

Supplemental Material: The $s(g)$ -Metric and Assortativity

Lun Li, David Alderson, John C. Doyle, and Walter Willinger

Following the development of Newman [Newman 02], let $P(\{D_i = k\}) = P(k)$ be the node degree distribution over the ensemble of graphs, and define $Q(k) = (k+1)P(k+1)/\sum_{j \in D} jP(j)$ to be the normalized distribution of *remaining degree* (i.e., the number of “additional” connections for each node at either end of the chosen link). Let $\bar{D} = \{d_1 - 1, d_2 - 1, \dots, d_n - 1\}$ denote the remaining degree sequence for g . This remaining degree distribution is $Q(k) = \sum_{k' \in \bar{D}} Q(k, k')$, where $Q(k, k')$ is the *joint probability distribution* among remaining nodes, i.e., $Q(k, k') = P(\{D_i = k+1, D_j = k'+1 | (i, j) \in \mathcal{E}\})$. In a network where the remaining degree of any two vertices is independent, i.e., $Q(k, k') = Q(k)Q(k')$, there is no degree-degree correlation, and this defines a network that is neither assortative nor disassortative (i.e., the “center” of this view into the ensemble). In contrast, a network with $Q(k, k') = Q(k)\delta[k - k']$ defines a perfectly assortative network. Thus, graph assortivity r is quantified by the *average* of $Q(k, k')$ over all the links

$$r = \frac{\sum_{k, k' \in \bar{D}} k k' (Q(k, k') - Q(k)Q(k'))}{\sum_{k, k' \in \bar{D}} k k' (Q(k)\delta[k - k'] - Q(k)Q(k'))}, \quad (1)$$

with proper centering and normalization according to the value of perfectly assortative network, which ensures that $-1 \leq r \leq 1$. Many stochastic graph generation processes can be understood directly in terms of the correlation distributions among these so-called remaining nodes, and this functional form facilitates the direct calculation of their assortativity. In particular, Newman [Newman 02] shows that both Erdős-Renyí random graphs and Barabási-Albert preferential attachment growth processes yield ensembles with zero assortativity.

Newman [Newman 05] also develops the following sample-based definition of