



Quantum-inspired measures of network distinguishability

Athanasia Polychronopoulou¹ · Jumanah Alshehri¹ · Zoran Obradovic^{1,2}

Received: 27 June 2022 / Revised: 13 March 2023 / Accepted: 15 March 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2023

Abstract

The recent emergence of information physics as a theoretical foundation for complex networks has inspired the utilization of measures, initially developed for use with quantum mechanical systems, for the solution of graph theory research problems. Network comparison is one such research problem that arises often in all domains, when entities that interact with each other possibly with more than one discrete interaction types are studied. A network similarity measure is required for any data mining application on graphs, such as graph clustering, classification, or outlier detection. A natural starting point for the identification of such a network similarity measure is information physics, offering a series of measures typically used to quantify the distance of quantum states. These quantum-inspired methods satisfy the mathematical requirements for graph similarity while offering high interpretability. In this work, we characterize these measures for use with monoplex and multiplex networks on experiments with synthetic data, and we report results on real-world applications to compare with a series of state-of-the-art and well-established methods of graph distinguishability.

Keywords Graph classification · Graph similarity · Information physics · Multiplex comparison

1 Introduction

Complex networks have become a natural representation of entities and their interactions in systems of all domains. They can be used to describe social networks, genetic and protein interaction networks, airline and road traffic networks, brain connectivity networks, web graphs as well as any other system composed of entities, represented as graph nodes, interacting with each other thereby forming the edges of the graph. In many applications, the same set of nodes may participate in multiple different types of interactions each forming a different and independent layer of

the network. Then, the system is represented as a multiplex graph, and the term monoplex can be used for single layer networks. Alternative terminology has been used over the years for such complex networks, with examples including multidimensional networks, multinetworks, multirelational networks, or multislice networks. More general representations are used in cases where layers are ordered or otherwise correlated, or in cases where the nodes across layers are not aligned, or even represent entities of different types with inter-layer connections. The resulting networks are frequently called multilayer networks, but other terms such as multilevel, multitype, or networks of networks are used, as discussed in detail in Kivelä et al. (2014).

A variety of network theoretic tools are used for the analysis of these complex systems (Bunke et al. 2007), and the question of graph distinguishability or graph similarity is often a central aspect of study. Examples include analysis of networks for classification purposes, such as comparison of networks created by brain fMRI images of patients, but also analysis of networks for clustering or outlier detection purposes, such as comparisons of daily traffic networks to facilitate the detection of abnormal change in traffic patterns. 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 essential.

✉ Athanasia Polychronopoulou
n.polychr@temple.edu

Jumanah Alshehri
shehri.j@temple.edu

Zoran Obradovic
zoran.obradovic@temple.edu

¹ Computer and Information Sciences Department, Temple University, 1925 N. 12th Street, Philadelphia, PA 19122, USA

² Center for Data Analytics and Biomedical Informatics, Temple University, 1925 N. 12th Street, Philadelphia, PA 19122, USA

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 (Bunke 2000).

Several approaches have been proposed to solve different variations of the problem of calculating graph similarity. The most straightforward 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. The 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 (Bunke et al. 2007; Wilson and Zhu 2008), or measures inspired by document similarity (Papadimitriou et al. 2010), and other intuitive approaches (Koutra et al. 2013; Monnig and Meyer 2018). More recently, spectral distances, as well as distances based on node affinities, have been studied in more depth (Wills et al. 2020); the usage of Quantum Jensen Divergence (QJSDiv) as a measure of graph distinguishability has been proposed (Domenico and Biamonte 2016); and recently a series of measures that have been effectively used in quantum mechanics and quantum information theory were proposed (Polychronopoulou et al. 2021).

Quantum mechanics has proven to be a valuable resource for the investigation of the behavior of complex networks (Biamonte et al. 2019). 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 (Javarone et al. 2013). Quantum transport probability and state fidelity have been implemented as a closeness function and used for community detection (Mauro et al. 2014), while quantum random walks have been shown to offer significant performance improvements on traditional computer science algorithms with respect to the classical random walks (Kempe 2003), and have been implemented for a quantum-inspired ranking algorithm (Sánchez-Burillo et al. 2018). 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 (Trávníšek et al. 2019). Given this interconnection between quantum mechanics and complex networks, quantum network theory is a natural place to look for similarity measures.

In the work of Polychronopoulou et al. (2021), four measures that have been broadly used in quantum information theory for the comparison of quantum states (Michael and Chuang 2002), systems, and processes (Jerzy et al. 2011; Gilchrist et al. 2005) were proposed for use with monoplex networks. These quantum-inspired methods offer a mathematical method of graph comparison, they satisfy the mathematical properties of a metric, offer intelligible results with high interpretability, and were shown in Polychronopoulou et al. (2021) to satisfy intuitive graph similarity properties for monoplex networks. In this work, the techniques are further extended for use with multiplex networks. An analysis of the intuitive requirements of multiplex graph comparison is presented, and we elucidate the properties of the four distance measures on both monoplex and multiplex networks. A series of artificial data is utilized to characterize their effectiveness, and showcase their behaviour. We then report results on real-world applications, answering different data mining problems from a variety of domains. The quantum-inspired measures are compared to a series of state-of-the-art and well established methods, adding in this work new baseline techniques based on the most recent publications in the domain. The code is available on GitHub (https://github.com/NancyPol/Quantum_Graph_Distances).

2 Methods

Mathematically, the structure of an undirected single layer network, or monoplex, $G = \{V, E\}$, on a set of N vertices $V = \{1, 2, \dots, n\}$ and a set of edges $E \subseteq V \times V$, is represented by a binary matrix A , known as adjacency matrix, whose element $A_{u,v} = 1$ if $u, v \in V$ and $A_{u,v} = 0$ otherwise. In order to extend distance measures that were initially developed for quantum states, to be used with graphs, we utilise 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 ρ of a graph can be defined through the combinatorial Laplacian of the graph (Braunstein et al. 2006), as $\rho_G = \frac{\Delta - A}{\text{Tr}(\Delta)}$. In this equation A is the adjacency matrix, Δ is the degree matrix, which is a diagonal matrix with elements equal to the degree $d(u)$ of each node u . The normalization using the trace of Δ 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 $\rho_G = \frac{\Delta - W}{\text{Tr}(\Delta)}$, where W is the weights matrix with $W_{u,v} = w_{uv}$ and $W_{u,v} = 0$ if the nodes u and v are not connected. The degree matrix Δ 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 the following sections we present and evaluate measures of quantum states comparison that are particularly important in Quantum Information Science (Michael and Chuang 2002; Jerzy et al. 2011; Gilchrist et al. 2005). They have been shown to be metrics, are mathematically well established and have high interpretability; thereby satisfying the mathematical requirements of distance measures for graphs. The metric character of a distance for the comparison of two complex networks G_σ and G_ρ , guarantees 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 $D(G_\sigma, G_\rho) \leq D(G_\sigma, G_\tau) + D(G_\tau, G_\rho)$.

2.1 Trace distance

The trace distance between two quantum states, or two networks, with density matrices σ and ρ is given by:

$$d_{\text{trace}}(G_\sigma, G_\rho) = \frac{1}{2}[\text{trace}(\sqrt{(\sigma - \rho)^2})] \quad (1)$$

It is a metric and is bounded to be $0 \leq D_{\text{trace}} \leq 1$ with the equality to 0 holding if and only if $\rho = \sigma$ and the equality to 1 holding if and only if ρ and σ 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.

2.2 Hilbert–Schmidt distance

The Hilbert-Schmidt distance

$$d_{\text{HS}}(G_\sigma, G_\rho) = \sqrt{\text{trace}(\sigma - \rho)^2} \quad (2)$$

is a Riemannian metric that is bounded to be $0 \leq D_{\text{HS}} \leq 2D_{\text{trace}}$. As it is defined on the space of operators it is unclear how to impose an operational interpretation, however, in Jinhyoung and Brukner (2003) the authors suggest that it can be seen as an information distance between two quantum states. It has been recently used in Trávníšek et al. (2019) 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.

2.3 Hellinger distance

The Hellinger distance is given by:

$$d_H(G_\sigma, G_\rho) = \sqrt{2[1 - A(\sigma, \rho)]} \quad (3)$$

with

$$A(\sigma, \rho) = \text{trace}(\sqrt{\sigma} \sqrt{\rho}) \quad (4)$$

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 $0 \leq D_H \leq \sqrt{2}$ (Muthuganesan et al. 2020).

2.4 Bures distance

The Bures distance is expressed through quantum Fidelity (Jozsa 1994), which is a measures of overlap between two quantum states:

$$F(\sigma, \rho) = [\text{trace}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})]^2 \quad (5)$$

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:

$$d_B(G_\sigma, G_\rho) = \sqrt{2[1 - \sqrt{F(\sigma, \rho)}]} \quad (6)$$

2.5 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:

$$S(\rho) = -\text{Tr}(\rho \log_2 \rho) \quad (7)$$

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

$$S_N(\sigma, \rho) = \text{Tr}(\sigma(\ln \sigma - \ln \rho)) \quad (8)$$

where $\ln \sigma$ is the matrix base 2 logarithm of σ . 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 σ is a subset of the support of ρ (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 ($QJSDiv$) (Majtey et al. 2005), given by:

$$QJSDiv(G_\sigma, G_\rho) = \frac{1}{2} [S_N(\sigma, \frac{\sigma + \rho}{2}) + S_N(\rho, \frac{\sigma + \rho}{2})] \quad (9)$$

where S_N is the quantum relative entropy of equation 8. Quantum Jensen-Shannon divergence was introduced as a measure of distinguishability between mixed quantum states (Majtey et al. 2005; Lamberti 2008). It is 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 ρ and σ now lifted. The authors of Domenico and Biamonte (2016) 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 multi-layer system. In this work, we use Quantum Jensen-Shannon as a quantum-inspired baseline and aim to compare its performance with the rest of the measures presented here.

2.6 Comparison of multiplex networks

Mathematically, a Multiplex Network G with L layers can be seen as a collection of L single-layer networks: $G = \{G_l \mid l \in \{1, \dots, L\}\}$. Each of these networks has a set of edges, that can either be weighted or not, and they all share the same set of N nodes V , while some nodes may not have connections in all layers. Each of the layers is represented by a weighted or non-weighted adjacency matrix A_l . A natural representation of the entire multiplex uses the supra-adjacency matrix $\mathcal{A}$ of the network (Kivelä et al. 2014; Stefano et al. 2014; Domenico et al. 2013). $\mathcal{A}$ is obtained from the intra-layer adjacency matrices A_l and a coupling matrix K , whose elements K_{lm} represent the inter-layer connections between layers l and m . Then the supra-adjacency matrix becomes $\mathcal{A} = \bigoplus_l A_l + K \otimes I_N$, with a term representing the direct sum of A_l and I_N the $N \times N$ identity matrix. Notice that, in multiplex networks, the only possible type of inter-layer connection is the one in which a given node is connected to its counterpart node in the remaining layers, resulting in a coupling matrix K whose non-diagonal elements are typically equal to 1.

This formulation allows multiplex networks to be compared using the distance measures introduced for single layer graphs, with the supra-adjacency matrix substituting the monoplex adjacency. However, as the size of the networks under study grow, either in terms of nodes or in terms of layers, so does the size of the supra-adjacency, rendering the use of complex algorithms computationally expensive and infeasible. In addition, the complexity of the representation

is not necessarily increasing the informative power. As the addition of layers increases the effective dimension of the model, the intrinsic information becomes more sparse, and may remain hidden within the noise of the representation (Tiago 2015).

Therefore, alternative representations have been proposed for multiplex networks, typically constructing a monoplex network by aggregating the various layers. An example of such an aggregation is the average network (Albert et al. 2013), where the aggregated monoplex is a simple average of the layers. Similarly, weighted or non-weighted overlapping networks (Battiston et al. 2014) have been constructed by summarising the various layers. The authors of Domenico et al. (2013) have used the projected monoplex, and the overlay monoplex networks. Both are obtained from a multiplex network by summing the edges over all layers, but the overlay network ignores the interlayer connections of nodes. Finally, the authors of Sánchez-García et al. (2014) introduced the notion of graph quotients and an aggregate network that summarizes the connections of nodes among the different layers, normalised by the multiplexity degrees of the nodes.

In this work we introduce an alternative representation, that can be seen as a very intuitive approach to layer aggregation, the noisy-OR function. In this case, the weight of the edge between nodes u and v of the aggregate monoplex is calculated as: $1 - \prod_{l=1}^L (1 - w_{u,v}^l)$. This function has been used in many domains and real-world applications for approximating relationships in which there is a set of two or more sources of information that can potentially explain a single variable (Sampath 1993). If the monoplex aggregate is seen as the network describing the absolute connection between nodes, and the layers are seen as evidence of this connection, then the extension for use with multiplex networks is natural. Alternatively, the aggregate monoplex describes the reliability of connections between nodes, as affected by multiple sources of information, the layers of the multiplex network. In the case of non-weighted networks, this function acts as a projection to a flattened network, but in the case of weighted networks, the final weight of an edge between two nodes is affected by the weights of the various layers.

The monoplex aggregations presented above can be used as adjacency matrices for the representation of the multiplex, and networks can be compared using the corresponding density matrices. In the case of Projected and Quotient monoplex, that take into account interlayer connections of nodes, the degree matrix should also include those, so that the trace of the density matrix will remain equal to 1. In the following sections, we utilize a series of experiments on artificial multiplex data to quantify the ability of these multiplex representations and similarity measures to distinguish between graphs, and capture

intuitive similarity aspects. Then we use real-world complex applications to evaluate the performance of alternative methods of graph comparison.

3 Evaluation on artificial data

3.1 Monoplex networks

A similarity measure needs to satisfy several requirements in order to be considered as a means of single layer network comparison. First, the measure should satisfy mathematical requirements (be a metric), and have high interpretability. But some more important challenges are graph-specific and are associated with the ability of the measure to capture changes in important structural characteristics such different kinds of centrality measures. Examples include the fact that a modification of a graph is more important in smaller graphs than larger ones, also that targeted modifications should be given higher importance, and that edge weights should be considered.

The performance of the proposed measures is evaluated on artificial data, created as Erdős-Rényi random graphs with given number of nodes (100) and edges (1000). These artificial networks are then continuously modified using a discrete-time process in which one elementary graph edit operation is applied at every step. The distance between the modified and original network is calculated for every step and the results are summarised in Fig. 1. More detailed results, accommodating all possible elementary graph edit operations, and different graphical structures are included in Polychronopoulou et al. (2021).

In the simple case of repeated node removals, the monotonically increasing behavior of the distance measures is evident and confirms that they all act as measures of graph distinguishability. Several other intuitive aspects are evident on this figure. First, with the exception of Bures distance, the measures have identified that the removal of one node is more crucial for smaller networks, as is evident by the slope of the curves. Interestingly, Hellinger distance is equally sensitive to modifications of large and small networks. Furthermore, all distances are sensitive to distinguishing targeted from random operations, both for the case of continuous edge removals targeting popular nodes, and continuous edge removals targeting edges of higher weights. The sensitivity of the methods differs, although they all exhibit distance values that are lower for random modification, compared to targeted ones. The behavior of trace distance is notable, with a lower sensitivity when the compared graphs are either very similar or very dissimilar.

3.2 Multiplex networks

A similar analysis is applied to the case of multiplex networks. Different network generation techniques are used to create multiplex networks with two or three layers following Poisson multidegree distribution and scale-free multidegree distribution (Bianconi 2013), as well as duplex networks with linear (Vincenzo et al. 2013) and nonlinear kernels (Vincenzo et al. 2014). Each network is then continuously modified, applying at each step one elementary graph edit operation, and the distance between the modified and original network is calculated. The process is repeated for all multiplex representations.

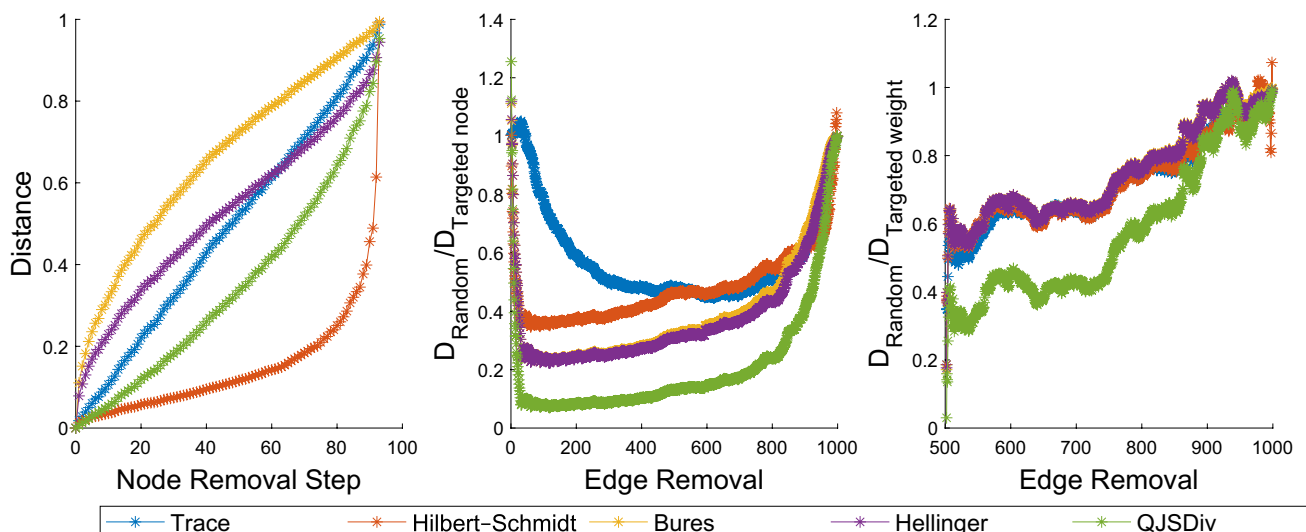


Fig. 1 The progression of each of the distance measures between the original graph and the graph after applying a series of elementary graph edit operations. For edge removals, two operations are com-

pared: random vs targeted edge removals, targeting the popular nodes (middle) or edges of higher weight (right)

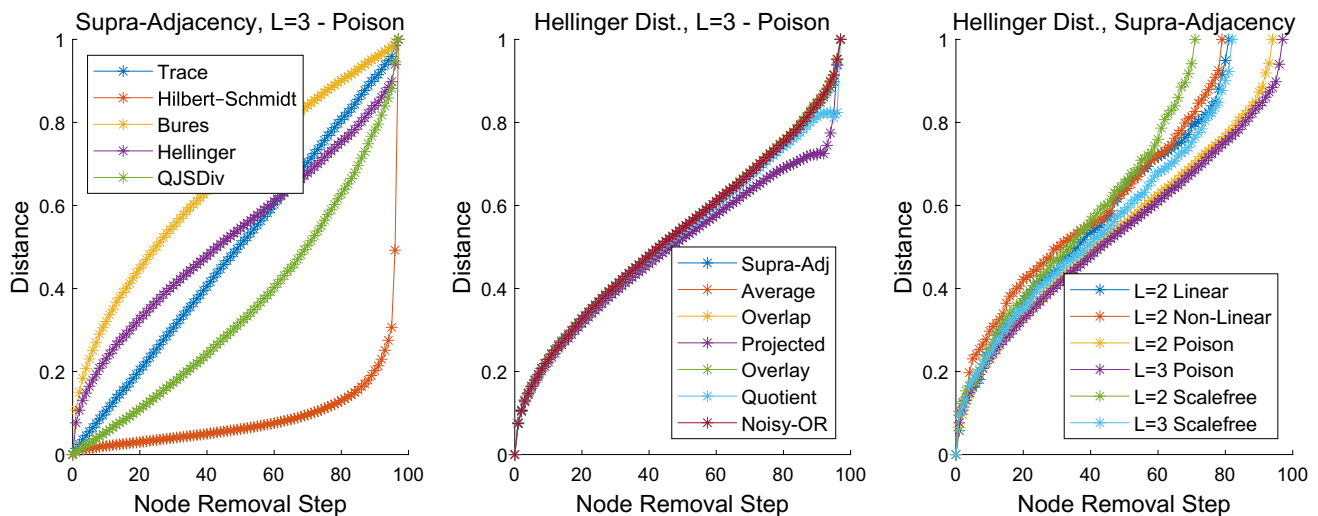


Fig. 2 The progression of the distance measures between the original graph and the graph after applying a series of elementary graph edit operations for: Left: supra-adjacency representation, 3-layer poisson multidegree distribution network, and various distance measures.

Middle: hellinger distance, 3-layer poisson multidegree distribution network, and various multiplex representations. Right: hellinger distance, supra-adjacency representation, and various multiplex artificial networks of L number of layers and different generation techniques

The first results are created by applying continuous random node removals and are presented in Fig. 2. In each part of the plot, two of the three variables of study (distance, multiplex representation, multiplex generation technique) remain constant, while the third one varies. Similarly to the case of monoplex networks, quantum distances satisfy intuitive similarity aspects: they continuously increase as the networks are further modified and this applies to most multiplex representations and all types of data. As observed in the middle plot of Fig. 2, in the case of Quotient and Projected monoplex networks that incorporate inter layer connections of nodes, this is not always the case. The inter layer connections become prevalent at very small size networks

affecting the expected continuous increase of distance in the plots of Fig. 2.

In the case of multiplex networks the measures used for their comparison should be able to capture some additional intuitive aspects. For example, removal of nodes with higher centrality should create graphs more distinguishable from the original one. Even more so, when the nodes that are removed act as a bridge between layers (Albert et al. 2014), such as node n_3 in Fig. 3. Similarly, the effect of an edge removal should depend on whether the adjacent nodes have a connection in other layers or not, such as the case with edge e_2 in Fig. 3.

In this small scale example, we can calculate the distances of some possible modifications using the supra-adjacency

Fig. 3 Examples of multiplex networks, where each layer is represented by a different color of edges. The removal of each of the nodes n_1, n_2 or n_3 should create graphs with different distances from the original one. The same applies to the removal of each of single layer edges e_1 and e_2

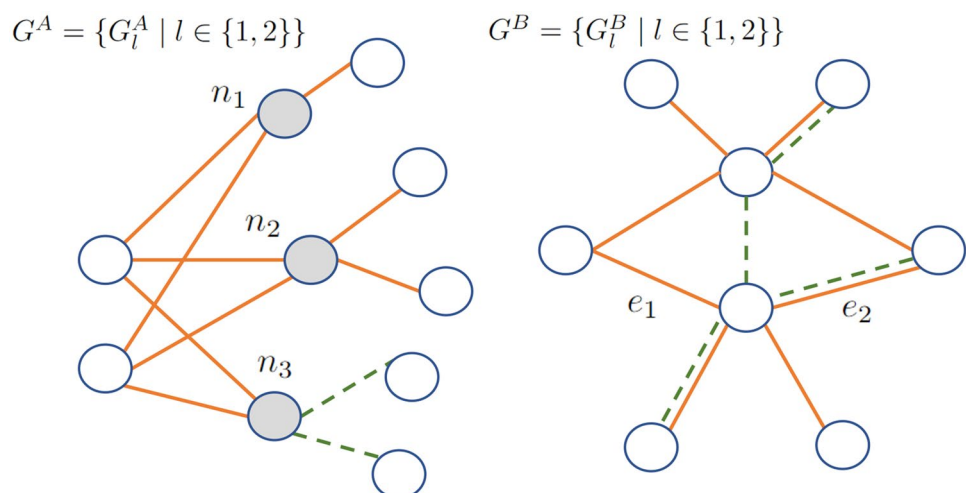


Table 1 Distances between different sample multiplex networks G and G'

	GED	Hellinger	Bures	Hilbert-Schmidt	Trace	QJSDiv
$G = G^A$						
$G' = G - \{n_1\}$:	0.1232	0.3247	0.4430	0.1535	0.2123	0.1100
$G' = G - \{n_2\}$:	0.1884	0.4004	0.5402	0.1999	0.2992	0.1680
$G' = G - \{n_3\}$:	0.1884	0.4006	0.5401	0.1912	0.3004	0.1683
$G = G^B$						
$G' = G - \{e_1\}$:	0.0140	0.0653	0.0880	0.0485	0.046189	0.0063
$G' = G - \{e_2\}$:	0.0140	0.0620	0.0849	0.0485	0.046187	0.0056

representation of the multiplex, and directly compare the results of Table 1. In the cases of node removals from graph G^A , when node n_1 is removed all distance measures calculate a distance that is smaller than when nodes n_2 and n_3 are removed. Then, comparing the effect of removing nodes n_2 and n_3 , we can see that GED is the same in both cases, as expected since it only incorporates structural changes, while Hellinger, Trace and QJSDiv, report that the removal of the bridge node is more substantial. In the cases of edge removals, all quantum measures, except Hilbert-Schmidt, agree that the removal of edge e_1 increases the distinguishability of the graph more than the removal of edge e_2 , since in that second case the two adjacent nodes maintain their connection over the other layer.

Similar conclusions can be reached using alternative representations for the multiplex. Table 2 shows the values of Hellinger Distance for the same modifications and different aggregated monoplex representations. All representation identify the removal of node n_1 as less substantial. However, with the exception of the projected and quotient monoplex, they are not able to differentiate between the removals of nodes n_2 and n_3 . In the case of edge removals most aggregate representations agree that the distance of graphs G and G' is larger when edge e_1 is removed, while since this is a non-weighted network, the noisy-OR function does not distinguish between G and G' when e_2 is removed.

We can generalise these findings, using larger scale artificial networks, generated as two and three layer artificial multiplex networks, with the multiplex generation techniques reported above. Then, at every node removal step, the distance of the new graph with the previous, and

the MultiRank centrality of the removed node are calculated. For this centrality, an algorithm by Christoph et al. (2018) is used, as it considers the multiplex structure and assigns higher centrality values to nodes that are linked to central nodes in highly influential layers. There is a clear correlation between these values, as evident by the Pearson correlation of the distances and node centralities, reported for the supra-adjacency representation and each of the distance measures, in Table 3. Other multiplex representations produce similar results, and an example for the correlations of Hellinger distance is presented in the same table for all other multiplex representations.

We can generalise the findings of the edge removal examples, using again larger scale artificial networks, while for every removal of an edge $e_{u,v}^l$ from a layer l , we monitor the number of layers in which an edge between the adjacent nodes u and v exists. The resulting distributions of distance values, across all types of networks and for the case of the supra-adjacency representation and Hellinger distance is reported in Fig. 4. The three plots, when normalised to the same integral, present distributions of distances that can be fitted with functions holding different exponential parameters, shown on the plots. These values have 95% Confidence Intervals given in this example by: $CI_b^{1L} = (-0.1573, -0.1303)$, $CI_b^{2L} = (-0.3306, -0.2752)$, $CI_b^{3L} = (-0.6758, -0.4288)$, indicating that the removal of an edge is creating less distinguishable graphs, when the adjacent nodes are also connected in different layers. It is also evident that the larger the number of layers in which the adjacent nodes u and v were originally connected, the more narrow the distribution is, or the smaller

Table 2 Hellinger distance between sample multiplex networks G and G' , for different aggregate monoplex representations

	Supra	Average	Overlap	Projected	Overlay	Quotient	Noisy-OR
$G = G^A$							
$G' = G - \{n_1\}$:	0.3247	0.3295	0.3356	0.3210	0.3295	0.3169	0.3295
$G' = G - \{n_2\}$:	0.4004	0.4017	0.4134	0.3905	0.4017	0.3868	0.4017
$G' = G - \{n_3\}$:	0.4006	0.4017	0.4134	0.3975	0.4017	0.4223	0.4017
$G = G^B$							
$G' = G - \{e_1\}$:	0.0653	0.0839	0.0509	0.0762	0.0839	0.0635	0.0994
$G' = G - \{e_2\}$:	0.0620	0.0593	0.0258	0.0539	0.0593	0.0387	0

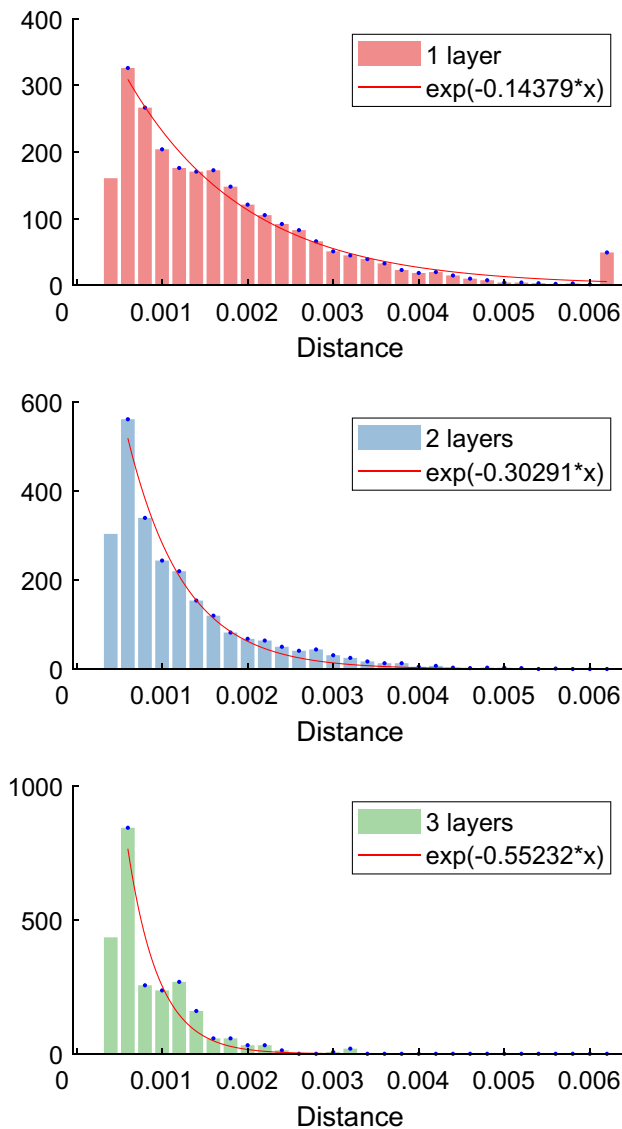


Fig. 4 Distribution of the values of Hellinger distance between G and G' when an edge $e_{u,v}^l$ from layer l is removed, while the adjacent nodes u and v are connected in 1, 2 or 3 layers

the expected distance value will be. Similar trends are observed on the plots of all other quantum distances, as well as the majority of multiplex representations, with the exception of overlap and Noisy-OR representations similar to the results of the small scale example.

4 Applications on real data

In this section, we evaluate the effectiveness of the proposed distance measures on real world applications. For each application we utilise data from different domains and we compare our results with several well-established and state of the art baseline methods.

4.1 Baselines

The first of our baselines is based on the Normalized Laplacian spectral distance of two graphs $G_\sigma = (V_\sigma, E_\sigma)$ and $G_\rho = (V_\rho, E_\rho)$, as calculated by:

$$d_{\mathcal{L}}(G_\sigma, G_\rho) = \sqrt{\sum_{i=1}^N (\lambda_i^\sigma - \lambda_i^\rho)^2} \quad (10)$$

with λ_i^σ and λ_i^ρ the i^{th} eigenvalue of the Normalised Laplacian of graphs G_σ and G_ρ correspondingly. The spectral distances of graphs have been studied (Wilson and Zhu 2008; Wills et al. 2020) and proven to be effective means of graph comparison, although they are not true metrics, since two graphs may exist that are co-spectral but not isomorphic.

We also compare our results with the resistance-perturbation distance, introduced in Monnig and Meyer (2018). In this work the authors represent each graph G by its effective graph resistance R , with elements: $R_{ij} = L_{ii}^\dagger + L_{jj}^\dagger - 2L_{ij}^\dagger$, where $L^\dagger$ denotes the Moore-Penrose pseudoinverse of the combinatorial Laplacian matrix L . Then the resistance-perturbation distance between two graphs $G_\sigma = (V_\sigma, E_\sigma)$ and $G_\rho = (V_\rho, E_\rho)$ is defined as the element-wise p-norm of the difference between their effective resistance matrices:

$$d_{PR}(G_\sigma, G_\rho) = \|R_\sigma - R_\rho\|_p \quad (11)$$

Table 3 Pearson Correlation of node centrality and Distance between G and G'

Supra adjacency representation						
	Hellinger	Bures	H-S	Trace	QJSDiv	
	0.974	0.969	0.967	0.983	0.967	
Hellinger Distance						
Supra.	Avg	Overlap	Proj.	Ovy	Quot.	OR
0.974	0.973	0.974	0.973	0.973	0.966	0.974

The last of our state-of-the-art baselines is DELTACON, an intuitive algorithm proposed by Koutra et al. (2013), that uses fast belief propagation to model the diffusion of information throughout the graph and represent each graph by an affinity matrix S , defined as: $S = [I + \epsilon^2 D + A]^{-1}$. Then the two graphs $G_\sigma = (V_\sigma, E_\sigma)$ and $G_\rho = (V_\rho, E_\rho)$ are compared via the Matusita difference of their representations:

$$d_{DELTA}(G_\sigma, G_\rho) = \sqrt{\sum_{ij} (\sqrt{S_{ij}^\sigma} - \sqrt{S_{ij}^\rho})^2} \quad (12)$$

We also compare with two well-established measures of graphs similarity, the graph edit distance (GED) and the edge-weight distance (DEW), studied in detail in Bunke et al. (2007). 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:

$$d_{GED}(G_\sigma, G_\rho) = \frac{|V_\sigma| + |V_\rho| - 2|V_\sigma \cap V_\rho| + |E_\sigma| + |E_\rho| - 2|E_\sigma \cap E_\rho|}{|V_\sigma| + |V_\rho| + |E_\sigma| + |E_\rho|} \quad (13)$$

while the edge-weight distance for the two graphs, on the simplified case of $V_\sigma = V_\rho = V$ is defined in Bunke et al. (2007) as:

$$d_{DEW}(G_\sigma, G_\rho) = \sum_{u,v \in V} \frac{|w_{uv}^\sigma - w_{uv}^\rho|}{\max(w_{uv}^\sigma, w_{uv}^\rho)} \quad (14)$$

with w_{uv}^i the weight of the edge between nodes u and v of graph G_i .

4.2 Single layer graph classification

As a first application, we evaluate the effectiveness of the distance measures on classification problems with real world single layer networks. The data sets used, are available

online, under (SNAP datasets), or (Austin datasets) and the results are reported in terms of the F1 score in Table 4.

For each data set, we construct labeled graphs using the ego-networks of each node. For the classification task, the dissimilarities between all possible combinations of graphs are calculated, a leave-one-out approach is implemented and each ego-network is assigned to the label that is on average most similar. The classification is based on the assumptions that researcher who belong in the same department will have similar email ego-networks (Leskovec et al. 2007), butterflies who belong in the same class will have similar visual similarities ego-networks (Bo et al. 2018), individuals from the same class have similar interaction ego-networks (Mastrandrea et al. 2015; Juliette et al. 2011; Philip et al. 2021), US Counties with a majority vote of the same party in the U.S. presidential elections will have similar Facebook social connectedness ego-networks, or physical proximity ego-networks (Jia and Benson 2022), and that Congresspersons with the same political party affiliation have similar bill co-sponsoring egonetworks (James 2006; Philip et al. 2021).

The results indicate that the quantum-inspired distance measures in most cases outperform the more traditionally used baselines, with Hellinger achieving the highest classification accuracy most often.

4.3 Multiplex layer clustering

As described earlier, in multiplex networks entities are connected to each other via multiple types of connections. The increased complexity of such systems has motivated the study of layer aggregation (Manlio et al. 2015) or layer clustering. Through this process it is possible to minimize the total number of layers in a system, resulting in a network that can be analysed more efficiently, but not without some loss of information content. Although the benefit of aggregating layers with structural similarities is not always clear,

Table 4 F1 scores for each of the classification tasks and distinguishability methods

	Dataset						
	Email-EU	Butterfly	High-School	Primary	USC-FB	USC	Senate
d_{trace}	0.5560	0.7566	0.9786	0.9051	0.8016	0.8218	0.5691
d_{HS}	0.4056	0.6950	0.9900	0.9173	0.8212	0.8007	0.5691
d_H	0.6123	0.7806	0.9869	0.9173	0.8065	0.8168	0.5691
d_B	0.5610	0.7678	0.9898	0.9025	0.4615	0.4615	0.5051
d_{qjs}	0.5946	0.7742	0.9869	0.9173	0.8062	0.7658	0.5481
d_{GED}	0.5726	0.7554	0.9931	0.9024	0.8062	0.8268	0.5582
d_{DEW}	0.5728	0.7595	0.9931	0.8986	0.8062	0.8168	0.5617
d_{DELTA}	0.4179	0.7433	0.9287	0.9168	0.6621	0.7924	0.5720
d_C	0.0471	0.2119	0.2097	0.1667	0.6135	0.4328	0.5444
d_{PR}	0.0463	0.7265	0.9297	0.7117	0.8034	0.7658	0.5704

Table 5 Evaluation of layer hierarchical clustering in terms silhouette scores

Data	N	L	L_G	d_H	d_B	d_{trace}	d_{HS}	d_{qjs}	d_{GED}	d_{DEW}	d_C	d_{DELT}	d_{PR}
FAO	214	364	250	0.575	0.575	0.589	0.582	0.580	0.569	0.570	0.381	0.527	0.514
EUAir	450	37	25	0.456	0.456	0.453	0.479	0.472	0.470	0.364	0.415	0.616	0.472
Auger	514	16	10	0.435	0.435	0.412	0.492	0.435	0.437	0.437	0.256	0.482	0.469
candida	367	7	4	0.555	0.414	0.555	0.492	0.555	0.555	0.555	0.461	0.489	0.437
London	369	6	4	1	1	1	1	1	1	1	1	1	1
Rattus	2640	6	4	0.543	0.376	0.376	0.543	0.376	0.543	0.377	0.494	0.465	0.494
Gallus	313	6	4	0.435	0.435	0.469	0.525	0.469	0.469	0.469	0.257	0.538	0.534
XG	461	5	4	0.596	0.596	0.631	0.631	0.631	0.631	0.631	0.592	0.592	0.592
HIV1	1005	5	4	0.656	0.656	0.656	0.534	0.656	0.657	0.656	0.657	0.657	0.558
D.Rerio	155	5	3	0.281	0.281	0.281	0.448	0.281	0.281	0.281	0.419	0.417	0.345
Arhus	61	5	3	0.561	0.561	0.561	0.391	0.561	0.561	0.561	0.270	0.292	0.266
KT	39	4	3	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.546	0.547
herpes	216	4	3	0.619	0.619	0.619	0.394	0.619	0.619	0.619	0.619	0.619	0.619
Bos	325	4	3	0.478	0.545	0.478	0.545	0.478	0.545	0.545	0.478	0.478	0.478
Hepc	105	3	2	0.511	0.511	0.511	0.511	0.511	0.511	0.511	0.511	0.268	0.268
KHigh	21	3	2	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.583	0.261	0.583
LLaw	71	3	2	0.608	0.608	0.608	0.608	0.608	0.608	0.608	0.608	0.608	0.608
VC7th	29	3	2	0.562	0.562	0.562	0.562	0.562	0.562	0.562	0.562	0.562	0.562

it typically depends on the system under study (Tiago 2015), and remains an open research problem.

Following this motivation, we evaluate the effectiveness of the quantum-inspired graph distance measures in the problem of computing pairwise similarities between layers, and then performing hierarchical clustering of structurally similar layers in multiplex networks. We utilise real-world data from various domains (Manlio Datasets), and the result are presented in Table 5. The same table includes the number of nodes N for each data set, the original number of layers L , and the number of layers after their hierarchical grouping L_G .

For the evaluation of the clusters created based on each of the dissimilarities of layers, the silhouette is utilized, which calculates how similar the data are within the cluster, as compared to other clusters. More specifically, for each layer, we calculate a silhouette score given by: $S_l = (Dist_{out} - Dist_{in}) / (\max(Dist_{out}, Dist_{in}))$, where $Dist_{out}$ is the distance of layer l with layers that are clustered in the closest cluster of layer l but not the same, and A_{in} is the intra cluster distance defined as the average distance to all other layers in the same cluster as l . For this distance we used GED, accounting for the structural changes. The results of Table 5 indicate that in most cases the quantum-inspired distance methods are forming clusters of higher silhouette values, thereby clustering the layers in more concise clusters.

4.4 Multiplex graph classification

In this section we utilize data collected by the Food and Agriculture Organization of the United Nations (FAO), that

describe trade relationships among countries for 2010, available under (Manlio Datasets). This is a multiplex network consisting of 214 nodes, each representing a country, and 364 layers, each holding import/export relationships of a specific food product. For each node, the trade ego-network is extracted, and the classification problem is the prediction of the country's development status, according to the composition of economies available from United Nations Conference on Trade and Development (UNCTADStat (as of June 2021)), using only the information of the ego-networks. The classification task is completed using all possible multiplex representations and distance measures listed above, while the F1 score is used as a measure of the performance of each classifier. In order to evaluate the ability of each representation to incorporate information from multiple layers, we repeat the classifications each time including a different number of layers, in decreasing size order. The results are reported in Figs. 5 and 6.

The results of Fig. 5 focus on presenting how the F1 score changes as more layers are used for the classification, for different multiplex representations and distance measures. The plot indicates that most of the distances are able to achieve an accurate classification when two or more layers are used, with an F1 score that is increasing with the addition of more layers. The results are similar for all possible multiplex representations.

As the number of the layers increases even more, the use of the supra-adjacency becomes impractical, however the monoplex aggregation techniques, with a smaller computational complexity, are able to achieve comparable accuracy scores, that continuously increases with the addition of more

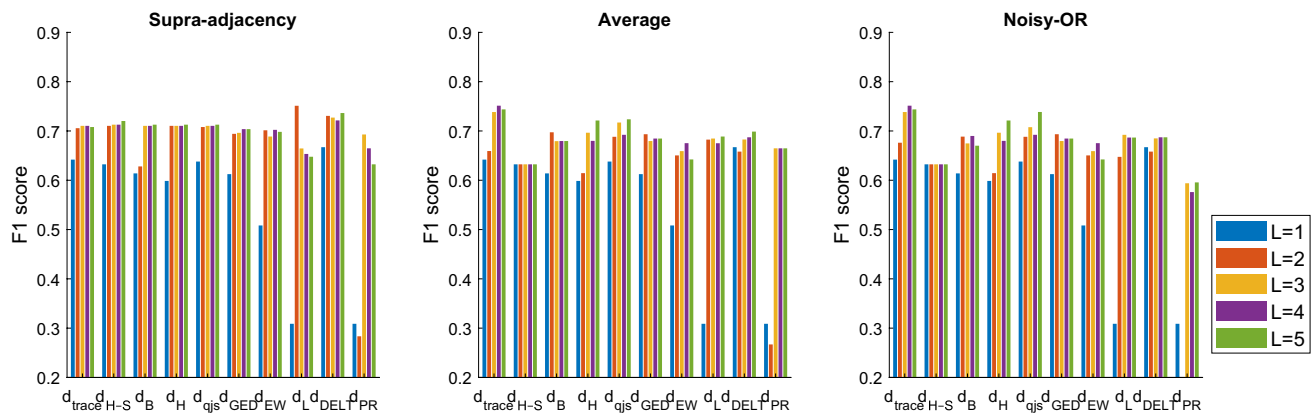


Fig. 5 Bar plots, for different multiplex representations, of the change in F1 scores as more layers are included in the classification, and for different graph distance methods. The plots for the natural representa-

tion of Supra-adjacency, the most common representation of Average and the newly-introduced Noisy-OR are included, while for all other representations the plots are very similar

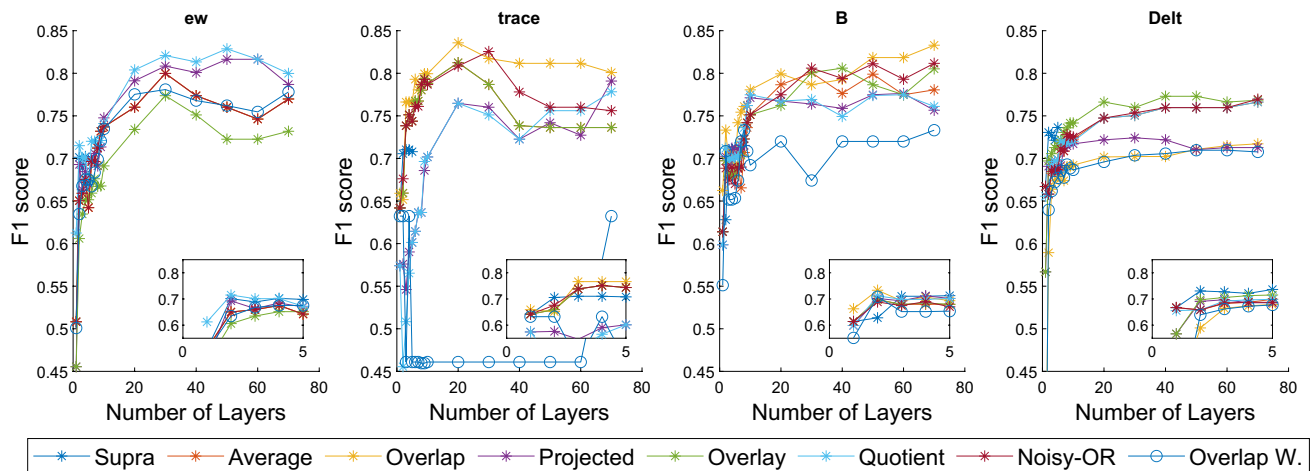


Fig. 6 Comparison of F1 scores, for each of the distance measures, when different multiplex representations are used, and as more layers are incorporated in the calculation. Some distances are omitted

for space but their results are discussed in the main text. The inserts focus on the use of only 1, to 5 layers

layers. Figure 6 illustrates the change of F1 score as even more layers are used for the classification. What is evident is that traditional methods, and the baselines, such as DELTACON, as well as quantum distances are increasing their accuracy with the addition of more layers. Out of the traditional methods, for this weighted dataset the edge-weight distance outperforms GED, while DELTACON outperforms the normalized Laplacian spectral distance and the resistance-perturbation distance. For the quantum measures, trace and Bures distances present the highest F1-scores, with Quantum JSDiv and Hellinger distances following closely. For many of the multiplex representations the accuracy is starting to drop after the addition of a given number of layers, defining a point at which the complexity of the problem increases,

more evidently for the edge-weight and trace distances. The same behavior is observed for the Hellinger distance, while the Hilbert-Schmidt distance presents the lowest F1-score overall and the added complexity affects it the most.

The previously reported findings are specific to the FAO dataset and cannot be generalised for all possible applications. In general, the increased complexity of multiplex systems does not always guarantee an increase in the informative power of these structures. Layers are informative only if their incorporation into the network yields a more detailed description of the data and if they unveil structural patterns not otherwise present (Tiago 2015). If this is not the case, additional layers increase the complexity of the system, and as dictated by Occam's razor, eventually obscure useful

Table 6 Average classification accuracy, using Supra-Adjacency representation in terms of F1 scores, when using one, combinations of two, or three layer

	d_H	d_B	d_{trace}	d_{HS}	d_{qjs}	d_{GED}	d_{DEW}	$d_{\mathcal{L}}$	d_{DELT}	d_{PR}
KHigh-Age										
All Layers:	0.6585	0.6585	0.6585	0.6479	0.5868	0.6585	0.6585	0.6119	0.6585	0.6119
2 Layers	0.6125	0.5973	0.5973	0.5866	0.6326	0.5662	0.5422	0.4899	0.5901	0.5416
1 Layer	0.6290	0.5809	0.5010	0.5886	0.5421	0.6064	0.5748	0.5136	0.5842	0.5909
LLaw-Practice										
All Layers	0.8932	0.8647	0.8767	0.8767	0.8767	0.8767	0.8767	0.5633	0.8767	0.8700
2 Layers	0.8866	0.8818	0.8921	0.8919	0.8937	0.8866	0.9001	0.5619	0.8745	0.8683
1 Layer:	0.9361	0.9359	0.9311	0.9321	0.9223	0.9361	0.5359	0.4492	0.9101	0.8605
LLaw-Status										
All Layers	0.7393	0.7752	0.7887	0.7393	0.7182	0.7265	0.7265	0.5502	0.7895	0.7992
2 Layers:	0.7554	0.8094	0.7934	0.7447	0.7661	0.7554	0.7606	0.6296	0.7567	0.8098
1 Layer	0.6996	0.6715	0.6832	0.7122	0.6057	0.6841	0.2825	0.5970	0.7098	0.7622

structural information. Such examples are seen in the classification results of Table 6, for two datasets available under (Manlio Datasets), the Krackhardt High-Tech company and the Lazega law firm.

In the Lazega law firm data set the three layers represent “advice”, “friendship”, and “co-work”. For the prediction of Practice, all 3 layers independently perform well, with “co-work” outperforming the other two, while use of combinations of these layers reduces the resulting accuracy. In the same data set, the prediction of node status behaves differently. In this case, as is intuitively expected, layers “advice”, and “friendship” appear to perform better than “co-work”, while their combination outperforms all other options. Notice that the monoplex aggregate representations provide similar results with F1 scores for the prediction of Practice (when using the Hellinger distance) ranging between 0.902, and 0.928.

It becomes evident that the benefit of utilization of additional layers depends on the system under study, and is beneficial only when the layers are informative, meaning that for classification purposes the layers need to be well correlated with the label and contain structural information not otherwise present. In most cases the quantum-inspired methods perform better than the baselines, and all multiplex representations are able to capture the underlying structure.

5 Discussion

In this work, measures for graph similarity inspired by information physics were introduced. These measures are well-established mathematical methods that incorporate the intrinsic structure of the entire network and have high interpretability. They can be used with both weighted and unweighted networks, are shown to effectively distinguish

monoplex and multiplex networks, and capture graph intuitive aspects of structural changes, such as node centrality. In the case of multiplex networks, the noisy-OR aggregate monoplex is introduced and contrasted with aggregates previously proposed in the bibliography. We utilized artificial data to characterise the methods and real-world data sets to showcase that they can be effectively used on a variety of applications outperforming previously well-established and state-of-the-art methods.

Quantum and information physics undoubtedly compose a rich resource of mathematically established and interpretable measures, used for the evaluation or comparison of natural systems. Many of these measures have been translated into meaningful complex network tools, in this work and previous ones (Biamonte et al. 2019). For many other measures their graph-interpretation remains to be seen. For example, quantum error correction methods, typically used to remove environmental interactions and other forms of noise from quantum states, could inspire methods eliminating different forms of noise from complex systems and highlighting the information they carry. Another possible example would be measures of quantum entanglement such as the entropy of entanglement, quantum concurrency, or quantum negativity, which are typically used to measure how entangled two quantum states are, or how far these states are from achieving separability. For complex networks the quantum separability can be translated into network distinguishability. In addition, in the case of multiplex networks, quantum states separability can be seen as layer separability, a possible measure of the layers’ informative power.

Acknowledgements This research is supported in part by the U.S. Army Research Laboratory subaward 555080-78055 under Prime Contract No.W911NF2220001, the U.S. Army Corp of Engineers

Engineer Research and Development Center under Cooperative Agreement W9132V-22-2-0001, and Temple University office of the Vice President for Research 2022 Catalytic Collaborative Research Initiative Program AI & ML Focus Area.

Author contributions AP conceived the presented study, performed the design and implementation of the research, and the analysis of the results. JA processed data and prepared tables. Both AP and ZO contributed to the final version of the manuscript. ZO supervised the project.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Austin R. Benson data sets <https://www.cs.cornell.edu/~arb/data/>
- Battiston F, Nicosia V, Latora V (2014) Structural measures for multiplex networks. *Phys Rev E* 89(3):032804
- Biamonte J, Faccin M, De Domenico M (2019) Complex networks from classical to quantum. *Commun Phys* 2(1):1–10
- Bianconi G (2013) Statistical mechanics of multiplex networks: entropy and overlap. *Phys Rev E* 87(6):062806
- Boccaletti S et al (2014) The structure and dynamics of multilayer networks. *Phys Rep* 544(1):1–122
- Braunstein SL, Ghosh S, Severini S (2006) The Laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states. *Ann Comb* 10(3):291–317
- Bunke H, Dickinson PJ, Kraetzl M, Wallis WD (2007) A graph theoretic approach to enterprise network dynamics. Birkhäuser, Boston
- Chodrow PS, Veldt N, Benson AR (2021) Generative hypergraph clustering: from blockmodels to modularity. *Sci Adv* 7(28):eabh1303
- Dajka J, Iużcka J, Hänggi P (2011) Distance between quantum states in the presence of initial qubit-environment correlations: a comparative study. *Phys Rev A* 84(3):032120
- De Domenico M et al (2015) Structural reducibility of multilayer networks. *Nat Commun* 6(1):1–9
- De Domenico M et al. (2013) Mathematical formulation of multilayer networks. *Phys Rev X* 3(4): 041022
- De Domenico M, Biamonte J (2016) Spectral entropies as information-theoretic tools for complex network comparison. *Phys Rev X* 6(4):041062
- Faccin M et al (2014) Community detection in quantum complex networks. *Phys Rev X* 4(4):041012
- Fowler JH (2006) Connecting the congress: a study of cosponsorship networks. *Polit Anal* 14(4):456–487
- Gilchrist A, Langford NK, Nielsen MA (2005) Distance measures to compare real and ideal quantum processes. *Phys Rev A* 71(6):062310
- Horst B (2000) Graph matching: theoretical foundations, algorithms, and applications. *Proc. Vision Interface*, vol 2000
- Javarone MA, Armano G (2013) Quantum-classical transitions in complex networks. *J Stat Mech Theory Exp* 2013(04): P04019
- Jia J, Benson AR (2022) A unifying generative model for graph learning algorithms: label propagation, graph convolutions, and combinations. *SIAM J Math Data Sci* 4(1):100–125
- Jozsa R (1994) Fidelity for mixed quantum states. *J Mod Opt* 41(12):2315–2323
- Kempe J (2003) Quantum random walks: an introductory overview. *Contemp Phys* 44(4):307–327
- Kivelä M et al (2014) Multilayer networks. *J Complex Netw* 2(3):203–271
- Koutra D, Vogelstein JT, Faloutsos C (2013) Deltacon: a principled massive-graph similarity function. In: *Proceedings of the 2013 SIAM international conference on data mining. society for industrial and applied mathematics*
- Lamberti PW et al (2008) Metric character of the quantum Jensen-Shannon divergence. *Phys Rev A* 77(5):052311
- Lee J, Kim MS, Brukner Č (2003) Operationally invariant measure of the distance between quantum states by complementary measurements. *Phys Rev Lett* 91(8):087902
- Leskovec J, Kleinberg J, Faloutsos C (2007) Graph evolution: densification and shrinking diameters. *ACM Trans Knowl Discov Data (ACM TKDD)*, 1(1)
- Leskovec J, Krevl A SNAP Datasets: Stanford large network dataset collection <http://snap.stanford.edu/data>
- Majtey AP, Lamberti PW, Prato DP (2005) Jensen–Shannon divergence as a measure of distinguishability between mixed quantum states. *Phys Rev A* 72(5):052310
- Manlio De Domenico data sets <https://manliodenedomenico.com/data.php>
- Mastrandrea R, Fournet J, Barrat A (2015) Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS One* 10(9):e0136497
- Monnig ND, Meyer FG (2018) The resistance perturbation distance: a metric for the analysis of dynamic networks. *Discrete Appl Math* 236:347–386
- Muthuganesan R., Chandrasekar VK, Sankaranarayanan R (2020) Quantum coherence and correlation measures based on affinity. *arXiv preprint arXiv:2003.13077*
- Nicosia V et al (2013) Growing multiplex networks. *Phys Rev Lett* 111(5):058701
- Nicosia V et al (2014) Nonlinear growth and condensation in multiplex networks. *Phys Rev E* 90(4):042807
- Nielsen MA, Chuang I (2002) Quantum computation and quantum information, pp 558–559
- Papadimitriou P, Dasdan A, Garcia-Molina H (2010) Web graph similarity for anomaly detection. *J Internet Serv Appl* 1(1):19–30
- Peixoto TP (2015) Inferring the mesoscale structure of layered, edge-valued, and time-varying networks. *Phys Rev E* 92(4):042807
- Polychronopoulou A, Alshehri J, Obradovic Z (2021) Distinguishability of graphs: a case for quantum-inspired measures. In: *Proceedings of the 2021 IEEE/ACM International conference on advances in social networks analysis and mining*
- Rahmede C et al (2018) Centralities of nodes and influences of layers in large multiplex networks. *J Complex Netw* 6(5):733–752
- Sánchez-Burillo E et al (2012) Quantum navigation and ranking in complex networks. *Sci Rep* 2(1):1–8
- Sánchez-García RJ, Cozzo E, Moreno Y (2014) Dimensionality reduction and spectral properties of multilayer networks. *Phys Rev E* 89(5):052815
- Solé-Ribalta, Albert, et al. (2014) Centrality rankings in multiplex networks. In: *Proceedings of the 2014 ACM conference on Web science*
- Sole-Ribalta A et al (2013) Spectral properties of the Laplacian of multiplex networks. *Phys Rev E* 88(3):032807
- Srinivas S (1993) A generalization of the noisy-or model. *Uncertainty in artificial intelligence*. Morgan Kaufmann, Burlington
- Stehlé J et al (2011) High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS One* 6(8):e23176
- Trávníšek V et al (2019) Experimental measurement of the Hilbert-Schmidt distance between two-qubit states as a means for

- reducing the complexity of machine learning. *Phys Rev Lett* 123(26):260501
- United Nations Conference on Trade and Development data http://unctadstat.unctad.org/EN/Classifications/DimCountries_DevStatus_Hierarchy.pdf
- Wang B et al (2018) Network enhancement as a general method to denoise weighted biological networks. *Nat Commun* 9(1):1–8
- Wills P, Meyer FG (2020) Metrics for graph comparison: a practitioner's guide. *PloS One* 15(2):e0228728
- Wilson RC, Zhu P (2008) A study of graph spectra for comparing graphs and trees. *Pattern Recogn* 41(9):2833–2841

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.