Quantum Inf Process (2017) 16:25
DOI 10.1007/s11128-016-1456-z

Comparing classical and quantum PageRanks

T. Loke[1] · J. W. Tang[1,2] · J. Rodriguez[1] · M. Small[3] · J. B. Wang[1]

Received: 28 October 2015 / Accepted: 19 October 2016 / Published online: 19 December 2016

Abstract Following recent developments in quantum PageRanking, we present a comparative analysis of discrete-time and continuous-time quantum-walk-based PageRank algorithms. Relative to classical PageRank and to different extents, the quantum measures better highlight secondary hubs and resolve ranking degeneracy among peripheral nodes for all networks we studied in this paper. For the discrete-time case, we investigated the periodic nature of the walker's probability distribution for a wide range of networks and found that the dominant period does not grow with the size of these networks. Based on this observation, we introduce a new quantum measure using the maximum probabilities of the associated walker during the first couple of periods. This is particularly important, since it leads to a quantum PageRanking scheme that is scalable with respect to network size.

Keywords Szegedy quantum walk · Open system quantum walk · Network analysis · PageRanking

✉ J. B. Wang
jingbo.wang@uwa.edu.au

T. Loke
21065783@student.uwa.edu.au

1 School of Physics, The University of Western Australia, Crawley, WA 6009, Australia

2 School of Physics, Nanjing University, Nanjing, Jiangsu, China

3 School of Mathematics and Statistics, The University of Western Australia, Crawley, WA 6009, Australia

1 Introduction

Characterising the relative importance of nodes in a graph is a key element in network analysis. A ubiquitous application of such centrality measures is Google's PageRank algorithm [1,2], whereby the World Wide Web (WWW) is considered as a network of webpages (nodes) connected by hyperlinks (directed edges) between them. By ranking each webpage according to its PageRank centrality, the search engine's results are ordered based on their approximated quality.

There has been recent interest in formulating a quantum version of PageRank. Since the intuition behind Google's PageRank is a classical "random surfer" crawling the WWW, a quantum walker traversing the associated directed network can be expected to provide an analogous measure of PageRank. As the quantum analogue of classical random walks, quantum walks serve as building blocks for quantum algorithms that can outperform their classical counterparts [3]. It is thus interesting to study whether their quantum mechanical properties afford an advantage over Google's classical PageRank algorithm.

Paparo and Martin-Delgado [4] and Sánchez-Burillo et al. [5] have separately proposed two quantum PageRank measures. The former is based on a discrete-time quantum walk (DTQW), whereas the latter uses a continuous-time quantum walk (CTQW). While quantum walks on arbitrary undirected graphs have been well defined, extending this framework to include directed quantum walks is non-trivial due to the requirements of unitarity and reversibility of the walk [6]. To deal with this difficulty, the discrete-time quantum PageRank uses a non-fully directed but unitary walk, whereas the continuous-time algorithm forgoes unitarity in using an open-system quantum walk.

Paparo et al. [7] performed further analysis of their proposed quantum PageRank on complex networks, specifically on hierarchical graphs, directed scale-free graphs, and Erdős–Rényi random graphs. The quantum PageRank algorithm not only distinguished clearly between the three graph classes, but also exhibited distinct characteristics in terms of highlighting secondary hubs and lifting the degeneracy of low-lying nodes. While it displayed a smoother power-law behaviour on scale-free networks, it was more sensitive to coordinated attacks on hubs than the classical PageRank algorithm.

Nevertheless, the number of time steps required for the underlying discrete-time quantum walk to yield a reliable quantum PageRank is yet to be considered. We seek to address this by investigating the oscillatory nature of the walker's probability amplitudes across nodes in the network. Such a consideration is crucially important, should an efficient quantum-system-based implementation of the PageRank scheme become realisable.

The open-system-quantum-walk-based PageRank in [5] modelled directionality as the walker's non-unitary interaction with the environment. Similar to the discrete-time case, the open-system PageRank lifted classical rank degeneracy of lowly connected nodes, while preserving identification of the most central nodes. By extension, it is useful to ascertain whether the other characteristics found in [7] for the discrete-time quantum PageRank are reflected in the open-system scheme.

In this article, we largely follow the analysis in [7], but extend it in three ways. Firstly, we consider the timescale involved for discrete-time-quantum-walk-based

PageRank. For the network types considered here, we gauge a suitable number of time steps for the walker's evolution after which reliable PageRanks can be obtained. We propose such an upper bound that does not scale significantly with increasing network size. Secondly, rather than taking the time average of the walker's probability distribution, we propose an alternative indicator of PageRank based on the maximum probability amplitude achieved by the walker on each node. This has previously been studied as a centrality measure on undirected graphs in [8]. Thirdly, we concurrently analyse an open-system-based PageRank algorithm. In our comparative study of three quantum PageRank schemes, we discuss their relative performance in extracting practically useful information about the networks under consideration. This provides a better understanding of each scheme as tools for quantum-walk-based complex network analysis. Our results suggest that as per classical PageRank, quantum PageRanking distinguishes clearly between the outerplanar hierarchical, scale-free, and Erdős–Rényi network families. While the quantum measures pick out more secondary hubs and remove degeneracies among low-lying nodes [5,7], each exhibits such quantum advantage to different extents.

This article is organised as follows: Sect. 2 outlines the theoretical framework underlying the classical and quantum PageRank algorithms. In Sect. 3, we present our numerical results for the algorithms on three types of directed networks, namely outerplanar hierarchical, scale-free, and Erdős–Rényi networks. We continue our comparative analysis on the algorithms in terms of secondary hub resolution on scale-free networks, localisation–delocalisation of the walker, and power-law behaviour on scale-free networks. Finally, Sect. 4 contains discussion and conclusions.

2 Theory

2.1 Classical PageRank

Google's PageRank algorithm is a variant of eigenvector centrality. The PageRank vector I_{cl} is given by

$$G I_{\mathrm{cl}} = I_{\mathrm{cl}}, \tag{1}$$

where G is the Google matrix, defined as

$$G := \alpha E + \frac{(1-\alpha)}{N}\mathbf{1}. \tag{2}$$

Here N is the number of nodes in the network, E is a (patched) connectivity matrix, α is the damping parameter (typically $\alpha = 0.85$), and $\mathbf{1}$ is the matrix of all ones. Intuitively, the second term represents the possibility of the walker randomly hopping to any other node in the network [2].

Define the connectivity (or adjacency) matrix C of the network as $C_{jk} = 1$ if there is an edge from k to j, and $C_{jk} = 0$ otherwise. To obtain the patched E, C is modified such that each column k containing all zeroes (corresponding to a node k with zero out-degree) is replaced by a column with all entries set to $\frac{1}{N}$. The remaining

columns corresponding to nodes with outgoing link(s) are normalised to sum to one by dividing by the out-degree of the node. Denote the out-degree of a node k by D_k, with $D_k = \sum_j C_{jk}$. Mathematically, E is then

$$E_{jk} = \begin{cases} \frac{1}{N} & \text{if } D_k = 0 \\ \frac{C_{jk}}{D_k} & \text{if } D_k \neq 0 \end{cases} \tag{3}$$

and is in general column stochastic.

2.2 Szegedy-Google PageRank via discrete-time quantum walk

Szegedy's formalism of the discrete-time quantum walk is a quantisation of the Markov chain corresponding to a classical random walk [9–12]. Classically, for an N-node graph, such a process is described by an N-by-N matrix P of transition probabilities, where each entry P_{jk} denotes the transition probability from node k to node j. Szegedy's walk takes place on the Hilbert space $\mathcal{H}^{N^2} = \mathcal{H}^N \otimes \mathcal{H}^N$. This space is the span of all vectors $|j, k\rangle$, where each vector represents a directed edge in the graph from node j to node k.

First, we define the state vector

$$\begin{aligned} |\psi_j\rangle &:= |j\rangle \otimes \sum_{k=1}^{N} \sqrt{P_{kj}}\, |k\rangle \\ &= \sum_{k=1}^{N} \sqrt{P_{kj}}\, |j, k\rangle \end{aligned} \tag{4}$$

for each node $j = 1, \ldots, N$ of the graph. This represents a superposition of edge states $|j\rangle_1 |k\rangle_2$ outgoing from the jth vertex, weighted by P. The projection operator is given by

$$\hat{\Pi} := \sum_{j=1}^{N} |\psi_j\rangle \langle\psi_j|, \tag{5}$$

and

$$\hat{S} := \sum_{j,k=1}^{N} |j, k\rangle \langle k, j| \tag{6}$$

is the swap operator. Then a step of the quantum walk is the unitary operator

$$\hat{U} := \hat{S}(2\hat{\Pi} - \hat{\mathbb{1}}), \tag{7}$$

whereas a two-step evolution operator takes the form

$$\hat{U}^2 := (2\hat{S}\hat{\Pi}\hat{S} - \hat{\mathbb{1}})(2\hat{\Pi} - \hat{\mathbb{1}}). \tag{8}$$

As proposed in [4], using the Google matrix G as the stochastic matrix P implements a quantum version of the classical PageRank algorithm. Unitarity of the quantum walk is maintained since G is stochastic; moreover, information on the directionality of the network is preserved in G.

The corresponding quantum walk is initialised as

$$|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} |\psi_j\rangle \,, \tag{9}$$

that is, an equal superposition across all nodes, but weighted among the edge states at each node by G. Taking $\hat{U}^2$ as the discrete-time evolution operator of the walk, the instantaneous quantum PageRank is then

$$I_q(P_i, t) = \langle\psi_0| \hat{U}^{\dagger 2t} |i\rangle_2 \langle i| \hat{U}^{2t} |\psi_0\rangle \,, \tag{10}$$

which is defined as the walker's probability distribution of the P_i pages in the network after t time steps. This value does not converge in time to any stationary distribution due to the unitarity and reversibility of the quantum-walk operator defined by Eq. (7).

Since a quantum PageRank measure must provide a unique ranking to each node in the graph, Paparo et al. defined it as the walker's time-averaged probability distribution [4,7]:

$$I_{\mathrm{TA}}(P_i) := \left\langle I_q(P_i, t) \right\rangle = \frac{1}{t_{\max}} \sum_{t=0}^{t_{\max}-1} I_q(P_i, t), \tag{11}$$

which converges for large enough $t_{\max}$. This will be referred to as the time-averaged (TA) PageRank measure in this paper.

We propose an alternative PageRank measure based on the peak probability of finding the walker on the node. We use the maximum $I_q(P_i, t)$ reached after $t_{\max}$ to be the quantum PageRank of a node:

$$I_{P_{\max}}(P_i) := \max\{I_q(P_i, t) : 1 \leq t \leq t_{\max}, t \in \mathbb{Z}\}. \tag{12}$$

We seek to gauge a suitable timescale $t_{\max}$ based on the oscillatory evolution of $I_q(P_i, t)$ according to Eq. (10). First, we apply $t = 500$ time steps of $\hat{U}^2$ onto the initial state (9). Performing a Fourier transform on the time series $I_q(P_i, t)$ yields a power spectrum of the oscillation frequencies present in it. We define $\omega(P_i)$ to be the lowest frequency present above noise using a threshold of 10% of the highest peak in the power spectrum [8]. The "period" of $I_q(P_i, t)$ is then $T_q(P_i) = \frac{2\pi}{\omega(P_i)}$. In general, each node i in the network, corresponding to page P_i, has a different period $T_q(P_i)$.

Denote the mean period of all nodes as $\langle T_q^{all} \rangle$:

$$\langle T_q^{all} \rangle := \frac{1}{N} \sum_{i=1}^{N} T_q(P_i). \tag{13}$$

Table 1 Mean periods $\langle T_q^5 \rangle$ and $\langle T_q^{all} \rangle$ of the three network types considered, with number of nodes N

Network type	N	$\langle T_q^5 \rangle$	$\langle T_q^{all} \rangle$
Outerplanar hierarchical	32	21.0	19.7
	64	24.0	19.7
	128	21.4	19.8
	256	20.2	20.5
	512	21.0	20.2
Scale-free	32	95.3	109.0
	64	100.2	120.7
	128	85.8	118.9
	256	91.0	111.0
	512	99.1	113.3
Erdős–Rényi	32	37.7	68.3
	64	56.6	64.4
	128	54.5	72.5
	256	20.4	33.8
	512	11.0	17.5

We use $t_{\max} = 2\langle T_q^5 \rangle$ as the number of time steps to obtain I_{TA} and $I_{P_{\max}}$ for a given network type and size

Let $\langle T_q^5 \rangle$ be the mean period of the five nodes whose instantaneous quantum PageRanks $I_q(P_j, t)$ reach the highest peak values within their respective periods $1 \leq t \leq T_q(P_j), t \in \mathbb{Z}$.

Following the above steps, we compute $\langle T_q^{all} \rangle$ and $\langle T_q^5 \rangle$ for the directed network families relevant to this study, namely outerplanar hierarchical, scale-free, and Erdős–Rényi networks with sizes $N = 32, 54, 128, 256, 512$ nodes. We use an ensemble of ten scale-free and Erdős–Rényi random networks for each N, generated using NetworkX [13]. For each Erdős–Rényi network here and throughout this article, the probability for edge creation is set to $p = 0.07$. We use $t_{\max} = 2\langle T_q^5 \rangle$ as the required timescale for our PageRank analyses, reasoning that the periods of the most central nodes should figure more strongly over those of the peripheral nodes in determining the general timescale for each network.

Numerical results are shown in Table 1, and Fig. 1 plots the scaling of the mean periods with network size. Our results suggest that $t_{\max} = 2\langle T_q^5 \rangle$ does not scale linearly upward with N, rather it remains stable for the network types considered here. In the case of the deterministically constructed outerplanar hierarchical networks, the mean period plateaus at approximately $\langle T \rangle = 20$ time steps for successive generations. Overall, the mean periods are highest for scale-free networks. We see that larger Erdős–Rényi networks (with same edge probability $p = 0.07$) tend to have smaller mean periods. We expect the timescale for higher N to remain similarly bounded.

This is an important observation since it provides the possibility of implementing a quantum PageRank scheme efficiently, which is a crucial issue not yet addressed previously in the literature. Note that Chiang et al. [14] presented an efficient quantum circuit to implement Szegedy walks on arbitrary sparse networks. Although the use of the Google matrix in I_{TA} and I_{OS} causes the associated unitary evolution operator $\hat{U}$ to be dense, Loke and Wang [15] exploited partitioning the Google matrix into

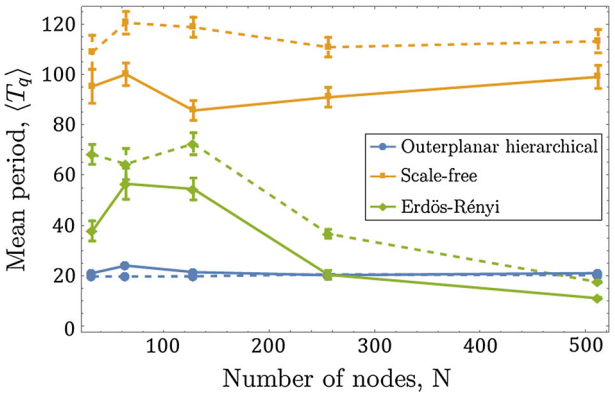

Fig. 1 Scaling of mean periods with network size for $N = 32, 64, 128, 256, 512$ nodes. Mean periods $\langle T_q^{\delta} \rangle$ (*solid*) and $\langle T_q^{all} \rangle$ (*dashed*) for outerplanar hierarchical, scale-free, and Erdős–Rényi networks are plotted in *blue*, *orange*, and *green*, respectively. For both the scale-free and Erdős–Rényi random networks, an ensemble of ten graphs is used for each N. Each *error bar* corresponds to the standard error of the mean of the ten $\langle T_q \rangle$ values from each ensemble (Color figure online)

manageable subsets and extended the scheme of Chiang et al. [14] to implement the Google matrix efficiently, as long as the original network is sparse.

2.3 Open-system PageRank via continuous-time quantum walk

The continuous-time quantum walk was originally proposed by Farhi and Gutmann [16] out of a study of computational problems reformulated in terms of decision trees. Following the Schrödinger equation, such evolution is described by

$$\frac{\mathrm{d}\,|\Psi(t)\rangle}{\mathrm{d}t} = -i\hat{H}\,|\Psi(t)\rangle\,, \tag{14}$$

where $\hat{H}$ is the transition rate matrix. Requiring unitary evolution operators in quantum mechanics implies that $\hat{H}$ must be Hermitian, which is generally not the case for a directed walk. To introduce directionality into CTQWs, we employ the open-system method using the Lindblad–von Neumann equation, which accounts for the non-unitary nature of the directed walk through coupling with an external environment.

To work with open quantum systems, the concept of a density operator is used as a substitution for wave functions in quantum mechanics. The density operator for the system is defined by [17]

$$\rho = \sum_{i=1}^{N} p_i\,|\Psi_i\rangle\,\langle\Psi_i| \tag{15}$$

where p_i are constants that represent how much of state $|\Psi_i\rangle$ is in the final mixed state, with $\sum_{i=1}^{N} p_i = 1$.

The Lindblad–von Neumann equation describes how a quantum system evolves after tracing out the environment, and can be written in the form [18]:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i\hbar[\hat{H}, \rho] + \sum_k \gamma_k \left(\hat{L}_k \rho \hat{L}_k^\dagger - \frac{1}{2} \left\{ \hat{L}_k \hat{L}_k^\dagger, \rho \right\} \right), \tag{16}$$

where $\hat{L}_k$ are unitary operators on the space that ρ is in. The set of all $\hat{L}_k$ forms a basis for this space. The matrix γ describes how non-energy-conserving phenomena such as temperature affect the system.

To parameterise interpolation between classical (undirected) and classical (directed) behaviours, a damping parameter β is introduced into the Lindblad–von Neumann equation:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i(1-\beta)[\hat{H}, \rho] + \beta \sum_{i,j} \gamma_{ij} \left(\hat{L}_{ij} \rho \hat{L}_{ij}^\dagger - \frac{1}{2} \left\{ \hat{L}_{ij} \hat{L}_{ij}^\dagger, \rho \right\} \right), \tag{17}$$

where $\hat{H}$ is a symmetrized form of the adjacency matrix C representing its corresponding undirected graph. γ is taken as the patched connectivity matrix E as per Eq. (3)—this a specific case of the Google matrix G in Eq. (2) with $\alpha = 1$. Our approach is equivalent to the quantum PageRank algorithm developed by Sánchez-Burillo et al. [5], where they set $\gamma = G$ with $\alpha = 0.9$ instead.

We solve the master equation via an eigenoperator method used by Saalfrank [19]. This is a linearisation method that turns a nonlinear equation into a linear one, whereby Eq. (17) becomes

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -i(1-\beta)\mathcal{L}_H \rho + \beta \mathcal{L}_D \rho = \mathcal{L}_{SO} \rho. \tag{18}$$

This takes the eigenoperators $\mathcal{L}_H$ and $\mathcal{L}_D$ from the original N-dimensional space to a space of N^2 dimensions, and the density matrix is vectorised in this set-up.

As per Eq. (18), $\mathcal{L}_H$ and $\mathcal{L}_D$ can be combined into one operator $\mathcal{L}_{SO}$. For time-independent $\mathcal{L}_{SO}$, this form is readily solved for any time t by taking the matrix exponential of $\mathcal{L}_{SO}$, i.e.

$$\rho = \rho_0 e^{\mathcal{L}_{SO} t}. \tag{19}$$

Convergence to a stationary result for large enough t is guaranteed [5], upon which the occupation probabilities of each node indicate the open-system quantum PageRank, namely

$$I_{\mathrm{OS}}(P_i) := \langle i|\rho|i\rangle = \rho_{ii}. \tag{20}$$

3 Results

To recap, the four PageRank measures considered here are as follows:

- I_{cl} (1)—classical Google PageRank

- I_{TA} (11)—DTQW-based PageRank using the time average of the instantaneous quantum PageRank I_q
- $I_{P_{\max}}$ (12)—DTQW-based PageRank using the maximum I_q reached
- I_{OS} (20)—CTQW-based PageRank using the open-system method

We set $\alpha = 0.85$ in the Google matrix G (2) for I_{cl}, I_{TA}, and $I_{P_{\max}}$. For I_{OS}, we use $\alpha = 1$ in G and $\beta = 0.85$ in the master equation (17).

For all numerical results below, the DTQW-based PageRank measures I_{TA} (11) and $I_{P_{\max}}$ (12) are computed using $t_{\max} = 2\langle T_q^5 \rangle$ steps of the evolution operator $\hat{U}^2$, with values for $\langle T_q^5 \rangle$ in Table 1. For I_{OS}, we use sufficiently large t for convergence to the stationary result.

3.1 Three types of directed networks considered

To allow meaningful comparison between the four PageRank measures on various networks, we normalise the PageRanks by dividing through by the maximum value obtained so that the most central node has a PageRank value of 1.

3.1.1 Outerplanar hierarchical networks

As introduced by Comellas and Miralles [20], outerplanar hierarchical networks are a family of modular, self-similar, small-world graphs with zero clustering. This family mirrors social, technological, and biological systems with a low clustering [21]. Outerplanarity refers to the network having an embedding where all nodes lie on the boundary of the exterior face, whereas the hierarchical structure is realised by using a recursive method of network construction. Each generation n has $N = 2^{n+1}$ nodes; thus the network doubles in size for successive n, with newly added nodes being those indexed $2^n < i \leq 2^{n+1}$. We follow [7] in giving directions to the edges.

As shown in Fig. 2, all three quantum PageRank measures have a step-like behaviour that reflects the network's hierarchical structure. Within the same hierarchical level, edge directionality gives rise to non-degenerate PageRank values. This intra-level non-degeneracy has a smaller amplitude for the newly added nodes (evidenced by nodes 17–32 in $n = 4$, nodes 33–64 in $n = 5$, nodes 65–128 in $n = 6$, and nodes 129–256 in $n = 7$), thus enabling the PageRank measures to distinguish between pre-existing and newly added nodes in successive generations.

Comparing I_{TA} and $I_{P_{\max}}$, the latter measure tends to give more closely valued PageRanks within a hierarchical level, particularly for the newly added nodes. As depicted in Fig. 3, nodes whose oscillatory I_q never peak above the value subsequent to the initial state receive near-similar $I_{P_{\max}} = I_q(P_i, t = 1)$ after one time step. For example, $I_{P_{\max}}$ almost plateaus for nodes 21–24 in $n = 4$ and nodes 41–48 in $n = 5$. Taking the time average resolves this degeneracy as each node's I_q evolves slightly differently in time.

Furthermore and as would be expected, we see that pairs of automorphically equivalent nodes have identical time evolution of I_q. (Two nodes a and b are said to be automorphically equivalent if there exists an isomorphism in which the labels of a and b are interchanged [22,23].) Such nodes are thus ranked identically by I_{TA} and $I_{P_{\max}}$,

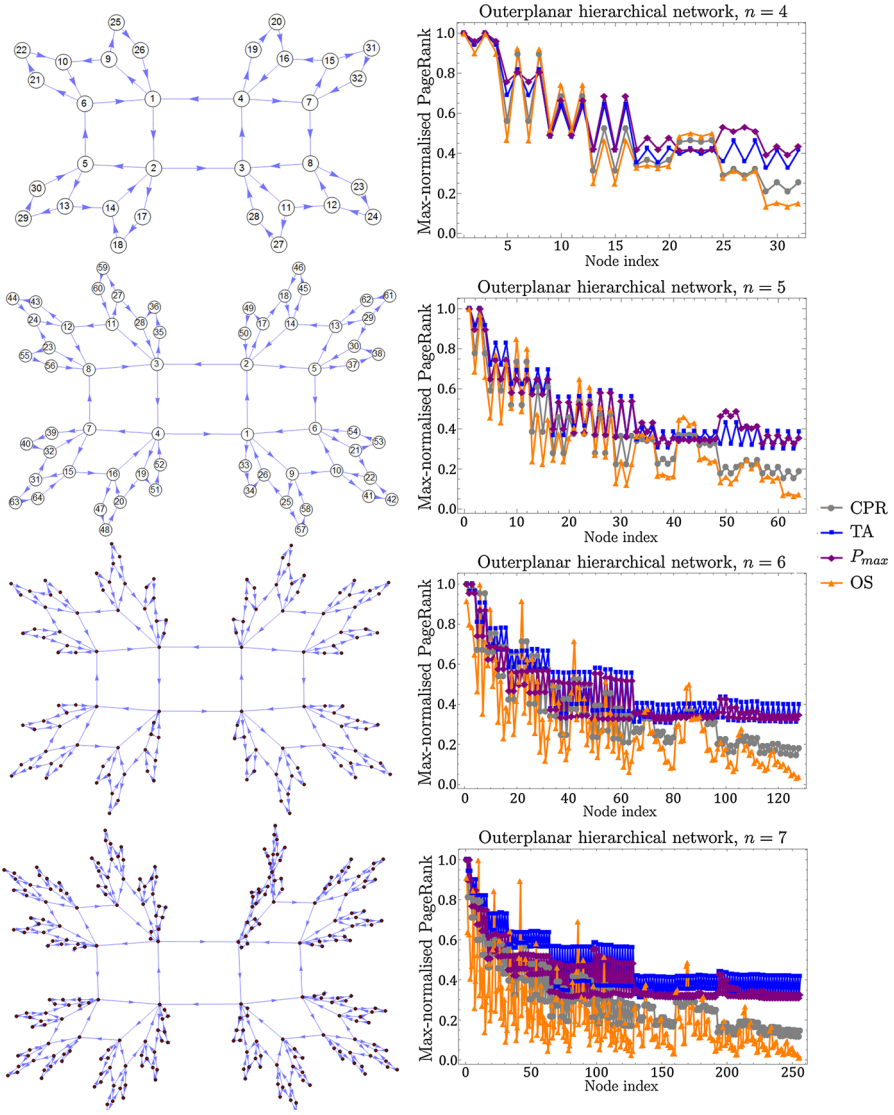

Fig. 2 PageRanks on outerplanar hierarchical networks for generations $n = 4, 5, 6, 7$. I_{cl}, I_{TA}, $I_{P_{\max}}$, and I_{OS}—normalised by their maximum values—are plotted in *grey*, *blue*, *purple*, and *orange*, respectively (Color figure online)

preserving the equivalence also present in I_{cl}. On the other hand, this identical ranking of automorphically equivalent nodes is not preserved by I_{OS}. We instead observe a falling pattern of I_{OS} values within the hierarchical levels.

In summary, the four PageRank measures considered here are able to uncover the hierarchical structure present in each network, but disagree on the relative rankings of

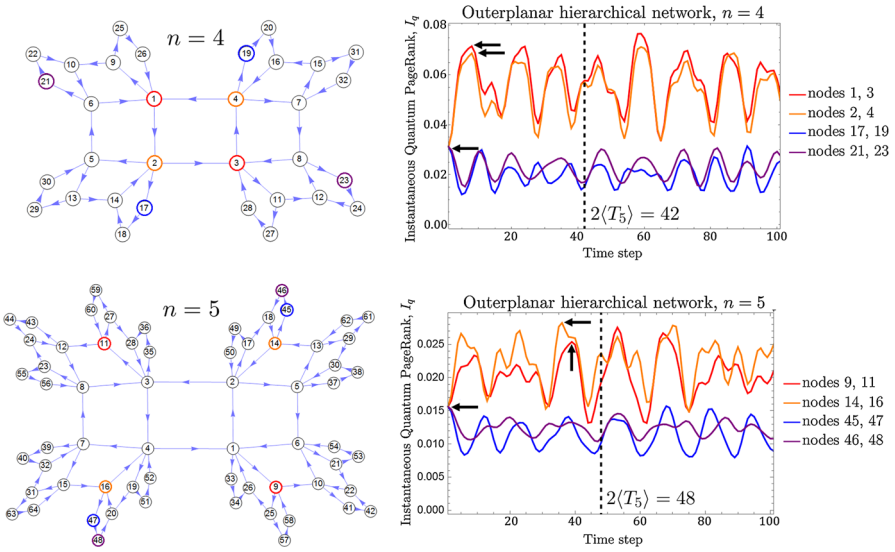

Fig. 3 Time evolution of I_q on $n = 4$ and $n = 5$ outerplanar hierarchical networks. *Left* the selected *like-coloured nodes* are automorphically equivalent and have I_q that evolve identically over time. *Right* time evolution of I_q. *Vertical dashed lines* bound $t_{\max} = 2\langle T_q^5 \rangle$ for the network, within which we determine $I_{P_{\max}}$ as indicated by *arrows*, and I_{TA} by taking the time-averaged probabilities (Color figure online)

these levels. Here I_{OS} most closely resembles I_{cl} in ranking each level, but, especially for higher n, its intra-level behaviour is in stark contrast to the other measures.

3.1.2 Scale-free networks

A power-law degree distribution typifies scale-free networks, that is, the probability $P(k)$ that a node is connected to k other nodes decays as $P(k) \sim k^{-\gamma}$ [24]. Such behaviour was first observed by Albert et al. [25] in their analysis of the topology of the World Wide Web, in that the numbers of incoming and outgoing hyperlinks of a webpage both follow a power law over several orders of magnitude. Numerous other real-world networks have since been found to be scale-free, from functional networks in the brain [26] and protein interactions in cells [27], to social networks and their technical derivatives such as the Internet, e-mail networks, and business collaboration [28].

Inherently linked to the scale-free property of a network is its evolution over time [29], which can be modelled by preferential attachment [30]. Starting with a small number of nodes, new nodes are added with a higher probability of being connected to pre-existing nodes that are already well connected—conceptually, "the rich get richer". In this study, we perform PageRank analysis on Bollobás et al.'s scheme for directed scale-free networks [31] as implemented in NetworkX [13]. This scheme allows multiple edges and loops.

Based on Fig. 4, all PageRank methods largely agree on identifying the most central nodes, or hubs. Figure 5 provides an example case for obtaining I_{TA} and $I_{P_{\max}}$ based on the time evolution of I_q.

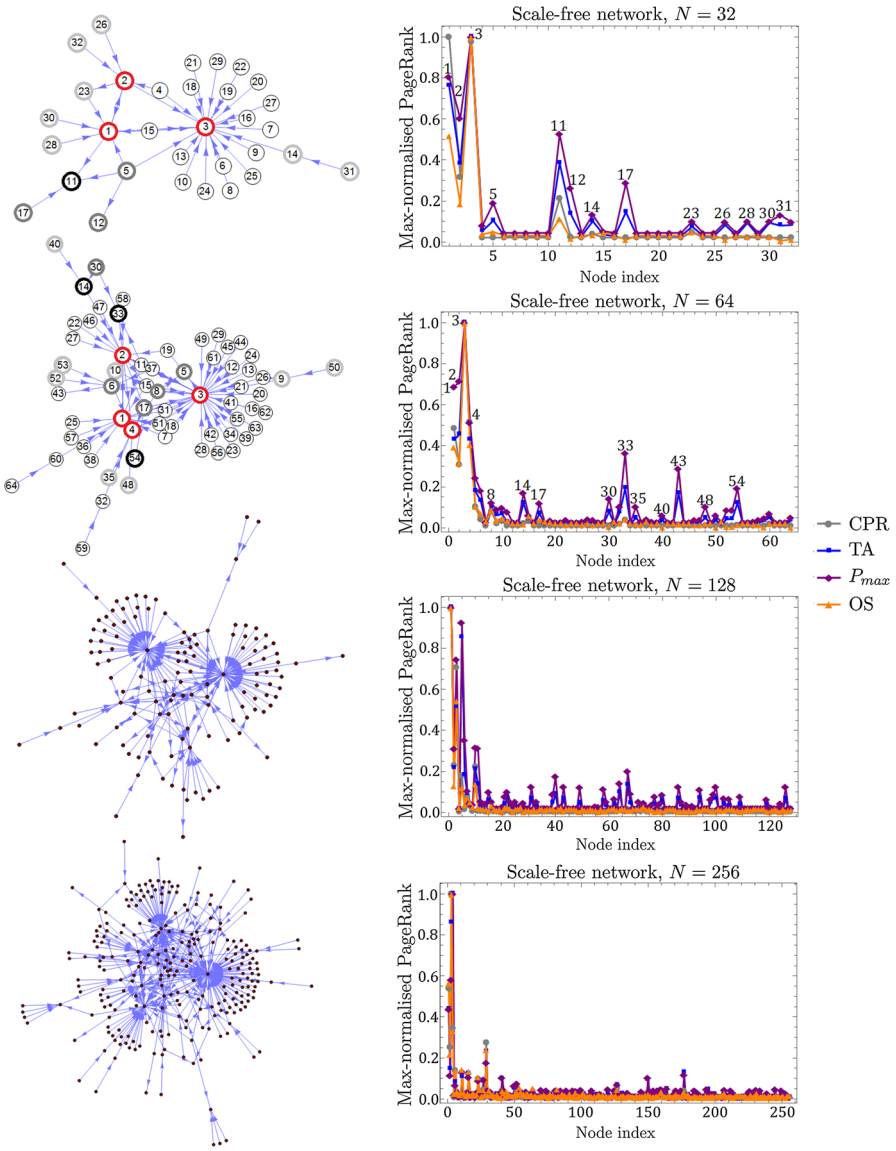

Fig. 4 PageRanks on directed scale-free networks of sizes $N = 32, 64, 128, 256$. I_{cl}, I_{TA}, $I_{P_{\text{max}}}$, and I_{OS}—normalised by their maximum values—are plotted in *grey*, *blue*, *purple*, and *orange*, respectively. For the $N = 32$ and $N = 64$ cases, nodes marked in *red* are the most central nodes, or hubs; secondary hubs identified by I_{TA} and $I_{P_{\text{max}}}$ are marked in *greyscale* (Color figure online)

As previously observed and noted in [7], I_{TA} better highlights the secondary hubs compared to I_{cl} rather than concentrating all the importance on the most important hubs. This improved ranking capability of I_{TA} over classical PageRank is even more strongly featured in $I_{P_{\text{max}}}$. We investigate this further in Sect. 3.2.

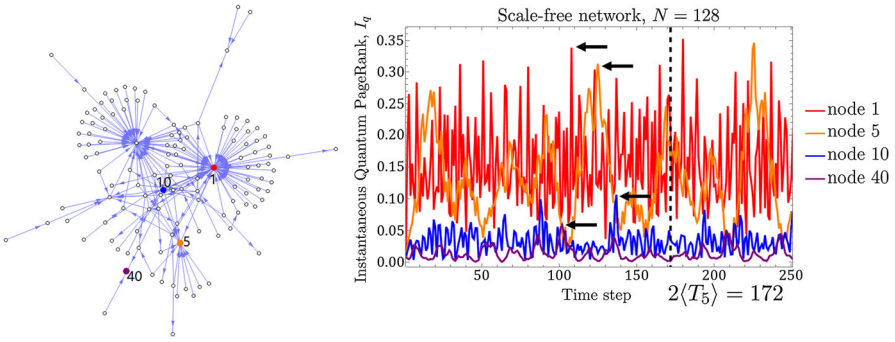

Fig. 5 Time evolution of I_q on a directed scale-free network of size $N = 128$. The *black vertical dashed line* bounds $t_{\max} = 2\langle T_q^5 \rangle$ for the network, within which we determine $I_{P_{\max}}$ as indicated by *arrows*, and I_{TA} by taking the time-averaged probabilities

For the examples considered here, nodes with little or no in-degree receive the lowest importance according to I_{cl} and I_{OS}, forming a near-plateau of low-lying nodes.

I_{TA} and $I_{P_{\max}}$ are additionally affected by which nodes a given node is pointing to. In the case of the 32-node scale-free network in Fig. 4, node 17 has unexpectedly high I_{TA} and $I_{P_{\max}}$ by virtue of its pointing to secondary hub node 11. Similarly, the nodes pointing to central hubs nodes 1 and 2 have low I_{cl} and I_{OS}, but have improved I_{TA} and $I_{P_{\max}}$. The more nodes there are that point to a given hub, the less advantage those nodes gain, as is the case for those surrounding node 3. This accords with the observation made by Sánchez-Burillo et al. [5] for their open-system-quantum-walk-based PageRank, in that nodes connected to hubs distributing their influence among a large number of connections receive lower rankings.

Therefore, low-lying nodes with degenerate I_{cl} can be distinguished by I_{TA} and $I_{P_{\max}}$ as these measures are more sensitive to the nodes' positions in the network. In particular, a node's ranking is increased by being linked to a network hub, provided it is among a few neighbours of the hub. This ability of quantumness to resolve classical PageRank degeneracies among peripheral nodes is also noted in [5,7].

Finally, we observe that I_{OS} closely resembles I_{cl} on these networks. Considering that I_{OS} as implemented here uses damping parameter $\alpha = 1 \Rightarrow G = E$ as the matrix coding classical behaviour (directionality) into the walk, this suggests that β in the master equation (18) (parameterising the classicality of the walk, here set to $\beta = 0.85$) can play a similar role to that of α in classical PageRanking (parameterising random hops to any node, set to $\alpha = 0.85$ for I_{cl}).

3.1.3 Erdős–Rényi random networks

An Erdős–Rényi random network of N nodes is constructed by choosing, with common edge probability p, whether or not to connect pairs of nodes, with the choices being independent for each node pair [32,33]. For large N, its degree distribution follows a Poisson distribution $P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$, where $\langle k \rangle$ is the mean degree. Despite having random edge positions, such a network is rather homogeneous as most nodes have the same degree [34]. In this study, we use directed Erdős–Rényi random networks generated using NetworkX [13].

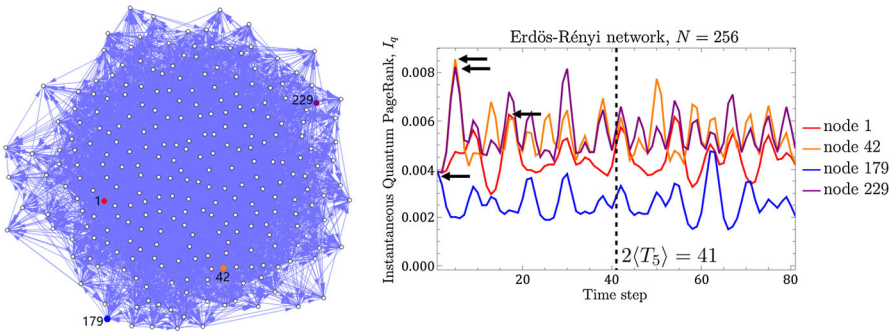

Fig. 6 Time evolution of I_q on a directed Erdős–Rényi network of size $N = 256$. The *vertical dashed line* indicates $t_{\max} = 2\langle T_q^5 \rangle$ for the network, within which we determine $I_{P_{\max}}$ as indicated by *arrows*, and I_{TA} by taking the time-averaged probabilities

An example case of obtaining the DTQW-based I_{TA} and $I_{P_{\max}}$ is presented in Fig. 6. Overall results in Fig. 7 demonstrate that for each PageRank measure, most nodes receive similar PageRanks with no discernible hubs. This is an ostensibly different distribution of PageRanks in contrast to the directed scale-free networks analysed earlier. Section 3.3 further studies this localised/delocalised behaviour of the walker on these two network types.

In summary, consistent with our analysis on scale-free networks, we find that $I_{P_{\max}}$ provides an alternative measure to I_{TA} as both exhibit similar features, whereas I_{OS} more closely resembles I_{cl}.

3.2 Detection of secondary hubs on scale-free graphs

The preliminary PageRank analyses on scale-free graphs suggest that the quantum-walk-based PageRank measures tend to give higher ranks to secondary hubs compared to classical PageRank. To quantify such secondary hub detection for each PageRank scheme, we analyse an ensemble of 30 directed scale-free networks of size $N = 256$ by categorising the nodes according to their PageRanks.

As before, for each network and PageRank method, we normalise the nodes' PageRanks by dividing through by the maximum so that the most central node has a PageRank of 1. Let the mean of these maximum-normalised PageRanks be a. Then we classify a node with PageRank x as a

- main hub if $x \geq ca$,
- secondary hub if $a \leq x < ca$,
- low-importance node hub if $x < a$,

where $c > 1$ is a fixed constant. Here we choose $c = 10$.

Results plotted in Fig. 8 affirm our earlier observations on scale-free networks. Comparing all PageRank measures considered, $I_{P_{\max}}$ identifies the most number of nodes as secondary hubs, followed by I_{TA}, I_{OS}, and finally I_{cl}. Based on Fig. 8c, the set of nodes classified as secondary hubs by I_{TA} is most likely a subset of those identified by $I_{P_{\max}}$. In particular, $I_{P_{\max}}$ never picks out less secondary hubs than I_{TA}; moreover,

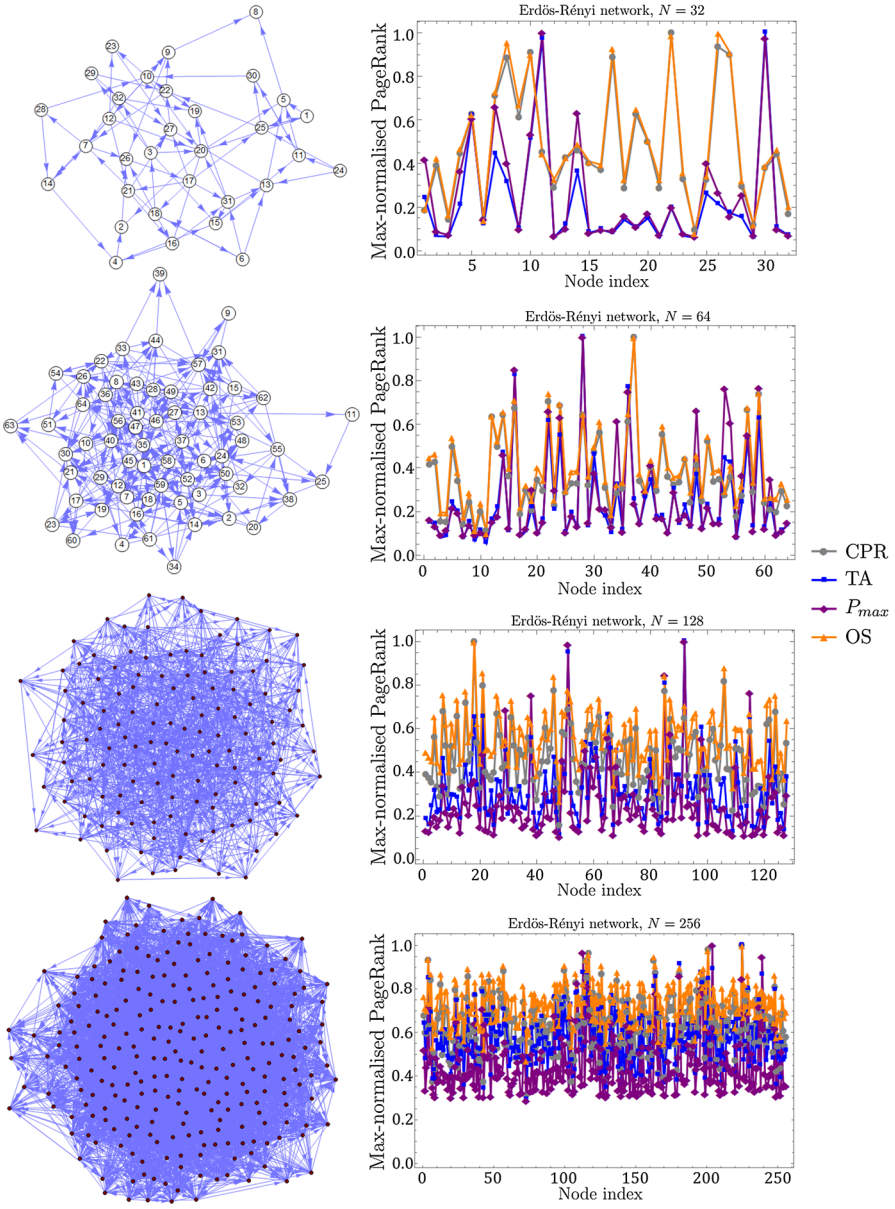

Fig. 7 PageRanks on directed Erdős–Rényi random networks of sizes $N = 32, 64, 128, 256$. I_{cl}, I_{TA}, $I_{P_{\max}}$, and I_{OS}—normalised by their maximum values—are plotted in *grey*, *blue*, *purple*, and *orange*, respectively (Color figure online)

both largely follow the same trend in quantifying secondary hubs across the network ensemble. On the other hand, I_{OS} exhibits a similar trend to I_{cl}. Although not as stark a difference as the other two quantum measures, I_{OS} also outperforms I_{cl} in secondary hub detection.

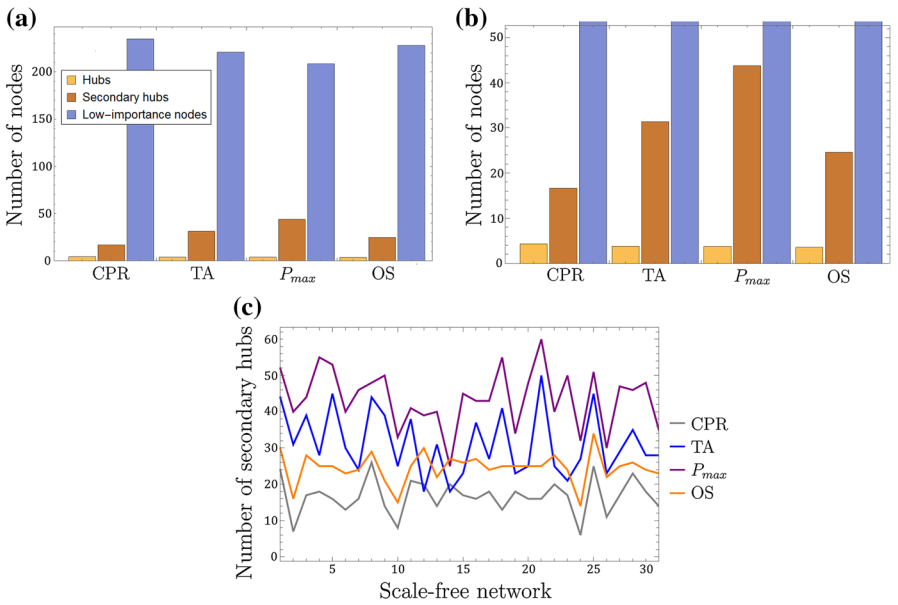

Fig. 8 Secondary hub resolution by PageRank algorithms on an ensemble of 30 directed scale-free networks of size $N = 256$. **a** A histogram of nodes classified as main hubs, secondary hubs, or low-importance nodes based on their (from *left* to *right*) I_{cl}, I_{TA}, $I_{P_{\max}}$, and I_{OS} values, respectively. **b** Zooming into **a**. The quantum PageRanks I_{TA}, $I_{P_{\max}}$, and I_{OS}, respectively, identify approximately 1.9, 2.6, 1.5 times more secondary hubs than I_{cl}. **c** The number of secondary hubs as measured by I_{cl} (*grey*), I_{TA} (*blue*), $I_{P_{\max}}$ (*purple*), and I_{OS} (*orange*) for each of the scale-free networks in the ensemble (Color figure online)

3.3 Localisation–delocalisation transition

To further compare the abilities of the various PageRank schemes to distinguish between scale-free and Erdős–Rényi random networks, we study their localisation behaviour on these network types. Since each PageRank scheme corresponds to a classical or quantum walk along nodes, we can infer the walker's degree of localisation based on the PageRank distribution across the network.

Within scale-free networks, the presence of hubs—nodes of unusually high degree—is the fundamental cause of localisation. Such localisation poses a problem for conventional eigenvector centrality because most of the weight of the centrality concentrates on a small number of nodes, thus necessitating the random-hop term in the Google matrix [35]. On the other hand, the relatively homogeneous degree distribution in Erdős–Rényi random networks is expected to favour a delocalised phase of the walker, as observed in [7].

To quantify localisation, we use the inverse participation ratio (IPR) defined as

$$\xi := \sum_{i=1}^{N} [\Pr(X = i)]^{2r} , \tag{21}$$

where $r > 0$ is a freely chosen integer and is fixed.

As in [7], we appropriate ξ to the case of PageRanks by considering X as a random variable whose realisations are the nodes of the network. Then the probability $Pr(X = i)$ corresponds directly to node i's PageRank I_{cl}, I_{TA}, or I_{OS}—for each of these measures, the sum of all PageRanks over the network is one. For a consistent probability interpretation of $I_{P_{\max}}$, we divide its values by the mean $\langle I_{P_{\max}} \rangle$, so that all such normalised values sum to one.

In the case of complete delocalisation, the walker's probability distribution would be uniform across the network, that is, $\Pr(X = i) = \frac{1}{N} \forall i$. If the walker is fully localised on one node j, $\Pr(X = i) = \delta_{ij}$, the Kronecker delta. Hence the IPR for these limiting cases is

$$\xi = \begin{cases} 1 & \text{if the walker is localised} \\ N^{1-2r} & \text{if the walker is delocalised.} \end{cases} \tag{22}$$

Rewrite the IPR as $\xi = N^{-\tau_{2r}}$ with

$$\tau_{2r} := (2r - 1) + \Delta_{2r}, \tag{23}$$

where Δ_{2r} is the normalised anomalous dimension, which interpolates between $\Delta_{2r} = 1 - 2r$ for a localised phase and $\Delta_{2r} = 0$ for a delocalised one. Then

$$\log \xi \sim (1 - 2r - \Delta_{2r}) \log N. \tag{24}$$

Choosing $r = 1$ and plotting $\log \xi$ against $\log N$, analytical values for the gradient of the plot are $a = 0$ and $a = -1$ for complete localisation and delocalisation, respectively.

Based on Eq. (24), we analyse the localisation–delocalisation transition of each PageRank measure. We compute the IPR for scale-free and Erdős–Rényi networks of size $N = 32, 64, 128, 256$, and 512, using an ensemble of ten graphs for each N.

As shown in Fig. 9, performing linear fits on the resulting log–log plots yield numerical gradient values a that are near-zero for the scale-free plots, and close to -1 for the Erdős–Rényi plots. These values, respectively, correspond to localised and delocalised phases of the walker, as was expected based on the presence and absence of hubs in these networks. All four PageRank measures are thus able to distinguish between scale-free and Erdős–Rényi random networks by virtue of the walker's localisation behaviour.

3.4 Power-law behaviour on scale-free networks

Analyses of classical PageRank on real-world Web graphs have revealed a power-law distribution of PageRank values [36,37]. This reflects a characteristic property of such scale-free networks, in that only a few main hubs account for much of the PageRank allocation in such scale-free networks. In particular,

$$I_j \sim j^{-\lambda}, \tag{25}$$

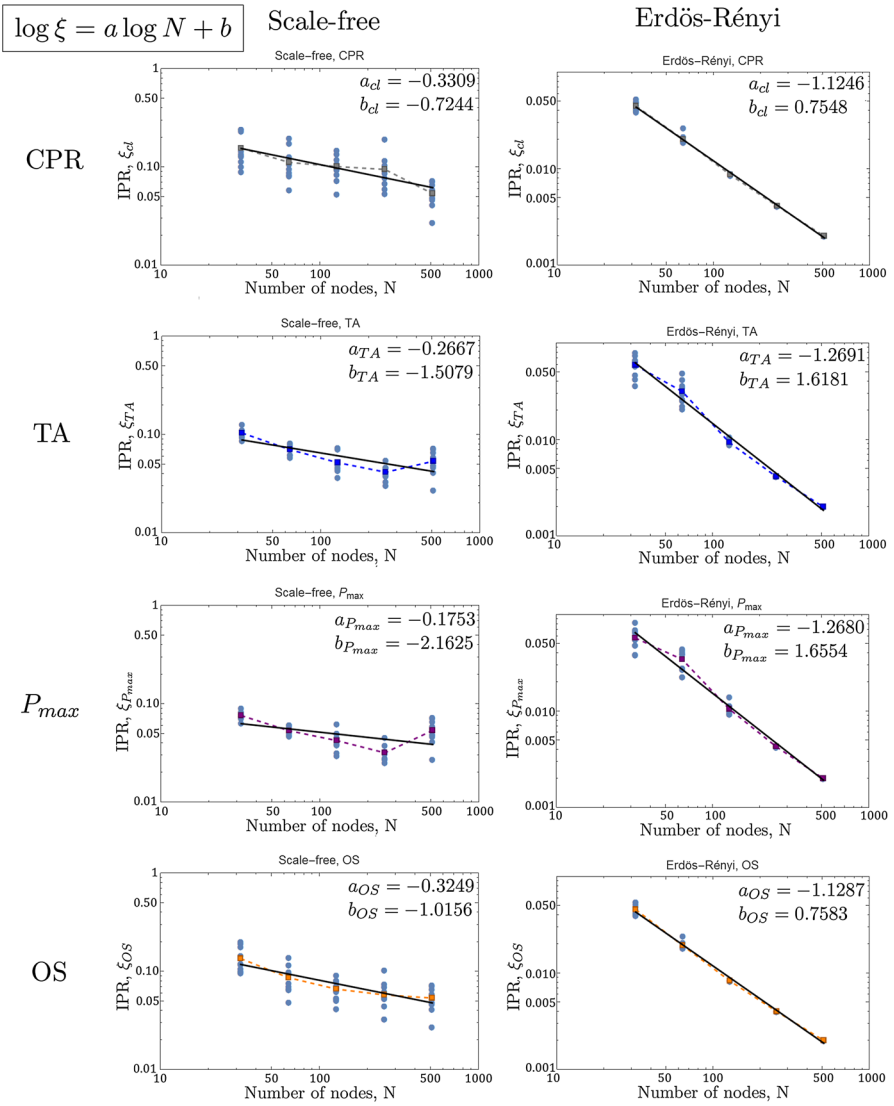

Fig. 9 Log–log plots of IPR (with $r = 1$) against the number of nodes in the network for the four PageRank schemes. Successive rows correspond to I_{cl}, I_{TA}, $I_{P_{\mathrm{max}}}$, and I_{OS}; *left* and *right columns* correspond to directed scale-free and Erdős–Rényi networks, respectively. An ensemble of ten graphs is used for each network type and size, with a *dotted line* joining the mean IPR values for each size. Linear model fits are performed according to $\log \xi = a \log N + b$ and plotted in *black*, with parameter values a and b as indicated

where the I_j are the PageRanks of nodes sorted in descending order, and λ is the power-law scaling coefficient. Such a power-law behaviour confirms that the PageRank algorithm is able to reveal a network's scale-free nature, with λ measuring the relative importance of hubs with respect to the other less important nodes [7].

Here we verify and compare such power-law behaviour for the PageRank measures $I_j = I_{\mathrm{cl}}, I_{\mathrm{TA}}, I_{P_{\mathrm{max}}}$, and I_{OS}. As per Sect. 3.3, the $I_{P_{\mathrm{max}}}$ values given by Eq. (12) are

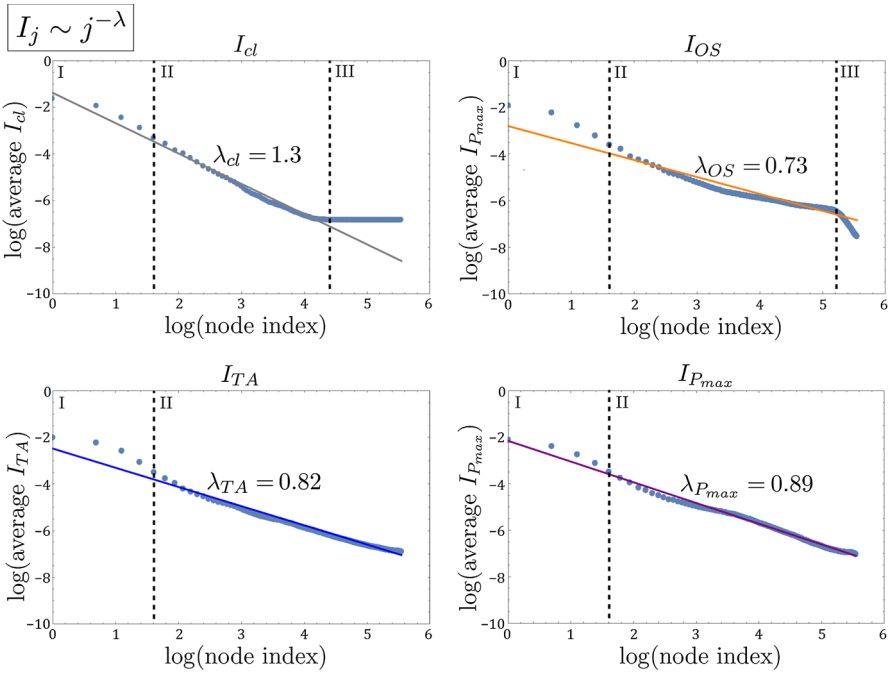

Fig. 10 Plots of the logarithms of the mean PageRanks for each measure over an ensemble of 30 scale-free networks against the logarithm of the node index (nodes sorted in descending PageRank order). Particularly over region II, the PageRanks follow a power-law distribution $I_j \sim j^{-\lambda}$ across the nodes with fitting parameter λ. For I_{cl} and I_{OS}, nodes in the tail region III are ignored in performing the linear fit

normalised to sum to one. For each measure, a plot of the logarithm of the sorted PageRank values I_j against the logarithm of the node index j has slope λ. We analyse an ensemble of 30 scale-free networks with $N = 256$ generated using NetworkX.

Figure 10 contains plots of the logarithm of the ensemble mean PageRanks against j, with linear fits yielding β for each PageRank measure. All PageRank measures display a power-law behaviour with scaling coefficients $\lambda_{\mathrm{OS}} < \lambda_{\mathrm{TA}} < \lambda_{P_{\mathrm{max}}} < \lambda_{\mathrm{cl}}$. The quantum-walk-based measures thus tend to concentrate less importance on the hubs compared to classical PageRank.

As noted in [7], the power-law behaviour of I_{TA} interpolates over a larger portion of the data compared to I_{cl}, marked as region II in Fig. 10. We observe that such a smoother power-law behaviour is preserved by $I_{P_{\mathrm{max}}}$. Therefore, I_{TA} and $I_{P_{\mathrm{max}}}$ can both better distinguish the low-lying nodes on scale-free networks.

In contrast to I_{cl}, and as observed in [37], the plot flattens out for nodes in region III with very low PageRank. Such a tail region without the general power-law behaviour is also present for I_{OS}, but is characterised instead by a sharp decrease in PageRank. Nodes of lowest importance are thus penalised by I_{OS}.

Finally, Fig. 11 presents an overall plot of each PageRank measure's distribution on scale-free networks. As shown in region II of the plot, intermediate nodes (including the secondary hubs discussed in Sect. 3.2) are given higher ranks according to (in

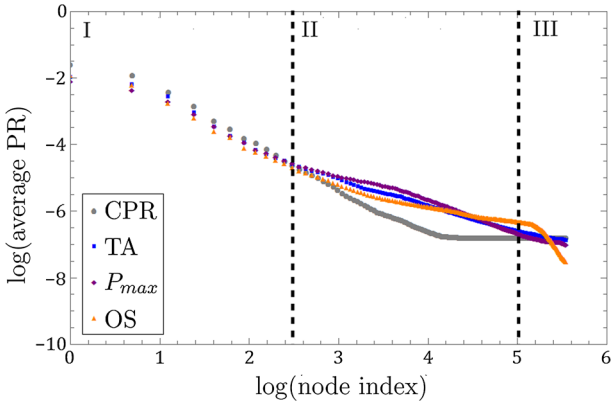

Fig. 11 Combined plot of the logarithms of the mean PageRanks from each measure over an ensemble of 30 scale-free networks versus the logarithm of the node index (nodes sorted in descending PageRank order)

descending order) $I_{P_{\max}}$, I_{TA}, and I_{OS} compared to I_{cl}. To compensate for this improved ranking, main hubs in region I receive lower quantum than classical PageRanks.

4 Discussion and conclusions

In this paper, we have presented a comparative investigation of three quantum PageRank measures, two of which are based on DTQWs, while the third uses an open-system CTQW. Extending the work in [7], we further studied the periodic nature of the instantaneous quantum PageRank I_q. In particular, we utilised it to propose a suitable timescale $t_{\max} = 2\langle T_q^5 \rangle$ for the corresponding DTQW based on the periods of the hub nodes. Such a timescale was observed to not scale with network size, making it feasible even use on for larger networks. This is particularly important, since it leads to a quantum PageRanking scheme that is scalable with respect to network size.

We have demonstrated that all three quantum PageRank measures are able to distinguish between outerplanar hierarchical, scale-free, and Erdős–Rényi directed networks. Through a comparative view, we observed similarities in rankings given by the DTQW-based PageRanks I_{TA} and $I_{P_{\max}}$, and between the open-system-based PageRank I_{OS} and the classical PageRank I_{cl}.

When applied to scale-free networks, the quantum PageRank schemes were better able to highlight secondary hubs in scale-free networks than the classical scheme, which tends to concentrate PageRanks on a few main hubs. This affirmed the results in [5,7]; moreover, we found that this quantum advantage is most apparent in $I_{P_{\max}}$, followed by I_{TA} and I_{OS}.

We used the inverse participation ratio (IPR) to characterise the walker's localisation on scale-free and Erdős–Rényi networks. We showed that for all four PageRank methods, the walker is in a localised phase on the hub-containing scale-free networks, and is delocalised on the hubless Erdős–Rényi random networks. Therefore, each PageRank scheme is shown to clearly distinguish between these two types of networks.

Lastly, the distribution of quantum PageRanks was observed to follow a power-law behaviour on scale-free networks—a property present in classical PageRank. In particular, we showed that $I_{P_{\max}}$ preserves the smoother power-law distribution over a larger portion of nodes as found in I_{TA} [7] compared to the classical case. We observed that I_{OS} scales according to a power law for most nodes, but displays a sharp drop in average PageRank for the nodes of lowest importance, thus indicating that these nodes are penalised by I_{OS} compared to I_{cl}.

Acknowledgements This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. TL thanks the Okinawa Institute of Science and Technology as much background learning was done there under the research internship programme, supervised by Thomas Busch and Chandrashekar Madaiah.

Author contributions TL made the most significant contribution to this work, including the development of methodology, detailed simulation and analysis especially for DTQW, as well as preparation of the manuscript. JT and JR worked on the open-system CTQW formulation, simulation, and analysis; MS provided very helpful feedback on network analysis in general; and JW provided direction and guidance.

References

1. Brin, S., Page, L.: Reprint of: The anatomy of a large-scale hypertextual web search engine. Comput. Netw. **56**, 3825 (2012)
2. Page, L., Brin, S., Motwani, R., Winograd, T.: The Pagerank Citation Ranking: Bringing Order to the Web, Technical Report, Stanford University, Stanford, CA (1998)
3. Ambainis, A.: Quantum walks and their algorithmic applications. Int. J. Quantum Inf. **1**, 507 (2003)
4. Paparo, G.D., Martin-Delgado, M.: Google in a quantum network. Sci. Rep. **2**, 444 (2012)
5. Sánchez-Burillo, E., Duch, J., Gómez-Gardeñes, J., Zueco, D.: Quantum navigation and ranking in complex networks. Sci. Rep. **2**, 605 (2012)
6. Montanaro, A.: Quantum walks on directed graphs. Quantum Inf. Comput. **7**, 93 (2007)
7. Paparo, G.D., Müller, M., Comellas, F., Martin-Delgado, M.A.: Quantum Google in a complex network. Sci. Rep. **3** (2013)
8. Berry, S.D., Wang, J.B.: Quantum-walk-based search and centrality. Phys. Rev. A **82**, 042333 (2010)
9. Szegedy, M.: Quantum speed-up of Markov chain based algorithms. In: Proceedings of 45th Annual IEEE Symposium on Foundations of Computer Science. IEEE, pp. 32–41 (2004)
10. Szegedy, M.: Spectra of quantized walks and a $\sqrt{\delta\epsilon}$ rule. arXiv: quant-ph/0401053 (2004)
11. Aharonov, D., Ambainis, A., Kempe, J., Vazirani, U.: Quantum walks on graphs. In: Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing. ACM, pp. 50–59 (2001)
12. Childs, A.: Lecture 11: Discrete-time quantum walk. http://www.cs.umd.edu/~amchilds/teaching/w13/l11.pdf, 14 Aug 2015
13. Schult, D.A., Swart, P.: Exploring network structure, dynamics, and function using NetworkX. In: Proceedings of the 7th Python in Science Conferences (SciPy 2008), vol. 2008, pp. 11–16 (2008)
14. Chiang, C.-F., Nagaj, D., Wocjan, P.: Efficient circuits for quantum walks. Quantum Inf. Comput. **10**, 420 (2010)
15. Loke, T., Wang, J.B.: Efficient quantum circuits for Szegedy quantum walks. arXiv:1609.00173 (2016)
16. Farhi, E., Gutmann, S.: Quantum computation and decision trees. Phys. Rev. A **58**, 915 (1998)
17. Nielsen, M., Chuang, I.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2010)
18. Rivas, Á., Huelga, S.F.: Open Quantum Systems. Springer, Berlin (2012)
19. Saalfrank, P.: Open system density matrix theory: numerics and chemical applications. Course notes (1998–2000)
20. Comellas, F., Miralles, A.: Modeling complex networks with self-similar outerplanar unclustered graphs. Phys. A **388**, 2227 (2009)
21. Comellas, F., Miralles, A.: Vertex labeling and routing in self-similar outerplanar unclustered graphs modeling complex networks. J. Phys. A Math. Theor. **42**, 425001 (2009)
22. Everett, M.G.: Role similarity and complexity in social networks. Soc. Netw. **7**, 353 (1985)

23. Borgatti, S.P., Everett, M.G.: Notions of position in social network analysis. Soc. Methodol. **22**, 1 (1992)
24. Barabáasi, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**, 509 (1999)
25. Albert, R., Jeong, H., Barabási, A.-L.: Internet: diameter of the world-wide web. Nature **401**, 130 (1999)
26. Eguiluz, V.M., Chialvo, D.R., Cecchi, G.A., Baliki, M., Apkarian, A.V.: Scale-free brain functional networks. Phys. Rev. Lett. **94**, 018102 (2005)
27. Albert, R.: Scale-free networks in cell biology. J. Cell Sci. **118**, 4947 (2005)
28. Hein, D.-I.O., Schwind, D.-W.-I.M., König, W.: Scale-free networks. Wirtschaftsinformatik **48**, 267 (2006)
29. Barabási, A.-L., et al.: Scale-free networks: a decade and beyond. Science **325**, 412 (2009)
30. Barabási, A.-L., Albert, R., Jeong, H.: Scale-free characteristics of random networks: the topology of the world-wide web. Phys. A **281**, 69 (2000)
31. Bollobás, B., Borgs, C., Chayes, J., Riordan, O.: Directed scale-free graphs. In: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, pp. 132–139 (2003)
32. Gilbert, E.N.: Random graphs. Ann. Math. Stat. **30**, 1141 (1959)
33. Erdös, P., Rényi, A.: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960)
34. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 47 (2002)
35. Martin, T., Zhang, X., Newman, M.: Localization and centrality in networks. Phys. Rev. E **90**, 052808 (2014)
36. Donato, D., Laura, L., Leonardi, S., Millozzi, S.: Large scale properties of the webgraph. Eur. Phys. J. B Condens. Matter Complex Syst. **38**, 239 (2004)
37. Pandurangan, G., Raghavan, P., Upfal, E.: Using pagerank to characterize web structure. In: Proceedings of the 8th Annual International Computing and Combinatorics Conference (COCOON), 2002