White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/174008/

Version: Accepted Version

Article:
Wang, Jianjia, Wilson, Richard Charles and Hancock, Edwin R (2021) Network Edge Entropy Decomposition with Spin Statistics. Pattern Recognition. 108040. ISSN: 0031-3203

https://doi.org/10.1016/j.patcog.2021.108040

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

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 including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Manuscript(Changes marked) Click here to view linked References

Network Edge Entropy Decomposition with Spin Statistics

Jianjia Wang[a,b], Richard C. Wilson[c], Edwin R. Hancock[c]

[a]*School of Computer Engineering and Science, Shanghai University, Shanghai, China*
[b]*Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai, P.R.China, 200444*
[c]*Department of Computer Science, University of York, York, YO10 5DD, UK*

Abstract

In a previous study, we have explored how to decompose the global entropy of a network into edge components using a graph-spectral decomposition technique. Here, we develop this work in more depth to understand the role of edge entropy as an efficient and effective tool in analysing network structure. We use the edge entropy distribution as a network feature or characterisation and combine it with linear discriminant analysis to distinguish different types of network model and structure. Interpreting the normalised Laplacian matrix as the network Hamiltonian (or energy) operator, the network is assumed to be in thermodynamic equilibrium with a heat bath where the energy states correspond to the normalised Laplacian eigenvalues. To model the way in which particles occupy the energy states, we explore the use of three different spin-dependent statistical models to determine the thermodynamic entropy of the network. These are a) the classical spinless Maxwell-Boltzmann distribution, and two models based on quantum mechanical spin-statistics, namely b) the Bose-Einstein model for particles with integer spin, and c) the Fermi-Dirac model for particles with half-integer spin. By using the spectral decomposition of the Laplacian, we illustrate how to project out the edge-entropy components from the global network entropy. In this way, the detailed distribution of entropy across the edges of a

*Corresponding author:
Email address: jianjiawang@shu.edu.cn (Jianjia Wang)

network can be constructed. Compared to our previous study of the von Neumann edge entropy, where the edge entropy just depends on the degrees of the nodes forming an edge, in the case of the new statistical mechanical model, there is a more subtle dependence of the edge entropy on the structure of a network. We illustrate how this new edge entropy distribution can be used to more effectively identify variations in network structure, in particular for edges incorporating nodes of large degree. Numerical experiments on synthetic and real-world data-sets are presented to evaluate the qualitative and quantitative differences in performance.

Keywords: Network Edge Entropy, Spin Statistics, Partition Function

1. Introduction

Recently, network entropy has attracted increased attention because of its capacity to distinguish the structural properties of different types of networks [1, 2, 3]. Different varieties of entropy have been extensively used to characterise the salient features of networks, not only in the static domain but also in the domain of time varying or dynamic networks, such as the biological, social and financial networks [4, 5, 6]. One of the most sophisticated studies involves the von Neumann entropy, which has been successfully used as an effective characterisation to describe the structural properties of random, small-world and scale-free networks [4, 7]. The von Neumann entropy derives from a quantum mechanical analogue where the network Laplacian matrix plays the role of the density matrix in quantum physics [8].

In quantum mechanics, the density matrix describes the statistical state, potentially either pure or mixed, of a system. The outcome probability for any well-defined measurement upon this system can be calculated from the density matrix [9]. The density matrices for non-pure states are mixed states, and can be represented as a convex combination of pure states. As a result, the density matrices are helpful for dealing with statistical ensembles of different possible preparations of a quantum system [10]. Describing a quantum state

by its density matrix is a fully general alternative formalism to describing a quantum state by its state vector or a mixture of state vectors. However, it is the most convenient for calculations involving mixed states. Such states arise when the observer does not know how the states of the system are excited, as in the case of a system in thermal equilibrium at a temperature above absolute zero. The density matrix is self-adjoint (or Hermitian), positive semi-definite and of trace one [11]. In simple terms, the desnity matrix is a complex valued representation of a system being in a state or mixture of states, and provides a means of calculating measurable quantities associated with this system.

Here, in our network description when the density matrix is related to the Laplacian matrix (i.e. the diagonal degree matrix minus the adjacency matrix), it provides a connection between the network and quantum domains which allows the von Neumann entropy to be computed using just the degrees of pairs of nodes linked by edges in a network [4, 5]. The eigenvalues of the density matrix reflect the energy states of a system, and the analogy provides a novel way for mapping the heat bath in statistical physics to network structure [12]. According to the heat bath analogy at a given temperature the energy states are populated by particles which are in thermal equilibrium with the heat bath. Such a method provides a convenient route for characterising network entropy. When the energy states are populated with particles in thermal equilibrium with the heat bath, it allows us to calculate the distribution of energy and entropy associated with different occupation statistics for the energy states at different temperatures [13]. Thus, the network can be regarded as a heat bath with a corresponding temperature. The properties of this physical system are described by a partition function (i.e. the sum of the probabilities of the different micro-states of the system), where the micro-states of a network are viewed as the energy states resulting from choosing a suitable Hamiltonian operator [12, 14].

The Hamiltonian operator specifies the energy states in a network which are occupied with particles in thermal equilibrium with the heat bath [15]. To pursue the analysis of this system using the apparatus of statistical mechanics, we require a model for the occupation statistics of the different energy levels [16].

Classically, when the only effects are the thermalisation of the energy states, the Maxwell-Boltzmann statistics can be used to describe the dependence of the energy state occupation upon temperature. According to this model, the particles are weakly interacting and distinguishable [17, 13]. However, if we admit quantum mechanical spin, then the particles become indistinguishable and the pattern of energy state occupation obeys the relevant spin-statistics. Here, we must distinguish whether the particles are fermions or bosons, i.e. whether they have half-integer or integer spin. In the case of fermions with half-integer spin, the particles follow Fermi-Dirac statistics and obey the Pauli exclusion principle. The Cayley tree and other types of geometric networks follow these statistics [18]. On the other hand, bosons do not obey the Pauli exclusion principle and follow Bose-Einstein statistics. Bosons can aggregate in the same energy state. In the low temperature limit, they condensate at the lowest energy state giving rise to a so-called Bose-Einstein condensation [19]. This interesting phenomenon has been extended to the study of network structure. For example, by mapping the network model to a Bose gas, Bose-Einstein condensation has been shown to be closely related to phase transitions in the network evolution [19].

Although different types of spin-statistics provide a sophisticated tool for structural network analysis, they do not lend themself easily either to the characterisation of network entropy or to the decomposition of the global entropy into edge or subnetwork structure. Our previous study shows that thermodynamic characterisations of network entropy can be projected onto the edges of a network using Maxwell-Boltzmann [20, 12], Bose-Einstein and Fermi-Dirac statistics [21]. Here, we perform a novel systematic study of the resulting edge-entropy components, which provides the detailed distribution of entropy across the edges of a network.

To this end, in this paper we aim to first consolidate our prior work which has appeared in the fragmented form of several conference and workshop papers [20, 12, 21], and then to explore the analysis of edge entropy in greater depth. Firstly, we consolidate our theoretical results and present them in more

detail. Secondly, we conduct more extensive experiments on both synthetic and real-world data to explore both the effectiveness and the potential uses of thermodynamic edge entropy in more depth. Specifically, we compare the differences in thermodynamic edge entropy obtained with classical occupation statistics on the one hand and quantum occupation statistics on the other. We further provide a more comprehensive analysis of the role of thermal parameters in controlling the entropy. At high temperatures, the effects of quantum spin-statistics are disrupted by thermalisation and behave identically to the classical Maxwell-Boltzmann case. However, at low temperature, the Bose-Einstein system condenses into a state where the particles coalesce into the lowest energy state. The Fermi-Dirac system, on the other hand, admits only one particle per energy state. These two quantum spin models produce quite different entropic characterisations of the network structure and are, therefore, appropriate for characterising different types of network structure. By applying linear discriminant analysis to a vector of extracted edge entropies, it is possible to distinguish different network structure. This reveals that both classical and quantum statistics determine different facets of the distribution of edge-entropy and how entropy encodes the intrinsic differences in various types of network structure [3].

Although it draws on quite sophisticated concepts from statistical mechanics and quantum physics, there are some simple intuitions underpinning our work. The original von Neumann treatment of graph entropy leads to expressions for the edge entropy which are determined purely by the number of nodes and edges in a graph and the degrees of the nodes connected by the edge in question. By introducing the concept of a graph being in thermal equilibrium with a heat bath, we have the possibility of allowing for thermal agitation of the edges which can be controlled by a temperature parameter. The Hamiltonian of this thermal system is the Laplacian of the graph, with the node degree matrix playing the role of potential energy and the adjacency matrix determining the kinetic energy. The edge-structure of the graph (i.e., the kinetic energy) responds to changes in temperature. An important aspect of our study is to explore how spin

statistics control the types of edge-structure of a graph. We explore two different cases where particles are allowed to occupy the energy states of the thermal system (given by the eigenvalues of the Laplacian) according to Bose-Einstein and Fermi-Dirac statistics. In the former case the particles are bosons and at low temperatures can condense into multiple occupancy of the lower energy states. In the latter case the particles are fermions and only a single particle can occupy each energy state. At low temperatures, in the Bose-Einstein case only the lower eigenvalues of the Laplacian determine the entropy of individual edges, and this can be interpreted as being associated with cluster formation in the graph. In the Fermi-Dirac case, on the other hand, more of the eigenvalue spectrum participates and this means that the edge entropy is sensitive to the type of network structure under study; specifically, it is capable of distinguishing small-world, scale-free and random (Erdős-Rényi) networks. At high temperatures these different behaviours are destroyed by thermal agitation, and our thermal system follows Maxwell-Boltzmann statistics.

2. Preliminaries

Let $G(V, E)$ be a network with node set V and edge set $E \subseteq V \times V$. The total number of nodes is $|V|$ with the adjacency matrix A being defined as

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

The degree matrix D has diagonal elements $D(u, u) = d_u = \sum_{v \in V} A_{uv}$ and zero off diagonal elements.

The normalised Laplacian matrix $\tilde{L}$ is then defined as

$$\tilde{L} = D^{-\frac{1}{2}} L D^{\frac{1}{2}} = D^{-\frac{1}{2}}(D - A)D^{\frac{1}{2}} \tag{2}$$

where $L = D - A$ is the Laplacian matrix. The normalised Laplacian has eigen-decomposition

$$\tilde{L} = \Phi \tilde{\Lambda} \Phi^T \tag{3}$$

where $\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_{|V|})$ is the matrix with the ordered eigenvectors as columns and $\tilde{\Lambda} = diag(\lambda_1, \lambda_2, \ldots \lambda_{|V|})$ is the diagonal matrix with the ordered eigenvalues as elements.

2.1. The Laplacian as a Density Matrix

The density matrix of a quantum system is defined as the sum of the probability for its pure quantum states $|\psi_i\rangle$,

$$\boldsymbol{\rho} = \sum_{i=1}^{|V|} p_i |\psi_i\rangle\langle\psi_i| \tag{4}$$

where the "ket" $|\psi_i\rangle$ uses the Dirac notation to represent a pure quantum state as a complex-valued column vector and p_i is the probability for finding the system in a statistical mixture of pure states $|\psi_i\rangle$.

This definition has been extended to the network domain by using the scaled normalised Laplacian matrix [8] as the network density matrix

$$\boldsymbol{\rho} = \frac{\tilde{L}}{|V|} \tag{5}$$

The density matrix is Hermitian with the properties that $\boldsymbol{\rho} = \boldsymbol{\rho}\dagger$ and $\boldsymbol{\rho} \geq 0, \mathrm{Tr}[\boldsymbol{\rho}] = 1$. It is a quantum operator that measures the expected value of the observable states of the network, i.e. its eigenvalues.

2.2. The von Neumann Entropy

The interpretation of the normalised Laplacian as a density matrix opens up the possibility of computing the von Neumann entropy of the network. In terms of the density matrix, the von Neumann entropy is defined to be

$$S_{VN} = -\mathrm{Tr}[\boldsymbol{\rho} \log \boldsymbol{\rho}] \tag{6}$$

which can be written in terms of the Shannon entropy of the density matrix eigenvalues as

$$S_{VN} = -\sum_{i=1}^{|V|} \frac{\hat{\lambda}_i}{|V|} \log \frac{\hat{\lambda}_i}{|V|} \tag{7}$$

However, the numerical computation of the Laplacian eigenvalues is in general cubic in the number of nodes in the graph $|V|$. To overcome this bottleneck, Han et al.[4] find an approximate expression for the von Neumann entropy and in so doing reduce the computation to being quadratic in the number of nodes. They make use of the approximation $x \log x \approx x(1-x)$ to replace the Shannon entropy $\frac{\hat{\lambda}_i}{|V|} \log \frac{\hat{\lambda}_i}{|V|}$ by the quadratic approximation $\frac{\hat{\lambda}_i}{|V|}(1 - \frac{\hat{\lambda}_i}{|V|})$. Expressing the sums of the eigenvalues and the sum of their squares by $Tr[\hat{L}]$ and $Tr[\hat{L}^2]$, respectively, the final expression for the approximate von Neumann entropy is

$$S_{VN} = \frac{1}{|V|} Tr[\hat{L}] - \frac{1}{|V|^2} Tr[\hat{L}^2] \tag{8}$$

Furthermore, the traces of the normalised Laplacian and its square, i.e. $Tr[\hat{L}]$ and $Tr[\hat{L}^2]$, can be expressed in terms of the number of nodes in the graph and the degrees of pairs of nodes connected by edges. As a result,

$$S_{VN} = 1 - \frac{1}{|V|} - \frac{1}{|V|^2} \sum_{(u,v)\in E} \frac{1}{d_u d_v} \tag{9}$$

This approximation allows the von Neumann entropy to be computed without explicitly solving the eigensystem for the normalised Laplacian. Thus, the von Neumann entropy can be computed in quadratic time using the node-degrees for pairs of nodes connected by edges. Moreover, the global network entropy is just a sum of contributions from individual edges, and the entropy of the edge connecting nodes u and v is

$$S_{VN}^{edge}(u,v) = \frac{1}{|E|} - \frac{1}{|V||E|} - \frac{1}{|E||V|^2} \frac{1}{d_u d_v} \tag{10}$$

and the global network entropy is $S_{VN} = \sum_{(u,v)\in E} S_{VN}^{edge}(u,v)$.

Therefore, Han et al's approximation to the von Neumann entropy straightforwardly decomposes into contributions from the individual edges. In Eq.(10), the edge entropy depends on the degrees of the nodes at both ends of an edge. A high value of node degree gives a large value of the approximate von Neumann entropy, and a low degree a small value. The minimum and maximum values of the entropy of each edge relates to the structure of network. For cycles or

string-like network structures, each node has a fixed minimum value of degree. The edge entropy for this kind of network structure is therefore a minimum for possible connected structures. On the other hand, for the fully connected network, all of the nodes have the fixed maximum number of degrees. This type of network structure has the maximum value of edge entropy for a connected structure.

3. Thermodynamic Representation

3.1. Hamiltonian Operator

This work on the efficient computation of von Neumann entropy provides one route to the entropy of network edges [8], through an analogy with quantum mechanics, and based on the density matrix. Another route is to use ideas from statistical mechanics. This commences by defining a partition function over the micro-states for a network, and then computing thermodynamic quantities from the partition function by appealing to a thermodynamic analogy in which the network is in thermal equilibrium with a heat bath. From a quantum perspective, this is equivalent to associating a Hamiltonian with the network. The Hamiltonian operator contains two components, i.e., kinetic energy and potential energy.

According to this picture, the kinetic energy measures the state of internode edge connection, which is equal to the negative value of elements in the adjacency matrix. The potential energy, on the other hand, is determined by the diagonal elements of the degree matrix. The physical picture here is that the the higher the degree of the nodes in a network, the larger their potential energydue to their propensity to connect to other vertices. Thus, the Laplacian matrix in the network can be regarded as the Hamiltonian operator. Similarly, the normalised form of Laplacian matrix is also give rise to an equivalent representation since $\hat{H} = \tilde{L}$.

Our assumption is that the network is in contact with a heat reservoir, and a set of particles can occupy the thermalised energy states defined by the

Hamiltonian, i.e. the Laplacian eigenvalues. In other words, the particles occupy the energy states subject to thermal agitation by the heat bath. The energy states, i.e., the eigenvalues of the Hamiltonian operator are the solution of the relevant Schrödinger equation. Thus, the energy states within the network $\{\varepsilon_i\}$ are the eigenvalues of the normalised Laplacian matrix, and satisfy

$$\hat{H}|\psi_i\rangle = \tilde{L}|\psi_i\rangle = E_i|\psi_i\rangle \tag{11}$$

where all eigenvalues are greater than or equal to zero, with the multiplicity of zeros representing the number of connected components within the network.

3.2. Thermodynamic Quantities

The interpretation of von Neumann entropy opens up the possibility of directly characterising network entropy by using spectral graph theory [7]. However, thermodynamic analogies of the sort outlined above also provide powerful tools for network analysis [14]. The statistical mechanical basis of thermodynamics combined with a graph spectral network characterisation of network structure provides a microscopic perspective for viewing network network structure [22].

In the heat bath analogy, where the network is viewed as a thermal system in equilibrium with a heat reservoir, the statistical occupation of the microstates of a network by particles is described by a partition function associated with a suitably chosen Hamiltonian [14]. The corresponding entropy and average particle energy can be derived from the partition function for the network energy microstates.

Here, we consider the thermal system with consisting of N particles which occupy the energy states associated with a network in thermal equilibrium with a heat bath at the temperature T. Let $\beta = 1/k_BT$, where k_B is the Boltzmann constant which we set to the unity. The partition function $Z(\beta, N)$ can be used to compute the following thermodynamic characteristics of the network system, a) the average energy

$$U = \mathrm{Tr}\left(\rho H\right) = k_B T^2 \left[\frac{\partial}{\partial T}\log Z\right]_N \tag{12}$$

b) the thermodynamic entropy

$$S = k_B \left[\frac{\partial}{\partial T} T \log Z \right]_N \tag{13}$$

c) the chemical potential

$$\mu = -k_B T \left[\frac{\partial}{\partial N} \log Z \right]_\beta \tag{14}$$

Both energy and entropy are weighted functions for the energy eigenvalues, i.e. network normalised Laplacian. The particle occupation statistics for the different energy states are governed by the partition function and the thermal parameter of the system, i.e. temperature T. Here, we explore in detail the thermodynamic entropy and how it can be used to represent the intrinsic structure of networks.

4. Quantum Spin Statistics and Network Entropy

4.1. Maxwell-Boltzmann Entropy

Classically, for weakly interacting distinguishable particles, the probability of finding a particle in the different energy states specified by the network Hamiltonian is governed by Maxwell-Boltzmann statistics. The partition function describing these occupation statistics is

$$Z_{MB} = \mathrm{Tr}\left[\exp(-\beta \tilde{L})^N \right] = \left(\sum_{i=1}^{|V|} e^{-\beta \lambda_i} \right)^N \tag{15}$$

where $\beta = 1/k_B T$ is the reciprocal of temperature, k_B is the Boltzmann constant, N is the total number of particles and λ_i are the eigenvalues of normalised Laplacian matrix.

From Eq.(13), the corresponding entropy for the network is

$$S_{MB} = \log Z - \beta \frac{\partial \log Z}{\partial \beta} = -N \sum_{i=1}^{|V|} \frac{e^{-\beta \lambda_i}}{\sum_{i=1}^{|V|} e^{-\beta \lambda_i}} \log \frac{e^{-\beta \lambda_i}}{\sum_{i=1}^{|V|} e^{-\beta \lambda_i}} \tag{16}$$

For a single particle, the equivalent density matrix is

$$\boldsymbol{\rho}_{MB} = \frac{e^{-\beta \lambda_i}}{\sum_{i=1}^{|V|} e^{-\beta \lambda_i}} \tag{17}$$

Since the density matrix and Hamiltonian operator commute, the network is in equilibrium and as a result the thermodynamic entropy is N times the corresponding entropy for a single particle network.

4.2. Bose-Einstein Entropy

Bose-Einstein statistics describe indistinguishable bosons, i.e. particles with integer spin. The occupation number in each energy state is unlimited, and as a result particles can aggregate in the same state. For a varying number of particles, the chemical potential μ specifies the network Hamiltonian with the partition function given by

$$Z_{BE} = \det\left(I - e^{\beta\mu}\exp[-\beta\tilde{L}]\right)^{-1} = \prod_{i=1}^{|V|}\left(\frac{1}{1-e^{\beta(\mu-\lambda_i)}}\right) \tag{18}$$

From Eq.(13), the corresponding entropy is given by

$$S_{BE} = -\sum_{i=1}^{|V|}\log\left(1-e^{\beta(\mu-\lambda_i)}\right) - \beta\sum_{i=1}^{|V|}\frac{(\mu-\lambda_i)e^{\beta(\mu-\lambda_i)}}{1-e^{\beta(\mu-\lambda_i)}} \tag{19}$$

The thermodynamic variables depend on both the chemical potential that controls the number of particles and the temperature appearing in the partition function. At the reciprocal temperature β, the number of particles occupying the energy state indexed i is,

$$n_i = \frac{1}{\exp[\beta(\lambda_i-\mu)]-1} \tag{20}$$

and so the total number of particles in the network is

$$N = \sum_{i=1}^{|V|} n_i = \sum_{i=1}^{|V|}\frac{1}{\exp[\beta(\lambda_i-\mu)]-1} = \mathrm{Tr}\left[\frac{1}{\exp(-\beta\mu)\exp[\beta\tilde{L}]-I}\right] \tag{21}$$

The chemical potential must be less than the minimum value of the energy of the state, i.e. $\mu < \min\lambda_i$ to ensure that the occupation number for each state is non-negative. The equivalent density matrix is given by

$$\boldsymbol{\rho}_{BE} = \frac{1}{\mathrm{Tr}(\boldsymbol{\rho}_1)+\mathrm{Tr}(\boldsymbol{\rho}_2)}\begin{pmatrix}\boldsymbol{\rho}_1 & 0\\ 0 & \boldsymbol{\rho}_2\end{pmatrix} \tag{22}$$

where $\boldsymbol{\rho}_1 = -\left(\exp[\beta(\tilde{L}-\mu I)]-I\right)^{-1}$ and $\boldsymbol{\rho}_2 = \left(I-\exp[-\beta(\tilde{L}-\mu I)]\right)^{-1}$.

The aggregation of particles into the lower energy states at low temperature is strongly determined by the lower end of the Laplacian spectrum. As a result, it is sensitive to the spectral gap and the number of connected components in the network.

4.3. Fermi-Dirac Entropy

Fermi-Dirac statistics describe the behaviour of fermions, i.e. indistinguishable particles with half integer spin which obey the Pauli exclusion principle. The maximum number of particles that can occupy a single energy state is unity. The corresponding partition function is

$$Z_{_{FD}} = \det\left(I + e^{\beta\mu}\exp[-\beta\tilde{L}]\right) = \prod_{i=1}^{|V|}\left(1+e^{\beta(\mu-\lambda_i)}\right) \tag{23}$$

From Eq.(13), the associated entropy can be achieved as

$$S_{_{FD}} = \sum_{i=1}^{|V|}\log\left(1+e^{\beta(\mu-\lambda_i)}\right) - \beta\sum_{i=1}^{|V|}\frac{(\mu-\lambda_i)e^{\beta(\mu-\lambda_i)}}{1+e^{\beta(\mu-\lambda_i)}} \tag{24}$$

The occupation number at the ith energy state is

$$n_i = \frac{1}{\exp[\beta(\lambda_i-\mu)]+1} \tag{25}$$

and the total number of particles is

$$N = \sum_{i=1}^{|V|} n_i = \sum_{i=1}^{|V|}\frac{1}{\exp[\beta(\lambda_i-\mu)]+1} = \mathrm{Tr}\left[\frac{1}{\exp(-\beta\mu)\exp[\beta\tilde{L}]+I}\right] \tag{26}$$

The chemical potential corresponds to the energy of the level in question, i.e. $\mu = \lambda_n$ for a single particle per energy state. The equivalent density matrix is

$$\boldsymbol{\rho}_{_{FD}} = \frac{1}{\mathrm{Tr}(\boldsymbol{\rho}_3)+\mathrm{Tr}(\boldsymbol{\rho}_4)}\begin{pmatrix}\boldsymbol{\rho}_3 & 0\\ 0 & \boldsymbol{\rho}_4\end{pmatrix} \tag{27}$$

where $\boldsymbol{\rho}_3 = \left(I+e^{-\beta\mu}\exp[\beta\tilde{L}]\right)^{-1}$ and $\boldsymbol{\rho}_4 = \left(I+e^{\beta\mu}\exp[-\beta\tilde{L}]\right)^{-1}$.

At low temperature, unlike bosons where multiple occupation of the same energy state is permitted and the entropy is largely determined by the lowest

few eigen-values, the entropy for Fermi-Dirac statistics is sensitive to a greater portion of the Laplacian spectrum since the occupation of the same energy state is limited to a single particle. This difference in energy state occupation can reflect subtle differences in a network structure, as well as to distinguish different network models.

4.4. Edge Entropy Analysis

Rather than focussing on the global thermodynamic entropy for a network, in this paper, we are interested in how the entropy is distributed across its different edges, and how this distribution can be used to characterise and probe network structure.

We therefore adopt a spectral approach and decompose the global network entropy into components residing on the individual edges. We conduct this by decomposing the normalised Laplacian into eigenvalues and eigenvectors and then projecting global entropy onto this basis [21]. The idea is as follows. We make use of the eigen-decomposition of the Laplacian matrix to write the edge-entropy matrix, i.e. the symmetric square matrix with edge entropies as off-diagonal elements, as the matrix function $\mathbf{S} = \Phi F(\Lambda)\Phi^T$, where $F(\Lambda)$ is a diagonal matrix with elements $\sigma(\lambda_1), ..., \lambda_{|V|}$, and $\sigma(\lambda)$ is a real-valued function of λ. The entropy associated with the edge $(u, v) \in E$ is then given by the element indexed (uv) of the matrix $\mathbf{S}$, i.e.

$$\mathbf{S}_{u,v} = \sum_{i=1}^{|V|} \sigma(\lambda_i)\varphi_i(u)\varphi_i(v) \tag{28}$$

With an appropriate choice of the function σ, we can compute edge-entropy for each of the different set of occupation statistics. Specifically, for Maxwell-Boltzmann statistics:

$$\sigma_{_{MB}}(\lambda_i) = -N\frac{e^{-\beta\lambda_i}}{\sum_{i=1}^{|V|} e^{-\beta\lambda_i}} \log \frac{e^{-\beta\lambda_i}}{\sum_{i=1}^{|V|} e^{-\beta\lambda_i}}$$

for Bose-Einstein statistics:

$$\sigma_{BE}(\lambda_i) = -\sum_{i=1}^{|V|} \log\left(1 - e^{\beta(\mu-\lambda_i)}\right) - \beta \sum_{i=1}^{|V|} \frac{(\mu-\lambda_i)e^{\beta(\mu-\lambda_i)}}{1-e^{\beta(\mu-\lambda_i)}}$$

and for Fermi-Dirac statistics:

$$\sigma_{FD}(\lambda_i) = \sum_{i=1}^{|V|} \log\left(1 + e^{\beta(\mu-\lambda_i)}\right) - \beta \sum_{i=1}^{|V|} \frac{(\mu-\lambda_i)e^{\beta(\mu-\lambda_i)}}{1+e^{\beta(\mu-\lambda_i)}}$$

As a result, the spectral projection of the global entropy allows the entropy of individual edges to be computed. As we will demonstrate later it provides a useful and relatively straightforward local entropic characterisation of network structure.

4.5. Linear Discriminant Analysis

Finally, we consider various network analysis problems involving samples of networks on a fixed set of uniquely labelled nodes. If there are $|V|$ node labels, and each label refers to a single node in a graph, then there are at most $|V| \times (|V| - 1)$ distinct edges between different nodes. Here, for each network, we take the edge entropies as the ordered components of a feature vector for that network. This fixed ordinal arrangement places the edges connected to the same pair of labelled nodes in the same element of the feature vector. Here, we focus on the graph collections with the same number of vertices. For the graph with inflexible number of nodes, the variance of vertices can work as a salient feature, which is not a challenge issue to discriminate different network structures.

Let $F = \Phi\Sigma\Phi^T$ be symmetric square $|V| \times |V|$ matrix with the edge entropies as elements. We vectorise this edge entropy matrix by concatenating the columns of the upper triangle of F to form a vector of length $1/2|V| \times (|V|-1)$, which we denote by $\vec{f}_i$ for the ith graph. Then, we can apply linear discriminant analysis to classify networks using their associated feature vector in a supervised manner [23].

Suppose we have a sample of n networks, each of which is known to belong to one of C different classes. Let K_c be the index-set of a set of networks with

entropic edge features known to belong to class c, and let $\vec{f_i}$ be the entropic feature vector of the network indexed i. The mean entropic feature vector for the class c is given by

$$\mu_c = \frac{1}{|K_c|} \sum_{i \in K_c} \vec{f_i} \tag{29}$$

and the overall population mean is given by

$$\mu = \frac{1}{n} \sum_{i=1}^{n} \vec{f_i} \tag{30}$$

Thus, the between class covariance matrix for the edge entropy feature vectors is

$$B = \frac{1}{n} \sum_{c=1}^{C} (\mu_c - \mu)(\mu_i - \mu)^T \tag{31}$$

The within-class variance W, on the other hand, is given by

$$W = \frac{1}{n} \sum_{c=1}^{C} \frac{1}{|K_c|} \hat{X}_c \hat{X}_c^T \tag{32}$$

where X_c is the matrix with the entropic feature vectors for class c as columns.

For jointly maximising the between-class covariance and minimising the within-class variance, we use the joint criterion

$$J = \frac{u^T B u}{u^T W u} \tag{33}$$

This separation criterion is maximised by the eigenvectors u of the matrix $W^{-1}B$ when the separation criterion will be equal to the corresponding eigenvalue.

If $W^{-1}B$ is diagonalizable, the variability between feature vectors will be contained in the subspace spanned by the eigenvectors corresponding to the $C-1$ largest eigenvalues. These eigenvectors can be used in feature reduction, as in principal component analysis (PCA). The eigenvectors corresponding to the smaller eigenvalues will tend to be very sensitive to the exact choice of training data, and it is often necessary to use regularisation.

For network classification, we apply regularised linear discriminant analysis (LDA) [24] to quantify the classification accuracy obtained with the features extracted using edge entropy. These correspond to the eigenvectors associated

with the eigenvalues falling into the top 10 percentile. The discriminant analysis model is based on the assumption that the edge features follow a multivariate normal distribution with an identical covariance matrix for each class [25].

According to the Bayes theorem, the posterior probability that an entropic network feature f in a class c is

$$P(c|f) = \frac{P(f|c)P(c)}{\sum_{c=1}^{C} P(f|c)P(c)} \tag{34}$$

where $P(c)$ is the prior probability of class c, and $p(f|c)$ is the probability density function of the multivariate normal distribution for class c. The class density function is assumed to follow the Gaussian distribution

$$P(f|c) = \frac{1}{\sqrt{2\pi|\Sigma_c|}} \exp\left(-\frac{1}{2}(f-\mu_c)^T\Sigma_c^{-1}(f-\mu_c)\right) \tag{35}$$

where μ_c is the mean feature vector for class c, Σ is the class covariance matrix for class c and Σ_c^{-1} is its inverse.

The classification strategy is to minimise the expected classification error

$$\hat{y} = \arg\min_{y=1,2,..c} \sum_{c=1}^{C} E(y|c)P(c|f) \tag{36}$$

where y is the network class label, C is the number of classes, $P(c|f)$ is the posterior probability of class c for entropic feature vector f and $E(y|c)$ is the confusion probability of classifying a network as class y when its true class label is c.

The expected confusion probability of a network misclassification $E(y|c)$ is given by

$$E(y|c) = \sum_{k=1}^{C} P(k|f)E(c|k) \tag{37}$$

where $P(k|f)$ is the posterior probability of a network in the class k with entropic feature vector f. $E(c|k)$ is the misclassification error probability for assigning a network in class k when its true class label is c. Here $E(c|k) = 1$ if $k = c$, and $E(c|k) = 0$, if $k \neq c$.

5. Experimental Evaluation

In this section, we conduct experiments to demonstrate the application of network entropy for analysing interregional connectivity. We first perform a qualitative analysis on the components of the quantum edge entropy and their dependence on degree and temperature. We compare their performance with the standard von Neumann entropy. Then, we conduct a quantitative analysis on real-world networks, including protein interaction networks, financial networks and fMRI brain connectivity networks. We evaluate whether the edge-entropy decomposition can identify significant structural variance in samples of networks. To simplify the calculations, we set the Boltzmann constant to be unity.

5.1. Data-sets

In our experiments, we explore four different kinds of the dataset to evaluate the performance of the edge entropy in the networks. The first data-set consists of synthetic networks generated using three typical complex network models (Erdős-Rényi random graphs, Watts-Strogatz small world and Barabási-Albert scale-free networks). The remaining three network datasets come from real-world situations in different domains.

Data-set 1: The synthetic data contains three different kinds of network structure which are generated according to typical complex network models, namely, a) the Erdős-Rényi random graph, b) the Watts-Strogatz small-world network [1], and c) the Barabási-Albert scale-free network [2, 26]. All of these networks have the same number of nodes, which is set to $N = 100$. The parameter in the random graphs, i.e., the probability of connection between two nodes, is set to 0.5. Similarly, the parameter of rewiring probability in the small-world network is 0.8, and the connecting parameter in the scale-free network is 10.

Data-set 2: The Protein-Protein Interactions (PPIs) dataset comes from STRING-8.2 [27] involving networks of interaction between histidine kinase and other proteins. The total number of 173 PPIs come from 4 different kinds

of bacteria, namely, a) 8 Aquifex and Thermotoga from Aquifex Aelicus and Thermotoga Maritima separately, b) 40 Proteobacteria from Acidovorax Avenae, c) 73 Cyanobacteria from Anabaena Variabilis, and d) 52 Gram-Positive from Staphylococcus Aureus.

Data-set 3: The time-evolving network dataset comes from the New York Stock Exchange with the daily prices of 3,799 stocks. The trading period is from January 1986 to February 2011 covering six thousand trading days. In the network, each stock is represented as a node and the edges indicate the statistical similarity between the stock closing price time series with a 20-day window [28]. To obtain a network time series, we shift the time window of 20 days over the closing price time series to compute the cross-correlation coefficients between the windowed time series for each pair of stocks. An empirically determined correlation threshold is set to 0.85 to identify the edges. This yields a stock market network time series with a fixed number of 347 nodes and varying edge structure over 6,000 trading days.

Data-set 4: The fMRI data is supplied by the ADNI initiative [29]. The fMRI images are scanned every two seconds and record the detail of the Blood-Oxygenation-Level-Dependent(BOLD) signals with different anatomical brain regions. Ninety-six anatomical regions are identified as regions of interest (ROIs) by aggregating the voxels in the fMRI images. The correlation between different BOLD signals in the pairs of ROIs represents their functional connectivity driven by neural activities [17]. There are four categories of patients according to the degree of disease severity, i.e., a) full Alzheimer's (AD), b) Late Mild Cognitive Impairment (LMCI), c) Early Mild Cognitive Impairment (EMCI) and d) Normal Healthy Controls (HC). The dataset contains 53 AD, 83 LMCI, 119 EMCI, and 96 HC subjects.

5.2. Experiments on Synthetic Data

We first investigated how well the edge entropy performs as a network feature to distinguish between synthetic networks from different models in Data-set 1. Here, we compare the edge entropy distribution for the three different network

models. We also compare the distribution for the entropy resulting from the three different types of occupation statistics (Maxwell-Boltzmann, Fermi-Dirac and Bose-Einstein).

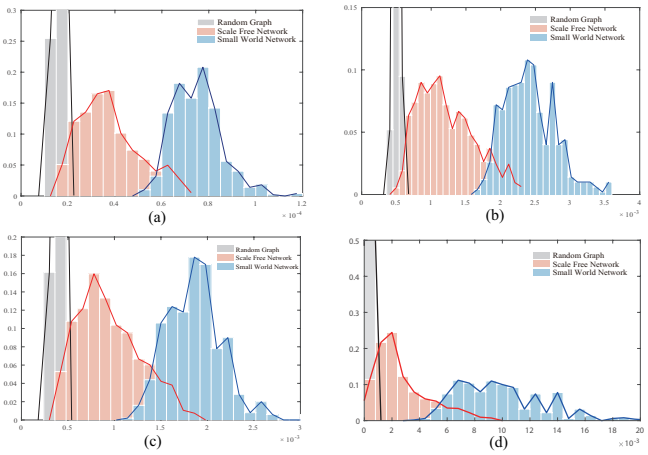


Figure 1: Histograms of edge entropy for three different network models. (a) Maxwell-Boltzmann statistics; (b) Bose-Einstein statistics; (c) Fermi-Dirac statistics; (d) von Neumann entropy. The networks are generated with the number of nodes $N = 100$. The grey area represents Erdős-Rényi random graphs with connection probability $p = 0.5$; the red area is scale-free networks with edges $m = 10$ to attach at every step; the blue area is small-world networks with the link rewiring probability $p = 0.8$. Temperature $\beta = 0.1$ and the number of particles $N = 10$.

In Fig.1, each subfigure shows the histograms of the edge entropy from a different statistical occupation model. The coloured areas in red, grey and blue correspond to the distribution of edge entropy value for the three network models (random graph, small-world and scale-free). It is clear that, in each case, the edge entropy for the Erdős-Rényi random graph model always has a lower value compared to the alternative network models. Small-world networks, on the other hand, give much larger edge entropy. The random graph and scale-free network edge entropy distribution are closer for all three types of occupation statistics.

Taking this study one step further, we explore the relationship between edge entropy and the edge degree configuration. To this end, we plot the three-

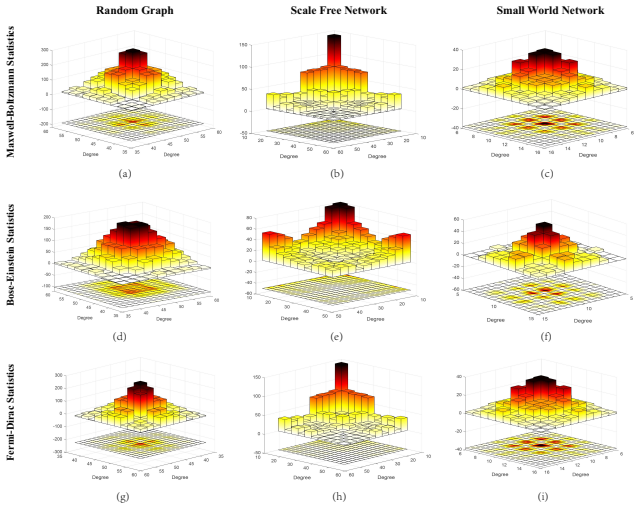


Figure 2: Distributions of edge entropies with corresponding node degree combinations for the three statistical mechanical models: (a)-(c) Maxwell-Boltzmann statistics, (d)-(f) Bose-Einstein statistics and (g)-(i) Fermi-Dirac statistics.

dimensional distribution of the edge entropy distribution versus the two-node degrees defining the edge. As shown in Fig.2, the three network models exhibit different shapes for the entropy-degree distribution. In the case of Bose–Einstein statistics, the area of the non-overlapping distribution for the three network models is more significant than that in the Maxwell-Boltzmann and Fermi-Dirac cases.

This is consistent with our intuition that the three different distributions are related to the network spectra of the Laplacian matrix. For example, the eigenvalues of Erdős-Rényi random graph follow a semicircular distribution, the scale-free networks exhibit a triangular distribution and the small-world networks present a more complex form determined by a parametric model. It is known that the spectral gap between the lowest and the second-lowest eigenvalues is related to the cluster structure of the model. The edge entropy for Bose-Einstein statistics is more sensitive to the spectral gap and better reflects the strong community structure, since the particles preferentially sample the lower energy states (eigenvalues).

The above results indicate that the edge entropy resulting from the three different statistical models is a useful tool that reflects the network structure for the synthetic data. In particular, the Bose-Einstein edge entropy is strongly correlated with the edge degree configuration.

5.3. Experiments on Real-world Data

5.3.1. PPI Network

Turning our attention to the real-world datasets, we first consider the Protein-Protein Interactions (PPI) networks. We illustrate the difference between the distribution of edge entropy for both the von Neumann and Maxwell-Boltzmann cases. The edge entropy distribution for two different PPIs, i.e., Anabaena Variabilis and Aquifex Aelicus, are shown in Fig.3. In the Maxwell-Boltzmann case,

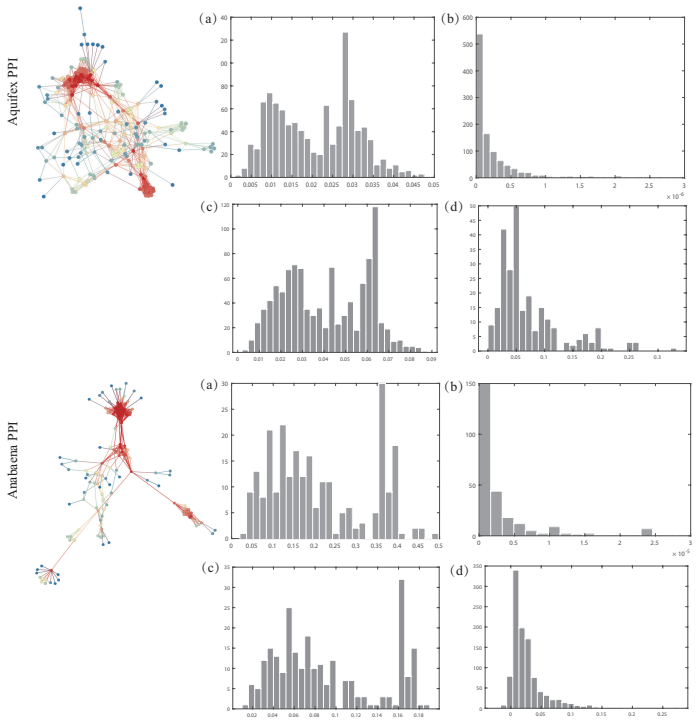


Figure 3: The distribution of edge entropy for two different PPI networks (Aquifex and Acidovorax) (a) Maxwell-Boltzmann entropy; (b) Von Neumann entropy; (c) Bose-Einstein entropy; (d) Fermi-Dirac entropy

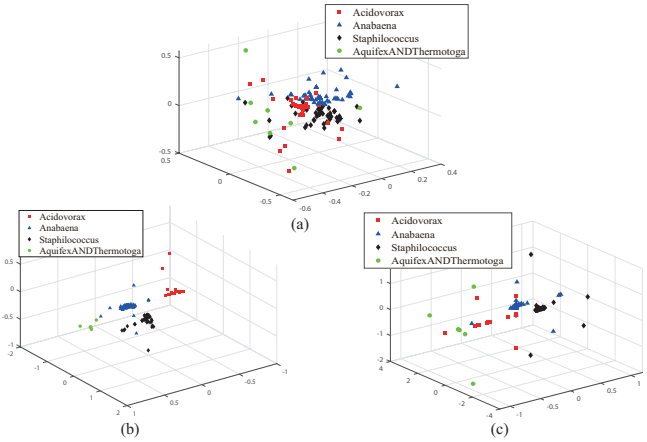


Figure 4: 3D visualisation of LDA performance on PPI Networks. (a) Maxwell-Boltzmann edge entropy, (b) Bose-Einstein edge entropy, (c) Fermi-Dirac edge entropy.

the histogram of edge entropy provides better discrimination between the two PPIs. The reason is that it is more sensitive to edges with a high degree. This discriminative power of edge entropy is evident in identifying network differences. Compared to the von Neumann case, whose histogram is concentrated at lower values with a single peak, the histogram of Maxwell-Boltzmann edge entropy shows two separated peaks. This means that the von Neumann edge entropy is not particularly effective as a tool to detect salient network structure when compared to the Maxwell-Boltzmann edge entropy.

Next, we perform LDA on the PPI data by concatenating the columns of edge entropy matrix to construct a feature vector. With the known labels for different groups of PPIs, we compute the within-class and between-class covariance matrices, respectively. Then, we select the feature vectors for those individual edge entropies associated with the largest variances. To do this, we apply Linear Discriminant Analysis as described in Section 4.5 to obtain optimal projection axes for class separation. In Fig.4 the four groups of PPIs are visualised using a 3D scatter plot of the three leading principle components for each PPI network. Each of the entropies derived from either of three statistical

models work well as features to separate the different PPI networks into four different clusters according to their structural characteristics. The four clusters are overlapped in Maxwell-Boltzmann case. The Bose-Einstein statistics, on the other hand, exhibit the clear and best performance of separation for different the groups compared to the alternatives.

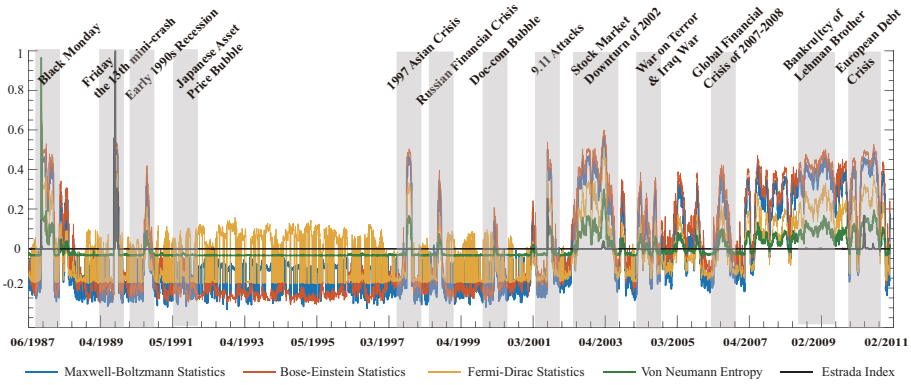


Figure 5: The thermodynamic entropy associated with partition functions identifies critical financial events in NYSE ($N = 2$ and $\beta = 7$).

5.3.2. Financial Network

Next, we focus on analysing time-evolving networks. Here, we conduct experiments on the financial networks extracted from the New York Stock Exchange data [28]. We aim to explore the entropic changes related to variations in network structure.

Fig.5 plots the four different entropies coming from the different classical and quantum statistical models, together with the Estrada index. Each of these entropies can detect the temporal anomalies in network structure related to the occurrences of significant global financial events. The value of the entropy undergoes significant fluctuations corresponding to the dramatic internal changes in network structural change during the financial crises. By contrast, the Estrada index is not sensitive to the fluctuations in the stock markets. It only shows a few peaks that indicate the whereabouts of the most critical financial events.

Compared to the von Neumann entropy, the Maxwell-Boltzmann entropy is more sensitive to the market fluctuations. For example, the Japanese Asset Price Bubble in 1992 shows obvious differences in the behaviour of these two entropies.

Similarly, the variance in financial market network structure can also be clearly observed in the entropy derived from quantum statistics. During the financial crises, in each case, the value of entropy undergoes a sharp increase corresponding to the dramatic fluctuation in network structure. In fact, the Bose-Einstein entropy exhibits the greatest variation during the crises. This indicates that the critical network structure undergoing change is the cluster-structure of the network (or modularity), and that this is undergoing profound changes during these extreme financial episodes.

5.3.3. fMRI Brain Connectivity Network

Finally, we explore the class structure of the fMRI brain activation networks using the different edge entropies. We aim to identify those anatomical regions that are crucial in the development of Alzheimer's disease [29]. As shown in Fig.6, two groups of patients, i.e., Alzheimer's disease (AD) and the healthy control group (Normal), exhibit a difference in the shape of the edge entropy distribution. It is clear that the distribution of von Neumann entropy cannot distinguish the two groups of patients. By comparison, the statistical mechanical methods are more adept at reflecting the details of the distribution in edge entropy. In Fig.6(b), the Maxwell-Boltzmann edge entropy distribution for the Alzheimer's subjects peaks at a lower value compared to the normal control subgroup. This observation is more palpable in the cases of Bose-Einstein and Fermi-Dirac distributions, as shown in Fig.6(c) and Fig.6(d) respectively. The Bose-Einstein edge entropy has a better separation for the normal (healthy) and Alzheimer's groups. The non-overlapping entropic area is much larger than that for the remaining statistical models.

Furthermore, identifying the regions in the brain associated with Alzheimer's disease is also helpful to understand the development of the disease [30]. Neuro-

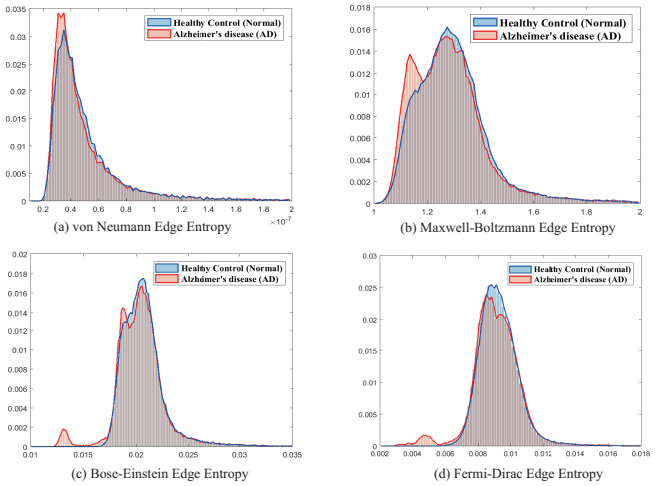


Figure 6: The edge entropy distribution between patients in Alzheimer's disease (AD) and healthy control (Normal). (a) von Neumann entropy, (b) Maxwell-Boltzmann entropy, (c) Bose-Einstein entropy and (d) Fermi-Dirac entropy.

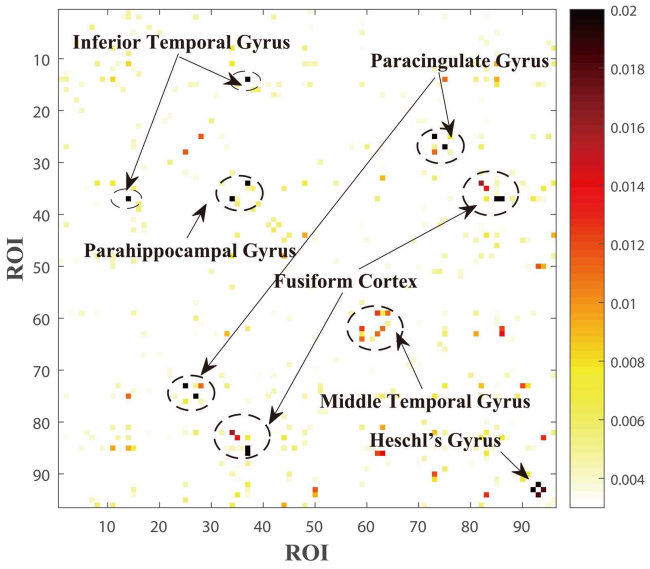


Figure 7: The disease regions in the brain. We apply the edge entropy to identify the significant divergence between AD and HC groups.

anatomical studies show that the anatomical structures in different brain regions are important for understanding different brain disorders [31, 32]. Here, we use the edge entropy to identify those brain regions that give rise to the most significant differences in edge entropy for Alzheimer's patients and normal samples. Fig.7 lists the pairs of regions for edges associated with the largest difference in edge-entropy between the AD and normal samples. Specifically, anatomical regions such as the Paracingulate Gyrus, Parahippocampal Gyrus, Inferior Temporal Gyrus and Temporal Fusiform Cortex, are known to be associated with the loss of interconnection in the brain network structure for AD patients.

We explore whether LDA can be used to perform supervised classification of subjects belonging to the different groups of patients based on the inter-region edge entropies. To this end, we vectorise the edge entropy matrix by concatenating the columns of the upper triangle to form a vector. From the sample of vectors for different subjects, we compute a within class and between class covariance matrices using the known class labels for the four classes of subject. To simplify the analysis we select only those edges connecting different anatomical regions that are associated with the largest entropy variances. We use 200 samples for training and 151 samples for testing. LDA is the applied to the within and between class covariance matrices computed from the training data to obtain optimal projection axes for class separation for the set of training data. We then project the test data onto these axes and then perform classification for the four groups of patients. The corresponding projected edge-entropy vectors are used classify subjects belonging to the different classes. Fig.8 plots the 3D visualisation of the leading three principle components for the four groups of patients. The cluster-centre for each group is represented by the principle eigenvectors in the Fisher discriminant analysis (LDA) [33]. These results show that each of the edge entropies resulting from the three different statistical mechanical models can be used to separate patients effectively. However, the Bose-Einstein model exhibits the best overall performance compared to the alternatives.

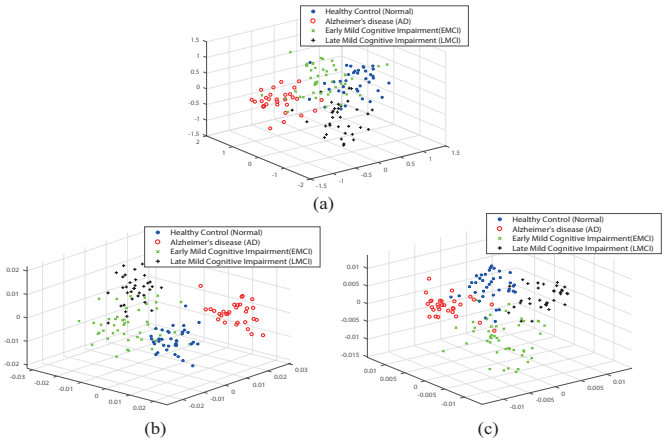


Figure 8: 3D visualisation of LDA performance with principle components in AD, LMCI, EMCI and HC. (a) Maxwell-Boltzmann edge entropy, (b) Bose-Einstein edge entropy, (c) Fermi-Dirac edge entropy.

5.4. Evaluation of Thermal Parameters

We now investigate the thermal parameters and their effects on the performance of the edge entropy distributions resulting from the different models. We first investigate how the parameters in the synthetic network model control the distribution of thermal entropy. For the Erdős-Rényi random graph, we vary the connection probability from 0.1 to 0.9; for the scale free network, we vary the node attachment parameter from 1 to 15; and for the small-world network we vary the average connection parameters from 2 to 30. For each different value of the parameter, we generate a single synthetic network and compute the total network entropy. Fig.9 plots the thermodynamic entropy with different values of parameters for three synthetic networks. All three statistical mechanical entropies decrease as the network parameter is increased, and this corresponds to an increasing edge density giving rise to denser patterns of connection in the network.

To explore the temperature dependence we study how the edge entropy distribution depends on the degree distribution for the different models. Specifically, we consider edges connecting nodes of large degree, median degree and

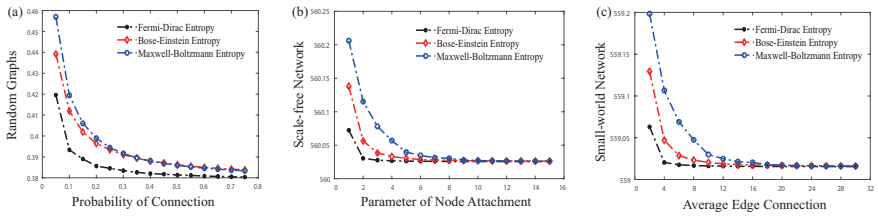


Figure 9: The change of entropy with varying the parameters of three synthetic network models.

small degree, and explore how the edge entropy depends on temperature.

Fig.10(a), (b) and (c) respectively plot the Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac edge entropies versus temperature. The three different curves in the different plots are for edges connecting high, medium and low degree vertices. For the Maxwell-Boltzmann case, the entropy always has a maximum value at a temperature that depends on the node-degree of the edge. For the Bose-Einstein and Fermi-Dirac cases, the low temperature (high β) behaviour depends on the node-degree configuration. In the case of the quantum entropies, there is a local maximum of edge entropy at a particular temperature. In the case of Bose-Einstein statistics, the entropy approaches limiting value with increasing temperature, and this value increases with node degree. For Fermi-Dirac statistics, on the other hand, there is a similar pattern, but the difference in edge entropy at low temperature is more marked.

Next, we explore the relationship between the von Neumann edge entropy and the edge entropies resulting from the three different statistical mechanical models. Fig.11 shows the scatter plots of the statistical mechanical edge entropy versus the corresponding von Neumann edge entropy for samples of edges drawn from a network. The different colours are for different values of the temperature parameter. All three statistical mechanical entropies exhibit a transition in behaviour with respect to the von Neumann entropy with temperature. For instance, in the Maxwell-Boltzmann case, there is a roughly linear relationship between the Maxwell-Boltzmann edge entropy and the von Neumann edge entropy at a high temperature. At low temperature though, the relationship is

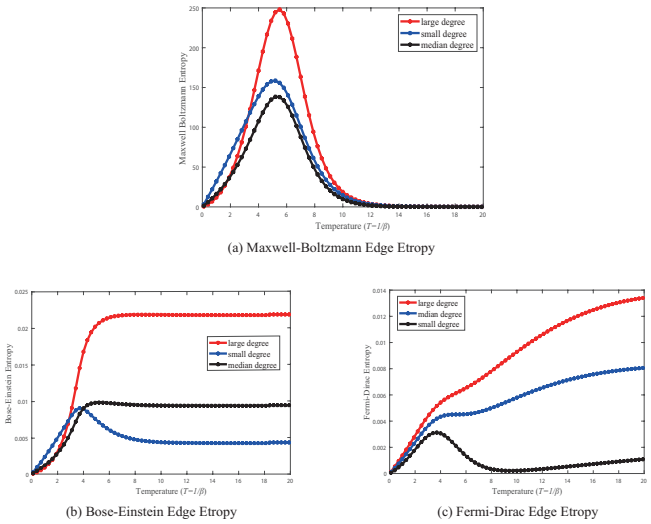
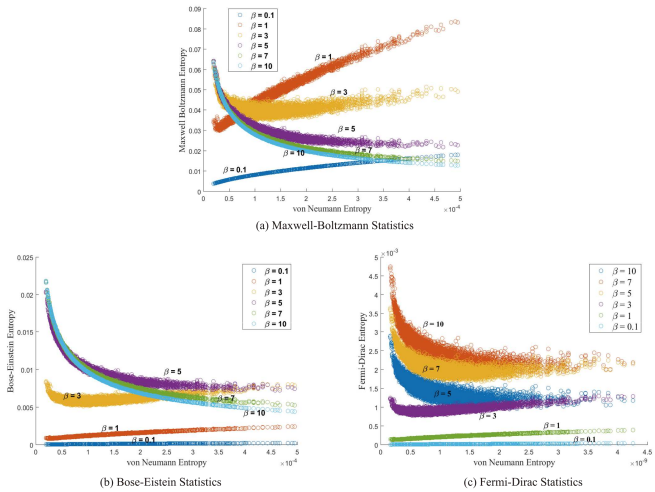


Figure 10: The temperature dependence of edge entropy with different degree configurations. High-degree edge (in red); low-degree edge (in blue); and the median-degree edge (in black). The value of high-degree is 300 and the value of low-degree is 30.

approximately inverse exponential. Whereas the Maxwell-Boltzmann edge entropy increases monotonically with the von Neumann edge entropy. This pattern is repeated for the Bose-Einstein and Fermi-Dirac edge entropies.

The spread of the statistical mechanical edge entropy for a fixed value of the von Neumann edge entropy reveals a number of interesting phenomena. Recall that the von Neumann entropy is completely determined by the degrees of the two nodes defining an edge. In the Bose-Einstein case, the spread is narrow, while in the Fermi-Dirac case it exhibits a broader and more scattered pattern.

This effect is more obvious at high temperature. The narrower the spread of statistical mechanical edge entropies for a given von Neumann entropy, the stronger the dependence on the node degree. At high temperature, the spread of the edge entropy is small, which means it is strongly determined by the degree configuration. This is consistent with our expectation that thermalisation effects disrupt the statistical occupation in the energy states. It is also interesting



(a) Maxwell-Boltzmann Statistics

(b) Bose-Eistein Statistics

(c) Fermi-Dirac Statistics

Figure 11: Scatter plot of edge entropies compared to von Neumann case varying with temperature.

to note that the spread of the spin-dependent entropy is the greatest at low values of the von Neumann entropy. This case corresponds to low degree edge configurations which in turn are associated with the larger variations of entropy with the degree.

In conclusion, all of the statistical mechanical methods for computing edge entropy can be used to effectively represent changes in network structure. The Maxwell-Boltzmann edge entropy is particularly sensitive to structural variance with node degree. The quantum Bose-Einstein and Fermi-Dirac edge entropies exhibit similar properties at high temperature. But at low temperature, the Bose-Einstein edge entropy is more sensitive to the presence of strong community structure in the edge connections, while the Fermi-Dirac edge entropy is more sensitive to the details of the degree distribution.

6. Conclusion

This paper has explored the thermodynamic characterisations of networks using edge entropies computed with both classical Maxwell-Boltzmann statistics

and quantum mechanical Bose-Einstein and Fermi-Dirac spin-statistics. Specifically, we focus on the entropic analysis in both static and time-series network data. This is achieved by decomposing the corresponding global network entropy into contributions for individual edges based on a spectral analysis technique.

Numerical simulations and real-world experiments reveal the comparison between the quantum edge entropies and von Neumann case. Both of the entropic characterisations resulting from Bose-Einstein and Fermi-Dirac statistics effectively identify the detailed variations in network structure. Both of them outperform the traditional von Neumann entropy. Moreover, the corresponding results demonstrate the efficiency of statistical edge entropies for temporal variations in the evolving network structures, and also in distinguishing different network models.

There are a number of ways in which the work reported can be developed further. For instance, the network can be decomposed into motifs, which are frequently recurring subgraphs. It would be interesting to apply our methodology to motif structures, and determine the distribution of edge-entropy over the motif structures. We also intend to explore whether the apparatus of statistical mechanics can be deployed to compute ensemble entropies for individual motifs. One possible approach here would be to use the cluster expansion to compute a motif based partition function and to derive the corresponding entropies.

References

[1] D. Watts, S. Strogatz, Collective dynamics of 'small world' networks, Nature 393 (1998) 440–442.

[2] A.-L. Barabasi, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.

[3] U. Harush, B. Barzel, Dynamic patterns of information flow in complex networks, Nature communications 8 (1) (2017) 2181.

[4] L. Han, E. R. Hancock, R. C. Wilson, Characterizing graphs using approximate von neumann entropy, Pattern Recognition Letter 33 (2012) 1958.

[5] C. Ye, R. C. Wilson, C. H. Comin, L. da F. Costa, E. R. Hancock, Approximate von neumann entropy for directed graphs, Physical Review E 89 (2014).

[6] J. Wang, C. Lin, Y. Wang, Thermodynamic entropy in quantum statistics for stock market networks, Complexity 2019 (2019).

[7] N. De Beaudrap, V. Giovannetti, S. Severini, R. Wilson, Interpreting the von neumann entropy of graph laplacians, and coentropic graphs, A Panorama of Mathematics: Pure and Applied 658 (2016) 227.

[8] F. Passerini, S. Severini, International journal of agent technologies and systems, The von Neumann entropy of networks (2008) 58–67.

[9] P. A. M. Dirac, On the theory of quantum mechanics, Proc. R. Soc. Lond. A 112 (762) (1926) 661–677.

[10] G. Bianconi, Quantum statistics in complex networks, Physical Review E 66 (5) (2002) 056123.

[11] L. Bai, L. Rossi, A. Torsello, E. R. Hancock, A quantum jensen–shannon graph kernel for unattributed graphs, Pattern Recognition 48 (2) (2015) 344–355.

[12] J. Wang, R. C. Wilson, E. R. Hancock, Spin statistics, partition functions and network entropy, Journal of Complex Networks 5 (6) (2017) 858–883.

[13] J. Wang, R. C. Wilson, E. R. Hancock, Thermodynamic network analysis with quantum spin statistics, Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR) (2016) 153–162.

[14] C. Ye, C. H. Comin, T. K. D. Peron, F. N. Silva, F. A. Rodrigues, L. d. F. Costa, A. Torsello, E. R. Hancock, Thermodynamic characterization of networks using graph polynomials, Physical Review E 92 (3) (2015) 032810.

[15] K. Zuev, F. Papadopoulos, D. Krioukov, Hamiltonian dynamics of preferential attachment, Journal of Physics A: Mathematical and Theoretical 49 (10) (2016) 105001.

[16] T. Hartonen, A. Annila, Natural networks as thermodynamic systems, Complexity 18 (2) (2012) 53–62.

[17] J. Wang, R. C. Wilson, E. R. Hancock, Network entropy analysis using the maxwell-boltzmann partition function, 2016 23rd International Conference on Pattern Recognition (ICPR) (2016) 1321–1326.

[18] G. Bianconi, Growing cayley trees described by a fermi distribution, Physical Review E 036116 66 (2002).

[19] G. Bianconi, A.-L. Barabasi, Bose-einstein condensation in complex networks, Physical Review Letter 88 (2001) 5632.

[20] J. Wang, R. C. Wilson, E. R. Hancock, Network edge entropy from maxwell-boltzmann statistics, Image Analysis and Processing - ICIAP 2017 10484 (2017) 254–264.

[21] J. Wang, R. C. Wilson, E. R. Hancock, Quantum edge entropy for alzheimer's disease analysis, in: Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), Springer, 2018, pp. 449–459.

[22] F. Manessi, A. Rozza, M. Manzo, Dynamic graph convolutional networks, Pattern Recognition 97 (2020) 107000.

[23] Y. Zhang, Q. Wang, D.-w. Gong, X.-f. Song, Nonnegative laplacian embedding guided subspace learning for unsupervised feature selection, Pattern Recognition 93 (2019) 337–352.

[24] F. Hu, Y. Liu, Multi-index algorithm of identifying important nodes in complex networks based on linear discriminant analysis, Modern Physics Letters B 29 (03) (2015) 1450268.

[25] L. Wu, C. Shen, A. Van Den Hengel, Deep linear discriminant analysis on fisher networks: A hybrid architecture for person re-identification, Pattern Recognition 65 (2017) 238–250.

[26] A.-L. Barabasi, R. Albert, H. Jeong, Mean-field theory for scale free random networks, Physics A, 272 (1999) 173–187.

[27] C. Von Mering, L. J. Jensen, B. Snel, S. D. Hooper, M. Krupp, M. Foglierini, N. Jouffre, M. A. Huynen, P. Bork, String: known and predicted protein–protein associations, integrated and transferred across organisms, Nucleic acids research 33 (2005) 433–437.

[28] F. Silva, C. Comin, T. Peron, F. Rodrigues, C. Ye, R. C. Wilson, E. R. Hancock, L. Costai, Modular dynamics of financial market networks, Physics and Society arXiv:1501.05040 (2015).

[29] R. C. Petersen, P. Aisen, L. Beckett, M. Donohue, A. Gamst, D. Harvey, C. Jack, W. Jagust, L. Shaw, A. Toga, et al., Alzheimer's disease neuroimaging initiative (adni) clinical characterization, Neurology 74 (3) (2010) 201–209.

[30] G. Del Ferraro, A. Moreno, B. Min, F. Morone, Ú. Pérez-Ramírez, L. Pérez-Cervera, L. C. Parra, A. Holodny, S. Canals, H. A. Makse, Finding influential nodes for integration in brain networks using optimal percolation theory, Nature communications 9 (1) (2018) 2274.

[31] M. Rubinov, O. Sporns, Complex network measures of brain connectivity: uses and interpretations, Neuroimage 52 (3) (2010) 1059–1069.

[32] S. A. Rombouts, F. Barkhof, R. Goekoop, C. J. Stam, P. Scheltens, Altered resting state networks in mild cognitive impairment and mild alzheimer's disease: an fmri study, Human brain mapping 26 (4) (2005) 231–239.

[33] H. Yu, J. Yang, A direct lda algorithm for high-dimensional data—with application to face recognition, Pattern recognition 34 (10) (2001) 2067–2070.