QBMK: Quantum-based Matching Kernels for Un-attributed Graphs

Lu Bai [1] **Lixin Cui** [2] **Ming Li** [3 4] **Yue Wang** [2] **Edwin R. Hancock** [5]

Abstract

In this work, we develop a new Quantum-based Matching Kernel (QBMK) for un-attributed graphs, by computing the kernel-based similarity between the quantum Shannon entropies of aligned vertices through the Continuous-time Quantum Walk (CTQW). The theoretical analysis reveals that the proposed QBMK kernel not only addresses the shortcoming of neglecting the structural correspondence information between graphs arising in existing R-convolution graph kernels, but also overcomes the problem of neglecting the structural differences between pairs of aligned vertices arising in existing vertex-based matching kernels. Moreover, the proposed QBMK kernel can simultaneously capture both global and local structural characteristics through the quantum Shannon entropies. Experimental evaluations on standard graph datasets demonstrate that the proposed QBMK kernel is able to outperform state-of-the-art graph kernels and graph deep learning approaches.

1. Introduction

There are increasing interests in employing graph kernels associated with specified kernel machines (e.g., the C-Support Vector Machine (C-SVM), etc) for graph classification (Gärtner et al., 2003; Jebara et al., 2004).

Generally speaking, graph kernels are defined as the similarity measure between graph structures. Most state-of-the-art graph kernels essentially fall into the concept of R-convolution (Haussler, 1999). This is a generic manner of defining new graph kernels based on measuring the similarity between decomposed substructures of graphs, e.g., the R-convolution graph kernels based on the decomposed paths (Borgwardt & Kriegel, 2005), cycles (Aziz et al., 2013), walks (Sugiyama & Borgwardt, 2015), subgraphs (Kriege & Mutzel, 2012), subtrees (Azaïs & Ingels, 2020), etc. With the scenario of R-convolution, Kalofolias et al., (Kalofolias et al., 2021) have introduced a Structural Similarity Random Walk Kernel by measuring the similarity between random walks. Costa and Grave (Costa & Grave, 2010) have introduced a Pairwise-Distance Expansion Subgraph Kernel based on layer-wise expansion subgraphs around pairwise rooted vertices, that have specified distance between each other. Shervashidze et al., (Shervashidze et al., 2010) have introduced a Weisfeiler-Lehman Subtree Kernel based on the subtree invariants, that are extracted by the classical Weisfeiler-Lehman Isomorphism Test Method (Weisfeiler & Lehman, 1968). Other classical R-convolution graph kernels also include: the Wasserstein Weisfeiler-Lehman Subtree Kernel (Togninalli et al., 2019), the Pyramid-Quantized Shortest-Path Kernel (Gkirtzou & Blaschko, 2016), the Isolation Graph Kernel (Xu et al., 2021a), the Graph Filtration Kernel (Schulz et al., 2022), etc.

Unfortunately, most of the above R-convolution graph kernels suffer from some common drawbacks, influencing the effectiveness. First, to guarantee the computational efficiency, the R-convolution graph kernels tend to utilize the local substructures of small sizes. Thus, the R-convolution graph kernels fail to capture characteristics of the global graph structures. Second, the R-convolution graph kernels only focus on whether a pair of substructures are isomorphic, ignoring the structural correspondence information between the substructures. For example, Fig. 1 exhibits two Delaunay graphs, that are both extracted from the same house object through different viewpoints. Since the two triangle-based substructures are isomorphic, the R-convolution graph kernels will directly add one unit kernel value, no matter whether the substructures are structurally aligned to each other in terms of the vision background. As a result, the R-convolution graph kernels can not reflect precise similarity measures between graphs.

To address the drawback of ignoring global structure in-

[1]School of Artificial Intelligence, Beijing Normal University, Beijing 100875, China. [2]School of Information, Central University of Finance and Economics, Beijing 100081, China. [3]Zhejiang Institute of Optoelectronics, Jinhua 321004, China. [4]Zhejiang Key Laboratory of Intelligent Education Technology and Application, Zhejiang Normal University, Jinhua 321004, China. [5]Department of Computer Science, University of York, YO10 5GH York, UK.. Correspondence to: Lixin Cui <cuilixin@cufe.edu.cn>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

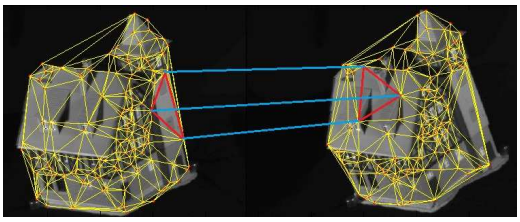


Figure 1. Delaunay graphs extracted with different viewpoints.

formation that arising in R-convolution graph kernels, a number of global-based graph kernels based on the global graph structures have been developed. For instance, Johansson et al., (Johansson et al., 2014) have introduced a Global Geometric Embedding Kernel associated with the Lovász numbers as well as their orthonormal representations, that are computed through the adjacency matrix. Xu et al., (Xu et al., 2018) have introduced a Global-based Reproducing Graph Kernel by measuring the Basic Reproducing Kernel between the approximated von Neumann entropies of global graph structures. On the other hand, Bai et al., (Bai et al., 2015a) have developed a family of quantum-inspired global-based graph kernels, namely the Quantum Jensen-Shannon Graph Kernels, by measuring the Quantum Jensen-Shannon Divergence (QJSD) between the density operators of the Continuous-Time Quantum Walk (CTQW) (Emms et al., 2009) evolved on global graph structures. Since the CTQW can not only better discriminate different graphs, but also reflect intrinsic structural characteristics. The resulting Quantum Jesen-Shannon Graph Kernels can capture more complicated intrinsic structure information residing on global graphs. Unfortunately, these global-based graph kernels focus more on the global characteristics, and ignore the inherent characteristics residing on local structures. Furthermore, similar to the R-convolution kernels, these global-based graph kernels cannot identify the structural correspondence information between pairwise graphs, influencing the effectiveness.

To address the drawback of overlooking the correspondence information that arises in both the R-convolution graph kernels and the global-based graph kernels, Bai and Xu et al., (Bai et al., 2015b; Xu et al., 2021b) have developed a family of vertex-based matching kernels by counting the pairs of aligned vertices, that are identified by evaluating the distance between the vectorial Depth-based Representations (Bai & Hancock, 2014) of vertices. Moreover, they show that these matching kernels are theoretically equivalent to Aligned Subgraph Kernels, that calculate the number of pairwise structurally-aligned subgraphs, encapsulating the structural correspondence information between graphs. However, similar to the R-convolution kernels, these matching kernels cannot capture global structure information through the local aligned vertices, influencing the effectiveness of the kernel measures. Furthermore, since each pair of aligned vertices will dedicate the same one unit kernel value, the matching kernels cannot identify structural differences between different pairs of aligned vertices in terms of global graph structures, influencing the precise kernel measure. Finally, these matching kernels only identify the correspondence information based on specified dimensional vectorial vertex representations, lacking multilevel alignment information between graphs.

The main objective of this work is to resolve the theoretical problems of the existing graph kernels discussed above. To this end, we develop a novel Quantum-based Matching Kernel (QBMK) for un-attributed graphs. One key innovation of the proposed QBMK kernel is to compute the kernel value by measuring the similarity between the quantum Shannon entropies of aligned vertices, through the Average Mixing Matrix (AMM) of the CTQW. Since these vertex-based entropies are usually different, the QBMK kernel can discriminate the structural difference between different pairs of aligned vertices based on the entropic similarity. Overall, the contributions of this work are summarized as follows.

First, for each pair of graphs, we commence by performing the CTQW on each of the graph structures. More specifically, we utilize the AMM matrix to describe the evolution of the CTQW. The reason of using the AMM matrix is due to the fact that it can not only reflect richer graph topological information in terms of the CTQW, but also assign each vertex an individual probability distribution that investigates how the CTQW visit all vertices when it departs from the vertex (see details in Section 3.1). Furthermore, we show how the AMM matrix allows us to compute a quantum Shannon entropy for each vertex, that not only reflects local structural information residing on the local vertex, but also captures complicated intrinsic structural characteristics of global graph structures.

Second, for pairwise graphs we compute a family of different h-level entropic correspondence matrices to reflect the multilevel alignment information between vertices. With the entropic correspondence matrices and the quantum Shannon entropies of vertices to hand, the proposed QBMK kernel is defined by computing the kernel-based similarities between the entropies over all pairs of aligned vertices. We theoretically demonstrate that the QBMK kernel not only encapsulates the structural correspondence information between graphs, but also identifies the structural differences between pairs of aligned vertices. Moreover, the QBMK kernel not only reflects multilevel alignment information through the entropic correspondence matrices, but also simultaneously captures both global and local characteristics through the quantum Shannon entropies. These advantages explain the effectiveness of the QBMK kernel.

Third, we experimentally demonstrate the graph classification performance of the proposed QBMK kernel asso-

ciated with the C-SVM on ten standard graph datasets. The QBMK kernel significantly outperforms state-of-the-art graph kernels and graph deep learning approaches.

This paper organizes as follows. Section 2 reviews related works. Section 3 defines the proposed kernel. Section 4 gives empirical evaluations. Section 5 gives conclusions.

2. Quantum Backgrounds and Related Works

In this section, we introduce the concept of the CTQW that will be used in this work. Moreover, we review the classical Depth-based Matching Kernel that is related to this work.

2.1. The Continuous-time Quantum Walks

One objective of this work is to explore the structural information and compute the entropy-based representations of graphs through the CTQW (Emms et al., 2009), that is the quantum analogy of the classical Continuous-time Random Walk (CTRW) (Watrous, 2001). Unlike the classical CTRW, the state vector of the CTQW is complex-valued rather than real-valued, and its evolution is governed by an unitary matrix rather than a stochastic matrix. Thus, the CTQW evolution is reversible, reducing the tottering problem arising in the CTRW. Moreover, since the evolution of the CTQW is not determined by the Laplacian spectrum with low frequency components, the CTQW can better discriminate different graph structures than the CTRW.

In this subsection, we introduce the concept of the C-TQW (Emms et al., 2009). For the CTQW evolving on a sample graph $G(V,E)$, its state space is defined over the vertex set V, and the corresponding basis state at each vertex $v \in V$ is defined as $|v\rangle$ based on the Dirac notation. Here, $|.\rangle$ is a $|V|$-dimensional orthonormal vector in a complex valued Hilbert space. Since the state $|\psi(t)\rangle$ at time t is defined as a complex linear combination of these basic state vectors $|v\rangle$, $|\psi(t)\rangle$ can be written as

$$|\psi(t)\rangle = \sum_{v\in V} \alpha_v(t)\,|v\rangle\,, \tag{1}$$

where $\alpha_v(t) \in \mathbb{C}$ is the complex amplitude. Unlike the classical counterpart, the evolution of the CTQW is based on the Schrödinger equation, i.e.,

$$\frac{\partial\,|\psi_t\rangle}{\partial t} = -i\bar{\mathcal{H}}\,|\psi_t\rangle\,, \tag{2}$$

where $\mathcal{H}$ represents the system Hamiltonian and accounts for the total energy of the system, and one can adopt the adjacency matrix as the Hamiltonian.

Remarks: In quantum mechanics, the state $|\psi(t)\rangle$ is the pure state. To further explore intrinsic complex structural information of a graph, a statistical ensemble (i.e., the mixed density matrix) of the pure states is computed as $\rho = \sum_t p_t\,|\psi_t\rangle\,\langle\psi_t|$. Specifically, the density matrix can be viewed as the structural matrix representation of the global graph strcture. For a pair of graphs, by measuring the QJSD between their density matrices, Bai and Rossi et al., (Bai et al., 2015a; Rossi et al., 2015) have developed a family of Quantum Jensen-Shannon Kernels (QJSK) that can reflect the intrinsic complex structural information for graphs at a global level in terms of the CTQW. However, these QJSK kernels focus more on global structures, ignoring the structural information on local structures.

2.2. The Classical Depth-based Matching Kernels

In this subsection, we review the definition of the classical Depth-based Matching Kernel (DBMK) (Bai et al., 2014a), that computes the kernel value by calculating the number of pairwise aligned vertices between graphs. For each pair of graphs $G_p(V_p,E_p)$ and $G_q(V_q,E_q)$, assume $\mathbf{R}^h(v_i)$ and $\mathbf{R}^h(v_j)$ are the h-dimensional vectorial representations of their vertices $v_i \in V_p$ and $v_j \in V_q$. We first compute the Euclidean distance as the affinity measure between v_i and v_j, and the (i,j)-th entry $R(i,j)$ of the resulting affinity matrix $R \in \mathcal{R}^{|V_p||V_q|}$ between G_p and G_q is

$$R(i,j) = \parallel \mathbf{R}^h(v_i) - \mathbf{R}^h(v_j) \parallel . \tag{3}$$

If $R(i,j)$ is the smallest one in both i-th row and j-th column, there is a one-to-one correspondence between $v_i \in V_p$ and $v_j \in V_q$, i.e., v_i and v_j are aligned. Specifically, we record the alignment information between G_p and G_q in the correspondence matrix $C \in \{0,1\}^{|V_p||V_q|}$ that satisfies

$$C(i,j) = \begin{cases} 1 & \text{if } R(i,j) \text{ is the smallest in both} \\ & i\text{−th row and } j\text{−th column;} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Eq.(4) indicates that the vertices $v_i \in V_p$ and $v_j \in V_q$ are aligned if $C(i,j) = 1$. To construct effective alignment information between G_p and G_q, Bai et al., (Bai et al., 2014a) suggest to employ Depth-based (DB) Representations as the h-dimensional vectorial vertex representations. Specifically, the DB representations are defined by gauging the entropy-based complexity on a family of $\bar{h}$-layer expansion subgraphs rooted at vertices (Bai & Hancock, 2014), where $\bar{h}$ varies from 1 to h (i.e., $\bar{h} = 1, 2, \ldots, h$) and the larger layer subgraph completely encapsulates the smaller layer one. Thus, the DB representation significantly encapsulates rich entropic content flow from each local vertex to the global graph structure, as a function of depth. Fig.2 shows the computation procedure of the DB representation. Finally, note that, the vertex of a graph may be aligned to two or more vertices of the other graph, and one can randomly preserve one correspondence entry (Bai et al., 2014a).

For the pair of graphs G_p and G_q and their associated correspondence matrix C defined by Eq.4, the DBMK kernel

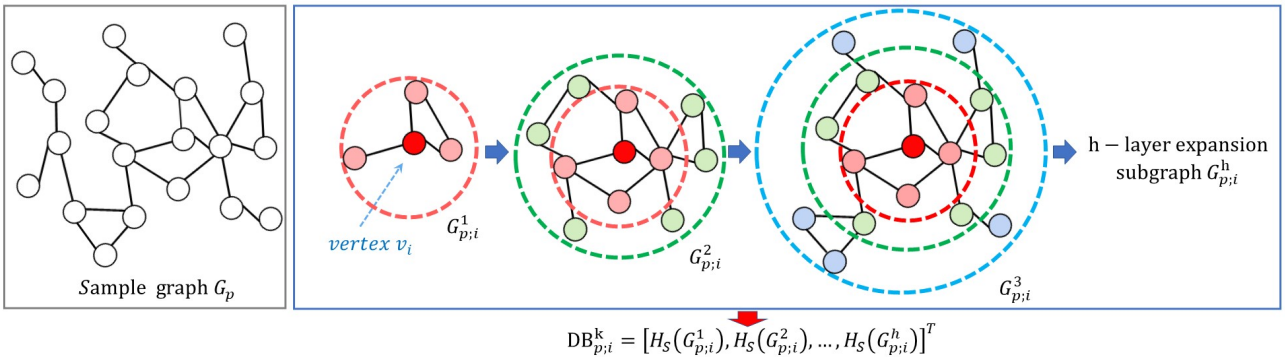


Figure 2. The procedure of constructing the DB representations. For a sample graph $G_p(V_p, E_p)$ and the associated i-th vertex $v_i \in V_p$ (marked with the color red), we commence by computing the $\bar{h}$-th order neighborhood set as $\mathcal{N}^{\bar{h}}_{p;i} = \{v_j \in V_p \mid d(v_i, v_j) \leq 1\}$, where $d(v_i, v_j)$ represents the length of the shortest path between the j-th vertex $v_j \in V_p$ and the i-th vertex $v_i \in V_p$. Then, the $\bar{h}$-layer expansion subgraph $H_S(G^{\bar{h}}_{p;i})$ rooted at $v_i \in V_p$ is defined as the substructure associated with the vertices in $\mathcal{N}^{\bar{h}}_i$ and their original topological structures in G_p, e.g., the 1-layer, 2-layer and 3-layer expansion subgraphs $G^1_{p;i}$, $G^2_{p;i}$ and $G^3_{p;i}$ surrounded by the red, green and blue broken line, respectively. Clearly, we can construct a family of $\bar{h}$-layer expansion subgraphs, if we vary $\bar{h}$ from 1 to h (i.e., $\bar{h} = 1, 2, \ldots, h$). The resulting h-dimensional DB representation of v_i is defined as $\mathrm{DB}^h_{p;i} = [H_S(G^1_{p;i}), H_S(G^2_{p;i}), \cdots, H_S(G^h_{p;i})]^T$, where $H_S(\cdot)$ is the Shannon entropy based on Steady State Random Walks (Bai & Hancock, 2014).

between G_p and G_q is formulated as

$$K_{\mathrm{DBMK}}(G_p, G_q) = \sum_{i=1}^{|V_p|} \sum_{j=1}^{|V_q|} C(i,j), \tag{5}$$

that counts the number of pairwise aligned vertices.

Remarks: Indeed, the DBMK kernel is theoretically related to a classical R-convolution kernel, namely the All Subgraph Kernel (ASK) (Gärtner et al.), explaining the effectiveness of the DBMK kernel. Specifically, for the pair of graphs G_p and G_q, the ASK kernel is defined as

$$K_{\mathrm{ASK}}(G_p, G_q) = \sum_{S_p \sqsubseteq G_p} \sum_{S_q \sqsubseteq G_q} \delta(S_p, S_q), \tag{6}$$

where S_p and S_q are pairwise subgraphs of G_p and G_q, and

$$\delta(S_p, S_q) = \begin{cases} 1 & \text{if } S_p \simeq S_q, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Here, δ is a Dirac kernel, and $\delta(S_p, S_q) = 1$ if S_p and S_q are isomorphic (i.e., $S_p \simeq S_q$), and 0 otherwise. Since the h-dimensional DB representation $\mathrm{DB}^h_{p;i}$ of vertex $v_i \in V_p$ can be theoretically considered as the vectorial representation of the associated h-layer expansion subgraph $G^h_{p;i}$ rooted at $v_i \in V_p$ (Bai et al., 2014a). Moreover, Eq.(3) and Eq.(4) indicate that $\mathrm{DB}^h_{p;i}$ and $\mathrm{DB}^h_{q;i}$ are closest to each other (i.e., $\mathrm{DB}^h_{p;i}$ and $\mathrm{DB}^h_{q;i}$ are structurally similar), if the the vertices $v_i \in V_p$ and $v_j \in V_q$ are aligned to each other. Hence, the DBMK kernel can be rewritten as

$$K^h_{\mathrm{DBMK}}(G_p, G_q) = \sum_{S_p \sqsubseteq G_p} \sum_{S_q \sqsubseteq G_q} \delta(S_p, S_q), \tag{8}$$

where

$$\delta(S_p, S_q) = \begin{cases} 1 & \text{if } S_p = G^h_{p;i} \text{ and } S_q = G^h_{q;j}, \\ & \text{and } v \text{ and } u \text{ are aligned,} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

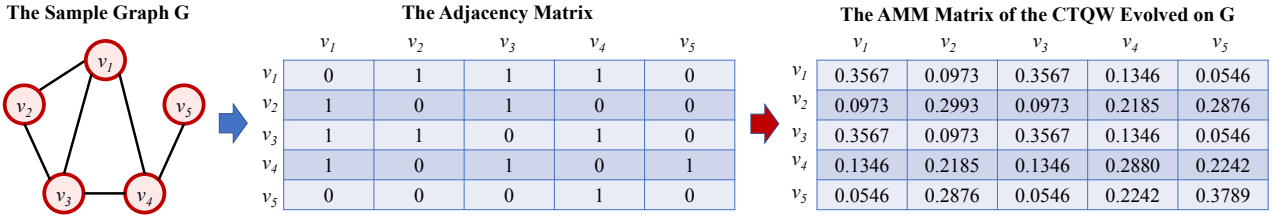
Clearly, the DBMK kernel can be theoretically considered as an Aligned Subgraph Kernel that calculates the number of pairwise isomorphic h-layer expansion subgraphs around aligned vertices, i.e., the DBMK kernel integrates the local correspondence information between the isomorphic subgraphs, addressing the shortcoming of classical R-convolution kernels that tend to overlook the structural correspondence information between substructures. Unfortunately, the DBMK kernel still suffers from some drawbacks. First, similar to the R-convolution kernels, the DBMK kernel cannot reflect global graph characteristics based on local aligned vertices. Second, the DBMK kernel cannot discriminate the structural differences between different pairs of aligned vertices in terms of the global graph structures, since any pair of aligned vertices will contribute the same kernel-based similarity value (i.e., the one unit value).

3. The Quantum-based Matching Kernel

In this section, we propose the Quantum-based Matching Kernel (QBMK) for un-attributed graphs. Moreover, we theoretically explain how the proposed kernel overcomes the drawbacks of aforementioned classical graph kernels.

3.1. Quantum Shannon Entropies through the CTQW

In this subsection, we define the quantum Shannon entropy based on the CTQW. We employ the AMM matrix (Godsil,



	v_1	v_2	v_3	v_4	v_5
v_1	0	1	1	1	0
v_2	1	0	1	0	0
v_3	1	1	0	1	0
v_4	1	0	1	0	1
v_5	0	0	0	1	0

	v_1	v_2	v_3	v_4	v_5
v_1	0.3567	0.0973	0.3567	0.1346	0.0546
v_2	0.0973	0.2993	0.0973	0.2185	0.2876
v_3	0.3567	0.0973	0.3567	0.1346	0.0546
v_4	0.1346	0.2185	0.1346	0.2880	0.2242
v_5	0.0546	0.2876	0.0546	0.2242	0.3789

Figure 3. An example of computing the AMM matrix.

2013) to describe the behavior of the CTQW evolving on the graph $G(V,E)$ at time t, and the matrix is defined as

$$Q_{\mathrm{M}}(t) = U(t) \circ U(-t) = e^{i\bar{\mathcal{H}}t} \circ e^{-i\bar{\mathcal{H}}t}, \tag{10}$$

where $\circ$ represents the Schur-Hadamard product operation between $e^{i\bar{\mathcal{H}}t}$ and $e^{-i\bar{\mathcal{H}}t}$. Since U is an unitary matrix, $Q_{\mathrm{M}}(t)$ denotes a doubly stochastic matrix and each of its entry $Q_{\mathrm{M}}(t)_{uv}$ corresponds to the probability of the CTQW visiting vertex $v \in V$ at time t and departing from vertex $u \in V$. To ensure the convergence for $Q_{\mathrm{M}}(t)$, we compute the time-averaged AMM matrix Q for the CTQW by taking the Cesàro mean, i.e.,

$$Q = \lim_{T\to\infty} \int_0^T Q_M(t)dt, \tag{11}$$

where each entry Q_{vu} of Q corresponds to the time-averaged probability of the CTQW arriving in $u \in V$ and departing from $v \in V$. Because each entry of Q is a rational number (Godsil, 2013), we can directly calculate Q based on the Hamiltonian spectrum. Assume the adjacency matrix A is the Hamiltonian $\bar{\mathcal{H}}$, $\lambda_1, \lambda_2, \ldots, \lambda_{|V|}$ be its eigenvalues, and $\mathbb{P}_j$ be its orthogonal projection on the eigenspace in association with λ_j (i.e., $\bar{\mathcal{H}} = \sum_{j=1}^{|V|} \lambda_j \mathbb{P}_j$), the AMM matrix Q of the CTQW is defined as

$$Q = \sum_{j=1}^{|V|} \mathbb{P}_j \circ \mathbb{P}_j, \tag{12}$$

where the v-th row or column of Q corresponds to a probability distribution of the CTQW visiting all vertices when it departs from the vertex $v \in V$.

Definition 3.1 (The Quantum Shannon Entropy of Vertices): With the AMM matrix Q of the graph $G(V,E)$ to hand, we define a quantum Shannon entropy $H_{\mathrm{QS}}(v)$ for each vertex $v \in V$ associated with the v-th row of Q, i.e.,

$$H_{\mathrm{Q}}(v) = -\sum_{u\in V} Q_{vu} \log Q_{vu}, \tag{13}$$

where $H_{\mathrm{Q}}(v)$ is the structural characteristics of each local vertex $v \in V$ in terms of the CTQW departing from v. □

Remarks: The proposed quantum Shannon entropy computed through the AMM matrix has some important theoretical properties. First, as we have discussed in Section 2.1, the CTQW can reflect the complicated intrinsic structural information of graph structures. Thus, the quantum Shannon entropy can well characterize the graph topological information through the AMM matrix of the CTQW. Second, since each v-th row of the AMM matrix corresponds to an individual probability distribution of the CTQW visiting all vertices when the CTQW departs from each vertex v. The quantum Shannon entropy computed through the AMM matrix not only encapsulates the global structure information through the CTQW probability distribution visiting all vertices, but also reflects the local structural information in terms of each local starting vertex. Fig.3 exhibits an example, it is clear that the probability distributions of the CTQW departing from different local vertices tend to be different, excluding the vertices v_1 and v_3 having the symmetric structural relationship. Thus, the quantum Shannon entropies of different vertices tend to be distinctive, simultaneously reflecting discriminative global and local information over different vertices. Overall, the quantum entropy provides an elegant way to develop novel entropic graph kernels, by measuring the similarity between the entropies.

3.2. The Proposed QBMK Kernel

In this subsection, we define the proposed QBMK kernel. To reflect multilevel vertex correspondence information between graphs, unlike the original DBMK kernel discussed in Section 2.2, we compute a family of different-level correspondence matrices rather than a single-level correspondence matrix between graphs. Specifically, for the pair of graphs $G_p(V_p,E_p)$ and $G_q(V_q,E_q)$, assume $\mathrm{DB}^h_{p;i}$ and $\mathrm{DB}^h_{q;j}$ are the associated h-dimensional DB representations of their vertices $v_i \in V_p$ and $v_j \in V_p$ (i.e., the vectorial vertex representations), respectively. The h-level affinity matrix R^h between G_p and G_q is defined as

$$R^h(i,j) = \| \mathrm{DB}^h_{p;i} - \mathrm{DB}^h_{q;j} \| . \tag{14}$$

To identify the vertex correspondence information through R^h, we not only request the entry $R^h(i,j)$ to be the smallest one in both row i and column j, but also request both

Table 1. Properties of the Proposed HAQJSK Kernels

Kernel Properties	**QBMK**	R-convolution Kernels	Matching-based Kernels	Global-based Kernels
Structural Alignment	**Yes**	No	**Yes**	No
Capture Local Information	**Yes**	**Yes**	**Yes**	No
Capture Global Information	**Yes**	No	No	**Yes**
Reflect Multilevel Alignments	**Yes**	No	No	No
Discriminate Different Alignments	**Yes**	No	No	No

the h-layer expansion subgraphs $G^{\bar{h}}_{p;i}$ and $G^{\bar{h}}_{q;j}$ for $\mathrm{DB}^h_{p;i}$ and $\mathrm{DB}^h_{q;j}$ to exist (i.e., $|\mathcal{N}^{\bar{h}}_{p;i}| > 0$ and $|\mathcal{N}^{\bar{h}}_{q;j}| > 0$). Thus, unlike the previous correspondence matrix C defined by Eq.(4), the h-level correspondence matrix $C^h \in \{0,1\}^{|V_p||V_q|}$ recording the vertex correspondence information between G_p and G_q based on R^h needs to satisfy

$$C^h(i,j) = \begin{cases} 1 & \text{if } R^h(i,j) \text{ is the smallest entry} \\ & \text{in both row } i \text{ and column } j, \text{ and} \\ & |\mathcal{N}^{\bar{h}}_{p;i}| > 0 \text{ and } |\mathcal{N}^{\bar{h}}_{q;j}| > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

When vary the parameter h from 1 to H (e.g., the greatest value of H), we can compute a family of different h-level correspondence matrices as $\mathbf{C} = \{C^1, \ldots, C^h, \ldots, C^H\}$, reflecting multilevel vertex correspondence information.

Definition 3.2 (The Quantum-based Matching Kernel): For the graphs G_p and G_q, and the associated h-level correspondence matrix set $\mathbf{C}$, the QBMK kernel is defined as

$$\mathrm{K}_{\mathrm{QBMK}}(G_p, G_q) = \sum_{h=1}^{H} \sum_{v_i \in V_p} \sum_{v_j \in V_q} C^h(i,j) \mathrm{K}_{\mathrm{r}}(v_i, v_j), \tag{16}$$

where $\mathrm{K}_{\mathrm{r}}(v_i, v_j)$ is defined as the Basic Reproducing Kernel (Xu et al., 2018) associated with the quantum Shannon entropies $H_{\mathrm{Q}}(v_i)$ and $H_{\mathrm{Q}}(v_j)$ of vertices $v_i \in V_p$ and $v_j \in V_p$ defined by Eq.(14), i.e.,

$$\mathrm{K}_{\mathrm{r}}(v_i, v_j) = \frac{1}{2} e^{-[H_{\mathrm{Q}}(v_i) - H_{\mathrm{Q}}(v_j)]}. \tag{17}$$

Clearly, the proposed QBMK kernel computes the sum of the reproducing kernel based similarities between the quantum Shannon entropies over all pairs of aligned vertices. □

Note that, Fröhlich et al., (Fröhlich et al., 2005) have pointed out that the transitivity between the aligned vertices is the necessary condition to guarantee the positive definiteness for a vertex-based matching kernel, i.e., if the vertices u and v are matched, the vertices v and w are matched, then the vertices u and w should also be matched. Since the required vertex correspondence information identified by the proposed QBMK kernel is evaluated between each individual pair of graphs, the proposed QBMK kernel cannot guarantee the transitive alignment and the positive definiteness. Nonetheless, Ong et al., (Ong et al., 2004) have proven that the non-positive definite kernels are still effective, and this is not necessarily a big issue for the use of these kernels in practice. In Section 4, we will indicate that the proposed QBMK kernel still significantly outperforms state-of-the-art methods, demonstrating the above theoretical statements in (Ong et al., 2004).

3.3. Theoretical Linkage to the QBMK Kernel

In this subsection, we theoretically indicates the advantages of the new QBMK kernel, by revealing the theoretical linkage to the classical DBMK kernel defined in Section 2.2. To this end, we commence by redefining the h-level correspondence matrix C^h described by Eq.(16) as a new h-level entropic correspondence matrix C^h_{E}, i.e.,

$$C^h_{\mathrm{E}}(i,j) = \begin{cases} \mathrm{K}_{\mathrm{r}}(v_i, v_j) & \text{if } R^h(i,j) \text{ is the smallest one} \\ & \text{both in row } i \text{ and in column } j, \\ & \text{and } |\mathcal{N}^{\bar{h}}_{p;i}| > 0 \text{ and } |\mathcal{N}^{\bar{h}}_{q;j}| > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

For the pair of graphs $G_p(V_p, E_p)$ and $G_q(V_q, E_q)$, we say that their vertices $v_i \in V_p$ and $v_j \in V_q$ are aligned to each other, if the entry $C^h_{\mathrm{E}}(i,j) > 0$. With the matrix C^h_{E} to hand, the QBMK kernel can be rewritten as

$$\mathrm{K}_{\mathrm{QBMK}}(G_p, G_q) = \sum_{h=1}^{H} \sum_{v_i \in V_p} \sum_{v_j \in V_q} C^h_{\mathrm{E}}(i,j). \tag{19}$$

As a result, the proposed QBMK kernel defined by Eq.(20) is similar with that of the classical DBMK kernel defined by Eq.(5), and can also be theoretically seen as a vertex-based matching kernel. Indeed, this theoretical linkage reveals the advantages of the proposed QBMK kernel, explaining the effectiveness. Specifically, the theoretical advantages are shown in Table 1 and briefly discussed as follows.

First, unlike the classical DBMK kernel that simply counts pairs of aligned vertices based on the correspondence matrix C defined by Eq.(4), the proposed QBMK kernel is able to discriminate the structural differences between different pairs aligned vertices through the entropic correspondence matrices C^h_{E}. This is because the associated quantum Shannon entropies of different vertices tend to be different (see details in Section 3.1). By contrast, for the DBMK kernel, $Eq.(4)$ indicates that any pair of aligned vertices will contribute the same one unit kernel value, ignoring the structural differences between the aligned vertices.

Table 2. Statistical Information of the Benchmark Datasets.

Datasets	MUTAG	D&D	PTC	PPIs	CATH2	BAR31	BSPHERE31	GEOD31	Shock	GatorBait
Max # vertices	28	5748	109	218	568	220	227	380	33	545
Min # vertices	10	30	2	3	143	41	43	29	4	239
Mean # vertices	17.93	284.4	25.60	109.63	308.03	95.43	99.83	57.42	13.16	348.70
# graphs	188	1178	344	219	190	300	300	300	150	100
# classes	2	2	2	5	2	15	15	15	10	30
Description	BIO	BIO	BIO	BIO	BIO	CV	CV	CV	CV	CV

Second, similar to the classical DBMK kernel, the proposed QBMK kernel can also be re-written as the similar manners of Eq.(8) and Eq.(9). In other words, the QBMK kernel can also be seen as the Aligned ASK kernel mentioned in Section 2.2, and is able to identify the isomorphism between pairs of h-layer expansion subgraphs $G^h_{p;i}$ and $G^h_{q;j}$ rooted at aligned vertices $v_i \in V_p$ and $v_i \in V_q$. However, unlike the classical DBMK kernel that is based on the single-level correspondence matrix C defined by Eq.(4) and can only identify the isomorphism between specified h-layer expansion subgraphs, the proposed QBMK kernel is based on the different h-level entropic correspondence matrices $C^h_{\rm E}$ defined by Eq.(19) and can identify the isomorphism between more different-level h-layer expansion subgraphs ($1 \leq h \leq H$). Thus, the proposed QBMK kernel not only addresses the problem of ignoring the correspondence information arising in aforementioned R-convolution kernels, but also reflects more multi-level structural information than the classical DBMK kernel.

Third, the proposed QBMK kernel is able to address the shortcoming of only focusing on local or global graph structure information that arises in existing R-convolution kernels, vertex-based matching kernels, and global-based kernels discussed in Section 1. This is because the QBMK kernel is based on measuring the similarity between the quantum Shannon entropies of aligned vertices, that can simultaneously reflect the global and local graph characteristics through the AMM matrix of the CTQW.

3.4. The Computational Complexity

For a pair of graphs each having n vertices, the computational complexity of the proposed QBMK kernel is mainly dominated by the procedure of computing the AMM matrix of the CTQW, that relies on the spectrum decomposition of the adjacency matrix. Thus, the whole time complexity of the QBMK kernel is $O(n^3)$, indicating a polynomial time.

4. Experiments

We evaluate the classification performance of the proposed QBMK kernel on ten benchmark graph datasets extracted from bioinformatics (BIO) (Kersting et al., 2016) and computer vision (CV) (Biasotti et al., 2003; Escolano et al., 2011), respectively. Table.2 shows the detailed statistical information of these CV and BIO datasets.

4.1. Comparisons with Graph Kernels

Experimental Setups: We empirically compare the proposed QBMK kernel against some state-of-the-art graph kernels, including: the Quantum Jensen-Shannon Kernel (QJSK) (Bai et al., 2015a), the Graphlet Counting Graph Kernel (GCGK) (Shervashidze et al., 2009) associated with the graphlets of size 4, the Jensen-Tsallis q-difference Kernel (JTQK) (Bai et al., 2014b), the Weisfeiler-Lehman Subtree Kernel (WL-SK) (Shervashidze et al., 2011), the Shortest Path Graph Kernel (SPGK) (Borgwardt & Kriegel, 2005), the Depth-based Matching Kernel (DBMK) (Bai et al., 2014a), the Extended Depth-based Matching Kernel (EDBMK) (Xu et al., 2021b), the Jensen-Shannon Subgraph Kernel (JSSK) (Bai & Hancock, 2016), and the Entropic Isomorphic Kernel (ISK) (Bai & Hancock, 2016). Detailed properties of these alternative graph kernels are shown in Table 3. For the proposed QBMK kernel, we set the parameter H as 10, this is because the required 10-layer expansion subgraphs rooted at each vertex for the QBMK kernel is able to cover most of the graph topological structures. For each kernel, we perform the 10-fold cross-validation to calculate the classification accuracy associated with the standard C-SVM (Chang & Lin, 2011). For each dataset, we utilize the optimal C-SVM parameters and run the experiment for 10 times, and calculate the averaged classification accuracies ($\pm$ standard errors) in Table 4. Note that, since the alternative kernels were evaluated with the same setup, we directly use the accuracies from the corresponding literatures.

Experimental Results and Analysis: Table 4 indicates that the proposed QBMK kernel can significantly outperform all competing graph kernels on seven of the ten datasets. Although, the QBMK kernel is not the best one on the D&D, PTC and PPIs datasets, the QBMK kernel is still competitive to most of the alternative graph kernels. The reasons for the effectiveness are threefold.

First, the proposed QBMK kernel not only encapsulates the structural correspondence information between graph structures, but also simultaneously reflects both and local structural characteristics through the quantum Shannon entropies, reflecting more comprehensive and complicated intrinsic structural information through the AMM matrix of the CTQW. By contrast, the alternative R-convolution, global-based, and local-global based kernels tend to over-

Table 3. Graph Kernels for Comparisons.

Kernel Methods	QBMK	QJSK	JTQK	GCCK	WLSK
Framework	Matching	Global	R-convolution	R-convolution	R-convolution
Based on CTQW	Yes	Yes	Yes	No	No
Capture Alignment Information	Yes	No	No	No	No
Capture Local Information	Yes	No	Yes	Yes	Yes
Capture Global Information	Yes	Yes	No	No	No
Discriminate Aligned Vertices	Yes	No	No	No	No
Kernel Methods	SPGK	EDBMK	DBMK	JSSK	ISK
Framework	R-convolution	Matching	Matching	Local-Global	R-convolution
Based on CTQW	No	No	No	No	No
Capture Alignment Information	No	Yes	Yes	No	No
Capture Local Information	Yes	Yes	Yes	Yes	Yes
Capture Global Information	No	No	No	Yes	No
Discriminate Aligned Vertices	No	No	No	No	No

Table 4. Comparisons with Different Graph Kernels.

Datasets	MUTAG	D&D	PTC	PPIs	CATH2
QBMK	**88.55** $\pm$ .43	77.60 $\pm$ .47	59.38 $\pm$ .36	84.47 $\pm$ .56	**84.36** $\pm$ .65
QJSK	82.72 $\pm$ .44	77.68 $\pm$ .31	56.70 $\pm$ .49	65.61 $\pm$ .77	71.11 $\pm$ .88
JTQK	85.50 $\pm$ .55	**79.89** $\pm$ .32	58.50 $\pm$ .39	**88.47** $\pm$ .47	68.70 $\pm$ .69
GCGK	82.04 $\pm$ .39	74.70 $\pm$ .30	55.41 $\pm$ .59	46.61 $\pm$ .47	73.68 $\pm$ 1.09
WLSK	82.88 $\pm$ .57	79.78 $\pm$ .36	58.26 $\pm$ .47	88.09 $\pm$ .41	67.36 $\pm$ .63
SPGK	83.38 $\pm$ .81	78.45 $\pm$ .26	55.52 $\pm$ .46	59.04 $\pm$ .44	81.89 $\pm$ .63
EDBMK	86.35	78.19	56.79	84.13	83.58
DBMK	85.27 $\pm$ .69	77.85	55.91	83.23	82.42
JSSK	83.77 $\pm$ .74	76.32 $\pm$ .46	56.94 $\pm$ .43	45.04 $\pm$ .88	75.42 $\pm$ .76
ISK	84.66 $\pm$ .56	75.32 $\pm$ .35	**60.26** $\pm$ .42	79.47 $\pm$ .32	67.55 $\pm$ .67
Datasets	BAR31	BSPHERE31	GEOD31	Shock	GatorBait
QBMK	**73.66** $\pm$ .57	**62.63** $\pm$ .41	**47.63** $\pm$ .45	**46.26** $\pm$ .74	**15.00** $\pm$ .89
QJSK	30.80 $\pm$ .61	24.80 $\pm$ .61	23.73 $\pm$ .66	40.60 $\pm$ .92	9.00 $\pm$ .89
JTQK	60.56 $\pm$ .35	46.93 $\pm$ .61	40.10 $\pm$ .46	37.73 $\pm$ .42	9.60 $\pm$ .87
GCGK	22.96 $\pm$ .65	17.10 $\pm$ .60	15.30 $\pm$ .68	26.63 $\pm$ .63	8.40 $\pm$ .83
WLSK	58.53 $\pm$ .53	42.10 $\pm$ .68	38.20 $\pm$ .68	36.40 $\pm$ 1.00	10.10 $\pm$ .61
SPGK	55.73 $\pm$ .44	48.20 $\pm$ .76	38.40 $\pm$ .65	37.88 $\pm$ .93	9.00 $\pm$ .75
EDBMK	70.08	57.36	43.57	33.24	14.40
DBMK	69.40 $\pm$ .56	56.43 $\pm$ .69	42.83 $\pm$ .50	26.73	13.76
JSSK	52.76 $\pm$ .47	43.33 $\pm$ .40	32.03 $\pm$ 1.02	37.66 $\pm$ .80	9.20 $\pm$ .65
ISK	62.80 $\pm$ .47	52.50 $\pm$ .74	39.76 $\pm$ .43	39.86 $\pm$ .68	11.40 $\pm$ .52

look the structural alignment information between graphs. Moreover, the alternative R-convolution, global-based, and vertex-based matching kernels only focus on one of the global and local structural information, reflecting limited structural information. **Second**, although the alternative EDBMK and DBMK kernels can integrate the structural alignment information, these two vertex-based matching kernels focus on simply counting the number of pairwise aligned vertices based on the single-level correspondence matrix. By contrast, the proposed QBMK kernel not only encapsulates multilevel correspondence information between graphs, but also discriminates the structural differences between different pairs aligned vertices through the AMM matrix. **Third**, although the local-global based JSSK kernel can also simultaneously capture both the global and local graph structure information, by measuring the entropy-based similarities between each pair of h-layer expansion subgraphs around the centroid vertex. However, the JSSK kernel relies on a limited number of subgraphs around the centroid vertex. By contrast, the QBMK kernel can measure the similarities between the expansion subgraphs around any vertex, reflecting more structure characteristics. Finally, the proposed QBMK kernel relies on the quantum Shannon entropies through quantum walks (i.e., the CTQW) rather than classical random walks (i.e., the C-TRW). Thus, the proposed QBMK kernel can reflect more complicated intrinsic structure characteristics.

4.2. Comparisons with Graph Deep Learning

Experimental Setups: We empirically compare the proposed QBMK kernel with some classical graph deep learning approaches, including: the EigenPooling based Graph Convolution Network (EigenPool) (Lee et al., 2019), the Specified Degree-based Graph Neural Network (DEMO-Net) (Wu et al., 2019), the Deep Graphlet Kernel (DGK) (Yanardag & Vishwanathan, 2015), the Deep Graph Convolution Neural Network (DGCNN) (Zhang et al., 2018), and the Diffusion-based Convolution Neural Network (DCNN) (Atwood & Towsley, 2016). Note that, all these alternative graph deep learning approaches are evaluated based on the same 10-fold cross-validation strategy with us. Thus, we straightforwardly report the accuracies from the original literatures in Table 5.

Experimental Results and Analysis: Table 5 indicates that the proposed QBMK kernel outperforms most of the deep learning approaches on two of the three datasets. Although the QBMK kernel is not the best on the D&D

Table 5. Comparisons with Graph Deep Learning Methods.

Datasets	MUTAG	D&D	PTC
QBMK	**88.55** $\pm$ .43	77.60 $\pm$ .47	**59.38** $\pm$ .36
EigenPool	79.50	76.60	–
DEMO-Net	81.40	70.80	57.20
DGK	82.66 $\pm$ 1.45	**78.50** $\pm$ 0.22	57.32 $\pm$ 1.13
DCNN	66.98	58.09 $\pm$ 0.53	56.60
DGCNN	85.83 $\pm$ 1.66	79.37 $\pm$ 0.94	58.59 $\pm$ 2.47

dataset, it is still competitive to most of the alternative deep learning approaches. In fact, the graph kernel associated with the C-SVM can be theoretically seen as a shallow learning approach, that may have lower classification performance than that of the deep learning approaches. Moreover, unlike the deep learning approach, the kernel computation cannot participate the end-to-end training process for the C-SVM. However, even under this disadvantageous context, the QBMK kernel still outperforms these alternative deep learning approaches, again demonstrating the effectiveness. The reason of the effectiveness may due to the fact that, the proposed QBMK kernel can capture more complicated intrinsic graph structure characteristics through the AMM matrix of the CTQW, that can be seen as a quantum-based graph structure representation. By contrast, the alternative deep learning approaches can only reflect graph structure information through the original graph structure representation (i.e., the adjacency matrix). Moreover, these graph deep learning approaches cannot encapsulate the structural correspondence information between graphs. This reveals that the framework of the alignment strategy associated with the CTQW for the QBMK kernel tremendously improves the performance of graph kernels.

5. Conclusions

In this work, we have developed a novel QBMK kernel for un-attributed graphs, through the AMM matrix of the CTQW. Theoretical analysis indicates that the proposed QBMK kernel can simultaneously address different theoretical drawbacks arising in the existing classical R-convolution, global-based, local-global based, and vertex-based matching kernels. The experimental evaluation demonstrates the effectiveness of the proposed kernel.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants T2122020, 61602535, 61976235. Ming Li acknowledged the supports from the Key R&D Program of Zhejiang Province (No. 2024C03262), the National Nature Science Foundation of China (No. U21A20473 and No. 62172370), and Jinhua Science and Technology Plan (No. 2023-3-003a). This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFC3302300. This work was also supported in part by the Program for Innovation Research in the Central University of Finance and Economics, and the Emerging Interdisciplinary Project of CUFE.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Atwood, J. and Towsley, D. Diffusion-convolutional neural networks. In *Proceedings of NIPS*, pp. 1993–2001, 2016.

Azaïs, R. and Ingels, F. The weight function in the subtree kernel is decisive. *J. Mach. Learn. Res.*, 21:67:1–67:36, 2020.

Aziz, F., Wilson, R. C., and Hancock, E. R. Backtrackless walks on a graph. *IEEE Transactions on Neural Networks and Learning Systems*, 24(6):977–989, 2013.

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

Bai, L. and Hancock, E. R. Fast depth-based subgraph kernels for unattributed graphs. *Pattern Recognition*, 50: 233–245, 2016.

Bai, L., Ren, P., Bai, X., and Hancock, E. R. A graph kernel from the depth-based representation. In *Proceedings of S+SSPR*, pp. 1–11, 2014a.

Bai, L., Rossi, L., Bunke, H., and Hancock, E. R. Attributed graph kernels using the jensen-tsallis q-differences. In *Proceedings of ECML-PKDD*, pp. I:99–114, 2014b.

Bai, L., Rossi, L., Torsello, A., and Hancock, E. R. A quantum jensen-shannon graph kernel for unattributed graphs. *Pattern Recognition*, 48(2):344–355, 2015a.

Bai, L., Rossi, L., Zhang, Z., and Hancock, E. R. An aligned subtree kernel for weighted graphs. In *Proceedings of ICML*, pp. 30–39, 2015b.

Biasotti, S., Marini, S., Mortara, M., Patanè, G., Spagnuolo, M., and Falcidieno, B. 3d shape matching through topological structures. In *Proceedings of DGCI*, pp. 194–203, 2003.

Borgwardt, K. M. and Kriegel, H. Shortest-path kernels on graphs. In *Proceedings of ICDM*, pp. 74–81, 2005.

Chang, C.-C. and Lin, C.-J. Libsvm: A library for support vector machines. *Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm*, 2011.

Costa, F. and Grave, K. D. Fast neighborhood subgraph pairwise distance kernel. In *Proceedings of ICML*, pp. 255–262, 2010.

Emms, D., Wilson, R. C., and Hancock, E. R. Graph matching using the interference of continuous-time quantum walks. *Pattern Recognition*, 42(5):985–1002, 2009.

Escolano, F., Hancock, E., and Lozano, M. Graph matching through entropic manifold alignment. In *CVPR*, pp. 2417–2424, 2011.

Fröhlich, H., Wegner, J. K., Sieker, F., and Zell, A. Optimal assignment kernels for attributed molecular graphs. In *Proceedings of ICML*, pp. 225–232, 2005.

Gärtner, T., Flach, P. A., and Wrobel, S. On graph kernels: Hardness results and efficient alternatives. In Schölkopf, B. and Warmuth, M. K. (eds.), *Proceedings of COLT*, volume 2777 of *Lecture Notes in Computer Science*, pp. 129–143.

Gärtner, T., Flach, P., and Wrobel, S. On graph kernels: hardness results and efficient alternatives. In *Proceedings of COLT*, pp. 129–143, 2003.

Gkirtzou, K. and Blaschko, M. B. The pyramid quantized weisfeiler-lehman graph representation. *Neurocomputing*, 173:1495–1507, 2016.

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

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

Jebara, T., Kondor, R., and Howard, A. Probability product kernels. *Journal of Machine Learning Research*, 5:819–844, 2004.

Johansson, F. D., Jethava, V., Dubhashi, D. P., and Bhattacharyya, C. Global graph kernels using geometric embeddings. In *Proceedings of ICML*, volume 32 of *JMLR Workshop and Conference Proceedings*, pp. 694–702. JMLR.org, 2014.

Kalofolias, J., Welke, P., and Vreeken, J. SUSAN: the structural similarity random walk kernel. In Demeniconi, C. and Davidson, I. (eds.), *Proceedings of SDM*, pp. 298–306, 2021.

Kersting, K., Kriege, N. M., Morris, C., Mutzel, P., and Neumann, M. Benchmark data sets for graph kernels, 2016. URL http://graphkernels.cs.tu-dortmund.de.

Kriege, N. and Mutzel, P. Subgraph matching kernels for attributed graphs. In *Proceedings of ICML*, 2012.

Lee, J., Lee, I., and Kang, J. Self-attention graph pooling. In *Proceedings of ICML*, pp. 3734–3743, 2019.

Ong, C. S., Mary, X., Canu, S., and Smola, A. J. Learning with non-positive kernels. In Brodley, C. E. (ed.), *Proceedings of ICML*, volume 69, 2004.

Rossi, L., Torsello, A., and Hancock, R. R. Measuring graph similarity through continuous-time quantum walks and the quantum jensen-shannon divergence. *Physical Review E*, 91(2):022815, 2015.

Schulz, T. H., Welke, P., and Wrobel, S. Graph filtration kernels. In *Proceedings of AAAI*, pp. 8196–8203, 2022.

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

Shervashidze, N., Schweitzer, P., van Leeuwen, E. J., Mehlhorn, K., and Borgwardt, K. M. Weisfeiler-lehman graph kernels. *Journal of Machine Learning Research*, 1:1–48, 2010.

Shervashidze, N., Schweitzer, P., van Leeuwen, E. J., Mehlhorn, K., and Borgwardt, K. M. Weisfeiler-lehman graph kernels. *Journal of Machine Learning Research*, 12:2539–2561, 2011.

Sugiyama, M. and Borgwardt, K. M. Halting in random walk kernels. In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R. (eds.), *Proceedings of NIPS*, pp. 1639–1647, 2015.

Togninalli, M., Ghisu, M. E., Llinares-López, F., Rieck, B., and Borgwardt, K. M. Wasserstein weisfeiler-lehman graph kernels. In *Proceedings of NeurIPS*, pp. 6436–6446, 2019.

Watrous, J. Quantum simulations of classical random walks and undirected graph connectivity. *J. Comput. Syst. Sci.*, 62(2):376–391, 2001.

Weisfeiler, B. and Lehman, A. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsia*, Ser.2(9), 1968.

Wu, J., He, J., and Xu, J. Demo-net: Degree-specific graph neural networks for node and graph classification. In Teredesai, A., Kumar, V., Li, Y., Rosales, R., Terzi, E., and Karypis, G. (eds.), *Proceedings of ACM SIGKDD*, pp. 406–415. ACM, 2019.

Xu, B., Ting, K. M., and Jiang, Y. Isolation graph kernel. In *Proceedings of AAAI*, pp. 10487–10495, 2021a.

Xu, L., Jiang, X., Bai, L., Xiao, J., and Luo, B. A hybrid reproducing graph kernel based on information entropy. *Pattern Recognit.*, 73:89–98, 2018.

Xu, L., Bai, L., Jiang, X., Tan, M., Zhang, D., and Luo, B. Deep rényi entropy graph kernel. *Pattern Recognit.*, 111: 107668, 2021b.

Yanardag, P. and Vishwanathan, S. V. N. Deep graph kernels. In *Proceedings of KDD*, pp. 1365–1374, 2015.

Zhang, M., Cui, Z., Neumann, M., and Chen, Y. An end-to-end deep learning architecture for graph classification. In *Proceedings of AAAI*, 2018.