

Web Note B. The yeast transcriptional regulatory network viewed as a directed random graph with specified connectivity distributions.

A random graph is a set of  $N$  vertices with edges connecting them pairwise at random<sup>1</sup>. A directed graph has oriented edges and may consequently exhibit differing numbers of *incoming* and *outgoing* edges for any given vertex. Let  $p_{in\ out}$  denote the fraction of vertices that have *in* arriving and *out* departing edges. Newman *et al.*<sup>2</sup> defined a generating function for the joint probability distribution of  $p_{in\ out}$  as the function  $G$  of two independent variables,  $x$  and  $y$  such that  $G(x,y)=\sum p_{in\ out} x^{in} y^{out}$ . They proved the following analytical expressions (parameters are defined and used in Table 1):

$$\langle in_1 \rangle = G_x(1,1) = G_y(1,1) = \langle out_1 \rangle$$

$$\langle in_2 \rangle = G_{xy}(1,1) = \langle out_2 \rangle,$$

$$\langle in_\infty \rangle = \langle out_\infty \rangle = 1 + \langle in_1 \rangle / (1 - (\langle in_2 \rangle / \langle in_1 \rangle)) = 1 + \langle out_1 \rangle / (1 - (\langle out_2 \rangle / \langle out_1 \rangle)).$$

where  $G_x$  and  $G_y$  are the first-order derivatives and  $G_{xy}$ ,  $G_{xx}$  and  $G_{yy}$  are the second-order derivatives of  $G$  with respect to  $x$  or  $y$  as indicated. Please note that the values of  $\langle in_\infty \rangle = \langle out_\infty \rangle$  diverge when  $\langle in_1 \rangle = \langle out_1 \rangle = \langle in_2 \rangle = \langle out_2 \rangle$ : a phase transition occurs as the giant component first appears.

"Cliquishness", the high local clustering of edges among cliques, is measured in undirected graphs by the clustering coefficient, *i.e.* three times the ratio of actual triangular interactions over all possible triangles<sup>2,3</sup>. Here we propose for directed graphs with  $N$  vertices the notion of "semi-cliquishness" and separately define for arriving and departing edges semi-clustering coefficients  $SC$  as  $N\Delta / N^3$  (between 0 and 1).  $N\Delta$  is the number of oriented triangles and  $N^3$  is the maximal number of potential oriented triangles (circuits excluded). Indeed, if the arriving and departing connectivities are differently distributed, upstream  $N^3$  may differ from downstream  $N^3$ . For the yeast network, fewer potential triangles are expected upstream ( $N^3_{in}$ ) than downstream ( $N^3_{out}$ ), as the upstream distribution has a shorter tail. In contrast,  $N\Delta$  must be identical, since each downstream

oriented triangle corresponds to an upstream one. As an extension of the results of Newman *et al.*<sup>2</sup>, we analytically evaluate these coefficients for directed random graphs by:

$$N3in / N = \langle in_1 (in_1 - 1) \rangle = G_{yy}(1,1)$$

$$N3out / N = \langle out_1 (out_1 - 1) \rangle = G_{xx}(1,1)$$

$$N\Delta = (\langle in_1 (in_1 - 1) \rangle / \langle in_1 \rangle) (\langle out_1 (out_1 - 1) \rangle / \langle out_1 \rangle) (\langle in_1 out_1 \rangle / \langle in_1 \rangle)$$

$$N\Delta = G_{xx}(1,1) G_{yy}(1,1) G_{xy}(1,1) / G_x(1,1)^3$$

$$SCin = N\Delta / N3in = G_{xx}(1,1) G_{xy}(1,1) / N G_x(1,1)^3$$

$$SCout = N\Delta / N3out = G_{yy}(1,1) G_{xy}(1,1) / N G_x(1,1)^3$$

We further analytically evaluate the expected number of feedback circuits comprising one vertex ( $FC_1$ ) or any number of vertices ( $FC_\infty$ ) as:

$$FC_1 = \langle in_1 out_1 \rangle / \langle in_1 \rangle = \langle in_2 \rangle / \langle in_1 \rangle = G_{xy}(1,1) / G_x(1,1)$$

$$FC_\infty = FC_1 / (1 - FC_1) = G_{xy}(1,1) / (G_x(1,1) - G_{xy}(1,1))$$

The calculated number of circuits is under 1 before phase transition, and diverges thereafter, as the giant component grows. Please note that  $FC$ , like  $N\Delta$ , does not depend on  $N$ , while  $SC$  coefficients are inversely proportional to  $N$ .

## References

1. Erdős, P. & Rényi, A. On random graphs. *Publicationes Mathematicae* **6**, 290-297 (1959).
2. Newman, M.E.J., Strogatz, S.H. & Watts, D.J. Random graphs with arbitrary degree distribution and their applications. *Phys. Rev. E* **64**, 2 (2001).
3. Watts, D.J. & Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440-442 (1998).