Technical Report CSTN-051

Eigenvalue Spectra Measurements of Complex Networks

K.A. Hawick
Computer Science, Institute for Information and Mathematical Sciences,
Massey University, North Shore 102-904, Auckland, New Zealand
k.a.hawick@massey.ac.nz
Tel: +64 9 414 0800 Fax: +64 9 441 8181

Abstract

Complex network models such as Kauffman's NK model have been shown to have interesting phase transitional properties as the connectivity is varied. Conventional network and graph analysis metrics concerning path-lengths, numbers of components and even circuits have been used to explore these transitions. This paper presents use of matrix analysis techniques to relate the shape and form of the eigenvalue spectrum of the adjacency matrices of complex networks to the phase transition. A number of computational experiments with different network realisations yield results for the Kauffman NK model at connectivity $K = 1, 2, ..., 7$ are these are discussed in terms of the transition at $K = 2$.

Keywords: complex networks; Kauffman network; eigenvalue densities; semi-circle rule.

1 Introduction

Complex networks arise in many applications arenas including: the study of physical networks such as power distribution systems or the Internet; social networks between groups of people in society; and many physics and biological systems [1]. A particularly simple but interesting network model is the Kauffman NK model [2], originally proposed to explore the dynamics of gene regulatory networks. The Kauffman network has drawn most attention due to its dynamical properties but it also has interesting static structural properties and a study of its static network properties have recently been used to shed some light on the growth of attractor dynamics [3]

A computational science approach to complex networks involves generating many different sample networks from a suitable statistical distribution and running simulations on them. Various quantitative metrics can be used to identify emergent patterns in the behaviour of such systems and explain or at least locate particular behaviours like phase transitions. A fast and efficient network simulator [4] and associated graph metric measurements were implemented and have helped explore the nature of some of the scaling behaviour of the Kauffman fixed-K and mixed-K models [5].

This present paper presents an exploration of a matrix algebra analysis technique as it relates to the $K = 2$ phase transition in the Kauffman model. The density spectrum of eigenvalues of the network adjacency matrix is studied and is shown to change form markedly at the connectivity $K = 2$ phase transition.

Section 2 summarises the Kauffman complex network model and the adjacency matrix structures that arise from it, as well as the algorithm used to generate sample networks; section 3 gives a description of the eigenvalue density spectrum calculations; and section 4 contains a presentation of results for the averaged eigen-spectra of various NK networks. A discussion of these ideas and some conclusions are presented in section 5

2 Kauffman NK Networks

The Kauffman model [6] is known to have a phase transition at exactly $K = K_C \equiv 2$, above which value the system makes a transition from a stable regime to a chaotic one [7].

Kauffman's original model makes some simplifying approximations to the way genetic networks are thought to behave. The genome of an organism is represented by N genes, each of which is described by a binary variable specifying whether that gene is expressed or not. To model the effect of genes controlling the expression of other genes, Kauffman's model takes the genome to be a directed network in which a link from a given gene a to another gene

b, means that a controls the expression of b. Since real genetic networks may have a high degree of inhomogeneity, Kauffman's simplifying approximation in the NK model is that each gene is only controlled by precisely K other genes. Lacking further information it is also convenient to initialise the model network randomly and uniformly, so that the K inputs to gene a are chosen with equal probability from the entire system. Gene expression can be biased by choosing a subset of the 2^{2^K} possible boolean functions of K inputs. The probability p that each gene is expressed can thus also be adjusted.

For every value of p, there is a known [8] critical value of the connectivity K such that $K_c(p) = \frac{1}{2p(1-p)}$, and in the (stable) regime $K < K_c(p)$ all perturbations in an initial state of the network will (eventually) die out. Above the critical value, small perturbations propagate across the entire system and continue as periodic oscillations or attractors. This is the chaotic regime and the number of attractors grows very rapidly with network size N.

There are a number of ways to construct appropriate network realisations. The quenched NK model involves firstly wiring up N nodes with an appropriate topology determined by the desired connectivities K_{in} and K_{out}, then decorating nodes with a random initial bit state and a randomly chosen Boolean function of the correct number of inputs.

The simulation codes used in this work were implemented in the D programming language. This is a relatively recent systems-oriented language with some excellent optimizing compilers available [9].

The connectivity of the network itself can be stored in a number of ways. It is a directed graph and a convenient storage structure makes use of the dynamic D arrays. The `neighbours[i][n]` structure lists the n neighbours of the i'th node. D arrays have the '`length`' property which can be assigned a value to invoke an appropriate dynamic reallocation of memory (and copying of old values) as necessary. This structure is suited to fixed-K networks as discussed here, or those with a distribution of K values, added dynamically [5].

The algorithm for assigning the network neighbours array so that each node has exactly K inputs is shown in figure 1.

While this algorithm avoids self-arcs it does allow multi-arc neighbours – i.e. the same node can be picked as an input more than once.

This generation algorithm gives rise to the connectivity patterns as shown by the typical sample realisations in figure 2. As the figure illustrates, at low connectivities even small networks may be fragmented into islands of disconnected or unreachable components. Increasing connectivity gives

```
// generate an NK network from N and K
int [][] neighbours;
neighbours.length = N;
for(int i=0;i<N;i++){
  neighbours[i].length = K;
  for(int n=0;n<neighbours[i].length;n++){
    int choice;  // avoid self-arcs
    do{
      choice = randomInt() % N;
    }while( choice == i );
    neighbours[i][n] = choice;
  }
}
```

Figure 1: Network Generation Algorithm

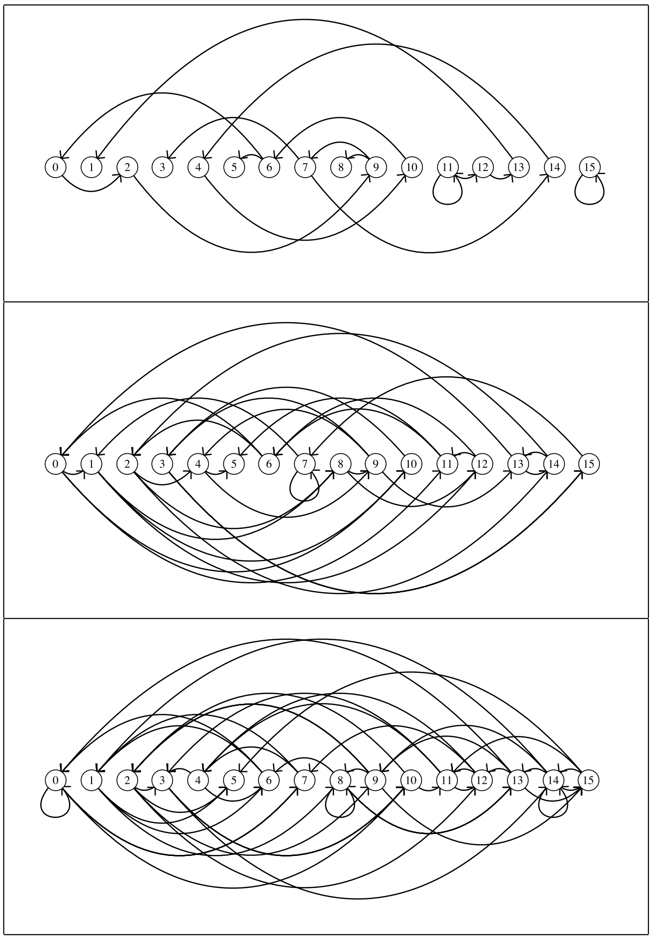

Figure 2: Typical 16-node networks (allowing self-arcs) generated for $K = 1, 2 3$ (top to bottom) with 16, 32 and 48 arcs respectively and 3, 73 and 966 circuits respectively (of longest lengths 8, 14 and 15)

rise to a reachability transition at $K = 1.5$ [5] and the well-studied Kauffman dynamical transition at $K = 2$. Circuits or loops can be studied to indicate how the internal path-lengths and reachabilities vary within the network. However circuit enumeration is a very expensive calculation to perform as part of a regular network analysis operation.

The work reported in this paper used network sizes of 200 or 500 nodes, the only practical limitation being the memory and computational limits of the eigen calculations.

3 Eigenvalue Calculations

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0	1	.	1	.	.	.	1	.	.	.	.	.	.	.	.	.	3
1	.	1	.	.	.	.	.	.	.	.	.	.	.	1	.	.	2
2	1	.	1	.	.	.	.	.	.	1	.	.	.	.	.	.	3
3	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	.	2
4	.	.	.	.	1	.	.	.	.	.	1	.	.	.	1	.	3
5	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.	2
6	1	.	.	.	.	1	1	.	.	.	1	.	.	.	.	.	4
7	.	.	.	1	.	.	.	1	.	1	.	.	.	.	1	.	4
8	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	2
9	.	.	1	.	.	.	.	1	1	1	.	.	.	.	.	.	4
1.	.	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.	3
11	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	2
12	.	.	.	.	.	.	.	.	.	.	.	1	1	1	.	.	3
13	.	1	.	.	.	.	.	.	.	.	.	.	1	1	.	.	3
14	.	.	.	.	1	.	.	1	.	.	.	.	.	.	1	.	3
15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1
																	===
																	44

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0	1	1	.	.	.	.	1	.	.	.	1	1	.	1	.	.	6
1	1	1	.	.	.	.	.	1	1	.	1	.	1	.	.	.	6
2	.	.	1	.	1	.	1	.	1	.	.	.	.	.	1	.	5
3	.	.	.	1	.	.	.	.	.	1	1	.	.	.	.	1	4
4	.	.	1	.	1	1	.	.	.	1	.	.	.	.	.	.	4
5	.	.	.	.	1	1	.	.	.	.	.	1	.	.	.	.	3
6	1	.	1	.	.	.	1	.	.	.	.	1	1	.	.	.	5
7	.	1	.	.	.	.	.	1	.	.	.	.	.	.	.	1	3
8	.	1	1	.	.	.	.	.	1	1	.	.	1	.	.	.	5
9	.	.	.	1	1	.	.	.	1	1	.	.	.	1	.	.	5
1.	1	1	.	1	.	.	.	.	.	.	1	.	.	.	.	.	4
11	1	.	.	.	.	1	1	.	.	.	.	1	1	.	.	.	5
12	.	1	.	.	.	.	1	.	1	.	.	1	1	.	.	.	5
13	1	.	.	.	.	.	.	.	.	1	.	.	.	1	1	.	4
14	.	.	1	.	.	.	.	.	.	.	.	.	.	1	1	.	3
15	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	1	3
																	===
																	70

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0	1	.	.	.	.	.	1	1	.	.	1	.	.	1	.	.	5
1	.	1	.	.	.	1	1	1	1	.	.	.	.	.	1	.	6
2	.	.	1	1	.	1	.	.	.	1	1	.	1	.	.	.	6
3	.	.	1	1	1	.	.	.	.	.	1	1	.	.	1	.	6
4	.	.	.	1	1	.	1	1	.	.	.	1	1	.	.	.	6
5	.	1	1	.	.	1	.	.	.	.	.	.	.	.	.	1	4
6	1	1	.	.	1	.	1	.	1	.	.	.	.	.	.	.	5
7	1	1	.	.	1	.	.	1	.	.	.	.	1	.	.	.	5
8	.	1	.	.	.	.	1	.	1	1	.	.	.	1	.	.	5
9	.	.	1	.	.	.	.	.	1	1	.	.	.	1	1	1	6
1.	1	.	1	1	.	.	.	.	.	.	1	1	.	.	.	.	5
11	.	.	.	1	1	.	.	.	.	.	1	1	1	.	.	1	6
12	.	.	1	.	1	.	.	1	.	.	.	1	1	1	.	.	6
13	1	.	.	.	.	.	.	.	1	1	.	.	1	1	1	1	7
14	.	1	.	1	.	.	.	.	.	1	.	.	.	1	1	1	6
15	.	.	.	.	.	1	.	.	.	1	.	1	.	1	1	1	6
																	===
																	90

Figure 3: Adjacency matrices for the sample networks shown in figure 2 of $K = 1, 2, 3$ (top to bottom). Row sums show the node-degrees

Although our network structures are held as neighbour lists, we can readily construct an adjacency matrix representation of the connectivities. This is a binary matrix where each ij'th element is unity if node i connects to node j, and zero otherwise.

The network models discussed here have directed arcs, so the matrix need not be symmetric, but it is of course wholly real. If this adjacency matrix is A, we are interested in the eigenvalues λ that correspond to the eigenvectors $\mathbf{e}$ in the prototypical eigen equation:

$$A\mathbf{e} = \lambda\mathbf{e} \tag{1}$$

This can be solved directly using classic algorithms such as the Jacobi rotation method [10] for dense matrices A. While generally one is not guaranteed sensible (real) eigen solutions for non-symmetric matrices, for the particular matrices generated by our network model one does obtain such solutions.

It is possible to build a histogram of the density of eigenvalues and this can be averaged over many network samples or realisations. This can be written for an N-node network as a sum of delta functions:

$$\rho(\lambda) = \frac{1}{N}\sum_{j=1}^{N}\delta(\lambda - \lambda_j) \tag{2}$$

Where λ_j is the j'th largest eigenvalue of the adjacency matrix. This summation will converge to a continuous function in the limit of large network sizes

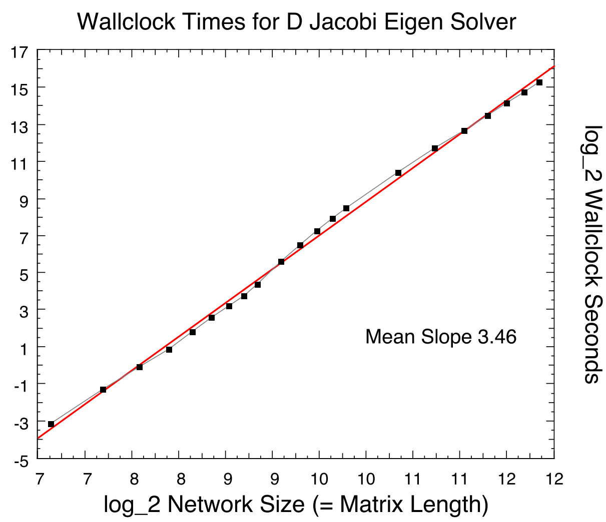

Figure 4: Plot of the empirically measured wall clock times to run the dense matrix Jacobi eigenvalue calculations. Error bars are approximately equal to the plot symbol sizes.

The Jacobi algorithm is well known and it was readily implemented as a space efficient code in D with versions using real and complex arithmetic.

As figure 4 shows from the $\log_2 - \log_2$ plot of wall-clock times (in seconds) for the dense matrix Jacobi eigen-solver the mean slope on log-log plot is approximately 3.46 which therefore implies an approximate power law dependency $t \approx N^{3.46}$. It can be seen there is a distortion in the curve at around $N = 600$ nodes.

This is likely due to a paging or memory system property of the MacPro 2.66GHz Intel Xeon workstation used for these measurements. It is possible to solve systems up to around $N = 4,600$ on this typical workstation, and this takes around 10 hours of wall-clock time. It requires a multi-processor system however to obtain useful multiple samples of this network size and the data reported in this paper are use the smaller more practical network size of $N = 200, 500$ nodes.

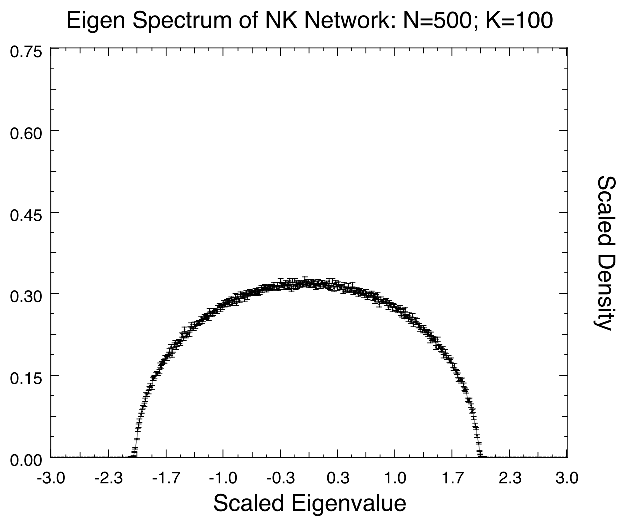

Figure 5: Scaled Eigenvalue density for an NK Network with $N = 500$ nodes and connectivity $K = 100$. Averages and error bars based on 1000 sample networks.

Farkas et al [11] have reported the use of eigen-spectral methods for small-world network graphs and in particular they make some observations about density patterns that go beyond the well-known Wigner law [12] for the semi-circular pattern that appears in the eigen-density. This manifests itself as

$$\rho(\lambda) = \frac{1}{2\pi\sigma^2}\sqrt{(4\sigma^2 - \lambda^2)}, |\lambda| < 2\sigma \qquad (3)$$

and zero otherwise, where 2σ is the Wigner radius of the semi-circle and can be empirically measured ($\sigma \approx 1$, here.)

Figure 5 shows how we can approach a large random network size that does exhibit the Wigner so-called semi-circle law by using a relatively large connectivity $K = 100$, with respect to the network size in this case of $N = 500$ nodes.

As the data shows with 1000 sample networks the semi-circle effect is obvious and quite smooth. The secondary peak corresponding to the largest eigenvalues is to the right (not shown for clarity of the semi-circle.)

It is generally convenient to scale both the eigenvalues and the density by $\sqrt{N}$. This allows us to compare different network sample sizes and in particular to maintain sensible density scales independent of the number of matrix connectivity entries.

4 Results

The density of eigenvalues $\rho(\lambda)$ can be measured empirically by histogramming values from many independently generated network realisations.

The data presented accrues from 1000 independently generated networks, with the scaled eigenvalues histogrammed into 121 bins over a range of ± 3. This gives a practical bin size of 0.05. The plotted data are based on average values and the error bars presented are simply based on standard deviations calculated over the 1000 sample networks.

Figure 6 shows the conventional approximately semi-circular spectrum of eigen values for "high-K" NK networks. A sharp peak superimposed on the semi-circle at the smallest (in magnitude) eigenvalue range appears due to the particular characteristics of the NK model compared to a random network or other model. Although this seems to be exacerbated by the small finite network size of 200 nodes, Farkas et al [11] also observe a deviant central peak in their eigen density that arises from non-random networks – in their case small-world networks.

The largest eigenvalues for the NK networks behave as described in [11] and manifest themselves as a small separate peak to the right of the main semi-circle structure. This peak shifts further to the right as K is increased, although its amplitude appears independent of K.

However a new and unexpected result arises for low connectivity networks. Figure 7 shows the spectrum lobes that arise at low-K values. At $K = 3$ slight dips appear at either side of the peak in the eigen-spectrum. At $K = 2$ the semi-circle has clearly split into two separate peaks, and with some severe lobes and other peaks superposed – apparently symmetrically about eigenvalues of zero. Lowering the network connectivity further to $K = 1$ the density curve has split into several smaller amplitude peaks. This clearly emphasises the phase transition in the network model that is known to occur at $K = 2$.

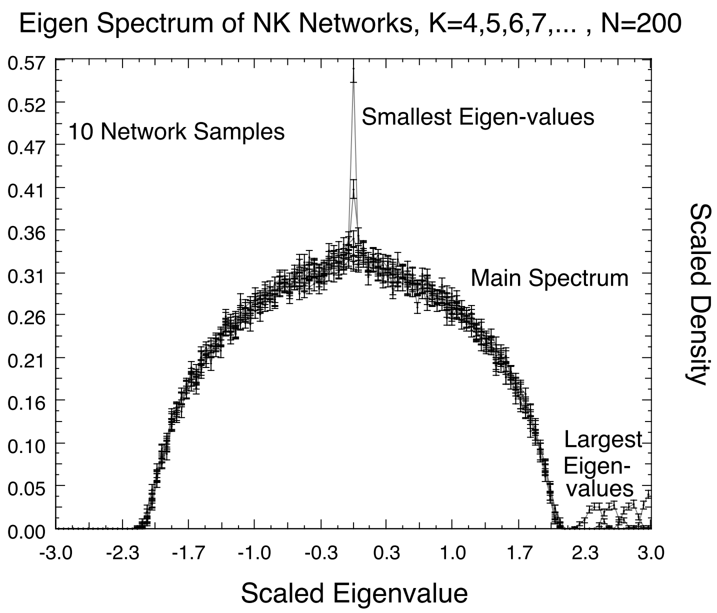

Figure 6: Scaled Eigen spectrum of NK Networks for $K = 4, 5, 6, 7, 8, 9, ...$

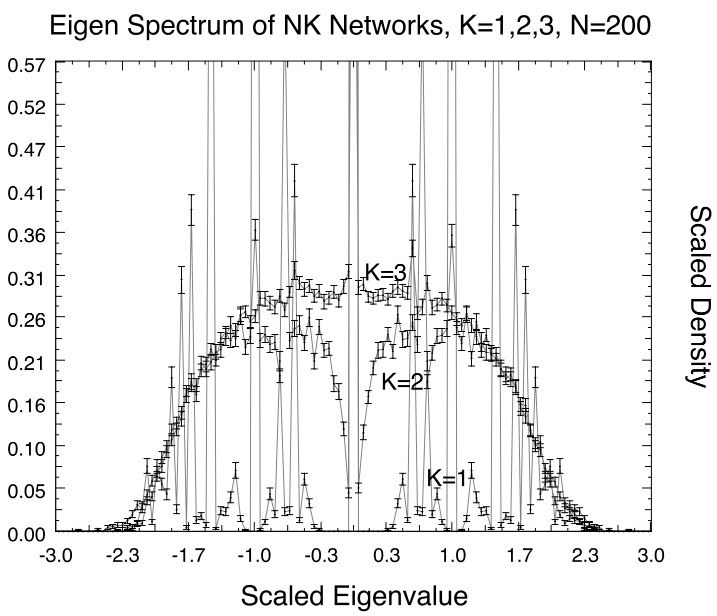

Figure 7: Scaled Eigen spectrum of NK Networks for $K = 1, 2, 3$

5 Discussion and Conclusions

This paper has shown the value of performing numerical simulation experiments with an associated set of quantitative measurements in studying phase transition effects in complex networks.

The results suggest that at low connectivities the eigenvalue density spectrum deviates from the Wigner semi-circle law in a systematic and interesting manner for certain complex networks. While this appeared to be initially due to the adjacency matrix no longer being a simple real symmetric structure, this does not necessarily appear to be the case and real eigenvalues are found. A further detailed study using a two dimensional complex eigenvalue histogramming technique is currently underway.

The multi-lobed patterns in the eigen density are systematic and show a dramatic change between the Kauffman network connectivities of $K = 1$ and $K = 3$. In another work [5] Hawick et al have investigated intermediate connectivities by using a mixed-K model. These spectral techniques may play an important part in further understanding the location and behaviour of the structural transitions found in that model.

As the timing data presented shows, the network size limitations in this work are mainly due to the computational cost of the dense matrix eigen-solver algorithm used. A sparse matrix eigenvalue approach such the thick-restart Lanczos methods studied by Wu et al [13] may allow allow exploration of very much larger systems. Sparse iterative approaches typically only find the eigenvalues in order and unfortunately may be ill-suited to our problem where we are most interested in the main spectrum.

The use of an eigen-spectral analysis technique with directed graphs appears to have some merits and holds some promise for the study of other complex network phenomena.

References

[1] Newman, M., Barabasi, A.L., Watts, D.J.: The Structure and Dynamics of Networks. Princeton University Press (2006)

[2] Kauffman, S.A.: Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of Theoretical Biology **22** (1969) 437–467

[3] Hawick, K., James, H., Scogings, C.: Structural Circuits and Attractors in Kauffman Networks. In Abbass, H.A., Randall, M., eds.: Proc. Third Australian Conference on Artificial Life. Volume 4828 of LNCS., Springer (2007) 189–200 978-3-540-76930-9.

[4] Hawick, K.A., James, H.A., Scogings, C.J.: Simulating large random boolean networks. Technical Report CSTN-039, Information and Mathematical Sciences, Massey University, Albany, North Shore 102-904, Auckland, New Zealand (2007)

[5] Hawick, K.A., James, H.A., Scogings, C.J.: Circuits, Attractors and Reachability in Mixed-K Kauffman Networks. Technical Report CSTN-046; arXiv:0711.2426, Massey University (2007)

[6] Kauffman, S.A.: The Origins of Order. Oxford University Press (1993)

[7] Kadanoff, L., Coppersmith, S., Aldana, M.: Boolean dynamics with random couplings. In Kaplan, E., Marsden, J., Sreenivasan, K., eds.: Perspectives and Problems in Nonlinear Science. Springer (2003)

[8] Derrida, B., Pomeau, Y.: Random networks of automata: A simple annealed approximation. Europhys. Lett. **1** (1986) 45–49

[9] Bright, W.: D Programming Language. Digital Mars. (2008)

[10] Wilkinson, J., Reinsch, C.: Linear Algebra, Volume II of Handbook for Automatic Computation. Springer (1971)

[11] Farkas, I.J., Derenyi, I., Barabasi, A.L., Vicsek, T.: Spectra of “real-world” graphs: Beyond the semi-circle law. Phys. Rev. E **64** (2001) 026704

[12] Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. **62** (1955) 548–564

[13] Wu, K., Canning, A., Simon, H.: Thick-restart Lanczos method for electronic structure calculations. Technical Report LBNL-42917, Lawrence Berkeley National Lab. (LBNL), USA (1999)