



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/136391/>

Version: Accepted Version

---

**Article:**

Ye, Cheng, Wilson, Richard Charles, Rossi, Luca et al. (Accepted: 2018) Thermodynamic Analysis of Time Evolving Networks. Entropy. ISSN: 1099-4300 (In Press)

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Thermodynamic Analysis of Time Evolving Networks

Cheng Ye<sup>1</sup>, Richard C. Wilson<sup>2</sup>, Luca Rossi<sup>3</sup>, Andrea Torsello<sup>4</sup>, and Edwin R. Hancock<sup>2,\*</sup> 

<sup>1</sup> Department of Computer Science, Royal Holloway, University of London, Egham, TW20 0EX, UK. 1; Cheng.Ye@rhul.ac.uk

<sup>2</sup> Department of Computer Science, University of York, York, YO10 5GH, UK. 2; Richard.Wilson@york.ac.uk

<sup>3</sup> School of Engineering and Applied Science, Aston University, Birmingham, B4 7ET, UK. 3;l.rossi@aston.ac.uk

<sup>4</sup> Dipartimento di Scienze Ambientali, Informatica, Statistica Università Ca' Foscari Venezia via Torino 155, 30172 Venezia Mestre, Italy. 4; Andrea.Torsello@unive.it

\* Correspondence: Edwin.Hancock@york.ac.uk; Tel.: +44 1904 325497

Version September 12, 2018 submitted to Entropy

1 The problem of how to represent networks, and from this representation derive succinct  
2 characterizations of network structure and in particular how this structure evolves with time, is  
3 of central importance in complex network analysis. This paper tackles the problem by proposing  
4 a thermodynamic framework to represent the structure of time-varying complex networks. More  
5 importantly, such a framework provides a powerful tool for better understanding the network time  
6 evolution. Specifically, the method uses a recently developed approximation of the network von  
7 Neumann entropy, and interprets it as the thermodynamic entropy for networks. With an appropriately  
8 defined internal energy to hand, the temperature between networks at consecutive time points can be  
9 readily derived, which is computed as the ratio of change of entropy and change in energy. It is critical  
10 to emphasize that one of the main advantages of the proposed method is that all these thermodynamic  
11 variables can be computed in terms of simple network statistics, such as network size and degree  
12 statistics. To demonstrate the usefulness of the thermodynamic framework, the paper uses real-world  
13 network data, which are extracted from time-evolving complex systems in the financial and biological  
14 domains. The experimental results successfully illustrate that critical events, including abrupt changes  
15 and distinct periods in the evolution of complex networks can be effectively characterized.

## 16 1. Introduction

17 There has been a vast amount of effort expended on the problems of how to represent networks,  
18 and from this representation derive succinct characterizations of network structure and in particular  
19 how this structure evolves with time [1–3]. Broadly speaking the representations and the resulting  
20 characterizations are goal-directed, and have centred around ways of capturing network substructure  
21 using clusters, or notions such as hubs and communities [4–7]. Here the underlying representations are  
22 based on the connectivity structure of the network, or statistics that capture the connectivity structure  
23 such as degree distributions [2,8,9]

24 A more principled approach is to try to characterize the properties of networks using ideas from  
25 statistical physics [10,11]. Here the network can be succinctly described using a partition function, and  
26 thermodynamic characterizations of the network such as entropy [12], total energy and temperature  
27 can be derived from the partition function [13–15]. Specifically, statistical thermodynamics can be  
28 combined with both graph theory and kinetics to provide a practical framework for handling highly  
29 structured and highly interactive time-evolving complex systems [13]. By using a random walk that  
30 maximizes the Ruelle-Bowens free-energy rate on weighted graphs, a novel centrality measure can  
31 be computed, and this has been successfully applied to both connected and disconnected large-scale  
32 networks [14]. Recently, it has been demonstrated that the subgraph centrality can be interpreted as a

33 partition function of a network [16], and as a result the entropy, internal energy and the Helmholtz  
34 free energy can be defined using spectral graph theory. The authors have also argued that the  
35 thermodynamic quantities are intimately related to the complex network dynamics. This approach  
36 combines the theoretical tools developed for studying graph spectra in the context of statistical  
37 mechanics of complex networks and clearly points out the potentials of the current approach to study  
38 real-world time-varying networks.

39 More recently, Minello et al. [17] have presented a quantum thermodynamic approach to study  
40 time-varying networks, in which the thermodynamic variables are developed through an unknown  
41 Hamiltonian operator governing the free evolution through the Schrödinger equation. Here, motivated  
42 by our recent work [18], we adopt a different theoretical foundation, namely the statistical mechanics,  
43 to establish our thermodynamic framework to analyze the time evolution of dynamic networks. We  
44 commence by studying undirected networks and define the network thermodynamic entropy, based  
45 on a recently developed approximation for the von Neumann entropy. We then focus on developing  
46 additional thermodynamic variables, i.e., internal energy and temperature for time-varying networks.  
47 We further show that our framework can be readily applied to directed networks, by taking into  
48 consideration the difference between the in- and out-degree of network nodes. We evaluate the  
49 usefulness of the proposed method using real-world time-varying complex system data from both  
50 financial and biological domains.

### 51 1.1. Related Literature

52 Although the bulk of existing network theory is concerned with static networks, most realistic  
53 networks are in reality dynamic in nature. Generally speaking, most existing methods for analyzing  
54 time evolution of complex networks have centered on studying structural measures of static networks,  
55 and then applying these quantities to each snapshot of the time-varying network in order to understand  
56 the evolutionary patterns. For instance, Holme et al. [19] have analyzed the time evolution of a number  
57 of well-known network features, including clustering coefficient, degree-degree correlations, average  
58 geodesic length and reciprocity of a large-scale online social network. Moreover, in [20], the authors  
59 have analyzed how the social networks of Flickr and Yahoo!360 evolve over different time periods using  
60 measures such as network density and average distance between nodes in the network components.  
61 Although such methods have proved to be efficient in reflecting the time evolution of some structural  
62 properties of evolving networks, they have a significant drawback, namely the lacking use of structure  
63 information between temporal networks at two consecutive time steps, e.g., the node degree change  
64 and edge number change.

65 In order to overcome this problem and to incorporate the missing structure information, a number  
66 of alternative techniques to capture the structure and evolution of networks have been proposed.  
67 For instance, Palla et al. [21] have developed a method for investigating the time dependence of the  
68 overlapping communities on a social network, using clique percolation method. Specifically, they take  
69 into consideration both the group size and age, and propose a measure for quantifying the relative  
70 overlap between two states of the same community at different time steps. Also they have developed  
71 a new network indicator called stationarity in order to quantify the changing rate of communities  
72 based on their size and age. In this way the authors have managed to exploit the community structure  
73 information between subsequent states of a time-evolving network. More recently, Peel and Clauset  
74 [22] have formalized the problem of identifying change points during network evolution within an  
75 online probabilistic learning framework, and have utilized generative network models and statistical  
76 hypothesis tests to solve it. This method has proved to successfully detect if, when and how change  
77 points occur in two high-resolution dynamic social networks.

78 Comparing to the existing evolving network analysis approaches, our thermodynamics analysis  
79 provides an advantageous approach in that the thermodynamic quantities, especially the temperature,  
80 fully exploit the information related to the structural changes of networks at subsequent time steps.  
81 More importantly, our approach does not require computationally complicated Bayesian probabilistic

frameworks such as the generalized hierarchical random graph (GHRG) model [22], but only uses a number of simple but important network characteristics, i.e., node degree statistics, edge number and degree information of some simple substructures such as triangles. This yields a low computational complexity to our thermodynamic analysis.

The structure of the remainder of the paper is as follows. Sec. II gives a detailed development of the thermodynamic framework for network evolution analysis. Sec. III tests the proposed method on a number of real-world time-varying networks, i.e., the New York Stock Exchange (NYSE) network and fruit fly life cycle gene expression network. Sec. IV summarizes the main contributions of this paper and also suggests a few research directions for the future.

## 2. Thermodynamic Framework for Time Evolving Complex Networks

In this section, we provide a detailed development of the thermodynamic framework for analyzing the time evolution of complex networks. In particular, the framework consists of three thermodynamic variables, namely the thermodynamic entropy, internal energy and temperature. Mathematically, the thermodynamic takes the same form as the network von Neumann entropy, when we associate the microscopic configurations of a network with the eigenstates of the normalized Laplacian spectrum. By defining an appropriate internal energy, the temperature is determined by measuring fluctuations in entropy and internal energy. We show that computationally, the framework is effective since each of these thermodynamic variables can be calculated using a few important graph statistics including number of nodes and edges and node degree statistics.

### 2.1. Initial Considerations

Let  $G(V, E)$  be an undirected graph with node set  $V$  and edge set  $E \subseteq V \times V$ . The adjacency matrix  $A$  of graph  $G(V, E)$  is defined as

$$A_{uv} = \begin{cases} 1 & \text{if } (u, v) \in E \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The degree of node  $u$  is  $d_u = \sum_{v \in V} A_{vu}$ . The normalized Laplacian matrix is  $\tilde{L} = D^{-1/2}(D - A)D^{-1/2}$  where  $D$  is the degree diagonal matrix whose elements are given by  $D_{uu} = d_u$  and zeros elsewhere. The elementwise expression of  $\tilde{L}$  is

$$\tilde{L}_{uv} = \begin{cases} 1 & \text{if } u = v \text{ and } d_u \neq 0 \\ -\frac{1}{\sqrt{d_u d_v}} & \text{if } u \neq v \text{ and } (u, v) \in E \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The normalized Laplacian matrix  $\tilde{L}$  and its spectrum yield a number of very useful graph invariants for a finite graph. For example, the eigenvalues for the graph normalized Laplacian are real numbers, bounded between 0 and 2 [23].

According to [24], the normalized Laplacian matrix  $\tilde{L}$  can be interpreted as the density matrix of an undirected graph. With this choice of density matrix, the von Neumann entropy of the undirected graph is defined as the Shannon entropy associated with the normalized Laplacian eigenvalues, i.e.,

$$H_{VN} = - \sum_{i=1}^{|V|} \frac{\tilde{\lambda}_i}{|V|} \ln \frac{\tilde{\lambda}_i}{|V|} \quad (3)$$

where  $\tilde{\lambda}_i, i = 1, \dots, |V|$ , are the eigenvalues of  $\tilde{L}$ .

In this paper, we aim at developing a thermodynamic characterization of network structure. We commence by assuming that at any instant in time a network  $G(V, E)$ , is statistically distributed across an ensemble of  $|V|$  microstates. The probability that the system occupies a microstate indexed  $s$  is given by  $p_s = \tilde{\lambda}_s / \sum_{s=1}^{|V|} \tilde{\lambda}_s$ , where  $\tilde{\lambda}_s, s = 1, 2, \dots, |V|$  are the eigenvalues of the normalized

118 Laplacian matrix of graph  $G$ . Noting that the trace of a matrix is the sum of its eigenvalues, we have  
 119  $\sum_{s=1}^{|V|} \tilde{\lambda}_s = \text{Tr}[\tilde{L}] = |V|$ , so the microstate occupation probability is simply  $p_s = \tilde{\lambda}_s / |V|$ .

We define the thermodynamic entropy of a network using the Shannon formula that is exclusively dependent on the probabilities of the microstates:

$$H_S = -k \sum_{s=1}^{|V|} p_s \ln p_s = -k \sum_{s=1}^{|V|} \frac{\tilde{\lambda}_s}{|V|} \ln \frac{\tilde{\lambda}_s}{|V|}, \quad (4)$$

120 where  $k$  is the Boltzmann constant and is set to be 1 to simplify matters.

121 It is clear that the thermodynamic entropy Eq. (4) and the von Neumann entropy Eq. (3) take the  
 122 same form. Both depend on the graph size and the eigenvalues of the normalized Laplacian matrix.  
 123 It is reasonable to suggest that the von Neumann entropy can be interpreted as the thermodynamic  
 124 entropy of a complex network.

## 125 2.2. Approximate Von Neumann Entropy for Undirected Graphs

In prior work [25], we have shown how the von Neumann entropy of an undirected graph Eq. (3) can be simplified by making use of the quadratic approximation (i.e.,  $-x \ln x \approx x(1-x)$ ),

$$H_Q = \sum_{i=1}^{|V|} \frac{\tilde{\lambda}_i}{|V|} \left(1 - \frac{\tilde{\lambda}_i}{|V|}\right). \quad (5)$$

For undirected graphs this quadratic approximation allows the von Neumann entropy to be expressed in terms of the trace of the normalized Laplacian and the trace of the squared normalized Laplacian, with the result that

$$H_{VN} = \frac{\text{Tr}[\tilde{L}]}{|V|} - \frac{\text{Tr}[\tilde{L}^2]}{|V|^2}. \quad (6)$$

The two traces appearing in the above expression are given in terms of node degree statistics [25], leading to

$$H_{VN} = 1 - \frac{1}{|V|} - \frac{1}{|V|^2} \sum_{(u,v) \in E} \frac{1}{d_u d_v}. \quad (7)$$

126 This formula contains two measures of graph structure, the first is the number of nodes of graph,  
 127 while the second is based on degree statistics for pairs of nodes connected by edges. Moreover, the  
 128 expression for the approximate entropy has computational complexity that is quadratic in graph size,  
 129 which is simpler than the original von Neumann entropy that is cubic since it requires enumeration of  
 130 the normalized Laplacian spectrum.

131 In order to obtain a better understanding of the entropic measure of graphs, it is interesting to  
 132 explore how the von Neumann entropy is bounded for graphs of a particular size, and in particular  
 133 which topologies give the maximum and minimum entropies. From Eq. (7) it is clear that when the  
 134 term under the summation is minimal, the von Neumann entropy reaches its maximal value. This  
 135 occurs when each pair of graph nodes is connected by an edge, and this means that the graph is  
 136 complete. On the other hand, when the summation takes on its maximal value, the von Neumann  
 137 entropy is minimum. This occurs when the structure is a string.

The maximum and minimum entropies corresponding to these cases are as follows. For a complete graph  $K_n$ , in which each node has degree  $n-1$ , it is straightforward to show that

$$H_{VN}(K_n) = 1 - \frac{1}{n} - \frac{1}{n^2} \cdot \frac{n(n-1)}{2(n-1)^2} = 1 - \frac{2n-1}{2n(n-1)}.$$

In the case of a string  $P_n$  ( $n \geq 3$ ), in which two terminal nodes have degree 1 while the remainder have degree 2, we have

$$H_{VN}(P_n) = 1 - \frac{1}{n} - \frac{1}{n^2} \cdot \frac{n+1}{4} = 1 - \frac{5n+1}{4n^2}.$$

As a result, the graph von Neumann entropy is bounded as follows:

$$1 - \frac{5|V|+1}{4|V|^2} \leq H_{VN}(G) \leq 1 - \frac{2|V|-1}{2|V|(|V|-1)}$$

138 where the lower boundary is obtained for strings, which are the simplest regular graph, and the upper  
139 bound is reached for complete graphs.

### 140 2.3. Internal Energy and Temperature

The internal energy of a network is defined as the mean value of the total energy, i.e., the sum of all microstate energies, each weighted by its occupation probability:

$$U = \sum_{s=1}^{|V|} p_s U_s, \quad (8)$$

141 where  $U_s$  is the energy of microstate  $s$ . Here we take the internal energy to be the total number of edges  
142 in the graph i.e.,  $U = |E|$ . From the properties of the Laplacian and normalized Laplacian matrices, we  
143 have  $|E| = \text{Tr}[L] = \text{Tr}[D^{1/2}\tilde{L}D^{1/2}] = \text{Tr}[D\tilde{L}]$ . This can be achieved if we set the microstate energies to  
144 be  $U_s = |V|d_s$ , i.e., proportional to the node degrees.

145 Suppose that the graphs  $G = (V, E)$  and  $G' = (V', E')$  represent the structure of a time-varying  
146 complex network at two consecutive epochs  $t$  and  $t'$  respectively. The reciprocal of the thermodynamic  
147 temperature  $T$  is the rate of change of entropy with internal energy, subject to the condition that  
148 the volume and number of particles are held constant, i.e.,  $1/T = dH_{VN}/dU$ . This definition can be  
149 applied to evolving complex networks which do not change size during their evolution.

#### 150 2.3.1. Undirected Edges

We approximate the change of the von Neumann entropy  $H_{VN}$  between undirected graphs  $G$  and  $G'$  as

$$dH_{VN} = H_{VN}(G') - H_{VN}(G) = \sum_{(u,v) \in E, E'} \frac{d_u \Delta_v + d_v \Delta_u + \Delta_u \Delta_v}{d_u(d_u + \Delta_u) d_v(d_v + \Delta_v)},$$

where  $\Delta_u$  is the change of degree of node  $u$ :  $\Delta_u = d'_u - d_u$ ;  $\Delta_v$  is similarly defined. The change in internal energy, is equal to the change in the total number of edges:  $dU = U(G') - U(G) = |E'| - |E| = \Delta|E|$ . Hence the reciprocal temperature  $T$  is:

$$\frac{1}{T(G, G')} = \sum_{(u,v) \in E, E'} \frac{d_u \Delta_v + d_v \Delta_u + \Delta_u \Delta_v}{\Delta|E| d_u(d_u + \Delta_u) d_v(d_v + \Delta_v)}. \quad (9)$$

When the changes in node degree are small compared to the node degree, i.e.,  $|\Delta_u| \ll d_u$ , then

$$\frac{1}{T(G, G')} = \sum_{(u,v) \in E, E'} \frac{d_u \Delta_v + d_v \Delta_u}{\Delta|E| d_u^2 d_v^2}. \quad (10)$$

151 The temperature measures fluctuations in the internal structure of the time evolving network, and  
152 depends on two properties of the network. The first is the overall or global change of the number of  
153 edges  $\Delta|E|$ , while the second property is a local one which measures the change in degree for pairs of  
154 nodes connected by edges, i.e.,  $d_u \Delta_v + d_v \Delta_u$ . Both quantities measure fluctuations in network structure,  
155 but at different levels of detail. The temperature is greatest when there are significant differences in

156 the global number of edges, and smallest when there are large local variations in edge structure which  
 157 do not result in an overall change in the number of edges.

158 Turning our attention in more detail to the term  $d_u \Delta_v + d_v \Delta_u$  appearing in the numerator of the  
 159 inverse temperature, it clearly measures the correlations between the degree of a node at one end of  
 160 an edge and the change in degree at the other. When the correlation is large, then the reciprocal of  
 161 the temperature is large, i.e., the temperature is low. On the other hand, low correlation corresponds  
 162 to high temperature. So at low temperature we can expect highly correlated changes in node degree,  
 163 while at high temperature these correlations are disrupted.

### 164 2.3.2. Directed Edges

165 We can extend this analysis to directed graphs. According to Ye et al. [26], the approximate von  
 166 Neumann entropy of a graph consisting entirely of directed edges, i.e., with no bidirectional edges is

$$H_D = 1 - \frac{1}{|V|} - \frac{1}{|V|^2} \sum_{(u,v) \in E_D} \frac{d_u^{in}}{d_u^{out}} \cdot \frac{1}{d_u^{out} d_v^{in}}. \quad (11)$$

167 where  $d_u^{in}$  is number of directed edges incident on node  $u$ , i.e., the in-degree of node  $u$ ,  $d_u^{out}$  is the  
 168 number of edge exiting node  $u$ , i.e., the out-degree of node  $u$  and  $E_D$  the directed edge set of graph  
 169  $G$ . The edge commences at node  $u$  and ends at node  $v$ . It should be noted that the out-degree of the  
 170 terminal node  $v$  does not participate in the expression for directed edge entropy. In terms of causality,  
 171 this means it is determined by the causal past but not the future of node  $v$ .

172 We can now repeat the incremental analysis for the directed version of the entropy. Considering  
 173 only terms of first order in the change in in-degree and out-degree, we find

$$dH_D = H_D(G') - H_D(G) = \sum_{(u,v) \in E_D} \frac{d_u^{out} d_v^{in} \Delta_u^{in} - 2d_u^{in} d_v^{in} \Delta_u^{out} - d_u^{in} d_u^{out} \Delta_v^{in}}{(d_u^{out})^3 (d_v^{in})^2} \quad (12)$$

174 where  $d_u^{in}$  is the in-degree at node  $u$ ,  $d_u^{out}$  the out-degree, and  $\Delta_u^{in}$  and  $\Delta_u^{out}$  the changes in in- and  
 175 out-degree of node  $u$ . Again the change in entropy takes the form of a correlation between the change  
 176 in the in-degree or out-degree of a node, and the product of the remaining two partial degrees. When  
 177 the entropy change is substituted into the expression for reciprocal temperature, we again find that  
 178 high correlation corresponds to low temperature.

### 179 2.4. Section Summary

180 In this section we have detailed the development of the thermodynamic framework for network  
 181 evolution analysis. In particular, we have employed three thermodynamic quantities, namely  
 182 thermodynamic entropy, internal energy and temperature, to characterize the structure of time-varying  
 183 complex networks. By analyzing how these quantities change over time, we are able to track the time  
 184 evolution of complex networks. It is also important to point out that one of the main advantages of  
 185 the proposed framework is that these thermodynamic variables can be simply computed using graph  
 186 statistics including graph size and node degree changes.

187 Another point worth noting is that when applying our approach to study real-world dynamic  
 188 systems, it is critical to take into consideration how the corresponding dynamic network is built and  
 189 particularly, how the links connecting nodes in the network are established. For instance, according to  
 190 Gorban et al. [27], given financial time series, the connections between financial entities can be assessed  
 191 by correlations between either two individuals or two time moments. The two different measures,  
 192 described as “varieties” and “volatilities”, respectively, have been shown to have different statistical  
 193 properties, e.g., the latter does not require averaging in time when calculating correlation coefficients  
 194 (locality), and thus could lead to different interpretations of our approach. A similar behaviour can be  
 195 observed in the process of cell fate decision as well [28].

### 196 3. Experiments and Evaluations

197 In this section we test the performance of the proposed thermodynamic framework by applying  
198 it to analyze the time evolution of realistic complex networks. In particular, we aim to apply  
199 the thermodynamic variables, i.e., the entropy, energy as well as temperature to a few real-world  
200 time-varying networks in order to explore whether abrupt changes in structure or different stages in  
201 network evolution can be efficiently characterized.

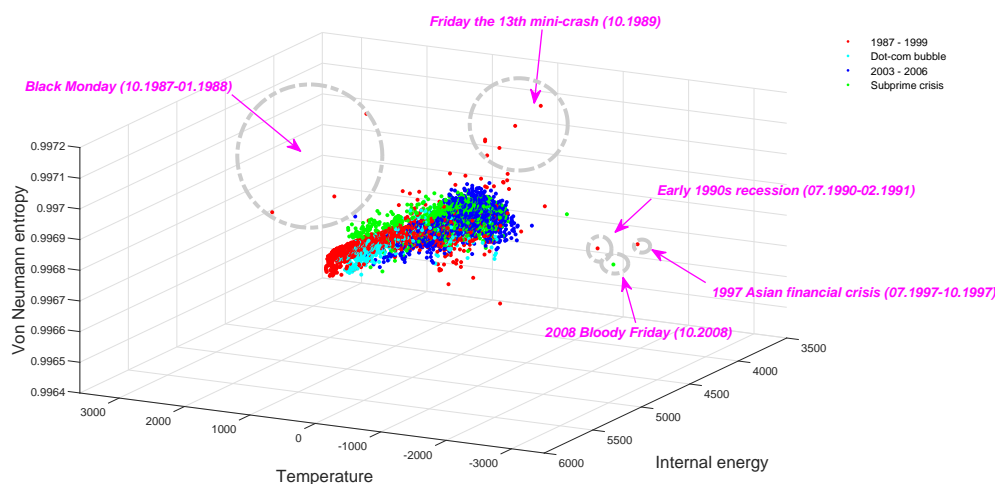
202 The data we will analyze in the experiments are summarized as follows.

203 *NYSE Stock Market Network Dataset.* Is extracted from a database consisting of the daily prices of  
204 3799 stocks traded on the New York Stock Exchange (NYSE). This data has been well analyzed in [29],  
205 which has provided an empirical investigation studying the role of communities in the structure of the  
206 inferred NYSE stock market. The authors have also defined a community-based model to represent  
207 the topological variations of the market during financial crises. Here we make use of a similar  
208 representation of the financial database. Specifically, we employ the correlation-based network to  
209 represent the structure of the stock market since many meaningful economic insights can be extracted  
210 from the stock correlation matrices [30–32]. To construct the dynamic network, 347 stocks that have  
211 historical data from January 1986 to February 2011 are selected [29,33]. Then, we use a time window of  
212 28 days and move this window along time to obtain a sequence (from day 29 to day 6004) in which  
213 each temporal window contains a time series of the daily return stock values over a 28-day period.  
214 We represent trades between different stocks as a network. For each time window, we compute the  
215 cross-correlation coefficients between the time series for each pair of stocks, and create connections  
216 between them if the maximum absolute value of the correlation coefficient is among the highest 5% of  
217 the total cross correlation coefficients. This yields a time-varying stock market network with a fixed  
218 number of 347 nodes and varying edge structure for each of 5976 trading days.

219  
220 *Drosophila Melanogaster Gene Network Dataset.* Is extracted from DNA microarrays that contain  
221 the transcriptional profiles for nearly one-third of all predicted fruit fly (*Drosophila melanogaster*)  
222 genes through the complete life cycle, from fertilization to adult. The data is sampled at 66 sequential  
223 developmental time points. The fruit fly life cycle is divided into four stages, namely the embryonic  
224 (samples 1-30), larval (samples 31-40) and pupal (samples 41-58) periods together with the first 30 days  
225 of adulthood (samples 59-66). Early embryos are sampled hourly and adults are sampled at multiday  
226 intervals according to the speed of the morphological changes. At each time point, by comparing each  
227 experimental sample to a reference pooled mRNA sample, the relative abundance of each transcript  
228 can be measured, which can further be used as a gene's expression level [34]. To represent this gene  
229 expression measurements data using a time-evolving network, the following steps are followed  
230 [35]. At each developmental point the 588 genes that are known to play an important role in the  
231 development of the *Drosophila* are selected. These genes are the nodes of the network. The edges  
232 are established based on the distribution of the gene expression values, which can be modeled as a  
233 binary pair-wise Markov Random Field (MRF) whose parameter indicates the strength of undirected  
234 interactions between two genes. In other words, two genes are connected when their model parameter  
235 exceeds a threshold. This dataset thus yields a time-evolving *Drosophila* gene-regulatory network  
236 with a fixed number of 588 nodes, sampled at 66 developmental time points.

#### 237 238 3.1. Thermodynamic Measures for Analyzing Network Evolution

239 To evaluate how well our thermodynamic characterization method can be used to analyze the  
240 time evolution of complex systems, we thoroughly study the dynamic networks in both datasets. In  
241 particular, given the network structure at each time step, we compute the thermodynamic entropy  
242 together with internal energy according to Eq. (7), Eq. (8), respectively. Also we compute the  
243 temperature between networks at consecutive time steps using Eq. (9). By investigating how these  
244 network thermodynamic variables evolve with time, it is interesting to see whether some critical events



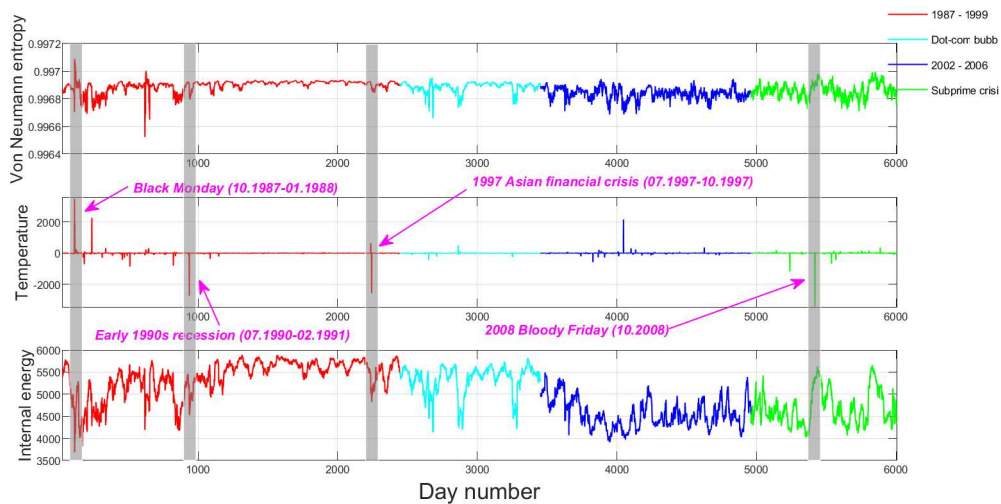
**Figure 1.** 3D scatter plot of the dynamic stock correlation network in the thermodynamic space. Red dots: 1987 - 1999 data; cyan dots: Dot-com bubble; blue dots: 2003 - 2006 background data; green dots: Subprime crisis.

245 can be detected in the network evolution. These include financial crises or crashes in the stock market,  
 246 and the essential morphological transformations that occur in the development of the *Drosophila*.

### 247 3.1.1. Financial Networks

248 In Fig. 1 we show a 3-dimensional scatter plot with each dimension representing a thermodynamic  
 249 variable for the time-evolving stock correlation network. Essentially, such a plot represents a  
 250 thermodynamic space spanned by entropy, internal energy and temperature. The most striking  
 251 feature here is that the thermodynamic distribution of the time-evolving financial network shows a  
 252 strong manifold structure with different phases of network evolution occupying different volumes of  
 253 the thermodynamic space. More interestingly, the outliers, which indicate significant global events  
 254 such as financial crises and stock market crashes, appear as peaks and troughs in the individual time  
 255 series (see Fig. 2). Examples include Black Monday, 1997 Asian Financial Crisis and the 24 October  
 256 2008 stock market crash (Bloody Friday). Another interesting observation in Fig. 1 is that the Dot-com  
 257 bubble period (approximately from 1999 to 2002), which is represented by cyan dots, is separated from  
 258 the background data points and occupies a distinct region in the 3-dimensional thermodynamic space.  
 259 Theoretically, this is due to the fact that during the Dot-com bubble period, a significant number of  
 260 Internet-based companies were founded, leading to a rapid increase of both stock prices and market  
 261 confidence. This considerably changed both the inter-relationships between stocks and the resulting  
 262 structure of the entire market.

263 To explore how our approach (especially the approximate von Neumann entropy) is compared  
 264 with existing graph characteristics in terms of revealing the network structural evolution across  
 265 different phases, we pause here to investigate two well-known measurements for networks, namely  
 266 1) degree assortativity coefficient, originally developed by Newman [36], and 2) Estrada index [37].  
 267 Theoretically speaking, the main difference between the von Neumann entropy and the degree  
 268 assortativity lies in that, the former quantifies the network structural complexity, i.e., how far a given  
 269 network deviates from a regular one, whereas the latter estimates the preference of nodes with different  
 270 degrees being connected, although both their mathematical expressions contain the product of degree  
 271 of nodes that are linked in the network. On the other hand, the difference between the Estrada index  
 272 and the von Neumann entropy is that the Estrada index exploits the spectrum of the network adjacency  
 273 matrix instead of the Laplacian. The individual time series of the two measurements are reported in Fig.  
 274 3. Clearly, both plots show significant fluctuations over the entire time period. Although critical events

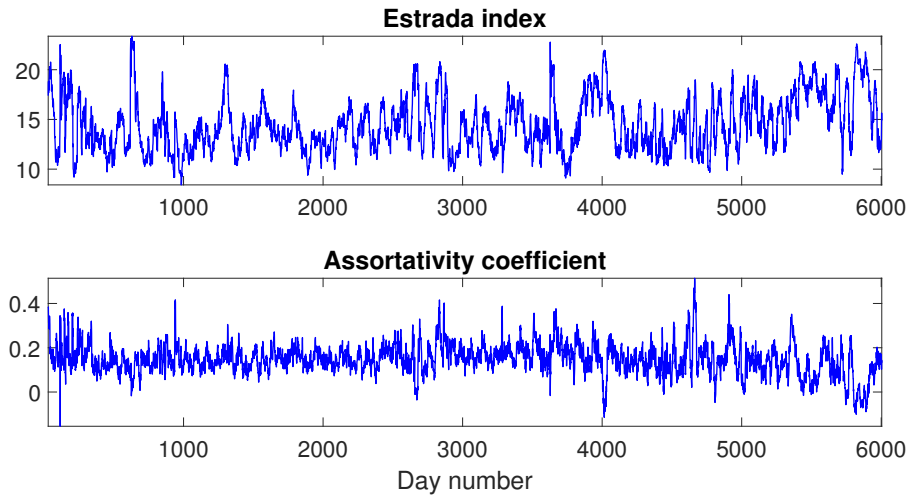


**Figure 2.** Top to bottom: a) The von Neumann entropy versus time for the dynamic stock correlation network. b) The temperature versus time for the dynamic stock correlation network. c) The internal energy versus time for the dynamic stock correlation network.

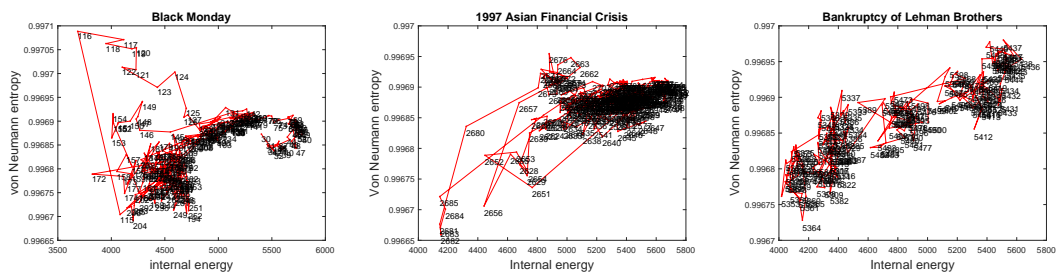
275 such as Black Monday and 1997 Asian financial crisis appear to be peaks and troughs in the figure, they  
 276 cannot be easily distinguished from a large number of other fluctuations. Moreover, compared with  
 277 Fig. 2, the time periods in which the network structure remains relatively stable, cannot be identified as  
 278 both time series display continuous fluctuations. These interesting observations together suggest that,  
 279 by viewing critical event networks as the outliers that deviate far from the regular network, which  
 280 corresponds to the stable phase in the time evolution, our thermodynamic framework turns out to be a  
 281 more appropriate option for analyzing the structural changes of dynamic networks. This is because  
 282 the von Neuman entropy can measure the distance between a given network and a regular one, which  
 283 cannot be readily estimated by other existing network characteristics.

284 We now study three financial crises in detail, and explore how the thermodynamic variables can  
 285 be used to unravel how the stock market network structure changes with time. In Fig. 4 we show the  
 286 trace of the stock network on the entropy-energy plane during Black Monday (left panel), the Asian  
 287 Financial Crisis (Middle panel) and the Lehman Brothers bankruptcy (right panel) respectively. The  
 288 number beside each data point represents the day number in the time series. From the figure, before  
 289 Black Monday, the network structure remains relatively stable, neither the network entropy nor the  
 290 internal energy changes significantly. However, when Black Monday takes place (day 115 and 116 in  
 291 the time series), the network undergoes a considerable change in structure since the entropy increases  
 292 dramatically. Then, the network entropy slowly decreases after the stock market crash, which implies  
 293 that the stock correlation network gradually returns to its normal state (before crisis). A similar pattern  
 294 can be observed concerning the 1997 Asian Financial Crisis which is shown in the middle panel as well.  
 295 In short, the stock market undergoes a significant crash in which the network structure undergoes a  
 296 significant change, as signalled by a large drop in network entropy. The crash is followed by a slow  
 297 recovery. It is interesting to note that for the Lehman Brothers bankruptcy case, as the time series  
 298 evolves, both the network entropy and the internal energy continue to grow gradually, which yields  
 299 a very different pattern as compared to previous cases. So the difference in the network structure  
 300 behaviour during different financial crises implies that our thermodynamic representation can be used  
 301 to both characterize and distinguish between different critical events in the network evolution.

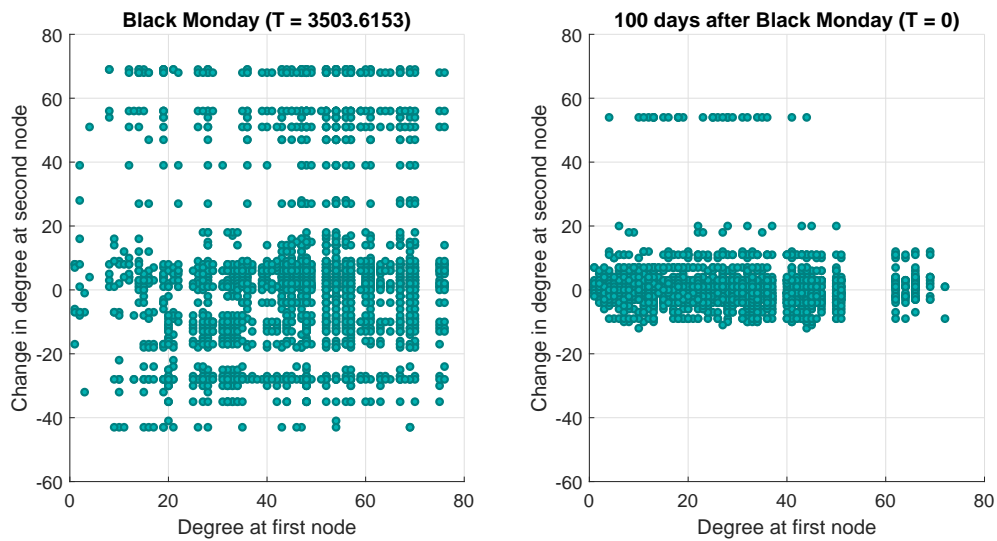
302 Next we particularly concentrate the temperature variable, which measures the structural  
 303 difference of networks at consecutive time steps. From the definition of temperature Eq. (10), clearly  
 304 the temperature depends on changes in node degree. Mathematically, for an undirected graph the  
 305 reciprocal of the temperature is determined by the quantity  $d_u \Delta_v + d_v \Delta_u$ . So in order to investigate



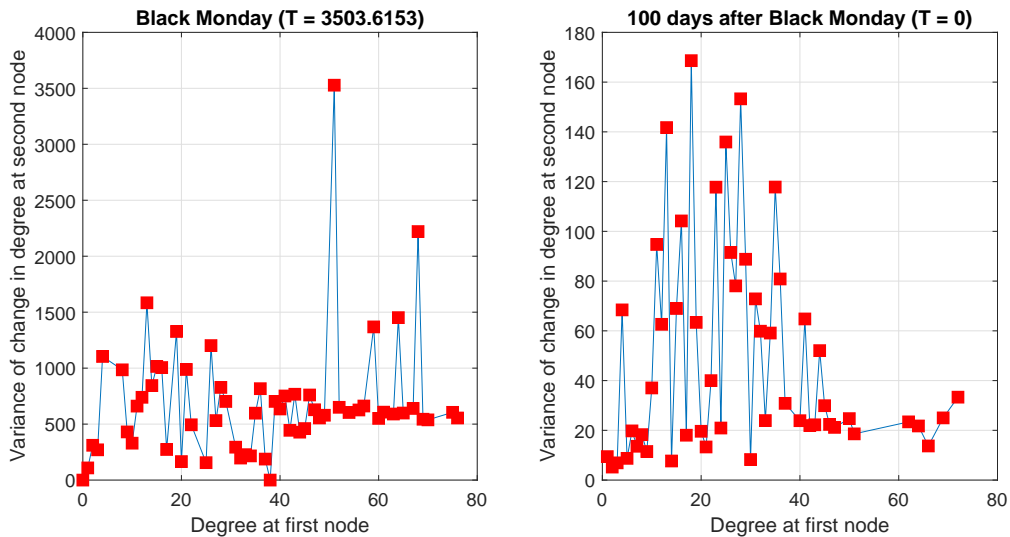
**Figure 3.** Top to bottom: a) The Estrada index versus time for the dynamic stock correlation network. b) The assortativity coefficient versus time for the dynamic stock correlation network.



**Figure 4.** Trace of the time-evolving stock correlation network in the entropy-energy plane during financial crises (the number beside data point is the day number in the time series). Left: Black Monday (from day 30 to 300); middle: Asian Financial Crisis (from day 2500 to 2800); right: Bankruptcy of Lehman Brothers (from day 5300 to 5500).

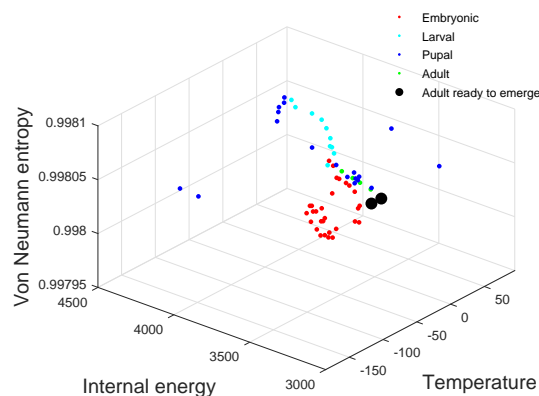


**Figure 5.** Scatter plots of  $\Delta_v$  versus  $d_u$  for high and low temperature networks.



**Figure 6.** Scatter plots of variance of  $\Delta_v$  versus  $d_u$  for high and low temperature networks.

306 the correlation between node degree and node degree change, we show a scatter plot of  $\Delta_v$  versus  $d_u$   
 307 for nodes  $u$  and  $v$  connected by an edge in Fig. 5. We consider two pairs of consecutive networks,  
 308 respectively: the first contains networks in the proximity of the Black Monday epoch (left panel)  
 309 whereas the second consists of networks far away from it (right panel). The main difference between  
 310 the two subplots lies in that for the case of the Black Monday networks, there is no correlation between  
 311  $d_u$  and  $\Delta_v$ , while in the case of the second pair there is a regression line of approximately zero slope. The  
 312 temperature between networks in the former pair is particularly high, whereas the latter corresponds  
 313 to a very low temperature. Another feature to note from the two plots is that for a given degree the  
 314 variance of the degree changes is greatest at high temperature. To illustrate this point, Fig. 6 shows the  
 315 variance of the degree change as a function of degree. In the case of the Black Monday networks the  
 316 variance is much larger than in the case of the second network pair, far away from it.



**Figure 7.** 3D scatter plot of the *Drosophila melanogaster* gene regulatory network in the thermodynamic space. Red dots: embryonic period; cyan dots: larval period; blue dots: pupal period; green dots: adulthood; black dot: adult ready to emerge.

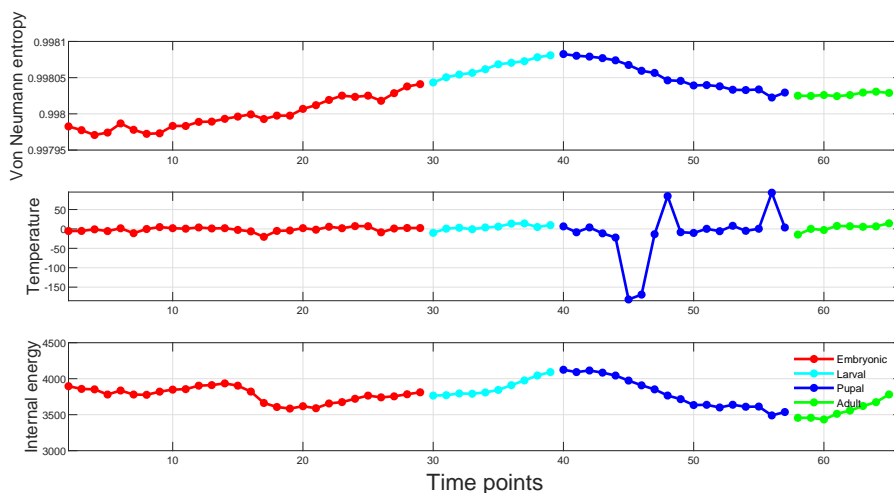
### 317 3.1.2. Gene Regulatory Network

318 We now apply the thermodynamic framework to the fruit fly network, i.e., the *Drosophila* gene  
 319 regulatory network in the second dataset. Similar to the experiments performed on the financial data,  
 320 we again show the 3-dimensional scatter plot of the thermodynamic variables of the time-varying  
 321 network in the thermodynamic space (Fig. 7), together with the entropy, energy and temperature times  
 322 series (Fig. 8). The four developmental stages are shown in different colours. Some key observations  
 323 can be made. First, the different stages of evolution are easily distinguished by the thermodynamic  
 324 variables. For instance from Fig. 8, due to the early development of an embryo, the red curve  
 325 (embryonic period) shows some fluctuations. This is attributable to strong and rapidly changing gene  
 326 interactions, because of the need for rapid development. Secondly, in Fig. 7, the pupal stage data  
 327 points are relatively sparsely distributed in the thermodynamic space. This is attributable to the fact  
 328 that during this period, the pupa undergoes a number of significant pupal-adult transformations.  
 329 Moreover, as the organism evolves into an adult, the gene interactions which control its growth begin  
 330 to slow down. Hence the green points (adulthood) remain stable. Finally, the black data points are  
 331 well separated from the remainder of the developmental samples, and correspond to the time when  
 332 the adult emerges.

333 To summarize, in this section we have implemented computational experiments on two realistic  
 334 time-evolving complex systems extracted from financial and biological domains, respectively. For the  
 335 stock market data, we have particularly analyzed a few well-known stock market crashes and have  
 336 demonstrated that the thermodynamic entropy, internal energy together with temperature provide a  
 337 powerful tool for detecting abrupt events and characterizing different stages in the network evolution.  
 338 The same conclusion can also be drawn based on the results of the fruit fly life cycle network analysis.

## 339 4. Conclusions

340 It is of fundamental importance to have methods to hand to characterize and understand the  
 341 time evolution of time-varying complex systems. To tackle this particular problem, in this paper we  
 342 have developed a few global variables for networks, namely the thermodynamic entropy, internal  
 343 energy and temperature, and have united them as a whole to analyze the structural properties of  
 344 time-evolving networks. In other words, we have adopted a thermodynamic framework to visualise  
 345 and understand the network evolution. Specifically, based on statistical thermodynamics, this method  
 346 starts with a recently derived expression for the von Neumann entropy of a network. The method then  
 347 connects the microscopic configurations of a network with the normalized Laplacian eigenstates. In  
 348 this way we have shown that the von Neumann entropy can be interpreted as the thermodynamic



**Figure 8.** Top to bottom: a) The von Neumann entropy versus time for the *Drosophila melanogaster* gene regulatory network. b) The temperature versus time for the *Drosophila melanogaster* gene regulatory network. c) The internal energy versus time for the *Drosophila melanogaster* gene regulatory network.

349 entropy of a network. The method further defines the network internal energy, which is determined  
 350 by the number of edges in the network. Finally, the thermodynamic temperature is a measure that  
 351 gauges the structural fluctuations between networks at consecutive time points, via changes in the  
 352 number of edges and individual node degree changes.

353 To demonstrate that the proposed framework serves as a powerful tool for detecting critical  
 354 events and distinct periods in the time evolution of real-world complex systems, we have evaluated the  
 355 method experimentally using data taken from the financial and biological domains. The experimental  
 356 results have confirmed that the thermodynamic variables together provide an efficient framework for  
 357 analyzing the evolutionary properties of dynamic networks.

358 In the future, in order to improve the thermodynamic characterizations so that they can become  
 359 more effective in identifying critical events and significant time stages in the evolution of time-varying  
 360 networks, we could turn our attention to the quantum physics. In particular, we would be interested in  
 361 exploring whether partition functions from different quantum statistics, such as Bose-Einstein partition  
 362 function and Fermi-Dirac partition function, can be used for the purpose of providing a more efficient  
 363 way to probe dynamic network structure.

## 364 5. Patents

365 **Author Contributions:** Conceptualization, Edwin Hancock; Formal analysis, Cheng Ye; Investigation, Cheng Ye;  
 366 Methodology, Cheng Ye; Supervision, Edwin Hancock; Visualization, Cheng Ye; Writing – original draft, Cheng  
 367 Ye; Writing – review & editing, Richard Wilson, Luca Rossi, Andrea Torsello and Edwin Hancock.

368 **Conflicts of Interest:** The authors declare no conflict of interest.

369

- 370 1. Hofstad, R.v.d. *Random Graphs and Complex Networks.*; Eindhoven University of Technology, 2010.
- 371 2. Anand, K.; Bianconi, G. Entropy measures for networks: Toward an information theory of complex  
 372 topologies. *Physical Review E* **2009**, *80*.
- 373 3. Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Reviews of Modern Physics* **2002**,  
 374 *74*, 47–97.
- 375 4. Newman, M. The Structure and Function of Complex Networks. *SIAM Review* **2003**, *45*, 167–256.
- 376 5. Estrada, E. *The Structure of Complex Networks: Theory and Applications.*; Oxford University Press, 2011.

- 377 6. Feldman, D.; Crutchfield, J. Measures of statistical complexity: Why? *Physics Letters A* **1998**, *238*, 244–252.
- 378 7. Dehmer, M.; Mowshowitz, A.; Emmert-Streib, F. *Advances in Network Complexity*; Wiley-Blackwell, 2013.
- 379 8. Anand, K.; Bianconi, G.; Severini, S. Shannon and von Neumann entropy of random networks with  
380 heterogeneous expected degree. *Physical Review E* **2011**, *83*.
- 381 9. Anand, K.; Krioukov, D.; Bianconi, G. Entropy distribution and condensation in random networks with a  
382 given degree distribution. *Physical Review E* **2014**, *89*.
- 383 10. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Reviews of Modern Physics*  
384 **2009**, *81*, 591–646.
- 385 11. Mantegna, R.N.; Stanley, H.E. *Introduction to Econophysics: Correlations and Complexity in Finance*; Cambridge  
386 University Press, 1999.
- 387 12. Bianconi, G. The entropy of randomized network ensembles. *Europhysics Letters* **2008**, *81*.
- 388 13. Mikulecky, D.C. Network thermodynamics and complexity: a transition to relational systems theory.  
389 *Computers & Chemistry* **2001**, *25*, 369–391.
- 390 14. Delvenne, J.C.; Libert, A.S. Centrality measures and thermodynamic formalism for complex networks.  
391 *Physical Review E* **2011**, *83*.
- 392 15. Fronczak, A.; Fronczak, P.; Holyst, J.A. Thermodynamic forces, flows, and Onsager coefficients in complex  
393 networks. *Physical Review E* **2007**, *76*.
- 394 16. Estrada, E.; Hatano, N. Statistical-mechanical approach to subgraph centrality in complex networks.  
395 *Chemical Physics Letters* **2007**, *439*, 247–251.
- 396 17. Minello, G.; Torsello, A.; Hancock, E.R. Quantum thermodynamics of time evolving networks. 2016 23rd  
397 International Conference on Pattern Recognition (ICPR), 2016, pp. 1536–1541.
- 398 18. Ye, C.; Torsello, A.; Wilson, R.C.; Hancock, E.R. Thermodynamics of Time Evolving Networks. Graph-Based  
399 Representations in Pattern Recognition; Liu, C.L.; Luo, B.; Kropatsch, W.G.; Cheng, J., Eds.; Springer  
400 International Publishing: Cham, 2015; pp. 315–324.
- 401 19. Holme, P.; Edling, C.R.; Liljeros, F. Structure and time evolution of an Internet dating community. *Social  
402 Networks* **2004**, *26*, 155–174.
- 403 20. Kumar, R.; Novak, J.; Tomkins, A. Structure and evolution of online social networks. *Proceedings of the  
404 Twelfth ACM SIGKDD international conference on Knowledge discovery and data mining* **2006**, pp. 611–617.
- 405 21. Palla, G.; Barabási, A.L.; Vicsek, T. Quantifying social group evolution. *Nature* **2007**, *446*, 664–667.
- 406 22. Peel, L.; Clauset, A. Detecting change points in the large-scale structure of evolving networks. AAAI, 2015.
- 407 23. Chung, F. *Spectral Graph Theory*; American Mathematical Society, 1997.
- 408 24. Passerini, F.; Severini, S. The Von Neumann Entropy of Networks. *International Journal of Agent Technologies  
409 and Systems* **2008**, pp. 58–67.
- 410 25. Han, L.; Escolano, F.; Hancock, E.R.; Wilson, R.C. Graph Characterizations from Von Neumann Entropy.  
411 *Pattern Recognition Letters* **2012**, *33*, 1958–1967.
- 412 26. Ye, C.; Wilson, R.C.; Comin, C.H.; Costa, L.d.F.; Hancock, E.R. Approximate von Neumann entropy for  
413 directed graphs. *Physical Review E* **2014**, *89*.
- 414 27. Gorban, A.N.; Smirnova, E.V.; Tyukina, T.A. Correlations, risk and crisis: From physiology to finance.  
415 *Physica A: Statistical Mechanics and its Applications* **2010**, *389*, 3193 – 3217.
- 416 28. Mojtahedi, M.; Skupin, A.; Zhou, J.; Castaño, I.G.; Leong-Quong, R.Y.Y.; Chang, H.; Trachana, K.; Giuliani,  
417 A.; Huang, S. Cell Fate Decision as High-Dimensional Critical State Transition. *PLOS Biology* **2016**, *14*, 1–28.
- 418 29. Silva, F.N.; Comin, C.H.; Peron, T.K.D.; Rodrigues, F.A.; Ye, C.; Wilson, R.C.; Hancock, E.R.; Costa, L.d.F.  
419 On the Modular Dynamics of Financial Market Networks. *ArXiv e-prints* **2015**.
- 420 30. Battiston, S.; Caldarelli, G. Systemic Risk in Financial Networks. *Journal of Financial Managements Markets  
421 and Institutions* **2013**, *1*, 129–154.
- 422 31. Bonanno, G.; Caldarelli, G.; Lillo, F.; Miccichè, S.; Vandewalle, N.; Mantegna, R.N. Networks of equities in  
423 financial markets. *European Physical Journal B* **2004**, *38*, 363–372.
- 424 32. Caldarelli, G.; Battiston, S.; Garlaschelli, D.; Catanzaro, M. Emergence of Complexity in Financial Networks.  
425 *Lecture Notes in Physics* **2004**, *650*, 399–423.
- 426 33. Peron, T.K.D.; Rodrigues, F.A. Collective behavior in financial markets. *Europhysics Letters* **2011**, *96*.
- 427 34. Arbeitman, M.N.; Furlong, E.E.; Imam, F.; Johnson, E.; Null, B.H.; Baker, B.S.; Krasnow, M.A.; Scott, M.P.;  
428 Davis, R.W.; White, K.P. Gene expression during the life cycle of *Drosophila melanogaster*. *Science* **2002**,  
429 *297*, 2270–2275.

- 4.30 35. Song, L.; Kolar, M.; Xing, E.P. KELLER: estimating time-varying interactions between genes. *Bioinformatics*  
4.31 **2009**, *25*, 128–136.
- 4.32 36. Newman, M. Assortative mixing in networks. *Physical Review Letters* **2002**, *89*.
- 4.33 37. Estrada, E. Characterization of 3D molecular structure. *Chemical Physics Letters* **2000**, *319*, 713 – 718.

4.34 © 2018 by the authors. Submitted to *Entropy* for possible open access publication under the terms and conditions  
4.35 of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).