Electronic Appendices for: The brainstem reticular formation is a small-world, not scale-free, network

M. D. Humphries, K. Gurney, T. J. Prescott

Address: Adaptive Behaviour Research Group,
Department of Psychology, University of Sheffield,
Sheffield. S10 2TP. UK

Corresponding author: m.d.humphries@sheffield.ac.uk

Electronic Appendix A provides a summary of the functional roles of nuclei within the reticular formation other than the medial structures which are the focus of the main text.

Electronic Appendix B derives the expected number of synaptic connections for the projection and inter-neurons given a parameter set. These values form the basis of the pruning model algorithm.

Electronic Appendix C reports further values of interest from the small-world analysis of the anatomical models.

Electronic Appendix D details the results of fitting curves to the probability distribution functions corresponding to the cumulative degree distributions fits reported in the main text.

Electronic Appendix E considers the plausibility of ever finding a true scale-free network within neural tissue.

Electronic Appendix F discusses the implications of bounds from existing biological data on the parameter-dependency of the small-world topology.

A Neuroanatomy: Functional roles of the constituent fields and nuclei

Within the rostral-most third of the reticular formation, the cholinergic nuclei form the so-called mesencephalic locomotor region (Whelan, 1996) and are part of the substrate for REM sleep control (Rechtschaffen & Siegel, 2000). The medial field in this area forms part of the oculomotor circuit which also incorporates parts of the rostral (or oral) pons (Moschovakis, Scudder, & Highstein, 1996). Above the oral pons lies the norepinephrinergic nuclei which are thought to modulate general arousal and attention via cortical projections (Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999). Continuing caudally, the midline of the pons and medulla contains the serotonergic raphe nuclei, also thought to function as a general arousal system (Aghajanian & Sanders-Bush, 2002). Finally, the lateral RF regions of the pons and medulla contain fields which are specialised as cerebellar input-output arrays (Brodal & Bjaalie, 1992), as substrates for specific motor pathways (e.g. chewing (Lund, Kolta, Westberg, & Scott, 1998)) and, classically, as integrating centres for collaterals of spinal-originating ascending sensory pathways (Scheibel, 1984).

B Expected numbers of connections in the anatomical models

The pruning model algorithm runs until it reaches a target value of remaining number of synapses, which is computed according to the expected synapse totals for a given target parameter set: t_p is the target value of $P(p)$ and t_l is the target value of $P(l)$. Expected synapse total $E(N_s) = E(N_p) + E(N_l)$ is the sum of expected totals of projection $E(N_e)$ and inter-neuron $E(N_l)$ originating synapses. For both collateral variants, $E(N_l)$ is the same:

$$E(N_l) = n^-(n-1)N_c t_l, \tag{1}$$

where $n^- = n(1-\rho)$ is the number of inter-neurons in a cluster (remembering that n is the total number of neurons in a cluster).

As there are two collateral variants, there are two definitions of $E(N_p)$. For the spatially uniform variant, this is straightforward:

$$E(N_p) = N_c\ n^p\ (N_c - 1)\ P(c)\ t_p\ n, \tag{2}$$

where $n^p = n\ \rho$ is the number of projection neurons in a cluster. However, for the distance-dependent variant, the expected number of cluster contacts per projection neuron is dependent on the position of its parent cluster. Thus, we know that for a given projection neuron in cluster c, its expected number of cluster contacts $E(N_q^c)$ is

$$E(N_q^c) = 2\left(\sum_{i=1}^{d_{min}} i^{-a}\right) + \sum_{i=d_{min}+1}^{d_{max}} i^{-a} \tag{3}$$

where d_{min} and d_{max} are

$$d_{min} = \min(N_c - c, (N_c - 1) - (N_c - c)) \tag{4}$$

$$d_{max} = \max(N_c - c, (N_c - 1) - (N_c - c)) \tag{5}$$

in words, d_{min} is the minimum and d_{max} is the maximum number of intervening clusters to either end of the model from cluster c. Thus, $E(N_p)$ for the distance-dependent model is given by

$$E(N_p) = n^p\ \left(\sum_{c=1}^{N_c} E(N_q^c)\right)\ t_p\ n. \tag{6}$$

C Small-world topologies of the stochastic model

By ranking parameter combinations by S_{max} the same 6 combinations appear in descending order for both the distance-dependent and spatially-uniform collateral versions (see Table 1).

Table 1: First six parameter combinations ordered by maximum S for spatially-uniform (U) and distance-dependent (D) collateral probability models.

ρ	$P(l)$	$P(p)$	S^u_{max}	S^d_{max}
0.7	0.9	0.1	4.6612	10.0513
0.7	0.5	0.1	3.1197	6.9934
0.8	0.9	0.1	2.6488	5.8910
0.8	0.5	0.1	1.9645	4.2939
0.7	0.9	0.5	1.6500	3.4214
0.9	0.9	0.1	1.4694	3.0680

D Degree distribution curve-fitting

We used the curve equations and initial conditions given in Table 2. As is common practice, these were fitted to the *inverted* cumulative degree distribution $1 - F(\beta)$ (we retain the $F(\beta)$ term in the main text to avoid unnecessary confusion) using the MatLab function lsqcurvefit. The inverted distribution was used in line with previous work (Amaral, Scala, Barthelemy, & Stanley, 2000; Strogatz, 2001): this ensures, among other things, that $\tau > 0$ as required. Before fitting the exponential curve, all values in the data vector $\mathbf{x}$ were shifted by the first element($\mathbf{x} = \mathbf{x} - x_1$) so that the $x_1 = 0$, and thus the exponential fit would start at 1.

Table 2: Fitted curves to the degree distribution and initial search parameters. Data point is x.

distribution		initial parameters
exponential	$\mathrm{e}^{-\lambda x}$	$\lambda = 2$
power law	$ax^{-\tau}$	$a = 1, \tau = 2$
truncated power law	$ax^{-\tau}\mathrm{e}^{-\lambda x}$	$a = 1, \tau = \lambda = 2$
Gaussian	$1 - (0.5(1 + \mathrm{erf}\frac{x-\mu}{\sigma\sqrt{2}}))$	$\mu = \mathrm{mean}(x), \sigma = 1$

The curve-fits were quantitatively compared using Akaike's Information Criterion (corrected). The basic form is calculated according to

$$AIC = N \ln\left(\frac{SS}{N}\right) + 2K \tag{7}$$

where N is the number of data points, SS is the sum-of-squares resulting from the curve fit, and K is the number of parameters in the curve plus one (so, for example, $K = 2$ for the exponential curve). The corrected version adds a term to account for situations in which $N \to K$:

$$AIC_c = AIC + \frac{2K(K+1)}{N-K-1}. \tag{8}$$

We also fitted curves to the degree distribution $P(\beta)$ of the stochastic model, primarily to investigate the poor fits of the curves to the output $F(\beta)$ illustrated in main text. The

curve family used included exponential, Gaussian, double Gaussian, quadratic, and power law fits (see Table 3). A quadratic fit was included to fit $P(\beta)$ that the other curves could not.

Table 3: Fitted curves to the degree distribution $P(\beta)$ and initial search parameters. G_1 and G_2 refer to separately parameterised Gaussians G.

distribution		initial parameters
exponential	$\lambda e^{-\lambda x}$	$a = \lambda = 2$
power law	$ax^{-\tau}$	$a = 1, \tau = 2$
Gaussian	$G = (1/\sigma\sqrt{(2\pi)})\exp(-(x-\mu)^2/(2\sigma^2))$	$\mu = \text{mean}(x), \sigma = 1$
double Gaussian	$G_1 + G_2$	$\mu_1 = \min(x), \mu_2 = \max(x), \sigma_1 = \sigma_2 = 1$
quadratic	$a + bx + cx^2$	$a = b = c = 1$

Most of the curve fits were Gaussian or double Gaussian (Table 4) - examples are shown in Figure 1. The double Gaussian fits corresponded to the cumulative degree distributions $F(\beta)$ poorly fitted by the tested curves, as illustrated in the main text. There is little evidence to suggest that a Gaussian-based distribution of connectivity is not to be expected for neural networks, but we acknowledge that the work on large-scale distributions of axonal or synaptic connectivity is sparse. The observed bimodal (double Gaussian) distributions are due to the different distributions of connectivity for each of the two neuron classes (inter- and projection-neurons). Numerous power law best-fits to $P(\beta)$ were found, which taken at face value would imply the presence of a scale-free topology. However, the corresponding distributions were not characteristically heavy-tailed (example in Figure 1c) and the corresponding $F(\beta)$ were best-fit by a Gaussian. We use this result to sound a cautionary note to the reader: despite the generally used, loose definition of a scale-free network, a power-law best fit to a degree distribution is not sufficient evidence to conclude that a network is scale-free (see Li, Alderson, Tanaka, Doyle, and Willinger (2004) for a more rigorous treatment).

Table 4: Percentage best-fit of each curve-type to the $P(\beta)$ of input, output, and undirected links in the spatially-uniform and distance-dependent collateral versions of the stochastic model.

	spatially-uniform			distance-dependent		
Curve type	inputs	outputs	undirected	inputs	outputs	undirected
exponential	0.0	2.0	0.0	0.0	0.5	0.0
power law	0.0	22.2	11.4	0.0	20.7	0.0
Gaussian	99.6	24.4	50.4	98.1	26.4	73.8
double Gaussian	0.4	46.9	11.6	1.9	42.0	1.0
quadratic	0.0	4.5	26.6	0.0	10.4	25.2

These results have thrown up an interesting general question about scale-free topologies: is it possible for a network to be scale-free in one direction only? As we found,

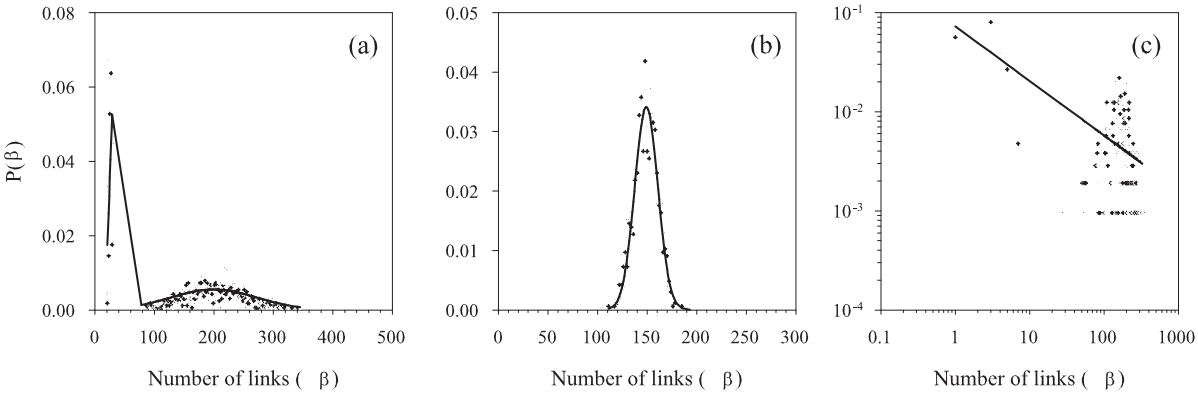

Figure 1: Best-fit curves to degree distribution $P(\beta)$, all from the stochastic model. (a) An example of a double Gaussian fit to an output distribution, suggesting the presence of two independent populations within the model. The double Gaussian fits to $P(\beta)$ seem to correspond to the poorly fitting curves of $F(\beta)$, as illustrated in the main text. (b) A Gaussian fit to an input distribution from the distance-dependent collateral model - again the dominant best-fit of the tested model curves. (c) Log-log plot showing that the output distribution from the same model instantiation as (b) was best-fit by a power-law. The fit is evidently poor, and the corresponding $F(\beta)$ was best-fit by a Gaussian. Nevertheless, the input and output distributions of this particular model differ considerably.

based on the quantitative fit results alone, some output $P(\beta)$ were best fit by different curves to the corresponding input and undirected distributions. Compare the input and output distributions of Figure 1b, c - these are taken from the same instantiation of the spatially-uniform collateral version of the stochastic model. They clearly show that the $P(\beta)$s are different (as are the corresponding $F(\beta)$s). As noted by Newman, Strogatz, and Watts (2001) input and output degree distributions of networks are often assumed to be correlated, but there is no *a priori* requirement that this is the necessary. A randomly chosen node in a network has probability p_{ij} that it has a particular in-degree i and out-degree j: if the underlying physical process that generates the connection between nodes is independent of the input-output relationship (if the number of links into a node does not effect the number of links coming out, and vice-versa) then the separate distributions are independent and therefore $p_{ij} = p_i p_j$. This in turn would imply that that the undirected distribution would tend towards a Gaussian (or multiple Gaussian). Thus, in general it is possible that input and output degree distributions are not the same, and therefore that a network could be scale-free in one direction only.

E Power-law degree distributions in neural tissue

It has not escaped our notice that the concepts of over-growth and pruning of connectivity could form the basis for a general inverse model of scale-free network generation (Barabasi & Albert, 1999). The general properties of neural development have similarities to the truncated power-law distribution algorithm proposed by Amaral et al. (2000). They suggest that network growth is limited by "aging", where a node becomes inactive after

a set period, and "cost", where a node's ability to support links is physically limited. Both properties apply to neural network development: initial axonal growth and synapse formation is time-limited; the quantity of connections a neuron can develop and sustain is physically limited by the metabolic cost of development, maintenance, and signalling (Cline, 2003; Laughlin & Sejnowski, 2003). The intriguing parallels suggest that it is possible a truncated scale-free topology exists within the nervous system of some species - at some level of neural organisation (neuronal, areal, and so on) - and thus it was not improbable that the medial RF had such a topology.

It is an open question as to whether or not a scale-free network can be developed within neural tissue. The limitations of neural development noted above would suggest that only partially scale-free topologies would be possible. Moreover, although the existence of super-connected hubs would make neural tissue more resistant to random damage - by maintaining network connectivity - the death of such neurons would be catastrophic, and any targeted disease would be highly effective. Such neurons or neuron populations (in particular) should be revealed by systematic anterograde or retrograde staining studies, due to the sheer number of connections they maintain, and would thus be amenable to discovery. On the one hand, as it appears that no such super-connected neurons are reported in the neuroscience literature, it is tempting to conclude that the characteristic hubs of scale-free networks do not exist in neural tissue. On the other hand, given the minute proportion of neurons stained in a typical study compared to the amount in the originating structure, the sampling bias that results in only certain neuron classes being stained within a structure, and the inconsistent uptake of the staining agent, it is possible that either: (a) there has been insufficient sampling to consistently reveal super-connected neurons or (b) that most reported stained neurons *are* super-connected, which is why they stained consistently in the first place.

F Dependence of the small-world topology on connection probabilities

The extent to which the medial RF conforms to a small-world topology is partly dependent upon the determination of synaptic connection probabilities (we discuss its dependence on the type of axon collateral distribution and on the validity of the anatomical models in the main text). There are no existing direct estimates of connection probabilities for the medial RF. Data from mouse cortex (Schuz, 1995) and the neural network of *C. elegans* (Albert & Barabasi, 2002) suggests that the probability of connection between any randomly chosen pair of neurons is $p \leq 0.1$ and therefore $P(p), P(l) \leq 0.1$. If these probabilities were applicable to the medial RF, they would rule out a small-world topology if the combination of stochastic anatomical model and spatially-uniform collaterals were verified (see Figure 2, main text).

However, extrapolating from these unrelated instances is probably unsafe. The dimensions of the projection neuron dendritic tree suggests, rather, that any axon collateral terminating in a cluster will contact the majority of projection neurons there, and thus $P(p)$ at least should be higher than this estimate. We can speculate on the relative connectivity from the reviewed data. An absolute maximum of 30% of the cell population

are GABAergic, and putative inter-neurons, but $\sim 45\%$ of the synapses on a giant neuron are GABAergic with no other main afferent GABAergic source, suggesting that the interneurons are proportionally more densely connected within a cluster than the incoming projection neurons collaterals. Thus, $P(l) > P(p)$, as required for the most small-world like combinations of the stochastic model - the pruning model would, of course, also have a small-world topology if this relationship held.

References

Aghajanian, G. K., & Sanders-Bush, E. (2002). Serotonin. In K. L. Davis, D. Charney, J. T. Coyle, & C. Nemeroff (Eds.), *Neuropsychopharmacology: The fifth generation of progress* (p. 15-34). Philadelphia: Lippincott Williams & Wilkins.

Albert, R., & Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, *74*, 47–97.

Amaral, L. A., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proc. Natl. Acad. Sci. U.S.A.*, *97*(21), 11149–11152.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512.

Brodal, P., & Bjaalie, J. G. (1992). Organization of the pontine nuclei. *Neurosci. Res.*, *13*(2), 83–118.

Cline, H. (2003). Sperry and Hebb: oil and vinegar? *Trends Neurosci.*, *26*(12), 655–661.

Laughlin, S. B., & Sejnowski, T. J. (2003). Communication in neuronal networks. *Science*, *301*(5641), 1870–1874.

Li, L., Alderson, D., Tanaka, R., Doyle, J. C., & Willinger, W. (2004). *Towards a theory of scale-free graphs: definition, properties, and implications (extended version).* (preprint cond-mat/0501169)

Lund, J. P., Kolta, A., Westberg, K. G., & Scott, G. (1998). Brainstem mechanisms underlying feeding behaviors. *Curr. Opin. Neurobiol.*, *8*(6), 718–724.

Moschovakis, A. K., Scudder, C. A., & Highstein, S. M. (1996). The microscopic anatomy and physiology of the mammalian saccadic system. *Prog. Neurobiol.*, *50*(2-3), 133–254.

Newman, M. E., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, *64*(2 Pt 2), 026118.

Rechtschaffen, A., & Siegel, J. M. (2000). Sleep and dreaming. In E. Kandel, J. Schwartz, & T. Jessel (Eds.), *Principles of neuroscience* (p. 936-947). New York: McGraw-Hill.

Scheibel, A. B. (1984). The brainstem reticular core and sensory function. In J. M. Brookhart & V. B. Mountcastle (Eds.), *Handbook of physiology. Section 1: The nervous system* (p. 213-256). Bethesda, Maryland: American Physiological Society.

Schuz, A. (1995). Neuroanatomy in a computational perspective. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (p. 622-626). Cambridge, MA: MIT Press.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268–276.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999).

The role of locus coeruleus in the regulation of cognitive performance. *Science*, *283*(5401), 549–554.

Whelan, P. J. (1996). Control of locomotion in the decerebrate cat. *Prog. Neurobiol.*, *49*(5), 481–515.