# Distinguishability of graphs: a case for quantum-inspired measures

Athanasia Polychronopoulou
Temple University
Philadelphia, PA 19122 USA
Email: n.polychr@temple.edu

J. Alshehri
Temple University
Philadelphia, PA 19122 USA
Email: shehri.j@temple.edu

Zoran Obradovic
Temple University
Philadelphia, PA 19122 USA
Email: zoran.obradovic@temple.edu

***Abstract*—The question of graph similarity or graph distinguishability arises often in natural systems and their analysis over graphical networks. In many domains, graph similarity is used for graph classification, outlier detection or the identification of distinguished interaction patterns. Several methods have been proposed on how to address this topic, but graph comparison still presents many challenges. Recently, information physics has emerged as a promising theoretical foundation for complex networks. In many applications, it has been demonstrated that natural complex systems exhibit features that can be described and interpreted by measures typically applied in quantum mechanical systems. Therefore, a natural starting point for the identification of network similarity measures is information physics and a series of measures of distance for quantum states. In this work, we report experiments on synthetic and real-world data sets, and compare quantum-inspired measures to a series of state-of-the-art and well-established methods of graph distinguishability. We show that quantum-inspired methods satisfy the mathematical and intuitive requirements for graph similarities, while offering high interpretability.**

## I. Introduction

Most of the natural systems that surround us can be seen as a collection of entities interacting with each other. Social networks, genetic and protein interaction networks, airline and road traffic networks, brain connectivity networks and web graphs are only some of the examples. These systems are typically analyzed using network theory, where they are represented as complex networks (graphs), whose nodes (entities) are connected through edges (interactions).

The question of graph distinguishability or graph similarity often arises in these natural systems [1]. For example, social networks are compared to identify distinguished interaction patterns, daily traffic networks are compared to facilitate the detection of abnormal change in traffic patterns, brain networks are compared to extract meaningful connectivity information at the population level, and web graphs are compared for anomaly detection. In all these settings the similarity between two graphs with overlapping sets of nodes is assessed and the detection of changes in the connectivity patterns is important. This problem is different from the problem of inexact graph matching, the graph isomorphism or the maximum common subgraph problem, where the node correspondence is unknown [2].

Several approaches have been proposed to solve different variations of the problem of calculating graph similarity. The simplest approach would be the calculation of a variety of the graphs' structural properties (diameter, edges distribution, degree and eigenvalues) and their subsequent comparison, but this approach is not able to capture every aspect of the graphs. More commonly, methods that estimate the graph edit distance (GED) are used. GED measures the dissimilarity between two graphs as the minimal cost of a sequence of elementary operations transforming one graph into another. Exact computation of GED is NP-hard and typically approximate or tangent solutions are implemented. Other solutions have been proposed, such as the usage of the graph spectra [1], or measures inspired by document similarity [3] and other intuitive approaches [4]. More recently, spectral distances as well as distances based on node affinities have been studied in more depth [5], and the usage of Quantum Jensen Divergence (QJSDiv) as a measure of graph distinguishability, has also been proposed [6]. However, graph comparison remains an open problem. Most of the solutions proposed so far are empirical in nature and a more mathematical method with high interpretability is required. Toward this goal, we study a series of measures that have been effectively used in quantum mechanics and quantum information theory, they satisfy the mathematical properties of a metric and similarity measure and offer intelligible results.

Quantum mechanics is a valuable resource for the investigation of the behavior of complex networks [7]. A variety of quantum systems have been seen as metaphors for natural systems described by complex networks. Quantum gases have been used to describe network evolution and the emergence of different structures in complex networks, has been represented in terms of a quantum–classical transition for quantum gases [8]. Quantum transport probability and state fidelity have been implemented as a closeness function and used for community detection [9], while quantum random walks have


ASONAM '21, November 8–11, 2021, Virtual Event, Netherlands © 2021 Association for Computing Machinery. ACM ISBN 978-1-4503-9128-3/21/11…$15.00 https://doi.org/10.1145/3487351.3488330

been shown to offer significant performance improvements on traditional computer science algorithms with respect to the classical random walks [10], and have been implemented for a quantum-inspired ranking algorithm [11]. Quantum distances are used to reduce the complexity of calculating Euclidean distances between multidimensional points,and improve the performance of algorithms such as k-means [12]. Given this interconnection between quantum mechanics and complex networks, quantum network theory is a natural place to look for similarity measures.

In this work we evaluate the use of quantum inspired measures for graph distinguishablity. We consider four possible measures, that have been broadly used in quantum information theory as measures of distance between quantum states [13]. They have been extended for use with quantum processes and have been studied in detail for quantum systems [14], [15]. In this work we evaluate their ability to distinguish two complex networks and we elucidate the properties of the four distance measures. We first utilize a series of artificial data to study in detail the effectiveness of these measures with various graphical structures and for complex network distances of various types and significance. We then report experiments on real-world data sets, from a variety of domains, and compare the quantum-inspired measures to a series of state-of-the-art and well established methods.

## II. Methods

In an attempt to define a measure of distance for complex networks that is mathematically well established and has high interpretability, we consider a series of distance measures for quantum states, that are particularly important in Quantum Information Science [13], [14], [15]. In this work we limit our consideration to measures which have been shown to be a metric, or in the case of the Quantum Jensen-Shannon Divergence are expected to be metrics [16]. The metric character of a distance measure between two complex networks $G_\sigma$ and $G_\rho$ requires three fundamental properties: $D(G_\sigma, G_\rho) \geq 0$ , with the equality to zero occurring if and only if $G_\sigma = G_\rho$, $D(G_\sigma, G_\rho) = D(G_\rho, G_\sigma)$ i.e. the distance measure is symmetric and finally the triangle inequality is satisfied and $D(G_\sigma, G_\rho) \leq D(G_\sigma, G_\tau) + D(G_\tau, G_\rho)$.

The quantum-inspired distance measures considered in this work are based on the definition of a density matrix. In quantum mechanics, a density matrix is a matrix that describes the statistical state of a quantum mechanical system. Mathematically, it is a hermitian matrix that is positive semidefinite with trace equal to 1. In graph theory, the density matrix $\rho$ of a graph can be defined through the combinatorial Laplacian of the graph [17]. Let $G = (V, E)$ be a simple undirected graph with a set of vertices $V(G) = 1, 2, ..., n$ and a set of edges $E(G) \subseteq V(G) \times V(G)$. Then the density matrix is defined as:

$$\rho_G = \frac{\Delta(G) - A(G)}{Tr(\Delta(G))} \tag{1}$$

In this equation $A(G)$ is the adjacency matrix with $[A(G)]_{u,v} = 1$ if $u, v \in E(G)$ and $[A(G)]_{u,v} = 0$ otherwise. $\Delta(G)$ is the degree matrix, which is a diagonal matrix with elements equal to the degree $d(u)$ of each node $u$ i.e. the number of edges adjacent to each vertex. The normalization using the trace of $\Delta(G)$ guarantees that the density matrix will have a trace of 1. For a weighted graph $G = (V, E, W)$ each edge is associated with an edge weight and in this case the density matrix becomes:

$$\rho_G = \frac{\Delta(G) - W(G)}{Tr(\Delta(G))} \tag{2}$$

where $W(G)$ is the weights matrix with $[W(G)]_{u,v} = w_{uv}$ and $[W(G)]_{u,v} = 0$ if the nodes $u$ and $v$ are not connected. The degree matrix $\Delta(G)$ is again a diagonal matrix holding for each node $u$ the value $d_u = \sum_{v=1}^{n} w_{uv}$.

Using this formulation we can then compare networks the same way we would compare states in quantum mechanical systems. In this work we evaluate the following measures, from quantum information processing, for use with complex networks:

### A. Trace Distance

The trace distance between two quantum states, or two networks, with density matrices $\sigma$ and $\rho$ is given by:

$$D_{trace}(\sigma||\rho) = \frac{1}{2}[trace(\sqrt{(\sigma - \rho)^2})] \tag{3}$$

It is a metric and is bounded to be $0 \leq D_{trace} \leq 1$ with the equality to 0 holding if and only if $\rho = \sigma$ and the equality to 1 holding if and only if $\rho$ and $\sigma$ have orthogonal supports. The trace distance is the quantum generalization of the Kolmogorov distance for classical probability distributions and, as it's classical counterpart, the trace distance can be interpreted to represent the maximum probability of distinguishing between two quantum systems, or in our case two networks.

### B. Hilbert–Schmidt distance

The Hilbert-Schmidt distance

$$D_{HS}(\sigma||\rho) = \sqrt{trace(\sigma - \rho)^2} \tag{4}$$

is a Riemannian metric that is bounded to be $0 \leq D_{HS} \leq 2D_{trace}$. As it is defined on the space of operators it is unclear how to impose an operational interpretation, however, in [18] the authors suggest that it can be seen as an information distance between two quantum states. It has been recently used in [12] to reduce the complexity of calculating Euclidean distances between multidimensional points and thereby reduce the complexity of classification algorithms such as k-means but it has not been studied so far as a measure for the distinguishability of graphs.

### C. Hellinger distance

The Hellinger distance is given by:

$$D_H(\sigma||\rho) = \sqrt{2[1 - A(\sigma||\rho)]} \tag{5}$$

with

$$A(\sigma||\rho) = trace(\sqrt{\sigma}\sqrt{\rho}) \tag{6}$$

representing the quantum affinity, a measure that characterizes the closeness of two quantum states and whose classical analog is the Bhattacharya coefficient between two classical probability distribution. The Hellinger distance is a metric, bounded to be $0 \leq D_H \leq \sqrt{2}$ [19].

### D. Bures distance

The Bures distance is expressed through quantum Fidelity, which is a measures of overlap between two quantum states:

$$F(\sigma||\rho) = [trace(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})]^2 \tag{7}$$

It can be shown that the fidelity is symmetric and is bounded to be $0 \leq F(\sigma||\rho) \leq 1$, with $F(\sigma||\rho) = 1$ if and only if $\sigma = \rho$. Although not a metric, the fidelity can easily be turned into a metric and the most common approach is the Bures metric:

$$D_B(\sigma||\rho) = \sqrt{2[1 - \sqrt{F(\sigma||\rho)}]} \tag{8}$$

### E. Quantum JSDiv

The Von Neumann entropy of a network, similarly to the Von Neumann entropy of a quantum system, can be interpreted as a measure of regularity and is given by the expression:

$$S(\rho) = -Tr(\rho \log_2 \rho) \tag{9}$$

Then, the distinguishability of two quantum states or two networks with density matrices $\sigma$ and $\rho$ can be measured using the von Neumann relative entropy (or quantum relative entropy) defined as:

$$S_N(\sigma||\rho) = Tr\sigma(ln\sigma - ln\rho) \tag{10}$$

where $ln\sigma$ is the matrix base 2 logarithm of $\sigma$. Notice that the quantum relative entropy is the quantum mechanical analog of relative entropy, otherwise called Kullback–Leibler divergence, commonly used in statistics as a measure of comparison for probability distributions. Relative entropy is not symmetric, does not satisfy the triangle equality and it is well defined only if the support of $\sigma$ is a subset of the support of $\rho$ (the support being the subspace spanned by the eigenvectors of the density matrix with non-zero eigenvalues). However, it can be extended to provide the Quantum Jensen-Shannon divergence ($QJSDiv$) [20], given by:

$$QJSDiv(\sigma||\rho) = \frac{1}{2}[S_N(\sigma||\frac{\sigma+\rho}{2}) + S_N(\rho||\frac{\sigma+\rho}{2})] \tag{11}$$

where $S_N$ is the quantum relative entropy of equation 10 and can be alternatively written as:

$$QJSDiv(\sigma||\rho) = S(\frac{\sigma+\rho}{2}) - \frac{1}{2}S(\rho) - \frac{1}{2}S(\sigma) \tag{12}$$

where $S$ is the quantum entropy of equation 9. Quantum Jensen-Shannon divergence was introduced as a measure of distinguishability between mixed quantum states [20], and it's properties have been extensively studied [16]. It bounded to be $0 \leq QJSDiv \leq 1$, with the equality to 0 holding if and only if $\rho = \sigma$, and it is always well-defined, with the restriction previously imposed on the supports of $\rho$ and $\sigma$ now lifted. The authors of [6] have shown that Quantum Jensen-Shannon divergence can be used to quantify the distance between pairs of networks, and they have applied it to successfully cluster the layers of a multilayer system. In this work, we aim to compare the performance of Quantum Jensen-Shannon divergence with the rest of the measures presented here as well as a series of baseline measures commonly used in the domain.

## III. Results

The experimental evaluation of the proposed graph distance measures is divided in two major parts, each concerned with applications on different types of data. The first part utilizes a series of experiments on artificial data aiming to answer three important research questions: 'Can each of the measures act as a measure of distinguishability between graphs?', 'Can each of the measures capture intuitive similarity aspects?' and finally 'Can each of the measures capture structural characteristics?'. In the second part we use the proposed distance measures in real-world applications using data from different domains and answer the question: 'Can each of the measures be effectively used as a similarity measure for real-world complex data?'.

### A. Experiments on Artificial Data

*1) 'Can each of the measures act as a measure of distinguishability between graphs?':* To answer this research question, we run a series of experiments following the general concept of graph edit distance. Artificial networks with 100 nodes are generated and then modified using a discrete-time process in which one elementary graph edit operation is applied at every step. The artificial networks are undirected and unweighted, and they are created using several graphical models or canonical structures: Random graphs, Preferential attachment, a symmetric Forest Fire model, binary trees, lines, circles, and a fully connected clique network. Three types of elementary graph edit operations are studied: the removal of a randomly selected edge, the removal of a randomly selected node and the rewiring of two randomly selected edges. For every step, we calculate the distance between the modified and original network and the results are presented in Figure 1. The monotonically increasing behavior of the distance measures is evident and confirms that they act as measures of graph distinguishability, accommodating all possible elementary graph edit operations and all types of networks.

As the elementary graph edit step increases and the graph becomes smaller the curvature of the plots increases. As intuitively expected, the same modification is more crucial for smaller networks. Notice that this is not the case for node removals and Bures Distance, while the Hellinger distance is equally sensitive to modifications of large and small networks, something that is also very intuitive. It is interesting to notice that, as intuitively expected, in the cases of edge rewiring and removal, the distance increases more rapidly for graphs like lines and circles and less rapidly for random graphs.

*2) 'Can each of the measures capture intuitive similarity aspects?':* Graph similarity measures should take into account that targeted operations are expected to affect the network

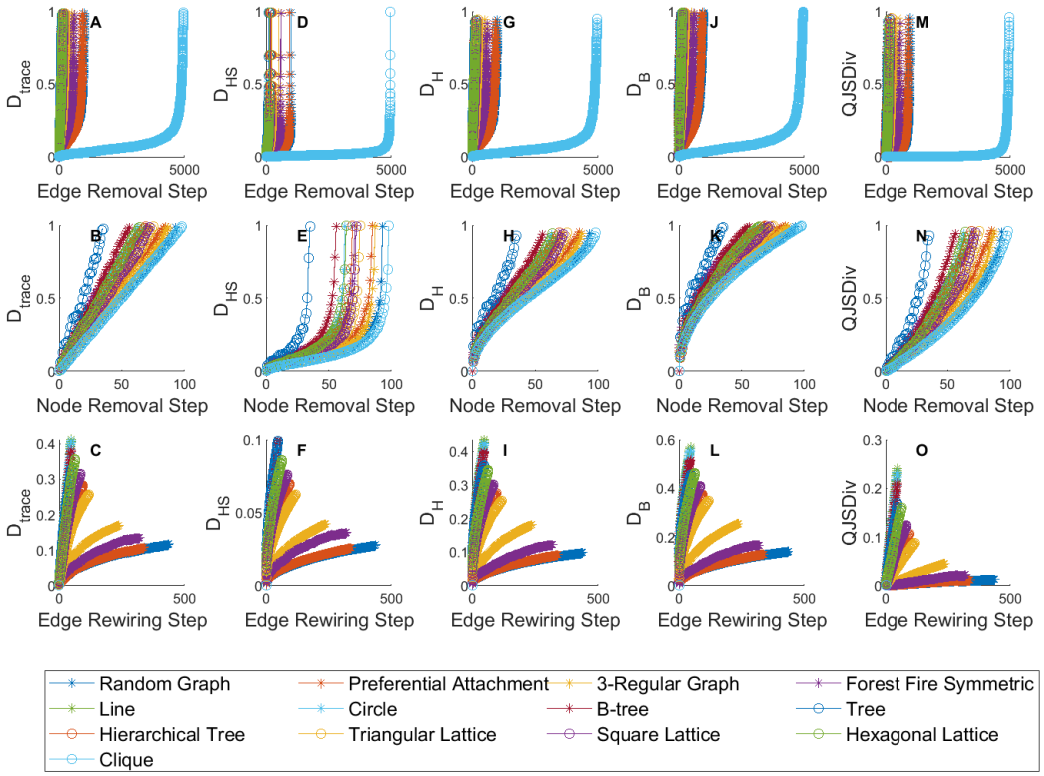


Fig. 1. The progression of each of the distance measures between the original graph and the graph after applying a series of elementary graph edit operations. Each elementary graph edit operation corresponds to a random edge removal, a random node removal, or the rewiring of two randomly chosen edges.

more than random ones, and operations that create disconnected components in the graph should affect it more than operations that do not. To answer our second research question we generate a random graph, and apply a similar discrete-time process with repeated edge removals. These edge removals occur in three different ways: with a random process, where the edge to be removed is randomly selected, with a semi-targeted process, by randomly selecting a node and removing in a random order all of it's edges before proceeding to the next node, and with a fully-targeted process, targeting again one node at a time, but now the targeted node is the one with the maximum degree in the graph. The results are presented on Figure 2. It is clear that targeted modifications result in higher distance values. The two modes of targeted edge selection do not appear to be significantly different, and a closer look at the plots reveals the reason. In the fully targeted approach all the popular nodes are removed in order, while the graph is still large enough and the effect of their removal is still small. In the targeted random approach, some popular nodes are removed at the beginning while others are removed in later steps, when the effect is more significant. The bottom part of Figure 2 answers a different aspect of the same question. The steps of the discrete-time edge removal process are separated in two categories based on whether the step's edge removal created a new disconnected component or not. Then we compare the values of each of the distance measures between the two categories. The results show that, the distance values are in general larger and more varied when the edge removal generates a disconnected component. For a more precise analysis a two sample t-test is applied, testing the null hypothesis that the two data samples come from two distributions with the same mean, without assuming equality of the variances. The hypothesis is in each case rejected at 5% significance level, with p values (averaged over 10 repetitions): $p_{trace} = 6.4 * 10^{-29}$, $p_{HS} = 0.0081$, $p_H = 1.8 * 10^{-48}$, $p_B = 1.3 * 10^{-32}$, and $p_{QJSDiv} = 1.2 * 10^{-21}$, indicating a smaller separation for $QJSDiv$ and $D_{HS}$.

The case of weighted networks is studied by expanding our elementary graph edit operations to include edge re-weighting and edge weight redistribution. In the first case, the weight of a random edge is substituted by a new random value and in the second case two random edges are chosen and their weights are interchanged. We use a random graph with weights generated randomly from a uniform distribution and the discrete-time process with elementary graph edit operations is repeated. These operations are random node removals, random edge removals, targeted removal of the edge with the highest weight, edge rewiring, and edge weight redistribution. The results are presented in Figure 3. As intuitively expected, node removal is much more effective than edge removal. Furthermore, removing higher weight edges is more effective than removing edges randomly. Finally, in all cases the removal of an edge causes a bigger change to the graph than edge re-weight or rewiring. For these two cases $QJSDiv$ and $D_{HS}$ maintain low

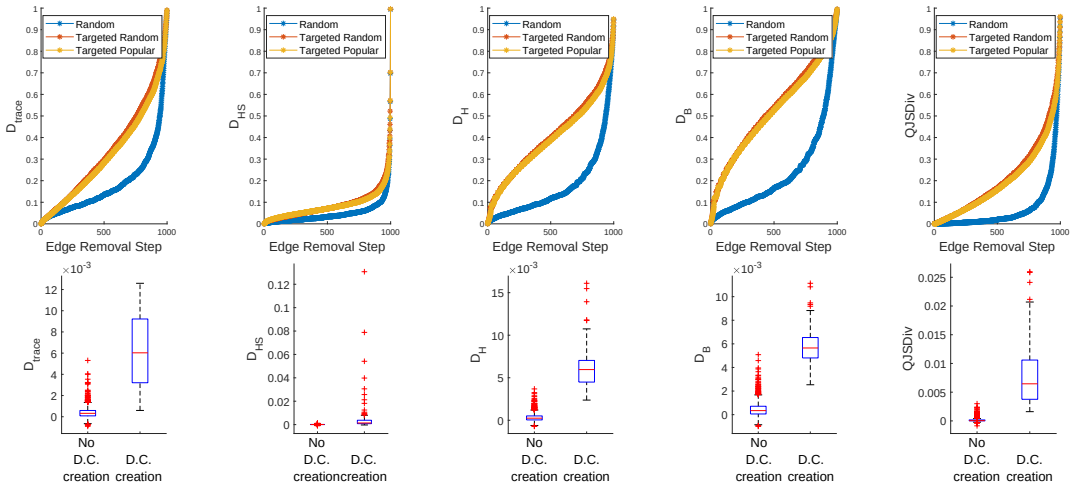


Fig. 2. Top:The progression of the distance measures with continuous edge removals. The edge selection can be random, targeted with random node selection or targeted with popular node selection. Bottom: Comparison of the values of the distance measures between two categories of edge removal steps: those that created a new disconnected component and those that did not.

values for the entire process and seem to be less sensitive.

*3) 'Can each of the measures capture structural characteristics?':* For the last research question, a similar experimental process is applied on a random graph, but this time at every step several structural properties of the graph along with the distance measures are calculated. The correlation between the evolution of the value of each structural measure and the value of each distance is calculated and reported on Table I. A series of structural properties are considered, each describing a different aspect of the network and representing both global and local graph characteristics. The closeness centrality of a node measures the mean distance of it to other nodes and can be seen as evidence of the access that the node has to information or the influence it has on other nodes. It has been shown that the closeness centrality distance can be used to effectively distinguish between randomly generated and actual evolutionary paths of two dynamic social networks [21]. Eigenvector centrality is also a measure of the importance of the node, measured now with respect to both the number and the importance of the node's neighbors [22]. In graph theory, the clustering coefficient is a well-established measure of the degree to which nodes in a graph tend to cluster together. The local clustering coefficient acts complementary to it, as it can be used as a probe for the existence of so-called 'structural holes' in a network and, in this sense, it represents a local betweenness centrality, measuring the extent to which a vertex lies on paths between its immediate neighbors. The eccentricity of a node [23] is it's maximum graph distance to any other node and is a measure commonly used to extract graph characteristics such as the radius and diameter. The friends of a friend measure counts the nodes that are connected to one's neighbors but are not connected to the one, while the common neighbors [24], the Adamic-Adar [25] and the Jaccard Coefficient [26] provide an actual and a normalized count of the common neighbors between two nodes. Finally, graph entropy can be interpreted as a measure of graph regularity as it tends to increase with the number of connected components, long paths, and nontrivial symmetries [27]. All distances present high correlations with these measures making it clear that they can all capture structural characteristics and hence can be used as measures of structural change. Averaging these results for each distance measure we get $\mu_{trace} = 0.7150$, $\mu_{HS} = 0.6099$, $\mu_H = 0.7160$, $\mu_B = 0.7135$, $\mu_{QJSDiv} = 0.7014$, showing a very consistent correlation of the Hellinger distance.

## B. Experiments on Real Data

*1) Graph Classification:* We evaluate the effectiveness of the distance measures using three real world data sets available under SNAP data repository [28]. The first data set describes all incoming and outgoing email between members of a large

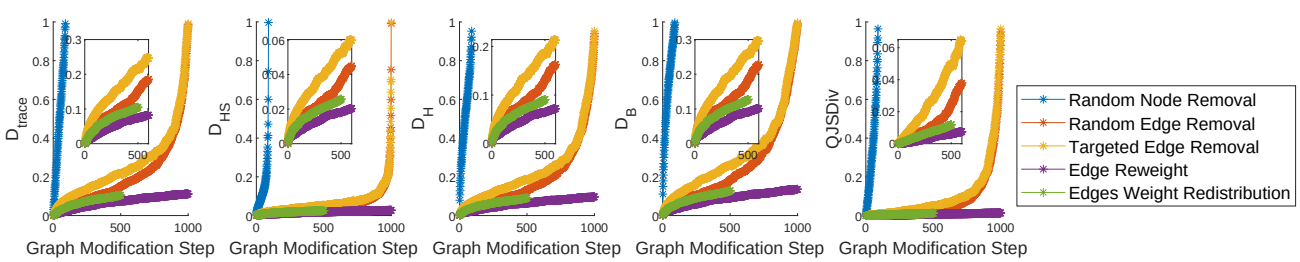


Fig. 3. The progression of the distance measures after applying a series of elementary graph edit operations on the original graph. Different types of graph edit operations are tested for the case of weighted networks.

| | Node Removal | | | | | Edge Removal | | | | | Edge Rewiring | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Structural Measure | $D_{trace}$ | $D_{HS}$ | $D_H$ | $D_B$ | $QJSDiv$ | $D_{trace}$ | $D_{HS}$ | $D_H$ | $D_B$ | $QJSDiv$ | $D_{trace}$ | $D_{HS}$ | $D_H$ | $D_B$ | $QJSDiv$ |
| Closeness Centrality | 0.14 | 0.07 | 0.18 | **0.21** | 0.11 | 0.71 | 0.48 | 0.72 | **0.76** | 0.56 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 |
| Clustering Coef | 0.56 | **0.73** | 0.55 | 0.44 | 0.67 | 0.80 | 0.60 | 0.80 | **0.82** | 0.67 | 0.74 | 0.74 | 0.74 | 0.74 | **0.82** |
| Local Clustering Coef | 0.97 | 0.76 | 0.95 | 0.92 | **0.98** | 0.87 | 0.62 | 0.88 | **0.91** | 0.74 | 0.73 | 0.74 | 0.74 | 0.74 | **0.80** |
| Entropy Difference | 0.89 | **0.96** | 0.89 | 0.81 | 0.95 | 0.92 | 0.96 | 0.90 | 0.86 | **0.97** | 0.61 | 0.62 | 0.62 | 0.62 | **0.69** |
| Eccentricity | 0.89 | 0.84 | 0.88 | 0.82 | **0.94** | 0.11 | 0.26 | 0.07 | 0.01 | **0.29** | 0.59 | 0.59 | 0.59 | 0.59 | **0.62** |
| Eigenvector Centrality | 0.98 | 0.81 | 0.97 | 0.93 | **0.99** | 0.96 | 0.74 | 0.97 | **0.98** | 0.92 | **0.63** | 0.62 | 0.62 | 0.62 | 0.55 |
| Friend of a Friend | 0.92 | 0.57 | 0.94 | **0.98** | 0.84 | 0.77 | 0.53 | 0.78 | **0.82** | 0.61 | 0.74 | 0.74 | 0.74 | 0.74 | **0.82** |
| Adamic Adar | 0.95 | 0.61 | 0.96 | **0.99** | 0.87 | 0.82 | 0.57 | 0.83 | **0.87** | 0.68 | 0.26 | 0.28 | 0.28 | 0.28 | **0.41** |
| Common Neighbor | 0.92 | 0.56 | 0.93 | **0.97** | 0.83 | 0.75 | 0.51 | 0.76 | **0.80** | 0.59 | NaN | NaN | NaN | NaN | NaN |
| Jaccard Coef | 0.98 | 0.67 | 0.98 | **0.99** | 0.93 | 0.92 | 0.68 | 0.92 | **0.95** | 0.80 | 0.71 | 0.71 | 0.71 | 0.71 | 0.70 |

TABLE I

THE PEARSON CORRELATION BETWEEN THE EVOLUTION OF DIFFERENT STRUCTURAL MEASURES AND THE EVOLUTION OF THE DISTANCE MEASURES, FOR THREE DIFFERENT GRAPH MODIFICATION APPROACHES. THE RESULTS ARE AVERAGES OVER 5 REPETITIONS.

European research institution [29]. It contains information on 1005 researchers, and their communications creating a network of 1005 nodes and 25571 edges, where an edge $(u, v)$ exists if researchers $u$ and $v$ have exchanged at least one email. Using these data we extract the ego-network of each node, creating 1005 networks with 1005 nodes each (self-edges are not considered). Each individual belongs to exactly one of 42 departments and the purpose of this experiment is to predict the department of each researcher using only the intrinsic properties of the email exchange network. The assumption is that researchers who belong in the same department will have similar email ego-networks.

The second data set is a butterfly similarity network. Nodes represent organisms and edges represent visual similarities between them [30], calculated using butterfly images. The data set contains information on 832 butterflies in 10 classes, with each class consisting of between 55 and 100 samples. Using these data we extract the ego-network of each node, creating 832 networks with 832 nodes each. The purpose of this experiment is to predict the class of each butterfly using again only the structure of their ego-networks. The assumption is that butterflies who belong in the same class will have similar ego-networks.

The third data set describes the ego-networks of Eastern European users collected from the music streaming service Deezer in February 2020. Nodes are users and edges are mutual follower relationships [31]. The ego-networks of the 9,629 users have been extracted, and the related classification task is the prediction of gender for each of the ego-nodes.

Using all the distance measures reported in this work, as well as several baselines, we calculate the dissimilarities between all pairs of ego-networks for the three different data sets. Then for each data set and each ego-node we follow a leave-one-out classification approach, where we calculate the average similarity of it with the ego-networks of each of the labels. We assign the ego-node to the label that is on average most similar.

We compare the results of all the distance measures with a series of state-of-the-art and well-established methods of graph comparison. The authors of [3] study five similarity measures, the best of which is the Signature Similarity (SS), based on the SimHash algorithm. We also compare our results with DELTACON, an intuitive algorithm proposed by [4], that uses fast belief propagation to model the diffusion of information throughout the graph, and is able to capture both local and global structures. Finally, we compare with the graph edit distance (GED) and the edge-weight distance (DEW), two well established measures of graphs similarity, studied in detail in [1]. Mathematically, the graph edit distance (GED) between two graphs $G_\sigma = (V_\sigma, E_\sigma)$ and $G_\rho = (V_\rho, E_\rho)$ is given by:

$$GED = \frac{|V_\sigma| + |V_\rho| - 2|V_\sigma \cap V_\rho| + |E_\sigma| + |E_\rho| - 2|E_\sigma \cap E_\rho|}{|V_\sigma| + |V_\rho| + |E_\sigma| + |E_\rho|} \tag{13}$$

while the edge-weight distance for the two graphs, on the simplified case of $V_\rho = V_\sigma = V$ is defined in [1] as:

$$DEW = \sum_{u,v \in V} \frac{|w^\sigma_{uv} - w^\rho_{uv}|}{max(w^\sigma_{uv}, w^\rho_{uv})} \tag{14}$$

with $w^i_{uv}$ the weight of the edge between nodes $u$ and $v$ of graph $G_i$.

Using the above data, experimental setup and baselines, we predict the label of each ego-node, using only the similarity between the ego-networks. The results are reported in terms of the F1 score for all methods and all data sets in Table II and they indicate that most of the distance measures presented in this work are able to complete the classification task outperforming the more traditionally used baselines. With the exception of the Deezer data set, that appears to pose a difficult classification problem for all the methods, Hellinger Distance outperforms the rest, with $QJSDiv$ following closely.

*2) Graph Outlier Detection:* For this experiment, we utilize data from the Correlates of War Project [32], that track total national trade and bilateral trade flows on a global level from 1870-2014. The data can be represented in the form of 145 weighted networks, one for each year of data. Then each network will be of 207 nodes, one for each country (as defined in [33]), connected with each other with edges, whose weights represent the bilateral trade reported in the data.

The aim of this experiment is to identify outlier years, years of unexpected change in national trade. We compare the net-

| Dissimilarity | Dataset | | |
|---|---|---|---|
| | Email-EU | Butterfly | Deezer |
| $D_{trace}$ | 0.5560 | 0.7566 | 0.5211 |
| $D_{HS}$ | 0.4056 | 0.6950 | 0.5009 |
| $D_H$ | **0.6123** | **0.7806** | 0.5195 |
| $D_B$ | 0.5610 | 0.7678 | 0.4925 |
| QJSDiv | 0.5946 | 0.7742 | 0.5242 |
| GED | 0.5726 | 0.7554 | 0.5149 |
| DEW | 0.5728 | 0.7595 | 0.5158 |
| DELTACON | 0.4179 | 0.7433 | **0.5294** |
| SS | 0.3262 | 0.2716 | 0.5195 |

TABLE II
F1 SCORES FOR EACH OF THE CLASSIFICATION TASKS AND EACH OF THE METHODS APPLIED.

work structure of each year to the network of the year before and report their dissimilarity for all distances and the baselines. For each data point we use the $z_{score} = (d - median)/\sigma$ as a means to describe its relationship to the $median$ and standard deviation of the distribution of the dissimilarity values for each of the distance measures (the median is used instead of the mean, since the data do not follow normal distribution). Then, the years that exhibit a value that is smaller than $median + 3\sigma$ (have a $z_{score}$ smaller than 3) are considered non-anomalous and the years that are outside this limit are considered anomalous or outliers. The dissimilarity results are presented in Figure 4, where the outlier years, as identified by each of the methods, are also clearly marked.

In their majority, the years that are identified as anomalous, correspond to significant historical events that are expected to affect the international trade. All of the methods have identified the most prominent events, years 1914, 1919, 1920, 1930, 1939 - 1946 (inclusive), 1948. The year 1914 marks the start of WWI and the economic recession of 1913-1914. In 1919 and 1920 the Treaty of Versailles is held, the League of Nations is founded and Germany begins reparation payments. Then 1930 corresponds to the Great Depression, the years 1939 - 1945 are the years of WWII and 1946 is the year following the end of WWII with post war recovery and civil unrest in Europe. In 1948 the Cold War begins, the Marshall Plan takes place and the Arab-Israeli War began. Furthermore, all the quantum-inspired methods agree on a series of other outlier years:1918, 1955, 1960, and 1990. The year 1918 marks the end of WWI and is also associated with the Russian revolution and the dissolution of Austria-Hungary. The year 1955 is the start of the space race, 1960 marks the first phase of the Vietnam war, and 1990 is associated with the fall of the Berlin Wall (Nov. 1989), the fall of communism in the Eastern Europe, and the beginning of Gulf War.

One notable difference is that all the quantum-inspired distances appear to have a favorable signal-to-noise ratio, providing more confidence in their outlier reporting. The signal-to-noise ratio, defined as the ratio of the average value of dissimilarity for outlier years, over the average value of dissimilarity for non-outlier years, is reported on Table III. Another, more subtle difference, is seen in the confidence at which each method reports some of the outliers, empirically seen as the height of each point in the dissimilarity plots of Figure 4. While all methods recognise the importance of the

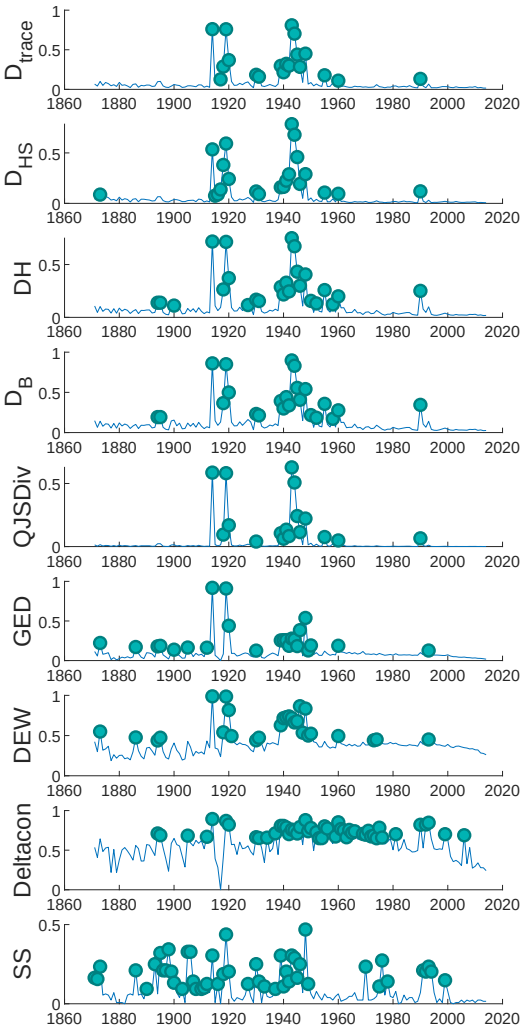


Fig. 4. The dissimilarity between the networks of each year y and the year before. The anomalous years identified by each method are also marked.

years of WWI and WWII, other events, such as the Great Depression of 1930 and the Gulf War of 1990, are reported with varying confidence. Using the value of $z_{score}$ for each data point, that represents the distance of the data point from the $median$ of the distribution in terms of standard deviations, we report in Table III for each method the confidence at which it identifies some of the prominent historical events. For both these years Hellinger and Bures distance appear to be more confident, while QJSDiv, trace and Hilbert-Schmidt are under-emphasising them.

Following the same rational, we can study the differences in other years,such as 1894, 1900 and 1952. The year 1894 marks the start of the first Sino-Japanese war and also follows the Panic of 1893, an economic recessions in the United States that began in the second half of 1893 and affected every sector of the economy. The 1900 has the Boxer Rebellion in which Russia, Japan, US, China, UK, France participated and finally, 1952 corresponds to the years of Korean War (1950-1953), the Sino-Japanese Peace Treaty is signed, the year that Elizabeth II becomes queen regnant, and the Great Smog of 1952 in London. The events of these years are significant enough to

| | $SNR$ | $z_{1930}$ | $z_{1990}$ | $z_{1894}$ | $z_{1900}$ | $z_{1952}$ |
|---|---|---|---|---|---|---|
| $D_{trace}$ | 8.21 | 6.87 | 4.36 | 2.65 | 0.92 | 1.24 |
| $D_{HS}$ | 10.95 | 5.91 | 6.00 | 2.64 | 0.89 | 1.04 |
| $D_H$ | 5.81 | 6.08 | 10.68 | 4.51 | 3.02 | 4.15 |
| $D_B$ | 5.62 | 5.44 | 9.46 | 4.06 | 2.74 | 3.76 |
| QJSDiv | 39.68 | 3.35 | 5.77 | 1.71 | 0.98 | 1.38 |
| GED | 4.24 | 3.16 | 1.56 | 6.61 | 3.91 | 1.49 |
| DEW | 1.77 | 3.14 | 2.44 | 3.03 | 1.45 | 1.63 |
| DELTACON | 1.55 | 3.51 | 9.35 | 5.27 | 2.95 | 5.87 |
| SS | 5.59 | 18.15 | 1.45 | 0.73 | 7.26 | 2.18 |

TABLE III
RESULTS OF THE SIGNAL-TO-NOISE RATIO AND THE CONFIDENCE IN OUTLIER DETECTION BASED ON THE $z_{score}$ FOR VARIOUS YEARS, FOR ALL GRAPH SIMILARITY MEASURES

have affected the international trade, however these years are not identified as outliers by all of the methods. Focusing on the quantum-inspired methods, years 1894 and 1952 are identified by Hellinger and Bures distance, while 1900 is only identified by Hellinger distance. These results emphasize the previously observed trend, that Hellinger distance is able to highlighting events of a wider rage of significance.

## IV. DISCUSSION

In this work, we introduce measures for graph similarity inspired by distances on a set of quantum states. We have shown that these measures can effectively distinguish graphs, and can be used with both weighted and unweighted networks, while identifying graph structure changes, such as the introduction of disconnected components. These measures intuitively capture and incorporate, several structural characteristics, that are often used to describe and compare networks, providing a holistic approach. We utilized real-world data sets to showcase that they can be effectively used on a variety of applications, domains, and problems outperforming previously well-established and state-of-the-art methods. One of the measures, Hellinger Distance, has proven to have the most consistent response and in most cases outperforms the rest. Two additional important features distinguish these methods from previously published approaches: they are well-established mathematical methods that incorporate the intrinsic structure of the entire network and have high interpretability.

## V. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Katherine Barbieri for useful discussions and insights on the international trade data.

## REFERENCES

[1] Bunke H, Dickinson PJ, Kraetzl M, Wallis WD (2007) A graph theoretic approach to enterprise network dynamics. Birkhäuser, Boston
[2] Bunke, Horst. "Graph matching: Theoretical foundations, algorithms, and applications." Proc. Vision Interface. Vol. 2000. 2000.
[3] Papadimitriou, Panagiotis, Ali Dasdan, and Hector Garcia-Molina. "Web graph similarity for anomaly detection." Journal of Internet Services and Applications 1.1 (2010): 19-30.
[4] Koutra, Danai, Joshua T. Vogelstein, and Christos Faloutsos. "Deltacon: A principled massive-graph similarity function." Proceedings of the 2013 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2013.
[5] Wills, Peter, and François G. Meyer. "Metrics for graph comparison: a practitioner's guide." PloS one 15.2 (2020): e0228728.
[6] De Domenico, Manlio, and Jacob Biamonte. "Spectral entropies as information-theoretic tools for complex network comparison." Physical Review X 6.4 (2016): 041062.
[7] Biamonte, Jacob, Mauro Faccin, and Manlio De Domenico. "Complex networks from classical to quantum." Communications Physics 2.1 (2019): 1-10.
[8] Javarone, Marco Alberto, and Giuliano Armano. "Quantum–classical transitions in complex networks." Journal of Statistical Mechanics: Theory and Experiment 2013.04 (2013): P04019.
[9] Faccin, Mauro, et al. "Community detection in quantum complex networks." Physical Review X 4.4 (2014): 041012.
[10] Kempe, Julia. "Quantum random walks: an introductory overview." Contemporary Physics 44.4 (2003): 307-327.
[11] Sánchez-Burillo, Eduardo, et al. "Quantum navigation and ranking in complex networks." Scientific reports 2.1 (2012): 1-8.
[12] Trávníček, Vojtěch, et al. "Experimental measurement of the hilbert-schmidt distance between two-qubit states as a means for reducing the complexity of machine learning." Physical review letters 123.26 (2019): 260501.
[13] Nielsen, Michael A., and Isaac Chuang. "Quantum computation and quantum information." (2002): 558-559.
[14] Dajka, Jerzy, Jerzy Łuczka, and Peter Hänggi. "Distance between quantum states in the presence of initial qubit-environment correlations: A comparative study." Physical Review A 84.3 (2011): 032120.
[15] Gilchrist, Alexei, Nathan K. Langford, and Michael A. Nielsen. "Distance measures to compare real and ideal quantum processes." Physical Review A 71.6 (2005): 062310.
[16] Lamberti, P. W., et al. "Metric character of the quantum Jensen-Shannon divergence." Physical Review A 77.5 (2008): 052311.
[17] Braunstein, Samuel L., Sibasish Ghosh, and Simone Severini. "The Laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states." Annals of Combinatorics 10.3 (2006): 291-317.
[18] Lee, Jinhyoung, M. S. Kim, and Časlav Brukner. "Operationally invariant measure of the distance between quantum states by complementary measurements." Physical review letters 91.8 (2003): 087902.
[19] Muthuganesan, R., V. K. Chandrasekar, and R. Sankaranarayanan. "Quantum coherence and correlation measures based on affinity." arXiv preprint arXiv:2003.13077 (2020).
[20] Majtey, A. P., P. W. Lamberti, and D. P. Prato. "Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states." Physical Review A 72.5 (2005): 052310.
[21] Roy, Matthieu, Stefan Schmid, and Gilles Tredan. "Modeling and measuring graph similarity: The case for centrality distance." Proceedings of the 10th ACM international workshop on Foundations of mobile computing. 2014.
[22] Newman, Mark. Networks. Oxford university press, 2018.
[23] Hage, Per, and Frank Harary. "Eccentricity and centrality in networks." Social networks 17.1 (1995): 57-63.
[24] Lorrain, Francois, and Harrison C. White. "Structural equivalence of individuals in social networks." The Journal of mathematical sociology 1.1 (1971): 49-80.
[25] Adamic, Lada A., and Eytan Adar. "Friends and neighbors on the web." Social networks 25.3 (2003): 211-230.
[26] Jaccard P., Bulletin de la Societe Vaudoise des Science Naturelles 37, 5471901
[27] Passerini, Filippo, and Simone Severini. "The von Neumann entropy of networks." arXiv preprint arXiv:0812.2597 (2008).
[28] Leskovec J. and Krevl A., SNAP Datasets: Stanford Large Network Dataset Collection, http://snap.stanford.edu/data
[29] J. Leskovec, J. Kleinberg and C. Faloutsos. Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data (ACM TKDD), 1(1), 2007.
[30] Wang, Bo, et al. "Network enhancement as a general method to denoise weighted biological networks." Nature communications 9.1 (2018): 1-8.
[31] Rozemberczki, Benedek, Oliver Kiss, and Rik Sarkar. "An API Oriented Open-source Python Framework for Unsupervised Learning on Graphs." arXiv preprint arXiv:2003.04819 (2020).
[32] Barbieri, Katherine and Omar M. G. Omar Keshk. 2016. Correlates of War Project Trade Data Set Codebook, Version 4.0. Online: http://correlatesofwar.org.
[33] Barbieri, Katherine, Omar M. G. Keshk, and Brian Pollins. 2009. "TRADING DATA: Evaluating our Assumptions and Coding Rules." Conflict Management and Peace Science. 26(5): 471-491.