Graph Motif Entropy For Understanding Time-Evolving Networks

Zhihong Zhang, Dongdong Chen, Lu Bai*, Jianjia Wang and Edwin R Hancock, *Fellow, IEEE*

Abstract—The structure of networks can be efficiently represented using motifs, which are those subgraphs that recur most frequently. One route to understanding the motif structure of a network is to study the distribution of subgraphs using statistical mechanics. In this paper, we address the use of motifs as network primitives using the cluster expansion from statistical physics. By mapping the network motifs to clusters in the gas model, we derive the partition function for a network and this allows us to calculate global thermodynamic quantities, such as energy and entropy. We present analytical expressions for the numbers of certain types of motifs, and compute their associated entropy. We conduct numerical experiments for synthetic and real-world data-sets and evaluate the qualitative and quantitative characterizations of the motif entropy derived from the partition function. We find that the motif entropy for real-world networks, such as financial stock market networks, is sensitive to the variance in network structure. This is in line with recent evidence that network motifs can be regarded as basic elements with well defined information-processing functions.

Index Terms—Cluster Expansion, Motif, Network Entropy

1 Introduction

Complex networks can be succinctly described using simple statistical models which describe their global properties. Examples include the Erdos-Renyi, small-world and scale-free models. The parameters of these models can in turn be used to characterize network structure [1], [2]. Moreover, these parametric characterizations implicitly represent the key structural mechanisms determining the global features of a network [3]. Recently, though network motifs have provided a powerful alternative route to the statistical analysis of network characteristics and structure. They emphasize the importance of studying the small scale aspects of network structure in order to gain a better understanding of their global structure and function [4], [5]. According to this picture, recurring patterns, termed network motifs, provide the basic building blocks for different subgraphs which perform specific functional roles in a larger network structure [6]. Moreover,these motifs reflect the underlying physical processes or interactions that generate each type of network [7]. For example, in *gene regulation network* motifs (or different functional subgraphs) have a variable frequency of occurrence during different phases of development. *Large Scale Evolving Graphs*, on the other hand, where edges and vertices are appended to the network as they arrive in time, sometimes contain bursty links. These links which aim characterize anomalous objects and relationships [8]. Graph structure can also be used in adversarial training [9].

Although several algorithms and graph-theoretic tools have been developed for efficiently detecting and counting network motifs [10], [11], there is little analysis of the statistical features of networks from the motif perspective. This task requires an understanding of the basic structural elements constituting the motifs and the processes which give rise to them from a microscope point of view [5], [6]. To embark on this type of analysis, tools from statistical mechanics provide a convenient route to the characterization of network structure [12]. Thermodynamic characterizations, such as entropy, total energy and temperature, derived from a partition function, can be used to succinctly describe the network statistics [13]. By adopting an analogy in which the network consists of a statistical system of interacting particles with a partition function, the *cluster expansion* provides deep insights into network behavior related to the occurrence of different motifs.

The cluster expansion is a powerful computational tool which can be used to express the partition function in terms of an approximating series [14]. It allows us to write the grand-canonical thermodynamic potential as a convergent perturbation over the interactions between particles. Commencing from the general principles of perturbation theory for particle systems, the cluster expansion allows us to understand complex systems of interactions in terms of the motif topologies in a diagrammatic expansion of the partition function. The cluster expansion is a general integral expression posed at the motif level [15]. For a graph, the partition function can be written as the exponential of a sum over connected subgraphs, namely the network motifs [16]. Thus the interactions present in the partition function are expressed in terms of motifs and these in turn express the statistical properties of the network.

Here, we use the motif content of a network and the corresponding partition function to compute the thermodynamic entropy as a statistical characterization of network

- *Zhihong Zhang and Dongdong Chen are with School of Informatics, Xiamen University, Xiamen, Fujian, China. E-mail: zhihong@xmu.edu.cn, chendongdong@stu.xmu.edu.cn*
- *Corresponding author: Lu Bai is with Central University of Finance and Economics, Beijing, China. E-mail: bailucs@cufe.edu.cn*
- *Jianjia Wang is with 1.School of Compputer Engineering and Science, Shanghai University; 2.Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai, P.R.China. E-mail: jianjiawang@shu.edu.cn*
- *Edwin R. Hancock is with University of York, York, UK. E-mail: edwin.hancock@york.ac.uk*

structure. We first outline the link between the partition function, the thermodynamic characterization and the cluster expansion formalism. Then we review the classical cluster expansion method in the gas model with particle interactions. We present an analytical solution for the numbers of network motifs and the scaling of all types of motifs with the partition functions. Finally, we find invariants that can be used to represent the statistical topology of networks. This specifically focuses on entropy for the cluster expansion in terms of network motifs.

2 Related Work

2.1 Network Motifs

There is a substantial literature on analysing network motifs, starting from a subgraph description of the frequently occurring patterns within a network structure. It has been shown that motifs are recurring patterns that can be used in representation of of complex structure [6]. Indeed, motifs reflect not only the structural properties of a network, but can also capture its functional properties too. They thus provide an efficient way to uncover the structural design principles of a system represented by a complex network. Examples include positive and negative autoregulation [5], positive and negative cascades [17], positive and negative feedback loops [18], feedforward loops (FFLs) [19], single input modules [20], and combinations of these too [21]. Motifs can be regarded as the fundamental building blocks of complex networks, since identical network motifs exist in diverse network structures fields as diverse as biology and sociology. For example in the biological domain, network motifs have been implicated in signalling [22] and neuronal activities [23], and also account for the integration of transcriptional regulation and protein-protein interactions [24].

Although network motifs provide deep insights into the characterisation of network structure and function, their detection is a computationally challenging problem. There are a variety of techniques for solving the motif discovery problem. These algorithms include exact counting methods, sampling methods, pattern growth methods, and have been developed under a number of different paradigms. Network-motif patterns can be identified when the nodes and edges in a network are annotated with quantitative features. The expected number of appearances of a motif can be determined using a Null-model [25], which is an ensemble of random networks with some of common properties with the original network.

2.2 Cluster Expansion

In statistical mechanics, the cluster expansion is usually a power series expansion of the partition function. It describes the pattern of interactions in a system with large number of particles. Mayer and his collaborators [26] first carried out a systematic study of alternative expansions, in the case of real gases obeying classical statistics. Kahn and Uhlenbeck [27] generalise the cluster expansion to gases obeying quantum statistics. Lee and Yang [28] have explored the application of these ideas to real world application.

The cluster expansion has become a standard tool in the analysis of solid state structures. Examples include the theory of two-dimensional solids, which are a kind of regular network structure. The seminal work of Peierls and Landau [29] use to the cluster expansion to how harmonic thermal fluctuations can destroy the long-range crystalline order in solids, and this idea has provided the foundation for an abundant literature on using the cluster expansion in understanding solids from a lattice or regular network perspective. [30], [31].

TABLE 1
Important notation used in this paper and their descriptions.

Symbol	Definition
N	number of particles or nodes
$\langle U_\nu \rangle$	energy of the network motif of type ν
S_ν	entropy of the motif of type ν
Q	configuration integral of the entire network
β	$\beta = \frac{1}{T}$ inverse temperature
λ	$\lambda = \sqrt{\frac{2\pi\beta h^2}{\mu}}$, average de Broglie wavelength of the particles
$f_{i,j}$	interaction strength of Mayer function between nodes i and j
l	the number of particles in an l-cluster
m_l	the number of l-clusters of type l
b_l	the classical cluster integral for l^{th} cluster type
χ_l	index set of nodes that constitute the network motif that consists of l nodes
ν	the individual motif index
n_ν	frequency of occurrence of the ν^{th} motif
q_ν	the configuration integral for the ν^{th} motif
l_ν	the number of nodes in the ν^{th} motif
ζ_ν	configuration integral of motif
Z	the partition function for the entire network
z_ν	the partition function associated with the ν^{th}

3 Preliminaries

The underlying aim of our method is to explore how microscopic characterisations can be used to describe the macroscopic structure of a network. The algorithm commences by detecting the motifs and describing their entropic properties. Such a mechanism is a novel way to describe the statistical characteristics of time-varying networks from a more microscopic perspective. In order to make our description clearer, we list the notation used in this paper in Table 1. The individual quantities will be defined in more detail as we develop our motif analysis.

3.1 Network Motif Representation

Motifs are representative subgraphs that frequently appear in a graph. They can be regarded as the basic building blocks of networks, which provide a better understanding of the global network structure from a statistical point of view [5]. Typically there are numerous network motifs corresponding to small sub-graph structures but which appear more frequently than would be expected in random networks [6]. Motifs can define the properties of broad classes of networks, each with a specific type of elementary structure [16]. Fig.1 illustrates typical types of motifs for graphs, which are n-node subgraphs (n = 2,3,4 in the figure). The number of occurrences of each subgraph is the motif frequency which are used as the basic structural pattern decomposition of a network [32]. Different types of motifs have specific

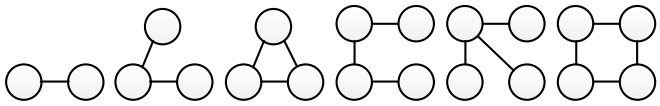

Fig. 1. Typical motifs for graph

functions as elementary patterns of interaction and allow us to interpret the structural properties of the overall network [33]. This similarity of the motifs reflects a fundamental similarity in the organization of network structure, and the statistical significance of different motifs can be regarded as a unifying property of networks too [5]. Thus, we use the network motifs as the elementary structural constructs to interpret the local and global topological information in a network.

3.2 Cluster Expansion and the Classical Gas

We characterize the statistical content of the different motif structure of the network using a partition function suggested by an analogy with the particle gas in statistical mechanics [13]. Specifically, we use motifs in a manner similar to groups of interacting particles in the thermodynamic gas model [6].

We begin by briefly reviewing the cluster expansion for the classical gas [34]. To commence we consider a system of N identical non-interacting particles occupying a volume V for which the Hamiltonian is composed purely of kinetic energy terms and is given by

$$\mathcal{E}_0 = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2\mu} \tag{1}$$

where $\vec{p}_i$ is the momentum of particle indexed i and μ is the mass of each particle. The partition function for this system of particles is given by

$$\begin{aligned} Z_0 &= \frac{1}{N!h^{3N}} \int d^3\vec{p}_i d^3\vec{r}_{i,j} \cdot \left\{ \exp\left[-\beta \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2\mu} \right] \right\} \\ &= \frac{V^N}{N!h^{3N}} \left[\frac{2\pi\mu}{\beta} \right]^{3N/2} \end{aligned} \tag{2}$$

where h is Planck's constant, $\vec{r}_{i,j}$ is the vector separation between particles i and j and $\beta = 1/T$, where T is the temperature.

To introduce interactions between pairs of particles we add a potential energy term into the Hamiltonian.

$$U(\{\vec{r}_i\}) = \sum_{i=1, i<j}^{N} v(\vec{r}_i - \vec{r}_j) = \sum_{i=1, i<j}^{N} v(\vec{r}_{i,j}) \tag{3}$$

where $\vec{r}i$ is the position vector of the particle indexed i. With the interaction potential included in the Hamiltonian, the partition function takes on the modified form

$$Z = Z_0 Q \tag{4}$$

where Q is the configuration integral given by integral over the positions for all the particles,

$$Q = \frac{1}{V^N} \int \prod_i d^3\vec{r}_i \exp\left\{ -\beta \sum_{1 \le i < j \le N} v(\vec{r}_{i,j}) \right\} \tag{5}$$

The sum inside the exponential can be replaced by a product over exponentials, with the result that

$$Q = \frac{1}{V^N} \int \prod_i d^3\vec{r}_i \prod_{1 \le i < j \le N} \exp\left\{ -\beta v(\vec{r}_{i,j}) \right\} \tag{6}$$

Unfortunately, the configuration integral is intractable for general pairwise potentials. To overcome this problem, the Boltzmann factor [35] $\exp[-\beta v(\vec{r}_{i,j})]$ appearing in the configuration integral is re-expressed in terms of the Mayer function $f_{i,j}$

$$\exp[-\beta v(\vec{r}_{i,j})] \equiv 1 + f_{i,j} \tag{7}$$

The Mayer function incorporates a hard core repulsion where the particles are forbidden from approaching closer than a fixed distance by imposing an infinite potential [14]. As a result the configuration integral can be rewritten as:

$$Q = \frac{1}{V^N} \int \prod_i d^3\vec{r}_i \prod_{1 \le i < j \le N} (1 + f_{i,j}) \tag{8}$$

The product over $1 + f_{i,j}$ can be expanded as a polynomial in Mayer functions on different edges. So grouping terms of different order together, we have the expansion

$$\begin{aligned} Q = \frac{1}{V^N} \int \prod_i d^3\vec{r}_i \Big\{ & 1 + (f_{1,2} + f_{1,3} + \cdots) \\ & + (f_{1,2}f_{1,3} + f_{1,2}f_{1,4} + \cdots) + \cdots + f_{1,2}f_{1,3} \ldots f_{N-1,N} \Big\} \end{aligned} \tag{9}$$

We can interpret the terms of each order in the above expansion as a graph representing particle interactions. Each edge represents an interaction between a pair of particles, determined by the strength of the interaction potential as determined by the Mayer function. So Q can be interpreted as the sum of all possible graphs representing the different combinations of pairwise interactions of N particles.

To proceed we separate the configuration integral from the set of particle interaction configurations and write

$$Q = \frac{1}{V^N} \int \prod_i d^3\vec{r}_i \mathcal{F} \tag{10}$$

where

$$\begin{aligned} \mathcal{F} = \{ & 1 + (f_{1,2} + f_{1,3} + \cdots) + (f_{1,2}f_{1,3} + f_{1,2}f_{1,4} + \cdots) \\ & + \cdots + f_{1,2}f_{1,3} \ldots f_{N-1,N} \} \end{aligned}$$

is the sum of products of Mayer functions over the edge configurations forming the sets of admissible pairwise interactions.

To establish the link with motifs we consider a connected graph representing the interactions between l particles, which we refer to as a l-cluster. Every N particle graph can be represented as the decomposition involving several l-clusters. Each l-cluster is a configuration of edges representing the interactions existing between the l component particles. The set of possible node combinations that can form an l-cluster is obtained by permuting the node labels. As a result the configuration integral is a sum over all cluster configurations and a product over the edges contained within them. This leads to a potentially exponential

growth in the summation. To simplify matters we note that in practice the edge configurations on different sets of particles of the same size are topologically identical. These constitute the set of motifs used to represent the network in hand. The edge configurations of each subgraph is only decided by the topological structure of the graph instead of by specific labelled nodes configurations. So it is easy for us to detect topologically equivalent edge configurations. As a result we can rewrite the sum over configurations as a sum over topologically equivalent motif edge configurations weighted by their frequency of occurrence.

The topologically equivalent edge configurations are the motifs for the network of particle interactions.

Suppose that the set of motifs is χ. The individual motif index is ν and the ν^{th} motif has node-set $\mathcal{L}_\nu$ and edge-set $\mathcal{M}_\nu$. Since the cluster configuration integral for each particle graph is independent of its node or particle labels, the number of terms in the multiple summation in the partition function is thus

$$C(N) = \frac{N!}{\prod_{l=1}^{N}\Big\{m_l!(l!)^{m_l}\Big\}}$$

where l is the number of particles in the lth-cluster and m_l is the number of l-clusters in the network, and we have the constraint $\sum_l lm_l = N$. As a result the normalisation of the expansion is given by

$$B(N) = \prod_{l=1}^{N}(l!V\lambda^{3l-3}b_l)^{m_l}$$

where $\lambda = \sqrt{\frac{2\pi\beta h^2}{\mu}}$ is the average de Broglie wavelength of the particles, and b_l is the cluster integrals as is given in terms of the co-efficients.

$$b_l = \frac{1}{l!\lambda^{3l-3}V}\int ... \int \sum \prod_{i<j\leq l} f_{i,j} d^3\vec{r}_1 d^3\vec{r}_2 ... d^3\vec{r}_l \quad (11)$$

For example, the first three cluster integrals can be calculated based on Eq. 11

$$\begin{aligned} b_1 =& \frac{1}{V}\int d^3r = 1 \\ b_2 =& \frac{1}{2!\lambda^3 V}\int f_{12} d^3\vec{r}_1 d^3\vec{r}_2 = \frac{1}{2\lambda^3}\int f_{12} d^3\vec{r}_{12} \\ b_3 =& \frac{1}{3!\lambda^6 V}\int (f_{12}f_{23} + f_{12}f_{13} + f_{13}f_{23} + f_{12}f_{23}f_{13}) \\ & d^3\vec{r}_1 d^3\vec{r}_2 d^3\vec{r}_3 \end{aligned} \quad (12)$$

Here for b_1, the cluster has only one particle when $l = 1$. For b_2 the cluster is composed of two connected particles when $l = 2$. For b_3 the cluster consists of three particles and the three particles are connected to each other to form four different connected graphs topologies, and so on.

As a result the configuration integral becomes the product of the number of items and each item:

$$\begin{aligned} Q &= N!\sum_{\nu\in\chi}\prod_{l=1}^{|\mathcal{L}_\nu|}\frac{\Big\{V\lambda^{3l-3}b_l\Big\}^{m_l}}{m_l!} \\ &= N!\lambda^{3N}\sum_{\nu\in\chi}\prod_{l=1}^{|\mathcal{L}_\nu|}\frac{1}{m_l!}\Big\{\frac{V}{\lambda^3}b_l\Big\}^{m_l} \end{aligned} \quad (13)$$

Finally, the motif-based partition function can be obtained by combining Eq.4 and Eq.13

$$Z = \sum_{\nu\in\chi}\prod_{l=1}^{|\mathcal{L}_\nu|}\frac{1}{m_l!}\Big\{\frac{V}{\lambda^3}b_l\Big\}^{m_l} \quad (14)$$

When specified in this way, the various global thermodynamic characterizations of the gas can be computed. For instance, the average energy of the network can be expressed in terms of the Hamiltonian operator and the partition function as

$$\langle U\rangle = -\frac{\partial}{\partial\beta}\ln Z \quad (15)$$

and the thermodynamic entropy by

$$S = k\Big\{\ln Z + \beta\langle U\rangle\Big\} \quad (16)$$

where k_B s the Boltzman constant, which later on we set to unity since it is a property of matter, i.e. a physical constant, which does not apply to the analogy exploited here.

Both the energy and the entropy can be regarded as characterizations of the network structure having different properties. In the following sections, we will explore these statistical properties in more detail, and in particular the entropy corresponding to the network motif decomposition.

4 Classical Cluster Expansion and Network Motifs

The classical cluster expansion described above can be used to describe the motif structure of a network using the terms in the expansion to represent different interaction topologies between particles. In the case of motif-based graph representation, we can interpret the term $f_{i,j}$ as the weight between two nodes i and j. The different terms in the configuration integral can thus be expressed in terms of subgraphs.

Therefore, we map the network motifs to the classical cluster expansion. We do this by treating the motifs as the expansion coefficients of the partition function of network structure [4]. We simplify the partition function to the one-dimensional case by replacing the multiple volume integrals by a single scalar radial variable r and by ignoring the dependence on λ which is a constant related to the physical properties of particles. We let n_ν be the number of motifs of type ν analogously to the number of edges m_l in the classical cluster expansion. The corresponding motif configuration integral q_ν as the configuration integral for the ν-th motif, and this plays a similar role to b_l in cluster expansion. The partition function Z for the network can be written as a sum over the individual motif contributions z_ν as,

$$Z = \sum_{\nu}\prod_{n_\nu}^{|\mathcal{L}_\nu|}\frac{1}{n_\nu!}\Big\{rq_\nu\Big\}^{n_\nu} = \sum_{n_\nu} z_\nu \quad (17)$$

which is the sum of all possible motifs for the N particles.

The global motif-based partition function Z is simply the sum of partition function contributions from all motifs with different index values ν, denoted by z_ν. The motif partition function is obtained through a product of the integrals for all possible configurations. Consider the partition function contribution from the motif indexed ν. We can view the nodes

of the graph as being drawn from a binomial distribution, i.e. those belonging to the ν^{th} motif type and those that are disjoint, i.e. singleton nodes. We combine their contributions to the network partition function by calculating the motif configuration integral and the single node integral respectively. As a result the configuration integral contribution from the ν^{th} set of motif configurations is

$$\frac{1}{n_\nu!}\{rq_\nu\}^{n_\nu}$$

where n_ν is the frequency of the occurrence of the ν^{th} motif, r is the radial variable, q_ν is the configuration integral for the ν^{th} motif, which plays a similar role to b_l in the cluster expansion. For disjoint nodes, on the other hand, the frequency of occurrence is $N - l_\nu n_\nu$ where l_ν is the number of nodes in the ν^{th} motif. The configuration integral for the singleton node is q_0 which we set to unity to allow such nodes without penalty. As a result, the configuration integral for the non-disjoint nodes is

$$\frac{1}{(N-l_\nu n_\nu)!}(rq_0)^{N-l_\nu n_\nu}, \quad \text{if} N - l_\nu n_\nu \geq 0$$

. The partition function for motif ν is thus

$$z_\nu = \frac{1}{n_\nu!}(rq_\nu)^{n_\nu}\frac{1}{(N-l_\nu n_\nu)!}(rq_0)^{N-l_\nu n_\nu}, \quad \text{if} N - l_\nu n_\nu \geq 0 \tag{18}$$

According to the classical cluster expansion, b_l represents the configuration integral of l-th cluster in Eq.11, and q_ν is the configuration integral of the ν-th motif in the network. In this setting the quantity analogous to b_l defined above is

$$q_\nu = \frac{1}{l_\nu! r}\int ... \int \sum \prod_{i<j\leq l} f_{i,j} d^3\vec{r}_1 d^3\vec{r}_2 ... d^3\vec{r}_l = \frac{1}{l_\nu! r}\zeta_\nu \tag{19}$$

where ζ_ν is the configuration integral obtained through the product over all edges connecting nodes in the ν-th motif.

4.1 Motif Expansion

For graphs, we use symbol d_ν to represent the number of edges in the ν-th motif, thus the number of the nodes without connecting edges is $l_\nu - d_\nu$. Since the ζ_ν is the integral over all edges in the ν-th motif, we separate the configuration integral into contributions from those nodes connected by edges and those that are disjoint. The value of the integral for the edges is ϵ, while that for disjoint nodes is ϵ_0. We allow the disjoint nodes without penalty. The main reason for this is that balancing an additional parameter proves difficult, and can lead to over fragmented networks with many disjoint nodes. So we operate with $\epsilon_0 = 1$. So the motif configuration integral is:

$$\begin{aligned}\zeta_\nu &= \epsilon_0^{l_\nu - d_\nu}(\epsilon)^{d_\nu} \\ &= (\epsilon)^{d_\nu}\end{aligned} \tag{20}$$

According to the Mayer function in Eq.7, the configuration integral ϵ for an edge is given by

$$\epsilon = \int_0^\infty (\exp[-\beta v(r)] - 1)dr \tag{21}$$

where

$$v(r) = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6]$$

is the Lennard-Jones potential function, and d_ν and l_ν are the number of edges and nodes in the motif indexed ν respectively.

We employ Simpson's Method [36] to evaluate the edge-integral numerically. as follows

$$\begin{aligned}\epsilon &= \sum_{r=r_{min}}^{r_{max}} (\exp[-\beta v(r)] - 1) \\ &= \exp[\beta \sum_{r=r_{min}}^{r_{max}} e^{-v(r)}] - \frac{r_{max} - r_{min}}{\Delta r}\end{aligned} \tag{22}$$

where $\Delta r \to 0$ is the increment size for r and $[r_{min}, r_{max}]$ is the interval of integration. As a result the Simpson rule approximation is

$$\epsilon = p\exp[\beta] + R \tag{23}$$

where

$$\begin{aligned}p &= \sum_{r=r_{min}}^{r_{max}} \exp[-v(r)] \\ &\text{and} \\ R &= -\frac{r_{max} - r_{min}}{\Delta r}\end{aligned} \tag{24}$$

Taking the first three motifs in Fig.1 as examples, the motif cluster integrals based on Eq. 19 Eq. 20 and Eq. 21 are

$$\begin{aligned}q_1 &= \frac{1}{2r}\int_0^\infty f_{12} d\vec{r}_1 d\vec{r}_2 = \frac{1}{2}\int_0^\infty f_{12} d\vec{r}_{12} = \frac{1}{2}\epsilon \\ q_2 &= \frac{1}{6r}\int_0^\infty (f_{12}f_{23} + f_{12}f_{13} + f_{13}f_{23}) d\vec{r}_1 d\vec{r}_2 d\vec{r}_3 \\ &= \frac{3}{6}\int_0^\infty f_{12}f_{23} d\vec{r}_{12} d\vec{r}_{23} = \frac{1}{2}(\int_0^\infty f_{12} d\vec{r}_{12})^2 = \frac{1}{2}\epsilon^2 \\ q_3 &= \frac{1}{6r}\int_0^\infty (f_{12}f_{23}f_{13}) d\vec{r}_1 d\vec{r}_2 d\vec{r}_3 \\ &= \frac{1}{6r}(\int_0^\infty f_{12} d\vec{r}_{12})^3 = \frac{1}{6r}\epsilon^3\end{aligned} \tag{25}$$

In the case of q_1, the first motif is composed of two nodes connected to each other by an edge. For q_2, the second motif contains three nodes connected by two edges. For q_3, the third motif is a fully connected graph with three nodes and three edges.

We render the floating point calculations tractable and avoid overflows in the computation of the partition function in Eq.(18) by taking logarithms

$$\begin{aligned}\log z_\nu = &-\log n_\nu! + n_\nu(\log r + \log q_\nu) - \log(N - l_\nu n_\nu)! \\ &+ (N - l_\nu n_\nu)\log r\end{aligned} \tag{26}$$

where we have q_0=1, since we do not penalise disjoint single nodes. We use Stirling's approximation [37] to find the approximate value of the logarithms of factorials of large numbers ($\log N! \simeq N\log N$). This approximation scheme applied to compute $\log n_\nu!$ and $\log(N - l_\nu n_\nu)!$. Substituting for q_ν, we find after algebra

$$\log z_\nu = n_\nu\Big\{\log rq_\nu - \log n_\nu\Big\} + (N - l_\nu n_\nu)\log\frac{r}{N - l_\nu n_\nu} \tag{27}$$

Substituting Equation (20) for ζ_ν, into Equation (19) for q_ν we find $q_\nu = \epsilon^{d_\nu}/(l_\nu! r)$, and hence

$$\log z_\nu = n_\nu \left\{ (d_\nu \log \epsilon - l_\nu \log l_\nu - \log n_\nu) \right\} + (N - l_\nu n_\nu) \log \frac{r}{N - l_\nu n_\nu} \tag{28}$$

The only quantity in the above expression for the log-partition function which depends on β is ϵ. As a result, from Eq.(15) the average energy of the νth motif is given by

$$\langle U_\nu \rangle = -\frac{\partial \ln z_\nu}{\partial \beta} = \frac{n_\nu d_\nu p e^\beta}{p e^\beta + R} \tag{29}$$

So the average motif energy is proportional to d_ν, the number of edges in the motif. When the temperature is low, i.e. β is large, then $\langle U_\nu \rangle \simeq \frac{d_\nu}{p}$. In other words it depends on the number of edges in the motif and the edge-strength parameter p derived from the potential. When the temperature is large, i.e. β is small, then $\langle U_\nu \rangle \simeq \frac{d_\nu}{p+R}$. So as temperature increases the average energy of the motif decreases, and in the case of large networks is n_ν/R, which is independent of the potential strength.

Turning our attention to the corresponding entropy from the νth motif, from Eq.(16)

$$\begin{aligned} S_\nu &= \log z_\nu + \beta \langle U \rangle \\ &= n_\nu \left\{ (d_\nu \log \epsilon - l_\nu \log l_\nu - \log n_\nu) \right\} \\ &+ (N - l_\nu n_\nu) \log \frac{r}{N - l_\nu n_\nu} + \beta \frac{n_\nu d_\nu p e^\beta}{p e^\beta + R} \end{aligned} \tag{30}$$

Letting $\epsilon_0 = r q_0$, and writing ϵ for $e^\beta + R$ and $\epsilon = R$ for e^β, this becomes

$$S_\nu = n_\nu \left\{ l_\nu \log \frac{\epsilon_0}{l_\nu} + d_\nu \left[\log \frac{\epsilon}{\epsilon_0} - \beta \frac{(\epsilon - R)}{\epsilon} \right] - \log n_\nu \right\} + (N - l_\nu n_\nu) \log \frac{r q_0}{N - l_\nu n_\nu} \tag{31}$$

We can re-express the entropy in terms of the number of nodes appearing in motifs of type ν, i.e. $V_\nu = n_\nu l_\nu$ and the total number of edges $E_\nu = n_\nu d_\nu$ appearing in the motif, with the result

$$S_\nu = N \log \frac{r q_0}{N - V_\nu} + V_\nu \log \left[\frac{\epsilon_0}{l_\nu} \frac{N - V_\nu}{r q_0} \right] - n_\nu \log n_\nu + E_\nu \log \left[\frac{\epsilon}{\epsilon_0} \exp \left[-\beta (1 - \frac{R}{\epsilon}) \right] \right] \tag{32}$$

There are four terms appearing in this expression for the network entropy. The first three terms are independent of temperature. The first of these terms is proportional to the number of nodes in the network and increases as the number of nodes in the motif of type ν increases. The second term is proportional to the number of nodes contained within the clique of type ν, and decreases as the size l_ν of the motif increases. The third term decreases with the increasing frequency of the motif of type ν and controls the distribution of motif frequencies. Finally, the fourth term is proportional to the number of edges contained in motifs of the type ν, and is controlled by the reciprocal temperature β and the value of the edge-strength ϵ which is itself temperature dependant. As β approaches zero, i.e. the temperature becomes large $\epsilon = p + R$ and the fourth term approaches $E_\nu \ln \frac{p+R}{\epsilon_0}$, while as β becomes large and the temperature approaches zero, then the term approaches the lower value $E_\nu \log \frac{p}{\epsilon}$. Entropy is hence minimised by an equitable distribution of nodes and edges among the available motifs, and an equitable population of different motifs type. The effect of the distribution of edges among motifs weakens with increasing temperature. In other words at low temperature the variations in the distributions of edges among different motifs is more important than at high temperature.

Algorithm 1 gives the pseudocode for the overall method.

Algorithm 1: Summary of the Graph Motif Entropy Method

Input: graph $G = (V, E)$, type of motifs m, inverse temperature β, scale parameter σ
Output: a set of motif entropy for each graph $G\_i = \{S\_1, ..., S_\nu, ..., S\_m\}$

initialization, Adjacency matrix of the Graph A, Node number of the Graph: N, Node number of the ν^{th} motif: l_ν, Edge number of the ν^{th} motif: d_ν;
Detecting graph motifs(Detailed algorithm in Appendix A). Get a set of motif number for each type of motif $\{n_1, ..., n_\nu, ..., n_m\}$;
Compute the edge configuration integral with formula(20): ζ_ν;
Get the configuration integral of the ν^{th} motif with formula(19):q_ν;
Calculate the partition function of the graph i in ν^{th} motif with formula(18):Z_ν;
Compute the average energy with partition function in formula(15): $< U_\nu >$;
Get motif entropy S_ν with formula(16);
Compute each type of motif entropy for graph i and get $\{S\_1, ..., S_\nu, ..., S\_m\}$.;
return G_i;

5 EXPERIMENTS

In this section, we use the thermodynamic expressions for the motif energy and entropy to analyse data for the motif representation of time-evolving complex networks. There are four aspects to our experimental evaluation. First, we construct principle component analysis for motif entropy vectors. Second, we perform entropy component analysis on the vector of motif entropies computed using our method. Third, we apply the C-SVM to the motif entropy kernel and then compare it with the results obtained with alternative kernel methods. Finally, we evaluate the influence of the physical parameters of our method on the motif entropy and the influence of the motif frequency on the final network entropy.

5.1 Principal Component Analysis for Motif Entropy Vectors

5.1.1 DataSets

Synthetic Networks: We generate synthetic graphs according to three different and widely studied complex network models, namely, a) the Erdős-Rényi random graph model, b) the Watts-Strogatz small-world model, c) the Barabási-Albert scale-free model. These graphs are created with a fixed number of vertices with time varying network parameters. For the Erdős-Rényi random graph, the connection probability monotonically increases at the uniform rate of 0.005 per unit time over an interval of 200 time units. Similarly, the link rewiring probability in the small-world model increases uniformly between 0 to 1 as the network evolves over the time interval studied. For scale-free model, one vertex is added to the connection at each time step. The aim in this study is to explore whether our method can distinguish graphs generated by the different network models.

AIDS Networks: The AIDs dataset [38] consists of 2000 graphs constructed from the AIDS Antiviral Screen Database. The compound molecules are converted into graphs in a straightforward manner by representing atoms as nodes and the covalent bonds as edges. Each graph can be classified into one of two classes, namely a) active and b) inactive, which respectively represent molecules with and without activity against HIV. Our method has also been tested to determine whether it can accurately classify the compound data into the two activity classes.

Finacial Networks: We test our method on dataset extracted from the New York Stock Exchange (NYSE) database [39]. This encapsulates the performance of 347 stock using their associated daily closing prices over 6004 trading days from January 1986 to February 2011. To extract the network representations, we use a time window of 28 days and move this window along time to obtain a sequence (from day 29 to day 6004), in which each time window contains a series of daily stock prices over a period of 28 days. For each time window, we compute the cross correlation coefficients between the closing price time series for each pair of stock. We create connections between pairs of stock if the maximum absolute value of the correlation coefficient is among the highest 5 % of the cumulative cross correlation coefficient distribution. By doing this, the trades between different stock are represented as a network with a fixed number of 347 nodes and varying edge structure for each of 5976 trading days [40]. Here, the nodes in the network corresponds to the stock of companies trading in the stock market over the complete 6000 days period. In our dataset, the 347 stocks contained in the NYSE database are the stocks of listed companies that have existed from January 1986 to February 2011. Therefore, the time series financial networks that we construct with these 347 stocks as nodes has a fixed number of 347 nodes. The set of stock considered, i.e. the node set of the financial network is constant over time, and stock (nodes) neither appear or disappear. The edge-structure on the other hand varies with time, and edges appear and disappear between different pairs of stock. In real-world applications, the financial market can be considered as a complex time-varying system consisting of multiple interacting financial components. To analyze the time-varying financial market crisis or risk, change point detection has played an important role to identify abrupt changes in the time series properties. Unfortunately, detecting such crucial points remains challenging, since it is difficult to detect the changes that cannot be easily observed for a system consisting of complex interactions between its constituent co-evolving time series. One way to overcome this problem is to represent multiple co-evolving financial time series as a family of dynamic time-varying networks [41], [42]. The aim in this study is to determine whether our network characterisation can be used to detect fluctuations in trading network structure due to global political or economic events, such as the attacks on the World Trade Centre or the Lehman Brothers collapse.

5.2 Entropy Component Analysis

We use the motif entropies to compute a feature vector for each graph, and then perform principal components analysis (PCA) on the set of vectors for a sample of graphs. These samples are generated by network time series. Here, we choose PCA [43] not only to reduce the dimensionality of the data, but also to facilitate the visualization of the embedding results. PCA is a standard and easy way to analyse the data. To be more formal for each graph we construct a feature vector $\vec{x} = (S_1, ..., S_\nu, ...)^T$ whose components S_ν are the total entropies of the different kinds of motifs occurring in a graph. Suppose we have a sample of M graphs $G_1, ..., G_M$ and that the total motif entropy vector for the graph G_i is $\vec{x}_i$. The mean motif total entropy vector is $\hat{\vec{x}} = \frac{1}{M}\sum_{i=1}^{M}\vec{x}_i$, and the sample covariance matrix is

$$\Sigma = \frac{1}{M}\sum_{i=1}^{M}(\vec{x}_i - \hat{\vec{x}})(\vec{x}_i - \hat{\vec{X}})^T.$$

The eigen-decomposition of the covariance matrix is $\Sigma = \Phi\Lambda\Phi^T$, where Λ is a diagonal eigenvalue matrix with ordered eigenvalues on the diagonal and Φ the eigenvector matrix with the correspondingly ordered eigenvectors as columns. We project the centred feature vectors onto the space spanned by the covariance matrix eigenvectors. This gives rotated feature vectors $\vec{\vec{x}}_i = \Phi^T(\vec{x}_i - \hat{\vec{x}})$, and display the leading three components of the centred and rotated vectors.

5.2.1 Experimental Settings

To count the frequencies of the motif types shown in Fig.1, we use the method outlined in [5]. We compute the motif entropy for different datasets, and the parameter settings used are listed in Table 2. The two physical parameters of the model are the inverse temperature β and the scale parameter σ appearing in the potential.

In our experiments, these parameters are set to the default values $\beta = 100$ and $\sigma = 9$. We will discuss the impact of varying the values of these parameters in detail through the experiments reported in Section 5.3. We evaluate our method on synthetic networks and AIDS networks in Section 5.2.2 and then further validate it on financial networks extracted from NYSE time series data in Section 5.2.3.

TABLE 2
Parameters for different data-sets

Datasets	Number of graph	Motifs used	Number of nodes	β	σ
Synthetic graph	200	8	500	100	9
AIDS	2000	8	Varies in different graphs	100	9
Finance Graph	5976	8	347	100	9

5.2.2 Graph characterization for synthetic and AIDS networks

We aim to explore whether the motif entropy can be used to distinguish between graphs using principal components analysis on samples of motif entropy vectors, i.e. entropy component analysis. For synthetic networks, i.e. the Erdos-Renyi, small world and Watts-Strogatz graphs, the embedding corresponding to the leading three eigenvectors is shown in Figure 2. Here the red points represent the small-world networks, the green ones scale-free networks and the blue ones represent the Erdos-Reyni random graphs. The three populations are clearly distinguished from each other by the entropy component embedding. In other words, the three different graph models lead to different populations with well separated means in the leading three principal components for the motif entropy vectors.

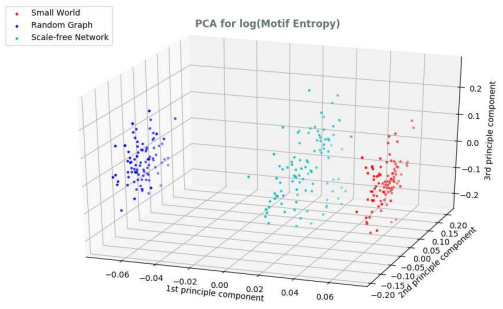

Fig. 2. PCA on the motif entropy vectors for synthetic networks.

For the AIDS dataset, the embedding corresponding to the first three motif entropy eigenvectors is shown in Figure 3. Here the red points represent the active compounds and the blue points represent the inactive compounds. The two classes are clearly separated into two distinct clusters by the PCA embedding of the motif entropy principal components representation.

5.2.3 Financial Networks

Thermodynamic measures for network evolution analysis: We continue our study by exploring whether the motif entropy can be used for better understanding the time evolution of realistic complex networks. To this end, we first compare the evolutionary behavior of motif entropy and Von Neumann entropy for the NYSE stock market data. At each time step, we compute both the motif entropy and Von Neumann entropy. For a graph G with adjacency matrix A, according to Passerini and Severini [44] the von-Neumann entropy is $S_{vN}(G) = -Tr[\frac{\tilde{L}}{|V|}\ln\frac{\tilde{L}}{|V|}]$ where

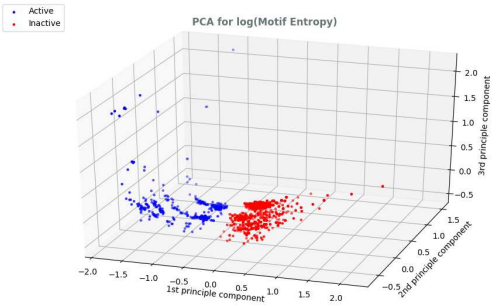

Fig. 3. PCA on the motif entropy vectors for the AIDS graphs.

$\tilde{L} = D^{-1/2}(D-A)D^{-1/2}$ is the normalised Laplacian matrix of the graph.

After processing the original closing price sequence data, we obtain 5976 different daily samples of the stock correlation network as it evolves with time. At each time step (or trading day) t, there is a sample of the network. Here, we calculate the first motif entropy and the von Neumann entropy for each sample of the graph, and plot the two entropy values for the stock network as it evolves with time. The basic idea is to explore whether the resulting motif entropy can be used to measure the changes in structure of the stock network and compare this to the corresponding result obtained using the Von Neumann entropy. Figure 4 shows the time series for the motif entropy(upper) and the Von Neumann entropy(lower) for 5976 trading days. When a financial crisis occurs, the stock market network experiences dramatic structural changes. We have annotated the plot with labels indicating some of these events. Figure 4 indicates that these structural changes can be effectively detected by both the time series representations of motif entropy and Von Neumann entropy, since the fluctuations in the two time series correspond closely to most of the well defined financial crises. In addition, the motif entropy outperforms Von Neumann entropy in the sense that the motif entropy appears more stable (i.e. is less noisy) during the intervals between crises.

Time Series Embeddings To better explore the performance of the proposed method for characterizing time-evolving networks, especially for detecting temporal anomalies, we perform embeddings of the graph time series. For the motif entropy vectors, we use PCA and for the Von Neumann Entropy, we use kernel PCA.

Kernel PCA on Von Neumann Entropy: We use the graph von Neumann entropies to construct a kernel matrix, and then perform linear kernel principal components analysis to embed the sample of graphs into a vector

Fig. 4. The motif entropy (upper, blue) and Von Neumann entropy (lower, green) versus time for the dynamic stock correlation network. The known financial crisis periods are identified by ellipses. e.g. the Madrod Bombings in 2004.3, the Hurricane Katrina in 2005.8.

space. Let H be the matrix of entropy differences with element $H(i,j) = ||S_{VN}(i) - S_{VN}(j)||$. We use the entropy similarity matrix to compute a symmetric kernel matrix $K = -1/2(I - J/M)H(I - J/M)$ where I is the $M \times M$ identity matrix and $J = ee^T$ where $e = (1, 1, ..., 1)^T$ is the all-ones vector of length M. We perform kernel embedding on the matrix K. To this end let Y be the matrix with the embedding co-ordinates of the graphs as columns, then $K = X^T X$. Performing the eigen-decomposition $K = U\Lambda U^T$, the matrix of embedding co-ordinates is $X = \sqrt{\Lambda}U^T$. We visualise the distribution of the graphs using the first three rows of X corresponding to the leading three eigenvectors of the kernel matrix.

The results are shown in Figure 5 and Figure 6. Figure 5 shows the entire time series during the 5976 trading days with the different financial crisis shown with different coloured markers. However, only the Black Monday event (black triangles) can be identified in all four embeddings, while the remaining financial events cannot be detected easily. When we compare 5 (a) and (b), it is clear that the motif entropy gives a relatively compact manifold structure which reflects the characteristics of time evolution. The Von Neumann entropy, on the other hand, does not give an easily interpreted manifold structure.

To take our study one step further, we show the embeddings during a short time interval around two different financial crisis to compare the two methods in more detail. Figure 6 illustrates the structural changes in the embedding spaces before and after crucial events generated from motif entropy and Von Neumann entropy, respectively. The blue star represent the exact day before the crisis occurred while the red stars represent the period during the crises. From the figure, it is clear that both types of event can be effectively detected by the two entropies. However, the embeddings of the graphs before and after the crisis generated by Von Neumann entropy are overlapped. On the other hand,

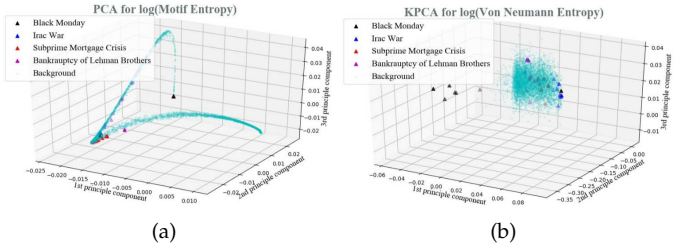

Fig. 5. PCA embedding plots for the time evolving stock correlation network characterization. The time series covers the 5976 days. The financial crisis period is represented by triangular symbols of different colors, while the remaining periods are represented by green dots, and these form the background. a) shows the PCA embedding for the 8-component vector representing the motif entropies for each graph, b) is obtained from the kernel PCA of Von Neumann entropy.

those generated from motif entropy are clearly separated into distinct clusters. This indicates that our motif entropy outperforms Von Neumann entropy in terms of its capacity to generate a clear manifold structure.

To take the analysis one step further, in Figure 7, we show a set of points indicating the path of the stock network in the space spanned by graph entropy graph energy and time (ordinal number of the graph in the time series sequence). We explore three time intervals namely in the proximity of (a) Black Monday(1987.10.19) , (b) Friday the 13th mini-crash(1989.10.13) and (c) Asian Financial Crisis(1997.7-1997.10). The colored bar beside each plot gives the colour coding of the ordinal numbers of the days spanned by the time series. The top panel shows that before Black Monday (blue and green triangles), the network structure remains stable. During Black Monday, on the other hand, the network undergoes an abrupt jump away from the trajectory prior to the crisis in the embedding space. This is followed by a gradual return after the crisis which still shows a compact manifold structure. Different

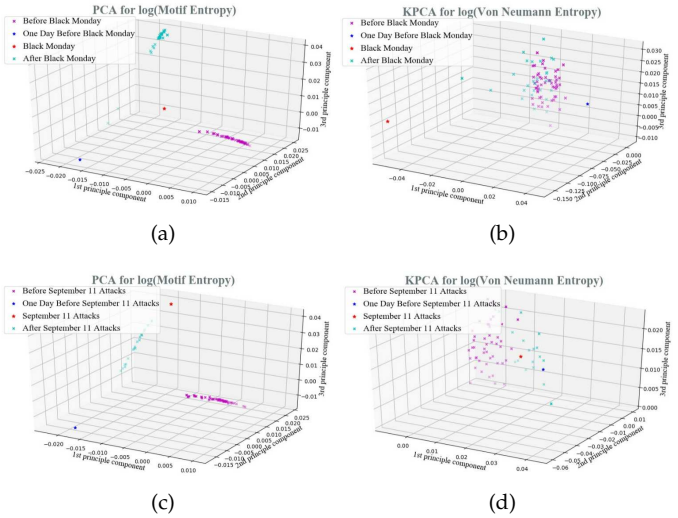

Fig. 6. Embeddings in the proximity of the Black Monday event. Figure 6 (a) and (b) depict the spatial distribution of embeddings obtained by PCA of the Motif entropy vectors and kernel PCA on the von Neumann entropy during Black Monday, respectively. While Figure 6 (c) and (d) illustrates the distribution around the period of the September 11 attacks.

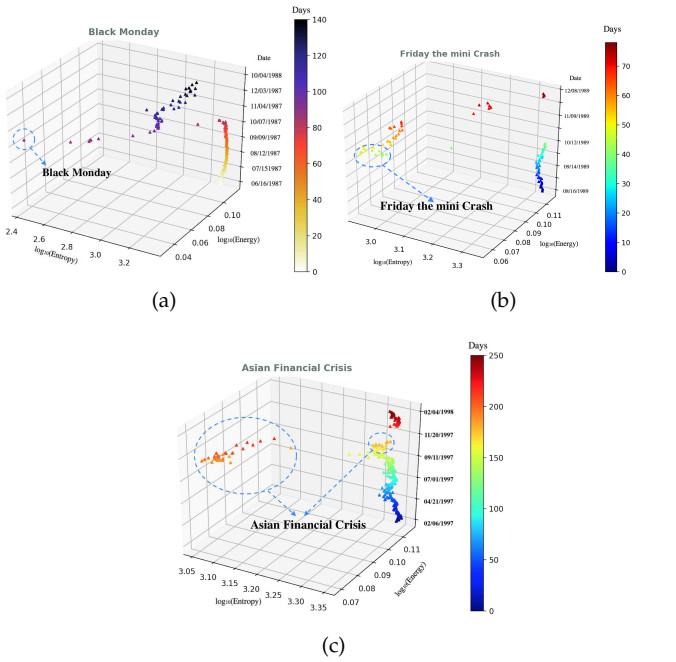

Fig. 7. Path of the time-evolving stock correlation network in the entropy-energy-time space during different financial crises. The panel are for a) Black Monday, b) Friday the 13th mini Crash and c) Asian Financial Crisis, respectively. The colored bar on the right hand side of each plot represents the colour code corresponding to positions of graphs in the time series.

behaviors can be observed concerning the Friday 13th mini-Crash and the Asian Financial Crisis. The former gives rise to a gradual change in the manifold structure, rather than an abrupt jump (as in the case of Black Monday) and then returns progressively to its normal state. In the latter case though, the stock market network suffers a significant change in structure during the crisis. This is signaled by a large decrease in both network energy and entropy. The market crash is followed by a rapid recovery. Hence, in addition of detecting crucial events, our thermodynamic motif representation can be utilized to distinguish between different types of financial crises, and probe their temporal dependence in more detail.

5.3 C-SVM on Graph Classification

In this section, we use the motif entropy to construct the entropy kernel for graphs, and compare it with several kernel methods on graph classification tasks.

5.3.1 DataSets

MUTAG is a data set of 188 chemical compounds where the class label is as either aromatic or heteroaromatic with seven node features. **PPIs** protein-protein interaction (PPI) networks, whose structure is represented by undirected graphs. There are 219 PPIs in this dataset and they are collected from 5 different kinds of bacteria. We select two kinds of bacteria, i.e. Proteobacteria40 PPIs and Acidobacteria46 PPIs. **PTC** comprises 344 compounds where the class label indicates whether they are carcinogenic or not in rats with 19 node features. The **NCI1** dataset made publicly available by the National Cancer Institute (NCI) is a subset of balanced datasets of chemical compounds screened for the ability to suppress or inhibit the growth of tumours. It consists of 4100 graphs that represent chemical compounds and each node is assigned one of 37 possible labels.

5.3.2 Entropy Kernel

Here, we design the motif entropy kernel based on an entropy kernel associated with dynamic time warping framework. For each graph with index i, we have a feature vector $\vec{x}_i = (S_1, ..., S_8)$ whose components are the eight kinds of motif entropy. We compute kernel matrix through $K_{ab} = \vec{x}_a \vec{x}_b^T$ where $\vec{x}_a$ and $\vec{x}_b$ is the motif entropy vector for graph a and b. We perform C-SVM on the output of our entropy kernel for the purposes of classification, and compare with several alternative state-of-art graph kernels. The alternative kernels are the dot product kernel (DP)([45]), the Jensen-Shannon kernel(JS)([45]), the Weisfeiler-Lehman subtree kernel(WLSK)([46]), the quantum Jensen-Shannon kernel associated with continues time(QJSK)([47]), the quantum Jensen-Shannon kernel associated with discrete time(QJSKT)([48]), the shortest path graph kernel(SPGK)([49]), the Jensen-Shannon graph kernel(JSGK)([50]), and the back-trackless version of the random walk kernel(BRWK)([51]). For each kernel, we compute the kernel matrix on each graph dataset. We perform a 10-fold cross-validation where the classification accuracy is computed using a C-Support Vector Machine(C-SVM). In particular, we make use of the LIBSVM library. For each dataset and each kernel, we compute the optimal C-SVM parameters. We repeat the whole experiment 10 times and report the average classification accuracies ($\pm$ standard error).

5.3.3 Classification Accuracy

The result of classification accuracy is shown in Table.3 and the corresponding running time is shown in Table.4. We can find that our method performed better than other methods in classification accuracy at the expense of a slight

TABLE 3
Classification accuracy(in %± standard error) runtime in second

Datasets	Mutag	PPIs	PTC	NCI1
K_{DP}	82.38 ± 0.55	88.62 ± 0.86	58.62 ± 0.69	85.03 ± 0.12
K_{JS}	85.44 ± 0.58	90.62 ± 0.90	60.70 ± 0.62	85.18 ± 0.10
$WLSK$	82.05 ± 0.57	78.50 ± 1.40	56.05 ± 0.51	80.68 ± 0.27
$QJSK$	83.83 ± 0.49	70.57 ± 1.40	58.23 ± 0.80	67.40 ± 0.20
$QJSKT$	81.55 ± 0.53	68.12 ± 0.84	57.44 ± 0.36	67.00 ± 0.15
$SPGK$	83.38 ± 0.81	61.12 ± 1.09	56.55 ± 0.53	74.21 ± 0.30
$JSGK$	83.11 ± 0.80	57.87 ± 1.36	57.29 ± 0.41	62.50 ± 0.33
$BRWK$	77.50 ± 0.75	53.50 ± 1.47	53.97 ± 0.31	60.34 ± 0.17
KE	$\mathbf{87.06 \pm 0.77}$	$\mathbf{90.83 \pm 1.23}$	$\mathbf{61.59 \pm 0.74}$	$\mathbf{85.76 \pm 0.21}$

TABLE 4
Runtime for various kernels(second)

Datasets	Mutag	PPIs	PTC	NCI1
K_{DP}	2.7×10^1	5.5×10^1	2.7×10^1	5.4×10^2
K_{JS}	1.2×10^1	1.7×10^2	2.7×10^2	4.1×10^4
$WLSK$	0.4×10^1	1.3×10^1	1.1×10^1	1.5×10^2
$QJSK$	1.2×10^1	1.4×10^4	1.1×10^2	1.6×10^4
$QJSKT$	2.9×10^1	1.5×10^2	1.7×10^2	1.4×10^4
$SPGK$	0.1×10^1	0.7×10^1	0.1×10^1	8.3×10^1
$JSGK$	0.1×10^1	0.1×10^1	0.1×10^1	0.1×10^1
$BRWK$	0.1×10^1	8.6×10^2	0.3×10^1	4.1×10^2
KE	$\mathbf{3.9 \times 10^1}$	$\mathbf{2.3 \times 10^2}$	$\mathbf{3.6 \times 10^2}$	$\mathbf{4.4 \times 10^4}$

increase in computation time. We also study the effect of Temperature($\frac{1}{\beta}$) on the classification accuracy and the results are shown in Fig.8

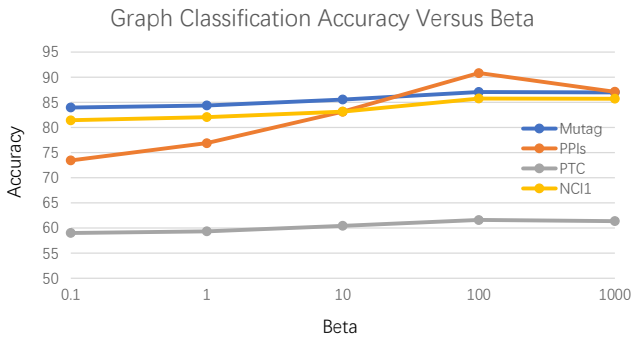

Fig. 8. Classification accuracy varying with inverse temperature β.

5.4 Parameter Settings

The aim in this section is to explore the parameter dependence of the motif model presented in this paper. We commence by exploring how the entropy is distributed among motifs of different size. We plot the motif entropies for the Financial Networks versus time, and observe how they behave during the different crises. In Fig.9, we vary the motif size from 1 to 8, and plot the different motif entropies as a function of time. Each of the motifs is sensitive to the complete set of crises. However motifs of sizes 1, 2 and 7 performs better when measured in terms of the amplitude of entropy variation in the proximity of the crises, and the smoothness in non-crises regions. This may be that motifs 1 and 2 are the most frequent, and can thus represent the vast majority of the topological information variance residing in a graph. Motif 7, on the other hand, may vary greatly in frequency of occurrence in different graphs. It is thus sensitive to the differences between different graphs. As for motif 8, there are few occurrences in any of the graphs studied.

We also explored the effect of varying the temperature. We varied temperature from 0.01 to 1000, and investigated the effect on the first motif entropy. From Fig.10, it is clear that the temperature plays the role of a smoothing parameter, controlling the amplitude of the background entropy variations but not affecting the behaviour in the proximity of the crises.

In order to explore the relationship between network energy and entropy, we studied the evolution of energy and entropy of different motifs in the Financial time series dataset. As is shown in Fig.11, the behaviour of the energy of the network follows that of the motif entropy, no matter the size. Both quantities can identify the financial crisis events. It also indicates that the motif entropy subsumes the information contained in the energy and vice-versa.

We further explore whether the detection motif method that we proposed can effectively reveal the basic mechanisms of the networks through the number of different motifs and their variations over time. As shown in Fig.12, the number of motifs also exhibits significant variation with time during the financial crisis. In particular, variation in the frequencies of the first two motifs is much greater than the remainder. As expected this coincides with the behaviour of the motif entropies in Fig.9. We can represent each graph as a vector $\vec{X} = (S_1, ..., S_\nu, ...)^T$, whose components S_ν are the total entropies of the different kinds of motifs occurring in the graph. As we increase the number of selected motif types, the dimensionality of the embedding associated with the graph entropy vector increases. This means that the entropy components more accurately represent the total graph entropy. Moreover, the crises are more cleanly separated by

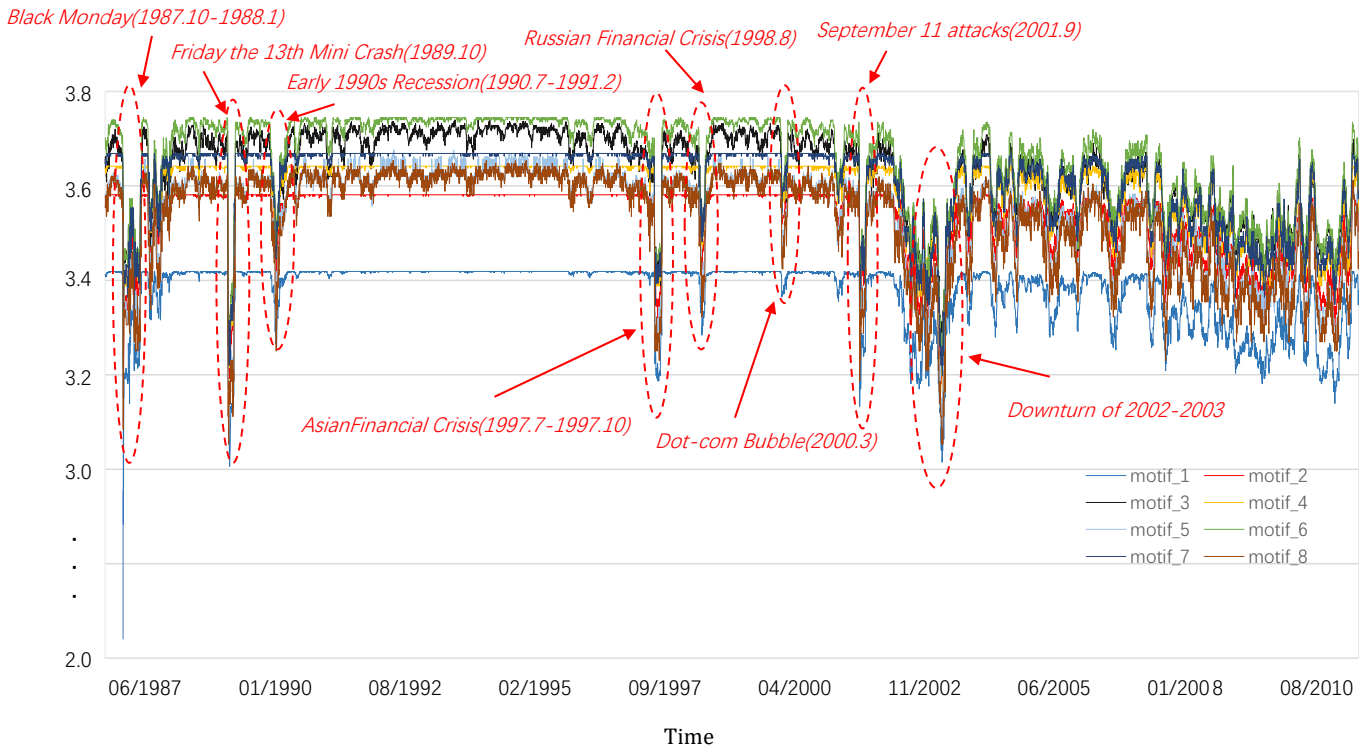

Fig. 9. Lines of eight colors represent the time-varying entropies of the 1-st motif to the 8-th motif, respectively.

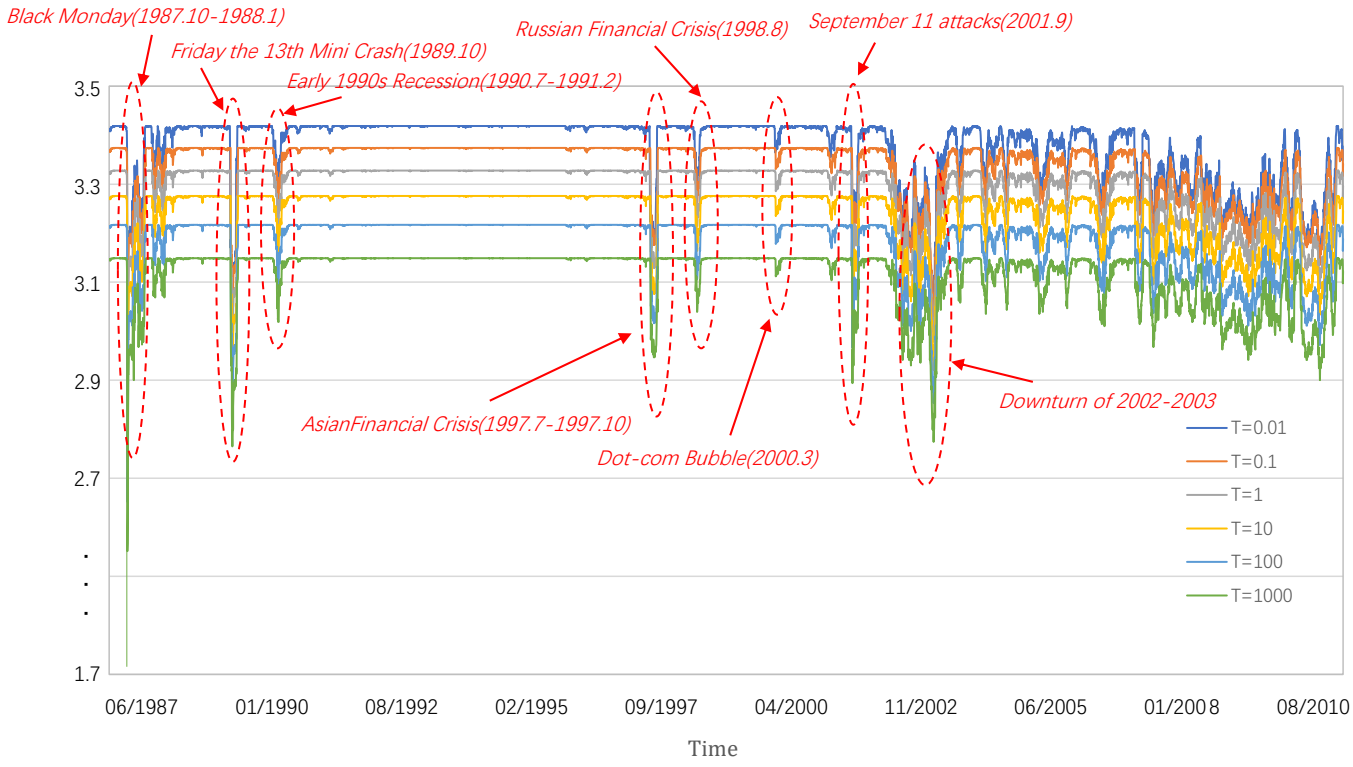

Fig. 10. The first motif entropy on Financial Networks versus time under different temperatures.

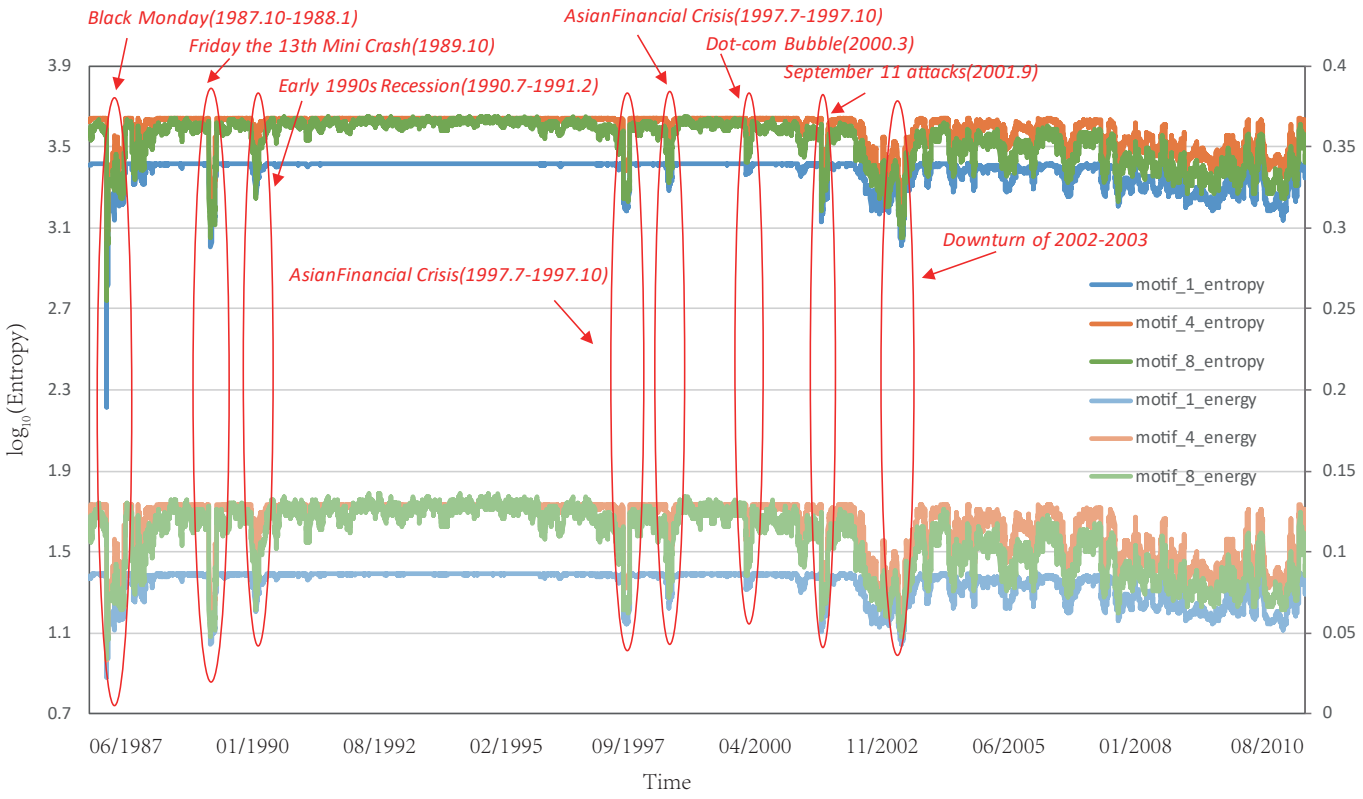

Fig. 11. The network energy and motif entropy versus time.

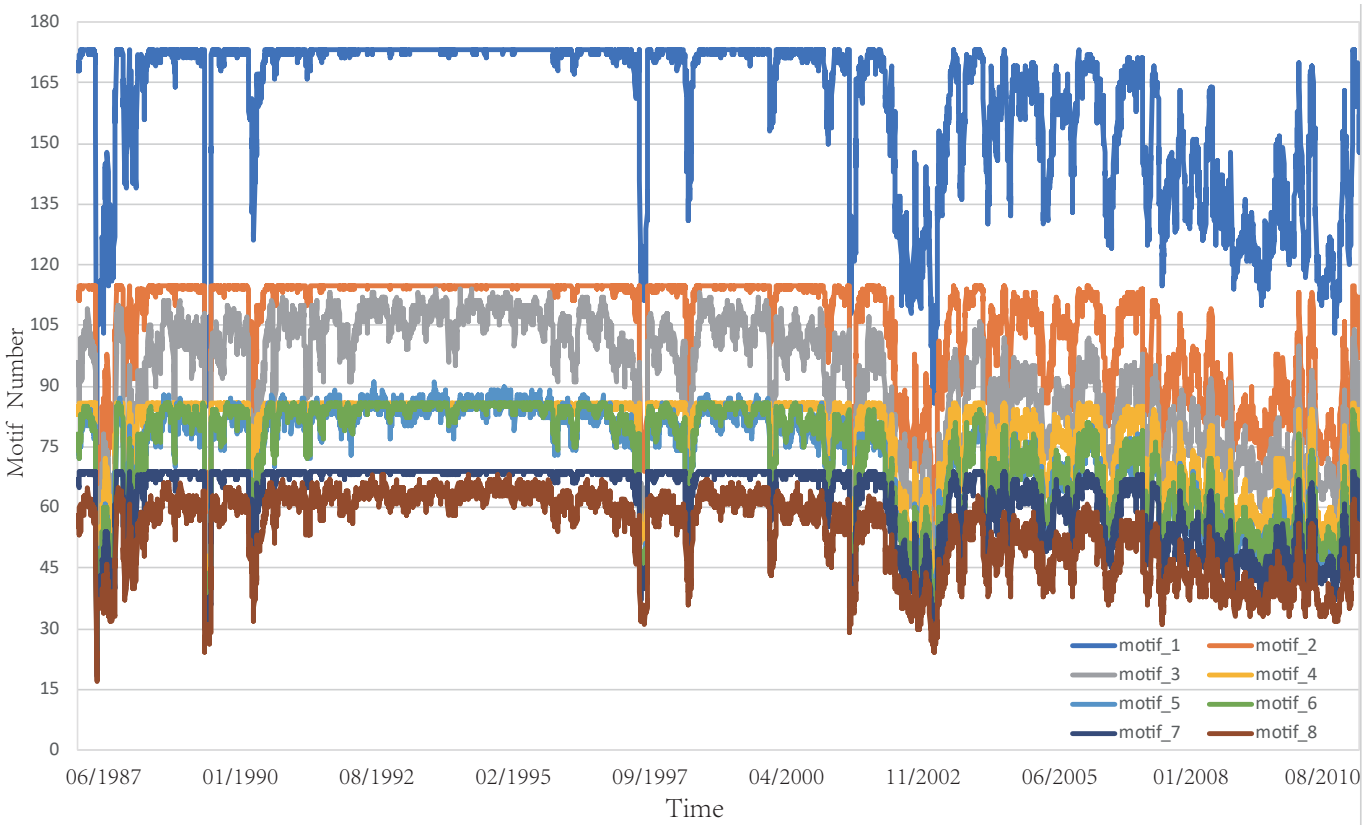

Fig. 12. The motif number versus time.

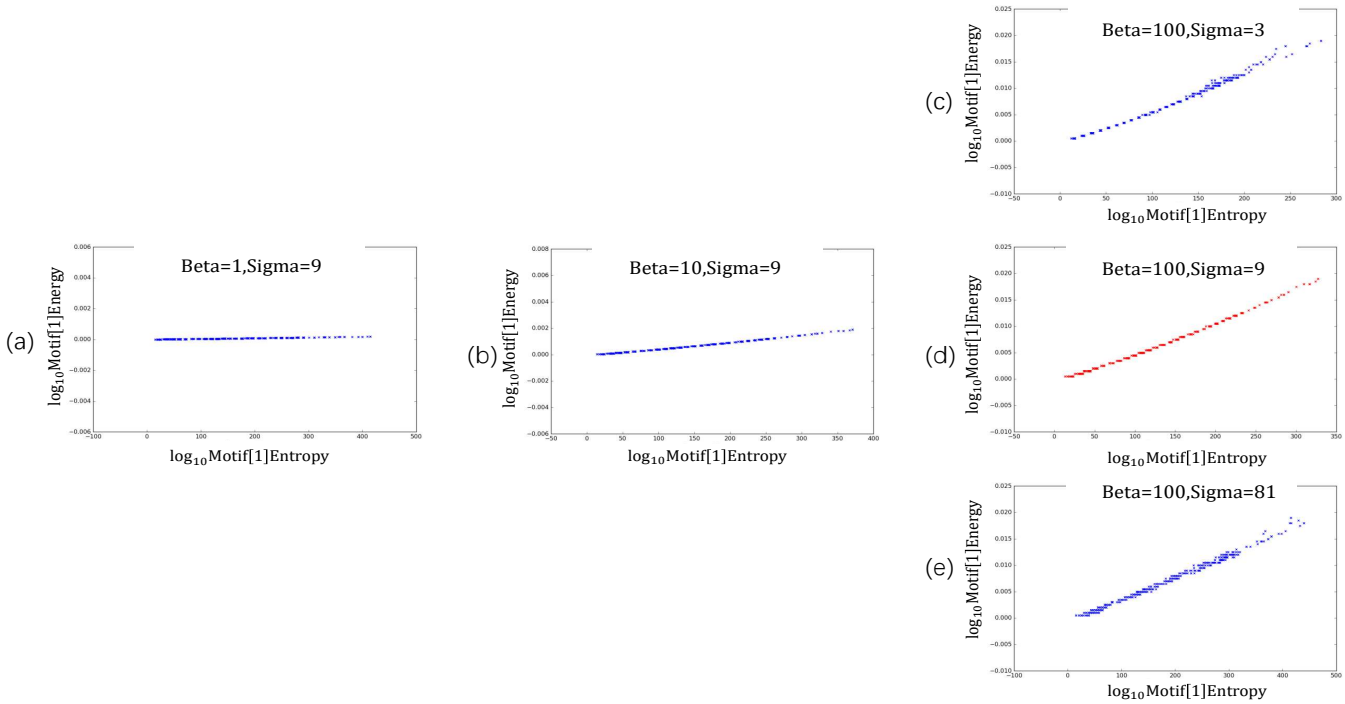

Fig. 13. Energy versus Motif Entropy on AIDS.

the entropies than by the raw motif frequencies.

We also evaluate the effect of parameters β and σ on the energy and motif entropy in the AIDS dataset. In Fig.13, the fitted curve in the scatter plot of energy and motif entropy, we varied σ and β separately to show their influence to the energy and entropy. As we can see in Figs.13(b), 13(c) and Figs.13(a), the fitted curve tends to rise with β becomes larger, which also means that the range of energy becomes wider. Since the larger the range of energy and entropy value describing the network, the finer the granularity of description and the more information it conveys. This is a desirable effect. These plots also show that at low temperature(inverse β) the variations in the distributions of number of edges is more important than at high temperature (as expected from our preliminary analysis). When $\beta = 100$, the slope of the fitted energy curve is maximum, and even if β increases significantly it remains unchanged. We continue by fixing β and varying σ. Comparing Fig.13(a),Fig.13(d) and Fig.13(e) when σ increases, the spread of entropy for a fixed value of energy increases. One energy value corresponds to a wider range motif entropies, while the regression relationship between energy and entropy becomes more scattered when σ is small.

6 Conclusion

In this paper, we have explored how to model the role of network motifs in determining network energy and entropy. To do this we make use of an analogy with the cluster expansion in statistical physics. We capture the motif content of a network using a partition function. This treatment leads to expressions for the thermodynamic energy and entropy, which can be used to characterize the structural properties of the network. We present an analytical solution for the numbers of network motifs and the scaling of all types of motifs using the partition functions. This allows us to compute energy and entropy using the cluster expansion in terms of network motifs. We conduct the experiments on both synthetic networks and real-world networks. For real-world networks, we focus on network time-series representing stock trades on the NYSE. Our model is capable of detecting abrupt changes or anomalies in network structure and distinguishing different types of time-dependency for different types of anomaly.

The work reported in this paper can clearly be extended in 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 on a greater variety of data. Another interesting line of investigation would be to explore the uses of network motifs in classification problem to identify different types of structures. Finally, we plan to extend this work to quantum statistics to further investigate the properties of network motifs in the quantum domain, where effects such as particle spin become important.

Acknowledgment

This work is supported by the Research Funds of State Grid Shaanxi Electric Power Company and State Grid Shaanxi Information and Telecommunication Company (contract no.SGSNXT00GCJS1900134), the National Natural Science Foundation of China (Grant no.61976235 and 61602535) and the Program for Innovation Research in Central University of Finance and Economics.

REFERENCES

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[2] M. Zitnik, R. Sosic, and J. Leskovec, "Prioritizing network communities," in *Nature communications*, 2018.

[3] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[4] S. Itzkovitz and U. Alon, "Subgraphs and network motifs in geometric networks," *Physical Review E*, vol. 71, no. 2, p. 026117, 2005.

[5] U. Alon, "Network motifs: theory and experimental approaches," *Nature Reviews Genetics*, vol. 8, no. 6, p. 450, 2007.

[6] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[7] S. Mangan and U. Alon, "Structure and function of the feed-forward loop network motif," *Proceedings of the National Academy of Sciences*, vol. 100, no. 21, pp. 11 980–11 985, 2003.

[8] H. Y. L. S. Yifeng Zhao, Xiangwei Wang and J. Tang, "Large scale evolving graphs with burst detection," *Proceedings of the 28th International Joint Conference on Artificial Intelligence.*, 2019.

[9] F. Feng, X. He, J. Tang, and T.-S. Chua, "Graph adversarial training: Dynamically regularizing based on graph structure," *IEEE Transaction on Knowledge and Data Engineering.*, 2019.

[10] E. Wong, B. Baur, S. Quader, and C.-H. Huang, "Biological network motif detection: principles and practice," *Briefings in bioinformatics*, vol. 13, no. 2, pp. 202–215, 2011.

[11] H. S. Wilf, *generatingfunctionology*. AK Peters/CRC Press, 2005.

[12] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[13] J. Wang, R. C. Wilson, and E. R. Hancock, "Spin statistics, partition functions and network entropy," *Journal of Complex Networks*, vol. 5, no. 6, pp. 858–883, 2017.

[14] E. E. Salpeter, "On mayer's theory of cluster expansions," *Annals of Physics*, vol. 5, no. 3, pp. 183–223, 1958.

[15] K. Husimi, "Note on mayers' theory of cluster integrals," *The Journal of Chemical Physics*, vol. 18, no. 5, pp. 682–684, 1950.

[16] C. E. Tsourakakis, J. Pachocki, and M. Mitzenmacher, "Scalable motif-aware graph clustering," in *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2017, pp. 1451–1460.

[17] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, "Superfamilies of designed and evolved networks," *Science*, vol. 303, no. 5663, pp. 1538–1542, 2004.

[18] Y. T. Maeda and M. Sano, "Regulatory dynamics of synthetic gene networks with positive feedback," *Journal of molecular biology*, vol. 359, no. 4, pp. 1107–1124, 2006.

[19] S. Kalir, S. Mangan, and U. Alon, "A coherent feed-forward loop with a sum input function prolongs flagella expression in escherichia coli," *Molecular systems biology*, vol. 1, no. 1, 2005.

[20] M. Ronen, R. Rosenberg, B. I. Shraiman, and U. Alon, "Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics," *Proceedings of the national academy of sciences*, vol. 99, no. 16, pp. 10 555–10 560, 2002.

[21] N. Rosenfeld and U. Alon, "Response delays and the structure of transcription networks," *Journal of molecular biology*, vol. 329, no. 4, pp. 645–654, 2003.

[22] A. Awan, H. Bari, F. Yan, S. Moksong, S. Yang, S. Chowdhury, Q. Cui, Z. Yu, E. Purisima, and E. Wang, "Regulatory network motifs and hotspots of cancer genes in a mammalian cellular signalling network," *IET Systems Biology*, vol. 1, no. 5, pp. 292–297, 2007.

[23] L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii, "Structural properties of the caenorhabditis elegans neuronal network," *PLoS computational biology*, vol. 7, no. 2, p. e1001066, 2011.

[24] E. Yeger-Lotem, S. Sattath, N. Kashtan, S. Itzkovitz, R. Milo, R. Y. Pinter, U. Alon, and H. Margalit, "Network motifs in integrated cellular networks of transcription–regulation and protein–protein interaction," *Proceedings of the National Academy of Sciences*, vol. 101, no. 16, pp. 5934–5939, 2004.

[25] T. Milenković, I. Filippis, M. Lappe, and N. Pržulj, "Optimized null model for protein structure networks," *PLoS One*, vol. 4, no. 6, p. e5967, 2009.

[26] J. E. Mayer, "The statistical mechanics of condensing systems. i," *The Journal of chemical physics*, vol. 5, no. 1, pp. 67–73, 1937.

[27] B. Kahn and G. E. Uhlenbeck, "On the theory of condensation," *Physica*, vol. 5, no. 5, pp. 399–416, 1938.

[28] T. D. Lee and C. N. Yang, "Many-body problem in quantum statistical mechanics. i. general formulation," *Physical Review*, vol. 113, no. 5, p. 1165, 1959.

[29] R. Peierls, "Bemerkungen über umwandlungstemperaturen," *Helv. Phys. Acta*, vol. 7, no. 2, p. 81, 1934.

[30] B. Halperin and D. R. Nelson, "Theory of two-dimensional melting," *Physical Review Letters*, vol. 41, no. 2, p. 121, 1978.

[31] A. Young, "Melting and the vector coulomb gas in two dimensions," *Physical Review B*, vol. 19, no. 4, p. 1855, 1979.

[32] E. A. Wong and B. Baur, "On network tools for network motif finding: a survey study," *Online Google Scholar*, 2010.

[33] A. Arenas, A. Fernandez, S. Fortunato, and S. Gomez, "Motif-based communities in complex networks," *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 22, p. 224001, 2008.

[34] H. Yin, A. R. Benson, and J. Leskovec, "Higher-order clustering in networks," *Physical review. E*, vol. 97 5-1, p. 052306, 2018.

[35] D. Zwanziger, "Fundamental modular region, boltzmann factor and area law in lattice theory," *Nuclear Physics B*, vol. 412, no. 3, pp. 657–730, 1994.

[36] P. J. Davis and P. Rabinowitz, *Methods of numerical integration*. Courier Corporation, 2007.

[37] N. Mermin, "Stirling's formula!" *American Journal of Physics*, vol. 52, p. 362–365, 1984.

[38] H. Bunke and K. Riesen, "Proc. 12th iberoamerican congress on pattern recognition," ser. LNCS 4756, L. Rueda, D. Mery, and J. Kittler, Eds., 2007, pp. 20–31.

[39] F. N. Silva, C. H. Comin, T. K. D. Peron, F. A. Rodrigues, Y. Cheng, R. C. Wilson, E. Hancock, and L. D. F. Costa, "Modular dynamics of financial market networks," 2015.

[40] C. Ye, C. H. Comin, T. K. D. Peron, F. N. Silva, F. A. Rodrigues, L. d. F. Costa, A. Torsello, and E. R. Hancock, "Thermodynamic characterization of networks using graph polynomials," *Physical Review E Statistical Nonlinear and Soft Matter Physics*, vol. 92, no. 3, p. 032810, 2015.

[41] Y. Yin, P. Shang, and J. Xia, "Compositional segmentation of time series in the financial markets," *Appl. Math. Comput.*, vol. 268, pp. 399–412, 2015.

[42] L. Song, M. Yamada, N. Collier, and M. Sugiyama, *Change-Point Detection in Time-Series Data by Relative Density-Ratio Estimation*.

[43] K. M. Lever, J. and N. Altman, *Principal component analysis*.

[44] F. Passerini and S. Severini, "The von neumann entropy of networks," *SSRN Electronic Journal*, no. 12538, 2008.

[45] L. Bai, L. Rossi, L. Cui, J. Cheng, and E. R. Hancock, "A quantum-inspired similarity measure for the analysis of complete weighted graphs," *IEEE transactions on cybernetics*, 2019.

[46] N. Shervashidze, P. Schweitzer, E. Jan, V. Leeuwen, and K. M. Borgwardt, "Weisfeiler-lehman graph kernels," *Journal of Machine Learning Research*, vol. 12, no. 3, pp. 2539–2561, 2011.

[47] B. Lu, L. Rossi, R. Peng, Z. Zhang, and E. R. Hancock, "A quantum jensen-shannon graph kernel using discrete-time quantum walks," in *International Workshop on Graph-based Representations in Pattern Recognition*, 2015.

[48] B. Lu, L. Rossi, L. Cui, Z. Zhang, R. Peng, B. Xiao, and E. Hancock, "Quantum kernels for unattributed graphs using discrete-time quantum walks," *Pattern Recognition Letters*, vol. 87, no. C, pp. 96–103, 2017.

[49] K. M. Borgwardt and H. P. Kriegel, "Shortest-path kernels on graphs," in *IEEE International Conference on Data Mining*, 2006.

[50] L. Bai and E. R. Hancock, "Graph kernels from the jensen-shannon divergence," *Journal of Mathematical Imaging and Vision*, vol. 47, no. 1-2, pp. 60–69, 2013.

[51] F. Aziz, R. C. Wilson, and E. R. Hancock, "Backtrackless walks on a graph," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 6, pp. 977–989, 2013.