DYNAMICS DAYS SOUTH AMERICA 2010
International Conference on Chaos and Nonlinear Dynamics

INPE – National Institute for Space Research
São José dos Campos – SP – Brazil – July 26-30, 2010

IDENTIFICATION OF FUNCTIONAL INFORMATION SUBGRAPHS IN COMPLEX NETWORKS

Vadas Gintautas[1], *Luís M. A. Bettencourt*[2], *Michael I. Ham*[3]

[1]T-5 and CNLS; MS B284, Los Alamos National Laboratory, Los Alamos NM 87545, USA, vadasg@lanl.gov
[2]T-5; MS B284, Los Alamos National Laboratory, Los Alamos NM 87545, USA, lmbett@lanl.gov
[3]P-21; MS D454, Los Alamos National Laboratory, Los Alamos NM 87545, USA, mikeh@lanl.gov

keywords: Neuronal Dynamics, Time Series Analysis, Complex Networks, Dynamics of Biosystems

Information plays a central role in conditioning structure and determining collective dynamics in many complex systems. For example, the ability to process and react to information certainly influences how neurons and synapses interact in large numbers to generate the complexity of cognitive processes. However, systematic methodologies for identifying functional relations between units of successive complexity, involved in information processing and storage, are still largely missing. New interest has arisen in applications of information theory to dynamical and statistical systems with many degrees of freedom. Specifically, information quantities have been shown to identify and classify spatial [1] and temporal [2] correlations, and reveal if a group of variables may be mutually redundant or synergetic [3, 4]. In this way an information theoretic treatment of groups of correlated degrees of freedom can reveal their functional roles as memory structures or those capable of processing information.

Here we present an approach based on information theory applied to the correlated statistical state of many variables. The uncertainty in the state of X can be quantified by its Shannon entropy $S(X)$; measuring a correlated variable Y contributes reduces its uncertainty, thus the conditional entropy of X given Y is $S(X|Y) \leq S(X)$ [5]. We also consider the optimization problem of minimizing the uncertainty in X given k measurements $\{Y\}_k$ within a larger set, where $\{Y\}_k$ refers to the set $Y_1, \ldots, Y_k$. Specifically, if a set exists such that $S(X|\{Y\}_k) = 0$, then it fully determines the state of X and no uncertainty remains. We seek the measurement set of size k resulting in the smallest remaining uncertainty. The computational complexity of this search grows combinatorially with the number of arrangements of size k within n variables, which quickly becomes prohibitive. To evade this problem, we introduce the exact expansion

$$S(X|\{Y\}_k) - S(X) = -I(X;\{Y\}_k) \quad (1)$$
$$= \sum_i \frac{\Delta S(X)}{\Delta Y_i} + \sum_{i>j} \frac{\Delta^2 S(X)}{\Delta Y_i \Delta Y_j} + \ldots + \frac{\Delta^k S(X)}{\Delta Y_1 \ldots \Delta Y_k}.$$

The variational operators in Eq. (1) define the change in entropy resulting from a measurement as

$$\frac{\Delta S(X)}{\Delta Y_i} \equiv S(X|Y_i) - S(X) = -I(X;Y_i) \quad (2)$$
$$\frac{\Delta^2 S(X)}{\Delta Y_i \Delta Y_j} \equiv -\frac{\Delta I(X;Y_i)}{\Delta Y_j} = I(X;Y_i) - I(X;Y_i|Y_j), \quad (3)$$

and so on. Higher order variations follow automatically from the successive application of the first variation. This expansion has two important properties. First, each term in the expansion at order k accounts for an irreducible set of correlations among a size-k group of Y_i nodes with the target X. Statistical independence among any of the Y_i results in a vanishing contribution to that order and terminates the expansion. If the Y_i are correlated in pairs, but not in higher order multiplets, then only terms with $k \leq 2$ will be present, and so on. Thus, for a system where not all correlations are realized, expression Eq. (1) allows the identification of correlated submultiplets, and determines their mutual organization in specifying the state of X. The second important property of this expansion is that the sign of each nonvanishing variation reveals the informational character of the corresponding multiplet. Specifically, negative indicates that the k-multiplet contributes to the state of X with more information than the sum of all its subgroups (synergy), while positive indicates the opposite (redundancy) [4]. We define a synergetic (redundant) core as a set $\{Y\}_k$ such that its variation and the variations of all its subgroups of two or more nodes are negative (positive). Here we use the expansion in Eq. (1) to define the problem of determining the set and decomposition of the Y_i in terms of functional information arrangements that best account for the stochastic behavior of a target X.

To illustrate this methodology, we apply it to temporal action potential activity from murine frontal cortex neuronal cultures grown in vitro on non-invasive microelectrode arrays (MEAs) [6]. Details of MEA fabrication and culture preparation are described elsewhere [4, 7]. These experimental platforms have become model systems for studying living neuronal networks in controlled environments. Results presented here refer to 62 cells of a 42 day old cortical network.

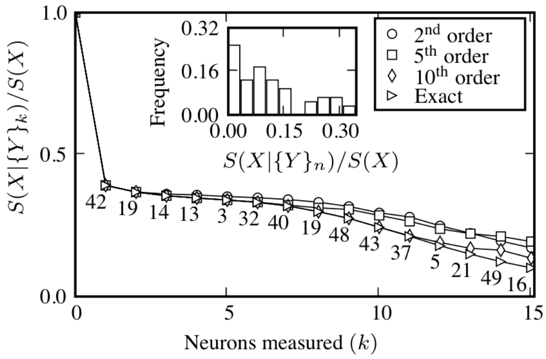

Figure 1 – Joint entropy of neuron 46 and a set of other neurons of size k. The next neuron measured is chosen by maximizing the variation to various orders; the neuron numbers appear for the exact curve. Inset: Histogram of entropy fraction for each neuron remaining after all possible measurements.

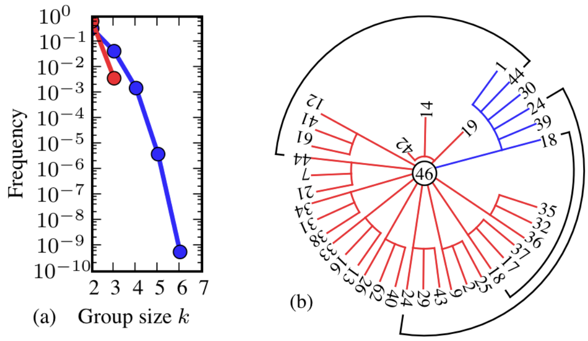

Figure 2 – (a) Frequency of redundant (red) and synergetic (blue) cores versus size k. (b) Purely redundant (red) and purely synergetic (blue) circuits relative to neuron 46. Neurons and groups with the most information about 46 are closest to the center. Arcs identify neurons in multiple functional groups.

We analyze a 5 hour recording (2.3 million spikes) of network activity. We construct binary states for each neuron's time series and use these to estimate probability distributions for firing pattern states of k neurons. Fig. 1 shows the relative entropy reduction of a target neuron, due to successive measurements of other neurons. Different lines correspond to searches for the optimal sequence of measurements at different orders of approximation in the expansion in Eq. (1). A search to exact order means that all $\{Y\}_k$ are considered, given the previous $\{Y\}_{k-1}$, and the set $\{Y\}_k$ with greatest information gain is chosen. Fig. 1(inset) shows the histogram of the ratio of final to initial entropy for all 62 neurons. Final entropy refers to the fraction of a neuron's initial entropy left unaccounted for once all other neurons are measured. Remarkably, the stochastic patterns of most cells can be nearly fully predicted by the activity of others, even if most degrees of freedom in the actual network remain unobserved.

Fig. 2(a) shows the frequency of synergetic and redundant cores, while Fig. 2(b) shows the reconstruction of circuits from functional subgraphs which account for the activity of target neuron 46 of Fig. 1. Evidently the target neuron is part of both redundant and synergetic functional multiplets, with the former being substantially more abundant. The target neuron is also part of a synergetic circuit with other neurons, several of which are part of smaller mutually redundant subgraphs. Some of these can, at least partially, be interchanged with other neurons carrying the same information, resulting globally in an interconnected ensemble where specific synergetic functional relationships are embedded on robust redundant cell arrangements.

In summary, we present a new information theoretic approach to constructing functional subgraphs in complex networks where nodes display observable stochastic dynamics. By performing targeted searches guided by expected information gain, we avoid some of the combinatorial issues of other methods. We apply this approach to action potential time series from cortical networks and find that the activity of most neurons is largely determined by observing other cells. An important fraction of a neuron's entropy is contained in multiple cell arrangements of varying size. These findings agree with recent neuronal network reconstructions from binary correlations [8] and small multiplets [4], and also illuminate the contribution of higher order correlations in a new way. The identification of functional subgraphs in neuronal cultures is critical for designing experiments that promote computational tasks within neural networks, and should find applications generally in other complex systems. Additional information and examples can be found in [9]. We thank G. W. Gross for sharing his extensive expertise with neuronal cultures and J. Crutchfield, J. Gutfriend, and A. Hagberg for helpful discussions. This work is supported by LANL's LDRD project 20050411ER.

References

[1] W. Bialek, I. Nemenman, and N. Tishby, Physica A **302**, 89 (2001).

[2] J. P. Crutchfield and D. P. Feldman, Chaos **15**, 25 (2003).

[3] E. Schneidman, W. Bialek, and M. J. Berry II, J. Neurosci. **23**, 11539 (2003).

[4] L. M. A. Bettencourt, G. J. Stephens, M. I. Ham, and G. W. Gross, Phys. Rev. E **75**, 021915 (2007).

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[6] C. Haldeman and J. M. Beggs, Phys. Rev. Lett. **94**, 058101 (2005).

[7] G. W. Gross and F. U. Schwalm, J. Neurosci. Meth. **52**, 73 (1994).

[8] E. Schneidman, M. J. Berry II, R. Segev, and W. Bialek, Nature **440**, 1007 (2006).

[9] L. M. A. Bettencourt, V. Gintautas, and M. I. Ham, Phys. Rev. Lett. **100**, 238701 (2008).