. To better comprehend the multifractal behavior in the real networks, we propose the weighted multifractal graph model to characterize the spatiotemporal complexity and heterogeneity encoded in the interaction weights. We provide analytical tools to verify the multifractal properties of the proposed model. By varying the parameters in the initial unit square, the model can reproduce a diverse range of multifractal spectrums with different degrees of symmetry, locations, support and shapes. We estimate and investigate the weighted multifractal graph model corresponding to two real-world complex systems, namely (i) the chromosome interactions of yeast cells in quiescence and in exponential growth, and (ii) the brain networks of cognitively healthy people and patients exhibiting late mild cognitive impairment leading to Alzheimer disease. The analysis of recovered models show that the proposed random graph model provides a novel way to understand the self-similar structure of complex networks and to discriminate different network structures. Additionally, by mapping real complex networks onto multifractal generating measures, it allows us to develop new network design and control strategies, such as the minimal control of multifractal measures of real systems under different functioning conditions or states.

Technological advances and information digitization contribute not only to richer complex multi-modal heterogeneous and noisy datasets in diverse fields such as social, geoscience, brain and biological networks<sup>1,2</sup>, but at the same time call for advanced mathematical techniques for mining and investigating complex multiscale and spatiotemporal relationships. Much of the complex network literature focuses on developing mathematical network models that characterize one or few pairwise interactions<sup>3–6</sup>. More recently, several network modelling approaches study the network complexity from a geometrical perspective<sup>7–12</sup>.

Since, in various settings, we deal with networks with unknown rules of network growth and dynamic evolution (also referred as non-stationary interactions), these approaches prove beneficial to characterize some new network properties. However, in majority of the real-world experiments, we only have access to partial observations, or we notice that the interaction rules are highly heterogeneous across space and time. Modelling a system with such limitations requires a new mathematical formalism that can not only extract the generalized geometric signatures across scales, but also provide tools for scaling analysis of motifs and patterns. Additionally, the framework is desired to enable new control strategies that go beyond node state and target geometric characteristics of the network (e.g., network curvature, multifractal spectrum).

From a geometrical perspective, many large-scale complex networks from sociology and biology exhibit self-similar and multifractal characteristics<sup>13,14</sup>. Multifractal geometric analysis makes it possible to capture the heterogeneous and multiscale interaction rules of large networks<sup>15,16</sup>. It efficiently characterizes large-scale complex systems<sup>16</sup> and can be employed to measure nodes similarity<sup>17</sup> and detect community structures<sup>18</sup>. For instance, the multifractality of geochemistry mapping explains the element concentration values distribution and spatial covariance structure in rock samples<sup>19</sup>. In multifractal formalism, the renormalization procedure of coarse-graining networks into boxes (i.e., box-counting method) helps at examining the existence of scaling and heterogeneity of motifs and rules in networks. The self-similar exponent is defined by the renormalization and a

Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, 90089, United States. \*email: [pbogdan@usc.edu](mailto:pbogdan@usc.edu)

noticeable power-law relationship is found between the number of boxes needed to cover the nodes and the scale of boxes<sup>18,20</sup>.

Building on the idea of self-repeating structures, the de Wijs model<sup>21</sup> partitions a unit segment into two sub-segments of equal size and further repeats this operation recursively, which leads to the Mandelbrot multifractal<sup>22,23</sup>. The model construction from multiplicative cascade results in the discrete log-binomial distribution<sup>19,24–28</sup>. In contrast, the stochastic Kronecker graph represents a random graph obtained from the Kronecker product of probability matrices, thus displaying a self-similarity. It is able to reproduce many properties in real networks including heavy tail degree distribution, low diameter and densification power law<sup>29–31</sup>. Along the same lines, the multifractal network generator generates random graphs with a wide variety of statistical properties<sup>32,33</sup>. The multiplicative attribute graph generalizes the stochastic Kronecker graph and the multifractal network generator<sup>34,35</sup>. By characterizing the interaction between nodes with different attributes, it models network structures as well as node property and information.

Prior works on modelling self-similar networks focus on unweighted graphs. However, as demonstrated in the literature<sup>7</sup>, much of the complexity is encoded in the heterogeneous weights of the interactions. Pure topological models are binary and they are inadequate to capture complicated properties of real complex networks<sup>36</sup>. In lots of situations, we deal with large complex networks where the interactions are heterogeneous and possibly varying. Introducing weights/strength can ameliorate existing binary random models to better depict and understand networks in real systems. Furthermore, there is a lack of analytical multifractal analysis for random graphs. Box-covering methods over graphs is equivalent to graph coloring problem, which is a known NP-complete problem. It's even more difficult to apply this node-based analysis on random graphs. Moreover, we do not only want to estimate and analyze network multifractality, but also want to control the multifractality and reflect the performance of robustness and resilience in complex networks.

To address these research gaps, in this paper, we propose the *weighted multifractal graph* (WMG) model. The WMG model captures and generates weighted multifractal networks by mapping from recursively constructed measures of linking probability. Rather than covering graph nodes with minimum number of boxes, we analyze and analytically quantify the multifractal properties of networks via counting edges that satisfy similar generating rules. We analytically show the multifractal properties of the WMG model. The WMG model could produce and account for a wide variety of multifractal spectrums with varying degrees of symmetry, locations, support and shapes. It can be employed to model the complexity of a wide array of networks from brain, genomics, proteomics, social and geoscience disciplines. The benefits of the WMG formalism is twofold: (1) It helps to recover mathematical models which are capable to discriminate between different structures of a complex system. (2) It enables us to understand and control the multifractality of complex networks with minimal intervention. It provides comprehension on how to explore the mechanism of evolving networks and how to regulate complex networks under different states.

## Results

**Weighted multifractal graph model.** In order to generate weighted networks with multifractal characteristics and to minimally control complex networks, we introduce the *weighted multifractal graph* (WMG) model to generate and capture random weighted graphs with multifractal topology. By choosing a few parameters, the model can generate a wide variety of weighted multifractal topologies with arbitrary statistical properties, such as generalized degree distribution and clustering coefficient. This multifractal model can be used to fit diverse real-world datasets in fields such as biological systems, geoscience and financial markets. Also, it enables to control multifractality of networks with mild modifications and adjustments.

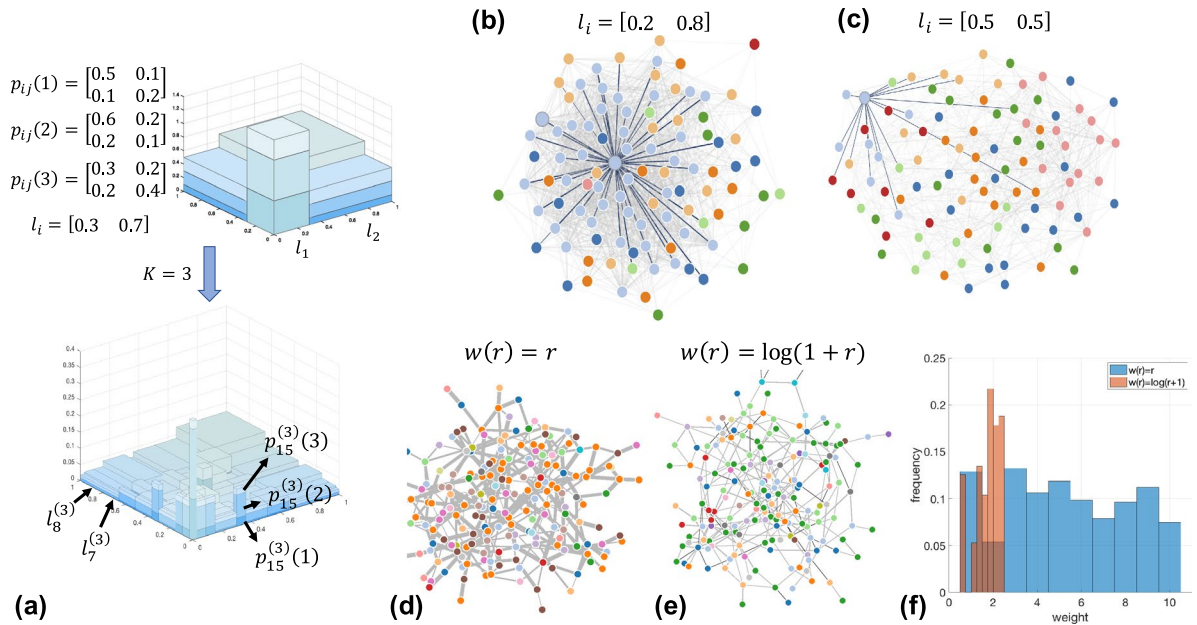
Along the same lines as in the multifractal network generator<sup>32</sup>, we generate the WMG model recursively from simple square structures. First, we define simple generating measures on  $R$ -layer unit squares. The  $x$ -axis and  $y$ -axis of each layer are identically divided into  $M$  intervals, and each interval length is denoted as  $l_i$ ,  $i = 1 \dots M$ . The unit square is divided into  $M^2$  rectangles, each rectangle  $(i, j)$  is assigned with a set of probabilities  $p_{ij}(r)$ ,  $i, j = 1 \dots M$ ,  $r = 1 \dots R$ . The WMG model is developed from this initial generating measure. The self-similarity appears when we divide each rectangle in the unit square into  $M^2$  sub-rectangles, which has the same structure as the unit one. The self-similar WMG model is formed after repeating the operation  $K$  times. Meanwhile, the assigned probabilities  $p_{ij}^{(K)}(r)$  and interval lengths  $l_i^{(K)}$  in the WMG model are the product of probabilities and interval lengths in every division. The expression of  $p_{ij}^{(K)}(r)$  and  $l_i^{(K)}$  can be written as

$$p_{ij}^{(K)}(r) = \prod_{q=0}^{K-1} p_{i_q j_q}(r), \quad (1)$$

$$l_i^{(K)} = \prod_{q=0}^{K-1} l_{i_q}, \quad (2)$$

where  $i_q = \left\lfloor \frac{i-1}{M^q} \right\rfloor \bmod M + 1$ ,  $j_q = \left\lfloor \frac{j-1}{M^q} \right\rfloor \bmod M + 1$ , and 'mod' represents the modulo operation. Of note, the values of  $l_i$  are identical for all layers. The  $l_i$  values are independent of the weight layer index  $r$ . Furthermore, this generating measure is defined on unit squares and satisfies  $\sum_{i_0=1}^M l_{i_0} = 1$  and  $\sum_{i=1}^M l_i^{(K)} = 1$ .

When generating a random multifractal graph, we first select a discrete weight set  $\{w(r)\}_{r=1:R}$ . The  $r$  and  $w(r)$  are the index of the discrete weight level and the corresponding weight value, respectively.  $p_{ij}^{(K)}(r)$  and  $l_i^{(K)}$  are the corresponding generating measures of the model at  $r$ -th layer. The weight set  $\{w(r)\}_{r=1:R}$  can be any desired positive and real values according with specific distributions for specific (desired) network features. Here  $R$  is the

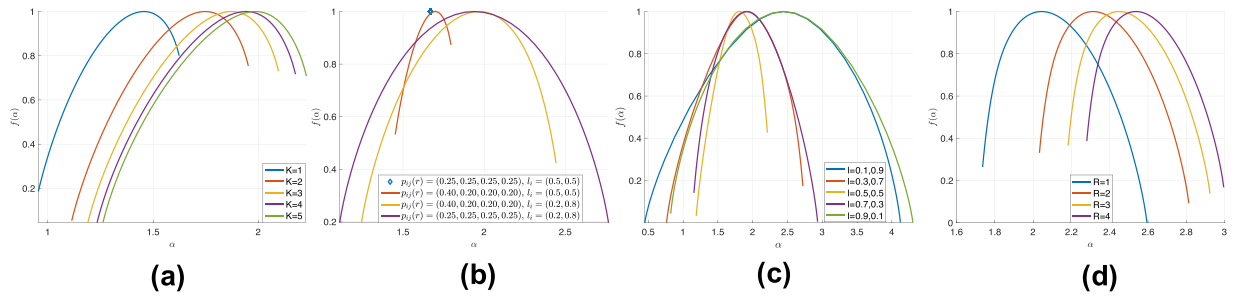


**Figure 1.** (a) Schematic illustration of a simple recursive WMG model generation with  $M = 2, R = 3$ . (b) A network generated with model interval length  $l_i = [0.2 \ 0.8]$ . (c) A network generated with model interval length  $l_i = [0.5 \ 0.5]$ . (d) A network generated with weight distribution  $w(r) = r$  for  $r = 1 \dots 10$ . (e) A network generated with weight distribution  $w(r) = \log(1 + r)$  for  $r = 1 \dots 10$ . (f) Normalized histogram of the weights in (d,e).

resolution level of the discrete weights. Choosing larger  $R$  leads to higher graph resolution and more parameters  $\{p_{ij}(r)\}$  will be needed to be introduced. Like De Wijs model<sup>24,25</sup>, Kronecker random graph<sup>29–31</sup>, stochastic block model<sup>34</sup> and multifractal network generator<sup>32</sup>, a weighted random graph with  $N$  nodes is generated from the linking probability measures using the following strategy: (1) Generate a series of uniformly distributed random variables  $U(n)$  for each node  $n = 1 \dots N$  independently. (2) Find the category  $i(n)$  of each node  $n$  such that  $\sum_{t=1}^{i(n)-1} l_t^{(K)} < U(n) \leq \sum_{t=1}^{i(n)} l_t^{(K)}$ . In the recursive model with  $K$  iterations, this suggest that the node  $n$  falls into sub-interval  $i(n)$  on the unit edge. Since  $l_i$  and  $l_i^{(K)}$  are identical for all layers, the categories  $i(n)$  assigned for the node  $n$  in every weight layer are also identical. (3) For each pair of nodes  $(u, v)$ , generate an uniformly distributed random variable  $L(u, v)$ , and find index  $r$  such that it satisfies  $\sum_{t=1}^{r-1} p_{i(u)i(v)}^{(K)}(t) < L(u, v) \leq \sum_{t=1}^r p_{i(u)i(v)}^{(K)}(t)$ . Next, we assign the weight value  $w(r)$  to the link between nodes  $u$  and  $v$ . If there is no  $r$  satisfying this condition, then there will be no edge between this pair of nodes  $(u, v)$ . It is noteworthy that when  $M^K \approx N$  and  $\{l_i\}_{i=1:M}$  is equally sized, if each node  $n$  corresponds with one category  $i(n)$  without repetition, this special case of WMG model would reduce to Kronecker graph<sup>29–31</sup>.

The iterative WMG model building procedure is illustrated in Fig. 1(a). The presented model example has 3 layers. In this example,  $\{l_i^{(K)}\} = [0.027 \ 0.063 \ 0.063 \ 0.147 \ 0.063 \ 0.147 \ 0.147 \ 0.343]$  is built from Eq. (2) with  $\{l_i\} = [0.3 \ 0.7]$ . Two networks in Fig. 1(b,c) are generated under the same model parameters  $M = 2, R = 2, K = 3, p_{ij}(1) = \begin{bmatrix} 0.9 & 0.3 \\ 0.3 & 0.3 \end{bmatrix}, p_{ij}(2) = \begin{bmatrix} 0.1 & 0.2 \\ 0.2 & 0.9 \end{bmatrix}$  and  $w(r) = [1 \ 2]$ . The color of nodes represents generating category and the stroke width of edges displays the weight. We highlight one node in each graph, its neighbours and the edges connecting them. As shown in Fig. 1(b), 45 of 100 nodes are colored with light blue and they are in the same category  $i = 8$  because  $l_8^{(3)} = 0.512$ . The neighbours of the highlighted node in the middle are more likely to have the same category (34 of 45 nodes, category  $i = 8$  in light blue). As a result of the relatively large value of  $p_{88}^{(3)}(2) = 0.729$ , the edges connecting them have weight  $w(2) = 2$ . By contrast, in Fig. 1(c), the categories of nodes are more evenly distributed because the interval lengths are  $l_i^{(3)} = 0.125$  for  $i = 1 \dots 8$ . The highlighted node in Fig. 1(c) has category  $i = 1$ . In this example,  $p_{1i}^{(3)}(1) > p_{1i}^{(3)}(2)$  for  $i = 1 \dots 8$ . Therefore, in this generated graph, all the links containing the highlighted node have weight  $w(1) = 1$ . Fig. 1(d,e) show two networks generated by different weight distributions  $w(r) = r$  and  $w(r) = \log(1 + r)$ . While all other parameters are kept the same, from Fig. 1(d,e), we can see that a change in  $w(r)$  from  $r$  to  $\log(r + 1)$  leads to networks with different weight features. In Fig. 1(f) we show the normalized histogram of the weights in two networks (d,e). Next, we present the multifractal analysis of this WMG model.

**Multifractal analysis.** Traditional approaches for quantifying the multifractal properties of networks rely on box-counting and box-covering method<sup>18,20,37</sup>. The fractal dimension can also be obtained with the help of closeness centrality<sup>38</sup>. However, finding the minimum number of boxes of a given radius required to cover the graph is as hard as the graph coloring problem<sup>20</sup>. Introducing randomness in the graph model brings uncertainty to the box-counting problem. Although there exist few analytical results that determine inequality constraints on



**Figure 2.** (a) Multifractal spectrums with  $K = 1 \dots 5$ ,  $M = 2$ ,  $R = 3$ ,  $p_{ij}(r)$  and  $l_i$  remain the same. (b) Multifractal spectrums with four different sets of initial unit square model parameters.  $M = 2$ ,  $R = 3$ ,  $K = 4$ . (c) Multifractal spectrums with  $\{l_i\} = (0.1, 0.9)$ ,  $(0.3, 0.7)$ ,  $(0.5, 0.5)$ ,  $(0.7, 0.3)$ ,  $(0.9, 0.1)$ .  $K = 5$ ,  $M = 2$ ,  $R = 3$ ,  $p_{ij}(r)$  remains the same. (d) Multifractal spectrums with  $R = 1 \dots 4$ .

the chromatic number for the graph coloring problem in Erdős-Rényi model<sup>39–41</sup>, these inequalities cannot be analytically extended to networks with higher-order correlations and so determine their box covering.

To elucidate the multifractal properties of the above-mentioned weighted network construction strategy, we describe next an analytical framework capable of estimating the partition function, the mass exponent, the generalized fractal dimension, the Lipschitz-Hölder exponent and the multifractal spectrum. Since the link generating process contains both node attributes and edge formation characteristics, the proposed mathematical formalism estimates the number of edges which are generated under the same rule. While this formalism bears some similitudes with the multiplicative cascade model<sup>26–28</sup>, it counts the number of edges in the generated graph rather than counting sub-blocks with same structure. Our edge-based strategy enables us to analytically compute the partition function and derive the multifractal related metrics (e.g., generalized fractal dimension, multifractal spectrum). For simplicity, let  $\{p_{ij}(r)\}_{i,j=1:M}$  denote the edge generating probability in the initial unit square model and  $\{l_i\}_{i=1:M}$  denote the length of the intervals, respectively. With these conventions, the derivation of the partition function reads:

$$Z_\epsilon(q) = \sum_{r=1}^R \left\{ \sum_{i,j=1}^M (l_i l_j p_{ij}(r))^q \right\}^K \tag{3}$$

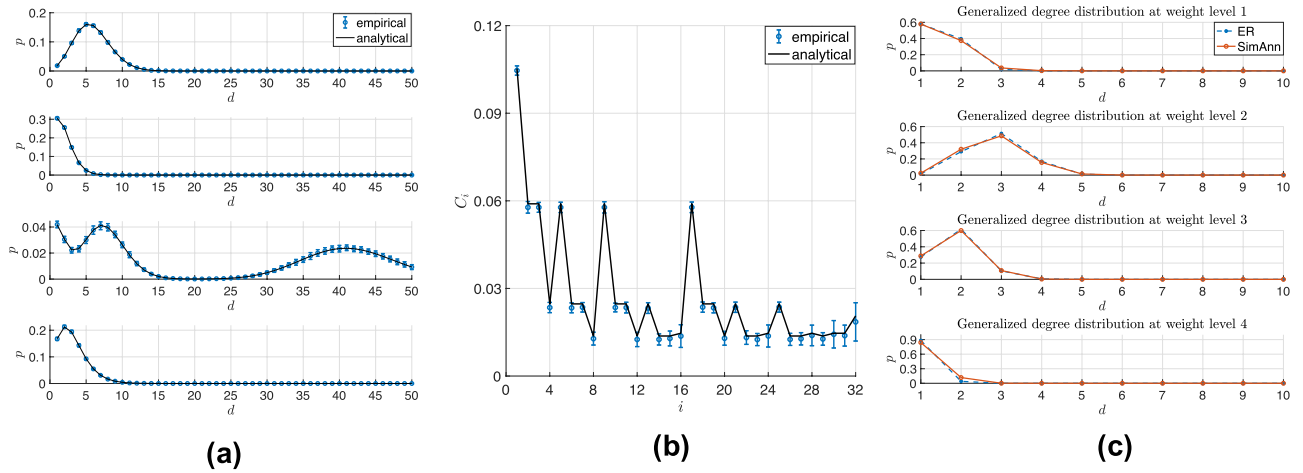
The analytical derivation of the partition function allows us to calculate the mass exponent as  $\tau(q) = \frac{\log Z_\epsilon(q)}{\log \epsilon}$ , the generalized fractal dimension from  $D(q) = \frac{\tau(q)}{1-q}$ , the Lipschitz-Hölder exponent (also refer to coarse Hölder exponent or singularity index in prior work)  $\alpha(q) = \frac{d\tau(q)}{dq}$  and the multifractal spectrum from  $f(\alpha) = \alpha(q)q - \tau(q)$ .

The networks generated by the proposed WMG model inherit a rich variety of multifractality compared with the unweighted MFNG model<sup>32</sup>. For example, when  $R = 1$  (corresponding to MFNG model), the partition function in equation (3) shows that the variation in  $K$  does not influence significantly the mass exponent  $\tau(q)$ , the generalized multifractal dimension  $D(q)$ , the Lipschitz-Hölder exponent  $\alpha(q)$  and the multifractal spectrum  $f(\alpha)$  because of the log operation. However, as shown in Fig. 2(a), for  $M = 2$ ,  $R = 3$  and same  $\{p_{ij}(r)\}$  and  $\{l_i\}$ , the variation in the model iteration step  $K$  causes a shift of multifractal spectrum towards a support on higher Lipschitz-Hölder exponents. One can observe that the shift in the support of the multifractal spectrum is smaller with increasing changes  $\Delta K$  in the model iteration steps  $K$ .

The parabola-like curve shows the multifractality of the WMG model. Fig. 2(b) presents more results of multifractal spectrum curves with four different sets of initial unit square model parameters  $\{p_{ij}(r)\}_{i,j=1:M, r=1:R}$  and  $\{l_i\}_{i=1:M}$ . Here, we consider  $M = 2$ ,  $R = 3$  and  $K = 4$ . The blue diamond is generated by equally assigning model parameters  $p_{ij}(r) = (0.25, 0.25, 0.25, 0.25)$ ,  $l_i = (0.5, 0.5)$ . The red curve corresponds to assigning  $p_{ij}(r) = (0.40, 0.20, 0.20, 0.20)$ ,  $l_i = (0.5, 0.5)$ . The yellow curve corresponds to  $p_{ij}(r) = (0.40, 0.20, 0.20, 0.20)$ ,  $l_i = (0.2, 0.8)$ . The purple curve is given by  $p_{ij}(r) = (0.25, 0.25, 0.25, 0.25)$ ,  $l_i = (0.2, 0.8)$ . Each set of parameters are the same for every weight layer  $r = 1 \dots R$ .

Of note, for identical parameters (i.e.  $p_{ij}(r) = (0.25, 0.25, 0.25, 0.25)$ ), the WMG generates a monofractal network (represented in Fig. 2(b) by a blue diamond marker). More precisely, this implies that a more evenly assigned  $p_{ij}(r)$  leads to more symmetric and narrower curve of the multifractal spectrum (which corresponds to a point for identical  $p_{ij}(r)$ ). A shift and reshape of the curve can also be caused by changing  $l_i$ .

We observe that there is an interesting effect of the sub-interval length  $\{l_i\}$  on the multi-fractal spectrum  $f(\alpha)$ . For example, as presented in Fig. 2(c), when node generating probabilities are identical ( $\{l_i\} = (0.5, 0.5)$  - yellow curve), the maximum value of multifractal spectrum  $f(\alpha)$  places at leftmost compared with other curves given by asymmetric  $\{l_i\}$ . While interchanging the  $\{l_i\}$  (i.e.,  $\{l_i\} = (0.3, 0.7)$  - red vs.  $\{l_i\} = (0.7, 0.3)$  - magenta, or  $\{l_i\} = (0.1, 0.9)$  - blue vs.  $\{l_i\} = (0.9, 0.1)$  - green), the maximum value of the multifractal spectrum remains unchanged but the left and right hand side end points of the spectrum change. Similarly, the influence of  $\{p_{ij}(r)\}$  can also be explored.



**Figure 3.** (a) Empirical and analytical generalized degree distribution. (b) Empirical and analytical results of clustering coefficients. This simulation is generated under parameters  $M = 2, K = 5, R = 4$ . The scale (number of nodes) of graph  $N = 5000$ . Graph generation is repeated 100 times. (c) The WMG model can represent networks generated by ER model.

Fig. 2(d) shows the influence of the weight resolution  $R$ . For different  $R$ , we assign the parameters  $p_{ij}(r) = \frac{c_{ij}}{R}$ ,  $r = 1 \dots R$ , where  $c_{ij} = \begin{bmatrix} 0.25 & 0.29 \\ 0.29 & 0.17 \end{bmatrix}$ . Therefore, the probability that an edge exists  $\sum_{r=1}^R p_{ij}^{(K)}(r)$  remains the same among different  $R$ . It shows that increasing weight resolution  $R$  causes a right shift of multifractal spectrum.

**Statistical properties.** To provide a first order characterization of the heterogeneity of complex weighted networks, we analyze the generalized degree  $\{D_r\}_{r=1 \dots R}$  of a node, which quantifies the number of connections (edges) with an associated weight  $w_r$ . The generalized degree distribution  $p_r(k_r) = P\{D_r = k_r\}$  retains more information of the network topology compared to the straightforward strength distribution, which helps us to better understand the network structure and its multifractality.

Via the extension of the multifractal network generator<sup>32</sup>, in the WMG model, the generalized degree distribution is expressed as

$$p_r(k_r) = \sum_{i=1}^{M^K} p_{i,r}(k_r) l_i^{(K)}, \tag{4}$$

where  $p_{i,r}(k_r)$  is the generalized degree distribution of nodes falling into interval  $i$ ,  $p_{i,r}(k_r) \approx \text{poisson}\left(N \sum_{j=1}^{M^K} p_{ij}^{(K)}(r) l_j^{(K)}\right)$ . Detailed derivation of Eq. (4) is provided in Methods section ‘Generalized degree distributions’.

Fig. 3(a) illustrates the comparison between the analytical and empirical generalized degree distribution. The comparison analysis considers 100 runs, where networks of size  $N = 5000$  nodes are generated from a set of random parameters  $\{p_{ij}(r)\}$  and  $\{l_i\}$ , with  $M = 2, K = 5$ , and  $R = 4$ . As one can notice, the WMG model can generate a wide variety of generalized degree distributions.

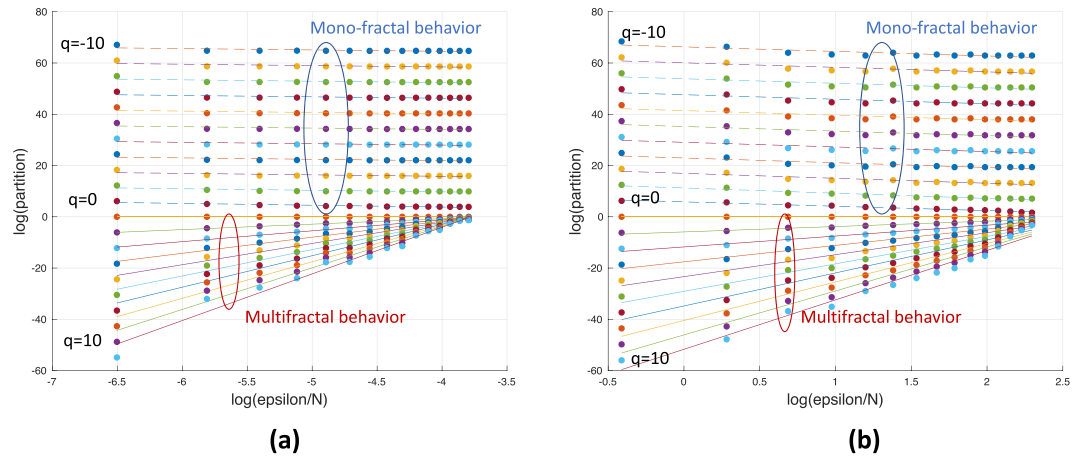
The clustering coefficients measure the local cliquishness and neighborhood connectivity<sup>5</sup>. Concerned with the density of connected triangles, it reflects the extent to which the neighbors of a certain node are also connected. In weighted graphs, the clustering coefficient of a node can be computed as the mean weights of triangles containing the node, divided by the mean weights of triplets (link pairs) originating from that node<sup>42</sup>. Following the derivation in the work of Palla *et al.*<sup>32</sup>, the average local clustering coefficient of a node falling into interval  $i$  can be calculated as

$$c_i = \frac{\sum_{j,q=1}^{M^K} \sum_{r_1,r_2,r_3=1}^R \hat{w}_1^{(K)} l_j^{(K)} l_q^{(K)} p_{ij}^{(K)}(r_1) p_{iq}^{(K)}(r_2) p_{jq}^{(K)}(r_3)}{\sum_{j,q=1}^{M^K} \sum_{r_1,r_2=1}^R \hat{w}_2^{(K)} l_j^{(K)} l_q^{(K)} p_{ij}^{(K)}(r_1) p_{iq}^{(K)}(r_2)}, \tag{5}$$

where  $\hat{w}_1 = \frac{w(r_1) + w(r_2) + w(r_3)}{3}$  and  $\hat{w}_2 = \frac{w(r_1) + w(r_2) + w(r_3)}{3}$  are the arithmetic mean of triangles and triplets edge weights. Index  $i$  marks the interval which the node falls into when generating the graph.  $i = 1 \dots M^K$ . Comparison between empirical and analytical clustering coefficients is shown in Fig. 3(b). For most of the intervals, the empirical result fits well with the analytical one. However, smaller clustering coefficient corresponds with sparser local graph structure, thus standard deviation tends to be larger than the ones with comparative larger clustering coefficients.

The proposed WMG model can reproduce the statistical properties of classic theoretical models. We test our model on a weighted graph generated by the Erdős-Rényi model. The weighted version of Erdős-Rényi model is





**Figure 4.** (a) Mono-fractal and multifractal scaling behavior in brain networks of cognitive normal person. (b) Mono-fractal and multifractal scaling behavior in chromosome-contact maps of exponentially growing genome.

given by a set of linking probabilities  $\{p(r)\}_{r=1, \dots, R}$ , where  $p(r)$  is the corresponding linking probability with the discrete weight  $w(r)$ . We apply the WMG model to a network with 500 nodes generated by a 4-layer Erdős-Rényi model. Fig. 3(c) illustrates the generalized degree distribution of the simulated network and the recovered WMG model. The blue asterisks denote the target distribution from the simulated ER graph, while the red circles represent the generalized degree distribution obtained from the identified WMG model. It shows that the proposed WMG model can capture the statistical properties of the classic random graph model.

**Multifractality in biological systems.** In two recent works<sup>18,43</sup>, the authors show that the interactions among different chromosomal regions and brain networks display a non-trivial multifractal property. We also applied the box-covering method proposed in<sup>18</sup> to show the existence of multifractality in real networks. The box covering algorithm returns the mass of each boxes with radius  $\epsilon$  to fully cover the network with  $N$  nodes. As illustrated in Fig. 4(a,b), the plots of partition functions and the scales reveal the multifractality behavior in the brain networks and the chromosome networks.

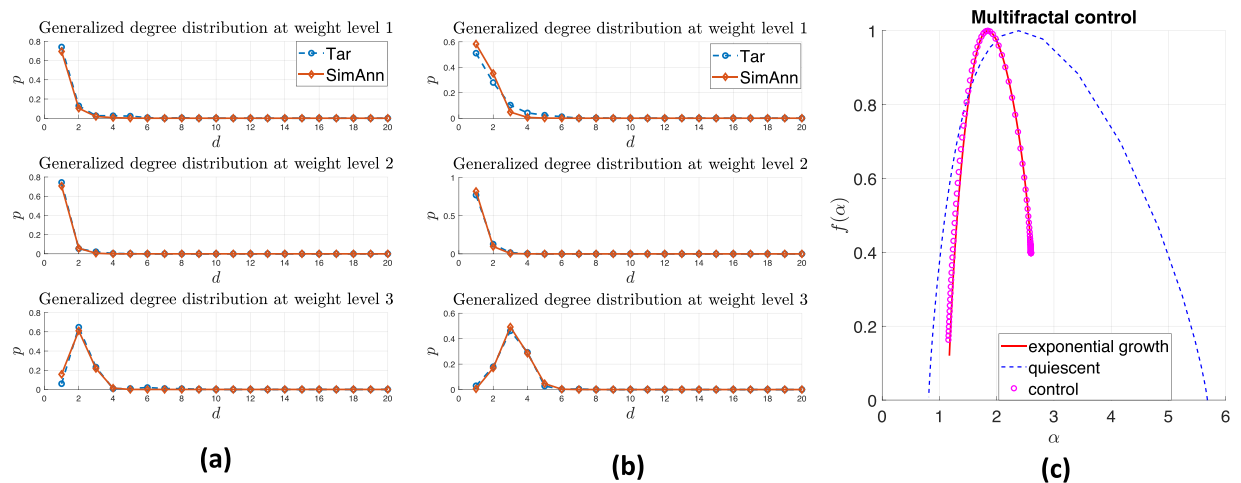
However, as discussed in section ‘Multifractal Analysis’, finding boxes to cover the graph is as hard (from a computational complexity perspective) as graph coloring problem and analytical expressions cannot be introduced. By mapping the real-world networks to multifractal generating measures that allow to construct weighted graph models, it allows us to characterize the complexity of real networks in function level and also makes it possible to better understand and control the topology structure. In this work, instead of counting boxes and calculating the multifractality in a given graph, we consider the multifractality in the recovered generating measures and its corresponding graph model.

Therefore, we exploit the proposed WMG model for investigating two biological systems, namely the chromosome interactions within the yeast genome and brain functionality networks during various stages of Alzheimer’s disease. Our goal is not only to characterize the generalized degree distributions in these real biological networks, but also to exemplify how a new network wide control strategy can be employed for ensuring specific mathematical network characteristics.

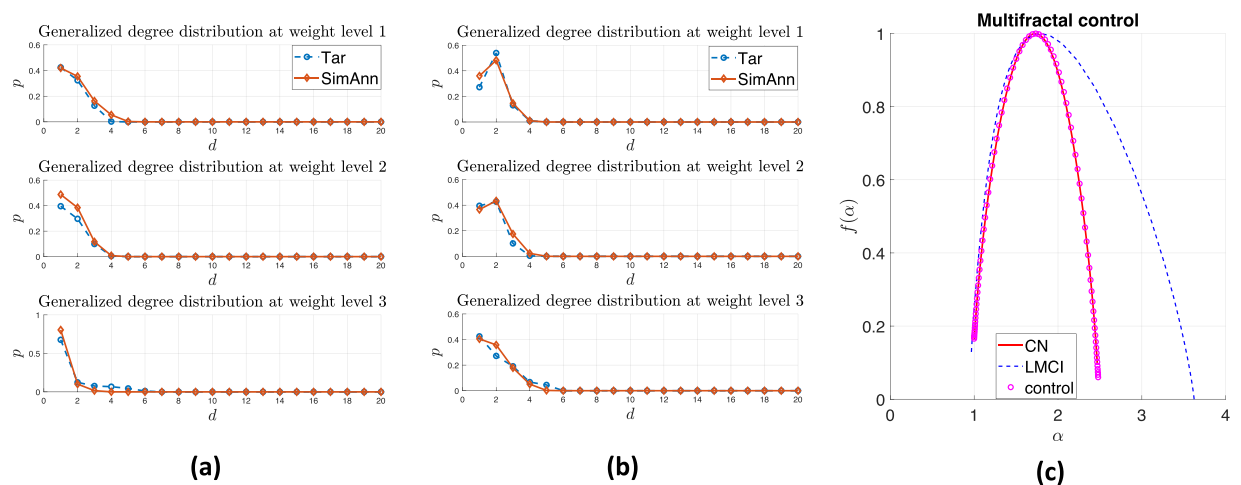
**Yeast genome.** Capturing the chromatin interactions via the Hi-C technique allows us to study the three-dimensional structure of the genome and understanding their observed behavior<sup>44–46</sup>. For instance, significant topological reorganization of yeast cell in the quiescence state has been observed<sup>47</sup>. Starting from these premises, we consider the yeast chromosome interaction data from Rutledge *et al.*<sup>47</sup>, where the Hi-C experiments are conducted on yeast cells in exponential growth on glucose-containing medium and on yeast cells in quiescence induced by glucose starvation. The data of Yeast genome are from the library of Juicebox software<sup>44,45</sup>. By interpreting the Hi-C matrix as an adjacency matrix corresponding to a complex weighted network, we estimate the WMG model from the yeast chromosome interaction data for both the exponential growth and quiescence states. We then regulate the multifractal spectrum of yeast cell in quiescence and transit it to the state of exponential growth using the proposed multifractality control strategy (see Methods section on ‘Multifractal control’).

Fig. 5(a,b) illustrate the generalized degree distribution for the yeast cells in the exponential growth state and quiescence state, respectively. The blue asterisks denote the target distribution from chromosome interaction matrix, while the red circles represent the generalized degree distribution obtained from the identified WMG model. The simulated annealing algorithm can reproduce the model with similar statistical properties (in terms of shape and trend) to the real-world data (see Methods section on ‘Reconstructing weighted multifractal model’). Eq. (16) can also be redefined with any other desired network structure properties such as strength distribution and average nearest neighbour strength, but this is left for future work.

As discussed in Methods section ‘Multifractal control’, by solving Eq. (15), we modulate quiescent yeast cell to the desired state which yeast cell is growing exponentially. In Fig. 5(c), the red line represents the desired multifractal spectrum from actively growing cell. The curve with blue dots is the multifractal spectrum of quiescent



**Figure 5.** Case study: controlling yeast cell activity condition. (a) Generalized degree distributions of yeast cell Hi-C data in exponential growth. (b) Generalized degree distributions of yeast cell Hi-C data in quiescence. (c) Controlling multifractality of yeast cell in quiescent condition to exponentially growing condition.



**Figure 6.** Case study: controlling brain structure of patients with Alzheimer's disease. (a) Generalized degree distributions of cognitively normal (CN) cohorts brain interactions. (b) Generalized degree distributions of late mild cognitive impairment (LMCI) cohorts brain interactions. (c) Controlling multifractality of LMCI and modify it to CN.

yeast cell. The magenta asterisks are the optimal modulated solution given by simulated annealing algorithm. By controlling real-world network multifractality, we could enforce or track cell evolution procedure and therefore regulate bio-feature mechanisms yeast growth.

**Alzheimer's disease.** Alzheimer's disease (AD) is a neurodegenerative disease that leads to progressive cognitive decline. While it's widely known that no cure exists for AD or terminating the neurodegeneration, early medical treatment might help to relieve the symptoms and slow the deterioration. To investigate and exemplify the benefit of proposed formalism within the context of this problem, we use data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). The rfMRI subject data are processed by BrainSuite fMRI Pipeline (BFP) and grayordinate representations are generated<sup>48,49</sup>.

The cerebral cortex is modeled as a surface mesh and globular subcortical nuclei are modeled as volume parcels<sup>50</sup>. The grayordinate data were downsampled to 445 nodes. The Pearson correlation matrix was computed using fMRI time-series<sup>51</sup>.

To control the brain interactions matrix from late mild cognitive impairment (LMCI) patients and improve their brain structure, we apply our model and modulate its multifractality to fit the one from cognitively normal (CN) people. We follow the same procedure as discussed in Hi-C case study. Fig. 6(a,b) illustrate the generalized degree distributions of CN matrix and LMCI matrix, respectively. The blue asterisks represent the target distribution from CN/LMCI matrix and the red circles are the distribution of recovered model via the simulated annealing-based reconstruction algorithm (see section 'Reconstructing weighted multifractal model' in

Methods). Simulated annealing (SA) algorithm provides a probabilistic approach for searching a large discrete space and finding approximate global optimum (e.g., parameters of a model) for a nonconvex optimization problem within a limited amount of time. Although it does not return the global optimum and it might not return identical results when running the SA algorithm multiple times, we find that one can distinguish between different networks from the retrieved WMG parameters. We individually run the simulated annealing algorithm on CN and LMCI matrix 200 times. We use 150 results of CN and 150 results of LMCI to train an SVM classifier with linear kernel, and using the remaining 50 of each class to test the dataset. The classification accuracy on the test dataset is 94%. It shows that the WMG model characterizes the network patterns and it could also discriminate between different traits of these complex networks.

Fig. 6(c) shows the multifractal spectrum of LMCI matrix before and after the control. The red curve is the desired multifractal spectrum from CN data. The blue dots and magenta asterisks represent the multifractal spectrum curves of LMCI and optimal regulated solution obtained by adapting the LMCI network structure. The controlled multifractal spectrum also fits well with the target curve.

## Discussion

We propose the *weighted multifractal graph* (WMG) model to capture the nodes attributes and interactions in weighted complex networks. We show that it fits a wide variety of multifractal spectrum and statistical properties of real-world networks including generalized degree distribution, clustering coefficients and joint degree distribution. Beyond the self-similar constructing procedure of the recursive generating model, we present the multifractal analysis of the random graph model and show that the WMG model could generate a variety of multifractal spectrum curves with different shapes and varying degree of asymmetry. More importantly, the introduction of weights in the WMG model brings flexibility in multifractal spectrum. The proposed model has potential applications in many disciplines including biological systems, geoscience, financial markets and social networks. In order to ground this extended model into real-world applications, one needs to develop rigorous, efficient and accurate model identification techniques that need to take into account the relationship between the complexity of real networks and the generating measures, assess the impact of model parameters on network properties (e.g., information flow performance, robustness), all of which are not trivial mathematical tasks.

Apart from modelling real networks, an important implication of the WMG model is to minimally control the multifractality of complex networks. The experiments on real-world datasets of yeast genome and Alzheimer's disease reveal that it is possible to regulate the multifractal spectrum of complex networks with minimal adjustments in the WMG model parameters. The proposed model is a step towards exploring the underlying growing mechanism of evolving networks such as brain structure of Alzheimer's disease patients, or actively growing yeast cells. Moreover, the control of multifractality can provide a novel way to treat brain diseases, control growing states or recover from potential attacks in biological and social systems.

Future extensions of the WMG model can consider the generalization of  $l_r$ . In the proposed WMG model, we consider that the values of the interval length  $l_r$  are identical for all  $R$  layers. However, one can consider that  $l_r$  depends on  $r$ , but this will introduce more complexity and model identification strategies. Moreover, each node can have different category  $i_r$  in different weight layer  $r$  and the linking probability would become  $p_{i_r}^{(K)}(r)$ . Another important addition would be to have temporal dimension into the proposed model. This corresponds to modelling and controlling the time varying complex networks by regulating the multifractality of graphs. It will help to answer research questions such as how does network structure evolve over time from one state to another. For example, in decoding human behavior, what are the causes of changing genome or brain structure and what does it lead to. It is also crucial to understand the physical meaning of network multifractality and deciphering the hidden information of real-world complex networks with self-similar patterns.

## Methods

**Multifractal analysis.** We study the multifractal analysis of the WMG model based on the partition function. Let  $\{p_{ij}(r)\}_{i,j=1:M}$  denote links generating probabilities in the initial unit square model. For simplicity, we reshape it as  $\{p_i(r)\}_{i=1:M^2}$ .  $\{l_i l_j\}_{i,j=1:M}$  is the area of each sub-rectangles in the unit square. Similarly, we reshape it as  $\{a_i\}_{i=1:M^2}$ . The partition function at an average sub-block size  $\epsilon = \frac{1}{R} \left(\frac{1}{M}\right)^{2K}$  can be defined as

$$Z_\epsilon(q) = \sum_{r, \{k_j\}_{j=1:M^2}} \binom{K}{k_1 \dots k_j \dots k_{M^2}} \left\{ \prod_{i=1}^{M^2} [a_i p_i(r)]^{k_i} \right\}^q. \quad (6)$$

Here the first term  $\binom{K}{k_1 \dots k_j \dots k_{M^2}}$  is number of sub-blocks which have the same area proportion  $\prod_{i=1}^{M^2} a_i^{k_i}$  and link generating probability  $\prod_{i=1}^{M^2} p_i(r)^{k_i}$ . The second term  $\prod_{i=1}^{M^2} [a_i p_i(r)]^{k_i}$  is the proportion of edges with weight  $w_r$  which are generated by those sub-blocks.  $\{k_j\}_{j=1:M^2}$  is subjected to  $\sum_{j=1}^{M^2} k_j = K$ . We continue to simplify the summation and multinomial coefficients.



$$\begin{aligned}
 Z_\epsilon(q) &\stackrel{A}{=} \sum_{r, \{k_j\}_{j=3:M^2}} \prod_{i=3}^{M^2} [a_i p_i(r)]^{q k_i} \binom{K}{k_3 \dots k_{M^2}} \sum_{k_2} \binom{K - k_3 \dots - k_{M^2}}{k_2} [a_1 p_1(r)]^{q k_1} \\
 &\quad \cdot [a_2 p_2(r)]^{q k_2} \\
 &\stackrel{B}{=} \sum_{r, \{k_j\}_{j=3:M^2}} \prod_{i=3}^{M^2} [a_i p_i(r)]^{q k_i} \binom{K}{k_3 \dots k_{M^2}} \{ [a_1 p_1(r)]^q + [a_2 p_2(r)]^q \}^{K - k_3 \dots - k_{M^2}} \\
 &\stackrel{C}{=} \sum_{r=1}^R \left\{ \sum_{i=1}^{M^2} a_i p_i(r)^q \right\}^K = \sum_{r=1}^R \left\{ \sum_{i,j=1}^M (l_i l_j p_{ij}(r))^q \right\}^K.
 \end{aligned} \tag{7}$$

In equation (7), the partition function is further simplified with binomial theorem and multinomial coefficients properties. Continuing with Eq. (6), the first equal label  $\stackrel{A}{=}$  is given by expanding multinomial coefficient  $\binom{K}{k_1 \dots k_j \dots k_{M^2}}$ .  $B$  uses the constraint of  $\sum_{i=1}^{M^2} k_j = K$ . Further simplifications are continued on  $k_3, k_4 \dots k_{M^2}$  with mathematical induction (C).

Other functions and measures including the mass exponent  $\tau(q) = \frac{\log Z_\epsilon(q)}{\log \epsilon}$ , the generalized fractal dimension  $D(q) = \frac{\tau(q)}{1-q}$ , the Lipschitz-Hölder exponent (also refer to coarse Hölder exponent or singularity index in some materials)  $\alpha(q) = \frac{d\tau(q)}{dq}$  and the multifractal spectrum  $f(\alpha) = \alpha(q)q - \tau(q)$  can be derived from the partition function in equation (7). Here we give the expression of the Lipschitz-Hölder exponent.

$$\alpha(q) = \frac{1}{Z_\epsilon(q) \log \epsilon} \sum_{r=1}^R \left\{ K \left[ \sum_{i,j=1}^M (l_i l_j p_{ij}(r))^q \right]^{K-1} \sum_{i,j=1}^M (l_i l_j p_{ij}(r))^q \ln [l_i l_j p_{ij}(r)] \right\}. \tag{8}$$

**Generalized degree distribution.** We introduce the generalized degree distribution to specify the network structure with more detailed information compared to the strength distribution. Generalized degree  $D_r$  is defined as the number of edges or connections a node has with weight  $w_r$ . When generating the nodes, we classify each node into  $M^K$  categories with probability  $l_i^{(K)}$ ,  $i = 1 \dots M^K$ . Given by the recursive model, the link generating probability of a node in category  $i$  and a node in category  $j$  is  $p_{ij}^{(K)}(r)$  and the edges are generated independently. Thus, the number of edges connecting a node in category  $i$  and all nodes in category  $j$  follows binomial distribution. As discussed in the publication of Palla *et al.*<sup>32</sup>, when the network scale  $N$  is large, it can be further approximated by Poisson distribution with rate  $\lambda_{i,j}(r) = p_{ij}^{(K)}(r) N l_j^{(K)}$ . Followed by summation property of Poisson distributed independent random variables, the generalized degree distribution of a node falling into interval  $i$  is given by

$$p_{i,r}(k_r) \approx \text{Poisson} \left( \sum_{j=1}^{M^K} \lambda_{i,j}(r), k_r \right) = \text{Poisson} \left( N \sum_{j=1}^{M^K} p_{ij}^{(K)}(r) l_j^{(K)}, k_r \right). \tag{9}$$

Therefore, the average generalized degree distribution is expressed as

$$p_r(k_r) = \sum_{i=1}^{M^K} p_{i,r}(k_r) l_i^{(K)}. \tag{10}$$

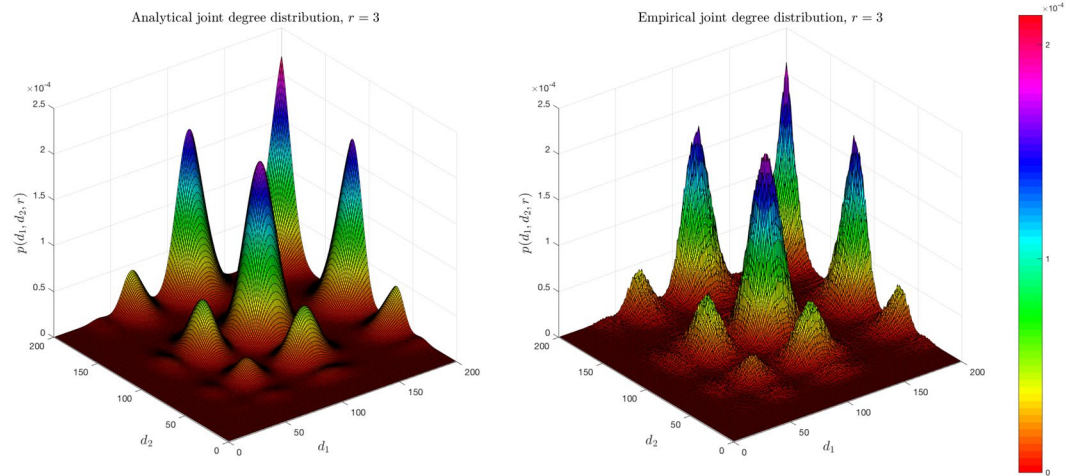
As shown in the main text, the analytic expression in (10) coordinates well with empirical simulation results when  $N \gg M^K$ .

**Clustering coefficients.** Clustering coefficient is a measure of graph cliquishness and neighborhood connectivity<sup>5</sup>. Following the definition of clustering coefficients in weighted graphs<sup>42</sup> and the multifractal network generator<sup>32</sup>, the clustering coefficient of a node in category  $i$  is written as

$$c_i = \frac{\sum_{j,q=1}^{M^K} \sum_{r_1, r_2, r_3=1}^R \hat{w}_1 l_j^{(K)} l_q^{(K)} p_{ij}^{(K)}(r_1) p_{iq}^{(K)}(r_2) p_{jq}^{(K)}(r_3)}{\sum_{j,q=1}^{M^K} \sum_{r_1, r_2=1}^R \hat{w}_2 l_j^{(K)} l_q^{(K)} p_{ij}^{(K)}(r_1) p_{iq}^{(K)}(r_2)}, \tag{11}$$

where  $\hat{w}_1 = \frac{w(r_1) + w(r_2) + w(r_3)}{3}$  and  $\hat{w}_2 = \frac{w(r_1) + w(r_2) + w(r_3)}{3}$  are the arithmetic mean of triangles and triplets edge weights. The numerator is the mean weights of triangles containing a node falling into sub-interval  $i$ ,  $i = 1 \dots M^K$ , the denominator is the mean weights of triplets (link pairs) originating from that node.

**Joint degree distribution.** The joint degree distribution of graphs with discrete weight set  $w(r)$  is defined as  $p(d_1, d_2, r) = \frac{m(d_1, d_2, r)}{M(r)}$ , where  $M(r)$  is the total number of edges with weight  $w(r)$ ,  $m(d_1, d_2, r)$  is the number of



**Figure 7.** Analytical and empirical joint degree distribution. This simulation is generated under parameters  $M = 2, K = 3, R = 4, N = 5000$ . Graph generation is repeated 10 times.

edges connecting nodes with the generalized degree  $d_1$  and  $d_2$ , and the weight is  $w(r)$ . The number of nodes which have  $d$  edges with weight  $w(r)$  is:

$$N_i(d, r) = Nl_i \times \text{Poisson} \left( \sum_{j=1}^{M^K} p_{ij}^{(K)}(r) Nl_j^{(K)}, d \right). \tag{12}$$

Therefore, we could obtain the expression of the joint degree distribution as

$$p(d_1, d_2, r) = \frac{\sum_{i,j=1}^{M^K} N_i(d_1, r) N_j(d_2, r) p_{ij}^{(K)}(r)}{\sum_{i,j=1}^{M^K} Nl_i^{(K)} Nl_j^{(K)} p_{ij}^{(K)}(r)}. \tag{13}$$

The comparison between the analytical and empirical joint degree distribution is shown in Fig. 7.

**Edge distribution.** In the WMG model, we focus on weighted complex networks. Therefore, we also consider the edge distribution in the proposed model. Let  $X_{ij}(r)$  denote the number of edges connecting nodes in category  $i$  and nodes in category  $j$  with weight  $w(r)$ . The vector of edge numbers  $\mathbf{X}_{ij}$  follows the multinomial distribution  $Multi(p_{ij}^{(K)}(0), p_{ij}^{(K)}(1), \dots, p_{ij}^{(K)}(R))$ . The total number of edges with weight  $w(r)$  is the sum of the independent random variables  $X(r) = \sum_{ij}^{M^K} X_{ij}(r)$ . Its moment generating function can be written as

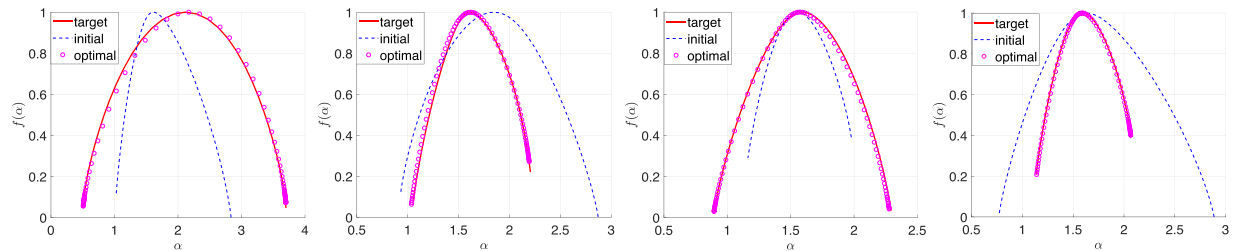
$$MGF_{\mathbf{X}}(\mathbf{t}) = \prod_{i,j=1}^{M^K} MGF_{X_{ij}}(\mathbf{t}) = \prod_{i,j=1}^{M^K} \left[ \sum_{r=0}^R p_{ij}^{(K)}(r) e^{t_r} \right]^{N^2 l_i^{(K)} l_j^{(K)}}. \tag{14}$$

**Controlling multifractality.** The proposed WMG model can quantify the emergence and self-organization<sup>52</sup> and complexity of network in terms of number of rules that are at play when growing a network. Introducing the WMG model makes it possible to minimally control complex network. As discussed in previous section, by computing its partition function, multifractal spectrum and Lipschitz-Hölder exponent, we could analytically obtain network multifractalities. While aiming to transit or change a multifractal network to some certain networks with desired multifractality, we also expect minimal adjustments comparing with original network.

This control problem can be written as the following optimization

$$\left\{ p_{i,j}(r) \right\}_{i,j,r}, \left\{ l_i \right\}_i = \underset{p, l}{\operatorname{argmin}} \left( \|f(p, l) - f_{Tar}\|_2 + \|\alpha(p, l) - \alpha_{Tar}\|_2 + \lambda (\|p - p_0\|_2 + \|l - l_0\|_2) \right), \tag{15}$$

where  $f_{Tar}, \alpha_{Tar}$  are the target multifractal spectrum and the Lipschitz-Hölder exponent,  $f(p, l)$  and  $\alpha(p, l)$  are given by proposed multifractal network model,  $\left\{ p_{i,j}(r) \right\}_{i,j,r}, \left\{ l_i \right\}_i$  are the optimal parameters solutions, and  $p_0, l_0$  are the starting point of model parameters. To focus on the influence of minimal changes in model parameters, here we fix  $M, R$  and  $K$  to some certain integers. To minimize the cost function defined in equation (15), here we use simulated annealing algorithm<sup>53</sup> to recover optimal model parameters  $\left\{ p_{i,j}(r) \right\}_{i,j,r}$  and  $\left\{ l_i \right\}_i$ .



**Figure 8.** Multifractal control simulations with four different perturbations of target multifractal spectrum.

Four multifractal control problem simulations are shown in Fig. 8. The red lines are the desired multifractal spectrums, blue dot lines are the starting multifractal spectrums given by random parameters. Magenta asterisks are the optimal solutions given by simulated annealing algorithm. Different deviations of shifting and scaling are shown here. In each case, the multifractal spectrums are almost identical to the desired ones. It shows that by minimizing the cost function in equation (15), we could control network multifractality with minimal changes.

**Reconstructing weighted multifractal model.** We use simulated annealing algorithm<sup>53</sup> to retrieve the optimal parameters  $\{p_{ij}(r)\}$  and  $\{l_i\}$  in the real networks. We define the distance of statistical measures in the WMG model and the real networks as energy function,

$$E = \sum_{r=1}^R \sum_{d=1}^N \left\| p(d, r) - p^*(d, r) \right\|_2 + \lambda \left\| \langle c_i \rangle - c^* \right\|_2, \quad (16)$$

where  $p^*(d, r)$  and  $p(d, r)$  are the generalized degree distributions generated from the real-world network and the proposed WMG model by equation (4),  $\langle c_i \rangle$  is the average clustering coefficient in equation (5),  $c^*$  is the empirical clustering coefficient. Equation (16) could also be redefined by other desired statistical properties of the networks. Minimizing Eq. (16) is a non-convex problem. And as the weight resolution in the WMG model increases, more parameters need to be estimated. Other optimization methods may also be applied.

Received: 4 November 2019; Accepted: 12 March 2020;

Published online: 26 March 2020

## References

- Newman, M. E. The structure and function of complex networks. *SIAM review* **45**, 167–256 (2003).
- Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. of modern physics* **74**, 47 (2002).
- Erds, P. & Rényi, A. On random graphs. *Publ. Math. Debrecen* **6**, 290–297 (1959).
- Erdős, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* **5**, 17–60 (1960).
- Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *nature* **393**, 440 (1998).
- Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *science* **286**, 509–512 (1999).
- Xue, Y. & Bogdan, P. Reconstructing missing complex networks against adversarial interventions. *Nat. communications* **10**, 1738 (2019).
- Sia, J., Jonckheere, E. & Bogdan, P. Ollivier-ricci curvature-based method to community detection in complex networks. *Sci. reports* **9**, 9800 (2019).
- Boccaletti, S. *et al.* The structure and dynamics of multilayer networks *Physics Reports* **544**, 1–122 (2014).
- Benson, A. R., Abebe, R., Schaub, M. T., Jadbabaie, A. & Kleinberg, J. Simplicial closure and higher-order link prediction. *Proc. Natl. Acad. Sci.* **115**, E11221–E11230 (2018).
- Bian, T. & Deng, Y. Identifying influential nodes in complex networks: A node information dimension approach *Chaos: An Interdisciplinary J. Nonlinear Sci.* **28**, 043109 (2018).
- Wen, T. & Deng, Y. Identification of influencers in complex networks by local information dimensionality. *Inf. Sci.* **512**, 549–562 (2020).
- Song, C., Havlin, S. & Makse, H. A. Self-similarity of complex networks. *Nature* **433**, 392 (2005).
- Song, C., Havlin, S. & Makse, H. A. Origins of fractality in the growth of complex networks. *Nat. physics* **2**, 275 (2006).
- Bunde, A. & Havlin, S. Percolation i. In *Fractals and disordered systems*, 59–114 (Springer, 1996).
- Gallos, L. K., Song, C. & Makse, H. A. A review of fractality and self-similarity in complex networks. *Phys. A: Stat. Mech. its Appl.* **386**, 686–691 (2007).
- Wen, T., Duan, S. & Jiang, W. Node similarity measuring in complex networks with relative entropy *Commun. Nonlinear Sci. Numer. Simul.* 104867 (2019).
- Xue, Y. & Bogdan, P. Reliable multi-fractal characterization of weighted complex networks: algorithms and implications *Sci. reports* **7**, 7487 (2017).
- Agterberg, F. P. Mixtures of multiplicative cascade models in geochemistry. *Nonlinear Process. Geophys.* **14**, 201–209, <https://doi.org/10.5194/npg-14-201-2007>. (2007).
- Song, C., Gallos, L. K., Havlin, S. & Makse, H. A. How to calculate the fractal dimension of a complex network: the box covering algorithm *J. Stat. Mech. Theory Exp.* **2007**, P03006 (2007).
- De Wijs, H. Statistics of ore distribution. part i: frequency distribution of assay values. *J. Royal Neth. Geol. Min. Soc.* **13**, 365–375 (1951).
- Mandelbrot, B. B. *The fractal geometry of nature*, vol. 173 (WH freeman New York, 1983).
- Mandelbrot, B. B. Multifractal measures, especially for the geophysicist. In *Fractals in geophysics*, 5–42 (Springer, 1989).
- Ford, A. & Blenkinsop, T. G. An expanded de wijs model for multifractal analysis of mineral production data. *Mineralium Deposita* **44**, 233 (2009).
- Agterberg, F. New applications of the model of de wijs in regional geochemistry. *Mathematical Geology* **39**, 1 (2007).

26. Marshak, A., Davis, A., Cahalan, R. & Wiscombe, W. Bounded cascade models as nonstationary multifractals. *Phys. Rev. E* **49**, 55 (1994).
27. Cheng, Q. Generalized binomial multiplicative cascade processes and asymmetrical multifractal distributions. *Nonlinear Process. Geophys.* **21**, 477–487 (2014).
28. Cheng, Q. Multifractal distribution of eigenvalues and eigenvectors from 2d multiplicative cascade multifractal fields. *Math. Geol.* **37**, 915–927 (2005).
29. Leskovec, J. & Faloutsos, C. Scalable modeling of real graphs using kronecker multiplication. In *Proceedings of the 24th international conference on Machine learning*, 497–504 (ACM, 2007).
30. Mahdian, M. & Xu, Y. Stochastic kronecker graphs. In *International workshop on algorithms and models for the web-graph*, 179–186 (Springer, 2007).
31. Leskovec, J., Chakrabarti, D., Kleinberg, J., Faloutsos, C. & Ghahramani, Z. Kronecker graphs: An approach to modeling networks. *J. Mach. Learn. Res.* **11**, 985–1042 (2010).
32. Palla, G., Lovász, L. & Vicsek, T. Multifractal network generator. *Proc. Natl. Acad. Sci.* **107**, 7640–7645 (2010).
33. Benson, A. R., Riquelme, C. & Schmit, S. Learning multifractal structure in large networks. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1326–1335 (ACM, 2014).
34. Kim, M. & Leskovec, J. Modeling social networks with node attributes using the multiplicative attribute graph model *arXiv preprint arXiv:1106.5053* (2011).
35. Kim, M. & Leskovec, J. Multiplicative attribute graph model of real-world networks. *Internet mathematics* **8**, 113–160 (2012).
36. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics *Physics reports*, **424**, 175–308 (2006).
37. Wei, D.-J. *et al.* Box-covering algorithm for fractal dimension of weighted networks *Sci. reports* **3**, 3049 (2013).
38. Wei, B. & Deng, Y. A cluster-growing dimension of complex networks: From the view of node closeness centrality. *Phys. A: Stat. Mech. its Appl.* **522**, 80–87 (2019).
39. Frieze, A. M. On the independence number of random graphs. *Discret. Math.* **81**, 171–175 (1990).
40. Bollobás, B. & Erdős, P. Cliques in random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 80, 419–427 (Cambridge University Press, 1976).
41. Mulet, R., Pagnani, A., Weigt, M. & Zecchina, R. Coloring random graphs. *Phys. review letters* **89**, 268701 (2002).
42. Opsahl, T. & Panzarasa, P. Clustering in weighted networks. *Soc. networks* **31**, 155–163 (2009).
43. Pigolotti, S., Jensen, M. H. & Tiana, G. Hierarchical domain model explains multifractal scaling of chromosome contact maps *arXiv preprint arXiv:1906.12090* (2019).
44. Durand, N. C. *et al.* Juicebox provides a visualization system for hi-c contact maps with unlimited zoom. *Cell systems* **3**, 99–101 (2016).
45. Robinson, J. T. *et al.* Juicebox. js provides a cloud-based visualization system for hi-c data. *Cell systems* **6**, 256–258 (2018).
46. Van Berkum, N. L. *et al.* Hi-c: a method to study the three-dimensional architecture of genomes *JoVE (Journal of Visualized Experiments)* e1869 (2010).
47. Rutledge, M. T., Russo, M., Belton, J.-M., Dekker, J. & Broach, J. R. The yeast genome undergoes significant topological reorganization in quiescence. *Nucleic acids research* **43**, 8299–8313 (2015).
48. Smith, S. M. *et al.* Functional connectomics from resting-state fmri. *Trends in cognitive sciences* **17**, 666–682 (2013).
49. Barch, D. M. *et al.* Function in the human connectome: task-fmri and individual differences in behavior. *Neuroimage* **80**, 169–189 (2013).
50. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–124 (2013).
51. Joshi, A. A., Chong, M., Li, J., Choi, S. & Leahy, R. M. Are you thinking what iam thinking? synchronization of resting fmri time-series across subjects. *NeuroImage* **172**, 740–752 (2018).
52. Balaban, V., Lim, S., Gupta, G., Boedicker, J. & Bogdan, P. Quantifying emergence and self-organisation of enterobacter cloacae microbial communities *Sci. reports* **8**, 12416 (2018).
53. Van Laarhoven, P. J. & Aarts, E. H. Simulated annealing. In *Simulated annealing: Theory and applications*, 7–15 (Springer, 1987).

## Acknowledgements

The authors gratefully acknowledge the support by the National Science Foundation Career award under Grant No. CPS/CNS-1453860, CCF-1837131, MCB-1936775, CNS-1932620, the U.S. Army Research Office (ARO) under Grant No. W911NF-17-1-0076, the Okawa Foundation award, and the Defense Advanced Research Projects Agency (DARPA) Young Faculty Award and Director's Fellowship award under Grant No. N66001-17-1-4044 support. P.B. thanks to S. Stanton for insightful discussions on concepts related to multifractals and complex networks that contributed to and inspired this problem definition of controlling complex networks. The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied by the Defense Advanced Research Projects Agency, the Air Force Research Lab, the Department of Defense or the National Science Foundation.

## Author contributions

R.Y. and P.B. contributed to the design of the research. R.Y. contributed to the simulations and experiments. R.Y. and P.B. analyzed the results and wrote the manuscript.

## Competing Interests

The authors declare that they have no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to P.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020