

# Supporting Information for “Scaling theory of transport in complex networks”

L.K. Gallos, C. Song, S. Havlin, and H.A. Makse

February 21, 2007

## 1 Protein-protein interaction database

We analyzed the filtered yeast interactome developed by Han *et al.* [1] which consists of the intersection of three different databases of proteins interactions: the 5249 interactions found in high-throughput yeast-two-hybrid experiments, 6630 interactions appearing in proteome-wide pull-down experiments followed by mass spectrometry analysis, and the 7446 predicted interactions from in silico computational searches by gene neighborhood, gene co-occurrence or gene fusion methods. The final filtered yeast interactome dataset is obtained by including the physical interactions in the above databases and those of the MIPS database [2]. This is done in an attempt to remove the large number of false positives in high-throughput yeast two hybrid methods. However, the purity of such a database may still be affected by others factors such as variation in the species number of the constituents and other false positives.

## 2 Calculation of the exponents through renormalization

The plots in Figs. 1(d) and 1(e) in the main text are constructed as follows: First we tile the original network  $G$  with boxes (see for example Fig. 1(a)). We assume that each box can be represented by its largest degree node, so we choose a local hub from each box. We then calculate the diffusion time  $T$  and the resistance  $R$  in  $G$  between two hub-nodes A and B that belong

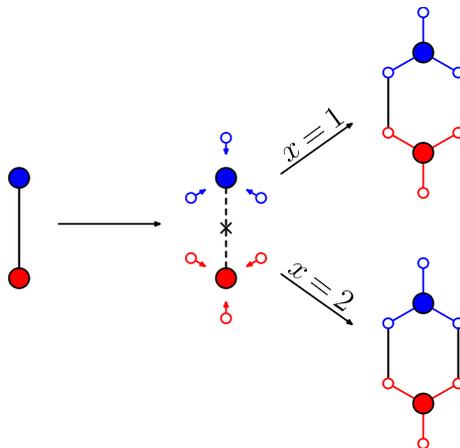


Figure 1: Generalization of the fractal network generation model. The model grows from a small network, usually two nodes connected to each other. During each step and for every link in the system, a node (which will be the center of a module) produces  $m$  offspring nodes (in this drawing  $m = 3$ ). The original link is then removed and we connect these two modules via  $x$  new links between randomly selected nodes of the new generation. Notice that the case of  $x = 1$  results in a tree structure, while loops appear for  $x = 2$ .

to different boxes. We renormalize the structure into a new network  $G'$  (all primed quantities refer to the renormalized system) where a box in  $G$  is now transformed into a node in  $G'$ . We then calculate the diffusion time  $T'$  and the resistance  $R'$  between the new nodes  $A'$  and  $B'$  in the renormalized network. We repeat the same procedure for all pairs of hubs in the network and for each pair we plot the values of  $T$  (or  $R$ ) vs the corresponding values of  $T'$  (or  $R'$ ). The distribution of these points on the  $T' - T$  ( $R' - R$ ) plane are shown in Figs. 1(b) and 1(c). We observe that practically all points fall on the same straight line. This means that the ratio  $T'/T$  is almost constant for all pairs in the network (and can be calculated from the slope of this curve). The value of this ratio represents a point in Fig. 1(d) or 1(e) as a function of  $N'/N$ . We also find that this is true between any two successive renormalization stages, so we compute more points for different ratios  $N'/N$  by applying the renormalization scheme many times on the same network.

### 3 Fractal network generation model

A general method for generating a fractal network with self-similar properties has been presented in Ref. [3]. This model leads to a tree (loopless) structure. Here, we briefly review this model and extend it to include loops preserving at the same time the property of fractality.

According to the model, we start with a formation of a small number of nodes. For example, we could use 2 or 3 nodes linked to each other. The model evolves by assigning  $m$  new offsprings to each node of the previous generations and at the same time an existing link is substituted by a link between two newly added nodes. It has been shown [3] that the main feature of fractal networks is the ‘repulsion’ between hubs, i.e. a small probability that any two hubs (nodes created in the early generation stages) in the network are directly linked to each other.

In the original version of this model the resulting structure is a tree. It is possible to extend this model by introducing more links during the growth process, which results in a structure with a predetermined number of loops. So, instead of adding  $x = 1$  new link between boxes at each growth step, we now add  $x$  random links between these boxes. For the current study, these  $x$  links always connect a non-hub to another non-hub node, so that fractality is preserved. The model is presented in Fig. 1.

### 4 Calculation of $d_M$ for the model

The diameter in the network model, as described above, increases multiplicatively as  $L(t+1) = 3L(t)$  (see also [3] and Supplementary Material therein). The number of links within a module grows with  $2m + x$  (each node on the side of one link gives rise to  $m$  new links and  $x$  extra links connect the new nodes), while the number of links pointing out of a module is by definition proportional to  $x$ . This means that the modularity  $M(\ell_B)$  of a network, defined as the average ratio of links within the module over links towards other modules, is proportional to  $(2m + x)/x$ . The relation

$$M(\ell_B) \sim \ell_B^{d_M} \tag{1}$$

can thus be used to calculate  $d_M$  for the model:

$$\frac{2m + x}{x} \sim 3^{d_M}, \tag{2}$$

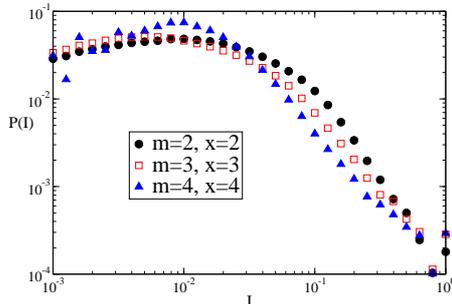


Figure 2: Probability distribution  $P(I)$  for fractal models with the same  $d_M = 1$  value, but varying  $m$  and  $x$  values (shown in the plot).

which finally yields

$$d_M = \frac{\ln\left(2\frac{m}{x} + 1\right)}{\ln 3}. \quad (3)$$

#### 4.1 Current distribution for fixed $d_M$

In the main text of the paper we describe how the exponents of resistance and diffusion in the fractal model network depend on static exponents, through Eqs. (13) and (16), reproduced here for convenience:

$$\zeta = 1 - dx, \quad d_w = 1 + d_M, \quad (4)$$

where we have used the notation  $dx \equiv \ln x / \ln 3$ . The above equations imply that diffusion in the model depends only on the degree of modularity, while the resistance is influenced by the number of loops. We have already displayed in Fig. 3 of the main text the  $d_w$  dependence on  $d_M$  while in the inset of Fig. 4 we have seen how the parameter  $x$  influences the current distribution. Similarly, here we provide direct numerical evidence that the influence of modularity is not as important as that of loops. In Fig. 2 we present the current distribution  $P(I)$  for three network models, where  $d_M = 1$  is fixed but the value of  $x$  varies from  $x = 2$  to  $x = 4$ . As we increase  $x$  there are continuously less links that carry large currents and the distribution shifts towards smaller  $I$  values. Although this may be expected (including more loops leads to a more uniform distribution of the current in the network with more links tending to carry similar current to each other), the important message is that modularity alone cannot control resistance.

Combining the above equations with the Einstein relation we finally get the modified Einstein form

$$d_B = d_M + d_x, \quad (5)$$

where all the exponents are now static.

## 5 Scaling of $R(\ell; k_1, k_2)$ and $T(\ell; k_1, k_2)$

Here we use scaling arguments and the renormalization property of the self-similar networks in order to predict the dependence of the resistance  $R$  and the diffusion time  $T$  on the distance  $\ell$  between two nodes and their corresponding degrees  $k_1$  and  $k_2$ . In earlier work it has been shown that under a transformation of the length  $\ell_B^{-1}\ell \rightarrow \ell'$  the degree of a node becomes  $\ell_B^{-d_k}k \rightarrow k'$ , where  $d_k$  is the exponent describing the scaling of the degree. According to the result of Fig. 1 of the main text and Eqs. (8) and (9) it follows that,

$$R'(\ell'; k'_1, k'_2) = \ell_B^{-\zeta} R(\ell; k_1, k_2). \quad (6)$$

Using the invariance property of the network under renormalization, where we assume that the form of the resistance also remains invariant, we can substitute the renormalized quantities and get:

$$R(\ell_B^{-1}\ell; \ell_B^{-d_k}k_1, \ell_B^{-d_k}k_2) = \ell_B^{-\zeta} R(\ell; k_1, k_2). \quad (7)$$

The above equation holds for all values of  $\ell_B$ , so we can select this quantity to be  $\ell_B = k_2^{1/d_k}$ . This constraint allows us to reduce the number of variables in the equation, with the final result:

$$R\left(\frac{\ell}{k_2^{1/d_k}}; \frac{k_1}{k_2}, 1\right) = k_2^{-\zeta/d_k} R(\ell; k_1, k_2). \quad (8)$$

This equation suggests a scaling for the resistance  $R$ :

$$R(\ell; k_1, k_2) = k_2^{\zeta/d_k} f_R\left(\frac{\ell}{k_2^{1/d_k}}, \frac{k_1}{k_2}\right), \quad (9)$$

where  $f_R()$  is an undetermined function. All the above arguments can be repeated for the diffusion time, with a similar expression:

$$T(\ell; k_1, k_2) = k_2^{d_w/d_k} f_T\left(\frac{\ell}{k_2^{1/d_k}}, \frac{k_1}{k_2}\right), \quad (10)$$

where the form of the right-hand function may be different. The scaling in the above equations is supported by the numerical data collapse in Fig. 2 of the main text.

Notice that the distance  $\ell$  measures the distance between any two nodes, while the distance  $\ell_B$  measures the maximum distance in a box that is used to renormalize the network. This distinction is important because at the scaling of  $R$  with  $\ell_B$  there is no explicit dependence on the scale transformation of  $k$ , since the degree is transformed implicitly when we renormalize the network with the boxes. Thus, there is no dependence on  $k$  in Eqs. (5) or (7) of the main text. However, in Eq. (6) above (Eq. 8 in the main text) we are looking at the transformation of any length, and not just the boxes, therefore the degrees  $k$  are not automatically scaled and we have to add the proper scaling transformation and dependence on  $k_1$  and  $k_2$ . Thus, we finally arrive at a more general form of Eq. (6) that is valid for any length  $\ell$  and not just the box size  $\ell_B$ .

## References

- [1] Han, J.-D.J., et al (2004) *Nature* **430** 88.
- [2] Mewes H.W., et al (2002) *Nucleic Acids Res.* **30** 31.
- [3] Song, C., Havlin, S., & Makse, H.A., (2006) *Nature Physics* **2**, 275.