UNIVERSITY of York

This is a repository copy of *Spin Statistics, Partition Functions and Network Entropy*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/118694/

Version: Accepted Version

Article:
Wang, Jianjia, Wilson, Richard Charles orcid.org/0000-0001-7265-3033 and Hancock, Edwin R orcid.org/0000-0003-4496-2028 (2017) Spin Statistics, Partition Functions and Network Entropy. Journal of Complex Networks. pp. 1-25. ISSN 2051-1329

https://doi.org/10.1093/comnet/cnx017

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

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

IMA Journal of Complex Networks (2017) Page 1 of 25
doi:10.1093/comnet/xxx000

Spin Statistics, Partition Functions and Network Entropy

JIANJIA WANG*
Department of Computer Science,University of York, York, YO10 5DD, UK
*jw1157@york.ac.uk

RICHARD C. WILSON
Department of Computer Science,University of York, York, YO10 5DD, UK

AND

EDWIN R. HANCOCK
Department of Computer Science,University of York, York, YO10 5DD, UK

[Received on 2 May 2017]

This paper explores the thermodynamic characterization of networks using the heat bath analogy when the energy states are occupied under different spin statistics, specified by a partition function. Using the heat bath analogy and a matrix characterization for the Hamiltonian operator, we consider the cases where the energy states are occupied according to Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics. We derive expressions for thermodynamic variables, such as entropy, for the system with particles occupying the energy states given by the normalised Laplacian eigenvalues. The chemical potential determines the number of particles at a given temperature. We provide the systematic study of the entropic measurements for network complexity resulting from the different partition functions and specifically those associated with alternative assumptions concerning the spin statistics. Compared with the network von Neumann entropy corresponding to the normalised Laplacian matrix, these entropies are effective in characterising the significant structural configurations and distinguishing different types of network models (Erdős-Rényi random graphs, Watts-Strogatz small world networks, Barabási-Albert scale-free networks). The effect of the spin statistics is a) in the case of bosons to allow the particles in the heat bath to congregate in the lower energy levels and b) in the case of fermions to populate higher energy levels. With normalised Laplacian energy states, this means that bosons are more sensitive to the spectral gap and hence to cluster or community structure, and fermions better sample the distribution of path lengths in a network. Numerical experiments for synthetic and real-world datasets are presented to evaluate the qualitative and quantitative differences of the thermodynamic network characterizations derived from the different occupation statistics, and these confirm these qualitative intuitions.

Keywords: Complex Networks, Network Entropy, Statistical Mechanics, Partition Function

1. Introduction

The literature contains many accounts of work aimed at developing effective characterizations of complex network structure. These characterizations have been widely exploited to both cluster and classify different types of network structure, and also to analyse how networks evolve with time [3, 21, 22, 33]. Broadly speaking, most of the available characterizations have centred around ways of capturing network substructure using clusters, hubs and communities [3, 22, 33]. The underlying representations are usually based on simple degree statistics that capture the connectivity structures [31, 40]. Although

© The author 2017. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.

many of the methods available are goal directed, most promising approaches are to draw on ideas from physics, using analogies based on statistical mechanics [3, 21, 33], thermodynamics [43] or quantum information [5].

One of the most powerful of these approaches is to use thermodynamics analogies suggested by statistical physics. For instance, by maximising the ensemble entropy in exponential random graphs, the Boltzmann distribution from classical statistical mechanics can be used to predict the network properties of time-evolving networks [33]. Tools from statistical mechanics can also be used to characterise the degree distribution for different types of complex networks [3]. Furthermore, by using a heat bath analogy from thermodynamics, principled physical measures of communicability and balance in networks can be defined [22]. Ideas from quantum information theory are also useful in the understanding network structure. For instance, the preferential attachment can lead to the phenomenon of condensation exhibited in growing networks [11]. Both Bose-Einstein and Fermi-Dirac statistics have been used to account for the quantum geometries associated with different types of networks [10]. Although these different physical analogies are useful, they are not always easily related to the graph spectral representation.

Another closely related approach is the heat bath analogy which provides a convenient route to network characterization. Here the energy states of a network are captured using the eigenvalues of a matrix representation of network structure. The energy states are then populated by particles which are in thermal equilibrium with the heat bath. As a result of this thermalization, the energy states are occupied according to the Boltzmann distribution [21, 43]. Formally, this physical heat bath system can be described by a partition function with the energy microstates of the network represented by a suitably chosen Hamiltonian. Usually, the Hamiltonian is computed from the adjacency or Laplacian matrix of the network, but recently, Ye et al.[43], have shown how the partition function can be computed from a characteristic polynomial instead.

To embark on this type of analysis, partition functions can be succinctly used to describe the network statistics and evolution. Thermodynamic characterizations of the network, such as entropy, total energy, and temperature can then be derived from the partition functions [31, 43]. By specifying the microstates of the network system, statistical thermodynamics can provide deep insights into network behaviour. For example, by using the Maxwell-Boltzmann partition function describing a thermalized network, the entropy, internal energy, and the Helmholtz free energy can be computed from the graph spectra, and this leads to natural definitions of notions such a centrality [21, 43].

However, the Boltzmann distribution does not take into account particle spin-statistics and their effects on the population of the thermalized energy levels. Unlike the classical case where particles are distinguishable, in quantum statistics particles are indistinguishable. Particles with integer spin are subject to Bose-Einstein statistics and do not obey the Pauli exclusion principle. As a result, they can aggregate in the same energy state. At low temperature, this leads to the phenomenon of Bose-Einstein condensation. There has been work aimed at extending the heat-bath model to take such effects into account. For instance, Bianconi and Barabási [11] have constructed a network model based on a Bose gas, and have studied the phase transitions in network structure associated with Bose-Einstein condensation [11]. This model has also been extended to understand processes such as supersymmetry in networks [10, 22]. On the other hand, particles with half-integer spin are subject to Fermi-Dirac statistics and obey the Pauli exclusion principle. They thus give rise to very different models of network structure, and these have been exploited to model situations where there are constraints on the occupancy of the nodes and edges of a network. Examples include traffic flow and also the modelling of certain types of geometric networks such as the Cayley tree [9, 36].

Despite the interest in alternative models of the thermalized distribution of energy states under dif-

ferent particle spin statistics, there has been no systematic study of the various thermodynamic characterizations resulting from different choices of partition functions, and specifically those associated with alternative assumptions concerning the spin statistics. Here we consider the effects of occupation statistics on the populations of energy states when the Hamiltonian operator is the normalised network Laplacian, and the energy states are then given by its spectrum. Commencing from the heat bath analogy with the Laplacian matrix playing the role as the Hamiltonian, the energy states of the system are occupied according to a) Maxwell-Boltzmann, b) Bose-Einstein and c) Fermi-Dirac statistics respectively. From the relevant partition function, we use the statistical mechanical properties of the networks to compute various thermodynamic quantities when the energy levels are occupied by particles in thermal equilibrium with the heat bath. Making different choices for the partition function, we obtain different occupation statistics for the energy levels. The network then can be characterised using thermodynamic quantities such as the entropy and energy derived from the relevant partition function [21, 43]. In qualitative terms, the Pauli exclusion principle means that particles subject to Fermi-Dirac statistics are populating the energy states less densely than that in the classical Maxwell-Boltzmann case. On the other hand, since particles obeying Bose-Einstein are indistinguishable, they populate the energy states more densely.

The thermodynamic picture offered by quantum Bose-Einstein and Fermi-Dirac statistics differs from that offered by classical Maxwell-Boltzmann statistics in a number of important ways. Both quantum statistics additionally require a chemical potential to specify the distribution of states in the partition function. The chemical potential is determined by the heat reservoir, and modifies the occupation probability of the energy levels [12]. In the case of both Bose-Einstein and Fermi-Dirac statistics, for energy levels greater than the chemical potential, the occupation probability is increased. In other words, at a given temperature the higher energy levels are more likely to be occupied in the quantum case than in the classical case. The difference between Fermionic and Bosonic statistics also manifests itself in important ways. For instance, at low temperatures where there is little thermal disruption of the occupation pattern dictated by the Pauli exclusion principle, Bosons tend to condense in the lowest energy states, while there is just one Fermion per energy state [12]. As a result, thermodynamic quantities such as the total energy or entropy of the system sample the spectrum of Laplacian energy states in different ways, and potentially convey different aspects of network structure. For instance, at low temperature under Bose-Einstein statistics the particles in the heat bath are are likely to respond more strongly to the spectral gap (the difference between the zero and first non-zero normalised Laplacian eigenvalues) and are thus sensitive to cluster or community structure[10]. Fermi-Dirac statistics, on the other hand, are sensitive to a larger portion of the spectrum and are more sensitive to the density of energy states [9]. As a result, they are more sensitive to the details of the degree distribution and also to structural artifacts requiring more information concerning the Laplacian spectrum such as the path length and cycle length distributions. Viewed from the perspective of the distribution of energy levels, as represented by the Laplacian spectrum, different types of commonly studied network model, e.g. random or Erdos-Renyi, small-world and scale-free, have very different degree distributions and this is reflected in both their Laplacian and normalised Laplacian spectra. The different energy level population statistics resulting from the different partition functions can be expected to result in them having different sensitivity to different network models.

The aim of this paper is to explore the behaviour of the entropy and total energy of networks resulting from different choices of partition functions. We compare four different entropic network characterizations. The first three result from the partition functions for a) Maxwell-Boltzmann, b) Bose-Einstein and c) Fermi-Dirac occupation statistics, while the fourth is the von Neumann entropy associated with the normalised Laplacian matrix of the network [26, 34, 42]. We explore how these different entropies

can be used to characterise the changes of network structure with time, and distinguish different types of network models (Erdős-Rényi random graphs, small world networks [40], and scale-free networks [7]).

The remainder of the paper is organised as follows. In Sec.II, we provide some background material concerning the choice of Hamiltonian operator and the von Neumann entropy. With these concepts to hand, in Sec.III, we provide a review of the relationship between the partition function and the thermodynamic variables, i.e. the average energy, thermodynamic entropy, Helmholtz free energy, temperature and chemical potential. In Sec.IV, we provide a detailed analysis of the entropies resulting from the three different choices of partition functions and explore their low and high-temperature limits. In Sec.V, numerical experiments on synthetic and real-world datasets are used to evaluate the effectiveness of the different thermodynamic network characterizations. Finally, in Sec.VI, we conclude the paper and make suggestions for future work.

2. Graph Representation

2.1 *Density Matrix*

In quantum mechanics the density matrix is used to describe a system whose state is an ensemble of pure quantum states $|\psi_i\rangle$, each with probability p_i. The density matrix is defined as

$$\boldsymbol{\rho} = \sum_{i=1}^{V} p_i |\psi_i\rangle\langle\psi_i| \tag{2.1}$$

Severini et al. [5, 34] have extended this idea to the graph domain. Specifically, they show that a density matrix for a graph or network can be obtained by scaling the combinatorial Laplacian matrix by the reciprocal of the number of nodes in the graph.

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

$$A = \begin{cases} 0 & \text{if } (u,v) \in E \\ 1 & \text{otherwise.} \end{cases} \tag{2.2}$$

Then the degree of node u is $d_u = \sum_{v \in V} A_{uv}$.

The normalised Laplacian matrix $\tilde{L}$ of the graph G is defined as

$$\tilde{L} = D^{-\frac{1}{2}} L D^{\frac{1}{2}} \tag{2.3}$$

where $L = D - A$ is the Laplacian matrix and D denotes the degree diagonal matrix whose elements are given by $D(u,u) = d_u$ and zeros elsewhere. The element-wise expression of $\tilde{L}$ is

$$\tilde{L}_{uv} = \begin{cases} 1 & \text{if } u = v \text{ and } d_u \neq 0 \\ -\frac{1}{\sqrt{d_u d_v}} & \text{if } u \neq v \text{ and } (u,v) \in E \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

With this notation, the specified density matrix is obtained by scaling the normalised Laplacian matrix by the number of nodes, i.e.

$$\boldsymbol{\rho} = \frac{\tilde{L}}{|V|} \tag{2.5}$$

When defined in this way the density matrix is Hermitian i.e. $\boldsymbol{\rho} = \boldsymbol{\rho}\dagger$ and $\boldsymbol{\rho} \geqslant 0, \mathrm{Tr}\boldsymbol{\rho} = 1$. It plays an important role in the quantum observation process, which can be used to calculate the expectation value of measurable quantity.

2.2 *Hamiltonian Operator of a Graph*

In quantum mechanics, the Hamiltonian operator is the sum of the kinetic energy and potential energy of all the particles in the system, and it dictates the Schrödinger equation for the relevant system. The Hamiltonian is given by

$$\hat{H} = -\nabla^2 + U(r,t) \tag{2.6}$$

The Hamiltonian operator of a graph can be defined in a number of ways. For example, if we specify the node potential energy as the degree matrix, i.e. $U(r,t) = D$, and replace the Laplacian by its combinational counterpart $L = D - A$ with $L = -\nabla^2$, then $\hat{H} = -A$. This Hamiltonian operator is often used in the Hückel molecular orbital (HMO) method [38].

Alternatively, we can consider the case where the graph is assumed to be in contact with a heat reservoir. In this case, the eigenvalues of the Laplacian matrix can be viewed as the energy eigenstates, and these determine the Hamiltonian and hence the relevant Schrödinger equation which governs a system of particles. The particles occupy the energy states of the Hamiltonian subject to thermal agitation by the heat bath. The number of particles in each energy state is determined by the temperature, the assumed model of occupation statistics and the relevant chemical potential.

From the Hückel picture above, the Hamiltonian operator is the Laplacian matrix on graph. Similarly, the normalised form of the graph Laplacian can be viewed as the Hamiltonian operator

$$\hat{H} = \tilde{L} \tag{2.7}$$

In this case, the energy states of the network $\{\varepsilon_i\}$ are then the eigenvalues of the Hamiltonian

$$\hat{H}|\psi_i\rangle = \tilde{L}|\psi_i\rangle = E_i|\psi_i\rangle \tag{2.8}$$

The eigenvalues are all greater than or equal to zero, and the multiplicity of the zero eigenvalue is the number of connected components in the network. Furthermore, the density matrix commutes with the Hamiltonian, i.e. the associated Poisson bracket is zero,

$$[\hat{H}, \rho] = [\tilde{L}, \frac{\tilde{L}}{|V|}] = 0 \tag{2.9}$$

which means that the network is in equilibrium when there are no changes in the density matrix which describes the system.

2.3 *von Neumann Entropy*

From the information theoretic perspective, entropy has proved to be an effective means of characterising the structure of complex networks. There have been many attempts to compute the entropy of a graph, and a recent review of the available methods is given in [34]. One of the obstacles to this endeavour is that methods based on a combinatorial analysis of the graph, soon become intractable as its size becomes large.

However, if a probability distribution can be simply constructed over the graph so that its Shannon entropy can be evaluated, then the computation of entropy is more tractable. Structures such as cycles and fully connected graphs have been shown to correspond to the extremal values of certain network entropies.

The interpretation of the scaled normalised Laplacian as a density operator, opens up the possibility of characterising a graph using the von Neumann entropy from quantum information theory. The von Neumann entropy is defined as the entropy of the density matrix associated with the state vector of a system. As noted above, Severini et al. [5, 34] suggest how the von Neumann entropy can commuted by scaling the normalised discrete Laplacian matrix for a network. As a result the von Neumann entropy is given in terms of the eigenvalues λ_1, ..., λ_V of the density matrix $\boldsymbol{\rho}$,

$$S = -\mathrm{Tr}(\boldsymbol{\rho}\log\boldsymbol{\rho}) = -\sum_{i=1}^{|V|}\frac{\hat{\lambda}_i}{|V|}\log\frac{\hat{\lambda}_i}{|V|} \tag{2.10}$$

The von Neumann entropy [5] computed from the normalised Laplacian spectrum has been shown to be effective for network characterization. In fact, Han et al.[26] have shown how to approximate the calculation of von Neumann entropy in terms of simple degree statistics. Their approximation allows the cubic complexity of computing the von Neumann entropy from the Laplacian spectrum, to be reduced to one of quadratic complexity using simple edge degree statistics, i.e.

$$S = 1 - \frac{1}{|V|} - \frac{1}{|V|^2}\sum_{(u,v)\in E}\frac{1}{d_u d_v} \tag{2.11}$$

This expression for the von Neumann entropy allows the approximate entropy of the network to be efficiently computed and has been shown to be an effective tool for characterising structural property of networks, with extremal values for the cycle and fully connected graphs. Ye et al. [42] have extended this result to directed graphs by distinguishing between the in-degree and out-degree of nodes.

3. Thermodynamic Representation of Networks

Thermodynamic analogies provide powerful tools for analysing complex networks. The underpinning idea is that statistical thermodynamics can be combined with network theory to characterise both static and time-evolving networks [31].

A complex network can be viewed as a grand canonical ensemble, which not only exchanges energy but also particles with a heat reservoir. In general, the energy and entropy of the network depend on the assumptions concerning the Hamiltonian for the system and the corresponding partition function.

In this section we present the relationship between the choice of partition function and the resulting energy and entropy of the network when characterised in terms of the heat-bath analogy. These results are largely standard bookwork and make no direct reference specifically to network structure other than making use of the normalised Laplacian as Hamiltonian , but are presented here for completeness since they are used in our numerical experiments. That said, the matrix forms of the expressions for energy and entropy less well known, and are important since they allow us to make easy contact between statistical mechanics and spectral graph theory.

3.1 *Statistical Mechanics*

Here we consider the thermodynamic system specified by a system of N particles with energy states given by the network Hamiltonian and immersed in a heat bath with temperature T. The ensemble is represented by a partition function $Z(\beta,N)$, where $\beta = 1/k_BT$ is an inverse of temperature parameter. When specified in this way, the various thermodynamic characterizations of the network can be computed. For instance, the average energy of the network can be expressed in terms of the density matrix and the Hamiltonian operator,

$$\langle U\rangle = \langle H\rangle = \mathrm{Tr}\,(\rho H) = \left[-\frac{\partial}{\partial\beta}\log Z\right]_N \tag{3.1}$$

and the thermodynamic entropy by

$$S = k_B\left[\log Z + \beta\langle U\rangle\right] \tag{3.2}$$

To compute quantum occupation statistics, we also require the chemical potential, which is given by

$$\mu = -\frac{1}{\beta}\left[\frac{\partial}{\partial N}\log Z\right]_\beta \tag{3.3}$$

The chemical potential is a measure of how resistive the system is to the addition of new particles. It acts to offset the energy levels of the Hamiltonian. In the case of Fermi-Dirac statistics, the chemical potential is equal to the Fermi level and at zero temperature determines the highest occupied energy state. In Bose-Einstein statistics, the chemical potential tends to zero at zero temperature, and this leads to the formation of the Bose-Einstein condensate. In the remainder of the paper, we set the Boltzmann constant $k_B = 1$.

Both the energy and the entropy can be regarded as weighted functions of the Laplacian eigenvalues which characterise the network structure in different ways. In the following sections, we will explore these differences in more detail, and in particular to which parts of the Laplacian spectrum they are most sensitive to different choices of the partition function resulting from different occupation statistics.

4. Partition Functions and Occupation Statistics

According to the picture adopted in this paper, the normalised Laplacian of the graph specifies a series of energy states that can be occupied by particles. At a given temperature, there are a number of alternative ways in which the energy levels can be occupied, depending on the spin-statistics of the particles.

Here we consider the different situations that arise when the occupation of the energy levels is governed by Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics. The Maxwell-Boltzmann distribution applies when spin statistics are ignored, and the population of the different energy levels is governed by thermalization. Bose-Einstein statistics apply to bosons of integer spin, and which are indistinguishable. Finally, Fermi-Dirac statistics apply when the particles are fermions with half-integer spin and are subject to the Pauli exclusion principle.

For each distribution, we capture the statistical mechanical properties of particles in the system using the partition function associated with the different occupation statistics. The network can then be characterised using thermodynamic quantities computed from the partition function, and these include the entropy, energy, and temperature.

4.1 *Maxwell-Boltzmann Statistics*

In statistical mechanics, the Maxwell-Boltzmann distribution relates the microscopic properties of particles to the macroscopic thermodynamic properties of matter [43]. It applies to systems consisting of a fixed number of weakly interacting distinguishable particles. These particles occupy the energy levels associated with a Hamiltonian and in our case the Hamiltonian of the network, which is in contact with a thermal bath [33].

Taking the Hamiltonian to be the normalised Laplacian of the network, the canonical partition function for Maxwell-Boltzmann occupation statistics of the energy levels is

$$Z_{MB} = \mathrm{Tr}\left[\exp(-\beta\tilde{L})^N\right] = \left(\sum_{i=1}^{|V|} e^{-\beta\varepsilon_i}\right)^N \tag{4.1}$$

where $\beta = 1/k_BT$ is the reciprocal of the temperature T with k_B as the Boltzmann constant; N is the total number of particles and ε_i denotes the microscopic energy of system at each microstate i. Furthermore, from Eq.(3.1), the average energy of the network is

$$\langle U\rangle_{MB} = N\frac{\mathrm{Tr}[\tilde{L}\exp(-\beta\tilde{L})]}{\mathrm{Tr}[\exp(-\beta\tilde{L})]} = N\frac{\sum_{i=1}^{|V|}\varepsilon_i e^{-\beta\varepsilon_i}}{\sum_{i=1}^{|V|}e^{-\beta\varepsilon_i}} \tag{4.2}$$

and similarly derived from Eq.(3.2), the entropy of the system with N particles is

$$S_{MB} = -N\mathrm{Tr}\left\{\frac{\exp(-\beta\tilde{L})}{\mathrm{Tr}[\exp(-\beta\tilde{L})]}\log\frac{\exp(-\beta\tilde{L})}{\mathrm{Tr}[\exp(-\beta\tilde{L})]}\right\} = -N\sum_{i=1}^{|V|}\frac{e^{-\beta\varepsilon_i}}{\sum_{i=1}^{|V|}e^{-\beta\varepsilon_i}}\log\frac{e^{-\beta\varepsilon_i}}{\sum_{i=1}^{|V|}e^{-\beta\varepsilon_i}} \tag{4.3}$$

For a single particle, the density matrix is

$$\boldsymbol{\rho}_{MB} = \frac{\exp(-\beta\tilde{L})}{\mathrm{Tr}[\exp(-\beta\tilde{L})]} \tag{4.4}$$

Since the density matrix commutes with the Hamiltonian operator, we have $\partial\rho/\partial t = 0$ and the system can be viewed as in equilibrium. So the entropy in the Maxwell-Boltzmann system is simply N times the von Neumann entropy of a single particle, as we might expect.

4.2 *Bose-Einstein Statistics*

The Bose-Einstein distribution applies to indistinguishable bosons. Each energy state can accommodate an unlimited number of particles specified by the network Hamiltonian. Bosonic particles subject to Bose-Einstein statistics do not obey the Pauli exclusion principle and can aggregate in the same energy state. Complex networks have been successfully characterised using systems of bosons to capture network topology. For instance, Bianconi and Barabási [11] have constructed a network model based on a Bose gas, and have studied the phase transitions in network structure associated with the Bose-Einstein condensation of the gas. This model has also been extended to understand processes such as supersymmetry in networks [10].

For a grand-canonical ensemble with a varying number of particles and a chemical potential μ, the Bose-Einstein partition function is

$$Z_{BE} = \det\left(I - e^{\beta\mu}\exp[-\beta\tilde{L}]\right)^{-1} = \prod_{i=1}^{|V|}\left(\frac{1}{1-e^{\beta(\mu-\varepsilon_i)}}\right) \tag{4.5}$$

which gives the average energy as

$$\langle U\rangle_{BE} = -\mathrm{Tr}\left\{\left[I - e^{\beta\mu}\exp(-\beta\tilde{L})\right]^{-1}(\mu I - \tilde{L})e^{\beta\mu}\exp(-\beta\tilde{L})\right\} = -\sum_{i=1}^{|V|}\frac{(\mu-\varepsilon_i)e^{\beta(\mu-\varepsilon_i)}}{1-e^{\beta(\mu-\varepsilon_i)}} \tag{4.6}$$

and entropy

$$\begin{aligned} S_{BE} &= -\mathrm{Tr}\left\{\log[I - e^{\beta\mu}\exp(-\beta\tilde{L})]\right\} - \mathrm{Tr}\left\{\beta[I - e^{\beta\mu}\exp(-\beta\tilde{L})]^{-1}(\mu I - \tilde{L})e^{\beta\mu}\exp(-\beta\tilde{L})\right\} \\ &= -\sum_{i=1}^{|V|}\log\left(1-e^{\beta(\mu-\varepsilon_i)}\right) - \beta\sum_{i=1}^{|V|}\frac{(\mu-\varepsilon_i)e^{\beta(\mu-\varepsilon_i)}}{1-e^{\beta(\mu-\varepsilon_i)}} \end{aligned} \tag{4.7}$$

As a result the average energy is the average difference between the Laplacian energy states and the chemical potential, weighted by the Bose-Einstein factor $\exp[-\beta(\varepsilon_i-\mu)]/(1-\exp[-\beta(\varepsilon_i-\mu)])$. The weighted energy difference therefore decreases with energy. The entropy also decreases with the energy of the states.

4.3 *Fermi-Dirac Statistics*

The Fermi-Dirac distribution applies to indistinguishable fermions with a maximum occupancy of one particle in each energy state. Particles cannot be added to states that are already occupied, and hence obey the Pauli exclusion principle. These particles behave like a set of free fermions in the complex network with energy states given by the network Hamiltonian. The statistical properties of the networks are thus given by the Fermi-Dirac distribution of the equivalent quantum system. The corresponding partition function is

$$Z_{FD} = \det\left(I + e^{\beta\mu}\exp[-\beta\tilde{L}]\right) = \prod_{i=1}^{|V|}\left(1+e^{\beta(\mu-\varepsilon_i)}\right) \tag{4.8}$$

giving average energy

$$\begin{aligned} \langle U\rangle_{FD} &= -\mathrm{Tr}\left\{\left[I + e^{\beta\mu}\exp(-\beta\tilde{L})\right]^{-1}(\mu I - \tilde{L})e^{\beta\mu}\exp(-\beta\tilde{L})\right\} \\ &= -\sum_{i=1}^{|V|}\frac{(\mu-\varepsilon_i)e^{\beta(\mu-\varepsilon_i)}}{1+e^{\beta(\mu-\varepsilon_i)}} \end{aligned} \tag{4.9}$$

and entropy

$$\begin{aligned} S_{FD} &= \mathrm{Tr}\left\{\log[I + e^{\beta\mu}\exp(-\beta\tilde{L})]\right\} - \mathrm{Tr}\left\{\beta[I + e^{\beta\mu}\exp(-\beta\tilde{L})]^{-1}(\mu I - \tilde{L})e^{\beta\mu}\exp(-\beta\tilde{L})\right\} \\ &= \sum_{i=1}^{|V|}\log\left(1+e^{\beta(\mu-\varepsilon_i)}\right) - \beta\sum_{i=1}^{|V|}\frac{(\mu-\varepsilon_i)e^{\beta(\mu-\varepsilon_i)}}{1+e^{\beta(\mu-\varepsilon_i)}} \end{aligned} \tag{4.10}$$

As the result the average energy is the average difference between the Laplacian energy states and the chemical potential, this time weighted by the Fermi-Dirac factor $\exp[-\beta(\varepsilon_i-\mu)]/(1+\exp[-\beta(\varepsilon_i-\mu)])$. As a result , for a given chemical potential the higher energy levels receive more weight than in the case of the Bose-Einstein statistics. Moreover, the entropy associated with the states peaks at the chemical potential.

4.4 *Particle Population and Chemical Potential*

We need to specify how the system associated with different partition functions is populated at various temperatures, and how we set the chemical potential in the case of Bose-Einstein and Fermi-Dirac statistics. Our approach is to compute the number of particles occupying each energy state, and sum over the different energy states.

In the case of Maxwell-Boltzmann statistics, the number of particles in the state with energy ε is

$$n_i = N\frac{e^{-\beta\varepsilon_i}}{Z_{MB}} = N\frac{\exp(-\beta\tilde{L})}{\mathrm{Tr}[\exp(-\beta\tilde{L})]} \tag{4.11}$$

and so the total number of particles is

$$N = \sum_{i=1}^{|V|} n_i \tag{4.12}$$

In the case of both Bose-Einstein and Fermi-Dirac occupation statistics, the partition function and hence both the average energy and entropy, depend on the chemical potential. This parameter is determined by the number of particles in the system and the temperature.

For Bose-Einstein statistics at the temperature corresponding to β, in order for the number of particles in each energy state to be non-negative, the chemical potential must be less than the minimum energy level, i.e. $\mu < \min\varepsilon_i$. Under Fermi-Dirac statistics, on the other hand. with a single particle per energy state, the chemical potential is hence just the nth energy level, and so $\mu = \varepsilon_n$.

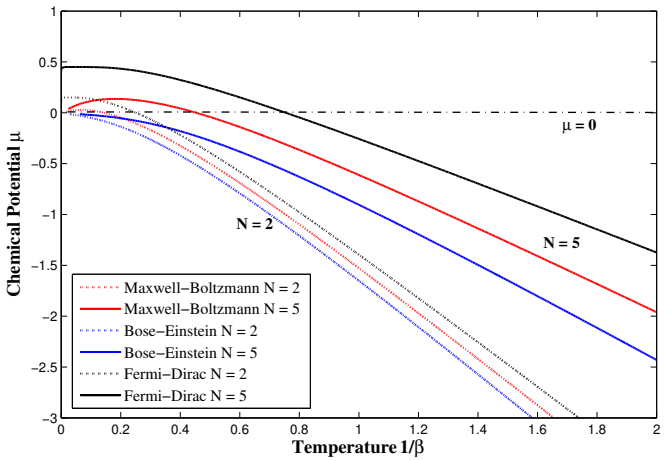


FIG. 1: (Colour online) Plot of the chemical potential μ versus temperature T for Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics. For convenience in performing the comparison, the numbers of particles are N = 2 and N = 5, with energy states uniformly distributed between 0 and 2. In the high-temperature region, the three chemical potentials exhibit similar behaviour. In the low-temperature region, the chemical potential for Bose-Einstein statistics is always less than 0. However, with Fermi-Dirac statistics, it is larger than 0 and increases with the number of particles N.

4.5 *High and Low Temperature Limits*

4.5.1 *High Temperature Limits* ($\beta \to 0$) At high temperature, i.e. when β approaches zero, thermalization disrupts the effects of the occupation statistics captured by different partition functions, and both the Bose-Einstein and Fermi-Dirac models are equivalent to the Maxwell-Boltzmann case. For the Maxwell-Boltzmann distribution, the high-temperature limit of the average energy is $\lim_{\beta\to 0}\langle U\rangle_{MB} = \frac{N}{|V|}\mathrm{Tr}[\tilde{L}]$, which is as expected proportional to the trace of the normalised Laplacian, giving an average energy per particle of $\frac{1}{|V|}Tr[\tilde{L}]$. The corresponding high-temperature limit of the entropy is

$$\lim_{\beta\to 0} S_{MB} = N\log|V| + \frac{N\beta^2}{2|V|}\left\{\frac{1}{|V|}\mathrm{Tr}[\tilde{L}] - \mathrm{Tr}[\tilde{L}^2]\right\} \tag{4.13}$$

This is similar to the result obtained by Han et al. [26] for the von Neumann entropy. As a result, the entropy at high temperature is a constant for all three models of the occupation statistics.

4.5.2 *Low Temperature Limits* ($\beta \to \infty$) The low temparature limits of the energy and entropy under Maxwell-Boltzmann statistics, when $\beta \to \infty$ are $\lim_{\beta\to+\infty}\langle U\rangle_{MB} = 0$ and $\lim_{\beta\to+\infty} S_{MB} = N\log c$, where c is the number of connected components in the network. We usually deal with graphs having a single connected component and as a result we have that the limit of entropy in Maxwell-Boltzmann case at the low temperature tends to zero.

In the case of the Bose-Einstein and Fermi-Dirac partition functions, the chemical potential plays a pivotal role in determining the low-temperature limit.

For Fermi-Dirac statistics, the constant chemical potential μ is equal to the energy of highest state occupied by one of the N particles at zero temperature. With a single particle per energy state, this is hence just the Nth energy level, namely ε_N. As the temperature approaches zero, the chemical potential μ approaches the Fermi energy ε_N, so that $\mu = \varepsilon_N$. There is only one configuration for each identical particle occupies at each energy state, and the corresponding entropy is $\lim_{\beta\to+\infty} S_{FD} = 0$.

For Bose-Einstein occupation statistics, at $T = 0$ all particles are in the ground state and it is straightforward to show that $\lim_{\beta\to+\infty} S_{BE} = (N+1)\ln(N+1) - N\ln N$. As N goes to infinity, the limits of entropy tends to $\lim_{N\to+\infty}\lim_{\beta\to+\infty} S_{BE} = \ln N$. The main difference between the thermal quantities of the classical statistical system and that of the quantum spin systems is that the partition function results in different occupation of the energy levels according to the relevant population statistics.

In the Maxwell-Boltzmann case, without thermalization of the levels at zero temperature, all particles occupy the zero energy ground state. But in the case of Bose-Einstein and Fermi-Dirac statistics, this pattern is modified by the chemical potential, and this modified the way which the higher energy levels are populated. For Bose-Einstein statistics, the effect is to shift the occupation number from the zero energy Maxwell-Boltzmann ground-state by an amount proportional to the chemical potential. In other words, the particles are found with higher probabilities at lower energy levels. In the case of the Fermi-Dirac statistics, the effect is exaggerated since the chemical potential is the energy of the state corresponding to the number of particles in the system.

5. Physical Intuitions

The graph Laplacian defines a set of energy levels for a system which is in thermodynamic equilibrium with a heat bath of known temperature. The different partition functions govern how a system of non-interacting particles populate these energy levels at a particular temperature. From the partition

functions, we can calculate the energy and entropy associated with the system of particles, at a particular temperature and for different numbers of particles. Our idea is to use these two thermodynamic quantities to characterise the network from which the Laplacian was computed.

Of course, different networks will have different graph spectra (i.e. distributions of Laplacian eigenvalues or energy levels of our thermodynamic system), and this, in turn, will give rise to a different population of energy levels with temperature. More importantly, in this study, the choice of partition function will also control how the different energy levels can be populated depending on the spin statistics of the particles, and the number of particles added to the system. When we work with the classical Maxwell-Boltzmann distribution, then the temperature is the only controlling parameter. By increasing temperature, we simply thermalize the population of the energy levels. On the other hand, when we envoke non-classical spin statistics, quantum effects become evident. In the case of Fermi-Dirac statistics, only one particle can occupy each energy state. For Bose-Einstein statistics, on the other hand, particles can condense in the lower energy states, particularly at low temperatures, but these particles are indistinguishable, leading to different statistics. In the quantum cases, the effect of changing the number of particles can be modelled by adding a chemical potential which effectively shifts the energy levels.

We use the entropy and energy associated with the distribution of energy levels and their different occupation probabilities to explore whether the different partition functions allow us to probe differences in graph structure in different and hopefully more useful ways. The main interest here lies in the low-temperature behaviour since at high temperature the effects of the quantum statistics are disturbed by thermal effects all three partition functions give identical results. At low temperature, we are more likely to find bosonic particles in the low energy states when compared to the Maxwell-Boltzmann distribution. On the other hand, because of the Pauli exclusion principle, we are more likely to find fermions at higher energies. Hence by populating the energy states in different ways, the particles respond to the Laplacian spectrum in different ways depending on which of the three partition functions governs their behaviour. The question we seek to answer in this paper is when measured in terms of their entropy or energy do the different partition functions allow us to probe graph structure in different ways.

It is well known that different types of network have different degree distributions, and this is reflected in their Laplacian spectra. For instance, Erdős-Rényi or random graphs the eigenvalues follow a semi-circular (or Wigner) distribution, with mean controlled by the connection probability. Scale-free networks have a triangular distribution and graphs of the Watts-Strogatz type have a more complex spectrum which depends on the parameters and may contain sharp peaks. For cluster structure, the distribution of the lowest eigenvalues and the spectral gap are most important. Hence, the choice of how the eigenvalues are sampled, or choice of the partition function, can be sensitive to the type of structure. One might, for instance, expect Bose-Einstein statistics to be better suited to detecting networks with strong community structure because they preferentially sample the lower energy levels. Fermi-Dirac statistics, on the other hand, may be better for distinguishing different network models because they probe a wider range of energy levels, and are hence more sensitive to the mean and variance of the eigenvalue distribution.

6. Experiments and Evaluations

We explore whether the thermodynamic characterizations resulting from the three alternative models for the energy level occupation statistics can be employed as a useful tool for better understanding the structural properties and the evolution of networks. Specifically, we numerically simulate the effects of the three different models and examine whether the resulting entropies can distinguish different structures,

and compare their relative performance. Furthermore, we compute the thermodynamic characterizations for a number of real-world time-evolving networks in order to investigate whether they can be used to detect abrupt changes in network structure at different time epochs. Finally, we use the different entropies to classify tumour mutation networks and protein to protein interaction networks resulting from different groups. To simplify the calculation, we set the Boltzmann constant to unity throughout our experiments.

This section is structured as follows. We commence by describing the data-sets used in our experiments and the kernel principal components method used to visualise the entropy differences between different networks. We then explore how the different entropies depend on their free parameters, name temperature of a number of particles for network drawn from different models (Erdős-Rényi, small world and scale-free). Using kernel PCA, we visualise how the networks from different models distribute themselves in three dimensions using the different partition functions and comment on which gives the best separation. Finally, we report results on real world data-sets.

6.1 *Data Sets*

Here, we use four different data sets. The first contains synthetically generated artificial networks, while the remaining three are extracted from real-world complex systems.

Data-set 1: Contains a large number of graphs which are randomly generated according to one of three different complex network models, namely, a) the classical Erdős-Rényi random graph model, b) the small-world model introduced by Watts and Strogatz [40], and c) the scale-free model, developed by Barabási-Albert model [7, 8]. These are created using a variety of model parameters, e.g., the graph size and the connection probability in the random graph model, the link rewiring probability [40] in the small-world model and the number of added connections at each time step [7] in the scale-free model.

Data-set 2: The New York Stock Exchange dataset consists of the daily prices of 3,799 stocks traded continuously on the New York Stock Exchange over 6000 trading days. The stock prices were obtained from the Yahoo! financial database (http://finance.yahoo.com) [37]. A total of 347 stock were selected from this set, for which historical stock prices from January 1986 to February 2011 are available. In our network representation, the nodes correspond to stock and the edges indicate that there is a statistical similarity between the time series associated with the stock closing prices [37]. To determine the edge structure of the network, we use a time window of 20 days is to compute the cross-correlation coefficients between the time-series for each pair of stock. Connections are created between a pair of stock if the cross-correlation exceeds an empirically determined threshold. In our experiments, we set the correlation coefficient threshold to the value to $\xi = 0.85$. This yields a time-varying stock market network with a fixed number of 347 nodes and varying edge structure for each of 6,000 trading days. The edges of the network, therefore, represent how the closing prices of the stock follow each other.

Data-set 3: Contains tumour mutation data for three major cancers taken from the Cancer Genome Atlas (TCGA). These are a) ovarian cancer b) uterine cancer and c) lung adenocarcinoma [1]. There are 356 patients with mutations in 9,850 genes in the ovarian cancer cohort, 248 patients with mutations in 17,968 genes in the uterine endometrial cancer cohort and 381 patients with mutations in 15,967 genes in the lung adenocarcinoma cohort [28]. The raw patient mutation data are binary vectors, with elements corresponding to different genes. The binary numbers indicate if the relevant gene is mutated or not (1 indicates the presence of a mutation, 0 that a mutation is absent). So each individual is characterised by a 0-1 binary gene sequence of mutation indicators. Patient mutation networks were mapped onto gene interaction networks by aggregating information from several pathways and interaction databases,

describing physical protein-protein interactions (PPIs) and functional relationships between genes in both regulatory, signalling and metabolic pathways [23].

Data-set 4: The PPIs dataset extracted from STRING8.2 [2] consisting of networks which describe the interaction relationships between histidine kinase and other proteins. Histidine kinase is a key protein in the development of signal transduction. If two proteins have direct (physical) or indirect (functional) association, they are connected by an edge. There are 173 PPIs in this dataset and they are collected from 4 different kinds of bacteria with the following evolution order (from older to more recent). Aquifex and Thermotoga-8 PPIs from Aquifex aelicus and Thermotoga maritima, Gram-Positive-52 PPIs from Staphylococcus aureus, Cyanobacteria-73 PPIs from Anabaena variabilis and Proteobacteria-40 PPIs from Acidovorax avenae [20].

6.2 *Visualising the Distribution of Networks using Jensen-Shannon Divergence*

We require a tool for visualising the similarity of sets of graphs measured by the entropies computed from the different partition functions. To this end, we measure similarity using the Jensen-Shannon divergence [30], which is asymmetric information theoretic divergence measure computed from the entropies of pairs of graphs. We characterise the similarities of a set of graphs using a kernel matrix and then embed the graphs into a vector space using kernel-embedding for the purposes of visualisation.

Here we deal with the case where the nodes in the graphs are labelled, and at each time step, the node-sets are identical. Only the edge-set varies between time-steps. Moreover, since the nodes are labelled it is straightforward to determine which edges have been added, removed or remained unchanged between different time steps. Suppose that G_i and G_j are the graphs at time steps i and j, and that $G_i \oplus G_j$ is the union graph with the set of edges formed from those edges that are present at either time step i or time step j. With the union graph to hand, the Jensen-Shannon divergence for the pair of graphs G_i and G_j is

$$D_{JS}(G_i,G_j) = S(G_i \oplus G_j) - \frac{S(G_i)+S(G_j)}{2} \tag{6.1}$$

where $S(G_i)$ is the entropy associated with the graph G_i, and $S(G_i \oplus G_j)$ is the entropy associated with the corresponding union graph G_U. Then the Jensen-Shannon kernel[6] is given by

$$k_{JS}(G_i,G_j) = \log 2 - D_{JS}(G_i,G_j) \tag{6.2}$$

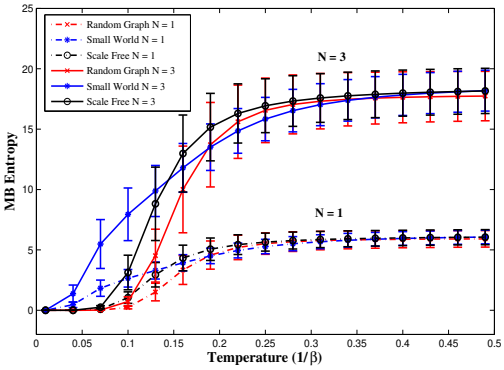
With the graph kernel to hand, we embed the graphs into a vector space. To compute the embedding, we commence by computing the eigendecomposition of the kernel matrix, which will reproduce the Hilbert space with a non-linear mapping. In such a case, graph features can be mapped to low dimensional feature space with linear separation. The graph kernel decomposition is

$$k_{JS} = \Phi \Lambda \Phi^T \tag{6.3}$$

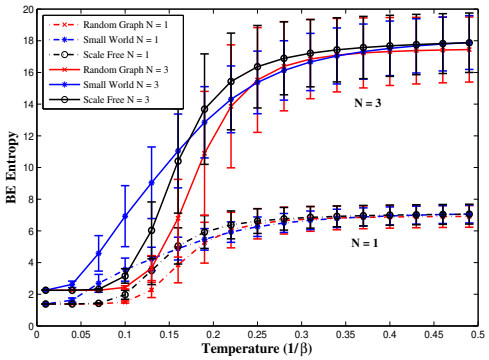
where Λ is the diagonal eigenvalue matrix and Φ is the matrix with eigenvectors as columns. To recover the matrix X with embedding coordinate vectors as columns, we write the kernel matrix in Gram-form, where each element is an inner product of embedding coordinate vectors

$$k_{JS} = XX^T \tag{6.4}$$

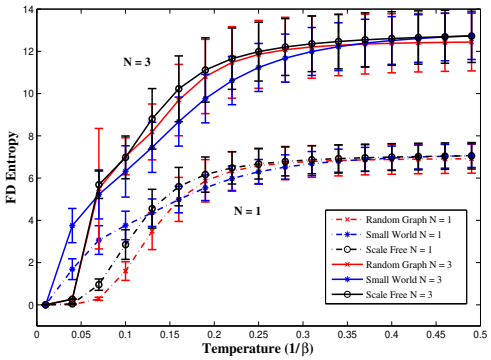
and as a result $X = \sqrt{\Lambda}\Phi^T$. In practice, we embed the samples of graphs into a three-dimensional space and hence use just the three leading eigenvalues and corresponding eigenvectors of k_{JS} to compute the embedding.



(a) Maxwell-Boltzmann Statistics



(b) Bose-Einstein Statistics



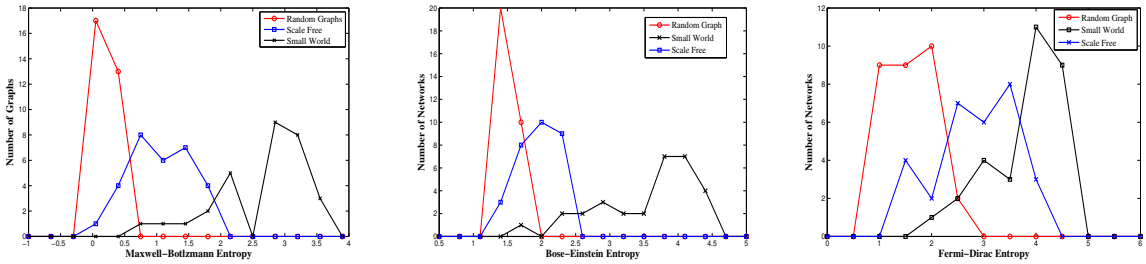
(c) Fermi-Dirac Statistics

FIG. 2: (Colour online) Mean and standard deviations of the entropies for three different network models versus temperature. Number of particles N = 1 and N = 3. Red cross line: Erdős-Rényi random graphs; blue star line: Watts-Strogatz small world networks; black circle line: Barabási-Albert scale free networks.

6.3 *Parameter Dependence*

In this section, we investigate how well the different models of the energy level occupation statistics can be used to distinguish synthetic networks generated using the Erdős-Rényi random graphs, Watts-Strogatz small-world models [40] and Barabási-Albert scale-free network models [7, 8]. We conduct numerical experiments to evaluate whether the thermodynamic variables, especially entropy, can represent differences in the structure and topology of networks.

Fig.2(a) shows the behaviour of the entropies resulting from Maxwell-Boltzmann occupation statistics as a function of temperature $(1/\beta)$. We explore the effect of varying the number of particles occupying the system and explore the cases where $N = 1$ and $N = 3$. From Eq.(4.3), it is clear that the effect of varying N is simply to scale the entropy by a multiplicative factor. For the three different graph models (Erdős-Rényi random graph model, Watts-Strogatz small-world model and Barabási-Albert scale-free



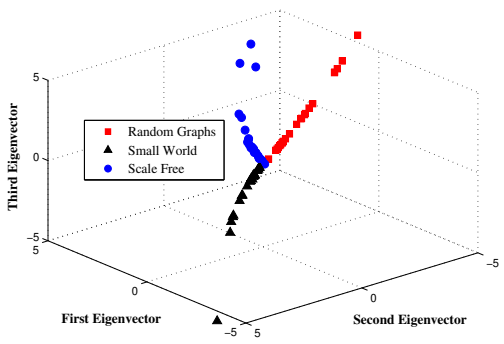
(a) Maxwell-Boltzmann Statistics (b) Bose-Einstein Statistics (c) Fermi-Dirac Statistics

FIG. 3: (Colour online) Histograms of entropy for Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics. The networks are randomly generated with the number of nodes generated from a normal distribution with the number of nodes between 100 and 1,000. The red line represents Erdős-Rényi random graphs; the black line small-world networks and the blue line scale-free networks. Temperature $\beta = 10$ and the number of particles $N = 1$.

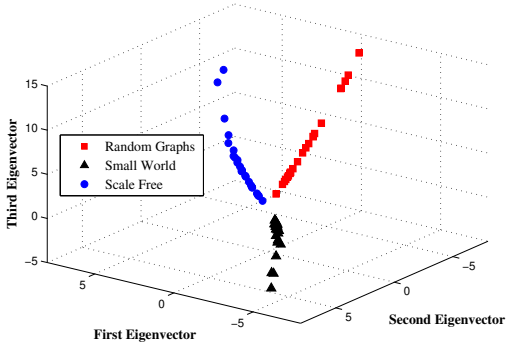
model), there is different behaviour with temperature. For small-world networks, the entropy increases fastest at low values of temperature. But it is quickly overtaken by the scale-free networks at intermediate temperatures. The Erdős-Rényi random graph model shows the slowest rate of increasements. The common feature is that all three entropies increase monotonically with temperature. However, the detailed dependence on $1/\beta$ depends on the partition function and the underlying occupation statistics. Specifically, at the low-temperature region ($0.07 \sim 0.12$), the entropy distinguishes strongly among the different types of network models.

Fig.2(b) and Fig.2(c) respectively show similar plots for the entropies derived from the Bose-Einstein and Fermi-Dirac partition functions. In the case of the Bose-Einstein entropy, the curves for the three different graph-modes exhibit the same pattern as in the Maxwell-Boltzmann case. As a result, at low temperatures, the ordering of the Bose-Einstein entropy can be used to separate the different network models. In both the Bose-Einstein and Fermi-Dirac, the number of particles N affects the entropy via the chemical potential μ. Hence, the entropy is not simply scaled by changing N. In the case of the Fermi-Dirac partition function, the pattern of entropies for the different modes is more complex for the various network models. Firstly, for different values of N, the behaviour is very different with temperature. For $N = 1$, we see a similar pattern to the Maxwell-Boltzmann and Bose-Einstein cases, but with $N = 3$ the behaviour is different with the scale-free and random graphs having small separation for all values of temperature. Additionally, the small-world model is overtaken by the random graphs and scale-free models at a lower value of temperature. This is a consequence of the exclusion principle manifesting itself at low temperature, and hence modifying the distribution of entropy for the different models.

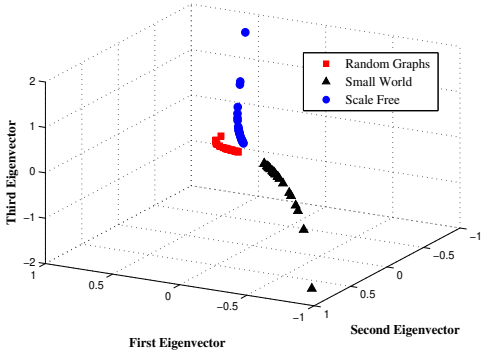
Comparing the plots for the Bose-Einstein and Fermi-Dirac entropies, the following features should also be noted: a) in each case for the different models approach the same limiting value for a given value of N, b) in the case of the Fermi-Dirac partition function all networks have zero entropy as temperature approaches zero, c) in the case of the Bose-Einstein model the entropy approaches the finite value $\ln N$ at zero temperature determined by the number of particles in the system, d) the Fermi-Dirac entropy increases more rapidly with increasing temperature than the Bose-Einstein entropy. On the other hand,



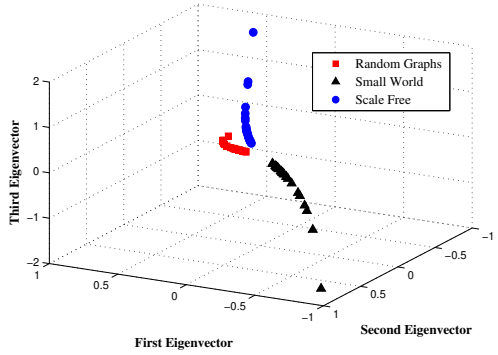
(a) Maxwell-Boltzmann Statistics $N = 5$



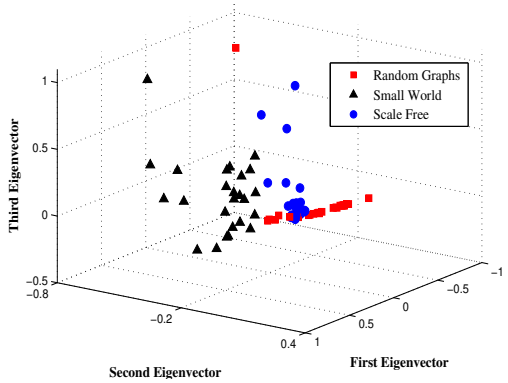
(b) Maxwell-Boltzmann Statistics $N = 10$



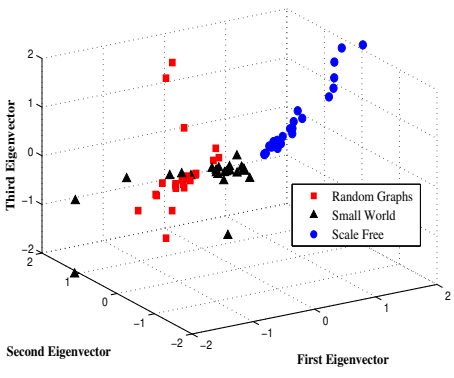
(c) Bose-Einstein Statistics $N = 5$



(d) Bose-Einstein Statistics $N = 10$



(e) Fermi-Dirac Statistics $N = 5$



(f) Fermi-Dirac Statistics $N = 10$

FIG. 4: (Colour online) Kernel embedding from Jensen-Shannon divergence computed with Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac entropies. We compare the effect of different numbers of particles ($N = 5$ and $N = 10$) with fixed temperature $\beta = 10$.

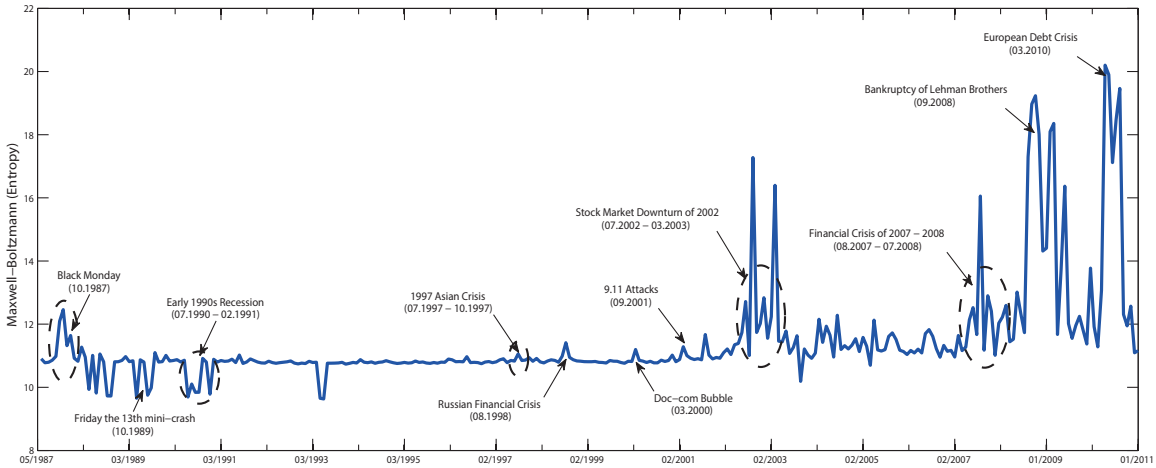


FIG. 5: (Colour online) Entropy from Maxwell-Boltzmann occupation statistics for NYSE (1987-2011). Critical financial events, i.e., Black Monday, Friday the 13th mini-crash, Early 1990s Recession, 1997 Asian Crisis, 9.11 Attacks, Downturn of 2002-2003, 2007 Financial Crisis, the Bankruptcy of Lehman Brothers and the European Debt Crisis, all appear as distinct events. Particle number $N = 5$ and temperature $\beta = 7$.

as the temperature increases, the occupation probability for the higher energy states increases and particles begin to occupy to the higher energy states. Moreover, the occupation probabilities for the three different partition functions become identical.

As expected, the differences between the different models are most evident at low temperature. These observations also fit with the intuitions outlined in Section 5. The faster rise of the Fermi-Dirac entropy with temperature is a consequence of the greater probability of finding fermions in the higher energy levels. For Bose-Einstein entropy, the greater separation between the different network models as low temperature is a consequence of the different shape of their degree distributions.

6.4 *Distinguishing Different Network Models*

We now explore the ability of the different entropies, resulting from the three different partition functions (Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac) to distinguish the three types of complex networks (random graphs, small-world networks and scale-free networks). Fig.3 shows histograms of the entropy for data generated from the three network models in Dataset 1. Each figure shows the entropy computed using a different partition function. The differently coloured curves in the histograms correspond to the distribution from three network models. In each plot, the Erdős-Rényi random graphs occupy the low entropy region while the small-world networks stay at the high entropy-area. The distributions of random graphs and scale- free networks are closer in Maxwell-Boltzmann and Bose-Einstein cases when compared to the small-world networks. However, using entropy simply as a unary feature is insufficient to obtain good separation between the different network models (Erdős-Rényi random graphs, Watts-Strogatz small-world networks and Barabási-Albert scale-free networks).

Better separation can though be obtained if we analyse the pattern of entropy differences between

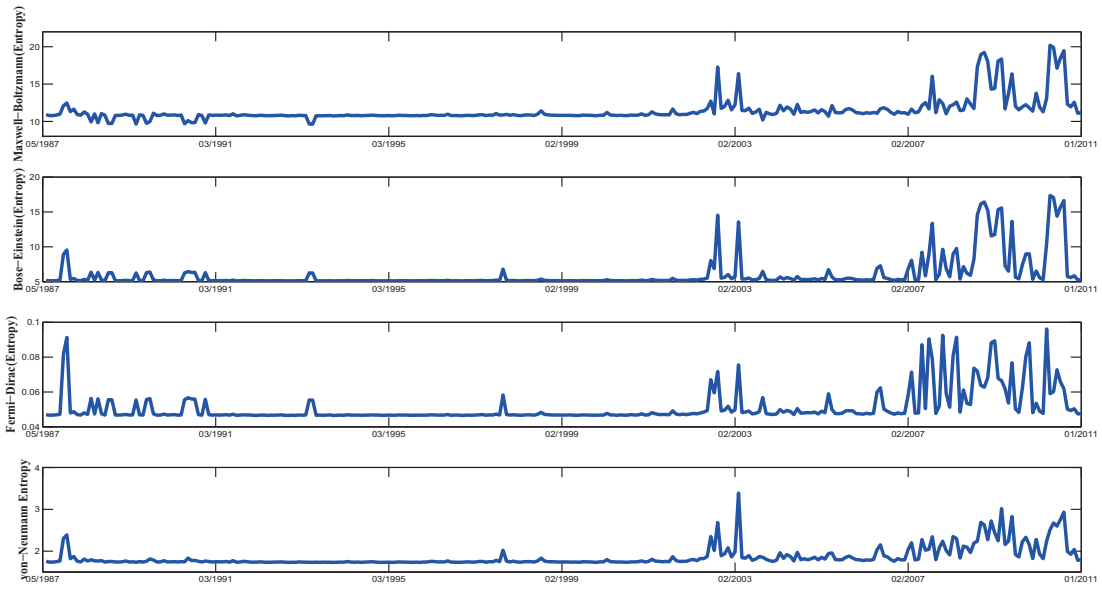


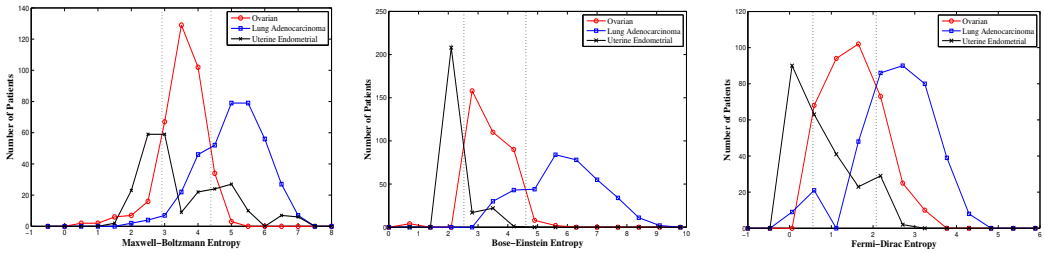
FIG. 6: (Colour online) von Neumann Entropy and thermodynamic entropy compared for NYSE (1987-2011): (a) Maxwell-Boltzmann occupation statistics, (b) Bose-Einstein occupation statistics and (c) Fermi-Dirac occupation statistics. (d) von Neumann entropy.

pairs of graphs. Fig. 4 shows the results of applying the kernel embedding technique outlined in Section B to the entropies computed from the three different partition functions. The differently coloured points correspond to the data generated from the three different network models (red - Erdős-Rényi random graphs, blue - small world networks, back - scale-free networks). In the case of Maxwell-Boltzmann and Bose-Einstein, the different models from non-overlapping subspaces and can be easily separated. In the case of Bose-Einstein statistics, the effect of changing the number of particles is negligible. In the case of Fermi-Dirac statistics, on the other hand, although more scattered when the number of particles is low, they form tightly clustered subspaces when a larger number of particles are used. This is in line with our physical intuition, since if the number of particles is increased, then so the number of energy levels populated increases, even at low temperature. This is in contrast to the Bose-Einstein case, where particles congregate at low energy levels.

The results above are obtained, using entropies derived from Maxwell-Boltzmann and Bose-Einstein partition functions, the Jensen-Shannon divergence with kernel embedding provides a better visualisation of the separation of the different numerical network models.

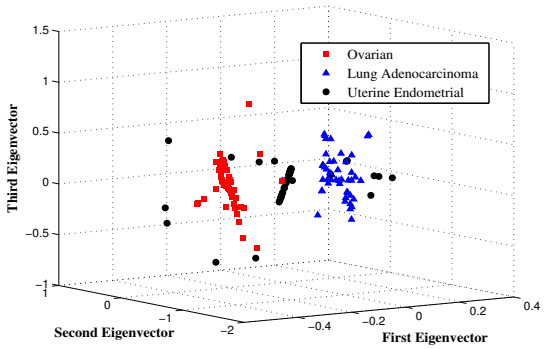
6.5 *Real World Data*

We now compute the entropy characterizations obtained from the three different partition functions on real world data. Specifically, we explore whether the entropy can be used as an effective tool for better understanding the evolution of real-world complex networks. First, we focus on the detail of New York Stock Exchange in Dataset 2 and then provide analysis for the tumour mutation networks in Dataset 3.

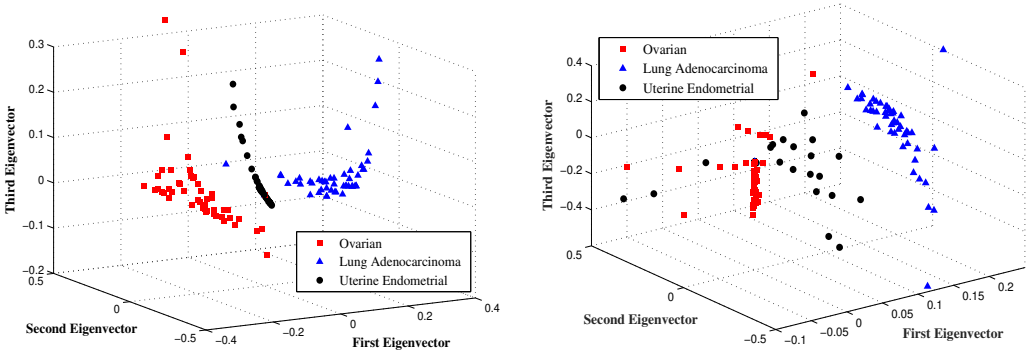


(a) Maxwell-Boltzmann Statistics (b) Bose-Einstein Statistics (c) Fermi-Dirac Statistics

FIG. 7: (Colour online) Histograms of entropy from three statistics for tumour mutation networks (ovarian, uterine and lung adenocarcinoma). Particle number $N = 2$, temperature $\beta = 10$.



(a) Maxwell-Boltzmann Statistics



(b) Bose-Einstein Statistics (c) Fermi-Dirac Statistics

FIG. 8: (Colour online) Kernel embedding with Jensen-Shannon divergence computed from tumour mutation network entropies (ovarian, uterine and lung adenocarcinoma) for different partition functions. Particle number $N = 3$, temperature $\beta = 10$.

6.5.1 *Stock Market Data* Fig.5 and Fig.6 show the entropy time-series for the NYSE data obtained from different partition functions. In Fig.5, the entropy is derived from the Maxwell-Boltzmann partition function. It is annotated to show the positions of significant financial events such as Black Monday, Friday the 13th mini-crash, Early 1990s Recession, 1997 Asian Crisis, 9.11 Attacks, Downturn of 2002-2003, 2007 Financial Crisis, the Bankruptcy of Lehman Brothers and the European Debt Crisis. In each case, the entropy undergoes significant fluctuations during the financial crises, associated with dramatic structural changes. A good example is the downturn of 2002-2003. After the 9.11 attacks, investors became unsure about the prospect of terrorism affecting the United States economy. Following the subsequent collapse of many internet companies, numerous large corporations were forced to restate earnings and investor confidence suffered. This considerably altered the inter-relationships among stocks and resulted in significant variance in the structure of the entire market.

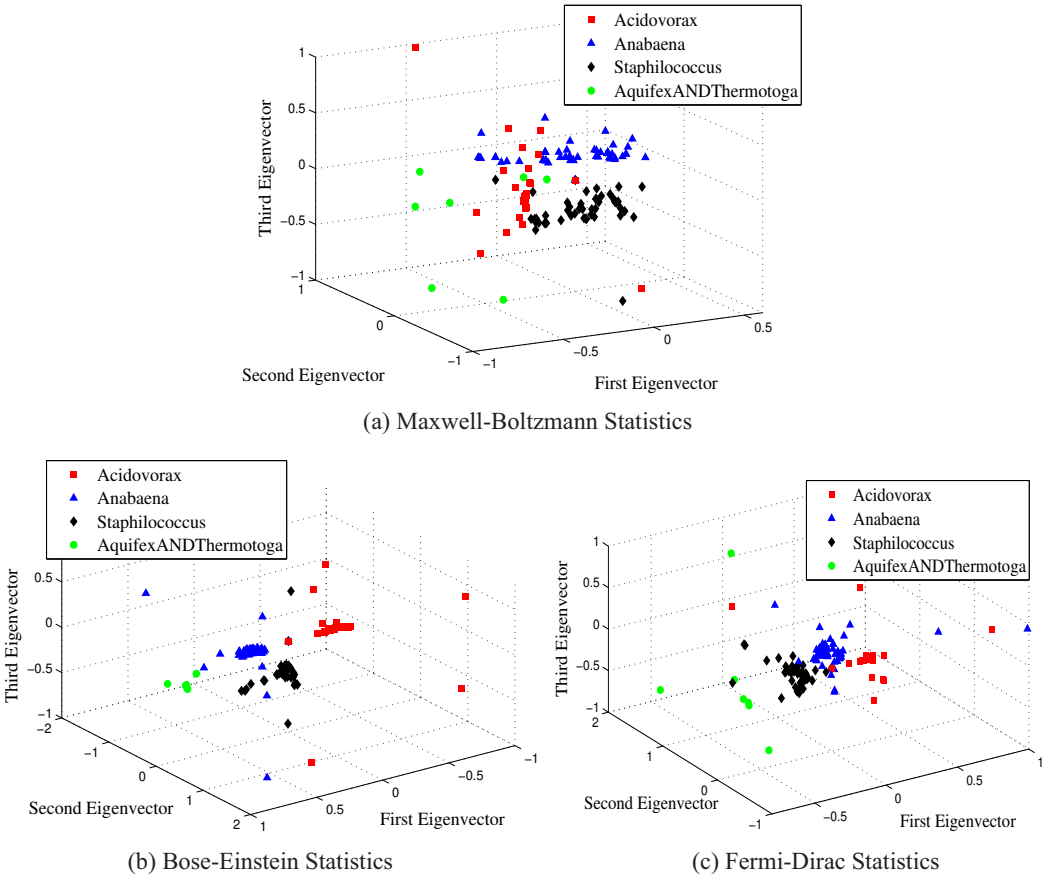
Fig.6 compares the entropy derived from the three different partition functions with the von Neumann entropy. In the figure, entropies coming from the three partition functions perform better in evaluating the structural changes in the network time-series when compared to the von Neumann entropy. Further exploration shows that entropies, derived from Bose-Einstein and Fermi-Dirac partition functions, exhibit the similar behaviour in the evolution of stock markets. Compared to the Maxwell-Boltzmann case, the Bose-Einstein and Fermi-Dirac entropies are more sensitive to the critical events in the financial data, such as Black Monday in 1987 and the Asian Financial Crisis in 1997.

6.5.2 *Tumour Mutation Networks* Next, we turn our attention to the tumour mutation networks for the three different cancers, i.e. a) ovarian cancer, b) uterine cancer and c) lung adenocarcinoma. In Fig.7(a), we provide the histogram of the entropy computed from the Maxwell-Boltzmann partition function. The different colour of curves represent the three types of cancers. The most striking feature of this plot is that the three kinds of tumour networks dominate different entropy intervals. By applying two separate thresholds to the entropy histogram, we can assign the patients to three classes. We have searched for the two thresholds which give the maximum pooled classification accuracy over the three cancer classes. We find that the best result is given when the uterine and ovarian classes are separated using an entropy threshold at $S_{MB} = 2.92$, and the ovarian and lung adenocarcinoma with a threshold at $S_{MB} = 4.38$. The resulting classification accuracies are 33.87% for uterine cancer, 83.71% for ovarian cancer and 78.48% for lung adenocarcinoma.

Table 1: Classification accuracy with three different partition functions. The thermodynamic entropic thresholds for Maxwell-Boltzmann statistics are 2.92 and 4.38. The values of entropy separation for Bose-Einstein statistics are 2.49 and 4.52. And the corresponding thresholds of entropy for Fermi-Dirac statistics are 0.56 and 2.08.

Accuracy	Uterine Cancer	Ovarian Cancer	Lung Adenocarcinoma	Total [1]
Maxwell-Boltzmann	33.87% (84/248)	83.71% (312/356)	78.48% (300/381)	70.66% (696/985)
Bose-Einstein	75.00% (186/248)	93.54% (333/356)	80.84% (308/381)	83.96% (827/985)
Fermi-Dirac	63.71% (153/248)	74.16% (264/356)	79.53% (303/381)	73.10% (720/985)

Fig.7(b) repeats the analysis using the entropy derived from the Bose-Einstein partition function. Here the corresponding thresholds are 2.49 and 4.52, giving correct classification rates of 75.00%, 93.54%, and 83.96% for the uterine, ovarian, and lung adenocarcinoma classes respectively. For the case of the Fermi-Dirac entropy, as shown in Fig.9(c), the thresholds are $S_{FD} = 0.56$ and $S_{FD} = 2.08$ giving classification accuracies of 63.71%, 74.16%, 73.10% for the uterine, ovarian and lung adenocarcinoma groups.



(a) Maxwell-Boltzmann Statistics

(b) Bose-Einstein Statistics

(c) Fermi-Dirac Statistics

FIG. 9: (Colour online) Kernel embedding with Jensen-Shannon divergence computed from PPI network entropies (Acidovorax, Anabaena, Staphilococcus and Aquifex&Thermotoga) for different partition functions. Particle number $N = 5$, temperature $\beta = 10$.

To improve the separation of the data, we use the kernel embedding based on the Jensen-Shannon divergence to measure network similarity, as outlined in Section B. The results of the tumour networks, embedding into the three-dimensional space spanned by the first three leading eigenvectors of the kernel matrix, are shown in Fig.8. The plot sheds light on the three different classes of data (shown in different colours) exhibiting a compact manifold structure for three statistics. For each entropy, the different groups of tumour mutation networks are well separated in the embedding space, and this is especially so in the case of the Bose-Einstein entropy. In the Maxwell-Boltzmann and Fermi-Dirac cases, although, the groups of lung adenocarcinoma and ovarian cancer are well separated, the outliers of uterine tumour networks are interspersed among remaining two classes. The best results are obtained in the Bose-Einstein case where the individual networks of the uterine group form the most compact cluster.

6.5.3 *Protein-Protein Interaction Networks* Our final example is based on Protein-protein interaction networks. We perform kernel embedding on the protein-protein interaction networks to visualise their distribution and to provide a comparison between the entropic discrimination obtained with different partition functions. To this end, we show the distribution of the PPI's in the space spanned by the leading three kernel principle components in Fig.9. In each case, the embedded data exhibits a manifold structure which results in good separation of the different classes of PPI. Moreover, the Bose-Einstein entropy provides a better separation of more tightly formed clusters and fewer outliers. The reason for this is that Bose-Einstein statistics encourage particles to aggregate in the lower energy states at low temperature. This amplifies the influence of the number of connected components and the spectral gap in determining the entropy. The former is reflected by the multiplicity of the zero eigenvalues, and the latter relates to the degree of bi-partivity in the network. The particle occupation of the low energy states produces a stronger entropic separation in the Bose-Einstein case. By contrast neither the Maxwell-Boltzmann nor Fermi-Dirac statistics strongly reflect the lower part of Laplacian spectrum, since they do not give a similar particle concentration in the lower energy states. As a result, Bose-Einstein statistics are more sensitive to the cluster structure of networks, and in the case of PPI's where there is a strong inhomogeneity of node degree which leads to better separation of different classes.

6.5.4 *Conclusions from the real world data study* In the case of the tumour mutation networks, overall, the best-pooled performance comes from the Bose-Einstein entropy. Compared with the Maxwell-Boltzmann case, the entropies derived from spin statistical partition functions appear to be more sensitive to differences in network structure and more accurately reflect the structural differences between distinct types of tumour mutation networks. The same pattern emerges with protein-protein interactions networks; This is not surprising since the PPI's have a strong cluster (community) structure. This again fits with the intuitions given in Section 5.

7. Conclusions

In this paper, we have explored different thermodynamic characterizations resulting from alternative energy level occupation statistics in the heat bath analogy. The effects of the different occupation statistics are captured using a partition function.

Our study uses the normalised Laplacian matrix as the Hamiltonian operator of the network, and the associated energy states are given by the eigenvalues of the normalised Laplacian. We explore the case where the particle occupations correspond to Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics. From the related partition functions, we can compute the thermodynamic entropy and energy. Motivated by an interest in revealing the nontrivial properties of the network structure, we have compared the three resulting entropic characterizations and with the von Neumann entropy. We provide a detailed analysis of the three different partition functions, expressed both in terms of the normalised Laplacian matrix and its eigenvalues.

We evaluate the network models resulting from the three different partition functions on both synthetic and real-world data sets. This study investigates how the different entropies can be used characterise changes in network structure, and distinguish different types of network structure. Studies with synthetic data show that the entropies can distinguish Erdős-Rényi random graphs, Watts-Strogatz small-world networks, Barabási-Albert scale-free networks. Experiments with real-world data, on the other hand, show that the thermodynamic variables not only can be used to detect both abrupt changes in network structure but also distinguish different classes of networks.

The main conclusion from this study is that for distinguishing different network models, the Fermi-

Dirac entropy appears best. The reason for this is that it is most sensitive to the higher eigenvalues of the normalised Laplacian and this allows it to better probe differences in the degree distributions for different models. Our real world data, on the other hand, comes mainly from problems where there is a strong community or cluster structure. Here the Bose-Einstein model performs best, and the reason for this is that it is most sensitive to the eigenvalue gap.

The work reported in this paper can clearly be extended to a number of different ways. First, we acknowledge that we have explored a relatively limited quantity of real-world data. It would, for example, be interesting to see if the thermodynamic variables can be used to detect temporal anomalies and disturbances in the evolution of networks in a greater variety of data. Another interesting line of investigation would be to explore whether phase transitions can be detected with thermodynamic quantities such as entropy and temperature. Finally, we plan to extend this work to the low-temperature limits to explore the phenomenon of Bose-Einstein condensation observed in the networks.

REFERENCES

1. (2010) *The International Cancer Genome Consortium, International network of cancer genome projects, Nature***464**, 993–996.
2. (2010) STRING - Known and Predicted Protein-Protein Interactions. .
3. Albert, R. & Barabasi, A.-L. (2002) Statistical Mechanics of Complex Networks. *Review Moden Physics*, **74**, 47.
4. Alstott, J., Pajevic, S., Bullmore, E. & Plenz, D. (2015) Opening bottlenecks on weighted networks by local adaptation to cascade failures. *Journal of Complex Networks*, **3**, 552–565.
5. Anand, K., Bianconi, G. & Severini, S. (2011) Shannon and von Neumann entropy of random networks with heterogeneous expected degree. *Physical Review E 036109*, **83(3)**.
6. Bai, L. & Hancock, E. (2012) Graph Kernels from the Jensen-Shannon Divergence. *Journal of Mathematical Imaging and Vision*, **47**, 60–69.
7. Barabasi, A.-L. & Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.
8. Barabasi, A.-L., Albert, R. & Jeong, H. (1999) Mean-field theory for scale free random networks. *Physics A,*, **272**, 173–187.
9. Bianconi, G. (2002) Growing Cayley trees described by a Fermi distribution. *Physical Review E 036116*, **66**.
10. Bianconi, G. (2015) Supersymmetric multiplex networks described by coupled Bose and Fermi statistics. *Physical Review E*, **91**.
11. Bianconi, G. & Barabasi, A.-L. (2001) Bose-Einstein Condensation in Complex Networks. *Physical Review Letter*, **88**, 5632.
12. Blundell, S. J. & Blundell, K. M. (2006) *Concepts in Thermal Physics*. Oxford University Press.
13. Boginski, V., Butenko, S. & Pardalos, P. M. (2006) Mining market data: A network approach. *Computers and Operations Research*, **33**, 3171–3184.
14. Braunstein, S., Ghosh, S. & Severini, S. (2006) The laplacian of a graph as a density matrix: A basic combinatorial approach to separability of mixed states. *Annals of Combinatorics*, **10(3)**, 291–317.
15. Chung, F. (1997) Spectral Graph Theory. *CBMS Regional Conference Series in Mathematics*, **92**.
16. Clough, J. R., Gollings, J., Loach, T. V. & Evans, T. S. (2015) Transitive reduction of citation networks. *Journal of Complex Networks*, **3**, 189–203.
17. Domenico, D., Lancichinetti, A., Arenas, A. & Rosvall, M. (2015) Structural reducibility of multilayer networks. *Nature Communications*, **6**, 7864.
18. Escolano, F., Bonev, B. & Hancock, E. R. (2012a) Heat Flow: Thermodynamic Depth Complexity in Directed Networks. *Structural, Syntactic, and Statistical Pattern Recognition*, **85**.
19. Escolano, F., Hancock, E. R. & Lozano, M. A. (2012b) Heat diffusion: Thermodynamic depth complexity of networks. *Physical Review E 036206*, **85**, 190–198.
20. Escolano, F., Lozano, M. A., Hancock, E. R. & Giorgi, D. (2010) What is the complexity of a network? the

heat flow-thermodynamic depth approach. *Joint IAPR International Workshops on Structural and Syntactic Pattern Recognition and Statistical Techniques in Pattern Recognition*, pages 286–295.

21. Estrada, E. & Hatano, N. (2007) Statistical- mechanical approach to subgraph centrality in complex networks. *Chemical Physics Letters*, **439**, 247–251.

22. Estrada, E. & Hatano, N. (2008) Communicability in Complex Networks. *Physical Review E*, **77**.

23. Ethan, C., Benjamin, G., Emek, D., Igor, R., Ozgun, B., Nadia, A., Nikola, S., Gary, B. & Chris., S. (2011) Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res*, **39**, D695–D690.

24. Garlaschelli, D. & Loffredo, M. I. (2006) Multispecies grand-canonical models for networks with reciprocity. *Physical Review E*, **73**.

25. Gutfraind, A., Bradonji, M. & Novikoff, T. (2015) Modelling the neighbour aid phenomenon for installing costly complex networks. *Journal of Complex Networks*, **3**, 249–263.

26. Han, L., Hancock, E. & Wilson, R. (2012) Characterizing Graphs Using Approximate von Neumann Entropy. *Pattern Recognition Letter*, **33**, 1958.

27. Hernndez, J. M., Li, Z. & Mieghem, P. V. (2014) Weighted betweenness and algebraic connectivity. *Journal of Complex Networks*, **2**, 272–287.

28. Hofree, M., Shen, J. P., Carter, H., Gross, A. & Ideker, T. (2013) Network-based stratification of tumor mutations. *Nature Methods*, **10**, 11081115.

29. Kuehn, C., Marstens, E. A. & Romero, D. M. (2014) Critical transitions in social network activity. *Journal of Complex Networks*, **2**, 141–152.

30. Martins, A., Smith, N., Xing, E., Aguiar, P. & Figueiredo, M. (2009) Nonextensive Information Theoretic Kernels on Measures. *Journal of Machine Learning Research*, **10**, 935–975.

31. Mikulecky (2001) Network thermodynamics and complexity: a transition to relational systems theory. *Computers & Chemistry*, **25**, 369.

32. Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J. & Kanto, A. (2003) Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E 056110*, **68**.

33. Park, J. & Newman, M. (2004) Statistical mechanics of networks. *Physical Review E 066117*, **70(6)**.

34. Passerini, F. & Severini, S. (2008) International Journal of Agent Technologies and Systems. *The von Neumann entropy of networks*, pages 58–67.

35. Perseguers, S., Lewenstein, M., Acn, A. & Cirac, J. (2009) Quantum complex networks. *Nature Physics*, **6**, 539 – 543.

36. Shen, Y., Zhu, D. & Liu, W. (2004) Fermi-Dirac Statistics of Complex Networks. *Chinese Phys. Lett.*, **22**, 1281.

37. Silva, F., Comin, C., Peron, T., Rodrigues, F., Ye, C., Wilson, R., Hancock, E. & Costai, L. (2015) Modular Dynamics of Financial Market Networks. *Physics and Society*, **arXiv:1501.05040**.

38. Streitwieser, A. (2013) *Molecular Orbital Theory for Organic Chemists*, volume 9. American Chemical Society.

39. Waclaw, B. (2013) Statistical mechanics of complex networks. *arXiv:0704.3702*.

40. Watts, D. & Strogatz, S. (1998) Collective dynamics of small world networks. *Nature*, **393**, 440–442.

41. Ye, C., Torsello, A., Wilson, R. C. & Hancock, E. R. (2015a) Thermodynamics of Time Evolving Networks. *GbRPR 2015*, pages 315–324.

42. Ye, C., Wilson, R. C., Comin, C. H., da F. Costa, L. & Hancock, E. R. (2014) Approximate von Neumann entropy for directed graphs. *Physical Review E*, **89**.

43. Ye, C., Wilson, R. C., Comin, C. H., da F. Costa, L. & Hancock, E. R. (2015b) Thermodynamic Characterization of Networks Using Graph Polynomials. *Physical Review E*, **92**.