Clustering of random scale-free networks

Pol Colomer-de-Simón* and Marián Boguñá
UB, Universitat de Barcelona, Departament de Física fonamental
Martí i Franquès 1 08028 Barcelona, Spain

We derive the finite size dependence of the clustering coefficient of scale-free random graphs generated by the configuration model with degree distribution exponent $2 < \gamma < 3$. Degree heterogeneity increases the presence of triangles in the network up to levels that compare to those found in many real networks even for extremely large nets. We also find that for values of $\gamma \approx 2$, clustering is virtually size independent and, at the same time, becomes a *de facto* non self-averaging topological property. This implies that a single instance network is not representative of the ensemble even for very large network sizes.

In the absence of high degree nodes the Clustering coefficient of the resulting network is given by [1]:

$$C \sim \frac{(\langle k(k-1)\rangle)^2}{\langle k\rangle^3} \tag{1}$$

That vanishes very fast for large system sizes. However in the case of a scale-free networks it predicts a behaviour

$$C \sim N^{\frac{7-3\gamma}{\gamma-1}} \tag{2}$$

that diverges for $\gamma < 7/3$. This is because this derivation does not account for the structural correlations among degrees of connected nodes that appear in order to be able to close the network for degree distributions with $\gamma < 3$ [2].
Here we explain how to derive the correct scaling behaviour of clustering for scale-free random graphs with $2 < \gamma < 3$. We also show that when clustering is very high and becomes nearly size independent.
Using a Canoncial version of the configuration model (hidden variables) we were able to derive that the Clustering coefficient depends with the size of the system as follows:

$$C \sim N^{2-\gamma} \ln N \tag{3}$$

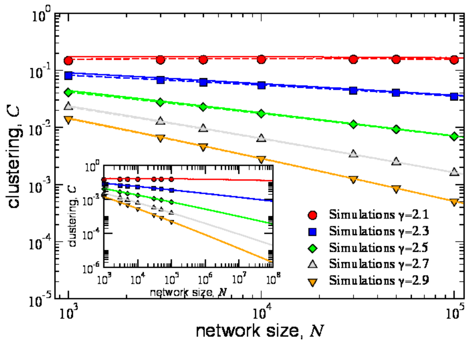

FIG. 1. Clustering coefficient as measured in numerical simulations for different values of γ and size N with $\bar{k}_{min} = 2$ and $\kappa_c = N^{1/(\gamma-1)}$. Each point is an average over 10^4 different network realizations. Dashed lines are the numerical solution and solid lines are the approximate solution. The inset shows an extrapolation up to size $N = 10^8$.

* polcolomerdesimon@gmail.com
[1] M.E.J. Newman, Random graphs as models of networks. In Handbook of graphs and Networks,S.Bornholdt and H.G. Shuster (eds.) (Wiley-VCH, Berlin, 2003)
[2] M. Boguñá, R.Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B 38, 205 (2004)