A Quantum-inspired Entropic Kernel for Multiple Financial Time Series Analysis

Lu Bai[1,2], **Lixin Cui**[1,2*], **Yue Wang**[1,3], **Yuhang Jiao**[1] and **Edwin R. Hancock**[4]
[1]Central University of Finance and Economics, Beijing, China
[2]China Center for Fintech Studies, Central University of Finance and Economics, Beijing, China
[3]State Key Laboratory of Cognitive Intelligence, iFLYTEK, Hefei, China
[4]Department of Computer Science, University of York, York, UK
{bailucs, cuilixin, wangyuecs}@cufe.edu.cn, edwin.hancock@york.ac.uk

Abstract

Network representations are powerful tools for the analysis of time-varying financial complex systems consisting of multiple co-evolving financial time series, e.g., stock prices, etc. In this work, we develop a new kernel-based similarity measure between dynamic time-varying financial networks. Our ideas is to transform each original financial network into quantum-based entropy time series and compute the similarity measure based on the classical dynamic time warping framework associated with the entropy time series. The proposed method bridges the gap between graph kernels and the classical dynamic time warping framework for multiple financial time series analysis. Experiments on time-varying networks abstracted from financial time series of New York Stock Exchange (NYSE) database demonstrate that our approach can effectively discriminate the abrupt structural changes in terms of the extreme financial events.

1 Introduction

Network representations are powerful tools to analyze the financial market that can be considered as a time-varying complex system consisting of multiple co-evolving financial time series [Zhang and Small, 2006; Nicolis *et al.*, 2005; Shimada *et al.*, 2008; Silva *et al.*, 2015], e.g., the stock market with the trade price. This is based on the idea that the structure of the so-called time-varying financial networks [Bullmore and Sporns, 2009] inferred from the corresponding time series of the system can represent richer physical interactions between system entities than the original individual time series. One main objective of existing approaches is to detect the extreme financial events that can significantly influence the network structures [Bai *et al.*, 2020].

In machine learning, graph kernels have been widely employed for analyzing structured data represented by graphs or networks [Xu *et al.*, 2018]. The main advantage of employing graph kernels is that they can offer us an effective way of mapping the network structures into a high dimensional space so that the standard kernel machinery for vectorial data is applicable to the network analysis. Most existing graph kernels are based on the idea of decomposing graphs or networks into substructures and then measuring pairs of isomorphic substructures [Haussler, 1999], e.g., graph kernels based on counting pairs of isomorphic a) paths [Borgwardt and Kriegel, 2005], b) walks [Kashima *et al.*, 2003], and c) subgraphs [Bai *et al.*, 2015b] or subtrees [Shervashidze *et al.*, 2009]. Unfortunately, directly adopting these graph kernels to analyze the time-varying financial networks inferred from original vectorial time series tends to be elusive. This is because these financial network structures are by nature complete weighted graphs [Ye *et al.*, 2015; Bai *et al.*, 2020], where each vertex represents an individual time series of a stock and is adjacent to all remainder vertices, and each edge represents the interaction (e.g., the correlation or distance) between a pair of coevolving financial time series. It is difficult to decompose a complete weighted graph into the required substructures, and thus influences the effectiveness of most existing graph kernels for financial network analysis.

One way to address the aforementioned problem is to construct sparse structures of the original time-varying financial networks. With this scenario, Cui et al. [Cui *et al.*, 2018] have used the well-known threshold-based approach to preserve the weighted edges falling into the larger 10% of the weights, and employed the classical graph kernels associated with the resulting sparse structures for financial network analysis. Bai et al. [Bai *et al.*, 2020] have abstracted the minimum or maximum spanning trees associated with the commute time matrix of the original complete weighted financial networks, and developed a novel quantum graph kernel over the spanning trees of the financial networks. Although, both the approaches overcome the restriction of employing graph kernels for time-varying financial network analysis, their required sparse structures also lead to significant information loss. Since many weighted edges of the original complete weighted financial networks are discarded. In summary, analyzing time-varying financial networks associated with graph kernels still remains challenges.

The aim of this paper is to overcome the aforementioned problems by developing a new kernel measure between timevarying networks for multiple co-evolving financial time series analysis. Overall, the main contributions are threefold.

First, for a family of time-varying financial network-

*Corresponding Author

s, we commence by computing the average mixing matrix [Godsil, 2013] to summarize the time-averaged behaviour of continuous-time quantum walks (CTQW) evolved on the network structures. The reason of using the CTQW is that it not only accommodates complete weighted graphs, but also better reflects richer financial network characteristics than the classical random walks [Bai *et al.*, 2015a] (see details in Section II-A). We show how the average mixing matrix of the CTQW allows to compute a quantum-based entropy for each vertex of the financial networks and represents the original networks as quantum entropy time series.

Second, with each pair of time-varying financial networks to hand, we define a Quantum-inspired Entropic Kernel between their quantum entropy time series through the classical dynamic time warping framework. The proposed kernel not only accommodates the complete weighted graphs through the entropy time series, but also bridges the gap between graph kernels and the classical dynamic time warping framework for time series analysis (see details in Section III-B).

Third, we perform the proposed kernel on time-varying financial networks abstracted from multiple co-evolving financial time series of New York Stock Exchange (NYSE) database. Experiments demonstrate that the proposed approach can effectively discriminate the abrupt structural changes in terms of the extreme financial events.

2 Preliminary Concepts

In this section, we briefly review some preliminary concepts.

2.1 The Average Mixing Matrix of the CTQW

The continuous-time quantum walk (CTQW) is the quantum analogue of the classical continuous-time random walk (CTRW) [Farhi and Gutmann, 1998]. The CTQW models a Markovian diffusion process over the vertices of a graph through their transition information. Assume a sample graph is $G(V,E)$, where V is the vertex set and E is the edge set. Similar to the classical CTRW, the state space of the CTQW is the vertex set V and its state at time t is a complex linear combination of the basis states $|u\rangle$, i.e., $|\psi(t)\rangle = \sum_{u\in V} \alpha_u(t)\,|u\rangle$, where $\alpha_u(t) \in \mathbb{C}$ and $|\psi(t)\rangle \in \mathbb{C}^{|V|}$ are the amplitude and both complex. Furthermore, $\alpha_u(t)\alpha_u^*(t)$ indicates the probability of the CTQW visiting vertex u at time t. $\sum_{u\in V} \alpha_u(t)\alpha_u^*(t) = 1$ and $\alpha_u(t)\alpha_u^*(t) \in [0,1]$, for all $u \in V$, $t \in \mathbb{R}^+$. Unlike the classical CTRW, the CTQW evolves based on the Schrödinger equation

$$\partial/\partial t\,|\psi_t\rangle = -i\mathcal{H}\,|\psi_t\rangle\,, \tag{1}$$

where $\mathcal{H}$ denotes the system Hamiltonian. In this work, we employ the adjacency matrix as the Hamiltonian. When a CTQW evolves on the sample graph $G(V,E)$, the behaviour of the walk at time t can be summarized using the mixing matrix [Godsil, 2013]

$$M(t) = U(t) \circ U(-t) = e^{i\mathcal{H}t} \circ e^{-i\mathcal{H}t}, \tag{2}$$

where $\circ$ denotes the Schur-Hadamard product of two matrices, i.e., $[A \circ B]_{uv} = A_{uv}B_{uv}$. Since U is unitary, $M(t)$ is a doubly stochastic matrix and each entry $M(t)_{uv}$ indicates the probability of the CTQW visiting vertex v at time t when the walk initially starts from vertex u. However, $Q_M(t)$ cannot converge, because $U(t)$ is also norm-preserving. To overcome this problem, we can enforce convergence by taking a time average. Specifically, we take the Cesàro mean and define the average mixing matrix as $Q = \lim_{T\to\infty}\int_0^T Q_M(t)dt$, where each entry Q_{uv} of the average mixing matrix Q represents the average probability for a CTQW to visit vertex v starting from vertex u, and Q is still a doubly stochastic matrix. Godsil [Godsil, 2013] has indicated that the entries of Q are rational numbers. We can easily compute Q from the spectrum of the Hamiltonian $\mathcal{H}$ that can be the adjacency matrix A of G. Let $\lambda_1, \ldots, \lambda_{|V|}$ represent the $|V|$ distinct eigenvalues of $\mathcal{H}$ and $\mathbb{P}_j$ be the matrix representation of the orthogonal projection on the eigenspace associated with the λ_j, i.e., $\mathcal{H} = \sum_{j=1}^{|V|} \lambda_j \mathbb{P}_j$. Then, we rewrite the average mixing matrix Q as

$$Q = \sum_{j=1}^{|V|} \mathbb{P}_j \circ \mathbb{P}_j. \tag{3}$$

Remarks. The CTQW has been successfully employed to develop novel approaches in machine learning and data mining [Bai *et al.*, 2014; Bai *et al.*, 2016], because of the richer structure than their classical counterparts. The reason of utilizing the CTQW in this work is that the state vector of the CTQW is complex-valued and its evolution is governed by a time-varying unitary matrix. By contrast, the state vector of the classical CTRW is real-valued and its evolution is governed by a doubly stochastic matrix. As a result, the behaviour of the CTQW is significantly different from their classical counterpart and possesses a number of important properties. For instance, the CTQW allows interference to take place, and thus reduces the tottering problem arising in the classical CTRW. Furthermore, since the evolution of the CTQW is not dominated by the low frequency components of the Laplacian spectrum, it has better ability to distinguish different graph structures. Finally, the CTQW can accommodate the complete weighted graph, since the Hamiltonian of the CTQW can be the complete weighted adjacency matrix.

2.2 The Dynamic Time Warping Framework

We review the global alignment kernel based on the dynamic time warping framework proposed in [Cuturi, 2011]. Let $\mathbf{T}$ be a set of discrete time series that take values in a space $\mathcal{X}$. For a pair of discrete time series $\mathbf{P} = (p_1, \ldots, p_m) \in \mathbf{T}$ and $\mathbf{Q} = (q_1, \ldots, q_n) \in \mathbf{T}$ with lengths m and n respectively, the alignment π between $\mathbf{P}$ and $\mathbf{Q}$ is defined as a pair of increasing integral vectors (π_p, π_q) of length $l \leq m+n-1$, where $1 = \pi_p(1) \leq \cdots \leq \pi_p(l) = m$ and $1 = \pi_q(1) \leq \cdots \leq \pi_q(l) = n$ such that (π_p, π_q) is assumed to have unitary increments and no simultaneous repetitions. Note that, for $\mathbf{P}$ and $\mathbf{Q}$, each of their elements can be an observation vector with fixed dimensions at a time step. For any index $1 \leq i \leq l-1$, the increment vector of $\pi = (\pi_p, \pi_q)$ satisfies

$$\begin{pmatrix} \pi_p(i+1) - \pi_p(i) \\ \pi_q(i+1) - \pi_q(i) \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \tag{4}$$

Within the framework of the classical dynamic time warping [Cuturi, 2011], the coordinates π_p and π_q of the alignment π define the warping function. Assume $\mathcal{A}(m,n)$ corresponds to a set of all possible alignments between $\mathbf{P}$ and $\mathbf{Q}$, Cuturi [Cuturi, 2011] has proposed a dynamic time warping inspired kernel, namely the Global Alignment Kernel, by considering all the possible alignments in $\mathcal{A}(m,n)$. The kernel is defined as

$$k_{\mathrm{GA}}(\mathbf{P},\mathbf{Q}) = \sum_{\pi\in\mathcal{A}(m,n)} e^{-D_{\mathbf{P},\mathbf{Q}}(\pi)}, \quad (5)$$

where $D_{\mathbf{P},\mathbf{Q}}(\pi)$ is the alignment cost given by

$$D_{\mathbf{P},\mathbf{Q}}(\pi) = \sum_{i=1}^{|\pi|} \varphi(p_{\pi_p(i)}, q_{\pi_q(i)}), \quad (6)$$

and is defined through a local divergence φ that quantifies the discrepancy between each pair of elements $p_i \in \mathbf{P}$ and $q_i \in \mathbf{Q}$. In general, φ is defined as the squared Euclidean distance. Note that, the kernel k_{GA} measures the quality of both the optimal alignment and all other alignments $\pi \in \mathcal{A}(m,n)$, thus it is positive definite. Moreover, k_{GA} provides richer statistical measures of similarity by encapsulating the overall spectrum of the alignment costs $\{D_{\mathbf{P},\mathbf{Q}}(\pi), \pi \in \mathcal{A}(m,n)\}$.

Remarks. The dynamic time warping based Global Alignment Kernel k_{GA} is a powerful tool for analyzing vectorial time series [Mikalsen *et al.*, 2018; Jain, 2019]. To extend k_{GA} into graph kernel domains, Bai et al. [Bai *et al.*, 2018] have developed a nested graph kernel by measuring k_{GA} between the depth-based complexity traces of graphs [Bai and Hancock, 2014]. Specifically, the complexity trace of each graph is computed by measuring the entropies on a family of K-layer expansion subgraphs rooted at its centroid vertex. Although, the nested graph kernel outperforms local substructure based graph kernels [Johansson *et al.*, 2014] on graph classification tasks. Unfortunately, the financial networks are by nature complete weighted graphs and it is difficult to decompose such graphs into required expansion subgraphs rooted at the centroid vertex. Thus, directly preforming the dynamic time warping inspired graph kernel for time-varying financial networks still remains challenges.

3 Kernels for Time-varying Networks

In this section, we propose a Quantum-inspired Entropic Kernel between time-varying networks for multiple co-evolving financial time series analysis. We commence by characterizing each financial network as a discrete quantum entropy time series through the CTQW. Moreover, we define the new kernel associated with the entropy time series, in terms of the classical dynamic time warping framework [Cuturi, 2011].

3.1 The Quantum Entropy Time Series

We introduce how to characterize each financial network structure as the quantum entropy time series through the CTQW. Assume $\mathbf{G} = \{G_1, \ldots, G_p, \ldots, G_q, \ldots, G_T\}$ denotes a family of time-varying financial networks extracted from a complex financial system $\mathbf{S}$ with a specific set of N co-evolving financial time series, i.e., the system has a fixed number of components (e.g., stocks) co-evolving with time. $G_p(V_p, E_p, A_p)$ is the sample network extracted from the system at time step p. For G_p, each individual vertex $v \in V_p$ represents a corresponding time series of a different stock (e.g., the stock price), each edge $e \in E_p$ represents the interaction (e.g., distances or correlations) between a pair of time series, and A_p is the interaction based weighted adjacency matrix. This is a popular way of modelling the multiple co-evolving financial time series as network structures [Silva *et al.*, 2015; Bai *et al.*, 2020]. Note that, since the vertices of each financial network $G_p \in \mathbf{G}$ correspond to the same N components of the system $\mathbf{S}$, all the networks in $\mathbf{G}$ have the same vertex set, whereas the edge sets E_t are quite different with time t.

Specifically, for each financial network $G_p(V_p, E_p, A_p)$ from $\mathbf{G}$ at time p, we first compute the average matrix matrix Q^p associated with the CTQW evolved on G_p. For each i-th vertex $v_i \in V_p$, the i-th row of Q^p gives the time-averaged probability distribution $\mathbf{P}_i$ for the CTQW to visit vertices $v_1, \ldots, v_N \in V$ ($|V_p| = N$) starting from v_i, i.e.,

$$\mathbf{P}_i = \{\mathcal{P}_i(v_1), \ldots, \mathcal{P}_i(v_j), \ldots, \mathcal{P}_i(v_N)\}. \quad (7)$$

where $\mathcal{P}_i(v_j) = Q^p_{i,j}$ is the time-averaged probability of the CTQW visiting v_j from v_i. The quantum based Shannon entropy [Bai *et al.*, 2016] of vertex v_i can be defined as

$$H_S(v_i) = -\sum_{v_j\in V_p} \mathcal{P}_i(v_j)\log\mathcal{P}_i(v_j). \quad (8)$$

As a result, the entropy characteristic vector of G_p associated with the entropies over all its vertices can be defined as

$$E_p = \{H_S(v_1), \ldots, H_S(v_i), \ldots, H_S(v_N)\}^\top, \quad (9)$$

where $H_S(v_i)$ is the quantum Shannon entropy of the i-th vertex v_i of G_p associated with the time-averaged probability distribution residing on the i-th row of Q_p.

We move a time interval of w time steps over all the time-varying networks of the financial system $\mathbf{S}$ to construct a time-varying quantum entropy time series for each network G_p at time p. In this work, we set the value of w as 28. Specifically, for each network G_p, we compute its quantum entropy time series $\mathcal{S}_t$ associated with its time window as

$$\mathcal{S}_p = \{E_{p-w+1}|E_{t-p+2}|\ldots|E_s|\ldots|E_p\}, \quad (10)$$

where each column E_s of $\mathcal{S}_p$ is the entropy characteristic vector of each network $G_s \in \mathbf{G}$ at time s and is defined by Eq.(9). $s \in \{p-w+1, p-w+2, \ldots, p\}$. Obviously, the quantum entropy time series $\mathcal{S}_p$ of G_p encapsulates a family of w time-varying entropy characteristic vectors from G_{p-w+1} at time $p-w+1$ to G_p at time t.

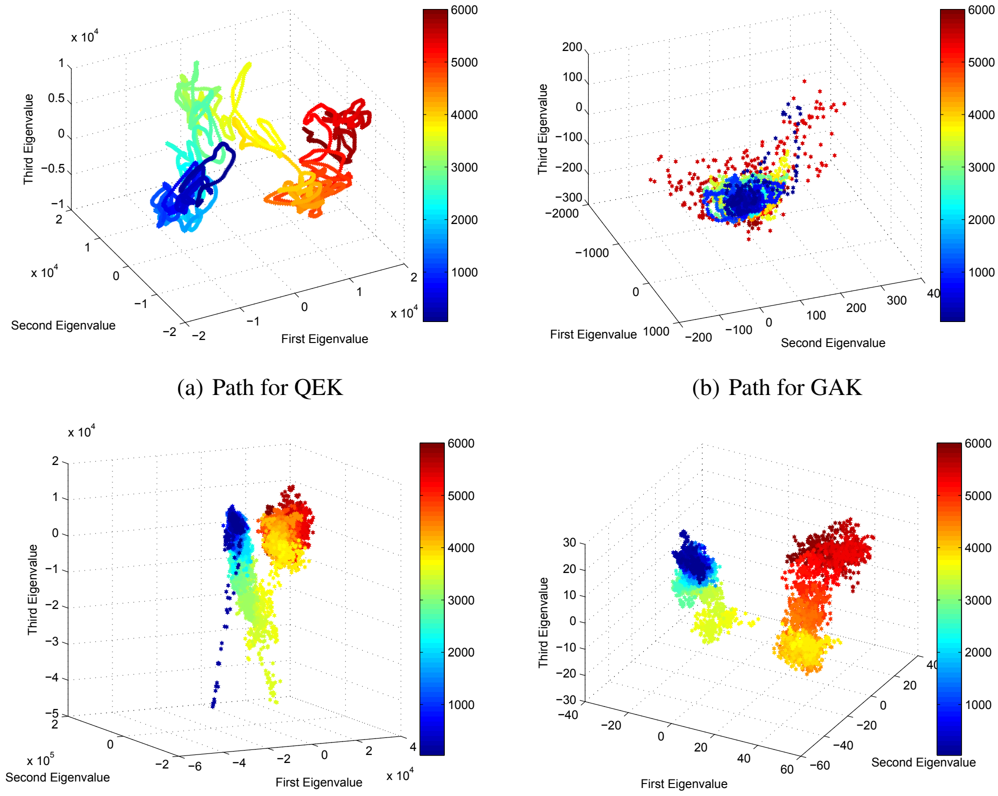
3.2 The Quantum-inspired Entropic Kernel

We develop a new kernel for analyzing time-varying financial networks based on the classical dynamic time warping framework. For a pair of time-varying networks $G_p \in \mathbf{G}$ and $G_q \in \mathbf{G}$ at time p and q respectively, we commence by computing their associated quantum entropy time series as

$$\mathcal{S}_p = \{E_{p-w+1}|E_{p-w+2}|\ldots|E_p\}$$

and

$$\mathcal{S}_q = \{E_{q-w+1}|E_{q-w+2}|\ldots|E_q\},$$



(a) Path for QEK (b) Path for GAK

(c) Path for WLSK (d) Path for DTQK

Figure 1: Color Path of Financial Networks Over All Trading Days.

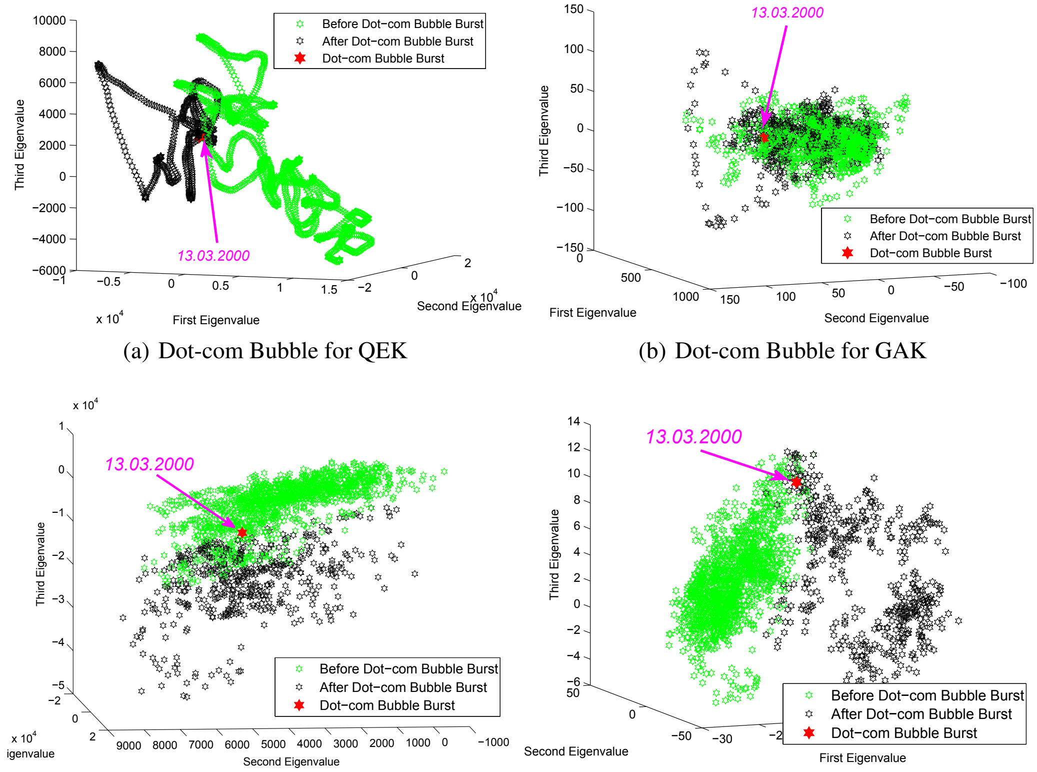
based on the definition in Section 3.1. The proposed Quantum Inspired Entropic Kernel k_{QEK} between G_p and G_q is

$$k_{\mathrm{QEK}}(G_p, G_q) = k_{\mathbf{GA}}(\mathcal{S}_p, \mathcal{S}_q) = \sum_{\pi \in \mathcal{A}(w,w)} e^{-D_{\mathbf{p},\mathbf{q}}(\pi)}, \tag{11}$$

where $k_{\mathbf{GA}}$ is the dynamic time warping inspired Global Alignment Kernel (GAK) defined in Eq.(5), π is the warping alignment between the entropy time series of G_p and G_q, $\mathcal{A}(w, w)$ is all possible alignments and $D_{\mathbf{p},\mathbf{q}}(\pi)$ refers to the alignment cost obtained via Eq.(6). Note that, the proposed kernel k_{QEK} is positive definite. This is because k_{QEK} is based on the positive definite kernel k_{GA}.

Remarks. Although the proposed kernel k_{QEK} is related to the general principles of the GAK kernel. The proposed kernel k_{QEK} still possesses two theoretical differences with the GAK kernel. F**irst**, the original GAK kernel is only developed for vectorial time series and thus cannot capture structural relationships between time series. By contrast, the proposed kernel k_{QEK} is explicitly proposed for time-varying financial networks that encapsulate physical interactions between pairs of time series. **Second**, unlike the GAK kernel, the proposed kernel k_{QEK} is defined based on the quantum entropy time series that is developed through the average mixing matrix of the CTQW. As we have stated in Section 2.1, the CTQW can accommodate the complete weighted graph and better distinguish different network structures in terms of the low frequency components of its Laplacian spectrum. Thus, the proposed kernel k_{QEK} can not only reflect the physical interactions between the original vectorial financial time series, but also capture richer structure information than the GAK kernel associated with the original time series. **On the other hand**, as we have stated, the state-of-the-art graph kernels mentioned in Section 1 and Section 2.2 cannot directly accommodate complete weighted graphs. Thus, it is difficult to directly perform these graph kernels on the complete weighted financial networks, unless one transforms these networks into sparse versions. By contrast, the proposed kernel k_{QEK} can encapsulate the whole structural information residing on all weighted edges. In summary, the proposed kernel k_{QEK} **bridges the gap between state-of-the-art graph kernels and the classical dynamic time warping framework for time-varying networks**, providing an effective way to analyze multiple coevolving financial time series.

Time Complexity. For a pair of networks each having n vertices, computing the kernel k_{QEK} associated with a time interval of w steps requires time complexity $O(n^3 + w^2)$. Because, computing the entropy time series relies on the spectral decomposition of CTQWs, thus has time complexity $O(n^3)$. Computing all possible alignments between the entropy time



(a) Dot-com Bubble for QEK (b) Dot-com Bubble for GAK

(c) Dot-com Bubble for WLSK (d) Dot-com Bubble for DTQK

Figure 2: The 3D kPCA Embeddings of Different Kernels for Dot-com Bubble Burst.

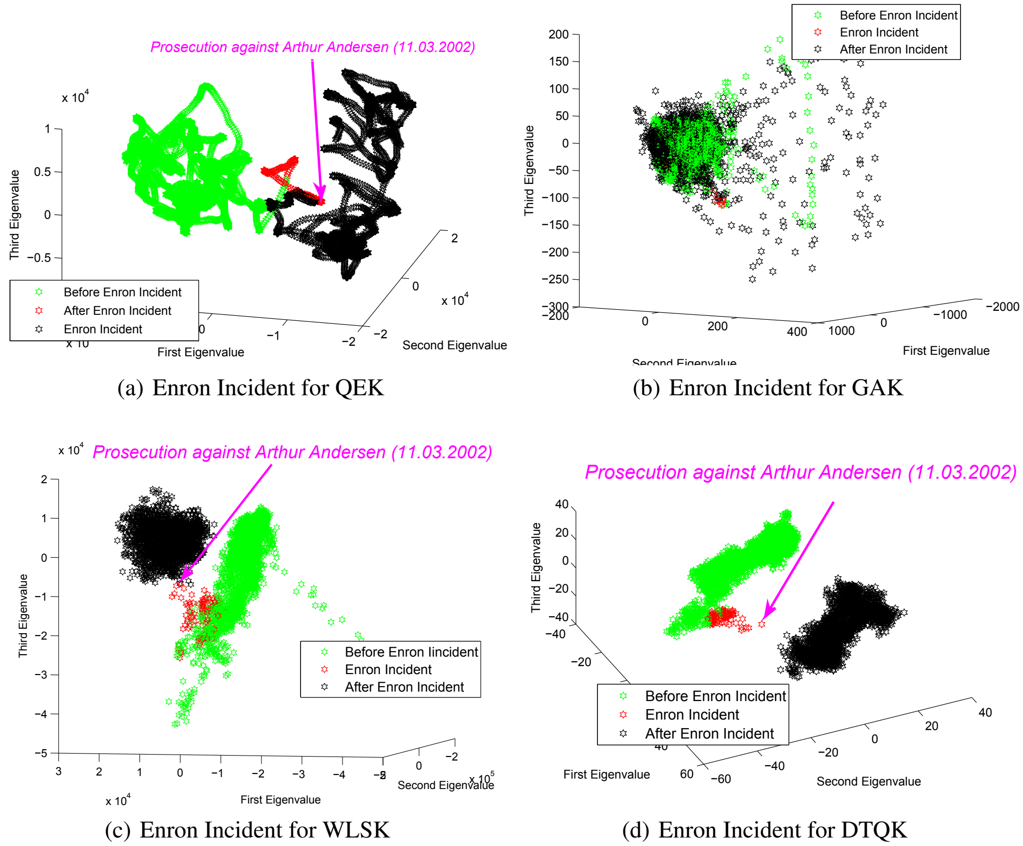
series over w time steps has time complexity $O(w^2)$. Thus, k_{QEK} has a polynomial time complexity $O(n^3 + w^2)$.

4 Experiments of Time Series Analysis

We establish a **NYSE dataset** that consists of a series of time-varying financial networks based on the New York Stock Exchange (NYSE) database [Silva *et al.*, 2015; Ye *et al.*, 2015]. The NYSE database encapsulates 347 stocks and their associated daily prices over 6004 trading days from January 1986 to February 2011, i.e., the market system has 347 co-evolving time series in terms of the daily stock prices. The prices are all corrected from the Yahoo financial dataset (http://finance.yahoo.com). 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 temporal window contains a time series of the daily return stock prices over a period of 28 days. To represent trades between different stocks as a network, for each window we compute the Euclidean distance between the time series of each pair of stocks as their connection (edge) weight, following the same setting in [Bai *et al.*, 2020]. It has been empirically shown that the financial networks associated with the Euclidean distance are more effective than those associated with the Pearson correlation. Clearly, this operation yields a time-varying financial network with a fixed number of 347 vertices and varying edge weights for each of the 5976 trading days. Each network is a complete weighted graph.

4.1 Kernel Embeddings from kPCA

We evaluate the performance of the proposed Quantum-inspired Entropic Kernel (QEK) on time-varying networks of the NYSE dataset. Specifically, we analyze whether the proposed QEK kernel can distinguish the structural changes of the network evolution with time. Furthermore, we also compare the proposed QEK kernel with three state-of-the-art kernel methods, that is, the dynamic time warping inspired Global Alignment Kernel (GAK) for original vectorial time series [Cuturi, 2011] and two graph kernels for time-varying financial networks. The graph kernels for comparisons include the Weisfeiler-Lehman Subtree Kernel (WLSK) [Shervashidze *et al.*, 2009], and the Discrete-time Quantum Walk Kernel (DTQK) [Bai *et al.*, 2020]. For the GAK kernel, we also utilize a time window of 28 days for each trading day. For the WLSK kernel, since it can only accommodate undirected and unweighted graphs, we transform each original network into a minimum spanning tree and ignore the weights on the preserved edges, following the same setting in the work [Bai *et al.*, 2020]. Since the DTQK kernel can accommodate edge weights, we straightforwardly perform this kernel on the original financial networks. We perform kernel Principle Component Analysis (kPCA) [Witten *et al.*, 2011] on the kernel matrices associated with different kernels, and



(a) Enron Incident for QEK

(b) Enron Incident for GAK

(c) Enron Incident for WLSK

(d) Enron Incident for DTQK

Figure 3: The 3D kPCA Embeddings of Different Kernels for Enron Incident.

embed the financial networks or the original time series into a vectorial pattern space. We visualize the embedding results using the first three principal components in Fig.1(a), Fig.1(b), Fig.1(c), and Fig.1(d) respectively.

Fig.1 exhibits the paths of the time-varying financial networks (or the original vectorial time series) in different kernel spaces, and the color bar of each subfigure indicates the date in the time series. We observe that the embeddings from the proposed QEK kernel exhibit a better manifold structure. Moreover, only the proposed QEK kernel generates a clear time-varying trajectory and the neighboring networks with time are close together in the embedding principal space. By contrast, the alternative methods hardly result in a trajectory and their embeddings tend to distribute as clusters. To further demonstrate the effectiveness of the QEK kernel, we compare the distance stress (DS) of the network embeddings from different kernels. Specifically, the DS is defined as

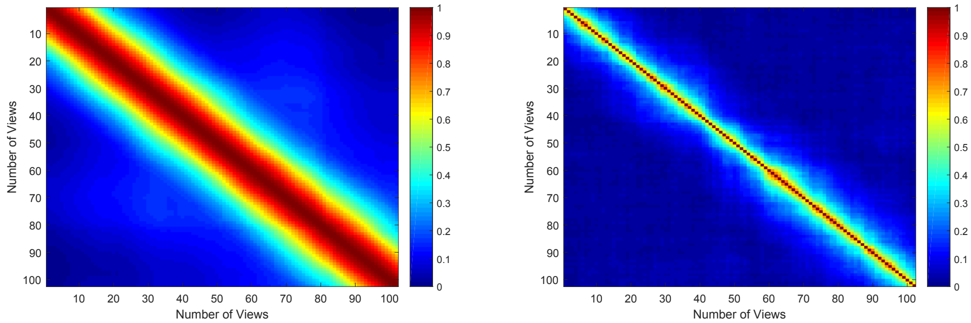
$$\mathrm{DS} = \frac{\sum_t \parallel x_t - x_{t-1} \parallel^2}{\sum_t \parallel x_t - x_{\mathrm{tn}} \parallel^2}, \quad (12)$$

where $t = 2, 3, \ldots, n$, x_t is the network embedding vector at time t, and x_{tn} is the nearest network embedding vector of x_t in the pattern space. For each embedding vector x_t at time t, if the nearest embedding vector is always the embedding vector at last time step (i.e., x_{t-1}), the value of DS will be 1. In other words, the DS value nearer to 1 indicates the better performance of the embeddings to form a clear time-varying trajectory. The DS value of each kernel is shown in Table 1. Clearly, only the DS value of the proposed QEK kernel is nearer to 1, indicating the better performance of preserving the ordinal arrangement of the time-varying networks.

Methods	**QEK**	GAK	WLSK	DTQK
Distance Stress	**1.0992**	2.9677	5.7053	4.7174

Table 1: The Distance Stress of the Network Embeddings

To take our study one step further, we explore the embeddings during different periods of three well-known financial events, i.e., the Black Monday period (***from 15th Jun 1987 to 17th Feb 1988***), the Dot-com Bubble period (***from 3rd Jan 1995 to 31st Dec 2001***), and the Enron Incident period (***from 16th Oct 2001 to 11th Mar 2002***). For differen kernels, Fig.2 corresponds to the Dot-com Bubble period and Fig.3 to the Enron Incident period. Due to the limit space, we do not exhibit the embeddings for Black Monday. However, we will observe the similar phenomenon with Fig.2 and Fig.3. These figures indicate that the Black Monday (***17th Oct, 1987***), the Dot-com Bubble Burst (***13rd Mar, 2000***) and the Enron Incident period are all crucial financial events, significantly influencing the structural time-varying evolution of the financial networks or the original vectorial financial time series. Excluding the GAK kernel, the embedding points of



(a) For QEK on Dot-com (b) For DTQK on Dot-com

Figure 4: Kernel Matrix Visualizations.

Methods	**QEK**	GAK	WLSK	DTQK
Black Monday	**0.9050**	0.6673	0.5667	0.6506
Dotcom Bubble	0.6473	**0.8882**	0.5279	0.7903
Enron Incident	**0.8504**	0.4992	0.5001	0.7042
Average Rand	**0.8009**	0.6849	0.5315	0.7150

Table 2: The Rand Index for K-means Clustering on the Embedding Points of 100 Trading Days around Each Financial Crisis.

the remaining kernels before and after these events are well separated into distinct clusters, and the points corresponding to the crucial events are midway between the clusters.

To place our analysis of the kernel embedding clusters on a more quantitative footing, for each kernel we select the kernel embedding points of **100** trading days around each financial crisis, i.e., we select embedding points of 50 trading days before and after each crisis date respectively. We apply the K-means method to the kernel embeddings of 100 trading days for each kernel to explore whether the clusters can be correctly separated in terms of the trading days before and after each financial crisis. We calculate the Rand Index for the resulting clusters and the Rand indicating each kernel is listed in Table 2. The results indicate that the embedding points associated with the proposed QEK kernel can produce the best clusters, i.e., the embedding points before and after the financial crisis are separated better than other kernels.

4.2 Evaluations of the Kernel Matrix

Based on the earlier evaluation, we find that the DTQK kernel is the most competitive kernel with the proposed QEK kernel. To further reveal the effectiveness of the proposed QEK kernel, we visualize the kernel matrices of both the kernels.

Due to the limited space, we only compute the kernel matrices between the networks belonging to the Dot-com Bubble period, and the period encapsulate 100 trading days. In fact, we will observe similar phenomenons if we compute the kernel matrices for other financial event periods. Specifically, the kernel matrices are visualized in Fig.4, where both the x-axis and y-axis represent the time steps. Note that, to compare the two kernels in the same scaled Hilbert space, we consider the normalized version of both the kernels as

$$k_n(G_p, G_q) = \frac{k(G_p, G_q)}{\sqrt{k(G_p, G_p)k(G_q, G_q)}}, \quad (13)$$

where k_n is the normalized kernel, and k is either the EDTWK or the WLSK kernel. As a result, the kernel values are all bounded between 0 to 1, and the colour bar beside each subfigure indicates the kernel value of the kernel matrix. Fig.4 indicates that the kernel values tend to decrease when the elements of the kernel matrix are far away from the matrix trace. This is because such elements are computed between time-varying networks having long time spans and there are more structure changes when the network evolves with a long time variation. Thus, both the QEK and DTQK kernels reflect structural evolutions of financial networks with time. However, on the other hand, the kernel value of the DTQK kernel tends to drop down more quickly when the element is a little far from the trace. By contrast, the kernel value of the QEK kernel tend to decrease more slowly when the element gets farer away from the trace. This observation explains why only the proposed QEK kernel can form a clear trajectory with time variation and generate better clusters before and after financial crisis, i.e., the proposed QEK kernel can better distinguish and understand the structural changes of the network structures evolving with a long time period.

5 Conclusion

In this paper, we have developed a new Quantum-inspired Entropic Kernel for time-varying complex networks. The proposed kernel bridges the gap between graph kernels and the classical dynamic time warping framework for time series analysis. Experimental analysis of NYSE financial time series demonstrates the effectiveness of the new kernel.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant no. 61976235 and 61602535), the Program for Innovation Research in Central University of Finance and Economics, the Youth Talent Development Support Program by Central University of Finance and Economics (No. QYP1908) and the Foundation of State Key Laboratory of Cognitive Intelligence (Grant No. COGOSC-20190002), iFLYTEK, China. Corresponding Author: Lixin Cui (cuilixin@cufe.edu.cn). The fourth author Yuhang Jiao mainly participated partial discussions for this work.

References

[Bai and Hancock, 2014] Lu Bai and Edwin R. Hancock. Depth-based complexity traces of graphs. *Pattern Recognition*, 47(3):1172–1186, 2014.

[Bai *et al.*, 2014] Lu Bai, Luca Rossi, Horst Bunke, and Edwin R. Hancock. Attributed graph kernels using the jensen-tsallis q-differences. In *Proceedings of ECML-PKDD*, pages 99–114, 2014.

[Bai *et al.*, 2015a] Lu Bai, Luca Rossi, Andrea Torsello, and Edwin R. Hancock. A quantum jensen-shannon graph kernel for unattributed graphs. *Pattern Recognition*, 48(2):344–355, 2015.

[Bai *et al.*, 2015b] Lu Bai, Zhihong Zhang, Chaoyan Wang, Xiao Bai, and Edwin R. Hancock. A graph kernel based on the jensen-shannon representation alignment. In *Proceedings of IJCAI*, pages 3322–3328, 2015.

[Bai *et al.*, 2016] Lu Bai, Luca Rossi, Lixin Cui, and Edwin R. Hancock. A novel entropy-based graph signature from the average mixing matrix. In *Proceedings of ICPR*, pages 1339–1344, 2016.

[Bai *et al.*, 2018] Lu Bai, Lixin Cui, Luca Rossi, Lixiang Xu, Xiao Bai, and Edwin Hancock. Local-global nested graph kernels using nested complexity traces. *Pattern Recognition Letters*, 2018.

[Bai *et al.*, 2020] Lu Bai, Luca Rossi, Lixin Cui, Jian Cheng, and Edwin R. Hancock. A quantum-inspired similarity measure for the analysis of complete weighted graphs. *IEEE Trans. Cybern.*, 50(3):1264–1277, 2020.

[Borgwardt and Kriegel, 2005] Karsten M. Borgwardt and Hans-Peter Kriegel. Shortest-path kernels on graphs. In *Proceedings of the IEEE International Conference on Data Mining*, pages 74–81, 2005.

[Bullmore and Sporns, 2009] Ed Bullmore and Olaf Sporns. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[Cui *et al.*, 2018] Lixin Cui, Lu Bai, Luca Rossi, Zhihong Zhang, Yuhang Jiao, and Edwin R. Hancock. A preliminary survey of analyzing dynamic time-varying financial networks using graph kernels. In *Proceedings of S+SSPR*, pages 237–247, 2018.

[Cuturi, 2011] Marco Cuturi. Fast global alignment kernels. In *Proceedings of ICML*, pages 929–936, 2011.

[Farhi and Gutmann, 1998] E. Farhi and S. Gutmann. Quantum computation and decision trees. *Physical Review A*, 58:915, 1998.

[Godsil, 2013] Chris Godsil. Average mixing of continuous quantum walks. *Journal of Combinatorial Theory, Series A*, 120(7):1649–1662, 2013.

[Haussler, 1999] David Haussler. Convolution kernels on discrete structures. In *Technical Report UCS-CRL-99-10*, Santa Cruz, CA, USA, 1999.

[Jain, 2019] Brijnesh J. Jain. Making the dynamic time warping distance warping-invariant. *Pattern Recognition*, 94:35–52, 2019.

[Johansson *et al.*, 2014] Fredrik D. Johansson, Vinay Jethava, Devdatt P. Dubhashi, and Chiranjib Bhattacharyya. Global graph kernels using geometric embeddings. In *Proceedings of ICML*, pages 694–702, 2014.

[Kashima *et al.*, 2003] Hisashi Kashima, Koji Tsuda, and Akihiro Inokuchi. Marginalized kernels between labeled graphs. In *Proceedings of ICML*, pages 321–328, 2003.

[Mikalsen *et al.*, 2018] Karl Øyvind Mikalsen, Filippo Maria Bianchi, Cristina Soguero-Ruíz, and Robert Jenssen. Time series cluster kernel for learning similarities between multivariate time series with missing data. *Pattern Recognition*, 76:569–581, 2018.

[Nicolis *et al.*, 2005] G. Nicolis, A. G. Cantu, and C. Nicolis. Dynamical aspects of interaction networks. *International Journal of Bifurcation and Chaos*, 15:3467, 2005.

[Shervashidze *et al.*, 2009] N. Shervashidze, S.V.N. Vishwanathan, K. Mehlhorn T. Petri, and K. M. Borgwardt. Efficient graphlet kernels for large graph comparison. *Journal of Machine Learning Research*, 5:488–495, 2009.

[Shimada *et al.*, 2008] Y. Shimada, T. Kimura, and T. Ikeguchi. Analysis of chaotic dynamics using measures of the complex network theory. In *Proceedings of ICANN*, pages 61–70, 2008.

[Silva *et al.*, 2015] Filipi N. Silva, Cesar H. Comin, Thomas K. Peron, Francisco A. Rodrigues, Cheng Ye, Richard C. Wilson, Edwin R. Hancock, and Luciano da F. Costa. Modular dynamics of financial market networks. *arXiv preprint arXiv:1501.05040*, 2015.

[Witten *et al.*, 2011] Ian H. Witten, Eibe Frank, and Mark A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2011.

[Xu *et al.*, 2018] Lixiang Xu, Xiaoyi Jiang, Lu Bai, Jin Xiao, and Bin Luo. A hybrid reproducing graph kernel based on information entropy. *Pattern Recognition*, 73:89–98, 2018.

[Ye *et al.*, 2015] Cheng Ye, César H. Comin, Thomas K. Peron, Filipi N. Silva, Francisco A. Rodrigues, Luciano da F. Costa, Andrea Torsello, and Edwin R. Hancock. Thermodynamic characterization of networks using graph polynomials. *Physical Review E*, 92(3):032810, 2015.

[Zhang and Small, 2006] J. Zhang and M. Small. Complex network from pseudoperiodic time series: Topology versus dynamics. *Physical Review Letters*, 96:238701, 2006.