

Deviation from Criticality for High Level Cognition

Tom Lorimer, Florian Gomez, and Ruedi Stoop

Institute of Neuroinformatics, University of Zurich and ETH Zurich
Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Email: lorimert@ini.phys.ethz.ch, fgomez@ini.phys.ethz.ch, ruedi@ini.phys.ethz.ch

Abstract—Claims of “brain criticality” suggest that the generation of activity on all scales in a critical state (e.g. neuronal avalanches) may underlie complex, adaptive cognition. However the relationship of the power-law distributed network observables behind these claims to brain functional output is not known. Here we use a novel approach to this problem by considering functional output directly: the networks generated by *Drosophila melanogaster* during pre-copulatory courtship. This courtship body language, treated as a symbolic dynamics, has been shown to exhibit Context-Free and Context-Sensitive grammars in the Chomsky hierarchy; here we show that the underlying networks corresponding to such grammars deviate from scale-free structure. We provide a simple network growth model, which matches the degree distributions of the non-scale-free networks by breaking the preferential attachment paradigm with a second internal linking process. From these observations, we suggest that the higher level cognition associated with such grammatical structure may not be compatible with a critical state.

1. Introduction

Fundamental to abstract reasoning is the ability to place cognitive “objects” in relation to one another. The explicit manifestation of this ability in language, formalised through grammar classes (the “Chomsky hierarchy”), has been conjectured to be the defining characteristic of human intelligence [1]. By solving a classic undecidability problem with statistical argument, recent work has shown that this is not the case: higher language classes can even be expressed by the simple fruit fly [1]. Here we examine this recent surprising example from *Drosophila melanogaster* pre-copulatory courtship (DPC) in light of claims (e.g. Ref. [2]) that “brain criticality” underlies complex adaptive cognition.

Claims of brain criticality have mostly relied on the observation of power laws, though some authors have also suggested that deviations from power law behaviour either side of the claimed critical state further enhance this evidence. Experimental limitations hinder this line of argument: there is seldom sufficient resolution of an order parameter to observe the necessary cusp at the proposed critical control parameter value (e.g. [3]), and isolated experimental cases which deviate from a power law serve only

to demonstrate that non-critical behaviour is also possible (e.g. [4]). Here we focus not on the existence of critical states, but rather on their utility in cognition.

Intuitively, it may seem that the availability of structure at all scales in a critical state should be useful for complex adaptive function, however we question the supporting evidence behind this intuition on two levels: 1) the dependence of simulation results on the network structures and measurements used; and 2) the unknown relationship between observables and function in biological networks. On the simulation level for instance, “reservoir computing” systems may perform calculations on time series when the “reservoir” is near a critical state [5], however this is contingent on learning occurring on feed-forward inputs and outputs, which are not considered in the criticality claim. In another example, a measure of information propagation based on avalanche size distributions (a classic critical order parameter) was maximised in the critical state, but when noise was added (as is always the case in biological systems), the critical state no longer maximised this measure [6]. Even the maximal dynamic range claimed in the critical state is not so clearly useful. Indeed, in modelling the mammalian cochlea, the correct sensitivity profile is obtained when the Hopf oscillator elements are tuned *away* from the critical bifurcation point (see e.g. [7] for a single element, and [8] for the networked context). On the biological level, the first invasive multielectrode array observations of power law distributed neuronal avalanches in rat cortex slices [9] were recorded *in vitro* (thus with no clear functional relevance), but even *in vivo*, neuronal avalanche sizes [10] or high activity brain regions [11] do not have clearly defined functional outputs (the search for such relationships, in fact, partially motivates the brain criticality conjecture). The above examples are by no means an exhaustive or conclusive rejection of brain criticality, and this is not our intention. Instead, we wish to motivate the approach used in this paper, which examines real biological networks, but avoids the function issue by working directly with functional output.

2. Networks Underlying DPC

From visual inspection of high speed video courtship recordings, DPC was decomposed into 37 fundamental actions: body language acts which are non-overlapping, occur on timescales well beyond the limit of neural refrac-

Table 1: Definitions of the DPC sequence groups. Here “females” comprises: *immature virgin female*; *mature virgin female*; *mated female*. See text for explanation of *female behaviour* group.

group name	included sequence types
<i>male behaviour</i>	<i>male</i> vs. females
<i>female behaviour</i>	females vs. <i>male</i> <i>male</i> vs. <i>fruitless</i> <i>mature v. fem.</i> vs. <i>fruitless</i>
<i>male and female behaviour</i>	seq. in <i>male behaviour</i> seq. in <i>female behaviour</i>

tory periods, do not generally have inherent physical restrictions on their ordering, and for which further decomposition yields no further information about the courtship sequence [12]. Thus the actions performed by each individual *Drosophila* during courtship pairing can be represented as a finite sequence $S_a = \{a_1, a_2, \dots, a_m\}$, where each $a_i \in S_a$ is one of the 37 fundamental actions.

The undirected topological network underlying such a DPC sequence has vertices defined by the actions present, and edges defined by adjacency of actions (vertices) in the sequence. i.e., vertex set $V = \{v_i : v_i \in S_a\}$ and edge set $E = \{(v_i, v_j) : v_i = a_n, v_j = a_{n+1}; a_n, a_{n+1} \in S_a\}$. Generally we deal with networks composed from more than one sequence, which are defined in the usual graph union sense. This topological, undirected representation avoids introducing bias due to incomplete data, while still elucidating the structure on which the sequences were (necessarily) produced by some network walk.

In all, 10 fundamental action sequence types were recorded [12], each corresponding to one pairing-protagonist combination. For instance, the pairing of *male* and *mature virgin female* drosophila gives the two sequence types (protagonist vs. antagonist): *male* vs. *mature virgin female*; and *mature virgin female* vs. *male*. 3 main pairings give 6 of the sequence types: *male* paired with each of *immature virgin female*, *mature virgin female* and *mated female*. Further pairing of males with a genetic mutation, here termed *fruitless*, with each of (normal) *male* and *mature virgin female* give a further 4 sequence types (a total of 10). Here we examine the sequence types in 3 groups: *male behaviour*; *female behaviour*; and *male and female behaviour* (Table 2). We use *behaviour* in the sense defined in [12, 13], so that the *female behaviour* includes the gender switching observed in *male* vs. *fruitless* (however sequences produced by *fruitless* as protagonist are excluded).

2.1. Truncated Power Law Fitting

If the neural system producing the DPC networks is operating in a critical state, the network degree distribution,

as the key structural observable, could be expected to follow a (truncated) power law. Great statistical care must be taken when testing for power laws, particularly in small datasets, where truncation effects become increasingly important. We follow closely the method outlined in [14] (which in turn closely follows [15]). Maximum likelihood estimation of the discrete truncated power law exponent is used to define a fitted distribution, from which 1000 surrogate datasets are sampled. Each of these surrogate datasets is then also fit by maximum likelihood estimation, and their Kolmogorov-Smirnov (KS) distances are compared to the KS distance of the original data to its fit. A “*p*-value” (here, a measure of goodness of fit) is then calculated as the fraction of surrogate datasets with a larger KS distance (a worse fit) than the original data. A high *p*-value is not sufficient to demonstrate that the data follow a power law, but only to indicate feasibility. A low *p*-value is sufficient for rejection of a power law, though the selection of a rejection threshold is rather arbitrary.

The lower and upper truncation bounds were selected as the minimum and maximum degrees present due to the small size of the DPC networks, except in the *male and female behaviour* case, where the second largest degree was selected, avoiding a large gap in the degree distribution. To avoid additional noise and binning issues, we plot results as the complement of the cumulative distribution $P(k)$, called the *survival function*, $SF(k) = 1 - P(k)$. Readers may expect that a power law should appear as a straight line in log-log space, however we emphasise that for a truncated survival function this is not the case, as the distribution must reach zero at the upper bound. Over a small range, power law and exponential distributions are very similar; we also tested exponential distributions on our data and obtained similar fits and *p*-values (not shown).

3. Results, Modeling, Discussion

The truncated power law fits to the DPC network degree distributions (Fig.1) reveal a structural difference which coincides with previous results regarding the grammatical structure of the language underlying the networks. The *female behaviour* network (Fig. 1a), which is generated from courtship sequences showing limited recursion, shows a plausible power law fit ($p = 0.62$). The *male behaviour* and *male and female behaviour* networks on the other hand (Fig. 1b,c), which are generated from sequences showing greater recursion, deviate sharply from a power law at higher degree (both $p = 0.04$). Notwithstanding the small size of the datasets, the distinctive way in which the datasets containing more recursion deviate warrants further investigation. In fact, another example of a similar deviation has previously been observed in human language networks, or “word-webs”. Ferrer i Cancho and Solé [16] showed that Zipf’s “law” for the frequency of occurrence of words in English breaks into two scaling regimes for sufficiently large *corpi*; an effect later reproduced by a simple

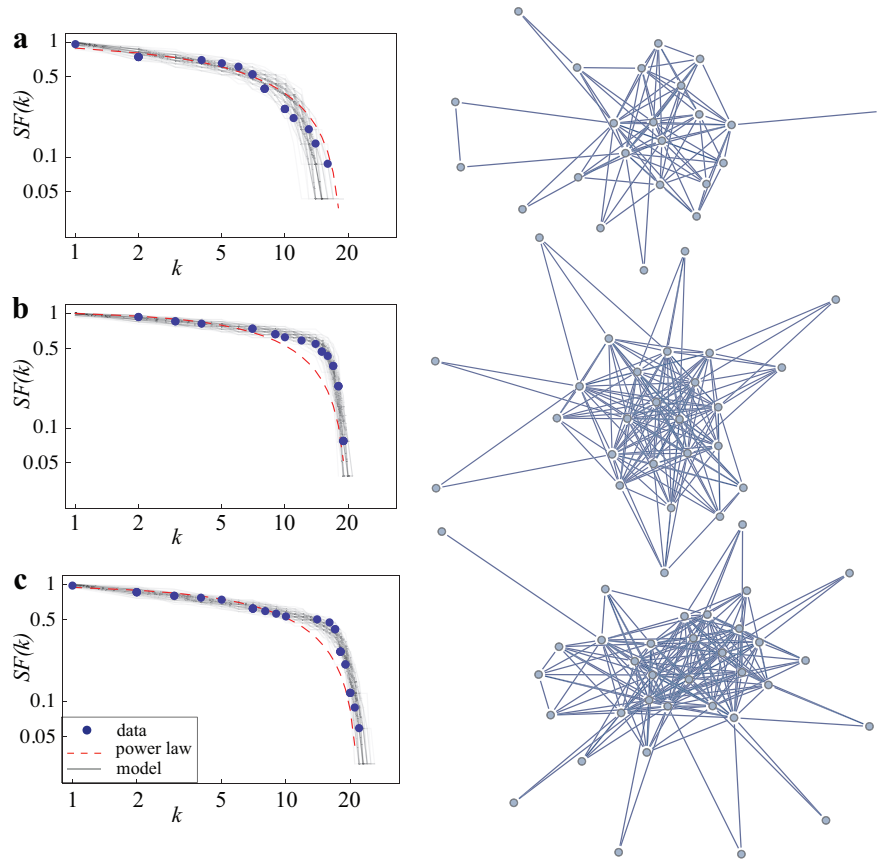


Figure 1: Top panel: log-log plots of truncated power law fits (*dashed lines*) and 100 realisations of network growth model (*continuous translucent lines*) with network degree distribution survival functions (*circles*, plotted only for degrees present). Bottom panel: corresponding networks. **a** *female behaviour*: $p = 0.62$; power law exponent 0.37; degree bounds $k_{min} = 1, k_{max} = 19$; KS difference 0.10. **b** *male behaviour*: $p = 0.04$ (rejected); power law exponent 0.001 (minimum permitted); degree bounds $k_{min} = 2, k_{max} = 20$; KS difference 0.22. **c**. *male and female behaviour*: $p = 0.04$ (rejected); power law exponent 0.076; degree bounds $k_{min} = 1, k_{max} = 22$ (2^{nd} highest degree); KS difference 0.18.

network growth algorithm for the undirected topological word network by Dorogovtsev and Mendes [17]. This original algorithm in Ref. [17] added an internal linking process to the standard preferential attachment growth algorithm (from Ref. [18]) whereby “internal” edges between vertices v_i and v_j were added with probability $p \propto k_i k_j$, where k denotes vertex degree. The DPC networks are far more densely connected than these word webs however, so a “saturation” effect appears, whereby the probability of a particular new edge appearing under such a scheme would be strongly modified by the probability that it already exists.

Taking this saturation effect into consideration, Dorogovtsev and Mendes’ algorithm can be modified such that the rate of addition of edges is no longer explicit, but instead implicit, depending on the “success” of adding a new connection. We use here a growth algorithm which instead selects a *constant* number r of possible edges at each step (an edge can be selected more than once) between existing

vertices with $p \propto k_i k_j$, and these are added to the network only if they are not already present. This is a fundamentally different mechanism to the Dorogovtsev and Mendes algorithm: the edge addition rate truncates at high degree, which tends to distort the distribution towards these values. The addition of new vertices occurs in the same way as the standard preferential attachment algorithm, with a new vertex joined to old vertices v_i by m new edges at each step with $p \propto k_i$. The parameters r and m can be defined directly from the DPC networks and data. It has been observed that DPC can be characterised by periodic orbits [12], with mean orbit length > 2 , so each new vertex is most likely to join the network by connecting to 2 existing vertices, i.e. $m = 2$. The parameter r can be experimentally varied until the number of edges in the synthetic network matches the number of edges in the real network on average (for the same number of vertices). This *a priori* parameter setting yields a strikingly good match to the degree distributions of the *male behaviour* and *male and female behaviour* net-

works without need for parameter fitting (Fig. 1b,c), and a less convincing fit to the *female behaviour* network degree distribution (Fig. 1a).

How can these results be interpreted in terms of brain criticality? The correspondence between the level of the underlying grammar and the breakdown of the power law fit, is further supported by the success of the simple growth algorithm in reproducing this effect. Grammars which express recursion require a second “book-keeping” function; the DPC networks to which they correspond are also matched by a two process algorithm, and not by a single scale. This suggests that the fingerprint of meso-scale structure on networks generated by recursive grammars may not be compatible with criticality. Recursion itself is essential for higher level abstract reasoning [1], which may generate functional network structure that is inconsistent with the notion of brain criticality.

In conclusion, we have presented a clear example of how power laws may not highlight essential higher functional processes. Power laws observed by other authors during brain function, need not be power laws of functional states themselves. Even so called “functional networks” (e.g. [11]) are only defined by activity correlations. It is not clear that the observables studied in claims of brain criticality are active and relevant in cognition. The DPC networks by contrast are networks *of* function, during specific goal directed behaviour. We suggest that this “network of function” approach may prove fruitful in the brain criticality debate.

Acknowledgments

The authors acknowledge the support of the Swiss National Science Foundation SNF, and the Swiss Federal Commission for Scholarships for Foreign Students FCS.

References

- [1] R. Stoop, P. Nüesch, R. L. Stoop, L. A. Bunimovich, “At Grammatical Faculty of Language, Flies Outsmarts Men,” *PLoS ONE*, vol.8, pp.e70284, 2013.
- [2] D. R. Chialvo, “Emergent Complex Neural Dynamics,” *Nature Phys.*, vol.6, pp.744–750, 2010.
- [3] W. L. Shew, H. Yang, S. Yu, R. Roy, D. Plenz, “Information Capacity and Transmission Are Maximized in Balanced Cortical Networks with Neuronal Avalanches,” *J. Neurosci.*, vol.31, pp.55–63, 2011
- [4] N. Friedman, S. Ito, B. A. W. Brinkman, M. Shimo-no, R. E. L. DeVille, K. A. Dahmen, J. M. Beggs, T. C. Butler, “Universal Critical Dynamics in High Resolution Neuronal Avalanche Data,” *Phys. Rev. Lett.*, vol.108, pp.208102, 2012.
- [5] N. Bertschinger, T. Natschläger, “Real-Time Computation at the Edge of Chaos in Recurrent Neural Networks,” *Neural Comput.*, vol.16, pp.1413–1436, 2004.
- [6] P. Rämö, S. Kauffman, J. Kesselia, O. Yliharjaa, “Measurements for Information Propagation in Boolean Networks,” *Physica D*, vol.227, pp.100–104, 2007.
- [7] A. Kern, R. Stoop, “Essential Role of Couplings between Hearing Nonlinearities,” *Phys. Rev. Lett.*, vol.91, pp.128101, 2003.
- [8] F. Gomez, R. Stoop, (in preparation)
- [9] J. M. Beggs, D. Plenz, “Neuronal Avalanches in Neocortical Circuits,” *J. Neurosci.*, vol.23, pp.11167–11177, 2003.
- [10] T. Petermann, T. C. Thiagarajan, M. A. Lebedev, M. A. L. Nicolelis, D. R. Chialvo, D. Plenz, “Spontaneous Cortical Activity in Awake Monkeys Composed of Neuronal Avalanches,” *Proc. Natl. Acad. Sci. U.S.A.*, vol.106, pp.15921–15926, 2009.
- [11] V. M. Eguíluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, A. V. Apkarian, “Scale-Free Brain Functional Networks,” *Phys. Rev. Lett.*, vol.94, pp.018102, 2005.
- [12] R. Stoop, B. I. Arthur Jr., “Periodic Orbit Analysis Demonstrates Genetic Constraints, Variability, and Switching in Drosophila Courtship Behavior,” *Chaos*, vol.18, pp.023123, 2008.
- [13] R. Stoop, J. Joller, “Mesoscopic Comparison of Complex Networks Based on Periodic Orbits,” *Chaos*, vol.21, pp.016112, 2011.
- [14] A. Deluca, A. Corral, “Fitting and Goodness-of-Fit Test of Non-Truncated and Truncated Power-Law Distributions,” *Acta Geophys.*, vol.61, pp.1351–1394, 2013.
- [15] A. Clauset, C. R. Shalizi, M. E. J. Newman, “Power-Law Distributions in Empirical Data,” *SIAM Rev.*, vol.51, pp.661–703, 2009.
- [16] R. Ferrer i Cancho, R. V. Solé, “Two Regimes in the Frequency of Words and the Origins of Complex Lexicons: Zipf’s Law Revisited,” *J. Quant. Linguist.*, vol.8, pp.165–173, 2001.
- [17] S. N. Dorogovtsev, J. F. F. Mendes, “Language as an Evolving Word Web,” *Proc. R. Soc. Lond. B*, vol.268, pp.2603–2606, 2001.
- [18] A.-L. Barabási, R. Albert, “Emergence of Scaling in Random Networks,” *Science*, vol.286, pp.509–512, 1999.